AWARD NUMBER: **W81XWH-18-1-0801**

TITLE: **Bioprospecting for Antiviral Peptides**

PRINCIPAL INVESTIGATOR: **Barney Bishop**

CONTRACTING ORGANIZATION: **George Mason University**
Fairfax, VA 22030-4422

REPORT DATE: **FEBRUARY 2021**

TYPE OF REPORT: **Final**

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release; Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE
FEB 2021

2. REPORT TYPE
Final

3. DATES COVERED
09/30/2018–10/31/2020

4. TITLE AND SUBTITLE

Bioprospecting for Antiviral Peptides

5a. CONTRACT NUMBER
W81XWH-18-1-0801

5b. GRANT NUMBER
204653

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Barney Bishop
Kylene Kehn-Hall

E-Mail: bbishop1@gmu.edu and kkehnhall@vt.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

GEORGE MASON UNIVERSITY
Karen Cohn
4400 University Dr.
Fairfax, VA 22030-4422

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Mosquito-borne alphaviruses (e.g. chikungunya virus and Venezuelan equine encephalitis virus-VEEV) are of great concern due to their potential to cause severe acute human disease. The absence of available vaccines or therapeutic agents for almost all arboviruses makes development of effective treatment options a significant national security and public health issue. Here, we focus on the study of VEEV as a model arbovirus. Cationic antimicrobial peptides (CAMPs) of innate immunity represent a promising resource for the development of new antiviral therapeutics. However, there is need for a system specifically designed for purpose of discovering AMPs with antiviral activity. We aim to develop a bioprospecting-inspired process for the identification of antiviral peptides against VEEV, using an integrated workflow that combines novel antiviral peptide harvesting microparticles, advanced mass spectrometry, and data analysis tools in order to establish peptide sequences and identify those with potential antiviral properties. We will focus on alligators and snakes CAMPs, since both have been suggested as potential arbovirus reservoir species.

15. SUBJECT TERMS

Reptile, Venezuelan equine encephalitis virus, innate immunity, cytokine, chemokines, host defense peptides, bioprospecting, antiviral, and antiviral peptide discovery.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	52	**19b. TELEPHONE NUMBER** *(include area code)*

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

TABLE OF CONTENTS

1. INTRODUCTION: *Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.*

The rapid, worldwide spread of arthropod-borne viruses (arboviruses) that were previously limited to remote tropical and sub-tropical regions, including mosquito-borne alphaviruses (e.g. chikungunya virus and Venezuelan equine encephalitis virus-VEEV), is of great concern due to their potential to cause severe acute human disease. The absence of FDA approved vaccines or therapeutic agents for almost all arboviruses makes development of effective treatment options for this class of pathogens a significant national security and public health issue. Here we focus on the study of VEEV as a model arbovirus. Cationic antimicrobial peptides (CAMPs) of innate immunity represent a promising resource for the development of new antiviral therapeutics. There is a need to develop a system for the specific purpose of discovering peptides with antiviral activity. We have aimed to develop a bioprospecting-inspired process for the identification of antiviral peptides against VEEV, using novel antiviral peptide harvesting microparticles. These particles incorporate elements from healthy host cells, infected host cells and virions in order to enable the identification of peptides that preferentially target infected cells and virus over healthy host cells. We have focused our efforts on identifying antiviral CAMPs from alligators and snakes, since these species have been suggested as potential arbovirus reservoir species. Our process employs an integrated workflow that begins with the preferential enrichment of CAMPs that target viral elements from reptile serum or plasma, followed by advanced mass spectrometry and data analysis in order to establish their sequences and identify those peptides with potential antiviral properties.

2. KEYWORDS: *Provide a brief list of keywords (limit to 20 words).*

Reptile, Venezuelan equine encephalitis virus, innate immunity, cytokine, chemokines, host defense peptides, bioprospecting, machine learning, antiviral, and antiviral peptide discovery.

3. ACCOMPLISHMENTS:

What were the major goals of the project?

Specific Aim 1: Develop microparticles for capturing antiviral peptides.
Major Task 1: Develop VLP-based capture particles (12 months: 09/30/18-09/29/19)
<u>Milestones:</u>
1.) Reproducible protocol for synthesizing and purifying PP2MA or similar click-chemistry-compatible monomer. (delivery date: 03/29/19 – completed)
2.) Preparations of VEEV virus like particles. (delivery date: 03/29/19 – completed)
3.) Protocol for preparing VLP-modified hydrogel particles using click chemistry and preliminary harvesting data for the particles. (delivery date: 09/29/19 – completed)
4.) GMU IACUC Approval (delivery date: 03/29/19 – completed)
5.) ACURO Approval (delivery date: 05/29/19 – completed)
Major Task 2: Develop cell-membrane based capture particles (9 months: 09/30/18-06/29/19)
<u>Milestones:</u>
1.) Protocols for preparing and assessing hydrogel particles encapsulated in membranes from healthy host cells, with preliminary harvesting data. (delivery date: 06/29/19 – completed)

2.) Protocols for preparing and assessing hydrogel particles encapsulated in membranes from infected host cells, with preliminary harvest data. (delivery date: 06/29/19 – completed)

Specific Aim 2: Capture, analysis and identification of peptides from reptile plasma.

Major Task 3: Perform plasma harvests (12 months: 03/30/19-03/29/20)

Milestones:

1.) VEEV challenge conditions for stimulating release/production of peptides by reptile blood cells. (delivery date: 09/29/19 - completed)

2.) Harvest protocol and conditions that afford efficient capture and identification of known model antiviral CAMPs from plasma. (delivery date: 09/29/19 – completed)

3.) Harvests performed from stimulated and unstimulated plasma using panel of particles developed in Aim 1 for the purpose of identifying antiviral peptides. (delivery date: 03/29/20 – completed)

Major Task 4: Analyze harvested peptides (12 months: 03/30/19-03/29/20)

Milestones:

1.) Workflow for efficiently analyzing harvest samples via tandem mass spectrometry and determination of the peptide sequences. (delivery date: 09/29/19 – completed)

2.) Mass spectrometry data and sequences from peptides harvested for the purpose of identifying novel antiviral peptides. (delivery date: 03/29/20 – completed)

3.) Statistical methods for analyzing harvested peptide sequences to afford statistical significance values of potential antiviral peptides. (delivery date: 03/29/20 – completed)

4.) Predictive model(s) based on machine learning for analyzing harvested peptides to aid in identification of peptides likely to have antiviral properties. (delivery date: 03/29/20 – completed)

Major Task 5: Assess performance of likely antiviral peptides (6 months: 09/30/19-03/29/20)

Milestones:

1.) One or more novel alligator or snake peptides that exhibit antiviral properties against VEEV. (delivery date: 03/29/20 – completed).

What was accomplished under these goals?

Specific Aim 1: Develop microparticles for capturing antiviral peptides.

Major Task 1: Develop VLP-based capture particles (12 months: 09/30/18-09/29/19)

Technical challenges associated with quantifying propargyl monomer incorporation in particles and in preparing suitable amounts of VLP for incorporation in virus-modified particles resulted in minor delays in the development of virus-particle coupling conditions and delivering Major Task 1 deliverables by 09/29/19 as originally projected in the SOW. Both of these challenges were overcome and virus-modified particles were produced, and they did not impact our ability to achieve the final goals of the project.

<u>1.) Preparation of propargyl-PEG monomer:</u> Incorporation of propargyl-PEG monomers in particle formulations allows chemoselective attachment of VEEV or VEEV VLP's to the particles under mild conditions via Click chemistry. Towards this end, we successfully synthesized and purified propargyl-PEG-methacrylamide and propargyl-PEG-methacrylate derivatives. We decided to focus our efforts on the methacrylate derivative propargyl-polyethylene-glycol-(4)-methacrylate (PP4ME), because the PEG-4 linker places the propargyl group further from the polymer backbone than is possible with shorter linkers. Initially we explored the use of flash chromatography to purify the completed

monomer, but we found that using Dowex anion-exchange resin provided a more efficient and scalable purification strategy. As seen in the NMR spectrum below, this approach affords very pure PP4ME (Figure 1). Accordingly, we developed a robust process for the synthesis and purification of PP4ME in amounts suitable to support the preparation of particle incorporating the monomer. We have used this process to prepare multiple batches of the PP4ME monomer.

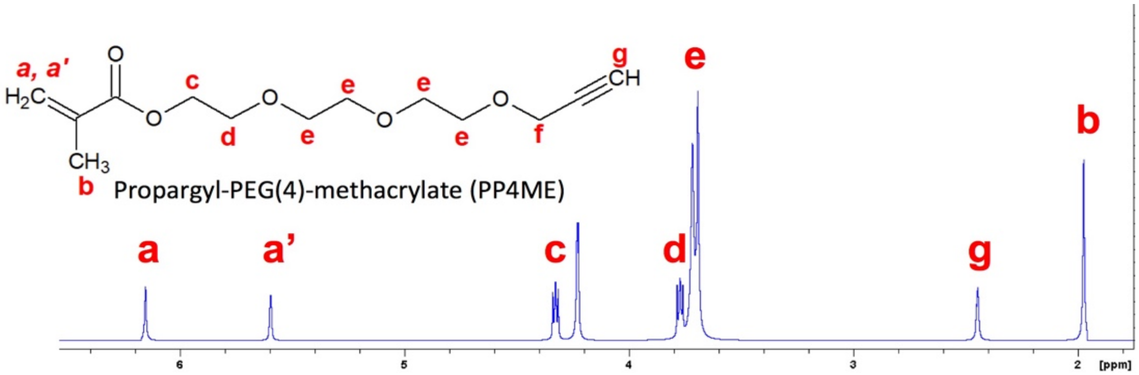


Figure 1. Assigned NMR spectra for PP4ME monomer establishing that the reaction was successful and purification scheme yielded very pure PP4ME.

<u>2.) Preparation of particles containing propargyl-PEG-monomer (PP4ME)</u>:

The **first generation** of VLP-based capture particles utilized core-shell microparticles incorporating PP4ME that were generated using RAFT polymerization. These particles consisted of an inert cross-linked tert-butylmethacrylamide (TBMAm) core with shells based on cross-linked copolymers of *N*-methacryloyl-6-aminohexanoic acid (MA6AHA) and PP4ME. We generated a series of particles that incorporated 10% PP4ME and 20, 10, 5 and 0% cross-linking with MA6AHA providing being the main component in their shells (N4.3, 4.4, 4.5, and 4.6 respectively) in order to identify shell architectures/formulations that are suitable for modification with virus (See Table 1). Analysis of the particles via NMR showed incorporation of the monomers with levels of PP4ME in the particles relative to MA6AHA being somewhat lower than they were in the feed stock. The sizes of the N4.3 (20% cross-linking : 10% PP4ME) and N4.4 (10% cross-linking : 10% PP4ME) particles, determined via dynamic light scattering (DLS) were consistent with the sizes we have seen with other particles with MA6AHA-based shells (Table 2).

A **second generation** of VLP-based particles were prepared based on core-shell hydrogel particles that combined inert cross-linked tert-butylmethacrylamide (TBMAm) cores and shells consisting of cross-linked copolymers combining hydroxypropylmethacrylamide (HPMA) with 10-20% PP4ME (with some also incorporating low amounts of MA6AHA in the shell formulation). These hydrogel particles were developed in order to reduce background peptide capture associated with the hydrogel particle component of the VLP-modified particles. The presence of peptides captured by the hydrogel particle of the VLP-modified particles could complicate the detection of peptides captured specifically by the VLP baits. Literature reports describe the use of HPMA in anti-fouling coatings (Langmuir, 2010 Nov 16;26(22):17375-82.). Furthermore, HPMA is a neutral polar monomer in contrast to MA6AHA, which is anionic, and thus HPMA-based hydrogel particles should exhibit different harvesting properties from the general harvest particles.

Particle		N4.3	N4.4	N4.5	N4.6	N4.9	N4.7	N4.8	N4.10
Monomers	**PP4ME**	10.0%	9.0%	10.0%	10.0%	10.0%	X	10.0%	25.0%
	MA6AHA	70.0%	81.0%	85.0%	90.0%	70.0%	X	X	10.0%
	BIS	20.0%	10.0%	5.0%	0.0%	20.0%	10.0%	10.0%	5.0%
	HPMA	X	X	X	X	X	90.0%	80.0%	60.0%

Table 1. Series of core-shell particles (1st and 2nd generation) produced for development of virus-modified particles.

DLS	Short Name		**N4.3**	**N4.4**	**N4.5**	**N4.6**	**N4.8**	**N4.8**	**N4.10**	**N4 (Core)**
	Z-Average	(d.nm)	294.3	302.2	ND	ND	356.1	389.6	322.2	215
	PdI		0.052	0.129	ND	ND	0.215	0.098	0.232	0.044

Table 2. Summary of DLS information indicating approximate size of particles. The 5% and 0% BIS crosslinking have difficulty being sized via DLS and are to be determined (TBD).

3.1 Preparations of VEEV virus like particles: We developed a semi-quantitative western blot methodology to quantitate our VLP preparations. For this we serially diluted VEEV TC83 (pfu/ml) and performed western blot analysis followed by densitometry analysis (Figure 2). Our standard curve gave us an R^2 value of 0.9852. This method allowed us to compare the amount of glycoprotein (GP) present in our VLPs to that found in our viral stocks; thereby giving us a semi-quantitative determination of the amount of particles present in our VLP preparations.

We successfully produced VEEV VLPs by transfecting 293T cells with the VEEV structural polyprotein plasmid (Figure 3). Viral E2 and E1 proteins (the major constituent of the VLPs) are detected in cellular lysates as well as in supernatants, indicating that VLPs are being released from these cells. Experiments determined that production of VLPs in the presence of fetal bovine serum (FBS) interfered with detection of our VLPs. VLPs were found to enter cells as determined through fluorescent microscopy (data not shown). We attempted to quantitate our VLPs through flow cytometry, but this approach was not successful likely due to the non-specific binding of antibodies to the negative control nanoparticles. As an alternative, VLPs were quantitated using semi-quantitative western blot analysis via densitometry measurements as described above.

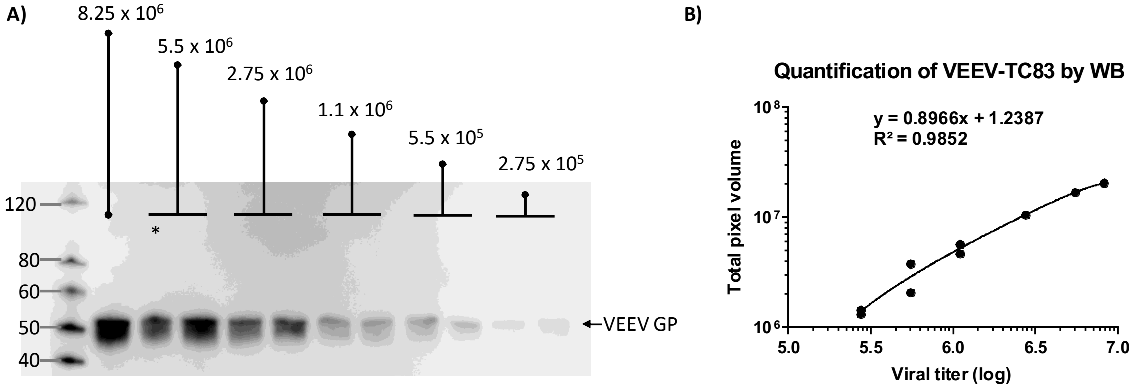


Figure 2: Semi-quantitative western blot analysis of GP concentrations. A) Western blot (WB) of serially diluted VEEV TC83 (pfu/ml). WB was performed with anti-VEEV glycoprotein antibody. B) Densitometry analysis was performed of the WB image and total pixel volume plotted vs. the viral titer (log pfu/ml).

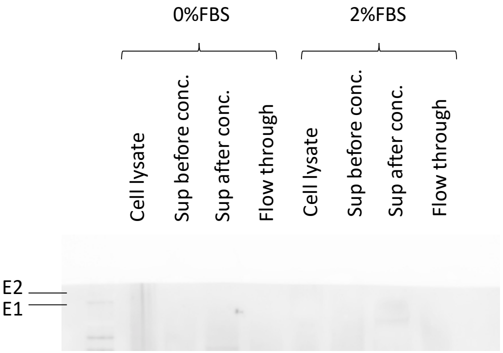


Figure 3: 293T cells were transfected with VEEV structural polyprotein in the presence of 0% or 2% FBS. Cell lysates and cell free supernatants (containing VLPs) were harvested at 48 hours post-transfection. Supernatants were concentrated using an Ambicon concentrator. Samples were separated via SDS-PAGE followed by western blot analysis using antibodies specific for VEEV E1 and E2 glycoproteins.

<u>3.1. Preparation VEEV TC-83 for Particle Incorporation:</u> There are some limitations of producing VLPs through transfection including the large expense of the transfection reagents and the difficulty in scaling-up production to enable large scale coupling of VLPs to the microparticles. To address these issues, we used VEEV TC-83 as a model to work on coupling VLP-based capture particles.

We optimized a sucrose-cushion based purification method for virus purification. The method uses 20-50% sucrose-cushion where following ultracentrifugation, VEEV TC-83 can be found at the interface between these sucrose layers. Analysis of fractions from the sucrose cushion indicated that VEEV did accumulate in the lower fractions, F11 and F12 as expected (Figure 4). This is indicated by VEEV glycoprotein detection (panel A) and infectious virus detected by plaque assays (panel B).

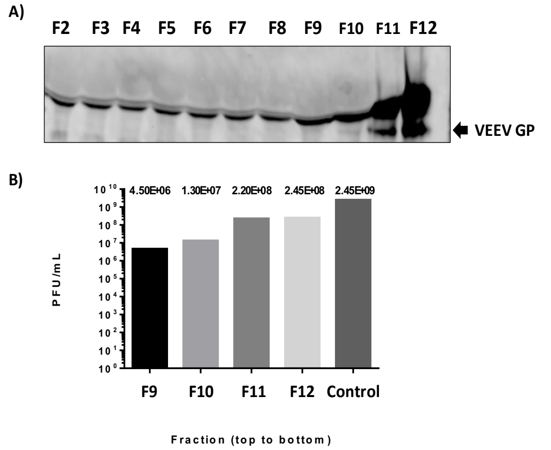


Figure 4: Sucrose-cushion purification of VEEV. 293T cells were infected with VEEV and supernatants collected at 48 hours post-infection. Supernatants were layered on top of a 20-50% sucrose cushion and centrifuged at 36,000 RPM for 1.5 hours at 4^0C. One mL fractions were collected from the top and an aliquot separated via SDS-PAGE followed by western blot analysis using antibodies specific for VEEV glycoproteins (panel A). Upper non-specific band is FBS. Fractions F9-F12 were also subjected to plaque assays to confirm that the lower fractions had infectious virus and not just glycoproteins (panel B).

<u>4. Preparation of VLP/Virus-modified Particles:</u> An initial pilot experiment was performed to determine our ability to produce VLP-particles using sucrose-cushion purified VEEV TC-83. Following the click-chemistry reaction, samples were analyzed by RT-qPCR and western blot analysis. Initial PEG modified VEEV (VLP-PEG Azide) had high levels of both viral RNA and VEEV GP present (Figure 5). However, little viral RNA and no VEEV GP was detected in the VLP-

particles (ssCTA-1 and N4.3-2) and any of the particle washes, indicating that the click-chemistry reaction was not optimal.

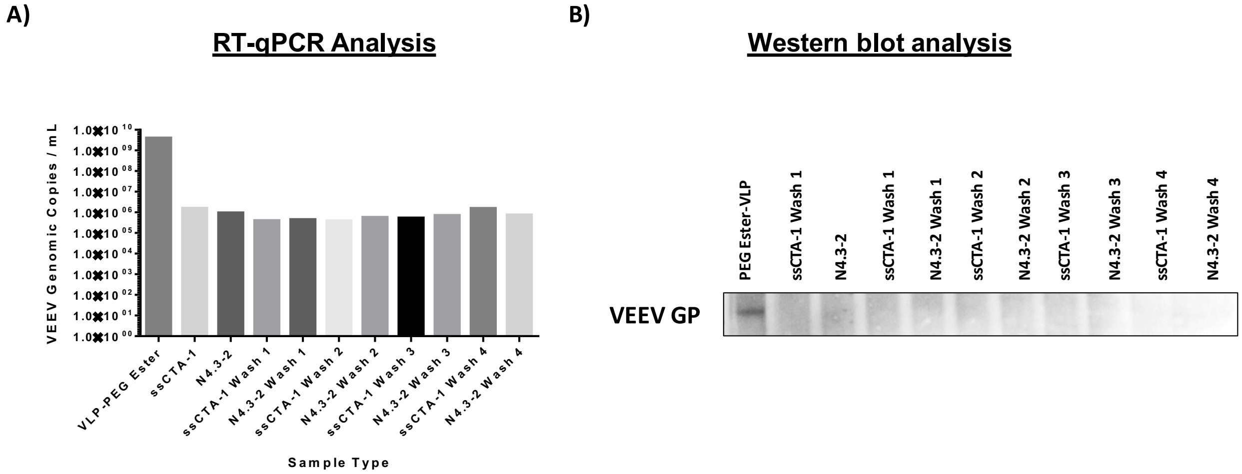


Figure 5: Pilot Study #1 on VLP-particle production. Sucrose-cushion purified VEEV TC-83 had PEG added and click-chemistry performed to link the PEG Azide modified virus to the particles. An aliquot of particles and washes were analyzed by RT-qPCR (panel A). An aliquot of particles and washes were also separated via SDS-PAGE followed by western blot analysis using antibodies specific for VEEV glycoproteins (panel B). In figure labels – "VLP-PEG Ester" and "PEG-Ester-VLP" refer to PEG Azide modified VEEV

A second pilot experiment was performed where a number of conditions were altered to help increase the coupling efficiency. All reactions were performed at room temperature (as opposed to 4^{0}C in the original experiment), the amount of PEG Azide used in modifying the virus was increase by 10-fold, and the particle concentration used in the subsequent click coupling reaction was increased from 1 mg of particles per reaction volume to 5 mg of particles per reaction volume. In addition, the sucrose was removed by dialysis in PBS prior to the addition of PEG Azide to the virus. There was detectable viral RNA and VEEV GP in the PEG-modified VEEV (Figure 6). Two different particles were used in this study (N4.4-3 and N4.3) which contained different amounts of cross-linking. Particle N4.4-3 had the greatest amount of viral RNA present with some VEEV GP detectable (albeit more of a smear on the gel). The results suggested that some virus had been coupled to the particles; however additional optimization of the process was needed to enable more stable virus incorporation.

