AWARD NUMBER: W81XWH-21-1-0483

TITLE: NF1-Associated Peripheral Nerve Sheath Tumors at Single-Cell Resolution: Heterogeneity, Tumor Growth, and Malignant Progression

PRINCIPAL INVESTIGATOR: EDUARD SERRA

CONTRACTING ORGANIZATION: Fundacio Institut d'Investigacio Germans Trias i Pujol
Carretera de Can Ruti. Camí de les Escoles s/n
08916 Badalona (Barcelona) Spain

REPORT DATE: July 2022

TYPE OF REPORT: Annual

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE:**	**3. DATES COVERED**
July 2022	**Annual**	**25Jun2021 – 24Jun2022**

4. TITLE AND SUBTITLE

NF1-Associated Peripheral Nerve Sheath Tumors at Single-Cell Resolution: Heterogeneity, Tumor Growth, and Malignant Progression

5a. CONTRACT NUMBER
W81XWH-21-1-0483

5b. GRANT NUMBER
NF200051

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Eduard Serra, PhD
Bernat Gel, PhD
Meritxell Carrió, PhD

E-Mail: eserra@igtp.cat

5d. PROJECT NUMBER
0011613883

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Fundacio Institut d'Investigacio En Ciencies de la Salut
Germans Trias i Pujol
Carretera de Can Ruti. Camí de les Escoles s/n
08916 Badalona (Barcelona) Spain

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

Background

The development of tumors of the peripheral nervous system (PNS) represents a major problem for persons with Neurofibromatosis type 1. Plexiform neurofibromas (pNFs) constitute a major source of morbidity. pNFs arise during development through the inactivation of the NF1 gene in a cell of the neural crest (NC) - Schwann cell (SC) lineage. In some cases, pNFs may undergo malignant transformation towards an aggressive and highly metastatic Malignant Peripheral Nerve Sheath Tumor (MPNST), normally through the previous development of a pre-malignant nodule termed atypical neurofibroma (aNF). Genomic analyses of pNF-aNF-MPNST progressions have demonstrated that these tumors share the same somatic NF1 inactivation, linking their cells of origin. We don't know whether within pNFs there are remaining cells with the same biological properties as the originating pNF cell, and if so, which role they might play in pNF growth, tumor progression or response to treatment. Our preliminary single cell RNA-seq data from different pNFs confirmed the diversity of cell types present in pNFs, and revealed the existence of cell subpopulations within specific cell type components, a previously unnoticed heterogeneity. The pNF SC component seems to contain at least two distinct groups of SCs, one expressing exclusively markers of precursor SCs (SCPs) and another group expressing in addition markers of SC commitment. There is also heterogeneity in the endoneurial fibroblast-like stromal (FB) component, with some subpopulations expressing key mesenchymal transcription factors also identified in MPNST cells. The FB component might play an important role in pNF growth. We have recently generated human neurofibroma-like tumors in mice by engrafting 3D spheroids in their sciatic nerve. Only when these spheroids contain NF1(-/-) iPSC-derived differentiating SCs plus primary endoneurial FBs neurofibroma-like tumors consistently develop.

Hypothesis

We hypothesize that some of the pNF cell populations expressing SCP markers could act as a pNF stem-like cell. We think these cells might have an important impact on pNF growth, progression to malignancy and response to treatment. In addition, we hypothesize that pNF FBs provide trophic and niche conditions essential to maintain this pNF stem-like cell population. Finally, a fine characterization of expression and mutation status at single cell level may uncover the cells undergoing tumor progression. These cells need to be present in faithful models to help characterize the impact of drug treatments on different tumor cell subpopulations.

Specific Aims

1) To generate a shared resource consisting in a comprehensive cell diversity map of the cellular composition, spatial distribution and genomic content of NF1-associated peripheral nerve sheath tumors (pNFs, aNFs and MPNSTs). 2) To biologically characterize and elucidate the contribution of specific subpopulation of cells of the SC and FB components, on pNF growth. 3) To identify and characterize the identity of the cell type within pNFs progressing towards aNF and MPNSTs. Analyze the presence of these cells in a human iPSC-based in vitro/in vivo 3D tumor model, and monitor the impact of drug treatment on their viability.

