Defense Threat Reduction Agency
8725 John J. Kingman Road, MS 6201
Fort Belvoir, VA 22060-6201

DTRA-TR-21-14

Comparative Computational Modeling of Immune Reponses to Vaccines

TECHNICAL REPORT

Distribution Statement A. Approved for public release. Distribution unlimited.

HDTRA1-17-1-0052
Performer Name: ioGenetics, LLC & Versiti, Inc.

Prepared by: Robert Bremel and Jane Homan

June 2021

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
06/21/2021	FINAL	08-16-2017 to 03-31-2021

4. TITLE AND SUBTITLE
Comparative Computational Modeling of Immune Responses to Vaccines

5a. CONTRACT NUMBER

5b. GRANT NUMBER
HDTRA11710052

5c. PROGRAM ELEMENT NUMBER
0601384BP

6. AUTHOR(S)
BREMEL, Robert D.
HOMAN, E. Jane
GORSKI Jack
CUI, Weiguo

5d. PROJECT NUMBER
CB10433

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
ioGenetics LLC, 3591 Anderson Street, Madison WI 53704 (mail PO Box 384 Sun Prairie WI 53590)
Versiti Inc, 8727 W Watertown Plank Rd, Milwaukee, WI 53226

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Threat Reduction Agency
8725 JOHN J. KINGMAN ROAD, MSC 6201
FORT BELVOIR VA 22060-6201

10. SPONSOR/MONITOR'S ACRONYM(S)
DTRA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DTRA-TR-21-14

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution A

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The project assessed the ability of a computational immunology platform as a predictor of multifaceted immune responses in humans and mice. An accurate computational predictor of immune responses in individuals of differing immunogenetics potentially enable the rapid development of vaccines for emerging pathogens with more confidence than depending on animal models, which do not always replicate human disease, nor the immune responses of a diverse human population. The project mapped epitopes in ten FDA approved vaccines and in SARS-CoV-2. B and T cell responses to SARS-CoV-2 were evaluated in convalescent human volunteers in comparison to computational predictions.

15. SUBJECT TERMS
immunoinformatics, vaccine, epitope, predictive modeling, immunology

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jane Homan
U	U	U	UU	819	19b. TELEPHONE NUMBER *(Include area code)*: 608 577 8829

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

TABLE OF CONTENTS

Objectives and Rationale 5
Implementation 6
Institutional Partners 6
Rationale for Changes in Scope of Work 7
Immunoinformatics Platform 9
Project Tasks 14
Task I: Computational analysis of approved vaccines and SARS-CoV-2 14
Modelling of FDA approved vaccines 14
Epitope Mapping of SARS-CoV-2 16
Selection of peptides for epitope mapping of yellow fever analysis in mice 26
Task II: Recruitment and collection from human volunteers recovered from COVID 29
Task III: Evaluation of responses in human volunteers to COVID 32
SARS-CoV-2 Antibody responses 32
SARS-CoV-2 T cell epitope responses 38
Comparative analysis of predictions to epitope specific immune responses to Sars-CoV-2 documented in literature reports. 47
Task IV: Evaluate responses of mice to Yellow fever vaccine virus. 55
Discussion and Conclusions 59
Appendices 61
Addenda 61
References 62

Tables:

Table 1. Components of EigenBio Analysis....11
Table 2: FDA approved vaccines evaluated....14
Table 3: Data provided for each sequential peptide in each protein (predicted values). 15
Table 4: Peptides selected to assay MHC IA binding....21
Table 5: Peptides selected to assay MHC IB binding....22
Table 6: Peptides selected to assay MHC II DRB binding....23
Table 7: YF Vax® peptides predicted affinity for murine MHC I alleles Kb and Db....26
Table 8: YF Vax® peptides predicted affinity for murine MHC II allele H2-1Ab....27
Table 9: Characteristics of convalescent SARS-CoV-2 donors....30
Table 10: MHC I peptides used for each volunteer....40
Table 11: MHC II peptides used for each volunteer....41
Table 12: Summary of assay components....42
Table 13: Interferon gamma responses IG-003....44
Table 14: Interferon gamma responses IG-0009....45
Table 15. Summary of responses from all patients....46

Figures:

Figure 1: Comparative overview spike glycoprotein of betacoronaviruses....17
Figure 2: Peptides from SARS-CoV-2 proteins selected for inclusion on the Pepperprint chips 19
Figure 3: Selection of predicted B cell linear peptides in YF-VAX®. 26
Figure 5: Variation between the antibody responses for the different subjects....33
Figure 4: Pepperprint® pattern for subject IG-002....33
Figure 6: ANOVA by predicted BEPI group....34
Figure 7: ANOVA individual subjects....35
Figure 8. Ranking of antibody binding to Spike glycoprotein linear B cell epitopes in Pepperprint®....37
Figure 9: Comparison of B cell epitope predictions to observations of Dahlke *et al*....47
Figure 10: Response rates in all MHC Class I responses curated by IEDB.org....48
Figure 11: TCEM motifs in the Spike protein of the reference SARS-CoV-2 protein. 50

Figure 12: Antigen presentation profiles of TCEM I from SARS-CoV-2 virus.....52

Figure 13: Distribution of TCEM in SARS-CoV-2 that elicit a T cell response.....53

Figure 15: Responses of $CD4^+CD44^+$ cells to interferon gamma stimulation by YF peptides predicted to produce a strong T cell stimulation. 56

Figure 14: Responses of $CD8^+CD44^+$ cells to interferon gamma stimulation by YF peptides predicted to produce a strong T cell stimulation 56

Figure 16: Statistical comparison of standardized binding predictions to IEDB curated in vitro T cell responses to MHC I and MHC II bound YF peptides.....58

Grant/Award #: HDTRA111710052

PI Name: Robert D. Bremel

Organization/Institution: ioGenetics LLC, SubAwardee: Versiti Inc.

Project Title: Computational Modeling of Immune Responses to Vaccines

Objectives and Rationale

The overall goal of the project was to assess the ability of the ioGenetics' computational immunology platform as a predictor of multifaceted immune responses. Having an accurate computational predictor of immune responses in individuals of differing immunogenetics would enable the rapid development of vaccines for emerging pathogens with more confidence than depending on animal models, which do not always replicate human disease, or the immune responses of a diverse human population. Immunoinformatics modelling conducted at ioGenetics was compared with B and T cell responses in human leukocyte antigen (HLA)-typed human volunteers sampled, and assayed by Versiti, as a sub-awardee (formerly the Blood Research Institute and Blood Center of Wisconsin). Predictions were also compared to results published by others.

As a result of difficulties in obtaining the FDA approved vaccines originally selected, the advent of the 2019 coronavirus (COVID) pandemic, and based on discussions with the Defense Threat Reduction Agency (DTRA), the goals were amended to address the immunological modeling of responses to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus infection.

Immunoinformatic models that provide accurate, comprehensive predictions of the adaptive immune response can enable them to play a greater role in the design of vaccines. The need for rapid response in vaccine development was clearly demonstrated in 2020 with the emergence of SARS-CoV-2. The ability to model the immunological features of any protein within a very short time (hours-minutes depending on complexity) once genome and then *fasta* amino acid sequences are available enables fast-tracking of vaccine development and so enables ongoing vigilance of efficacy in the face of mutations.

Determinations of efficacy in prior vaccine development have rested heavily on the detection of neutralizing antibodies, while largely ignoring T cell responses. Yet it is increasingly clear that immunological memory and efficacy rests as much, or more, on T cell responses as antibodies [1, 2]. It is clear that many factors play into vaccine efficacy and responses. Prior exposure to similar epitopes, T cell precursor frequency, epitope mimics, and especially individual HLA alleles are all key factors. Immunoinformatic modelling allows these to be assessed rapidly in a way that *in vivo* testing cannot do. Phase I and II clinical trials are necessarily limited in size and hence the diversity of HLA and prior immunological experience that they encompass. In this project, the time differential in obtaining informatic analysis compared to the extended time frame needed to obtain even limited human data is testament to that.

Assays of antibodies based on neutralization or *in vitro* responses to selected peptide antigens tend to overlook the response to a microorganism by non-neutralizing antibodies. Non-neutralizing antibodies may be present in higher titers than neutralizing and have many significant effects on host-pathogen interaction. Computational analysis offers the opportunity to envision the immune response in more granular detail and to understand the impact on the immune response arising from single non-synonymous nucleotide changes. In addition, computational analysis can anticipate some potential adverse effects, which may not affect all individuals and may not affect animals. Such adverse effects can arise, for example, from off-target binding causing autoimmune reactions. Potential off-targets can be identified by searching a pre-computed database of the human proteome and related to the HLA for whom they may have an impact, whereas this convergence of risk factors would be hard to identify *in vivo* due to the sporadic incidence.

The output of this project clearly demonstrates the speed and detail that immunoinformatic analysis can provide relative to human studies. Given the obstacles encountered in obtaining vaccines, and the emergence of the COVID pandemic, it has only partially completed the original goals of comparing *in vivo* to *in silico* analysis. However, additional comparative data has been derived from the literature and for SARS-Cov-2 as it continues to emerge. Also evident from the output of this study is the very great degree of individual (personal) variability in responses, which can be linked to the informatic responses.

Implementation

Institutional Partners

ioGenetics LLC is a small business, based in Madison WI, which has developed the EigenBio immunoinformatic platform. This platform has been extensively applied in collaborative projects with universities and companies in the US, Europe, South America, and Asia. While early work focused primarily on infectious organisms, the company has collaborated in research on autoimmune diseases and cancer. Current collaborative projects include large scale mapping of fungal antigens, hookworm vaccine development, and multiple sclerosis. A major initiative is underway in the mapping of cancer neoepitopes and design of personalized neoantigen vaccines. In this project, as prime contractor ioGenetics implemented all immunoinformatics mapping and analysis.

Project personnel at ioGenetics:

Robert Bremel, PhD	Principal Investigator. Chief Scientific Officer, lead bioinformatician
Jane Homan, PhD, MRCVS	Chief Executive Officer. Bioinformatics analysis

Versiti (BRI/ARL). In the present project Versiti was a subawardee to ioGenetics and responsible for conducting the human and mouse studies. Versiti is a non-profit organization that provides blood products, hematology related services, cell and tissue banking, and diagnostic lab services in five states. Versiti WI, formerly known as Blood Center of Wisconsin, played a lead role in the formation of what is now Versiti. The Applied Research Lab (ARL) is a

part of Versiti that provides contract research support, generally based on blood products or associated with blood banking, to other organizations. It was well placed to coordinate subject accrual and blood sample preparation. The Blood Research Institute is the basic research division of Versiti. Since its origin, Versiti WI, known then as the Milwaukee Blood Center, has had an active research program in immunohematology and transfusion medicine. Currently the BRI has researchers in the fields of benign hematology, blood banking, immunology, cardiovascular biology, and stem cell biology. Most Versiti researchers have academic appointments at the Medical College of Wisconsin (MCW), which is located on the same medical center campus as is the BRI. MCW provides animal facilities and provides Institutional Review Board service for Versiti researcher who perform human research. Weiguo Cui, MD Ph.D. is one of the latest immunology recruits and has a successful program in the study of immune memory, immune metabolism, gene regulation and immuno-oncology. Dr Gorski, Ph.D. has concentrated on the human immune system examining T cell responses to influenza. He has also been interested in how the polymorphism of HLA affects peptide binding and selection. Dr. Cui applied his expertise to the mouse model, and Dr. Gorski to the testing of predicted T cell epitopes from SARS-CoV-2.

Project Personnel at Versiti:

Jack Gorski, PhD	Sub Awardee PI, Senior Investigator, BRI Assoc Professor, MCW (now Emeritus), human studies.
Weiguo Cui. MD, PhD	Investigator, BRI, Assoc Professor, MCW, mouse studies
Wendy Winker, BS	Laboratory Manager/Research Technologist II, BRI. Antibody and T cell assays human.
Christine Nguyen	Graduate Student, Cui lab.
Sybil Heldke	Applied Research Lab Versiti, Applied Research Specialist. Human volunteer coordination and sample preparation.

Rationale for Changes in Scope of Work

The initial conception of the project was to identify two Food and Drug Administration (FDA) approved vaccines, model the predicted T and B epitopes and to recruit volunteers to receive the vaccines and undergo monitoring of their immune responses. Yellow Fever (YF) vaccine was proposed as the first vaccine, with the choice of a second one to be made following immunoinformatic monitoring. Actually obtaining a suitable vaccine for study proved much more difficult than anticipated.

While all the vaccines originally listed were modelled and epitopes mapped, the selection of FDA vaccines suitable for a study of epitope-specific T cell responses study is fairly narrow and constrained the choice of vaccines we could use. Over 80 vaccines are currently listed as having received FDA approval, spanning 27 diseases (Appendix 1). However, this list includes vaccines that are inactivated (e.g. Japanese encephalitis), conjugated (e.g. *Hemophilus*), toxoids (e.g. tetanus), very large bacterial vaccines in which individual epitopes can be mapped

but responses are clinically difficult to separate (e.g. Bacillus Calmette-Guerin (BCG)), multi-valent (e.g. Human Papilloma Virus), vaccines not easily accessible or of restricted use (e.g. anthrax), and others with responses likely to be confounded by recent exposure (e.g. influenza). In many cases, the package insert does not make clear the exact source or the microbial strain used in preparation of the vaccine; a problem as single amino acid differences impact epitope prediction. Acceptability to the volunteers and logistics of administration was also a consideration, with a single dose vaccine being an ideal choice. Inactivated vaccines were considered unsuitable for evaluation of the informatics platform as formaldehyde, betapropiolactone and other methods of inactivation cross link or denature amino acids precluding natural binding of potential T cell epitopes to major histocompatibility complex (MHC) molecules [3, 4]. These considerations narrowed the candidate vaccines for application in volunteers to a handful.

Yellow fever was considered an optimal choice as a single dose live viral vaccine, with an extensive safety record, and for a disease to which the volunteer donor population in Milwaukee would not have been exposed. Unfortunately, Sanofi Pasteur's YF-Vax® became unavailable in the United States in late 2017 and remains unavailable in 2021 as they transition to a new manufacturing facility [5, 6]. A near-identical yellow fever vaccine produced by Sanofi in France, Stamaril® was made available to US travelers under an investigational new drug application (IND). However, the scope of that IND did not cover research studies and had the vaccine even been possible to obtain, volunteers in the present study would have had to also be enrolled under the Sanofi Pasteur IND.

Our second choice was a herpes zoster vaccine. We modeled Shingrix® a recombinant protein and planned to use that. As soon as the Centers for Disease Control and Prevention (CDC) recommended Shingrix® as the preferred vaccine for shingles over Zostavax®, demand for the vaccine immediately outstripped manufacturing capacity, making it impossible to obtain for research purposes, a situation that continued into 2020.

An Institutional Review Board (IRB) approval was therefore sought for use of MMR II® (measles mumps rubella). This has the advantages of being a live vaccine, with known strains of virus and accessibility. The disadvantage is that most volunteers had prior exposure due to childhood vaccinations or prior infection. We elected to focus on the mumps component as T cells in mumpsvirus have been less studied. The Jeryl Lynn vaccine strains are proprietary to Merck, meaning that lab work would have to be done with an alternative isolate and account for the mutations unique to Jeryl Lynn, including mutations in the key neutralizing epitope [7].

At this point our project officer recommended we use the Ebola vaccine Erbevo® which received FDA approval in December 2019. This vaccine fit the criteria for our use and our study would have provided valuable information. An IRB amendment was initiated. However, Erbevo® proved to be unavailable to the Medical College of Wisconsin Travel Clinic and also unavailable to us through DTRA channels.

IRB approval for the MMR vaccine was finally received in March 2020, shortly after a lockdown was declared due to COVID, making it impossible to bring volunteers in for vaccination and sample collection.

After discussion with the project team at DTRA, the focus was amended in June 2020, to evaluation of immune responses in individuals recovered from COVID infection. This then necessitated the resubmission of an IRB proposal which was approved 27 August, 2020 and the Human Research Protection Office (HRPO) approval followed on 1 September, 2020 permitting volunteers to be enrolled from late September to late November 2020.

Immunoinformatics Platform

Technology platform

ioGenetics' EigenBio immunoinformatics platform is a novel systems approach to immunome analysis which integrates analytics for multiple components of the adaptive immune response to enable understanding of the host-antigen interface. It is a holistic approach deploying a combination of machine learning, artificial intelligence technology and statistics that interface with multiple different proprietary and public databases.

The immune system response to proteins, regardless of their origin, is embedded in physicochemical properties of the amino acid sequence. The alphabetic sequence of the amino acids cannot be used in numerical computation because there is no direct mathematical transformation of alphabetic to numeric values. Many schemes such as the point accepted mutation (PAM) matrices and other substitution matrices [8, 9] have been used for specific purposes, but all involve a number of compromises. In particular, there is no direct transform of alphabetics to physical properties. Plus, the substitution matrices are derived from evolutionary substitution of amino acids in protein structures in mammals.

A key, and unique, feature of the EigenBio system is the 'first principles' approach. The predictive platforms are derived from multiple physical properties of the amino acids [10, 11] as the fundamental features that are recognized by various components of the immune system. EigenBio achieves a direct transformation of amino acid alphabetics to numerical values by the use of principal component analysis (PCA) [10, 12]. Principal component analysis of the physical properties allows diverse physical properties, measured in many different units, to be used to produce a matrix of numbers that, through matrix algebra, can be shown to account for a defined fraction of the variance of the physical properties themselves. PCA starting with 31 physical properties of the amino acids, generates 31x31x20 combinations. A 20x3 matrix of the first 3 components accounts for 90% of the variance. Seven principal components (PC) comprise approximately 95%, and eleven PC 99% of the variance in 31 diverse amino acid physical properties widely used in structural biologic modeling. Thus, this is a tunable system depending on the level of precision needed and speed required. EigenBio routinely uses 3 PC. In addition, by virtue of the PCA matrix algebra, the matrix elements are appropriately weighted for regression predictions and neural network use. Because the mathematical properties are related directly to the physical properties, an alphabetic sequence can thus be directly converted into a matrix of physical properties, a process for which there is no equivalent with

substitution matrices. Thus, peptide alphabetic sequences are represented as small matrices and various matrix manipulations can be used to describe helices, beta sheets, flexibility, etc. In addition, these matrices are directly usable in the widening array of machine learning and artificial intelligence tools.

EigenBio takes advantage of the enormous increase in available machine learning tools and neural network algorithms that has occurred in the past few years. We selected the JMP® statistical package from SAS Institute on which to develop the EigenBio platform for several reasons. SAS Institute has produced robust statistical tools over four decades, including JMP®. JMP® is written in JSL, the javascript-like scripting language (translatable to JAVA and Javascript). JMP® contains a series of artificial intelligence (AI) internals analogous to the TensorFlow library from Google. It interfaces with other commonly used tools such as R, Python, Matlab, BioConductor, Genbank, and other databases as needed. JMP® is recognized by both FDA and International Organization for Standardization (ISO) as validatable. EigenBio currently comprises approximately 80,000 lines of code and has been maintained under version control over the last 8 years.

The components that EigenBio routinely analyzes, shown in Table 1, include topology, linear B cell epitopes, predicted binding to human MHC I (70 alleles) and MHC II (65 alleles), as well as those of several strains of inbred mice, cleavage by cathepsin S, L and B, and T cell exposed motif frequency. The platform is high-throughput, allowing detailed analysis of every peptide in individual proteins, as well as characterization of whole organism proteomes and epitope patterns within organisms and proteins [7, 13, 14]. By evaluating every sequential peptide, each with a single amino acid displacement, EigenBio avoids the pitfalls of analyzing peptides with longer overlaps. Overlapping peptides, frequently used in the lab for cost and expediency, typically use a five amino acid displacement, which may unwittingly interrupt optimal MHC binders and overlook the need for endopeptidase processing for natural presentation.

Using EigenBio, ioGenetics has established extensive reference databases of proteins computed in detail to allow comparison of the patterns found in newly analyzed organisms and proteins. The databases include the human proteome (~80,000 proteins including isoforms), immunoglobulin variable regions (over 40 million), the gastrointestinal microbiome, several thousand cancer-associated proteins, and thousands of pathogenic microorganisms and parasites. Cross-referencing of B and T cell epitopes against these databases makes it possible to identify epitope networking made possible by polyspecificity and the occurrence of epitope mimics [15-17]. Epitope frequency and binding are then combined as a predictor of cognate T cell population response and to generate predictors of stimulation vs suppression.

Table 1. Components of EigenBio Analysis		
Immunological Metric	**Method**	**Prediction**
MHC I affinity 9-mer	Neural Network Ensembles trained on binding data 9-aa sliding window indexed by 1 amino acid 31 HLA-A; 31 HLA-B; 8 HLA-C plus 6 murine	LN (IC_{50}) std dev LN(IC_{50}) (ave = 0.5 LN (IC_{50})) standardized affinity within protein or proteome; relative binding probability thresholds for each peptide
MHC II affinity 15-mer	Neural Network Ensembles trained on binding data sliding window of 15-amino acids indexed by 1 aa 24 DR; 13 DP; 28 DQ plus 3 murine	LN (IC_{50}) std dev LN(IC_{50}) (ave = 0.5 – 0.7 LN (IC_{50})) standardized affinity within protein or proteome; relative binding probability thresholds for each peptide
MHC + MHC II cross presentation	MHC I + MHC II simultaneous binding to peptides in a protein	Number of high affinity MHC I 9-mers within each high affinity MHC II 15-mer
Linear B-cell epitope 9-mer	Neural network trained on B-cell epitopes Sliding window of 9 amino acids indexed by 1 aa	Probability of B cell binding standardized within protein Relative probability among peptides
Endosomal and proteasomal processing Combined cathepsin cleavage-MHC binding affinity	Neural Network Ensembles trained on large proteomic cleavage database mass spectrometry of cleaved peptides for cathepsins. Proteasomal shuffling of amino acid primary sequences of neoantigens. Combination of 3 prediction method outputs	Human cathepsin B, L and S: Two independent predictions: probability of cleavage + probability of non-cleavage between aa4 and aa5 (P1P1') of an octamer. Peptides generated by proteasomal shuffling to generate novel neoantigens. Probability of excision of various length peptides between 15-21 amino acids in length for MHC II and exactly 9 amino acids for MHC I.
T-cell exposed motif (TCEM) relative to normal human repertoires	Frequency comparison to database for: continuous pentamer within a bound MHC I 9-mer two discontinuous pentamers (9-mer core of a bound MHC II 15-mer)	Specific motifs exposed to T cell by peptide bound in MHC Frequency relative to IGHV germline and somatic mutation Frequency relative to human proteome Frequency relative to gastrointestinal microbiome
Combination of MHC I and MHC II binding with TCEM T-cell exposure and cathepsin cleavage	Graphical/interactive combination of 5 different prediction outputs	Interactive graphical platform for evaluation Treg potential based on combination of: Binding predictions and TCEM frequency data using additivity of variance
Mimicry of B-cell receptors (immunoglobulins)	Pentamer (core of 9-mer BEPI prediction) exposure comparison between proteins of interest (e.g. viral protein vs multiple isoforms of all proteins in human proteome) in combination with MHC binding predictions	Proximity of MHC binding regions to BEPIs Pentamer matches (random probability of match = 20^{-10}) BEPI probability classification matches between selected proteins UniProt keyword screens URL connection to internet resources
Protein topology	Neural Network Ensembles, 5 models Slider, graphical interface	Amino acids of protein comprising extra-cellular, intracellular domain and trans-membrane domains Signal peptides cleavage point
Repertoire analysis	Unique approaches to B cell and T cell repertoire analysis	Characteristics of clonal frequency distributions Time-dependent changes in frequency distributions Machine learning techniques applied to frequency distribution clustering

A unique feature of the EigenBio platform is the recognition that because the T cells respond to exposed pentamer motifs in the peptide:MHC complex there is a limited repertoire of T cell exposed motifs which engage T cell receptors [18]. The twenty natural amino acids can only be configured in 3.2 million possible combinations (20^5 = 3.2 x 10^6) for each of MHC I, and two possible pentameric engagement patterns for MHC II [15]. We refer to these exposed amino acid motifs as T cell exposed motifs (TCEM). Moreover, we have determined that different motifs in proteins have characteristic frequency distributions and have created databases of these motif patterns. The obverse face of the peptide:MHC complex engages with the side chains of the amino acids in the MHC binding groove and is responsible for the affinity and stability of the complex. This is termed the groove exposed motifs (GEM) or pocket positions [19]. It is implicit that, like the TCEM, these motifs also have a finite range and characteristic frequencies. This enables us to characterize a protein (or a peptide) by its content of uncommon motifs with potential to upregulate the immune response, or by its content of very common motifs which may elicit a T regulatory response, provided in both cases the peptides are flanked by MHC binding motifs. Building on this, we can generate graphic outputs identifying proteins in a proteome which are more, or less, likely to stimulate an immune response identified by a stimulation or suppressive index.

EigenBio's integration of the many components of the immune response and application of statistics recognizes that the immune response is a multivariate system which operates with fuzzy logic [20]. The primary computational workflow produces a mastermatrix for a protein, or a proteome, comprising nearly a thousand numerical metrics for each peptide. This large set of metrics makes it possible to provide an end-user with simple ways to analyze their data and produce various graphics from the protein-specific scripts produced during the mastermatrix workflow.

The predictions made by EigenBio have been validated against reference proteins and through prospective predictions [21-24]. Where possible, benchmark testing has been done against the reductionist web site tools [10]. Many of these have serious technical issues and none have an ability to link the results to other necessary protein features that are needed to contextualize the results.

EigenBio has established a track record not only with microbial immunology, but also in cancer neoepitope analysis, autoimmunity and B and T cell repertoire analysis. Examples of additional publications based on EigenBio are referenced at its website. The initial platform has been expanded to include workflows and immune-analytics for B cell and T cell repertoires in diffuse large B-cell lymphoma (DLBCL), chronic lymphocytic leukemia (CLL), and human stem cell transfer (hSCT) as well as for predictions of personal neoepitopes in oncology, with several major clinical centers in the US, Japan, and Europe.

A key aspect of the EigenBio platform is its high throughput and speed in execution of the mapping of the features shown in Table 1. For example, the mapping of these features in a protein of 500 amino acids can be computed in ~15 minutes. This enables the evaluation of novel agents as soon as a sequence is available and hence the potential to start design of a vaccine within the same day. This was evidenced in the case of SARS-CoV-2 where we

modelled the virus proteins as soon as the initial sequence became available in early January 2020 and it had become apparent that a potentially rapidly spreading pathogen was at large. We compared the epitope profile to that of severe acute respiratory syndrome (SARS) and MERS and determined that these viruses provided an excellent blueprint for an approach to vaccine development. In other instances, however, with an unknown agent, where such similar and previously characterized viruses may not be available, the ability to rapidly identify epitopes of interest could be critical to rapid intervention.

The EigenBio immunoinformatics platform is a proprietary platform. Issued and pending patents pertaining to its overall operation and application to infectious agents were filed prior to the inception of the present project. ioGenetics is open to discussion of licensing in specific fields of use.

Project Tasks

Task I: Computational analysis of approved vaccines (ioGenetics)

Task II: Recruitment and collection from human volunteers recovered from COVID (Versiti)

Task III: Evaluation of responses in human volunteers to COVID (Versiti)

Task IV: Evaluate responses of mice to Yellow fever vaccine virus. (Versiti)

Task I: Computational analysis of approved vaccines and SARS-CoV-2

Modelling of FDA approved vaccines

Modeling of the approved vaccines was completed as shown in Table 2. This includes the originally proposed vaccines plus others explored in the process of trying to identify a suitable and accessible candidate vaccine to work with.

Graphical representations of each and predicted B and T cell epitopes are compiled and available as pdf and Excel spreadsheets as an Addendum to this report.

Table 2: FDA approved vaccines evaluated		
Vaccine/Tradename	Format	**Isolate/s**
Yellow fever (YF Vax®)	Live attenuated	17D-204
Hepatitis B (Recombivax®)	Virus like particles (VLP)	*adw* subtype of HBsAg
Human papillomavirus (Gardasil 9®)	Multivalent VLP	Types 6, 11, 16, 18, 31, 33, 45, 52, and 58
Measles, mumps, and rubella (MMR-II®)	Live attenuated and multivalent	Enders' attenuated Edmonston strain measles; JL5 JL2 mumpsvirus; Wistar RA 27/3 rubella
Rotavirus (Rotarix®)	Live attenuated	human 89-12
Diphtheria and tetanus -Sanofi	Toxoids	“selected strains”
Mycobacterium tuberculosis BCG	Live	TICE
Ebola (Erbevo®)	VSV chimera	Zaire Kikwit 1995 strain
Hepatitis A (Havrix®)	Inactivated	strain HM175
Rabies (RabAdvert®)	Inactivated	Flury-LEP
Human herpes virus 3 (varicella zoster) (Shingrix®)	recombinant	Glycoprotein E; strain unspecified

The output of mapping of these vaccine strains is shown in the Addenda, which include both graphical representations and spreadsheets. As the EigenBio immunoinformatics platform operates in JMP® predictions for every sequential peptide in the proteins for the key metrics shown in the table below have been exported in Excel format to facilitate access and provided to DTRA.

Table 3: Data provided for each sequential peptide in each protein (predicted values).

Permuted average MHC I binding
Permuted average MHC II binding
Predicted linear B cell epitopes
Individual MHC I alleles: A_0101, A_0201, A_0202, A_0203, A_0206, A_0211, A_0212, A_0216, A_0217, A_0219, A_0250, A_0301, A_0801, A_1101, A_2301, A_2402, A_2403, A_2501, A_2601, A_2602, A_2603, A_2902, A_3001, A_3002, A_3101, A_3201, A_3301, A_6801, A_6802, A_6901, A_8001, B_0702, B_0801, B_0802, B_0803, B_1501, B_1502, B_1503, B_1509, B_1517, B_1542, B_1801, B_2703, B_2705, B_3501, B_3801, B_3901, B_4001, B_4002, B_4402, B_4403, B_4501, B_4506, B_4601, B_4801, B_5101, B_5301, B_5401, B_5701, B_5801, B_7301, B_8301, C_0303, C_0401, C_0501, C_0602, C_0702, C_1203, C_1402, C_1502
Individual MHC II alleles: DRB1_0101, DRB1_0101 C30S, DRB1_0301, DRB1_0401, DRB1_0404, DRB1_0405, DRB1_0701, DRB1_0801, DRB1_0802, DRB1_0901, DRB1_1001, DRB1_1101, DRB1_1201, DRB1_1301, DRB1_1302, DRB1_1454, DRB1_1501, DRB1_1602, DRB3_0101, DRB3_0202, DRB3_0301, DRB4_0101, DRB4_0103, DRB5_0101 (plus DP DQ alleles)
Murine alleles: MHC I H-2-Db, H-2-Dd, H-2-Kb, H-2-Kd, H-2-Kk; MHC II H-2-IAb, H-2-IAd, H-2-IAg7
Predicted cathepsin cleavage /excision: Predicted excision of 9mer and of 12-18mer (15mer +/- 3)

Additional data included in the interactive JMP® files includes T cell exposed motif frequencies compared to human immunoglobulinome, proteome and gastrointestinal microbiome; suppressive and stimulation indices and multiple other derivatives.

Epitope mapping of the TICE strain *M. tuberculosis* BCG was completed, but is not included in the Addenda, as it comprises >4000 proteins and it was not selected for further use in the project. These files are available on request.

Consideration was given to including the FDA approved seasonal influenza vaccine in our selection of vaccines for detailed mapping. This was not a desirable choice given the extensive prior exposure to influenza vaccines and/or natural infection of any possible volunteers. Nevertheless, predictions of the seasonal influenza vaccines were done in each year as the recommended strains for the year were made known. These were periodically checked against circulating wild type isolates of the year, which were also reviewed for any indicators of increased pathogenicity. Differences between sources of vaccine seedstock for the same reference isolate were noted. For instance, in 2017 which was a particularly bad year for H3N2 influenza, the cell-based vaccine shared B cell epitopes with the egg derived vaccine, but differed in its predicted T cell epitopes. Notably, our observations showed that the cell-based vaccine in 2017 more closely matched the circulating wild type viruses that year. The cell-based vaccine was subsequently reported to have a greater efficacy, as we predicted [25, 26]. This analysis is shown in Appendix 2. With the occurrence of a single human infection with influenza A H10N3 in China in the spring of 2021, we mapped this influenza strain as soon as its sequence was available and noted the presence of a very common, potentially CD8+ suppressive TCEM motif in the hemagglutinin, similar to one seen in 1918, but absent any evidence of human transmissibility.

Epitope Mapping of SARS-CoV-2

Initial mapping – January 2020

The initial sequences for SARS-CoV-2, then known as Wuhan seafood market virus, were posted to Genbank on January 5, 2020 by Wu *et al* [27], subsequently being designated as the Reference Sequence NC_045512. When the first reported case occurred in the United States on January 21, 2020, we mapped all proteins in the virus and compared to reference strains of SARS-CoV-1 and MERS. This allowed a comparison within hours. Comparative overview illustrations of the three virus surface glycoproteins are shown below (Figure 1). Note that the plots shown in this Figure provide only a general overview; this is underpinned by detailed analysis of each protein in each virus and a complete set of graphics for SARS-CoV-2 is in Appendix 4. We also evaluated potential peptide epitopes for vaccine inclusion. In addition, we evaluated the potential for off-target adverse B cell and T cell responses by mapping linear B cell epitopes and T cell against the human proteome.

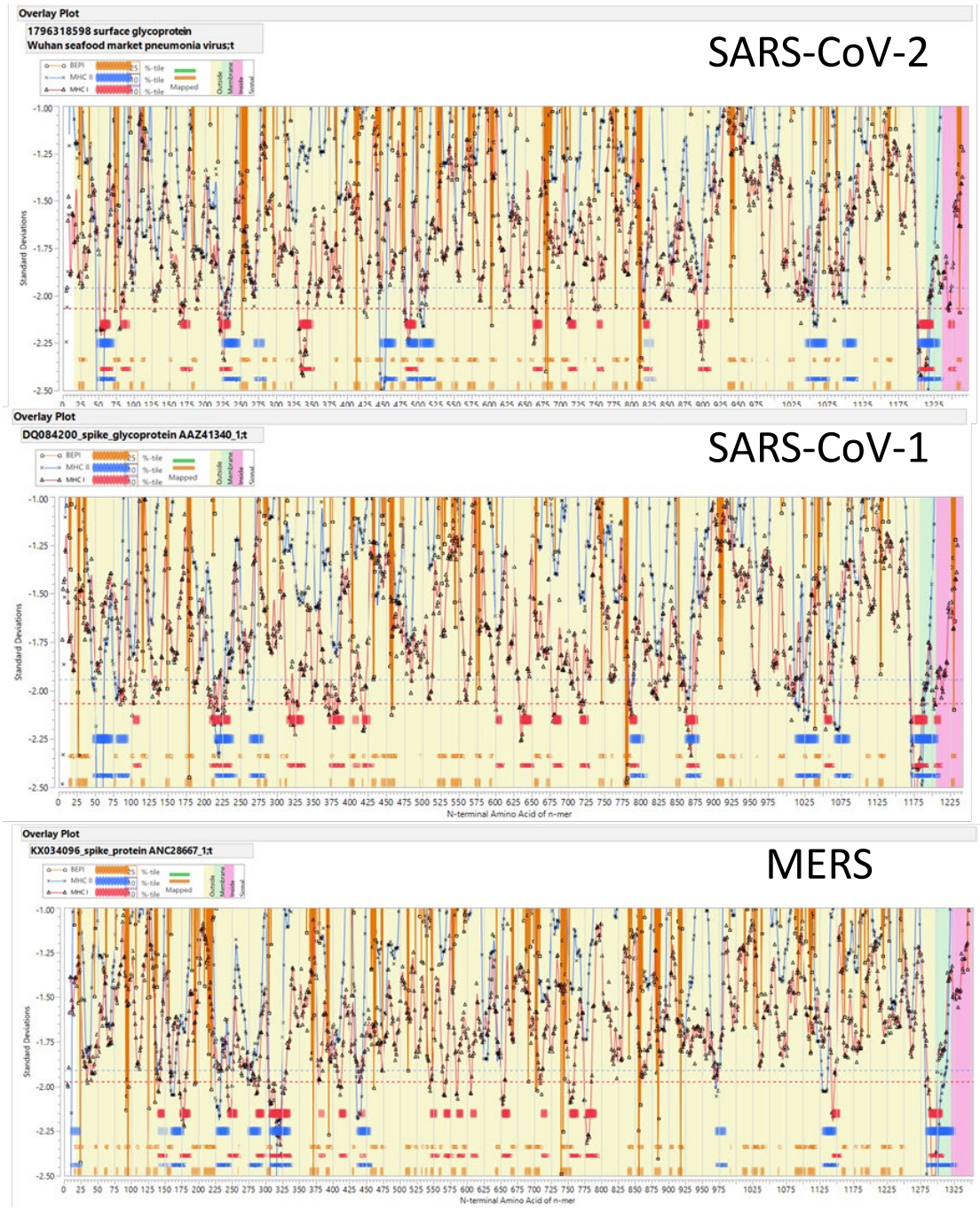


Figure 1: Comparative overview spike glycoprotein of betacoronaviruses.

Sequential peptides are arrayed and numbered N-C. Positions are for the index amino acid of each sequential peptide, with a single amino acid displacement. The Red line shows the permuted average predicted MHC-IA and B (37 alleles) binding affinity by index position of sequential 9-mer peptides. The Blue line shows the permuted average predicted MHC-II DRB allele (16 alleles) binding affinity by index position of sequential 15-mer peptides. Both are plotted in standard deviation units, shown on the Y axis. Orange lines show the predicted probability of B-cell binding ("BEPI") for an amino acid centered in each sequential 9-mer peptide. The standardized B-cell metric has been inverted in order to overlay the standardized probability onto the same scale as the MHC data. Low numbers for MHC data represent high binding affinity, whereas low numbers equate to high BEPI contact probability. Ribbons (Red: MHC-I, Blue: MHC-II) indicate the 10% highest predicted MHC affinity binding. Orange ribbons indicate the top 25% predicted probability B-cell binding. Background shading shows membrane (green) extramembrane (yellow), intramembrane (pink) location. Signal peptides have white background. The Y axis shows predicted affinity in standard deviation units below the mean. Horizontal dotted lines demarcate the top 5% of binding affinity for the protein (red MHC I, blue MHC II).

Selection of epitopes for comparative Versiti assays

The SARS-CoV-2 proteins were remapped in greater detail once a decision was made to pursue this model (Appendix 4). As described below, volunteers who had recovered from SARS-CoV-2 infection were recruited and peripheral blood mononuclear cells (PBMCs) and serum collected for assay of antibody and T cell responses. In preparation for these assays, immunoinformatic analysis of the SARS-CoV-2 Reference Strain was used to select a set of peptides for use in each assay.

It should be emphasized that the goal of the study, and hence the choice of peptides, was to evaluate the immunoinformatics platform as a predictor of *in vivo* immune responses. It was not a study designed to provide a complete profile of the immune response to SARS-CoV-2. Nevertheless, some interesting observations on COVID did emerge, as discussed below.

B cell epitopes (BEPI)

SARS CoV-2 B cell epitopes were assayed using the Pepperprint® microarray system (www.Pepperprint.com). This microarray utilizes sequential 15-mer peptides anchored at their C terminal end. Pepperprint® offers full COVID proteome arrays, but also custom arrays which permit analysis of 5 samples on one slide. For budgetary reasons the latter approach was used. Each array accommodates 480 peptides in duplicate plus controls. We therefore elected to focus on virus structural proteins plus nsp3, which had already been shown by others to elicit antibodies [28]. The structural proteins (spike glycoprotein, membrane, envelope, nucleocapsid plus ORF3a and nsp3 proteins) together comprise 4209 sequential peptides. To down-select to within the 480 limit, 15-mer peptides were selected which were predicted to lie in the top 10% of linear B cell epitope probability (-1.3 standard deviation units below the mean) for each protein and then a random subset of predicted negative peptides (those with probability above the mean), This resulted in 312 peptides predicted as spanning medium to strong probability B cell linear epitopes plus 165 predicted negatives (35%). A listing of peptides in the array is provided in Appendix 3.

All peptides in a protein are evaluated for their potential of being recognized by B cell receptors and thus being a B cell epitope (BEPI). The predictions are derived using neural networks and training sets of known B cell epitopes found in public databases using a sliding window of 9 amino acids. Thus, the predictions are for linear B cell epitopes and not for structurally derived epitopes. It is known that the majority of B cell epitopes detected are linear epitopes and further that a linear amino acid pentamer spans the 3 complementarity determining regions (CDR) of a BCR and thus forms the smallest recognizable protein unit [29]. From structural models it is clear that molecular protrusions are scored highest by the neural networks and further that in antigens whose structures are known, adjacent protrusions scored highly as linear epitopes will often comprise a structural epitope.

As described above the parameters used for predictions is the first three physical property PC. The metric predicted by the neural networks is the probability that the central amino acid of each 9-mer is likely to be recognized by a B cell (i.e. amino acid 5 of each 9-mer, being flanked by the

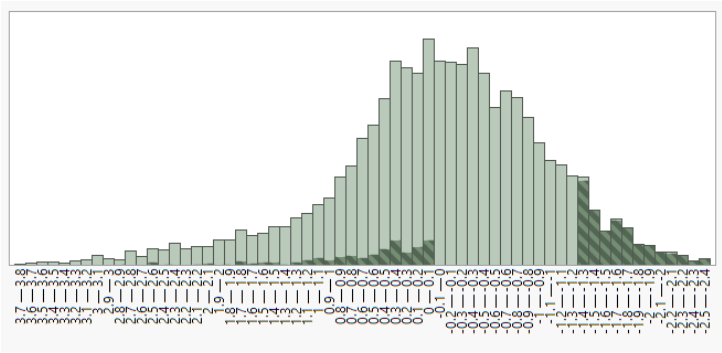


Figure 2: Peptides from SARS-CoV-2 proteins selected for inclusion on the Pepperprint chips to comprise the top 10% (<-1.3 SD) and a random subset of predicted negatives (>0 SD)

particular four amino acids on the N and C terminal sides). This probability is then standardized to a zero mean unit variance distribution within-protein using a Johnson Distribution algorithm. For convenience in plotting and analysis alongside MHC binding the number is also inverted (multiplied by -1). This process puts all proteins on a common Y-axis scale that ranks all the peptides within the molecule. Prior work [13] has demonstrated that using a -1σ threshold, (i.e. approximately 16%-tile) is an effective means of selecting potential protruding regions that might be recognized by a B cell.

MHC I and II binding peptides.

The cost and assay logistics of conducting epitope specific *in vitro* assays with volunteers of many different alleles required that we select a library of peptides for use in assays from the overall MHC binding maps of the SARS-CoV-2 proteins. Peptides were selected representing the surface glycoprotein and nucleocapsid and also the orf3A and orf1a proteins. While high affinity MHC II binding peptides tend to cluster together for all alleles, those binding MHC I alleles are typically more distributed.

A listing of 24 peptides comprising seventeen 9 mers and seven 10 mers including high predicted MHC I binding affinity for common alleles and a small set of predicted non binders for common alleles were identified to assay CD8+ responses. Similarly, a set of seventeen 16-mer and one 15-mer peptides corresponding to predicted high affinity binders and negative controls for common MHC II alleles were identified to assay CD4+ responses.

The peptides are shown in Tables 4-6. MHC binding was identified for 31 MHC IA, 31MHC IB 8 MHC IC, 24 MHC II DRB. Given the limited space only those alleles subsequently identified in the volunteers are shown in the table. As predictions in EigenBio are made for 9-mer and 15-mer peptides, pairs of peptides are shown where a 9mer or 15mer is extended by one peptide position to be used as a 10 or 16 mer. Units are standard deviation (SD) units about the mean for the particular protein of origin, a recognition that binding is a competitive process. More negative numbers are thus higher predicted affinity. While each protein varies typically -1SD represents a 200 nanomolar and -2SD units represents 25-50 nanomolar binding depending on allele

This library of peptides was synthesized (Mimotope Inc) and held in readiness to use in the *in vitro* assays described in Task III. As the HLA alleles of each volunteer were identified, we

selected a set of peptides from within the library of synthetized peptides to assay for that patient to ensure predicted positive and negative epitopes were included for that patient's alelles. The selection was made to include binding alleles for MHC IA and IB and DRB alleles and to include some negative controls. The list of peptides to use for each volunteer was provided to the lab without the associated binding predictions to ensure an unbiased assay.

Comparative analysis of emerging mutant strains

With the emergence of the United Kingdom (B.1.1.7 20I/501Y.v1 VOC 202012/01) and South African (B.1351 20H/501Y.v2) mutant strains in late 2020/early 2021 we compared epitope profiles of the spike glycoproteins and nucleocapsid from representatives of these strains with the corresponding Reference Sequences (gi 1796318598, 1798174255). This is unrelated to the modeling of responses in the volunteers, but was considered relevant to evaluate potential cross protection.

The mutations in the region of the receptor binding domain are known to bring about conformational changes which appears to increase binding affinity to the angiotensin-converting enzyme-2 (ACE2) receptor and thus increase infectivity [30]. We looked at how they may affect dominance and qualitative aspects of the B and T cell epitopes. The N501Y mutation decreases the relative dominance of a linear B cell epitope, linked to diminished antibody reactivity. No significant changes occur in the predicted binding affinity to MHC I and MHC II molecules of the mutated peptides, although there are minor differences in predicted binding affinity for individual HLA alleles, as shown in Appendix 5. Nevertheless, the amino acid substitution results in subtle qualitative differences, as the T cell clones cognate to the T cell exposed motifs for one strain do differ from those of the other strain. The T cell exposed motifs of the mutant strains comprise less common T cell exposed motifs than the Reference Sequence which indicates the probability of a lower precursor frequency of T cell clones which can respond to these strains in a naïve individuals, potentially slowing the T cell response. The mutations do not affect regions which emerge as the dominant B and T cell epitopes and thus we predict are unlikely to significantly affect vaccine efficacy for those vaccines which include all or most of the spike glycoprotein molecule.

In the nucleocapsid, the R203K, G204R (UK) and T205I (S. Africa) mutations lie in the most dominant B cell epitope and may impact cross binding of antibody to this epitope, but do not differ much in quantitative dominance. The T cell epitopes in the nucleocapsid that are highly conserved across coronaviruses are unaffected [31]. Currently FDA approved vaccines for SARS-CoV-2 do not include the nucleocapsid protein, as do very few in development (https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines).

The literature accumulating as of the date of this report indicates that while the UK and South African mutants may increase infectivity, there is still significant cross protection to these strains by antibodies from earlier infection and vaccination based on the reference sequence [32-34]

Illustrations and tabulations of the comparative mutant strain analysis is provided in Appendix 5.

Table 4: Peptides selected to assay MHC IA binding

Peptides in white cells are 9mers; pairs of peptides in grey cells were used as a continuous 10mer. Predicted binding is shown in standard deviation units relative to all peptides in that protein.

pos	curation	9-mer	A_0101	A_0201	A_0301	A_1101	A_2601	A_3001	A_3101	A_3201	A_3301	A_6801	
192	surface glycoprotein	FVFKNIDGY	-0.63	-0.29	-2.19	-1.63	-4.05	0.39	-0.20	-0.73	-0.40	-1.35	
194	surface glycoprotein	FKNIDGYFK	-0.46	-0.10	-0.56	-1.78	1.29	-0.87	-1.64	1.15	-0.55	-0.85	
269	surface glycoprotein	YLQPRTFLL	-0.77	-3.37	-0.29	-0.26	0.75	-0.56	-0.23	-1.33	0.73	0.76	
270	surface glycoprotein	LQPRTFLLK	-0.28	0.57	-2.17	-2.64	-0.07	-1.14	-2.12	1.49	-1.71	-1.21	
348	surface glycoprotein	ASVYAWNRK	0.54	0.95	-2.22	-2.95	0.17	-0.34	-2.02	1.74	-0.78	-1.87	
349	surface glycoprotein	SVYAWNRKR	1.15	0.16	-1.62	-2.47	-0.80	-0.02	-3.05	-0.22	-3.05	-1.93	
438	surface glycoprotein	SNNLDSKVG	0.53	1.63	0.43	0.44	1.94	-0.37	1.01	0.53	1.12	1.58	neg control
449	surface glycoprotein	YNYLYRLFR	-0.88	0.09	-2.05	-2.41	-0.78	-1.34	-3.20	1.38	-3.57	-2.36	
450	surface glycoprotein	NYLYRLFRK	0.57	0.58	-1.82	-2.21	0.24	-0.40	-2.52	0.46	-2.05	-1.08	
484	surface glycoprotein	EGFNCYFPL	-0.37	-1.62	-1.19	-1.00	-1.95	-1.09	-0.87	-2.14	-1.13	-1.41	
509	surface glycoprotein	RVVVLSFEL	-0.43	-2.17	-0.87	-0.95	-1.06	-0.15	-0.37	-2.40	-0.04	-0.65	
554	surface glycoprotein	ESNKKFLPF	-1.76	0.38	0.71	0.11	-3.19	0.21	0.12	-1.23	-0.25	-0.56	
626	surface glycoprotein	ADQLTPTWR	-0.01	0.61	-1.30	-1.93	-0.19	-0.41	-2.96	3.16	-2.51	-1.39	
826	surface glycoprotein	VTLADAGFI	0.20	-1.27	0.94	0.04	-1.65	1.15	-0.07	-0.65	0.63	-0.13	
827	surface glycoprotein	TLADAGFIK	-0.03	-0.09	-3.01	-2.88	0.22	-0.74	-0.93	0.88	-1.21	-1.69	
888	surface glycoprotein	FGAGAALQI	-1.48	-1.84	0.50	1.41	-0.39	0.86	1.27	-1.09	0.34	-0.01	
889	surface glycoprotein	GAGAALQIP	-1.87	0.78	-2.13	-2.13	-0.54	0.00	1.13	0.87	0.81	-0.51	
1026	surface glycoprotein	ATKMSECVL	0.58	-0.23	1.68	0.12	1.07	-0.34	0.13	-0.86	0.32	0.34	neg control
89	ORF3a protein	TVYSHLLLV	-0.48	-2.39	-0.22	-0.77	-1.97	-1.63	-1.24	-1.25	-0.43	0.56	
90	ORF3a protein	VYSHLLLVA	-0.47	-0.39	0.95	0.49	0.64	-0.48	0.22	0.54	-0.07	0.69	
103	ORF3a protein	APFLYLYAL	0.32	-1.10	0.16	-0.49	-0.59	0.01	-0.49	-1.74	-0.49	0.69	
2098	orf1a polyprotein	AYVDNSSLT	1.02	0.33	1.08	0.69	0.66	0.51	0.24	0.56	1.86	1.15	neg control
2321	orf1a polyprotein	FGLVAEWFL	-0.68	-2.17	-0.33	-0.87	-0.09	-0.66	-0.12	-1.36	-0.86	-1.22	
2328	orf1a polyprotein	FLAYILFTR	-0.93	-0.58	-2.76	-2.03	-0.70	0.78	-2.97	0.88	-3.35	-2.22	
2366	orf1a polyprotein	WLIINLVQM	-0.56	-2.23	0.43	-0.17	-1.34	0.54	-0.33	-0.49	-0.61	0.85	
102	nucleocapsid	KDLSPRWYF	-0.02	-1.25	0.08	-0.31	0.57	-0.04	-0.76	-2.18	-0.74	-0.62	

112	nucleocapsi	YLGTGPEAG	-0.20	-2.46	0.02	0.47	0.42	0.54	0.84	-0.19	0.42	0.95	
114	nucleocapsid	GTGPEAGLP	-1.16	-0.13	-1.65	-1.64	-2.07	1.14	0.69	0.10	1.05	-0.17	
115	nucleocapsid	TGPEAGLPY	-1.46	-0.16	-1.68	-1.89	-2.45	-0.67	0.06	-0.24	-0.63	-1.53	
197	nucleocapsid	STPGSSRGT	1.42	0.28	0.59	0.58	0.65	0.57	1.42	-0.33	0.18	0.42	neg control
218	nucleocapsid	ALALLLLDR	-0.42	-0.30	-2.54	-2.11	-0.27	1.01	-2.20	0.92	-1.72	-1.03	

Table 5: Peptides selected to assay MHC IB binding

Peptides in white cells are 9mers; pairs of peptides in grey cells were used as a continuous 10mer. Predicted binding is shown in standard deviation units relative to all peptides in that protein.

pos	curation	9-mer	B_0702	B_0801	B_1501	B_1801	B_2703	B_2705	B_3501	B_3901	B_4402	B_5101	B_5701	B_5801	
192	surface glycoprotein	FVFKNIDGY	-1.17	-1.02	-2.72	-0.69	-0.56	-0.39	-2.21	0.13	0.03	-0.22	-1.82	-2.00	
194	surface glycoprotein	FKNIDGYFK	1.19	0.83	0.84	0.12	0.38	-2.49	0.12	0.34	-0.04	-0.52	0.41	1.22	
269	surface glycoprotein	YLQPRTFLL	-1.90	-2.91	-1.18	-0.21	0.02	0.35	-1.40	-1.23	-0.29	-0.08	-0.20	0.93	
270	surface glycoprotein	LQPRTFLLK	0.53	0.20	-0.50	-0.42	0.59	-1.41	0.36	-0.01	-0.09	0.41	-0.09	0.84	
348	surface glycoprotein	ASVYAWNRK	1.25	0.41	1.13	0.34	-0.46	-0.08	1.15	1.89	-0.04	1.05	-0.38	-0.53	
349	surface glycoprotein	SVYAWNRKR	0.12	0.89	1.58	0.21	0.00	-0.15	0.15	2.24	0.24	0.73	-1.14	1.22	
438	surface glycoprotein	SNNLDSKVG	0.93	0.10	-0.54	5.02	0.43	1.53	1.92	0.47	0.64	0.28	1.20	0.82	neg control
449	surface glycoprotein	YNYLYRLFR	0.22	0.71	-0.70	-1.38	-0.45	-3.06	-0.61	0.94	-0.97	0.56	-1.66	-0.42	
450	surface glycoprotein	NYLYRLFRK	1.03	-0.37	1.99	-0.32	0.07	-0.55	1.11	1.04	0.56	0.20	0.25	1.19	
484	surface glycoprotein	EGFNCYFPL	-1.72	-1.72	-0.15	-1.58	-0.13	-0.54	-1.36	-1.45	-0.13	-2.10	-3.10	-2.91	
509	surface glycoprotein	RVVVLSFEL	-0.65	-0.17	-1.13	-0.36	-1.03	0.20	-1.41	-2.22	-0.80	-0.54	-1.34	-2.75	
554	surface glycoprotein	ESNKKFLPF	-0.71	-2.37	-1.86	-0.68	0.11	-0.24	-1.57	0.26	-0.28	-0.25	-2.94	-2.55	
626	surface glycoprotein	ADQLTPTWR	0.69	2.00	0.43	-0.18	0.06	-0.63	0.99	0.64	-1.41	0.20	-0.27	0.07	
826	surface glycoprotein	VTLADAGFI	0.49	-0.46	-0.76	-0.36	-0.23	0.68	-0.76	-0.44	0.04	-1.53	-0.32	-1.01	
827	surface glycoprotein	TLADAGFIK	0.83	0.62	0.22	0.30	-0.56	0.86	-0.09	0.83	0.37	0.32	0.14	0.63	
888	surface glycoprotein	FGAGAALQI	-0.89	-0.25	-1.38	-0.13	0.10	1.50	-1.99	-1.39	-0.59	-2.95	-1.98	-3.12	
889	surface glycoprotein	GAGAALQIP	0.35	0.90	1.04	0.44	1.26	0.57	0.07	0.50	-0.42	-0.15	-0.51	-0.62	
1026	surface glycoprotein	ATKMSECVL	-1.05	-0.39	-0.81	0.53	-0.13	0.46	0.55	-1.32	2.02	-0.09	-0.46	-0.57	
89	ORF3a protein	TVYSHLLLV	-0.62	-2.27	-0.67	0.01	-0.36	-0.45	0.43	-0.28	0.21	-0.05	-0.33	0.26	

90	ORF3a protein	VYSHLLLVA	0.58	-1.08	0.89	0.01	0.50	0.84	-0.22	-0.67	-0.65	-0.97	2.02	0.37	
103	ORF3a protein	APFLYLYAL	-3.53	-2.27	0.24	-0.62	-0.27	-0.71	-1.17	-1.45	0.09	-3.31	-0.91	-0.41	
2098	orf1a polyprotein	AYVDNSSLT	0.43	0.76	1.25	0.32	0.39	1.29	0.68	0.00	0.96	0.36	1.70	1.45	neg control
2321	orf1a polyprotein	FGLVAEWFL	-1.20	-0.38	-0.62	-1.14	-1.17	-0.13	-1.92	-2.17	-0.70	-2.11	-3.19	-3.07	
2328	orf1a polyprotein	FLAYILFTR	0.50	0.98	0.16	-0.91	0.31	-0.19	-1.26	1.44	-0.66	-0.10	-0.51	-1.11	
2366	orf1a polyprotein	WLIINLVQM	-0.18	-1.58	-2.25	-0.97	0.26	-0.08	-2.05	-0.28	-0.64	-1.50	-0.71	-1.94	
102	nucleocapsid	KDLSPRWYF	-0.24	-0.62	-2.21	-2.23	-0.40	-2.68	-1.19	-1.03	-3.40	-0.81	-3.60	-2.99	
112	nucleocapsid	YLGTGPEAG	-1.08	-0.81	-0.20	1.27	-0.33	0.80	-0.88	-0.84	-0.72	0.08	0.56	-0.42	
114	nucleocapsid	GTGPEAGLP	0.61	0.22	0.38	0.02	-0.35	1.34	0.32	0.67	0.70	1.10	0.63	-0.14	
115	nucleocapsid	TGPEAGLPY	0.33	1.21	-2.17	-1.19	-0.31	0.20	-2.10	1.28	-1.14	-0.17	-2.86	-2.92	
197	nucleocapsid	STPGSSRGT	0.64	1.66	0.57	0.72	-0.47	0.62	0.70	0.08	-0.52	-0.40	0.17	0.13	neg control
218	nucleocapsid	ALALLLLDR	1.58	1.26	0.72	-1.53	-0.22	0.54	-0.34	1.11	-0.48	-0.09	-0.64	0.17	

Table 6: Peptides selected to assay MHC II DRB binding

Peptides paired in white cells together comprise 16mers; one peptide shown in grey is a 15merr. Predicted binding is shown in standard deviation units relative to all peptides in that protein.

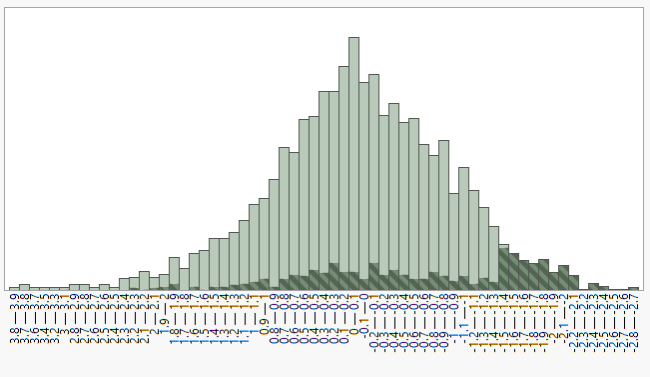
pos	curation	peptide	DRB1 0101	DRB1 0301	DRB1 0401	DRB1 0404	DRB1 0701	DRB1 0801	DRB1 0901	DRB1 1101	DRB1 1301	DRB1 1501	DRB1 1602	Column 11
269	surface glycoprotein	YLQPRTFLLKYNENG	-1.15	0.05	-1.38	-2.26	-0.91	-2.32	-0.83	-1.85	-1.65	-1.33	-0.48	
270	surface glycoprotein	LQPRTFLLKYNENGT	-0.18	0.43	-1.15	-1.01	0.32	-1.74	-0.40	-1.66	-1.37	-0.92	1.25	
435	surface glycoprotein	AWNSNNLDSKVGGNY	1.43	1.48	0.84	1.02	1.54	1.17	0.60	0.93	0.79	2.61	-0.32	neg control
436	surface glycoprotein	WNSNNLDSKVGGNYN	1.33	1.13	0.86	1.76	0.85	1.77	1.18	1.05	1.55	0.92	1.51	
449	surface glycoprotein	YNYLYRLFRKSNLKP	-0.81	-1.74	-2.46	-1.13	-0.95	-2.58	-0.41	-2.80	-2.90	-2.03	0.70	
450	surface glycoprotein	NYLYRLFRKSNLKPF	-1.29	-1.58	-2.74	-2.17	-2.27	-2.00	-0.53	-4.15	-1.87	-1.61	-0.84	
504	surface glycoprotein	GYQPYRVVVLSFELL	-1.15	-0.64	-2.64	-1.67	-2.53	-1.24	-2.43	-1.20	-0.96	-1.06	-0.53	
505	surface glycoprotein	YQPYRVVVLSFELLH	-1.57	-1.13	-0.70	-0.76	-1.98	-1.57	-0.74	-1.10	-0.65	-1.44	-0.93	

555	surface glycoprotein	SNKKFLPFQQFGRDI	-0.25	-0.06	-0.85	-0.34	-0.90	-2.10	-0.89	-1.97	-2.11	-0.88	-1.93	
556	surface glycoprotein	NKKFLPFQQFGRDIA	-1.47	-1.46	-0.85	-0.26	-1.16	-1.72	-0.55	-1.73	-1.41	-2.15	-0.47	
672	surface glycoprotein	ASYQTQTNSPRRARS	1.50	0.30	0.98	1.08	1.63	1.13	1.95	0.30	-0.80	1.11	0.42	neg control
673	surface glycoprotein	SYQTQTNSPRRARSV	2.13	0.54	1.60	1.43	1.92	0.63	1.45	0.01	-0.35	1.78	-0.57	
819	surface glycoprotein	EDLLFNKVTLADAGF	-1.34	-0.02	-2.38	-0.36	-1.38	-1.06	-0.35	-0.75	-0.76	-0.62	-0.83	
820	surface glycoprotein	DLLFNKVTLADAGFI	-2.60	0.20	-2.00	-1.64	-2.87	-0.93	-2.46	-1.75	-0.88	-0.58	-0.65	
882	surface glycoprotein	ITSGWTFGAGAALQI	-1.86	0.73	-0.33	-1.54	-1.79	-0.17	-2.69	-0.47	0.37	-0.36	-0.23	
883	surface glycoprotein	TSGWTFGAGAALQIP	-3.48	0.38	-1.85	-1.60	-2.26	-0.21	-2.16	-0.34	0.17	-0.10	-0.49	
93	ORF3a protein	HLLLVAAGLEAPFLY	-1.86	-0.94	-1.26	-0.51	-0.60	-0.36	-1.10	-0.28	-0.54	-0.21	-0.41	
94	ORF3a protein	LLLVAAGLEAPFLYL	-0.48	-0.72	0.09	-0.80	-1.30	-0.58	-0.86	-0.02	-1.12	-0.62	-0.22	
104	ORF3a protein	PFLYLYALVYFLQSI	-0.36	-1.00	-1.79	-1.79	-1.09	-2.29	-0.63	-1.46	-3.57	-0.46	-1.60	
105	ORF3a protein	FLYLYALVYFLQSIN	-0.31	-0.50	-3.12	-2.46	-0.49	-1.95	-1.38	-1.61	-1.70	-1.40	-0.54	
2101	orf1a polyprotein	DNSSLTIKKPNELSR	1.79	0.56	1.27	1.31	1.51	0.32	1.61	0.15	0.03	0.33	0.82	neg control
2102	orf1a polyprotein	NSSLTIKKPNELSRV	0.56	-0.79	0.49	0.30	0.67	-0.11	0.98	-0.91	-0.17	0.10	0.99	
2325	orf1a polyprotein	AEWFLAYILFTRFFY	-0.59	-1.91	-1.36	-1.92	-1.50	-2.16	-1.58	-2.03	-2.22	-2.33	-0.79	
2326	orf1a polyprotein	EWFLAYILFTRFFYV	-1.39	-2.74	-1.15	-2.12	-2.39	-2.71	-1.85	-2.75	-2.57	-2.53	-0.94	
2362	orf1a polyprotein	SWLMWLIINLVQMAP	-0.62	-2.47	-2.41	-2.99	-2.75	-2.54	-2.62	-2.72	-1.85	-2.01	-0.84	
2363	orf1a polyprotein	WLMWLIINLVQMAPI	-1.43	-2.26	-2.48	-2.69	-2.58	-2.61	-1.46	-2.94	-2.44	-2.06	-0.25	
2383	orf1a polyprotein	MYIFFASFYYVWKSY	-1.29	-2.10	-3.26	-2.04	-1.27	-2.10	-0.97	-2.29	-2.74	-2.05	1.02	
2384	orf1a polyprotein	YIFFASFYYVWKSYV	-0.85	-1.61	-1.57	-2.26	-1.02	-2.41	-2.30	-2.58	-2.65	-2.59	-0.80	

106	nucleocapsid	PRWYFYYLGTGPEAG	-1.95	-0.41	-2.96	-2.56	-2.25	-2.50	-2.46	-2.63	-1.70	-2.29	-2.63	
108	nucleocapsid	WYFYYLGTGPEAGLP	-3.27	-0.68	-2.69	-1.87	-2.83	-1.95	-3.33	-2.38	-1.13	-1.74	-0.70	
109	nucleocapsid	YFYYLGTGPEAGLPY	-2.19	-1.27	-2.02	-1.36	-1.71	-1.25	-1.73	-1.48	-1.50	-1.37	-1.86	
196	nucleocapsid	NSTPGSSRGTSPARM	0.74	0.46	0.30	0.79	0.49	1.34	0.51	1.21	0.62	1.83	-1.18	
197	nucleocapsid	STPGSSRGTSPARMA	0.48	1.03	1.14	0.10	0.57	0.70	-0.28	1.48	0.96	0.78	-1.37	
216	nucleocapsid	DAALALLLLDRLNQL	-1.24	-0.62	-1.47	-2.41	-1.38	-2.50	-0.60	-1.89	-2.75	-2.31	1.06	
217	nucleocapsid	AALALLLLDRLNQLE	-1.25	-1.59	-1.11	-2.23	-1.50	-2.42	-0.71	-2.48	-2.01	-2.06	-1.59	

Selection of peptides for epitope mapping of yellow fever analysis in mice

The epitope maps of YF VAX® were evaluated to select peptides for analysis of T cell responses and antibody responses.



gi	N Rows
core_protein_C	16
envelope_protein	71
matrix_protein_M	2
non_structural_protein_NS1	41
non_structural_protein_NS2a	24
non_structural_protein_NS2b	11
non_structural_protein_NS3	115
non_structural_protein_NS4a	13
non_structural_protein_NS4b	29
RNA_dependent_RNA_polymerase	132

Figure 3: Selection of predicted B cell linear peptides in YF-VAX®.

As with SARS-CoV-2, budgetary considerations dictated that we use the custom arrays each comprising 480 unique peptides in duplicate. We therefore selected peptides from all viral proteins with probabilities of being in a B cell linear epitope in the top 10% plus a random subset of all peptides with lower binding affinity (Figure 3).

For assessment of T cell responses a subset of peptides was made based on their predicted binding to Kb Db and H2IAb alleles of the C57BL/6 mice, but also with consideration of whether the same peptides had a likelihood of MHC binding to future human donors. Peptides selected are included in Table 7 (MHC I) and Table 8 (MHC II). Predicted binding affinity is shown in standard deviation units below the mean for that protein, with values below -1 being considered positives (equivalent to <200 nanomolar).

Table 7: YF Vax® peptides predicted affinity for murine MHC I alleles Kb and Db

Peptides -	gi	pos	peptide MHC II	Db	Kb	Negative controls
E106	envelope_protein	106-114	GLFGKGSIV	-1.90	-1.36	
E197	envelope_protein	197-205	EMETESWIV	-0.86	0.02	
E213	envelope_protein	213-221	LTLPWQSGS	0.92	-1.61	
E264	envelope_protein	264-272	VTKDTNDNN	2.84	0.82	neg
E298	envelope_protein	298-306	KICTDKMFF	-2.19	-0.04	
E326	envelope_protein	326-334	KGAPCRIPV	-1.46	-0.12	
E327	envelope_protein	327-335	GAPCRIPVI	-1.49	-2.28	
E336	envelope_protein	336-344	VADDLTAAI	-0.47	-1.84	

E354	envelope_protein	354-362	PIASTNDDE	0.19	1.94	neg
E398	envelope_protein	398-306	KLFTQTMKG	-0.85	-0.34	
1NS116	non_structural_protein_NS1	116-124	KTWGKNLVF	-1.39	-2.17	
1NS128	non_structural_protein_NS1	128-136	RKNGSFIID	0.98	1.51	neg
1NS211	non_structural_protein_NS1	211-220	WMIHTLEAL	-2.56	-1.76	
1NS243	non_structural_protein_NS1	243-251	MPRSIGGPV	0.46	-1.26	
1NS253	non_structural_protein_NS1	253-261	SHNHIPGYK	2.00	-0.02	neg
1NS48	non_structural_protein_NS1	48-56	ASFEEGKCG	-1.44	0.02	
3NS157	non_structural_protein_NS3	157-165	ILVGDNSFV	-2.74	-1.04	
3NS161	non_structural_protein_NS3	161-169	DNSFVSAIS	0.87	1.31	neg
3NS267	non_structural_protein_NS3	267-275	HATLTYRML	-1.27	-0.75	
3NS287	non_structural_protein_NS3	287-295	IMDEAHFLD	1.04	0.74	neg
3NS382	non_structural_protein_NS3	382-390	KAGKSVVVL	-1.19	0.34	
3NS457	non_structural_protein_NS3	457-465	SAAQRRGRI	-0.84	-2.39	
3NS540	non_structural_protein_NS3	540-548	LVRNCDLPV	-0.58	-1.74	
POL126	RNA_dependent_RNA_polymerase	126-134	FKDKTDIHR	2.12	1.48	neg
POL167	RNA_dependent_RNA_polymerase	167-175	TVEKWLACG	1.68	-1.70	
POL175	RNA_dependent_RNA_polymerase	175-183	GVDNFCVKV	-0.28	-1.87	
POL25	RNA_dependent_RNA_polymerase	25-33	ELYKRTDIV	-0.72	-1.18	
POL251	RNA_dependent_RNA_polymerase	251-259	VTLEADVIL	-1.15	-1.70	
POL308	RNA_dependent_RNA_polymerase	308-316	DPKFWELVD	1.97	0.04	
POL712	RNA_dependent_RNA_polymerase	712-720	FCSHHFHEL	-1.49	-1.36	
POL856	RNA_dependent_RNA_polymerase	856-862	GMTNRATWA	-0.06	-1.24	
Ova	Ovalbumin		PINFQTAAD	2.02	2.42	

Table 8: YF Vax® peptides predicted affinity for murine MHC II allele H2-1Ab

Peptides-	gi	pos	peptide MHC II	IAb	Negative controls
E357	envelope_protein	357-371	STNDDEVLIEVNPPF	1.73	neg
E261	envelope_protein	261-275	AMRVTKDTNDNNLYK	1.19	neg
E231	envelope_protein	231-245	LVEFEPPHAATIRVL	-1.18	
E209	envelope_protein	209-228	WAQDLTLPWQSGSGGVWREM	-2.91	
E238	envelope_protein	238-255	HAATIRVLALGNQEGSLK	-1.93	
E301	envelope_protein	301-318	TDKMFFVKNPTDTGHGTV	-1.82	
E345	envelope_protein	345-360	NKGILVTVNPIASTND	-1.62	
E383	envelope_protein	383-400	SRLTYQWHKEGSSIGKLF	-1.79	

E414	envelope_protein	414-429	GDTAWDFSSAGGFFTS	-2.23	
E75	envelope_protein	75-91	PSTGEAHLAEENEGDNA	1.07	neg
1NS109	non_structural_protein_NS1	109-125	DGLQYGWKTWGKNLVFS	-1.11	
1NS112	non_structural_protein_NS1	112-128	QYGWKTWGKNLVFSPGR	-2.46	
1NS208	non_structural_protein_NS1	208-225	NGTWMIHTLEALDYKECE	-1.20	
1NS238	non_structural_protein_NS1	238-255	ESEMFMPRSIGGPVSSHN	-1.79	
1NS265	non_structural_protein_NS1	265-279	NGPWMQVPLEVKREA	-0.79	
1NS290	non_structural_protein_NS1	290-306	NCDGRGKSTRSTTDSGK	-0.16	neg
1NS30	non_structural_protein_NS1	30-46	NKYSYYPEDPVKLASIV	-2.36	
1NS34	non_structural_protein_NS1	34-51	YYPEDPVKLASIVKASFE	-1.76	
1NS1307	non_structural_protein_NS1	307-321	VIPEWCCRSCTMPPV	1.96	neg
1NS1326	non_structural_protein_NS1	326-340	SDGCWYPMEIRPRKT	1.85	neg
3NS169	non_structural_protein_NS3	169-183	SQTEVKEEGKEELQE	1.53	neg
3NS469	non_structural_protein_NS3	469-482	PNRDGDSYYYSEPTS	1.43	neg
3NS127	non_structural_protein_NS3	127-246	GAVALDYPSGTSGSPIVNRN	-3.26	
3NS217	non_structural_protein_NS3	217-233	ARRRLRTLVLAPTRVVL	-0.91	
3NS314	non_structural_protein_NS3	314-330	SATILMTATPPGTSDEF	-2.14	
3NS359	non_structural_protein_NS3	359-375	DKRPTAWFLPSIRAANV	-1.21	
3NS404	non_structural_protein_NS3	404-421	QKKPDFILATDIAEMGAN	-1.73	
3NS64	non_structural_protein_NS3	64-80	GKKLIPSWASVKEDLVA	-1.73	
3NS75	non_structural_protein_NS3	75-89	KEDLVAYGGSWKLEG	-2.02	
POL436	RNA_dependent_RNA_polymerase	436-449	VDEERKLHQQGRCRT	1.71	neg
POL545	RNA_dependent_RNA_polymerase	545-559	EADLDDEQEILNYMS	1.40	neg
POL194	RNA_dependent_RNA_polymerase	194-212	KLELLQRRFGGTVIRNPLS	-1.33	
POL308	RNA_dependent_RNA_polymerase	308-329	GSYVTKTSGSAASMVNGVIKIL	-2.83	
POL481	RNA_dependent_RNA_polymerase	481-498	GARYLEFEALGFLNEDHW	-1.01	
POL577	RNA_dependent_RNA_polymerase	577-594	KVVKVLRPAPGGKAYMDV	-2.71	
POL671	RNA_dependent_RNA_polymerase	671-678	VRPIDDRFGLALSHLNAM	-2.09	
POL771	RNA_dependent_RNA_polymerase	771-789	FHKRDMRLLSLAVSSAVPT	-3.37	
OV324	ovalbumin	324-340	ISQAVHAAHAEINEAGR	1.07	neg

As it was not possible to access YFVax for human use we provide a discussion (below) of the correlation between our predictions and literature and Immune Epitope DataBase reports (IEDB – see iedb.org).

Task II: Recruitment and collection from human volunteers recovered from COVID

The Applied Research Laboratory at Versiti – Blood Center of Wisconsin was responsible for recruiting volunteer contributors for this study from a bank of post COVID research participants. The purpose of this bank was to develop a population of consented donors to contribute COVID Convalescent Plasma (CCP) via apheresis. Volunteers were recruited into the CCP group following physician-confirmed clinical signs of COVID plus either a molecular diagnostic test for SARS-CoV-2 virus (e.g. PCR) or 2 positive serologic tests performed by different companies. In addition, volunteers were recruited only after completing a recovery period of at least 14 days asymptomatic.

In order to use this bank, a study protocol was submitted to the Medical College of Wisconsin Institutional Review Board (MCW-IRB) outlining the goals of our study and the tools we would be using to reach out to potential contributors. Approval was subsequently sought from HRPO. All members of the COVID Bank provided data to support their diagnoses of an infection with the SARS-CoV-2 virus, the last date they experienced significant symptoms, and personal information regarding age and where to contact them. Contributors for this study were selected on the basis of three criteria:

1. If the potential contributor was female and tested positive for Human Leukocyte Antibodies (HLA), they were ineligible to donate CCP, but were still eligible for this study. Table 9, shows 20 of the 22 participants were in this population of contributors.
2. If the potential contributor donated CCP only 1 or 2 times because apheresis was a difficult process for the individual.
3. If the potential contributor had already completed the maximum allowable CCP donations. Table 9 shows that 2 of the 22 participants were from this population of contributors.

A trained phlebotomist collected the required whole blood (WB) in K_2EDTA tubes after each participant was consented to this specific study. Tubes of WB were maintained at room temperature for up to 24 hours, and then the mononuclear white cells were isolated using a specific gravity medium following a classic chemistry method. This was consistent with the prior validation of the tube use and storage periods and temperatures which was conducted in anticipation of the original vaccine studies, to permit transport from a different collection site. The isolated mononuclear white cells were then suspended in a sterile cryoprotectant solution containing 10% dimethyl sulfoxide (DMSO), aliquoted into sterile cryovials and frozen. The vials were then transferred to the lab where cells were propagated, and further work was completed.

The time lapse from last symptoms to sample collection ranged from 58-223 days (approximately 70-249 days post infection). Although 22 patients were enrolled, 13 viable sets of PBMCs were obtained. Sera were available from 18 patients. It is unclear whether factors that led to exclusion from the CCP collection may have contributed to the lower than expected recovery rate; typically it is the lipid content of blood that can interfere with PBMC collection. However, the recruitment goal into the analysis group was met.

Table 9: Characteristics of convalescent SARS-CoV-2 donors

SUBJECT ID	DATE OF CONSENT	COLLECT DATE	Year of Birth	Date of Last Symptoms	Level of Symptoms	Date of Positive Antibody Test	Ab Level	Month of Last Plasma Donation	HLA -A typing	HLA B typing	HLA-DR typing
IG-001	9/21/2020	9/21/2020	1969	4/1/2020	Unknown; not hospitalized	5/20/2020	UNK	Deferred due to anti-HLA Abs	A*11:01:01G A*26:01:01G	B*27:05:02G B*57:01:01G	DRB1*07:01:01G
IG-002	9/22/2020	9/22/2020	1980	3/28/2020	Unknown; not hospitalized	UNK (Unknown)	UNK	Deferred due to anti-HLA Abs	A*02:01:01G A*26:01:01G	B*15:01:01G B*51:01P	DRB1*04:01:01G DRB1*07:01:01G
IG-003	9/23/2020	9/23/2020	1960	5/8/2020	Severe; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*32:01:01G A*33:01:01G	B*39:06 B*78:01	DRB1*04:04:01G DRB1*11:02:01G
IG-004	9/24/2020	9/24/2020	1988	5/1/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*03:01:01G A*31:01:02G	B*07:02:01G B*44:02:01G	DRB1*01:03:01G DRB1*09:01:02G
IG-005	9/28/2020	9/28/2020	1973	8/1/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*03:01:01G A*31:01:02G	B*07:02:01G B*40:01:01G	DRB1*08:01:01G DRB1*15:01:01G
IG-006	9/29/2020	9/29/2020	1956	2/21/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*02:01:01G A*03:01:01G	B*15:01:01G B*44:02:01G	DRB1*04:01:01G
IG-007	10/1/2020	10/1/2020	1995	5/21/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*03:01:01G A*30:01:01G	B*13:02:01G B*27:02:01G	DRB1*07:01:01G DRB1*16:01:01G
IG-008	10/7/2020	10/7/2020	1962	4/16/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*01:01:01G A*02:01:01G	B*07:02:01G B*08:01:01G	DRB1*15:01:01G
IG-009	10/8/2020	10/8/2020	1964	7/27/2020	Unknown; not hospitalized	9/7/2020	UNK	Deferred due to anti-HLA Abs	A*26:01:01G A*32:01:01G	B*39:01:01G B*50:01:01G	DRB1*08:06 DRB1*13:01:01G
IG-010	10/13/2020	10/13/2020	1949	4/15/2020	Severe; on ventilator	UNK	UNK	September 2020	A*02:01:01G	B*15:01:01G B*35:01:01G	DRB1*11:04:01G DRB1*13:01:01G
IG-011	10/13/2020	10/13/2020	1950	3/25/2020	Unknown; not hospitalized	UNK	UNK	August or September 2020	A*01:01:01G A*02:01:01G	B*15:01:01G B*58:01:01G	DRB1*01:01:01G DRB1*04:01:01G

IG-012	10/19/2020	10/19/2020	1980	3/6/2020	Unknown; not hospitalized	5/28/2020	UNK	Deferred due to anti-HLA Abs			
IG-013	10/20/2020	10/20/2020	1961	3/20/2020	Unknown; not hospitalized	5/6/2020	UNK	Deferred due to anti-HLA Abs			
IG-014	10/21/2020	12/21/2020	1959	4/28/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs			
IG-015	10/28/2020	10/28/2020	1955	4/6/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs			
IG-016	11/3/2020	11/3/2020	1976	7/21/2020	Unknown; not hospitalized	7/20/2020	UNK	Deferred due to anti-HLA Abs			
IG-017	11/3/2020	11/3/2020	1968	4/20/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs			
IG-018	11/4/2020	11/4/2020	1957	9/1/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs			
IG-019	11/4/2020	11/4/2020	1966	3/26/2020	Unknown; not hospitalized	7/3/2020	UNK	Deferred due to anti-HLA Abs			
IG-020	11/10/2020	11/10/2020	1980	5/1/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs			
IG-021	11/19/2020	11/19/2020	1984	8/18/2020	Unknown; not hospitalized	UNK	UNK	Deferred due to anti-HLA Abs	A*01:01:01G A*01:01:01G	B*08:01:01G B*18:01:01G	DRB1*03:01:01G DRB1*04:03:01G
IG-022	11/23/2020	11/23/2020	1962	6/28/2020	Unknown; not hospitalized	8/13/2020	UNK	Deferred due to anti-HLA Abs	A*03:01:01G A*68:01P	B*07:02:01G B*15:01:01G	DRB1*01:01:01G DRB1*04:01:01G

Task III: Evaluation of responses in human volunteers to COVID

Versiti Blood Research Institute.

With the outbreak of COVID in the US in early 2020 treatment options were limited and fraught with misinformation. A well-established intervention for infectious disease is to use sera collected from convalescent patients known to have been exposed to the pathogen. As explained above Versiti became involved in such a project providing sera locally and to the Mayo Clinic. A subset of the volunteers was approached to provide research blood samples which were used for preparation of sera for antibody assays and PBMC for cellular assays.

In substituting the immune response to SARS-CoV-2 in place of the originally proposed Yellow Fever vaccine (YFV) or the substitute mumps vaccine, we were moving away from well-established vaccines with documented strong immune responses into an unknown situation. However, since all work on non-COVID studies was halted, this was the only path open, and as described above held out the possibility of providing useful public health data.

Makeup of the study population:
Table 9 describes the study population. In summary, the subjects were drawn predominantly from convalescent individuals who had enrolled to provide therapeutic sera but who were deferred due to measurable titers of anti-HLA antibodies or who had surpassed the permitted number of donations.

SARS-CoV-2 Antibody responses

Methods: Sera were tested for antibodies to selected peptides using Pepperprint® peptide arrays following Pepperprint® 's protocol. The negative control was commercial human AB serum (Gemini Bioproducts, West Sacramento, California).

Pepperprint® slides with 5 peptide arrays per slide were put into the PEPperCHIP Incubation Tray. The slides were equilibrated in Washing Buffer (PBS 0.05% Tween20, pH7.4) on an orbital shaker then blocked with Blocking Buffer (Rockland Buffer, Pepperprint® MB-070). After another incubation and washing, the secondary anti-body (pAb goat anti-human IgG (H+L) Cy5 (abcam ab97172) diluted in Staining Buffer (Washing Buffer with 10% of Blocking Buffer) was added to all the arrays to detect any cross-reactivity. After incubation on an orbital shaker and washing, the slides were dipped in Dipping Buffer (1mM Tris Buffer, pH 7.4) and dried. The slides were scanned with a LiCOR® Odessey imager. After seeing no cross-reaction, the serum samples were diluted in Staining Buffer and put onto the arrays. Arrays were incubated overnight at 4°C on the orbital shaker, washed, then the secondary antibody was added again. After another incubation and wash, the control antibody mouse anti-HA Cy5 was put on the arrays. Slides were incubated, washed, dipped, and dried. Final images were made with the LiCOR Odessey. Slide images were analyzed with MAPIX Analyzer software data package from Pepperprint® .

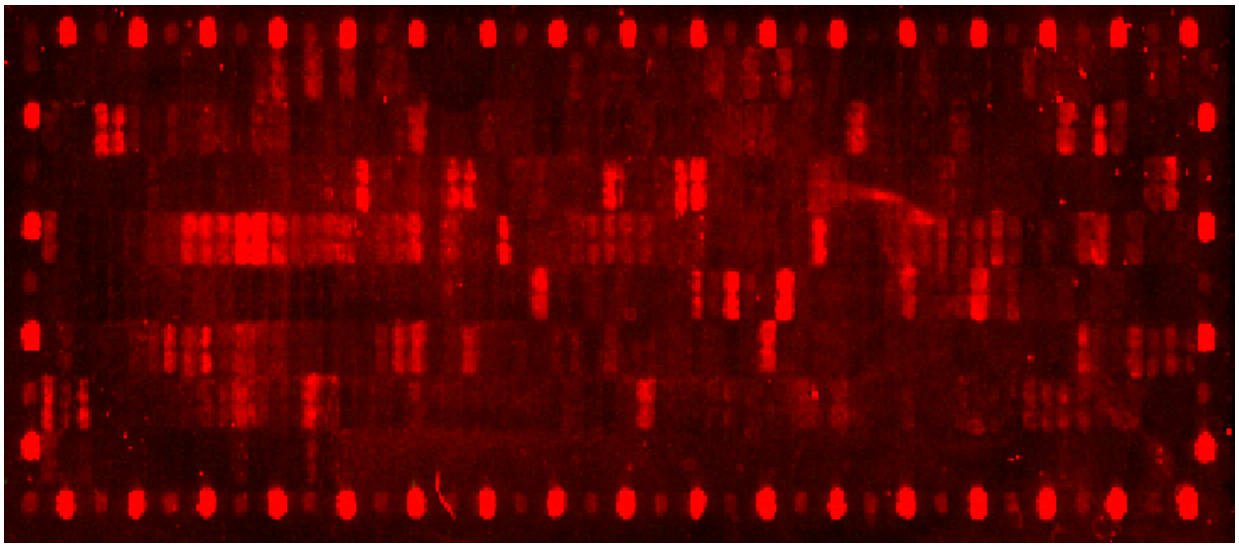


Figure 4: Pepperprint® pattern for subject IG-002 that contains 480 duplicates of SARS-CoV-2 peptides. The border of the chip contains alternating rectangles comprising attachment chemistry background and an influenza peptide widely recognized in the US population. It can clearly be seen that the subject has antibodies that recognize the influenza peptide and that there is a wide range of responses to the SARS-CoV-2 peptides. The instrument software makes multiple readings of the chip to arrive at a final output

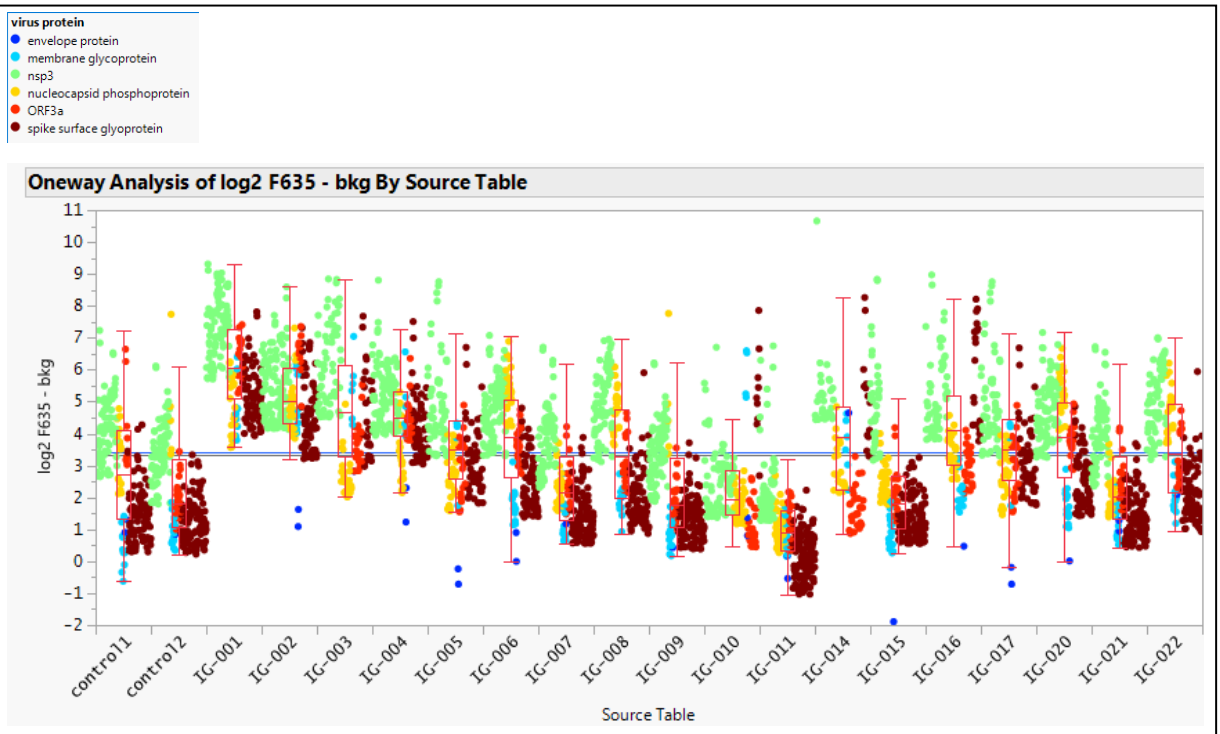


Figure 5: Variation between the antibody responses for the different subjects. It further shows that that the antibody response to the viral protein nsp3 is consistently greater than to the spike protein in all subjects. Even the two commercial antisera used as controls show a differential expression. These different responses complicate the analysis considerably.

ANOVA of peptide fluorescence and BEPI predictions.

As described above, BEPI predictions are converted to a zero mean unit variance distribution within-protein. For analysis the peptides were grouped into 5 subgroups (Figure 6). Those selected as strong BEPI candidates were in the -2σ→-3σ and -2σ→-1σ groups. This is the range described above that we have found to be most predictive of being recognized by B cells. Additionally, as negative controls, peptides (n=165) outside these groups were randomly selected within the 0σ→1σ, 1σ→2σ and 2σ→3σ groups (the most unlikely BEPI candidates). The immunoglobulin binding metric used was the log2 of the 635 nm fluorescence over background. Background is determined by the microarray positions interspersed between the influenza peptides on the perimeter of the microarray. A log2 conversion of the fluorescence data was used because the overall distribution fit of the data was lognormal. A simple linear ANOVA model with BEPI predicted group and fluorescent signal was used to test the difference between the different prediction groups. The plot of the least squares means of the fluorescence over background demonstrates that the two subgroups of predicted strong BEPI peptides exhibit a statistically significantly higher than the other

Response log2 F635 - bkg

Summary of Fit

RSquare	0.059341
RSquare Adj	0.058429
Root Mean Square Error	1.885295
Mean of Response	3.31522
Observations (or Sum Wgts)	4129

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	4	924.700	231.175	65.0403
Error	4124	14658.084	3.554	Prob > F
C. Total	4128	15582.784		<.0001*

Effect Tests

Source	Nparm	DF	Sum of Squares	F Ratio	Prob > F
BEPI Binned	4	4	924.70004	65.0403	<.0001*

Effect Details

BEPI Binned

Least Squares Means Table

Level	Least Sq Mean	Std Error	Mean
-3 — -2	3.9280906	0.07240431	3.92809
-2 — -1	3.4510517	0.03758581	3.45105
0 — 1	2.4941654	0.07219167	2.49417
1 — 2	2.6263574	0.13040857	2.62636
2 — 3	2.1027741	0.28421888	2.10277

Least Squares Means Plot

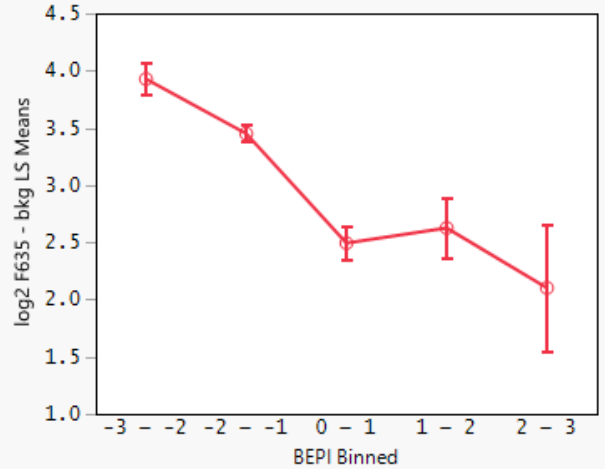


Figure 6: ANOVA by predicted BEPI group

three randomly selected groups. Further the analysis shows that among the BEPI predictions the higher predictions (the -2σ→-3σ subgroup) have a higher fluorescent values. Thus, despite the considerable variation in response among the different subjects this linear model confirms the ability of our neural network predictions in human subjects.

The consistency of the predictions is shown in a more detailed ANOVA (Figure 7). In this more detailed model the individual subjects (Source table) as well as an interaction term was included. In addition, two “region” groups of BEPI peptides were created by combining the -2σ→-3σ plus -1σ→-2σ subgroups (that are shown to be statistically different from the random set in the analysis in Figure 6 above)’. One of these ‘region’ subgroups was a simple singleton peptide ‘target’ and the other ‘region’ was a number of sequential peptides that were all part of a ”target BEPI” subgroup. The “Source Table” variable is the different subjects in the analysis. The results show that in some subjects there is a slight difference between the two region subgroups but for most a singleton peptide and a group of contiguous peptides are no different. Again, as above, despite there being considerable differences among the subjects with regard to their antibody responses, the negative regions are statistically significantly different from the positive regions.

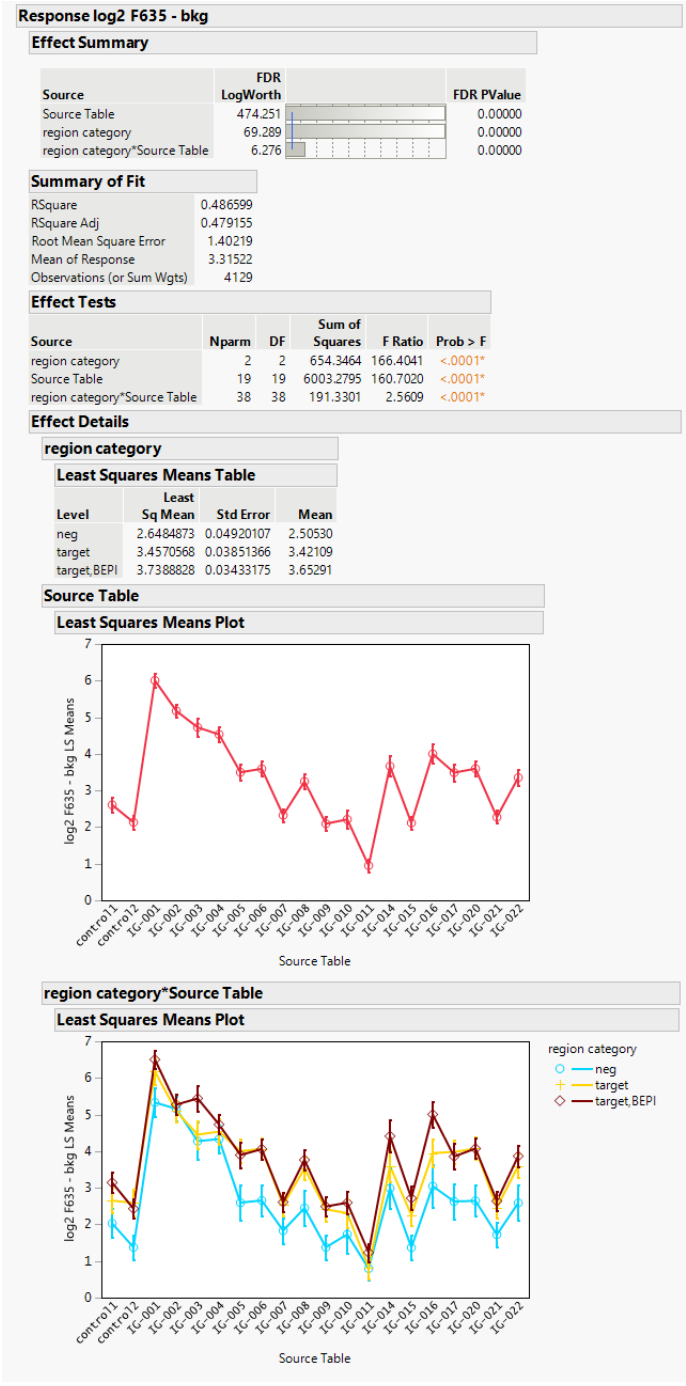


Figure 7: ANOVA individual subjects

In both analyses it is interesting that the controls (commercial heat-treated human sera) have antibodies that recognize epitopes in SARS-CoV-2 proteins. In particular nsp3 responses in one of the controls is in the same range as the subjects that had been infected. This was also observed by the Pepperprint team in a later study and may relate to the high concentration of highly polarized amino acids (aspartic and glutamic acids) in this protein [35]

(see also https://www.pepperprint.com/updates/webinars-and-movies/). A study of SARS-CoV-2 antibody cross reactivity with other human coronaviruses detected the presence of non-neutralizing antibodies to SARS-CoV-2 nucleocapsid in 23% of uninfected individuals [36]; this study did not evaluate antibodies to non-structural proteins.

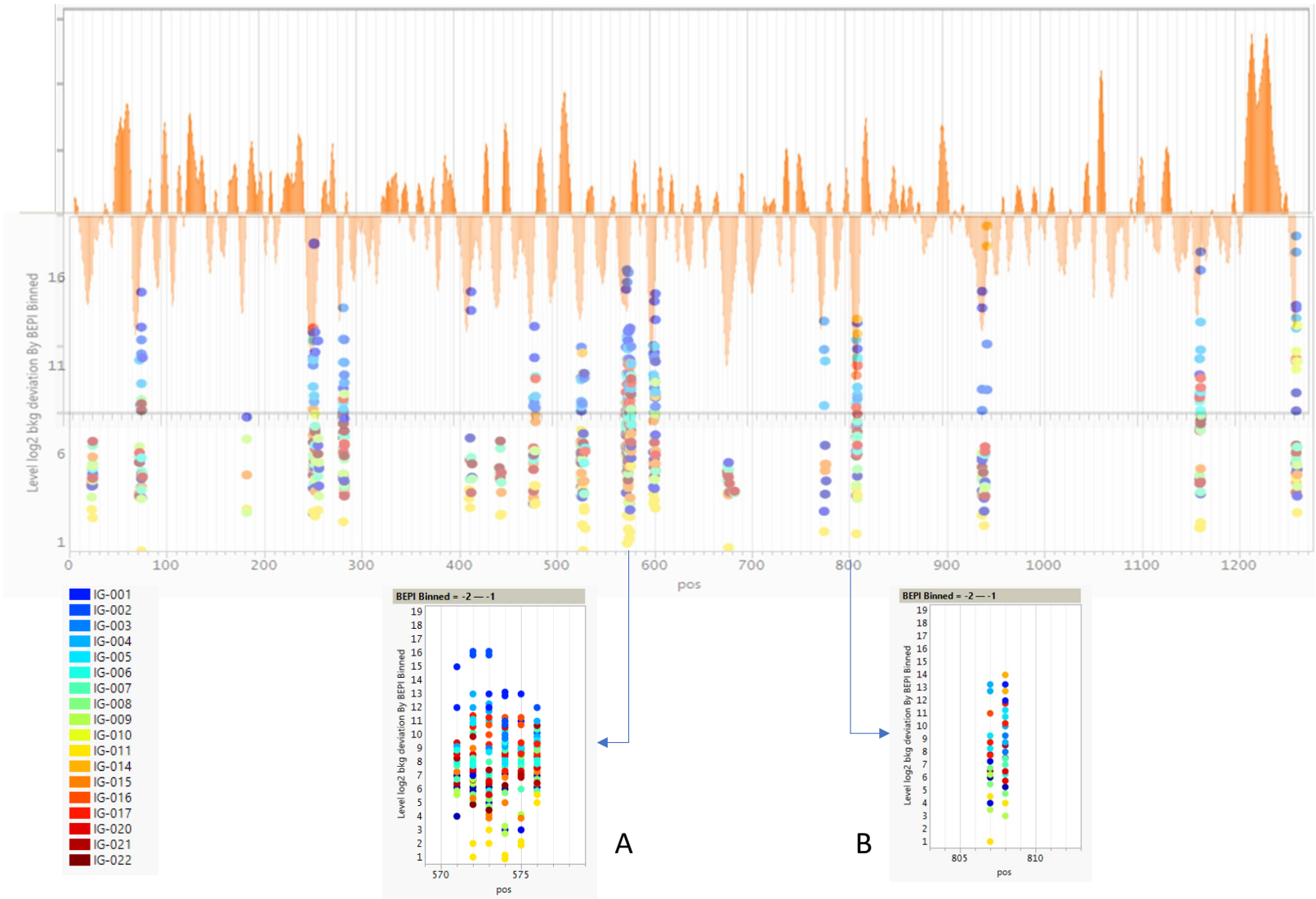


Figure 8. Ranking of antibody binding to Spike glycoprotein linear B cell epitopes in Pepperprint® – overlay of predictions and recorded results.
Top Panel. Predicted BEPI in standardized units displayed in orange for sequential peptides. On this scale -2σ corresponds to approximately the top 2.5% and the peptides exceeding the threshold are the highest probability BEPI candidates.
Lower panel. This shows the different subjects that have circulating antibodies that recognize the peptides. The fluorescence is ranked from low to high Boxes A and B show an expanded view shows that the antibodies are recognizing several sequential 15-mer peptides in these regions. This graphic shows that among the many subjects, even those with lower overall residual antibody responses have antibodies that respond to these peptides.

SARS-CoV-2 T cell epitope responses

PBMC and HLA-Typing:

PBMC were prepared by the Versiti ARL and frozen in aliquots of 10^7 cell/vial. A vial of cells was thawed and the viability of the PBMC was determined on an aliquot using a live cell count and by culture overnight. PBMC that showed normal *in vitro* growth characteristics were used for high-resolution HLA typing. The results of the HLA typing were used for identifying possible epitope peptides to be studied. In most cases, the remaining cells from the first vial were analyzed by flow cytometry to determine the proportion of lymphocytes by forward and side scatter characteristics. The CD3, CD4 and CD8 levels in the lymphocytes was measured using specific monoclonal antibodies (Ab). These data are provided in Table 12.

Testing for responder T cells:

In vitro stimulation of PBMC in peptide recall culture

PBMC numbers were limited, so PBMC were cultured at 2 x 10^6 cells/ml with 15 μM each of a pool of 4 MHC II or 6 MHC I peptide pools for 3 days. Cultures with MHC I peptides included 10 U/ml of recombinant IL-2. On day 3, interferon gamma (IFN-γ) production was detected by enzyme-linked immunosorbent assay (ELISA). The peptides used for each patient were from a list prepared by ioGenetics that predicted both positive and negative responses by the individual patients.

PBMC were then cultured at 2 x 10^6 cells/ml with 15 μM of the individual peptides from the pools that had detectable IFN-γ for 7 days in duplicate wells. Cultures with MHC I peptides included 10 U/ml of recombinant IL-2. At 7 days, the duplicate wells were pooled into 1 well and boosted with another 15 μM peptide. Seven days after the boost, ELISA was used to detect IFN-γ. Individual peptide response cultures were done following the same method with remaining PBMC from the patients who had responded to any peptides. The number of peptides tested per patient was limited by the number of PBMC per patient.

Materials:

Cells and serum:	PBMC and serum were obtained from convalescent SARS-CoV-2 donors under Blood Center Wisconsin, Versiti, IRB-approved protocols. Cells were stored frozen under liquid nitrogen until used. Serum was aliquoted and stored at -20°C. Thawed aliquots were heat inactivated and stored at 4°C until used.
Peptides:	Peptides were synthesized by Mimotopes (Mulgrave, Victoria, Australia). Tables 4-6 shows the panel of peptides used.
Peptide arrays	synthesized by Pepperprint® (Heidelberg, Germany) Appendix 3 shows the panel of peptides used
Monoclonal antibodies:	ELISA: IFN gamma anti-human antibody (ThermoFisher M700A) IFN gamma anti-human antibody biotinylated (ThermoFisher M701B) Streptavidin HRP (BioRad STAR5B IFN gamma human (ThermoFisher RIFNG)
	Pepperprint® Assay: pAb Goat anti-human IgG (H+L) Cy5 (Abcam ab97172)

	Mouse anti-HA Cy5 (Pepperprint® PPC.002.001)
Other reagents	IL2 (PHC0027 ThermoFisher), 10 unit/ml final concentration to all of the MHC I peptide wells on the day the culture is started, and again 2 days after. Add again the day of peptide boost and 2 days after boost.

Table 10: MHC I peptides used for each volunteer. Note alleles that are italicized are those for which we do not have training sets and do not make predictions. Alleles sharing the first two digits are used as a surrogate for such alleles.

Alleles	**IG 1**	**IG 2**	**IG 3**	**IG 4**	**IG 05**	**IG 6**	**IG 7**	**IG 8**	**IG 9**	**IG 10**	**IG 11**	**IG 21**	**IG 22**
MHC IA	1101	0201	3201	0301	0301	0201	0301	0101	2601	0201	0101	0101	0301
MHC IA	2601	2601	3301	3101	3101	0301	3001	0201	3201	0201	0201	0101	6801
MHC IB	2705	1501	*3906*	0702	0702	1501	1302	0702	3901	1501	1501	0801	0702
MHC IB	5701	5101	7801	4402	4101	4402	*2702*	0801	*5001*	3501	5801	1801	1501
9 or 10 mer													
ADQLTPTWR			X	X	X			X		X	X		
ALALLLLDR			X	X	X		X					X	X
APFLYLYAL		X		X	X		X	X	X		X		X
ASVYAWNRKR	X		X	X	X			X			X		
ATKMSECVL	X												
AYVDNSSLT						X	X	X	X	X	X		X
EGFNCYFPL	X	X	X	X	X			X	X		X	X	
ESNKKFLPF	X	X							X			X	
FGAGAALQIP		X	X			X	X	X			X	X	X
FGLVAEWFL	X		X			X			X	X			
FKNIDGYFK	X									X			
FLAYILFTR			X	X	X	X	X						X
FVFKNIDGY	X	X		X	X		X		X	X			X
GTGPEAGLPY	X	X	X			X	X		X	X		X	X
KDLSPRWYF	X	X	X	X	X		X		X	X		X	
RVVVLSFEL		X	X			X		X	X	X	X		
SNNLDSKVG			X	X	X	X	X		X				X
STPGSSRGT	X					X			X	X		X	X
TVYSHLLLVA		X					X	X	X	X	X	X	
VTLADAGFIK	X			X	X	X	X					X	X
WLIINLVQM		X				X		X		X	X	X	
YLGTGPEAG		X				X	X	X			X		
YLQPRTFLLK		X		X	X	X		X		X	X	X	X
YNYLYRLFRK	X		X	X	X			X			X	X	X

Table 11: MHC II peptides used for each volunteer. Note alleles that are italicized are those for which we do not have training sets and do not make predictions. Alleles sharing the first two digits are used as a surrogate for such alleles

Alleles	**IG 1**	**IG 2**	**IG 3**	**IG 4**	**IG 05**	**IG 6**	**IG 7**	**IG 8**	**IG 9**	**IG 10**	**IG 11**	**IG 21**	**IG 22**
MHC II DRB	1_0701	1_0401	1_0404	*1_0103*	1_0801	1_0401	1_0701	1_1501	*1_0806*	1_1104	1_0101	1_0301	1_0101
MHC II DRB	1_0701	1_0701	1_1102	1_0901	1_1501	1_0401	*1_1601*	1_1501	1_1301	1_1301	1_0401	1_0403	1_0401
15-16mer peptides													
AEWFLAYILFTRFFYV	X				X			X	X				
ASYQTQTNSPRRARSV		X			X	X					X		
AWNSNNLDSKVGGNYN	X				X		X		X	X		X	X
DAALALLLLDRLNQLE			X		X			X	X			X	
DNSSLTIKKPNELSRV			X					X		X			X
EDLLFNKVTLADAGFI		X				X	X			X	X	X	
GYQPYRVVVLSFELLH	X	X		X		X	X				X		X
HLLLVAAGLEAPFLYL				X									X
ITSGWTFGAGAALQIP	X			X		X	X				X		X
MYIFFASFYYVWKSYV		X		X				X		X			X
NSTPGSSRGTSPARMA	X	X		X		X		X			X	X	
PFLYLYALVYFLQSIN		X				X			X		X	X	
PRWYFYYLGTGPEAG	X	X	X	X	X		X		X				
SNKKFLPFQQFGRDIA			X		X			X	X	X			
SWLMWLIINLVQMAPI	X		X	X		X	X	X		X	X	X	
WYFYYLGTGPEAGLPY	X		X	X			X			X			X
YLQPRTFLLKYNENGT			X		X				X			X	
YNYLYRLFRKSNLKPF		X	X		X	X	X	X	X	X	X	X	

Table 12: Summary of assay components

SUBJECT ID	Date MNCs Isolated	P = Usable PBMC S = Usable serum	Vials collected	Cells/vial Viability	% CD3 positive of lymphocyte gate	% CD4 positive of CD3+ gate	% CD8 positive of CD3+ gate	MHC I Peptide pool test	MHC II peptide pool test
IG-001	9/22/2020	P and plasma	4	9.8E+06	81	84.3	12.3	no response	no response
IG-002	9/23/2020	PS	6	1.2E+07	80.9	76.5	17.6	positive	positive
IG-003	9/24/2020	PS	14	3.1E+06	60.5	80.1	11.5	positive	no response
IG-004	9/25/2020	PS	5	1.2E+07	71.5	68.5	22.7	positive	positive
IG-005	9/29/2020	PS	10	1.2E+07	83.9	82.8	12.6	positive	no response
IG-006	9/30/2020	PS	10	5.3E+06	78.1	80.7	14	positive	no response
IG-007	10/2/2020	PS	7	8.1E+06	73.9	37.7	16.4	no response	no response
IG-008	10/8/2020	PS	15	8.4E+06	47.2	65.9	11.8	no response	no response
IG-009	10/9/2020	PS	13	9.1E+06	84.7	69.2	22.3	positive	no response
IG-010	10/14/2020	PS	5	1.2E+07	81.8	NA (not analyzed)	15.1	no response	no response
IG-011	10/14/2020	PS	4	1.4E+07	72.7	NA	20.7	no response	no response
IG-012	10/20/2020	not sent to BRI	not sent to BRI						
IG-013	10/21/2020	not sent to BRI	not sent to BRI						
IG-014	10/22/2020	S	3	not viable					
IG-015	10/29/2020	S	4	not viable					
IG-016	11/4/2020	S	5	not viable					
IG-017	11/4/2020	S	7	not viable					
IG-018	11/4/2020	not sent to BRI	not sent to BRI						
IG-019	11/5/2020	not sent to BRI	not sent to BRI						
IG-020	11/11/2020	S	4	not viable					
IG-021	11/20/2020	PS	5	1.4E+07	NA	NA	NA	no response	no response
IG-022	11/24/2020	PS	5	1.1E+07	NA	NA	NA	no response	no response

Assay Implementation

Measuring T cell responses *in vitro* is still as much art as science, in that the culture conditions that mimic the *in-situ* environment of the lymph nodes or of infected tissues are not yet well developed. Our approach was to use the same conditions that we have successfully used for influenza studies. Assay of responses involving CD4+ T cell recognition of class II HLA molecules is performed in the absence of supplements, whereas those involving CD8+ T cells include IL2 supplementation of the culture. The readout was to measure the production of interferon gamma (IFNγ) by ELISA from supernatant samples obtained at day 3 and day 7 of culture.

After our initial analysis we realized that there may be an issue with precursor frequency as the responses were not as robust as those usually obtained from analyses of influenza recall responses. This may not be surprising as influenza is a virus that has been infecting subjects frequently and all adult subjects have undergone multiple exposures to infection or vaccines generating a robust T cell repertoire. Therefore, we implemented a first round analysis to identify robust responders by using one vial of cells to test a pool of half of each of the candidate peptides for each HLA class. If IFNγ production was observed, the peptides were then used individually in cultures to determine which peptide(s) were responsible for the pooled signal.

The original proposal with YFVax took advantage of the availability of the vaccine strain to use as a first week *in vitro* boost to T cells to ensure optimal stimulation of virus specific cells. A live virus infecting PBMC generally provides a very robust response as all the pathogen recognition systems (retinoic acid-inducible gene, toll-like receptor, interferon) are stimulated in the antigen presenting cell (APC). We were unable to do this for SARS-CoV-2 as a safe (BLS2) vaccine strain was not then available. Thus, the directed pooled peptide approach was not an optimal solution to the limitations posed by of precursor T cell frequency, but rather the best that could be implemented under the existing constraints.

A summary of the results of cultures stimulated using pooled MHC I and MHC II peptides are provided as columns in Table 12. PBMCs from most subjects did not give an IFNγ response to the pooled peptides. Of the thirteen sets of PBMCs tested, six gave positive results with MHC I peptides and two with MHC II peptides. While this could possibly reflect the absence of any stimulatory peptides in the pools used, it is more likely due to low numbers of responder cells in the circulation, i.e. precursor frequency below detectable limits.

The low precursor frequency could also be a function of the time samples were collected from the time of end of symptoms, which for the subjects studied here was an average of 5.4 months (StDev 1.9 mo). This is beyond the two months considered optimum for observing circulating T cells after a vaccination. The time post end-of-symptoms was not a variable that we had control over in accessing this group of subjects. It should also be pointed out that the temporal extent of protective immunity in convalescent subjects is still being studied. A recent study indicated that a combination of T cell and antibody immunity could be measured for up to eight months post-exposure [37]. Our measurement of IFNγ production is a more functional assay compared to the

surface marker changes employed in that study. Of the six subjects responding to the peptide pool, the average time post symptoms was 4.5 months and one of these was for a sample collected 7.4 months post end-of-symptoms. This makes it likely that the variation is more a matter of individual variation in the circulating T cell repertoire composition, which is a function of the extent of use of this arm of the immune system.

Analysis of individual peptide responses:

PBMC from subjects that were positive for the pooled peptides were subject to an additional recall analyses with individual peptides. Supernatants from the cultures were removed at day 3 and 7. For some cultures additional peptide was added at day 7 and the culture continued for another week. We completed the analysis on the subjects whose pooled peptide results were positive (Table 12) and the results are summarized in Table 14.

The results for IG-003 are shown in Table 13. Three of the eight peptides resulted in a reproducible IFNγ measurement at day 7. The responses are weak in that it required a peptide boost and seven days of additional culture to observe the IFN production. It is possible that YNYLYRLFRK also generated a response but the IFN levels were very low and could not be measure reproducibly. This interpretation is supported by the day three results in which only FGLVAEWFL showed any response, and this non-reproducibly.

The results for IG-009 are shown in Table 14. Four of the nine peptides examined generated responses. Again, there was one possible but weak response. One observes that there is an overlap of three peptides used in this analysis with those used in IG-003 analysis. There is a shared HLA-A*32 locus and a shared HLA-B*39 locus. Peptide FGAGAALQIP stimulated both cultures. Peptide FLAYILFTR was a strong stimulator of IG-009 but did not stimulate cells from IG-003. Peptide KDLSPRWYF stimulated PBMC from IG-003 but was a weak stimulator of IG-009 cells.

Overall, the signal strength from the IG-009 cultures was stronger than that of IG-003. This data will need to be placed in the context of the remaining cultures.

Table 13: Interferon gamma responses **IG-003**

protein	9-mer	pos	day7.1	day7.2	A_3201	A_3301	B_3901
nucleocapsid	KDLSPRWYF	102	0.92	1.01	-2.18	-0.74	-1.03
nucleocapsid	GTGPEAGLP	114	0.00	0.00	0.10	1.05	0.67
orf1a polyprotein	FGLVAEWFL	2321	0.83	0.80	-1.36	-0.86	-2.17
orf1a polyprotein	FLAYILFTR	2328	0.00	0.00	0.88	-3.35	1.44
spike	SNNLDSKVG	438	0.00	0.00	0.53	1.12	0.47
spike	YNYLYRLFR	449	0.35	0.00	1.38	-3.57	0.94
spike	RVVVLSFEL	509	0.00	0.00	-2.40	-0.04	-2.22
spike	FGAGAALQI	888	0.93	1.00	-1.09	0.34	-1.39

Table 13 shows interferon gamma responses (ng/ml in culture well) to peptides from different viral proteins IG_0003. The ‘pos’ is the amino acid in the protein that corresponds to the N-

terminus of the peptide. Day7.1 sample is the response at 7 days and the day 7.2 is the response at day 14 with a boost at day 7. The cells in the HLA genotype columns represent the standardized ic50 binding activity for the peptide tested.

Table 14: Interferon gamma responses **IG-0009**

protein	9-mer	pos	day7.1	day7.2	A_2601	A_3201	B_3901
nucleocapsid	KDLSPRWYF	102	na	0.65	0.57	-2.18	-1.03
nucleocapsid	GTGPEAGLP	114	0.00	0.00	-2.07	0.10	0.67
nucleocapsid	STPGSSRGT	197	1.49	1.43	0.65	-0.33	0.08
orf1a polyprotein	FLAYILFTR	2328	7.99	6.70	-1.70	0.88	1.44
spike	FVFKNIDGY	192	0.00	0.00	-4.05	-0.73	0.13
spike	SNNLDSKVG	438	0.00	0.00	1.94	0.53	0.47
spike	EGFNCYFPL	484	2.01	1.81	-1.95	-2.14	-1.45
spike	FGAGAALQI	888	2.52	2.96	-0.39	-1.09	-1.39

Table 14 shows interferon gamma responses (ng/ml in culture well) to peptides from different viral proteins IG_0009. The ‘pos’ is the amino acid in the protein that corresponds to the N-terminus of the peptide. Day7.1 sample is the response at 7 days and the day 7.2 is the response at day 14 with a boost at day 7. The cells in the HLA genotype columns represent the standardized ic50 binding activity for the peptide tested.

As can be seen, there is only a partial overlap of the immune dominance patterns of the peptides in the two subjects. The spike protein position 449 (peptide sequence) peptide response in subject IG-003 is interesting in that it is only a low response in one of the two cultures to a peptide with a very high predicted binding affinity. We have seen similar behavior in T cell responses tumor peptides where very high affinity binding seems to be suppress the T cell response rather than cause high levels of stimulation. Our experience is that there is an intermediate affinity that provides high levels of stimulation, but that very high- level binding leading to persistent stimulation of the cells does not necessarily cause high levels of response in the T cells. It is possible that the absence of response to some of the other high affinity peptides occurs for similar reasons.

Results were compiled for the remaining patients collected and are summarized in Table 15. Overall response rate of MHC Class I responses in subjects from whom PBMC isolates were compatible with T cells stimulation assays.

	Nucleocapsid		ORF1a		ORF3a		Spike	
Subject	Peptide	Response	Peptide	Response	Peptide	Response	Peptide	Response
IG002	20	3	10	6	15	0	40	12
IG003	12	6	6	2			30	16
IG004	8	0	4	0	4	1	48	1
IG005	6	0	3	0	3	0	36	5
IG006	16	1	16	0			32	1
IG009	12	0	4	0			16	0
Total	74	10	43	8	22	1	192	35
Response rate		0.135		0.186		0.045		0.182

Table 15. Summary of responses from all patients from experiments carried out as in Table 13 and 14. The ‘Peptide’ column represents the number of peptides with predicted high affinity for one or more alleles in the subject haplotype (equivalent to the red and purple colored cells in Tables 13 and 14. ‘Response’ is the aggregate number of cultures that generated IFNγ responses at 7 days and at 14 days. The 14 day samples were given a peptide-restimulation boost at 7 days.

As shown in Table 15 the overall IFNγ response rates to peptides was lower than anticipated. However, as discussed in the section below, these results are consistent with the published results.

Comparative analysis of predictions to epitope specific immune responses to Sars-CoV-2 documented in literature reports.

Since January 2020 over 146,000 peer reviewed papers on Sars-CoV-2 have been published as of this writing, in addition to a large number of non-reviewed papers in BioRx and medRxiv. A number of these address epitope-specific responses.

B cell responses

One paper published relatively early (April 2020) by researchers at the Max Planck Institute [28] utilized the Pepperprint® arrays. This study evaluated IgG, IgM, and IgA responses at various time points up to 22 days post infection in a clinically serious case of COVID and two mild cases plus controls. We compared raw data from that study to our predictions and this is provided in Appendix 6. Figure 9 shows the concordance of our predictions in the spike proteins across the clinical cases. The HLA of these patients, which may affect T helper responses is unknown.

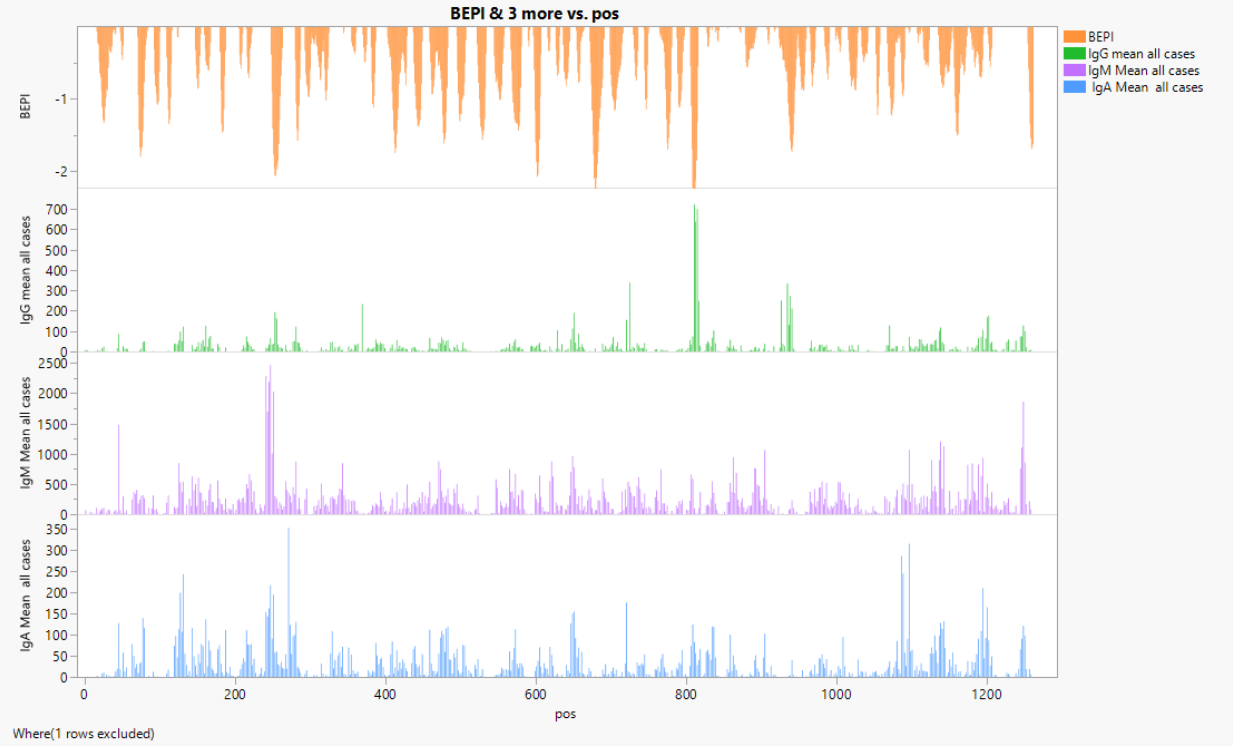


Figure 9: Comparison of B cell epitope predictions to observations of Dahlke *et al* [22]
Pepperprint array results mean of 3 clinical patients. Orange: B cell linear epitope probability (inverted); values below -1 considered positive. Green:-IgG, Purple: IgM; Blue: IgA.

Indications are that most neutralizing antibodies mapped to the receptor binding domain [38], however strong antibody responses are also reported to other regions of the spike protein [39, 40], other structural proteins [41] and to nonstructural proteins [35].

T cell responses

The La Jolla Institute are among those who have conducted extensive studies of T cell responses to SARS CoV-2. In addition, compiled and curated data on responses in the IEDB database managed by the Institute. Therefore, COVID cellular immune responses curated at IEDB.org were downloaded and analyzed. As of 7 May, 2021 IEDB catalogued a total of 21 publications with recorded responses by 12 different techniques amounting to over 1181 different response assays to peptides from the reference SARS-CoV-2virus [42-69]. A total of 581 different peptides were tested for MHC I stimulation of T cells. The number of times individual peptides were tested ranged from singletons (1 test in 1 subject) to 44 for a peptide with high affinity for the common A*02:01 peptide tested in multiple labs. As we have shown above the responses documented in the published data is highly variable. The median overall frequency of response in individuals that had recovered from COVID (as curated by IEDB.org) was approximately 25% (Figure 10 and panels A, B, C Figure 12). This figure is slightly higher than the median 15% for 'healthy' individuals who had not been infected by the virus and lower than 55% acutely ill subjects. In all cases researchers were testing peptides that were expected to have given T cell responses based on the subject's HLA haplotypes, but nevertheless the results ranged from zero to 100%. The results we obtained with our recovered individuals is thus comparable and equally as variable as the published data.

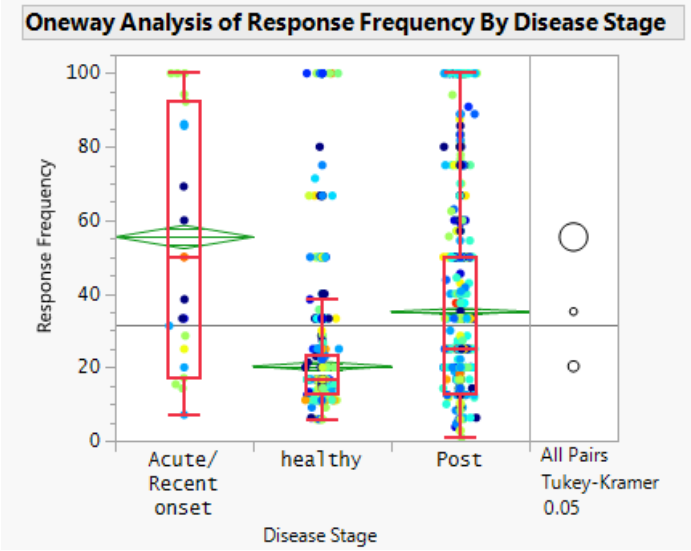


Figure 10: Response rates in all MHC Class I responses curated by IEDB.org in 21 publications (includes bioRx and medRxiv). The boxes show the median, the 25 & 75 percentile and the extended whiskers the 1.5 times the interquartile range. The diamonds show the mean and standard error. The comparison on the right shows that the differences between groups. Acute/Recent onset are subjects with active disease, "healthy" are non- infected control subjects and "Post" are subjects who had been infected with the virus at varying times prior to test.

Appendix 8 shows the positive results of HLA specific peptide responses documented by IEDB compared to the EigenBio predictions for the same peptide allele combinations.

Role of TCEM frequency

An aspect of T cell responses to which we have given considerable attention is the limited possible repertoire of T cell exposed motifs. This is a feature of T cell assays that is not widely appreciated. Being able to assess a T cell response requires that a responding T cell is present in the collected PBMCs. A peptide bound in an MHC groove comprises anchor or pocket position amino acids, which are not exposed to the T cell receptor (TCR) and pentameric motifs that in conjunction with the MHC histotope engage the TCR. In the case of MHC I and CD8+ engagement the pentamers are contiguous amino acids (P4,5,6,7,8 of a 9-mer) whereas in MHC II and CD4+ they are discontinuous pentamers (2,3,5,7,8 and -1,3,5,7,8 with respect to the central 9mer core of a 15mer or longer peptide) [15]

That a pentamer made up from 20 natural amino acids can only be arrayed in 20^5 possible configurations limits the repertoire of potential cognate T cells and necessarily results in polyspecificity - the situation where a particular T cell receptor has the capability of responding to multiple different amino acid combinations in a TCEM. When compared to the frequency of occurrence in various large databases distinct frequency patterns are noted; some TCEM are common and shared across many proteins; others are rare. TCEM frequency can be an indicator of precursor frequency, in essence indicating the likelihood that a cognate T cell clone may exist in a patient's T cell repertoire.

We compared the TCEM frequencies in all the SARS COV2 responding peptides compiled at IEDB with the TCEM frequencies in 3 large reference databases: the human proteome, the gastrointestinal microbiome and the human immunoglobulinome (all human IgV regions).

To make this comparison we used the central core amino acids at positions p4,p5,p6,p7,p8 that protrude from the 9-mer peptide bound in the MHC I groove that we call TCEM I. Atomic interactions of the buried amino acid side chains of p1,p2,p3, and p9 with the MHC binding groove and drive the binding affinity of the peptide for the MHC with p2 and p9 often referred to as "anchor positions" although analysis shows that the binding is the composite result of the 4 amino acids.

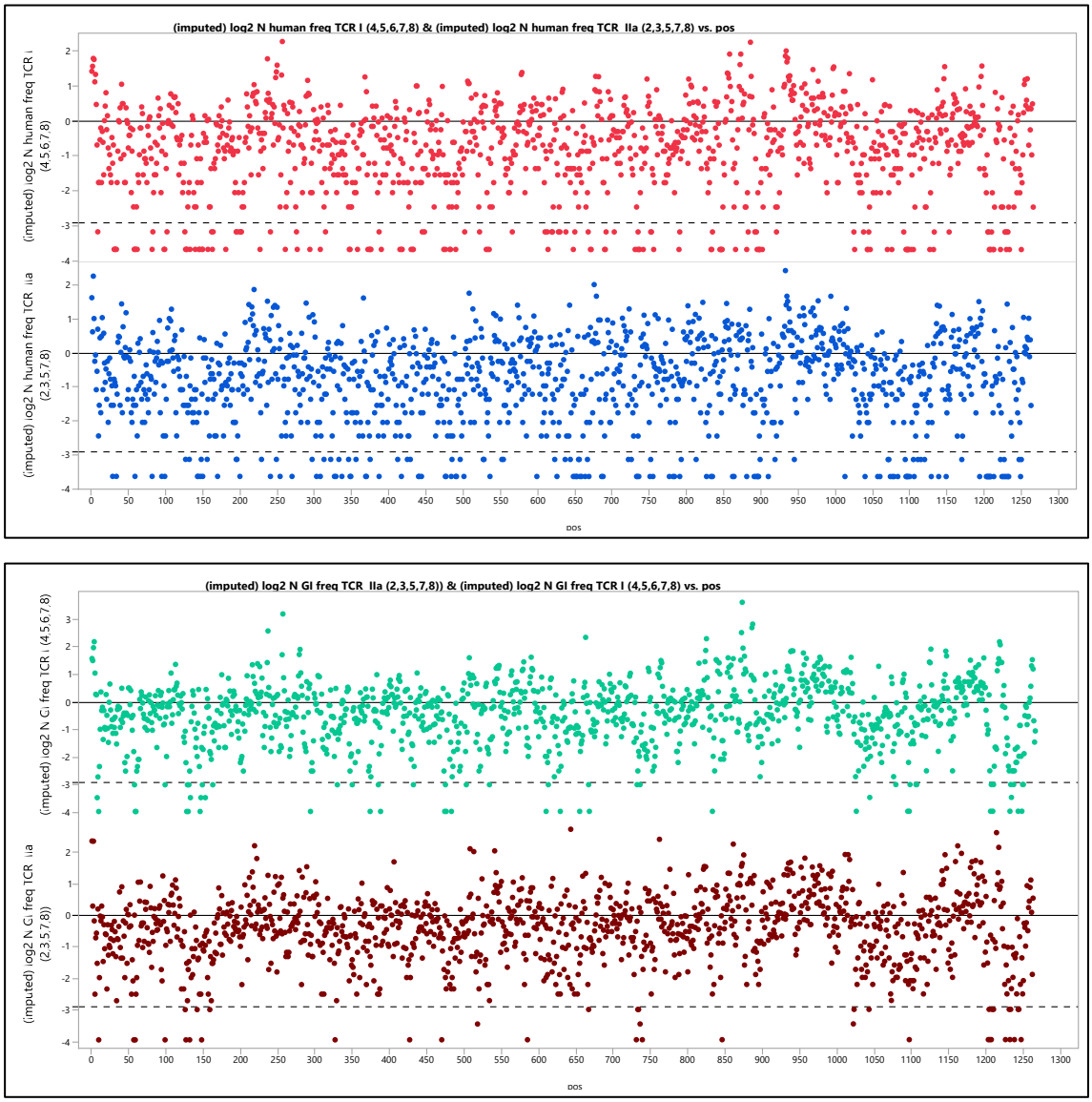


Figure 11: TCEM motifs in the Spike protein of the reference SARS-CoV-2 protein. Upper panels show the TCEM I (CD8+ engaging motifs) and lower panels show TCEM IIA (CD4+ engaging motifs) in the spike protein of SARS-CoV-2 protein compared to the frequency in the human proteome (top) and gastrointestinal microbiome represented by 130 bacterial species (bottom). Each pixel indicates the frequency of a sequential T cell exposed motif across the spike glycoprotein of SARS-CoV-2. The Y axis indicates the frequency relative to that in the human proteome, distributed about a mean of zero. Any TCEM shown below -3 is completely absent from the human proteome or the gastrointestinal microbiome database respectively.

Our work has also shown [15, 16] that the protruding TCEM I amino acid pentamers have characteristic frequencies of occurrence in the human proteome, the GI microbiome and the stochastically generated immunoglobulin variable regions. In order to make the comparisons statistically valid we have appropriately scaled the data. We have used a SHASH (Sinh-Arcsinh) distribution function to normalize the TCEM I frequencies of the human proteome (approx. 25 million peptides) and the GI microbiome (approx. 110 million peptides, from 132 species of organisms). This converts the distributions to a zero mean unit variance distribution commonly used in statistical analysis. The frequency distribution of the motifs in the immunoglobulins is more complex and is based on a 1/log2 frequency and is based on the observed frequency in approximately 40 million immunoglobulin clonotypes.

In total, the presentation of TCEM I throughout the body comprises the antigen presentation profile (APP). The APP is the composite of all proteins being differentially expressed by cells throughout the body and continuously surveilled by the stochastically generated T cell repertoire. The SARS CoV-2 APP represents all peptides of the virus as they are presented on the host MHC molecules. Viral peptides with high affinities will displace a fraction of the normally presented cellular TCEM and the composite change in pattern will be detected as non-self by the immune surveillance system.

Figure 12 compares the APP of SARS CoV-2 to that of the human proteome. The upper row in the figure (Panels A, B, and C) are the frequency of occurrence of all TCEM tested in the publications curated by IEDB, without regard for the HLA binding affinity in the context in the proteome. The colored contours represent a number of modal clusters of TCEM frequencies seen in the datasets. Each of the colored contours are graded and each contains 5 percentile of the total values. In this visualization the intensely red colored areas contain a large number of peptides within the contour. The clusters in panels A and C are distributed about the mean of zero in the distribution of the reference distribution. The distribution of the proteomic data is skewed toward more rare motifs that are expected to be found in viral protein sequences. The distribution of the GI microbiome is normally distributed about the mean.

The lower panels (D,E,F) shows the same data as (A,B,C) but weighted by the number of times each TCEM I is found in the human proteome. This gives rise to the scattered modal clusters with multiple contour lines where individual or groups of TCEM have a large presence in the proteome. Thus, the regions with multiple red-colored contours are regions that flag potential HLA specific epitope mimics in the human proteome – viral TCEM that are also common in the human proteome.

Epitope mimics have often been implicated in adverse physiologic responses to viral infections. The unique circumstances surrounding the COVID pandemic, with large numbers of contemporaneous subjects some of whom have long term effects of viral infection, offers a unique perspective to examine the potential for epitope mimic cross reactivity. We therefore combined the IEDB dataset with our dataset of T cell exposed motifs for the entire human proteome. All 9-mers in the human proteome are included in this comparison to the virus.

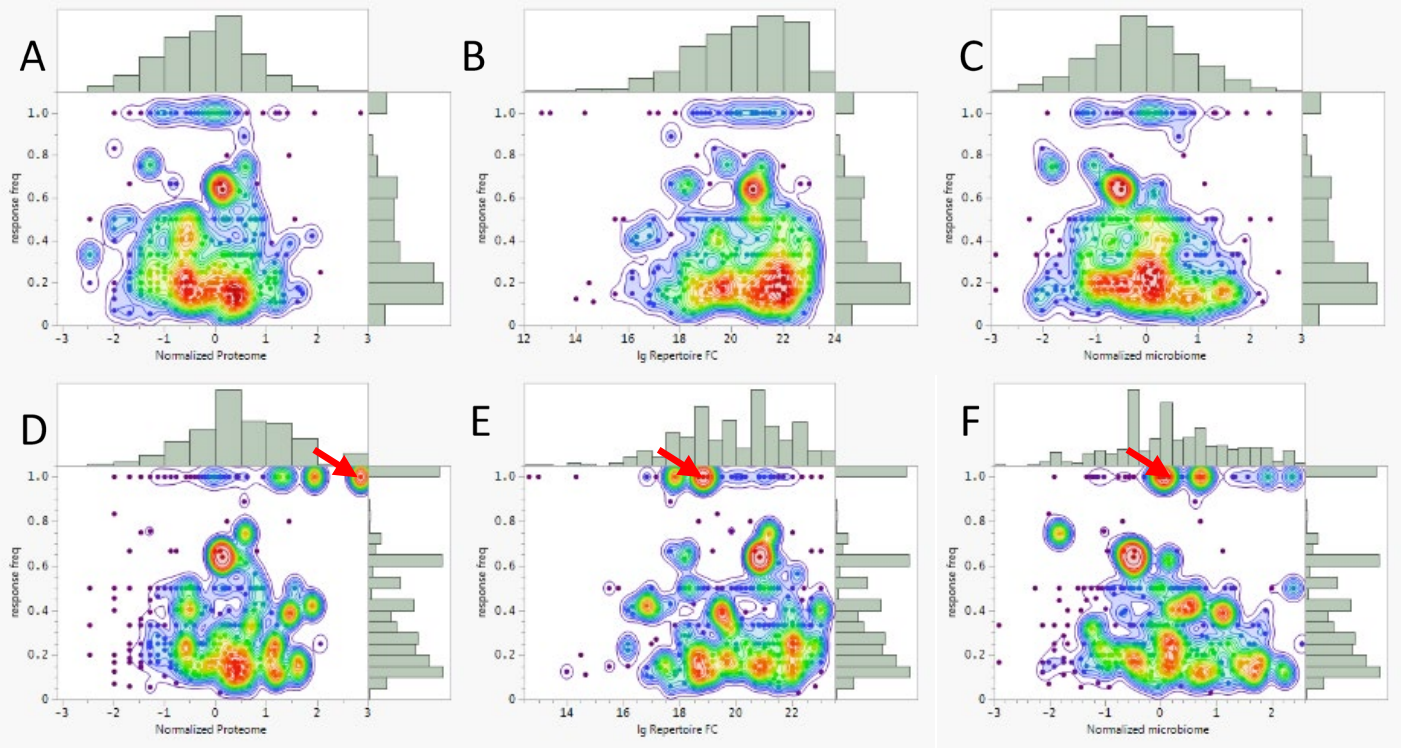


Figure 12: Antigen presentation profiles of TCEM I from SARS-CoV-2 virus. Panels A,B,C represent the overall frequency of occurrence in the human proteome (A) the immunoglobulin (B cell receptor) repertoire (B) and 132 species resident in the GI microbiome (C). The X-axis of panels A and C are in standard deviation units for a normal distribution with zero mean and unit variance produced using a SHASH distribution transformation of the logarithmic frequency data. The frequency distribution of the immunoglobulin repertoire is in reciprocal log2 units (i.e. 22 = $1/2^{22}$ = 1 per 4.2 million clonotypes). The Y axis it the response frequency of the IEDB-curated test results for the MHC I peptides. The lower panels D, E, F are a high affinity binding subset in the proteome context of the upper panel weighted by the number of matches to a particular TCEM I in the proteome. The red arrows indicate a TCEM I ~~~PGLPG~ having a total of 216 copies in the proteome in peptides with high affinity binding context of multiple HLA alleles. For specific patients with a particular HLA haplotype this binding may be particularly high or stimulatory.

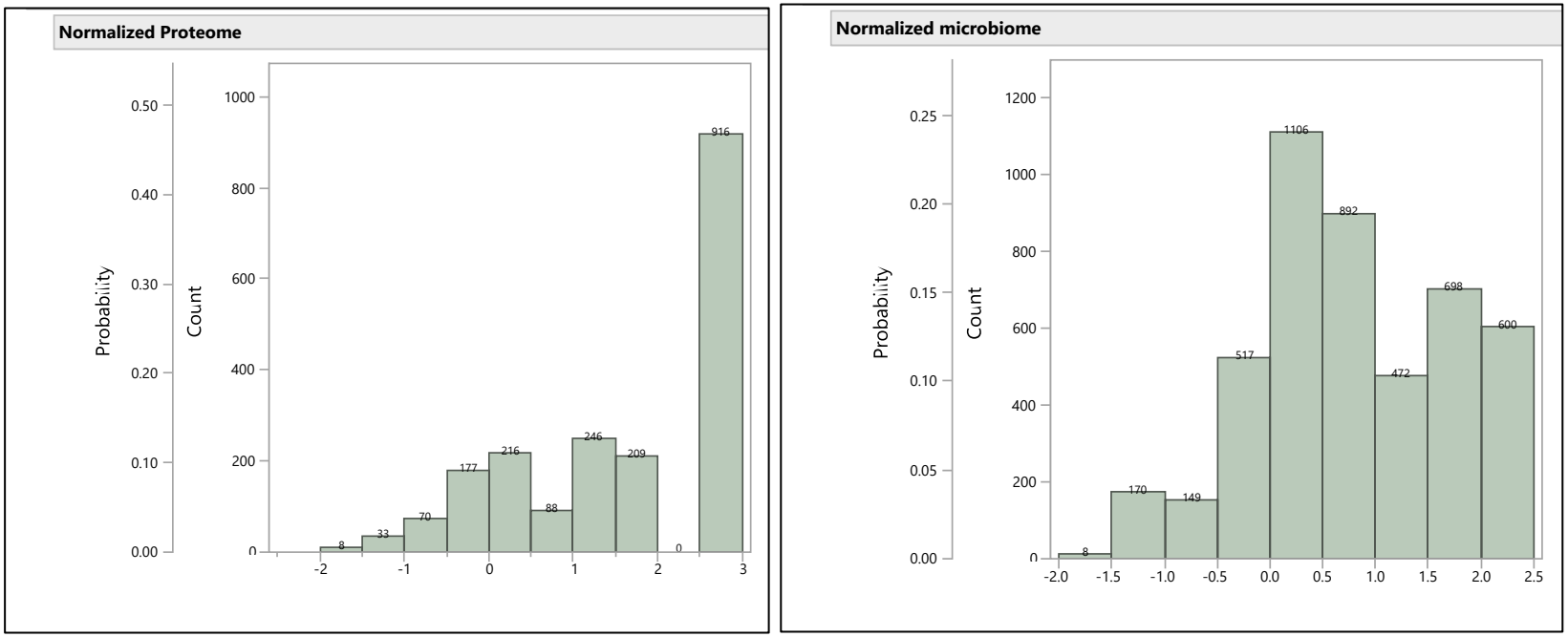


Figure 13: Distribution of TCEM in SARS-CoV-2 that elicit a T cell response documented by IEDB relative to their frequency in the human proteome (left) or gastrointestinal microbiome (right). In both cases the frequency in the database is distributed about a mean of zero. It is notable that more responders are to the more common TCEM motifs, which might be expected to have the larger cognate T cell populations at the time of infection. The counts indicated on each bar show the number of occurrences in the human proteome and representative GI microbiome respectively.

Several features are noteworthy in this visualization of the data in Figures 11 and 12.

1) As noted above in Figure 11 the T cell responses are low in the IEDB dataset and the responding TCEM are found at many different frequencies albeit with somewhat skewed distributions as is shown by the histograms on the borders. There is one cluster group, forming the horizontal band at the top of the figure which gave positive responses in every test (some by a number different research groups). But, overall the contours forming the major modal clusters of T cell responses are in the response frequency range of0.2 - 0.3, meaning that T cell responses were measured in only 20% - 30% of the tested subjects. The publications also note wide variation in the magnitude of responses in different subjects. The results in our own tested convalescent subjects are not inconsistent with these response frequencies and the differential responses among subjects.
2) The red arrows point to a high prominence TCEM I in a peptide for which multiple positive responses were recorded by multiple research groups (FVVPGLPGT, found at position 2864 of replicase polyprotein 1ab). The highlighted TCEM PGLPG is highly represented in the proteome, found 216 times in a collagen-alpha family of proteins that are major components of basement membranes. The motif is also in the microbiome and immunoglobulins (where its frequency is near the median).
3) Responding TCEM shown in panels D,E,F are skewed above the mean for both the proteome and the microbiome. There is a singleton outlier at -2 σ in the in the microbome with a 0.8 response frequency. That motif is the singleton at zero σ in the human proteome. But, overall only two modal clusters are below the mean at about -0.5 σ all others are above the mean

What can be deduced from these patterns?

1) The simplest explanation for the overall response pattern is that T cells responding to SARS CoV-2 viral motifs are predominantly those that are relatively common in the APP of both the human body as a whole and the GI microbiome. These are motifs for which we would expect there to be higher numbers of preexisting clones of T cells (high precursor frequency) in the immune surveillance system. This is a type of response that would be expected to occur from any type of vaccination or viral infection. It is the fundamental immune process needed to provide future protection.
2) It also follows from 1) that viral activation of these relatively common T cells provides the potential for autoimmune responses.
3) As an example, focusing just on the three clusters on the top right of panel D which in all tests were positive to peptides with these TCEM motifs. If we track the motifs to their counterparts in the human proteome we find that T cells responsive to these motifs are found in various subtypes of collagen IV, V, and XI, in an enzyme curated to be involved in cognitive function, and a major mitochondrial protein. Those with 10 or more family members in the proteome are UNIPROT_ID CO4A5_HUMAN, E7EUG2_HUMAN, NPB11_HUMAN, NPB13_HUMAN, NPIB3_HUMAN, NPIB4_HUMAN, NPIB5_HUMAN, Q49AM6_HUMAN). However, allele specific T cell responses to any on the list of proteins which comprise these TCEM may have an impact on their corresponding physiologic processes through autoimmunity.

Task IV: Evaluate responses of mice to Yellow fever vaccine virus.

Versiti Blood Research Institute

The purpose of this Task was to evaluate immunoinformatic predictions of immune responses in a mouse model. The original intent had been to compare with the human response, and hence the selection of YF vaccine which we intended to evaluate in both and determine if the computational modelling in mice was predictive of responses in humans. Based on a literature review of the responses to mice to YF [70] C57BL/6 mice were selected for vaccination.

To assess the magnitude of CD4 T Helper 1 (Th1) and CD8 effector response to MHC I and MHC II yellow fever virus (YFV) synthesized peptides, we vaccinated 12 C57BL/6 mice with YF17D vaccine strain ($1x10^5$ pfu/mouse by intraperitoneal injection). To study the primary response 7 days after vaccination, 3 wild type control mice and 3 vaccinated mice were sacrificed. At this time point, serum and spleen were collected for *in vitro* peptide stimulation. Serum was collected on day 27 before the recall challenge.

To assess the recall response, memory mice groups were administered with 1 $x10^5$ pfu YF17D vaccine. Seven days after the rechallenge, mice were sacrificed, serum and spleen were collected for *in vitro* peptide stimulation. Peptides were selected as described in Task I. Splenocytes were diluted to $1x10^7$/mL and 100 ul of cells were plated in flat-bottom 96 well plate leaving maximum room between no stim and peptide stimulated cells. Peptides were diluted to 25 uM in 250 ul in complete RPMI. 50 ul peptide to each well for a final concentration of 5 uM. Brefeldin A was diluted 1:250 and added 50 ul to each well. Cells were incubated at 37℃ for 6 hours. For cell surface staining, 100 ul cells were transferred into a U-bottom plate at a concentration of $1x10^7$/ml. Cells were stained in 50 ul of antibody mix (CD8 - FITC, CD4 – Percp (1:200) CD44 – APC Cy7 (1:200)) in flow cytometry staining buffer (FACS) (0.5-1% FBS in PBS + 0.1% NaN3) to each well. Cells were incubated at 4℃ for 30 minutes then washed in FACS buffer. For intracellular cytokine staining, 50 ul of Cytofix/Cytoperm buffer was added to each well for 15 minutes incubation in the dark. Cells were then washed with 50 ul perm wash. Intracellular antibodies (IFN-Gamma – PE (1:300)) were diluted in 50 ul perm wash and added to each well for 30 minutes incubation. Cells were then washed with perm wash, resuspended in FACs buffer, and analyzed on flow cytometer LSR-II Green. Consistent with previous publication [71], MHC I peptides 3NS267 and 3NS261 stimulated the highest frequency of $IFNg^+CD8^+$ cells, whereas POL308, POL712, POL856, and E1 moderately stimulated IFN production in $CD8^+$ cells. All MHC II peptides stimulated IFNg production in $CD4^+$ cells.

However, despite early results suggesting strong T cell responses to peptides predicted to bind and stimulate a both CD8 and CD4 T cell responses (Figures 13 and 14) we were unable to obtain consistent infections of the animals with identical amounts of virus and route of administration described in the paper. Moreover, further studies suggested that secondary immunizations appeared to affect YFV-specific CD8 and CD4 T cell responses. Splenocytes were harvested from B6 mice that were infected with YF17D 5 days after the second immunization. Cells were stimulated with MHC-I- or MHC-II-specific peptides such as H-2Db-restricted epitope from the E protein and H-2Kb restricted epitope (NS3). However, differences between animals were so great

that there was really no consistent response detected. The tests for antibody responses in the mouse were likewise not successful. In our efforts we largely confirmed the quotation from a recent paper: "*In this study, we used YFV-17D infection of mice to examine factors that play a role in viral pathogenesis. We found that identically infected mice developed … distinct responses … and we determined that the route of infection and host immune responses influenced YFV-17D pathogenesis more than the amount of virus* [72]. In light of this it was not surprising that evaluation of antibody responses using the Pepperprint® arrays likewise failed to differentiate test from background control sera.

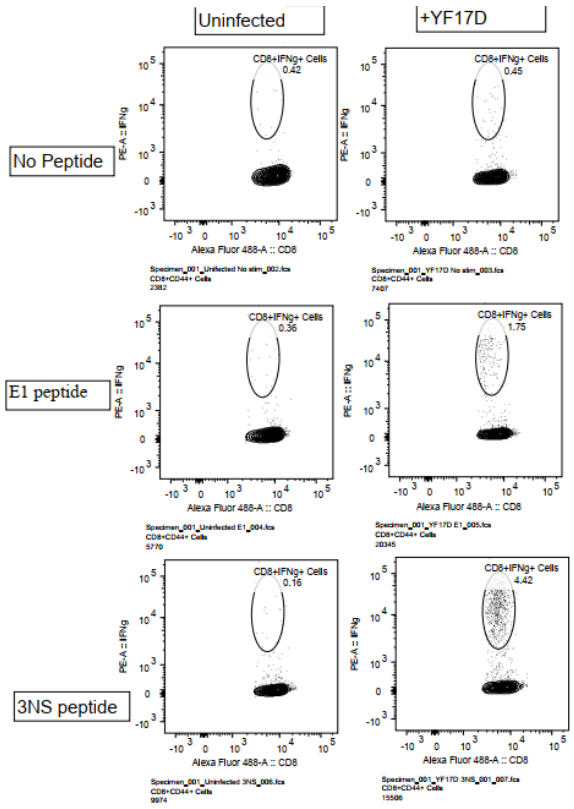


Figure 14: Responses of $CD8^{+}CD44^{+}$ cells to interferon gamma stimulation by YF peptides predicted to produce a strong T cell stimulation

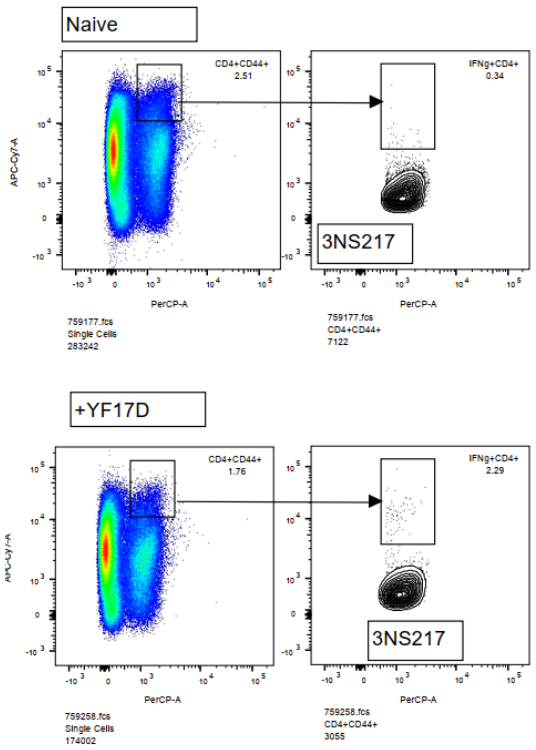


Figure 15: Responses of $CD4^{+}CD44^{+}$ cells to interferon gamma stimulation by YF peptides predicted to produce a strong T cell stimulation.

Analysis of Yellow fever epitope predictions relative to literature reports

In the absence of access to YFVax to conduct *in vivo* human studies, we conducted a review of data in the literature and IEDB and correlated this with our predictions.

A large number of B cell and T cell epitopes in yellow fever (YF) virus have been cataloged in the Immune Epitope Data Base (IEDB www.iedb.org). Epitopes submitted to IEDB are generally identified experimentally by systematically testing the peptides comprising the proteins in a sliding window with partial overlap between different peptides. This process is an empirical compromise due to the cost of peptide synthesis; so it inevitably misses peptides that might be good prospects. In contrast ioGenetics predicts for every sequential peptide and in a contextual way where competition between peptides is considered. The curated YF epitopes cataloged in IEDB are summarized in Table 16. The IEDB curation process assigns one of three to five different categorical values to the immune response (negative, positive, strong positive, weak positive). These categories are not consistently applied, however, most likely due to the curation process from the published scientific publications which typically use different methods of analysis. Some of the data is also indicated to be by "direct submission" indicating that a laboratory has run the analysis and submitted it to IEDB for curation, precluding a review of methods.

Table 16. T cell epitopes of Yellow Fever Virus cataloged in IEDB (Accession JN811140).

Protein	*Length of Protein*	*Curated Epitopes in IEDB*	*Peptides not tested at 1 amino acid offset*
Core protein C ± Anchor	222	25	197
Envelope protein	493	98	395
Matrix protein M	75	9	66
M precursor	164	22	142
Non-structural protein NS1	352	76	276
Non-structural protein NS2a	224	33	191
Non-structural protein NS2b	130	21	109
Non-structural protein NS3	623	152	471
Non-structural protein NS4a	126	19	107
Non-structural protein NS4b	250	41	209
RNA dependent RNA polymerase	905	209	696

Comparisons were made between the ioGenetics binding predictions and the qualitative T cell predictions curated on IEDB. Binding affinity predictions were made using neural network ensembles for different MHC I and MHC II alleles that were trained on 9-mer and 15-mer peptides, respectively. Training peptides with their associated affinities were obtained from the repository maintained at IEDB. The binding affinities were standardized to zero mean and unit variance scales within the YFVAX protein using a Johnson Sb distribution function in JMP. IEDB curated results only reports results for MHC I A and B alleles and MHC II DR alleles.

ioGenetics selected a value of -1 standard deviation below the mean affinity within protein to be the positive binding-activity threshold, based on prior testing with a large influenza data set [13]. The grey shaded area represents the peptides that will be deemed positive by this criterion. As is shown in these graphics this selection criterion show quite good discrimination with these data sets for MHC I. The concordance is best for the positive high category for MHC class II. Detailed inspection suggests that this is likely (at least in part) due to the optimum peptide

binding register not being coincident with the peptides tested by researchers submitting to IEDB. The actual peptide binding affinities in the “positive low” category are very low and overlap considerably with the negative peptide. We note the concordance between ioGenetics predictions and IEDB but given the limitations in data submitted to IEDB noted above there is no basis for evaluating discrepancies as false positive/negative (no gold standard).

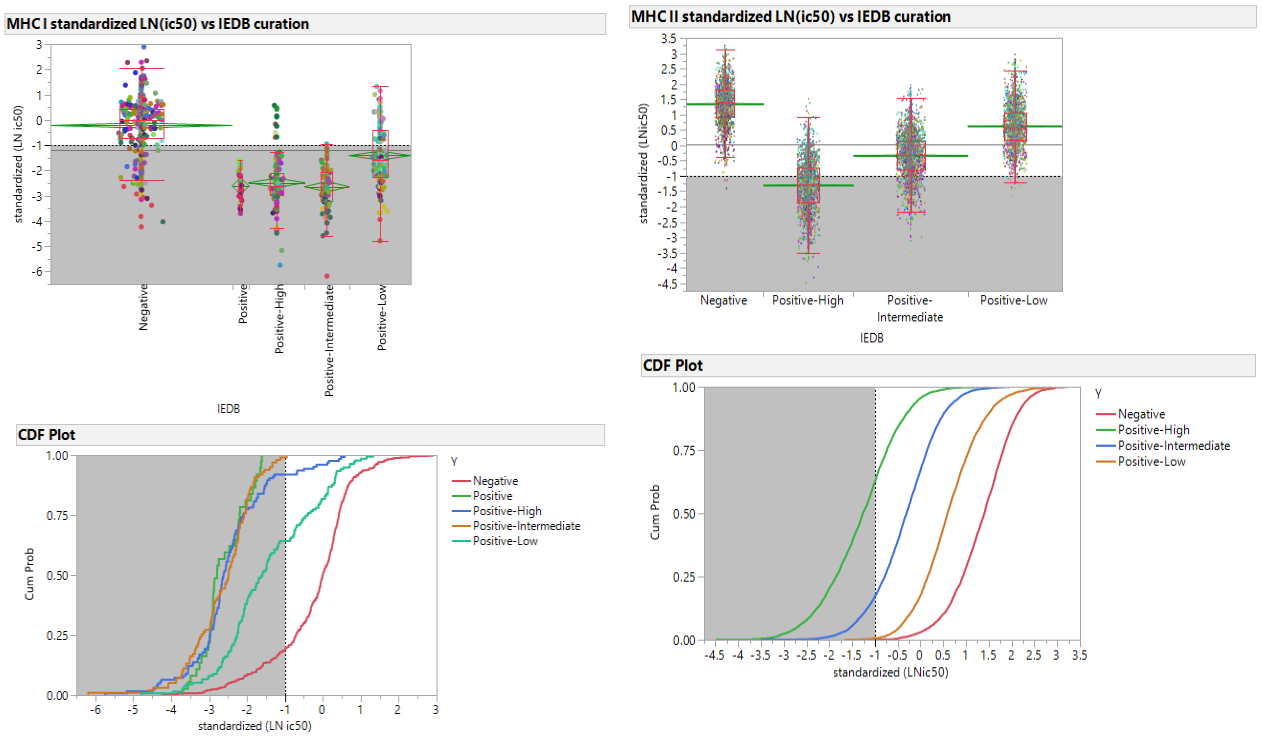


Figure 16: Statistical comparison of standardized binding predictions to IEDB curated in vitro T cell responses to MHC I and MHC II bound YF peptides. YFVAX® peptide binding affinities predicted and standardized as described in Table 1. IEDB curated responses are compiled from a number of studies using different methodologies. The shaded area encompasses NN-predicted affinities that are more than 1 standard deviation below the mean for the particular peptide:allele combination being compared. The width of the box plots are proportional to the number of data points in the data subset the whiskers indicate the outliers based on the inter-quartile range. The lower panels are the cumulative distribution plots of the data in the upper panels and show the overall concordance between the NN predictions and the qualitative curation.

Discussion and Conclusions

The EigenBio immunoinformatics platform has proven itself to be a valuable aid to understanding of overarching patterns in immunological recognition which enable hypothesis development and testing. It is also emerging as a valuable aid to vaccine design for both infectious diseases (including viruses, bacteria, fungi and parasites) and cancer, as demonstrated by on-going collaborative research. Unlike other immunoanalytics, EigenBio combines multiple analyses in one comprehensive platform, built from first principles.

The various logistical obstacles in vaccine access and the COVID pandemic have made achieving all the specific goals of this project challenging, however it has also underscored a few key points:

- The immunoinformatics platform can be implemented very rapidly to generate both high level overviews and very granular data describing the immunomic features of proteins and proteomes (whole organisms). This can be used to select peptides (or longer sequences) as vaccine antigen targets. This can also enable development of, and be supported by assays, such as the Pepperprint® microarray, for subsequent monitoring of outcome.
- Mapping of proteins to determine epitopes by computational means can be achieved based on only sequence information. No growth of organisms is needed, which would require biocontainment, development of optimized culture systems and considerable time.
- While *in vivo* trials and assays are constrained by enrollment of a limited number of human volunteers reflecting a limited array of immunogenetics, the immunoinformatics platform can look at a “whole population *in silico*”. No IRB approval is required to run a computational analysis of a microbial sequence.
- Monte Carlo simulations can rapidly assess the impact of mutations or different designs.
- Current lab assays require highly skilled personnel to obtain consistent results. Details of reagents, instrumentation, implementation etc. can all affect outcome. There remain questions about whether responses (e.g. cytokine profiles) measured in peripheral cells or following *in vitro* culture truly reflect the activity at the site of immune function. Obviously, samples from different individuals, even if of the same HLA, are affected by prior immunological exposure and hence repertoire frequency. Thus, there are major questions as to what really is the “gold standard” for assessing an immune response. The immunoinformatics approach is not affected by a change in subject, reagents or technician.
- Vaccine research has depended heavily on animal models. However, many, indeed most, animal models are outbred and lack training sets to model their HLA based responses. Yet it is clear that this is a critical factor in the variability of immune response and the ability to relate animal models to particular human outcomes.
- The analyses presented here (as illustrated in the Addenda which characterize several current vaccines) indicate that it is time to move beyond characterization of individual epitopes, which has been the traditional practice spearheaded by IEDB, and recognize

that understanding immune responses requires systems analysis. As we show in the discussion of COVID responses analyzed by ourselves and others, responses are multivariate. At a minimum they are conditioned by their relative binding affinity within a protein, by the host MHC alleles, and by the precursor frequency of T cells in the host's repertoire, which in turn is a function of prior immune exposure of that individual and the loss of repertoire due to immunosenescence. These are variables which cannot be replicated in animal models.

The SARS CoV 2 pandemic has provided a unique opportunity to look at the immune response to a pathogen for which the population is naïve. Clearly emerging for our work in this study is the multifaceted nature of the immune response. We had experienced low and variable T cell responses in our own patient studies, but it emerges more clearly from the composite of the analyses by many different groups that has been curated by IEDB. The predicted (and in some cases measured) affinity of peptide binding to the MHC does not in itself predict immunodominance or responses. Peptides with a high affinity for a subject's HLA haplotype predicted by our tools (and several websites referenced in the publications from which IEDB accumulated data) produced a wide range of responses among individuals. Overall, 20-30% of haplotype-matched individuals respond. These low T cell response frequencies clearly illuminate the critical role that the prior T cell repertoires play in an individual's response. A T cell repertoire is likely to contain a number of different TCRs that can recognize a peptide with varying affinities [73], but there is a significant repertoire sampling problem involved in collection of a relatively small volume of blood to carry out the assay. Ultimately the responding T cell population needs to be sufficiently large for a measurable response to be generated by the approx. 10^5 T cells used in an assay. The peptide boosting process we used, and also used in some of the published work, helps but cannot completely overcome this sampling issue.

T cell repertoires are stochastically generated during youth but as the thymus involutes, T cell responses are driven by the naïve and memory repertoire availability of responding TCRs [74, 75]. Overall, T cell repertoire diversity diminishes with age. Logically, this may explain the age-dependence of serious COVID illness.

These underlying concepts are underscored by the TCEM frequencies of the SARS-CoV-2 responding peptides; these are significantly skewed towards T cells for which there is repertoire 'familiarity'. T cell responses are more likely to be generated by TCEM that are common in the human proteome, the GI microbiome, and the immunoglobulinome (the overall antigen presentation profile or APP) (Figures 11 and 12) Given the diversity of individual T cell repertoires, based on early repertoire establishment and subsequent epitope exposure experience, both the CD8+ cytotoxic T cell and CD4+ T-helper responses are highly individualized. The skewing of the TCEM responses is also intertwined with the sampling issues described above. This all underscores the importance of maintaining a diverse T cell repertoire to confront an unknown pathogen.

Appendices

1	FDA approved vaccines
2	Epitope mapping Seasonal influenza vaccine H3N2
3	Peptides included in SARS-CoV-2 Pepperprint® array
4	Analysis of Ref Seq SARS-CoV-2 – graphical output
5	Comparative analysis Ref Seq SARS-CoV-2 and UK and South African mutants
6	Additional Pepperprint® analysis of SARS-Cov-2
7	Legends for EigenBio Graphics in Addenda
8	IEDB recorded and EigenBio predicted allele specific peptide responses to SARS-Cov-2 (Excel)

Addenda

Each Addendum comprises multiple graphics of each protein. Contact DTRA for access to the spreadsheets of predictive values for epitopes.

1	Yellow fever
2	Hepatitis B
3	Human papillomavirus
4	Measles, mumps, and rubella
5	Rotavirus
6	Diphtheria and tetanus toxoids
7	Ebola
8	Hepatitis A
9	Rabies
10	Human herpes virus 3 (varicella)

References

1. Wilkinson TM, Li CK, Chui CS, Huang AK, Perkins M, Liebner JC, et al. Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans. Nat Med. 2012;18(2):274-80. Epub 2012/01/31. doi: 10.1038/nm.2612. PubMed PMID: 22286307.

2. Altmann DM, Boyton RJ. SARS-CoV-2 T cell immunity: Specificity, function, durability, and role in protection. Sci Immunol. 2020;5(49). Epub 2020/07/19. doi: 10.1126/sciimmunol.abd6160. PubMed PMID: 32680954.

3. Uittenbogaard JP, Zomer B, Hoogerhout P, Metz B. Reactions of beta-propiolactone with nucleobase analogues, nucleosides, and peptides: implications for the inactivation of viruses. J Biol Chem. 2011;286(42):36198-214. Epub 2011/08/27. doi: 10.1074/jbc.M111.279232. PubMed PMID: 21868382; PubMed Central PMCID: PMCPMC3196120.

4. Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, et al. Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. J Biol Chem. 2004;279(8):6235-43. doi: 10.1074/jbc.M310752200. PubMed PMID: 14638685.

5. Gershman MD, Angelo KM, Ritchey J, Greenberg DP, Muhammad RD, Brunette G, et al. Addressing a Yellow Fever Vaccine Shortage - United States, 2016-2017. MMWR Morbidity and mortality weekly report. 2017;66(17):457-9. Epub 2017/05/05. doi: 10.15585/mmwr.mm6617e2. PubMed PMID: 28472025; PubMed Central PMCID: PMCPMC5687078.

6. Gershman MD, Sotir MJ. Update: Temporary Total Depletion of U.S. Licensed Yellow Fever Vaccine for Civilian Travelers Addressed by Investigational New Drug Use of Imported Stamaril Vaccine. MMWR Morbidity and mortality weekly report. 2017;66(29):780. Epub 2017/07/28. doi: 10.15585/mmwr.mm6629a4. PubMed PMID: 28749924; PubMed Central PMCID: PMCPMC5657809.

7. Homan EJ, Bremel RD. Are cases of mumps in vaccinated patients attributable to mismatches in both vaccine T-cell and B-cell epitopes? An immunoinformatic analysis. Human vaccines & immunotherapeutics. 2013;10(2). Epub 2013/11/28. PubMed PMID: 24275080.

8. Dayhoff MOS, R. M.; Orcutt, B. C. . A model of evolutionary change in proteins. Atlas of Protein Sequence and Structure. 1978;5(3):345-52.

9. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. ProcNatlAcadSciUSA. 1992;89(22):10915-9.

10. Bremel RD, Homan EJ. An integrated approach to epitope analysis I: Dimensional reduction, visualization and prediction of MHC binding using amino acid principal components and regression approaches. Immunome research. 2010;6:7. Epub 2010/11/04. doi: 10.1186/1745-7580-6-7. PubMed PMID: 21044289; PubMed Central PMCID: PMC2990731.

11. Bremel RD, Homan EJ. An integrated approach to epitope analysis II: A system for proteomic-scale prediction of immunological characteristics. Immunome Res. 2010;6:8. Epub 2010/11/04. doi: 10.1186/1745-7580-6-8. PubMed PMID: 21044290; PubMed Central PMCID: PMCPMC2991286.

12. Bremel RD, Homan EJ. An integrated approach to epitope analysis II: A system for proteomic-scale prediction of immunological characteristics. ImmunomeRes. 2010;6(1):8. doi: 1745-7580-6-8 [pii];10.1186/1745-7580-6-8 [doi].

13. Homan EJ, Bremel RD. Patterns of Predicted T-Cell Epitopes Associated with Antigenic Drift in Influenza H3N2 Hemagglutinin. PLoSOne. 2011;6(10):e26711. doi: 10.1371/journal.pone.0026711 [doi];PONE-D-11-07616 [pii].

14. Bremel RD, Homan EJ. Recognition of higher order patterns in proteins: immunologic kernels. PloS one. 2013;8(7):e70115. Epub 2013/08/08. doi: 10.1371/journal.pone.0070115. PubMed PMID: 23922927.

15. Bremel RD, Homan EJ. Frequency Patterns of T-Cell Exposed Amino Acid Motifs in Immunoglobulin Heavy Chain Peptides Presented by MHCs. Frontiers in immunology. 2014;5:541. doi: 10.3389/fimmu.2014.00541. PubMed PMID: 25389426; PubMed Central PMCID: PMC4211557.

16. Bremel RD, Homan J. Extensive T-cell epitope repertoire sharing among human proteome, gastrointestinal microbiome, and pathogenic bacteria: Implications for the definition of self. Frontiers in immunology. 2015;6. doi: 10.3389/fimmu.2015.00538.

17. Homan EJ, Bremel RD. A Role for Epitope Networking in Immunomodulation by Helminths. Frontiers in immunology. 2018;9:1763. Epub 2018/08/16. doi: 10.3389/fimmu.2018.01763. PubMed PMID: 30108588; PubMed Central PMCID: PMCPMC6079203.

18. Rudolph MG, Stanfield RL, Wilson IA. How TCRs bind MHCs, peptides, and coreceptors. Annu Rev Immunol. 2006;24:419-66. Epub 2006/03/23. doi: 10.1146/annurev.immunol.23.021704.115658. PubMed PMID: 16551255.

19. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanovic S. SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics. 1999;50(3-4):213-9. doi: 90500213.251 [pii].

20. Jamshidi M, Vadiee, V., Ross, T.J. Fuzzy Logic And Control. 2 ed. Jamshidi M, editor. Engelwood Cliffs NJ: PTR Prentice Hall Inc.; 1993. 392 p.

21. Hoglund RA, Lossius A, Johansen JN, Homan J, Benth JS, Robins H, et al. In Silico Prediction Analysis of Idiotope-Driven T-B Cell Collaboration in Multiple Sclerosis. Frontiers in immunology. 2017;8:1255. doi: 10.3389/fimmu.2017.01255. PubMed PMID: 29038659; PubMed Central PMCID: PMCPMC5630699.

22. Morais SB, Figueiredo BC, Assis NRG, Homan J, Mambelli FS, Bicalho RM, et al. Schistosoma mansoni SmKI-1 or Its C-Terminal Fragment Induces Partial Protection Against S. mansoni Infection in Mice. Frontiers in immunology. 2018;9:1762. Epub 2018/08/15. doi: 10.3389/fimmu.2018.01762. PubMed PMID: 30105029; PubMed Central PMCID: PMCPMC6077287.

23. Specht CA, Lee CK, Huang H, Hester MM, Liu J, Luckie BA, et al. Vaccination with Recombinant Cryptococcus Proteins in Glucan Particles Protects Mice against Cryptococcosis in a Manner Dependent upon Mouse Strain and Cryptococcal Species. mBio. 2017;8(6). doi: 10.1128/mBio.01872-17. PubMed PMID: 29184017; PubMed Central PMCID: PMCPMC5705919.

24. Hoglund RA, Torsetnes SB, Lossius A, Bogen B, Homan EJ, Bremel R, et al. Human Cysteine Cathepsins Degrade Immunoglobulin G In Vitro in a Predictable Manner. Int J Mol Sci. 2019;20(19). Epub 2019/10/02. doi: 10.3390/ijms20194843. PubMed PMID: 31569504; PubMed Central PMCID: PMCPMC6801702.

25. Klein NP, Fireman B, Goddard K, Zerbo O, Asher J, Zhou J, et al. Vaccine effectiveness of cell-culture relative to egg-based inactivated influenza vaccine during the 2017-18 influenza season. PloS one. 2020;15(2):e0229279. Epub 2020/02/27. doi: 10.1371/journal.pone.0229279. PubMed PMID: 32101582; PubMed Central PMCID: PMCPMC7043803 following competing interests: NP Klein reports research support for unrelated studies from Sanofi Pasteur, GlaxoSmithKline, Pfizer, Merck, Protein Sciences (now Sanofi Pasteur), MedImmune and Dynavax. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

26. Boikos C, Sylvester GC, Sampalis JS, Mansi JA. Relative Effectiveness of the Cell-Cultured Quadrivalent Influenza Vaccine Compared to Standard, Egg-Derived Quadrivalent Influenza Vaccines in Preventing Influenza-Like Illness in 2017-2018. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2020. Epub 2020/04/08. doi: 10.1093/cid/ciaa371. PubMed PMID: 32253431; PubMed Central PMCID: PMCPMC7745007.

27. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579(7798):265-9. Epub 2020/02/06. doi: 10.1038/s41586-020-2008-3. PubMed PMID: 32015508; PubMed Central PMCID: PMCPMC7094943.

28. Dahlke C, Heidepriem, J., Kobbe, R., Santer, R., Koch, T., Fathi, A., Ly, M.L., Schmeidel, S., Seeberger, P.H., Addo, M.M., Loeffler, F.F. Distinct early IgA profile may determine severity of COVID-19 symptoms: an immunological case series. medRxiv. 2020. doi: https://doi.org/10.1101/2020.04.14.20059733.

29. Robin G, Sato Y, Desplancq D, Rochel N, Weiss E, Martineau P. Restricted diversity of antigen binding residues of antibodies revealed by computational alanine scanning of 227 antibody-antigen complexes. J Mol Biol. 2014;426(22):3729-43. doi: 10.1016/j.jmb.2014.08.013. PubMed PMID: 25174334.

30. Santos JC, Passos, G.A. The high infectivity of SARS-CoV-2 B.1.1.7 is associated with increased interaction force between Spike-ACE2 caused by the viral N501Y mutation. bioRxiv. 2021. doi: https://doi.org/10.1101/2020.12.29.424708

31. Oliveira SC, de Magalhaes MTQ, Homan EJ. Immunoinformatic Analysis of SARS-CoV-2 Nucleocapsid Protein and Identification of COVID-19 Vaccine Targets. Frontiers in immunology. 2020;11:587615. Epub 2020/11/17. doi: 10.3389/fimmu.2020.587615. PubMed PMID: 33193414; PubMed Central PMCID: PMCPMC7655779.

32. Shahhosseini N, Babuadze GG, Wong G, Kobinger GP. Mutation Signatures and In Silico Docking of Novel SARS-CoV-2 Variants of Concern. Microorganisms. 2021;9(5). Epub 2021/05/01. doi: 10.3390/microorganisms9050926. PubMed PMID: 33925854.

33. Tada T, Dcosta BM, Samanovic-Golden M, Herati RS, Cornelius A, Mulligan MJ, et al. Neutralization of viruses with European, South African, and United States SARS-CoV-2 variant spike proteins by convalescent sera and BNT162b2 mRNA vaccine-elicited antibodies. bioRxiv. 2021. Epub 2021/02/11. doi: 10.1101/2021.02.05.430003. PubMed PMID: 33564768; PubMed Central PMCID: PMCPMC7872356.

34. Xie X, Liu Y, Liu J, Zhang X, Zou J, Fontes-Garfias CR, et al. Neutralization of SARS-CoV-2 spike 69/70 deletion, E484K and N501Y variants by BNT162b2 vaccine-elicited sera. Nat Med. 2021;27(4):620-1. Epub 2021/02/10. doi: 10.1038/s41591-021-01270-4. PubMed PMID: 33558724.

35. Schwarz T, Heiss K, Mahendran Y, Casilag F, Kurth F, Sander LE, et al. SARS-CoV-2 Proteome-Wide Analysis Revealed Significant Epitope Signatures in COVID-19 Patients. Frontiers in immunology. 2021;12:629185. Epub 2021/04/10. doi: 10.3389/fimmu.2021.629185. PubMed PMID: 33833755; PubMed Central PMCID: PMCPMC8021850.

36. Anderson EM, Goodwin EC, Verma A, Arevalo CP, Bolton MJ, Weirick ME, et al. Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. Cell. 2021;184(7):1858-64 e10. Epub 2021/02/26. doi: 10.1016/j.cell.2021.02.010. PubMed PMID: 33631096; PubMed Central PMCID: PMCPMC7871851.

37. Dan JM, Mateus J, Kato Y, Hastie KM, Yu ED, Faliti CE, et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. Science. 2021;371(6529). Epub 2021/01/08. doi: 10.1126/science.abf4063. PubMed PMID: 33408181.

38. Dong J, Zost, S.J., Greaney, A.J. Starr, T.N., Dingens, A.S., Chen, E.C., Chen, R.E., Case, J.B., Sutton, R.E., Gilchuk, P., Rodriguez, J., Armstrong, E., Gainza, C., Nargi, R.S., Binshtein, E., Xie, X., Zhang, X., Shi, P-Y., Logue, D., Weston, S., McGrath, M.E., Frieman, M.B., Brady, T., Tuffy, K., Bright, H., Loo, Y-M., McTamney, P., Esser, M., Carnahan, R.H., Diamond, M.S., Bloom, J.D., Crowe, J.E. . Genetic and structural basis for recognition of SARS-CoV-2 spike protein by a two-antibody cocktail. biRviv. 2021;2021.01.27.428529. doi: doi.org/10.1101/2021.01.27.428529.

39. Voss C, Esmail S, Liu X, Knauer MJ, Ackloo S, Kaneko T, et al. Epitope-specific antibody responses differentiate COVID-19 outcome and variants of concern. JCI Insight. 2021. Epub 2021/06/04. doi: 10.1172/jci.insight.148855. PubMed PMID: 34081630.

40. Suryadevara N, Shrihari S, Gilchuk P, VanBlargan LA, Binshtein E, Zost SJ, et al. Neutralizing and protective human monoclonal antibodies recognizing the N-terminal domain of the SARS-CoV-2 spike protein. bioRxiv. 2021. Epub 2021/01/28. doi: 10.1101/2021.01.19.427324. PubMed PMID: 33501445; PubMed Central PMCID: PMCPMC7836115.

41. Ye Q, Lu, S., Corbett, K.D. Structural basis for SARS-CoV-2 Nucleocapsid protein recognition by single-domain antibodies. BioRxiv. 2021. doi: https://doi.org/10.1101/2021.06.01.446591.

42. Mateus J, Grifoni A, Tarke A, Sidney J, Ramirez SI, Dan JM, et al. Selective and cross-reactive SARS-CoV-2 T cell epitopes in unexposed humans. Science. 2020. Epub 2020/08/06. doi: 10.1126/science.abd3871. PubMed PMID: 32753554.

43. Poran A, Harjanto D, Malloy M, Arieta CM, Rothenberg DA, Lenkala D, et al. Sequence-based prediction of SARS-CoV-2 vaccine targets using a mass spectrometry-based bioinformatics predictor identifies immunogenic T cell epitopes. Genome Med. 2020;12(1):70. Epub 2020/08/15. doi: 10.1186/s13073-020-00767-w. PubMed PMID: 32791978; PubMed Central PMCID: PMCPMC7425796.

44. Snyder TM, Gittelman RM, Klinger M, May DH, Osborne EJ, Taniguchi R, et al. Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. medRxiv. 2020. Epub 2020/08/15. doi: 10.1101/2020.07.31.20165647. PubMed PMID: 32793919; PubMed Central PMCID: PMCPMC7418734.

45. Peng Y, Mentzer AJ, Liu G, Yao X, Yin Z, Dong D, et al. Broad and strong memory CD4(+) and CD8(+) T cells induced by SARS-CoV-2 in UK convalescent individuals following

COVID-19. Nat Immunol. 2020;21(11):1336-45. Epub 2020/09/06. doi: 10.1038/s41590-020-0782-6. PubMed PMID: 32887977.

46. Habel JR, Nguyen THO, van de Sandt CE, Juno JA, Chaurasia P, Wragg K, et al. Suboptimal SARS-CoV-2-specific CD8(+) T cell response associated with the prominent HLA-A*02:01 phenotype. Proc Natl Acad Sci U S A. 2020;117(39):24384-91. Epub 2020/09/12. doi: 10.1073/pnas.2015486117. PubMed PMID: 32913053; PubMed Central PMCID: PMCPMC7533701.

47. Sekine T, Perez-Potti A, Rivera-Ballesteros O, Stralin K, Gorin JB, Olsson A, et al. Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. Cell. 2020;183(1):158-68 e14. Epub 2020/09/28. doi: 10.1016/j.cell.2020.08.017. PubMed PMID: 32979941; PubMed Central PMCID: PMCPMC7427556.

48. Le Bert N, Clapham HE, Tan AT, Chia WN, Tham CYL, Lim JM, et al. Highly functional virus-specific cellular immune response in asymptomatic SARS-CoV-2 infection. bioRxiv. 2020;2020.11.25.399139.

49. Nielsen SSF, Vibholm LK, Monrad I, Olesen R, Frattari GS, Pahus MH, et al. SARS-CoV-2 elicits robust adaptive immune responses regardless of disease severity. bioRxiv. 2020;2020.10.08.331645.

50. Saini SK, Hersby DS, Tamhane T, Povlsen HR, Amaya-Hernandez SP, Nielsen M, et al. SARS-CoV-2 genome-wide mapping of CD8 T cell recognition reveals strong immunodominance and substantial CD8 T cell activation in COVID-19 patients. bioRxiv. 2020;2020.10.19.344911.

51. Gangaev A, Ketelaars SLC, Isaeva OI, Patiwael S, Dopler A, Hoefakker K, et al. Identification and characterization of a SARS-CoV-2 specific CD8(+) T cell response with immunodominant features. Nature communications. 2021;12(1):2593. Epub 2021/05/12. doi: 10.1038/s41467-021-22811-y. PubMed PMID: 33972535; PubMed Central PMCID: PMCPMC8110804.

52. Le Bert N, Tan AT, Kunasegaran K, Tham CYL, Hafezi M, Chia A, et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. Nature. 2020;584(7821):457-62. Epub 2020/07/16. doi: 10.1038/s41586-020-2550-z. PubMed PMID: 32668444.

53. Prakash S, Srivastava R, Coulon PG, Dhanushkodi NR, Chentoufi AA, Tifrea DF, et al. Genome-Wide Asymptomatic B-Cell, CD4 (+) and CD8 (+) T-Cell Epitopes, that are Highly Conserved Between Human and Animal Coronaviruses, Identified from SARS-CoV-2 as Immune Targets for Pre-Emptive Pan-Coronavirus Vaccines. bioRxiv. 2020. Epub 2020/10/08. doi: 10.1101/2020.09.27.316018. PubMed PMID: 33024971; PubMed Central PMCID: PMCPMC7536874.

54. Keller MD, Harris KM, Jensen-Wachspress MA, Kankate V, Lang H, Lazarski CA, et al. SARS-CoV-2 specific T-cells Are Rapidly Expanded for Therapeutic Use and Target Conserved Regions of Membrane Protein. Blood. 2020. Epub 2020/10/27. doi: 10.1182/blood.2020008488. PubMed PMID: 33104167.

55. Ferretti AP, Kula T, Wang Y, Nguyen DMV, Weinheimer A, Dunlap GS, et al. Unbiased Screens Show CD8(+) T Cells of COVID-19 Patients Recognize Shared Epitopes in SARS-CoV-2 that Largely Reside outside the Spike Protein. Immunity. 2020;53(5):1095-107 e3. Epub 2020/11/01. doi: 10.1016/j.immuni.2020.10.006. PubMed PMID: 33128877; PubMed Central PMCID: PMCPMC7574860.

56. Schulien I, Kemming J, Oberhardt V, Wild K, Seidel LM, Killmer S, et al. Characterization of pre-existing and induced SARS-CoV-2-specific CD8(+) T cells. Nat Med. 2021;27(1):78-85. Epub 2020/11/14. doi: 10.1038/s41591-020-01143-2. PubMed PMID: 33184509.

57. Chen Z, Ruan P, Wang L, Nie X, Ma X, Tan Y. T and B cell Epitope analysis of SARS-CoV-2 S protein based on immunoinformatics and experimental research. J Cell Mol Med. 2021;25(2):1274-89. Epub 2020/12/17. doi: 10.1111/jcmm.16200. PubMed PMID: 33325143; PubMed Central PMCID: PMCPMC7812294.

58. Shomuradova AS, Vagida MS, Sheetikov SA, Zornikova KV, Kiryukhin D, Titov A, et al. SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. Immunity. 2020;53(6):1245-57 e5. Epub 2020/12/17. doi: 10.1016/j.immuni.2020.11.004. PubMed PMID: 33326767; PubMed Central PMCID: PMCPMC7664363.

59. Rha MS, Jeong HW, Ko JH, Choi SJ, Seo IH, Lee JS, et al. PD-1-Expressing SARS-CoV-2-Specific CD8(+) T Cells Are Not Exhausted, but Functional in Patients with COVID-19. Immunity. 2021;54(1):44-52 e3. Epub 2020/12/19. doi: 10.1016/j.immuni.2020.12.002. PubMed PMID: 33338412; PubMed Central PMCID: PMCPMC7834198.

60. Minervina AA, Komech EA, Titov A, Bensouda Koraichi M, Rosati E, Mamedov IZ, et al. Longitudinal high-throughput TCR repertoire profiling reveals the dynamics of T-cell memory formation after mild COVID-19 infection. Elife. 2021;10. Epub 2021/01/06. doi: 10.7554/eLife.63502. PubMed PMID: 33399535; PubMed Central PMCID: PMCPMC7806265.

61. Kared H, Redd AD, Bloch EM, Bonny TS, Sumatoh H, Kairi F, et al. SARS-CoV-2-specific CD8+ T cell responses in convalescent COVID-19 individuals. J Clin Invest. 2021;131(5). Epub 2021/01/12. doi: 10.1172/JCI145476. PubMed PMID: 33427749; PubMed Central PMCID: PMCPMC7919723.

62. Joag V, Wijeyesinghe S, Stolley JM, Quarnstrom CF, Dileepan T, Soerens AG, et al. Cutting Edge: Mouse SARS-CoV-2 Epitope Reveals Infection and Vaccine-Elicited CD8 T Cell Responses. J Immunol. 2021;206(5):931-5. Epub 2021/01/15. doi: 10.4049/jimmunol.2001400. PubMed PMID: 33441437.

63. Szeto C, Chatzileontiadou DSM, Nguyen AT, Sloane H, Lobos CA, Jayasinghe D, et al. The presentation of SARS-CoV-2 peptides by the common HLA-A(*)02:01 molecule. iScience. 2021;24(2):102096. Epub 2021/02/02. doi: 10.1016/j.isci.2021.102096. PubMed PMID: 33521593; PubMed Central PMCID: PMCPMC7825995.

64. Tarke A, Sidney J, Kidd CK, Dan JM, Ramirez SI, Yu ED, et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. Cell Rep Med. 2021;2(2):100204. Epub 2021/02/02. doi: 10.1016/j.xcrm.2021.100204. PubMed PMID: 33521695; PubMed Central PMCID: PMCPMC7837622.

65. Hansbro PM, Starkey MR, Kim RY, Stevens RL, Foster PS, Horvat JC. Programming of the lung by early-life infection. Journal of developmental origins of health and disease. 2012;3(3):153-8. Epub 2012/06/01. doi: 10.1017/s2040174412000050. PubMed PMID: 25102006.

66. Agerer B, Koblischke M, Gudipati V, Montano-Gutierrez LF, Smyth M, Popa A, et al. SARS-CoV-2 mutations in MHC-I-restricted epitopes evade CD8(+) T cell responses. Sci Immunol. 2021;6(57). Epub 2021/03/06. doi: 10.1126/sciimmunol.abg6461. PubMed PMID: 33664060.

67. Bilich T, Nelde A, Heitmann JS, Maringer Y, Roerden M, Bauer J, et al. T cell and antibody kinetics delineate SARS-CoV-2 peptides mediating long-term immune responses in COVID-19 convalescent individuals. Science translational medicine. 2021;13(590). Epub 2021/03/17. doi: 10.1126/scitranslmed.abf7517. PubMed PMID: 33723016; PubMed Central PMCID: PMCPMC8128286.

68. Poluektov Y, George M, Daftarian P, Delcommenne MC. Assessment of SARS-CoV-2 specific CD4(+) and CD8 (+) T cell responses using MHC class I and II tetramers. Vaccine. 2021;39(15):2110-6. Epub 2021/03/22. doi: 10.1016/j.vaccine.2021.03.008. PubMed PMID: 33744048; PubMed Central PMCID: PMCPMC7934687.

69. Sahin U, Muik, A., Vogler, I., Derhovanessian, E., Kranz, L.M., et al . BNT162b2 induces SARS-CoV-2-neutralising antibodies and T cells in humans. MedRxiv. 2020;12.09.20245175. doi: https://doi.org/10.1101/2020.12.09.20245175.

70. Julander JG. Animal models of yellow fever and their application in clinical research. Curr Opin Virol. 2016;18:64-9. Epub 2016/04/20. doi: 10.1016/j.coviro.2016.03.010. PubMed PMID: 27093699.

71. van der Most RG, Harrington LE, Giuggio V, Mahar PL, Ahmed R. Yellow fever virus 17D envelope and NS3 proteins are major targets of the antiviral T cell response in mice. Virology. 2002;296(1):117-24. Epub 2002/05/31. doi: 10.1006/viro.2002.1432. PubMed PMID: 12036323.

72. Erickson AK, Pfeiffer JK. Spectrum of disease outcomes in mice infected with YFV-17D. The Journal of general virology. 2015;96(Pt 6):1328-39. Epub 2015/02/04. doi: 10.1099/vir.0.000075. PubMed PMID: 25646269; PubMed Central PMCID: PMCPMC4635484.

73. Overall SA, Toor JS, Hao S, Yarmarkovich M, Sara MOR, Morozov GI, et al. High throughput pMHC-I tetramer library production using chaperone-mediated peptide exchange. Nature communications. 2020;11(1):1909. Epub 2020/04/22. doi: 10.1038/s41467-020-15710-1. PubMed PMID: 32312993; PubMed Central PMCID: PMCPMC7170893.

74. Palmer S, Albergante L, Blackburn CC, Newman TJ. Thymic involution and rising disease incidence with age. Proc Natl Acad Sci U S A. 2018;115(8):1883-8. Epub 2018/02/13. doi: 10.1073/pnas.1714478115. PubMed PMID: 29432166; PubMed Central PMCID: PMCPMC5828591.

75. Murray JM, Kaufmann GR, Hodgkin PD, Lewin SR, Kelleher AD, Davenport MP, et al. Naive T cells are maintained by thymic output in early ages but by proliferation without phenotypic change after age twenty. Immunology and cell biology. 2003;81(6):487-95. Epub 2003/11/26. doi: 10.1046/j.1440-1711.2003.01191.x. PubMed PMID: 14636246.

Defense Threat Reduction Agency
8725 John J. Kingman Road, MS 6201
Fort Belvoir, VA 22060-6201

DTRA-TR-21-14

TECHNICAL REPORT - APPENDICES

Comparative Computational Modeling of Immune Reponses to Vaccines

Distribution Statement A. Approved for public release. Distribution unlimited.

HDTRA1-17-1-0052
Performer Name: ioGenetics, LLC & Versiti, Inc.

Prepared by: Robert Bremel and Jane Homan

June 2021

Appendix 1

- **Vaccines Licensed for Use in the United States** as of 1/30/2021

Product Name	Trade Name	
Adenovirus Type 4 and Type 7 Vaccine, Live, Oral	No Trade Name	Vaccinal isolate/sequence not specified
Anthrax Vaccine Adsorbed	Biothrax	Restricted use
BCG Live	BCG Vaccine	Complex >4000 proteins
BCG Live	TICE BCG	Complex >4000 proteins
Cholera Vaccine Live Oral	Vaxchora	Complex,acceptability
Dengue Tetravalent Vaccine, Live	DENGVAXIA	Approved during project for endemic areas
Diphtheria & Tetanus Toxoids Adsorbed	No Trade Name	Toxoid, multi organism
Diphtheria & Tetanus Toxoids & Acellular Pertussis Vaccine Adsorbed	Infanrix	Toxoid, multi organism
Diphtheria & Tetanus Toxoids & Acellular Pertussis Vaccine Adsorbed	DAPTACEL	Toxoid, multi organism
Diphtheria & Tetanus Toxoids & Acellular Pertussis Vaccine Adsorbed, Hepatitis B (recombinant) and Inactivated Poliovirus Vaccine Combined	Pediarix	Toxoid, multi organism
Diphtheria and Tetanus Toxoids and Acellular Pertussis Adsorbed and Inactivated Poliovirus Vaccine	KINRIX	Toxoid, multi organism, inactivated
Diphtheria and Tetanus Toxoids and Acellular Pertussis Adsorbed and Inactivated Poliovirus Vaccine	Quadracel	Toxoid, multi organism, inactivated
Diphtheria and Tetanus Toxoids and Acellular Pertussis Adsorbed, Inactivated Poliovirus, Haemophilus b Conjugate [Meningococcal Protein Conjugate] and Hepatitis B [Recombinant] Vaccine	VAXELIS	Conjugated
Diphtheria and Tetanus Toxoids and Acellular Pertussis Adsorbed, Inactivated Poliovirus and Haemophilus b Conjugate (Tetanus Toxoid Conjugate) Vaccine	Pentacel	Toxoid, multi organism
Ebola Zaire Vaccine, Live	ERVEBO	**Unavailable at MCW Travel clinic
Haemophilus b Conjugate Vaccine (Meningococcal Protein Conjugate)	PedvaxHIB	Conjugated
Haemophilus b Conjugate Vaccine (Tetanus Toxoid Conjugate)	ActHIB	Conjugated
Haemophilus b Conjugate Vaccine (Tetanus Toxoid Conjugate)	Hiberix	Conjugated
Hepatitis A Vaccine, Inactivated	Havrix	Inactivated
Hepatitis A Vaccine, Inactivated	VAQTA	Inactivated
Hepatitis A Inactivated and Hepatitis B (Recombinant) Vaccine	Twinrix	Inactivated
Hepatitis B Vaccine (Recombinant)	Recombivax HB	Single protein, simple

Product Name	Trade Name	
Hepatitis B Vaccine (Recombinant)	Engerix-B	Single protein, simple
Hepatitis B Vaccine (Recombinant), Adjuvanted	HEPLISAV-B	Single protein, simple
Human Papillomavirus Quadrivalent (Types 6, 11, 16, 18) Vaccine, Recombinant	Gardasil	Multivalent, adult acceptability prior exposure
Human Papillomavirus 9-valent Vaccine, Recombinant	Gardasil 9	Multivalent, adult acceptability prior exposure
Human Papillomavirus Bivalent (Types 16, 18) Vaccine, Recombinant	Cervarix	Multivalent, adult acceptability prior exposure
Influenza A (H1N1) 2009 Monovalent Vaccine	No Trade Name	Not available
Influenza A (H1N1) 2009 Monovalent Vaccine	No Trade Name	
Influenza A (H1N1) 2009 Monovalent Vaccine	No Trade Name	
Influenza A (H1N1) 2009 Monovalent Vaccine	No Trade Name	
Influenza A (H1N1) 2009 Monovalent Vaccine	No Trade Name	
Influenza Virus Vaccine, H5N1 (for National Stockpile)	No Trade Name	
Influenza A (H5N1) Virus Monovalent Vaccine, Adjuvanted	No Trade Name	
Influenza A (H5N1) Monovalent Vaccine, Adjuvanted	AUDENZ	
Influenza Vaccine, Adjuvanted	FLUAD QUADRIVALENT	Multivalent, compounded by prior exposures
Influenza Vaccine, Adjuvanted	FLUAD	
Influenza Vaccine	AFLURIA QUADRIVALENT	
Influenza Vaccine	Flucelvax Quadrivalent	
Influenza Virus Vaccine (Trivalent, Types A and B)	Afluria	
Influenza Virus Vaccine (Trivalent, Types A and B)	FluLaval	
Influenza Vaccine, Live, Intranasal (Trivalent, Types A and B)	FluMist	
Influenza Virus Vaccine (Trivalent, Types A and B)	Fluarix	
Influenza Virus Vaccine (Trivalent, Types A and B)	Fluvirin	
Influenza Virus Vaccine (Trivalent, Types A and B)	Agriflu	
Influenza Virus Vaccine (Trivalent, Types A and B)	Fluzone, Fluzone High-Dose and Fluzone Intradermal	

Product Name	Trade Name	
Influenza Virus Vaccine (Trivalent, Types A and B)	Flucelvax	
Influenza Vaccine (Trivalent)	Flublok	
Influenza Vaccine (Quadrivalent)	Flublok Quadrivalent	
Influenza Vaccine,Live, Intranasal (Quadrivalent, Types A and Types B)	FluMist Quadrivalent	
Influenza Virus Vaccine (Quadrivalent, Types A and Types B)	Fluarix Quadrivalent	
Influenza Virus Vaccine (Quadrivalent, Types A and Types B)	Fluzone Quadrivalent	
Influenza Vaccine	Flulaval Quadrivalent	
Japanese Encephalitis Virus Vaccine, Inactivated, Adsorbed	Ixiaro	Inactivated
Measles, Mumps, and Rubella Virus Vaccine, Live	M-M-R II	*** planned to include
Measles, Mumps, Rubella and Varicella Virus Vaccine Live	ProQuad	See above
Meningococcal (Groups A, C, Y, and W-135) Oligosaccharide Diphtheria CRM197 Conjugate Vaccine	Menveo	polysaccharide
Meningococcal (Groups A, C, Y and W-135) Polysaccharide Diphtheria Toxoid Conjugate Vaccine	Menactra	polysaccharide
Meningococcal Group B Vaccine	BEXSERO	polysaccharide
Meningococcal Group B Vaccine	TRUMENBA	polysaccharide
Meningococcal Polysaccharide Vaccine, Groups A, C, Y and W-135 Combined	Menomune-A/C/Y/W-135	polysaccharide
Meningococcal (Groups A, C, Y, W) Conjugate Vaccine	MenQuadfi	polysaccharide
Plague Vaccine	No trade name	Complex - access
Pneumococcal Vaccine, Polyvalent	Pneumovax 23	polysaccharide
Pneumococcal 13-valent Conjugate Vaccine (Diphtheria CRM_{197} Protein)	Prevnar 13	polysaccharide
Poliovirus Vaccine Inactivated (Human Diploid Cell)	Poliovax	inactivated
Poliovirus Vaccine Inactivated (Monkey Kidney Cell)	IPOL	inactivated
Rabies Vaccine	Imovax	Inactivated. Multi dose/ acceptability
Rabies Vaccine	RabAvert	Inactivated. Multi dose/ acceptability
Rabies Vaccine Adsorbed	No Trade Name	Inactivated. Multi dose/ acceptability

Product Name	Trade Name	
Rotavirus Vaccine, Live, Oral	ROTARIX	Live attenuated; unable to determine exact sequences used
Rotavirus Vaccine, Live, Oral, Pentavalent	RotaTeq	Live attenuated; unable to determine exact sequences used – multiple human bovine reassortants
Smallpox and Monkeypox Vaccine, Live, Non-Replicating	JYNNEOS	Restricted use
Smallpox (Vaccinia) Vaccine, Live	ACAM2000	Toxoid
Tetanus & Diphtheria Toxoids, Adsorbed	TDVAX	Toxoid
Tetanus & Diphtheria Toxoids Adsorbed for Adult Use	TENIVAC	Toxoid
Tetanus Toxoid Adsorbed	No Trade Name	Toxoid
Tetanus Toxoid, Reduced Diphtheria Toxoid and Acellular Pertussis Vaccine, Adsorbed	Adacel	Toxoid
Tetanus Toxoid, Reduced Diphtheria Toxoid and Acellular Pertussis Vaccine, Adsorbed	Boostrix	Toxoid
Typhoid Vaccine Live Oral Ty21a	Vivotif	
Typhoid Vi Polysaccharide Vaccine	TYPHIM Vi	polysaccharide
Varicella Virus Vaccine Live	Varivax	Superceded by Shingrix
Yellow Fever Vaccine	YF-Vax	*** live attenuated first choice; unavailable
Zoster Vaccine, Live, (Oka/Merck)	Zostavax	Superceded by Shingrix
Zoster Vaccine Recombinant, Adjuvanted	SHINGRIX	*** recombinant

Appendix 2

Influenza Vaccine Efficacy – Immunoinformatics observations of T cell epitopes

Seasonal influenza continues to exact a huge annual toll of illness and deaths. The 2017/18 season was one of the worst 'flu seasons for a decade, with approximately 60% of the cases in the US attributed to H3N2 [1]. The burden of seasonal influenza and the threat of emergence of an influenza strain with pandemic potential persists, creating urgency for better vaccines, better early warning and better rapid response capability [2].

Vaccine efficacy: Influenza viruses undergo continual antigenic drift. The challenge to vaccine development is to predict immunogens that will match influenza circulating in the coming season, and ideally to identify "universal" immunogens which will obviate the need to change the vaccine annually. Protection against influenza is closely correlated with the titer of hemagglutinin binding antibody. Vaccines are selected based on the degree to which antisera raised in ferrets neutralize the most recently circulating strains[3].
In the 2017/18 season vaccine efficacy (VE) for H3N2 was only 25% overall [1, 4]. Vaccines derived from cell culture (based on A/Singapore/GP2050/2015) provided better protection against the prevalent H3N2 wildtype than egg-grown viruses (based on A/Hong Kong/4801/2014-like strains including, A/New Caledonia/71/2014)[4]. In past years inadequate VE has variously been attributed to mutations which may occur during egg propagation, structural changes, and changes in glycosylation[5-7]. Various indices of VE prediction beyond ferret antiserum neutralization have been proposed[8, 9]. Vaccine selection based on antibody neutralization does not take into account the degree of correlation of T cell epitopes between vaccine strains and circulating wildtype. As our analysis shows this may be an important factor in VE. The active specific T helper response observed with live attenuated vaccines supports this[10]. It also raises questions as to the integrity of T cell epitopes following virus inactivation or splitting by b-propiolactone, formaldehyde or detergents[3, 11-14].

Role of T cell epitopes T cells are essential players in the response to influenza [15-18]. $CD4^+$ cells function both as cognate helpers to B cells in antibody class switch, affinity maturation and memory, and as enhancers of $CD8^+$ responses [16]. $CD8^+$ cytotoxic T cells limit infection by eliminating infected cells and provide heterosubtypic protection[19]. Curtailing the duration of replication in an individual reduces the number of variants which could emerge[20, 21]. $CD8^+$ may also elicit cytokine responses exacerbating inflammation[22, 23]. However, $CD8^+$ also generate T regulatory responses[24]. IL-10 may limit inflammation but may also be detrimental to protection from influenza [25, 26]. Importantly, as T cell responses are influenced by host immunogenetics, individuals and sub-populations may respond differently and contribute differently to transmission.

Of critical importance for the design of vaccines, $CD4^+$ T cell help for influenza antibody responses is specific to a particular B cell epitope [10, 27, 28]. We have shown predicted T cell epitopes lie adjacent to and overlapping the principal protective B cell epitopes[29]. Vaccine derived neutralizing antibody directed to the HA is protective while the titer is maintained. However, antibody titers wane after vaccination in the absence of re-stimulation[30, 31]. Exposure to a wildtype virus which matches a neutralizing B cell epitope, but which does not match the vaccinal T cell epitope will result in a slower response as a *de novo* Th population is stimulated, so antibody titers to subdominant or heterosubtypic epitopes may rise faster than antibodies which bind the dominant neutralizing epitope [16]. Indeed, vaccination with a T cell epitope that is mismatched in one or more major protective epitopes of HA could actually exacerbate disease on the next exposure, by favoring antibody development to subdominant epitopes where vaccine and wildtype T helper epitopes coincide, rather than to the needed new protective epitope. By the same mechanism, repeated exposure to influenza virus or vaccine may explain the apparent dominance of the T cell epitopes in the conserved matrix M1 and NP proteins [32]. Unlike antibody responses, T cell responses depend on host immunogenetics. Despite

the dependence on ferrets in influenza research, only recently has attention been paid to the T cell responses of this outbred species [33] and no characterization of ferret MHC exists.

Every amino acid change, as evidenced in antibody antigenic drift, inevitably also contributes to drift of adjacent or overlapping T cell epitopes which account for the specific help. We have shown previously that T cell epitope patterns provide grouping and cartography of influenza viruses similar to those provided by antibodies [29, 34]. Structural and synthetic approaches to universal vaccine design overlook the essential role of specific T cell epitopes and underestimate the complexity of the immune response [35-37]. The apparent importance of ensuring a Th match in a vaccine also draws attention to the necessity of considering host immunogenetics. Differing HLA may lead to differential levels of protection, or extended virus shedding in the absence of protection, in different subpopulations. This also exerts different pressures favoring escape mutants. Fortunately, MHC II peptide binding tends to cluster in overlapping peptides [29, 38]. As we show in preliminary data (below), predicted T cell epitopes in H3N2 egg-derived vaccine in 2017/18 were poorly matched to the predominant circulating virus adjacent to a critical protective B cell epitope, whereas cell-derived vaccine had closely correlated predicted T cell epitopes.

Immune responses to viral infections comprise a balance of effector responses that clear infection vs regulatory or immune suppressive responses that modulate immunopathology [15, 23, 39]. A robust $CD8^+$ response is needed to limit the duration of virus shedding and to eliminate virus infected cells [17]; $CD4^+$ are needed to help both antigen and cytotoxic responses. But countering this a Tregulatory response may be essential to dampen the immunopathology arising from inflammation [23]. When these responses become inbalanced more severe diseases may result from excessive inflammation or from immune suppression and extended viral replication.

Observations on H3N2 2017 Vaccine efficiency

Note: *Because we analyze all HA types in parallel, based on raw FASTA sequences, we adopted an "absolute" numbering convention and have not used H3N2 numbering in the examples shown.*

Vaccine related observations: We have previously reported on the clustering of H3N2 isolates over time, based on HA T cell epitopes, showing how T cell epitopes drift with the amino acid changes that cause antibody antigen drift [29]. Hence, T cell mapping can offer an alternate approach to antigenic cartography. We have also described the close association of MHC I and II binding with B cell epitopes as a composite ordered, multifactorial antigenic signal [38] (now shown experimentally by others in the context of influenza [27, 28]). We have also shown how single amino acid changes in cell culture adaptation in development of a vaccine strain may result in a T cell epitope mismatch in mumpsvirus vaccine [40].

We mapped the HA of all 2017/18 vaccine strains [41]; relative to NYMC egg-based strain X263 A/Hong Kong/4801/2014, the alternate strains recommended vary by 1-6 amino acids. In particular, we compared egg-based isolates with A/Singapore/GP2050/2015 recommended for cell-based vaccines. As each amino acid mutation impacts epitopes upstream and downstream up to 23 aa (15-mer plus cathepsin cleavage octomer), a 6 aa difference in 566 amino acids actually results in a difference in 56% of its T cell epitopes. Figure 1 shows the population permuted average binding of the most common DRB and MHC I A and B alleles for the egg and cell-based vaccine HAs overlaid by the linear B cell epitopes. Within the population average there are of course differences in impact by allele; Figure 2 compares the predicted peptide binding for DRB1:1201 across epitope B region [42]. Importantly, while these graphics plot *quantitative* differences in predicted binding affinity, this reflects *qualitative* differences in the actual amino acids and TCEM from each sequential 15-mer, and thus which T cell clones are cognate for the pentamer motifs engaging the TCR. In Figure 2 we also show predicted MHC II binding for 100aa spanning Epitope B of the HA of both egg and cell-based vaccines and overlay one wild type isolate (A/New York/52/2017) which represents the majority of the US H3N2 isolates in 2017/18. The T cell epitopes of the wildtype more closely match the cell-based vaccine.

Hence, following vaccination and exposure to wildtype re-stimulation of cognate T cell clones would occur more rapidly for the cell vaccine than with the egg vaccine, which would be more likely to elicit only a heterosubtypic response.

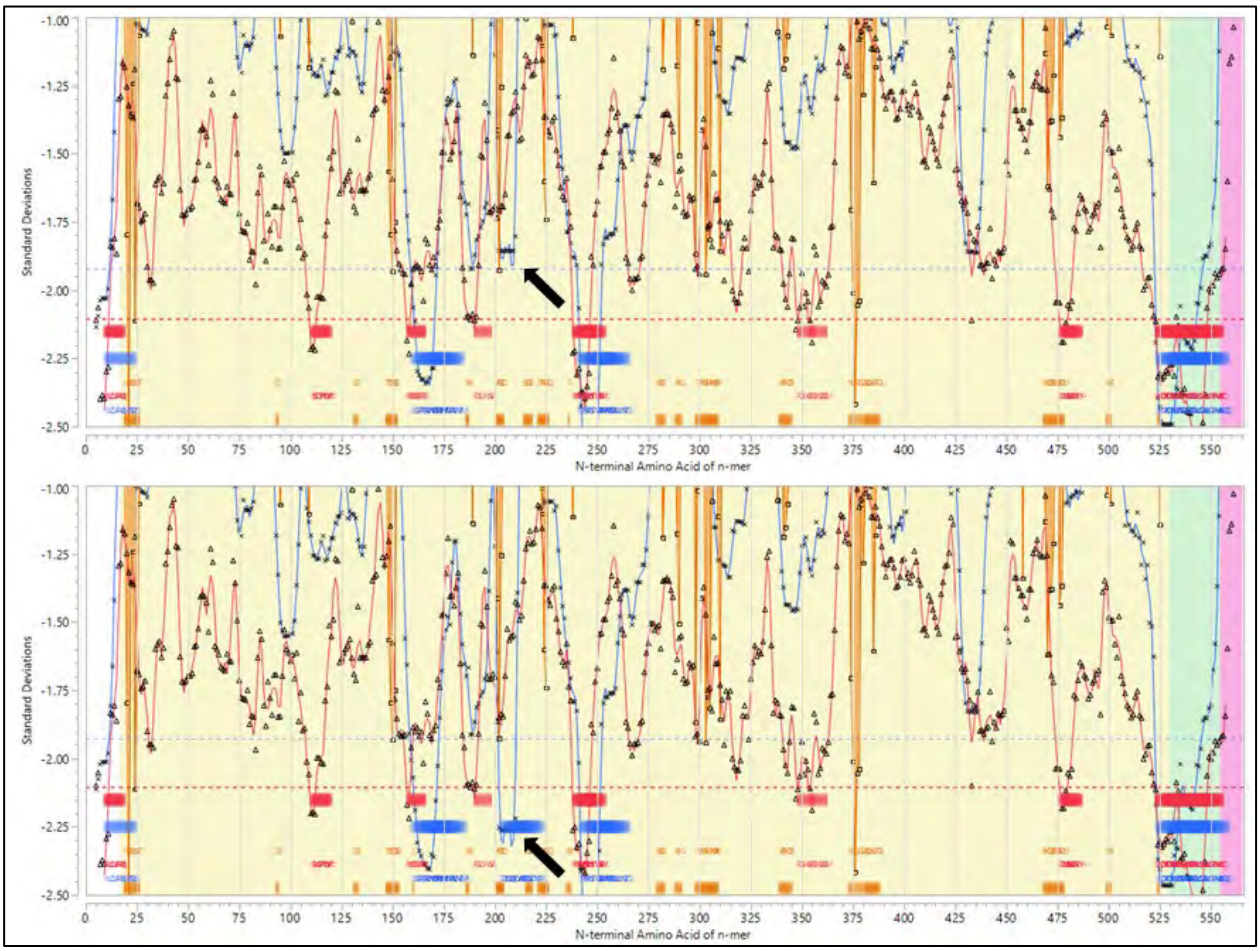
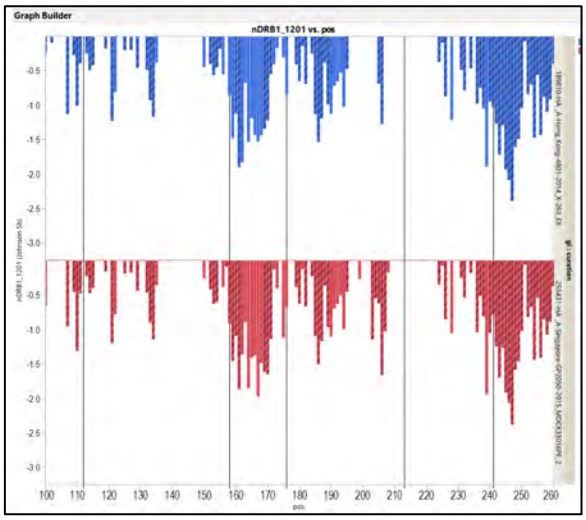


Figure 1: Comparison of egg, cell and wild type H3N2 2017/18: Egg vaccine (top) shows less MHC II binding (indicated by blue line and ribbon) than cell vaccine, in region centered on aa 205 (absolute numbering) for permuted average of common DR alleles.



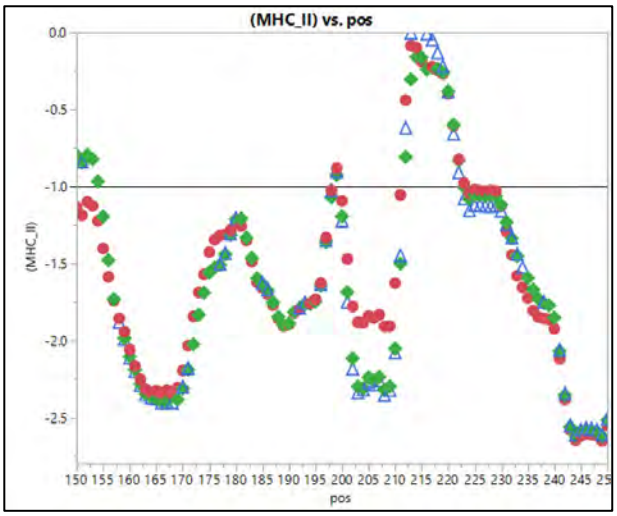


Figure 2: Left: Comparison of Egg vaccine (blue) vs Cell based vaccine (red) for DRB1:1201. Y axis shows predicted binding in standard deviation units; X axis shows successive 15mer peptides by index position from aa100-250. Vertical lines show amino acid differences; hashed bars are peptides predicted to be excised by cathepsins. Right: Shows how one exemplary wildtype isolate A/New York-52/2017 (representative of 70% 2017/18 H3N2 isolates (triangle) matches T cell epitope profile of the cell vaccine (green diamond) more closely than the egg-based vaccine (red circles).

Postscript

Publications in 2020 in which the 2017 seasonal influenza vaccine response was analyzed document the greater efficacy of the cell culture derived vaccine [43, 44].

References

1. Flannery B. Preliminary estimates of 2017-18 seasonal Influenza vaccine effectiveness agains laboratory confirmed influenza from the US Flu VE and HAIVEN networks: Advisory Committee on Immunization Practices (ACIP), Centers for Disease Control and Prevention; 2018 [06/20/2018]. Available from: https://www.cdc.gov/vaccines/acip/meetings/slides-2018-06.html.

2. Erbelding EJ, Post DJ, Stemmy EJ, Roberts PC, Augustine AD, Ferguson S, et al. A Universal Influenza Vaccine: The Strategic Plan for the National Institute of Allergy and Infectious Diseases. J Infect Dis. 2018;218(3):347-54. Epub 2018/03/06. doi: 10.1093/infdis/jiy103. PubMed PMID: 29506129.

3. Stohr K, Bucher D, Colgate T, Wood J. Influenza virus surveillance, vaccine strain selection, and manufacture. Methods Mol Biol. 2012;865:147-62. Epub 2012/04/25. doi: 10.1007/978-1-61779-621-0_9. PubMed PMID: 22528158.

4. Lu Y. Relative effectiveness of cell-cultured versus egg-based vaccines 2017-18: Advisory Committee on Immunization Practices (ACIP), Centers for Disease Control and Prevention; 2018 [06/20/2018]. Available from: https://www.cdc.gov/vaccines/acip/meetings/slides-2018-06.html.

5. Wu NC, Zost SJ, Thompson AJ, Oyen D, Nycholat CM, McBride R, et al. A structural explanation for the low effectiveness of the seasonal influenza H3N2 vaccine. PLoS pathogens. 2017;13(10):e1006682. Epub 2017/10/24. doi: 10.1371/journal.ppat.1006682. PubMed PMID: 29059230; PubMed Central PMCID: PMCPMC5667890.

6. Zost SJ, Parkhouse K, Gumina ME, Kim K, Diaz Perez S, Wilson PC, et al. Contemporary H3N2 influenza viruses have a glycosylation site that alters binding of antibodies elicited by egg-adapted vaccine strains. Proc Natl Acad Sci U S A. 2017;114(47):12578-83. Epub 2017/11/08. doi: 10.1073/pnas.1712377114. PubMed PMID: 29109276; PubMed Central PMCID: PMCPMC5703309.

7. Cobey S, Gouma S, Parkhouse K, Chambers BS, Ertl HC, Schmader KE, et al. Poor Immunogenicity, Not Vaccine Strain Egg Adaptation, May Explain the Low H3N2 Influenza Vaccine Effectiveness in 2012-2013. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2018;67(3):327-33. Epub 2018/02/23. doi: 10.1093/cid/ciy097. PubMed PMID: 29471464; PubMed Central PMCID: PMCPMC6051447.

8. Bonomo ME, Deem MW. Predicting Influenza H3N2 Vaccine Efficacy From Evolution of the Dominant Epitope. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2018;67(7):1129-31. Epub 2018/04/20. doi: 10.1093/cid/ciy323. PubMed PMID: 29672670.

9. Paessler S, Veljkovic V. Prediction of influenza vaccine effectiveness for the influenza season 2017/18 in the US. F1000Res. 2017;6:2067. Epub 2018/03/15. doi: 10.12688/f1000research.13198.1. PubMed PMID: 29535862; PubMed Central PMCID: PMCPMC5795273.

10. Aljurayyan A, Puksuriwong S, Ahmed M, Sharma R, Krishnan M, Sood S, et al. Activation and Induction of Antigen-Specific T Follicular Helper Cells Play a Critical Role in Live-Attenuated Influenza Vaccine-Induced Human Mucosal Anti-influenza Antibody Response. Journal of virology. 2018;92(11). Epub 2018/03/23. doi: 10.1128/JVI.00114-18. PubMed PMID: 29563292; PubMed Central PMCID: PMCPMC5952133.

11. Uittenbogaard JP, Zomer B, Hoogerhout P, Metz B. Reactions of beta-propiolactone with nucleobase analogues, nucleosides, and peptides: implications for the inactivation of viruses. J Biol Chem. 2011;286(42):36198-214. doi: 10.1074/jbc.M111.279232. PubMed PMID: 21868382; PubMed Central PMCID: PMCPMC3196120.

12. She YM, Cheng K, Farnsworth A, Li X, Cyr TD. Surface modifications of influenza proteins upon virus inactivation by beta-propiolactone. Proteomics. 2013;13(23-24):3537-47. Epub 2013/10/15. doi: 10.1002/pmic.201300096. PubMed PMID: 24123778; PubMed Central PMCID: PMCPMC4265195.

13. Astill J, Alkie T, Yitbarek A, Taha-Abdelaziz K, Bavananthasivam J, Nagy E, et al. Examination of the effects of virus inactivation methods on the induction of antibody- and cell-mediated immune responses against whole inactivated H9N2 avian influenza virus vaccines in chickens. Vaccine. 2018;36(27):3908-16. Epub 2018/06/02. doi: 10.1016/j.vaccine.2018.05.093. PubMed PMID: 29853199.

14. Kon TC, Onu A, Berbecila L, Lupulescu E, Ghiorgisor A, Kersten GF, et al. Influenza Vaccine Manufacturing: Effect of Inactivation, Splitting and Site of Manufacturing. Comparison of Influenza Vaccine Production Processes. PloS one. 2016;11(3):e0150700. Epub 2016/03/10. doi: 10.1371/journal.pone.0150700. PubMed PMID: 26959983; PubMed Central PMCID: PMCPMC4784929.

15. Sant AJ, DiPiazza AT, Nayak JL, Rattan A, Richards KA. CD4 T cells in protection from influenza virus: Viral antigen specificity and functional potential. Immunol Rev. 2018;284(1):91-105. Epub 2018/06/27. doi: 10.1111/imr.12662. PubMed PMID: 29944766; PubMed Central PMCID: PMCPMC6070306.

16. Sant AJ, Richards KA, Nayak J. Distinct and complementary roles of CD4 T cells in protective immunity to influenza virus. Curr Opin Immunol. 2018;53:13-21. Epub 2018/04/06. doi: 10.1016/j.coi.2018.03.019. PubMed PMID: 29621639.

17. Keating R, Morris MY, Yue W, Reynolds CE, Harris TL, Brown SA, et al. Potential killers exposed: tracking endogenous influenza-specific CD8(+) T cells. Immunology and cell biology. 2018. Epub 2018/07/05. doi: 10.1111/imcb.12189. PubMed PMID: 29972699.

18. McKinstry KK, Strutt TM, Buck A, Curtis JD, Dibble JP, Huston G, et al. IL-10 deficiency unleashes an influenza-specific Th17 response and enhances survival against high-dose challenge. J Immunol. 2009;182(12):7353-63. Epub 2009/06/06. doi: 10.4049/jimmunol.0900657. PubMed PMID: 19494257; PubMed Central PMCID: PMCPMC2724021.

19. Sridhar S, Begom S, Bermingham A, Hoschler K, Adamson W, Carman W, et al. Cellular immune correlates of protection against symptomatic pandemic influenza. Nat Med. 2013;19(10):1305-12. Epub 2013/09/24. doi: 10.1038/nm.3350. PubMed PMID: 24056771.

20. Murcia PR, Baillie GJ, Daly J, Elton D, Jervis C, Mumford JA, et al. Intra- and interhost evolutionary dynamics of equine influenza virus. JVirol. 2010;84(14):6943-54. doi: JVI.00112-10 [pii];10.1128/JVI.00112-10 [doi].

21. Iqbal M, Xiao H, Baillie G, Warry A, Essen SC, Londt B, et al. Within-host variation of avian influenza viruses. PhilosTransRSocLond B BiolSci. 2009;364(1530):2739-47. doi: 364/1530/2739 [pii];10.1098/rstb.2009.0088 [doi].

22. Perrone LA, Plowden JK, Garcia-Sastre A, Katz JM, Tumpey TM. H5N1 and 1918 pandemic influenza virus infection results in early and excessive infiltration of macrophages and neutrophils in the lungs of mice. PLoS pathogens. 2008;4(8):e1000115. Epub 2008/08/02. doi: 10.1371/journal.ppat.1000115. PubMed PMID: 18670648; PubMed Central PMCID: PMCPMC2483250.

23. Duan S, Thomas PG. Balancing Immune Protection and Immune Pathology by CD8(+) T-Cell Responses to Influenza Infection. Frontiers in immunology. 2016;7:25. Epub 2016/02/24. doi: 10.3389/fimmu.2016.00025. PubMed PMID: 26904022; PubMed Central PMCID: PMCPMC4742794.

24. Zou Q, Wu B, Xue J, Fan X, Feng C, Geng S, et al. CD8+ Treg cells suppress CD8+ T cell-responses by IL-10-dependent mechanism during H5N1 influenza virus infection. European journal of immunology. 2014;44(1):103-14. Epub 2013/10/12. doi: 10.1002/eji.201343583. PubMed PMID: 24114149; PubMed Central PMCID: PMCPMC4165276.

25. Sun K, Torres L, Metzger DW. A detrimental effect of interleukin-10 on protective pulmonary humoral immunity during primary influenza A virus infection. Journal of virology. 2010;84(10):5007-14. Epub 2010/03/05. doi: 10.1128/JVI.02408-09. PubMed PMID: 20200252; PubMed Central PMCID: PMCPMC2863832.

26. Sun J, Madan R, Karp CL, Braciale TJ. Effector T cells control lung inflammation during acute influenza virus infection by producing IL-10. Nat Med. 2009;15(3):277-84. Epub 2009/02/24. doi: 10.1038/nm.1929. PubMed PMID: 19234462; PubMed Central PMCID: PMCPMC2693210.

27. Knowlden ZA, Sant AJ. CD4 T cell epitope specificity determines follicular versus non-follicular helper differentiation in the polyclonal response to influenza infection or vaccination. Scientific reports. 2016;6:28287. Epub 2016/06/23. doi: 10.1038/srep28287. PubMed PMID: 27329272; PubMed Central PMCID: PMCPMC4916409.

28. DiPiazza A, Laniewski N, Rattan A, Topham DJ, Miller J, Sant AJ. CD4 T Cell Epitope Specificity and Cytokine Potential Are Preserved as Cells Transition from the Lung Vasculature to Lung Tissue following Influenza Virus Infection. Journal of virology. 2018;92(13). Epub 2018/04/20. doi: 10.1128/JVI.00377-18. PubMed PMID: 29669836; PubMed Central PMCID: PMCPMC6002724.

29. Homan EJ, Bremel RD. Patterns of Predicted T-Cell Epitopes Associated with Antigenic Drift in Influenza H3N2 Hemagglutinin. PLoSOne. 2011;6(10):e26711. doi: 10.1371/journal.pone.0026711 [doi];PONE-D-11-07616 [pii].

30. Ray GT, Lewis N, Klein NP, Daley MF, Wang SV, Kulldorff M, et al. Intra-season Waning of Influenza Vaccine Effectiveness. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2018. Epub 2018/09/12. doi: 10.1093/cid/ciy770. PubMed PMID: 30204855.

31. Belongia EA, Sundaram ME, McClure DL, Meece JK, Ferdinands J, VanWormer JJ. Waning vaccine protection against influenza A (H3N2) illness in children and older adults during a single season. Vaccine. 2015;33(1):246-51. Epub 2014/06/26. doi: 10.1016/j.vaccine.2014.06.052. PubMed PMID: 24962752.

32. Zens KD, Farber DL. Memory CD4 T cells in influenza. Current topics in microbiology and immunology. 2015;386:399-421. Epub 2014/07/10. doi: 10.1007/82_2014_401. PubMed PMID: 25005927; PubMed Central PMCID: PMCPMC4339101.

33. Reber AJ, Music N, Kim JH, Gansebom S, Chen J, York I. Extensive T cell cross-reactivity between diverse seasonal influenza strains in the ferret model. Scientific reports. 2018;8(1):6112. Epub 2018/04/19. doi: 10.1038/s41598-018-24394-z. PubMed PMID: 29666412; PubMed Central PMCID: PMCPMC5904180.

34. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus AD, et al. Mapping the antigenic and genetic evolution of influenza virus. Science. 2004;305(5682):371-6. doi: 10.1126/science.1097211 [doi];1097211 [pii].

35. Wang TT, Tan GS, Hai R, Pica N, Ngai L, Ekiert DC, et al. Vaccination with a synthetic peptide from the influenza virus hemagglutinin provides protection against distinct viral subtypes. ProcNatlAcadSciUSA. 2010;107(44):18979-84. doi: 1013387107 [pii];10.1073/pnas.1013387107 [doi].

36. Van Regenmortel MHV. Development of a Preventive HIV Vaccine Requires Solving Inverse Problems Which Is Unattainable by Rational Vaccine Design. Frontiers in immunology. 2017;8:2009. Epub 2018/02/02. doi: 10.3389/fimmu.2017.02009. PubMed PMID: 29387066; PubMed Central PMCID: PMCPMC5776009.

37. Moise L, B MB, Boyle CM, Kurt Yilmaz N, Jang H, Schiffer C, et al. T cell epitope engineering: an avian H7N9 influenza vaccine strategy for pandemic preparedness and response. Human vaccines & immunotherapeutics. 2018;14(9):2203-7. Epub 2018/07/18. doi: 10.1080/21645515.2018.1495303. PubMed PMID: 30015562.

38. Bremel RD, Homan EJ. Recognition of higher order patterns in proteins: immunologic kernels. PloS one. 2013;8(7):e70115. Epub 2013/08/08. doi: 10.1371/journal.pone.0070115. PubMed PMID: 23922927.

39. Clemens EB, van de Sandt C, Wong SS, Wakim LM, Valkenburg SA. Harnessing the Power of T Cells: The Promising Hope for a Universal Influenza Vaccine. Vaccines (Basel). 2018;6(2). Epub 2018/03/29. doi: 10.3390/vaccines6020018. PubMed PMID: 29587436; PubMed Central PMCID: PMCPMC6027237.

40. Homan EJ, Bremel RD. Are cases of mumps in vaccinated patients attributable to mismatches in both vaccine T-cell and B-cell epitopes? An immunoinformatic analysis. Human vaccines & immunotherapeutics. 2013;10(2). Epub 2013/11/28. PubMed PMID: 24275080.

41. WHO. Recommendations on the composition of influenza vaccines, 1998-present Geneva. Available from: http://www.who.int/influenza/vaccines/virus/recommendations/en/.

42. Wiley DC, Wilson IA, Skehel JJ. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature. 1981;289(5796):373-8.

43. Klein NP, Fireman B, Goddard K, Zerbo O, Asher J, Zhou J, et al. Vaccine effectiveness of cell-culture relative to egg-based inactivated influenza vaccine during the 2017-18 influenza season. PloS one. 2020;15(2):e0229279. Epub 2020/02/27. doi: 10.1371/journal.pone.0229279. PubMed PMID: 32101582; PubMed Central PMCID: PMCPMC7043803 following competing interests: NP Klein reports research support for unrelated studies from Sanofi Pasteur, GlaxoSmithKline, Pfizer, Merck, Protein Sciences (now Sanofi Pasteur), MedImmune and Dynavax. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

44. Boikos C, Sylvester GC, Sampalis JS, Mansi JA. Relative Effectiveness of the Cell-Cultured Quadrivalent Influenza Vaccine Compared to Standard, Egg-Derived Quadrivalent Influenza Vaccines

in Preventing Influenza-Like Illness in 2017-2018. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2020. Epub 2020/04/08. doi: 10.1093/cid/ciaa371. PubMed PMID: 32253431; PubMed Central PMCID: PMCPMC7745007.

Appendix 3: Peptides included in SARS-COV-3 Microarray

Peptides and B cell epitope probabilities are numbered by their index position, but reflect the B cell epitope centered at that index position; hence array was constructed. Note the chip payout is included in Addendum #11

gi	index position	curation	peptide	lag peptides	B_epi penta	BEPI	control
1796318598	12	surface glycoprotein	SSQCVNLTTRTQLPP	LVSSQCVNLTTRTQL	LVSSQ	0.25	neg
	25		PPAYTNSFTRGVYYP	QLPPAYTNSFTRGVY	QLPPA	-1.34	
	26		PAYTNSFTRGVYYPD	LPPAYTNSFTRGVYY	LPPAY	-1.31	
	44		RSSVLHSTQDLFLPF	VFRSSVLHSTQDLFL	VFRSS	0.05	neg
	52		QDLFLPFFSNVTWFH	STQDLFLPFFSNVTW	STQDL	0.19	neg
	55		FLPFFSNVTWFHAIH	DLFLPFFSNVTWFHA	DLFLP	1.18	neg
	61		NVTWFHAIHVSGTNG	FSNVTWFHAIHVSGT	FSNVT	1.25	neg
	66		HAIHVSGTNGTKRFD	WFHAIHVSGTNGTKR	WFHAI	1.67	neg
	69		HVSGTNGTKRFDNPV	AIHVSGTNGTKRFDN	AIHVS	0.67	neg
	72		GTNGTKRFDNPVLPF	VSGTNGTKRFDNPVL	VSGTN	-1.43	
	73		TNGTKRFDNPVLPFN	SGTNGTKRFDNPVLP	SGTNG	-1.68	
	74		NGTKRFDNPVLPFND	GTNGTKRFDNPVLPF	GTNGT	-1.80	
	75		GTKRFDNPVLPFNDG	TNGTKRFDNPVLPFN	TNGTK	-1.73	
	76		TKRFDNPVLPFNDGV	NGTKRFDNPVLPFND	NGTKR	-1.66	
	77		KRFDNPVLPFNDGVY	GTKRFDNPVLPFNDG	GTKRF	-1.45	
	89		GVYFASTEKSNIIRG	NDGVYFASTEKSNII	NDGVY	0.53	neg
	106		FGTTLDSKTQSLLIV	WIFGTTLDSKTQSLL	WIFGT	0.90	neg
	112		SKTQSLLIVNNATNV	LDSKTQSLLIVNNAT	LDSKT	-1.35	
	117		LLIVNNATNVVIKVC	QSLLIVNNATNVVIK	QSLLI	0.38	neg
	138		DPFLGVYYHKNNKSW	CNDPFLGVYYHKNNK	CNDPF	0.50	neg
	182		KQGNFKNLREFVFKN	EGKQGNFKNLREFVF	EGKQG	-1.35	
	183		QGNFKNLREFVFKNI	GKQGNFKNLREFVFK	GKQGN	-1.47	
	184		GNFKNLREFVFKNID	KQGNFKNLREFVFKN	KQGNF	-1.45	
	191		EFVFKNIDGYFKIYS	LREFVFKNIDGYFKI	LREFV	0.81	neg
	195		KNIDGYFKIYSKHTP	VFKNIDGYFKIYSKH	VFKNI	0.77	neg
	225		PLVDLPIGINITRFQ	LEPLVDLPIGINITR	LEPLV	0.19	neg
	231		IGINITRFQTLLALH	LPIGINITRFQTLLA	LPIGI	0.60	neg
	244		LHRSYLTPGDSSSGW	LALHRSYLTPGDSSS	LALHR	1.08	neg
	250		TPGDSSSGWTAGAAA	YLTPGDSSSGWTAGA	YLTPG	-1.42	
	251		PGDSSSGWTAGAAAY	LTPGDSSSGWTAGAA	LTPGD	-1.77	
	252		GDSSSGWTAGAAAYY	TPGDSSSGWTAGAAA	TPGDS	-1.96	
	253		DSSSGWTAGAAAYYV	PGDSSSGWTAGAAAY	PGDSS	-2.07	
	254		SSSGWTAGAAAYYVG	GDSSSGWTAGAAAYY	GDSSS	-2.06	
	255		SSGWTAGAAAYYVGY	DSSSGWTAGAAAYYV	DSSSG	-1.92	
	256		SGWTAGAAAYYVGYL	SSSGWTAGAAAYYVG	SSSGW	-1.98	
	257		GWTAGAAAYYVGYLQ	SSGWTAGAAAYYVGY	SSGWT	-1.86	
	258		WTAGAAAYYVGYLQP	SGWTAGAAAYYVGYL	SGWTA	-1.63	
	276		LLKYNENGTITDAVD	TFLLKYNENGTITDA	TFLLK	1.07	neg
	282		NGTITDAVDCALDPL	NENGTITDAVDCALD	NENGT	-1.41	
	283		GTITDAVDCALDPLS	ENGTITDAVDCALDP	ENGTI	-1.58	
	284		TITDAVDCALDPLSE	NGTITDAVDCALDPL	NGTIT	-1.45	
	339		GEVFNATRFASVYAW	PFGEVFNATRFASVY	PFGEV	0.58	neg
	352		AWNRKRISNCVADYS	VYAWNRKRISNCVAD	VYAWN	0.35	neg
	365		YSVLYNSASFSTFKC	ADYSVLYNSASFSTF	ADYSV	0.42	neg
	397		ADSFVIRGDEVRQIA	VYADSFVIRGDEVRQ	VYADS	0.60	neg
	411		APGQTGKIADYNYKL	QIAPGQTGKIADYNY	QIAPG	-1.44	
	412		PGQTGKIADYNYKLP	IAPGQTGKIADYNYK	IAPGQ	-1.63	
	413		GQTGKIADYNYKLPD	APGQTGKIADYNYKL	APGQT	-1.75	
	414		QTGKIADYNYKLPDD	PGQTGKIADYNYKLP	PGQTG	-1.69	
	415		TGKIADYNYKLPDDF	GQTGKIADYNYKLPD	GQTGK	-1.49	
	436		WNSNNLDSKVGGNYN	IAWNSNNLDSKVGGN	IAWNS	0.39	neg
	443		SKVGGNYNYLYRLFR	LDSKVGGNYNYLYRL	LDSKV	-1.38	
	444		KVGGNYNYLYRLFRK	DSKVGGNYNYLYRLF	DSKVG	-1.38	
	445		VGGNYNYLYRLFRKS	SKVGGNYNYLYRLFR	SKVGG	-1.35	
	453		YRLFRKSNLKPFERD	YLYRLFRKSNLKPFE	YLYRL	1.38	neg
	455		LFRKSNLKPFERDIS	YRLFRKSNLKPFERD	YRLFR	1.12	neg
	477		STPCNGVEGFNCYFP	AGSTPCNGVEGFNCY	AGSTP	-1.48	

	478		TPCNGVEGFNCYFPL	GSTPCNGVEGFNCYF	GSTPC	-1.62	
	479		PCNGVEGFNCYFPLQ	STPCNGVEGFNCYFP	STPCN	-1.56	
	480		CNGVEGFNCYFPLQS	TPCNGVEGFNCYFPL	TPCNG	-1.46	
	493		QSYGFQPTNGVGYQP	PLQSYGFQPTNGVGY	PLQSY	0.38	neg
	499		PTNGVGYQPYRVVVL	FQPTNGVGYQPYRVV	FQPTN	-1.31	
	512		VLSFELLHAPATVCG	VVVLSFELLHAPATV	VVVLS	1.70	neg
	527		PKKSTNLVKNKCVNF	CGPKKSTNLVKNKCV	CGPKK	-1.39	
	528		KKSTNLVKNKCVNFN	GPKKSTNLVKNKCVN	GPKKS	-1.45	
	529		KSTNLVKNKCVNFNF	PKKSTNLVKNKCVNF	PKKST	-1.57	
	530		STNLVKNKCVNFNFN	KKSTNLVKNKCVNFN	KKSTN	-1.50	
	531		TNLVKNKCVNFNFNG	KSTNLVKNKCVNFNF	KSTNL	-1.43	
	573		TDAVRDPQTLEILDI	DTTDAVRDPQTLEIL	DTTDA	-1.32	
	574		DAVRDPQTLEILDIT	TTDAVRDPQTLEILD	TTDAV	-1.34	
	575		AVRDPQTLEILDITP	TDAVRDPQTLEILDI	TDAVR	-1.34	
	576		VRDPQTLEILDITPC	DAVRDPQTLEILDIT	DAVRD	-1.36	
	577		RDPQTLEILDITPCS	AVRDPQTLEILDITP	AVRDP	-1.44	
	578		DPQTLEILDITPCSF	VRDPQTLEILDITPC	VRDPQ	-1.37	
	600		PGTNTSNQVAVLYQD	ITPGTNTSNQVAVLY	ITPGT	-1.61	
	601		GTNTSNQVAVLYQDV	TPGTNTSNQVAVLYQ	TPGTN	-1.88	
	602		TNTSNQVAVLYQDVN	PGTNTSNQVAVLYQD	PGTNT	-2.08	
	603		NTSNQVAVLYQDVNC	GTNTSNQVAVLYQDV	GTNTS	-2.04	
	604		TSNQVAVLYQDVNCT	TNTSNQVAVLYQDVN	TNTSN	-1.88	
	605		SNQVAVLYQDVNCTE	NTSNQVAVLYQDVNC	NTSNQ	-1.56	
	615		VNCTEVPVAIHADQL	QDVNCTEVPVAIHAD	QDVNC	0.12	neg
	666		IGAGICASYQTQTNS	IPIGAGICASYQTQT	IPIGA	0.12	neg
	676		TQTNSPRRARSVASQ	YQTQTNSPRRARSVA	YQTQT	-1.50	
	677		QTNSPRRARSVASQS	QTQTNSPRRARSVAS	QTQTN	-1.77	
	678		TNSPRRARSVASQSI	TQTNSPRRARSVASQ	TQTNS	-2.10	
	679		NSPRRARSVASQSII	QTNSPRRARSVASQS	QTNSP	-2.26	
	680		SPRRARSVASQSIIA	TNSPRRARSVASQSI	TNSPR	-2.25	
	681		PRRARSVASQSIIAY	NSPRRARSVASQSII	NSPRR	-2.09	
	682		RRARSVASQSIIAYT	SPRRARSVASQSIIA	SPRRA	-1.91	
	683		RARSVASQSIIAYTM	PRRARSVASQSIIAY	PRRAR	-1.73	
	684		ARSVASQSIIAYTMS	RRARSVASQSIIAYT	RRARS	-1.45	
	697		MSLGAENSVAYSNNS	YTMSLGAENSVAYSN	YTMSL	0.35	neg
	751		NLLLQYGSFCTQLNR	CSNLLLQYGSFCTQL	CSNLL	0.35	neg
	759		FCTQLNRALTGIAVE	GSFCTQLNRALTGIA	GSFCT	0.33	neg
	775		DKNTQEVFAQVKQIY	EQDKNTQEVFAQVKQ	EQDKN	-1.62	
	776		KNTQEVFAQVKQIYK	QDKNTQEVFAQVKQI	QDKNT	-1.70	
	777		NTQEVFAQVKQIYKT	DKNTQEVFAQVKQIY	DKNTQ	-1.54	
	778		TQEVFAQVKQIYKTP	KNTQEVFAQVKQIYK	KNTQE	-1.34	
	785		VKQIYKTPPIKDFGG	AQVKQIYKTPPIKDF	AQVKQ	0.47	neg
	787		QIYKTPPIKDFGGFN	VKQIYKTPPIKDFGG	VKQIY	0.09	neg
	808		DPSKPSKRSFIEDLL	LPDPSKPSKRSFIED	LPDPS	-1.69	
	809		PSKPSKRSFIEDLLF	PDPSKPSKRSFIEDL	PDPSK	-2.23	
	810		SKPSKRSFIEDLLFN	DPSKPSKRSFIEDLL	DPSKP	-2.49	
	811		KPSKRSFIEDLLFNK	PSKPSKRSFIEDLLF	PSKPS	-2.50	
	812		PSKRSFIEDLLFNKV	SKPSKRSFIEDLLFN	SKPSK	-2.39	
	813		SKRSFIEDLLFNKVT	KPSKRSFIEDLLFNK	KPSKR	-2.13	
	814		KRSFIEDLLFNKVTL	PSKRSFIEDLLFNKV	PSKRS	-1.85	
	815		RSFIEDLLFNKVTLA	SKRSFIEDLLFNKVT	SKRSF	-1.32	
	836		QYGDCLGDIAARDLI	IKQYGDCLGDIAARD	IKQYG	0.05	neg
	837		YGDCLGDIAARDLIC	KQYGDCLGDIAARDL	KQYGD	0.05	neg
	848		DLICAQKFNGLTVLP	ARDLICAQKFNGLTV	ARDLI	0.33	neg
	853		QKFNGLTVLPPLLTD	CAQKFNGLTVLPPLL	CAQKF	0.23	neg
	916		LYENQKLIANQFNSA	NVLYENQKLIANQFN	NVLYE	0.00	neg
	938		LSSTASALGKLQDVV	DSLSSTASALGKLQD	DSLSS	-1.36	
	939		SSTASALGKLQDVVN	SLSSTASALGKLQDV	SLSST	-1.59	
	940		STASALGKLQDVVNQ	LSSTASALGKLQDVV	LSSTA	-1.70	
	941		TASALGKLQDVVNQN	SSTASALGKLQDVVN	SSTAS	-1.74	
	942		ASALGKLQDVVNQNA	STASALGKLQDVVNQ	STASA	-1.66	
	943		SALGKLQDVVNQNAQ	TASALGKLQDVVNQN	TASAL	-1.49	
	975		SVLNDILSRLDKVEA	ISSVLNDILSRLDKV	ISSVL	0.13	neg
	1047		YHLMSFPQSAPHGVV	KGYHLMSFPQSAPHG	KGYHL	0.69	neg
	1095		FVSNGTHWFVTQRNF	GVFVSNGTHWFVTQR	GVFVS	0.05	neg
	1104		VTQRNFYEPQIITTD	WFVTQRNFYEPQIIT	WFVTQ	0.86	neg

	1128		VVIGIVNNTVYDPLQ	CDVVIGIVNNTVYDP	CDVVI	0.80	neg
	1160		TSPDVDLGDISGINA	NHTSPDVDLGDISGI	NHTSP	-1.46	
	1161		SPDVDLGDISGINAS	HTSPDVDLGDISGIN	HTSPD	-1.51	
	1162		PDVDLGDISGINASV	TSPDVDLGDISGINA	TSPDV	-1.49	
	1175		SVVNIQKEIDRLNEV	NASVVNIQKEIDRLN	NASVV	0.02	neg
	1200		LQELGKYEQYIKWPW	IDLQELGKYEQYIKW	IDLQE	0.19	neg
	1218		LGFIAGLIAIVMVTI	IWLGFIAGLIAIVMV	IWLGF	2.59	neg
	1219		GFIAGLIAIVMVTIM	WLGFIAGLIAIVMVT	WLGFI	2.37	neg
	1258		EDDSEPVLKGVKLHY	FDEDDSEPVLKGVKL	FDEDD	-1.35	
	1259		DDSEPVLKGVKLHYT	DEDDSEPVLKGVKLH	DEDDS	-1.55	
	1260		DSEPVLKGVKLHYT	EDDSEPVLKGVKLHY	EDDSE	-1.69	
	1261		SEPVLKGVKLHYT	DDSEPVLKGVKLHYT	DDSEP	-1.67	
1796318599	21	ORF3a protein	KDATPSDFVRATATI	EIKDATPSDFVRATA	EIKDA	-1.49	
	22		DATPSDFVRATATIP	IKDATPSDFVRATAT	IKDAT	-1.62	
	23		ATPSDFVRATATIPI	KDATPSDFVRATATI	KDATP	-1.68	
	24		TPSDFVRATATIPIQ	DATPSDFVRATATIP	DATPS	-1.66	
	25		PSDFVRATATIPIQA	ATPSDFVRATATIPI	ATPSD	-1.53	
	42		PFGWLIVGVALLAVF	SLPFGWLIVGVALLA	SLPFG	0.69	neg
	52		LLAVFQSASKIITLK	VALLAVFQSASKIIT	VALLA	1.83	neg
	55		VFQSASKIITLKKRW	LAVFQSASKIITLKK	LAVFQ	0.64	neg
	58		SASKIITLKKRWQLA	FQSASKIITLKKRWQ	FQSAS	0.04	neg
	64		TLKKRWQLALSKGVH	IITLKKRWQLALSKG	IITLK	0.13	neg
	75		KGVHFVCNLLLLFVT	LSKGVHFVCNLLLLF	LSKGV	0.27	neg
	103		APFLYLYALVYFLQS	LEAPFLYLYALVYFL	LEAPF	0.24	neg
	108		LYALVYFLQSINFVR	LYLYALVYFLQSINF	LYLYA	2.60	neg
	125		MRLWLCWKCRSKNPL	IIMRLWLCWKCRSKN	IIMRL	2.84	neg
	134		RSKNPLLYDANYFLC	KCRSKNPLLYDANYF	KCRSK	0.07	neg
	172		GDGTTSPISEHDYQI	TSGDGTTSPISEHDY	TSGDG	-1.65	
	173		DGTTSPISEHDYQIG	SGDGTTSPISEHDYQ	SGDGT	-1.91	
	174		GTTSPISEHDYQIGG	GDGTTSPISEHDYQI	GDGTT	-2.15	
	175		TTSPISEHDYQIGGY	DGTTSPISEHDYQIG	DGTTS	-2.30	
	176		TSPISEHDYQIGGYT	GTTSPISEHDYQIGG	GTTSP	-2.33	
	177		SPISEHDYQIGGYTE	TTSPISEHDYQIGGY	TTSPI	-2.15	
	178		PISEHDYQIGGYTEK	TSPISEHDYQIGGYT	TSPIS	-1.69	
	179		ISEHDYQIGGYTEKW	SPISEHDYQIGGYTE	SPISE	-1.31	
	198		KDCVVLHSYFTSDYY	GVKDCVVLHSYFTSD	GVKDC	0.02	neg
	200		CVVLHSYFTSDYYQL	KDCVVLHSYFTSDYY	KDCVV	0.89	neg
	201		VVLHSYFTSDYYQLY	DCVVLHSYFTSDYYQ	DCVVL	1.16	neg
	204		HSYFTSDYYQLYSTQ	VLHSYFTSDYYQLYS	VLHSY	1.22	neg
	205		SYFTSDYYQLYSTQL	LHSYFTSDYYQLYST	LHSYF	0.87	neg
	208		TSDYYQLYSTQLSTD	YFTSDYYQLYSTQLS	YFTSD	0.13	neg
	210		DYYQLYSTQLSTDTG	TSDYYQLYSTQLSTD	TSDYY	0.06	neg
	220		STDTGVEHVTFFIYN	QLSTDTGVEHVTFFI	QLSTD	-1.37	
	221		TDTGVEHVTFFIYNK	LSTDTGVEHVTFFIY	LSTDT	-1.50	
	222		DTGVEHVTFFIYNKI	STDTGVEHVTFFIYN	STDTG	-1.47	
	223		TGVEHVTFFIYNKIV	TDTGVEHVTFFIYNK	TDTGV	-1.35	
	236		IVDEPEEHVQIHTID	NKIVDEPEEHVQIHT	NKIVD	0.30	neg
1796 3186	13	envelope protein	IVNSVLLFLAFVVFL	TLIVNSVLLFLAFVV	TLIVN	0.57	neg
	40		CAYCCNIVNVSLVKP	RLCAYCCNIVNVSLV	RLCAY	1.69	neg
	48		NVSLVKPSFYVYSRV	IVNVSLVKPSFYVYS	IVNVS	0.98	neg
	49		VSLVKPSFYVYSRVK	VNVSLVKPSFYVYSR	VNVSL	0.77	neg
1796318601	12	membrane glycoprotein	ELKKLLEQWNLVIGF	VEELKKLLEQWNLVI	VEELK	0.30	neg
	13		LKKLLEQWNLVIGFL	EELKKLLEQWNLVIG	EELKK	0.46	neg
	20		WNLVIGFLFLTWICL	EQWNLVIGFLFLTWI	EQWNL	1.07	neg
	23		VIGFLFLTWICLLQF	NLVIGFLFLTWICLL	NLVIG	2.12	neg
	42		RNRFLYIIKLIFLWL	ANRNRFLYIIKLIFL	ANRNR	0.02	neg
	46		LYIIKLIFLWLLWPV	RFLYIIKLIFLWLLW	RFLYI	1.60	neg
	61		TLACFVLAAVYRINW	PVTLACFVLAAVYRI	PVTLA	1.79	neg
	78		GGIAIAMACLVGLMW	ITGGIAIAMACLVGL	ITGGI	0.36	neg
	87		LVGLMWLSYFIASFR	ACLVGLMWLSYFIAS	ACLVG	2.01	neg
	105		RTRSMWSFNPETNIL	FARTRSMWSFNPETN	FARTR	0.12	neg
	122		VPLHGTILTRPLLES	LNVPLHGTILTRPLL	LNVPL	0.81	neg
	124		LHGTILTRPLLESEL	VPLHGTILTRPLLES	VPLHG	0.52	neg
	133		LLESELVIGAVILRG	RPLLESELVIGAVIL	RPLLE	0.09	neg
	136		SELVIGAVILRGHLR	LESELVIGAVILRGH	LESEL	0.30	neg
	144		ILRGHLRIAGHHLGR	AVILRGHLRIAGHHL	AVILR	1.26	neg

	179		YKLGASQRVAGDSGF	SYYKLGASQRVAGDS	SYYKL	0.08	neg
	185		QRVAGDSGFAAYSRY	ASQRVAGDSGFAAYS	ASQRV	-1.33	
	186		RVAGDSGFAAYSRYR	SQRVAGDSGFAAYSR	SQRVA	-1.37	
	187		VAGDSGFAAYSRYRI	QRVAGDSGFAAYSRY	QRVAG	-1.36	
	188		AGDSGFAAYSRYRIG	RVAGDSGFAAYSRYR	RVAGD	-1.36	
	208		TDHSSSSDNIALLVQ	LNTDHSSSSDNIALL	LNTDH	-1.42	
	209		DHSSSSDNIALLVQ	NTDHSSSSDNIALLV	NTDHS	-1.77	
	210		HSSSSDNIALLVQ	TDHSSSSDNIALLVQ	TDHSS	-2.01	
1802476807	12	nsp3	VIEVQGYKSVNITFE	DTVIEVQGYKSVNIT	DTVIE	0.00	neg
	24		TFELDERIDKVLNEK	NITFELDERIDKVLN	NITFE	0.32	neg
	51		NEFACVVADAVIKTL	EVNEFACVVADAVIK	EVNEF	0.07	neg
	52		EFACVVADAVIKTLQ	VNEFACVVADAVIKT	VNEFA	0.34	neg
	55		CVVADAVIKTLQPVS	FACVVADAVIKTLQP	FACVV	0.96	neg
	59		DAVIKTLQPVSELLT	VADAVIKTLQPVSEL	VADAV	0.77	neg
	61		VIKTLQPVSELLTPL	DAVIKTLQPVSELLT	DAVIK	0.47	neg
	100		SHMYCSFYPPDEDEE	LASHMYCSFYPPDED	LASHM	0.75	neg
	102		MYCSFYPPDEDEEEG	SHMYCSFYPPDEDEE	SHMYC	1.16	neg
	109		PDEDEEEGDCEEEEF	YPPDEDEEEGDCEEE	YPPDE	-1.43	
	110		DEDEEEGDCEEEEFE	PPDEDEEEGDCEEEE	PPDED	-1.81	
	111		EDEEEGDCEEEEFEP	PDEDEEEGDCEEEEF	PDEDE	-2.04	
	112		DEEEGDCEEEEFEPS	DEDEEEGDCEEEEFE	DEDEE	-2.12	
	113		EEEGDCEEEEFEPST	EDEEEGDCEEEEFEP	EDEEE	-2.18	
	114		EEGDCEEEEFEPSTQ	DEEEGDCEEEEFEPS	DEEEG	-2.19	
	115		EGDCEEEEFEPSTQY	EEEGDCEEEEFEPST	EEEGD	-2.11	
	116		GDCEEEEFEPSTQYE	EEGDCEEEEFEPSTQ	EEGDC	-2.04	
	117		DCEEEEFEPSTQYEY	EGDCEEEEFEPSTQY	EGDCE	-1.61	
	118		CEEEEFEPSTQYEYG	GDCEEEEFEPSTQYE	GDCEE	-1.33	
	122		EFEPSTQYEYGTEDD	EEEFEPSTQYEYGTE	EEEFE	-1.32	
	123		FEPSTQYEYGTEDDY	EEFEPSTQYEYGTED	EEFEP	-1.32	
	124		EPSTQYEYGTEDDYQ	EFEPSTQYEYGTEDD	EFEPS	-1.31	
	125		PSTQYEYGTEDDYQG	FEPSTQYEYGTEDDY	FEPST	-1.33	
	130		EYGTEDDYQGKPLEF	QYEYGTEDDYQGKPL	QYEYG	-1.33	
	132		GTEDDYQGKPLEFGA	EYGTEDDYQGKPLEF	EYGTE	-1.35	
	133		TEDDYQGKPLEFGAT	YGTEDDYQGKPLEFG	YGTED	-1.37	
	134		EDDYQGKPLEFGATS	GTEDDYQGKPLEFGA	GTEDD	-1.50	
	135		DDYQGKPLEFGATSA	TEDDYQGKPLEFGAT	TEDDY	-1.71	
	136		DYQGKPLEFGATSAA	EDDYQGKPLEFGATS	EDDYQ	-1.69	
	137		YQGKPLEFGATSAAL	DDYQGKPLEFGATSA	DDYQG	-1.58	
	138		QGKPLEFGATSAALQ	DYQGKPLEFGATSAA	DYQGK	-1.43	
	151		LQPEEEQEEDWLDDD	AALQPEEEQEEDWLD	AALQP	-1.43	
	152		QPEEEQEEDWLDDDS	ALQPEEEQEEDWLDD	ALQPE	-1.70	
	153		PEEEQEEDWLDDDSQ	LQPEEEQEEDWLDDD	LQPEE	-1.91	
	154		EEEQEEDWLDDDSQQ	QPEEEQEEDWLDDDS	QPEEE	-2.06	
	155		EEQEEDWLDDDSQQT	PEEEQEEDWLDDDSQ	PEEEQ	-1.95	
	156		EQEEDWLDDDSQQTV	EEEQEEDWLDDDSQQ	EEEQE	-1.79	
	157		QEEDWLDDDSQQTVG	EEQEEDWLDDDSQQT	EEQEE	-1.54	
	158		EEDWLDDDSQQTVGQ	EQEEDWLDDDSQQTV	EQEED	-1.33	
	165		DSQQTVGQQDGSEDN	DDDSQQTVGQQDGSE	DDDSQ	-1.32	
	166		SQQTVGQQDGSEDNQ	DDSQQTVGQQDGSED	DDSQQ	-1.58	
	167		QQTVGQQDGSEDNQT	DSQQTVGQQDGSEDN	DSQQT	-1.76	
	168		QTVGQQDGSEDNQTT	SQQTVGQQDGSEDNQ	SQQTV	-1.95	
	169		TVGQQDGSEDNQTTT	QQTVGQQDGSEDNQT	QQTVG	-1.91	
	170		VGQQDGSEDNQTTTI	QTVGQQDGSEDNQTT	QTVGQ	-1.83	
	171		GQQDGSEDNQTTTIQ	TVGQQDGSEDNQTTT	TVGQQ	-1.80	
	172		QQDGSEDNQTTTIQT	VGQQDGSEDNQTTTI	VGQQD	-1.85	
	173		QDGSEDNQTTTIQTI	GQQDGSEDNQTTTIQ	GQQDG	-2.00	
	174		DGSEDNQTTTIQTIV	QQDGSEDNQTTTIQT	QQDGS	-2.24	
	175		GSEDNQTTTIQTIVE	QDGSEDNQTTTIQTI	QDGSE	-2.41	
	176		SEDNQTTTIQTIVEV	DGSEDNQTTTIQTIV	DGSED	-2.46	
	177		EDNQTTTIQTIVEVQ	GSEDNQTTTIQTIVE	GSEDN	-2.47	
	178		DNQTTTIQTIVEVQP	SEDNQTTTIQTIVEV	SEDNQ	-2.49	
	179		NQTTTIQTIVEVQPQ	EDNQTTTIQTIVEVQ	EDNQT	-2.33	
	180		QTTTIQTIVEVQPQL	DNQTTTIQTIVEVQP	DNQTT	-2.04	
	181		TTTIQTIVEVQPQLE	NQTTTIQTIVEVQPQ	NQTTT	-1.71	
	182		TTIQTIVEVQPQLEM	QTTTIQTIVEVQPQL	QTTTI	-1.41	
	190		VQPQLEMELTPVVQT	VEVQPQLEMELTPVV	VEVQP	0.03	neg

	199		TPVVQTIEVNSFSGY	ELTPVVQTIEVNSFS	ELTPV	0.31	neg
	242		AANVYLKHGGGVAGA	VNAANVYLKHGGGVA	VNAAN	0.44	neg
	243		ANVYLKHGGGVAGAL	NAANVYLKHGGGVAG	NAANV	0.35	neg
	261		TNNAMQVESDDYIAT	KATNNAMQVESDDYI	KATNN	-1.33	
	304		VNKGEDIQLLKSAYE	PNVNKGEDIQLLKSA	PNVNK	-1.40	
	305		NKGEDIQLLKSAYEN	NVNKGEDIQLLKSAY	NVNKG	-1.47	
	306		KGEDIQLLKSAYENF	VNKGEDIQLLKSAYE	VNKGE	-1.38	
	312		LLKSAYENFNQHEVL	IQLLKSAYENFNQHE	IQLLK	0.24	neg
	313		LKSAYENFNQHEVLL	QLLKSAYENFNQHEV	QLLKS	0.34	neg
	314		KSAYENFNQHEVLLA	LLKSAYENFNQHEVL	LLKSA	0.21	neg
	345		RVCVDTVRTNVYLAV	SLRVCVDTVRTNVYL	SLRVC	1.01	neg
	373		LEMKSEKQVEQKIAE	SFLEMKSEKQVEQKI	SFLEM	0.21	neg
	379		KQVEQKIAEIPKEEV	SEKQVEQKIAEIPKE	SEKQV	-1.41	
	380		QVEQKIAEIPKEEVK	EKQVEQKIAEIPKEE	EKQVE	-1.52	
	401		KPSVEQRKQDDKKIK	ESKPSVEQRKQDDKK	ESKPS	-1.44	
	402		PSVEQRKQDDKKIKA	SKPSVEQRKQDDKKI	SKPSV	-1.75	
	403		SVEQRKQDDKKIKAC	KPSVEQRKQDDKKIK	KPSVE	-1.90	
	404		VEQRKQDDKKIKACV	PSVEQRKQDDKKIKA	PSVEQ	-1.98	
	405		EQRKQDDKKIKACVE	SVEQRKQDDKKIKAC	SVEQR	-2.03	
	406		QRKQDDKKIKACVEE	VEQRKQDDKKIKACV	VEQRK	-1.90	
	407		RKQDDKKIKACVEEV	EQRKQDDKKIKACVE	EQRKQ	-1.84	
	408		KQDDKKIKACVEEVT	QRKQDDKKIKACVEE	QRKQD	-1.81	
	409		QDDKKIKACVEEVTT	RKQDDKKIKACVEEV	RKQDD	-1.88	
	410		DDKKIKACVEEVTTT	KQDDKKIKACVEEVT	KQDDK	-1.82	
	411		DKKIKACVEEVTTTL	QDDKKIKACVEEVTT	QDDKK	-1.62	
	412		KKIKACVEEVTTTLE	DDKKIKACVEEVTTT	DDKKI	-1.37	
	461		LKKDAPYIVGDVVQE	TFLKKDAPYIVGDVV	TFLKK	0.45	neg
	480		AVVIPTKKAGGTTEM	LTAVVIPTKKAGGTT	LTAVV	0.87	neg
	486		KKAGGTTEMLAKALR	PTKKAGGTTEMLAKA	PTKKA	-1.39	
	487		KAGGTTEMLAKALRK	TKKAGGTTEMLAKAL	TKKAG	-1.72	
	488		AGGTTEMLAKALRKV	KKAGGTTEMLAKALR	KKAGG	-1.93	
	489		GGTTEMLAKALRKVP	KAGGTTEMLAKALRK	KAGGT	-2.22	
	490		GTTEMLAKALRKVPT	AGGTTEMLAKALRKV	AGGTT	-2.25	
	491		TTEMLAKALRKVPTD	GGTTEMLAKALRKVP	GGTTE	-1.92	
	492		TEMLAKALRKVPTDN	GTTEMLAKALRKVPT	GTTEM	-1.37	
	533		SAFYILPSIISNEKQ	CKSAFYILPSIISNE	CKSAF	0.03	neg
	535		FYILPSIISNEKQEI	SAFYILPSIISNEKQ	SAFYI	0.58	neg
	558		REMLAHAEETRKLMP	NLREMLAHAEETRKL	NLREM	0.72	neg
	607		YTSKTTVASLINTLN	YFYTSKTTVASLINT	YFYTS	0.39	neg
	650		SLKVPATVSVSSPDA	MRSLKVPATVSVSSP	MRSLK	0.35	neg
	660		SSPDAVTAYNGYLTS	SVSSPDAVTAYNGYL	SVSSP	-1.33	
	661		SPDAVTAYNGYLTSS	VSSPDAVTAYNGYLT	VSSPD	-1.55	
	662		PDAVTAYNGYLTSSS	SSPDAVTAYNGYLTS	SSPDA	-1.60	
	663		DAVTAYNGYLTSSSK	SPDAVTAYNGYLTSS	SPDAV	-1.54	
	664		AVTAYNGYLTSSSKT	PDAVTAYNGYLTSSS	PDAVT	-1.33	
	674		SSSKTPEEHFIETIS	LTSSSKTPEEHFIET	LTSSS	-1.37	
	675		SSKTPEEHFIETISL	TSSSKTPEEHFIETI	TSSSK	-1.82	
	676		SKTPEEHFIETISLA	SSSKTPEEHFIETIS	SSSKT	-2.11	
	677		KTPEEHFIETISLAG	SSKTPEEHFIETISL	SSKTP	-2.25	
	678		TPEEHFIETISLAGS	SKTPEEHFIETISLA	SKTPE	-2.14	
	679		PEEHFIETISLAGSY	KTPEEHFIETISLAG	KTPEE	-1.86	
	680		EEHFIETISLAGSYK	TPEEHFIETISLAGS	TPEEH	-1.40	
	698		YSGQSTQLGIEFLKR	WSYSGQSTQLGIEFL	WSYSG	-1.30	
	699		SGQSTQLGIEFLKRG	SYSGQSTQLGIEFLK	SYSGQ	-1.47	
	700		GQSTQLGIEFLKRGD	YSGQSTQLGIEFLKR	YSGQS	-1.61	
	701		QSTQLGIEFLKRGDK	SGQSTQLGIEFLKRG	SGQST	-1.61	
	702		STQLGIEFLKRGDKS	GQSTQLGIEFLKRGD	GQSTQ	-1.42	
	749		TIKVFTTVDNINLHT	VRTIKVFTTVDNINL	VRTIK	0.50	neg
	762		HTQVVDMSMTYGQQF	NLHTQVVDMSMTYGQ	NLHTQ	0.16	neg
	792		HNSHEGKTFYVLPND	KPHNSHEGKTFYVLP	KPHNS	-1.53	
	793		NSHEGKTFYVLPNDD	PHNSHEGKTFYVLPN	PHNSH	-1.77	
	794		SHEGKTFYVLPNDDT	HNSHEGKTFYVLPND	HNSHE	-1.76	
	795		HEGKTFYVLPNDDTL	NSHEGKTFYVLPNDD	NSHEG	-1.53	
	826		GRYMSALNHTKKWKY	FLGRYMSALNHTKKW	FLGRY	0.13	neg
	866		QQIELKFNPPALQDA	TLQQIELKFNPPALQ	TLQQI	0.54	neg
	891		NFCALILAYCNKTVG	AANFCALILAYCNKT	AANFC	0.29	neg

	916		YLFQHANLDSCKRVL	MSYLFQHANLDSCKR	MSYLF	0.95	neg
	920		HANLDSCKRVLNVVC	FQHANLDSCKRVLNV	FQHAN	0.52	neg
	939		QQQTTLKGVEAVMYM	CGQQQTTLKGVEAVM	CGQQQ	-1.44	
	940		QQTTLKGVEAVMYMG	GQQQTTLKGVEAVMY	GQQQT	-1.43	
	941		QTTLKGVEAVMYMGT	QQQTTLKGVEAVMYM	QQQTT	-1.42	
	942		TTLKGVEAVMYMGTL	QQTTLKGVEAVMYMG	QQTTL	-1.34	
	989		MSAPPAQYELKHGTF	VMMSAPPAQYELKHG	VMMSA	0.73	neg
	993		PAQYELKHGTFTCAS	APPAQYELKHGTFTC	APPAQ	-1.45	
	994		AQYELKHGTFTCASE	PPAQYELKHGTFTCA	PPAQY	-1.54	
	1016		GHYKHITSKETLYCI	QCGHYKHITSKETLY	QCGHY	0.14	neg
	1038		SSEYKGPITDVFYKE	TKSSEYKGPITDVFY	TKSSE	-1.31	
	1039		SEYKGPITDVFYKEN	KSSEYKGPITDVFYK	KSSEY	-1.49	
	1040		EYKGPITDVFYKENS	SSEYKGPITDVFYKE	SSEYK	-1.60	
	1041		YKGPITDVFYKENSY	SEYKGPITDVFYKEN	SEYKG	-1.80	
	1042		KGPITDVFYKENSYT	EYKGPITDVFYKENS	EYKGP	-1.74	
	1043		GPITDVFYKENSYTT	YKGPITDVFYKENSY	YKGPI	-1.39	
	1048		VFYKENSYTTTIKPV	TDVFYKENSYTTTIK	TDVFY	0.44	neg
	1055		YTTTIKPVTYKLDGV	NSYTTTIKPVTYKLD	NSYTT	-1.39	
	1065		KLDGVVCTEIDPKLD	TYKLDGVVCTEIDPK	TYKLD	0.05	neg
	1100		QPYPNASFDNFKFVC	PNQPYPNASFDNFKF	PNQPY	-1.44	
	1101		PYPNASFDNFKFVCD	NQPYPNASFDNFKFV	NQPYP	-1.66	
	1102		YPNASFDNFKFVCDN	QPYPNASFDNFKFVC	QPYPN	-1.64	
	1103		PNASFDNFKFVCDNI	PYPNASFDNFKFVCD	PYPNA	-1.55	
	1130		KKPASRELKVTFFPD	GYKKPASRELKVTFF	GYKKP	-1.30	
	1131		KPASRELKVTFFPDL	YKKPASRELKVTFFP	YKKPA	-1.65	
	1132		PASRELKVTFFPDLN	KKPASRELKVTFFPD	KKPAS	-1.94	
	1133		ASRELKVTFFPDLNG	KPASRELKVTFFPDL	KPASR	-2.07	
	1134		SRELKVTFFPDLNGD	PASRELKVTFFPDLN	PASRE	-1.85	
	1154		YKHYTPSFKKGAKLL	IDYKHYTPSFKKGAK	IDYKH	0.30	neg
	1173		VWHVNNATNKATYKP	PIVWHVNNATNKATY	PIVWH	1.44	neg
	1180		TNKATYKPNTWCIRC	NATNKATYKPNTWCI	NATNK	-1.38	
	1181		NKATYKPNTWCIRCL	ATNKATYKPNTWCIR	ATNKA	-1.76	
	1182		KATYKPNTWCIRCLW	TNKATYKPNTWCIRC	TNKAT	-1.85	
	1183		ATYKPNTWCIRCLWS	NKATYKPNTWCIRCL	NKATY	-1.82	
	1184		TYKPNTWCIRCLWST	KATYKPNTWCIRCLW	KATYK	-1.73	
	1185		YKPNTWCIRCLWSTK	ATYKPNTWCIRCLWS	ATYKP	-1.55	
	1186		KPNTWCIRCLWSTKP	TYKPNTWCIRCLWST	TYKPN	-1.30	
	1189		TWCIRCLWSTKPVET	PNTWCIRCLWSTKPV	PNTWC	0.48	neg
	1196		WSTKPVETSNSFDVL	CLWSTKPVETSNSFD	CLWST	0.99	neg
	1200		PVETSNSFDVLKSED	TKPVETSNSFDVLKS	TKPVE	-1.42	
	1201		VETSNSFDVLKSEDA	KPVETSNSFDVLKSE	KPVET	-1.65	
	1202		ETSNSFDVLKSEDAQ	PVETSNSFDVLKSED	PVETS	-1.74	
	1203		TSNSFDVLKSEDAQG	VETSNSFDVLKSEDA	VETSN	-1.64	
	1204		SNSFDVLKSEDAQGM	ETSNSFDVLKSEDAQ	ETSNS	-1.35	
	1213		EDAQGMDNLACEDLK	KSEDAQGMDNLACED	KSEDA	-1.39	
	1214		DAQGMDNLACEDLKP	SEDAQGMDNLACEDL	SEDAQ	-1.72	
	1215		AQGMDNLACEDLKPV	EDAQGMDNLACEDLK	EDAQG	-1.77	
	1216		QGMDNLACEDLKPVS	DAQGMDNLACEDLKP	DAQGM	-1.61	
	1243		VLECNVKTTEVVGDI	KDVLECNVKTTEVVG	KDVLE	0.24	neg
	1259		LKPANNSLKITEEVG	IILKPANNSLKITEE	IILKP	0.19	neg
	1346		VCTNYMPYFFTLLLQ	NRVCTNYMPYFFTLL	NRVCT	0.46	neg
	1368		TNSRIKASMPTTIAK	RSTNSRIKASMPTTI	RSTNS	-1.36	
	1369		NSRIKASMPTTIAKN	STNSRIKASMPTTIA	STNSR	-1.31	
	1393		LEASFNYLKSPNFSK	FCLEASFNYLKSPNF	FCLEA	0.62	neg
	1396		SFNYLKSPNFSKLIN	EASFNYLKSPNFSKL	EASFN	0.49	neg
	1421		CLGSLIYSTAALGVL	SVCLGSLIYSTAALG	SVCLG	1.49	neg
	1427		YSTAALGVLMSNLGM	LIYSTAALGVLMSNL	LIYST	0.18	neg
	1434		VLMSNLGMPSYCTGY	LGVLMSNLGMPSYCT	LGVLM	0.73	neg
	1435		LMSNLGMPSYCTGYR	GVLMSNLGMPSYCTG	GVLMS	0.87	neg
	1460		IATYCTGSIPCSVCL	VTIATYCTGSIPCSV	VTIAT	0.00	neg
	1474		LSGLDSLDTYPSLET	VCLSGLDSLDTYPSL	VCLSG	0.52	neg
	1492		TISSFKWDLTAFGLV	QITISSFKWDLTAFG	QITIS	0.37	neg
	1505		LVAEWFLAYILFTRF	FGLVAEWFLAYILFT	FGLVA	1.12	neg
	1508		EWFLAYILFTRFFYV	VAEWFLAYILFTRFF	VAEWF	1.75	neg
	1511		LAYILFTRFFYVLGL	WFLAYILFTRFFYVL	WFLAY	2.46	neg
	1529		MQLFFSYFAVHFISN	AIMQLFFSYFAVHFI	AIMQL	1.80	neg

	1530		QLFFSYFAVHFISNS	IMQLFFSYFAVHFIS	IMQLF	1.95	neg
	1571		SFYYVWKSYVHVVDG	FASFYYVWKSYVHVV	FASFY	2.09	neg
	1595		YKRNRATRVECTTIV	MCYKRNRATRVECTT	MCYKR	0.87	neg
	1606		TTIVNGVRRSFYVYA	ECTTIVNGVRRSFYV	ECTTI	0.14	neg
	1630		HNWNCVNCDTFCAGS	KLHNWNCVNCDTFCA	KLHNW	1.05	neg
	1638		DTFCAGSTFISDEVA	NCDTFCAGSTFISDE	NCDTF	0.43	neg
	1663		INPTDQSSYIVDSVT	RPINPTDQSSYIVDS	RPINP	-1.40	
	1664		NPTDQSSYIVDSVTV	PINPTDQSSYIVDSV	PINPT	-1.77	
	1665		PTDQSSYIVDSVTVK	INPTDQSSYIVDSVT	INPTD	-1.87	
	1666		TDQSSYIVDSVTVKN	NPTDQSSYIVDSVTV	NPTDQ	-1.76	
	1667		DQSSYIVDSVTVKNG	PTDQSSYIVDSVTVK	PTDQS	-1.62	
	1668		QSSYIVDSVTVKNGS	TDQSSYIVDSVTVKN	TDQSS	-1.31	
	1691		GQKTYERHSLSHFVN	KAGQKTYERHSLSHF	KAGQK	-1.35	
	1692		QKTYERHSLSHFVNL	AGQKTYERHSLSHFV	AGQKT	-1.65	
	1693		KTYERHSLSHFVNLD	GQKTYERHSLSHFVN	GQKTY	-1.81	
	1694		TYERHSLSHFVNLDN	QKTYERHSLSHFVNL	QKTYE	-1.70	
	1695		YERHSLSHFVNLDNL	KTYERHSLSHFVNLD	KTYER	-1.38	
	1699		SLSHFVNLDNLRANN	RHSLSHFVNLDNLRA	RHSLS	0.13	neg
	1700		LSHFVNLDNLRANNT	HSLSHFVNLDNLRAN	HSLSH	0.38	neg
	1705		NLDNLRANNTKGSLP	FVNLDNLRANNTKGS	FVNLD	0.68	neg
	1707		DNLRANNTKGSLPIN	NLDNLRANNTKGSLP	NLDNL	0.34	neg
	1711		ANNTKGSLPINVIVF	LRANNTKGSLPINVI	LRANN	-1.38	
	1712		NNTKGSLPINVIVFD	RANNTKGSLPINVIV	RANNT	-1.64	
	1713		NTKGSLPINVIVFDG	ANNTKGSLPINVIVF	ANNTK	-1.73	
	1714		TKGSLPINVIVFDGK	NNTKGSLPINVIVFD	NNTKG	-1.64	
	1715		KGSLPINVIVFDGKS	NTKGSLPINVIVFDG	NTKGS	-1.67	
	1716		GSLPINVIVFDGKSK	TKGSLPINVIVFDGK	TKGSL	-1.55	
	1730		KCEESSAKSASVYYS	KSKCEESSAKSASVY	KSKCE	-1.65	
	1731		CEESSAKSASVYYSQ	SKCEESSAKSASVYY	SKCEE	-1.97	
	1732		EESSAKSASVYYSQL	KCEESSAKSASVYYS	KCEES	-2.24	
	1733		ESSAKSASVYYSQLM	CEESSAKSASVYYSQ	CEESS	-2.28	
	1734		SSAKSASVYYSQLMC	EESSAKSASVYYSQL	EESSA	-2.19	
	1735		SAKSASVYYSQLMCQ	ESSAKSASVYYSQLM	ESSAK	-1.96	
	1736		AKSASVYYSQLMCQP	SSAKSASVYYSQLMC	SSAKS	-1.75	
	1737		KSASVYYSQLMCQPI	SAKSASVYYSQLMCQ	SAKSA	-1.61	
	1738		SASVYYSQLMCQPIL	AKSASVYYSQLMCQP	AKSAS	-1.54	
	1739		ASVYYSQLMCQPILL	KSASVYYSQLMCQPI	KSASV	-1.36	
	1743		YSQLMCQPILLLDQA	VYYSQLMCQPILLLD	VYYSQ	0.45	neg
	1754		LDQALVSDVGDSAEV	LLLDQALVSDVGDSA	LLLDQ	1.45	neg
	1764		DSAEVAVKMFDAYVN	VGDSAEVAVKMFDAY	VGDSA	-1.40	
	1765		SAEVAVKMFDAYVNT	GDSAEVAVKMFDAYV	GDSAE	-1.31	
	1792		KTLVATAEAELAKNV	KLKTLVATAEAELAK	KLKTL	0.24	neg
	1917		KNNLPFKLTCATTRQ	AKKNNLPFKLTCATT	AKKNN	-1.34	
1798174255	13	nucleocapsid	PRITFGGPSDSTGSN	NAPRITFGGPSDSTG	NAPRI	0.09	neg
	21		SDSTGSNQNGERSGA	GPSDSTGSNQNGERS	GPSDS	-1.33	
	22		DSTGSNQNGERSGAR	PSDSTGSNQNGERSG	PSDST	-1.56	
	23		STGSNQNGERSGARS	SDSTGSNQNGERSGA	SDSTG	-1.69	
	24		TGSNQNGERSGARSK	DSTGSNQNGERSGAR	DSTGS	-1.67	
	25		GSNQNGERSGARSKQ	STGSNQNGERSGARS	STGSN	-1.67	
	26		SNQNGERSGARSKQR	TGSNQNGERSGARSK	TGSNQ	-1.51	
	27		NQNGERSGARSKQRR	GSNQNGERSGARSKQ	GSNQN	-1.43	
	28		QNGERSGARSKQRRP	SNQNGERSGARSKQR	SNQNG	-1.36	
	29		NGERSGARSKQRRPQ	NQNGERSGARSKQRR	NQNGE	-1.33	
	30		GERSGARSKQRRPQG	QNGERSGARSKQRRP	QNGER	-1.34	
	31		ERSGARSKQRRPQGL	NGERSGARSKQRRPQ	NGERS	-1.31	
	56		LTQHGKEDLKFPRGQ	TALTQHGKEDLKFPR	TALTQ	1.02	neg
	58		QHGKEDLKFPRGQGV	LTQHGKEDLKFPRGQ	LTQHG	0.09	neg
	87		YRRATRRIRGGDGKM	GYYRRATRRIRGGDG	GYYRR	0.73	neg
	131		IWVATEGALNTPKDH	GIIWVATEGALNTPK	GIIWV	1.59	neg
	171		FYAEGSRGGSQASSR	KGFYAEGSRGGSQAS	KGFYA	0.47	neg
	179		GSQASSRSSSRSRNS	RGGSQASSRSSSRSR	RGGSQ	-1.44	
	180		SQASSRSSSRSRNSS	GGSQASSRSSSRSRN	GGSQA	-1.41	
	181		QASSRSSSRSRNSSR	GSQASSRSSSRSRNS	GSQAS	-1.37	
	182		ASSRSSSRSRNSSRN	SQASSRSSSRSRNSS	SQASS	-1.31	
	183		SSRSSSRSRNSSRNS	QASSRSSSRSRNSSR	QASSR	-1.34	
	184		SRSSSRSRNSSRNST	ASSRSSSRSRNSSRN	ASSRS	-1.42	

	185		RSSSRSRNSSRNSTP	SSRSSSRSRNSSRNS	SSRSS	-1.44	
	186		SSSRSRNSSRNSTPG	SRSSSRSRNSSRNST	SRSSS	-1.46	
	187		SSRSRNSSRNSTPGS	RSSSRSRNSSRNSTP	RSSSR	-1.44	
	188		SRSRNSSRNSTPGSS	SSSRSRNSSRNSTPG	SSSRS	-1.37	
	189		RSRNSSRNSTPGSSR	SSRSRNSSRNSTPGS	SSRSR	-1.36	
	190		SRNSSRNSTPGSSRG	SRSRNSSRNSTPGSS	SRSRN	-1.40	
	195		RNSTPGSSRGTSPAR	SSRNSTPGSSRGTSP	SSRNS	-1.41	
	196		NSTPGSSRGTSPARM	SRNSTPGSSRGTSPA	SRNST	-1.58	
	197		STPGSSRGTSPARMA	RNSTPGSSRGTSPAR	RNSTP	-1.76	
	198		TPGSSRGTSPARMAG	NSTPGSSRGTSPARM	NSTPG	-1.87	
	199		PGSSRGTSPARMAGN	STPGSSRGTSPARMA	STPGS	-1.79	
	200		GSSRGTSPARMAGNG	TPGSSRGTSPARMAG	TPGSS	-1.78	
	201		SSRGTSPARMAGNGG	PGSSRGTSPARMAGN	PGSSR	-1.77	
	202		SRGTSPARMAGNGGD	GSSRGTSPARMAGNG	GSSRG	-1.55	
	203		RGTSPARMAGNGGDA	SSRGTSPARMAGNGG	SSRGT	-1.40	
	204		GTSPARMAGNGGDAA	SRGTSPARMAGNGGD	SRGTS	-1.33	
	223		LLDRLNQLESKMSGK	LLLLDRLNQLESKMS	LLLLD	2.53	neg
	226		RLNQLESKMSGKGQQ	LDRLNQLESKMSGKG	LDRLN	1.54	neg
	270		VTQAFGRRGPEQTQG	YNVTQAFGRRGPEQT	YNVTQ	0.38	neg
	293		RQGTDYKHWPQIAQF	LIRQGTDYKHWPQIA	LIRQG	0.17	neg
	321		GMEVTPSGTWLTYTG	RIGMEVTPSGTWLTY	RIGME	1.07	neg
	326		PSGTWLTYTGAIKLD	VTPSGTWLTYTGAIK	VTPSG	0.01	neg
	329		TWLTYTGAIKLDDKD	SGTWLTYTGAIKLDD	SGTWL	0.55	neg
	375		KADETQALPQRQKKQ	KKKADETQALPQRQK	KKKAD	-1.37	
	376		ADETQALPQRQKKQQ	KKADETQALPQRQKK	KKADE	-1.46	
	377		DETQALPQRQKKQQT	KADETQALPQRQKKQ	KADET	-1.31	
	401		DDFSKQLQQSMSSAD	DLDDFSKQLQQSMSS	DLDDF	0.31	neg

Appendix 4

SARS-CoV-2 RefSeq

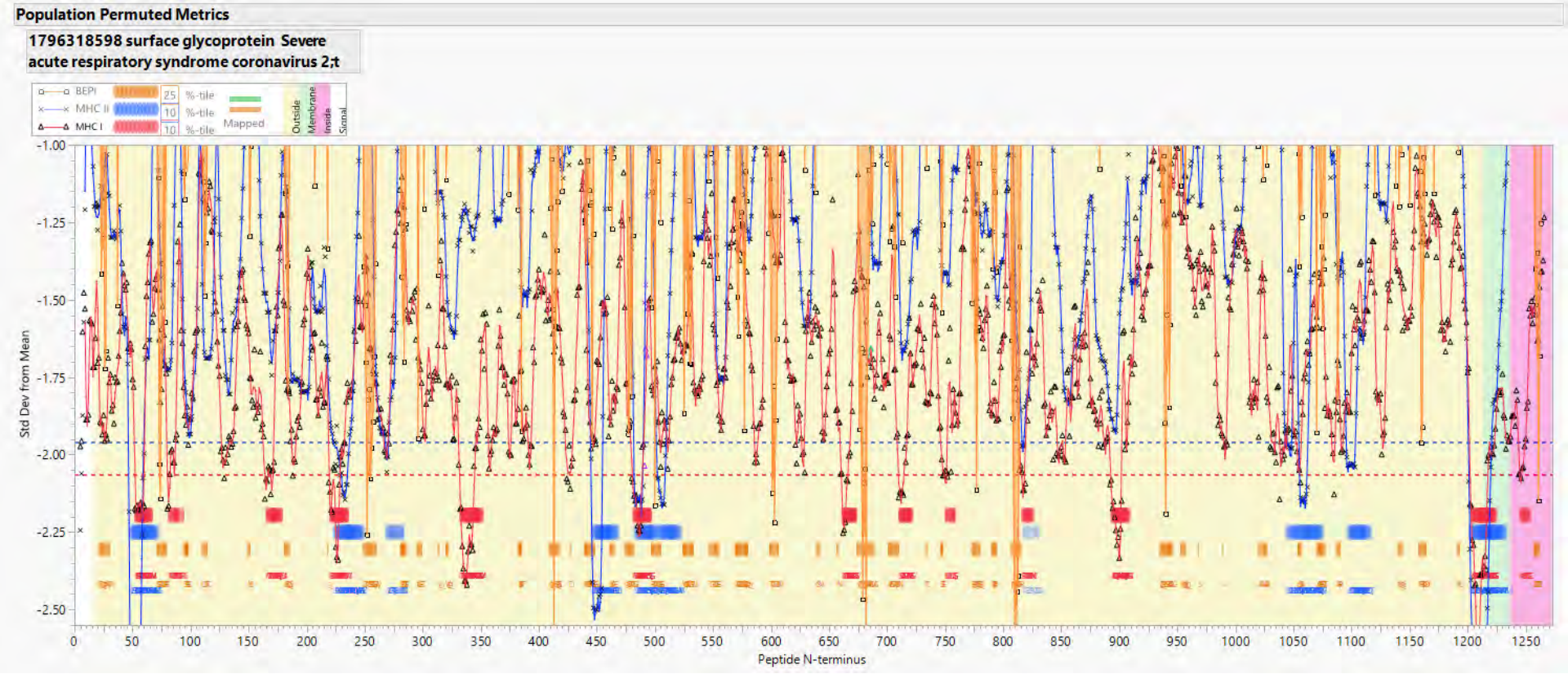
Structural and major nonstructural proteins

The following graphics are provided as an overview and orientation. For larger proteins some detail is hard to distinguish with the space available here. When viewed in the JMP source program the graphics are interactive. The Excel files exported from JMP and corresponding to each protein are provided in the Addendum.
Legends are provided for these and all similar files in Appendix 7

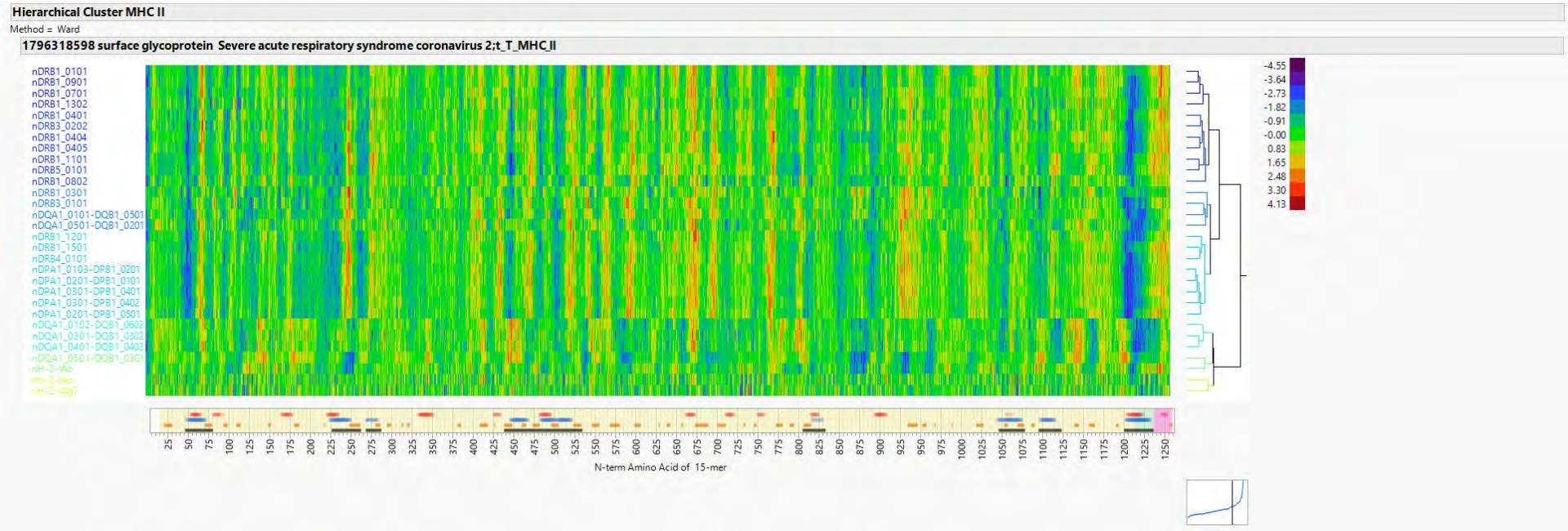
Surface glycoprotein “spike”