DLS	Short Name		**N5.1**	**N5.2**	**N5 (Core)**
	Z-Average	(d.nm)	301.0	266.7	256.5
	PdI		0.245	0.022	0.110

Table 3. Summary of DLS information indicating approximate size of particles based on N5 core. N5 core refers to a second preparation of the core particles with identical formulation to that of the preceding N4 core particles.

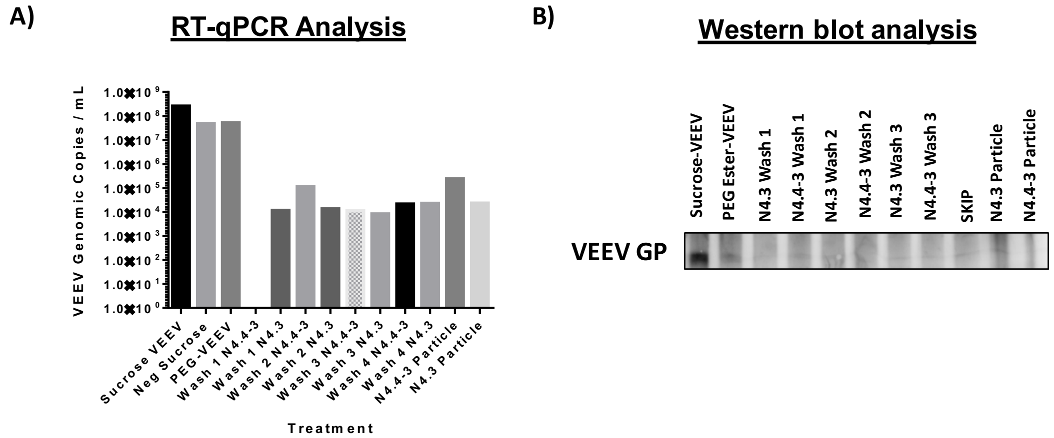


Figure 6: Pilot Study #2 on VLP-particle production. Sucrose-cushion purified VEEV TC-83 had PEG added and click-chemistry performed to link the PEG Azide modified virus to the particles. An aliquot of particles and washes were analyzed by RT-qPCR (panel A). Neg sucrose = sample after dialysis to remove the sucrose. An aliquot of particles and washes were also separated via SDS-PAGE followed by western blot analysis using antibodies specific for VEEV glycoproteins (panel B). In figure labels – "PEG-VEEV" and "PEG Ester-VEEV" both refer to PEG Azide modified VEEV

5. Perform Harvests using VLP/Virus-Modified Particles: Peptides were harvested from alligator plasma using the two-stage harvesting process to assess their harvesting properties and develop harvesting protocols. General harvests from healthy alligator plasma were performed in preparation for harvests using cell-membrane based capture particles. The first round of harvesting (the General Harvest) was performed using core-shell particles with cross-linked MA6AHA shells, and captured peptides were eluted using 1:1:0.1 TFE/H_2O/TFA. The resulting peptide eluents were aliquoted, and solvent then removed using a speed-vac. A second round of harvesting was then performed using both the virus modified and unmodified N4.9 and N4.8 particles, with harvests using each particle type being done in duplicate. The captured peptides were eluted from the particles and analyzed via mass spectrometry using the Thermo Fusion Tribrid mass spectrometer at GMU using our standard parameters for peptidomic analyses. The resulting mass spectrometry data (focusing on MS/MS data) were processed using PEAKS *de novo* sequencing software and our Python script in order to determine and organize the sequences of peptides captured in each harvest. The sequence and additional mass spectrometry data were transferred to the team at PNNL for statistical analyses.

6. Third generation of VLP/Virus-Modified Particles: We prepared a new generation of core-shell particles as a framework for constructing virus-modified particles. The formulations for these particles were designed to maximize virus loading. Both of these particles were built upon the same cross-linked N-*tert*-butylmethacrylamide polymer core. They differed in the architecture of the outer harvesting layer(s). The outer harvesting layers of both consisted of cross-linked HPMA-based copolymers that include MA6AHA and PP4ME in their formulations. Where they differed was the distribution of the co-monomers. In one case, N5.1 particles, the outer harvesting layer consisted of a random co-mixture of the different monomers. The N5.2 particles had a more complex architecture where PP4ME monomer was added later in the synthesis of the harvesting layer, resulting in a higher concentration of the monomer near the outer surface of the particle. This arrangement was intended to make the PP4ME more accessible for cross-linking with virus. These particles were coupled to

purified VEEV TC-83 as described above. VEEV coupling to the particles was assessed by RT-qPCR analysis (**Figure 7**). Particle N5.2 displayed the highest amount of bound virus (10^7 genomic copies/mL). However, after 3 washes, virus release by the particles was still observed, suggesting that a percentage of the VEEV virions were noncovalently trapped by the particles or that covalently bound virions were being released under wash conditions.

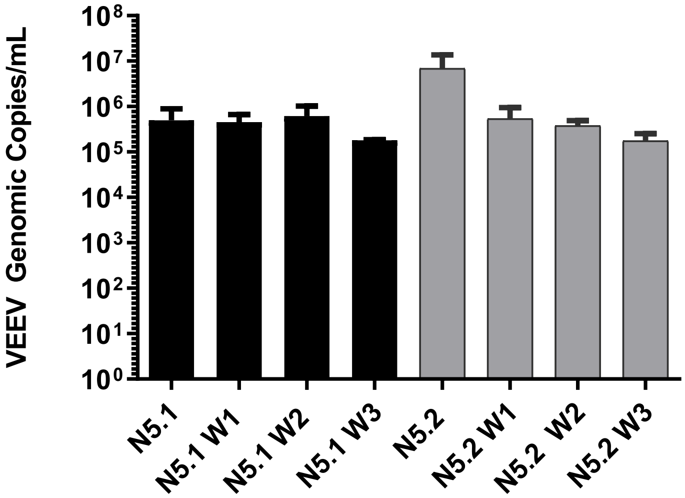


Figure 7: Third generation of VLP/Virus-Modified Particles. Sucrose-cushion purified VEEV TC-83 had PEG added and click-chemistry performed to link the PEG Azide modified virus to 2 different particle types, N5.1 and N5.2. An aliquot of particles and washes were analyzed by RT-qPCR. W1=wash 1, W2=wash 2, W3=wash 3

Major Task 2: Develop cell-membrane based capture particles (9 months: 09/30/18-06/29/19)

1.) <u>Protocols for preparing and assessing hydrogel particles encapsulated in membranes from healthy host cells, with preliminary harvesting data.</u>

2.) <u>Protocols for preparing and assessing hydrogel particles encapsulated in membranes from infected host cells, with preliminary harvesting data.</u>

Particles encapsulated in healthy and infected cell membranes were prepared and assessed simultaneously.

A.) In the first quarter, we synthesized two series of core-shell hydrogel particles with functionalized cross-linked polymer shells. These particles explored reaction conditions used to generate the particles (heating via oil bath and hot plate vs water and a circulating immersion heater) as well as shell formulation (MA6AHA vs. N-isopropylmethacrylamide (NIPMAM) and methacrylic acid (MAAc) copolymers). Replicate batches of particles were produced in order to establish reproducibility of particle syntheses and resulting particle physical properties. Based on these studies, it was determined that particles with harvesting outer shells based on cross-linked poly-MA6AHA afforded the most consistent performance. Therefore, we elected to focus on these particles for use in first round general harvests.

B.) We tested 4 different plasma membrane preparation protocols to find the one that was optimal for plasma membrane isolation from U937 cells. CD45 was used as a marker of successful plasma membrane isolation (data not shown). Based on CD45 expression in our preparations, we selected a method whereby U937 cells are incubated with a hypotonic buffer (50mM Tris-HCl, 5mM EDTA, 100mM NaCl, 2mM MgCl, 3mM CaCl2, 5mM KCl, and protease inhibitor cocktail) for 10 minutes on ice. Cells are then disrupted with a dounce glass homogenizer and centrifuged at 4ºC at 400g for 10 minutes. The pellet is discarded and supernatant containing plasma membranes is saved for analysis.

To determine the purity of our preparations we performed western blot analysis with other cellular organelle markers (Figure 8A). Cadherin and CD45 (both plasma membrane markers) were found in our preparations, with detection enhanced by not subjecting the reduced protein to boiling (lanes 1-4). Boiling can result in protein aggregation and sticking to tubes, thus resulting in loss of protein. KPNB1 (nuclear protein) and Calnexin (ER protein) were not detected in our plasma membrane preparations (lanes 4-8). However, we were able to detect Calnexin and KPNB1 in U937 whole cell extracts via western blot (Figure 8B). These results demonstrated our ability to isolate plasma membranes from U937 cells.

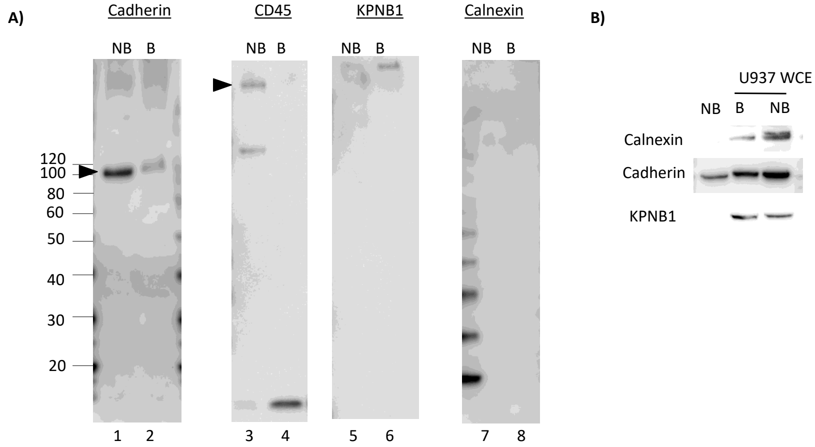
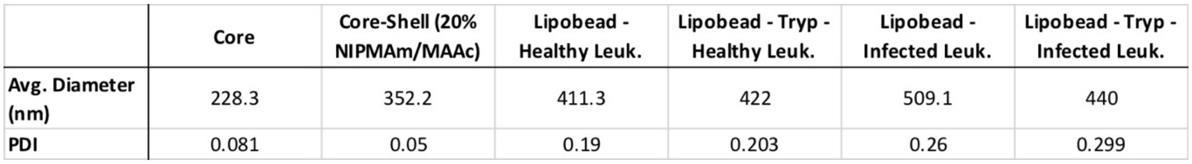


Figure 8: Plasma membrane preparation purity analysis. A) Plasma membranes were isolated from U937 cells and purity assessed by western blot analysis using anti-Cadherin, anti-CD45, anti-KPNB1 and anti-Calnexin antibodies. NB: no boiling; B: boiling at 95ºC B) Western blotting was performed with plasma membrane preps (lane 1) and U937 whole cell extracts (WCE). NB: no boiling; B: boiling at 95ºC

C.) Hydrophobically modified NIPMAm/MAAc particles (cross-linked NIPMAm/MAAc copolymer outer shells) were prepared and encapsulated within leukocyte membranes (lipobeads). In generating these lipobeads, leukocyte cytoplasmic membranes were passed through an extruder to generate liposomes 100 nm in diameter. The liposomes were then incubated overnight with the hydrophobically modified particles at 4°C. The lipobeads that were formed were then pelleted by centrifugation and the supernatant collected in order to remove any free liposomes remaining in solution. The particles were suspended in buffer and the process repeated so as to remove any remaining unbound liposomes. This process was used to generate lipobeads encapsulated in membranes from healthy leukocytes (Lipobead-Healthy Leuk.), healthy cells that had been treated with trypsin (Lipobead-Tryp-Healthy Leuk.), VEEV-infected leukocytes (Lipobead-Infected Leuk.) and infected cells that had been treated with trypsin (Lipobead-Tryp-Infected Leuk.). Light scattering data (sizes) for particles are provided below in Table 4. Results from initial harvests from commercial Alligator plasma are provided below in Aim 2 data below.



	Core	Core-Shell (20% NIPMAm/MAAc)	Lipobead - Healthy Leuk.	Lipobead - Tryp - Healthy Leuk.	Lipobead - Infected Leuk.	Lipobead - Tryp - Infected Leuk.
Avg. Diameter (nm)	228.3	352.2	411.3	422	509.1	440
PDI	0.081	0.05	0.19	0.203	0.26	0.299

Table 4. Comparison of Particle Sizes: The diameters (nm) of the core particles, core-shell particles and varied lipobeads as determined through light scattering are provided.

Specific Aim 2: Capture, analysis and identification of peptides from reptile plasma.

Major Task 3: Perform plasma harvests (12 months: 03/30/19-03/29/20)

<u>VEEV challenge conditions for stimulating release/production of peptides by reptile blood cells.</u>

Efforts to optimize conditions for stimulating reptile blood cells with VEEV took longer than originally projected due to complications in scheduling collection of blood samples due to Hurricane Dorian, which had been projected in different forecasts to hit Florida where our collaborators at the St. Augustine Alligator Farm Zoological Park are located. Their need to prepare for the potential storm to protect the animals disrupted our schedule for getting blood in September, which in turn delayed our efforts to optimize stimulation conditions beyond our original projected 09/29/12019 deliver date. That being said, we received a fresh shipment of alligator blood on 10/23/2019, PBMCs were collected and stimulation studies initiated. We had established conditions and a protocol for stimulating reptile blood with VEEV by the end of October, completing this Task 3 deliverable.

Pilot experiments were performed to determine the optimal conditions for stimulating antiviral peptides in reptile blood. First, primers were designed to enable RT-qPCR analysis of innate immune response genes of the American Alligator. OASL, CXCL10, Mx1, and ISG20 were selected based on a published study indicating these genes were altered in Crocodile cells in culture following poly I:C stimulation (PMID: 29213275). American Alligator blood was incubated without or with VEEV TC83 at 1:10, 1:100, 1:1000 dilutions. Samples were collected at 2, 8, and 24 hours post stimulation. RNA was extracted and RT-qPCR analysis performed. Issues with RNA sample quantity were encountered and therefore data regarding the induction of gene expression was inconclusive. However, our data did indicate that our primers were able to successfully detect OASL, ISG20, and Mx1 gene expression with no primer dimers detected (data now shown). Repeated RT-qPCR analysis showed that these genes were not being stimulated by VEEV when exposed in this manner.

Stimulation experiments were repeated, however this time we opted to stimulate peripheral blood mononuclear cells (PBMCs) isolated from American Alligator blood. PBMCs were isolate using a Ficoll gradient and cells stimulated with VEEV TC83 (MOI ~100). Cells were collected at 2, 8 and 24 hours post-VEEV addition and RNA isolated. Mock-infected cells were collected in parallel. Interestingly, we observed a strong induction of Mx1 and OASL at 8 and 24 hours post-VEEV exposure (Figure 9B and C) and CXCL10 only at 24 hours (Figure 9D). No change in ISG20 expression was observed (Figure 9A). These data indicate that innate immune response genes are stimulated in PBMCs exposed to VEEV. For harvesting experiments, we opted to utilize the 2- and 8-hour samples. The 2-hour samples may contain a population of antiviral peptides that are stimulated independent of transcriptional regulation, while the 8-hour samples should include peptides stimulated through transcriptional changes as the innate immune response genes we examined. We believed the 24-hour samples were too late in the time course, with cell death likely being induced, which would confound our analysis.

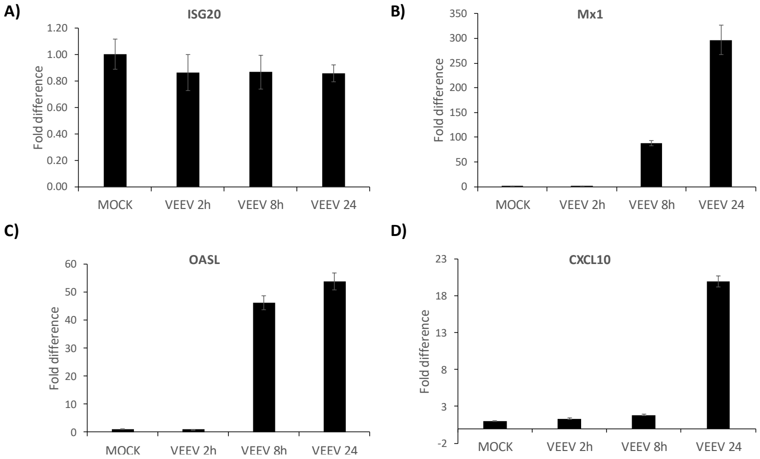


Figure 9: Innate immune response genes are stimulated in American Alligator PBMCs exposed to VEEV. RNA was extracted from PBMCs exposed to VEEV TC-83 (MOI ~100) and RT-qPCR analysis performed using a one-step SYBR Green kit including primers for ISG20, Mx1, OASL, CXCL10 and GAPDH (endogenous control). Gene expression changes were calculated using the ΔΔCt method, with Mock infected cells set to 1.

We also determined if VEEV RNA was present in the PBMCs, which would be indicative of a successful infection. RT-qPCR for viral RNA indicated that VEEV RNA levels increased over time; however, viral RNA levels were quite high at the start of the infection (7 logs) (Figure 10). Therefore, a repeat experiment with lower amounts of VEEV would be required to determine if viral replication is occurring. At this point, we can conclude that VEEV viral RNA persists in American Alligator PBMCs.

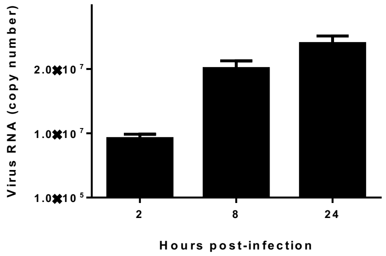


Figure 10: VEEV persists in American Alligator PBMCs. RNA was extracted from PBMCs exposed to VEEV TC-83 (MOI ~100) and RT-qPCR analysis performed using a one-step kit including primers and a probe for VEEV RNA. Virus RNA copy number was determined using absolute quantitation with known amounts of viral RNA used for the standard curve.

Harvest protocol and conditions that afford efficient capture and identification of known model antiviral CAMPs from plasma.

In the third quarter, peptides were harvested from alligator plasma using the two-stage harvesting process to assess their harvesting properties and develop harvesting protocols. The first round of harvesting was performed using core-shell particles with cross-linked MA6AHA shells, and captured peptides were eluted using 1:1:0.1 TFE/H_2O/TFA. The resulting peptide eluents were aliquoted, and solvent then removed using a speed-vac. A second round of harvesting was then performed from these aliquots, using the four different membrane encapsulated particles, by dissolving the dried aliquots in buffered aqueous particle suspension (10 mM Tris, pH 7.4). The harvested peptides were then analyzed by tandem mass spectrometry (LC-MS/MS with ETD and HCD). Harvests were performed in triplicate and duplicate injections of captured peptides were performed on the mass spectrometer. The results of these harvests are presented in Table 5 and Figure 11. These initial harvests yielded promising results.

	General Harvest (20% MA6AHA)	**Lipobead - Healthy Leuk.**	**Lipobead - Tryp - Healthy Leuk.**	**Lipobead - Infected Leuk.**	**Lipobead - Tryp - Infected Leuk.**
No. Unique Database Search Peptide Spectrum Match Sequences	138	19	66	13	26
No. Unique De Novo Only Sequences	569	161	393	66	103
No. Database Peptides Unique to Particle	121	8	53	5	12
No. De Novo Peptides Unique to Particle	566	152	387	61	98

Table 5. Results from initial round of harvests using new particles encapsulated within cell membranes. The initial general harvest using MA6AHA-based particles yielded the greatest number of identified peptide sequences, both those with matches in the reference proteomic database as well as *de novo* sequences. Harvests using membranes from healthy and infected leukocytes yielded the fewest peptide identifications, with harvests using membranes from cells treated with trypsin yielded an intermediate number of peptide identifications. In all cases, the number of de novo only peptides greatly outnumbered those with database matches.