Study design

We will generate a resource on single cell information of pNFs, aNFs and MPNSTs available to the scientific community. This resource will provide a comprehensive map at single cell resolution, combining four layers of information: gene expression (scRNA-seq), genetic and genomic content (target scDNA-seq), spatial distribution (spatial transcriptomics) and functional properties. Single cell information will be used to separate and functionally characterize (potency, trophic and niche capacities) specific subpopulations of cells within pNFs cell components. The identification of somatic mutations (NF1, CDKN2A) will be used to trace cells progressing from pNF to aNF and MPNST. We will characterize our iPSC-based 3D tumor model system for the presence of these cells that will be used to monitor tumor response to Selumetinib in a pilot treatment.

Impact

The present work may change the way we understand the cell composition and heterogeneity of pNFs. We will better understand the impact of specific cell subpopulations on pNF growth and progression to MPNST. Translating this cell composition information to faithful tumor models may provide a way to analyze tumor response to treatment based on the impact on specific cell

15. Subject Terms None listed.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					USAMRDC
a. REPORT	b. ABSTRACT	c. THIS PAGE	Unclassified	33	19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified			

TABLE OF CONTENTS

1. INTRODUCTION

The development of tumors of the peripheral nervous system (plexiform neurofibromas (pNFs); atypical neurofibromas (aNF); malignant peripheral nerve sheath tumor (MPNST)), represents a major clinical problem for persons with Neurofibromatosis type 1 (NF1). Plexiform neurofibromas (pNFs) constitute a major source of morbidity and can progress towards an aggressive and highly metastatic MPNST. The cell originating these tumors and their cellular composition are key aspects to understand how these tumors grow, progress and respond to therapies. In the present project we propose to perform an integrative analysis of different genomic techniques at a single cell level of pNFs, aNFs and MPNSTs, comparing them to nerves, their tissue of origin. We will generate a shared resource consisting in a comprehensive cell diversity map of the cellular composition, spatial distribution and genomic content of NF1-associated peripheral nerve sheath tumors. We will elucidate the contribution of specific subpopulations of cells within these tumors to their growth and will characterize the identity of the cell type within pNFs progressing towards aNF and MPNSTs. Finally, we will analyze the impact of drug treatments on specific subpopulations of cells using a human iPSC-based *in vitro*/*in vivo* 3D tumor model developed in our lab.

2. KEYWORDS

Plexiform neurofibroma; atypical neurofibroma; malignant peripheral nerve sheath tumor (MPNST); nerve; Schwann cell; fibroblast; cell-of-origin; cell heterogeneity; single cell analysis; genomics; bioinformatics

3. ACCOMPLISHMENTS

Major goals of the project within months 1-12

As stated in the SOW

Specific Aim 1: To generate a shared resource consisting in a comprehensive cell diversity map of the cellular composition, spatial distribution and genomic content of NF1-associated peripheral nerve sheath tumors (pNFs, aNFs and MPNSTs) and adult nerves	**Timeline (Months)**	**Site 1 GSU**
Major Task 1: Expression analysis: bulk RNAseq, scRNAseq, spatial transcriptomics		
Subtask 1: HRPO review and approval	1-4	
Subtask 2: Bulk RNAseq from primary tumors and nerves Samples used: Human samples of pNFs, aNFs/ANNUBP*, MPNSTs, adult nerve (collected and preserve in our IRB-approved collection of samples number C.0002242 of the National Biobank Registry of Instituto de Salud Carlos III, Spain). (*) aNFs/ANNUBP are partially being collected prospectively. We cannot guarantee the proposed timeframe for this specific tumor type. At least 4 samples of each. R1, R2	4-8	X
Subtask 3: Bulk RNAseq bioinformatic analysis	8-12	X
Subtask 4: scRNAseq from primary tumors and nerves Samples used: Same samples used for Subtask 1 At least 4 samples of each. R1, HH	4-8	X
Subtask 5: scRNAseq bioinformatic analysis	8-12	X
Subtask 6: Smart-seq from primary tumors and nerves Samples used: Same samples used for Subtask 1 At least 4 samples of each. R1, R2, HH	4-12	X
Subtask 7: Smart-seq bioinformatic analysis	8-18	X