1796318598

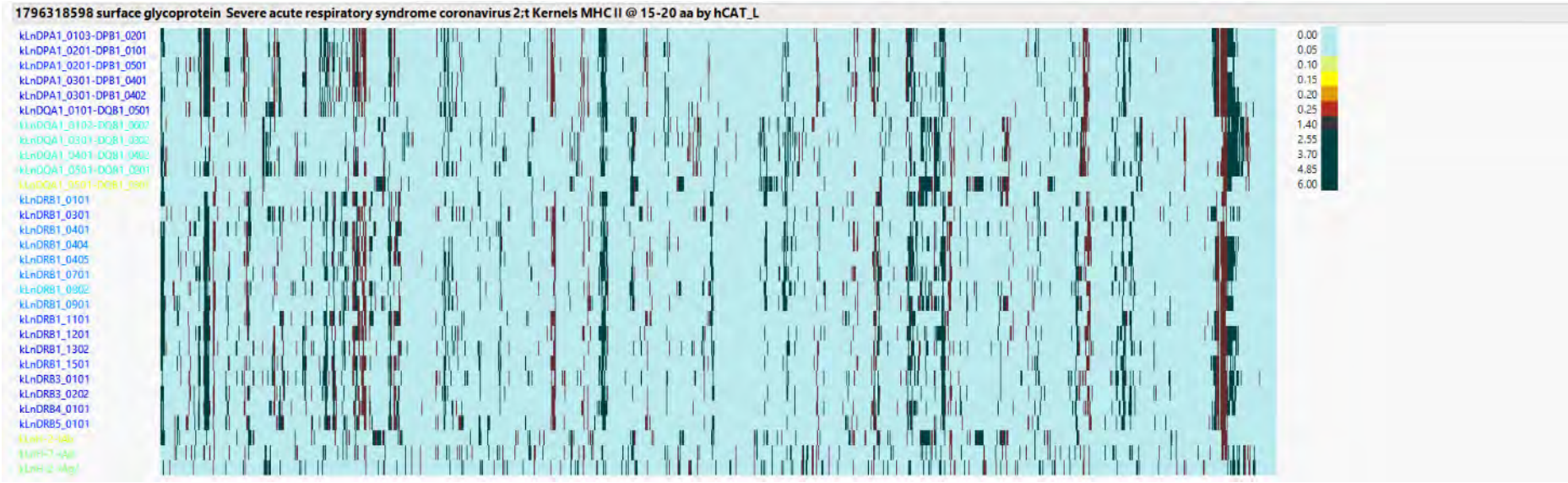
Surface glycoprotein

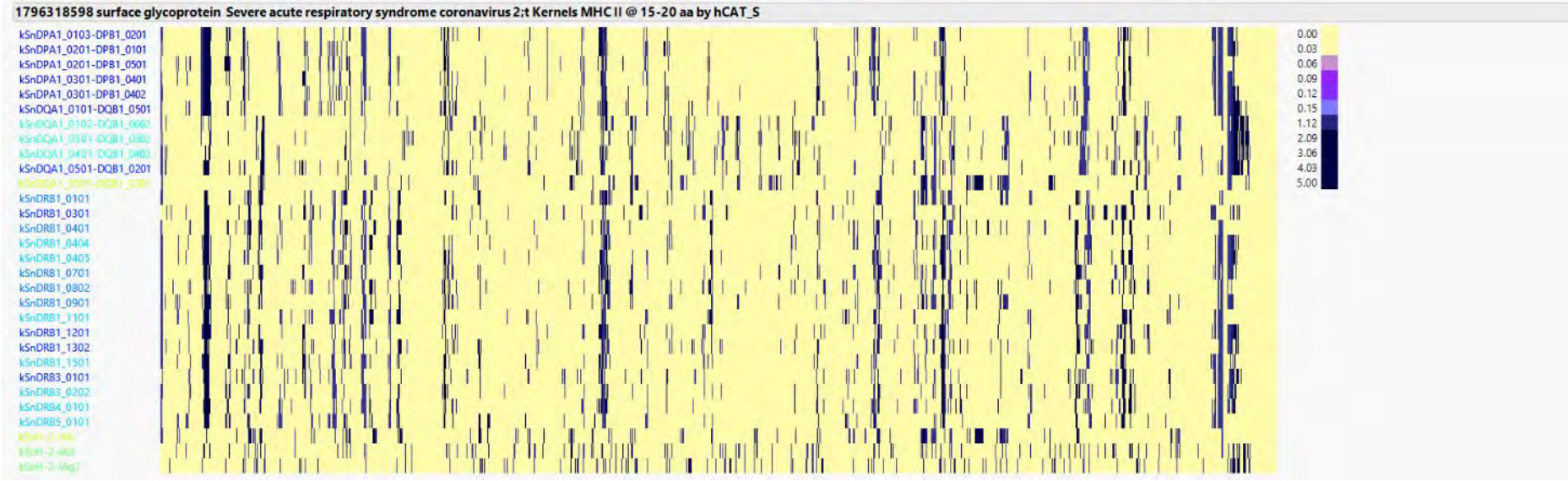


Surface glycoprotein

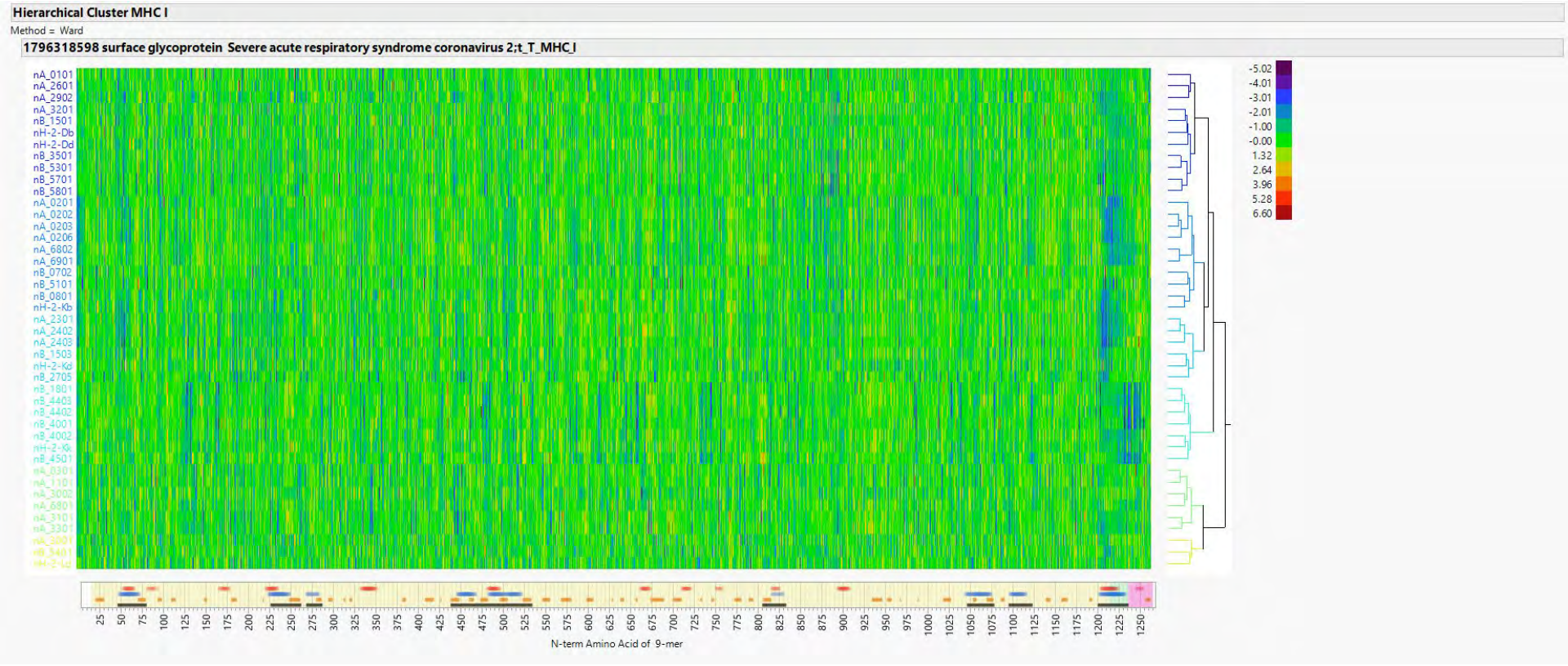


Surface glycoprotein

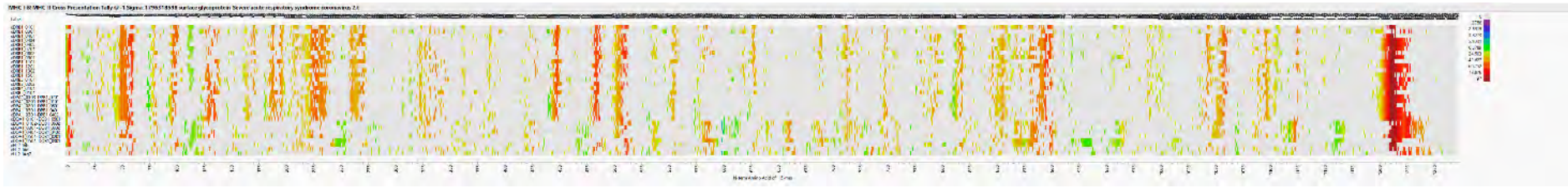


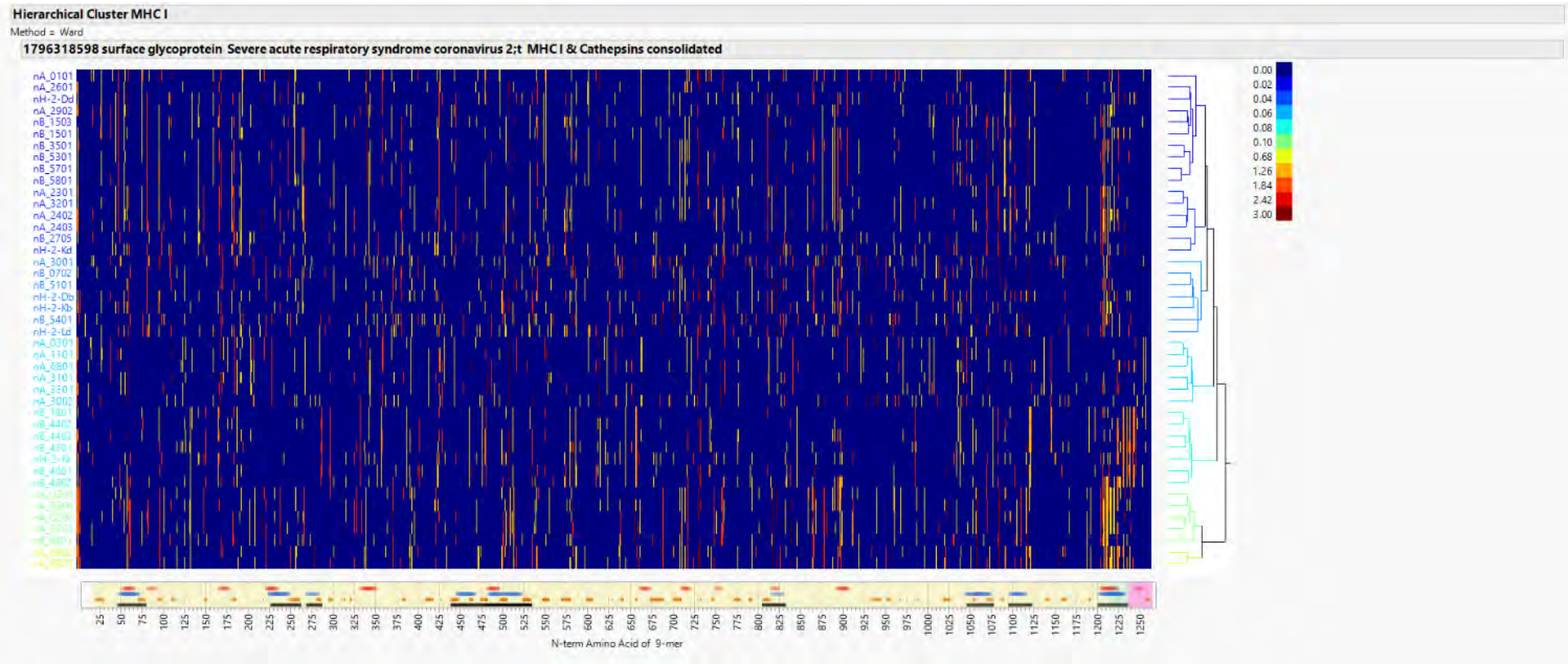


Surface glycoprotein

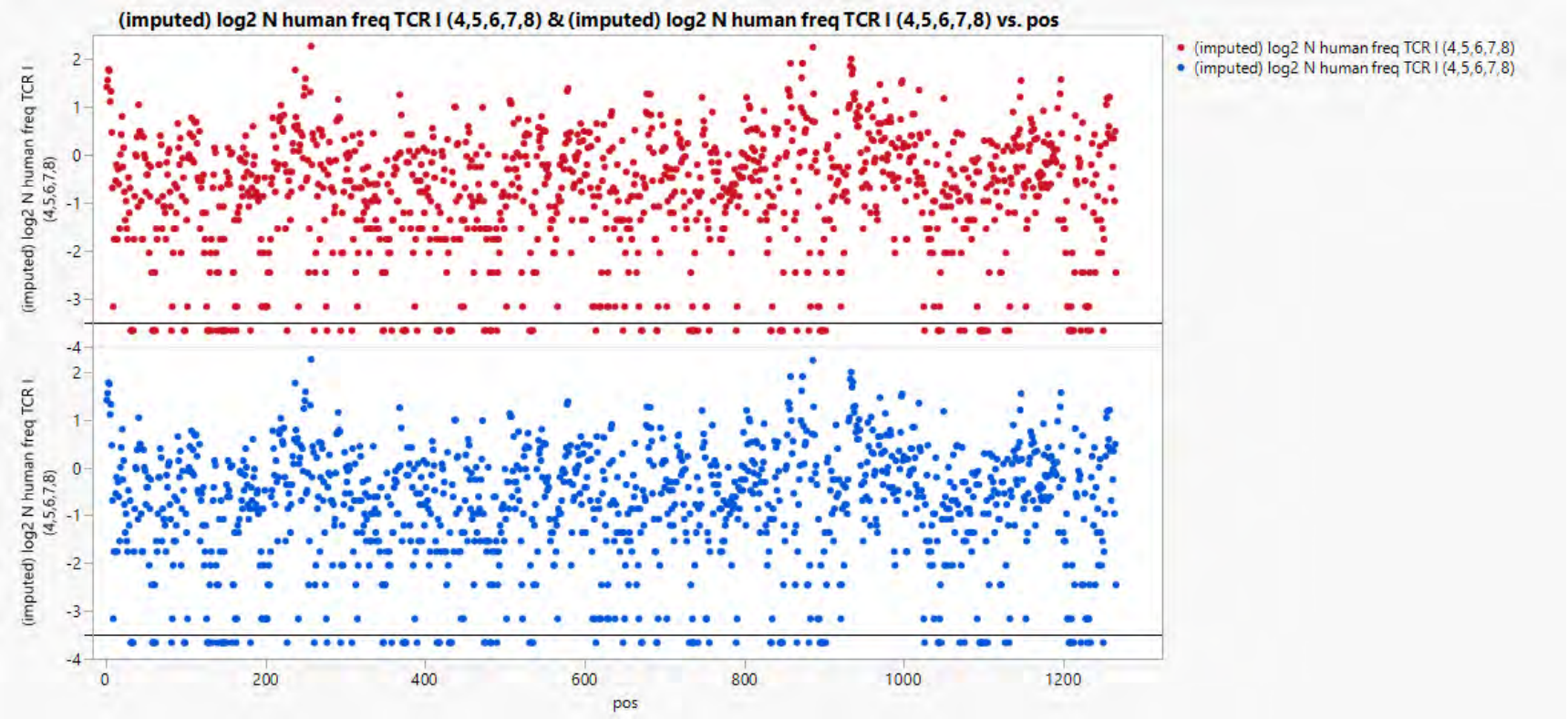


Surface glycoprotein





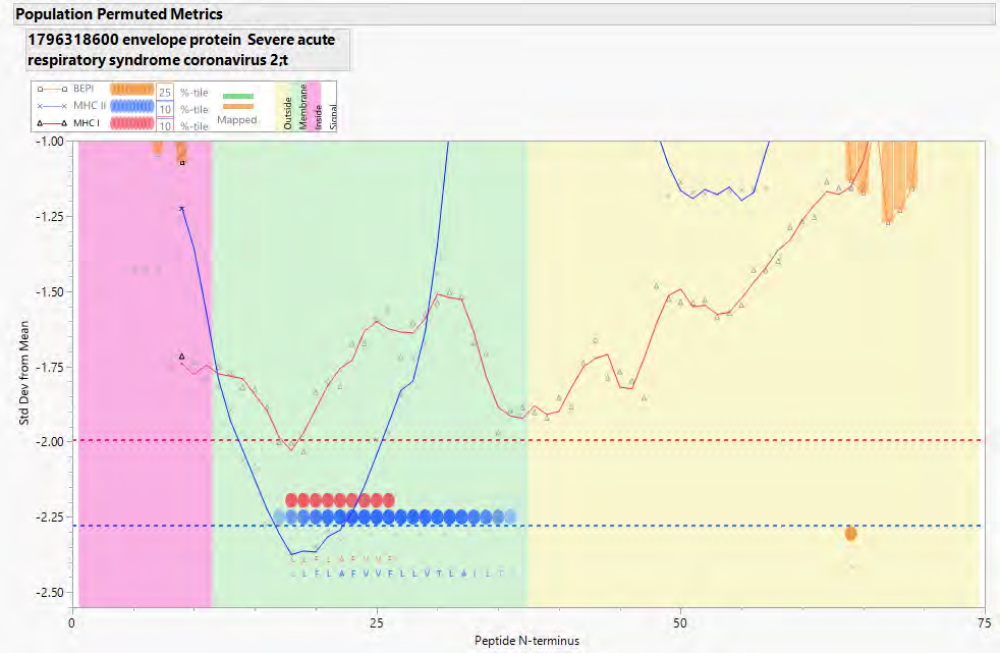
Surface glycoprotein



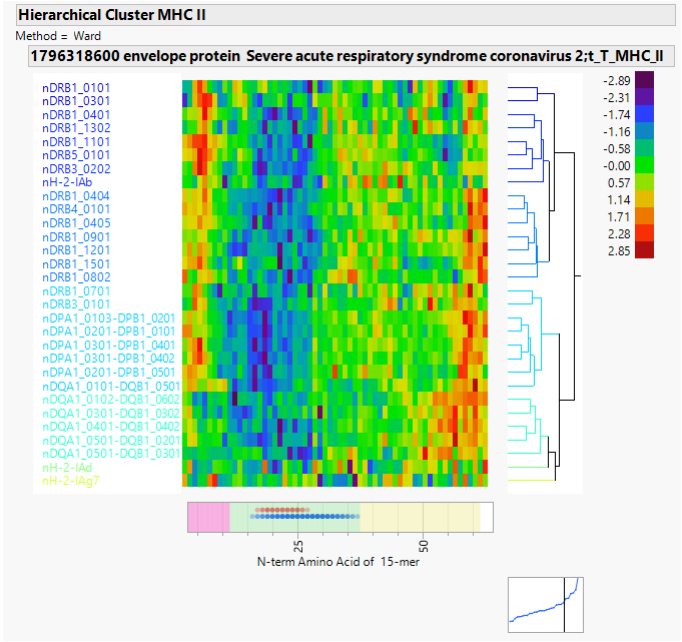
Envelope protein

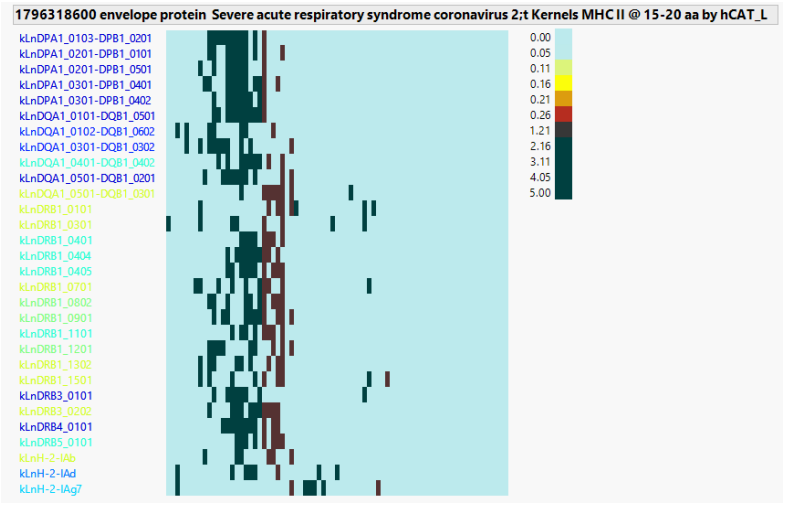
1796318600

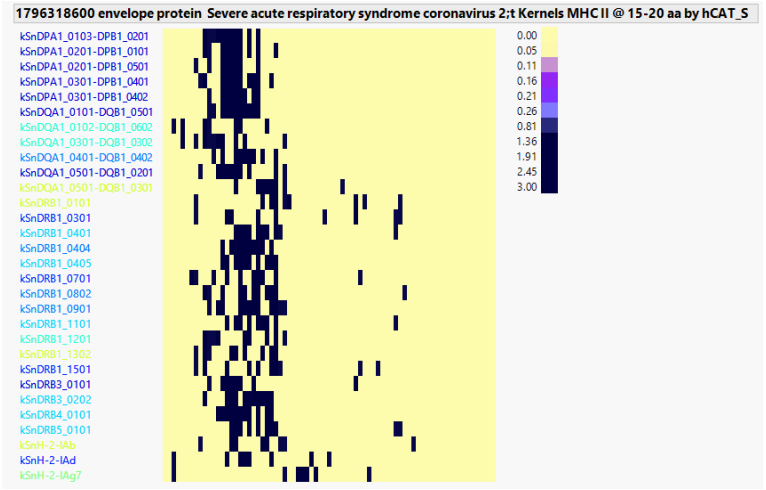
Envelope protein



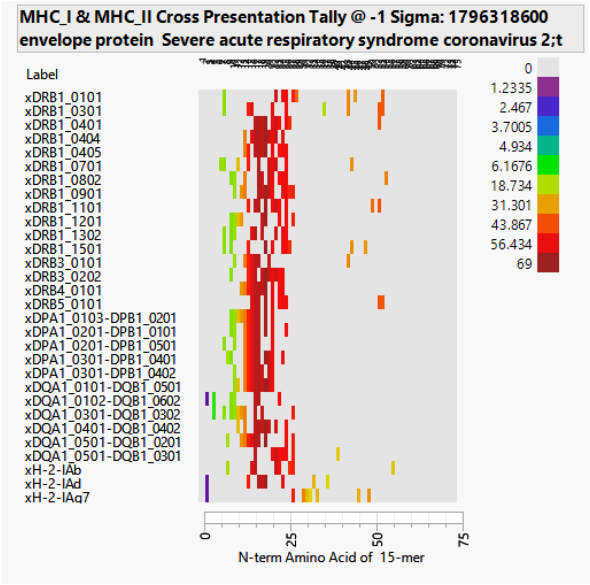
Envelope protein

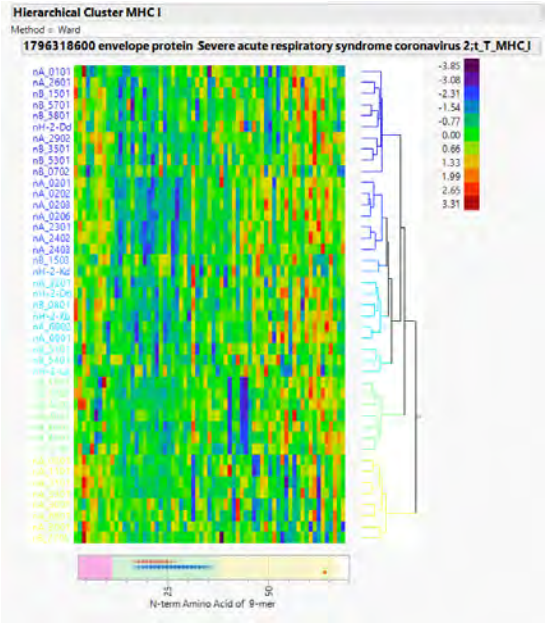


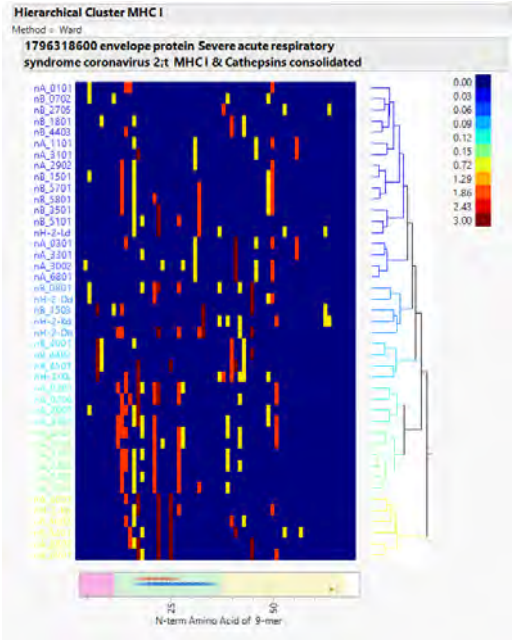


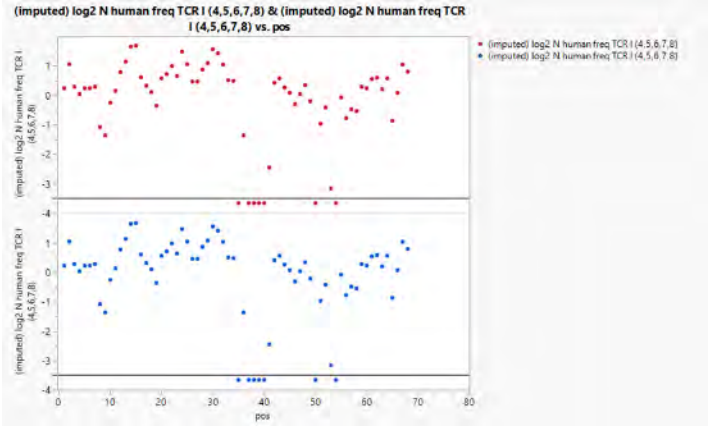


Envelope protein





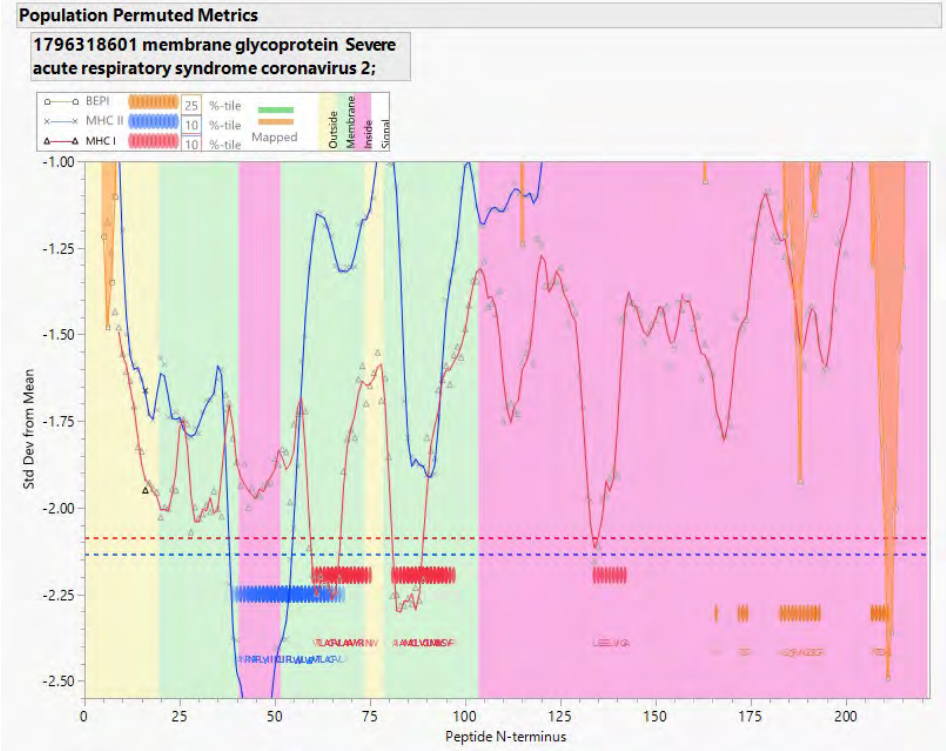




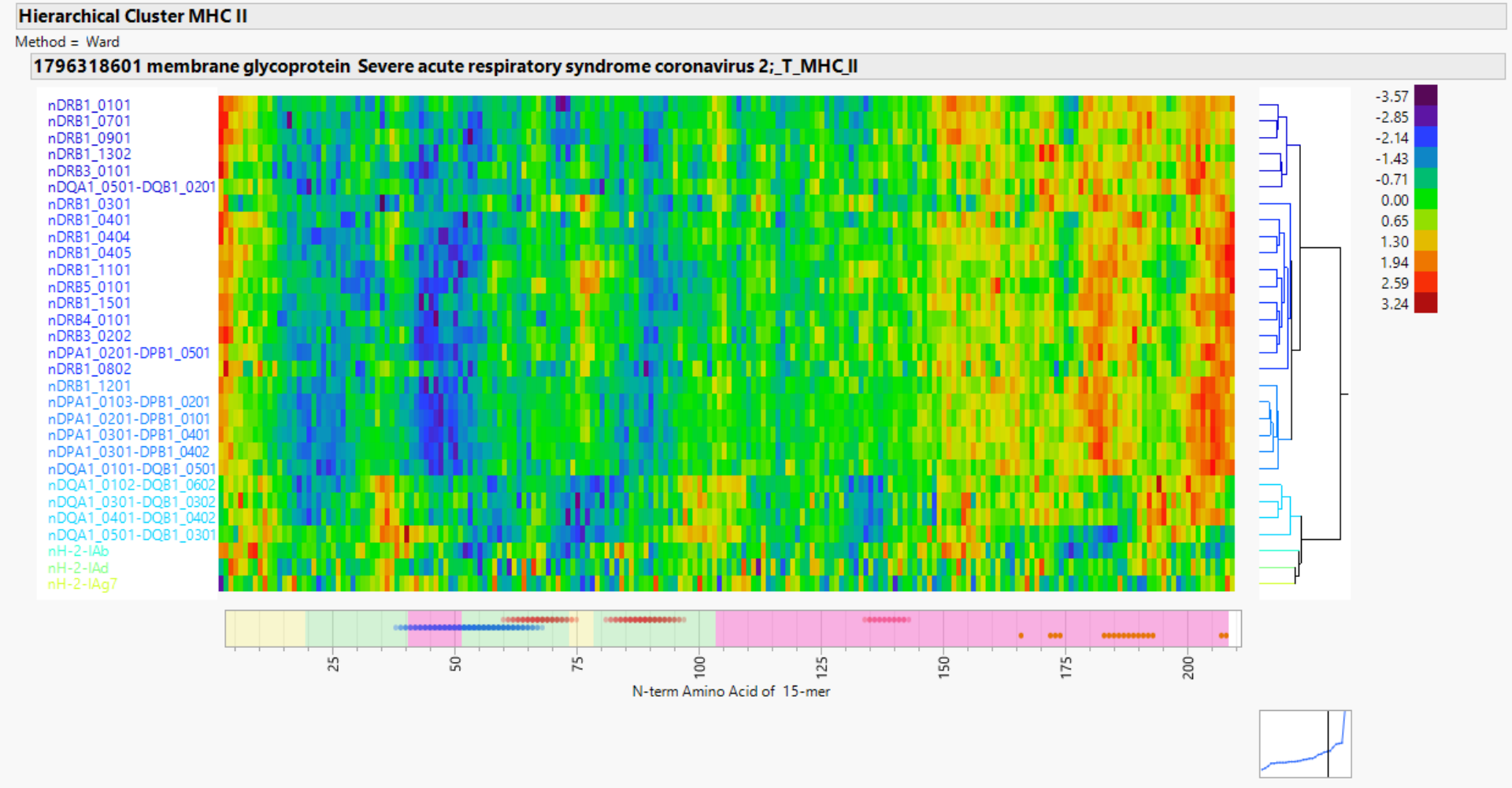
Membrane glycoprotein

1796318601

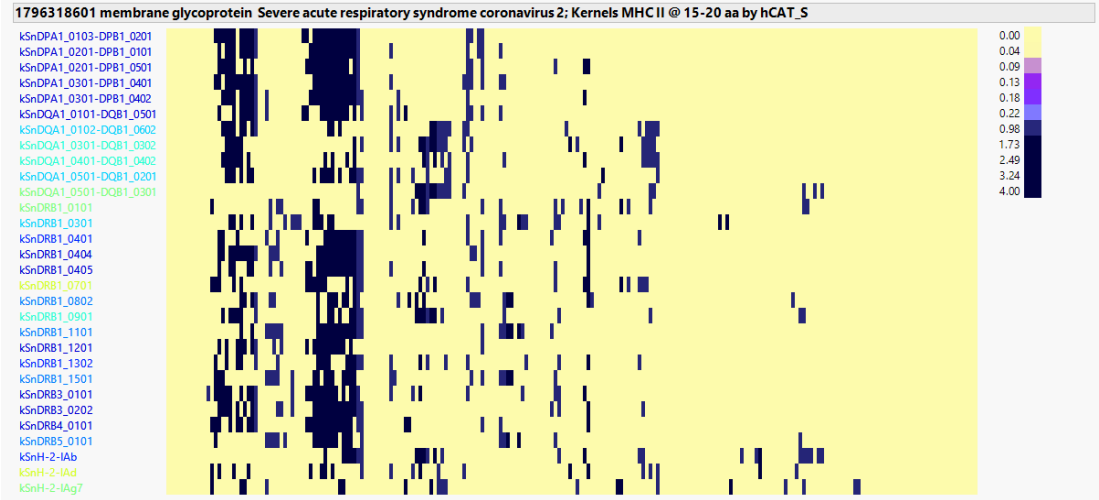
Membrane glycoprotein

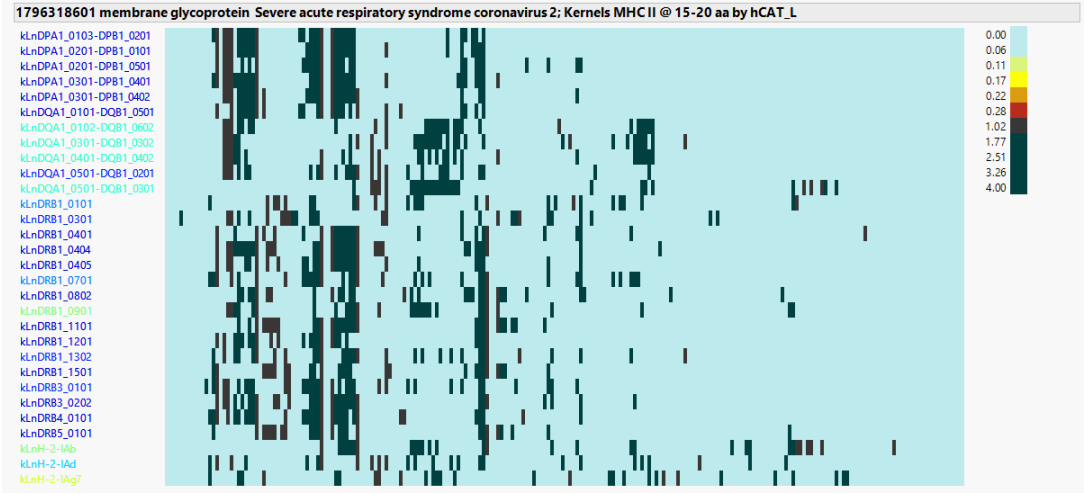


Membrane glycoprotein

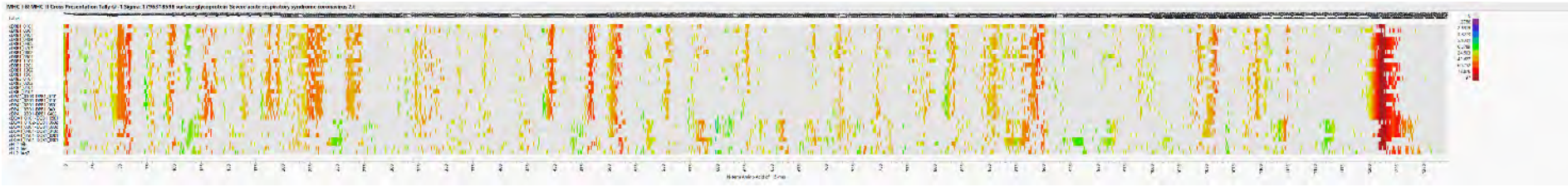


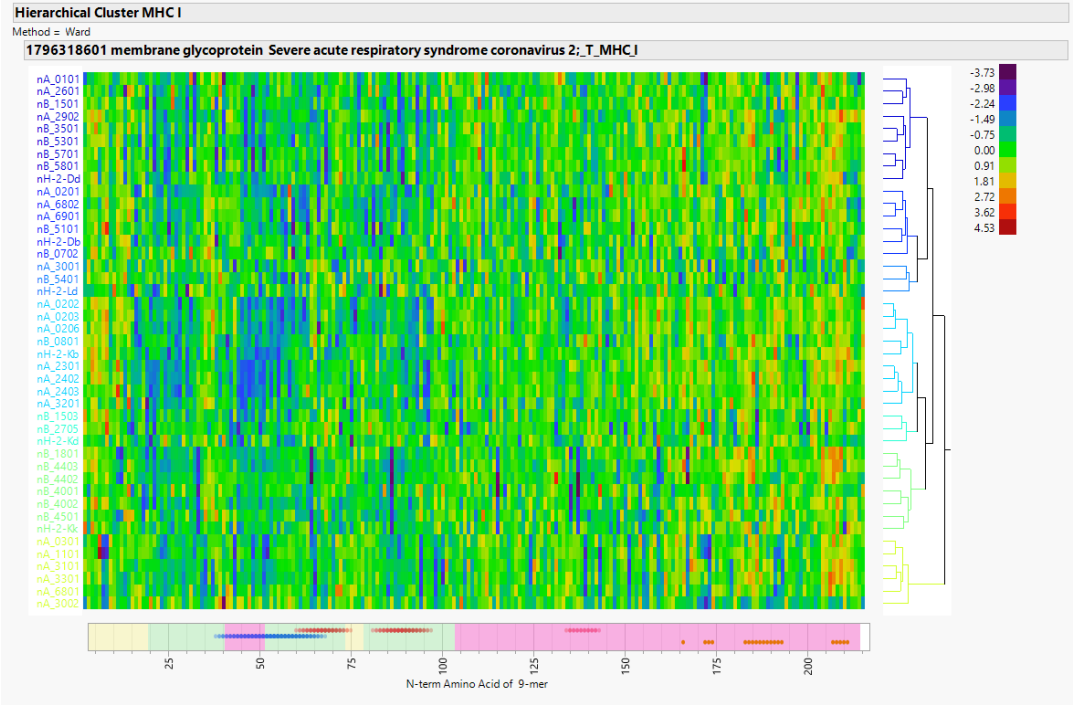
Membrane glycoprotein



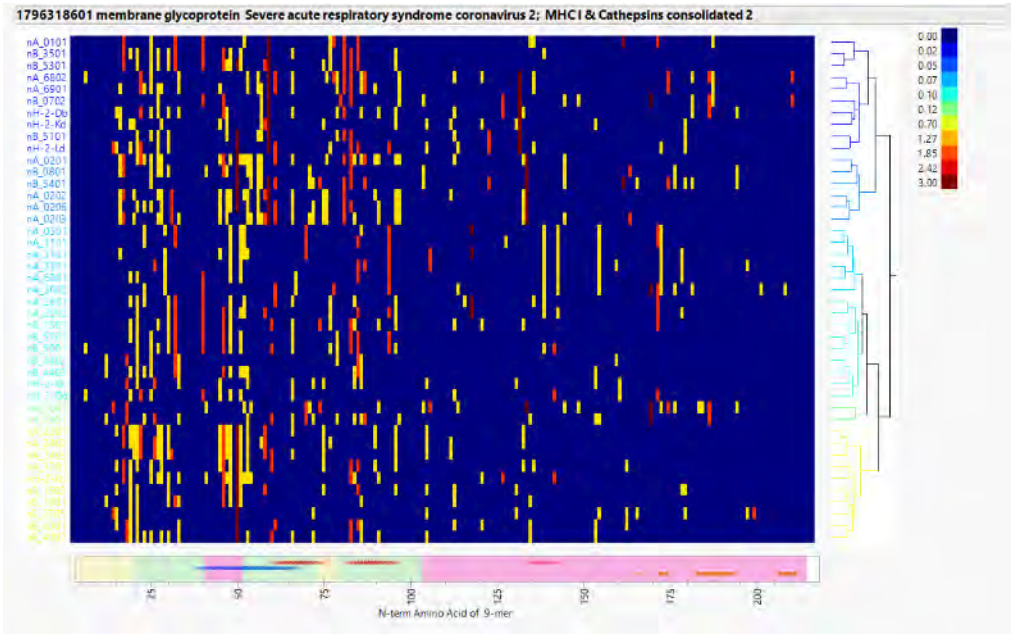


Membrane glycoprotein

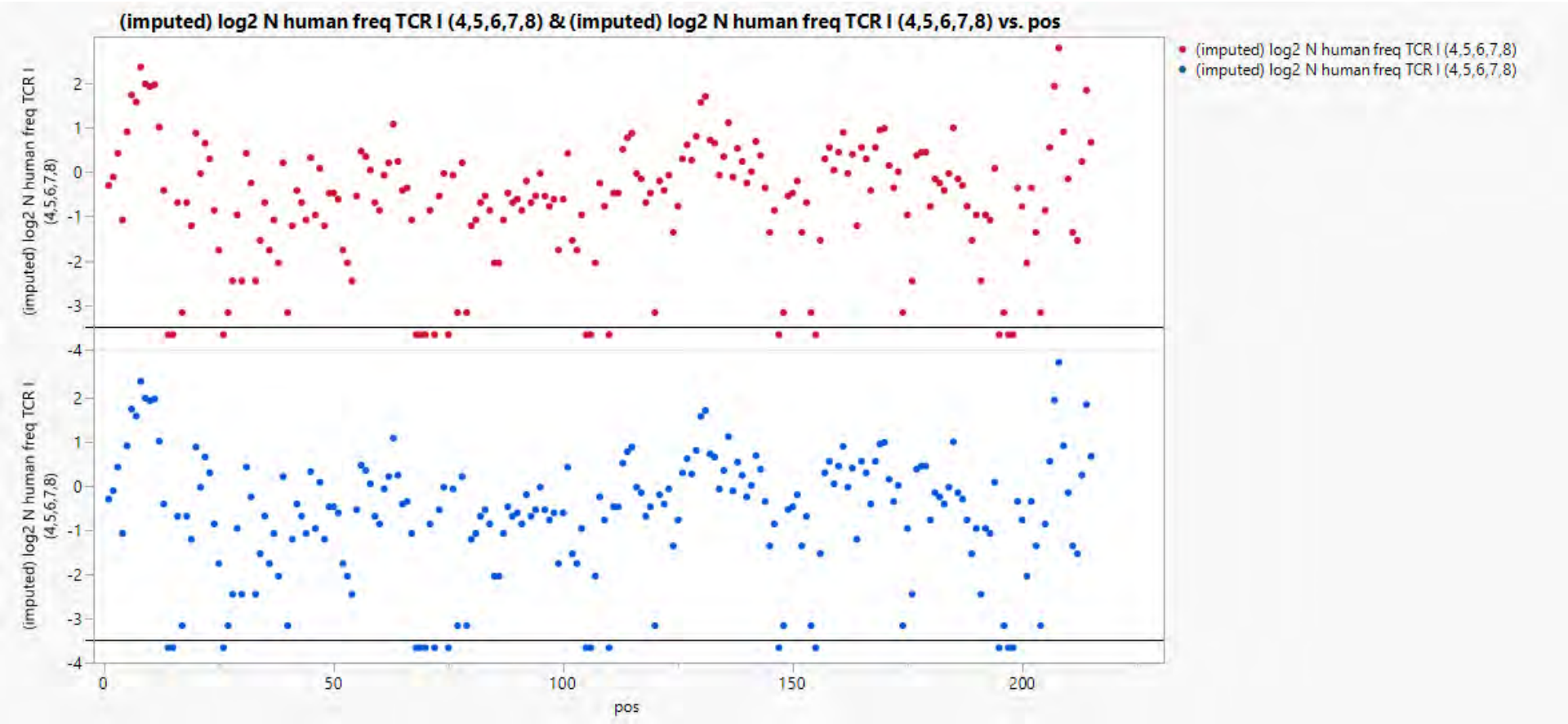




Membrane glycoprotein



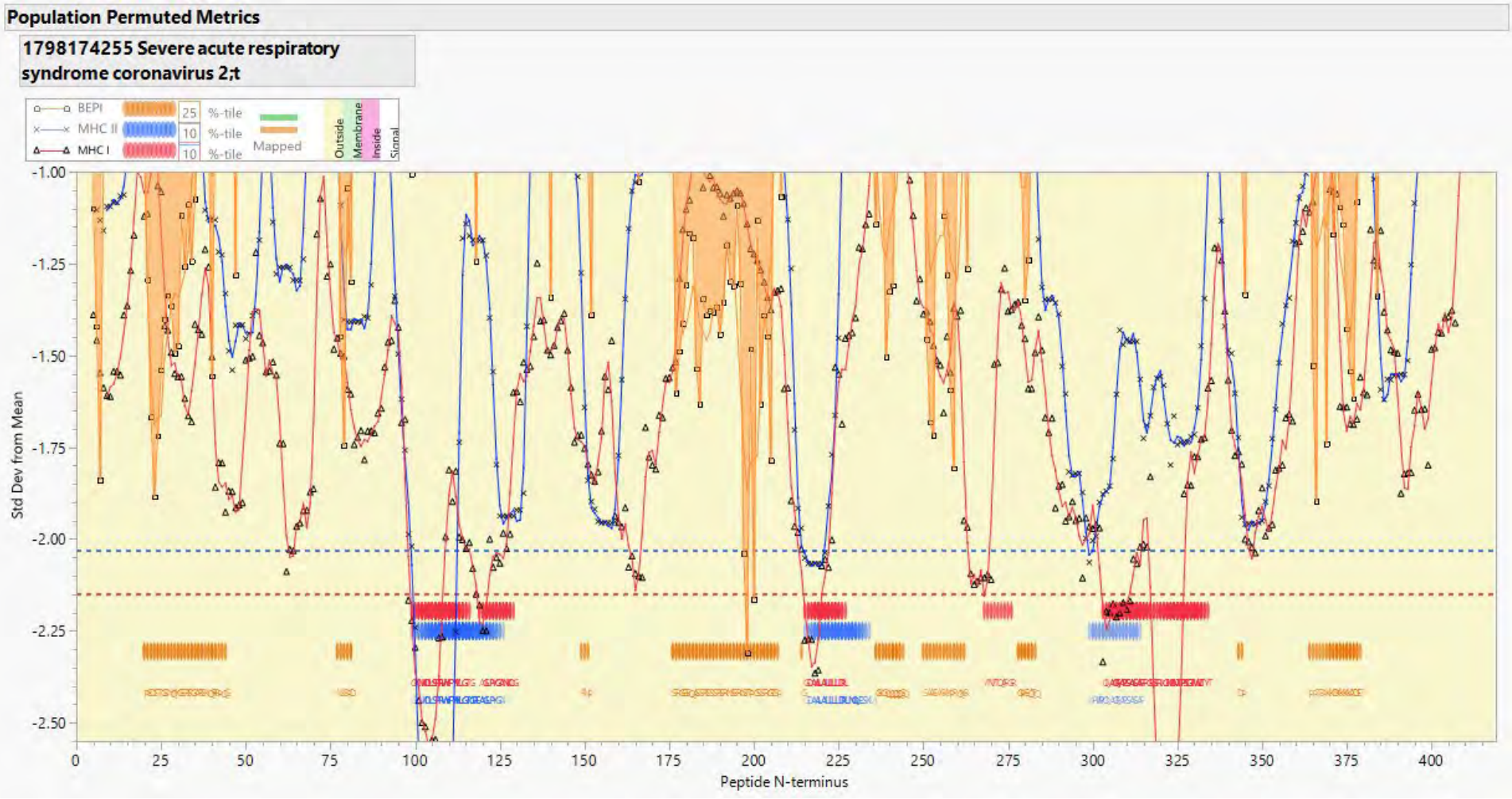
Membrane glycoprotein



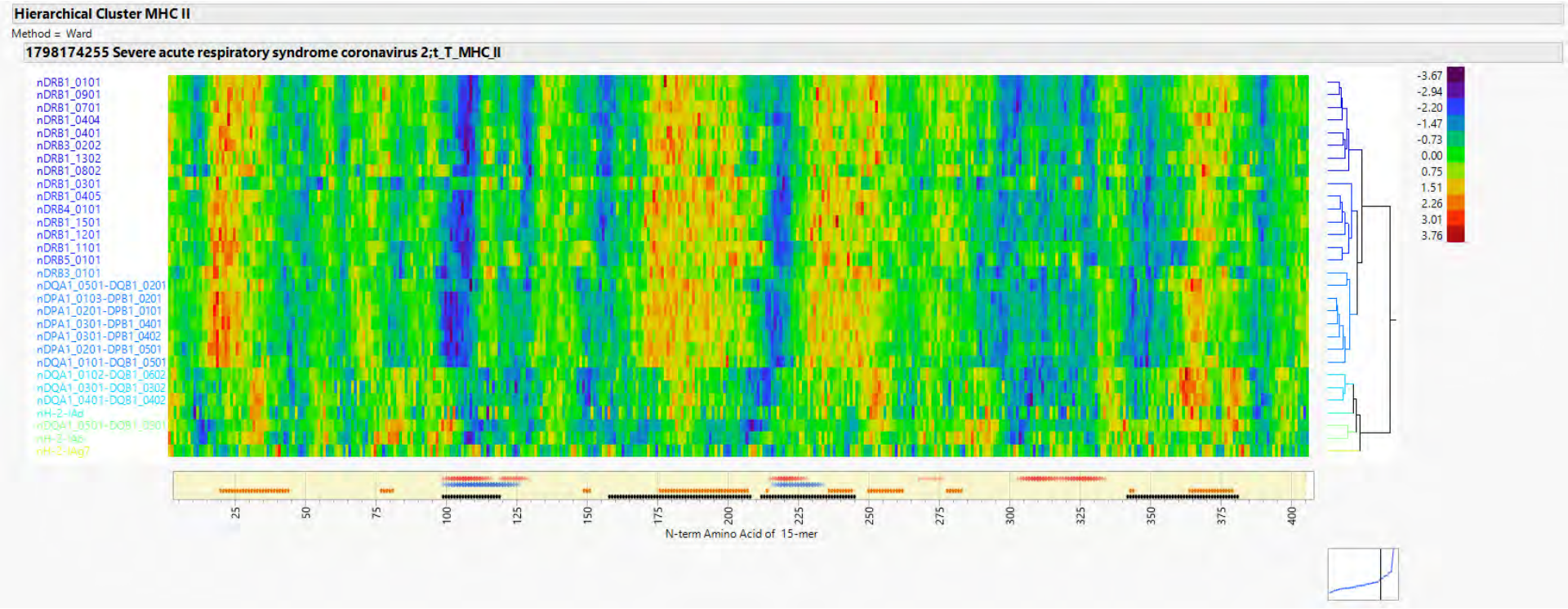
Nucleocapsid

1798174255

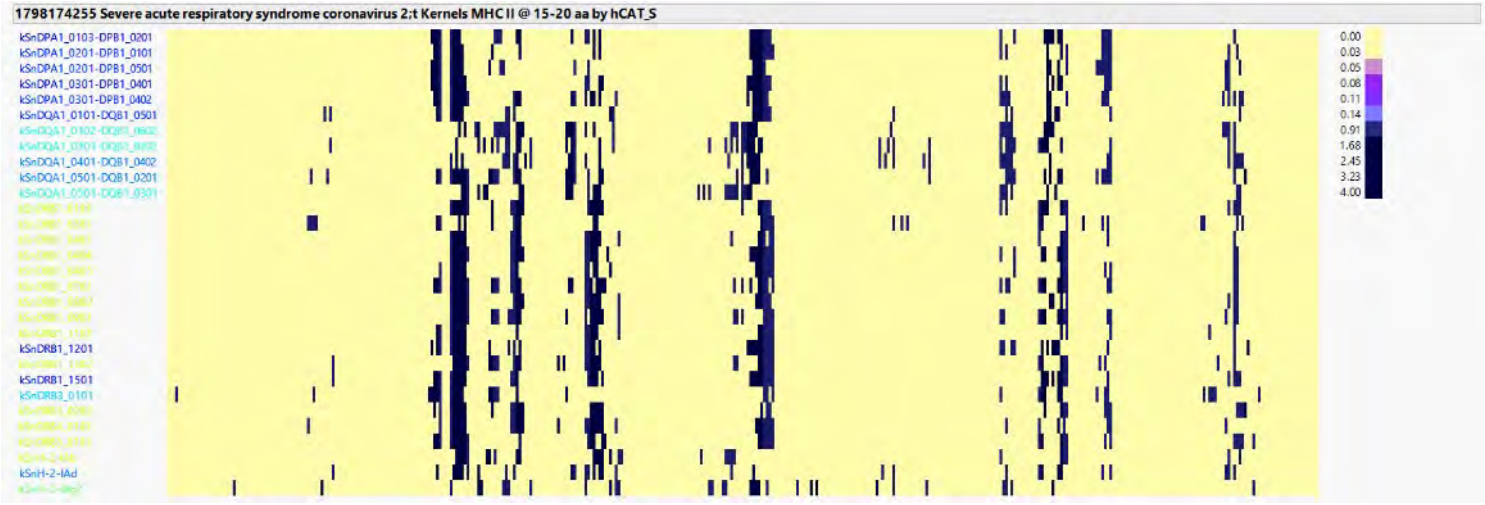
Nucleocapsid



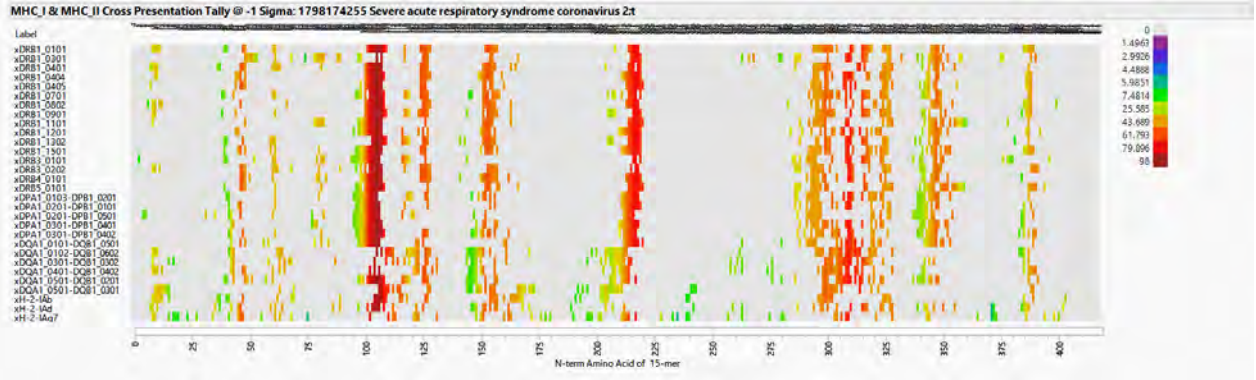
Nucleocapsid

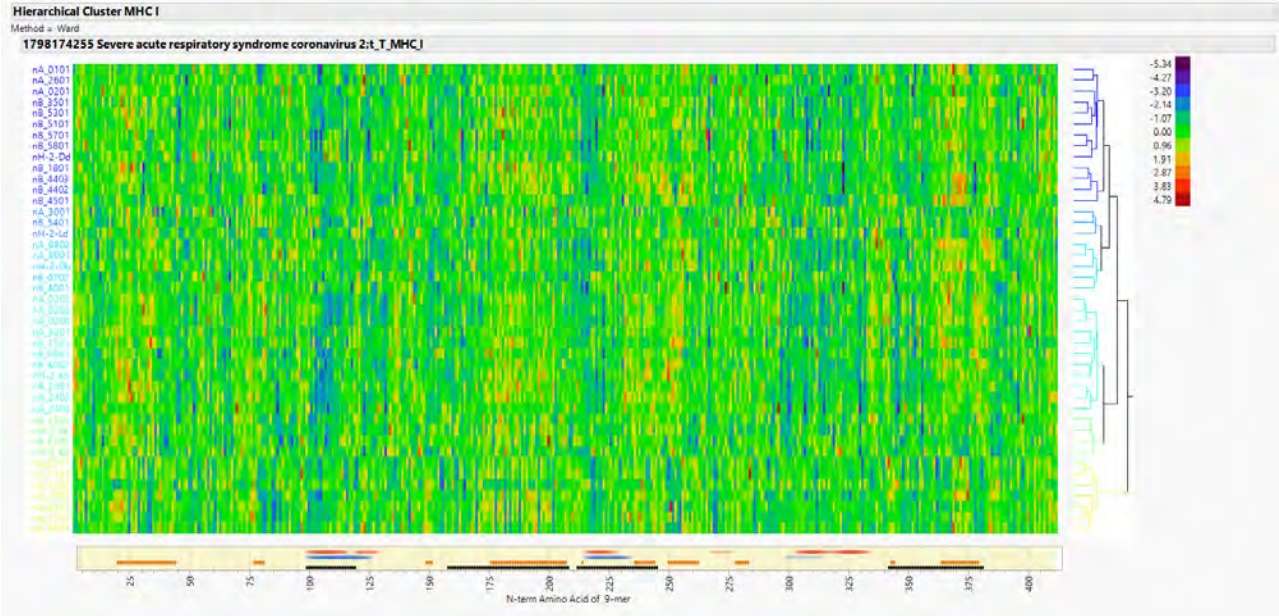


Nucleocapsid

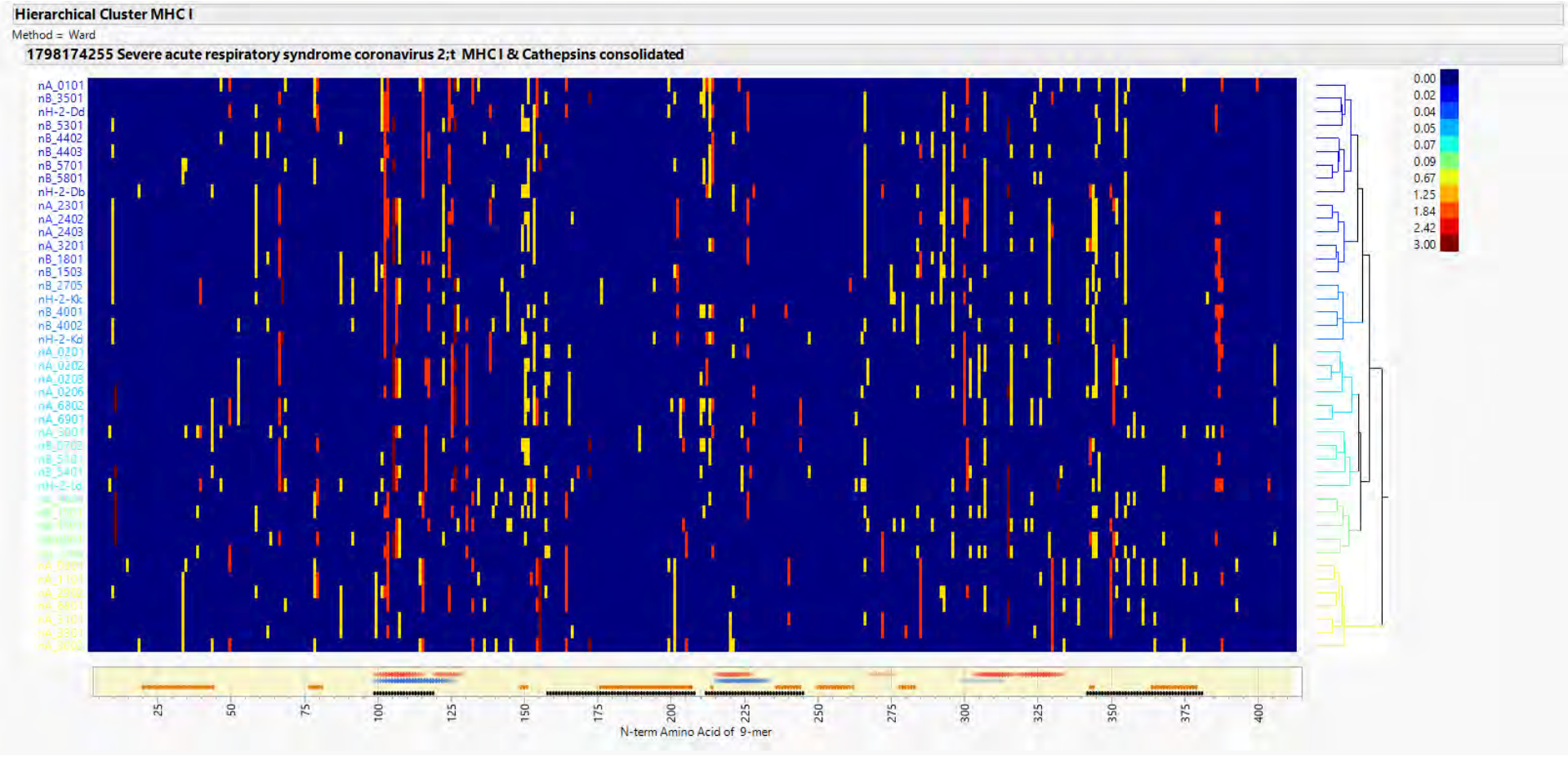


Nucleocapsid

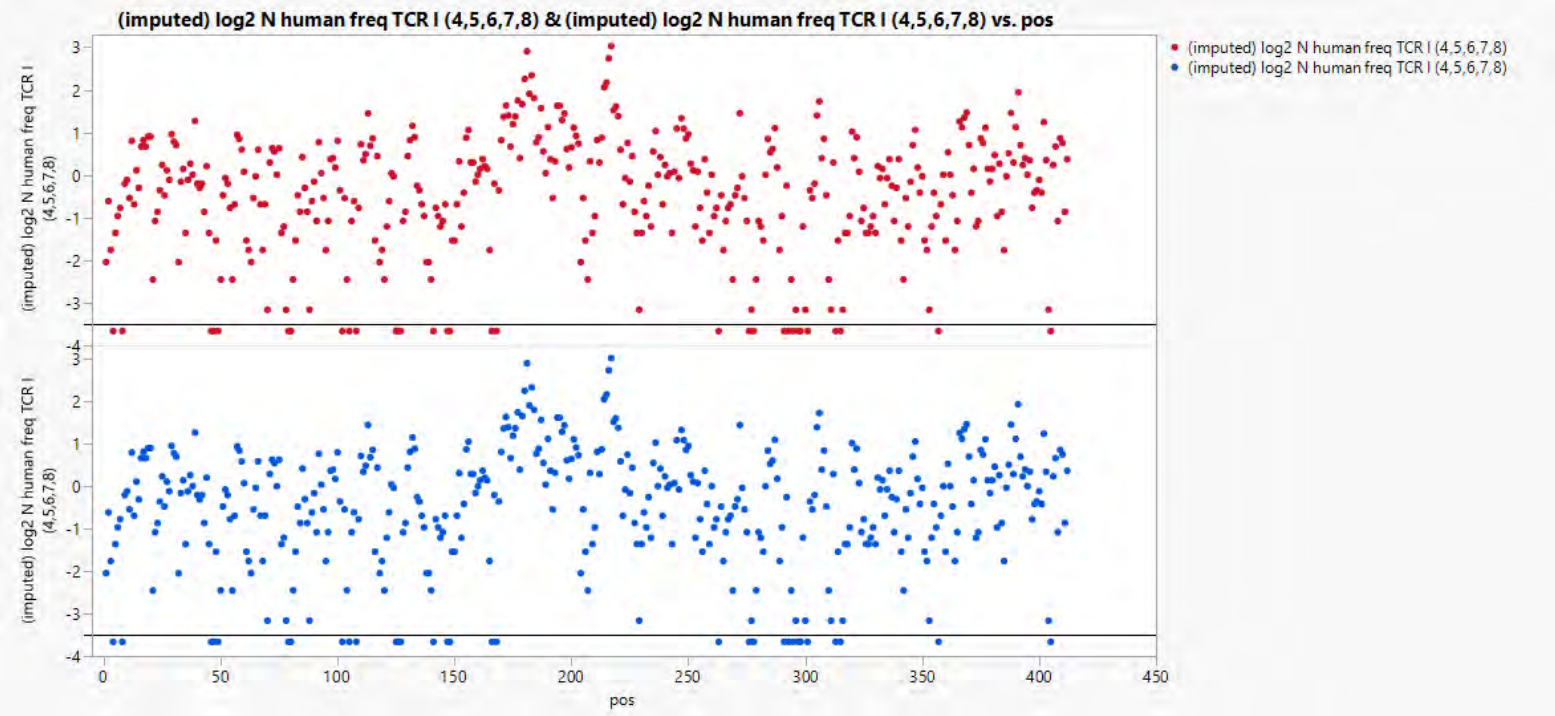




Nucleocapsid



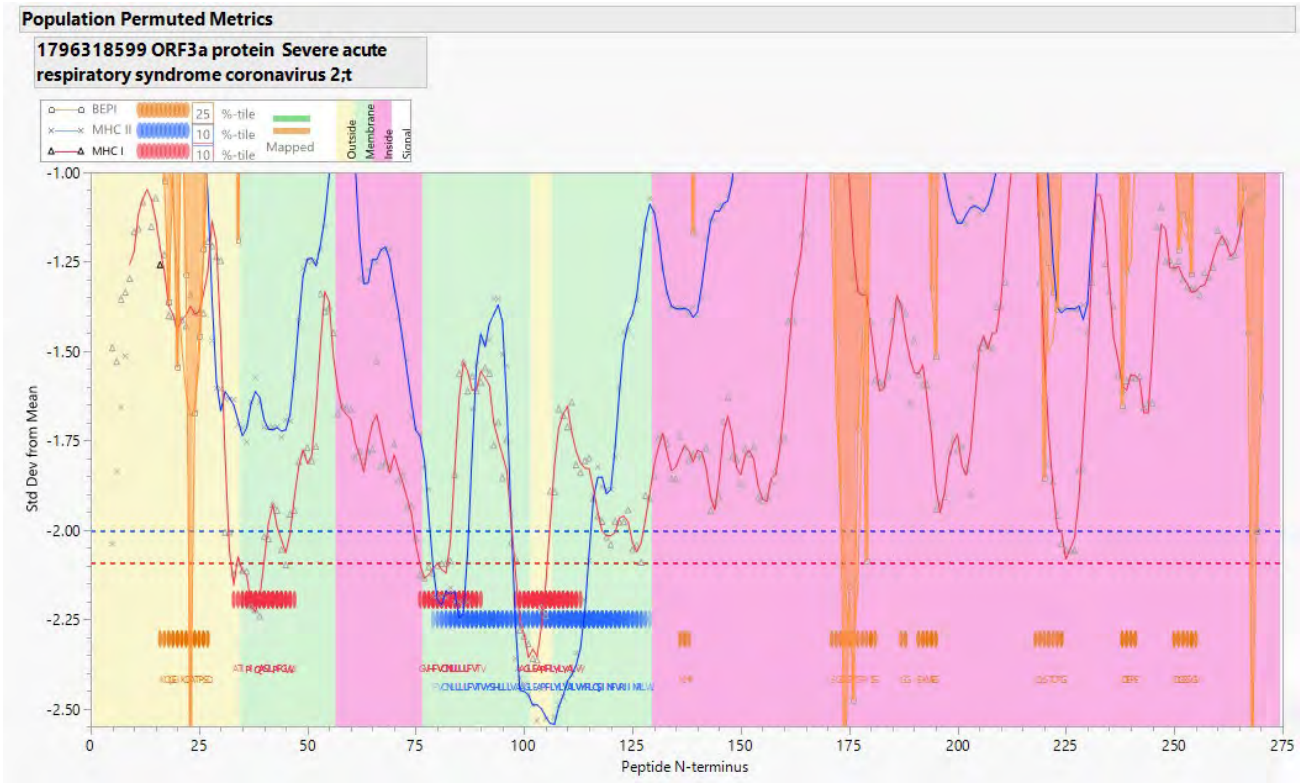
Nucleocapsid



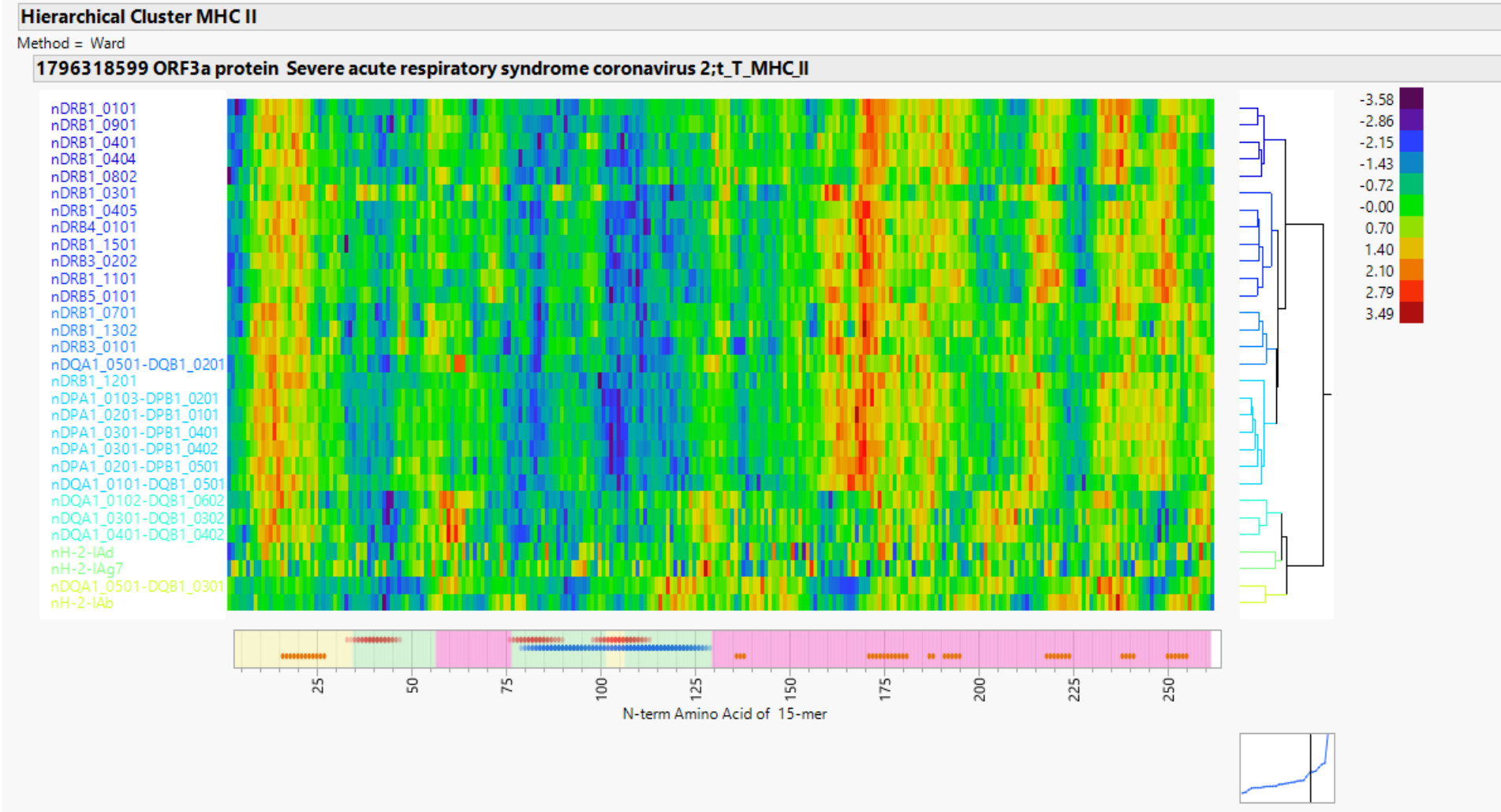
ORF3a

1796318599

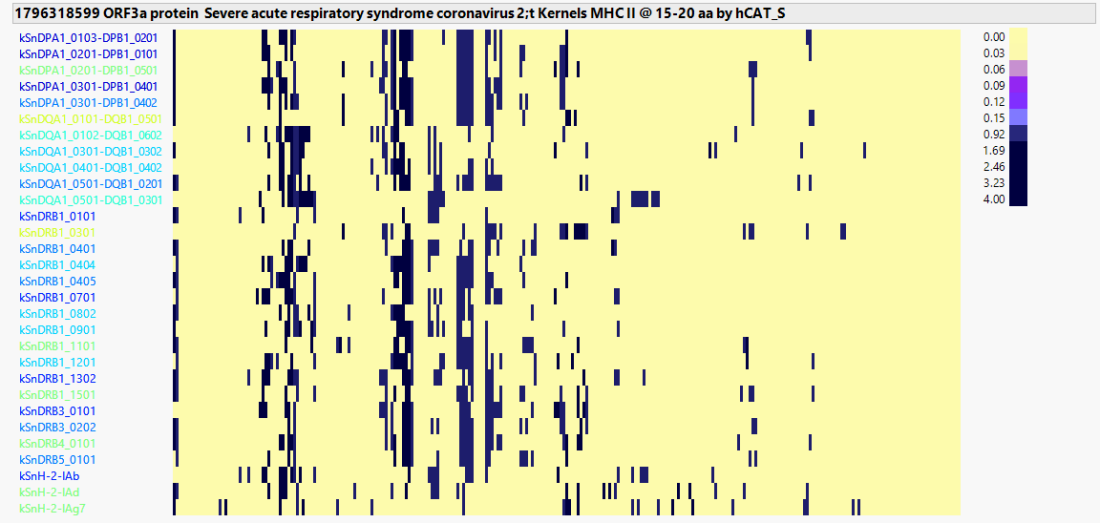
ORF3a

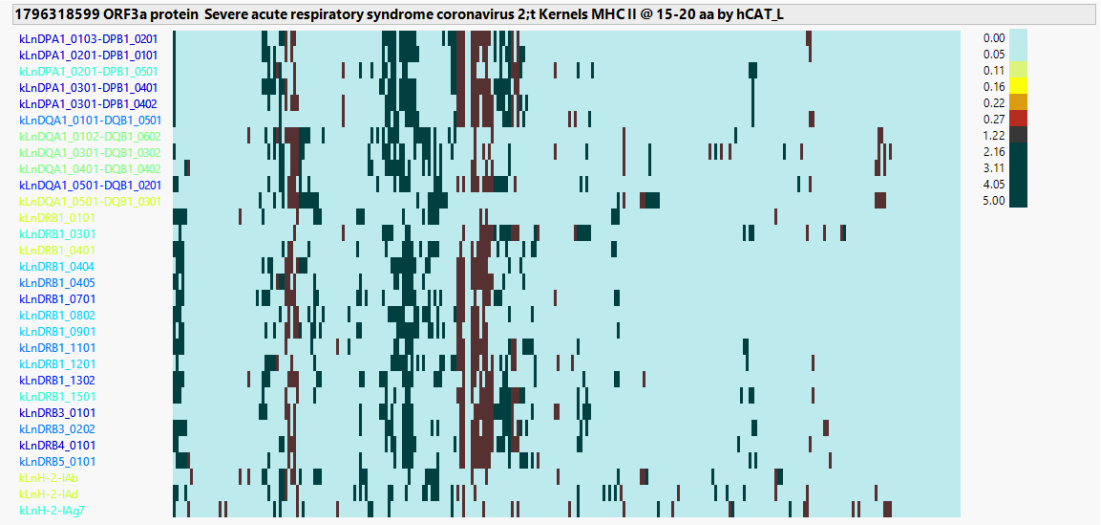

Population Permuted Metrics
1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Std Dev from Mean
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
Peptide N-terminus

ORF3a

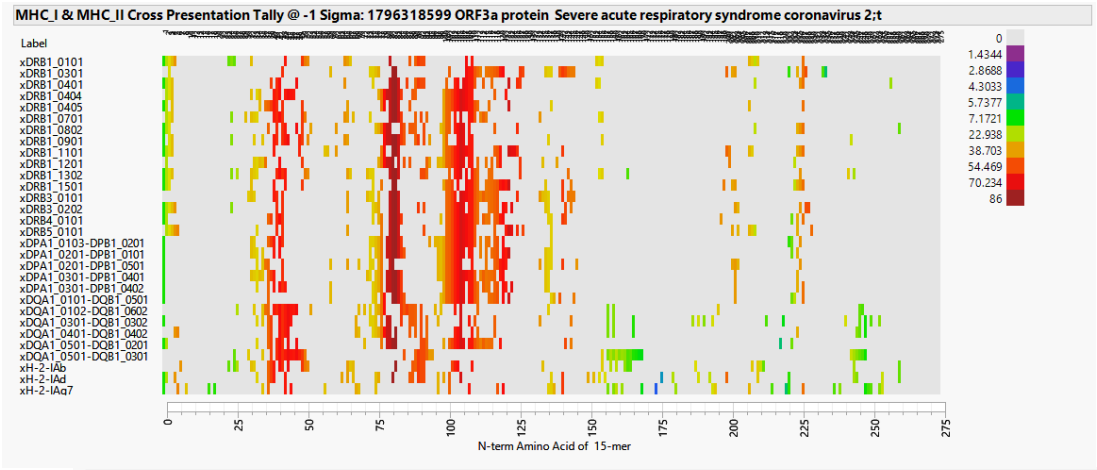

Hierarchical Cluster MHC II
Method = Ward
1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0401
nDRB1_0404
nDRB1_0802
nDRB1_0301
nDRB1_0405
nDRB4_0101
nDRB1_1501
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0701
nDRB1_1302
nDRB3_0101
nDQA1_0501-DQB1_0201
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.58
-2.86
-2.15
-1.43
-0.72
-0.00
0.70
1.40
2.10
2.79
3.49
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 15-mer

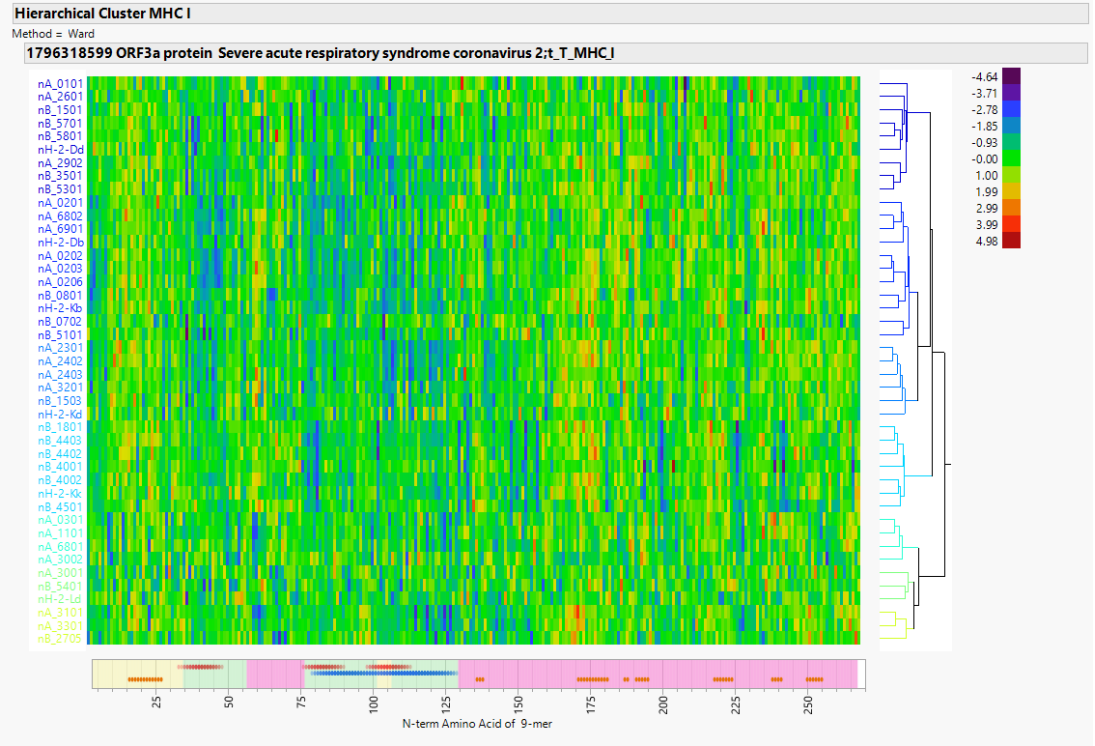
ORF3a


1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.06
0.09
0.12
0.15
0.92
1.69
2.46
3.23
4.00

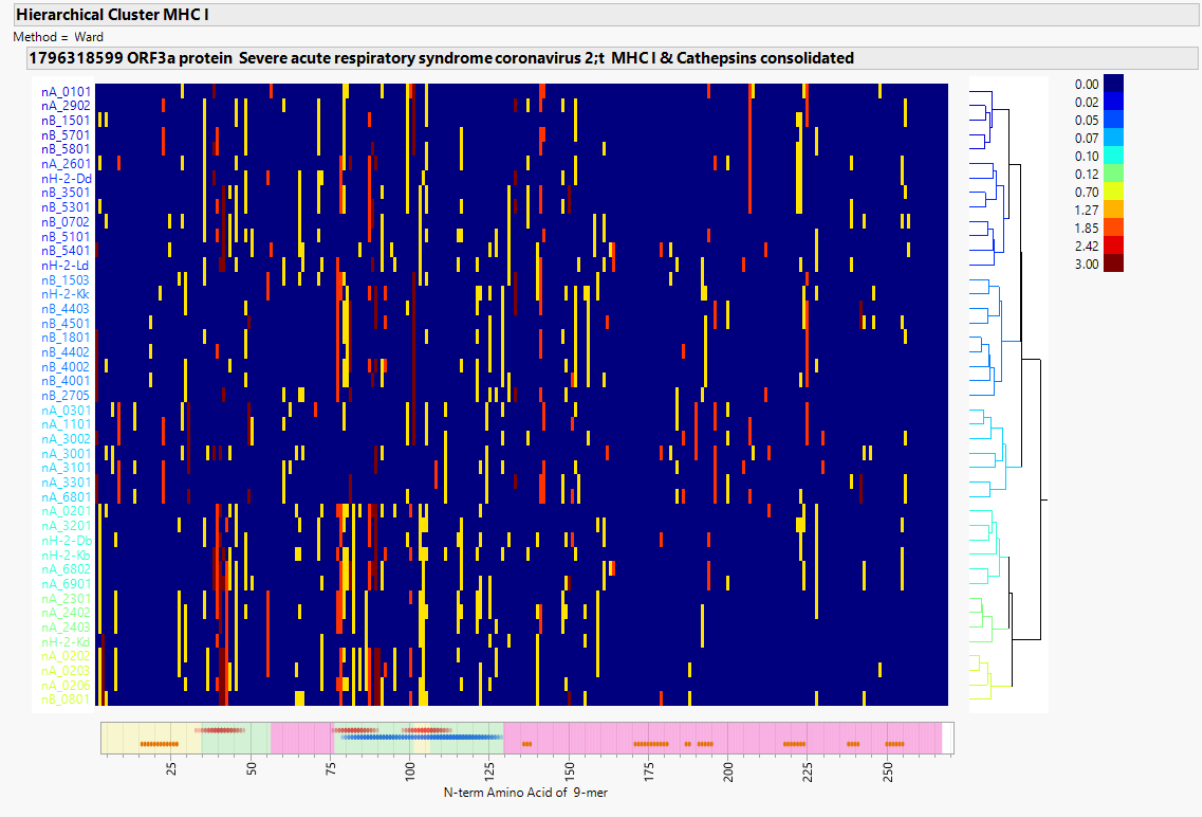

1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.05
0.11
0.16
0.22
0.27
1.22
2.16
3.11
4.05
5.00

ORF3a

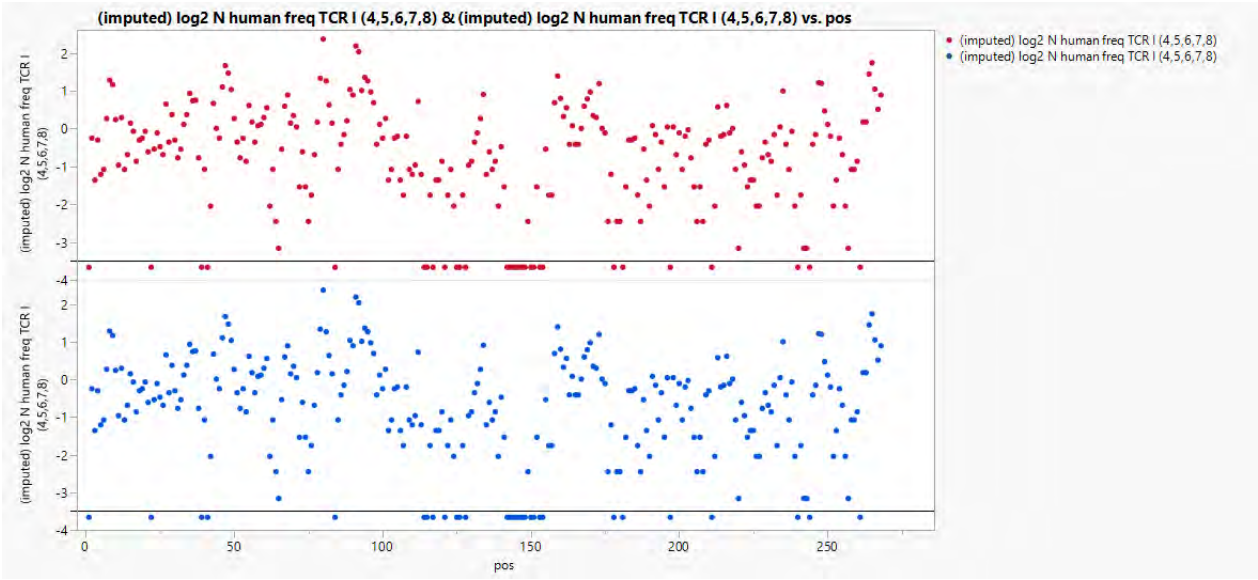

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.4344
2.8688
4.3033
5.7377
7.1721
22.938
38.703
54.469
70.234
86
0
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer


Hierarchical Cluster MHC I
Method = Ward
1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t_T_MHC_I
nA_0101
nA_2601
nB_1501
nB_5701
nB_5801
nH-2-Dd
nA_2902
nB_3501
nB_5301
nA_0201
nA_6802
nA_6901
nH-2-Db
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nB_0702
nB_5101
nA_2301
nA_2402
nA_2403
nA_3201
nB_1503
nH-2-Kd
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_6801
nA_3002
nA_3001
nB_5401
nH-2-Ld
nA_3101
nA_3301
nB_2705
-4.64
-3.71
-2.78
-1.85
-0.93
-0.00
1.00
1.99
2.99
3.99
4.98
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 9-mer

ORF3a


Hierarchical Cluster MHC I
Method = Ward
1796318599 ORF3a protein Severe acute respiratory syndrome coronavirus 2;t MHC I & Cathepsins consolidated
nA_0101
nA_2902
nB_1501
nB_5701
nB_5801
nA_2601
nH-2-Dd
nB_3501
nB_5301
nB_0702
nB_5101
nB_5401
nH-2-Ld
nB_1503
nH-2-Kk
nB_4403
nB_4501
nB_1801
nB_4402
nB_4002
nB_4001
nB_2705
nA_0301
nA_1101
nA_3002
nA_3001
nA_3101
nA_3301
nA_6801
nA_0201
nA_3201
nH-2-Db
nH-2-Kb
nA_6802
nA_6901
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_0202
nA_0203
nA_0206
nB_0801
0.00
0.02
0.05
0.07
0.10
0.12
0.70
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 9-mer

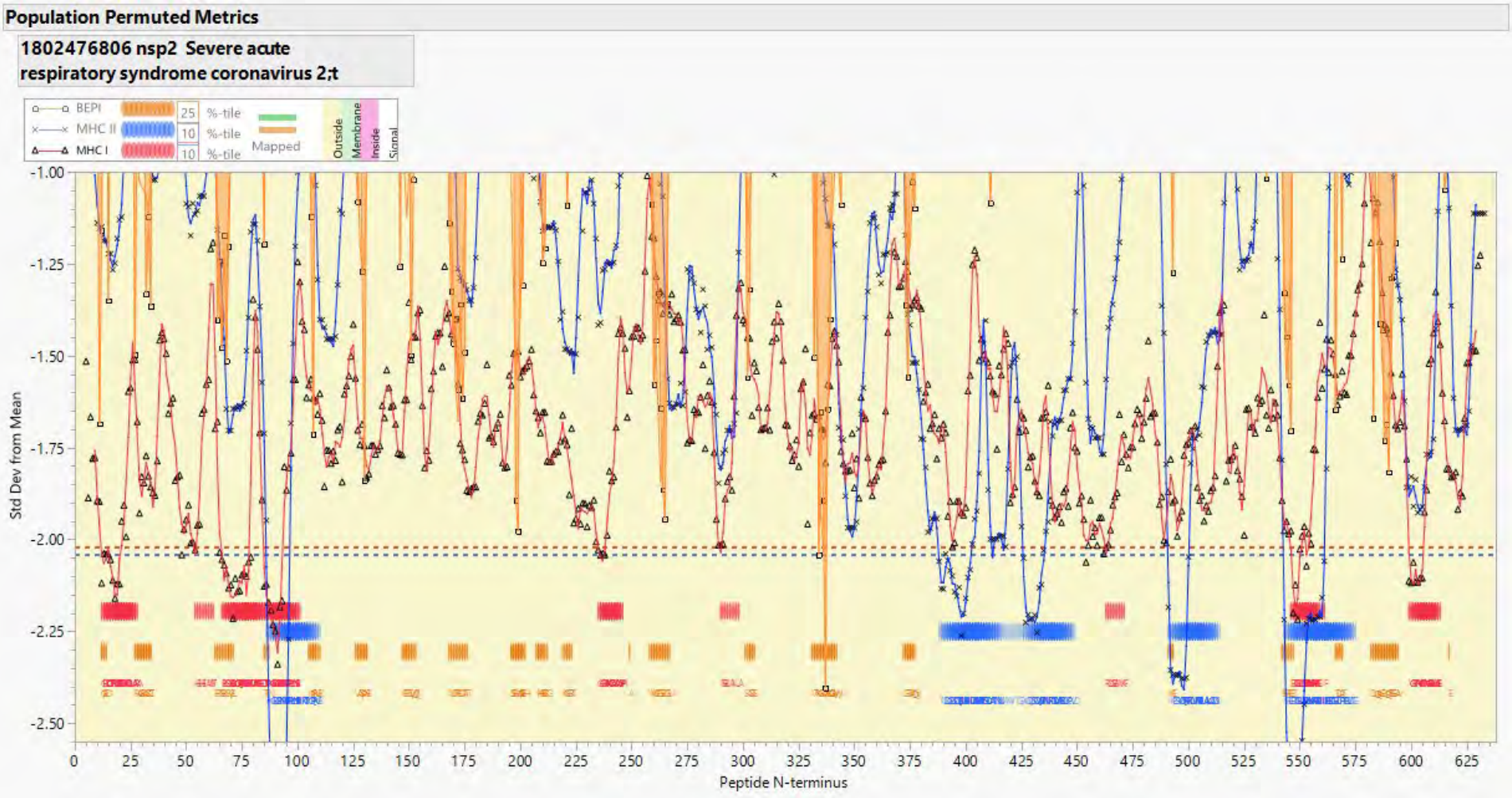
ORF3a


(imputed) log2 N human freq TCR I (4,5,6,7,8) & (imputed) log2 N human freq TCR I (4,5,6,7,8) vs. pos
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
pos
0
50
100
150
200
250

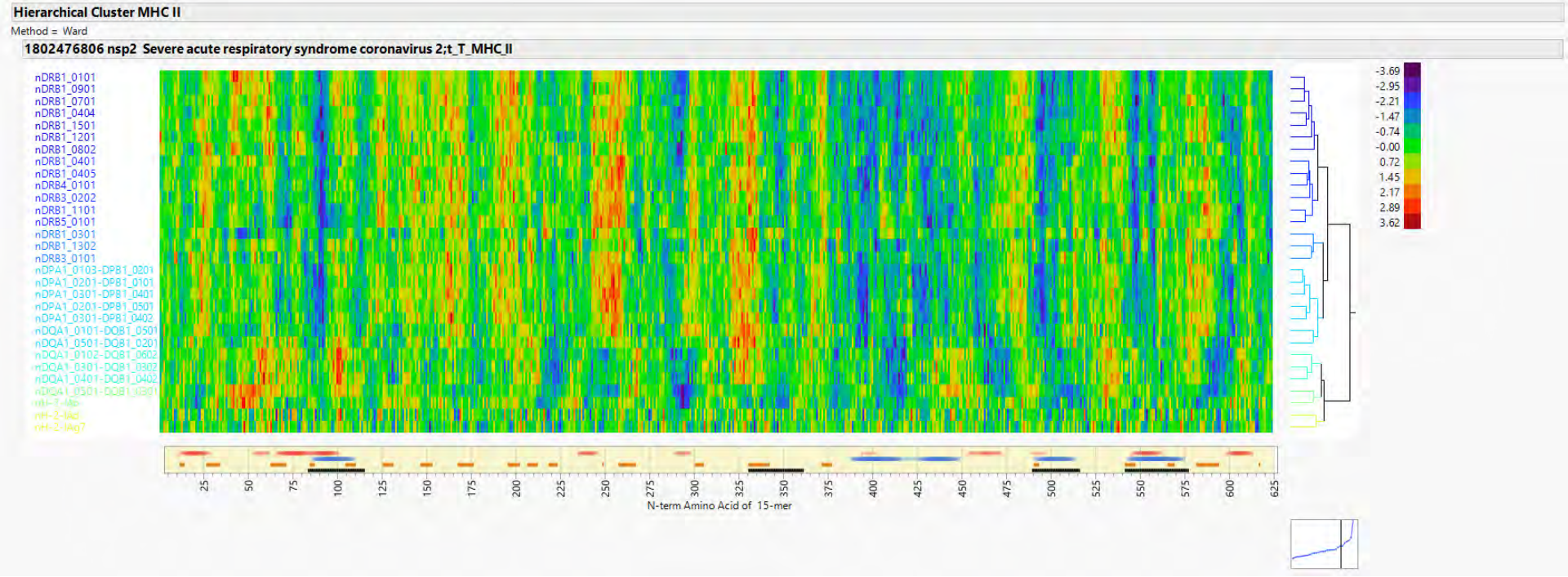
nsp2

1802476806

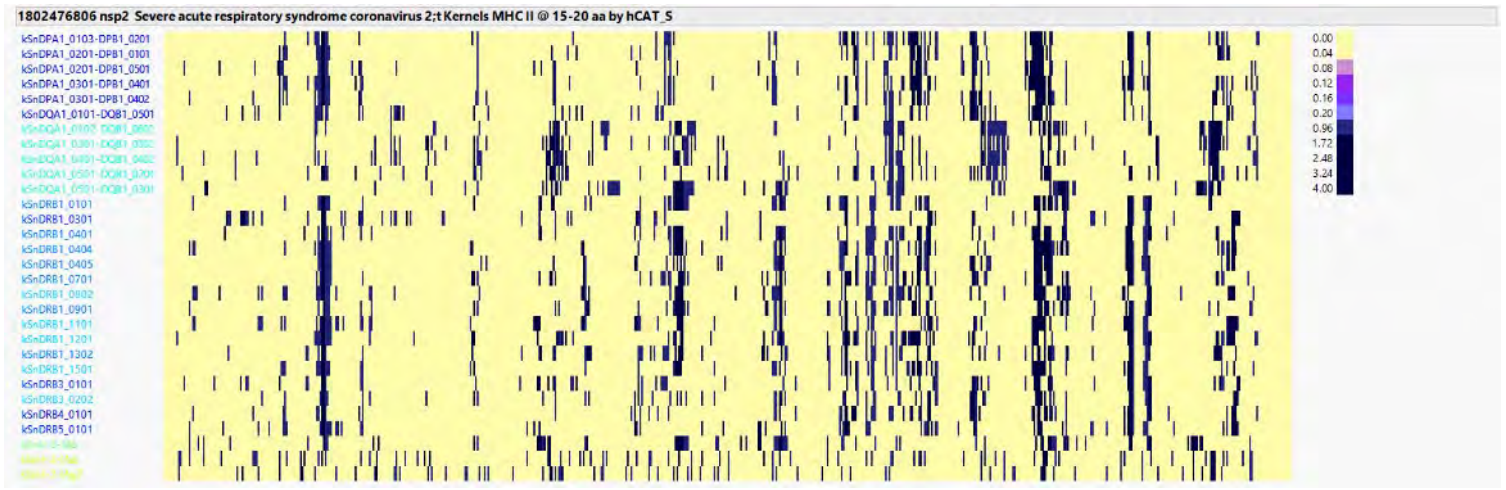
nsp2

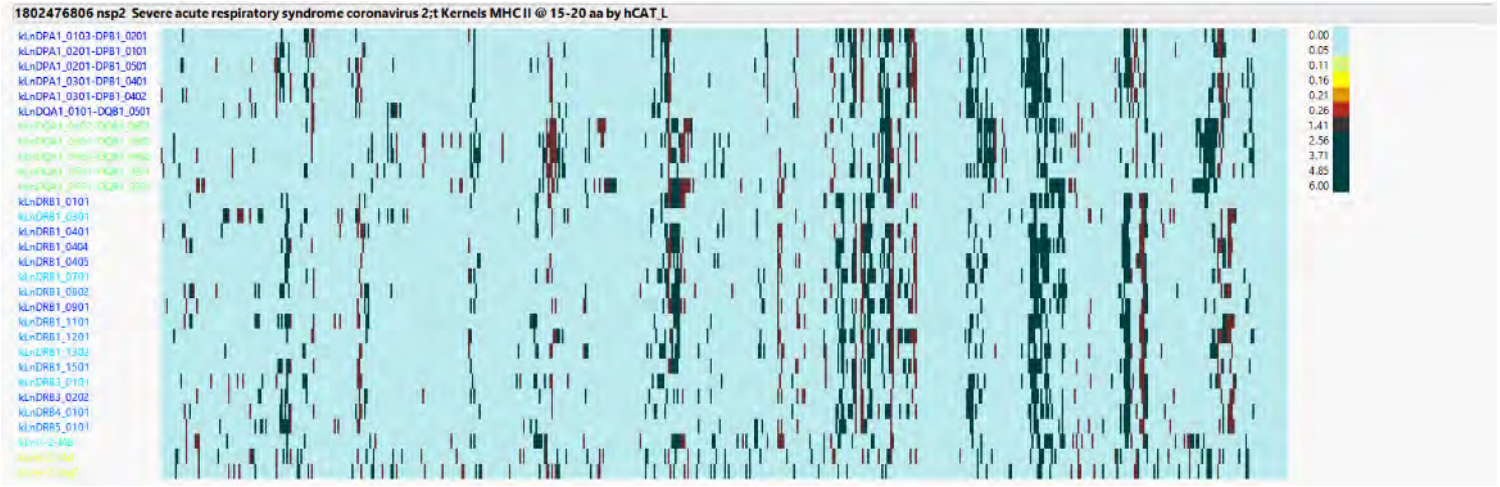

Population Permuted Metrics
1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Std Dev from Mean
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625
Peptide N-terminus

nsp2

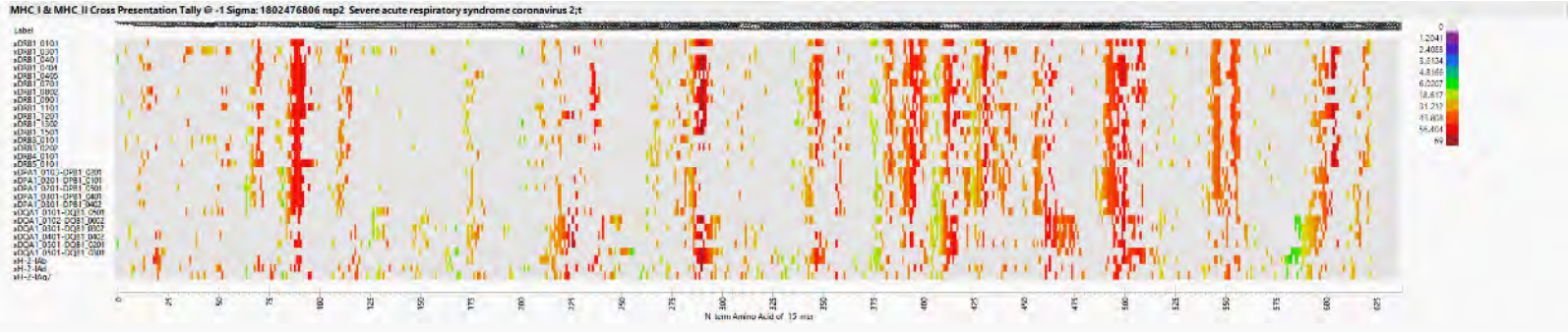

Hierarchical Cluster MHC II
Method = Ward
1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1501
nDRB1_1201
nDRB1_0802
nDRB1_0401
nDRB1_0405
nDRB4_0101
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB1_1302
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
-3.69
-2.95
-2.21
-1.47
-0.74
-0.00
0.72
1.45
2.17
2.89
3.62
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
N-term Amino Acid of 15-mer

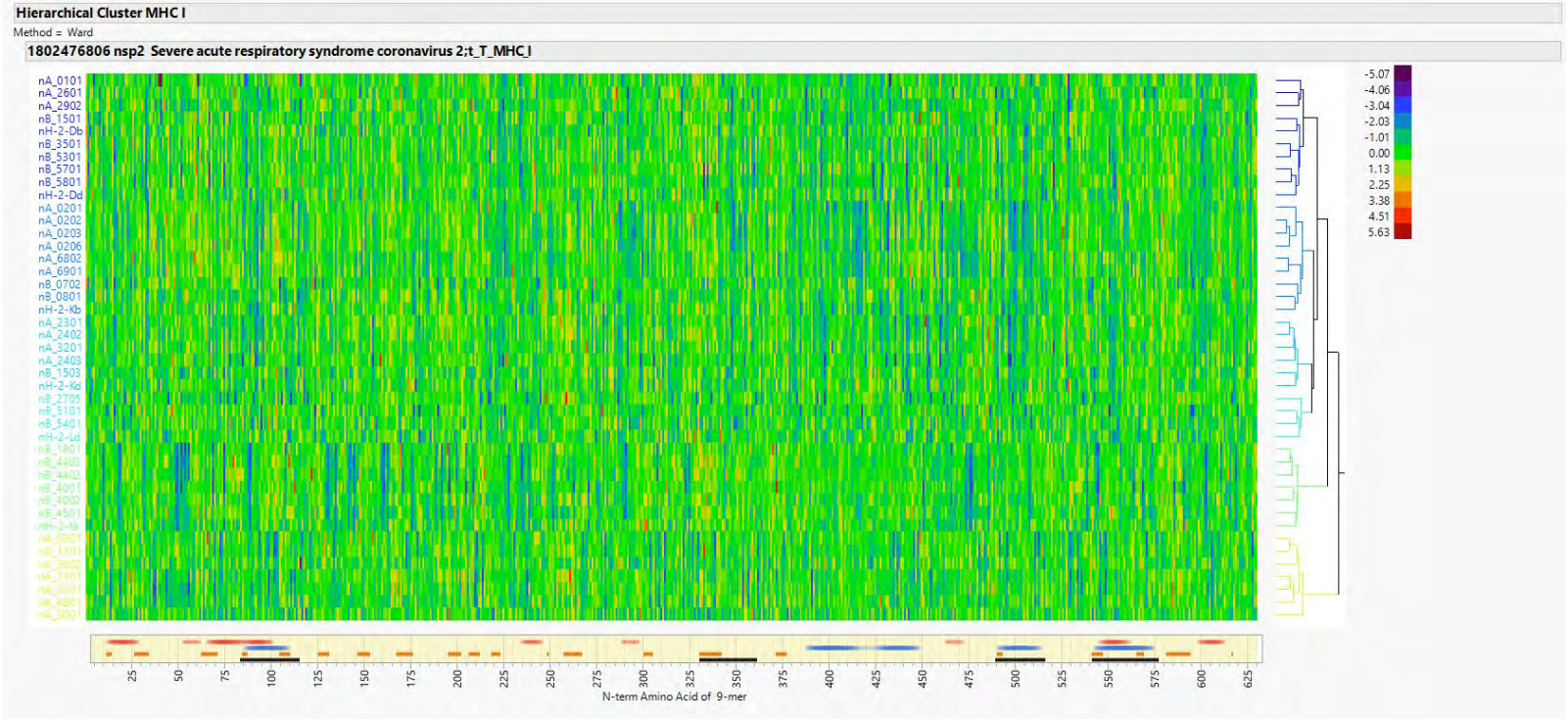
nsp2


1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
0.00
0.04
0.08
0.12
0.16
0.20
0.96
1.72
2.48
3.24
4.00

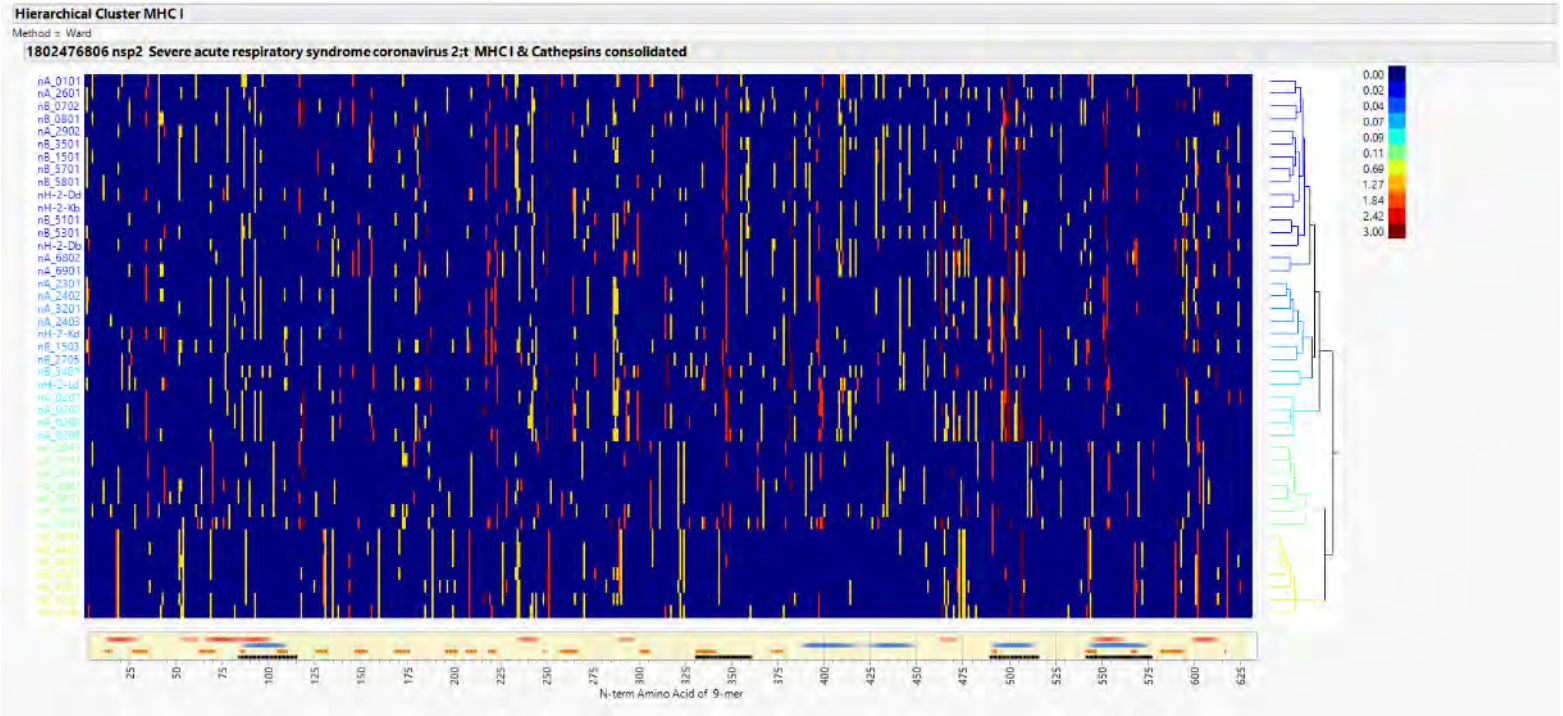

1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDRB1_0101
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1501
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
0.00
0.05
0.11
0.16
0.21
0.26
1.41
2.56
3.71
4.85
6.00

nsp2

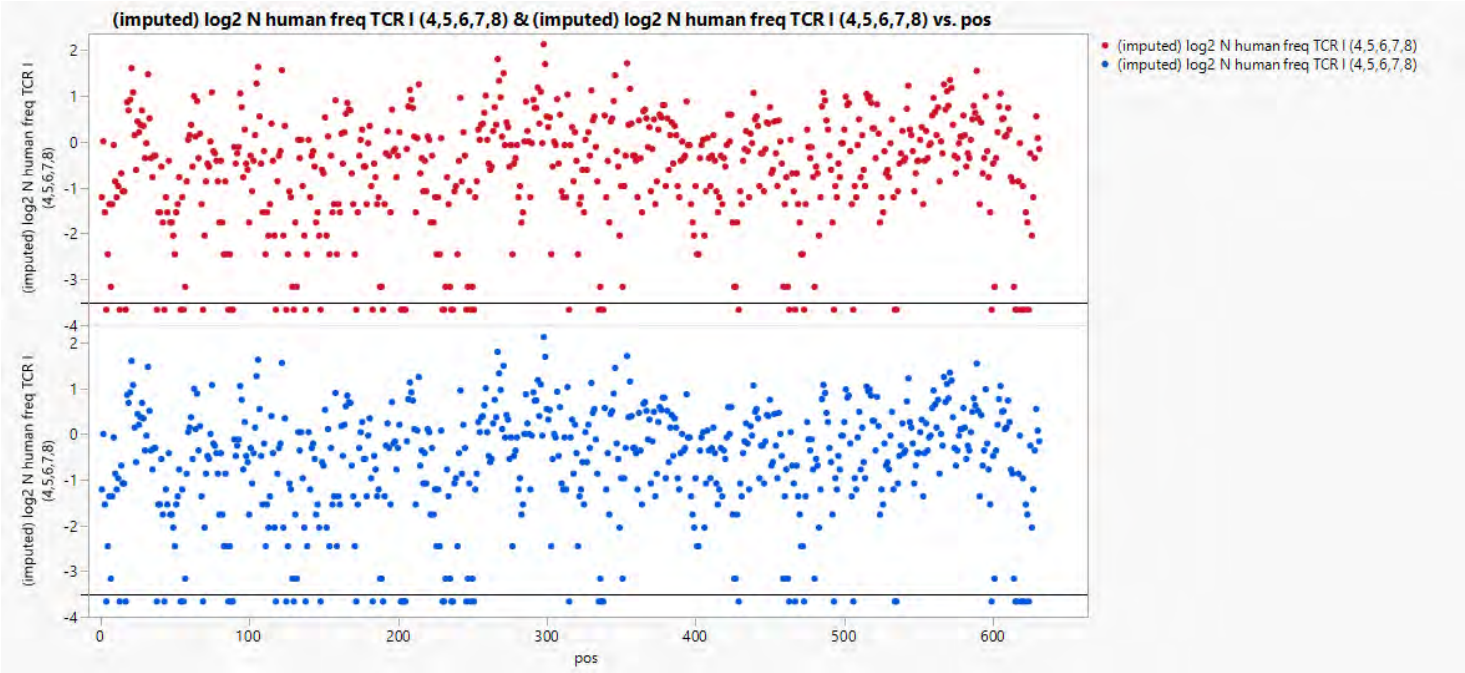

MHC I & MHC_II Cross Presentation Tally @ -1 Sigma: 1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t
Label
N-term Amino Acid of 15-mer


Hierarchical Cluster MHC I
Method = Ward
1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nB_1501
nH-2-Db
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2408
nB_1503
nH-2-Kd
nB_2705
nB_3101
nB_5401
nH-2-Ld
nB_1801
nB_4405
nB_4402
nB_4001
nB_4002
nB_4501
-5.07
-4.06
-3.04
-2.03
-1.01
0.00
1.13
2.25
3.38
4.51
5.63
N-term Amino Acid of 9-mer

nsp2


Hierarchical Cluster MHC I
Method = Ward
1802476806 nsp2 Severe acute respiratory syndrome coronavirus 2;t MHC I & Cathepsins consolidated
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
N-term Amino Acid of 9-mer

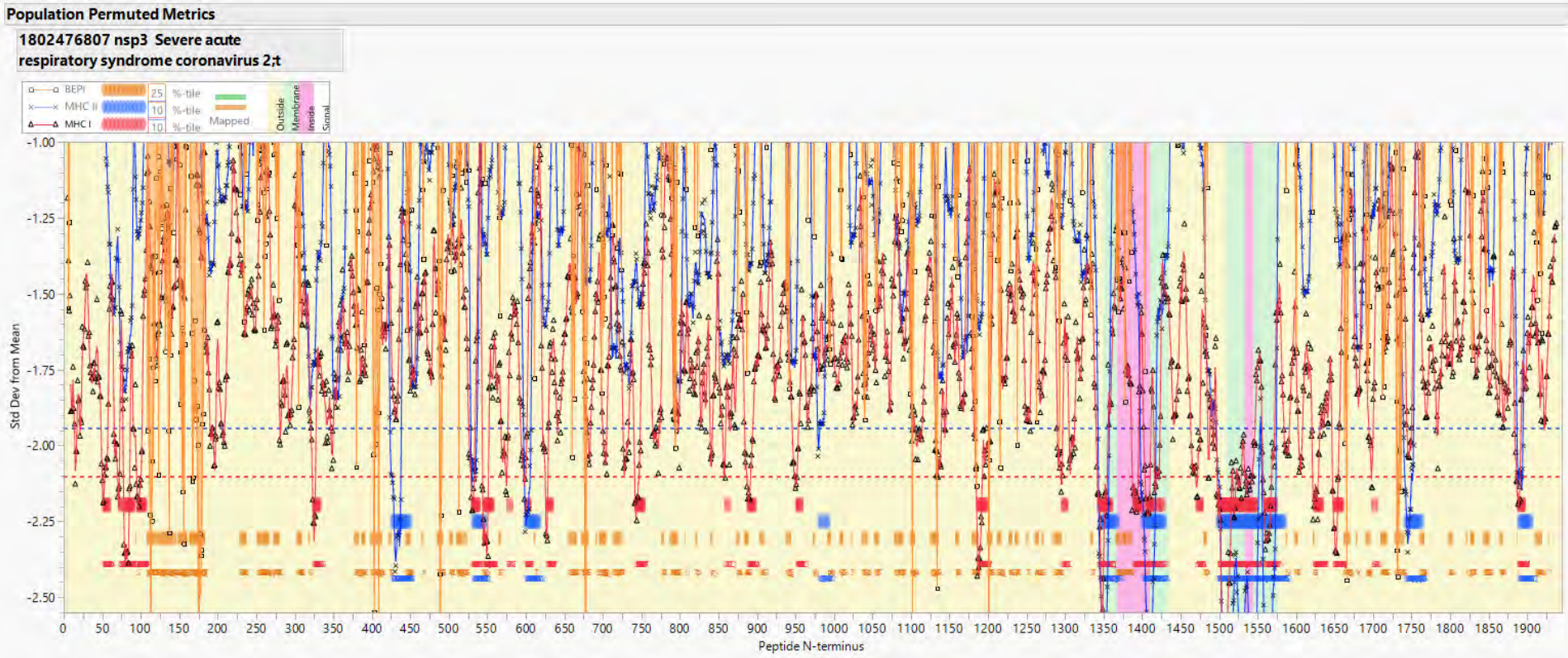
nsp2


(imputed) log2 N human freq TCR I (4,5,6,7,8) & (imputed) log2 N human freq TCR I (4,5,6,7,8) vs. pos
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
pos
0
100
200
300
400
500
600

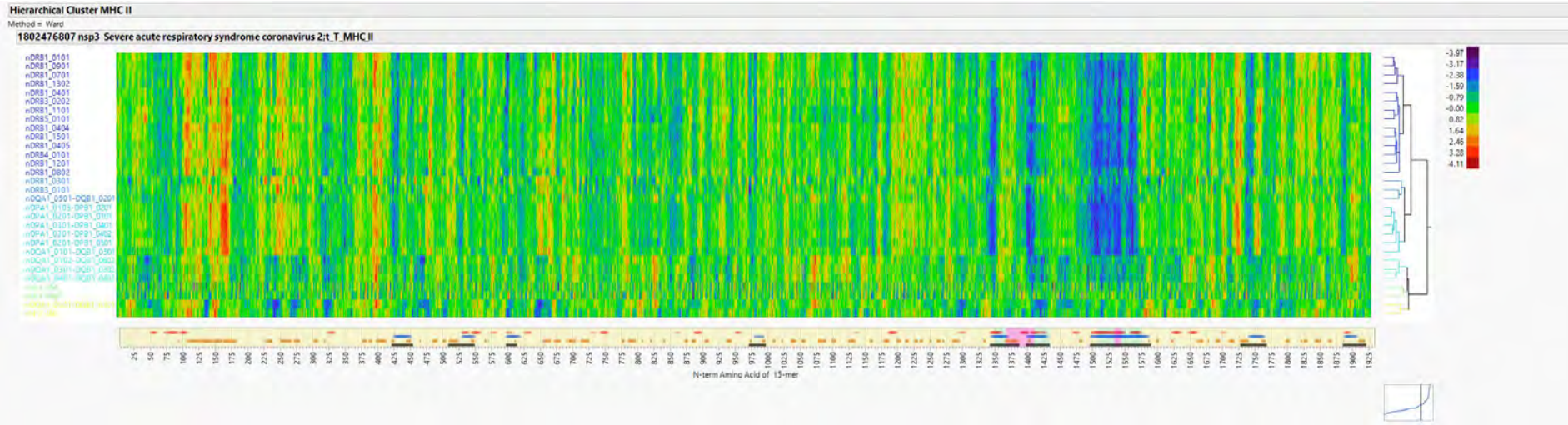
nsp3

1802476807

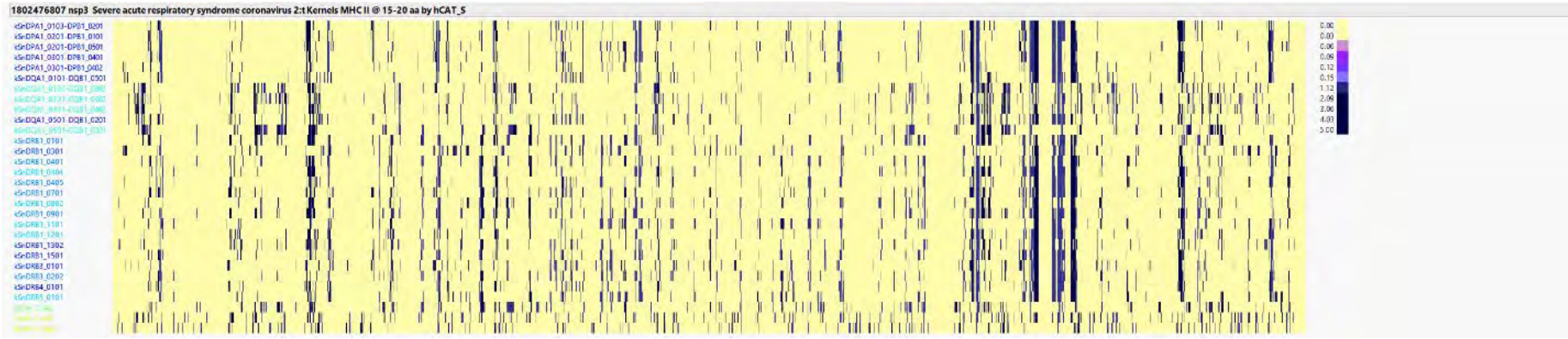
nsp3

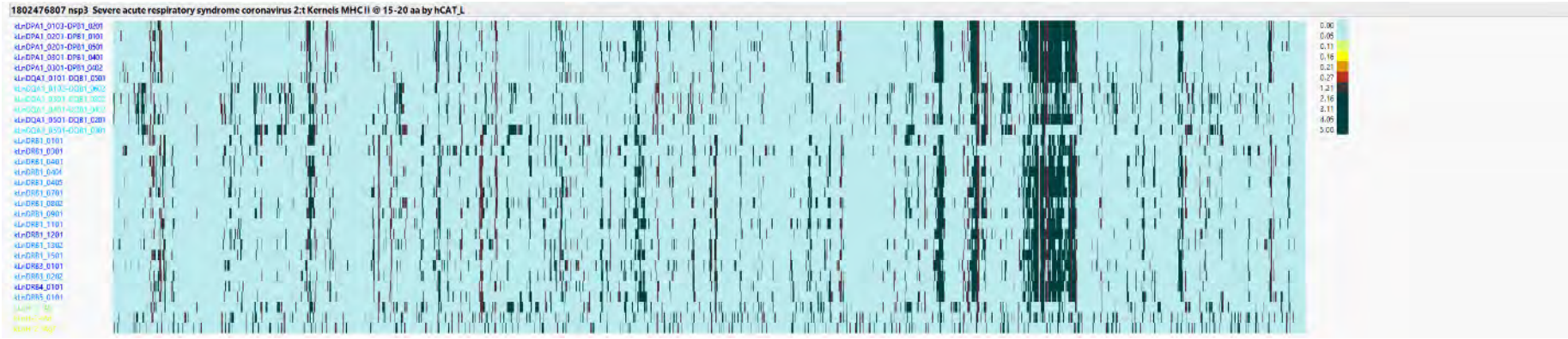

Population Permuted Metrics
1802476807 nsp3 Severe acute respiratory syndrome coronavirus 2;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Std Dev from Mean
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0 50 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250 1300 1350 1400 1450 1500 1550 1600 1650 1700 1750 1800 1850 1900
Peptide N-terminus

nsp3

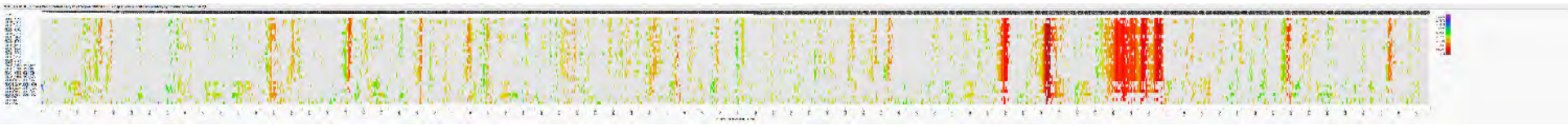

Hierarchical Cluster MHC II
Method = Ward
1802476807 nsp3 Severe acute respiratory syndrome coronavirus 2:t_T_MHC_II
N-term Amino Acid of 15-mer

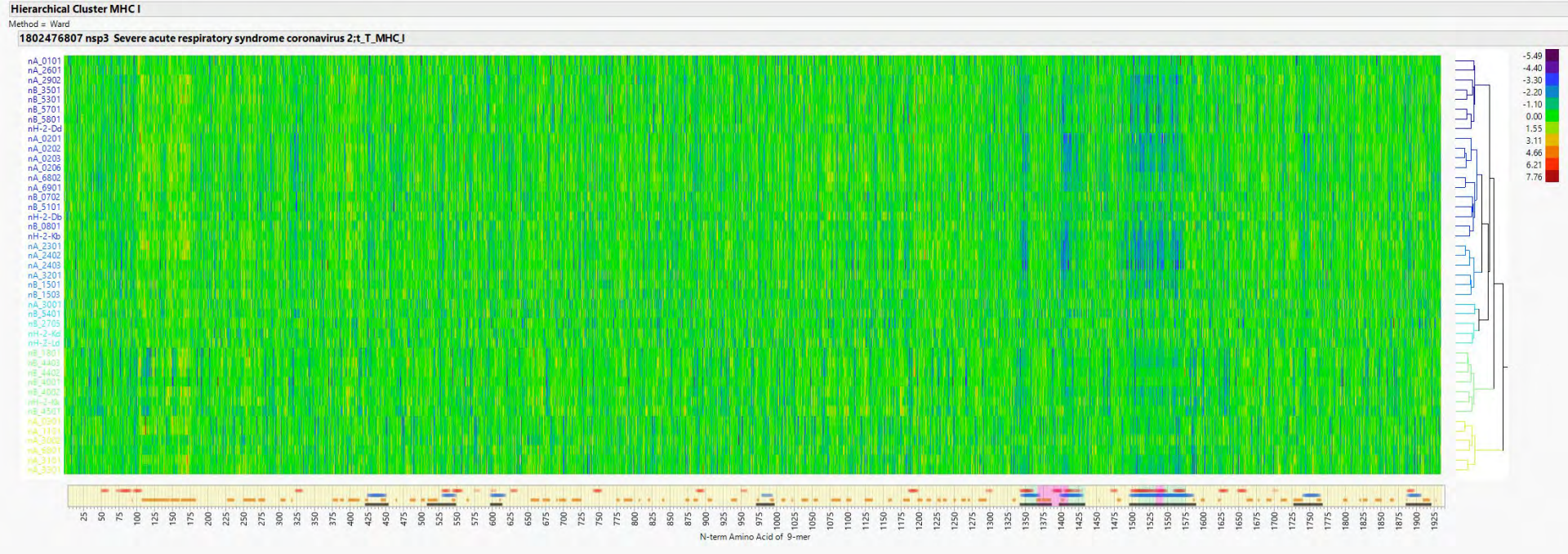
nsp3


1802476807 nsp3 Severe acute respiratory syndrome coronavirus 2:t Kernels MHC II @ 15-20 aa by hCAT_S

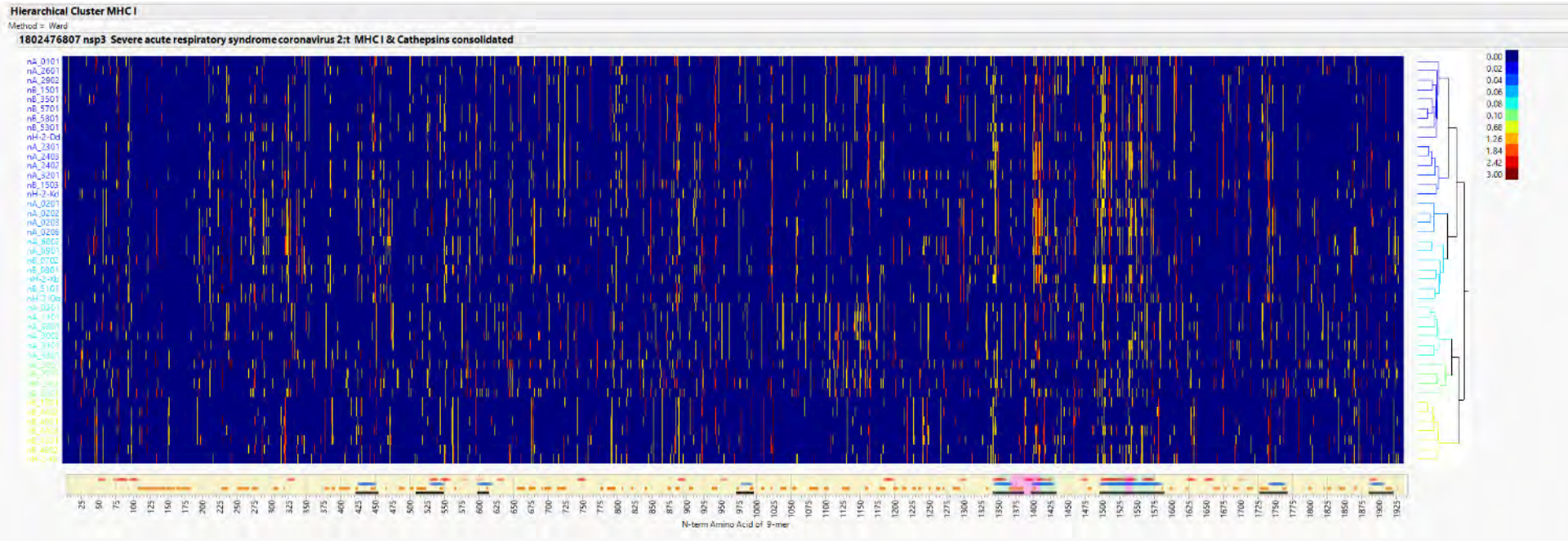

1802476807 nsp3 Severe acute respiratory syndrome coronavirus 2:t Kernels MHC II @ 15-20 aa by hCAT_L

nsp3

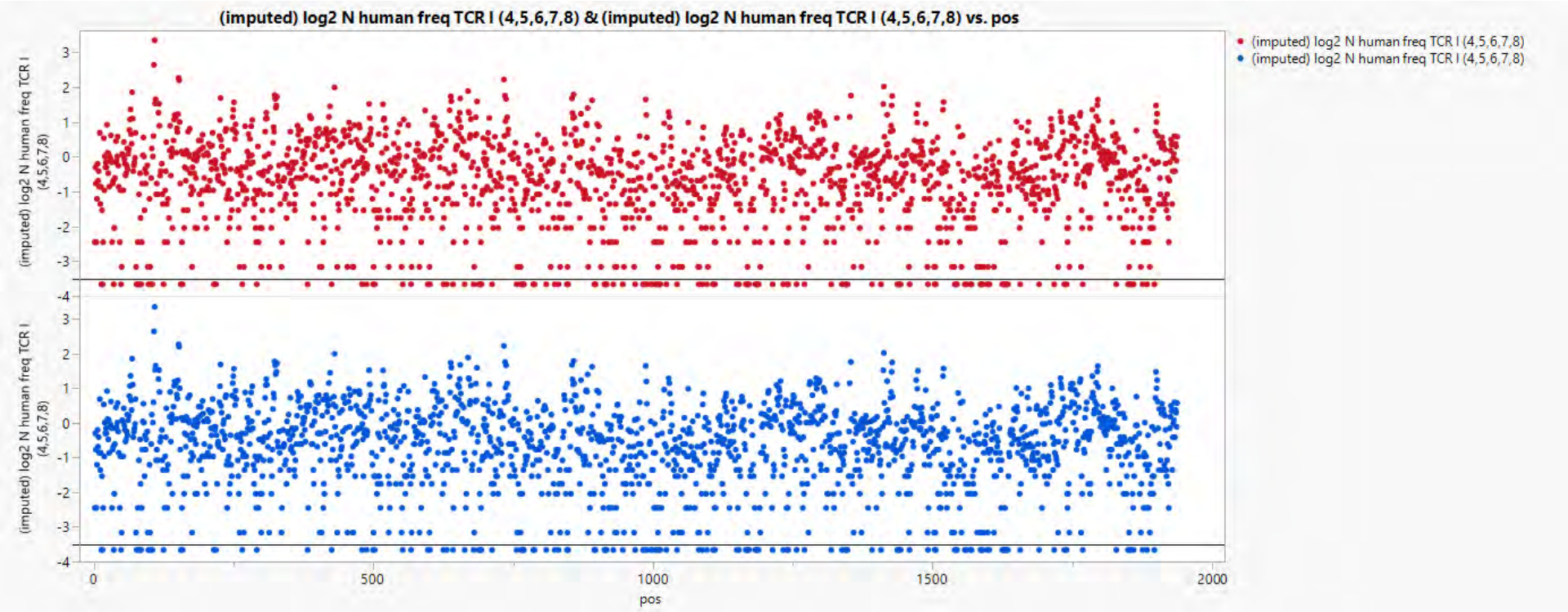



Hierarchical Cluster MHC I
Method = Ward
1802476807 nsp3 Severe acute respiratory syndrome coronavirus 2;t_T_MHC_I
N-term Amino Acid of 9-mer

nsp3


Hierarchical Cluster MHC I
Method = Ward
1802476807 nsp3 Severe acute respiratory syndrome coronavirus 2;t MHC I & Cathepsins consolidated
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
N-term Amino Acid of 9-mer

nsp3


(imputed) log2 N human freq TCR I (4,5,6,7,8) & (imputed) log2 N human freq TCR I (4,5,6,7,8) vs. pos
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
(imputed) log2 N human freq TCR I (4,5,6,7,8)
3
2
1
0
-1
-2
-3
-4
0
500
1000
1500
2000
pos

Appendix 5

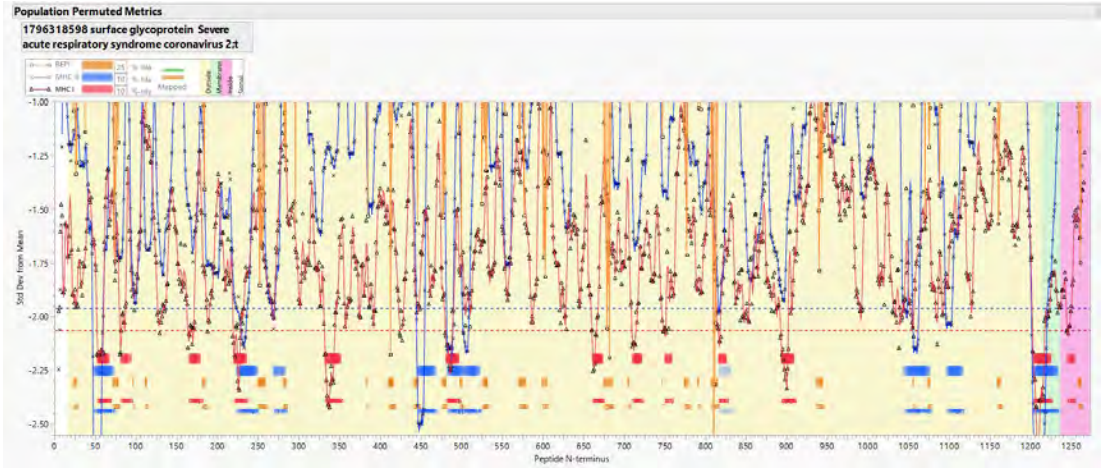
Comparison of SARS-Cov-2 RefSeq to UK and South African variants of concern – spike and nucleocapsid proteins

Mutations relative to reference strain

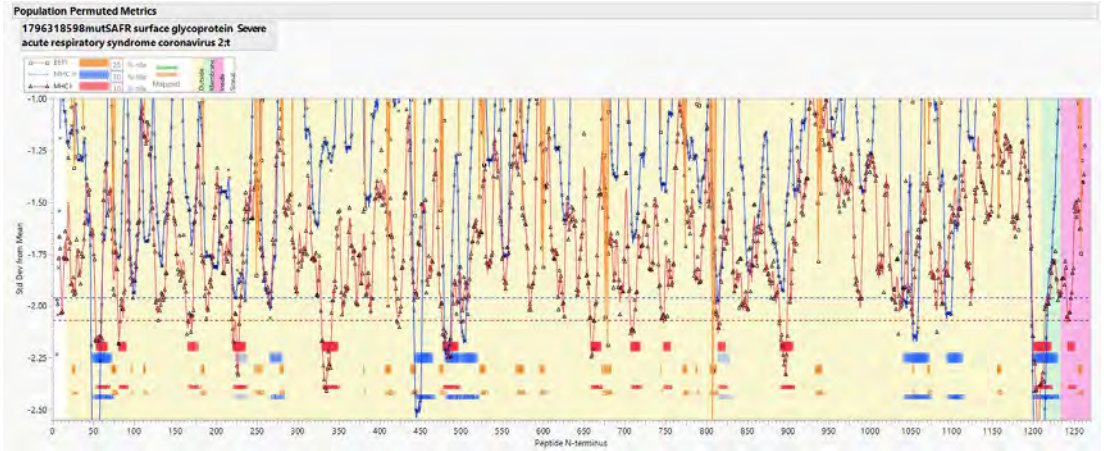
	B.1.1.7 UK	B.1.351 South Africa
Spike Glycoprotein	A570D, D614G, D1118H, H69del, N501Y, P681H, S982A, T716I, V70del, Y145del,	A243del, A701V, D80A, D215G, D624G, E484K, K417N, L242del, L244del, N501Y
Nucleocapsid	D3L, G204R, R203K, S235F	T205I

Spike Glycoprotein

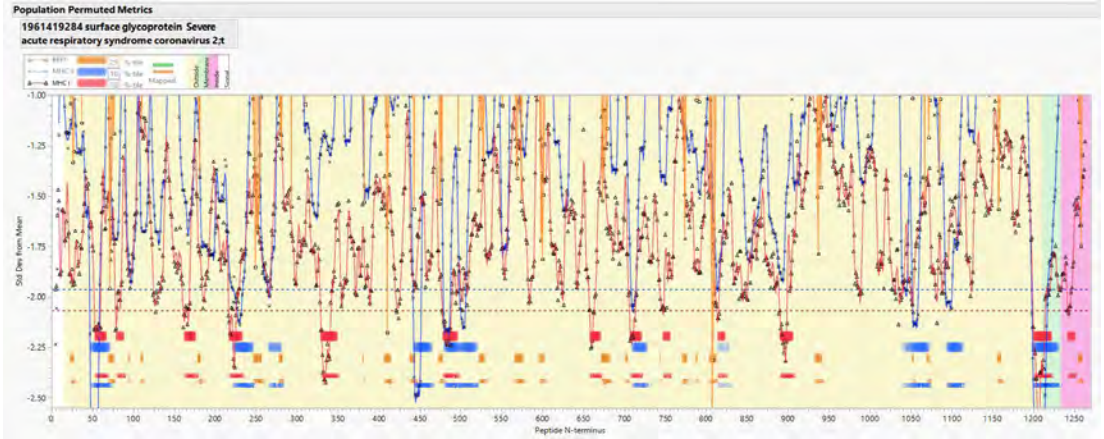
Ref



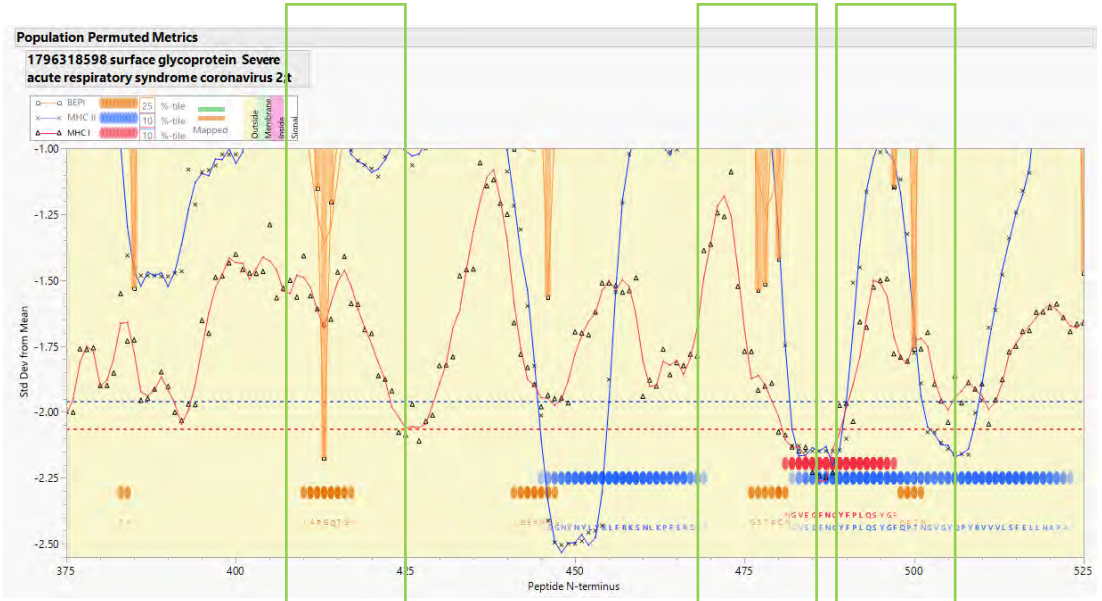
S Africa



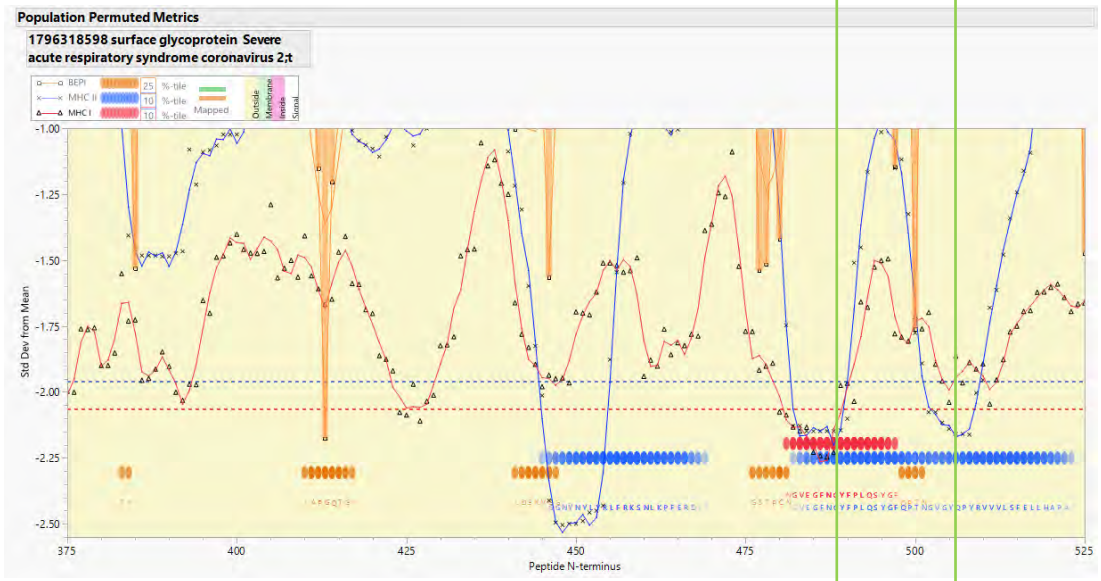
UK



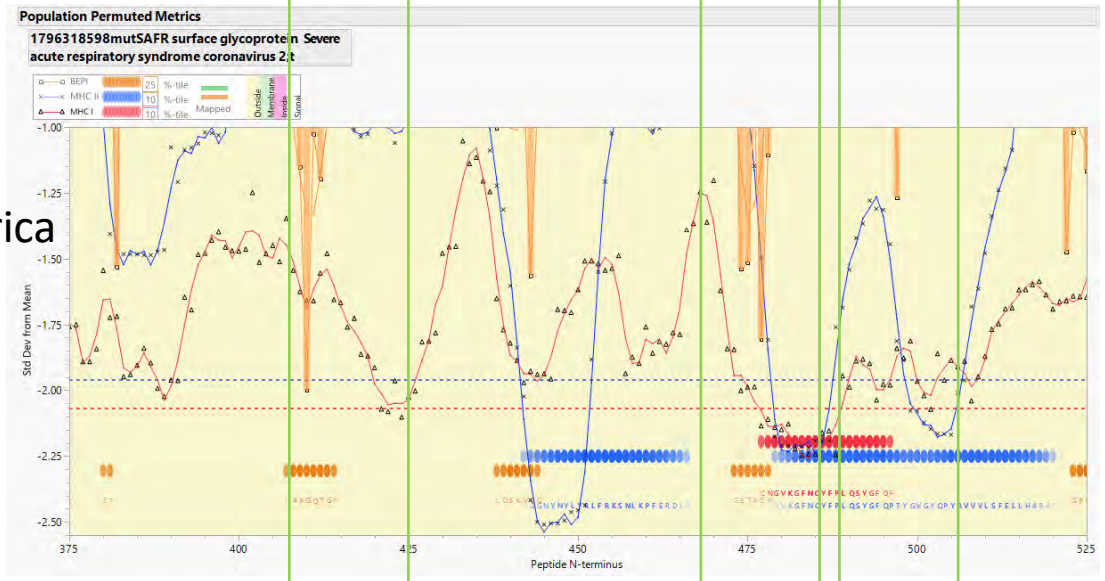
Ref



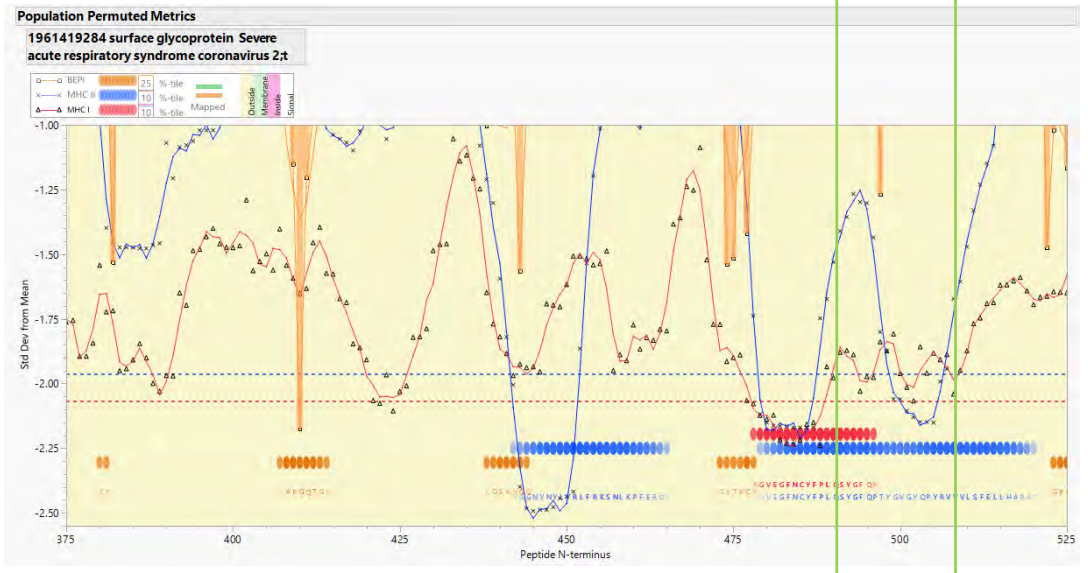
Ref



S Africa



UK



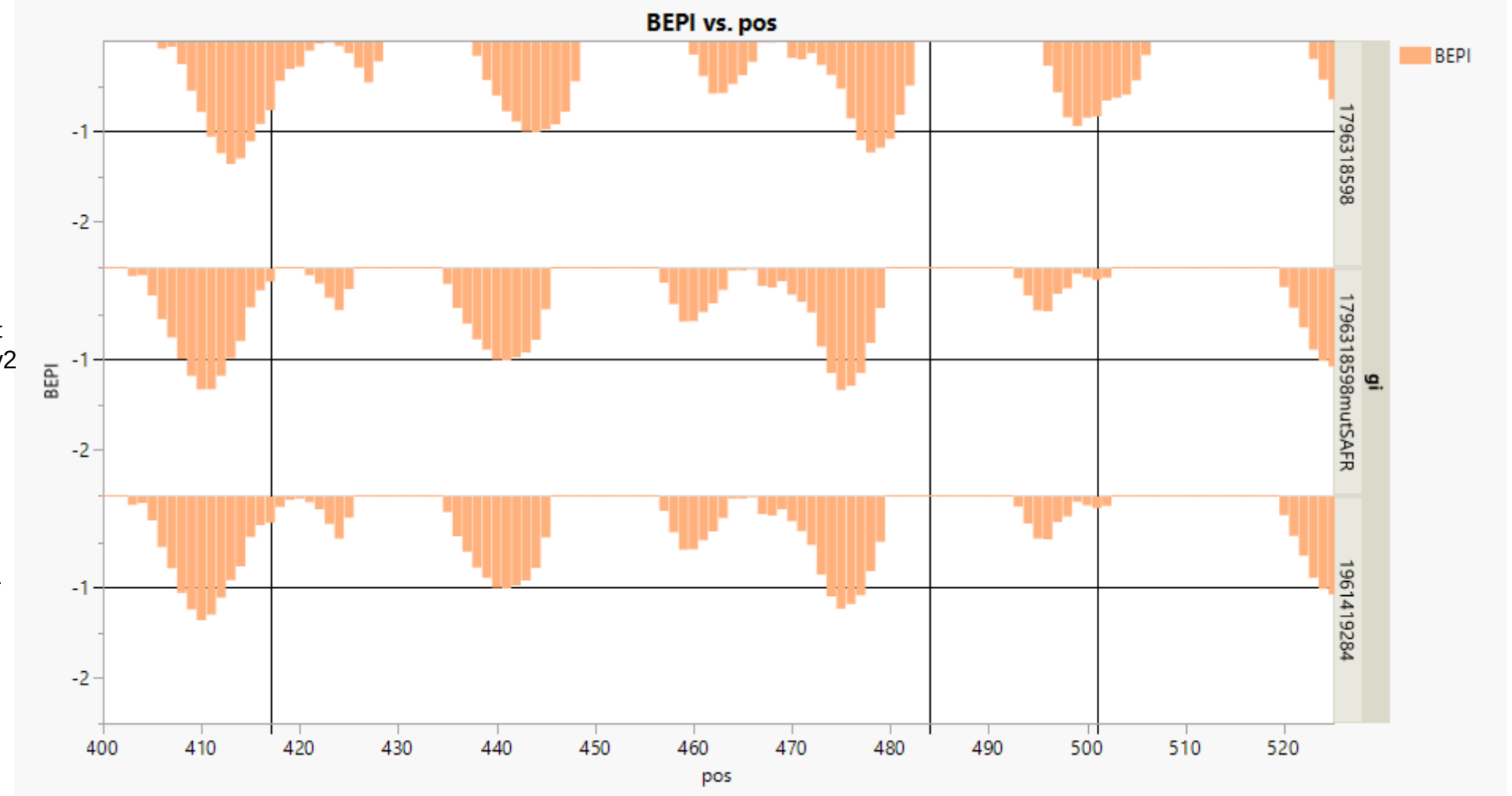
Green boxes show changes relative to positions 417, 484, and 501

B cell linear epitopes

Reference sequence

South African variant
B.1351 20H/501Y.v2

UK variant
B.1.1.7 20I/501Y.v1



Lines indicate positions 417, 484, and 501

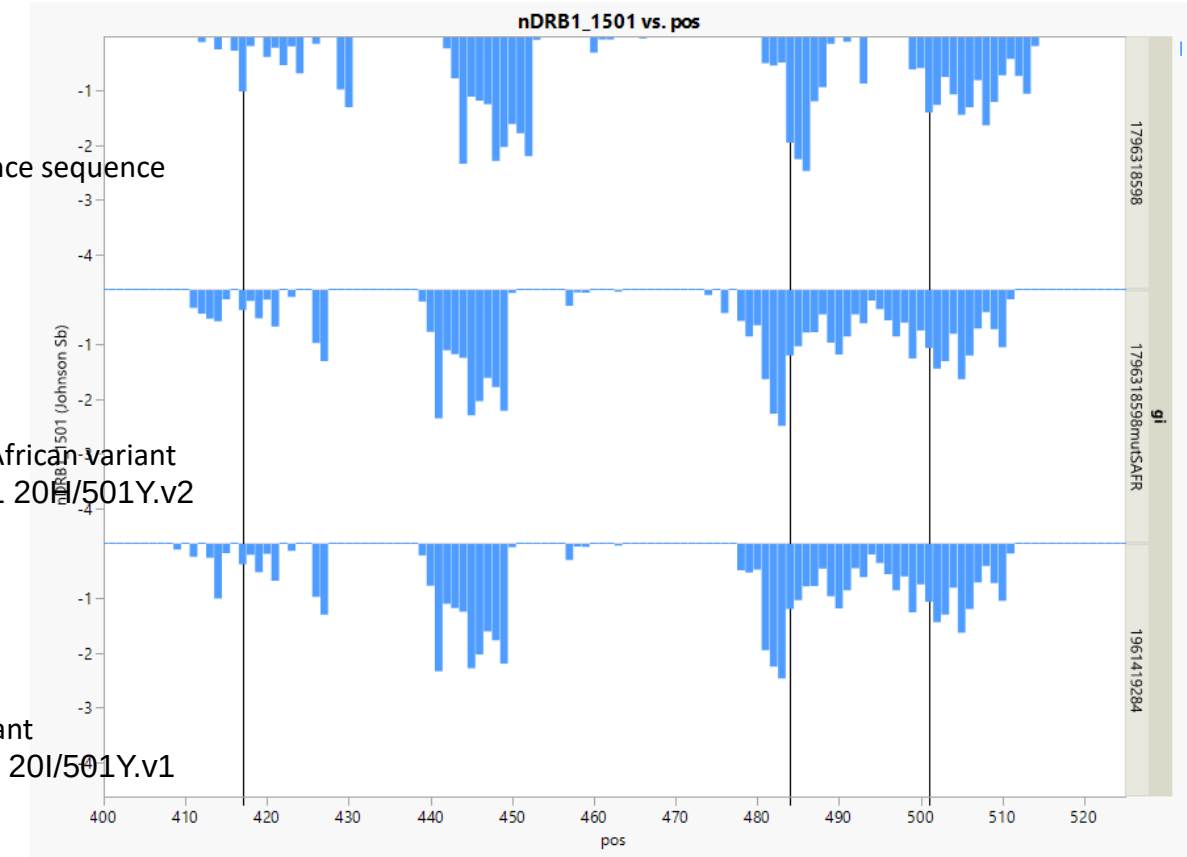
With the deletion of amino acids 69-70 the UK and S African sequences are displaced relative the Reference Sequence

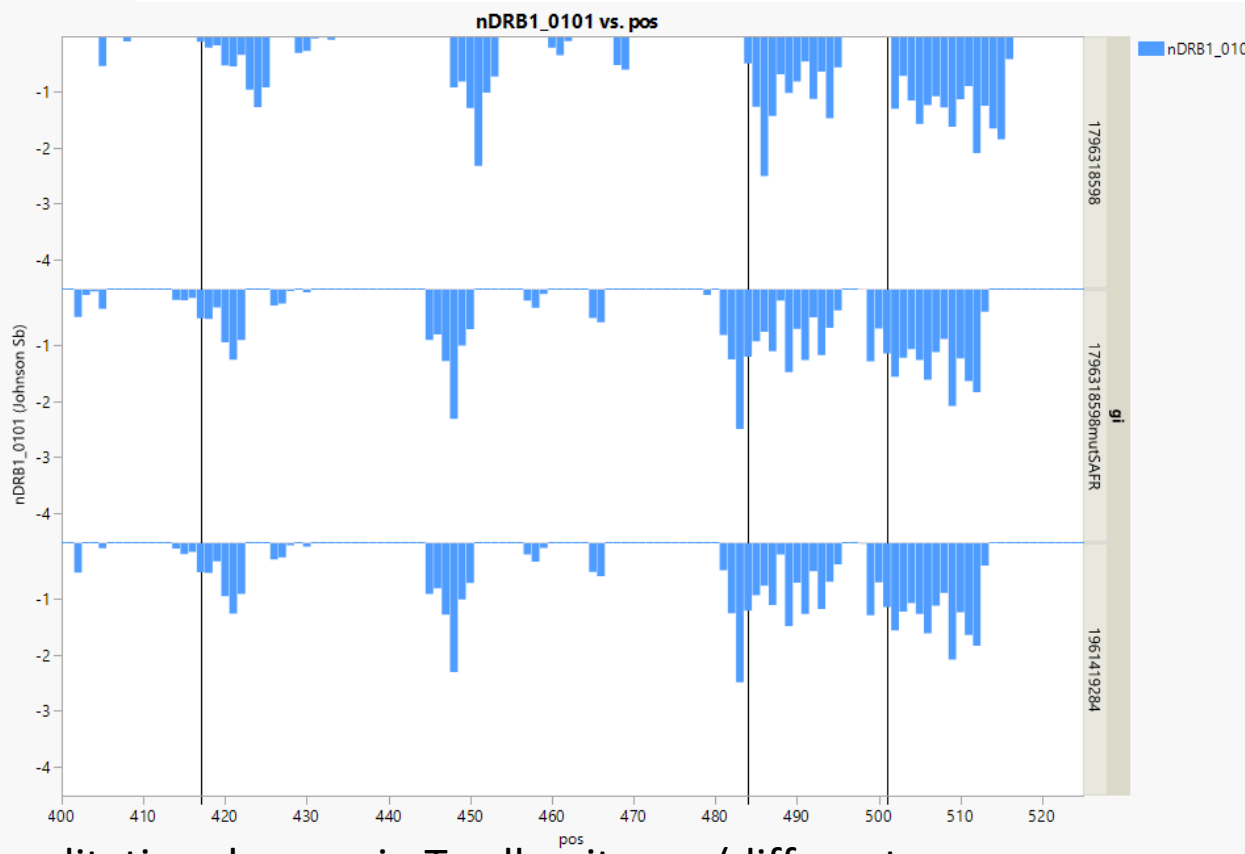
Examples of MHC II alleles

Reference sequence

South African variant
B.1351 20H/501Y.v2

UK variant
B.1.1.7 20I/501Y.v1





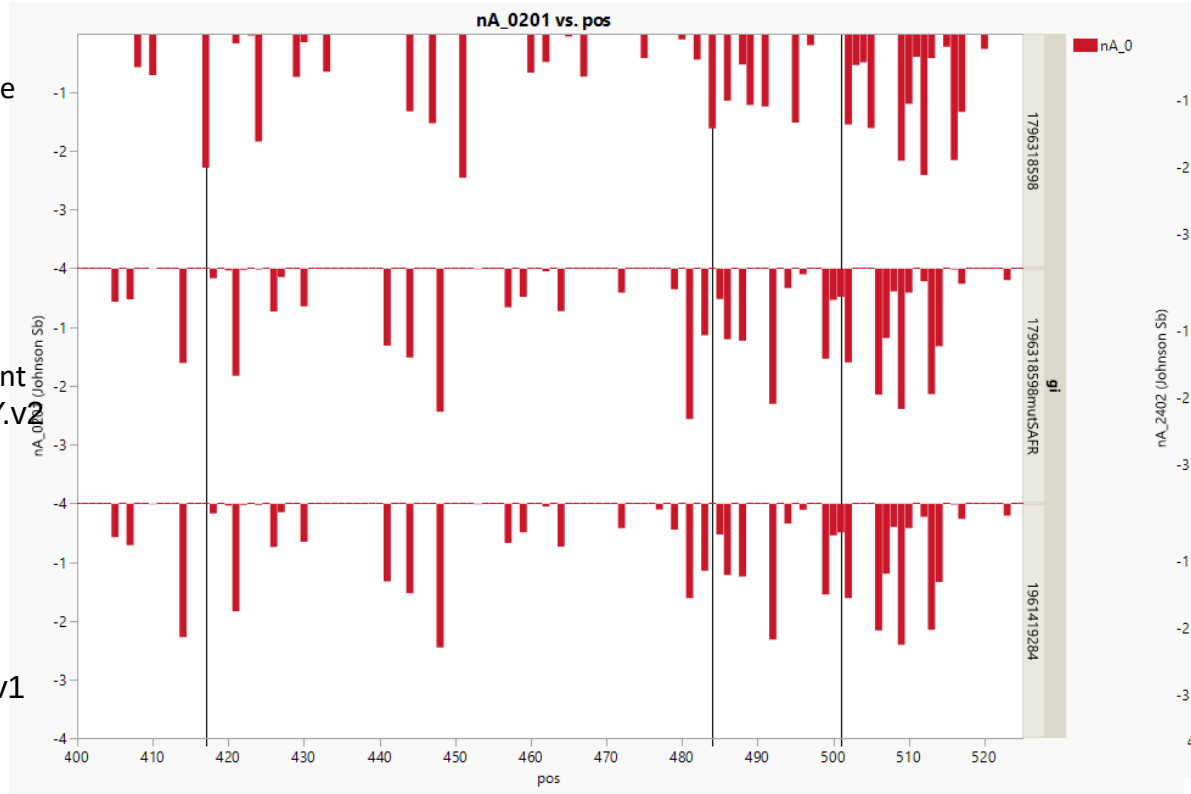
There is little difference in binding affinity; there are qualitative changes in T cell epitopes (different cognate T cells clones) arising from the amino acid differences but the 484 is in a low binding position and has essentially no impact on CD8 or CD4

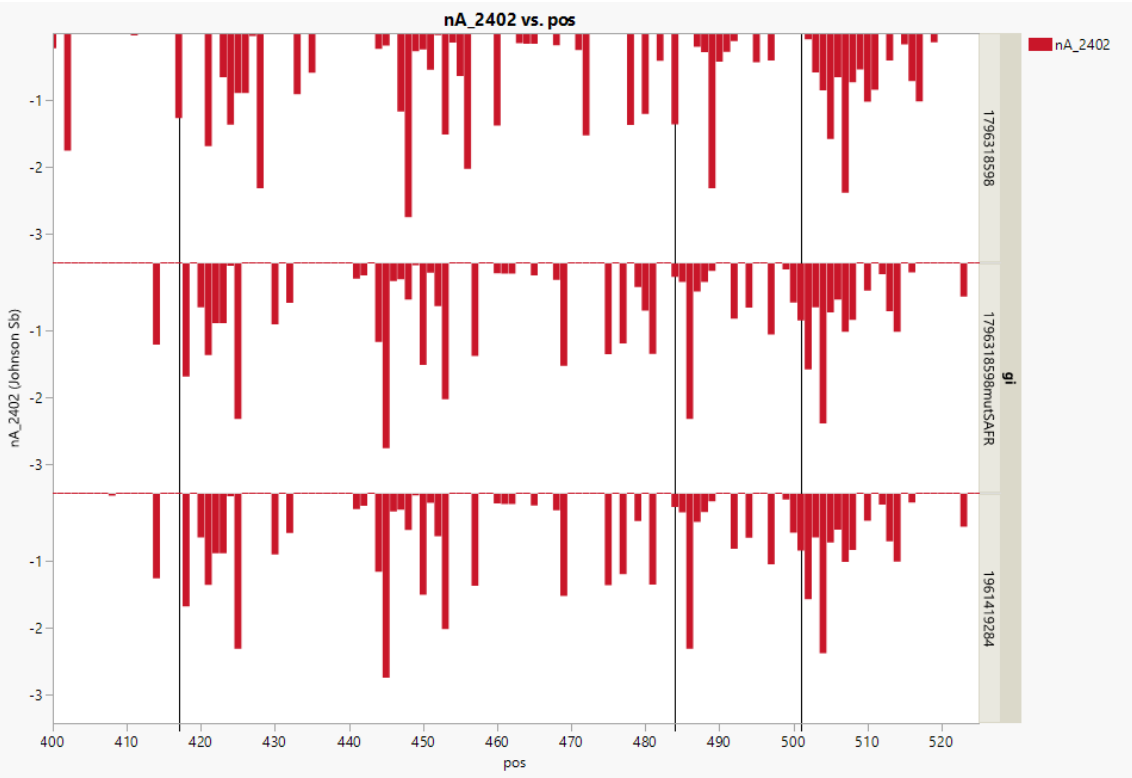
Example MHC I alleles

Reference sequence

South African variant
B.1351 20H/501Y.v2

UK variant
B.1.1.7 20I/501Y.v1





There is little difference in binding affinity; there are qualitative changes in T cell epitopes (different cognate T cells clones) arising from the amino acid differences but the 484 is in a low binding position and has essentially no impact on CD8 or CD4

Following slides…

- Show individual sequential peptide bidning affinity around the mutations. These are screen grabs of JMP files

Observations

- 501Y has more impact on MHC II binding – the Y is higher binding
- Plot numbering adjusted to place all at same position despite the 69-70 deletion
- BEPI at this point not strong/dominant but diminished probability - -0.94 to -0.47 – But has higher associated CD4 help
- Impact on MHC I is likewise small but motifs qualitatively different

gi	pos	peptide (15-mer)	aa_1	TCEM I	TCEM IIa	nDRB1_0101 ...	nDRB1_0301 ...	nDRB1_040...	nDRB1_0404 ...	nDRB1_0405 ...	nDRB1_0701 ...	nDRB1_0801 ...	nDRB1_0802 ...	nDRB1_0901 ...	nDRB1_1001 ...	nDRB1_1101 ...	nDRB1_1201 ...	nDRB1_1301 ...	nDRB1_1302 ...	nDRB1_1454 ...	nDRB1_1501 ...
1796318598	490	FPLQSYGFQPTN...	F	~~~QSYGF~	SY~F~PT	-0.81	1.36	-0.92	-1.32	-0.92	-0.77	-1.31	-1.09	-1.35	0.98	-0.31	0.88	-0.43	0.37	-1.09	0.33
1796318598mu...	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	-1.10	0.36	-1.45	-1.12	-1.01	-1.40	-1.37	-0.88	-1.72	0.81	-0.61	-0.80	-1.03	0.73	-1.68	-0.77
1961419284	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	-1.11	0.37	-1.45	-1.11	-1.00	-1.40	-1.37	-0.87	-1.71	0.79	-0.60	-0.79	-1.02	0.72	-1.66	-0.77
1796318598	491	PLQSYGFQPTNG...	P	~~~SYGFQ~	YG~Q~TN	-0.45	2.17	-0.64	-0.83	-0.29	-0.25	-0.23	-0.92	-1.10	-0.59	-0.62	-1.50	0.52	0.37	0.06	-0.10
1796318598mu...	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	-0.19	1.32	-0.82	-0.79	0.16	-0.34	-0.52	-1.08	-0.77	-0.80	-0.51	-2.15	0.05	0.65	-0.62	-0.45
1961419284	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	-0.20	1.33	-0.82	-0.78	0.16	-0.35	-0.51	-1.07	-0.77	-0.80	-0.50	-2.15	0.06	0.64	-0.62	-0.45
1796318598	492	LQSYGFQPTNGV...	L	~~~YGFQP~	GF~P~NG	-1.12	0.77	-0.51	-0.63	-0.12	0.51	-0.77	-0.46	-0.38	1.03	-0.71	1.12	0.05	1.44	-0.90	-0.01
1796318598mu...	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	-1.48	0.88	-0.36	-1.03	-0.24	0.08	-1.10	-0.64	-0.55	0.64	-0.92	-0.28	-0.41	0.92	-1.18	-0.96
1961419284	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	-1.48	0.89	-0.35	-1.03	-0.24	0.07	-1.10	-0.64	-0.55	0.63	-0.91	-0.28	-0.40	0.91	-1.17	-0.96
1796318598	493	QSYGFQPTNGVG...	Q	~~~GFQPT~	FQ~T~GV	-0.63	0.27	-1.18	-0.52	-0.13	-0.57	0.21	-0.65	-1.23	-0.51	-0.02	0.15	0.52	-0.55	0.79	-0.86
1796318598mu...	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	-0.70	0.59	-0.73	-1.12	-0.60	-0.93	-0.20	-1.26	-0.96	-0.60	-0.63	-0.42	0.13	-0.64	0.36	-1.18
1961419284	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	-0.71	0.60	-0.72	-1.11	-0.60	-0.94	-0.20	-1.25	-0.96	-0.60	-0.62	-0.42	0.13	-0.64	0.35	-1.18
1796318598	494	SYGFQPTNGVGY...	S	~~~FQPTN~	QP~N~VG	-1.47	-0.18	-0.78	-0.67	0.68	-0.35	0.41	0.01	-0.63	0.18	-0.16	-0.11	-0.21	-0.68	-0.17	0.09
1796318598mu...	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	-1.26	-0.00	-1.12	-1.10	-0.15	-0.74	-0.71	-0.58	-1.05	-0.02	-0.85	-1.15	-0.49	-0.88	-1.07	-0.85
1961419284	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	-1.26	0.01	-1.12	-1.09	-0.15	-0.75	-0.70	-0.58	-1.05	-0.02	-0.84	-1.13	-0.49	-0.87	-1.07	-0.85
1796318598	495	YGFQPTNGVGYQ...	Y	~~~QPTNG~	PT~G~GY	-0.55	0.21	-0.76	0.02	-0.32	-0.13	0.57	0.23	-0.03	0.30	0.23	0.43	-0.35	0.37	-0.06	0.11
1796318598mu...	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	-0.49	0.60	-0.86	-0.49	-0.62	-0.50	-0.04	-0.86	-0.36	0.28	-0.25	-0.27	-0.61	0.18	-0.31	-0.44
1961419284	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	-0.50	0.62	-0.86	-0.48	-0.62	-0.52	-0.04	-0.86	-0.36	0.28	-0.24	-0.27	-0.60	0.18	-0.31	-0.44
1796318598	496	GFQPTNGVGYQP...	G	~~~PTNGV~	TN~V~YQ	0.11	-0.91	0.26	-0.13	-0.05	-0.13	-0.07	0.87	0.35	-1.30	0.60	0.75	-0.84	-0.54	0.22	0.60
1796318598mu...	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	-1.17	-1.13	-0.10	-1.01	-0.31	-0.62	-0.73	0.42	-0.59	-1.31	-0.55	-0.32	-1.40	-0.86	-0.41	-0.60
1961419284	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	-1.18	-1.12	-0.10	-1.00	-0.31	-0.63	-0.73	0.41	-0.59	-1.32	-0.54	-0.32	-1.39	-0.86	-0.41	-0.61
1796318598	497	FQPTNGVGYQPY...	F	~~~TNGVG~	NG~G~QP	0.39	0.32	0.36	0.03	0.35	-0.66	-0.06	1.46	-0.06	-0.55	0.74	-0.77	-0.12	0.21	-0.63	0.65
1796318598mu...	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	-0.68	-0.72	-0.53	-0.42	-0.30	-1.38	-0.75	0.87	-0.93	-0.90	-0.24	-1.09	-0.81	-0.91	-1.32	-0.19
1961419284	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	-0.69	-0.71	-0.53	-0.42	-0.31	-1.39	-0.75	0.86	-0.93	-0.90	-0.23	-1.08	-0.80	-0.90	-1.31	-0.19
1796318598	498	QPTNGVGYQPYR...	Q	~~~NGVGY~	GV~Y~PY	0.68	0.04	1.75	1.11	0.28	-0.38	-0.09	1.27	0.83	0.48	0.15	0.14	-0.37	0.28	-0.61	0.36
1796318598mu...	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	-0.37	-1.04	0.39	0.09	-0.56	-0.96	-0.96	0.25	-0.74	0.63	-0.46	-0.54	-1.11	-0.63	-1.41	-0.34
1961419284	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	-0.38	-1.03	0.38	0.08	-0.56	-0.97	-0.95	0.24	-0.74	0.62	-0.45	-0.54	-1.10	-0.63	-1.40	-0.35
1796318598	499	PTNGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	0.69	-0.01	0.14	0.10	0.05	-0.23	0.18	0.74	0.10	-0.29	0.90	-0.65	0.09	0.32	-0.54	-0.61
1796318598mu...	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	0.32	-0.95	-0.07	-0.71	-0.12	-0.98	-0.22	0.70	-0.75	-1.13	0.09	-1.22	-0.22	-0.39	-1.43	-0.55
1961419284	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	0.31	-0.94	-0.07	-0.71	-0.12	-0.99	-0.21	0.69	-0.75	-1.13	0.10	-1.20	-0.22	-0.39	-1.42	-0.55
1796318598	500	TNGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	0.44	0.65	1.35	0.25	0.13	0.21	-0.22	-0.05	1.02	0.71	0.13	0.38	-0.62	0.54	-0.86	-0.58
1796318598mu...	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	0.22	-0.14	0.77	-0.32	-0.19	0.08	-0.56	0.17	0.57	0.67	-0.66	-0.00	-0.78	0.76	-1.16	-0.85
1961419284	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	0.21	-0.12	0.75	-0.32	-0.19	0.08	-0.56	0.16	0.56	0.66	-0.65	-0.01	-0.77	0.75	-1.16	-0.85
1796318598	501	NGVGYQPYRVVV...	N	~~~GYQPY~	YQ~Y~VV	0.28	-0.18	-0.08	0.30	-0.64	-1.24	-0.67	-1.21	-0.33	-1.08	-0.61	-1.60	-0.22	-0.48	-0.99	-1.39
1796318598mu...	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	0.00	-0.51	-0.32	-0.27	-1.21	-1.42	-0.96	-1.11	-1.14	-1.12	-0.77	-1.06	-0.75	-0.19	-1.40	-0.59
1961419284	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	-0.00	-0.50	-0.32	-0.27	-1.20	-1.42	-0.95	-1.10	-1.13	-1.12	-0.75	-1.04	-0.74	-0.19	-1.39	-0.59
1796318598	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	-1.30	1.22	-0.00	-0.64	-1.27	-0.58	0.02	-0.04	-0.48	-0.30	-0.34	-1.27	0.03	-0.37	-0.32	-1.26
1796318598mu...	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	-1.28	1.23	0.01	-0.63	-1.26	-0.56	0.03	-0.03	-0.46	-0.29	-0.34	-1.27	0.04	-0.36	-0.30	-1.25
1961419284	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	-1.29	1.24	0.01	-0.62	-1.25	-0.57	0.03	-0.04	-0.46	-0.29	-0.33	-1.25	0.04	-0.36	-0.30	-1.25

gi	pos	peptide (15-mer)	aa_1	TCEM I	TCEM IIa	9-mer	nA_0101 (Johns...	nA_0201 (Johns...	nA_0202 (Johns...	nA_0203 (Johns...	nA_0206 (Johns...	nA_0211 (Johns...	nA_0212 (Johns...	nA_0216 (Johns...	nA_0217 (Johns...	nA_0219 (Johns...	nA_0250 (Johns...	nA_0301 (Johns...	nA_0801 (Johns...
1796318598	490	FPLQSYGFQPTN...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	0.21	0.07	-0.28	-0.09	-0.45	0.43	0.35	-0.39	0.77	0.29	0.68	-0.27	0.69
1796318598mu...	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	0.22	0.07	-0.27	-0.09	-0.44	0.44	0.36	-0.39	0.77	0.31	0.69	-0.26	0.69
1961419284	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	0.22	0.07	-0.27	-0.09	-0.44	0.43	0.36	-0.38	0.77	0.30	0.69	-0.26	0.69
1796318598	491	PLQSYGFQPTNG...	P	~~~SYGFQ~	YG~Q~TN	PLQSYGFQP	-0.89	-1.24	-0.91	-0.67	-0.08	-1.45	-1.59	-0.89	-1.26	-1.04	-0.57	-1.57	-0.86
1796318598mu...	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	PLQSYGFQP	-0.88	-1.23	-0.90	-0.67	-0.07	-1.44	-1.59	-0.88	-1.26	-1.04	-0.57	-1.56	-0.87
1961419284	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	PLQSYGFQP	-0.89	-1.24	-0.89	-0.67	-0.08	-1.44	-1.58	-0.88	-1.26	-1.04	-0.58	-1.56	-0.86
1796318598	492	LQSYGFQPTNGV...	L	~~~YGFQP~	GF~P~NG	LQSYGFQPT	-0.36	0.50	-0.09	0.23	-0.60	0.27	0.59	0.51	0.08	0.81	-0.42	0.95	0.19
1796318598mu...	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	LQSYGFQPT	-0.35	0.50	-0.08	0.24	-0.58	0.28	0.59	0.51	0.08	0.82	-0.42	0.95	0.20
1961419284	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	LQSYGFQPT	-0.35	0.51	-0.09	0.24	-0.59	0.27	0.60	0.51	0.09	0.82	-0.42	0.95	0.20
1796318598	493	QSYGFQPTNGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTN	0.44	0.41	1.27	1.07	0.61	0.37	0.69	0.83	-0.90	0.65	2.05	-0.68	0.57
1796318598mu...	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTY	-2.80	0.19	1.00	1.91	0.76	0.72	1.02	0.67	-1.62	0.95	1.90	-1.78	0.40
1961419284	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTY	-2.80	0.19	0.98	1.92	0.75	0.71	1.02	0.66	-1.62	0.95	1.94	-1.79	0.40
1796318598	494	SYGFQPTNGVGY...	S	~~~FQPTN~	QP~N~VG	SYGFQPTNG	1.55	1.64	-0.61	-1.20	-0.23	-0.24	-1.22	-1.70	2.79	-0.84	-1.44	0.34	0.02
1796318598mu...	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	SYGFQPTYG	0.86	1.36	-0.93	-1.27	-0.23	-0.30	-1.14	-1.57	2.36	-0.61	-1.48	0.23	0.99
1961419284	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	SYGFQPTYG	0.86	1.38	-0.91	-1.28	-0.24	-0.30	-1.13	-1.58	2.38	-0.62	-1.48	0.23	0.99
1796318598	495	YGFQPTNGVGYQ...	Y	~~~QPTNG~	PT~G~GY	YGFQPTNGV	-0.04	-1.52	0.00	-0.81	-1.48	-0.68	-0.26	-0.51	-1.50	-1.09	-0.12	0.70	-1.43
1796318598mu...	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	YGFQPTYGV	-0.28	-2.30	-0.01	-0.52	-1.96	-0.83	-0.69	-0.48	-1.66	-1.45	-0.05	0.74	-1.98
1961419284	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	YGFQPTYGV	-0.28	-2.31	-0.02	-0.52	-1.97	-0.82	-0.68	-0.48	-1.66	-1.45	-0.05	0.74	-1.98
1796318598	496	GFQPTNGVGYQP...	G	~~~PTNGV~	TN~V~YQ	GFQPTNGVG	1.72	1.95	-0.71	-1.35	0.22	-1.51	-1.87	-1.87	1.27	-1.47	-1.67	0.94	-0.34
1796318598mu...	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	GFQPTYGVG	1.17	1.36	-1.21	-1.90	-0.68	-1.62	-1.84	-2.05	1.43	-1.59	-1.81	0.96	-0.27
1961419284	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	GFQPTYGVG	1.17	1.38	-1.20	-1.92	-0.68	-1.62	-1.83	-2.07	1.44	-1.60	-1.81	0.96	-0.27
1796318598	497	FQPTNGVGYQPY...	F	~~~TNGVG~	NG~G~QP	FQPTNGVGY	-0.54	-0.19	0.08	0.13	-0.16	0.14	0.12	1.53	0.03	0.55	0.71	0.70	0.34
1796318598mu...	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	FQPTYGVGY	-0.76	-0.32	-0.18	0.12	-0.25	0.52	0.54	1.77	-0.21	0.53	0.73	0.32	0.09
1961419284	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	FQPTYGVGY	-0.76	-0.33	-0.19	0.12	-0.25	0.52	0.54	1.76	-0.21	0.52	0.73	0.32	0.10
1796318598	498	QPTNGVGYQPYR...	Q	~~~NGVGY~	GV~Y~PY	QPTNGVGYQ	0.65	0.32	0.41	0.03	0.44	0.68	0.06	0.35	0.56	0.63	1.37	0.21	1.24
1796318598mu...	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	QPTYGVGYQ	0.63	0.36	0.49	0.39	0.91	0.86	0.21	0.68	0.65	0.75	1.29	0.04	1.40
1961419284	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	QPTYGVGYQ	0.63	0.36	0.46	0.38	0.90	0.85	0.21	0.67	0.66	0.75	1.31	0.04	1.39
1796318598	499	PTNGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTNGVGYQP	-1.13	0.47	0.61	2.20	1.18	0.18	0.92	-0.08	-0.32	0.34	0.35	-0.28	0.37
1796318598mu...	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTYGVGYQP	-0.57	-0.09	0.30	1.95	0.37	-0.22	0.63	-0.05	-0.45	-0.05	0.32	-1.51	-0.19
1961419284	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTYGVGYQP	-0.57	-0.09	0.28	1.95	0.37	-0.21	0.64	-0.05	-0.44	-0.05	0.32	-1.52	-0.19
1796318598	500	TNGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TNGVGYQPY	-1.23	0.64	0.35	1.44	0.35	0.80	0.79	0.49	0.07	1.48	0.99	-0.13	2.31
1796318598mu...	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TYGVGYQPY	0.54	0.57	-0.86	0.14	0.16	0.89	0.35	0.44	0.07	0.46	-0.15	0.28	1.85
1961419284	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TYGVGYQPY	0.54	0.58	-0.84	0.14	0.16	0.87	0.35	0.44	0.08	0.46	-0.16	0.28	1.83
1796318598	501	NGVGYQPYRVVV...	N	~~~GYQPY~	YQ~Y~VV	NGVGYQPYR	0.60	0.34	0.82	1.01	1.09	1.47	0.39	1.43	-1.11	0.79	3.03	-0.09	1.95
1796318598mu...	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	YGVGYQPYR	0.03	0.28	0.59	0.91	0.50	1.55	0.32	1.72	-1.55	0.91	2.67	-0.24	1.64
1961419284	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	YGVGYQPYR	0.03	0.28	0.57	0.91	0.49	1.52	0.32	1.70	-1.54	0.91	2.76	-0.25	1.63
1796318598	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	-0.26	-1.55	0.12	-0.90	-0.49	-1.66	-1.51	-2.11	-0.51	-1.44	-1.03	-0.09	-0.35
1796318598mu...	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	-0.25	-1.53	0.13	-0.90	-0.48	-1.66	-1.51	-2.11	-0.51	-1.44	-1.03	-0.08	-0.35
1961419284	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	-0.25	-1.54	0.11	-0.90	-0.48	-1.66	-1.51	-2.12	-0.50	-1.44	-1.04	-0.08	-0.35

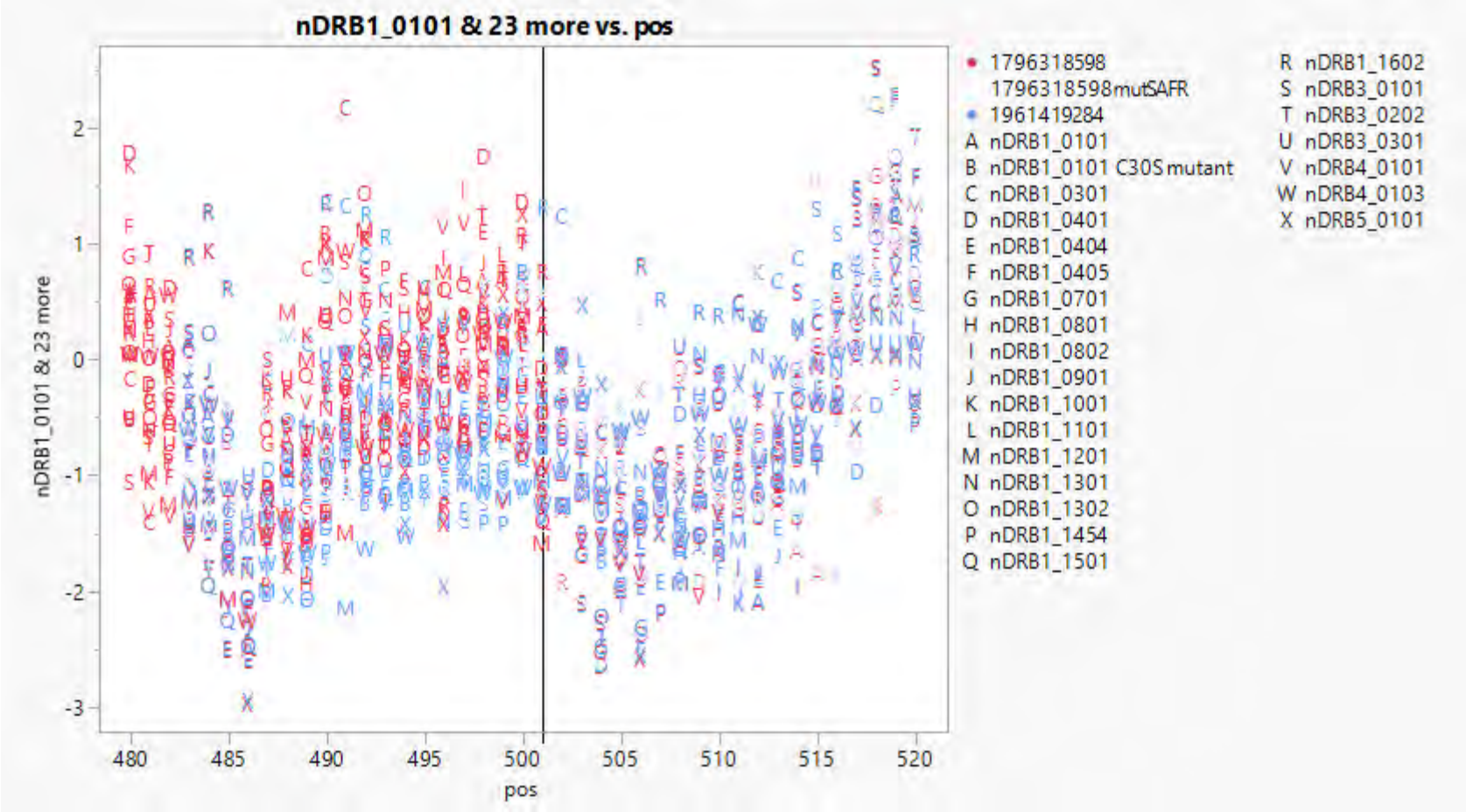
gi	pos	peptide (15-mer)	aa_1	TCEM I	TCEM IIa	9-mer	(imputed) log2 ...	(imputed) log2 ...	(imputed) log2 ...	(imputed) log2 ...
1796318598mu...	488	CYFPLQSYGFQPT...	C	~~~PLQSY~	LQ~Y~FQ	CYFPLQSYG	-0.54	-1.35	-0.61	-1.08
1961419284	488	CYFPLQSYGFQPT...	C	~~~PLQSY~	LQ~Y~FQ	CYFPLQSYG	-0.54	-1.35	-0.61	-1.08
1796318598	489	YFPLQSYGFQPTN...	Y	~~~LQSYG~	QS~G~QP	YFPLQSYGF	-0.05	-0.43	-1.54	-0.02
1796318598mu...	489	YFPLQSYGFQPTY...	Y	~~~LQSYG~	QS~G~QP	YFPLQSYGF	-0.05	-0.43	-1.54	-0.02
1961419284	489	YFPLQSYGFQPTY...	Y	~~~LQSYG~	QS~G~QP	YFPLQSYGF	-0.05	-0.43	-1.54	-0.02
1796318598	490	FPLQSYGFQPTN...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	-0.89	-0.74	-3.66	-3.62
1796318598mu...	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	-0.89	-0.74	-3.66	-3.62
1961419284	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	-0.89	-0.74	-3.66	-3.62
1796318598	491	PLQSYGFQPTNG...	P	~~~SYGFQ~	YG~Q~TN	PLQSYGFQP	-1.71	-1.02	-2.45	-3.62
1796318598mu...	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	PLQSYGFQP	-1.71	-1.13	-2.45	-0.61
1961419284	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	PLQSYGFQP	-1.71	-1.13	-2.45	-0.61
1796318598	492	LQSYGFQPTNGV...	L	~~~YGFQP~	GF~P~NG	LQSYGFQPT	-1.46	-0.56	-1.76	-0.77
1796318598mu...	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	LQSYGFQPT	-1.46	-0.43	-1.76	-0.10
1961419284	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	LQSYGFQPT	-1.46	-0.43	-1.76	-0.10
1796318598	493	QSYGFQPTNGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTN	-1.13	-0.71	-2.04	-1.54
1796318598mu...	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTY	-1.13	-0.71	-2.04	-1.54
1961419284	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTY	-1.13	-0.71	-2.04	-1.54
1796318598	494	SYGFQPTNGVGY...	S	~~~FQPTN~	QP~N~VG	SYGFQPTNG	-2.50	-1.46	-1.21	-1.36
1796318598mu...	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	SYGFQPTYG	-1.71	-1.65	-0.11	-3.62
1961419284	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	SYGFQPTYG	-1.71	-1.65	-0.11	-3.62
1796318598	495	YGFQPTNGVGYQ...	Y	~~~QPTNG~	PT~G~GY	YGFQPTNGV	-2.33	0.19	-1.36	-0.77
1796318598mu...	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	YGFQPTYGV	-1.79	0.19	-3.66	-0.77
1961419284	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	YGFQPTYGV	-1.79	0.19	-3.66	-0.77
1796318598	496	GFQPTNGVGYQP...	G	~~~PTNGV~	TN~V~YQ	GFQPTNGVG	-0.77	-1.17	-0.86	-3.12
1796318598mu...	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	GFQPTYGVG	-0.50	-2.19	0.37	-1.08
1961419284	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	GFQPTYGVG	-0.50	-2.19	0.37	-1.08
1796318598	497	FQPTNGVGYQPY...	F	~~~TNGVG~	NG~G~QP	FQPTNGVGY	0.09	-0.20	-0.86	-1.75
1796318598mu...	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	FQPTYGVGY	1.55	-0.98	0.11	-0.61
1961419284	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	FQPTYGVGY	1.55	-0.98	0.11	-0.61
1796318598	498	QPTNGVGYQPYR...	Q	~~~NGVGY~	GV~Y~PY	QPTNGVGYQ	0.43	-0.27	-0.77	-2.44
1796318598mu...	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	QPTYGVGYQ	0.05	-0.27	-3.66	-2.44
1961419284	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	QPTYGVGYQ	0.05	-0.27	-3.66	-2.44
1796318598	499	PTNGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTNGVGYQP	0.14	-0.66	-1.54	-0.77
1796318598mu...	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTYGVGYQP	0.14	-0.66	-1.54	-0.77
1961419284	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTYGVGYQP	0.14	-0.66	-1.54	-0.77
1796318598	500	TNGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TNGVGYQPY	0.68	-0.56	-0.03	-2.04
1796318598mu...	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TYGVGYQPY	0.68	-0.56	-0.03	-2.04
1961419284	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TYGVGYQPY	0.68	-0.56	-0.03	-2.04
1796318598	501	NGVGYQPYRVVV...	N	~~~GYQPY~	YQ~Y~VV	NGVGYQPYR	-0.89	-0.68	-1.08	-3.62
1796318598mu...	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	YGVGYQPYR	-0.89	-0.68	-1.08	-3.62
1961419284	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	YGVGYQPYR	-0.89	-0.68	-1.08	-3.62
1796318598	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	-1.79	-0.19	-3.16	-1.21
1796318598mu...	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	-1.79	-0.19	-3.16	-1.21

Impact on TCEM frequency relative to that in the human proteome

Gi motifs followed by human

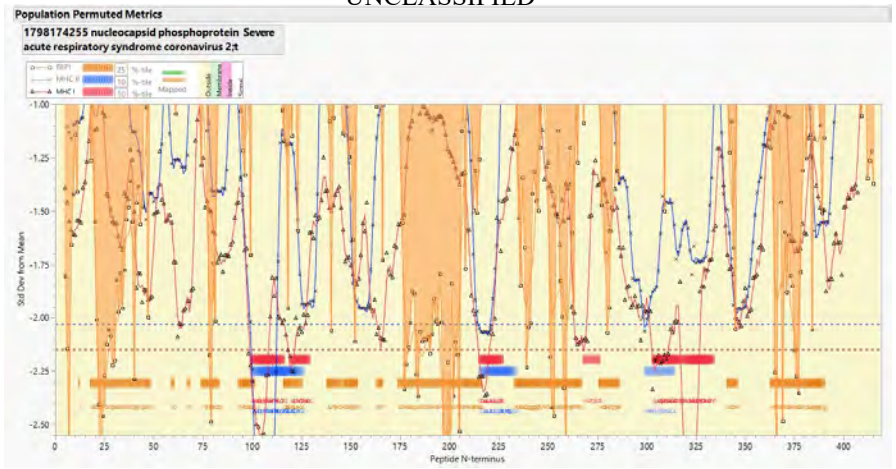
Y mutant has more rare motifs

gi	pos	peptide (15-mer)	aa_1	TCEM I	TCEM IIa	9-mer	imputed TCEM_I	imputed TCEM_IIa
1796318598mu...	488	CYFPLQSYGFQPT...	C	~~~PLQSY~	LQ~Y~FQ	CYFPLQSYG	18	22
1961419284	488	CYFPLQSYGFQPT...	C	~~~PLQSY~	LQ~Y~FQ	CYFPLQSYG	18	22
1796318598	489	YFPLQSYGFQPTN...	Y	~~~LQSYG~	QS~G~QP	YFPLQSYGF	14	22
1796318598mu...	489	YFPLQSYGFQPTY...	Y	~~~LQSYG~	QS~G~QP	YFPLQSYGF	14	22
1961419284	489	YFPLQSYGFQPTY...	Y	~~~LQSYG~	QS~G~QP	YFPLQSYGF	14	22
1796318598	490	FPLQSYGFQPTN...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	20	19
1796318598mu...	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	20	19
1961419284	490	FPLQSYGFQPTYG...	F	~~~QSYGF~	SY~F~PT	FPLQSYGFQ	20	19
1796318598	491	PLQSYGFQPTNG...	P	~~~SYGFQ~	YG~Q~TN	PLQSYGFQP	17	19
1796318598mu...	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	PLQSYGFQP	17	19
1961419284	491	PLQSYGFQPTYGV...	P	~~~SYGFQ~	YG~Q~TY	PLQSYGFQP	17	19
1796318598	492	LQSYGFQPTNGV...	L	~~~YGFQP~	GF~P~NG	LQSYGFQPT	21	14
1796318598mu...	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	LQSYGFQPT	21	16
1961419284	492	LQSYGFQPTYGV...	L	~~~YGFQP~	GF~P~YG	LQSYGFQPT	21	16
1796318598	493	QSYGFQPTNGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTN	21	20
1796318598mu...	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTY	21	20
1961419284	493	QSYGFQPTYGVG...	Q	~~~GFQPT~	FQ~T~GV	QSYGFQPTY	21	20
1796318598	494	SYGFQPTNGVGY...	S	~~~FQPTN~	QP~N~VG	SYGFQPTNG	23	23
1796318598mu...	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	SYGFQPTYG	22	22
1961419284	494	SYGFQPTYGVGY...	S	~~~FQPTY~	QP~Y~VG	SYGFQPTYG	22	22
1796318598	495	YGFQPTNGVGYQ...	Y	~~~QPTNG~	PT~G~GY	YGFQPTNGV	21	16
1796318598mu...	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	YGFQPTYGV	18	16
1961419284	495	YGFQPTYGVGYQ...	Y	~~~QPTYG~	PT~G~GY	YGFQPTYGV	18	16
1796318598	496	GFQPTNGVGYQP...	G	~~~PTNGV~	TN~V~YQ	GFQPTNGVG	19	16
1796318598mu...	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	GFQPTYGVG	16	14
1961419284	496	GFQPTYGVGYQP...	G	~~~PTYGV~	TY~V~YQ	GFQPTYGVG	16	14
1796318598	497	FQPTNGVGYQPY...	F	~~~TNGVG~	NG~G~QP	FQPTNGVGY	12	20
1796318598mu...	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	FQPTYGVGY	15	15
1961419284	497	FQPTYGVGYQPY...	F	~~~TYGVG~	YG~G~QP	FQPTYGVGY	15	15
1796318598	498	QPTNGVGYQPYR...	Q	~~~NGVGY~	GV~Y~PY	QPTNGVGYQ	16	15
1796318598mu...	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	QPTYGVGYQ	14	15
1961419284	498	QPTYGVGYQPYR...	Q	~~~YGVGY~	GV~Y~PY	QPTYGVGYQ	14	15
1796318598	499	PTNGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTNGVGYQP	18	21
1796318598mu...	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTYGVGYQP	18	21
1961419284	499	PTYGVGYQPYRV...	P	~~~GVGYQ~	VG~Q~YR	PTYGVGYQP	18	21
1796318598	500	TNGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TNGVGYQPY	19	18
1796318598mu...	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TYGVGYQPY	19	18
1961419284	500	TYGVGYQPYRVV...	T	~~~VGYQP~	GY~P~RV	TYGVGYQPY	19	18
1796318598	501	NGVGYQPYRVVV...	N	~~~GYQPY~	YQ~Y~VV	NGVGYQPYR	16	23
1796318598mu...	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	YGVGYQPYR	16	23
1961419284	501	YGVGYQPYRVVV...	Y	~~~GYQPY~	YQ~Y~VV	YGVGYQPYR	16	23
1796318598	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	22	13
1796318598mu...	502	GVGYQPYRVVVL...	G	~~~YQPYR~	QP~R~VV	GVGYQPYRV	22	13

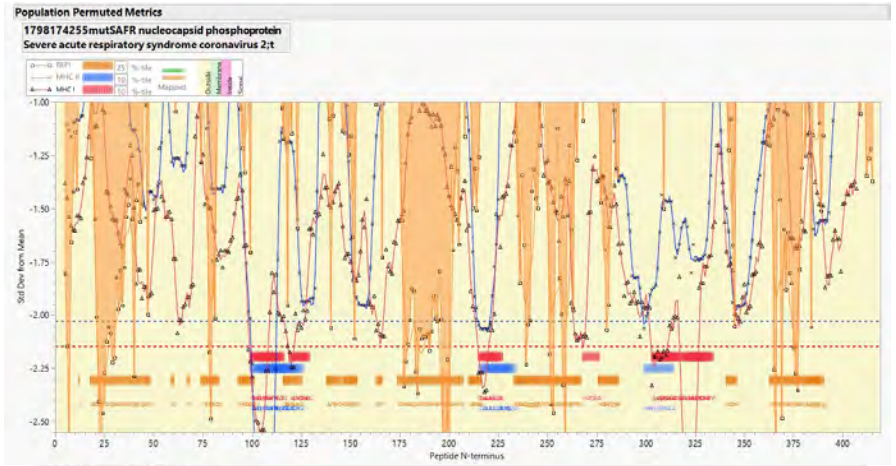

nDRB1_0101 & 23 more vs. pos
nDRB1_0101 & 23 more
pos
480
485
490
495
500
505
510
515
520
2
1
0
-1
-2
-3
1796318598
1796318598mutSAFR
1961419284
A nDRB1_0101
B nDRB1_0101 C30S mutant
C nDRB1_0301
D nDRB1_0401
E nDRB1_0404
F nDRB1_0405
G nDRB1_0701
H nDRB1_0801
I nDRB1_0802
J nDRB1_0901
K nDRB1_1001
L nDRB1_1101
M nDRB1_1201
N nDRB1_1301
O nDRB1_1302
P nDRB1_1454
Q nDRB1_1501
R nDRB1_1602
S nDRB3_0101
T nDRB3_0202
U nDRB3_0301
V nDRB4_0101
W nDRB4_0103
X nDRB5_0101

Nucleocapsid

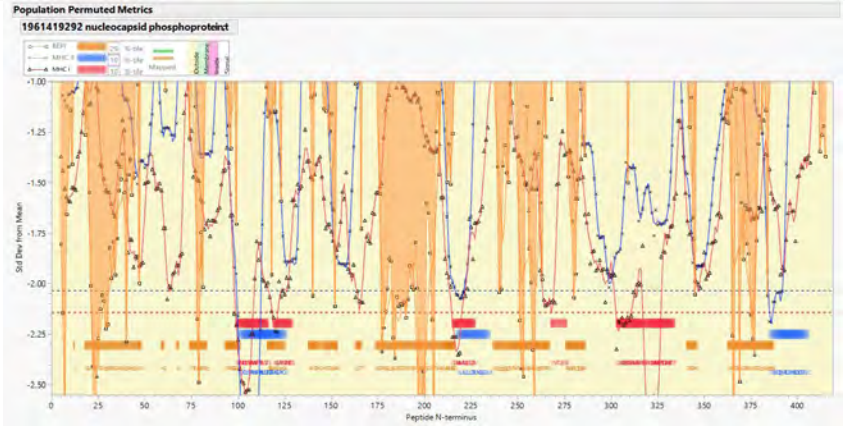
Ref



S Africa



UK

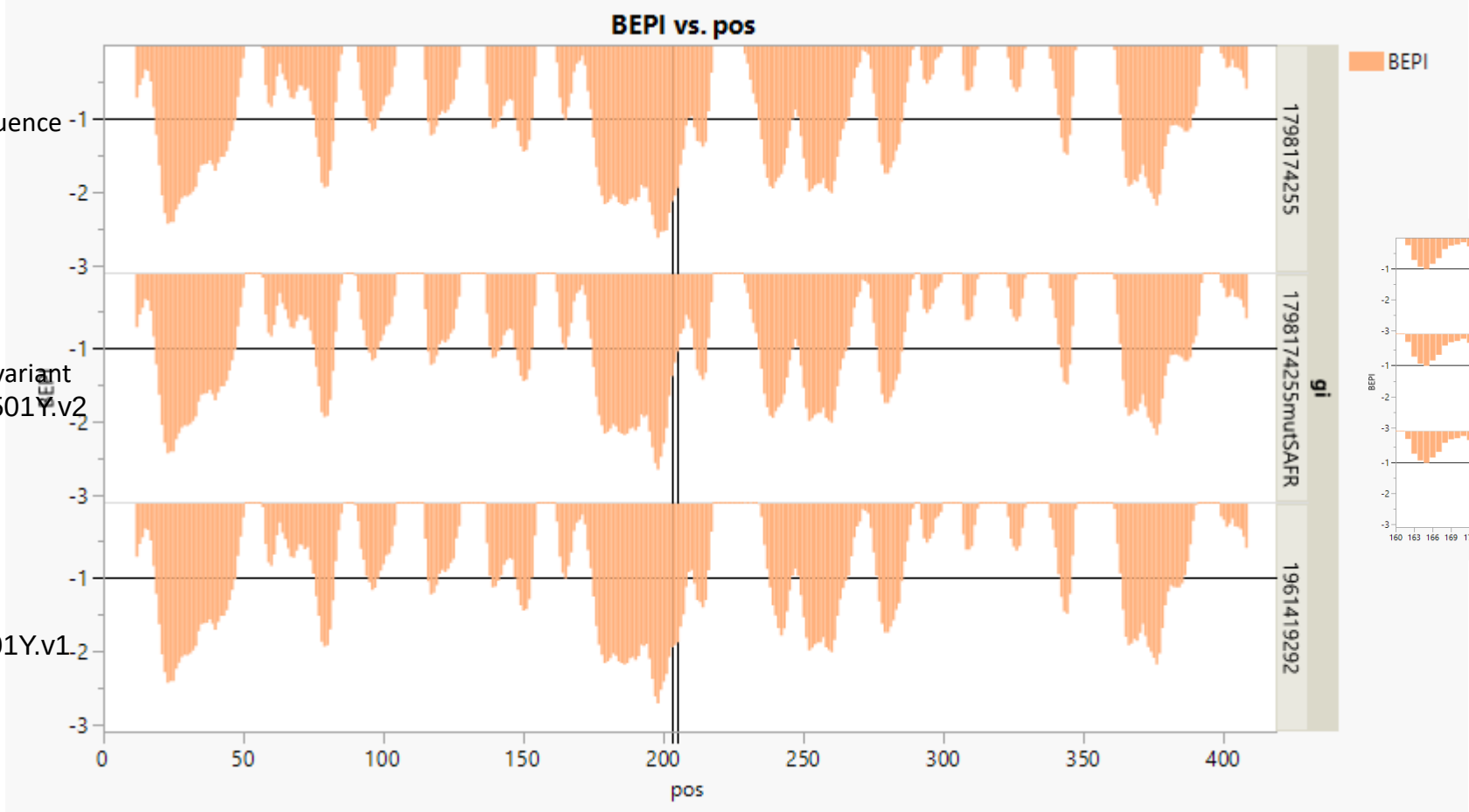


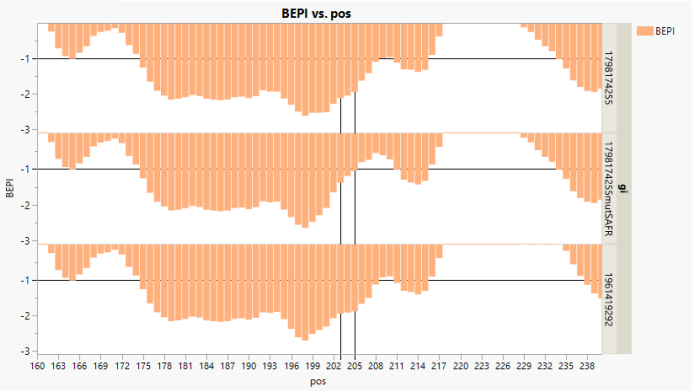
B cell linear epitopes

Reference sequence

South African variant
B.1351 20H/501Y.v2

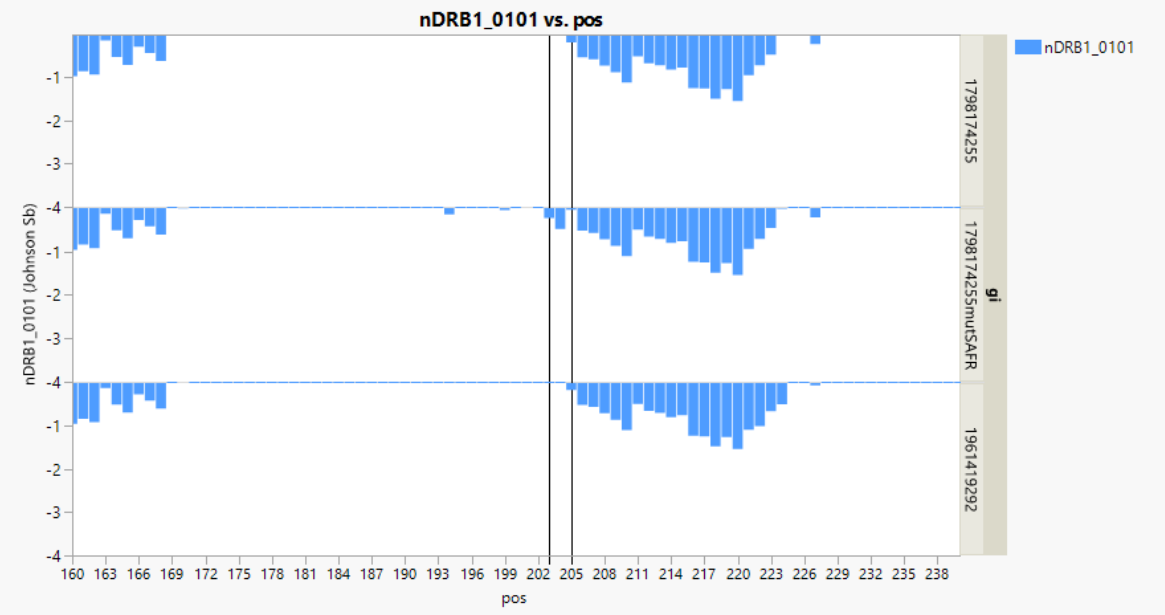
UK variant
B.1.1.7 20I/501Y.v1

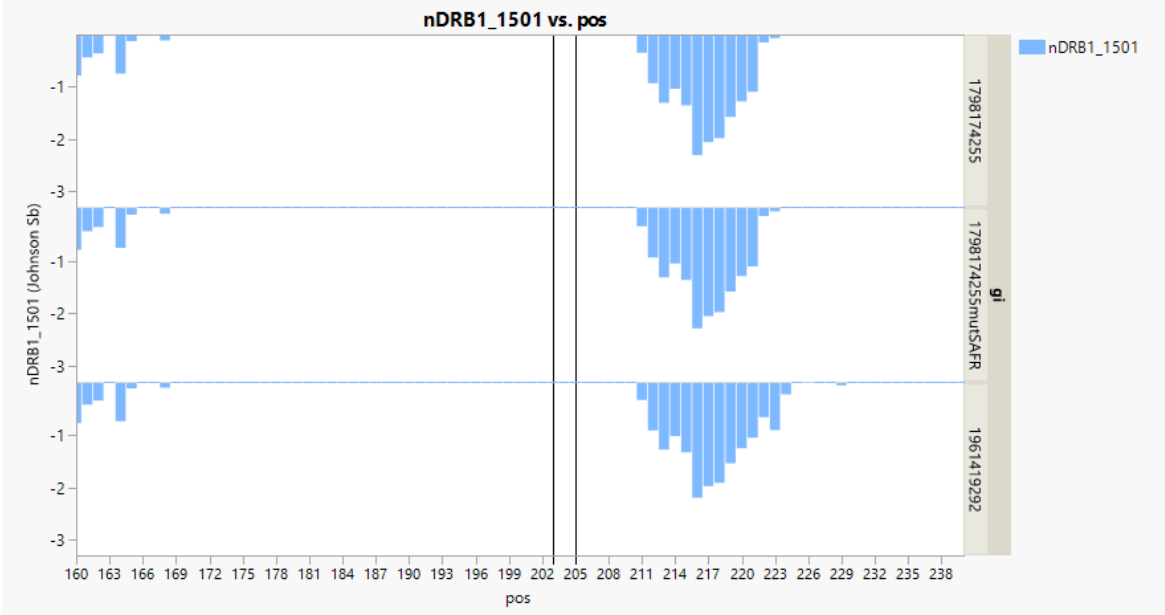


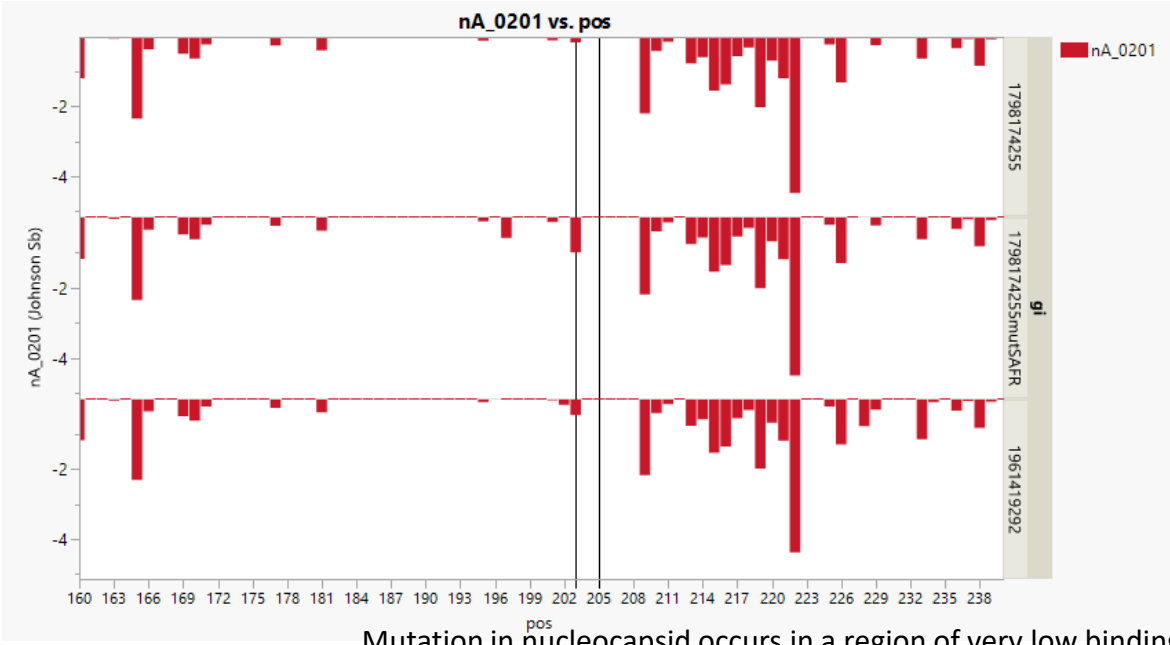


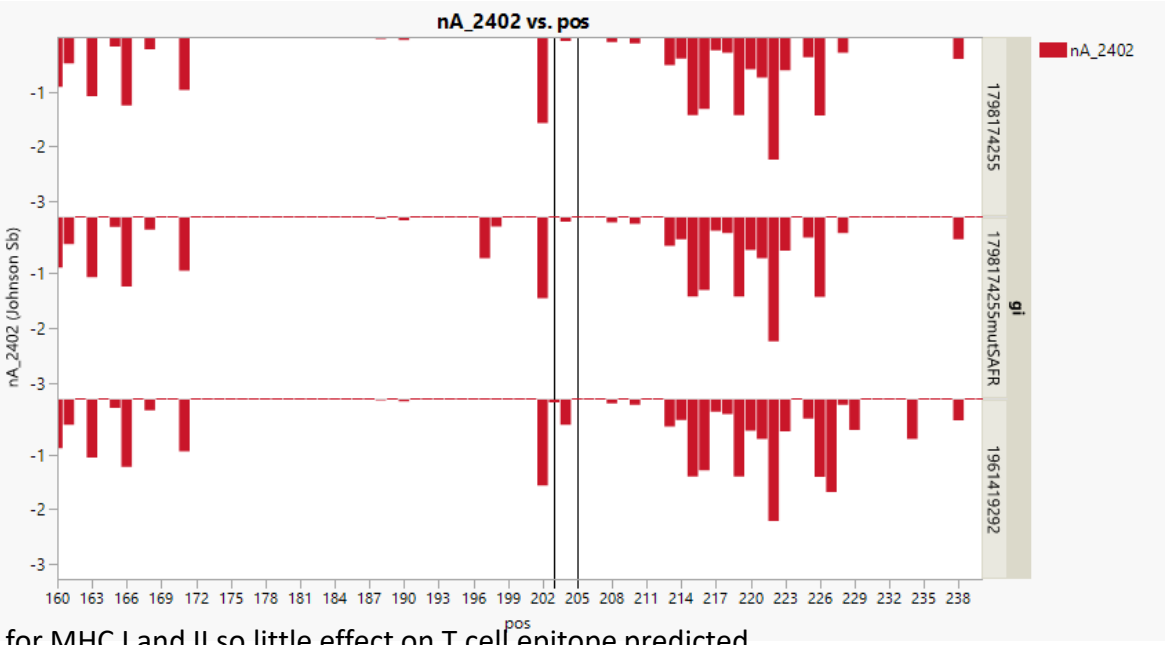
Mutation in nucleocapsid B cell epitope has very minor effect on probability but may change epitope conformation

Example MHC I and MHC II alleles









Mutation in nucleocapsid occurs in a region of very low binding for MHC I and II so little effect on T cell epitope predicted

Appendix 6
B cell epitope comparisons to Pepperprint Microarray on COVID patient

medRxiv preprint doi: https://doi.org/10.1101/2020.04.14.20059733. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Distinct early IgA profile may determine severity of COVID-19 symptoms: an immunological case series

Christine Dahlke[1,2,3*], Jasmin Heidepriem[4], Robin Kobbe[5], René Santer[5], Till Koch[1,2,3], Anahita Fathi[1,2,3], My L. Ly[1,2,3], Stefan Schmiedel[1,2], Peter H. Seeberger[4], ID-UKE COVID-19 study group, Marylyn M. Addo[1,2,3*], Felix F. Loeffler[4,6*]

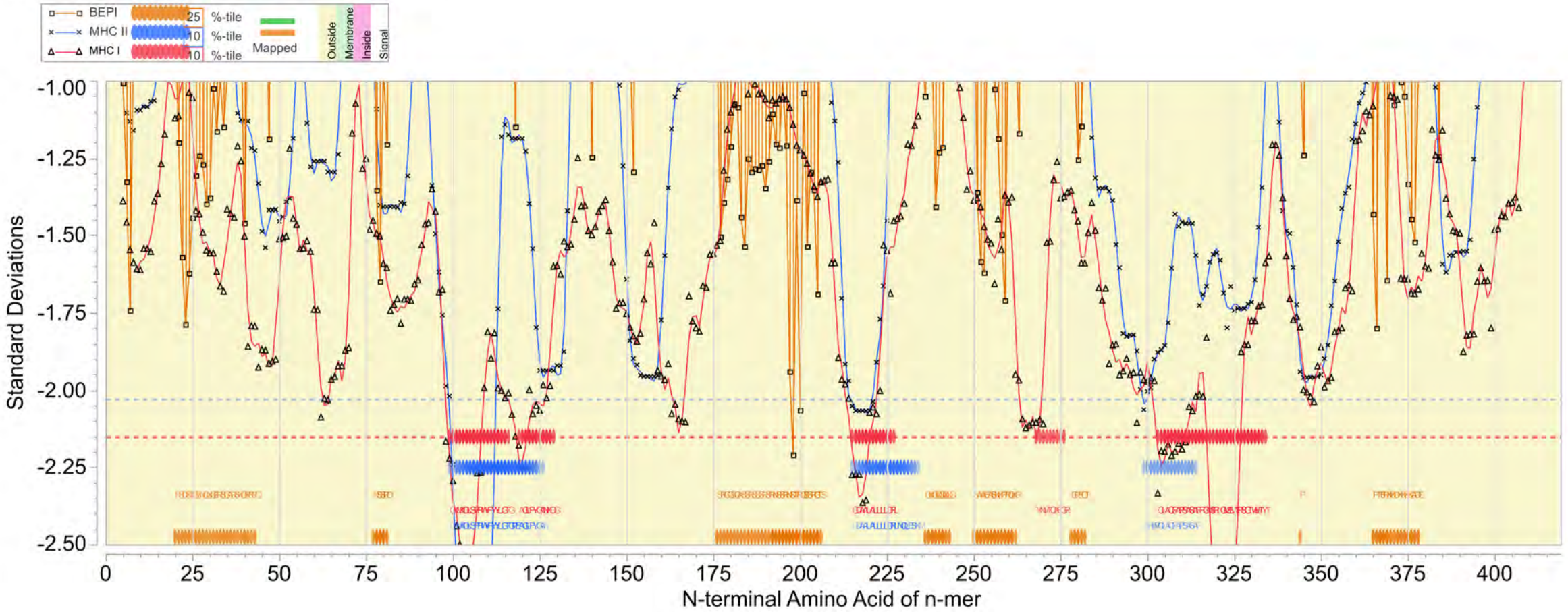
- Figures for each of three structural proteins show our permuted population plot (legend follows) followed by a comparison of EigenBio prediction of peptides which are part of B cell linear epitopes compared to the output of the Pepperprint peptide microarray for several cases documented by Dahkle et al
- The cases included 1 severe – case 101 - and 4 mild cases. Shown is the responses for case 101 and a mean of all cases/all samples (but with no standardization).
- Data was derived from the spreadsheet of microarray output found in the Supplementals to Dahlke *et al*. There are some minor offsets as we predict B cell epitopes as a 9 mer window centered around the index amino acid of a 15 mer and the microarray does not specify where in the 15 mers binding occurs.

Permuted plot legend

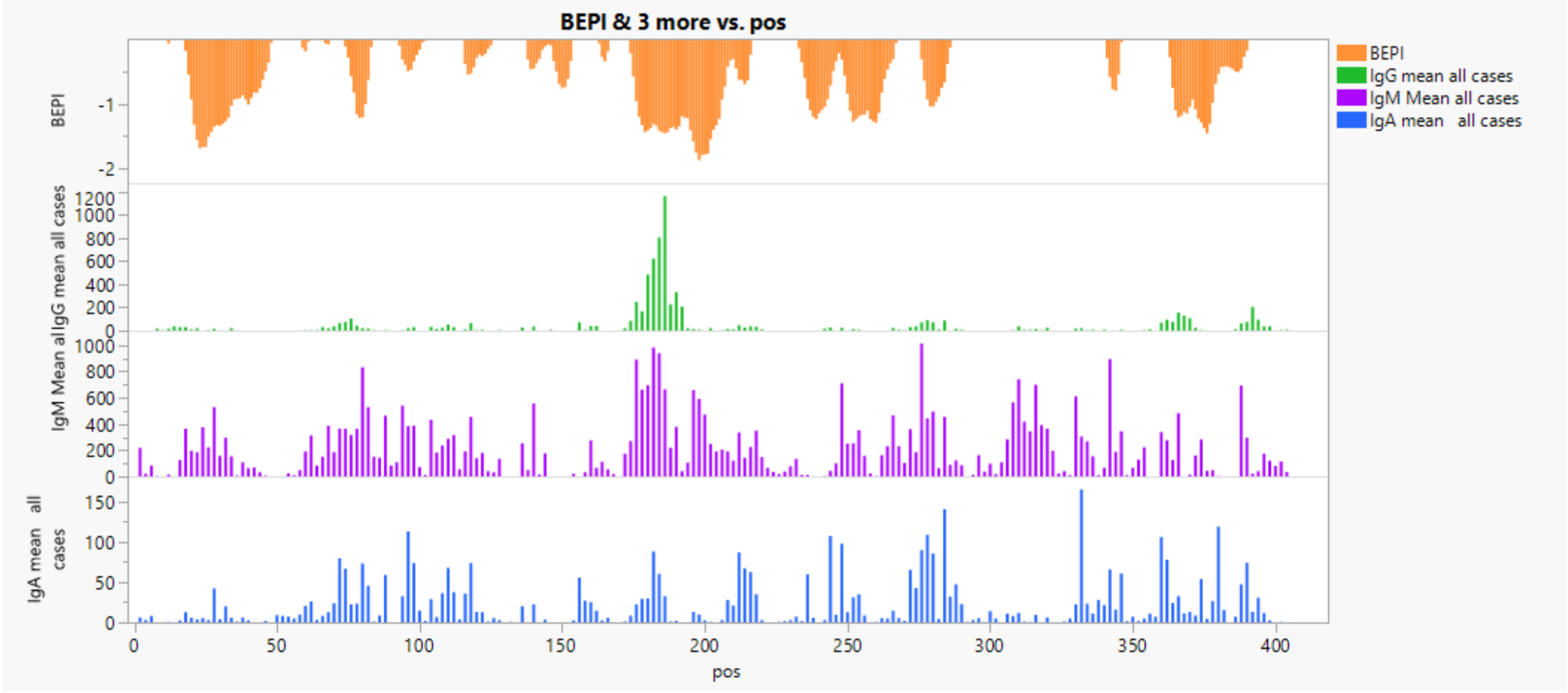
- Sequential peptides are arrayed and numbered N-C. Positions are for the index amino acid of each sequential peptide, with a single amino acid displacement. The Red line shows the permuted average predicted MHC-IA and B (37 alleles) binding affinity by index position of sequential 9-mer peptides. The Blue line shows the permuted average predicted MHC-II DRB allele (16 alleles) binding affinity by index position of sequential 15-mer peptides. Both are plotted in standard deviation units, shown on the Y axis. Orange lines show the predicted probability of B-cell binding ("BEPI") for an amino acid centered in each sequential 9-mer peptide. The standardized B-cell metric has been inverted in order to overlay the standardized probability onto the same scale as the MHC data. Low numbers for MHC data represent high binding affinity, whereas low numbers equate to high BEPI contact probability. Ribbons (Red: MHC-I, Blue: MHC-II) indicate the 10% highest predicted MHC affinity binding. Orange ribbons indicate the top 25% predicted probability B-cell binding. Background shading shows membrane (green) extramembrane (yellow), intramembrane (pink) location. Signal peptides have white background. The Y axis shows predicted affinity in standard deviation units below the mean. Horizontal dotted lines demarcate the top 5% of binding affinity for the protein (red MHC I, blue MHC II).

Overlay Plot

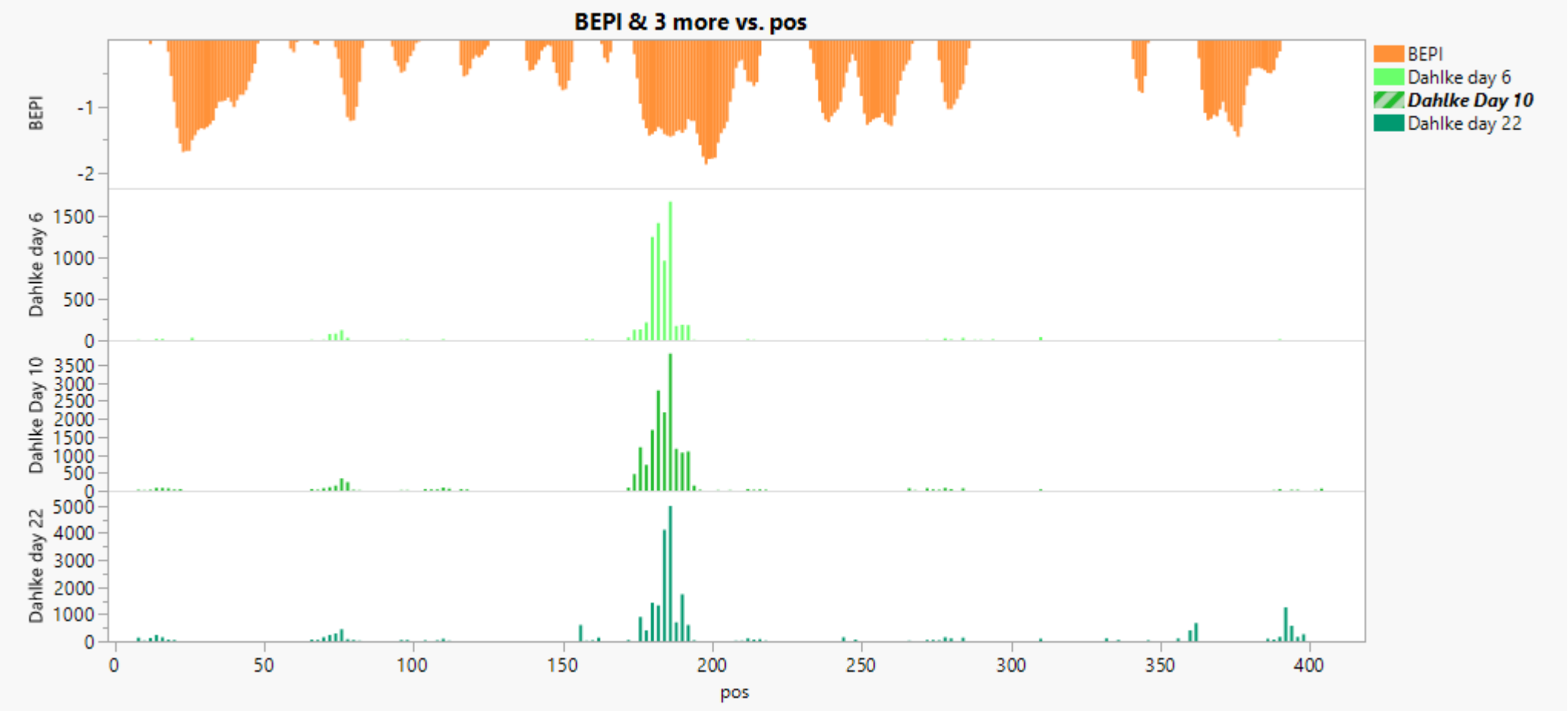
1798174255 nucleocapsid phosphoprotein Severe acute respiratory syndrome coronavirus 2;t



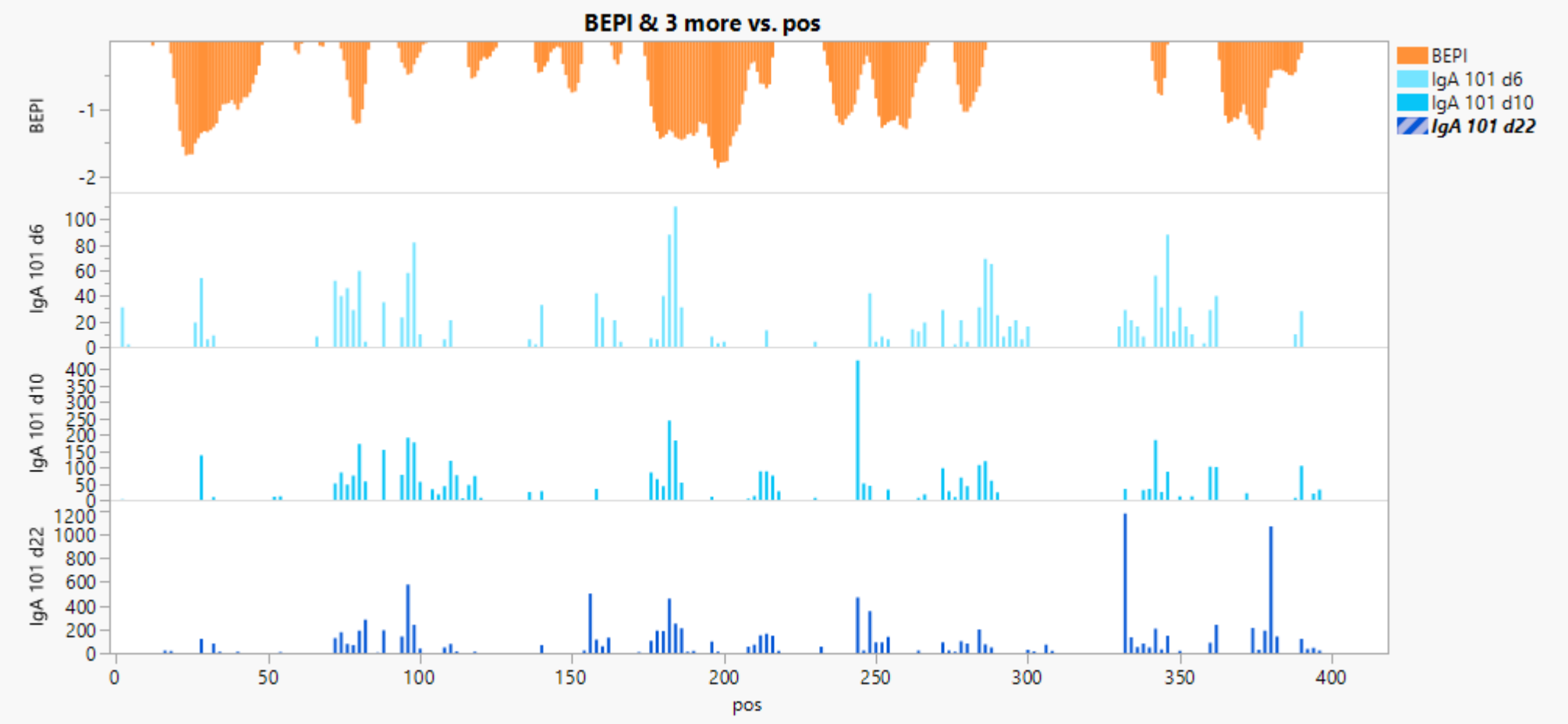
Nucleocapsid Mean all cases



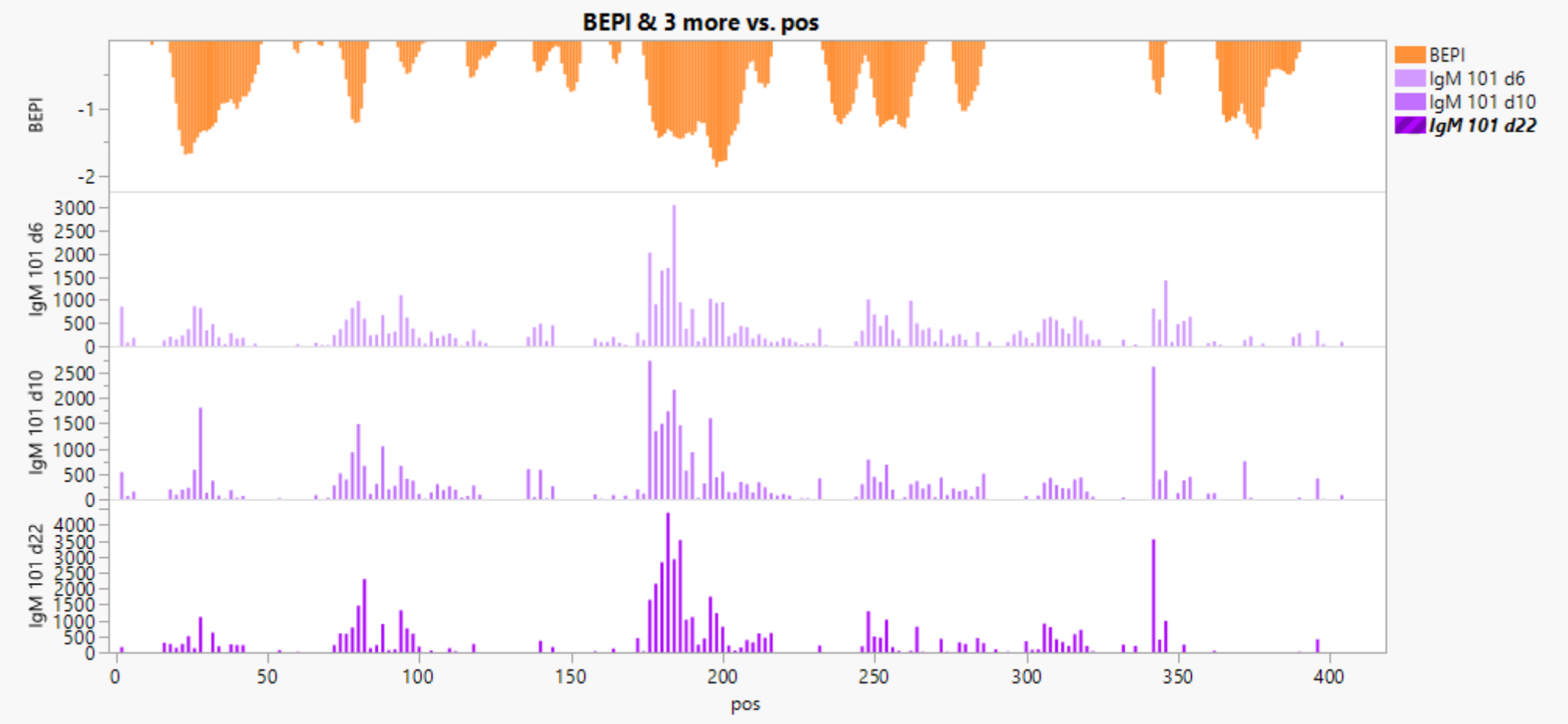
Nucleocapsid IgG -- Patient 101



Nucleocapsid IgA 101

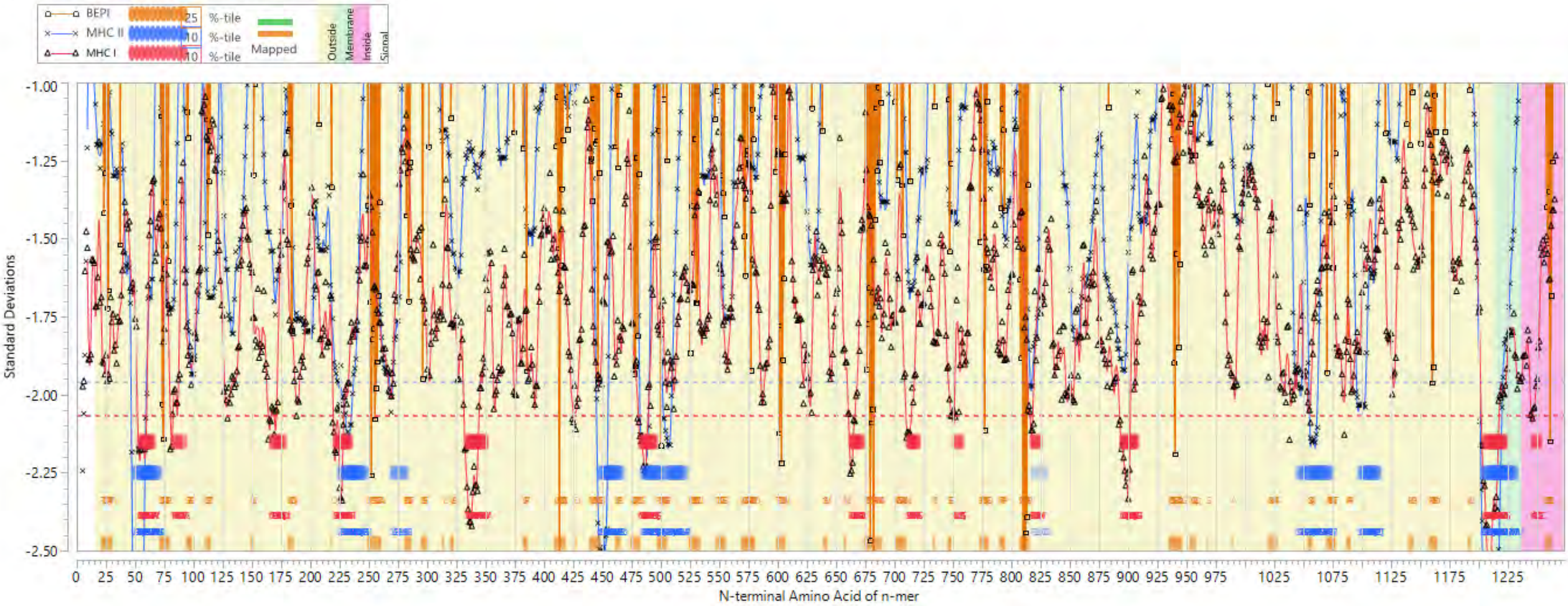


Nucleocapsid IgM 101

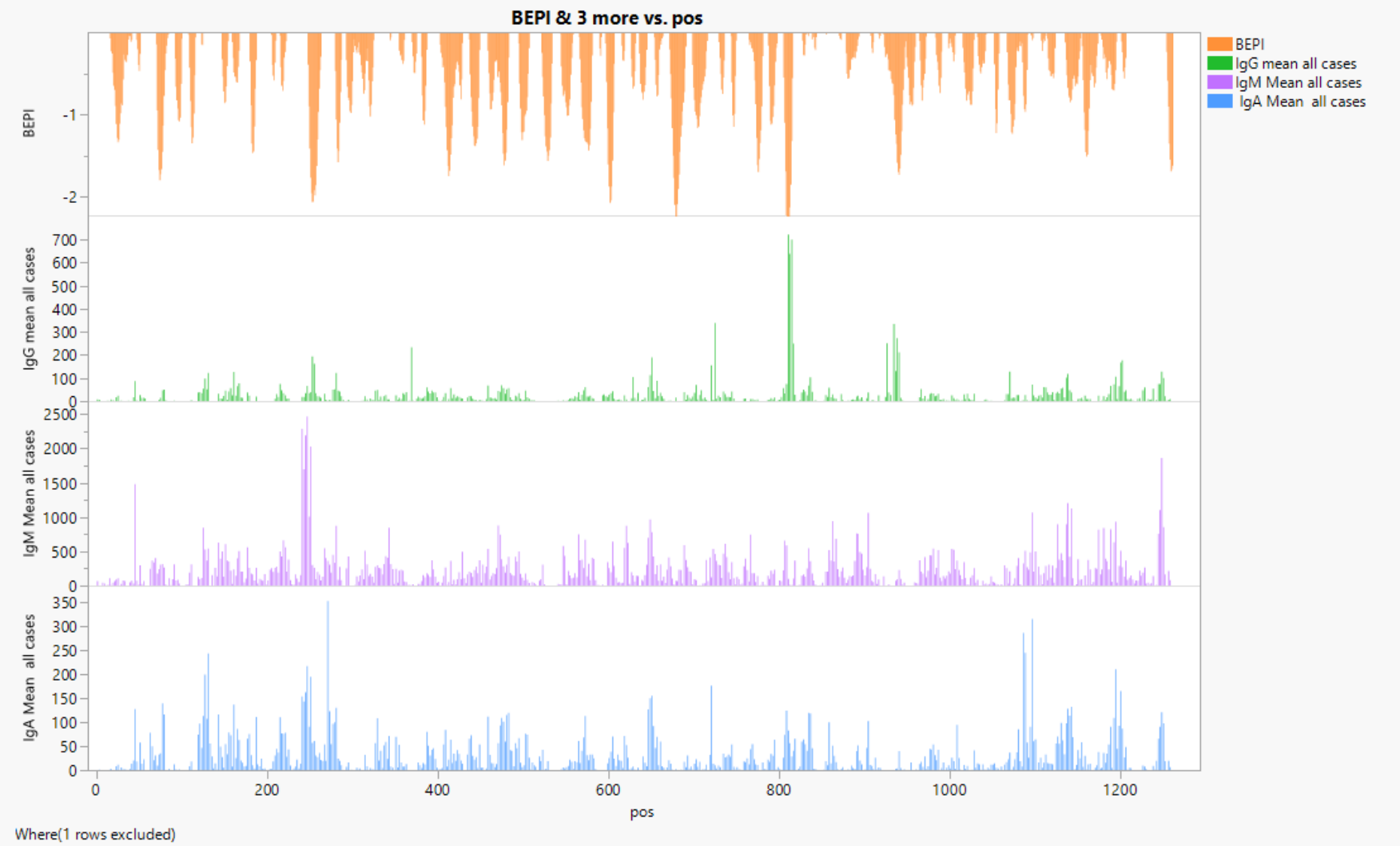


Overlay Plot

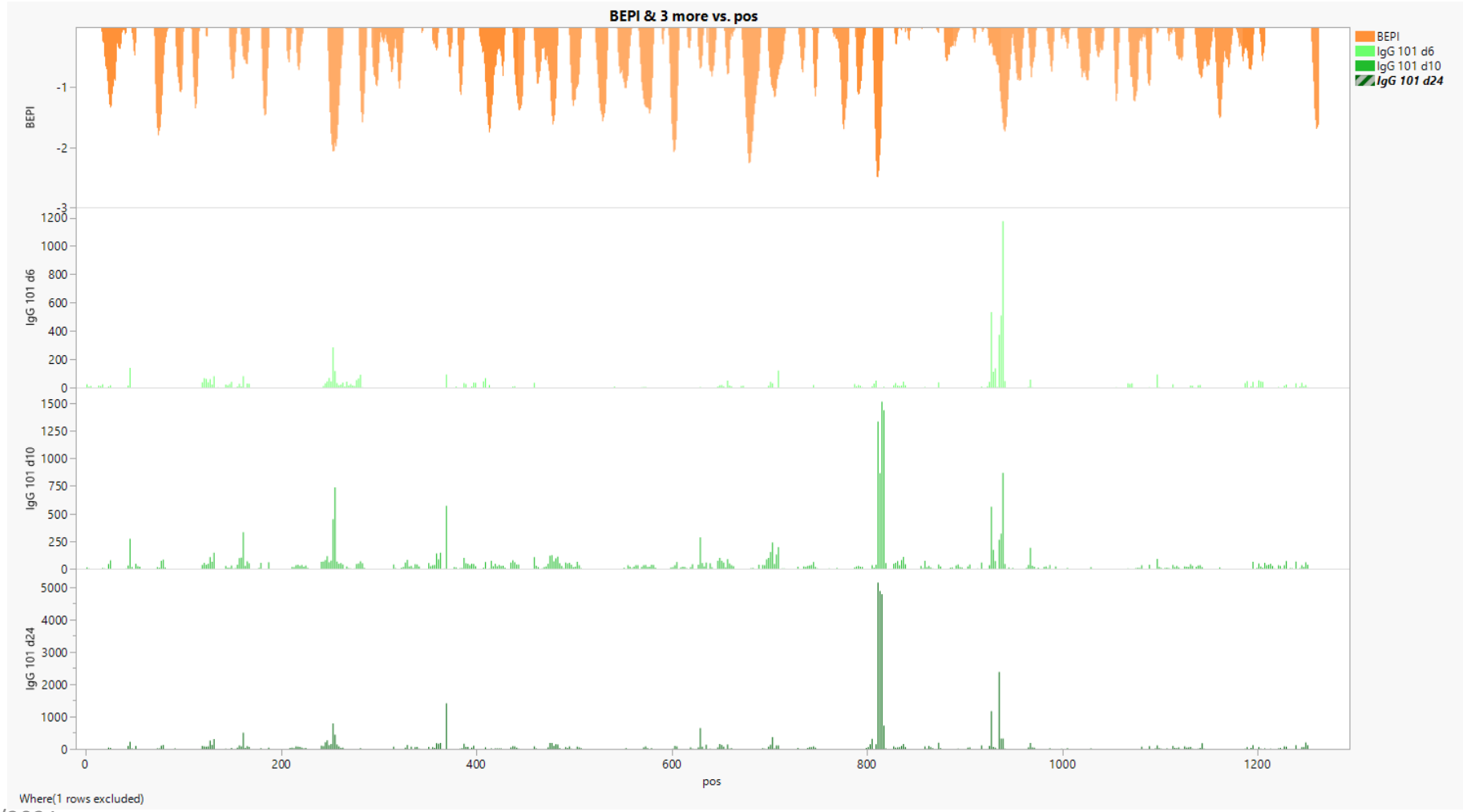
1796318598 surface glycoprotein Severe acute respiratory syndrome coronavirus 2;t



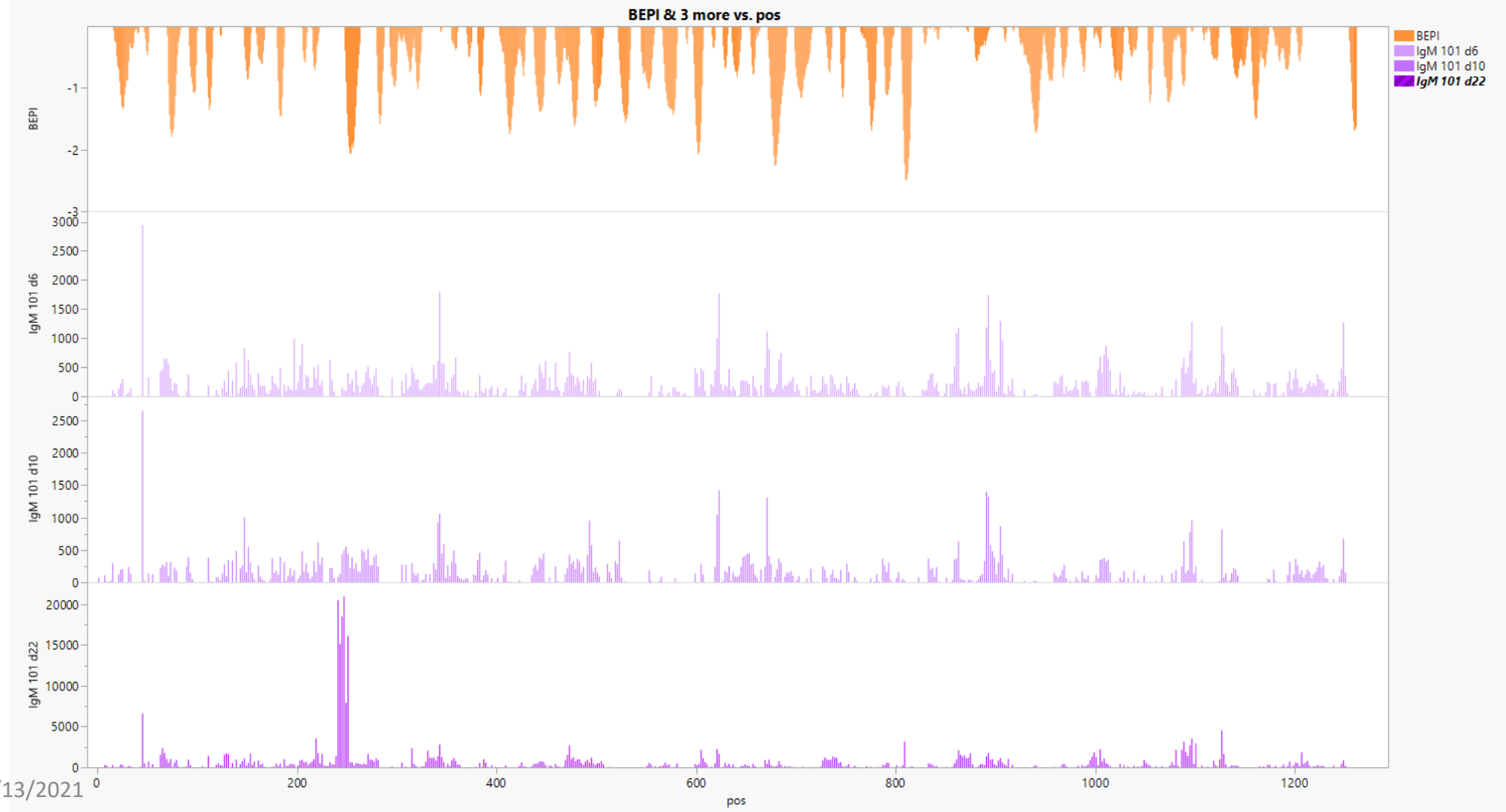
Spike Mean all cases



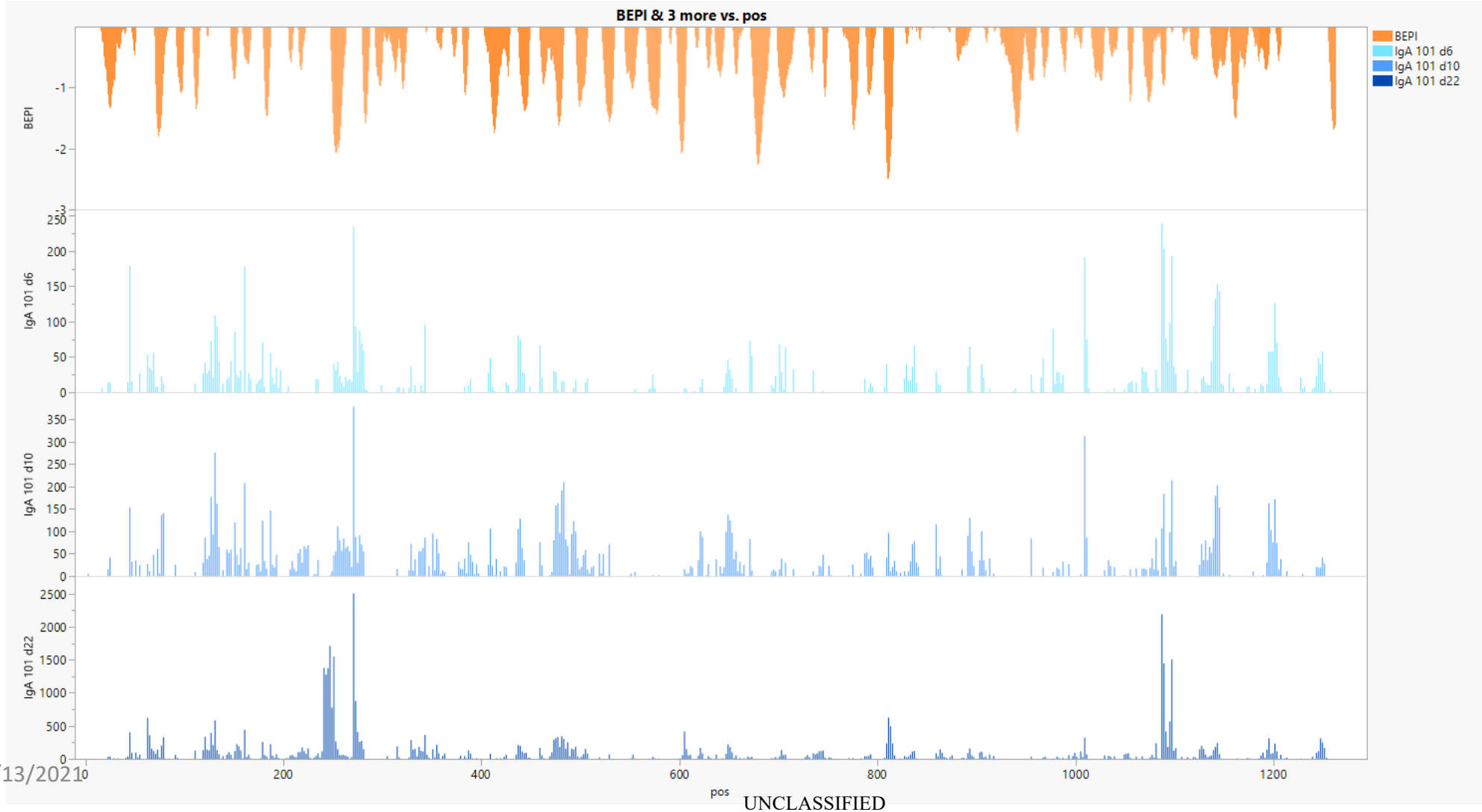
Spike IgG- Patient 101

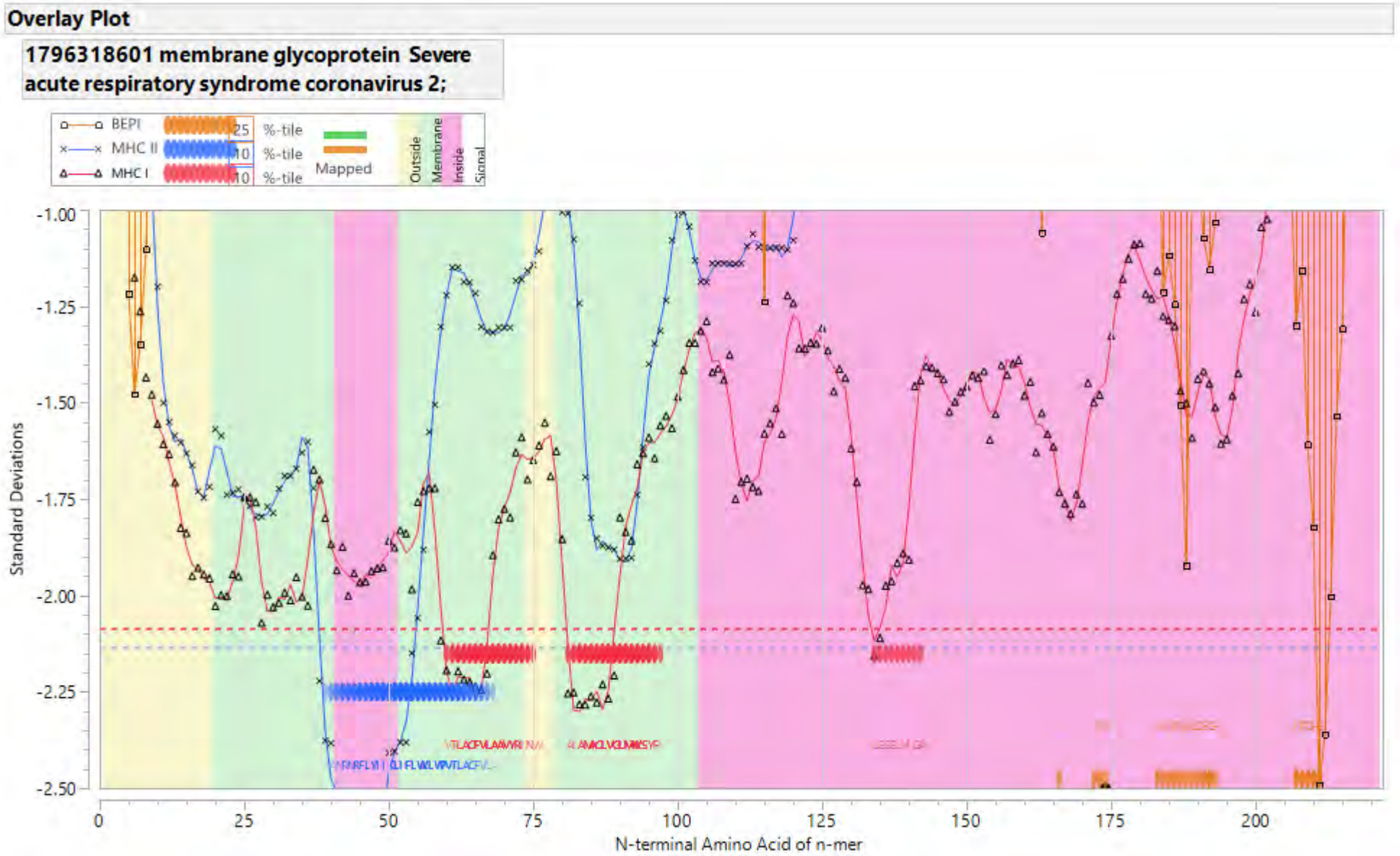


Spike IgM 101

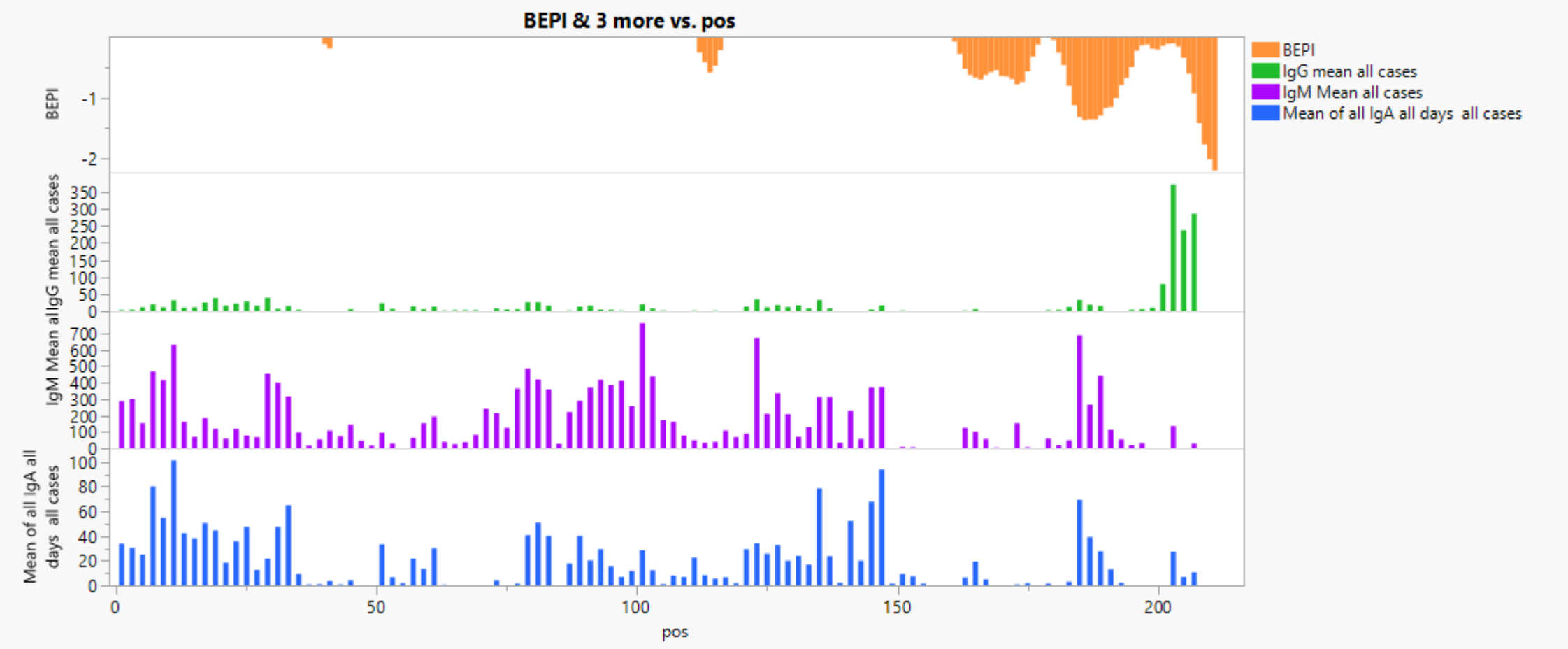


Spike IgA 101

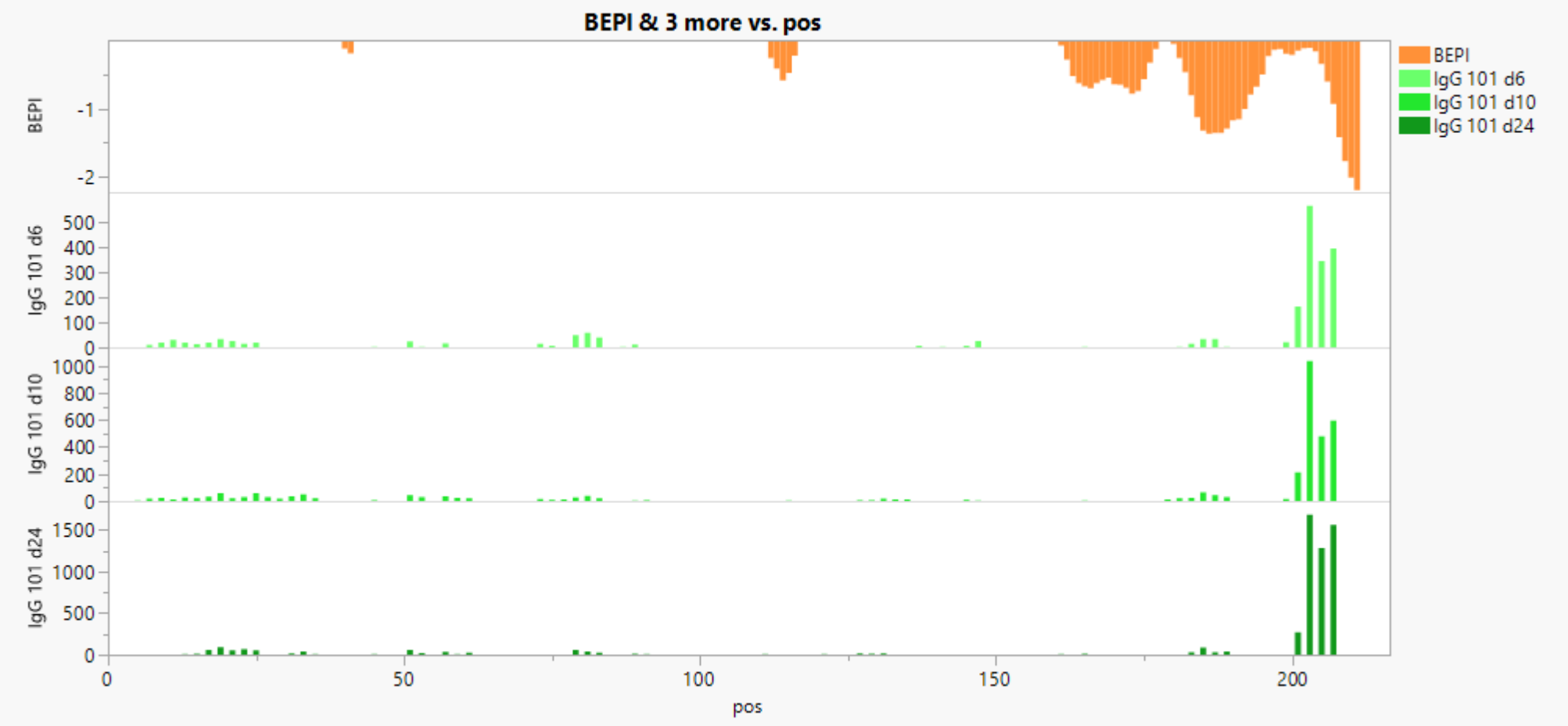



Overlay Plot
1796318601 membrane glycoprotein Severe acute respiratory syndrome coronavirus 2;
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
N-terminal Amino Acid of n-mer

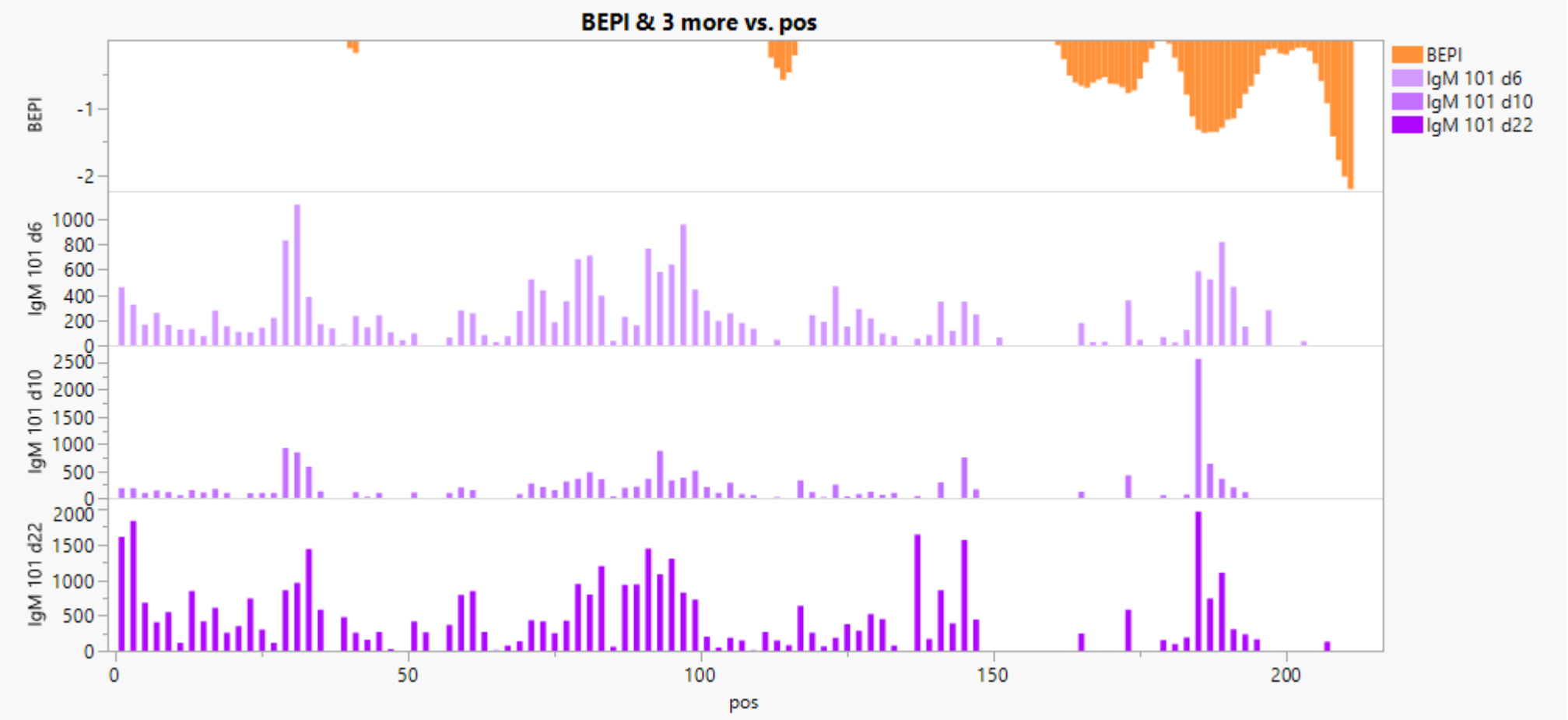
Membrane mean all cases



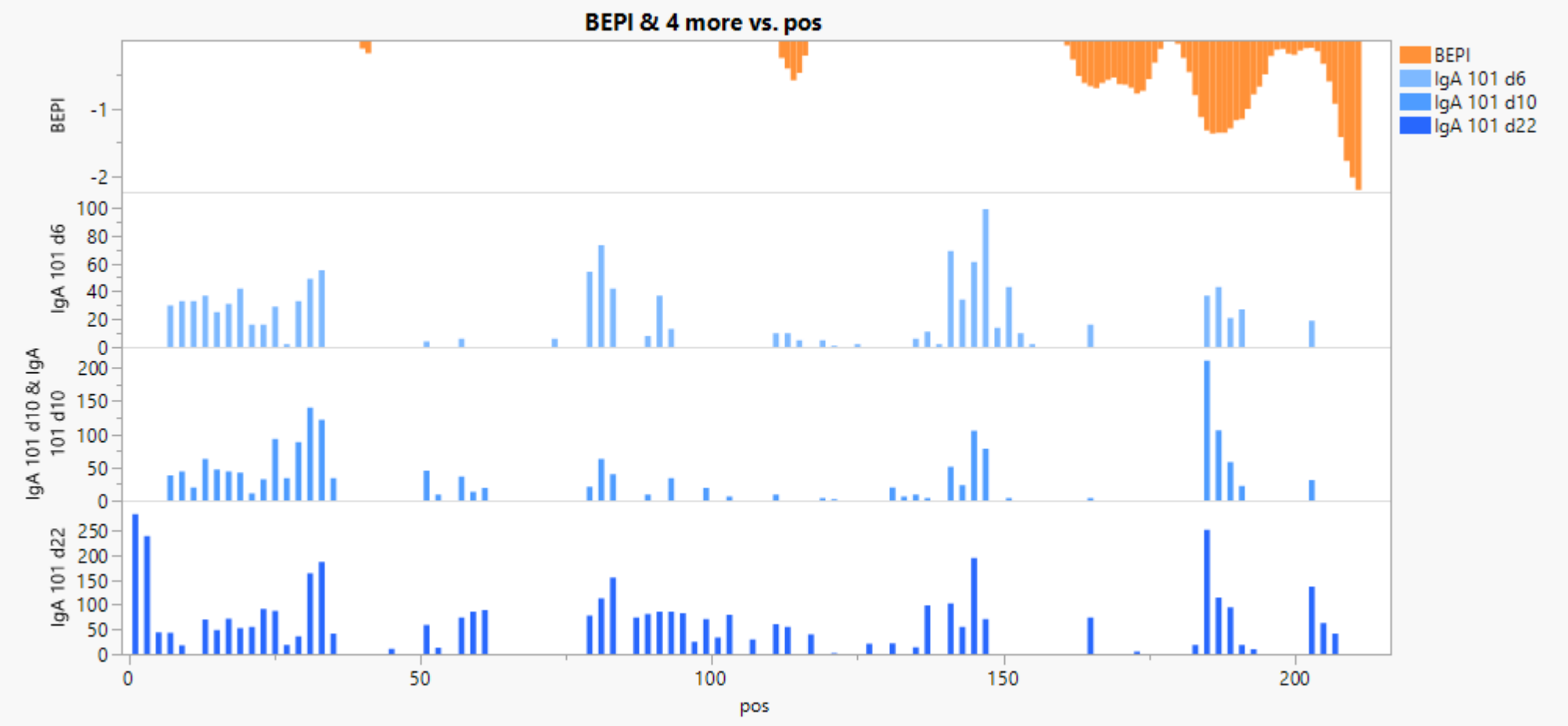
Membrane protein IgG - Patient 101



Membrane IgM 101

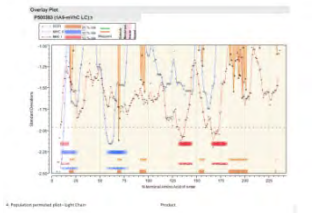


Membrane IgA 101



Appendix 7

General legends for graphics provided for each protein



Permuted population plot

This figure provides a representation of the overall immunome of the protein as seen by an “average” diverse human population. It combines predicted B-cell linear epitopes, predicted MHC-I and MHC-II (NOTE: for DR but not DP and DQ) binding affinities, and protein topology. It does not include the impact of cathepsin cleavage. As such it provides a good overview introduction to the immunologic features of a protein for the human population.

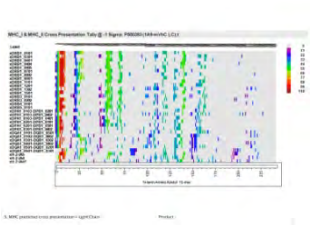
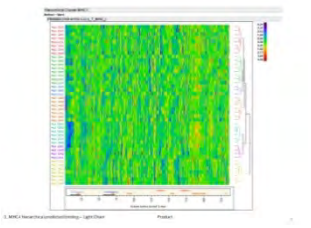
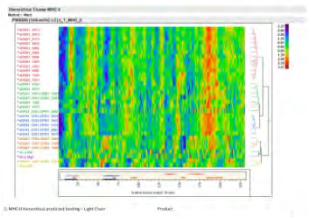
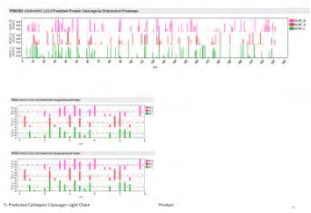
The permuted average of predicted MHC binding affinity is calculated for each sequential peptide. For MHC-I 9-mers are used, for MHC-II 15-mers, each with a single amino acid displacement. At each sequential peptide position we examine every possible heterozygous pair of HLA alleles and determine which allele of the pair has a higher predicted binding affinity. For each of the higher binding alleles of each pair, the mean and standard deviation of predicted binding affinities is calculated. The predicted binding affinity is represented in standard deviation units relative to the mean for all the peptides in the protein as a whole. As the lines on the graphic show the permuted average, any individual allele affinity may differ from the average.

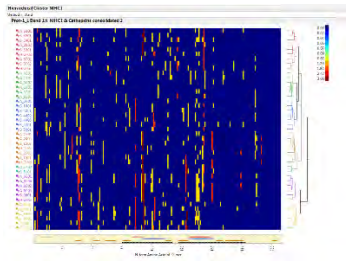
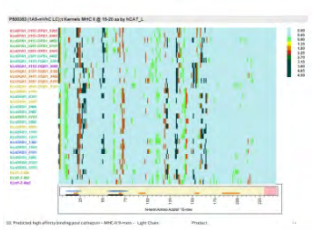
X axis: Sequential peptides are arrayed and numbered N-C. Positions are for the index amino acid of each sequential peptide, with a single amino acid displacement.

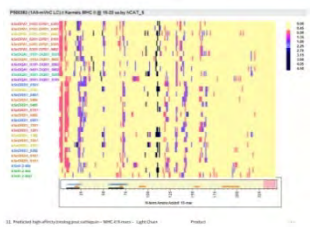
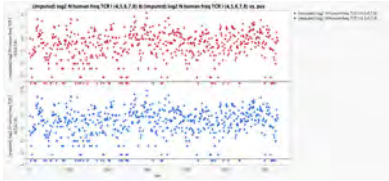
The Red line shows the permuted average predicted MHC-IA and B (37 alleles) binding affinity by index position of sequential 9-mer peptides. The Blue line shows the permuted average predicted MHC-II DRB allele (16 alleles) binding affinity by index position of sequential 15-mer peptides. Both are plotted in standard deviation units, shown on the Y axis. Orange lines show the predicted probability of B-cell binding (“BEPI”) for an amino acid centered in each sequential 9-mer peptide. Note that the standardized B-cell metric has been inverted (multiplied by -1) in order to overlay the standardized probability onto the same scale as the MHC data. Low numbers for MHC data represent high binding affinity, whereas low numbers equate to high BEPI contact probability. Ribbons (Red: MHC-I, Blue: MHC-II) indicate the 10% highest predicted affinity binding. Orange ribbons indicate the top 25% predicted probability B-cell binding.

Background shading shows membrane (green) extramembrane (yellow), intramembrane (pink) location. Signal peptides have white background.

Continuous Epitope Groups (“CEGS”) or Immunological Kernels refers to sequences which comprise overlapping MHC I and MHC II and a B cell epitope contact point occurring within 3 amino acids of the end of a peptide binding MHC molecules at ≤ -1σ below the mean for that protein. In some displays the CEGs have been selected in the Master table so they are indicated by a black underscore and black points in the graphic.

	**MHC predicted cross presentation plot.** X axis shows sequential 15-mer peptides indexed to their N terminal positions; Y axis shows MHC-II alleles. For each 15-mer having a predicted MHC-II binding <-1σ, the number of 9-mer peptides with binding to any MHC-I with a predicted affinity of <-1σ is computed. The number of such predicted MHC-I high binding peptides is indicated by the color, as shown in the thermometer. Hence red indicates a high probability of both MHC-I and MHC-II binding with high affinity to peptides at that location.
 MHC I	**MHC hierarchical predicted binding affinity (two plots – MHC-I and MHC-II)** X axis shows sequential 9-mer (for MHC-I) or 15-mer (for MHC-II) peptides indexed to their N terminal positions; each pixel indicates the peptide starting at that index position. Pixels are colored by the predicted MHC binding affinity for that peptide to the MHC indicated on the Y-axis. Thermometer shows the predicted binding affinity in standard deviation units. A ln(ic50) below the average indicates high binding; blue coloration shows highest predicted binding; red coloration have poor binding affinity. Values <-1 σ are indicative of peptides in top 15% Y axis indicates MHC (human and murine) and shows hierarchical groupings of MHC which have similar binding patterns for this protein. Baseline bar shows regions of highest binding affinity of permuted average of all MHC as ribbons indicating the top 10% affinity binding; Red=MHC-I, Blue=MHC-II. Orange ribbons indicate high probability B-cell binding (top 25%). Black bars show coincident epitope groups (CEGs) where MHC high affinity binding and B-cell epitope high probability overlap or lie within 3 amino acids. Background color shows which peptides are in the transmembrane domain (green), extra-membrane (yellow) or intra-membrane (pink). Signal peptides are white.
 MHC II	
	**Predicted cleavage sites for Cathepsin B, Cathepsin S, and Cathepsin L** X axis shows the amino acid positions from N-C. Y axis shows the predicted probability of cleavage by one of three different endosomal cathepsins, top to bottom cathespin B (pink), cathepsin S (red) and cathepsin L (green). The bar indicates predicted cleavage probability on the C-terminal side of the amino acid at that position. Only probabilities > 0.2 are shown.

	**Predicted high affinity MHC binding peptides following cleavage by cathepsin – MHC-I.** X axis shows sequential peptide index positions from N-C. Y Axis shows MHC-I alleles. Typeface is colored by cluster but alleles are ranked alphanumerically. Murine alleles at the bottom. Individual 9-mer peptides with high predicted binding affinity (<-1σ below the mean) for each allele which are also predicted to be excised by 1-3 cathepsins at their C-terminal are shown. Peptides released by cleavage by all three cathepsins (B, L, S) are shown in dark red, two of 3 in red and 1 of 3 in yellow. The uniform background coloration indicates regions where peptides do not bind with an affinity <1σ, although there may be numerous potential cleavages predicted. Baseline bar shows regions of highest binding affinity of permuted average of all MHC as ribbons indicating the top 10% affinity binding; Red=MHC-I, Blue-MHC-II. Orange ribbons indicate high probability B-cell binding (top 25%). Black bars show coincident epitope groups where MHC high affinity binding and B cell epitope high probability overlap o rile within 3 amino acids. Background color shows which peptides are in the transmembrane domain (green), extra-membrane (yellow) or intra-membrane (pink). Signal peptides are white. For MHC-I we show the consolidated effect of the three cathepsins; plots for individual cathepsins typically show very sparse distribution of excised high affinity peptides; MHC-I epitopes are relatively rare but diffusely distributed.
	**Predicted high affinity MHC binding peptides following cleavage by cathepsin – MHC-II** X axis shows sequential peptide index positions from N-C. Y Axis shows MHC alleles. Typeface is colored by cluster but alleles are ranked alphanumerically. Murine alleles at the bottom. Peptides with high predicted binding affinity (<-1σ below the mean) for each allele, that are between 15 and 20 amino acids in length and that are predicted to be excised by cathepsin cleavage are colored more intensely. The peptide pixel color intensity is indicative of the number of sites with > 0.5 probability of being cleaved i.e. the darkest colors might be cut in any one of six different positions yielding 15-20 mer peptides.

	As the MHC-II binding groove accommodates multiple lengths of peptide, any of the peptides is likely to bind. The uniform background coloration indicates that there are no high affinity peptides in that region, although there may be numerous potential cleavages predicted. The difference in background colors are only to aid in identification of the operative peptidase Baseline bar shows regions of highest binding affinity of permuted average of all MHC as ribbons indicating the top 10% affinity binding; Red=MHC-I, Blue-MHC-II. Orange ribbons indicate high probability B-cell binding (top 25%). Black bars show coincident epitope groups where MHC high affinity binding and B-cell epitope high probability overlap or lie within 3 amino acids. Background color shows which peptides are in the transmembrane domain (green), extra-membrane (yellow) or intra-membrane (pink). Signal peptides are white.
	Frequency of T cell exposed motifs compared to human proteome. X axis shows index position of the containing peptide. Y axis shows frequency of the T cell exposed motif in the human proteome relative the the mean (zero) in the proteome. Values below -3.6 do not occur in the human proteome (shown by line); values of -3.12 occur in only one instance in the human proteome. All isoforms of the human proteome components are considered in this analysis. Top tier shown in red is the TCEM I – positions 4,5,6,7,8 within a 9 mer bound by a MHC I; lower tier shown in blue is TCEM IIA the positions 2,3,5,7,8 relative to the core 9-mer of a 15-mer binding a MHC II

Appendix 8 IEDB Recorded and Eigen Predicted aAllele Specific Peptide Responses

Virus protein	Allele	peptide	TCEM_I	N-term pos peptide	MHC_I Ave	TCEM_I immunglobulins	TCEM_I proteome	N proteome TCEM_I	TCEM_I microbiome	Number of Subjects Tested	Number of Subjects Responded	Response Freq	N IEDB tests	A*02:01	A*24:02	B*07:02	B*08:01	B*15:01	C*06:02	C:07:02
envelope protein	C*0702	LVKPSFYV~	~~~PSFYV~	51	-1.55	22	-0.95	8	-1.05	7	1	0.14	1							-0.18
membrane glycoprotein	A*0201	KLLEQWNLV	~~~EQWNL~	15	-1.87	22	3.08	0	-1.71	13	6	0.46	2	-2.31						
membrane glycoprotein	A*0201	FLFLTWICL	~~~LTWIC~	26	-1.75	24	-3.10	0	-2.50	8	5	0.63	2	-1.75						
membrane glycoprotein	A*2401	AYANRNRFL	~~~NRNRF~	38	-1.70	22	-2.04	2	-1.52	5	1	0.20	1		-1.56					
membrane glycoprotein	B*0801	ANRNRFLY	~~~NRFLY~	40	-1.87	20	-3.10	7	-1.26	10	2	0.20	1				-2.68			
membrane glycoprotein	A*0201	WLLWPVTLA	~~~WPVTL~	55	-1.84	22	-0.54	10	-1.71	5	2	0.40	1	-1.86						
membrane glycoprotein	A*0201	LLWPVTLAC	~~~PVTLA~	56	-1.71	19	0.47	32	0.91	15	2	0.13	1	-1.37						
membrane glycoprotein	A*0201	FVLAAVYRI	~~~AAVYR~	65	-2.27	19	-0.41	10	0.17	15	5	0.33	3	-1.98						
membrane glycoprotein	A*0201	SMWSFNPET	~~~SFNPE~	108	-1.42	22	-0.25	13	-0.17	17	3	0.18	2	-1.16						
membrane glycoprotein	B*0801	SFNPETNIL	~~~PETNI~	111	-1.71	21	-0.47	13	-0.50	10	2	0.20	2				-0.98			
membrane glycoprotein	B*0801	HLRIAGHHI	~~~IAGHH~	148	-1.50	22	-3.10	2	-0.32	4	1	0.25	1				-0.54			
membrane glycoprotein	B*0702	LPKEITVAT	~~~EITVA~	164	-1.59	16	-1.21	7	0.22	4	2	0.50	1			-1.59				
membrane glycoprotein	A*0201	GLMWLSYFI	~~~WLSYF~	89	-2.09	19	-0.69	5	-2.07	59	33	0.56	4	-1.87						
nucleocapsid phosphoprotein	A*0201	NTASWFTAL	~~~SWFTA~	48	-1.92	20	-1.54	5	-2.50	5	3	0.60	1	-1.86						
nucleocapsid phosphoprotein	C*0701	MKDLSPRWY	~~~LSPRW~	101	-2.40	18	-0.36	19	-1.40	8	1	0.13	1							0.50
nucleocapsid phosphoprotein	A*0201	IIWVATEGA	~~~VATEG~	130	-1.59	17	0.45	28	-0.40	4	4	1.00	4	-2.40						
nucleocapsid phosphoprotein	A*0201	LALLLLDRL	~~~LLLDR~	219	-2.27	14	1.61	104	1.08	31	12	0.39	5	-2.01						
nucleocapsid phosphoprotein	A*0201	LLLDRLNQL	~~~DRLNQ~	222	-2.04	21	-0.69	12	0.18	190	48	0.25	11	-1.48						
nucleocapsid phosphoprotein	A*0201	RTATKAYNV	~~~TKAYN~	262	-1.79	22	-0.77	10	-0.92	4	4	1.00	4	-1.98						
nucleocapsid phosphoprotein	A*0201	AQFAPSASA	~~~APSAS~	305	-2.22	11	1.40	104	0.68	5	4	0.80	1	-1.40						
nucleocapsid phosphoprotein	A*0201	GMSRIGMEV	~~~RIGME~	316	-1.94	22	-3.10	1	-0.36	136	11	0.08	7	-3.76						
nucleocapsid phosphoprotein	A*0201	KLDDKDPNF	~~~DKDPN~	338	-1.22	21	-1.08	3	-0.44	3	2	0.67	1	-1.62						
nucleocapsid phosphoprotein	A*0201	ILLNKHIDA	~~~NKHID~	351	-1.95	23	-1.54	4	-0.77	160	9	0.06	5	-1.16						
nucleocapsid phosphoprotein	B*1501	LLNKHIDAY	~~~KHIDA~	352	-1.98	22	-1.76	5	-1.30	3	1	0.33	1					-2.97		
nucleocapsid phosphoprotein	C*0701	KKQQTVTLL	~~~QTVTL~	387	-1.44	16	0.51	34	0.67	8	1	0.13	1							0.04
orf1ab polyprotein	A*0201	VLSEARQHL	~~~EARQH~	38	-1.59	23	-1.21	8	-1.64	9	3	0.33	1	-1.75						
orf1ab polyprotein	B*0801	HLKDGTCGL	~~~DGTCG~	45	-1.49	15	-1.54	3	-1.02	4	1	0.25	1				-1.71			
orf1ab polyprotein	B*0702	VPHVGEIPV	~~~VGEIP~	108	-2.14	21	-1.36	4	0.23	7	1	0.14	1			-2.79				
orf1ab polyprotein	B*1501	HVGEIPVAY	~~~EIPVA~	110	-2.24	19	-1.03	18	0.54	5	2	0.40	2					-2.34		
orf1ab polyprotein	A*2401	TFNGECPNF	~~~GECPN~	265	-2.08	23	-2.45	2	-2.99	4	1	0.25	1		-2.01					
orf1ab polyprotein	C*0602	GFMGRIRSV	~~~GRIRS~	295	-1.76	8	0.40	19	-0.03	1	1	1.00	1						-0.36	
orf1ab polyprotein	A*0201	ILSPLYAFA	~~~PLYAF~	529	-1.77	18	-2.04	6	-0.32	6	1	0.17	1	-2.57						
orf1ab polyprotein	C*0602	VRSIFSRTL	~~~IFSRT~	544	-1.66	21	-1.54	4	-1.13	1	1	1.00	1						-0.12	
orf1ab polyprotein	A*0201	RLIDAMMFT	~~~DAMMF~	579	-1.93	24	-2.04	3	-2.19	6	1	0.17	1	-2.61						
orf1ab polyprotein	A*0201	TVYEKLKPV	~~~EKLKP~	619	-1.92	22	1.07	66	0.86	9	1	0.11	1	-1.70						
orf1ab polyprotein	B*1501	EIKESVQTF	~~~ESVQT~	670	-1.95	22	-0.30	7	-0.20	2	1	0.50	1					-2.33		
orf1ab polyprotein	A*0201	KLVNKFLAL	~~~NKFLA~	680	-1.72	28	0.27	37	0.21	23	6	0.26	2	-3.10						
orf1ab polyprotein	A*0201	FGDTVIEV	~~~GDTVI~	825	[illegible]	18	-0.69	10	0.59	18	5	0.28	2	-1.75						
orf1ab polyprotein	A*0201	ELDERIDKV	~~~ERIDK~	844	-1.63	23	0.08	17	0.52	8	1	0.13	1	-1.56						
orf1ab polyprotein	B*1501	VLNEKCSAY	~~~EKCSA~	852	-1.74	21	-0.69	10	-1.21	7	3	0.43	2					-2.33		
orf1ab polyprotein	B*1501	YLFDESGEF	~~~DESGE~	906	-1.68	18	0.55	35	0.60	2	1	0.50	1					-2.66		
orf1ab polyprotein	B*0702	TPVVQTEGV	~~~VQTE~	1017	-1.69	23	-0.96	5	-0.52	6	1	0.17	1			-2.69				
orf1ab polyprotein	C*0602	NSFSGYLKL	~~~SGYLK~	1026	-1.99	15	0.29	23	0.49	2	1	0.50	1						1.80	
orf1ab polyprotein	B*1501	YLKLTDNVY	~~~LTDNV~	1031	-1.54	21	-0.96	7	0.21	3	1	0.33	1					-2.28		
orf1ab polyprotein	A*2401	LYDKLVSSF	~~~KLVSS~	1182	-1.60	18	0.89	53	1.25	1	1	1.00	1		-2.54					
orf1ab polyprotein	B*1501	KLVSSFLEM	~~~SSFLE~	1185	-1.64	17	0.79	40	0.26	2	1	0.50	1					-2.11		
orf1ab polyprotein	C*0602	AHAEETRKL	~~~EETRK~	1380	-1.79	23	0.90	47	1.07	3	2	0.67	1						0.73	
orf1ab polyprotein	A*0201	KLMPVCVET	~~~PVCVE~	1387	-2.04	23	-1.21	6	-1.79	9	1	0.11	1	-1.95						
orf1ab polyprotein	A*0201	SLINTLNDL	~~~NTLND~	1433	-1.16	16	-0.86	6	0.28	9	1	0.11	1	-2.85						
orf1ab polyprotein	A*2401	YYTSNPTTF	~~~SNPTT~	1536	-1.56	19	-1.36	5	-0.69	10	2	0.20	2		-2.53					
orf1ab polyprotein	C*0702	HLDGEVITF	~~~GEVIT~	1545	-2.06	17	0.29	29	0.94	8	1	0.13	1							-1.58
orf1ab polyprotein	A*0201	RTIKVFTTV	~~~KVFTT~	1566	-2.03	22	-1.21	5	-0.29	17	4	0.24	2	-2.55						
orf1ab polyprotein	B*0801	NINLHTQVV	~~~LHTQV~	1576	-1.35	21	-0.03	13	-1.09	10	2	0.20	2				-2.06			
orf1ab polyprotein	C*0701	LVREAFEYY	~~~EAFEY~	1627	-1.53	21	-3.10	0	-0.66	6	1	0.17	1							-0.83
orf1ab polyprotein	C*0701	WKYPQVNGL	~~~PQVNG~	1656	-1.74	23	-1.08	8	-0.66	6	1	0.17	1							-1.40
orf1ab polyprotein	A*0201	VLATALLTL	~~~TALLT~	1675	-2.09	13	1.11	78	1.66	28	3	0.11	5	-3.59						
orf1ab polyprotein	C*0701	ARAGEAANF	~~~GEAAN~	1702	-1.99	16	-3.10	1	-0.32	6	1	0.17	1							-0.99
orf1ab polyprotein	C*0701	VRETMSYLF	~~~TMSYL~	1728	-1.64	21	-1.54	4	-0.92	6	1	0.17	1							-1.30
orf1ab polyprotein	C*0701	SAPPAQYEL	~~~PAQYE~	1808	-1.68	22	-3.10	0	-1.30	20	4	0.20	2							0.51
orf1ab polyprotein	C*0602	NYYKKDNSY	~~~KKDNS~	1898	-1.22	22	-3.10	18	-0.14	8	1	0.13	1							-2.65
orf1ab polyprotein	A*2401	YYKKDNSYF	~~~KDNSY~	1899	-1.18	14	-3.10	2	-0.03	2	1	0.50	1		-2.37					
orf1ab polyprotein	C*0702	KFADDLNQL	~~~DDLNQ~	1936	-1.65	22	-0.86	6	-0.12	5	1	0.20	1							-0.36
orf1ab polyprotein	B*1501	VVAIDYKHY	~~~IDYKH~	1967	-1.60	28	-1.36	2	-1.17	1	1	1.00	1					-1.30		
orf1ab polyprotein	A*2401	DYKHYTPSF	~~~HYTPS~	1971	-1.45	13	-0.96	6	-1.79	7	2	0.29	2		-2.32					
orf1ab polyprotein	C*0602	LHKDVYLPV	~~~PIVWH~	1986	-1.89	21	-2.86	0	-2.06	3	1	0.33	1						-0.39	
orf1ab polyprotein	B*0801	LTIKKPNEL	~~~KKPNE~	2105	-1.38	23	-0.77	8	-0.32	4	1	0.25	1				-1.44			
orf1ab polyprotein	C*0701	SRVLGLKTL	~~~LGLKT~	2114	-2.12	19	0.47	25	1.41	12	2	0.17	2							-0.14
orf1ab polyprotein	C*0602	YAKPFLNKV	~~~PFLNK~	2141	-1.58	23	-2.04	2	-0.52	3	1	0.33	1						0.78	
orf1ab polyprotein	A*2401	NYMPYFFTL	~~~PYFFT~	2167	-2.56	23	-3.10	0	-1.97	12	4	0.33	3		-2.62					
orf1ab polyprotein	A*0201	SLIYSTAAL	~~~YSTAA~	2242	-1.81	18	-0.69	11	0.34	2	1	0.50	1	-2.24						
orf1ab polyprotein	A*2401	AYILFTRFF	~~~LFTRF~	2330	-2.47	22	-0.36	9	-0.15	3	1	0.33	1		-2.72					
orf1ab polyprotein	A*0201	ILFTRFFYV	~~~TRFFY~	2332	-2.39	19	-2.04	2	-2.07	18	15	0.83	3	-3.59						
orf1ab polyprotein	B*0801	SAMVRMYIF	~~~VRMYI~	2378	-2.30	21	-1.76	3	-2.70	6	2	0.33	1				-2.12			
orf1ab polyprotein	C*0602	VNKGSIHLV	~~~GSIHL~	2496	-1.74	20	-1.08	5	-0.09	3	1	0.33	1						-0.71	
orf1ab polyprotein	C*0602	ERHSLSHFV	~~~SLSHF~	2514	-1.99	19	-2.45	2	-0.34	3	1	0.33	1						-0.29	
orf1ab polyprotein	B*1501	SSAKSASVY	~~~KSASV~	2552	-1.63	20	0.32	24	0.16	1	1	1.00	1					-1.78		
orf1ab polyprotein	B*0801	MFDAYVNTF	~~~AYVNT~	2590	-1.76	23	-0.86	11	-0.61	6	1	0.17	1				-1.84			
orf1ab polyprotein	A*0201	TPRDLGACI	~~~DLGAC~	2684	-1.72	23	-0.07	25	-0.69	7	1	0.14	1			-3.48				
orf1ab polyprotein	B*0702	VAKSHNIAI	~~~SHNIA~	2703	-2.04	22	-1.36	5	-0.64	6	1	0.17	1			-2.31				
orf1ab polyprotein	C*0702	FYLITPVHV	~~~ITPVH~	2787	-1.44	20	-3.10	2	-1.58	8	1	0.13	1							0.45
orf1ab polyprotein	A*0201	YLITPVHVM	~~~TPVHV~	2788	-1.38	21	-1.36	1	-1.05	20	4	0.20	3	-1.86						
orf1ab polyprotein	A*0201	VPGLPGTIL	~~~LPGTI~	2866	-1.76	18	-0.25	10	0.71	7	2	0.29	1			-3.29				
orf1ab polyprotein	A*0201	FLPRVSAV	~~~RVFSA~	2884	-1.95	16	0.01	19	0.11	73	21	0.29	6	-2.77						
orf1ab polyprotein	B*1501	VLAAECTIF	~~~AECTI~	2914	-1.72	24	-2.45	4	-1.09	2	1	0.50	1					-1.78		
orf1ab polyprotein	A*0201	SLPGVFCGV	~~~GVFCG~	3013	-1.54	18	-2.04	3	-0.80	9	2	0.22	1	-2.44						
orf1ab polyprotein	A*0201	IVAGGIVA	~~~GGIVA~	3047	-1.36	12	0.11	14	1.89	9	1	0.11	1	-1.59						
orf1ab polyprotein	A*2401	YFMSFVSFF	~~~RFMSF~	3063	-1.90	18	-0.16	14	-0.44	4	1	0.25	1		-1.66					
orf1ab polyprotein	A*2401	SFLPGVYSV	~~~PGVYS~	3099	-1.87	16	-1.08	6	0.77	4	2	0.50	1		-1.57					
orf1ab polyprotein	A*2401	IYLYLTFYL	~~~YLTFY~	3108	-1.84	20	-0.20	21	-1.13	2	1	0.50	1		-2.18					
orf1ab polyprotein	A*0201	FLAHIQWMV	~~~HIQWM~	3122	-2.12	24	-3.10	0	-1.52	15	4	0.27	3	-2.85						
orf1ab polyprotein	A*2401	MFTPLVPFW	~~~PLVPF~	3131	-1.88	21	-0.30	10	-0.64	2	1	0.50	1		-2.26					
orf1ab polyprotein	A*0201	FLLNKEMYL	~~~KEMYL~	3183	-1.73	20	-3.10	1	-1.13	81	27	0.33	3	-2.83						
orf1ab polyprotein	C*0701	TANPKTPKY	~~~PKTPK~	3356	-1.65	21	0.11	18	0.03	8	1	0.13	1							-0.10
orf1ab polyprotein	B*0702	NPKTPKYKF	~~~TPKYK~	3358	-1.74	22	-1.21	7	-0.92	6	1	0.17	1			-1.93				
orf1ab polyprotein	B*0801	TPKYKFVRI	~~~YKFVR~	3361	-1.72	23	-3.10	0	-2.19	11	3	0.27	1				-2.39			
orf1ab polyprotein	C*0701	VRIQGQTFF	~~~QPGQT~	3367	-1.55	22	0.04	18	-0.24	14	5	0.36	1							-1.57
orf1ab polyprotein	B*0702	QPGQTFSVL	~~~QTFSV~	3370	-1.87	23	-0.07	20	-1.13	5	1	0.20	1			-3.04				
orf1ab polyprotein	A*0201	ILTSLLVLV	~~~SLLVL~	3587	-1.20	18	1.89	126	2.30	9	1	0.11	1	-2.19						
orf1ab polyprotein	A*0201	FLYENAFLP	~~~ENAFL~	3605	-1.63	23	0.01	11	0.48	1	1	1.00	1	-2.03						
orf1ab polyprotein	C*0702	MFKHKHAAF	~~~KHKHA~	3627	-1.54	23	-3.10	0	-1.52	20	3	0.15	2							-1.34
orf1ab polyprotein	B*0801	FVKHKHAFL	~~~HKHAF~	3628	-1.64	22	-2.04	4	-1.26	13	2	0.15	2				-2.63			
orf1ab polyprotein	A*0201	FLLPSLATV	~~~PSLAT~	3639	-2.13	15	1.13	63	0.14	23	4	0.17	3	-2.25						
orf1ab polyprotein	A*2401	VYMPASWVN	~~~PASWV~	3653	-2.06	21	-2.45	1	-2.07	6	2	0.33	2		-1.33					
orf1ab polyprotein	A*0201	RIMTWLDMV	~~~TWLDM~	3662	-1.57	23	-3.10	0	-2.07	6	1	0.17	1	-1.67						
orf1ab polyprotein	A*0201	WLDMVDTSL	~~~MVDTS~	3666	-1.36	20	-3.10	0	-0.57	1	1	1.00	1	-1.30						
orf1ab polyprotein	A*0201	KLKDCVMYA	~~~DCVMY~	3678	-1.76	23	-3.66	0	-2.50	15	6	0.40	1	-1.88						
orf1ab polyprotein	C*0702	MYASAVVLL	~~~SAVVL~	3684	-1.93	16	1.24	70	1.57	5	1	0.20	1							0.29
orf1ab polyprotein	C*0602	YASAVVLLI	~~~AVVLL~	3685	-1.88	15	1.50	98	2.46	1	1	1.00	1						2.13	
orf1ab polyprotein	A*0201	TLMWLTLV	~~~NVLTL~	3710	-1.70	16	0.51	35	0.76	5	2	0.40	2	-2.27						
orf1ab polyprotein	A*0201	SMWALIISV	~~~ALIIS~	3732	-1.64	14	0.18	22	1.41	14	4	0.29	4	-2.52						
orf1ab polyprotein	A*0201	FLARGIVFM	~~~RGIVF~	3753	-1.59	19	-1.08	6	-0.24	14	2	0.14	2	-1.50						
orf1ab polyprotein	A*2401	IFFITGNTL	~~~ITGNT~	3768	-1.76	14	-0.36	12	0.33	3	1	0.33	1		-0.77					
orf1ab polyprotein	A*0201	GVYDYLVST	~~~DYLVS~	3809	-1.41	19	0.11	11	0.49	9	1	0.11	1	-0.18						
orf1ab polyprotein	A*0201	GLLPPKNIS	~~~PPKNS~	3827	-1.29	20	-0.77	15	-1.79	9	1	0.11	1	-0.81						
orf1ab polyprotein	A*0201	KLNIKLLGV	~~~IKLLG~	3839	-1.58	21	-0.54	9	1.25	15	2	0.13	1	-1.82						
orf1ab polyprotein	A*0201	KLWAQCVQL	~~~AQCVQ~	3886	-1.71	24	-2.04	1	-2.19	102	49	0.48	7	-2.01						
orf1ab polyprotein	B*0801	EAFEKMVSL	~~~EKMVS~	3906	-1.75	22	0.42	23	-0.25	23	2	0.09	2				-2.55			
orf1ab polyprotein	A*0201	SVLLSVLSM	~~~SVLLS~	3913	-1.68	19	1.70	208	1.80	17	3	0.18	2	-1.85						
orf1ab polyprotein	A*2401	TFTYASALW	~~~YASAL~	4088	-1.73	21	0.40	37	1.27	5	1	0.20	1		-2.40					
orf1ab polyprotein	A*0201	ALWEIQQVV	~~~EIQQV~	4094	-1.75	22	-1.21	5	-0.05	125	32	0.26	7	-3.31						
orf1ab polyprotein	A*2401	AWPLIVTAL	~~~LIVTA~	4123	-2.02	16	0.21	24	1.34	3	1	0.33	1		-2.09					
orf1ab polyprotein	B*0801	GPKVKYLYF	~~~VKYLY~	4222	-1.97	21	-0.86	6	-0.34	6	1	0.17	1				-2.38			
orf1ab polyprotein	C*0702	YFHKGLNNL	~~~KGLNN~	4229	-1.49	21	-0.47	13	0.56	8	1	0.13	1							-0.76
orf1ab polyprotein	B*0801	DLKGKYVQI	~~~GKYVQ~	4344	-1.67	22	-2.04	4	-0.69	27	7	0.26	3				-2.51			
orf1ab polyprotein	B*1501	VVYRAFDIY	~~~RAFDI~	4422	-1.64	9	-1.21	7	-0.40	2	1	0.50	1					-1.79		
orf1ab polyprotein	A*0201	NLIDSYFVV	~~~DSYFV~	4456	-1.52	22	-0.61	6	-0.52	16	1	0.06	1	-3.27						
orf1ab polyprotein	B*0702	VPHISRQRL	~~~ISRQR~	4503	-1.68	19	-1.47	17	0.44	1	1	1.00	1			-3.42				
orf1ab polyprotein	A*0201	TMADLVYAL	~~~DLVYA~	4515	-1.91	18	-1.54	3	0.32	14	3	0.21	2	-3.10						
orf1ab polyprotein	C*0602	VRQALLKTV	~~~ALLKT~	4574	-1.96	23	0.77	45	1.75	3	1	0.33	1						-0.51	
orf1ab polyprotein	A*0201	SLLMPILTL	~~~MPILT~	4631	-1.82	22	-3.66	0	-1.17	6	1	0.17	1	-3.10						
orf1ab polyprotein	B*0702	KPYIKWDLL	~~~IKWDL~	4655	-1.87	22	-2.45	0	-0.98	13	2	0.15	1			-2.99				
orf1ab polyprotein	A*2401	SFGPLVRKI	~~~PLVRK~	4717	-2.08	18	0.35	27	0.71	5	1	0.20	1		-2.32					
orf1ab polyprotein	A*0201	FVDGVPFVV	~~~GVPFV~	4726	-1.97	20	-0.69	12	0.05	63	32	0.51	5	-3.07						
orf1ab polyprotein	B*0801	NLHSSRLSF	~~~SSRLS~	4752	-1.54	15	1.72	138	0.67	6	1	0.17	1				-1.00			
orf1ab polyprotein	C*0602	YRYNLPTMC	~~~NLPTM~	4848	-1.54	22	-3.66	0	-2.33	3	1	0.33	1						-0.14	
orf1ab polyprotein	A*0201	RQLLFVFEV	~~~LFVFE~	4859	-1.90	16	-0.58	9	0.33	72	29	0.40	5	-2.31						
orf1ab polyprotein	B*1501	FVVEVVDKY	~~~EVVDK~	4863	-1.74	21	0.69	38	0.91	3	1	0.33	1					-2.15		
orf1ab polyprotein	C*0701	NVIPTITQN	~~~PTITQ~	4926	-1.84	22	-0.54	11	-0.66	6	1	0.17	1							0.96
orf1ab polyprotein	B*0702	IPTITQMNL	~~~ITQMN~	4928	-1.90	23	-1.54	5	-1.88	6	1	0.17	1			-2.99				
orf1ab polyprotein	C*0602	SAKNRARTV	~~~NRART~	4941	-1.27	20	-3.66	0	-0.66	3	1	0.33	1						-0.40	
orf1ab polyprotein	C*0702	LRIMASLVL	~~~MASLV~	5022	-2.03	23	-0.47	16	-0.02	7	1	0.14	1							-0.07
orf1ab polyprotein	A*0201	RLNAECAQV	~~~NECAQ~	5046	-1.81	24	-2.04	3	-2.99	6	1	0.17	1	-2.58						
orf1ab polyprotein	C*0701	VRNLQHRLY	~~~LQHRL~	5112	-1.37	22	0.59	33	0.29	6	1	0.17	1							-1.16
orf1ab polyprotein	A*0201	FVNEFYAYL	~~~EFYAY~	5133	-1.83	22	-2.45	1	-1.02	15	2	0.13	1	-2.73						
orf1ab polyprotein	A*2401	FYAYLRKHF	~~~YLRKH~	5137	-1.84	22	-0.77	7	-0.23	12	7	0.58	2		-1.82					
orf1ab polyprotein	C*0602	LRKHFSMM	~~~HFSMM~	5141	-1.68	22	-2.04	2	-2.50	2	1	0.50	1						-0.42	
orf1ab polyprotein	A*2401	LYYQNNVFM	~~~QNNVF~	5178	-1.44	23	-1.08	5	-1.13	4	1	0.25	1		-2.12					
orf1ab polyprotein	A*0201	KQGDDYVYL	~~~DDYVY~	5213	-1.46	18	-1.36	4	-0.46	8	1	0.13	1	-1.35						
orf1ab polyprotein	A*0201	LMIERFVSL	~~~ERFVS~	5246	-2.32	28	0.32	29	0.64	27	3	0.11	2	-3.28						
orf1ab polyprotein	C*0702	YADFFHLYL	~~~FFHLY~	5269	-2.33	22	0.00	0	-1.05	5	1	0.20	1							1.36
orf1ab polyprotein	C*0702	LYLQYIRKL	~~~QYIRK~	5275	-1.81	23	-1.36	5	-0.71	20	4	0.20	2							-1.54
orf1ab polyprotein	A*0201	KLSYGIATV	~~~YGIAT~	5470	-1.97	22	-0.54	7	0.28	70	26	0.37	5	-3.17						
orf1ab polyprotein	C*0602	YRGTTTYKL	~~~TTTYK~	5535	-1.41	19	0.37	22	-1.17	1	1	1.00	1						0.26	
orf1ab polyprotein	A*0201	KLNVGDYFV	~~~VGDYF~	5542	-1.76	14	-0.47	17	-0.01	63	17	0.27	4	-3.40						
orf1ab polyprotein	B*0801	VPQEHVVRI	~~~EHVVR~	5565	-2.05	23	-1.76	2	-1.64	6	1	0

Appendix 8 IEDB Recorded and Eigen Predicted aAllele Specific Peptide Responses

Virus protein	Allele	peptide	TCEM_I	N-term pos peptide	MHC_I Ave	TCEM_I immunglobulins	TCEM_I proteome	N proteome TCEM_I	TCEM_I microbiome	Number of Subjects Tested	Number of Subjects Responded	Response Freq	N IEDB tests	A*02:01	A*24:02	B*07:02	B*08:01	B*15:01	C*06:02	C:07:02
orf1ab polyprotein	A*0201	NLWNTFTRL	~~~NTFTR~	6443	-1.55	15	-0.16	14	-1.35	60	9	0.15	3	-2.99						
orf1ab polyprotein	C*0701	NTFTRLQSL	~~~TRLQS~	6446	-1.47	16	0.45	37	0.15	8	1	0.13	1							-0.29
orf1ab polyprotein	A*0201	SLENVAFNV	~~~NVAFN~	6453		24	-1.54	3	-0.44	9	1	0.11	1	-3.12						
orf1ab polyprotein	A*0201	KVDGVDVEL	~~~GVDVE~	6486	-1.28	16	-0.03	15	0.83	15	4	0.27	1	-1.26						
orf1ab polyprotein	B*1501	TICAPLTVF	~~~APLTV~	6566	-2.24	17	0.75	45	0.74	3	1	0.33	1					-2.68		
orf1ab polyprotein	C*0701	GRVDGQVDL	~~~DGQVD~	6577	-1.59	18	0.47	20	-0.32	8	1	0.13	1							-2.07
orf1ab polyprotein	A*0201	KVDGVVQQL	~~~GVVQQ~	6633	-1.31	21	-0.16	14	-0.08	15	2	0.13	1	-1.51						
orf1ab polyprotein	B*1501	VVQQLPETY	~~~QLPET~	6637	-1.36	12	0.29	26	-0.14	3	1	0.33	1					-1.75		
orf1ab polyprotein	A*0201	SVVSKVVKV	~~~SKVVK~	6764	-1.68	22	0.37	24	0.36	9	1	0.11	1	-2.35						
orf1ab polyprotein	B*0702	LPKGIMMNV	~~~GIMMN~	6834	-1.74	22	-3.67	0	-1.88	7	1	0.14	1			-2.79				
orf1ab polyprotein	C*0701	WRNTNPIQL	~~~TNPIQ~	7029	-1.36	22	-3.66	0	-0.50	6	1	0.17	1							-2.57
orf1ab polyprotein	B*0702	FPLKLRGTA	~~~KLRGT~	7048	-1.90	19	-0.61	10	0.22	2	1	0.50	1			-3.18				
orf1ab polyprotein	C*0701	LRGTAVMSL	~~~TAVMS~	7052	-1.73	22	-0.30	14	-0.50	8	1	0.13	1							-2.21
ORF3a protein	A*0201	ALLAVFQSA	~~~AVFQS~	51	-1.79	22	-0.36	12	-0.80	6	1	0.17	1	-2.21						
ORF3a protein	B*0801	QSASKIITL	~~~SKIIT~	57	-1.61	20	-0.36	11	0.02	13	3	0.23	3				-1.06			
ORF3a protein	B*0801	TLKRWQLA	~~~KRWQL~	64	-1.76	20	-2.45	8	-1.71	4	1	0.25	1				-2.66			
ORF3a protein	A*0201	ALSKGVHFV	~~~KGVHF~	72	-1.87	22	-1.54	5	-0.20	91	24	0.26	6	-2.26						
ORF3a protein	A*0201	NLLLLFVTV	~~~LLFVT~	82	-2.12	21	0.63	54	0.87	8	2	0.25	1	-2.10						
ORF3a protein	A*0201	FVTVYSHLL	~~~VYSHL~	87	-1.56	22	-0.16	17	0.16	9	1	0.11	1	-1.87						
ORF3a protein	A*0201	TVYSHLLLV	~~~SHLLL~	89	-1.61	21	1.03	36	-0.06	3	1	0.33	1	-2.39						
ORF3a protein	A*0201	GLEAPFLYL	~~~APFLY~	100	-2.29	18	-0.25	6	-0.50	3	1	0.33	1	-2.08						
ORF3a protein	C*0702	FLYLYALVY	~~~LYALV~	105	-2.15	21	-0.20	18	0.77	5	1	0.20	1							-0.28
ORF3a protein	A*2401	LYLYALVYF	~~~YALVY~	106	-1.99	15	-1.36	5	-1.35	2	1	0.50	1		-2.87					
ORF3a protein	A*0201	YLYALVYFL	~~~ALVYF~	107	-1.82	20	-1.76	4	0.33	36	13	0.36	3	-2.90						
ORF3a protein	A*2401	VYFLQSINF	~~~LQSIN~	112	-1.78	21	0.72	36	0.16	31	19	0.61	5		-2.35					
ORF3a protein	C*0702	VRIIMRLWL	~~~IMRLW~	121	-2.01	22	-3.69	0	-2.99	7	1	0.14	1							-0.51
ORF3a protein	C*0602	CRSKNPLLY	~~~KNPLL~	133	-1.79	16	0.27	26	1.22	2	1	0.50	1						0.08	
ORF3a protein	A*0201	LLYDANYFL	~~~DANYF~	139	-1.81	20	-2.04	3	-0.06	99	62	0.63	10	-2.47						
ORF3a protein	C*0702	YFTSDYYQL	~~~SDYYQ~	206	-1.46	18	-2.45	3	-0.77	7	1	0.14	1							-2.29
ORF3a protein	C*0702	IYDEPTTTT	~~~EPTTT~	263	-1.24	15	0.18	13	-1.05	8	1	0.13	1							-1.05
ORF6 protein	A*0201	HLVDFQVTI	~~~DFQVT~	3		20	-0.47	9	-0.77	61	13	0.21	5	-1.73						
ORF6 protein	B*0801	LIIMRTFKV	~~~MRTFK~	16	-2.12	24	-2.45	1	-2.50	6	1	0.17	1				-2.47			
ORF6 protein	B*0801	TFKVSIWNL	~~~VSIWN~	21	-1.75	22	-1.76	2	-1.58	4	1	0.25	1				-1.14			
ORF7a protein	B*0801	EGNSPFHPL	~~~SPFHP~	41	-1.24	20	-0.54	12	-1.79	6	1	0.17	1				-0.76			
ORF7a protein	B*0702	RARSVSPKL	~~~SVSPK~	78	-1.76	18	0.53	27	-0.66	5	4	0.80	1			-3.39				
ORF7a protein	B*0702	SPIFLIVAA	~~~FLIVA~	98	-2.05	16	0.11	16	1.39	6	1	0.17	1			-2.72				
ORF8 protein	B*0702	EPKLGSLVV	~~~LGSLV~	92	-1.93	15	1.29	87	1.69	7	1	0.14	1			-2.91				
ORF8 protein	A*0201	KLGSLVVRC	~~~SLVVR~	94	-2.01	17	-0.20	18	0.47	15	3	0.20	1	-1.64						
ORF8 protein	A*2401	EYHDVRVVL	~~~DVRVV~	110	-1.97	17	-0.61	10	0.40	3	1	0.33	1		-1.04					
surface glycoprotein	C*0702	TRTQLPPAY	~~~QLPPA~	20	-1.53	19	0.80	61	0.27	5	1	0.20	1							-1.19
surface glycoprotein	C*0701	FRSSVLHST	~~~SVLHS~	43	-1.42	11	0.49	31	-0.50	6	1	0.17	1							-1.23
surface glycoprotein	A*2401	TQDLFLPFF	~~~LFLPF~	51	-1.90	20	-0.41	17	0.50	5	1	0.20	1		-2.21					
surface glycoprotein	A*2401	PFFSNVTWF	~~~SNVTW~	57	-2.15	22	-2.45	3	-1.79	4	1	0.25	1		-2.46					
surface glycoprotein	A*0201	TLDSKTQSL	~~~SKTQS~	109	-1.02	17	0.37	27	-0.46	77	13	0.17	7	-1.44						
surface glycoprotein	A*2401	YYHKNNKSW	~~~KNNKS~	144	-1.40	23	-0.69	20	0.45	4	1	0.25	1		-1.90					
surface glycoprotein	A*2401	SWMESEFRV	~~~ESEFR~	151	-1.70	21	-0.36	15	-1.13	3	1	0.33	1		-1.69					
surface glycoprotein	A*2401	EYVSQPFLM	~~~SQPFL~	169	-2.06	20	-0.11	11	-0.30	4	1	0.25	1		-2.41					
surface glycoprotein	C*0701	FKNLREFVF	~~~LREFV~	186	-1.90	21	-0.03	17	0.40	6	1	0.17	1							-1.54
surface glycoprotein	A*2401	VFKNIDGYF	~~~NIDGY~	193	-1.88	19	-2.04	3	-0.29	4	1	0.25	1		-2.06					
surface glycoprotein	B*0801	KIYSKHTP	~~~SKHTP~	202	-1.37	22	-3.18	2	-1.97	5	1	0.20	1				-1.75			
surface glycoprotein	B*0801	INITRFQTL	~~~TRFQT~	233	-1.93	22	-0.36	9	-0.89	5	1	0.20	1				-2.52			
surface glycoprotein	A*0201	YLQPRTFLL	~~~PRTFL~	269	-2.01	21	-0.20	23	-0.46	360	265	0.74	43	-3.37						
surface glycoprotein	A*2401	IYQTSNFRV	~~~TSNFR~	312	-1.74	22	0.08	14	-0.92	4	2	0.50	1		-1.49					
surface glycoprotein	C*0702	TRFASVYAW	~~~ASVYA~	345	-2.00	19	-0.41	22	0.46	8	1	0.13	1							-2.39
surface glycoprotein	A*0201	KIADYNYKL	~~~DYNYK~	417	-1.52	21	-3.80	0	-1.71	26	9	0.35	6	-2.29						
surface glycoprotein	A*2401	NYNYLYRLF	~~~YLYRL~	448	-1.95	22	-3.68	3	0.31	10	5	0.50	3		-2.75					
surface glycoprotein	C*0602	VGYQPYRVV	~~~QPYRV~	503	-1.88	16	-0.30	13	-2.99	3	1	0.33	1						0.67	
surface glycoprotein	C*0602	YQPYRVVVL	~~~YRVVV~	505	-1.99	17	-0.86	17	0.29	3	1	0.33	1						-0.11	
surface glycoprotein	A*0201	ELLHAPATV	~~~HAPAT~	516	-1.70	22	-0.47	8	-1.46	6	1	0.17	1	-2.18						
surface glycoprotein	B*1501	LVKNKCVNF	~~~NKCVN~	533	-1.65	24	-2.90	0	-1.88	3	1	0.33	1					-2.68		
surface glycoprotein	B*1501	VLTESNKKF	~~~ESNKK~	551	-1.59	23	0.49	17	-0.02	2	1	0.50	1					-1.59		
surface glycoprotein	B*0702	SPRRARSVA	~~~RARSV~	680	-1.49	13	0.42	22	-0.17	13	5	0.38	2			-3.18				
surface glycoprotein	A*0201	SIIAYTMSL	~~~AYTMS~	691	-1.71	22	-3.49	0	-1.26	58	22	0.38	6	-2.63						
surface glycoprotein	A*2401	AYSNNSIA	~~~NNSIA~	706	-1.56	23	-0.47	10	0.17	4	1	0.25	1		-1.76					
surface glycoprotein	A*0201	LLFNKVTLA	~~~NKVTL~	822	-1.80	22	-0.30	13	0.26	52	11	0.21	3	-2.99						
surface glycoprotein	A*0201	GLTVLPPLL	~~~VLPPL~	857	-1.80	20	1.23	69	0.89	9	2	0.22	2	-2.59						
surface glycoprotein	A*0201	ALNTLVKQL	~~~TLVKQ~	958	-1.26	23	-1.21	3	0.36	34	2	0.06	4	-2.13						
surface glycoprotein	A*0201	VLNDILSRL	~~~DILSR~	976	-1.46	17	0.35	27	1.17	136	10	0.07	6	-2.73						
surface glycoprotein	A*0201	RLDKVEAEV	~~~KVEAE~	983	-1.62	12	0.79	32	0.92	24	5	0.21	2	-1.89						
surface glycoprotein	A*0201	LITGRLQSL	~~~GRLQS~	997	-1.45	18	0.74	39	-0.02	34	2	0.06	2	-2.05						
surface glycoprotein	A*0201	RLQSLQTYV	~~~SLQTY~	1001	-1.28	19	0.24	22	-0.66	27	18	0.67	10	-2.70						
surface glycoprotein	A*0201	HLMSFPQSA	~~~SFPQS~	1048	-1.65	21	-1.06	11	-0.95	29	4	0.14	3	-2.39						
surface glycoprotein	C*0702	QSAPHGVVF	~~~PHGVV~	1054	-2.05	18	-0.86	6	-0.85	5	1	0.20	1							0.88
surface glycoprotein	A*0201	VVFLHVTYV	~~~LHVTY~	1060	-1.72	22	-0.69	10	-2.19	41	5	0.12	5	-2.64						
surface glycoprotein	A*2401	VFVSNGTHW	~~~SNGTH~	1094	-1.86	23	-0.36	11	-1.21	4	1	0.25	1		-1.93					
surface glycoprotein	A*2401	HWFVTQRNF	~~~VTQRN~	1101	-1.59	21	0.29	22	-1.09	4	1	0.25	1		-2.32					
surface glycoprotein	A*0201	RLNEVAKNL	~~~EVAKN~	1185	-1.55	22	0.24	22	0.86	165	5	0.03	5	-1.75						
surface glycoprotein	A*0201	NLNESLIDL	~~~ESLID~	1192	-1.29	21	0.11	29	0.58	141	11	0.08	3	-2.31						
surface glycoprotein	A*2401	QYIKWPWY	~~~KWPWY~	1208	-2.69	22	-3.18	0	-2.99	58	27	0.47	10		-2.88					
surface glycoprotein	A*0201	FIAGLIAIV	~~~GLIAI~	1220	-2.00	20	-0.25	30	1.43	214	39	0.18	12	-3.19						
surface glycoprotein	B*0702	EPVLKGVKL	~~~LKGVK~	1262	-1.50	22	-0.25	19	1.52	12	2	0.17	2			-2.50				

DTRA-TR-21-14

TECHNICAL REPORT - ADDENDA

Comparative Computational Modeling of Immune Reponses to Vaccines

Distribution Statement A. Approved for public release. Distribution unlimited.

HDTRA1-17-1-0052
Performer Name: ioGenetics, LLC
& Versiti, Inc.