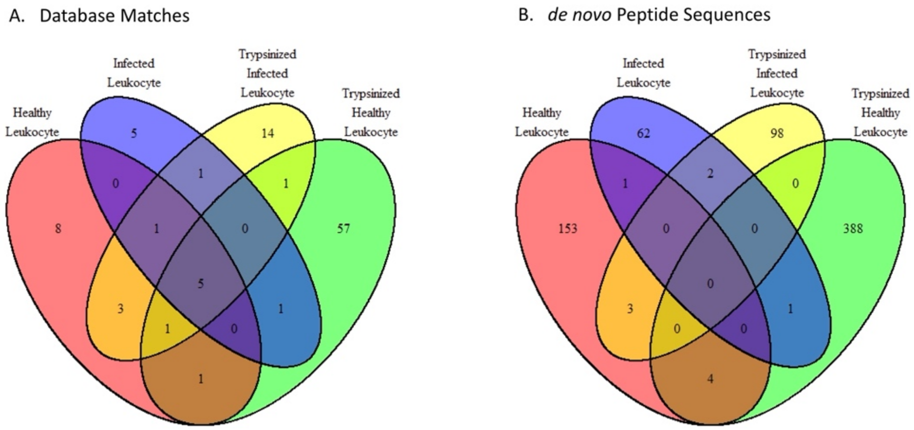


Figure 11. Venn diagrams illustrating distribution of peptides based on the four leukocyte membrane-encapsulated particle types. Diagram A.) peptide sequences with matches in the reference sequence database, and B.) *de novo* sequences with no identified database matches.

We detected reference AMPs in multiple harvests. While the reference peptides were not always detected in direct analyses of the general harvest MA6AHA-based particles, they were detected in subsequent harvests using the membrane-encapsulated particles, which verifies that the reference AMPs are indeed captured and enriched in the first-round harvests. The fact that these peptides were not detected in all of the harvests performed using membrane-encapsulated particles is not surprising as these peptides are anticipated to have a low propensity to bind to host cells.

<u>Harvests performed from stimulated and unstimulated plasma using panel of particles developed in Aim 1 for the purpose of identifying antiviral peptides.</u>

Multiple harvests had been performed and were at various stages of analysis at the time that GMU shut down non COVID-related research in spring 2020. Upon our return to the lab we completed analysis of the data that we had collected. We completed our remaining analyses of alligator plasma and PBMC-secreted peptides as well as analyses of snake (*Boa constrictor*) plasma peptides. Information regarding these harvests are indicated in Table 6 below.

Exp. ID #	Species	Sample	Harvests	Post-Harvest Processing	MS/MS Analysis	Data Analysis	Purpose/Results of study
8	*A. mississippiensis*	Media Supernatant – Control	🟦	🟦	🟦	🟦	Purpose: Identify potential alligator AMPs. Outcome: Afforded varied results. Due to limited access to the mass spectrometer at the time, these studies were not repeated and focused on Exp. ID# 13 and 14.
		Media Supernatant – 2h post VEEV	🟦	🟦	🟦	🟦	
		Media Supernatant – 8h post VEEV	🟦	🟦	🟦	🟦	
9	*A. mississippiensis*	Unstimulated Plasma - EDTA	🟦	🟦	🟦	🟦	Purpose: Evaluate impact of anticoagulant. Outcome: Afforded few peptide identifications complicating analysis. Study was not central to project – opted to focus on other harvests.
		Unstimulated Plasma - Heparin	🟦	🟦	🟦	🟦	
10[1]	*A. mississippiensis*	Unstimulated Plasma –Crazy	🟦	🟦	🟦	🟦	Purpose: Analyze alligator peptides from animal with inflammation. Outcome: Poor peptide identification – Repeated – see 13 below.
11	*A. mississippiensis*	Media Supernatant – 8h post VEEV	🟦	🟦	🟦	🟦	Purpose: Assess parameters and their impact. Outcome: Improvement - led to better results in **13** and **14.**
12	*B. constrictor*	Unstimulated – healthy membranes	🟦	🟦	🟦	🟦	Purpose: Identify potential snake AMPs. Outcome: Low number high confidence peptide identifications – repeated – see 14 below.
		Unstimulated – infected membranes	🟦	🟦	🟦	🟦	
13[1]	*A. mississippiensis*	Unstimulated Plasma –Crazy	🟪	🟪	🟪	🟪	Purpose: Analyze peptides from alligator with inflammation. Outcome: Data presented in **Table 7** and **Figure 12**.
14[2]	*B. constrictor*	Unstimulated – healthy membranes	🟪	🟪	🟪	🟪	Purpose: Identify snake peptides & assess direct harvest and MS/MS parameters. Outcome: Data presented in **Table 9** and **Figure 14**.
		Unstimulated – infected membranes	🟪	🟪	🟪	🟪	

Table 6. Harvests and analyses from third and fourth quarters of year 2, along with their purposes and outcomes. Blue bars indicate harvests whose sequence data were transferred to PNNL for subsequent analysis. Purple bars indicate harvests from experiments that did not require analysis by PNNL.

[1] Blood was collected from a large adult alligator named Crazy that had a large growth on its leg that was being removed by the veterinarian. Studies explore how altering MS/MS parameters on the mass spectrometer influences the number and size of peptides for whom sequences are determined.

[2] Experiment designed to explore direct harvesting from snake plasma using membrane-encapsulated particles and to assess the effect of MS/MS fragmentation parameters on peptide identification.

Results from repeated harvests from American alligator (Crazy) plasma: Blood was received that had been collected from a large male alligator as part of a medical examination. Initial harvests and analyses (Exp ID# 10) yielded poor peptide identification, and were repeated (Exp ID# 13). As described above (Table 6), completion of these analyses was delayed as a result of measures enacted by the university in response to the COVID-19 pandemic. However, they were completed by the end of the extended AIMM project period (with no-cost extension – 10/31/2021). These harvests were performed directly from plasma (unstimulated) using membrane-encapsulated particles with membranes from healthy and VEEV-infected human leukocytes (Table 7). In analyzing the captured peptides, we compared the sequences identified using our standard MS/MS parameters and a method that combined higher resolution HCD and ETD with increased injection times, which was recommended to afford improved MS/MS spectra and more complete fragmentation for analyzing larger peptides (Table 7 and Figure 12). Our interest in larger peptide was because many of the classic antimicrobial peptides that exhibit AVP properties such as defensins and cathelicidins are relatively large peptides (between 20-50 amino acids in length).

Sample Type	Particle Type	Experimental Replicate #	Mass Spec Method 1		Mass Spec Method 2	
			Database Search	De Novo	Database Search	De Novo
Gator plasma, no infection/no incubation ("Crazy") (Redo), reduced/alkylated PRE lipobead	Healthy leukocyte membrane lipobeads	1	11	325	43	18
Gator plasma, no infection/no incubation ("Crazy") (Redo), reduced/alkylated PRE lipobead	Infected leukocyte membrane lipobeads	1	15	141	53	113
Gator plasma, no infection/no incubation ("Crazy") (Redo), reduced/alkylated POST lipobead	Infected leukocyte membrane lipobeads	1	22	249	59	82
Gator plasma, no infection/no incubation ("Crazy") (Redo), reduced/alkylated POST lipobead	Infected leukocyte membrane lipobeads	2	17	230	61	96
Gator plasma, no infection/no incubation ("Crazy") (Redo), no reduction/alkylation	Infected leukocyte membrane lipobeads	1	5	266	99	116
Gator plasma, no infection/no incubation ("Crazy") (Redo), no reduction/alkylation	Infected leukocyte membrane lipobeads	2	10	238	90	136

Table 7. Results from repeated analyses of American alligator plasma. Harvests were performed from alligator plasma using hydrogel particles encapsulated in membranes form healthy and VEEV-infected human leukocytes. Two different sets of parameters were used during MS/MS analyses (Mass Spec Method 1 = original parameters and Mass Spec Method 2 = new parameters). We here report the number of unique peptide sequences determined for each harvest using each Mass Spec Method, differentiating between those with database matches and those that were determined purely de novo.

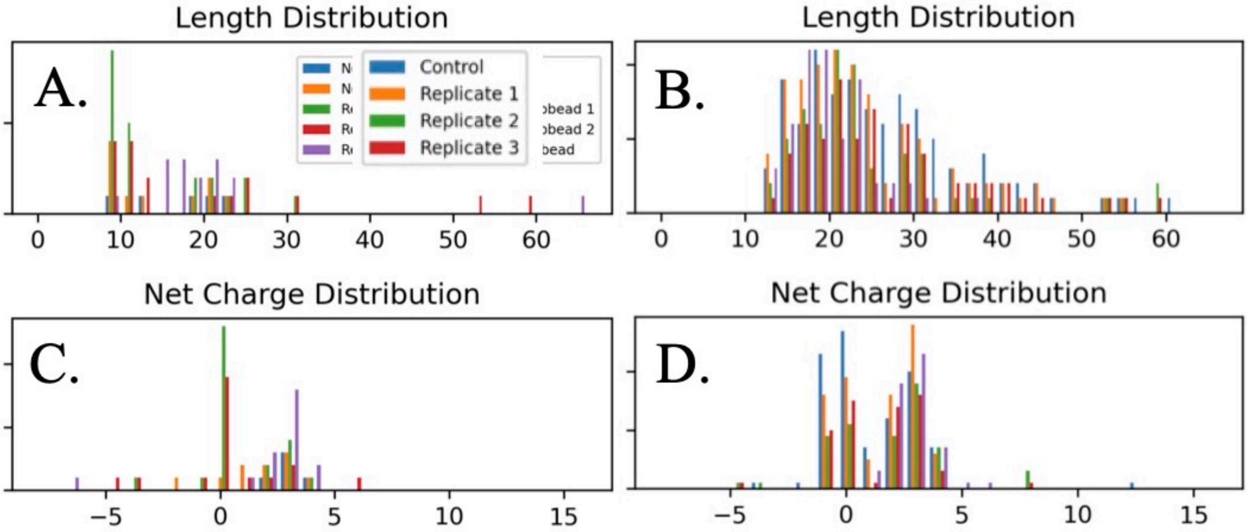


Figure 12 Representative results for comparison of repeated American alligator harvests – Harvests performed using membranes from infected leukocytes. Comparison of peptide lengths (A. and B.) and net charges (C. and D.) determined using the original MS/MS parameters (Method 1: A. and C.) and the newer parameters (Method 2: B. and D.). The new parameters resulted in a significant shift towards larger peptides, with multiple identifications of peptides between 20 and 60 amino acids in length.

Results from analyses of first round Boa constrictor harvests: Due to the low volume of plasma provided (~800 µL), harvests were performed at half normal scale (50µL of plasma) in order to allow experimental replicates for each harvest. Two rounds of harvest were performed. In the second round, harvests were performed directly from plasma diluted in buffer, using particles encapsulated in membranes from healthy and VEEV-infected cell membranes. Peptides captured in the first round of harvests were analyzed via LC-MS/MS immediately before they were shut down in compliance with GMU precautionary measures in response to COVID-19. The second round of snake harvests were analyzed by LC-MS/MS in September 2021, after the instrument was brought back online and the proteomic MS lab returned to operations. The sequences of the peptides from the first round of harvests were determined using PEAKS *de novo* sequencing software version X+ using a reference database consisting of predicted protein sequences from a *B. constrictor* genome assembly and annotation (Card et al., 2019). An additional homology match search was performed using PEAKS to identify peptides similar to database peptides that may have been initially missed in the more stringent database search stage. The results of these initial analyses of Boa constrictor peptides are given below in Table 8.

Particle Type	**Rep. #**	**# Features**	**# DB Search PSMs**	**# DB Search Unique Peptides > 7 aa**	**# *de novo* Spectra**	**# *de novo* Unique Peptides > 7 aa**
MA6AHA	1	4994	5	3	213	25
Healthy	1	25335			105	33
Healthy	2	4581	3	1	48	11
Healthy	3	4155			44	14
Infected	1	12049	7	3	493	126
Infected	2	13753			400	81
Infected	3	10435			480	114

Table 8. Peptide identification summary for all samples and replicates: Particle type indicates the particle type used to perform the harvest: "MA6AHA" = general harvest particle, "Healthy" = particles encapsulated in membranes form healthy leukocytes, and "Infected" refers to particles prepared using membrane from VEEV-infected leukocytes. The low number of peptides with database matches (# DB Search PSMs) could reflect gaps in our sequence database or poor fragmentation on the mass spectrometer.

In addition to analyzing the MS/MS spectra and peptide sequences, we analyzed the mass spectra and the features that were present, focusing on features we predict correspond to peptides. The results of these feature analyses are given in Figure 13. We determined that we do not have high quality MS/MS spectra for the many of these features. We investigated modifications to the MS/MS parameters in order to capture better data for large highly charged features/peptides (Table 9 and Figure 14).

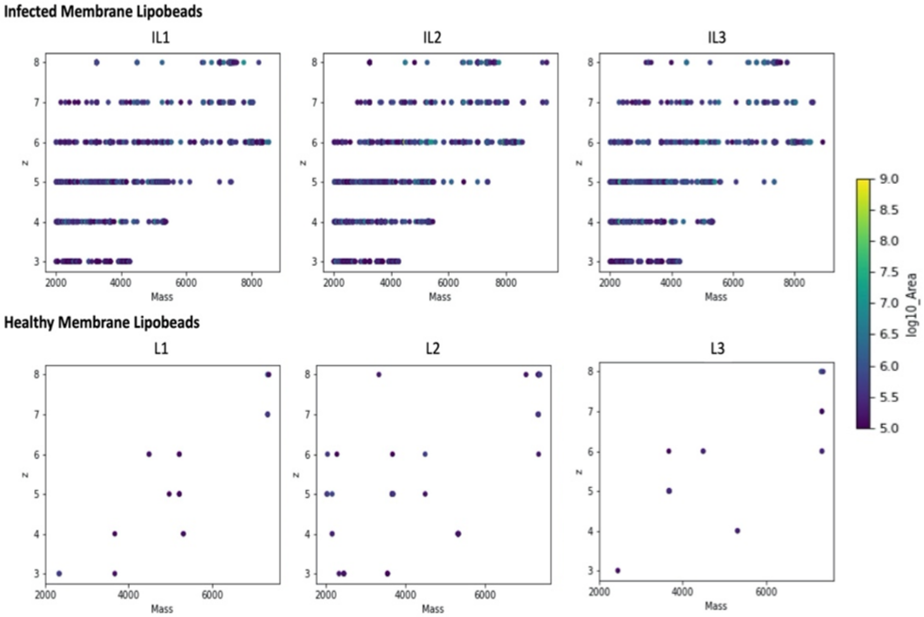


Figure 13. Peptide feature analyses of mass spectrometry data from snake samples: These analyses have focused on features with peak areas ≥ 1E5, and corresponding to peptides with masses of ≥ 2000 Da and charges of +3 - +8. **IL1**, **IL2** and **IL3** present the features identified from harvests performed using membranes from infected leukocytes**. L1**, **L2** and **L3** present the features identified from harvests performed using membranes from healthy leukocytes.

These were encouraging results for a first harvest of *Boa constrictor* host defense peptides; however further analysis of the first-round data was complicated by the limited number of confirmed peptide identifications from the database. Another round of harvests was performed directly from snake plasma using membrane-encapsulated particles (Exp. ID#4 in Table 6 above). The specific conditions, particles/membranes used and results for these harvests are outlined in Table 9. Here, we also evaluated changes to the MS/MS parameters in order to improve sequencing of larger peptides, as is described above. Focusing primarily on the de novo sequences, due to the low number of database identifications, we did observe that the new MS/MS method resulted in a greater number of larger peptide sequences relative to the original method (Figure 14). These results are encouraging for future efforts because it suggests that these new parameters will allow us to identify larger novel AVP’s, including peptides with sizes consistent with cathelicidin and defensin peptides.

Sample Type	Particle Type	Experimental Replicate #	Mass Spec Method 1		Mass Spec Method 2	
			Database Search	De Novo	Database Search	De Novo
Snake plasma, no infection/no incubation	Healthy leukocyte membrane lipobeads Hydro mod, MeOH washes (Direct lipobead harvest from plasma)	1	105	346	0	90
Snake plasma, no infection/no incubation	Healthy leukocyte membrane lipobeads Hydro mod, MeOH washes (Direct lipobead harvest from plasma)	2	42	364	9	76
Snake plasma, no infection/no incubation	Healthy leukocyte membrane lipobeads Hydro mod, MeOH washes (Direct lipobead harvest from plasma)	3	44	374	15	66
Snake plasma, no infection/no incubation	Infected leukocyte membrane lipobeads Hydro mod, MeOH washes (Direct lipobead harvest from plasma)	1	58	450	8	273
Snake plasma, no infection/no incubation	Infected leukocyte membrane lipobeads Hydro mod, MeOH washes (Direct lipobead harvest from plasma)	2	44	455	9	264
Snake plasma, no infection/no incubation	Infected leukocyte membrane lipobeads Hydro mod, MeOH washes (Direct lipobead harvest from plasma)	3	55	384	10	235
Snake plasma, no infection/no incubation	Infected leukocyte membrane lipobeads Hydro mod, DMF washes (Direct lipobead harvest from plasma)	1	56	339	4	181
Snake plasma, no infection/no incubation	Infected leukocyte membrane lipobeads Hydro mod, DMF washes (Direct lipobead harvest from plasma)	2	52	343	10	84
Snake plasma, no infection/no incubation	Infected leukocyte membrane lipobeads Hydro mod, DMF washes (Direct lipobead harvest from plasma)	3	55	344	2	68

Table 9. Results from second round of Boa constrictor analyses. Harvests were performed in triplicate from Snake plasma using hydrogel particles encapsulated in membranes form healthy and VEEV-infected human leukocytes. Two different sets of parameters were used during MS/MS analyses (Mass Spec Method 1 = original parameters and Mass Spec Method 2 = new parameters). We here report the number of unique peptide sequences determined for each harvest using each Mass Spec Method, differentiating between those that had database matches and those that were determined purely de novo.

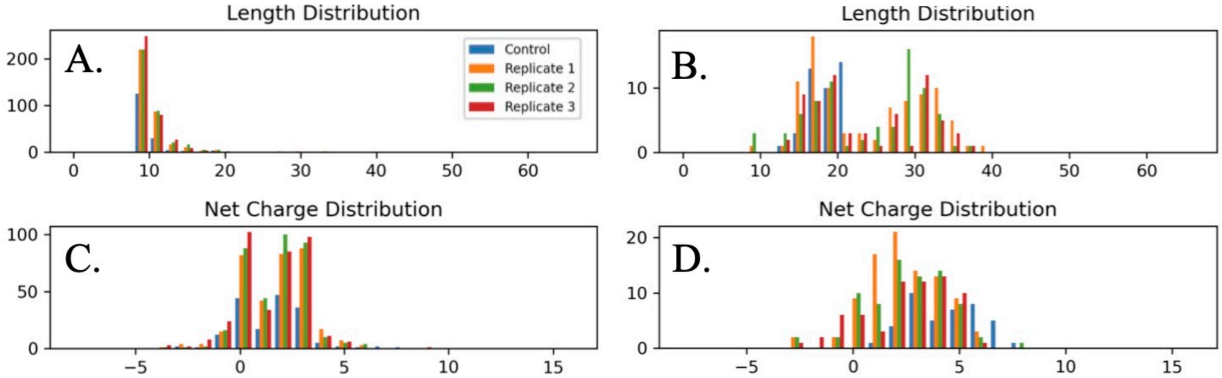


Figure 14. Representative results for comparison of Boa constrictor harvests – Harvests performed using membranes from healthy leukocytes. Comparison of peptide lengths (A. and B.) and net charges (C. and D.) determined using the original MS/MS parameters (Method 1: A. and C.) and the newer parameters (Method 2: B. and D.). The new parameters resulted in a significant shift towards larger peptides, with multiple identifications of peptides between 20 and 40 amino acids in length.

Major Task 4: Analyze harvested peptides (12 months: 03/30/19-03/29/20)

1. Statistical Analyses of Harvests:

1.) Workflow for efficiently analyzing harvest samples via tandem mass spectrometry and determination of the peptide sequences. We refined the sample processing workflow to optimize it for analysis via tandem mass spectrometry with ETD and HCD. The initial harvests begun in Major Task 3 provided an opportunity for refining the sample processing workflow. We have a working framework for performing harvests from plasma or culture media using the first-round general harvest core-shell particles as well as both VLP-based capture particles and membrane-encapsulated hydrogel particles.

2.) Mass spectrometry data and sequences for peptides harvested for the purpose of identifying novel antiviral peptides. In the third quarter of year 1, we analyzed peptides harvested from commercial plasma. Initial LC-MS/MS data yielded several peptide sequences of varied lengths. We observed a substantially greater number of *de novo* peptide sequences than sequences with matches in the reference sequence database. We used this harvest and mass spectrometry data to refine mass spectrometry, chromatography, and MS data analysis parameters.

3.) Statistical methods for analyzing harvested peptide sequences to afford statistical significance values of potential antiviral peptides.

Preliminary analyses: Peptide sequence data from initial harvests and replicate injections were sent to collaborators at PNNL and they used this data to begin framing their analyses. The team at PNNL analyzed data for these harvested peptides in the 3rd quarter evaluating run-to-run (five replicate runs on the mass spectrometer of the same sample) variability. Initial analyses indicate relatively high percentage of common peptides based on the Jaccard Index (Figure 15) and percentage of identical peptides. In respect to proteome coverage, two replicates will identify approximately ~73% of all of the peptides identified from the combined replicate injections. This increases dramatically to ~93% for 4 replicates.