What has been accomplished

Specific Aim 1: To generate a shared resource consisting in a comprehensive cell diversity map of the cellular composition, spatial distribution and genomic content of NF1-associated peripheral nerve sheath tumors (pNFs, aNFs and MPNSTs) and adult nerves	**Timeline (Months)**	**Site 1 GSU**
Major Task 1: Expression analysis: bulk RNAseq, scRNAseq, spatial transcriptomics		
Subtask 1: HRPO review and approval	1-4	
Achieved		
Subtask 2: Bulk RNAseq from primary tumors and nerves Samples used: Human samples of pNFs, aNFs/ANNUBP*, MPNSTs, adult nerve (collected and preserve in our IRB-approved collection of samples number C.0002242 of the National Biobank Registry of Instituto de Salud Carlos III, Spain). (*) aNFs/ANNUBP are partially being collected prospectively. We cannot guarantee the proposed timeframe for this specific tumor type. At least 4 samples of each. R1, R2	4-8	X
Partially achieved We have performed bulk RNAseq from 4 pNFs and 4 MPNSTs We have collected 3 aNF. As stated in the original SOW, these tumors are difficult to obtain. Accordingly, we have prioritized to perform scRNA-seq and SMART-seq analysis, before bulk RNA-seq, since this information was the most crucial to obtain for the project. Now that we confirmed that both single cell techniques work correctly in this type of samples, we will use the remaining parts of the tumors to perform bulk RNAseq. We have been collecting nerve samples and setting up conditions to dissociate them in order to perform at the same time both bulk RNAseq and single cell RNAseq (see subtask 4 for protocols details).		
Subtask 3: Bulk RNAseq bioinformatic analysis	8-12	X
Partially achieved. We have designed and implemented the main RNAseq data analysis pipeline. We have processed the first 8 available samples. We are designing and implementing the analysis strategy for the integration of RNA-seq and single cell-based technologies		

Subtask 4: scRNAseq from primary tumors and nerves Samples used: Same samples used for Subtask 1 At least 4 samples of each. R1, HH	4-8	X
Partially achieved Preparing the samples for single cell gene expression analysis is critical for maintaining cell integrity and preserving each cell's transcriptome. It is critical to obtain a well-cingulated cell suspension free of cell debris, with minimal cell aggregates and high viability (>70%). In general, cell preparation protocols vary depending on the tissue of origin. Each tissue type is unique and thus, it is critical to optimize sample preparation before starting single cell experiments. We have set up conditions and established robust protocols to process all tumor types (pNFs, aNFs and MPNSTs) and also nerves, for single cell RNA-seq analysis, for either 10x scRNA-seq and SMART-seq. These protocols include a dissociation step specific for each type of tissue that maintains cell integrity, a removal of cell debris and aggregates step by using the Death Cell Removal kit (Miltenyi Biotec); a removal of myelin step for nerve samples (Miltenyi Bitoec), and a final confirmation step of high cell viability using Flow cytometry. We below the section "Developed Protocols". We prioritized to set up the whole single cell analysis workflow (sample dissociation; single cell preparation; library constructions; sequencing; bioinformatic analysis) of few samples of different tumor/tissue types first, instead of performing sample preparation and sequencing for all samples before checking that the whole workflow was performing as expected. We performed scRNAseq from 2 pNFs, 3 aNFs and 3 MPNSTs, and are currently performing the remaining samples, that will be finished at the end of the summer.		
Subtask 5: scRNAseq bioinformatic analysis	8-12	X
Partially achieved As stated above, we prioritized having produced and analyzed a small number of pilot samples for the complete workflow, rather than performing only the sequencing part for all the samples upfront. Therefore, although we implemented the whole scRNAseq data analysis pipeline, we put a special emphasis on the quality control (QC) and batch effect analysis code. In particular: We have designed and implemented the infrastructure for **scRNAseq data analysis** including a data analysis pipeline with: - Standard well tested preprocessing and QC steps taking into account the most frequent error and noise sources found from initial library		

preparation to the final sequencing.