Prepared by: Robert Bremel and
Jane Homan

June 2021

ioGenetics®

Addendum 1

YF VAX

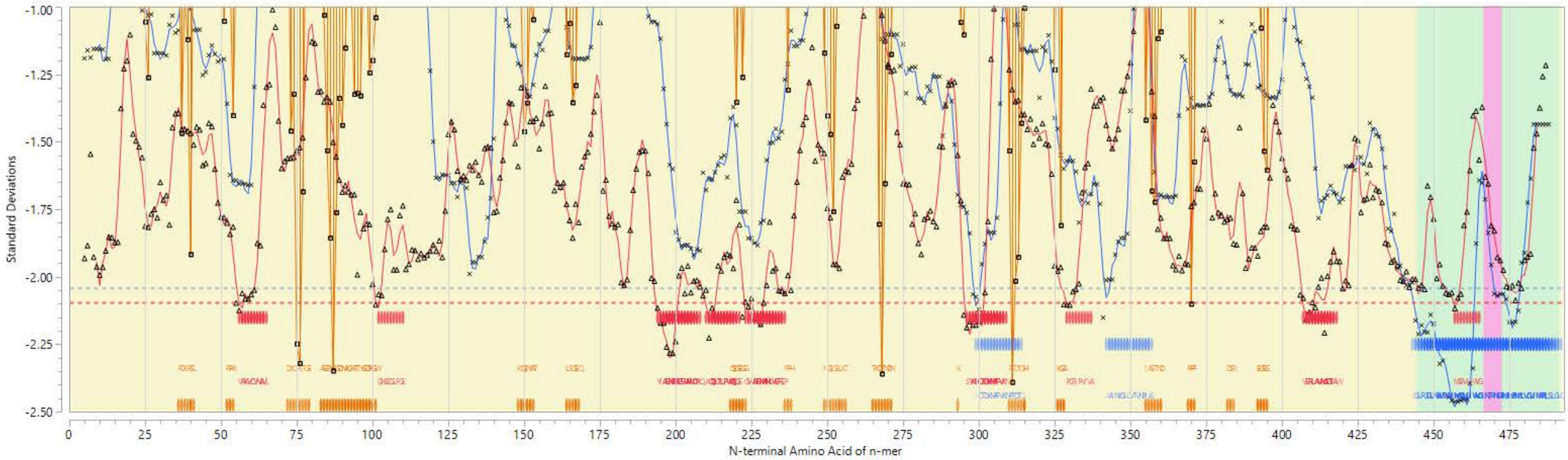
(see Appendix 2 for legends)

JN811140 envelope

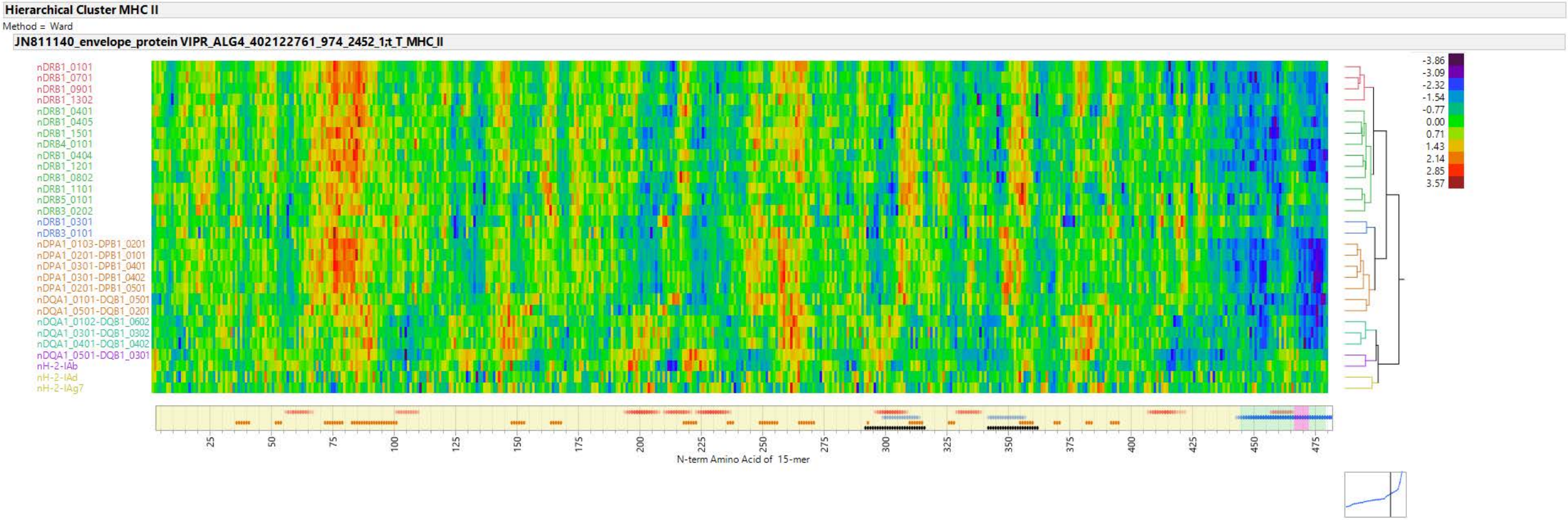
ioGenetics

Overlay Plot

JN811140_envelope_protein
VIPR_ALG4_402122761_974_2452_1;t

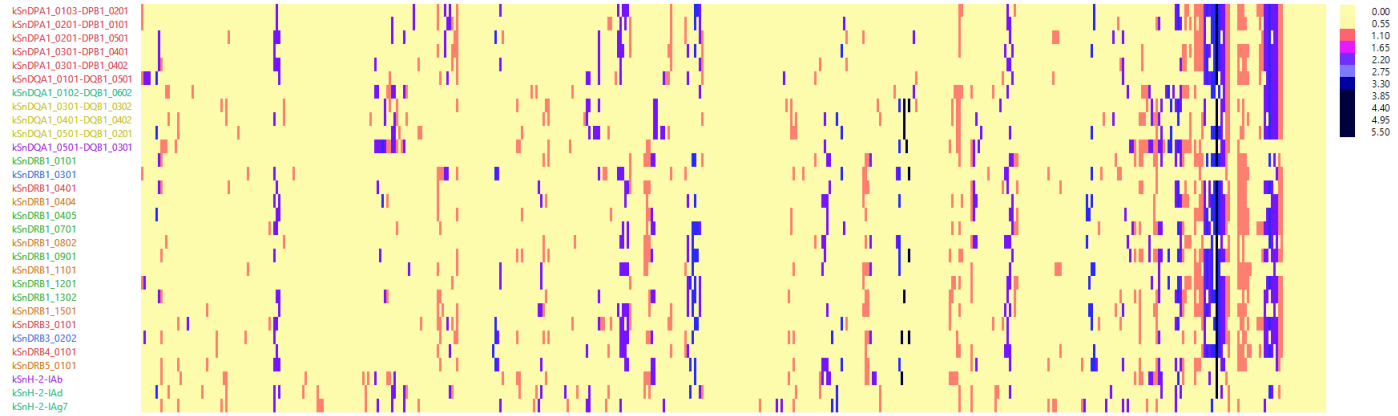


ioGenetics

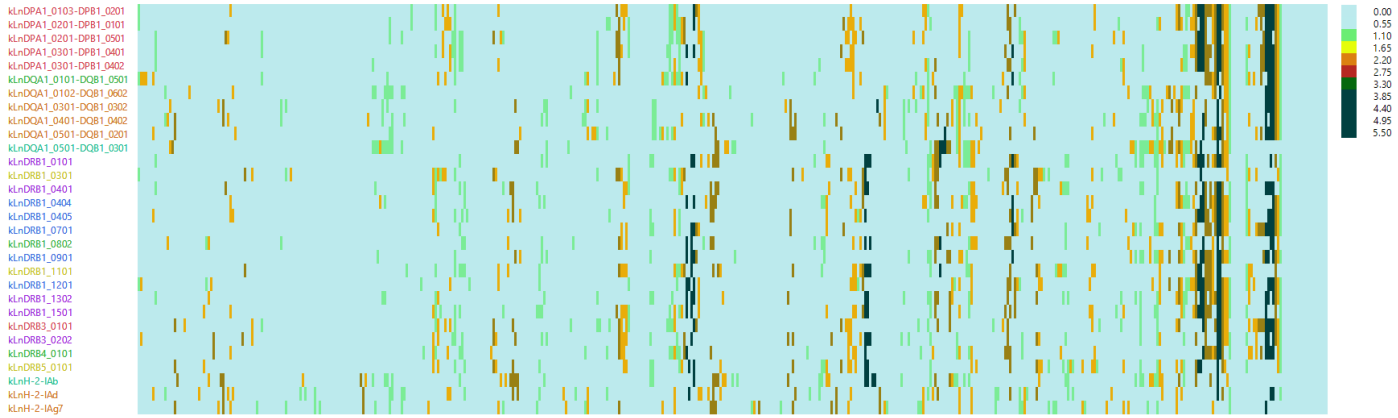

Hierarchical Cluster MHC II
Method = Ward
JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t_T_MHC_II
nDRB1_0101
nDRB1_0701
nDRB1_0901
nDRB1_1302
nDRB1_0401
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_0404
nDRB1_1201
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB3_0202
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.86
-3.09
-2.32
-1.54
-0.77
0.00
0.71
1.43
2.14
2.85
3.57
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

ioGenetics

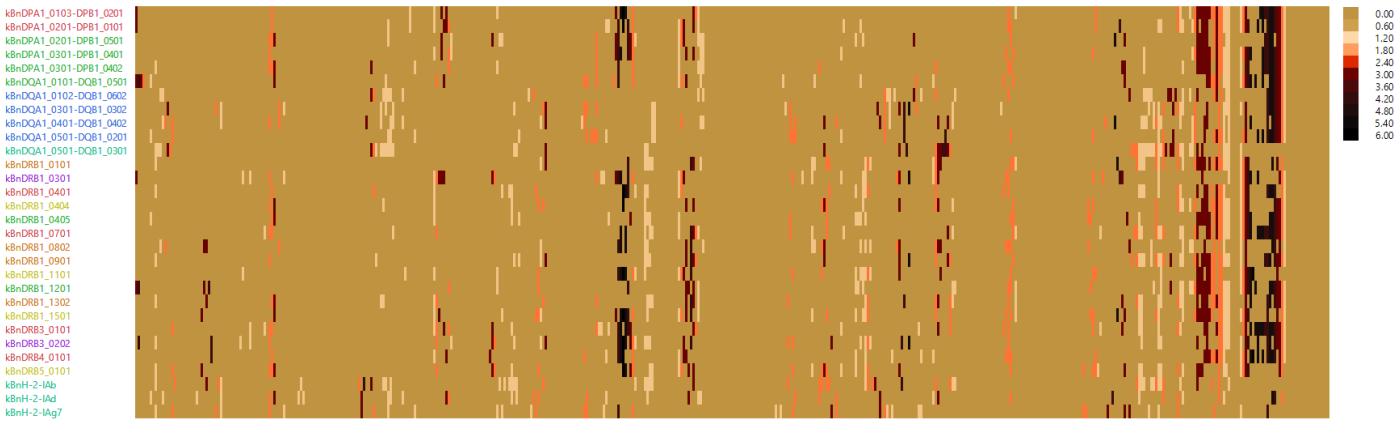
JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t Kernels MHC II @ 15-20 aa by hCAT_L



JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t Kernels MHC II @ 15-20 aa by hCAT_B

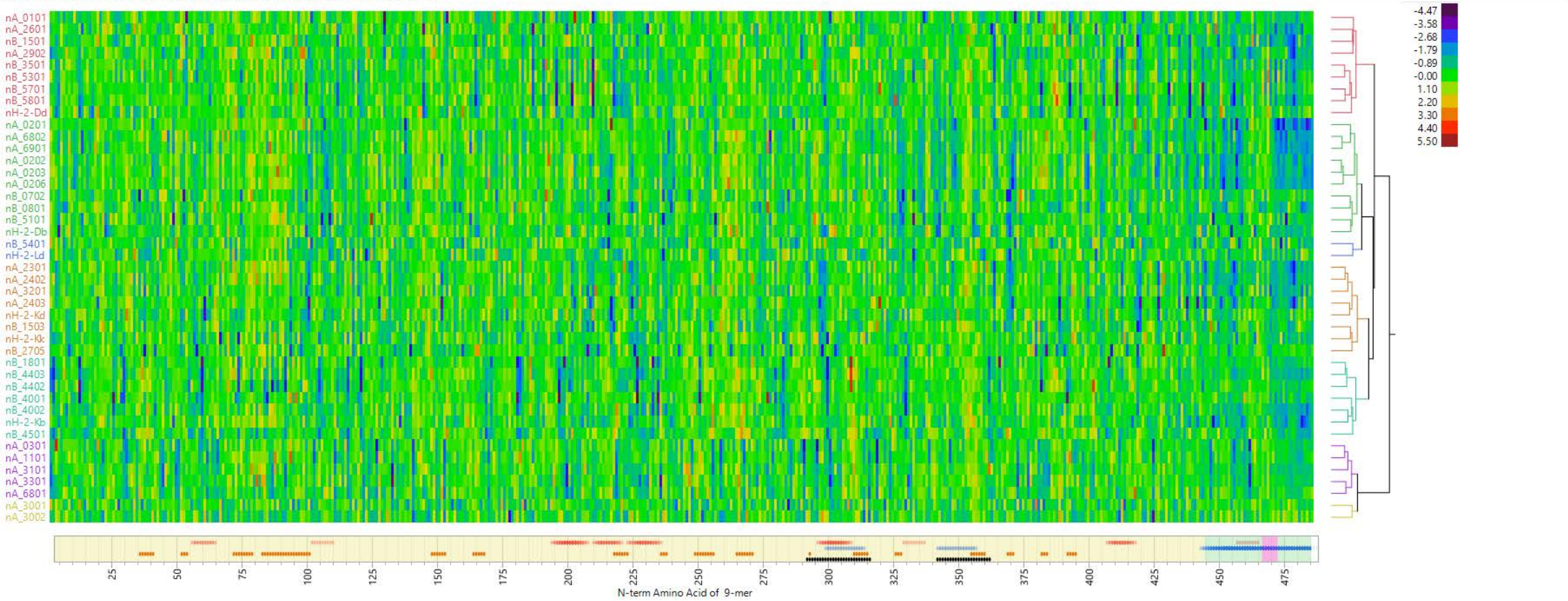


ioGenetics

Hierarchical Cluster MHC I

Method = Ward

JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t_T_MHC_I

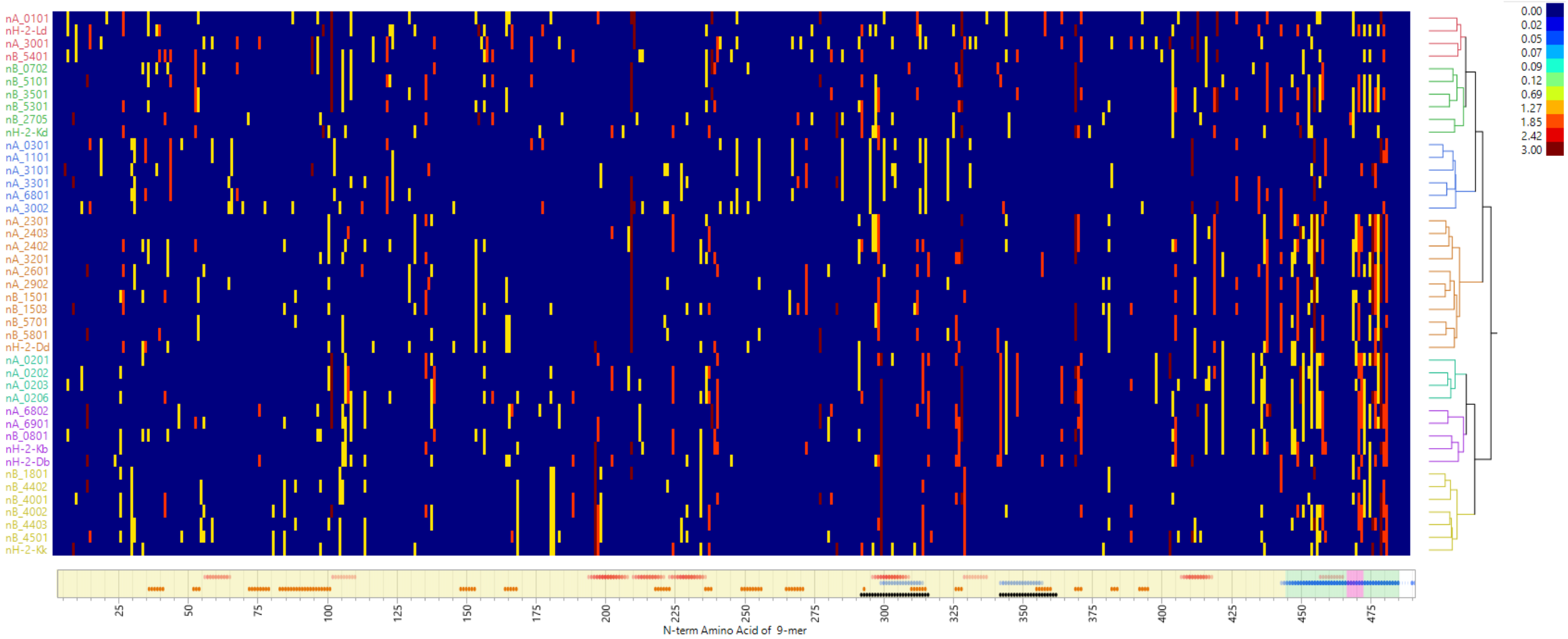


ioGenetics

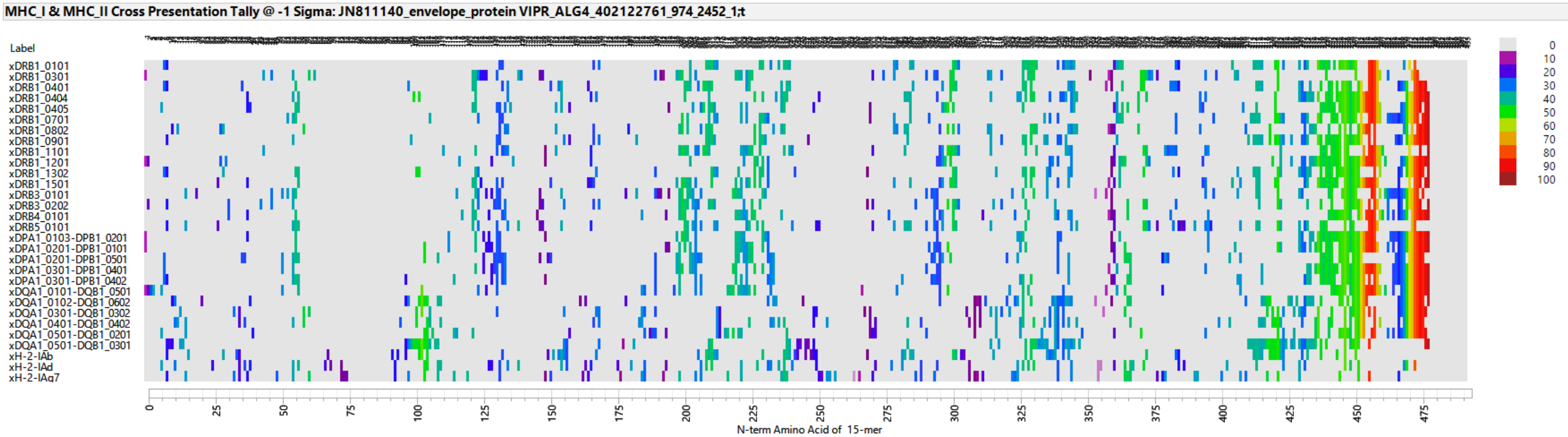
Hierarchical Cluster MHC I

Method = Ward

JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t MHC I & Cathepsins consolidated



ioGenetics


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_envelope_protein VIPR_ALG4_402122761_974_2452_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer
0
10
20
30
40
50
60
70
80
90
100

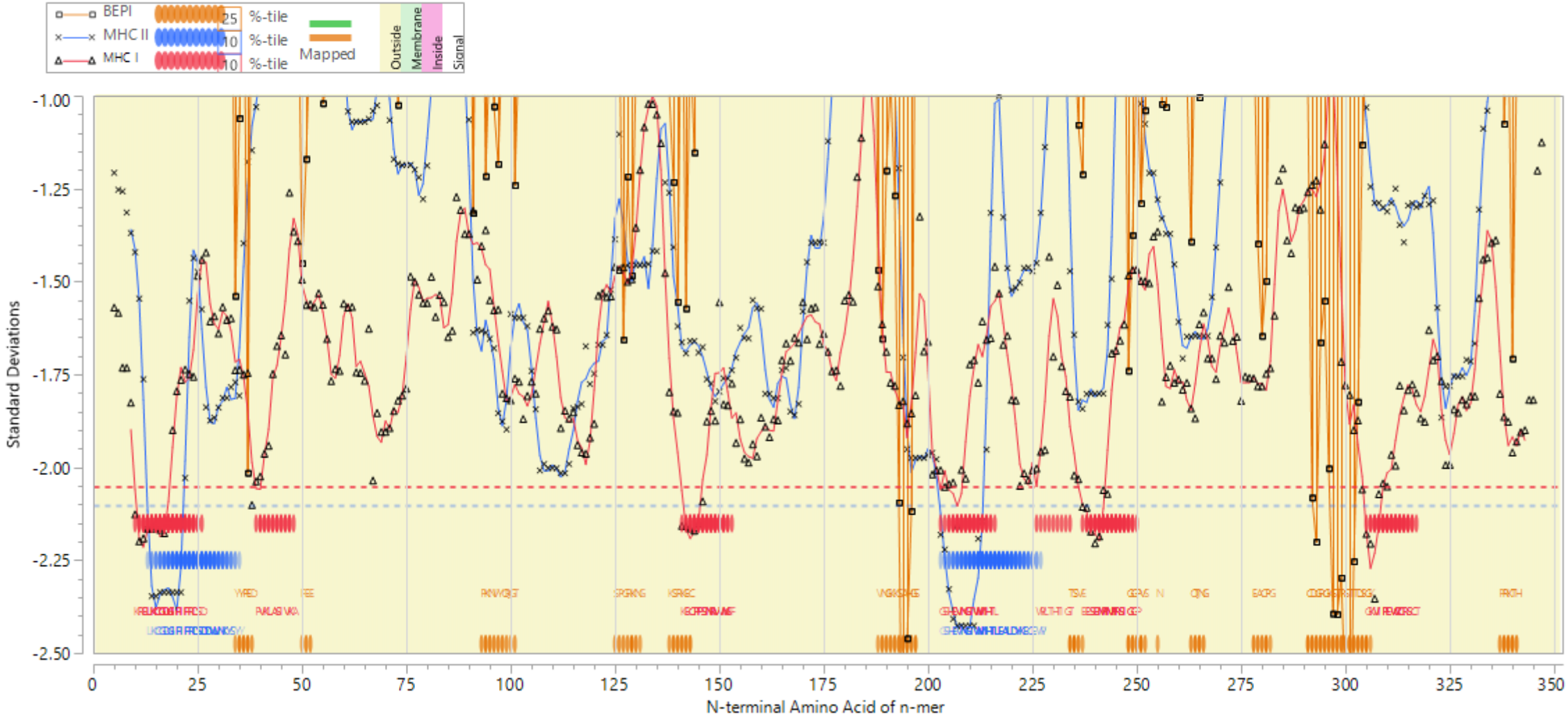
ioGenetics

JN811140_non_structural_protein_NS1

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS1
VIPR_ALG4_402122761_2453_3508_1;t

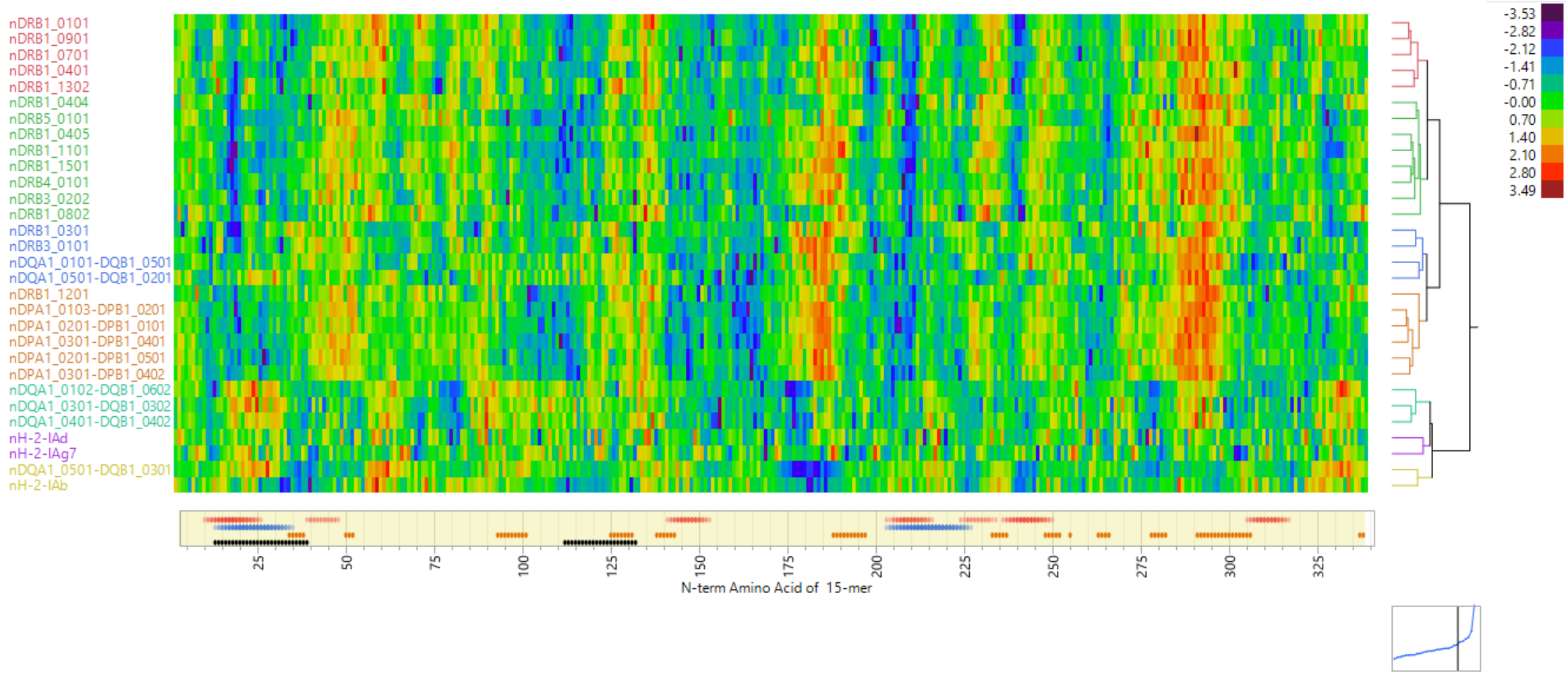


ioGenetics

Hierarchical Cluster MHC II

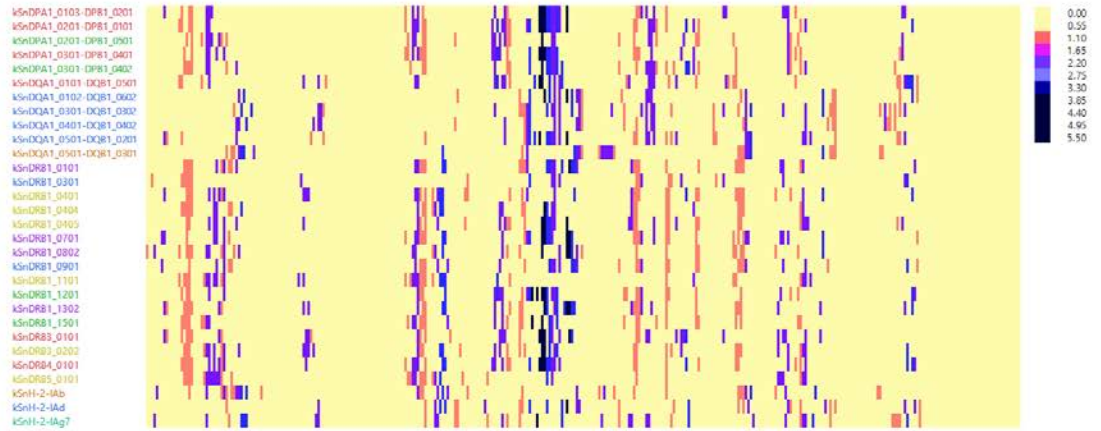
Method = Ward

JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t_T_MHC_II

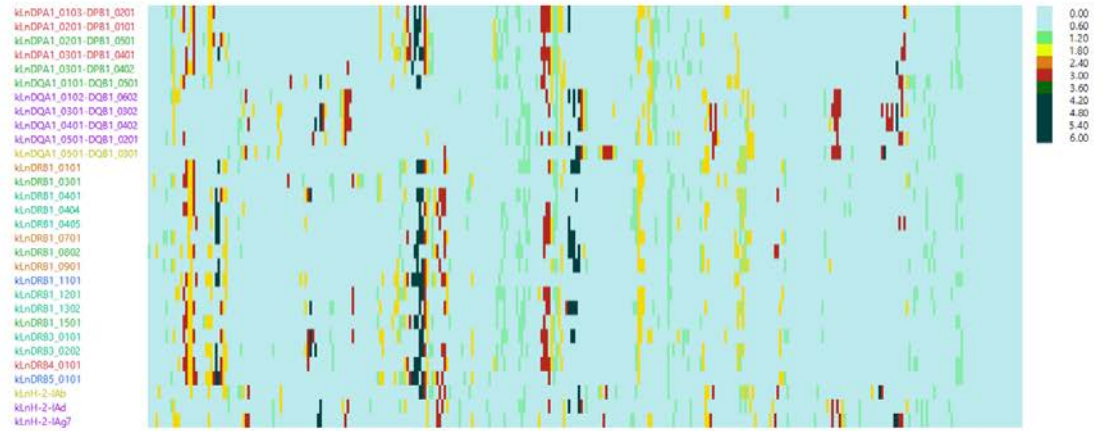


ioGenetics

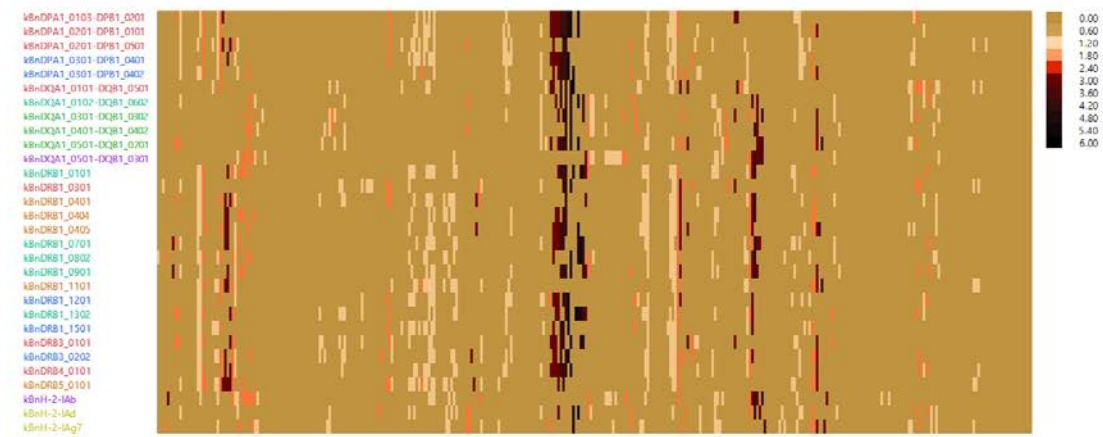
JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t Kernels MHC II @ 15-20 aa by hCAT_L



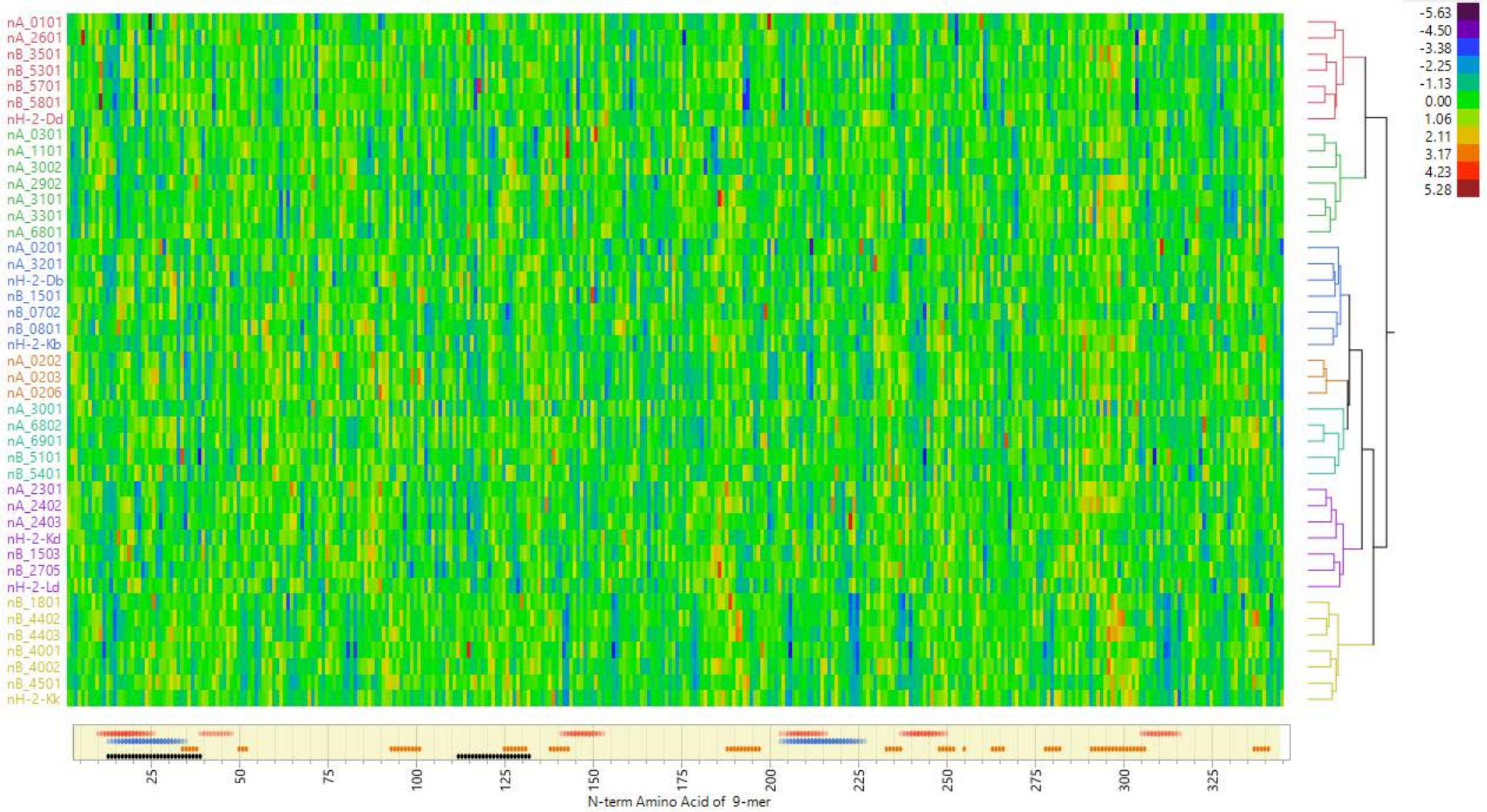
JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

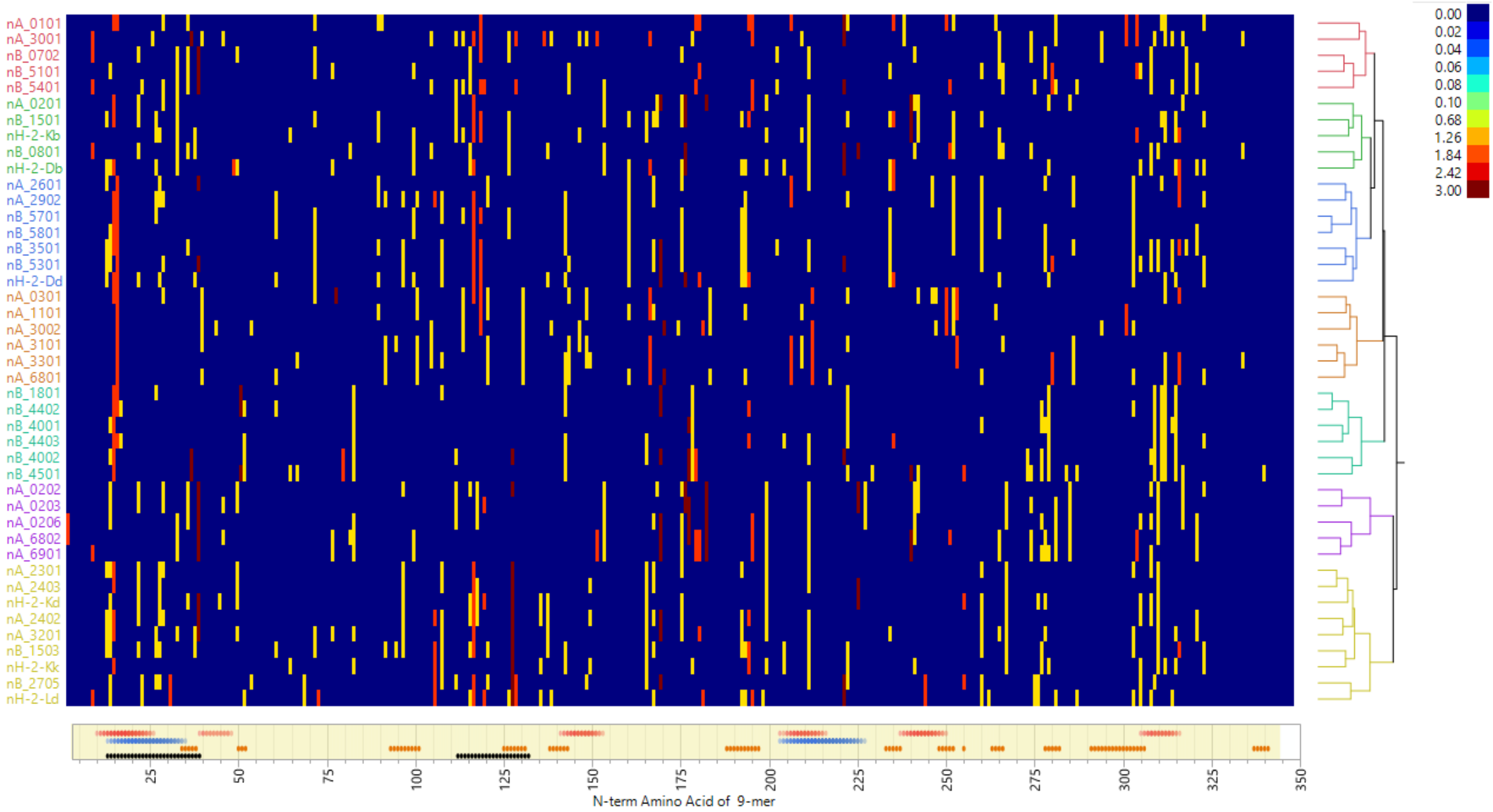
JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t_T_MHC_I

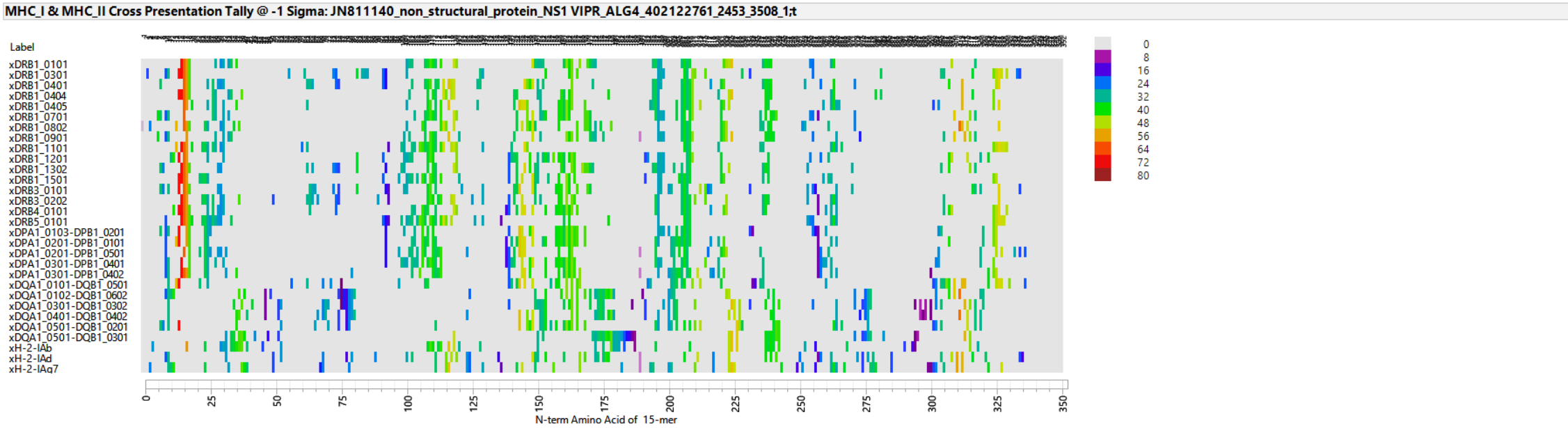


Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t MHC I & Cathepsins consolidated




MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS1 VIPR_ALG4_402122761_2453_3508_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

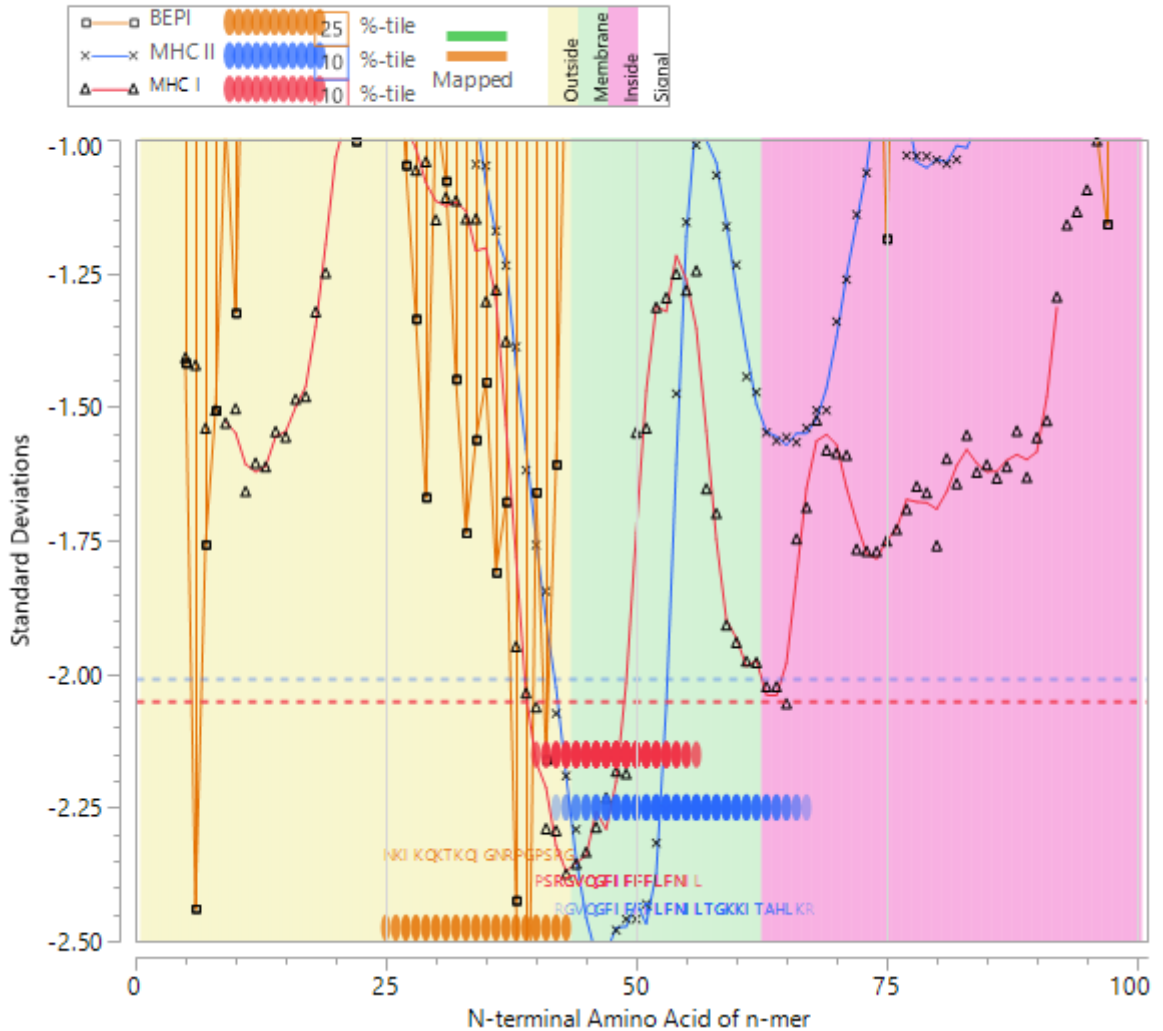
ioGenetics

JN811140_core_protein_C

ioGenetics

Overlay Plot

JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t

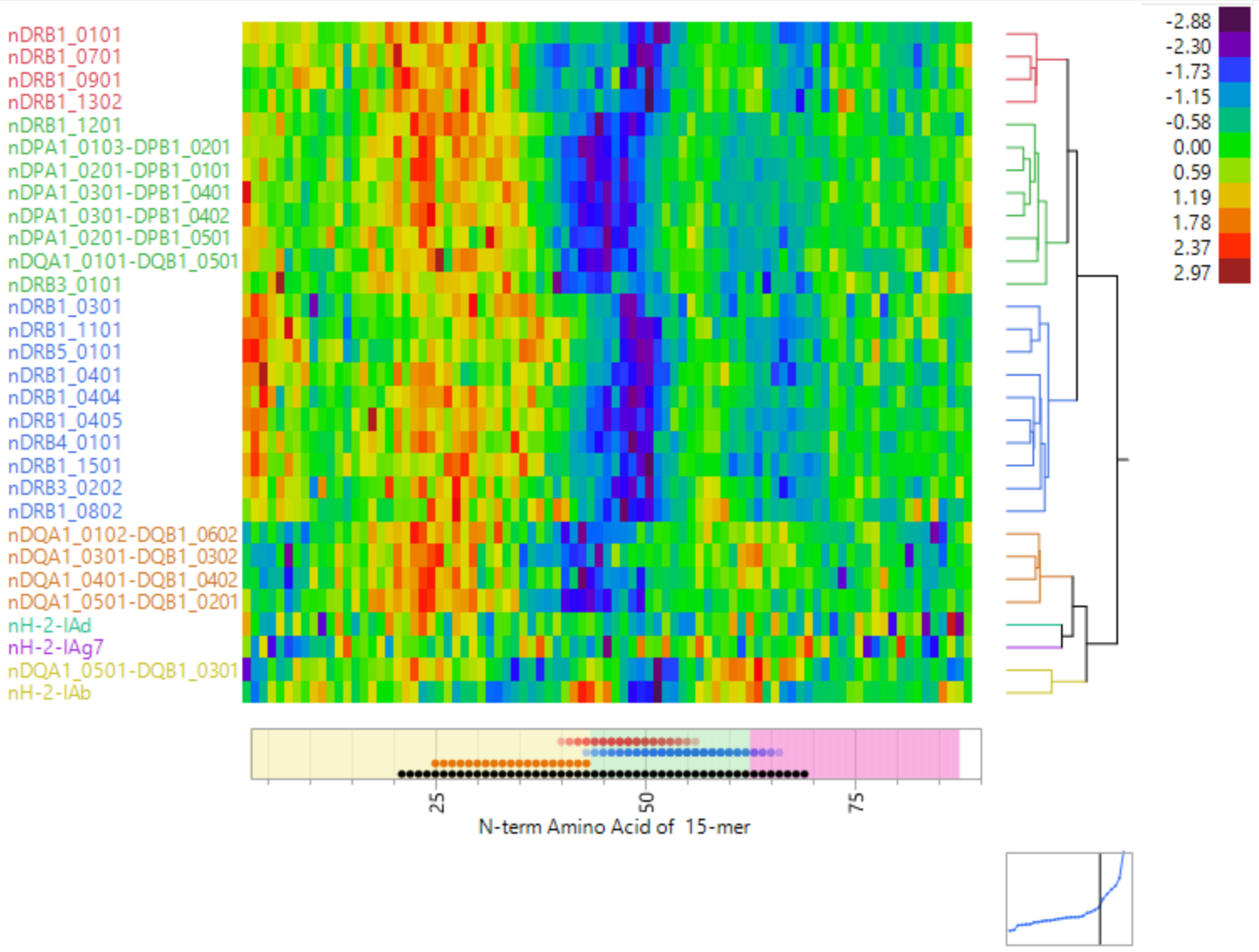


ioGenetics

Hierarchical Cluster MHC II

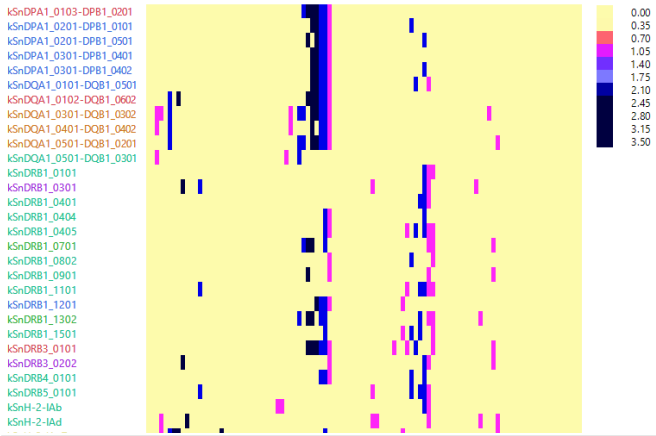
Method = Ward

JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t_T_MHC_II

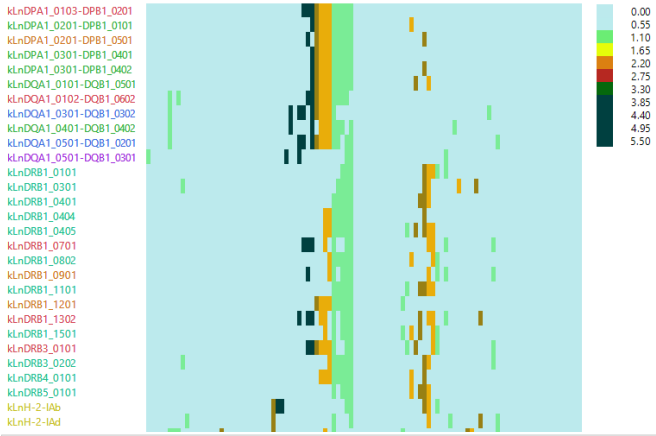


ioGenetics

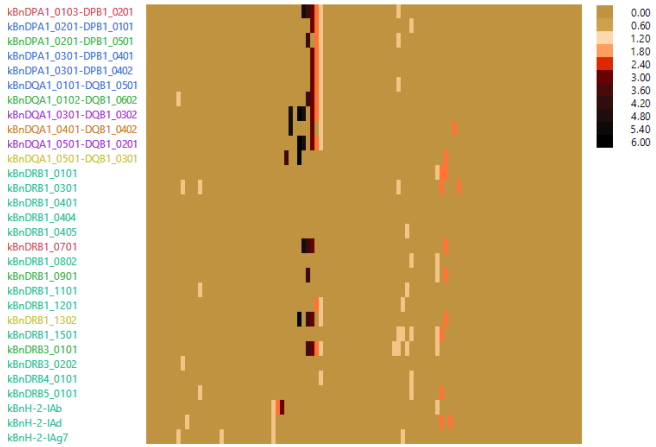
JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t Kernels MHC II @ 15-20 aa by hCAT_L



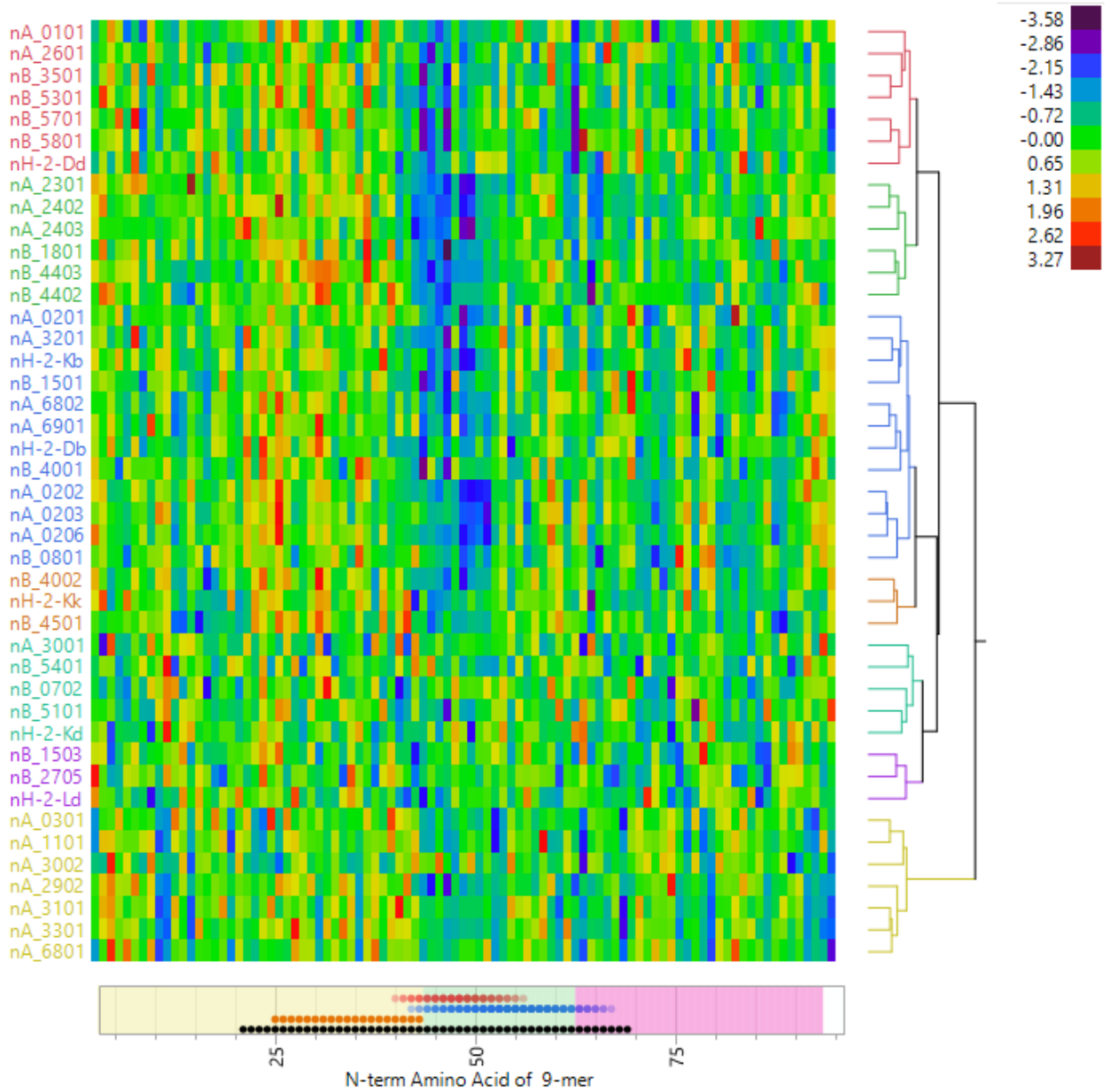
JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

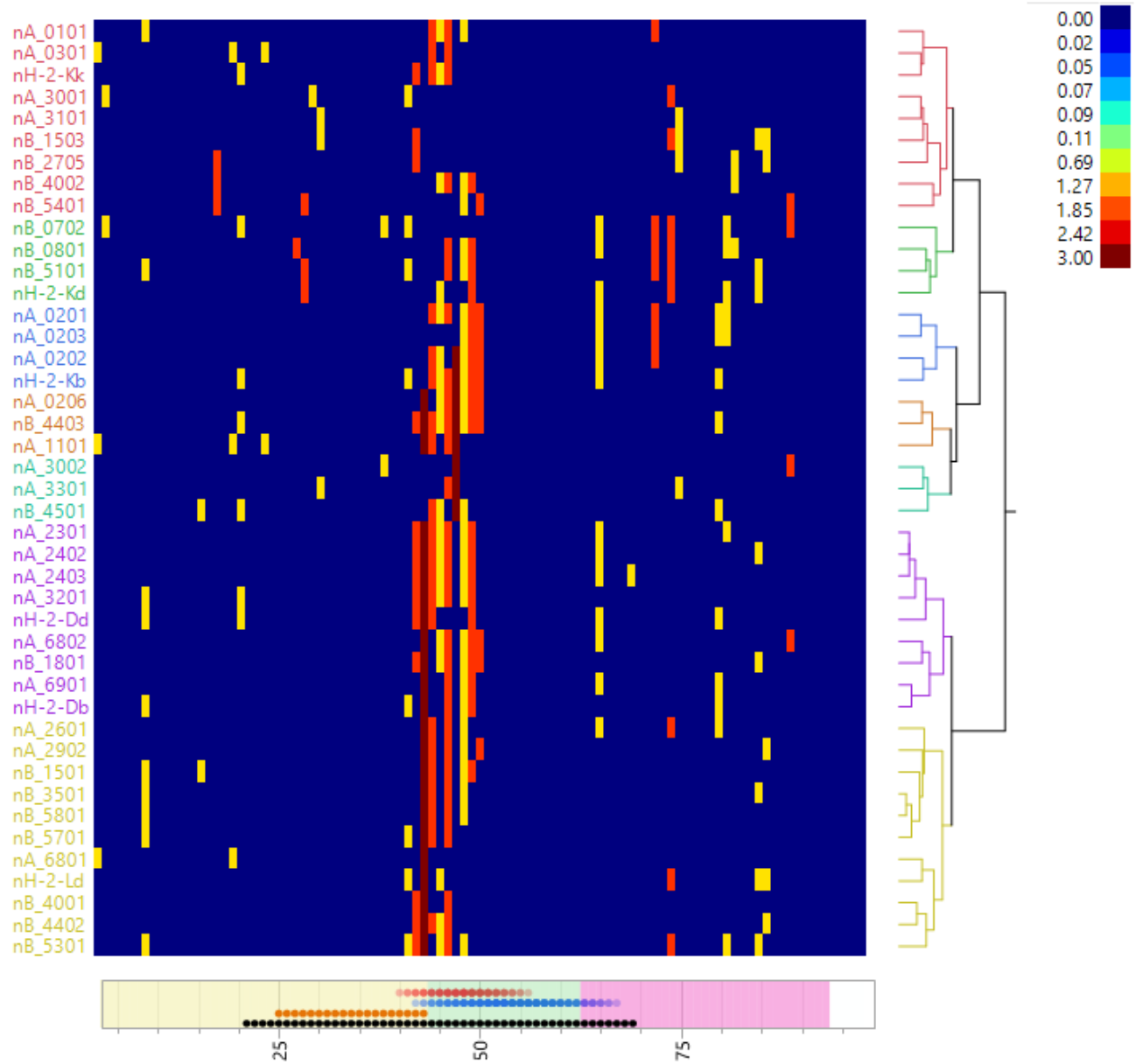
JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t_T_MHC I



Hierarchical Cluster MHC I

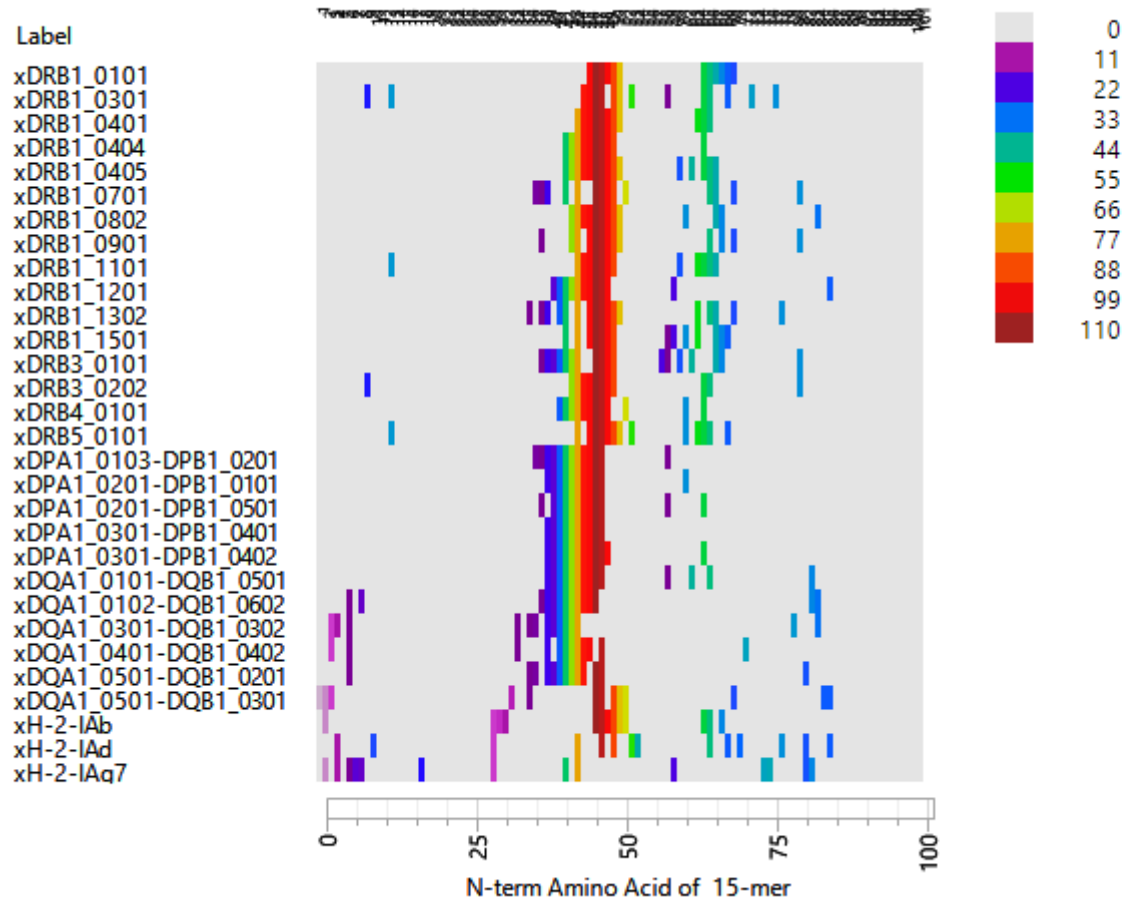
Method = Ward

JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t MHC I & Cathepsins consolidated



ioGenetics

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_core_protein_C VIPR_ALG4_402122761_119_421_1;t



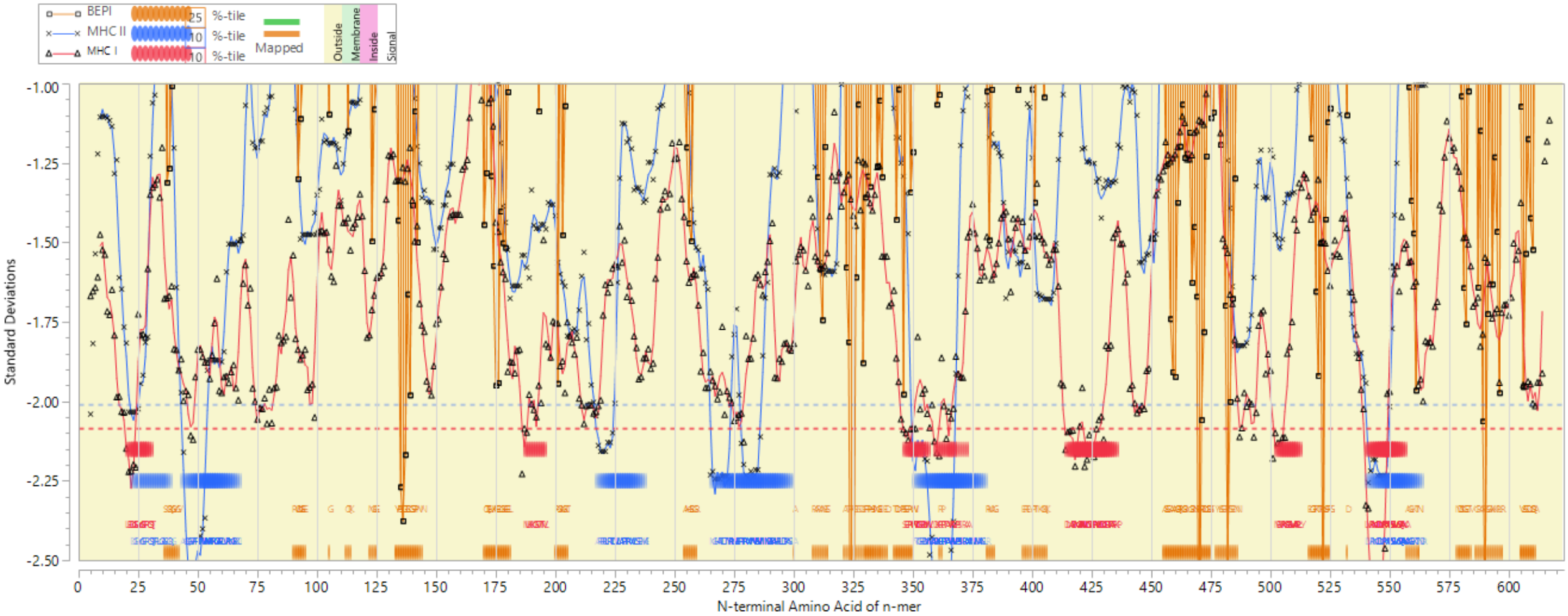
ioGenetics

JN811140_non_structural_protein_NS3

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS3
VIPR_ALG4_402122761_4571_6439_1;t

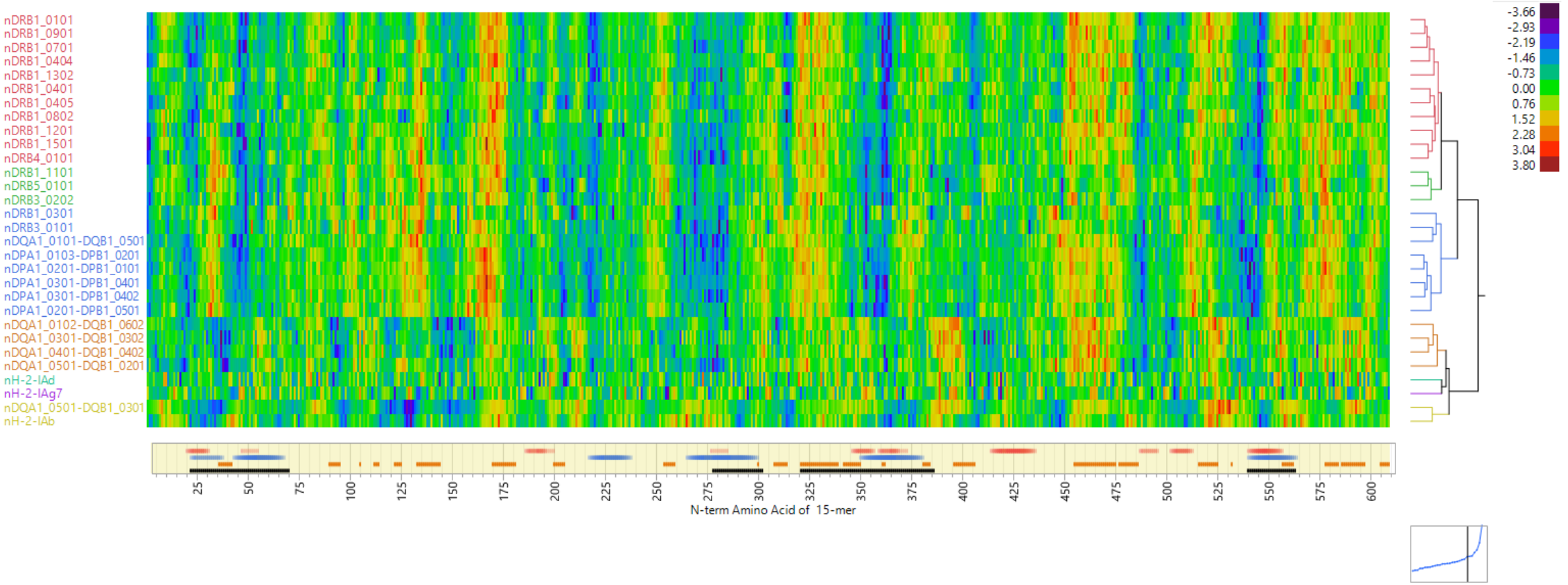


ioGenetics

Hierarchical Cluster MHC II

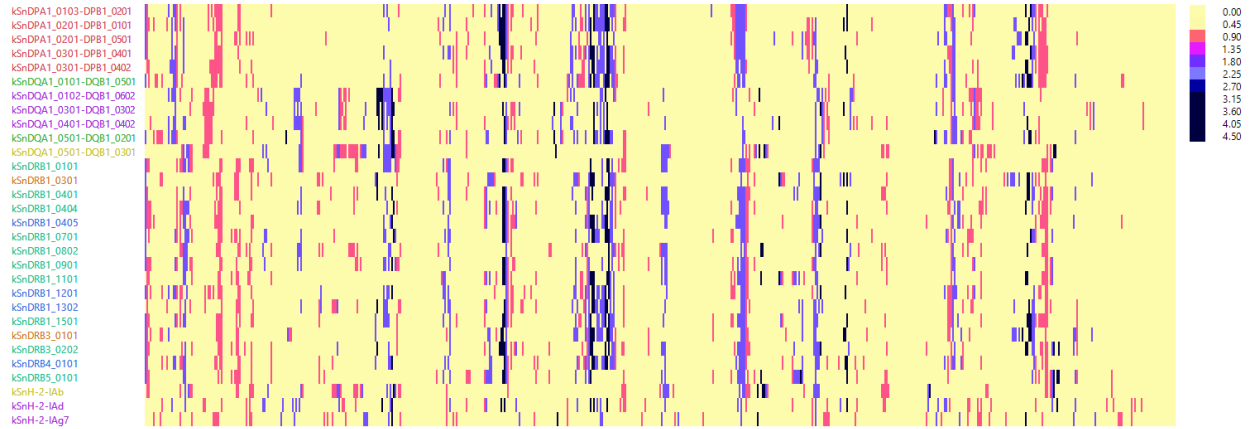
Method = Ward

JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t_T_MHC_II

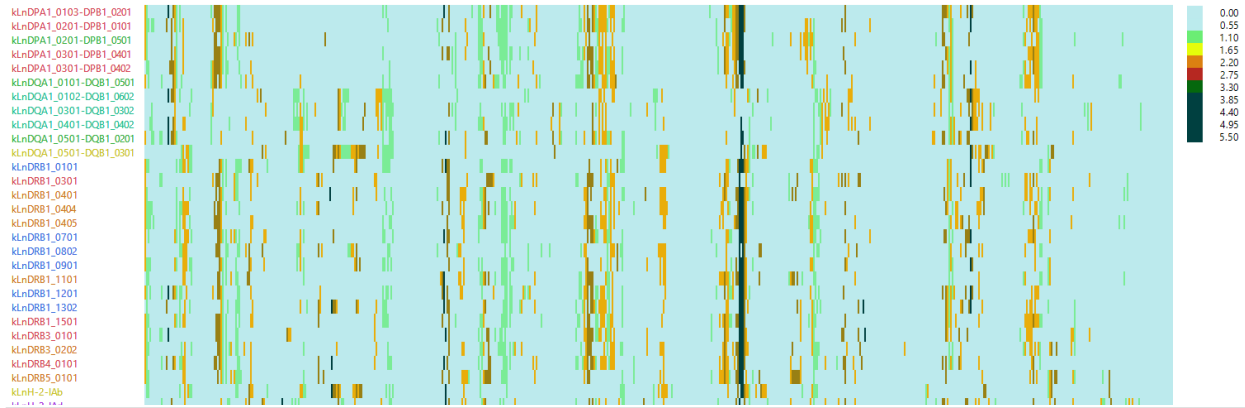


ioGenetics

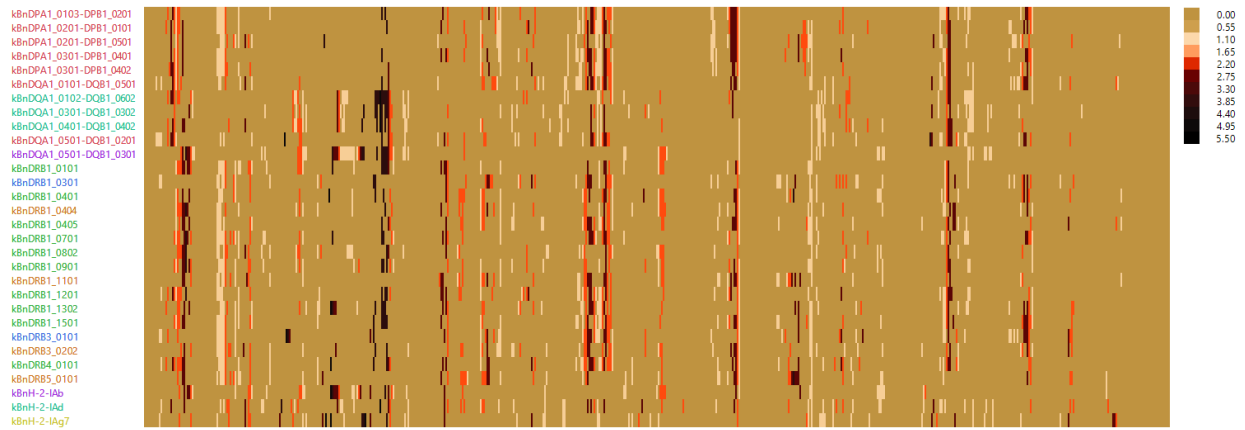
JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t Kernels MHC II @ 15-20 aa by hCAT_L



JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t Kernels MHC II @ 15-20 aa by hCAT_B

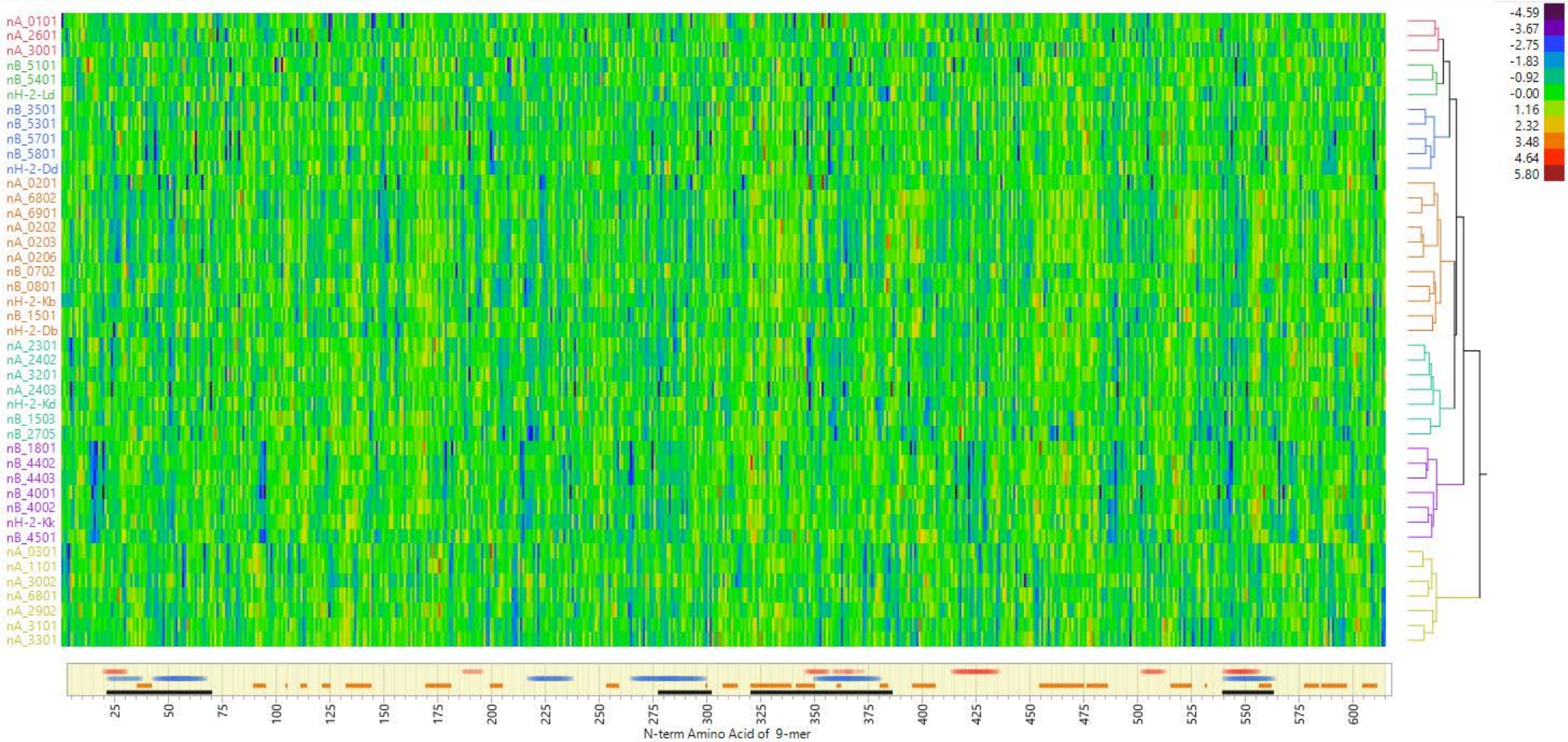


ioGenetics

Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t_T_MHC I

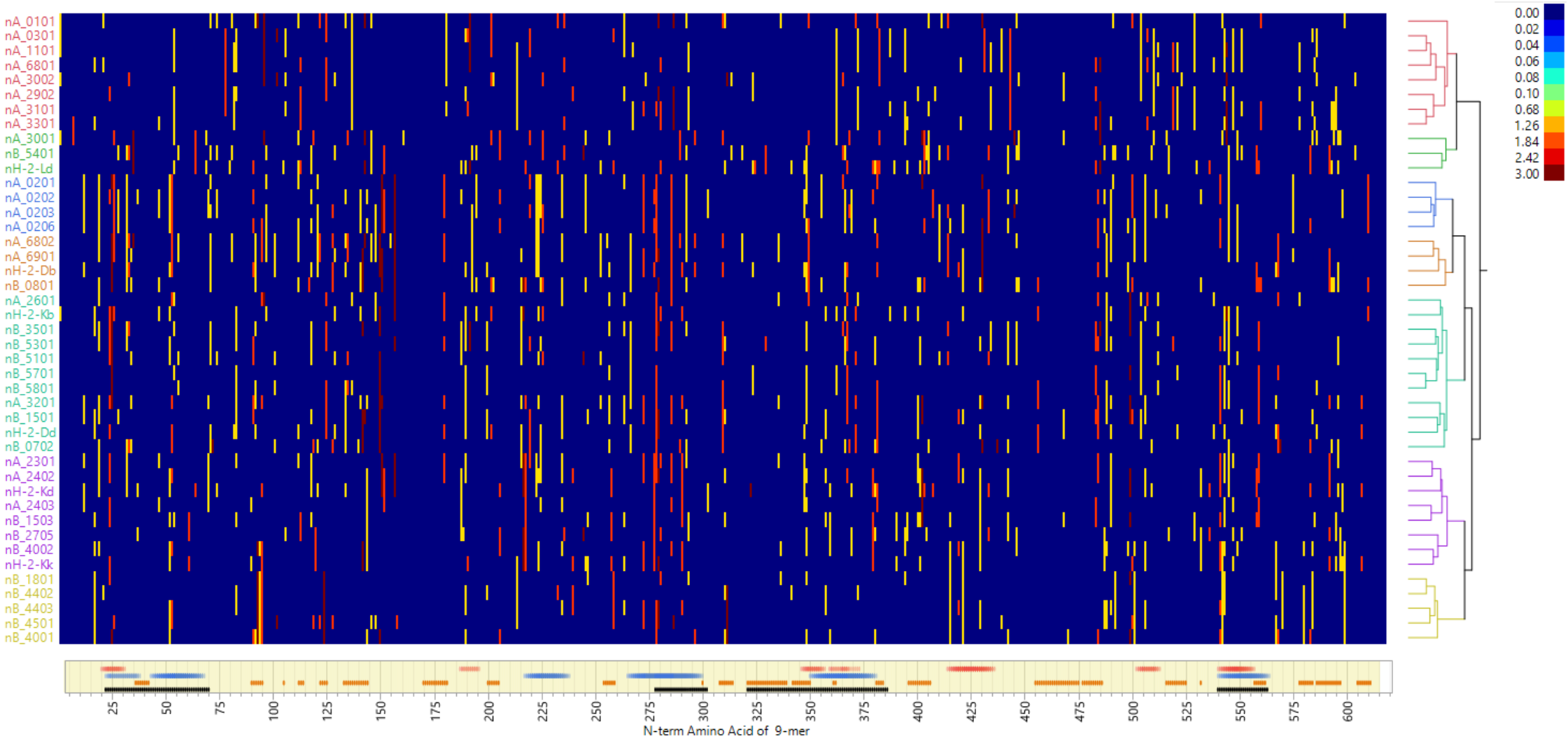


ioGenetics

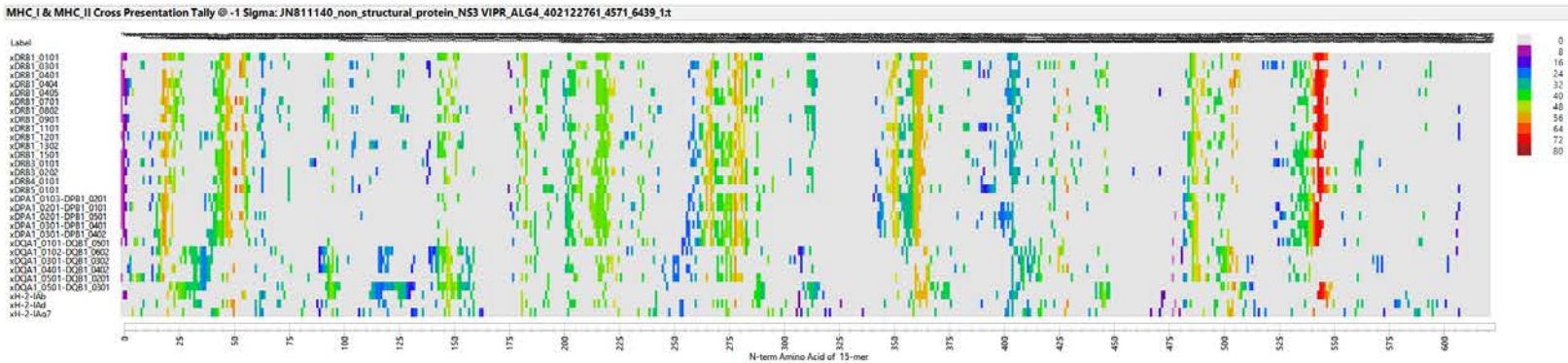
Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t MHC I & Cathepsins consolidated

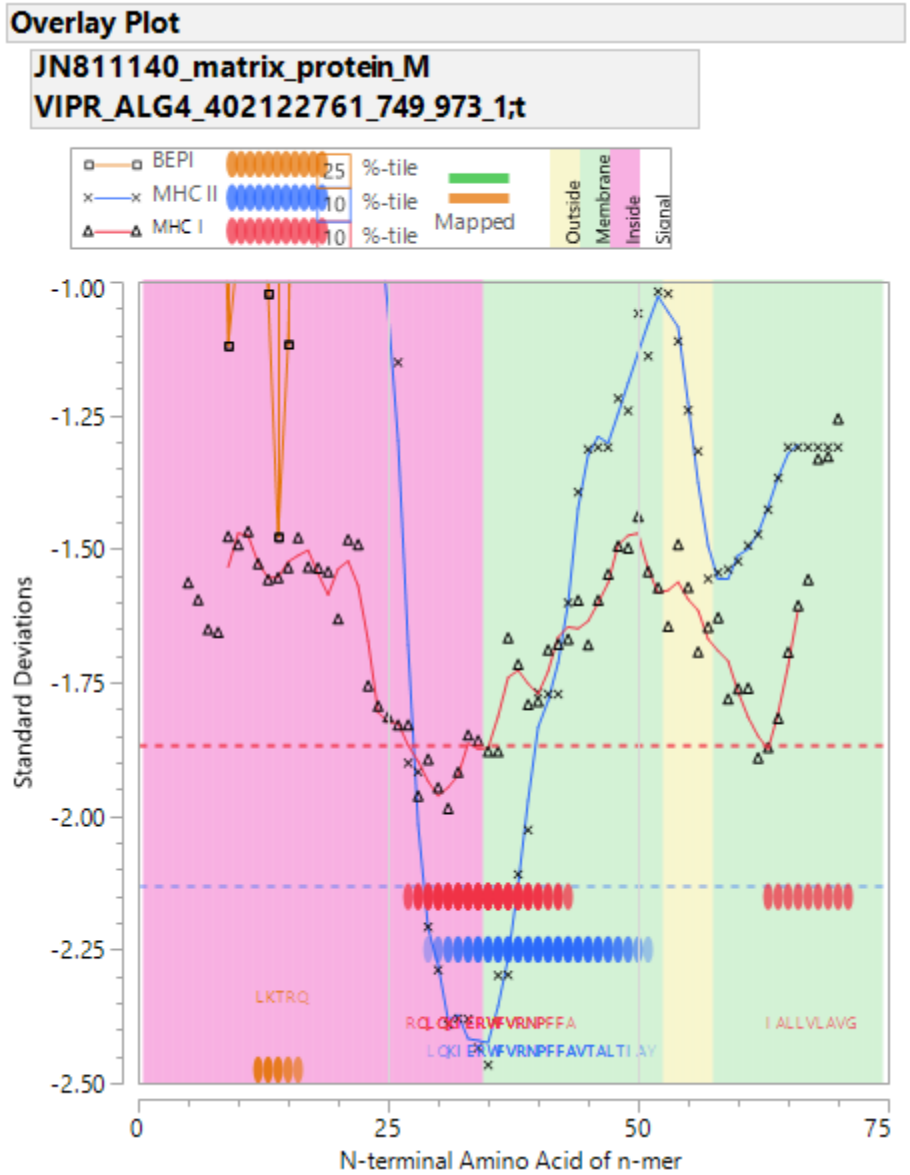


ioGenetics


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1:t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

JN811140_matrix_protein_M

ioGenetics

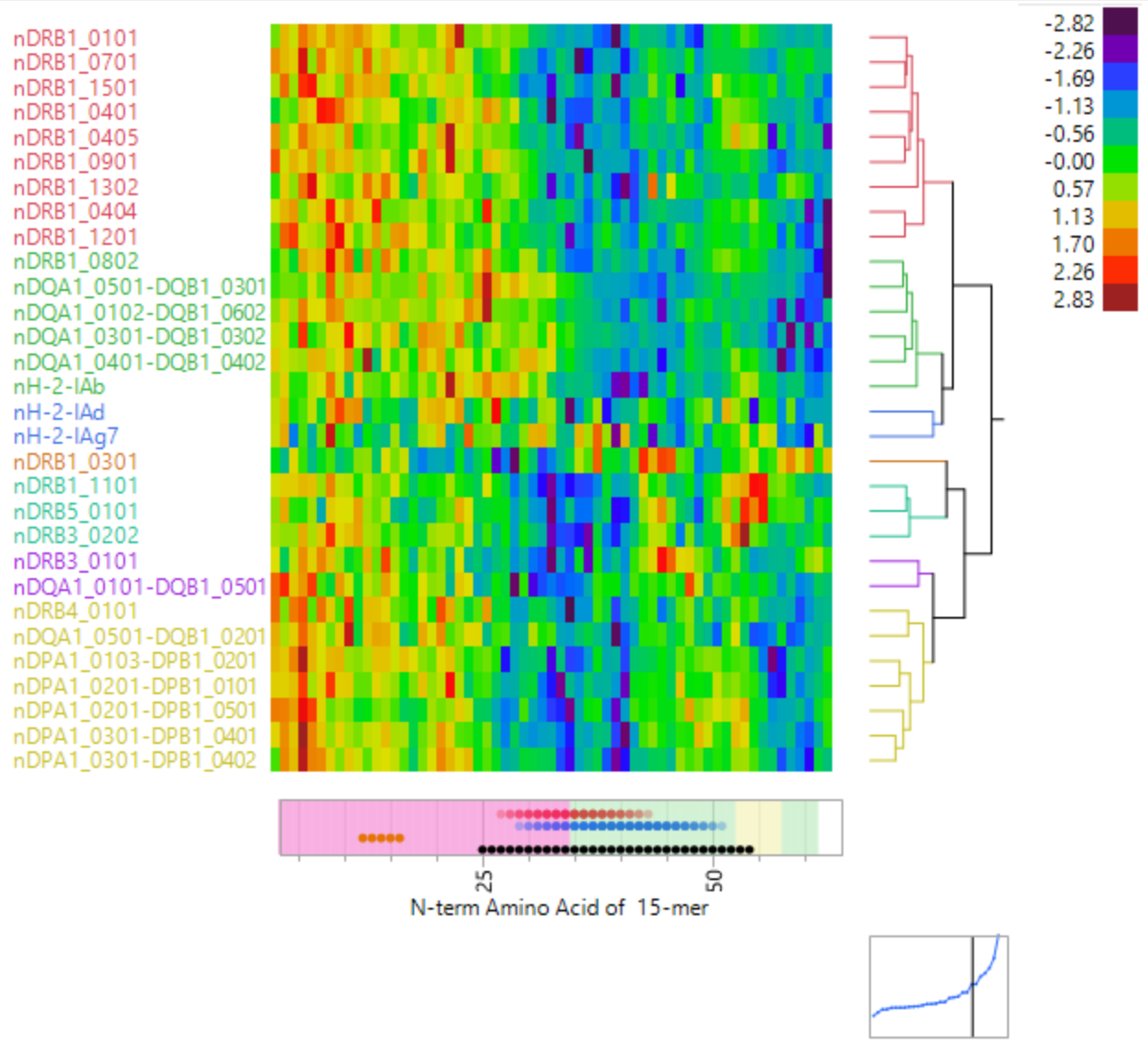

Overlay Plot
JN811140_matrix_protein_M
VIPR_ALG4_402122761_749_973_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
N-terminal Amino Acid of n-mer
LKTRQ
I ALLVLAVG

ioGenetics

Hierarchical Cluster MHC II

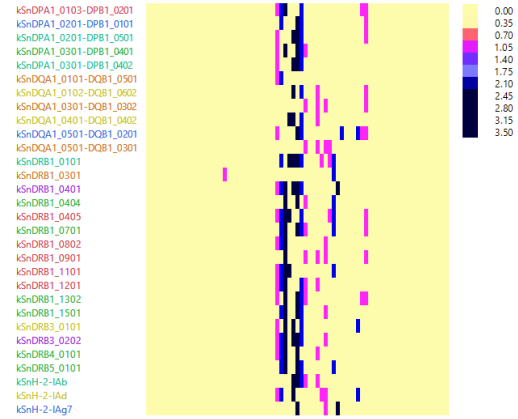
Method = Ward

JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t_T_MHC_II

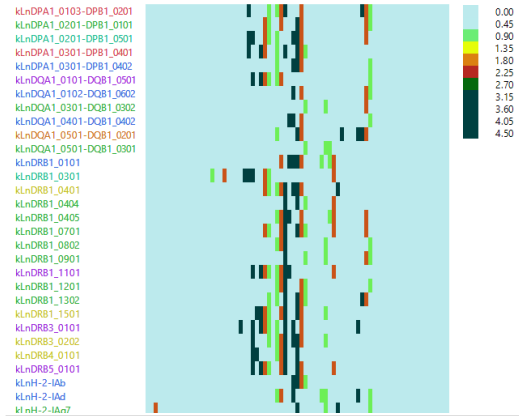


ioGenetics

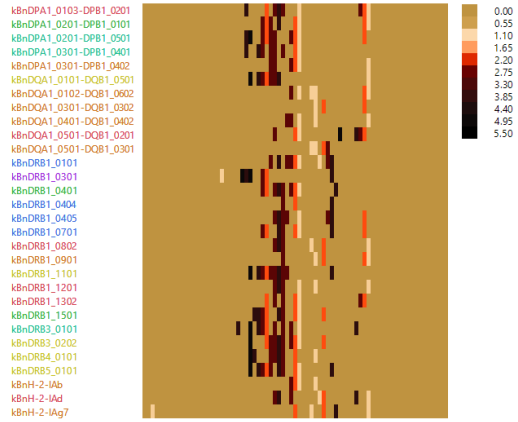
JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t Kernels MHC II @ 15-20 aa by hCAT_L



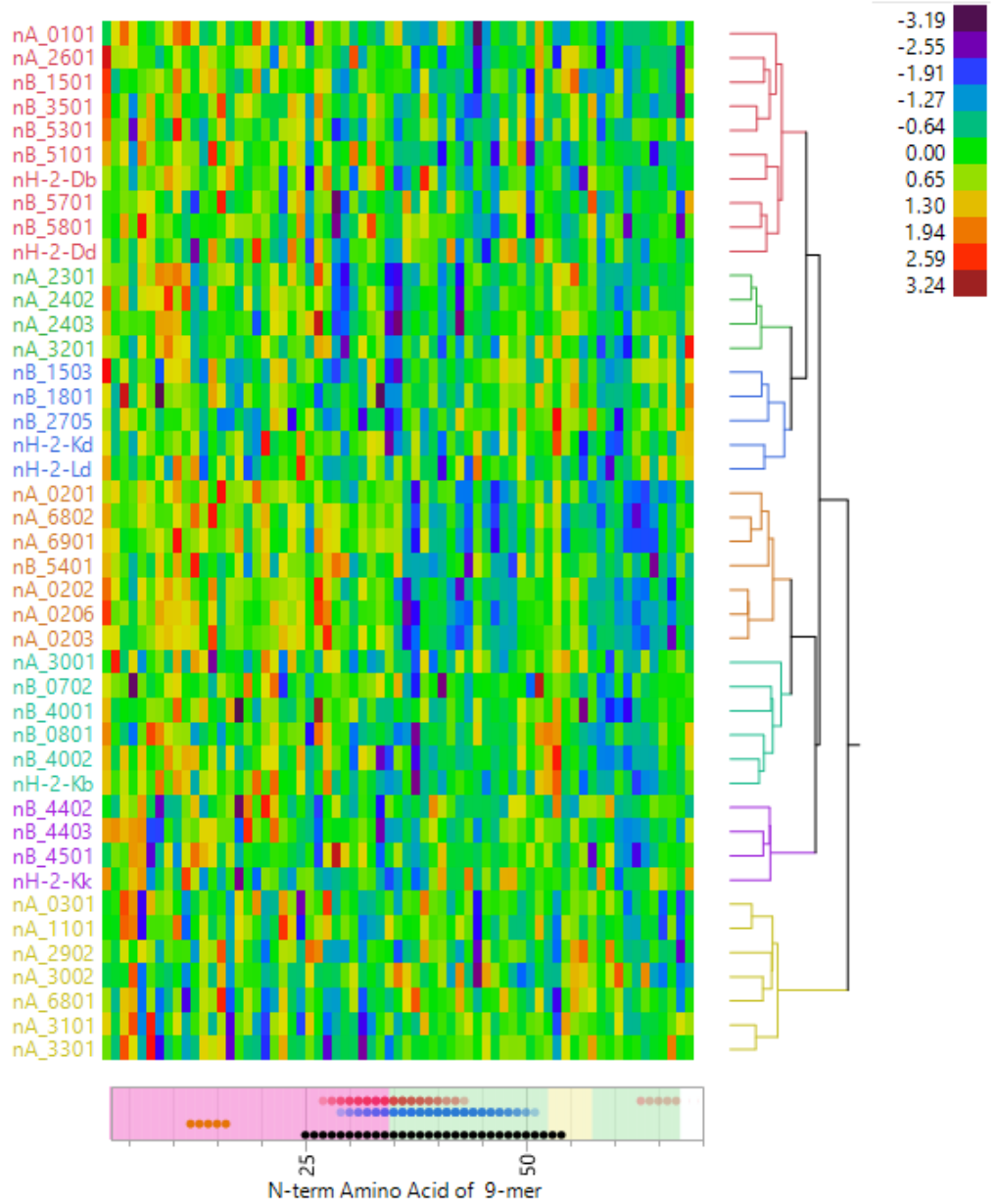
JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t_T_MHC_I

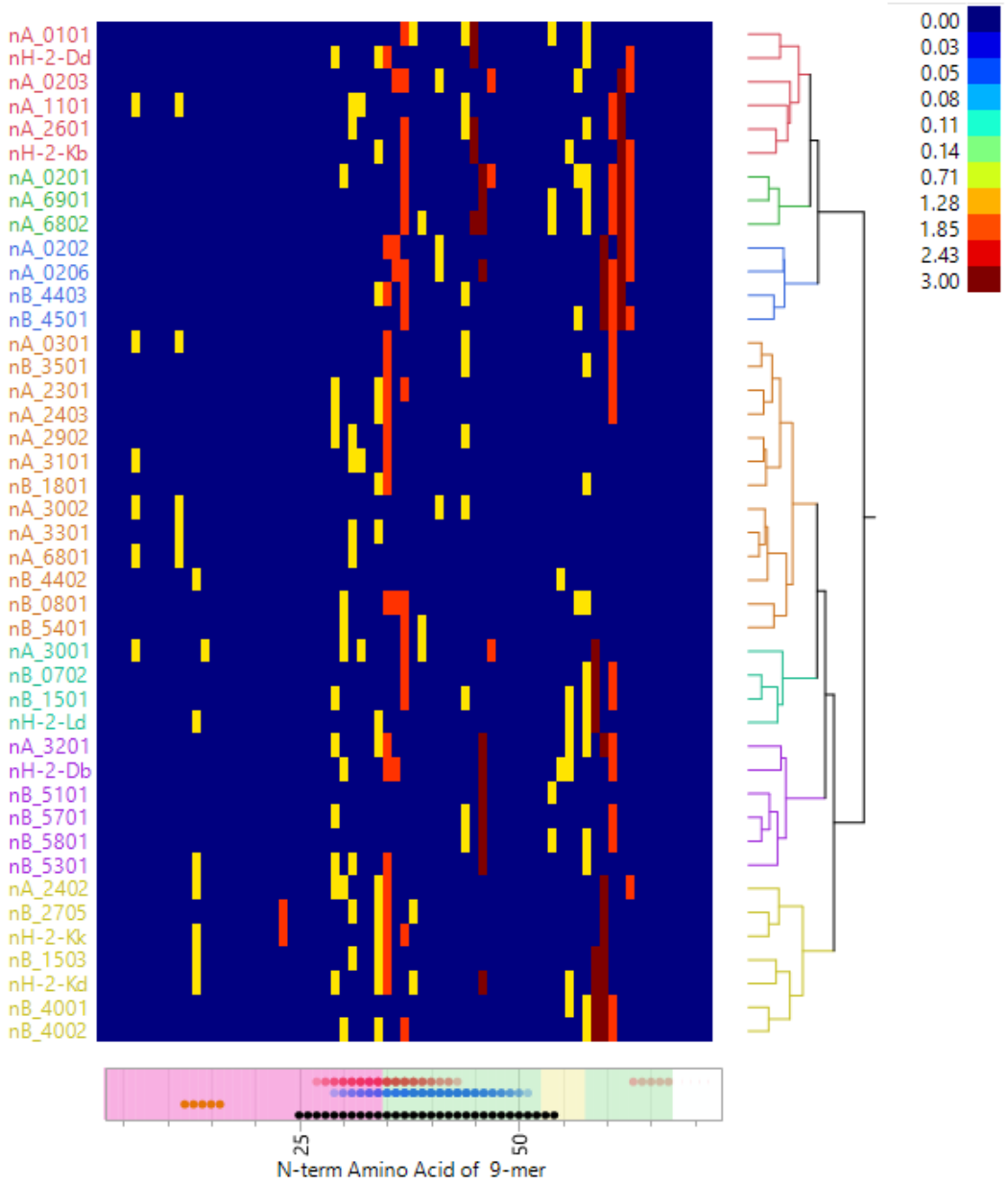


ioGenetics

Hierarchical Cluster MHC I

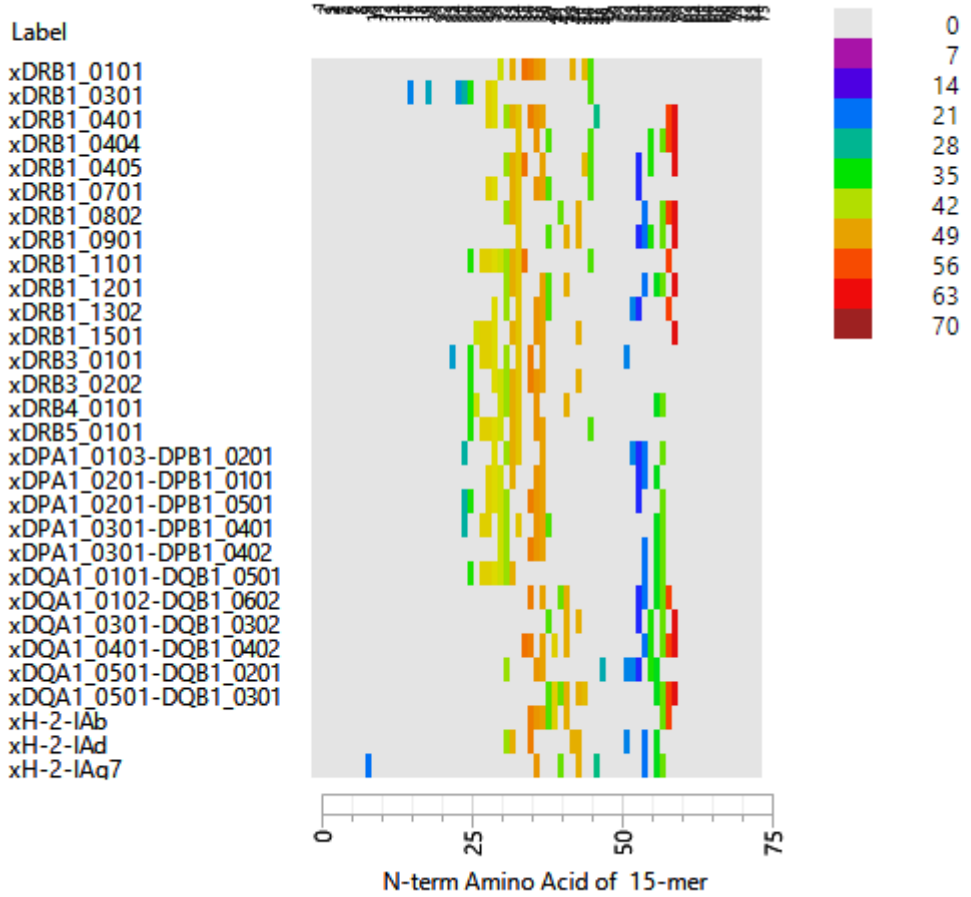
Method = Ward

JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t MHC I & Cathepsins consolidated 2



ioGenetics

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_matrix_protein_M VIPR_ALG4_402122761_749_973_1;t



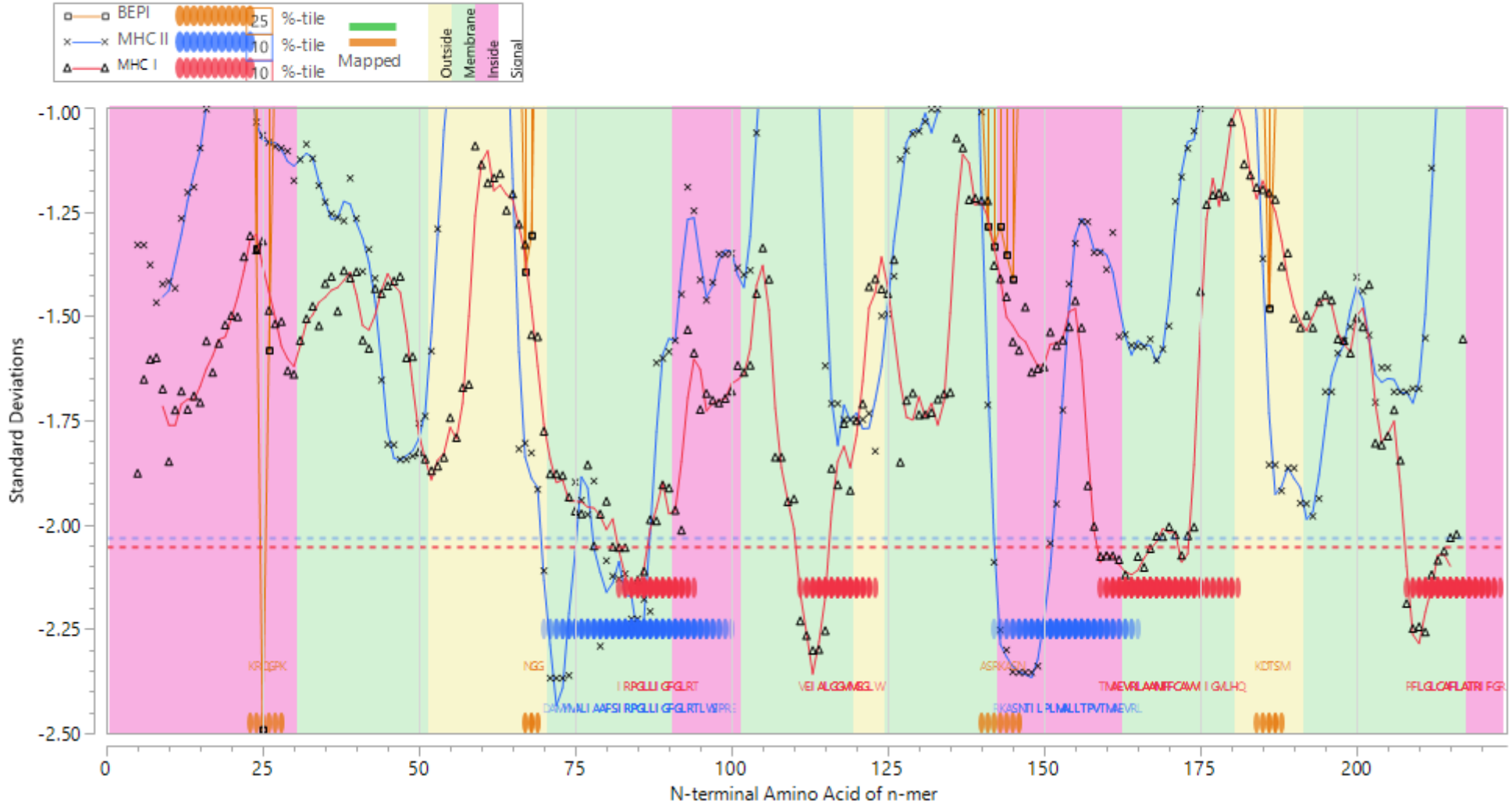
ioGenetics

JN811140_non_structural_protein_NS2a

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS2a
VIPR_ALG4_402122761_3509_4180_1;t

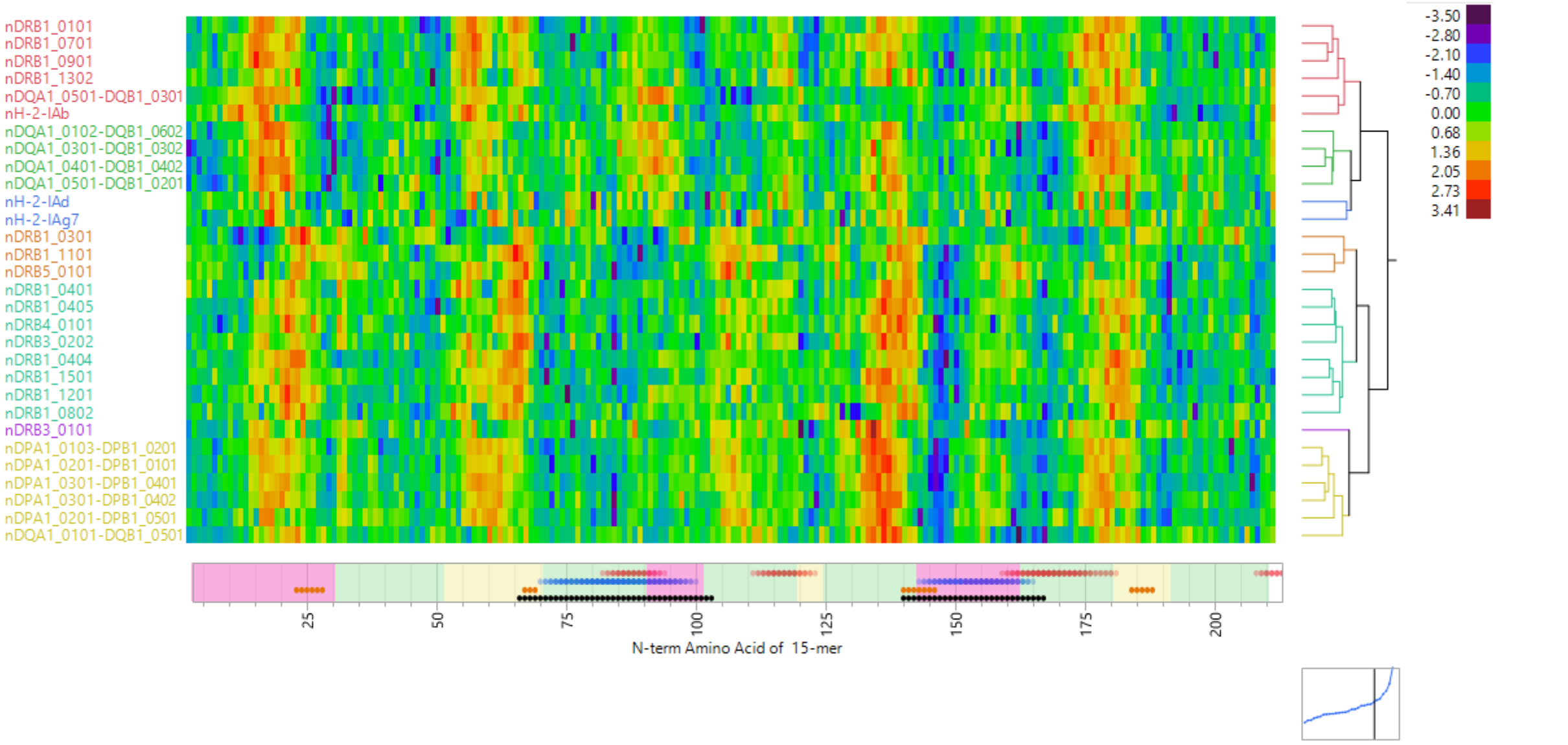


ioGenetics

Hierarchical Cluster MHC II

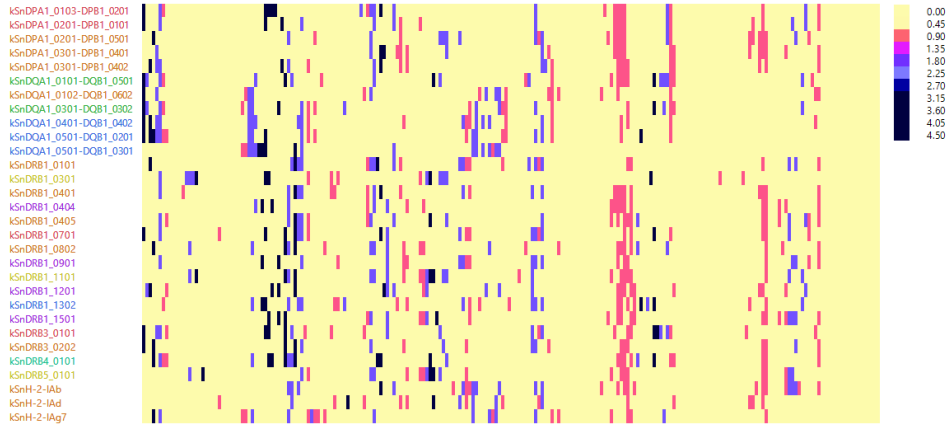
Method = Ward

JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t_T_MHC_II

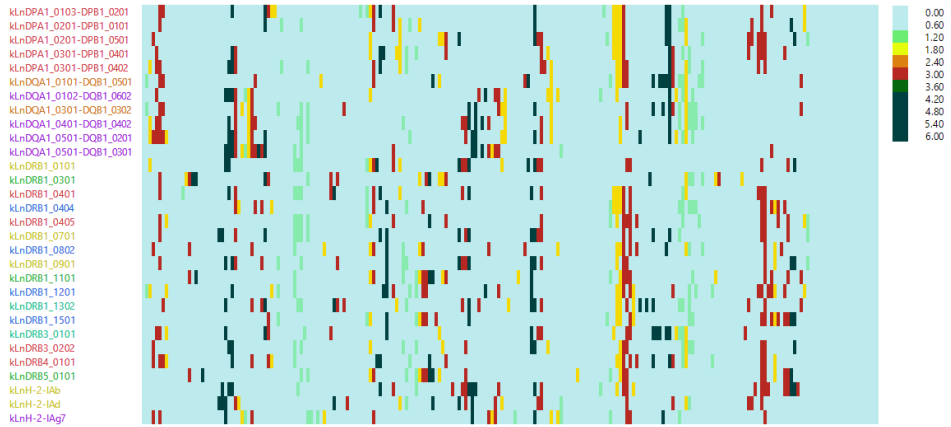


ioGenetics

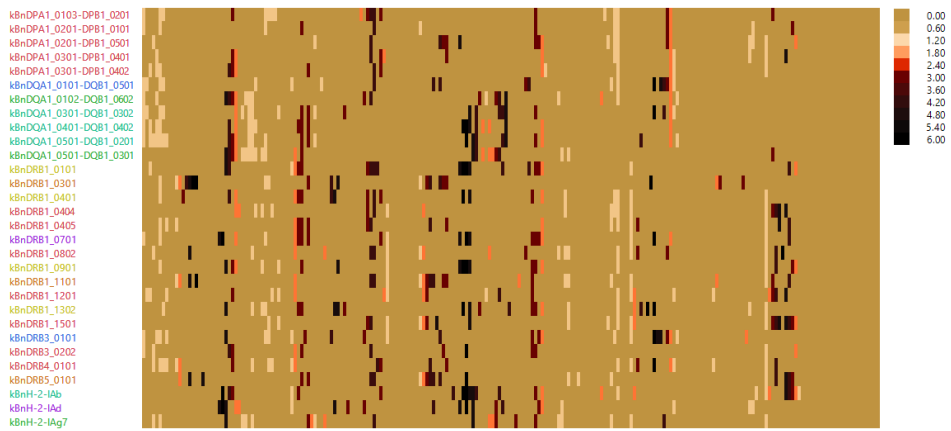
JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t Kernels MHC II @ 15-20 aa by hCAT_L



JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t Kernels MHC II @ 15-20 aa by hCAT_B

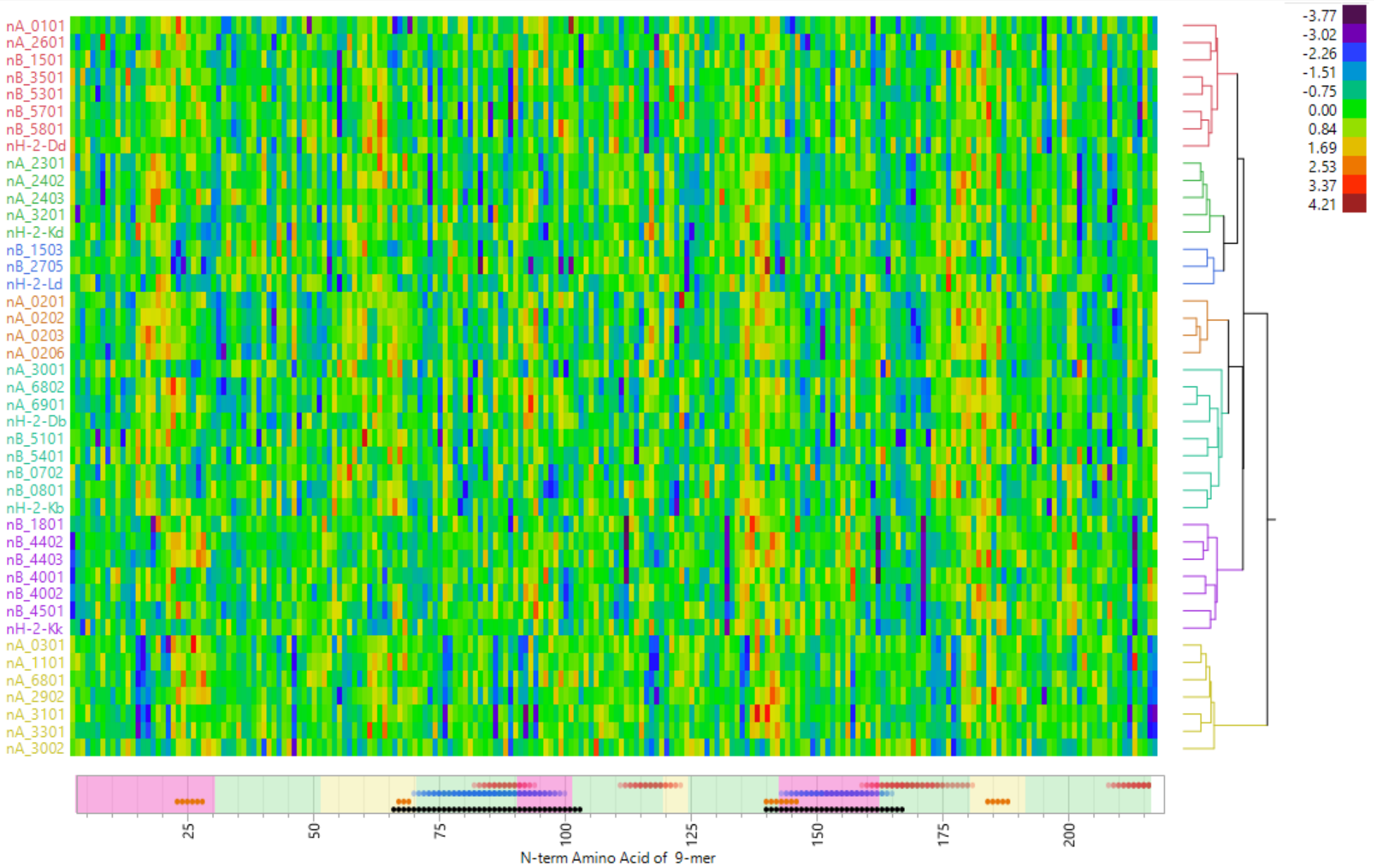


ioGenetics

Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t_T_MHC_I

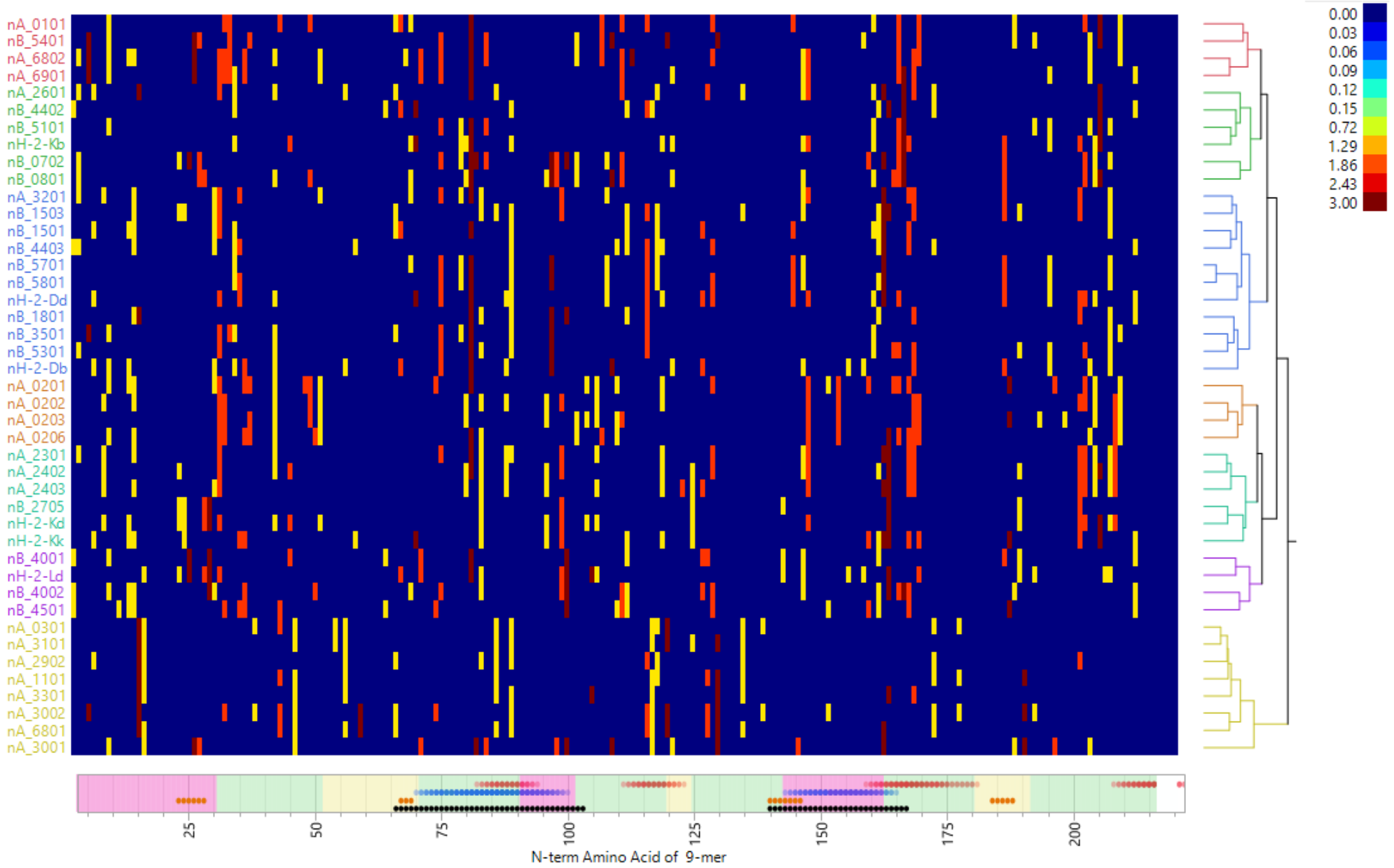


ioGenetics

Hierarchical Cluster MHC I

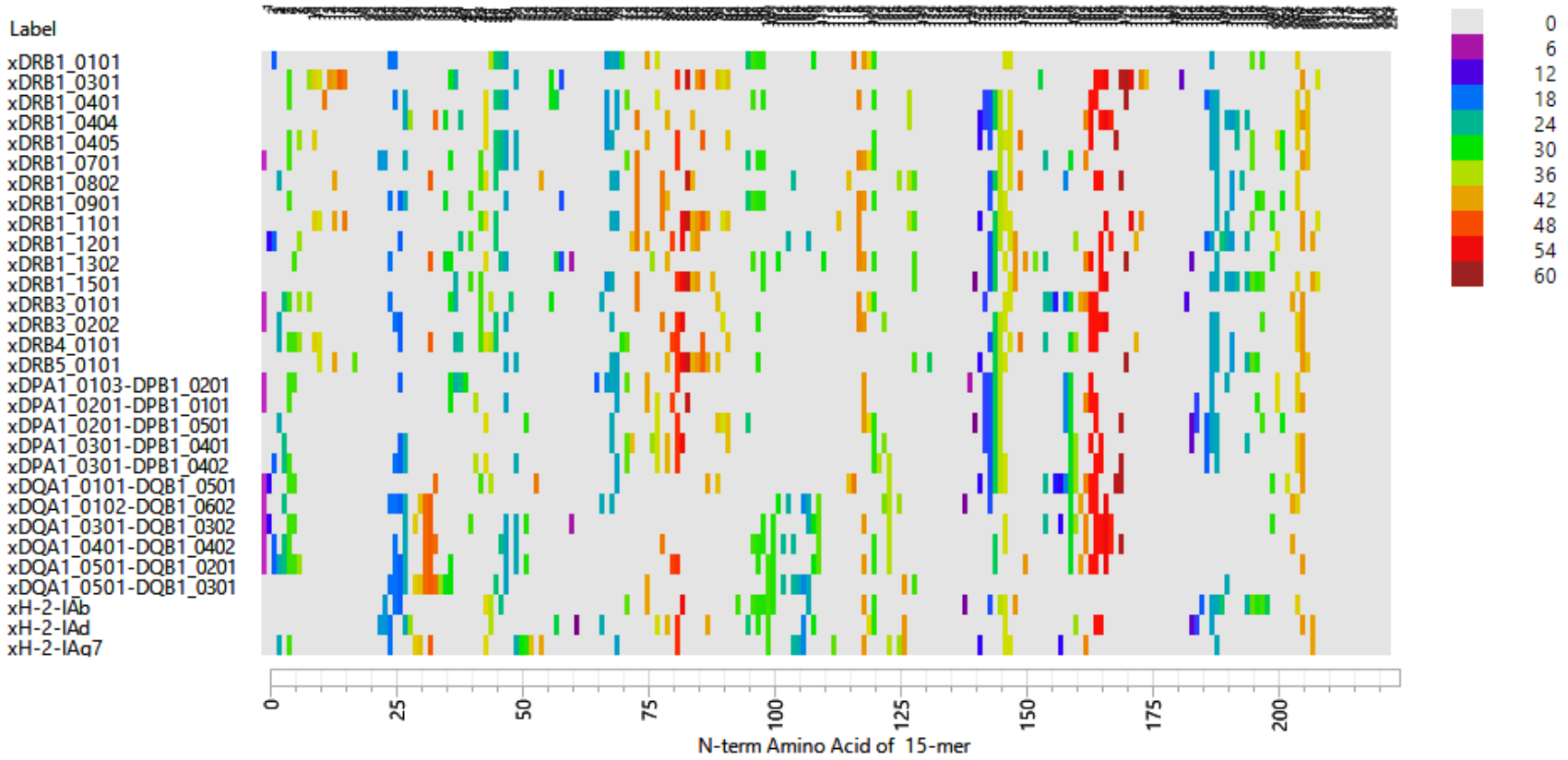
Method = Ward

JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t MHC I & Cathepsins consolidated



oGenetics

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS2a VIPR_ALG4_402122761_3509_4180_1;t



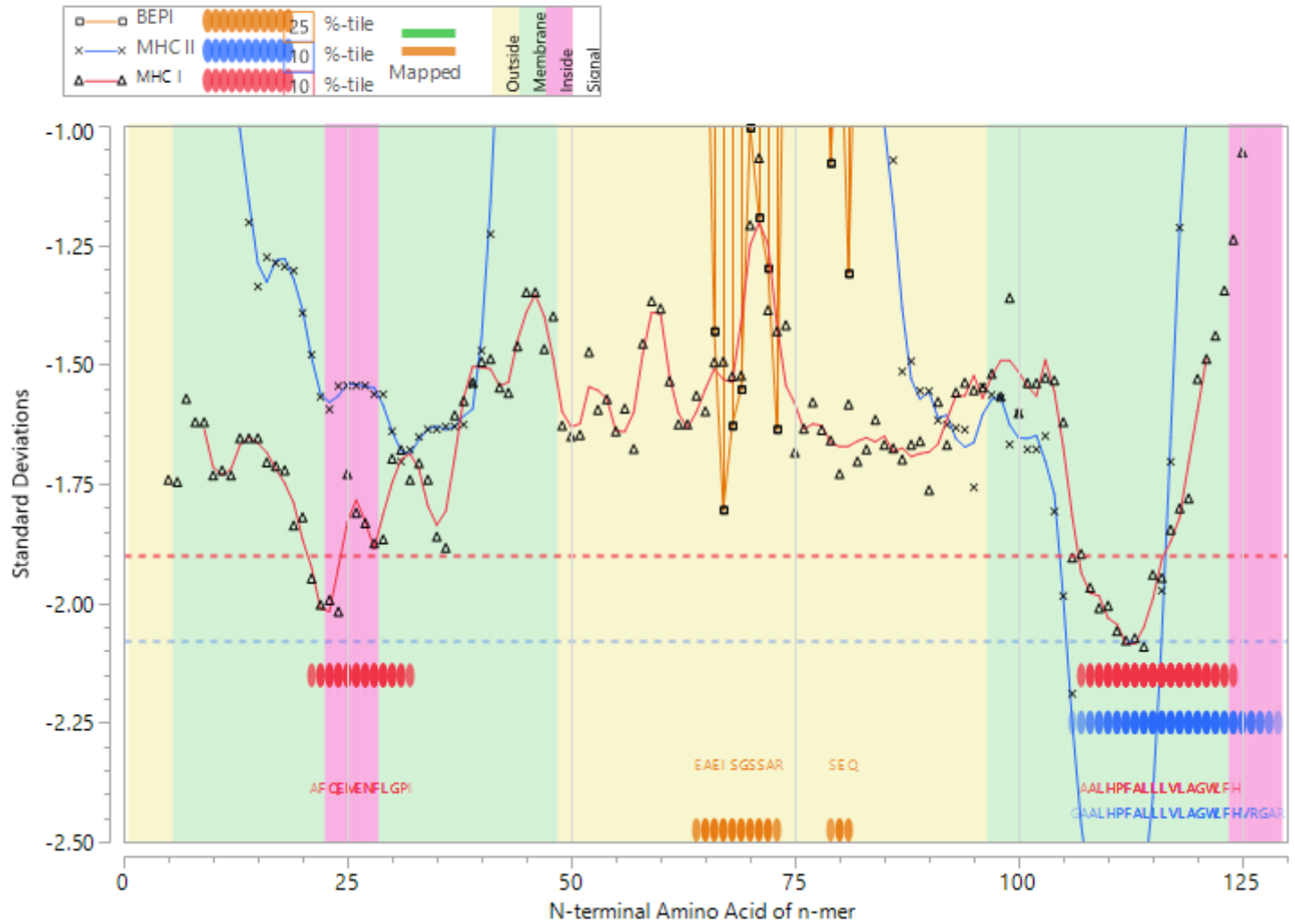
ioGenetics

JN811140_non_structural_protein_NS2b

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS2b
VIPR_ALG4_402122761_4181_4570_1;t

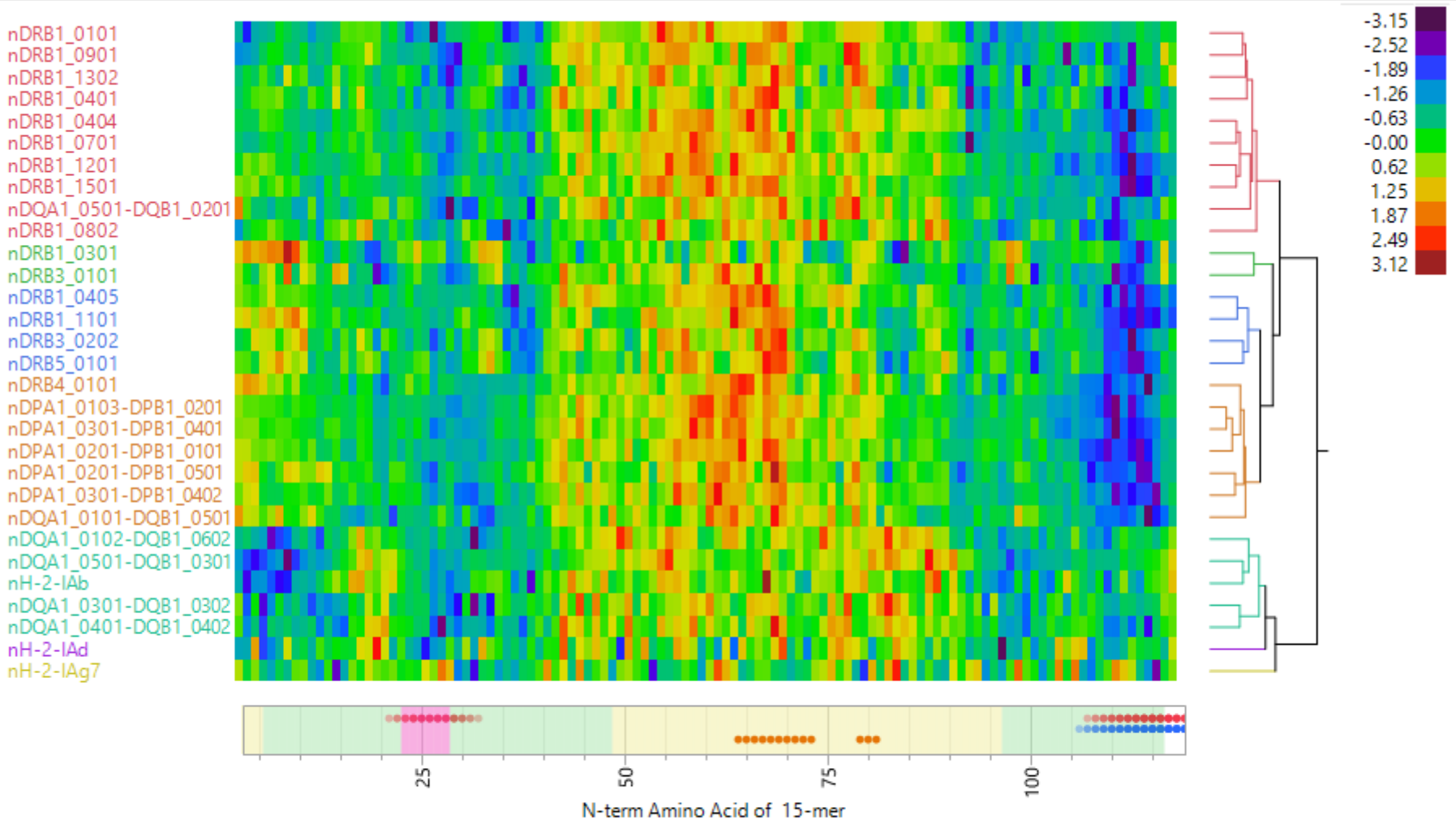


ioGenetics

Hierarchical Cluster MHC II

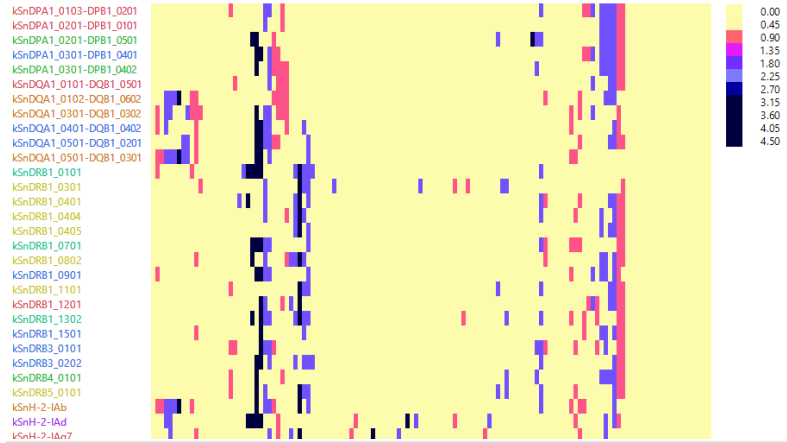
Method = Ward

JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t_T_MHC_II

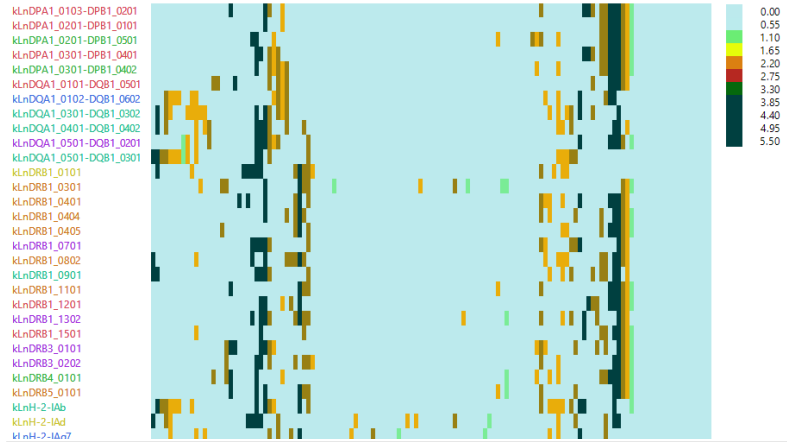


ioGenetics

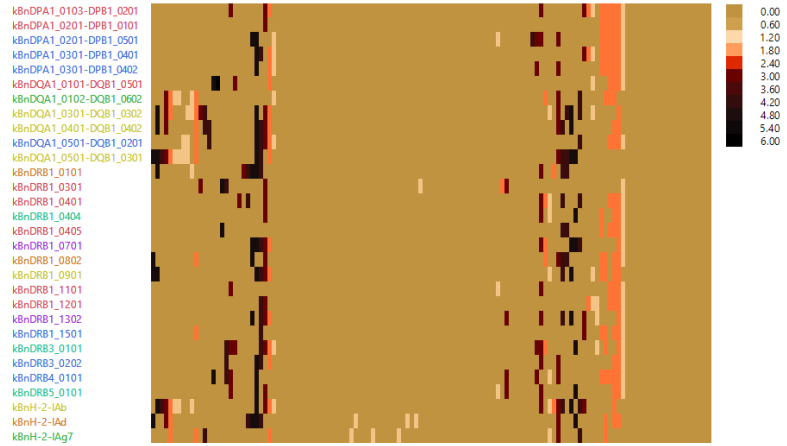
JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t Kernels MHC II @ 15-20 aa by hCAT_L



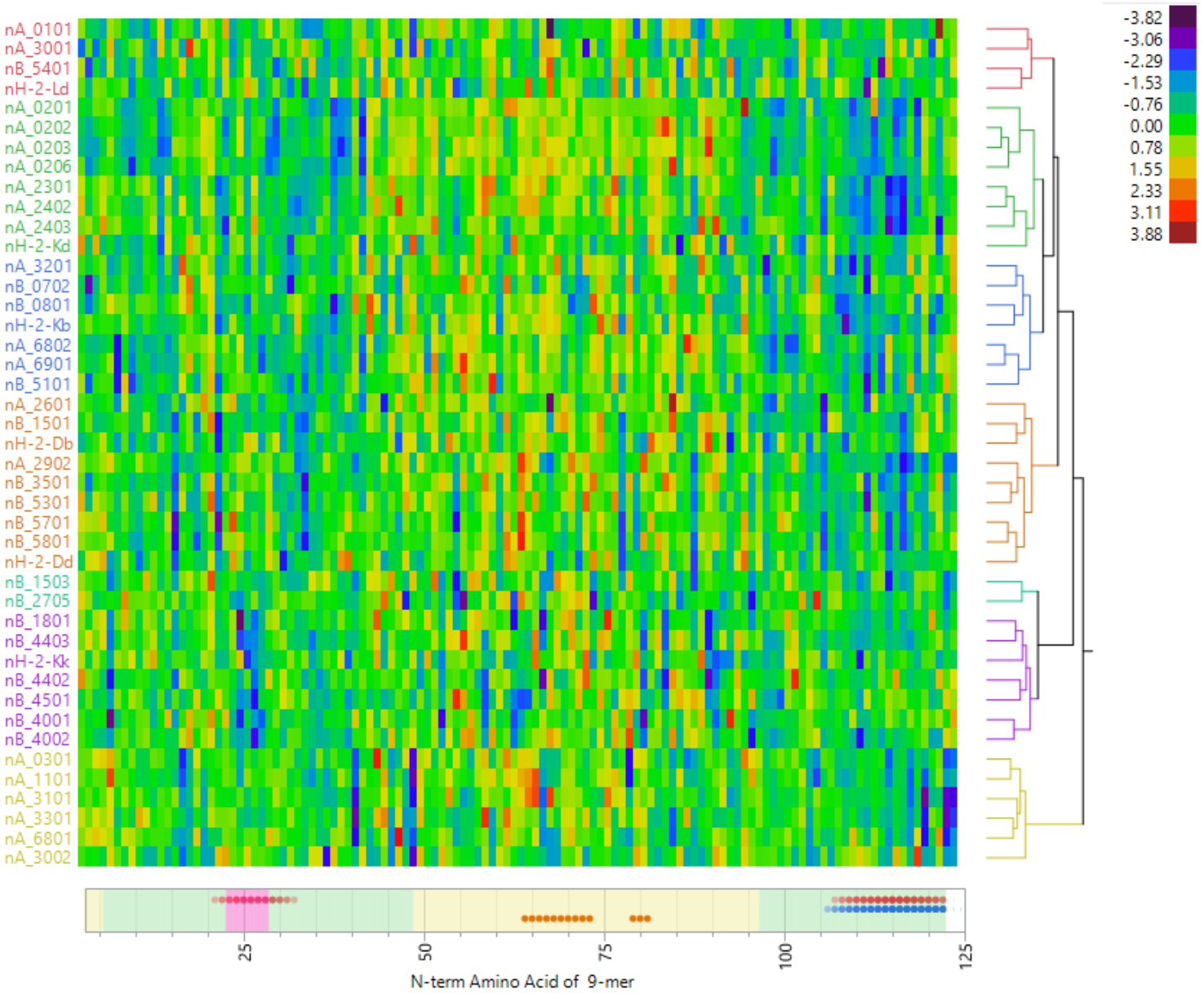
JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t_T_MHC_I

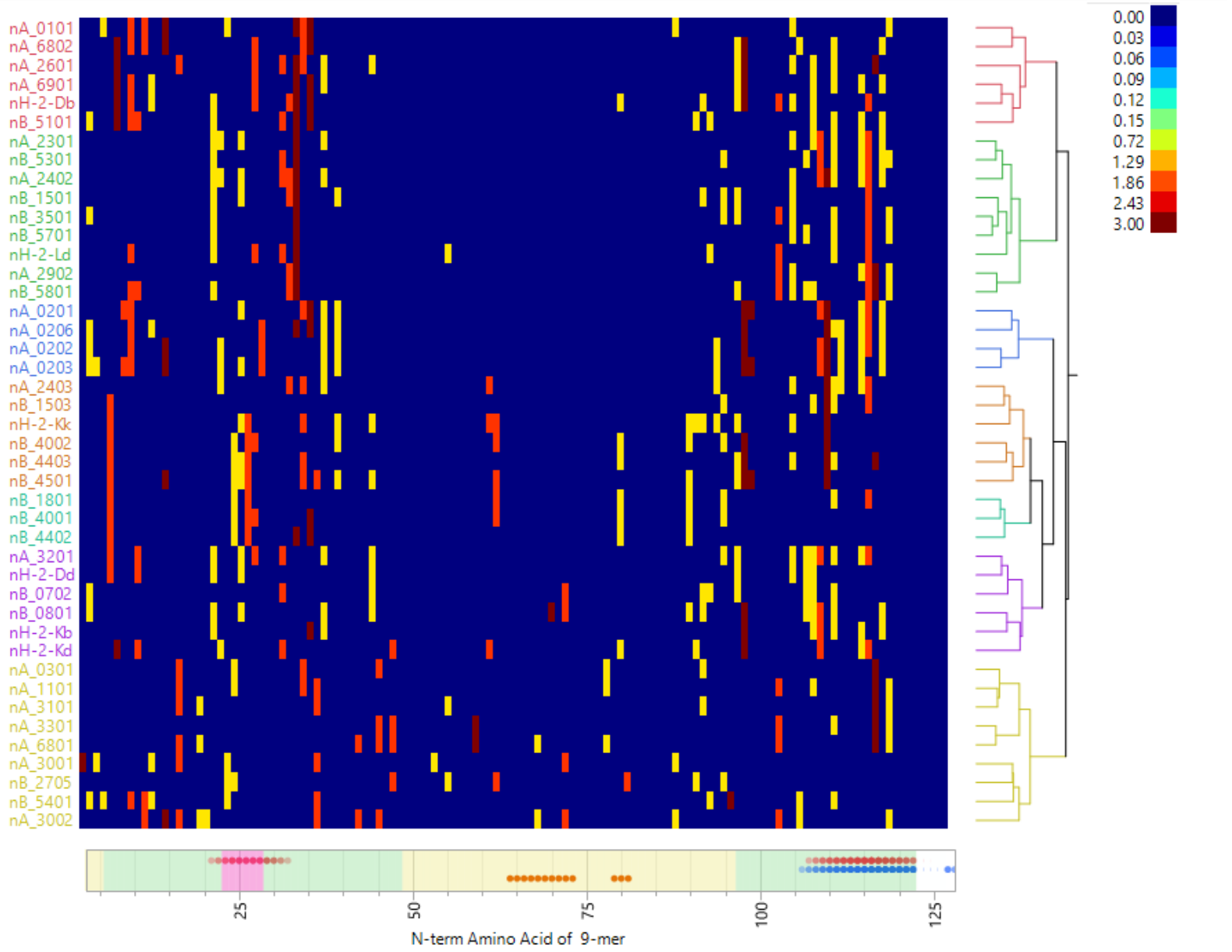


ioGenetics

Hierarchical Cluster MHC I

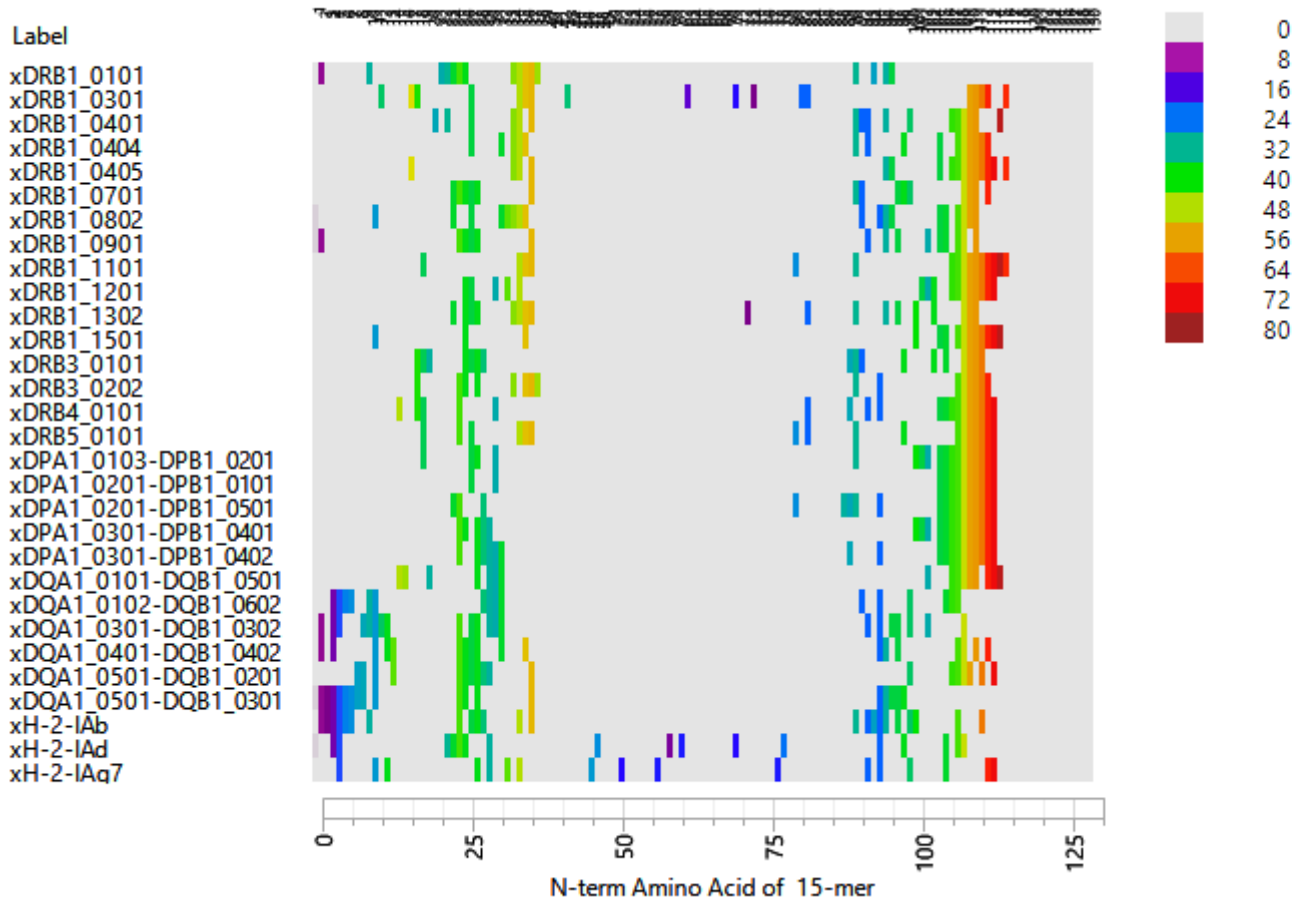
Method = Ward

JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t MHC I & Cathepsins consolidated



ioGenetics

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS2b VIPR_ALG4_402122761_4181_4570_1;t

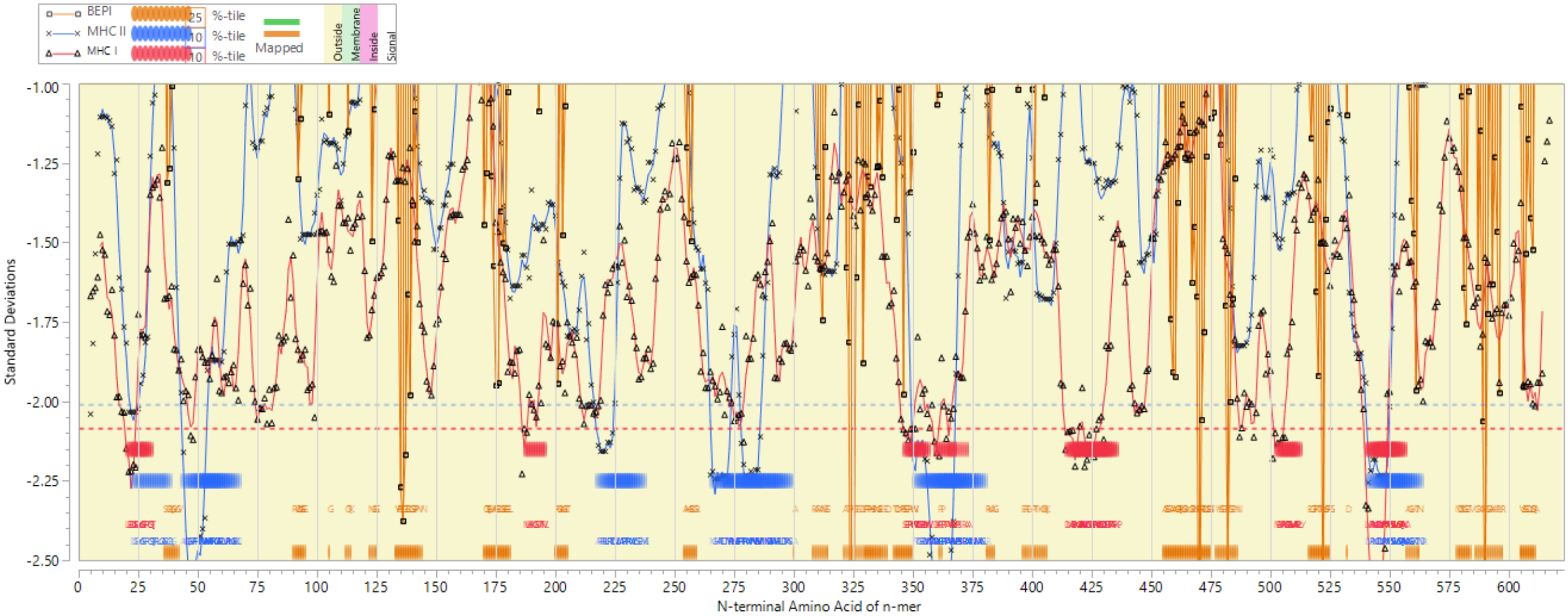


JN811140_non_structural_protein_NS3

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS3
VIPR_ALG4_402122761_4571_6439_1;t

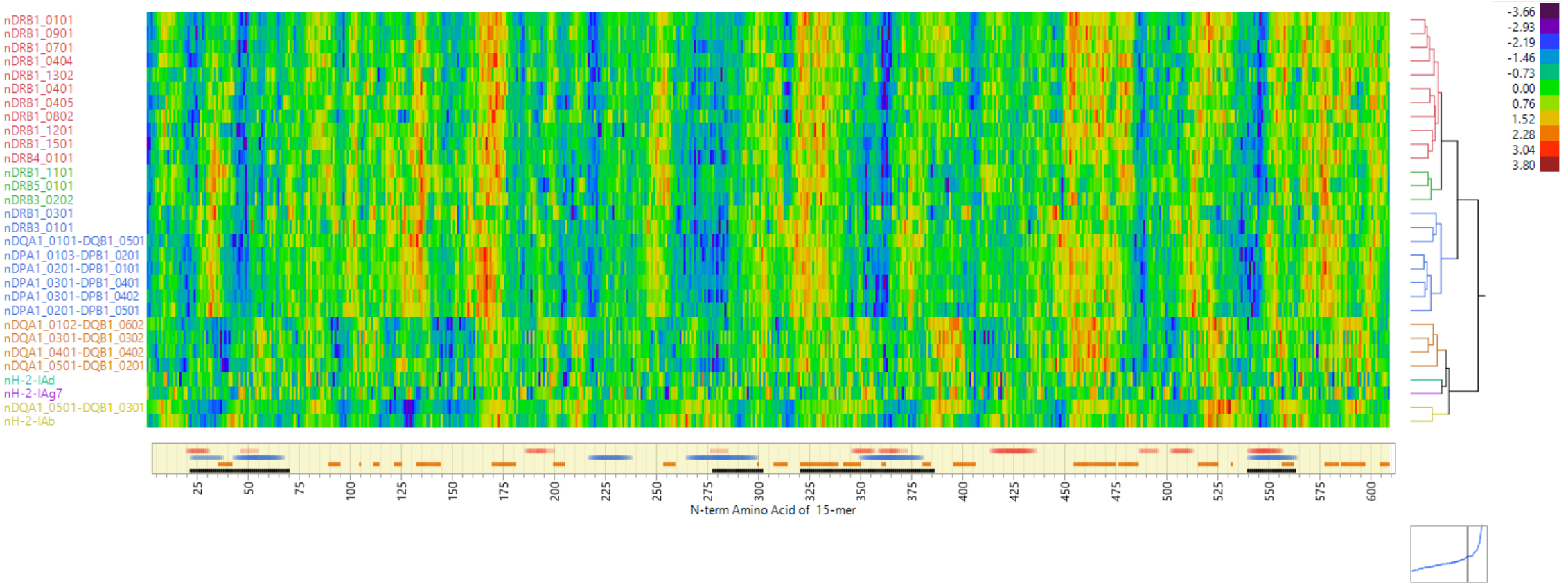


ioGenetics

Hierarchical Cluster MHC II

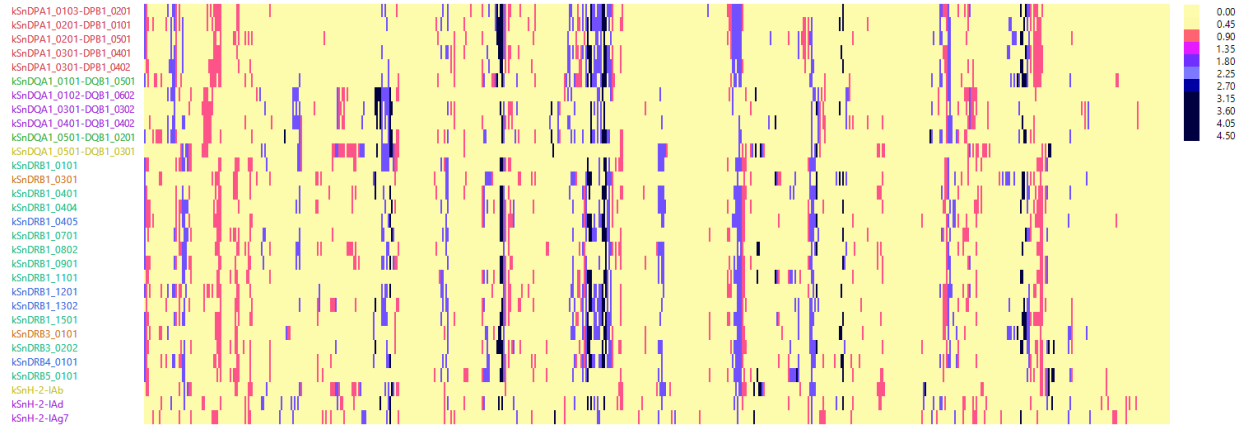
Method = Ward

JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t_T_MHC_II 2

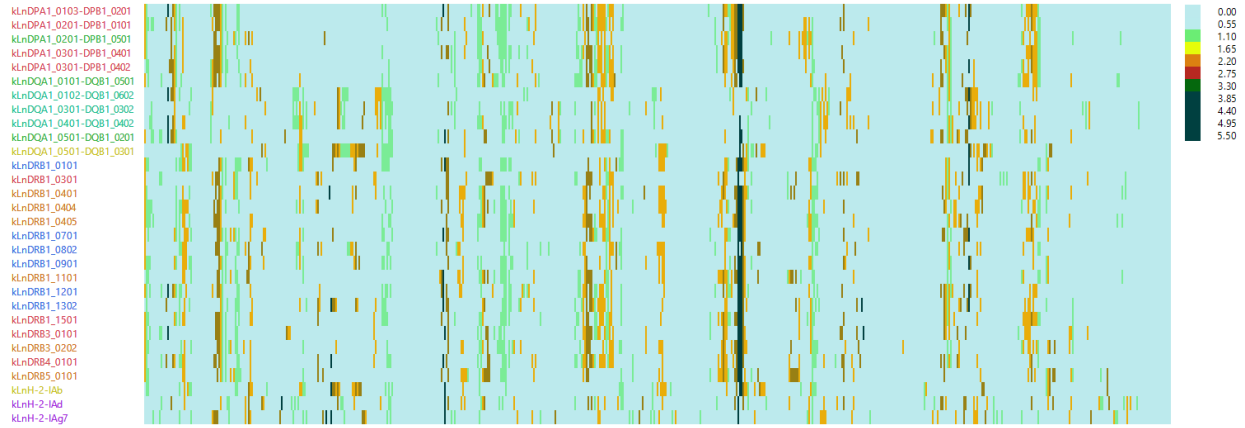


ioGenetics

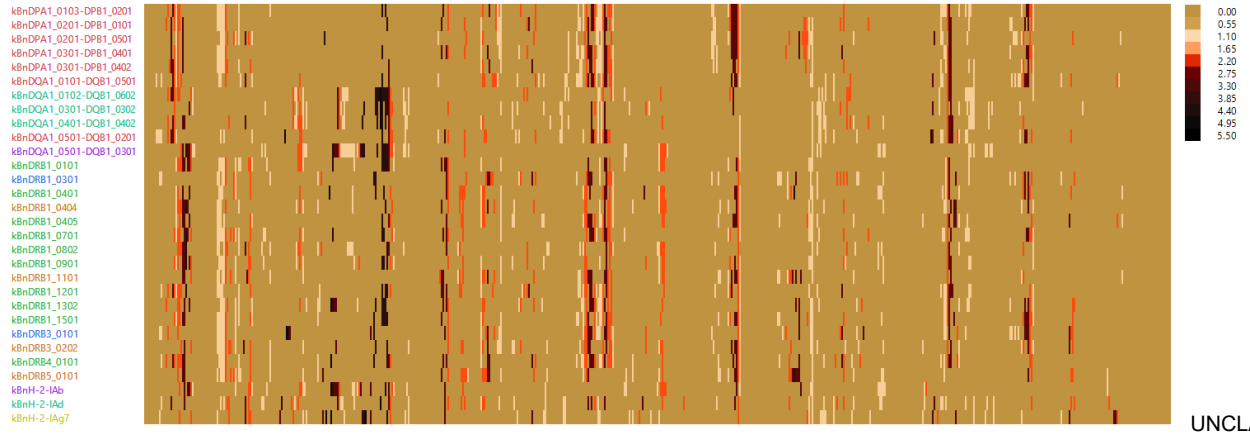
JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t Kernels MHC II @ 15-20 aa by hCAT_L



JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t Kernels MHC II @ 15-20 aa by hCAT_B

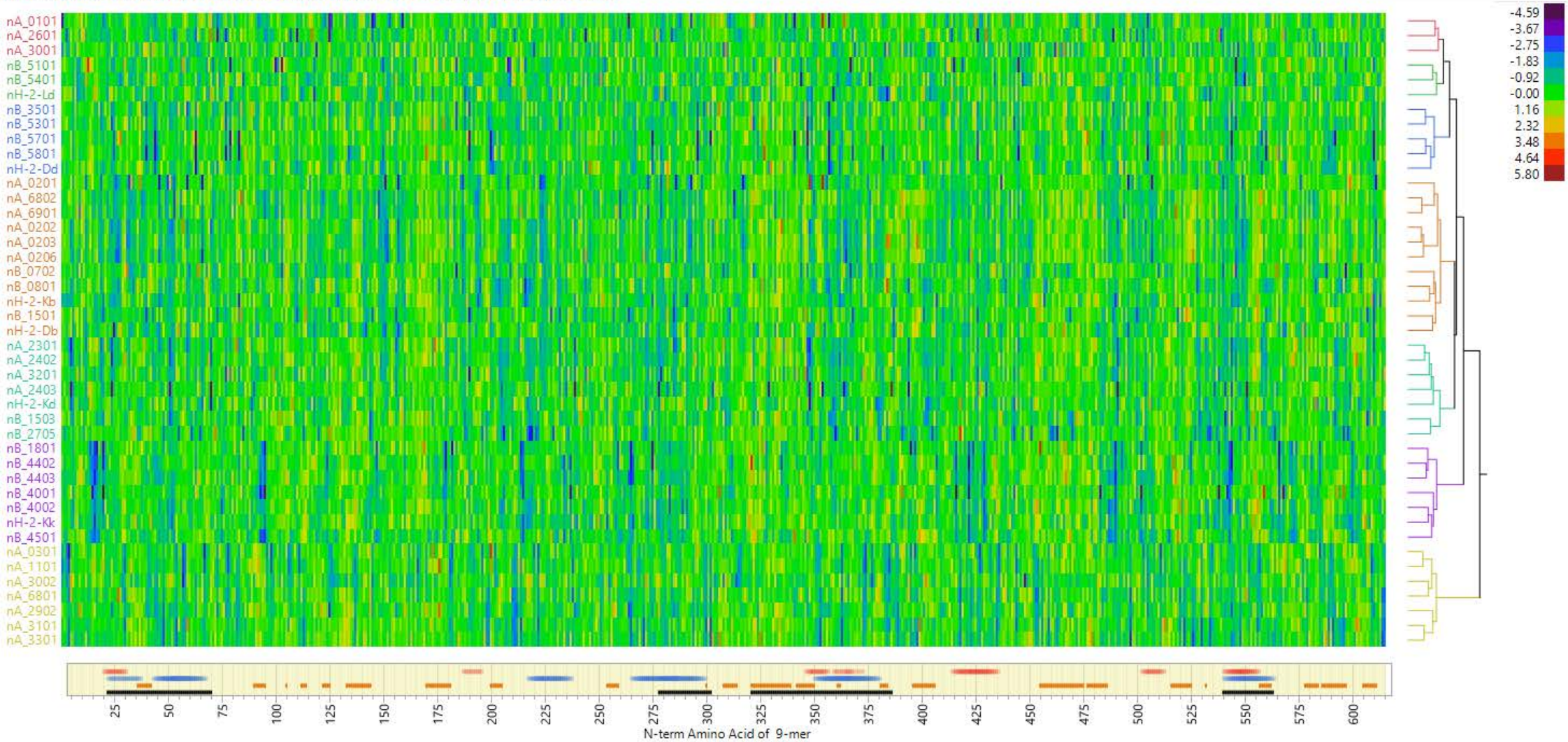


ioGenetics

Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t_T_MHC I

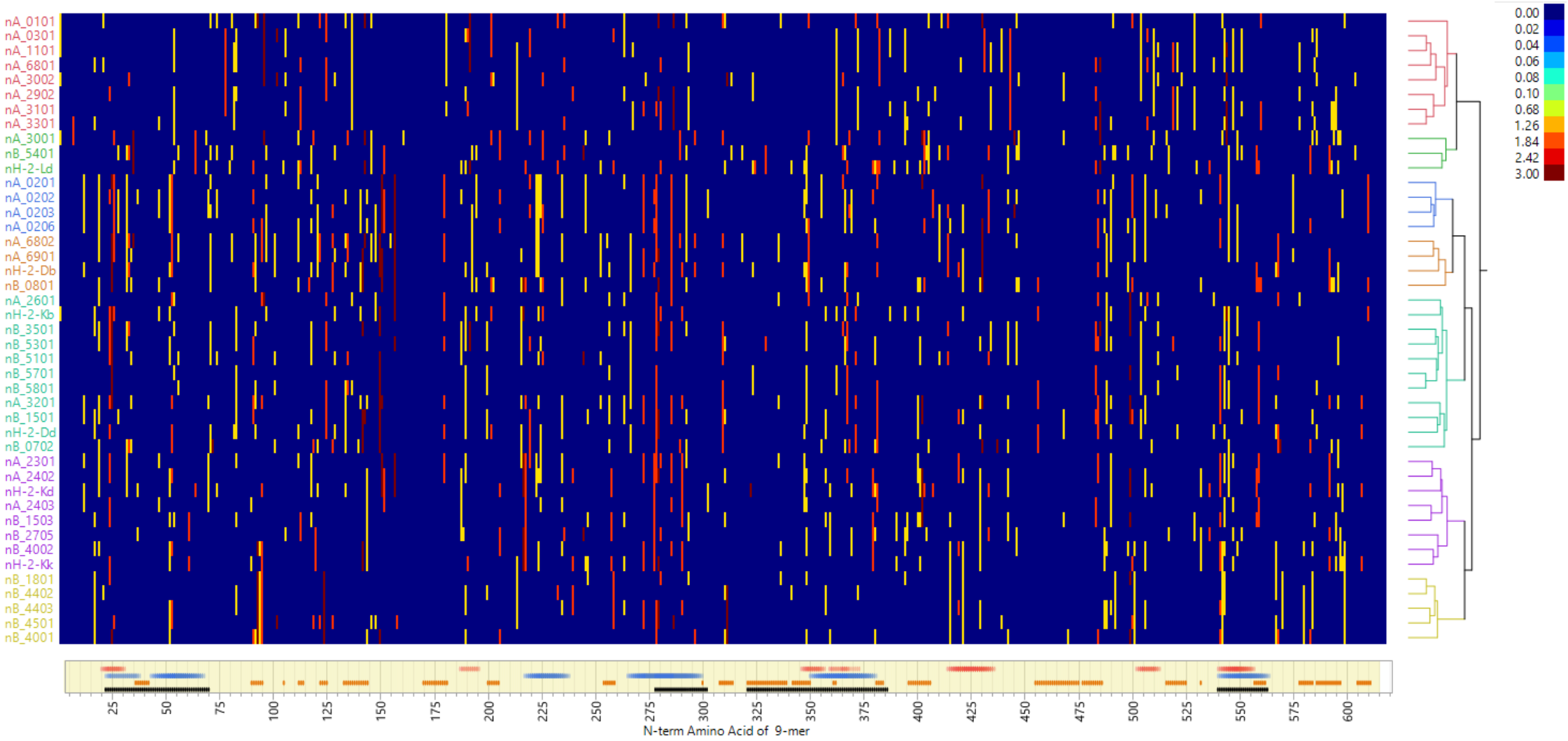


ioGenetics

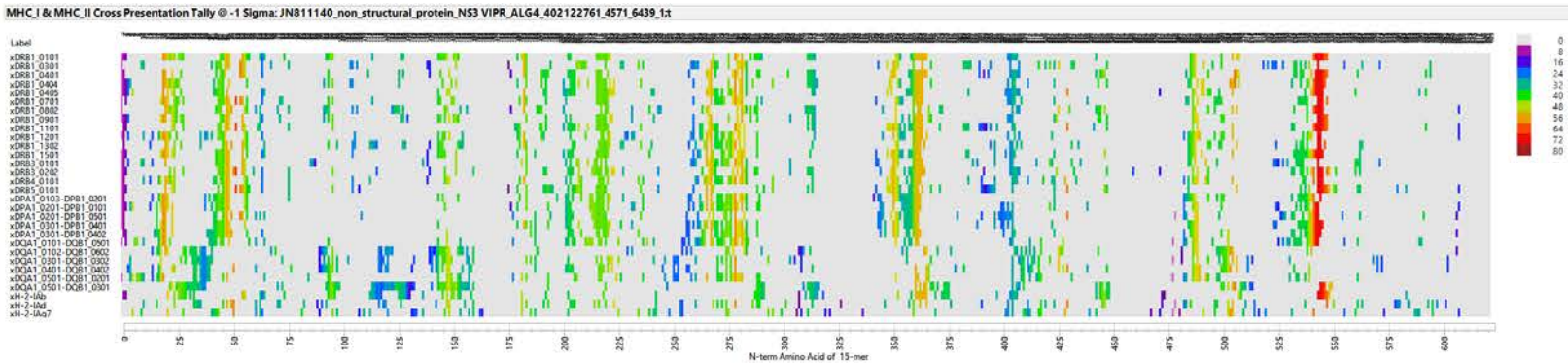
Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1;t MHC I & Cathepsins consolidated 2



ioGenetics

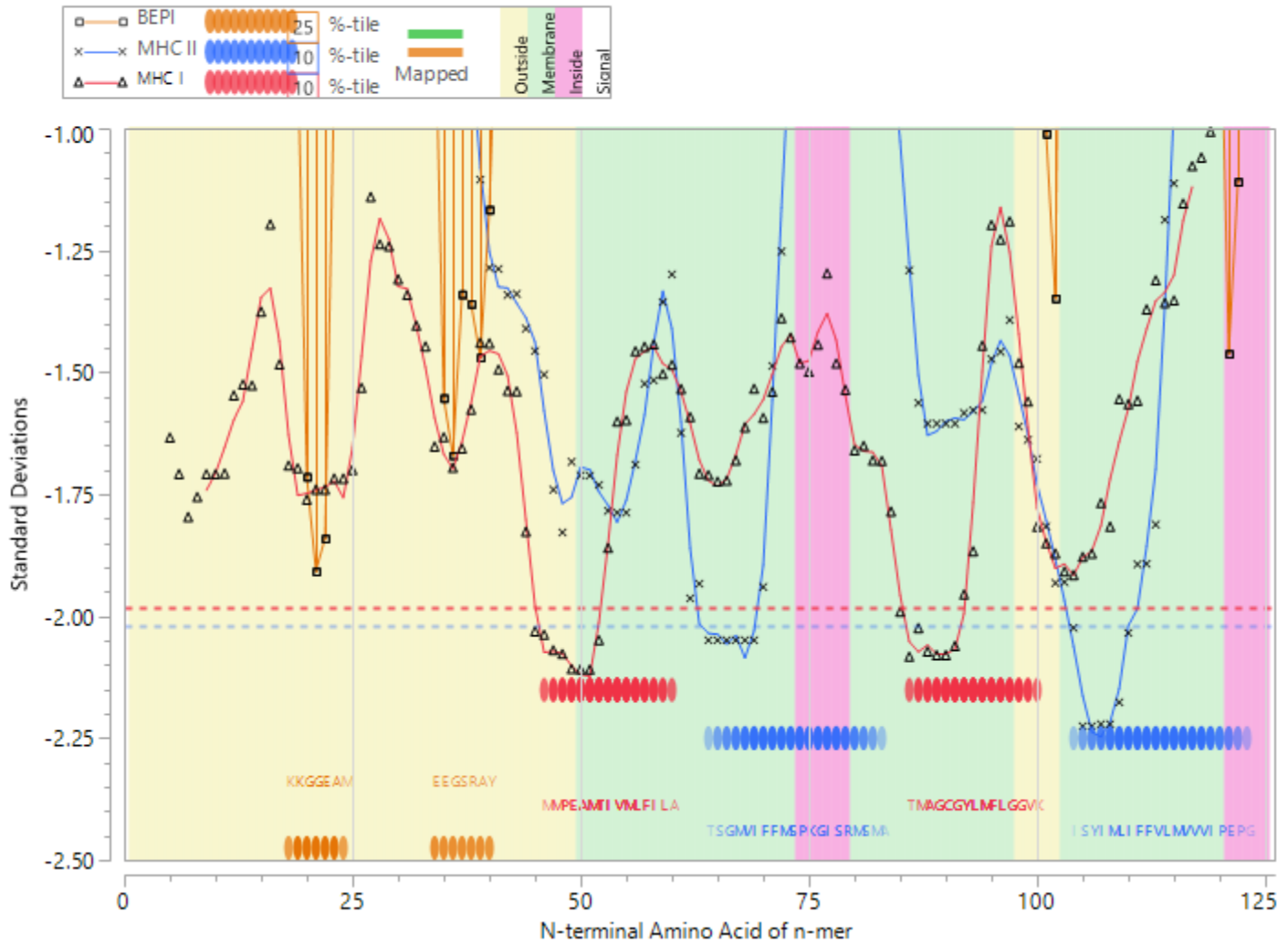

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS3 VIPR_ALG4_402122761_4571_6439_1:t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

JN811140_non_structural_protein_NS4a

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS4a
VIPR_ALG4_402122761_6440_6817_1;t

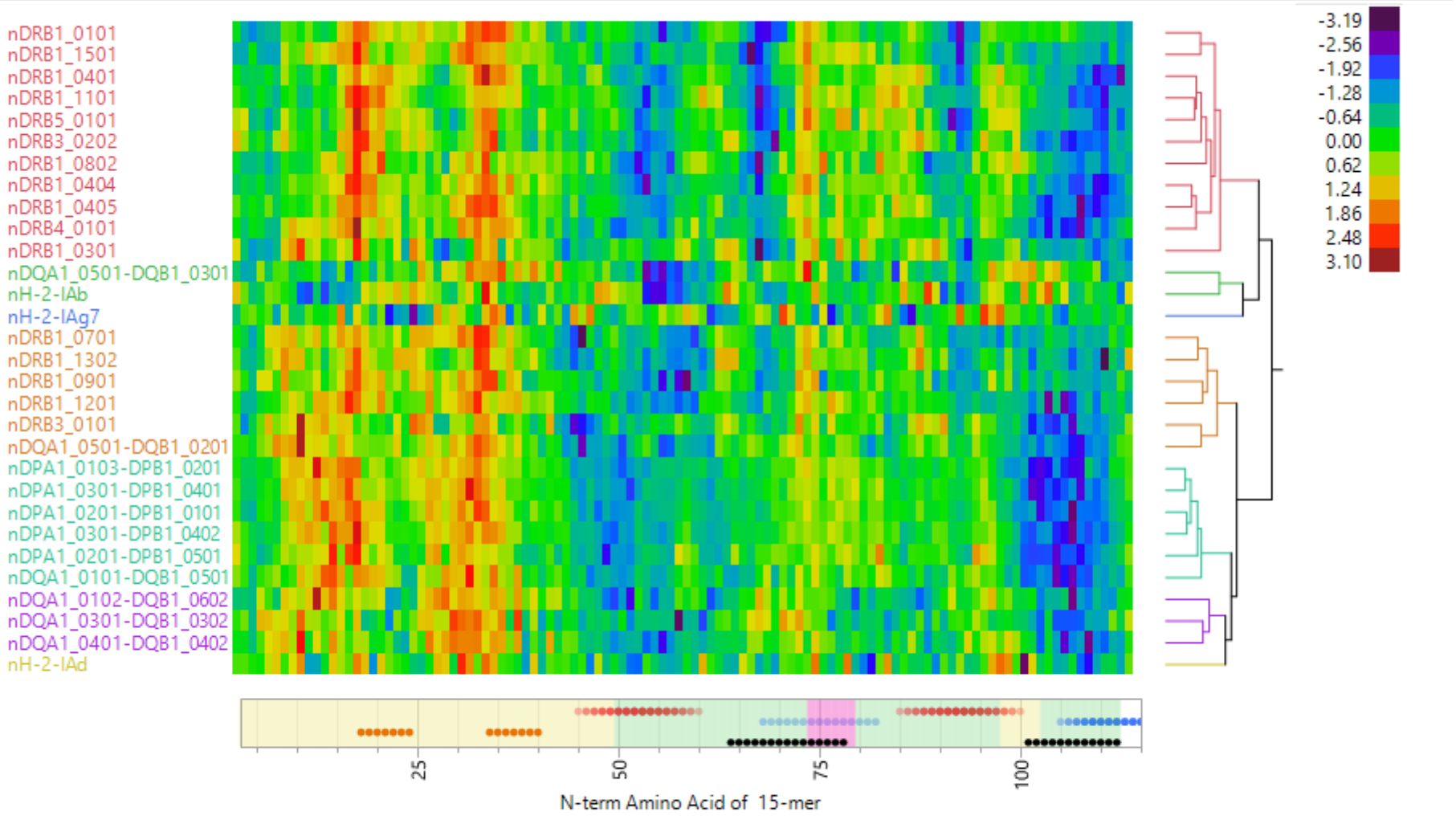


ioGenetics

Hierarchical Cluster MHC II

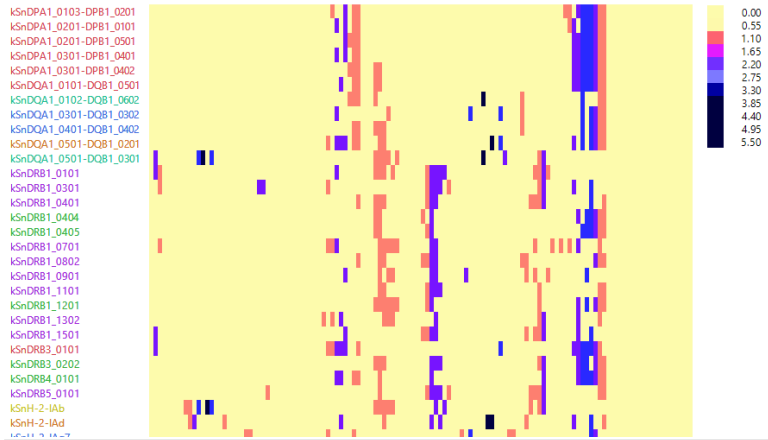
Method = Ward

JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t_T_MHC_II

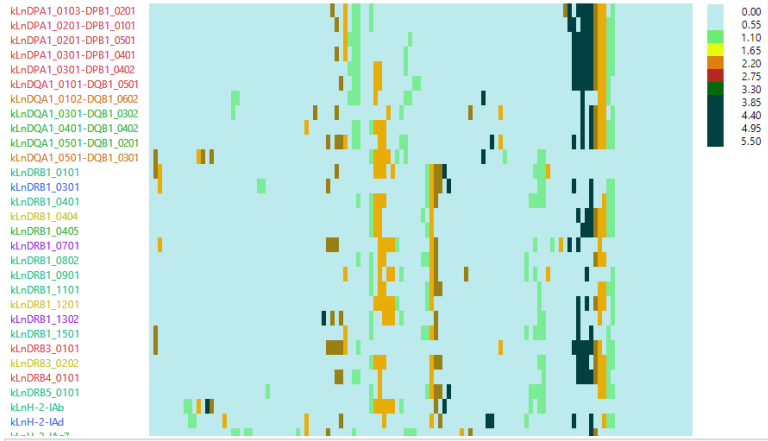


ioGenetics

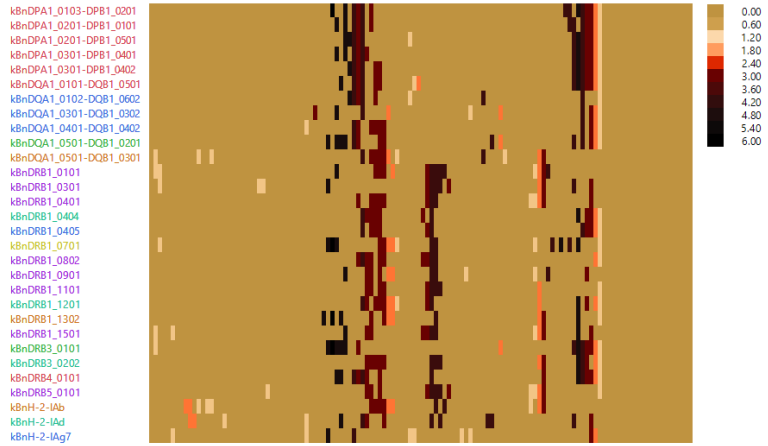
JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t Kernels MHC II @ 15-20 aa by hCAT_L



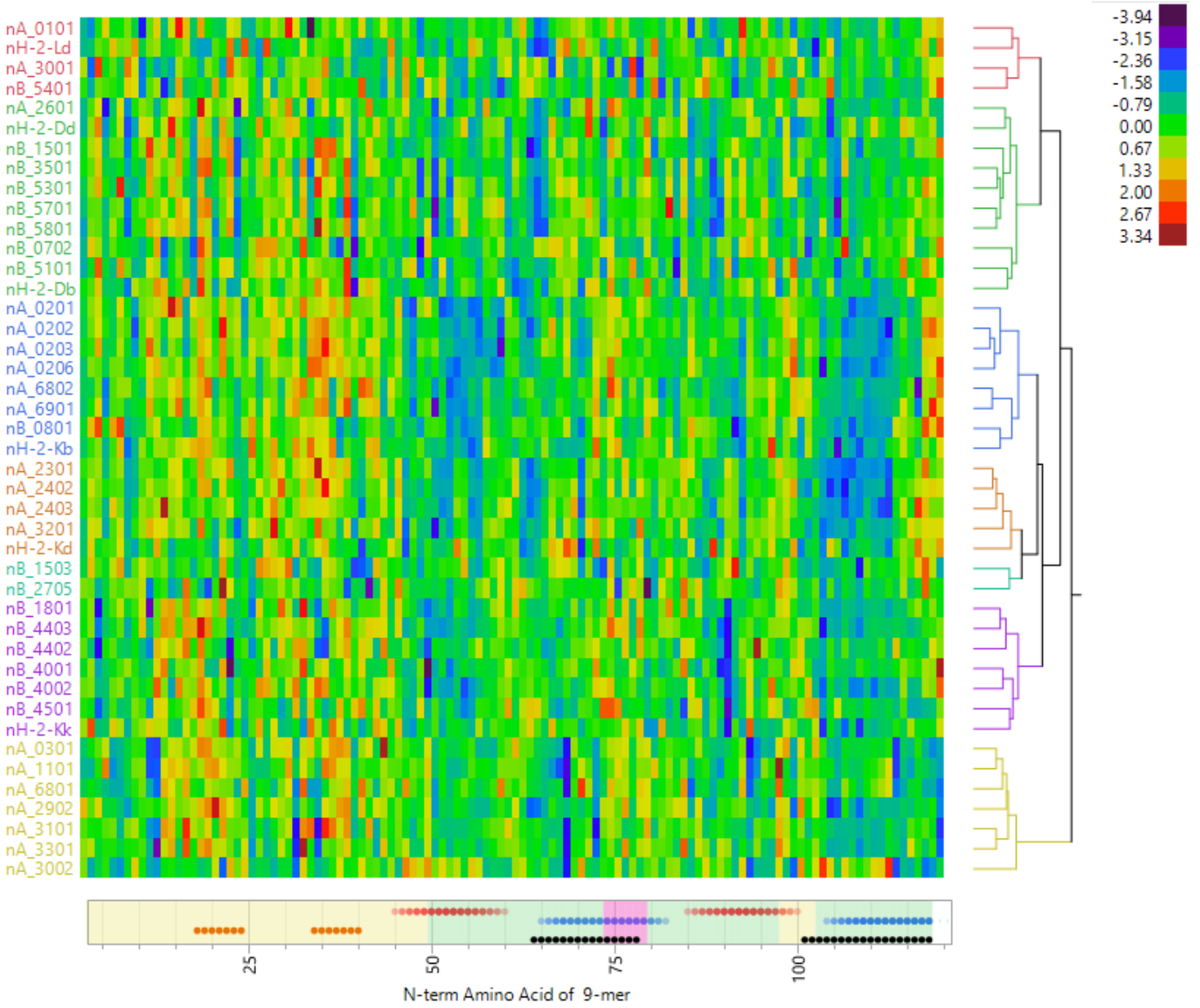
JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t_T_MHC_I

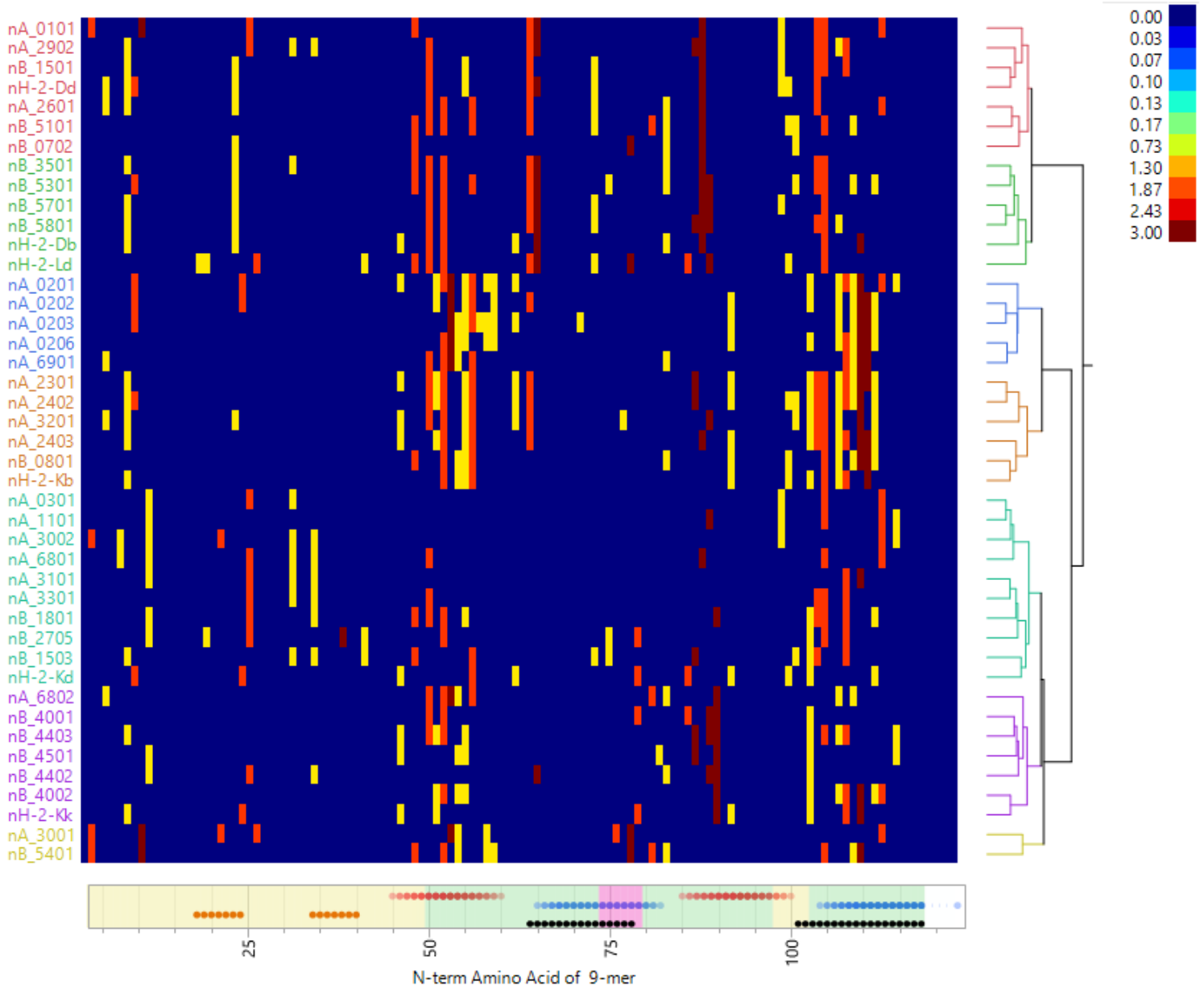


ioGenetics

Hierarchical Cluster MHC I

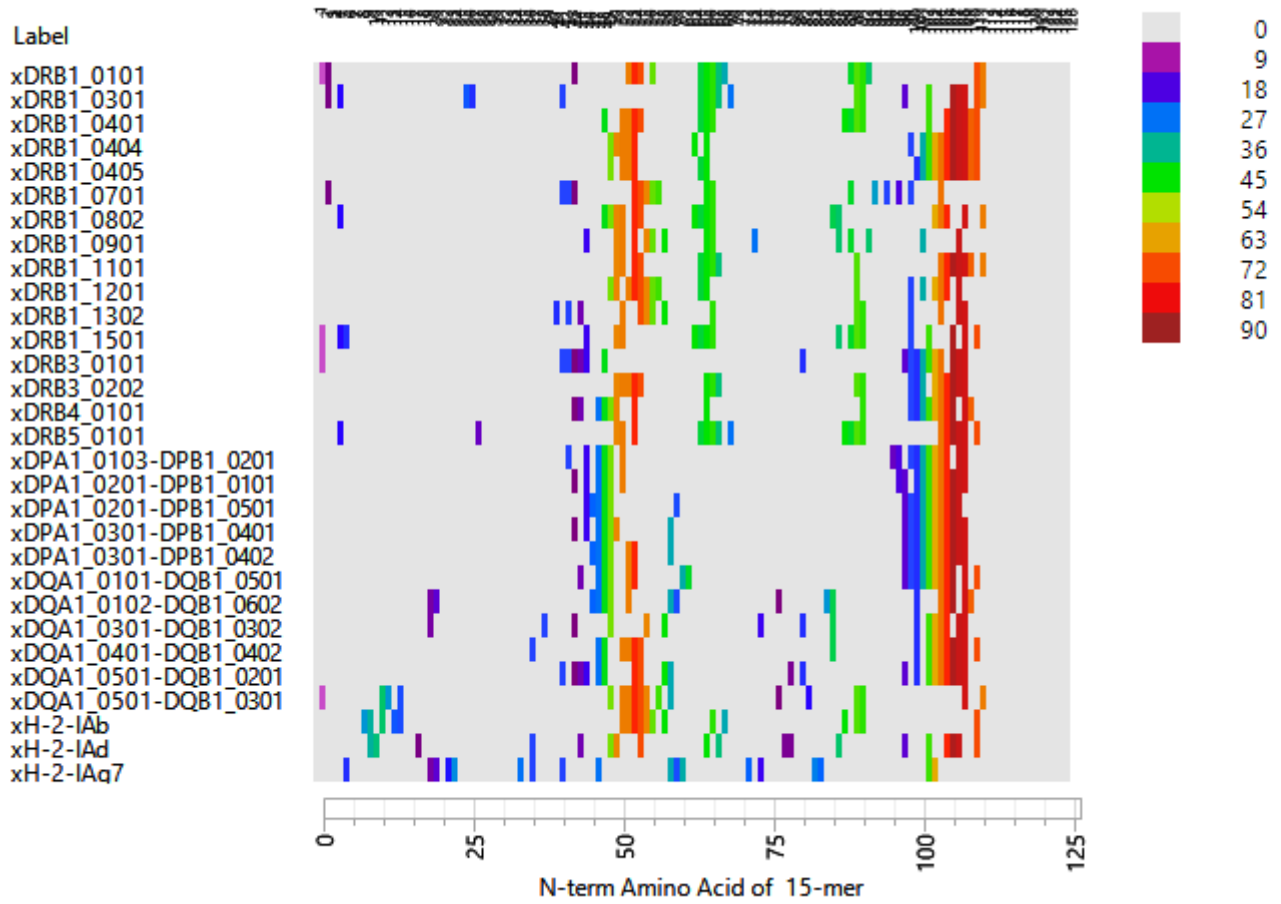
Method = Ward

JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t MHC I & Cathepsins consolidated



ioGenetics

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS4a VIPR_ALG4_402122761_6440_6817_1;t



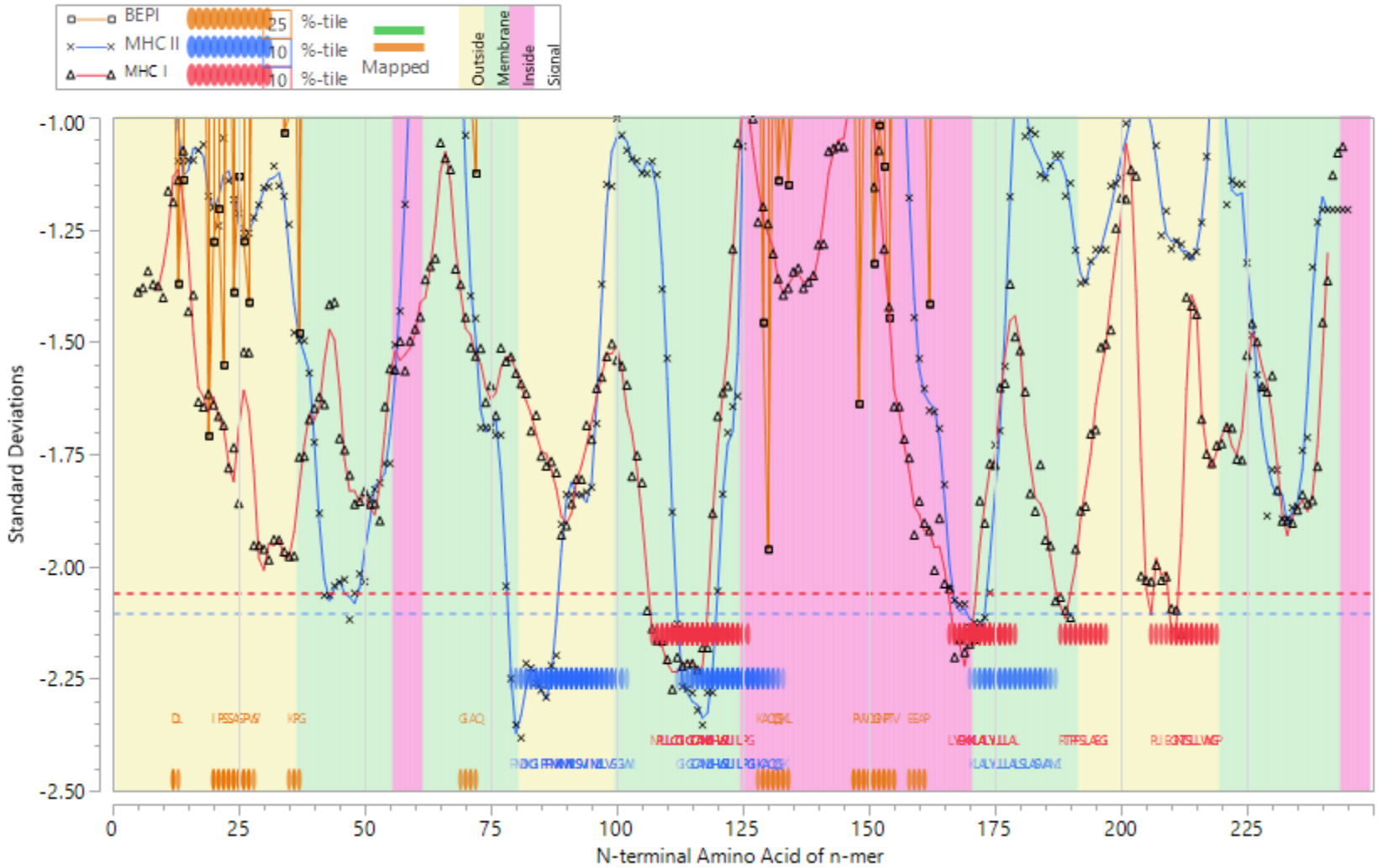
ioGenetics

JN811140_non_structural_protein_NS4b

ioGenetics

Overlay Plot

JN811140_non_structural_protein_NS4b
VIPR_ALG4_402122761_6887_7636_1;t

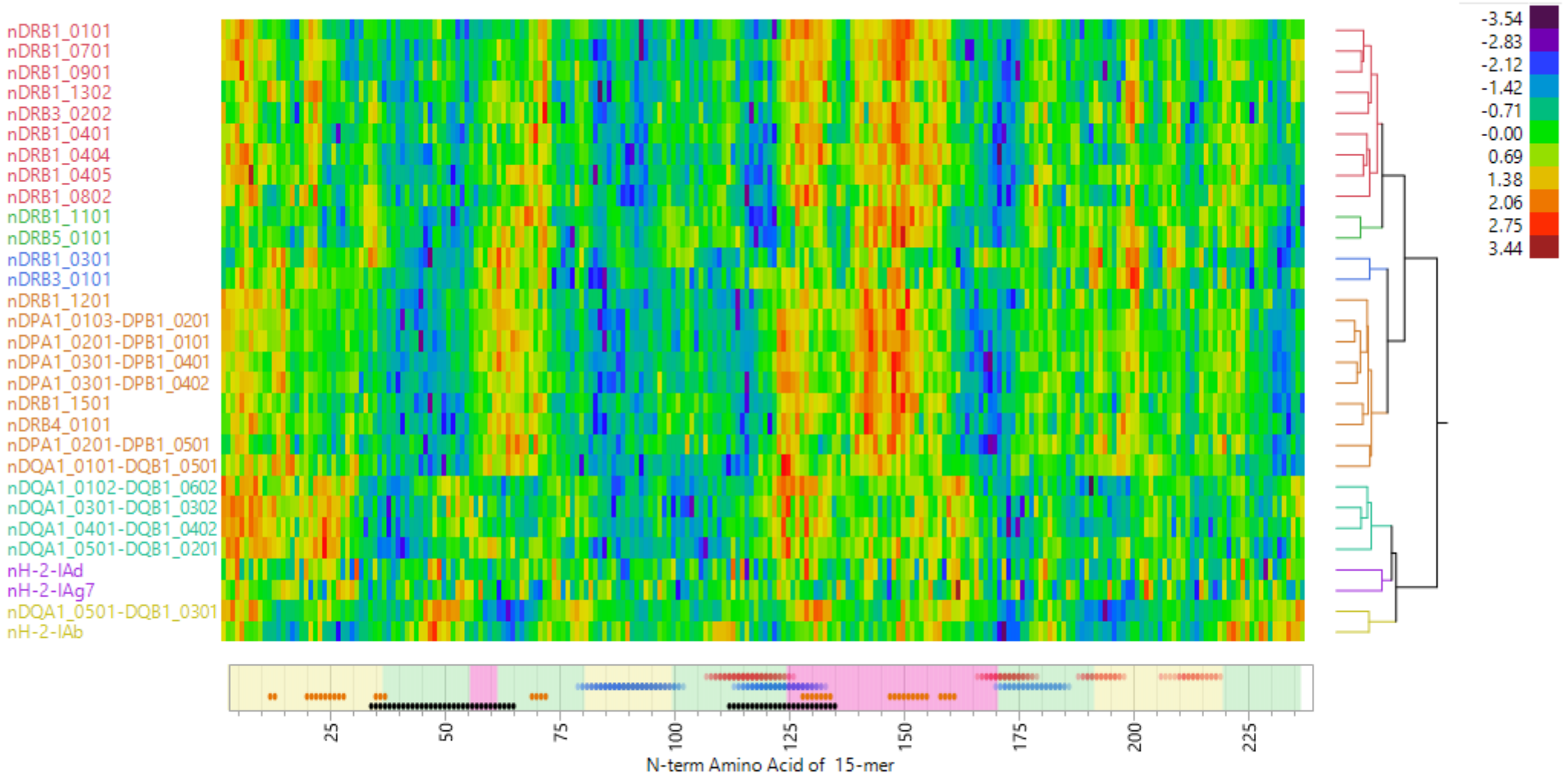


ioGenetics

Hierarchical Cluster MHC II

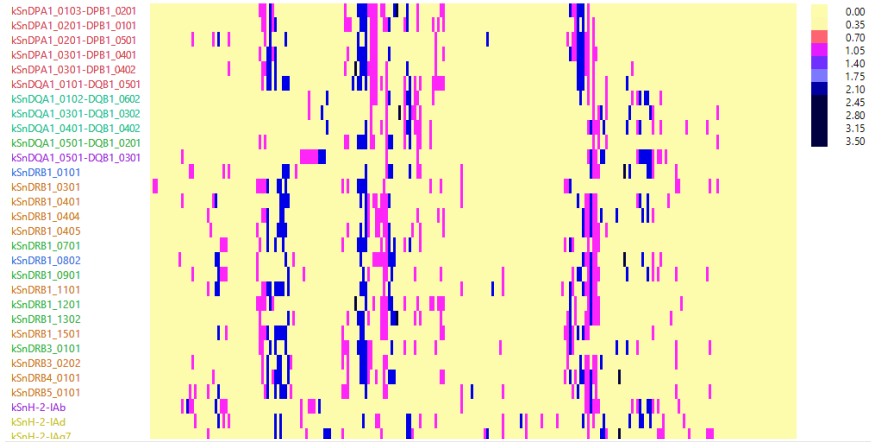
Method = Ward

JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t_T_MHC_II

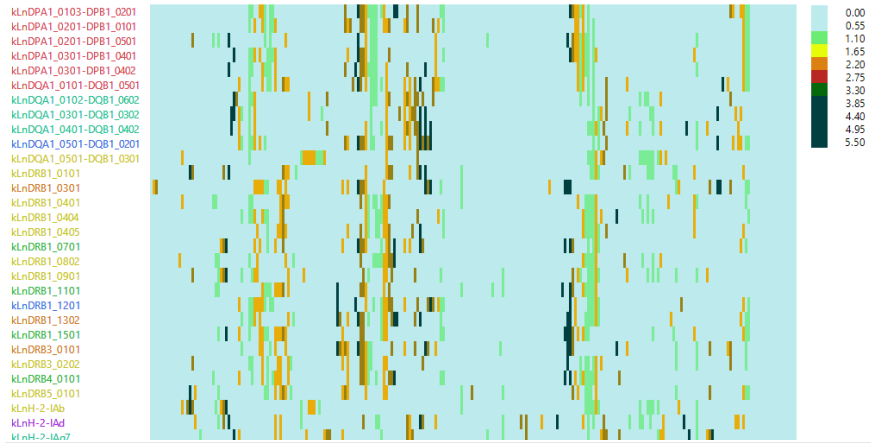


ioGenetics

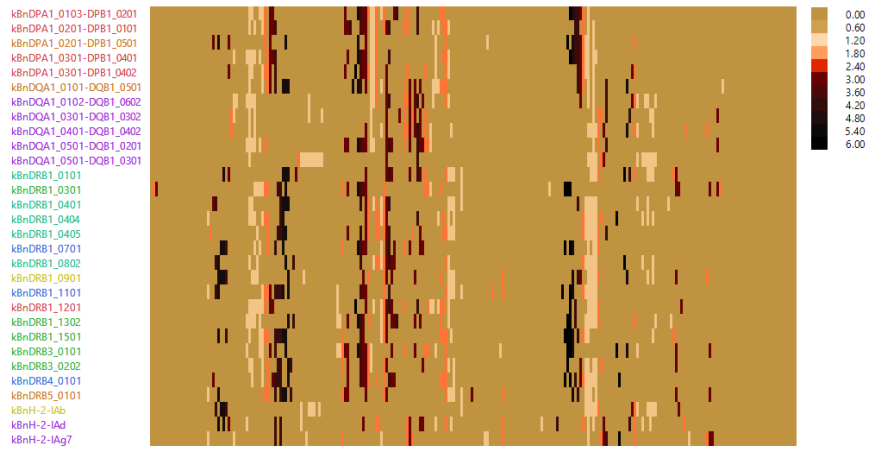
JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t Kernels MHC II @ 15-20 aa by hCAT_L



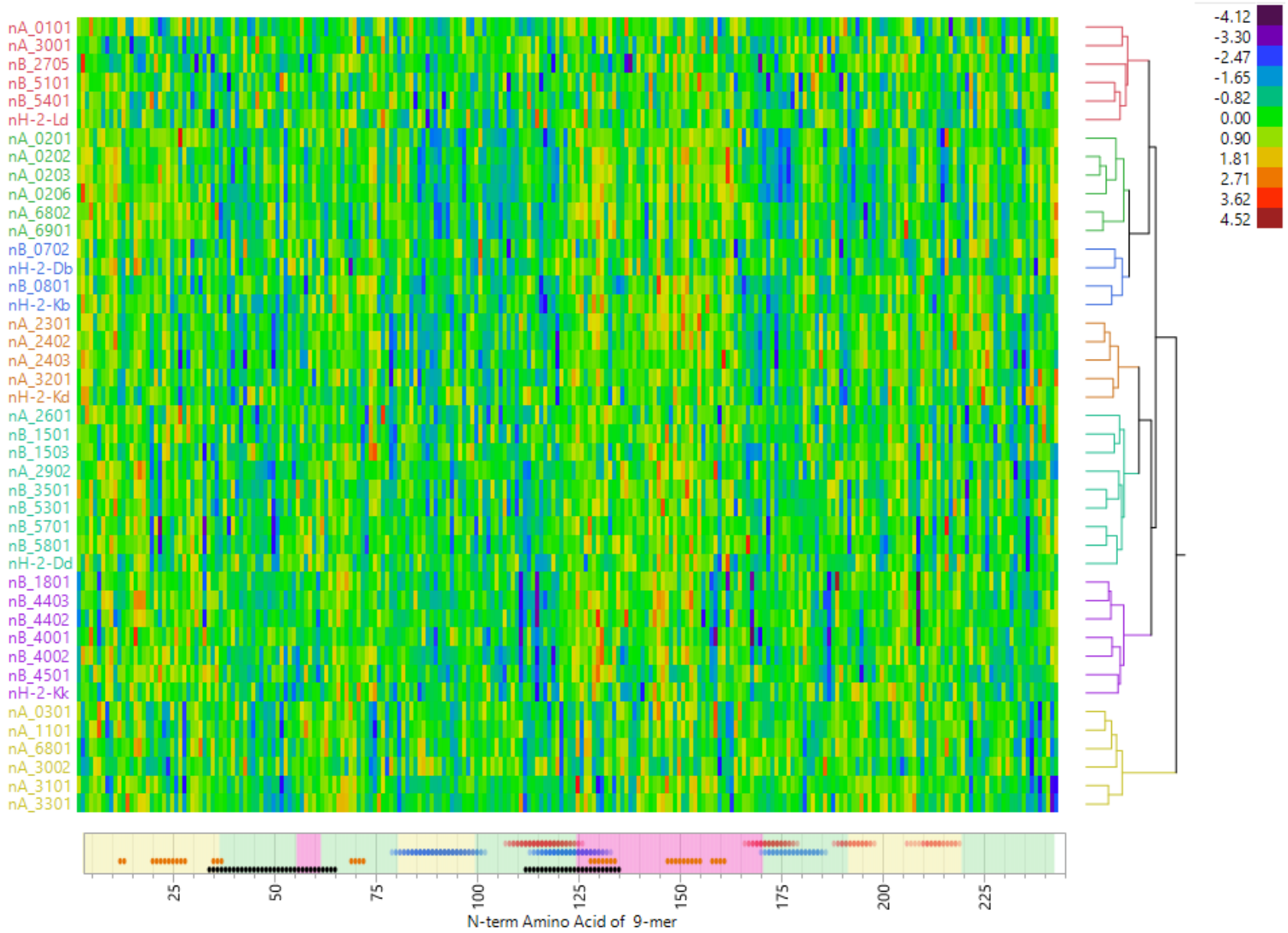
JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t_T_MHC_I

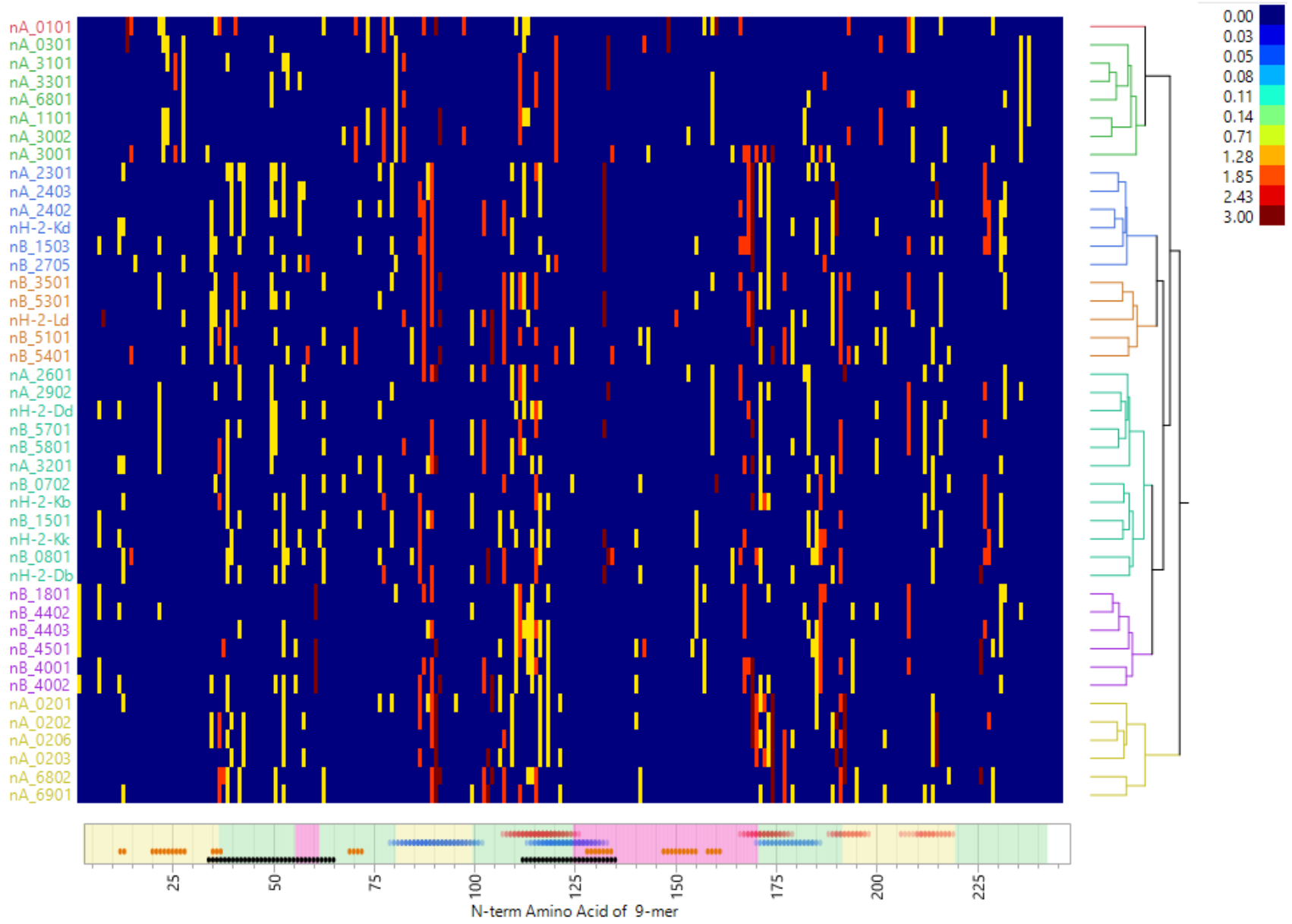


ioGenetics

Hierarchical Cluster MHC I

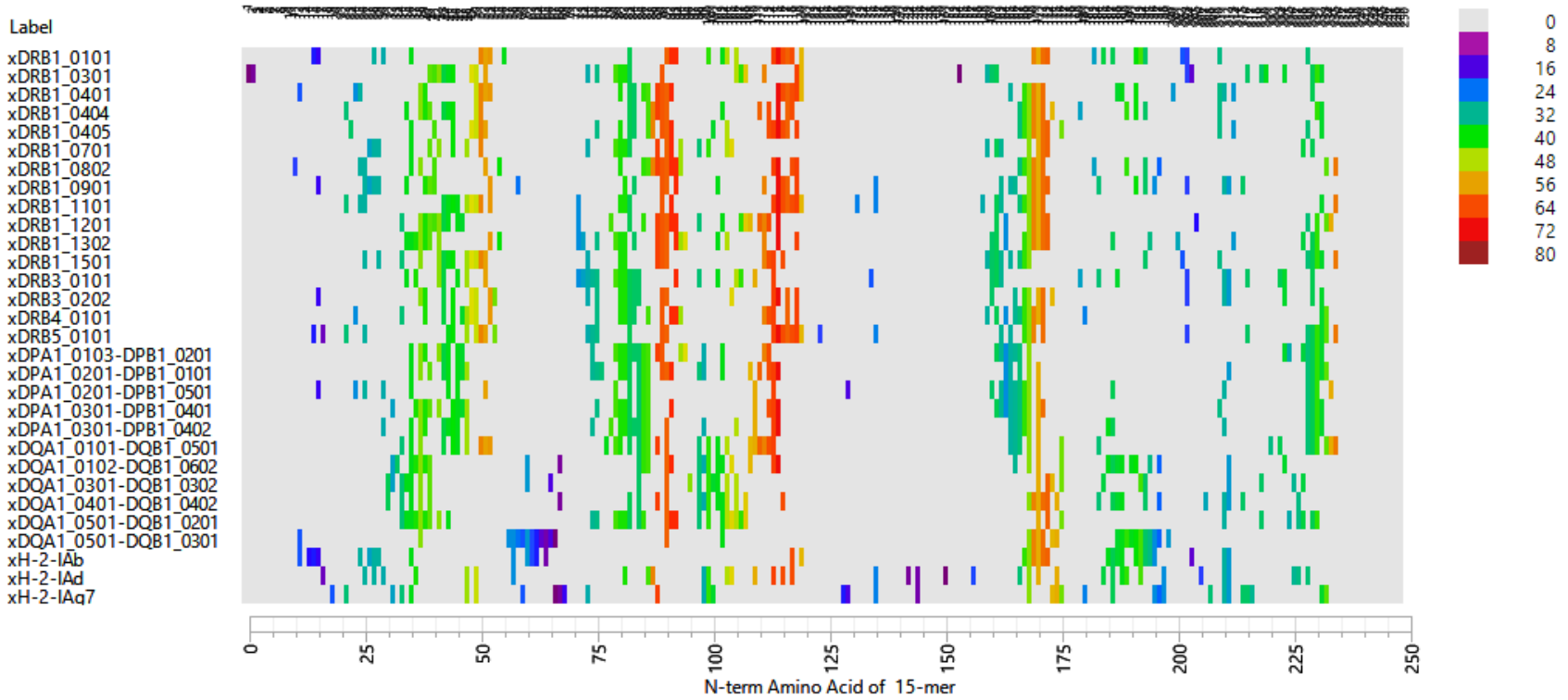
Method = Ward

JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t MHC I & Cathepsins consolidated 2



ioGenetics

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_non_structural_protein_NS4b VIPR_ALG4_402122761_6887_7636_1;t



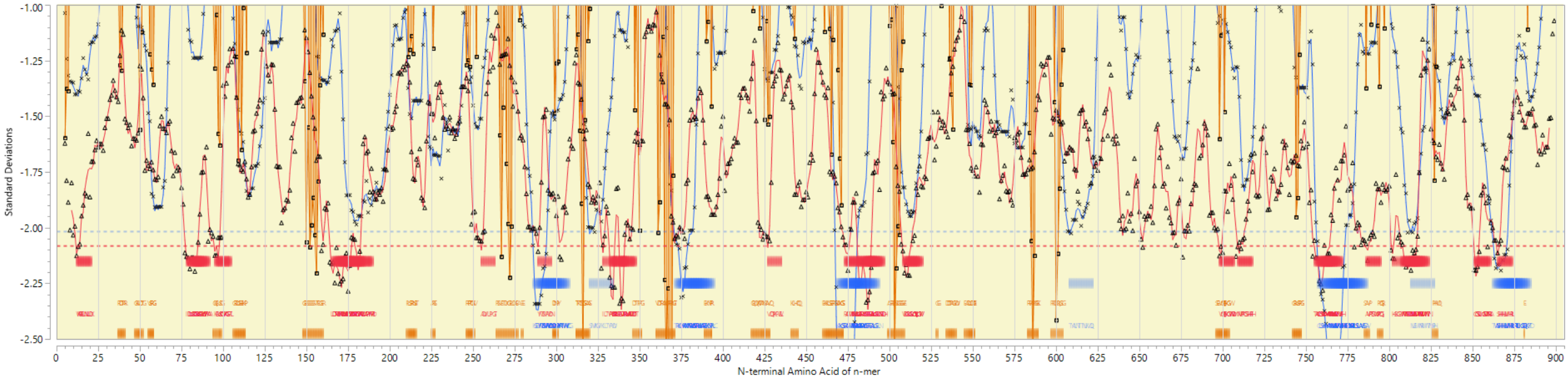
ioGenetics

JN811140_RNA_dependent_RNA_polymerase_NS5

ioGenetics

Overlay Plot

JN811140_RNA_dependent_RNA_polymerase
VIPR_ALG4_402122761_7637_10351_1;t

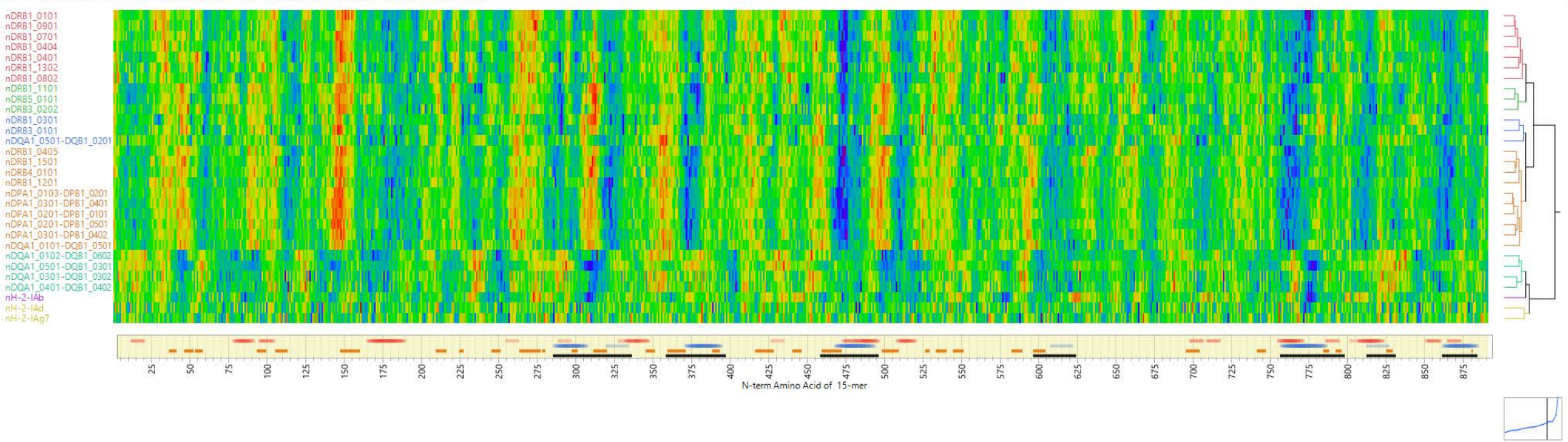


ioGenetics

Hierarchical Cluster MHC II

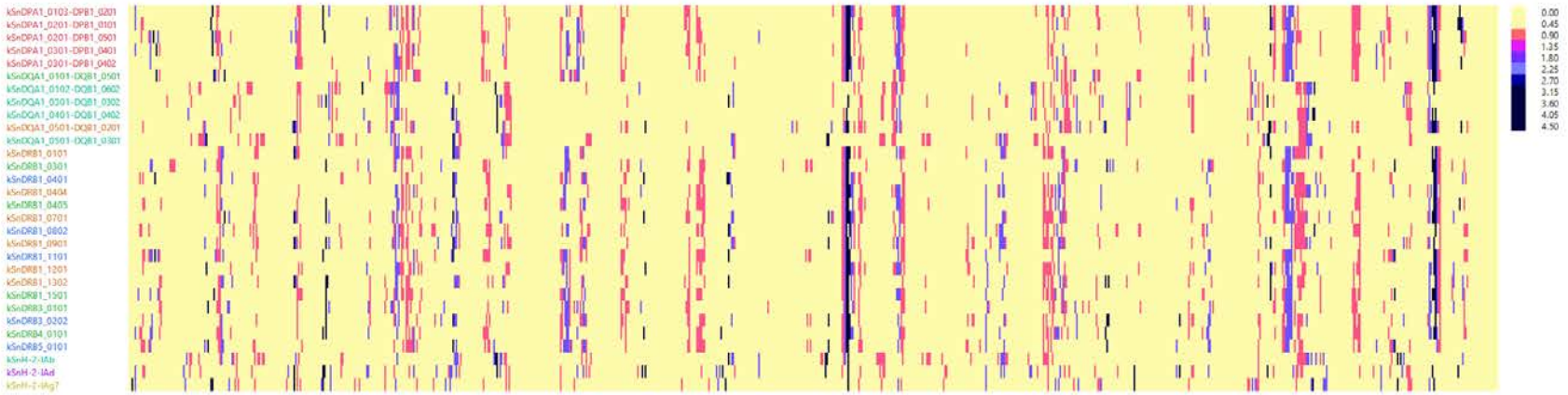
Method = Ward

JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1;t_T_MHC_II

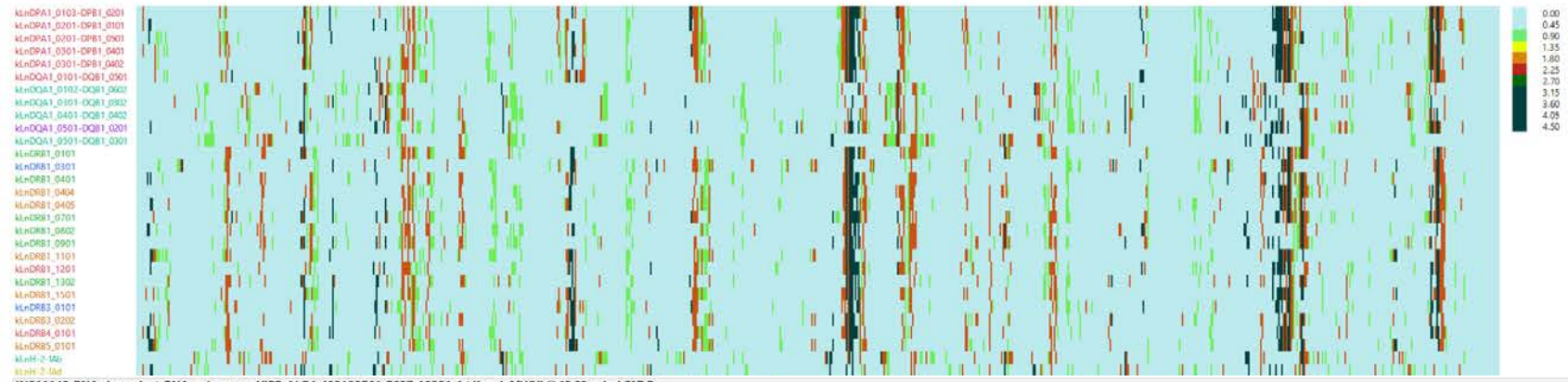


ioGenetics

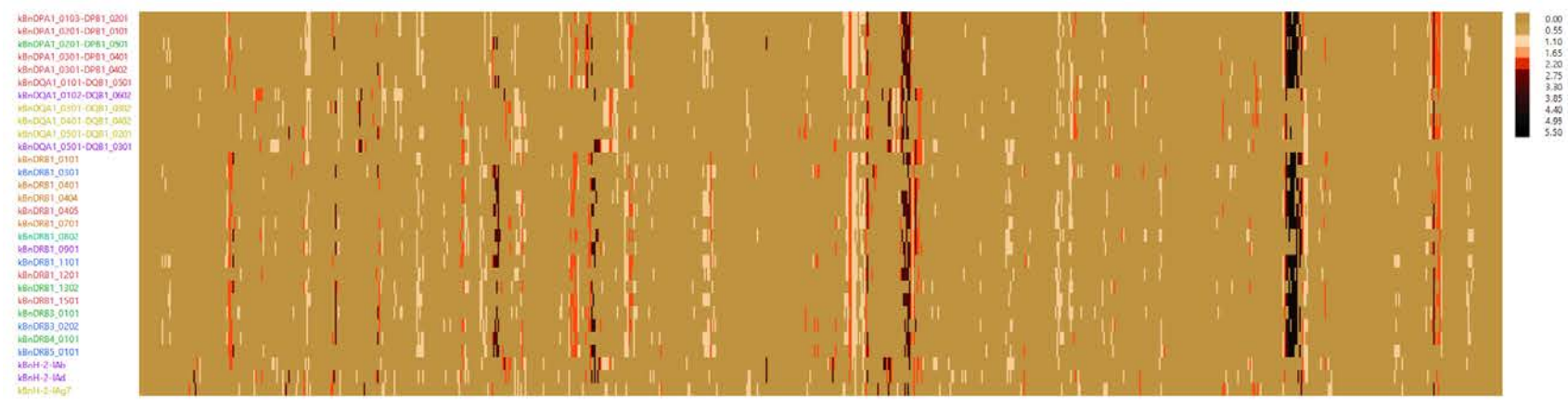
JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1:t Kernels MHC II @ 15-20 aa by hCAT_S



JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1:t Kernels MHC II @ 15-20 aa by hCAT_L



JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1:t Kernels MHC II @ 15-20 aa by hCAT_B

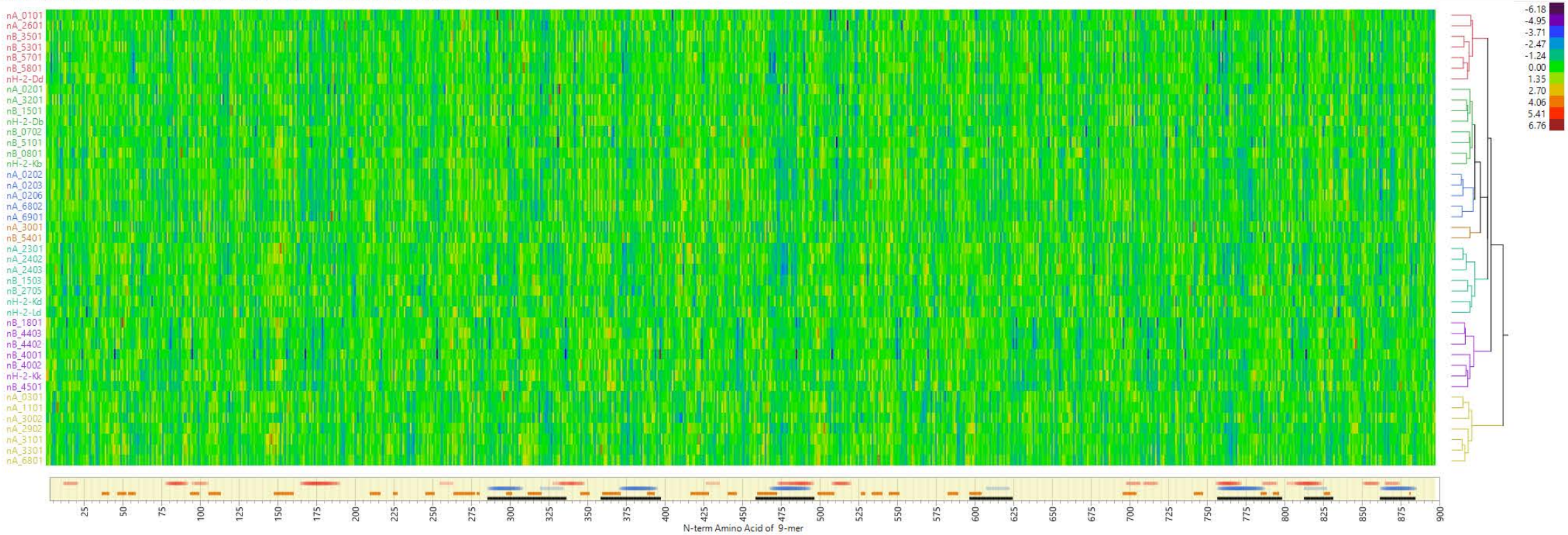


ioGenetics

Hierarchical Cluster MHC I

Method = Ward

JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1;t_T_MHC_I

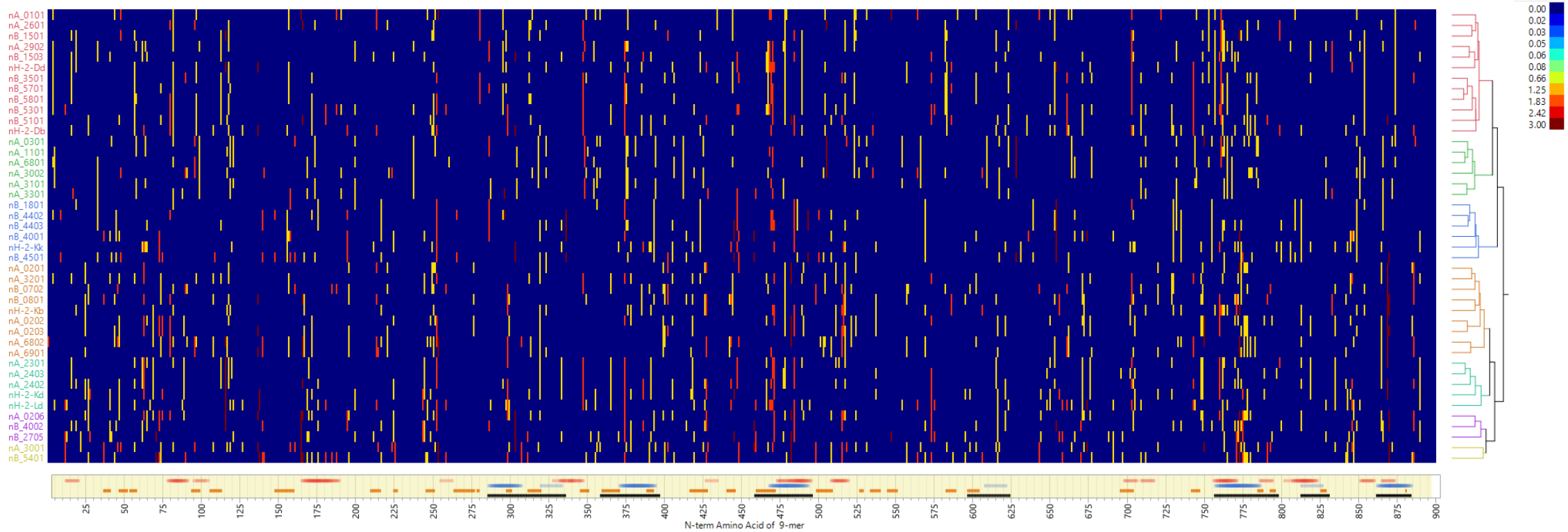


ioGenetics

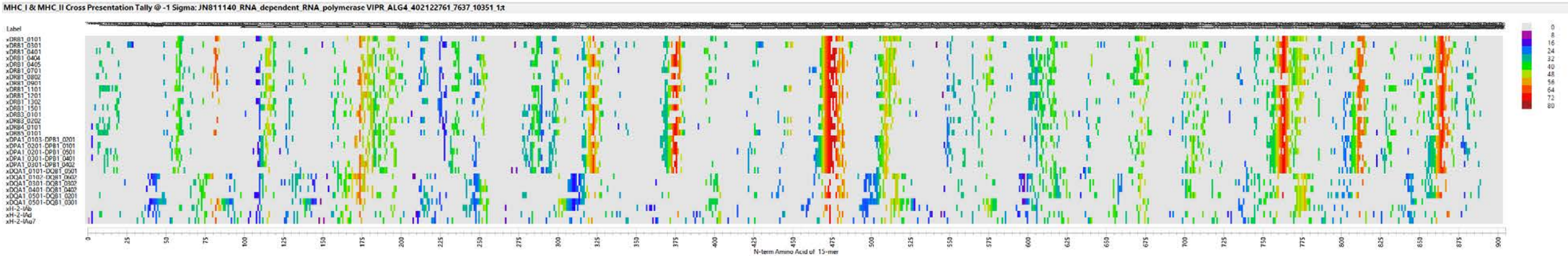
Hierarchical Cluster MHC I

Method = Ward

JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1;t MHC I & Cathepsins consolidated



ioGenetics


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: JN811140_RNA_dependent_RNA_polymerase VIPR_ALG4_402122761_7637_10351_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900
N-term Amino Acid of 15-mer
0 8 16 24 32 40 48 56 64 72 80

ioGenetics

ioGenetics®

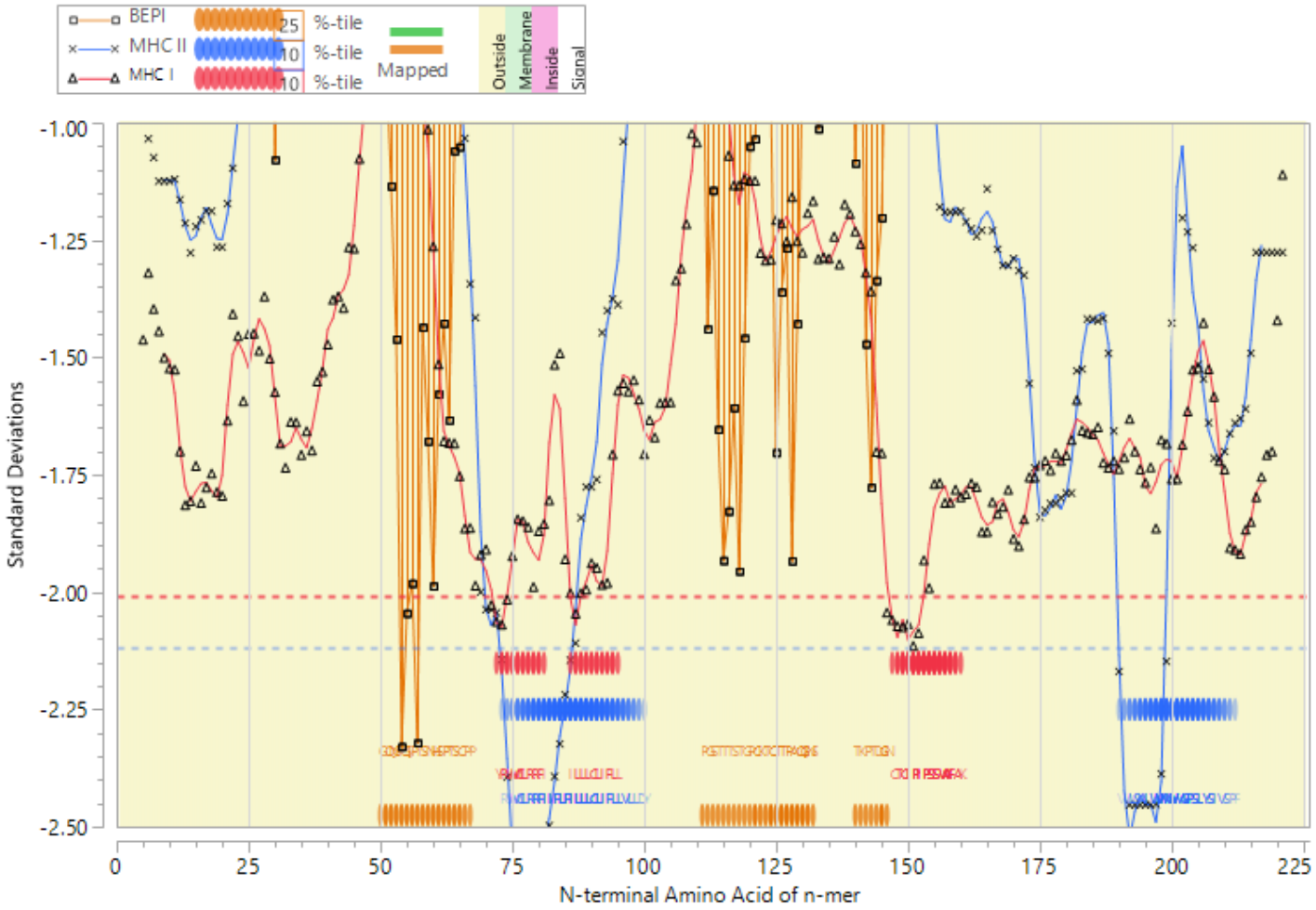
Addendum 2 Hepatitis B

HbSAg

Based on Valenzuela sequence

Overlay Plot

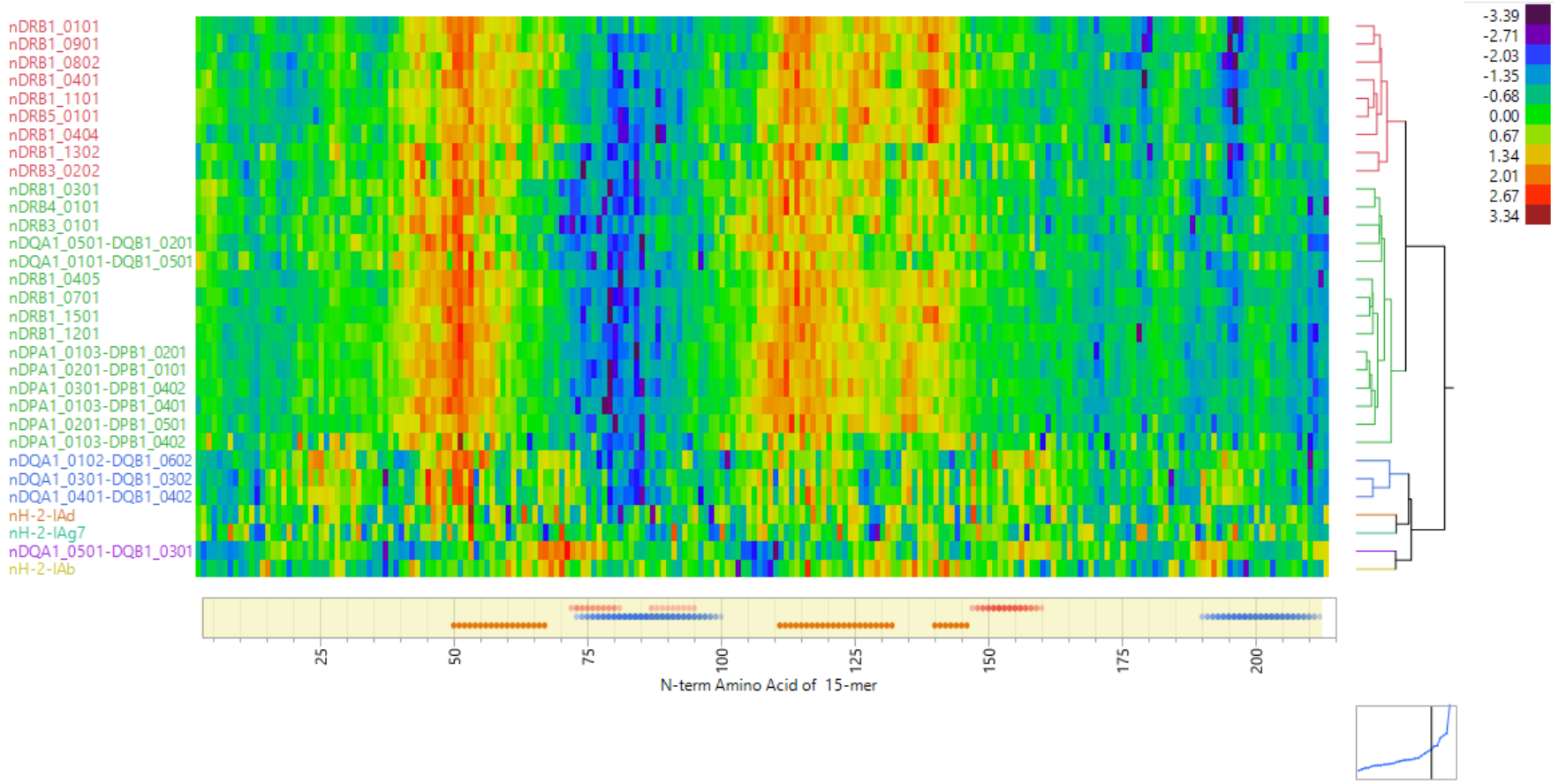
HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t



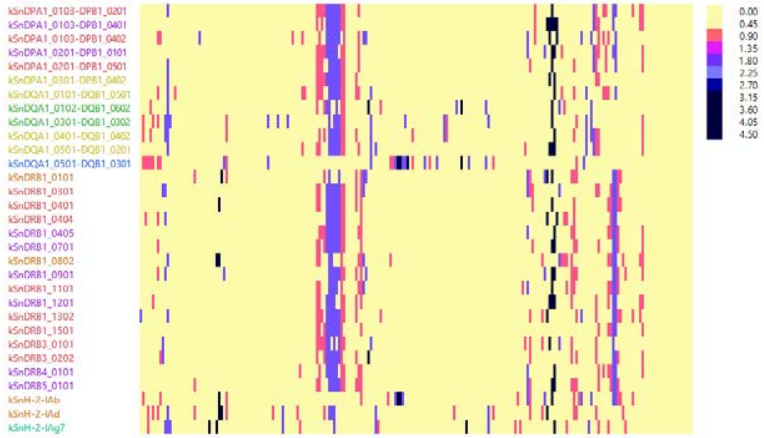
Hierarchical Cluster MHC II

Method = Ward

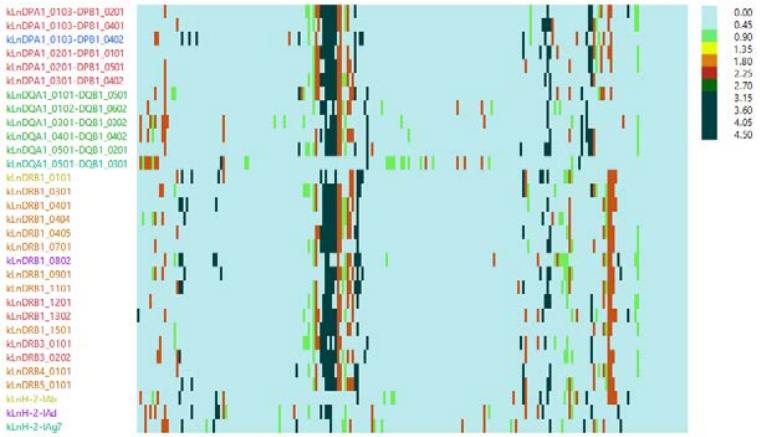
HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t_T_MHC_II



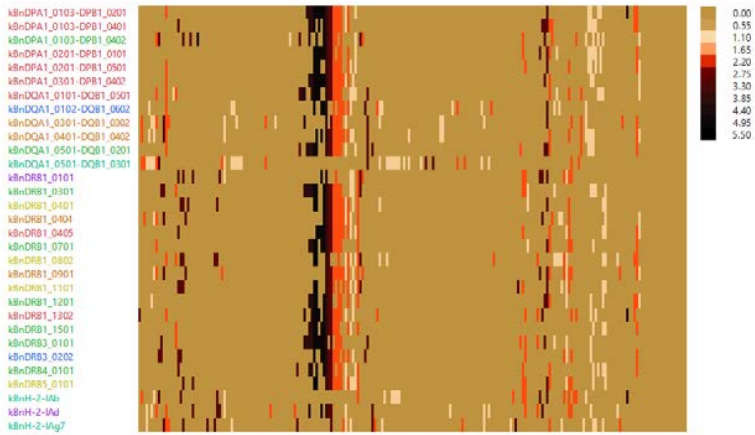
HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t Kernels MHC II @ 15-20 aa by hCAT_S



HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t Kernels MHC II @ 15-20 aa by hCAT_L



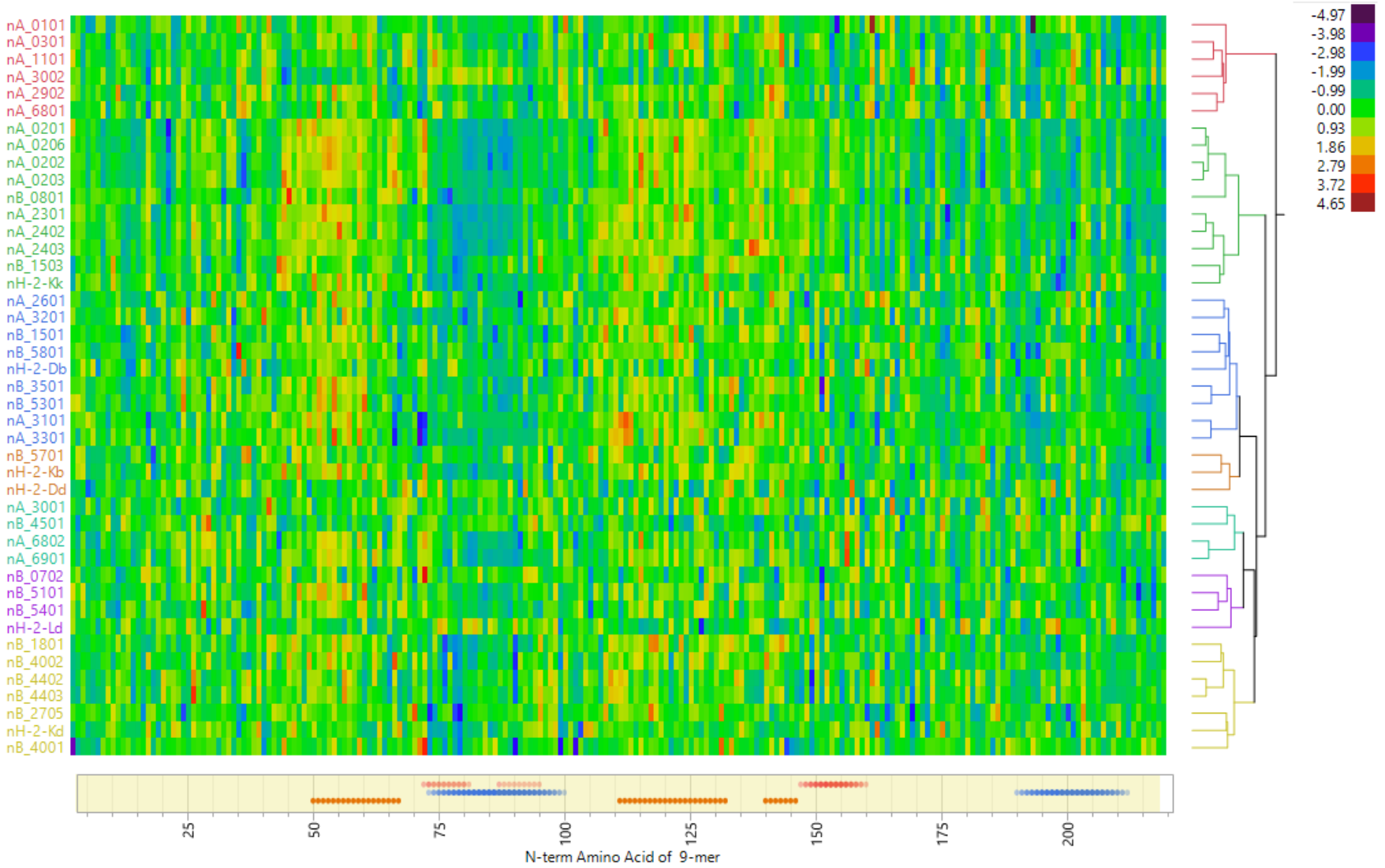
HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t Kernels MHC II @ 15-20 aa by hCAT_B



Hierarchical Cluster MHC I

Method = Ward

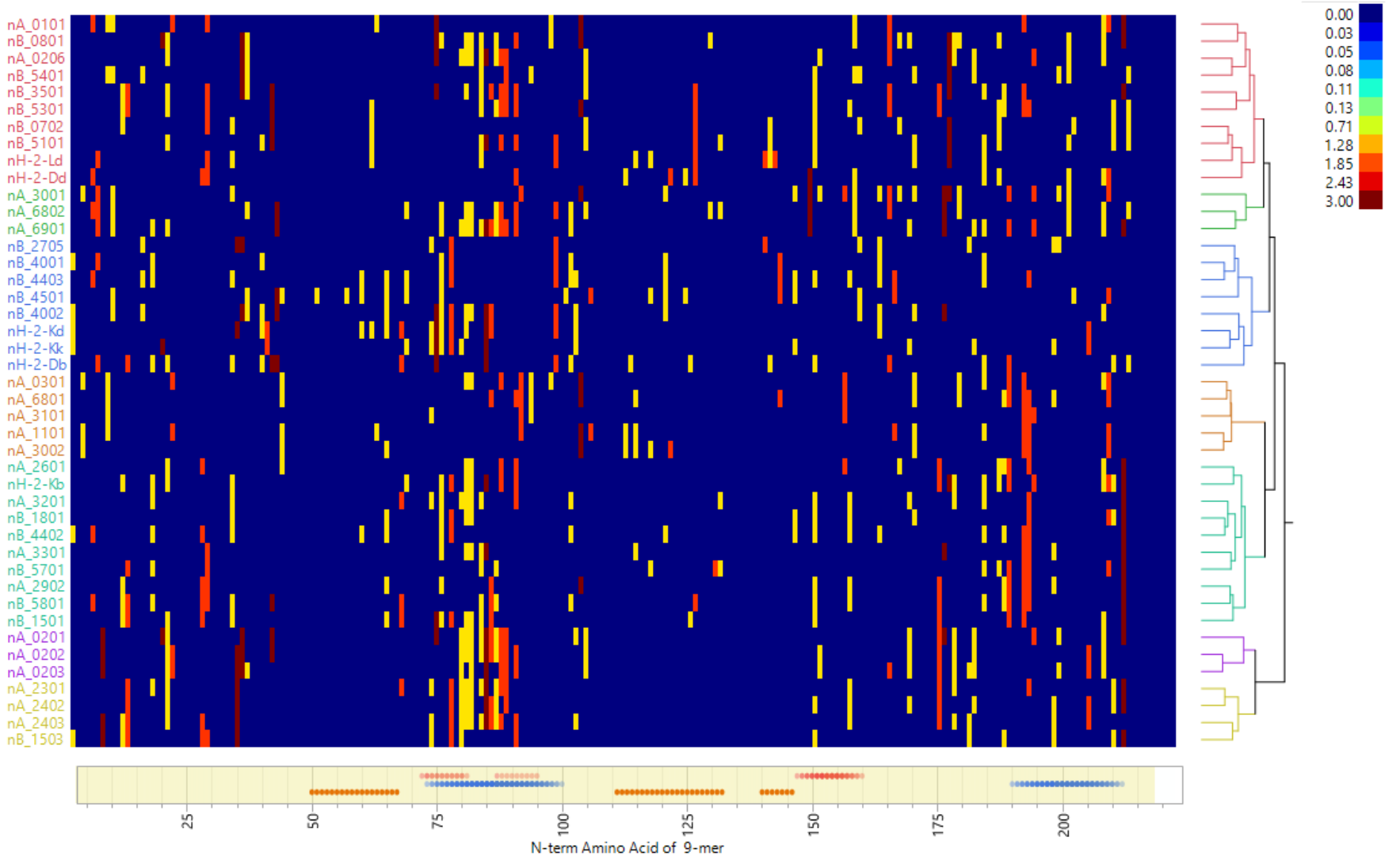
HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t_T_MHC_I 2

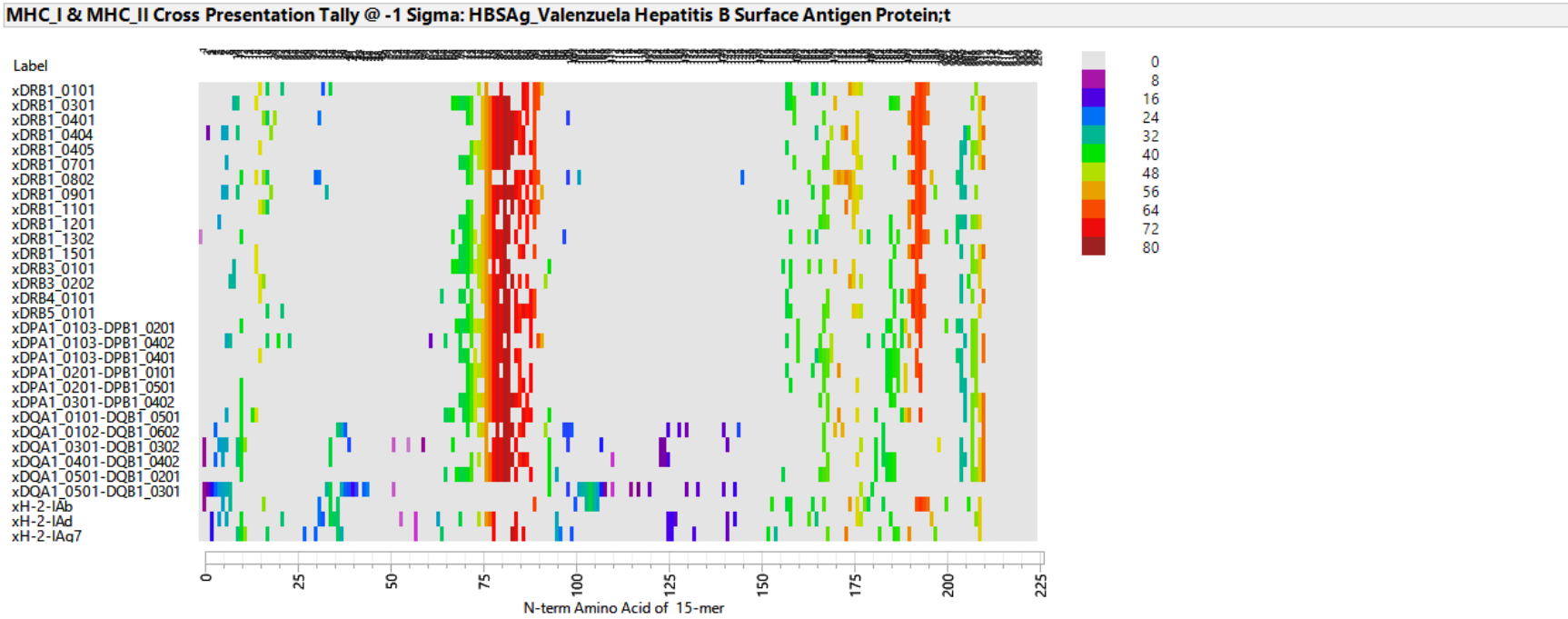


Hierarchical Cluster MHC I

Method = Ward

HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t MHC I & Cathepsins consolidated

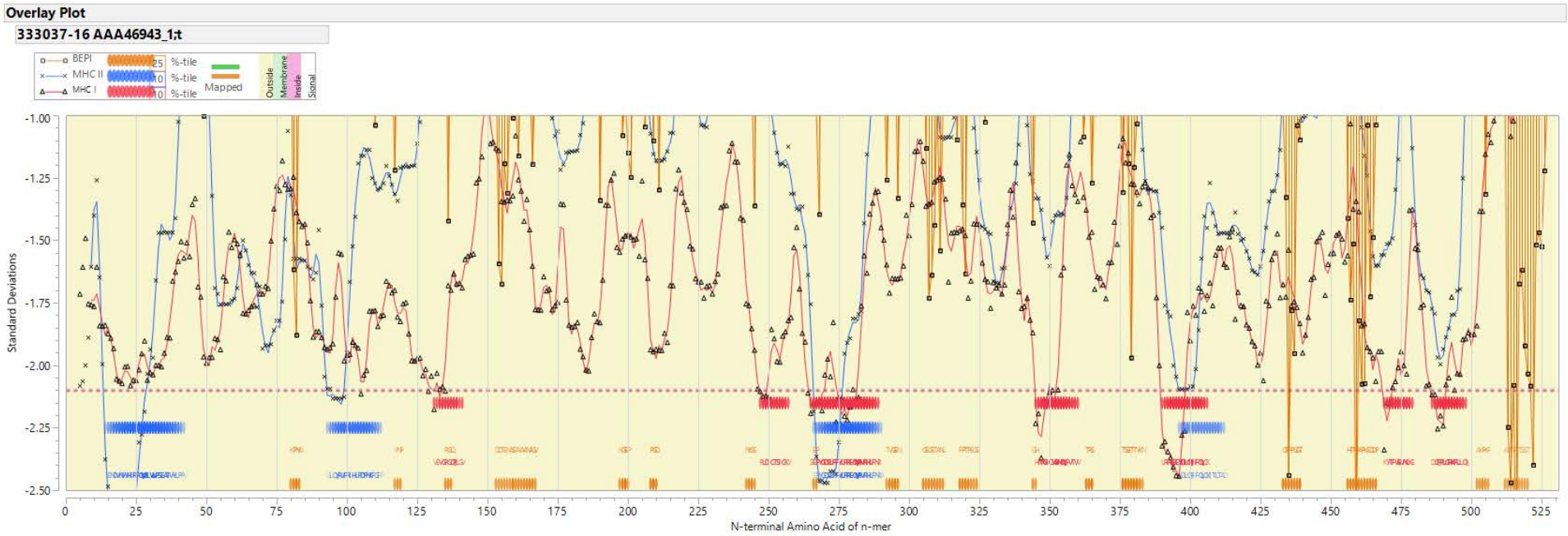


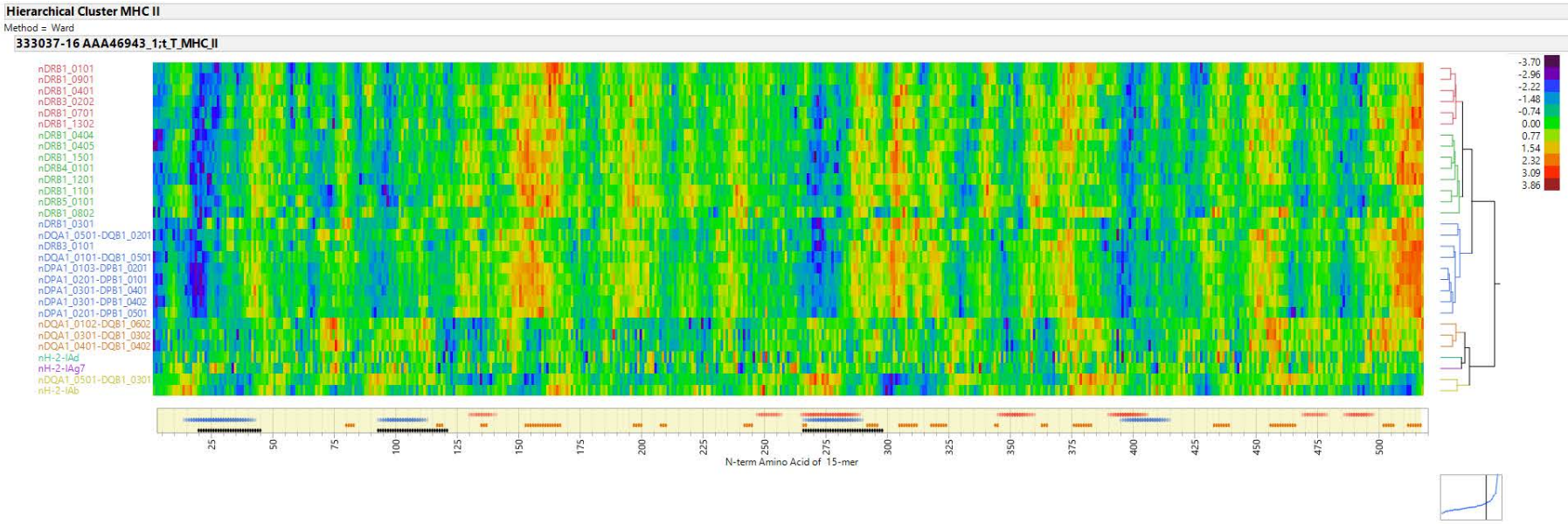

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: HBSAg_Valenzuela Hepatitis B Surface Antigen Protein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0103-DPB1_0402
xDPA1_0103-DPB1_0401
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

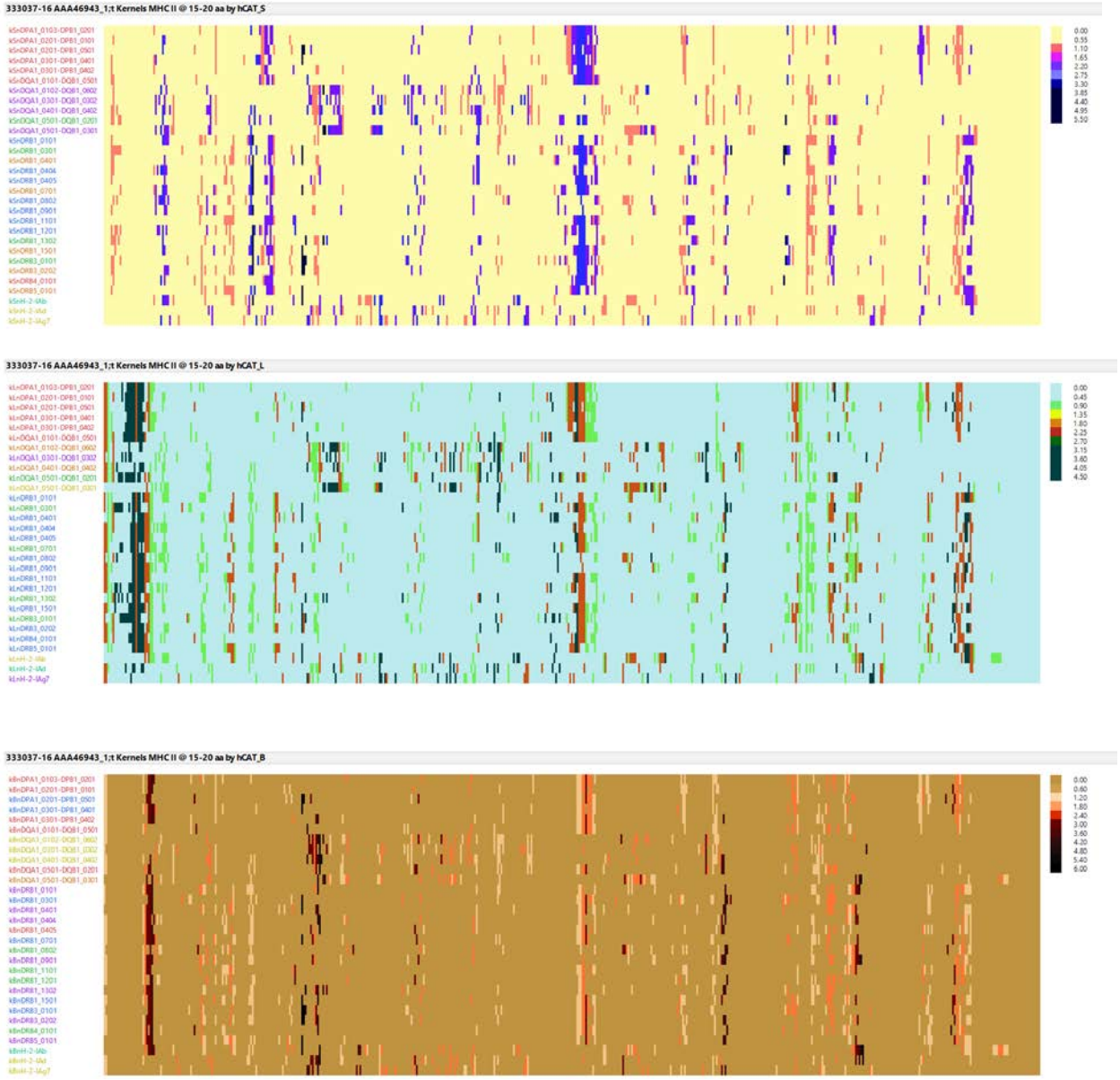
ioGenetics®

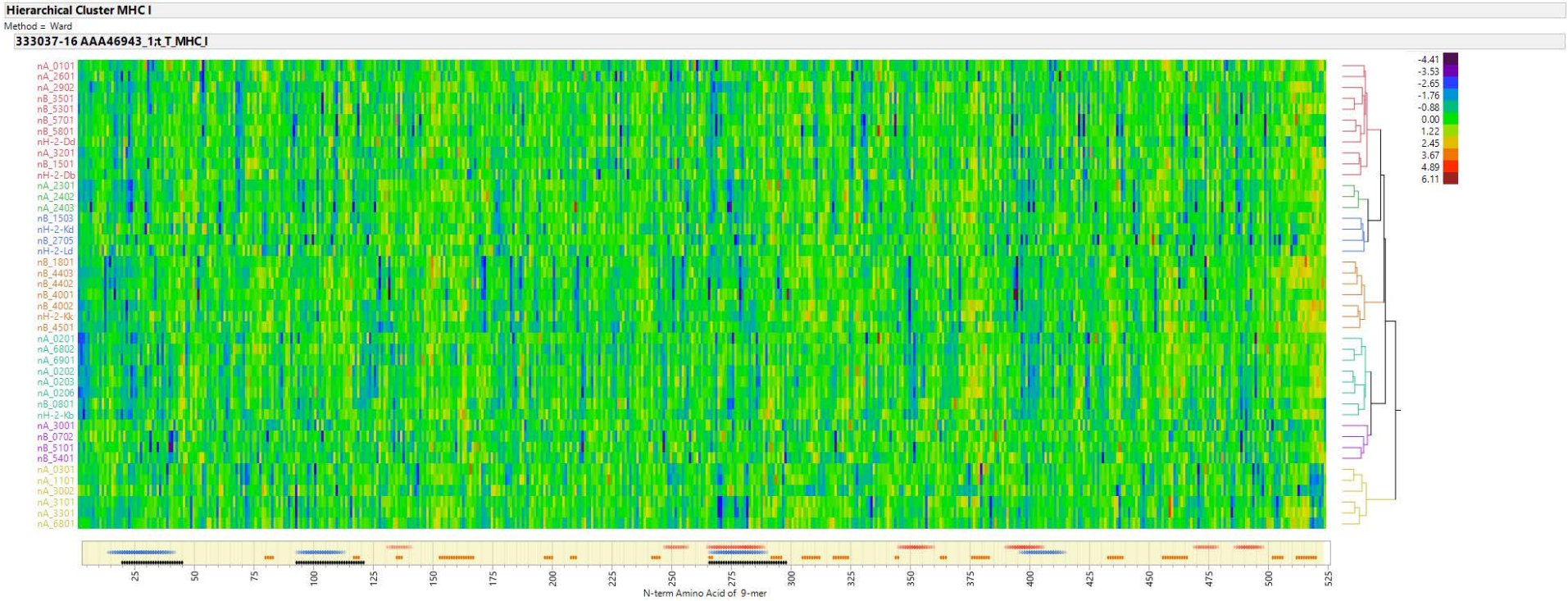
Addendum 3
Individual HPV L1

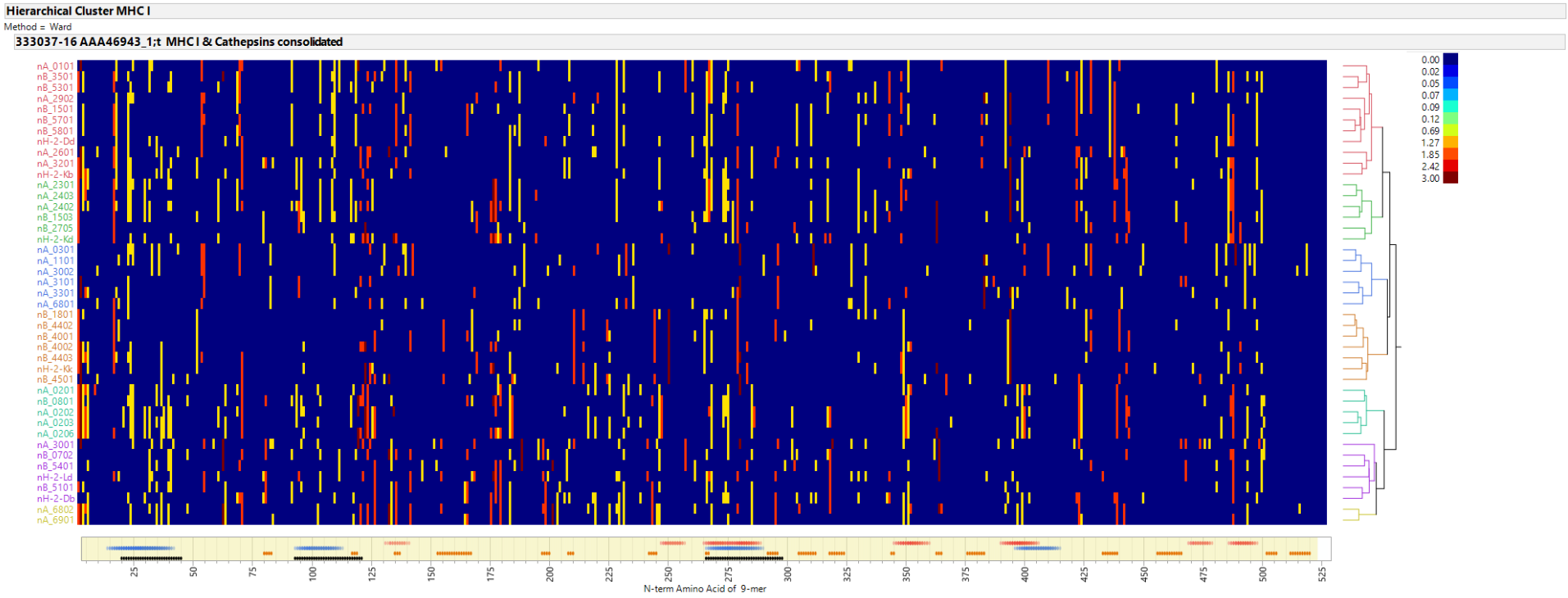
HPV 16

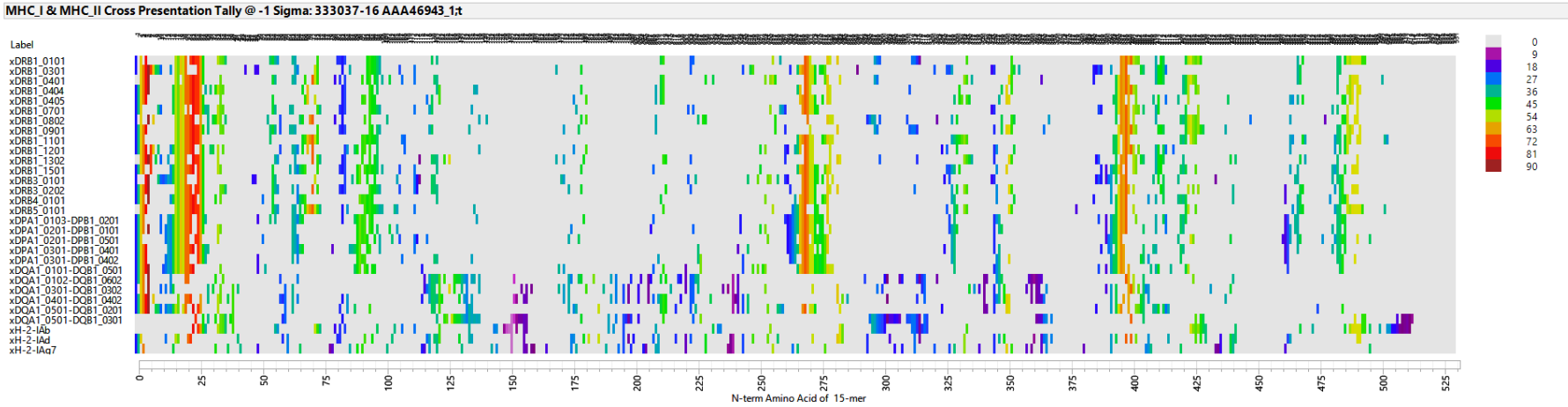

Overlay Plot
333037-16 AAA46943_1;t
BEPI
MHC II
MHC I
25 %-tile
50 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
333037-16 AAA46943_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0401
nDRB3_0202
nDRB1_0701
nDRB1_1302
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_0301
nDQA1_0501-DQB1_0201
nDRB3_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.70
-2.96
-2.22
-1.48
-0.74
0.00
0.77
1.54
2.32
3.09
3.86
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

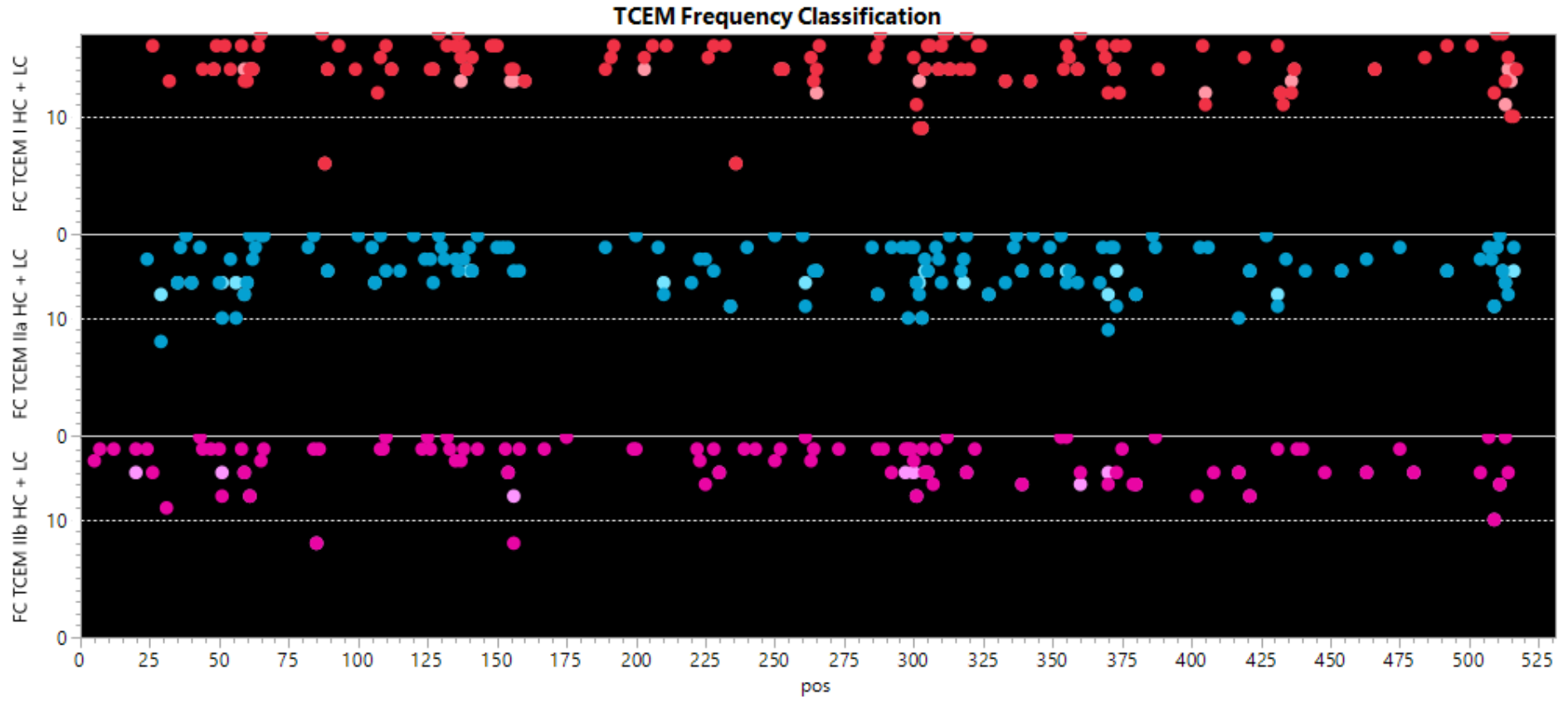

333037-16 AAA46943_1;t Kernels MHC II @ 15-20 aa by hCAT_S
333037-16 AAA46943_1;t Kernels MHC II @ 15-20 aa by hCAT_L
333037-16 AAA46943_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
333037-16 AAA46943_1;t_T_MHC I
nA_0101
nA_2601
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_3201
nB_1501
nH-2-Db
nA_2301
nA_2402
nA_2403
nB_1503
nH-2-Kd
nB_2705
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_3001
nB_0702
nB_5101
nB_5401
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
-4.41
-3.53
-2.65
-1.76
-0.88
0.00
1.22
2.45
3.67
4.89
6.11
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer

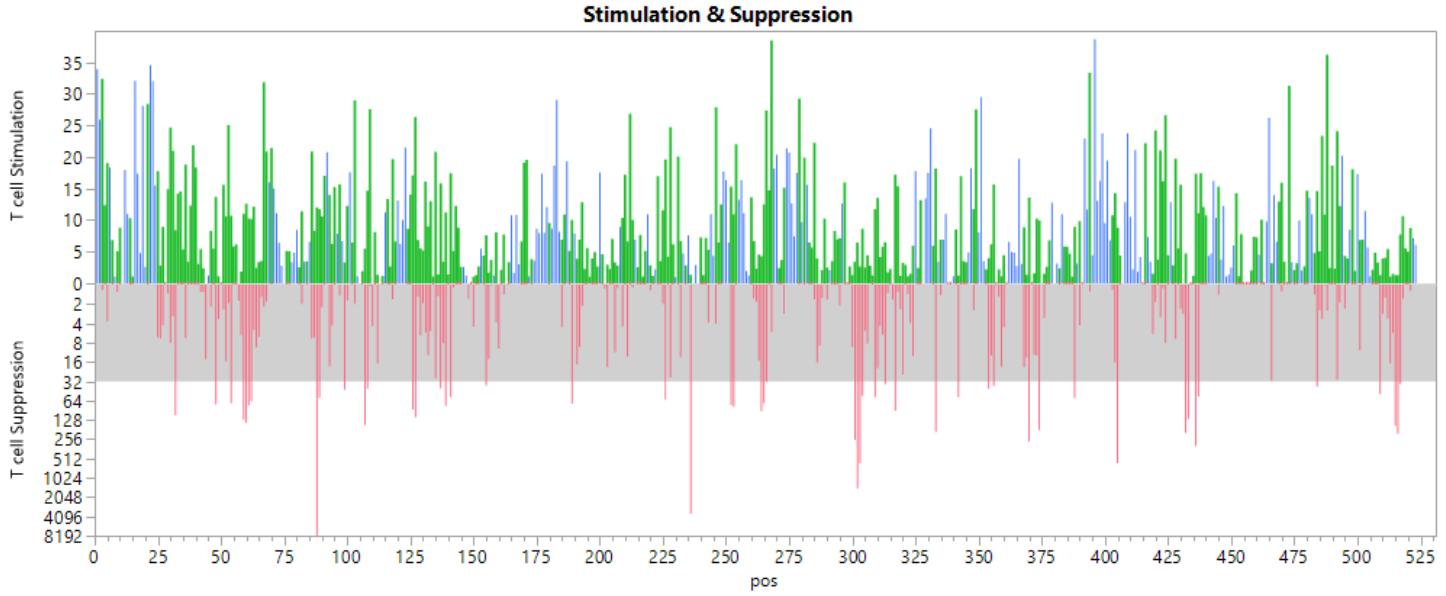

Hierarchical Cluster MHC I
Method = Ward
333037-16 AAA46943_1;t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nB_5301
nA_2902
nB_1501
nB_5701
nB_5801
nH-2-Dd
nA_2601
nA_3201
nH-2-Kb
nA_2301
nA_2403
nA_2402
nB_1503
nB_2705
nH-2-Kd
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nB_1801
nB_4402
nB_4001
nB_4002
nB_4403
nH-2-Kk
nB_4501
nA_0201
nB_0801
nA_0202
nA_0203
nA_0206
nA_3001
nB_0702
nB_5401
nH-2-Ld
nB_5101
nH-2-Db
nA_6802
nA_6901
0.00
0.02
0.05
0.07
0.09
0.12
0.69
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 333037-16 AAA46943_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525
N-term Amino Acid of 15-mer
0
9
18
27
36
45
54
63
72
81
90

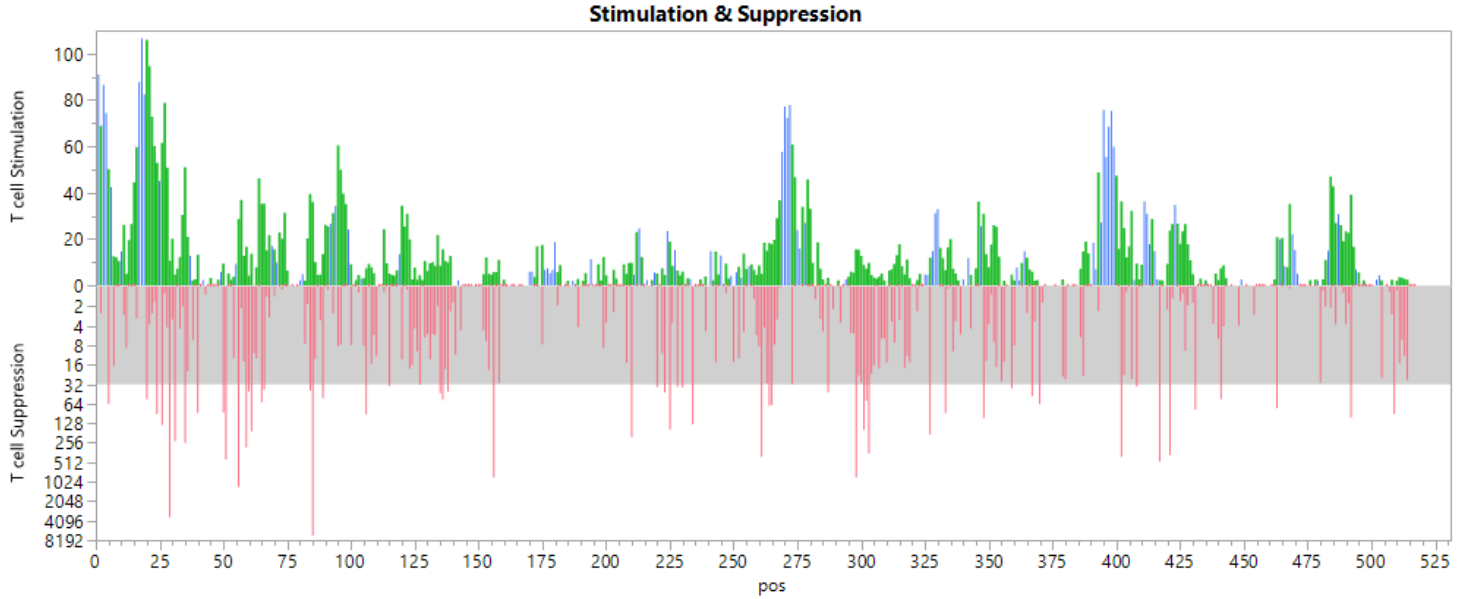
333037-16 AAA46943_1



333037-16 AAA46943_1



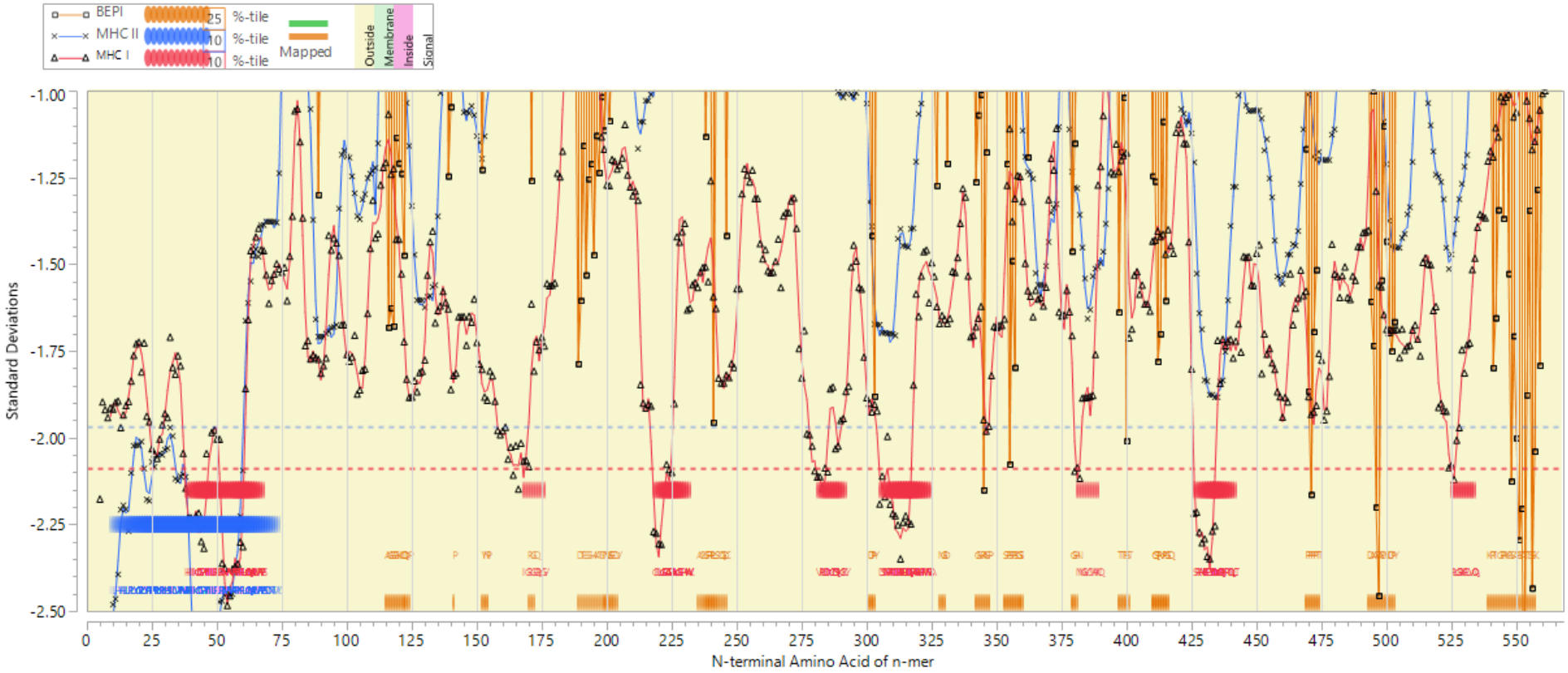
333037-16 AAA46943_1

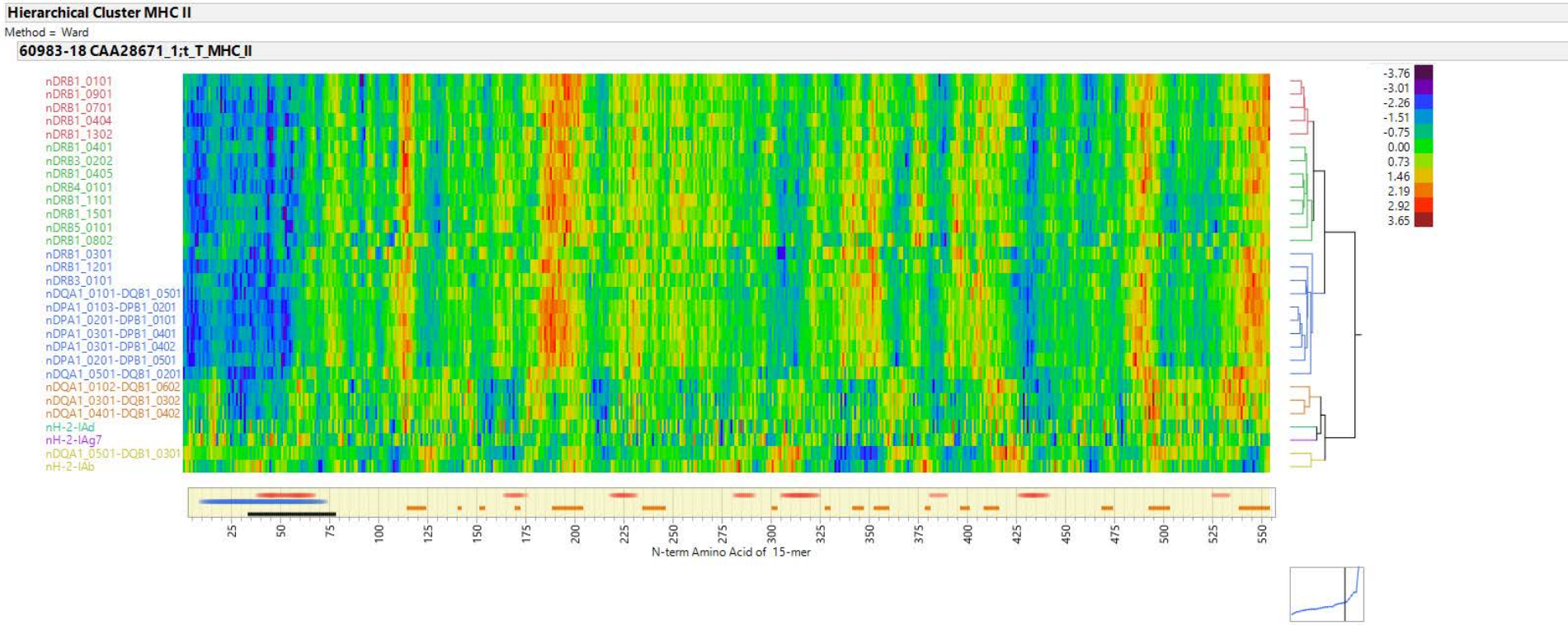


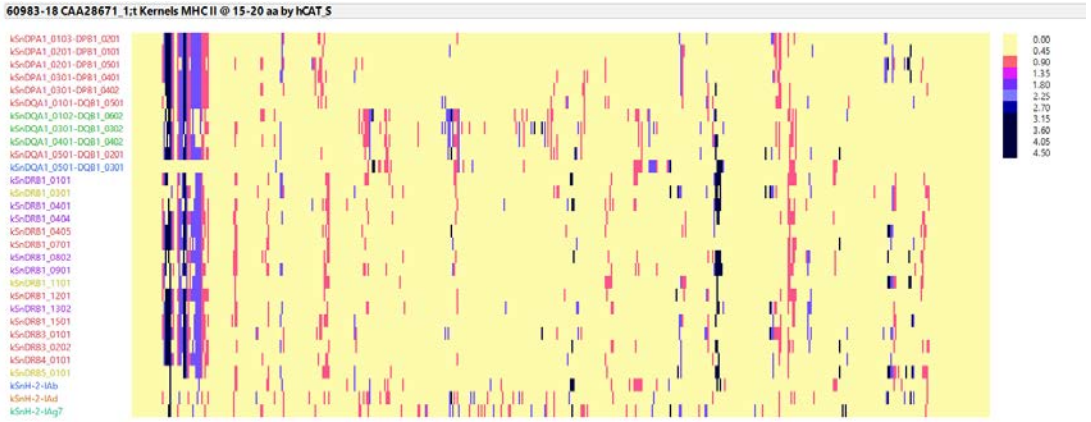
HPV 18

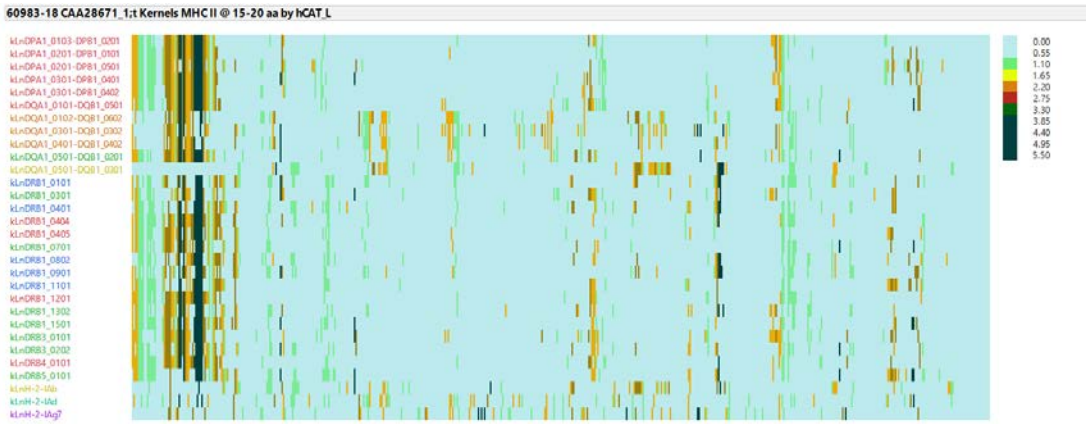
Overlay Plot

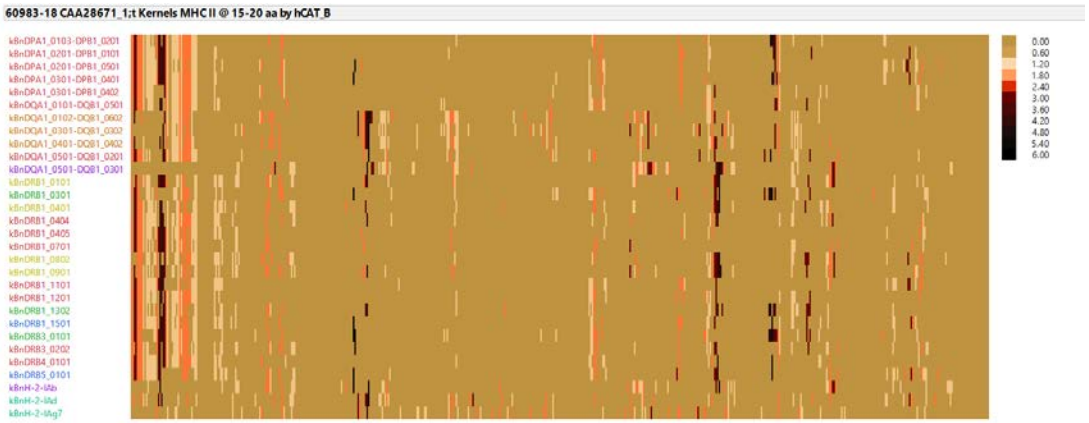
60983-18 CAA28671_1;t



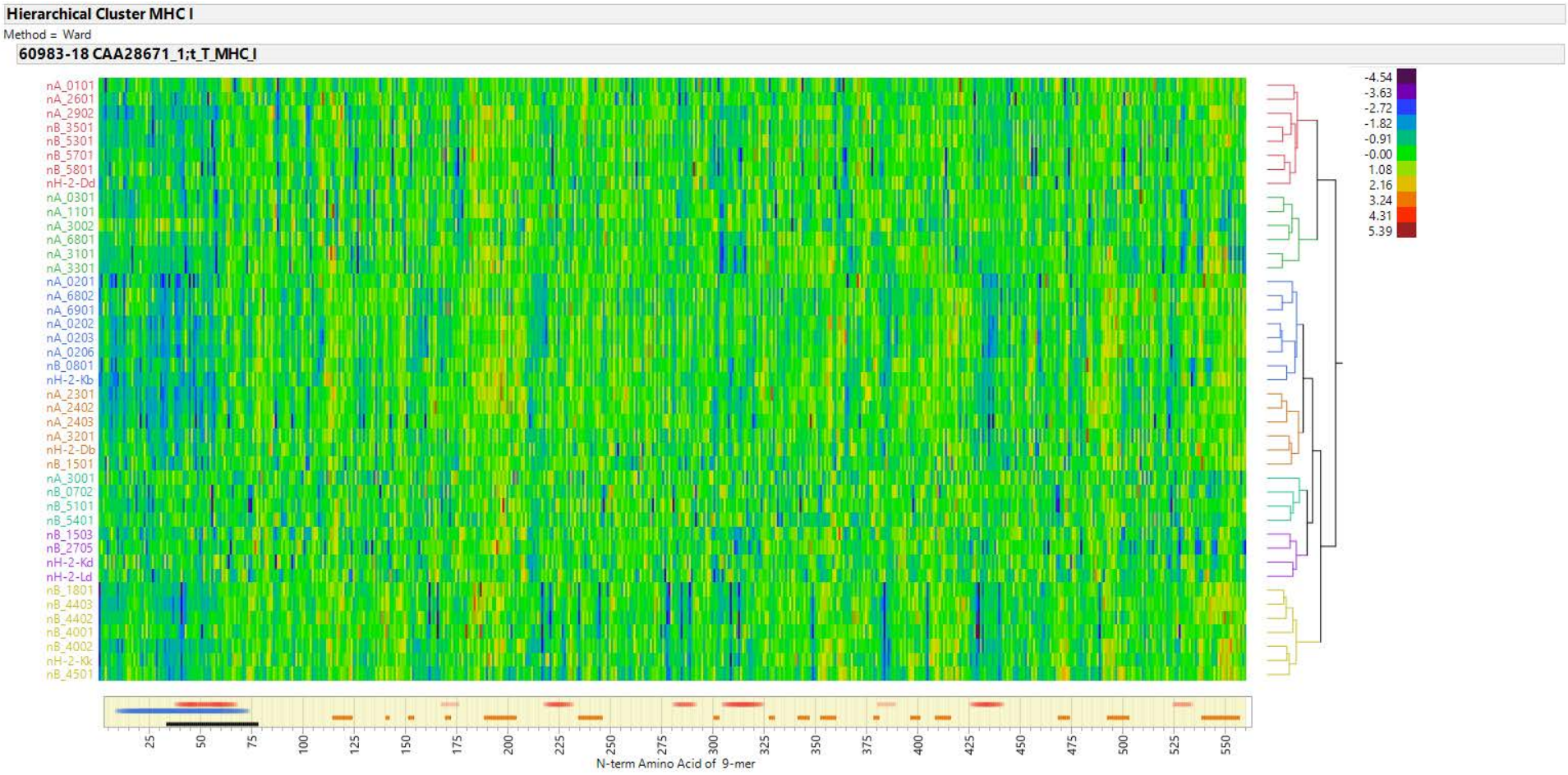

Hierarchical Cluster MHC II
Method = Ward
60983-18 CAA28671_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0405
nDRB4_0101
nDRB1_1101
nDRB1_1501
nDRB5_0101
nDRB1_0802
nDRB1_0301
nDRB1_1201
nDRB3_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.76
-3.01
-2.26
-1.51
-0.75
0.00
0.73
1.46
2.19
2.92
3.65
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 15-mer

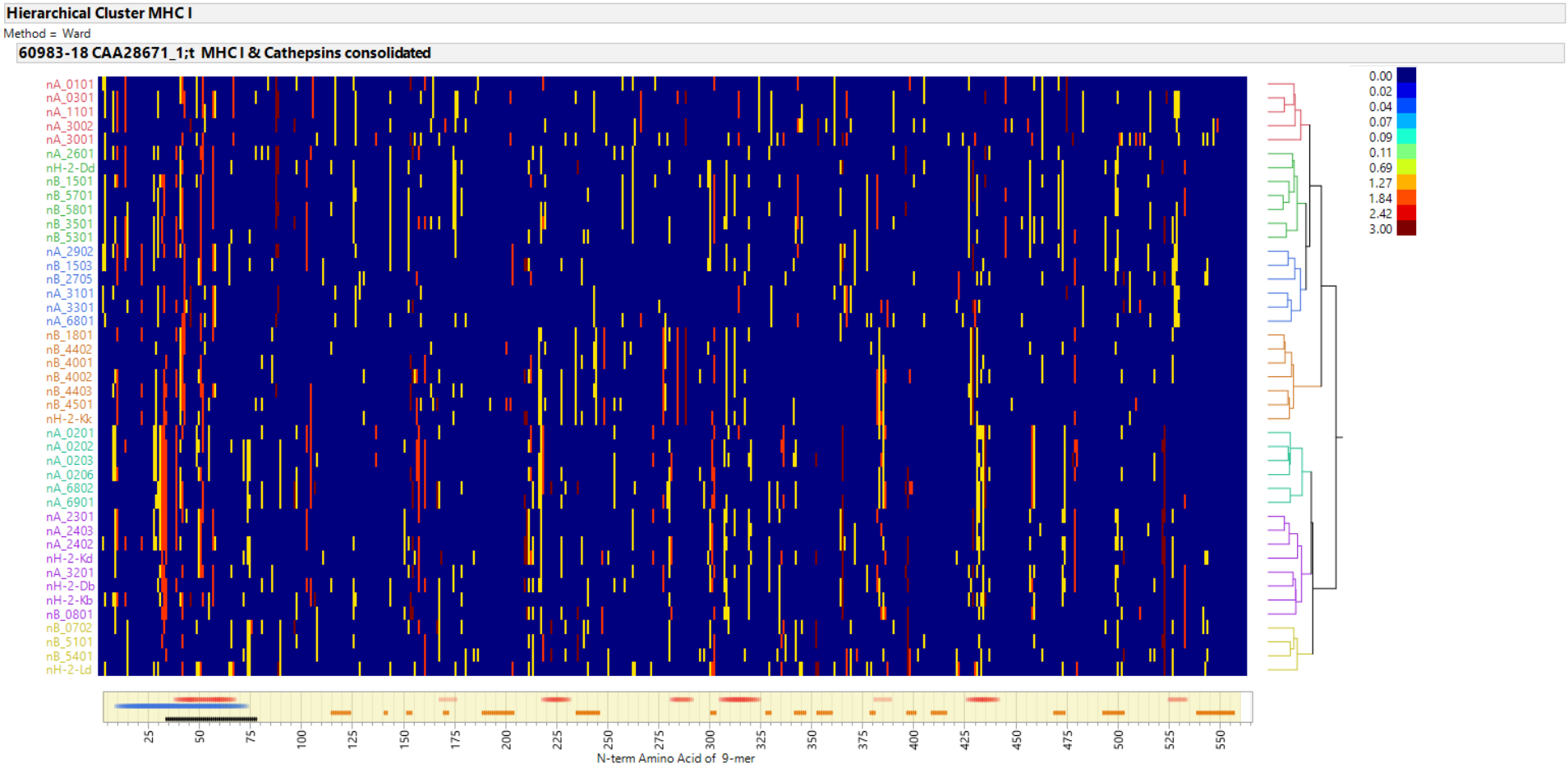

60983-18 CAA28671_1;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

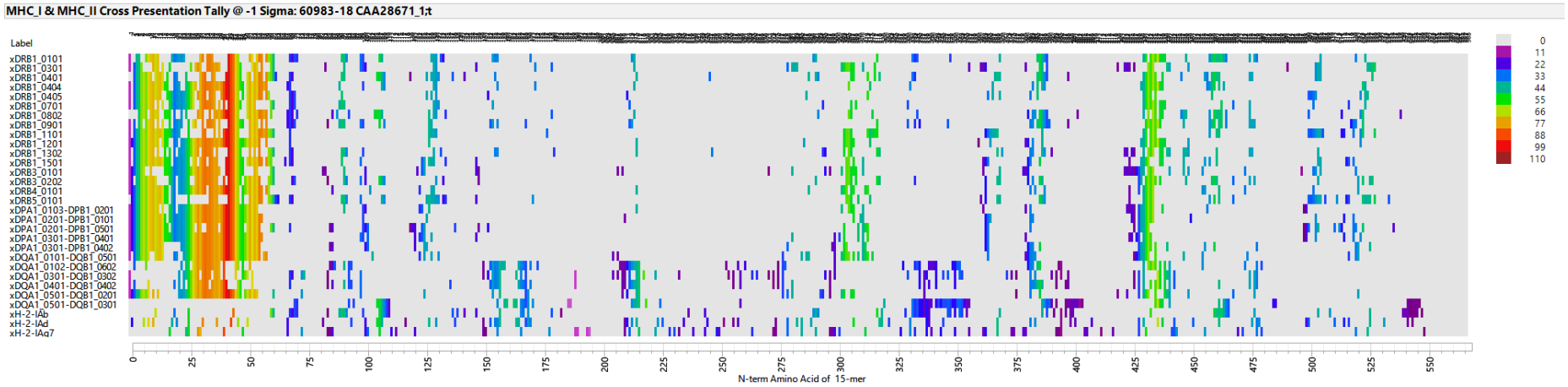

60983-18 CAA28671_1;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50


60983-18 CAA28671_1;t Kernels MHC II @ 15-20 aa by hCAT_B
kBnDPA1_0103-DPB1_0201
kBnDPA1_0201-DPB1_0101
kBnDPA1_0201-DPB1_0501
kBnDPA1_0301-DPB1_0401
kBnDPA1_0301-DPB1_0402
kBnDQA1_0101-DQB1_0501
kBnDQA1_0102-DQB1_0602
kBnDQA1_0301-DQB1_0302
kBnDQA1_0401-DQB1_0402
kBnDQA1_0501-DQB1_0201
kBnDQA1_0501-DQB1_0301
kBnDRB1_0101
kBnDRB1_0301
kBnDRB1_0401
kBnDRB1_0404
kBnDRB1_0405
kBnDRB1_0701
kBnDRB1_0802
kBnDRB1_0901
kBnDRB1_1101
kBnDRB1_1201
kBnDRB1_1302
kBnDRB1_1501
kBnDRB3_0101
kBnDRB3_0202
kBnDRB4_0101
kBnDRB5_0101
kBnH-2-IAb
kBnH-2-IAd
kBnH-2-IAg7
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

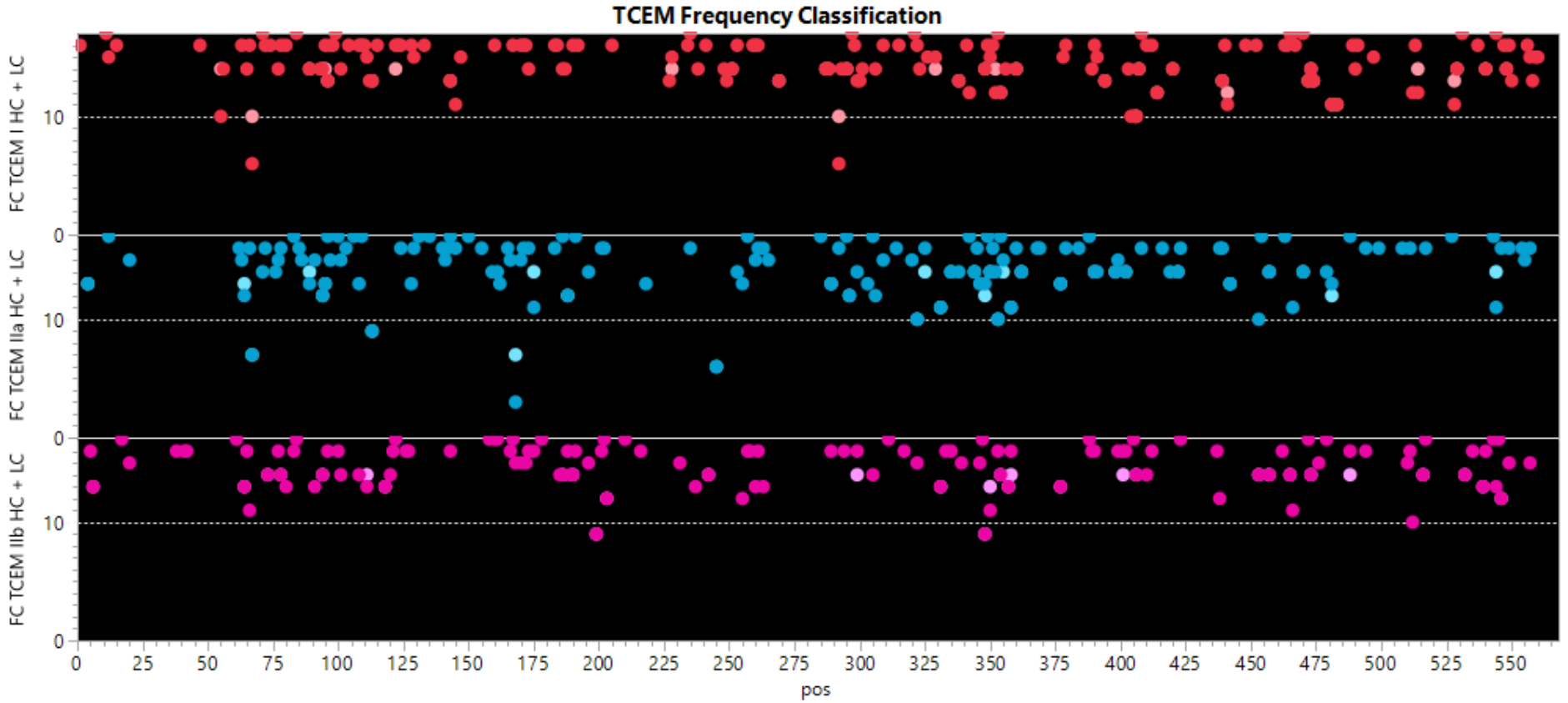
DTRA HPV

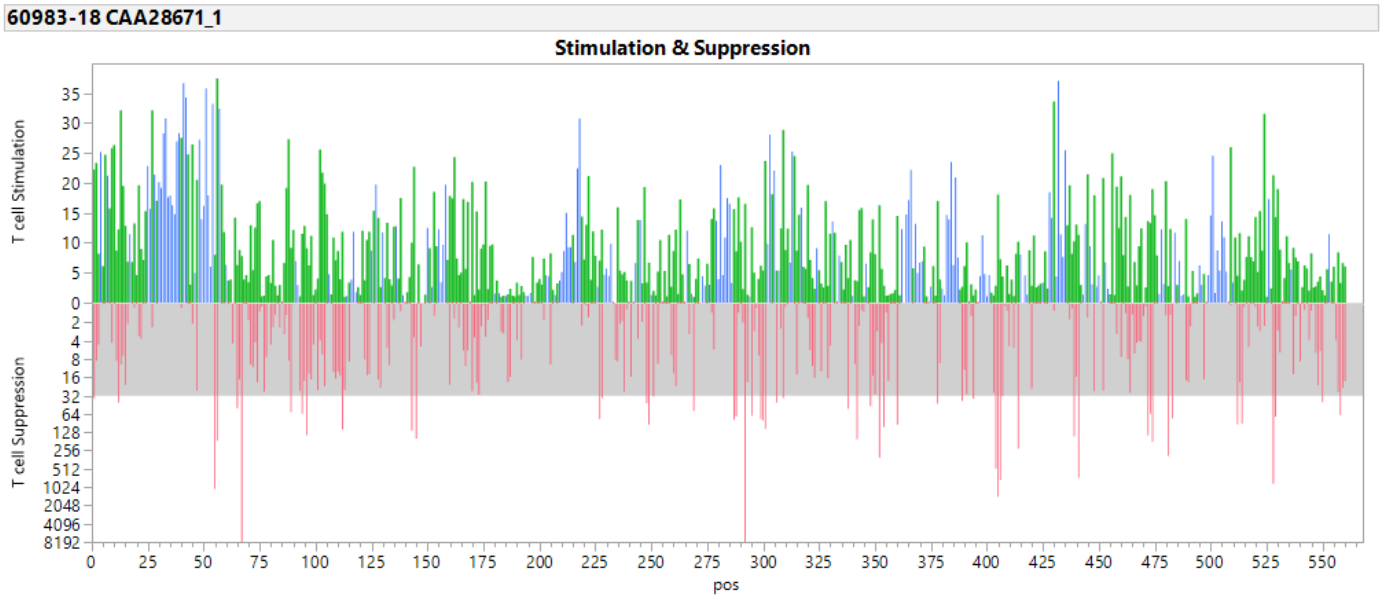

Hierarchical Cluster MHC I
Method = Ward
60983-18 CAA28671_1;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nA_3201
nH-2-Db
nB_1501
nA_3001
nB_0702
nB_5101
nB_5401
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
-4.54
-3.63
-2.72
-1.82
-0.91
-0.00
1.08
2.16
3.24
4.31
5.39
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 9-mer

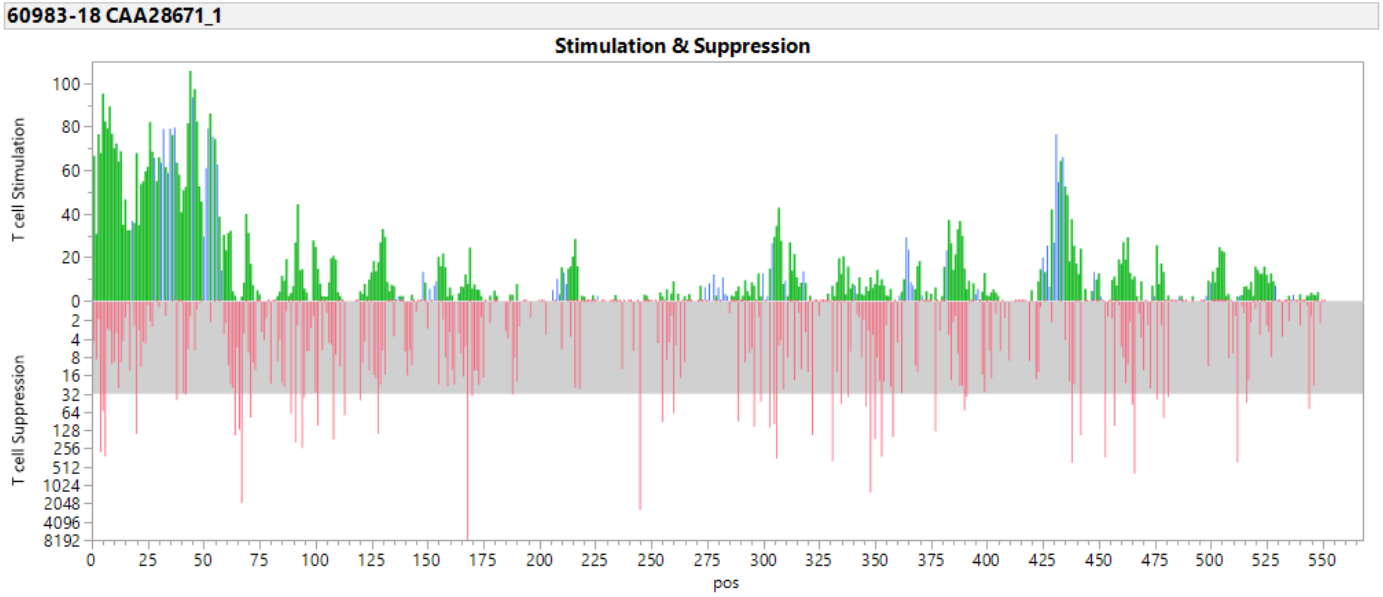

Hierarchical Cluster MHC I
Method = Ward
60983-18 CAA28671_1;t MHC I & Cathepsins consolidated
nA_0101
nA_0301
nA_1101
nA_3002
nA_3001
nA_2601
nH-2-Dd
nB_1501
nB_5701
nB_5801
nB_3501
nB_5301
nA_2902
nB_1503
nB_2705
nA_3101
nA_3301
nA_6801
nB_1801
nB_4402
nB_4001
nB_4002
nB_4403
nB_4501
nH-2-Kk
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_2301
nA_2403
nA_2402
nH-2-Kd
nA_3201
nH-2-Db
nH-2-Kb
nB_0801
nB_0702
nB_5101
nB_5401
nH-2-Ld
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 60983-18 CAA28671_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 15-mer
0
11
22
33
44
55
66
77
88
99
110

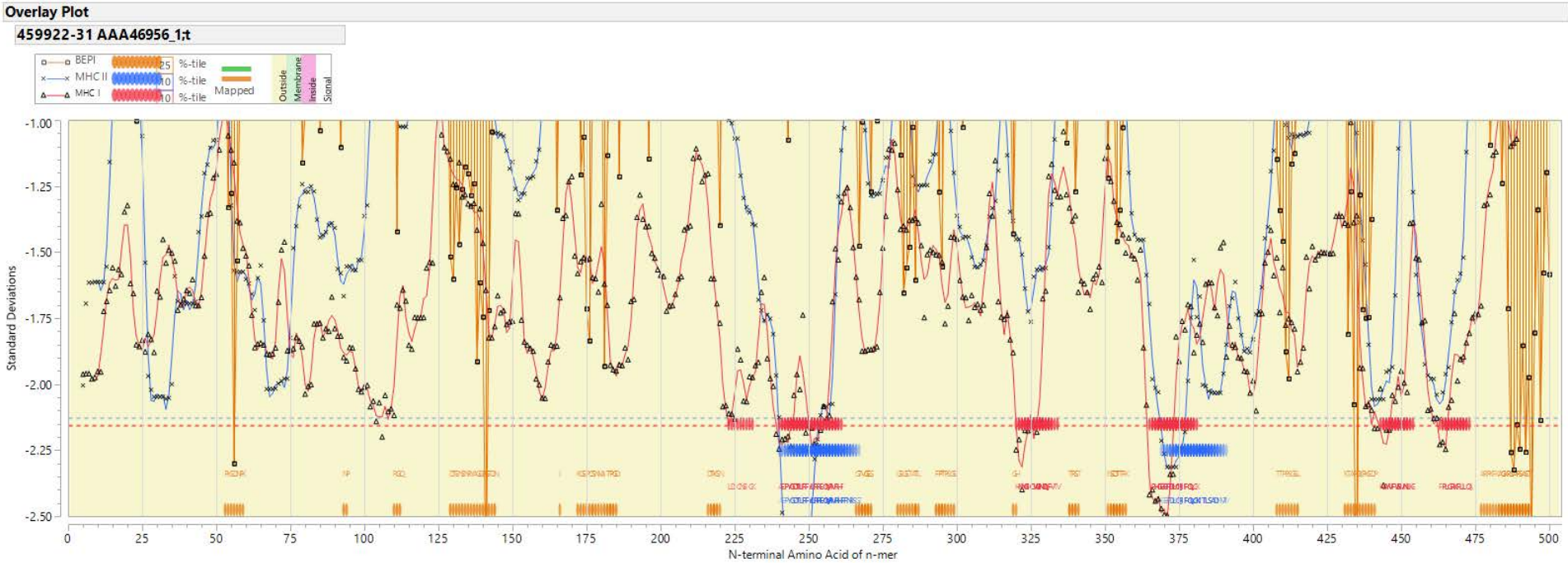
60983-18 CAA28671_1

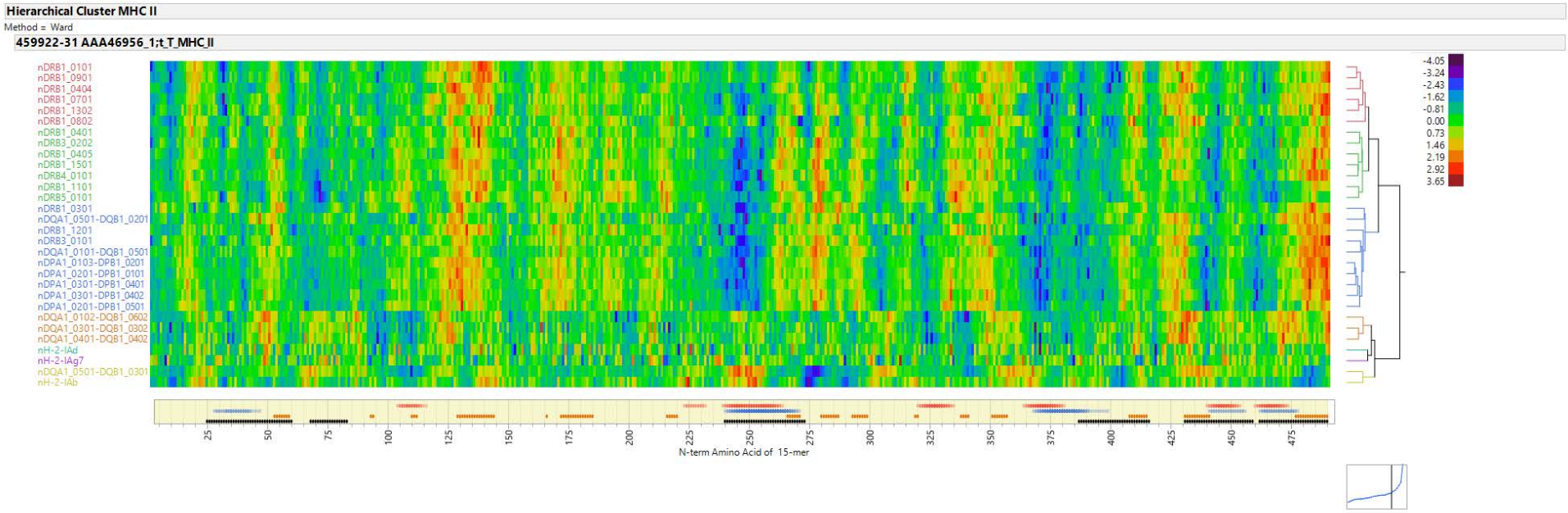


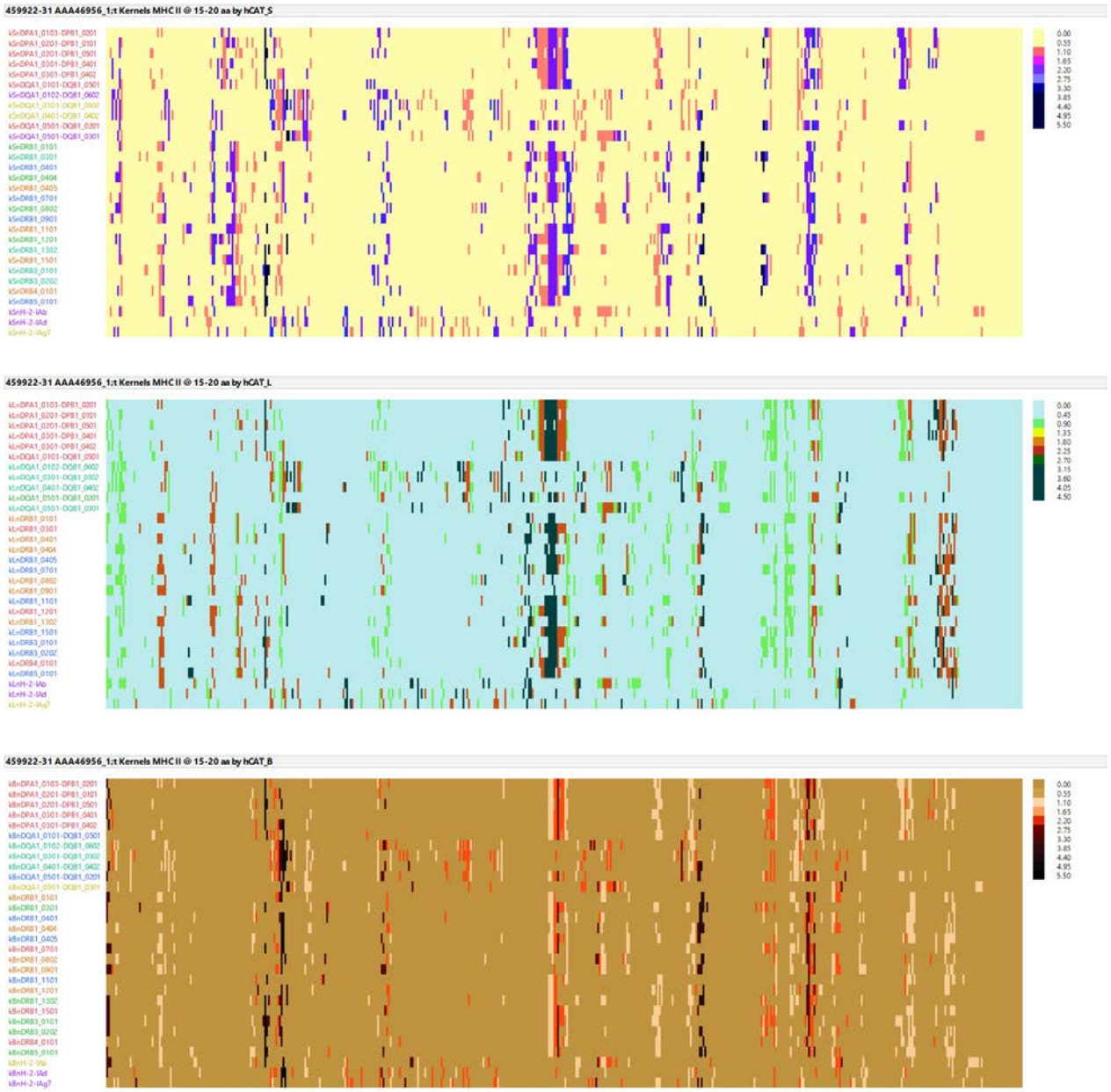

60983-18 CAA28671_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

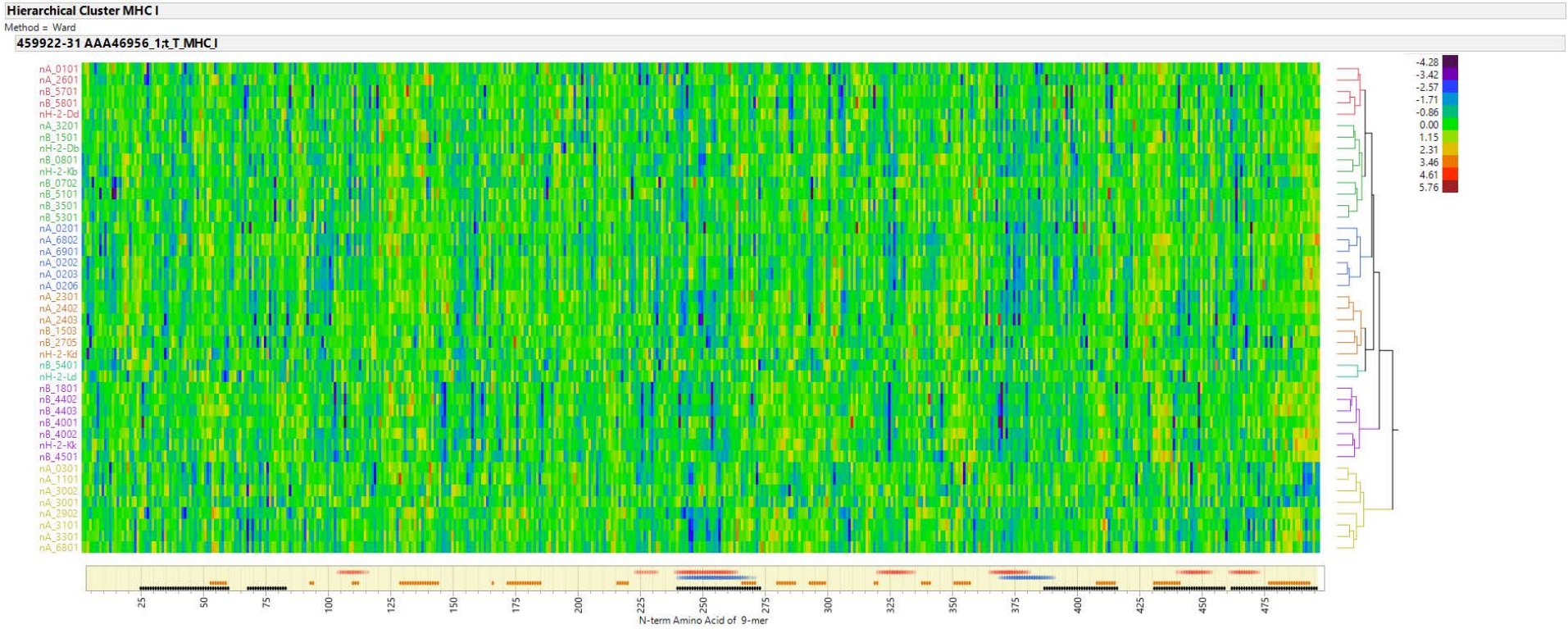

60983-18 CAA28671_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

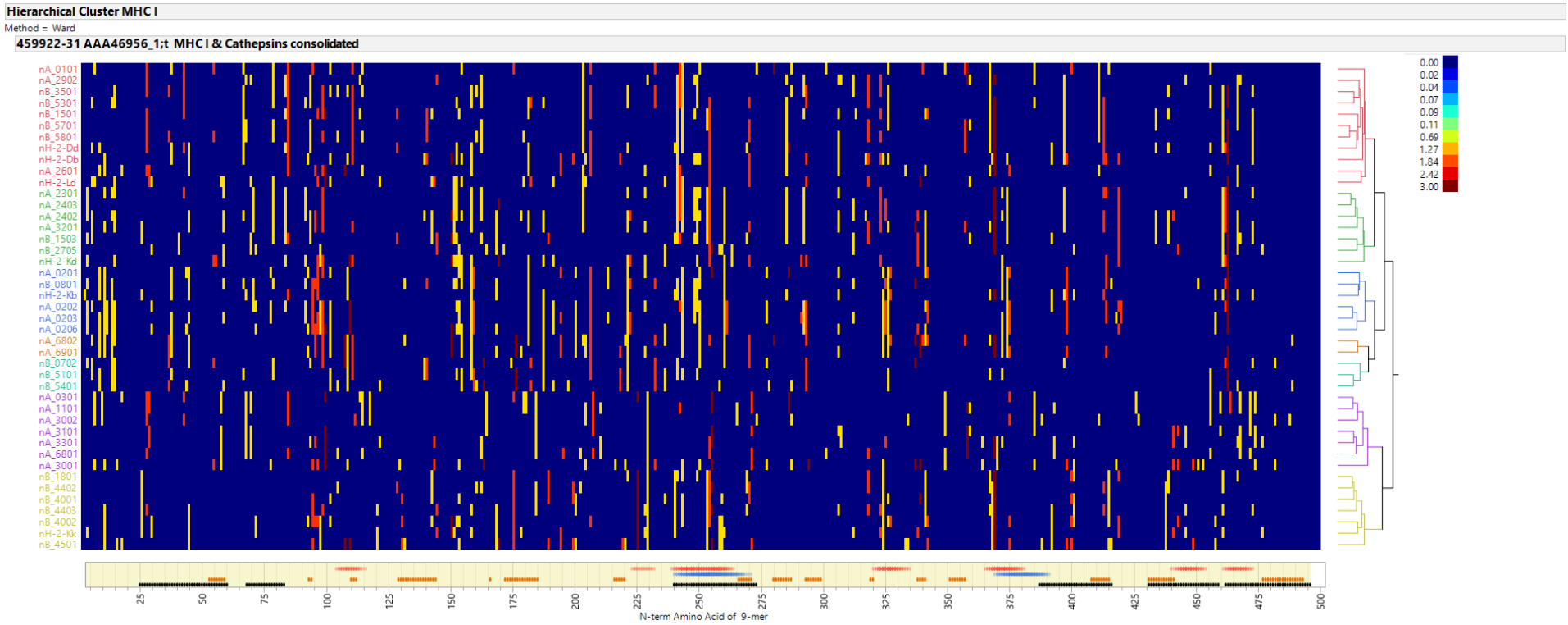
HPV 31

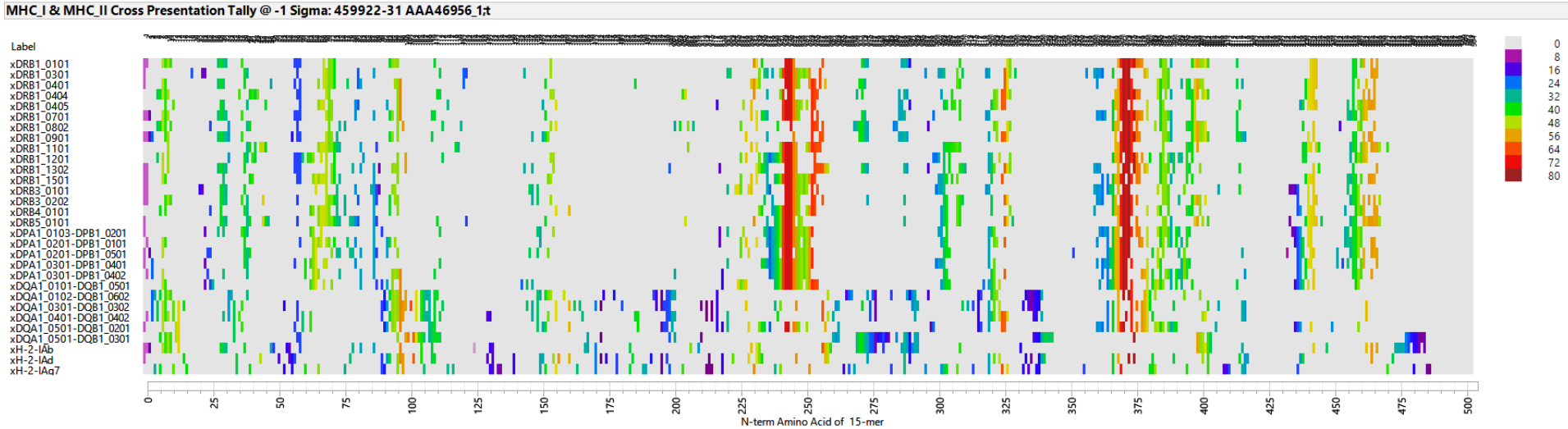

Overlay Plot
459922-31 AAA46956_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
459922-31 AAA46956_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0404
nDRB1_0701
nDRB1_1302
nDRB1_0802
nDRB1_0401
nDRB3_0202
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDQA1_0501-DQB1_0201
nDRB1_1201
nDRB3_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-4.05
-3.24
-2.43
-1.62
-0.81
0.00
0.73
1.46
2.19
2.92
3.65
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

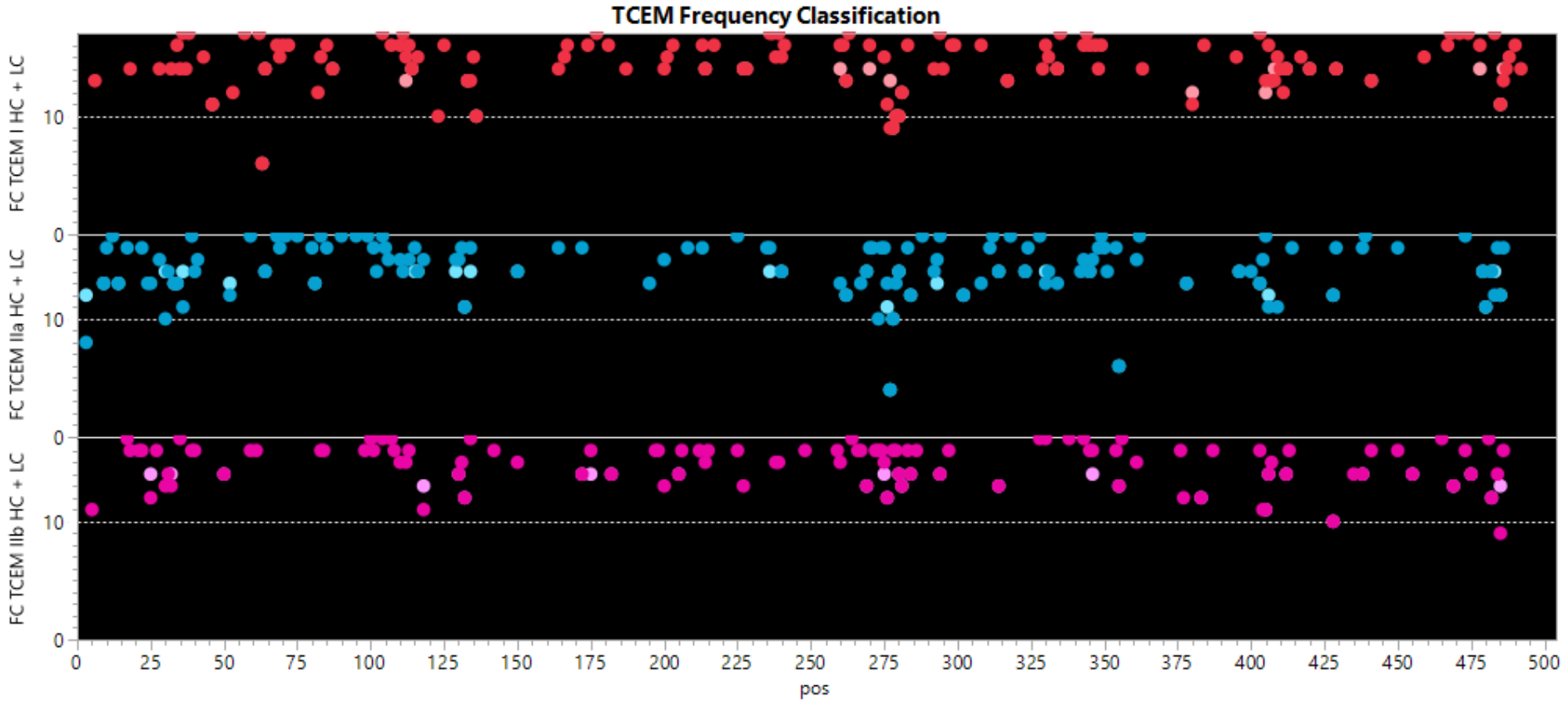

459922-31 AAA46956_1;t Kernels MHC II @ 15-20 aa by hCAT_S
459922-31 AAA46956_1;t Kernels MHC II @ 15-20 aa by hCAT_L
459922-31 AAA46956_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
459922-31 AAA46956_1;t_T_MHC_I
nA_0101
nA_2601
nB_5701
nB_5801
nH-2-Dd
nA_3201
nB_1501
nH-2-Db
nB_0801
nH-2-Kb
nB_0702
nB_5101
nB_3501
nB_5301
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nB_5401
nH-2-Ld
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_3001
nA_2902
nA_3101
nA_3301
nA_6801
-4.28
-3.42
-2.57
-1.71
-0.86
0.00
1.15
2.31
3.46
4.61
5.76
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

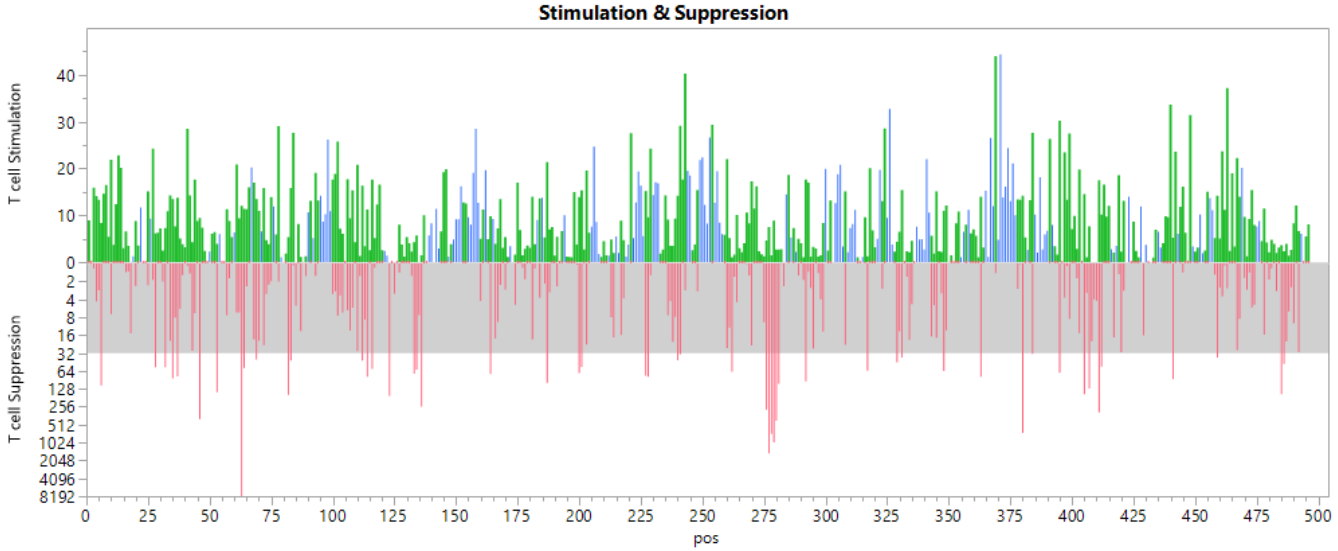

Hierarchical Cluster MHC I
Method = Ward
459922-31 AAA46956_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2902
nB_3501
nB_5301
nB_1501
nB_5701
nB_5801
nH-2-Dd
nH-2-Db
nA_2601
nH-2-Ld
nA_2301
nA_2403
nA_2402
nA_3201
nB_1503
nB_2705
nH-2-Kd
nA_0201
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_5101
nB_5401
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nB_1801
nB_4402
nB_4001
nB_4403
nB_4002
nH-2-Kk
nB_4501
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 459922-31 AAA46956_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
8
16
24
32
40
48
56
64
72
80
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

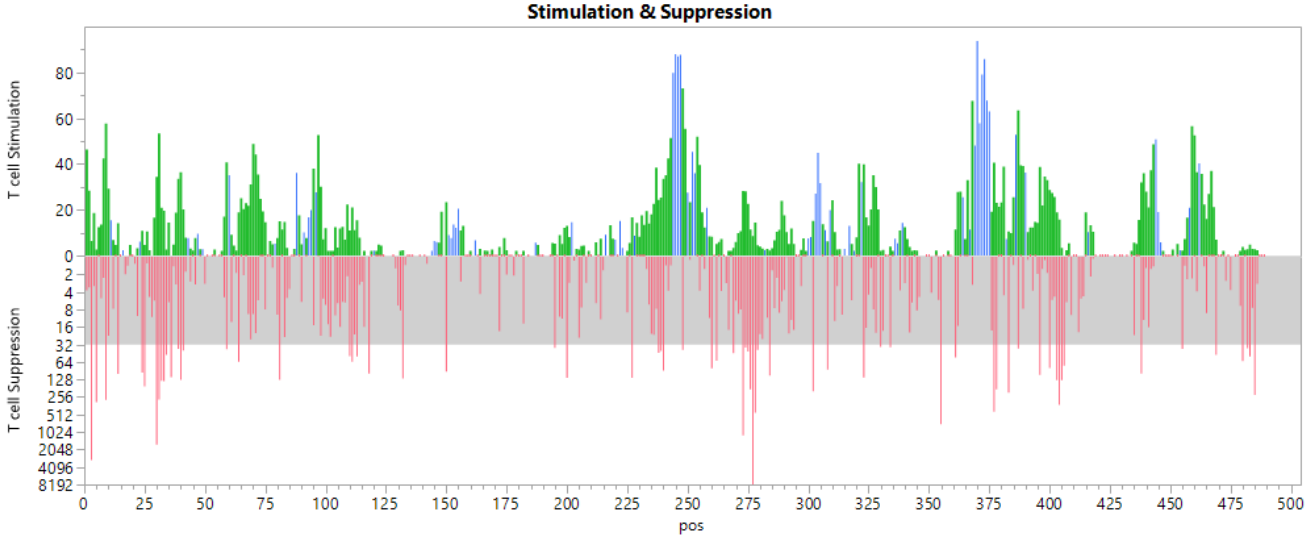
459922-31 AAA46956_1



459922-31 AAA46956_1



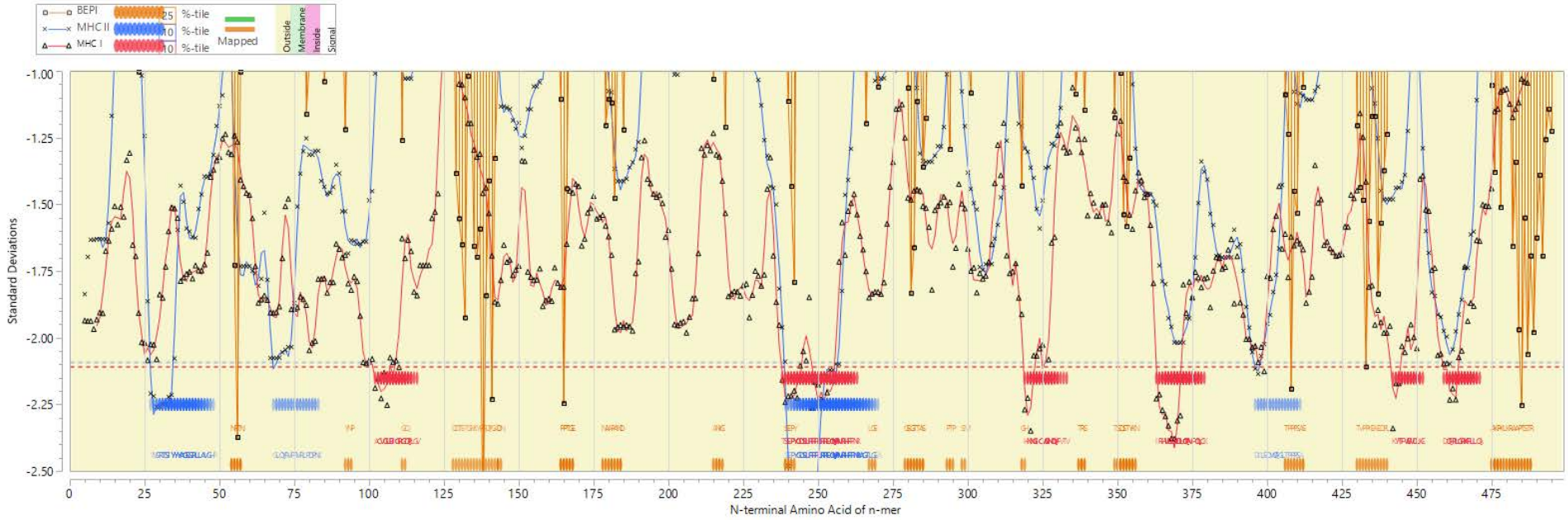
459922-31 AAA46956_1

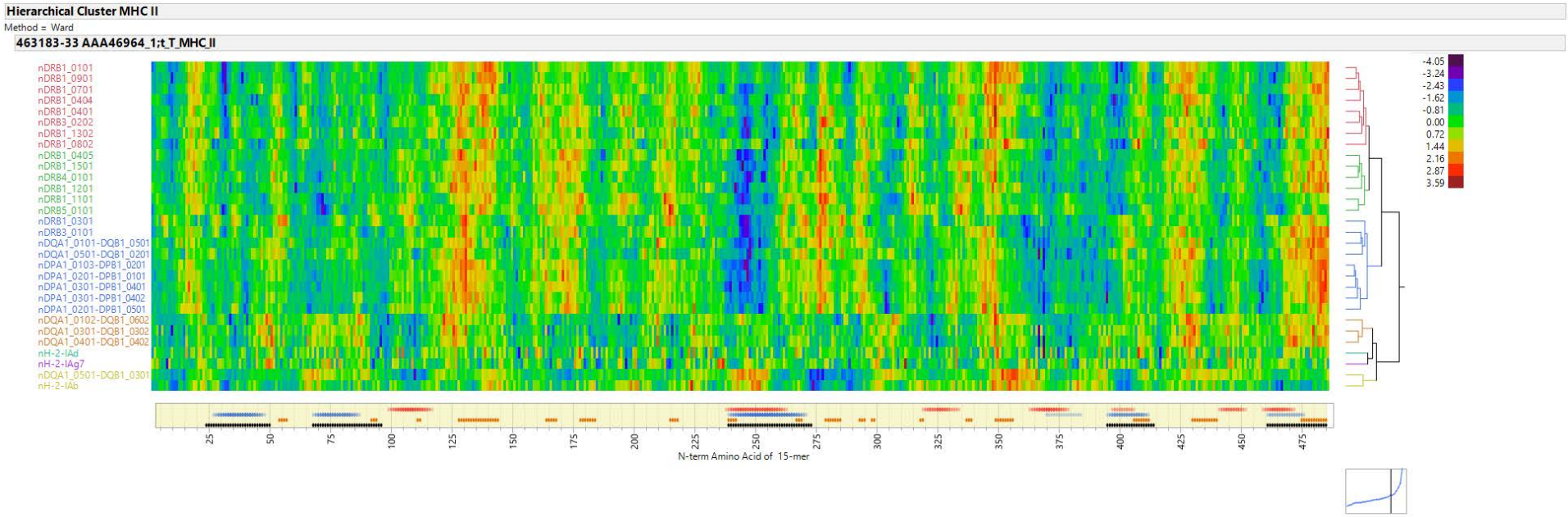


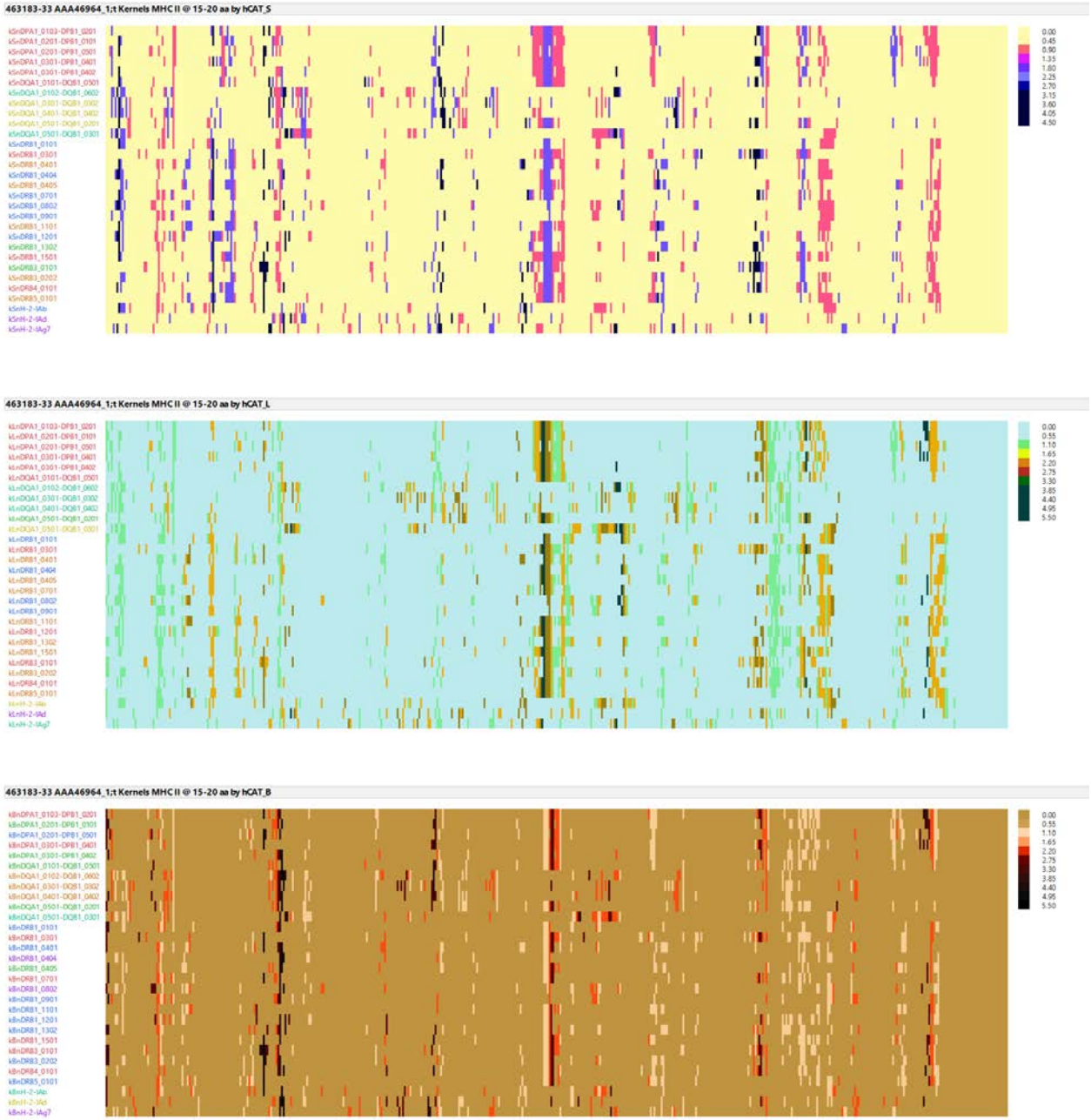
HPV 33

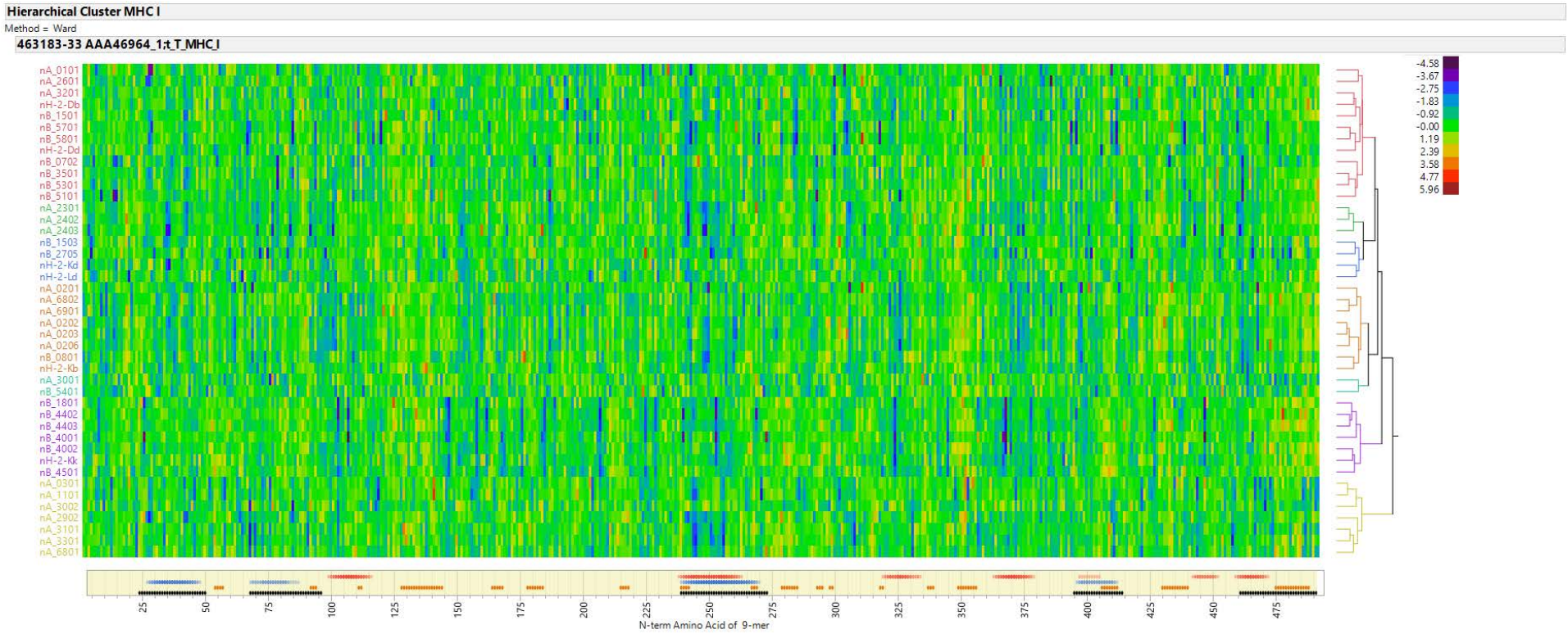
Overlay Plot

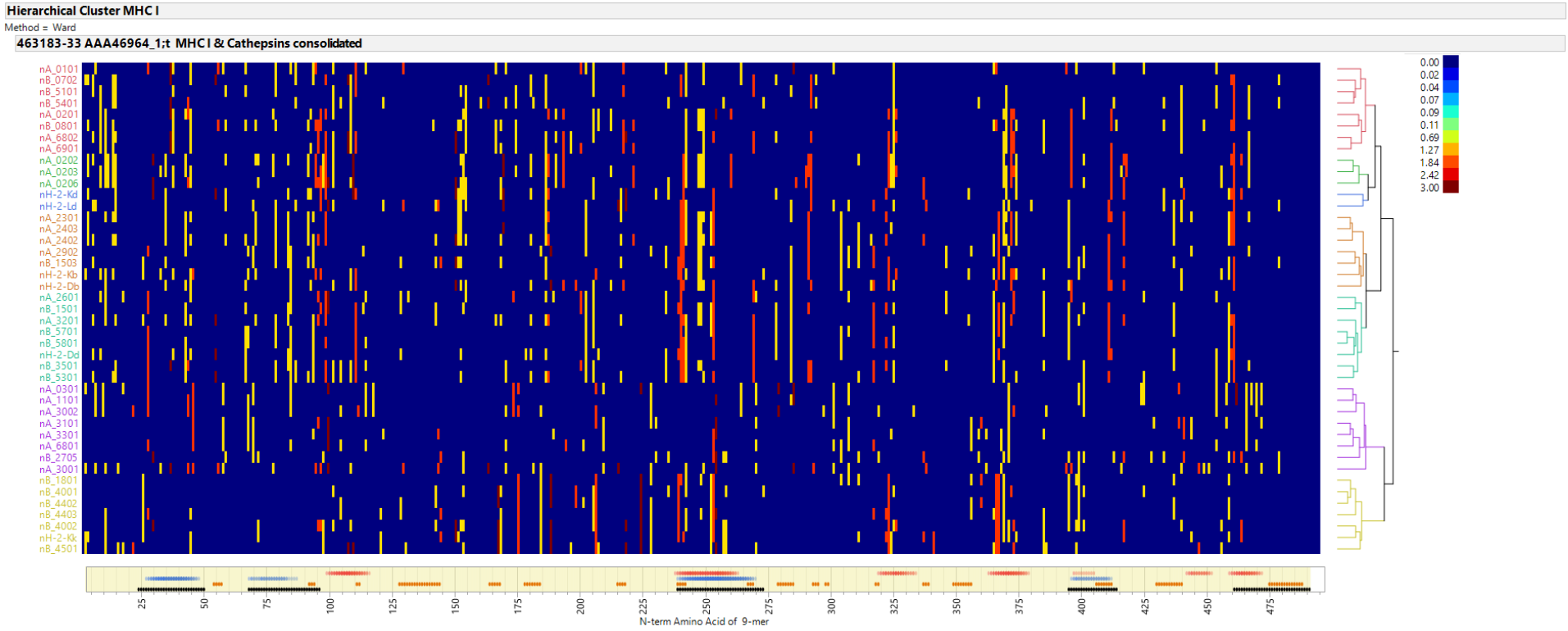
463183-33 AAA46964_1;t

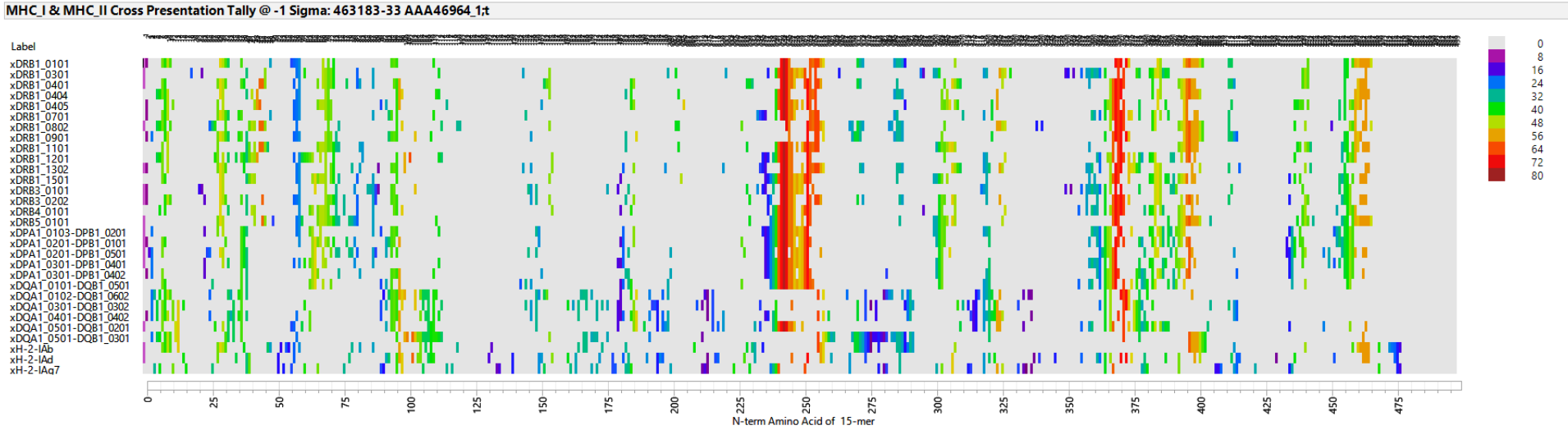



Hierarchical Cluster MHC II
Method = Ward
463183-33 AAA46964_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_0401
nDRB3_0202
nDRB1_1302
nDRB1_0802
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-4.05
-3.24
-2.43
-1.62
-0.81
0.00
0.72
1.44
2.16
2.87
3.59
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

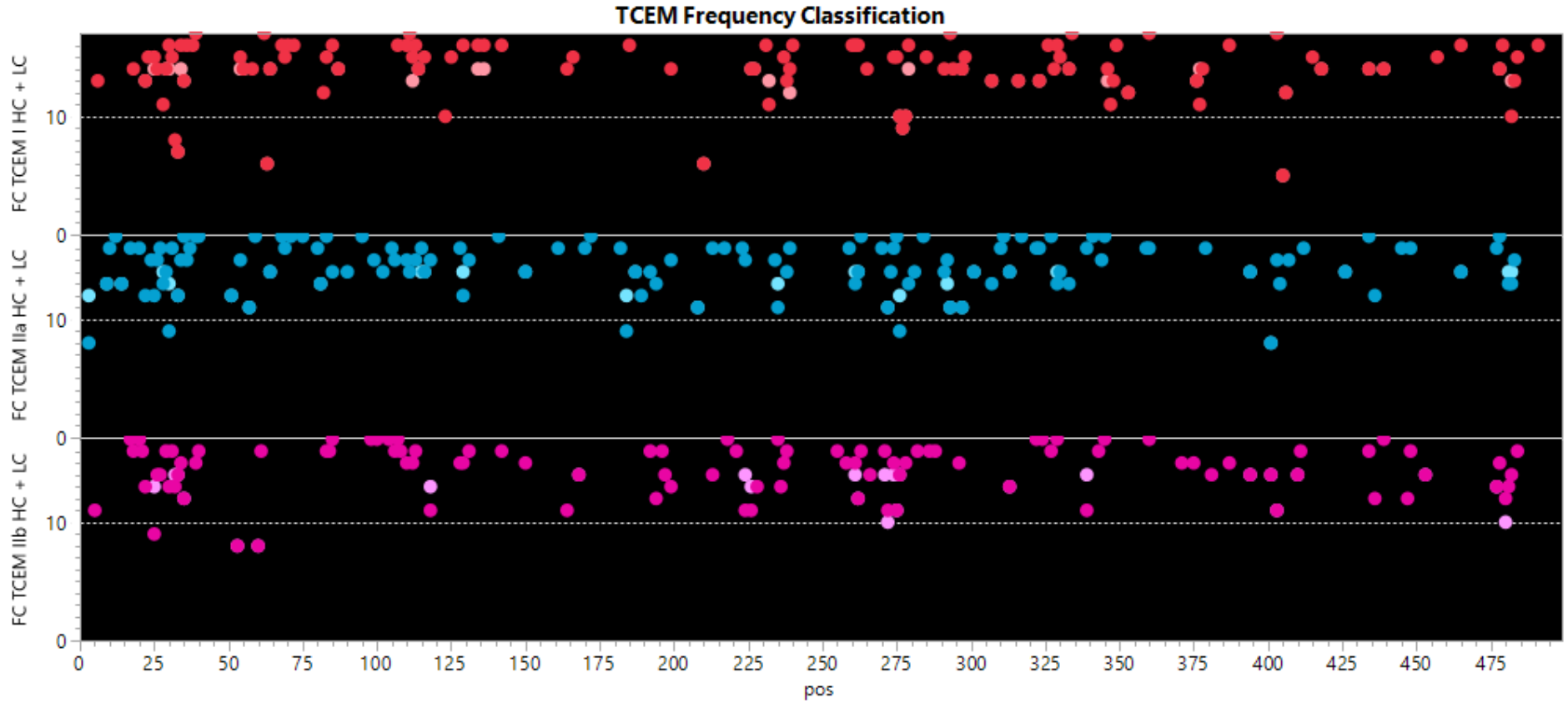

463183-33 AAA46964_1;t Kernels MHC II @ 15-20 aa by hCAT_S
463183-33 AAA46964_1;t Kernels MHC II @ 15-20 aa by hCAT_L
463183-33 AAA46964_1;t Kernels MHC II @ 15-20 aa by hCAT_B

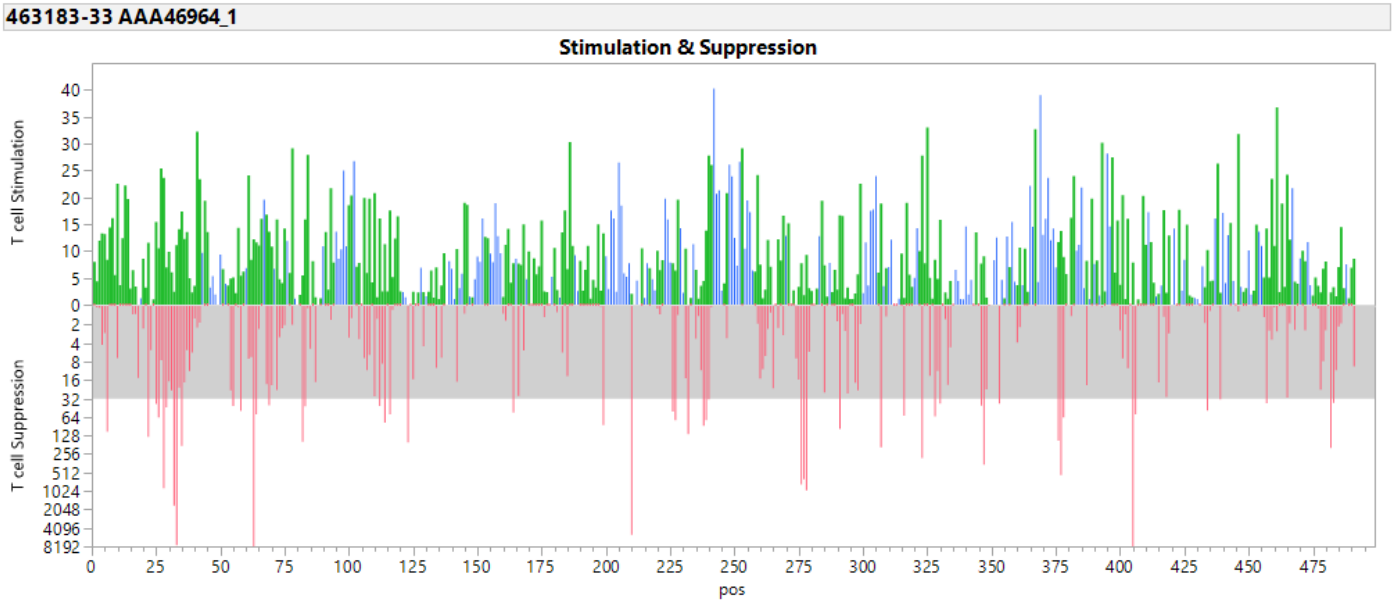

Hierarchical Cluster MHC I
Method = Ward
463183-33 AAA46964_1;t_T_MHC_I
nA_0101
nA_2601
nA_3201
nH-2-Db
nB_1501
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_3501
nB_5301
nB_5101
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_3001
nB_5401
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.58
-3.67
-2.75
-1.83
-0.92
-0.00
1.19
2.39
3.58
4.77
5.96
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

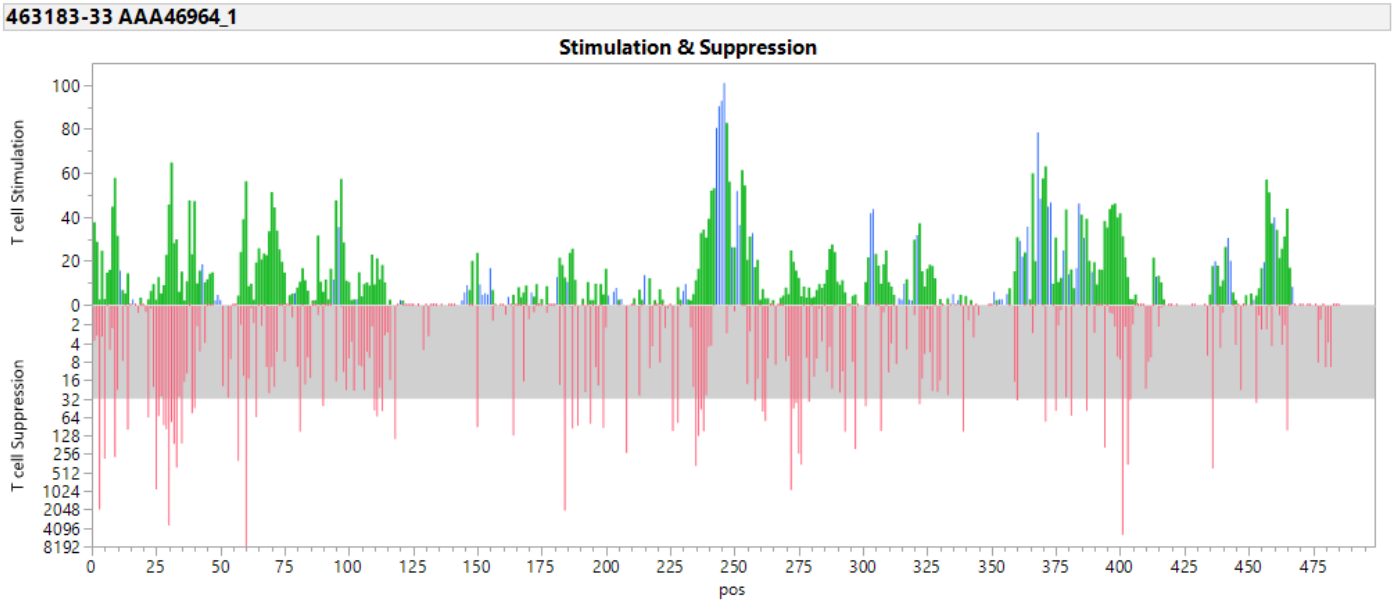

Hierarchical Cluster MHC I
Method = Ward
463183-33 AAA46964_1;t MHC I & Cathepsins consolidated
nA_0101
nB_0702
nB_5101
nB_5401
nA_0201
nB_0801
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nH-2-Kd
nH-2-Ld
nA_2301
nA_2403
nA_2402
nA_2902
nB_1503
nH-2-Kb
nH-2-Db
nA_2601
nB_1501
nA_3201
nB_5701
nB_5801
nH-2-Dd
nB_3501
nB_5301
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nB_2705
nA_3001
nB_1801
nB_4001
nB_4402
nB_4403
nB_4002
nH-2-Kk
nB_4501
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 463183-33 AAA46964_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