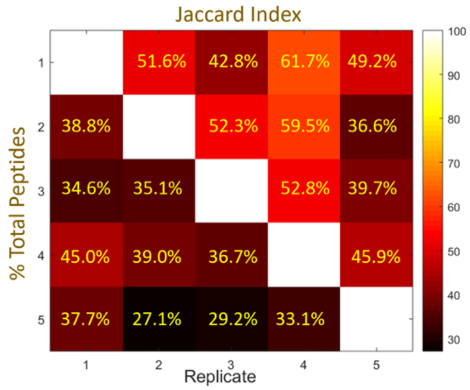


Figure 15: Image of two related similarity metrics (1) % total peptides and (2) Jaccard Index.

Healthy Alligator Plasma Initial Analysis: The team at PNNL analyzed peptide sequences identified from harvests performed on healthy alligator plasma beginning with a general using MA6AHA particles followed by harvests using the panel of membrane-encapsulated particles (described above under Major Task 3). Initially, the similarity and differences between harvests based solely on the observed peptides was evaluated. Again, the Jaccard index was used to evaluate similarity between samples, which evaluates the size of the intersection of observed peptides relative to the total peptides observed (Figure 16B). The data quality and similarity were excellent as seen in Figure 16C based on a standard hierarchical clustering based on the Jaccard similarity of the data. From this analysis it was observed that the healthy and infected leukocyte membranes were the most similar while the general harvest was the most distinct. Furthermore, we evaluated if there were differences in the total number of captured peptides for each of the harvests. As seen in Figure 17, the healthy trypsinized leukocyte membranes have the highest counts of observed peptides, but is only statistically larger than the infected leukocyte membranes based on a Friedman test ($\chi2$ statistic; p ~0.0013)

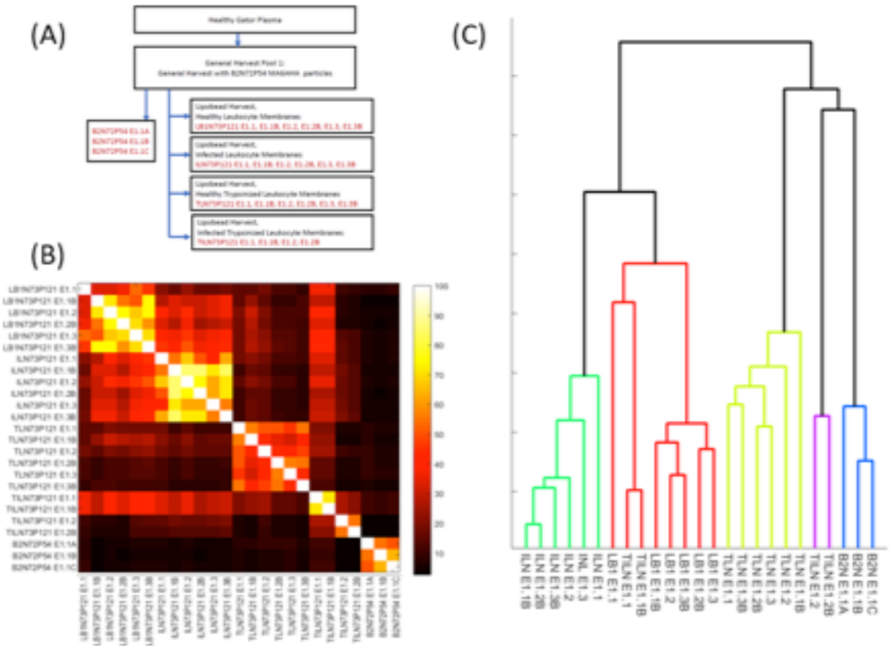


Figure 16: The (A) overall design showing relationship of the general harvest with B2N72P54 MA6AHA particles with the associated (B) JA measurements from the measured samples and (C) hierarchical clustering of samples into the expected groupings.

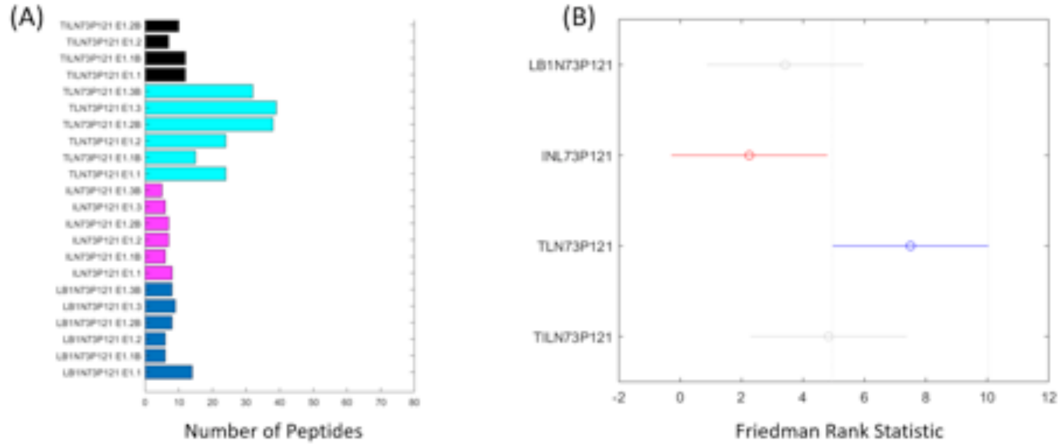


Figure 17: The (A) number of peptides per sample and (B) the post-hoc test for the statistical comparison of the number of peptides by harvest type.

A statistical analysis was performed that focused on the longest peptide variants (LPVs) in each harvest using a G-test (1) to evaluate if a specific peptide sequence was observed in one or more of the harvests than expected by random chance. This analysis identified several peptides of interest as unique to specific harvests. PEPAKSAPAPKKGSKKAVTK" and "GTGASGSFKLNKK" were both unique to infected leukocyte membranes with a p-values of approximately 0.002 and 0.007, respectively. Specifically, in this case PEPAKSAPAPKKGSKKAVTK was observed in 83.3% of the infected leukocyte membranes samples and none of the other samples. GTGASGSFKLNKK was observed in 66.7% of the infected leukocyte membranes samples and none of the other samples. The peptide NKRSTITSREIQTAVRLLLPGELAKHAVSEGTKAVTKY was unique to infected trypsinized leukocyte membranes samples (p-value approximately 0.007), observed 100% of the time in these samples and in none of the other groups. Lastly, ATPVKIRIENSNAFLSR was significantly higher in to infected cells (p-value approximately 0.032). It was observed in 100% of the infected samples, but only 62.5% of the samples. Thus, it was seen more frequently in the infected samples than expected by chance, but it was not unique. Expanded statistical analysis of the peptide sequences (not limited to LPVs) from each harvest revealed two additional peptide sequences (RPPGFTPFRS and KPMKDSTVLPHF) that were associated with harvests performed using membranes from infected cells (p-values of 0.0065 and 0.018, respectively).

[1.] Webb-Robertson et al., (2010) Combined statistical analyses of peptide intensities and peptide occurrences improves identification of significant peptides from MS-based proteomics data. J Proteome Res, 9(11):5748-56.

Supernatant from media with VEEV-infected alligator PBMCs (8 hours time point)

In this experiment, 20 samples were analyzed, 13 stimulated with VEEV and 7 that had not been - Table 10, with the various particles and membranes as shown. The short descriptor names for each harvest are given in the first column. Analysis of mass spectrometry data resulted in peptide sequences that were sorted into two groups based on whether they could be matched to sequences in a reference database (database) or whether were strictly determined de novo from the mass spectra and did not have a match in the reference database (de novo). These harvests yielded 83 database peptides, as well as 1903 de novo peptides. The total number of peptides per sample is shown in Figure 18. We first compared the peptides identified from the general harvest to harvests using membrane-encapsulated particles. No overlaps were observed in the database peptides identified from the general harvest (GH) and harvests using membrane-encapsulated particles (Membrane). While for de novo peptides, one peptide was identified from both GH and Membrane harvests, Figure 19.

Name	Sample Description	Particle	Membrane
8H-GH	Supernatant from media with VEEV-infected gator	MA6AHA	N/A
8H-o-H1	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Healthy
8H-o-H2	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Healthy
8H-o-H3	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Healthy
8H-o-I1	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Infected
8H-o-I2	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Infected
8H-o-I3	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Infected
8H-t-H1	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Trypsinized Healthy
8H-t-H2	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Trypsinized Healthy
8H-t-H3	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Trypsinized Healthy
8H-t-I1	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Trypsinized Infected
8H-t-I2	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Trypsinized Infected
8H-t-I3	Supernatant from media with VEEV-infected gator	NIPMAM Hydro Mod	Trypsinized Infected
MK-GH	Supernatant from media	MA6AHA	N/A
MK-o-H1	Supernatant from media	NIPMAM Hydro Mod	Healthy
MK -o-H2	Supernatant from media	NIPMAM Hydro Mod	Healthy
MK -o-H3	Supernatant from media	NIPMAM Hydro Mod	Healthy
MK -o-I1	Supernatant from media	NIPMAM Hydro Mod	Infected
MK -o-I2	Supernatant from media	NIPMAM Hydro Mod	Infected
MK -o-I3	Supernatant from media	NIPMAM Hydro Mod	Infected

Table 10. Description of experimental samples.

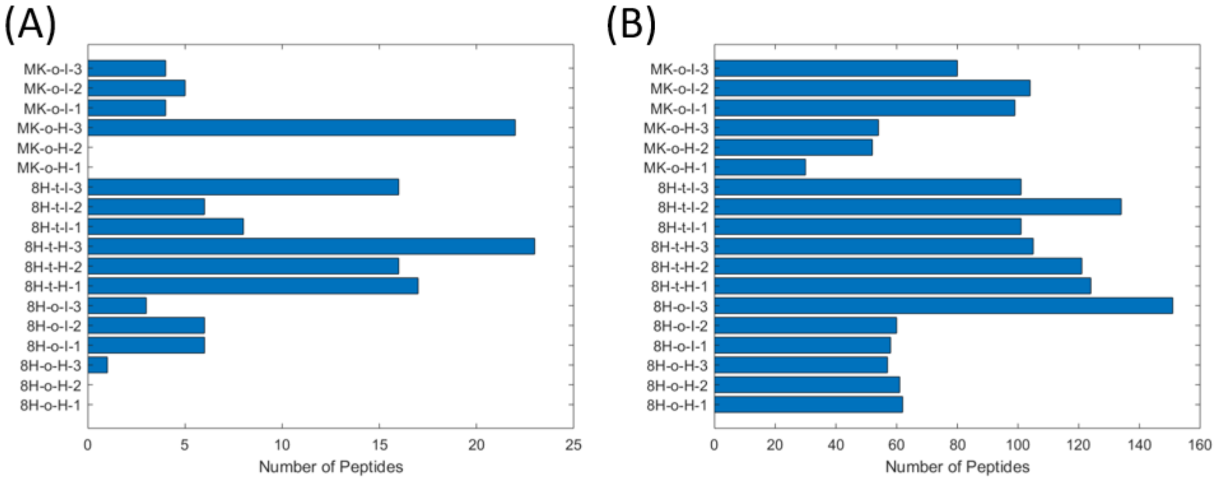


Figure 18. Total number of peptides per sample for the (A) database and (B) de novo datasets.

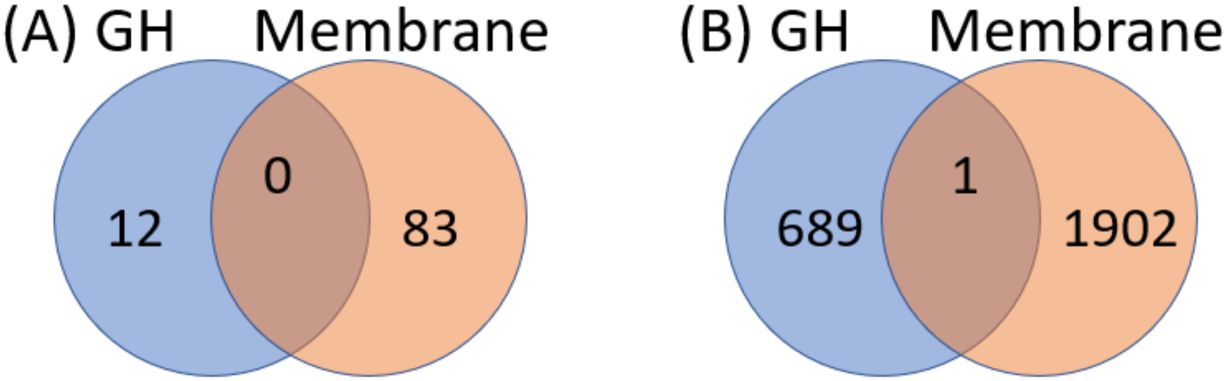


Figure 19. Overlap between peptide sequences from general and membrane harvests: (A) database and (B) de novo sequences.

We evaluated the similarity between each sample using the Jaccard Index (1), which is the number of peptides that two samples have in common scaled to the total number of peptides in the two samples. In Figure 20, the JA is shown in the bottom left triangle and the specific counts are shown in the remainder of the figure where the diagonal is the number of peptides per sample. Two samples with perfect identity will have a JA of 1. Figure 21 further demonstrates how the samples cluster based on the JA. In particular, they cluster on healthy versus infected, as well as a separation based on whether cell membrane proteins had been trypsinized prior to their use in generating the membrane-encapsulated particles (these particles have been referred to as “trypsinized”, and this term is also used to denote harvests using these particles).

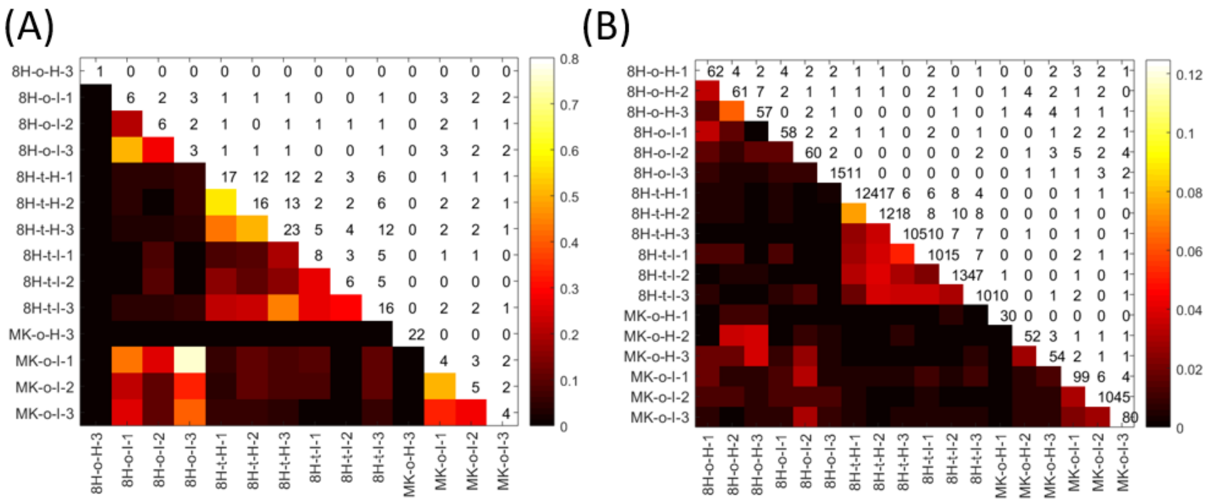


Figure 20. Jaccard Index for the (A) database and (B) de novo datasets with the peptide counts per sample given on the diagonal and the overlap in the upper triangle cells.

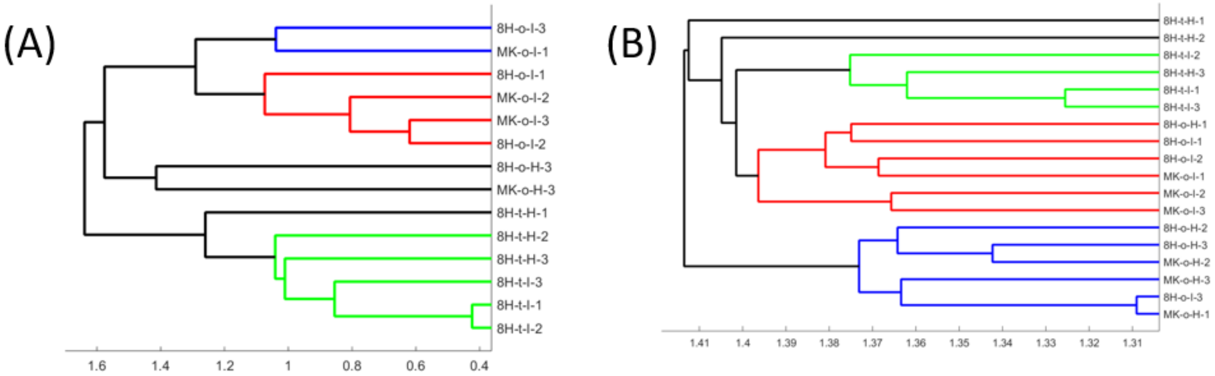
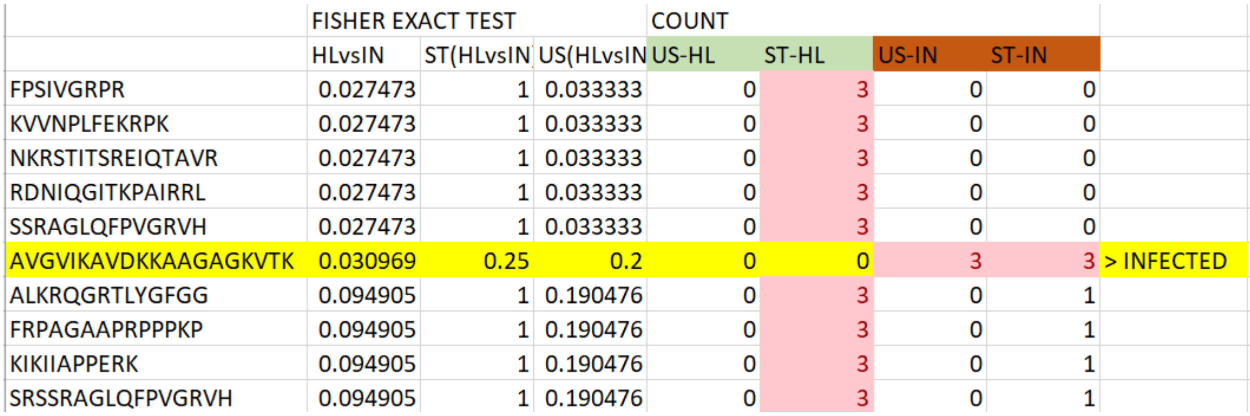


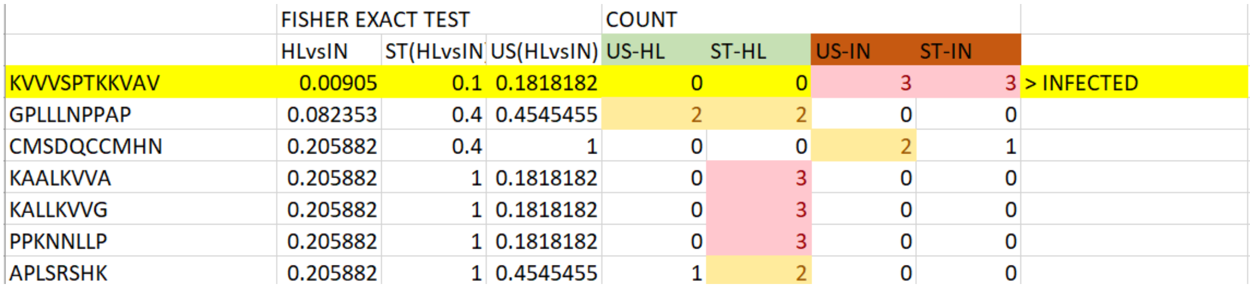
Figure 21. Hierarchal clustering based on Jaccard Index similarity: (A) database and (B) *de novo* datasets.

For statistical analyses, we used a Fisher's Exact test due to the small sample size. Similarly, significant p-values via the Fisher's Exact test are difficult to achieve due to the small sample sizes, however we can use them to sort the results. There were 10 and 7 peptides with p-values that can be utilized for prioritization from the database and de novo searches, respectively. ST and US are used to represent harvests from media collected from stimulated and unstimulated alligator PBMCs, respectively, and results are in Figures 22 and 23.



	FISHER EXACT TEST			COUNT				
	HLvsIN	ST(HLvsIN	US(HLvsIN	US-HL	ST-HL	US-IN	ST-IN	
FPSIVGRPR	0.027473	1	0.033333	0	3	0	0	
KVVNPLFEKRPK	0.027473	1	0.033333	0	3	0	0	
NKRSTITSREIQTAVR	0.027473	1	0.033333	0	3	0	0	
RDNIQGITKPAIRRL	0.027473	1	0.033333	0	3	0	0	
SSRAGLQFPVGRVH	0.027473	1	0.033333	0	3	0	0	
AVGVIKAVDKKAAGAGKVTK	0.030969	0.25	0.2	0	0	3	3	> INFECTED
ALKRQGRTLYGFGG	0.094905	1	0.190476	0	3	0	1	
FRPAGAAPRPPPKP	0.094905	1	0.190476	0	3	0	1	
KIKIIAPPERK	0.094905	1	0.190476	0	3	0	1	
SRSSRAGLQFPVGRVH	0.094905	1	0.190476	0	3	0	1	

Figure 22. Top antiviral peptide candidates for peptides from the database dataset (labeled ">INFECTED").