- Additional extended QC for the detection of cellular stress prior to library preparation. We have identified cellular stress as a major noise source in a few samples and developed a new bioinformatics methodology to identify stressed cells and samples and correct the problem. Addition information can be found below in the section entitled "Extended QC for stress detection and removal".

- The analysis of batch effects and the application of corrective batch effect removal measures to ensure full comparability between all samples in the project. Our results so far show differences between individual samples but not strong batch effects.

- Single-cell level cell-type identification with a two-tier cell-type assignment and analysis of sample-level cell-type proportions. We are still testing different reference single cell atlases to identify the one performing the best.

- Extensive gene expression analysis, from cluster-specific marker gene identification to individual gene expression evaluation and differential expression analysis between any groups of cells from the same or different samples.

- Gene-set expression analysis for the integrated analysis of the expression patterns of complete sets of genes.

- A specialized method based on multiple gene-set expression analysis to identify cells at different differentiation stages along the Neural Crest-Schwann Cell axis, from neural-crest to mature Schwann cells.

As it will be explained in the section "Changes/Problems", in order to make this single cell data set even more complete, we decided to include single cell epigenetic information generated in the context of another project, but for the same samples analyzed here.

We have designed and implemented the infrastructure for scATAC-seq data analysis including:

- Preprocessing, mapping, basic peak calling and integration of multiple samples.

- Additional peak refinement and peak-gene assignment.

- Differential accessibility analysis between any group of cell from the same or different samples.

- Analysis of enhancer activation/deactivation based on the accessibility of known enhancers.

- ATAC-based clustering and identification of cluster-level accessibility markers

- Integration of scATAs-eq with scRNA data including gene-level accessibility analysis based on cell clusters defined from scRNAseq data In addition to this, we have set up a specific pipeline for RNA-velocity based on spliced/unspliced transcripts detected in scRNAseq data and the integration of these results with the main scRNAseq and scATACseq pipelines. Finally, we are working on the estimation of cell-level broad copy-number profiles based on scRNAseq data, especially useful for atypical neurofibromas and MPNST. A short summary of a few analysis results and very-preliminary findings can be found below in the section entitled "Observations derived from the preliminary analysis of 4 pNFs and 3 aNFs".		
Subtask 6: Smart-seq from primary tumors and nerves Samples used: Same samples used for Subtask 1 At least 4 samples of each. R1, R2, HH	4-12	X
Partially achieved The samples processed for scRNAseq have been prepared at the same time for SMARTseq analysis. We have established a protocol to FACS-sort the dissociated cells from Subtask 4 in 96-well plates containing the lysis buffer from the SMARTseq preparation kit, provided by CNAG. So far, all quality controls are OK. The quantity of cDNA obtained is adequate and at a good concentration and size. Library preparation has been performed. We have performed SMARTseq from 2 pNFs, 3 aNFs and 3 MPNSTs		
Subtask 7: Smart-seq bioinformatic analysis	8-18	X
Partially achieved We have designed and set up a very basic data analysis pipeline, but so far we have not been able to test it with our own data and therefore identify specific challenges or steps that need to be fine-tuned.		

Serra Lab Protocols IGTP	Version 1 Prepared by: H. Mazuelas Date: January 2020	Version 2 Revised by: M.Carrió Date: July 2022

SOP nº8

PNF, ANNUBS (ANFs) and MPNST dissociation for scRNAseq

Objective

To dissociate PNFs, ANNUBS (ANFs) and MPNSTs for single cell RNAseq analysis

Materials and Reagents

Collagenase Type I (Worthington Cat# LS004196)
Neutral Protease (Dispase grade I) (Worthington Cat# LS02104)
Normocin cocktail (Invivogen Cat# ANT-NR-1)
DMEM (Gibco 2196035)
FBS (Gibco 10270106)
GlutaMax (Glx) (Gibco 35050038)
Penicillin/Streptomycin (P/S) (Gibco 15140122)
Dead Cell Removal Kit (Miltenyi Biotec 130-090-101)
MACS MS columns (Miltenyi Biotec 130-042-201)
MACS Multistand (Miltenyi Biotec 130-042-303)
MACS separator (Miltenyi Biotec)