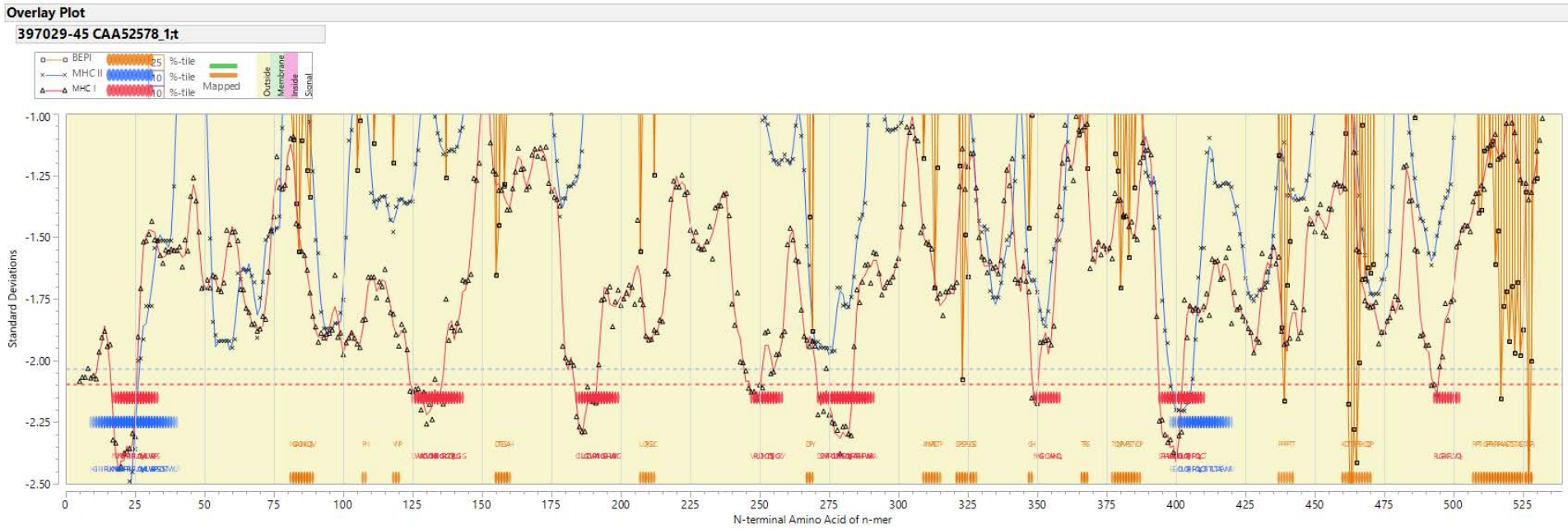
463183-33 AAA46964_1

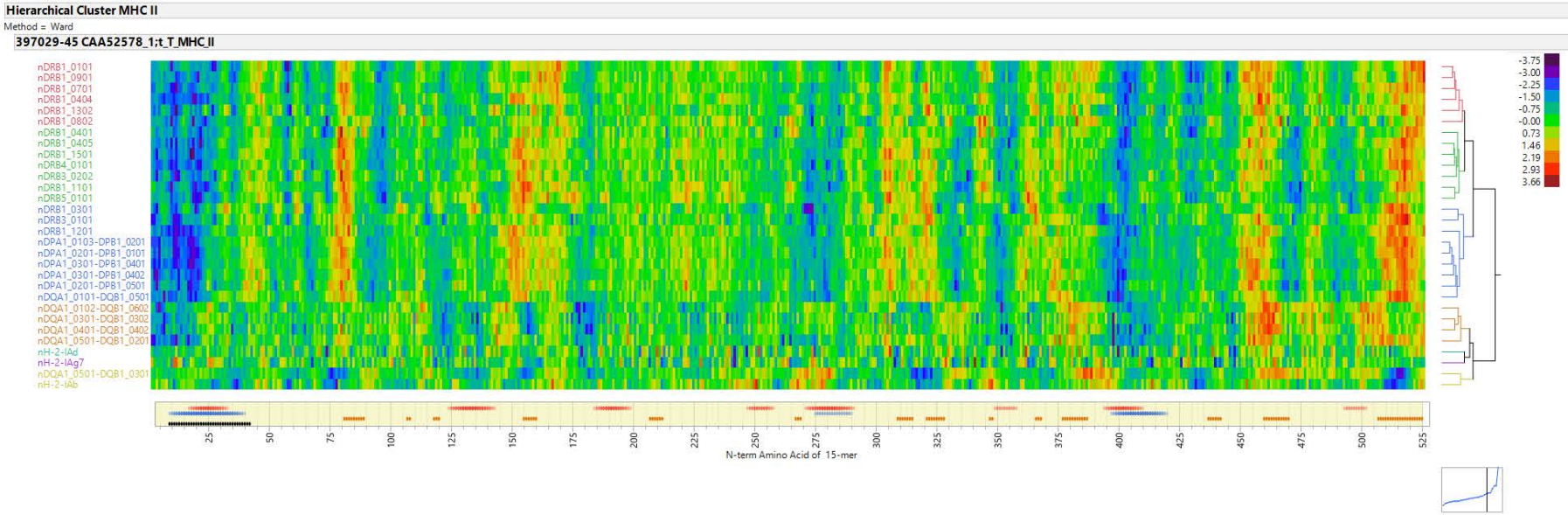


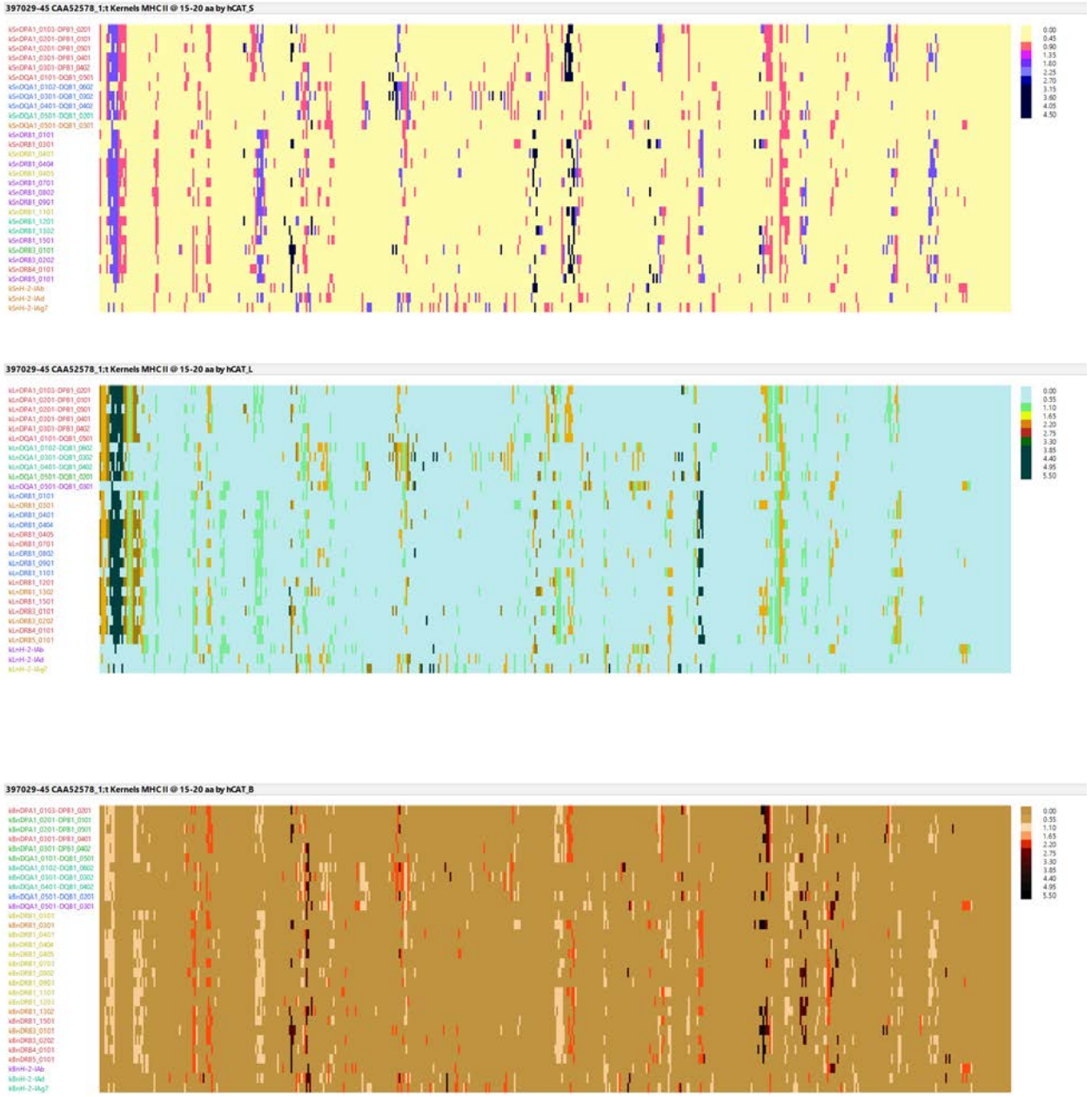

463183-33 AAA46964_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

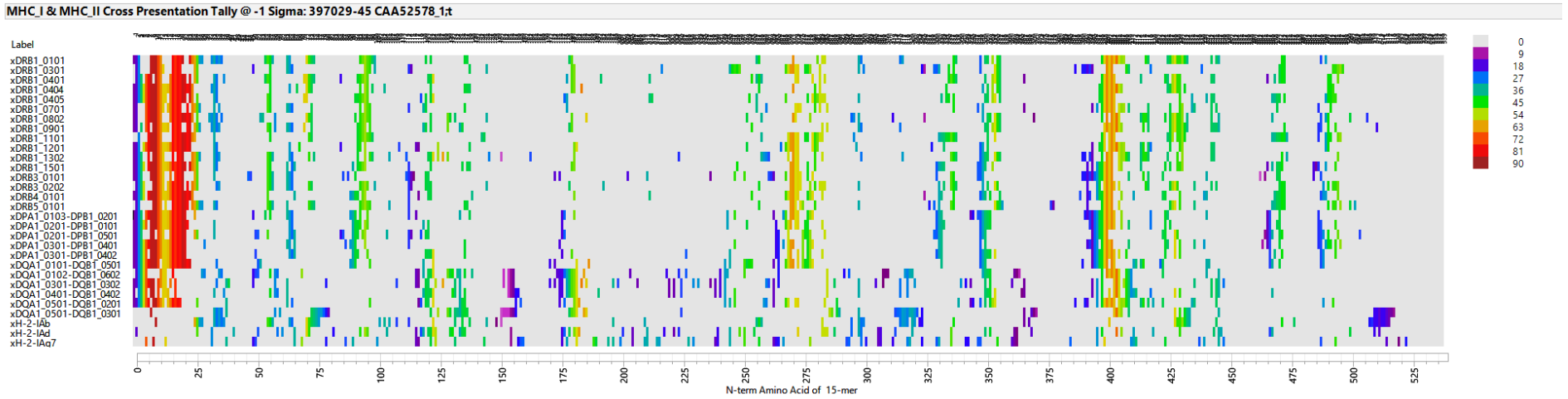

463183-33 AAA46964_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

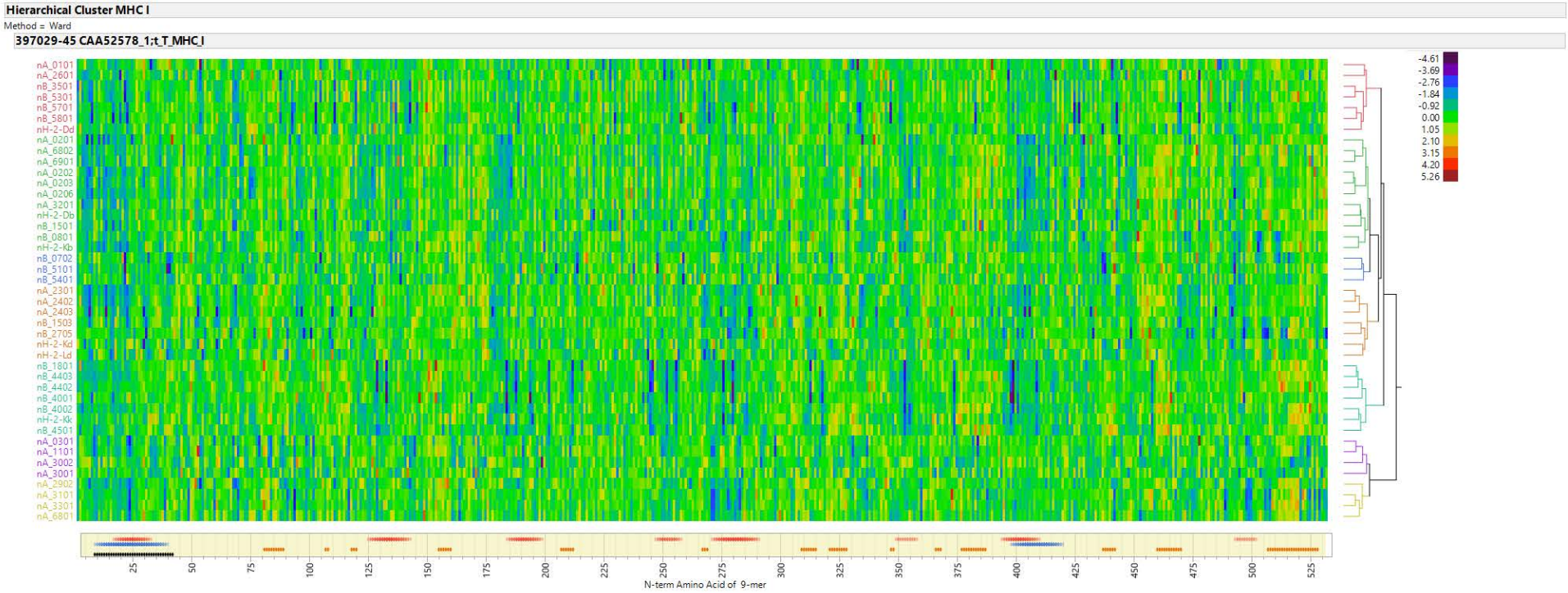
HPV 45

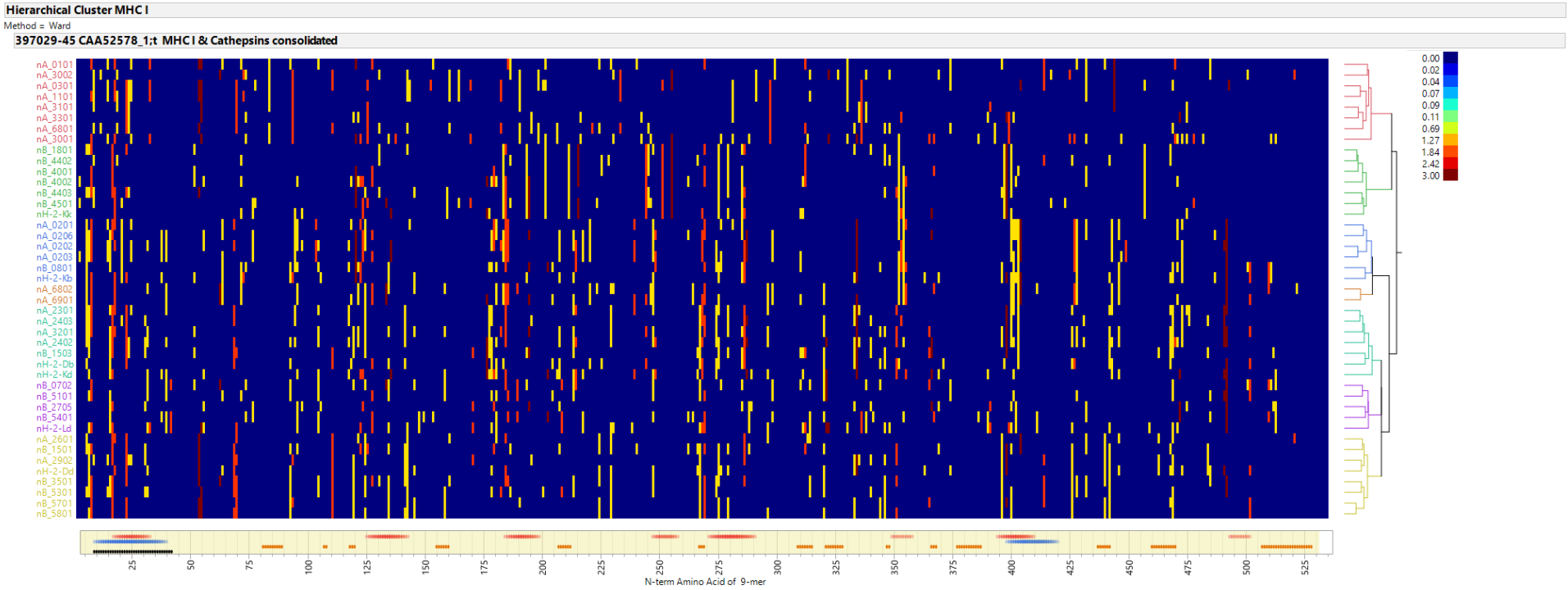

Overlay Plot
397029-45 CAA52578_1:t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
397029-45 CAA52578_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1302
nDRB1_0802
nDRB1_0401
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.75
-3.00
-2.25
-1.50
-0.75
-0.00
0.73
1.46
2.19
2.93
3.66
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

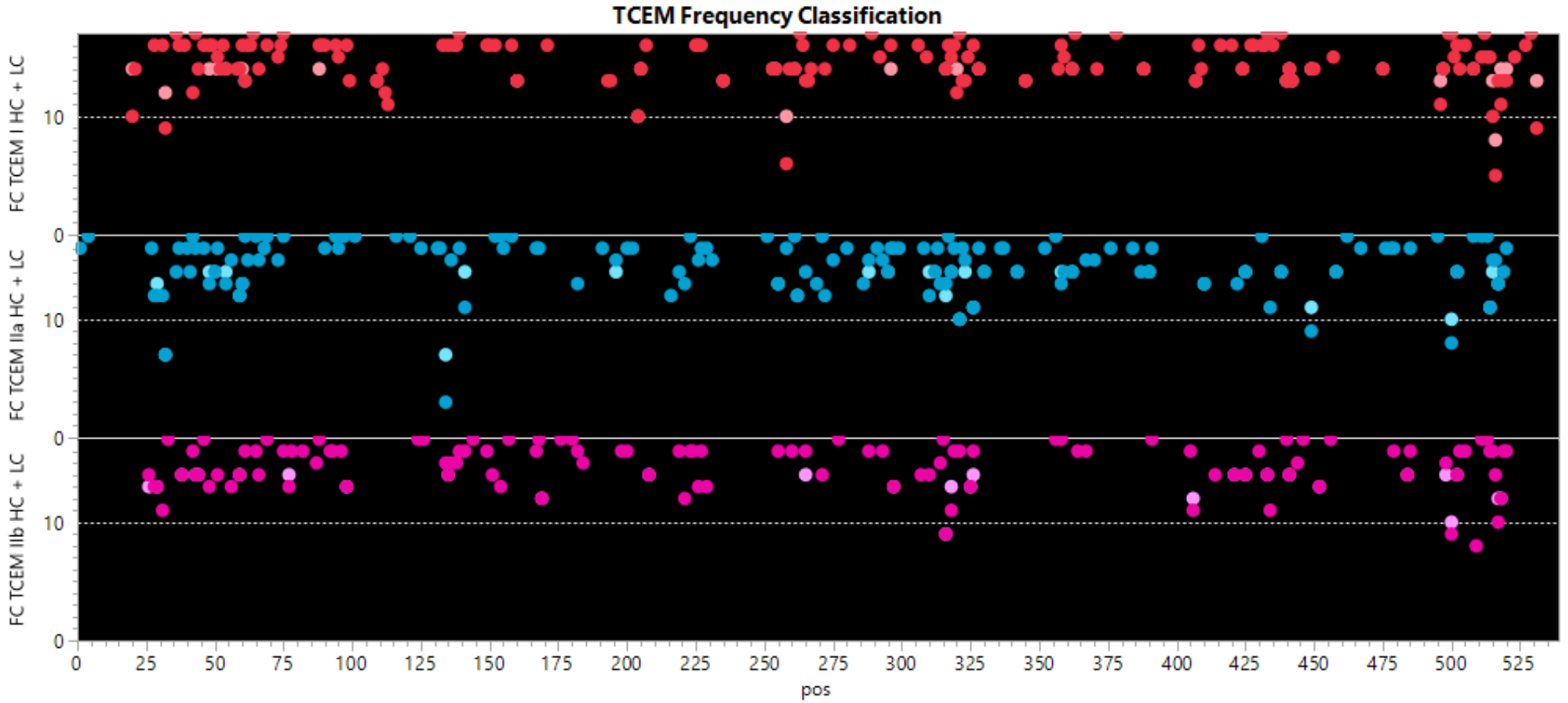

397029-45 CAA52578_1;t Kernels MHC II @ 15-20 aa by hCAT_S
397029-45 CAA52578_1;t Kernels MHC II @ 15-20 aa by hCAT_L
397029-45 CAA52578_1;t Kernels MHC II @ 15-20 aa by hCAT_B

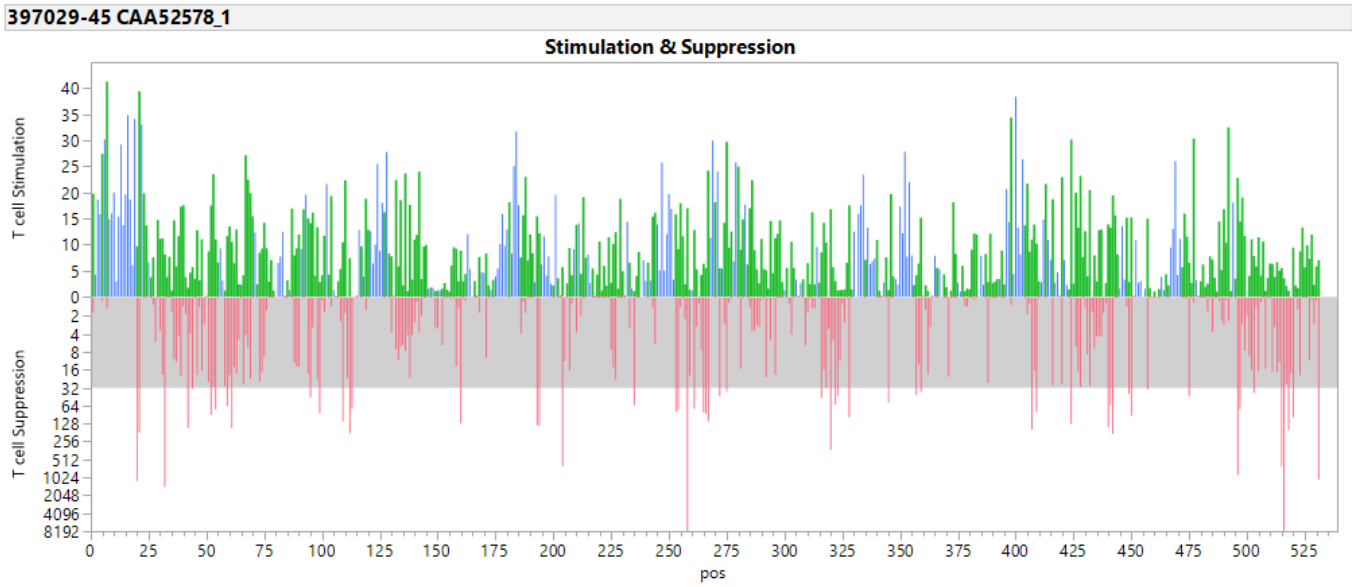

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 397029-45 CAA52578_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer
0
9
18
27
36
45
54
63
72
81
90

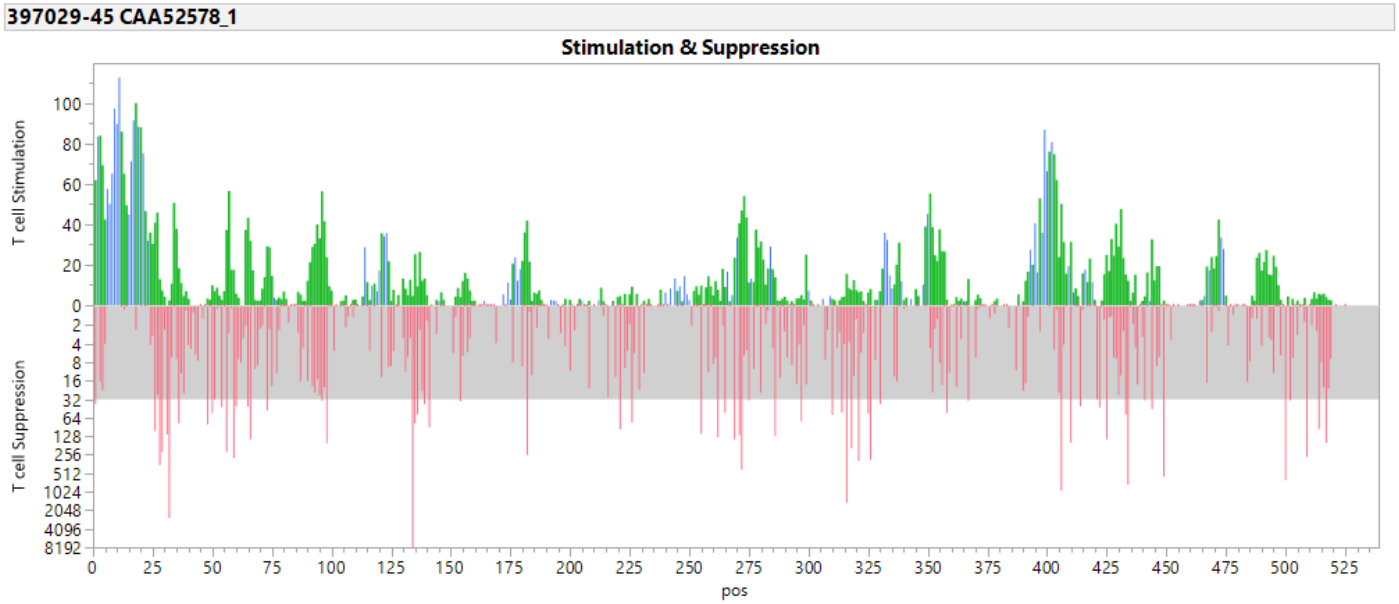

Hierarchical Cluster MHC I
Method = Ward
397029-45 CAA52578_1;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_3201
nH-2-Db
nB_1501
nB_0801
nH-2-Kb
nB_0702
nB_5101
nB_5401
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_3001
nA_2902
nA_3101
nA_3301
nA_6801
-4.61
-3.69
-2.76
-1.84
-0.92
0.00
1.05
2.10
3.15
4.20
5.26
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
397029-45 CAA52578_1;t MHC I & Cathepsins consolidated
nA_0101
nA_3002
nA_0301
nA_1101
nA_3101
nA_3301
nA_6801
nA_3001
nB_1801
nB_4402
nB_4001
nB_4002
nB_4403
nB_4501
nH-2-Kk
nA_0201
nA_0206
nA_0202
nA_0203
nB_0801
nH-2-Kb
nA_6802
nA_6901
nA_2301
nA_2403
nA_3201
nA_2402
nB_1503
nH-2-Db
nH-2-Kd
nB_0702
nB_5101
nB_2705
nB_5401
nH-2-Ld
nA_2601
nB_1501
nA_2902
nH-2-Dd
nB_3501
nB_5301
nB_5701
nB_5801
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer

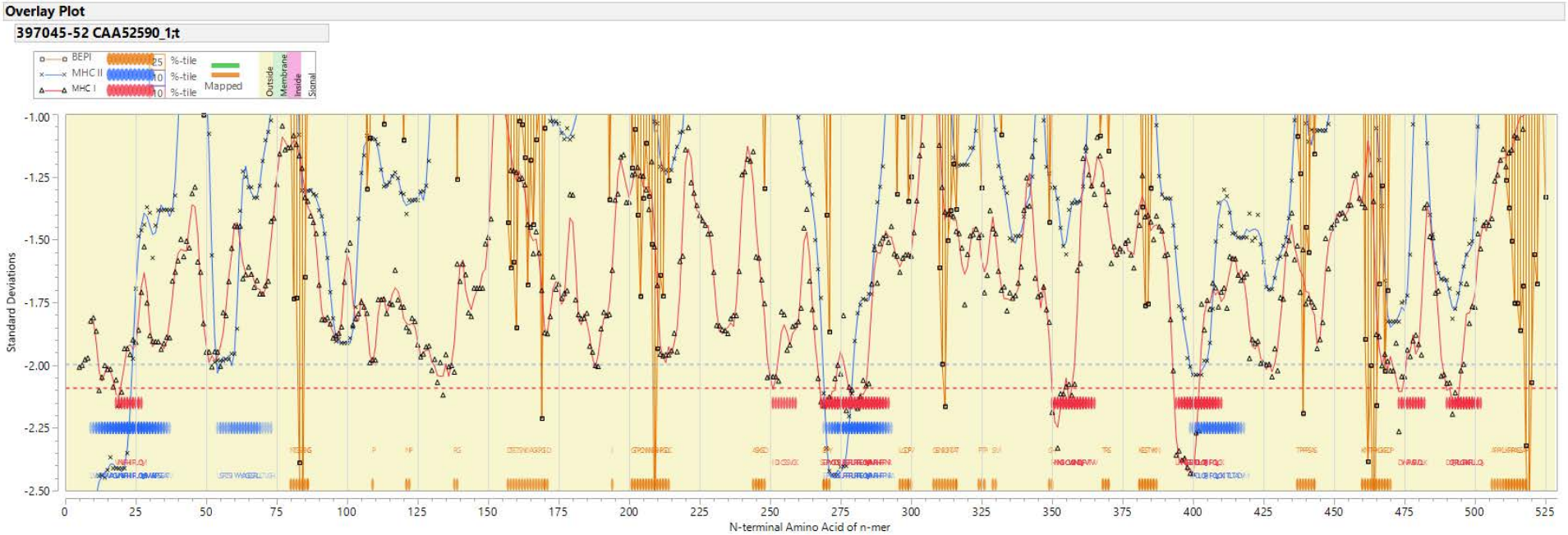
397029-45 CAA52578_1

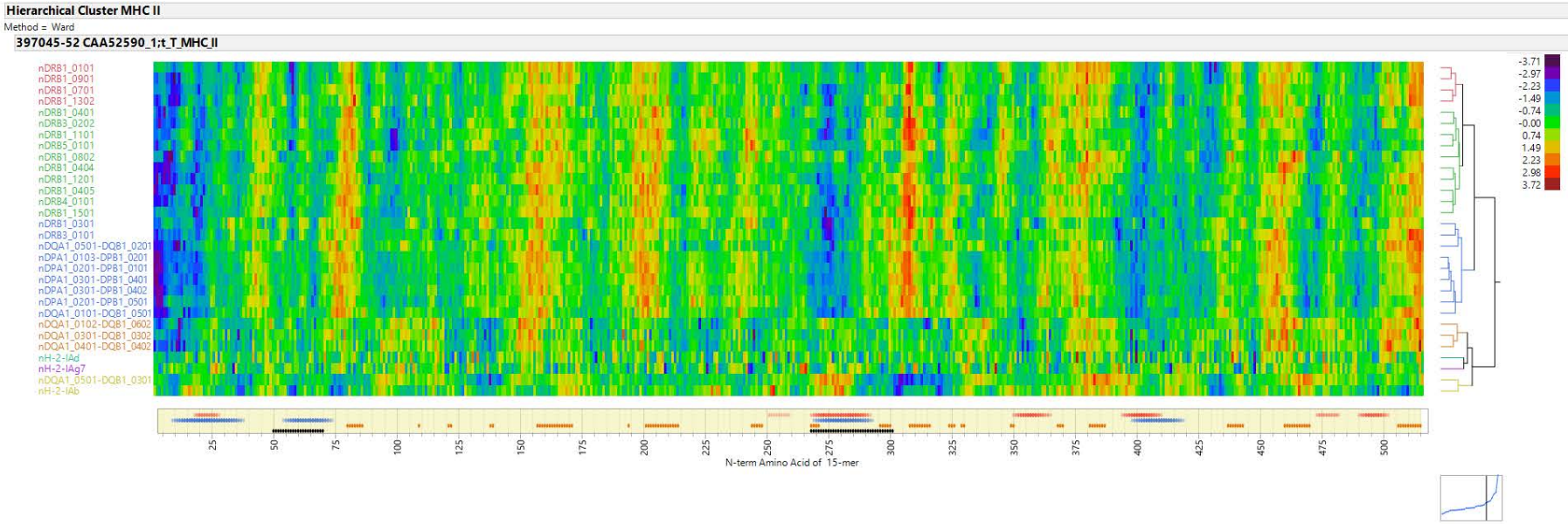


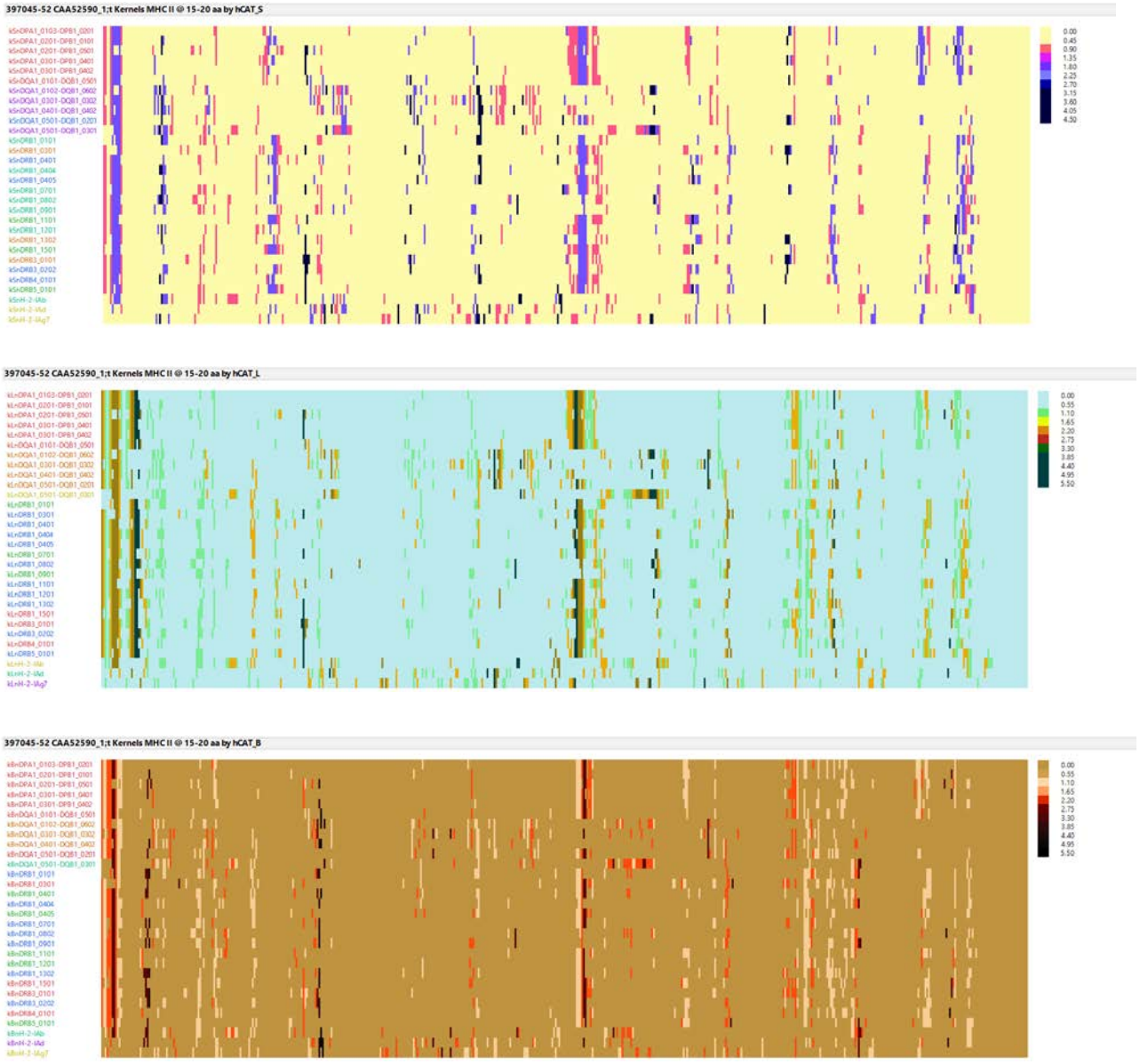

397029-45 CAA52578_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

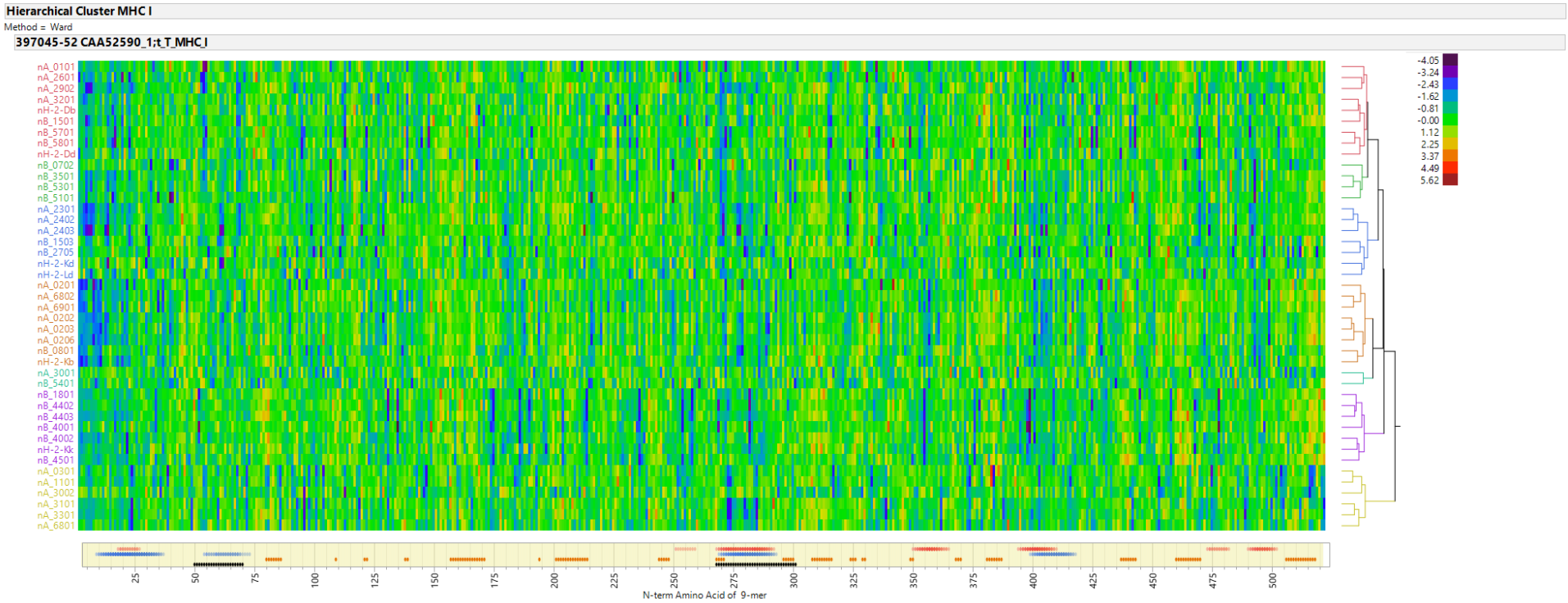

397029-45 CAA52578_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

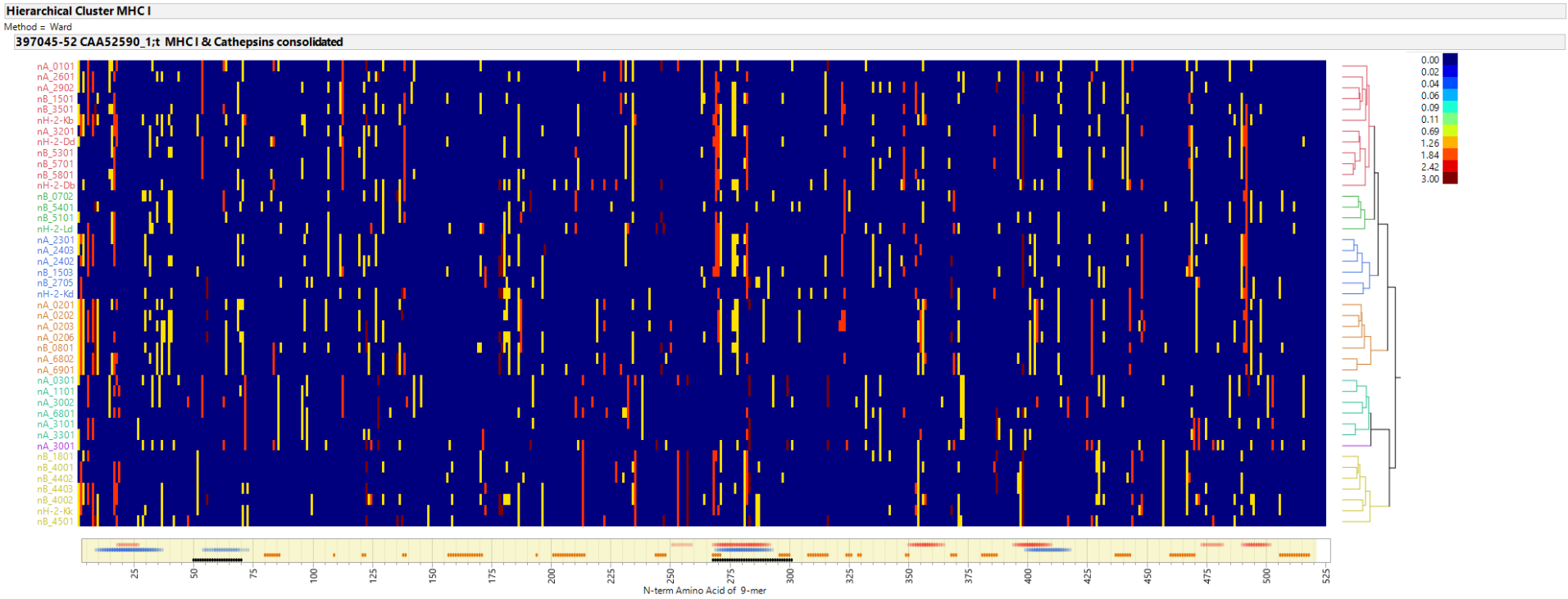
HPV 52

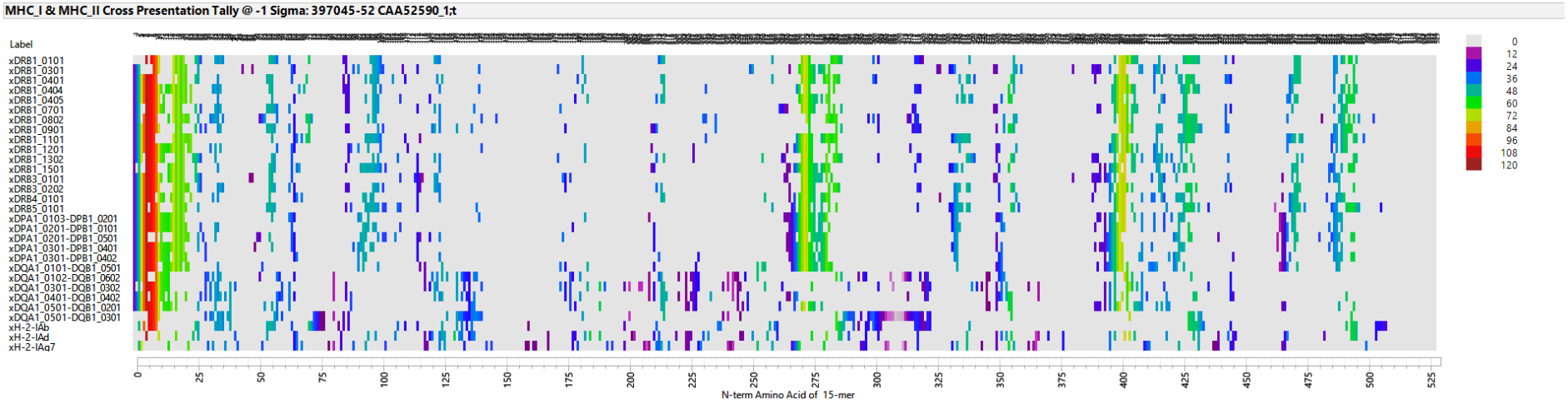

Overlay Plot
397045-52 CAA52590_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
397045-52 CAA52590_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_0404
nDRB1_1201
nDRB1_0405
nDRB4_0101
nDRB1_1501
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.71
-2.97
-2.23
-1.49
-0.74
-0.00
0.74
1.49
2.23
2.98
3.72
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

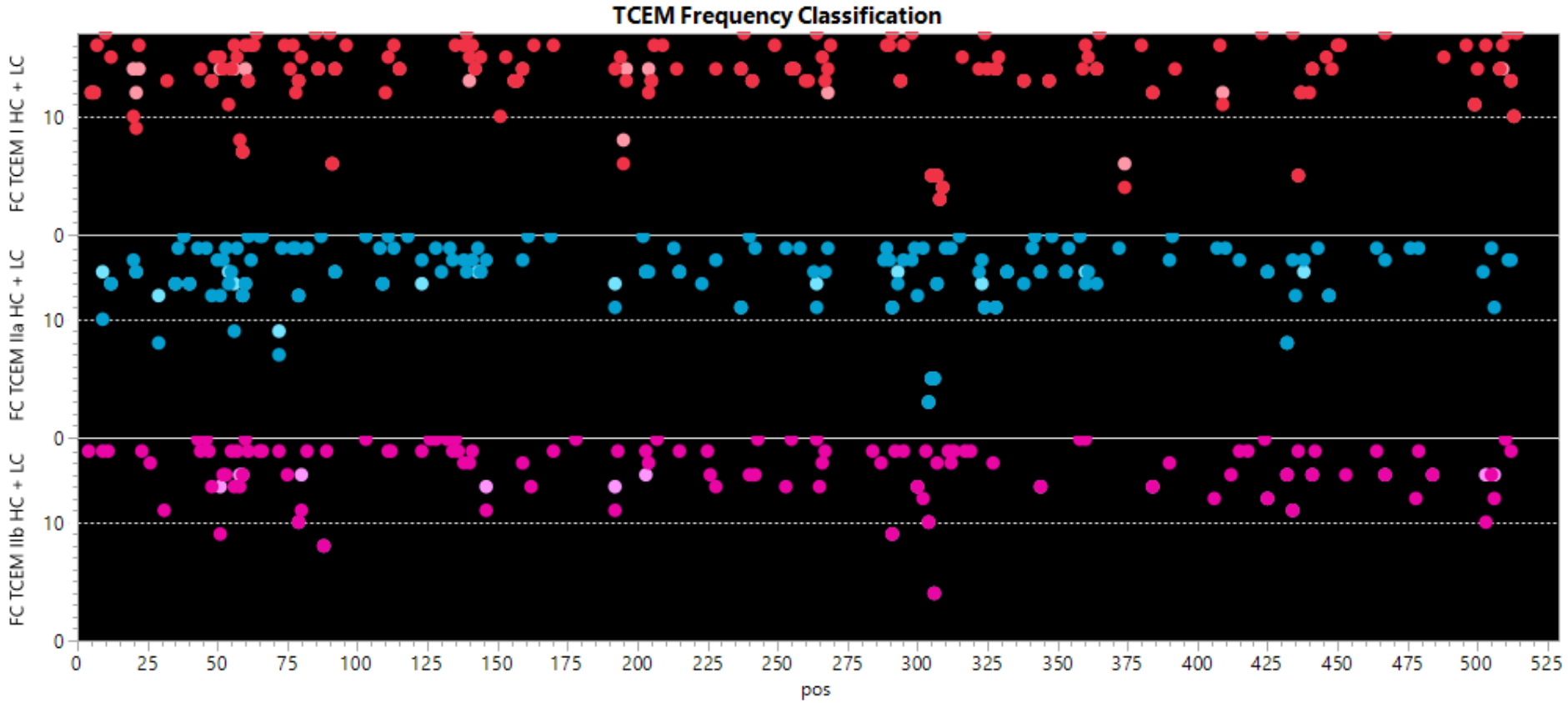

397045-52 CAA52590_1;t Kernels MHC II @ 15-20 aa by hCAT_S
397045-52 CAA52590_1;t Kernels MHC II @ 15-20 aa by hCAT_L
397045-52 CAA52590_1;t Kernels MHC II @ 15-20 aa by hCAT_B

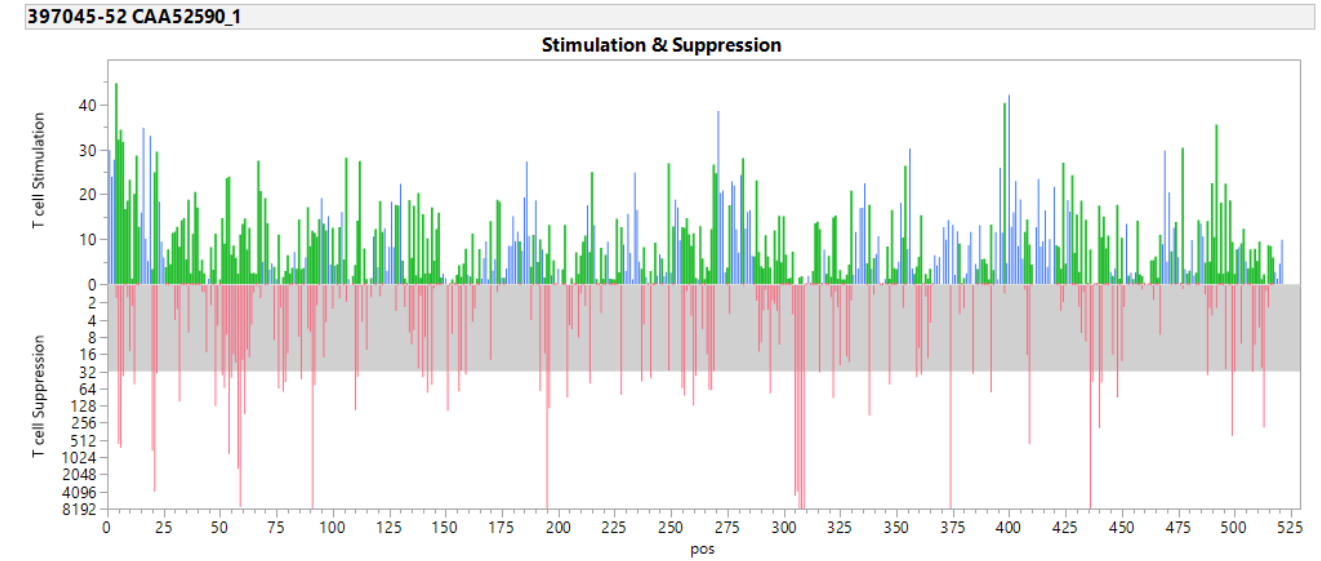

Hierarchical Cluster MHC I
Method = Ward
397045-52 CAA52590_1;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nA_3201
nH-2-Db
nB_1501
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_3501
nB_5301
nB_5101
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_3001
nB_5401
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
-4.05
-3.24
-2.43
-1.62
-0.81
-0.00
1.12
2.25
3.37
4.49
5.62
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer

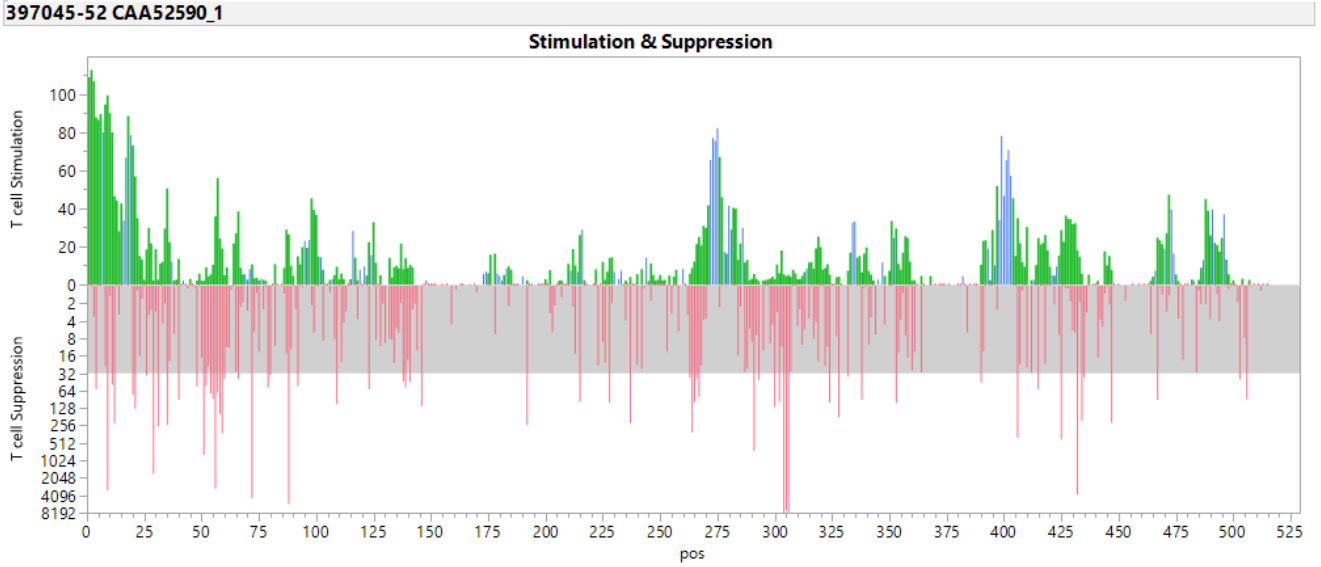

Hierarchical Cluster MHC I
Method = Ward
397045-52 CAA52590_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nA_2902
nB_1501
nB_3501
nH-2-Kb
nA_3201
nH-2-Dd
nB_5301
nB_5701
nB_5801
nH-2-Db
nB_0702
nB_5401
nB_5101
nH-2-Ld
nA_2301
nA_2403
nA_2402
nB_1503
nB_2705
nH-2-Kd
nA_0201
nA_0202
nA_0203
nA_0206
nB_0801
nA_6802
nA_6901
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_3001
nB_1801
nB_4001
nB_4402
nB_4403
nB_4002
nH-2-Kk
nB_4501
0.00
0.02
0.04
0.06
0.09
0.11
0.69
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 397045-52 CAA52590_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer
0
12
24
36
48
60
72
84
96
108
120

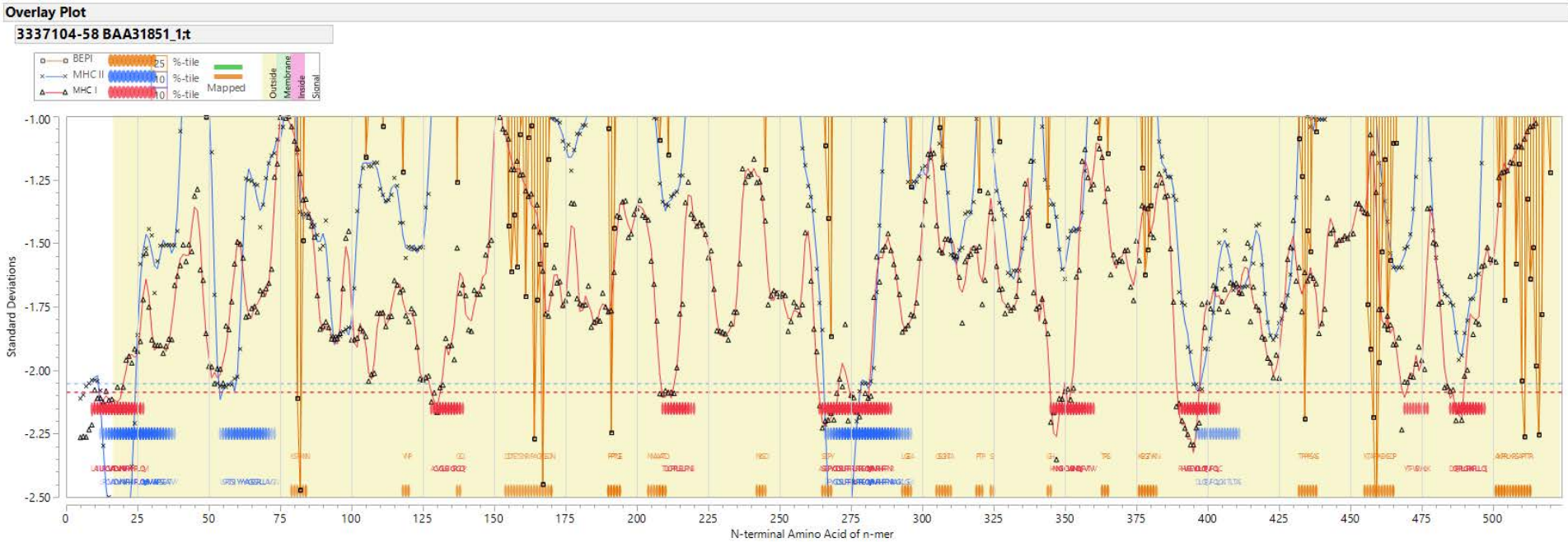
397045-52 CAA52590_1

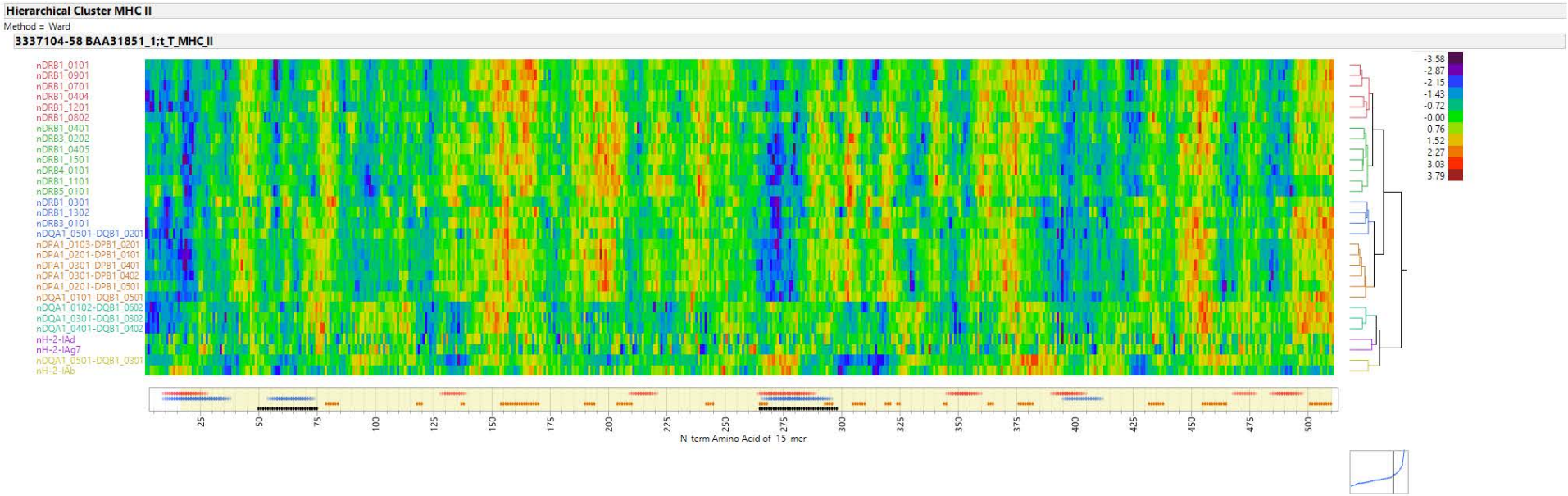


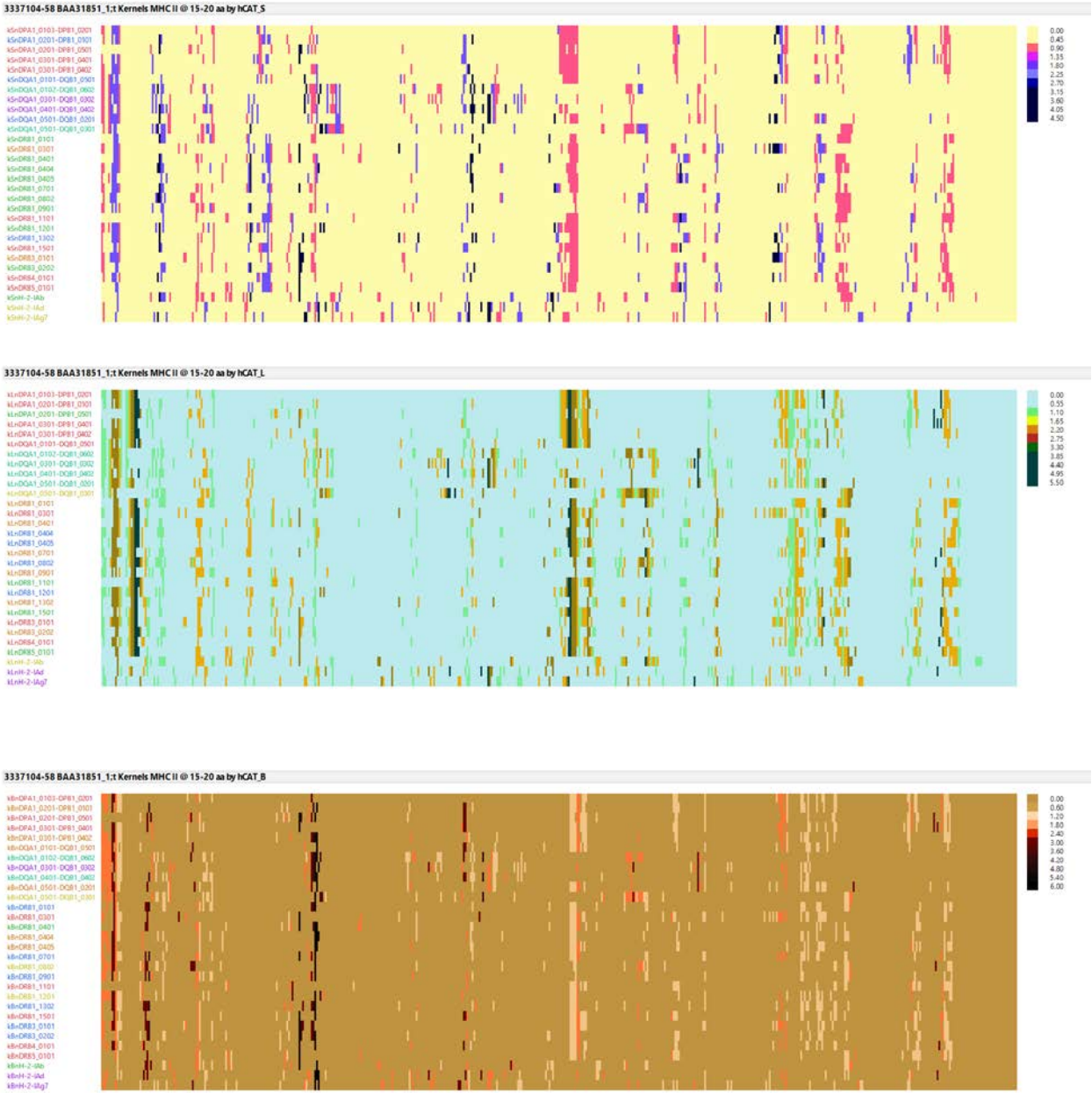

397045-52 CAA52590_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

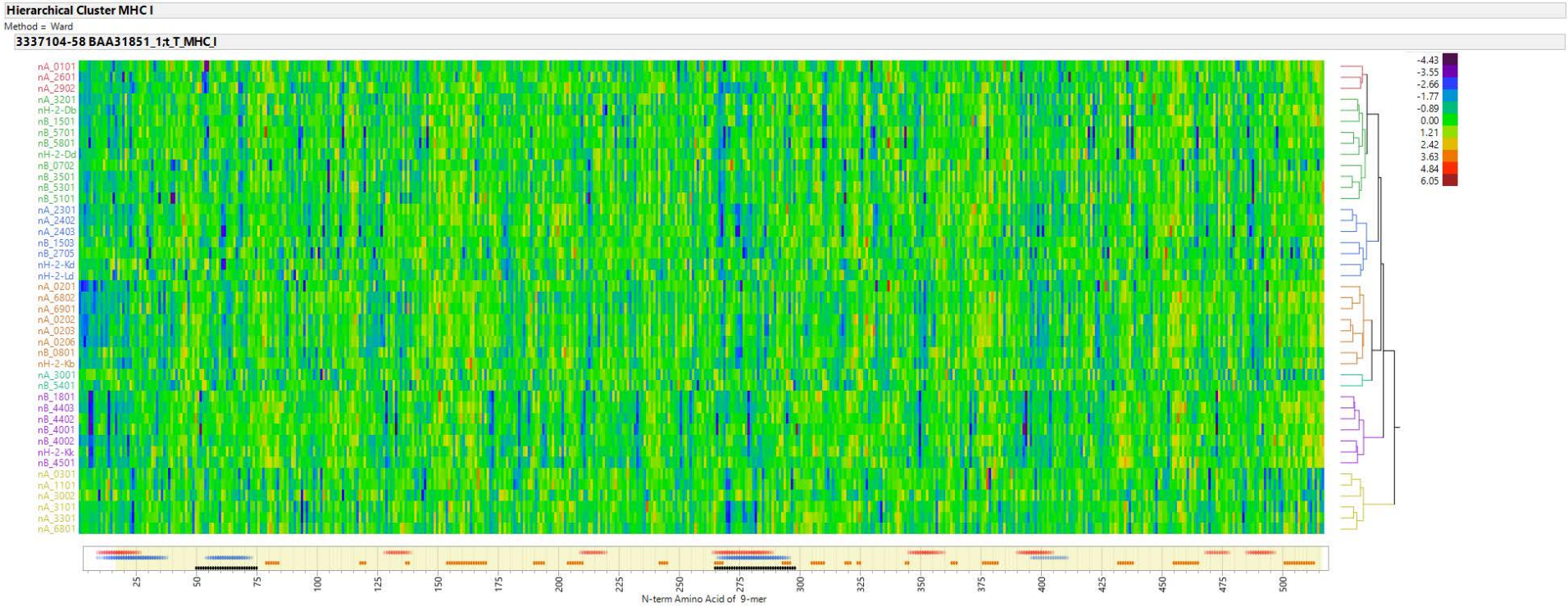

397045-52 CAA52590_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

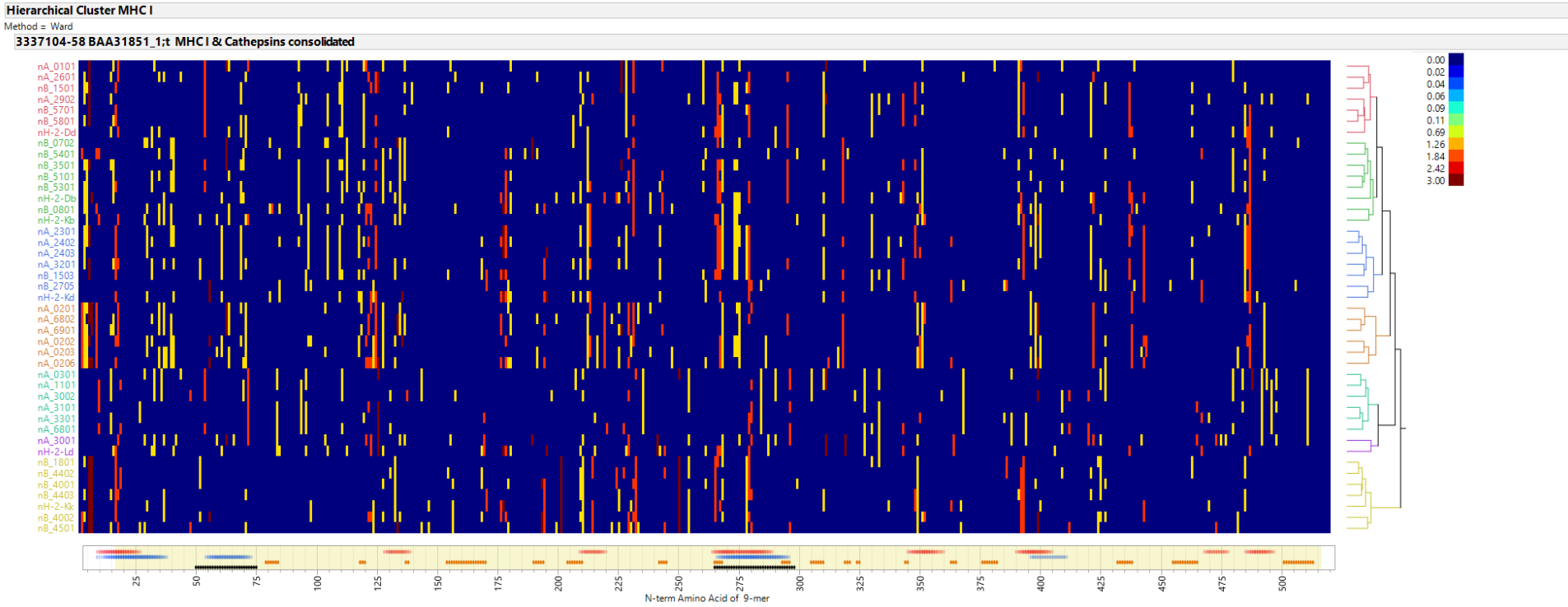
HPV 58

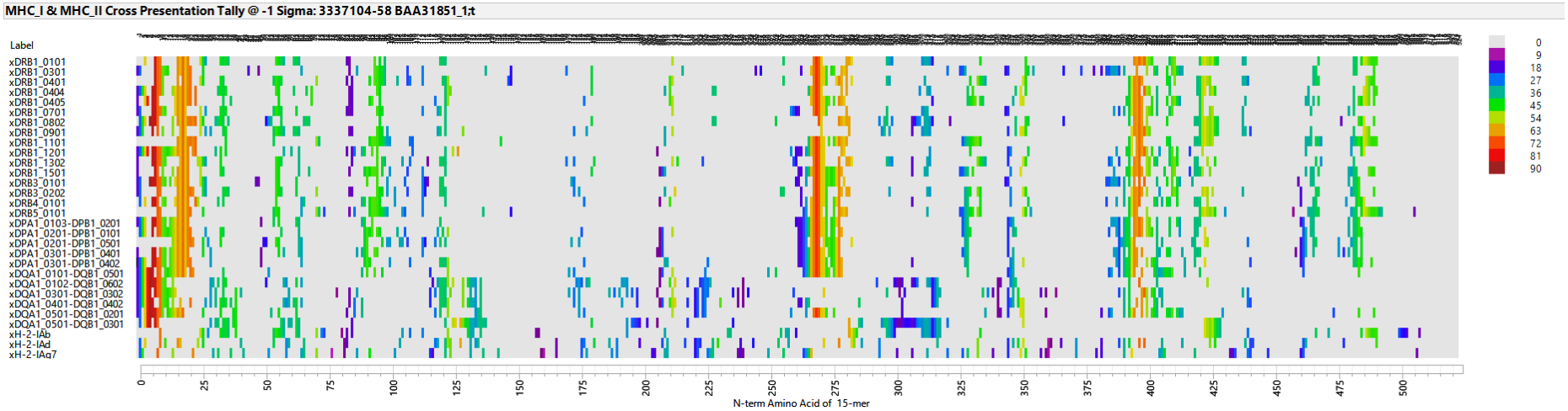

Overlay Plot
3337104-58 BAA31851_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
3337104-58 BAA31851_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1201
nDRB1_0802
nDRB1_0401
nDRB3_0202
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB1_1302
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.58
-2.87
-2.15
-1.43
-0.72
-0.00
0.76
1.52
2.27
3.03
3.79
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

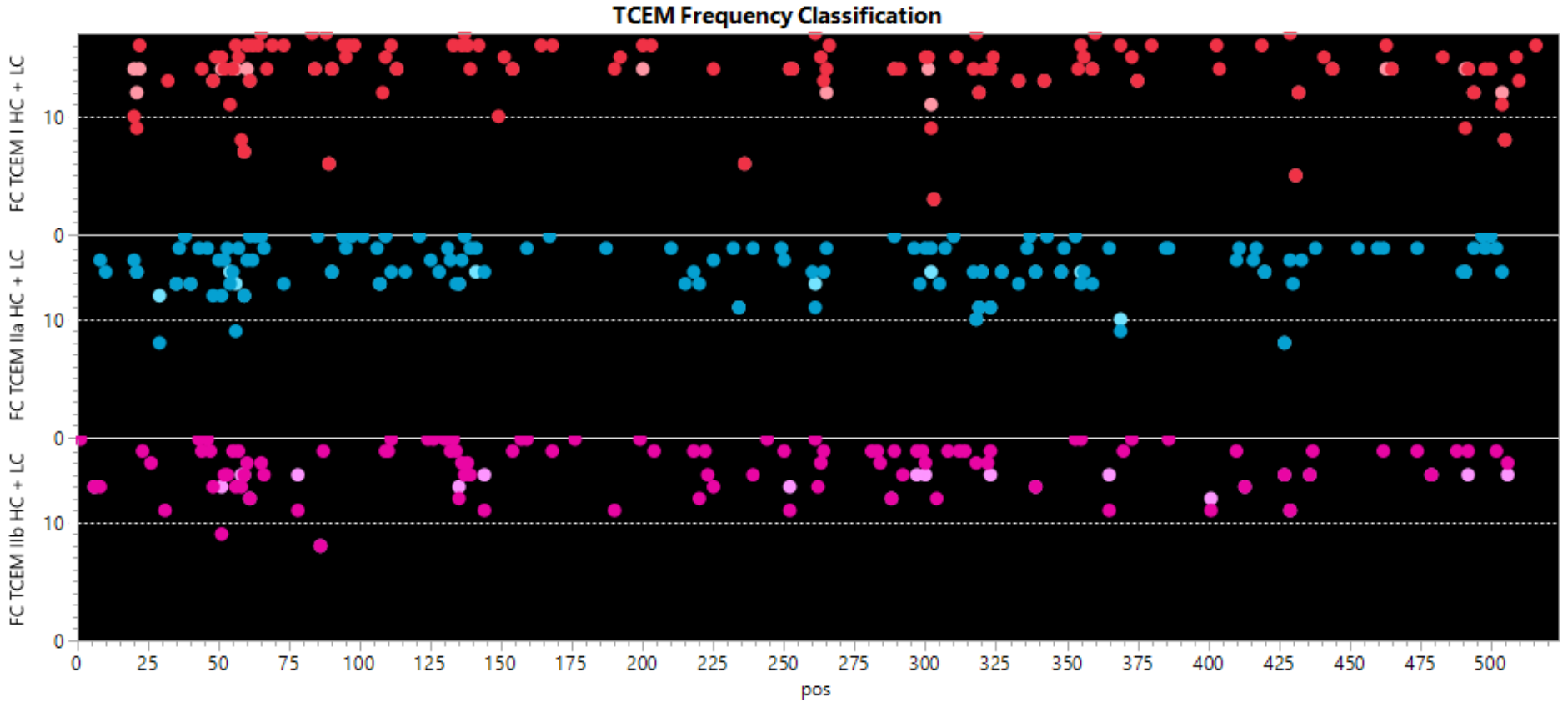

3337104-58 BAA31851_1;t Kernels MHC II @ 15-20 aa by hCAT_S
3337104-58 BAA31851_1;t Kernels MHC II @ 15-20 aa by hCAT_L
3337104-58 BAA31851_1;t Kernels MHC II @ 15-20 aa by hCAT_B

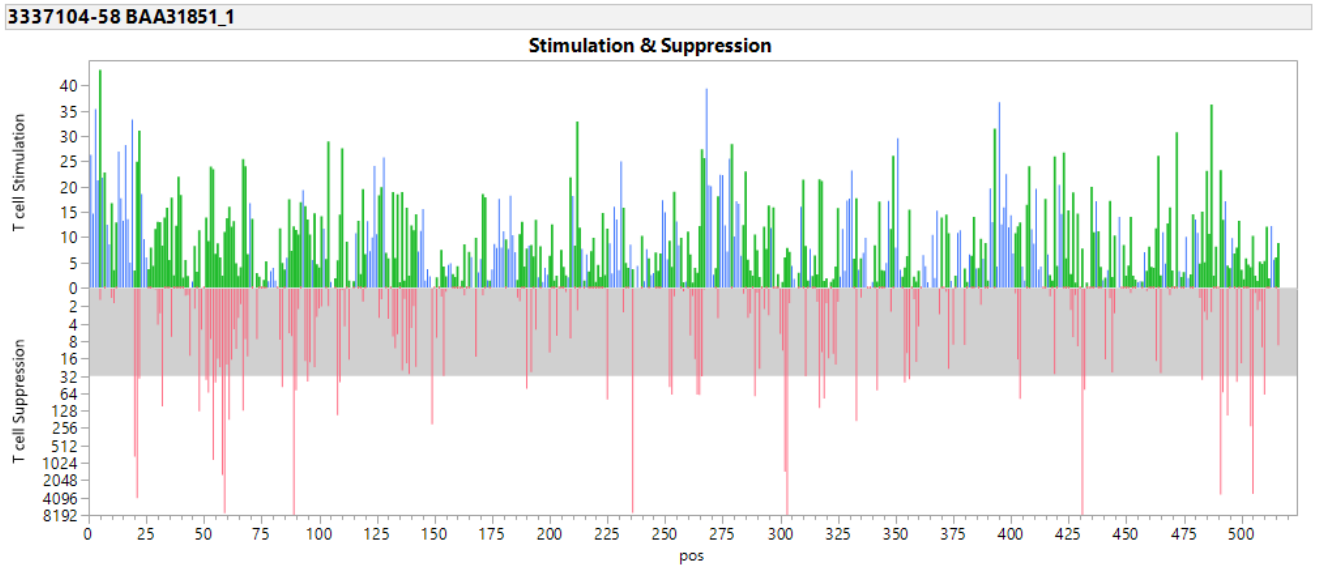

Hierarchical Cluster MHC I
Method = Ward
3337104-58 BAA31851_1;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nA_3201
nH-2-Db
nB_1501
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_3501
nB_5301
nB_5101
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_3001
nB_5401
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
-4.43
-3.55
-2.66
-1.77
-0.89
0.00
1.21
2.42
3.63
4.84
6.05
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer

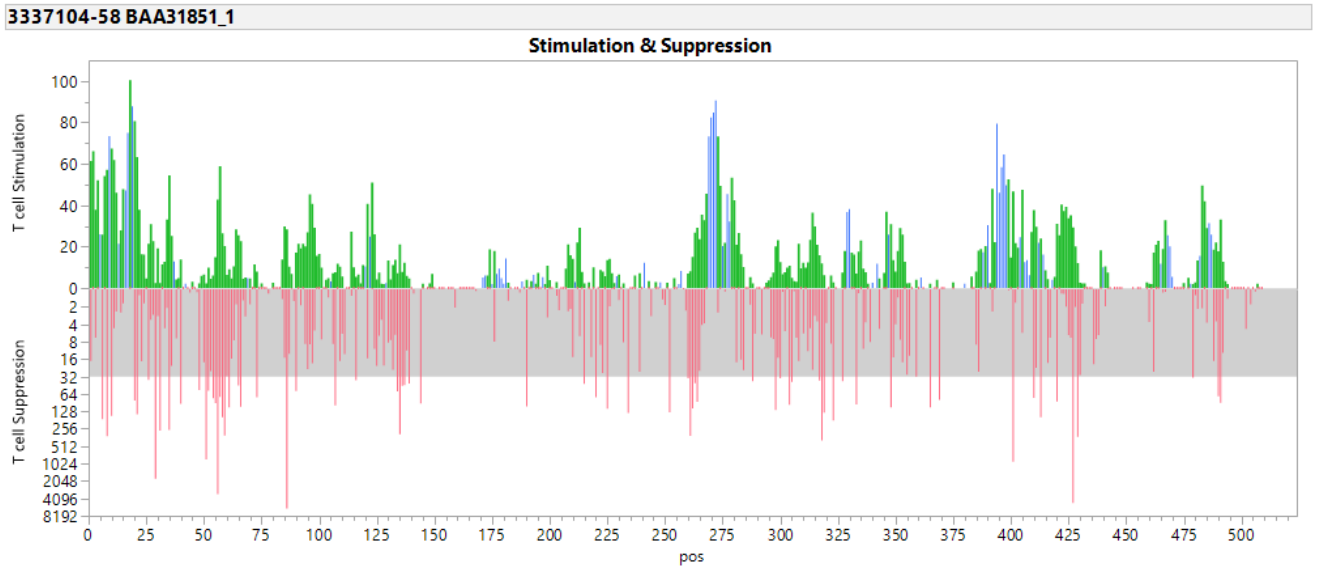

Hierarchical Cluster MHC I
Method = Ward
3337104-58 BAA31851_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nB_1501
nA_2902
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_5401
nB_3501
nB_5101
nB_5301
nH-2-Db
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nA_3201
nB_1503
nB_2705
nH-2-Kd
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nH-2-Ld
nB_1801
nB_4402
nB_4001
nB_4403
nH-2-Kk
nB_4002
nB_4501
0.00
0.02
0.04
0.06
0.09
0.11
0.69
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 3337104-58 BAA31851_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer
0
9
18
27
36
45
54
63
72
81
90

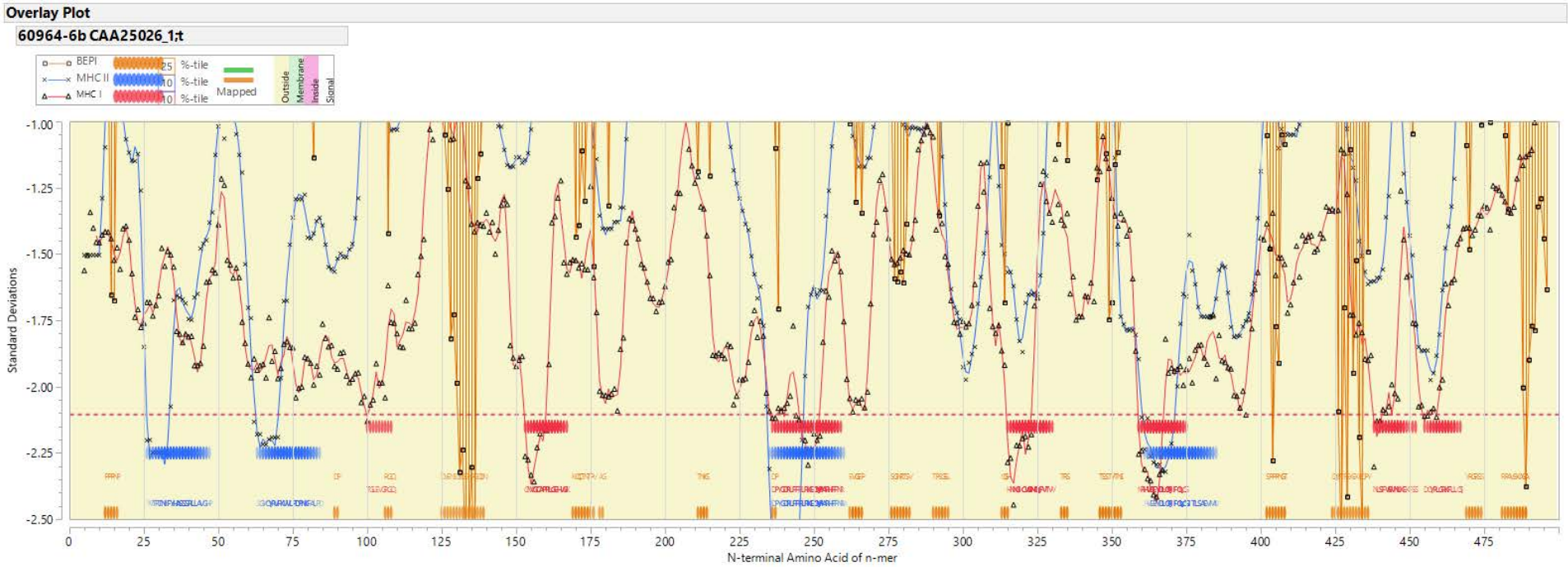
3337104-58 BAA31851_1

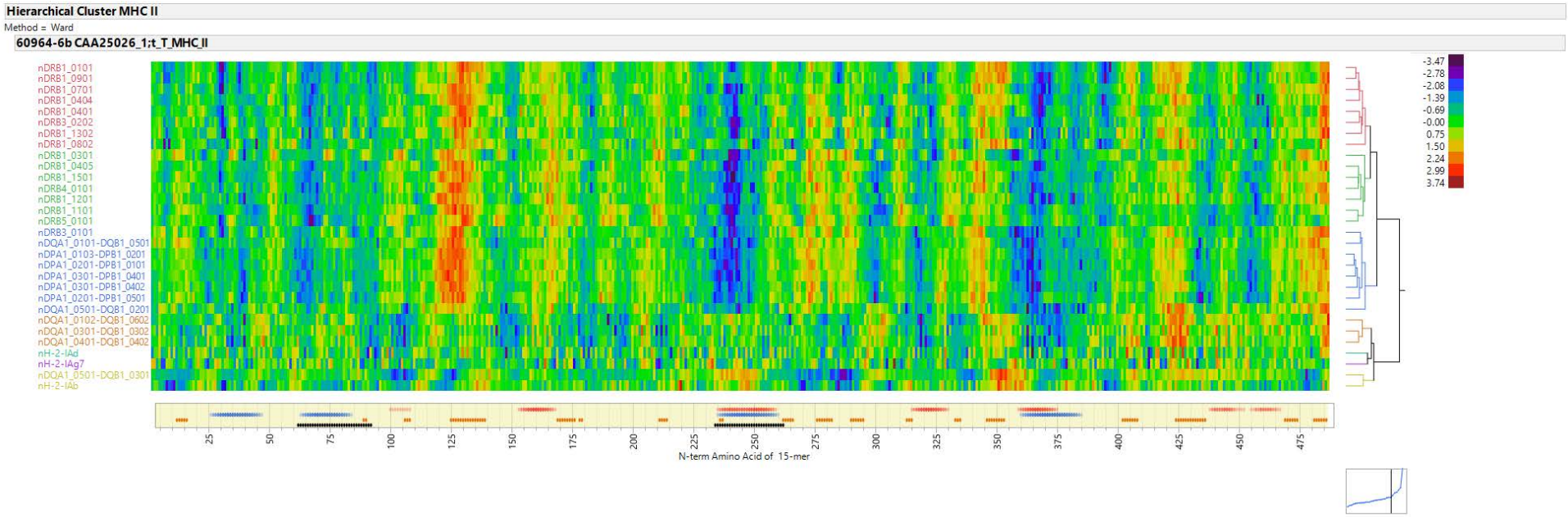


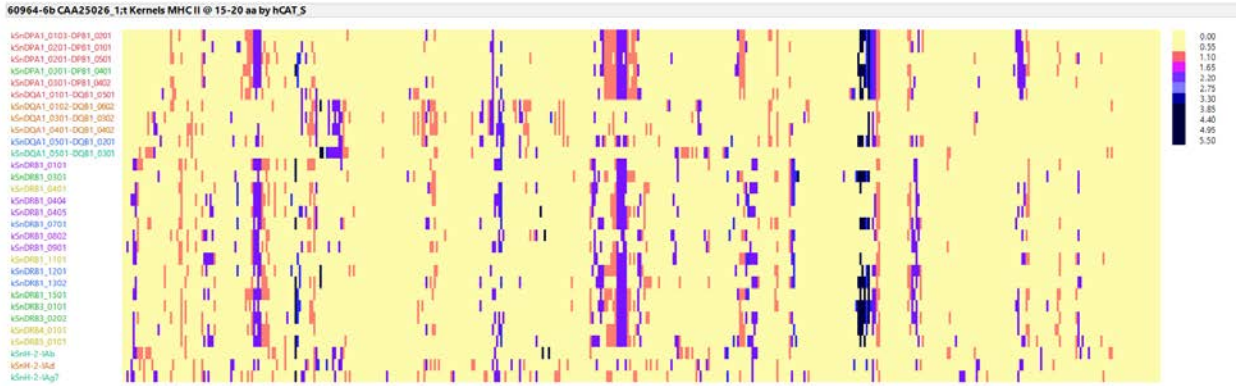

3337104-58 BAA31851_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

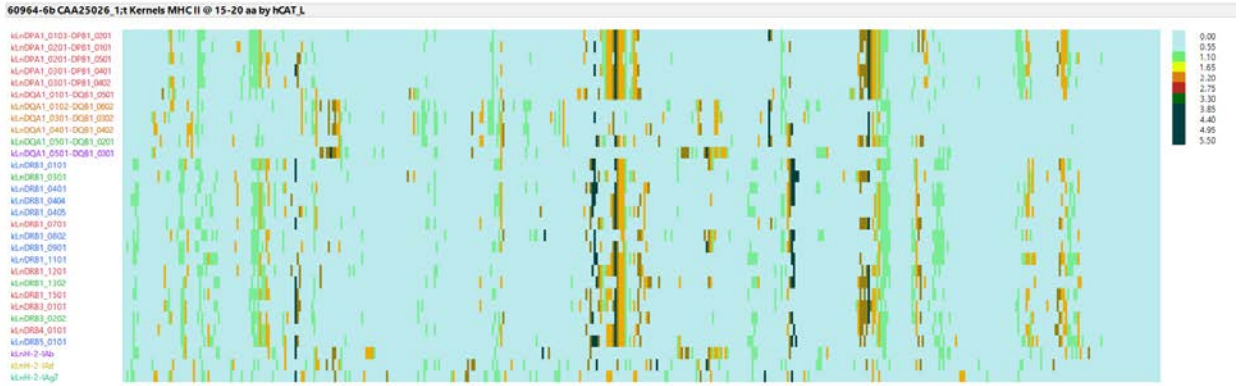

3337104-58 BAA31851_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

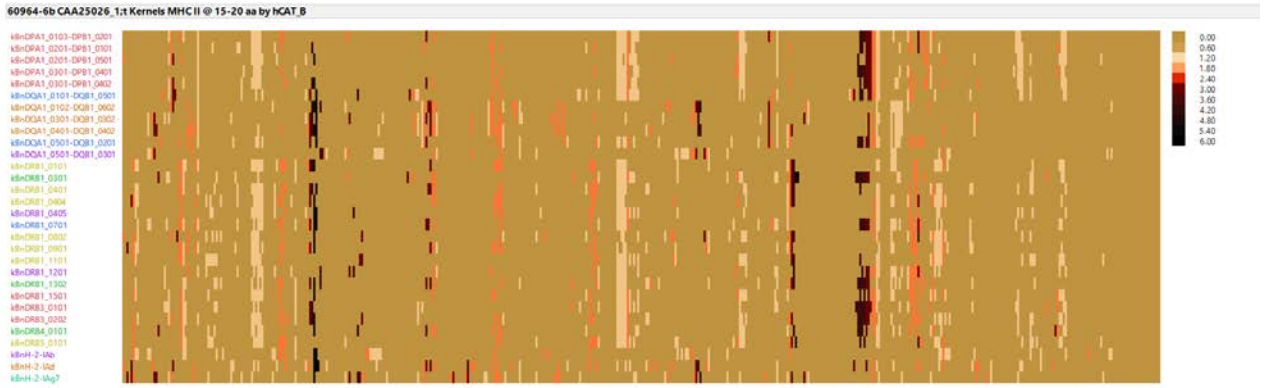
HPV 6b

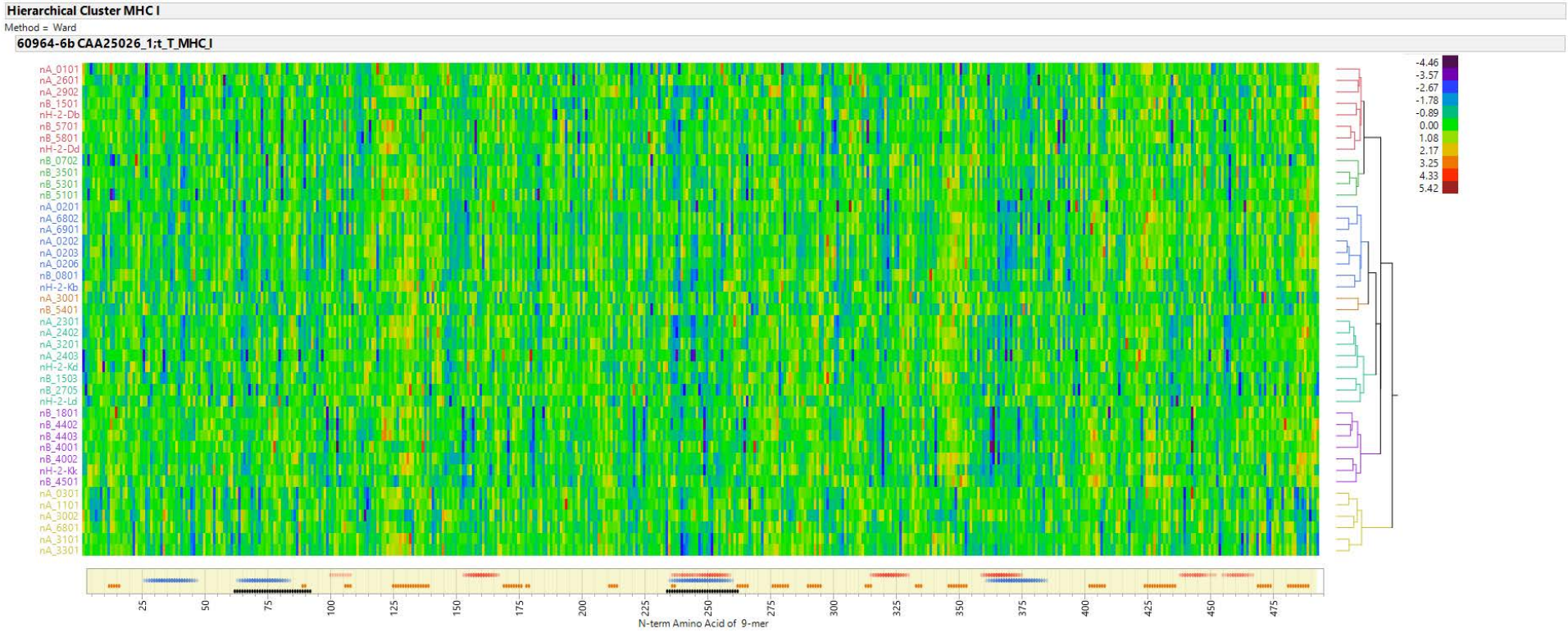

Overlay Plot
60964-6b CAA25026_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-terminal Amino Acid of n-mer

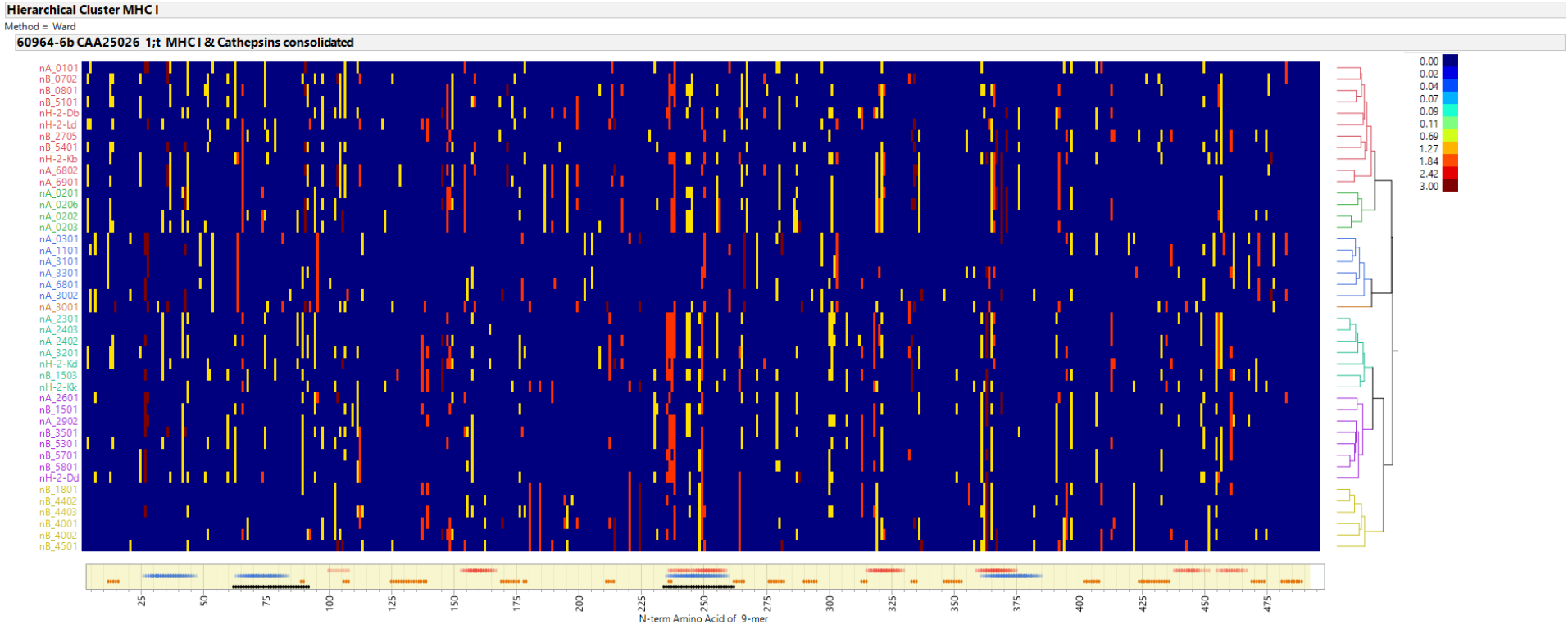

Hierarchical Cluster MHC II
Method = Ward
60964-6b CAA25026_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_0401
nDRB3_0202
nDRB1_1302
nDRB1_0802
nDRB1_0301
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB3_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.47
-2.78
-2.08
-1.39
-0.69
-0.00
0.75
1.50
2.24
2.99
3.74
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

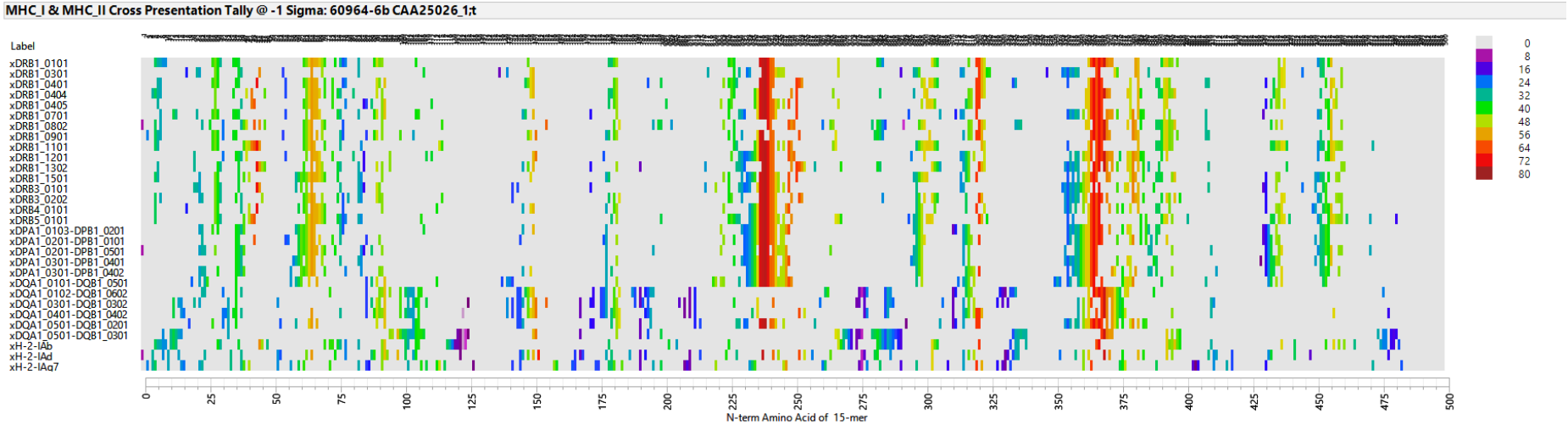

60964-6b CAA25026_1:t Kernels MHC II @ 15-20 aa by hCAT_S


60964-6b CAA25026_1:t Kernels MHC II @ 15-20 aa by hCAT_L

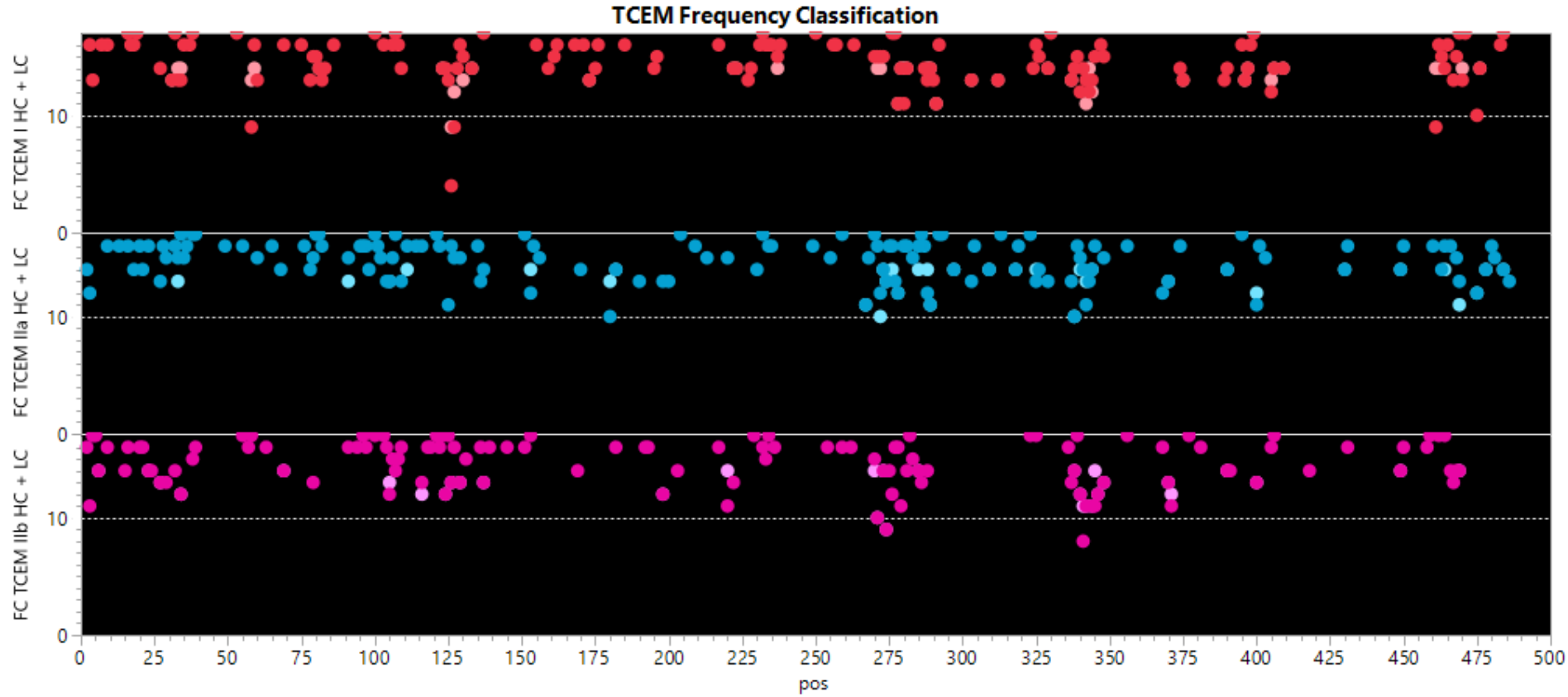

60964-6b CAA25026_1:t Kernels MHC II @ 15-20 aa by hCAT_B

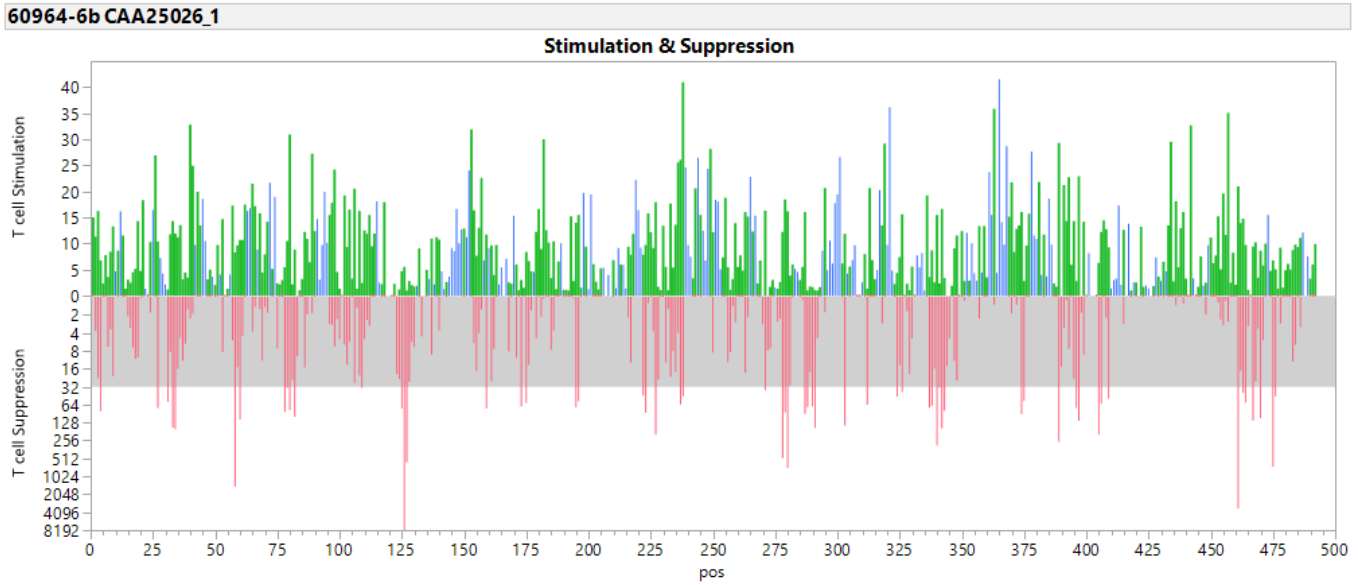

Hierarchical Cluster MHC I
Method = Ward
60964-6b CAA25026_1;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nB_1501
nH-2-Db
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_3501
nB_5301
nB_5101
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_3001
nB_5401
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nB_1503
nB_2705
nH-2-Ld
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
-4.46
-3.57
-2.67
-1.78
-0.89
0.00
1.08
2.17
3.25
4.33
5.42
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

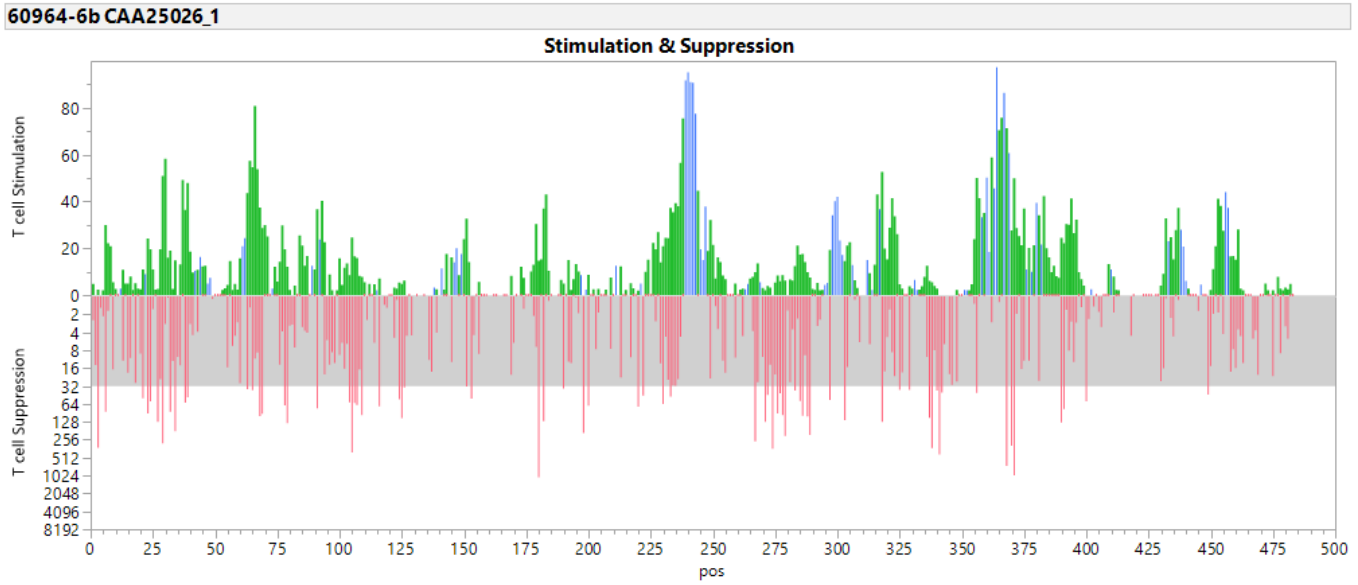

Hierarchical Cluster MHC I
Method = Ward
60964-6b CAA25026_1;t MHC I & Cathepsins consolidated
nA_0101
nB_0702
nB_0801
nB_5101
nH-2-Db
nH-2-Ld
nB_2705
nB_5401
nH-2-Kb
nA_6802
nA_6901
nA_0201
nA_0206
nA_0202
nA_0203
nA_0301
nA_1101
nA_3101
nA_3301
nA_6801
nA_3002
nA_3001
nA_2301
nA_2403
nA_2402
nA_3201
nH-2-Kd
nB_1503
nH-2-Kk
nA_2601
nB_1501
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 60964-6b CAA25026_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

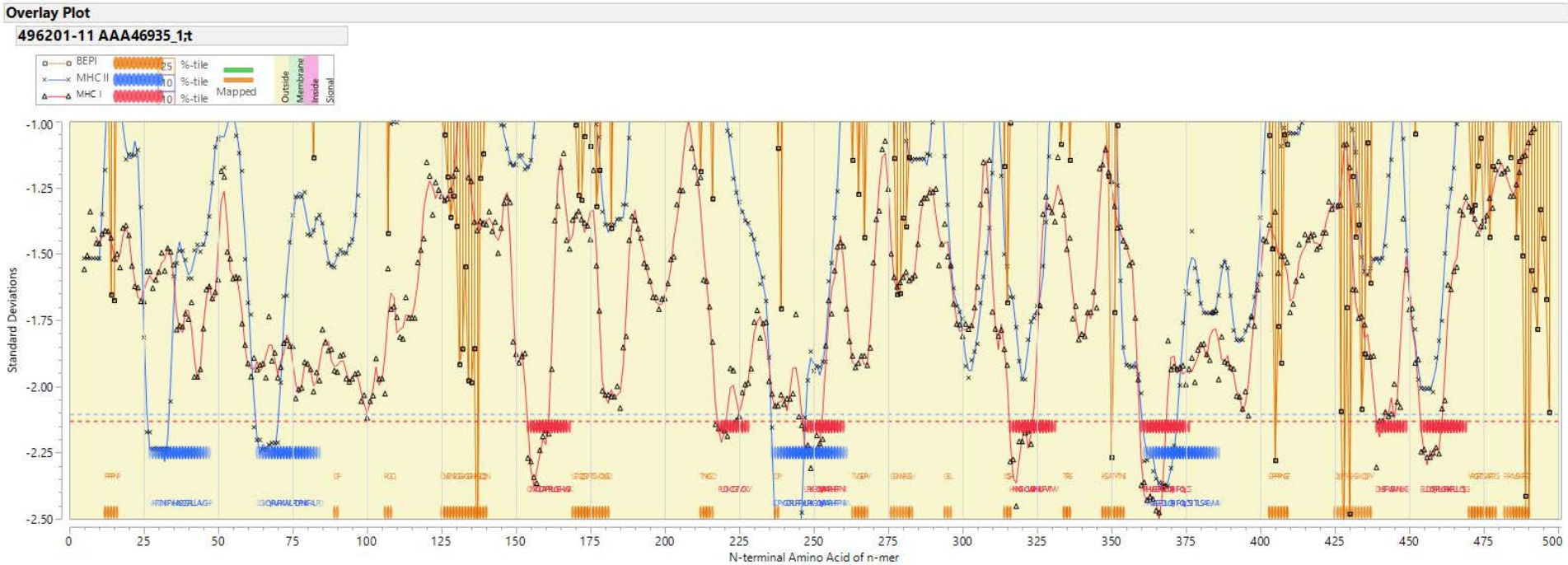
60964-6b CAA25026_1

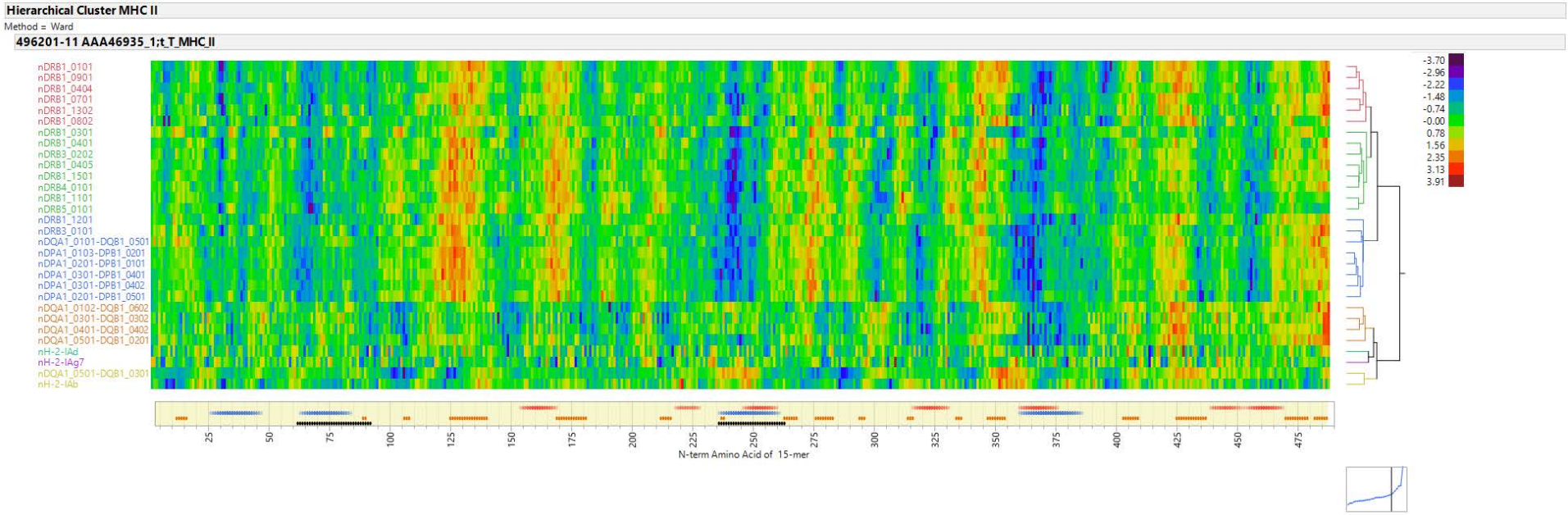


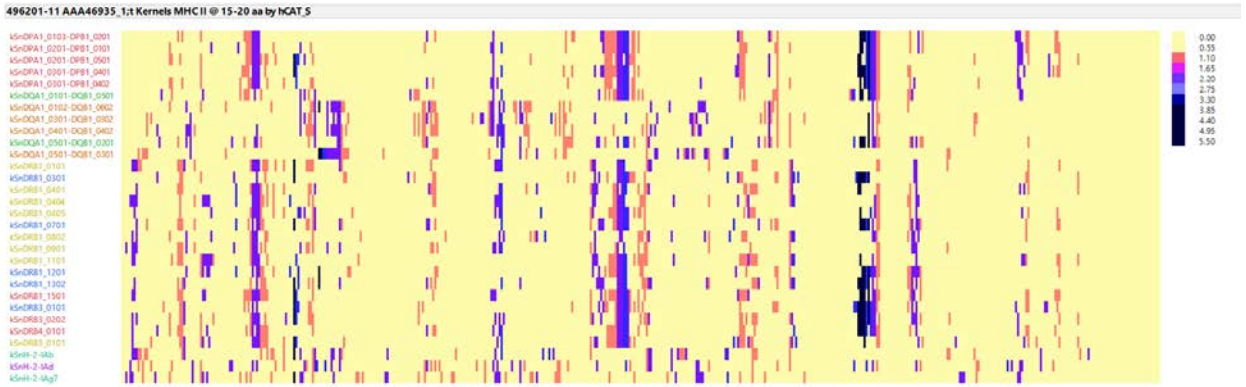

60964-6b CAA25026_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

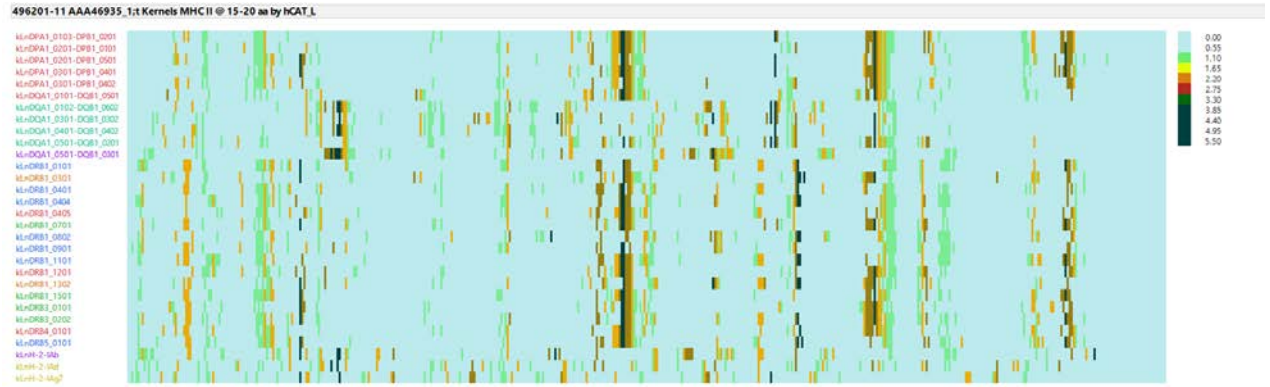

60964-6b CAA25026_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

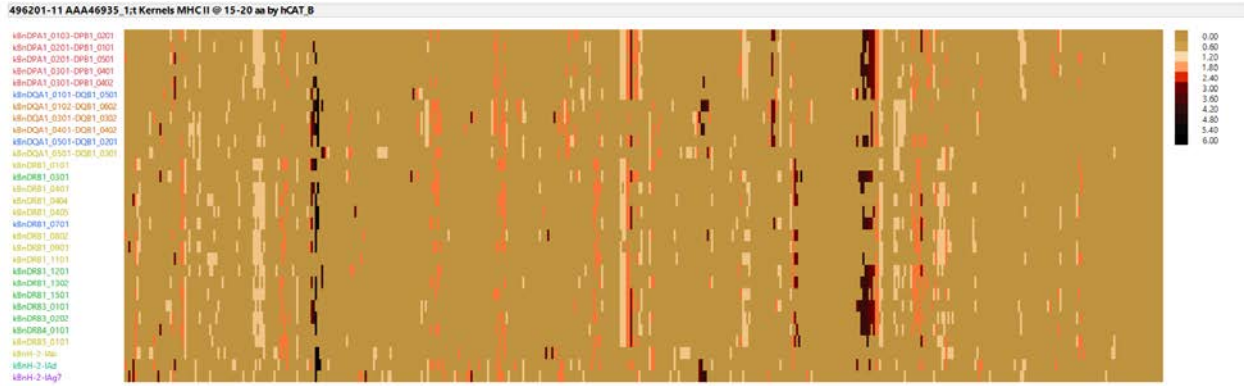
HPV 11

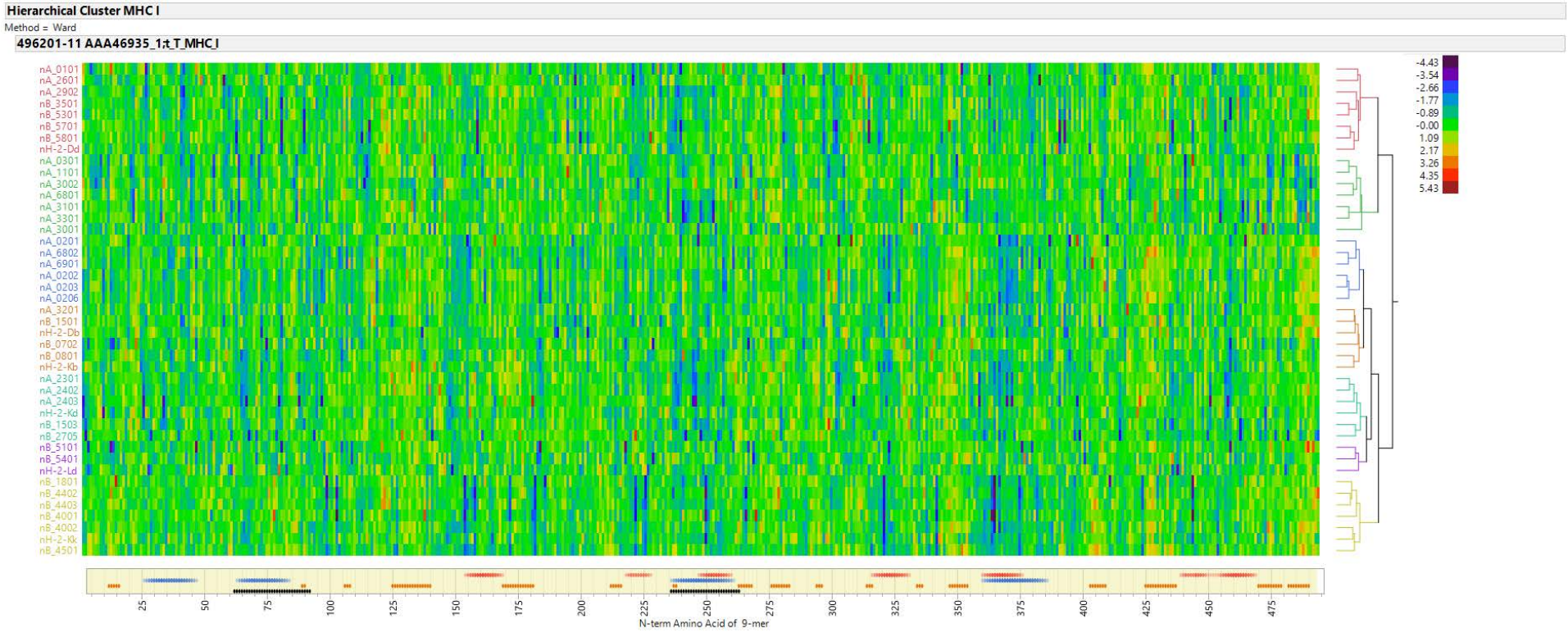

Overlay Plot
496201-11 AAA46935_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-terminal Amino Acid of n-mer

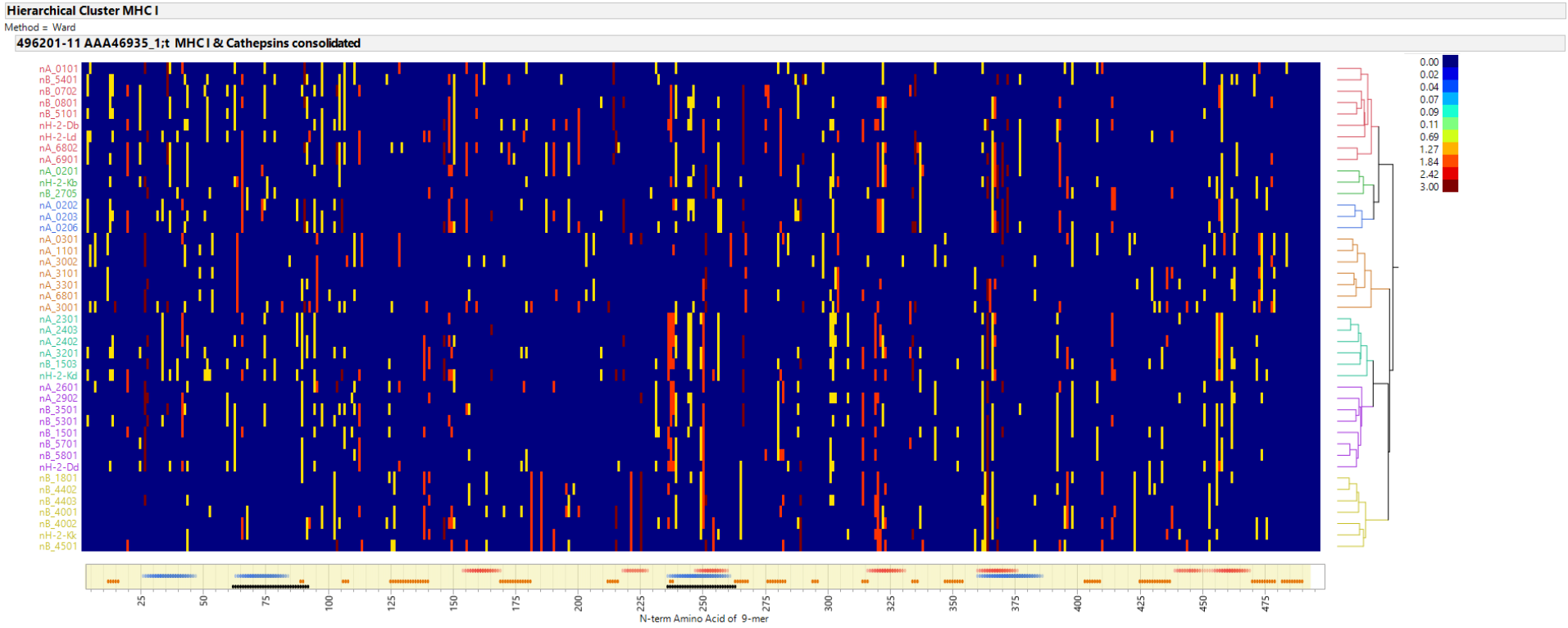

Hierarchical Cluster MHC II
Method = Ward
496201-11 AAA46935_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0404
nDRB1_0701
nDRB1_1302
nDRB1_0802
nDRB1_0301
nDRB1_0401
nDRB3_0202
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDRB1_1201
nDRB3_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.70
-2.96
-2.22
-1.48
-0.74
-0.00
0.78
1.56
2.35
3.13
3.91
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

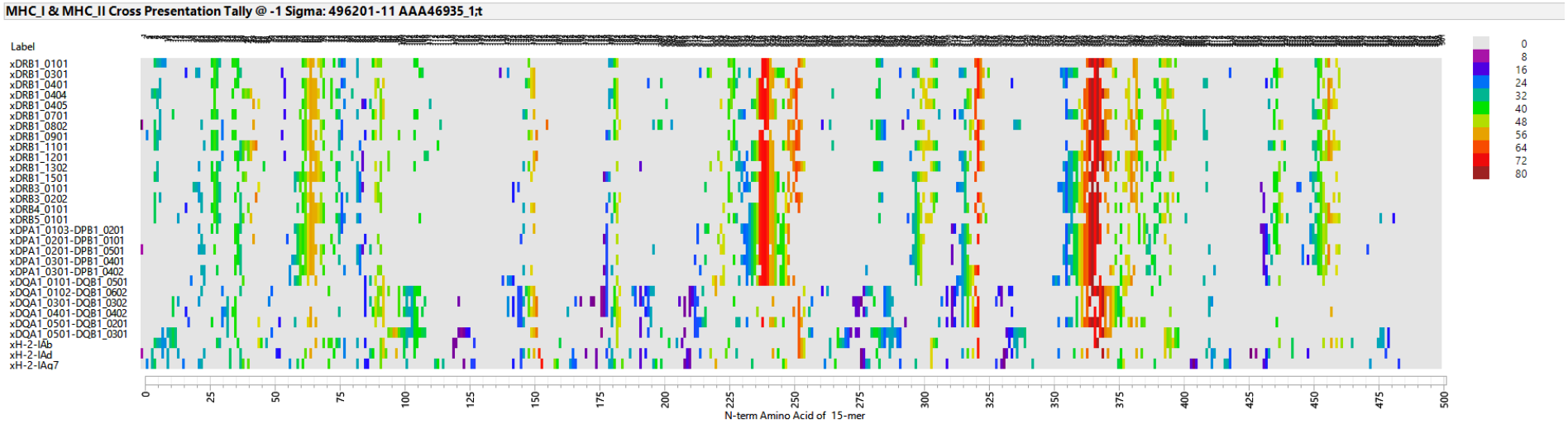

496201-11 AAA46935_1;t Kernels MHC II @ 15-20 aa by hCAT_S


496201-11 AAA46935_1;t Kernels MHC II @ 15-20 aa by hCAT_L

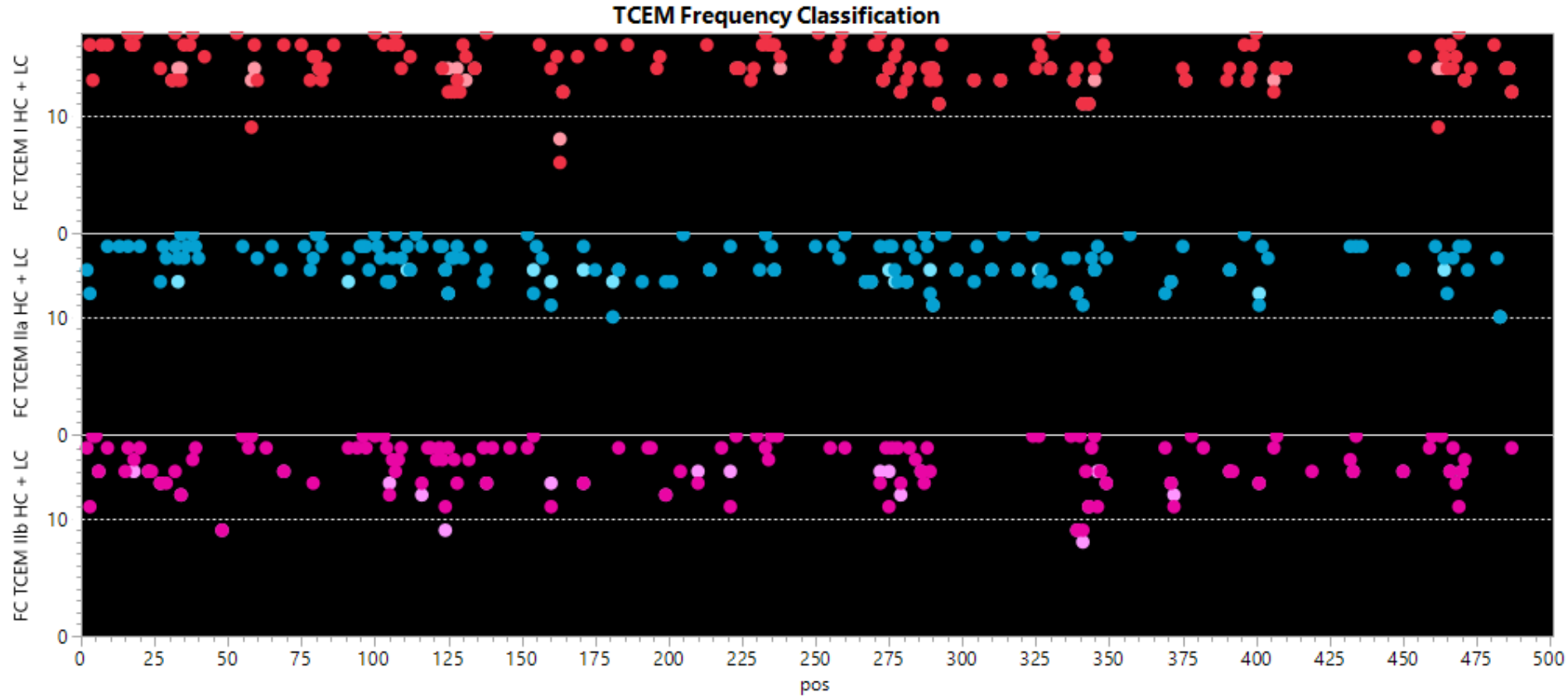

496201-11 AAA46935_1;t Kernels MHC II @ 15-20 aa by hCAT_B

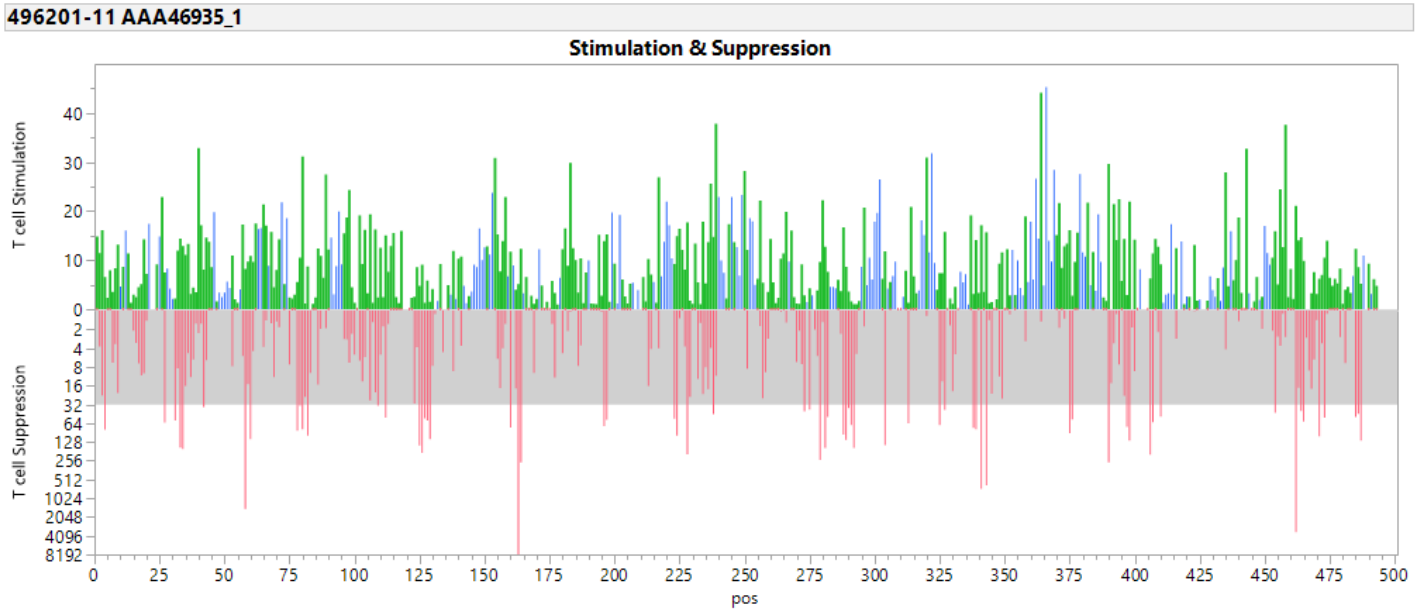

Hierarchical Cluster MHC I
Method = Ward
496201-11 AAA46935_1;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_3001
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_3201
nB_1501
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_5101
nB_5401
nH-2-Ld
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
-4.43
-3.54
-2.66
-1.77
-0.89
-0.00
1.09
2.17
3.26
4.35
5.43
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

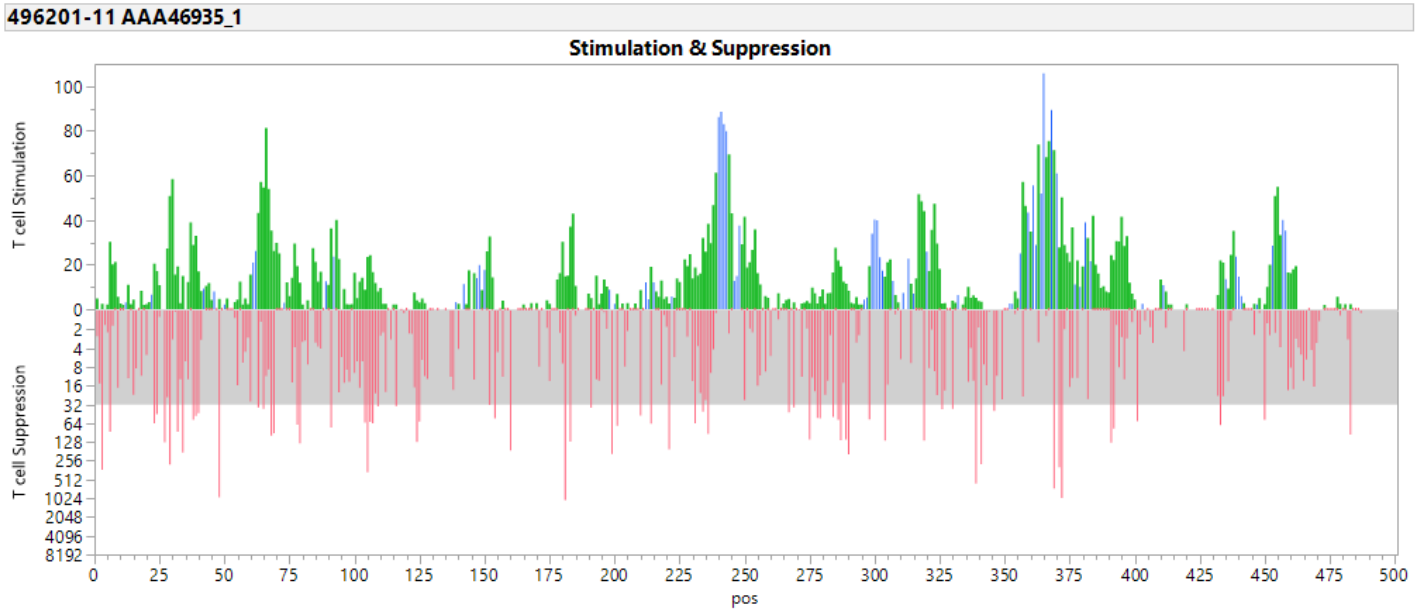

Hierarchical Cluster MHC I
Method = Ward
496201-11 AAA46935_1;t MHC I & Cathepsins consolidated
nA_0101
nB_5401
nB_0702
nB_0801
nB_5101
nH-2-Db
nH-2-Ld
nA_6802
nA_6901
nA_0201
nH-2-Kb
nB_2705
nA_0202
nA_0203
nA_0206
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nA_2301
nA_2403
nA_2402
nA_3201
nB_1503
nH-2-Kd
nA_2601
nA_2902
nB_3501
nB_5301
nB_1501
nB_5701
nB_5801
nH-2-Dd
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 496201-11 AAA46935_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

496201-11 AAA46935_1




496201-11 AAA46935_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos


496201-11 AAA46935_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

ioGenetics®

Addendum 4 Measles

Enders Edmonston

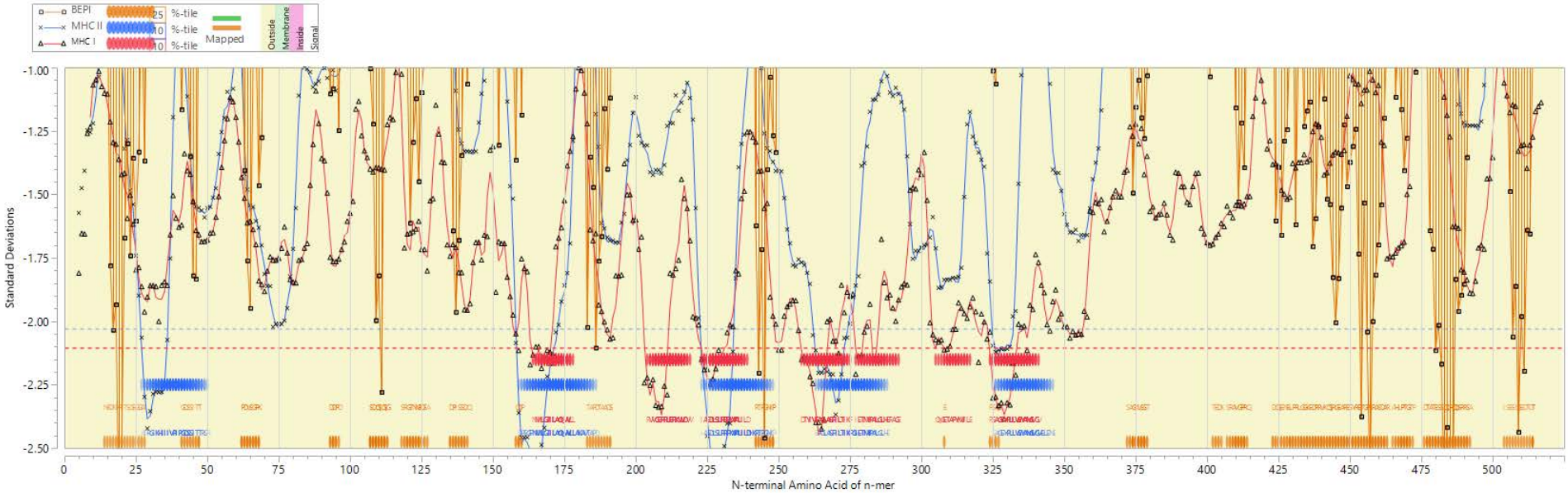
Gi: 224482351-224482358

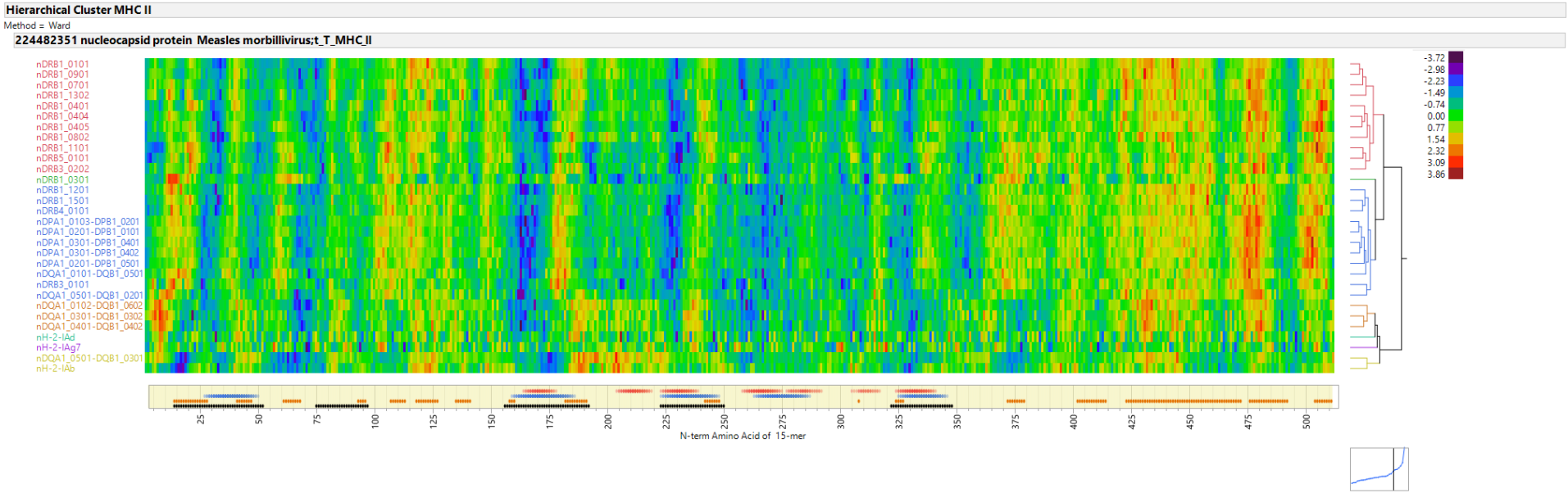
ATTENUVAX® (Measles Virus Vaccine Live), a more attenuated line of measles virus, derived from Enders' attenuated Edmonston strain and propagated in chick embryo cell culture

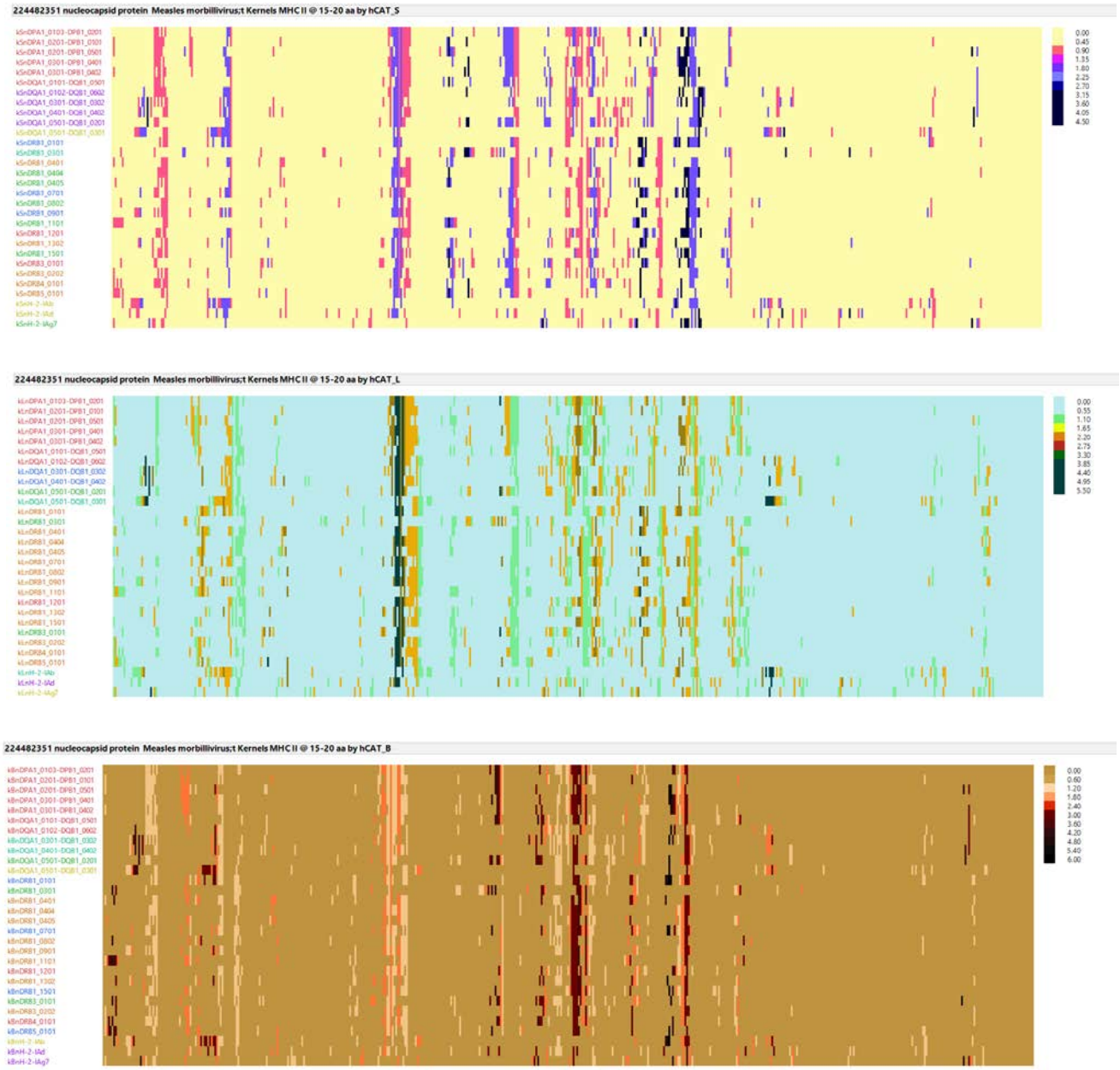
Nucelocpasid

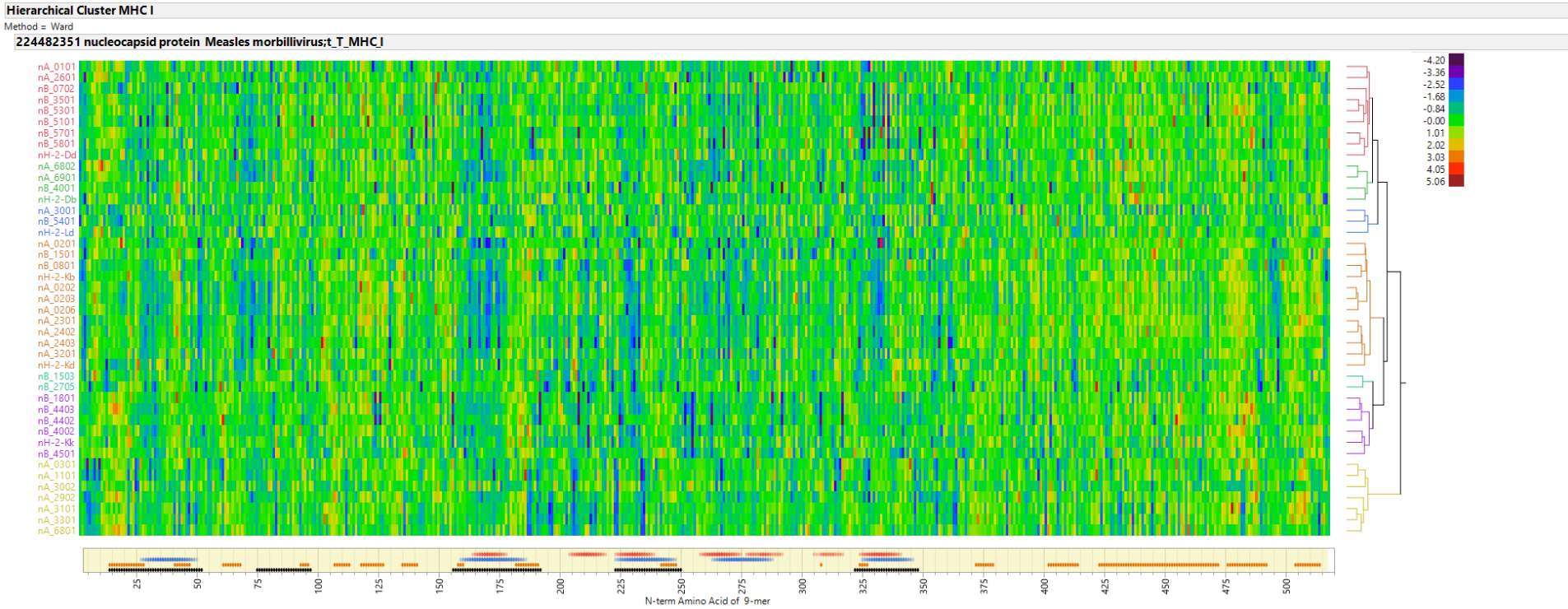
Overlay Plot

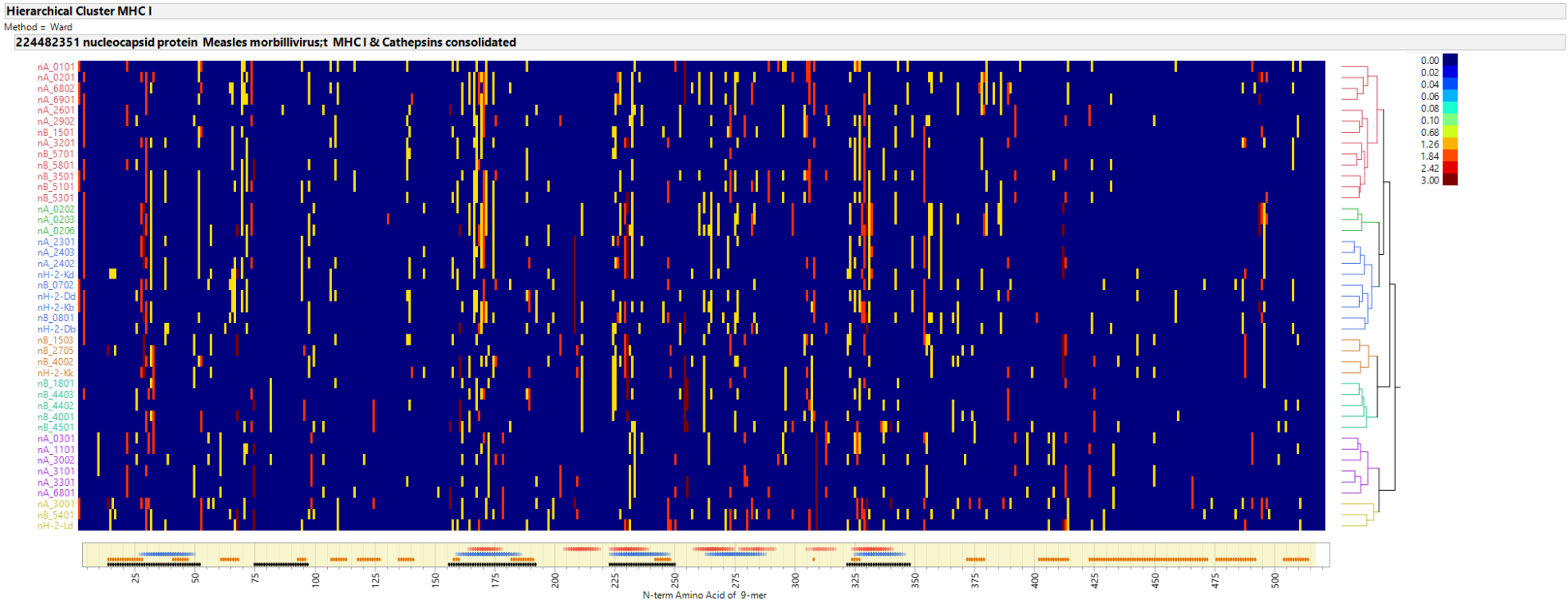
224482351 nucleocapsid protein Measles morbillivirus;t

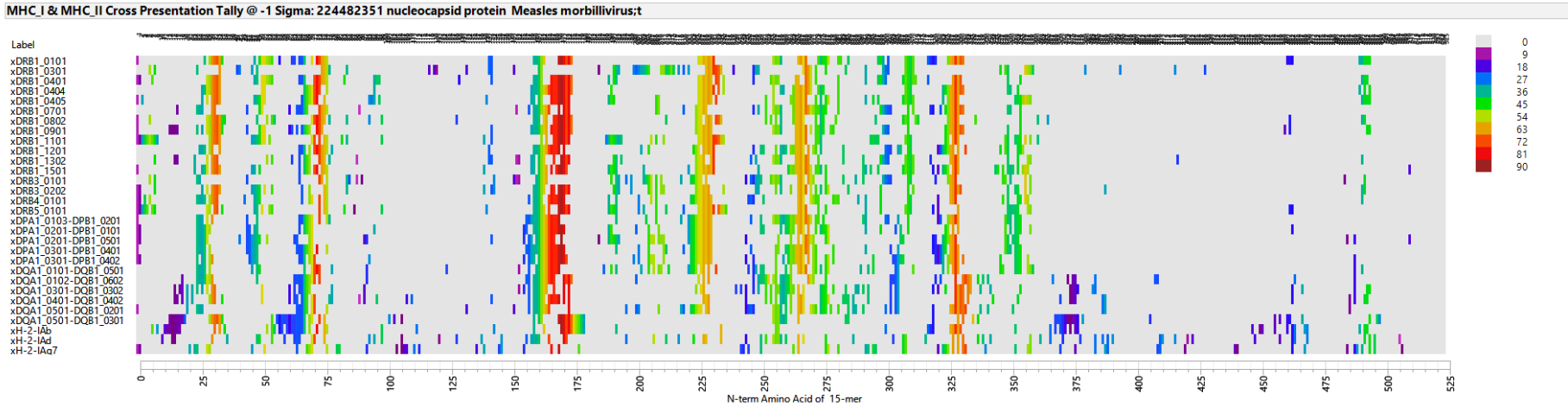



Hierarchical Cluster MHC II
Method = Ward
224482351 nucleocapsid protein Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB1_0404
nDRB1_0405
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB3_0202
nDRB1_0301
nDRB1_1201
nDRB1_1501
nDRB4_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDRB3_0101
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.72
-2.98
-2.23
-1.49
-0.74
0.00
0.77
1.54
2.32
3.09
3.86
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

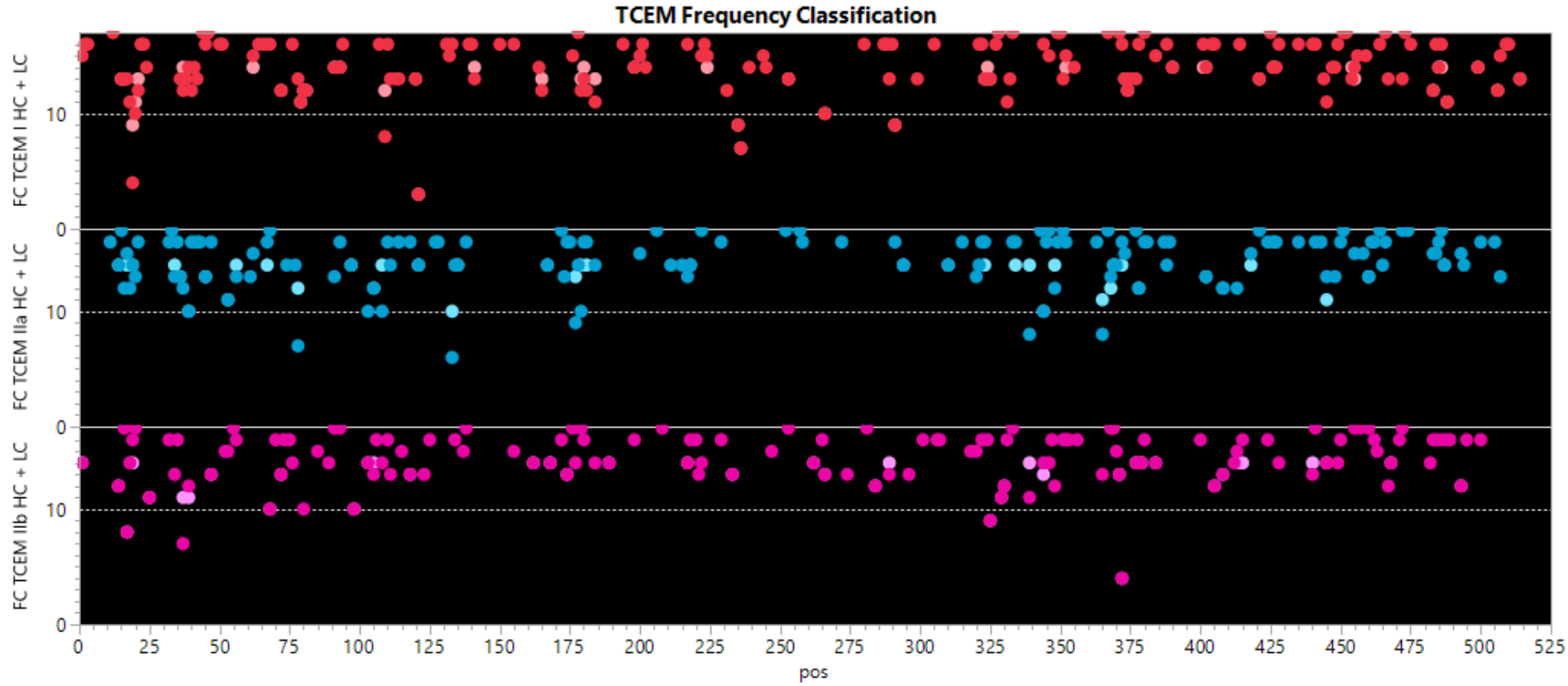

224482351 nucleocapsid protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S
224482351 nucleocapsid protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L
224482351 nucleocapsid protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
224482351 nucleocapsid protein Measles morbillivirus;t_T_MHC_I
nA_0101
nA_2601
nB_0702
nB_3501
nB_5301
nB_5101
nB_5701
nB_5801
nH-2-Dd
nA_6802
nA_6901
nB_4001
nH-2-Db
nA_3001
nB_5401
nH-2-Ld
nA_0201
nB_1501
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_2403
nA_3201
nH-2-Kd
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.20
-3.36
-2.52
-1.68
-0.84
-0.00
1.01
2.02
3.03
4.05
5.06
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer

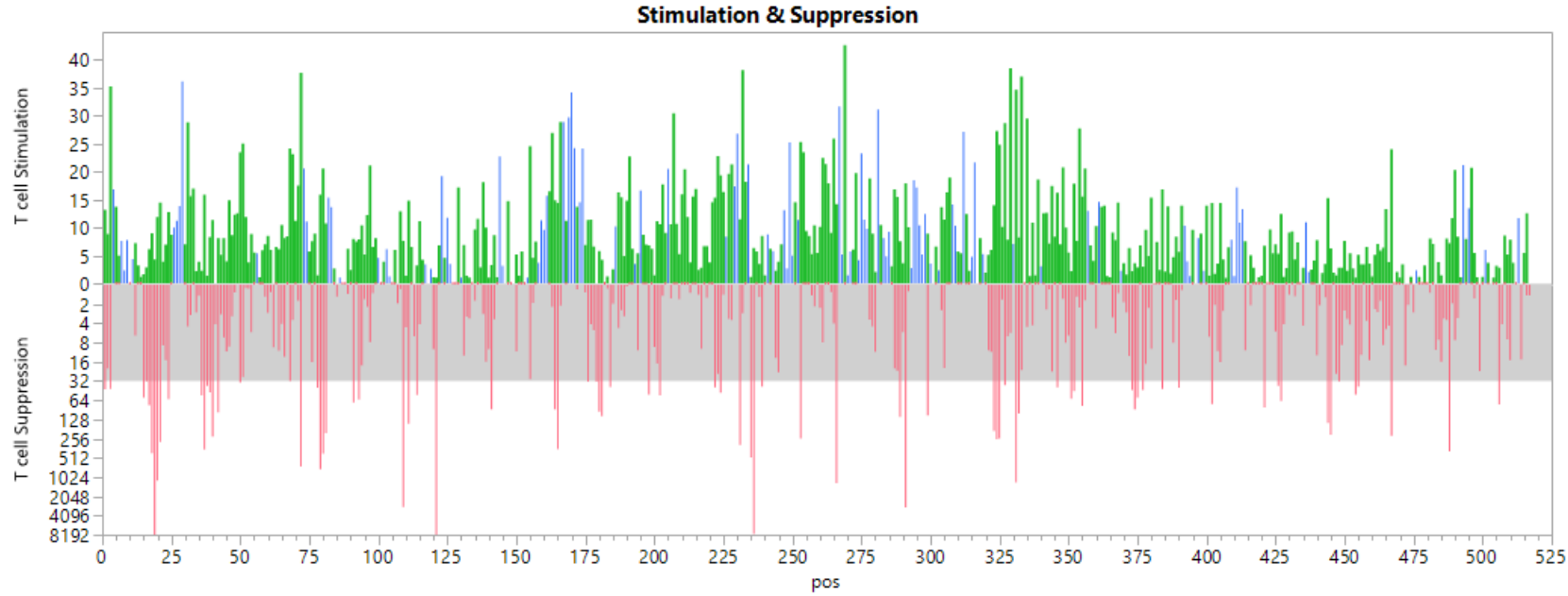

Hierarchical Cluster MHC I
Method = Ward
224482351 nucleocapsid protein Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nA_0201
nA_6802
nA_6901
nA_2601
nA_2902
nB_1501
nA_3201
nB_5701
nB_5801
nB_3501
nB_5101
nB_5301
nA_0202
nA_0203
nA_0206
nA_2301
nA_2403
nA_2402
nH-2-Kd
nB_0702
nH-2-Dd
nH-2-Kb
nB_0801
nH-2-Db
nB_1503
nB_2705
nB_4002
nH-2-Kk
nB_1801
nB_4403
nB_4402
nB_4001
nB_4501
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nB_5401
nH-2-Ld
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482351 nucleocapsid protein Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525
N-term Amino Acid of 15-mer
0 9 18 27 36 45 54 63 72 81 90

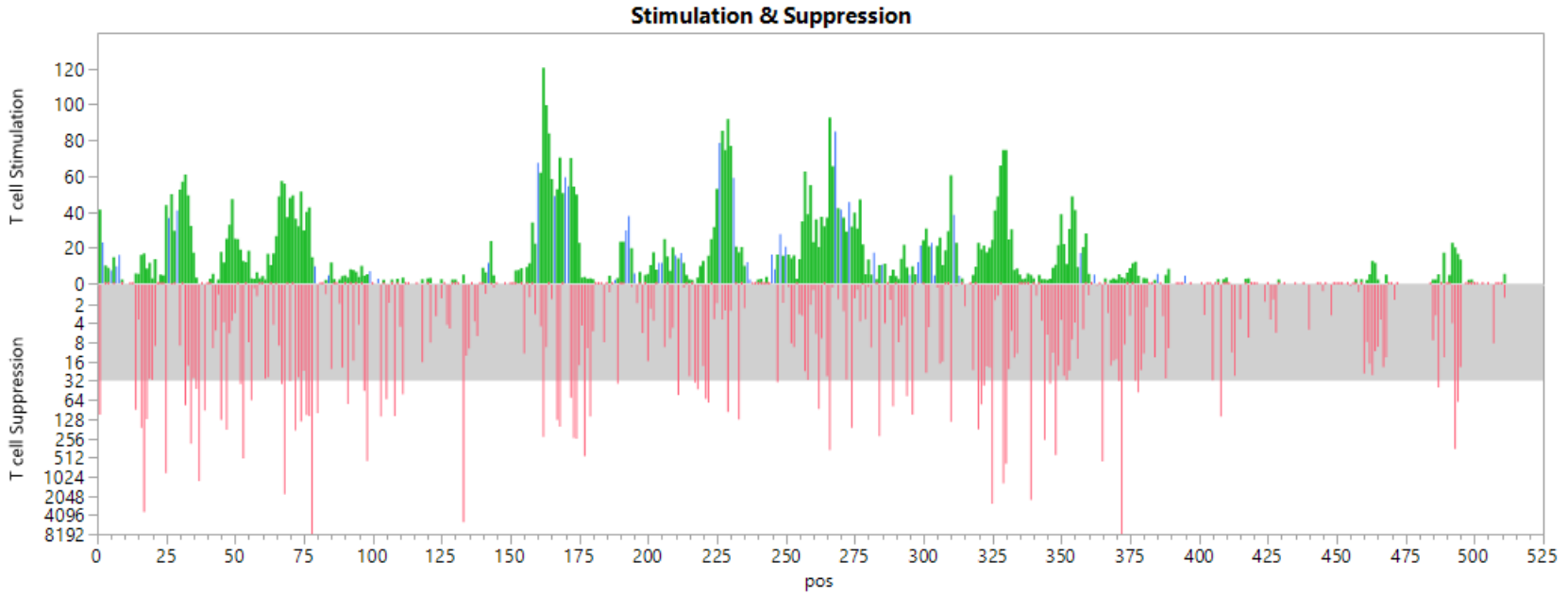
224482351 nucleocapsid protein Measles morbillivirus



224482351 nucleocapsid protein Measles morbillivirus



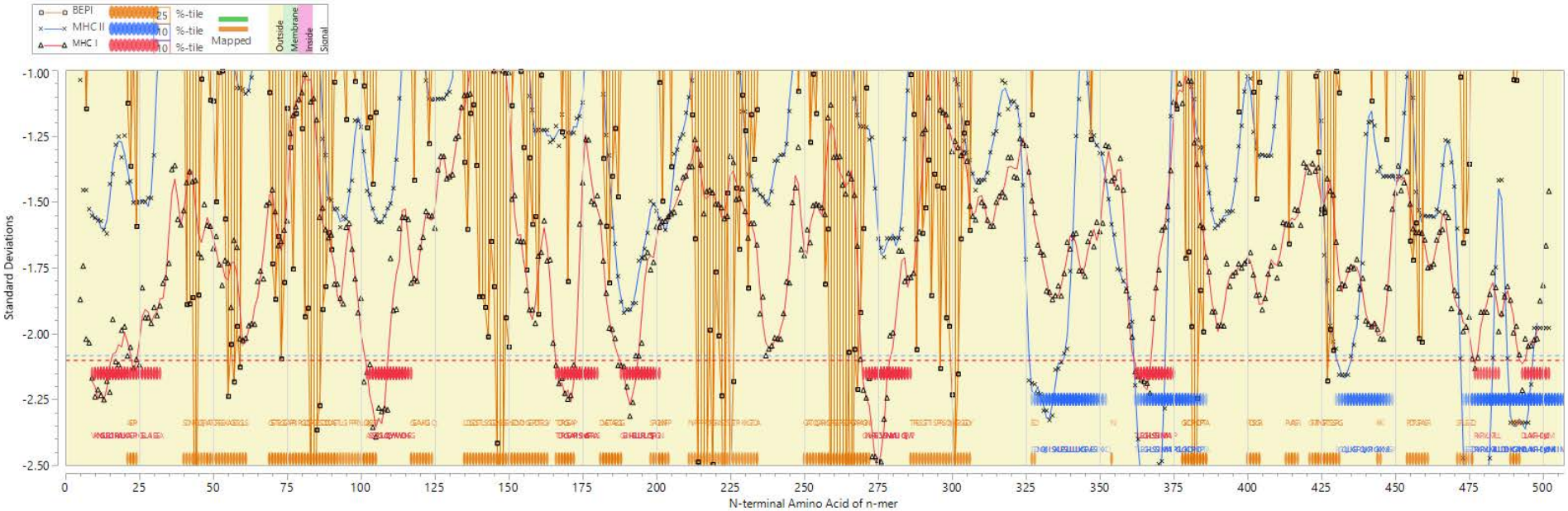
224482351 nucleocapsid protein Measles morbillivirus

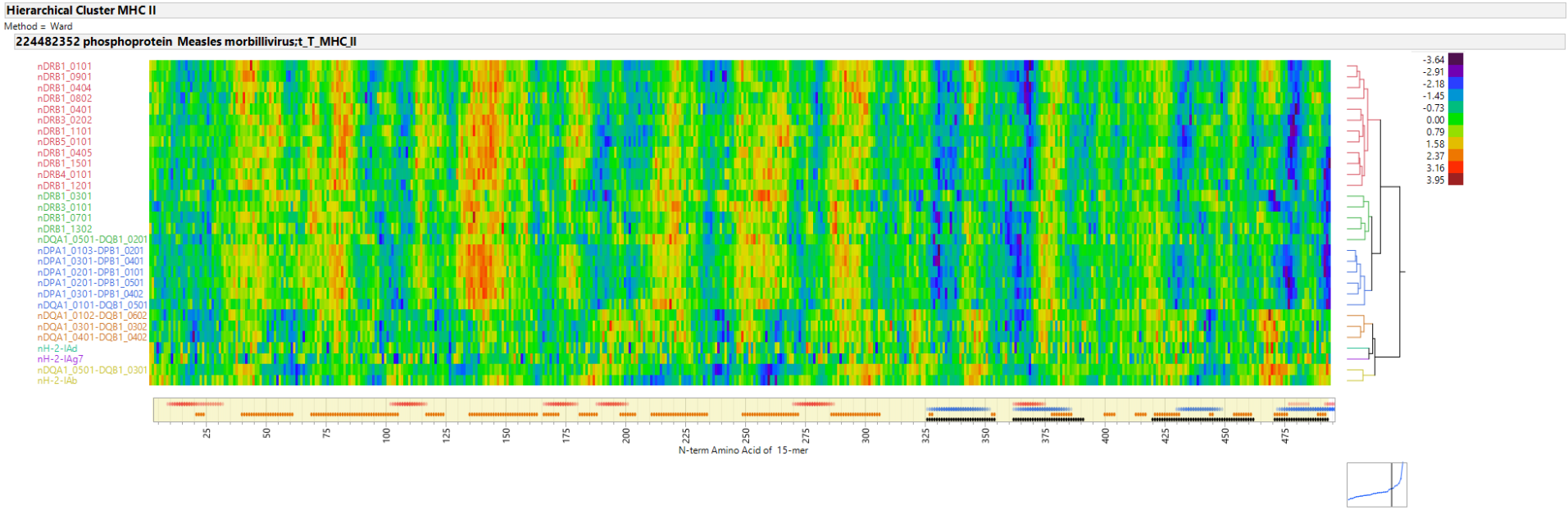


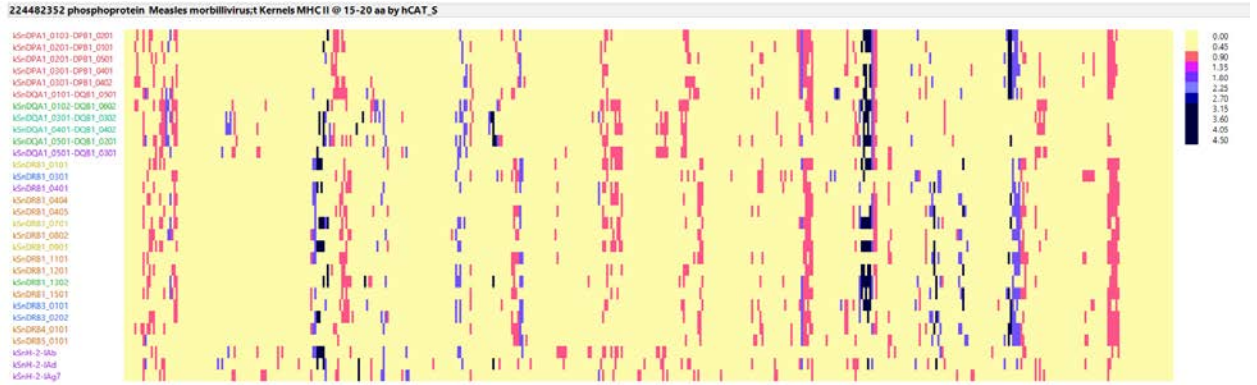
Phosphoprotein

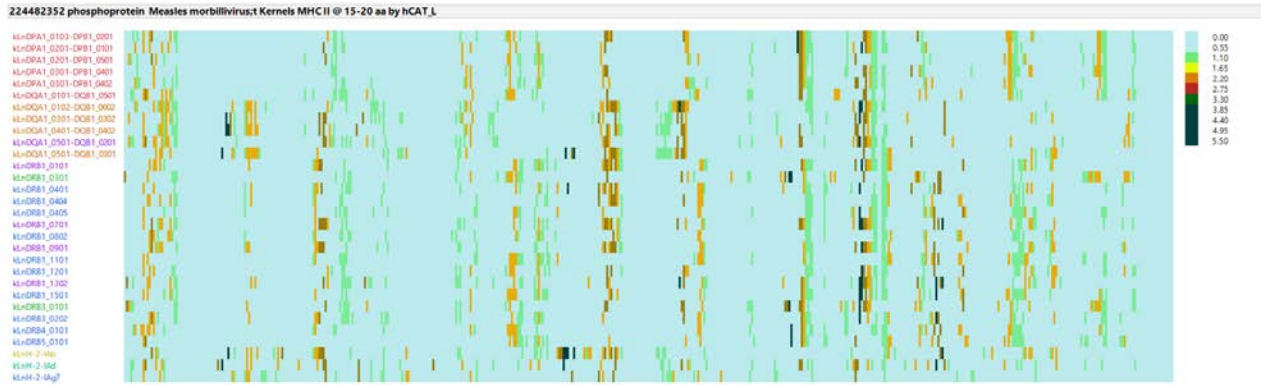
Overlay Plot

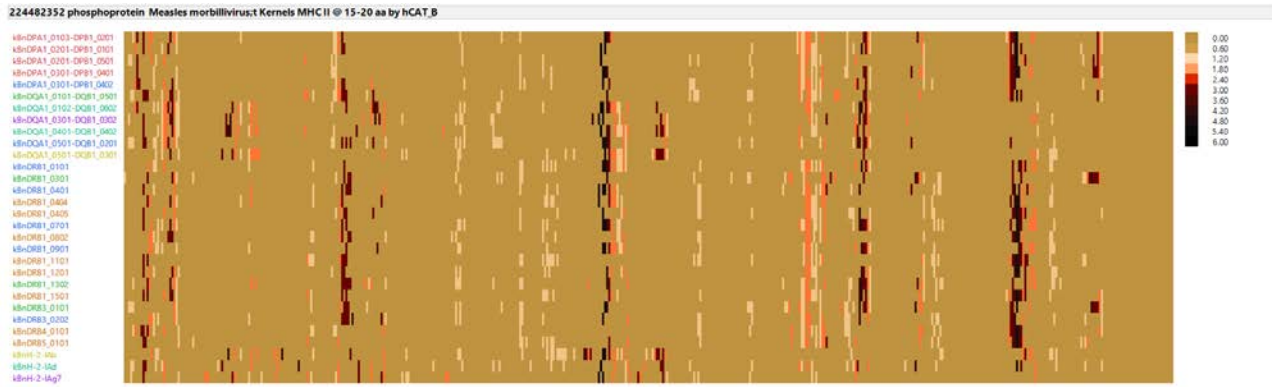
224482352 phosphoprotein Measles morbillivirus;t

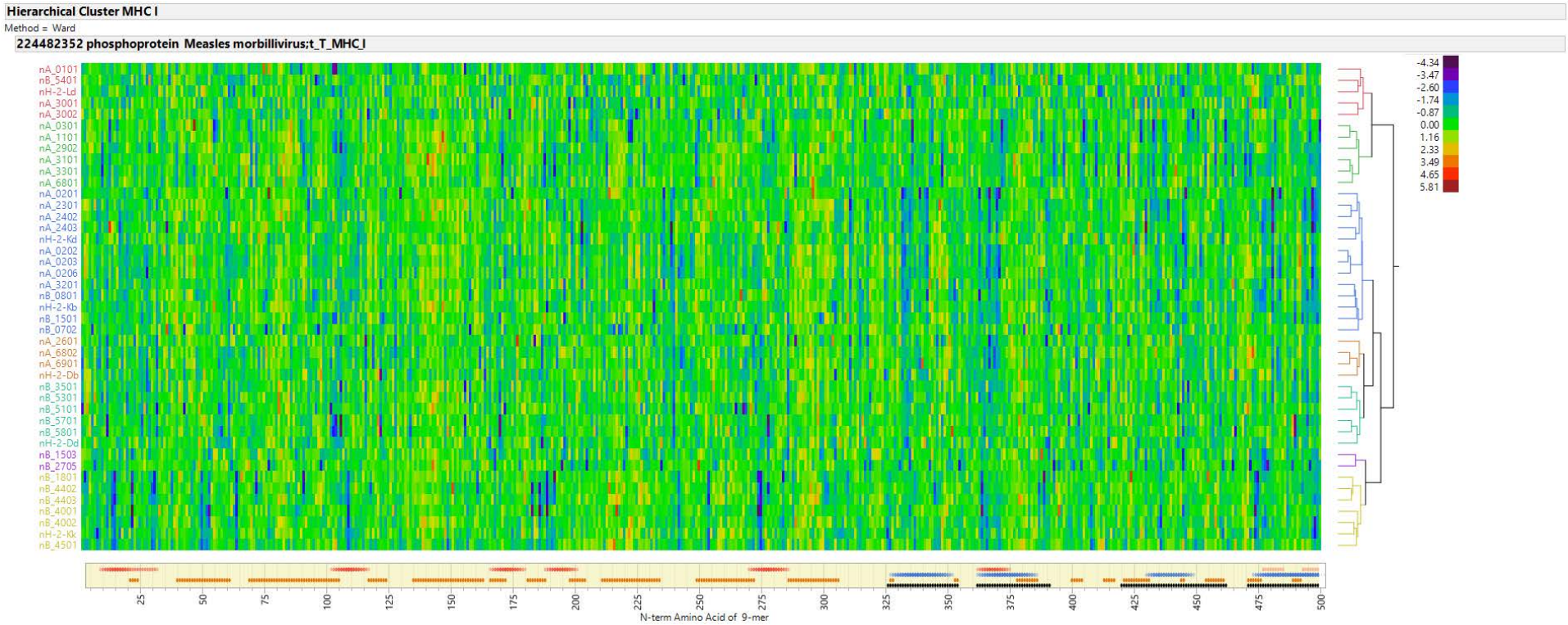


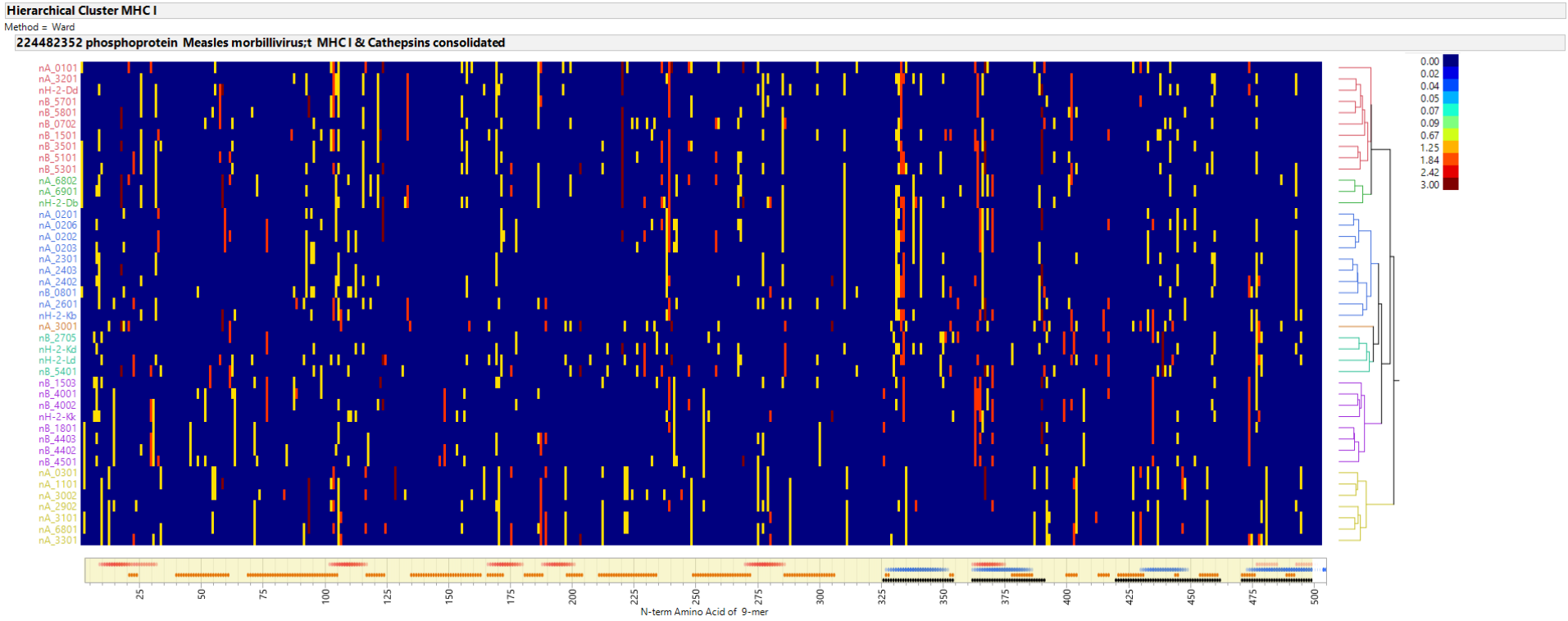

Hierarchical Cluster MHC II
Method = Ward
224482352 phosphoprotein Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0404
nDRB1_0802
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_0301
nDRB3_0101
nDRB1_0701
nDRB1_1302
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.64
-2.91
-2.18
-1.45
-0.73
0.00
0.79
1.58
2.37
3.16
3.95
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

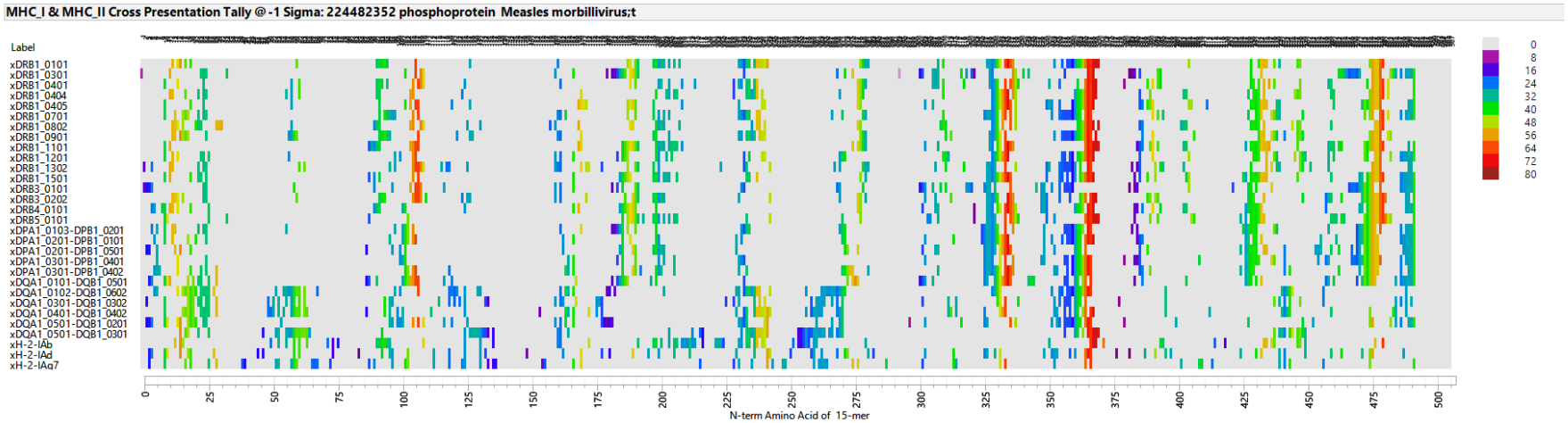

224482352 phosphoprotein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S


224482352 phosphoprotein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L

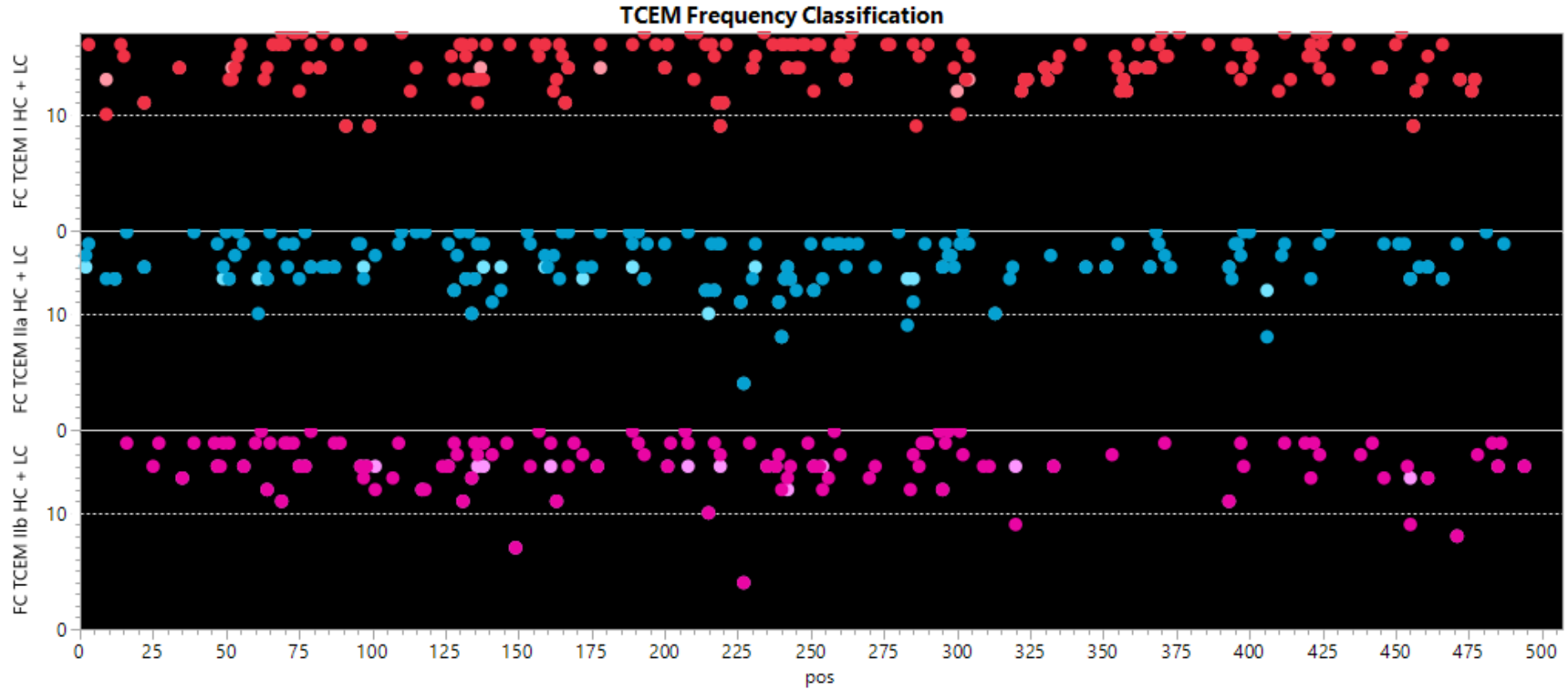

224482352 phosphoprotein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
224482352 phosphoprotein Measles morbillivirus;t_T_MHC_I
nA_0101
nB_5401
nH-2-Ld
nA_3001
nA_3002
nA_0301
nA_1101
nA_2902
nA_3101
nA_3301
nA_6801
nA_0201
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_0202
nA_0203
nA_0206
nA_3201
nB_0801
nH-2-Kb
nB_1501
nB_0702
nA_2601
nA_6802
nA_6901
nH-2-Db
nB_3501
nB_5301
nB_5101
nB_5701
nB_5801
nH-2-Dd
nB_1503
nB_2705
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
-4.34
-3.47
-2.60
-1.74
-0.87
0.00
1.16
2.33
3.49
4.65
5.81
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer

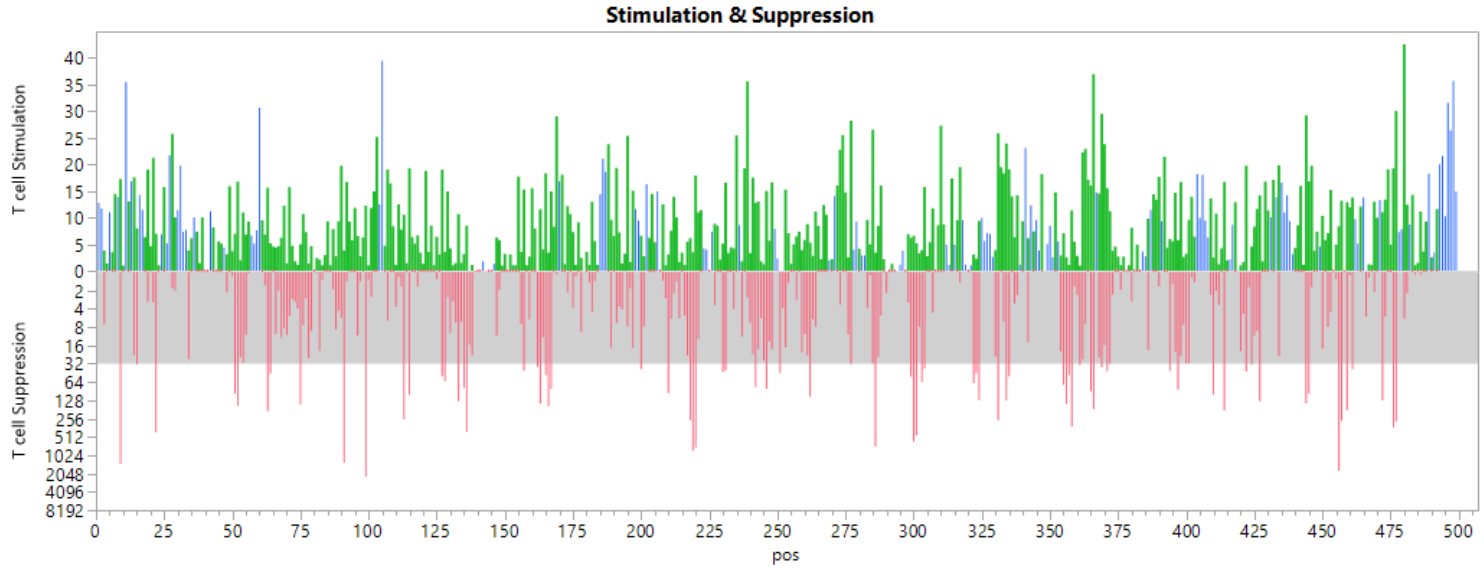

Hierarchical Cluster MHC I
Method = Ward
224482352 phosphoprotein Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nA_3201
nH-2-Dd
nB_5701
nB_5801
nB_0702
nB_1501
nB_3501
nB_5101
nB_5301
nA_6802
nA_6901
nH-2-Db
nA_0201
nA_0206
nA_0202
nA_0203
nA_2301
nA_2403
nA_2402
nB_0801
nA_2601
nH-2-Kb
nA_3001
nB_2705
nH-2-Kd
nH-2-Ld
nB_5401
nB_1503
nB_4001
nB_4002
nH-2-Kk
nB_1801
nB_4403
nB_4402
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_6801
nA_3301
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482352 phosphoprotein Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500
N-term Amino Acid of 15-mer
0 8 16 24 32 40 48 56 64 72 80

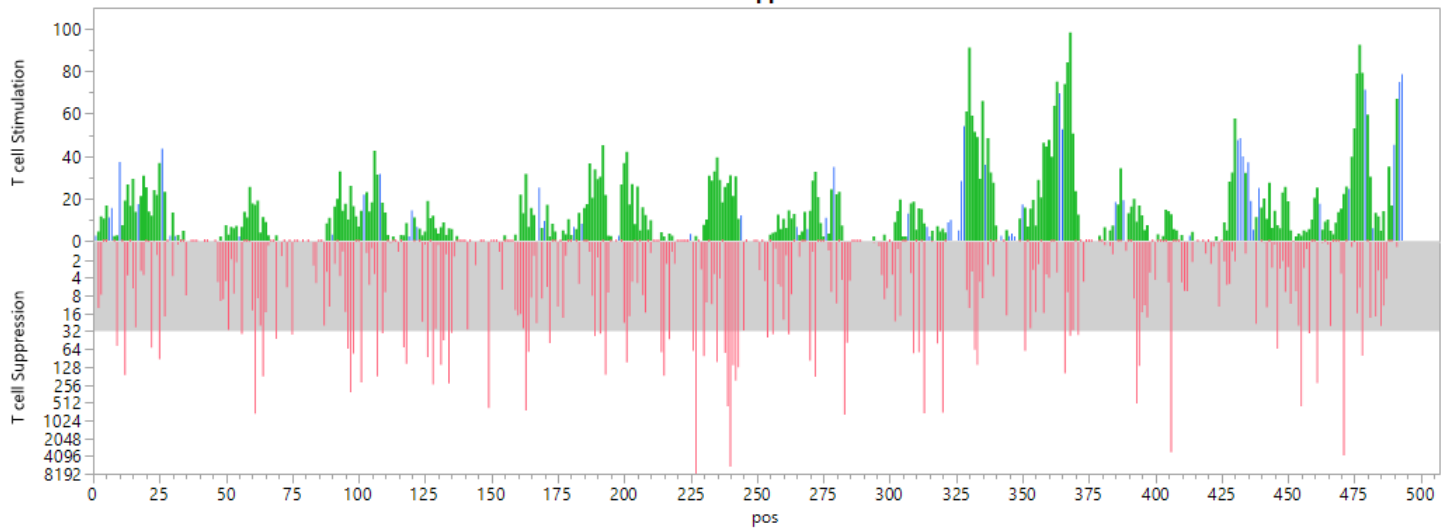
224482352 phosphoprotein Measles morbillivirus



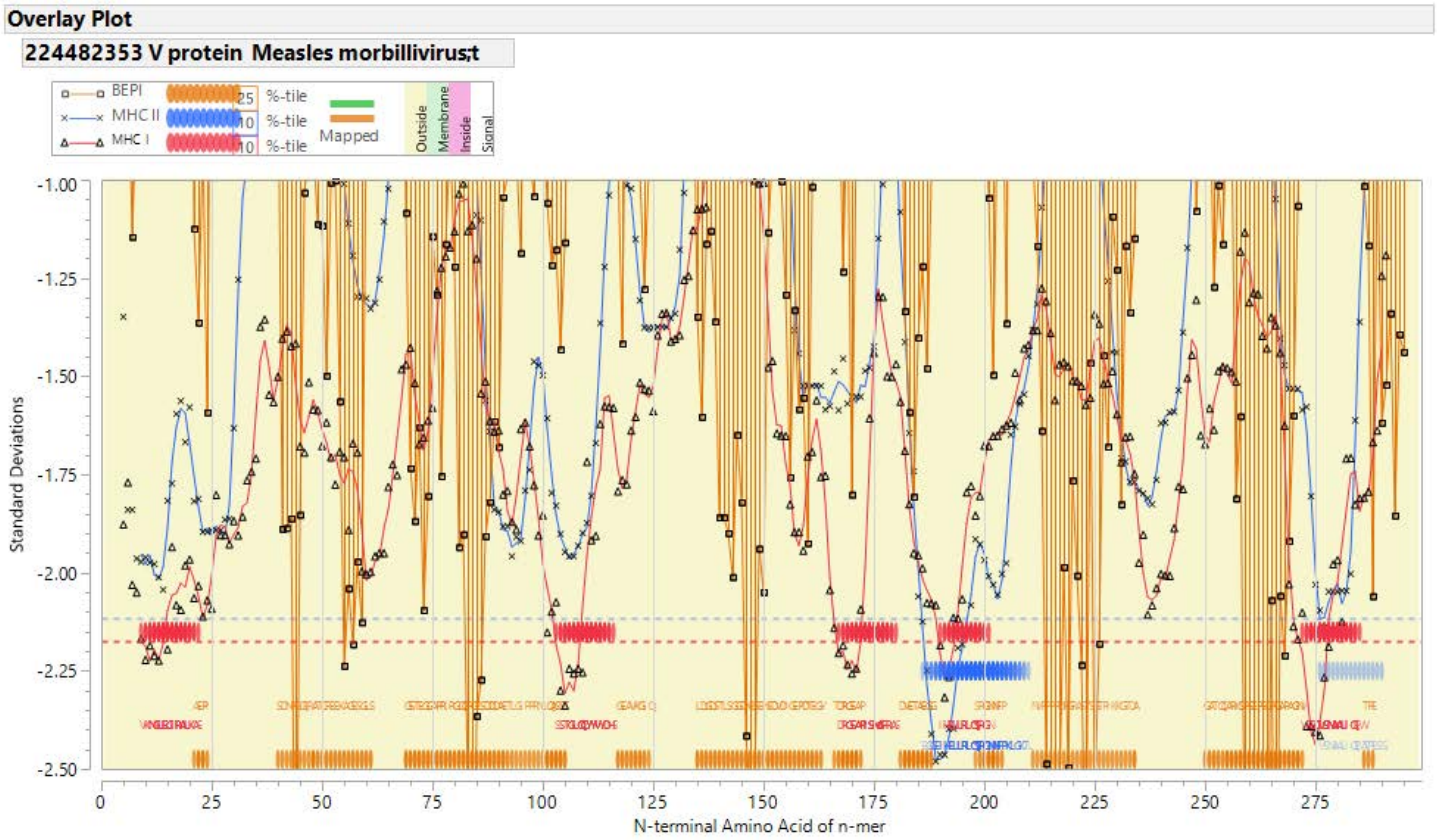
224482352 phosphoprotein Measles morbillivirus

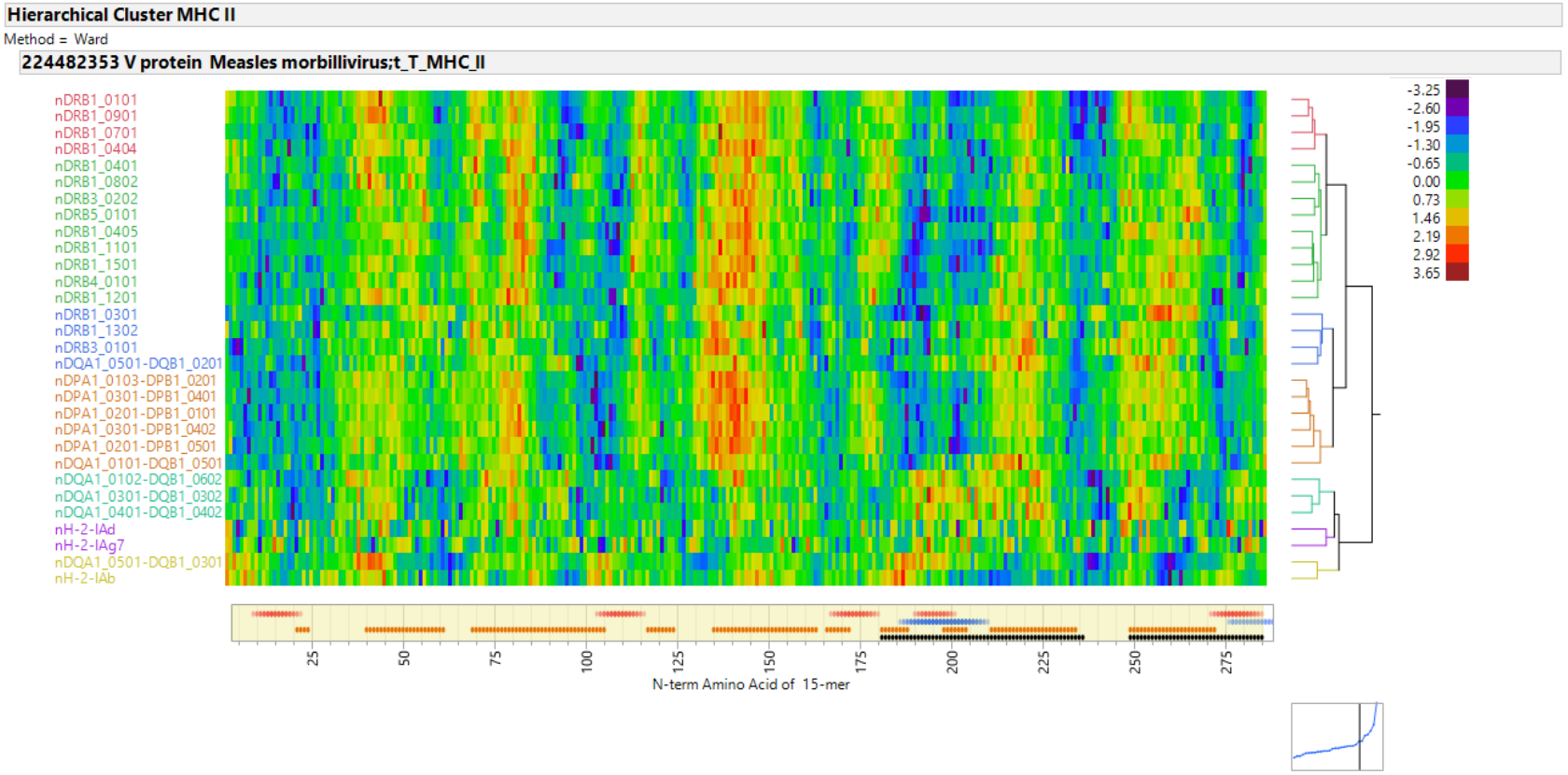


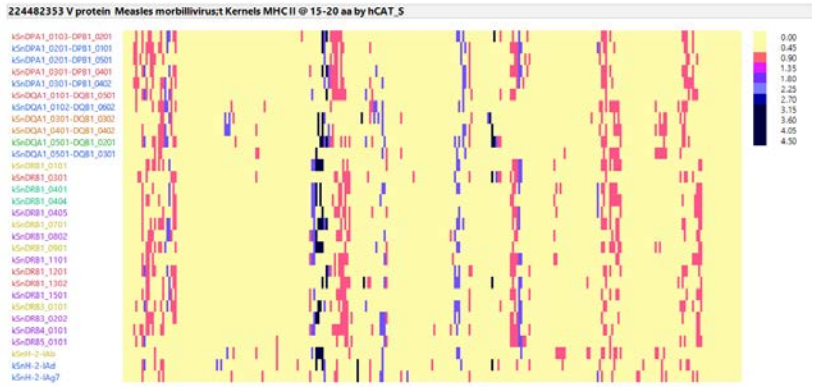
224482352 phosphoprotein Measles morbillivirus

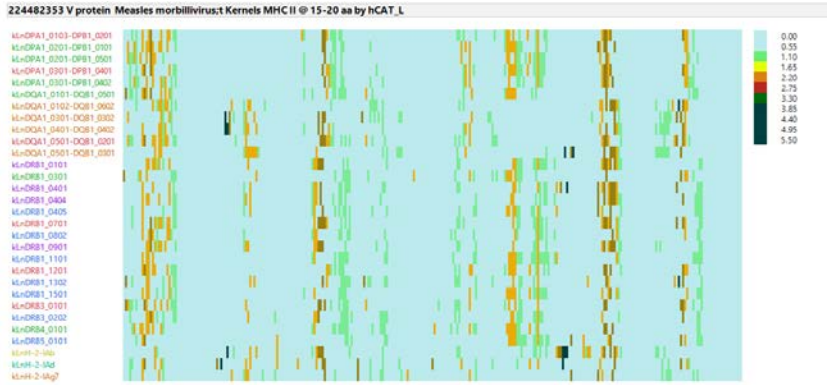


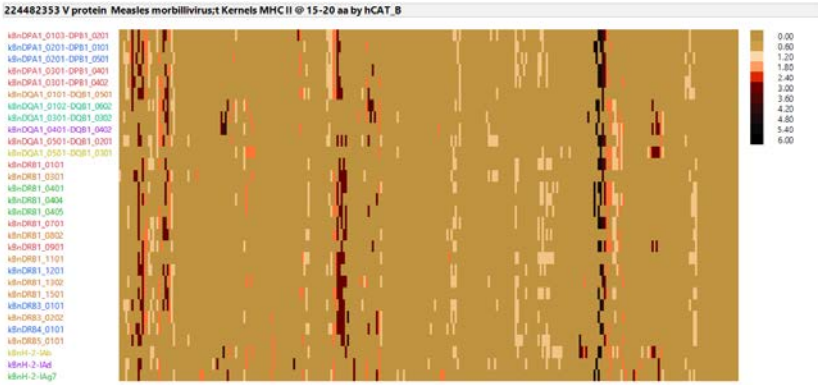
V protein

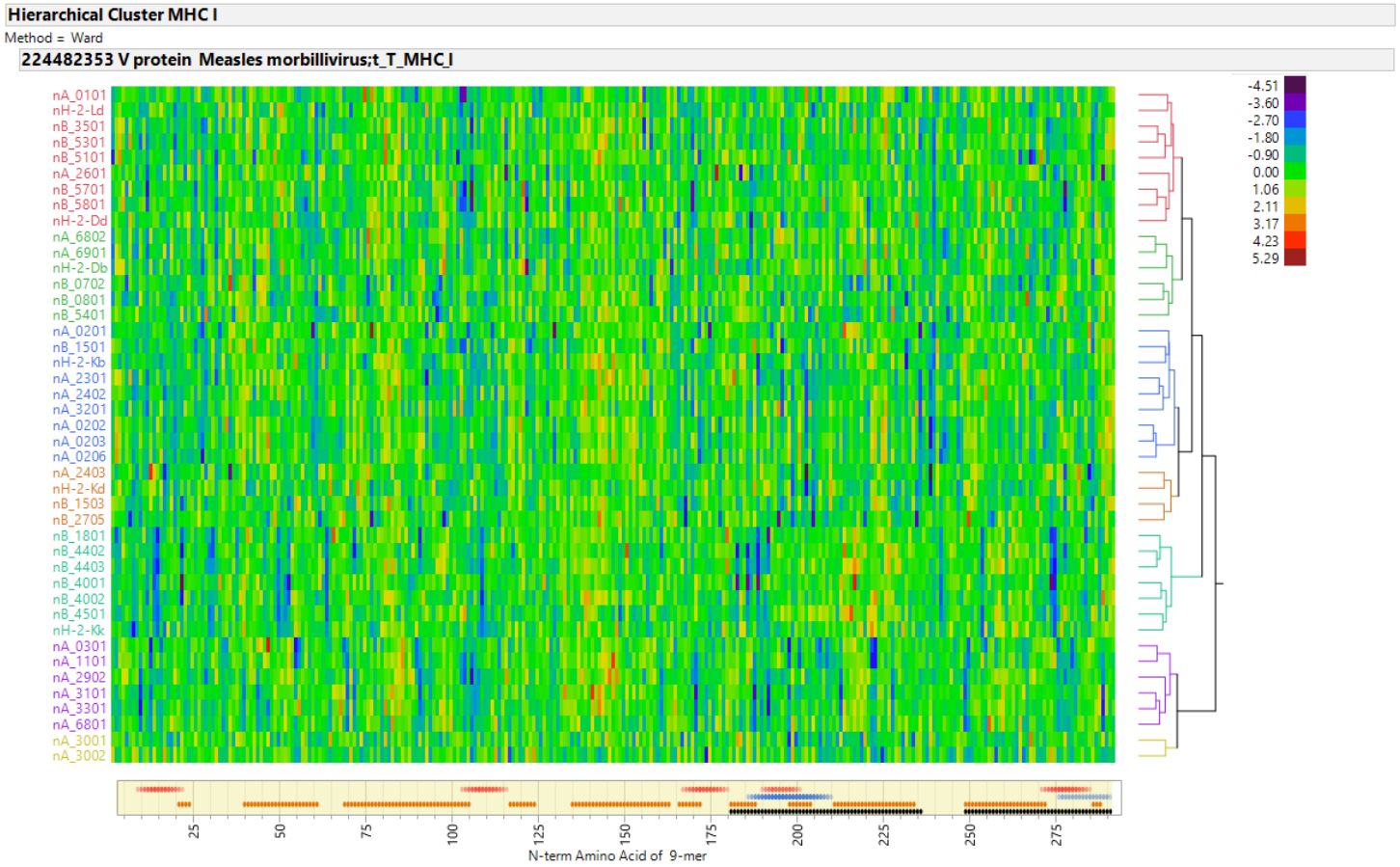

Overlay Plot
224482353 V protein Measles morbillivirus;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
N-terminal Amino Acid of n-mer

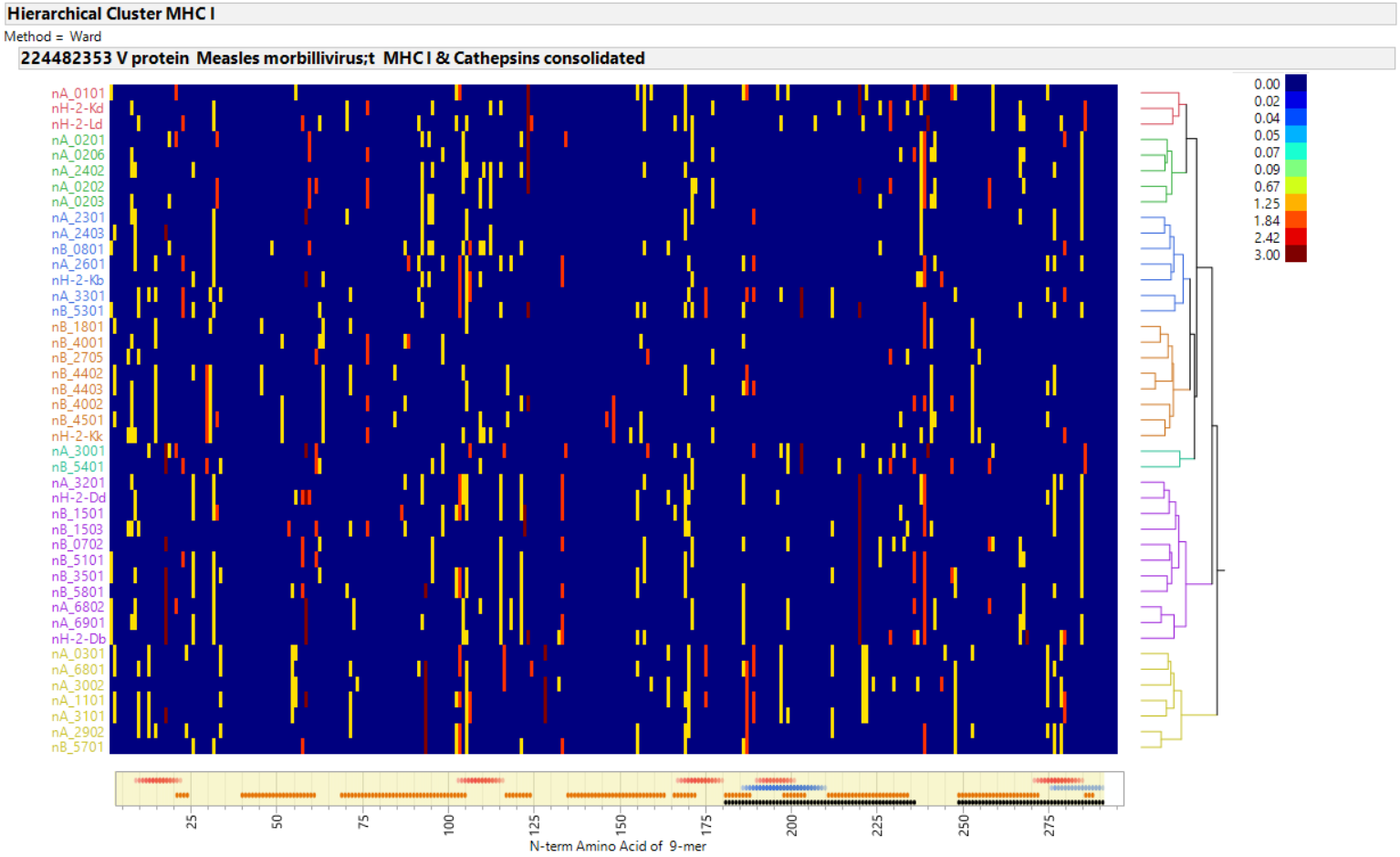

Hierarchical Cluster MHC II
Method = Ward
224482353 V protein Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_0401
nDRB1_0802
nDRB3_0202
nDRB5_0101
nDRB1_0405
nDRB1_1101
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_0301
nDRB1_1302
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.25
-2.60
-1.95
-1.30
-0.65
0.00
0.73
1.46
2.19
2.92
3.65
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer

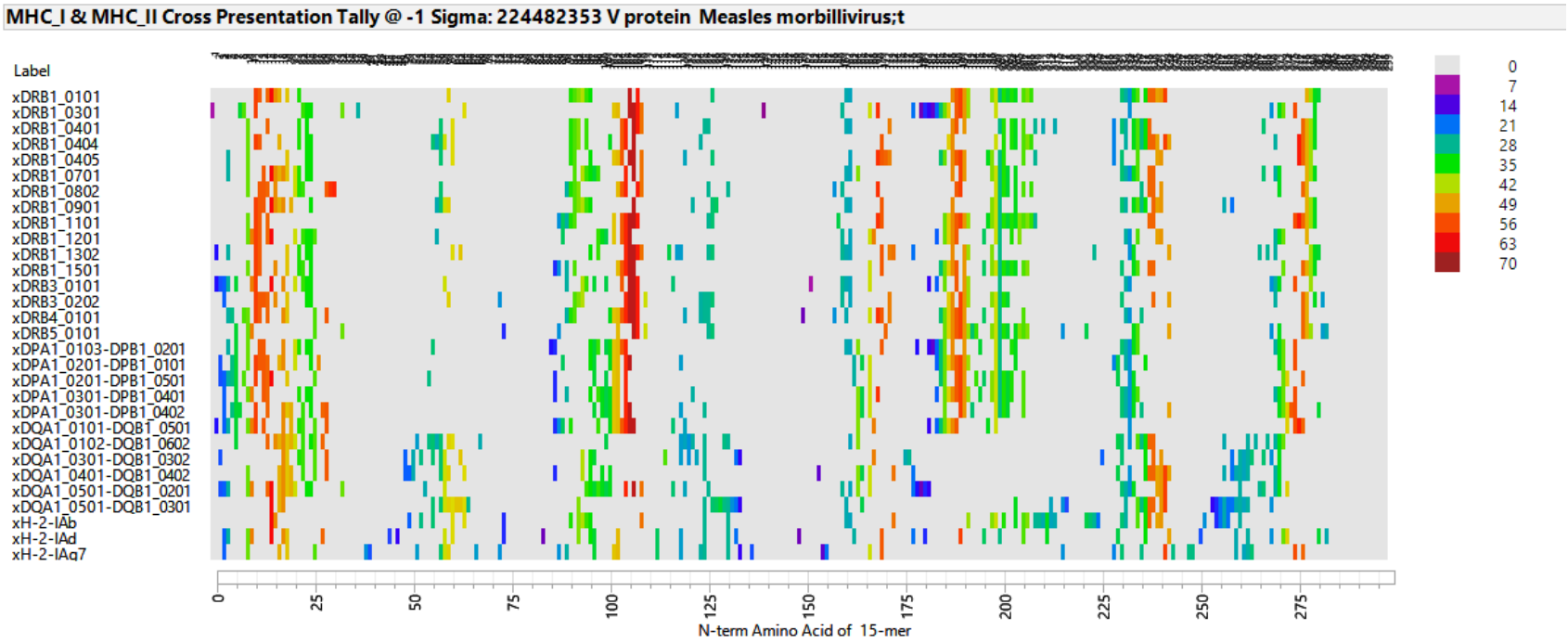

224482353 V protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S


224482353 V protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L

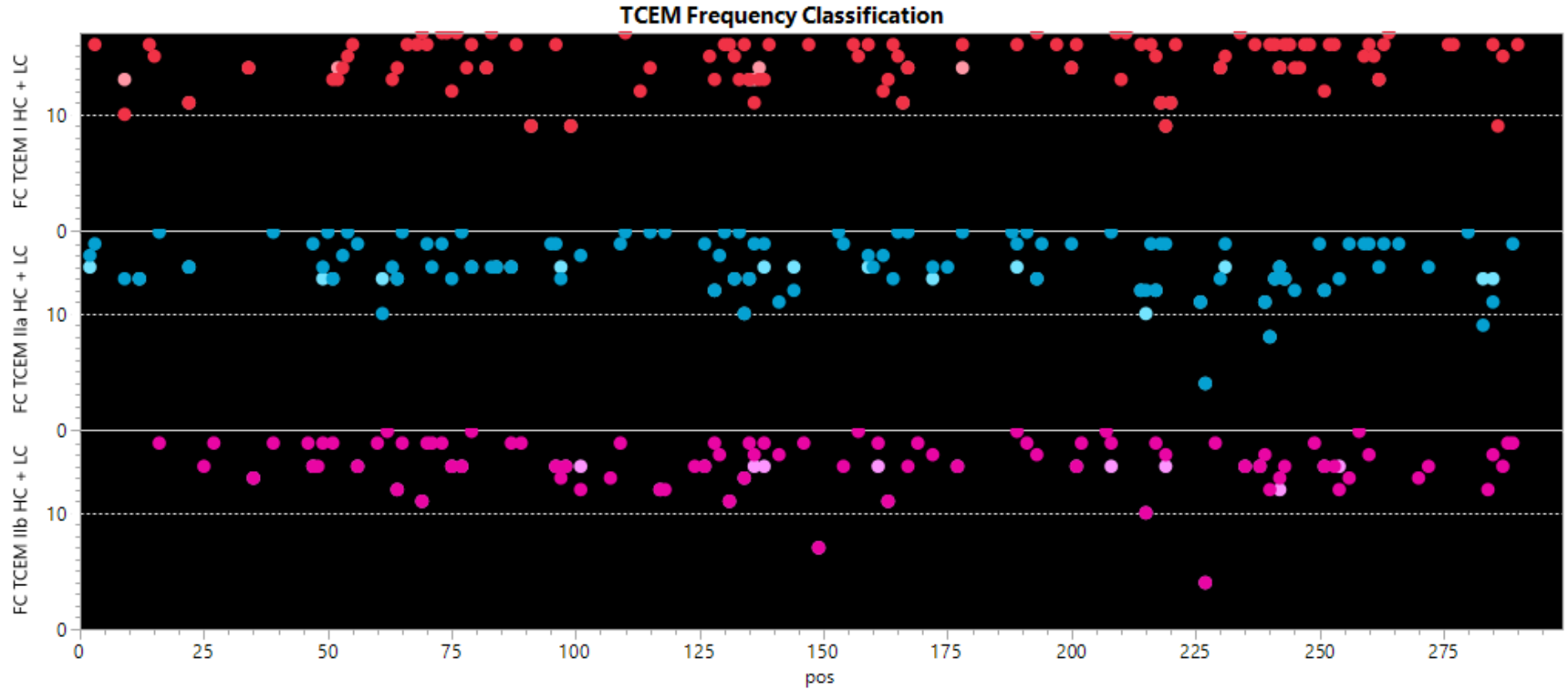

224482353 V protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
224482353 V protein Measles morbillivirus;t_T_MHC_I
nA_0101
nH-2-Ld
nB_3501
nB_5301
nB_5101
nA_2601
nB_5701
nB_5801
nH-2-Dd
nA_6802
nA_6901
nH-2-Db
nB_0702
nB_0801
nB_5401
nA_0201
nB_1501
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_0202
nA_0203
nA_0206
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nH-2-Kk
nA_0301
nA_1101
nA_2902
nA_3101
nA_3301
nA_6801
nA_3001
nA_3002
-4.51
-3.60
-2.70
-1.80
-0.90
0.00
1.06
2.11
3.17
4.23
5.29
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer

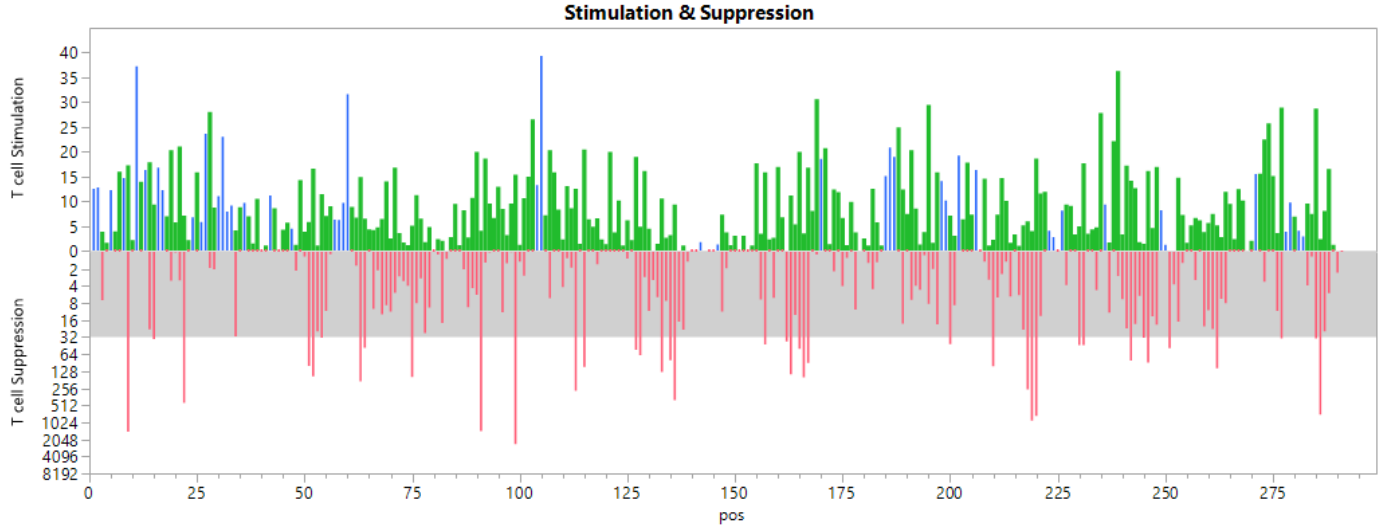

Hierarchical Cluster MHC I
Method = Ward
224482353 V protein Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nH-2-Kd
nH-2-Ld
nA_0201
nA_0206
nA_2402
nA_0202
nA_0203
nA_2301
nA_2403
nB_0801
nA_2601
nH-2-Kb
nA_3301
nB_5301
nB_1801
nB_4001
nB_2705
nB_4402
nB_4403
nB_4002
nB_4501
nH-2-Kk
nA_3001
nB_5401
nA_3201
nH-2-Dd
nB_1501
nB_1503
nB_0702
nB_5101
nB_3501
nB_5801
nA_6802
nA_6901
nH-2-Db
nA_0301
nA_6801
nA_3002
nA_1101
nA_3101
nA_2902
nB_5701
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482353 V protein Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
7
14
21
28
35
42
49
56
63
70
0
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer

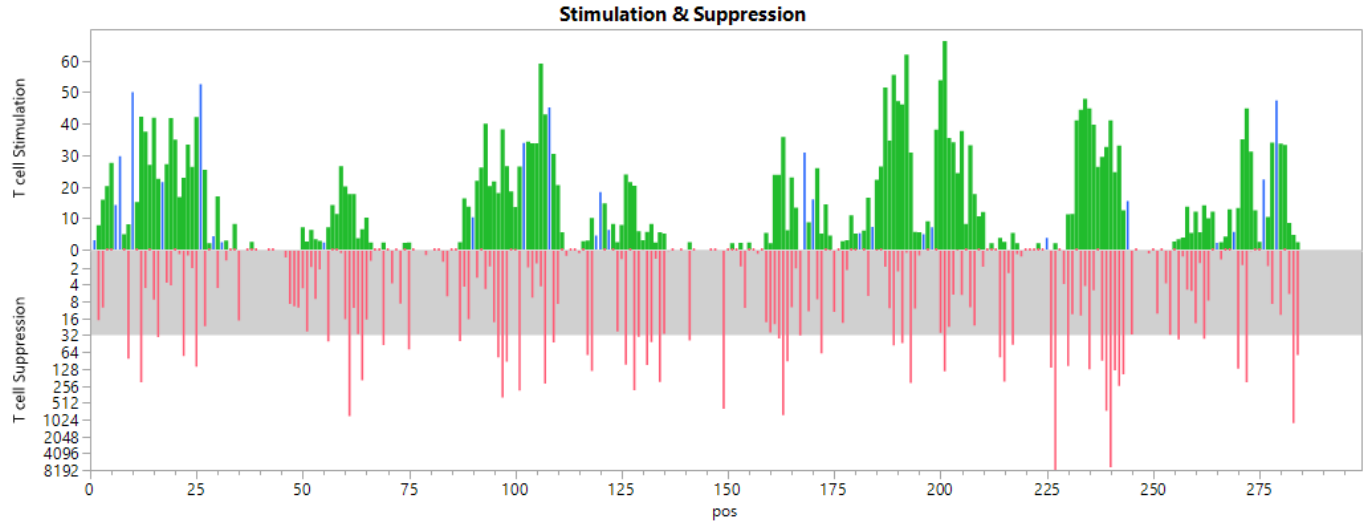
224482353 V protein Measles morbillivirus



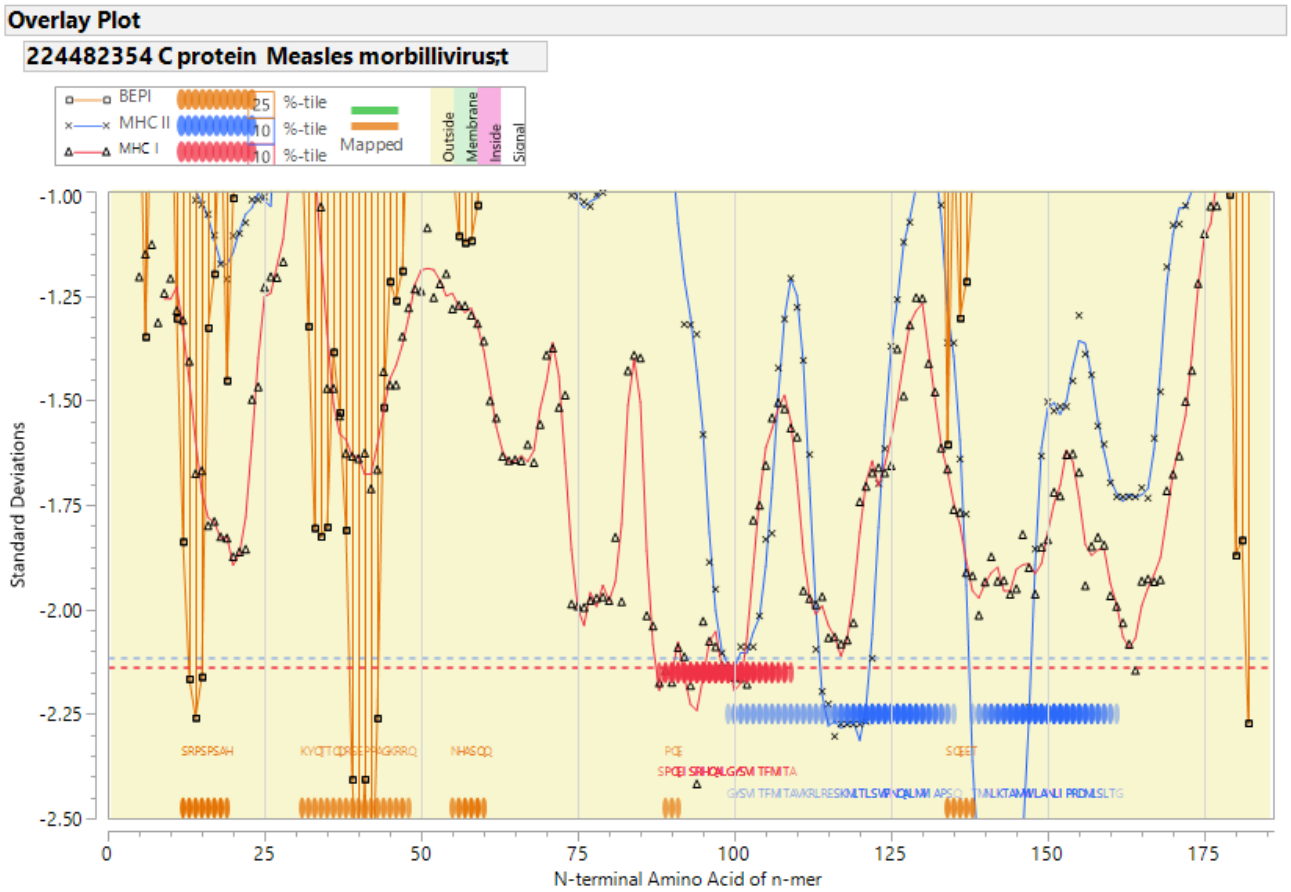
224482353 V protein Measles morbillivirus

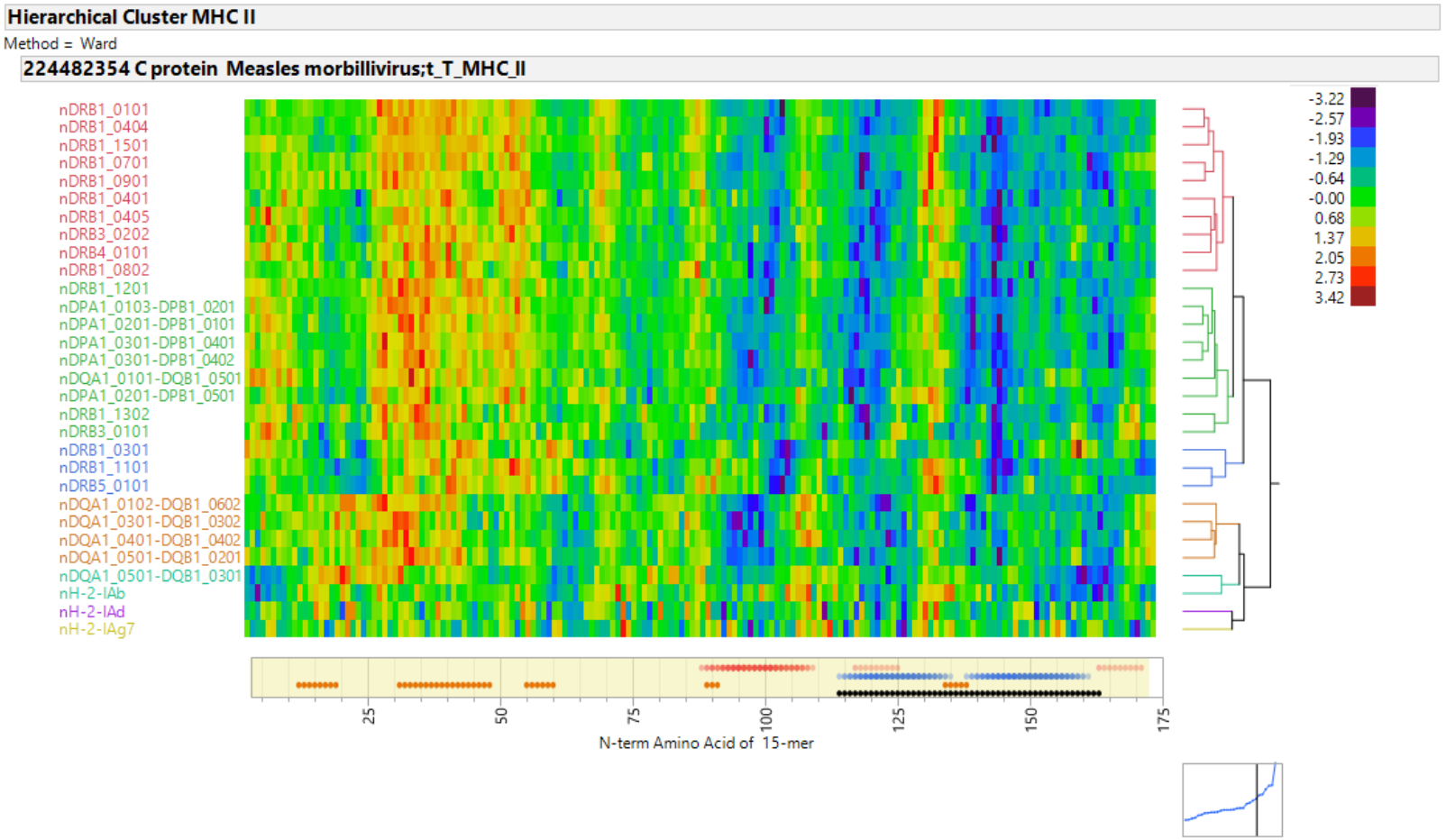


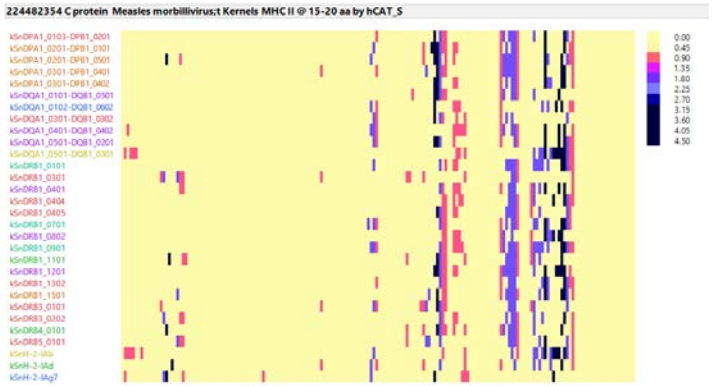
224482353 V protein Measles morbillivirus

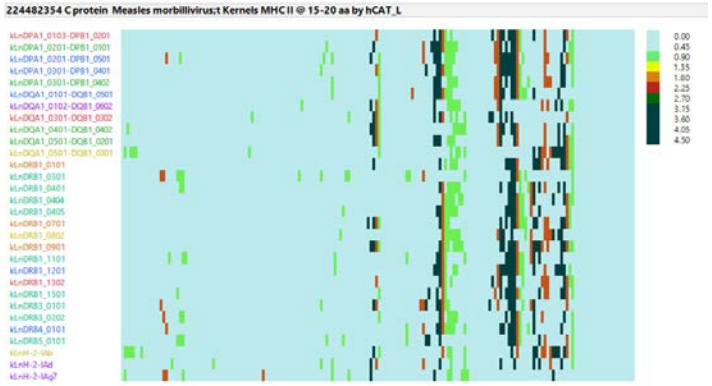


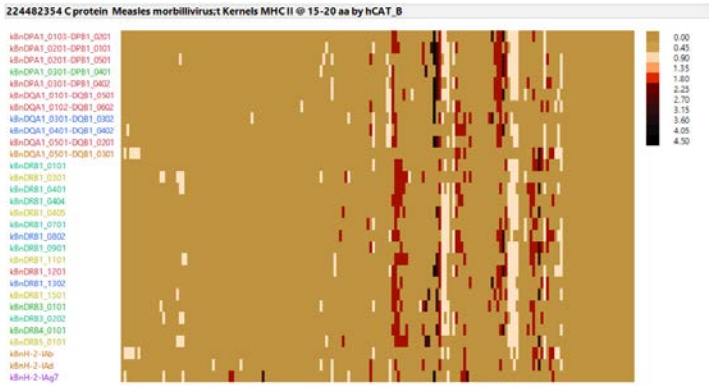
C protein

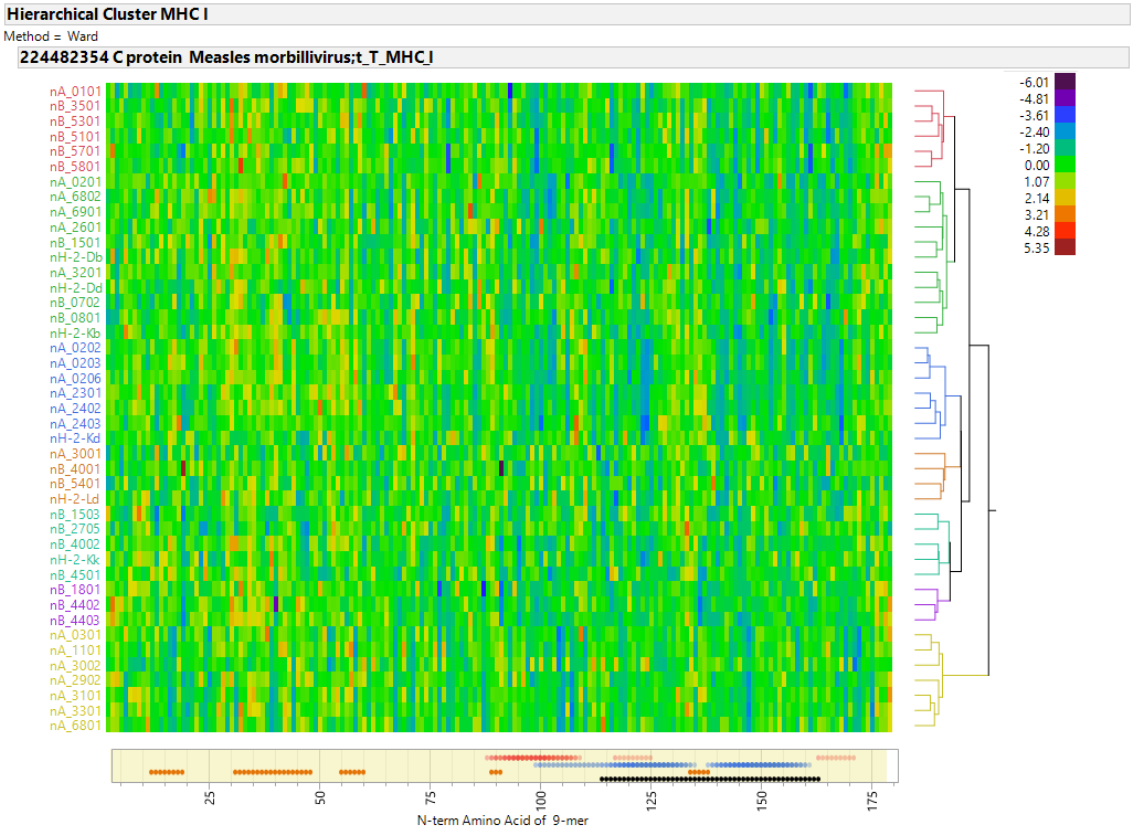

Overlay Plot
224482354 C protein Measles morbillivirus;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
N-terminal Amino Acid of n-mer

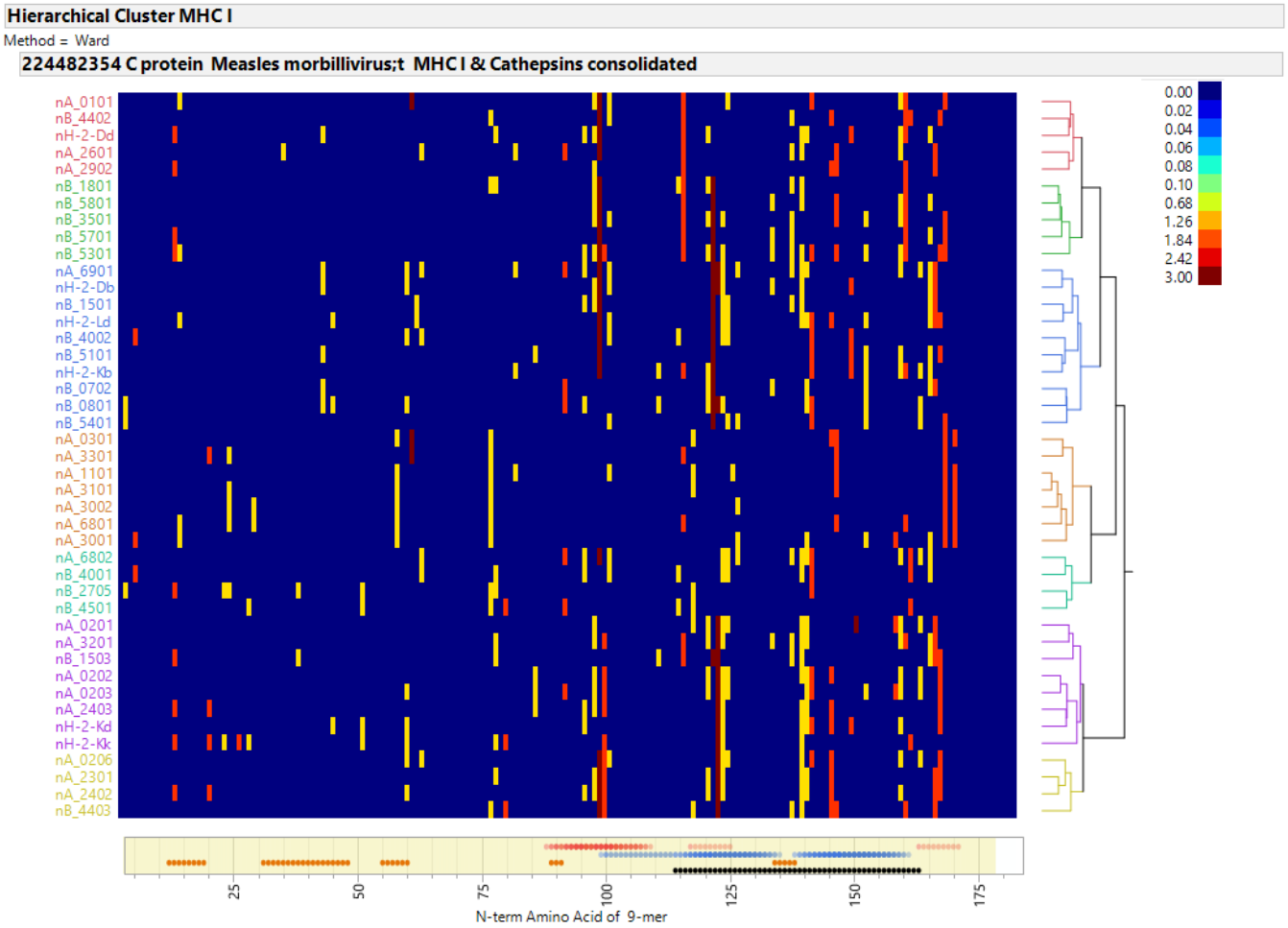

Hierarchical Cluster MHC II
Method = Ward
224482354 C protein Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0404
nDRB1_1501
nDRB1_0701
nDRB1_0901
nDRB1_0401
nDRB1_0405
nDRB3_0202
nDRB4_0101
nDRB1_0802
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDQA1_0101-DQB1_0501
nDPA1_0201-DPB1_0501
nDRB1_1302
nDRB3_0101
nDRB1_0301
nDRB1_1101
nDRB5_0101
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.22
-2.57
-1.93
-1.29
-0.64
-0.00
0.68
1.37
2.05
2.73
3.42
25
50
75
100
125
150
175
N-term Amino Acid of 15-mer

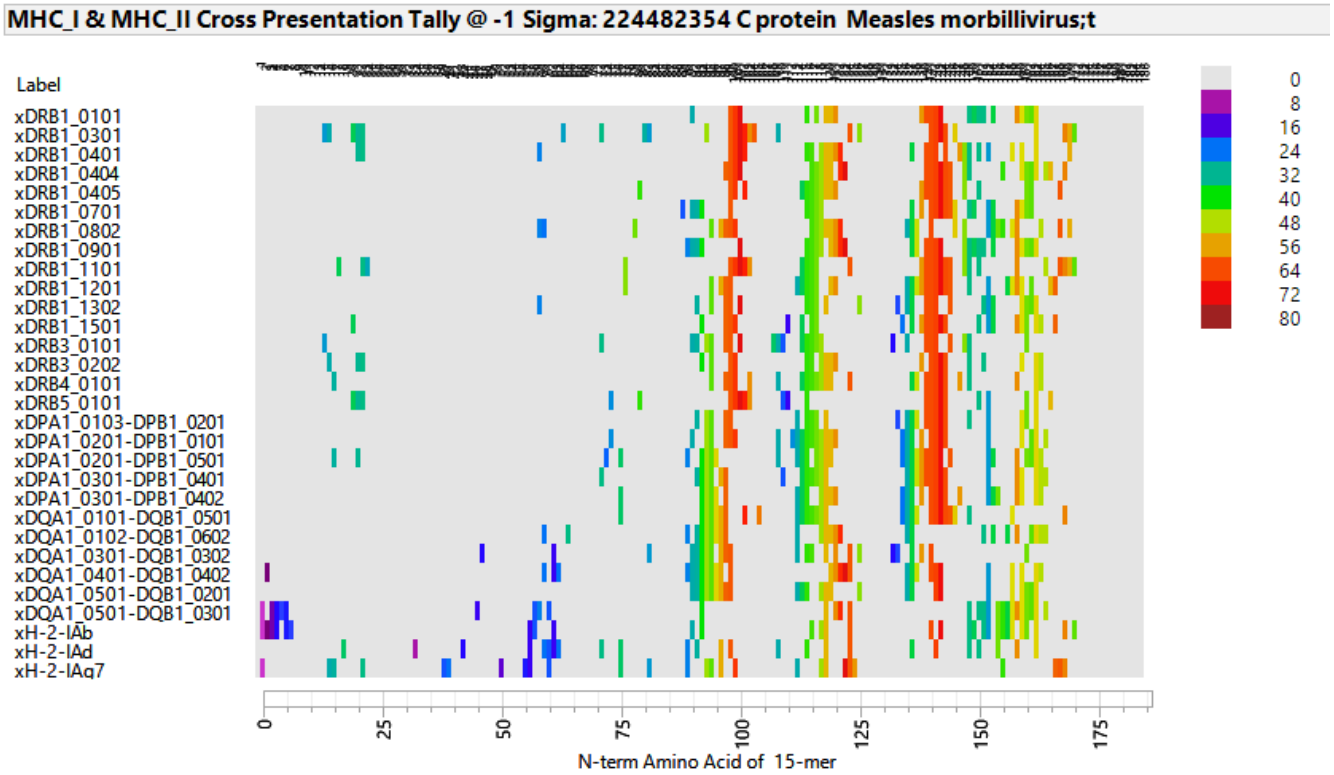

224482354 C protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50


224482354 C protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

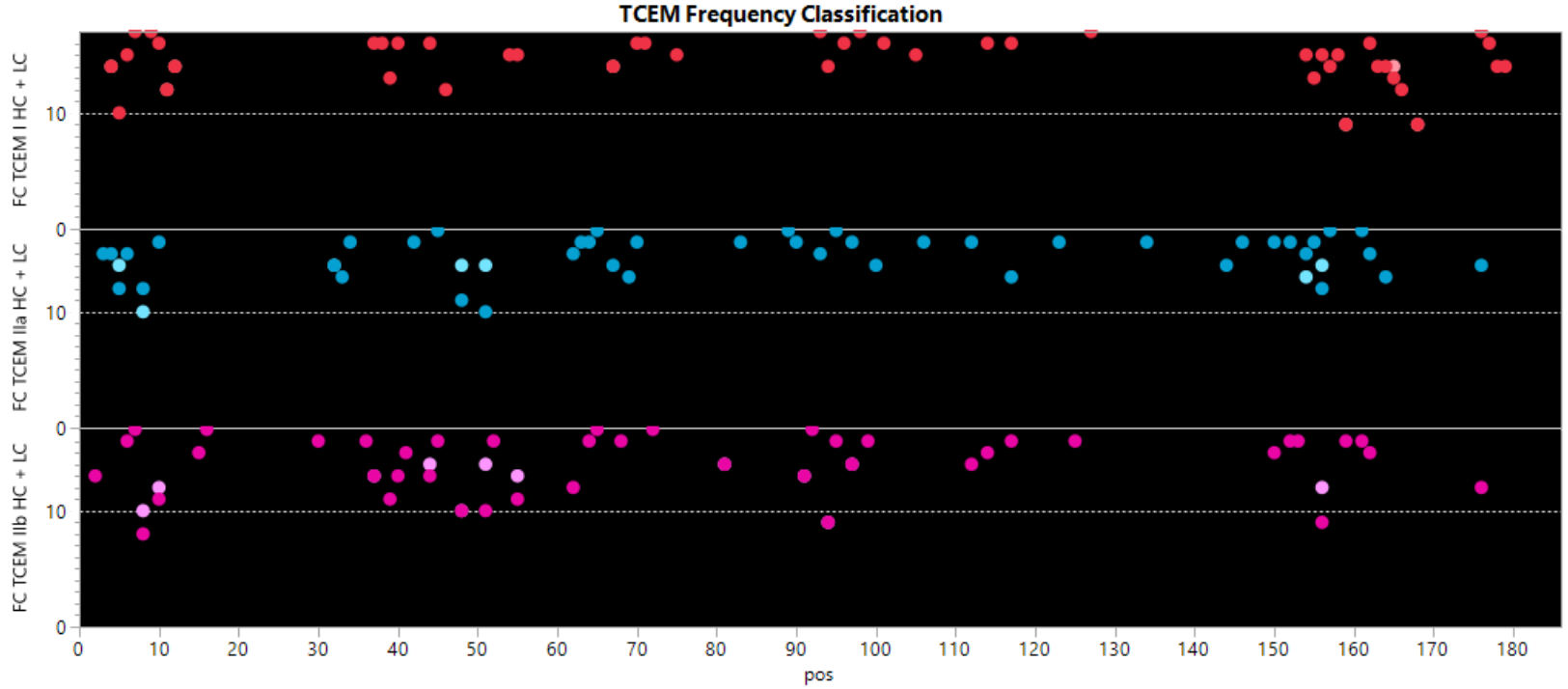

224482354 C protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

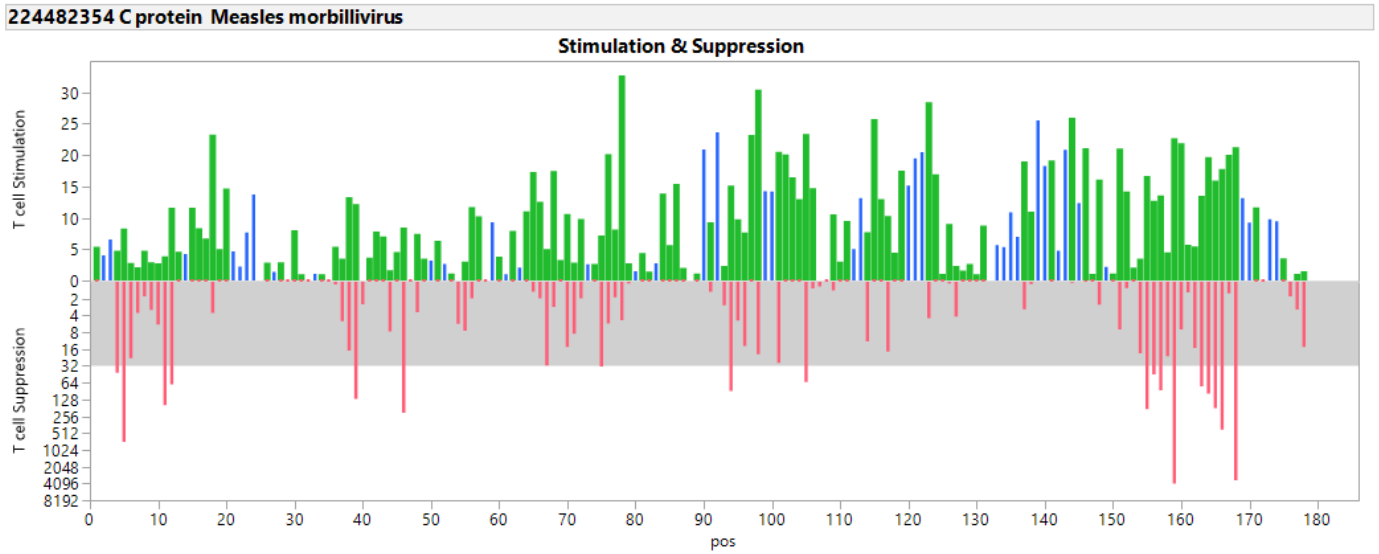

Hierarchical Cluster MHC I
Method = Ward
224482354 C protein Measles morbillivirus;t_T_MHC_I
nA_0101
nB_3501
nB_5301
nB_5101
nB_5701
nB_5801
nA_0201
nA_6802
nA_6901
nA_2601
nB_1501
nH-2-Db
nA_3201
nH-2-Dd
nB_0702
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_3001
nB_4001
nB_5401
nH-2-Ld
nB_1503
nB_2705
nB_4002
nH-2-Kk
nB_4501
nB_1801
nB_4402
nB_4403
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-6.01
-4.81
-3.61
-2.40
-1.20
0.00
1.07
2.14
3.21
4.28
5.35
25
50
75
100
125
150
175
N-term Amino Acid of 9-mer

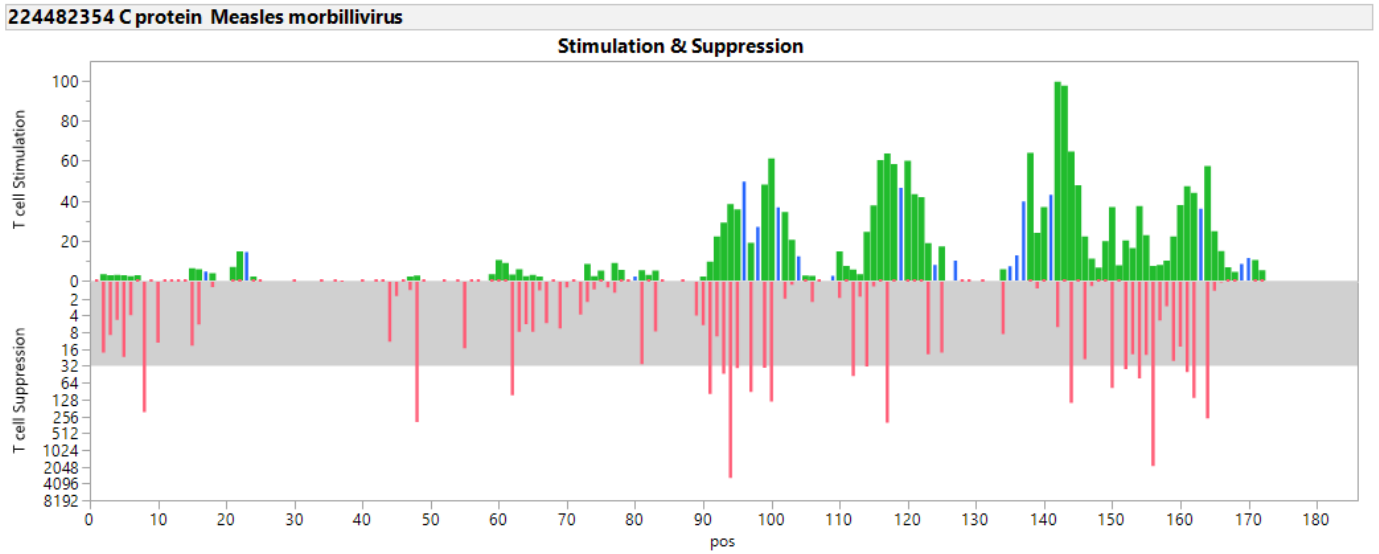

Hierarchical Cluster MHC I
Method = Ward
224482354 C protein Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nB_4402
nH-2-Dd
nA_2601
nA_2902
nB_1801
nB_5801
nB_3501
nB_5701
nB_5301
nA_6901
nH-2-Db
nB_1501
nH-2-Ld
nB_4002
nB_5101
nH-2-Kb
nB_0702
nB_0801
nB_5401
nA_0301
nA_3301
nA_1101
nA_3101
nA_3002
nA_6801
nA_3001
nA_6802
nB_4001
nB_2705
nB_4501
nA_0201
nA_3201
nB_1503
nA_0202
nA_0203
nA_2403
nH-2-Kd
nH-2-Kk
nA_0206
nA_2301
nA_2402
nB_4403
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482354 C protein Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

224482354 C protein Measles morbillivirus



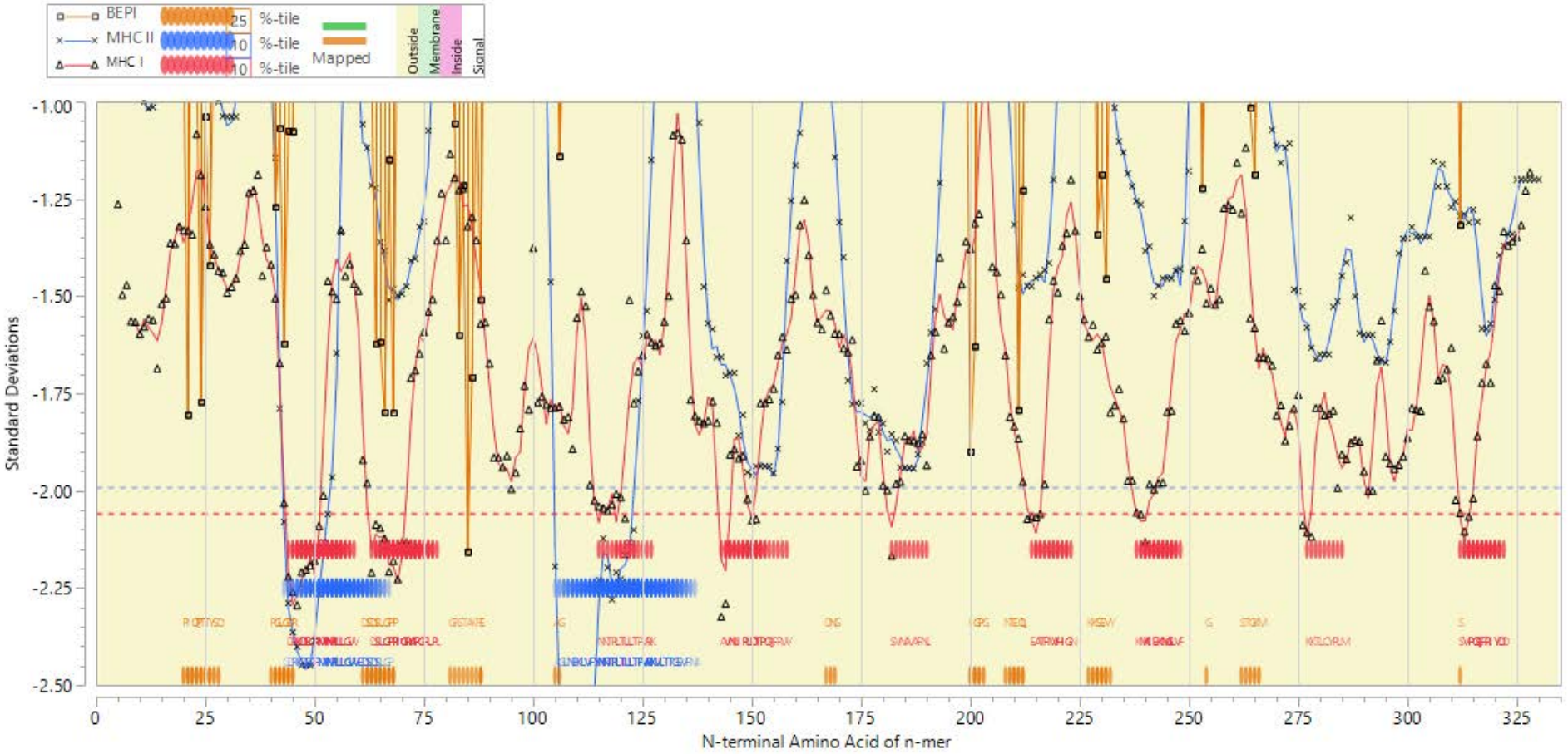

224482354 C protein Measles morbillivirus
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

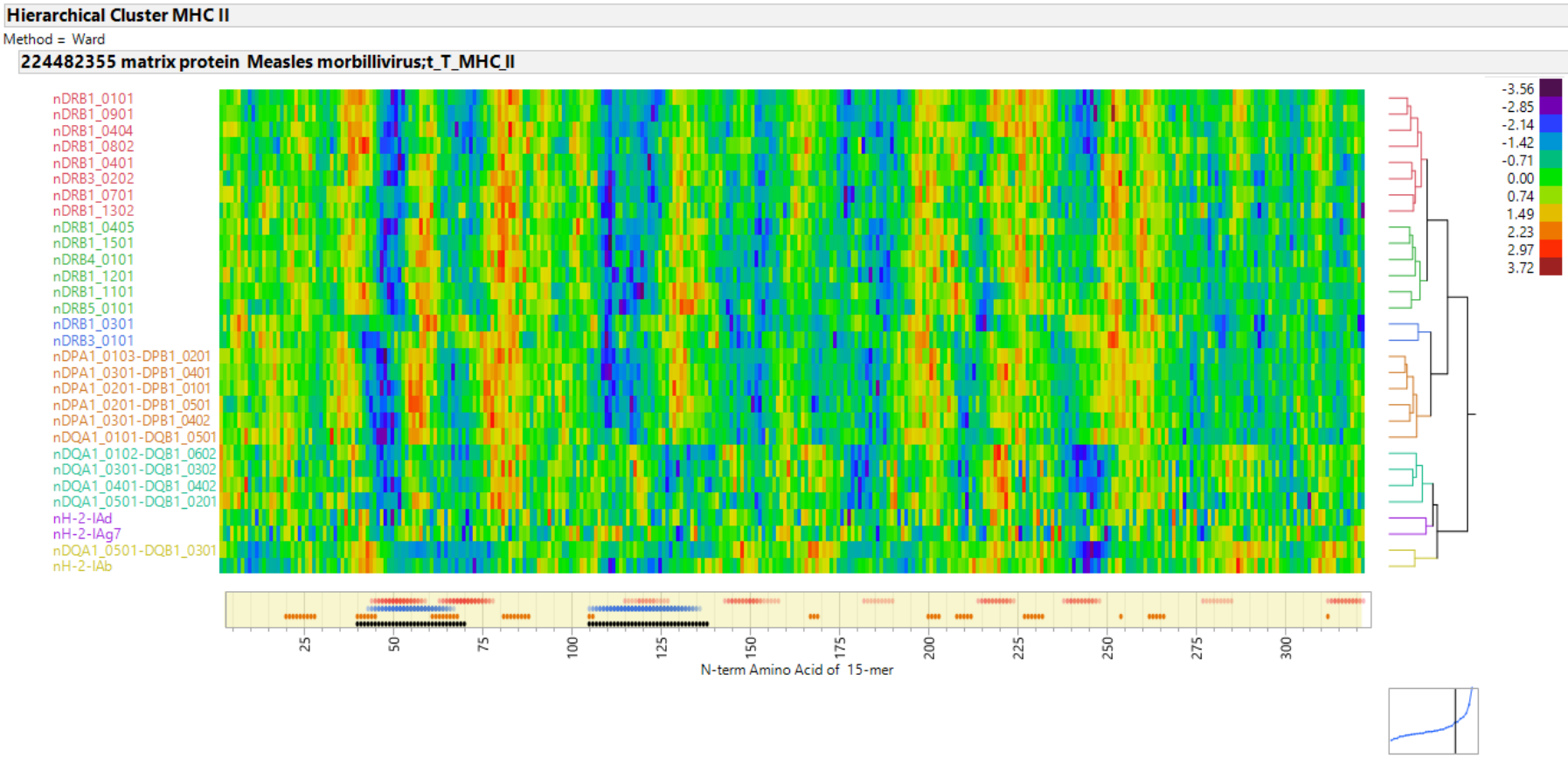

224482354 C protein Measles morbillivirus
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

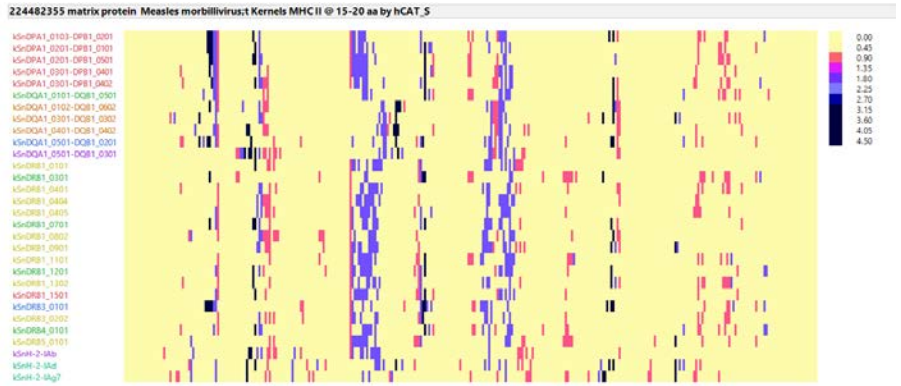
Matrix protein

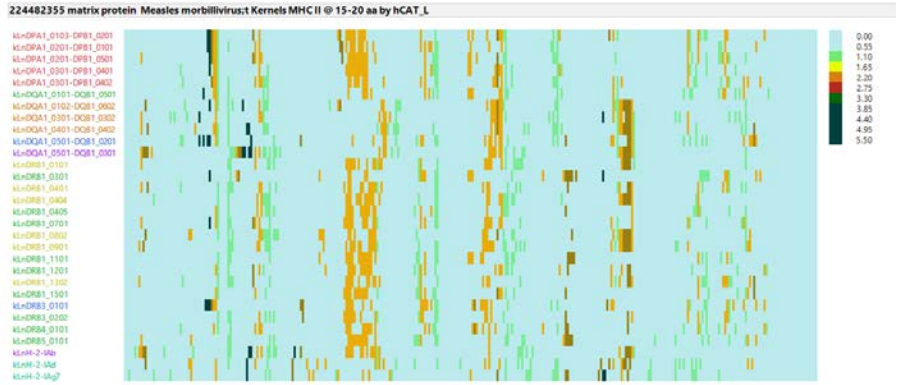
Overlay Plot

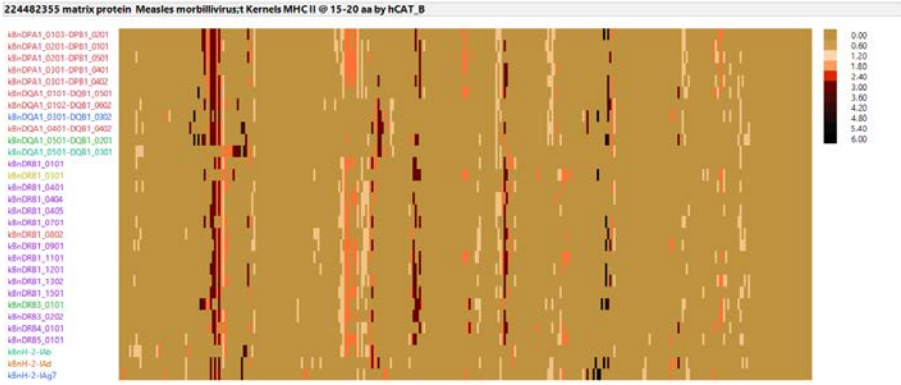
224482355 matrix protein Measles morbillivirus;t

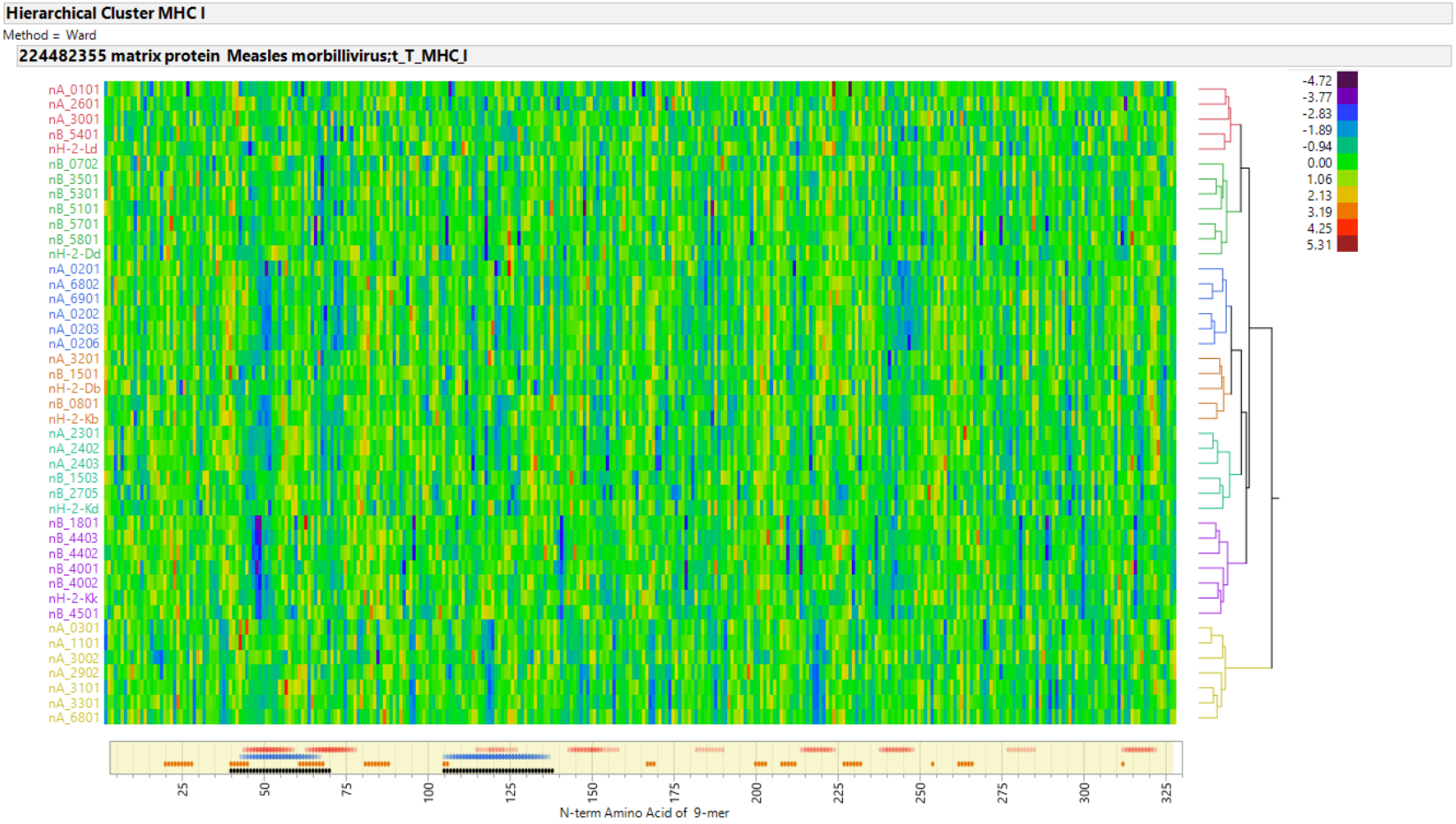


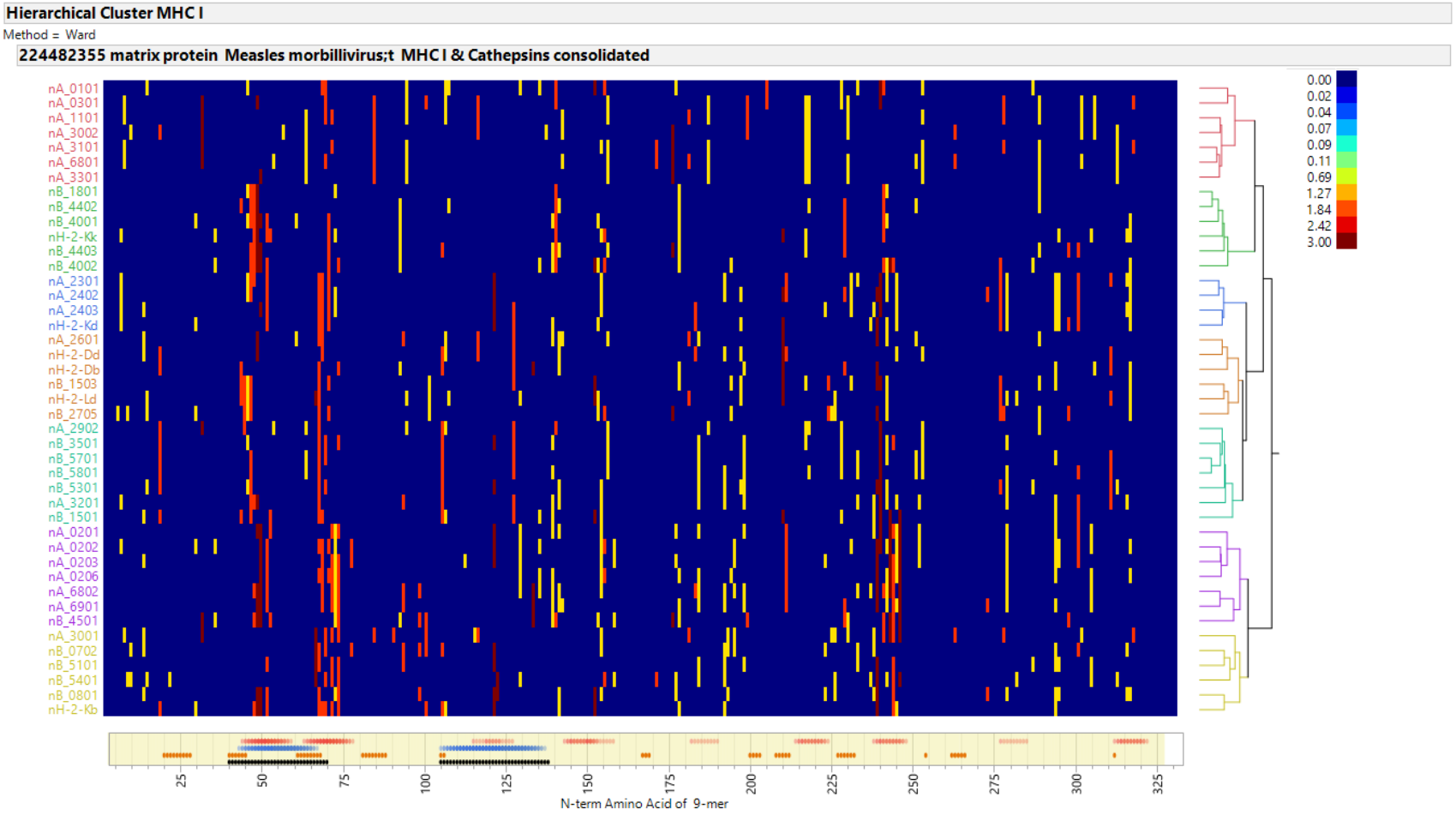

Hierarchical Cluster MHC II
Method = Ward
224482355 matrix protein Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0404
nDRB1_0802
nDRB1_0401
nDRB3_0202
nDRB1_0701
nDRB1_1302
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.56
-2.85
-2.14
-1.42
-0.71
0.00
0.74
1.49
2.23
2.97
3.72
25
50
75
100
125
150
175
200
225
250
275
300
N-term Amino Acid of 15-mer

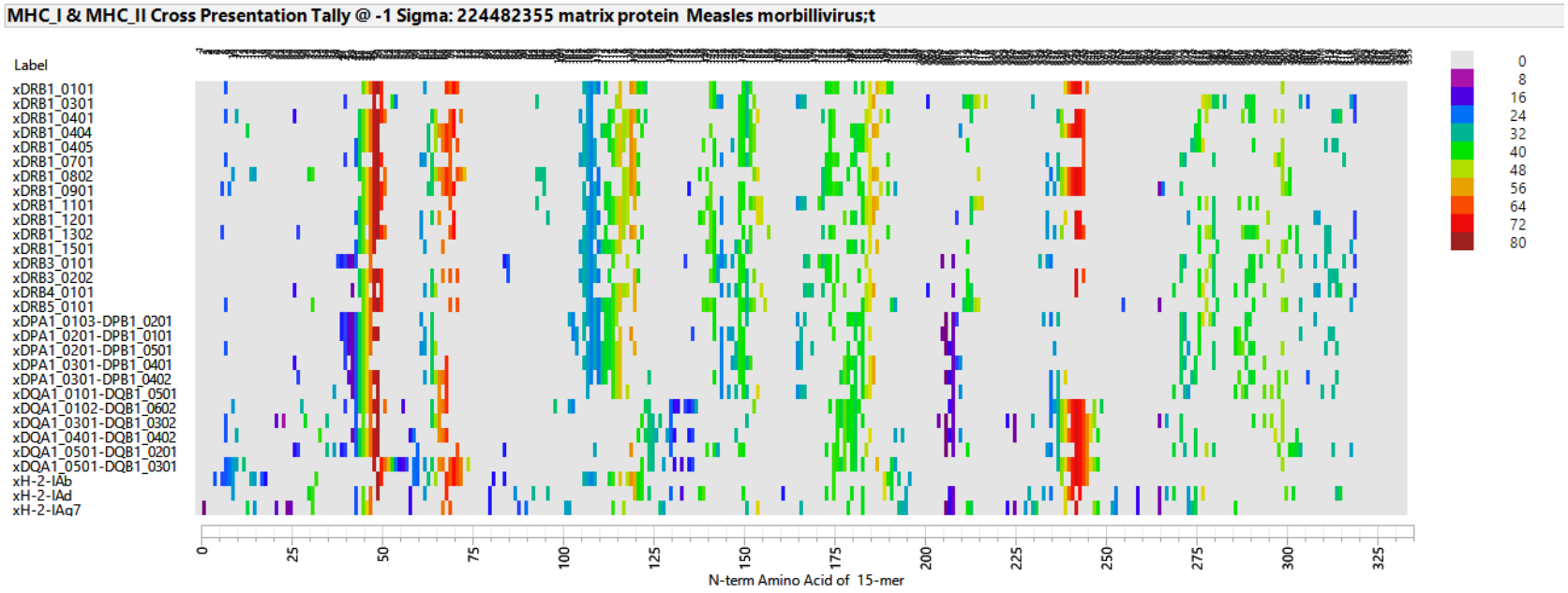

224482355 matrix protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

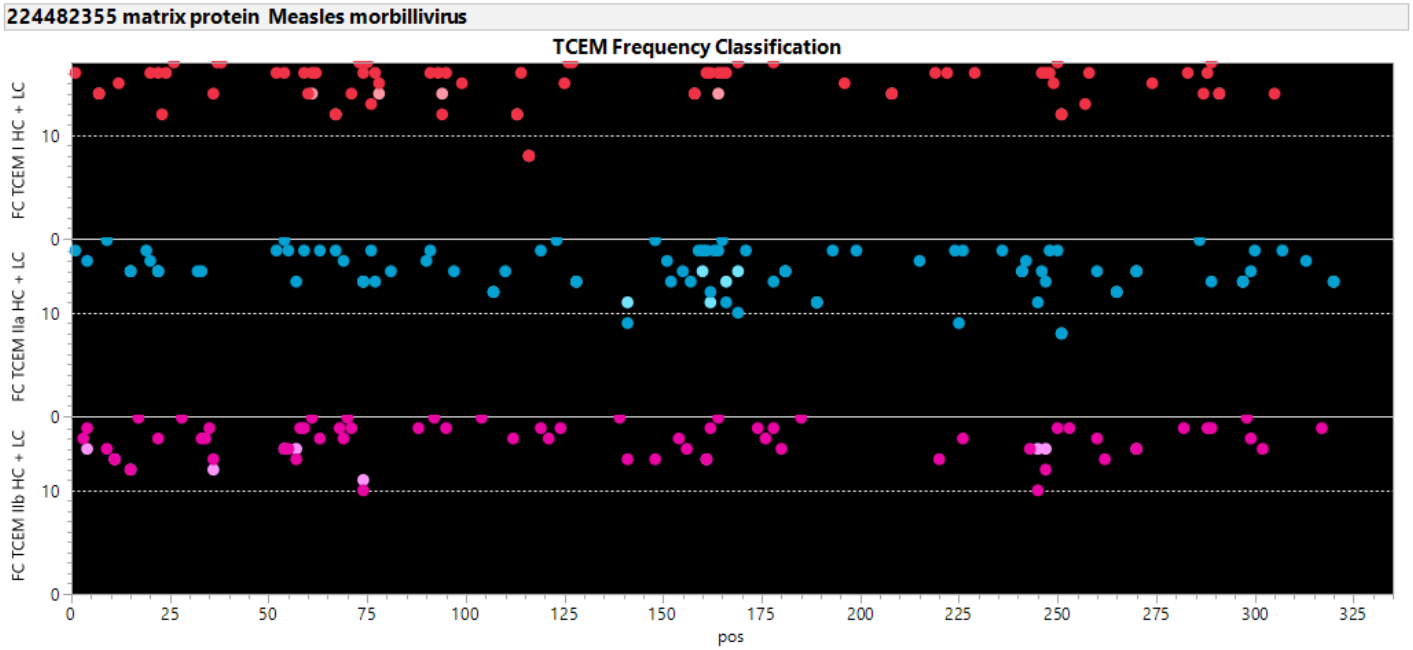

224482355 matrix protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

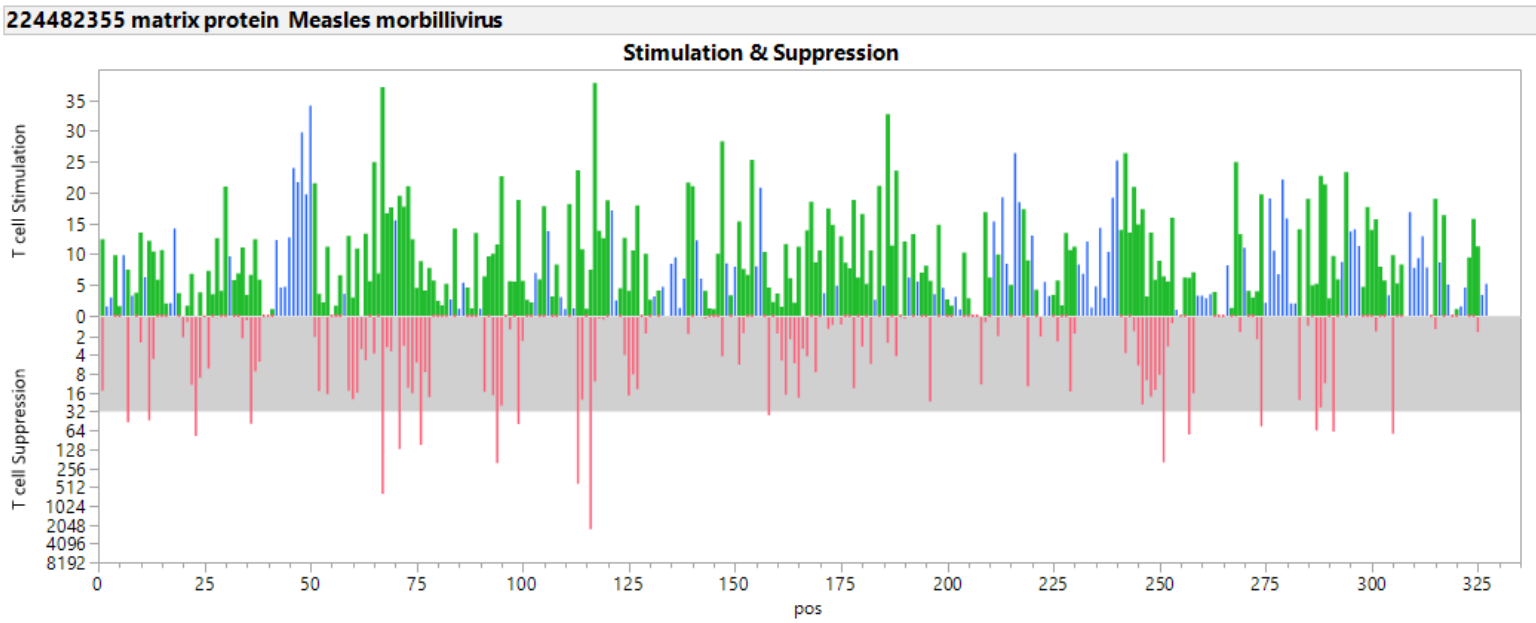

224482355 matrix protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

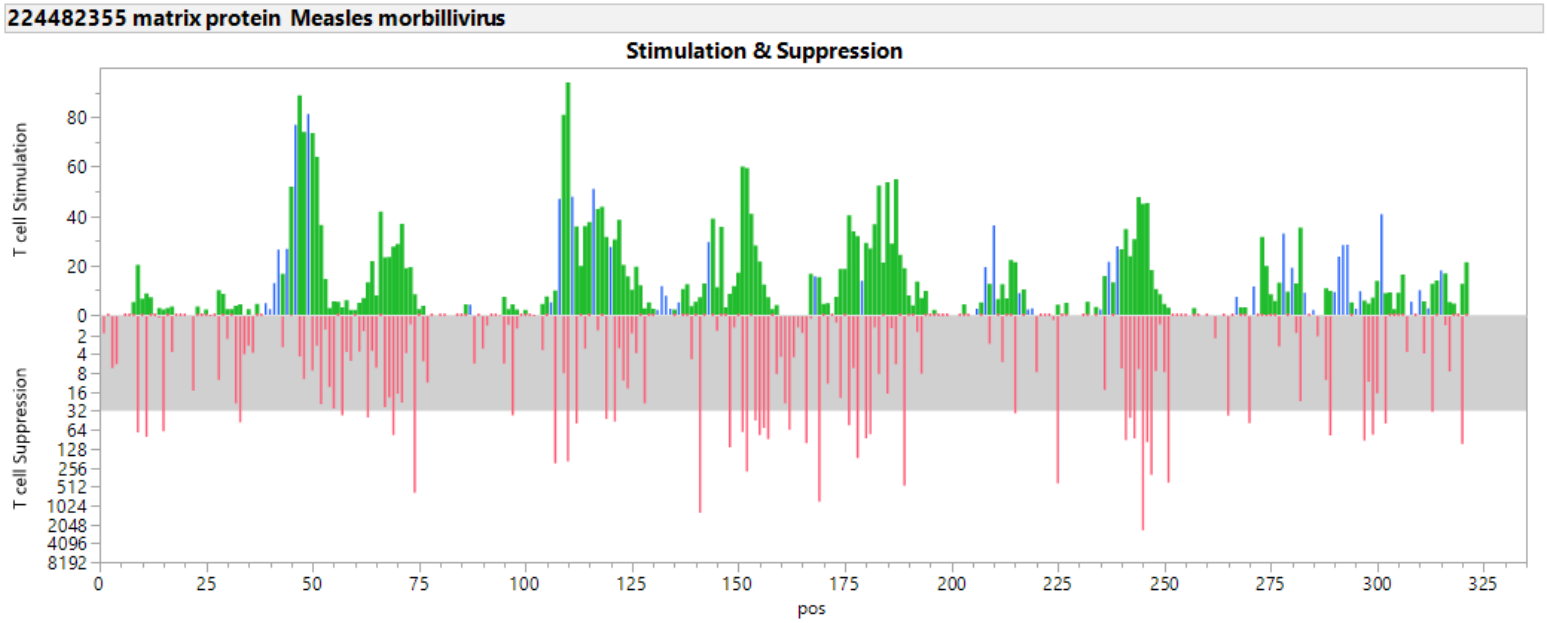

Hierarchical Cluster MHC I
Method = Ward
224482355 matrix protein Measles morbillivirus;t_T_MHC_I
nA_0101
nA_2601
nA_3001
nB_5401
nH-2-Ld
nB_0702
nB_3501
nB_5301
nB_5101
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_3201
nB_1501
nH-2-Db
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.72
-3.77
-2.83
-1.89
-0.94
0.00
1.06
2.13
3.19
4.25
5.31
25
50
75
100
125
150
175
200
225
250
275
300
325
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
224482355 matrix protein Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nA_0301
nA_1101
nA_3002
nA_3101
nA_6801
nA_3301
nB_1801
nB_4402
nB_4001
nH-2-Kk
nB_4403
nB_4002
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_2601
nH-2-Dd
nH-2-Db
nB_1503
nH-2-Ld
nB_2705
nA_2902
nB_3501
nB_5701
nB_5801
nB_5301
nA_3201
nB_1501
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_4501
nA_3001
nB_0702
nB_5101
nB_5401
nB_0801
nH-2-Kb
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482355 matrix protein Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
8
16
24
32
40
48
56
64
72
80
0
25
50
75
100
125
150
175
200
225
250
275
300
325
N-term Amino Acid of 15-mer


224482355 matrix protein Measles morbillivirus
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0
25
50
75
100
125
150
175
200
225
250
275
300
325
pos



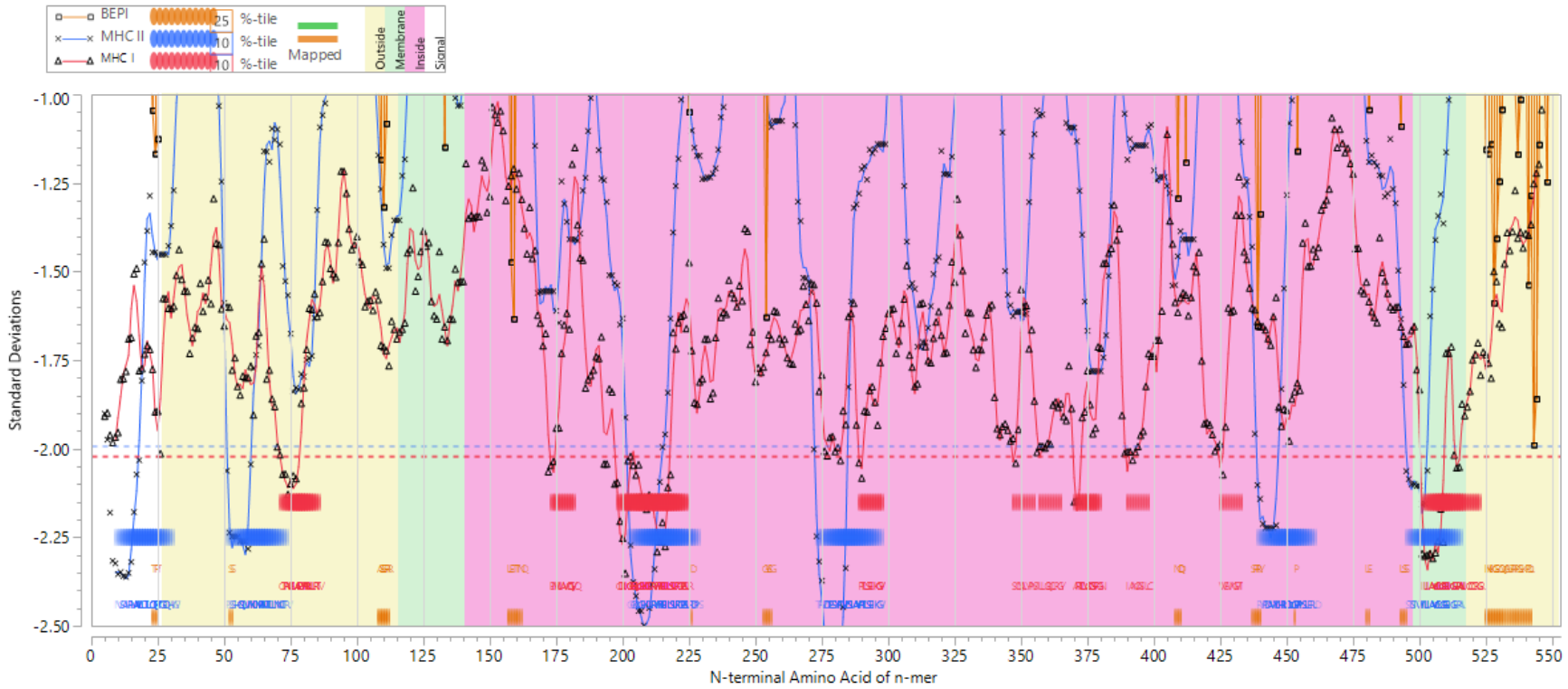


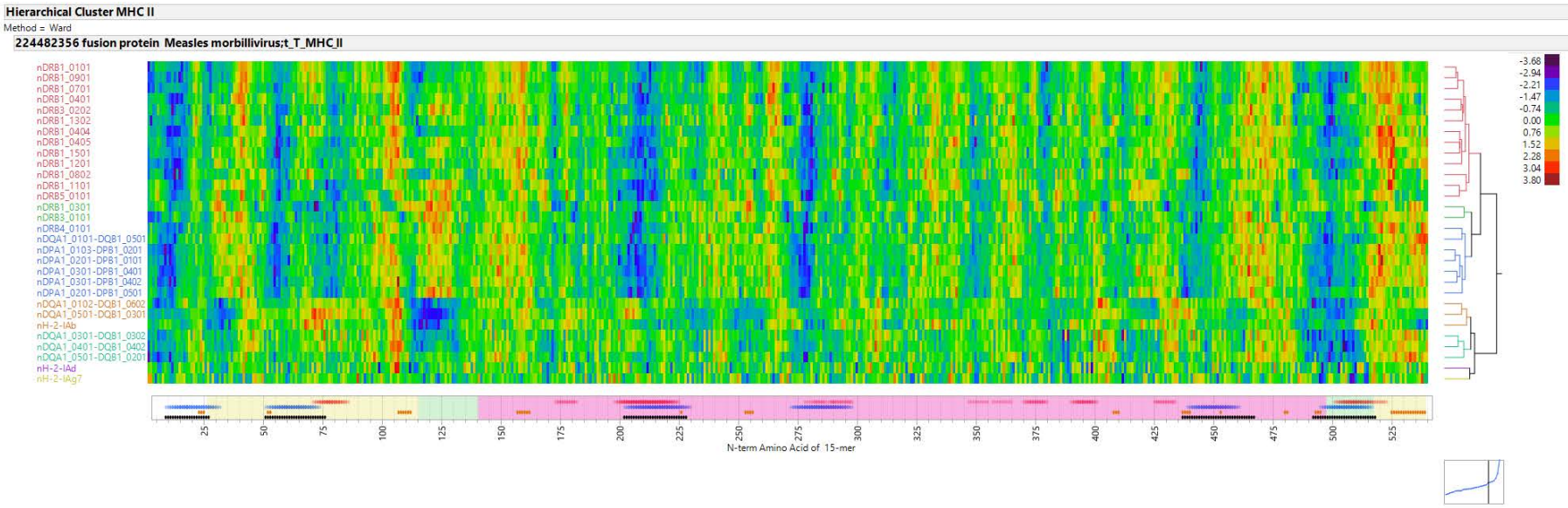
MHC II pos 131 suppressive index 172300

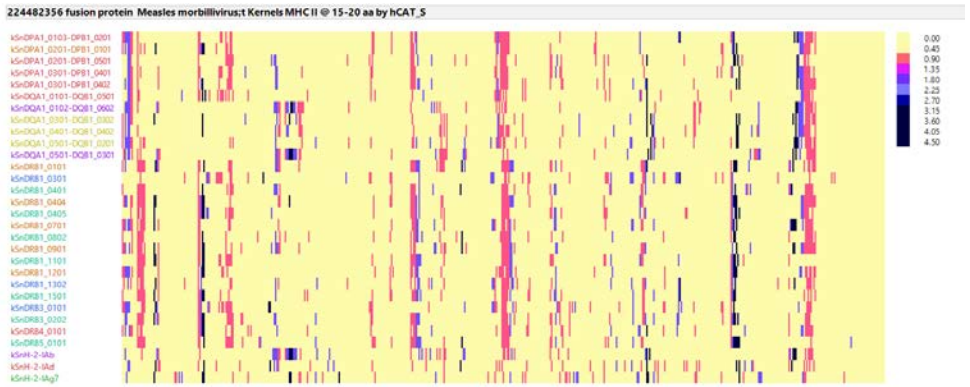
Fusion protein

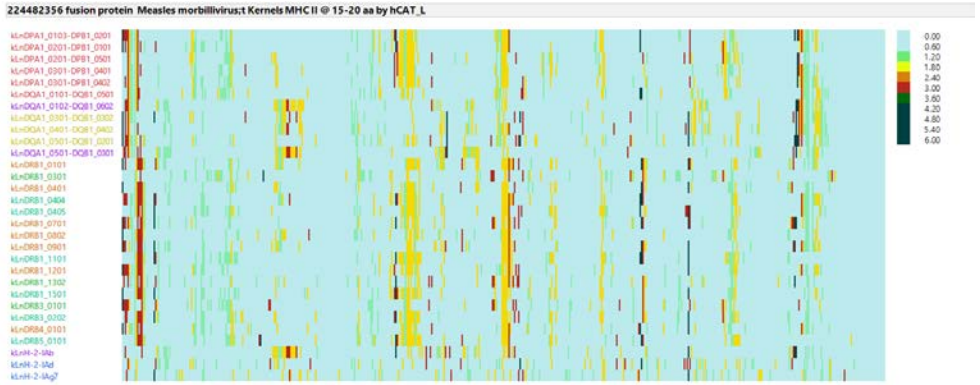
Overlay Plot

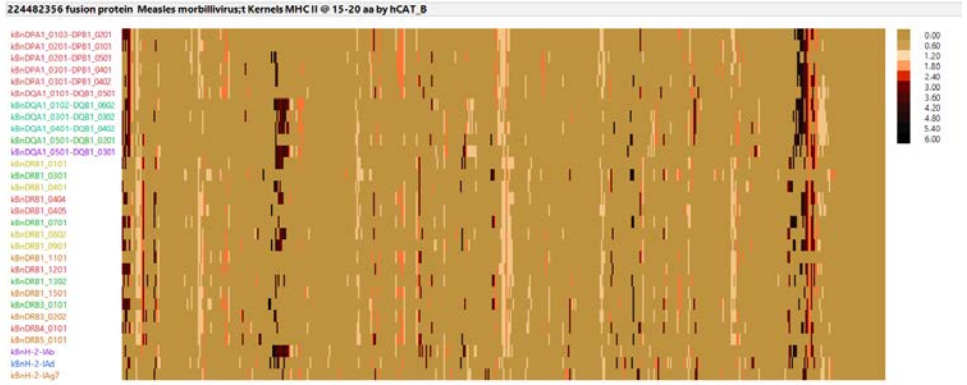
224482356 fusion protein Measles morbillivirus;t

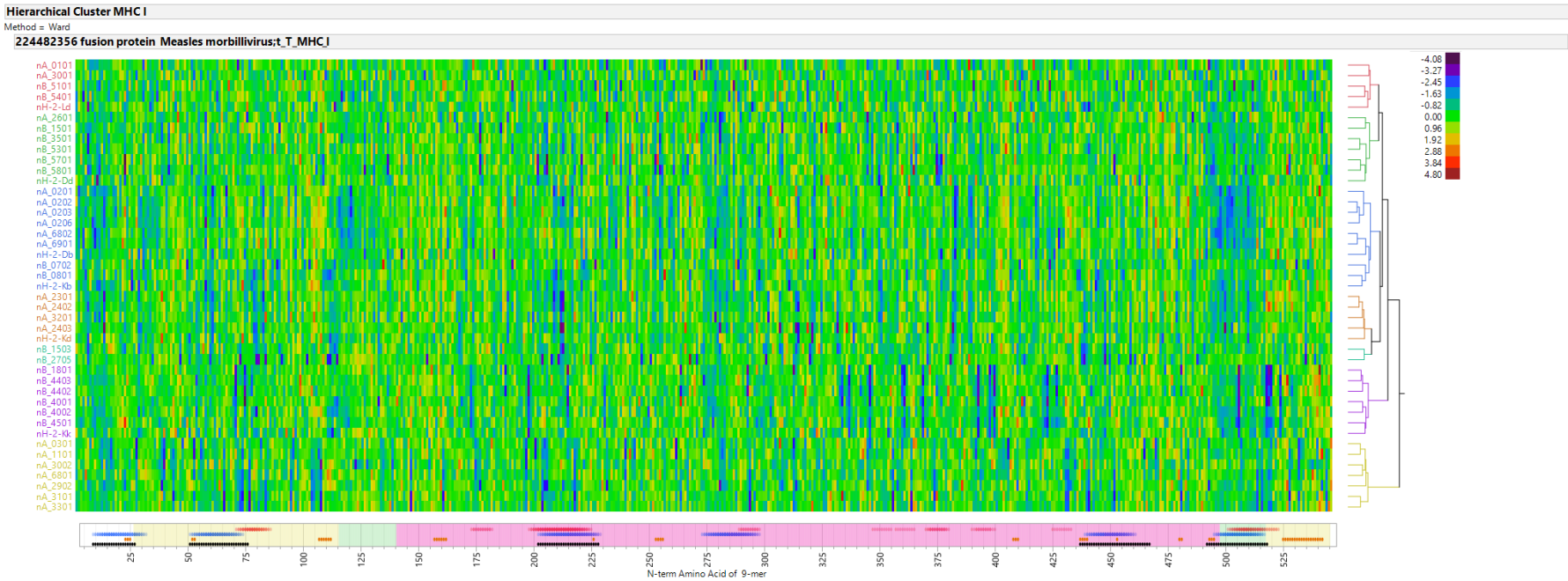


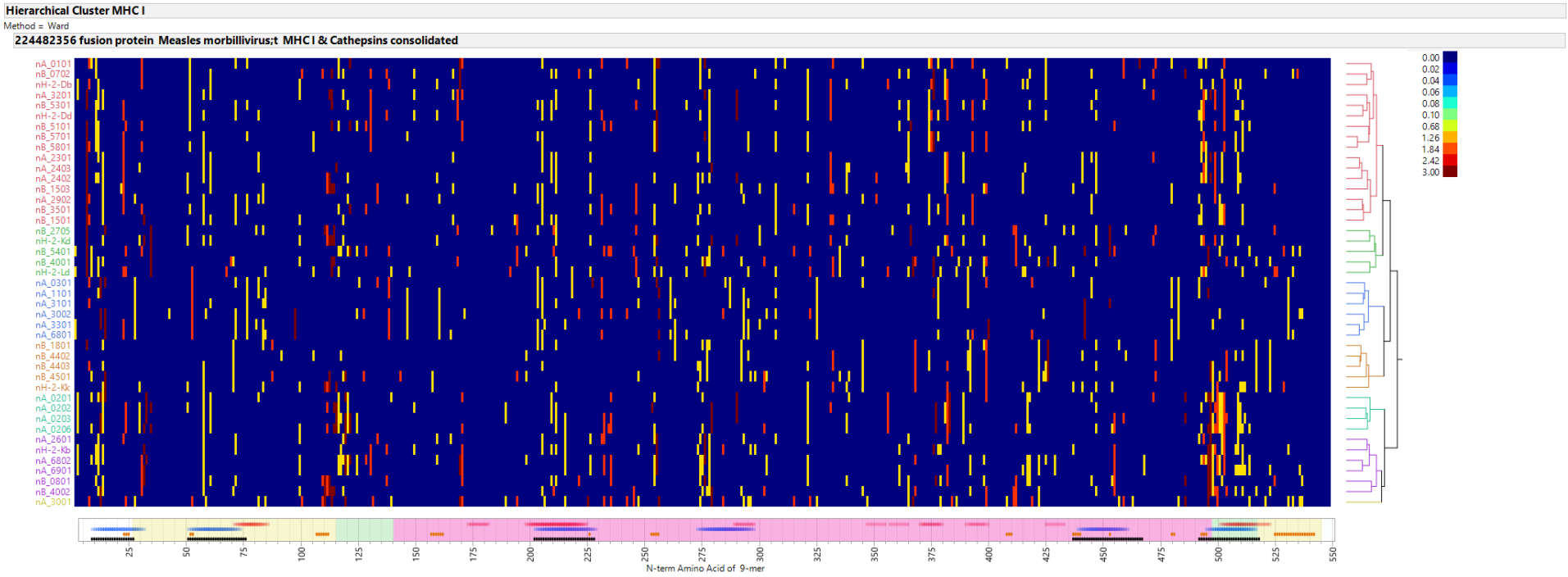

Hierarchical Cluster MHC II
Method = Ward
224482356 fusion protein Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0401
nDRB3_0202
nDRB1_1302
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB1_1201
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDRB4_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0501-DQB1_0301
nH-2-IAb
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
-3.68
-2.94
-2.21
-1.47
-0.74
0.00
0.76
1.52
2.28
3.04
3.80
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

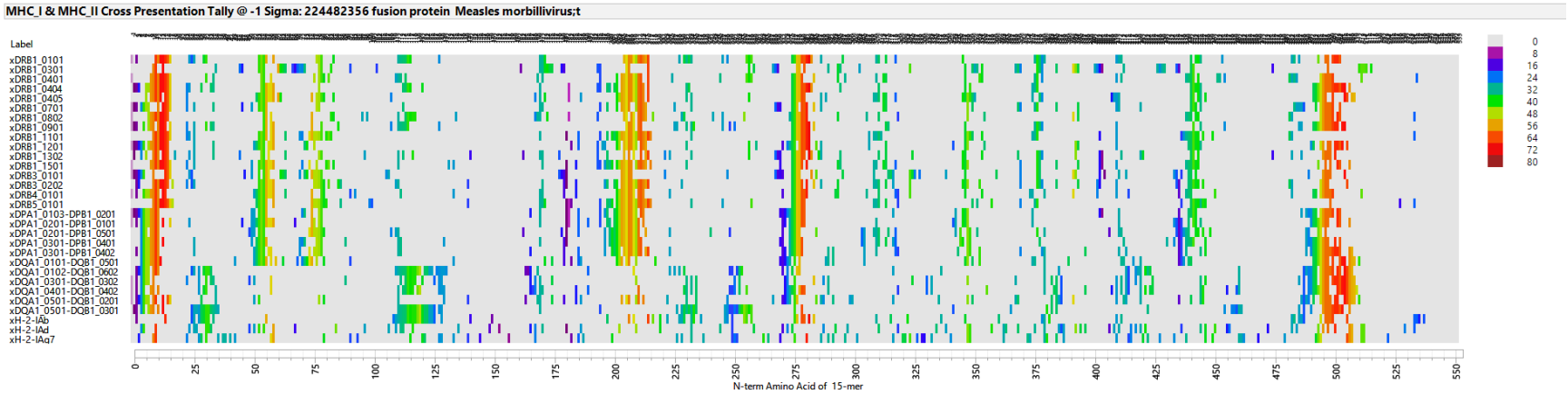

224482356 fusion protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

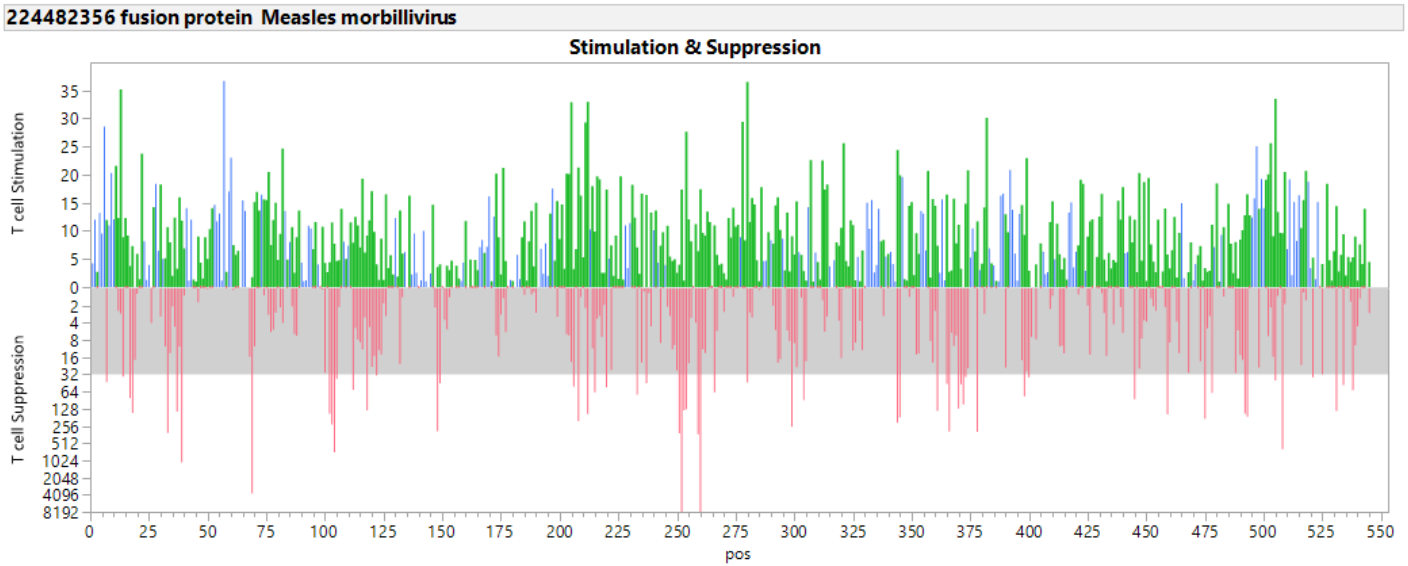

224482356 fusion protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

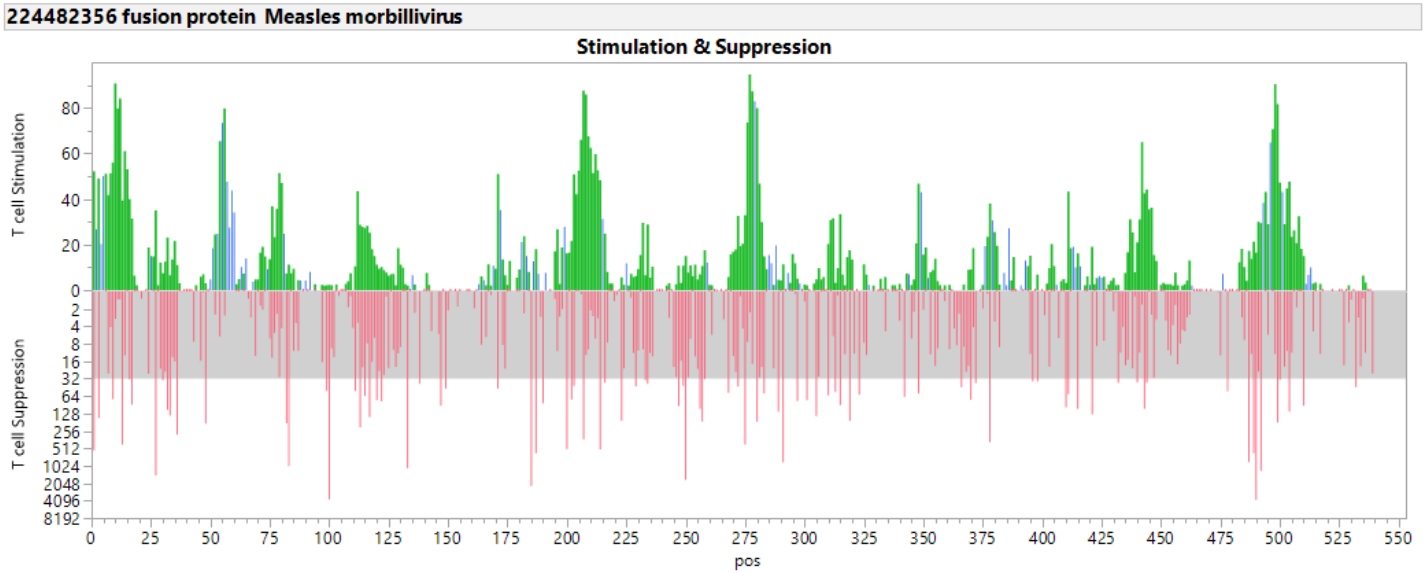

224482356 fusion protein Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00


Hierarchical Cluster MHC I
Method = Ward
224482356 fusion protein Measles morbillivirus;t_T_MHC_I
nA_0101
nA_3001
nB_5101
nB_5401
nH-2-Ld
nA_2601
nB_1501
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
-4.08
-3.27
-2.45
-1.63
-0.82
0.00
0.96
1.92
2.88
3.84
4.80
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
224482356 fusion protein Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nB_0702
nH-2-Db
nA_3201
nB_5301
nH-2-Dd
nB_5101
nB_5701
nB_5801
nA_2301
nA_2403
nA_2402
nB_1503
nA_2902
nB_3501
nB_1501
nB_2705
nH-2-Kd
nB_5401
nB_4001
nH-2-Ld
nA_0301
nA_1101
nA_3101
nA_3002
nA_3301
nA_6801
nB_1801
nB_4402
nB_4403
nB_4501
nH-2-Kk
nA_0201
nA_0202
nA_0203
nA_0206
nA_2601
nH-2-Kb
nA_6802
nA_6901
nB_0801
nB_4002
nA_3001
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 9-mer

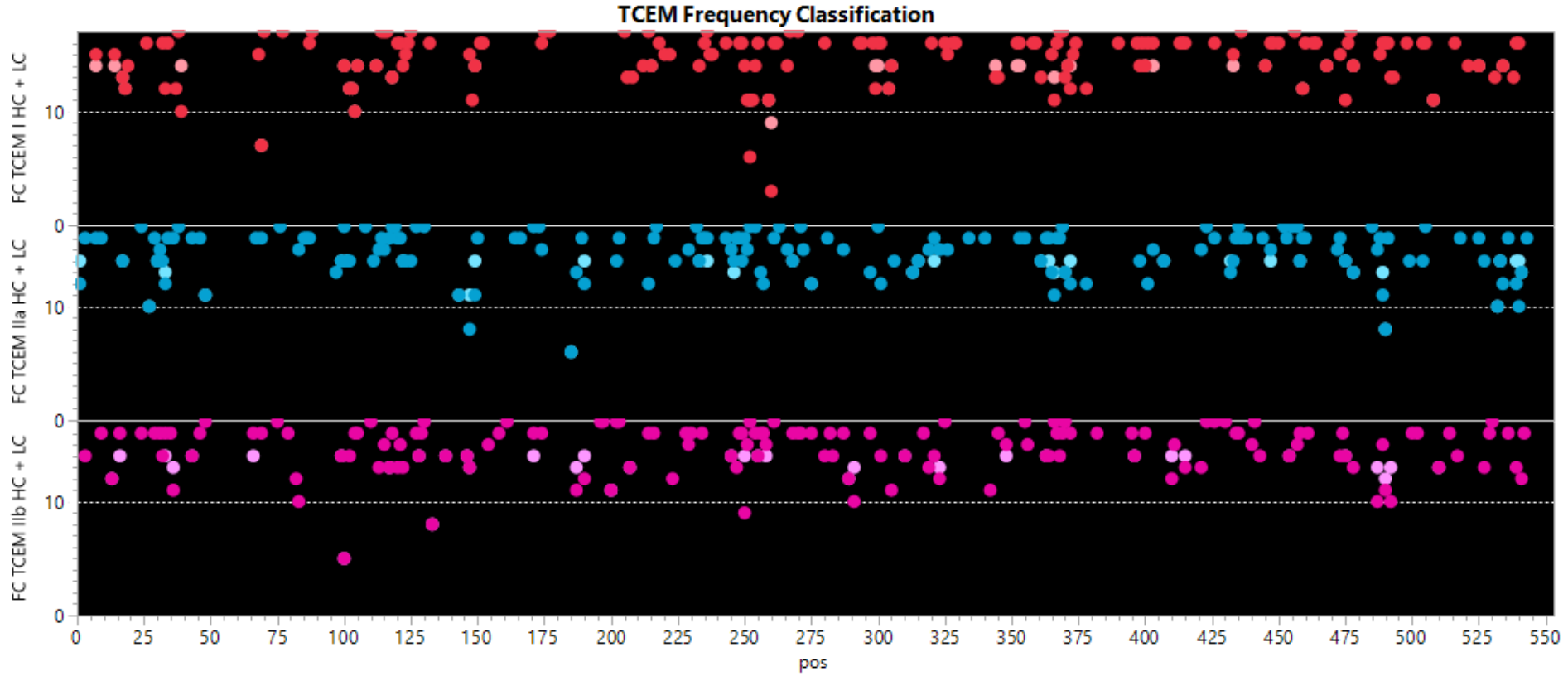

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482356 fusion protein Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550
N-term Amino Acid of 15-mer
0 8 16 24 32 40 48 56 64 72 80





MHC II pos 131 suppressive index 172300

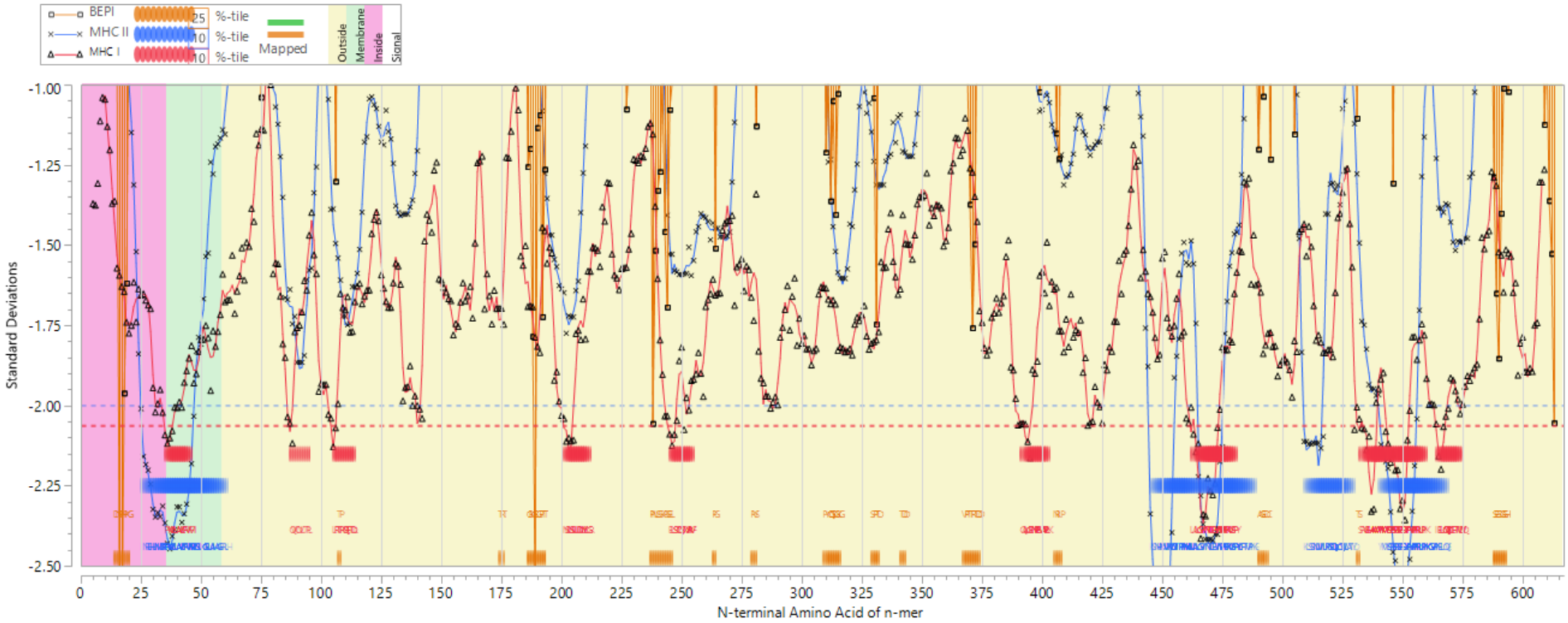
224482356 fusion protein Measles morbillivirus

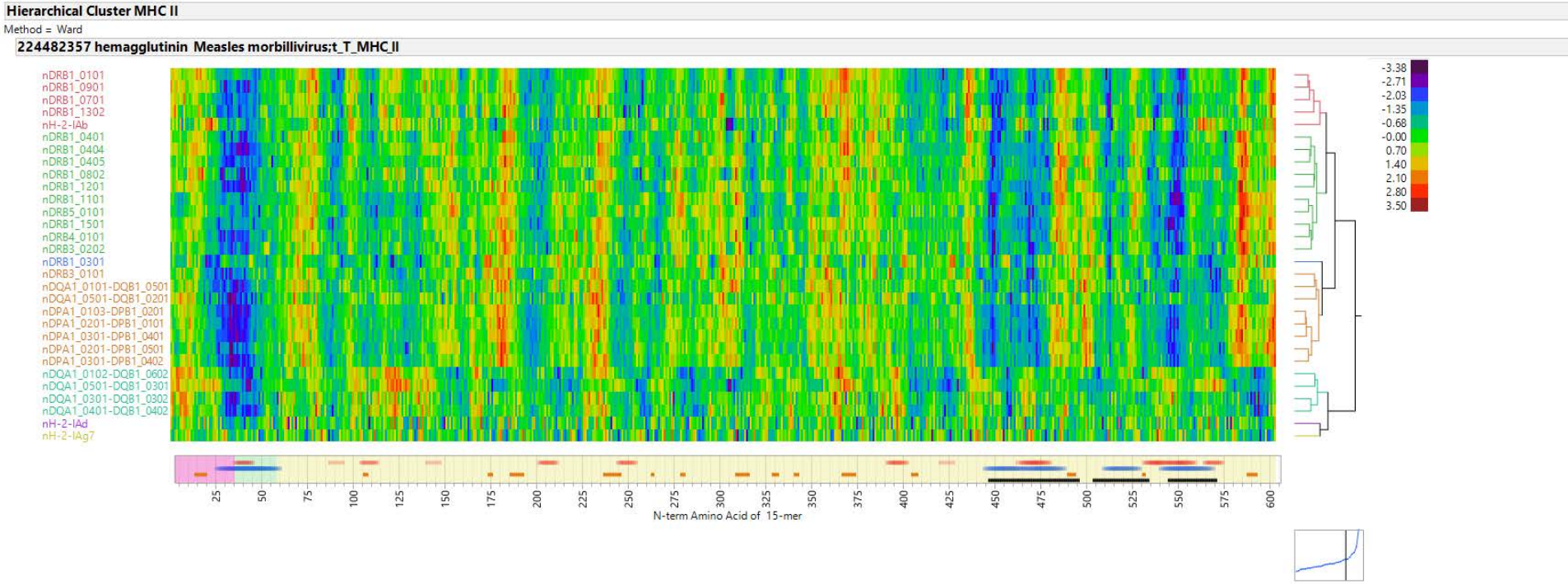


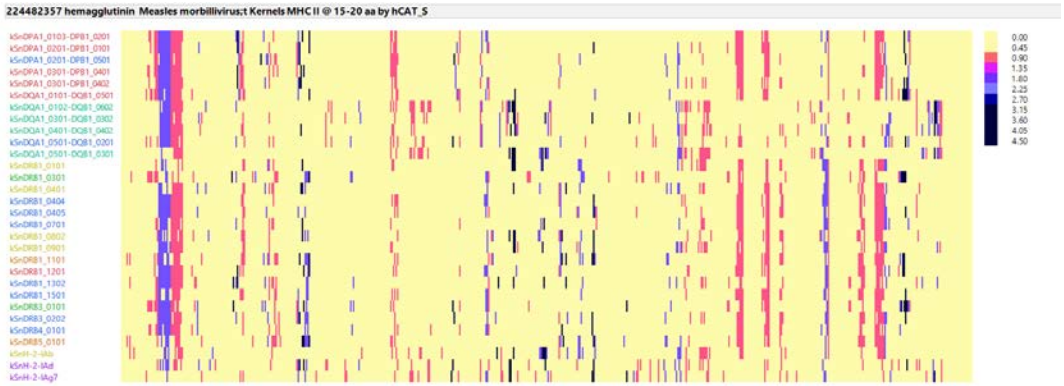
Hemagglutinin

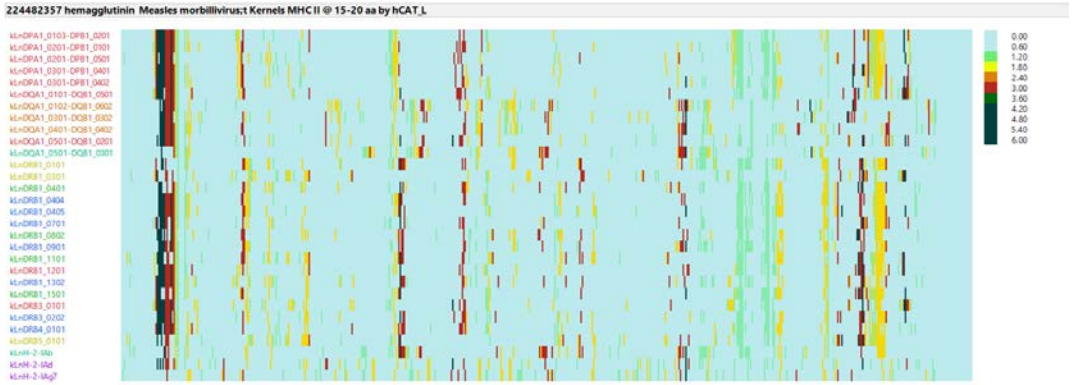
Overlay Plot

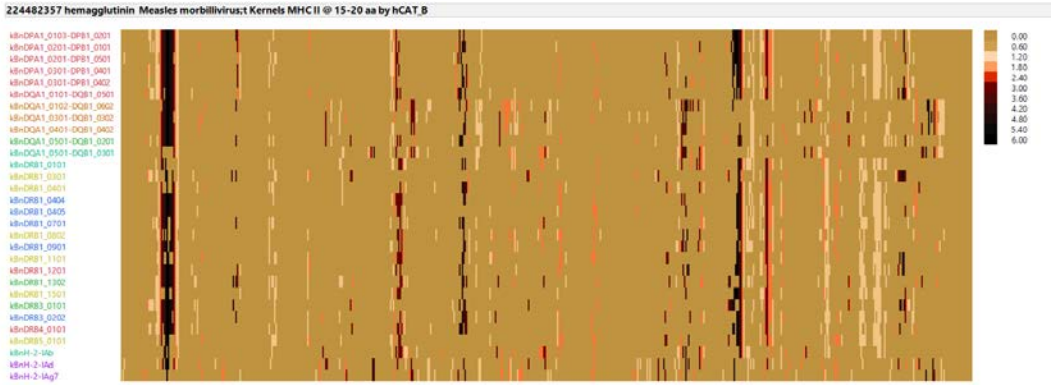
224482357 hemagglutinin Measles morbillivirus;t