	FISHER EXACT TEST			COUNT				
	HLvsIN	ST(HLvsIN	US(HLvsIN)	US-HL	ST-HL	US-IN	ST-IN	
KVVVSPTKKVAV	0.00905	0.1	0.1818182	0	0	3	3	> INFECTED
GPLLLNPPAP	0.082353	0.4	0.4545455	2	2	0	0	
CMSDQCCMHN	0.205882	0.4	1	0	0	2	1	
KAALKVVA	0.205882	1	0.1818182	0	3	0	0	
KALLKVVG	0.205882	1	0.1818182	0	3	0	0	
PPKNNLLP	0.205882	1	0.1818182	0	3	0	0	
APLSRSHK	0.205882	1	0.4545455	1	2	0	0	

Figure 23. Top antiviral peptide candidates for the de novo peptides (labeled ">INFECTED").

<u>Harvests from plasma isolated from healthy and VEEV-stimulated alligator blood (8 hour time point)</u>
In this experiment, 10 samples were analyzed, 5 stimulated with VEEV and 5 that had not been - Table 11, with the various particles and membranes as shown. The short descriptor names for each harvest are given in the first column. These harvests yielded identification of 137 database peptides, as well as 1840 de novo peptides. The total number of peptides per sample is shown in Figure 24. We first compared the peptides identified from the general harvest to those from harvests using membrane-encapsulated particles. For the database there were 5 peptides that overlap between GH and Membrane harvests, while the de novo data exhibited no overlap in peptide sequences between the two groups, Figure 25.

Name	Sample Description	Particle	Membrane
ST-GH	Gator plasma, incubated 8 hours, stimulated VEEV	MA6AHA	N/A
ST-o-HL	Gator plasma, incubated 8 hours, stimulated VEEV	NIPMAM Hydro Mod	Healthy
ST-o-IN	Gator plasma, incubated 8 hours, stimulated VEEV	NIPMAM Hydro Mod	Infected
ST-t-HL	Gator plasma, incubated 8 hours, stimulated VEEV	NIPMAM Hydro Mod	Trypsinized Healthy
ST-t-IN	Gator plasma, incubated 8 hours, stimulated VEEV	NIPMAM Hydro Mod	Trypsinized Infected
US-GH	Gator plasma, incubated 8 hours, no VEEV	MA6AHA	N/A
US-o-HL	Gator plasma, incubated 8 hours, no VEEV	NIPMAM Hydro Mod	Healthy
US-o-IN	Gator plasma, incubated 8 hours, no VEEV	NIPMAM Hydro Mod	Infected
US-t-HL	Gator plasma, incubated 8 hours, no VEEV	NIPMAM Hydro Mod	Trypsinized Healthy
US-t-IN	Gator plasma, incubated 8 hours, no VEEV	NIPMAM Hydro Mod	Trypsinized Infected

Table 11. Description of experimental samples

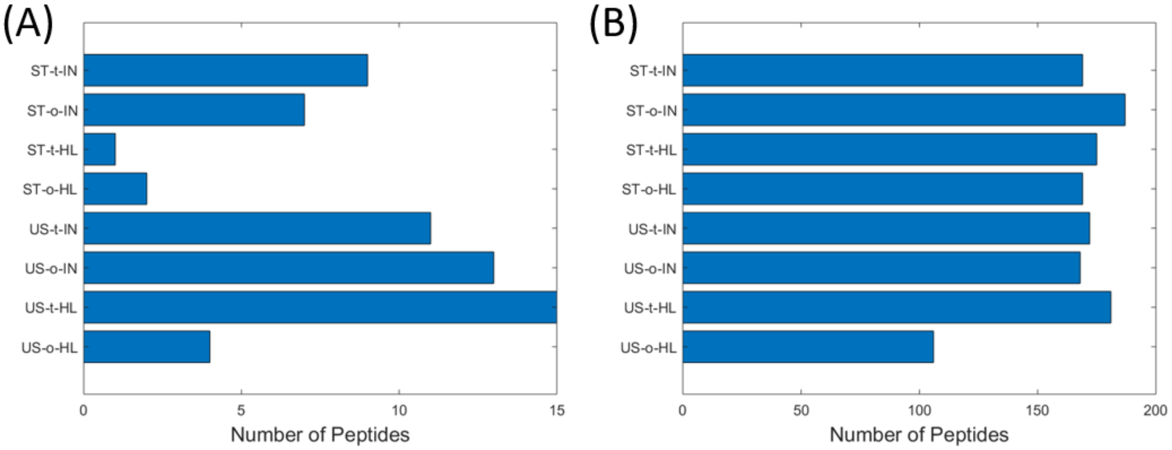


Figure 24. Total number of peptides per sample for the (A) database and (B) de novo datasets.

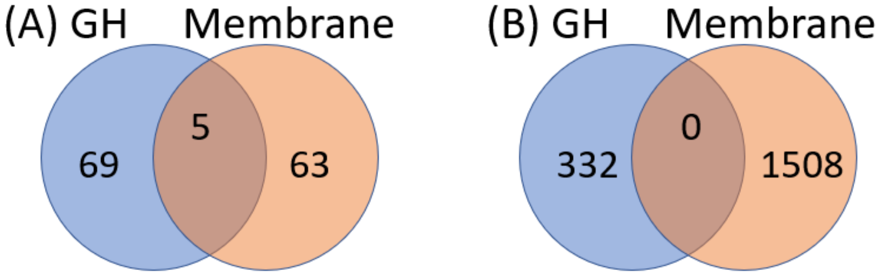


Figure 24. Overlap between peptide sequences from general and membrane harvests: (A) database and (B) de novo sequences.

We again evaluated the similarity between each sample using the Jaccard Index (1). In Figure 26, the JA is shown in the bottom left triangle and the specific counts are shown in the remainder of the figure where the diagonal is the number of peptides per sample. Figure 27 further demonstrates how the samples cluster based on the JA. In particular, they cluster on healthy versus infected, as well as a separation based on whether the cell membrane proteins had been trypsinization.

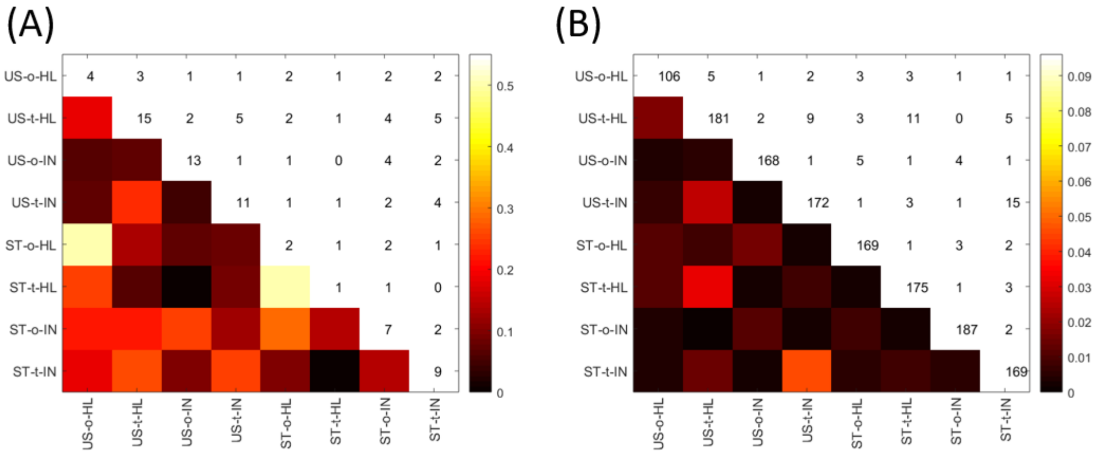


Figure 26. Jaccard Index for the (A) database and (B) de novo datasets with the peptide counts per sample given on the diagonal and the overlap in the upper triangle cells.

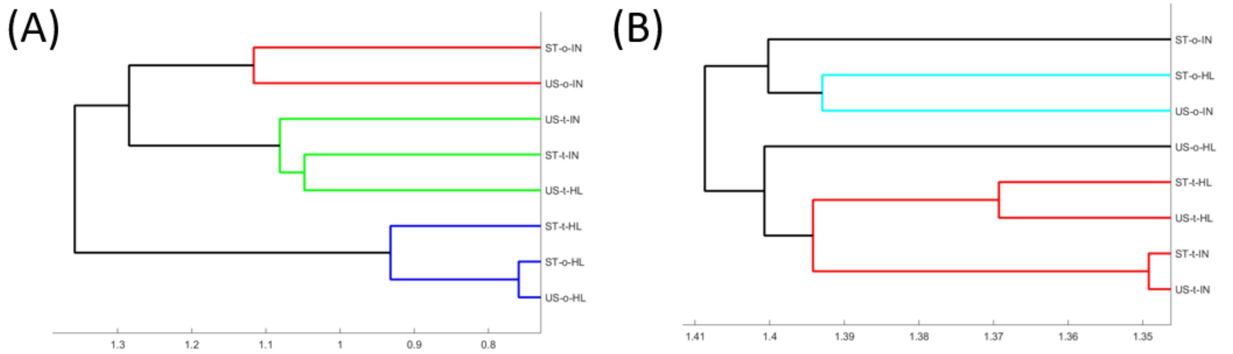
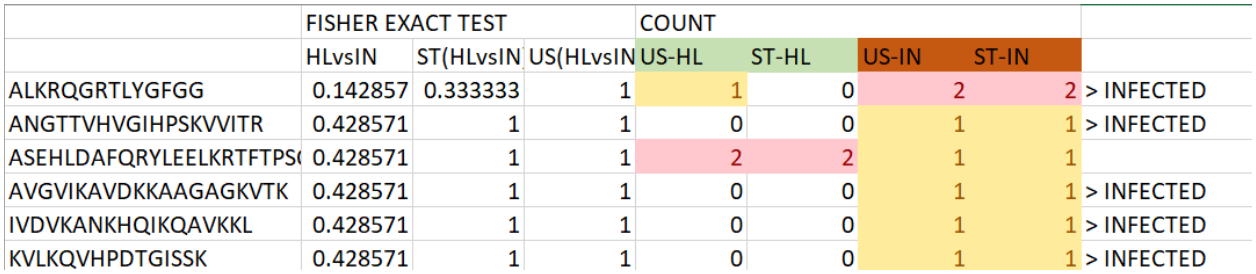


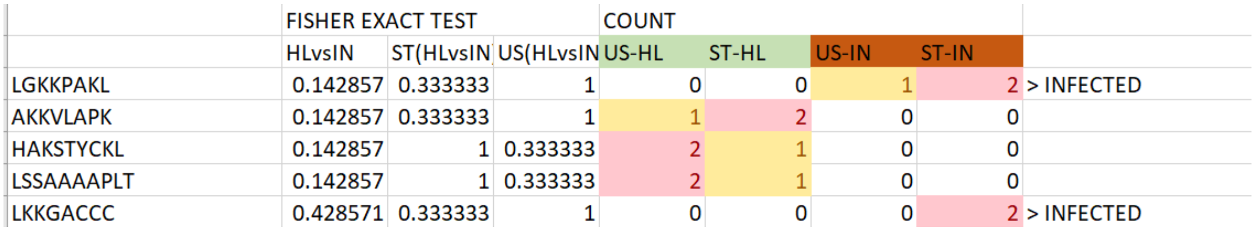
Figure 27. Hierarchal clustering based on Jaccard Index similarity: (A) database and (B) de novo datasets.

For statistical analyses, we used a Fisher's Exact test due to the small sample size. Similarly, significant p-values via the Fisher's Exact test are difficult to achieve due to the small sample sizes, however we can use them to sort the results. There were 6 and 5 peptides with p-values that can be utilized for prioritization from the database and de novo searches, respectively. ST and US are used to represented harvests from plasma derived from stimulated and unstimulated alligator blood, respectively, and results are in Figures 28 and 29.



	FISHER EXACT TEST			COUNT				
	HLvsIN	ST(HLvsIN	US(HLvsIN	US-HL	ST-HL	US-IN	ST-IN	
ALKRQGRTLYGFGG	0.142857	0.333333	1	1	0	2	2	> INFECTED
ANGTTVHVGIHPSKVVITR	0.428571	1	1	0	0	1	1	> INFECTED
ASEHLDAFQRYLEELKRTFTPS(	0.428571	1	1	2	2	1	1	
AVGVIKAVDKKAAGAGKVTK	0.428571	1	1	0	0	1	1	> INFECTED
IVDVKANKHQIKQAVKKL	0.428571	1	1	0	0	1	1	> INFECTED
KVLKQVHPDTGISSK	0.428571	1	1	0	0	1	1	> INFECTED

Figure 28. Top antiviral peptide candidates for peptides from the database search (labeled ">INFECTED").



	FISHER EXACT TEST			COUNT				
	HLvsIN	ST(HLvsIN	US(HLvsIN	US-HL	ST-HL	US-IN	ST-IN	
LGKKPAKL	0.142857	0.333333	1	0	0	1	2	> INFECTED
AKKVLAPK	0.142857	0.333333	1	1	2	0	0	
HAKSTYCKL	0.142857	1	0.333333	2	1	0	0	
LSSAAAAPLT	0.142857	1	0.333333	2	1	0	0	
LKKGACCC	0.428571	0.333333	1	0	0	0	2	> INFECTED

Figure 29. Top antiviral peptide candidates for the de novo peptides (labeled ">INFECTED").

<u>Harvests comparing general harvest with ones using VEEV-modified particles</u>

In this experiment, 9 samples were analyzed, 4 stimulated with VEEV and 5 that had not been - Table 12, with the various particle formulations and with/without covalently-bound VEEV as shown. The short descriptor names for each harvest are given in the first column. These harvests yielded 138 database peptides, as well as 1399 de novo peptides. The total number of peptides per sample is shown in Figure 30. We first compared the peptides identified from the general harvest to those from the harvests using VEEV-modified particles (VLP). For database peptides, 26 peptides were shared between general harvest (GH) and harvests using virus-modified particles (VLP), while de novo peptides afforded 119 shared peptides, Figure 31.

Name	Sample Description	Particle	Bound Virus
HP-HL-1	Harvest with HPMA	HPMA	No
HP-HL-2	Harvest with HPMA	HPMA	No
HP-IN-1	Harvest with HPMA coupled to virus	HPMA	Yes
HP-IN-2	Harvest with HPMA coupled to virus	HPMA	Yes
N4-HL-1	Harvest with MA6aHA	N4	No
N4-HL-2	Harvest with MA6aHA	N4	No
N4-IN-1	Harvest with MA6aHA coupled to virus	N4	Yes
N4-IN-2	Harvest with MA6aHA coupled to virus	N4	Yes

Table 12. Description of experimental samples

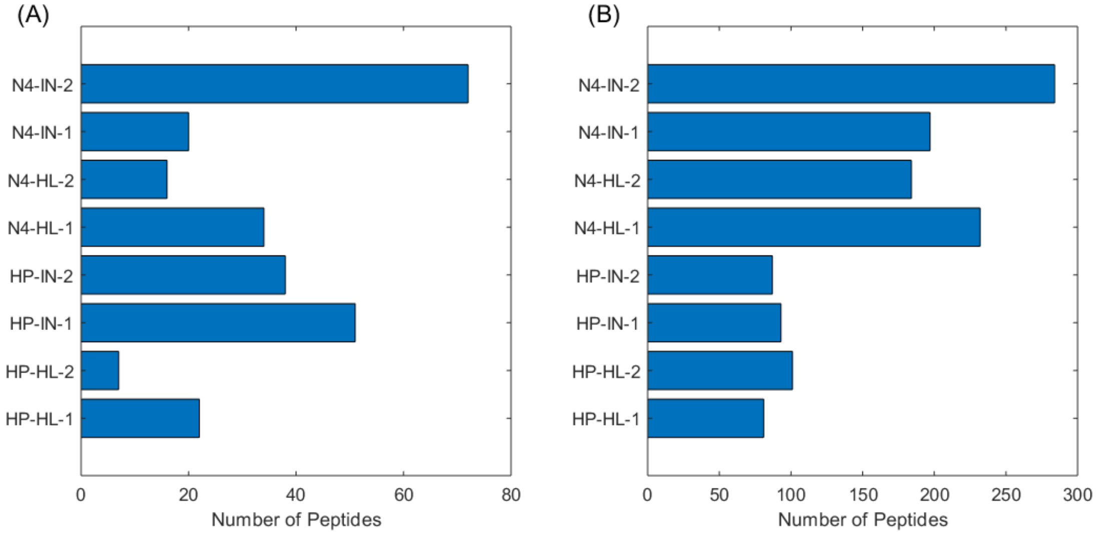


Figure 30. Total number of peptides per sample for the (A) database and (B) de novo datasets.

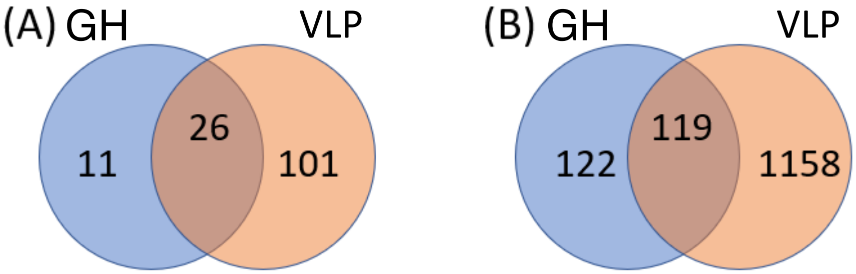


Figure 31. Overlap between peptide sequences from general and VLP harvests: (A) database and (B) *de novo* sequences.

In Figure 32, the JA is shown in the bottom left triangle and the specific counts are shown in the remainder of the figure where the diagonal is the number of peptides per sample. Figure 33 further demonstrates how the samples cluster based on the JA. In particular, they cluster on core particle type (HPMA vs N4) and unmodified versus virus-modified particles.

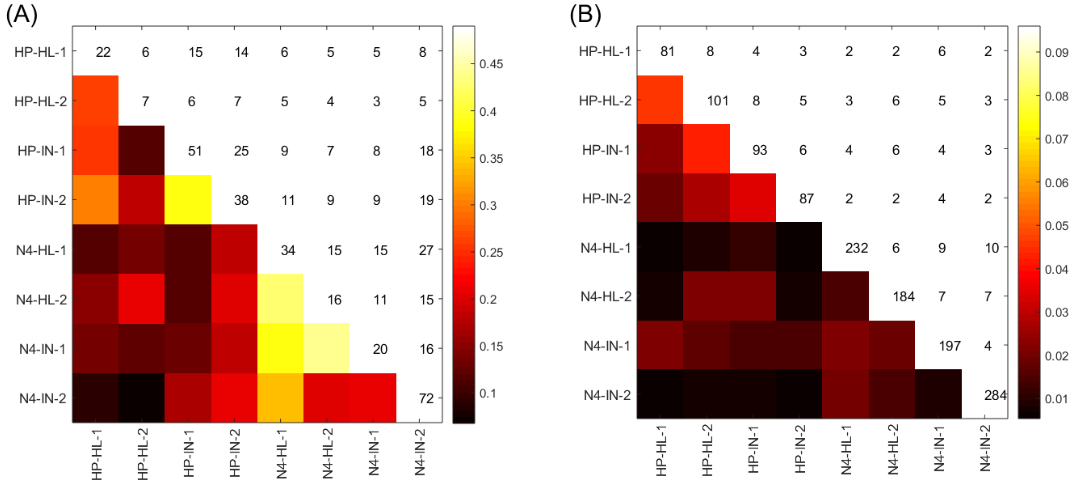


Figure 32. Jaccard Index for the (A) database and (B) de novo datasets with the peptide counts per sample given on the diagonal and the overlap in the upper triangle cells.