Methodology

Note: Samples need to be delivered at CNAG before 13:00h

To prepare:

- Dissociation media :
  160 U/ml Collagenase Type I
  *dilution 1:2 Collagenase lot 49HI9435 for PNF and ANNUB
  *dilution 1:5 Collagenase lot 49HI9435 for MPNST
  0.8 U/ml Neutral Protease (Dispase grade I)
  1x Normocin cocktail
  DMEM 10%FBS +1xGlx +1xP/S
  Mix and filter using a sterile 0,22µm PES filter unit

  10 mL 1:2 Collagenase (PNF/ANNUB)
  0,5 mL Collagenase type I (1600 U/mL stock)
  260µl Dispase (30 U/mL stock)
  20µl normocin
  9,2 mL DMEM 10%FBS +1xGlx +1xP/S
  Mix and filter using a sterile 0,22µm PES filter unit

10 mL 1:5 Collagenase (MPNST)
0,2 mL Collagenase (1600 U/mL stock)
260µl De dispase (30 U/mL stock)
20µl normocin
9,5 mL DMEM 10%FBS +1xGlx +1xP/S
Mix and filter using a sterile 0,22µm PES filter unit

- Binding buffer 1X from Dead Cell Removal Kit (in cell culture H_2O; prepare 3,5mL/sample + 0,5mL more to calibrate the column)
    - 10 mL binding buffer: 500 µl 20x binding buffer + 9,5 mL H_2O
- Prepare 15 ml conical centrifuge tubs to collect samples, SN1, SN2 for each tumor to dissociate
- Prepare 15 ml conical centrifuge tubs to collect alive cells (negative fraction)
- Prepare 15 ml conical centrifuge tubs to collect death cells (positive fraction. OPTIONAL)
- Prepare 1,5ml Low Binding Eppendorf's to aliquot cells to take to CNAG

Protocol:

Dissociation

- Transfer the piece of tumor into a sterile culture dish containing some drops of dissociation media and cut it into 2 mm x 2 mm pieces.
- Place 4-8 neurofibroma pieces per well of a 6-well plate and add 2 mL of dissociation buffer and incubate in a standard CO_2 incubator at 37ºC, 5% CO_2 for 16h
- Next day, mechanically dissociate the neurofibroma tissue into a suspension using a glass Pasteur pipette. Observe the sample under the microscope:
    - If there are aggregates, place the sample at 37 ºC for an extra hour.
    - If the sample is in single cell suspension proceed with the protocol
- Transfer the neurofibroma suspension into a 15 mL conical centrifuge tube and add 8 mL complete DMEM to stop the activity of dissociation enzymes.
- Centrifuge at 1400 rpm 5 min.
- Recover supernatant (SN1) and repeat the centrifugation step. Keep the pellet of cells apart.
    - The reason of this step is because sometimes many cells remain in the supernatant.
- Centrifuge SN1 at 1400 rpm 5 min.
    - If the supernatant 2 (SN2) is still cloudy, you can centrifuge SN2.
- Resuspend pellet 1 and pellet 2 in 0,5-1 mL (depending on the pellet) in DMEM 10% FBS +1xGLx + 1x P/S
- Count cells

    IMPORTANT: For PNFs that they contain a lot of fiber, if the number of cells counted is more than 3M, use two columns in the following steps to prevent collapsing them.
    IMPROTANT: If the cell pellet is very small, DO NOT count the cells and go straight to perform Death Cell Removal.
- Centrifuge at 1200 rpm 5 min
- Discard supernatant using a pipette tip.

Dead Cell Removal Kit (Miltenyi Biotec 130-090-101)

- Add 100 µl Dead Cell Removal Microbeads (enough for 10M cells).
- Incubate 15 min at room temperature (protected from light).