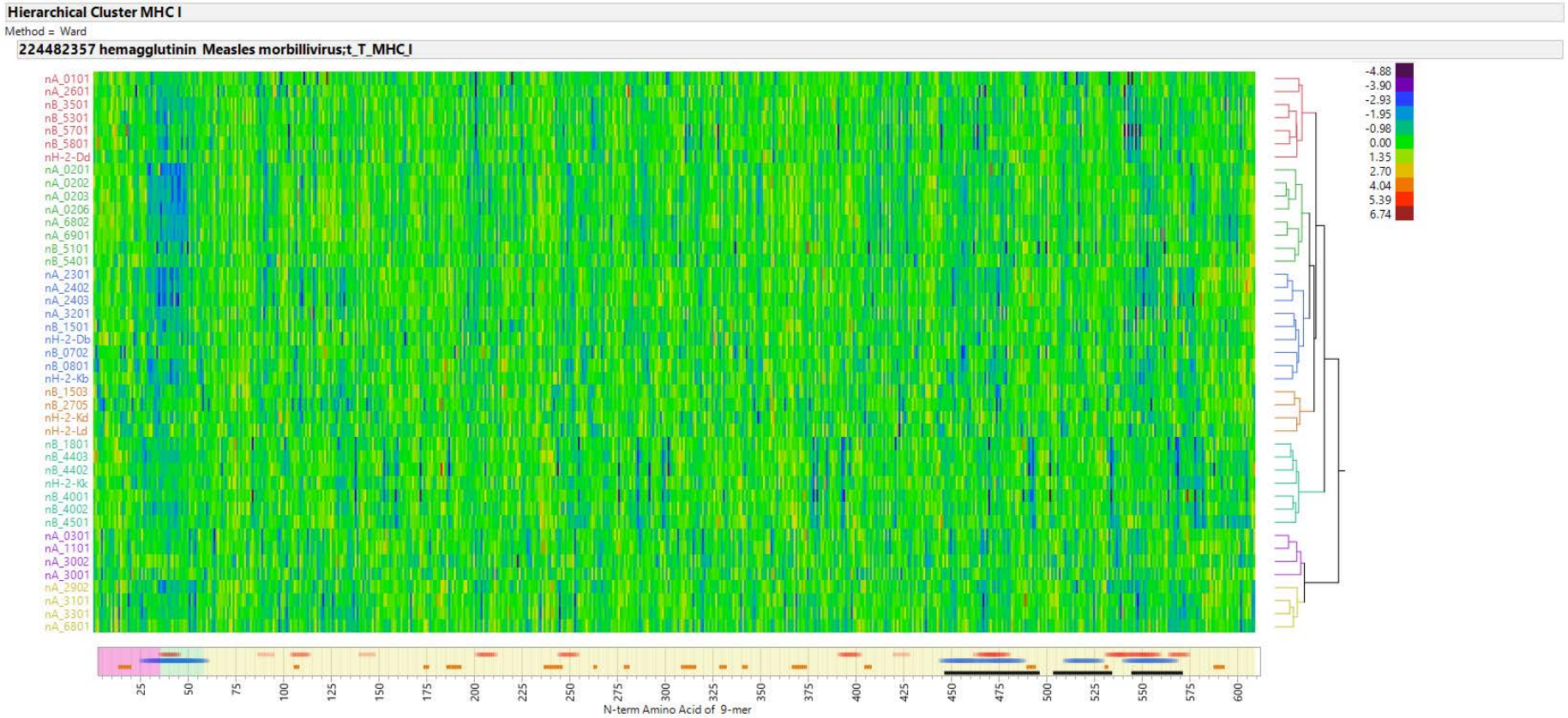


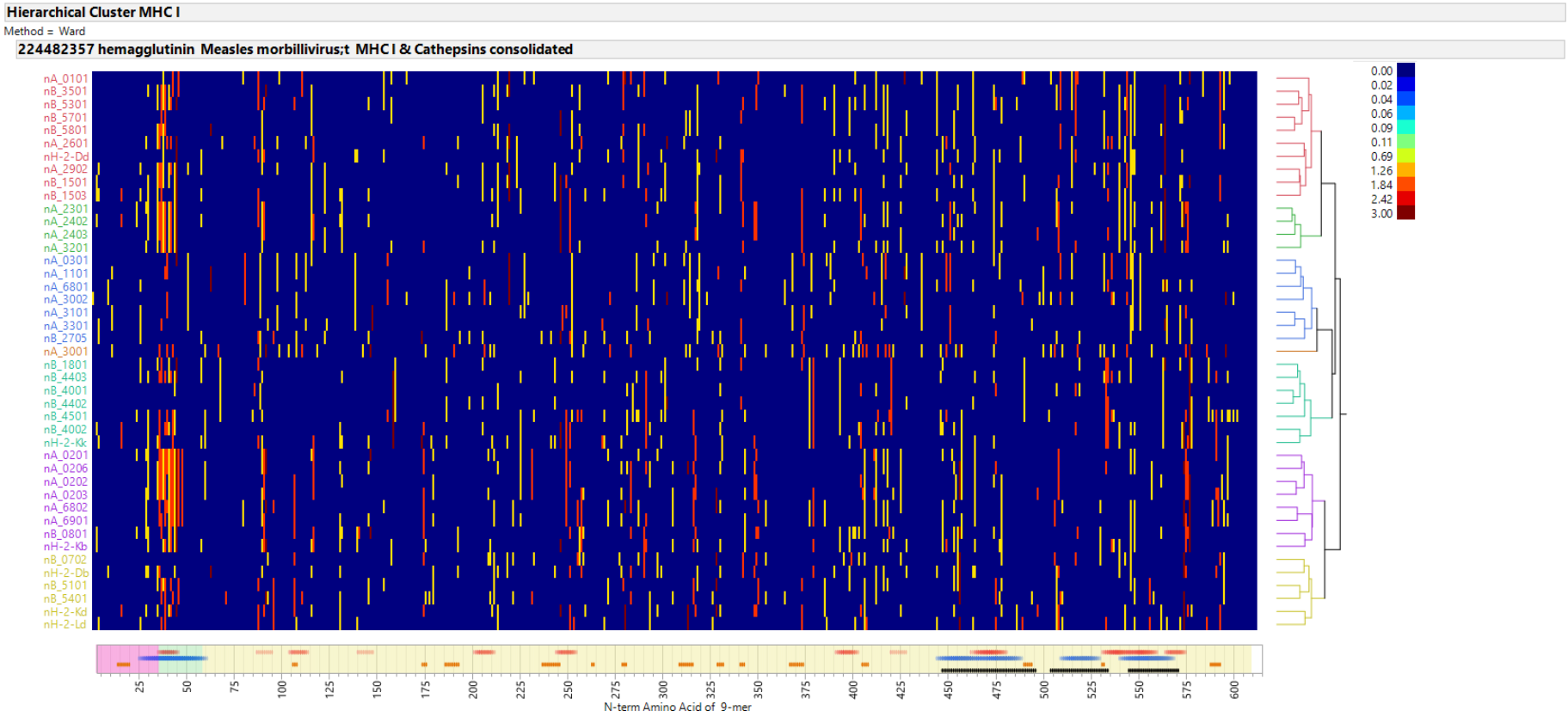

Hierarchical Cluster MHC II
Method = Ward
224482357 hemagglutinin Measles morbillivirus;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nH-2-IAb
nDRB1_0401
nDRB1_0404
nDRB1_0405
nDRB1_0802
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_0301
nDRB3_0101
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0102-DQB1_0602
nDQA1_0501-DQB1_0301
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
-3.38
-2.71
-2.03
-1.35
-0.68
-0.00
0.70
1.40
2.10
2.80
3.50
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
N-term Amino Acid of 15-mer

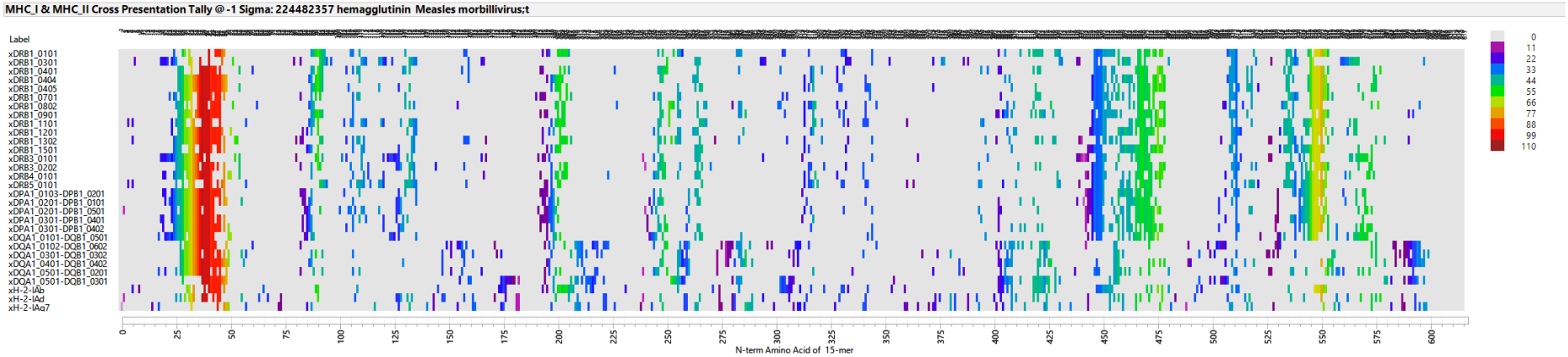

224482357 hemagglutinin Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

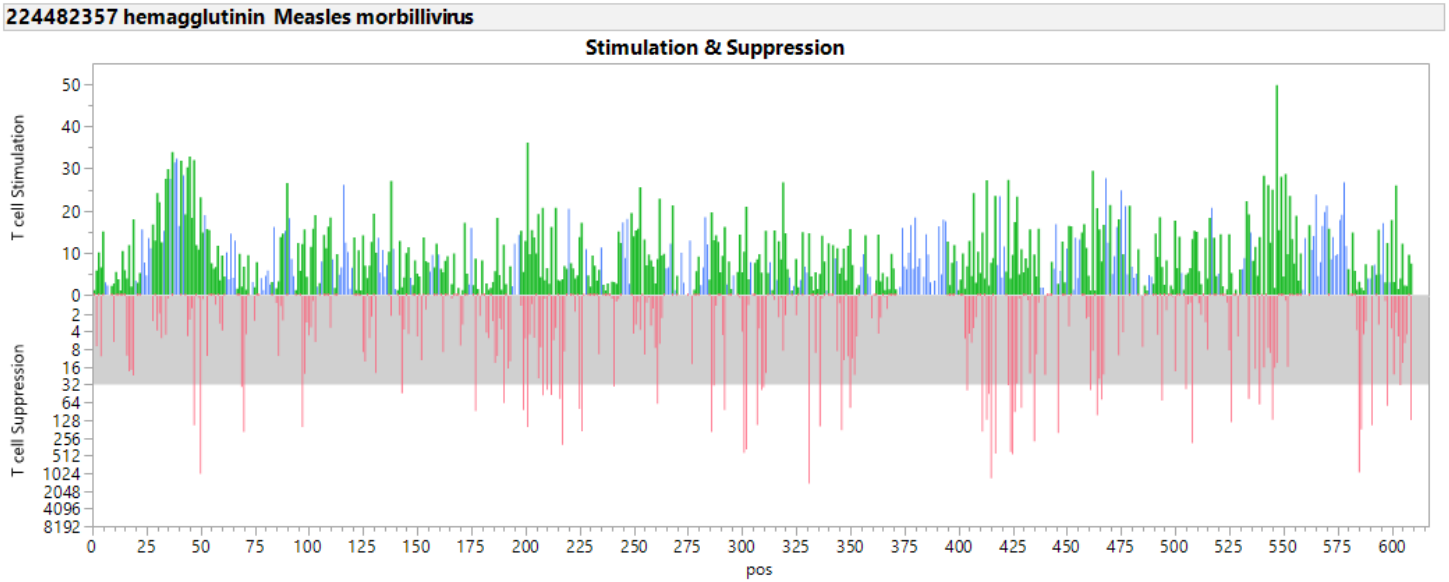

224482357 hemagglutinin Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

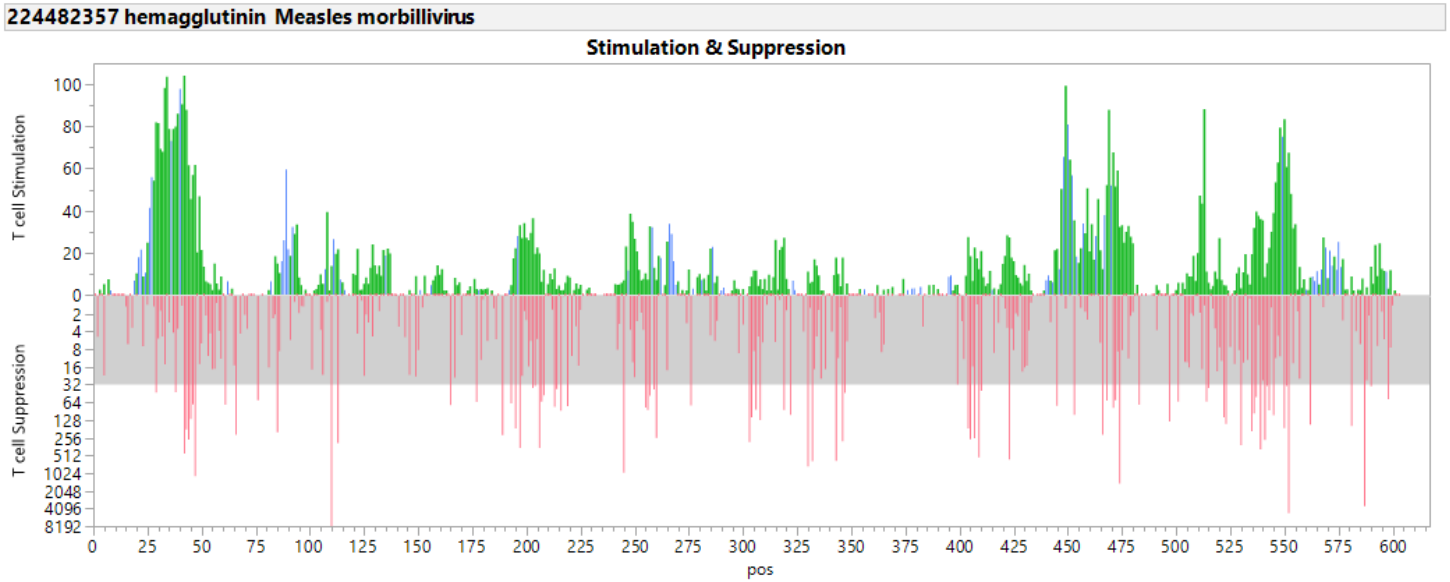

224482357 hemagglutinin Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00


Hierarchical Cluster MHC I
Method = Ward
224482357 hemagglutinin Measles morbillivirus;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nB_5401
nA_2301
nA_2402
nA_2403
nA_3201
nB_1501
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nH-2-Kk
nB_4001
nB_4002
nB_4501
nA_0301
nA_1101
nA_3002
nA_3001
nA_2902
nA_3101
nA_3301
nA_6801
-4.88
-3.90
-2.93
-1.95
-0.98
0.00
1.35
2.70
4.04
5.39
6.74
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
224482357 hemagglutinin Measles morbillivirus;t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nB_5301
nB_5701
nB_5801
nA_2601
nH-2-Dd
nA_2902
nB_1501
nB_1503
nA_2301
nA_2402
nA_2403
nA_3201
nA_0301
nA_1101
nA_6801
nA_3002
nA_3101
nA_3301
nB_2705
nA_3001
nB_1801
nB_4403
nB_4001
nB_4402
nB_4501
nB_4002
nH-2-Kk
nA_0201
nA_0206
nA_0202
nA_0203
nA_6802
nA_6901
nB_0801
nH-2-Kb
nB_0702
nH-2-Db
nB_5101
nB_5401
nH-2-Kd
nH-2-Ld
0.00
0.02
0.04
0.06
0.09
0.11
0.69
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
N-term Amino Acid of 9-mer

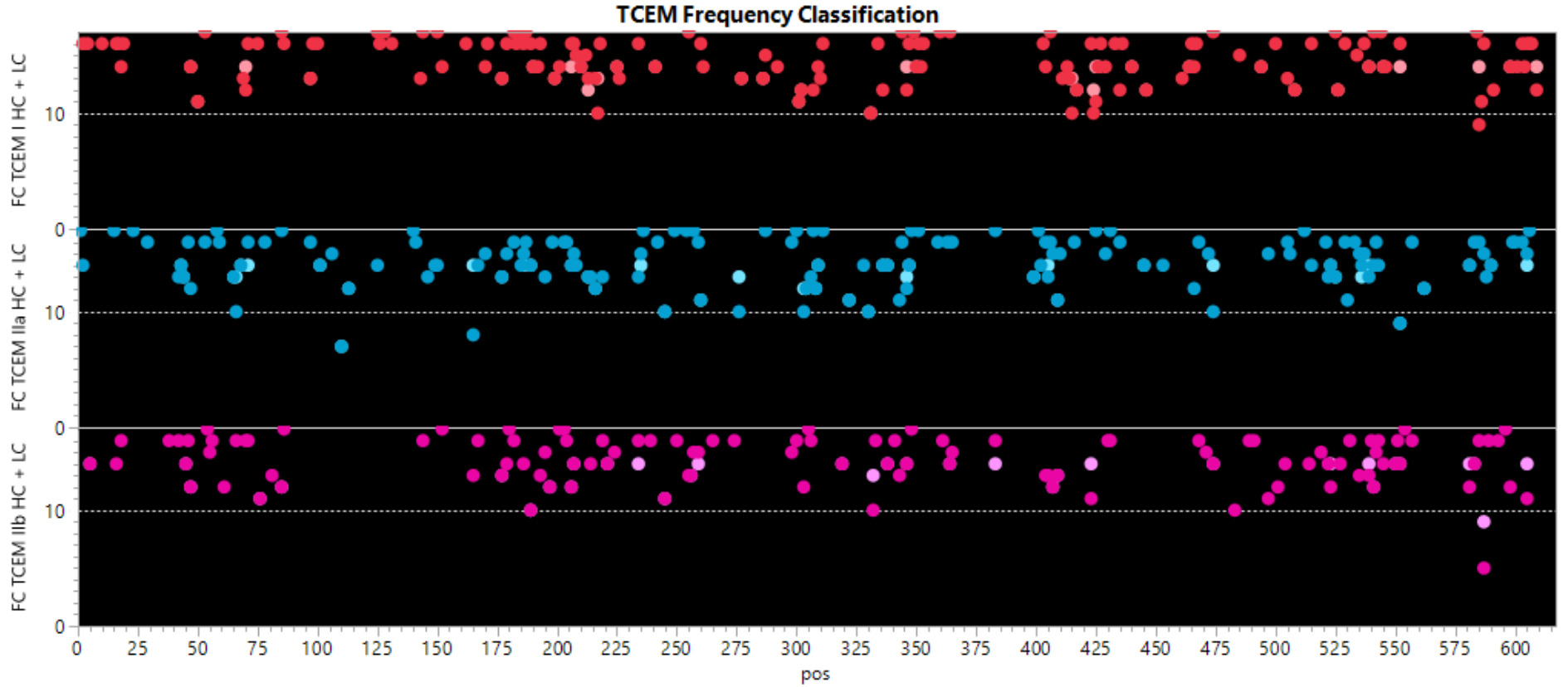

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 224482357 hemagglutinin Measles morbillivirus;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600
N-term Amino Acid of 15-mer
0
11
22
33
44
55
66
77
88
99
110



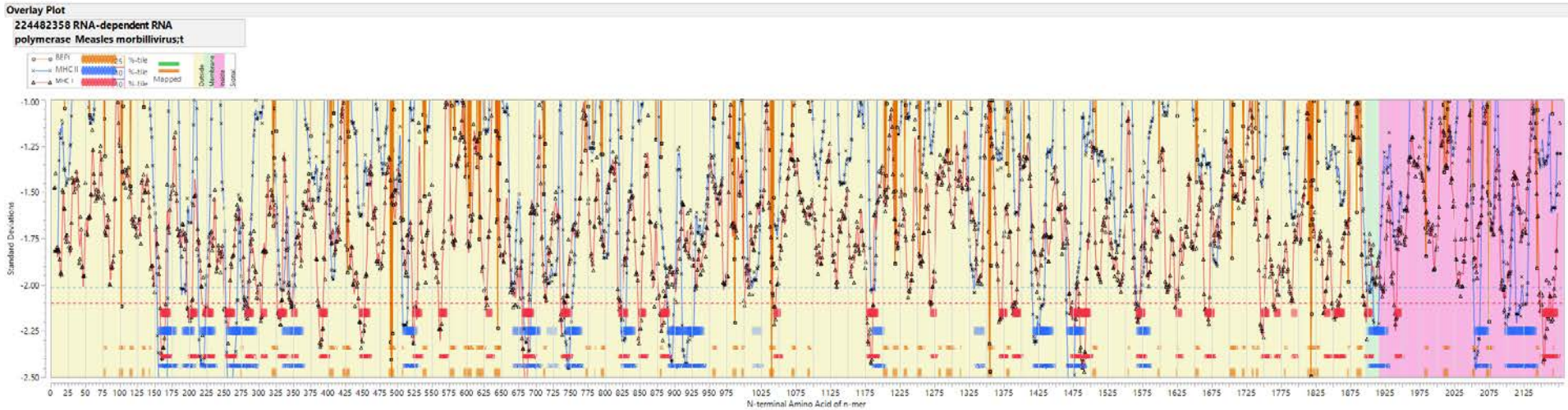


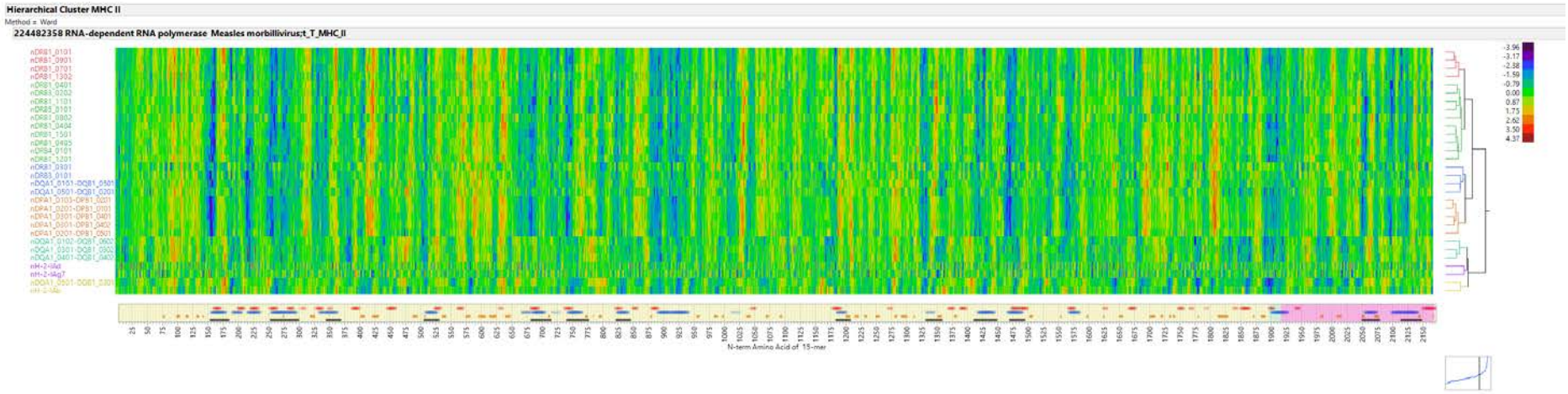
MHC II pos 131 suppressive index 172300

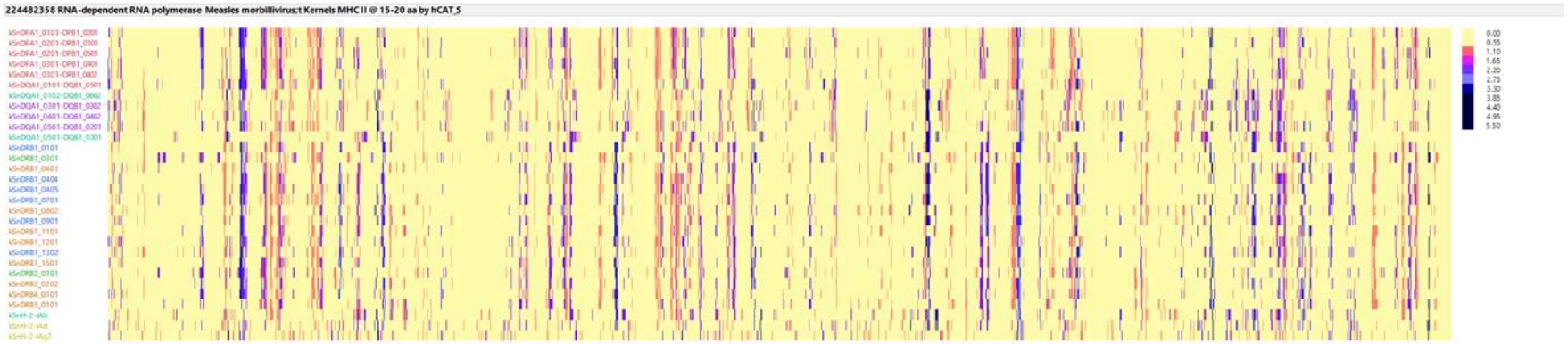
224482357 hemagglutinin Measles morbillivirus

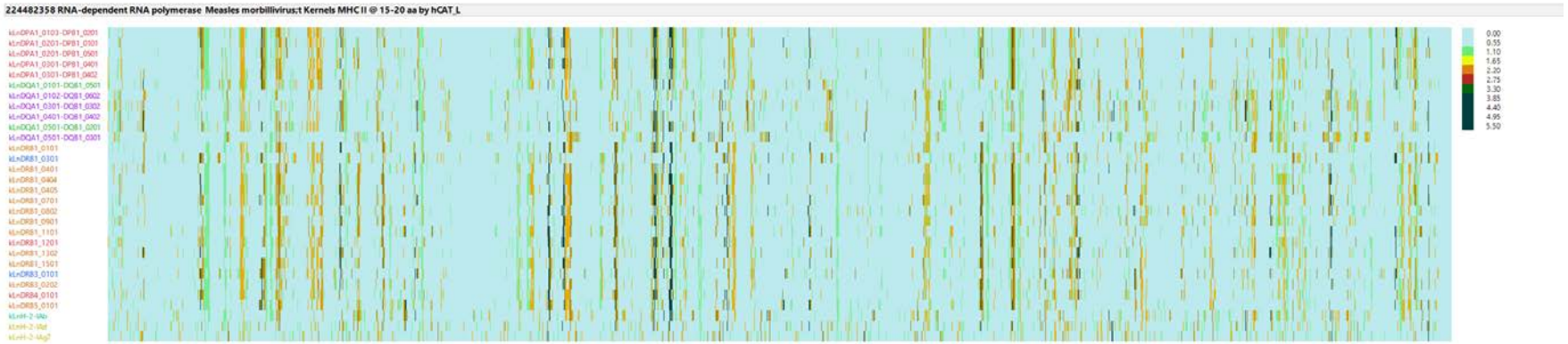


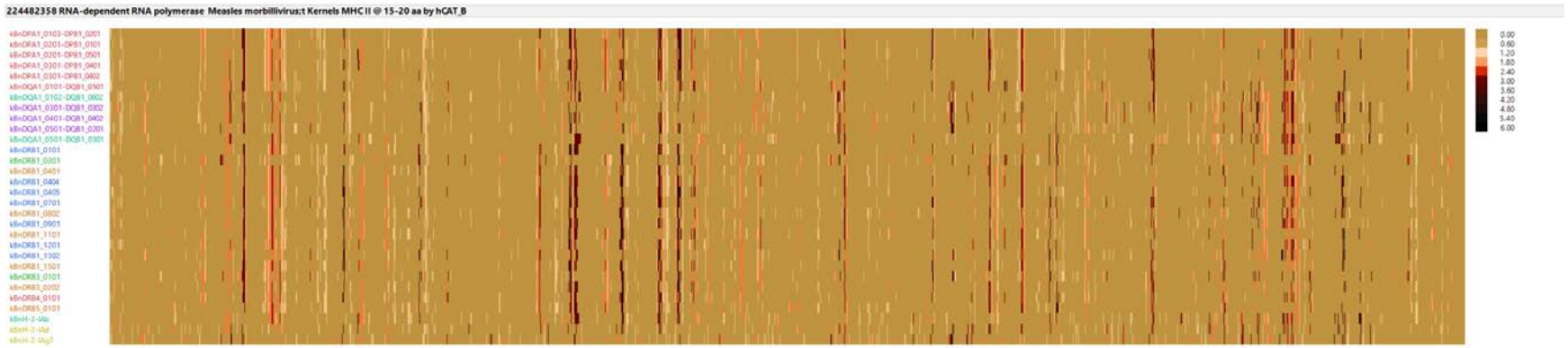
RNA polymerase

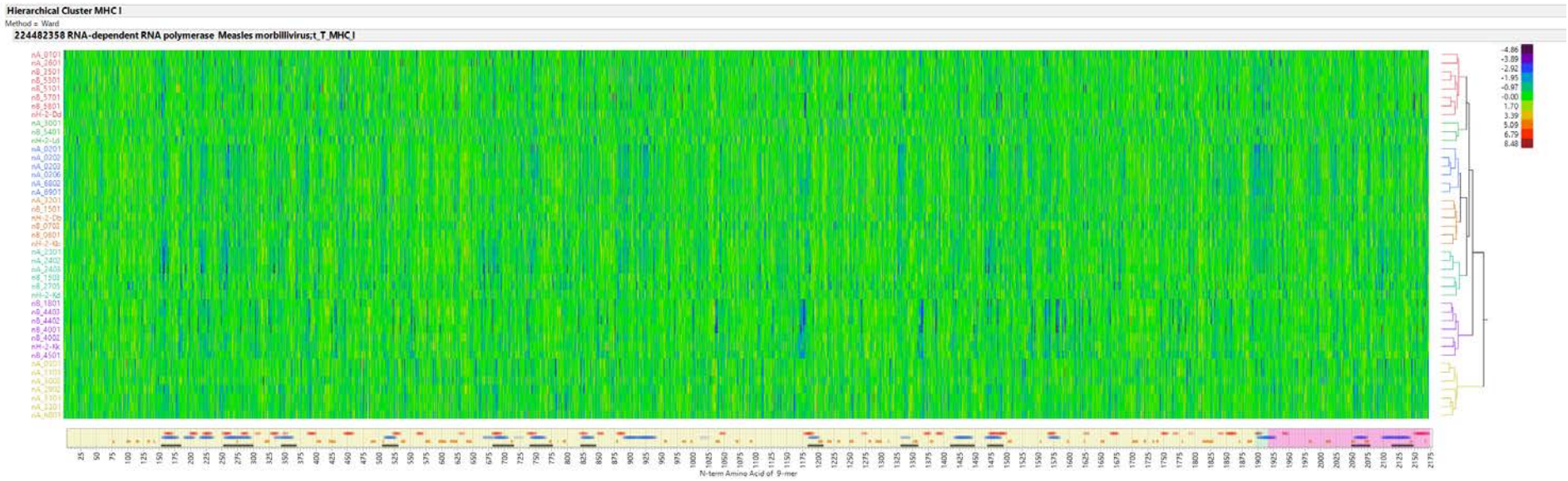

Overlay Plot
224482358 RNA-dependent RNA polymerase Measles morbillivirus;t
BEPI
MHC II
MHC I
%-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
N-terminal Amino Acid of n-mer

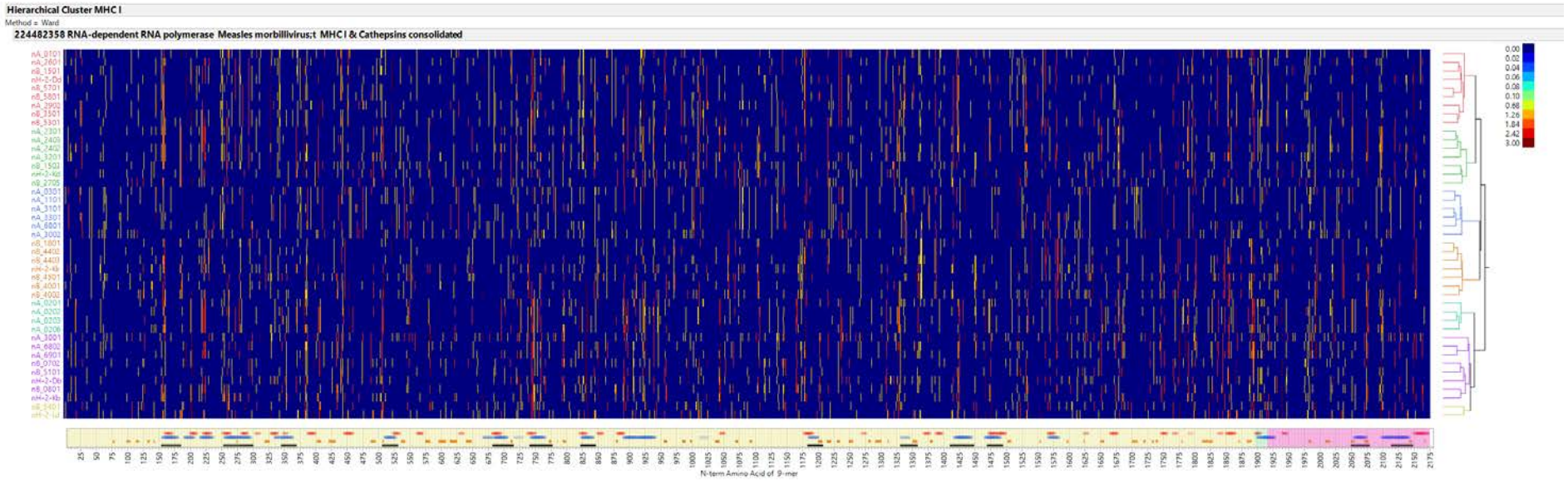

Hierarchical Cluster MHC II
Method = Ward
224482358 RNA-dependent RNA polymerase Measles morbillivirus;t_T_MHC_II
N-term Amino Acid of 15-mer

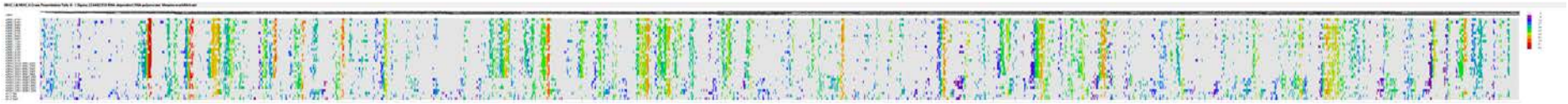

224482358 RNA-dependent RNA polymerase Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_S

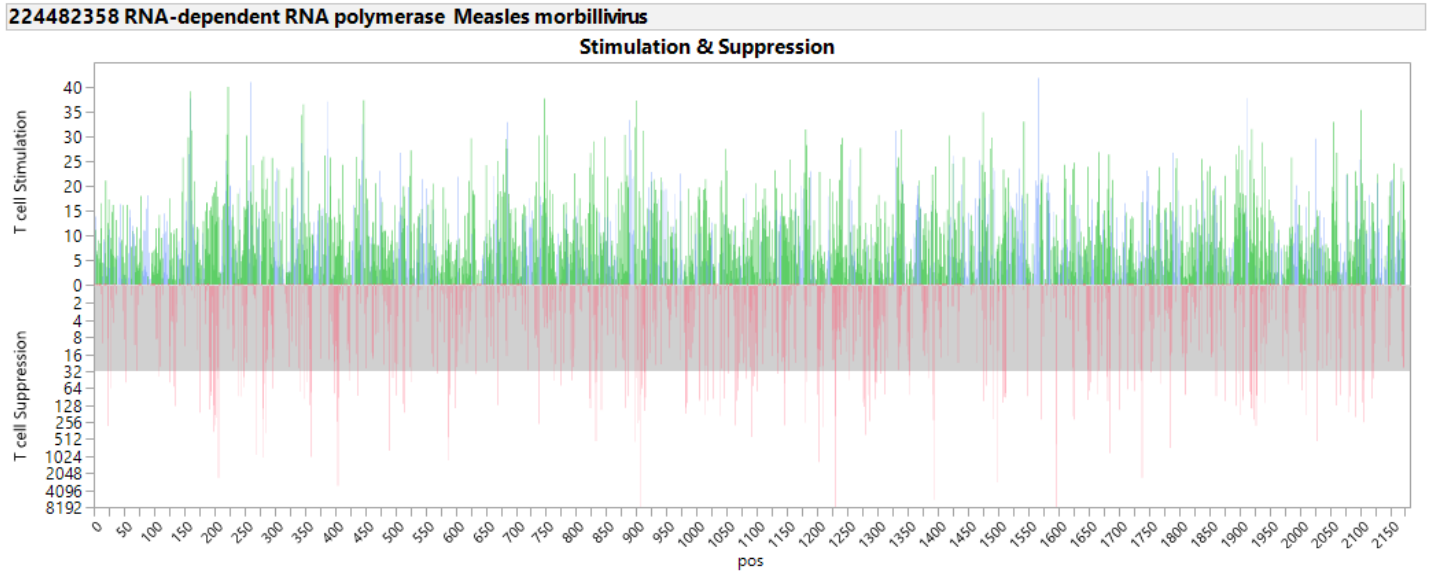

224482358 RNA-dependent RNA polymerase Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_L

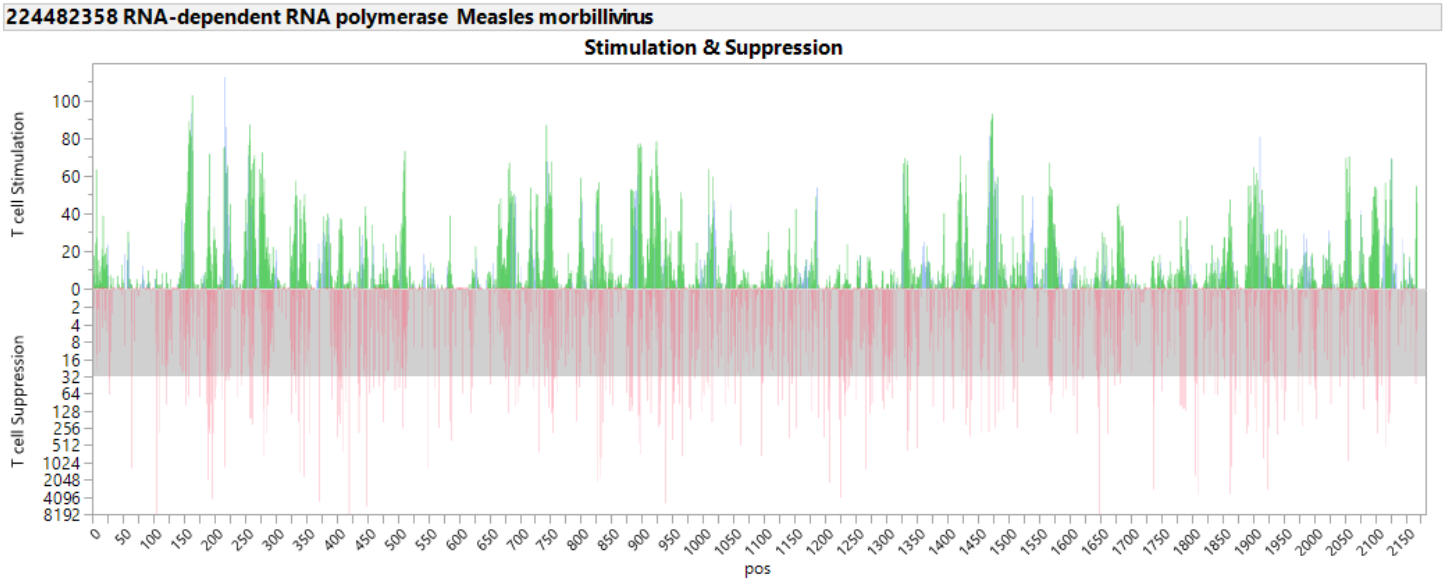

224482358 RNA-dependent RNA polymerase Measles morbillivirus;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
224482358 RNA-dependent RNA polymerase Measles morbillivirus;t_T_MHC I
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
224482358 RNA-dependent RNA polymerase Measles morbillivirus;t MHC I & Cathepsins consolidated
N-term Amino Acid of 9-mer

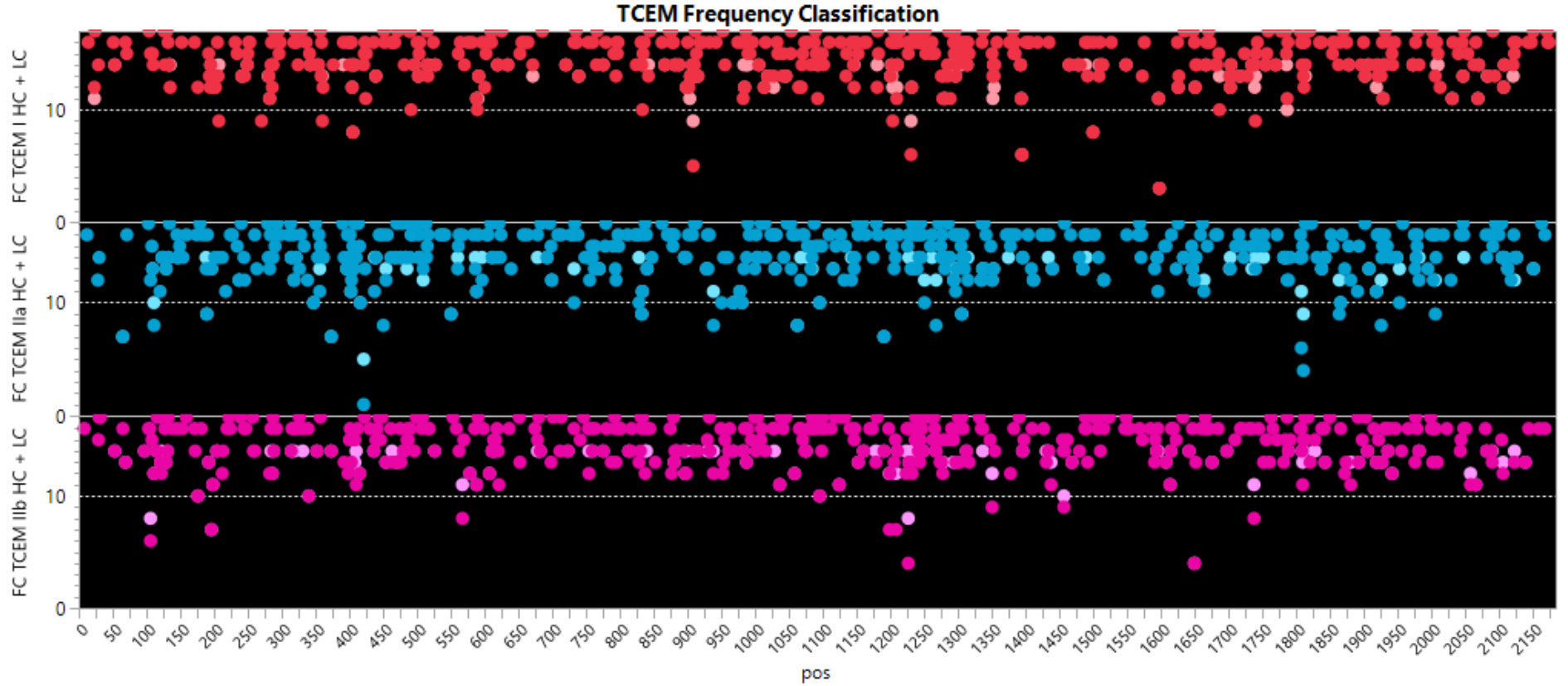






MHC II pos 131 suppressive index 172300

224482358 RNA-dependent RNA polymerase Measles morbillivirus



ioGenetics®

Mumps

MUMPSVAX® (Mumps Virus Vaccine Live), the Jeryl Lynn™ (B level) strain of mumps virus propagated in chick embryo cell culture;

JL2 “minor” gi 19070169-19070177

ioGenetics

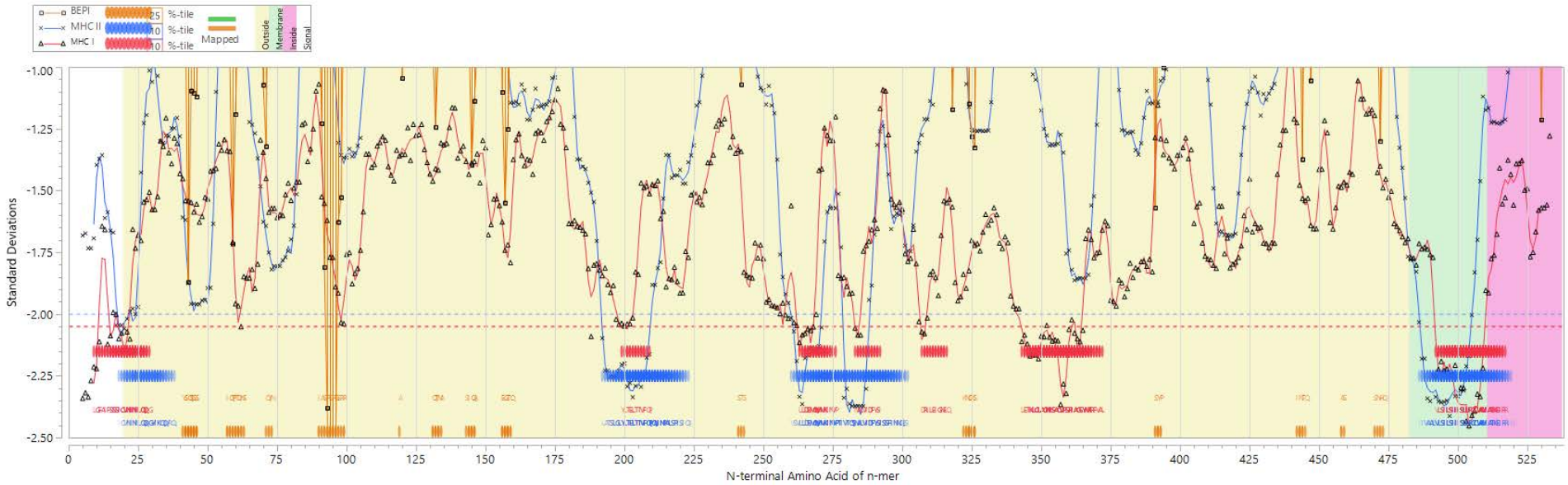
gi : curation	N Rows	protein
19070169-JL2-JL2 AAL83738_1	549	nucleocapsid protein
19070170-JL2-JL2 AAL83739_1	171	l protein
19070171-JL2-JL2 AAL83740_1	391	phosphoprotein
19070172-JL2-JL2 AAL83741_1	224	V protein
19070173-JL2-JL2 AAL83742_1	375	matrix protein
19070174-JL2-JL2 AAL83743_1	538	fusion protein
19070175-JL2-JL2 AAL83744_1	57	small hydrophobic protein
19070176-JL2-JL2 AAL83745_1	582	hemagglutinin/neuraminid...
19070177-JL2-JL2 AAL83746_1	2261	large protein

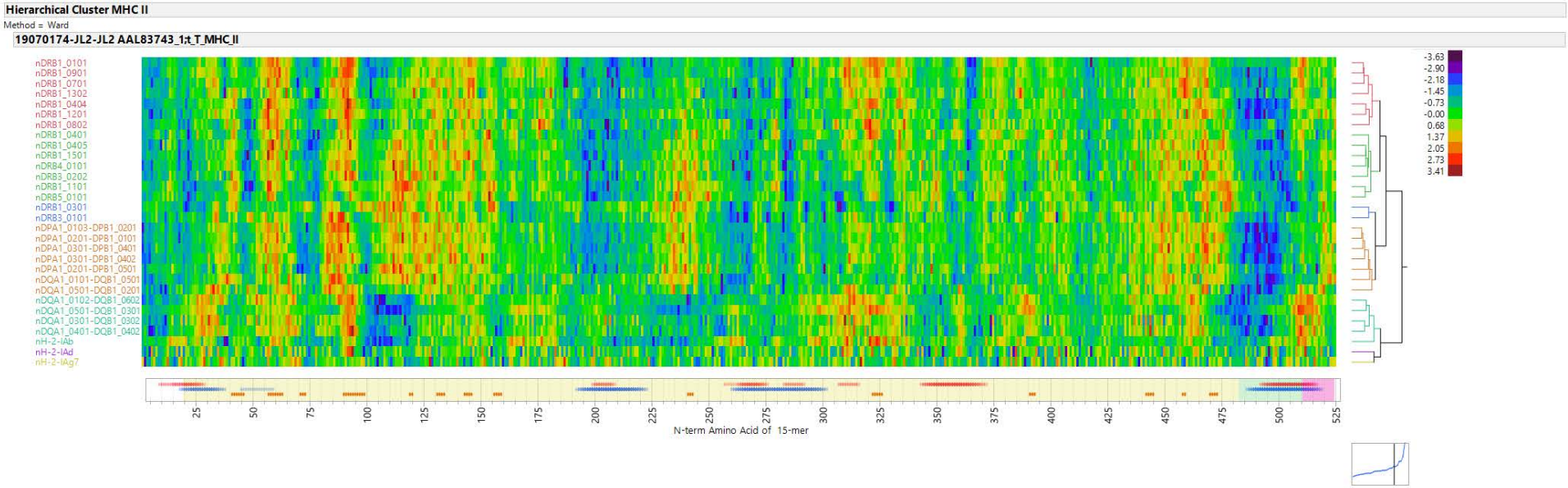
Fusion protein

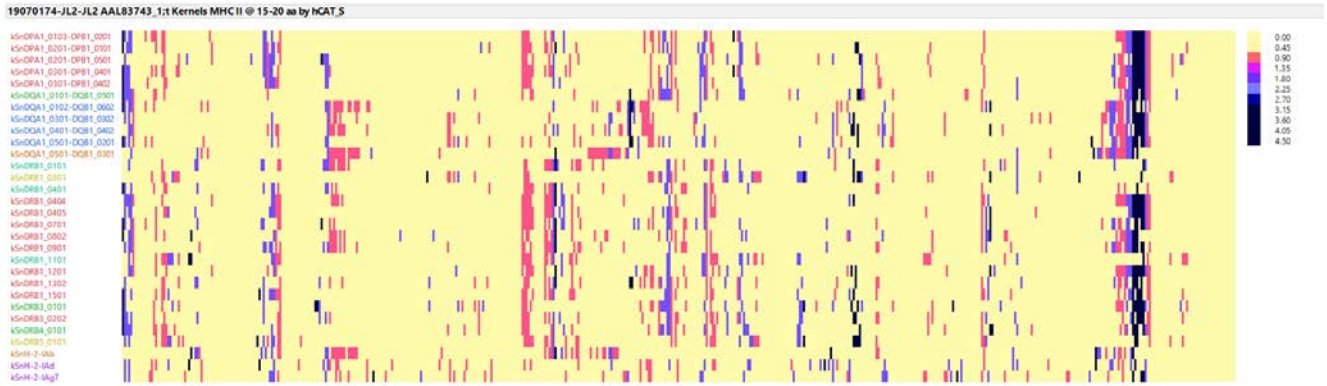
19070174-JL2-JL2 AAL83743_1| fusion protein

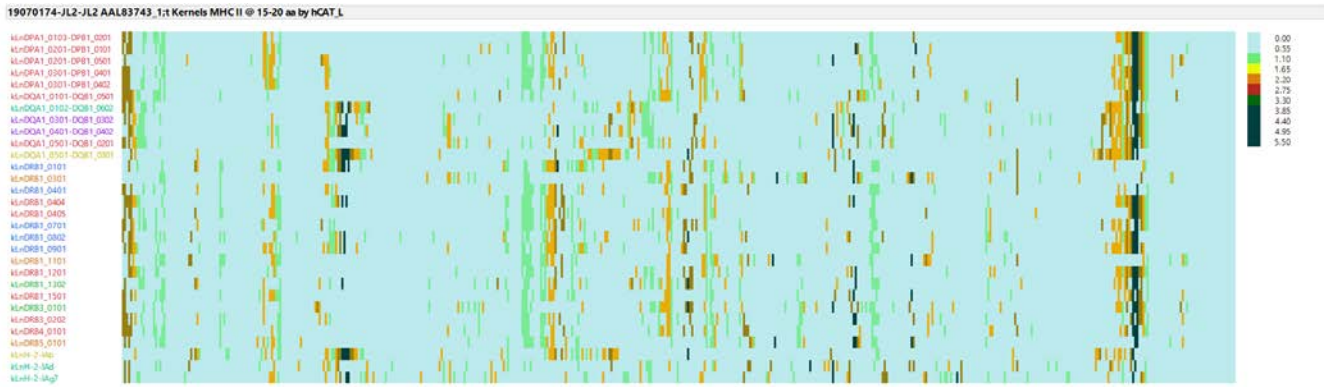
Overlay Plot

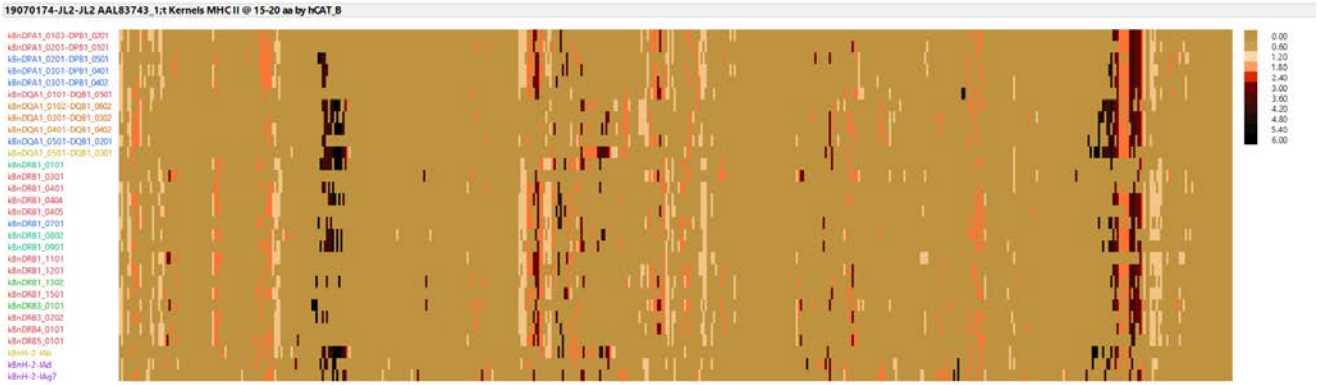
19070174-JL2-JL2 AAL83743_1:t

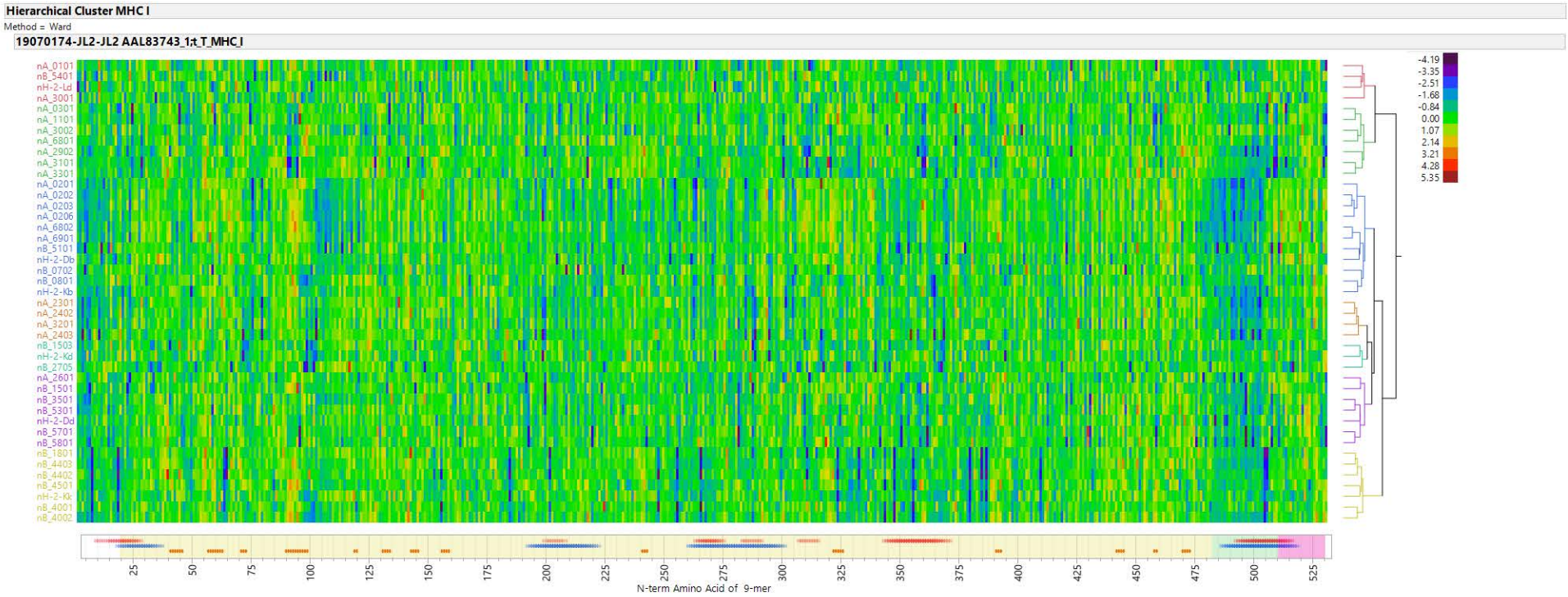


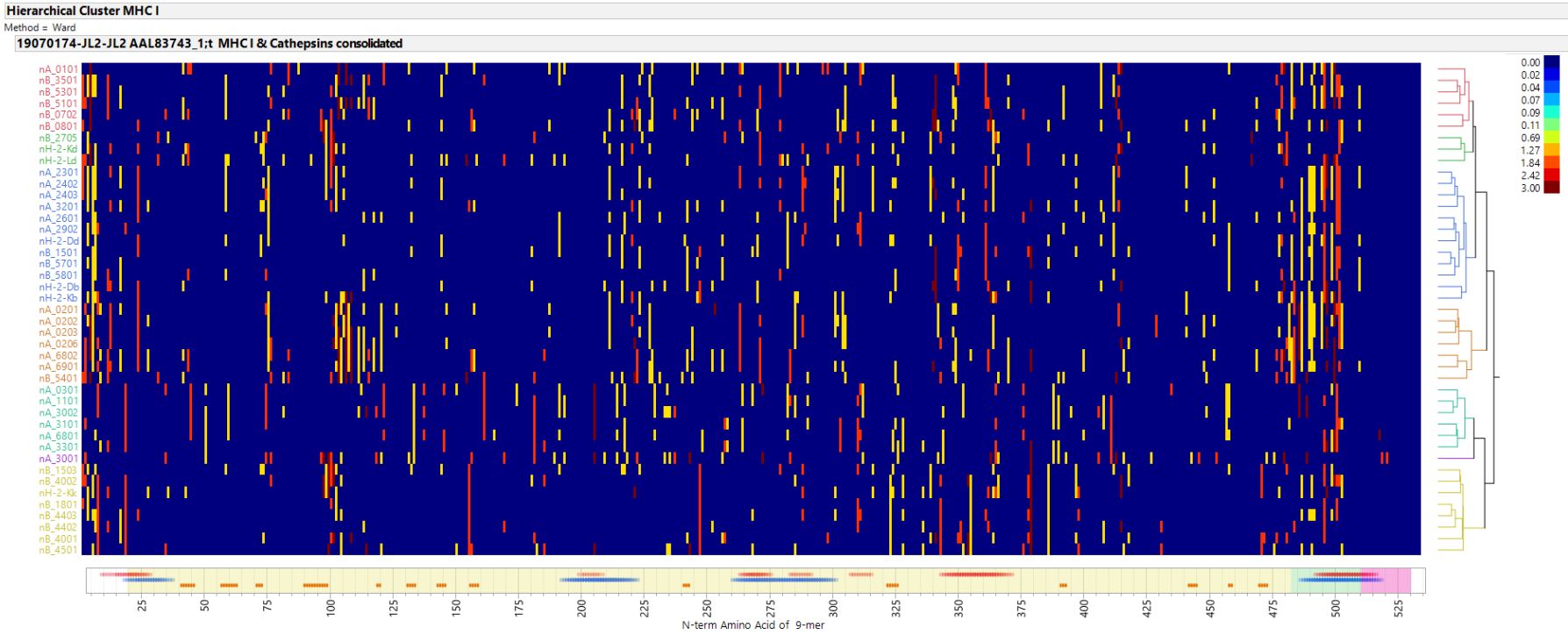

Hierarchical Cluster MHC II
Method = Ward
19070174-JL2-JL2 AAL83743_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0404
nDRB1_1201
nDRB1_0802
nDRB1_0401
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0501-DQB1_0301
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.63
-2.90
-2.18
-1.45
-0.73
-0.00
0.68
1.37
2.05
2.73
3.41
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

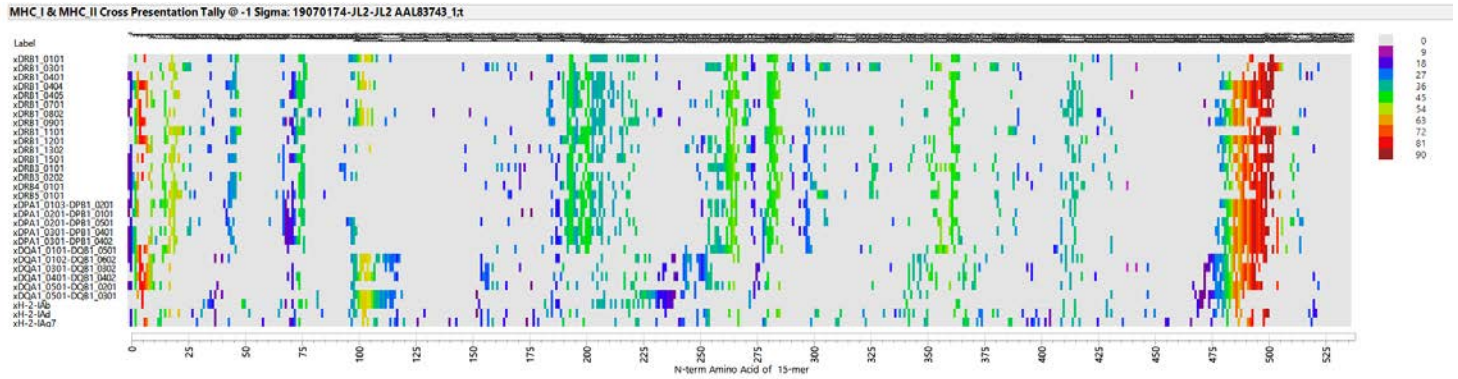

19070174-JL2-JL2 AAL83743_1;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-Ma
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50


19070174-JL2-JL2 AAL83743_1;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-Ma
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

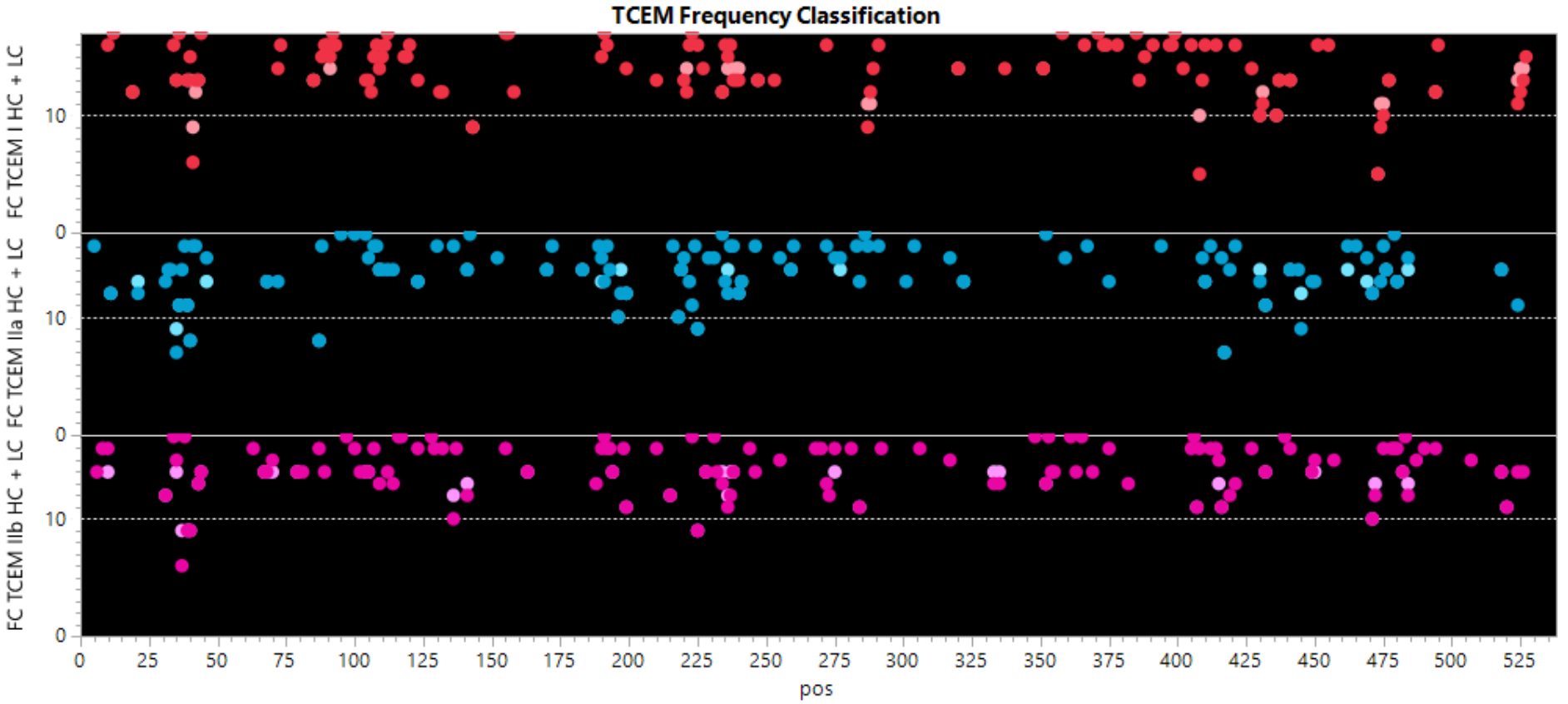

19070174-JL2-JL2 AAL83743_1;t Kernels MHC II @ 15-20 aa by hCAT_B
kBnDPA1_0103-DPB1_0201
kBnDPA1_0201-DPB1_0101
kBnDPA1_0201-DPB1_0501
kBnDPA1_0301-DPB1_0401
kBnDPA1_0301-DPB1_0402
kBnDQA1_0101-DQB1_0501
kBnDQA1_0102-DQB1_0602
kBnDQA1_0301-DQB1_0302
kBnDQA1_0401-DQB1_0402
kBnDQA1_0501-DQB1_0201
kBnDQA1_0501-DQB1_0301
kBnDRB1_0101
kBnDRB1_0301
kBnDRB1_0401
kBnDRB1_0404
kBnDRB1_0405
kBnDRB1_0701
kBnDRB1_0802
kBnDRB1_0901
kBnDRB1_1101
kBnDRB1_1201
kBnDRB1_1302
kBnDRB1_1501
kBnDRB3_0101
kBnDRB3_0202
kBnDRB4_0101
kBnDRB5_0101
kBnH-2-Ma
kBnH-2-IAd
kBnH-2-IAg7
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

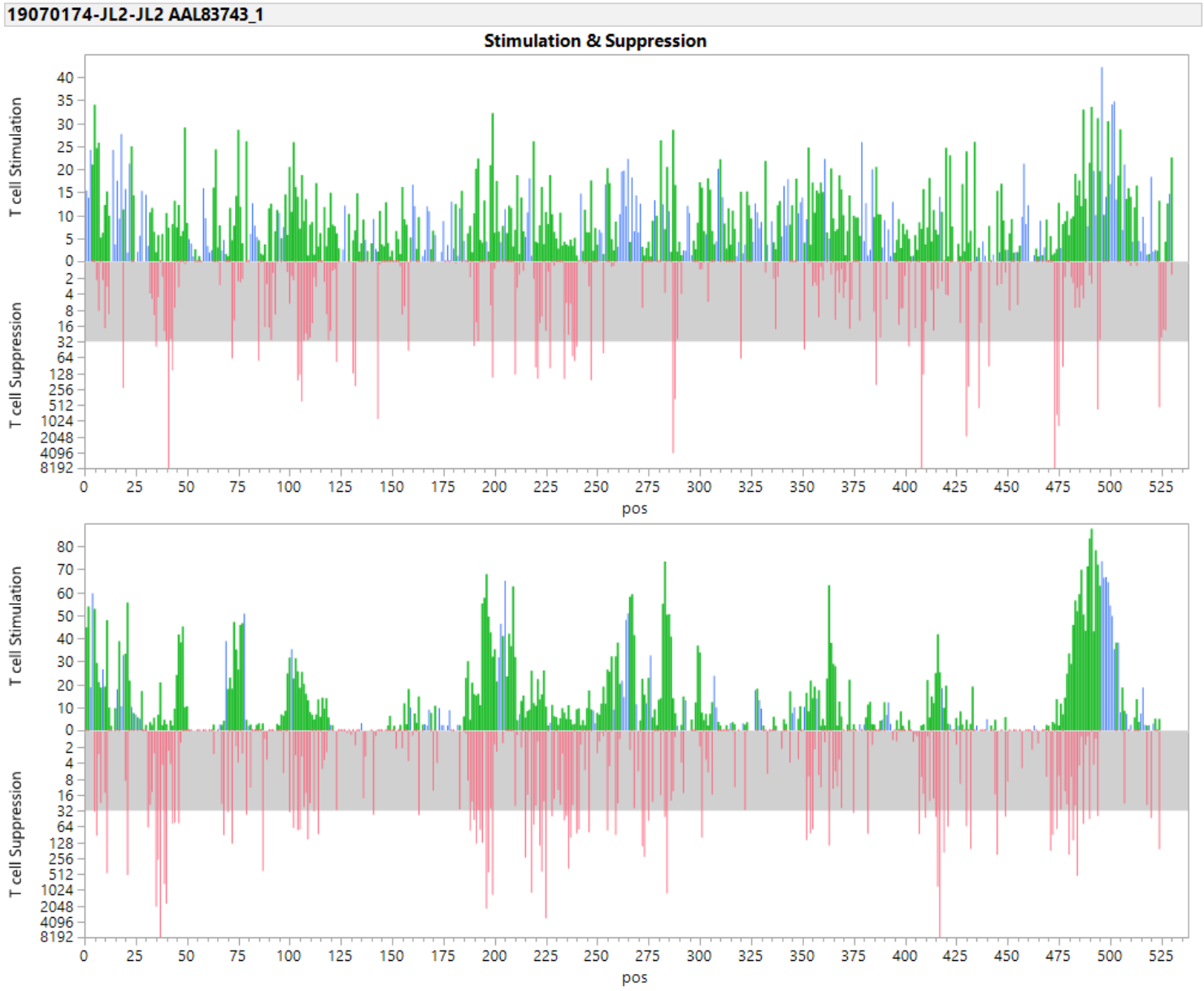

Hierarchical Cluster MHC I
Method = Ward
19070174-JL2-JL2 AAL83743_1;t_T_MHC_I
nA_0101
nB_5401
nH-2-Ld
nA_3001
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nB_1503
nH-2-Kd
nB_2705
nA_2601
nB_1501
nB_3501
nB_5301
nH-2-Dd
nB_5701
nB_5801
nB_1801
nB_4403
nB_4402
nB_4501
nH-2-Kk
nB_4001
nB_4002
-4.19
-3.35
-2.51
-1.68
-0.84
0.00
1.07
2.14
3.21
4.28
5.35
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
19070174-JL2-JL2 AAL83743_1;t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nB_5301
nB_5101
nB_0702
nB_0801
nB_2705
nH-2-Kd
nH-2-Ld
nA_2301
nA_2402
nA_2403
nA_3201
nA_2601
nA_2902
nH-2-Dd
nB_1501
nB_5701
nB_5801
nH-2-Db
nH-2-Kb
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5401
nA_0301
nA_1101
nA_3002
nA_3101
nA_6801
nA_3301
nA_3001
nB_1503
nB_4002
nH-2-Kk
nB_1801
nB_4403
nB_4402
nB_4001
nB_4501
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070174-JL2-JL2 AAL83743_1:t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer
0
9
18
27
36
45
54
63
72
81
90

19070174-JL2-JL2 AAL83743_1



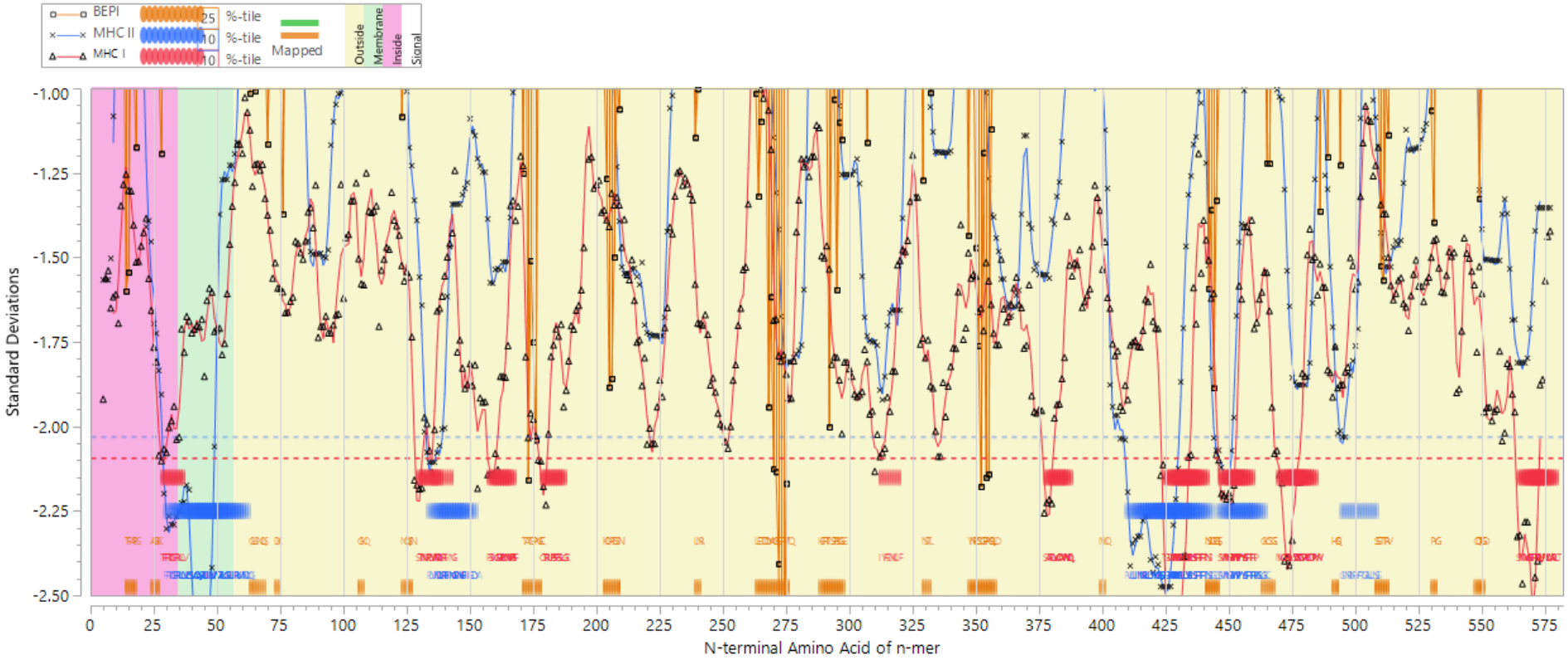

19070174-JL2-JL2 AAL83743_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos
T cell Stimulation
T cell Suppression
pos

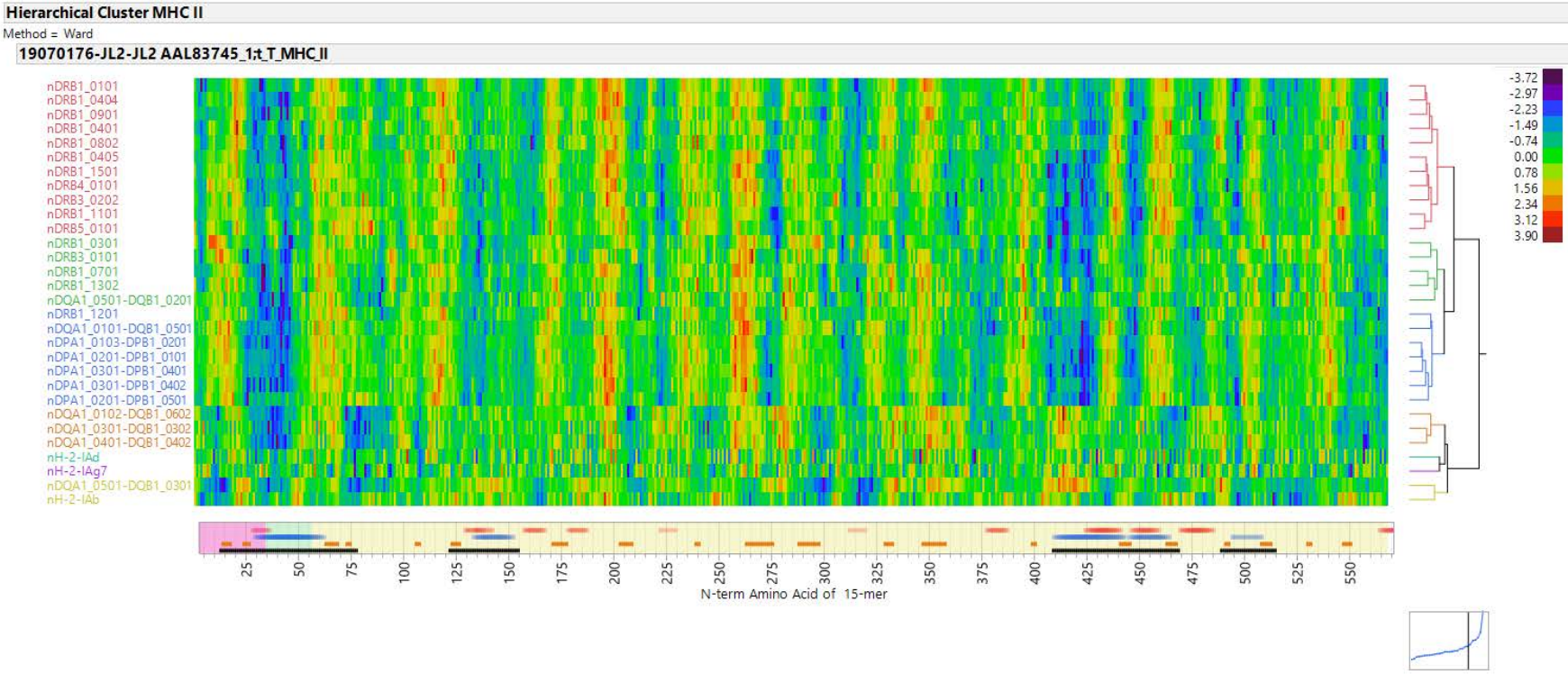
Hemagglutinin/Neuraminidase

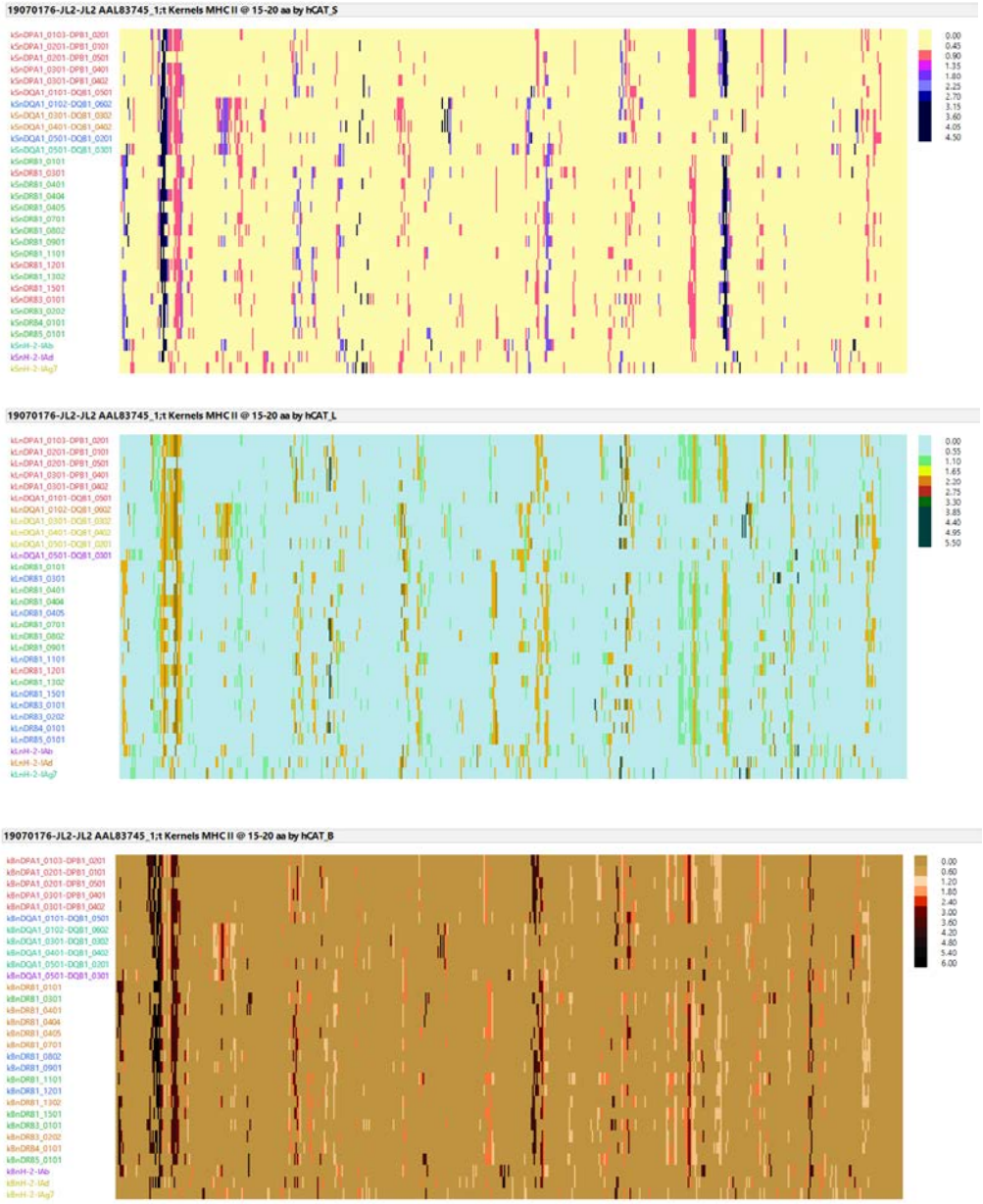
19070176-JL2-JL2 AAL83745_1| hemagglutinin/neuraminidase

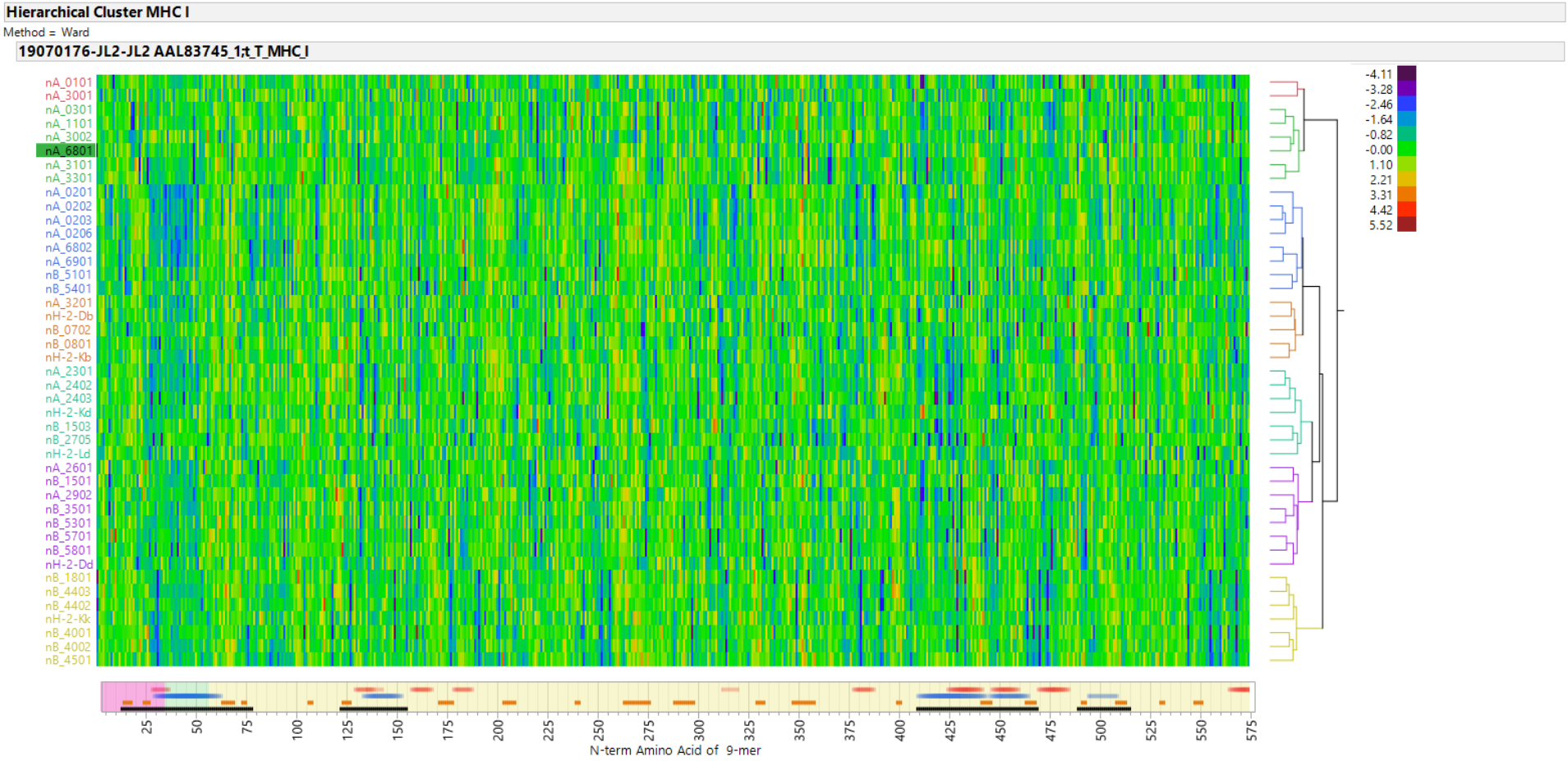
Overlay Plot

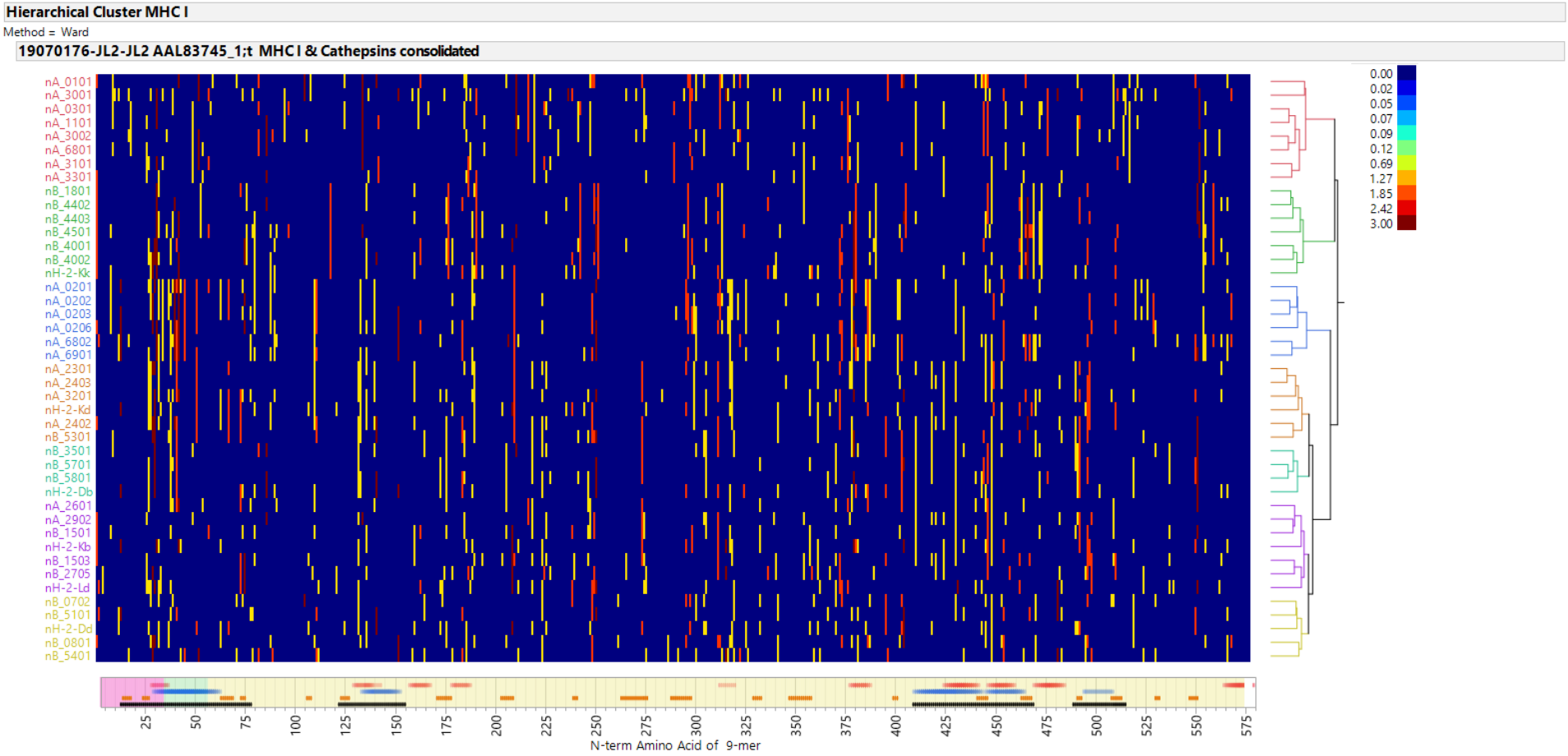
19070176-JL2-JL2 AAL83745_1;t

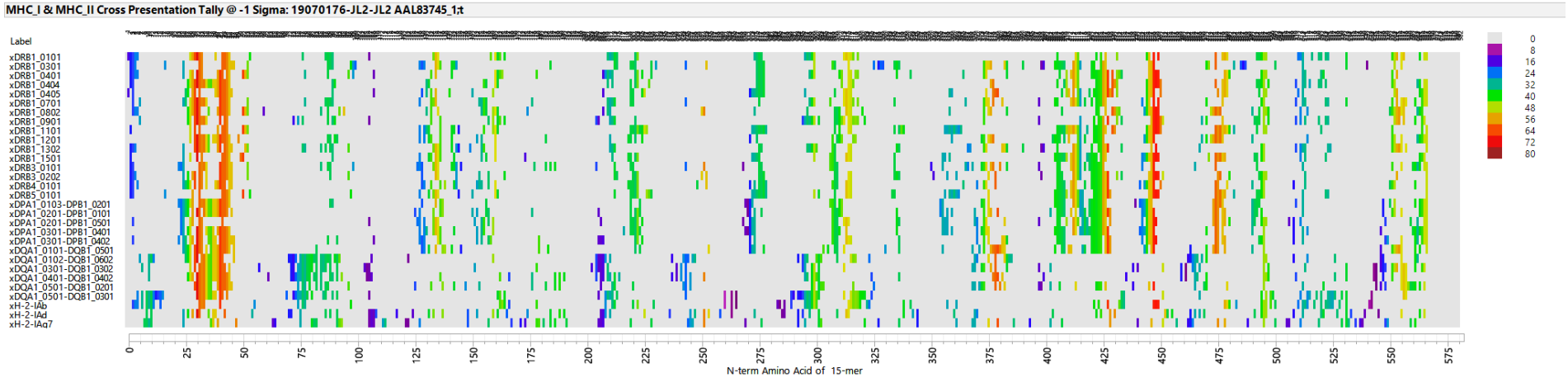



Hierarchical Cluster MHC II
Method = Ward
19070176-JL2-JL2 AAL83745_1;t_T_MHC_II
nDRB1_0101
nDRB1_0404
nDRB1_0901
nDRB1_0401
nDRB1_0802
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDRB1_0701
nDRB1_1302
nDQA1_0501-DQB1_0201
nDRB1_1201
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.72
-2.97
-2.23
-1.49
-0.74
0.00
0.78
1.56
2.34
3.12
3.90
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 15-mer

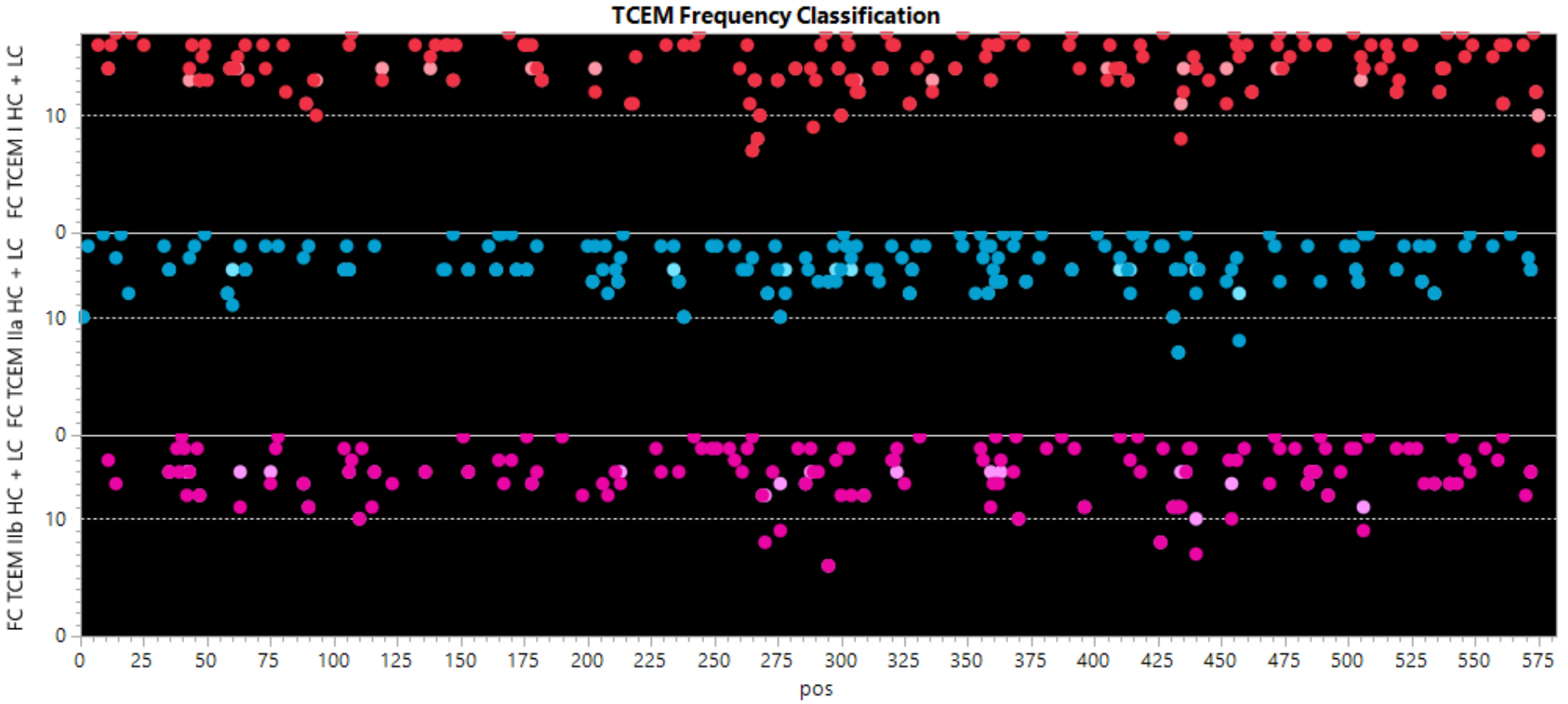

19070176-JL2-JL2 AAL83745_1;t Kernels MHC II @ 15-20 aa by hCAT_S
19070176-JL2-JL2 AAL83745_1;t Kernels MHC II @ 15-20 aa by hCAT_L
19070176-JL2-JL2 AAL83745_1;t Kernels MHC II @ 15-20 aa by hCAT_B

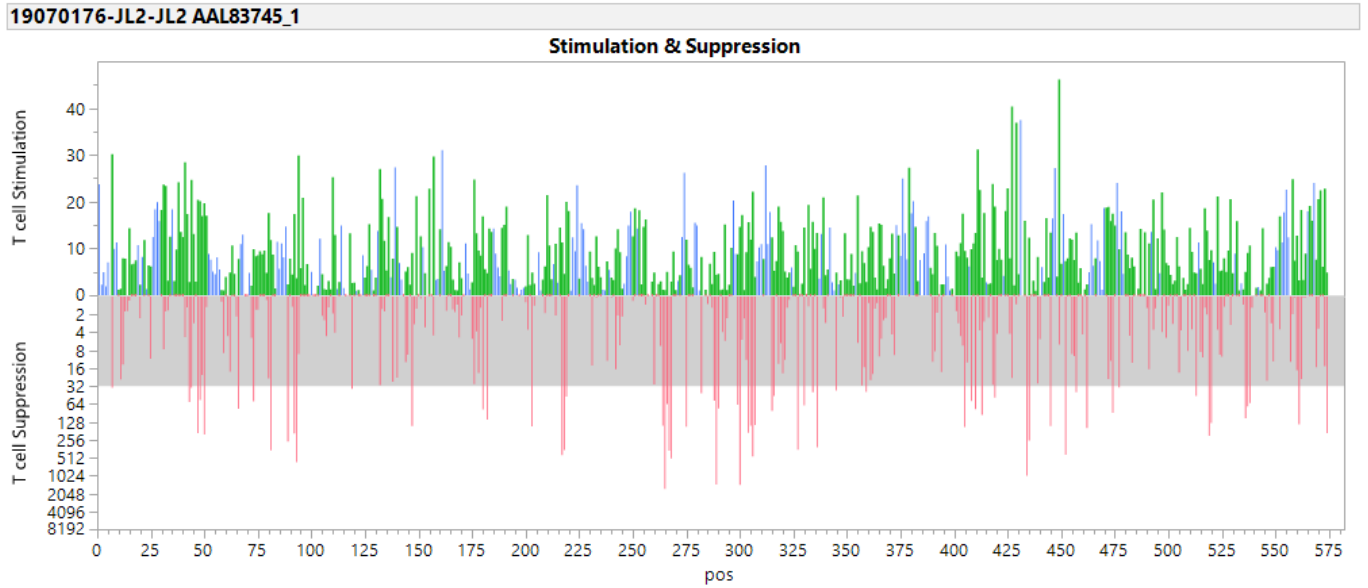

Hierarchical Cluster MHC I
Method = Ward
19070176-JL2-JL2 AAL83745_1;t_T_MHC_I
nA_0101
nA_3001
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nB_5401
nA_3201
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nH-2-Ld
nA_2601
nB_1501
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nB_1801
nB_4403
nB_4402
nH-2-Kk
nB_4001
nB_4002
nB_4501
-4.11
-3.28
-2.46
-1.64
-0.82
-0.00
1.10
2.21
3.31
4.42
5.52
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
N-term Amino Acid of 9-mer

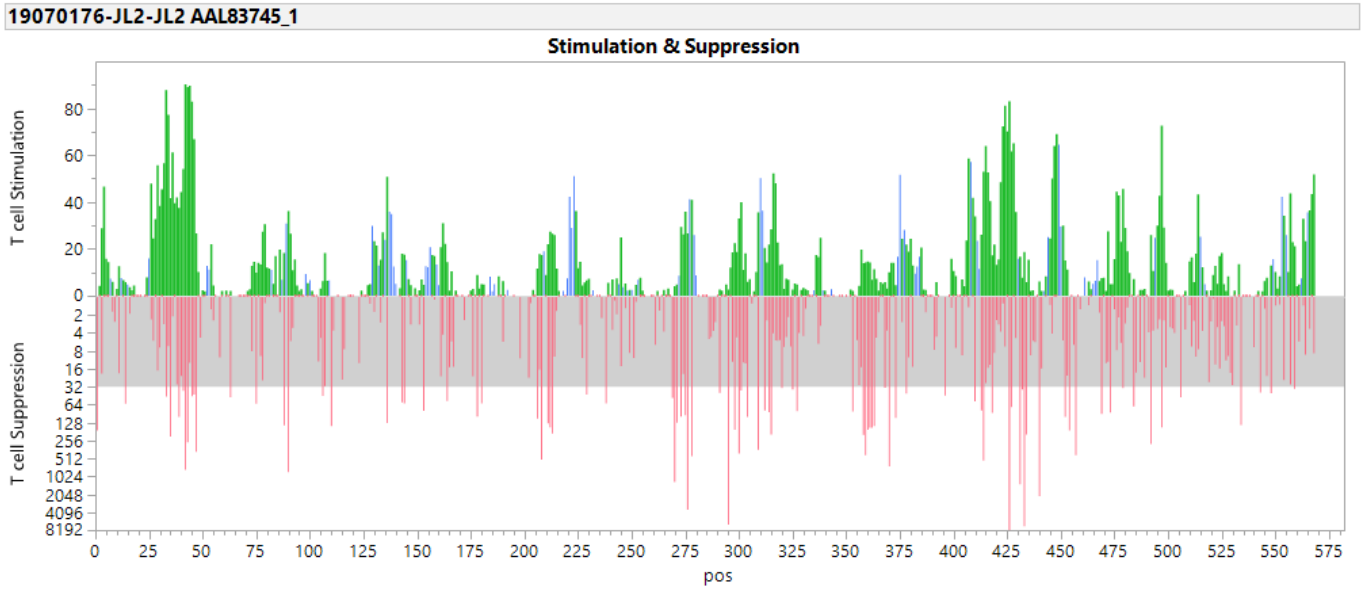

Hierarchical Cluster MHC I
Method = Ward
19070176-JL2-JL2 AAL83745_1;t MHC I & Cathepsins consolidated
nA_0101
nA_3001
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nB_1801
nB_4402
nB_4403
nB_4501
nB_4001
nB_4002
nH-2-Kk
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_2301
nA_2403
nA_3201
nH-2-Kd
nA_2402
nB_5301
nB_3501
nB_5701
nB_5801
nH-2-Db
nA_2601
nA_2902
nB_1501
nH-2-Kb
nB_1503
nB_2705
nH-2-Ld
nB_0702
nB_5101
nH-2-Dd
nB_0801
nB_5401
0.00
0.02
0.05
0.07
0.09
0.12
0.69
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070176-JL2-JL2 AAL83745_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

19070176-JL2-JL2 AAL83745_1




19070176-JL2-JL2 AAL83745_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

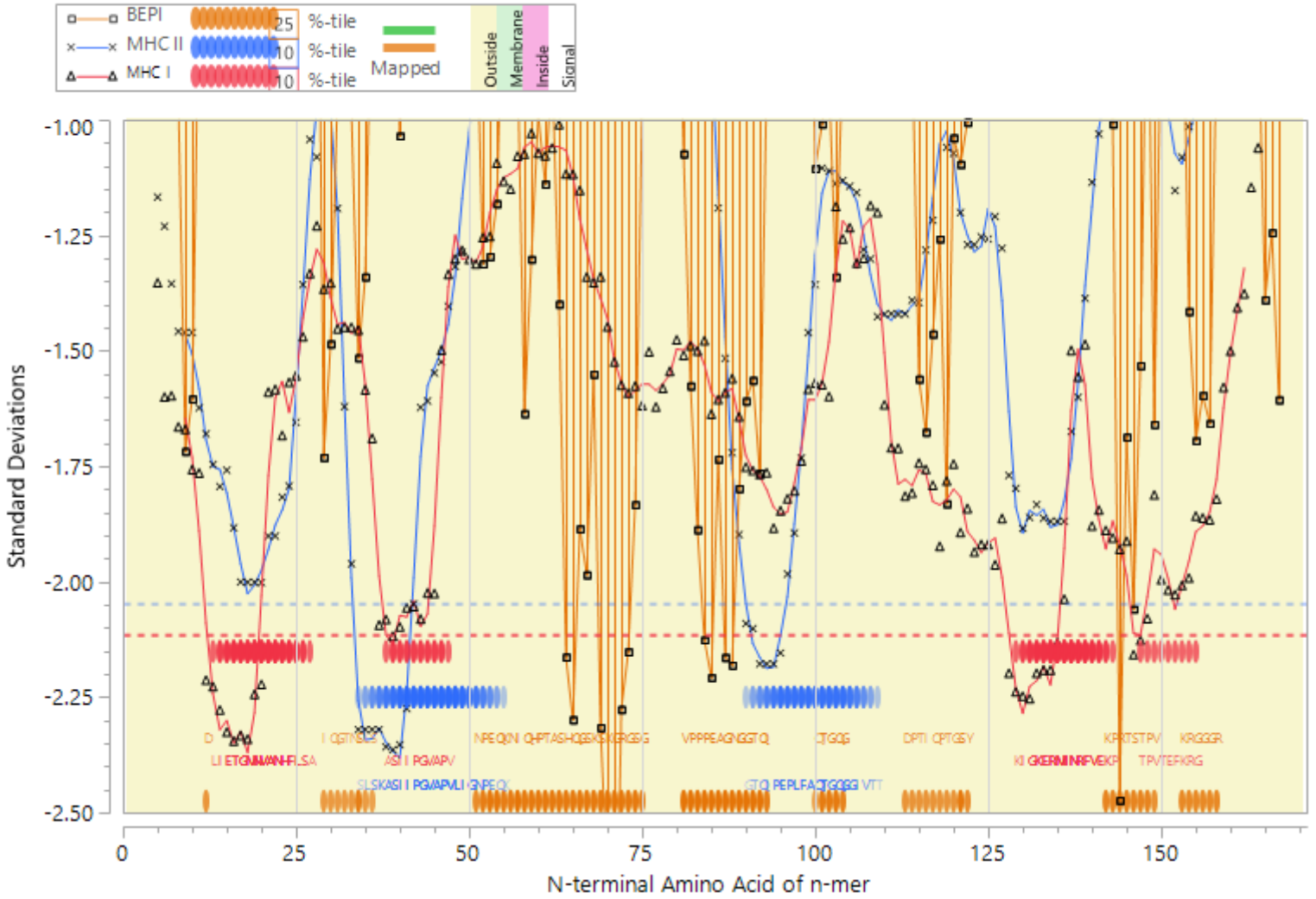

19070176-JL2-JL2 AAL83745_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

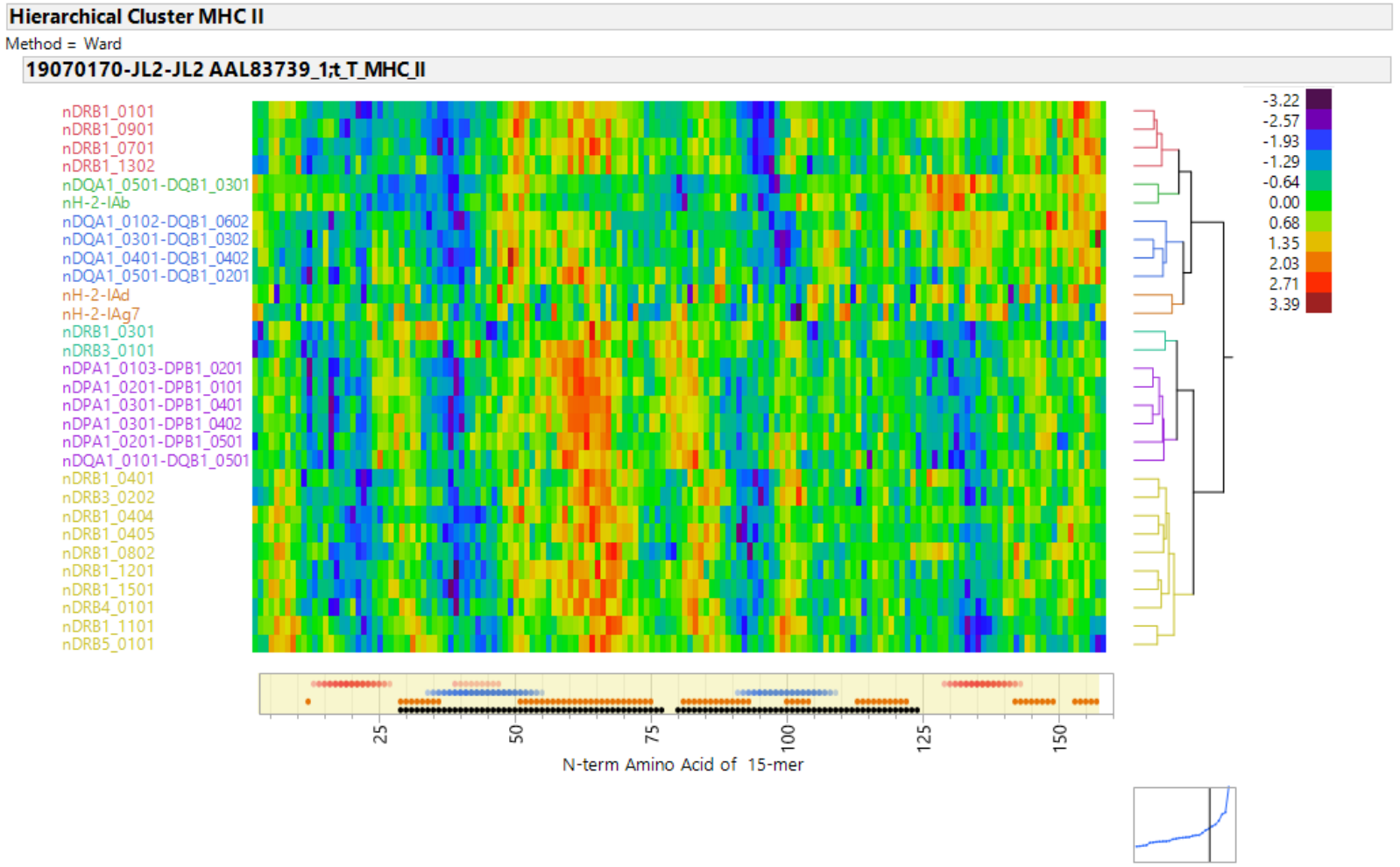
I protein

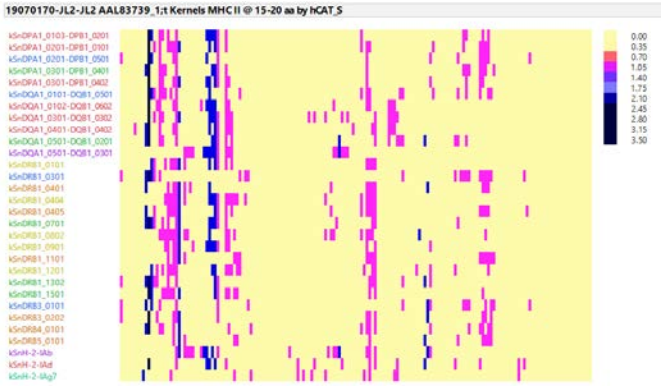
19070170-JL2-JL2 AAL83739_1| I protein

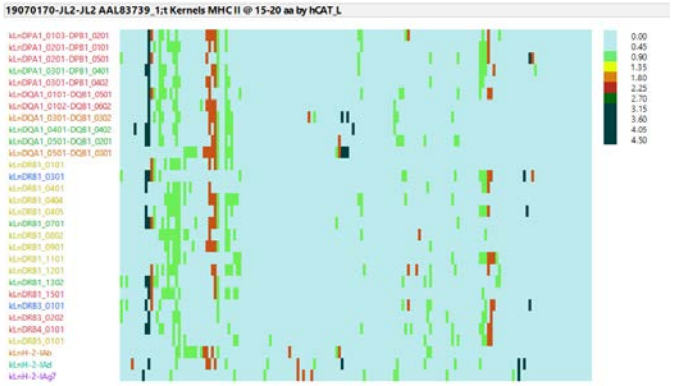
Overlay Plot

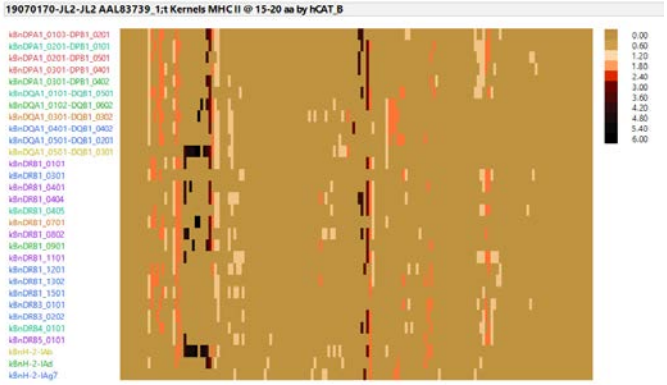
19070170-JL2-JL2 AAL83739_1;t

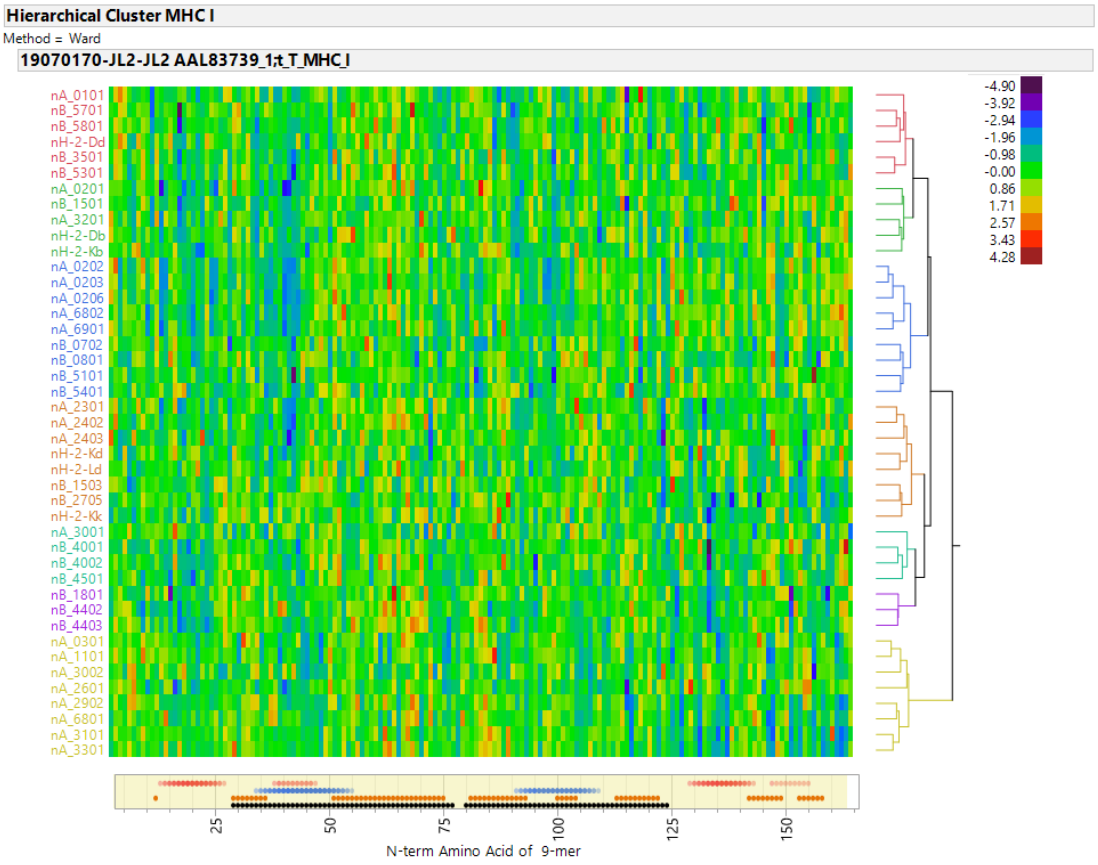


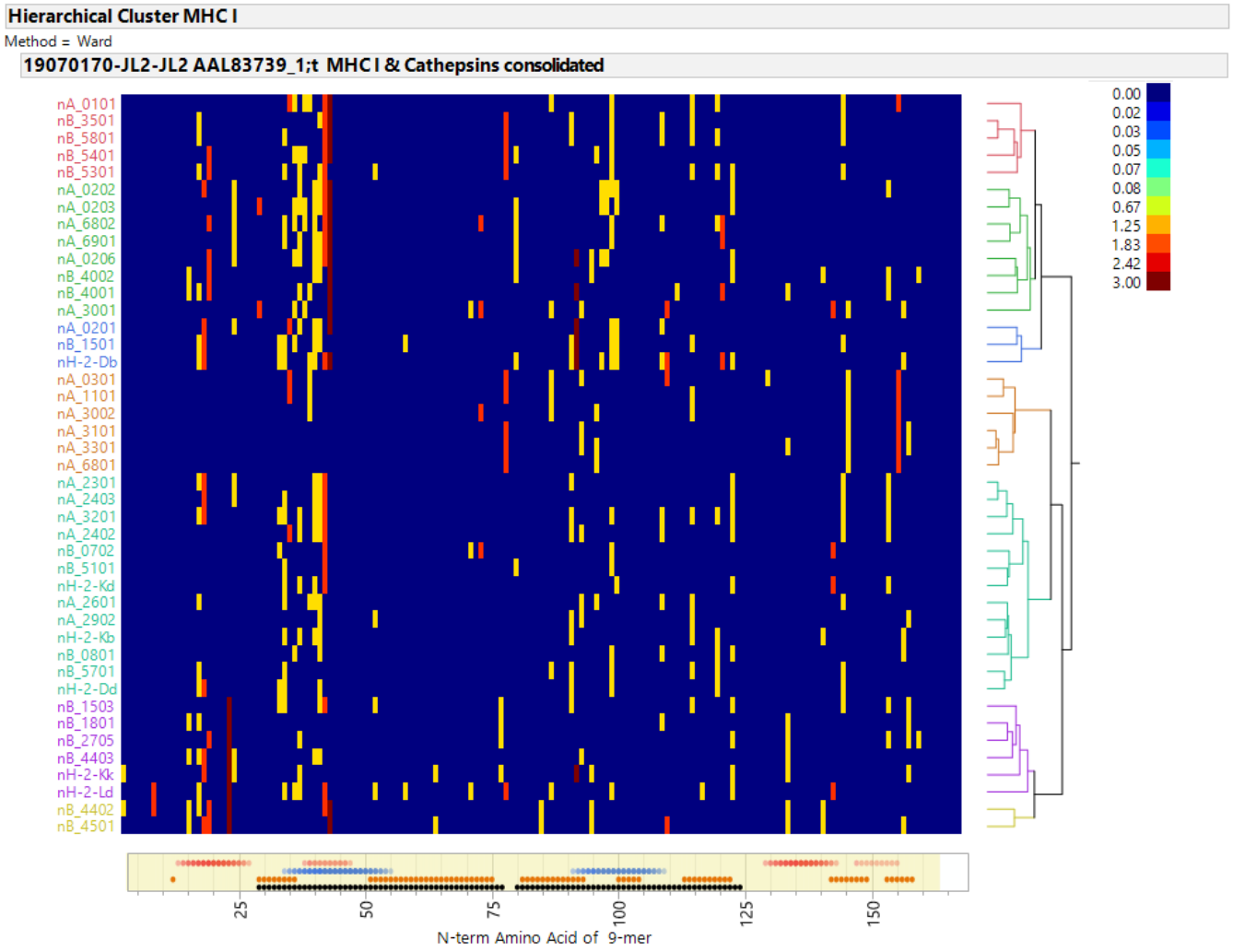

Hierarchical Cluster MHC II
Method = Ward
19070170-JL2-JL2 AAL83739_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDQA1_0501-DQB1_0301
nH-2-IAb
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDRB1_0401
nDRB3_0202
nDRB1_0404
nDRB1_0405
nDRB1_0802
nDRB1_1201
nDRB1_1501
nDRB4_0101
nDRB1_1101
nDRB5_0101
-3.22
-2.57
-1.93
-1.29
-0.64
0.00
0.68
1.35
2.03
2.71
3.39
25
50
75
100
125
150
N-term Amino Acid of 15-mer

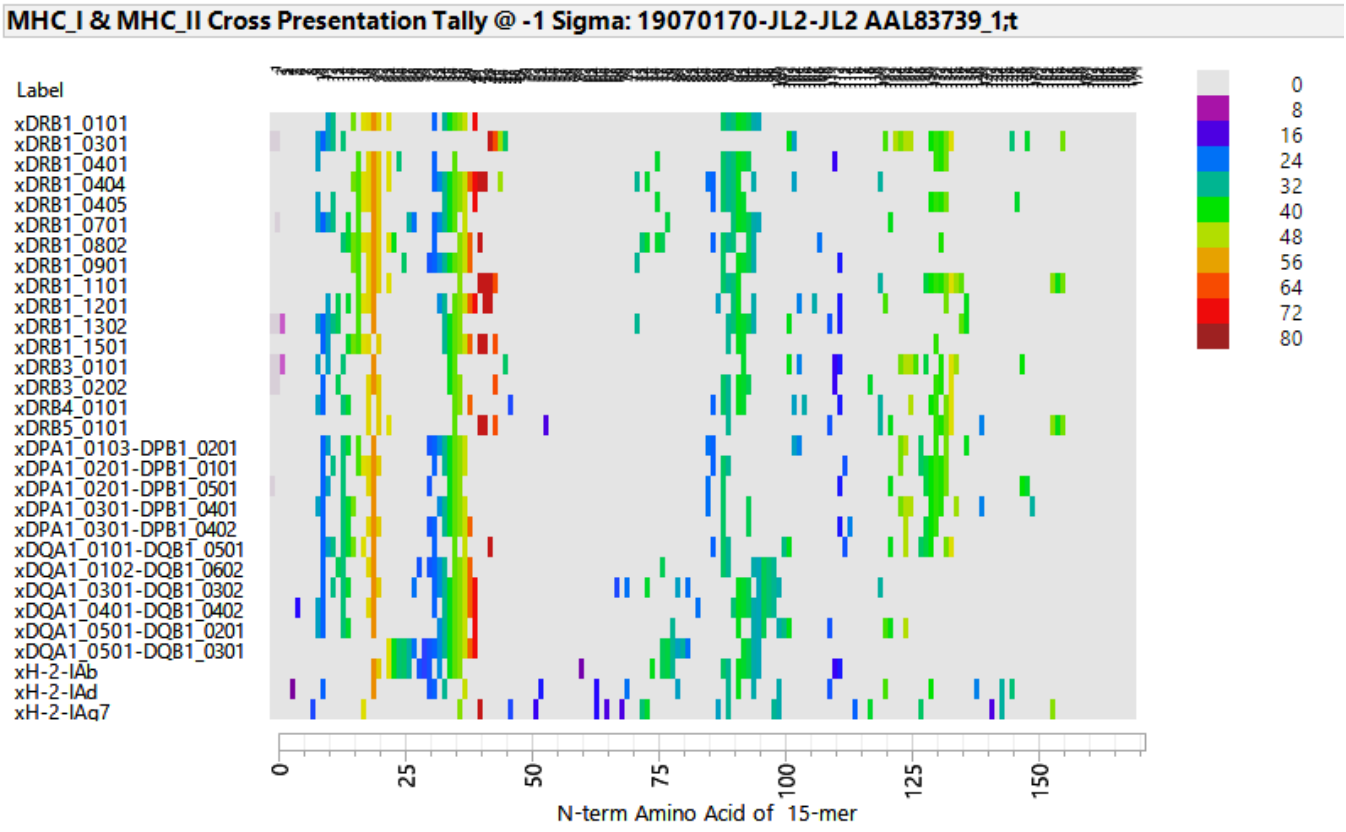

19070170-JL2-JL2 AAL83739_1;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.35
0.70
1.05
1.40
1.75
2.10
2.45
2.80
3.15
3.50


19070170-JL2-JL2 AAL83739_1;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

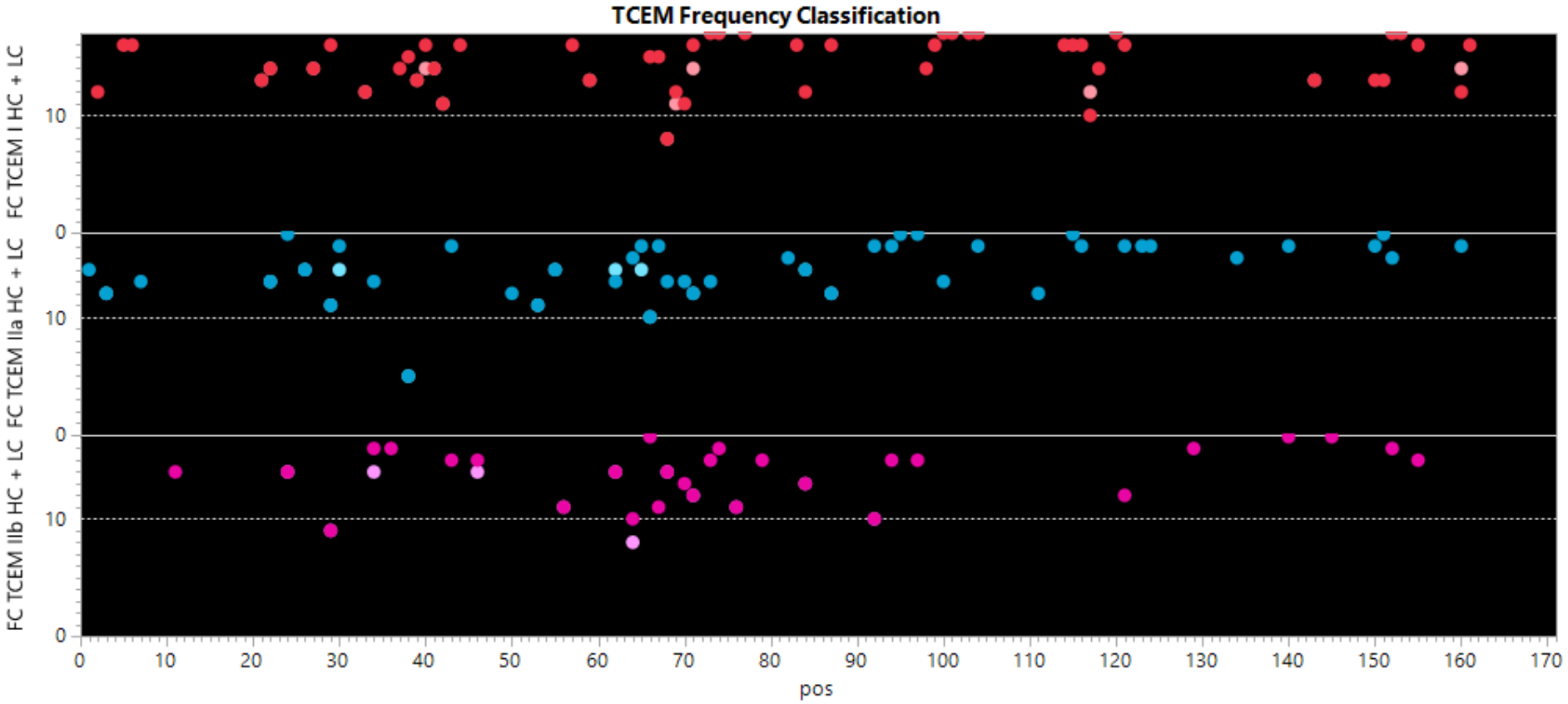

19070170-JL2-JL2 AAL83739_1;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

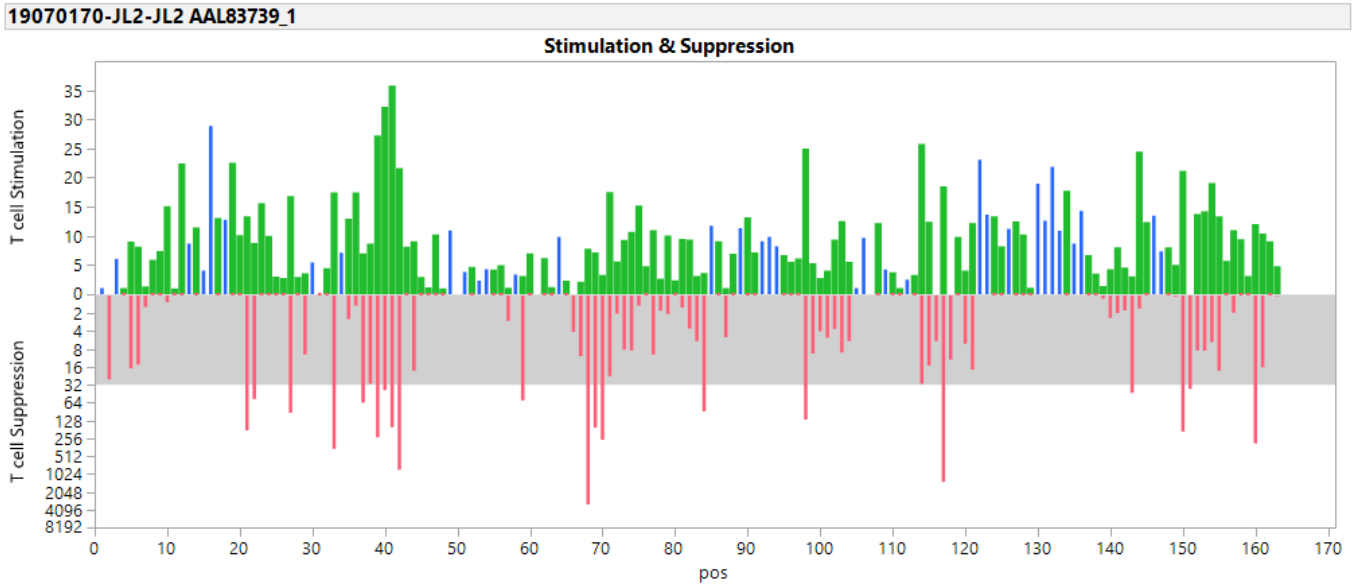

Hierarchical Cluster MHC I
Method = Ward
19070170-JL2-JL2 AAL83739_1;t_T_MHC_I
nA_0101
nB_5701
nB_5801
nH-2-Dd
nB_3501
nB_5301
nA_0201
nB_1501
nA_3201
nH-2-Db
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_0801
nB_5101
nB_5401
nA_2301
nA_2402
nA_2403
nH-2-Kd
nH-2-Ld
nB_1503
nB_2705
nH-2-Kk
nA_3001
nB_4001
nB_4002
nB_4501
nB_1801
nB_4402
nB_4403
nA_0301
nA_1101
nA_3002
nA_2601
nA_2902
nA_6801
nA_3101
nA_3301
-4.90
-3.92
-2.94
-1.96
-0.98
-0.00
0.86
1.71
2.57
3.43
4.28
25
50
75
100
125
150
N-term Amino Acid of 9-mer

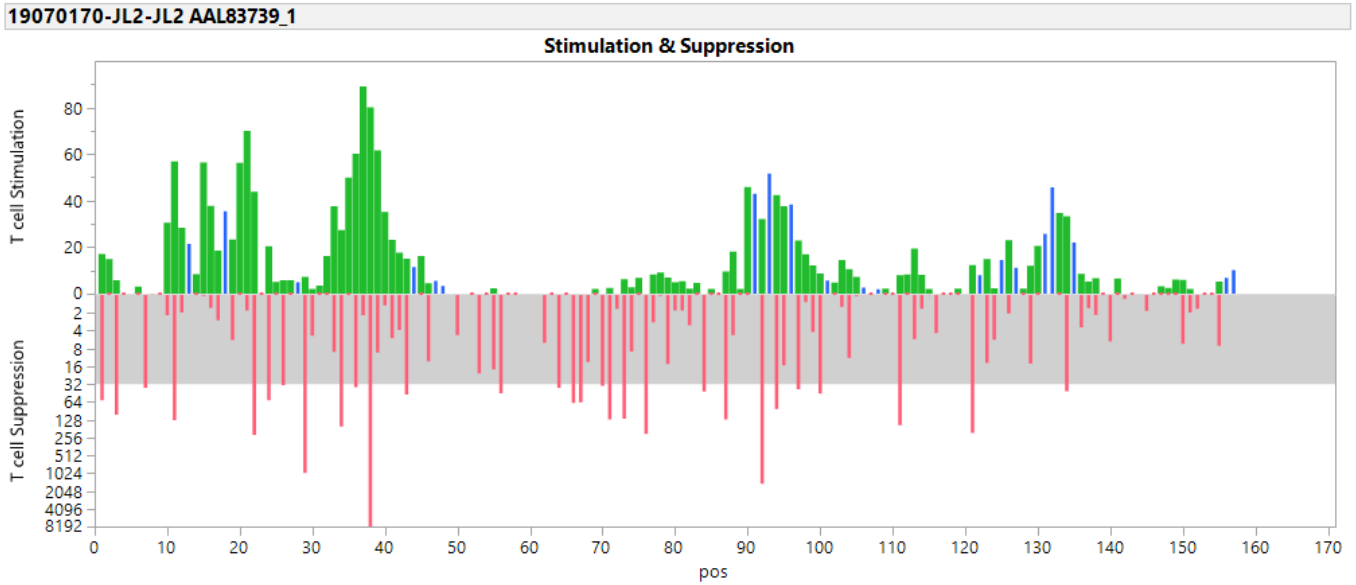

Hierarchical Cluster MHC I
Method = Ward
19070170-JL2-JL2 AAL83739_1;t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nB_5801
nB_5401
nB_5301
nA_0202
nA_0203
nA_6802
nA_6901
nA_0206
nB_4002
nB_4001
nA_3001
nA_0201
nB_1501
nH-2-Db
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_2301
nA_2403
nA_3201
nA_2402
nB_0702
nB_5101
nH-2-Kd
nA_2601
nA_2902
nH-2-Kb
nB_0801
nB_5701
nH-2-Dd
nB_1503
nB_1801
nB_2705
nB_4403
nH-2-Kk
nH-2-Ld
nB_4402
nB_4501
0.00
0.02
0.03
0.05
0.07
0.08
0.67
1.25
1.83
2.42
3.00
25
50
75
100
125
150
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070170-JL2-JL2 AAL83739_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
8
16
24
32
40
48
56
64
72
80
0
25
50
75
100
125
150
N-term Amino Acid of 15-mer

19070170-JL2-JL2 AAL83739_1

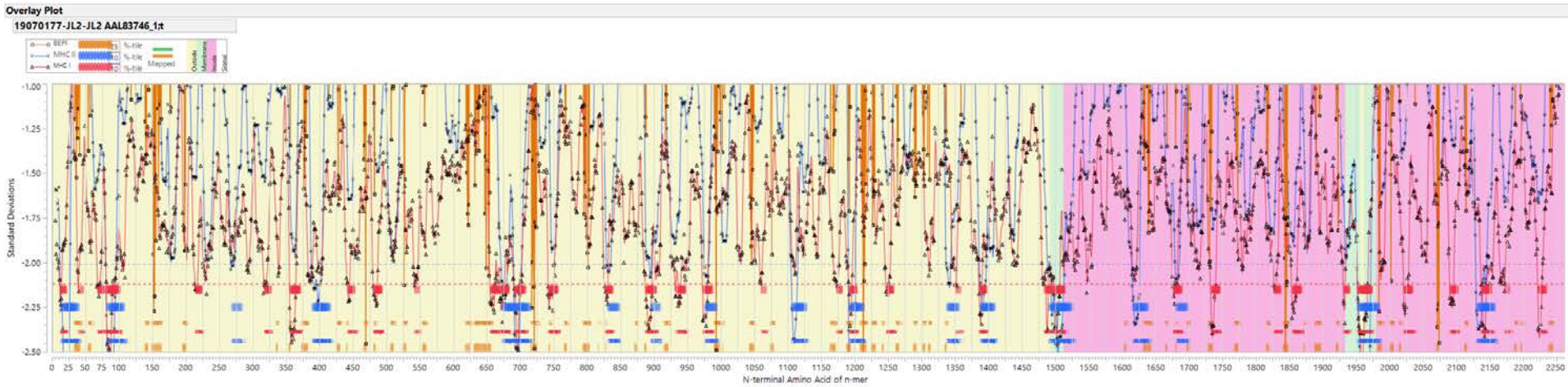


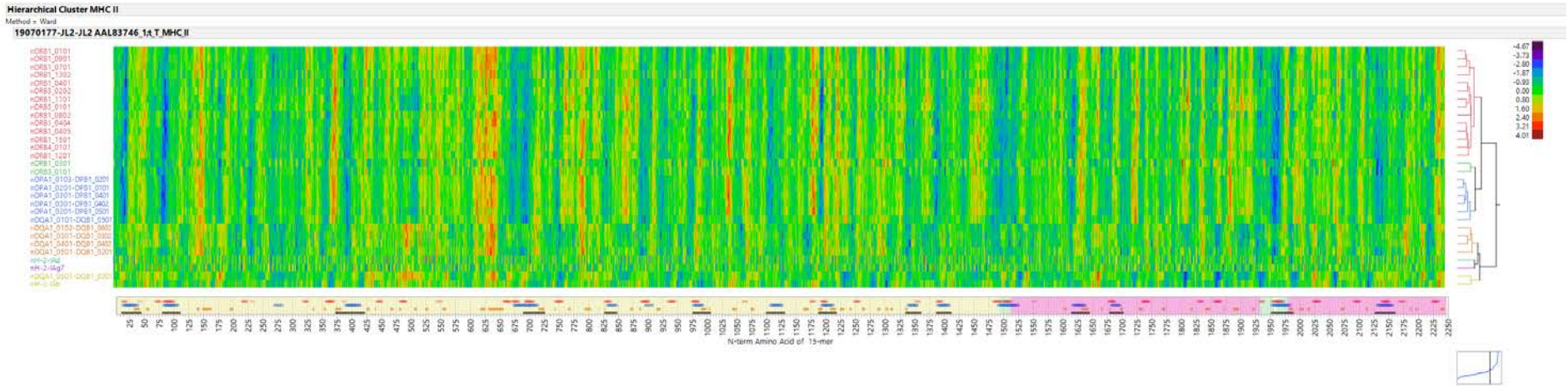

19070170-JL2-JL2 AAL83739_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

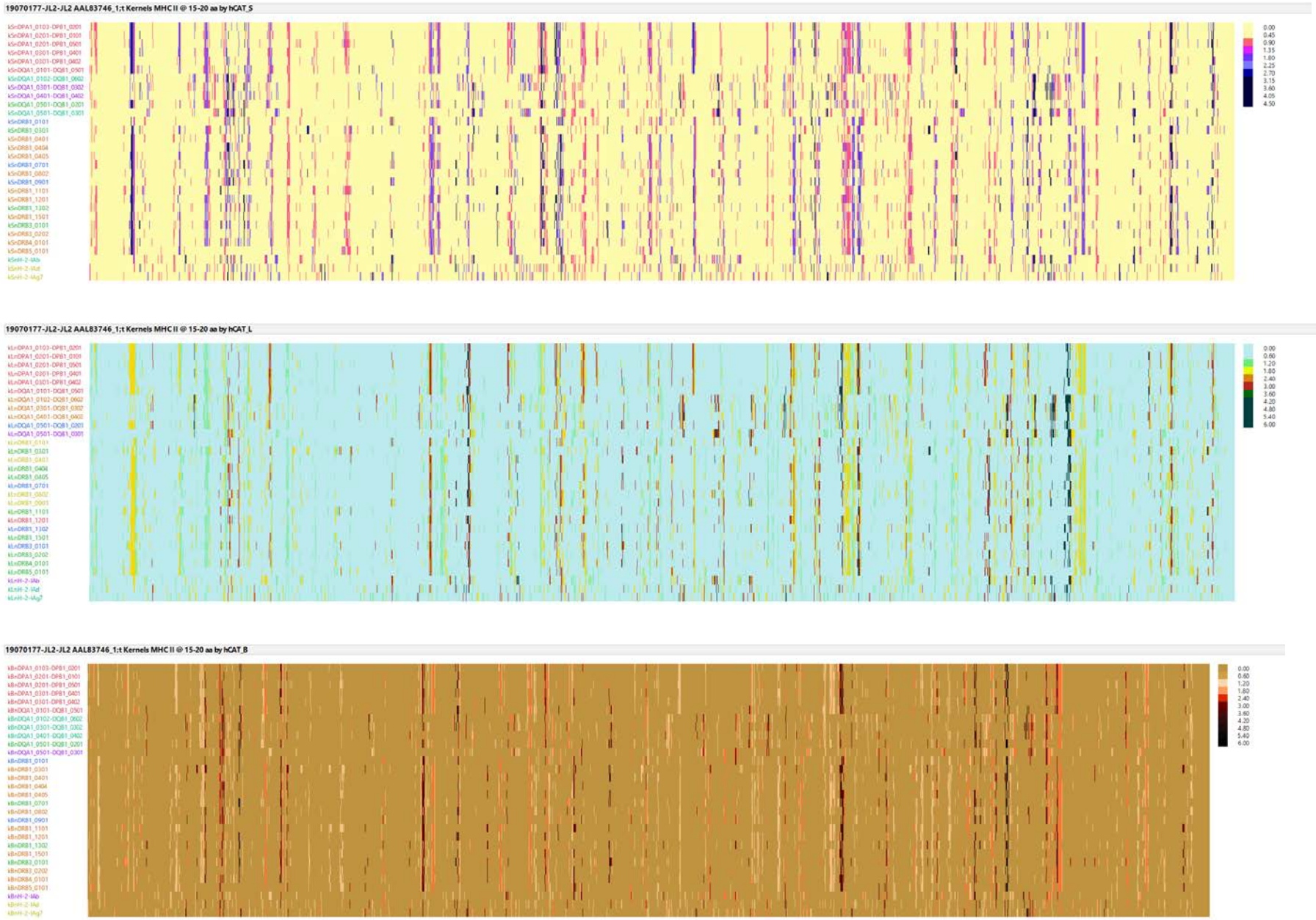

19070170-JL2-JL2 AAL83739_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

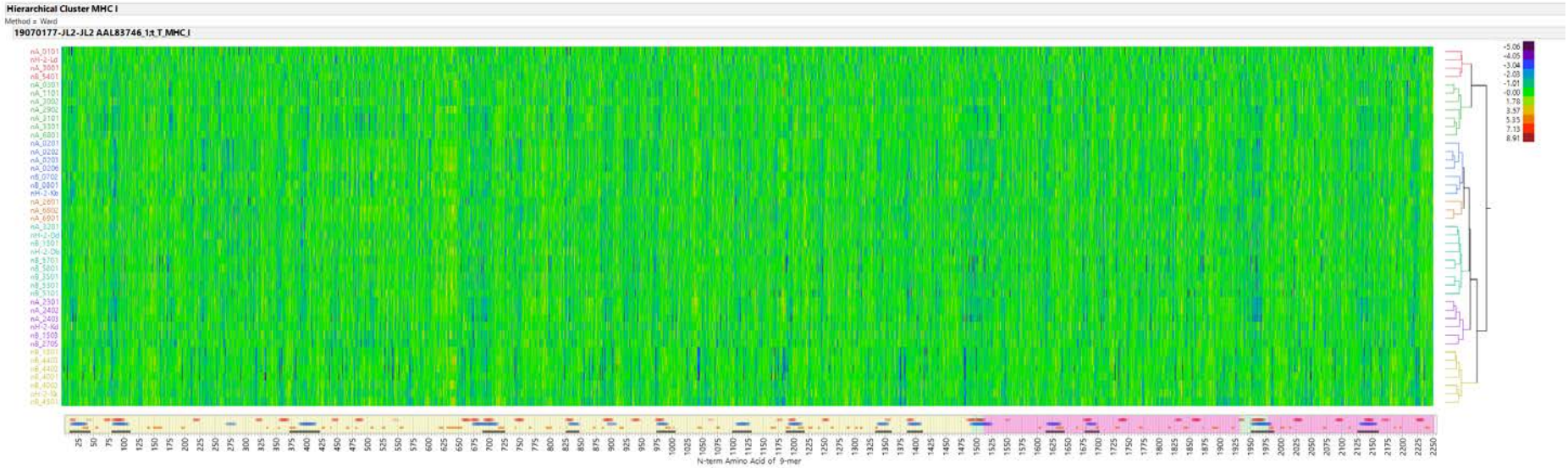
Large protein

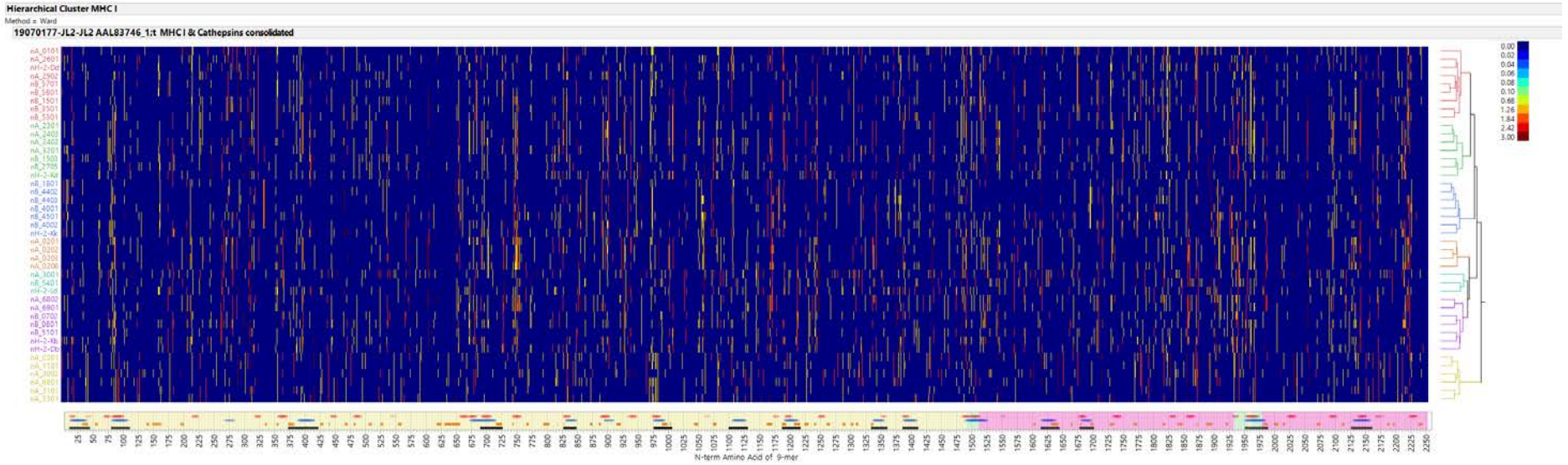
19070177-JL2-JL2 AAL83746_1| large protein

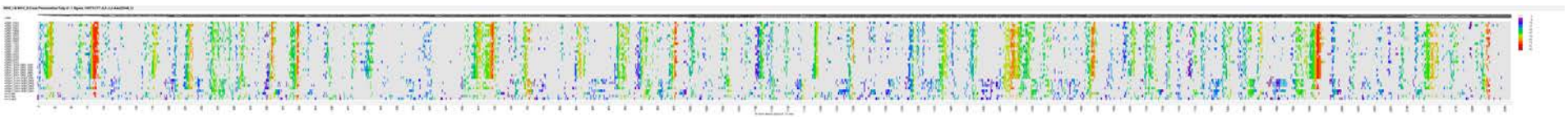

Overlay Plot
19070177-JL2-JL2 AAL83746_1;t
BEPI
MHC II
MHC I
%-tile
%-tile
%-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0 25 50 75 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250 1300 1350 1400 1450 1500 1550 1600 1650 1700 1750 1800 1850 1900 1950 2000 2050 2100 2150 2200 2250
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
19070177-JL2-JL2 AAL83746_1:t_T_MHC_II
N-term Amino Acid of 15-mer

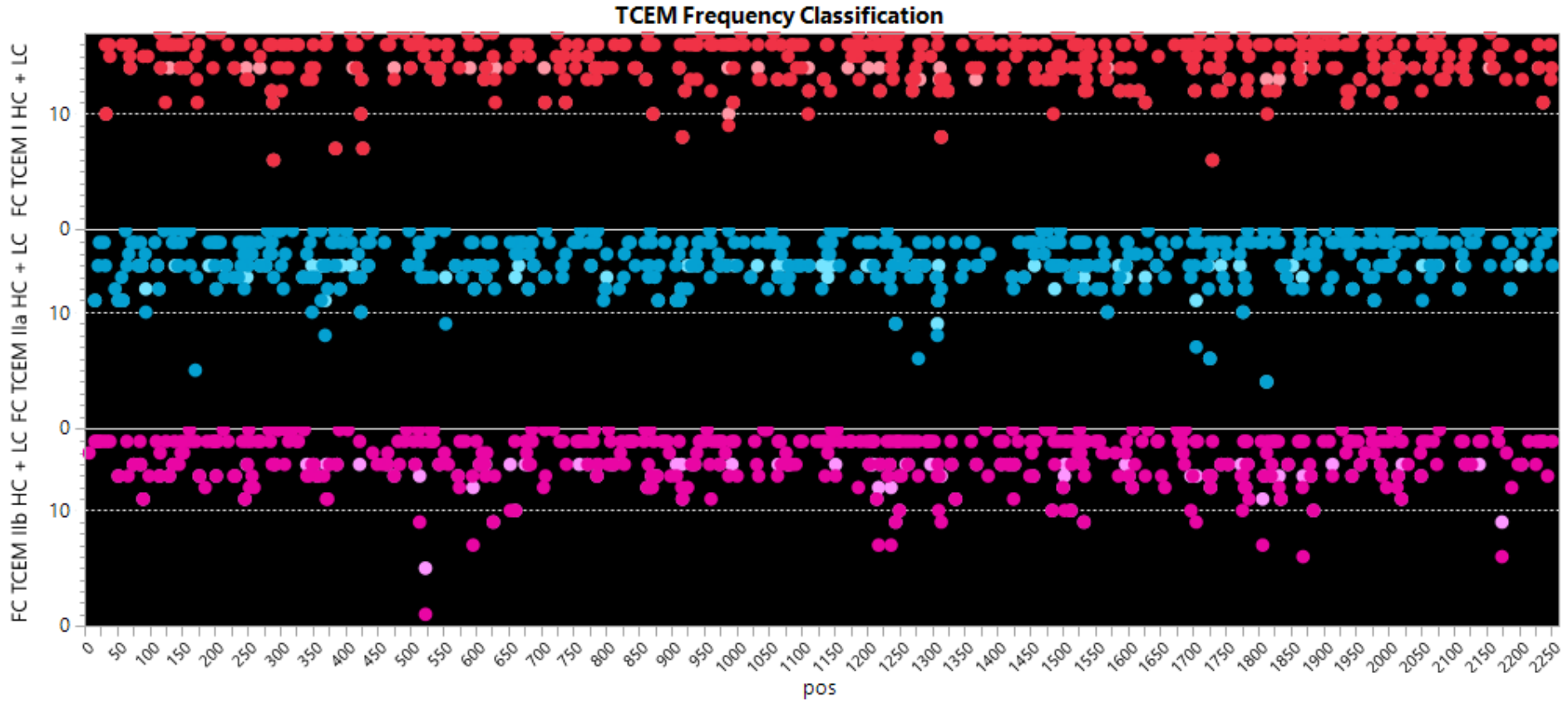

19070177-JL2-JL2 AAL83746_1:t Kernels MHC II @ 15-20 aa by hCAT_S
19070177-JL2-JL2 AAL83746_1:t Kernels MHC II @ 15-20 aa by hCAT_L
19070177-JL2-JL2 AAL83746_1:t Kernels MHC II @ 15-20 aa by hCAT_B

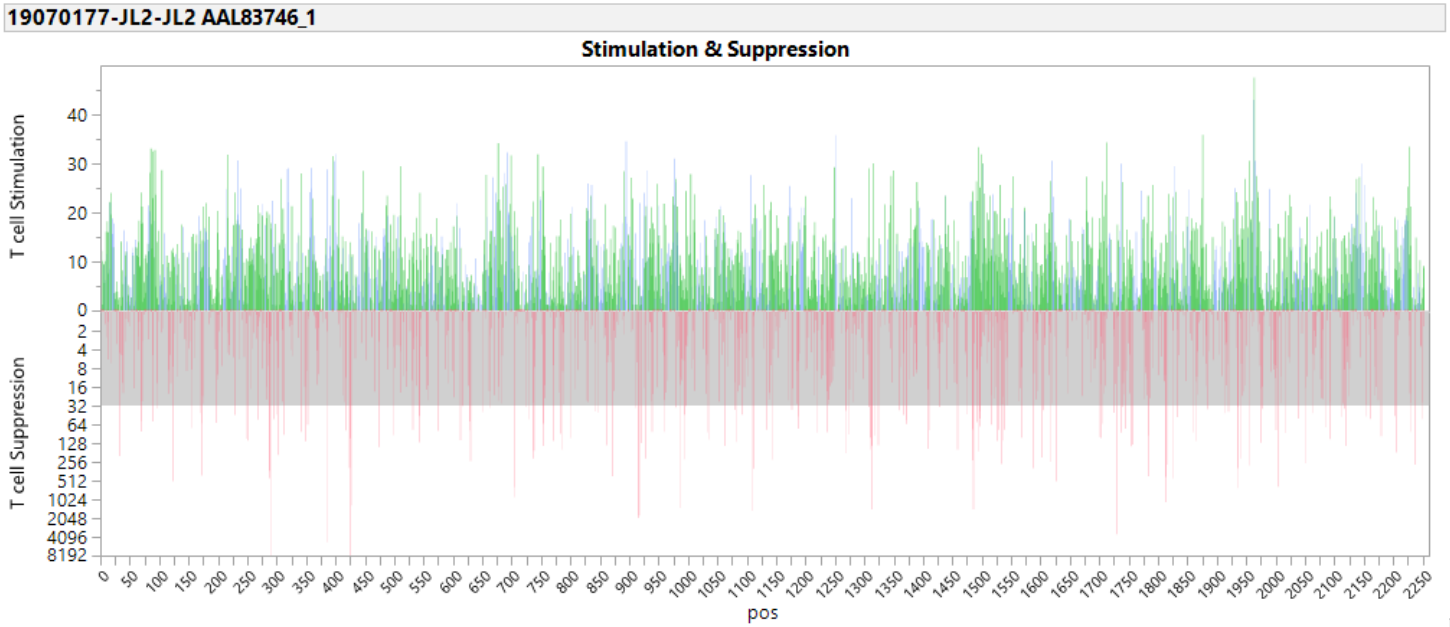

Hierarchical Cluster MHC I
Method = Ward
19070177-JL2-JL2 AAL83746_1:t_T_MHC_I
N-term Amino Acid of 9-mer

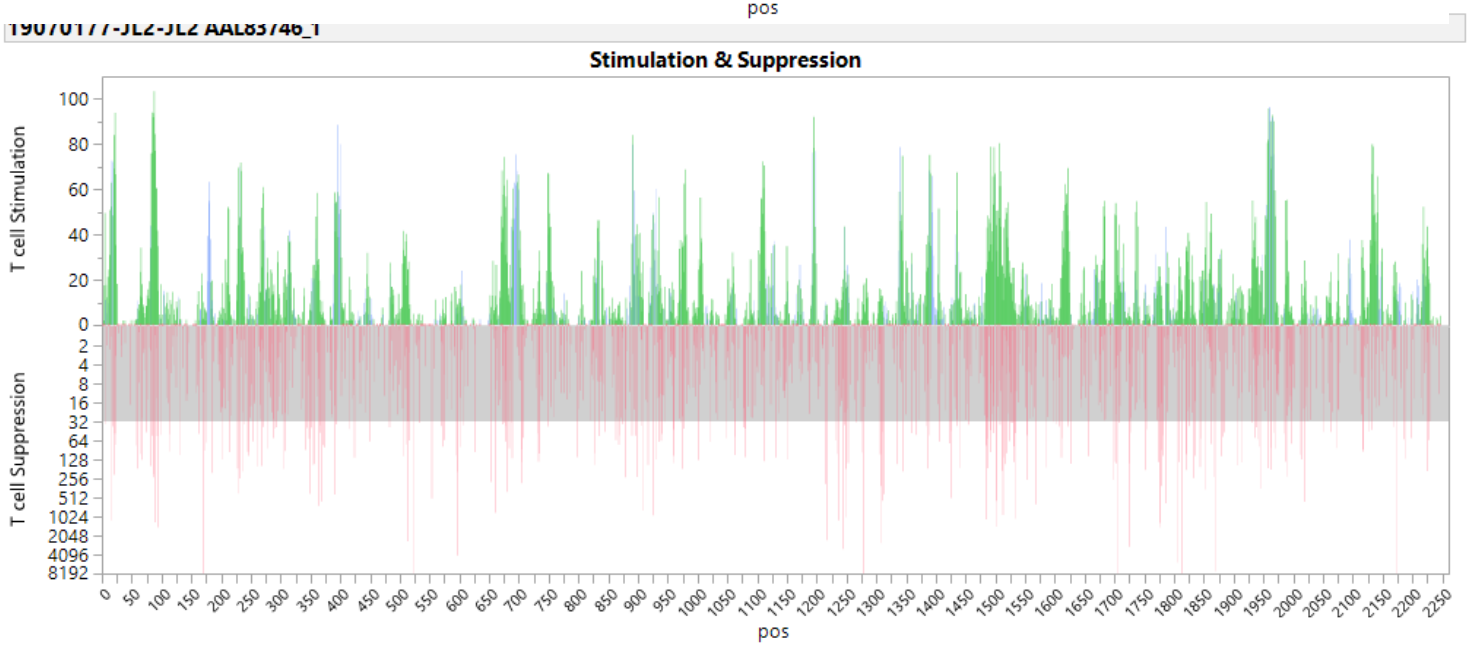

Hierarchical Cluster MHC I
Method = Ward
19070177-JL2-JL2 AAL83746_1;t MHC I & Cathepsins consolidated
0.00
0.02
0.04
0.06
0.08
0.10
0.66
1.26
1.84
2.42
3.00
N-term Amino Acid of 9-mer



19070177-JL2-JL2 AAL83746_1



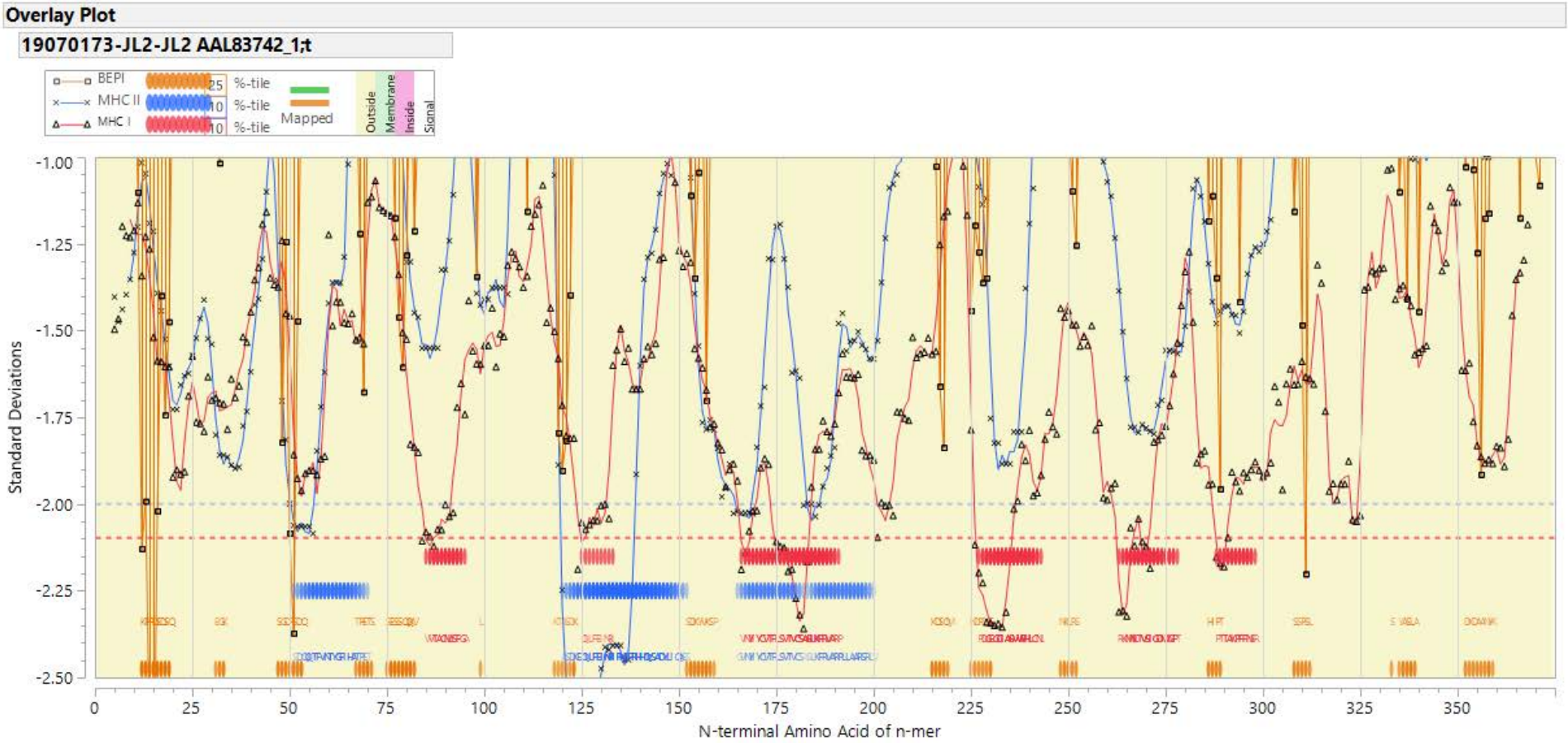


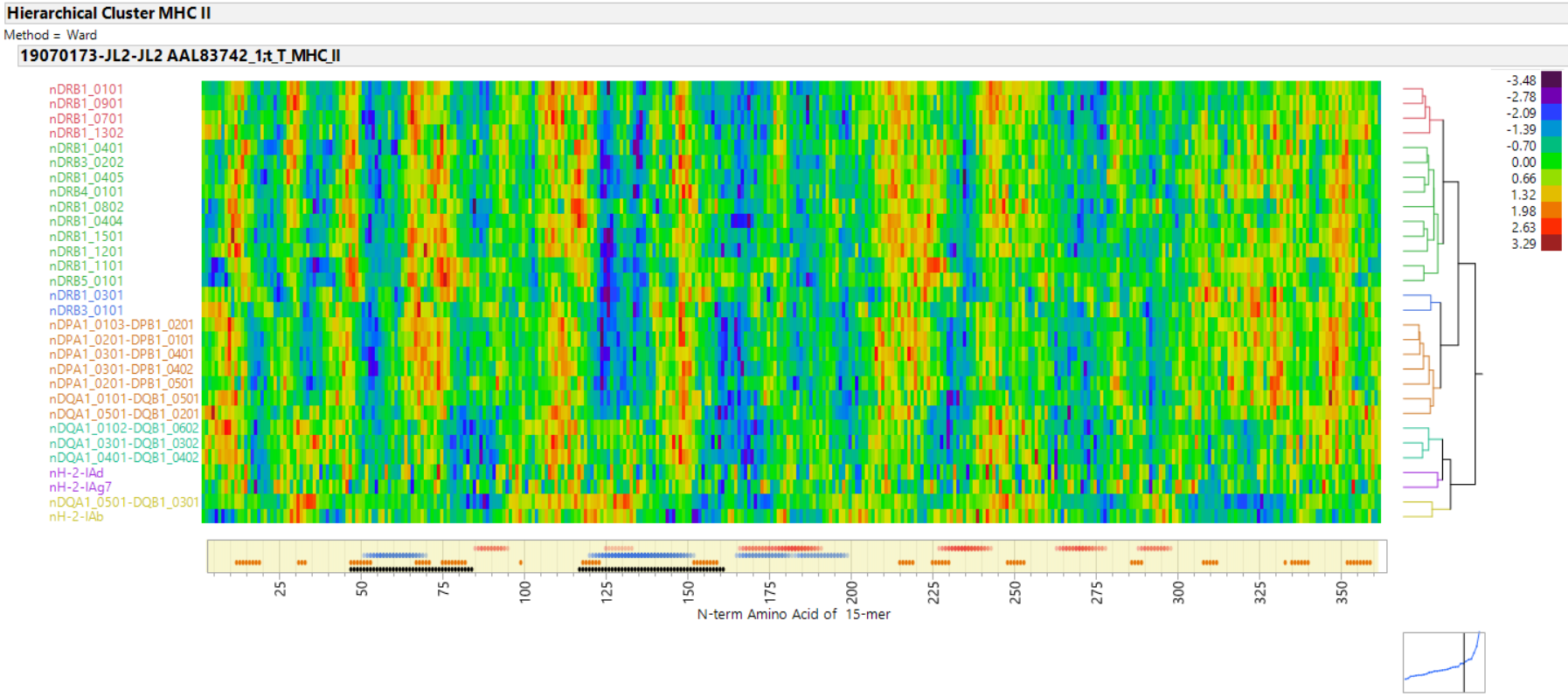


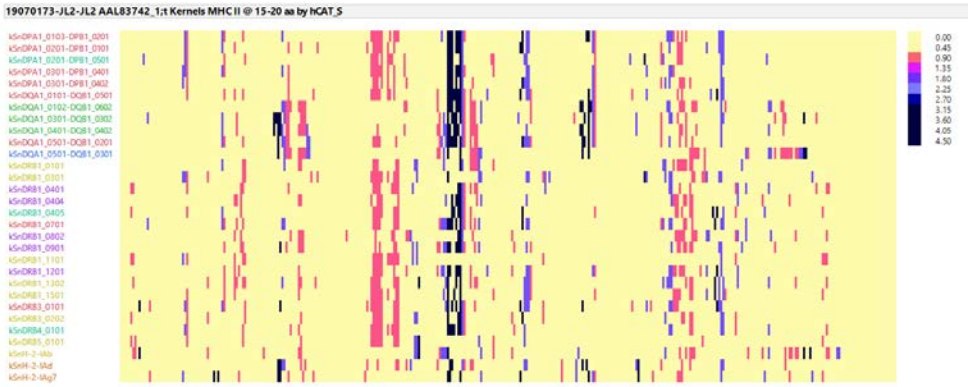
Pos 523
Supp II
204001

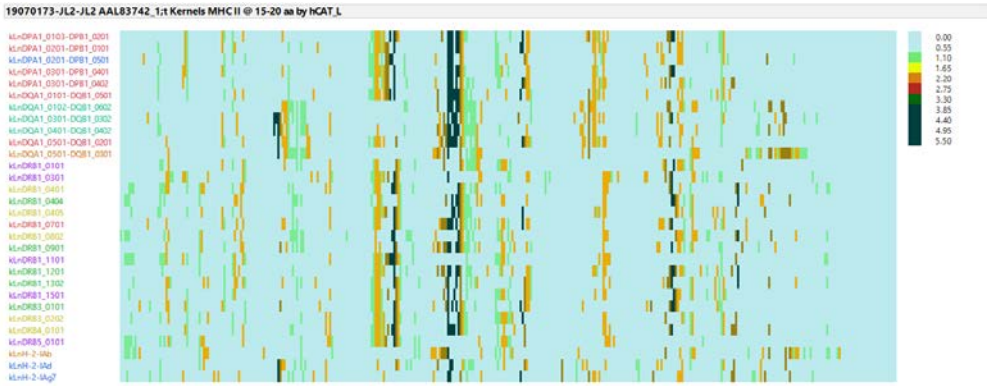
Matrix protein

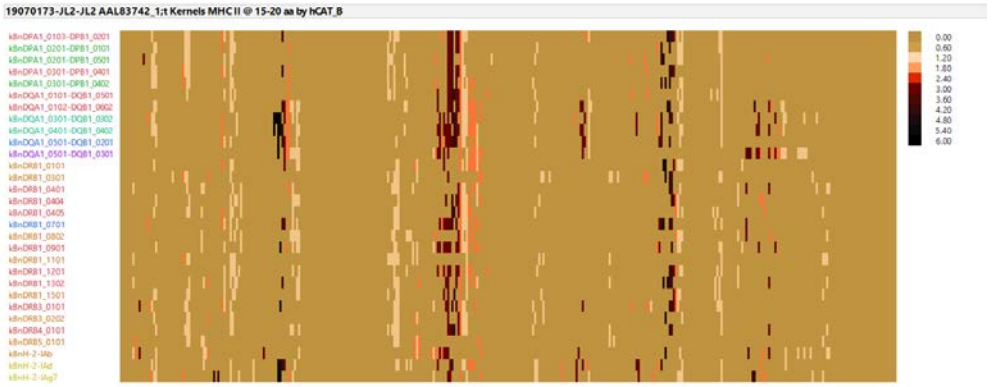
19070173-JL2-JL2 AAL83742_1| matrix protein

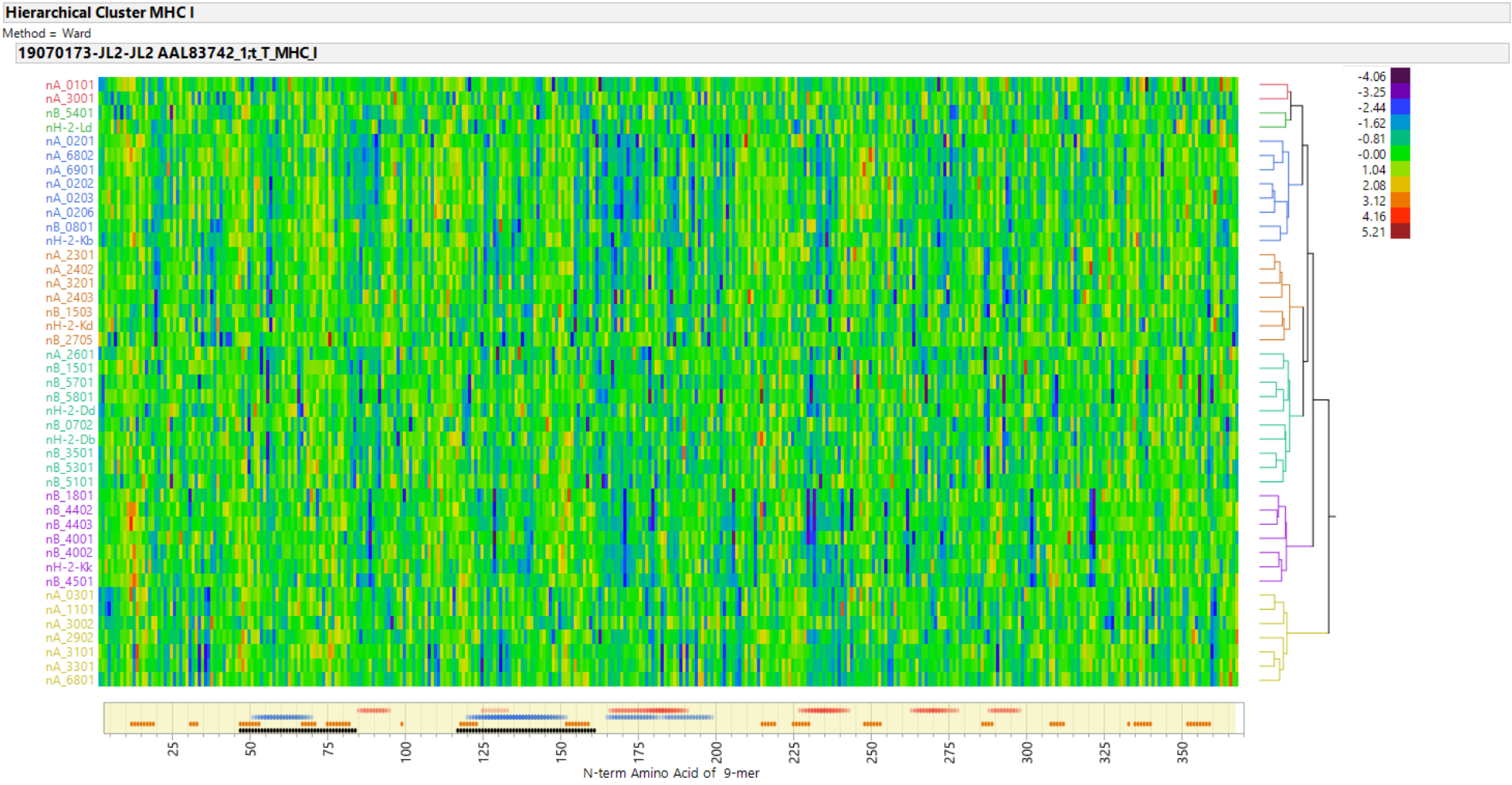

Overlay Plot
19070173-JL2-JL2 AAL83742_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-terminal Amino Acid of n-mer

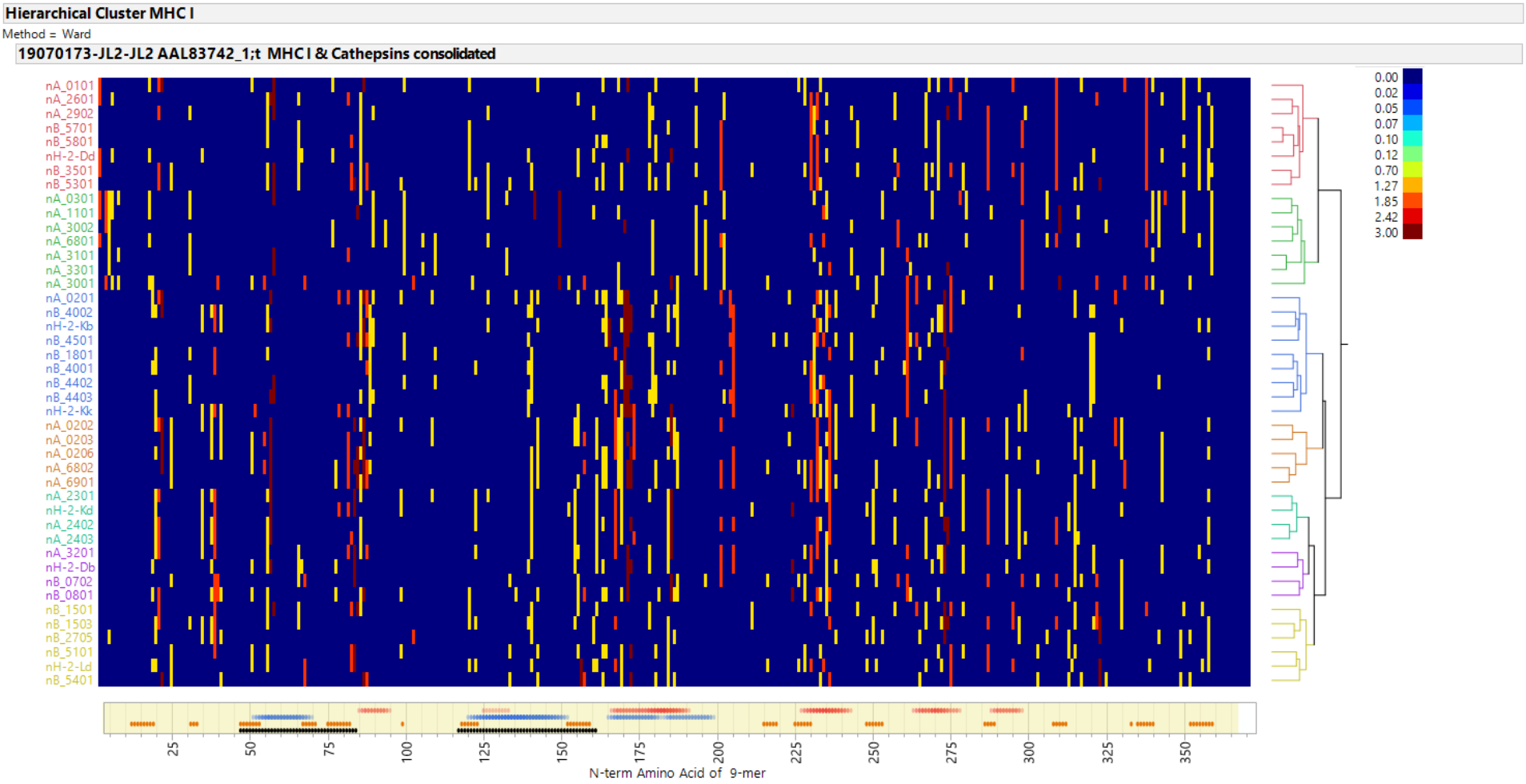

Hierarchical Cluster MHC II
Method = Ward
19070173-JL2-JL2 AAL83742_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0405
nDRB4_0101
nDRB1_0802
nDRB1_0404
nDRB1_1501
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.48
-2.78
-2.09
-1.39
-0.70
0.00
0.66
1.32
1.98
2.63
3.29
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 15-mer

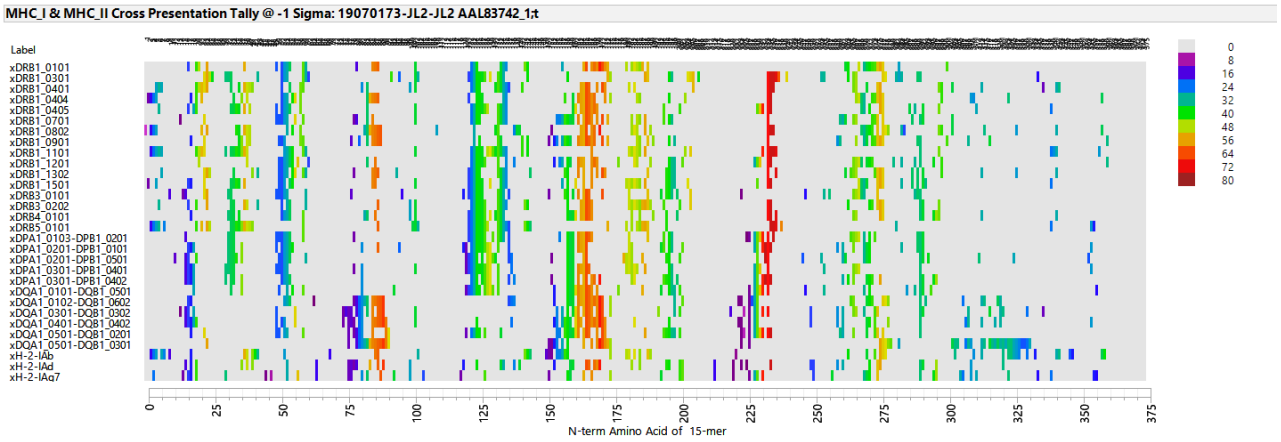

19070173-JL2-JL2 AAL83742_1;t Kernels MHC II @ 15-20 aa by hCAT_S


19070173-JL2-JL2 AAL83742_1;t Kernels MHC II @ 15-20 aa by hCAT_L

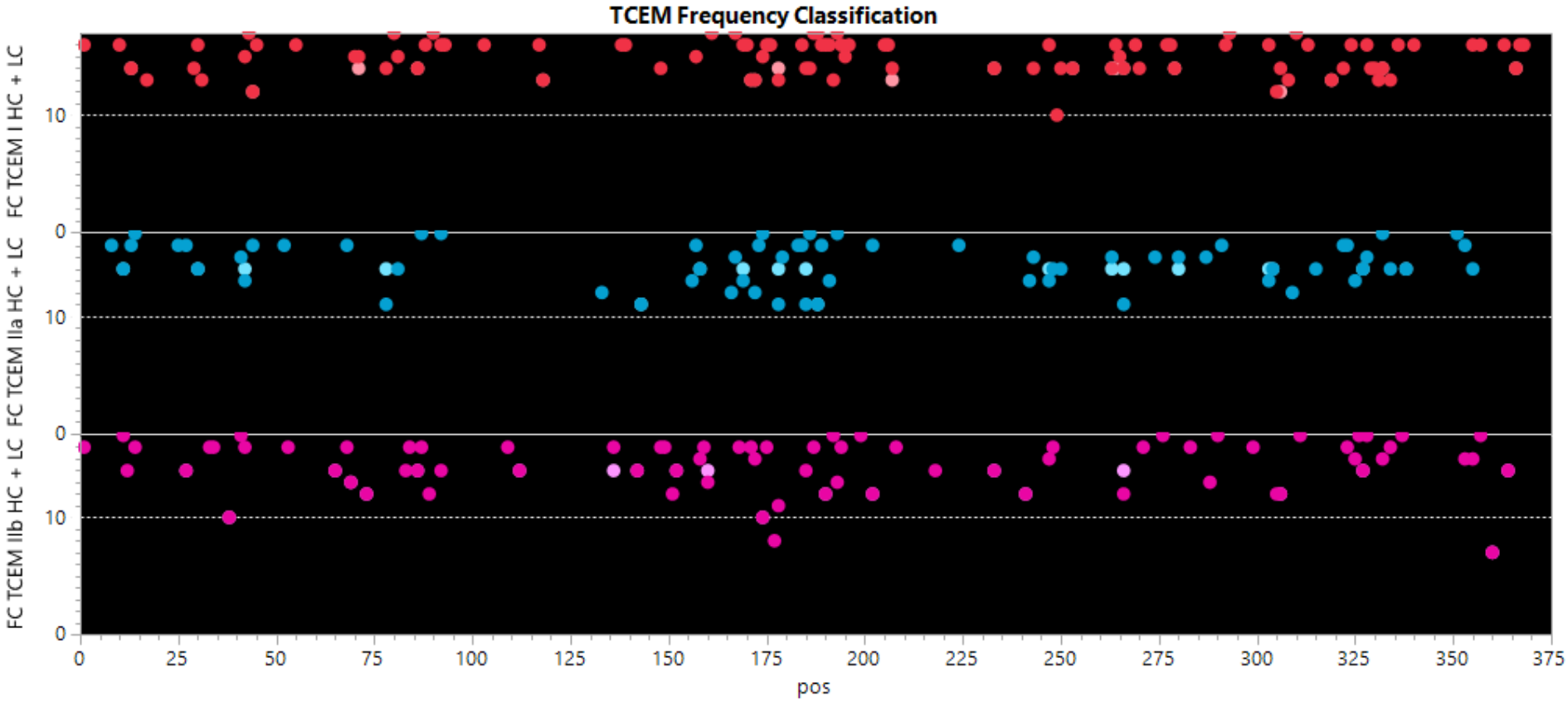

19070173-JL2-JL2 AAL83742_1;t Kernels MHC II @ 15-20 aa by hCAT_B

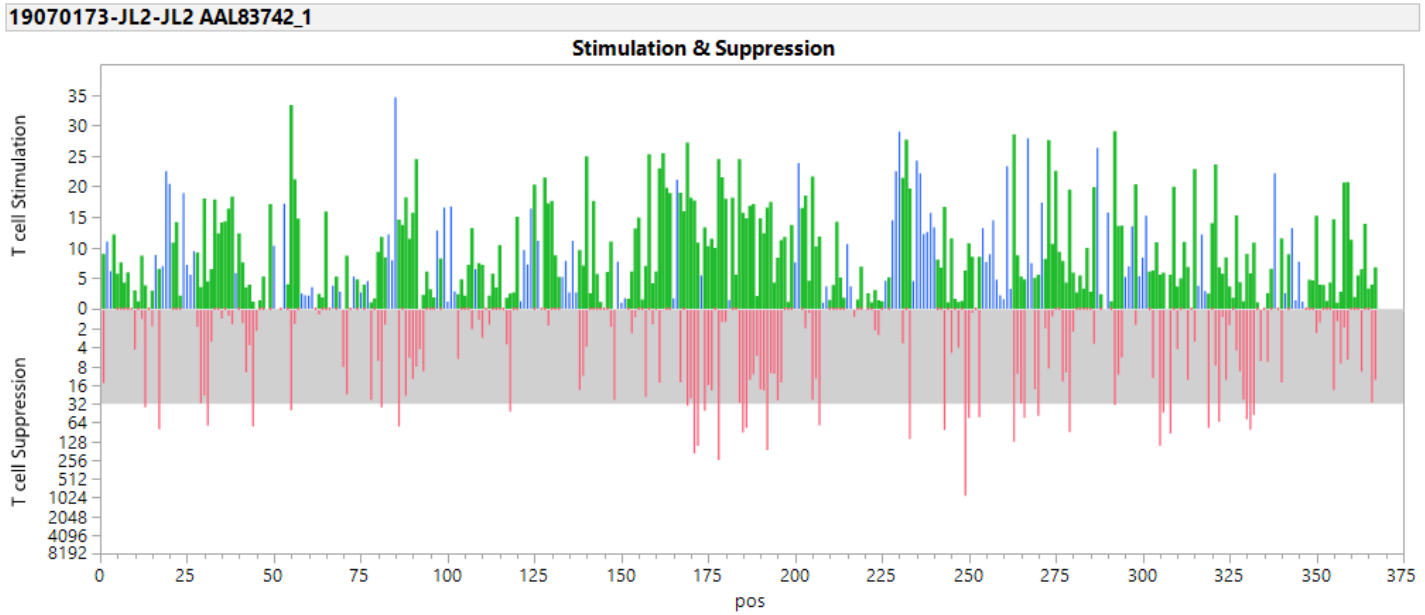

Hierarchical Cluster MHC I
Method = Ward
19070173-JL2-JL2 AAL83742_1;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nB_1503
nH-2-Kd
nB_2705
nA_2601
nB_1501
nB_5701
nB_5801
nH-2-Dd
nB_0702
nH-2-Db
nB_3501
nB_5301
nB_5101
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.06
-3.25
-2.44
-1.62
-0.81
-0.00
1.04
2.08
3.12
4.16
5.21
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 9-mer

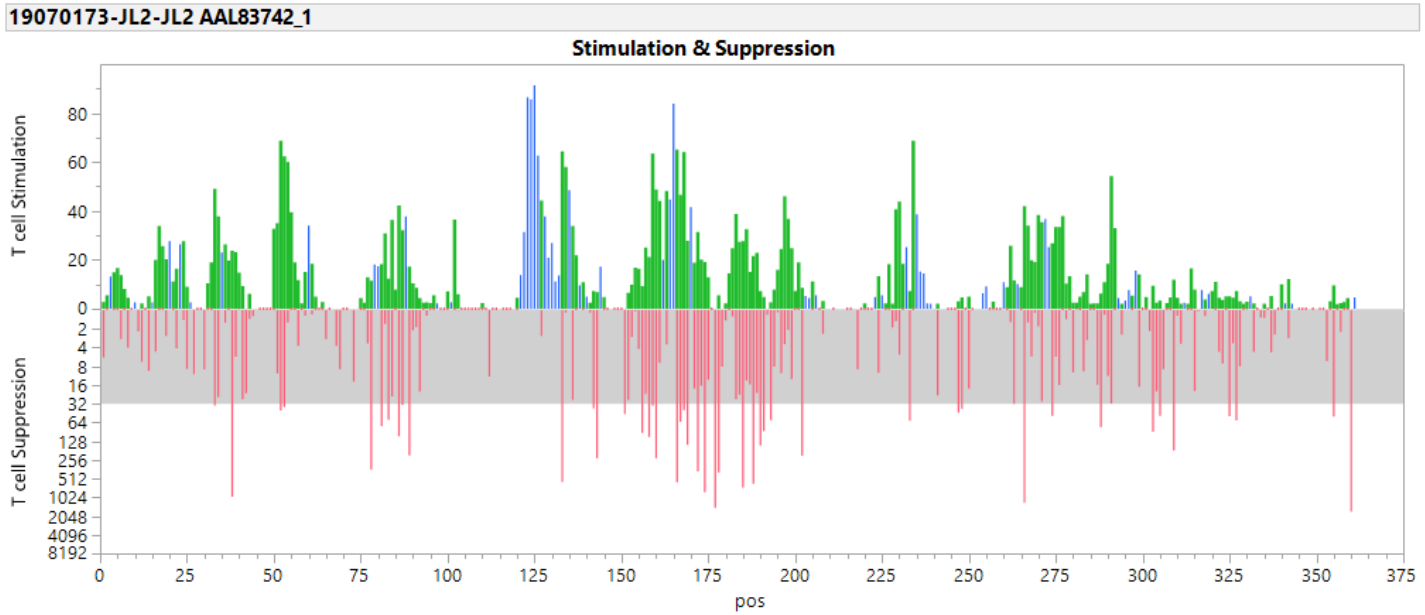

Hierarchical Cluster MHC I
Method = Ward
19070173-JL2-JL2 AAL83742_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nA_2902
nB_5701
nB_5801
nH-2-Dd
nB_3501
nB_5301
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_3001
nA_0201
nB_4002
nH-2-Kb
nB_4501
nB_1801
nB_4001
nB_4402
nB_4403
nH-2-Kk
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_2301
nH-2-Kd
nA_2402
nA_2403
nA_3201
nH-2-Db
nB_0702
nB_0801
nB_1501
nB_1503
nB_2705
nB_5101
nH-2-Ld
nB_5401
0.00
0.02
0.05
0.07
0.10
0.12
0.70
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070173-JL2-JL2 AAL83742_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

19070173-JL2-JL2 AAL83742_1

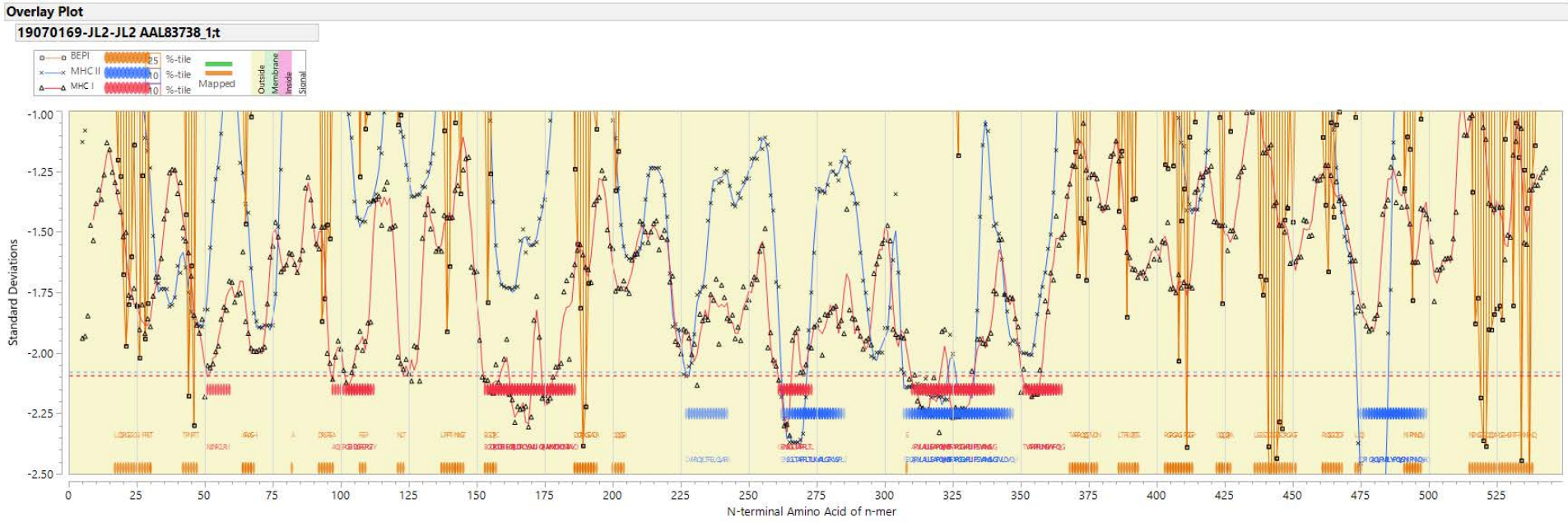


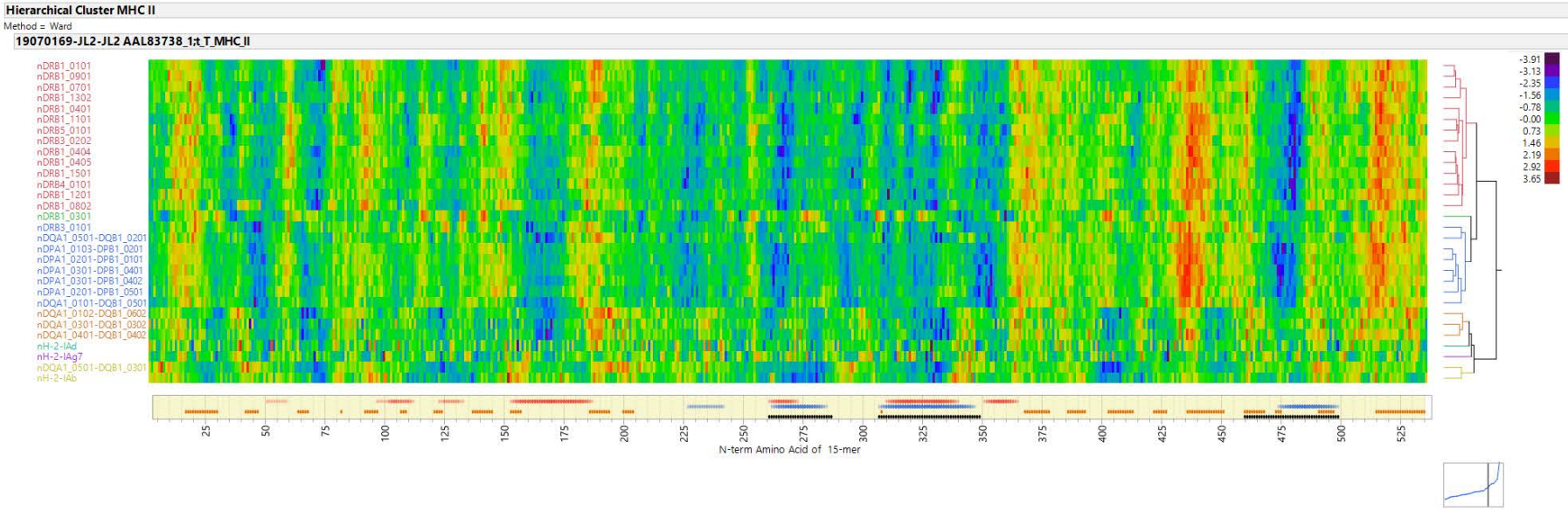

19070173-JL2-JL2 AAL83742_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

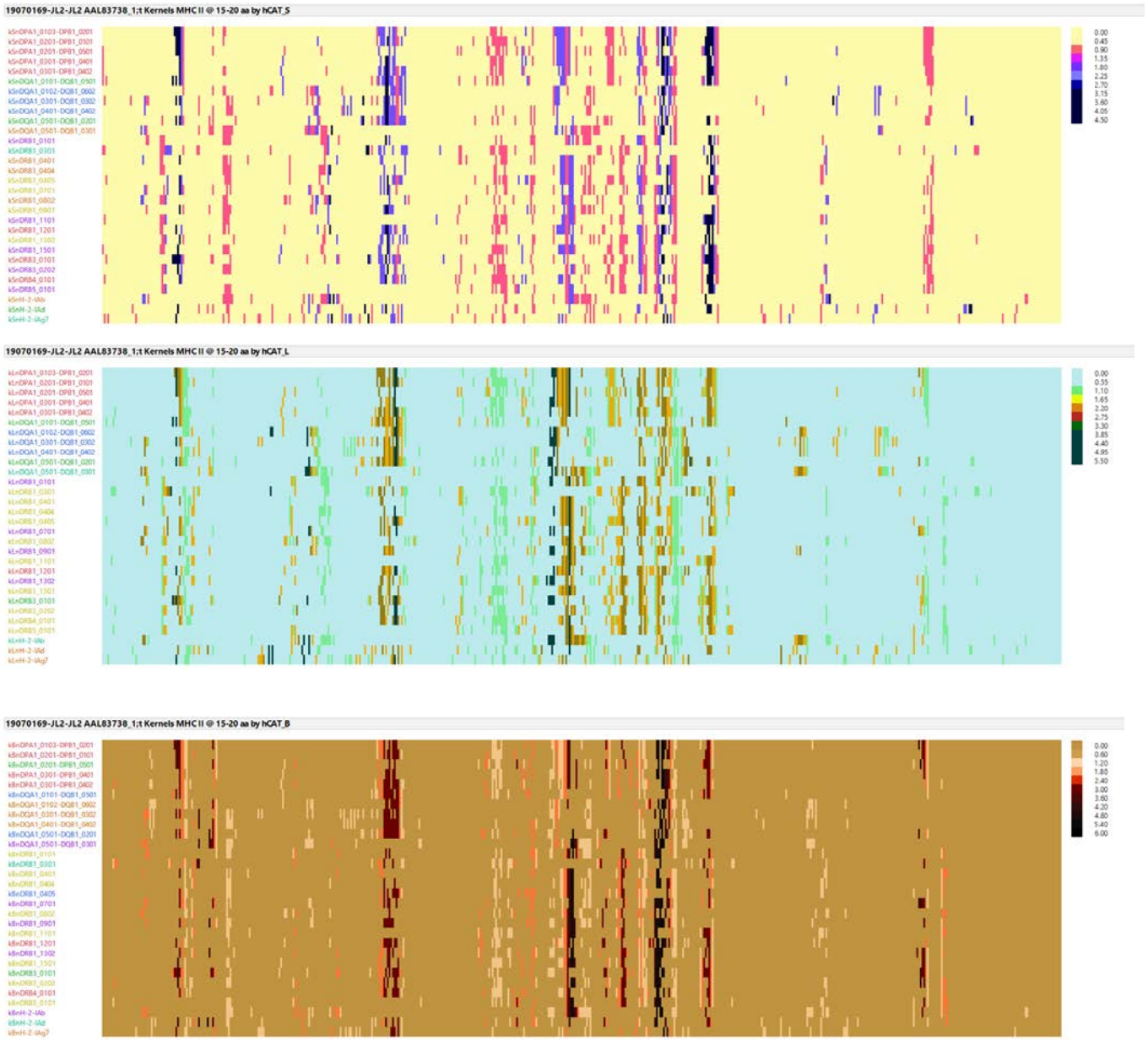

19070173-JL2-JL2 AAL83742_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

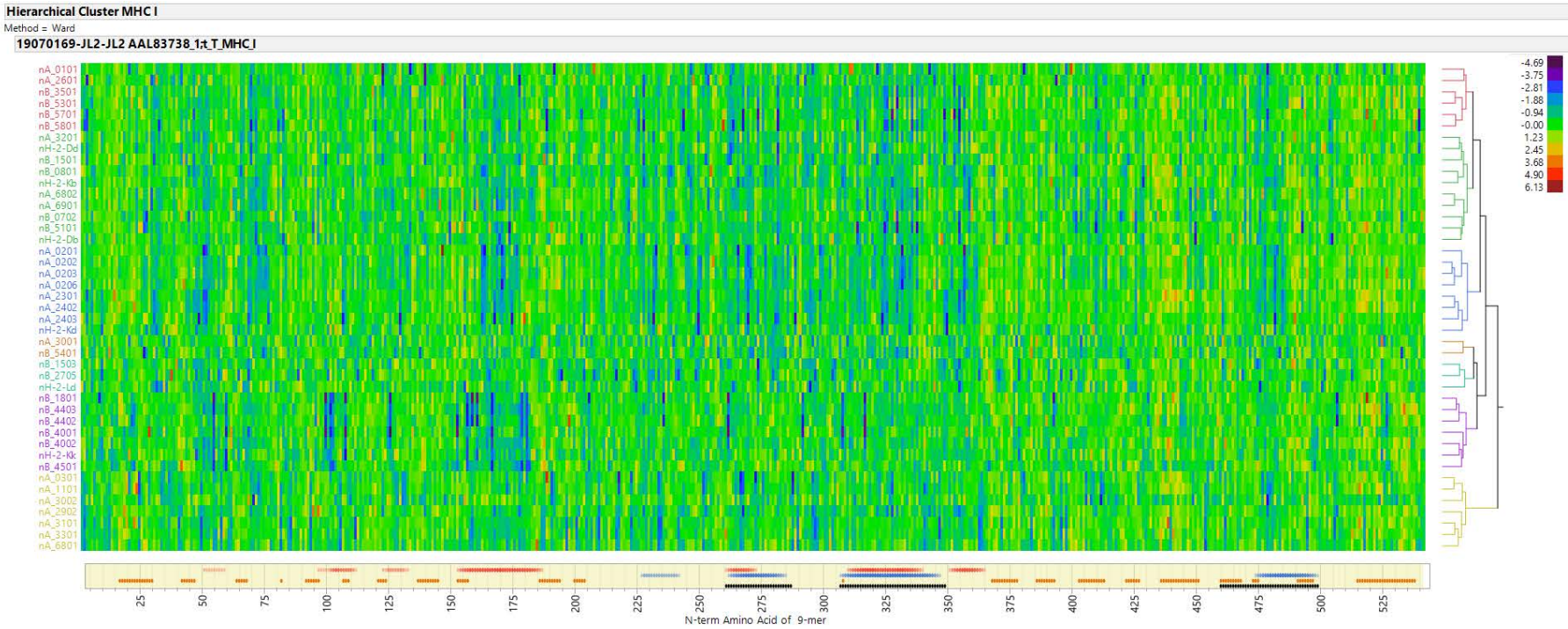
Nucelocapsid

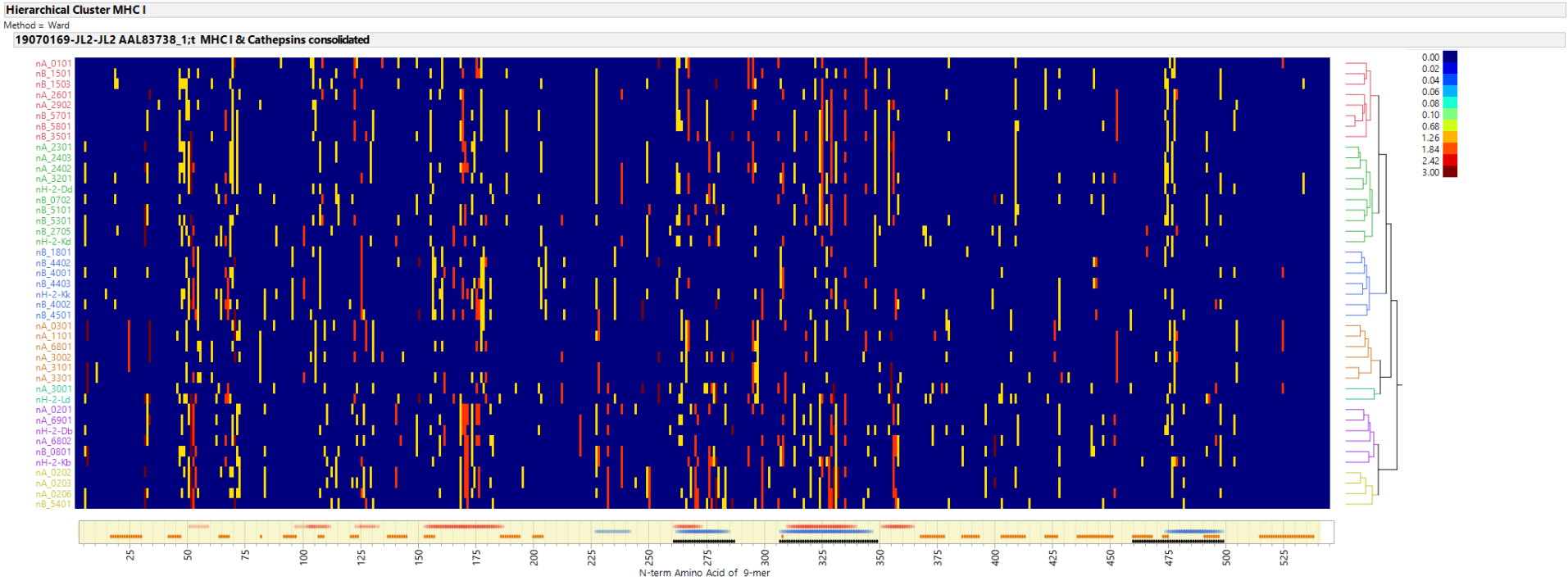
19070169-JL2-JL2 AAL83738_1| nucleocapsid protein

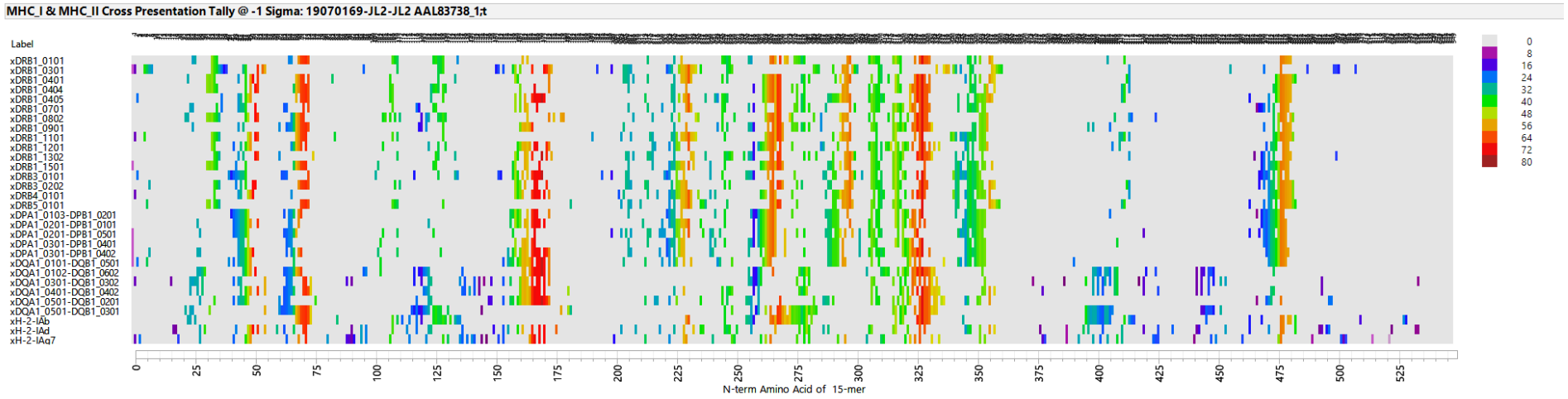

Overlay Plot
19070169-JL2-JL2 AAL83738_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
19070169-JL2-JL2 AAL83738_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB1_1101
nDRB5_0101
nDRB3_0202
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.91
-3.13
-2.35
-1.56
-0.78
-0.00
0.73
1.46
2.19
2.92
3.65
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

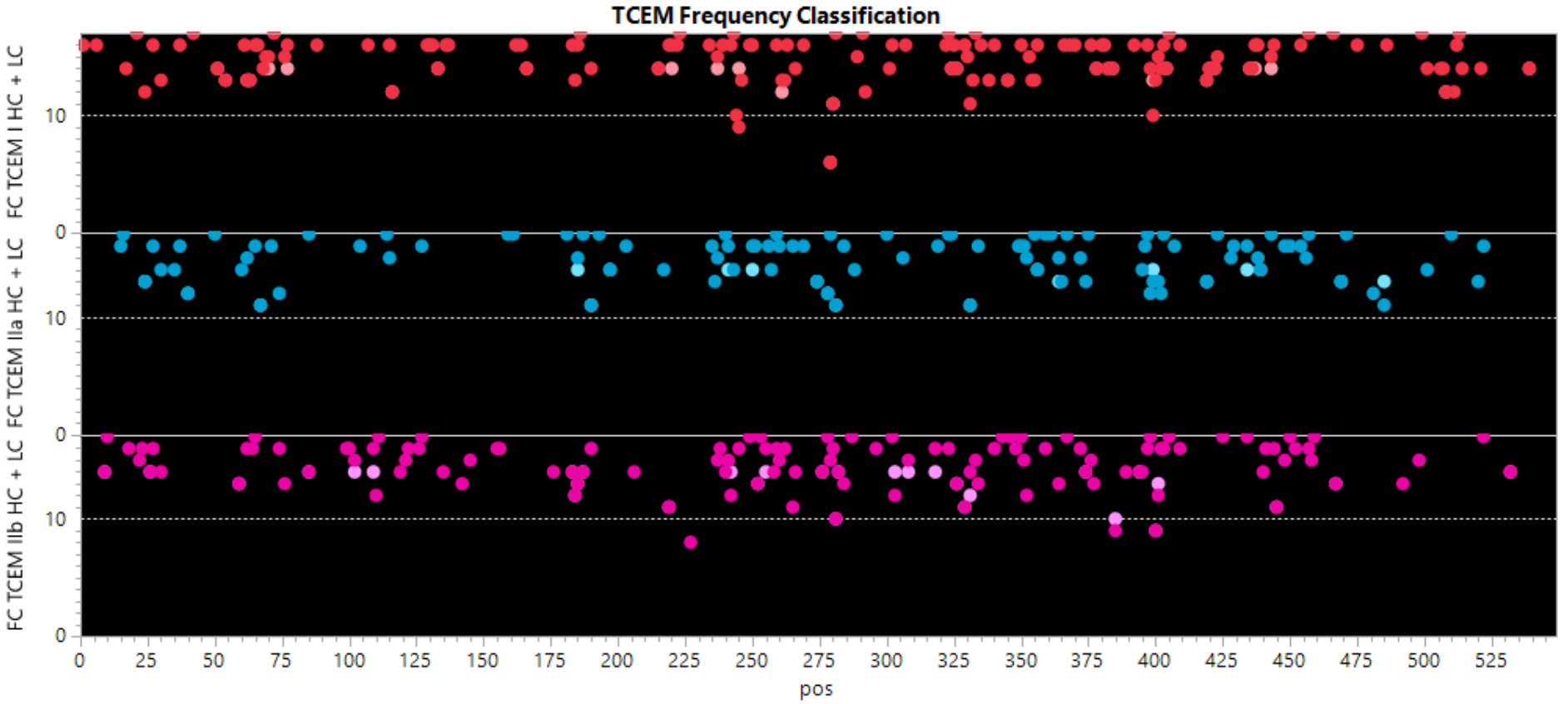

19070169-JL2-JL2 AAL83738_1;t Kernels MHC II @ 15-20 aa by hCAT_S
19070169-JL2-JL2 AAL83738_1;t Kernels MHC II @ 15-20 aa by hCAT_L
19070169-JL2-JL2 AAL83738_1;t Kernels MHC II @ 15-20 aa by hCAT_B

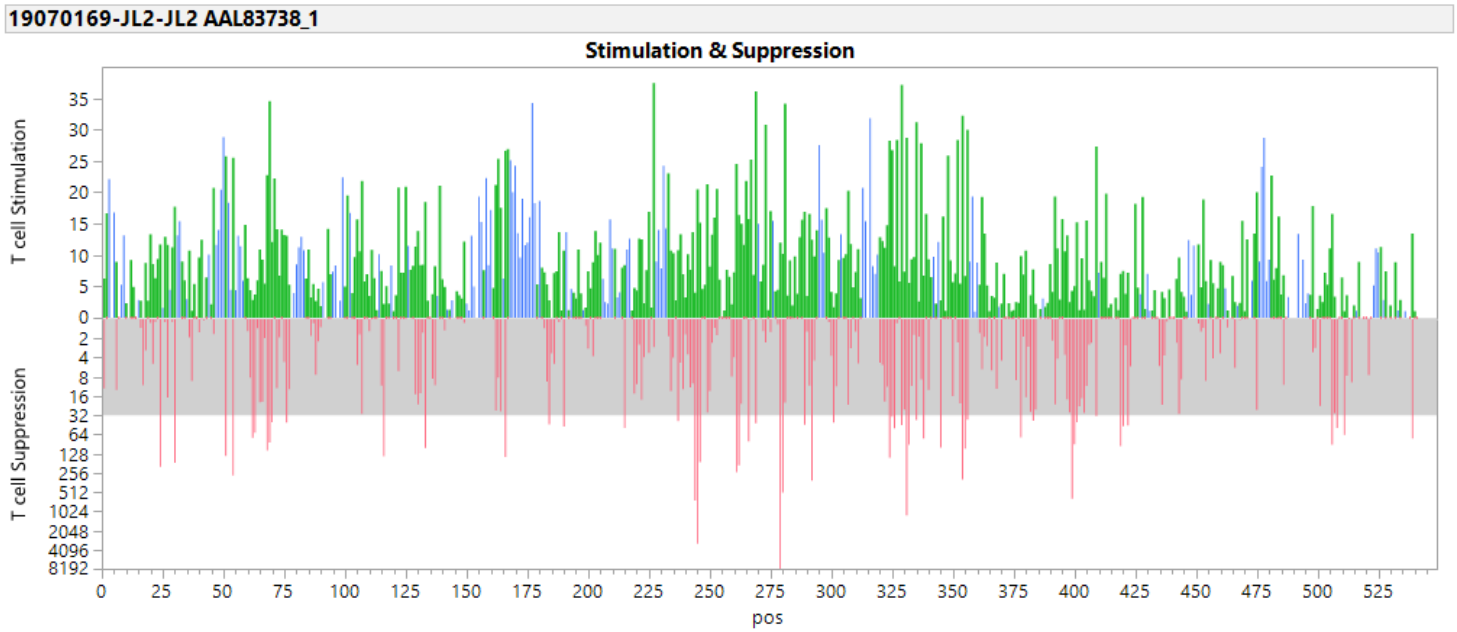

Hierarchical Cluster MHC I
Method = Ward
19070169-JL2-JL2 AAL83738_1;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nA_3201
nH-2-Dd
nB_1501
nB_0801
nH-2-Kb
nA_6802
nA_6901
nB_0702
nB_5101
nH-2-Db
nA_0201
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_3001
nB_5401
nB_1503
nB_2705
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.69
-3.75
-2.81
-1.88
-0.94
-0.00
1.23
2.45
3.68
4.90
6.13
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer

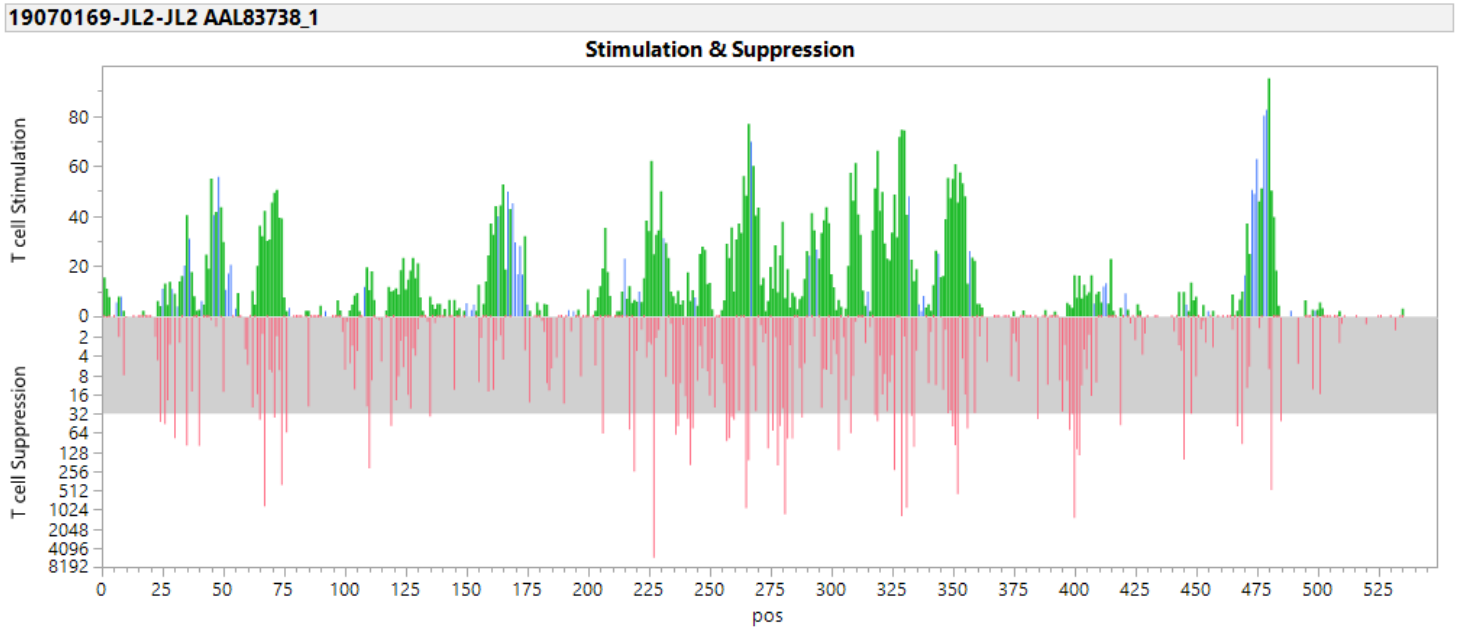

Hierarchical Cluster MHC I
Method = Ward
19070169-JL2-JL2 AAL83738_1;t MHC I & Cathepsins consolidated
nA_0101
nB_1501
nB_1503
nA_2601
nA_2902
nB_5701
nB_5801
nB_3501
nA_2301
nA_2403
nA_2402
nA_3201
nH-2-Dd
nB_0702
nB_5101
nB_5301
nB_2705
nH-2-Kd
nB_1801
nB_4402
nB_4001
nB_4403
nH-2-Kk
nB_4002
nB_4501
nA_0301
nA_1101
nA_6801
nA_3002
nA_3101
nA_3301
nA_3001
nH-2-Ld
nA_0201
nA_6901
nH-2-Db
nA_6802
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nB_5401
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070169-JL2-JL2 AAL83738_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

19070169-JL2-JL2 AAL83738_1




19070169-JL2-JL2 AAL83738_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

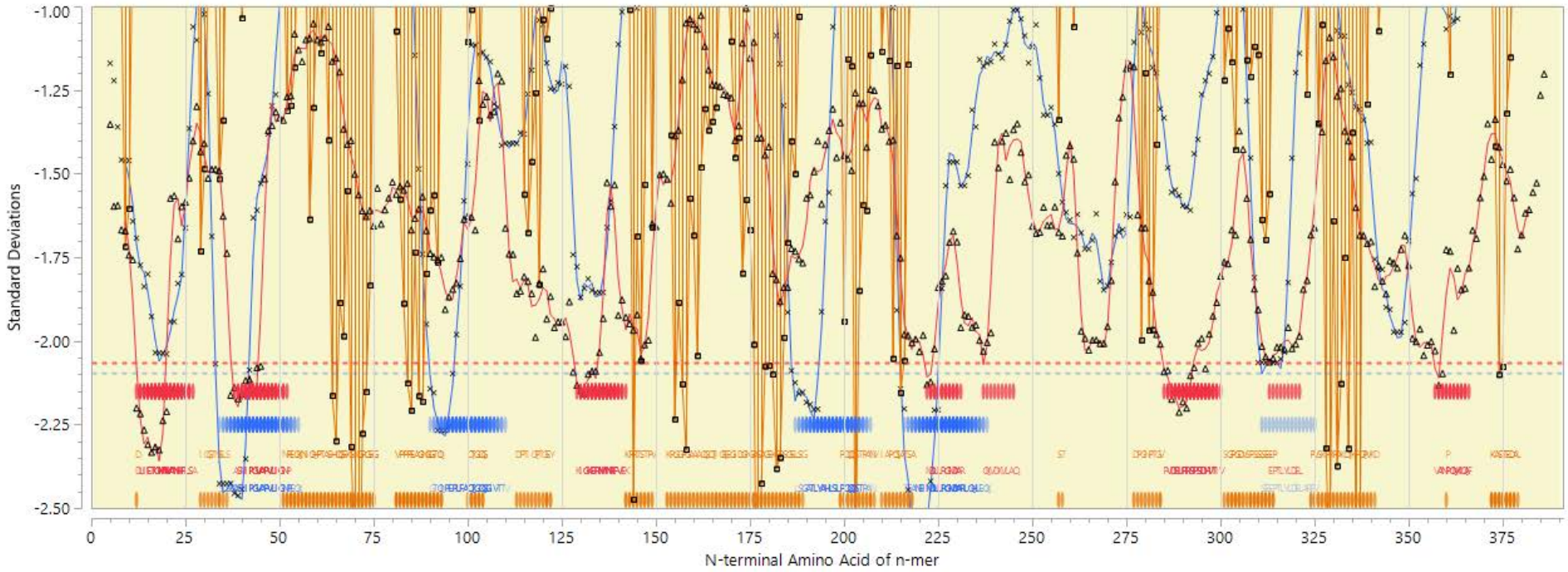

19070169-JL2-JL2 AAL83738_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

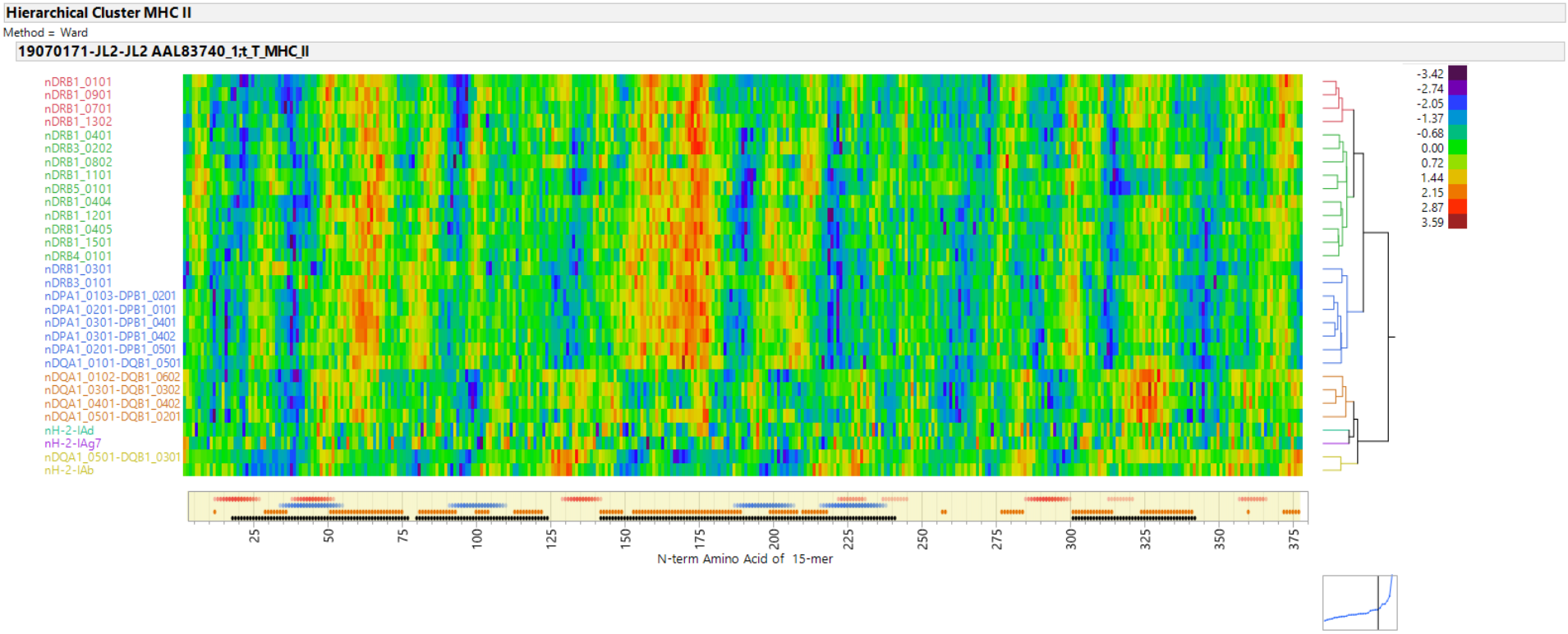
Phosphoprotein

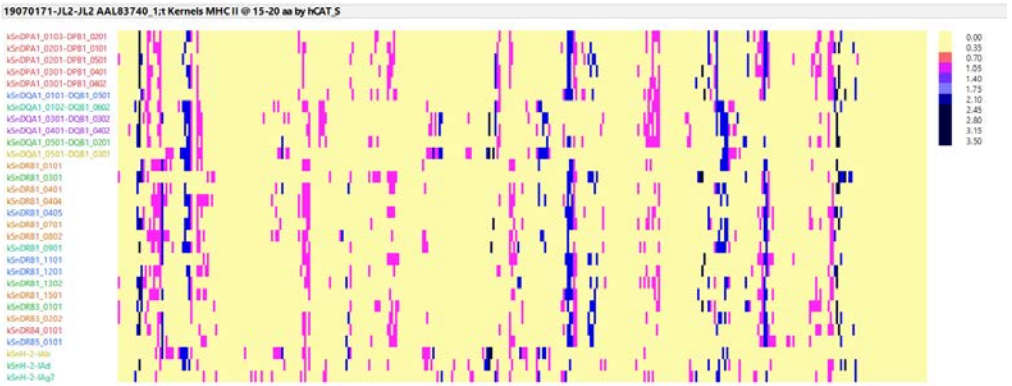
19070171-JL2-JL2 AAL83740_1| phosphoprotein

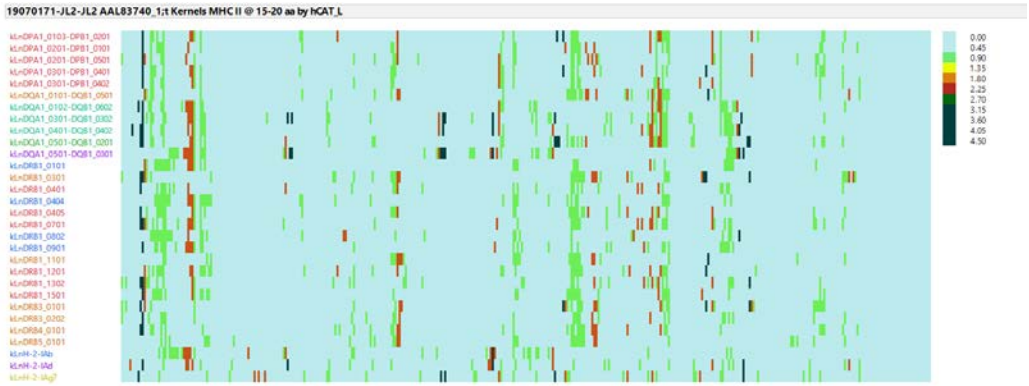
Overlay Plot

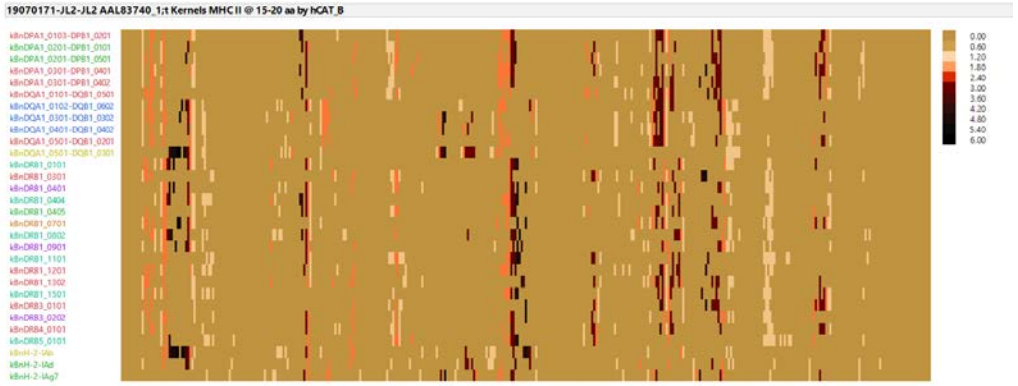
19070171-JL2-JL2 AAL83740_1;t

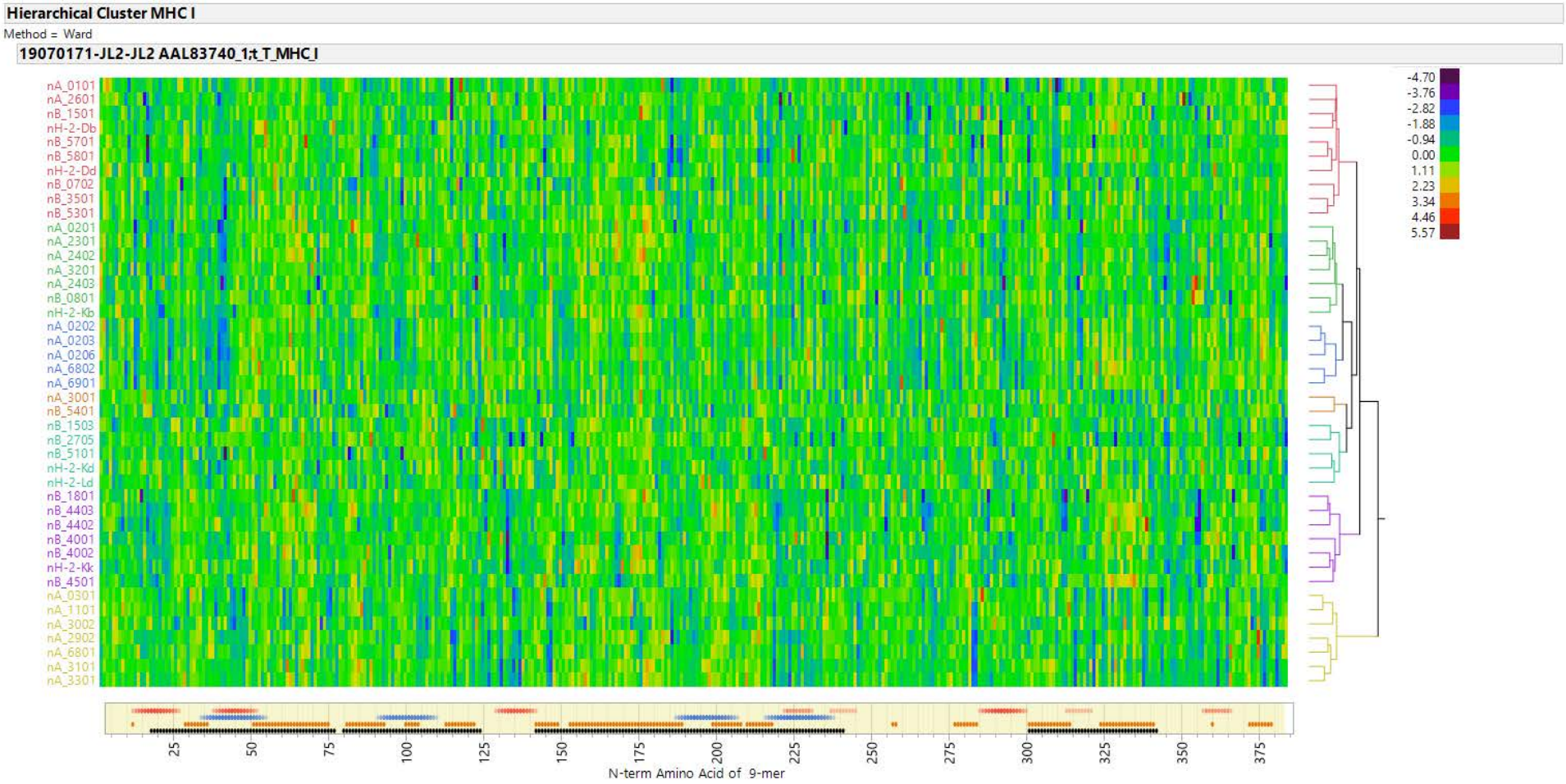


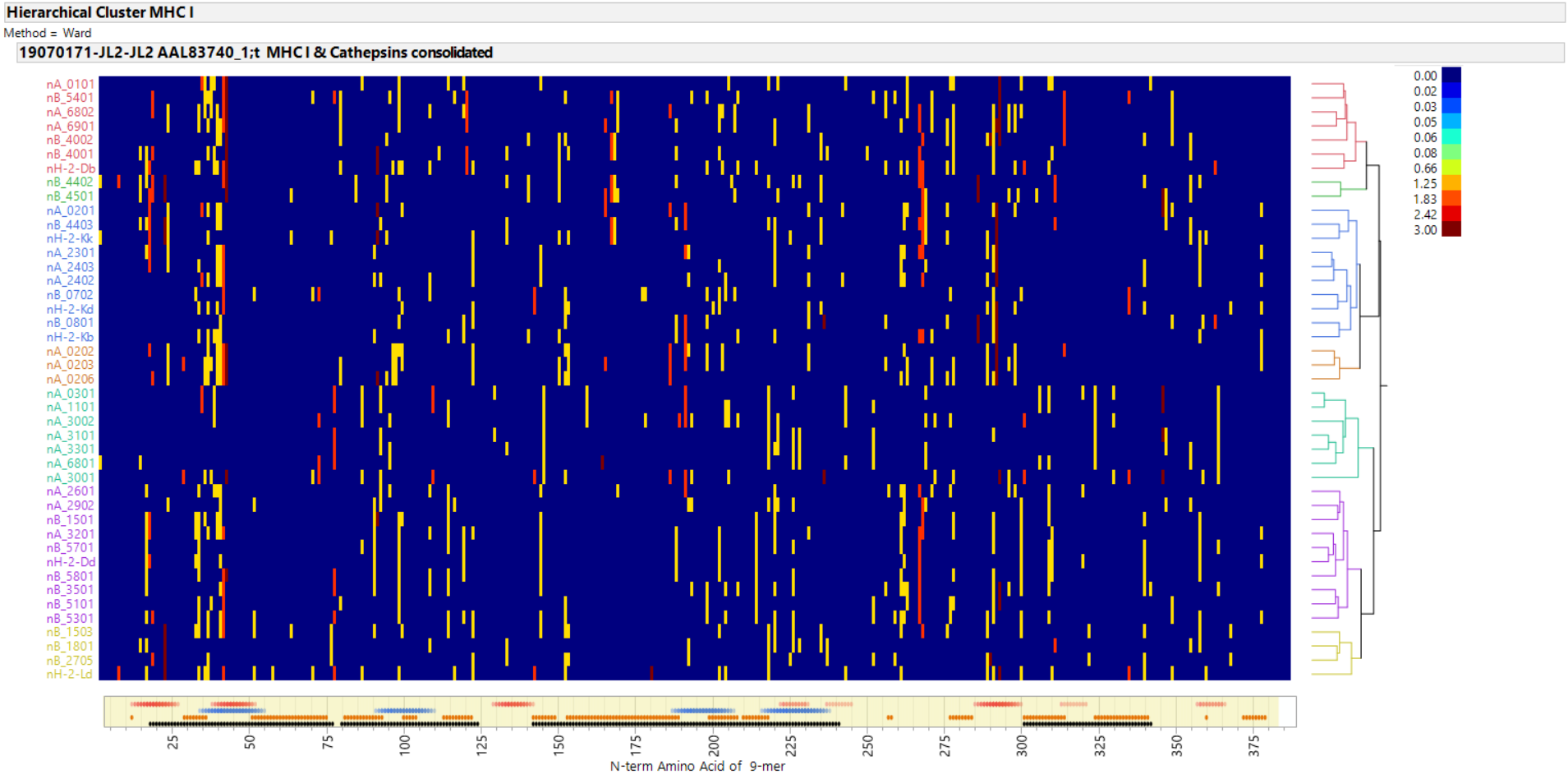

Hierarchical Cluster MHC II
Method = Ward
19070171-JL2-JL2 AAL83740_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_1201
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.42
-2.74
-2.05
-1.37
-0.68
0.00
0.72
1.44
2.15
2.87
3.59
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 15-mer

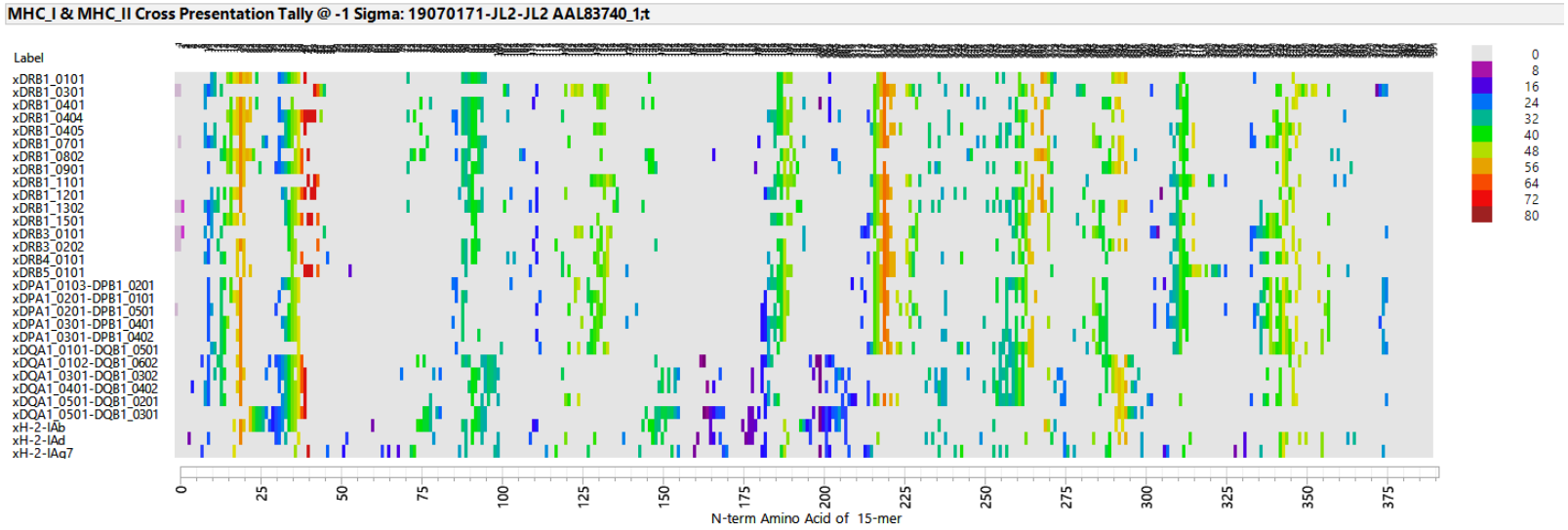

19070171-JL2-JL2 AAL83740_1;t Kernels MHC II @ 15-20 aa by hCAT_S


19070171-JL2-JL2 AAL83740_1;t Kernels MHC II @ 15-20 aa by hCAT_L

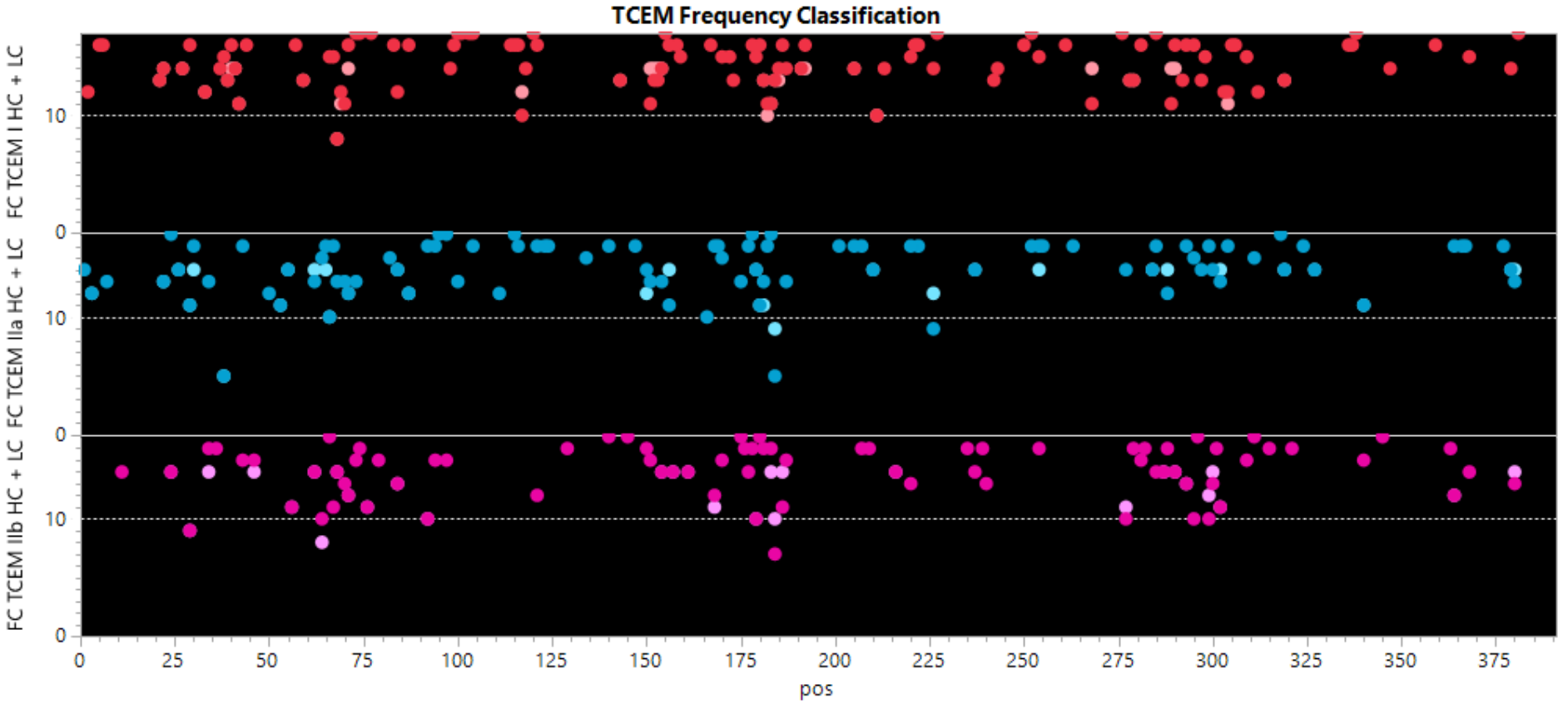

19070171-JL2-JL2 AAL83740_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
19070171-JL2-JL2 AAL83740_1;t_T_MHC_I
nA_0101
nA_2601
nB_1501
nH-2-Db
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_3501
nB_5301
nA_0201
nA_2301
nA_2402
nA_3201
nA_2403
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_3001
nB_5401
nB_1503
nB_2705
nB_5101
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_6801
nA_3101
nA_3301
-4.70
-3.76
-2.82
-1.88
-0.94
0.00
1.11
2.23
3.34
4.46
5.57
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 9-mer

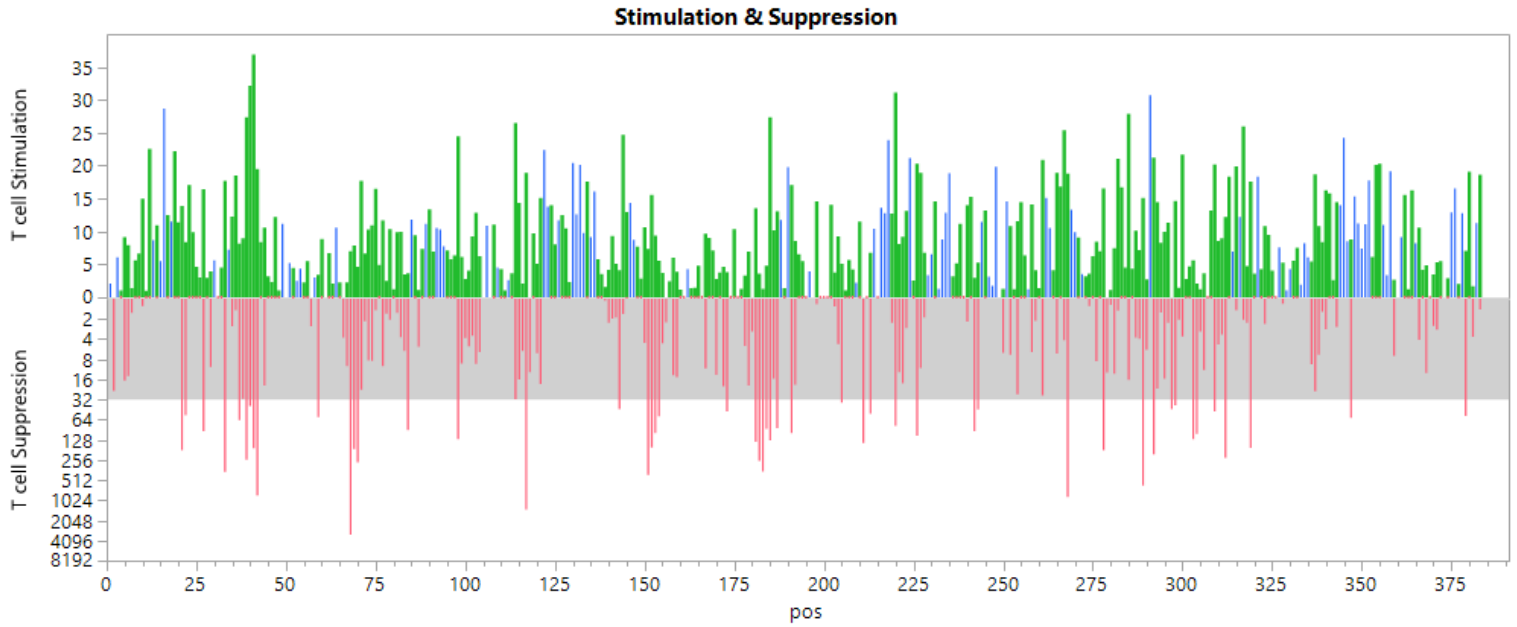

Hierarchical Cluster MHC I
Method = Ward
19070171-JL2-JL2 AAL83740_1;t MHC I & Cathepsins consolidated
nA_0101
nB_5401
nA_6802
nA_6901
nB_4002
nB_4001
nH-2-Db
nB_4402
nB_4501
nA_0201
nB_4403
nH-2-Kk
nA_2301
nA_2403
nA_2402
nB_0702
nH-2-Kd
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nA_2601
nA_2902
nB_1501
nA_3201
nB_5701
nH-2-Dd
nB_5801
nB_3501
nB_5101
nB_5301
nB_1503
nB_1801
nB_2705
nH-2-Ld
0.00
0.02
0.03
0.05
0.06
0.08
0.66
1.25
1.83
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070171-JL2-JL2 AAL83740_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

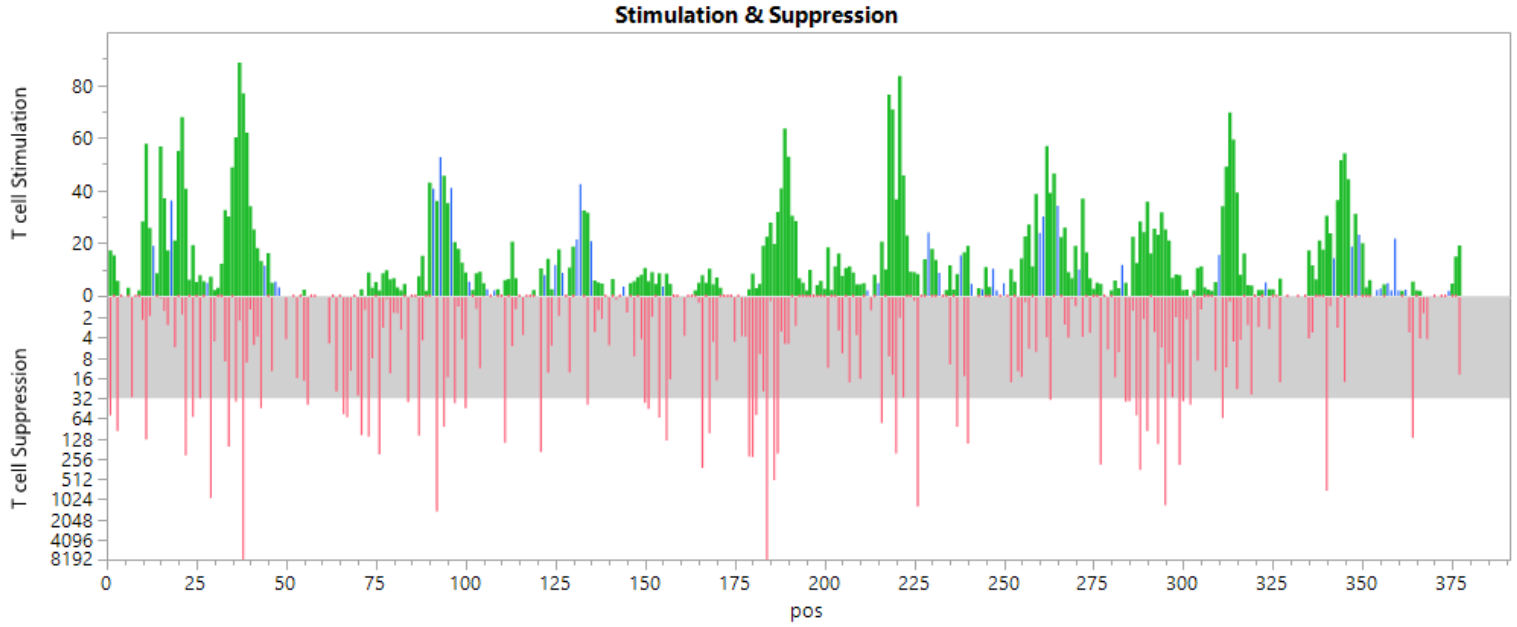
19070171-JL2-JL2 AAL83740_1



19070171-JL2-JL2 AAL83740_1



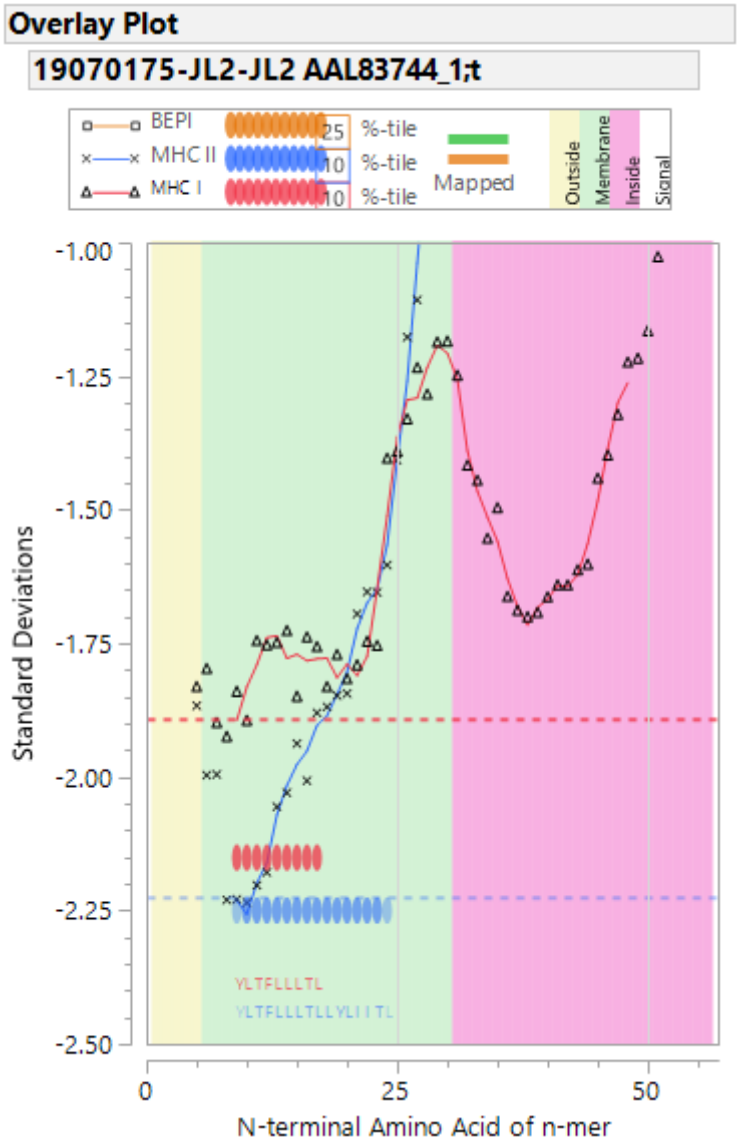
19070171-JL2-JL2 AAL83740_1

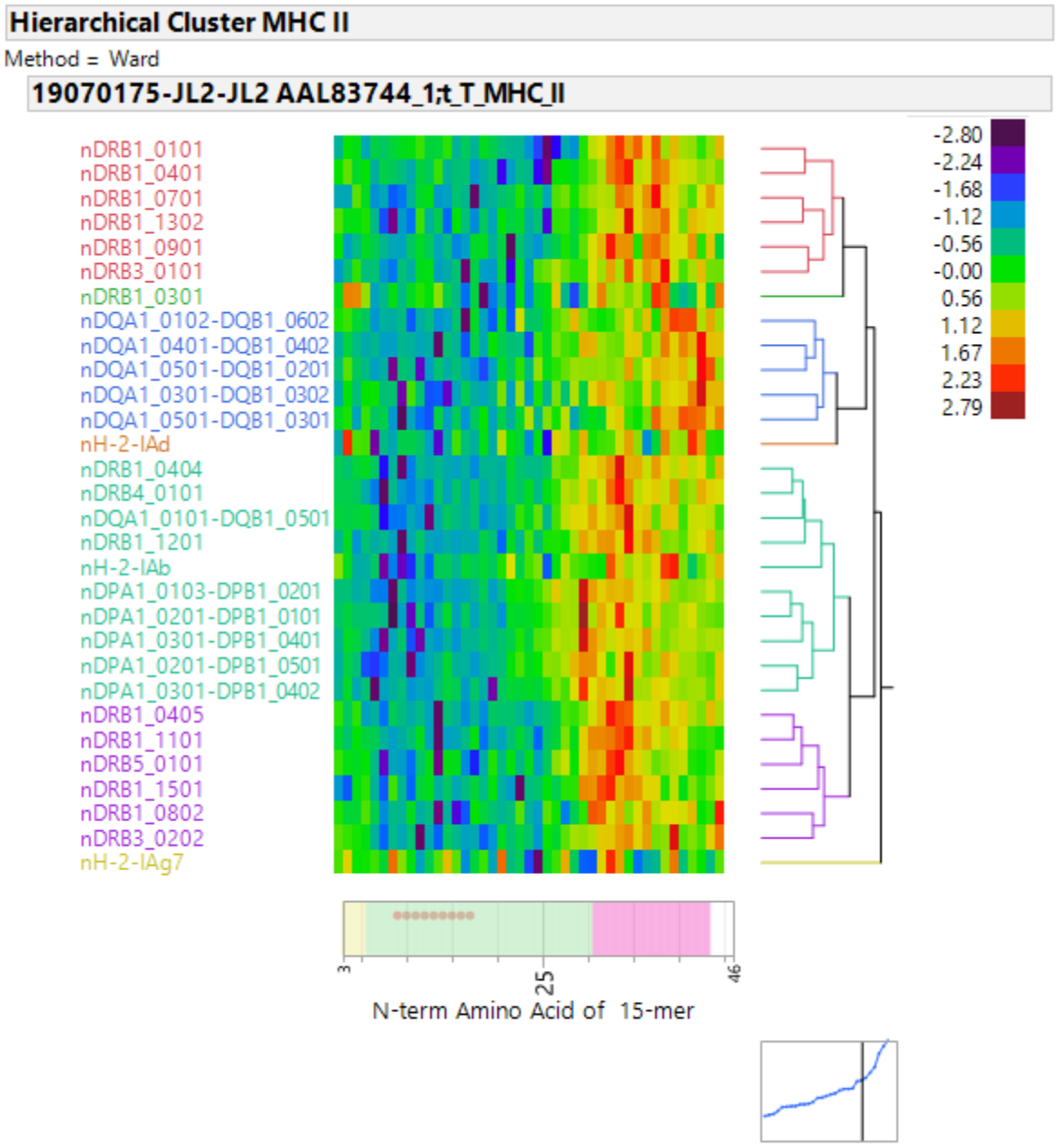


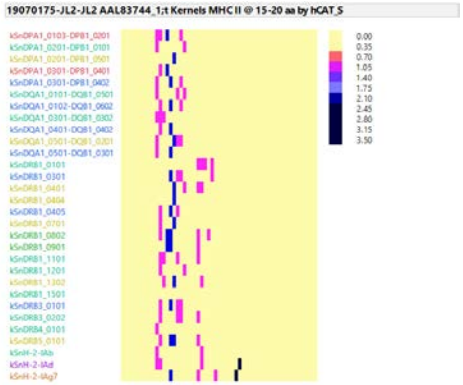
Small hydrophobic protein

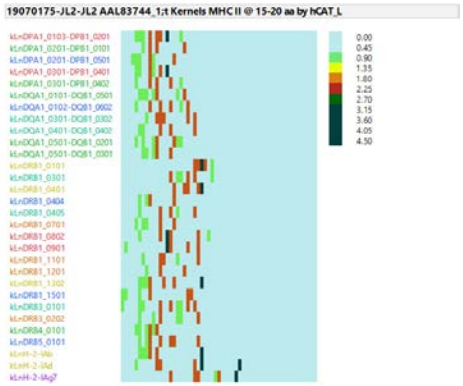
19070175-JL2-JL2 AAL83744_1| small hydrophobic protein

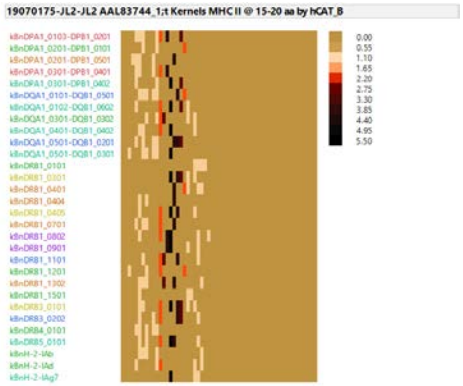
Note accuracy at C term is reduced in very short proteins – could be re run with “padding” of dummy amino acids at each end

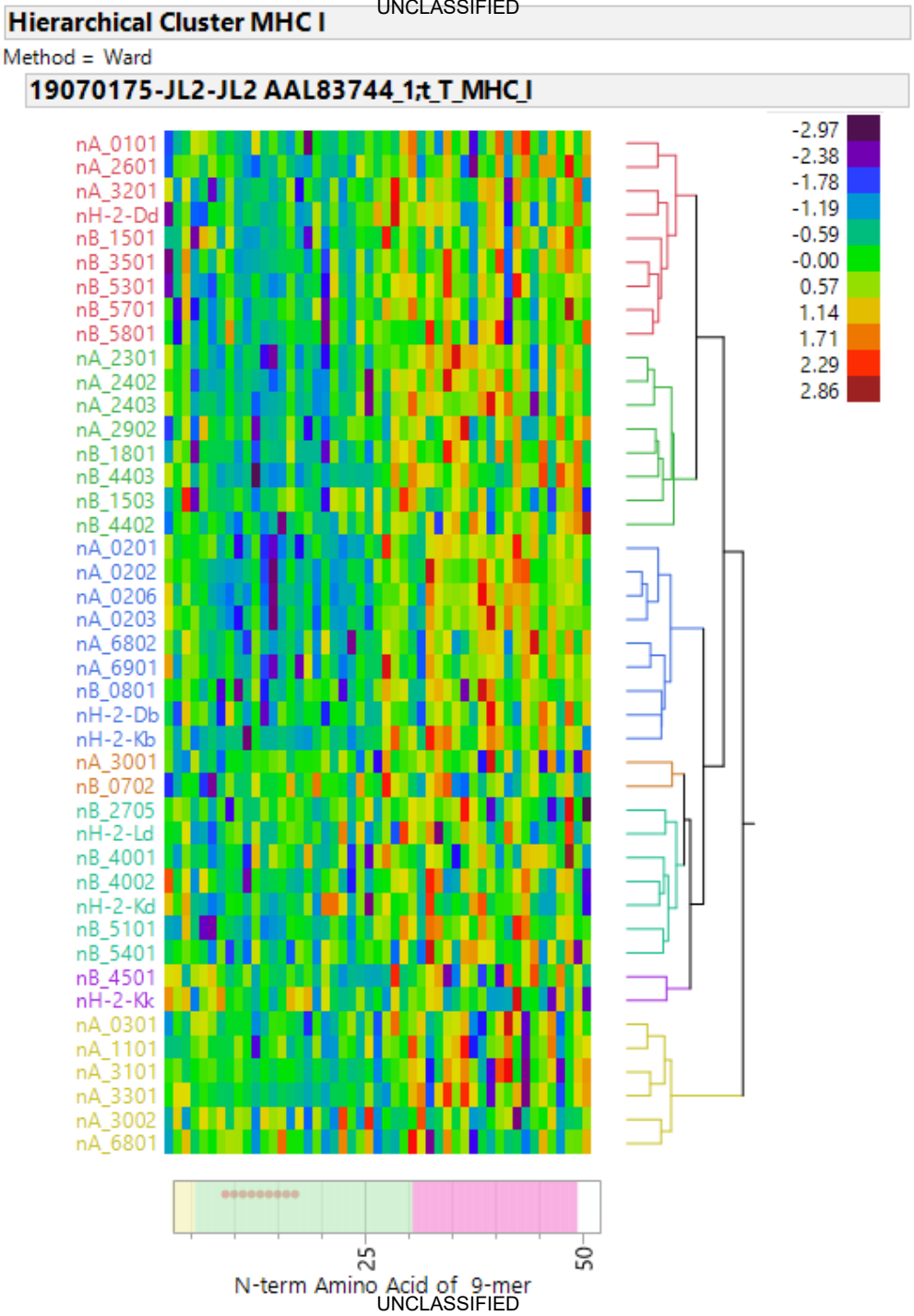

Overlay Plot
19070175-JL2-JL2 AAL83744_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
Standard Deviations
YLTFLLLTL
YLTFLLLTLLYLIITL
0
25
50
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
19070175-JL2-JL2 AAL83744_1;t_T_MHC_II
nDRB1_0101
nDRB1_0401
nDRB1_0701
nDRB1_1302
nDRB1_0901
nDRB3_0101
nDRB1_0301
nDQA1_0102-DQB1_0602
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nDQA1_0301-DQB1_0302
nDQA1_0501-DQB1_0301
nH-2-IAd
nDRB1_0404
nDRB4_0101
nDQA1_0101-DQB1_0501
nDRB1_1201
nH-2-IAb
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDRB1_0405
nDRB1_1101
nDRB5_0101
nDRB1_1501
nDRB1_0802
nDRB3_0202
nH-2-IAg7
-2.80
-2.24
-1.68
-1.12
-0.56
-0.00
0.56
1.12
1.67
2.23
2.79
3
25
46
N-term Amino Acid of 15-mer


19070175-JL2-JL2 AAL83744_1;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.35
0.70
1.05
1.40
1.75
2.10
2.45
2.80
3.15
3.50


19070175-JL2-JL2 AAL83744_1;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

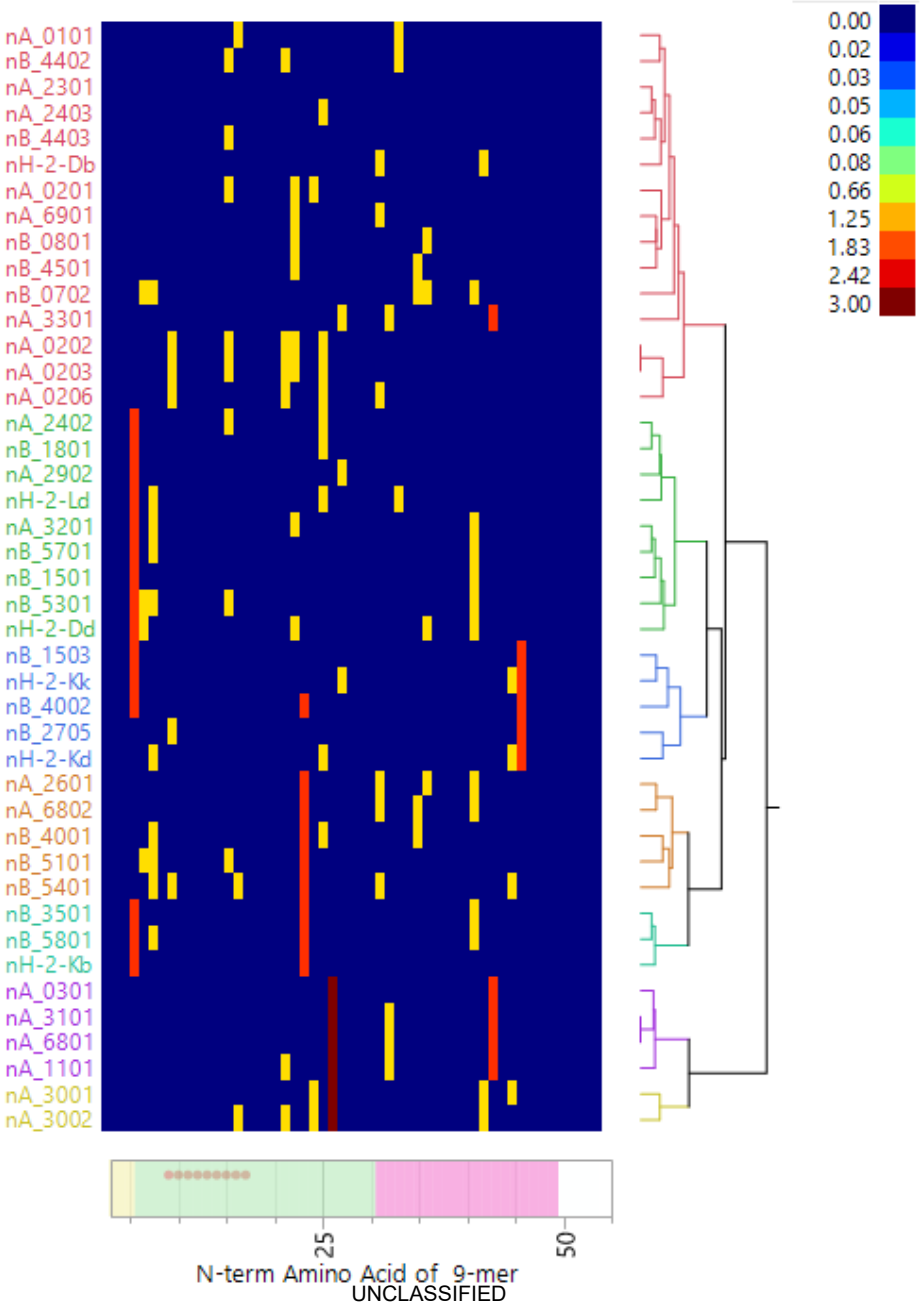

19070175-JL2-JL2 AAL83744_1;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

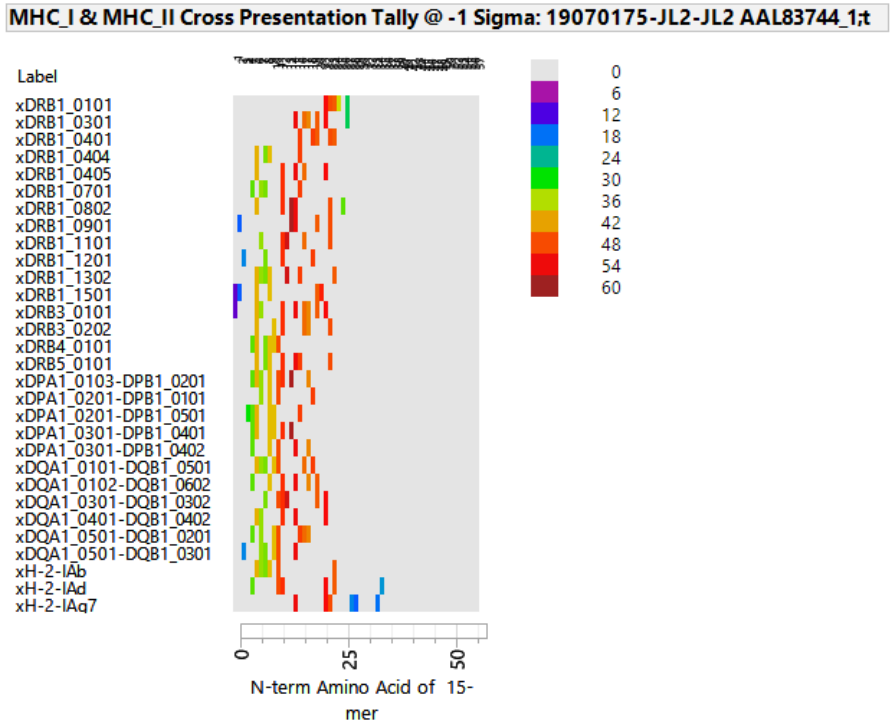

Hierarchical Cluster MHC I
Method = Ward
19070175-JL2-JL2 AAL83744_1;t_T_MHC I
nA_0101
nA_2601
nA_3201
nH-2-Dd
nB_1501
nB_3501
nB_5301
nB_5701
nB_5801
nA_2301
nA_2402
nA_2403
nA_2902
nB_1801
nB_4403
nB_1503
nB_4402
nA_0201
nA_0202
nA_0206
nA_0203
nA_6802
nA_6901
nB_0801
nH-2-Db
nH-2-Kb
nA_3001
nB_0702
nB_2705
nH-2-Ld
nB_4001
nB_4002
nH-2-Kd
nB_5101
nB_5401
nB_4501
nH-2-Kk
nA_0301
nA_1101
nA_3101
nA_3301
nA_3002
nA_6801
-2.97
-2.38
-1.78
-1.19
-0.59
-0.00
0.57
1.14
1.71
2.29
2.86
25
50
N-term Amino Acid of 9-mer

Hierarchical Cluster MHC I

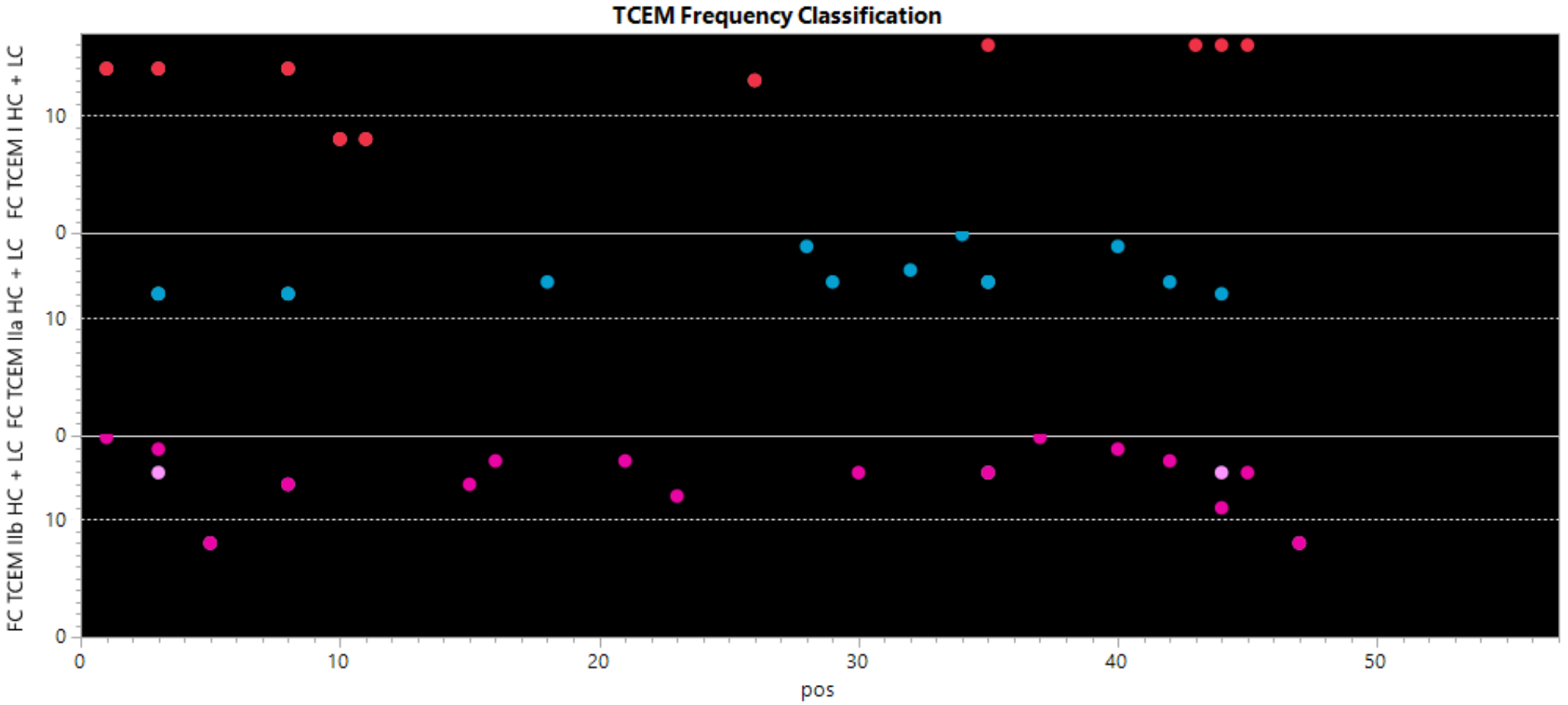
Method = Ward

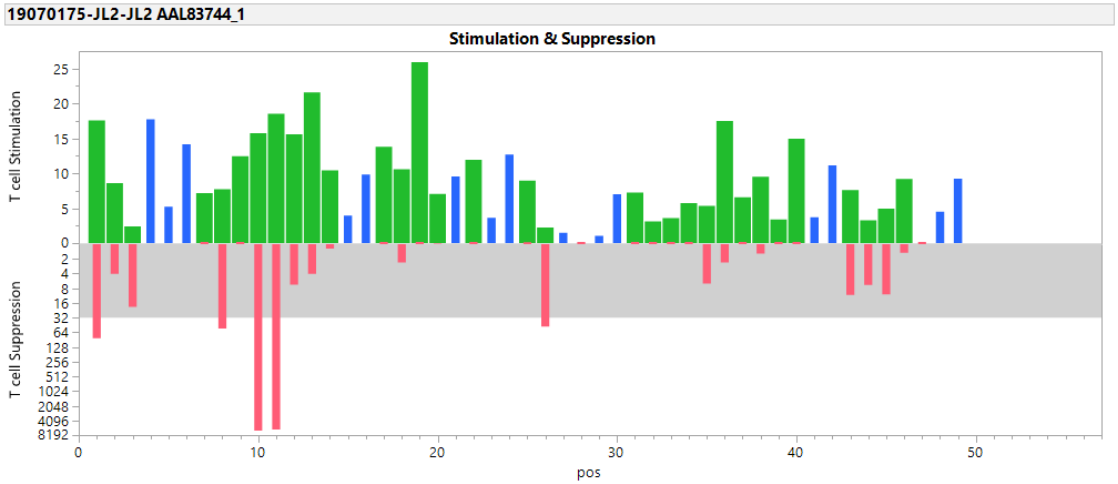
19070175-JL2-JL2 AAL83744_1;t MHC I & Cathepsins consolidated

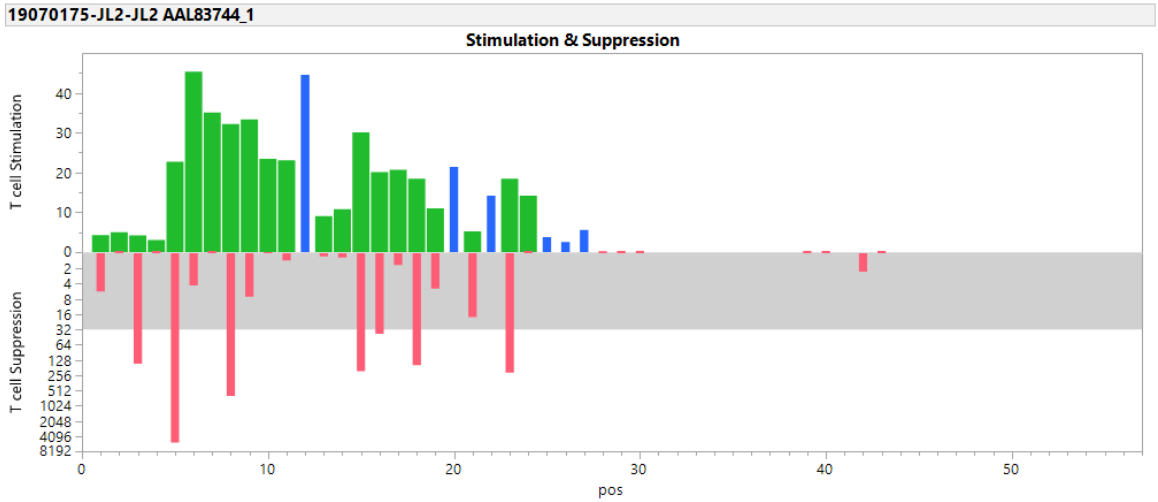



MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070175-JL2-JL2 AAL83744_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
6
12
18
24
30
36
42
48
54
60
0
25
50
N-term Amino Acid of 15-mer

19070175-JL2-JL2 AAL83744_1

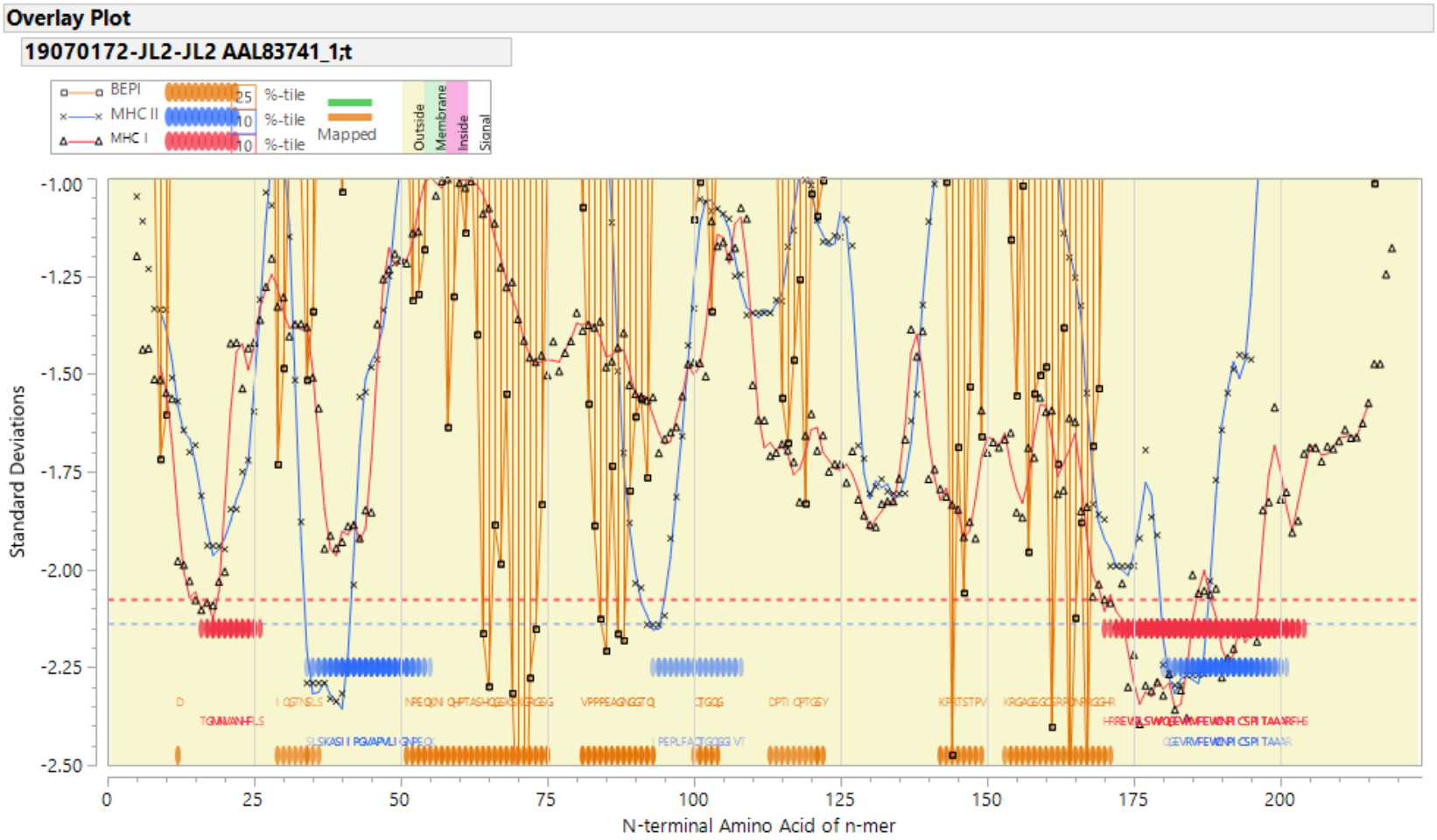


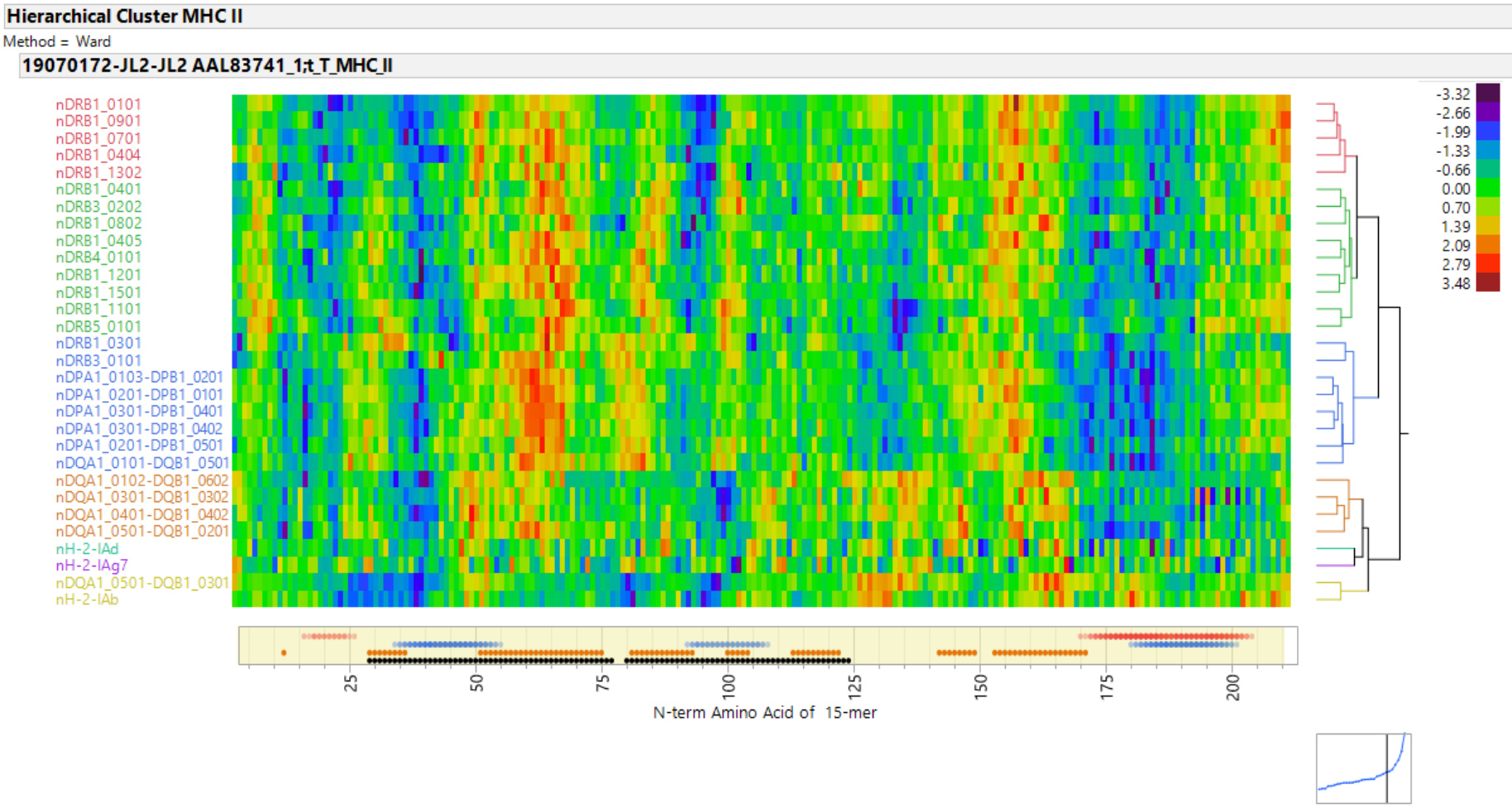

19070175-JL2-JL2 AAL83744_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

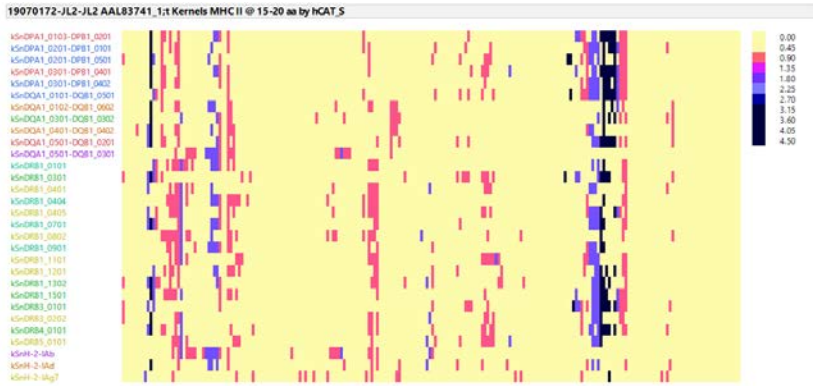

19070175-JL2-JL2 AAL83744_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

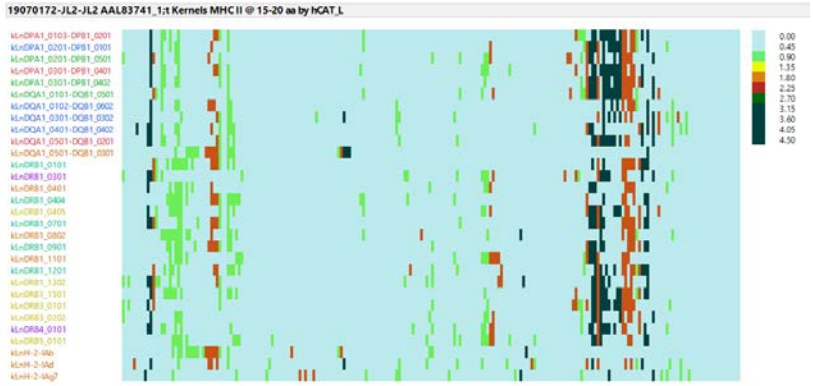
V protein

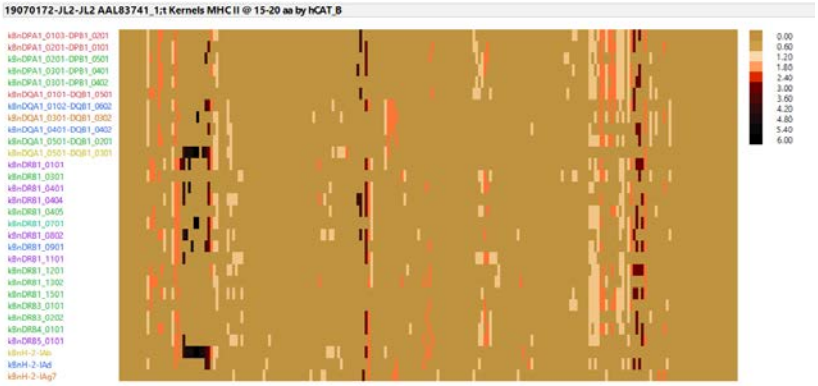
19070172-JL2-JL2 AAL83741_1| V protein

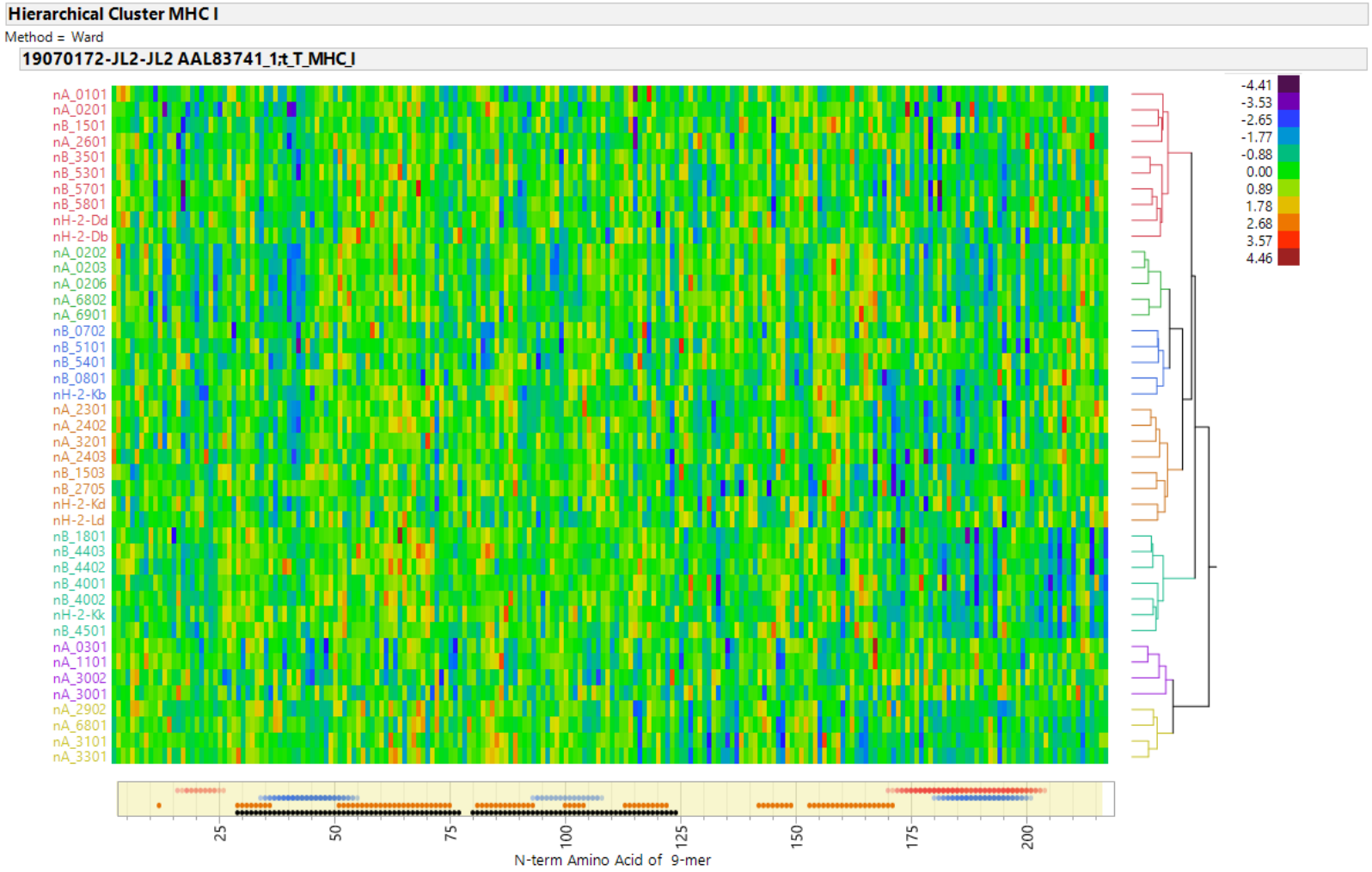

Overlay Plot
19070172-JL2-JL2 AAL83741_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
N-terminal Amino Acid of n-mer

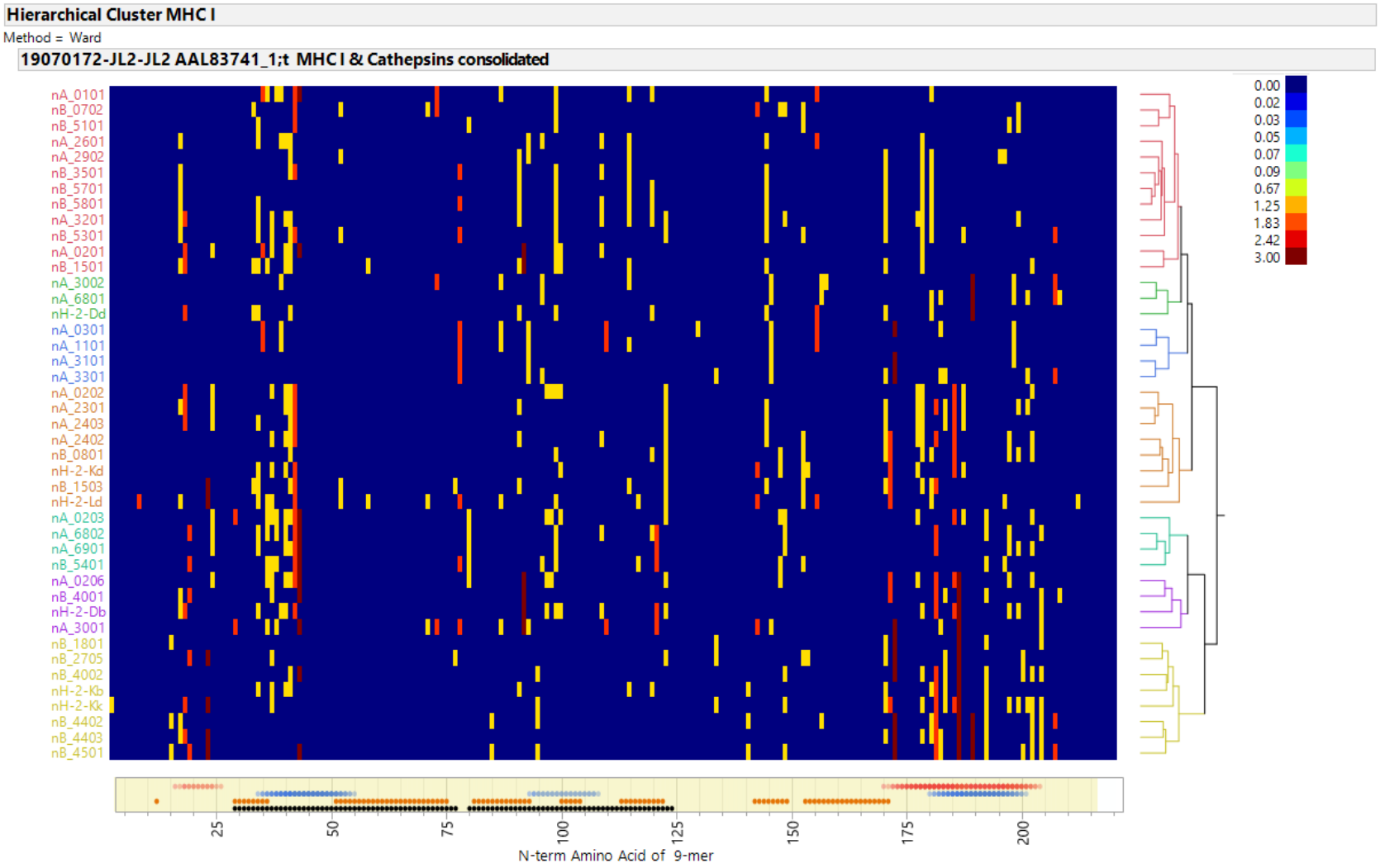

Hierarchical Cluster MHC II
Method = Ward
19070172-JL2-JL2 AAL83741_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_0405
nDRB4_0101
nDRB1_1201
nDRB1_1501
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.32
-2.66
-1.99
-1.33
-0.66
0.00
0.70
1.39
2.09
2.79
3.48
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

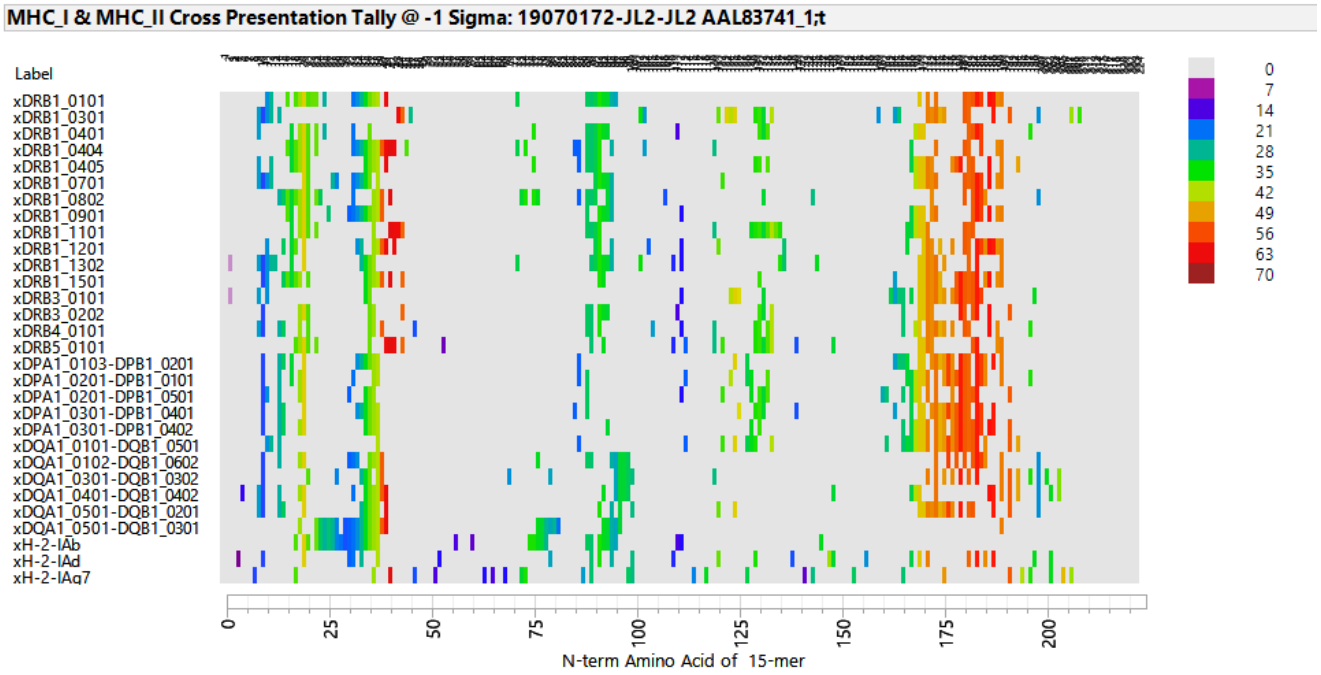

19070172-JL2-JL2 AAL83741_1;t Kernels MHC II @ 15-20 aa by hCAT_S


19070172-JL2-JL2 AAL83741_1;t Kernels MHC II @ 15-20 aa by hCAT_L

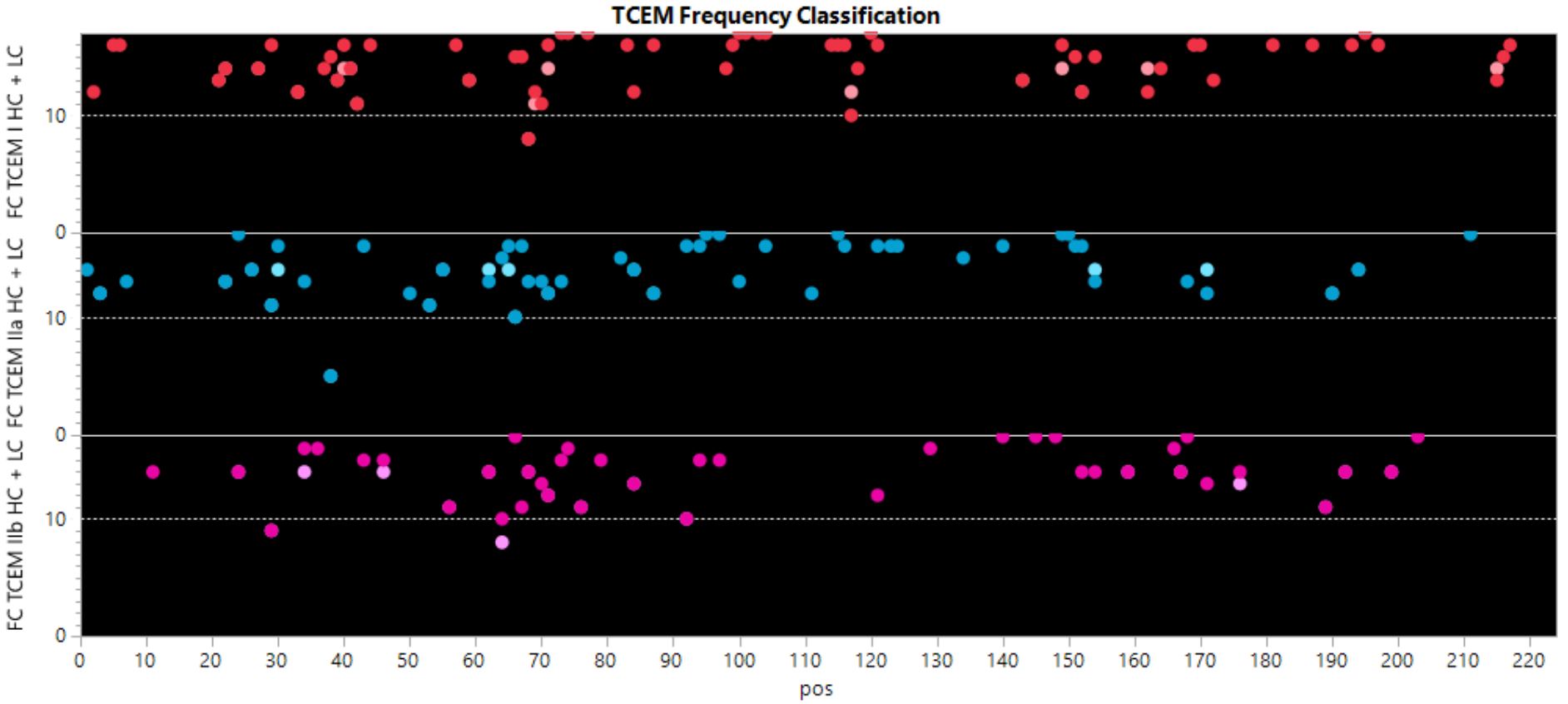

19070172-JL2-JL2 AAL83741_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
19070172-JL2-JL2 AAL83741_1;t_T_MHC_I
nA_0101
nA_0201
nB_1501
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nH-2-Db
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_5101
nB_5401
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_3001
nA_2902
nA_6801
nA_3101
nA_3301
-4.41
-3.53
-2.65
-1.77
-0.88
0.00
0.89
1.78
2.68
3.57
4.46
25
50
75
100
125
150
175
200
N-term Amino Acid of 9-mer

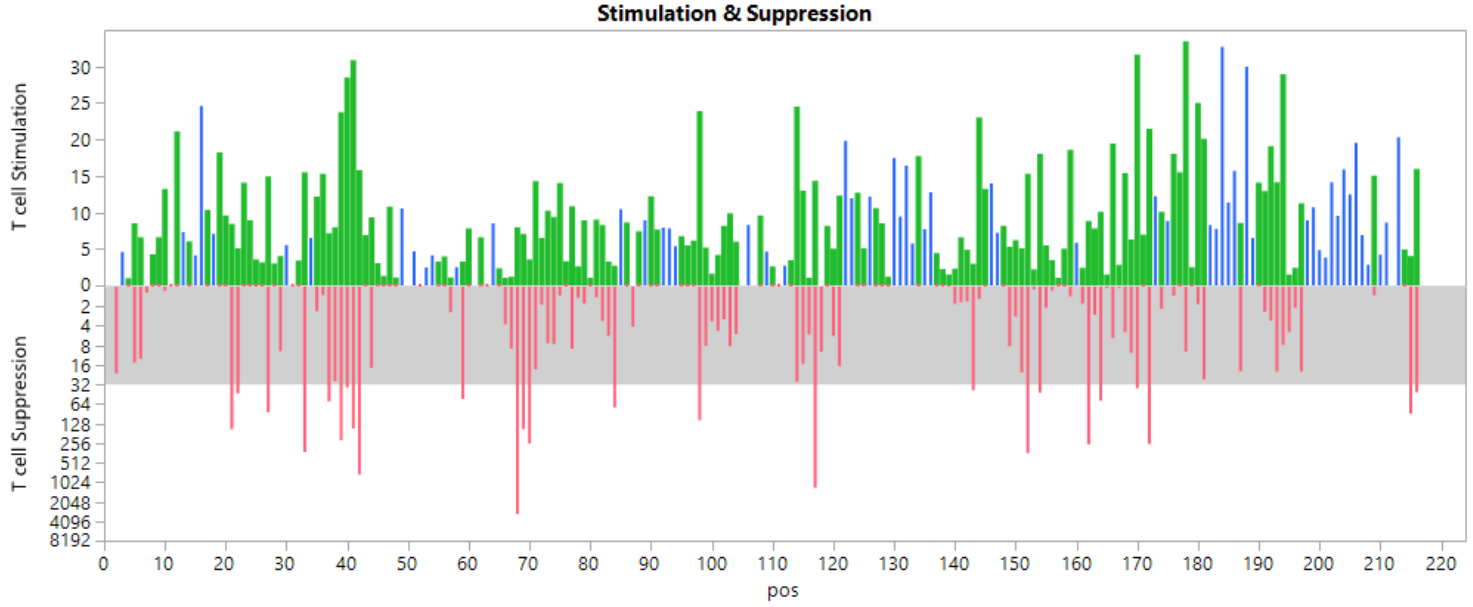

Hierarchical Cluster MHC I
Method = Ward
19070172-JL2-JL2 AAL83741_1;t MHC I & Cathepsins consolidated
nA_0101
nB_0702
nB_5101
nA_2601
nA_2902
nB_3501
nB_5701
nB_5801
nA_3201
nB_5301
nA_0201
nB_1501
nA_3002
nA_6801
nH-2-Dd
nA_0301
nA_1101
nA_3101
nA_3301
nA_0202
nA_2301
nA_2403
nA_2402
nB_0801
nH-2-Kd
nB_1503
nH-2-Ld
nA_0203
nA_6802
nA_6901
nB_5401
nA_0206
nB_4001
nH-2-Db
nA_3001
nB_1801
nB_2705
nB_4002
nH-2-Kb
nH-2-Kk
nB_4402
nB_4403
nB_4501
0.00
0.02
0.03
0.05
0.07
0.09
0.67
1.25
1.83
2.42
3.00
25
50
75
100
125
150
175
200
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 19070172-JL2-JL2 AAL83741_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer
0
7
14
21
28
35
42
49
56
63
70

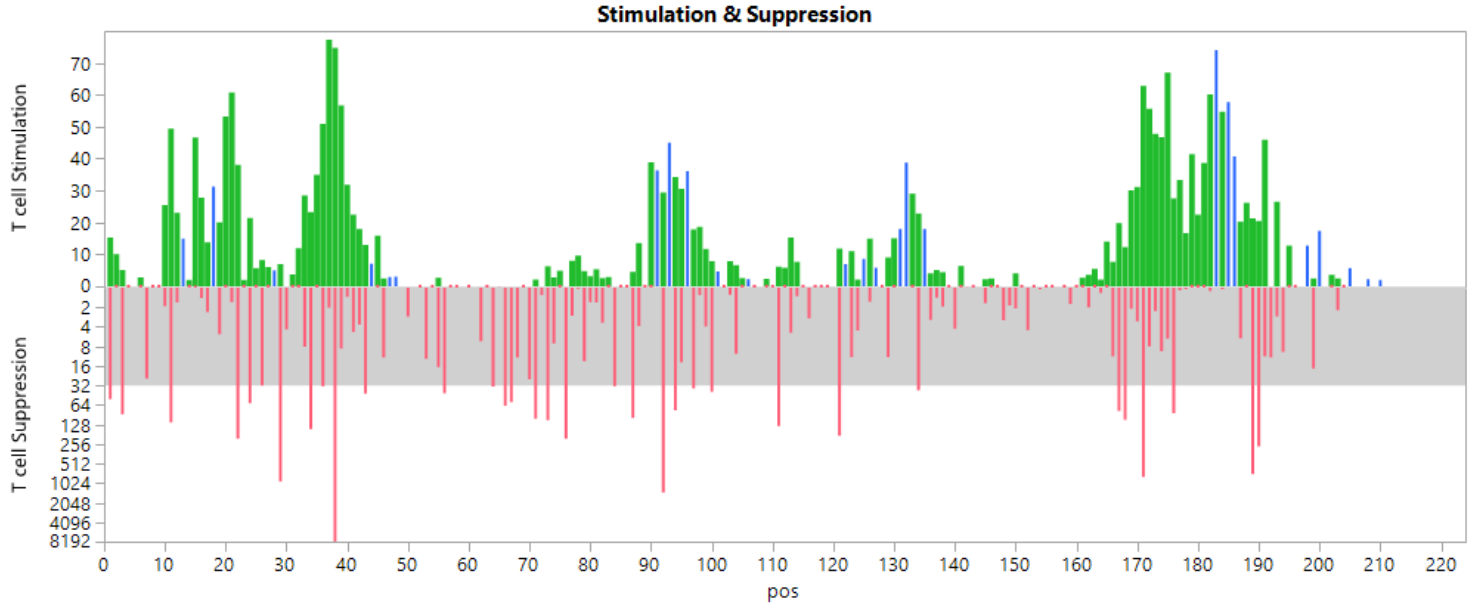
19070172-JL2-JL2 AAL83741_1



19070172-JL2-JL2 AAL83741_1



19070172-JL2-JL2 AAL83741_1



ioGenetics®

Mumps

MUMPSVAX® (Mumps Virus Vaccine Live), the Jeryl Lynn™ (B level) strain of mumps virus propagated in chick embryo cell culture;

Jeryl Lynn JL5 “major” 15077514

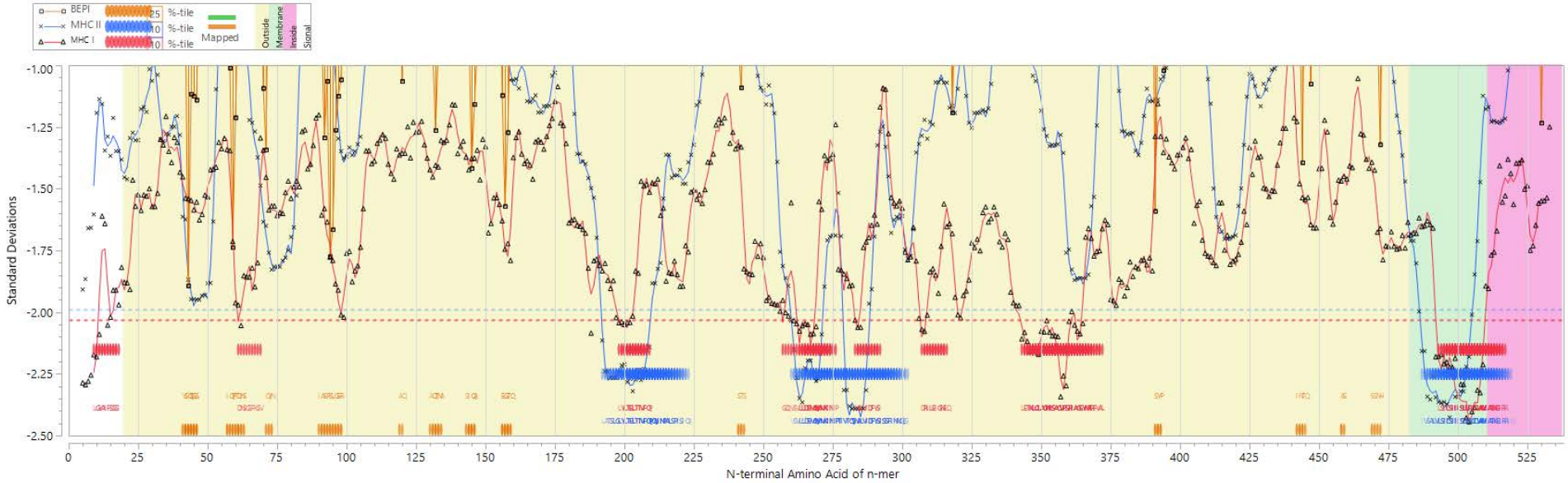
JL2 “minor”

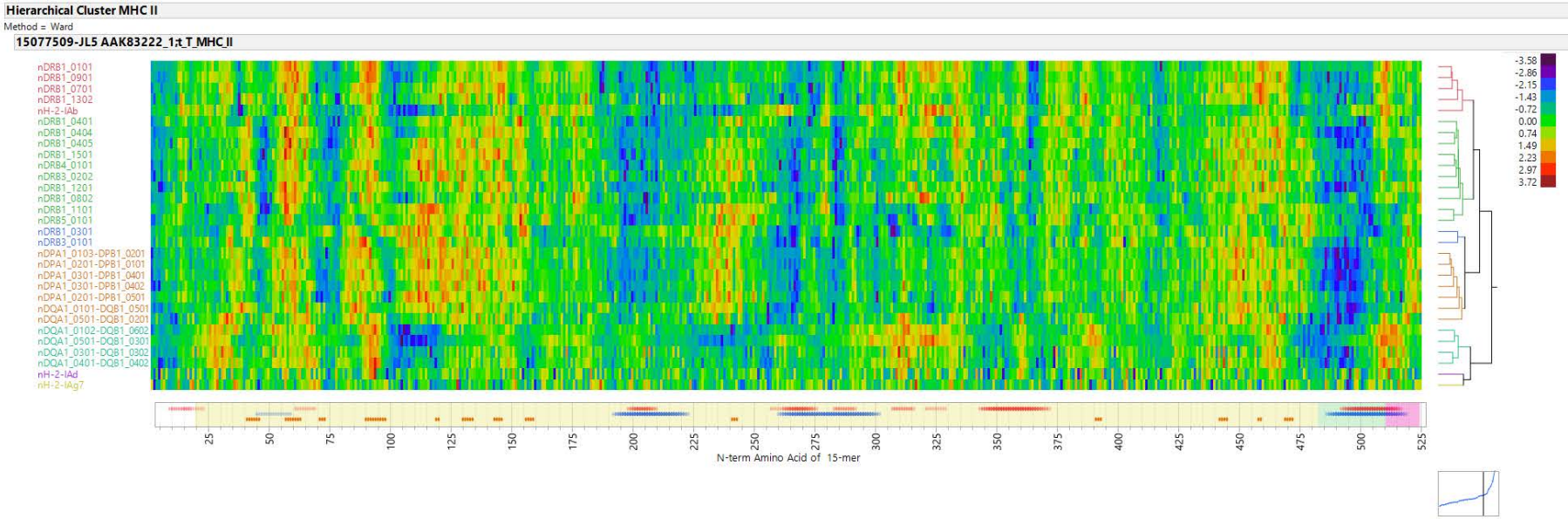
Fusion protein

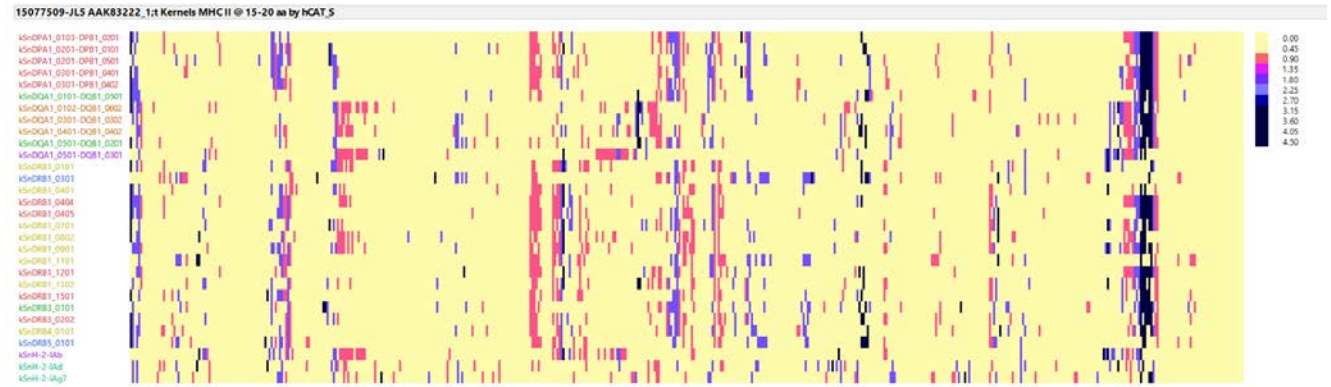
AAK83222.1| fusion protein

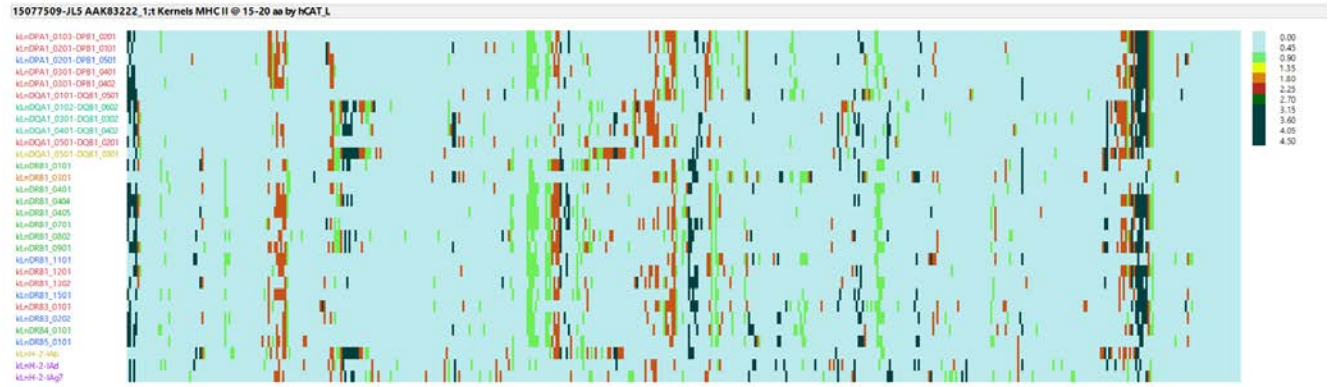
Overlay Plot

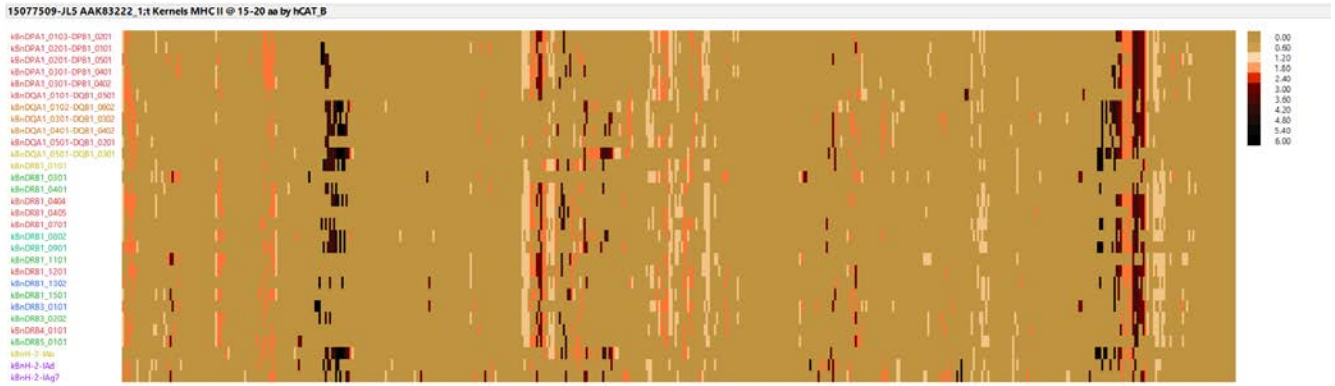
15077509-JL5 AAK83222_1;t

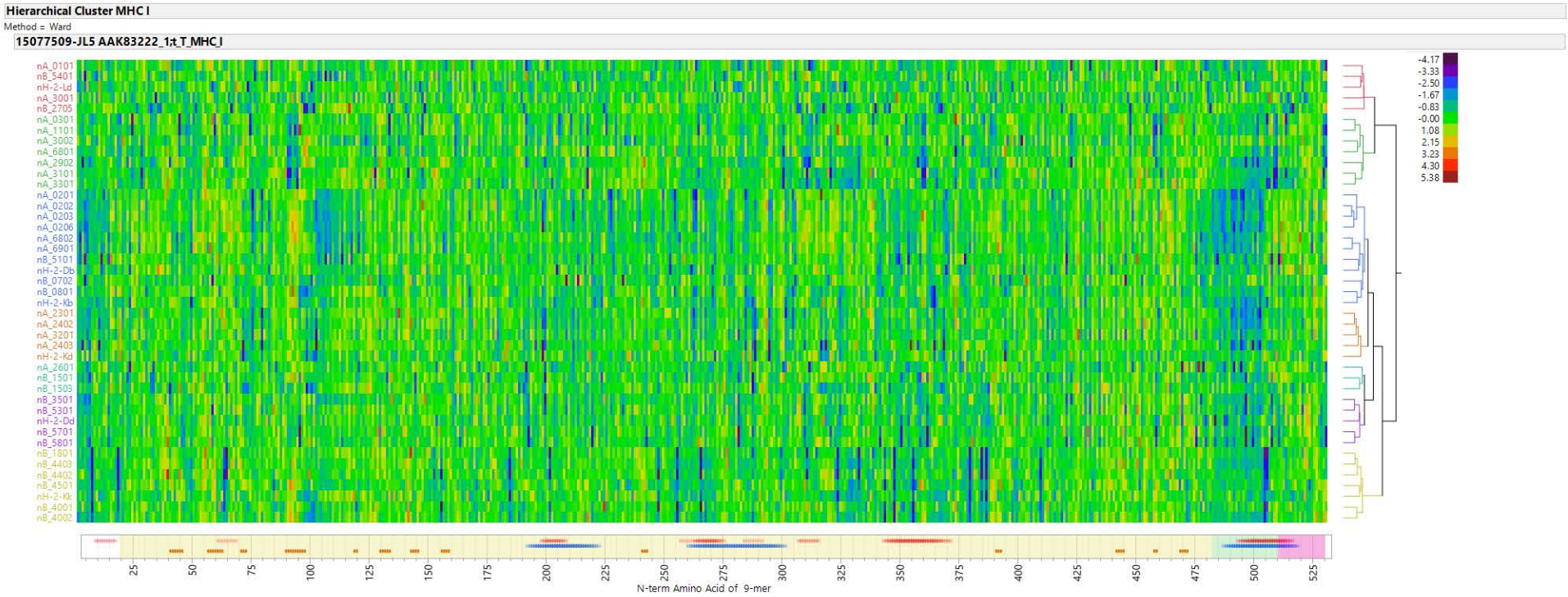


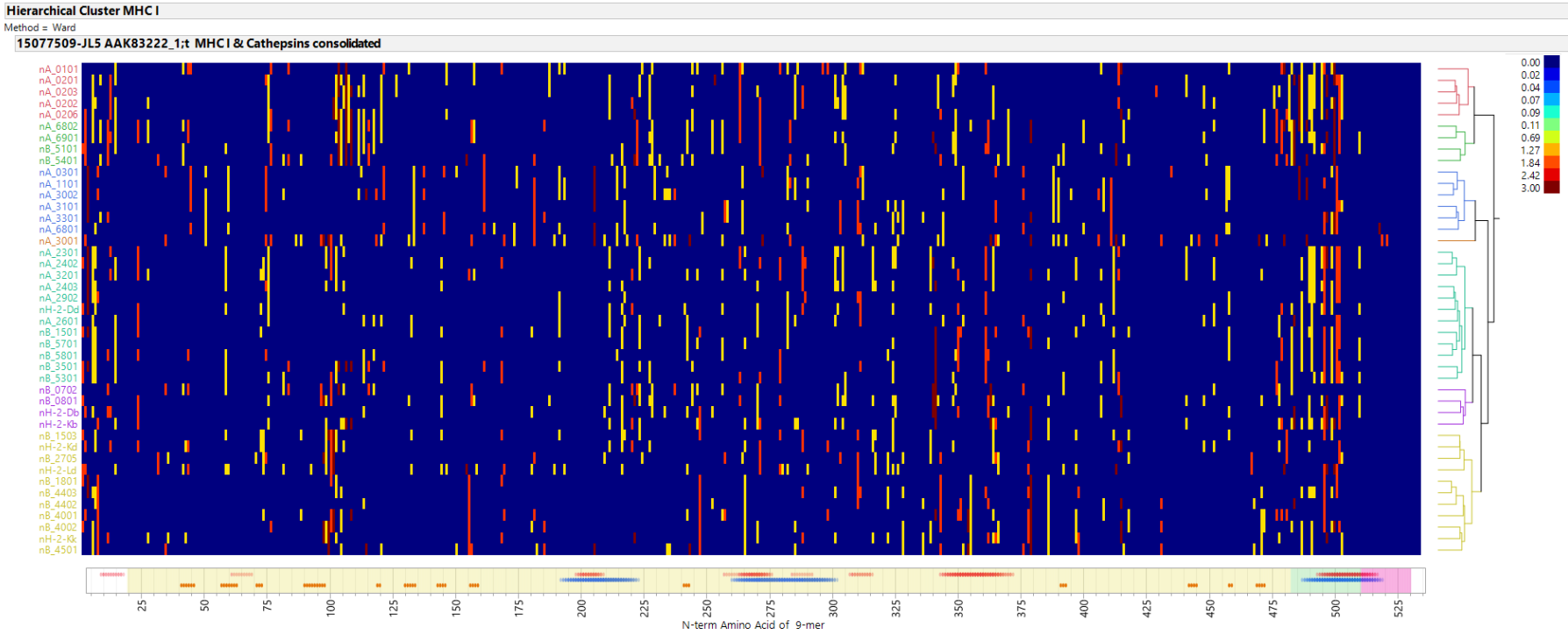

Hierarchical Cluster MHC II
Method = Ward
15077509-JL5 AAK83222_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nH-2-IAb
nDRB1_0401
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_1201
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0501-DQB1_0301
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
-3.58
-2.86
-2.15
-1.43
-0.72
0.00
0.74
1.49
2.23
2.97
3.72
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

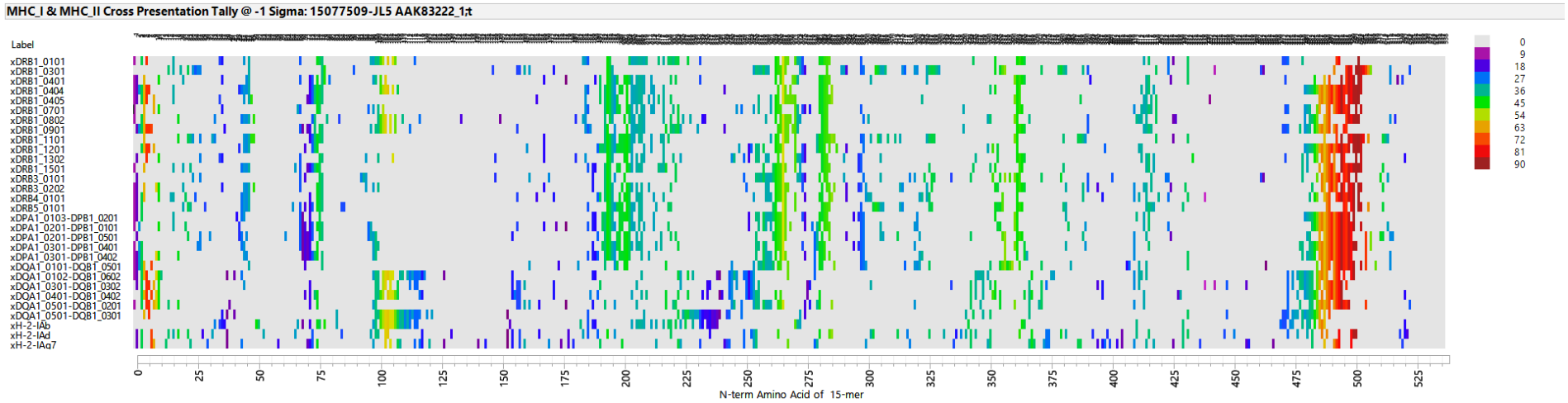

15077509-JL5 AAK83222_1;t Kernels MHC II @ 15-20 aa by hCAT_S


15077509-JL5 AAK83222_1;t Kernels MHC II @ 15-20 aa by hCAT_L

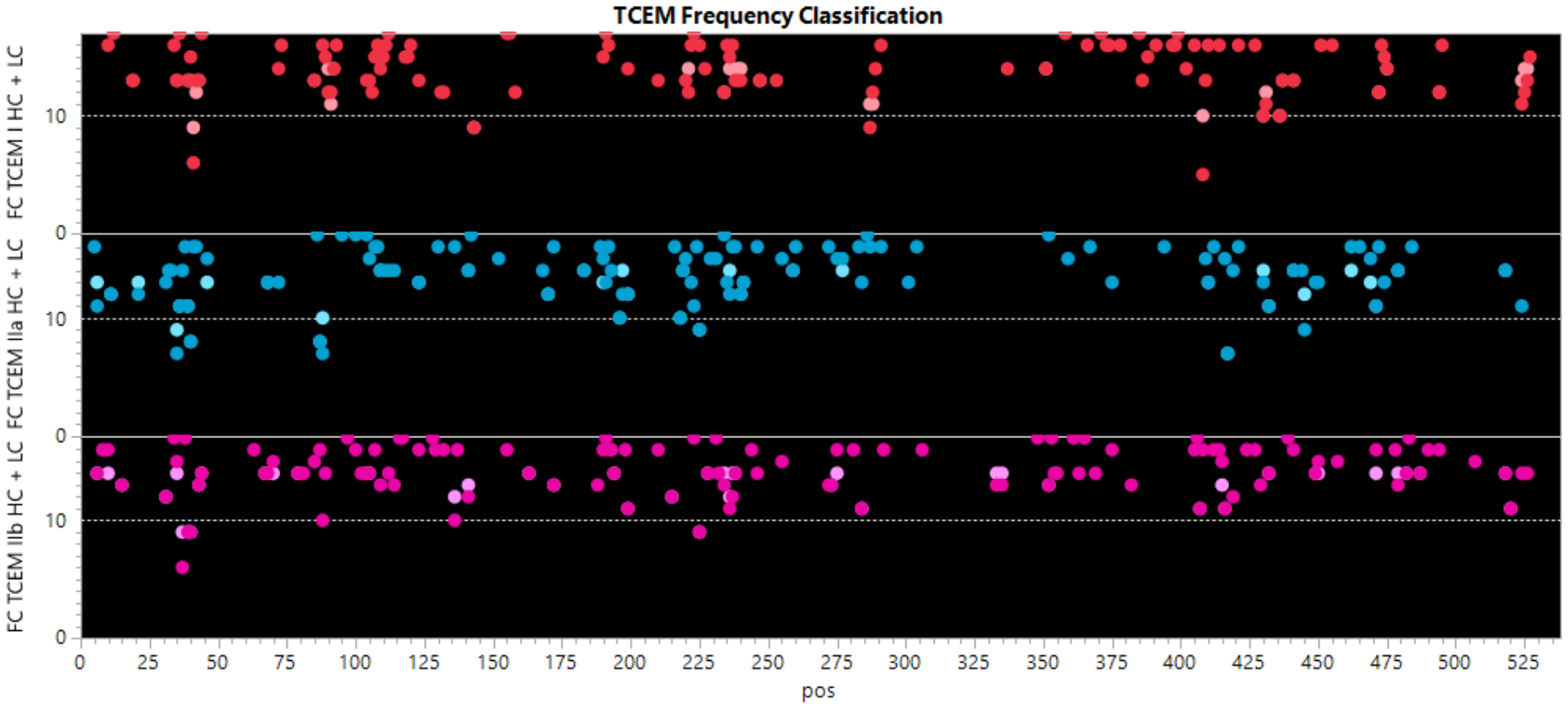

15077509-JL5 AAK83222_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
15077509-JL5 AAK83222_1:t_T_MHC_I
nA_0101
nB_5401
nH-2-Ld
nA_3001
nB_2705
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nA_2601
nB_1501
nB_1503
nB_3501
nB_5301
nH-2-Dd
nB_5701
nB_5801
nB_1801
nB_4403
nB_4402
nB_4501
nH-2-Kk
nB_4001
nB_4002
-4.17
-3.33
-2.50
-1.67
-0.83
-0.00
1.08
2.15
3.23
4.30
5.38
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
15077509-JL5 AAK83222_1;t MHC I & Cathepsins consolidated
nA_0101
nA_0201
nA_0203
nA_0202
nA_0206
nA_6802
nA_6901
nB_5101
nB_5401
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nA_2301
nA_2402
nA_3201
nA_2403
nA_2902
nH-2-Dd
nA_2601
nB_1501
nB_5701
nB_5801
nB_3501
nB_5301
nB_0702
nB_0801
nH-2-Db
nH-2-Kb
nB_1503
nH-2-Kd
nB_2705
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077509-JL5 AAK83222_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525
N-term Amino Acid of 15-mer
0
9
18
27
36
45
54
63
72
81
90