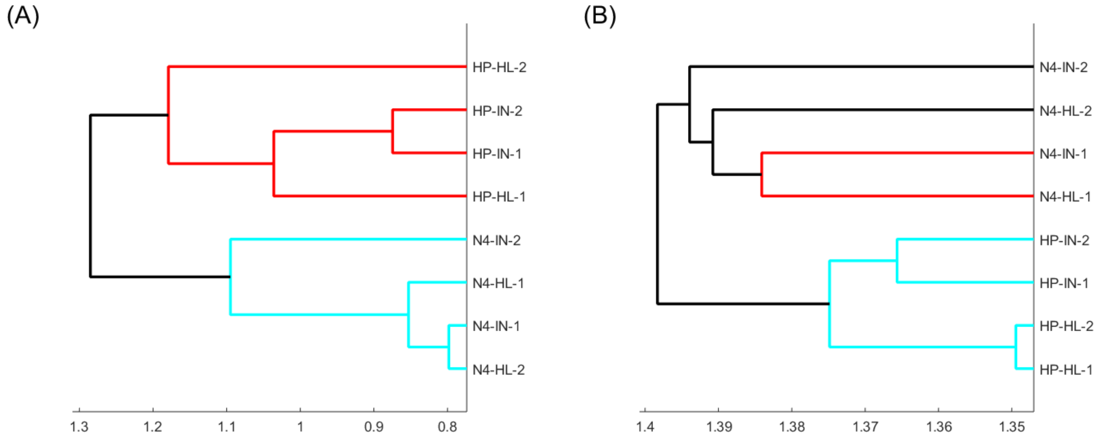
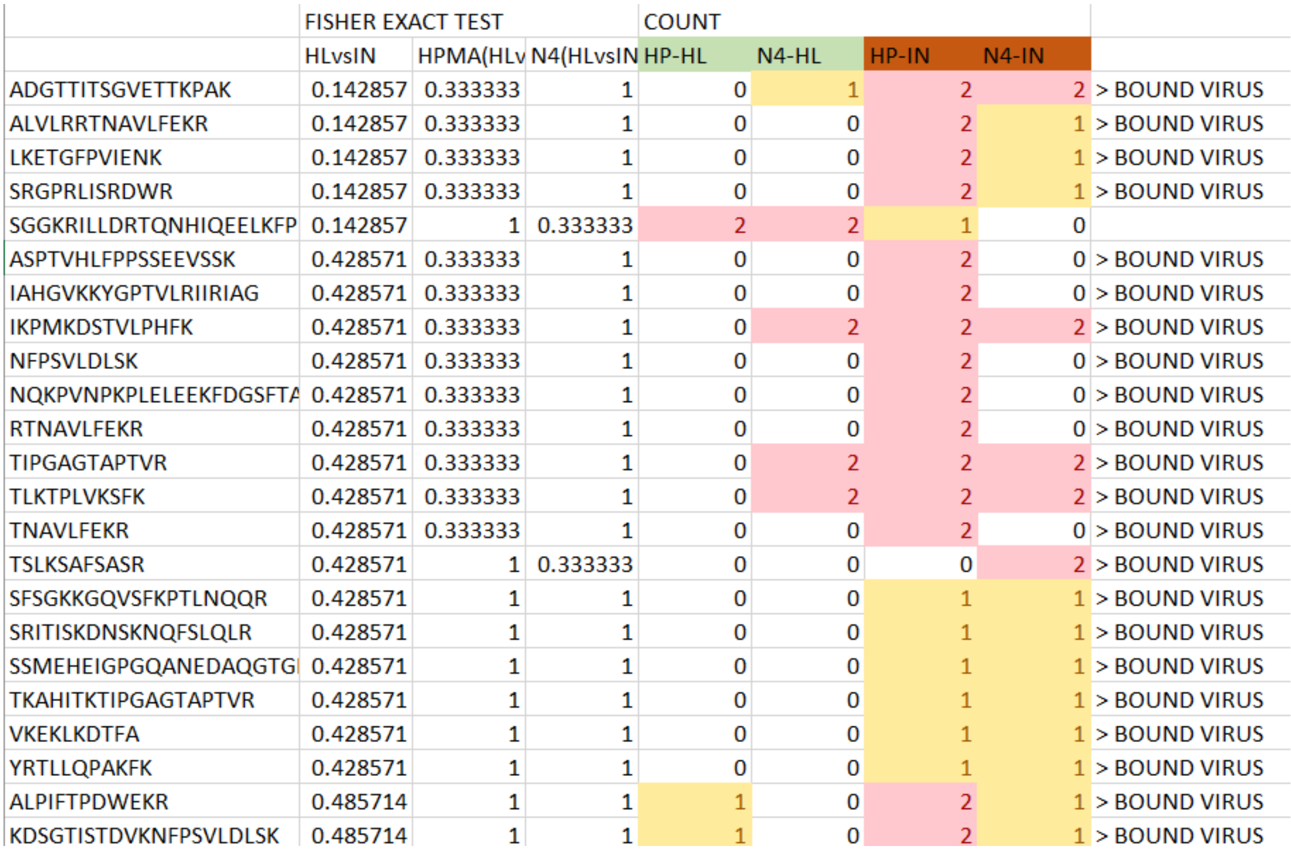


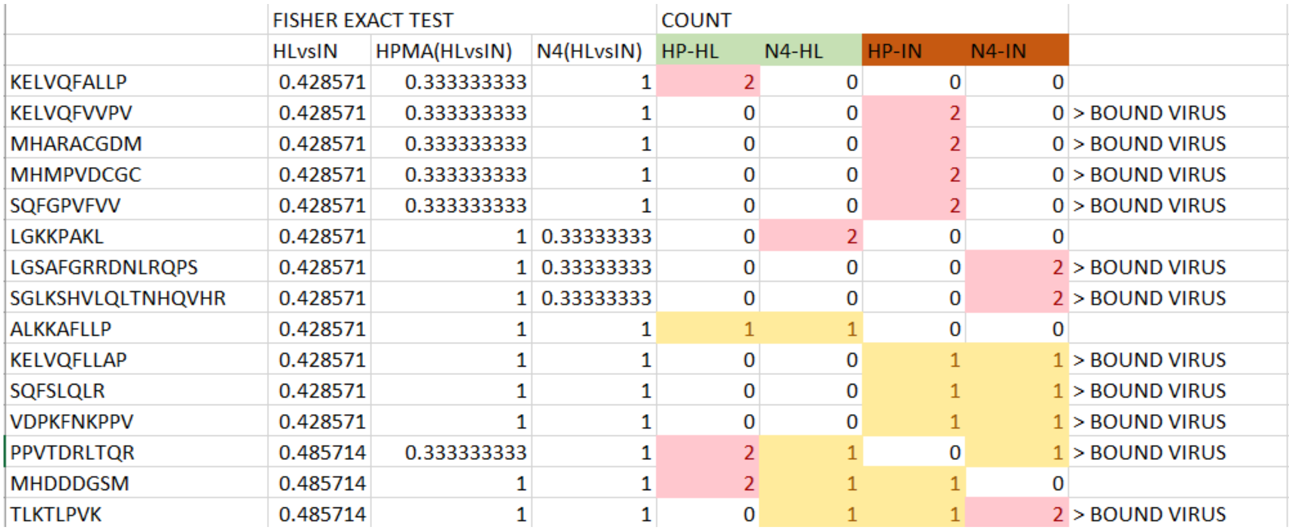
Figure 33. Hierarchal clustering based on Jaccard Index similarity: (A) database and (B) de novo sequences.

For statistical analyses, we used a Fisher's Exact test due to the small sample size. Similarly, significant p-values via the Fisher's Exact test are difficult to achieve due to the small sample sizes, however we can use them to sort the results. There were 15 and 12 peptides with p-values that can be utilized for prioritization from the database and de novo sequences, respectively; results are in Figures 34 and 35.



	FISHER EXACT TEST			COUNT				
	HLvsIN	HPMA(HLv	N4(HLvsIN	HP-HL	N4-HL	HP-IN	N4-IN	
ADGTTITSGVETTKPAK	0.142857	0.333333	1	0	1	2	2	> BOUND VIRUS
ALVLRRTNAVLFEKR	0.142857	0.333333	1	0	0	2	1	> BOUND VIRUS
LKETGFPVIENK	0.142857	0.333333	1	0	0	2	1	> BOUND VIRUS
SRGPRLISRDWR	0.142857	0.333333	1	0	0	2	1	> BOUND VIRUS
SGGKRILLDRTQNHIQEELKFP	0.142857	1	0.333333	2	2	1	0	
ASPTVHLFPPSSEEVSSK	0.428571	0.333333	1	0	0	2	0	> BOUND VIRUS
IAHGVKKYGPTVLRIIRIAG	0.428571	0.333333	1	0	0	2	0	> BOUND VIRUS
IKPMKDSTVLPHFK	0.428571	0.333333	1	0	2	2	2	> BOUND VIRUS
NFPSVLDLSK	0.428571	0.333333	1	0	0	2	0	> BOUND VIRUS
NQKPVNPKPLELEEKFDGSFTA	0.428571	0.333333	1	0	0	2	0	> BOUND VIRUS
RTNAVLFEKR	0.428571	0.333333	1	0	0	2	0	> BOUND VIRUS
TIPGAGTAPTVR	0.428571	0.333333	1	0	2	2	2	> BOUND VIRUS
TLKTPLVKSFK	0.428571	0.333333	1	0	2	2	2	> BOUND VIRUS
TNAVLFEKR	0.428571	0.333333	1	0	0	2	0	> BOUND VIRUS
TSLKSAFSASR	0.428571	1	0.333333	0	0	0	2	> BOUND VIRUS
SFSGKKGQVSFKPTLNQQR	0.428571	1	1	0	0	1	1	> BOUND VIRUS
SRITISKDNSKNQFSLQLR	0.428571	1	1	0	0	1	1	> BOUND VIRUS
SSMEHEIGPGQANEDAQGTGI	0.428571	1	1	0	0	1	1	> BOUND VIRUS
TKAHITKTIPGAGTAPTVR	0.428571	1	1	0	0	1	1	> BOUND VIRUS
VKEKLKDTFA	0.428571	1	1	0	0	1	1	> BOUND VIRUS
YRTLLQPAKFK	0.428571	1	1	0	0	1	1	> BOUND VIRUS
ALPIFTPDWEKR	0.485714	1	1	1	0	2	1	> BOUND VIRUS
KDSGTISTDVKNFPSVLDLSK	0.485714	1	1	1	0	2	1	> BOUND VIRUS

Figure 34. Top antiviral peptide candidates for peptides from the database search.



	FISHER EXACT TEST			COUNT				
	HLvsIN	HPMA(HLvsIN)	N4(HLvsIN)	HP-HL	N4-HL	HP-IN	N4-IN	
KELVQFALLP	0.428571	0.333333333	1	2	0	0	0	
KELVQFVVPV	0.428571	0.333333333	1	0	0	2	0	> BOUND VIRUS
MHARACGDM	0.428571	0.333333333	1	0	0	2	0	> BOUND VIRUS
MHMPVDCGC	0.428571	0.333333333	1	0	0	2	0	> BOUND VIRUS
SQFGPVFVV	0.428571	0.333333333	1	0	0	2	0	> BOUND VIRUS
LGKKPAKL	0.428571	1	0.33333333	0	2	0	0	
LGSAFGRRDNLRQPS	0.428571	1	0.33333333	0	0	0	2	> BOUND VIRUS
SGLKSHVLQLTNHQVHR	0.428571	1	0.33333333	0	0	0	2	> BOUND VIRUS
ALKKAFLLP	0.428571	1	1	1	1	0	0	
KELVQFLLAP	0.428571	1	1	0	0	1	1	> BOUND VIRUS
SQFSLQLR	0.428571	1	1	0	0	1	1	> BOUND VIRUS
VDPKFNKPPV	0.428571	1	1	0	0	1	1	> BOUND VIRUS
PPVTDRLTQR	0.485714	0.333333333	1	2	1	0	1	> BOUND VIRUS
MHDDDGSM	0.485714	1	1	2	1	1	0	
TLKTLPVK	0.485714	1	1	0	1	1	2	> BOUND VIRUS

Figure 35. Top antiviral peptide candidates for the de novo peptides.

4.) Predictive model(s) based on machine learning for analyzing harvested peptides to aid in identification of peptides likely to have antiviral properties. - **Develop Machine Learning Algorithms for AVP Prediction:**

The series of machine learning developments in AVP have to date focused on increasing the features that characterize a peptide and modifications of the machine learning algorithm. They have not included feature reduction techniques that would determine the most relevant and non-redundant features from the input features used. The performance of a machine learning model relies heavily on using most informative features and non-informative features can degrade classifier performance. We generated candidate features from the physicochemical and secondary structure properties of known AVP and non-AVP sequences, Table 13. We identified the most important features by estimating Pearson's correlation coefficient and mean decrease of Gini index (MDGI) for all candidate features. We then applied a recursive feature elimination (RFE) algorithm to determine the most important features to aide in interpretability of the model. Finally, we developed a machine learning predictive model based on a Support Vector Machine (SVM) to predict antiviral activity based on sequence.

Peptide Feature	Feature dimension
Amino acid composition	20D
Dipeptide composition	400D
Pseudo-amino acid composition	25D
Amphiphilic pseudo-amino acid composition	30D
Composition/transition/distribution	168D
Secondary structure sequence	6D

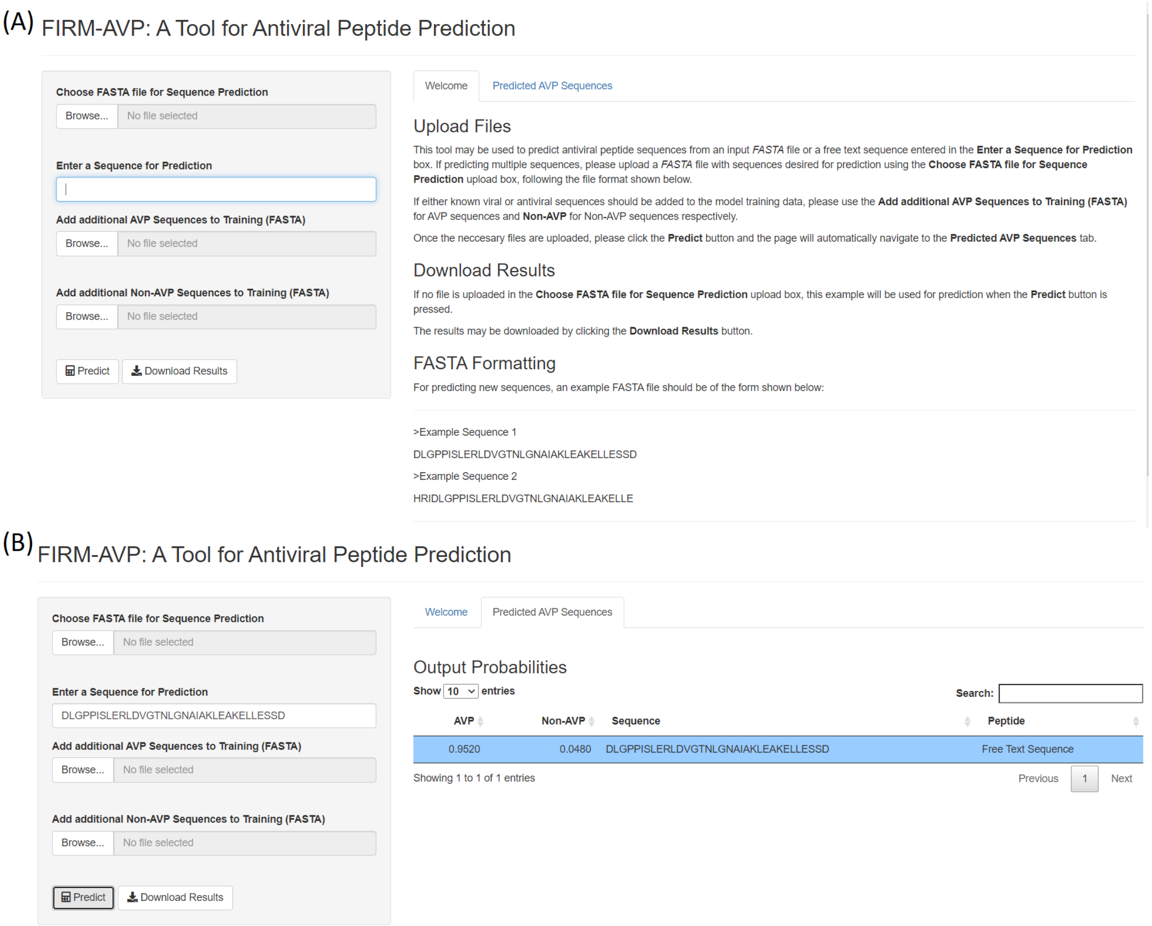
Table 13. List of 649 Peptide Features

The performance of the SVM classifiers with our best feature sets is evaluated by estimating sensitivity, specificity, accuracy and Mathew's Correlation Coefficient (MCC) values using Eqs.1-4, where TP, TN, FP, and FN are true positives (positives accurately classified), true negatives (negatives accurately classified), false positives (negatives classified as positives), and false negatives (positives classified as negatives), respectively. Table 14 shows that the new SVM approach achieved better prediction accuracy as compared to the state-of-the-art in all four core metrics reported by previous algorithms.

The Feature-Informed Machine Learning approach to AVP prediction (FIRM-AVP) has been published in the journal Scientific Reports. In addition to improving overall accuracy as seen in Table 8, mostly using the Support Vector Machine (SVM) approach versus a Random Forest (RF) or Deep Learning (DL) methods, we have also developed a web-service to enable users to evaluate AVP peptides (Figure 36).

Model	Sensitivity (%)	Specificity (%)	Accuracy (%)	MCC
FIRM-AVP (SVM)	93.3	91.1	92.4	0.84
FIRM-AVP (RF)	95.0	82.2	89.5	0.79
FIRM-AVP (DL)	91.7	80.0	86.7	0.73
AVP-649D (SVM)	95.0	82.2	89.5	0.79
AVP-649D (RF)	90.0	82.2	86.7	0.73
AVPcompo	83.3	88.9	85.7	0.72
AVPphysico	88.3	82.2	85.7	0.71
RFcompo+structure+agg	91.7	86.7	89.5	0.79

Table 14. Performance comparison of our models with existing models on validation data where MCC is the Matthew's Correlation Coefficient, for which the higher the value the more overall accuracy of the model.



FIRM-AVP: A Tool for Antiviral Peptide Prediction

Choose FASTA file for Sequence Prediction: No file selected

Enter a Sequence for Prediction

Add additional AVP Sequences to Training (FASTA): No file selected

Add additional Non-AVP Sequences to Training (FASTA): No file selected

Upload Files

This tool may be used to predict antiviral peptide sequences from an input *FASTA* file or a free text sequence entered in the **Enter a Sequence for Prediction** box. If predicting multiple sequences, please upload a *FASTA* file with sequences desired for prediction using the **Choose FASTA file for Sequence Prediction** upload box, following the file format shown below.

If either known viral or antiviral sequences should be added to the model training data, please use the **Add additional AVP Sequences to Training (FASTA)** for AVP sequences and **Non-AVP** for Non-AVP sequences respectively.

Once the neccesary files are uploaded, please click the **Predict** button and the page will automatically navigate to the **Predicted AVP Sequences** tab.

Download Results

If no file is uploaded in the **Choose FASTA file for Sequence Prediction** upload box, this example will be used for prediction when the **Predict** button is pressed.

The results may be downloaded by clicking the **Download Results** button.

FASTA Formatting

For predicting new sequences, an example FASTA file should be of the form shown below:

```
>Example Sequence 1
DLGPPISLERLDVGTNLGNAIAKLEAKELLESSD
>Example Sequence 2
HRIDLGPPISLERLDVGTNLGNAIAKLEAKELLE
```

FIRM-AVP: A Tool for Antiviral Peptide Prediction

Choose FASTA file for Sequence Prediction: No file selected

Enter a Sequence for Prediction: DLGPPISLERLDVGTNLGNAIAKLEAKELLESSD

Add additional AVP Sequences to Training (FASTA): No file selected

Add additional Non-AVP Sequences to Training (FASTA): No file selected

Output Probabilities

Show 10 entries

Search:

AVP	Non-AVP	Sequence	Peptide
0.9520	0.0480	DLGPPISLERLDVGTNLGNAIAKLEAKELLESSD	Free Text Sequence

Showing 1 to 1 of 1 entries

Previous 1 Next

Figure 36. Online FIRM-AVP software interface (https://msc-viz.emsl.gov/AVPR/). Where (A) is the starting page that allows users to either paste in a single peptide sequence or upload a FASTA file containing a collection of peptide sequences. Example sequences and files are given. (B) The probability of AVP versus non-AVP is returned for each sequence based on the pasted peptide sequence or the uploaded FASTA file.

Major Task 5: Assess performance of likely antiviral peptides (6 months: 09/30/19-03/29/20)

1.) One or more novel alligator or snake peptides that exhibit antiviral properties against VEEV.

Five peptides were identified as binding specifically to the cell-membrane particles from infected cells as compared to cell-membrane particles from uninfected cells. These five peptides were purchased and tested for in vitro cytotoxicity and anti-viral activity using Vero cells. LL37 was included as a positive control as it has been shown to inhibit VEEV replication (Ahmed et al., 2019, PMID: 30738837). We also included a scrambled LL-37 as a negative control. Cell viability was determined using CellTiter-Glo (Promega), which measures ATP within cells. Minimal cell toxicity was observed in Vero cells treated with up to 100 ug/mL of the novel peptides (Figure 37A). In particular, peptides KF, AR, and RS did not induce any cell death even at 100 ug/mL However, there was some toxicity (~25%) observed in cells treated with 100 ug/mL of LL-37, but not scrambled LL-37.

Antiviral activity was determined by pre-treating cells for 1 hour with peptides, infecting cells with a firefly luciferase reporter virus (VEEV-luc) for 1 hour, followed by post-treatment with peptides. Luminescent was measured at 16 hours post-infection as an indicator of viral replication. Peptide GK showed no antiviral activity (Figure 37B). Peptide PK only inhibited VEEV at 100 ug/mL Peptides KF, AR, and RS showed dose dependent inhibition of VEEV, with peptide AR showing the most potent inhibition. LL-37 potently inhibit VEEV replication down to background levels, but surprisingly the scrambled LL-37 also inhibited VEEV. The most potent novel peptide is AR, which is derived from Alligator complement-3.

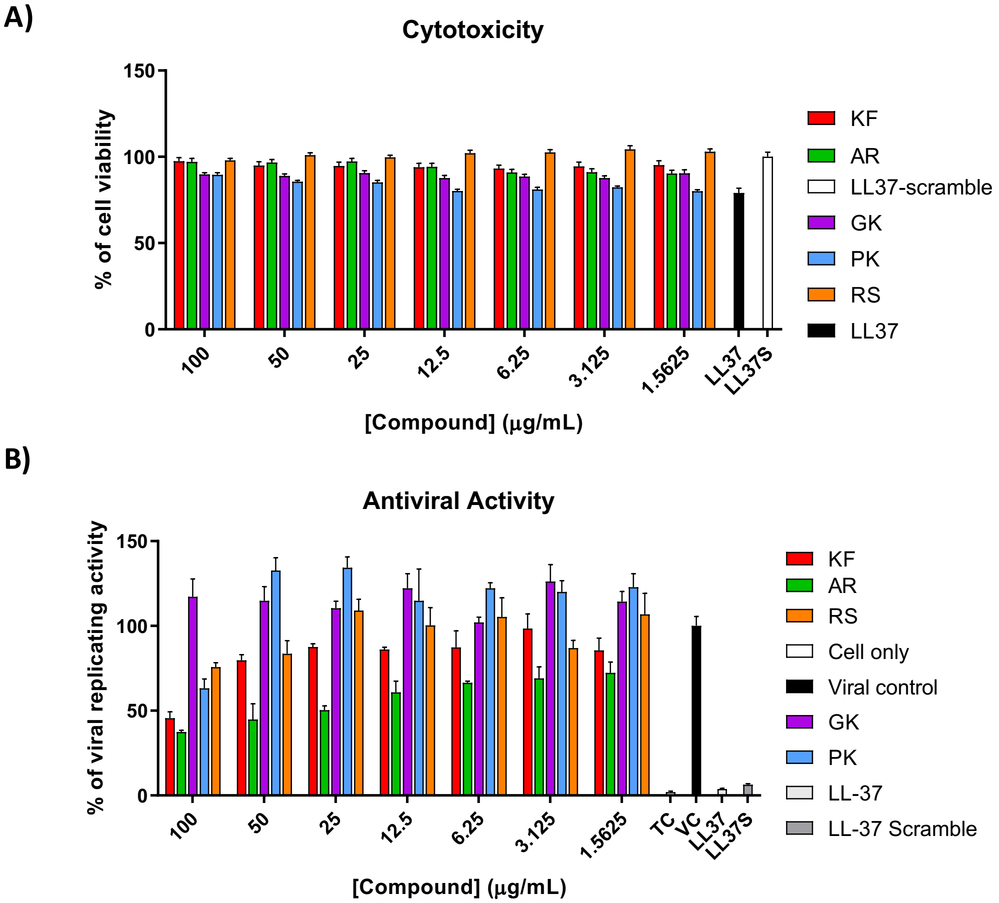


Figure 37: Cell toxicity and antiviral activity of novel peptides. A) Vero cells were treated with various concentrations of peptides (as indicated) for 16 hours. LL-37 and scrambled LL-37 were used at 100 ug/mL Cell viability was measured using CellTiter-Glo. Untreated cells were set to 100% and all treatments normalized to these values. Values graphed are averages plus standard deviations, N=3 **B)** Vero cells were pre-treated with the indicated concentrations of peptides for 1 hour, infected with VEEV-luciferase reporter virus (MOI 1) for 1 hour, and then post-treated with peptides. Luminescence was measured at 16 hours post-infection. LL-37 and scrambled LL-37 were used at 100 ug/mL Untreated infected cells (viral control) were set to 100% and all treatments normalized to these values. Cells only are uninfected cells which display only background luminescence signal. Values graphed are averages plus standard deviations, N=3.

Additional assays were performed to confirm the antiviral activity of AR-17. First, we tested the ability of AR-17 to suppress replication of VEEV-GFP reporter virus. As can be seen in Figure 38, treatment with 100 ug/mL and 50 ug/mL of AR-17 suppressed the expression of GFP as compared to the untreated virally infected samples (Figure 38).

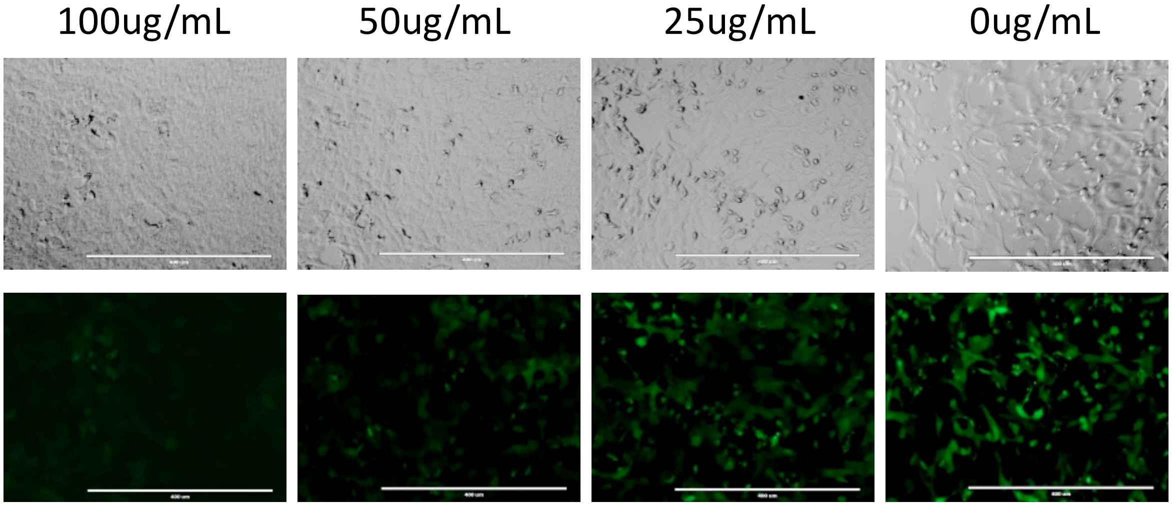


Figure 38. Peptide AR-17 inhibits VEEV replication. Vero cells were pre-treated with the indicated concentrations of peptide for 1 hour, infected with VEEV-GFP reporter virus (MOI 1) for 1 hour, and then post-treated with peptides.

Second, we tested the impact of AR-17 on VEEV replication through plaque assays which measure the amount of infectious virus being released from cells. At 8 hours post infection (hpi) AR-17 decreased VEEV titers by ~2 logs and by ~ 1 log at 16 hpi (Figure 39). These data confirm that AR-17 inhibits VEEV replication.

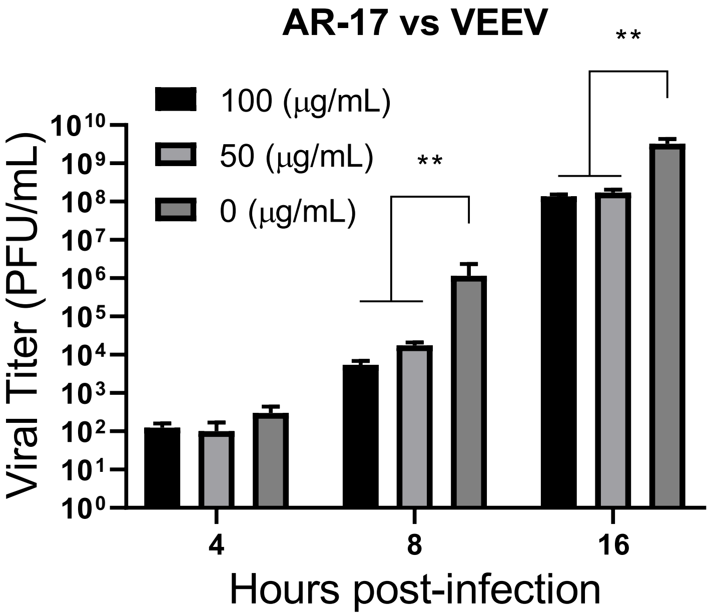


Figure 39. AR-17 inhibits VEEV infectious titers. Vero cells were pre-treated with the indicated concentrations of peptide for 1 hour, infected with VEEV (MOI 1) for 1 hour, and then post-treated with peptides. At 4, 8, and 16 hpi viral supernatants were collected and tittered by plaque assay.

A second set of nine peptides were tested (Table 15). These peptides were harvested preferentially using particles bearing either membranes from infected human leukocytes or covalently bound VEEV as baits. In laboratory testing, none of the peptides displayed any toxicity at up to 100 μg/mL towards

Vero cells (Figure 40). Antiviral testing of these peptides was initially performed using the VEEV luciferase reporter virus. However, there was extremely high variations in the viral replication observed among replicates and the peptides did not for the most part display a dose dependent response (Figure 41). Some peptides, such as VIT and Plasmin increase viral replication at higher concentrations. Three peptides (Comp, SerIg and Rib1) were chosen for additional antiviral testing based on a potential trend of antiviral activity. In order to minimize deviations, it was decided to test antiviral properties of the most efficient peptides (Comp, SerIg and Rib1) by plaque assay at 8h and 24 h time points. Rib1 (50 and 100 µg/mL) showed a minor but significant reduction of VEEV titers at 8 hours post-infection, but not difference in viral titers at 24 hours post-infection (Figure 42). Similarly, SerIg and Comp peptides (at only 100 µg/mL) showed a minor but significant reduction of VEEV titers at 8 hours post-infection, but no difference in viral titers at 24 hours post-infection (Figure 42). These results indicate that these peptides have minimal anti-VEEV activity.

Name	**Sequence**	**Molecular Weight**	**Charge**	**Origin**
AM-EF	AVGVIKAVDKKAAGAGKVTK	1911.3	+4	Elongation Factor
AM-VIT	SRGPRLISRDWR	1498.7	+3	Vitronectin
AM-Plasmin	ALVLRRTNAVLFEKR	1786.2	+3	Plasminogen Isoform
AM-Comp	YRTLLQPAKFK	1364.6	+3	Complement component C8
AM-SerIg	TKAHITKTIPGAGTAPTVR	1920.2	+3	Immunoglobulin
AM-IgG	SRITISKDNSKNQFSLQLR	2235.5	+3	Immunoglobulin
AM-Rib1	IVDVKANKHQIKQAVKKL	2060.5	+4	Ribosomal protein
AM-Rib2	ANGTTVHVGIHPSKVVITR	1985.3	+2	Ribosomal protein
AM-His	KVLKQVHPDTGISSK	1636.9	+2	Histone

Table 15. Second round of peptides selected from alligator plasma harvests. The peptide AM-EF was identified from two separate harvest studies and was strongly associated with harvests using membranes form infected cells, making it a promising candidate, and one of the few peptides that we identified that appears in its entirety in multiple harvests. The peptides AM-VIT, AM-Rib1, AM-Rib2 and AM-His were identified in harvests from plasma and stimulated PBMC's using particle bearing membranes from VEEV-infected human leukocytes. The remaining peptides, AM-Plasmin, AM-Comp, AM-SerIg and AM-IgG were identified specifically from harvests using particles bearing covalently bound VEEV as baits.

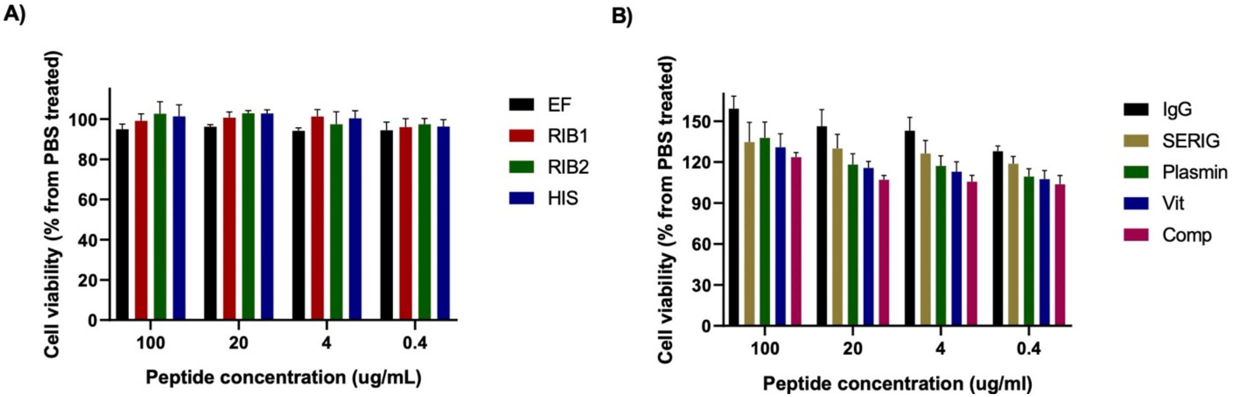


Figure 40: Cell toxicity of 2nd set of novel peptides. Vero cells were treated with various concentrations of peptides (as indicated) for 16 hours. Cell viability was measured using CellTiter-Glo. Untreated cells were set to 100% and all treatments normalized to these values. Values graphed are averages plus standard deviations, N=3

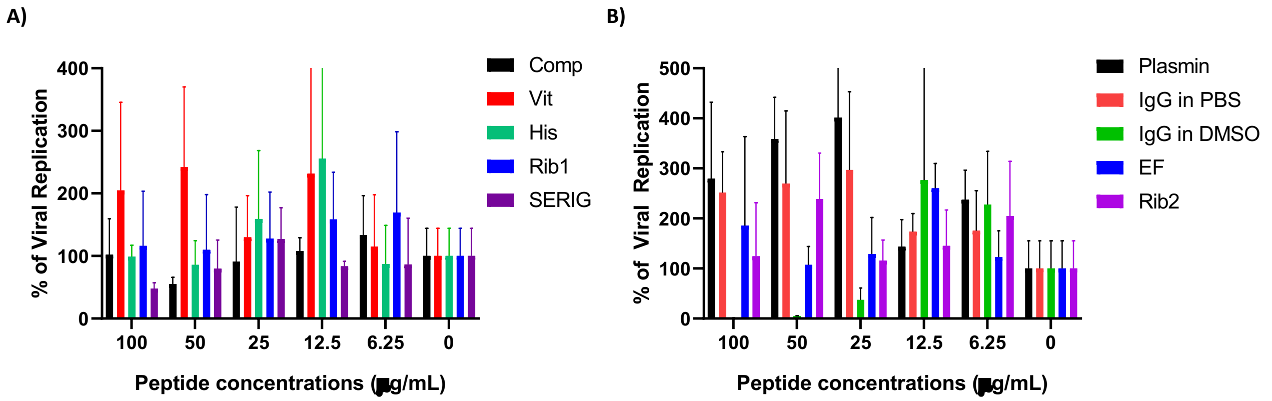
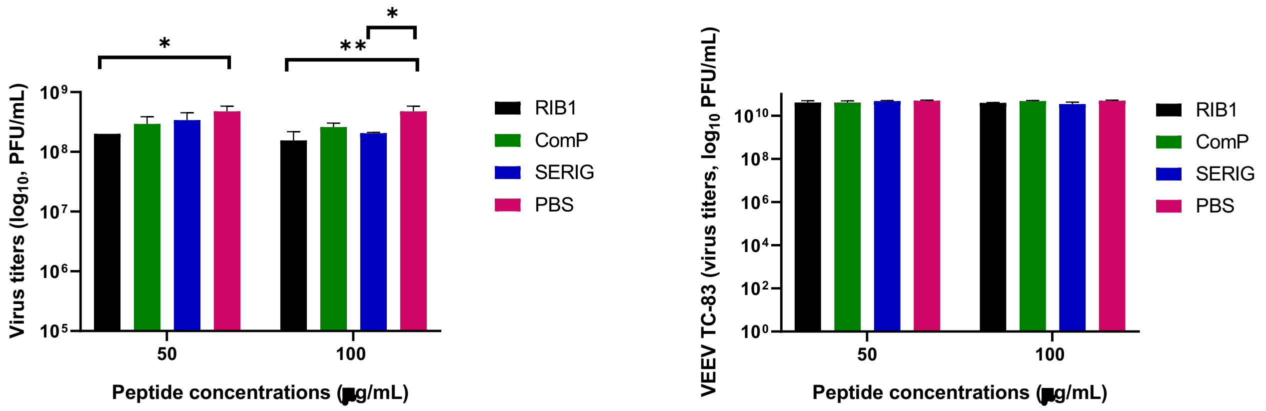


Figure 41. Antiviral testing of 2nd set of peptides via reporter virus assay. Cells were treated with peptides for 1 hour prior to infection with VEEV-luciferase at MOI 0.1. One hour post infection infectant was removed and fresh media containing peptide was added back to each well. Antiviral effect was estimated based on luminescent assay 20h post infection. Untreated infected cells were set to 100% and all treatments normalized to these values. IgG, when resuspended in DMSO, killed cells at higher concentration.



8 hours

24 hours

Figure 42. Antiviral testing of 2nd set of peptides via plaque assay. Cells were treated with peptides in two concentrations 50 and 100 ug/mL for 1 hour prior to infection with VEEV at MOI 0.1. One hour post infection infectant was removed and fresh media containing peptide was added back to each well. Supernatants were collected 8 and 24h post infection. Antiviral effect was estimated based on plaque assay.

What opportunities for training and professional development has the project provided?

Over the course of the project, one student (Victoria Callahan) and two research associates (Shih-Chao Lin and Ivan Akhrymuk) have been involved in this project from the Kehn-Hall Lab. Over the same period, two graduate students (Amy Carfagno and Samuel Garvey), a student research assistant (Sabrina Lamont) and a research associate (Liana Chafran) have been involved on this project from the Bishop Lab. Dr. Bishop, Dr. Kehn-Hall and researchers from both labs meet frequently to plan and coordinate their efforts on the project.

This project provided a unique training environment because it was a multidisciplinary project, which allowed researchers from the biology/virology side to be exposed to the application of chemistry to address a biological question. They gained experience stimulating reptile blood cells with virus, which is outside the scope of the typically human focused research in the Kehn-Hall Lab. Similarly, the project provided researchers with backgrounds in organic chemistry, materials science and proteomic mass spectrometry the opportunity to work with virologists and biologists to gain experience and understanding of the biology that is at the heart of the project.

Researchers on the project contributed in the preparation of the reports and have also be included in the preparation of manuscripts produced by the project and will continue to do so in future related manuscripts. They have presented their data at conferences and will do so at future conferences.

How were the results disseminated to communities of interest?

Our strategy for disseminating the results and findings of the project to communities of interest focus primarily on publication of technical manuscripts in peer-reviewed journals and presentations and national and other professional meetings. This year:

- We presented a poster on the project and our findings at the MHSRS conference in Orlando in August, 2019 2019 (see below under **"Other publications, conference papers, and presentations"**).
- We presented a poster on the project and our findings at the Defense Threat Reduction Agency CBD S&T conference in November of 2019 (see below under **"Other publications, conference papers, and presentations"**).
- We also presented on the project and our findings at ASM Biothreats in January, 2020. (see below under **"Other publications, conference papers, and presentations"**).
- Data from the project was also included as part of a DTRA Tech Watch Seminar presented by Dr. Bishop in April, 2020.
- Our manuscript reporting development of improved machine learning algorithms for predicting antiviral peptides has been published the scientific journal Nature Scientific Reports. (Please see "Products" for publication details.)

What do you plan to do during the next reporting period to accomplish the goals?

Nothing to Report

4. IMPACT:

What was the impact on the development of the principal discipline(s) of the project?

The ultimate goal of this project was to identify novel peptides that have antiviral activity against Venezuelan equine encephalitis virus (VEEV). VEEV is an important human and veterinary pathogen with no current treatment options. Therefore, discovery of antivirals is an urgent unmet need. Due to the nature of antiviral peptides, the discovered peptides will likely have broad-spectrum activity, being able to inhibit multiple viruses within the alphavirus family that cause significant disease, such as chikungunya virus (CHIKV) and eastern equine encephalitis virus (EEEV). The peptide discovery process developed here, one that tightly integrates advanced peptidomic analytical technologies, rigorous performance evaluation and machine learning, can in the future be adapted to include other viruses or classes of virus. In expanding the scope of the process and tools to other viruses, we anticipate that it can be developed into a platform for identifying peptides with broad-spectrum antiviral efficacy and basis for developing broad-spectrum antiviral therapeutics for use against emerging viral pathogens and future pandemic outbreaks.

The results of these studies have highlighted the unmet need for robust machine-learning algorithms for high confidence prediction and design of antiviral peptides and provided a framework for integrating experimental antiviral peptide discovery and assessment with development of predictive machine learning algorithms for this purpose. Such algorithms will be invaluable for developing future therapeutics based on antiviral peptides.

What was the impact on other disciplines?

We have developed novel particles that can be used to isolate peptides associate with the innate immune response. These particles can be used as discovery tools for multiple other disciplines including cancer biology, neurobiology and immunology. There are also implications for using these particles as sensors.

What was the impact on technology transfer?

Nothing to report at this time.

What was the impact on society beyond science and technology?