15077509-JL5 AAK83222_1



15077509-JL5 AAK83222_1

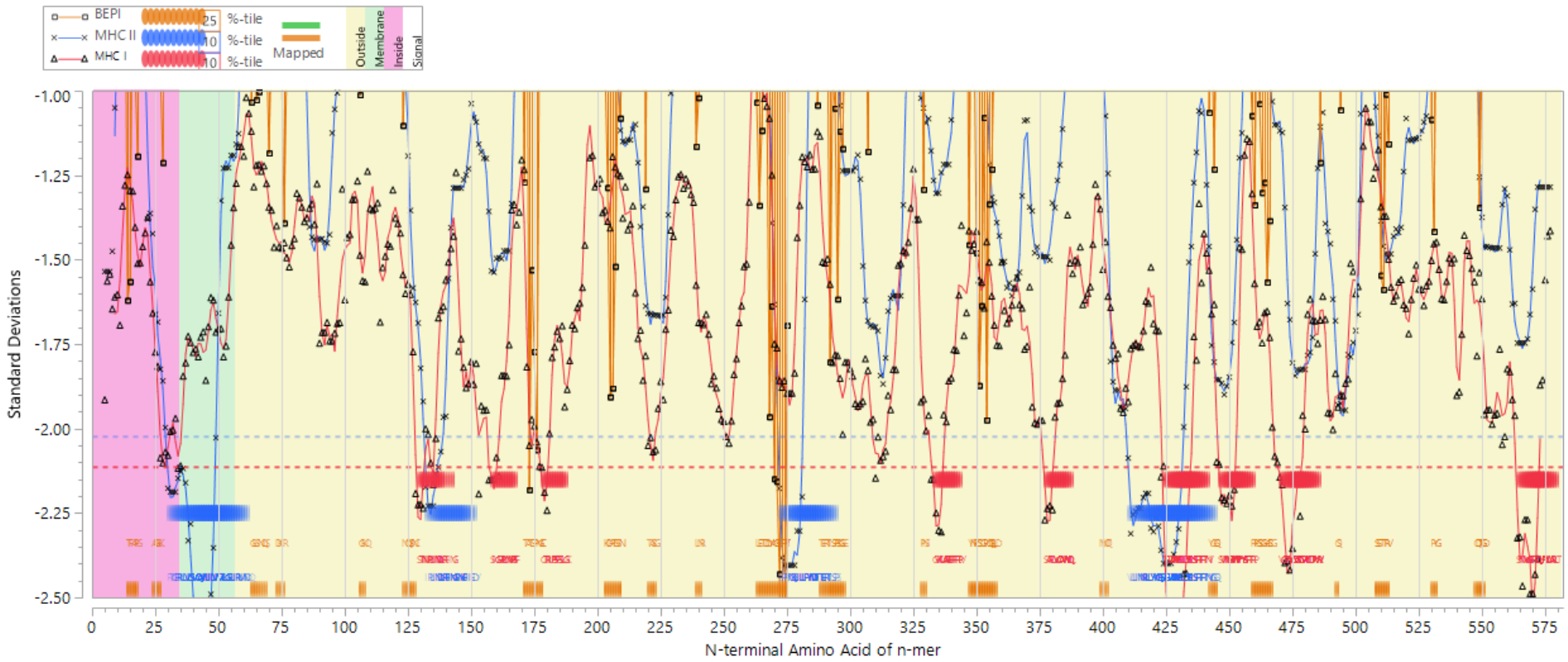


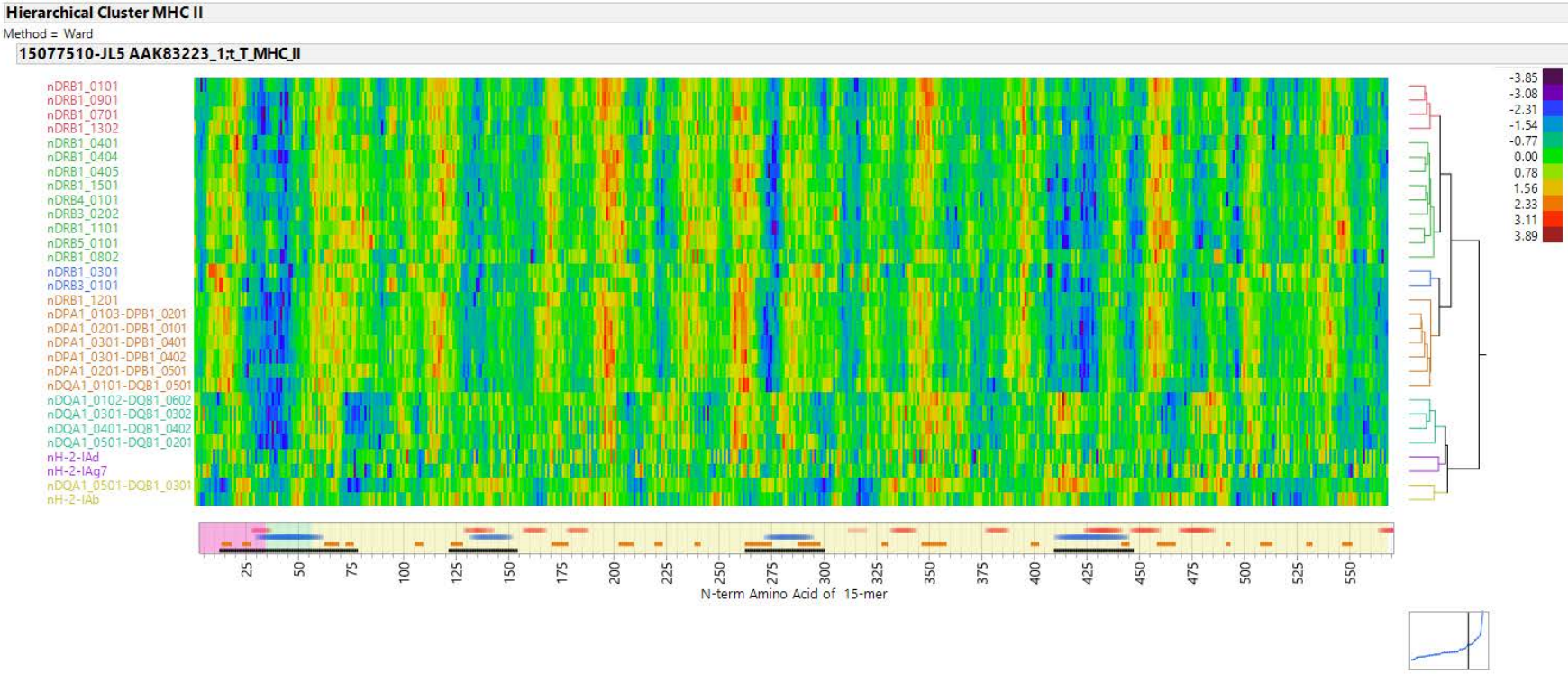
Hemagglutinin/Neuraminidase

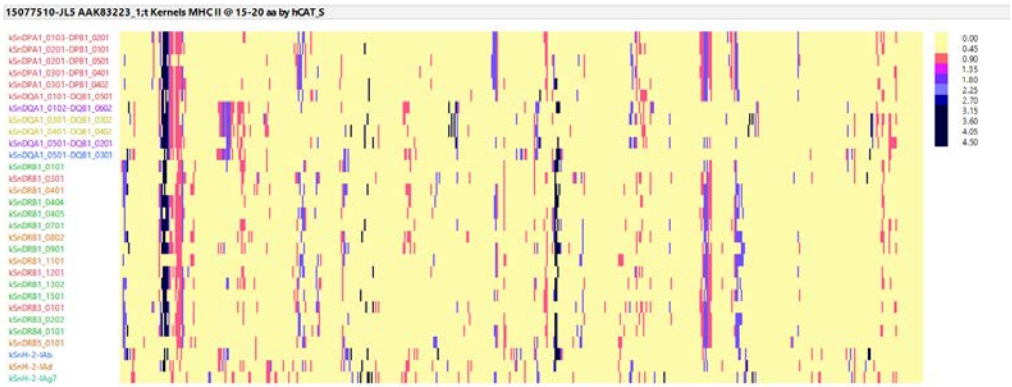
AAK83223.1| hemagglutinin/neuraminidase

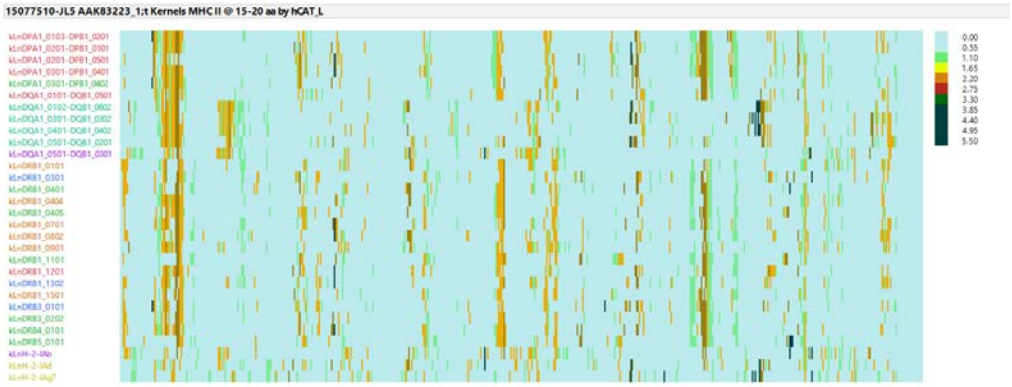
Overlay Plot

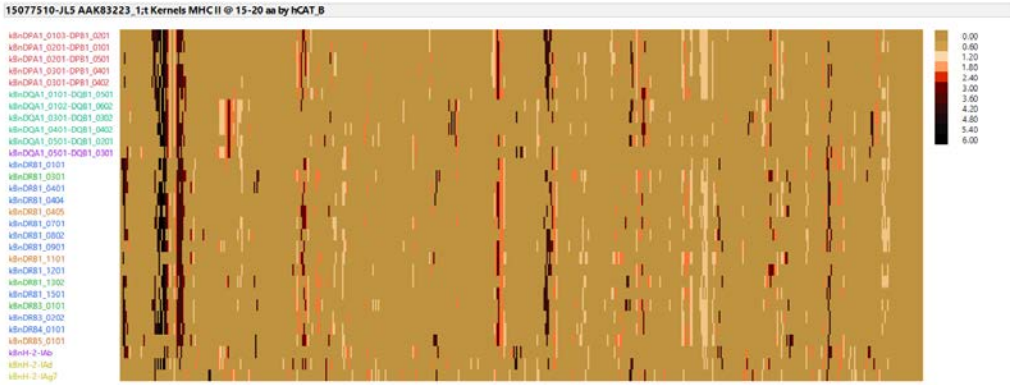
15077510-JL5 AAK83223_1;t

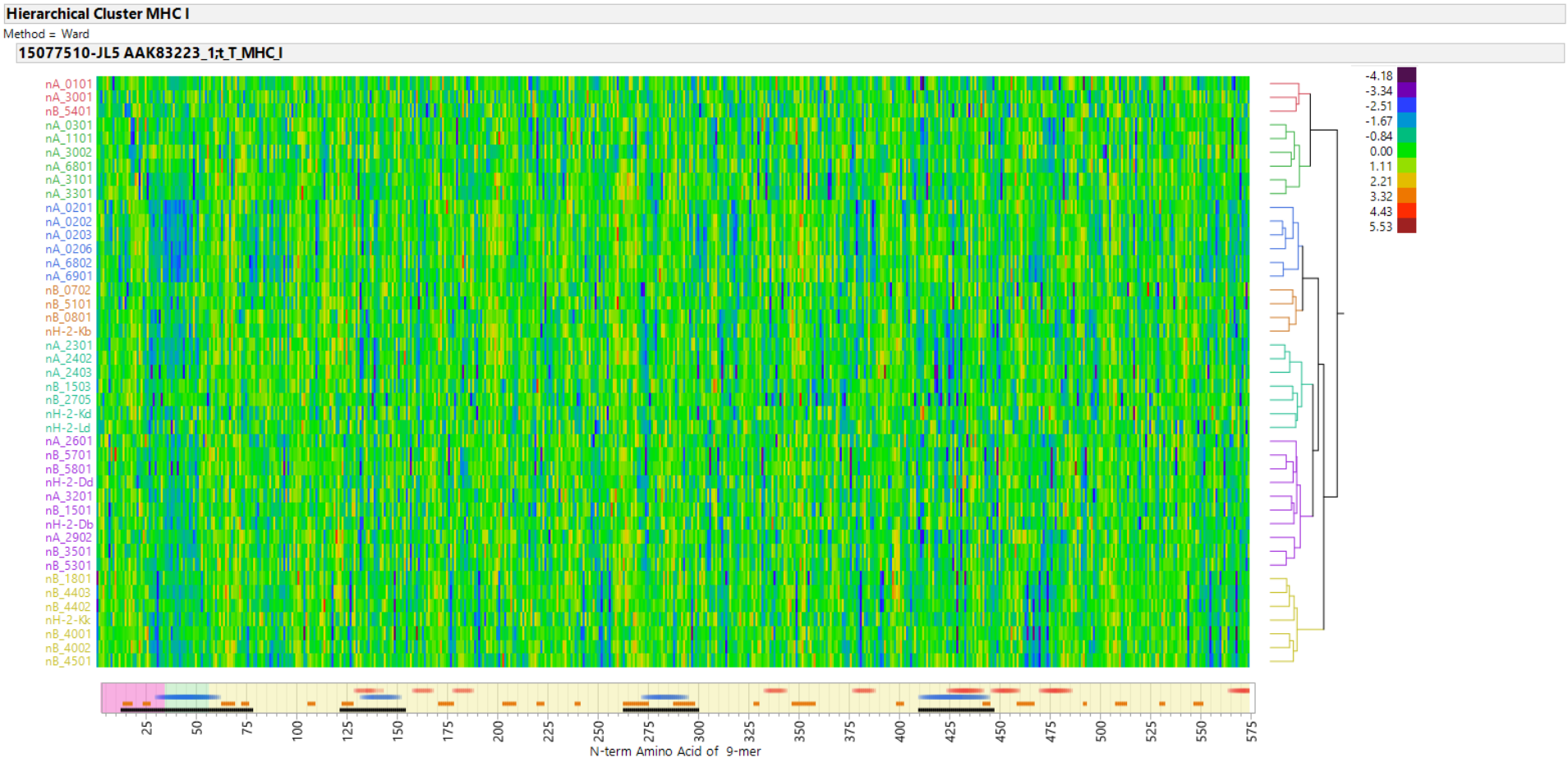


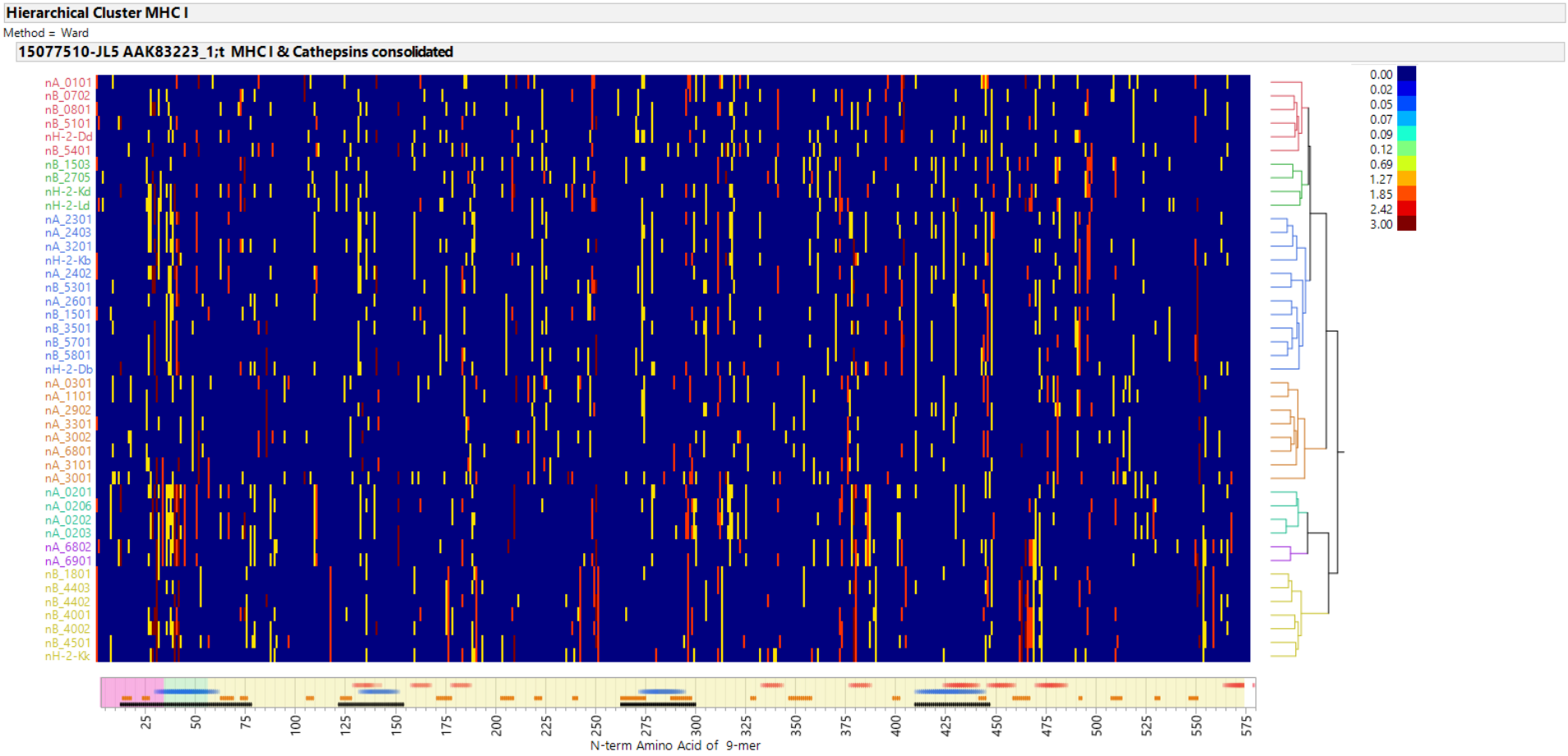

Hierarchical Cluster MHC II
Method = Ward
15077510-JL5 AAK83223_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.85
-3.08
-2.31
-1.54
-0.77
0.00
0.78
1.56
2.33
3.11
3.89
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 15-mer

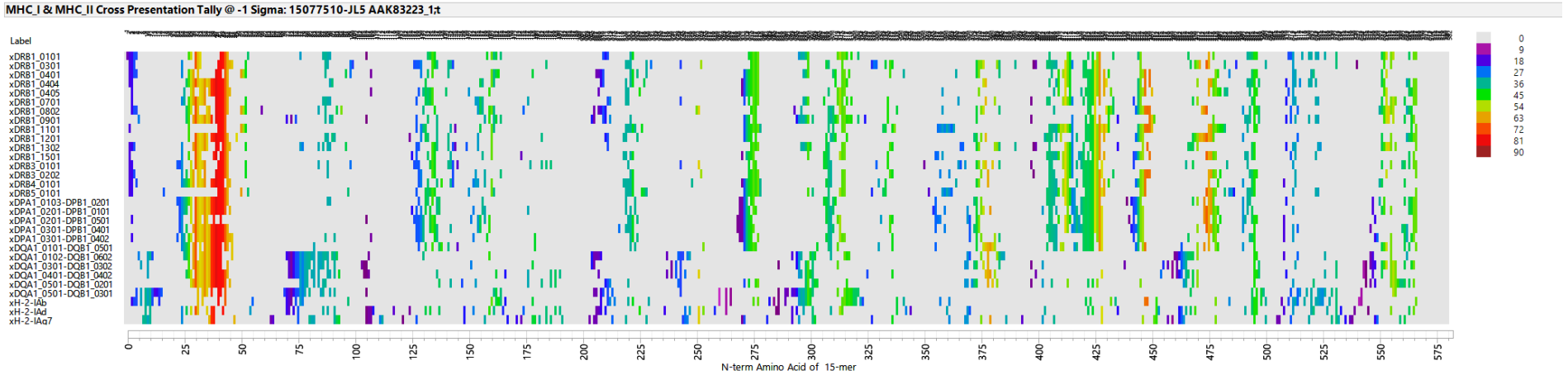

15077510-JL5 AAK83223_1;t Kernels MHC II @ 15-20 aa by hCAT_S
kSmDPA1_0103-DPB1_0201
kSmDPA1_0201-DPB1_0101
kSmDPA1_0201-DPB1_0501
kSmDPA1_0301-DPB1_0401
kSmDPA1_0301-DPB1_0402
kSmDQA1_0101-DQB1_0501
kSmDQA1_0102-DQB1_0602
kSmDQA1_0301-DQB1_0302
kSmDQA1_0401-DQB1_0402
kSmDQA1_0501-DQB1_0201
kSmDQA1_0501-DQB1_0301
kSmDRB1_0101
kSmDRB1_0301
kSmDRB1_0401
kSmDRB1_0404
kSmDRB1_0405
kSmDRB1_0701
kSmDRB1_0802
kSmDRB1_0901
kSmDRB1_1101
kSmDRB1_1201
kSmDRB1_1302
kSmDRB1_1501
kSmDRB3_0101
kSmDRB3_0202
kSmDRB4_0101
kSmDRB5_0101
kSmH-2-IAb
kSmH-2-IAd
kSmH-2-IAg7
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50


15077510-JL5 AAK83223_1;t Kernels MHC II @ 15-20 aa by hCAT_L
kLmDPA1_0103-DPB1_0201
kLmDPA1_0201-DPB1_0101
kLmDPA1_0201-DPB1_0501
kLmDPA1_0301-DPB1_0401
kLmDPA1_0301-DPB1_0402
kLmDQA1_0101-DQB1_0501
kLmDQA1_0102-DQB1_0602
kLmDQA1_0301-DQB1_0302
kLmDQA1_0401-DQB1_0402
kLmDQA1_0501-DQB1_0201
kLmDQA1_0501-DQB1_0301
kLmDRB1_0101
kLmDRB1_0301
kLmDRB1_0401
kLmDRB1_0404
kLmDRB1_0405
kLmDRB1_0701
kLmDRB1_0802
kLmDRB1_0901
kLmDRB1_1101
kLmDRB1_1201
kLmDRB1_1302
kLmDRB1_1501
kLmDRB3_0101
kLmDRB3_0202
kLmDRB4_0101
kLmDRB5_0101
kLmH-2-IAb
kLmH-2-IAd
kLmH-2-IAg7
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

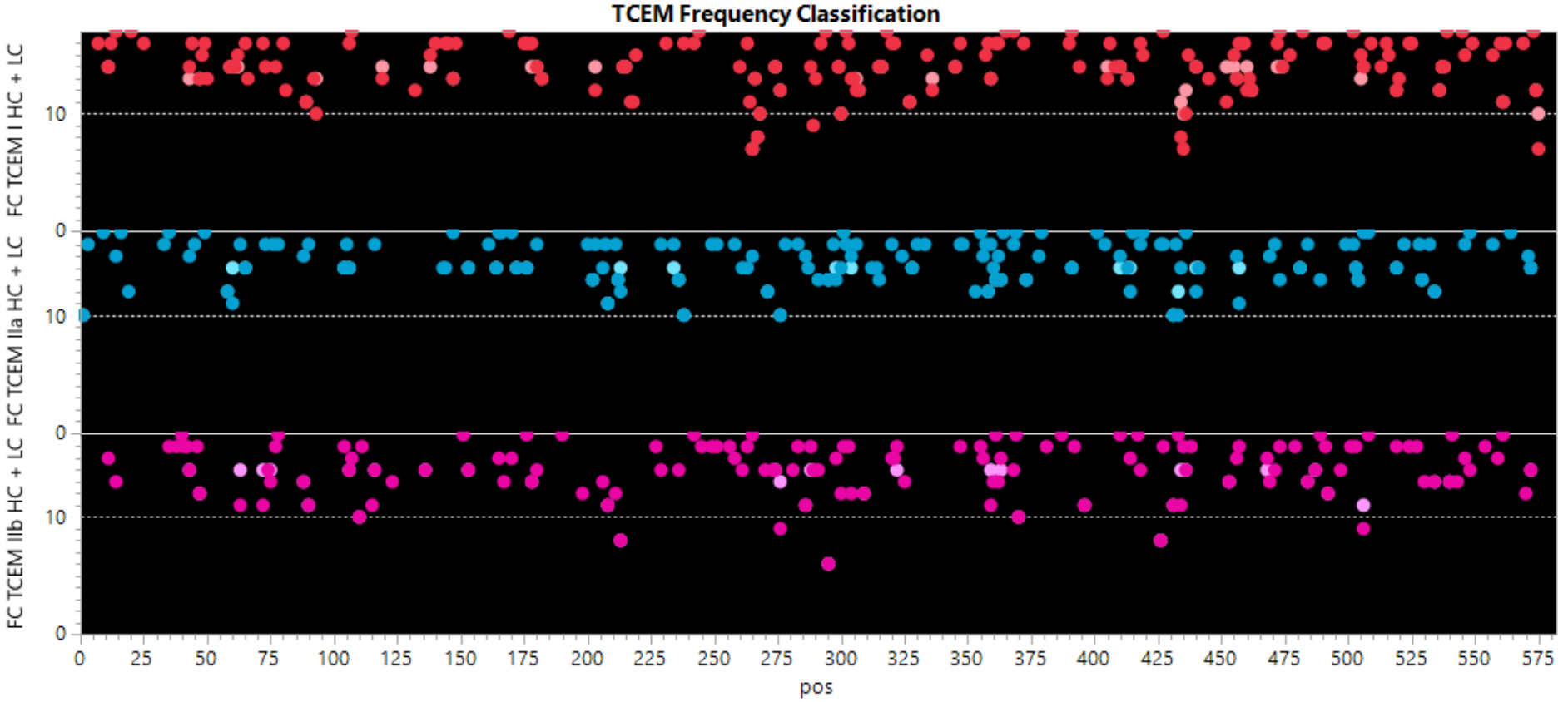

15077510-JL5 AAK83223_1;t Kernels MHC II @ 15-20 aa by hCAT_B
kBmDPA1_0103-DPB1_0201
kBmDPA1_0201-DPB1_0101
kBmDPA1_0201-DPB1_0501
kBmDPA1_0301-DPB1_0401
kBmDPA1_0301-DPB1_0402
kBmDQA1_0101-DQB1_0501
kBmDQA1_0102-DQB1_0602
kBmDQA1_0301-DQB1_0302
kBmDQA1_0401-DQB1_0402
kBmDQA1_0501-DQB1_0201
kBmDQA1_0501-DQB1_0301
kBmDRB1_0101
kBmDRB1_0301
kBmDRB1_0401
kBmDRB1_0404
kBmDRB1_0405
kBmDRB1_0701
kBmDRB1_0802
kBmDRB1_0901
kBmDRB1_1101
kBmDRB1_1201
kBmDRB1_1302
kBmDRB1_1501
kBmDRB3_0101
kBmDRB3_0202
kBmDRB4_0101
kBmDRB5_0101
kBmH-2-IAb
kBmH-2-IAd
kBmH-2-IAg7
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

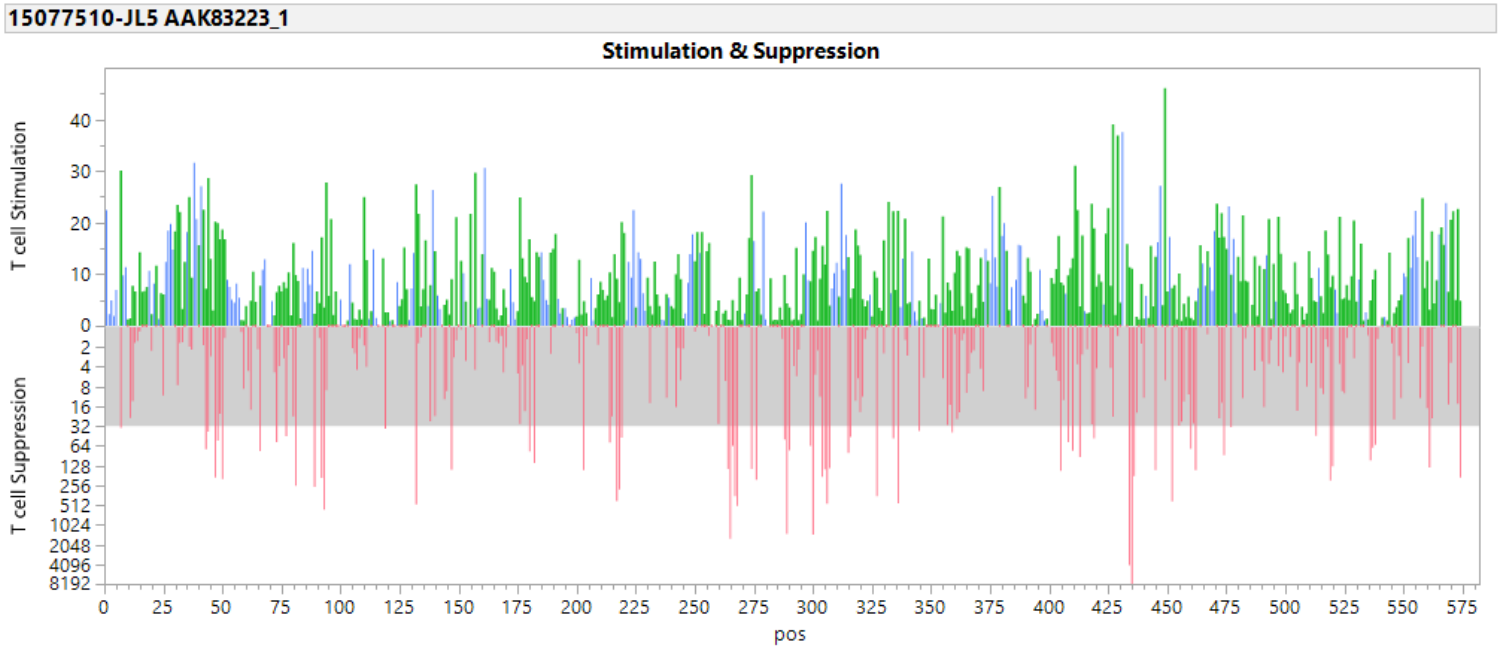

Hierarchical Cluster MHC I
Method = Ward
15077510-JL5 AAK83223_1;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_5101
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_2601
nB_5701
nB_5801
nH-2-Dd
nA_3201
nB_1501
nH-2-Db
nA_2902
nB_3501
nB_5301
nB_1801
nB_4403
nB_4402
nH-2-Kk
nB_4001
nB_4002
nB_4501
-4.18
-3.34
-2.51
-1.67
-0.84
0.00
1.11
2.21
3.32
4.43
5.53
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
N-term Amino Acid of 9-mer

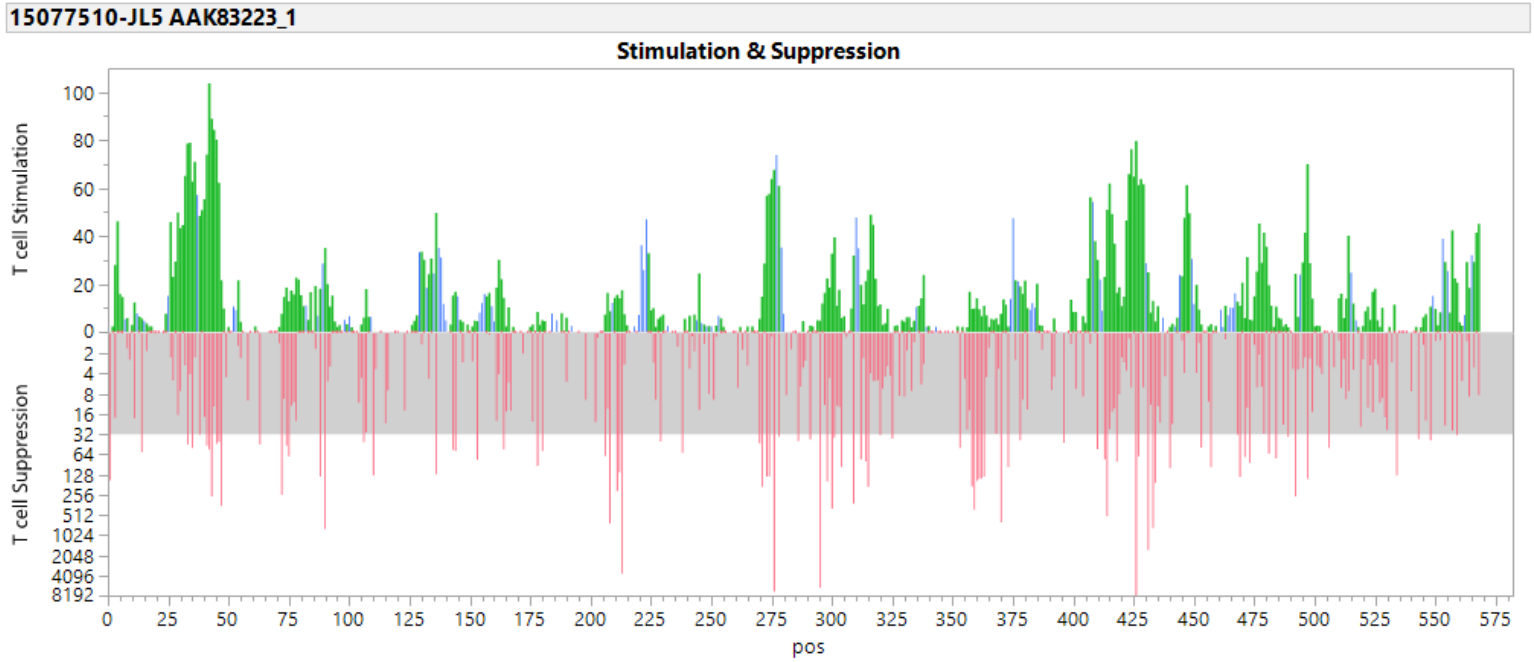

Hierarchical Cluster MHC I
Method = Ward
15077510-JL5 AAK83223_1;t MHC I & Cathepsins consolidated
nA_0101
nB_0702
nB_0801
nB_5101
nH-2-Dd
nB_5401
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_2301
nA_2403
nA_3201
nH-2-Kb
nA_2402
nB_5301
nA_2601
nB_1501
nB_3501
nB_5701
nB_5801
nH-2-Db
nA_0301
nA_1101
nA_2902
nA_3301
nA_3002
nA_6801
nA_3101
nA_3001
nA_0201
nA_0206
nA_0202
nA_0203
nA_6802
nA_6901
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
0.00
0.02
0.05
0.07
0.09
0.12
0.69
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077510-JL5 AAK83223_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575
N-term Amino Acid of 15-mer
0 9 18 27 36 45 54 63 72 81 90

15077510-JL5 AAK83223_1




15077510-JL5 AAK83223_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

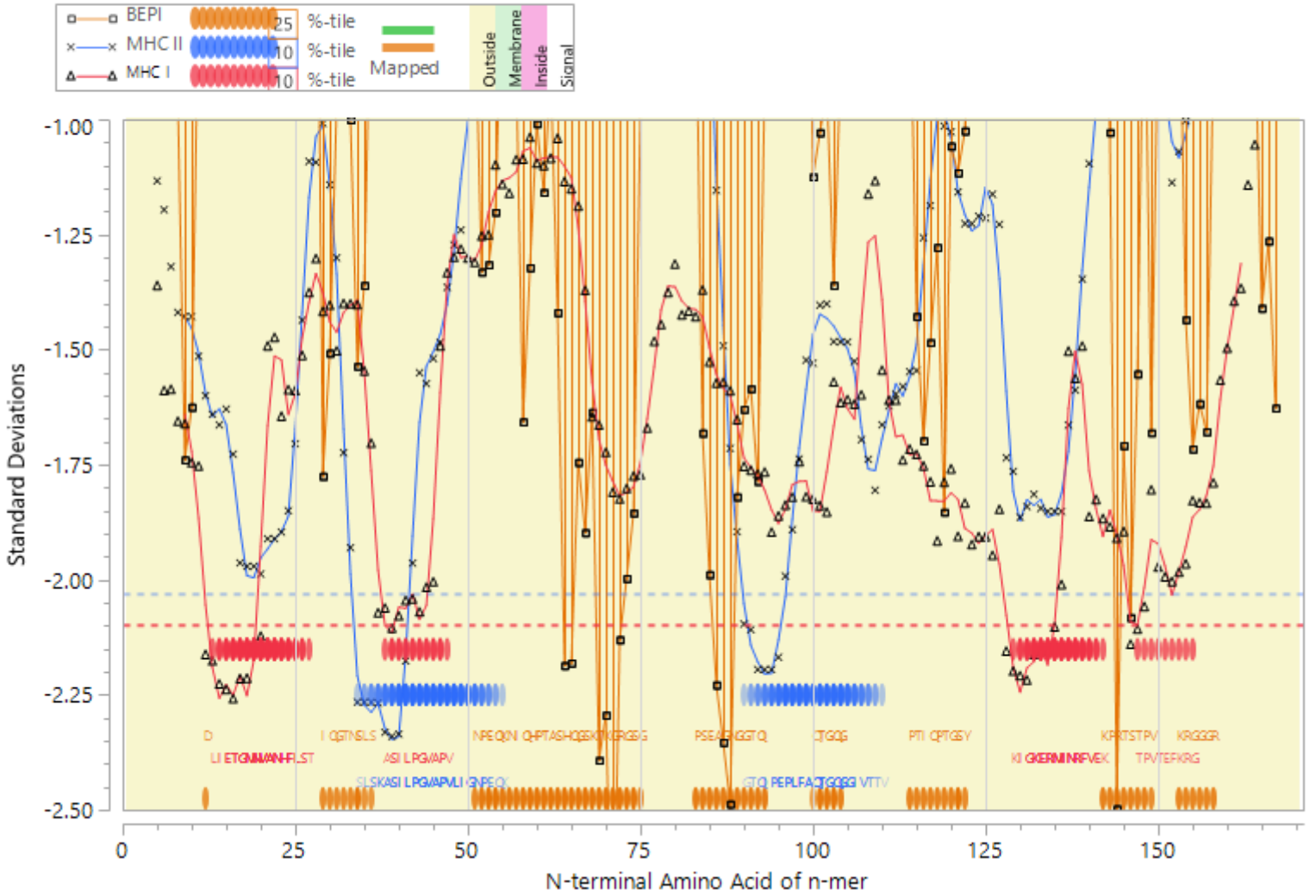

15077510-JL5 AAK83223_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

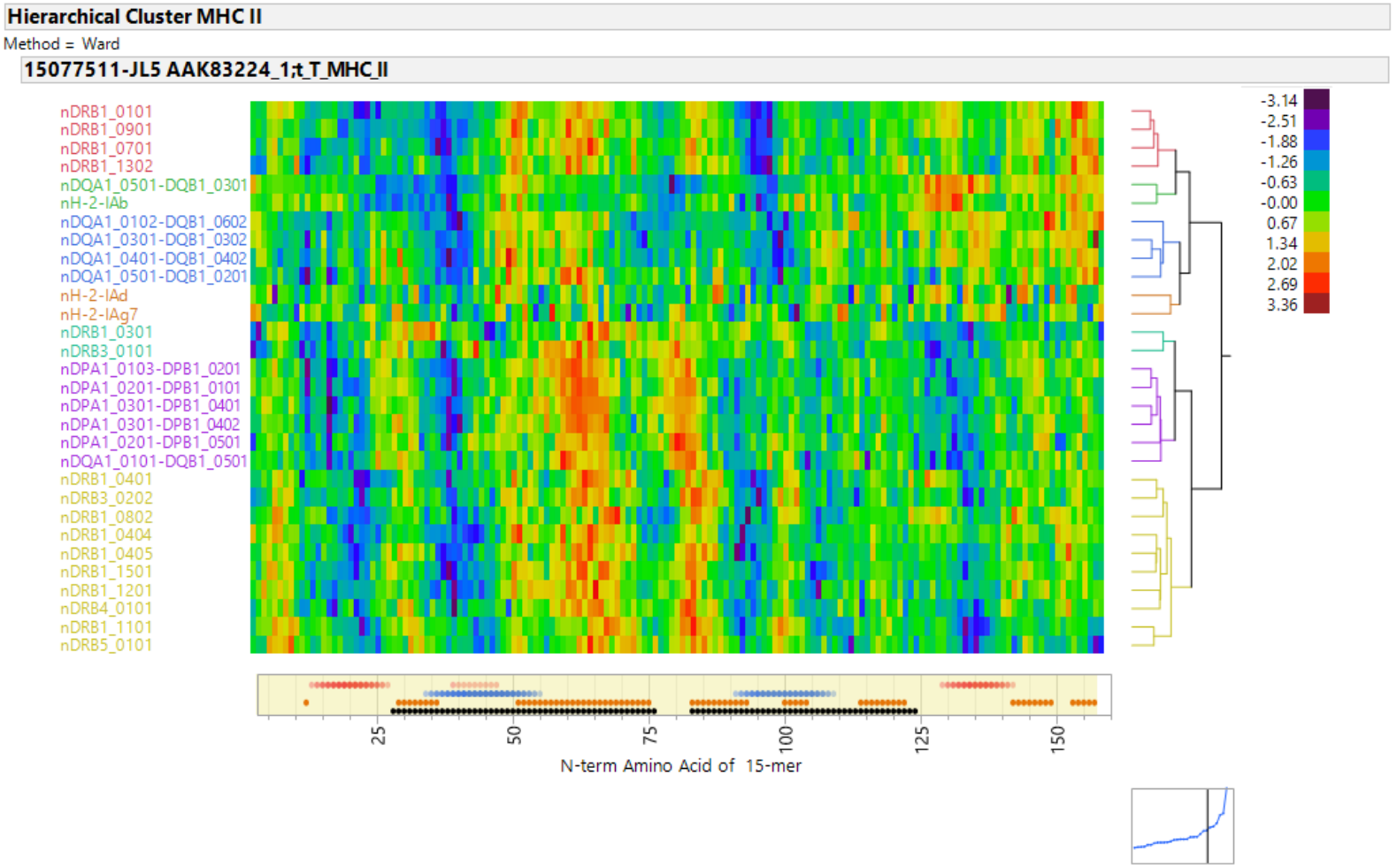
I protein

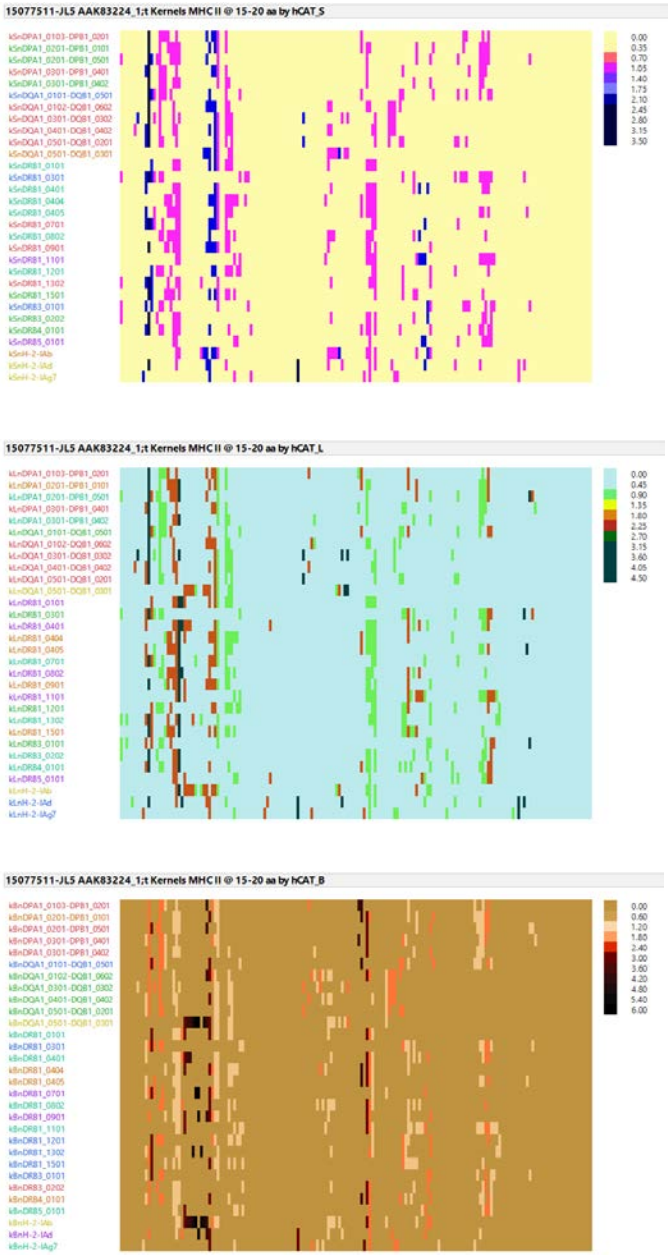
AAK83224.1| I protein

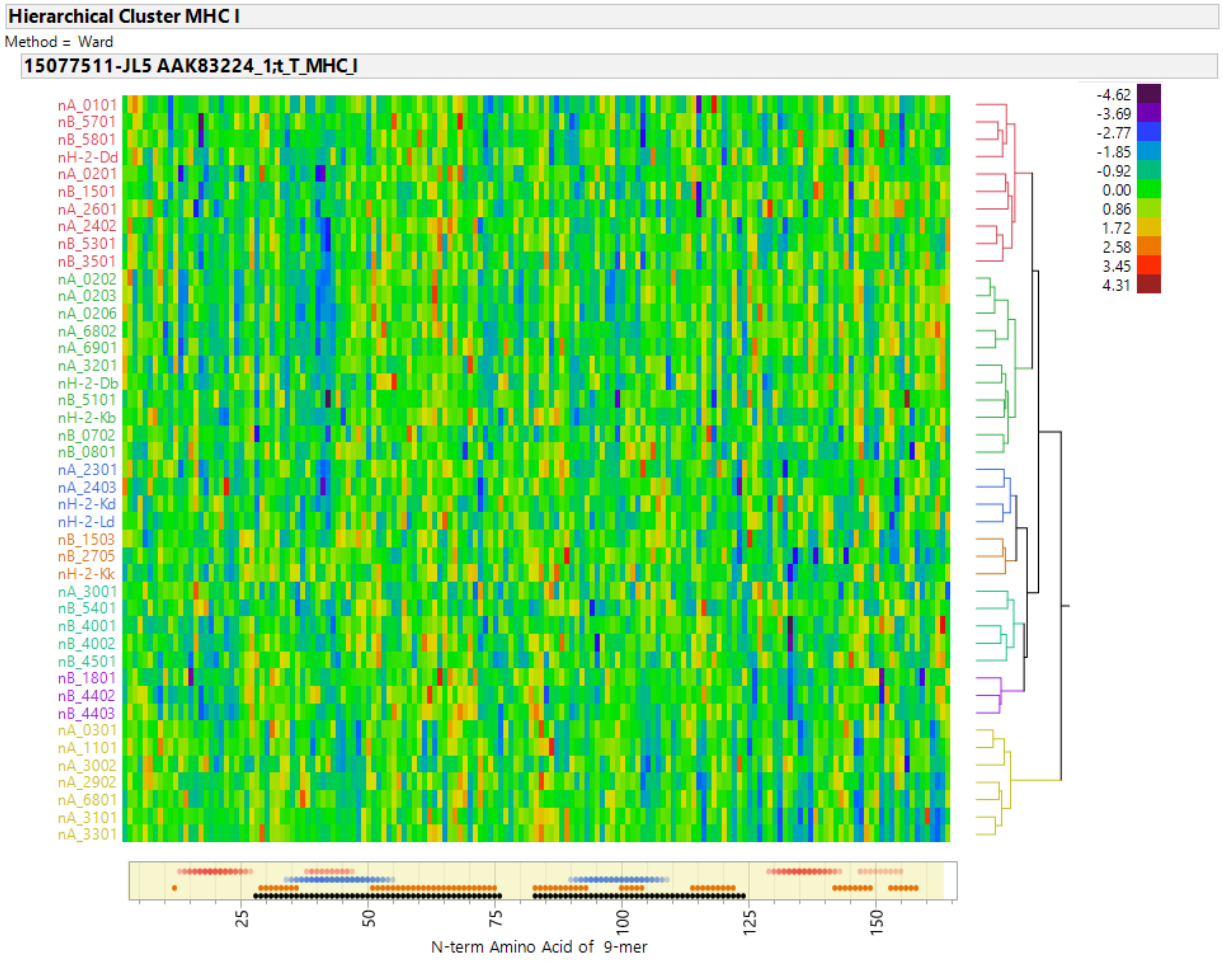
Overlay Plot

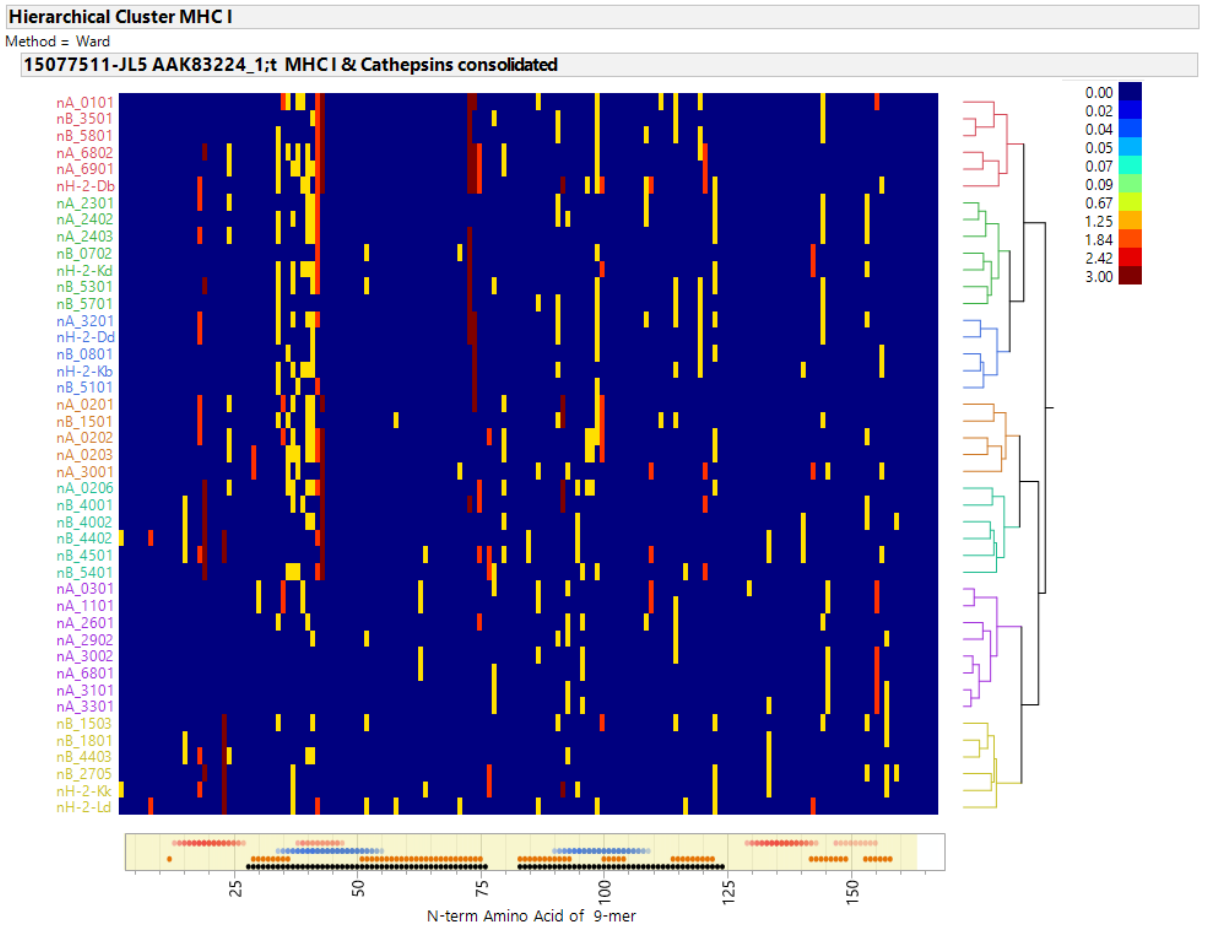
15077511-JL5 AAK83224_1;t

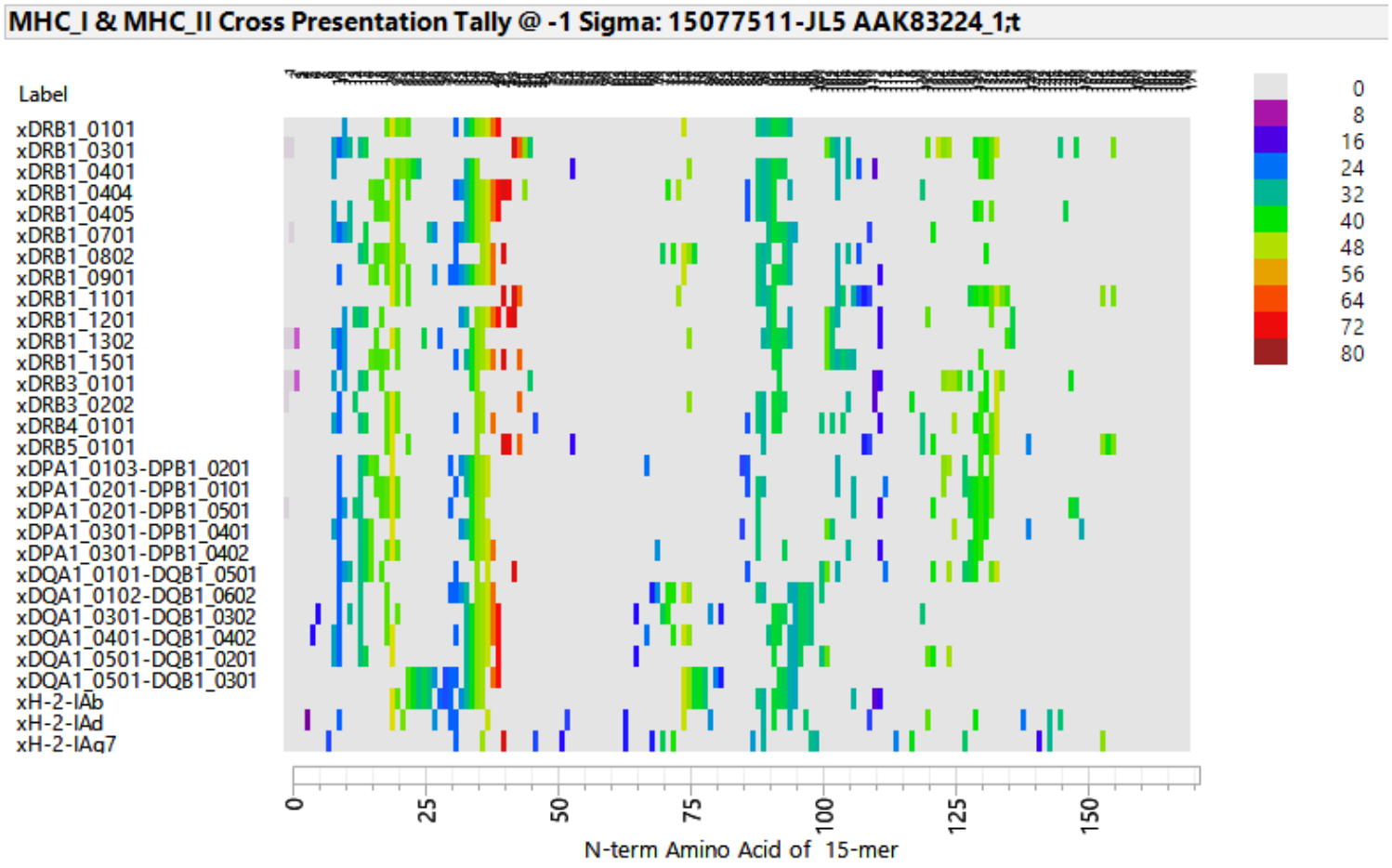



Hierarchical Cluster MHC II
Method = Ward
15077511-JL5 AAK83224_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDQA1_0501-DQB1_0301
nH-2-IAb
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB1_1201
nDRB4_0101
nDRB1_1101
nDRB5_0101
-3.14
-2.51
-1.88
-1.26
-0.63
-0.00
0.67
1.34
2.02
2.69
3.36
25
50
75
100
125
150
N-term Amino Acid of 15-mer

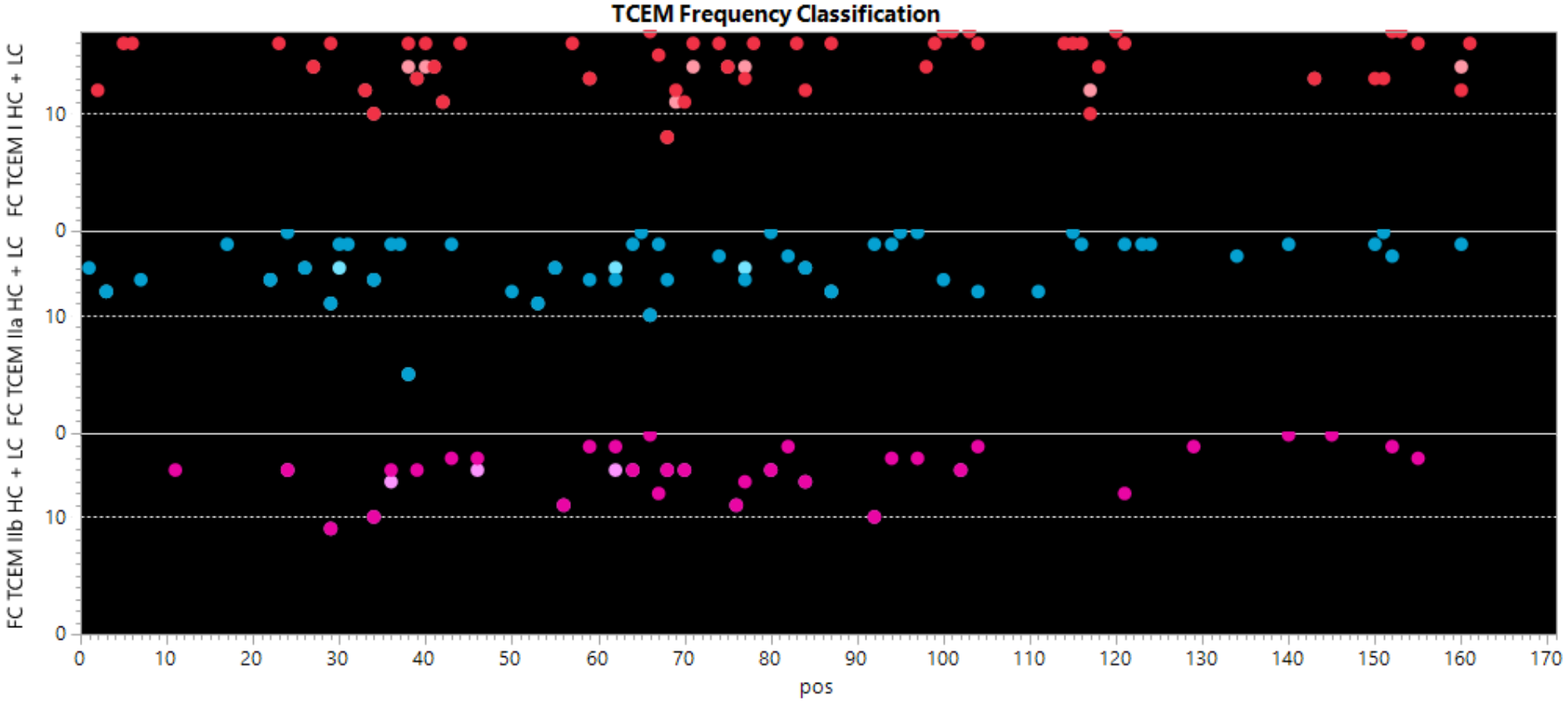

15077511-JL5 AAK83224_1;t Kernels MHC II @ 15-20 aa by hCAT_S
15077511-JL5 AAK83224_1;t Kernels MHC II @ 15-20 aa by hCAT_L
15077511-JL5 AAK83224_1;t Kernels MHC II @ 15-20 aa by hCAT_B

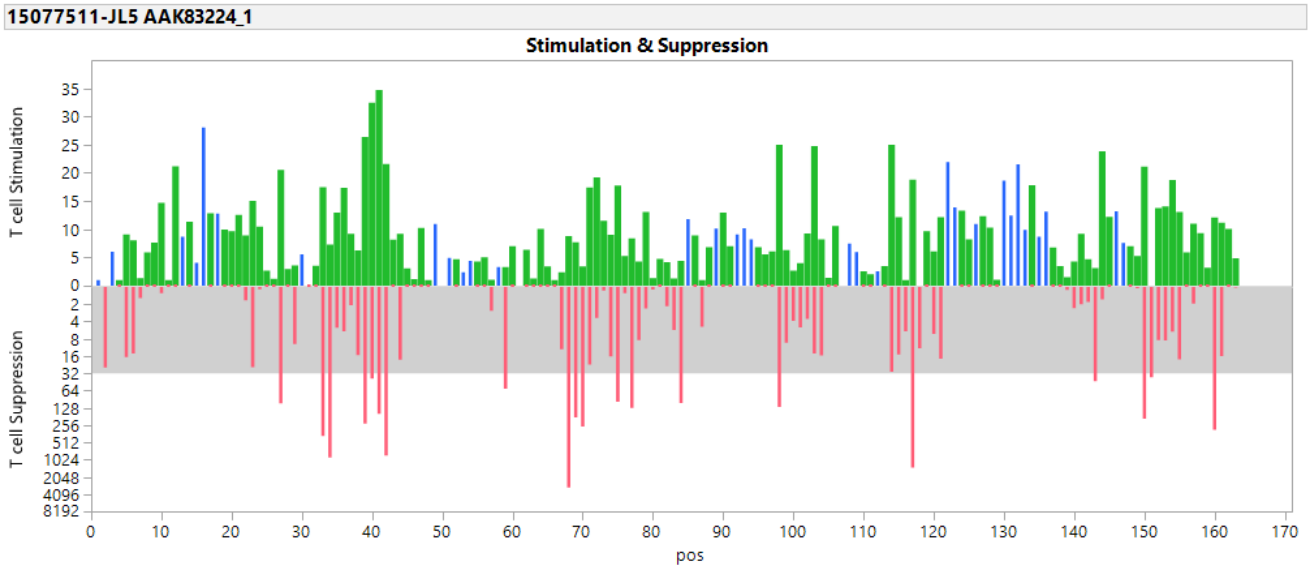

Hierarchical Cluster MHC I
Method = Ward
15077511-JL5 AAK83224_1;t_T_MHC_I
nA_0101
nB_5701
nB_5801
nH-2-Dd
nA_0201
nB_1501
nA_2601
nA_2402
nB_5301
nB_3501
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_3201
nH-2-Db
nB_5101
nH-2-Kb
nB_0702
nB_0801
nA_2301
nA_2403
nH-2-Kd
nH-2-Ld
nB_1503
nB_2705
nH-2-Kk
nA_3001
nB_5401
nB_4001
nB_4002
nB_4501
nB_1801
nB_4402
nB_4403
nA_0301
nA_1101
nA_3002
nA_2902
nA_6801
nA_3101
nA_3301
-4.62
-3.69
-2.77
-1.85
-0.92
0.00
0.86
1.72
2.58
3.45
4.31
25
50
75
100
125
150
N-term Amino Acid of 9-mer

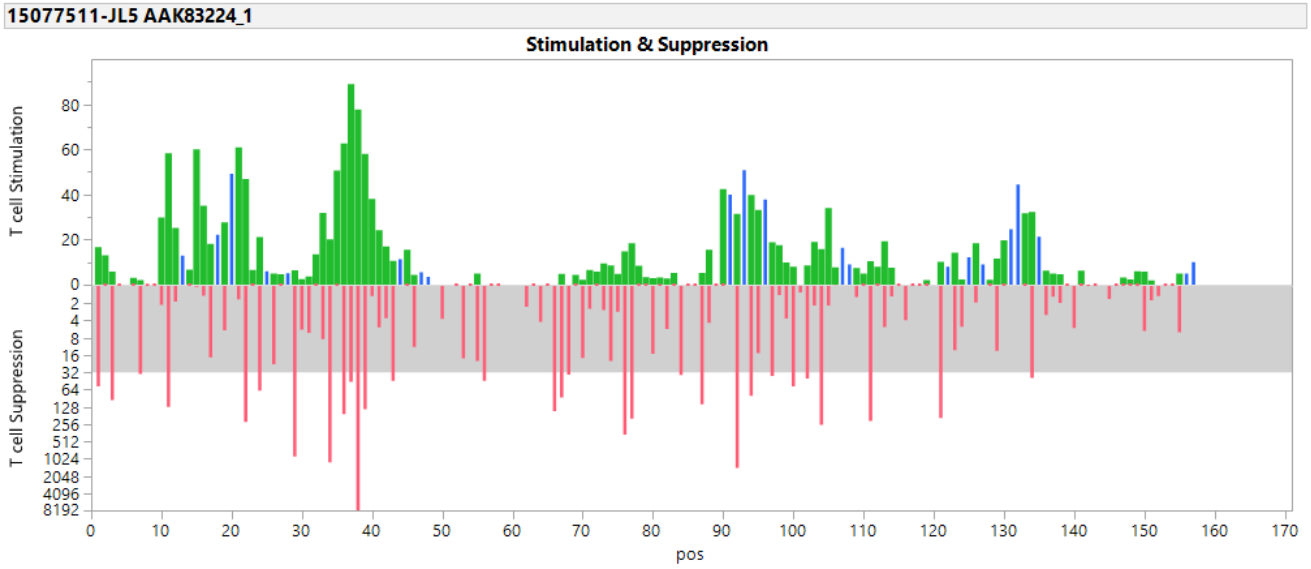

Hierarchical Cluster MHC I
Method = Ward
15077511-JL5 AAK83224_1;t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nB_5801
nA_6802
nA_6901
nH-2-Db
nA_2301
nA_2402
nA_2403
nB_0702
nH-2-Kd
nB_5301
nB_5701
nA_3201
nH-2-Dd
nB_0801
nH-2-Kb
nB_5101
nA_0201
nB_1501
nA_0202
nA_0203
nA_3001
nA_0206
nB_4001
nB_4002
nB_4402
nB_4501
nB_5401
nA_0301
nA_1101
nA_2601
nA_2902
nA_3002
nA_6801
nA_3101
nA_3301
nB_1503
nB_1801
nB_4403
nB_2705
nH-2-Kk
nH-2-Ld
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00
25
50
75
100
125
150
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077511-JL5 AAK83224_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
8
16
24
32
40
48
56
64
72
80
0
25
50
75
100
125
150
N-term Amino Acid of 15-mer

15077511-JL5 AAK83224_1

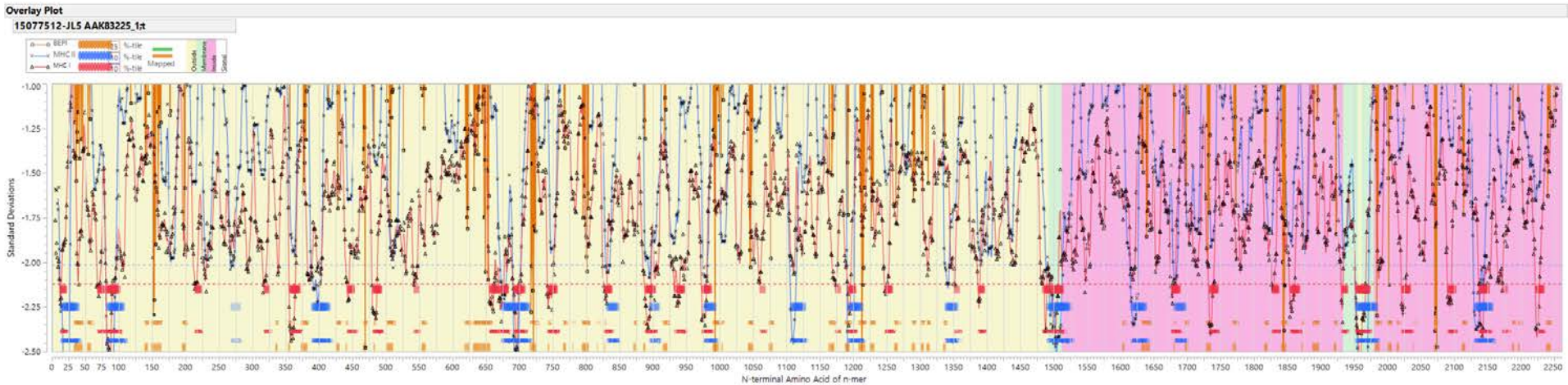


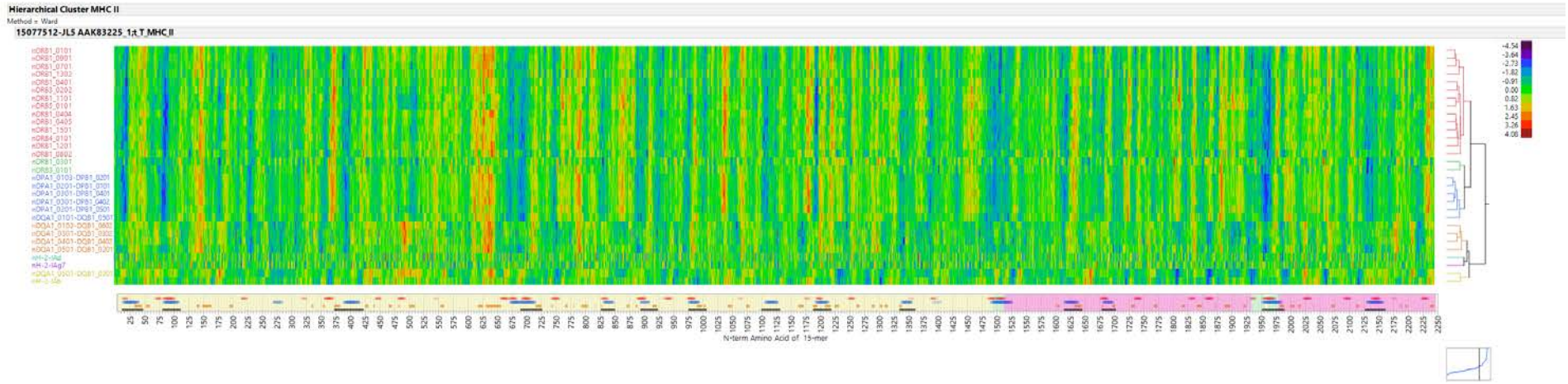

15077511-JL5 AAK83224_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

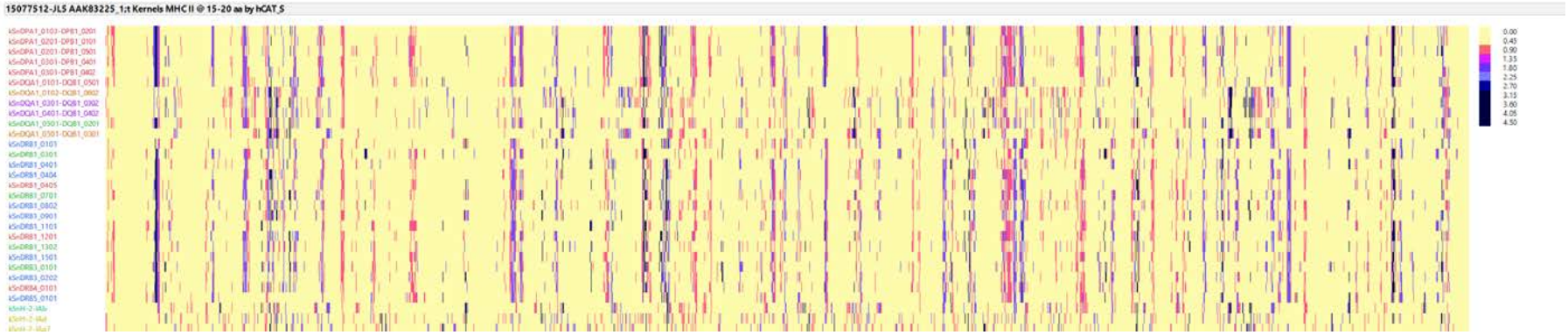

15077511-JL5 AAK83224_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

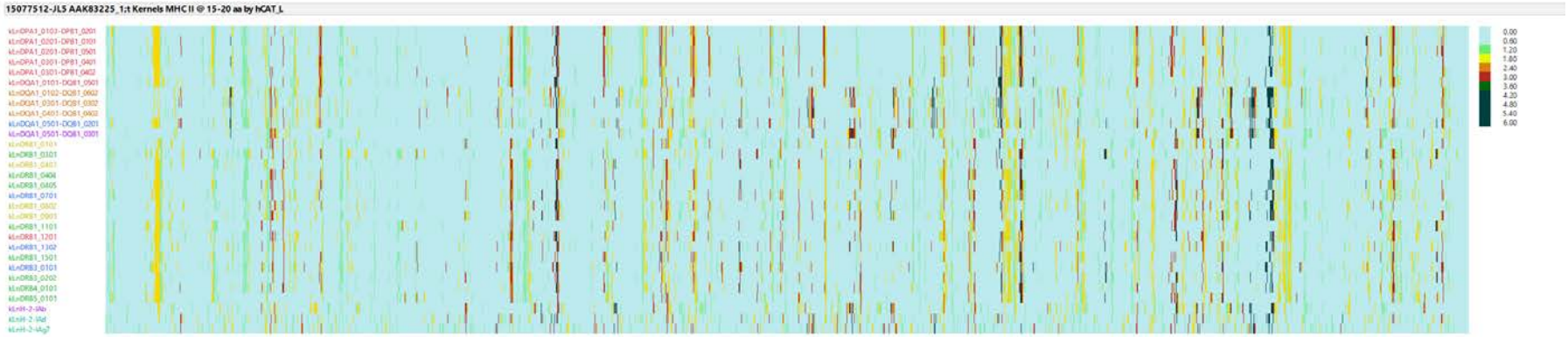
Large protein

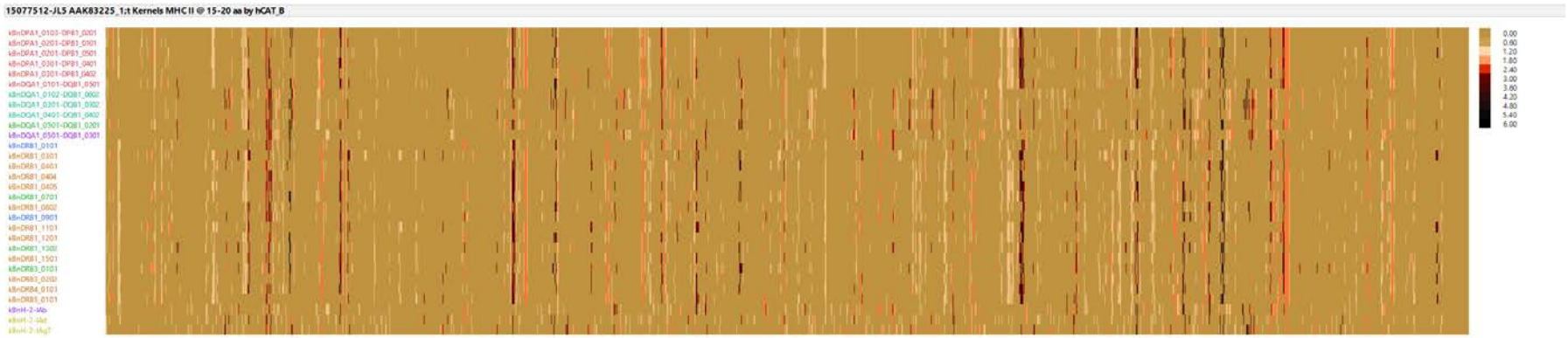
AAK83225.1| large protein

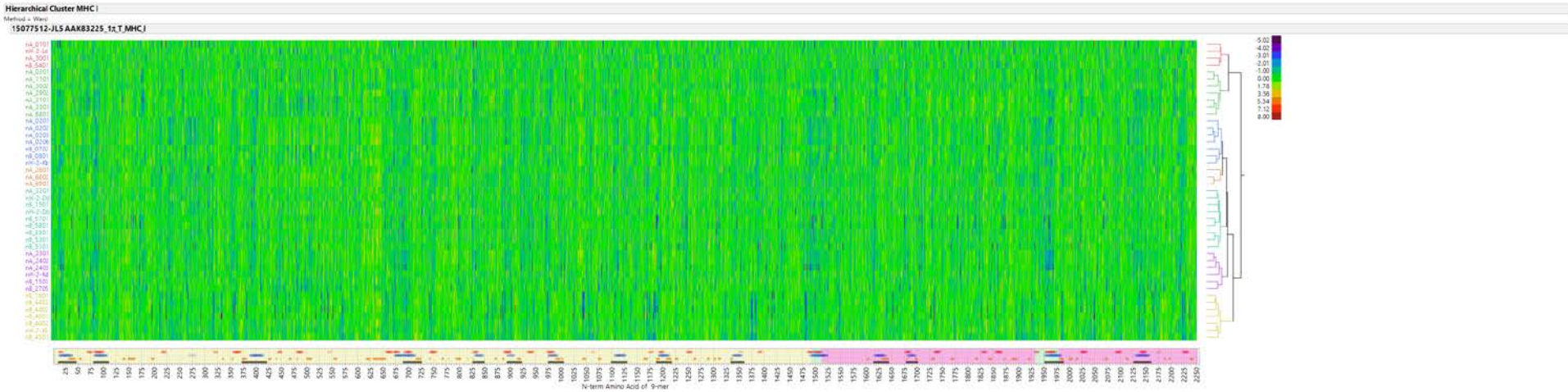

Overlay Plot
15077512-JL5 AAK83225_1:t
BEPI
MHC II
MHC I
%-tile
%-tile
%-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0 25 50 75 100 150 200 250 300 350 400 450 500 550 600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250 1300 1350 1400 1450 1500 1550 1600 1650 1700 1750 1800 1850 1900 1950 2000 2050 2100 2150 2200 2250
N-terminal Amino Acid of n-mer

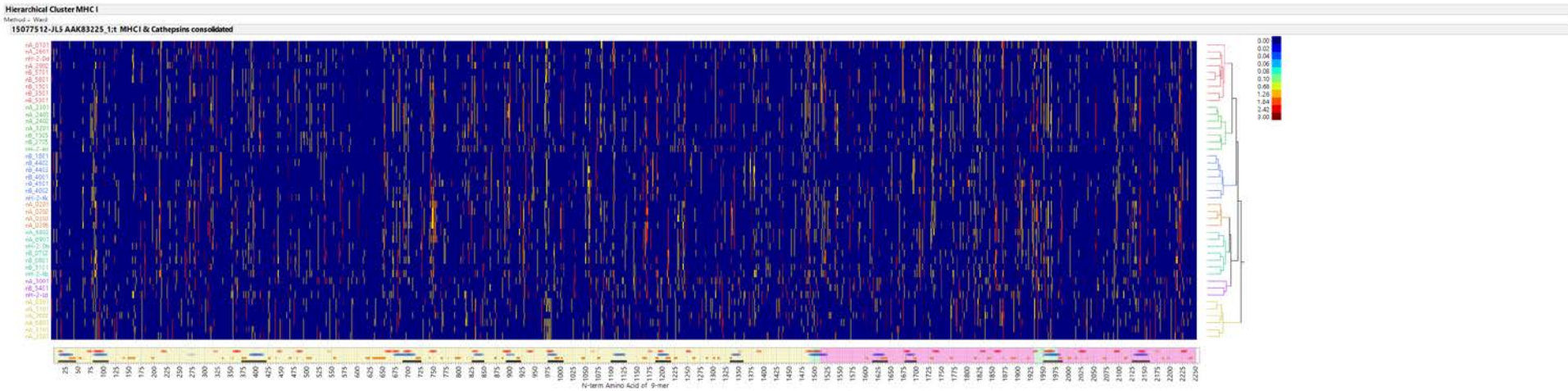

Hierarchical Cluster MHC II
Method = Ward
15077512-JL5 AAK83225_1;t T_MHC II
N-term Amino Acid of 15-mer

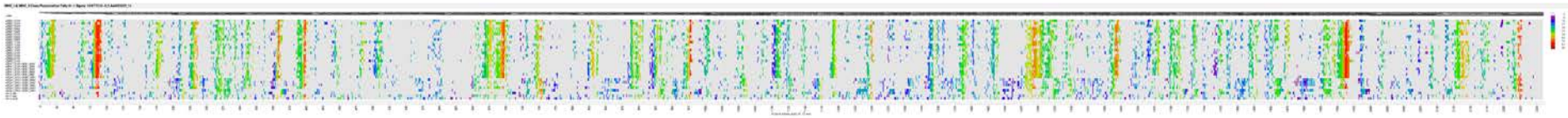

15077512-JL5 AAK83225_1:t Kernels MHC II @ 15-20 aa by hCAT_S


15077512-JL5 AAK83225_1:t Kernels MHC II @ 15-20 aa by hCAT_L

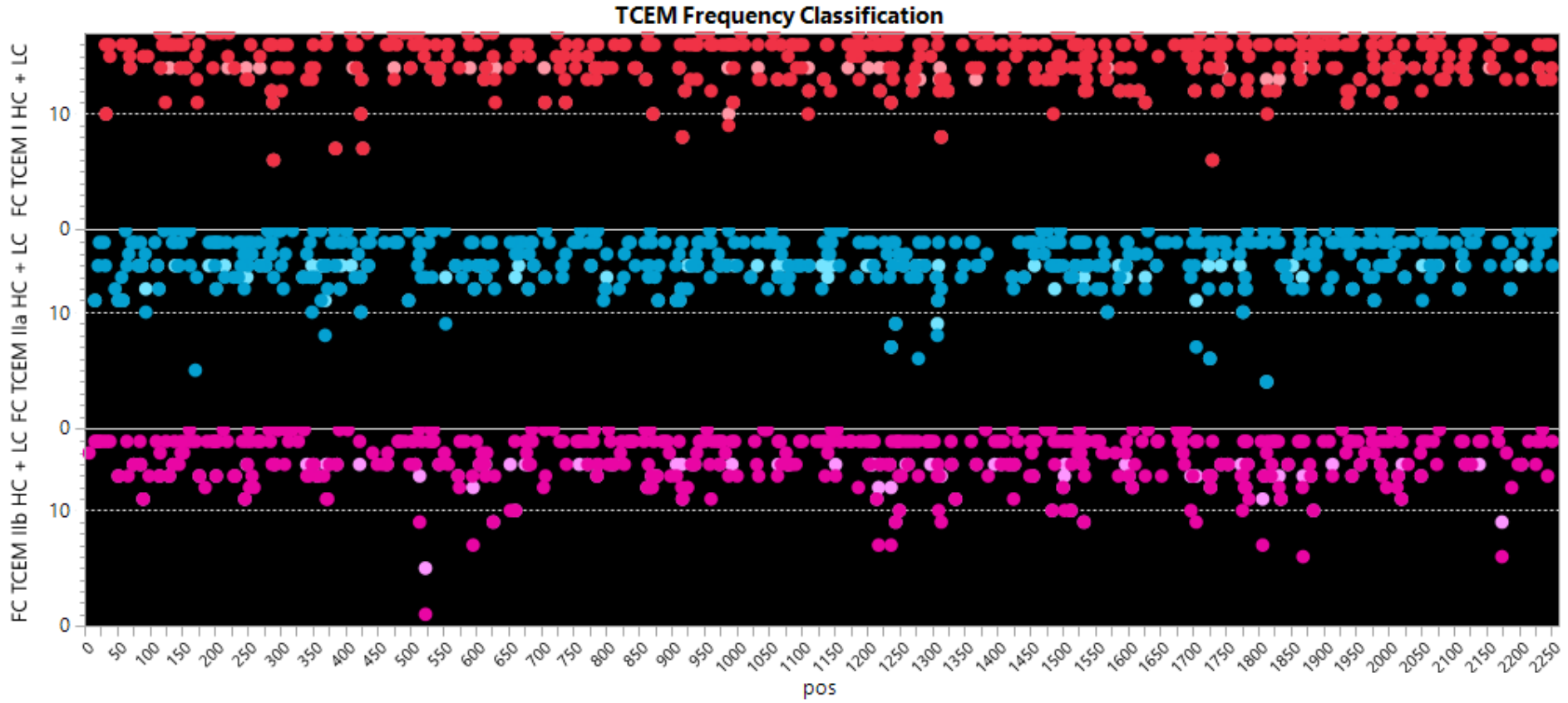

15077512-JL5 AAK83225_1:t Kernels MHC II @ 15-20 aa by hCAT_B

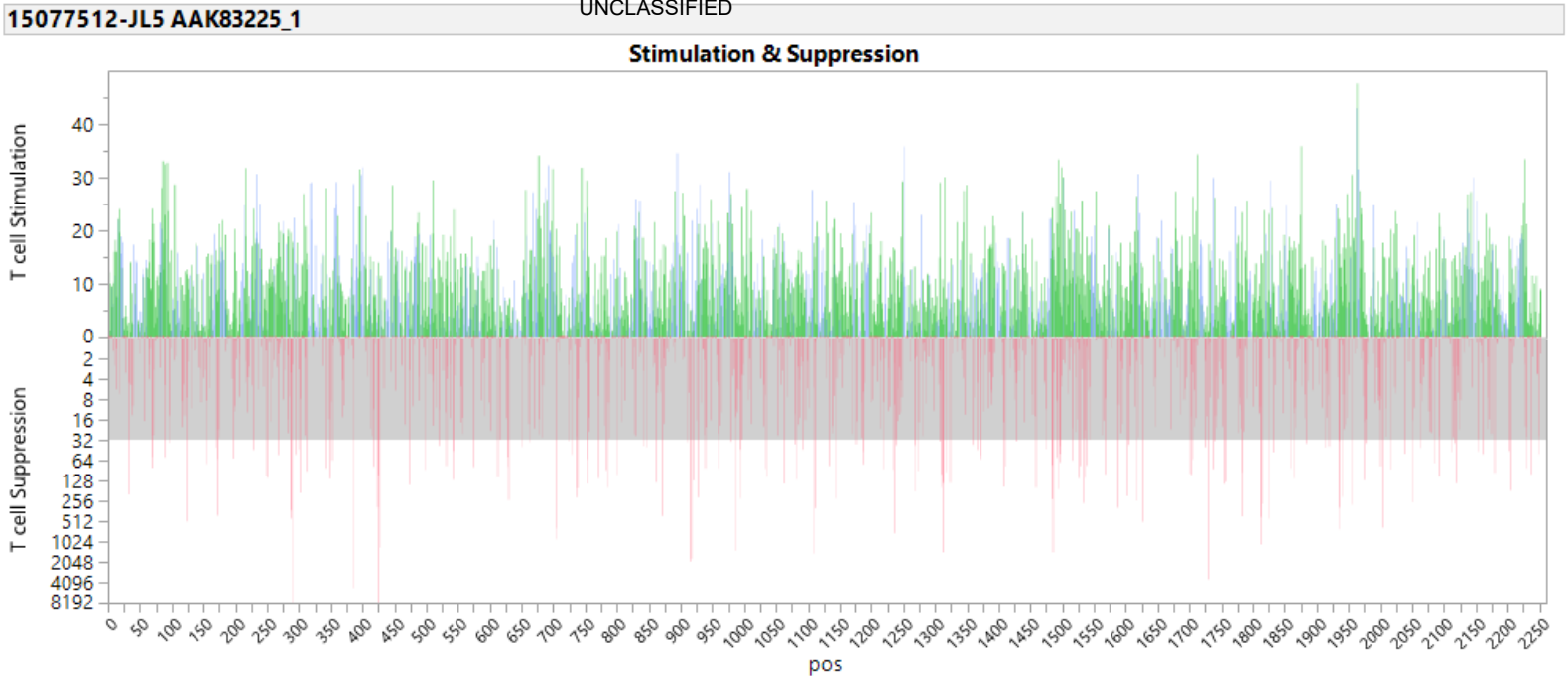

Hierarchical Cluster MHC I
15077512-JL5 AAK83225_1st_T_MHC_I
N-term Amino Acid of 9-mer

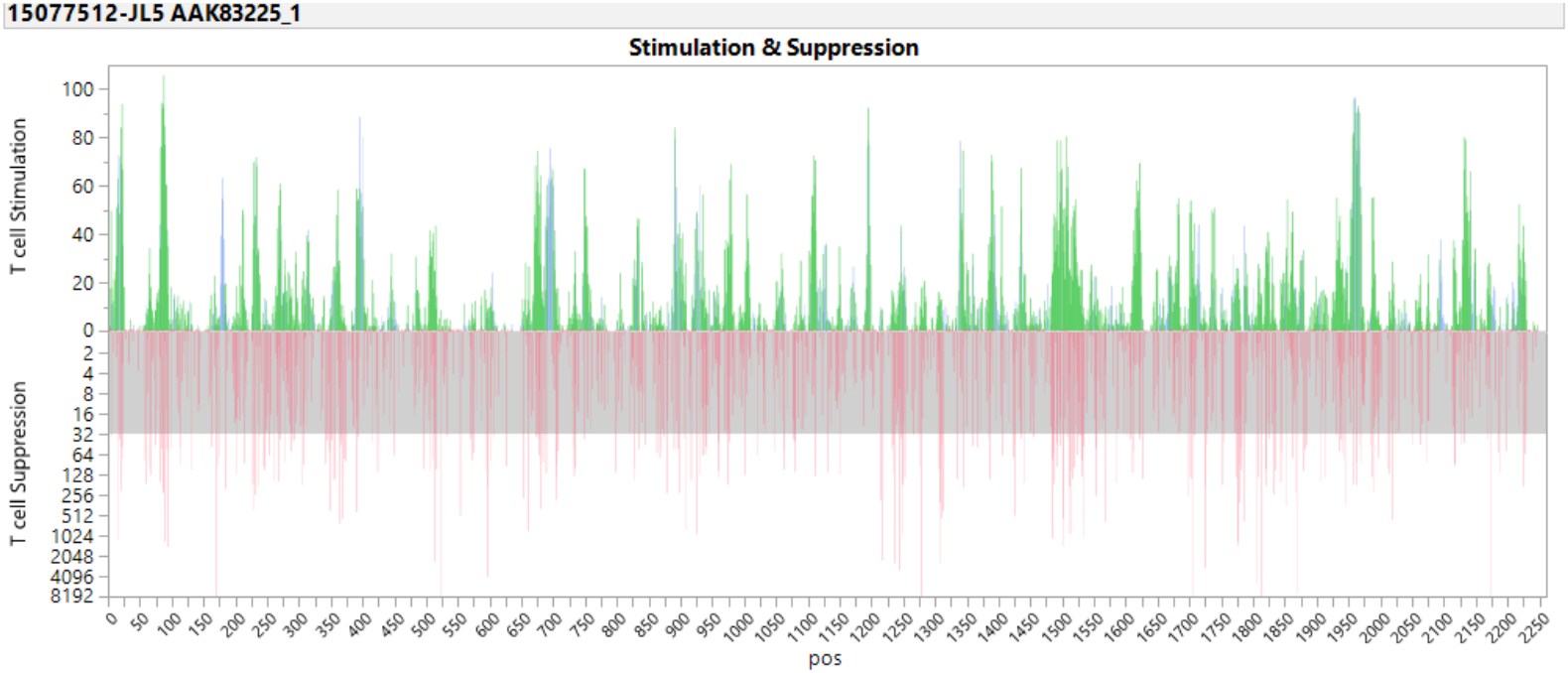

Hierarchical Cluster MHC I
Method - Ward
15077512-JL5 AAK83225_1:t MHC I & Cathepsins consolidated
N-term Amino Acid of 9-mer



15077512-JL5 AAK83225_1







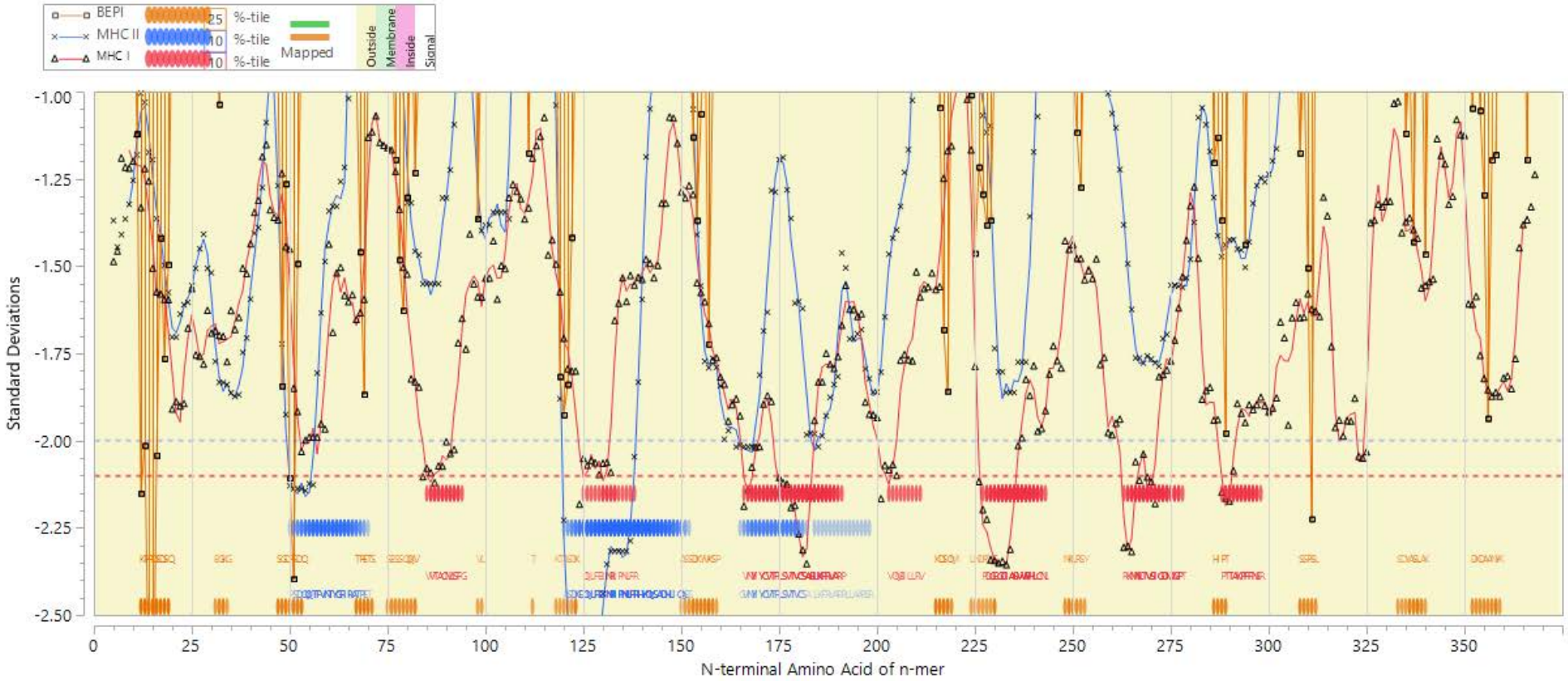
Pos 523
Supp II
303305

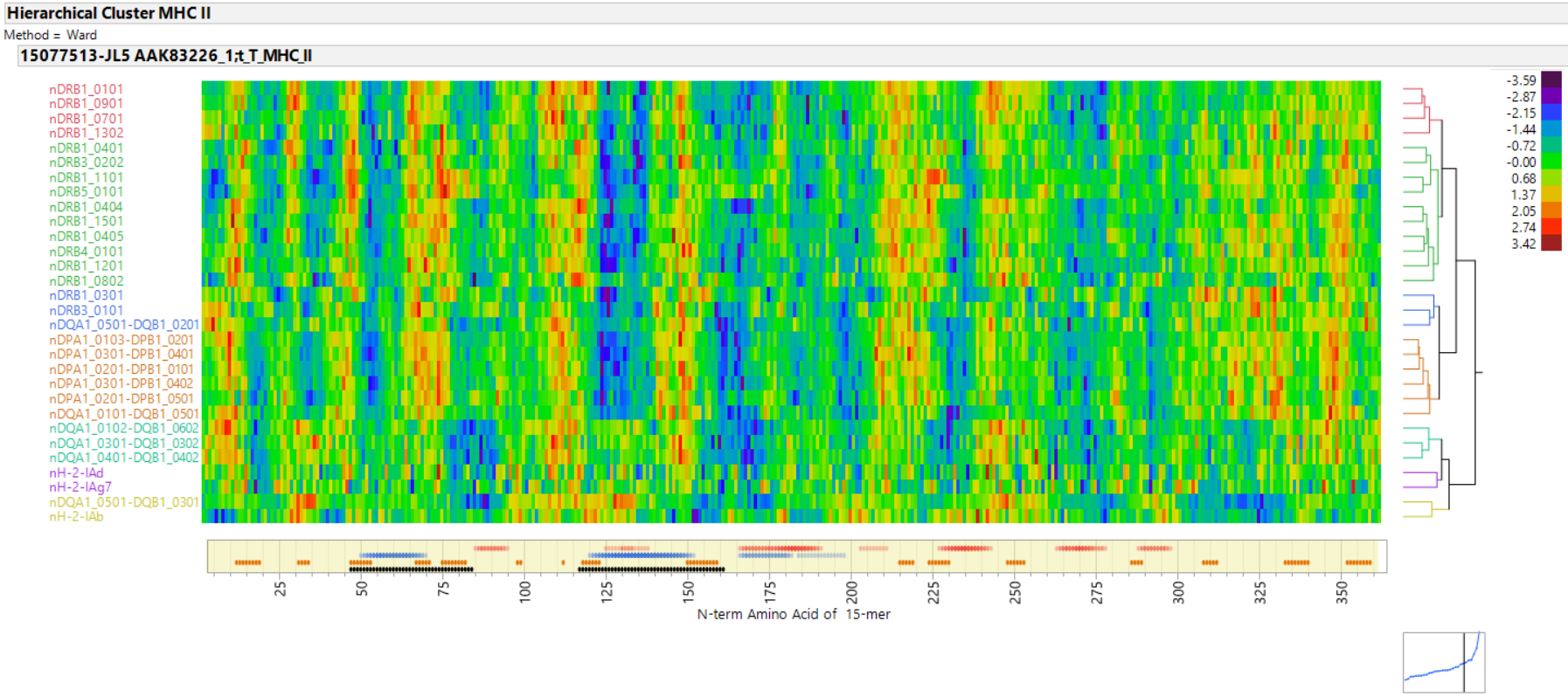
Matrix protein

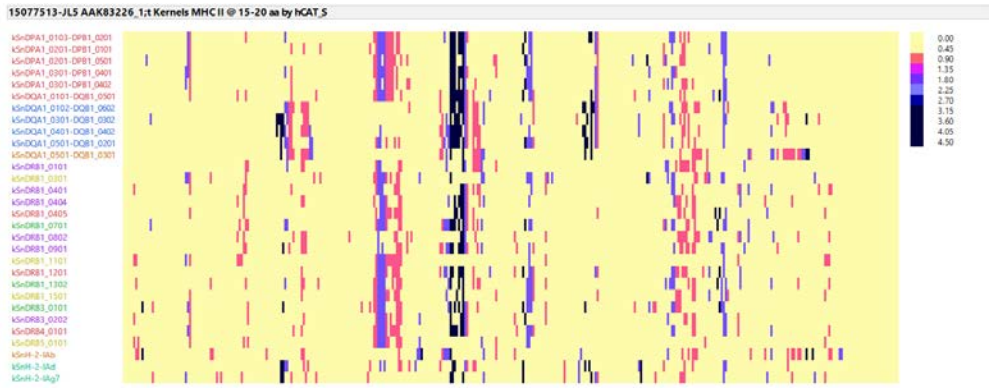
AAK83226.1| matrix protein

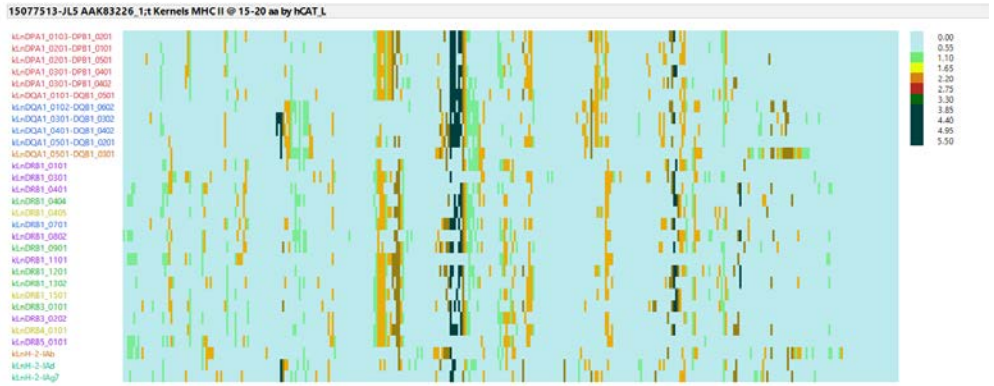
Overlay Plot

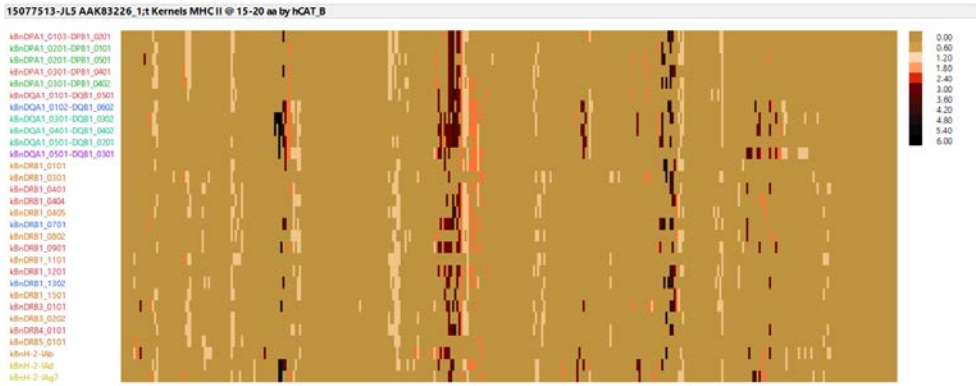
15077513-JL5 AAK83226_1;t

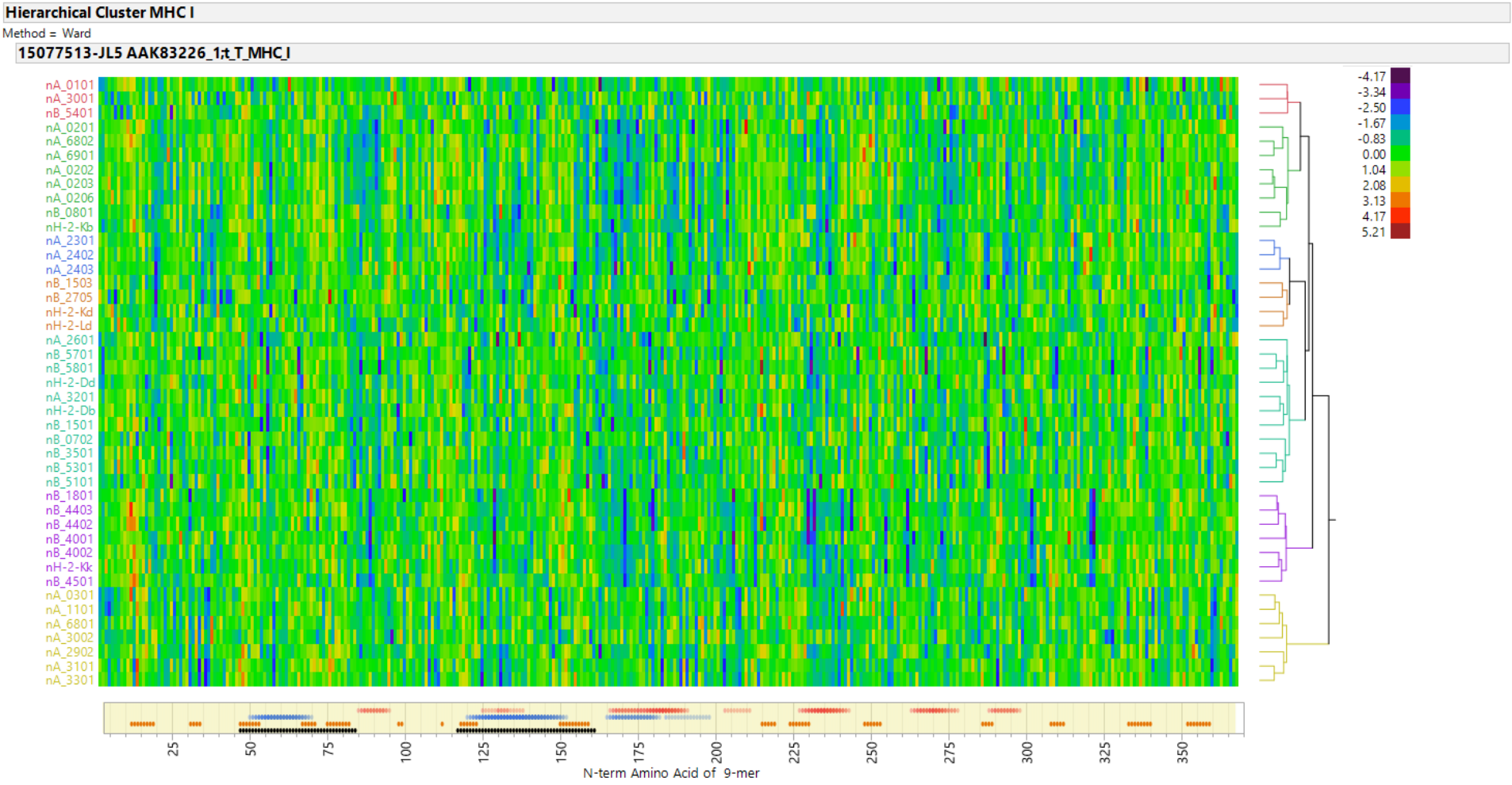


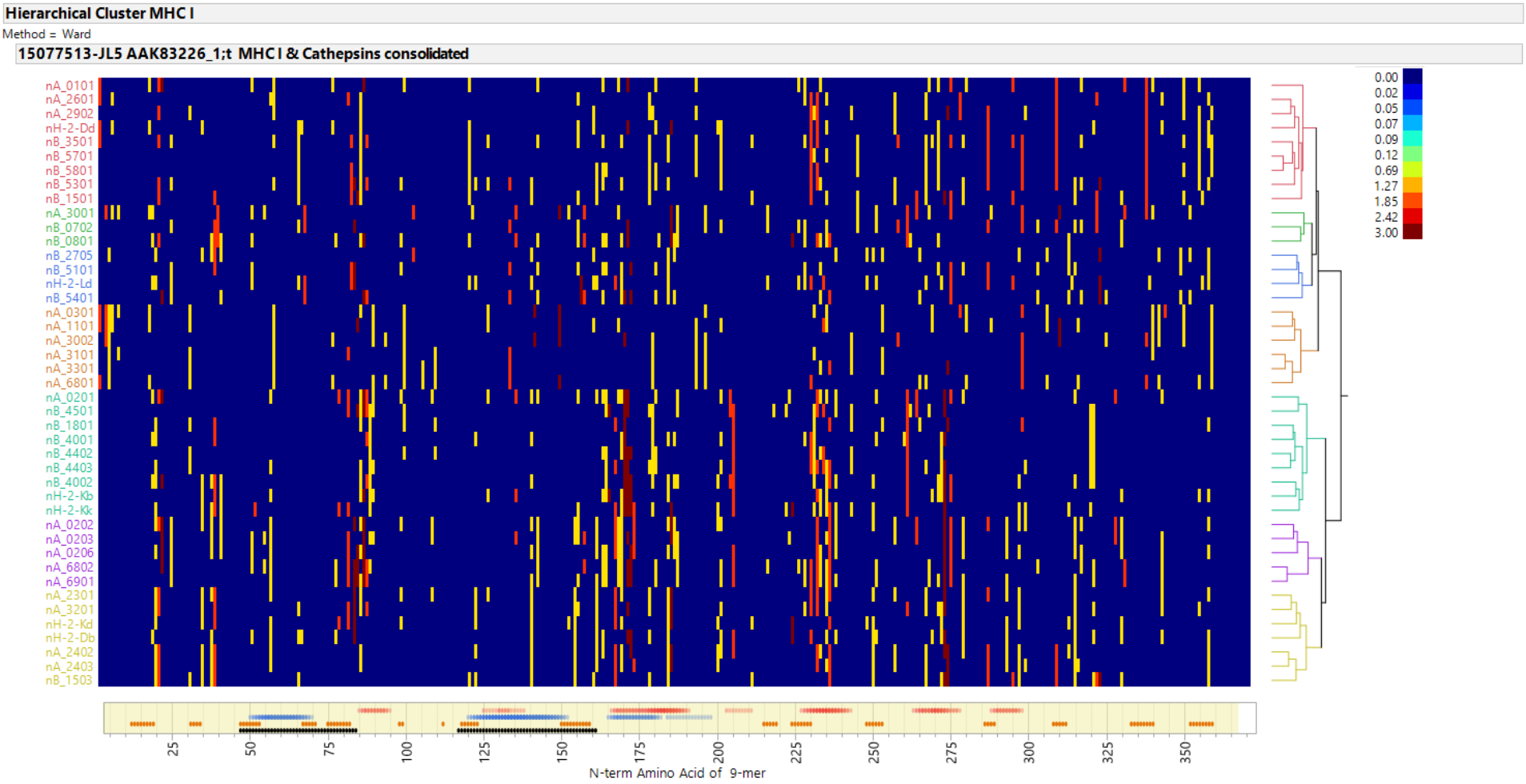

Hierarchical Cluster MHC II
Method = Ward
15077513-JL5 AAK83226_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDRB1_1201
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.59
-2.87
-2.15
-1.44
-0.72
-0.00
0.68
1.37
2.05
2.74
3.42
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 15-mer

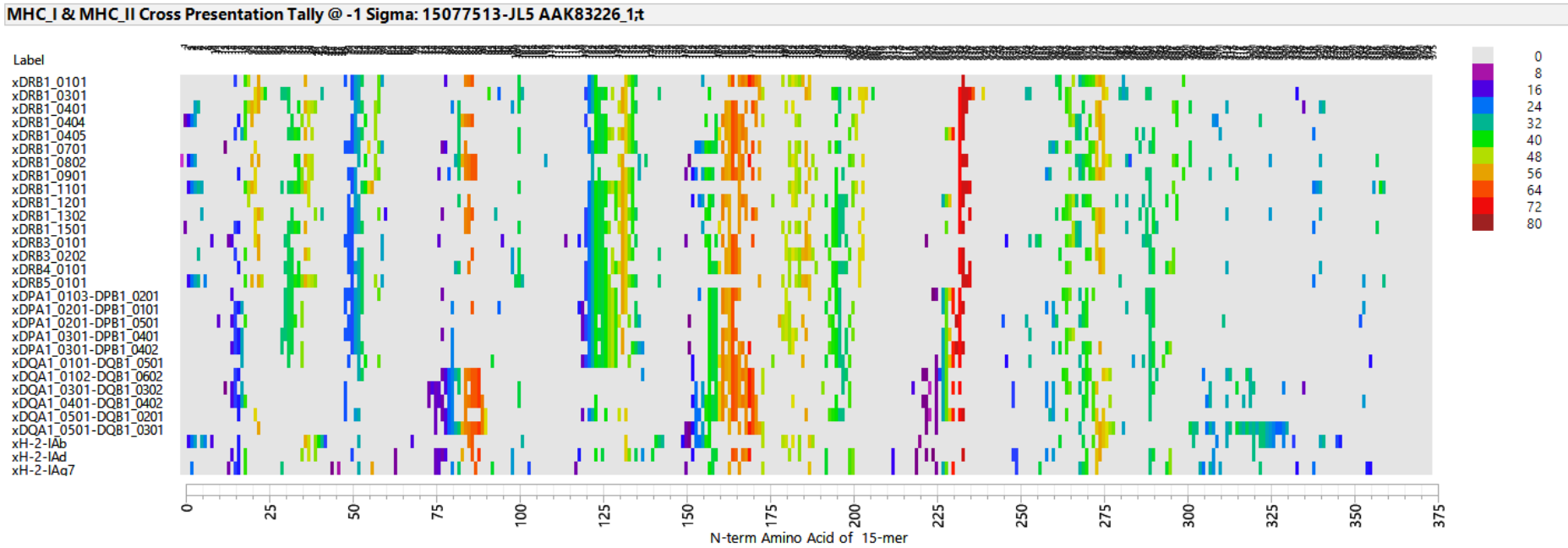

15077513-JL5 AAK83226_1;t Kernels MHC II @ 15-20 aa by hCAT_S


15077513-JL5 AAK83226_1;t Kernels MHC II @ 15-20 aa by hCAT_L

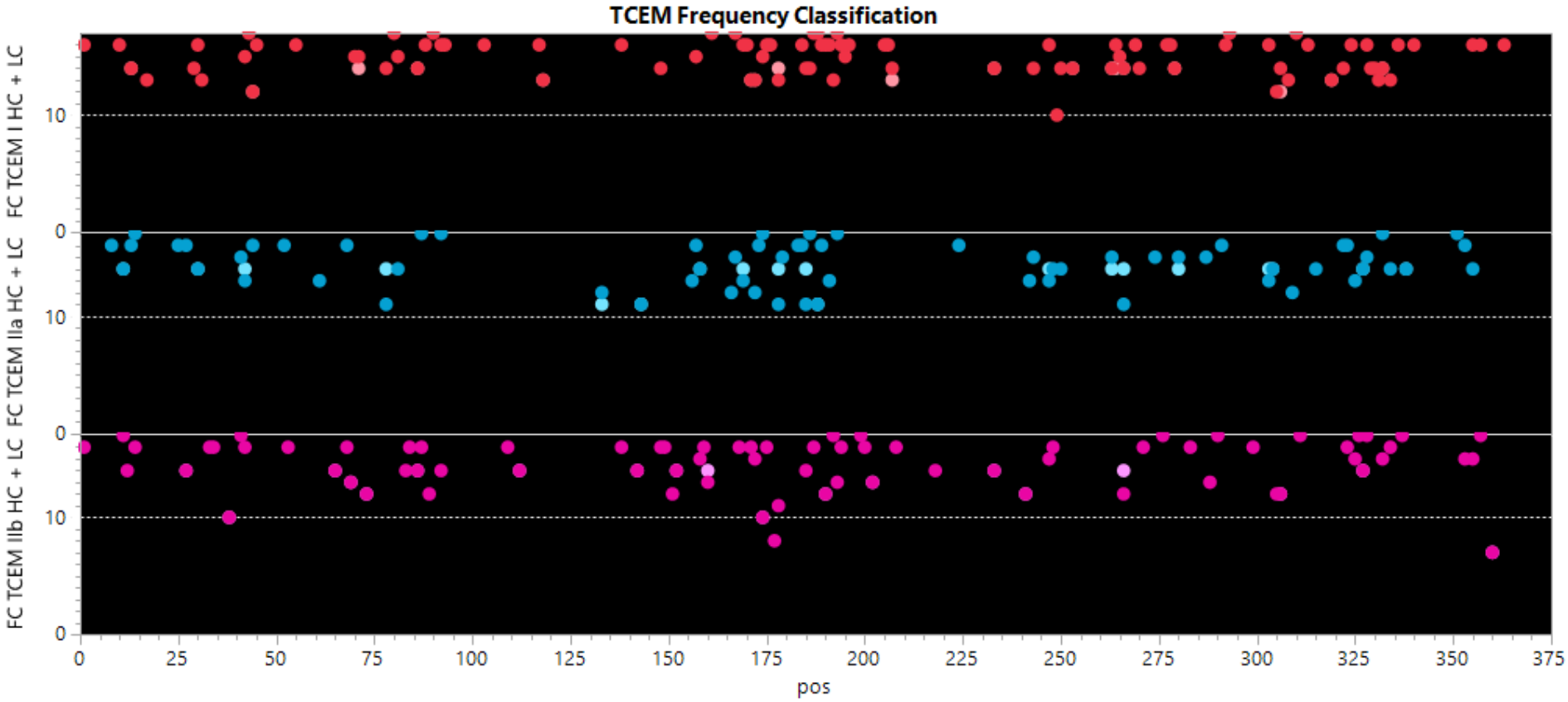

15077513-JL5 AAK83226_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
15077513-JL5 AAK83226_1;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_2601
nB_5701
nB_5801
nH-2-Dd
nA_3201
nH-2-Db
nB_1501
nB_0702
nB_3501
nB_5301
nB_5101
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_6801
nA_3002
nA_2902
nA_3101
nA_3301
-4.17
-3.34
-2.50
-1.67
-0.83
0.00
1.04
2.08
3.13
4.17
5.21
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 9-mer

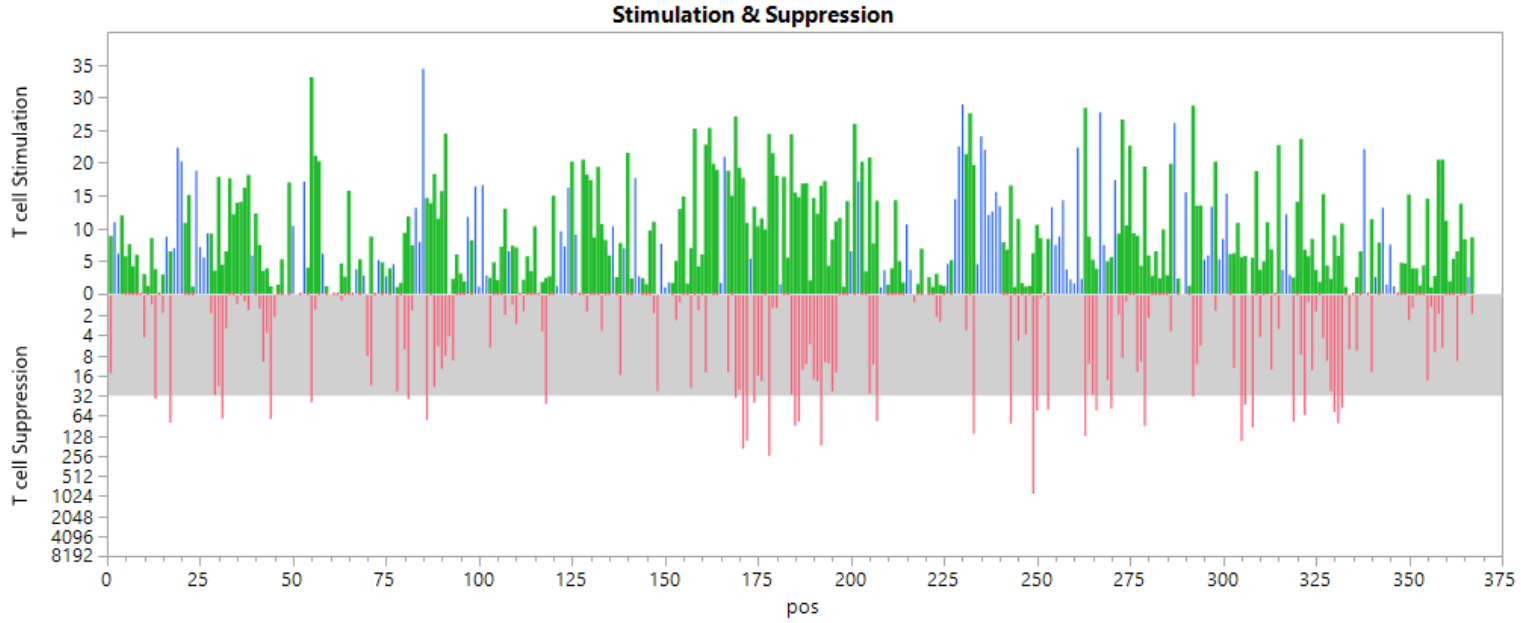

Hierarchical Cluster MHC I
Method = Ward
15077513-JL5 AAK83226_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nA_2902
nH-2-Dd
nB_3501
nB_5701
nB_5801
nB_5301
nB_1501
nA_3001
nB_0702
nB_0801
nB_2705
nB_5101
nH-2-Ld
nB_5401
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_0201
nB_4501
nB_1801
nB_4001
nB_4402
nB_4403
nB_4002
nH-2-Kb
nH-2-Kk
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_2301
nA_3201
nH-2-Kd
nH-2-Db
nA_2402
nA_2403
nB_1503
0.00
0.02
0.05
0.07
0.09
0.12
0.69
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077513-JL5 AAK83226_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

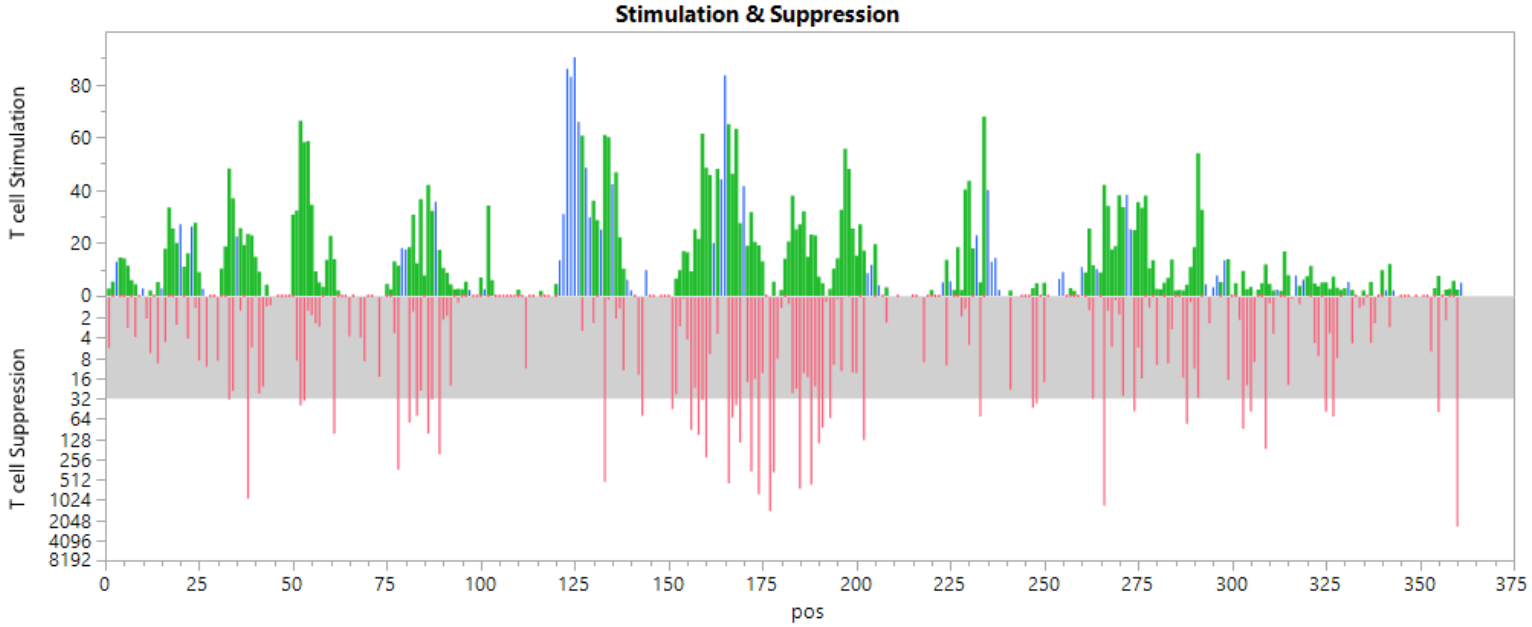
15077513-JL5 AAK83226_1



15077513-JL5 AAK83226_1

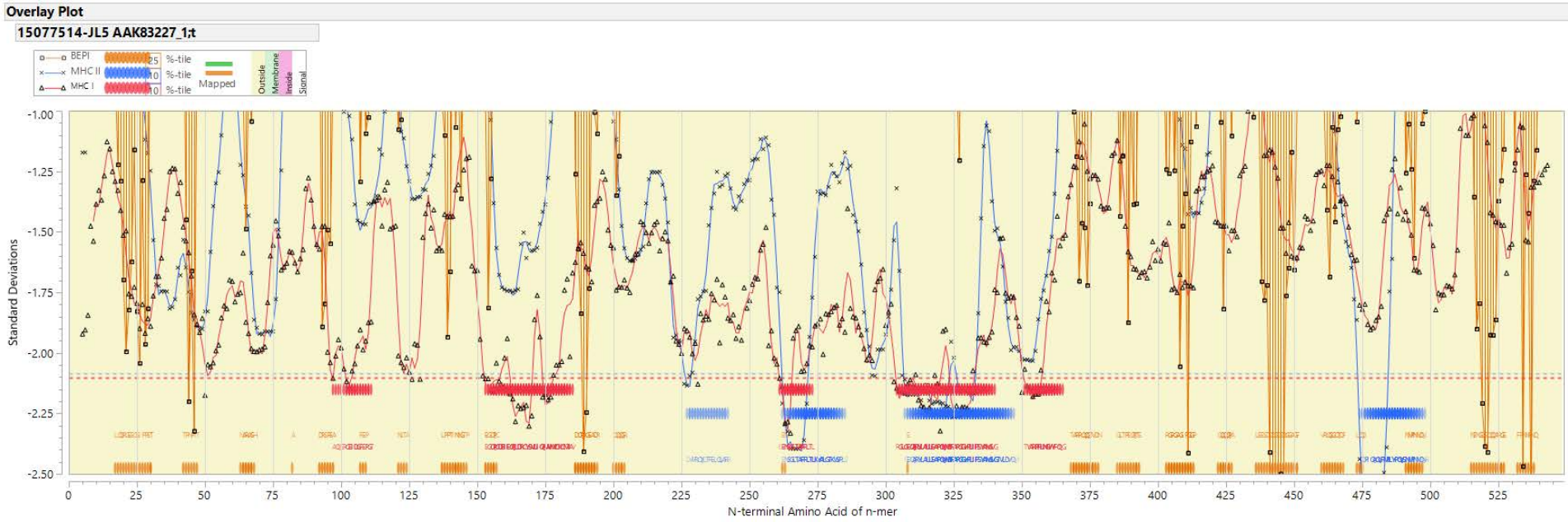


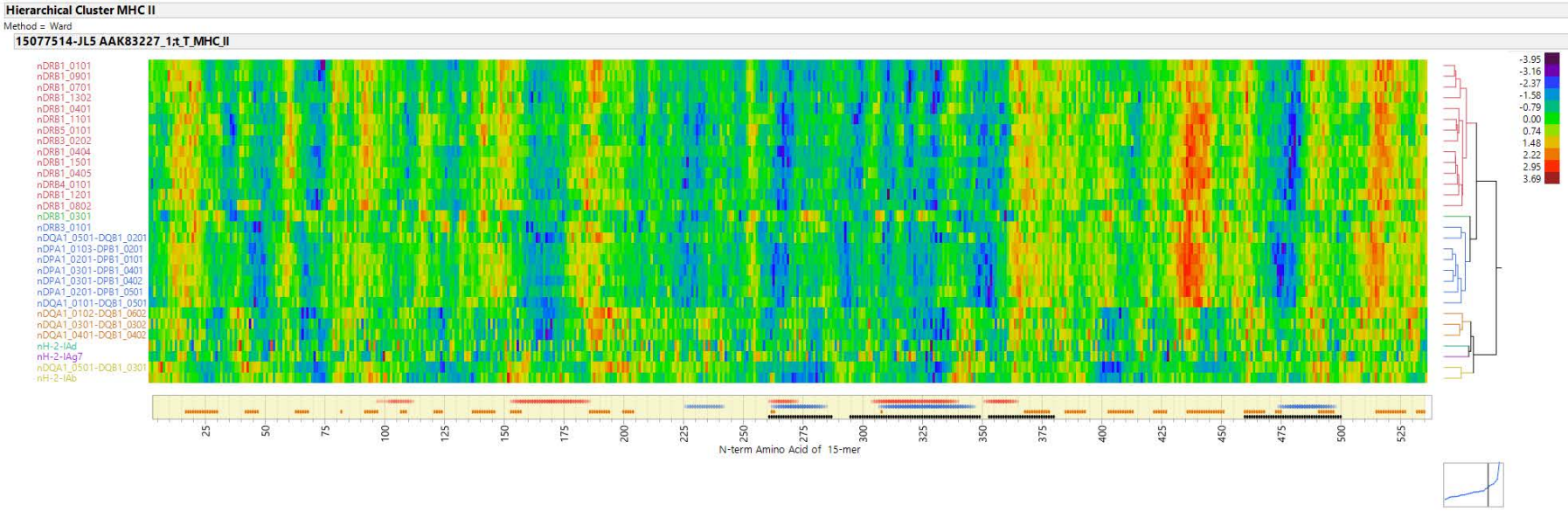
15077513-JL5 AAK83226_1

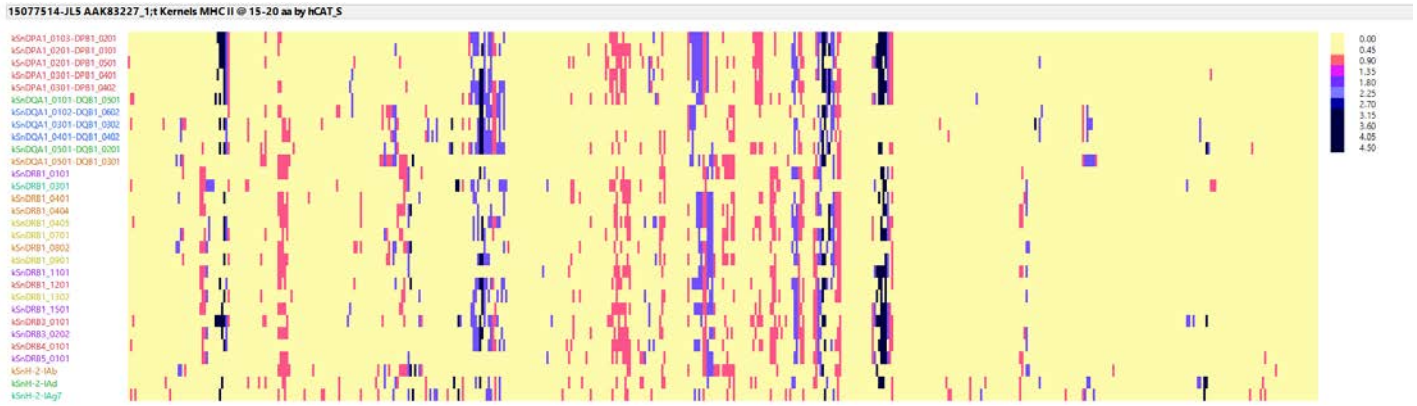


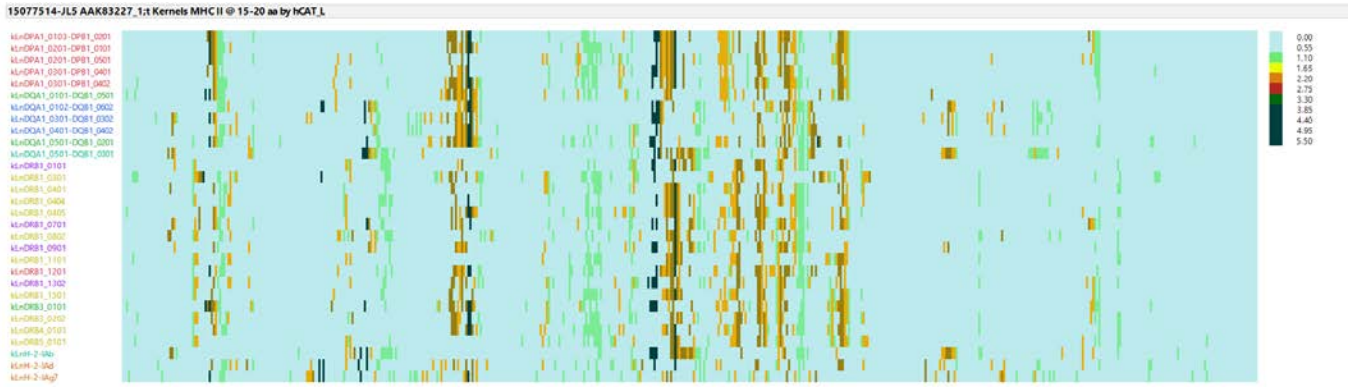
Nucelocapsid

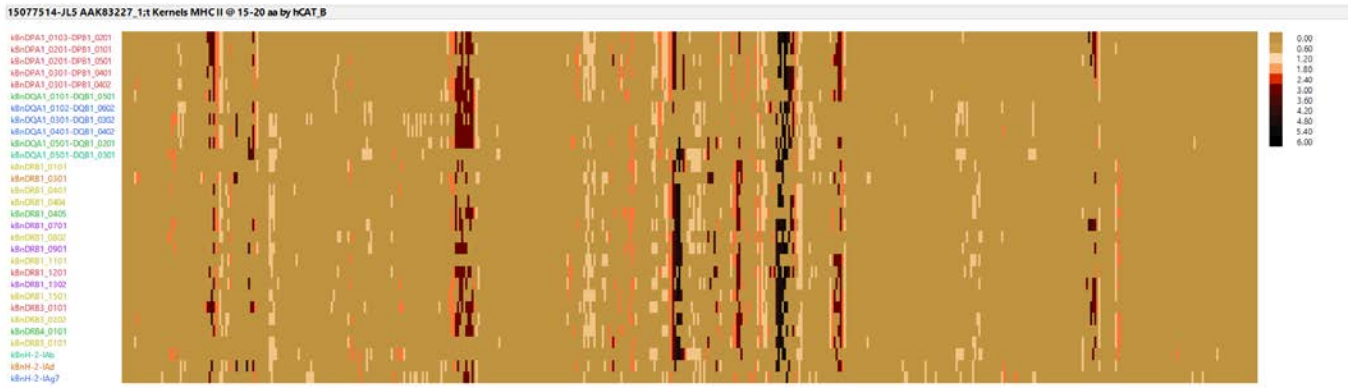
AAK83227.1| nucleocapsid protein

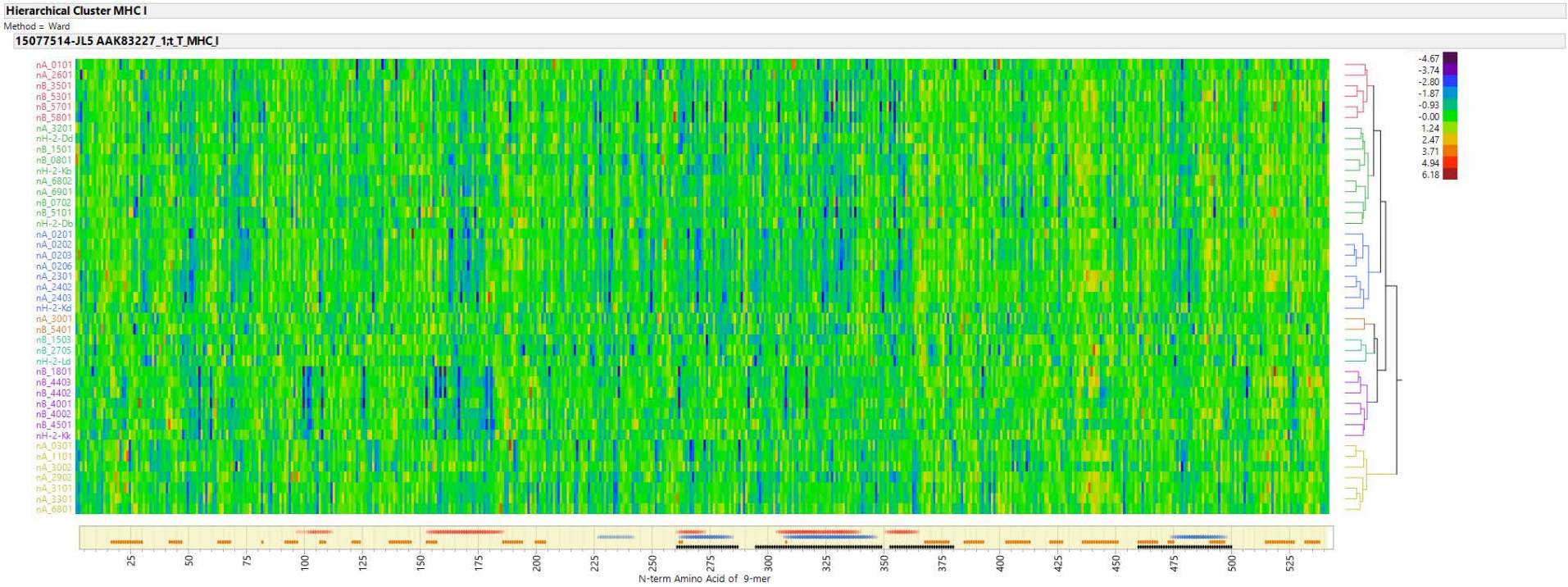

Overlay Plot
15077514-JL5 AAK83227_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-terminal Amino Acid of n-mer

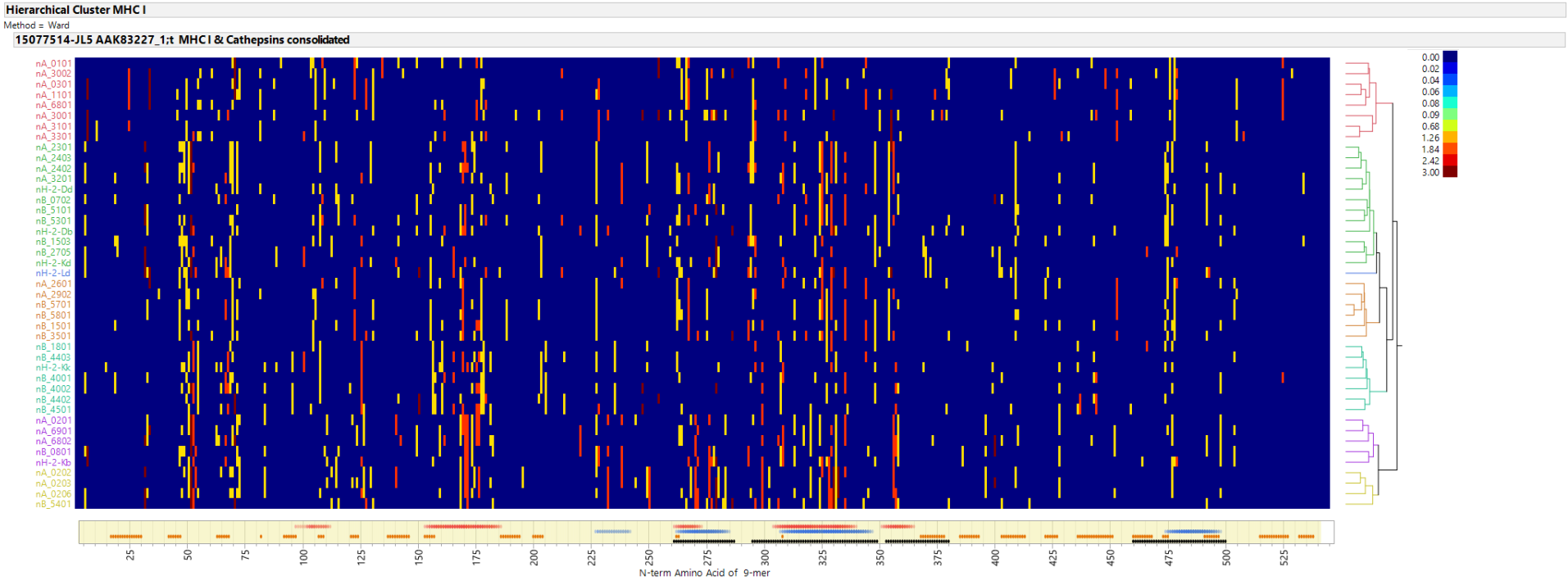

Hierarchical Cluster MHC II
Method = Ward
15077514-JL5 AAK83227_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB1_1101
nDRB5_0101
nDRB3_0202
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDRB1_1201
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.95
-3.16
-2.37
-1.58
-0.79
0.00
0.74
1.48
2.22
2.95
3.69
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

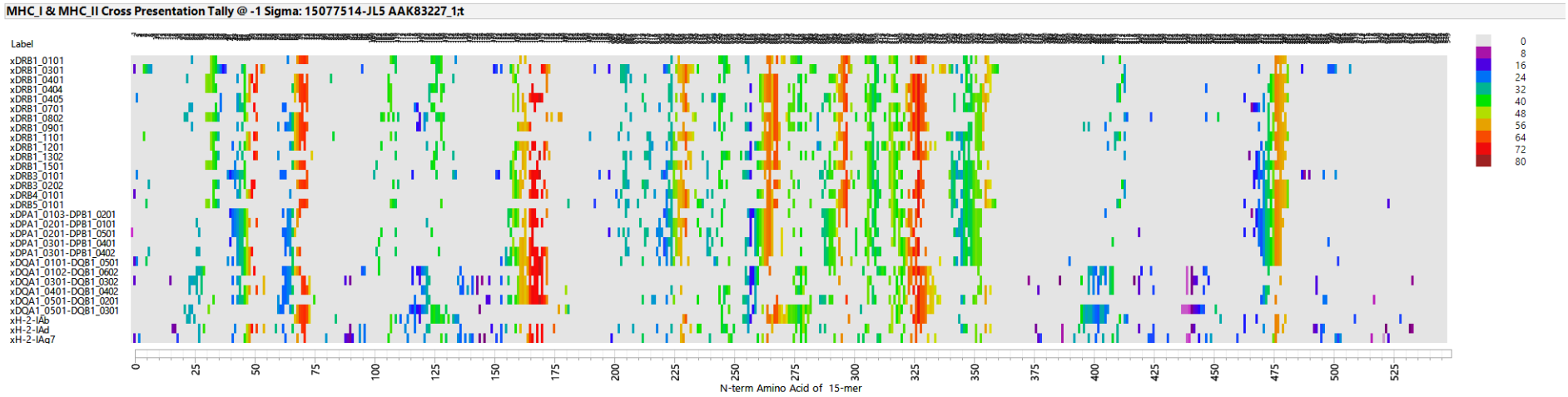

15077514-JL5 AAK83227_1;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50


15077514-JL5 AAK83227_1;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

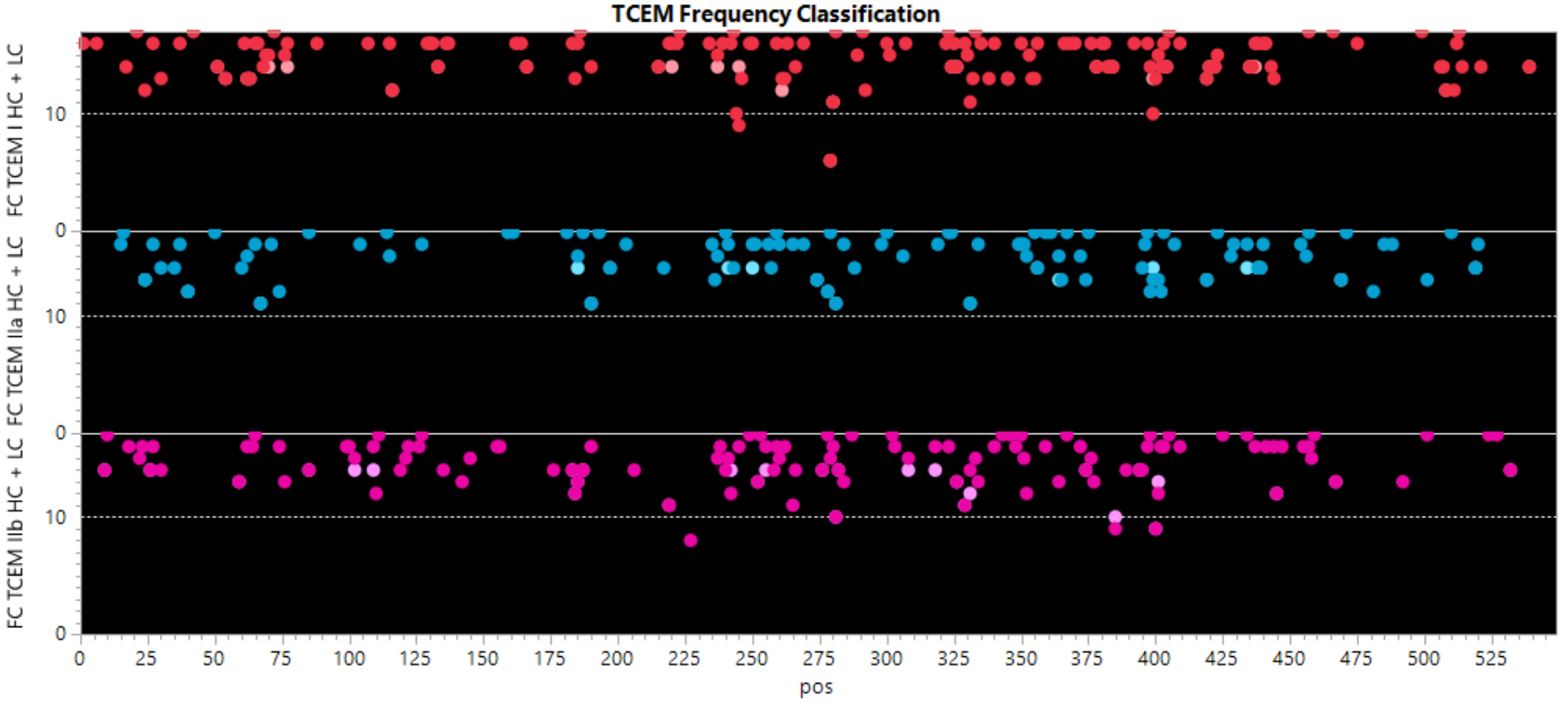

15077514-JL5 AAK83227_1;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

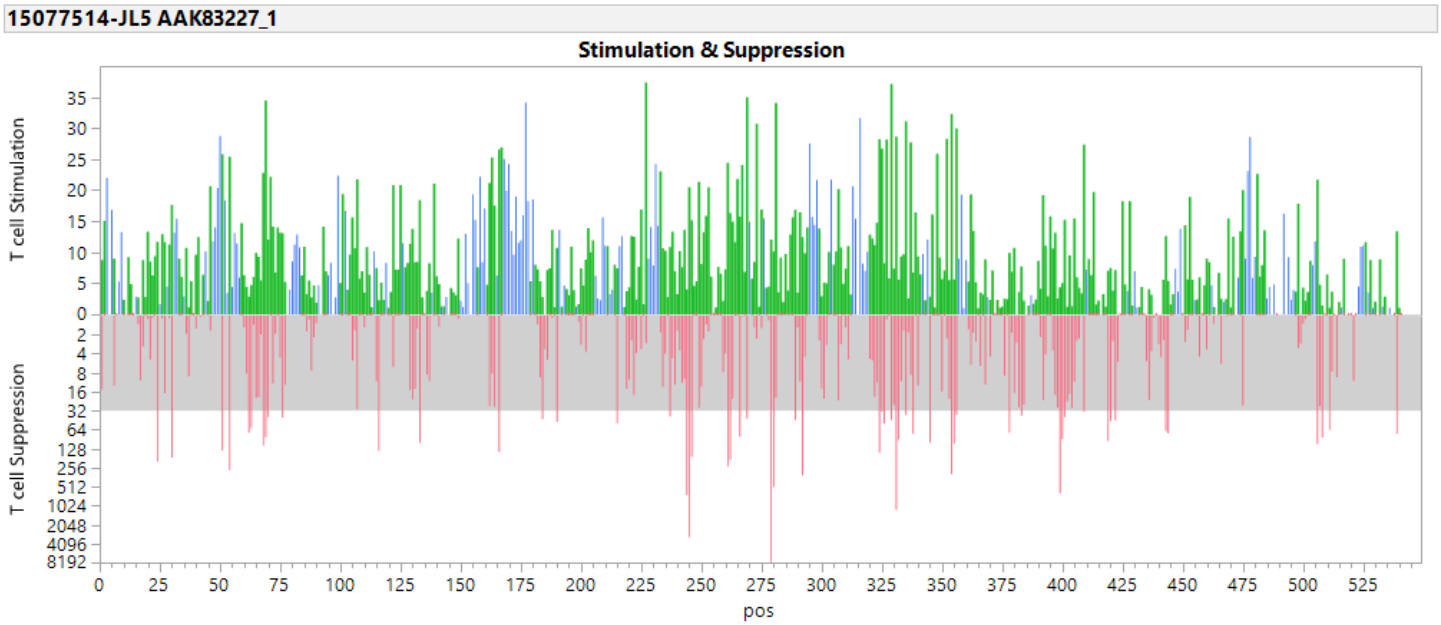

Hierarchical Cluster MHC I
Method = Ward
15077514-JL5 AAK83227_1;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nA_3201
nH-2-Dd
nB_1501
nB_0801
nH-2-Kb
nA_6802
nA_6901
nB_0702
nB_5101
nH-2-Db
nA_0201
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_3001
nB_5401
nB_1503
nB_2705
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.67
-3.74
-2.80
-1.87
-0.93
-0.00
1.24
2.47
3.71
4.94
6.18
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer

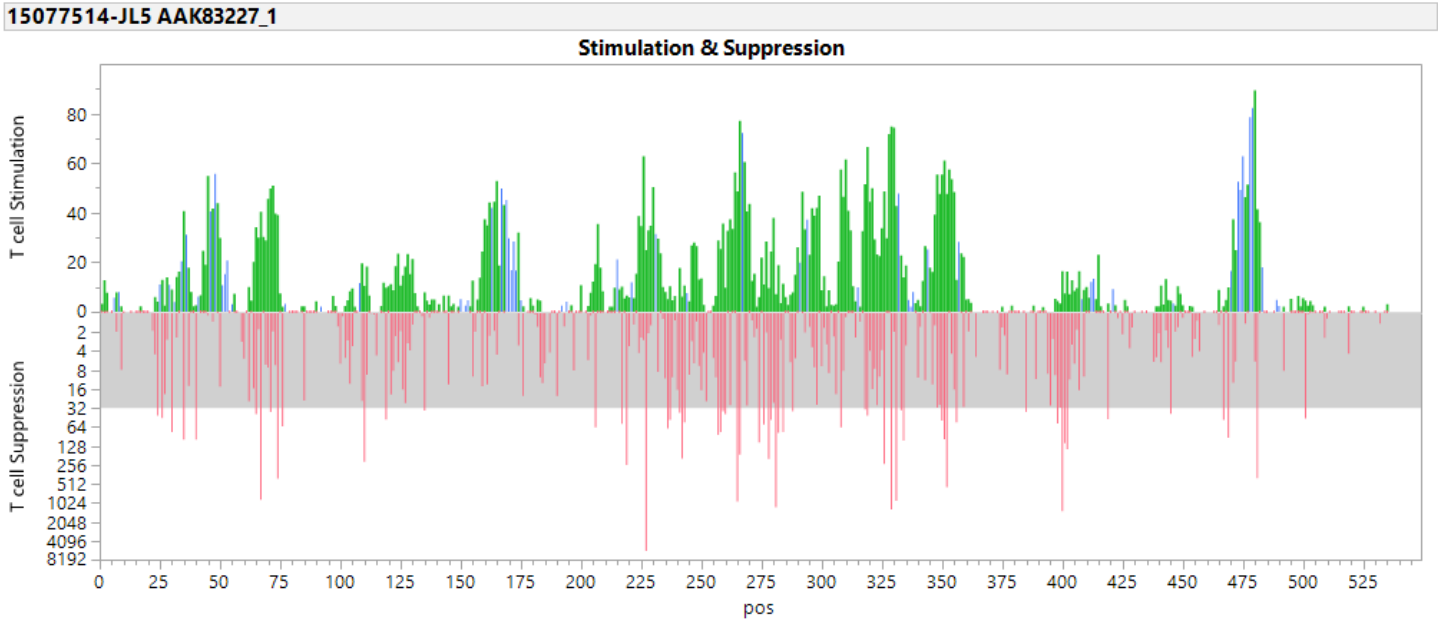

Hierarchical Cluster MHC I
Method = Ward
15077514-JL5 AAK83227_1;t MHC I & Cathepsins consolidated
nA_0101
nA_3002
nA_0301
nA_1101
nA_6801
nA_3001
nA_3101
nA_3301
nA_2301
nA_2403
nA_2402
nA_3201
nH-2-Dd
nB_0702
nB_5101
nB_5301
nH-2-Db
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_2601
nA_2902
nB_5701
nB_5801
nB_1501
nB_3501
nB_1801
nB_4403
nH-2-Kk
nB_4001
nB_4002
nB_4402
nB_4501
nA_0201
nA_6901
nA_6802
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nB_5401
0.00
0.02
0.04
0.06
0.08
0.09
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077514-JL5 AAK83227_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB2_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

15077514-JL5 AAK83227_1




15077514-JL5 AAK83227_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

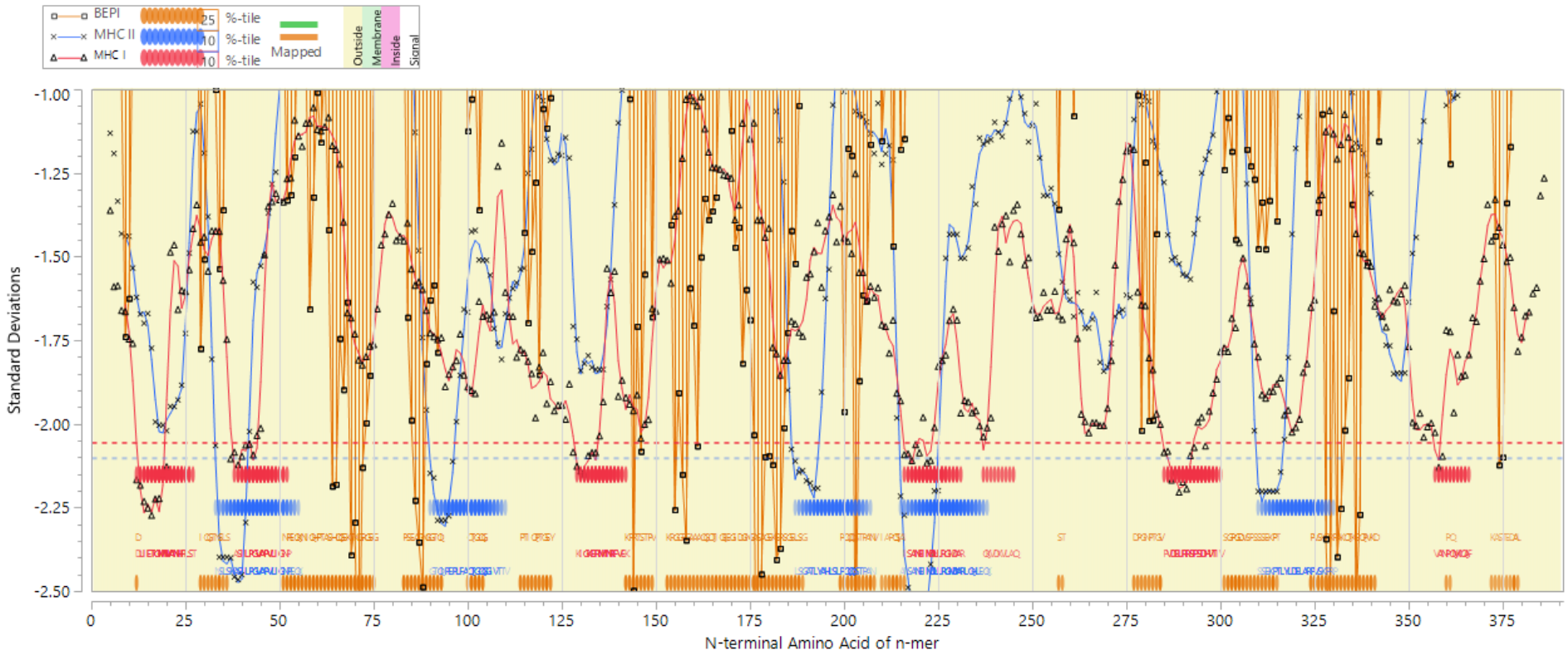

15077514-JL5 AAK83227_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

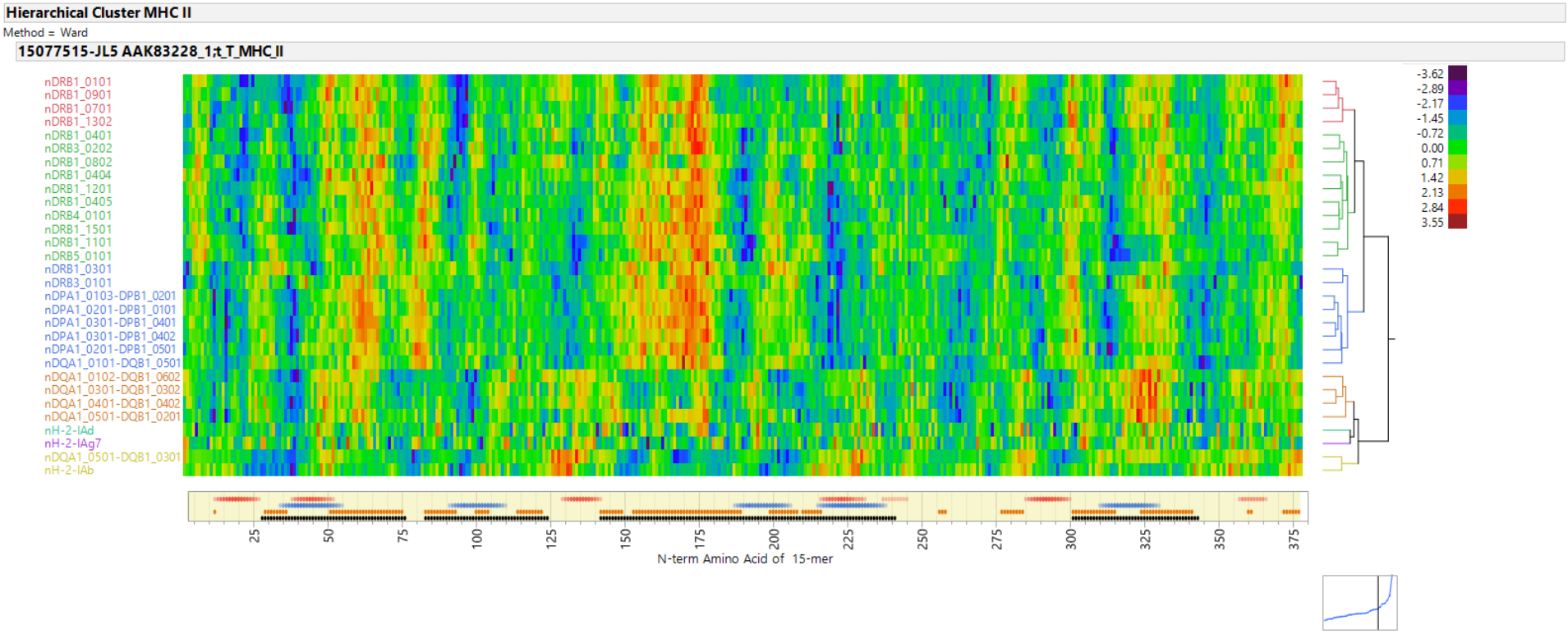
Phosphoprotein

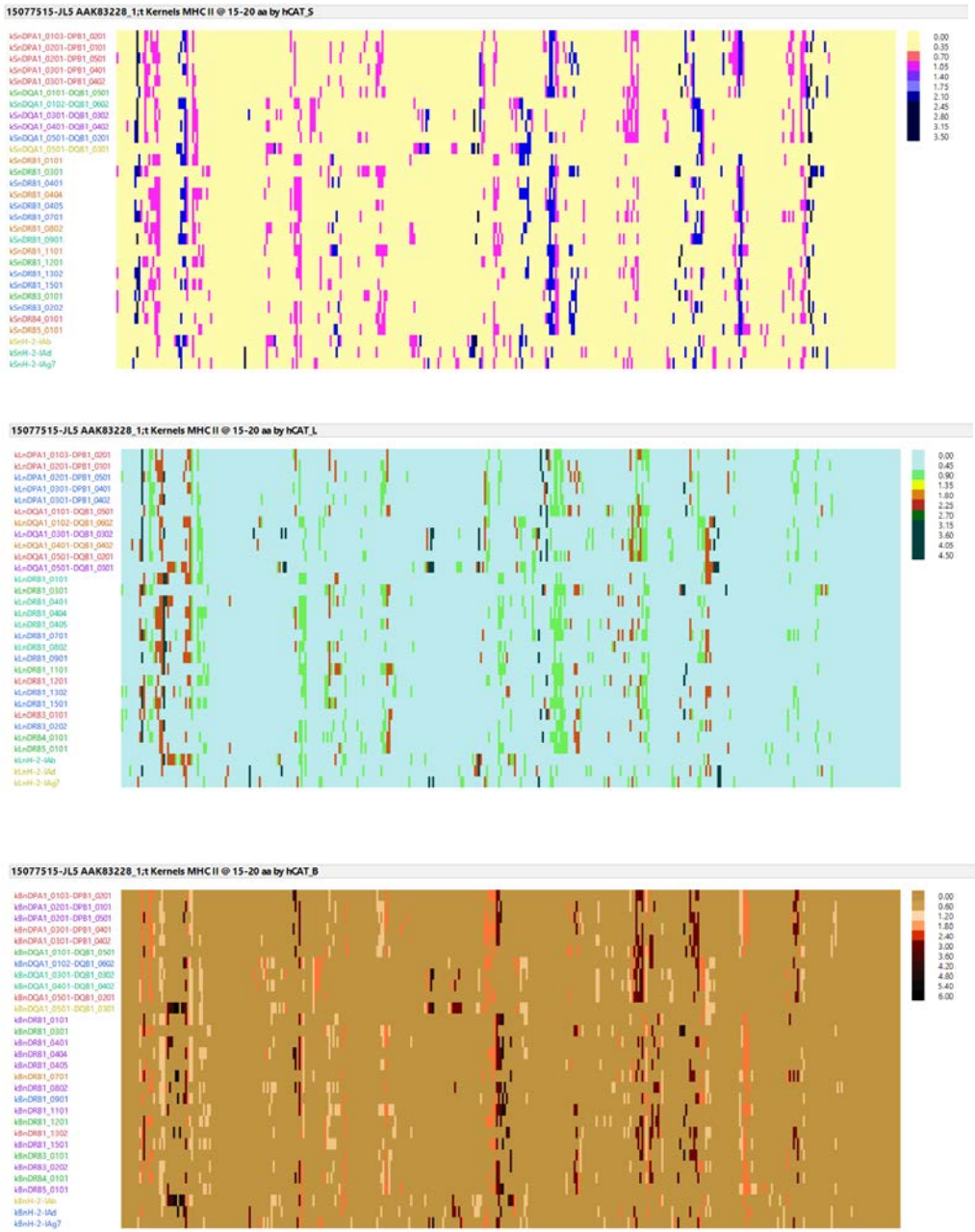
AAK83228.1| phosphoprotein

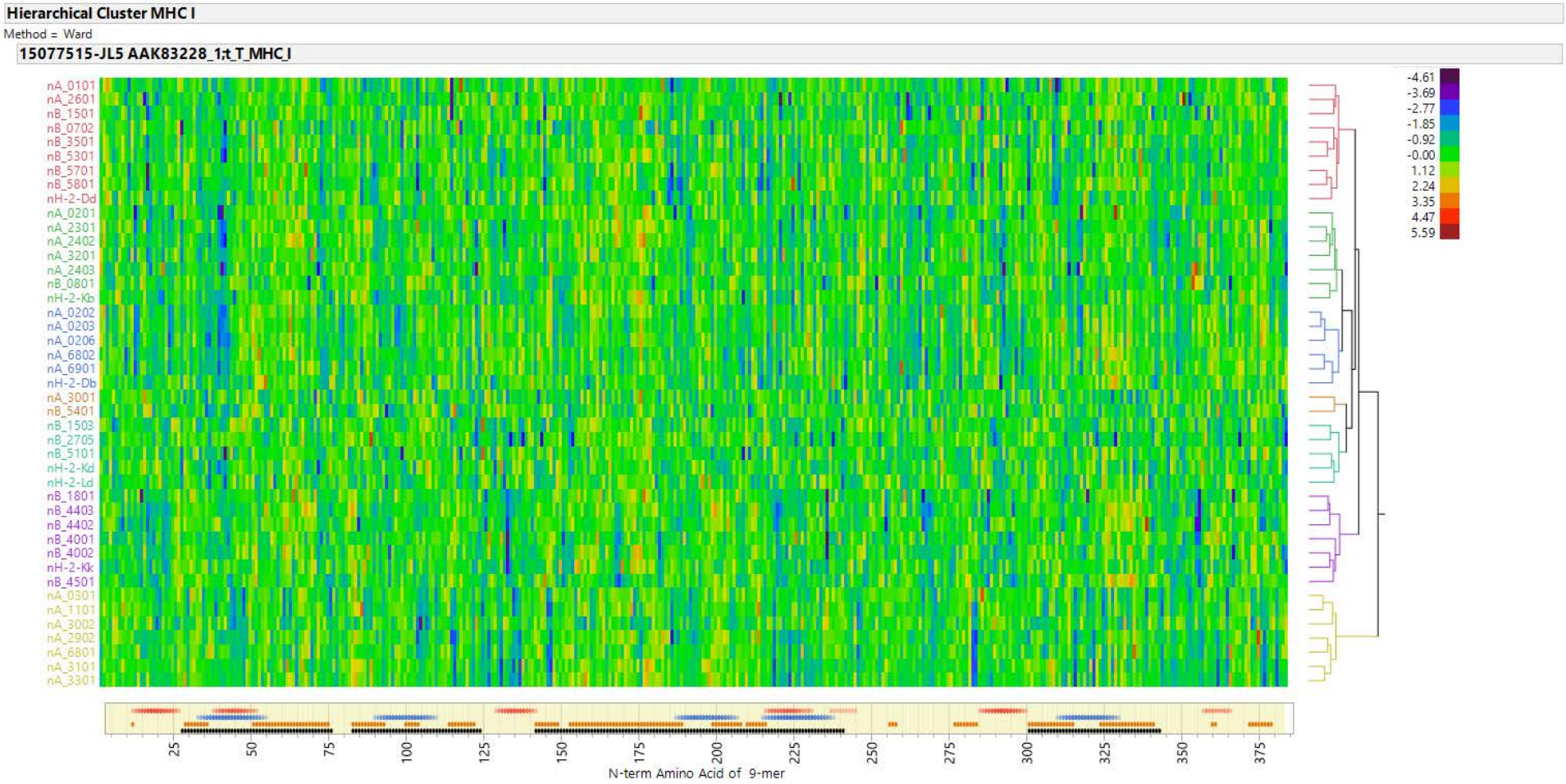
Overlay Plot

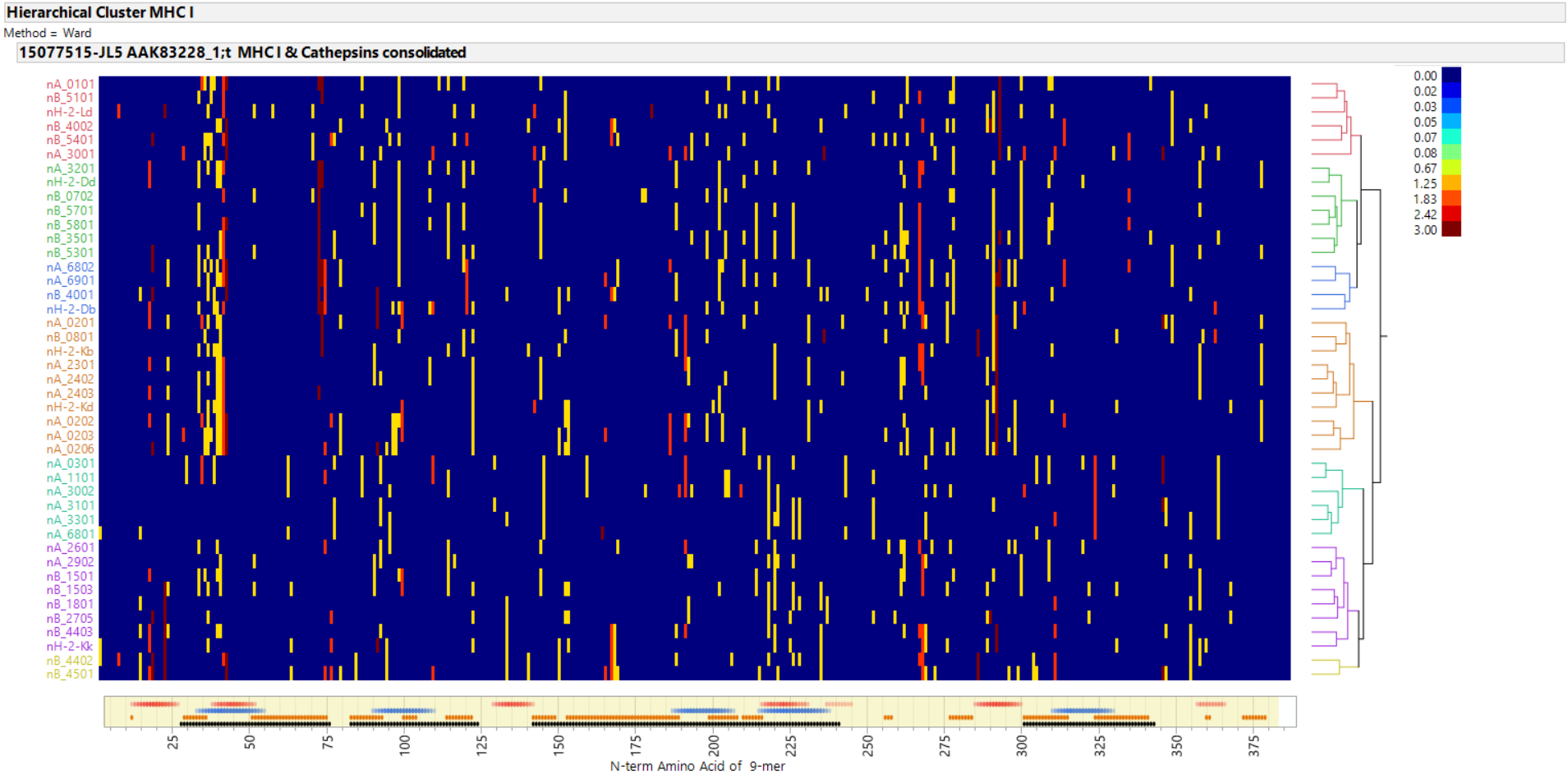
15077515-JL5 AAK83228_1;t

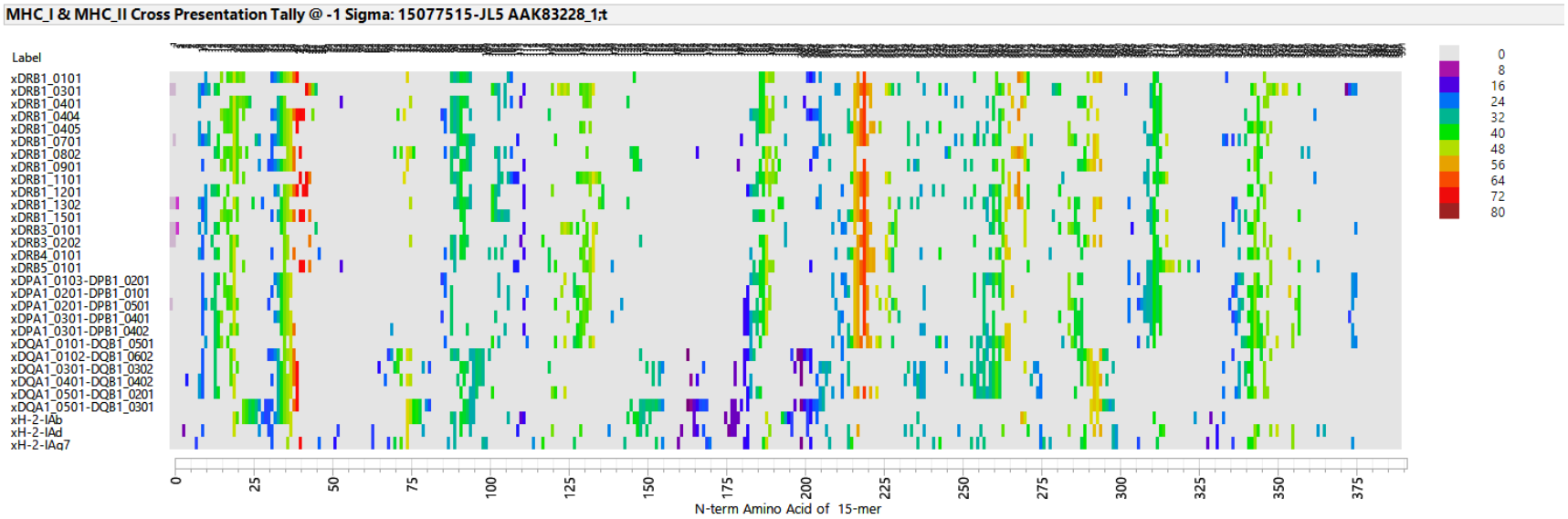



Hierarchical Cluster MHC II
Method = Ward
15077515-JL5 AAK83228_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_0404
nDRB1_1201
nDRB1_0405
nDRB4_0101
nDRB1_1501
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.62
-2.89
-2.17
-1.45
-0.72
0.00
0.71
1.42
2.13
2.84
3.55
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 15-mer

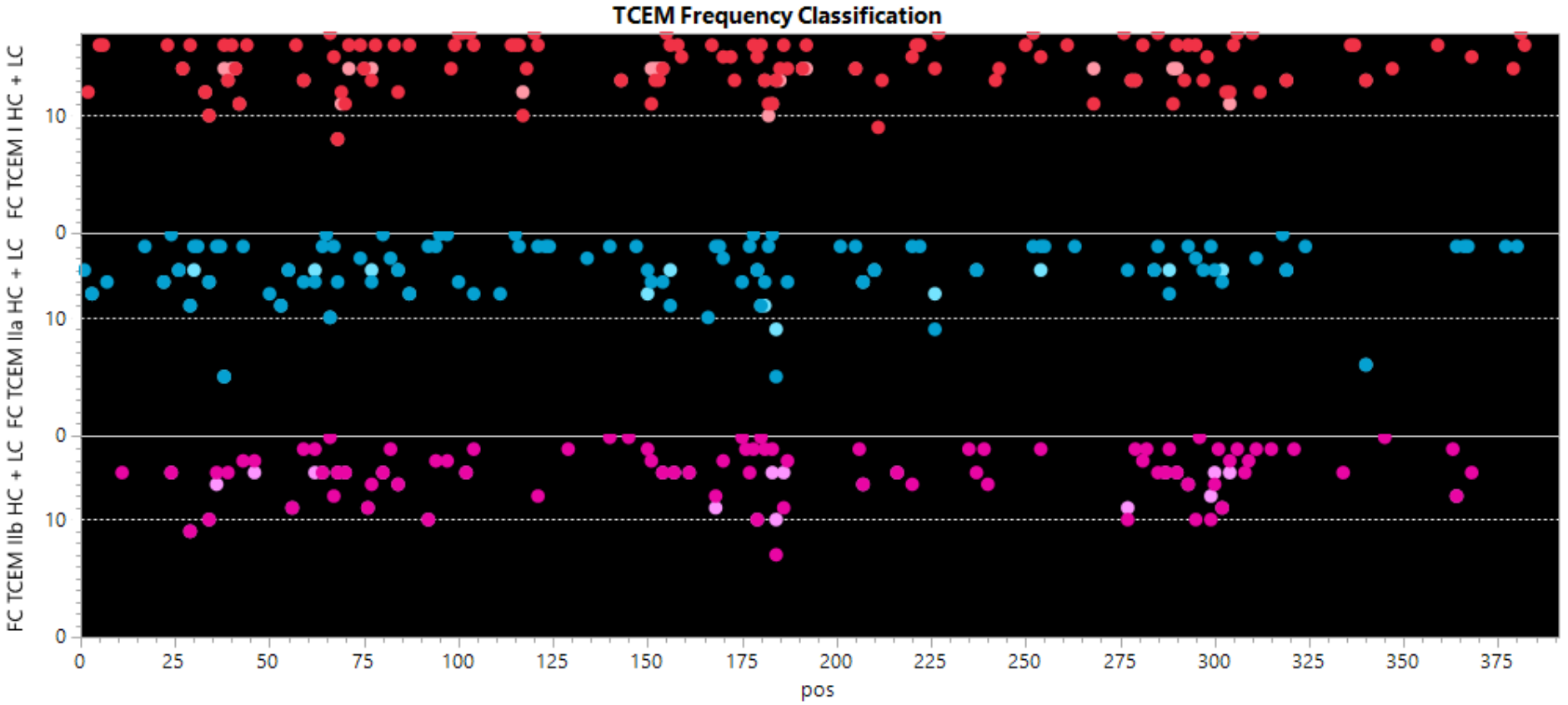

15077515-JL5 AAK83228_1;t Kernels MHC II @ 15-20 aa by hCAT_S
15077515-JL5 AAK83228_1;t Kernels MHC II @ 15-20 aa by hCAT_L
15077515-JL5 AAK83228_1;t Kernels MHC II @ 15-20 aa by hCAT_B

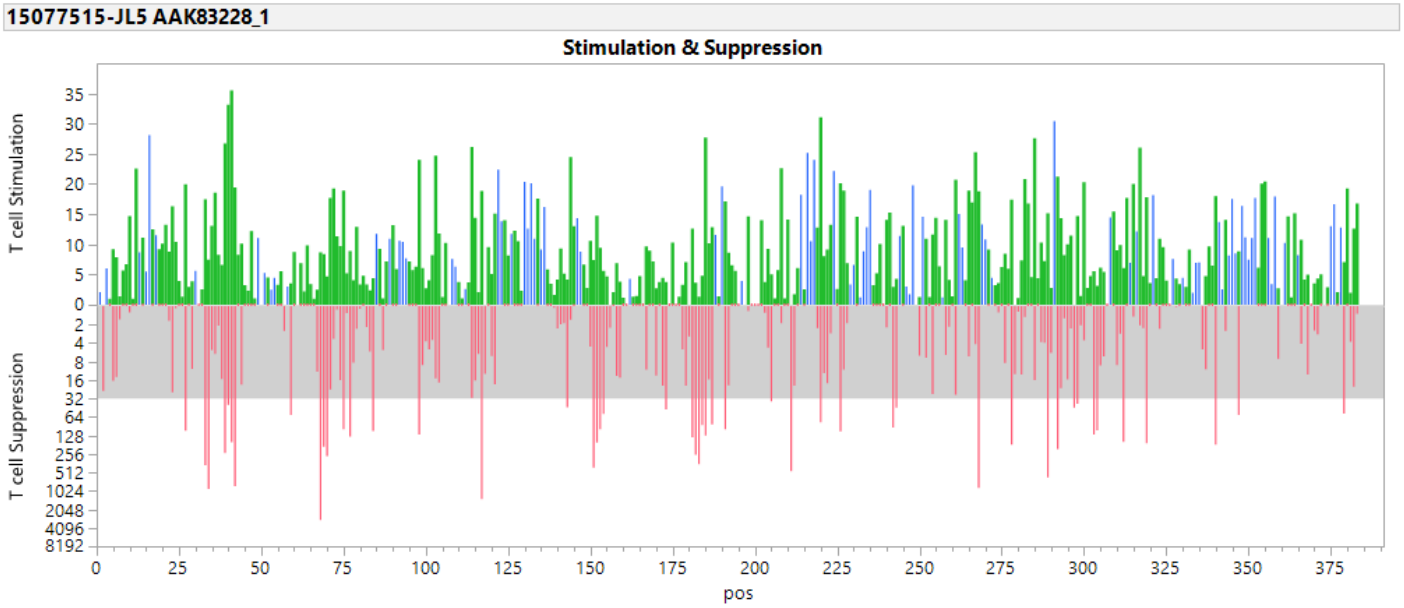

Hierarchical Cluster MHC I
Method = Ward
15077515-JL5 AAK83228_1;t T_MHC I
nA_0101
nA_2601
nB_1501
nB_0702
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_2301
nA_2402
nA_3201
nA_2403
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nH-2-Db
nA_3001
nB_5401
nB_1503
nB_2705
nB_5101
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_6801
nA_3101
nA_3301
-4.61
-3.69
-2.77
-1.85
-0.92
-0.00
1.12
2.24
3.35
4.47
5.59
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 9-mer

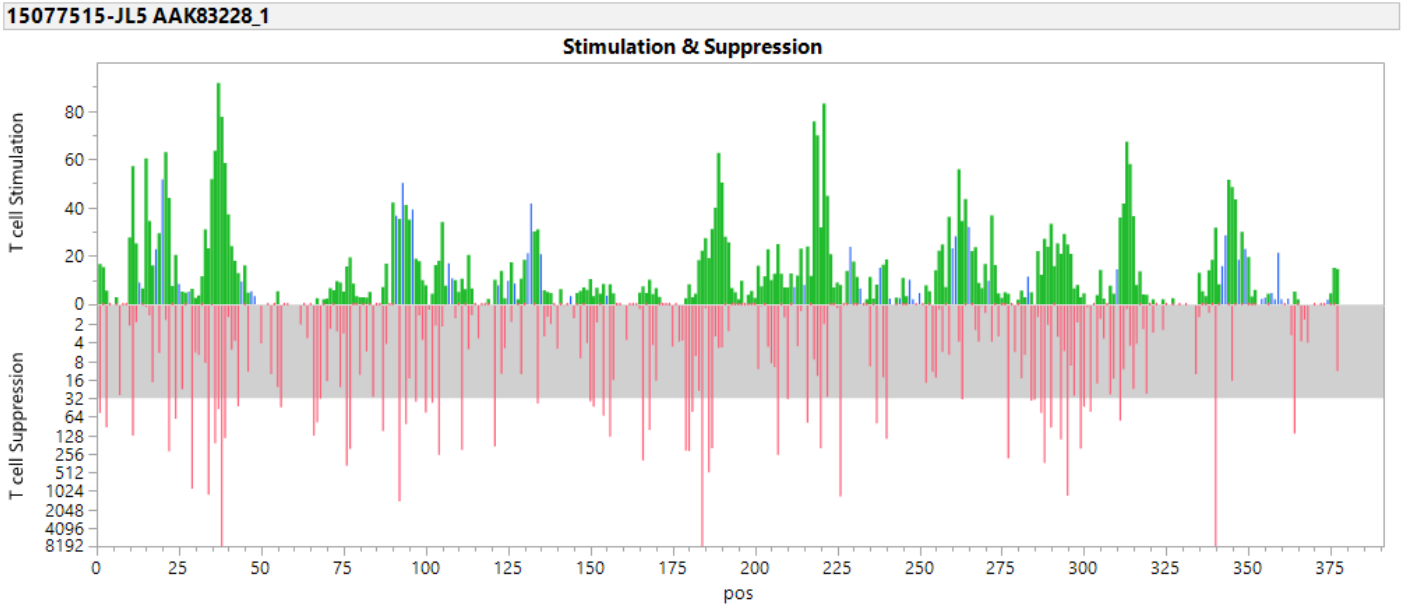

Hierarchical Cluster MHC I
Method = Ward
15077515-JL5 AAK83228_1;t MHC I & Cathepsins consolidated
nA_0101
nB_5101
nH-2-Ld
nB_4002
nB_5401
nA_3001
nA_3201
nH-2-Dd
nB_0702
nB_5701
nB_5801
nB_3501
nB_5301
nA_6802
nA_6901
nB_4001
nH-2-Db
nA_0201
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_0202
nA_0203
nA_0206
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_2601
nA_2902
nB_1501
nB_1503
nB_1801
nB_2705
nB_4403
nH-2-Kk
nB_4402
nB_4501
0.00
0.02
0.03
0.05
0.07
0.08
0.67
1.25
1.83
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077515-JL5 AAK83228_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
8
16
24
32
40
48
56
64
72
80
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
N-term Amino Acid of 15-mer

15077515-JL5 AAK83228_1



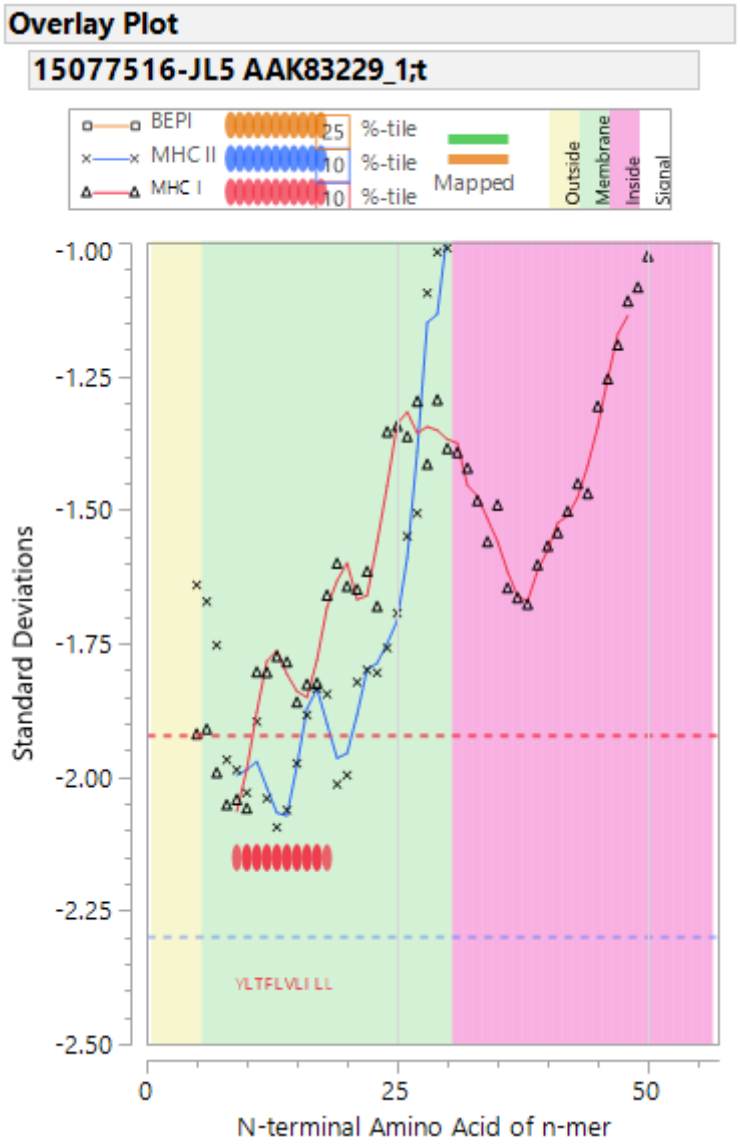

15077515-JL5 AAK83228_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

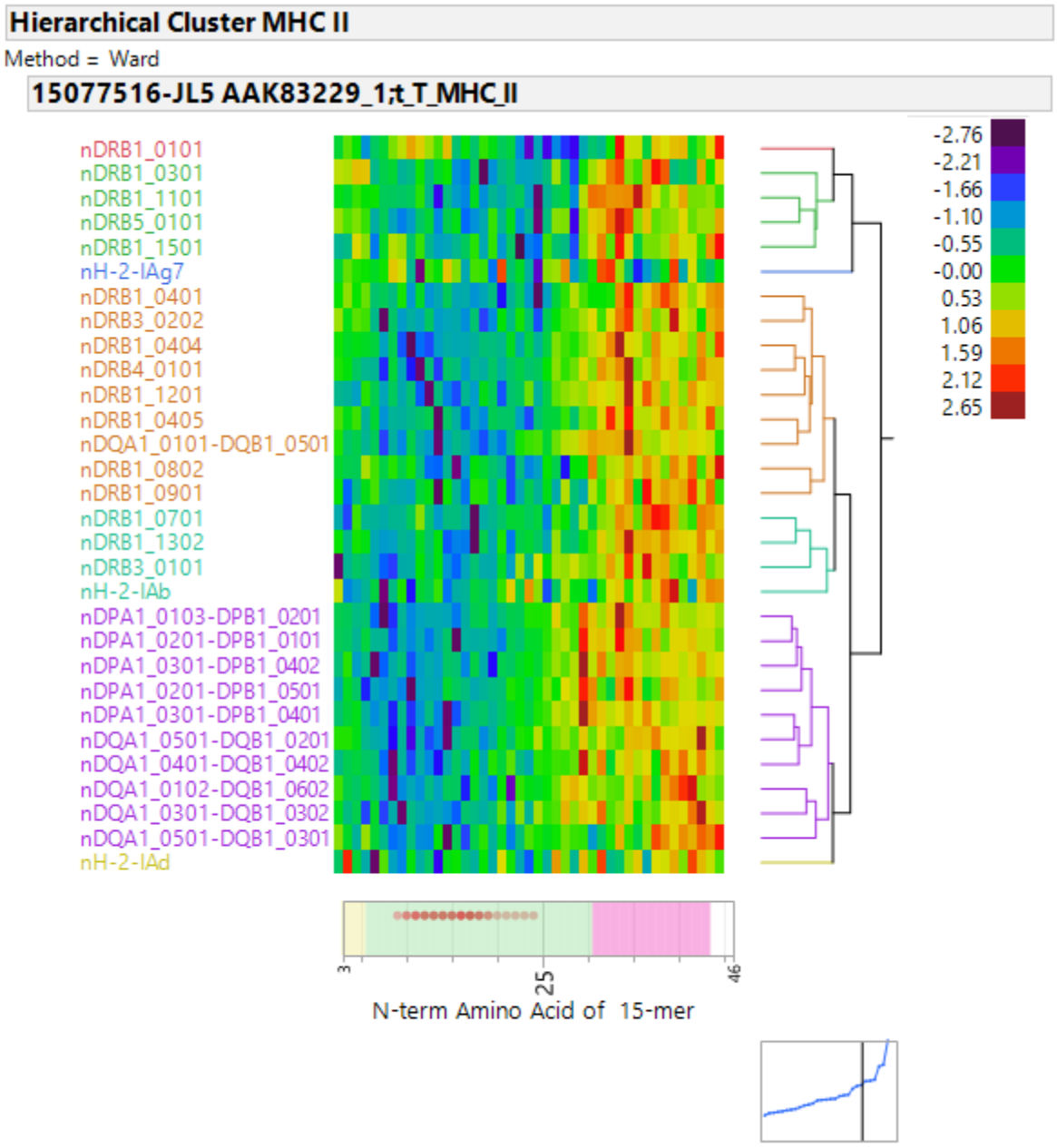

15077515-JL5 AAK83228_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

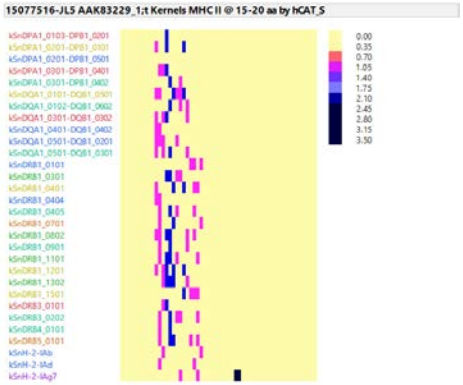
Small hydrophobic protein

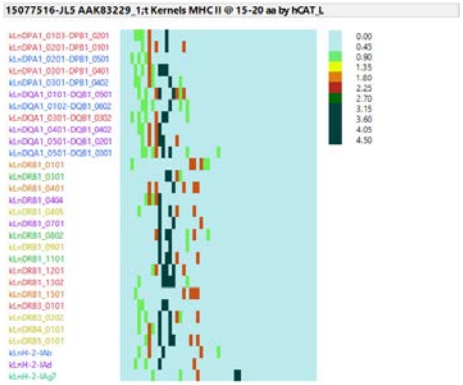
AAK83229.1| small hydrophobic protein

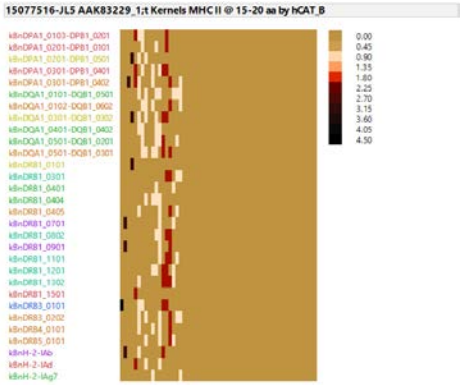
Note accuracy at C term is reduced in very short proteins – could be re run with “padding” of dummy amino acids at each end

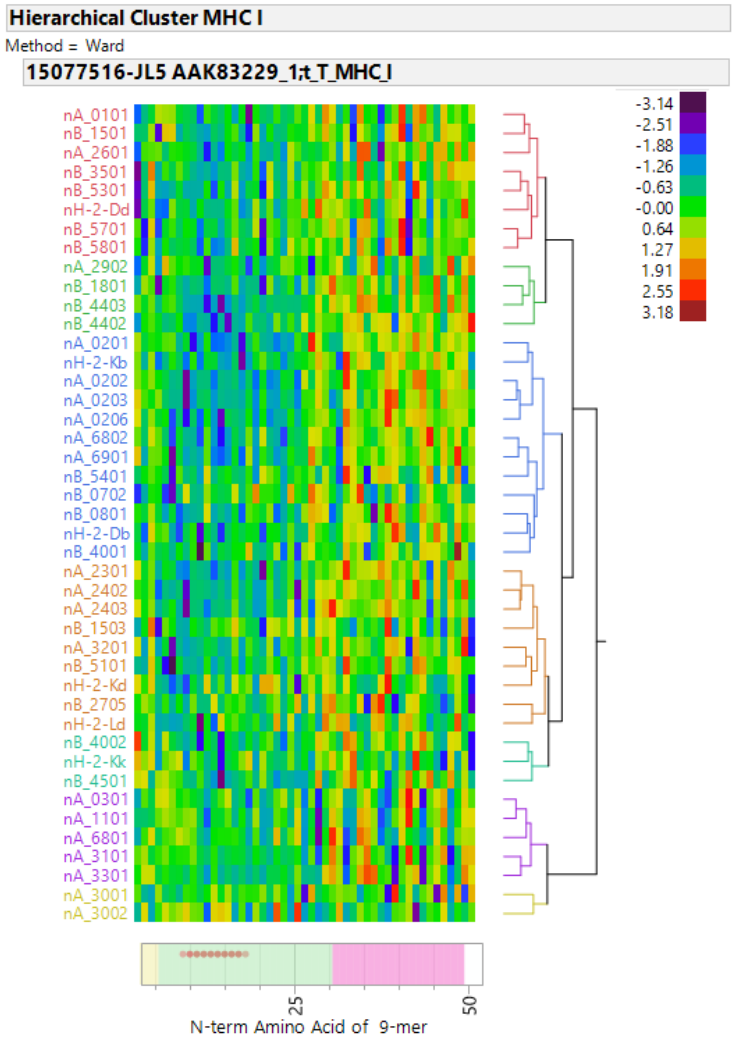

Overlay Plot
15077516-JL5 AAK83229_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
Standard Deviations
YLTFLVLI LL
0
25
50
N-terminal Amino Acid of n-mer

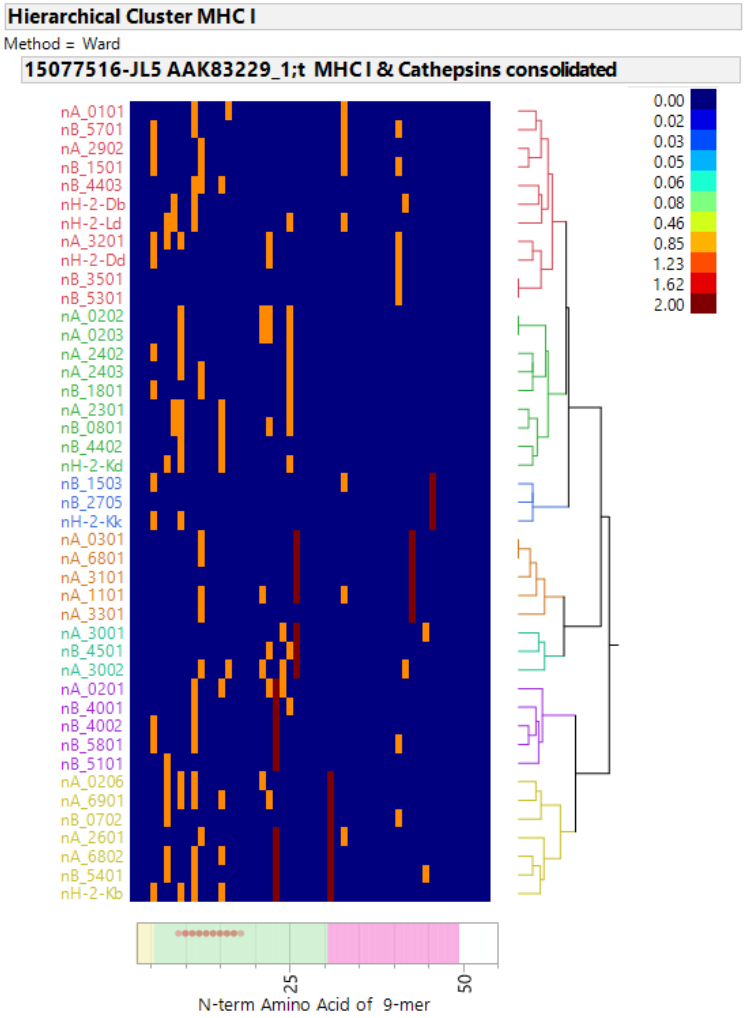

Hierarchical Cluster MHC II
Method = Ward
15077516-JL5 AAK83229_1;t_T_MHC_II
nDRB1_0101
nDRB1_0301
nDRB1_1101
nDRB5_0101
nDRB1_1501
nH-2-IAg7
nDRB1_0401
nDRB3_0202
nDRB1_0404
nDRB4_0101
nDRB1_1201
nDRB1_0405
nDQA1_0101-DQB1_0501
nDRB1_0802
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB3_0101
nH-2-IAb
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0401
nDQA1_0501-DQB1_0201
nDQA1_0401-DQB1_0402
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0501-DQB1_0301
nH-2-IAd
-2.76
-2.21
-1.66
-1.10
-0.55
-0.00
0.53
1.06
1.59
2.12
2.65
3
25
46
N-term Amino Acid of 15-mer

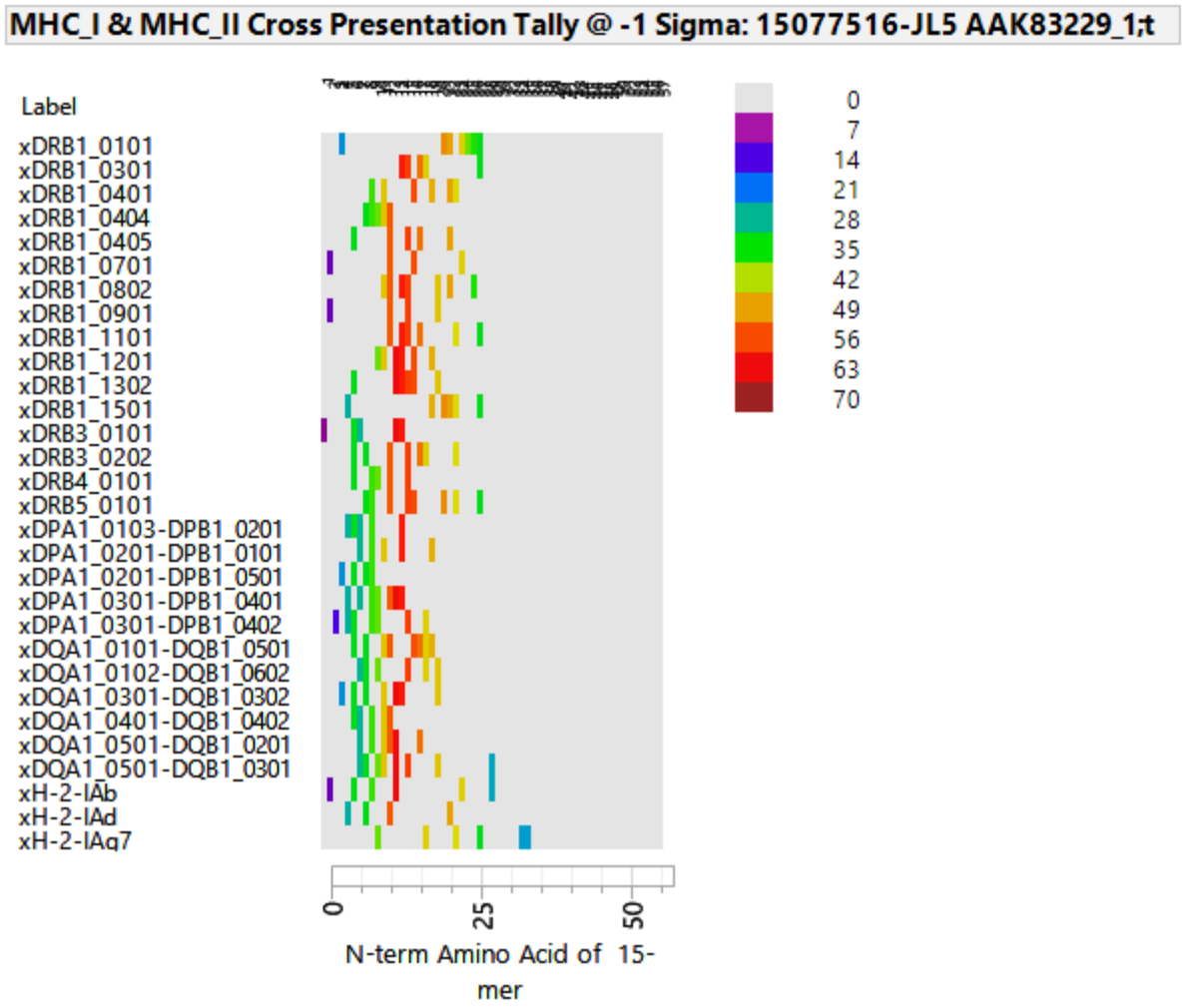

15077516-JL5 AAK83229_1;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.35
0.70
1.05
1.40
1.75
2.10
2.45
2.80
3.15
3.50

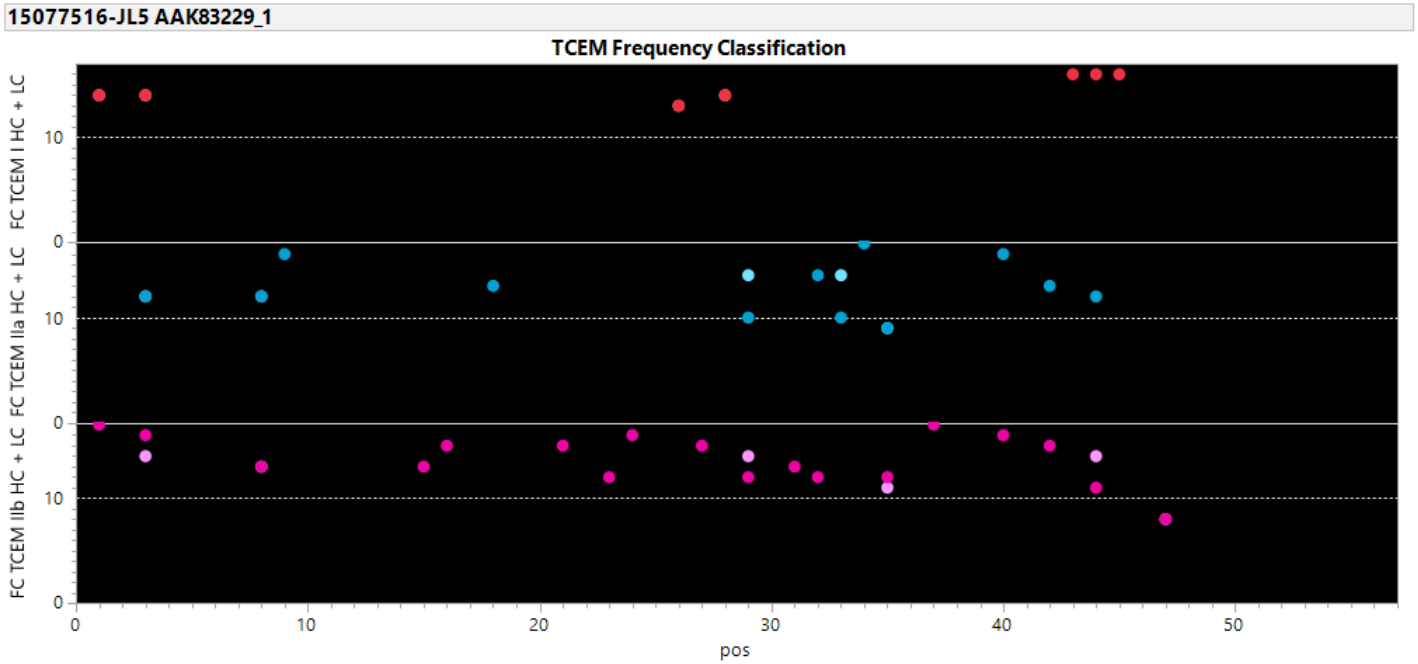

15077516-JL5 AAK83229_1;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

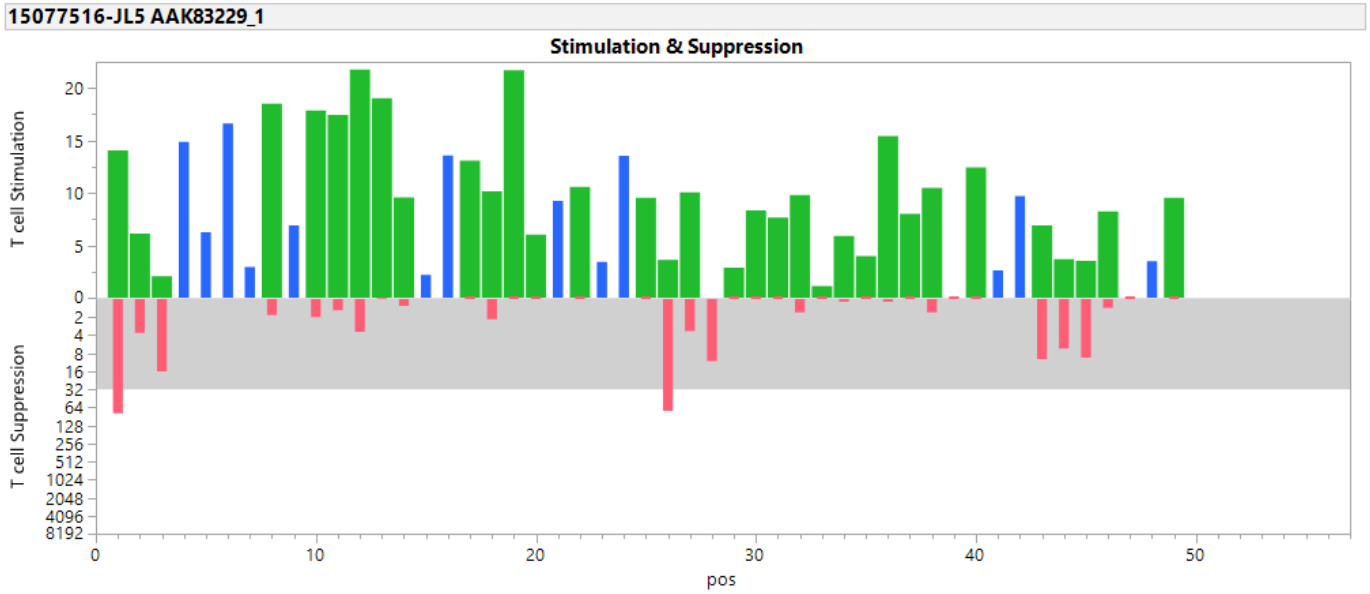

15077516-JL5 AAK83229_1;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

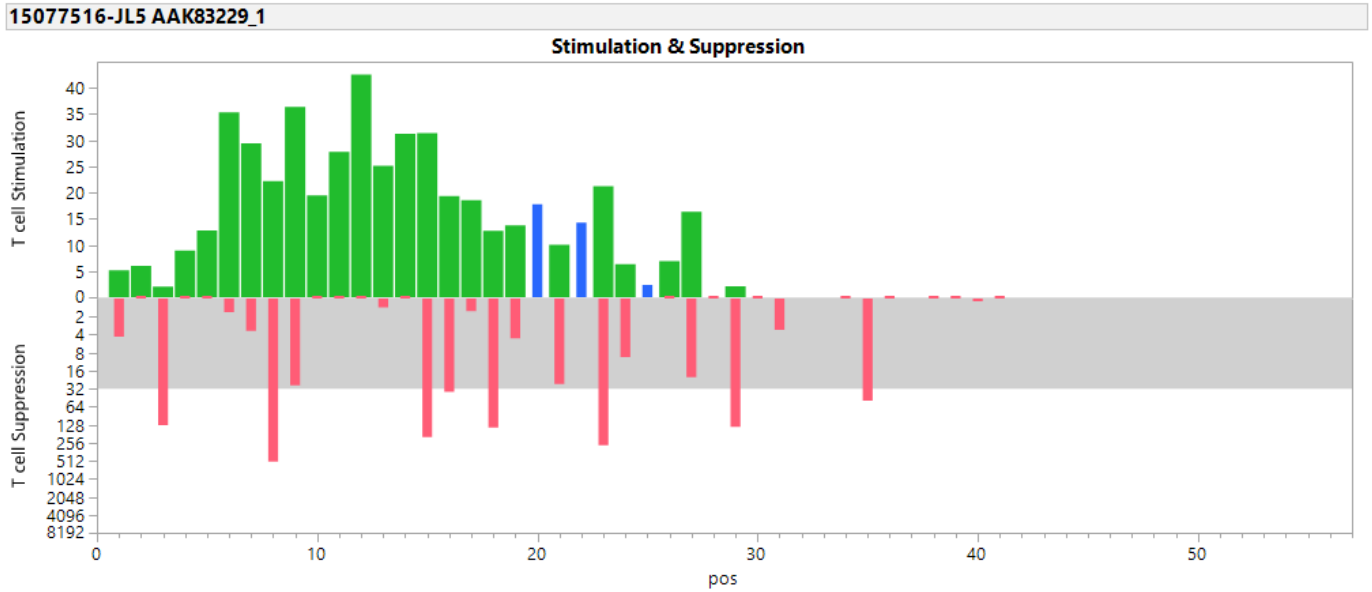

Hierarchical Cluster MHC I
Method = Ward
15077516-JL5 AAK83229_1;t_T_MHC_I
nA_0101
nB_1501
nA_2601
nB_3501
nB_5301
nH-2-Dd
nB_5701
nB_5801
nA_2902
nB_1801
nB_4403
nB_4402
nA_0201
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5401
nB_0702
nB_0801
nH-2-Db
nB_4001
nA_2301
nA_2402
nA_2403
nB_1503
nA_3201
nB_5101
nH-2-Kd
nB_2705
nH-2-Ld
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_6801
nA_3101
nA_3301
nA_3001
nA_3002
-3.14
-2.51
-1.88
-1.26
-0.63
-0.00
0.64
1.27
1.91
2.55
3.18
25
50
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
15077516-JL5 AAK83229_1;t MHC I & Cathepsins consolidated
nA_0101
nB_5701
nA_2902
nB_1501
nB_4403
nH-2-Db
nH-2-Ld
nA_3201
nH-2-Dd
nB_3501
nB_5301
nA_0202
nA_0203
nA_2402
nA_2403
nB_1801
nA_2301
nB_0801
nB_4402
nH-2-Kd
nB_1503
nB_2705
nH-2-Kk
nA_0301
nA_6801
nA_3101
nA_1101
nA_3301
nA_3001
nB_4501
nA_3002
nA_0201
nB_4001
nB_4002
nB_5801
nB_5101
nA_0206
nA_6901
nB_0702
nA_2601
nA_6802
nB_5401
nH-2-Kb
0.00
0.02
0.03
0.05
0.06
0.08
0.46
0.85
1.23
1.62
2.00
25
50
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077516-JL5 AAK83229_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
7
14
21
28
35
42
49
56
63
70
0
25
50
N-term Amino Acid of 15-mer


15077516-JL5 AAK83229_1
TCEM Frequency Classification
FC TCEM IIb HC + LC
FC TCEM IIa HC + LC
FC TCEM I HC + LC
10
0
10
0
10
0
0
10
20
30
40
50
pos


15077516-JL5 AAK83229_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

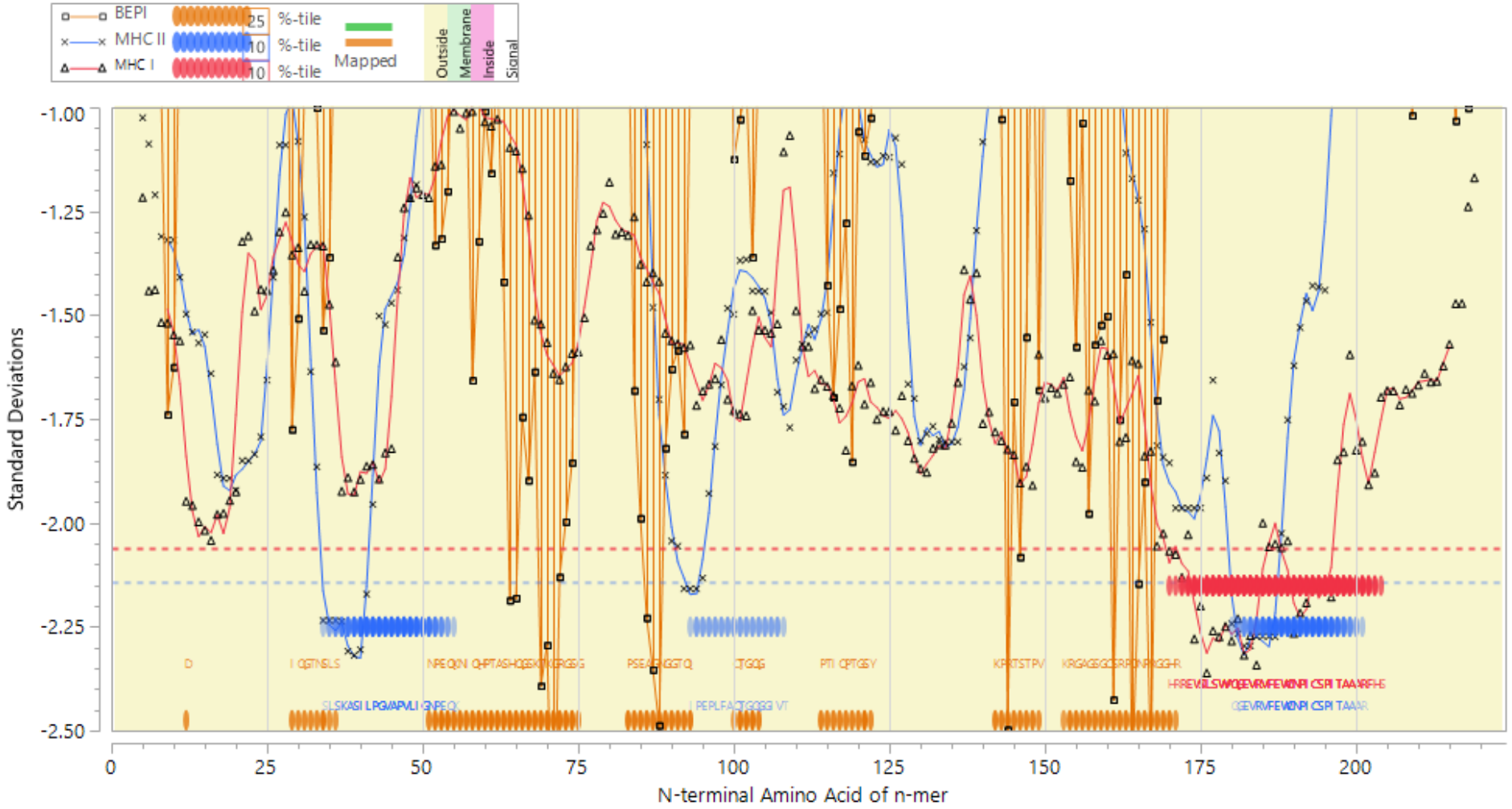

15077516-JL5 AAK83229_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

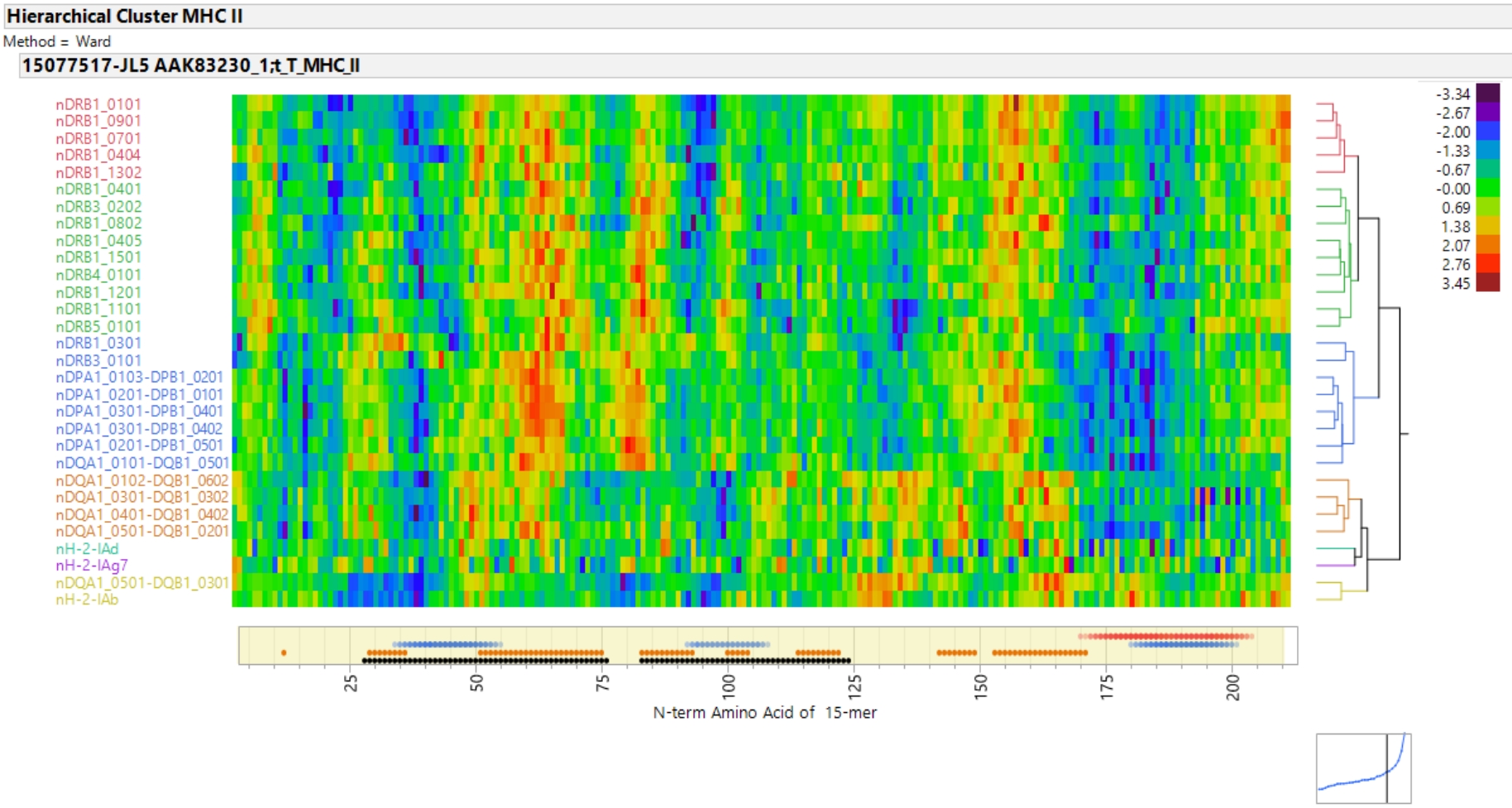
V protein

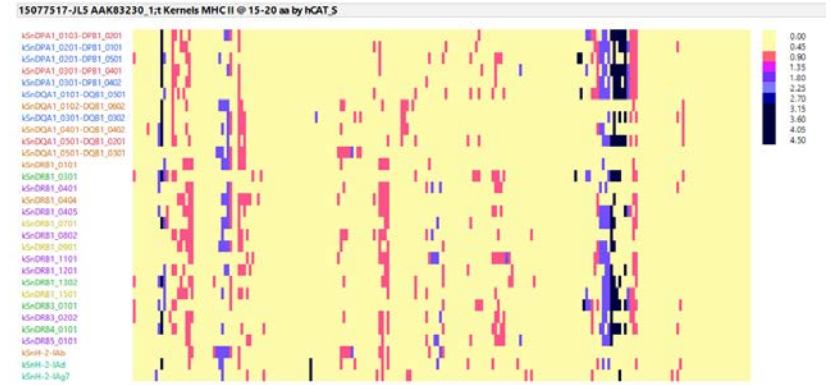
AAK83230.1| V protein

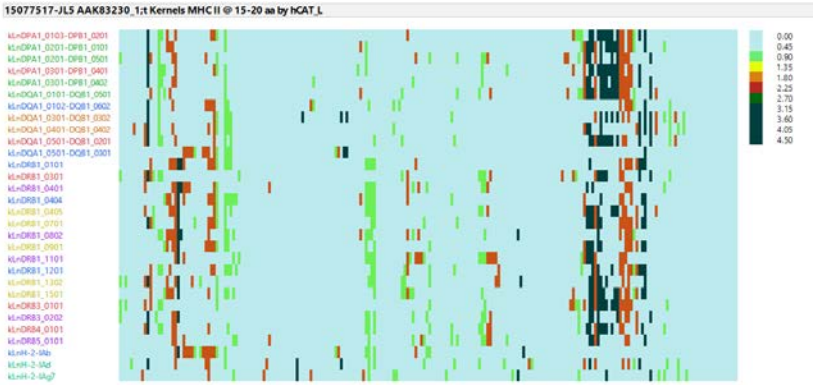
Overlay Plot

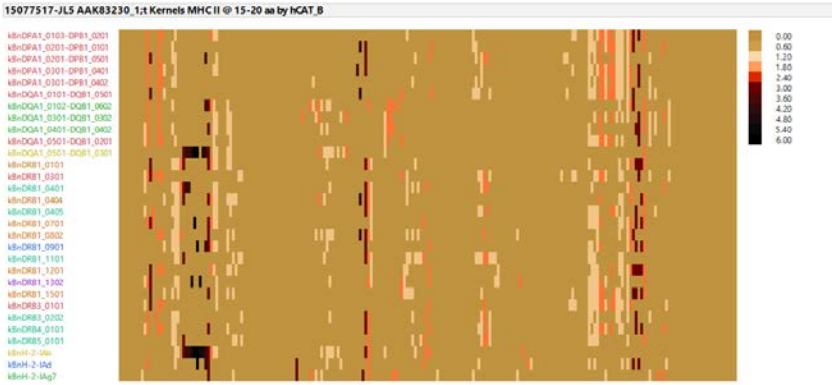
15077517-JL5 AAK83230_1;t

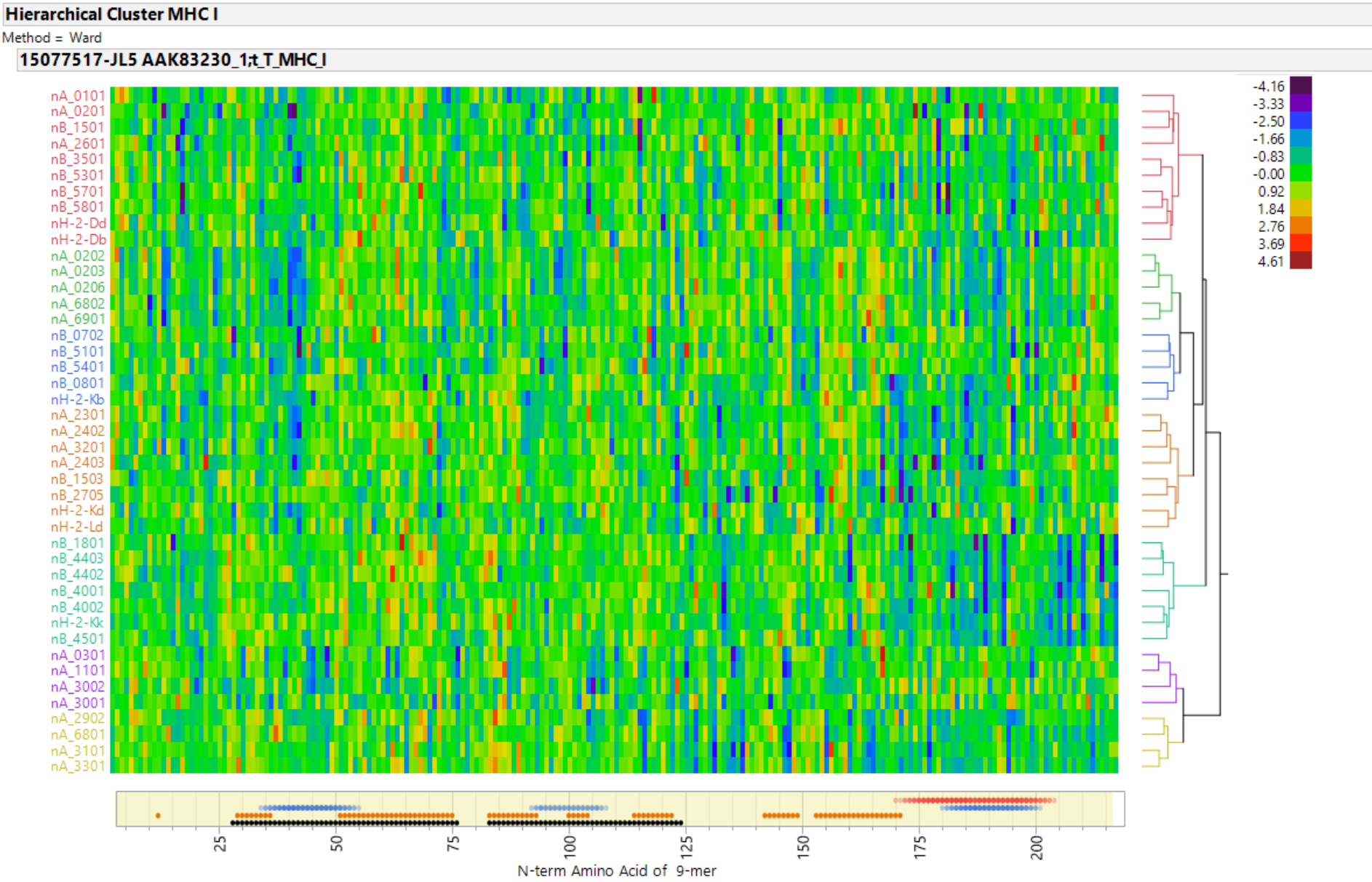



Hierarchical Cluster MHC II
Method = Ward
15077517-JL5 AAK83230_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.34
-2.67
-2.00
-1.33
-0.67
-0.00
0.69
1.38
2.07
2.76
3.45
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer


15077517-JL5 AAK83230_1;t Kernels MHC II @ 15-20 aa by hCAT_S


15077517-JL5 AAK83230_1;t Kernels MHC II @ 15-20 aa by hCAT_L

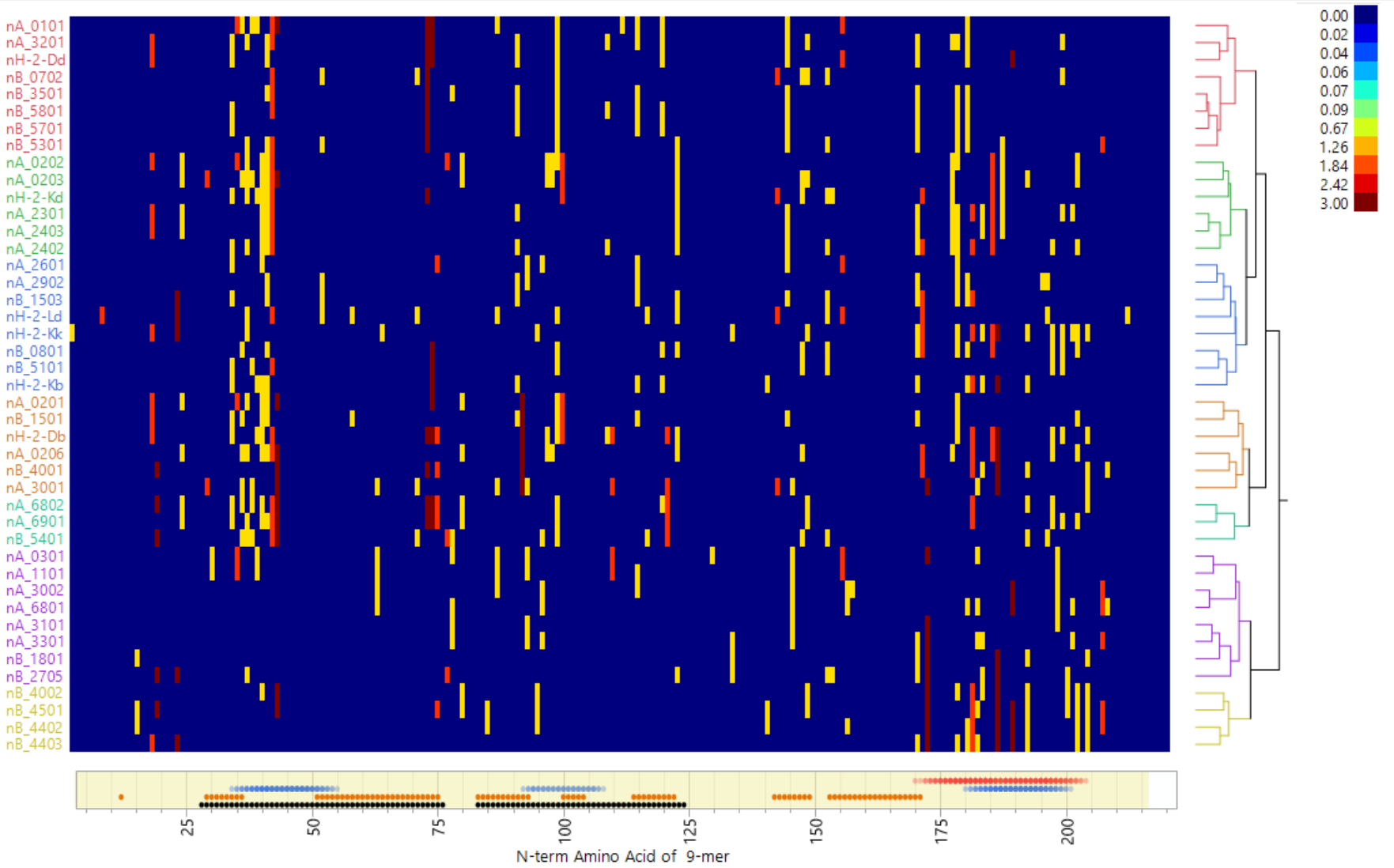

15077517-JL5 AAK83230_1;t Kernels MHC II @ 15-20 aa by hCAT_B

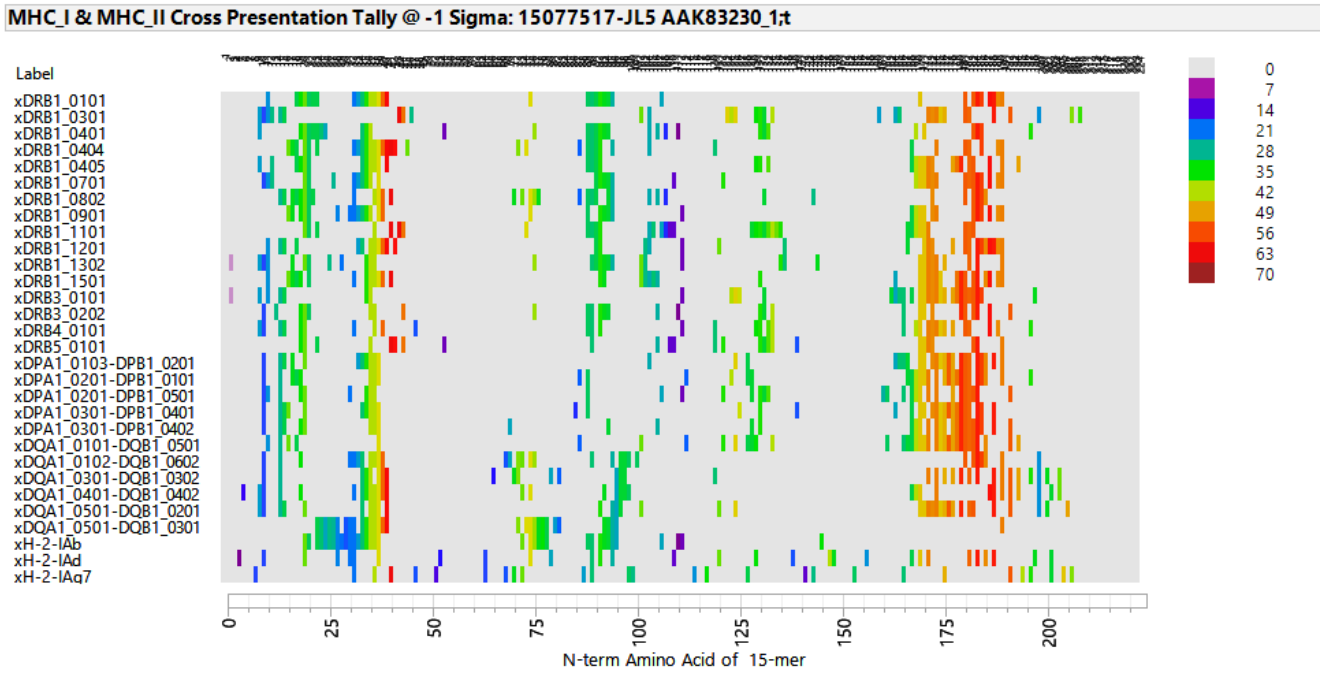

Hierarchical Cluster MHC I
Method = Ward
15077517-JL5 AAK83230_1;t_T_MHC_I
nA_0101
nA_0201
nB_1501
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nH-2-Db
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_5101
nB_5401
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_3001
nA_2902
nA_6801
nA_3101
nA_3301
-4.16
-3.33
-2.50
-1.66
-0.83
-0.00
0.92
1.84
2.76
3.69
4.61
25
50
75
100
125
150
175
200
N-term Amino Acid of 9-mer

Hierarchical Cluster MHC I

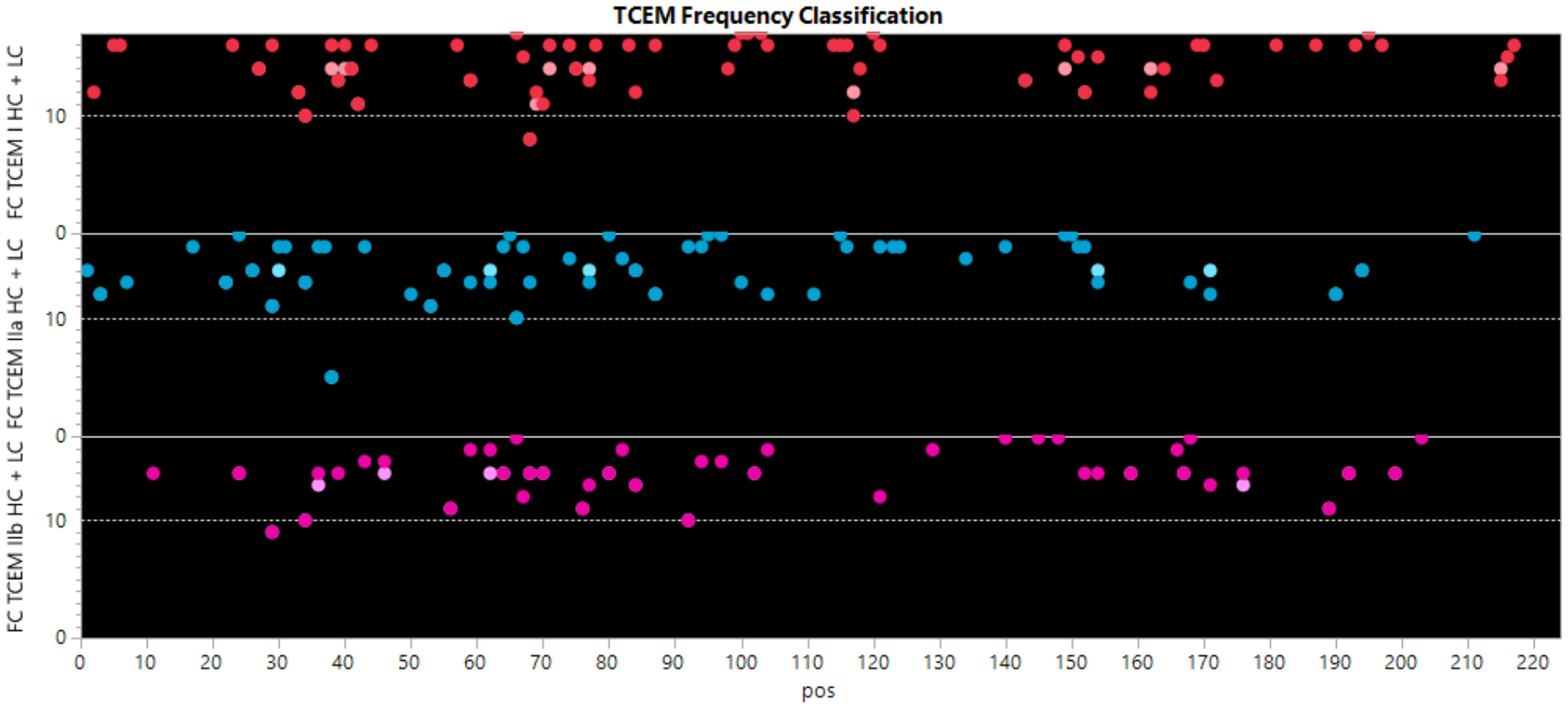
Method = Ward

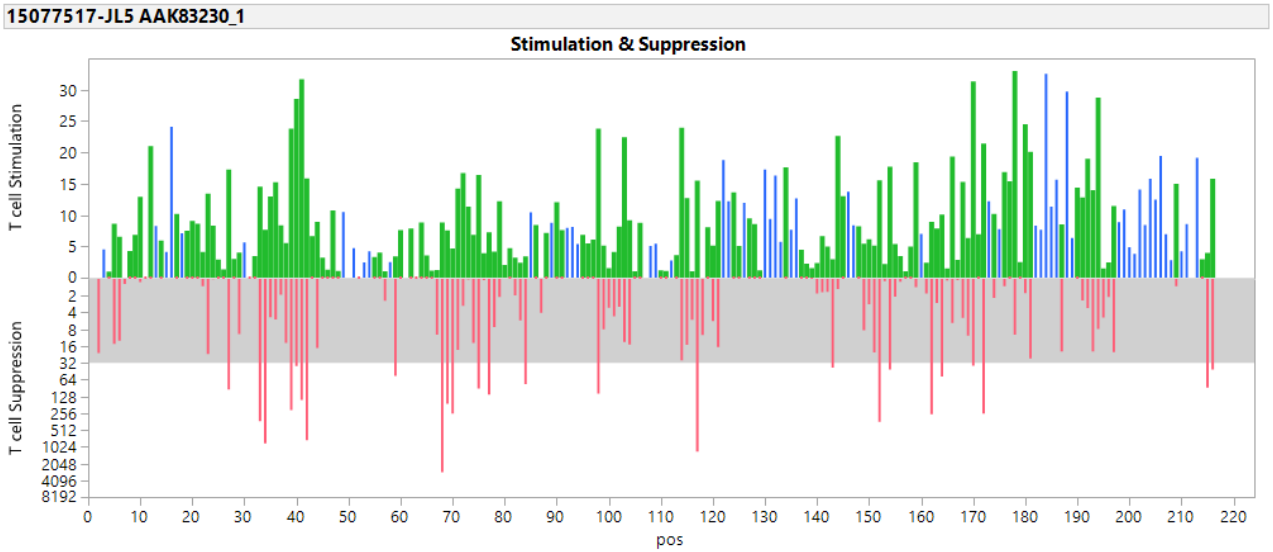
15077517-JL5 AAK83230_1;t MHC I & Cathepsins consolidated

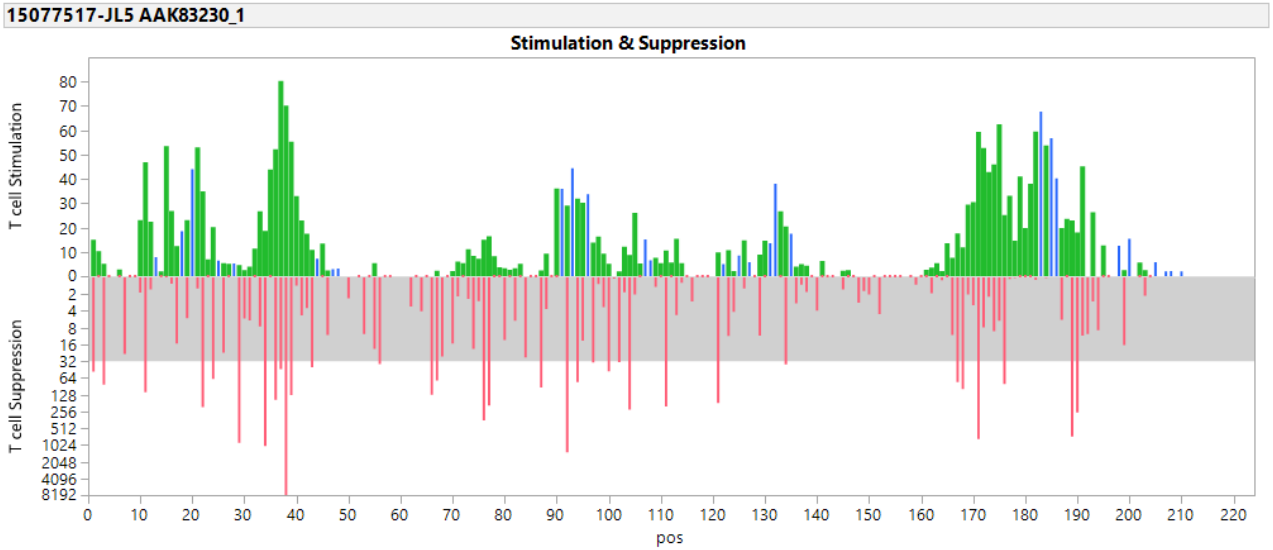



MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 15077517-JL5 AAK83230_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
7
14
21
28
35
42
49
56
63
70
0
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

15077517-JL5 AAK83230_1




15077517-JL5 AAK83230_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos


15077517-JL5 AAK83230_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

ioGenetics®

Rubella

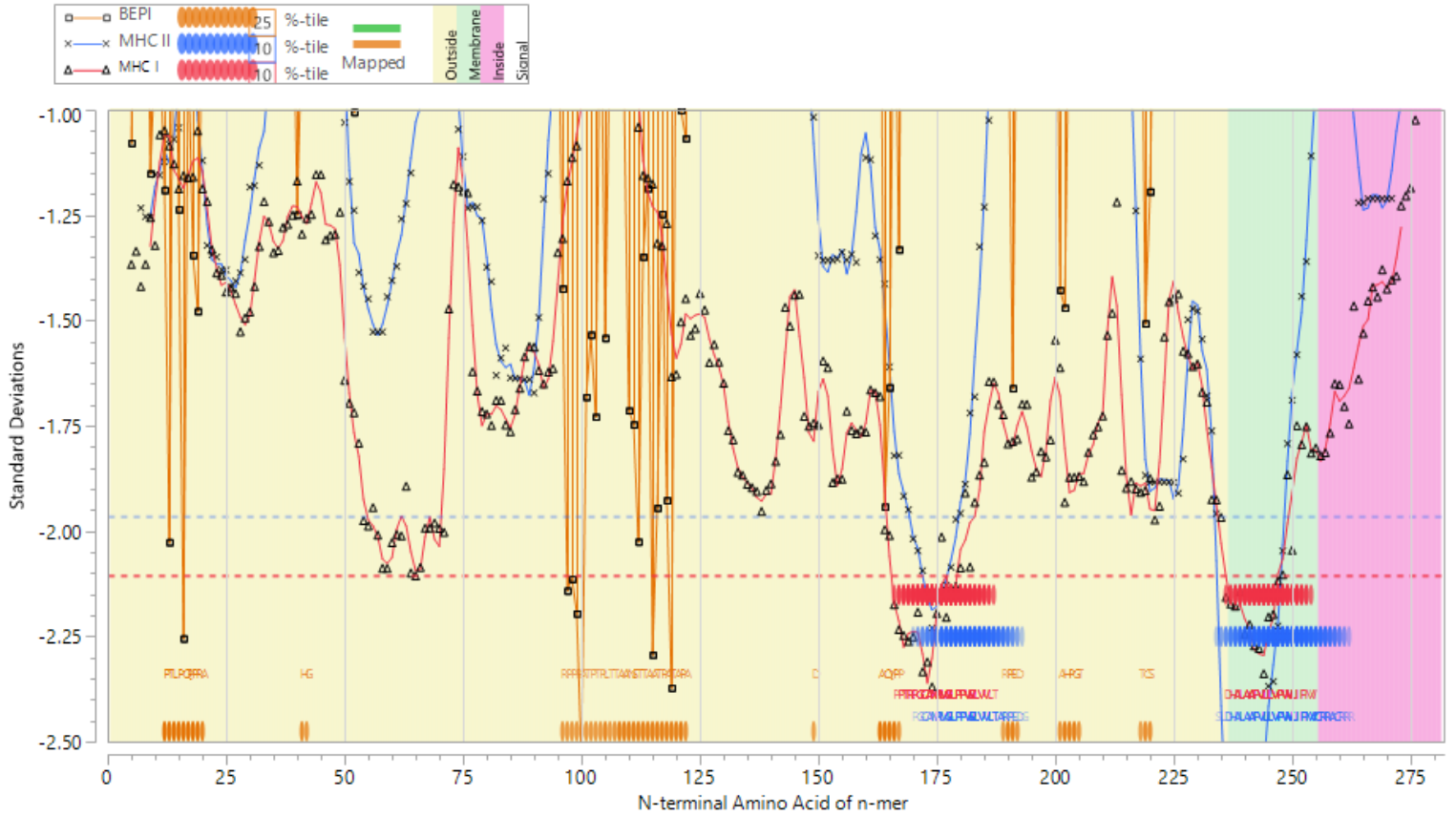
MERUVAX® II (Rubella Virus Vaccine Live), the Wistar RA 27/3 strain of live attenuated rubella virus propagated in WI-38 human diploid lung fibroblasts

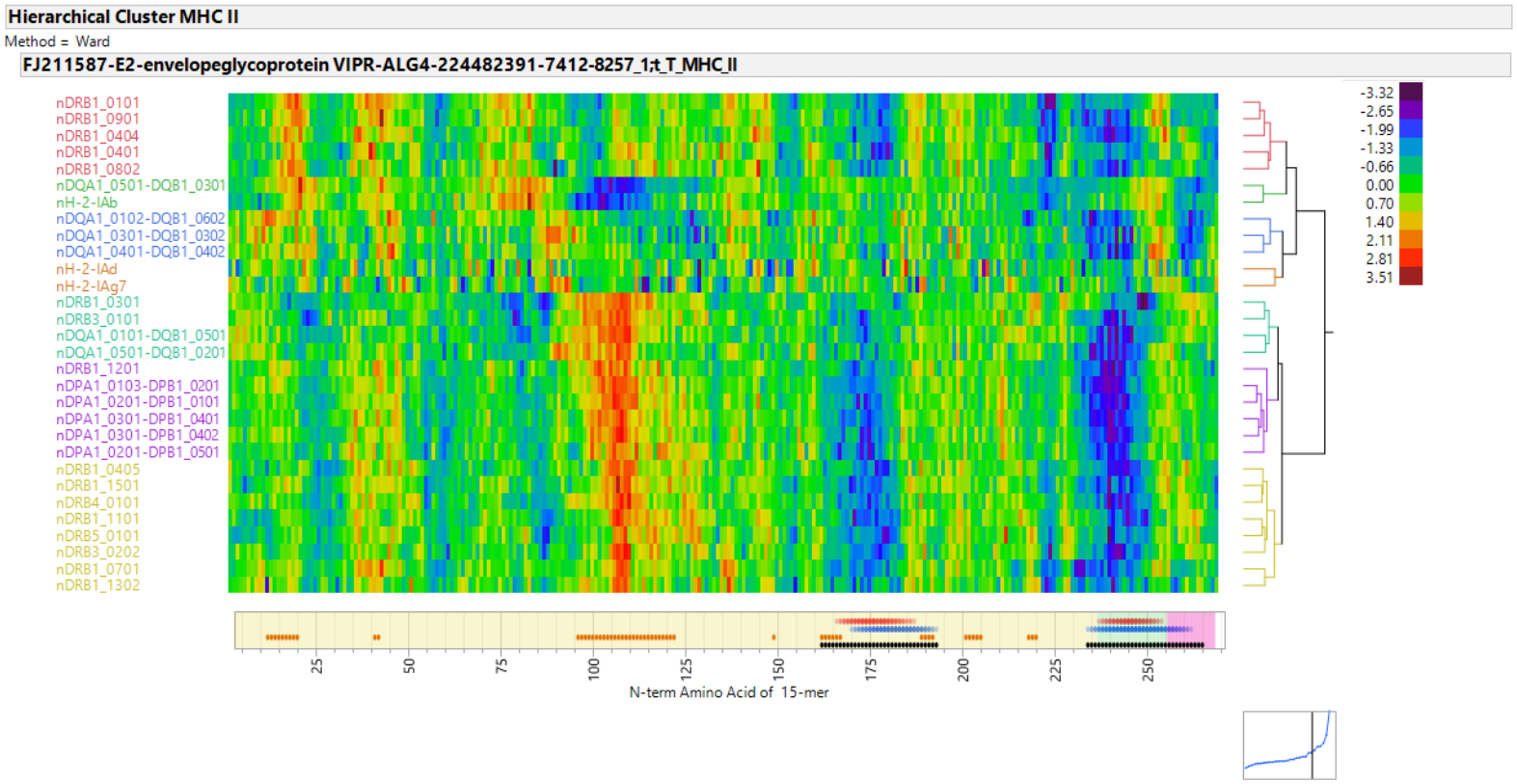
Based on proteins encoded from FJ211587

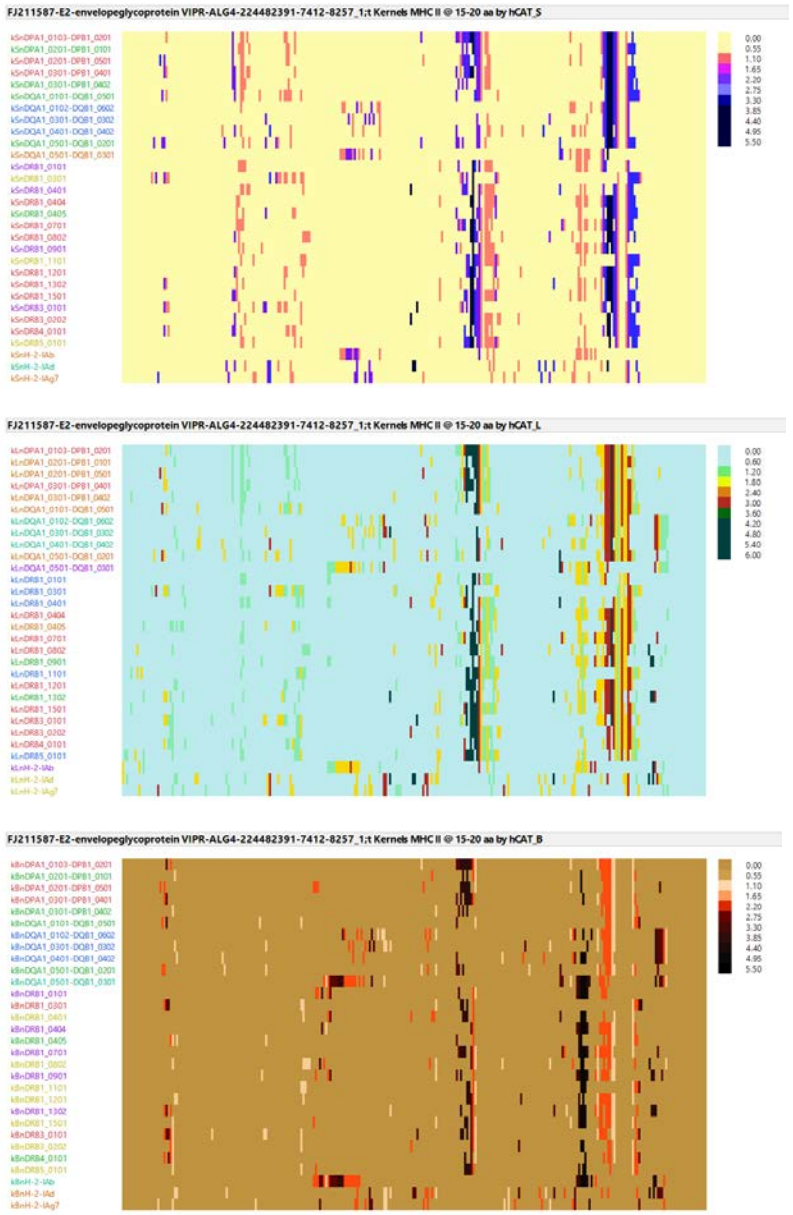
E2 Envelope glycoprotein

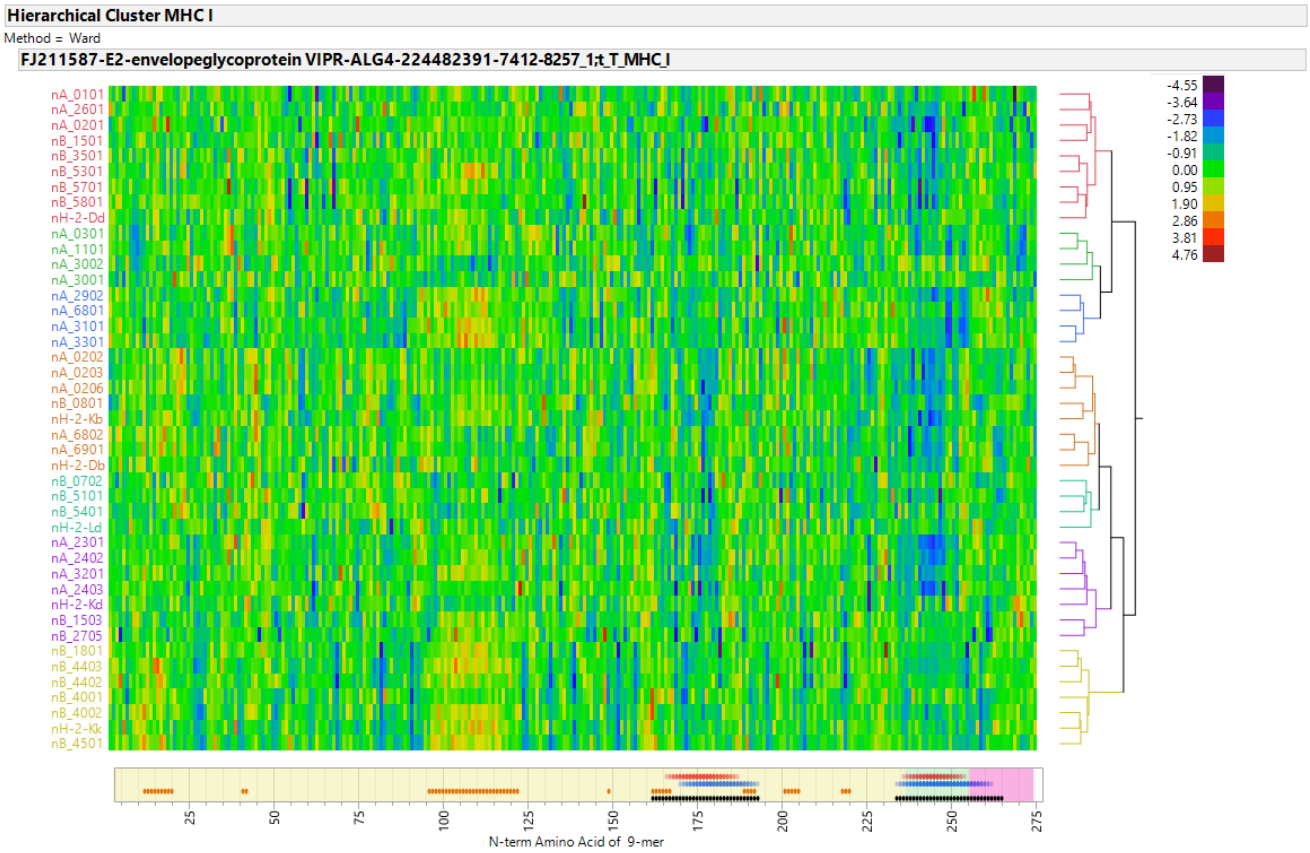
Overlay Plot

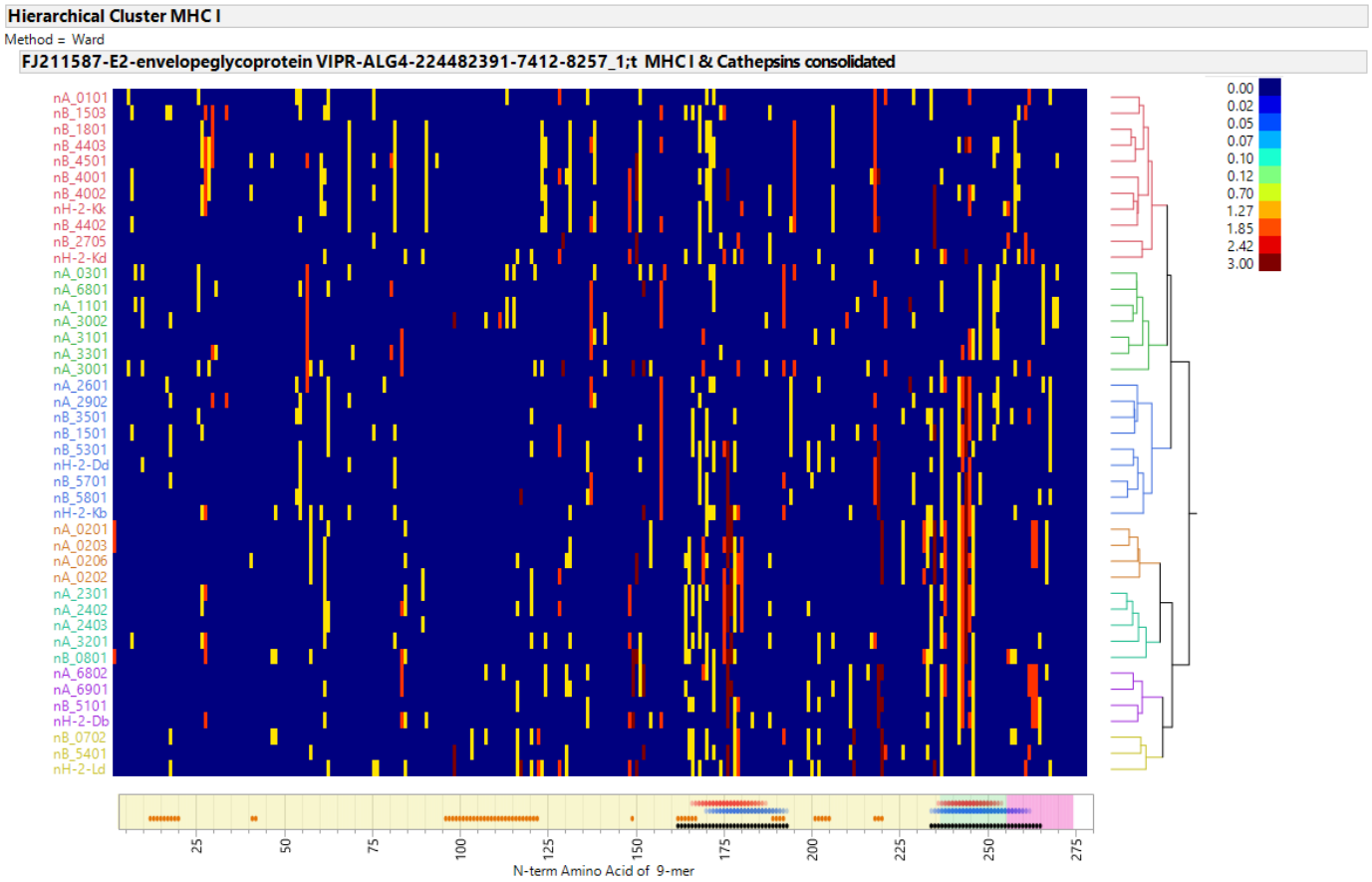
FJ211587-E2-envelopeglycoprotein
VIPR-ALG4-224482391-7412-8257_1;t

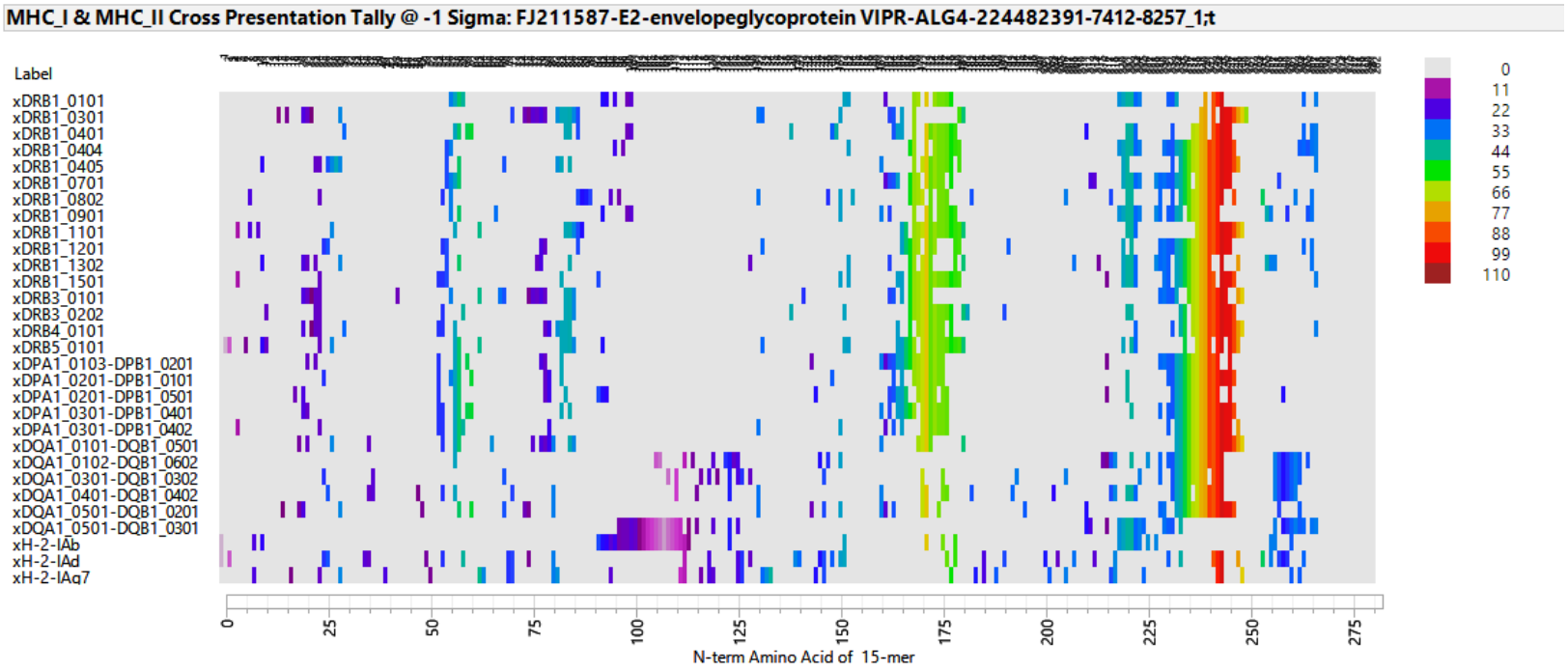


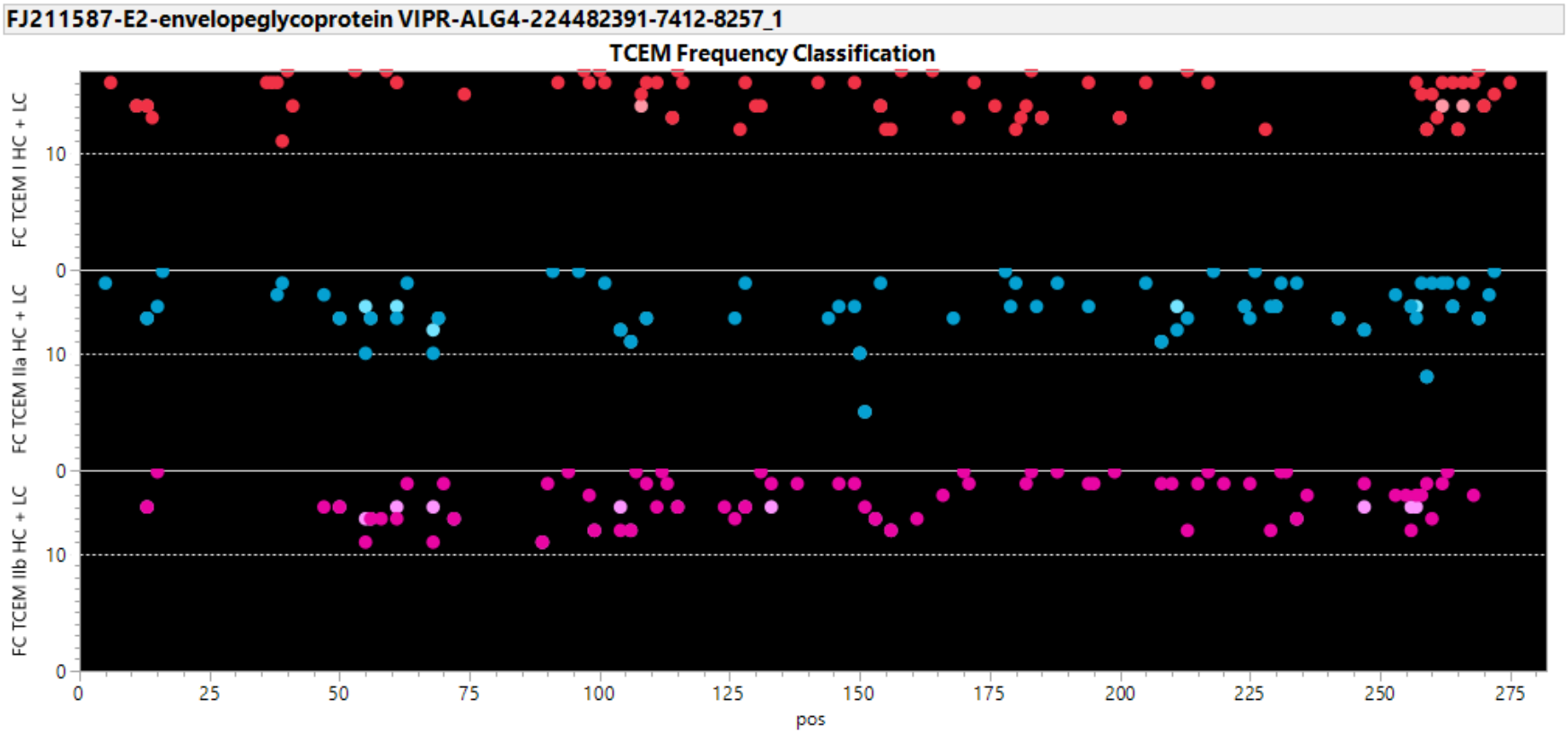

Hierarchical Cluster MHC II
Method = Ward
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0404
nDRB1_0401
nDRB1_0802
nDQA1_0501-DQB1_0301
nH-2-IAb
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDRB1_0301
nDRB3_0101
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDRB3_0202
nDRB1_0701
nDRB1_1302
-3.32
-2.65
-1.99
-1.33
-0.66
0.00
0.70
1.40
2.11
2.81
3.51
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 15-mer


FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t Kernels MHC II @ 15-20 aa by hCAT_S
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t Kernels MHC II @ 15-20 aa by hCAT_L
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t Kernels MHC II @ 15-20 aa by hCAT_B

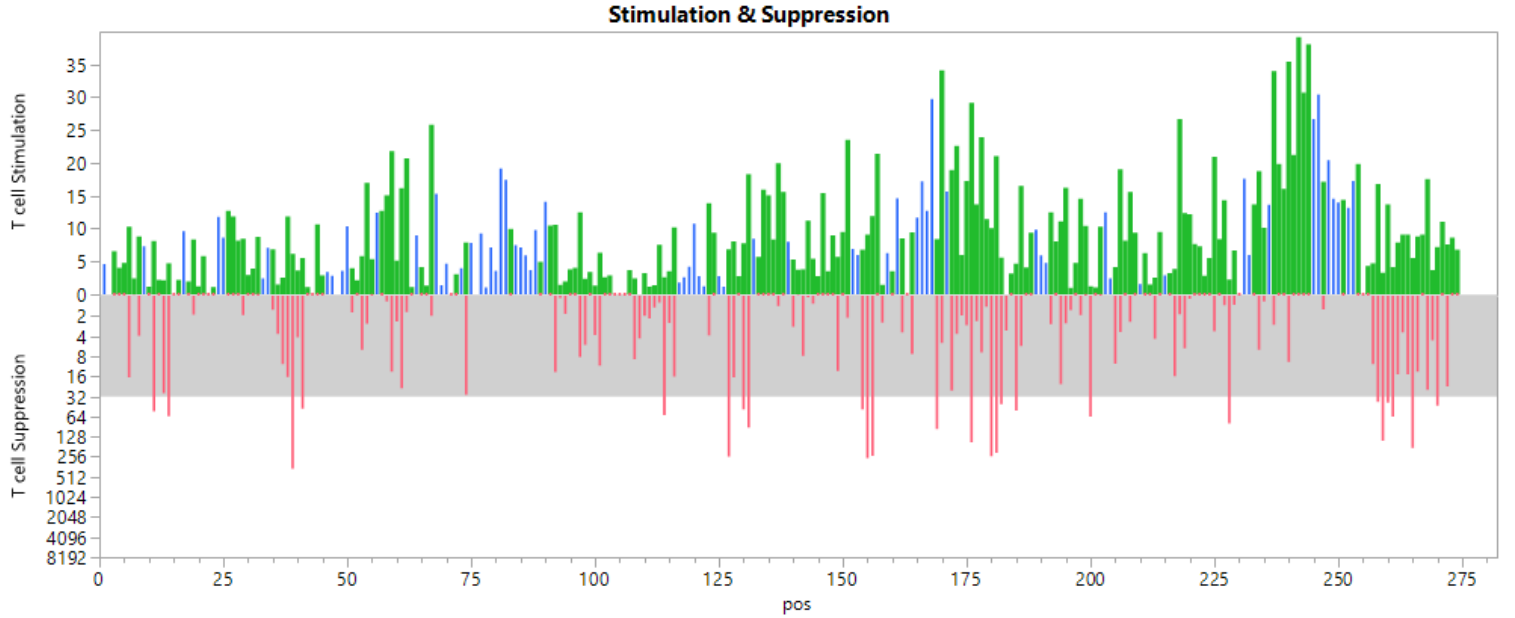

Hierarchical Cluster MHC I
Method = Ward
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t_T_MHC_I
nA_0101
nA_2601
nA_0201
nB_1501
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_3002
nA_3001
nA_2902
nA_6801
nA_3101
nA_3301
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_6802
nA_6901
nH-2-Db
nB_0702
nB_5101
nB_5401
nH-2-Ld
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
-4.55
-3.64
-2.73
-1.82
-0.91
0.00
0.95
1.90
2.86
3.81
4.76
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t MHC I & Cathepsins consolidated
nA_0101
nB_1503
nB_1801
nB_4403
nB_4501
nB_4001
nB_4002
nH-2-Kk
nB_4402
nB_2705
nH-2-Kd
nA_0301
nA_6801
nA_1101
nA_3002
nA_3101
nA_3301
nA_3001
nA_2601
nA_2902
nB_3501
nB_1501
nB_5301
nH-2-Dd
nB_5701
nB_5801
nH-2-Kb
nA_0201
nA_0203
nA_0206
nA_0202
nA_2301
nA_2402
nA_2403
nA_3201
nB_0801
nA_6802
nA_6901
nB_5101
nH-2-Db
nB_0702
nB_5401
nH-2-Ld
0.00
0.02
0.05
0.07
0.10
0.12
0.70
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer

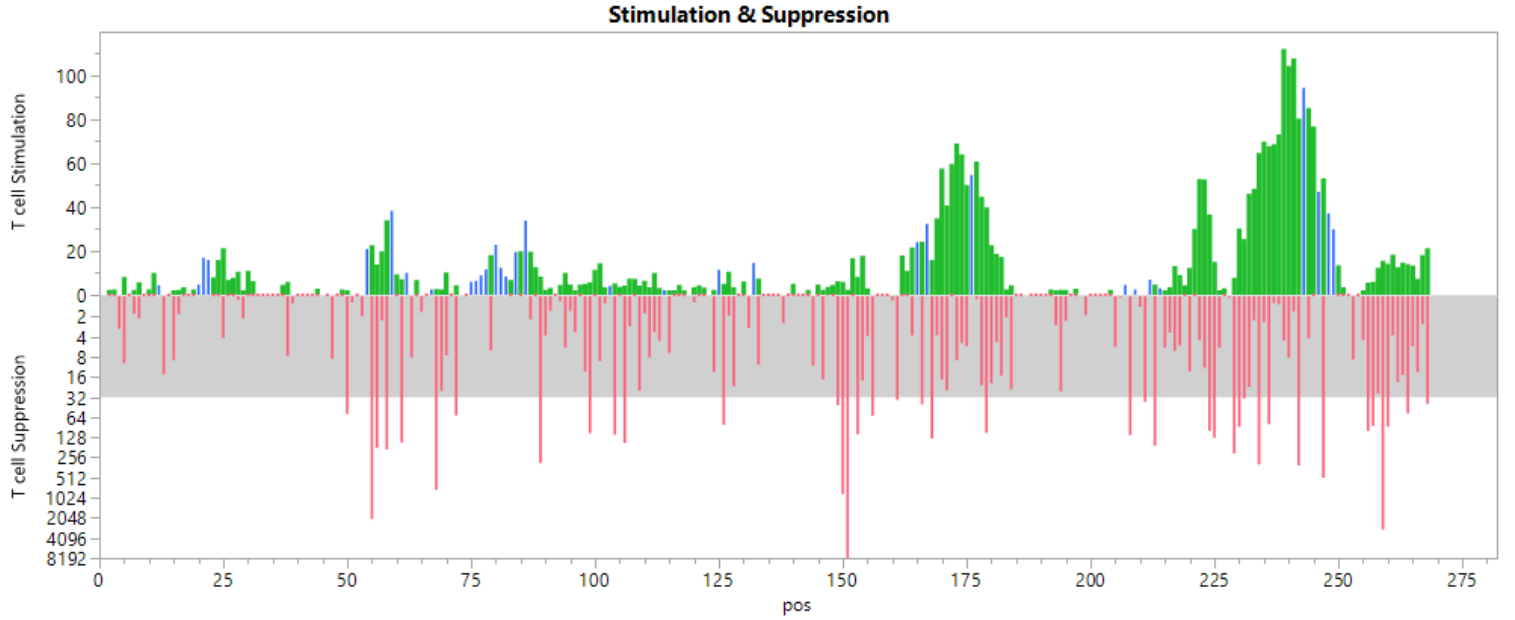

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer
0
11
22
33
44
55
66
77
88
99
110


FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0
25
50
75
100
125
150
175
200
225
250
275
pos

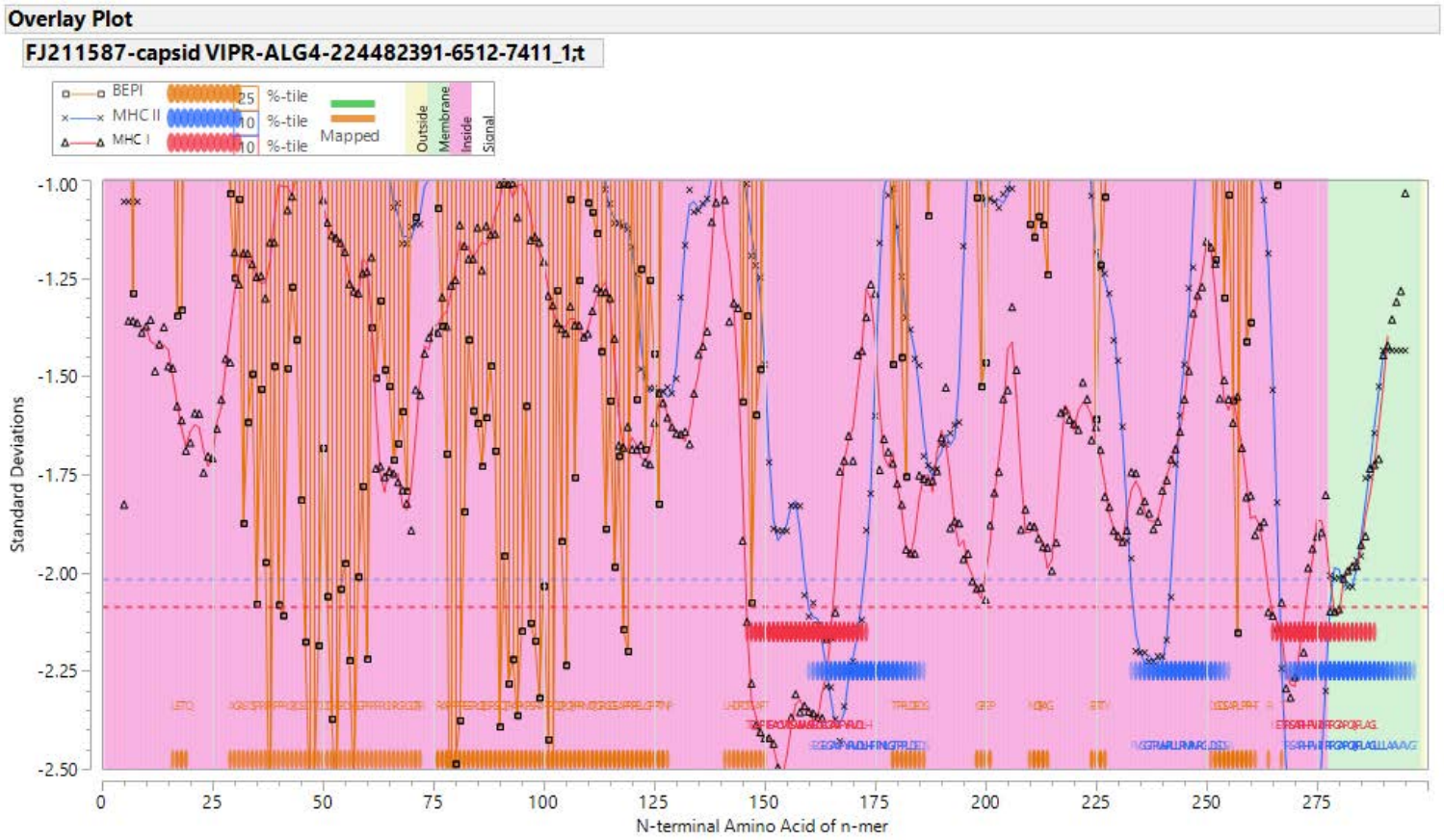
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1

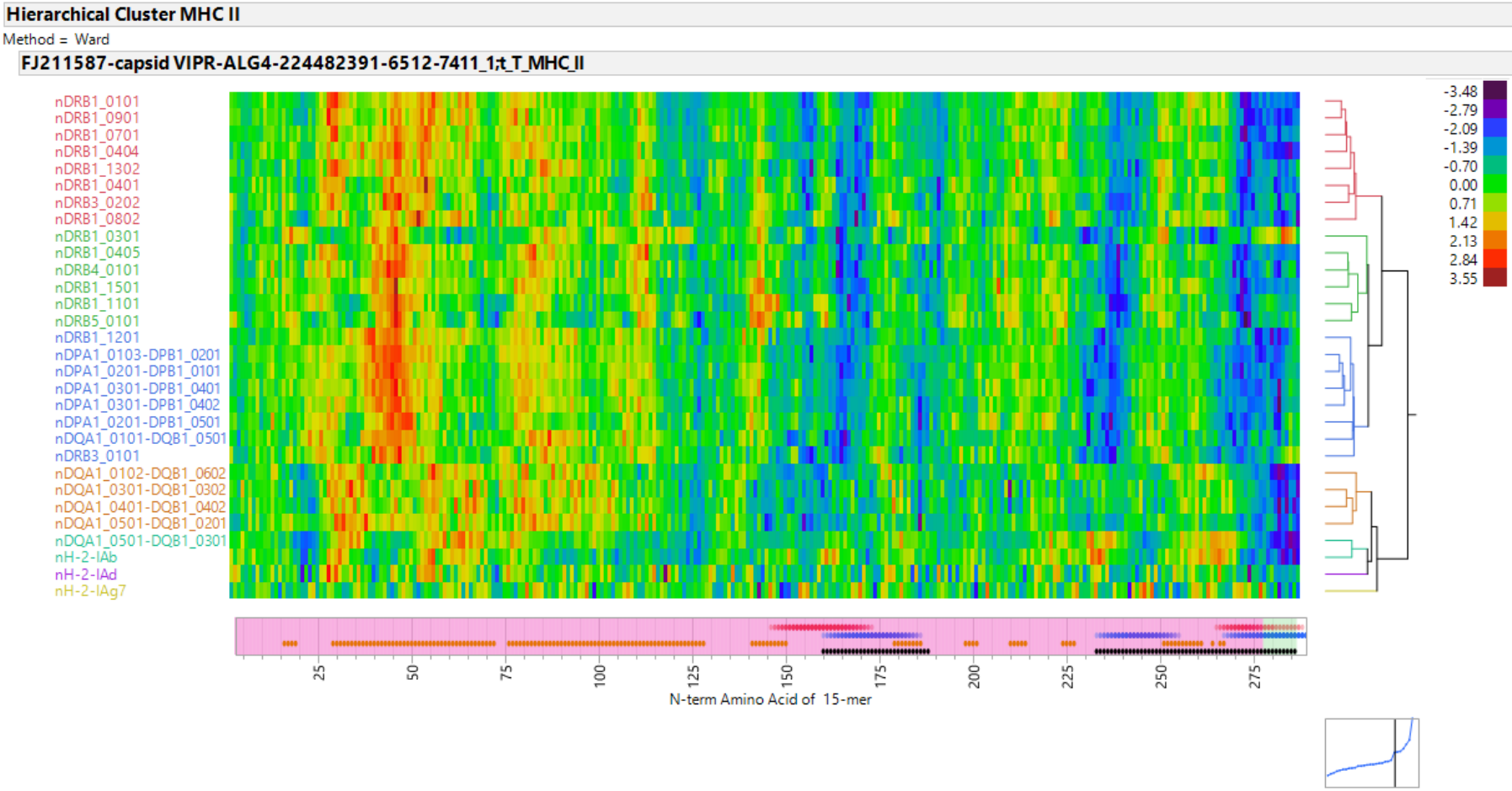


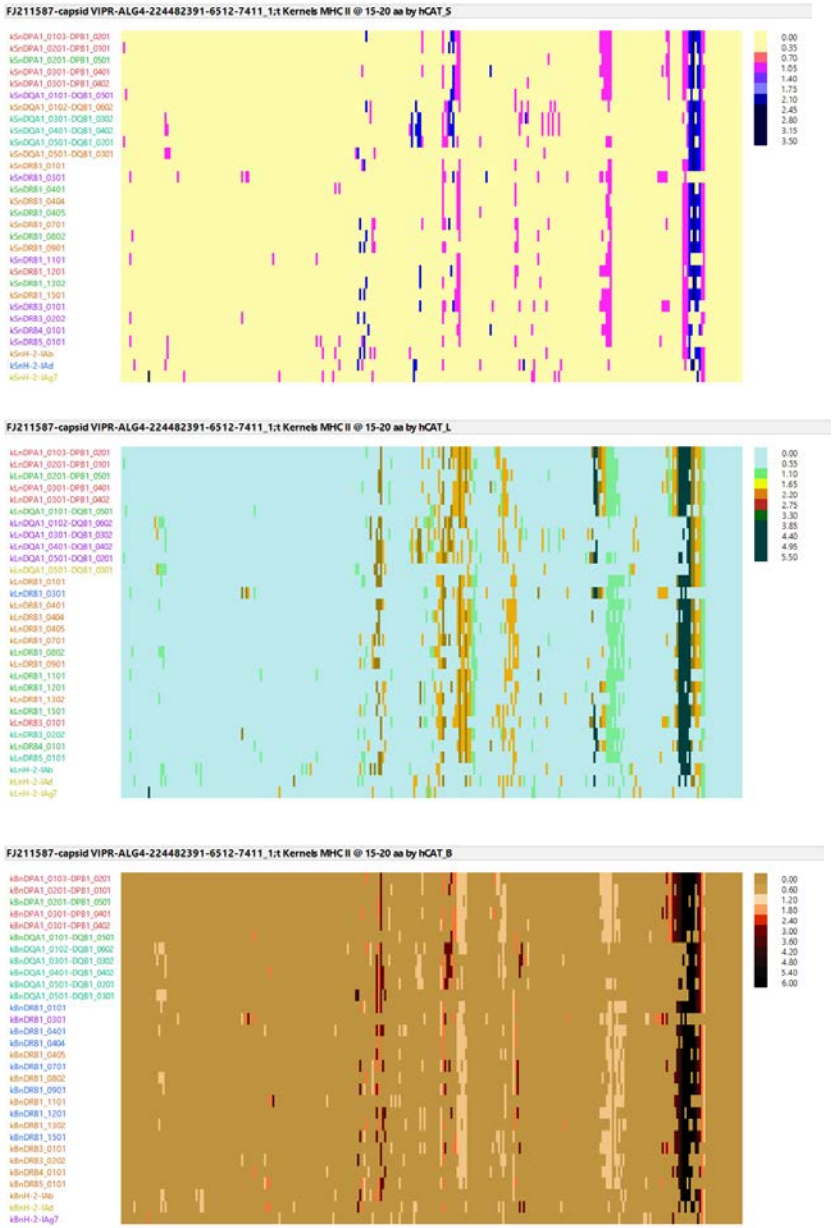
FJ211587-E2-envelopeglycoprotein VIPR-ALG4-224482391-7412-8257_1

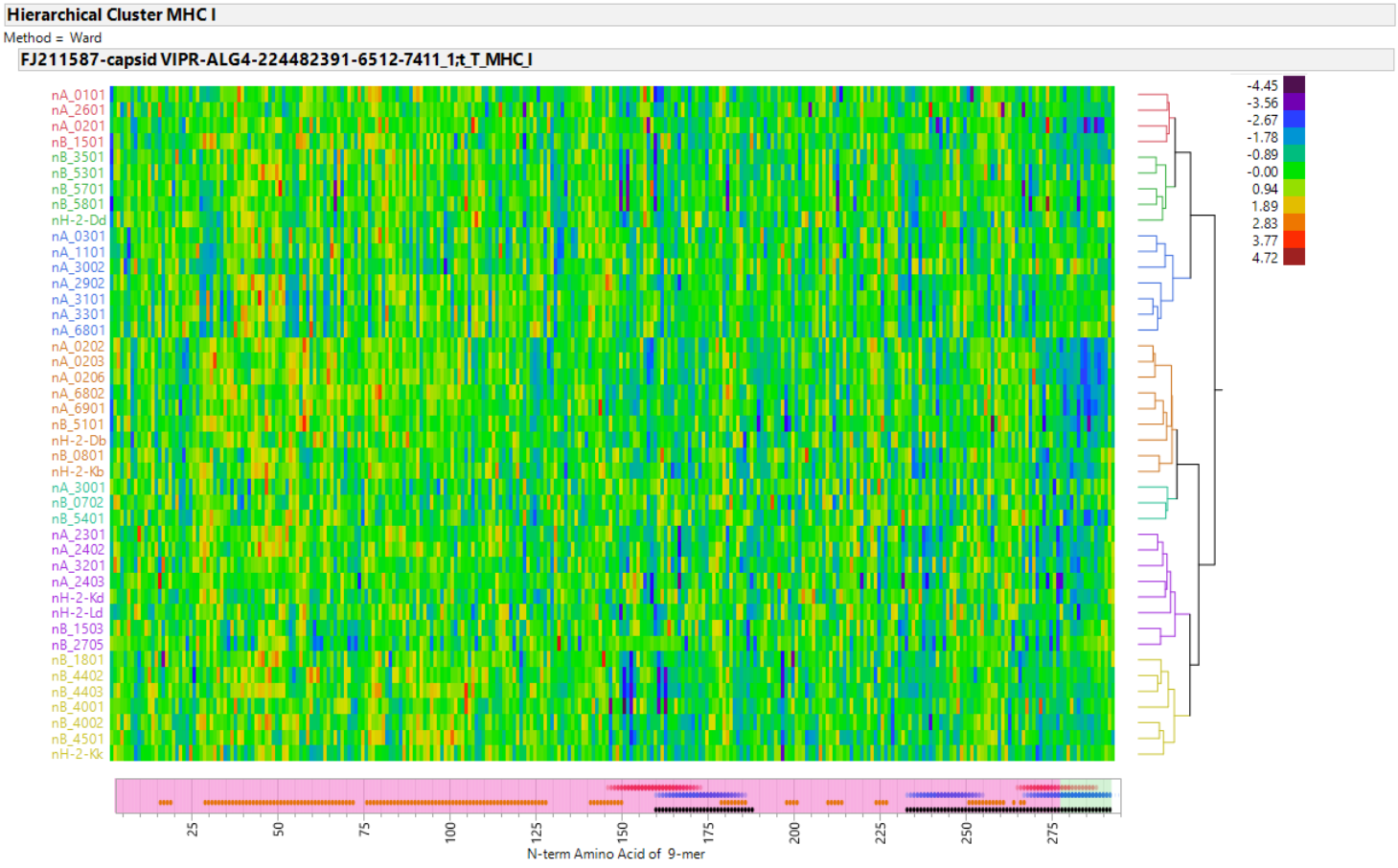


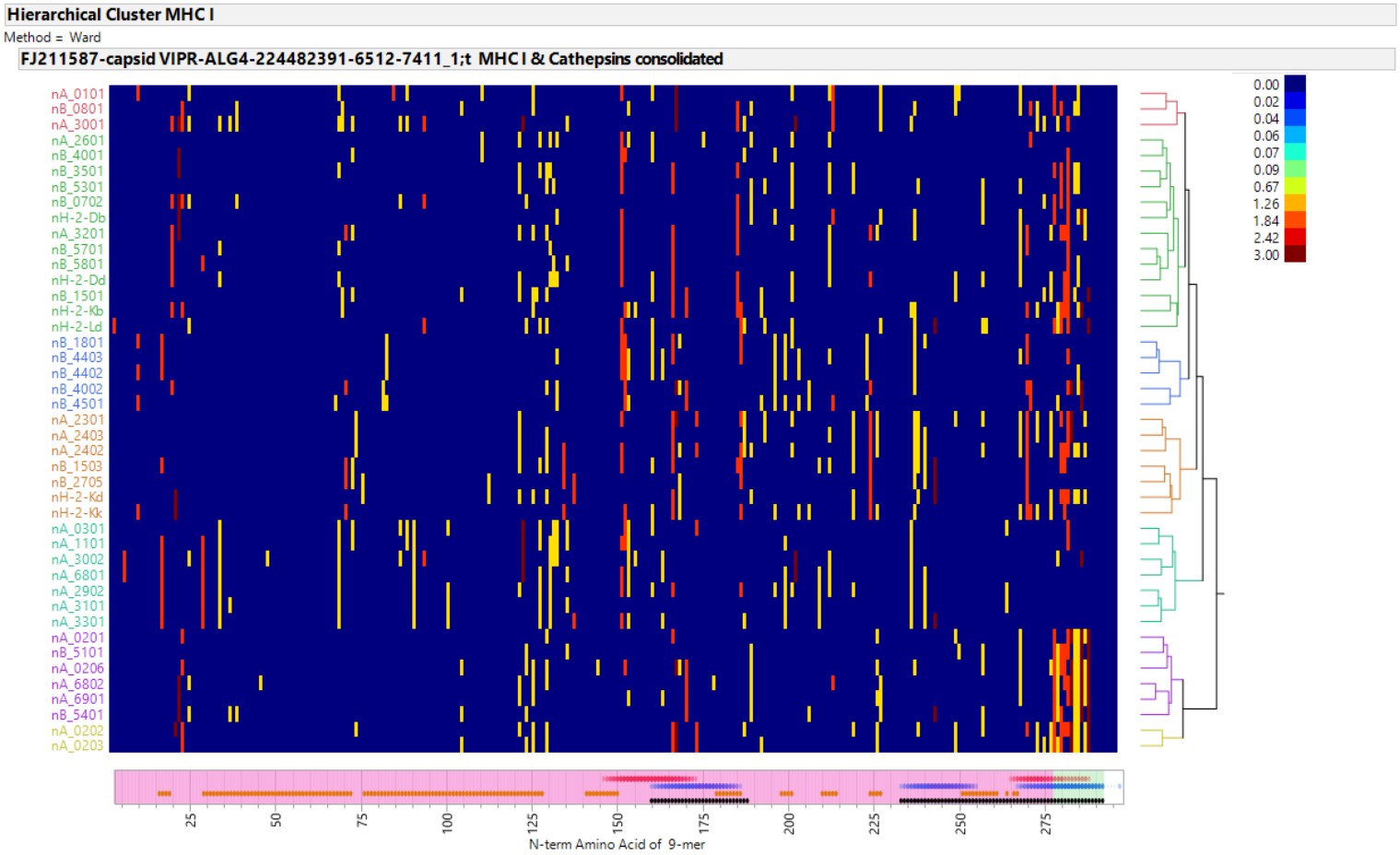
Capsid

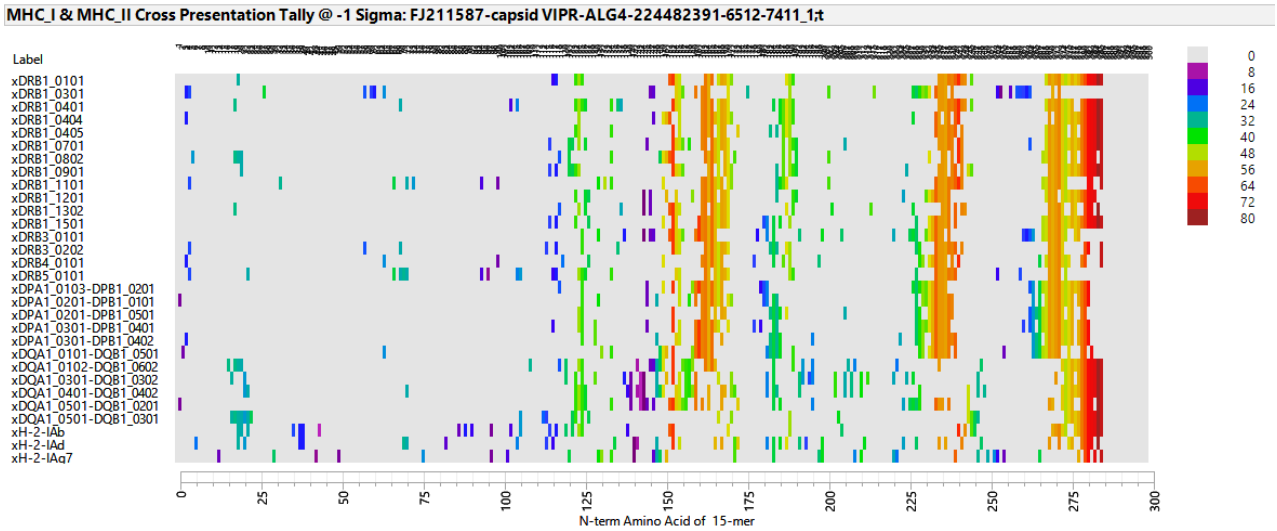

Overlay Plot
FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
275
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_0301
nDRB1_0405
nDRB4_0101
nDRB1_1501
nDRB1_1101
nDRB5_0101
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDRB3_0101
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.48
-2.79
-2.09
-1.39
-0.70
0.00
0.71
1.42
2.13
2.84
3.55
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer

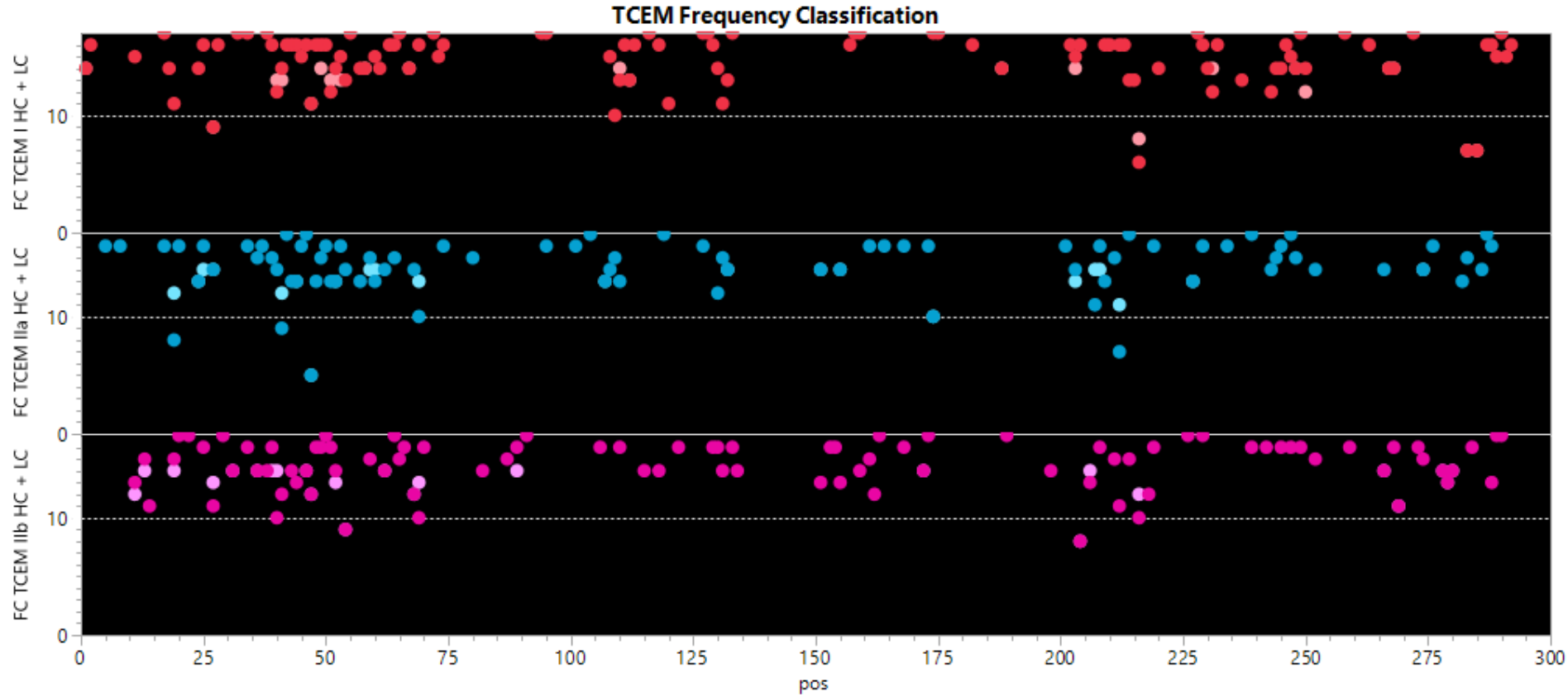

FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1:t Kernels MHC II @ 15-20 aa by hCAT_S
FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1:t Kernels MHC II @ 15-20 aa by hCAT_L
FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1:t Kernels MHC II @ 15-20 aa by hCAT_B

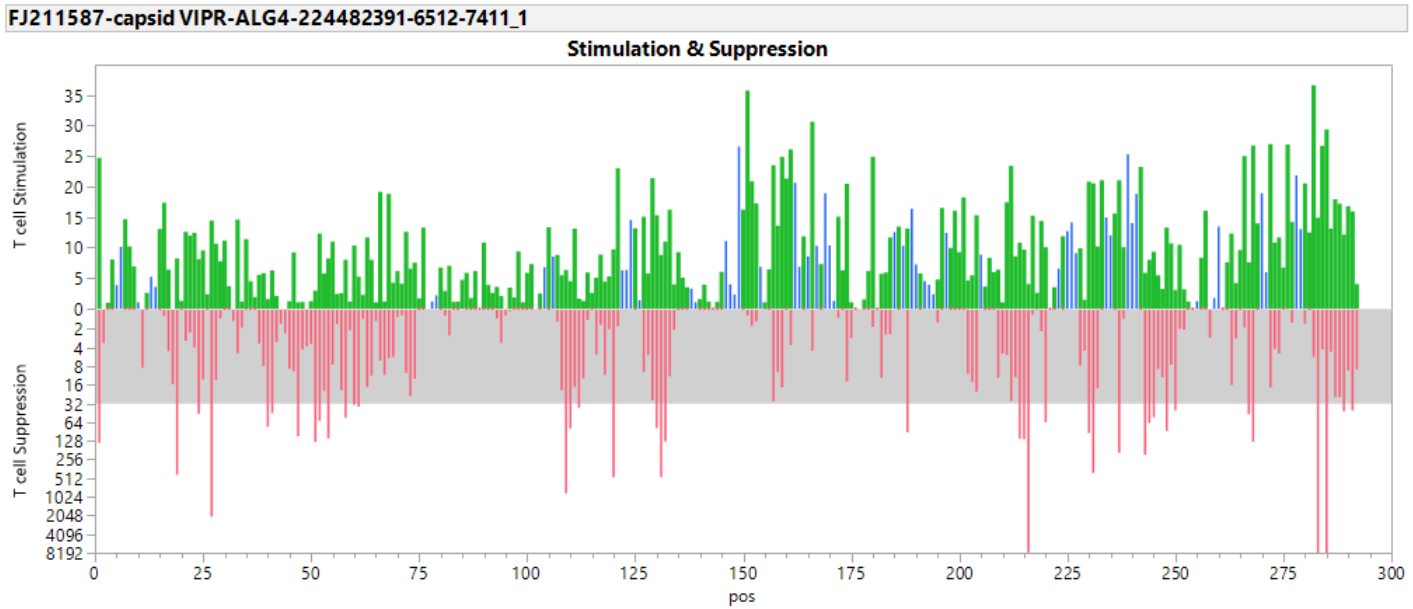

Hierarchical Cluster MHC I
Method = Ward
FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1;t_T_MHC_I
nA_0101
nA_2601
nA_0201
nB_1501
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nH-2-Db
nB_0801
nH-2-Kb
nA_3001
nB_0702
nB_5401
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nH-2-Ld
nB_1503
nB_2705
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nH-2-Kk
-4.45
-3.56
-2.67
-1.78
-0.89
-0.00
0.94
1.89
2.83
3.77
4.72
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer

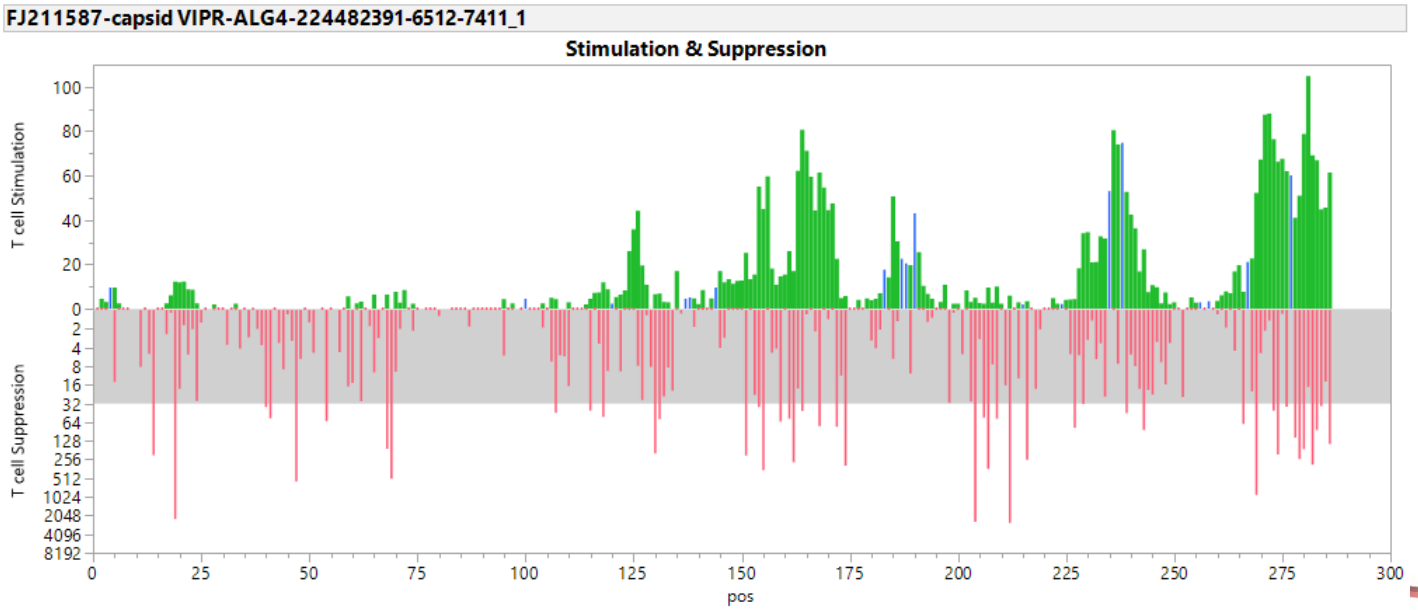

Hierarchical Cluster MHC I
Method = Ward
FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1;t MHC I & Cathepsins consolidated
nA_0101
nB_0801
nA_3001
nA_2601
nB_4001
nB_3501
nB_5301
nB_0702
nH-2-Db
nA_3201
nB_5701
nB_5801
nH-2-Dd
nB_1501
nH-2-Kb
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4002
nB_4501
nA_2301
nA_2403
nA_2402
nB_1503
nB_2705
nH-2-Kd
nH-2-Kk
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
nA_0201
nB_5101
nA_0206
nA_6802
nA_6901
nB_5401
nA_0202
nA_0203
0.00
0.02
0.04
0.06
0.07
0.09
0.67
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
25
50
75
100
125
150
175
200
225
250
275
300
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80

FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1



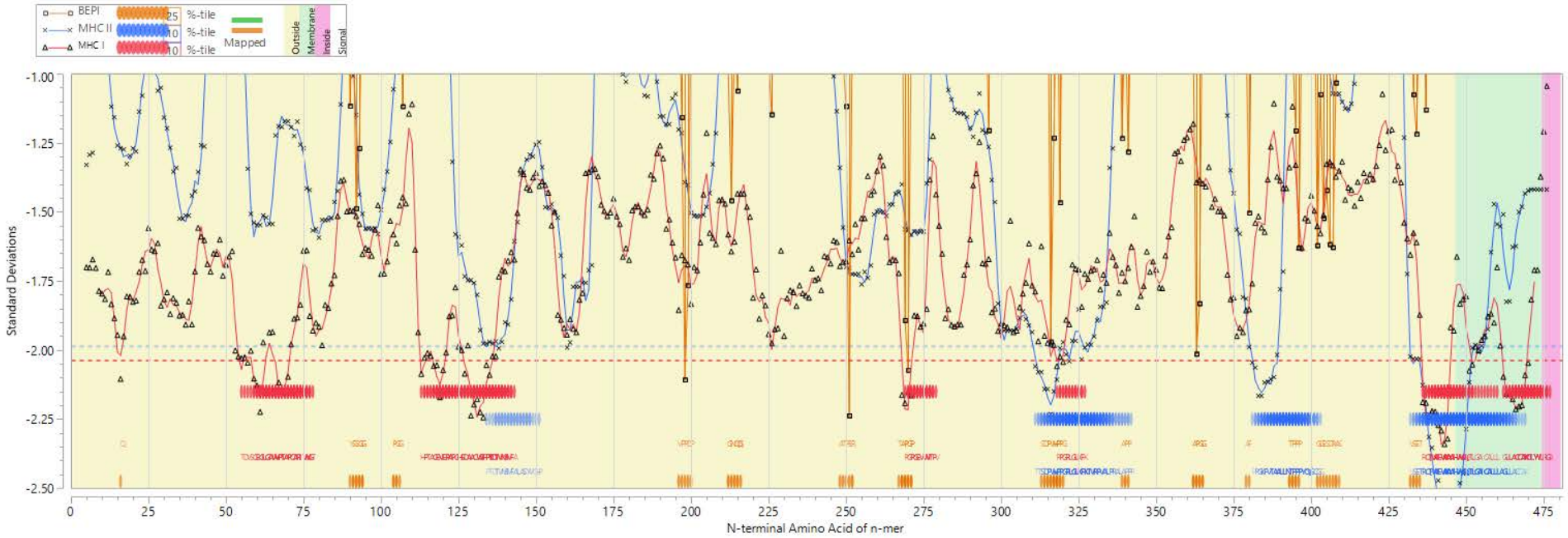

FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

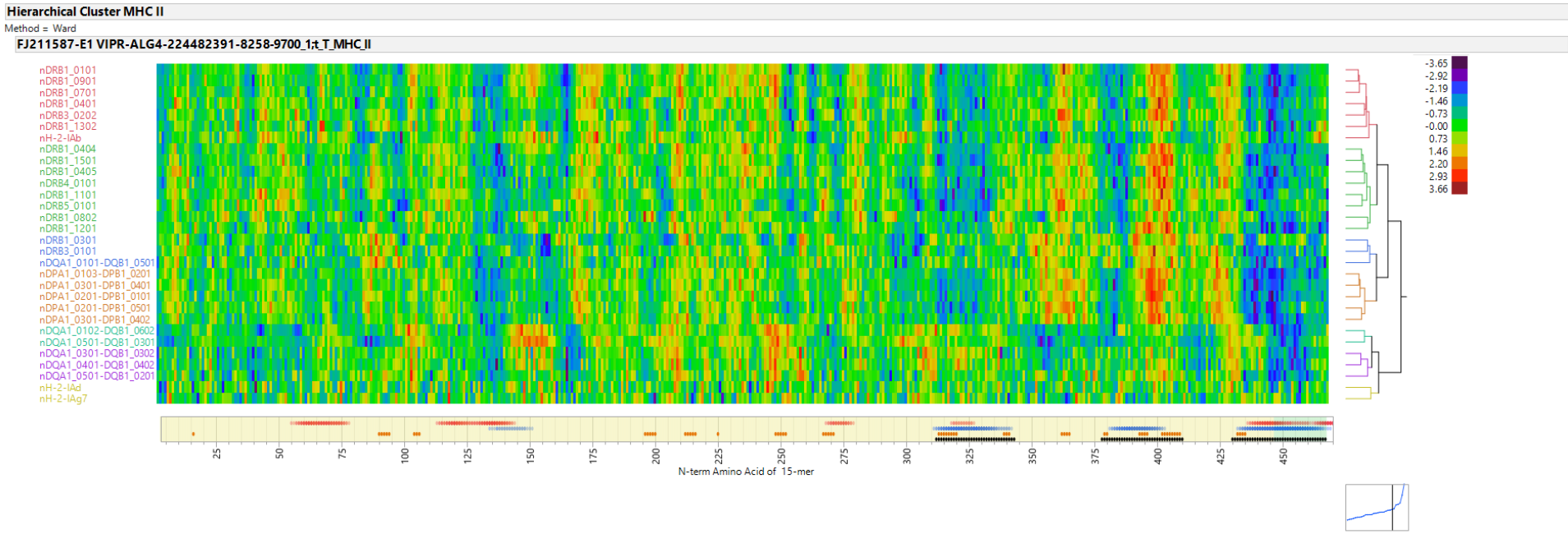

FJ211587-capsid VIPR-ALG4-224482391-6512-7411_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

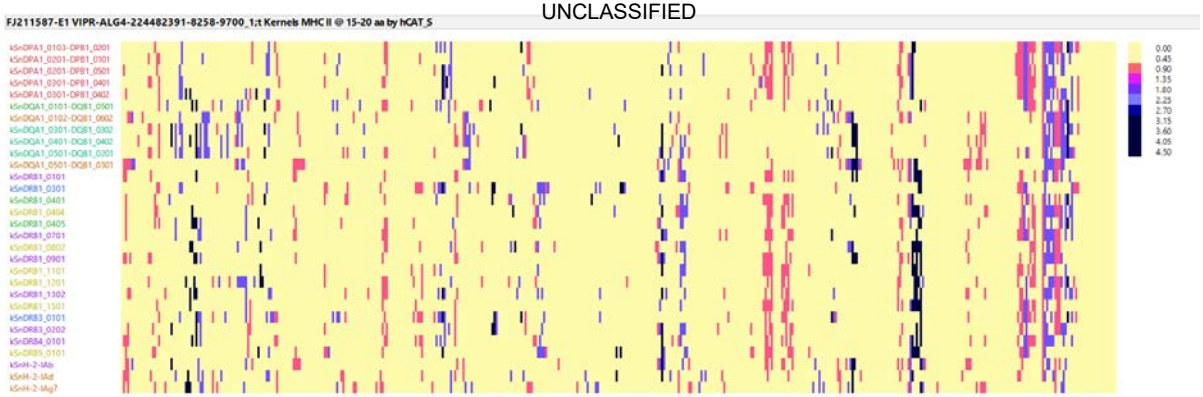
E1 Glycoprotein

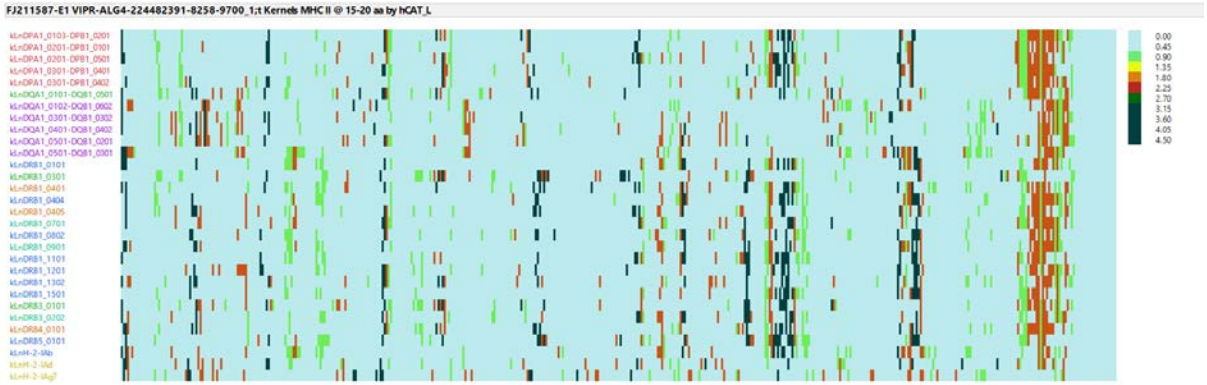
Overlay Plot

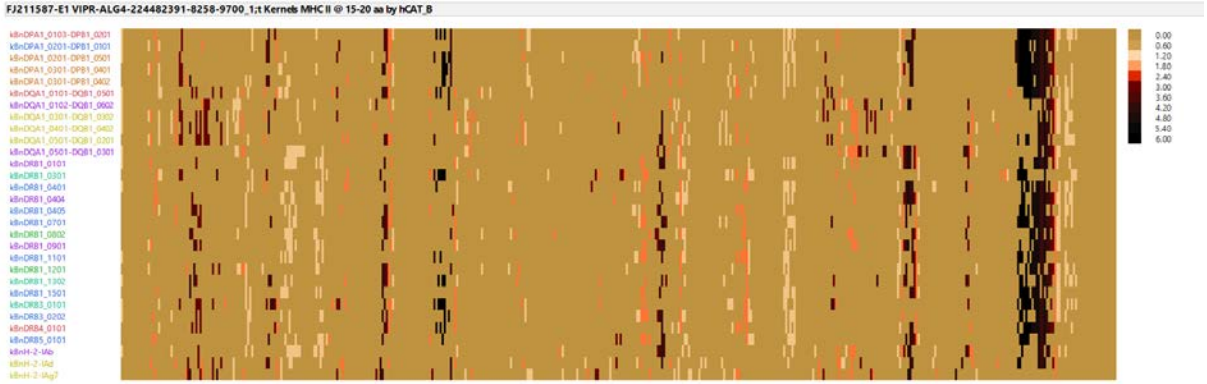
FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t

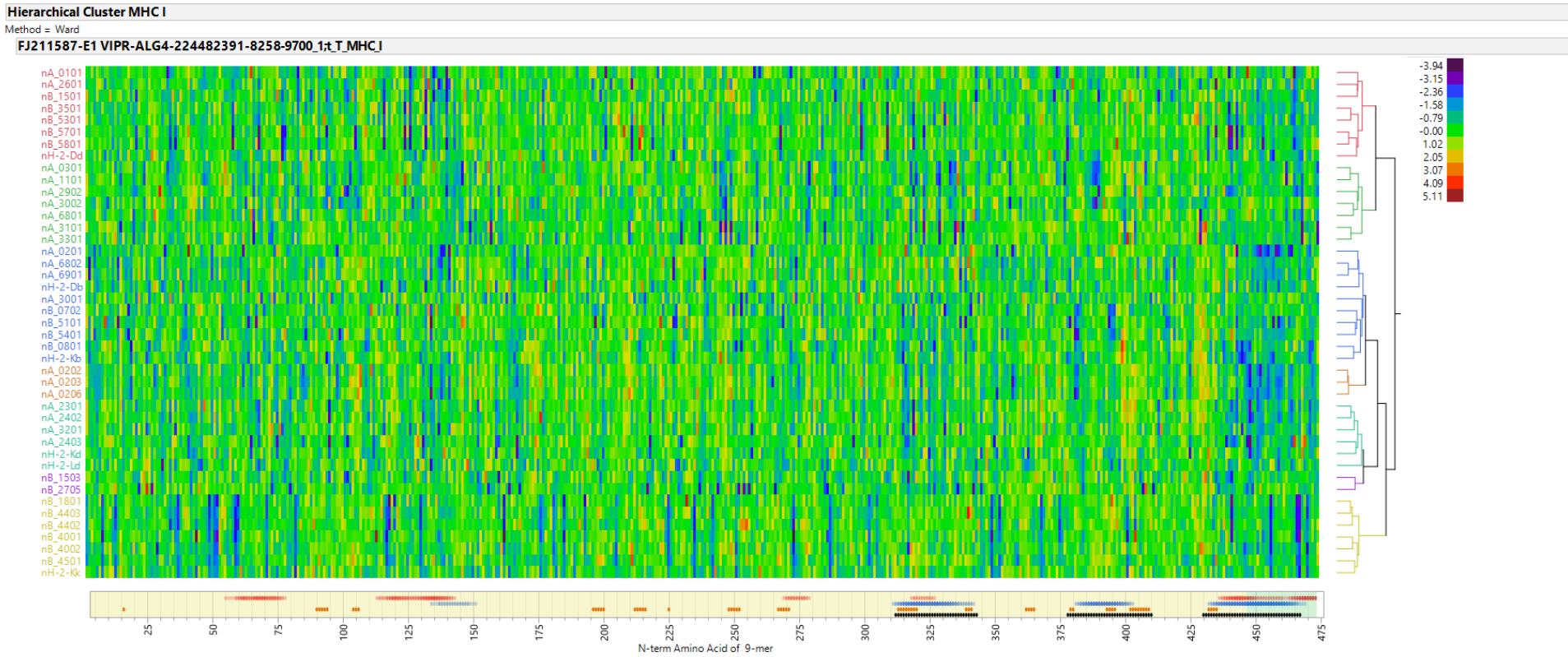


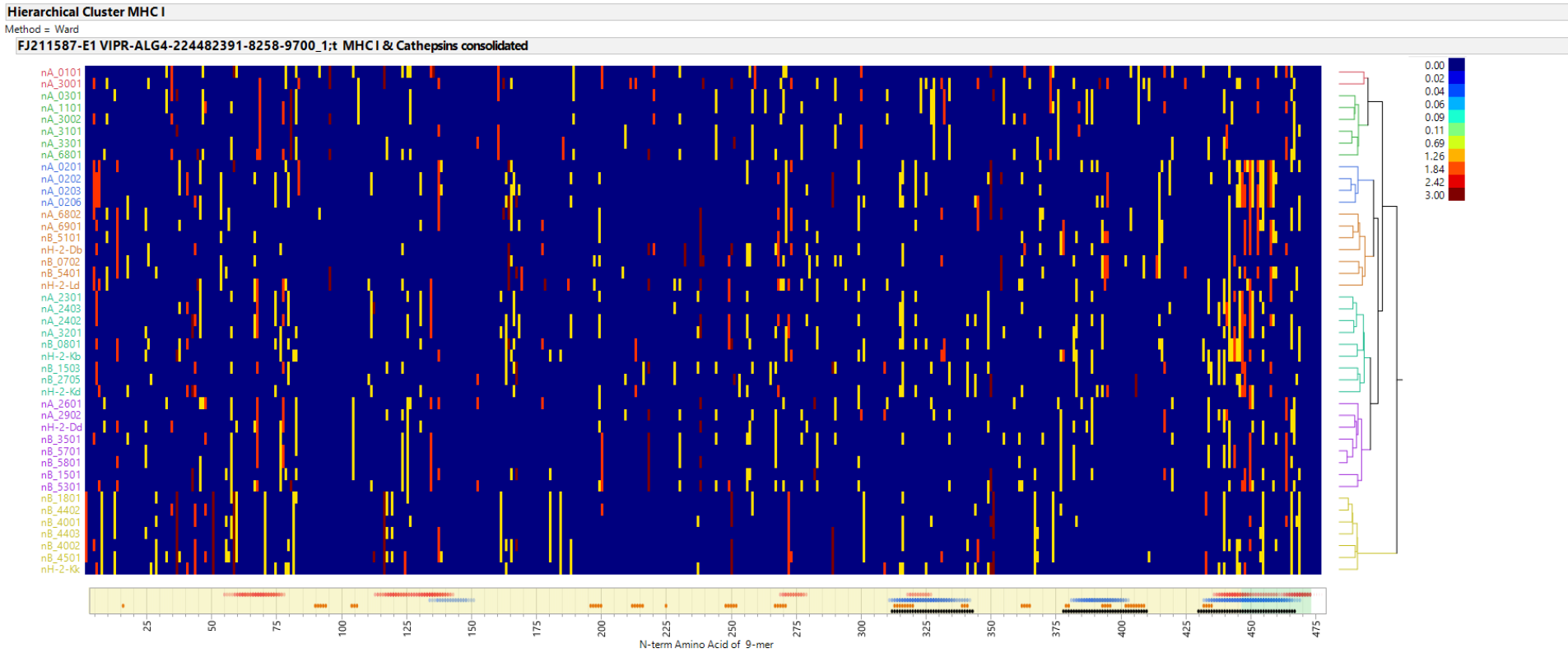

Hierarchical Cluster MHC II
Method = Ward
FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0401
nDRB3_0202
nDRB1_1302
nH-2-IAb
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_1201
nDRB1_0301
nDRB3_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0102-DQB1_0602
nDQA1_0501-DQB1_0301
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
-3.65
-2.92
-2.19
-1.46
-0.73
-0.00
0.73
1.46
2.20
2.93
3.66
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
N-term Amino Acid of 15-mer

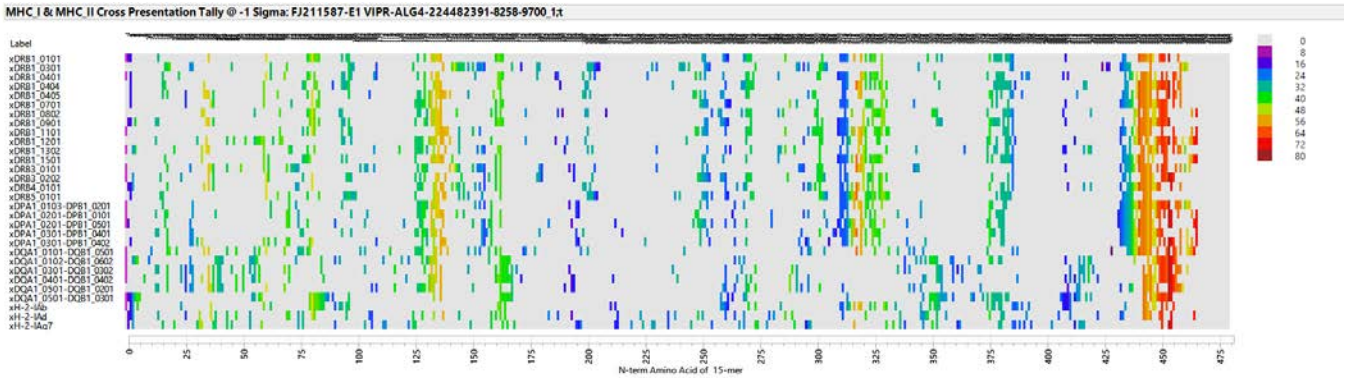

FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t Kernels MHC II @ 15-20 aa by hCAT_S

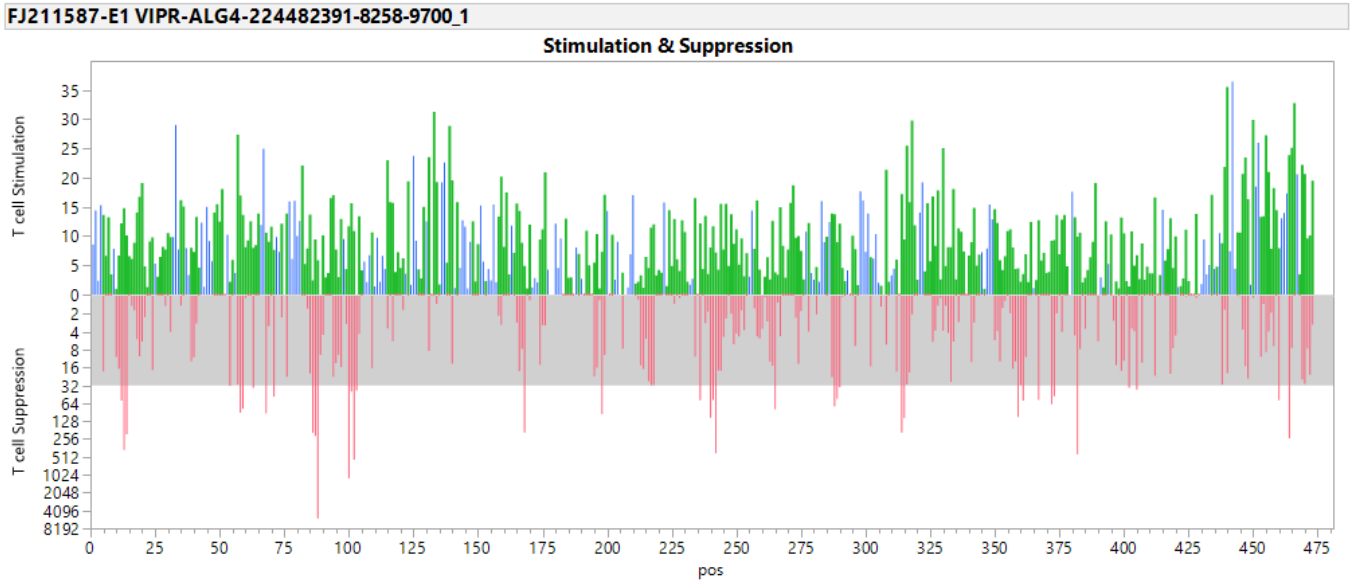

FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t Kernels MHC II @ 15-20 aa by hCAT_L

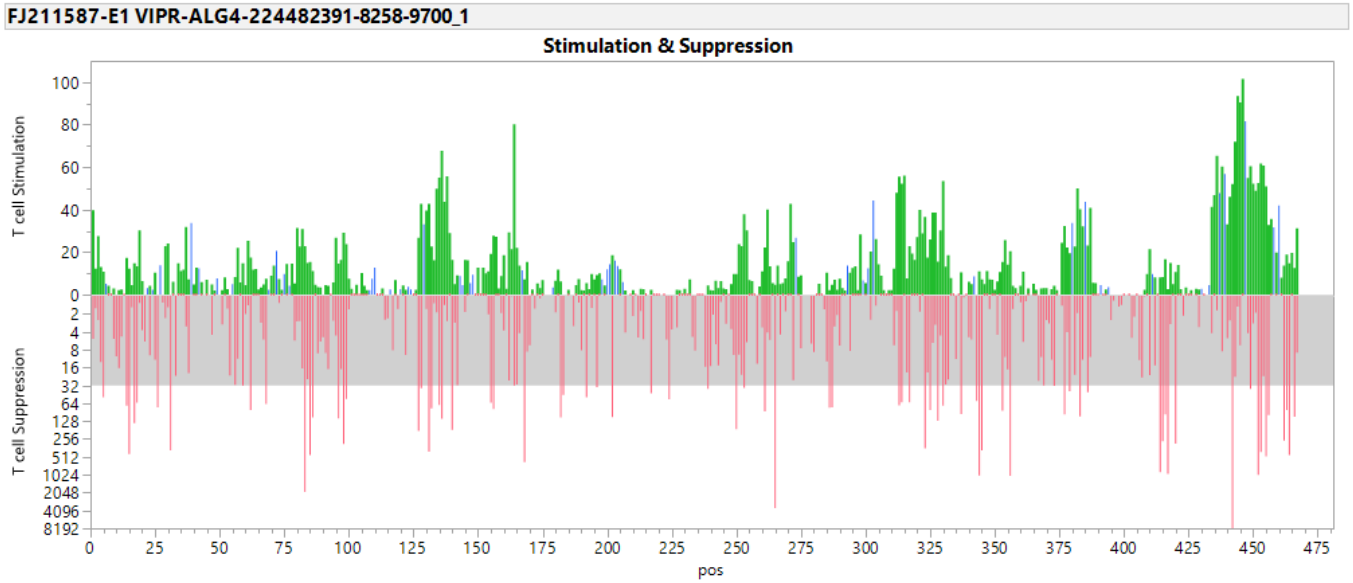

FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t_T_MHC_I
nA_0101
nA_2601
nB_1501
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_2902
nA_3002
nA_6801
nA_3101
nA_3301
nA_0201
nA_6802
nA_6901
nH-2-Db
nA_3001
nB_0702
nB_5101
nB_5401
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nH-2-Ld
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
-3.94
-3.15
-2.36
-1.58
-0.79
-0.00
1.02
2.05
3.07
4.09
5.11
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

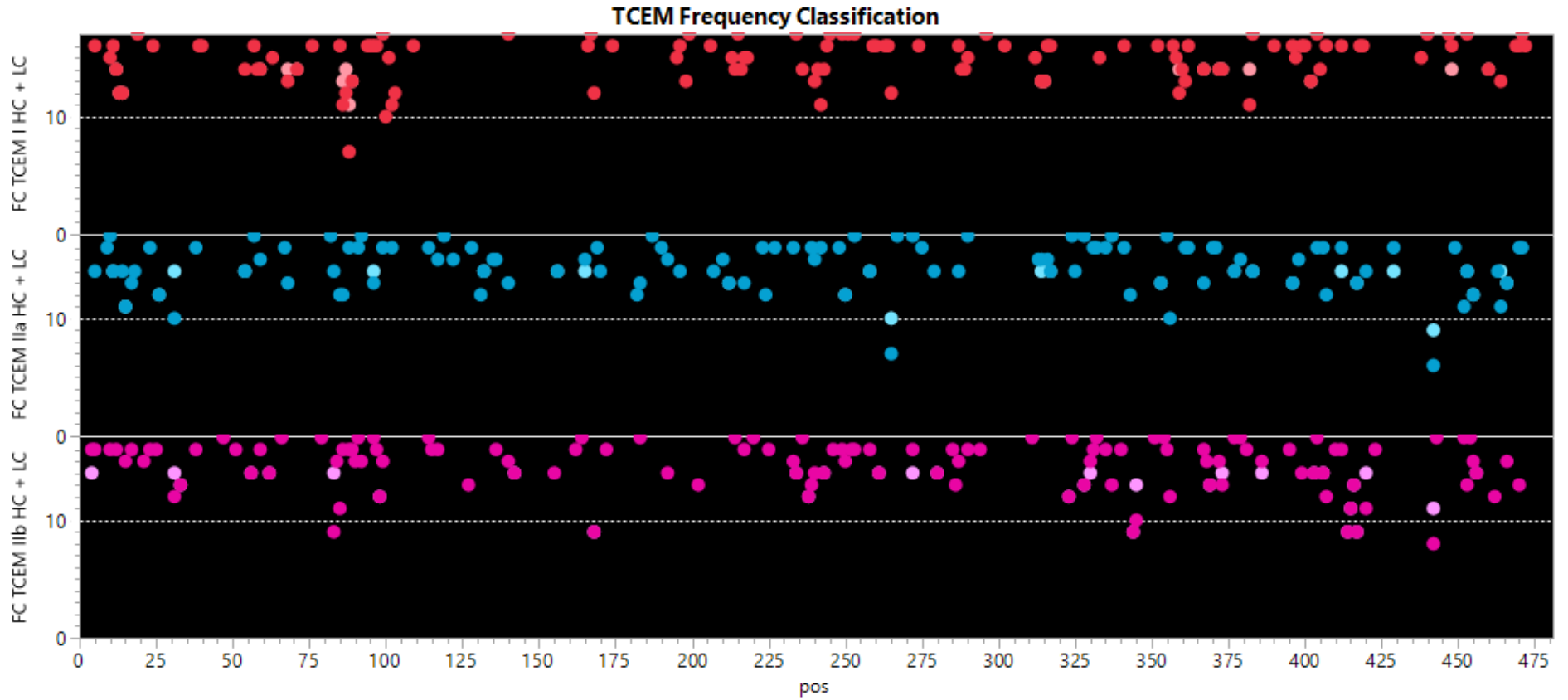

Hierarchical Cluster MHC I
Method = Ward
FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t MHC I & Cathepsins consolidated
nA_0101
nA_3001
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nH-2-Db
nB_0702
nB_5401
nH-2-Ld
nA_2301
nA_2403
nA_2402
nA_3201
nB_0801
nH-2-Kb
nB_1503
nB_2705
nH-2-Kd
nA_2601
nA_2902
nH-2-Dd
nB_3501
nB_5701
nB_5801
nB_1501
nB_5301
nB_1801
nB_4402
nB_4001
nB_4403
nB_4002
nB_4501
nH-2-Kk
0.00
0.02
0.04
0.06
0.09
0.11
0.69
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer


MHC I & MHC II Cross Presentation Tally @ -1 Sigma: FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAd7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer
0
8
16
24
32
40
48
56
64
72
80


FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos


FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

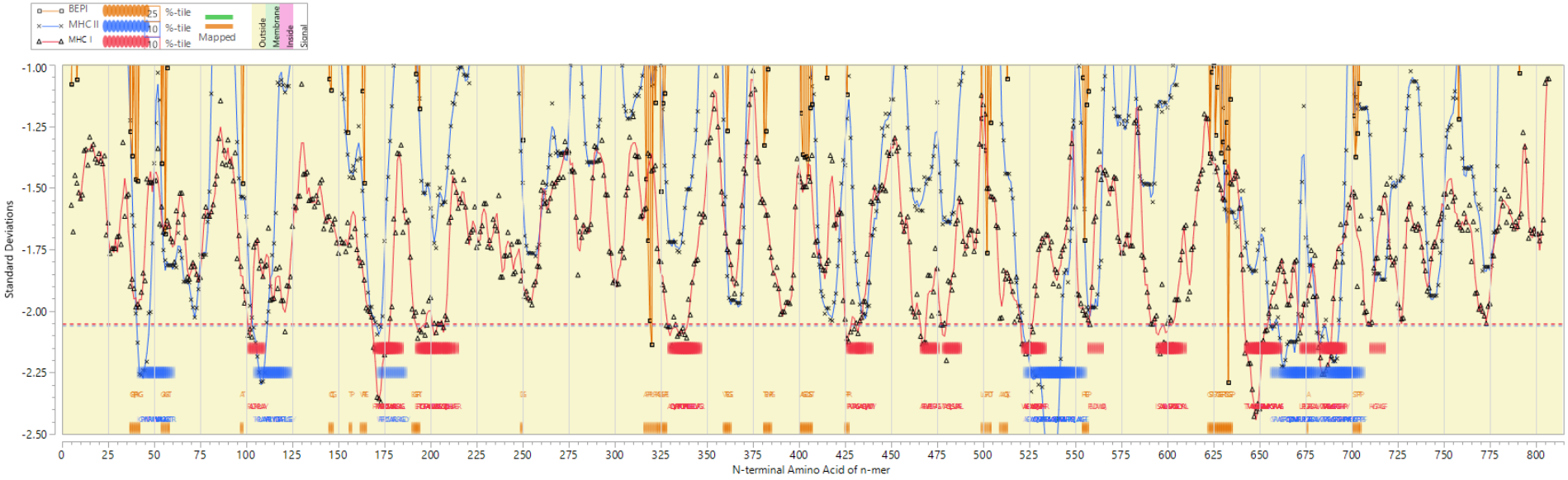
FJ211587-E1 VIPR-ALG4-224482391-8258-9700_1

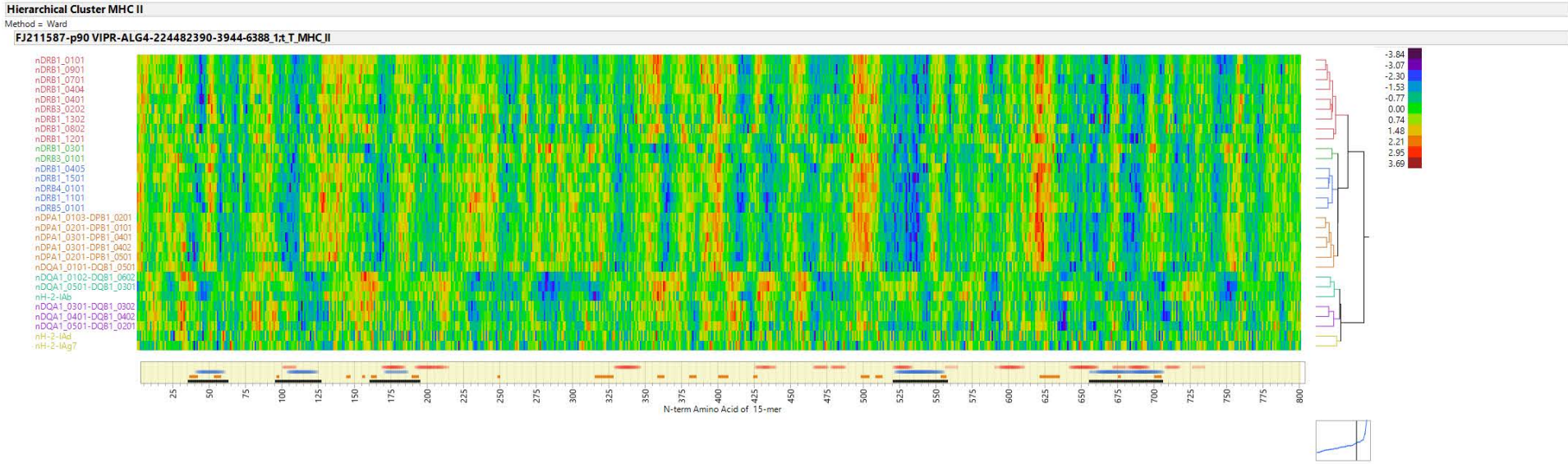


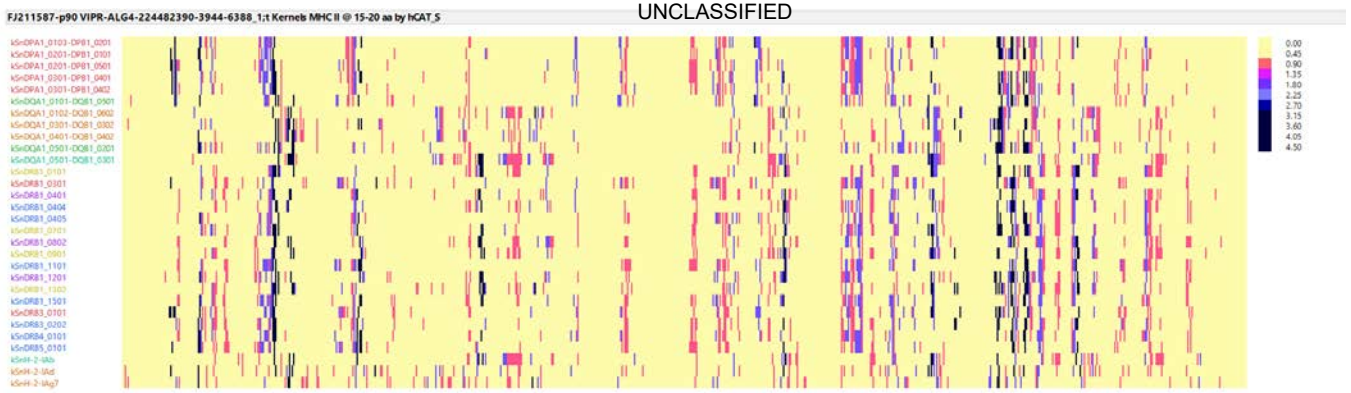
P90

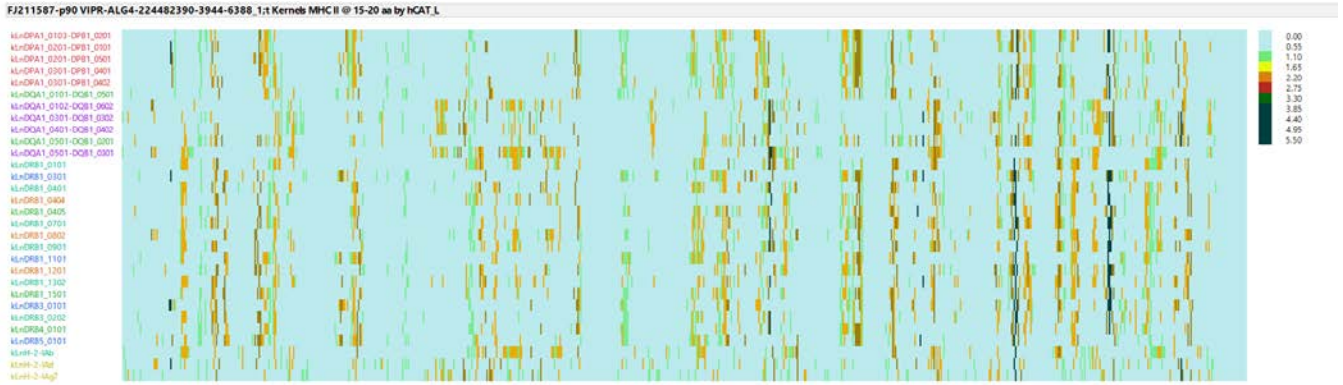
Overlay Plot

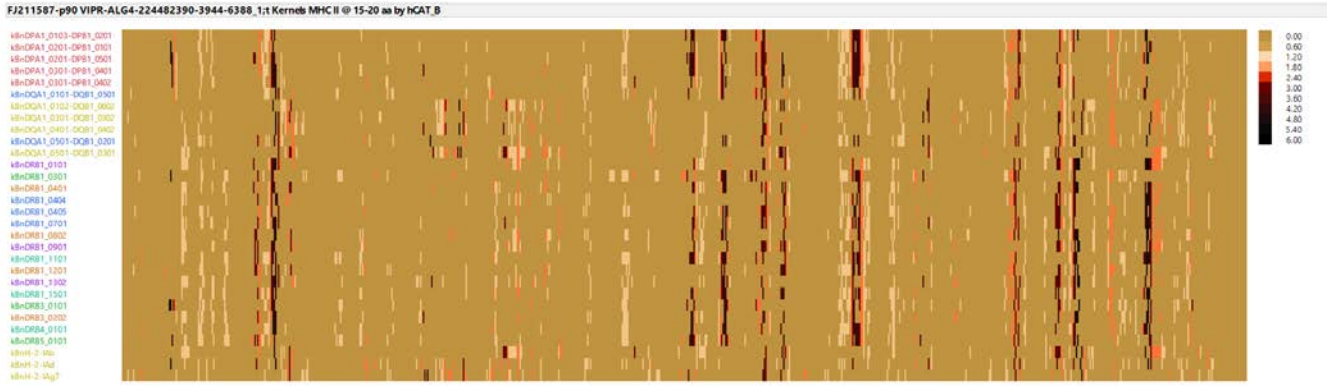
FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1:t

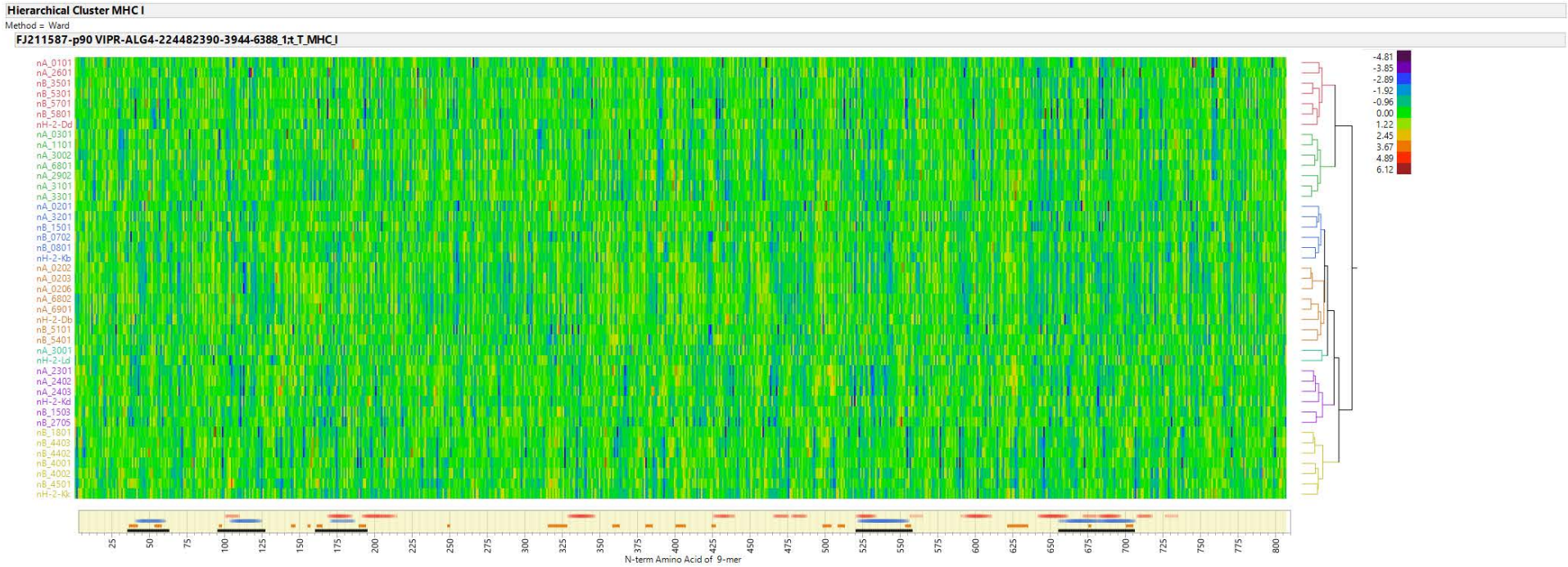


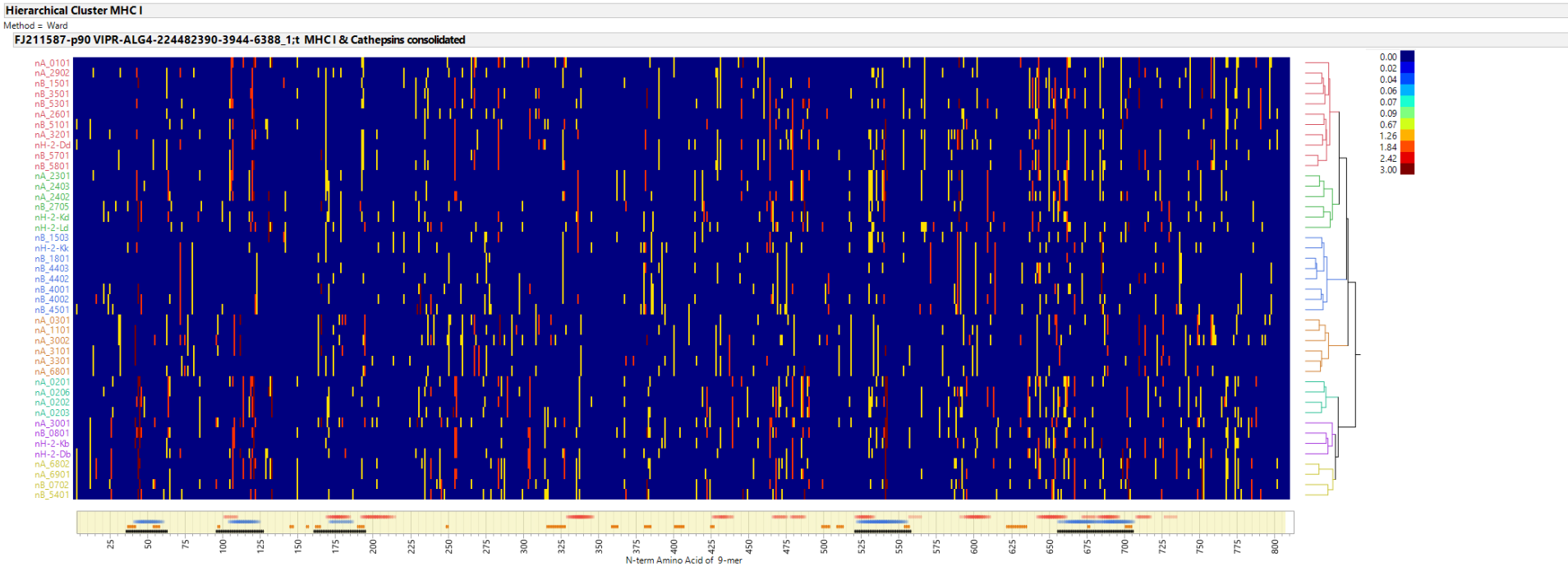

Hierarchical Cluster MHC II
Method = Ward
FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_0401
nDRB3_0202
nDRB1_1302
nDRB1_0802
nDRB1_1201
nDRB1_0301
nDRB3_0101
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0501-DQB1_0301
nH-2-IAb
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
-3.84
-3.07
-2.30
-1.53
-0.77
0.00
0.74
1.48
2.21
2.95
3.69
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
675
700
725
750
775
800
N-term Amino Acid of 15-mer

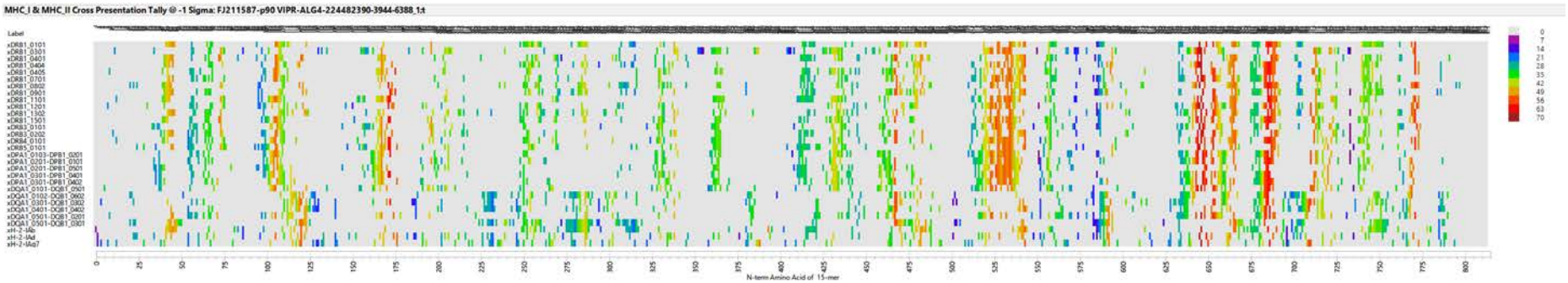

FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50

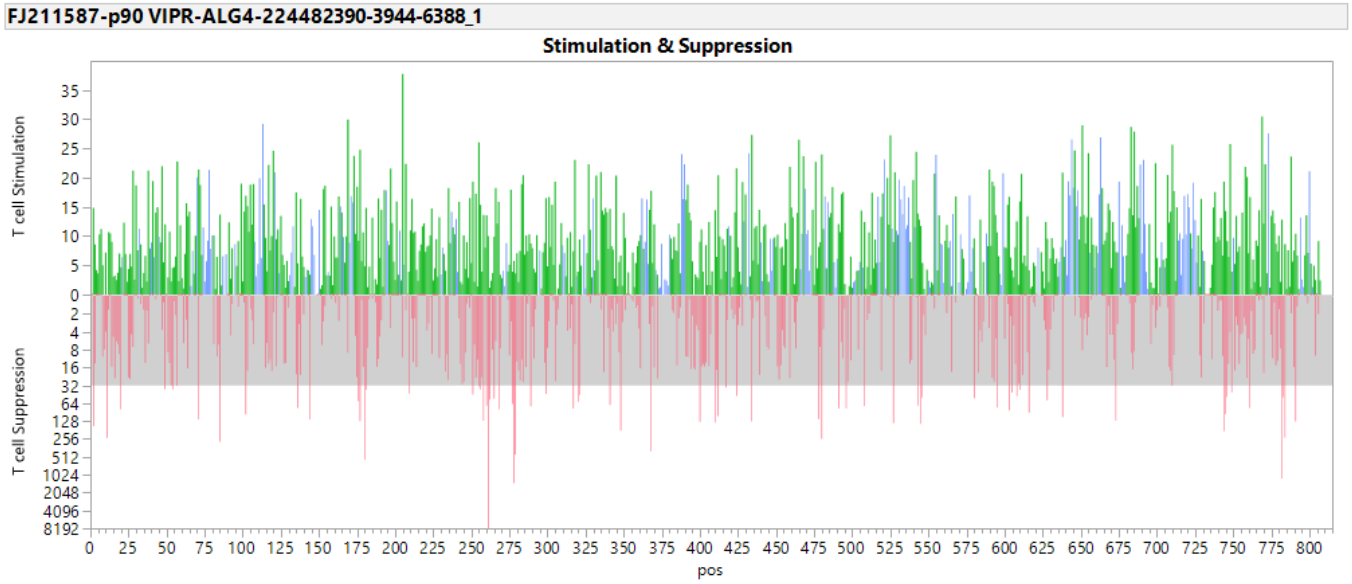

FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

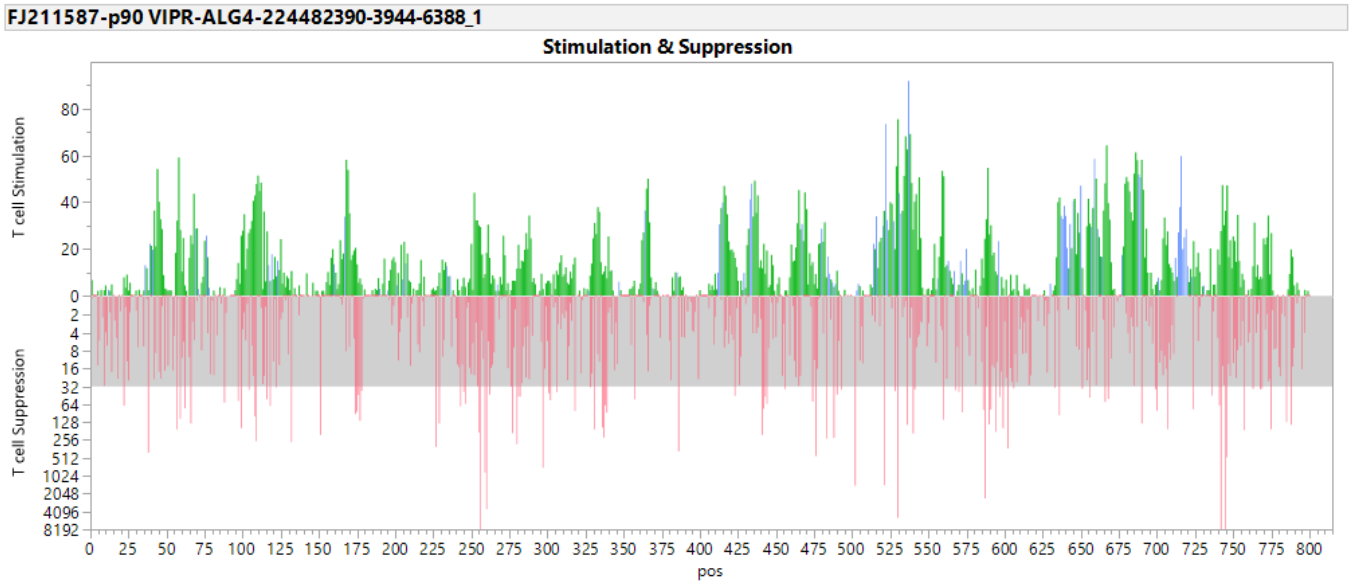

FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00


Hierarchical Cluster MHC I
Method = Ward
FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
nA_0201
nA_3201
nB_1501
nB_0702
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nH-2-Db
nB_5101
nB_5401
nA_3001
nH-2-Ld
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
-4.81
-3.85
-2.89
-1.92
-0.96
0.00
1.22
2.45
3.67
4.89
6.12
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
675
700
725
750
775
800
N-term Amino Acid of 9-mer

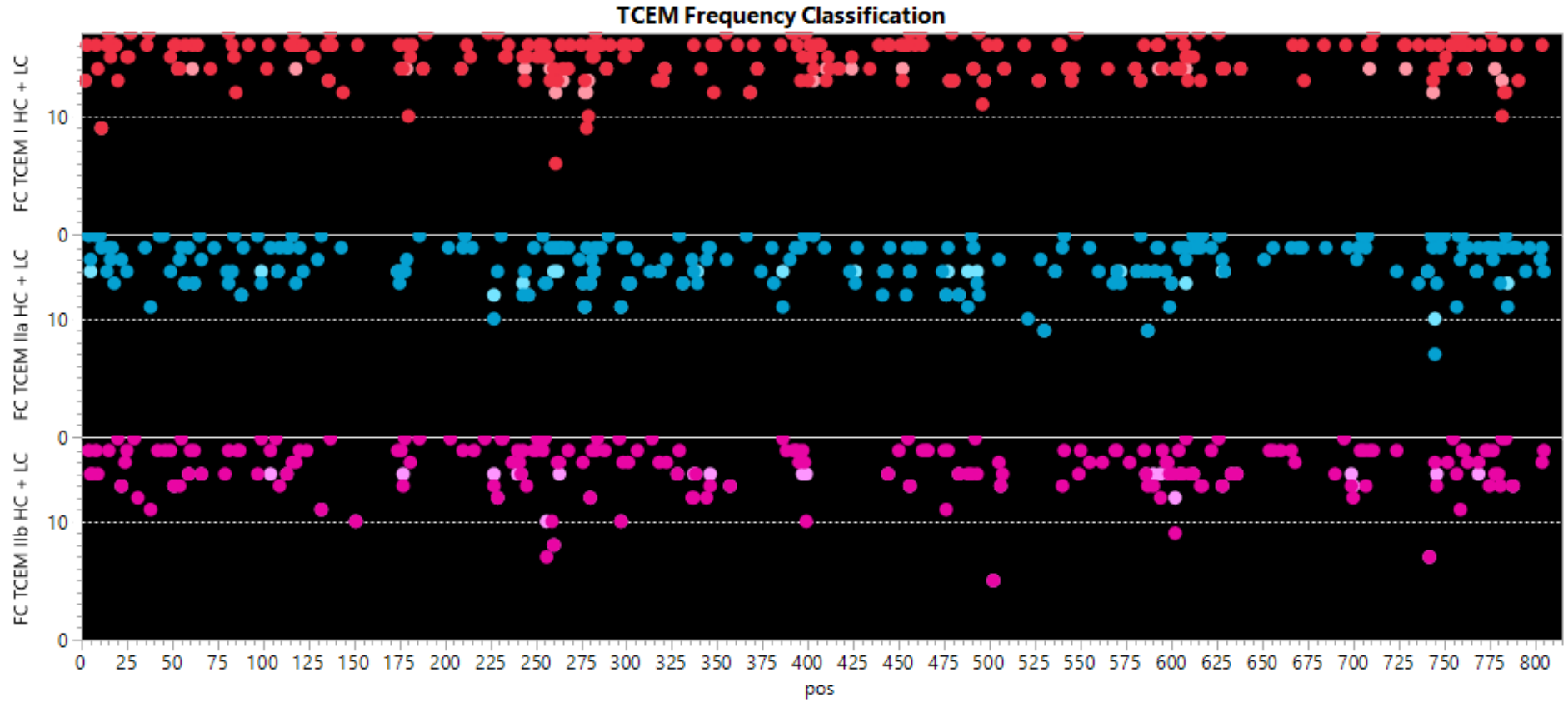

Hierarchical Cluster MHC I
Method = Ward
FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2902
nB_1501
nB_3501
nB_5301
nA_2601
nB_5101
nA_3201
nH-2-Dd
nB_5701
nB_5801
nA_2301
nA_2403
nA_2402
nB_2705
nH-2-Kd
nH-2-Ld
nB_1503
nH-2-Kk
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_0201
nA_0206
nA_0202
nA_0203
nA_3001
nB_0801
nH-2-Kb
nH-2-Db
nA_6802
nA_6901
nB_0702
nB_5401
0.00
0.02
0.04
0.06
0.07
0.09
0.67
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
675
700
725
750
775
800
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1.t
Label
N-term Amino Acid of 15-mer


FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1
Stimulation & Suppression
T cell Stimulation
0 5 10 15 20 25 30 35
T cell Suppression
0 2 4 8 16 32 64 128 256 512 1024 2048 4096 8192
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800
pos


FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1
Stimulation & Suppression
T cell Stimulation
0 20 40 60 80
T cell Suppression
0 2 4 8 16 32 64 128 256 512 1024 2048 4096 8192
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800
pos

FJ211587-p90 VIPR-ALG4-224482390-3944-6388_1

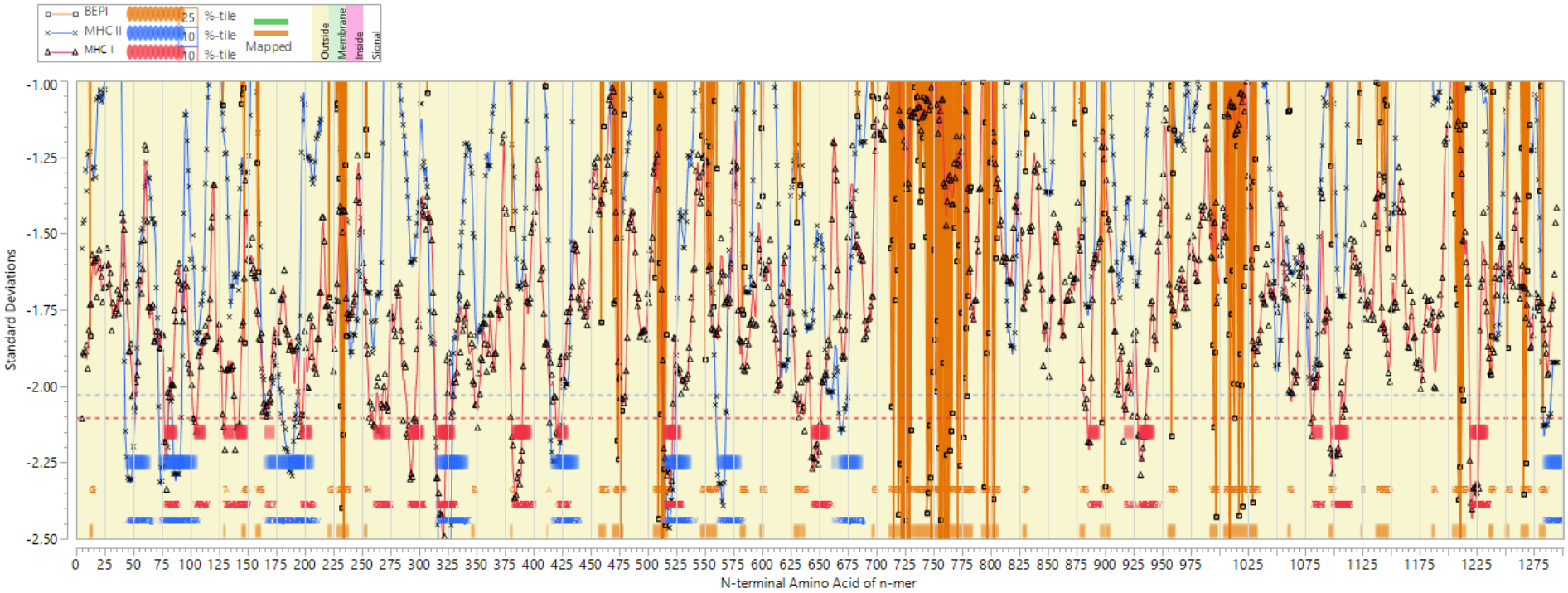


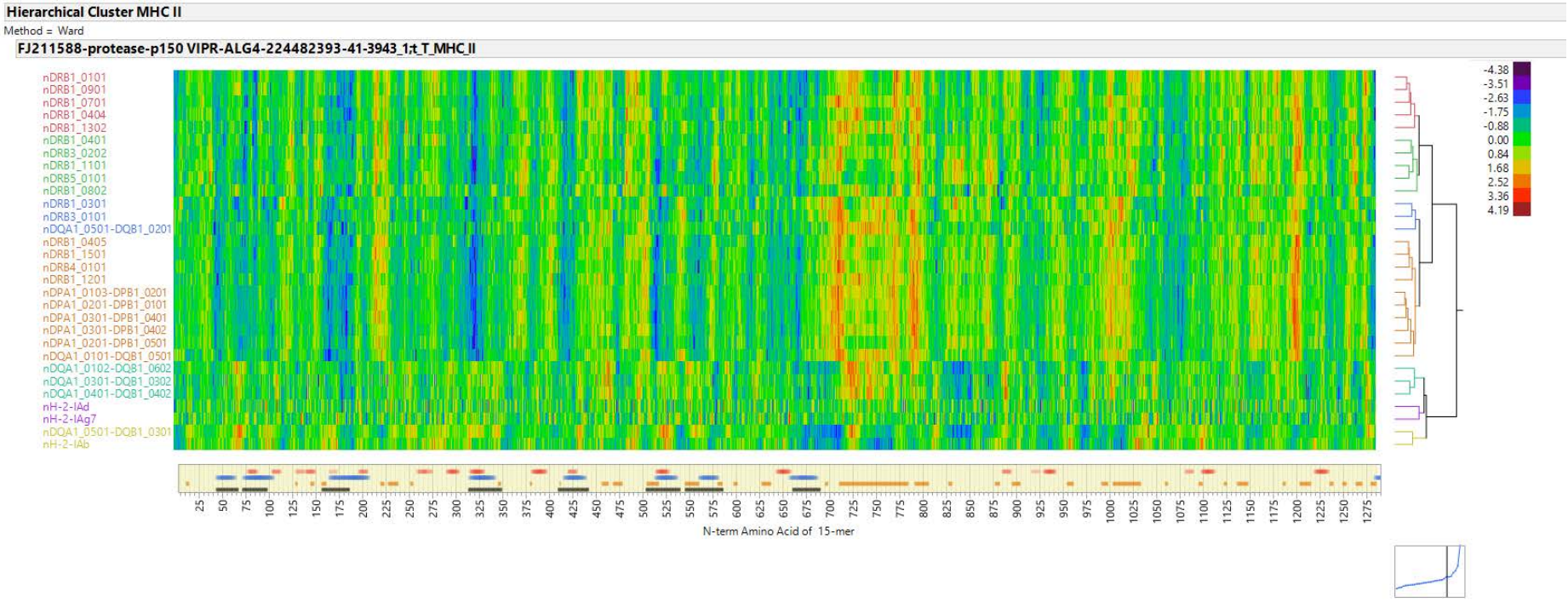
P150

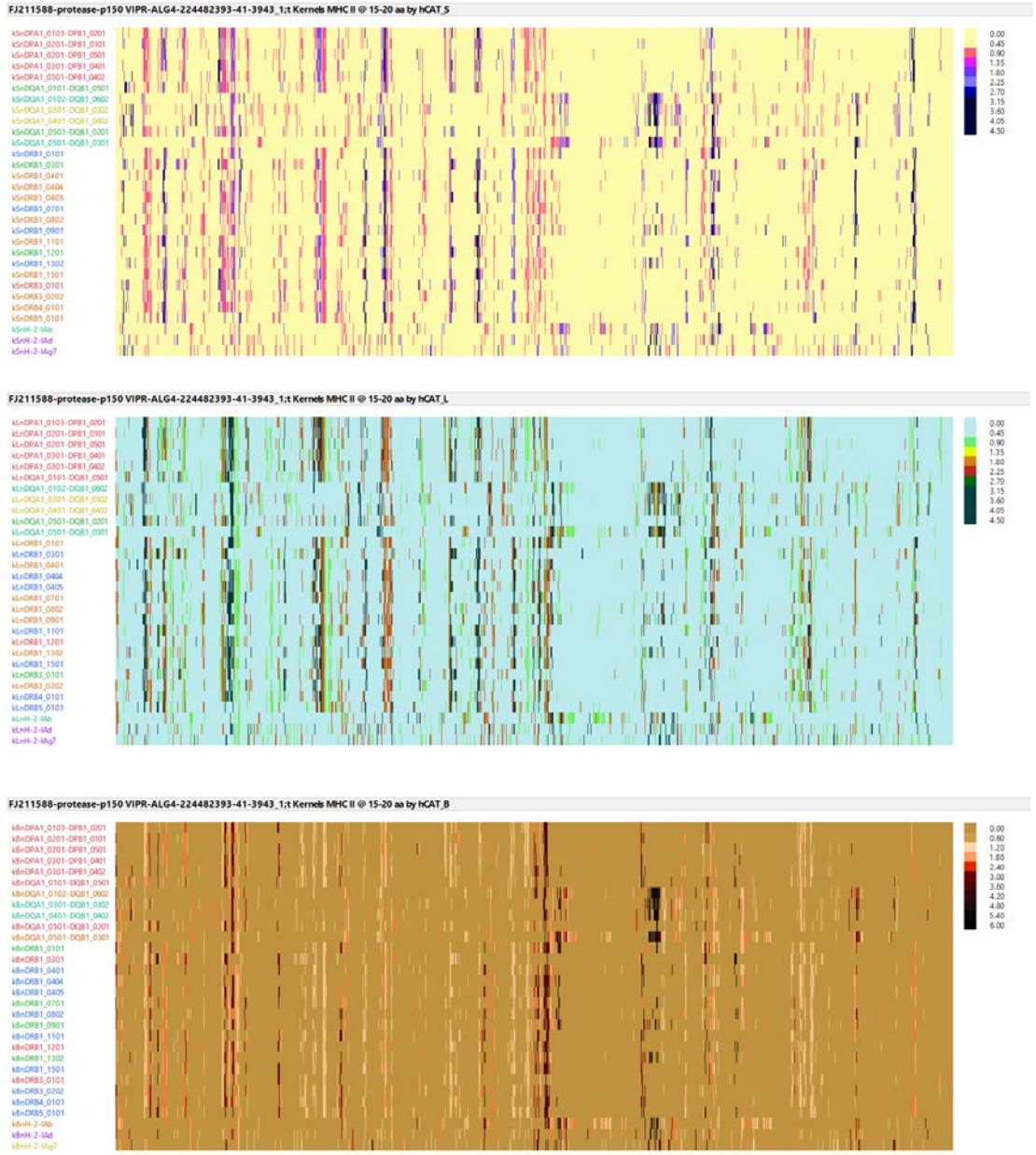
(note this is a different accession number series but also from a full genome Wistar RA 27/3)

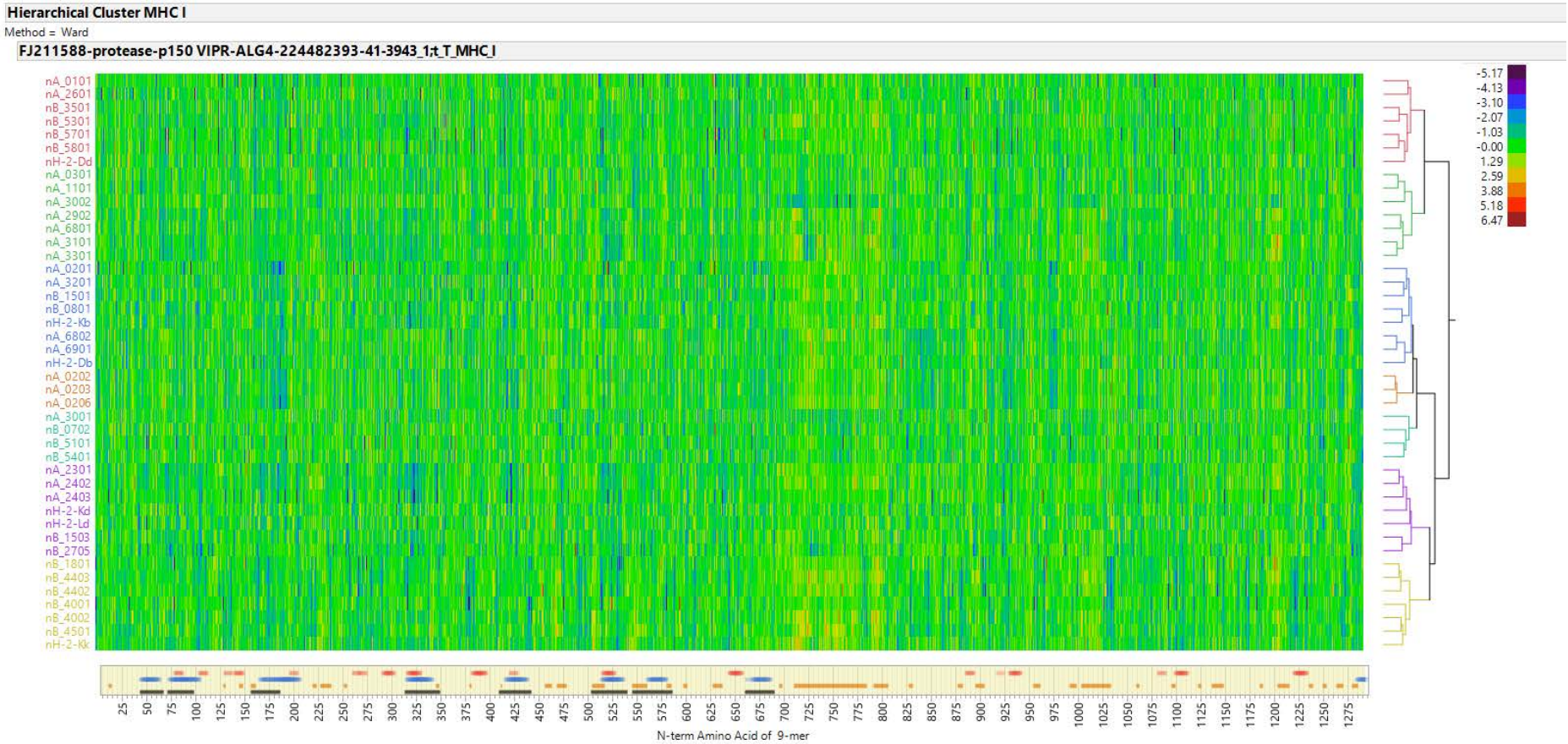
Overlay Plot

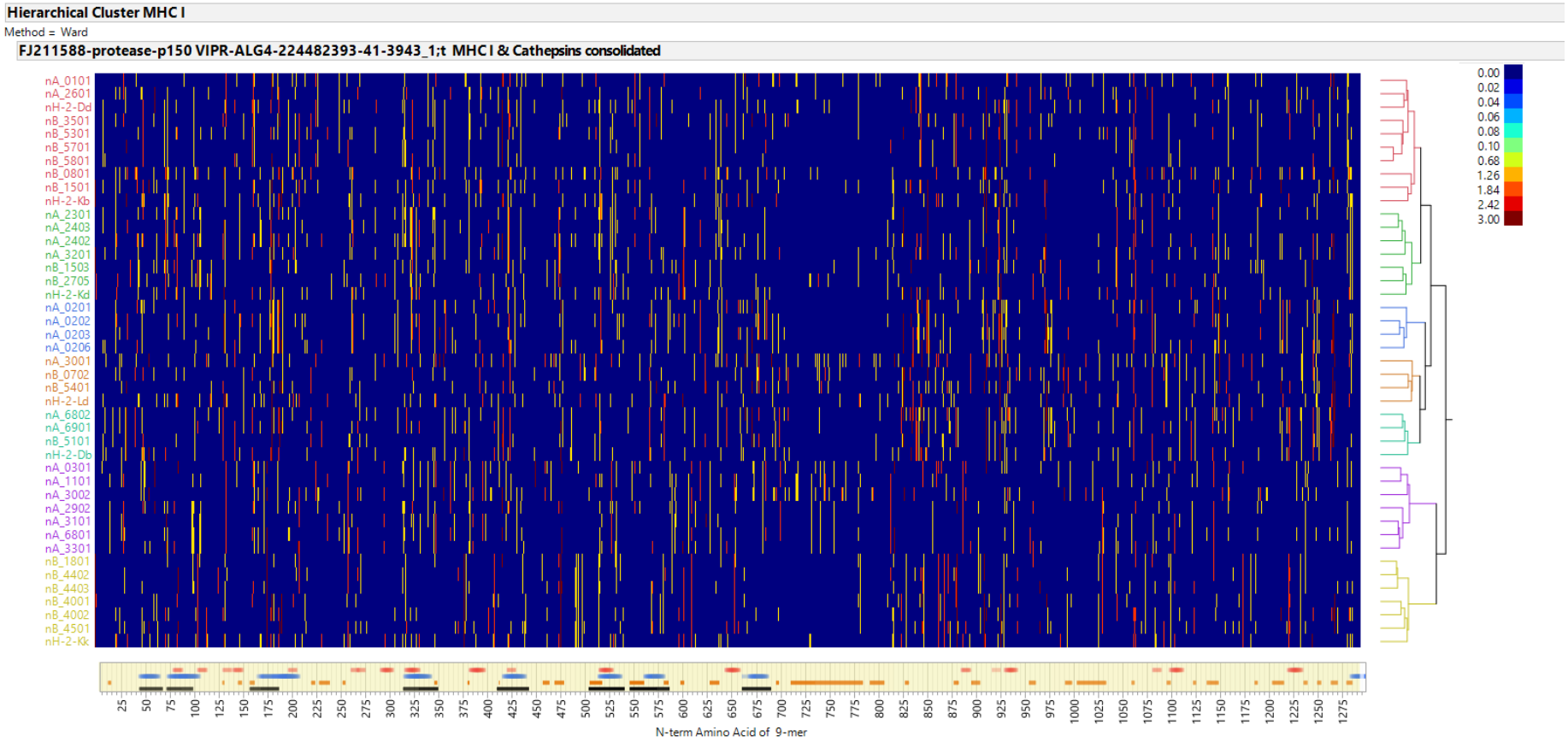
FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t

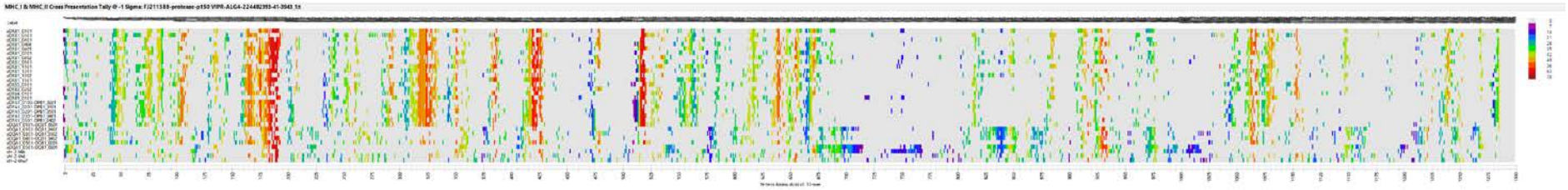


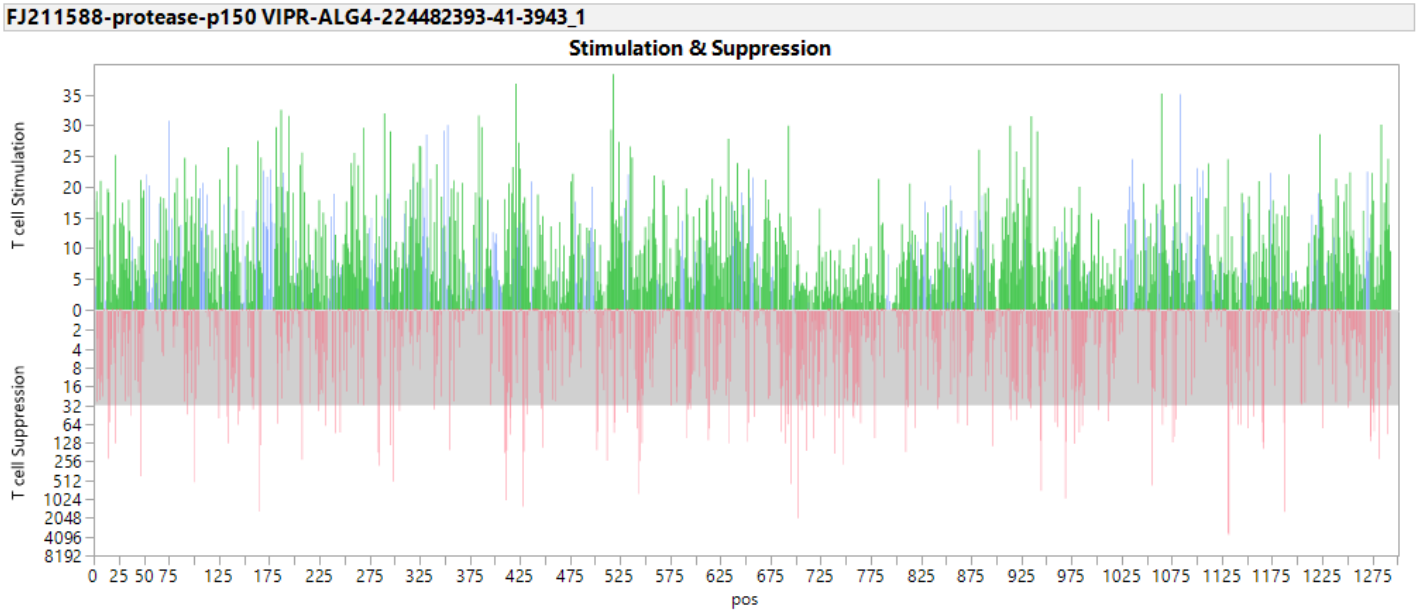

Hierarchical Cluster MHC II
Method = Ward
FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0404
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-4.38
-3.51
-2.63
-1.75
-0.88
0.00
0.84
1.68
2.52
3.36
4.19
25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200 1225 1250 1275
N-term Amino Acid of 15-mer

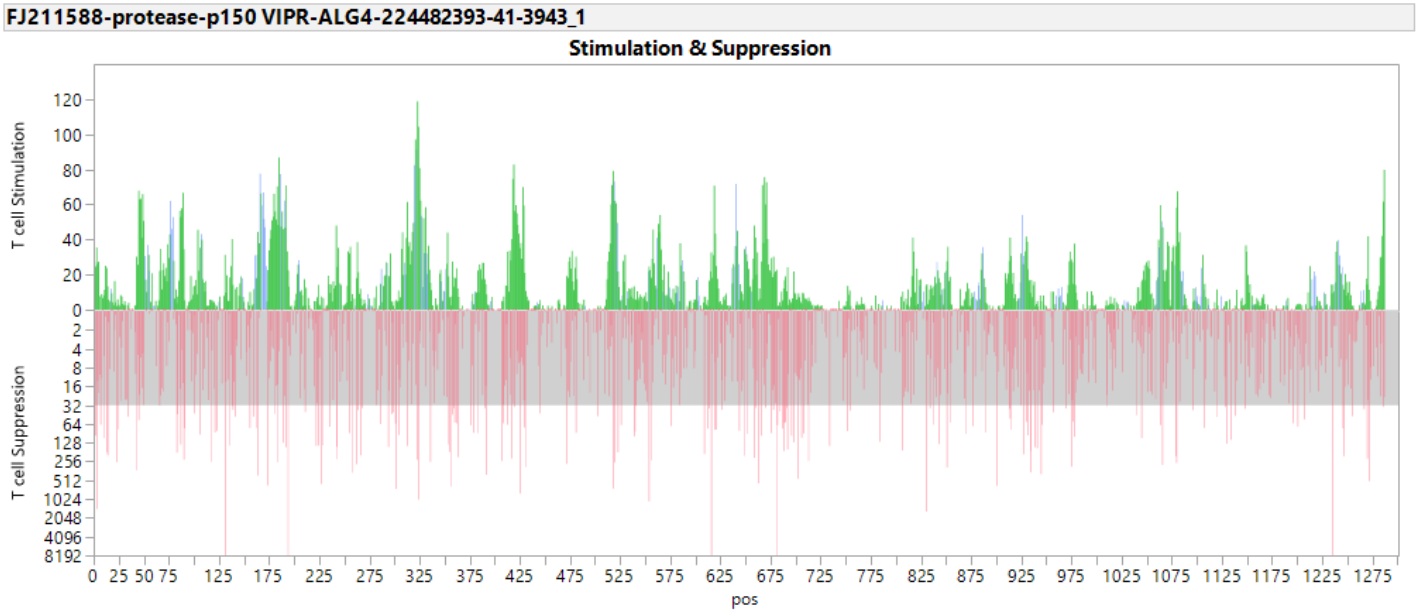

FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t Kernels MHC II @ 15-20 aa by hCAT_S
FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t Kernels MHC II @ 15-20 aa by hCAT_L
FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t Kernels MHC II @ 15-20 aa by hCAT_B


Hierarchical Cluster MHC I
Method = Ward
FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_1101
nA_3002
nA_2902
nA_6801
nA_3101
nA_3301
nA_0201
nA_3201
nB_1501
nB_0801
nH-2-Kb
nA_6802
nA_6901
nH-2-Db
nA_0202
nA_0203
nA_0206
nA_3001
nB_0702
nB_5101
nB_5401
nA_2301
nA_2402
nA_2403
nH-2-Kd
nH-2-Ld
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
-5.17
-4.13
-3.10
-2.07
-1.03
-0.00
1.29
2.59
3.88
5.18
6.47
25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200 1225 1250 1275
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nH-2-Dd
nB_3501
nB_5301
nB_5701
nB_5801
nB_0801
nB_1501
nH-2-Kb
nA_2301
nA_2403
nA_2402
nA_3201
nB_1503
nB_2705
nH-2-Kd
nA_0201
nA_0202
nA_0203
nA_0206
nA_3001
nB_0702
nB_5401
nH-2-Ld
nA_6802
nA_6901
nB_5101
nH-2-Db
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_6801
nA_3301
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nH-2-Kk
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200 1225 1250 1275
N-term Amino Acid of 9-mer

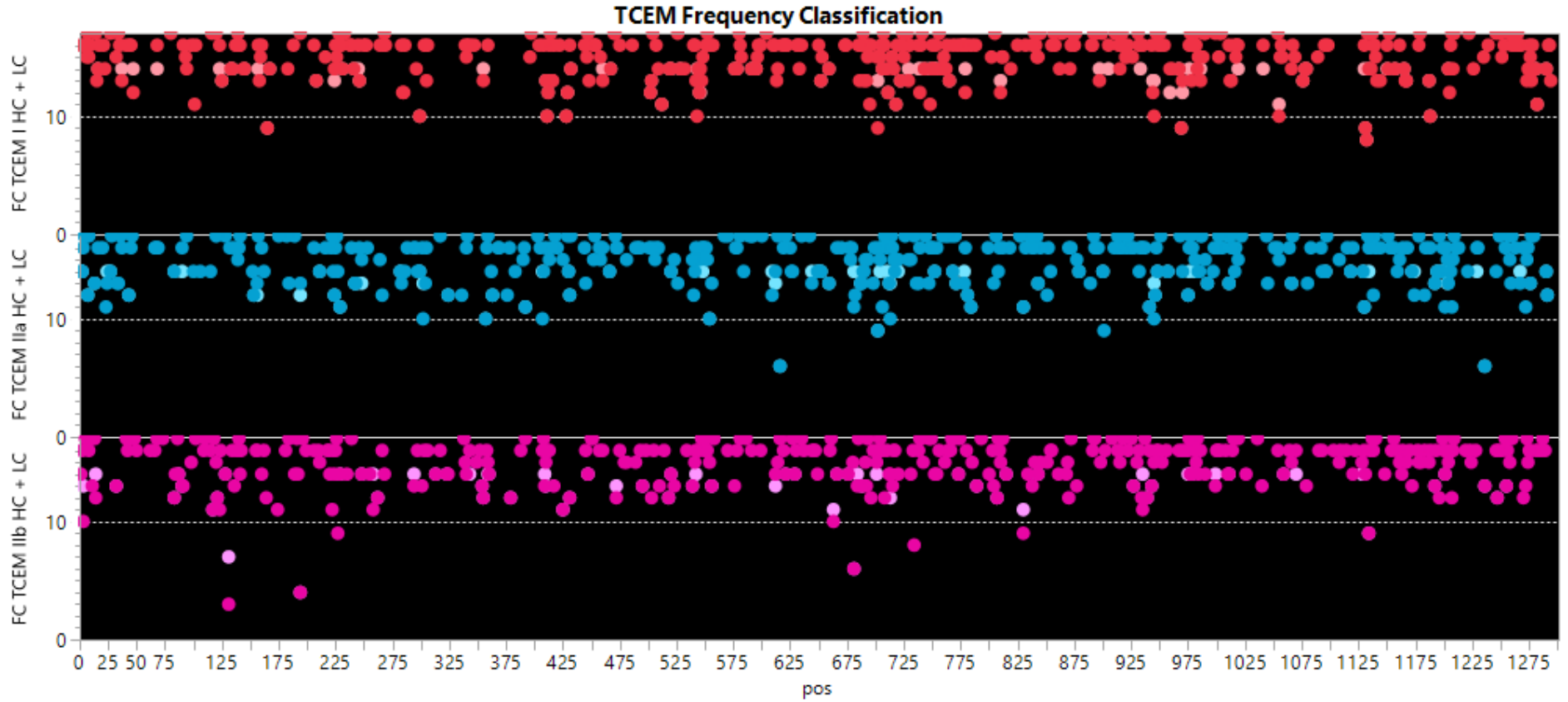






MHC II pos 131 suppressive index 172300

FJ211588-protease-p150 VIPR-ALG4-224482393-41-3943_1

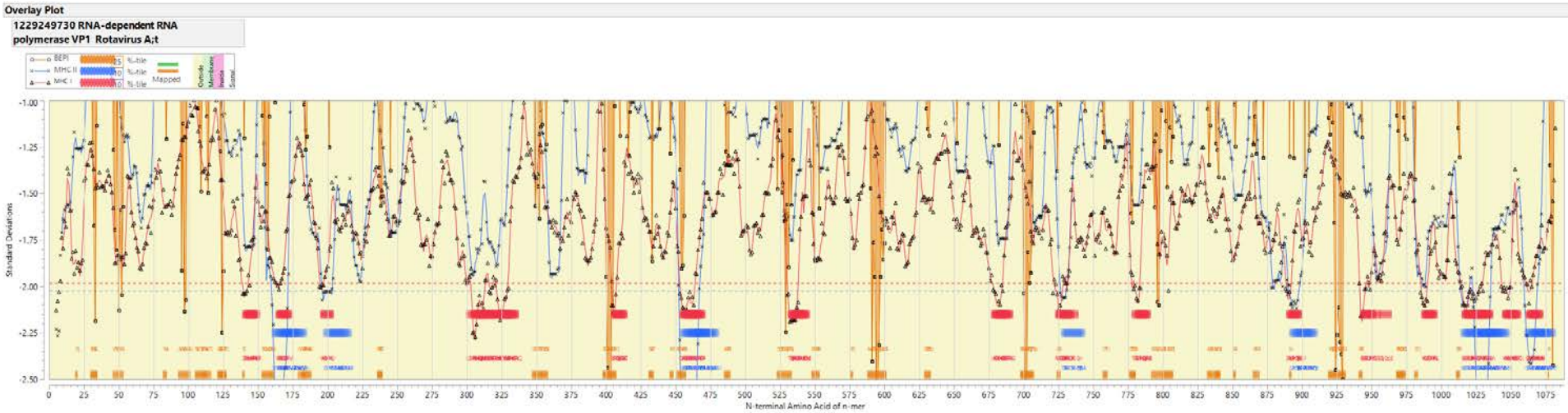


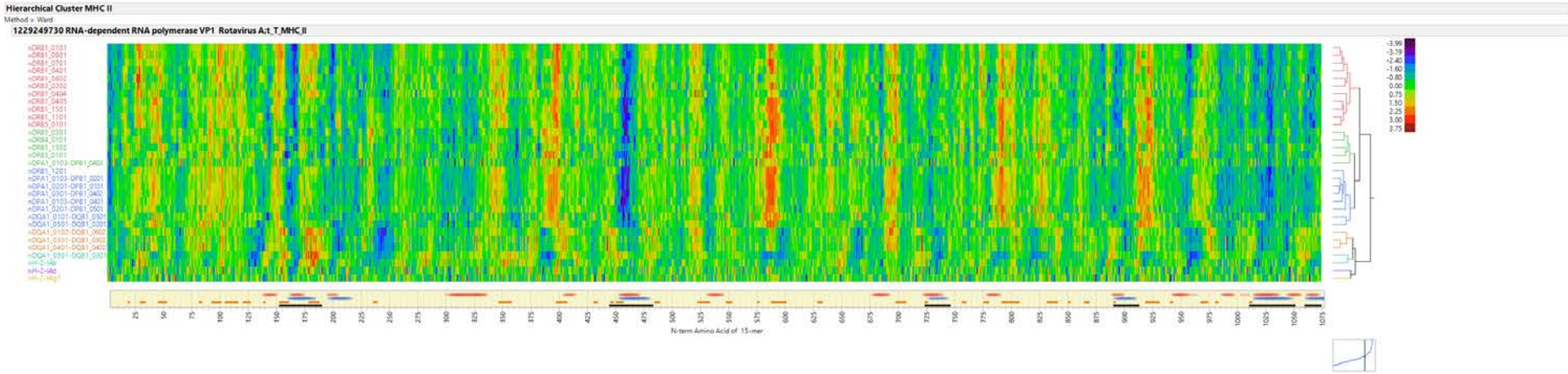
ioGenetics®

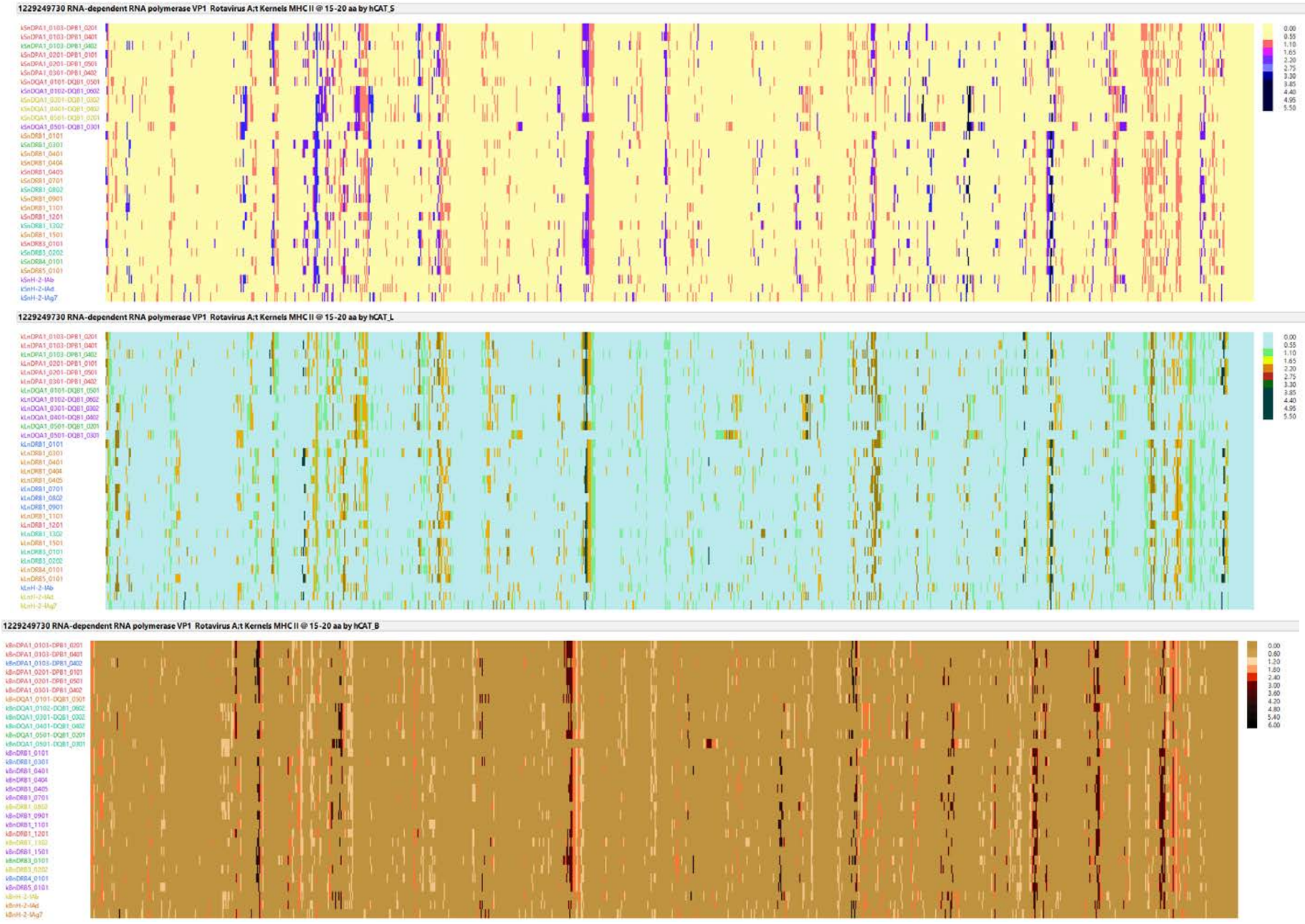
Addendum 5 Rotarix

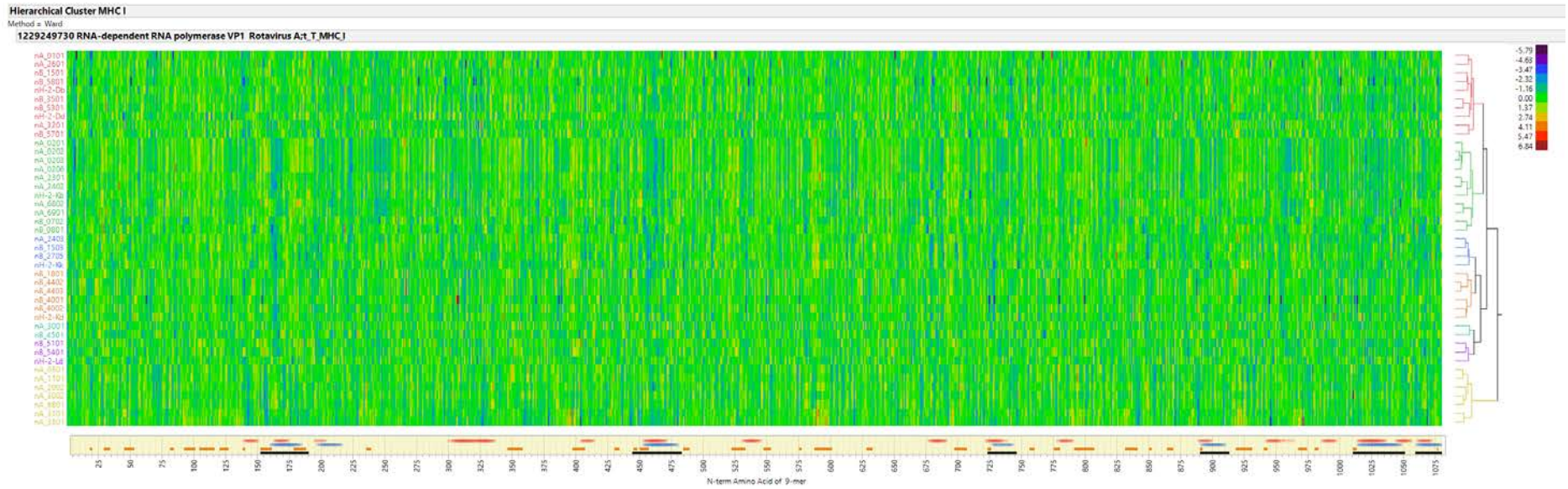
VP1, VP2, VP3, VP6, VP7, NS1, NS2, NS3, NS4, NS5, NS6

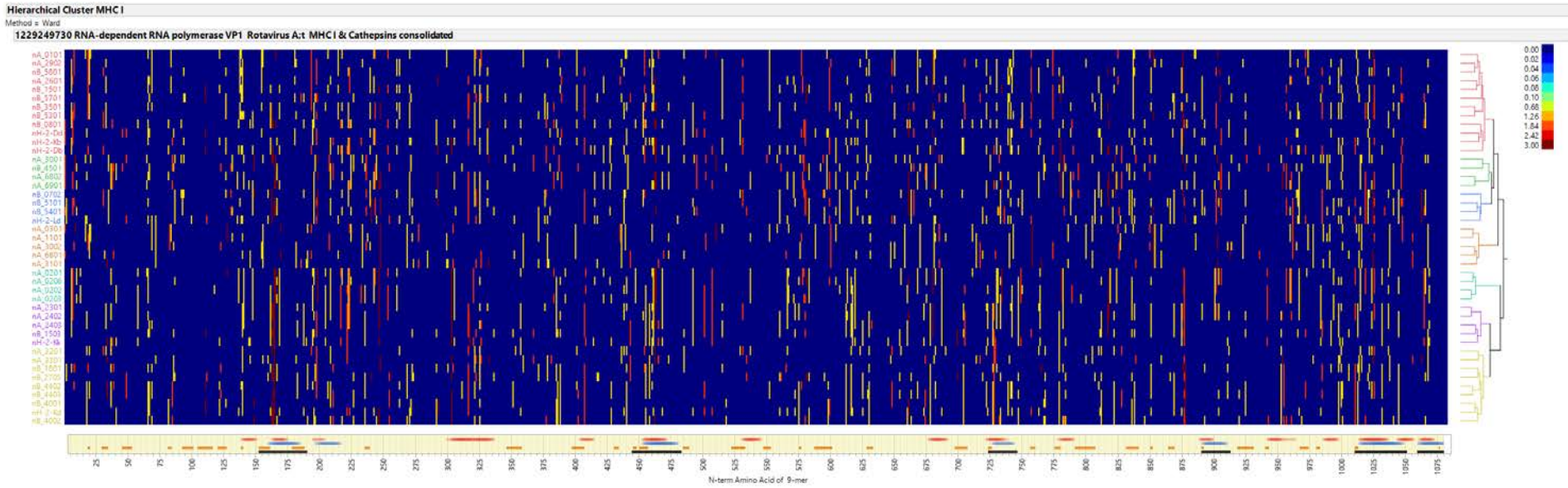
VP1

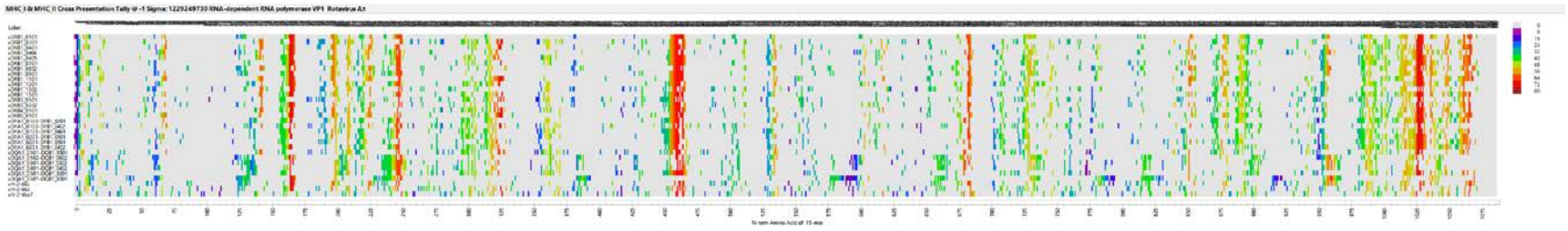

Overlay Plot
1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t
BEPI
MHC II
MHC I
%-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
N-terminal Amino Acid of n-mer


Hierarchical Cluster MHC II
Method = Ward
1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t_T_MHC_II
N-term Amino Acid of 15-mer

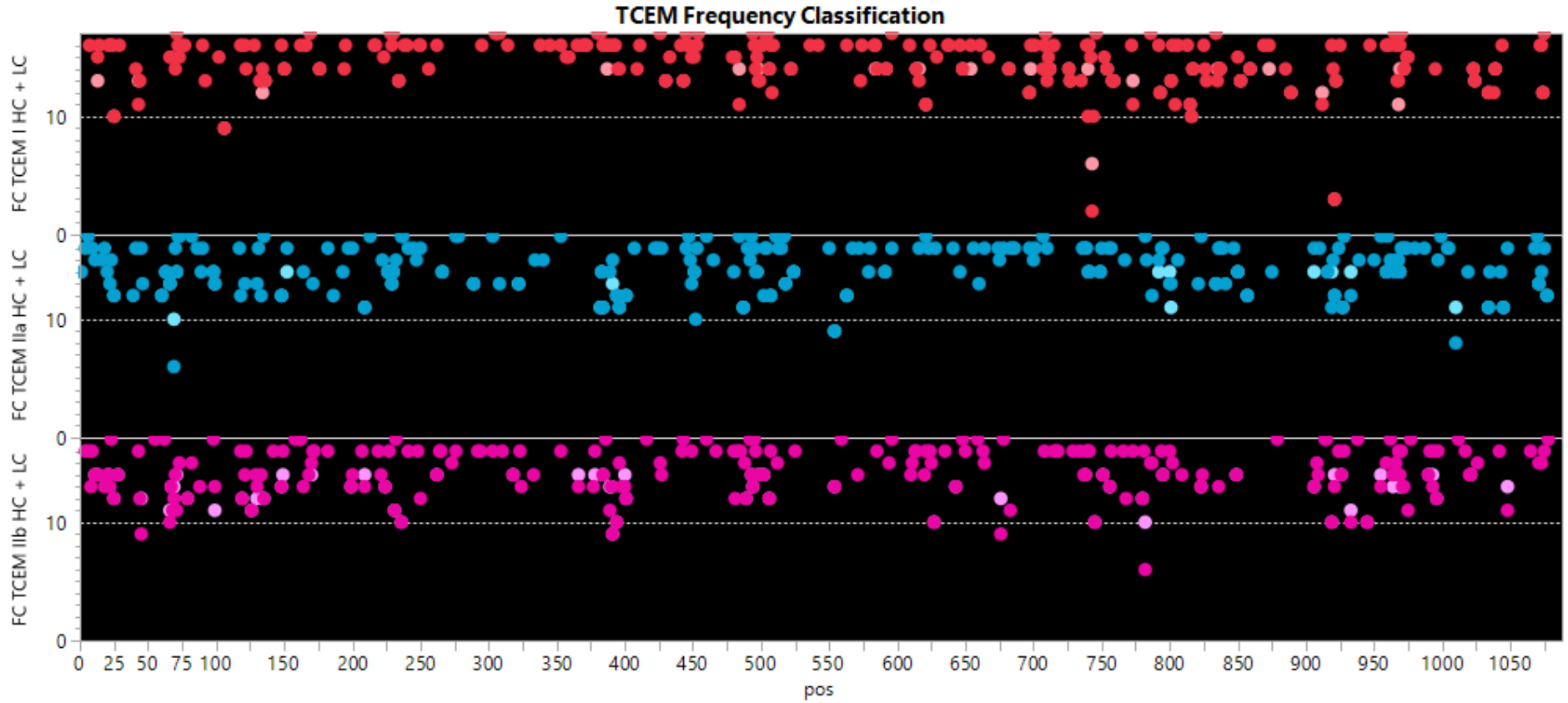

1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t Kernels MHC II @ 15-20 aa by hCAT_S
1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t Kernels MHC II @ 15-20 aa by hCAT_L
1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t Kernels MHC II @ 15-20 aa by hCAT_B

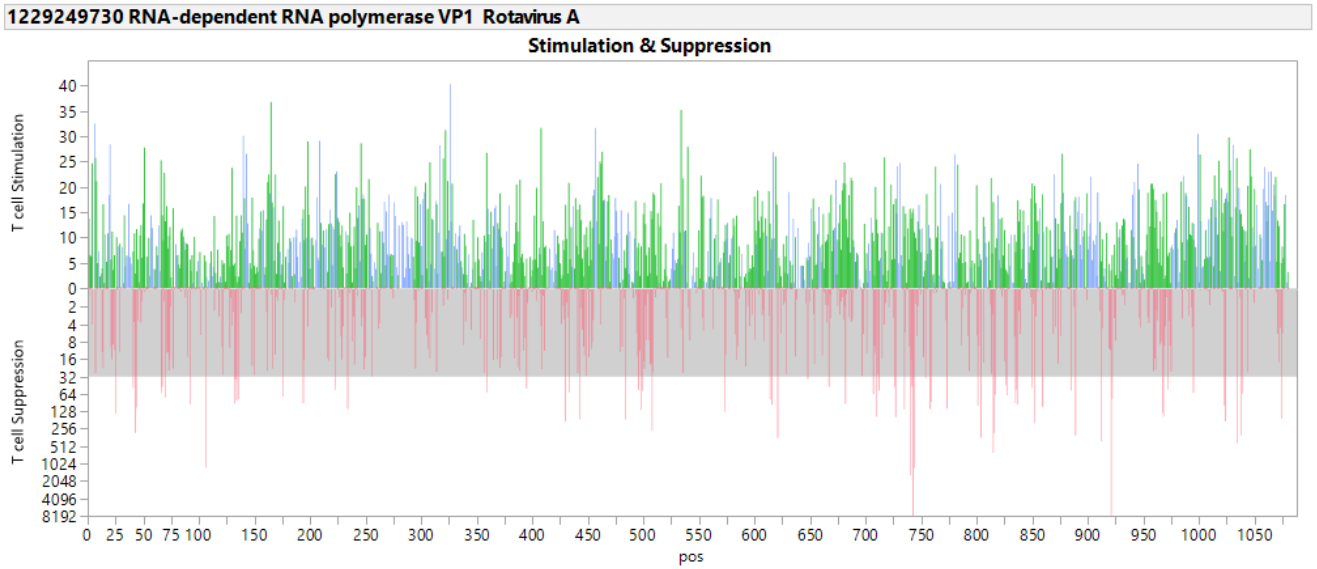

Hierarchical Cluster MHC I
Method = Ward
1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t_T_MHC_I
N-term Amino Acid of 9-mer

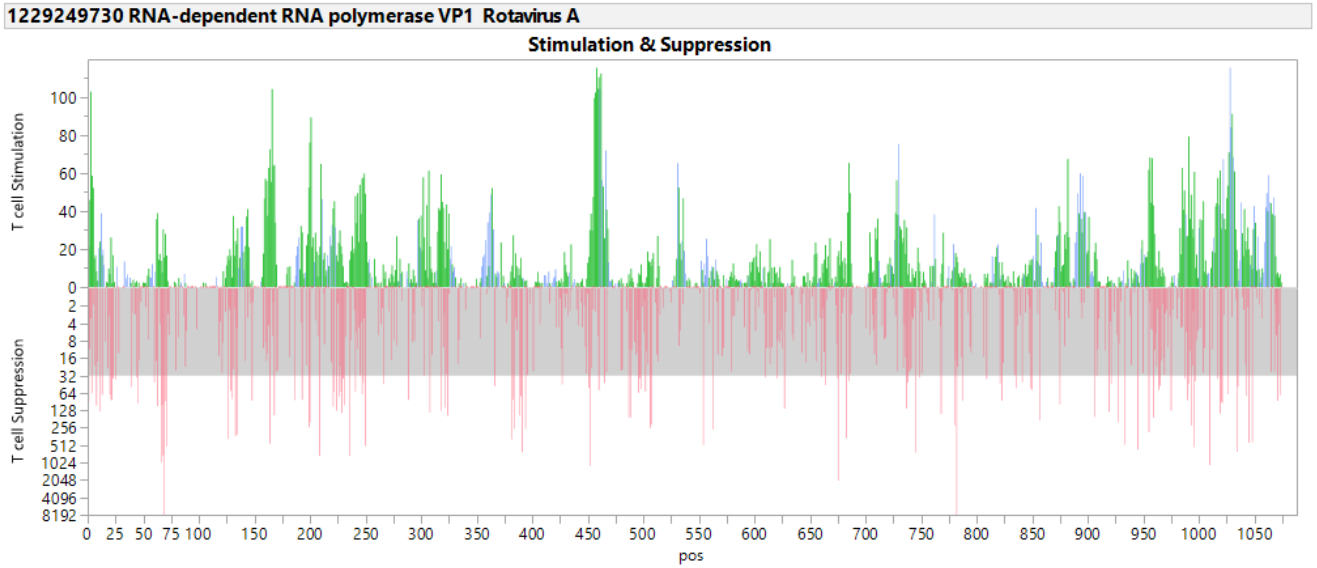

Hierarchical Cluster MHC I
Method = Ward
1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A;t MHC I & Cathepsins consolidated
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
N-term Amino Acid of 9-mer



1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A




1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

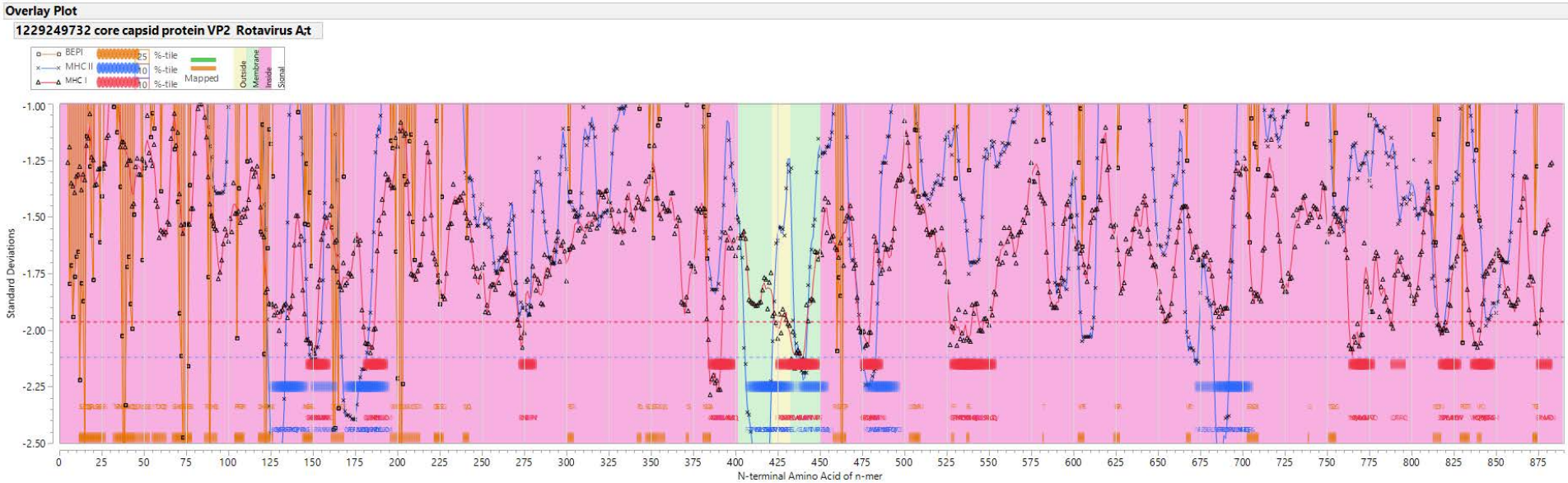

1229249730 RNA-dependent RNA polymerase VP1 Rotavirus A
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

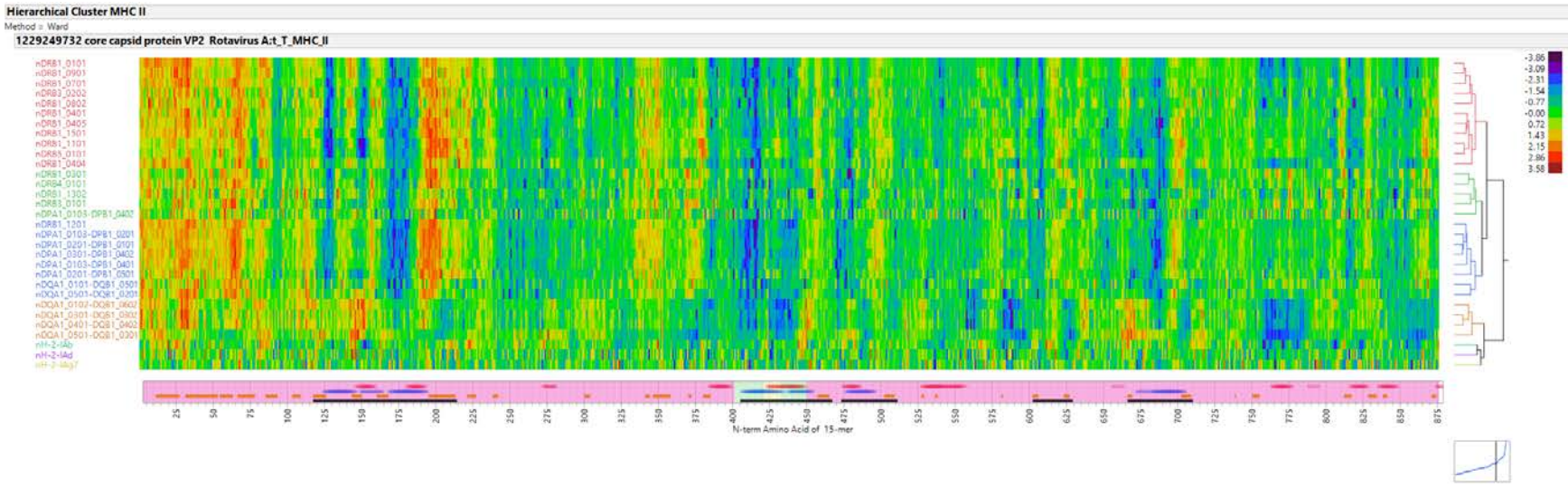
Notes on VP1

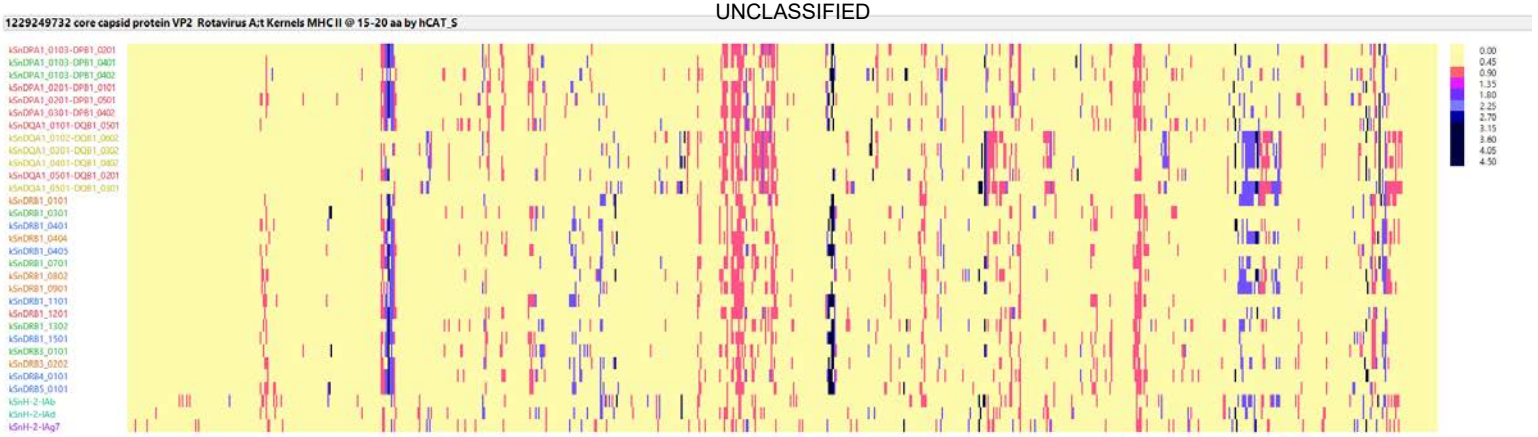
- Peptide at index position 743 comprises a TCEM I FC2 which leads to a suppressive index of 335,351
- No significant TCEM II suppressors

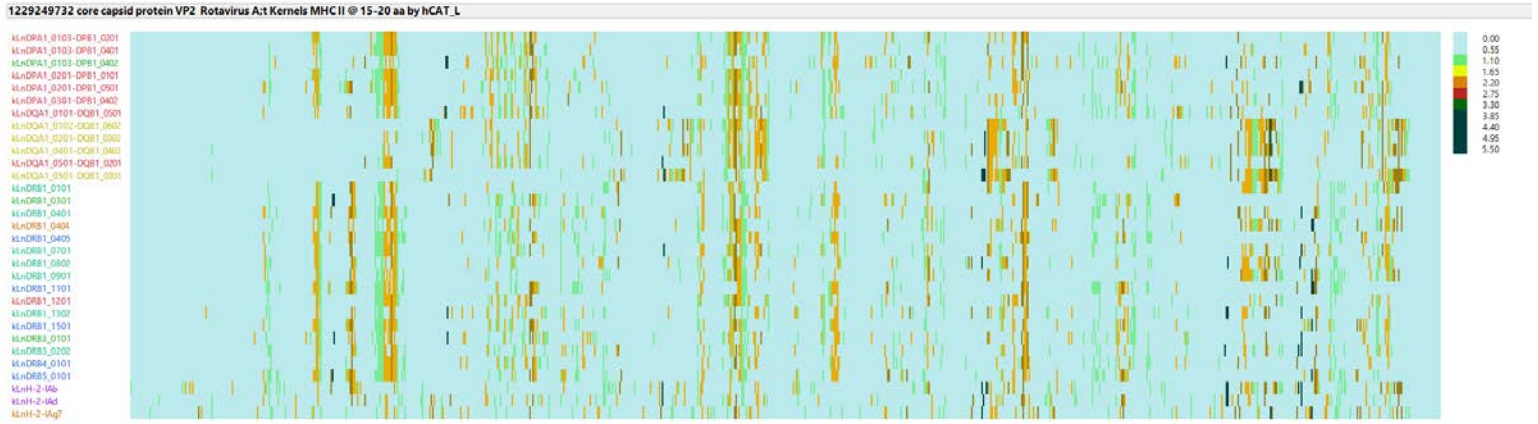
VP2

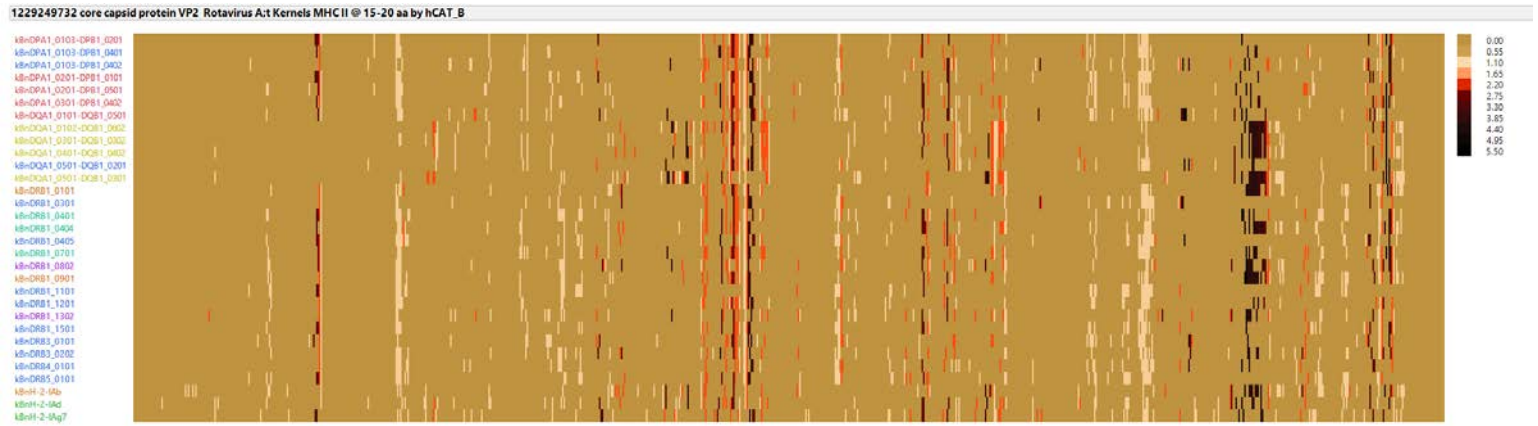
Core Capsid protein

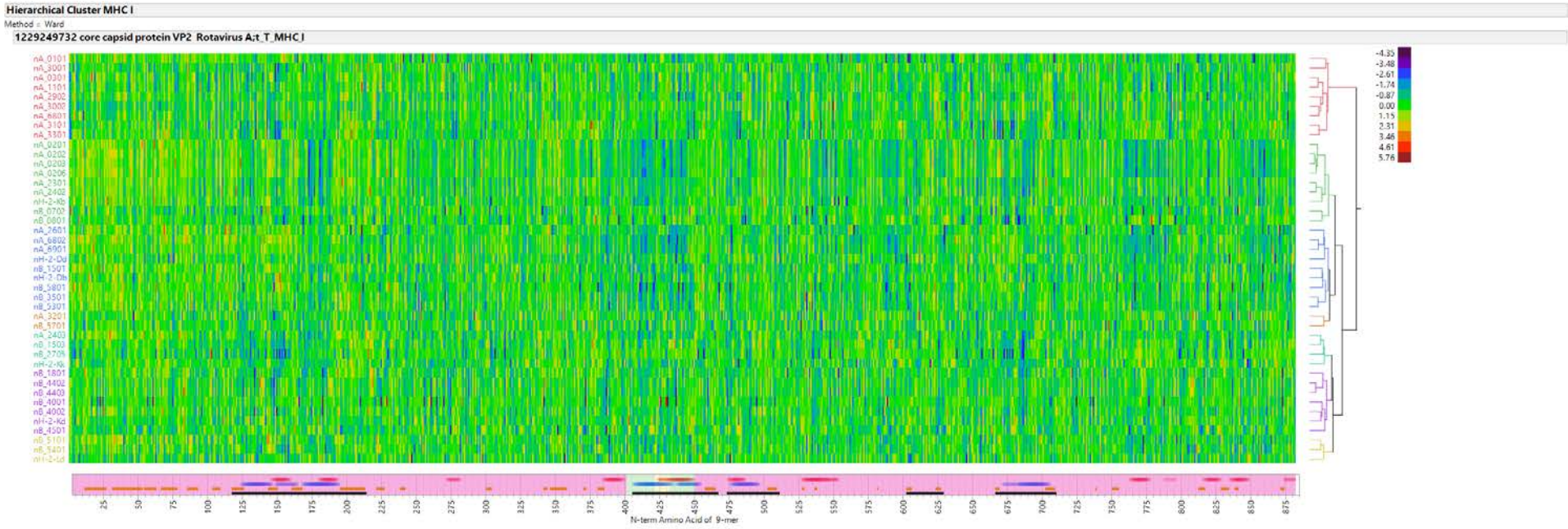

Overlay Plot
1229249732 core capsid protein VP2 Rotavirus A:t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875
N-terminal Amino Acid of n-mer

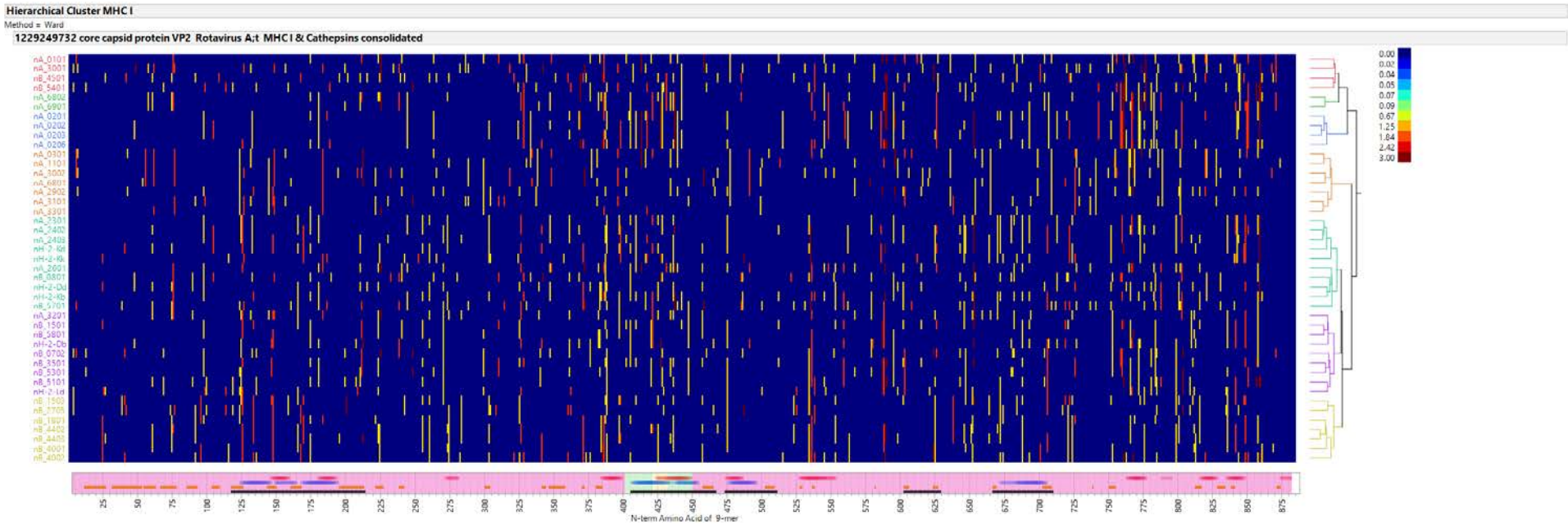

Hierarchical Cluster MHC II
Method = Ward
1229249732 core capsid protein VP2 Rotavirus A:t_T_MHC_II
-3.86
-3.09
-2.31
-1.54
-0.77
-0.00
0.72
1.43
2.15
2.86
3.58
N-term Amino Acid of 15-mer

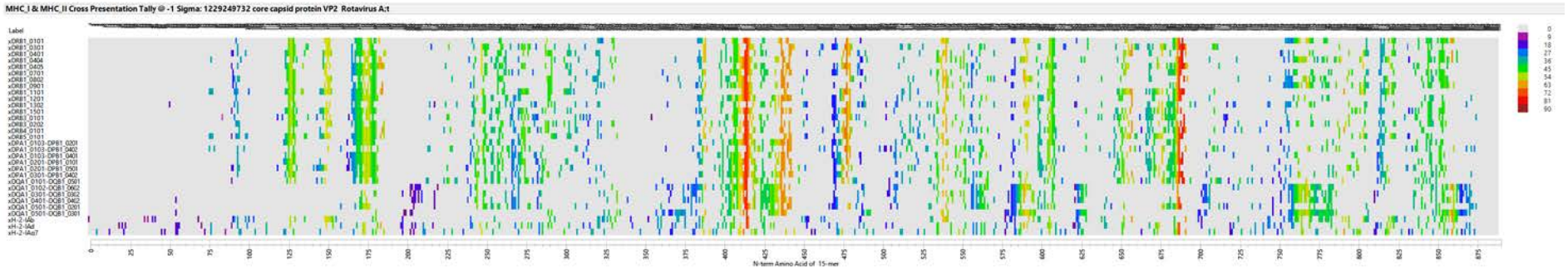

1229249732 core capsid protein VP2 Rotavirus A;t Kernels MHC II @ 15-20 aa by hCAT_S
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50


1229249732 core capsid protein VP2 Rotavirus A;t Kernels MHC II @ 15-20 aa by hCAT_L
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50

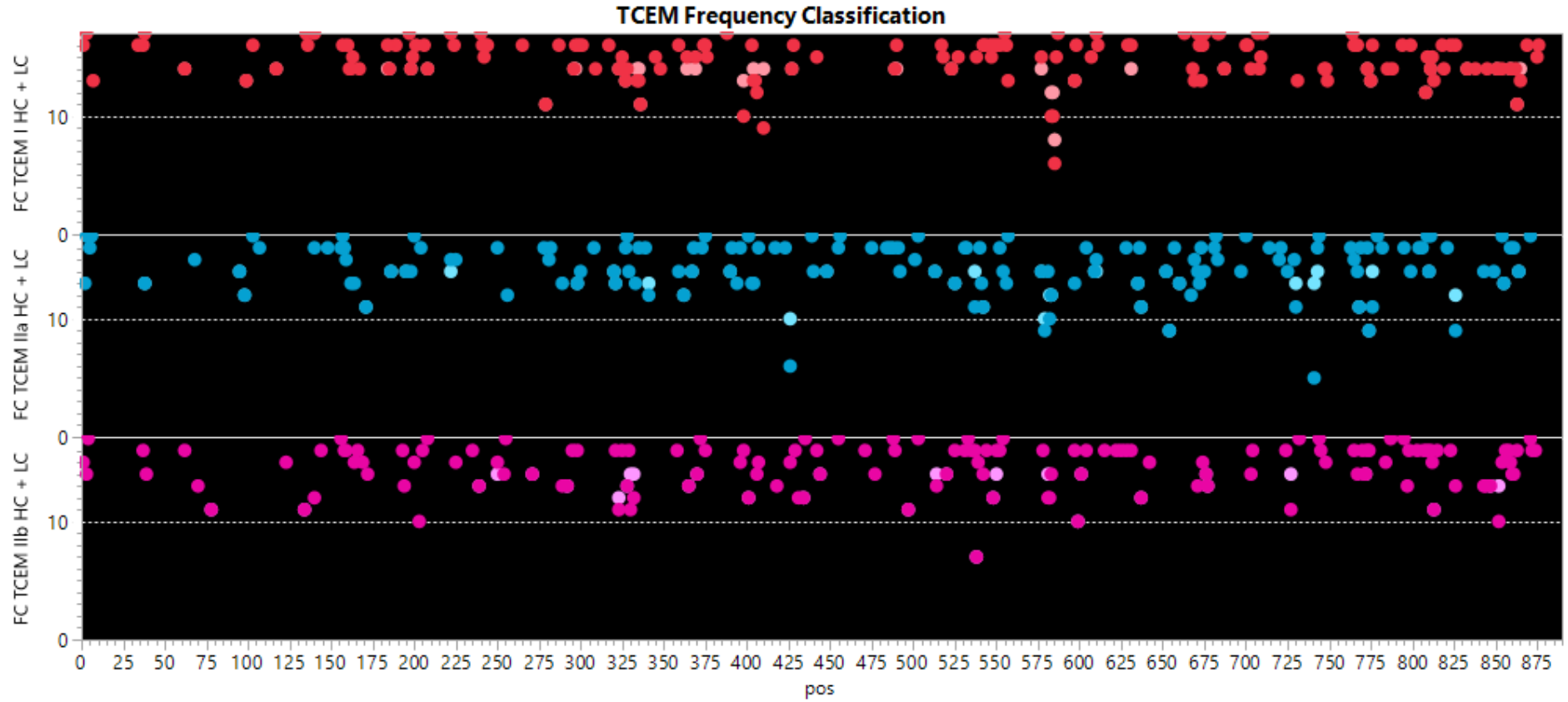

1229249732 core capsid protein VP2 Rotavirus A;t Kernels MHC II @ 15-20 aa by hCAT_B
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50


Hierarchical Cluster MHC I
Method = Ward
1229249732 core capsid protein VP2 Rotavirus A;t_T_MHC I
-4.35
-3.48
-2.61
-1.74
-0.87
0.00
1.15
2.31
3.46
4.61
5.76
N-term Amino Acid of 9-mer

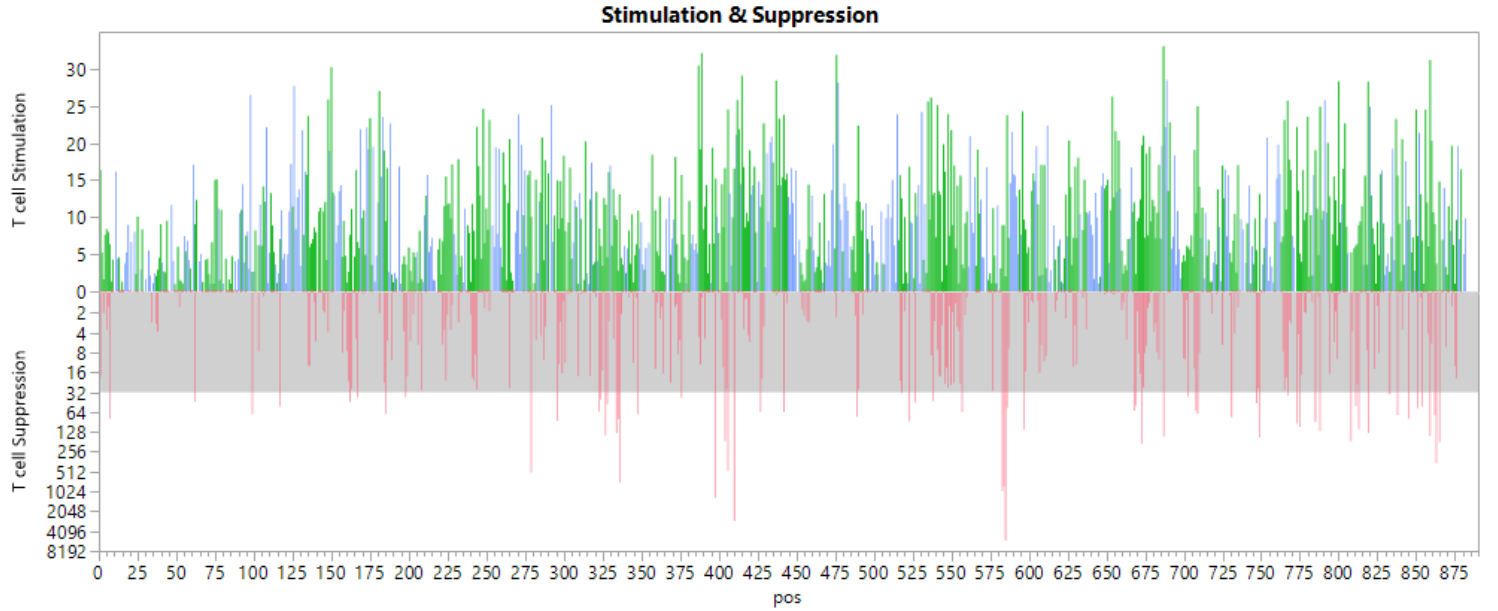

Hierarchical Cluster MHC I
Method = Ward
1229249732 core capsid protein VP2 Rotavirus A;t MHC I & Cathepsins consolidated
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00
N-term Amino Acid of 9-mer

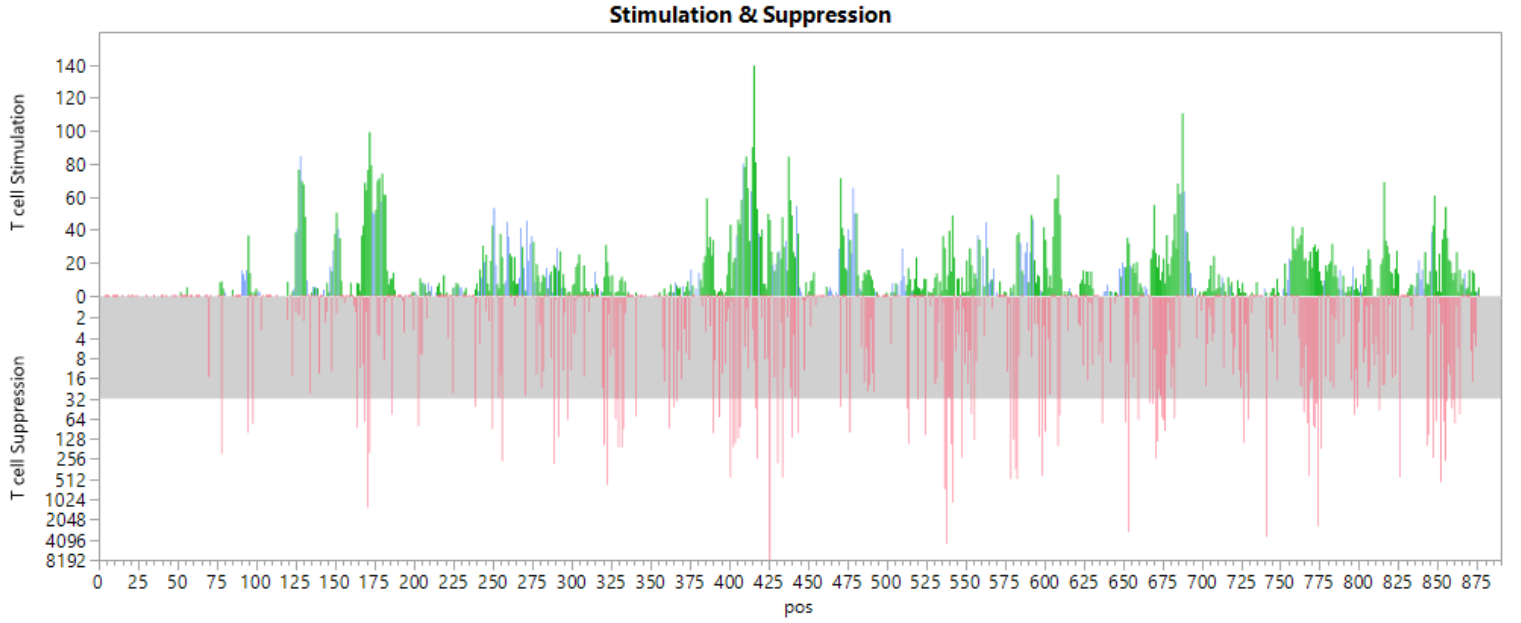

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 1229249732 core capsid protein VP2 Rotavirus A;t
Label
N-term Amino Acid of 15-mer

1229249732 core capsid protein VP2 Rotavirus A



1229249732 core capsid protein VP2 Rotavirus A





ioGenetics®

Addendum 6

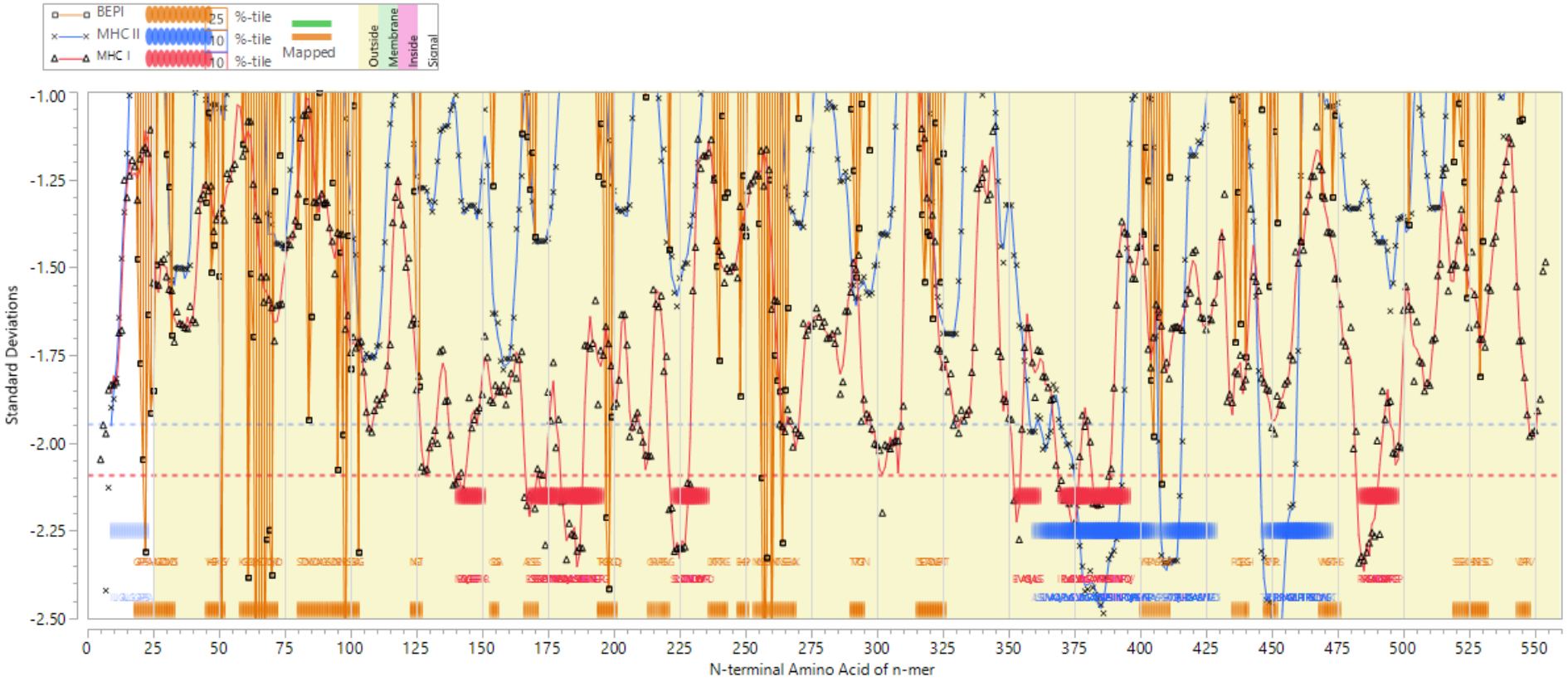
Corynebacterium diphtheriae toxin
Clostridium tetani toxin

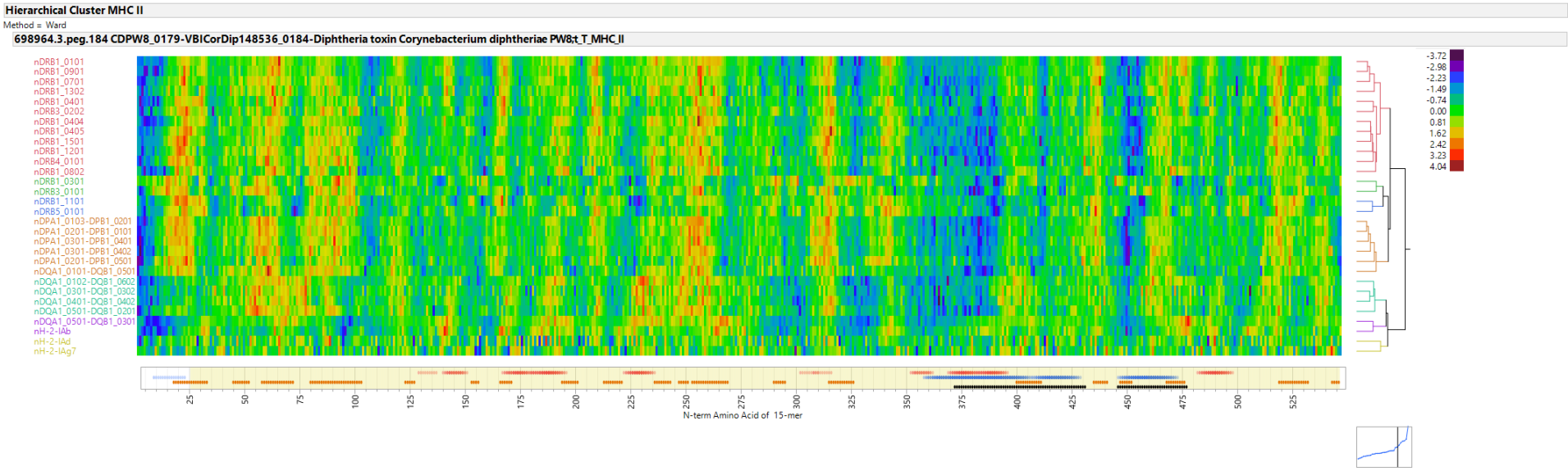
(note that vaccine comprises toxoids)

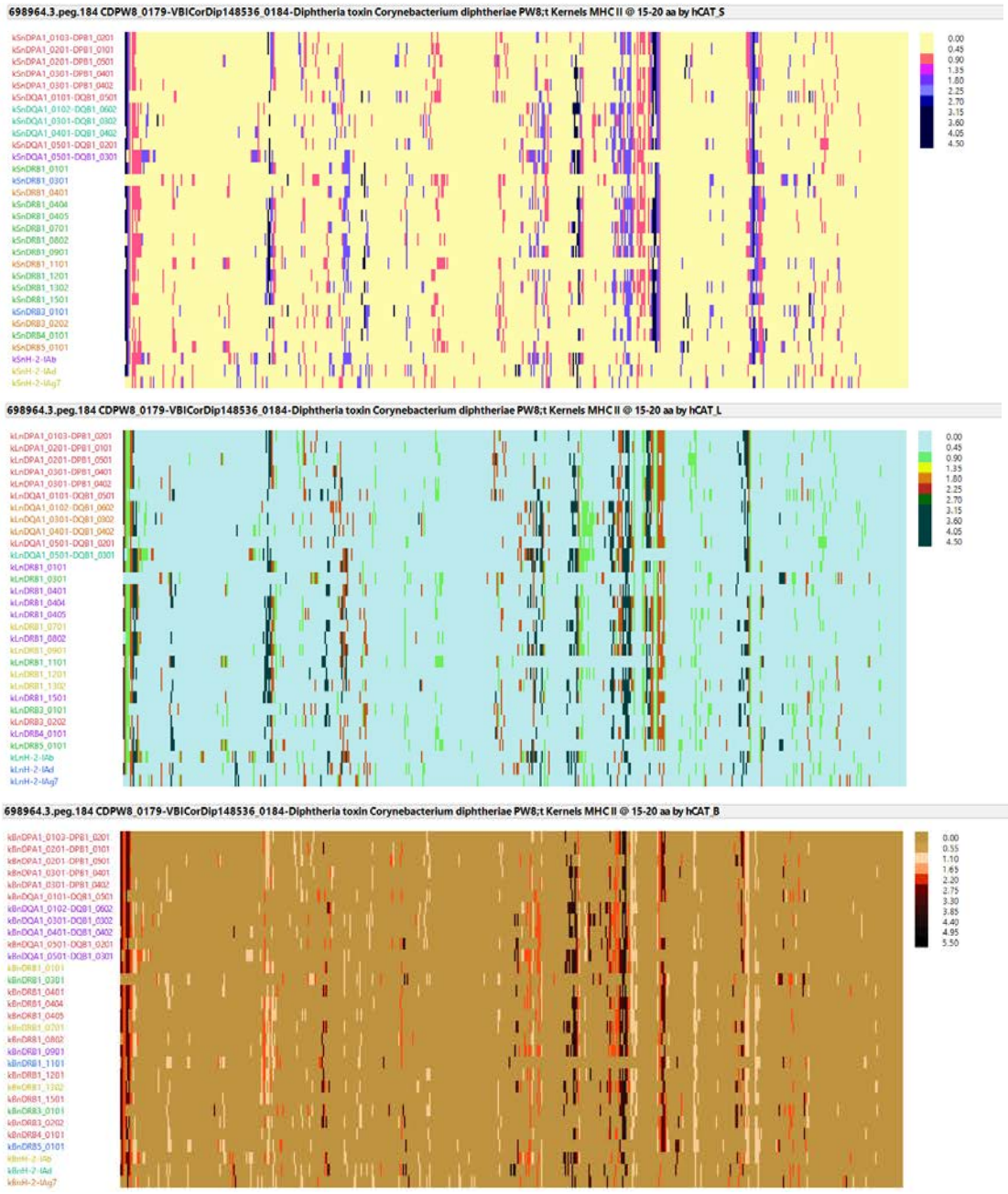
Corynebacterium diphtheriae toxin

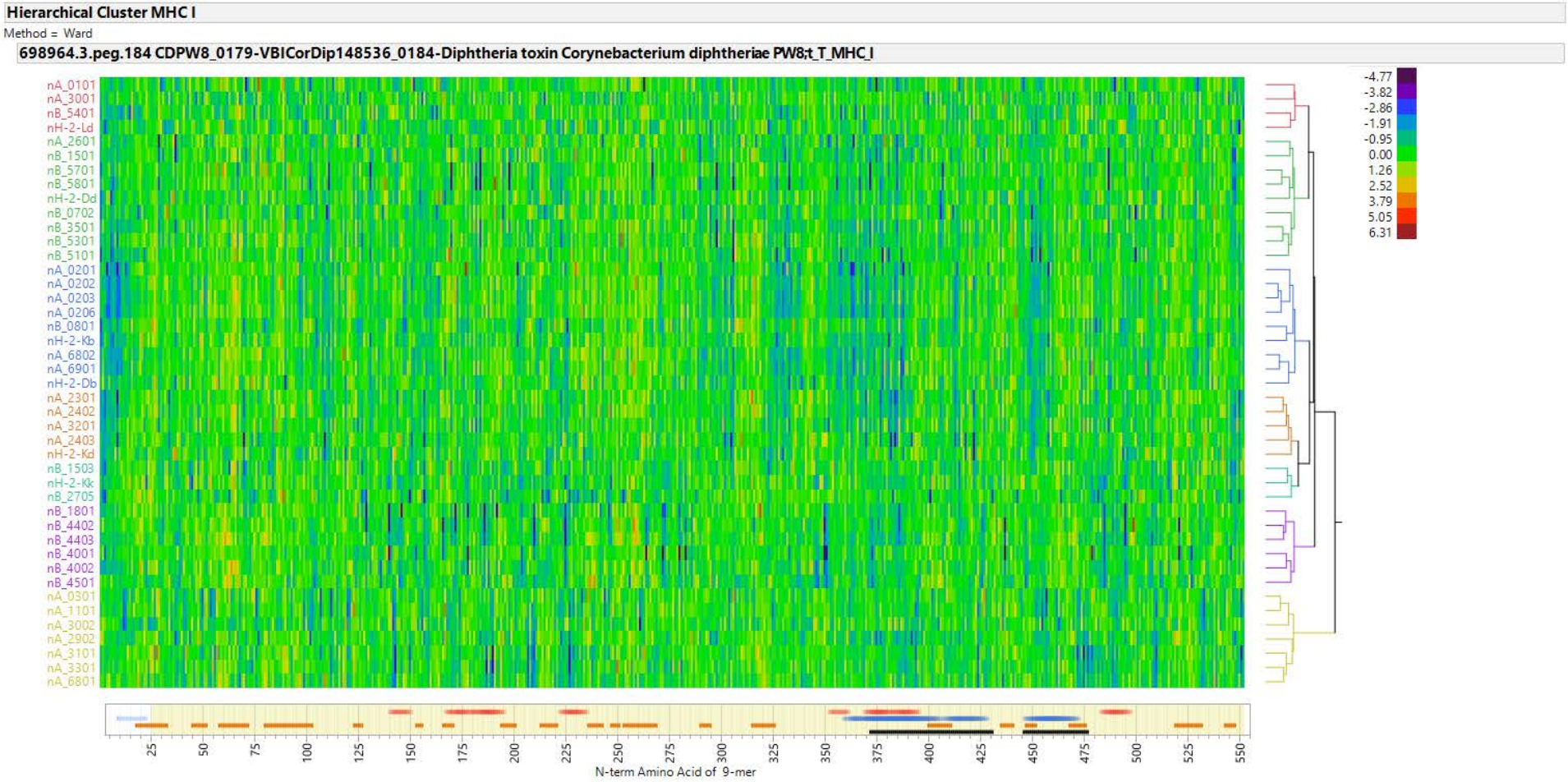
Overlay Plot

698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t




Hierarchical Cluster MHC II
Method = Ward
698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB1_1201
nDRB4_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDRB1_1101
nDRB5_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.72
-2.98
-2.23
-1.49
-0.74
0.00
0.81
1.62
2.42
3.23
4.04
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

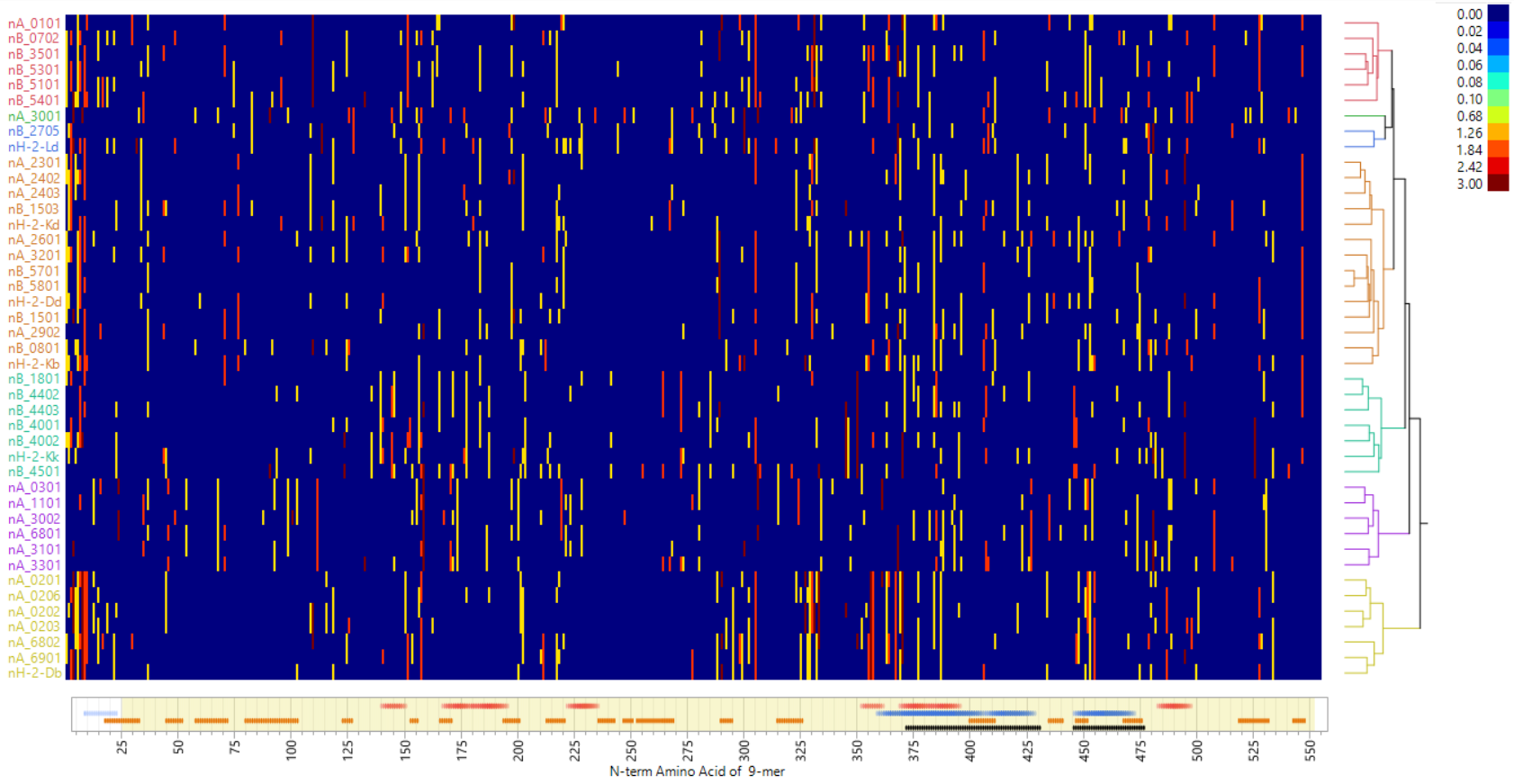

698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t Kernels MHC II @ 15-20 aa by hCAT_S
698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t Kernels MHC II @ 15-20 aa by hCAT_L
698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t Kernels MHC II @ 15-20 aa by hCAT_B

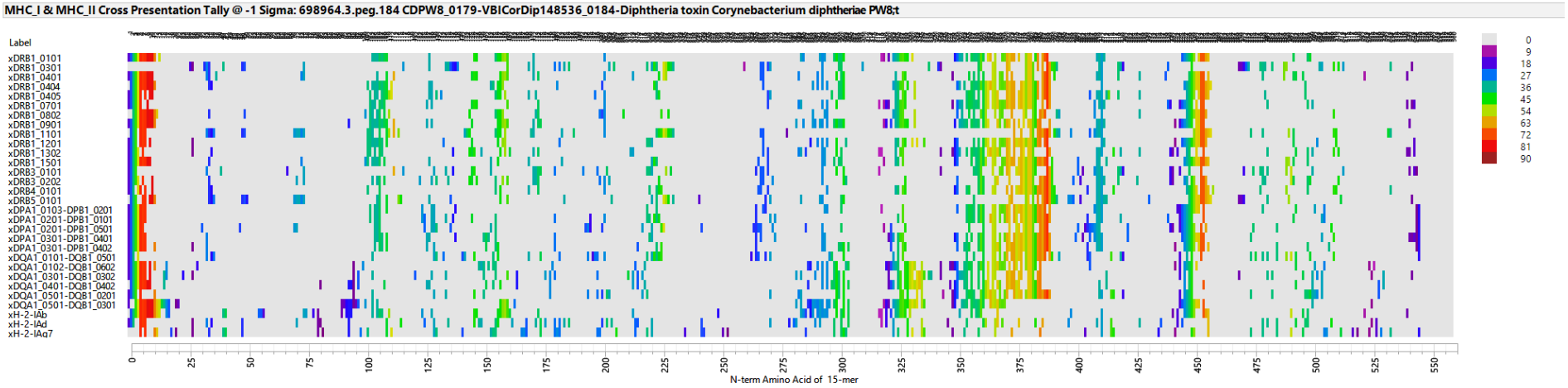

Hierarchical Cluster MHC I
Method = Ward
698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_2601
nB_1501
nB_5701
nB_5801
nH-2-Dd
nB_0702
nB_3501
nB_5301
nB_5101
nA_0201
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_6802
nA_6901
nH-2-Db
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nB_1503
nH-2-Kk
nB_2705
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.77
-3.82
-2.86
-1.91
-0.95
0.00
1.26
2.52
3.79
5.05
6.31
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
N-term Amino Acid of 9-mer

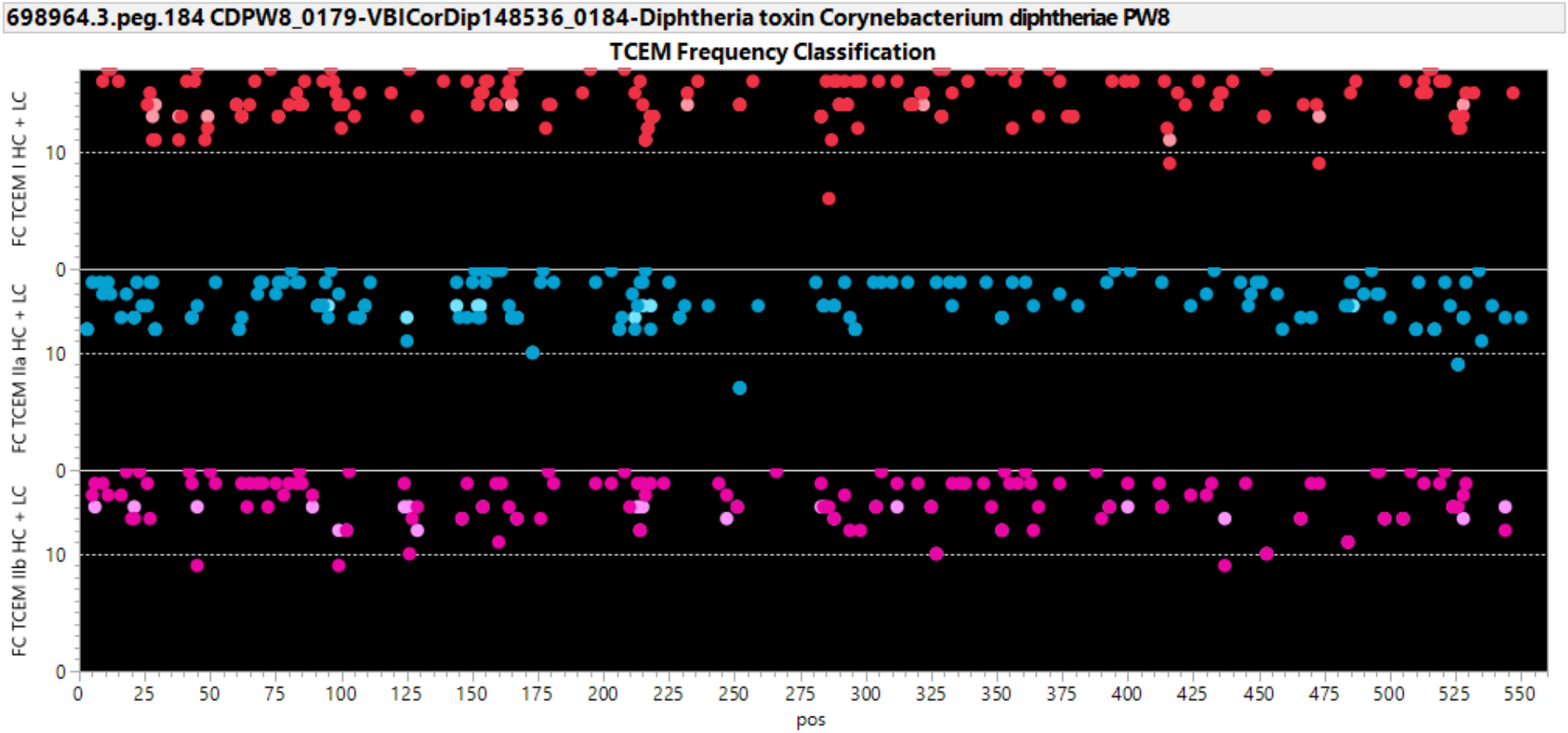
Hierarchical Cluster MHC I

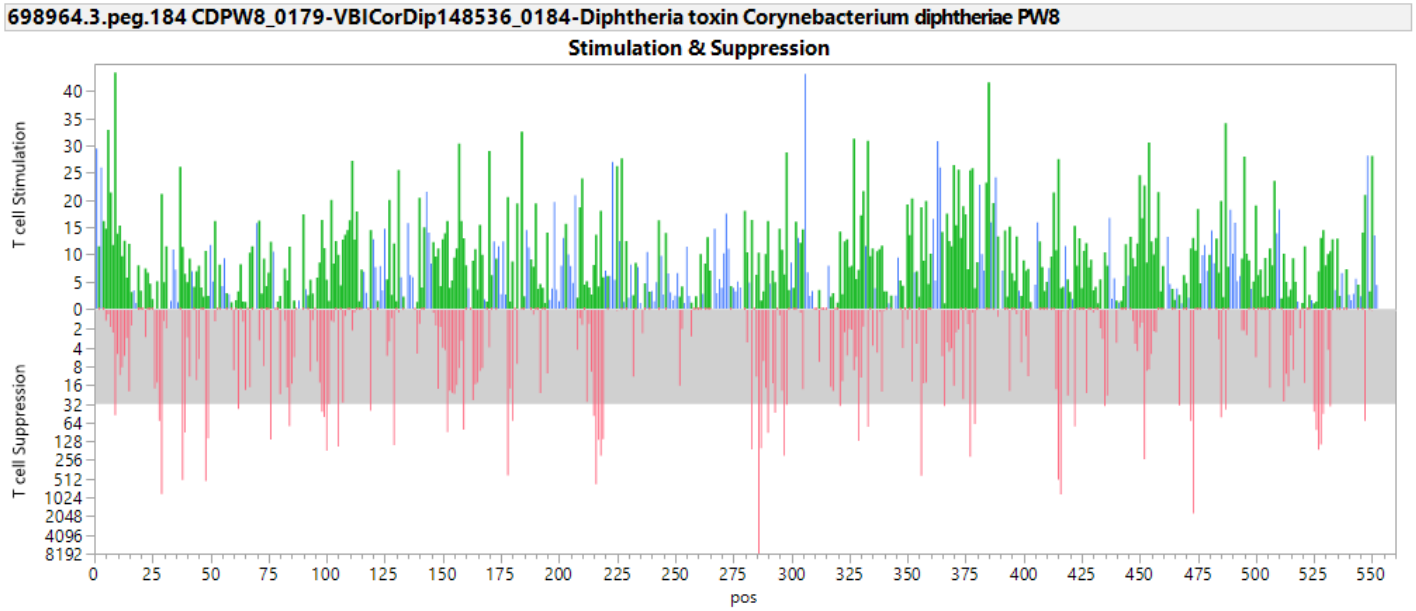
Method = Ward

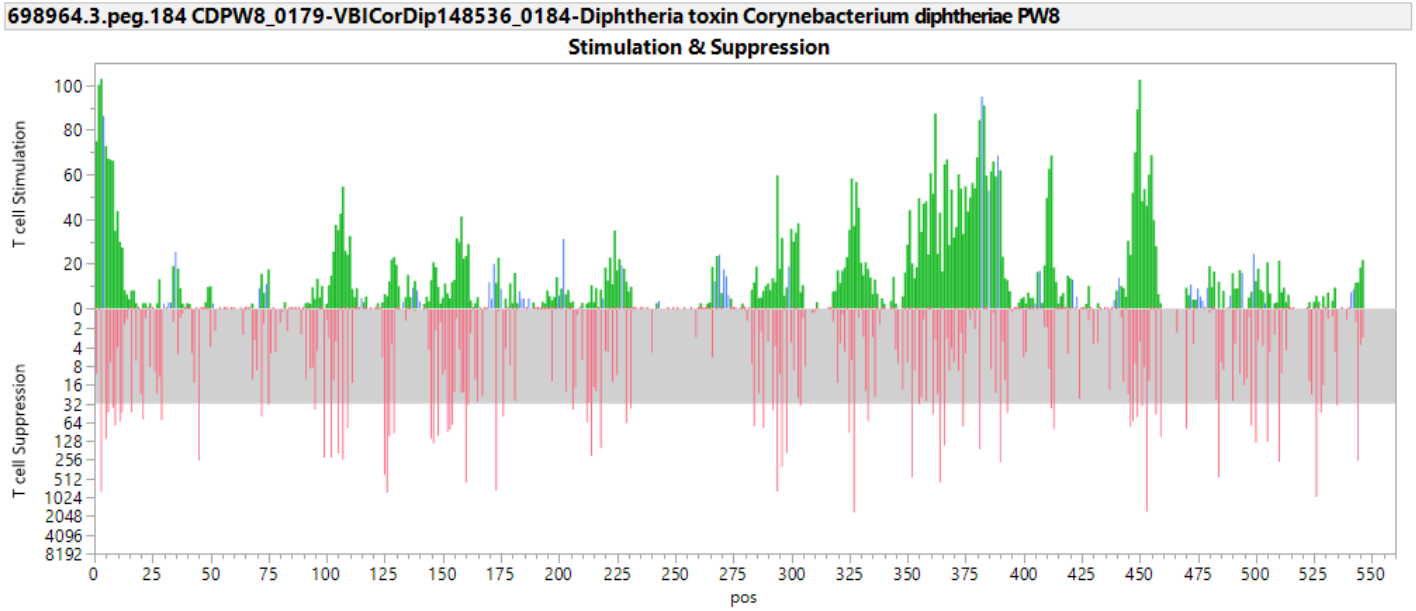
698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t MHC I & Cathepsins consolidated



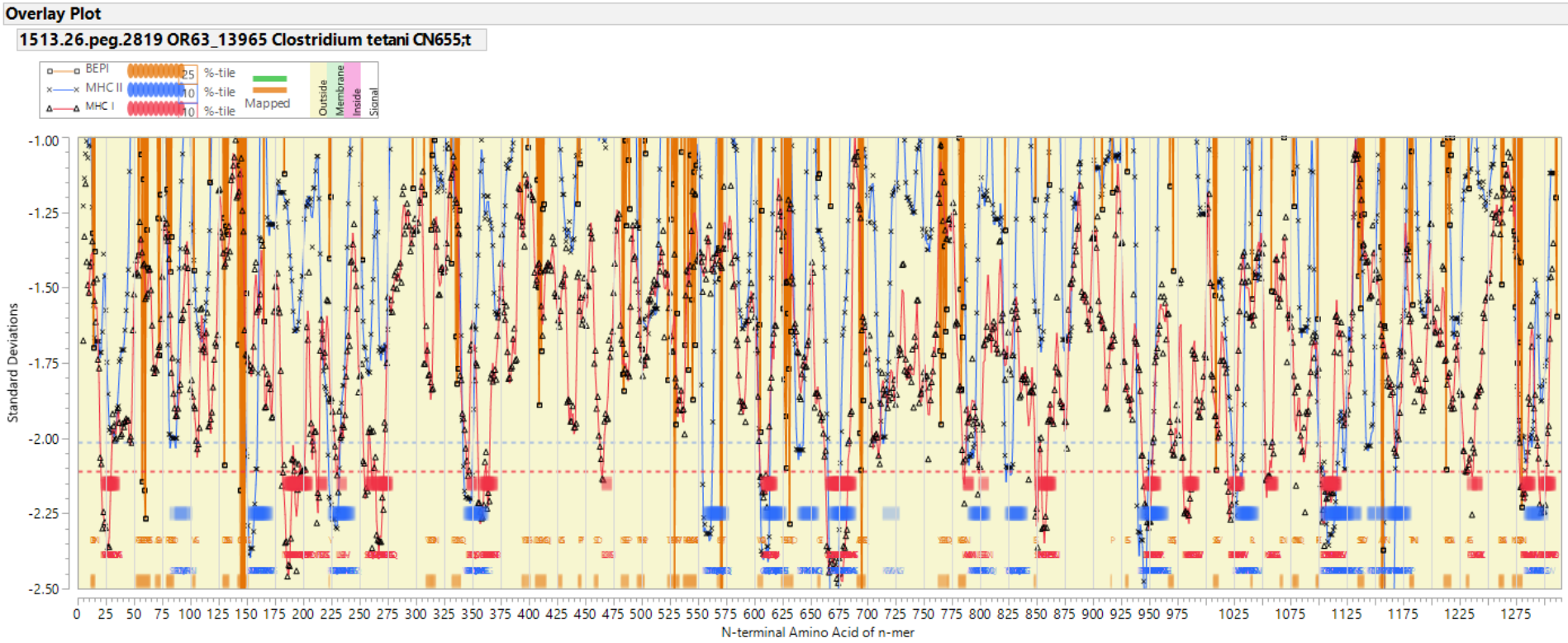

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550
N-term Amino Acid of 15-mer
0
9
18
27
36
45
54
63
72
81
90

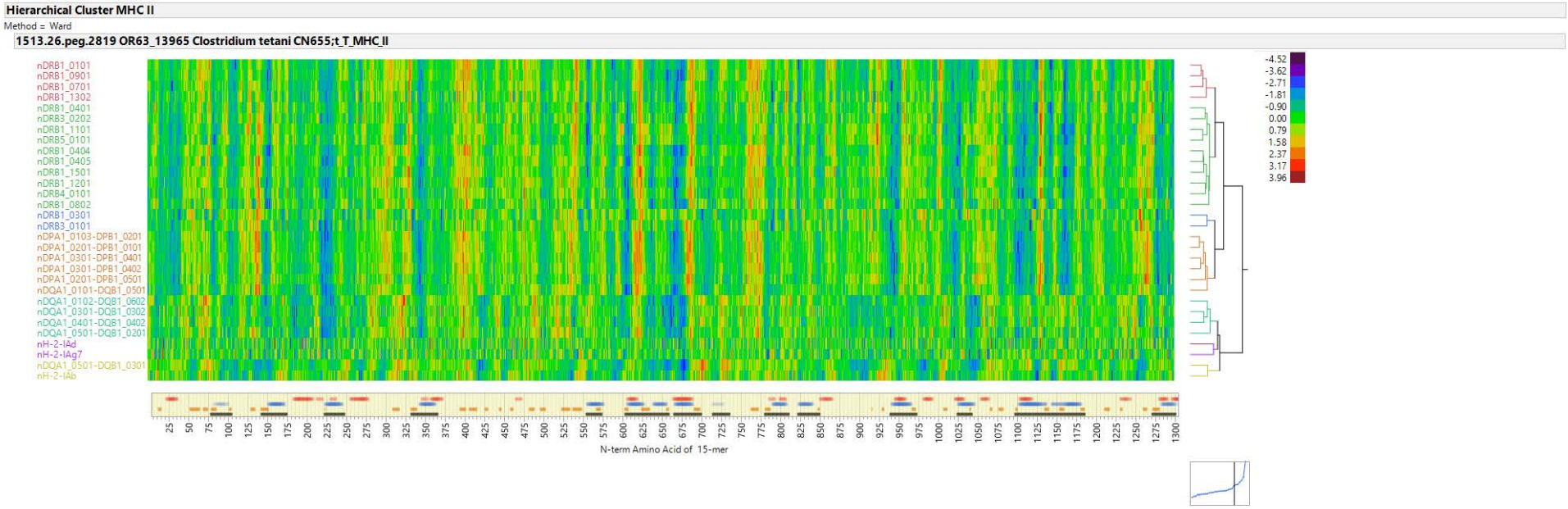

698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550
pos

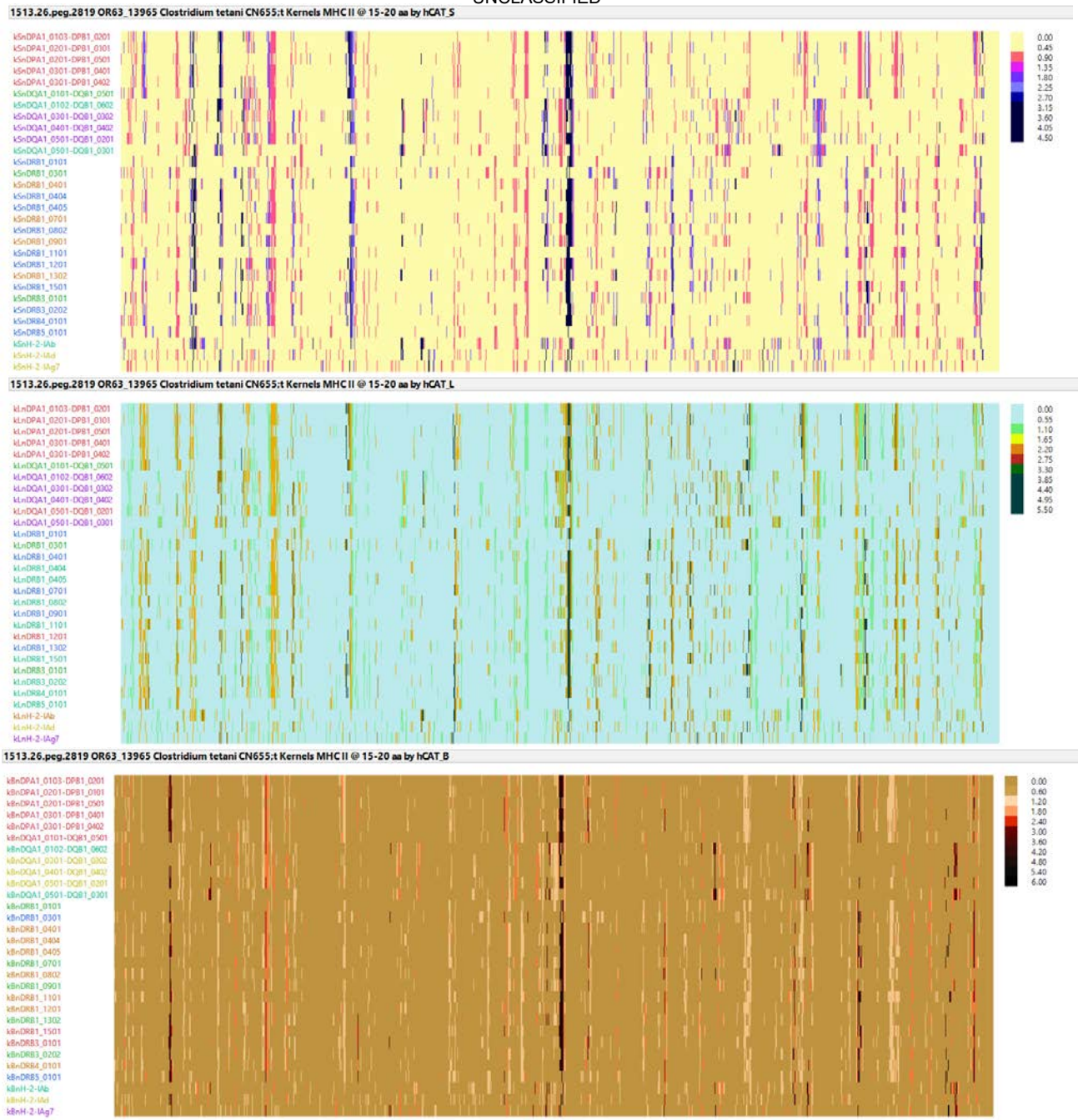

698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

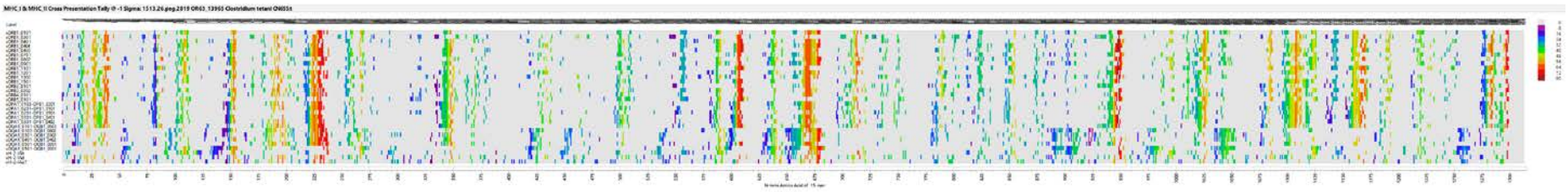

698964.3.peg.184 CDPW8_0179-VBICorDip148536_0184-Diphtheria toxin Corynebacterium diphtheriae PW8
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

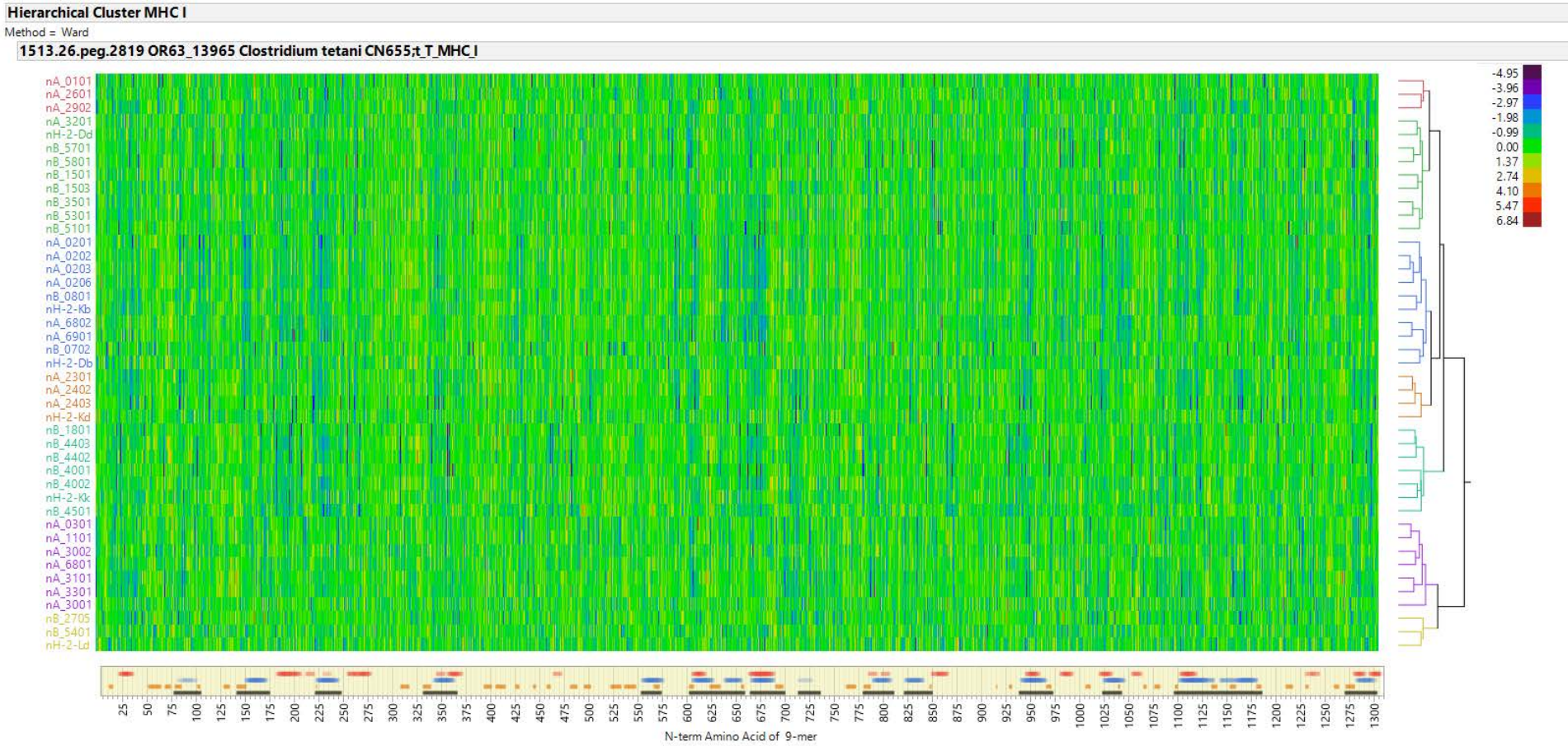
C.tetani toxin

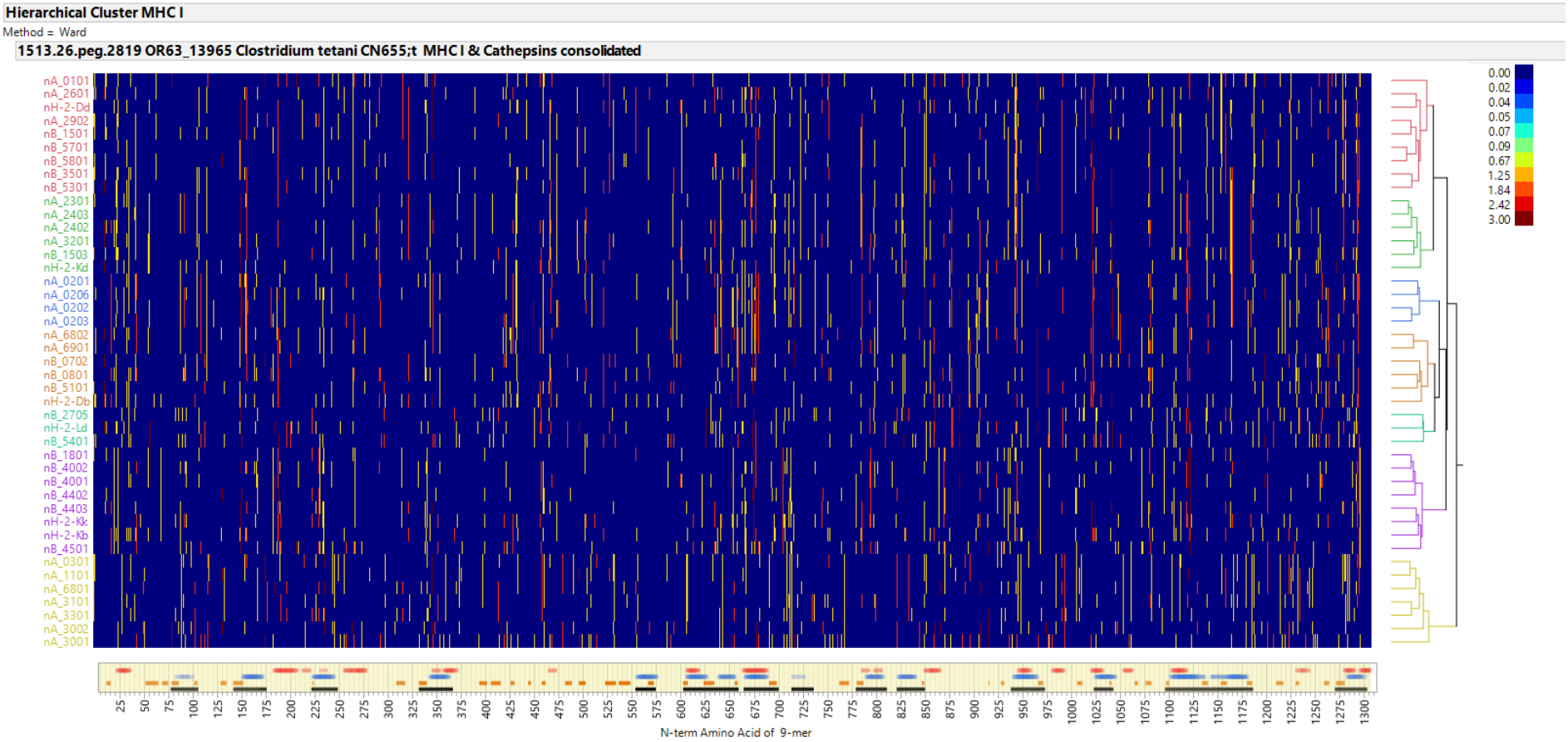

Overlay Plot
1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t
BEPI
MHC II
MHC I
25
10
10
%-tile
%-tile
%-tile
Mapped
Outside
Membrane
Inside
Signal
Standard Deviations
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1025 1075 1125 1175 1225 1275
N-terminal Amino Acid of n-mer

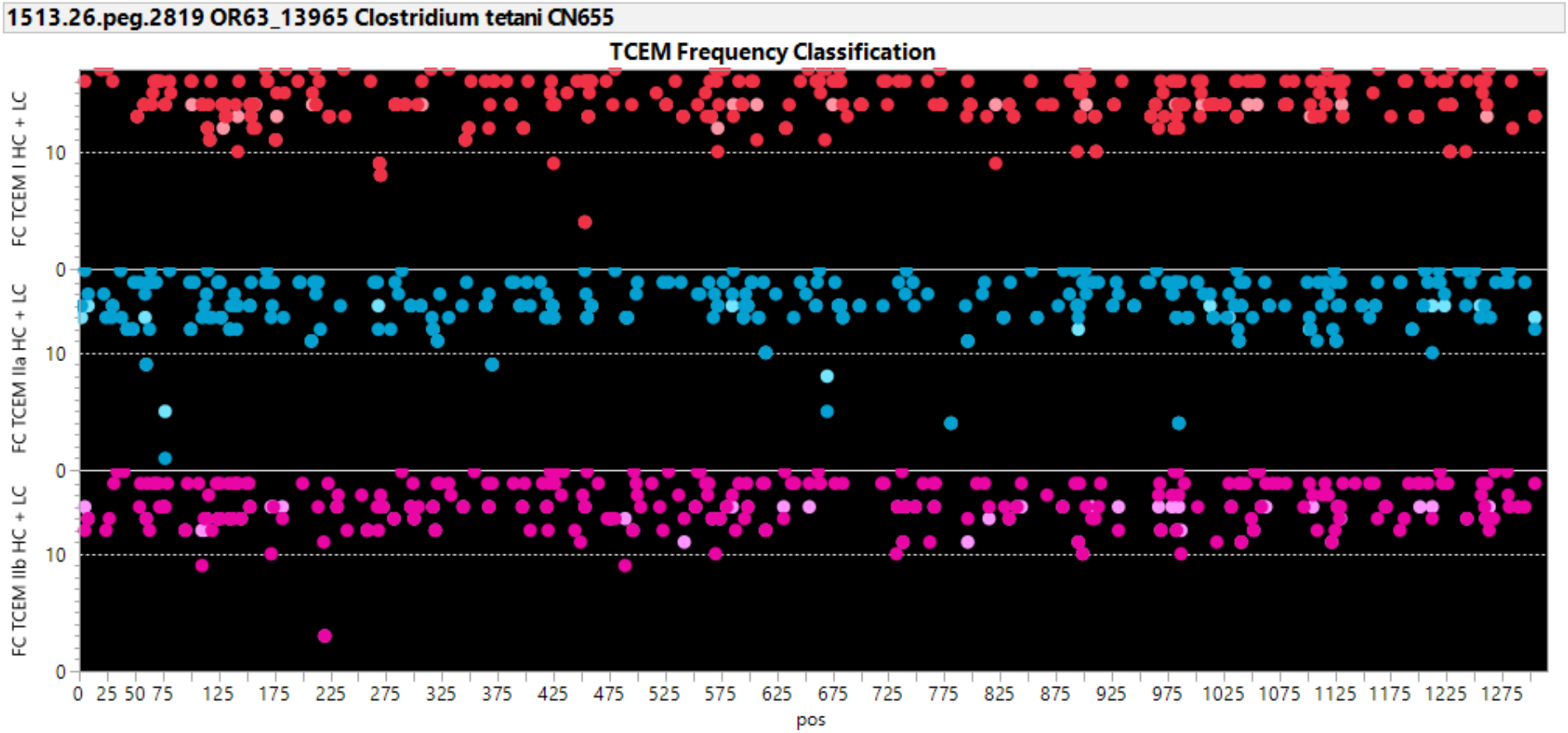

Hierarchical Cluster MHC II
Method = Ward
1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB1_1201
nDRB4_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-4.52
-3.62
-2.71
-1.81
-0.90
0.00
0.79
1.58
2.37
3.17
3.96
25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200 1225 1250 1275 1300
N-term Amino Acid of 15-mer

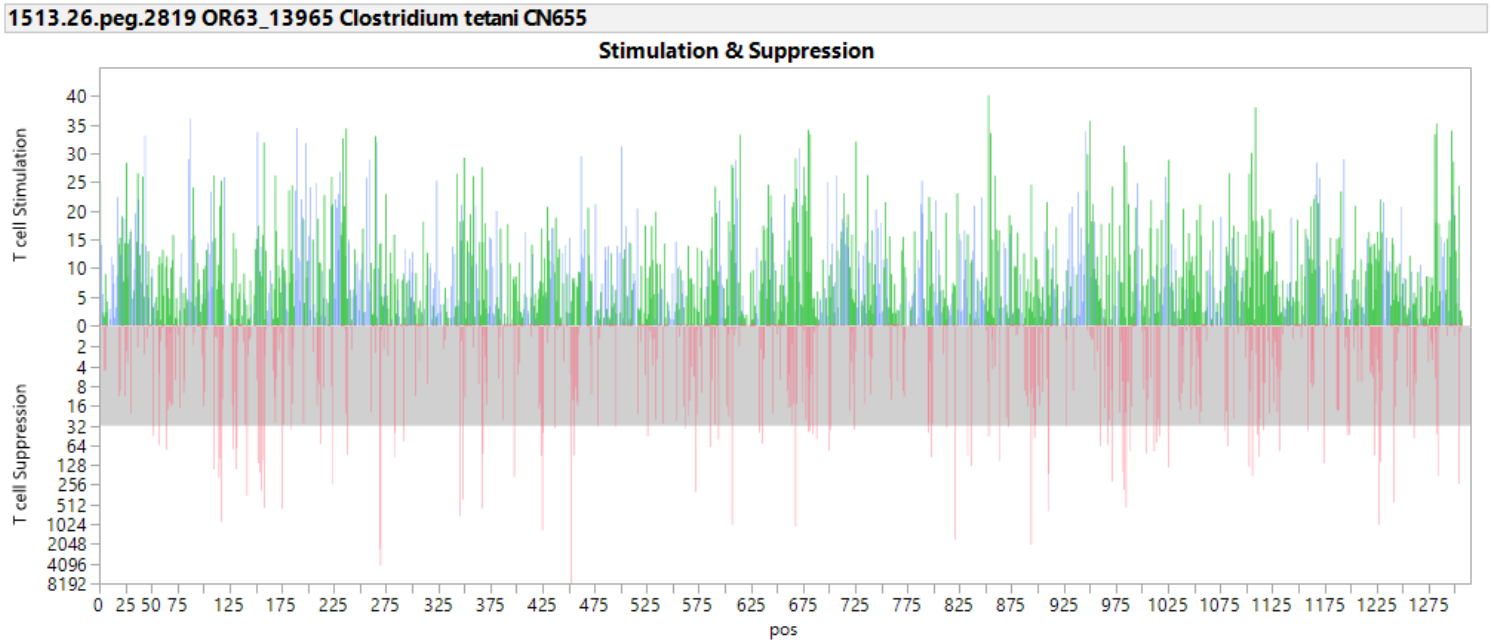

1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.45
0.90
1.35
1.80
2.25
2.70
3.15
3.60
4.05
4.50
1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.55
1.10
1.65
2.20
2.75
3.30
3.85
4.40
4.95
5.50
1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t Kernels MHC II @ 15-20 aa by hCAT_B
kBnDPA1_0103-DPB1_0201
kBnDPA1_0201-DPB1_0101
kBnDPA1_0201-DPB1_0501
kBnDPA1_0301-DPB1_0401
kBnDPA1_0301-DPB1_0402
kBnDQA1_0101-DQB1_0501
kBnDQA1_0102-DQB1_0602
kBnDQA1_0301-DQB1_0302
kBnDQA1_0401-DQB1_0402
kBnDQA1_0501-DQB1_0201
kBnDQA1_0501-DQB1_0301
kBnDRB1_0101
kBnDRB1_0301
kBnDRB1_0401
kBnDRB1_0404
kBnDRB1_0405
kBnDRB1_0701
kBnDRB1_0802
kBnDRB1_0901
kBnDRB1_1101
kBnDRB1_1201
kBnDRB1_1302
kBnDRB1_1501
kBnDRB3_0101
kBnDRB3_0202
kBnDRB4_0101
kBnDRB5_0101
kBnH-2-IAb
kBnH-2-IAd
kBnH-2-IAg7
0.00
0.60
1.20
1.80
2.40
3.00
3.60
4.20
4.80
5.40
6.00

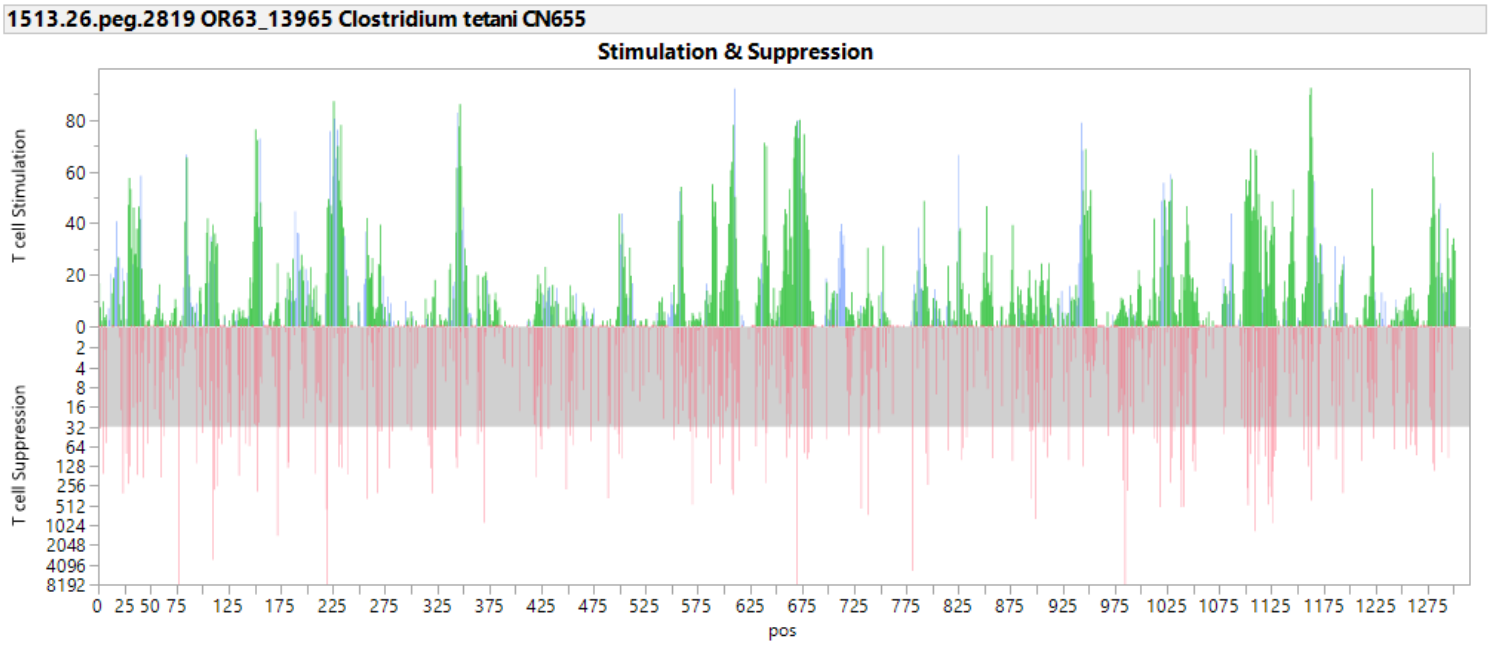



Hierarchical Cluster MHC I
Method = Ward
1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t_T_MHC_I
nA_0101
nA_2601
nA_2902
nA_3201
nH-2-Dd
nB_5701
nB_5801
nB_1501
nB_1503
nB_3501
nB_5301
nB_5101
nA_0201
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_6802
nA_6901
nB_0702
nH-2-Db
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_3001
nB_2705
nB_5401
nH-2-Ld
-4.95
-3.96
-2.97
-1.98
-0.99
0.00
1.37
2.74
4.10
5.47
6.84
25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200 1225 1250 1275 1300
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
1513.26.peg.2819 OR63_13965 Clostridium tetani CN655;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nH-2-Dd
nA_2902
nB_1501
nB_5701
nB_5801
nB_3501
nB_5301
nA_2301
nA_2403
nA_2402
nA_3201
nB_1503
nH-2-Kd
nA_0201
nA_0206
nA_0202
nA_0203
nA_6802
nA_6901
nB_0702
nB_0801
nB_5101
nH-2-Db
nB_2705
nH-2-Ld
nB_5401
nB_1801
nB_4002
nB_4001
nB_4402
nB_4403
nH-2-Kk
nH-2-Kb
nB_4501
nA_0301
nA_1101
nA_6801
nA_3101
nA_3301
nA_3002
nA_3001
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00
25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675 700 725 750 775 800 825 850 875 900 925 950 975 1000 1025 1050 1075 1100 1125 1150 1175 1200 1225 1250 1275 1300
N-term Amino Acid of 9-mer


1513.26.peg.2819 OR63_13965 Clostridium tetani CN655
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0 25 50 75 125 175 225 275 325 375 425 475 525 575 625 675 725 775 825 875 925 975 1025 1075 1125 1175 1225 1275
pos


1513.26.peg.2819 OR63_13965 Clostridium tetani CN655
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos


1513.26.peg.2819 OR63_13965 Clostridium tetani CN655
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

- C. tetani toxin has a common/highly bound MHC Ii motif at index position 220 which leads to a suppressive index of 241,241 indicating it is a likely Treg inducer.
- *Corynebacterium diphtheriae* toxin does not have such extreme motifs.

ioGenetics®

Addendum 7
Ebolavirus

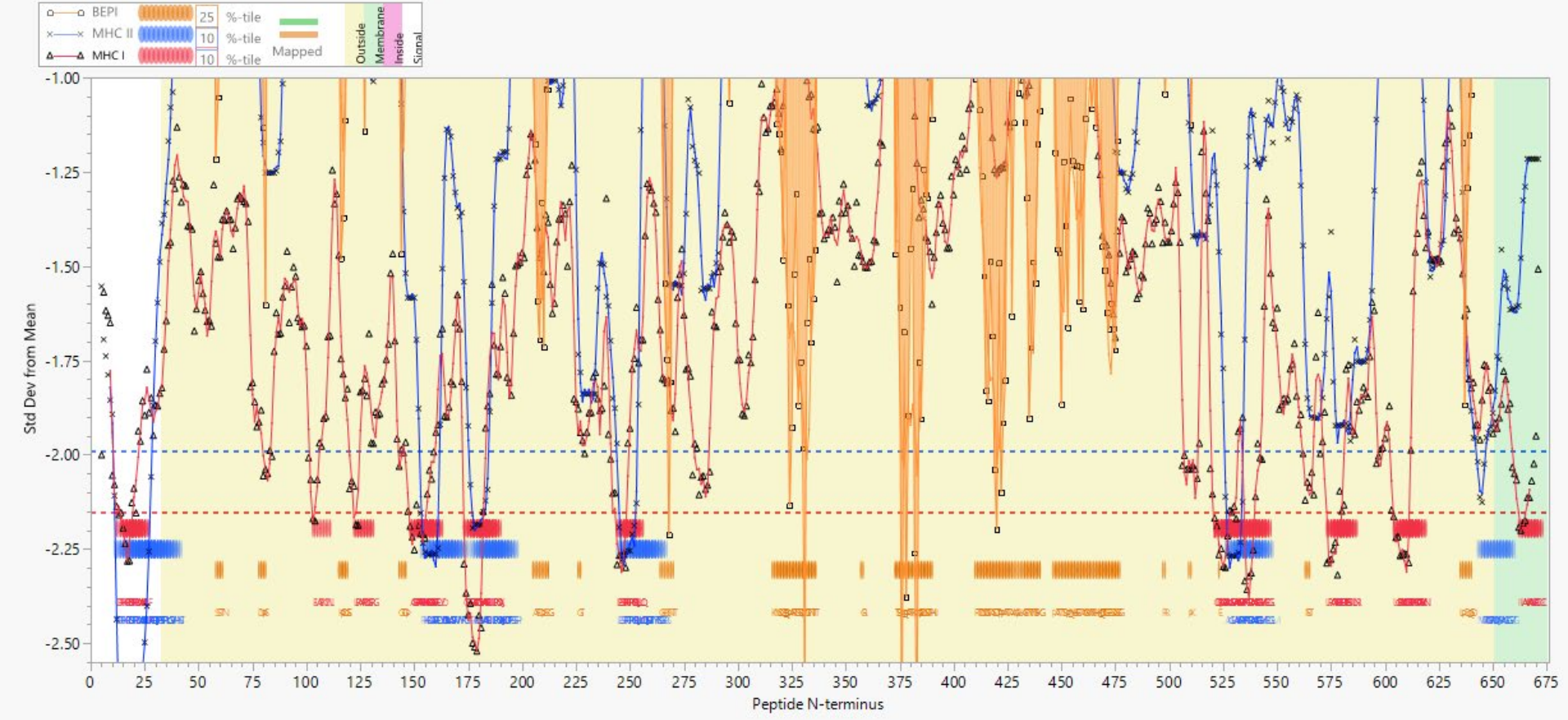
Envelope glycoprotein of Kikwit strain

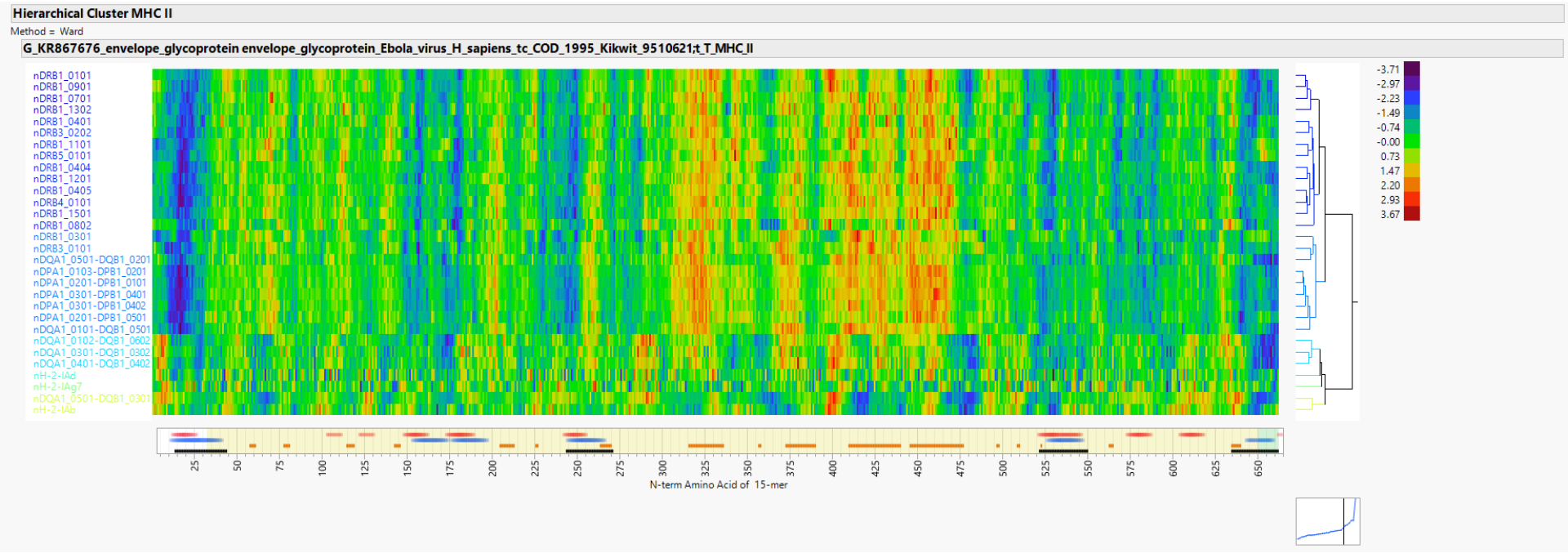
As used in ERBEVO®

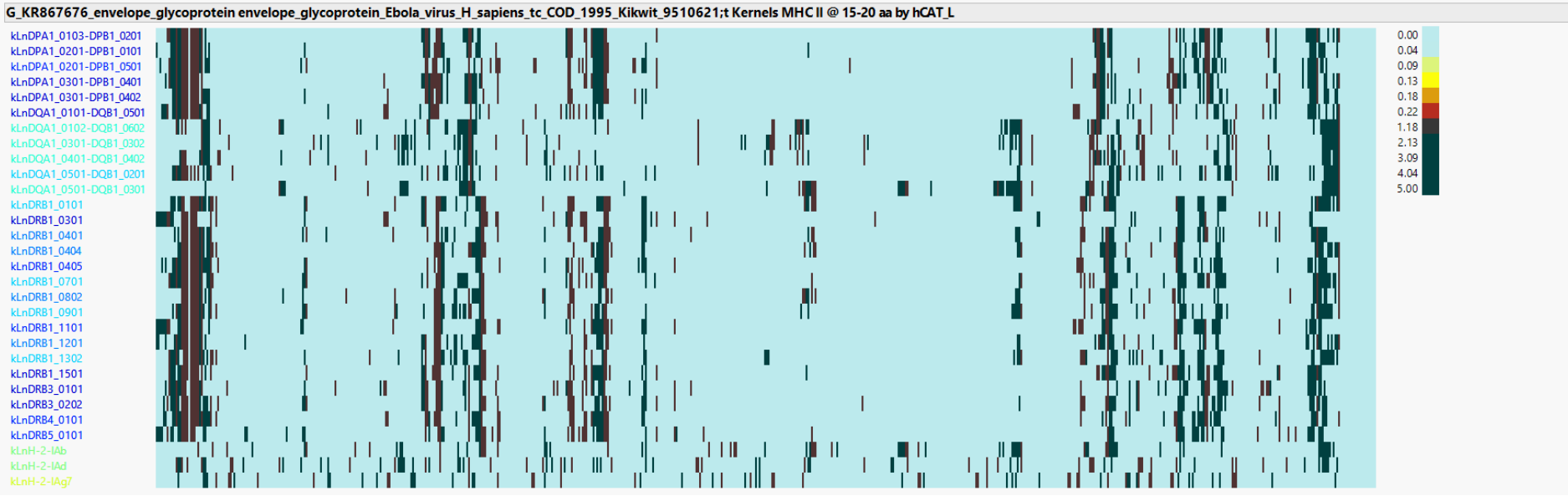
gi KR867676

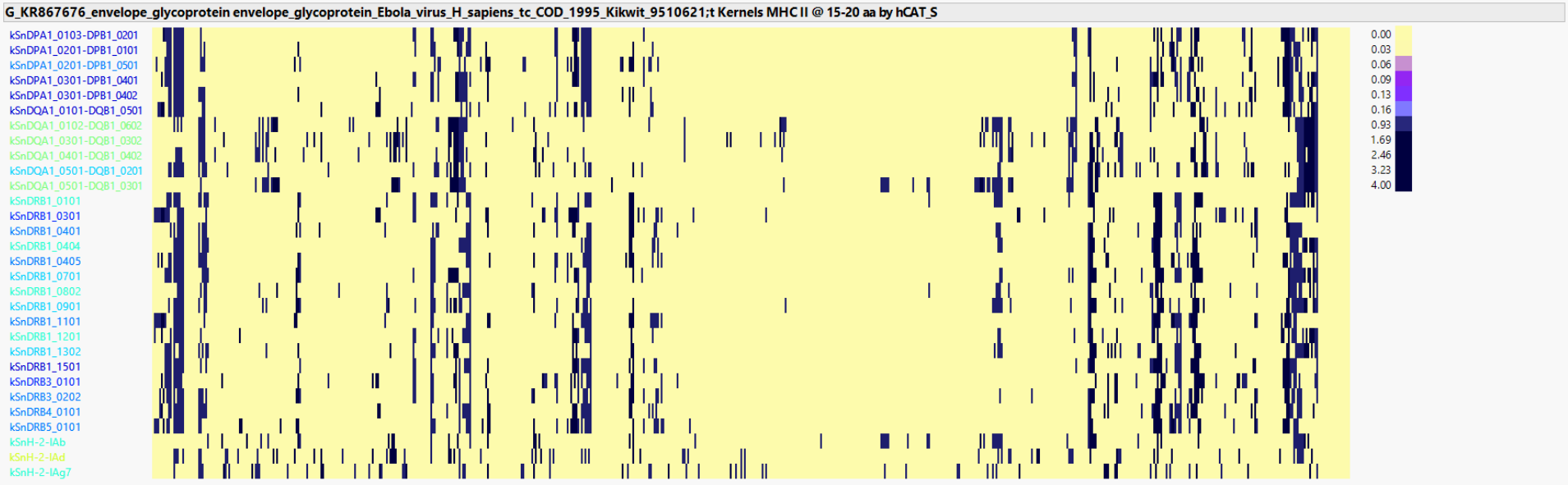
Population Permuted Metrics

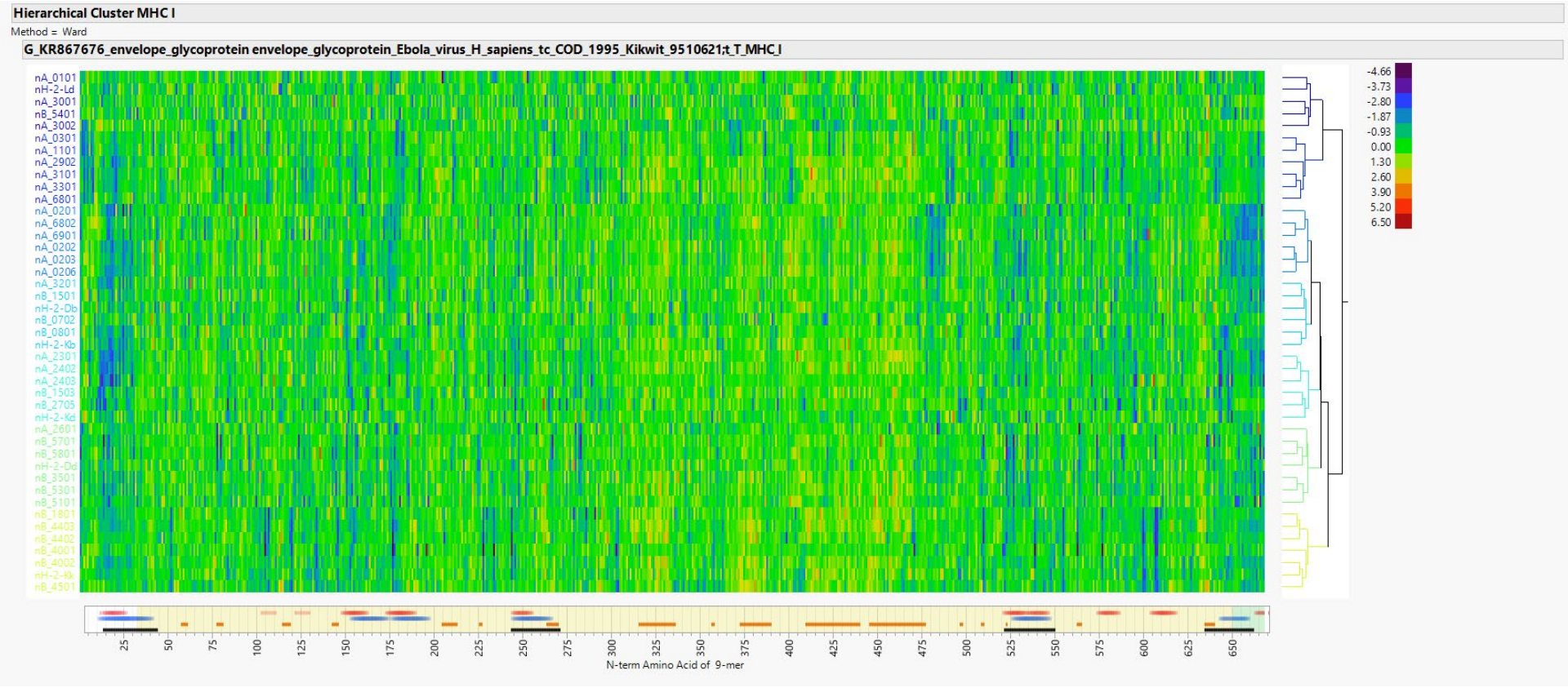
G_KR867676_envelope_glycoprotein
envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t

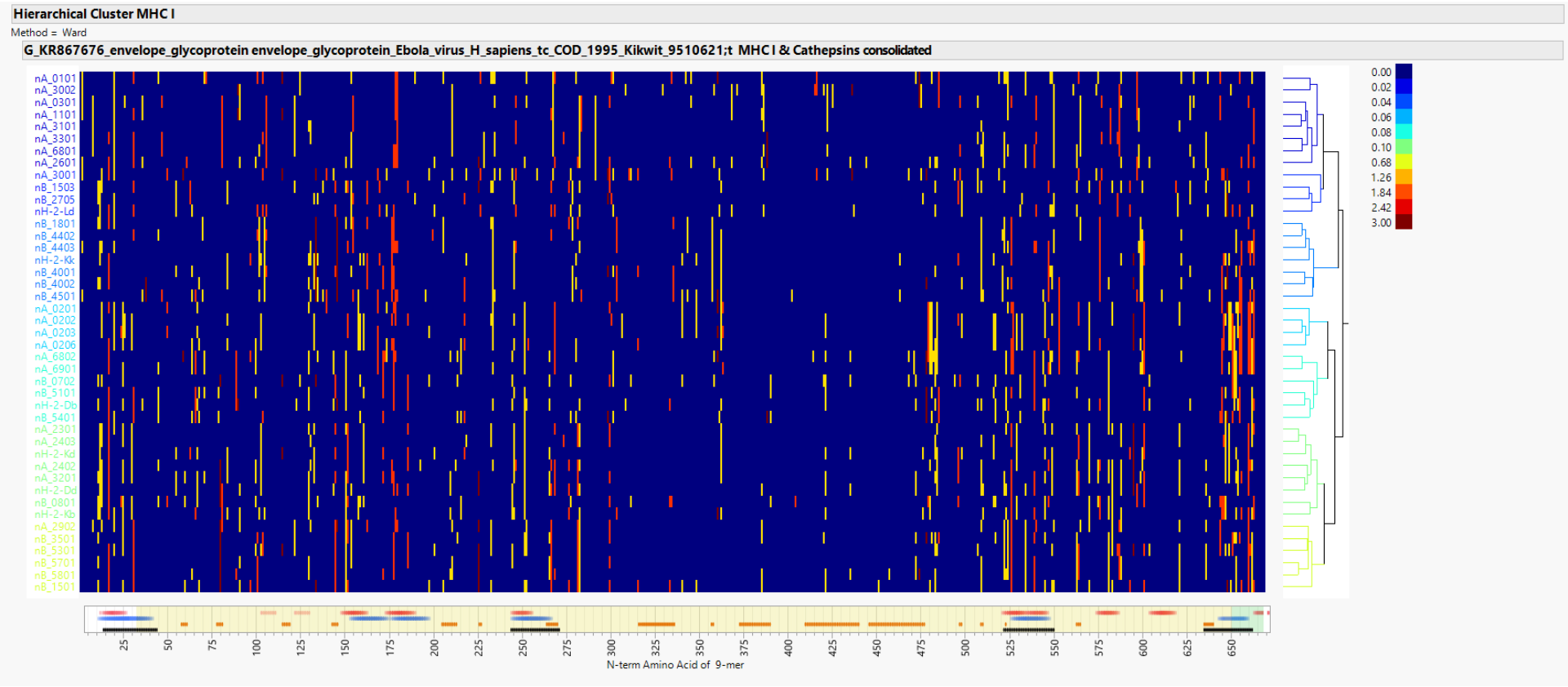


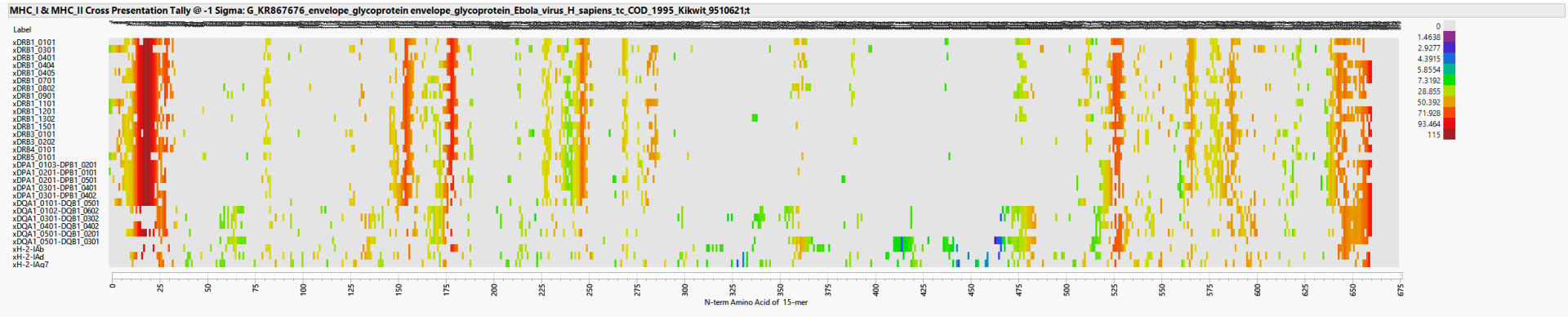

Hierarchical Cluster MHC II
Method = Ward
G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t T_MHC II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_1201
nDRB1_0405
nDRB4_0101
nDRB1_1501
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.71
-2.97
-2.23
-1.49
-0.74
-0.00
0.73
1.47
2.20
2.93
3.67
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
N-term Amino Acid of 15-mer

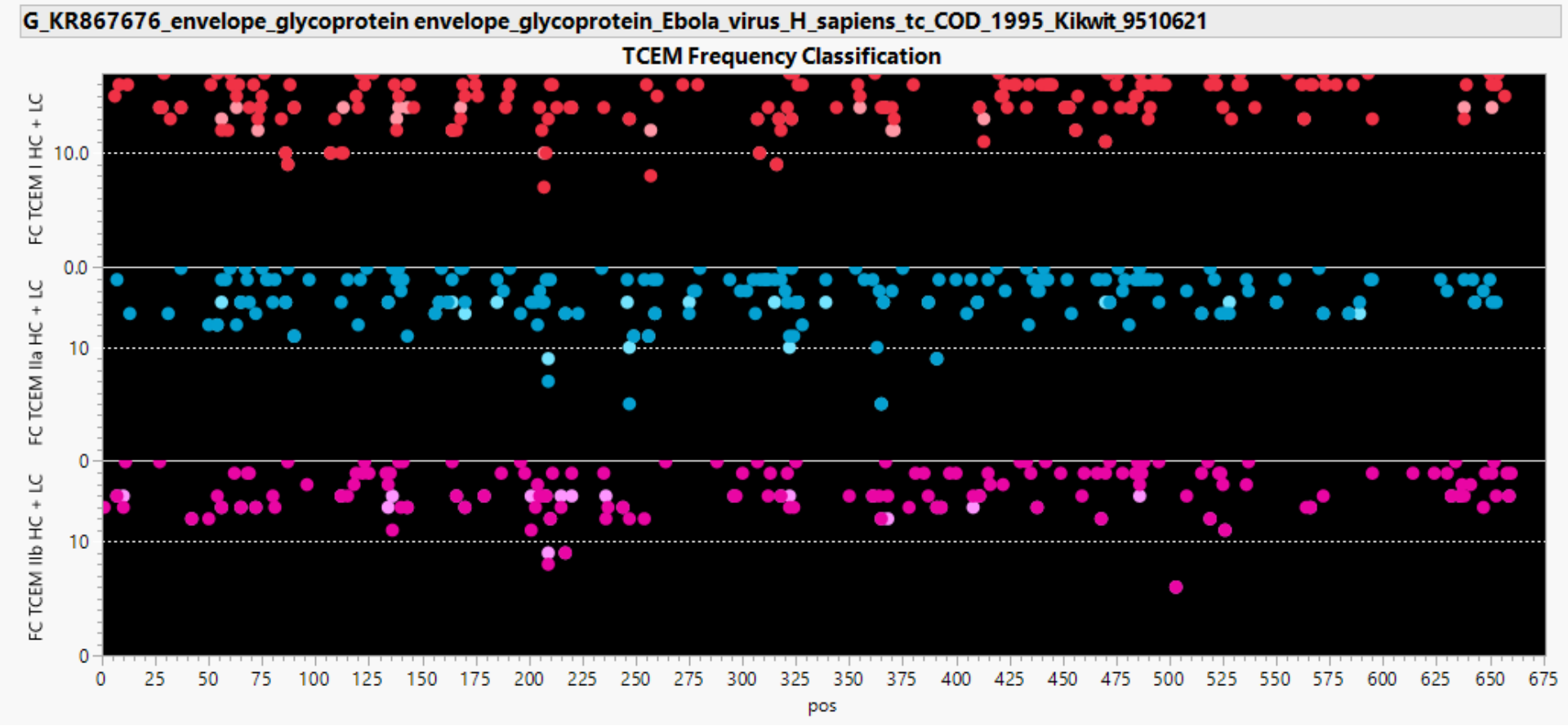

G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.04
0.09
0.13
0.18
0.22
1.18
2.13
3.09
4.04
5.00

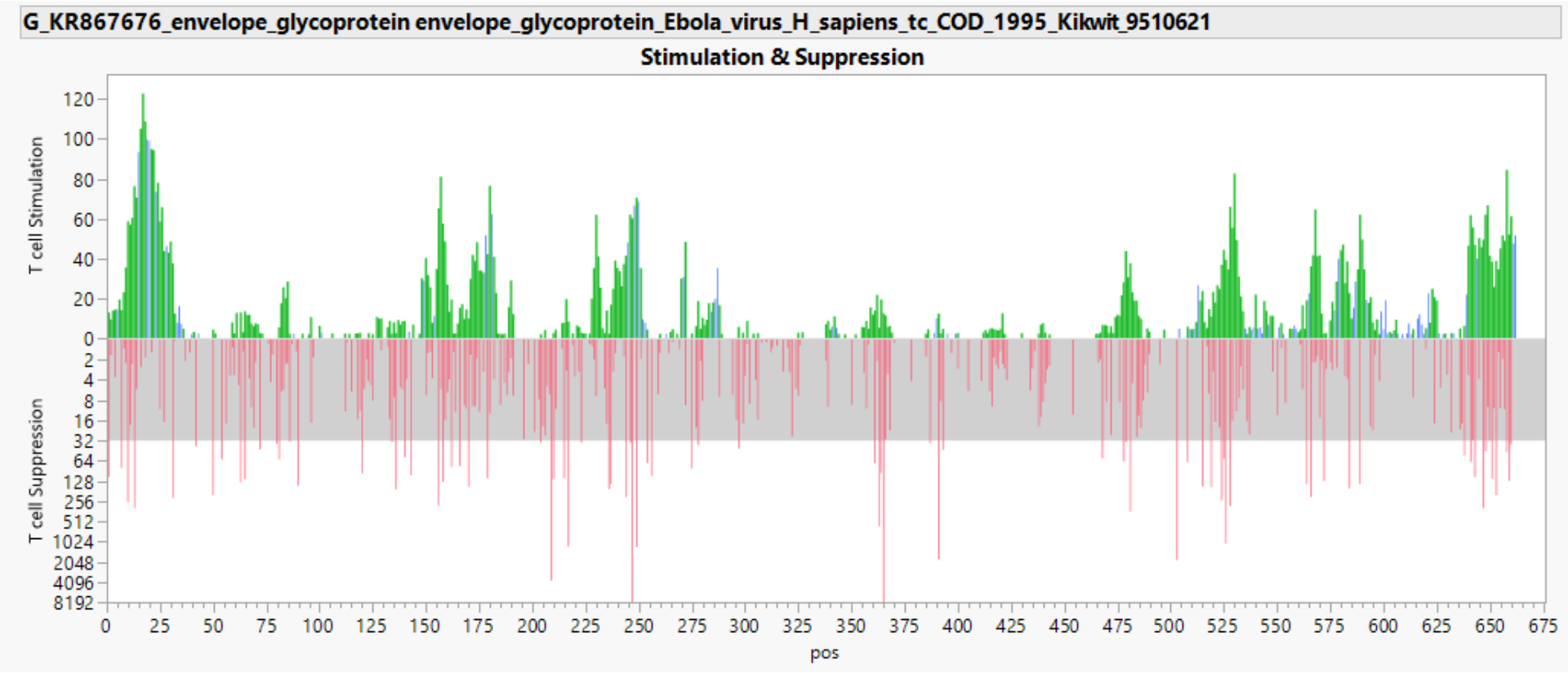

G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.06
0.09
0.13
0.16
0.93
1.69
2.46
3.23
4.00


Hierarchical Cluster MHC I
Method = Ward
G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t_T_MHC_I
nA_0101
nH-2-Ld
nA_3001
nB_5401
nA_3002
nA_0301
nA_1101
nA_2902
nA_3101
nA_3301
nA_6801
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_3201
nB_1501
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nA_2601
nB_5701
nB_5801
nH-2-Dd
nB_3501
nB_5301
nB_5101
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
-4.66
-3.73
-2.80
-1.87
-0.93
0.00
1.30
2.60
3.90
5.20
6.50
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t MHC I & Cathepsins consolidated
nA_0101
nA_3002
nA_0301
nA_1101
nA_3101
nA_3301
nA_6801
nA_2601
nA_3001
nB_1503
nB_2705
nH-2-Ld
nB_1801
nB_4402
nB_4403
nH-2-Kk
nB_4001
nB_4002
nB_4501
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nB_5101
nH-2-Db
nB_5401
nA_2301
nA_2403
nH-2-Kd
nA_2402
nA_3201
nH-2-Dd
nB_0801
nH-2-Kb
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nB_1501
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
675
N-term Amino Acid of 15-mer
0
1.4638
2.9277
4.3915
5.8554
7.3192
28.855
50.392
71.928
93.464
115


G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10.0
0.0
10
0
10
0
0 25 50 75 100 125 150 175 200 225 250 275 300 325 350 375 400 425 450 475 500 525 550 575 600 625 650 675
pos


G_KR867676_envelope_glycoprotein envelope_glycoprotein_Ebola_virus_H_sapiens_tc_COD_1995_Kikwit_9510621
Stimulation & Suppression
T cell Stimulation
120
100
80
60
40
20
0
T cell Suppression
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
550
575
600
625
650
675
pos

ioGenetics®

Addendum 8
Hepatitis A

Proteins derived from polyprotein of strain HM-175

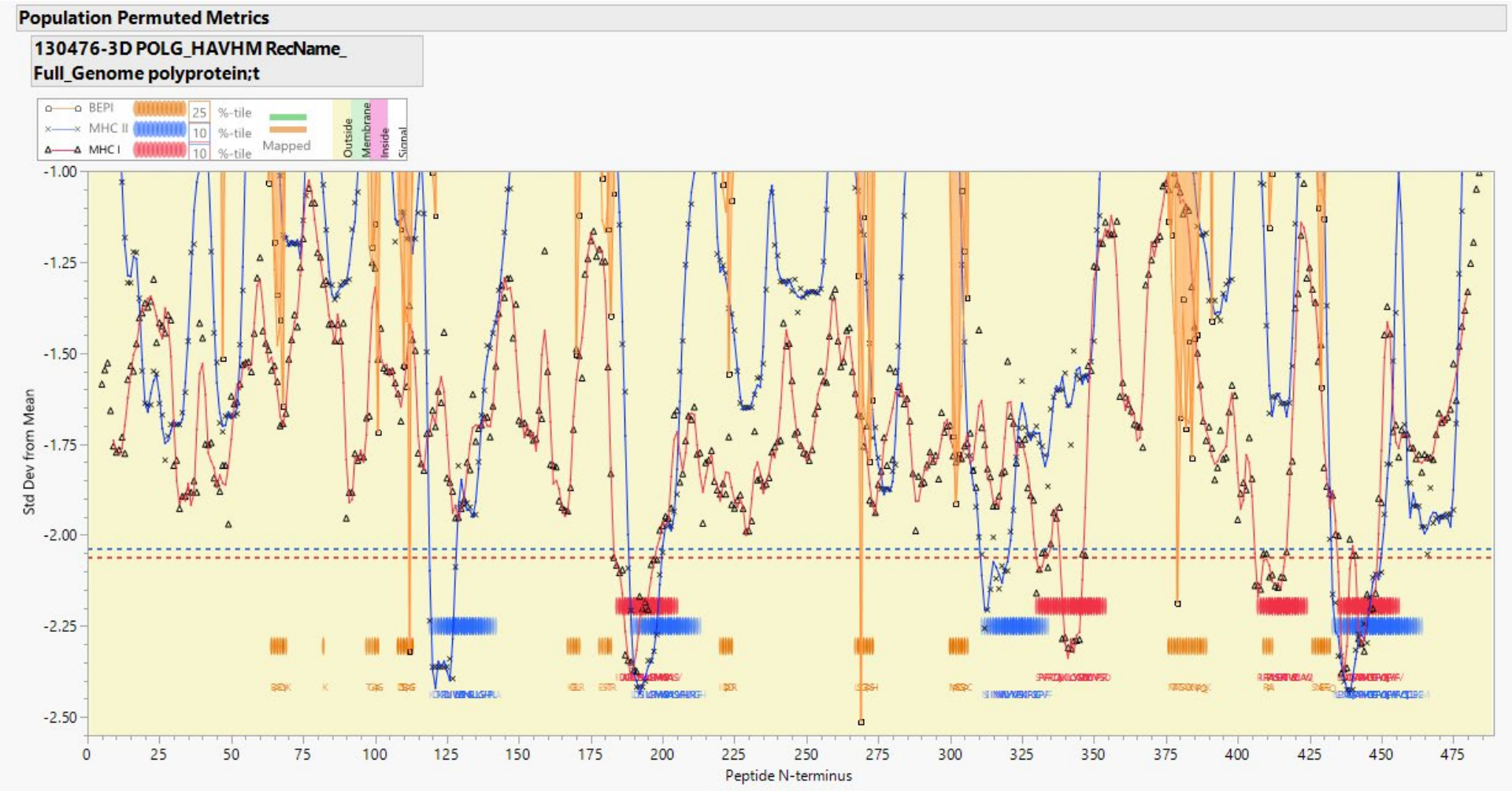
Havrix® comprises inactivated HM-175

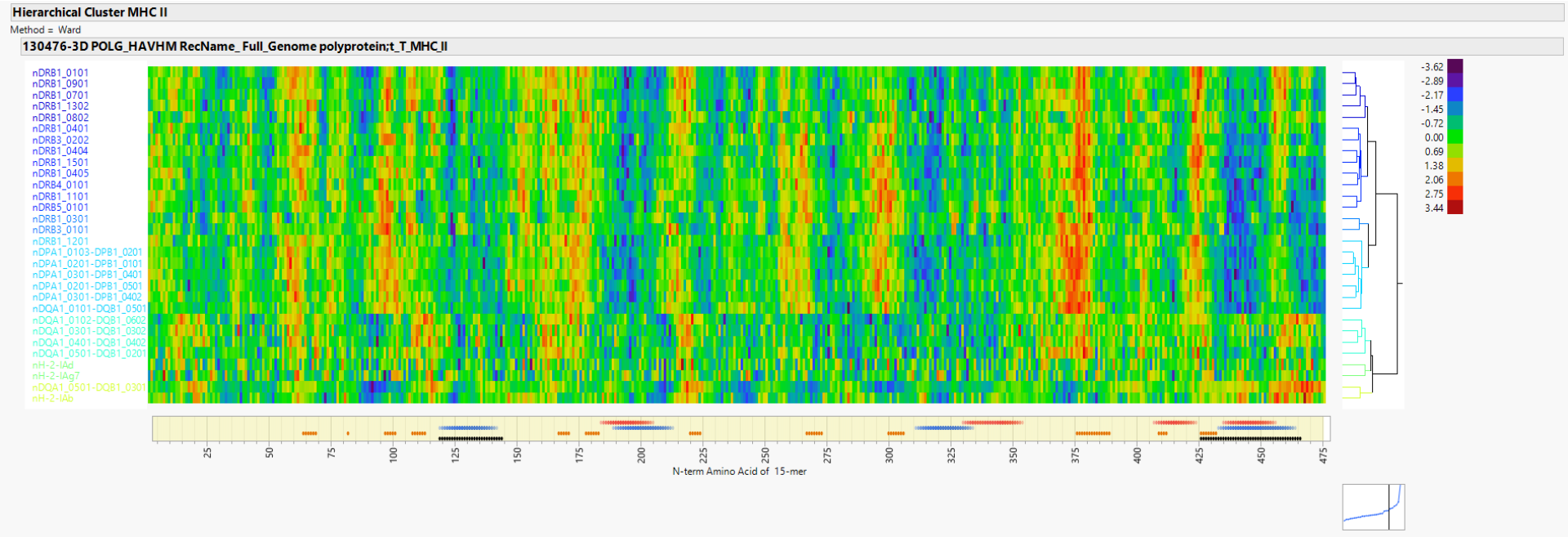
RefSeq 13076

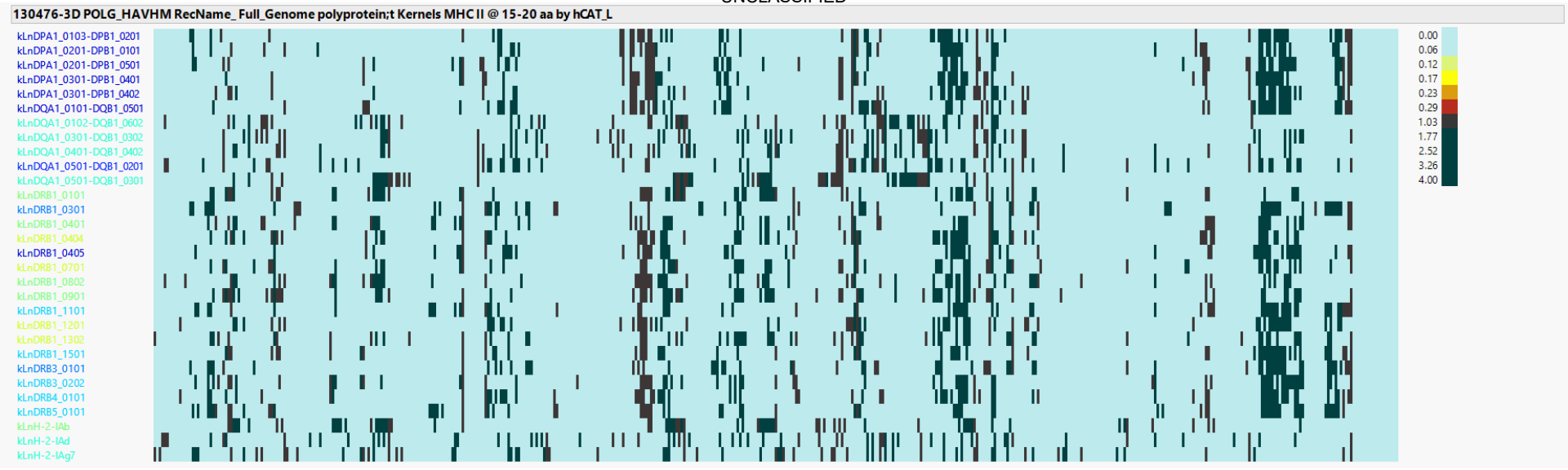
Protein 3D

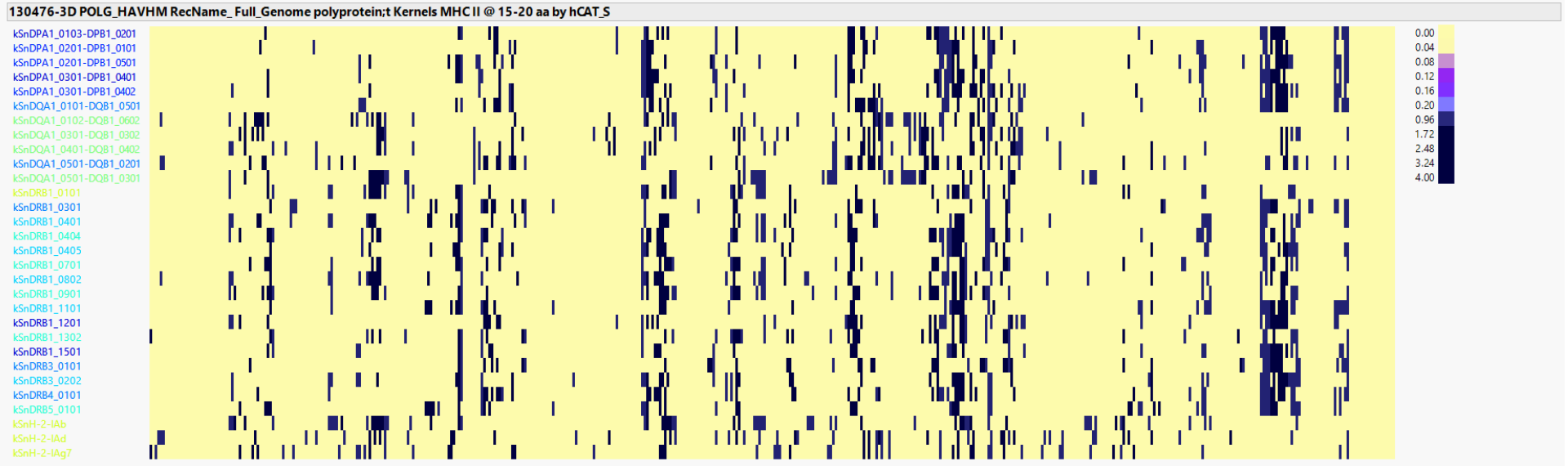
Population Permuted Metrics

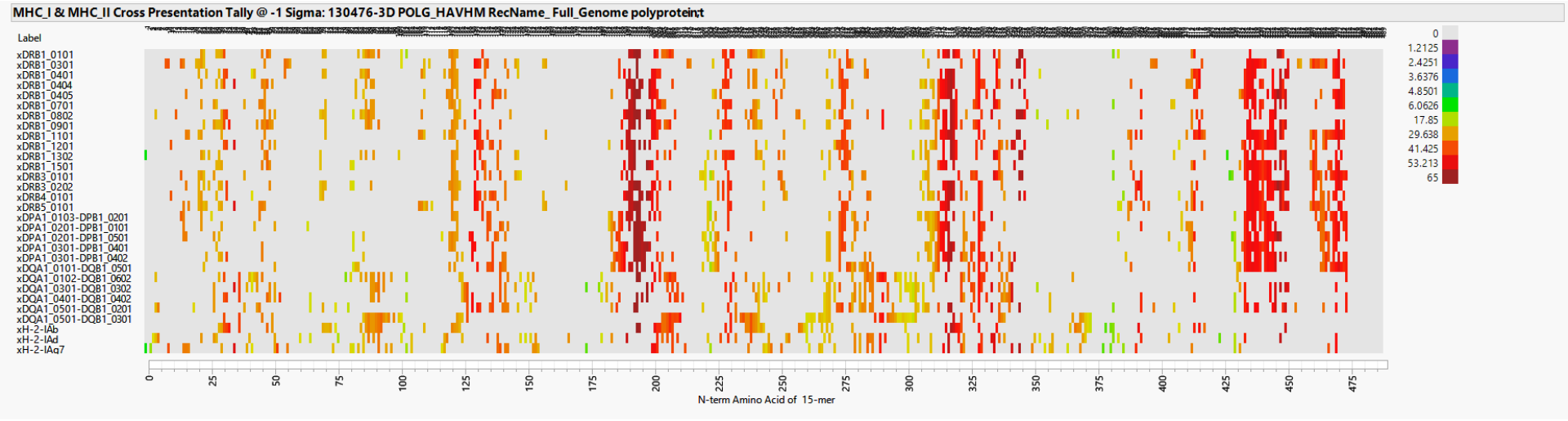
130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein;t

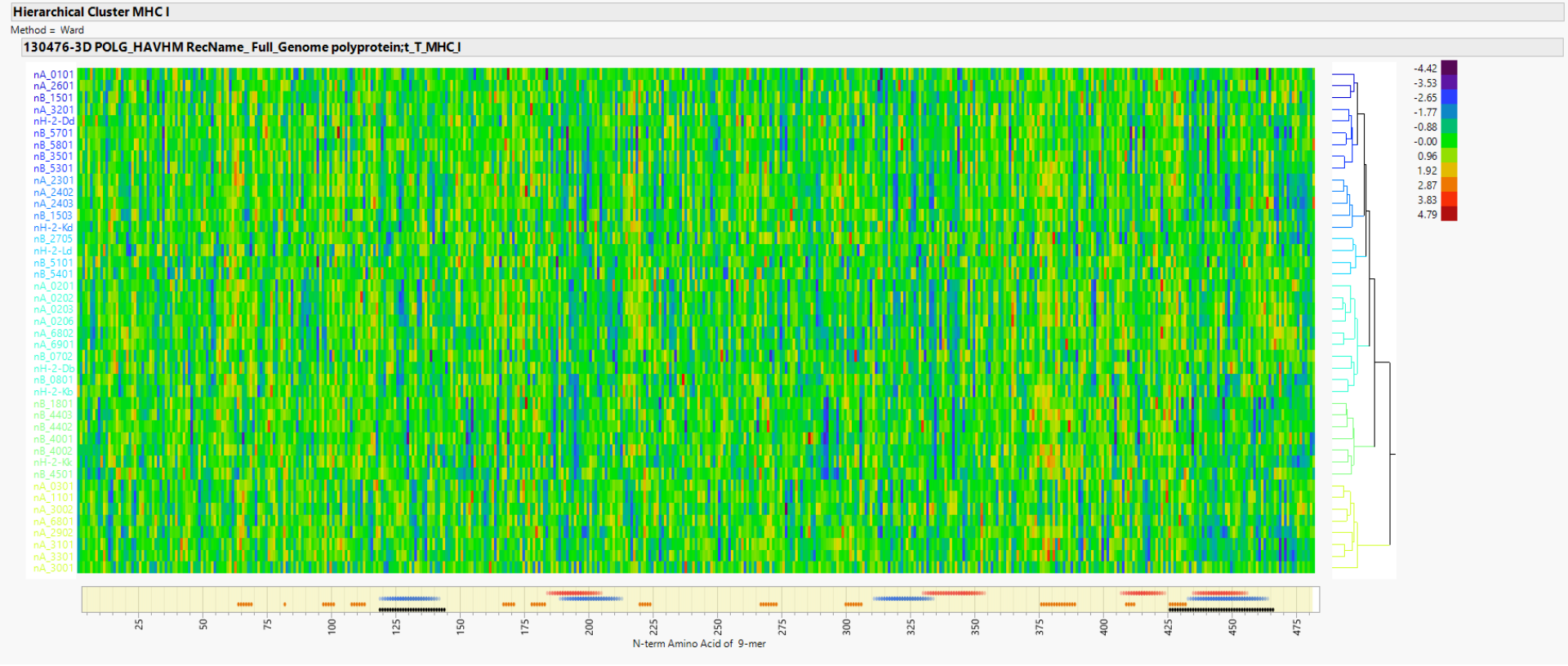


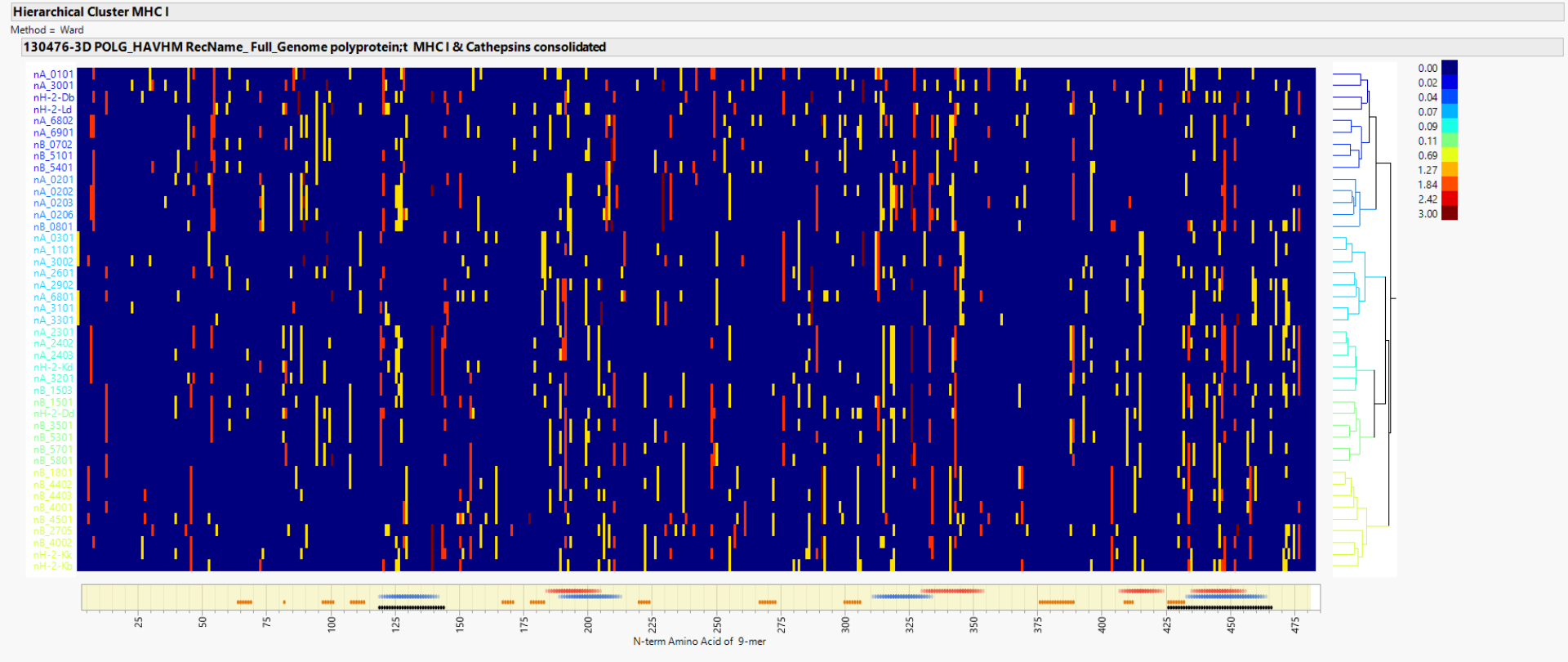

Hierarchical Cluster MHC II
Method = Ward
130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0802
nDRB1_0401
nDRB3_0202
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.62
-2.89
-2.17
-1.45
-0.72
0.00
0.69
1.38
2.06
2.75
3.44
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer

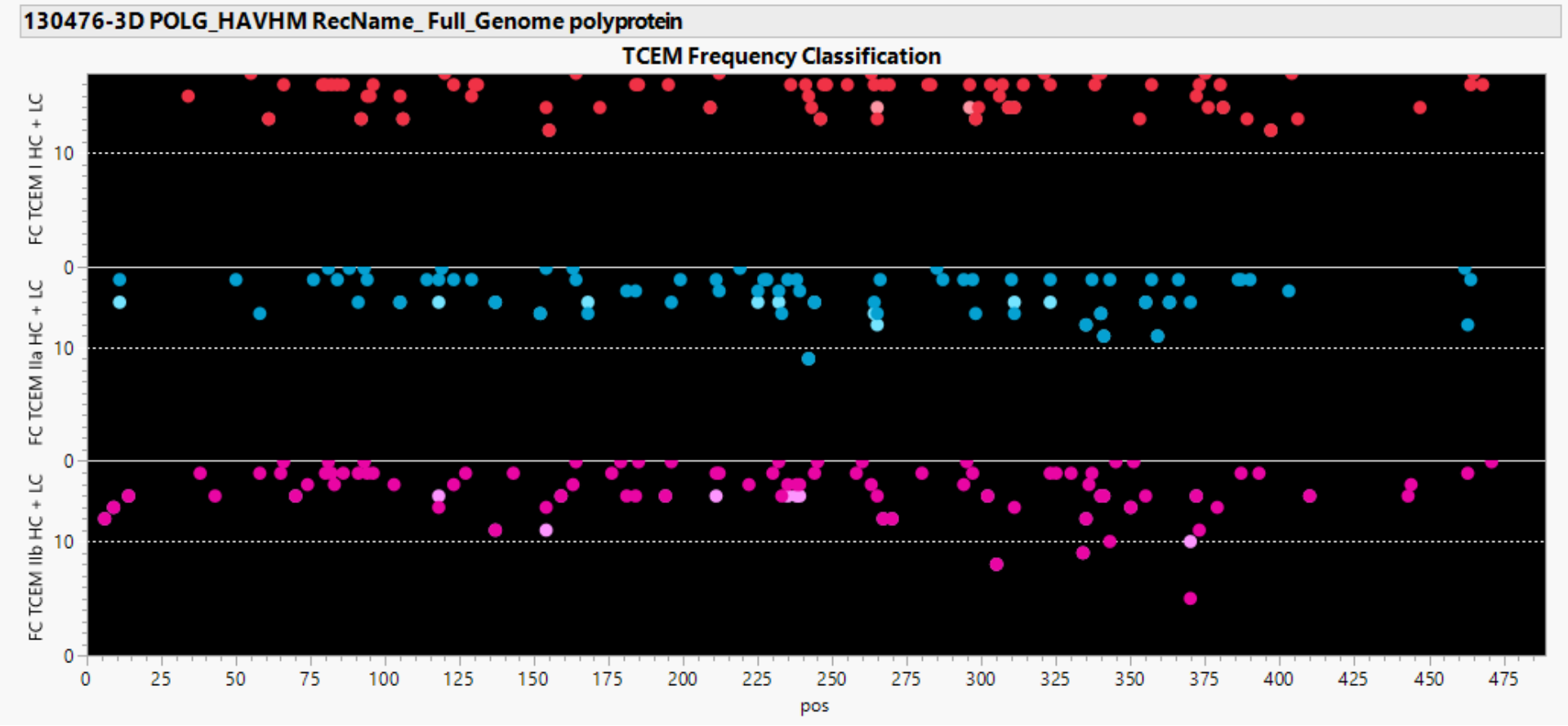

130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.06
0.12
0.17
0.23
0.29
1.03
1.77
2.52
3.26
4.00

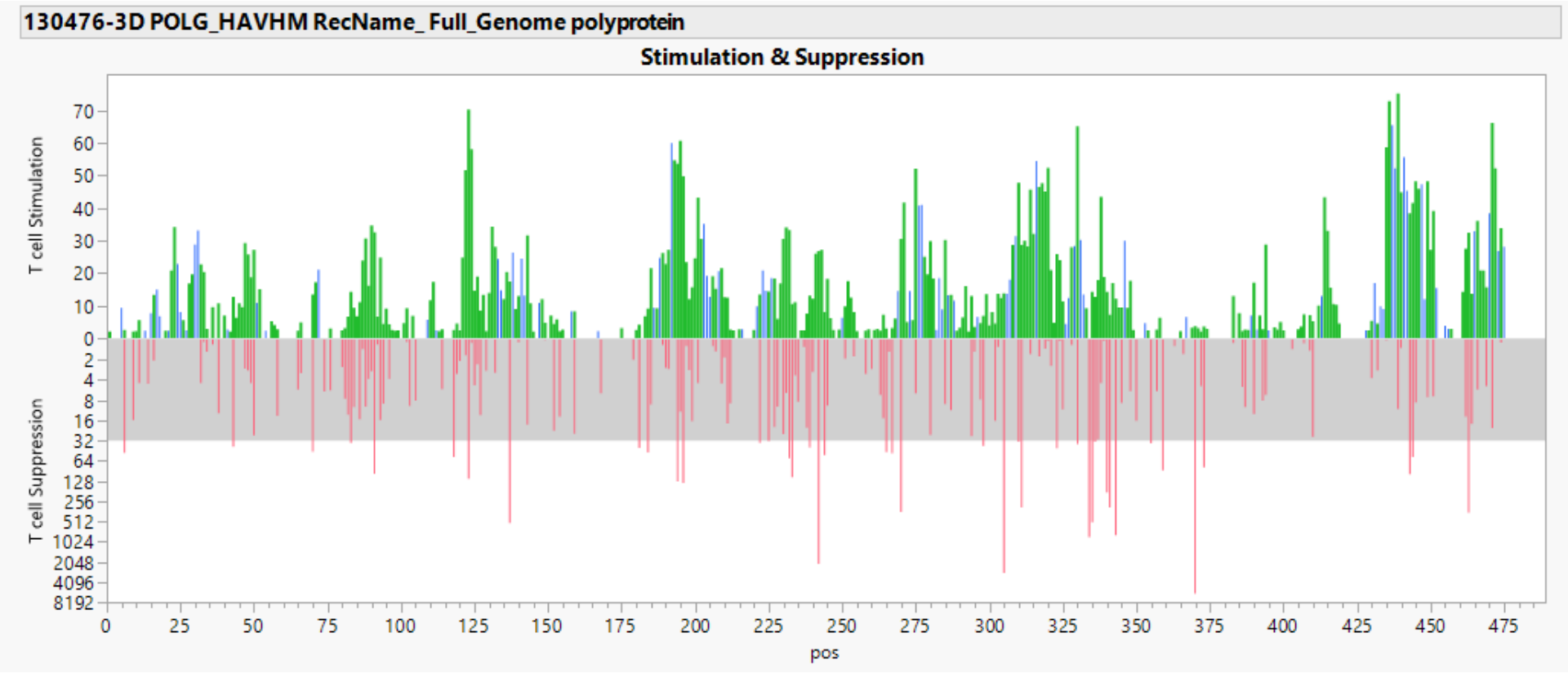

130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.04
0.08
0.12
0.16
0.20
0.96
1.72
2.48
3.24
4.00


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-3D POLG_HAVHM RecName_Full_Genome polyproteint
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 15-mer
0
1.2125
2.4251
3.6376
4.8501
6.0626
17.85
29.638
41.425
53.213
65

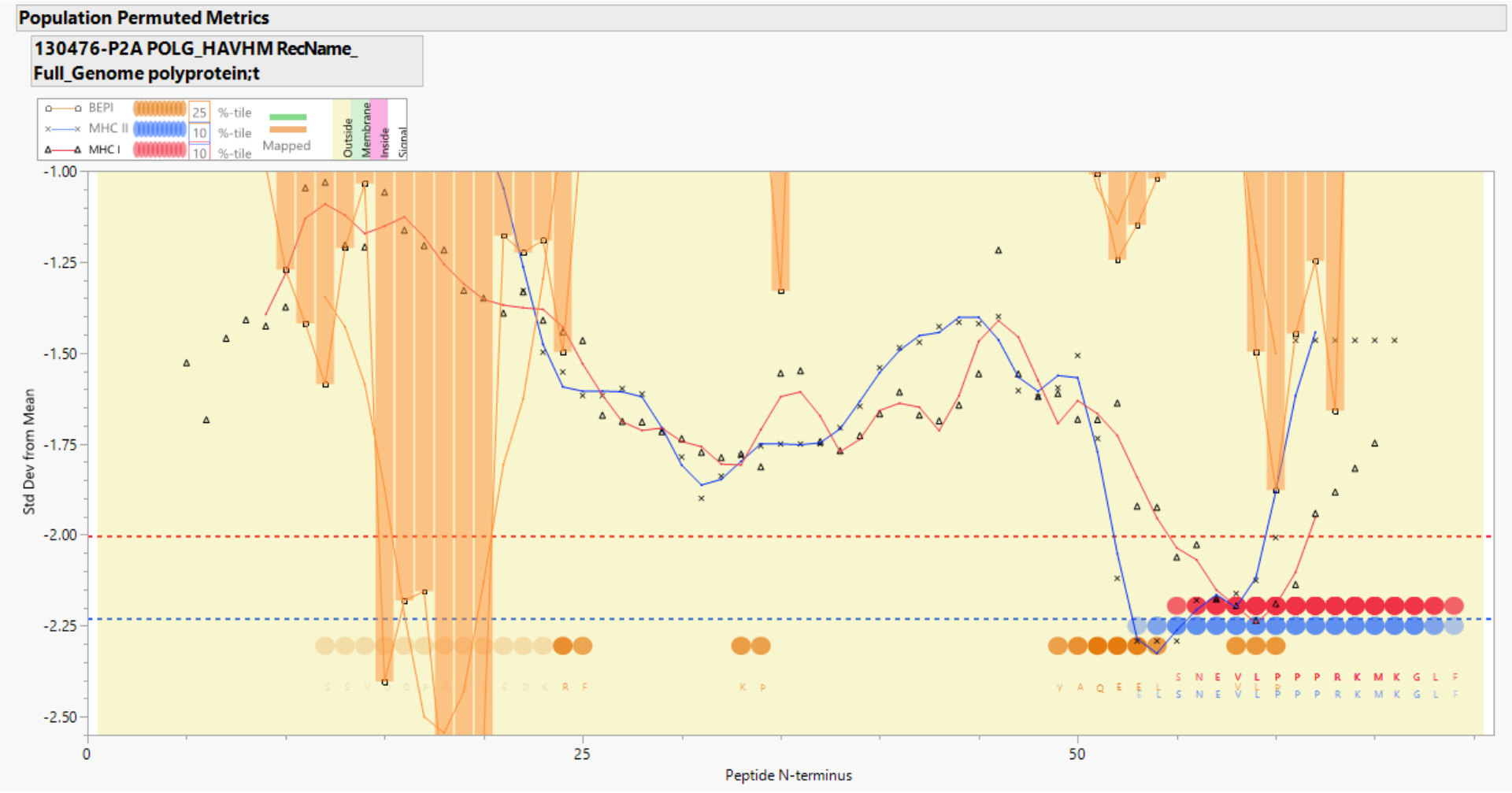

Hierarchical Cluster MHC I
Method = Ward
130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_2601
nB_1501
nA_3201
nH-2-Dd
nB_5701
nB_5801
nB_3501
nB_5301
nA_2301
nA_2402
nA_2403
nB_1503
nH-2-Kd
nB_2705
nH-2-Ld
nB_5101
nB_5401
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_0702
nH-2-Db
nB_0801
nH-2-Kb
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
nA_3001
-4.42
-3.53
-2.65
-1.77
-0.88
-0.00
0.96
1.92
2.87
3.83
4.79
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

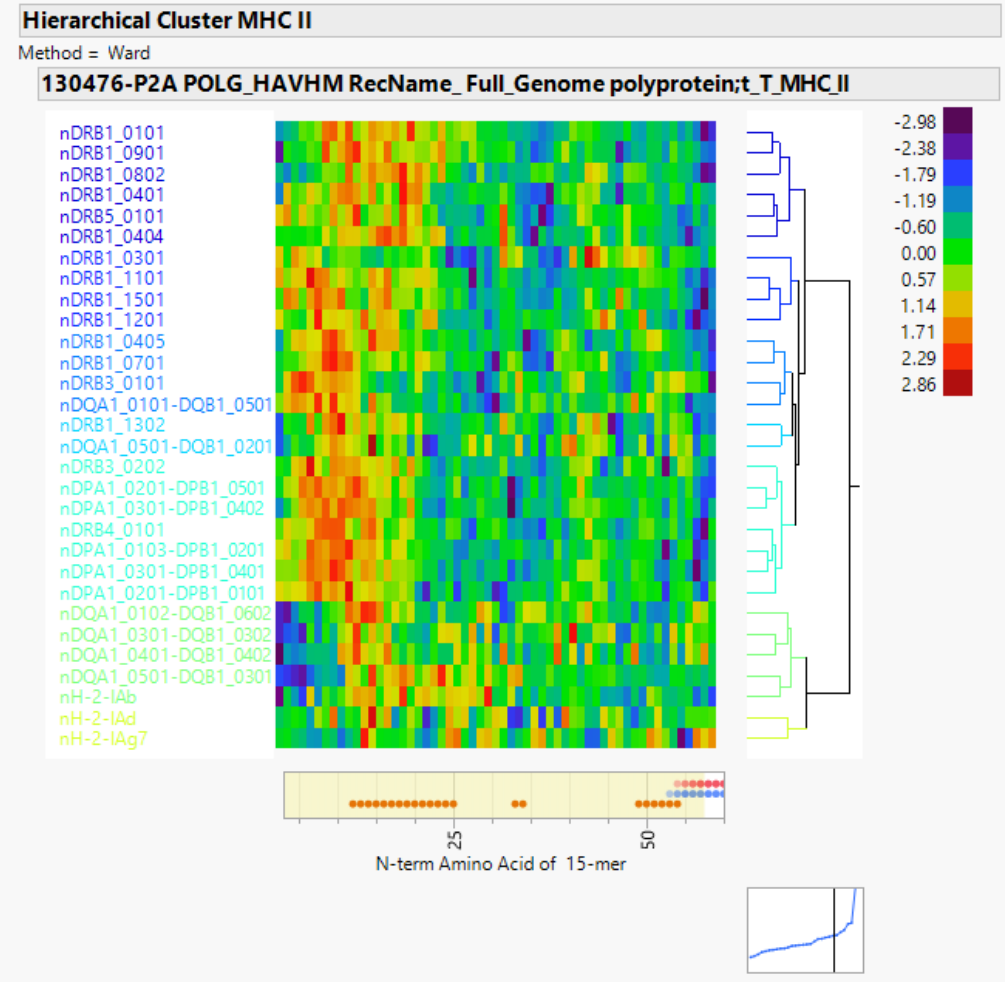

Hierarchical Cluster MHC I
Method = Ward
130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nA_3001
nH-2-Db
nH-2-Ld
nA_6802
nA_6901
nB_0702
nB_5101
nB_5401
nA_0201
nA_0202
nA_0203
nA_0206
nB_0801
nA_0301
nA_1101
nA_3002
nA_2601
nA_2902
nA_6801
nA_3101
nA_3301
nA_2301
nA_2402
nA_2403
nH-2-Kd
nA_3201
nB_1503
nB_1501
nH-2-Dd
nB_3501
nB_5301
nB_5701
nB_5801
nB_1801
nB_4402
nB_4403
nB_4001
nB_4501
nB_2705
nB_4002
nH-2-Kk
nH-2-Kb
0.00
0.02
0.04
0.07
0.09
0.11
0.69
1.27
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
N-term Amino Acid of 9-mer

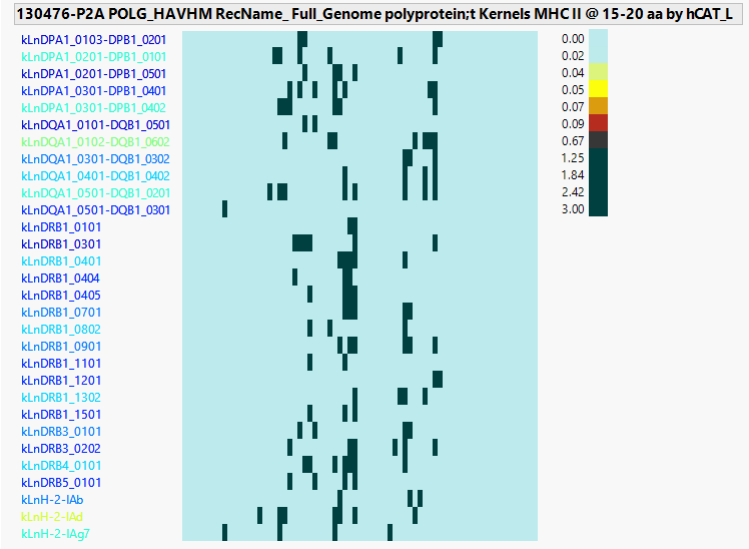

130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
pos

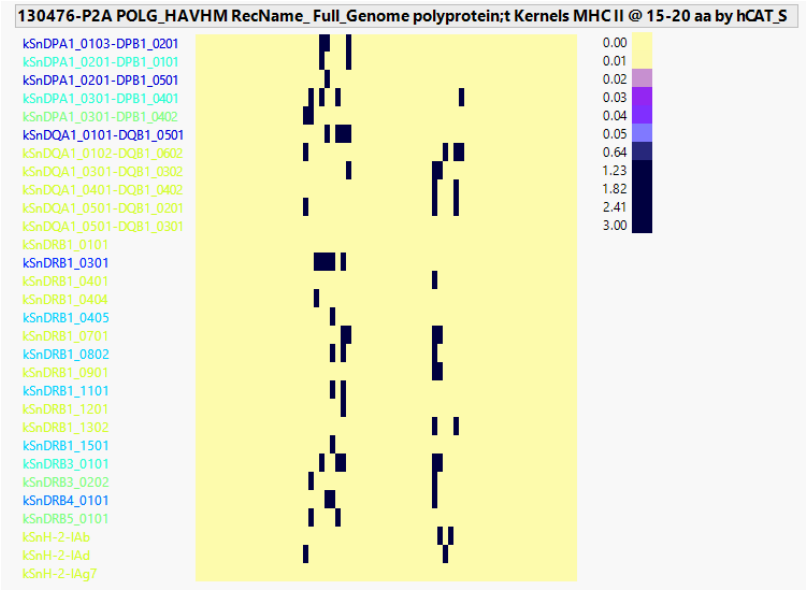

130476-3D POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475

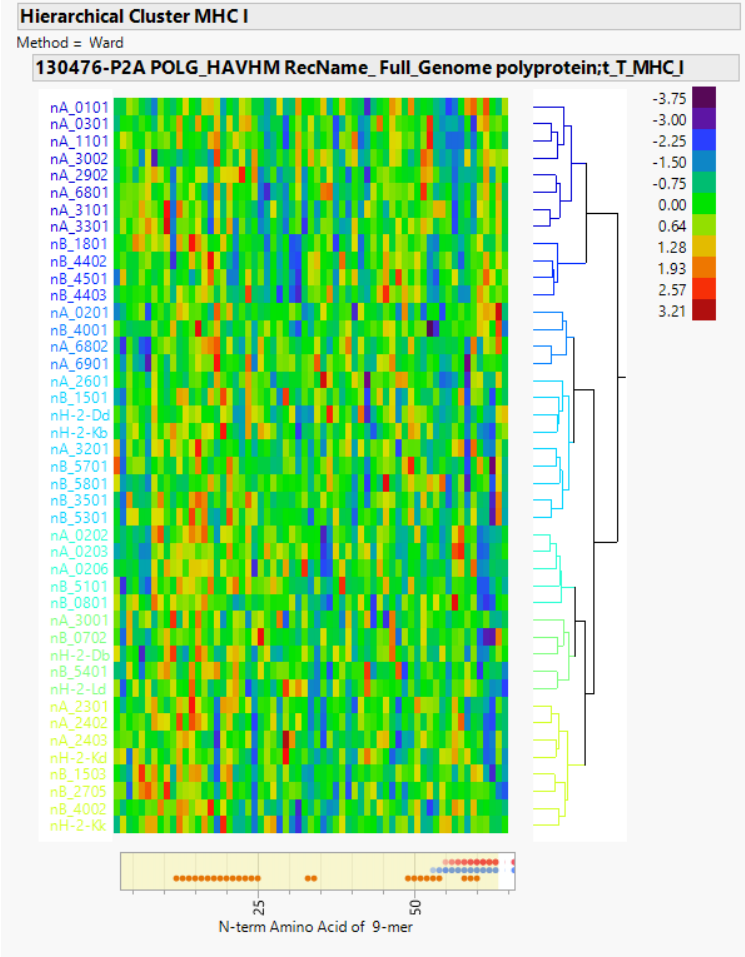
Protein 2A

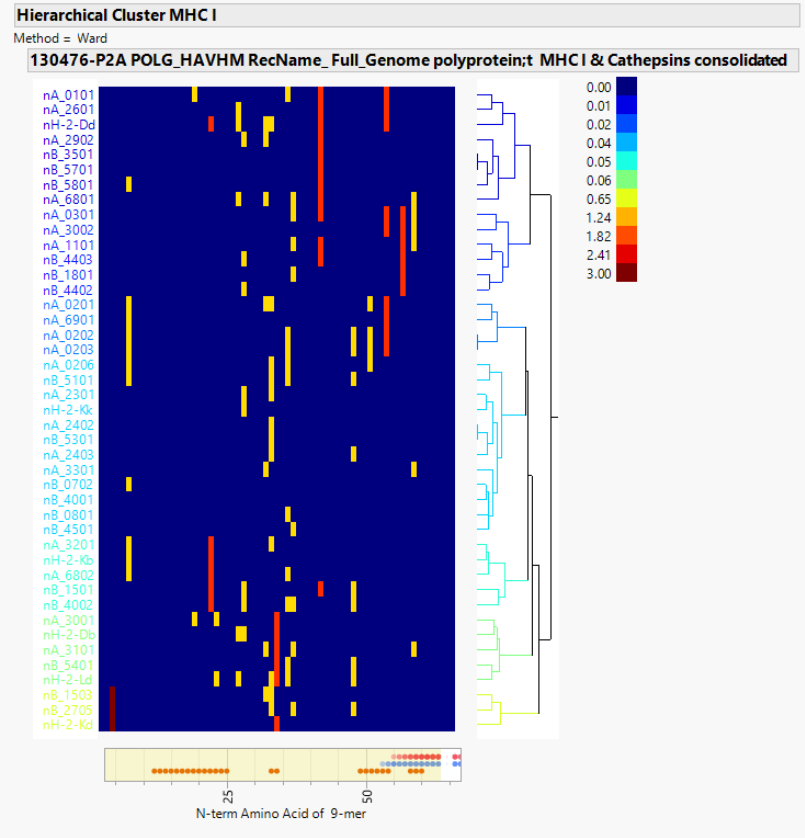

Population Permuted Metrics
130476-P2A POLG_HAVHM RecName_
Full_Genome polyprotein;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Std Dev from Mean
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
Peptide N-terminus

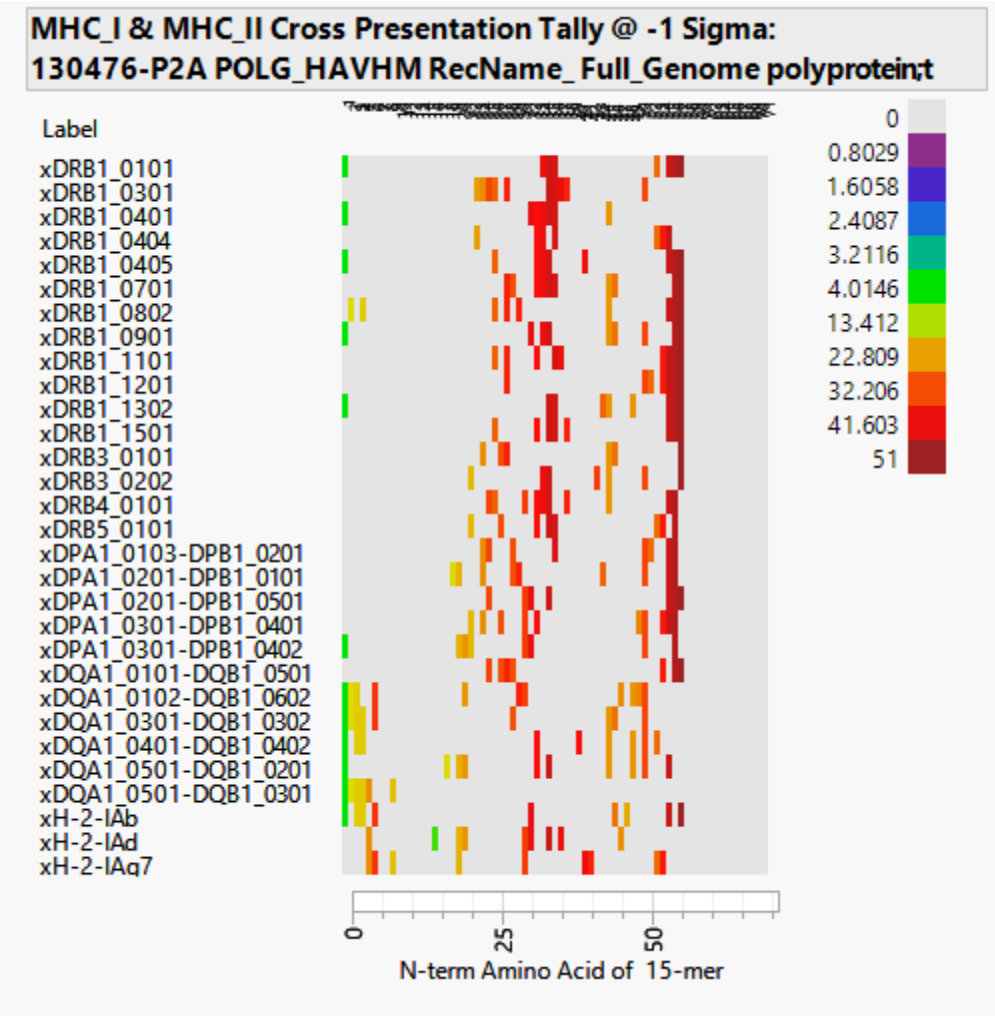

Hierarchical Cluster MHC II
Method = Ward
130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0802
nDRB1_0401
nDRB5_0101
nDRB1_0404
nDRB1_0301
nDRB1_1101
nDRB1_1501
nDRB1_1201
nDRB1_0405
nDRB1_0701
nDRB3_0101
nDQA1_0101-DQB1_0501
nDRB1_1302
nDQA1_0501-DQB1_0201
nDRB3_0202
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDRB4_0101
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-2.98
-2.38
-1.79
-1.19
-0.60
0.00
0.57
1.14
1.71
2.29
2.86
25
50
N-term Amino Acid of 15-mer

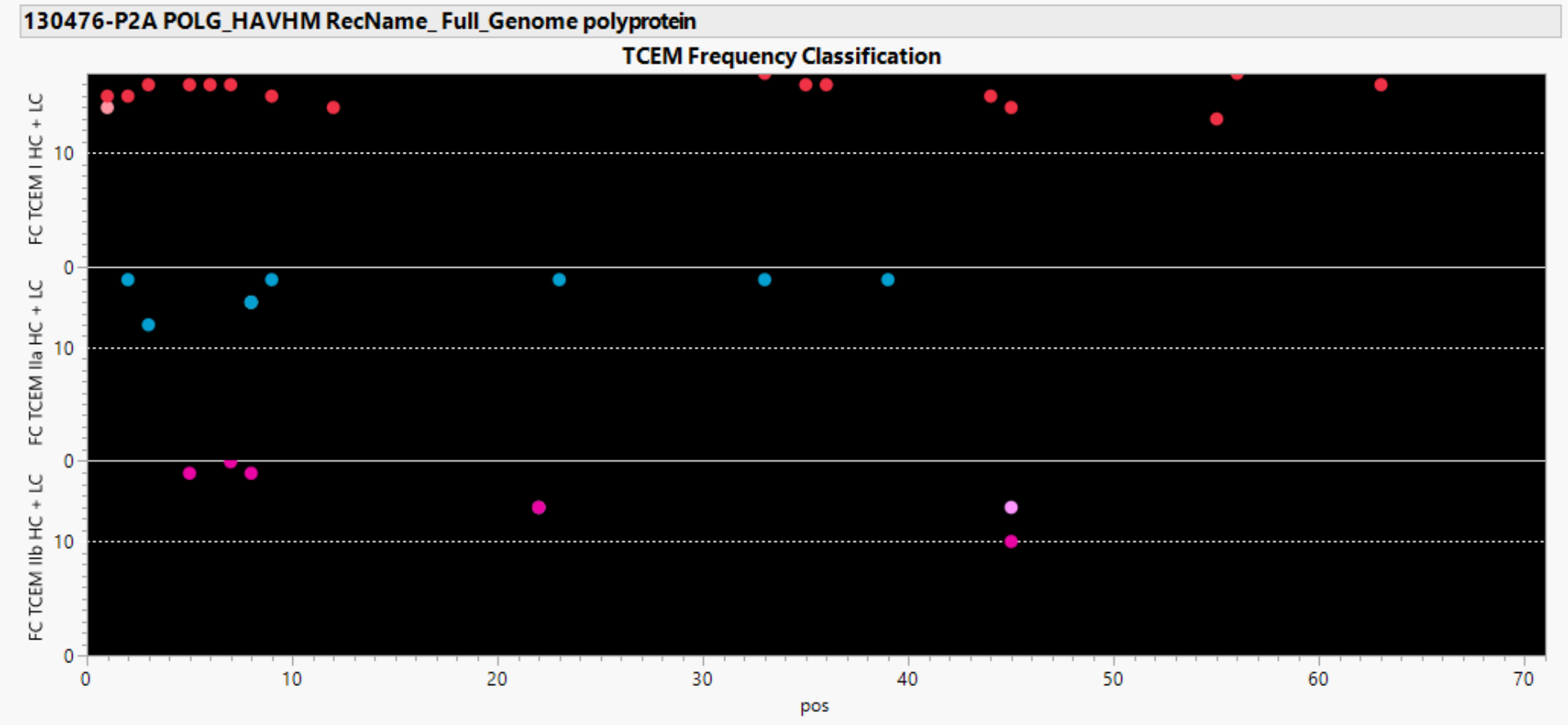

130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00

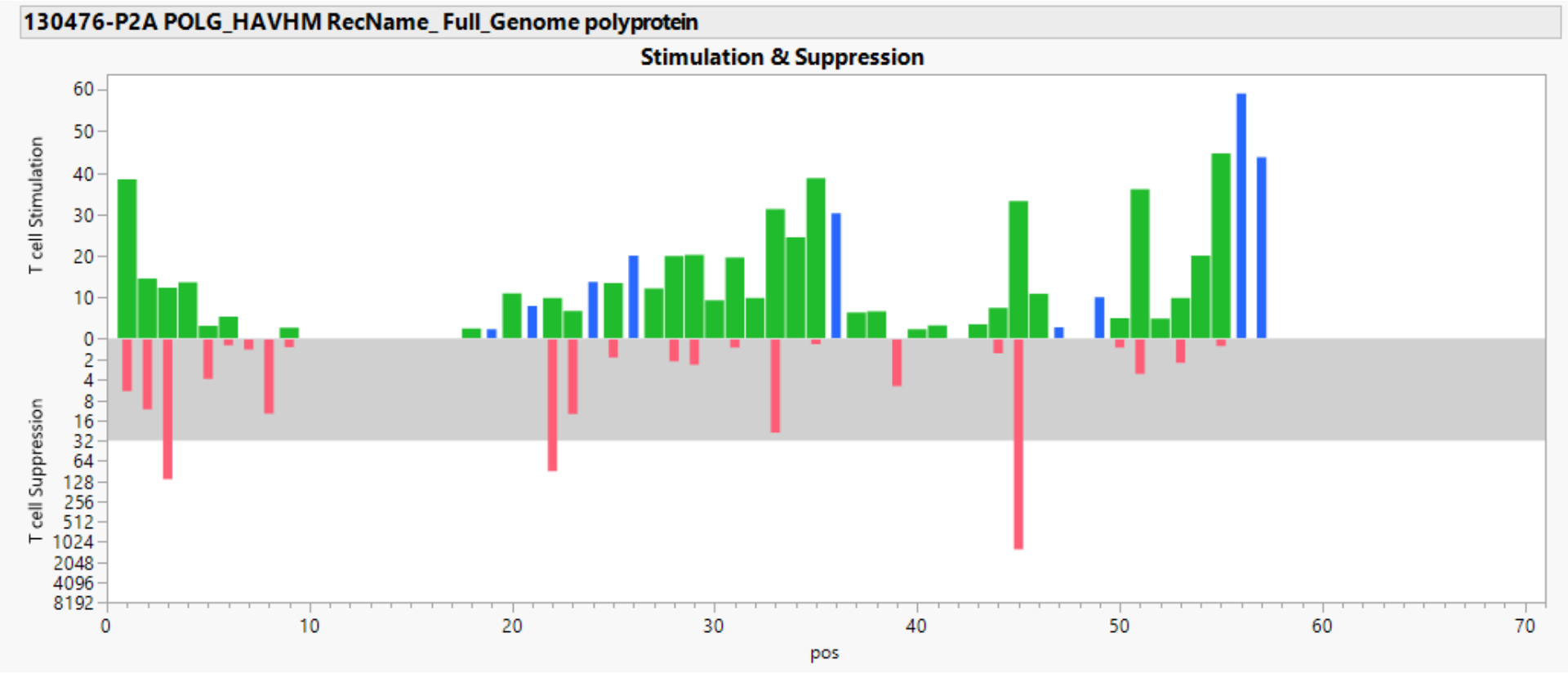

130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.01
0.02
0.03
0.04
0.05
0.64
1.23
1.82
2.41
3.00


Hierarchical Cluster MHC I
Method = Ward
130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_0301
nA_1101
nA_3002
nA_2902
nA_6801
nA_3101
nA_3301
nB_1801
nB_4402
nB_4501
nB_4403
nA_0201
nB_4001
nA_6802
nA_6901
nA_2601
nB_1501
nH-2-Dd
nH-2-Kb
nA_3201
nB_5701
nB_5801
nB_3501
nB_5301
nA_0202
nA_0203
nA_0206
nB_5101
nB_0801
nA_3001
nB_0702
nH-2-Db
nB_5401
nH-2-Ld
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_4002
nH-2-Kk
-3.75
-3.00
-2.25
-1.50
-0.75
0.00
0.64
1.28
1.93
2.57
3.21
25
50
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nH-2-Dd
nA_2902
nB_3501
nB_5701
nB_5801
nA_6801
nA_0301
nA_3002
nA_1101
nB_4403
nB_1801
nB_4402
nA_0201
nA_6901
nA_0202
nA_0203
nA_0206
nB_5101
nA_2301
nH-2-Kk
nA_2402
nB_5301
nA_2403
nA_3301
nB_0702
nB_4001
nB_0801
nB_4501
nA_3201
nH-2-Kb
nA_6802
nB_1501
nB_4002
nA_3001
nH-2-Db
nA_3101
nB_5401
nH-2-Ld
nB_1503
nB_2705
nH-2-Kd
0.00
0.01
0.02
0.04
0.05
0.06
0.65
1.24
1.82
2.41
3.00
25
50
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma:
130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
0.8029
1.6058
2.4087
3.2116
4.0146
13.412
22.809
32.206
41.603
51
0
25
50
N-term Amino Acid of 15-mer

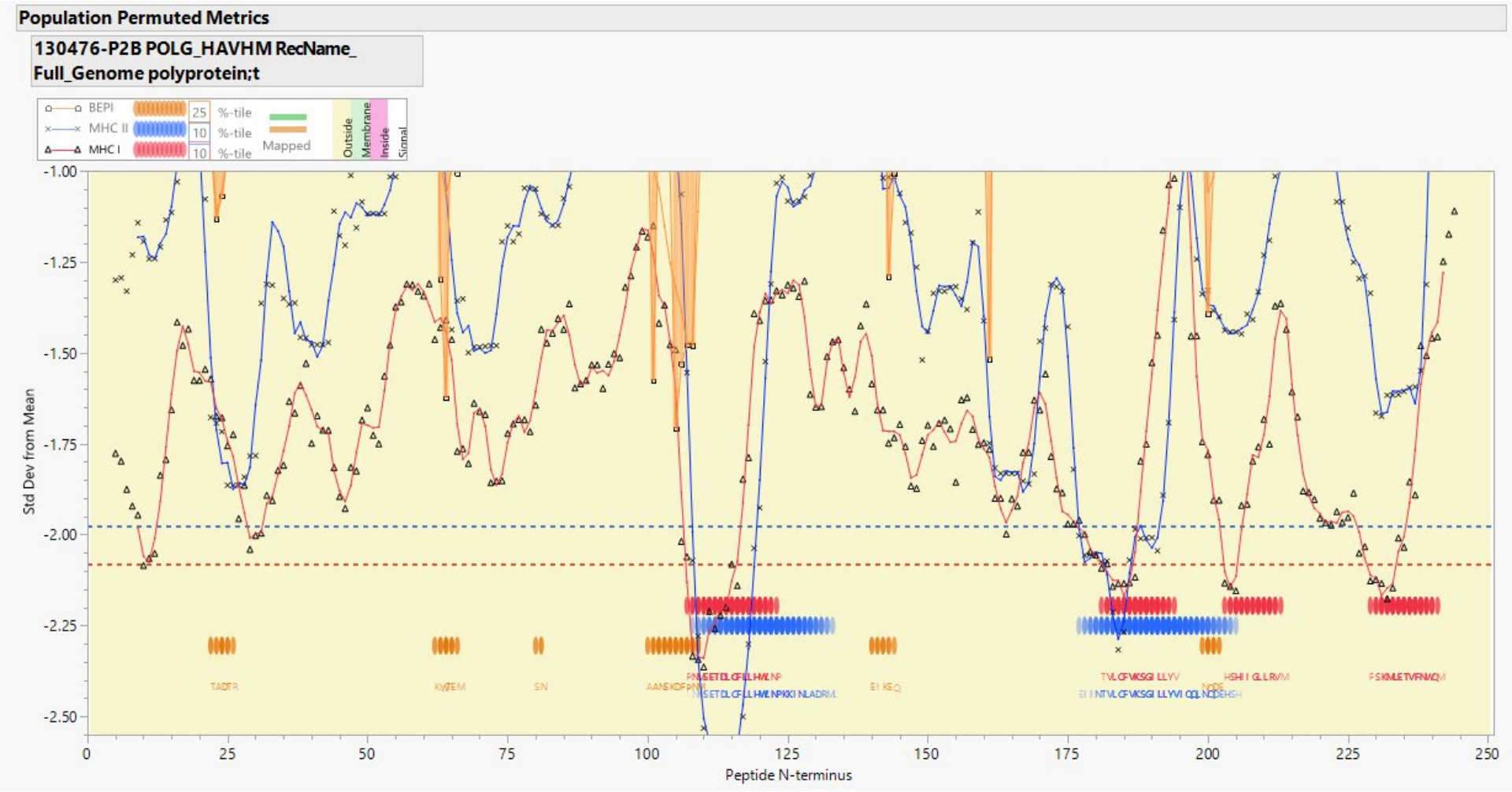

130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0
10
20
30
40
50
60
70
pos

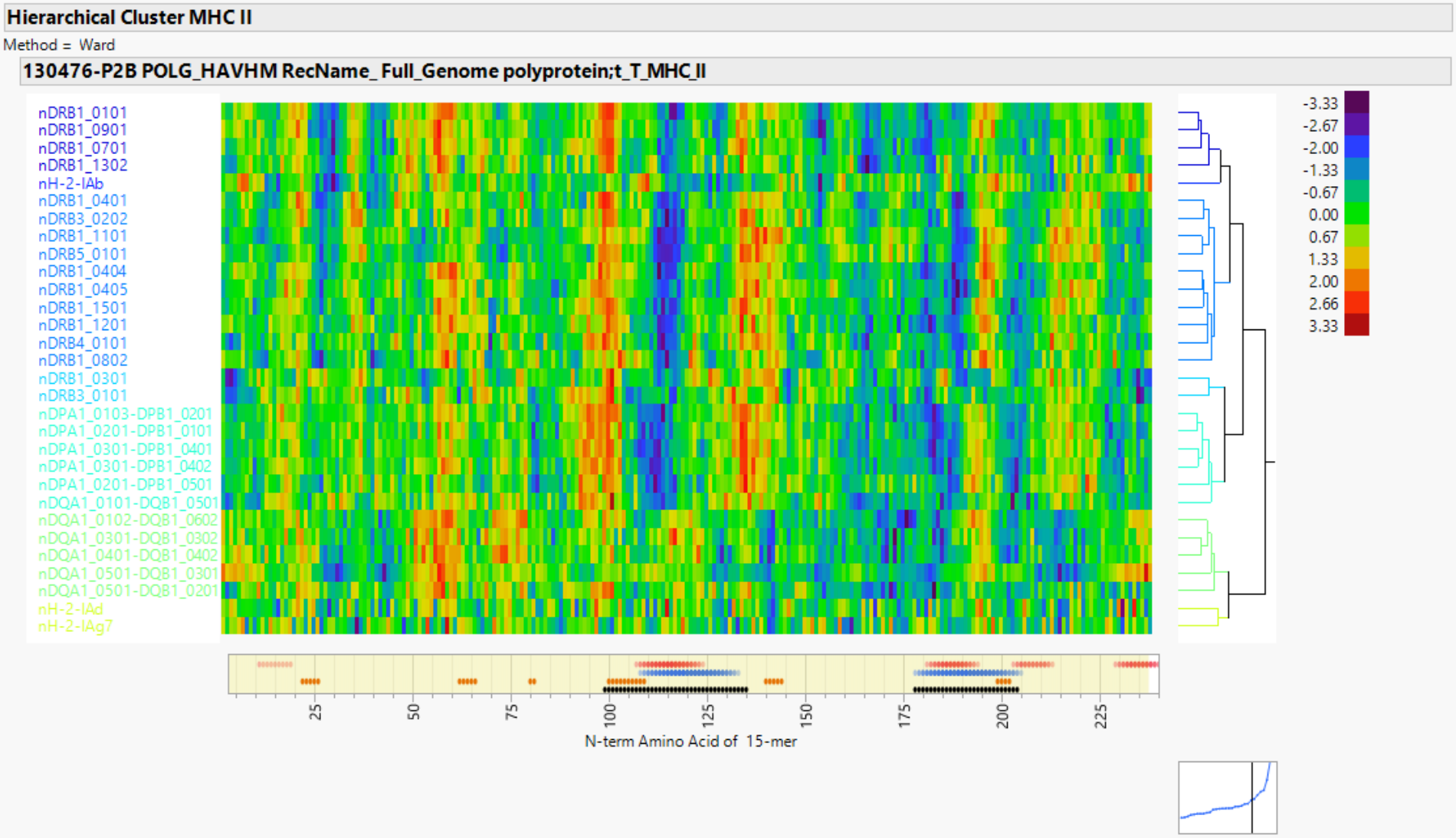

130476-P2A POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
60
50
40
30
20
10
0
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
10
20
30
40
50
60
70
pos

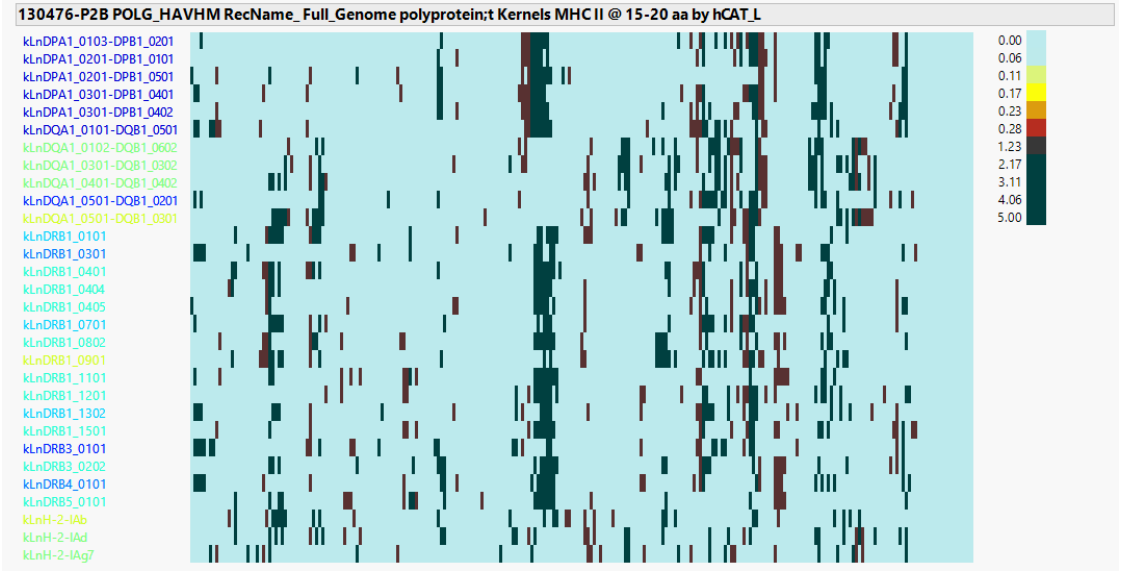
Protein 2B

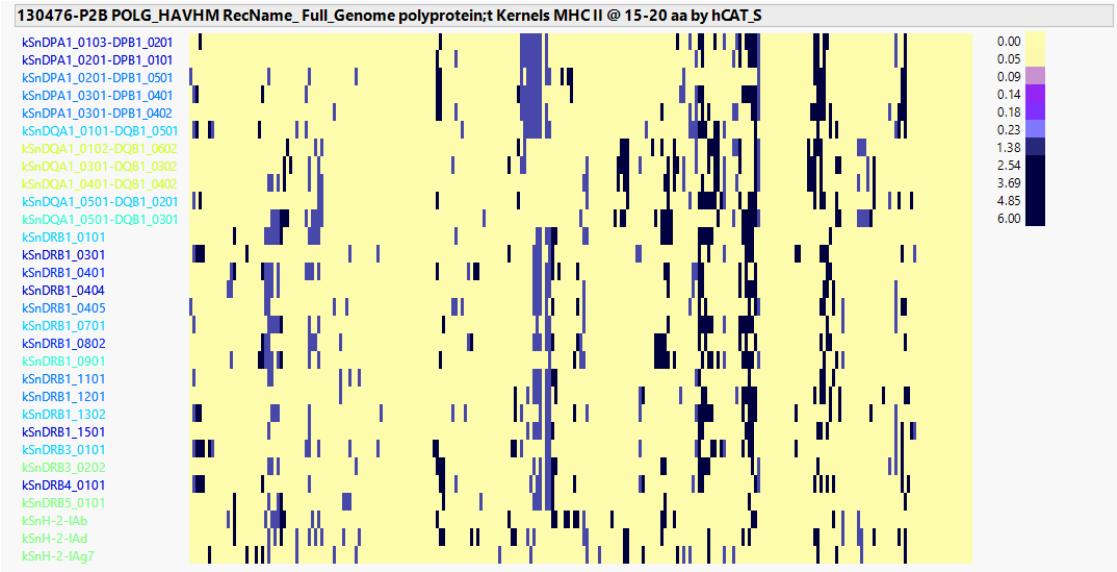
Population Permuted Metrics

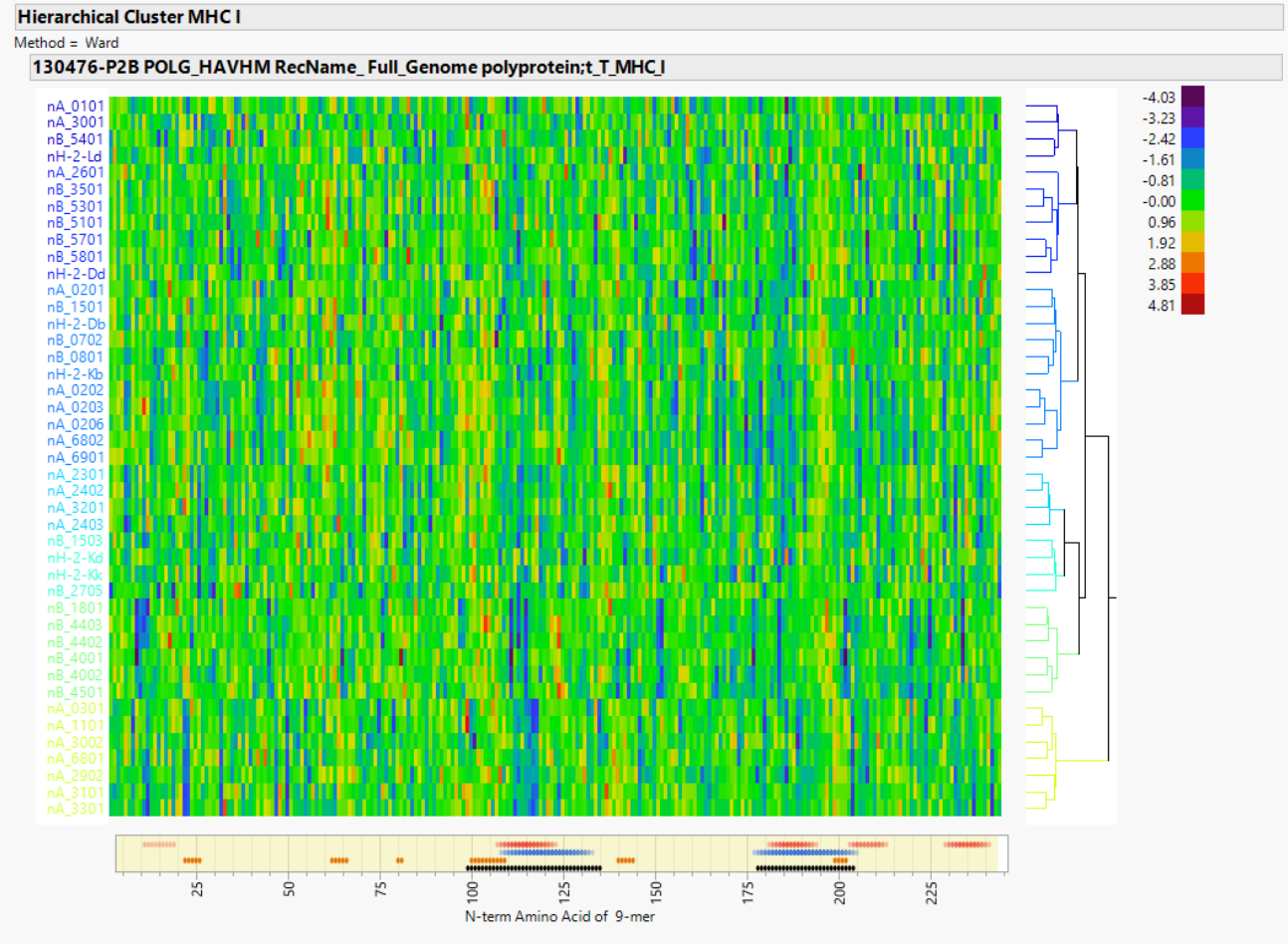
130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t

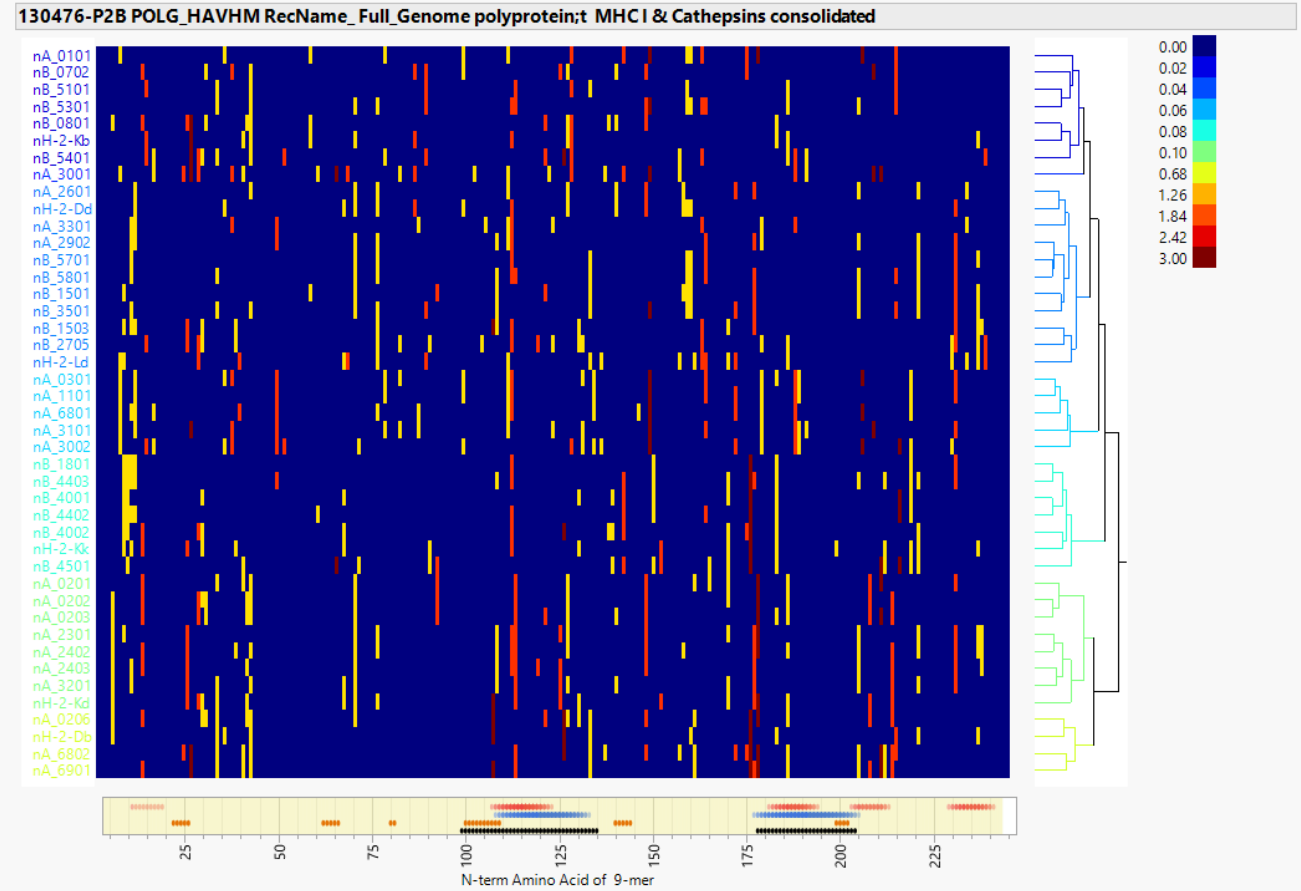


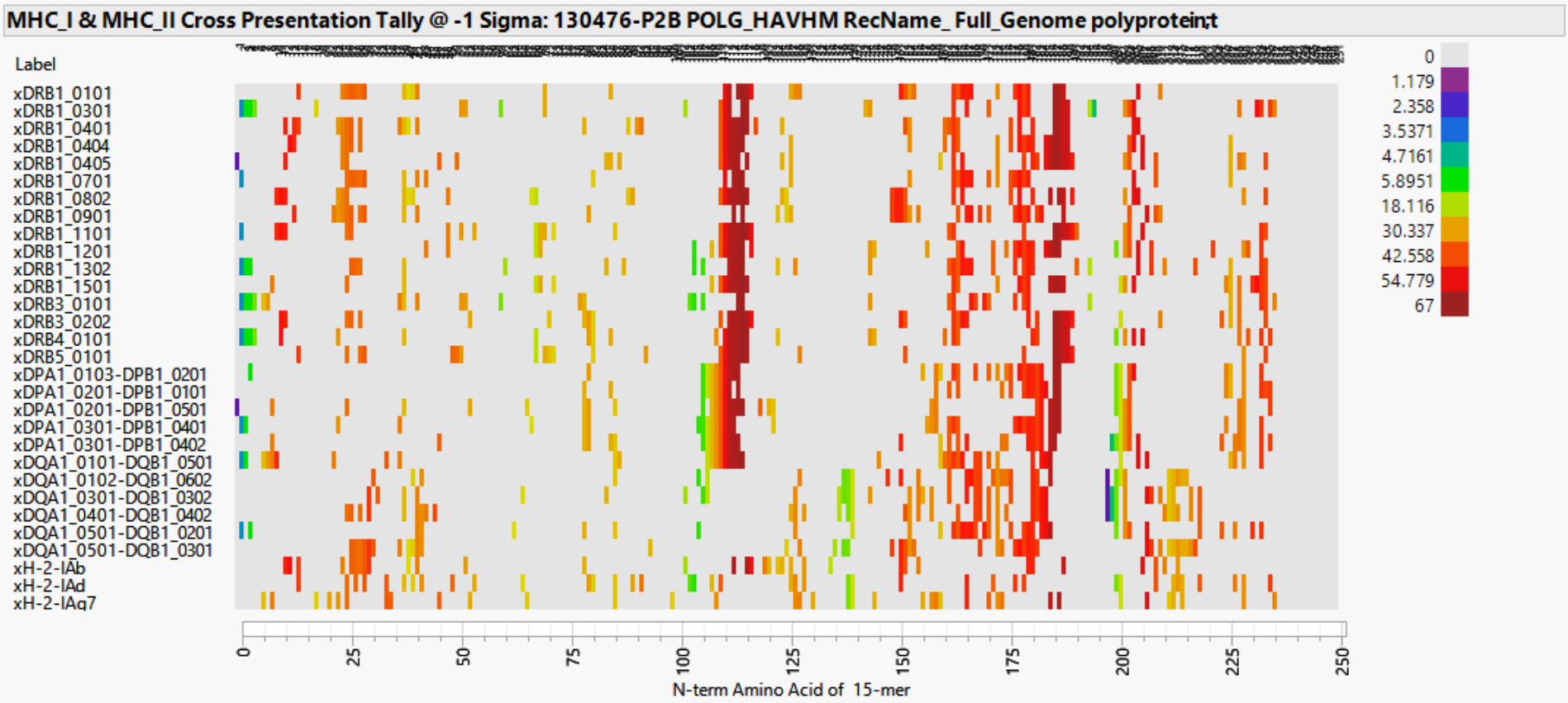

Hierarchical Cluster MHC II
Method = Ward
130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nH-2-IAb
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB1_1201
nDRB4_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0301
nDQA1_0501-DQB1_0201
nH-2-IAd
nH-2-IAg7
-3.33
-2.67
-2.00
-1.33
-0.67
0.00
0.67
1.33
2.00
2.66
3.33
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 15-mer

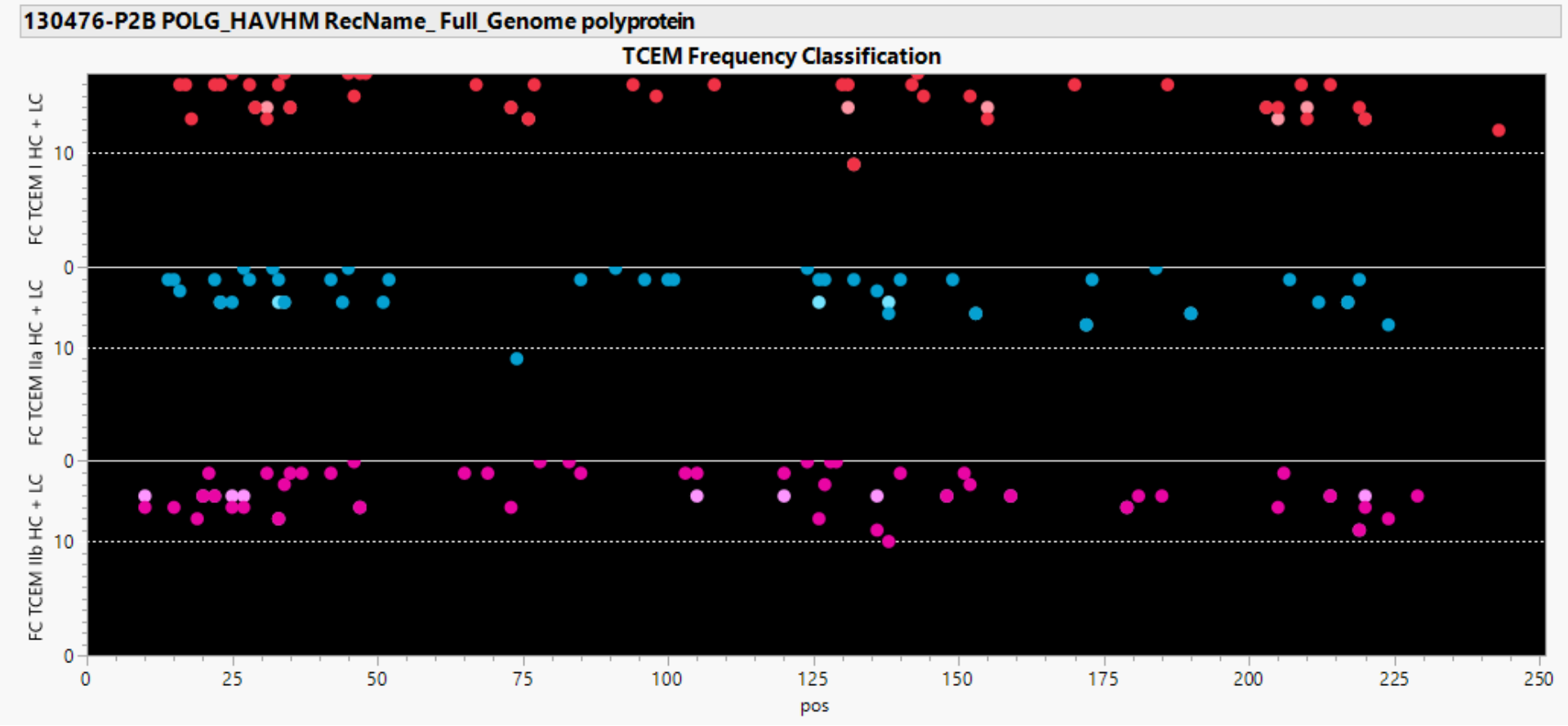

130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.06
0.11
0.17
0.23
0.28
1.23
2.17
3.11
4.06
5.00

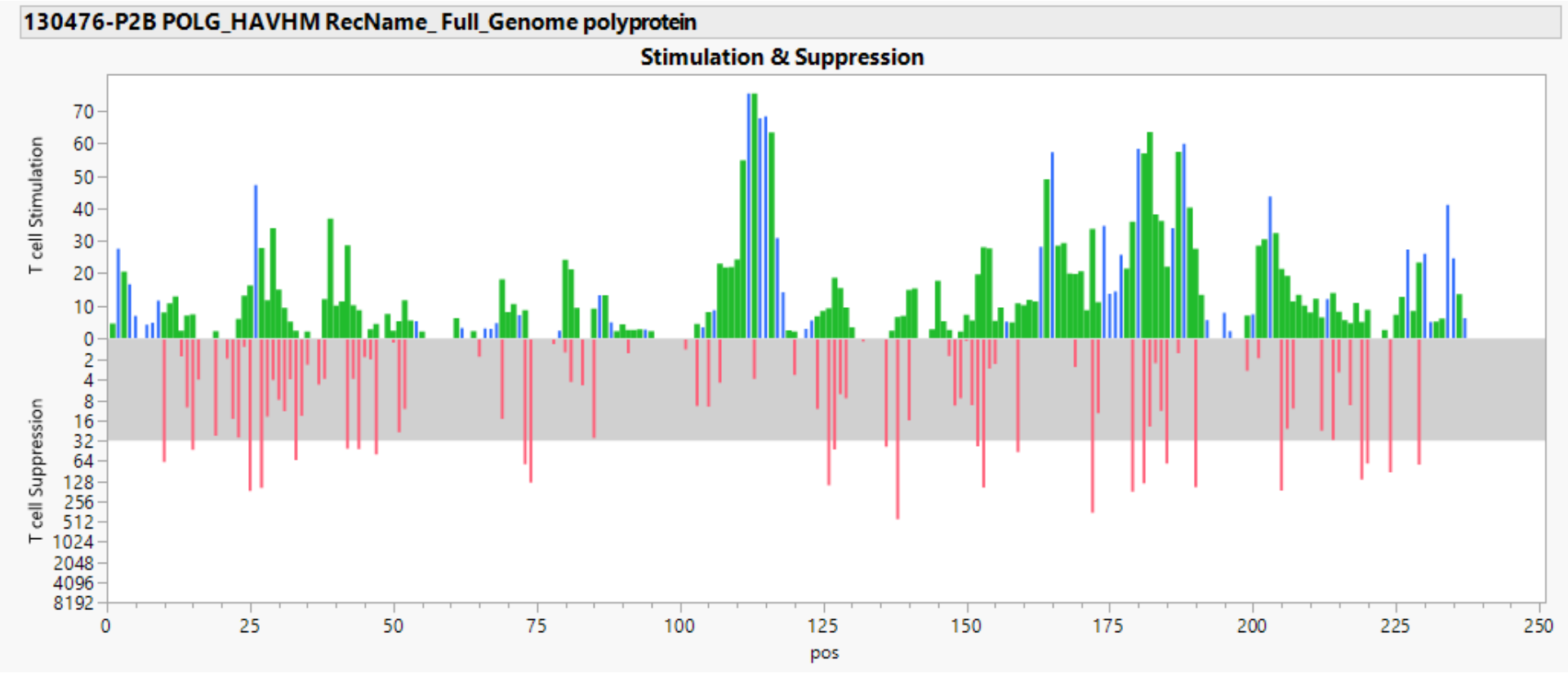

130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.05
0.09
0.14
0.18
0.23
1.38
2.54
3.69
4.85
6.00


Hierarchical Cluster MHC I
Method = Ward
130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_2601
nB_3501
nB_5301
nB_5101
nB_5701
nB_5801
nH-2-Dd
nA_0201
nB_1501
nH-2-Db
nB_0702
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_2301
nA_2402
nA_3201
nA_2403
nB_1503
nH-2-Kd
nH-2-Kk
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
-4.03
-3.23
-2.42
-1.61
-0.81
-0.00
0.96
1.92
2.88
3.85
4.81
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 9-mer


130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nB_0702
nB_5101
nB_5301
nB_0801
nH-2-Kb
nB_5401
nA_3001
nA_2601
nH-2-Dd
nA_3301
nA_2902
nB_5701
nB_5801
nB_1501
nB_3501
nB_1503
nB_2705
nH-2-Ld
nA_0301
nA_1101
nA_6801
nA_3101
nA_3002
nB_1801
nB_4403
nB_4001
nB_4402
nB_4002
nH-2-Kk
nB_4501
nA_0201
nA_0202
nA_0203
nA_2301
nA_2402
nA_2403
nA_3201
nH-2-Kd
nA_0206
nH-2-Db
nA_6802
nA_6901
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.179
2.358
3.5371
4.7161
5.8951
18.116
30.337
42.558
54.779
67
0
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 15-mer

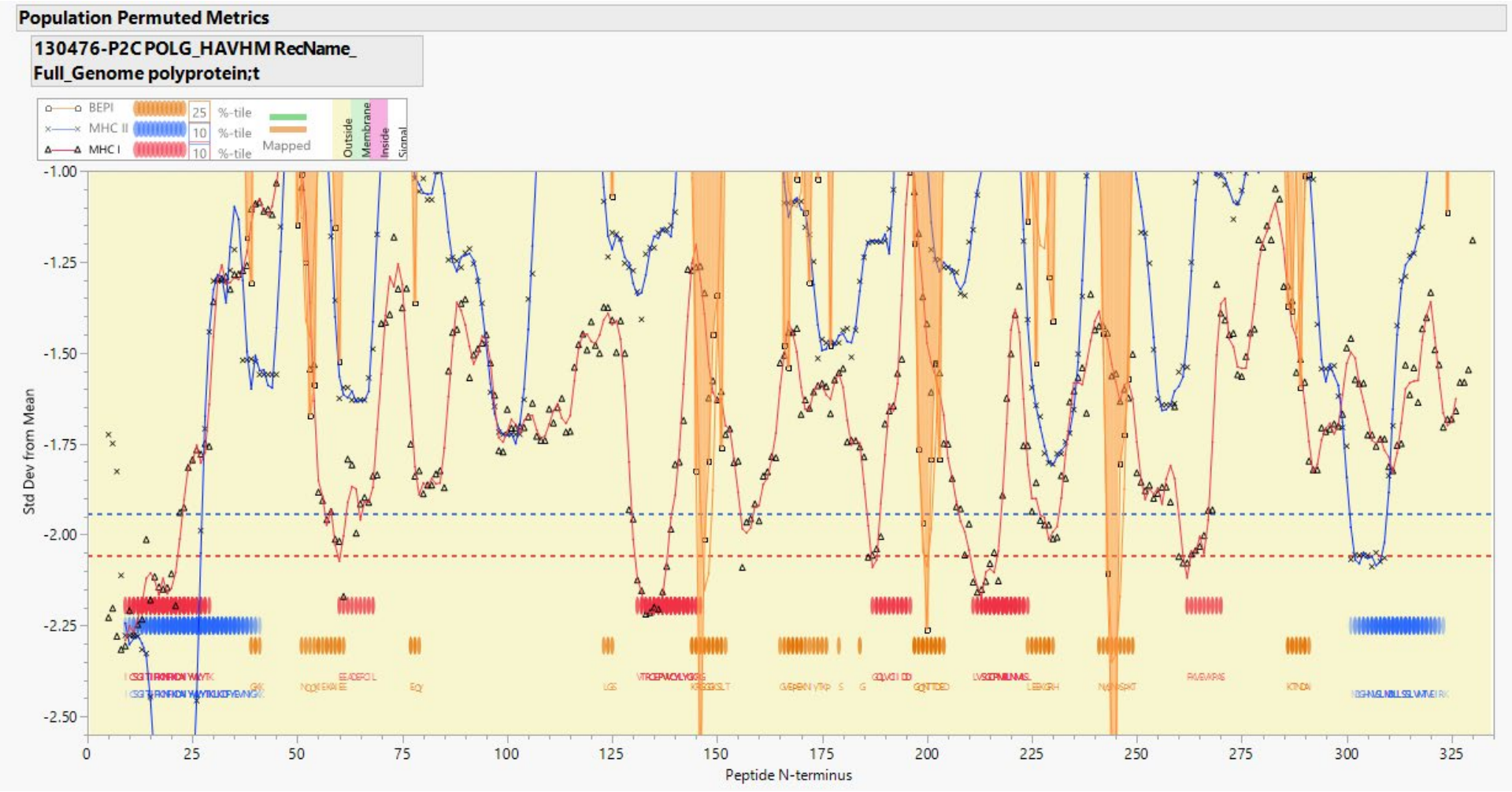

130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0
25
50
75
100
125
150
175
200
225
250
pos

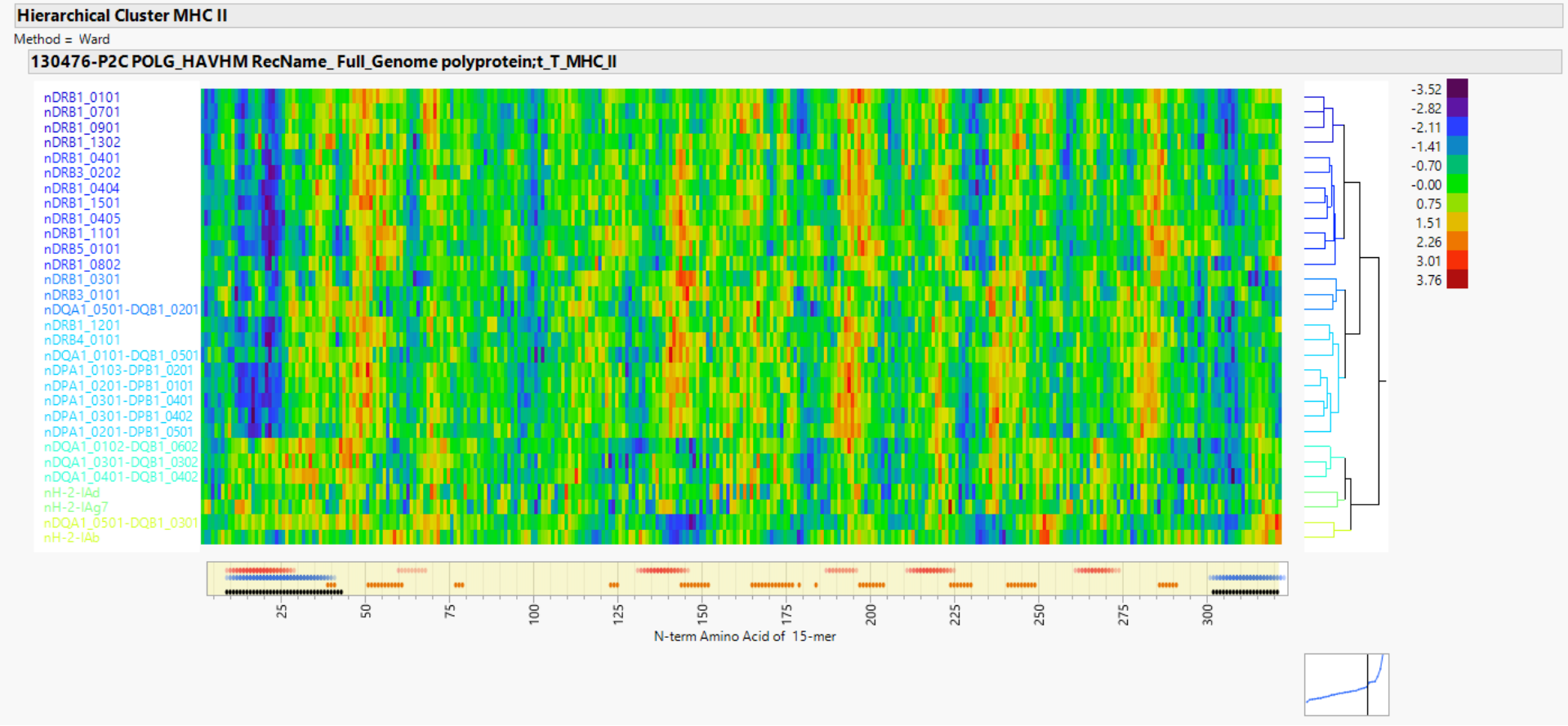

130476-P2B POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
70
60
50
40
30
20
10
0
T cell Suppression
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
25
50
75
100
125
150
175
200
225
250
pos

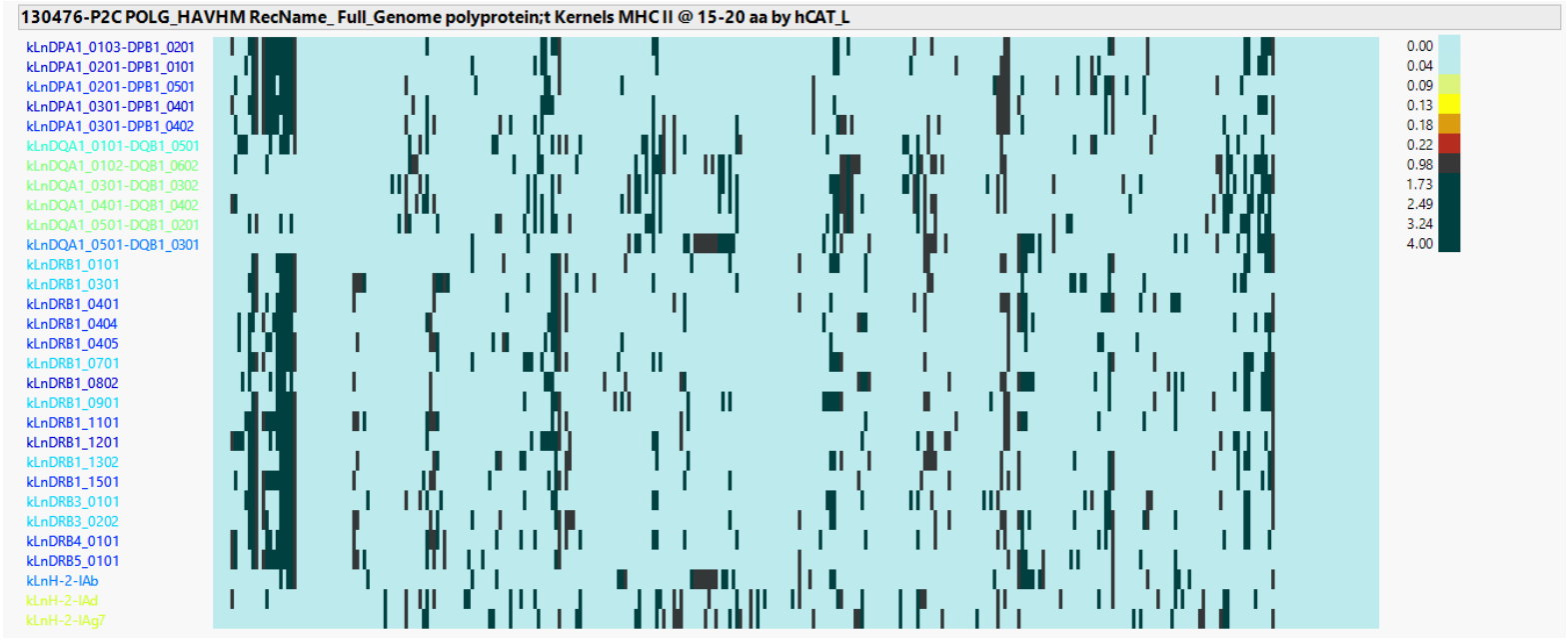
Protein 2C

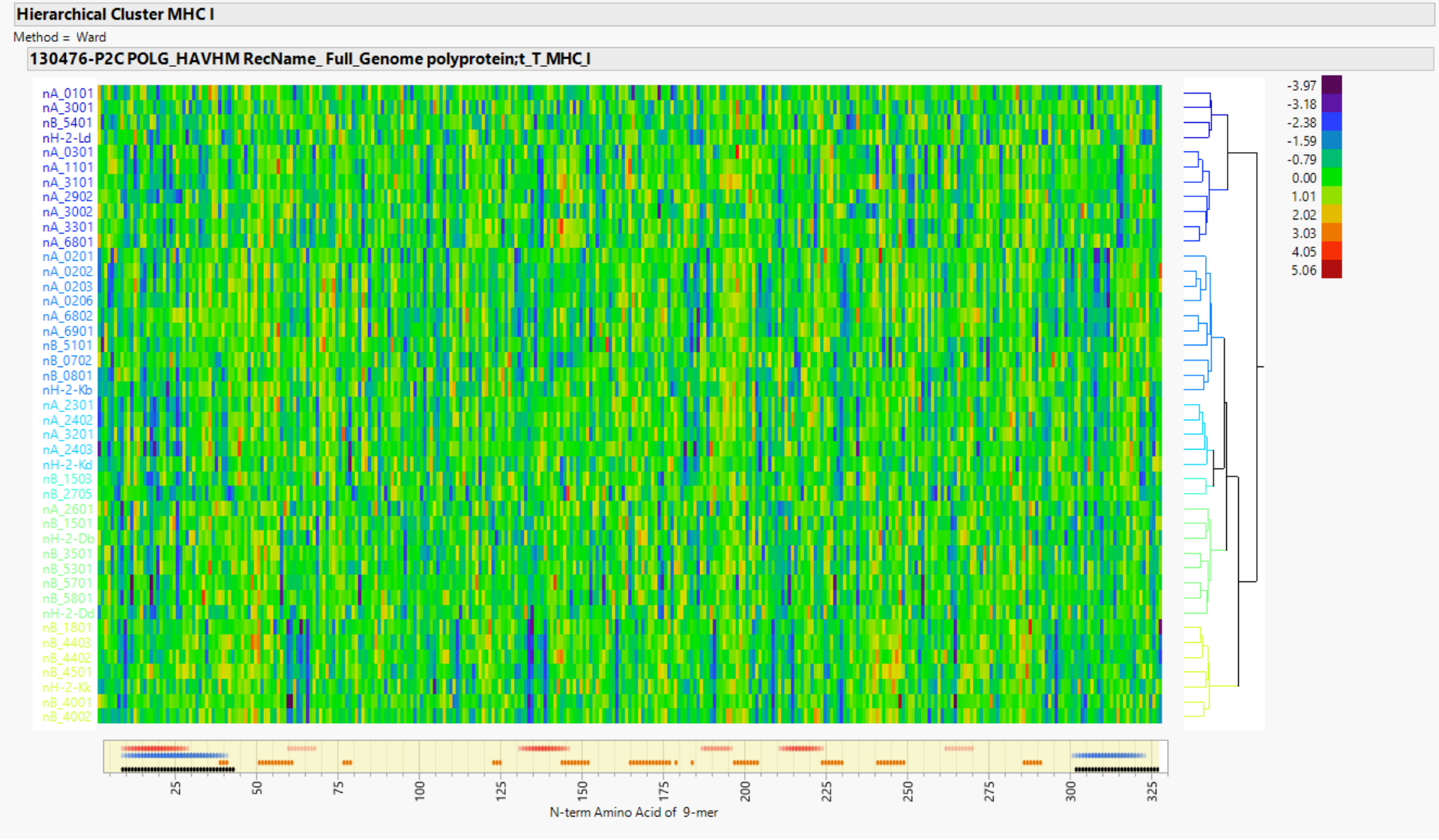
Population Permuted Metrics

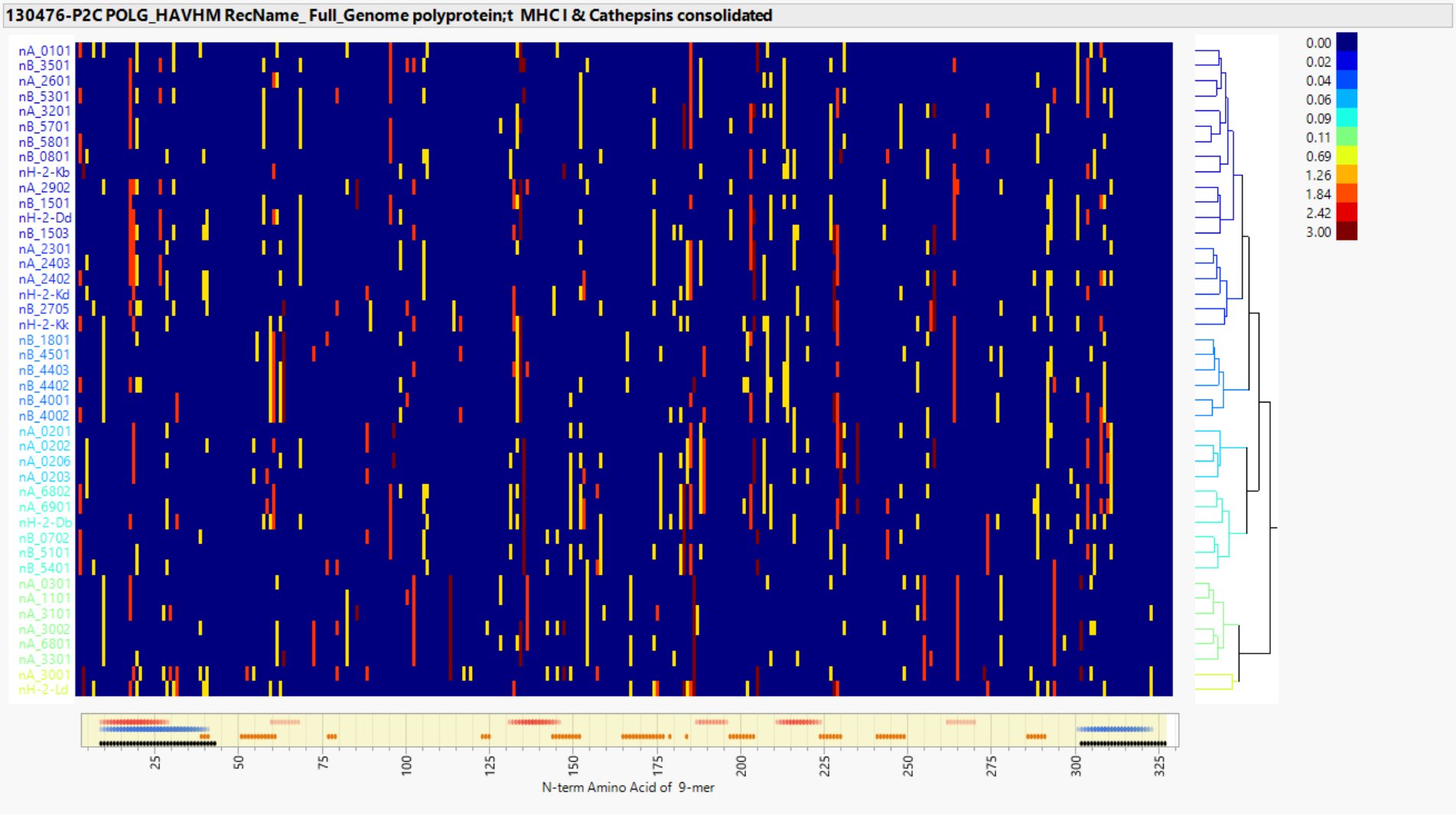
130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t

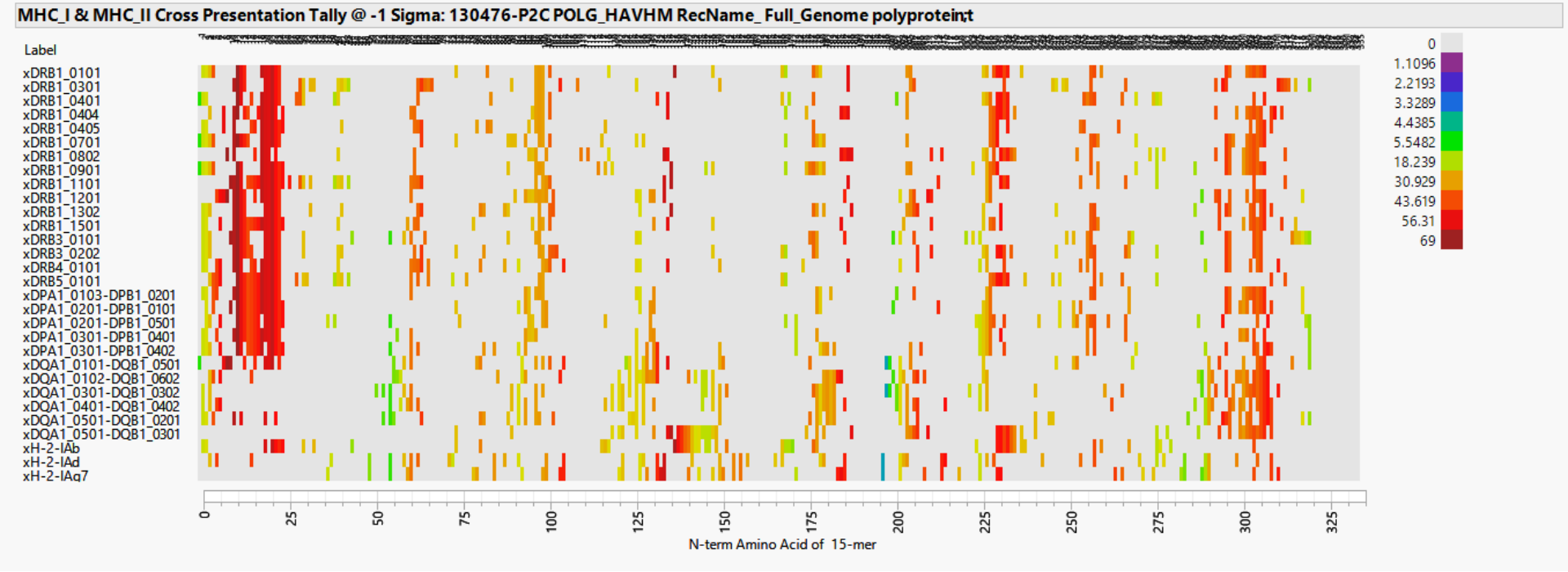


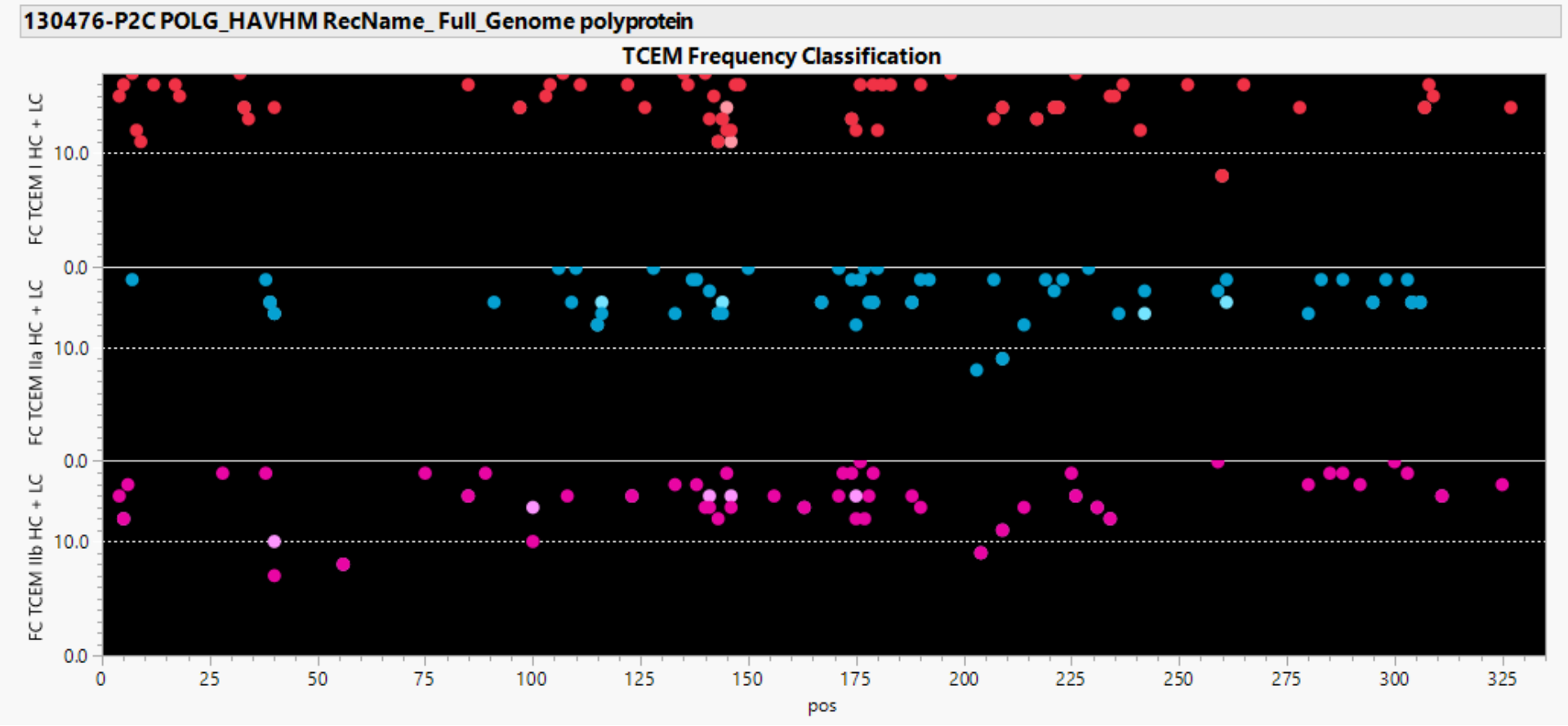

Hierarchical Cluster MHC II
Method = Ward
130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0701
nDRB1_0901
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDRB1_1201
nDRB4_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.52
-2.82
-2.11
-1.41
-0.70
-0.00
0.75
1.51
2.26
3.01
3.76
25
50
75
100
125
150
175
200
225
250
275
300
N-term Amino Acid of 15-mer

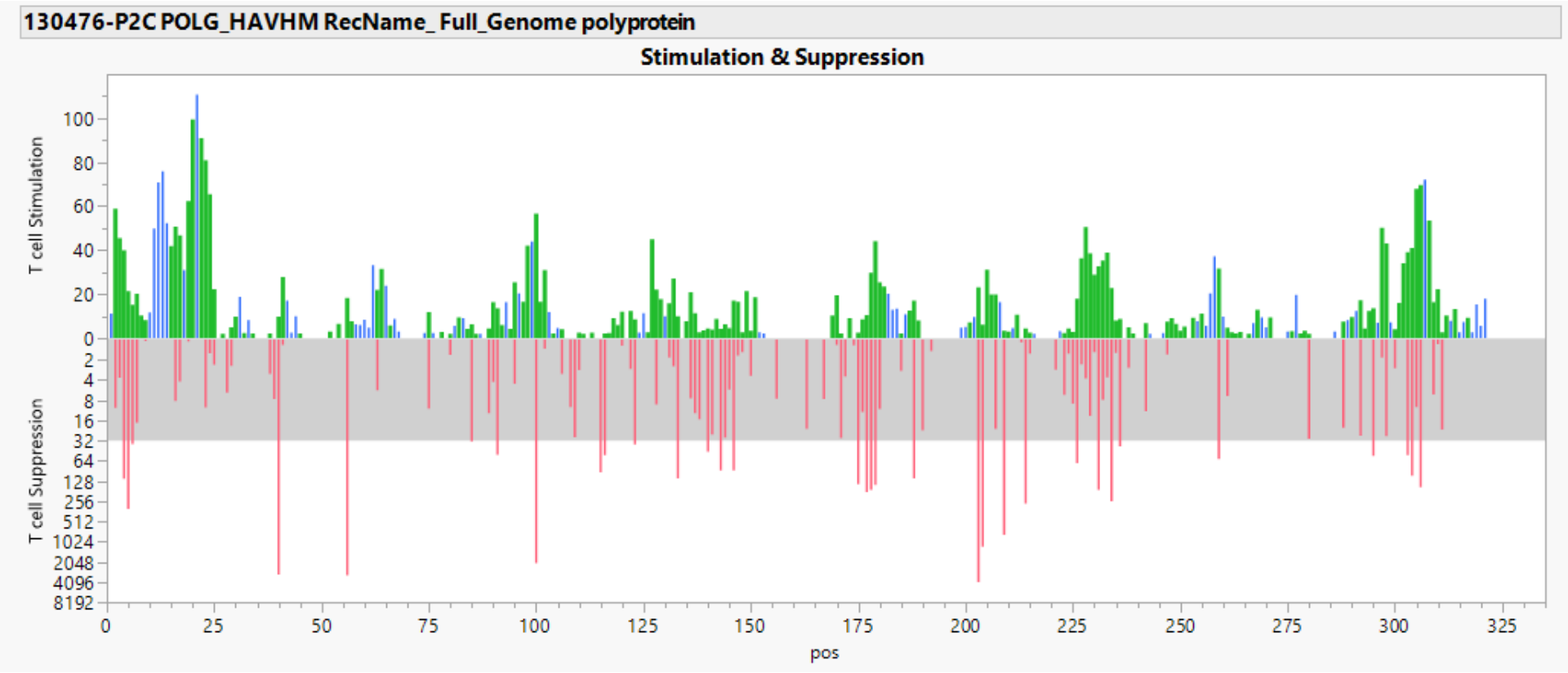

130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.04
0.09
0.13
0.18
0.22
0.98
1.73
2.49
3.24
4.00
130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.07
0.10
0.13
0.17
1.33
2.50
3.67
4.83
6.00


Hierarchical Cluster MHC I
Method = Ward
130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_0301
nA_1101
nA_3101
nA_2902
nA_3002
nA_3301
nA_6801
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nB_5101
nB_0702
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_3201
nA_2403
nH-2-Kd
nB_1503
nB_2705
nA_2601
nB_1501
nH-2-Db
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nB_1801
nB_4403
nB_4402
nB_4501
nH-2-Kk
nB_4001
nB_4002
-3.97
-3.18
-2.38
-1.59
-0.79
0.00
1.01
2.02
3.03
4.05
5.06
25
50
75
100
125
150
175
200
225
250
275
300
325
N-term Amino Acid of 9-mer


130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nA_2601
nB_5301
nA_3201
nB_5701
nB_5801
nB_0801
nH-2-Kb
nA_2902
nB_1501
nH-2-Dd
nB_1503
nA_2301
nA_2403
nA_2402
nH-2-Kd
nB_2705
nH-2-Kk
nB_1801
nB_4501
nB_4403
nB_4402
nB_4001
nB_4002
nA_0201
nA_0202
nA_0206
nA_0203
nA_6802
nA_6901
nH-2-Db
nB_0702
nB_5101
nB_5401
nA_0301
nA_1101
nA_3101
nA_3002
nA_6801
nA_3301
nA_3001
nH-2-Ld
0.00
0.02
0.04
0.06
0.09
0.11
0.69
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.1096
2.2193
3.3289
4.4385
5.5482
18.239
30.929
43.619
56.31
69
0
25
50
75
100
125
150
175
200
225
250
275
300
325
N-term Amino Acid of 15-mer

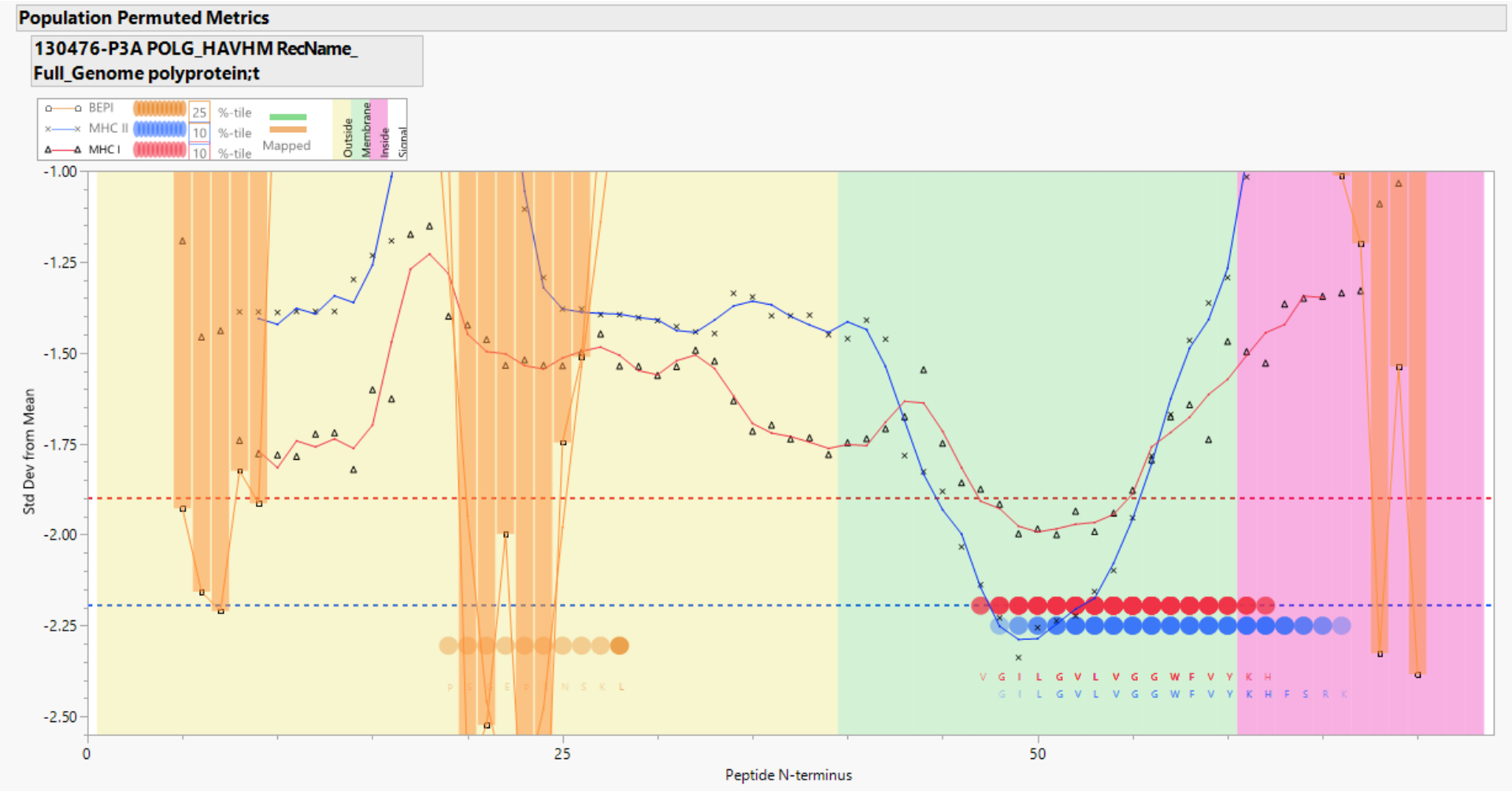

130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10.0
0.0
10.0
0.0
10.0
0.0
0
25
50
75
100
125
150
175
200
225
250
275
300
325
pos