The rapid spread of SARS-CoV-2, and its evolution into the COVID-19 pandemic that continues to grip the United States and the rest of the world has revealed the inadequacy of our preparations for such emerging viral threats and the need for new strategies new therapeutics. Our successful identification of novel peptides that exhibit antiviral properties against the target pathogen Venezuelan equine encephalitis virus (VEEV), demonstrates the potential of this strategy for antiviral peptide discovery. While our current efforts are focused on identifying antiviral peptides that are effective against alpha viruses, specifically targeting VEEV, the peptide discovery platform and tools that we are developing can be readily adapted and extended to include a broader spectrum of viruses, including SARS-CoV-2. Hence, the present research provides the foundation for the development of a powerful antiviral peptide discovery and design platform and an arsenal of novel effective antiviral peptides to combat emerging viruses and future outbreaks in their early stages. Moreover, the antiviral peptide discovery platform and tools that we are developing should be amenable to being adapted to target other pathogens including fungal and parasites. Thus, this research has potential broad biomedical applications and the potential to have a significant impact on the health of both warfighters and civilians.

5. CHANGES/PROBLEMS:

Changes in approach and reasons for change

Nothing to Report

Actual or anticipated problems or delays and actions or plans to resolve them

Technical challenges associated with quantifying propargyl monomer incorporation in particles and in preparing suitable amounts of VLP for incorporation in virus modified-particles have delayed development of virus-particle coupling conditions and achieving Major Task 1 deliverables by 09/29/2019 as originally projected in the SOW. This delay did not significantly impact our ability to achieve the final goals of the project as projected in the SOW. Despite the initial challenges, we ultimately, resorted to NMR for estimating propargyl monomer incorporating in the particles, addressing the first challenge. With respect to the challenges associated with producing sufficient amounts of VLP's for particle production, we explored using intact virus, which ended up providing a better route for generating particles presenting virus surface features as the affinity bait. Initial harvests performed using these particles resulted in the capture and identification of a large number of peptides. Our initial studies yielded promising results. Over the course of the project, we continued refining the virus-modified particle architecture. However, our efforts emphasized harvests using membrane-encapsulated particles based on our initial successes in identifying potential antiviral peptides using them.

Efforts to optimize conditions for stimulating reptile blood cells with VEEV took longer than originally projected due to complications in scheduling collection of blood samples due to Hurricane Dorian, which had been projected in different forecasts to hit Florida where our collaborators at the St. Augustine Alligator Farm Zoological Park are located. Their need to prepare for the potential storm to protect the animals disrupted our schedule for getting blood in September, which in turn delayed our efforts to optimize stimulation conditions beyond our original projected 09/29/12019 deliver date. That being said, we received a fresh shipment of alligator blood on 10/23/2019. PBMCs were collected and stimulation studies performed. Conditions and a protocol for stimulating reptile blood with VEEV were established by the end of October, completing this Task 3 deliverable.

The SARS-CoV-2 pandemic and its impact on GMU operations became the biggest challenge to us completing project tasks and achieving deliverables within the remaining time allocated for the project (reflecting original requested and approved no-cost-extension to May). In mid- March, the university shifted to remote learning/teaching for all Spring semester classes, and the preparations and changes in schedule impacted faculty, staff and student's work on the project. In late March, GMU effectively temporarily shut down all research that was not determined to be critical SARS-COV-2 research, which included our AIMM project. At the time of the shutdown, harvests from alligator plasma and stimulated media along with snake plasma were at various stages of processing and analysis. Many of these samples had not been analyzed on the mass spectrometer at the time of the shutdown and were placed in cold storage (-80C) until our return. However, we were able to remotely analyze mass spectrometry data from those samples that had been analyzed on the instrument. We identified promising peptides in the third quarter based on analyses performed by the team at PNNL.

At the time that GMU shut down non-SARS-CoV-2 research, we requested and were granted an additional no cost extension through August 30 to provide us added time to compensate for the lost time and effort associated with the shut-down. At the time, we believed that if we were able to return to the lab in early summer, we are positioned to wrap up these studies and deliver on our project deliverables by that time. However, when we were allowed to return to lab in June, the restrictions and time constraints put in place as part of GMU's safety precautions in response to the pandemic slowed our progress on the project. Accordingly, we requested and were granted an additional no cost extension to October 31. Based on our experience working under the new rules and conditions, we believed that this would allow us to complete the projects objectives and deliverables.

In August, we arranged for the synthesis of the identified antiviral peptide candidates from New England Peptides. These peptides arrived later that month, and we immediately began assessing their properties. The measures that we took positioned us to successfully these studies and deliver on our project deliverables by the extended project end date.

Changes that had a significant impact on expenditures

Nothing to Report

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Nothing to Report

Significant changes in use or care of human subjects

Not applicable

Significant changes in use or care of vertebrate animals

Protocol [ACURO Assigned Number]: ACURO Log Number = DM171301
Title: Translational Peptide Research for Personnel Protection
Target required for statistical significance: Proposed 20 crocodilians and 15 snakes
Target approved for statistical significance: Approved

Protocol Modification (1 OF 1)
Protocol modified to include cardiocentesis as means of blood collection in snakes.

SUBMITTED TO AND APPROVED BY:

Protocol originally submitted 01/15/2019 – Approved 03/20/2019
Modification Submitted 01/24/20 – Approved 02/06/2020

STATUS:

- Protocol modification was approved.

•. 30 ml of American alligator blood was collected from 2 animals at the St. Augustine Alligator Farm Zoological Park on 08/13/2019 and received at GMU on 08/14/2019.

• 50 ml of American alligator blood was collected from 4 animals at the St. Augustine Alligator Farm Zoological Park on 10/23/2019 and received at GMU on 10/24/2019.

• 32 ml of American alligator blood was collected from two animals at the St. Augustine Alligator Farm Zoological Park on 12/12/2019 and received at GMU on 12/13/2019.

• 2 ml of Boa constrictor blood was collected from one animal at the St. Augustine Alligator Farm Zoological Park on 02/20/2020 and received at GMU on 02/21/2020.

Rationale for modification: The resident veterinarian at the St. Augustine Alligator Farm Zoological Park advised us that an alternative method was preferred for collecting large volumes of blood from constrictor snakes than that which was included in our approved IACUC and ACURO protocol. We requested a modification to our IACUC protocol to include this method, which was approved by GMU IACUC. We then submitted our request for approval of this modification to ACURO, who in turn approved the modification.

Significant changes in use of biohazards and/or select agents

Nothing to report

6. PRODUCTS:

- **Publications, conference papers, and presentations**

Journal publications.

Chowdhury, A., Reehl, S., Kehn-Hall, K., Bishop, B., Webb-Robertson, B. "Better Understanding and Prediction of Antiviral Peptides through Primary and Secondary Structure Feature Importance", *Nature Scientific Reports*, 2020 Nov 6;10(1):19260. doi: 10.1038/s41598-020-76161-8.

Books or other non-periodical, one-time publications.

Nothing to report

Other publications, conference papers and presentations.

Poster Presentations:

Bishop, B. and Kehn-Hall, K. "Bioprospecting Host Defense for New Antiviral Agents" MHSRS, Orlando, FL, August, 2019

Bishop, B., Carfagno, A., Chafran, L., Callahan, V., Lin, S., Akhrymuk, I., Tsogtbayar, D., Po, M., Lamont, S., Webb-Robertson, B., and Kehn-Hall, K. "Reptilian Host Defense and New Antiviral Agents", presented at DTRA CBD S&T conference, Cincinnati, Ohio, November 18-21, 2020

Webb-Robertson, B., Bishop, B., Kehn-Hall, K. "A Biosprospecting Approach to Identify Host Defense Peptides that Target Viruses", presented at ASM Biotreats, Arlington, Virginia, January 28-20, 2020

Data from the project was included in a DTRA Tech Watch Seminar presented by Dr. Bishop in April, 2020.

- **Website(s) or other Internet site(s)**

  The antiviral peptide prediction tool that we developed can be accessed at https://msc-viz.emsl.gov/AVPR/.

- **Technologies or techniques**

  We have developed a novel Feature-Informed Machine Learning approach to AVP prediction (FIRM-AVP) and a manuscript describing our algorithms and their accuracy has been submitted to the journal Scientific Reports. In addition to improving overall accuracy using the Support Vector Machine (SVM) approach versus a Random Forest (RF) or Deep Learning (DL) methods, we have also developed a web-service (https://msc-viz.emsl.gov/AVPR/) to allow users to submit peptide sequences for analysis and predict whether they are likely to have antiviral properties.

- **Inventions, patent applications, and/or licenses**

  FIRM-AVP has been released open source on GitHub (https://github.com/pmartR/FIRM-AVP) and is licensed up the BSD 2-Clause "Simplified" License

- **Other Products**

  Nothing additional to report.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Name: *Barney Bishop*

Project Role: *Principle Investigator*
Researcher Identifier (e.g. ORCID ID): *0000-0002-6626-9251*
Nearest person month worked: *7*
Contribution to Project: *As principle investigator, Dr. Bishop manages the project and coordinates research efforts between the GMU teams and the team at PNNL. He is also directly involved with particle and peptide harvest development as well as analysis of harvested peptides via mass spectrometry.*

Name: *Kylene Kehn-Hall*
Project Role: *co-Principle Investigator*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *5*
Contribution to Project: *As co-principle investigator, Dr. Kehn-Hall has worked closely with Dr Bishop in coordinating research efforts. She is directly involved with microbiology/virology aspects of the project, including cell culture, preparation of virus like particles, blood stimulation and infection studies.*

Name: Bobbie-Jo Webb-Robertson
Project Role: *PNNL co-Principle Investigator*
Researcher Identifier (e.g. ORCID ID): *0000-0002-4744-2397*
Nearest person month worked: *2*
Contribution to Project: *Dr. Webb-Robertson is responsible for performing the statistical aspects of the projects and leads the PNNL efforts. She has been analyzing the technical replicate data.*

Name: Abu Chowdhury
Project Role: *PNNL post-doc*
Researcher Identifier (e.g. ORCID ID): *0000-0003-3454-5861*
Nearest person month worked: *2*
Contribution to Project: *Dr. Chowdhury is responsible for coding and validating the machine learning AVP prediction methods.*

Name: Sarah Reehl
Project Role: *PNNL data scientist*
Researcher Identifier (e.g. ORCID ID): *0000-0003-3727-5801*
Nearest person month worked: *1*
Contribution to Project: *Ms. Reehl was responsible for continuing code development after Dr. Chowdhury finished his post-doc at PNNL and generated the web-service and public respositories.*

Name: *Paul Russo*
Project Role: *Affiliated Faculty*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *3*
Contribution to Project: *Dr. Russo's area of expertise and contribution is in protein/peptide mass spectrometry. He is responsible for analyzing the harvested peptides via LC-MS/MS and developing methods/protocols for that purpose. He is also working with students in Dr. Bishop's lab in analyzing the mass spectrometry data for the harvested peptides.*

Name: *Sabrina Lamont*
Project Role: *Research Assistant*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *6*
Contribution to Project: *Ms. Lamont supported research efforts by Dr. Bishop and his graduate students in the area of particle development and evaluation. She also helped with coordinating research efforts in the lab.*

Name: *Liana Soares Chafran*
Project Role: *Research Associate*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *12*
Contribution to Project: *Dr. Soares is an experienced polymer chemist with experience in materials science and nanotechnology. Dr. Soares has been working with Dr. Bishop and his graduate students to improve particle production and develop new particle technologies for the targeted capture of peptides and proteins of interest from plasma. She has also contributed in performing harvests from Alligator plasma using the hydrogel particles and processing samples for mass spectrometry analysis. She has helped to provide much needed bandwidth in this area. Simultaneously, she is introducing multiple advancements in to the particle harvest technology, enhancing their performance and versatility.*

Name: *Shih-Chao Lin*
Project Role: *Research Associate*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *7*
Contribution to Project: *Dr. Lin is responsible for performing the virological aspects of the projects. He developed the VLP production, plasma membrane isolations and blood stimulation protocols. He has also been responsible for preparing the cell membranes for incorporation in particles. He recently began evaluating the antiviral properties of peptides identified from healthy alligator plasma.*

Name: *Amy Carfagno*
Project Role: *Graduate Student*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *3*
Funding/Support: *GMU Presidential Graduate Fellowship (over the summer received some wages support from the grant for efforts on the project).*
Contribution to Project: *Ms. Carfagno has been supporting research efforts by Dr. Bishop and his graduate students in the area of mass spectrometry data processing and analysis.*

Name: *Victoria Callahan*
Project Role: *Graduate Student*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *2*
Funding/Support: *Volunteer*
Contribution to Project: *Ms. Callahan has been assisting with development of the VLP-based particles, including virus preparation, modification, coupling to particles, and performing harvests.*

Name: *Ivan Akhrymuk*
Project Role: *Post-doctoral Researcher*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *3*
Contribution to Project: *Dr. Akhrymuk has been assisting with development of the VLP-based particles, including virus preparation, modification, coupling to particles, and performing harvests.*

Name: *Samuel Garvey*
Project Role: *Graduate Student*
Researcher Identifier (e.g. ORCID ID):
Nearest person month worked: *3*
Funding/Support: *GMU Graduate Teaching Assistantship.*
Contribution to Project: *Mr. Garvey developed the protocol for synthesizing the PP4ME monomer and is involved with developing hydrogel particles for use in generating virus-modified particles.*

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?

Dr. Kylene Kehn-Hall has received the following active support since the last reporting period:

NIH 1R01AI143817-01A1
Role: Multi-PI (Paige, Klimov, Kehn-Hall)
Project Title: Developing capsid-importin alpha inhibitors for the treatment of VEEV infection
Sponsor: George Mason University (NIH subcontract)
Total Award Amount: Total Award Period: 09/15/2020-08/30/2024
Commitment in Person-Months per Year: 0.75 summer months

Fund #460826
Role: PI
Project Title: Impact of Silicon Nitride on SARS-Cov-2
Sponsor: SINTX Technologies Corporation
Total Award Amount: Total Award Period: 08/10/2020-08/09/2021
Commitment in Person-Months per Year: 0.45 Academic and 0.15 Summer

HDTRA12010015
Role: Subcontract PI (PI: Bradfute)
Project Title: Long Non-Coding Ribonucleic Acids (lncRNA) Role in Pathogenesis
Sponsor: University of New Mexico (DTRA subcontract)
Total Award Amount: Total Award Period: 09/29/2020-09/30/2021
Commitment in Person-Months per Year: 0.45 Academic and 0.42 Summer

What other organizations were involved as partners?

Organization Name: Pacific Northwest National Laboratory (PNNL)

Location of Organization: 902 Battelle Blvd, Richland, WA 99354

Partner's Contribution: Dr. Bobbie-Jo Webb-Robertson, of the Biological Sciences Division at PNNL, is collaborating with us on the project via a subcontract. She is primarily contributing in Major Task 4 efforts and her focus is on statistical analyses of the peptides identified from the harvests in Major Task 3 in order to identify those peptides uniquely associated with harvests performed using virion-modified particles as well as particles encapsulated in membranes from VEEV-infected Leukocytes. These peptides are expected to be potential antiviral and will by synthesized and evaluated.

Organization Name: Virginia Tech

Location of Organization: 1981 Kraft Drive, Blacksburg, VA 24060

Partner's Contribution: Dr. Kylene Kehn-Hall, Co-PI on this project, left George Mason University in August to join the faculty in the Department of Biomedical Sciences and Pathobiology at Virginia Tech. Her laboratory continues to contribute to Major Task 1 and 5 including final optimization studies on the VLP-particles and analysis of peptides for antiviral activity.

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS: *For collaborative awards, independent reports are required from BOTH the Initiating Principal Investigator (PI) and the Collaborating/Partnering PI. A duplicative report is acceptable; however, tasks shall be clearly marked with the responsible PI and research site. A report shall be submitted to* ***https://ers.amedd.army.mil*** *for each unique award.*

QUAD CHARTS: *If applicable, the Quad Chart (available on https://www.usamraa.army.mil) should be updated and submitted with attachments.*

9. APPENDICES: *Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.*

APPENDIX 1.
Quad Chart

Quad Chart provided on following page

Bioprospecting for Antiviral Peptides

ERMS/Log Number: Log#DM171301 and Task Title: Accelerating Innovation in Military Medicine (AIMM).
Award Number: W81XWH-18-1-0801

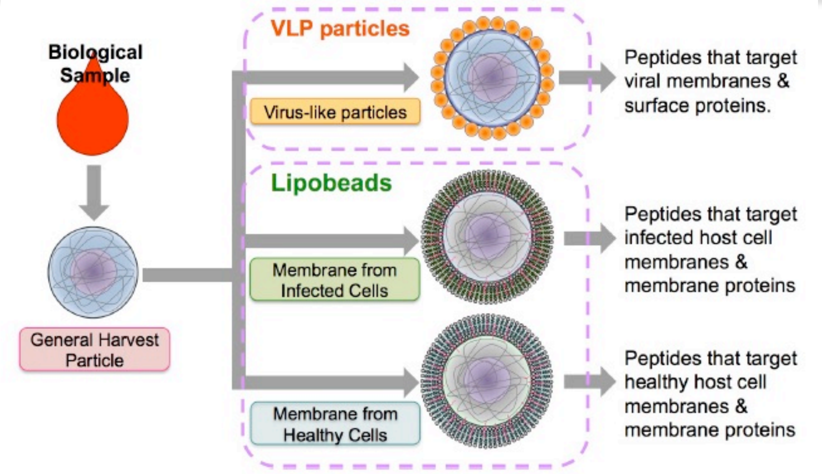
PI: Barney Bishop **Org:** George Mason University **Award Amount:** $506,750

Study/Product Aim(s)

- •Aim 1. Develop microparticles for capturing antiviral peptides.
- •Aim 2. Capture, analysis and identification of peptides from reptile plasma.

Approach

Reptiles (e.g. snakes and alligators) have been implicated as potential reservoirs for arboviruses. We propose to develop a hydrogel microparticle-based peptide enrichment strategy targeting antiviral peptides coupled with mass spectrometry and analytical tools to identify those captured peptides that are likely to be antiviral. We will focus on VEEV as a model arbovirus, and will apply our technology to identify anti-VEEV peptides from the blood of reptiles that has been stimulated using VEEV.



We proposed a new strategy based on differential enrichment of bioactive peptides using hydrogel particles to identify antiviral peptides in reptiles. Accomplishments: A) Performed harvests from plasma and cell culture media after stimulation. B) Completed analyses of peptides from multiple harvests. C) Identified anti-VEEV peptide candidates. D) Developed improved antiviral peptide prediction algorithm.

Timeline and Cost

Activities **CY**	**18***	**19**	**20**
Task 1: Dev. VLP-based capture particle.	🟩	🟩	
Task 2: Dev. cell-based capture particle.	🟩	🟩	
Task 3: Perform plasma harvests.		🟩	🟩 NCE
Task 4: Analyze harvested peptides.		🟩	🟩 NCE
Task 5: Assess anti-VEEV performance.		🟩	🟩 NCE
Estimated Budget ($K)	**$105**	**$318**	**$84**

* Start 09/30/2018

Updated: (02/26/2021)

Goals/Milestones

CY19 Goal–Develop microparticles for capturing antiviral peptides.

☑ GMU IACUC and ACURO Approval.
☑ Synthesize Click-monomer and microparticles.
☑ Synthesize microparticles encapsulated in cell membranes.
☑ Synthesize VLP-modified microparticles.

CY20 Goal–Capture and analysis of peptides from reptile plasma.

☑ Perform harvests from VEEV stimulated and unstimulated plasma.
☑ Develop workflow, statistical analyses, and predictive models.
☑ Identify and assess potential reptile anti-VEEV peptides.

Comments/Challenges/Issues/Concerns

- Completed tasks and deliverables by October 31, 2020.
- Developed bioprospecting process for antiviral peptide discovery.
- Developed improved antiviral peptide prediction machine learning algorithms.
- We identified two novel peptides with anti-VEEV properties.

Budget Expenditure to Date

Projected Expenditure: $506,750
Actual Expenditure: $506,750

APPENDIX 2.
Bibliography of Publications

Journal Publications.

Chowdhury, A., Reehl, S., Kehn-Hall, K., Bishop, B., Webb-Robertson, B. "Better Understanding and Prediction of Antiviral Peptides through Primary and Secondary Structure Feature Importance", *Nature Scientific Reports*, 2020 Nov 6;10(1):19260. doi: 10.1038/s41598-020-76161-8.

Conference Papers and Presentations.

Poster Presentations:
Bishop, B. and Kehn-Hall, K. “Bioprospecting Host Defense for New Antiviral Agents” MHSRS, Orlando, FL, August, 2019

Bishop, B., Carfagno, A., Chafran, L., Callahan, V., Lin, S., Akhrymuk, I., Tsogtbayar, D., Po, M., Lamont, S., Webb-Robertson, B., and Kehn-Hall, K. “Reptilian Host Defense and New Antiviral Agents”, presented at DTRA CBD S&T conference, Cincinnati, Ohio, November 18-21, 2020

Webb-Robertson, B., Bishop, B., Kehn-Hall, K. “A Biosprospecting Approach to Identify Host Defense Peptides that Target Viruses”, presented at ASM Biotreats, Arlington, Virginia, January 28-20, 2020

Data from the project was included in a DTRA Tech Watch Seminar presented by Dr. Bishop in April, 2020.

APPENDIX 3.
Personnel Supported by Grant

1. Barney Bishop, PhD
2. Kylene Kehn-Hall, PhD
3. Bobbie-Jo Webb-Robertson, PhD
4. Abu Chowdhury, PhD
5. Sarah Reehl, PhD
6. Paul Russo, PhD
7. Sabrina Lamont
8. Liana Soares Chafran, PhD
9. Shih-Chao Lin, PhD
10. Amy Carfagno
11. Ivan Akhrymuk