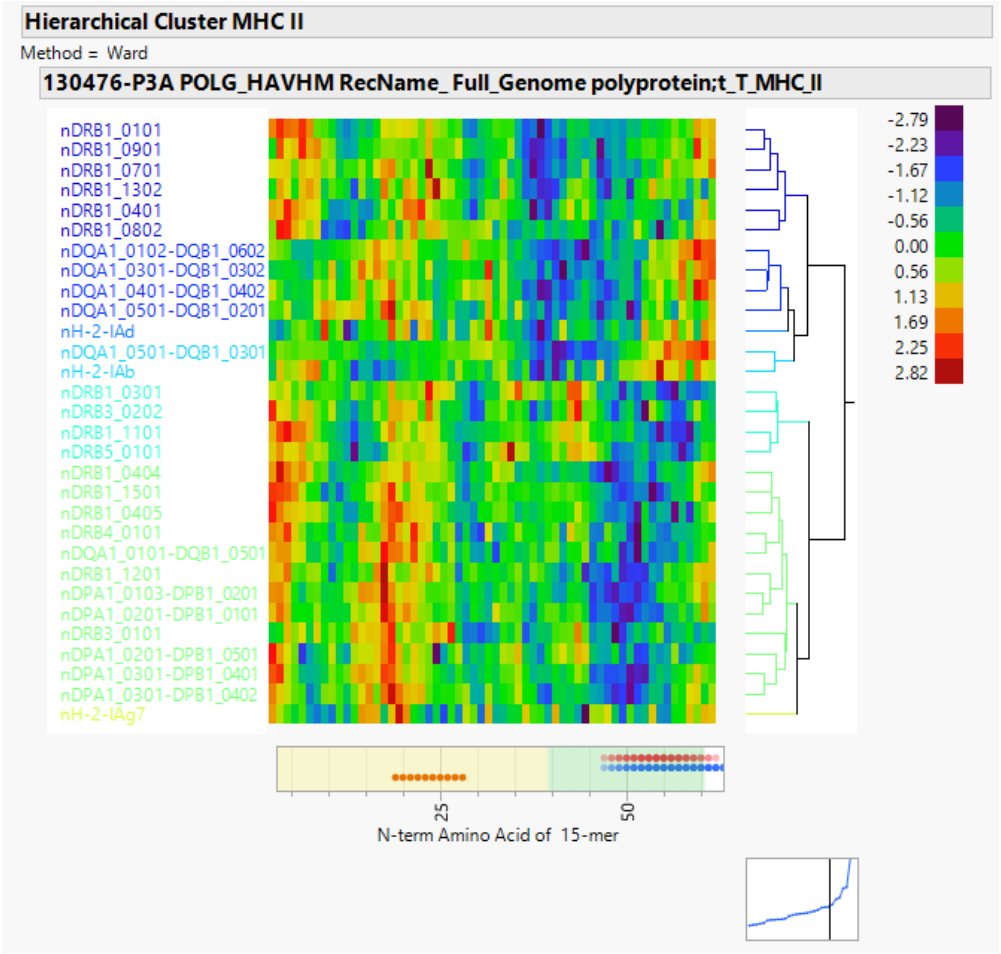

130476-P2C POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
100
80
60
40
20
0
T cell Suppression
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
25
50
75
100
125
150
175
200
225
250
275
300
325
pos

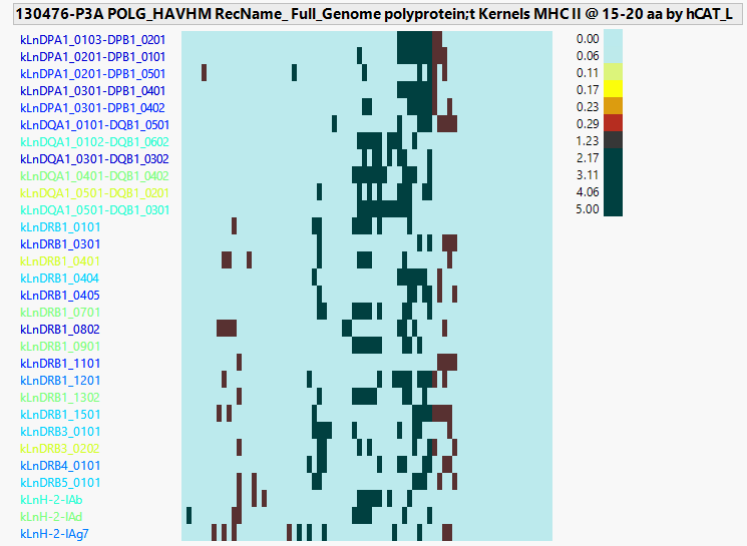
Protein 3A

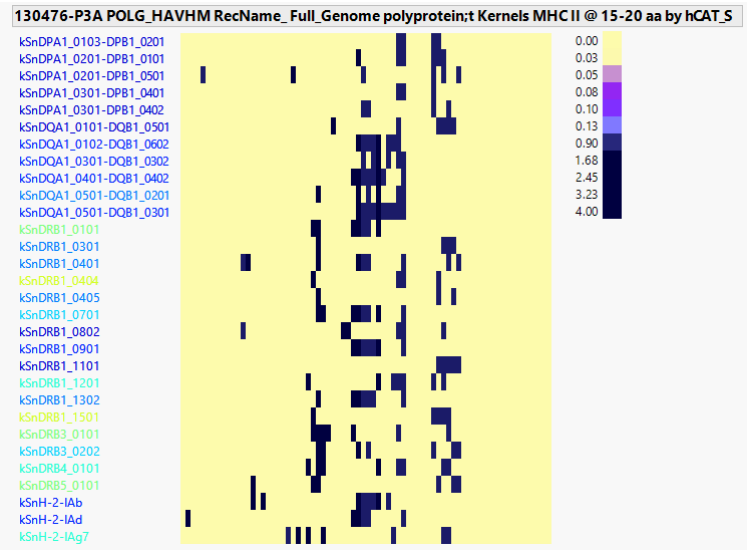
Population Permuted Metrics

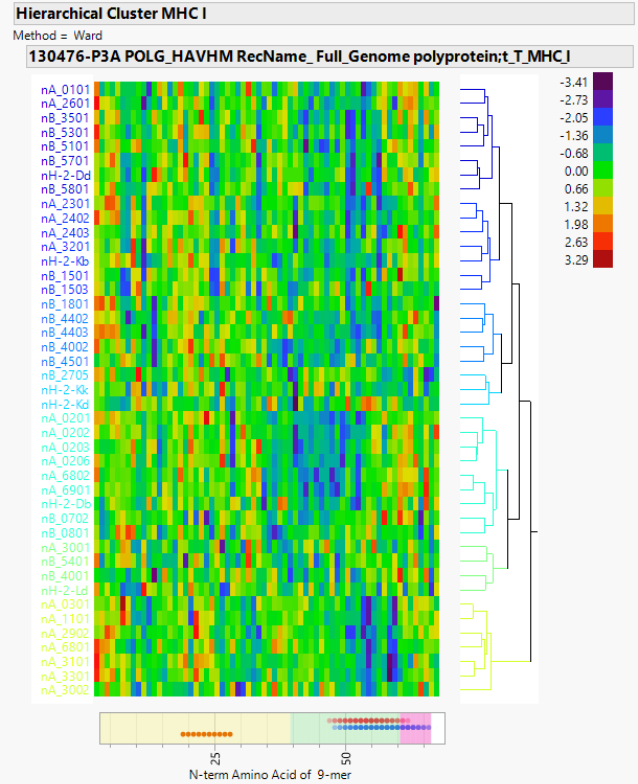
130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t

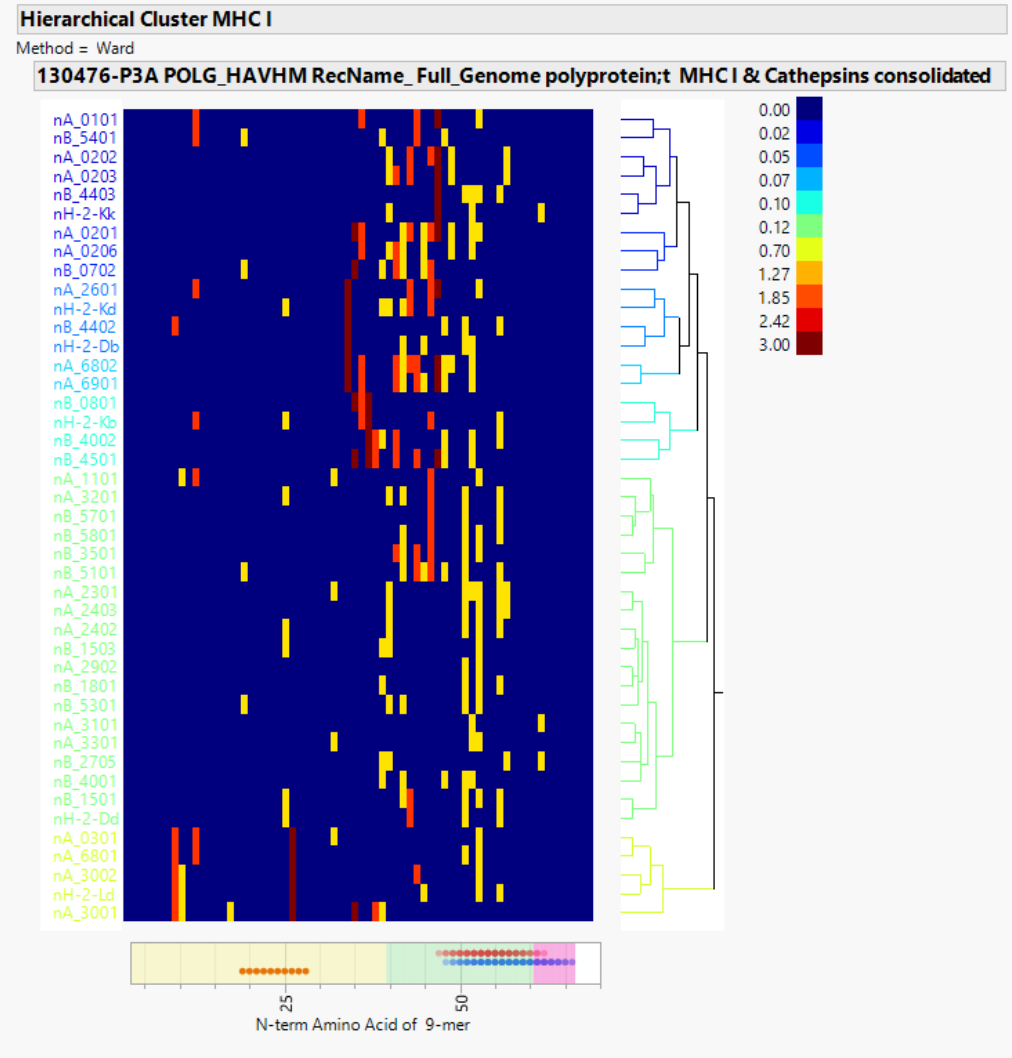


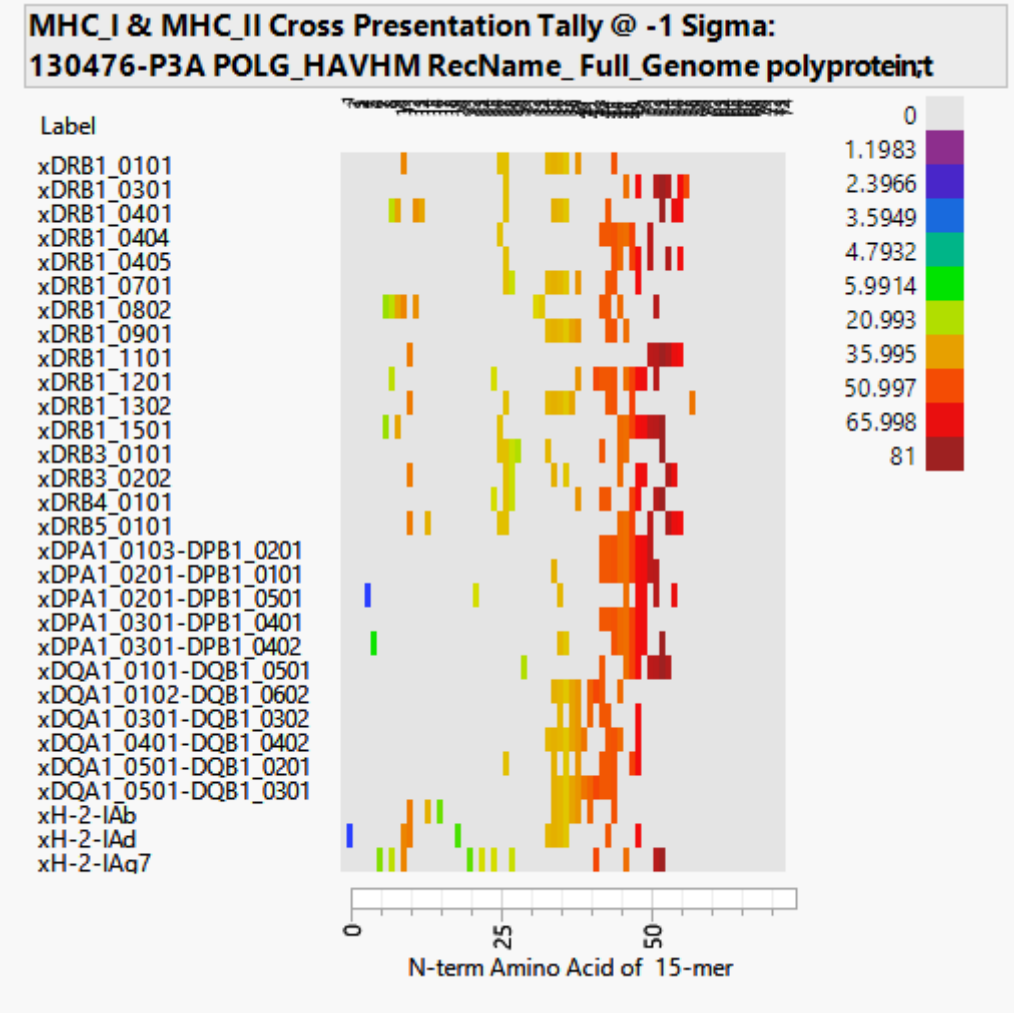

Hierarchical Cluster MHC II
Method = Ward
130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB1_0802
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0201
nH-2-IAd
nDQA1_0501-DQB1_0301
nH-2-IAb
nDRB1_0301
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDQA1_0101-DQB1_0501
nDRB1_1201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDRB3_0101
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nH-2-IAg7
-2.79
-2.23
-1.67
-1.12
-0.56
0.00
0.56
1.13
1.69
2.25
2.82
25
50
N-term Amino Acid of 15-mer

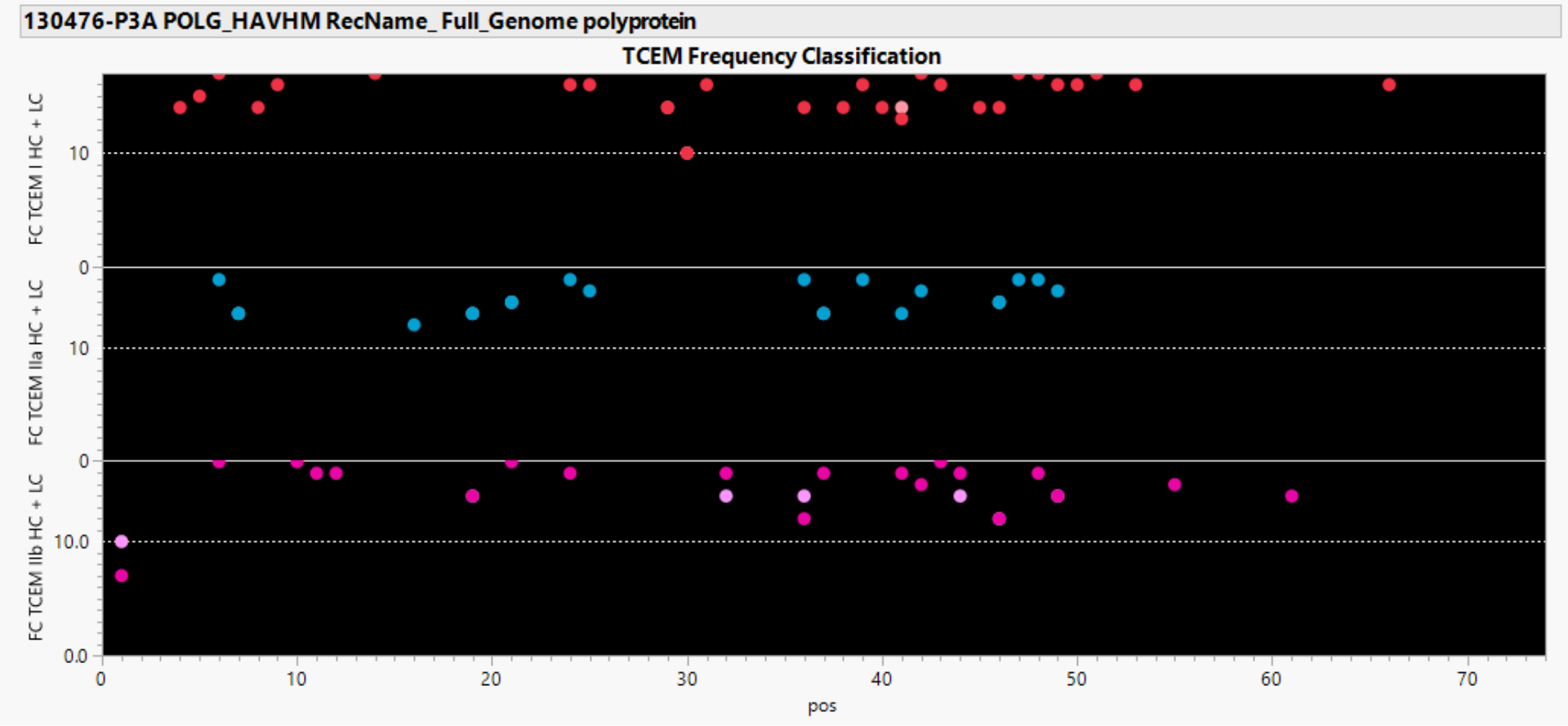

130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.06
0.11
0.17
0.23
0.29
1.23
2.17
3.11
4.06
5.00

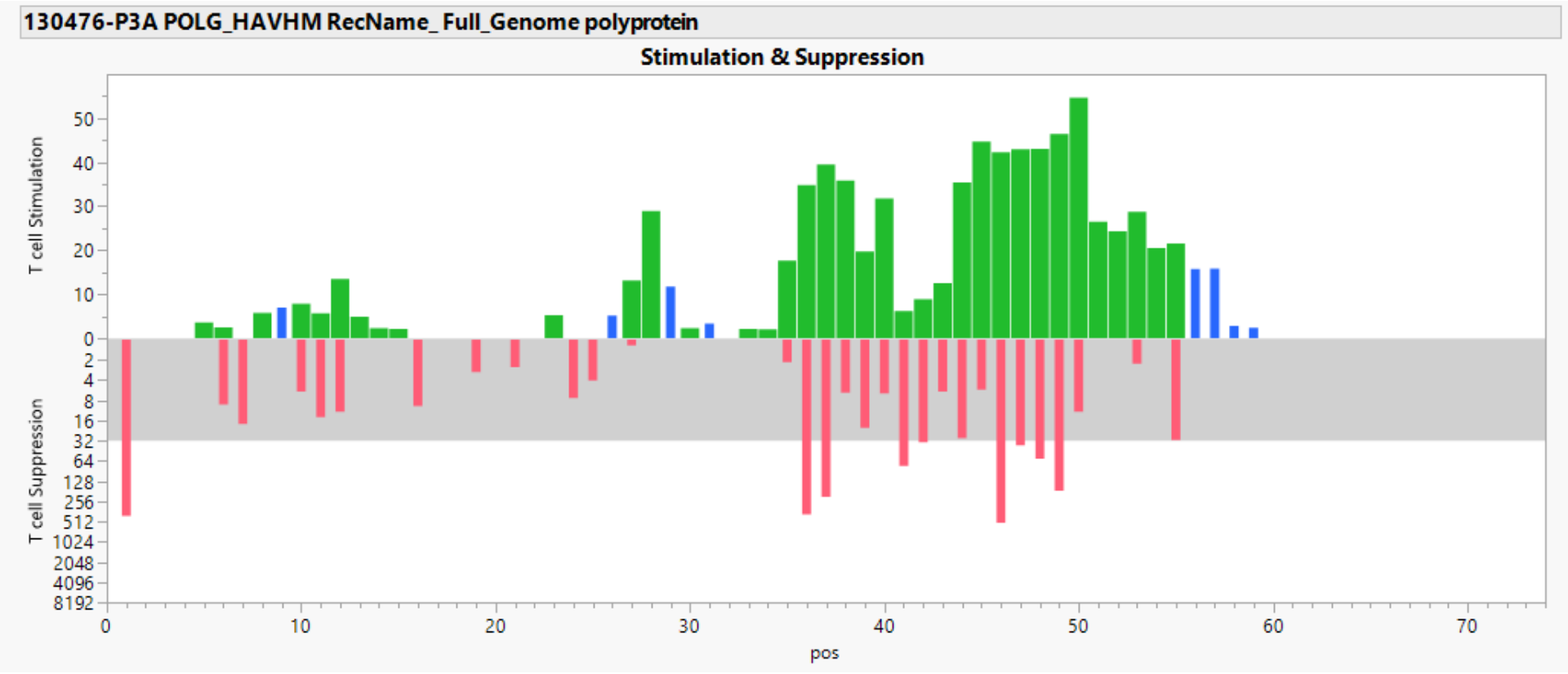

130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.05
0.08
0.10
0.13
0.90
1.68
2.45
3.23
4.00


Hierarchical Cluster MHC I
Method = Ward
130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5101
nB_5701
nH-2-Dd
nB_5801
nA_2301
nA_2402
nA_2403
nA_3201
nH-2-Kb
nB_1501
nB_1503
nB_1801
nB_4402
nB_4403
nB_4002
nB_4501
nB_2705
nH-2-Kk
nH-2-Kd
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nH-2-Db
nB_0702
nB_0801
nA_3001
nB_5401
nB_4001
nH-2-Ld
nA_0301
nA_1101
nA_2902
nA_6801
nA_3101
nA_3301
nA_3002
-3.41
-2.73
-2.05
-1.36
-0.68
0.00
0.66
1.32
1.98
2.63
3.29
25
50
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nB_5401
nA_0202
nA_0203
nB_4403
nH-2-Kk
nA_0201
nA_0206
nB_0702
nA_2601
nH-2-Kd
nB_4402
nH-2-Db
nA_6802
nA_6901
nB_0801
nH-2-Kb
nB_4002
nB_4501
nA_1101
nA_3201
nB_5701
nB_5801
nB_3501
nB_5101
nA_2301
nA_2403
nA_2402
nB_1503
nA_2902
nB_1801
nB_5301
nA_3101
nA_3301
nB_2705
nB_4001
nB_1501
nH-2-Dd
nA_0301
nA_6801
nA_3002
nH-2-Ld
nA_3001
0.00
0.02
0.05
0.07
0.10
0.12
0.70
1.27
1.85
2.42
3.00
25
50
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma:
130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.1983
2.3966
3.5949
4.7932
5.9914
20.993
35.995
50.997
65.998
81
0
25
50
N-term Amino Acid of 15-mer

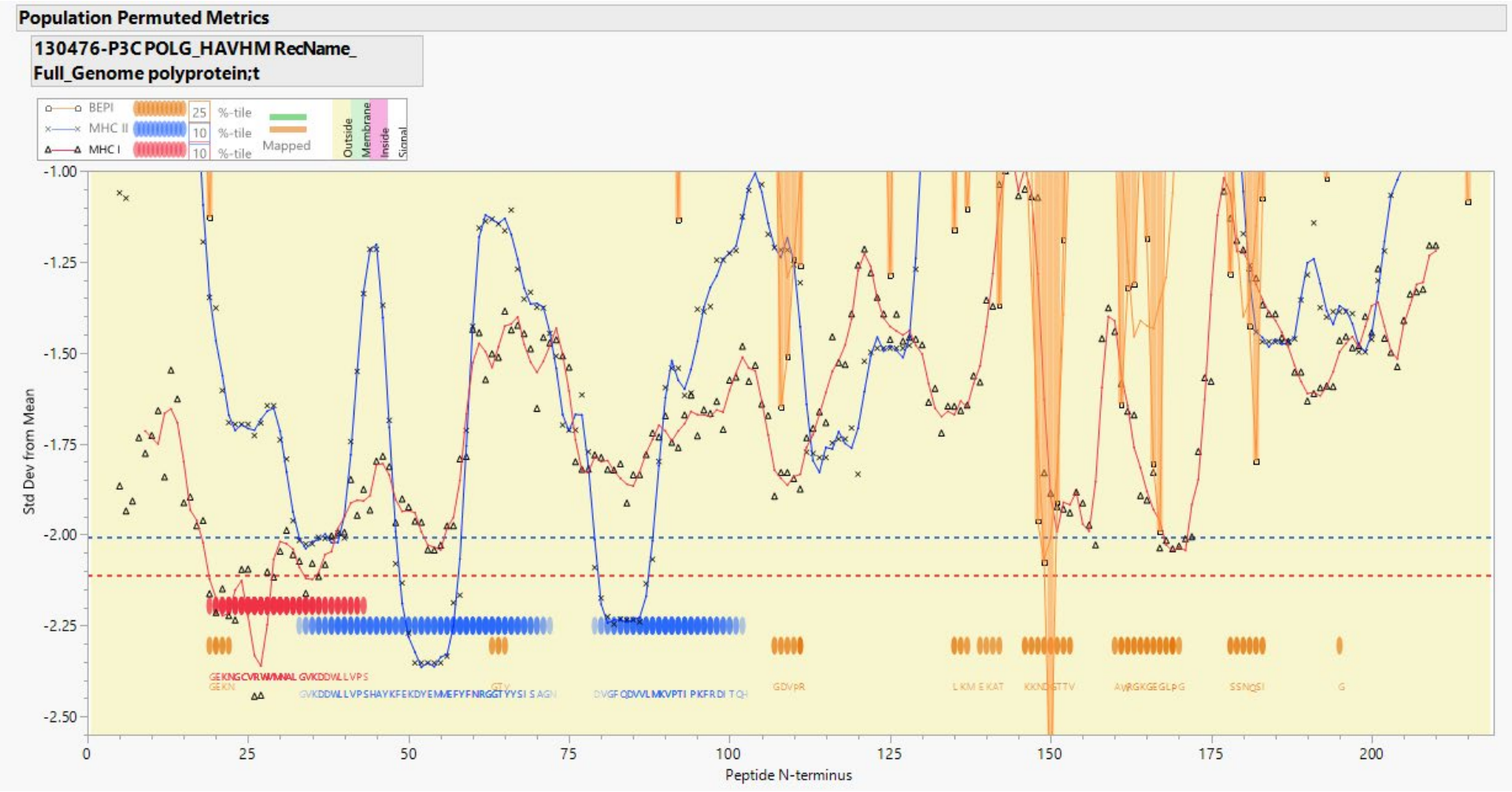

130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10.0
0.0
0
10
20
30
40
50
60
70
pos

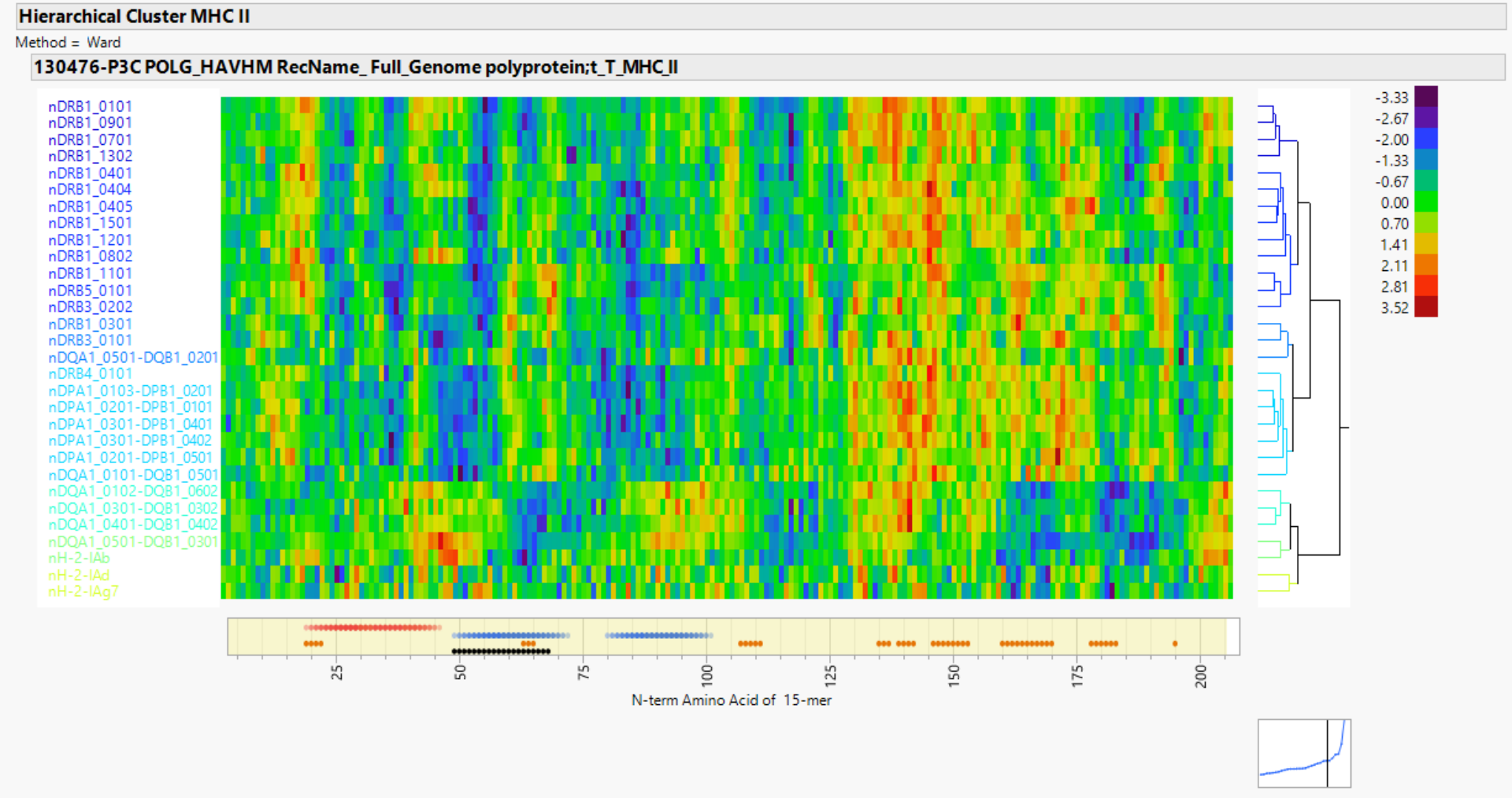

130476-P3A POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
50
40
30
20
10
0
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
10
20
30
40
50
60
70
pos

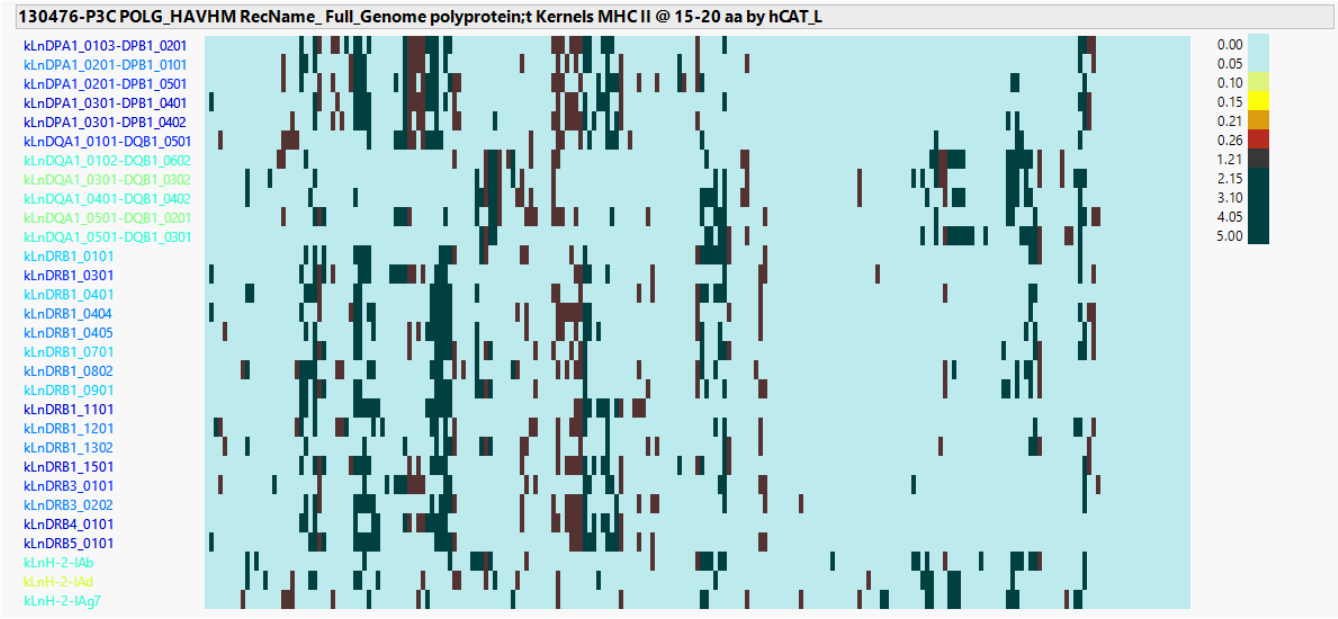
Protein 3C

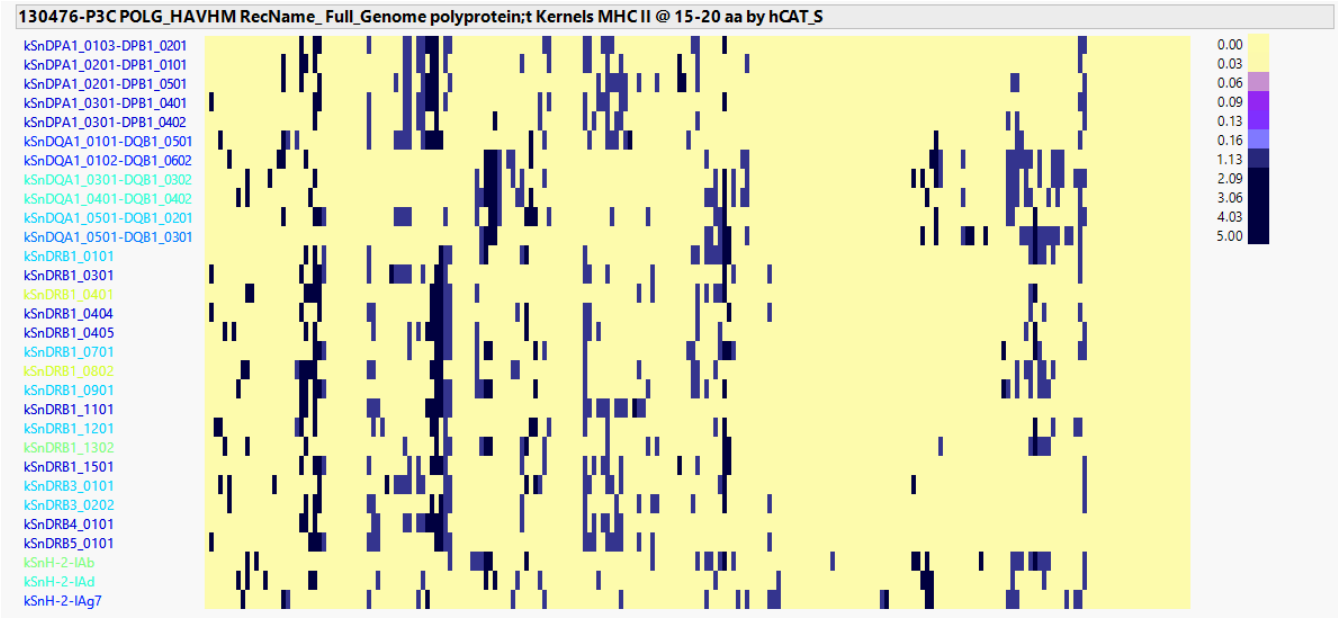
Population Permuted Metrics

130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t




Hierarchical Cluster MHC II
Method = Ward
130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB1_1201
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB3_0202
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDRB4_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.33
-2.67
-2.00
-1.33
-0.67
0.00
0.70
1.41
2.11
2.81
3.52
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

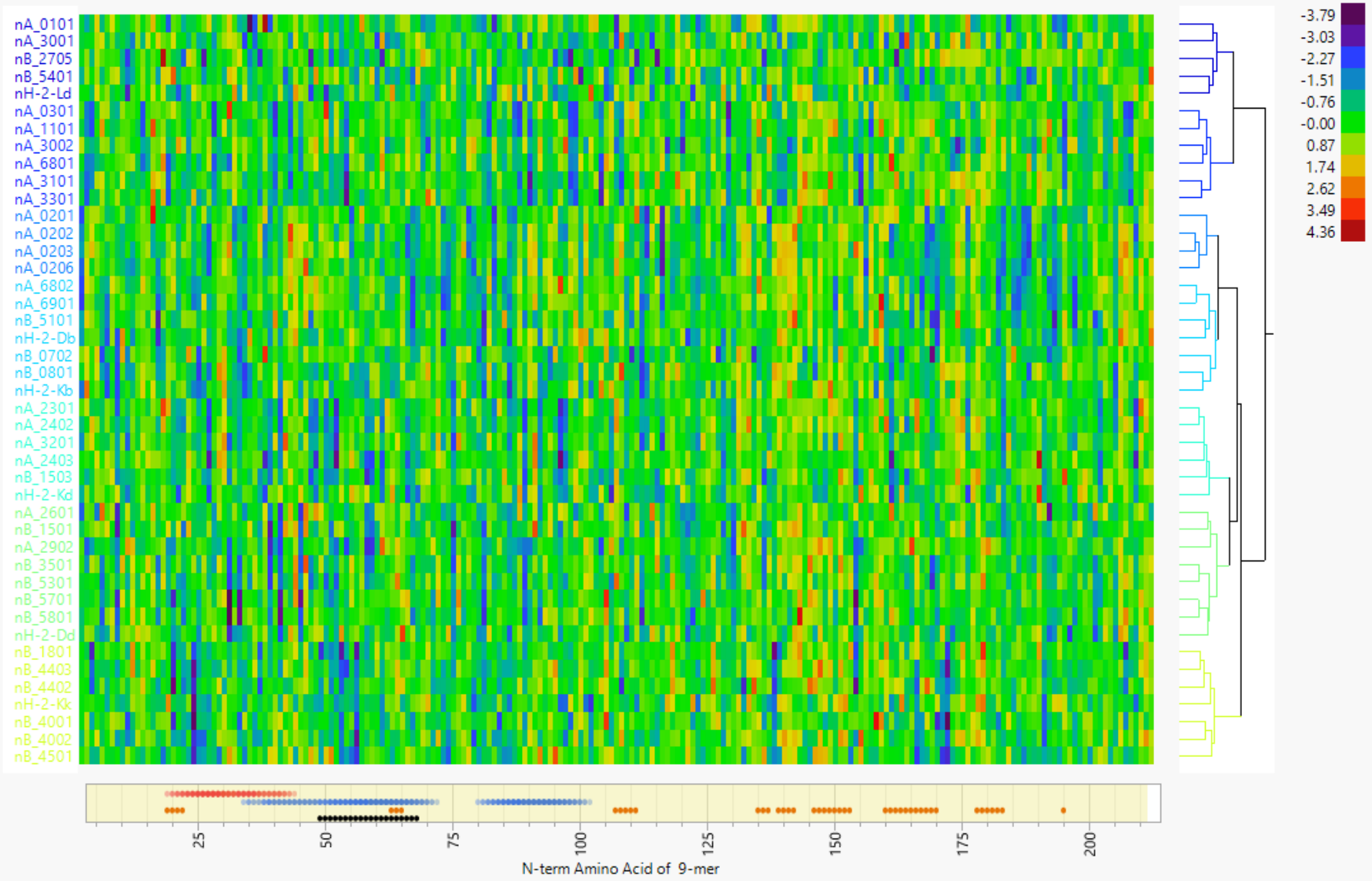

130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.05
0.10
0.15
0.21
0.26
1.21
2.15
3.10
4.05
5.00

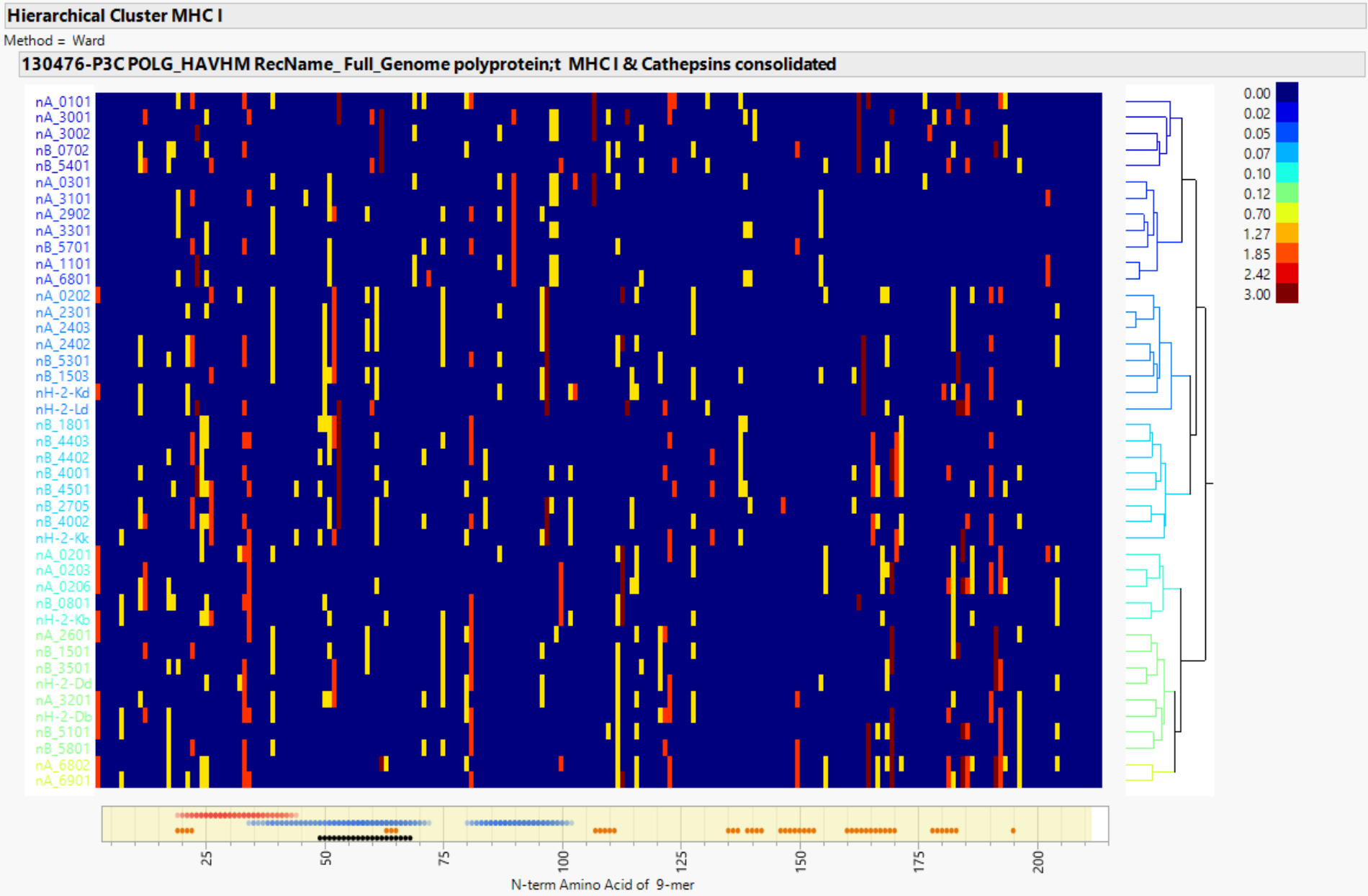

130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.06
0.09
0.13
0.16
1.13
2.09
3.06
4.03
5.00

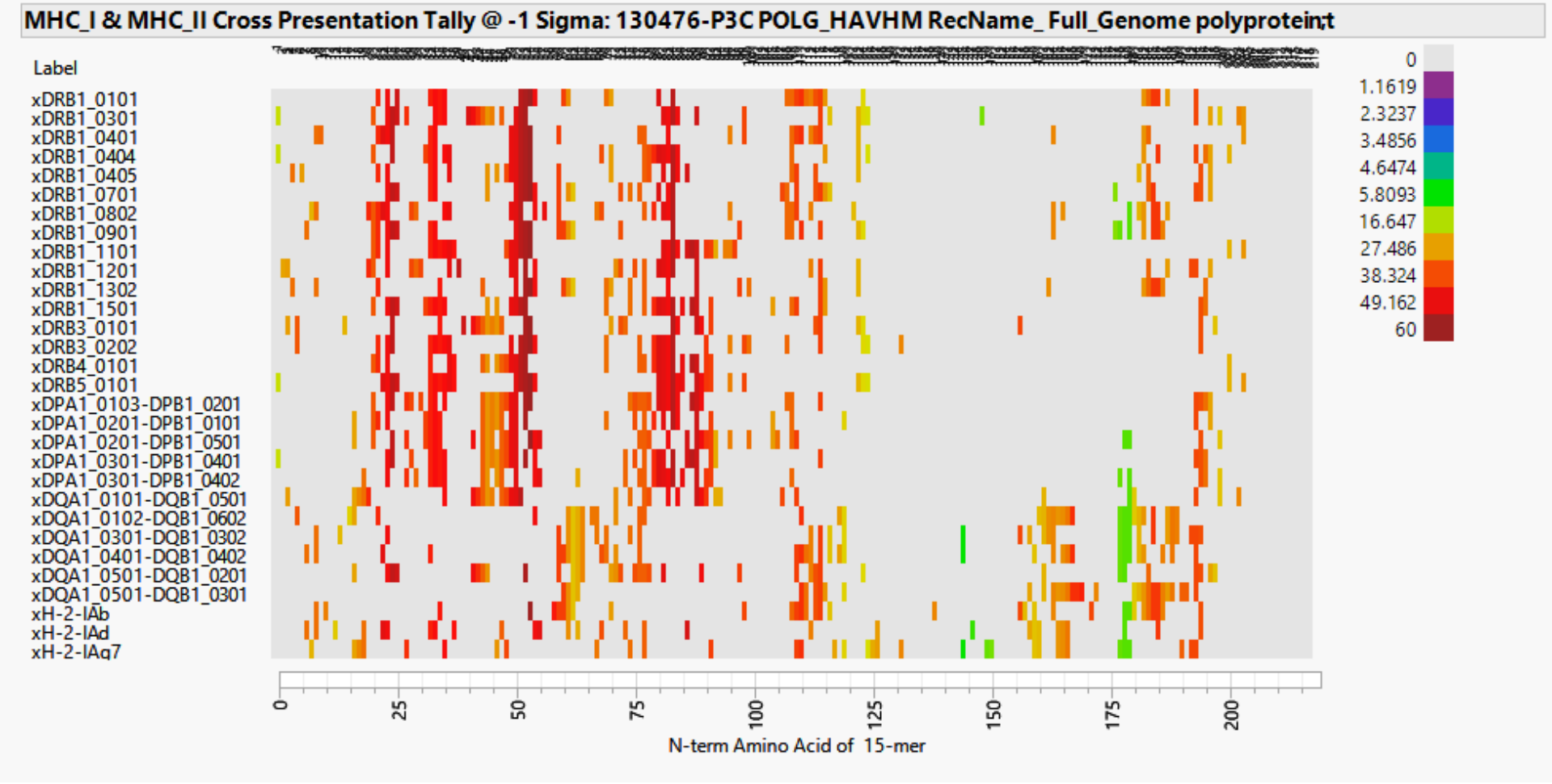
Hierarchical Cluster MHC I

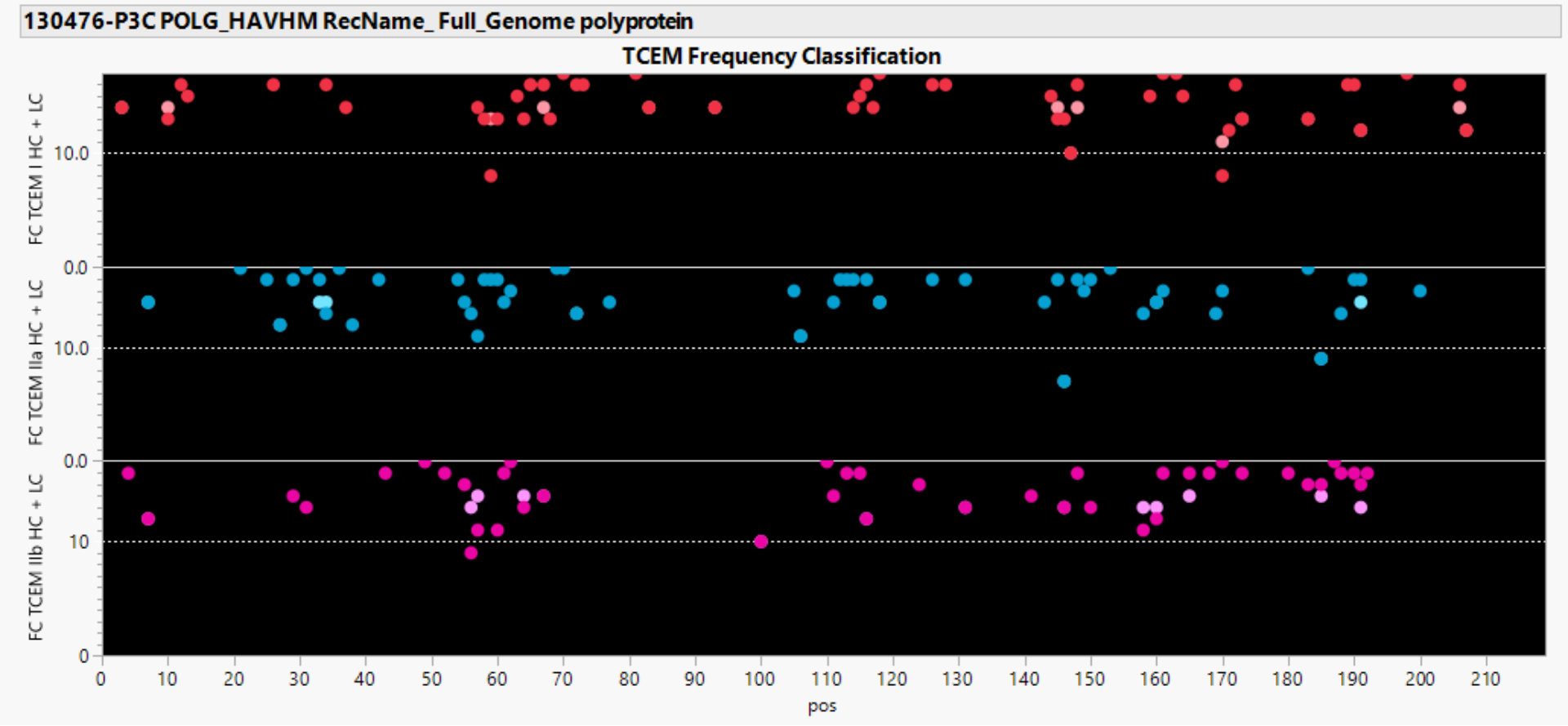
Method = Ward

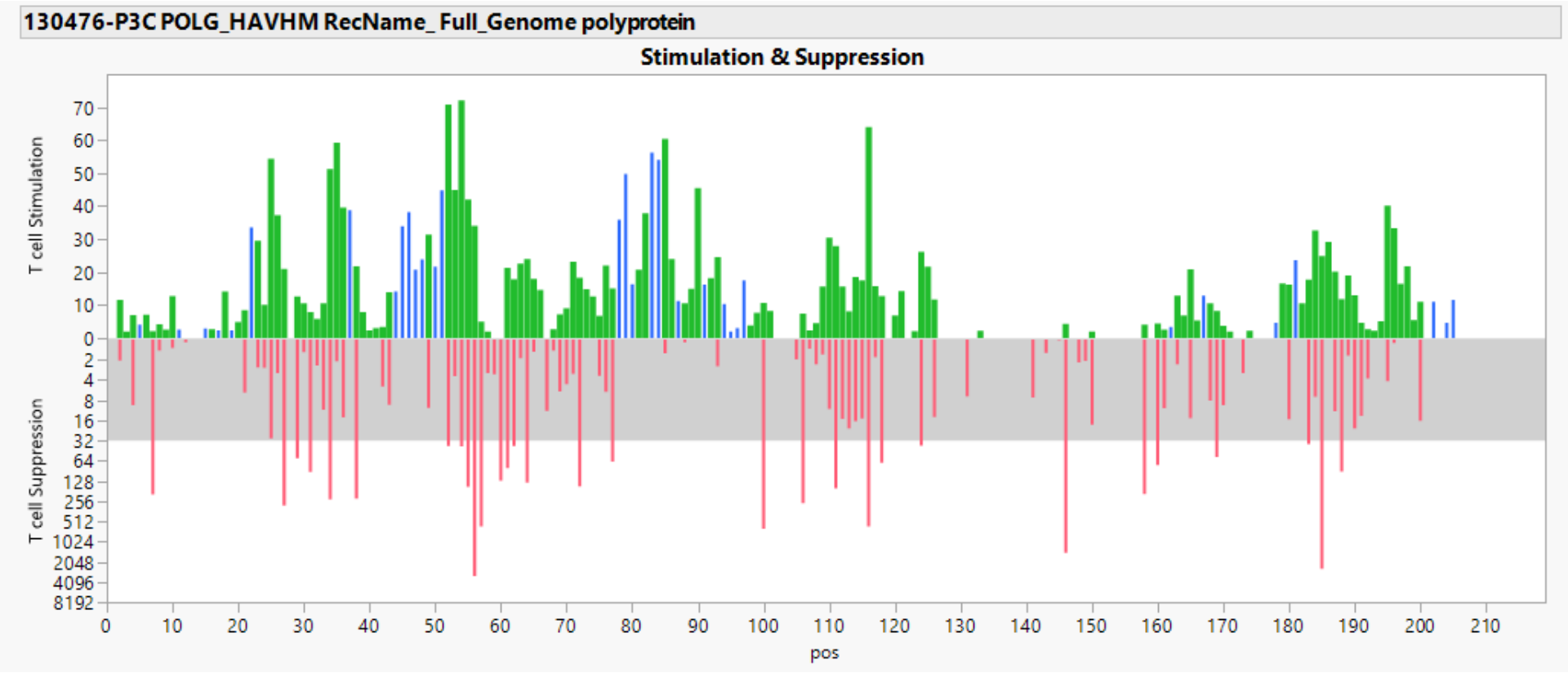
130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I



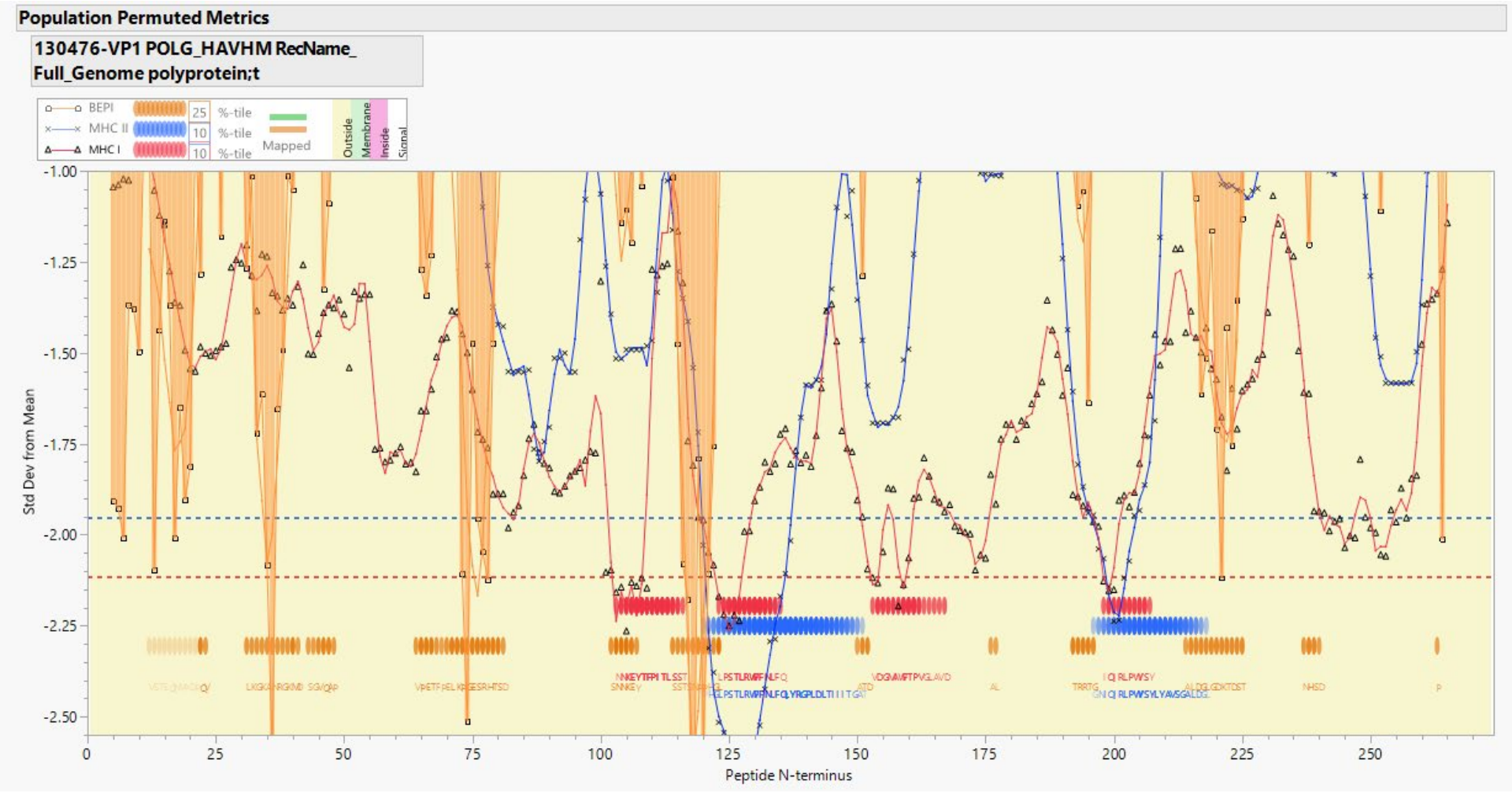

Hierarchical Cluster MHC I
Method = Ward
130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nA_3001
nA_3002
nB_0702
nB_5401
nA_0301
nA_3101
nA_2902
nA_3301
nB_5701
nA_1101
nA_6801
nA_0202
nA_2301
nA_2403
nA_2402
nB_5301
nB_1503
nH-2-Kd
nH-2-Ld
nB_1801
nB_4403
nB_4402
nB_4001
nB_4501
nB_2705
nB_4002
nH-2-Kk
nA_0201
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_2601
nB_1501
nB_3501
nH-2-Dd
nA_3201
nH-2-Db
nB_5101
nB_5801
nA_6802
nA_6901
0.00
0.02
0.05
0.07
0.10
0.12
0.70
1.27
1.85
2.42
3.00
25
50
75
100
125
150
175
200
N-term Amino Acid of 9-mer

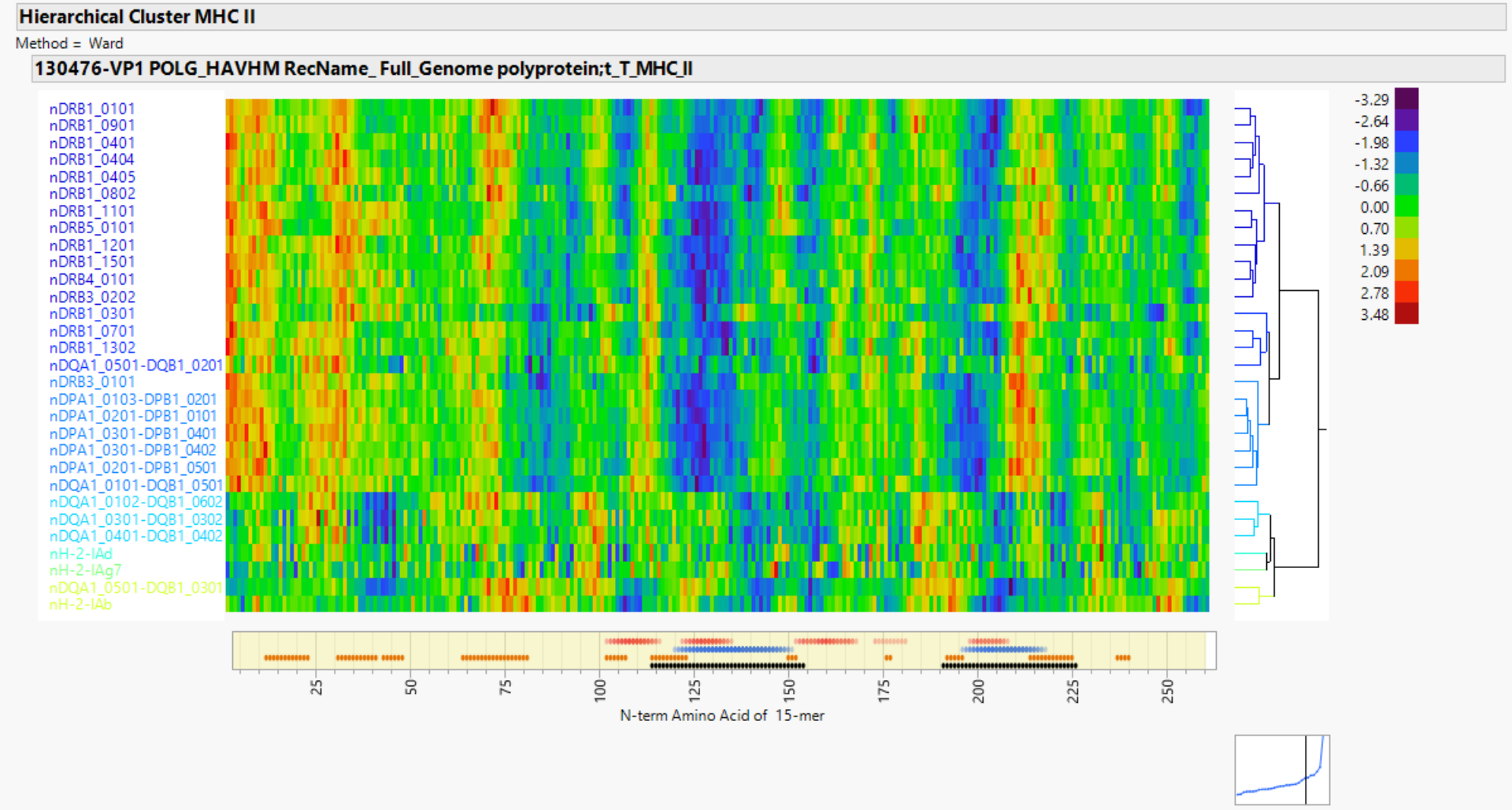

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.1619
2.3237
3.4856
4.6474
5.8093
16.647
27.486
38.324
49.162
60
0
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

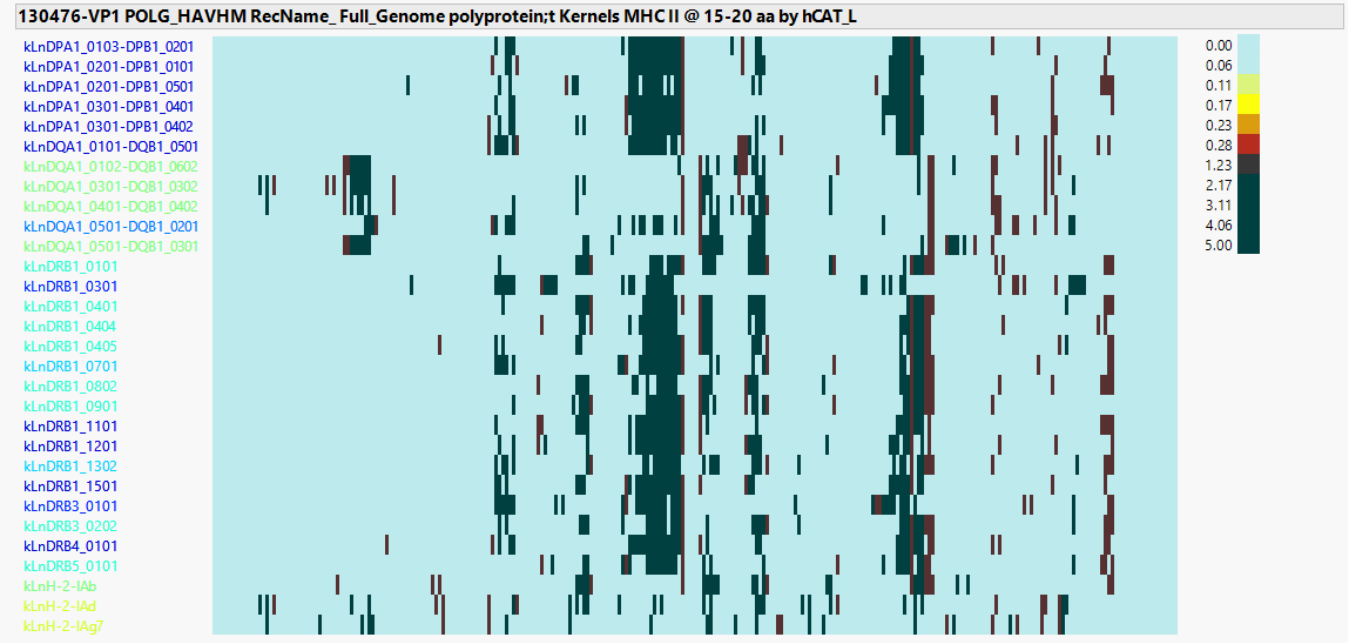

130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10.0
0.0
10.0
0.0
10
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210
pos

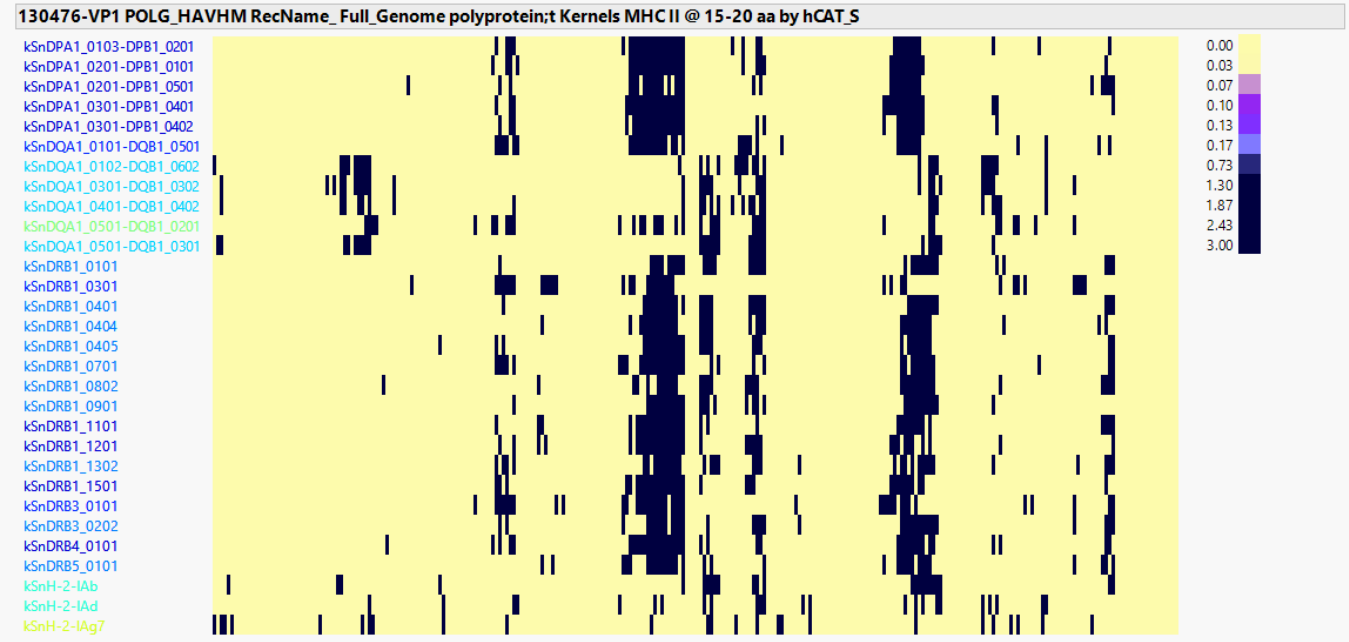

130476-P3C POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

Protein VP1


Population Permuted Metrics
130476-VP1 POLG_HAVHM RecName_
Full_Genome polyprotein;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Std Dev from Mean
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
225
250
Peptide N-terminus


Hierarchical Cluster MHC II
Method = Ward
130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0401
nDRB1_0404
nDRB1_0405
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB1_1201
nDRB1_1501
nDRB4_0101
nDRB3_0202
nDRB1_0301
nDRB1_0701
nDRB1_1302
nDQA1_0501-DQB1_0201
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.29
-2.64
-1.98
-1.32
-0.66
0.00
0.70
1.39
2.09
2.78
3.48
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 15-mer

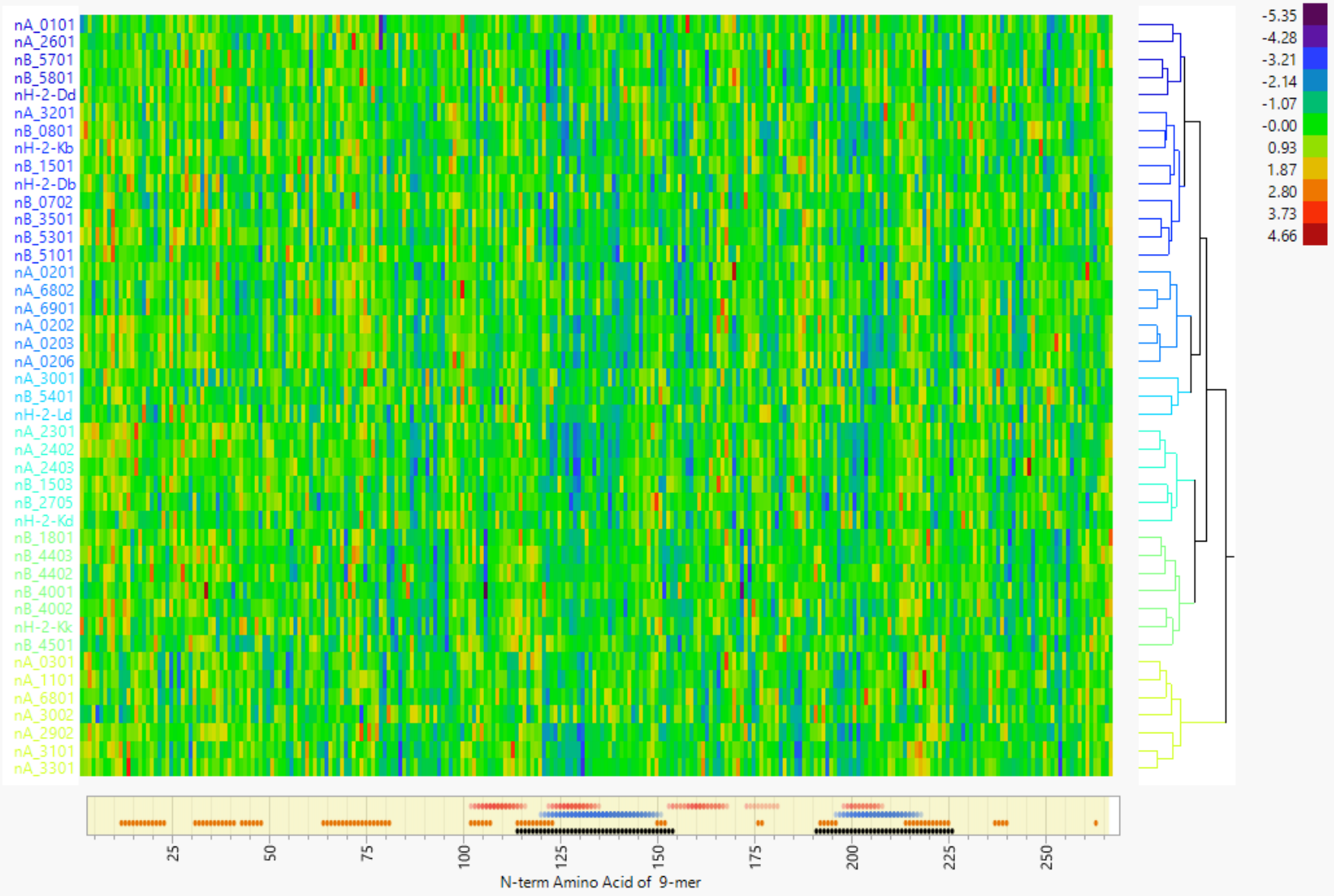

130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.06
0.11
0.17
0.23
0.28
1.23
2.17
3.11
4.06
5.00


130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.07
0.10
0.13
0.17
0.73
1.30
1.87
2.43
3.00

Hierarchical Cluster MHC I

Method = Ward

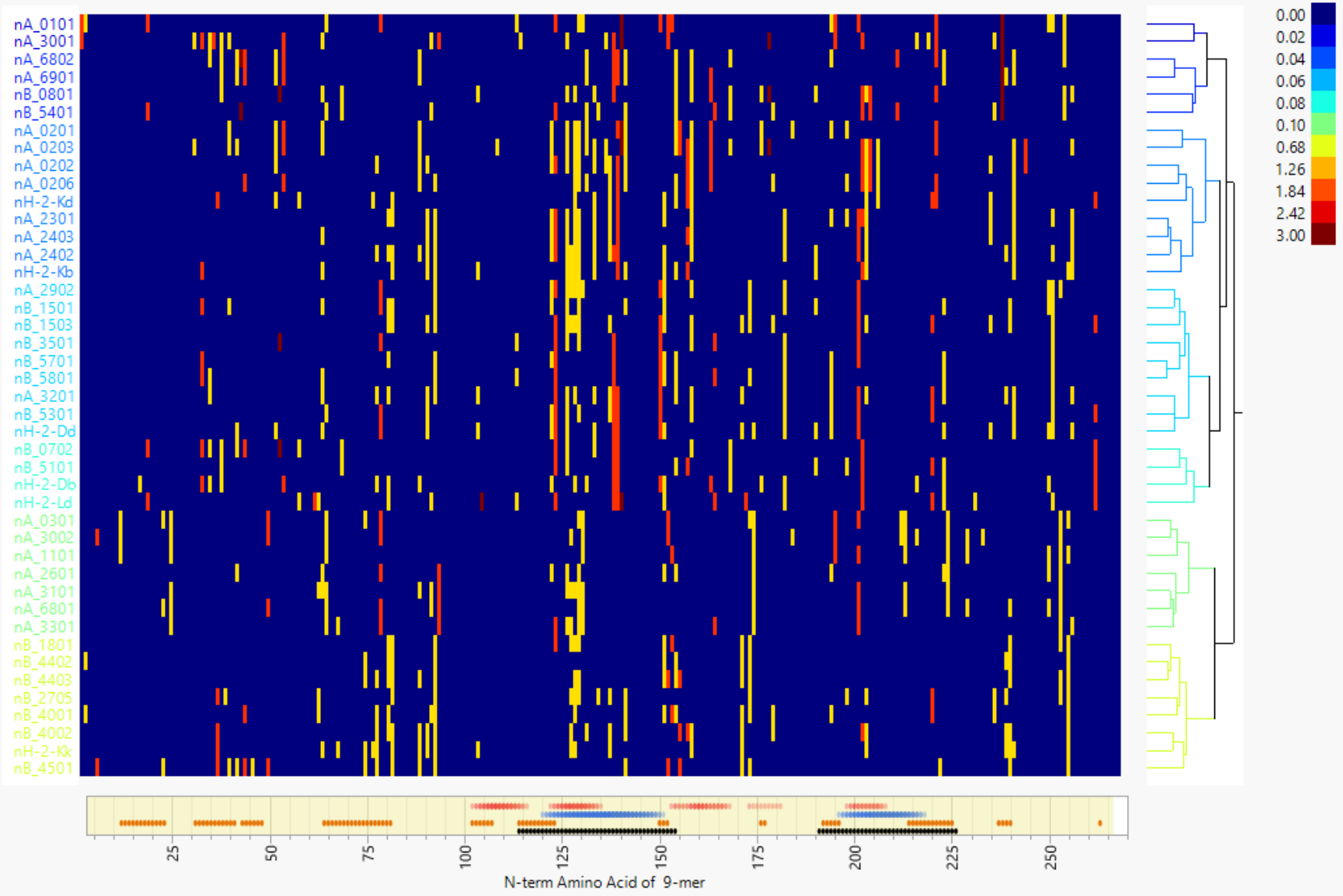
130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I

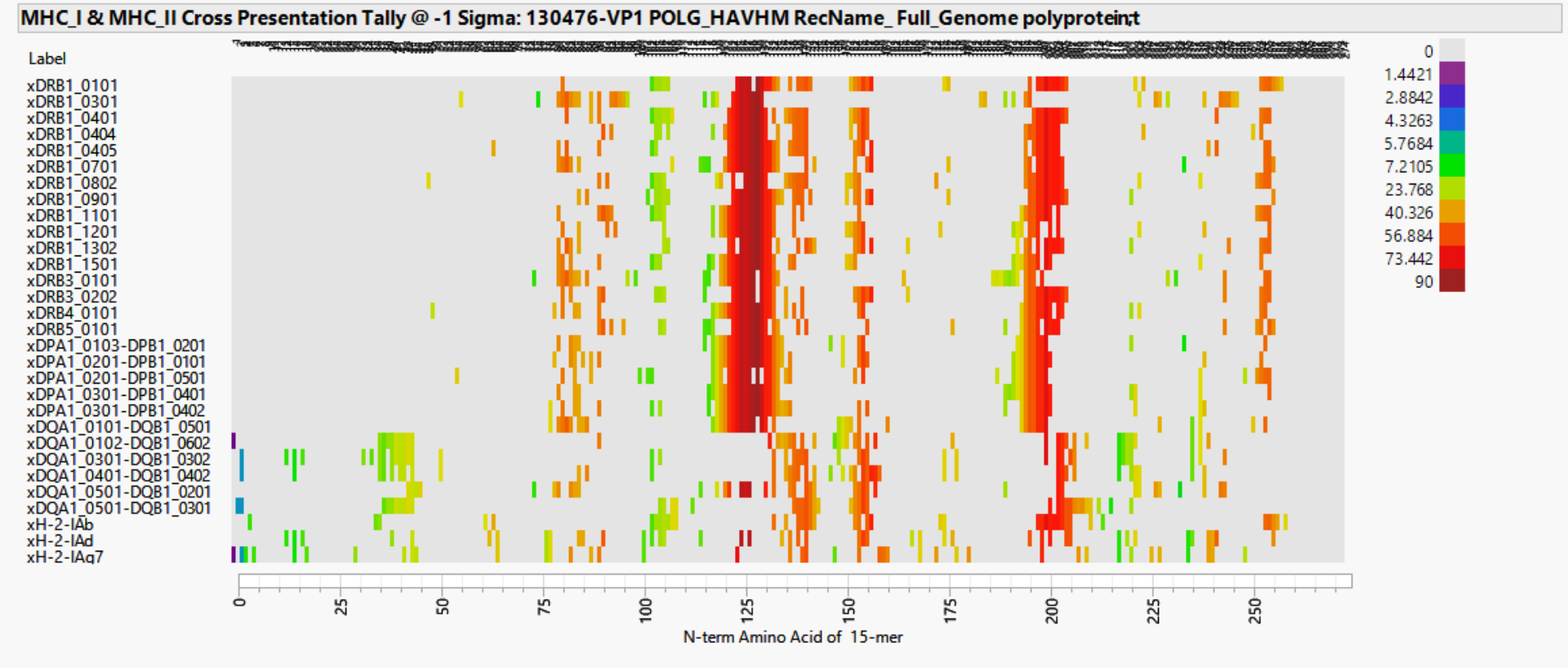


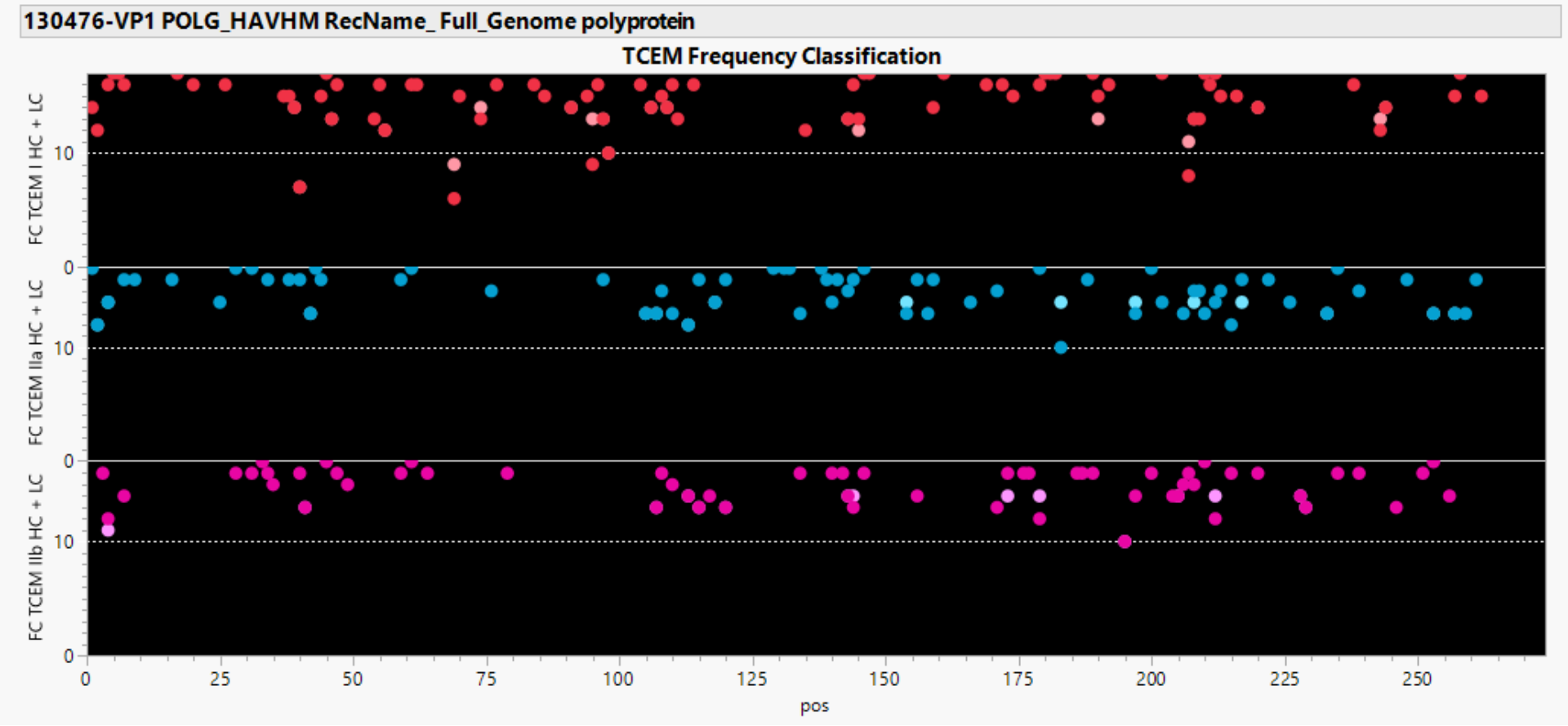
Hierarchical Cluster MHC I

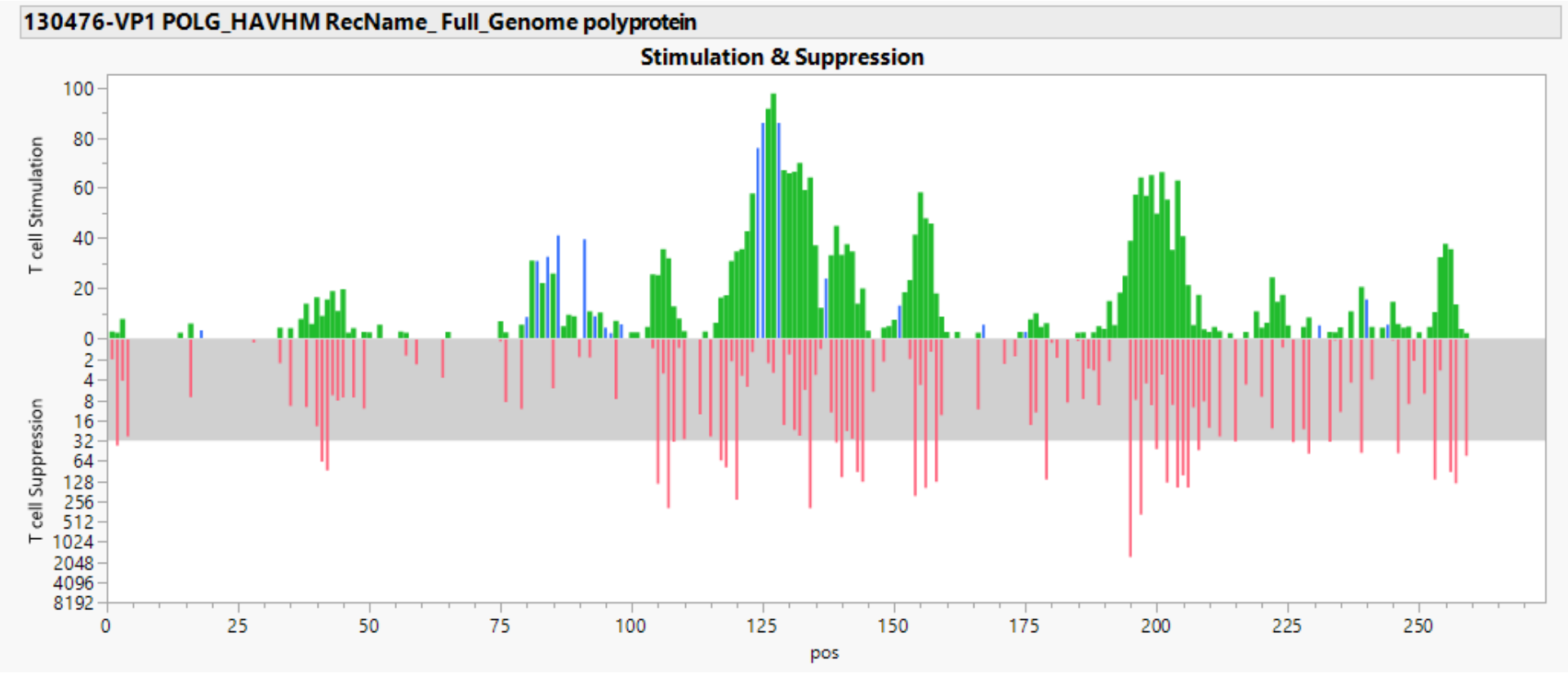
Method = Ward

130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated

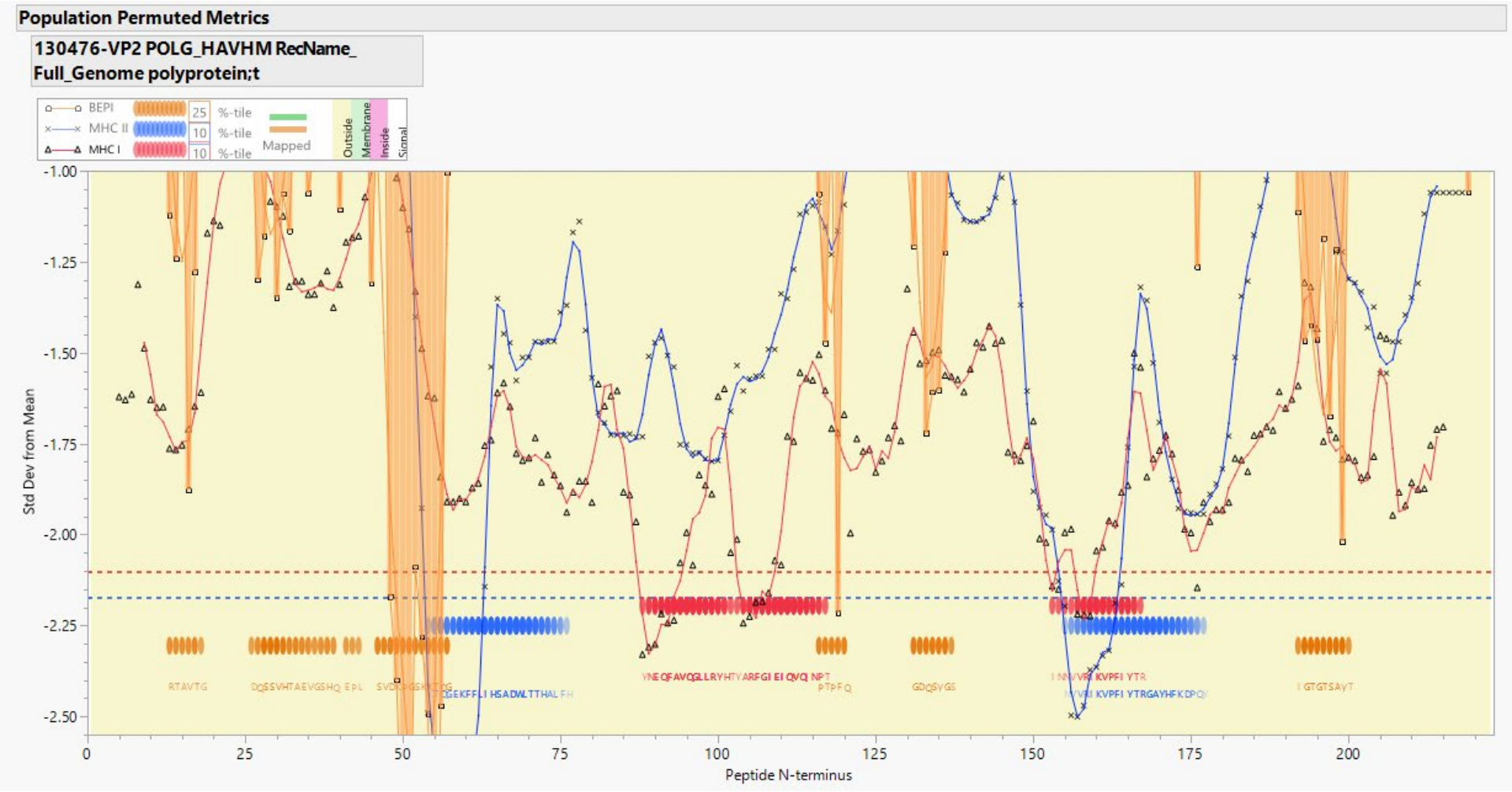


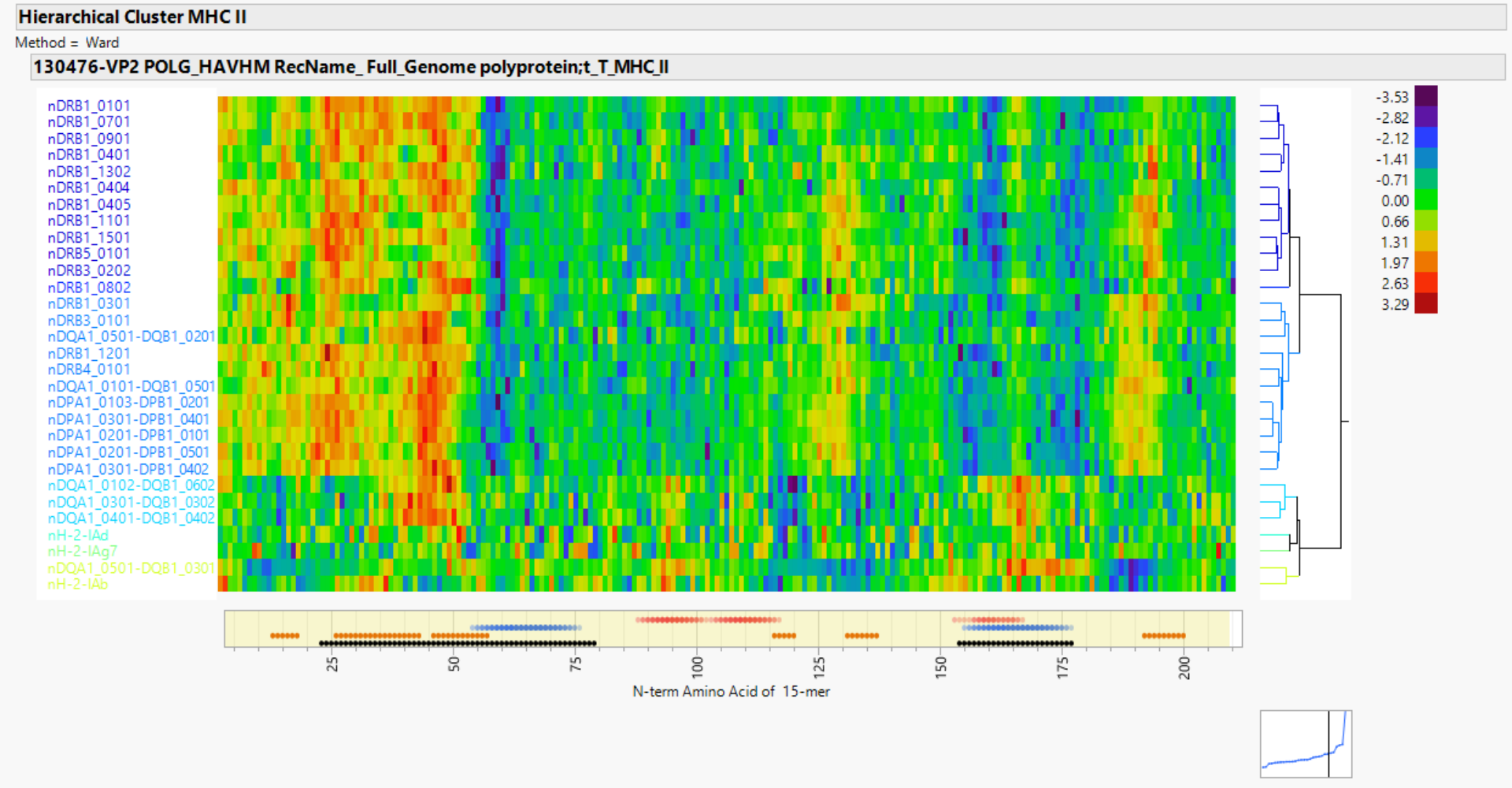

MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAq7
0
1.4421
2.8842
4.3263
5.7684
7.2105
23.768
40.326
56.884
73.442
90
0
25
50
75
100
125
150
175
200
225
250
N-term Amino Acid of 15-mer

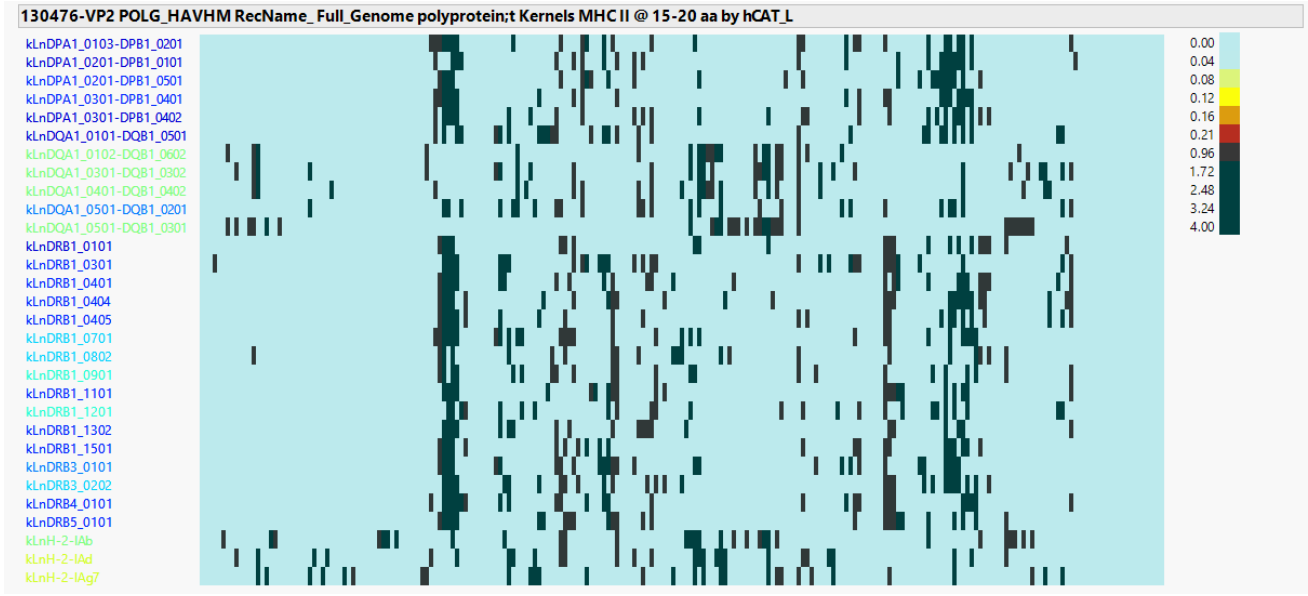

130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0
25
50
75
100
125
150
175
200
225
250
pos

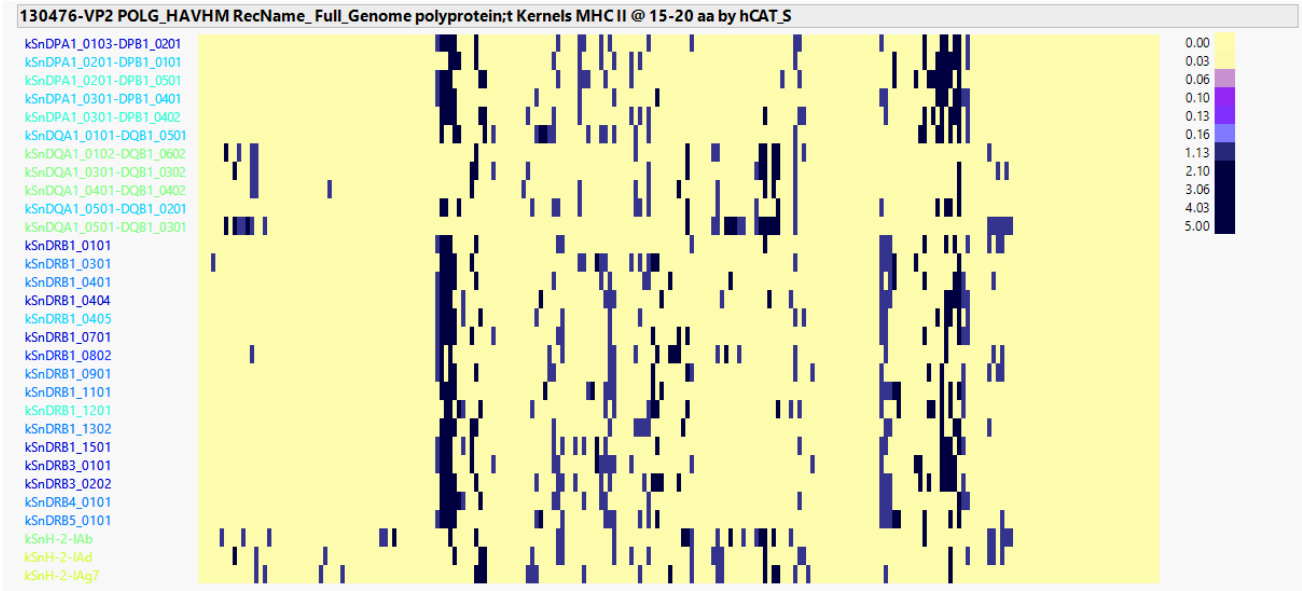

130476-VP1 POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
100
80
60
40
20
0
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
25
50
75
100
125
150
175
200
225
250
pos

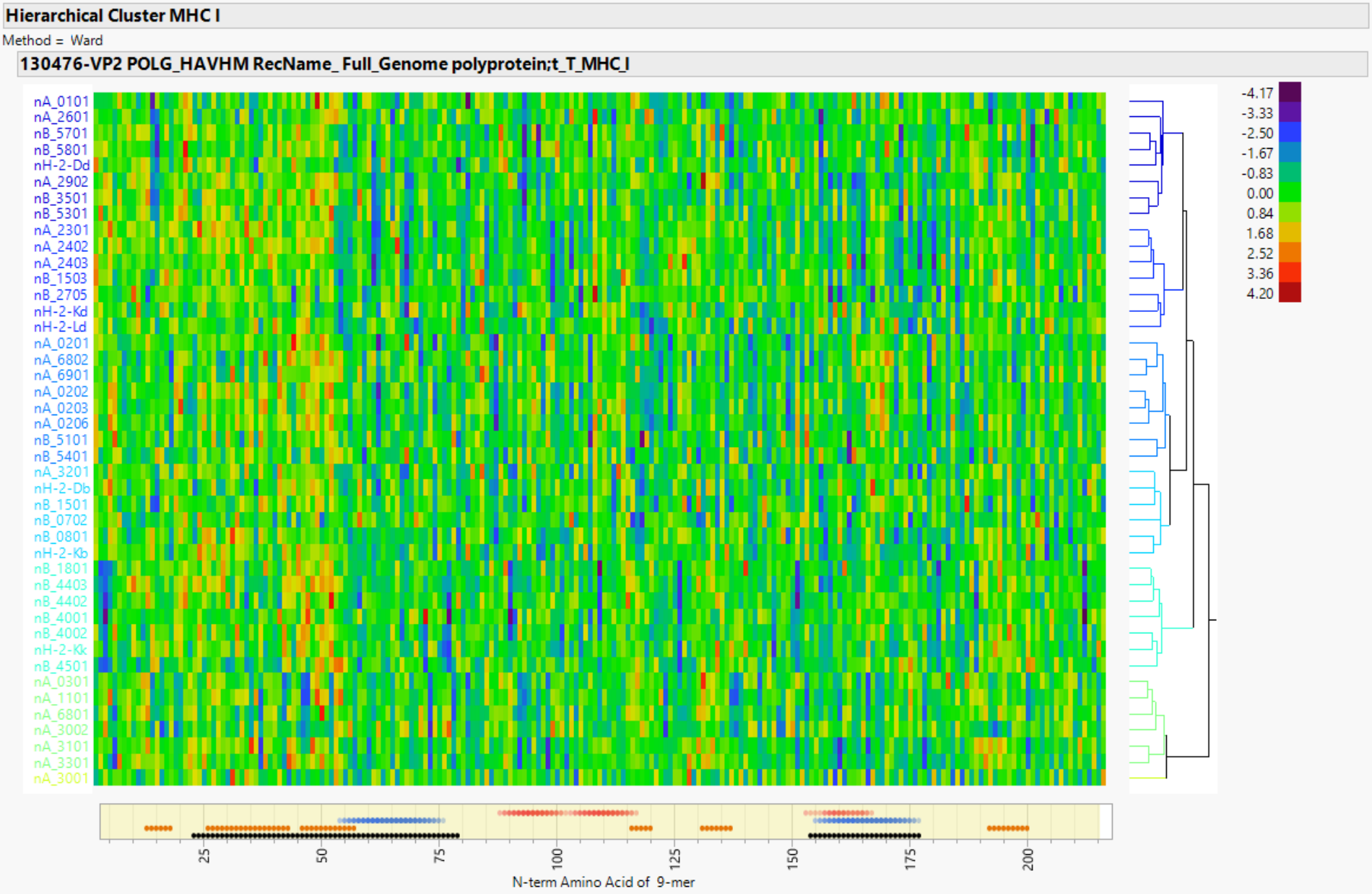
Protein VP2

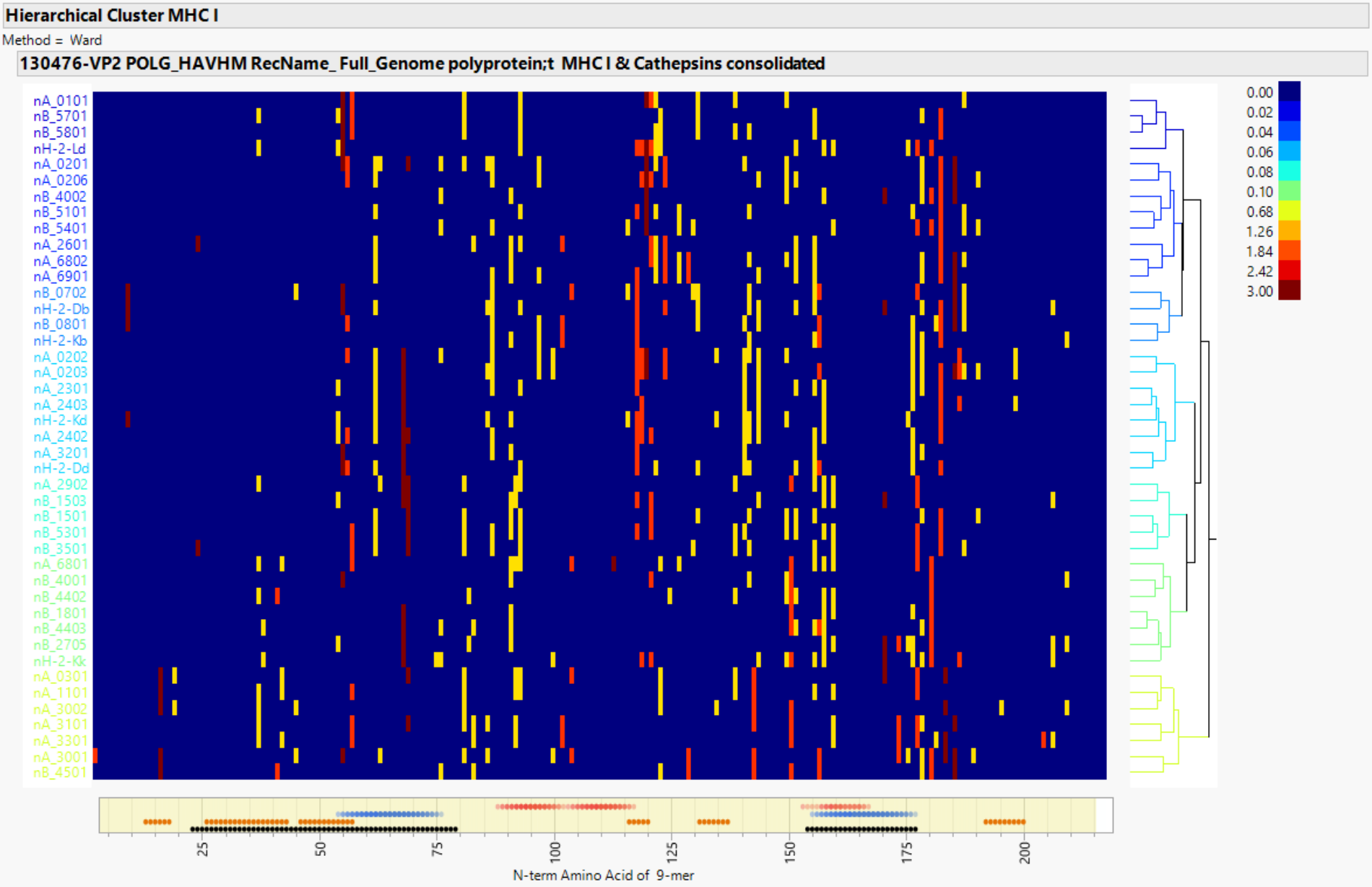

Population Permuted Metrics
130476-VP2 POLG_HAVHM RecName_
Full_Genome polyprotein;t
BEPI
MHC II
MHC I
25 %-tile
10 %-tile
10 %-tile
Mapped
Outside
Membrane
Inside
Signal
Std Dev from Mean
-1.00
-1.25
-1.50
-1.75
-2.00
-2.25
-2.50
0
25
50
75
100
125
150
175
200
Peptide N-terminus

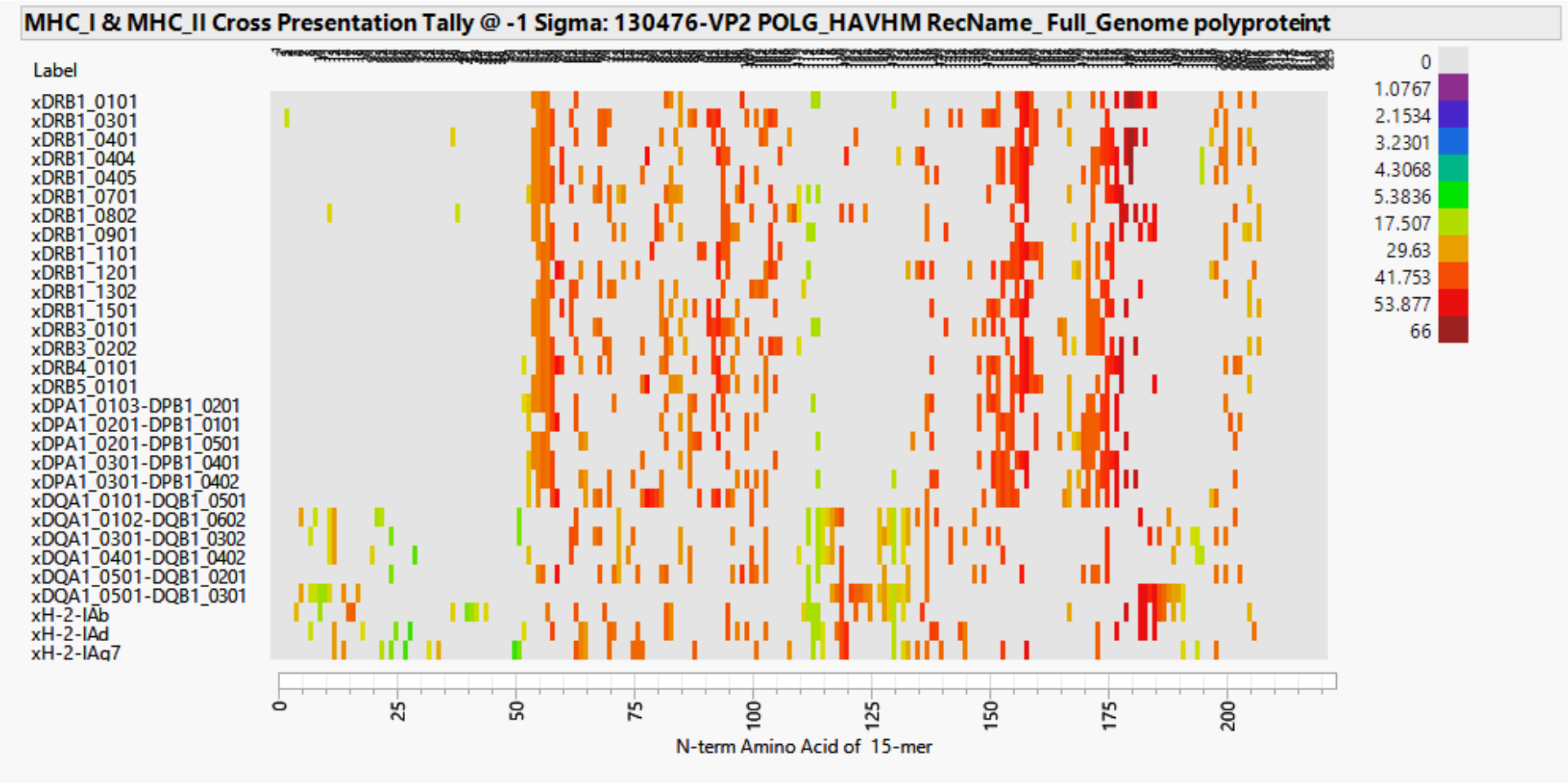

Hierarchical Cluster MHC II
Method = Ward
130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II
nDRB1_0101
nDRB1_0701
nDRB1_0901
nDRB1_0401
nDRB1_1302
nDRB1_0404
nDRB1_0405
nDRB1_1101
nDRB1_1501
nDRB5_0101
nDRB3_0202
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDRB1_1201
nDRB4_0101
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0301-DPB1_0401
nDPA1_0201-DPB1_0101
nDPA1_0201-DPB1_0501
nDPA1_0301-DPB1_0402
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.53
-2.82
-2.12
-1.41
-0.71
0.00
0.66
1.31
1.97
2.63
3.29
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

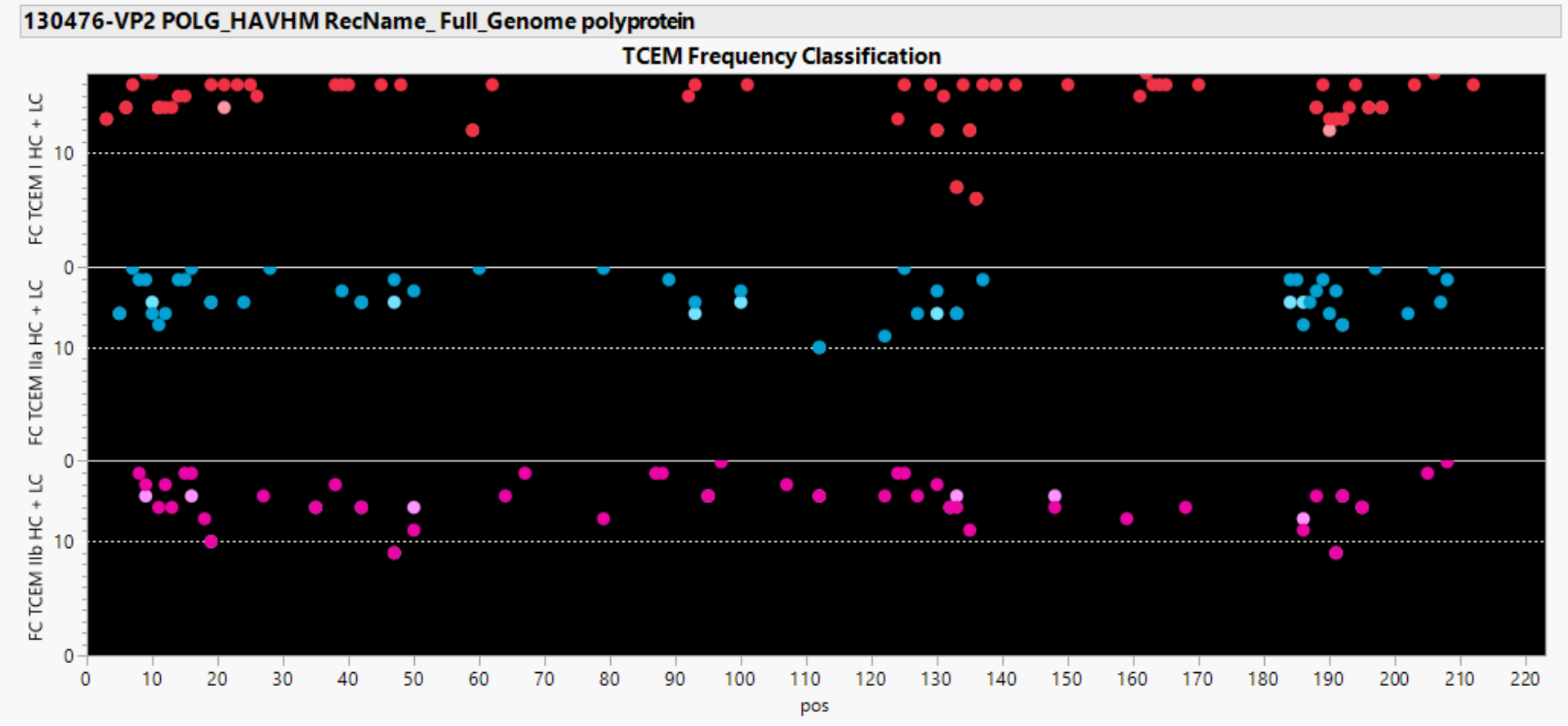

130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.04
0.08
0.12
0.16
0.21
0.96
1.72
2.48
3.24
4.00

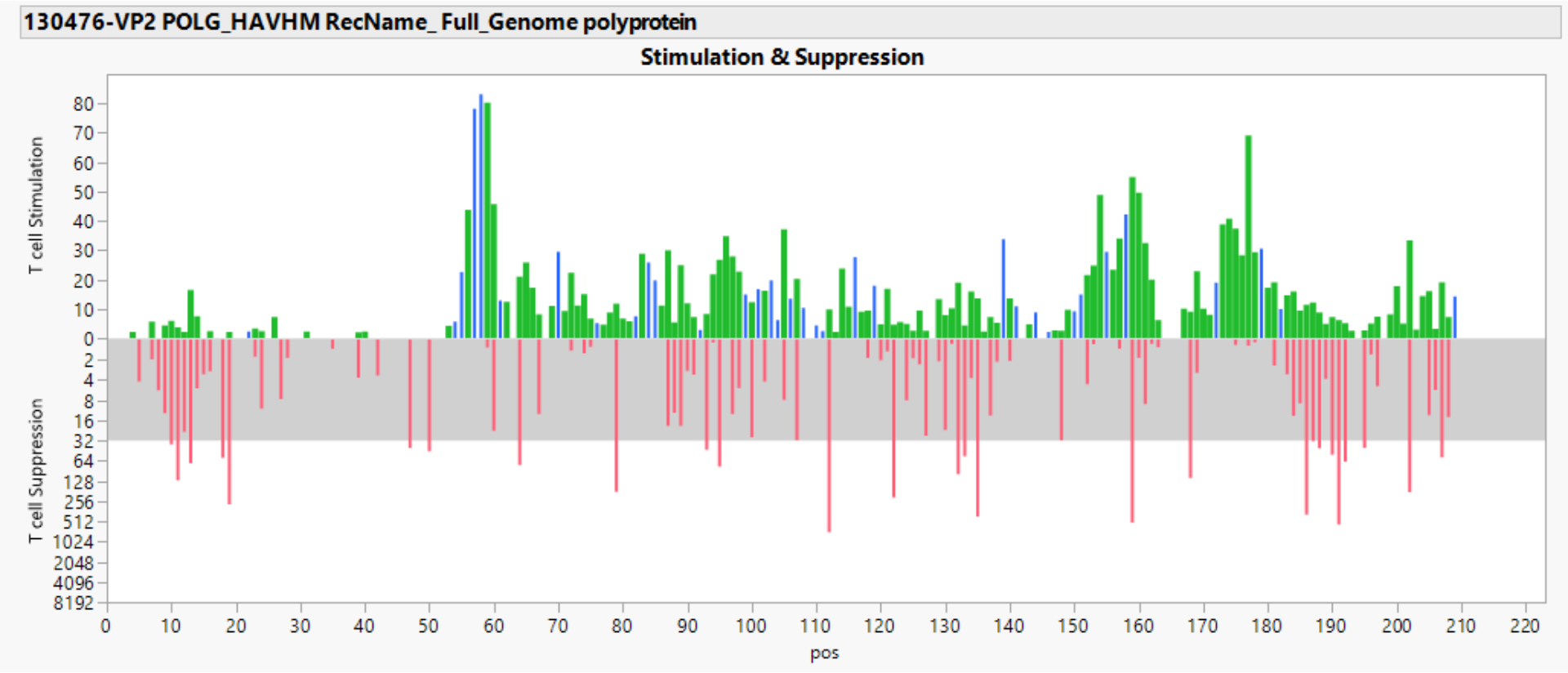

130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.06
0.10
0.13
0.16
1.13
2.10
3.06
4.03
5.00


Hierarchical Cluster MHC I
Method = Ward
130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_2601
nB_5701
nB_5801
nH-2-Dd
nA_2902
nB_3501
nB_5301
nA_2301
nA_2402
nA_2403
nB_1503
nB_2705
nH-2-Kd
nH-2-Ld
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nB_5101
nB_5401
nA_3201
nH-2-Db
nB_1501
nB_0702
nB_0801
nH-2-Kb
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_6801
nA_3002
nA_3101
nA_3301
nA_3001
-4.17
-3.33
-2.50
-1.67
-0.83
0.00
0.84
1.68
2.52
3.36
4.20
25
50
75
100
125
150
175
200
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nB_5701
nB_5801
nH-2-Ld
nA_0201
nA_0206
nB_4002
nB_5101
nB_5401
nA_2601
nA_6802
nA_6901
nB_0702
nH-2-Db
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_2301
nA_2403
nH-2-Kd
nA_2402
nA_3201
nH-2-Dd
nA_2902
nB_1503
nB_1501
nB_5301
nB_3501
nA_6801
nB_4001
nB_4402
nB_1801
nB_4403
nB_2705
nH-2-Kk
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_3001
nB_4501
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.0767
2.1534
3.2301
4.3068
5.3836
17.507
29.63
41.753
53.877
66
0
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

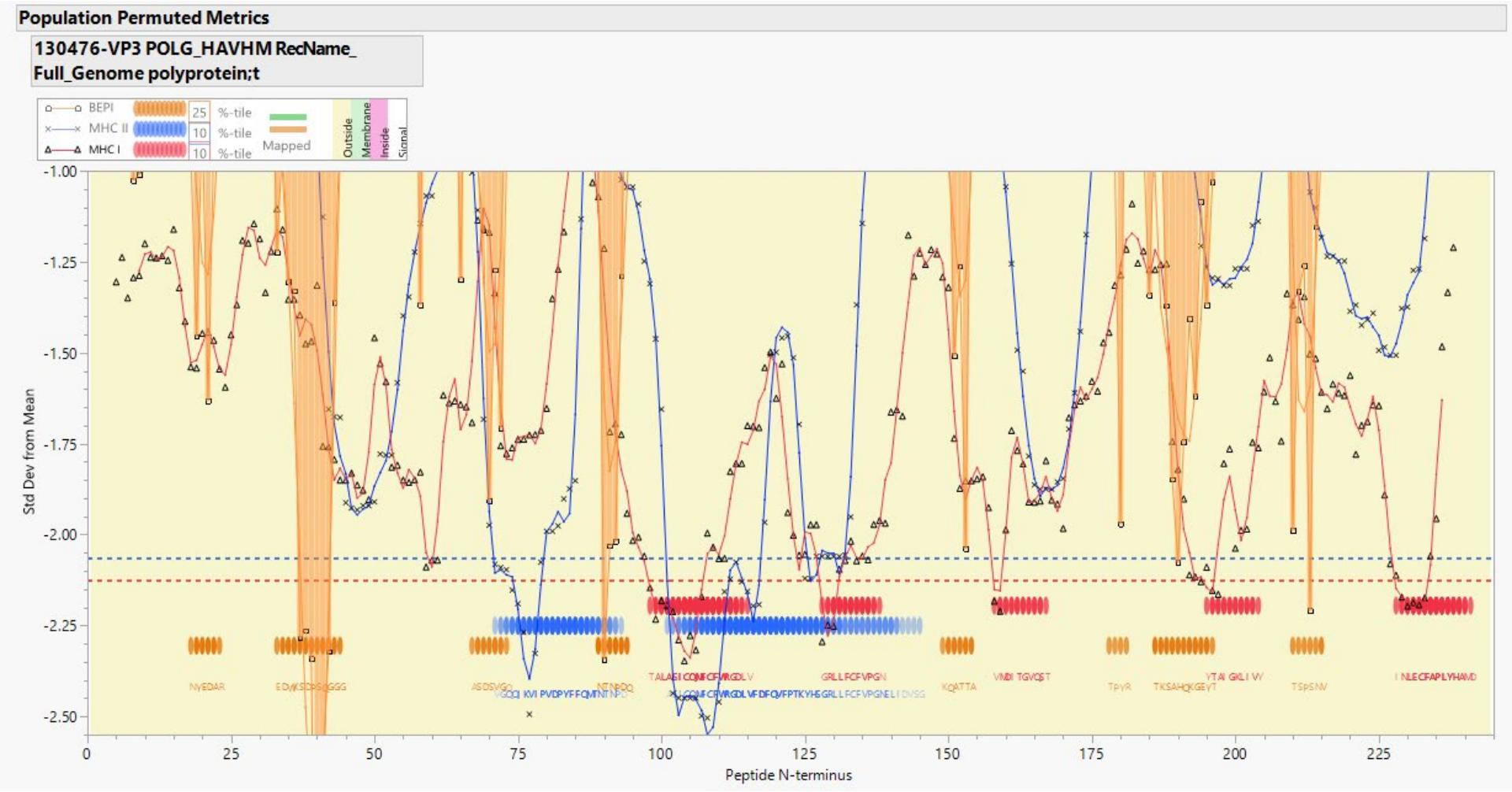

130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220
pos


130476-VP2 POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

Protein VP3

Population Permuted Metrics

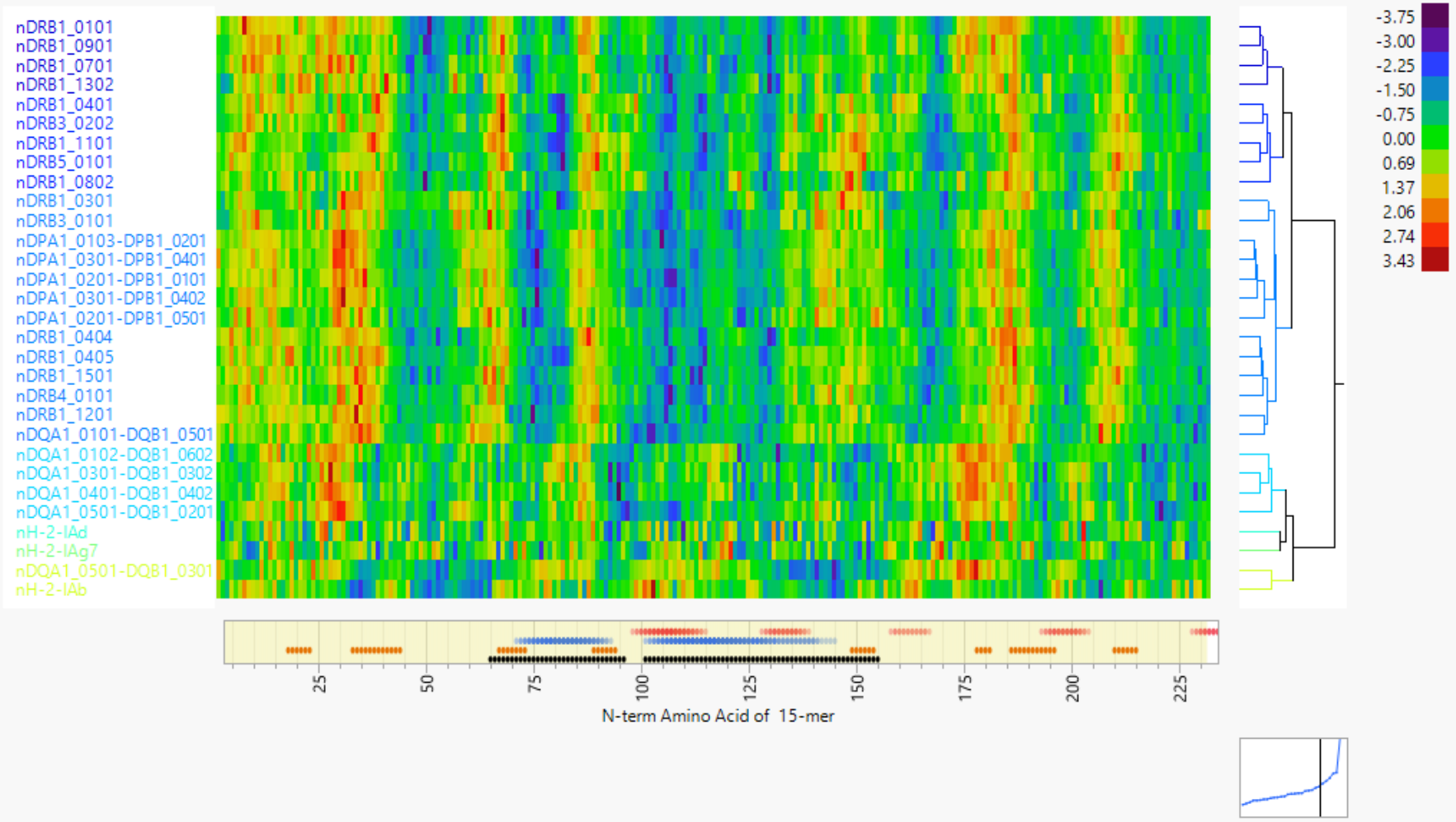
130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t



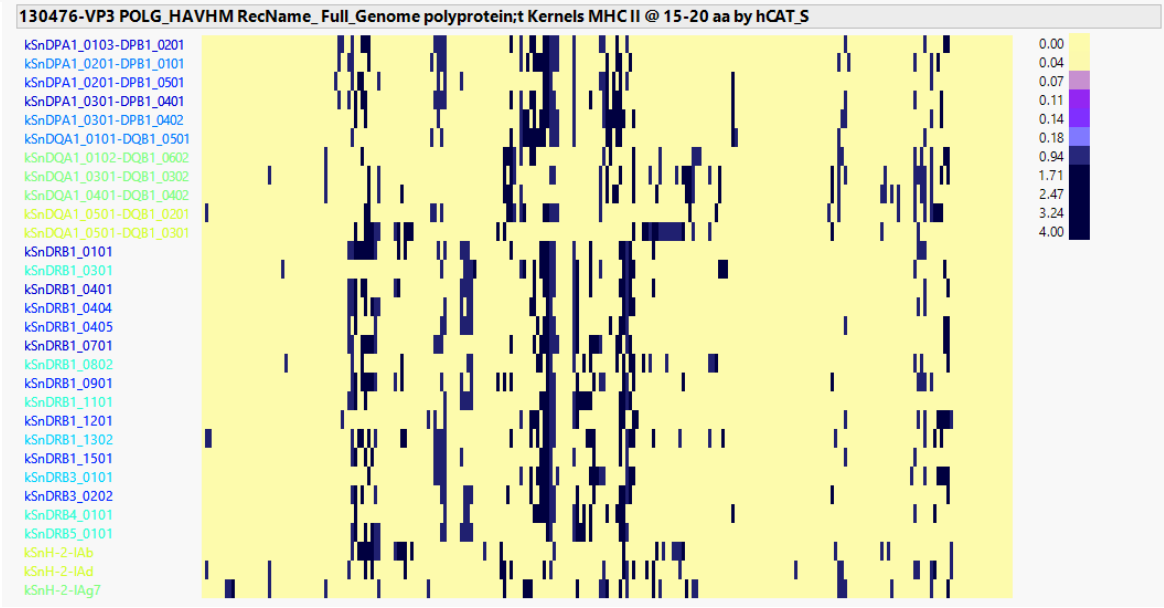
Hierarchical Cluster MHC II

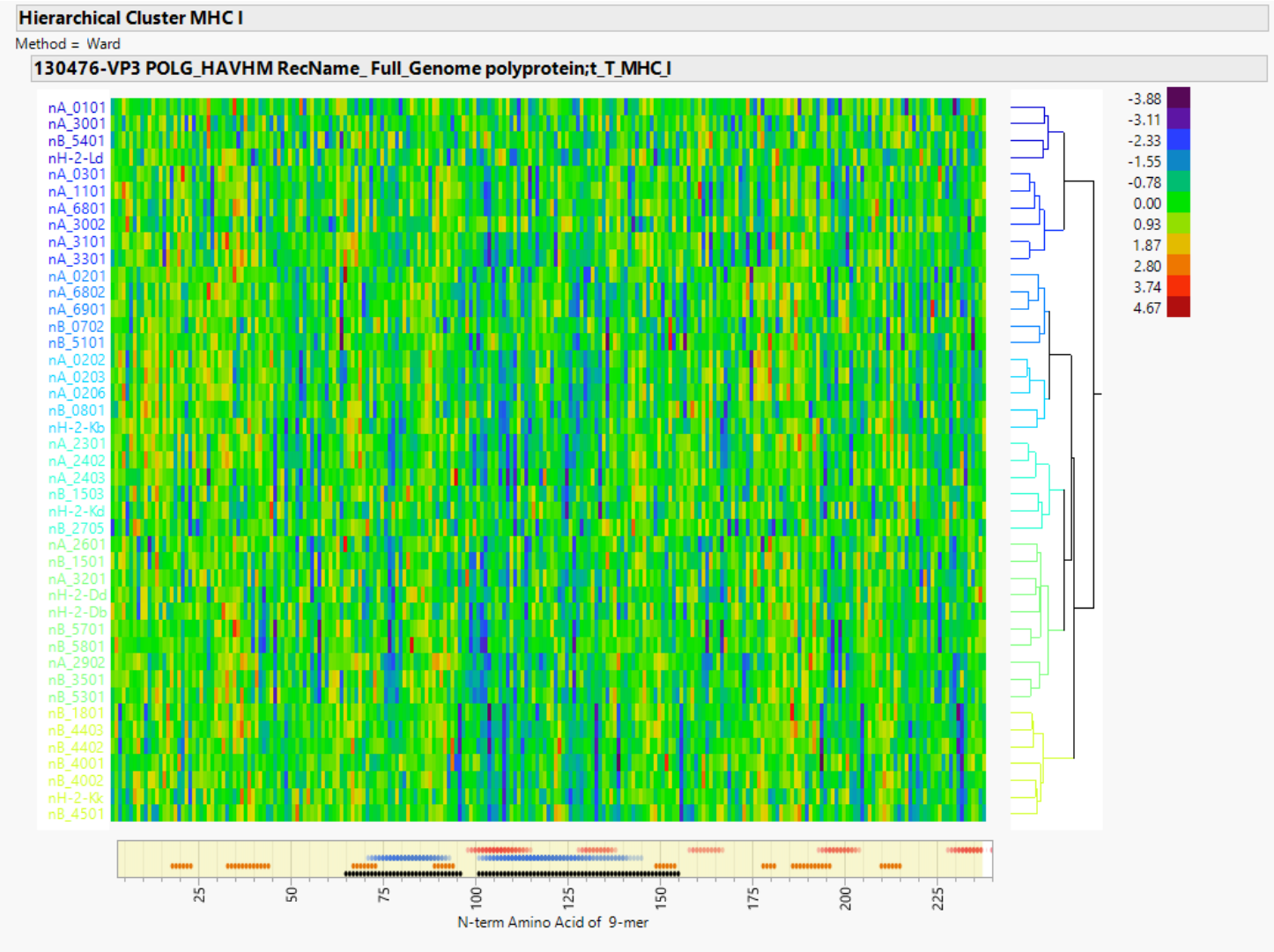
Method = Ward

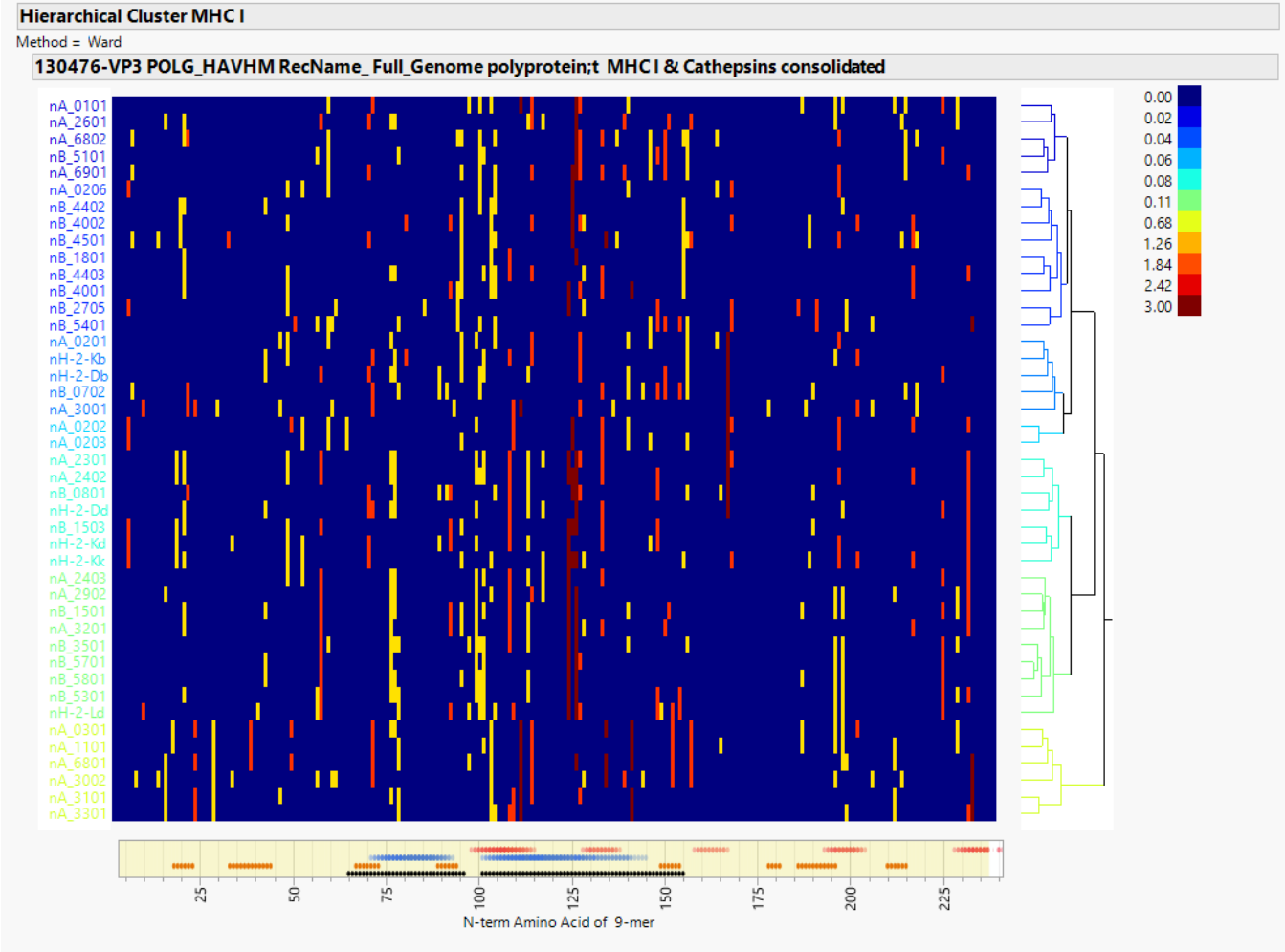
130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_II

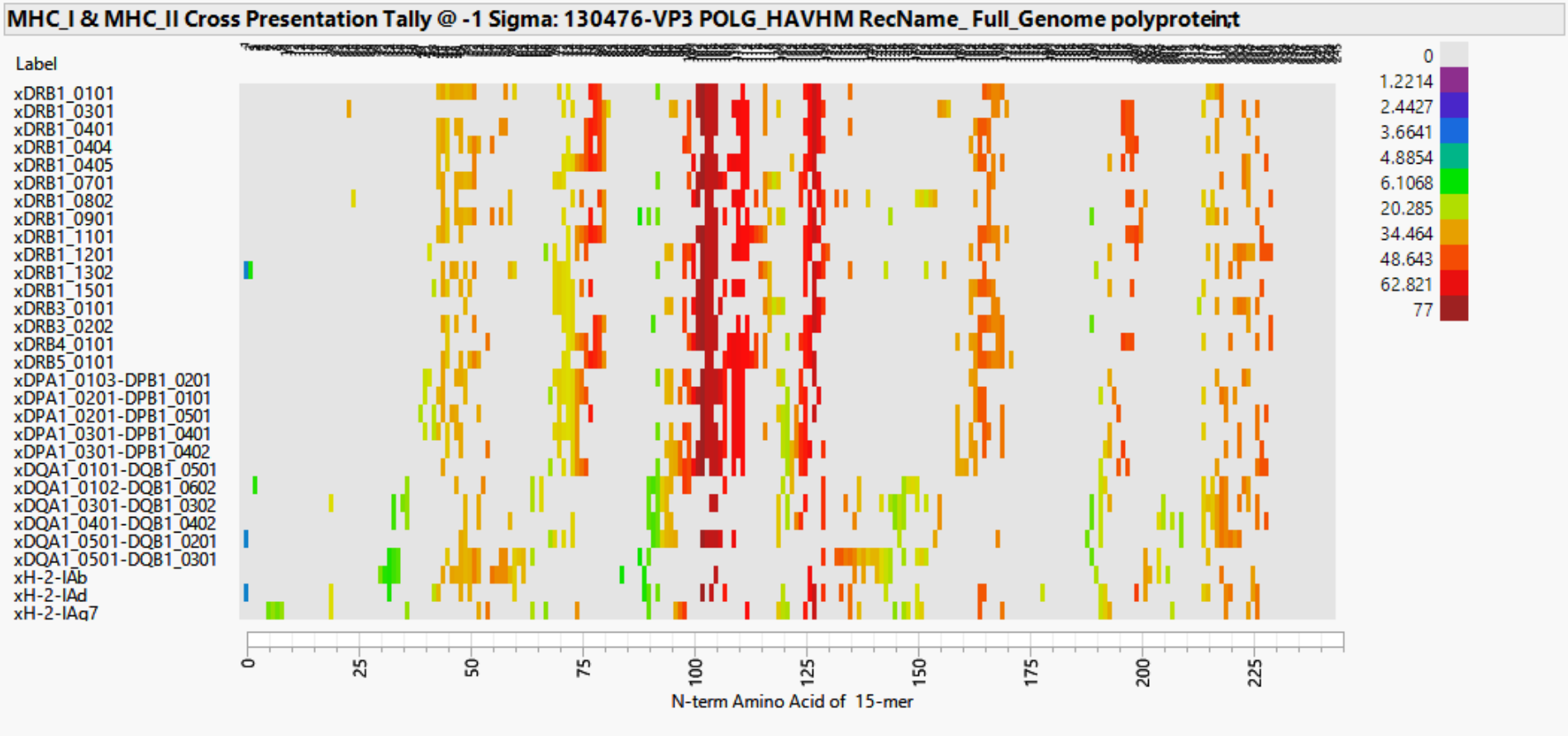


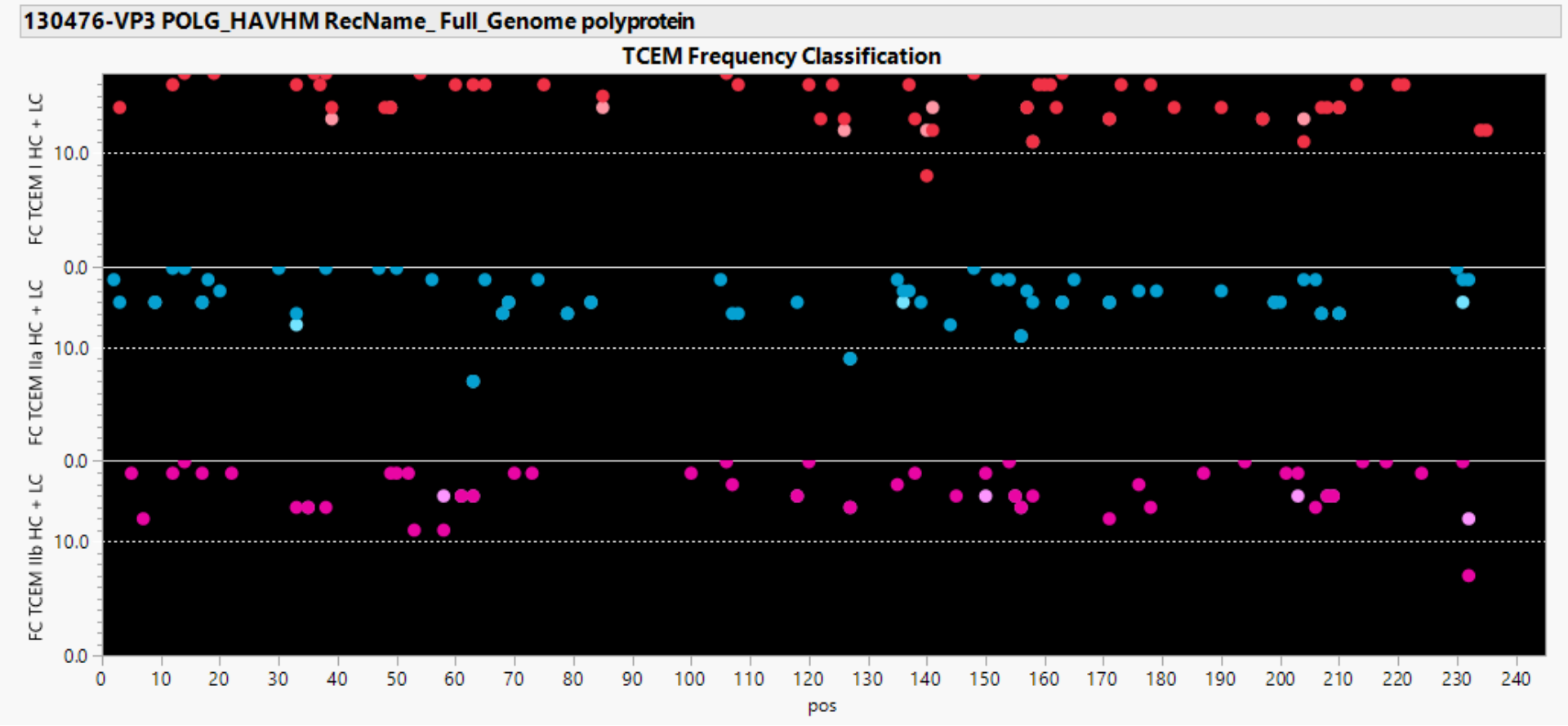

130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.05
0.10
0.15
0.21
0.26
1.21
2.15
3.10
4.05
5.00

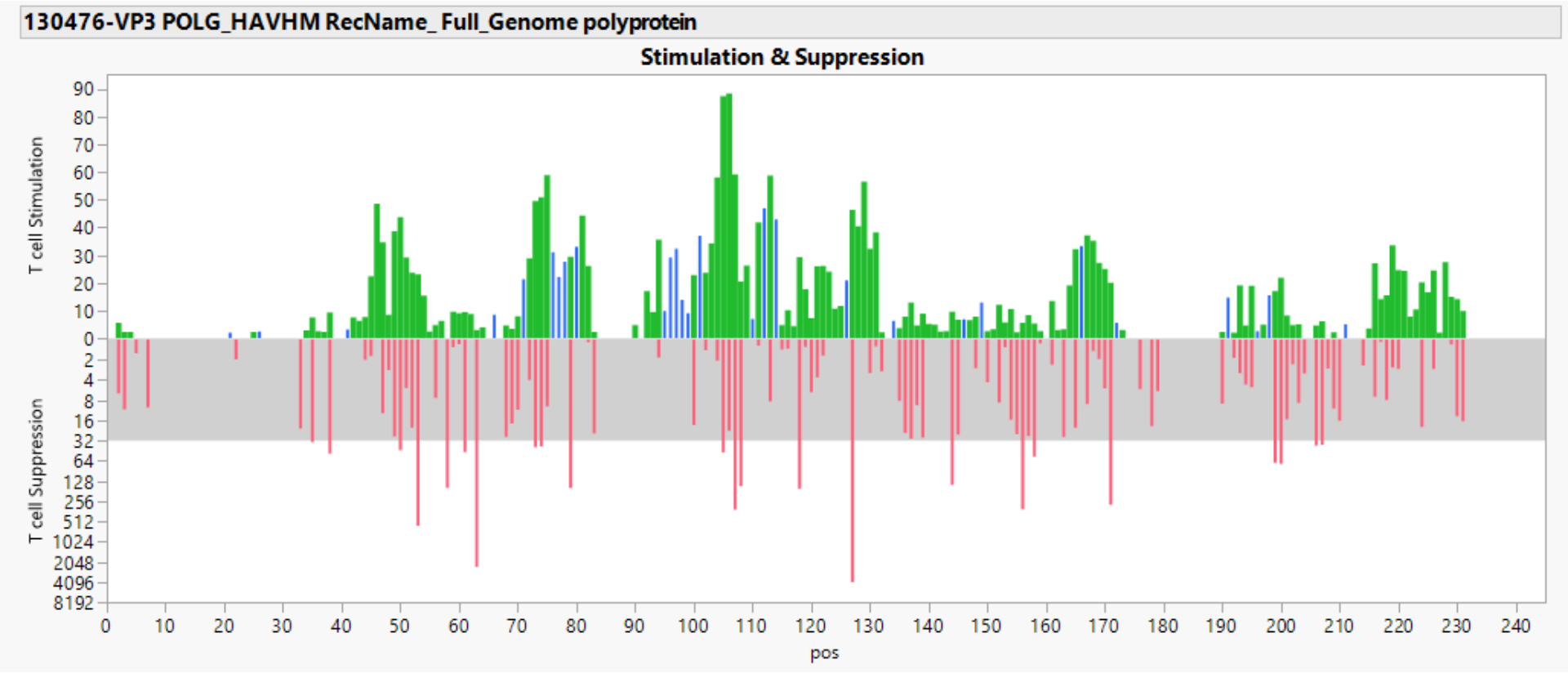

130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.04
0.07
0.11
0.14
0.18
0.94
1.71
2.47
3.24
4.00


Hierarchical Cluster MHC I
Method = Ward
130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_0301
nA_1101
nA_6801
nA_3002
nA_3101
nA_3301
nA_0201
nA_6802
nA_6901
nB_0702
nB_5101
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nB_1503
nH-2-Kd
nB_2705
nA_2601
nB_1501
nA_3201
nH-2-Dd
nH-2-Db
nB_5701
nB_5801
nA_2902
nB_3501
nB_5301
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
-3.88
-3.11
-2.33
-1.55
-0.78
0.00
0.93
1.87
2.80
3.74
4.67
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t MHC I & Cathepsins consolidated
nA_0101
nA_2601
nA_6802
nB_5101
nA_6901
nA_0206
nB_4402
nB_4002
nB_4501
nB_1801
nB_4403
nB_4001
nB_2705
nB_5401
nA_0201
nH-2-Kb
nH-2-Db
nB_0702
nA_3001
nA_0202
nA_0203
nA_2301
nA_2402
nB_0801
nH-2-Dd
nB_1503
nH-2-Kd
nH-2-Kk
nA_2403
nA_2902
nB_1501
nA_3201
nB_3501
nB_5701
nB_5801
nB_5301
nH-2-Ld
nA_0301
nA_1101
nA_6801
nA_3002
nA_3101
nA_3301
0.00
0.02
0.04
0.06
0.08
0.11
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.2214
2.4427
3.6641
4.8854
6.1068
20.285
34.464
48.643
62.821
77
0
25
50
75
100
125
150
175
200
225
N-term Amino Acid of 15-mer


130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10.0
0.0
10.0
0.0
10.0
0.0
0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210 220 230 240
pos


130476-VP3 POLG_HAVHM RecName_ Full_Genome polyprotein
Stimulation & Suppression
T cell Stimulation
T cell Suppression
90
80
70
60
50
40
30
20
10
0
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
pos

ioGenetics®

Addendum 9
Rabies Virus

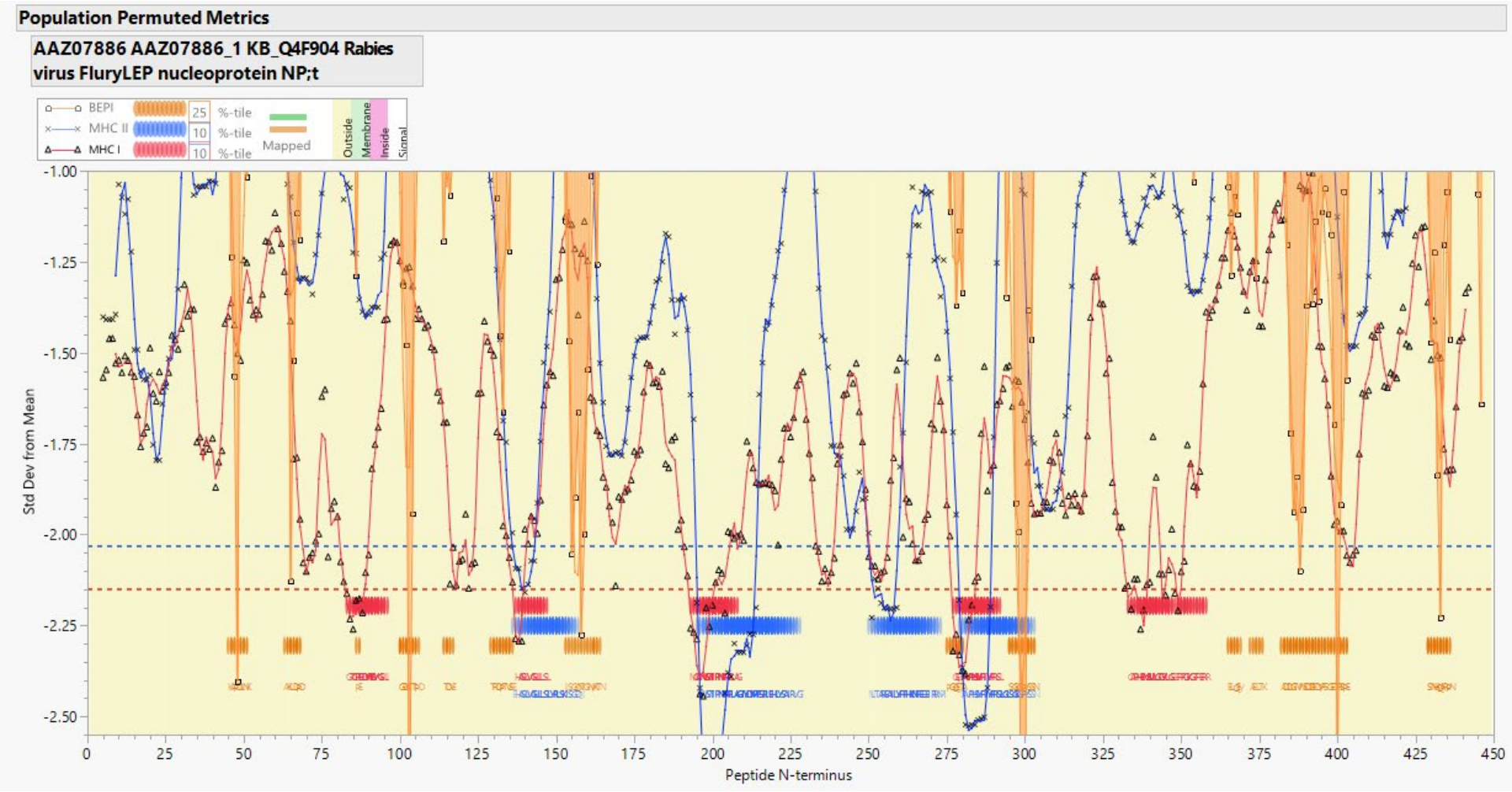
Flury LEP strain

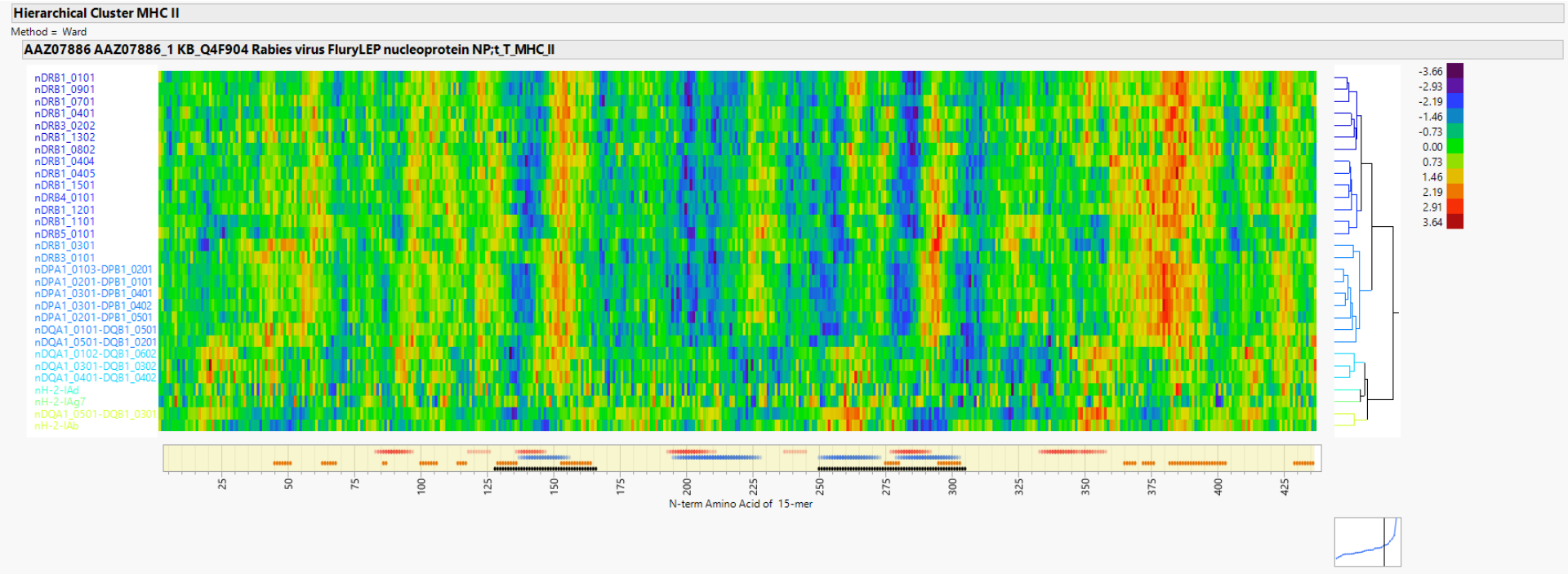
AAZ07886

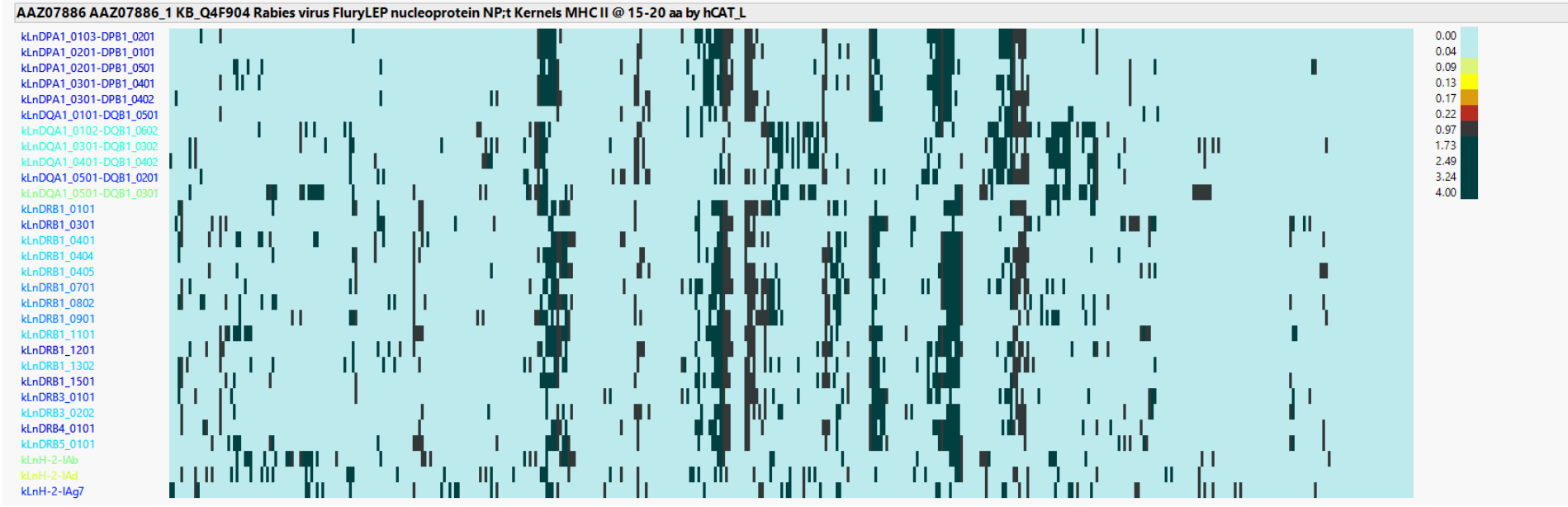
Nucleoprotein

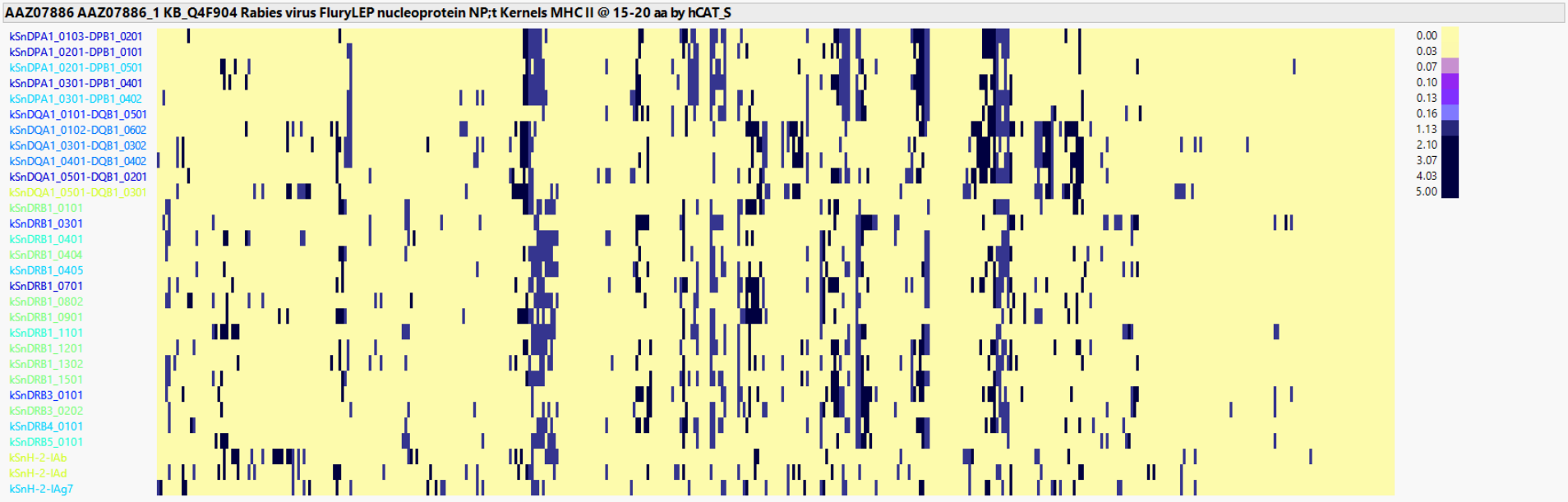
Population Permuted Metrics

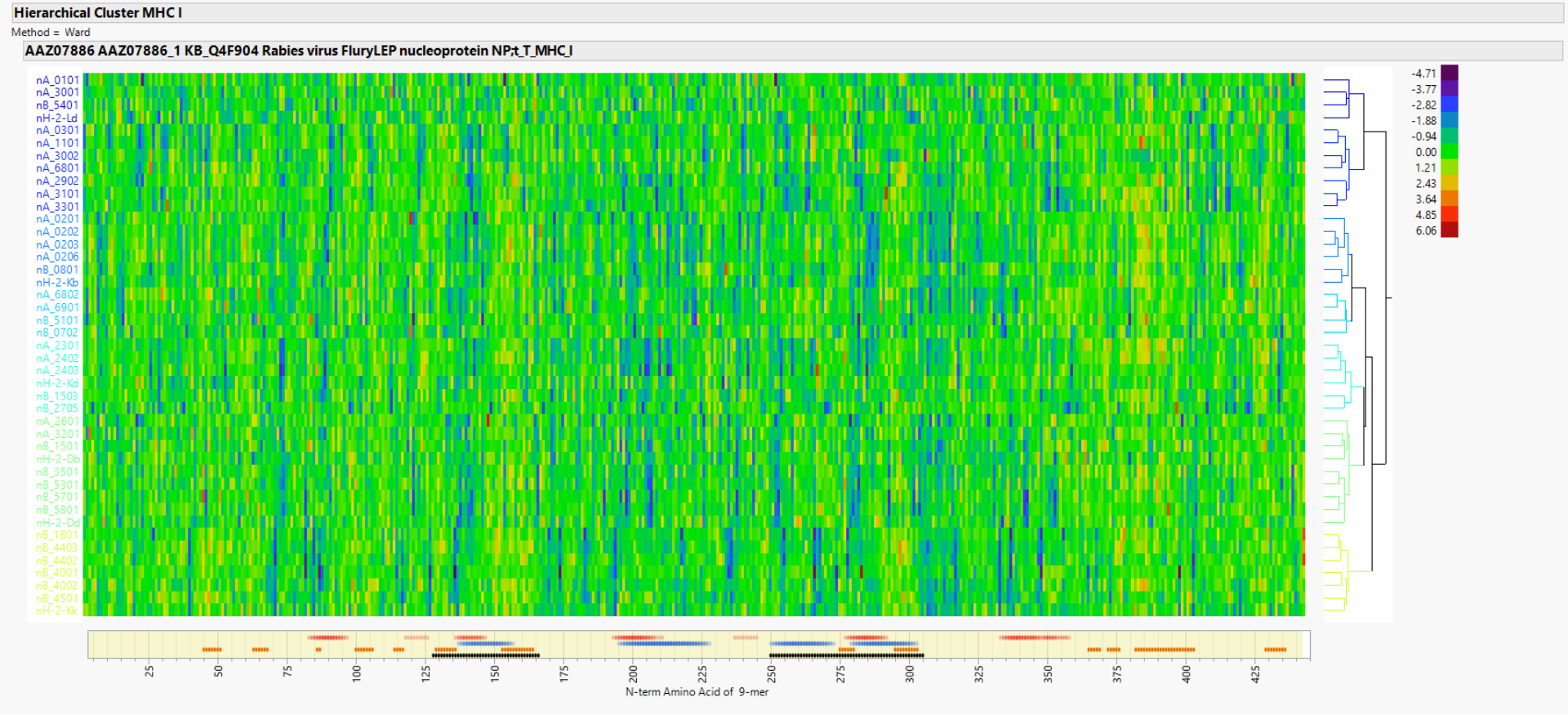
AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t

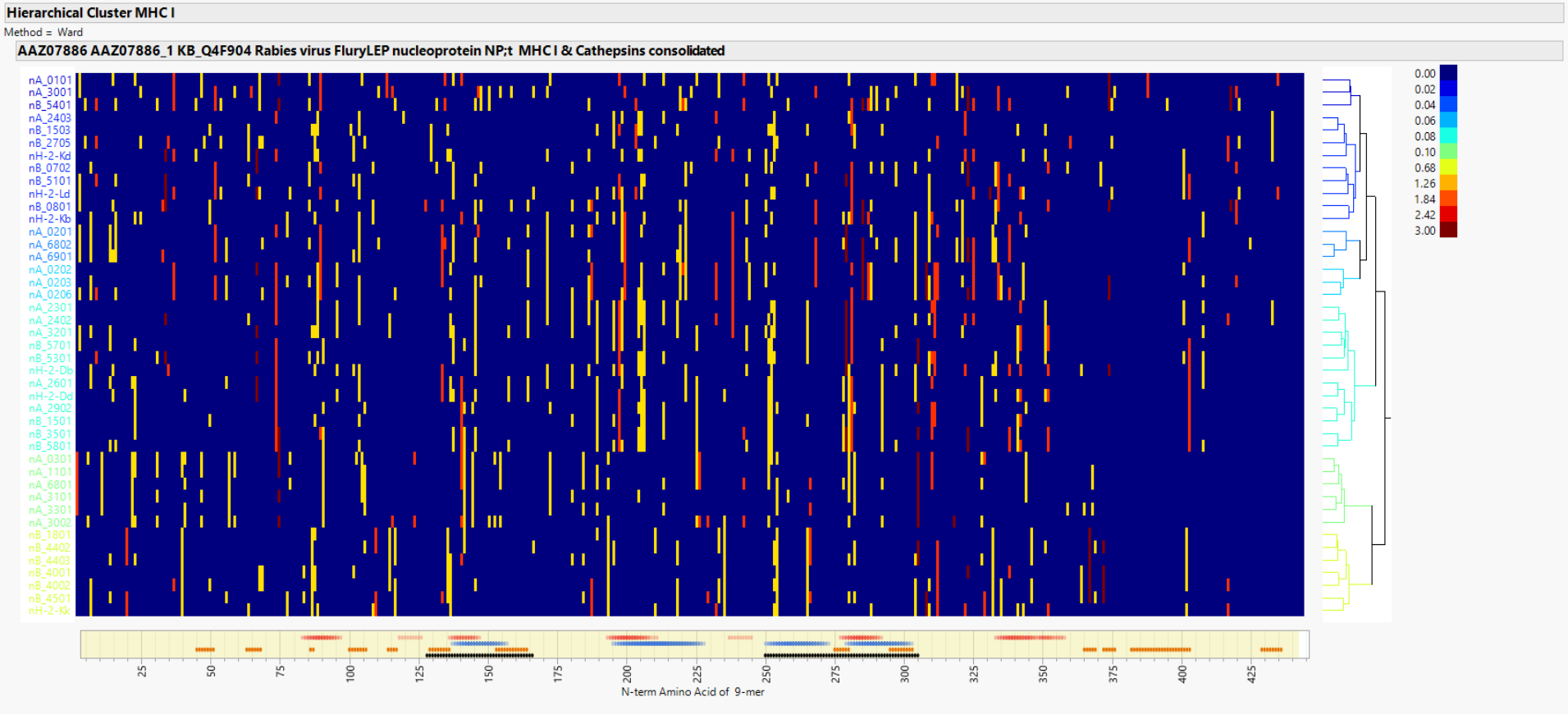


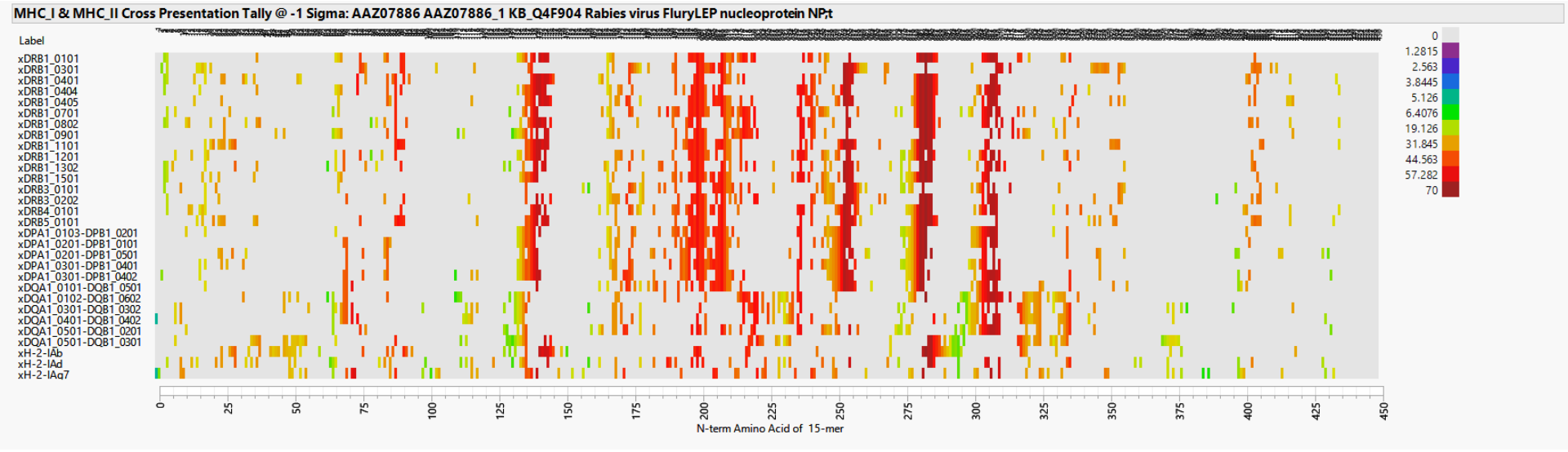

Hierarchical Cluster MHC II
Method = Ward
AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0401
nDRB3_0202
nDRB1_1302
nDRB1_0802
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.66
-2.93
-2.19
-1.46
-0.73
0.00
0.73
1.46
2.19
2.91
3.64
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
N-term Amino Acid of 15-mer

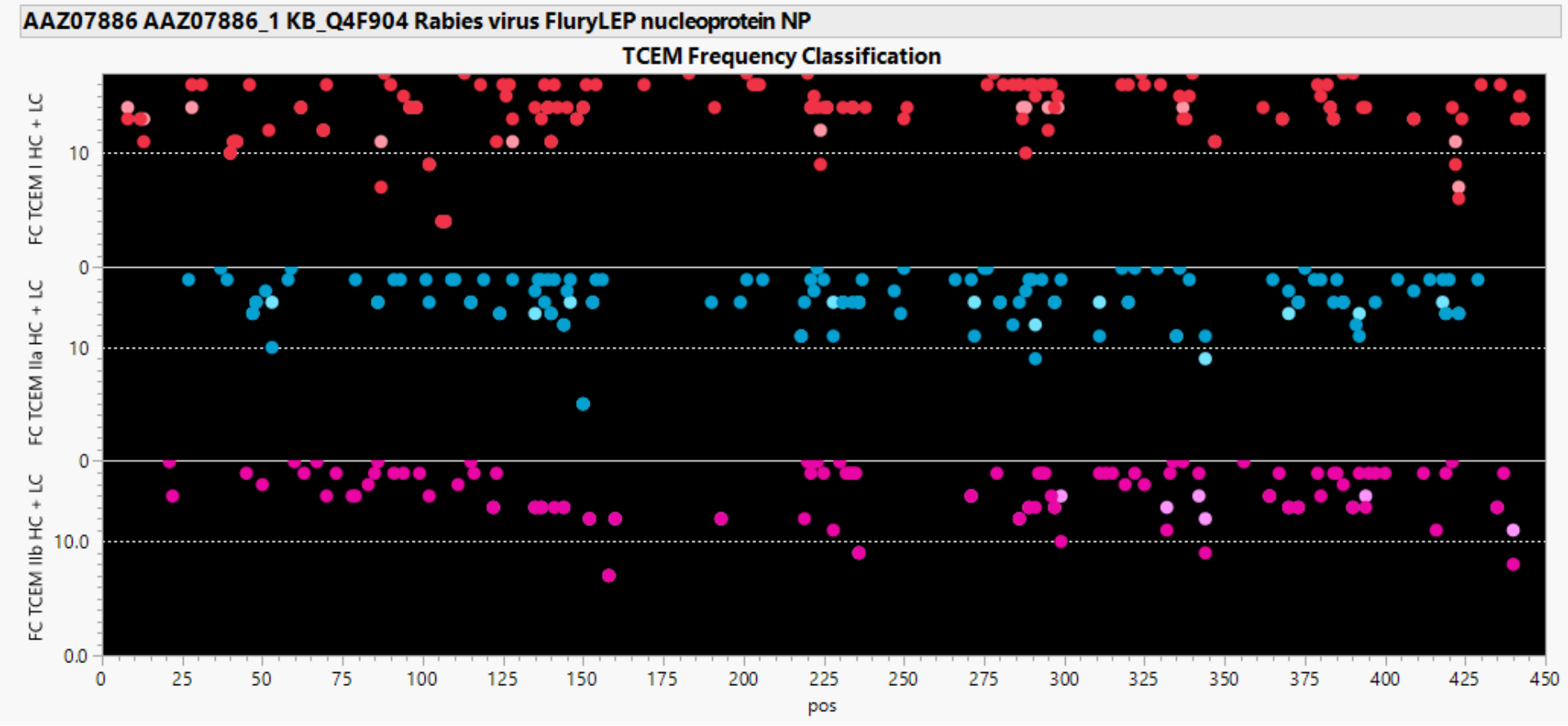

AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.04
0.09
0.13
0.17
0.22
0.97
1.73
2.49
3.24
4.00

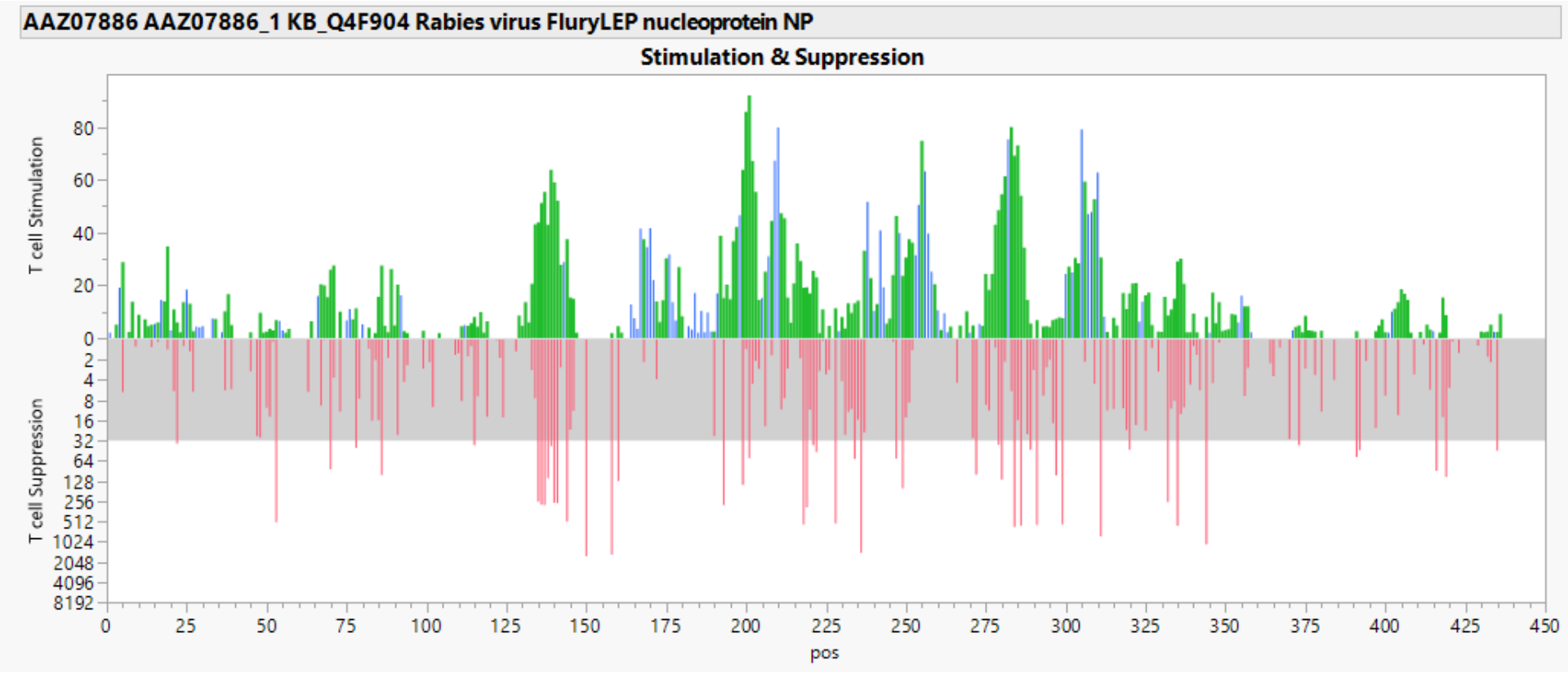

AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.07
0.10
0.13
0.16
1.13
2.10
3.07
4.03
5.00


Hierarchical Cluster MHC I
Method = Ward
AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
nA_0201
nA_0202
nA_0203
nA_0206
nB_0801
nH-2-Kb
nA_6802
nA_6901
nB_5101
nB_0702
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nA_2601
nA_3201
nB_1501
nH-2-Db
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nB_4501
nH-2-Kk
-4.71
-3.77
-2.82
-1.88
-0.94
0.00
1.21
2.43
3.64
4.85
6.06
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t MHC I & Cathepsins consolidated
nA_0101
nA_3001
nB_5401
nA_2403
nB_1503
nB_2705
nH-2-Kd
nB_0702
nB_5101
nH-2-Ld
nB_0801
nH-2-Kb
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_2301
nA_2402
nA_3201
nB_5701
nB_5301
nH-2-Db
nA_2601
nH-2-Dd
nA_2902
nB_1501
nB_3501
nB_5801
nA_0301
nA_1101
nA_6801
nA_3101
nA_3301
nA_3002
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nH-2-Kk
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.2815
2.563
3.8445
5.126
6.4076
19.126
31.845
44.563
57.282
70
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
N-term Amino Acid of 15-mer

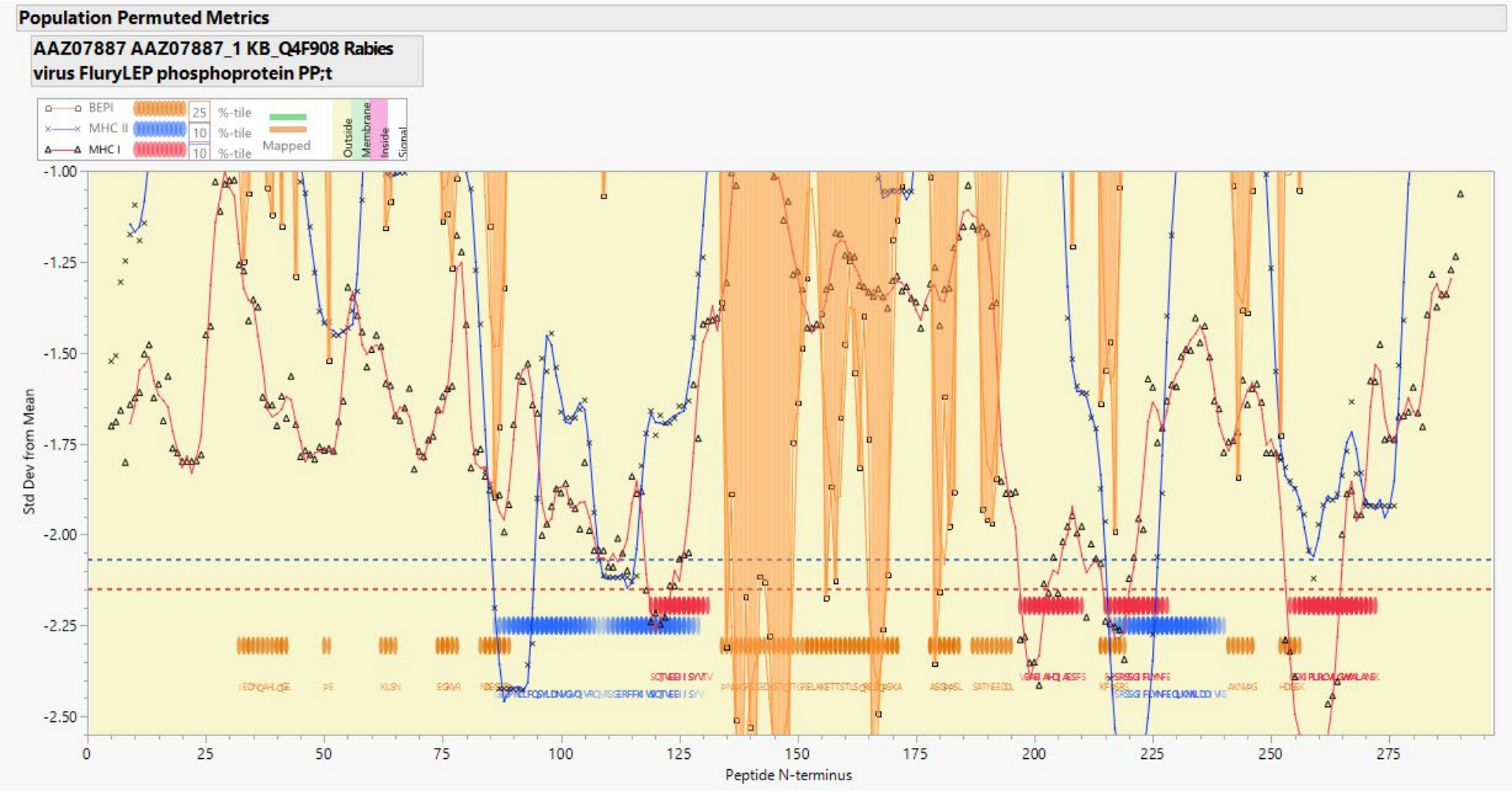

AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10.0
0.0
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
pos

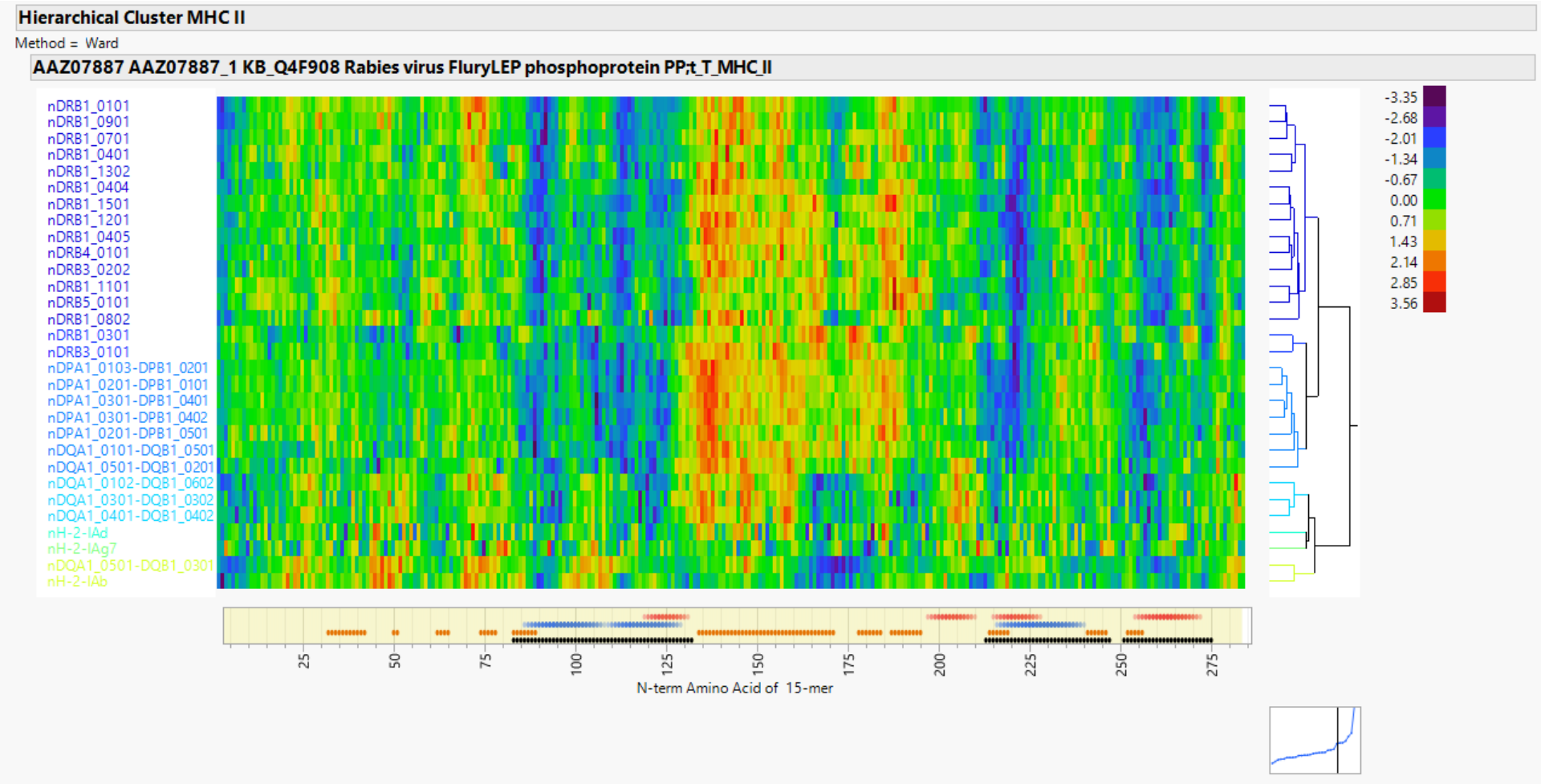

AAZ07886 AAZ07886_1 KB_Q4F904 Rabies virus FluryLEP nucleoprotein NP
Stimulation & Suppression
T cell Stimulation
80
60
40
20
0
T cell Suppression
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
pos

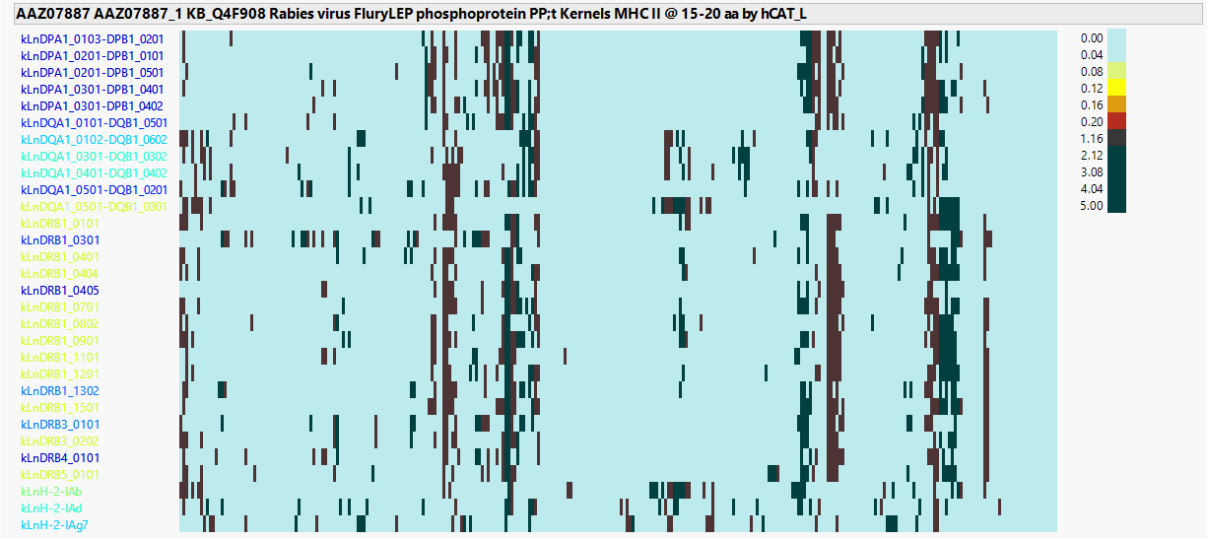
Phosphoprotein

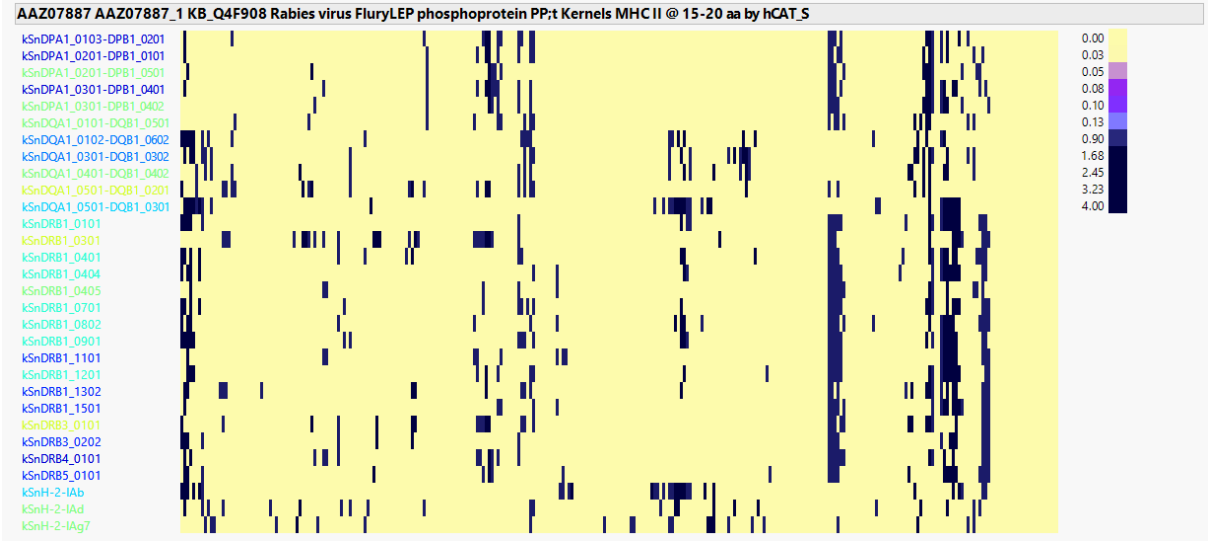
Population Permuted Metrics

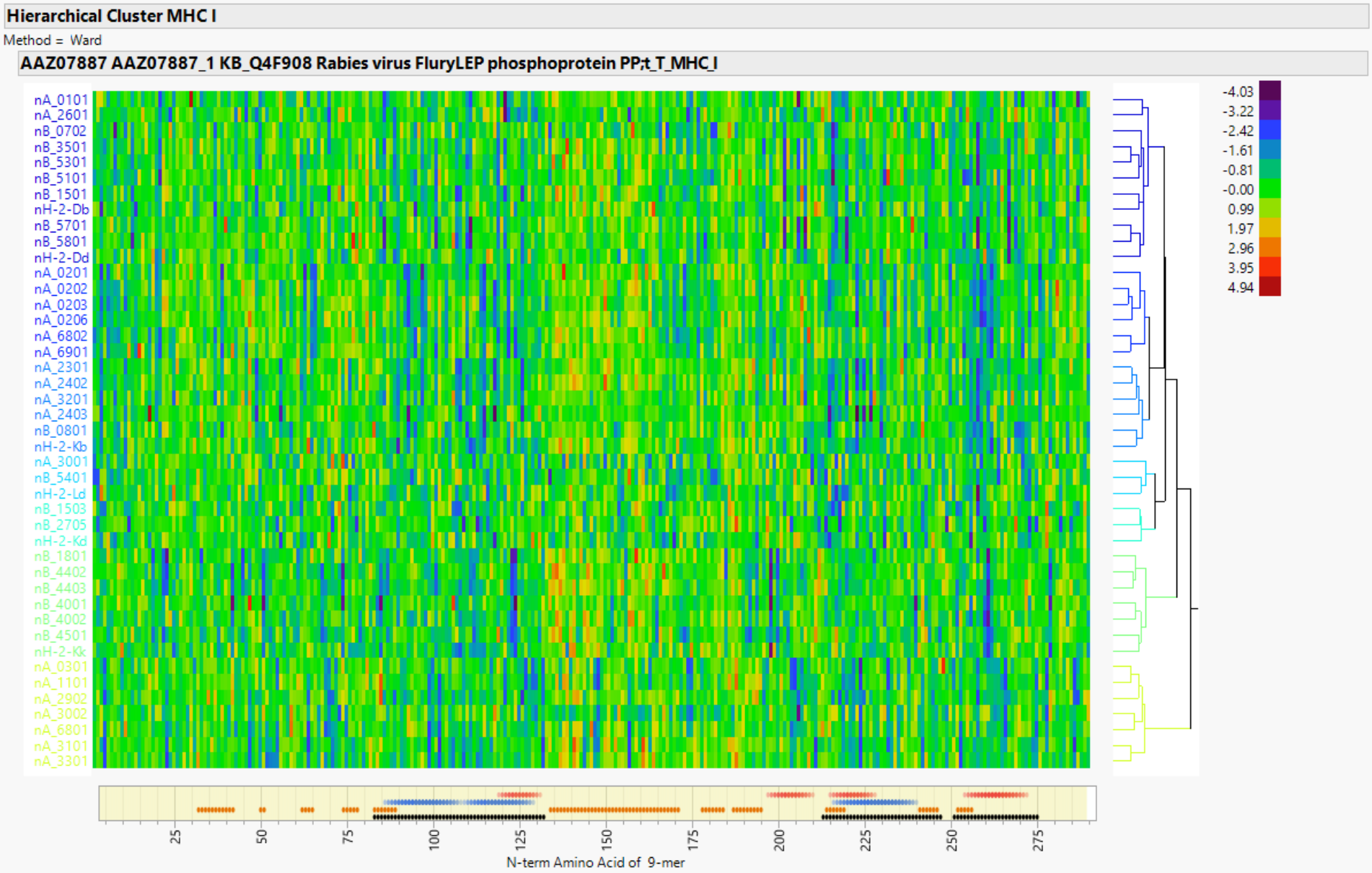
AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t

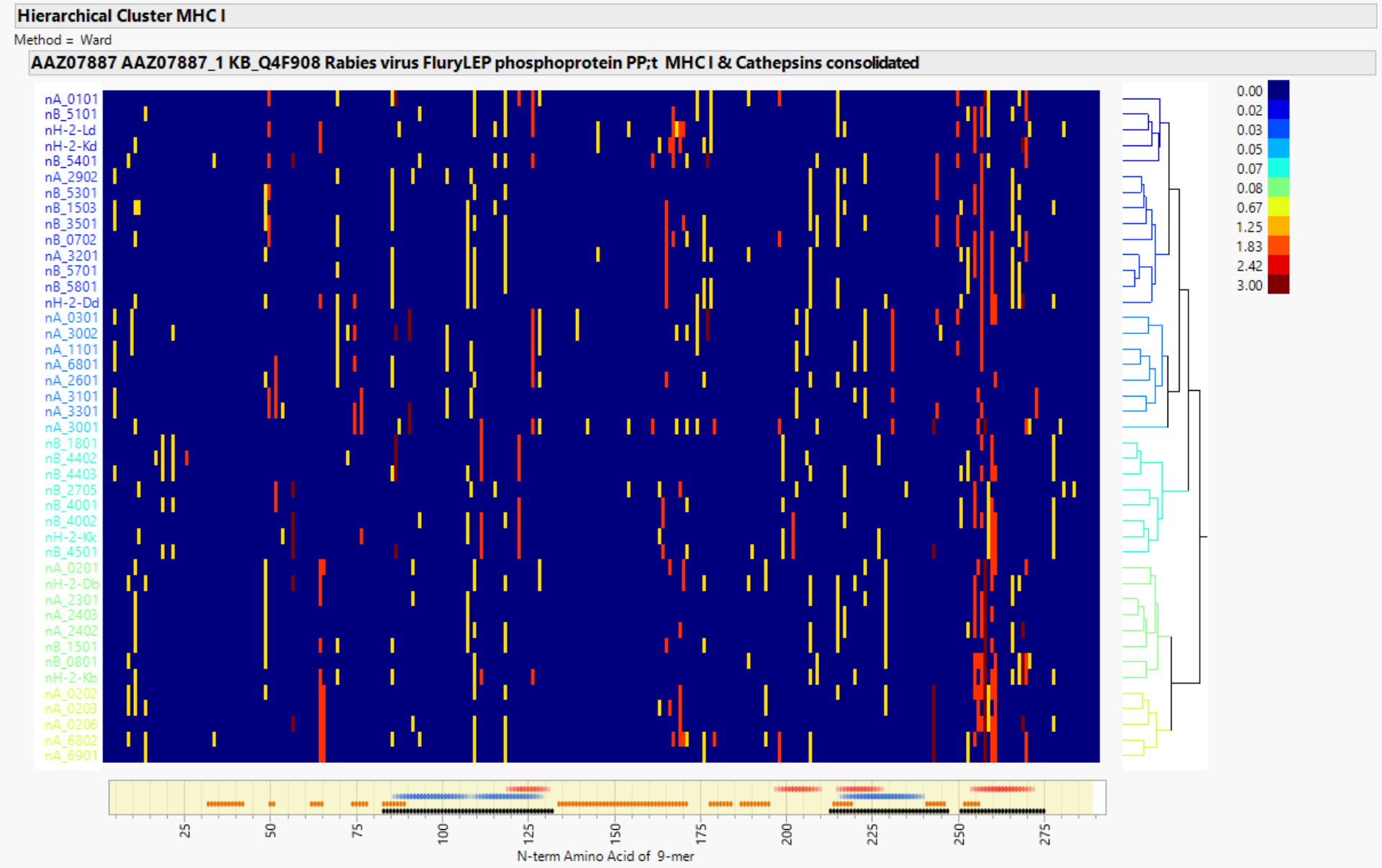


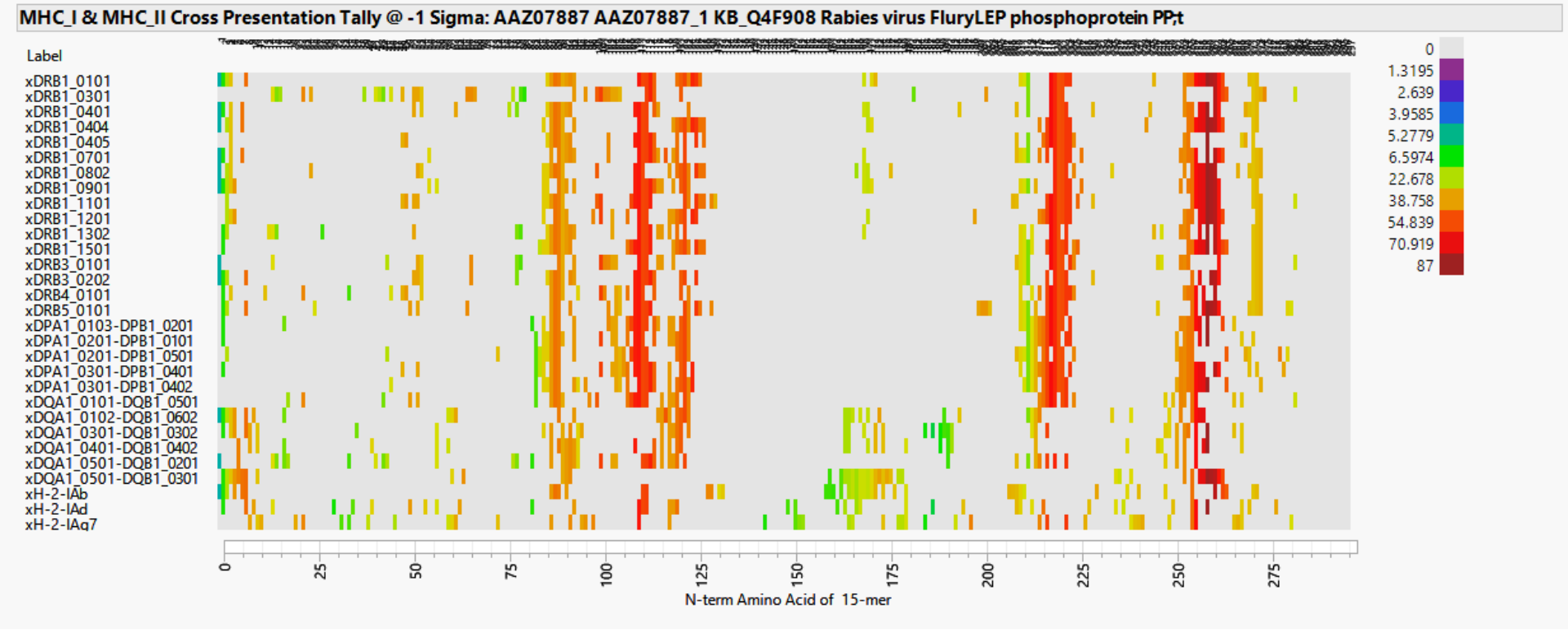

Hierarchical Cluster MHC II
Method = Ward
AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_0401
nDRB1_1302
nDRB1_0404
nDRB1_1501
nDRB1_1201
nDRB1_0405
nDRB4_0101
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.35
-2.68
-2.01
-1.34
-0.67
0.00
0.71
1.43
2.14
2.85
3.56
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer

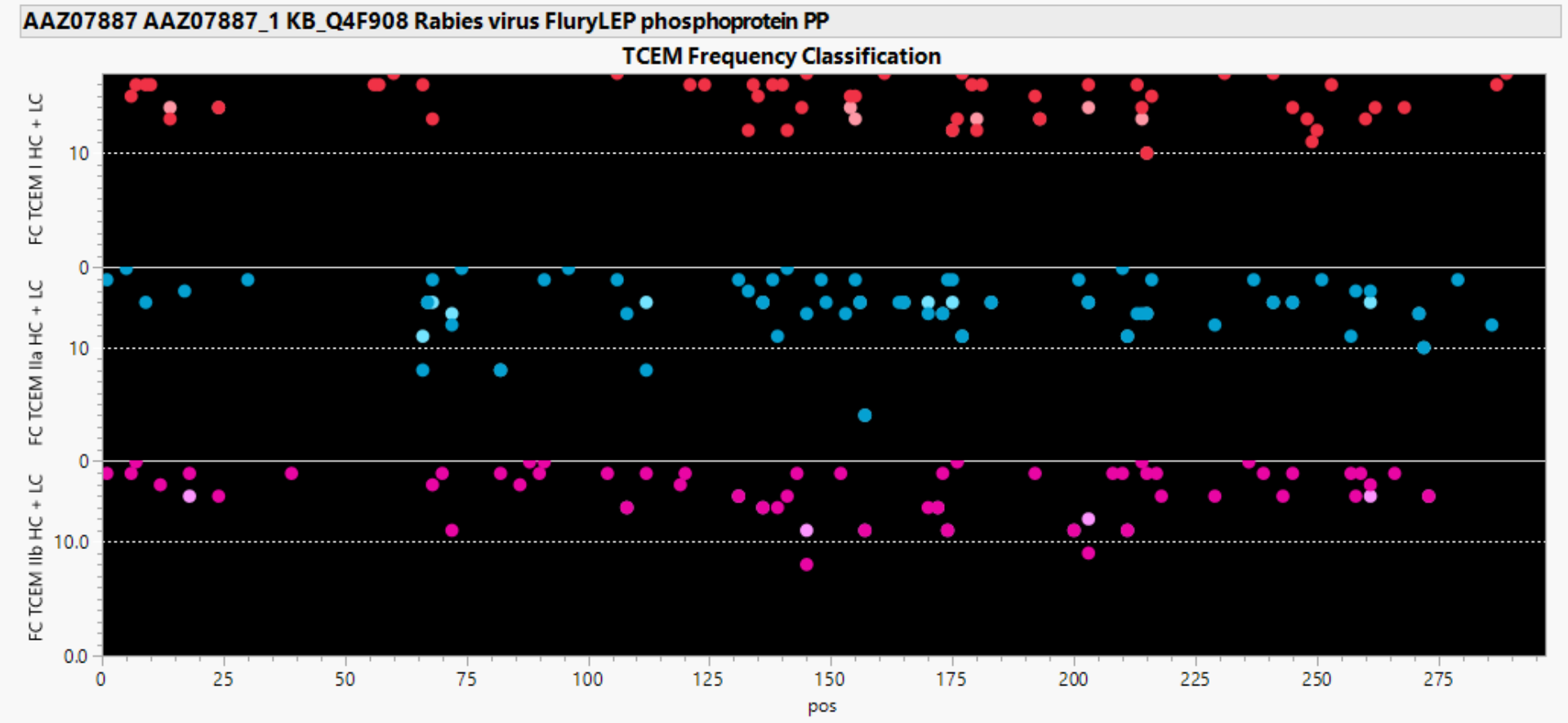

AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.04
0.08
0.12
0.16
0.20
1.16
2.12
3.08
4.04
5.00

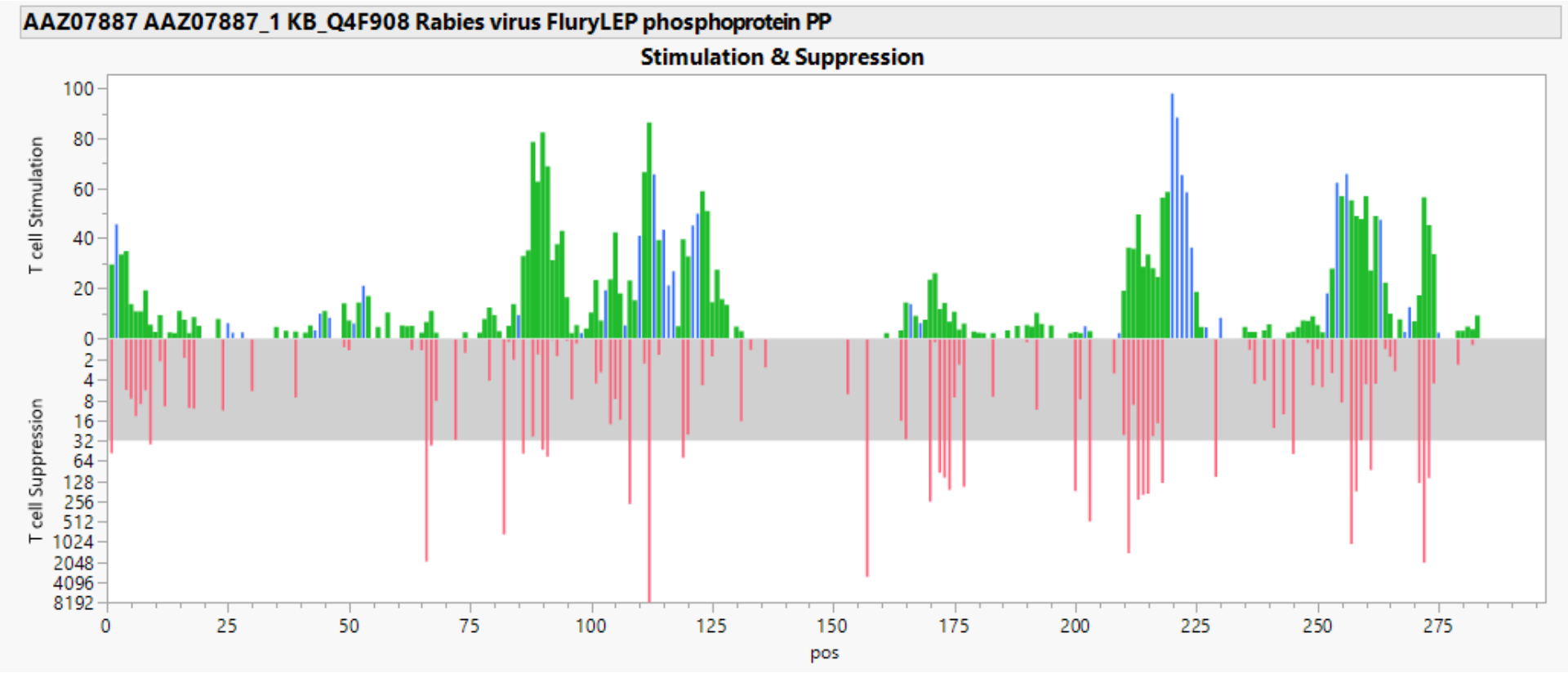

AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.05
0.08
0.10
0.13
0.90
1.68
2.45
3.23
4.00


Hierarchical Cluster MHC I
Method = Ward
AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t_T_MHC_I
nA_0101
nA_2601
nB_0702
nB_3501
nB_5301
nB_5101
nB_1501
nH-2-Db
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_2301
nA_2402
nA_3201
nA_2403
nB_0801
nH-2-Kb
nA_3001
nB_5401
nH-2-Ld
nB_1503
nB_2705
nH-2-Kd
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nH-2-Kk
nA_0301
nA_1101
nA_2902
nA_3002
nA_6801
nA_3101
nA_3301
-4.03
-3.22
-2.42
-1.61
-0.81
-0.00
0.99
1.97
2.96
3.95
4.94
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t MHC I & Cathepsins consolidated
nA_0101
nB_5101
nH-2-Ld
nH-2-Kd
nB_5401
nA_2902
nB_5301
nB_1503
nB_3501
nB_0702
nA_3201
nB_5701
nB_5801
nH-2-Dd
nA_0301
nA_3002
nA_1101
nA_6801
nA_2601
nA_3101
nA_3301
nA_3001
nB_1801
nB_4402
nB_4403
nB_2705
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0201
nH-2-Db
nA_2301
nA_2403
nA_2402
nB_1501
nB_0801
nH-2-Kb
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
0.00
0.02
0.03
0.05
0.07
0.08
0.67
1.25
1.83
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.3195
2.639
3.9585
5.2779
6.5974
22.678
38.758
54.839
70.919
87
0
25
50
75
100
125
150
175
200
225
250
275
N-term Amino Acid of 15-mer


AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10.0
0.0
0
25
50
75
100
125
150
175
200
225
250
275
pos

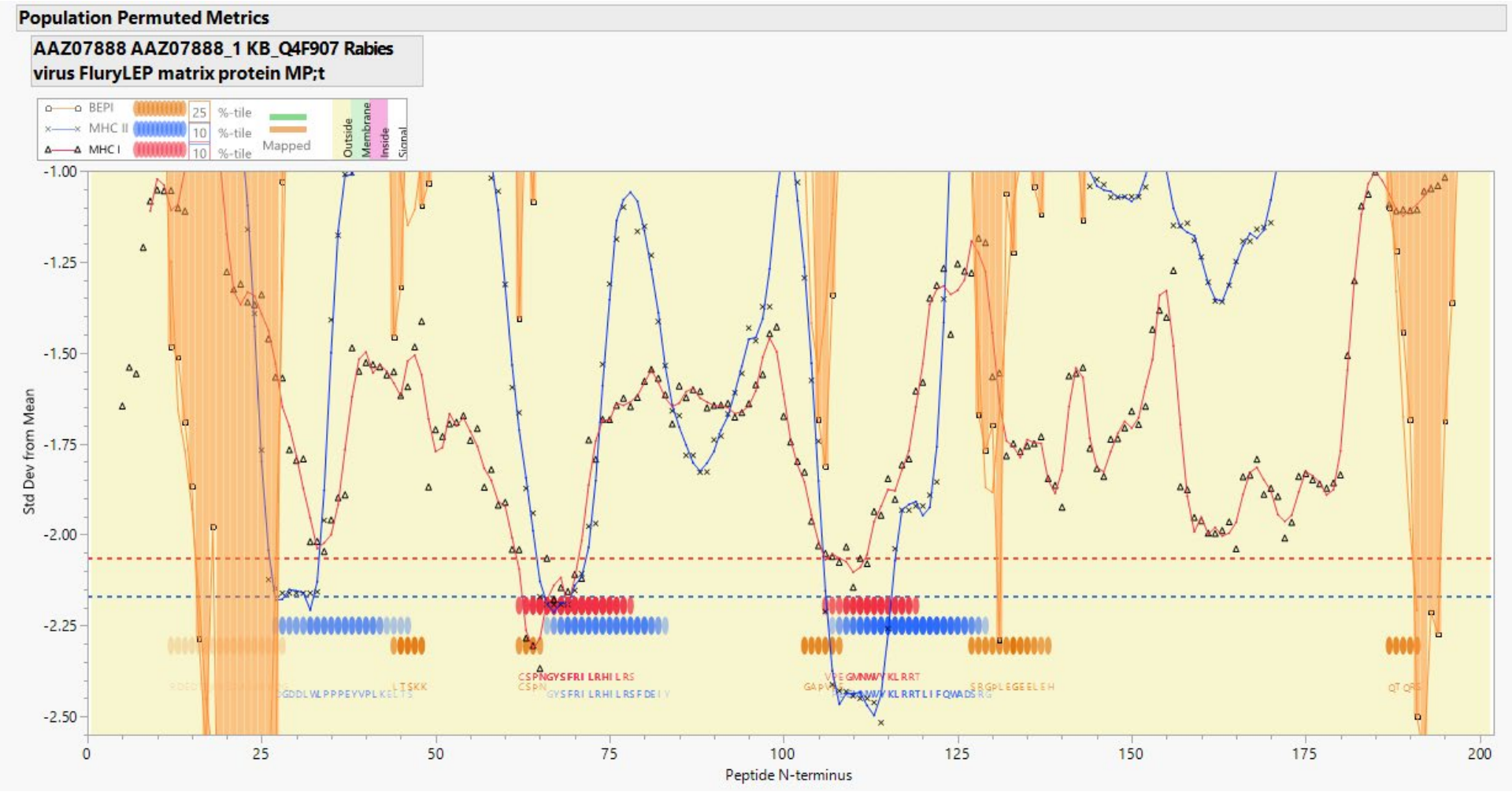

AAZ07887 AAZ07887_1 KB_Q4F908 Rabies virus FluryLEP phosphoprotein PP
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

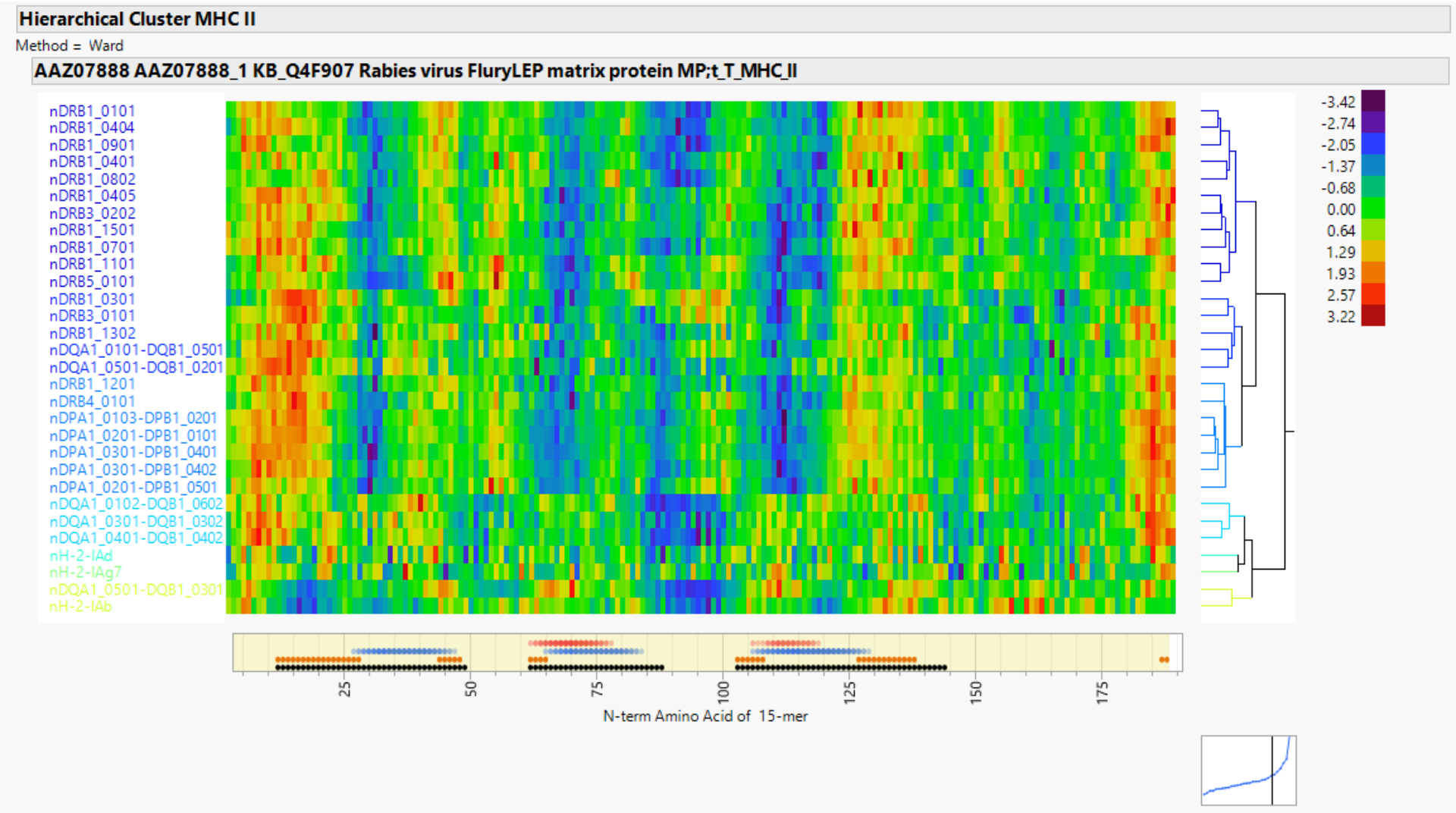


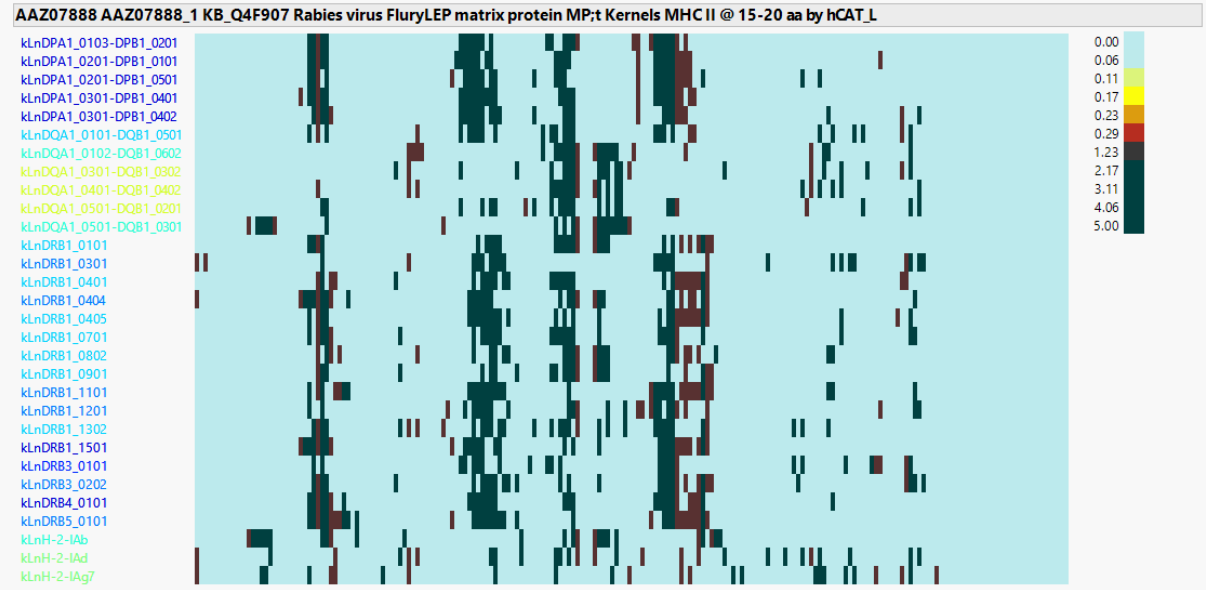
Matrix protein

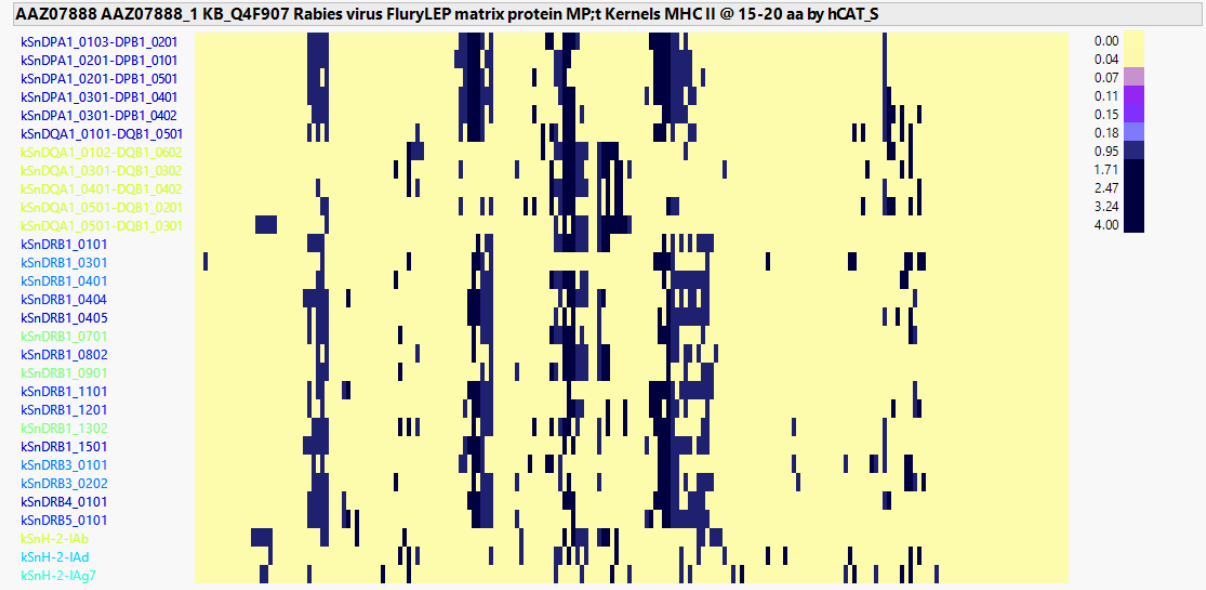
Population Permuted Metrics

AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t




Hierarchical Cluster MHC II
Method = Ward
AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t_T_MHC_II
nDRB1_0101
nDRB1_0404
nDRB1_0901
nDRB1_0401
nDRB1_0802
nDRB1_0405
nDRB3_0202
nDRB1_1501
nDRB1_0701
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDRB1_1302
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDRB1_1201
nDRB4_0101
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.42
-2.74
-2.05
-1.37
-0.68
0.00
0.64
1.29
1.93
2.57
3.22
25
50
75
100
125
150
175
N-term Amino Acid of 15-mer

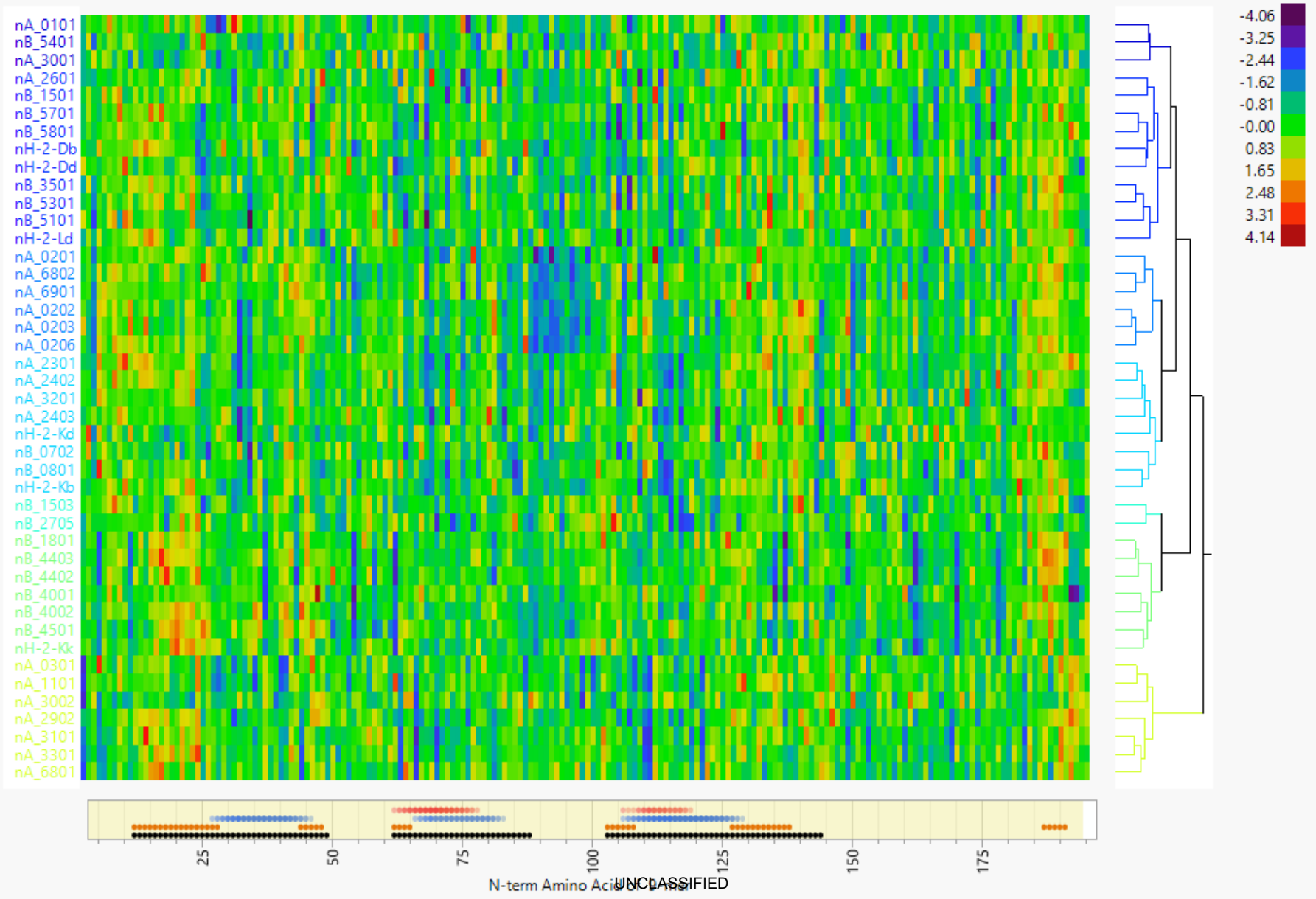

AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.06
0.11
0.17
0.23
0.29
1.23
2.17
3.11
4.06
5.00


AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.04
0.07
0.11
0.15
0.18
0.95
1.71
2.47
3.24
4.00

Hierarchical Cluster MHC I

Method = Ward

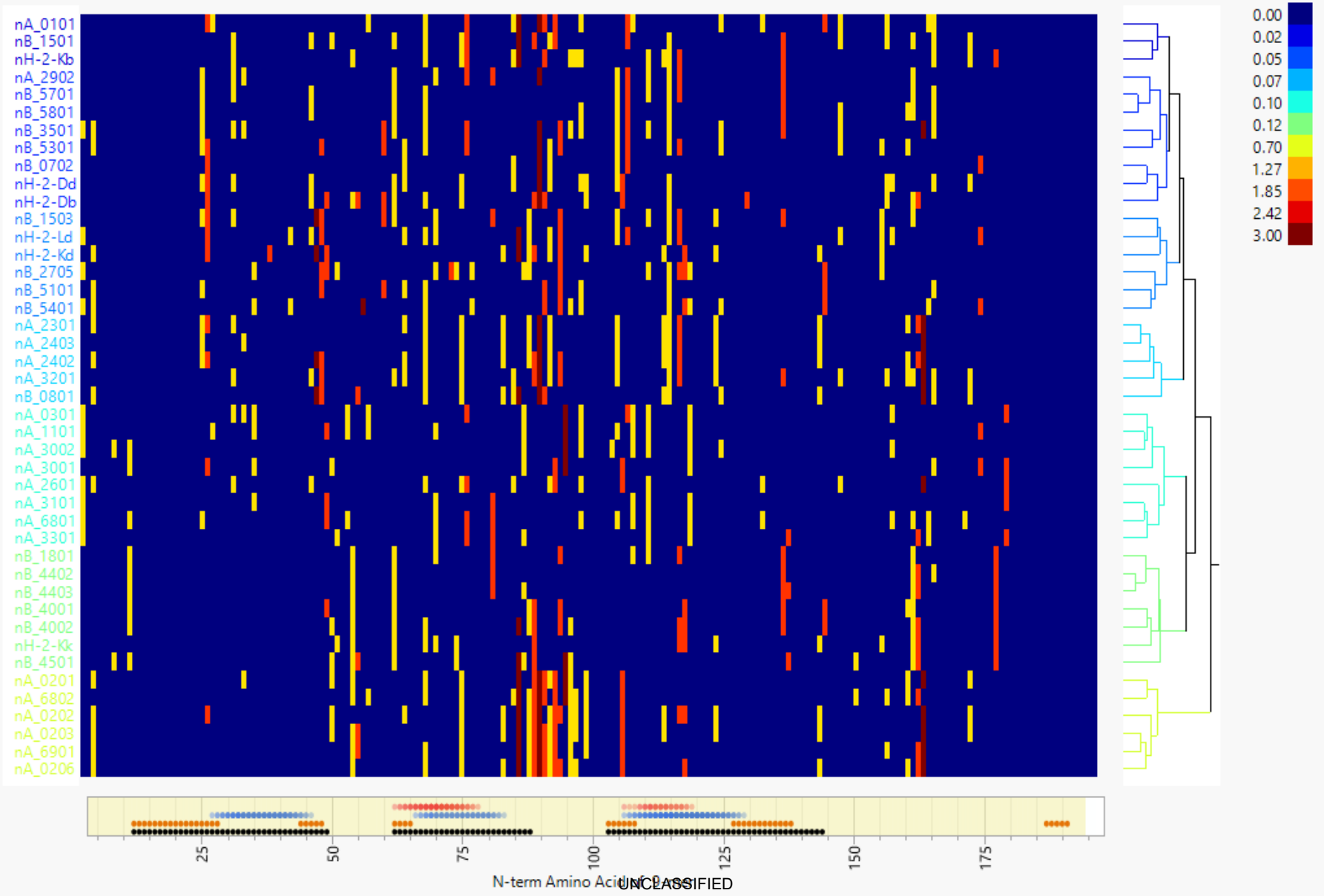
AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t_T_MHC_I



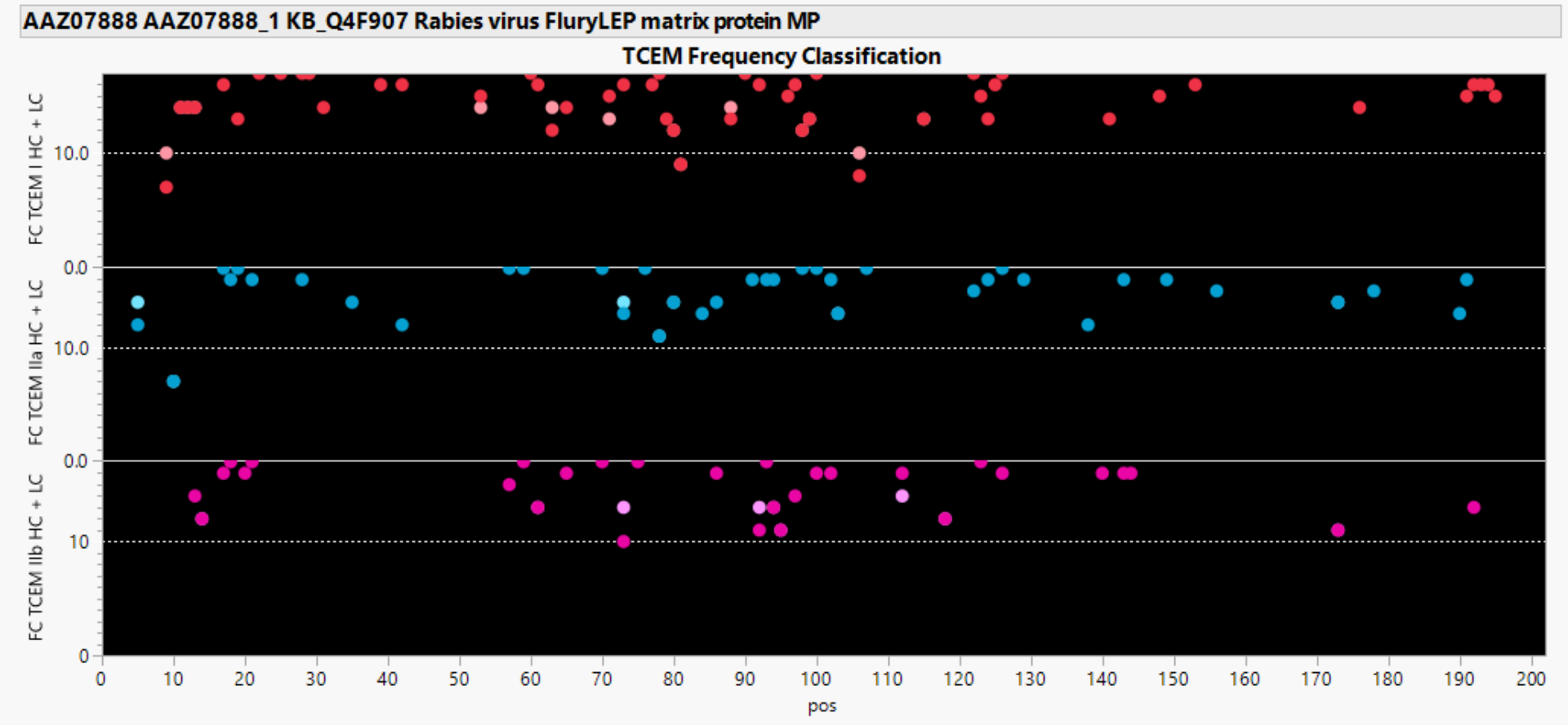
Hierarchical Cluster MHC I

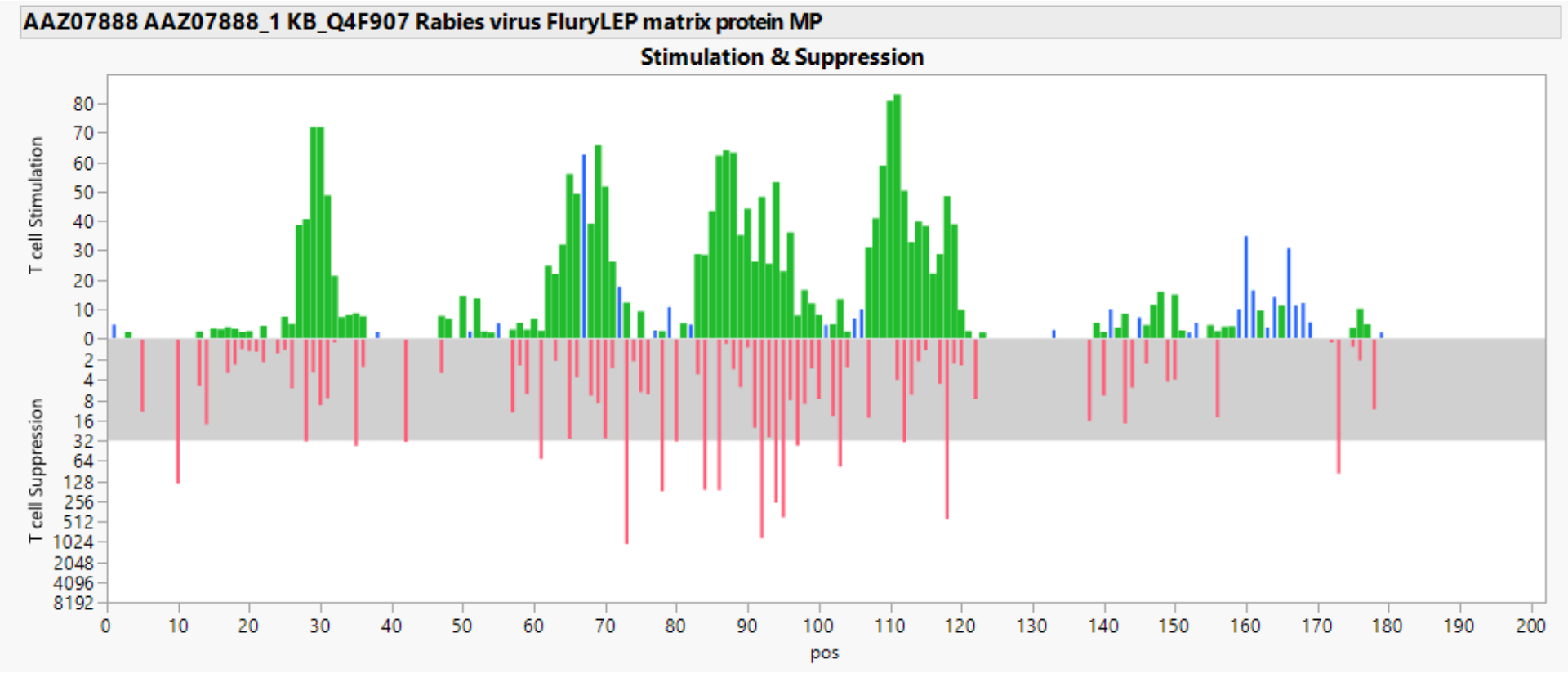
Method = Ward

AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t MHC I & Cathepsins consolidated




MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.3383
2.6766
4.015
5.3533
6.6916
19.953
33.215
46.477
59.738
73
0
25
50
75
100
125
150
175
200
N-term Amino Acid of 15-mer

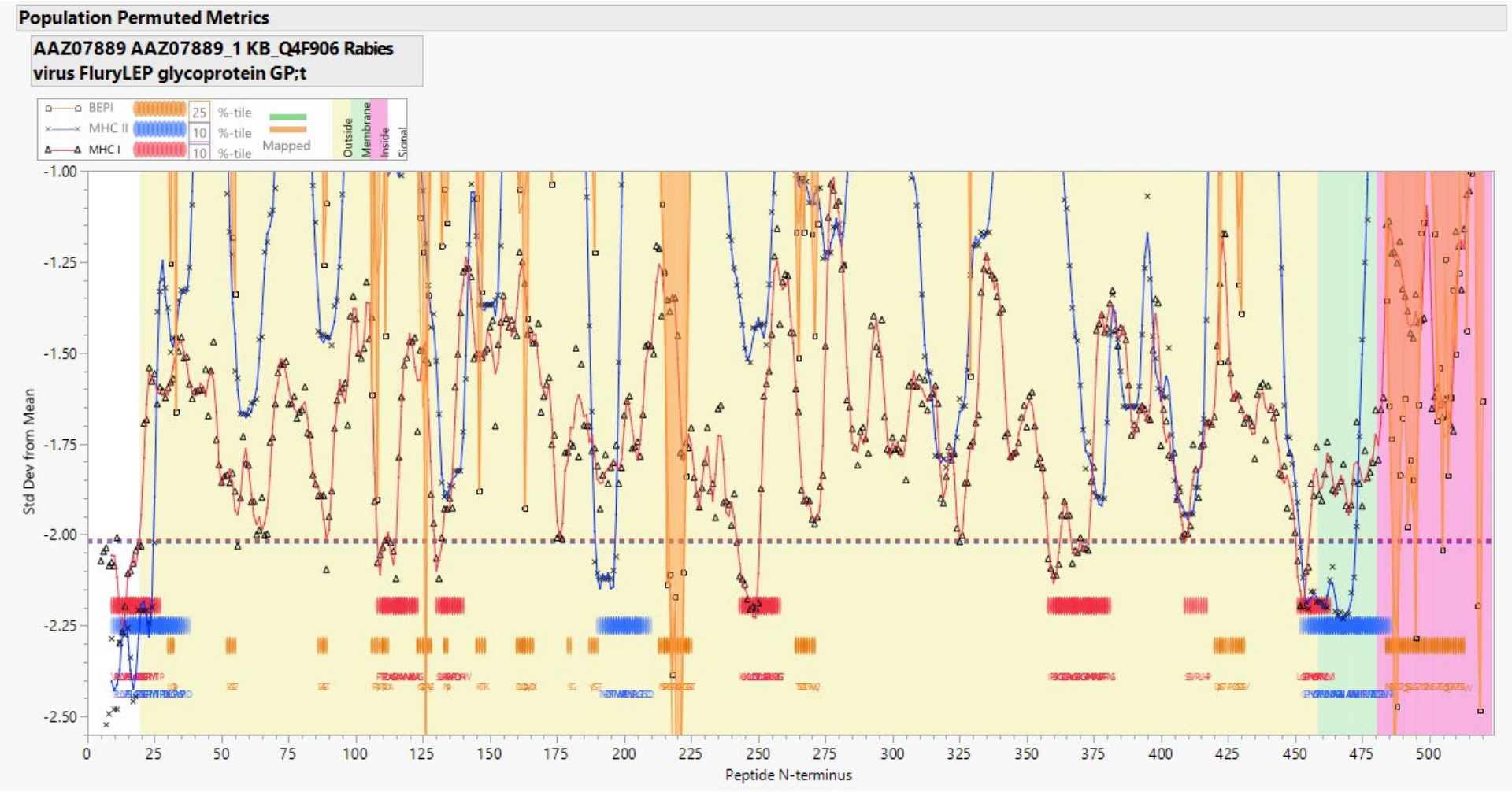

AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10.0
0.0
10.0
0.0
10
0
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
pos

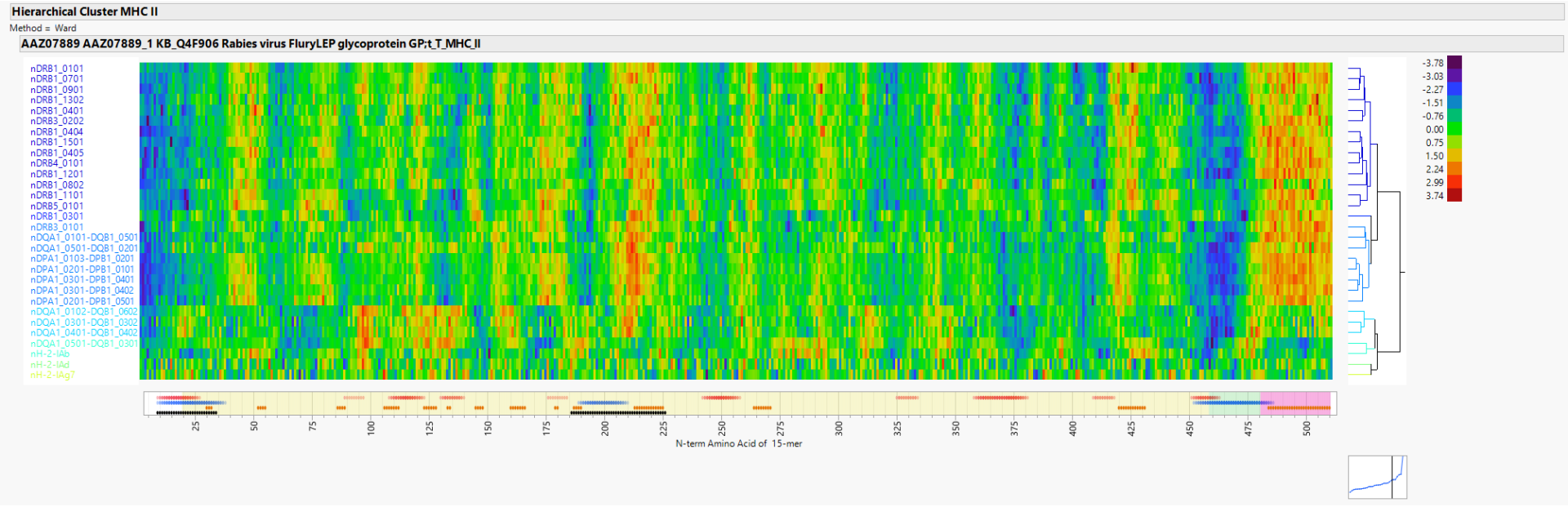

AAZ07888 AAZ07888_1 KB_Q4F907 Rabies virus FluryLEP matrix protein MP
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos

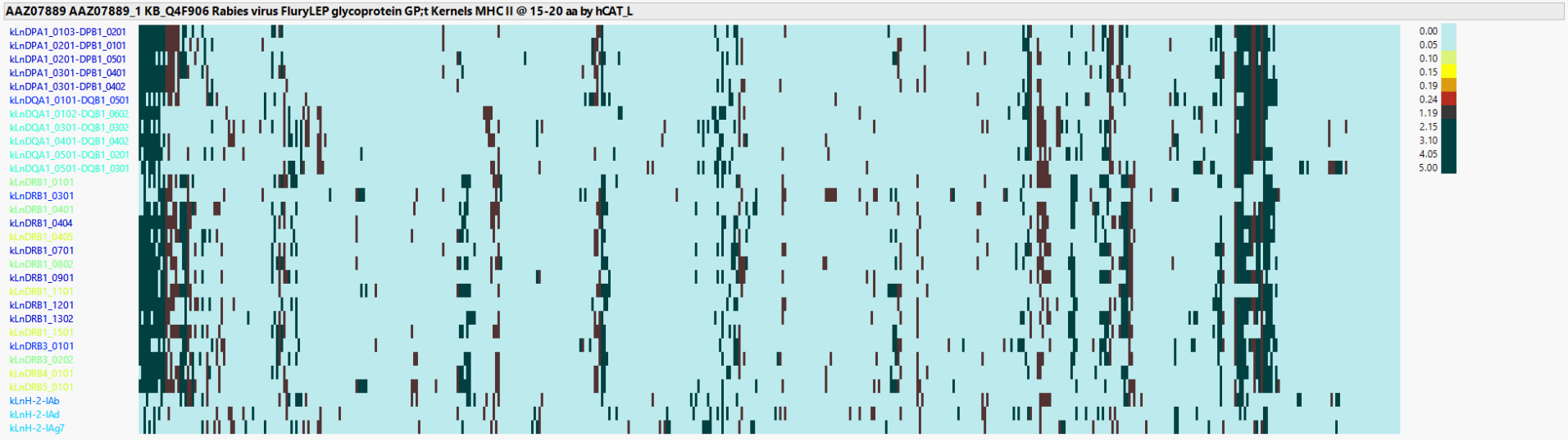
Glycoprotein

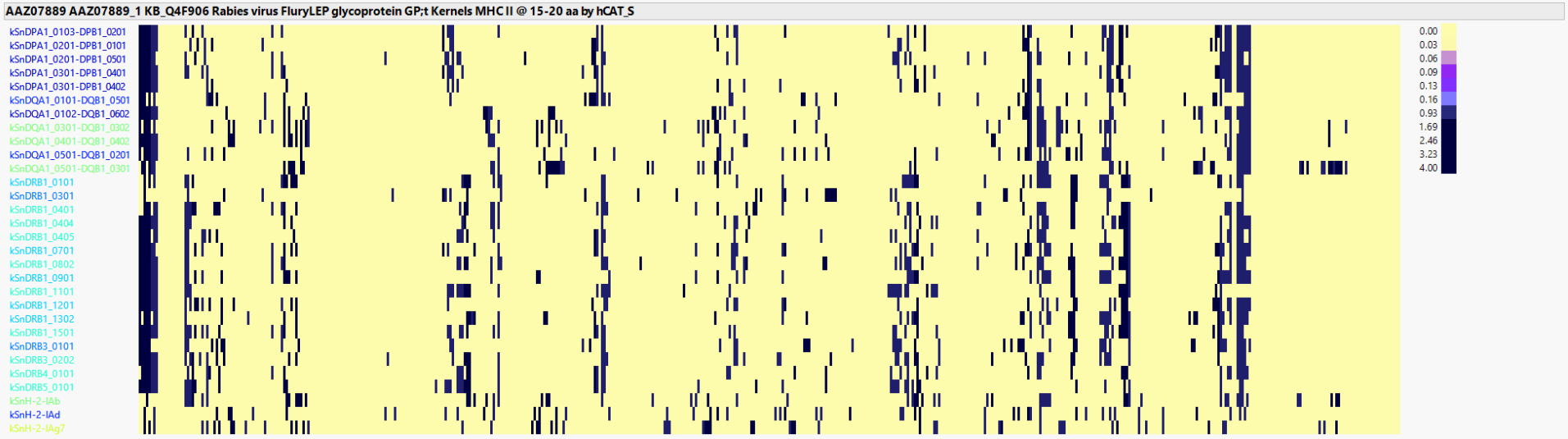
Population Permuted Metrics

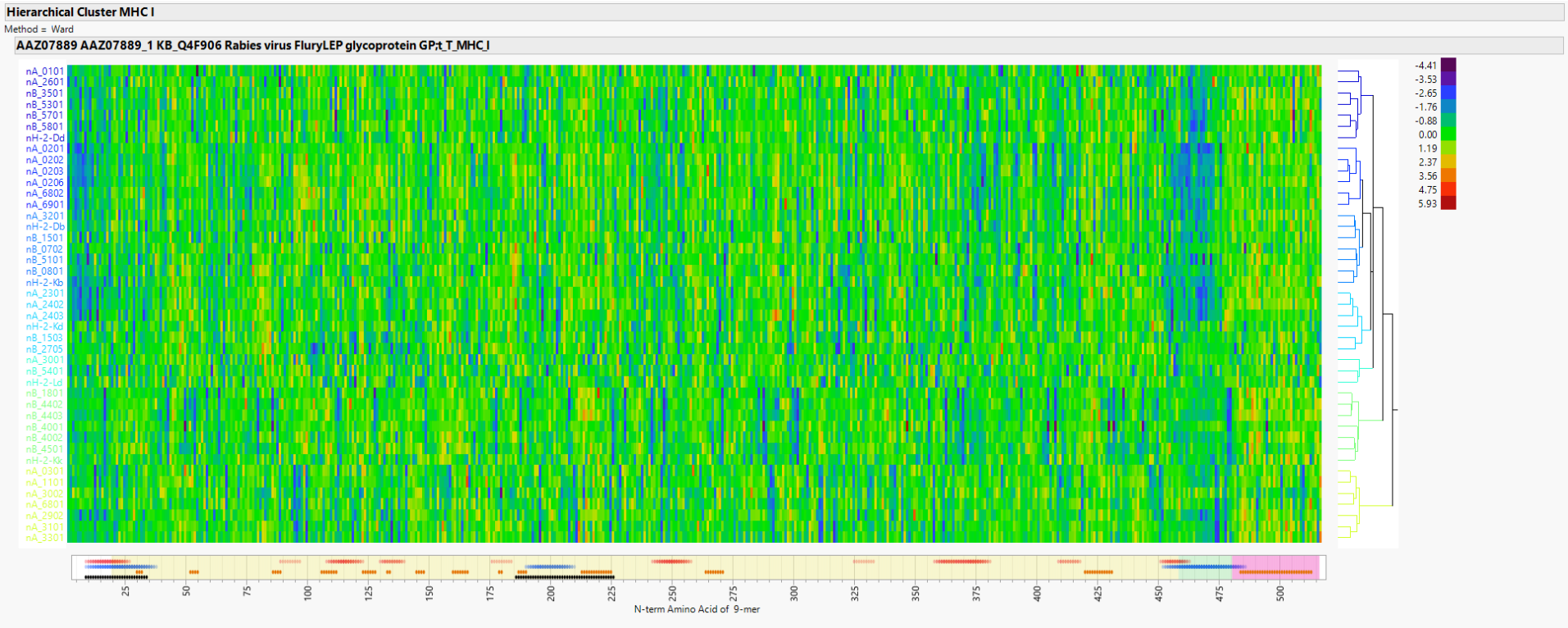
AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t

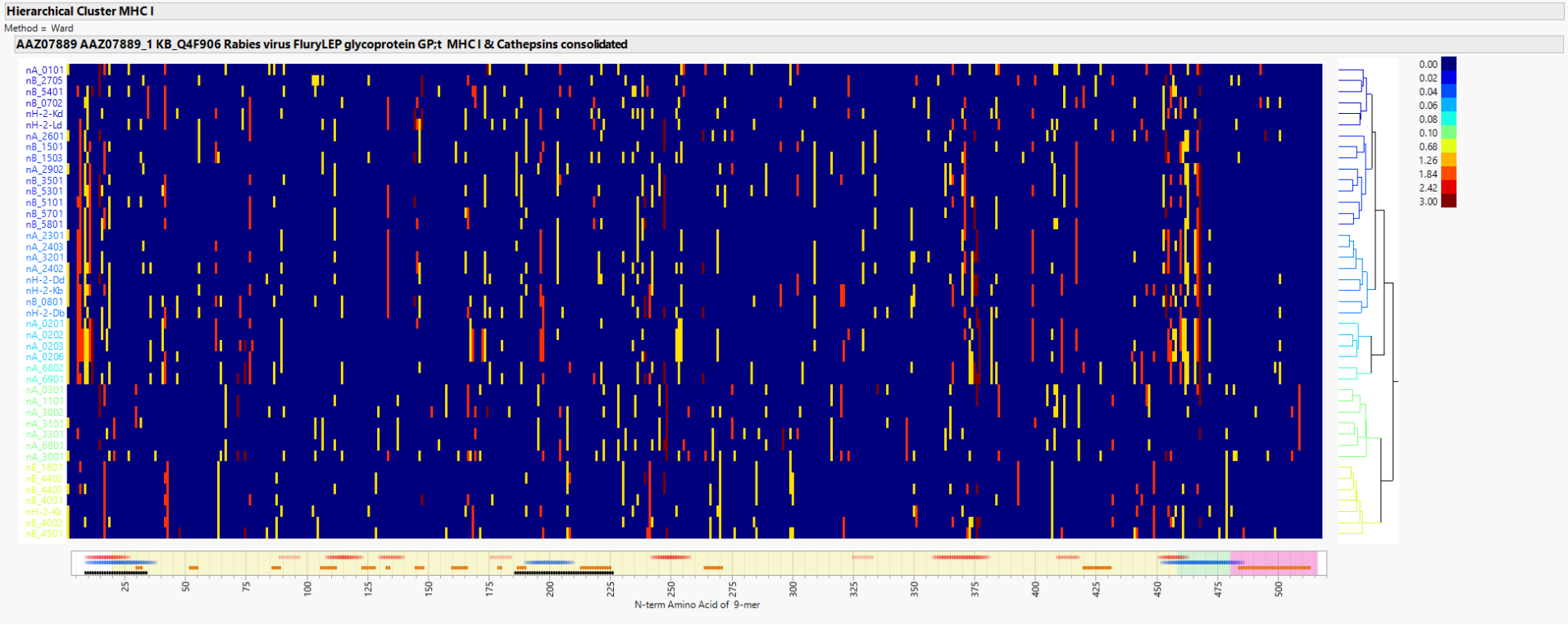


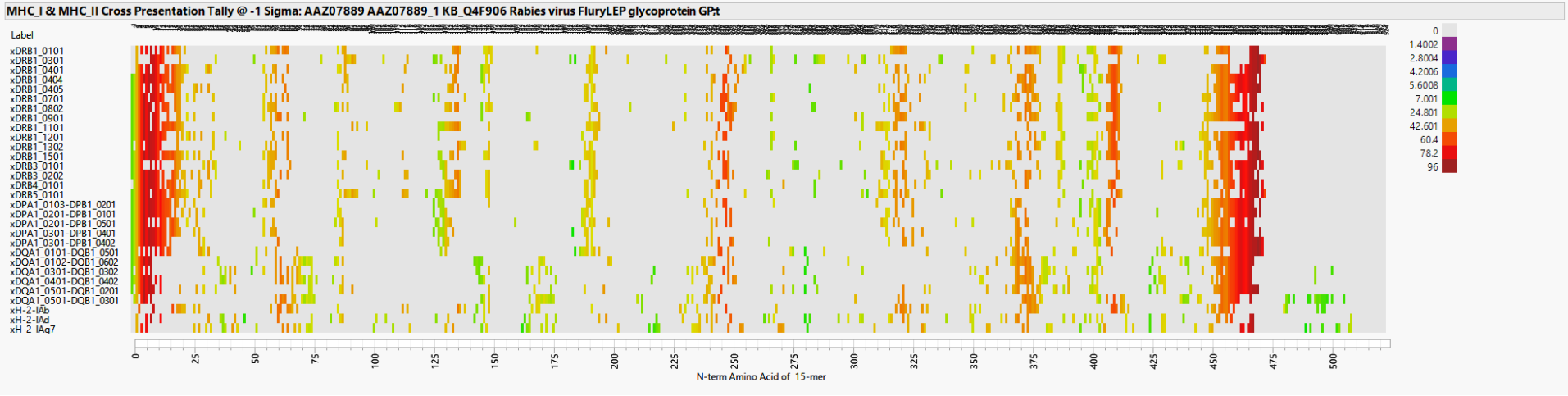

Hierarchical Cluster MHC II
Method = Ward
AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t_T_MHC_II
nDRB1_0101
nDRB1_0701
nDRB1_0901
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDRB1_1201
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB1_0301
nDRB3_0101
nDQA1_0101-DQB1_0501
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-3.78
-3.03
-2.27
-1.51
-0.76
0.00
0.75
1.50
2.24
2.99
3.74
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

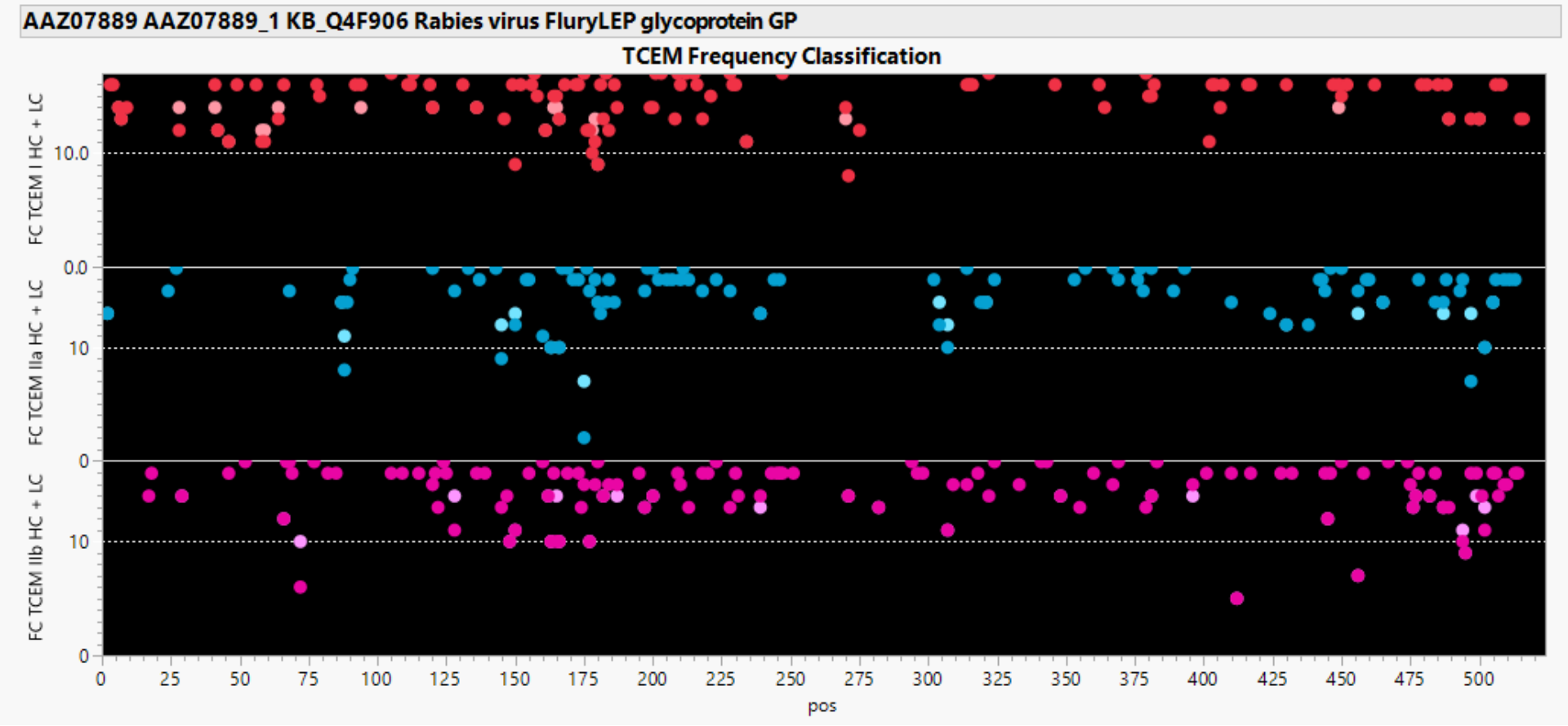

AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.05
0.10
0.15
0.19
0.24
1.19
2.15
3.10
4.05
5.00

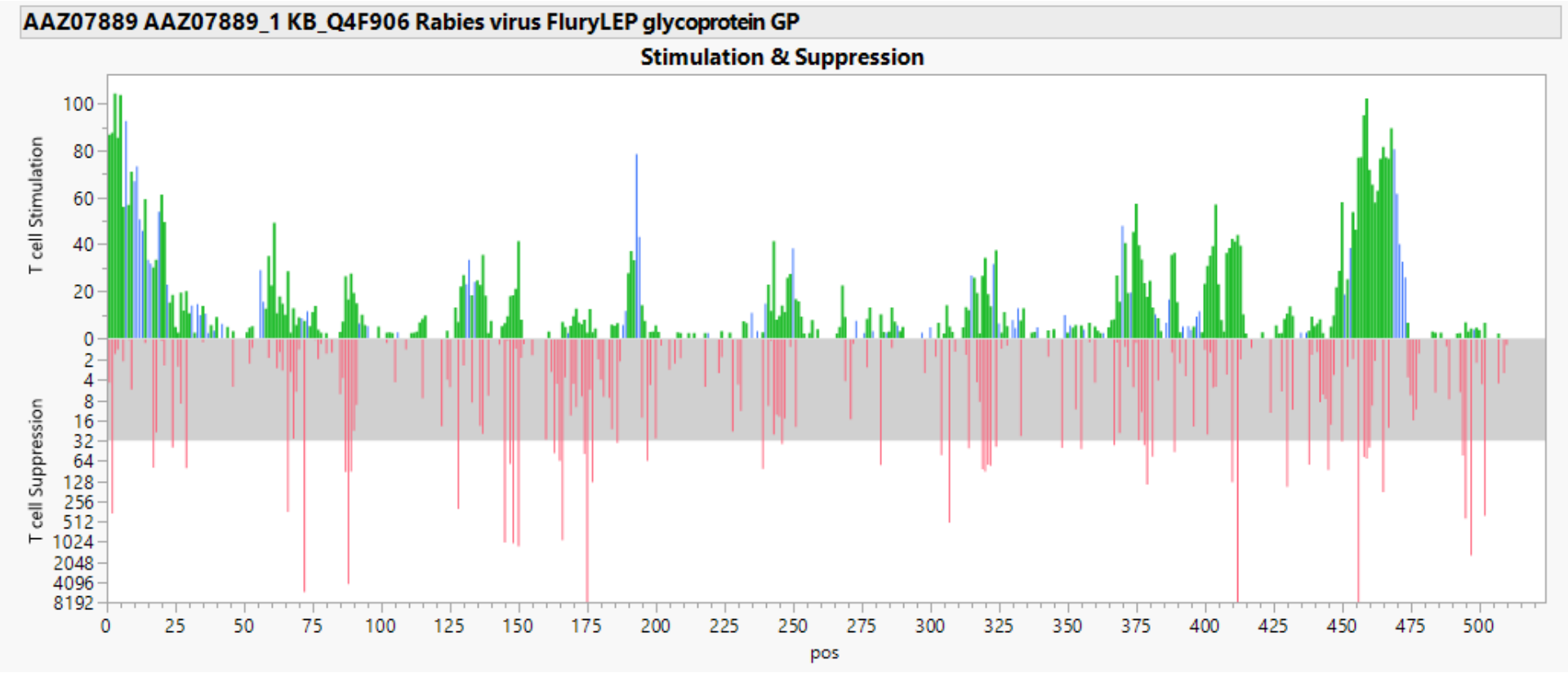

AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.06
0.09
0.13
0.16
0.93
1.69
2.46
3.23
4.00


Hierarchical Cluster MHC I
Method = Ward
AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_3201
nH-2-Db
nB_1501
nB_0702
nB_5101
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nA_3001
nB_5401
nH-2-Ld
nB_1801
nB_4402
nB_4403
nB_4001
nB_4002
nB_4501
nH-2-Kk
nA_0301
nA_1101
nA_3002
nA_6801
nA_2902
nA_3101
nA_3301
-4.41
-3.53
-2.65
-1.76
-0.88
0.00
1.19
2.37
3.56
4.75
5.93
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t MHC I & Cathepsins consolidated
nA_0101
nB_2705
nB_5401
nB_0702
nH-2-Kd
nH-2-Ld
nA_2601
nB_1501
nB_1503
nA_2902
nB_3501
nB_5301
nB_5101
nB_5701
nB_5801
nA_2301
nA_2403
nA_3201
nA_2402
nH-2-Dd
nH-2-Kb
nB_0801
nH-2-Db
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_0301
nA_1101
nA_3002
nA_3101
nA_3301
nA_6801
nA_3001
nB_1801
nB_4402
nB_4403
nB_4001
nH-2-Kk
nB_4002
nB_4501
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
1.4002
2.8004
4.2006
5.6008
7.001
24.801
42.601
60.4
78.2
96
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
N-term Amino Acid of 15-mer

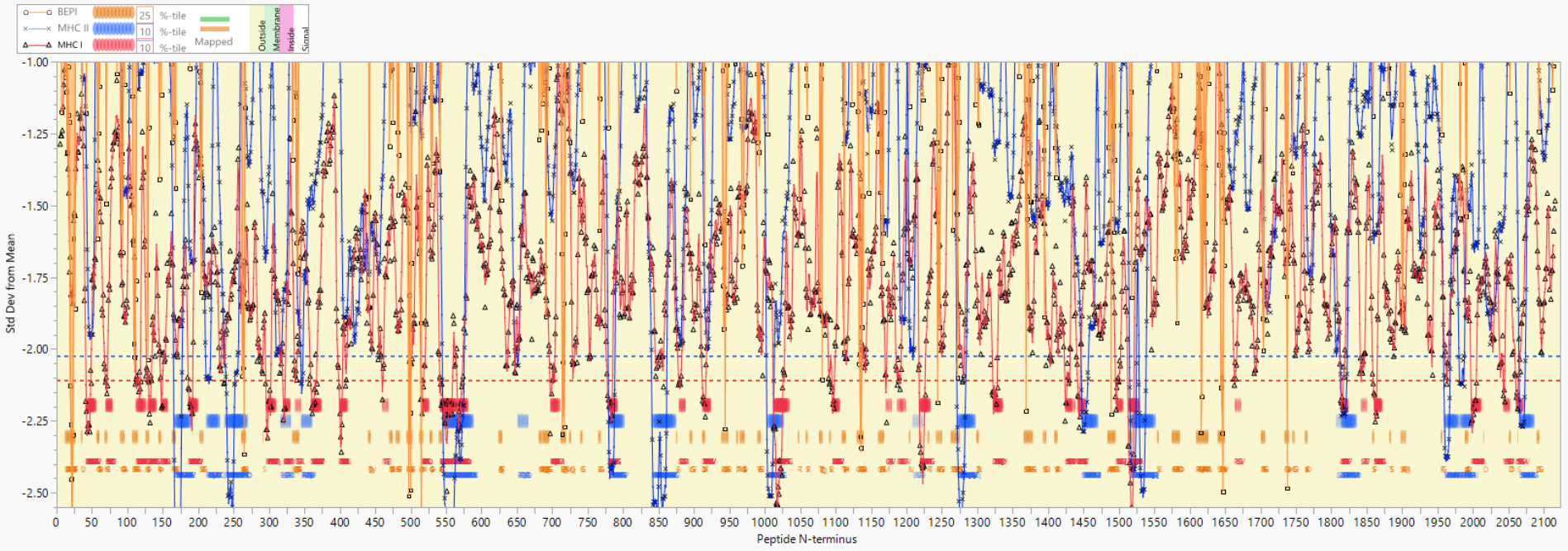

AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10.0
0.0
10
0
10
0
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
pos

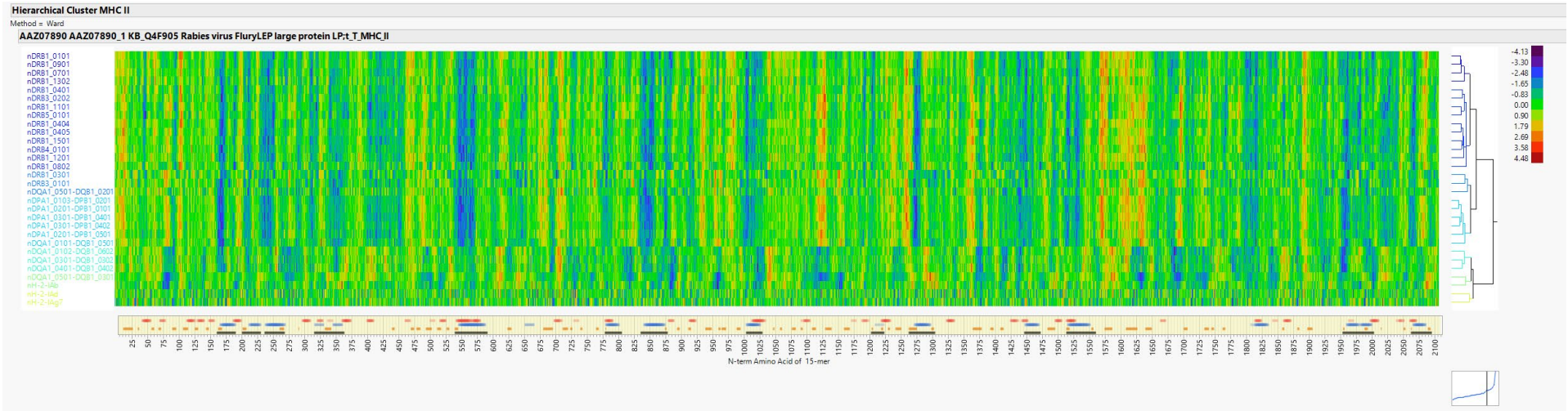

AAZ07889 AAZ07889_1 KB_Q4F906 Rabies virus FluryLEP glycoprotein GP
Stimulation & Suppression
T cell Stimulation
100
80
60
40
20
0
T cell Suppression
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
pos

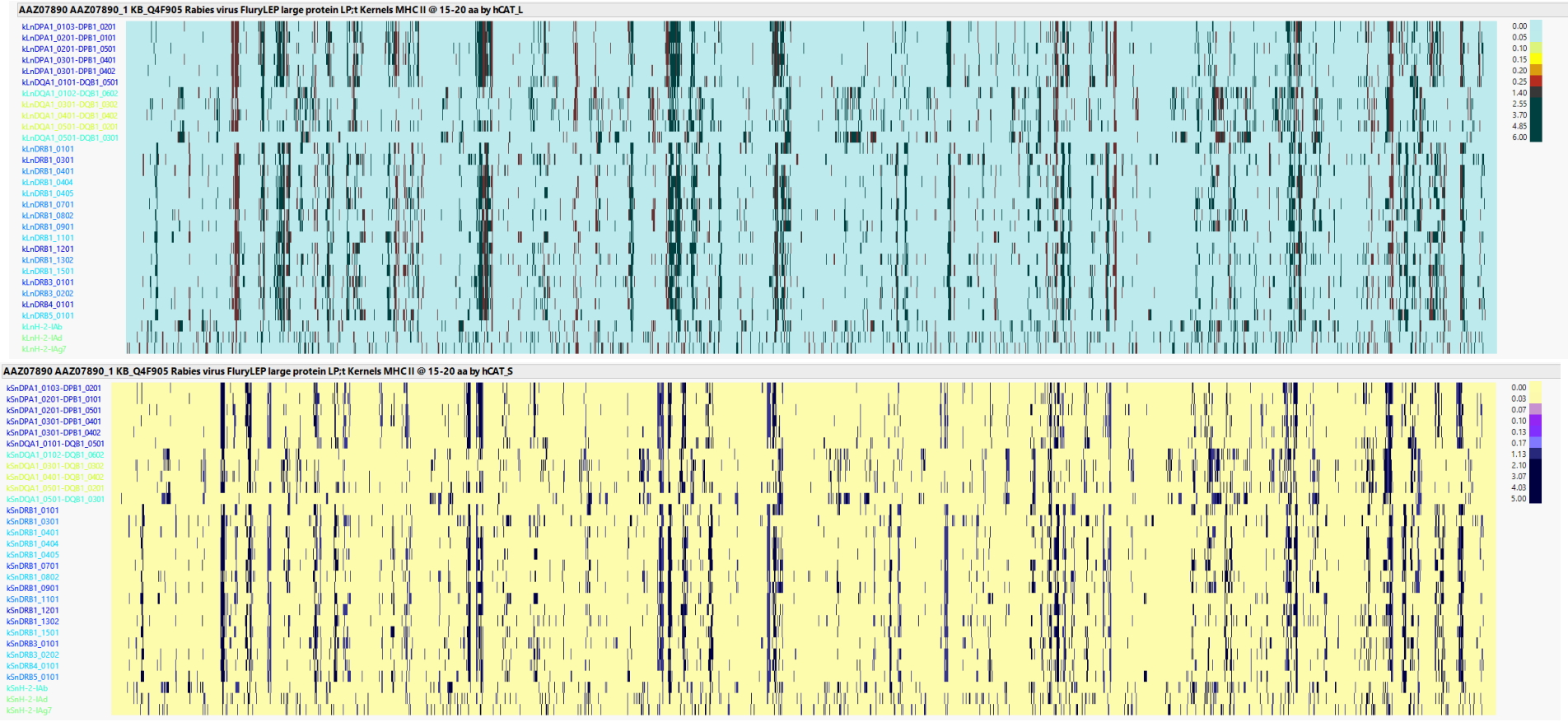
Large protein

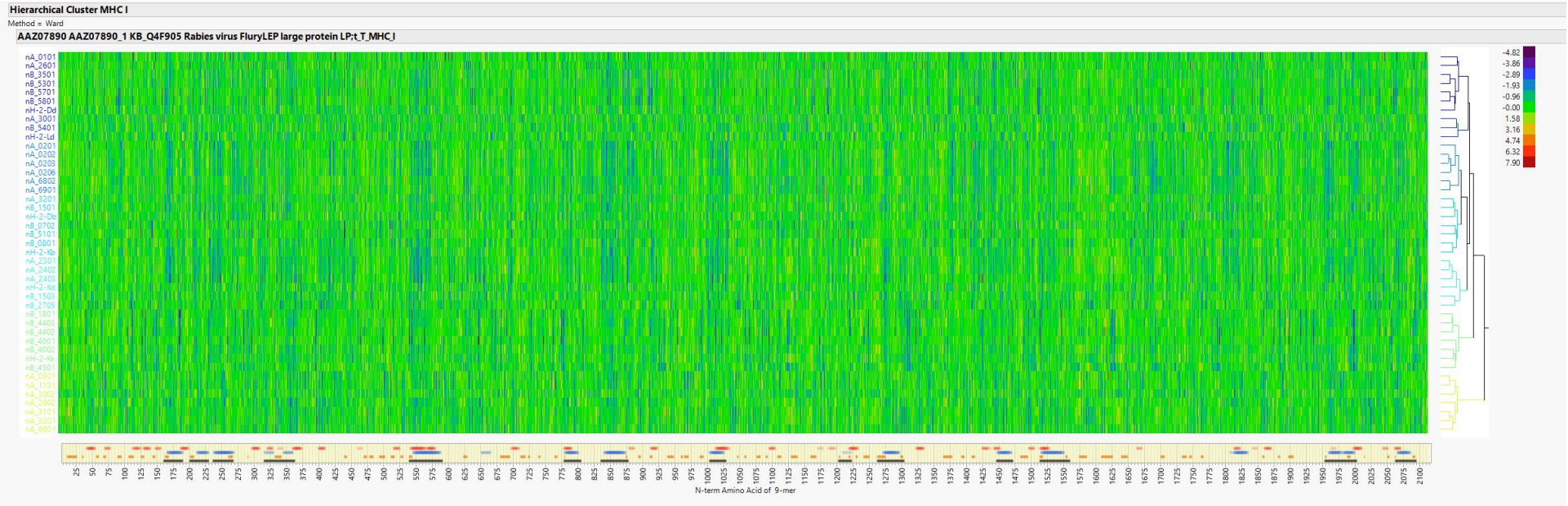
Population Permuted Metrics

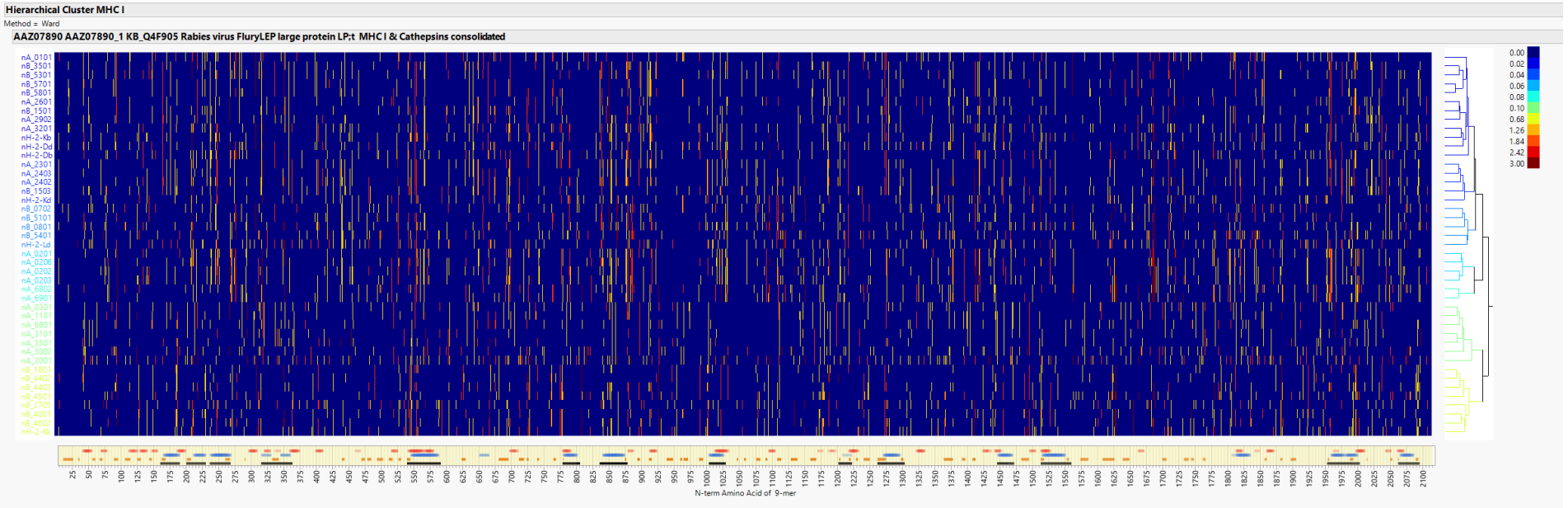
AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP;t

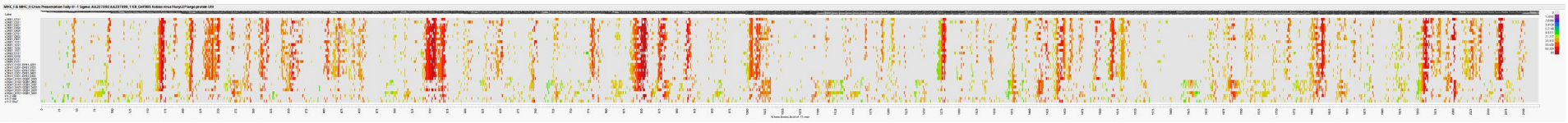


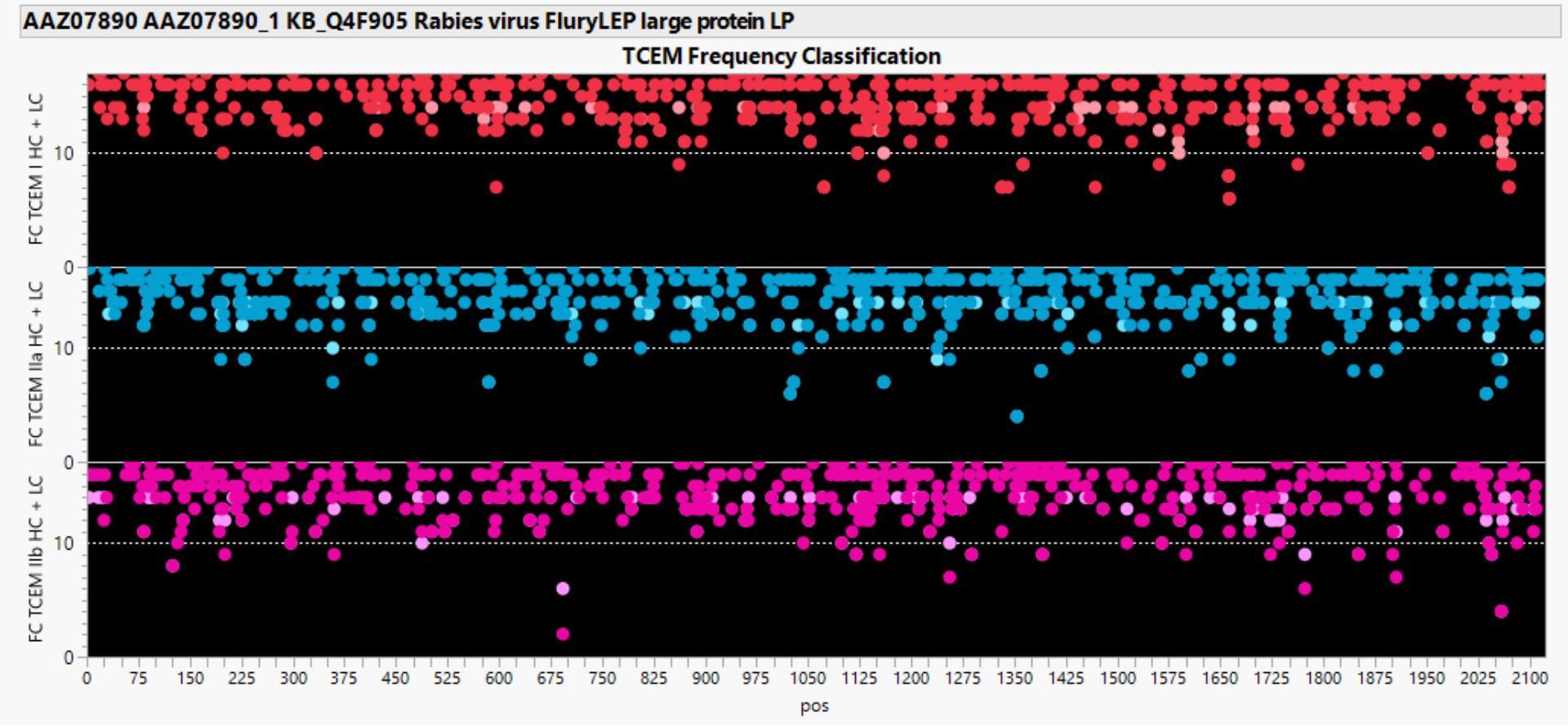

Hierarchical Cluster MHC II
Method = Ward
AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_0405
nDRB1_1501
nDRB4_0101
nDRB1_1201
nDRB1_0802
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0101-DQB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nDQA1_0501-DQB1_0301
nH-2-IAb
nH-2-IAd
nH-2-IAg7
-4.13
-3.30
-2.48
-1.65
-0.83
0.00
0.90
1.79
2.69
3.58
4.48
N-term Amino Acid of 15-mer

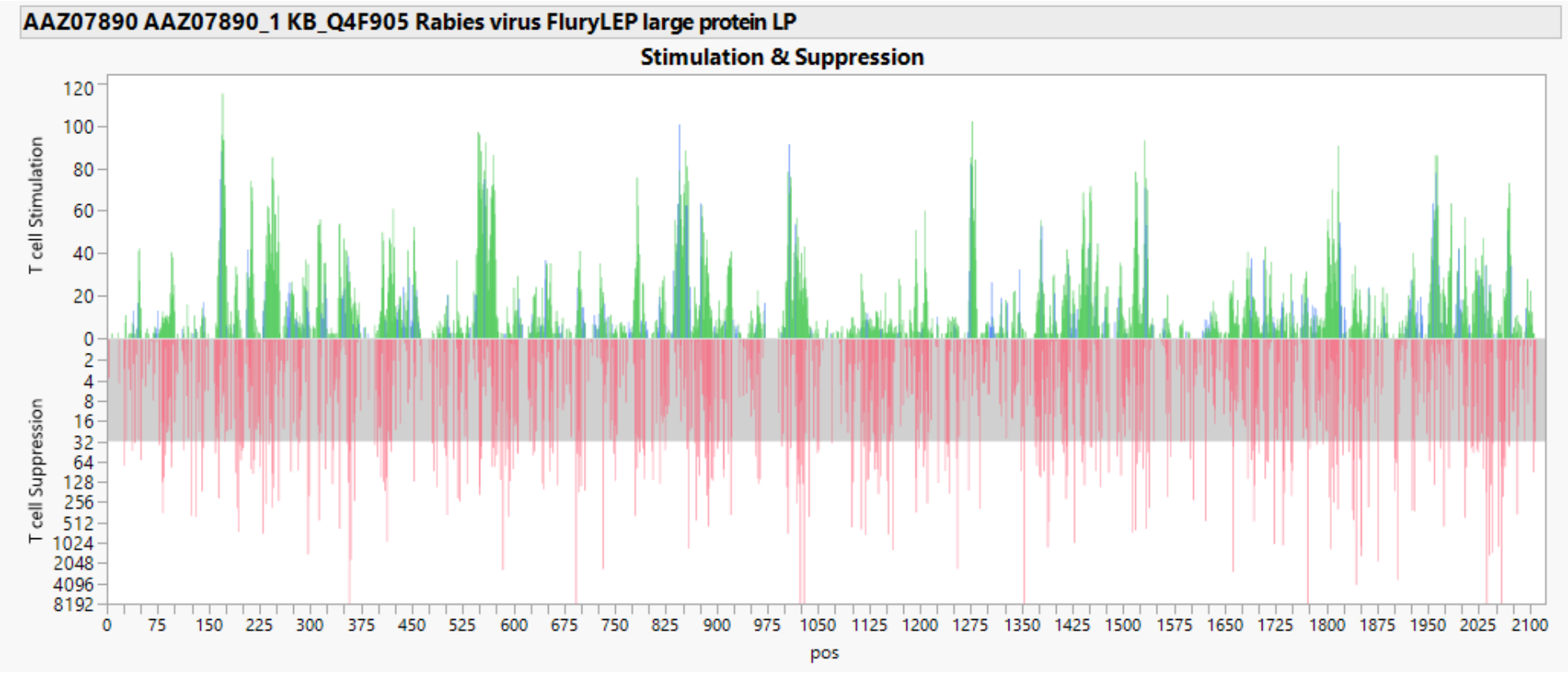

AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.05
0.10
0.15
0.20
0.25
1.40
2.55
3.70
4.85
6.00
AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.07
0.10
0.13
0.17
1.13
2.10
3.07
4.03
5.00


Hierarchical Cluster MHC I
Method = Ward
AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP;t_T_MHC_I
nA_0101
nA_2601
nB_3501
nB_5301
nB_5701
nB_5801
nH-2-Dd
nA_3001
nB_5401
nH-2-Ld
nA_0201
nA_0202
nA_0203
nA_0206
nA_6802
nA_6901
nA_3201
nB_1501
nH-2-Db
nB_0702
nB_5101
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
nA_0301
nA_1101
nA_3002
nA_2902
nA_3101
nA_3301
nA_6801
-4.82
-3.86
-2.89
-1.93
-0.96
-0.00
1.58
3.16
4.74
6.32
7.90
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP:t MHC I & Cathepsins consolidated
nA_0101
nB_3501
nB_5301
nB_5701
nB_5801
nA_2601
nB_1501
nA_2902
nA_3201
nH-2-Kb
nH-2-Dd
nH-2-Db
nA_2301
nA_2403
nA_2402
nB_1503
nH-2-Kd
nB_0702
nB_5101
nB_0801
nB_5401
nH-2-Ld
nA_0201
nA_0206
nA_0202
nA_0203
nA_6802
nA_6901
nA_0301
nA_1101
nA_6801
nA_3101
nA_3301
nA_3002
nA_3001
nB_1801
nB_4402
nB_4403
nB_4501
nB_2705
nB_4001
nB_4002
nH-2-Kk
0.00
0.02
0.04
0.06
0.08
0.10
0.68
1.26
1.84
2.42
3.00
N-term Amino Acid of 9-mer




AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10
0
0 75 150 225 300 375 450 525 600 675 750 825 900 975 1050 1125 1200 1275 1350 1425 1500 1575 1650 1725 1800 1875 1950 2025 2100
pos


AAZ07890 AAZ07890_1 KB_Q4F905 Rabies virus FluryLEP large protein LP
Stimulation & Suppression
T cell Stimulation
120
100
80
60
40
20
0
T cell Suppression
2
4
8
16
32
64
128
256
512
1024
2048
4096
8192
0
75
150
225
300
375
450
525
600
675
750
825
900
975
1050
1125
1200
1275
1350
1425
1500
1575
1650
1725
1800
1875
1950
2025
2100
pos

ioGenetics®

Addendum 10
Human alphaherpes virus 3 glycoprotein E

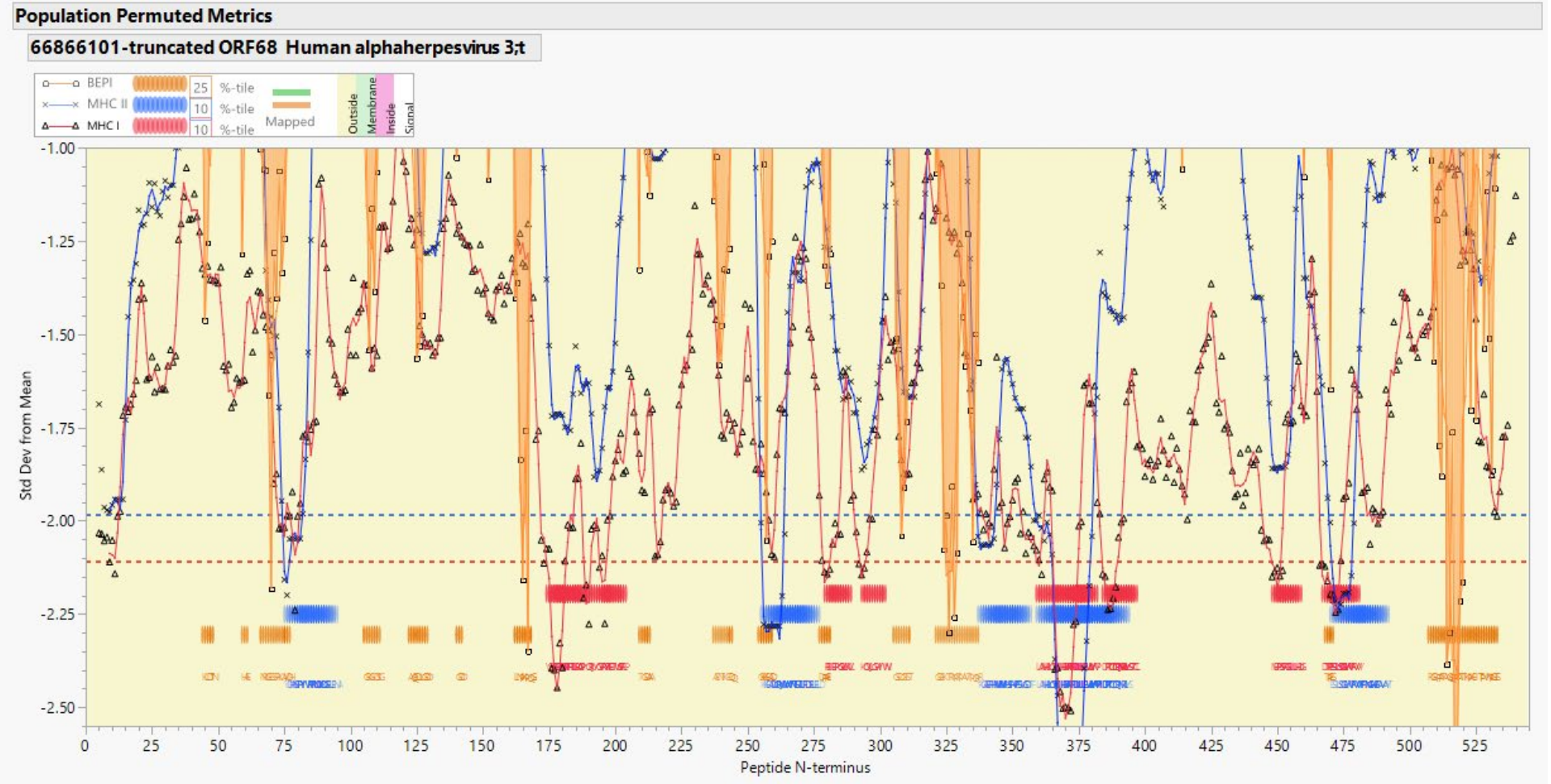
As used in Shingrix® vaccine

This protein is truncated relative to natural gE to remove the envelope insertion site

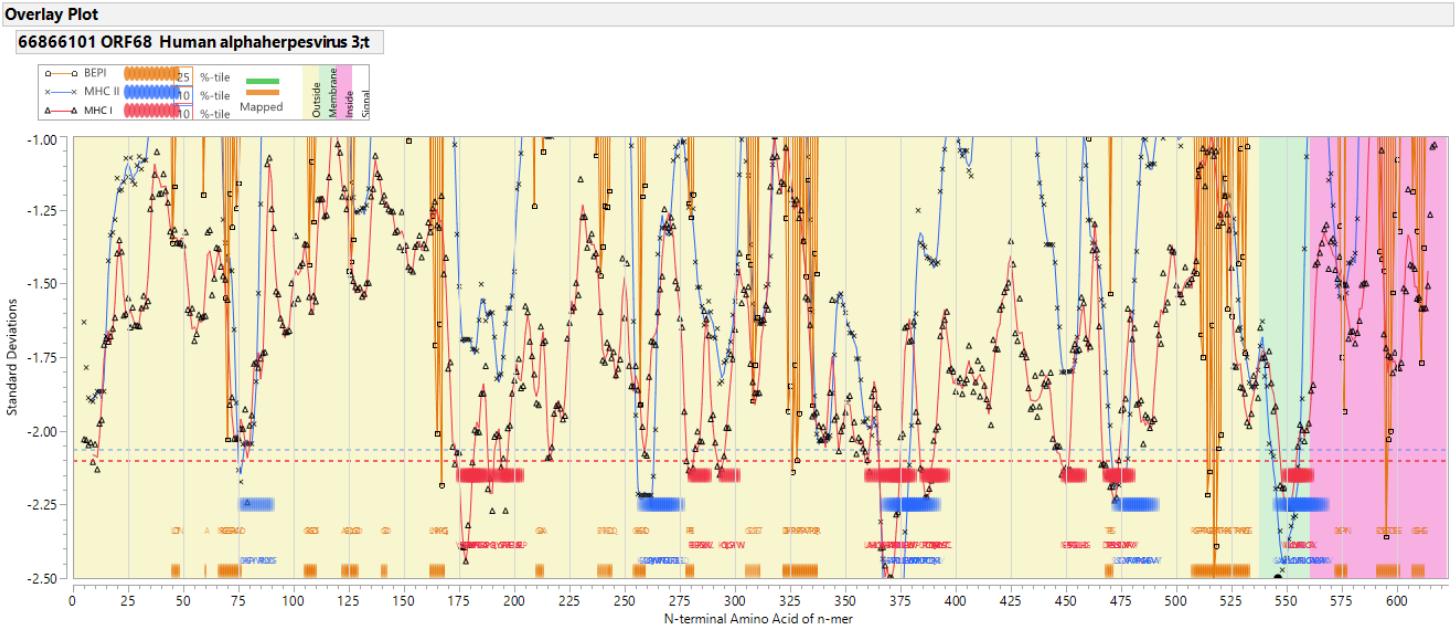
We show the natural configuration in slide 3; all other slides are the truncated form.

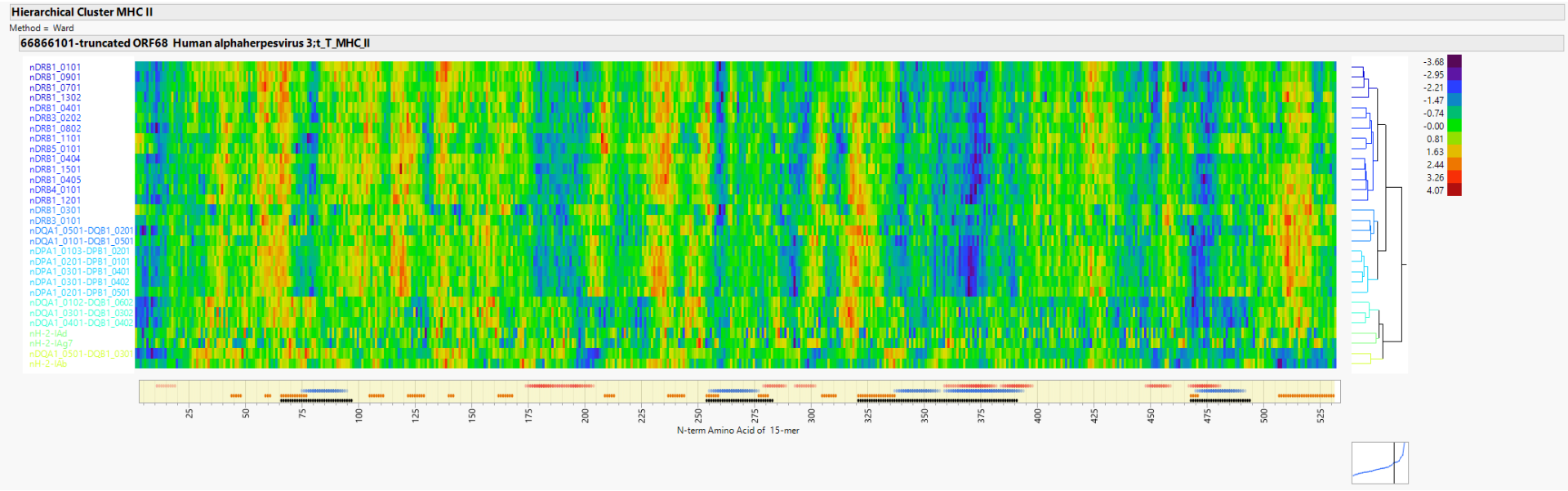
Population Permuted Metrics

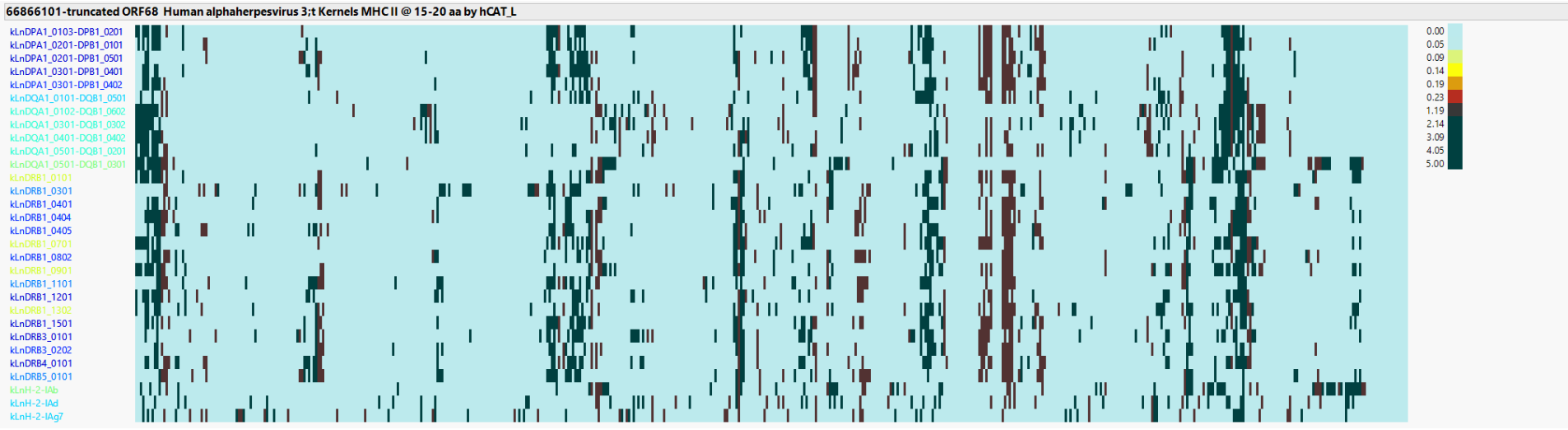
66866101-truncated ORF68 Human alphaherpesvirus 3;t

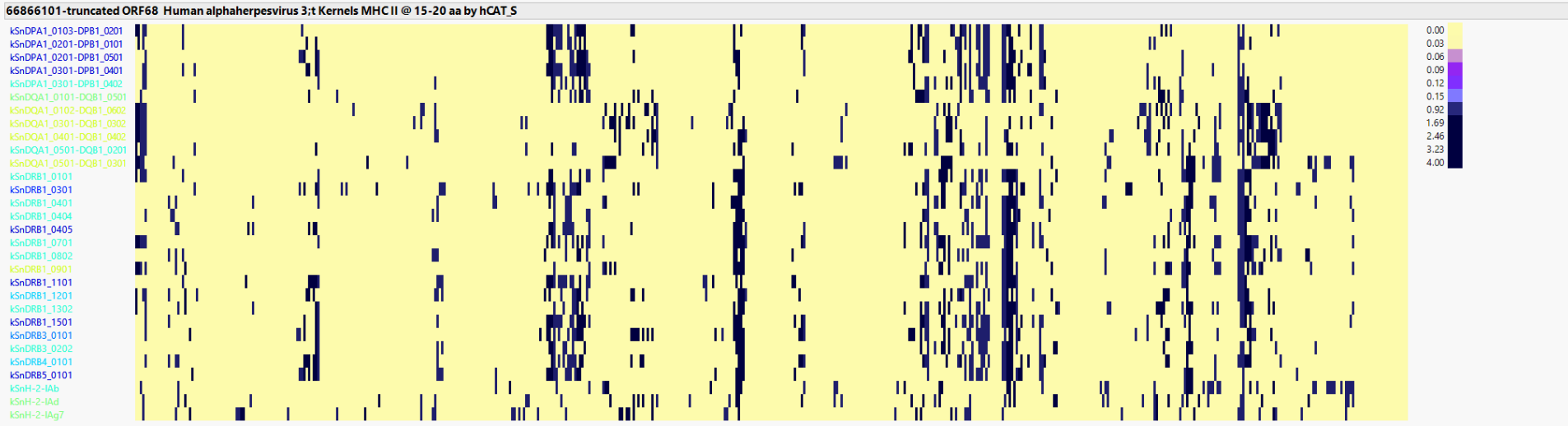


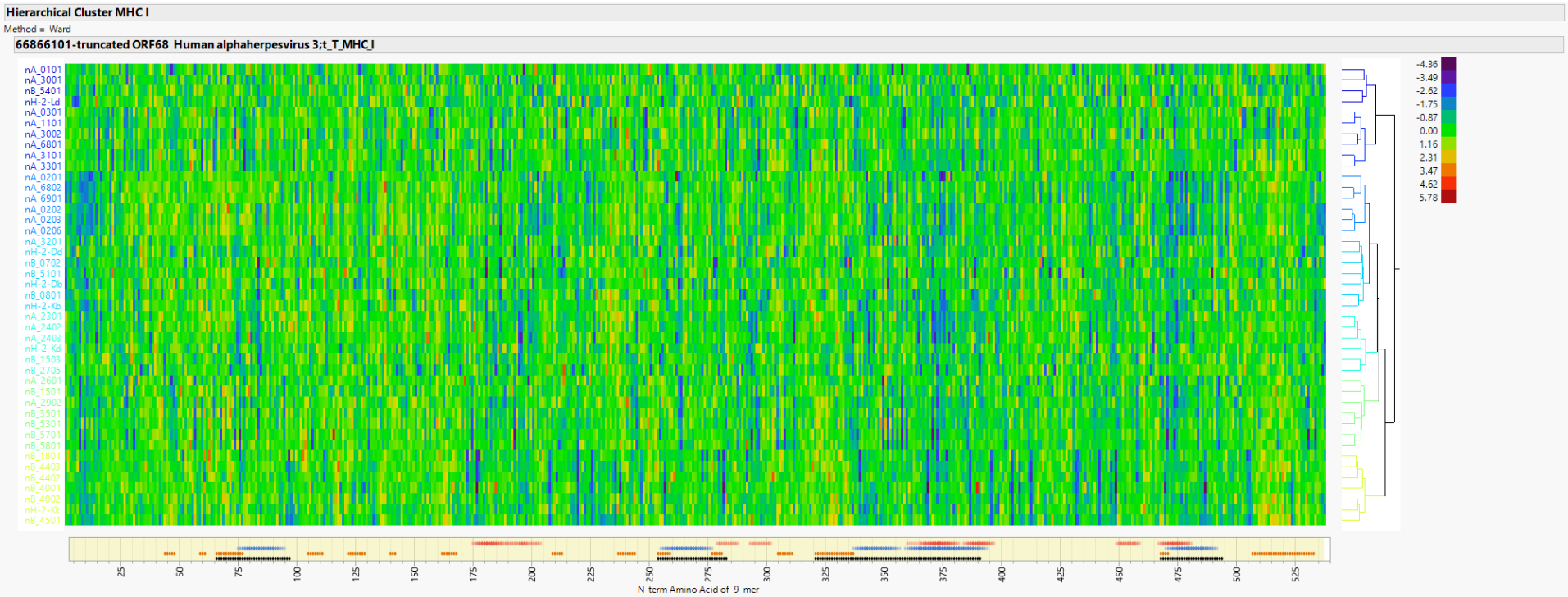
Natural form of gE

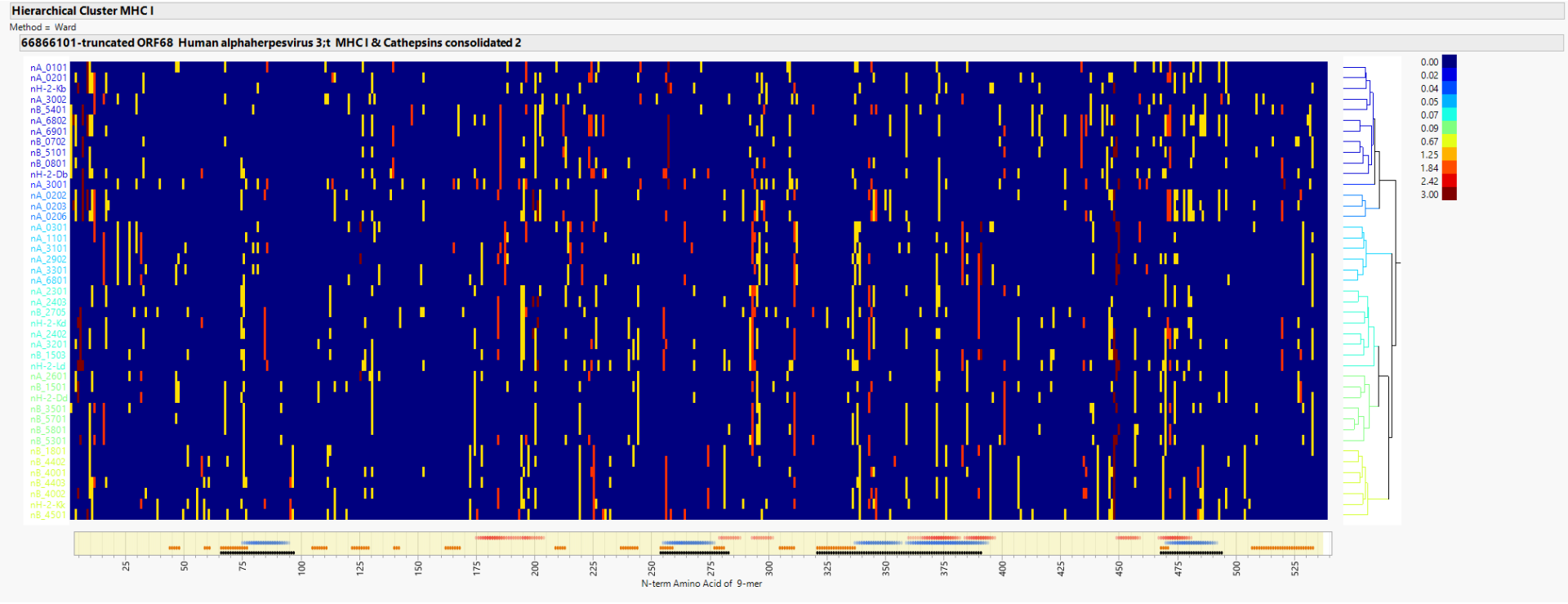


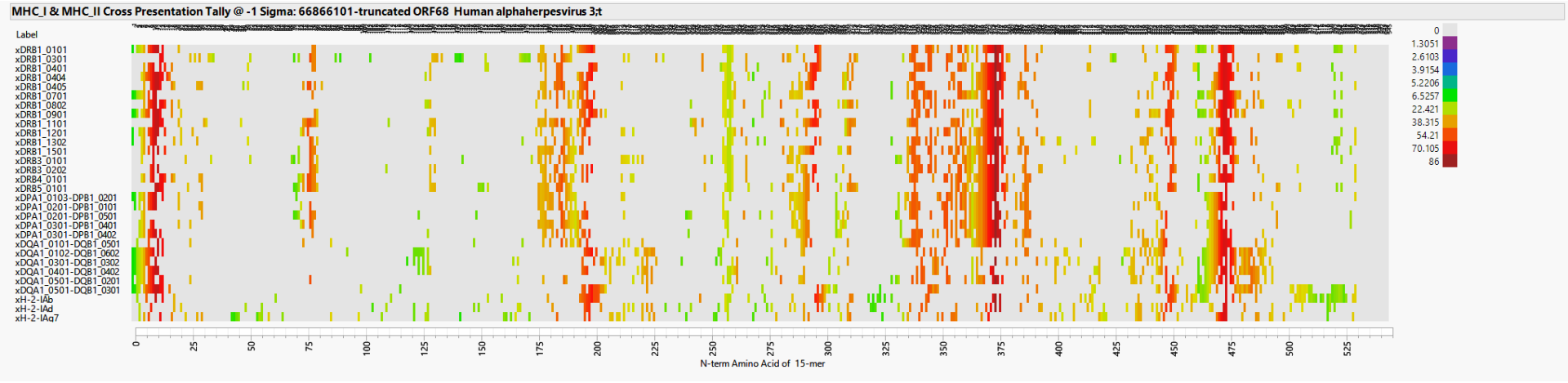

Hierarchical Cluster MHC II
Method = Ward
66866101-truncated ORF68 Human alphaherpesvirus 3;t_T_MHC_II
nDRB1_0101
nDRB1_0901
nDRB1_0701
nDRB1_1302
nDRB1_0401
nDRB3_0202
nDRB1_0802
nDRB1_1101
nDRB5_0101
nDRB1_0404
nDRB1_1501
nDRB1_0405
nDRB4_0101
nDRB1_1201
nDRB1_0301
nDRB3_0101
nDQA1_0501-DQB1_0201
nDQA1_0101-DQB1_0501
nDPA1_0103-DPB1_0201
nDPA1_0201-DPB1_0101
nDPA1_0301-DPB1_0401
nDPA1_0301-DPB1_0402
nDPA1_0201-DPB1_0501
nDQA1_0102-DQB1_0602
nDQA1_0301-DQB1_0302
nDQA1_0401-DQB1_0402
nH-2-IAd
nH-2-IAg7
nDQA1_0501-DQB1_0301
nH-2-IAb
-3.68
-2.95
-2.21
-1.47
-0.74
-0.00
0.81
1.63
2.44
3.26
4.07
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer

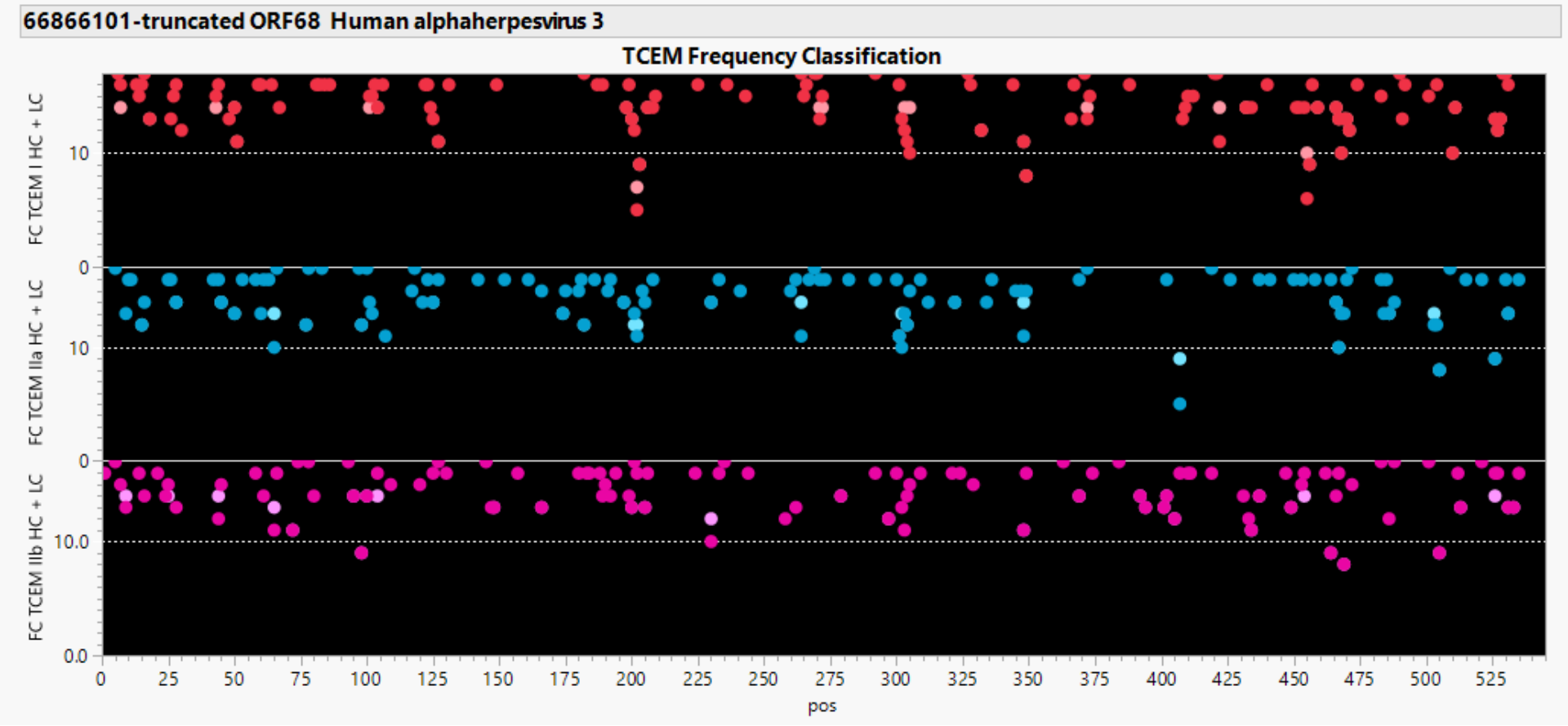

66866101-truncated ORF68 Human alphaherpesvirus 3;t Kernels MHC II @ 15-20 aa by hCAT_L
kLnDPA1_0103-DPB1_0201
kLnDPA1_0201-DPB1_0101
kLnDPA1_0201-DPB1_0501
kLnDPA1_0301-DPB1_0401
kLnDPA1_0301-DPB1_0402
kLnDQA1_0101-DQB1_0501
kLnDQA1_0102-DQB1_0602
kLnDQA1_0301-DQB1_0302
kLnDQA1_0401-DQB1_0402
kLnDQA1_0501-DQB1_0201
kLnDQA1_0501-DQB1_0301
kLnDRB1_0101
kLnDRB1_0301
kLnDRB1_0401
kLnDRB1_0404
kLnDRB1_0405
kLnDRB1_0701
kLnDRB1_0802
kLnDRB1_0901
kLnDRB1_1101
kLnDRB1_1201
kLnDRB1_1302
kLnDRB1_1501
kLnDRB3_0101
kLnDRB3_0202
kLnDRB4_0101
kLnDRB5_0101
kLnH-2-IAb
kLnH-2-IAd
kLnH-2-IAg7
0.00
0.05
0.09
0.14
0.19
0.23
1.19
2.14
3.09
4.05
5.00

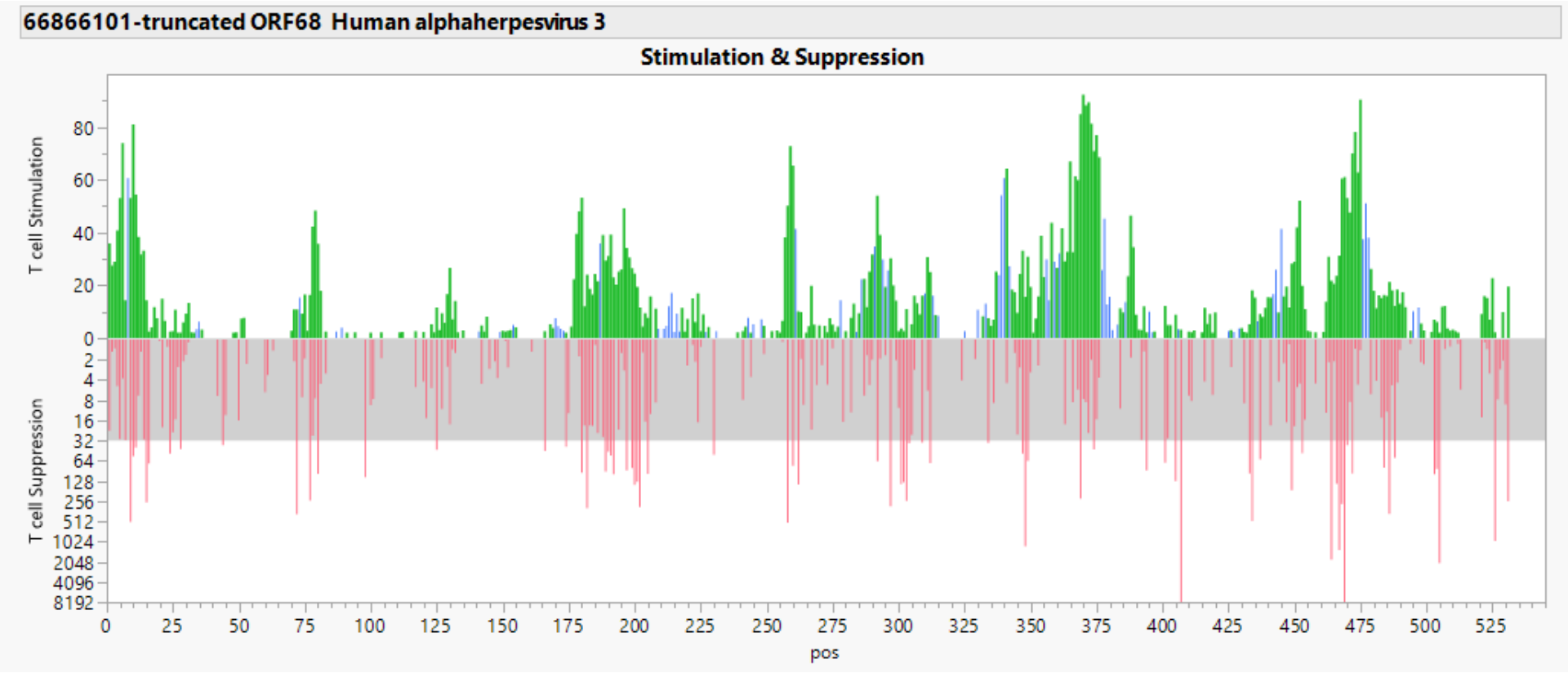

66866101-truncated ORF68 Human alphaherpesvirus 3;t Kernels MHC II @ 15-20 aa by hCAT_S
kSnDPA1_0103-DPB1_0201
kSnDPA1_0201-DPB1_0101
kSnDPA1_0201-DPB1_0501
kSnDPA1_0301-DPB1_0401
kSnDPA1_0301-DPB1_0402
kSnDQA1_0101-DQB1_0501
kSnDQA1_0102-DQB1_0602
kSnDQA1_0301-DQB1_0302
kSnDQA1_0401-DQB1_0402
kSnDQA1_0501-DQB1_0201
kSnDQA1_0501-DQB1_0301
kSnDRB1_0101
kSnDRB1_0301
kSnDRB1_0401
kSnDRB1_0404
kSnDRB1_0405
kSnDRB1_0701
kSnDRB1_0802
kSnDRB1_0901
kSnDRB1_1101
kSnDRB1_1201
kSnDRB1_1302
kSnDRB1_1501
kSnDRB3_0101
kSnDRB3_0202
kSnDRB4_0101
kSnDRB5_0101
kSnH-2-IAb
kSnH-2-IAd
kSnH-2-IAg7
0.00
0.03
0.06
0.09
0.12
0.15
0.92
1.69
2.46
3.23
4.00


Hierarchical Cluster MHC I
Method = Ward
66866101-truncated ORF68 Human alphaherpesvirus 3;t_T_MHC_I
nA_0101
nA_3001
nB_5401
nH-2-Ld
nA_0301
nA_1101
nA_3002
nA_6801
nA_3101
nA_3301
nA_0201
nA_6802
nA_6901
nA_0202
nA_0203
nA_0206
nA_3201
nH-2-Dd
nB_0702
nB_5101
nH-2-Db
nB_0801
nH-2-Kb
nA_2301
nA_2402
nA_2403
nH-2-Kd
nB_1503
nB_2705
nA_2601
nB_1501
nA_2902
nB_3501
nB_5301
nB_5701
nB_5801
nB_1801
nB_4403
nB_4402
nB_4001
nB_4002
nH-2-Kk
nB_4501
-4.36
-3.49
-2.62
-1.75
-0.87
0.00
1.16
2.31
3.47
4.62
5.78
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


Hierarchical Cluster MHC I
Method = Ward
66866101-truncated ORF68 Human alphaherpesvirus 3;t MHC I & Cathepsins consolidated 2
nA_0101
nA_0201
nH-2-Kb
nA_3002
nB_5401
nA_6802
nA_6901
nB_0702
nB_5101
nB_0801
nH-2-Db
nA_3001
nA_0202
nA_0203
nA_0206
nA_0301
nA_1101
nA_3101
nA_2902
nA_3301
nA_6801
nA_2301
nA_2403
nB_2705
nH-2-Kd
nA_2402
nA_3201
nB_1503
nH-2-Ld
nA_2601
nB_1501
nH-2-Dd
nB_3501
nB_5701
nB_5801
nB_5301
nB_1801
nB_4402
nB_4001
nB_4403
nB_4002
nH-2-Kk
nB_4501
0.00
0.02
0.04
0.05
0.07
0.09
0.67
1.25
1.84
2.42
3.00
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 9-mer


MHC_I & MHC_II Cross Presentation Tally @ -1 Sigma: 66866101-truncated ORF68 Human alphaherpesvirus 3;t
Label
xDRB1_0101
xDRB1_0301
xDRB1_0401
xDRB1_0404
xDRB1_0405
xDRB1_0701
xDRB1_0802
xDRB1_0901
xDRB1_1101
xDRB1_1201
xDRB1_1302
xDRB1_1501
xDRB3_0101
xDRB3_0202
xDRB4_0101
xDRB5_0101
xDPA1_0103-DPB1_0201
xDPA1_0201-DPB1_0101
xDPA1_0201-DPB1_0501
xDPA1_0301-DPB1_0401
xDPA1_0301-DPB1_0402
xDQA1_0101-DQB1_0501
xDQA1_0102-DQB1_0602
xDQA1_0301-DQB1_0302
xDQA1_0401-DQB1_0402
xDQA1_0501-DQB1_0201
xDQA1_0501-DQB1_0301
xH-2-IAb
xH-2-IAd
xH-2-IAg7
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
N-term Amino Acid of 15-mer
0
1.3051
2.6103
3.9154
5.2206
6.5257
22.421
38.315
54.21
70.105
86


66866101-truncated ORF68 Human alphaherpesvirus 3
TCEM Frequency Classification
FC TCEM I HC + LC
FC TCEM IIa HC + LC
FC TCEM IIb HC + LC
10
0
10
0
10.0
0.0
0
25
50
75
100
125
150
175
200
225
250
275
300
325
350
375
400
425
450
475
500
525
pos


66866101-truncated ORF68 Human alphaherpesvirus 3
Stimulation & Suppression
T cell Stimulation
T cell Suppression
pos