Preparation of MS MACS columns MACS

- Mount columns in the MACS separator attached to the MACS Multistand (avoid position 1)
- Equilibrate columns with 500 µl Binding buffer 1x

Magnetic separation

- Add 500µl Binding buffer 1x to the sample containing the Microbeads
- Apply cell suspension into the column

  IMPORTANT: if the column collapses, take the column and transfer it to a new equilibrated column.
- Collect alive cells (negative fraction) into a 15 mL conical centrifuge tube
- Wash the column with 4 x 500 µl Binding Buffer 1x
- Count alive cells: take 10 µl from the total volume of 2,5 mL without making any dilution
- Centrifuge alive cells at 1200 rpm 5 min
- Resuspend cells at concentration of 1.000 cells/ µL (1.000.000 cells/ mL) in **DMEM+10%FBS+Glx w/o antibiotic**.
    - Optional: Take column from the Multistand, add 1mL of 1x binding buffer and elute positive fraction of cells (death cells) into a 15 mL tube. Observe cells under a Neubauer chamber.

Sample preparation to take to CNAG

- Divide the sample for scRNAseq_ATAqseq (Multiome) (CNAG) and SMARTseq (FACS-sorting) into two low binding Eppendorf's:
    - Multiome: 0,5-1M cells
    - SMARTseq: Minimum 50.000 cells
- Place sample on ice and go:
    - Multiome: to CNAG straight away
    - SMARTseq: start de staining with DAPI to sort the cells (See SOP9).

SOP nº9
Sample preparation for SMARTseq

Objective

To prepare single-cell samples (from SOP8, SOP10) and FACS-sort them in 96 well plates for SMARTseq.

Materials and Reagents

SMARTseq2 plates (provided by CNAG)
DMEM 10%FBS 1xGlx No P/S
PBS
DAPI

Methodology

Tissues are dissociated for scRNAseq analysis following specific protocols (SOP8, SOP10)

Cells are resuspended at a concentration de 1.000 cells/µL (1M cells/ml) in DMEM+10%FBS+Glx -P/S

- Divide the sample in two:
    - For Multiome: 0,5-1M cells
    - For SMARTseq: Minimum 50.000cells

SMARTseq cell preparation:
Take an aliquot of cells and stain them with DAPI to asses cell viability
Ideal situation:

100.000 cells or more:
10.000 DAPI staining to assess viability
The rest for sorting

50.000 cells:
10.000 DAPI
The rest for sorting

Keep the rest of the cells to be sorted on ice
DAPI staining:

- Prepare DAPI solution (1:1000 in PBS)
- Centrifuge cells 300g 3min RT
- Add 100 µL DAPI solution
- Incubate 10 min at RT
- Add 900 µL PBS
- Centrifuge cells 300g 3 min RT
- Resuspend cells in 200 µL DMEM+10%FBS+Glx-P/S

Go to cytometry facility

Prepare plate for sorting

- Thaw SMART-seq2 plates on ice (follow CNAG instructions for preparing the plate).
- NAME THE PLATE:
    - Ex: SERRA_14 P2 DATA
        - SERRA_14 (project name) P2 (plate 2) Date
- Centrifuge the plate 2000 rpm 1min 4ºC (use centrifuge located at genomic facility)

- Keep plate on ice until sorting

Sort cells:

- Analyze cell viability using DAPI-stained cells
- Remove plate cover from the plate
- Sort non-stained cells into SMART-seq2 plates (follow CNAG suggestions for plate layout).
- Place a new plate cover (provided by CNAG)
- Once the sorting is finished, centrifuge the plate 2000 rpm 1 min 4ºC
- Snap freeze the plate on dry ice
- Store at -80 ºC until shipment to CNAG

SOP nº10

Nerve dissociation for scRNAseq

Objective

To dissociate nerves for single cell RNAseq analysis.

Materials and Reagents

Collagenase Type I (Worthington Cat# LS004196)
Neutral Protease (Dispase grade I) (Worthington Cat# LS02104)
Normocin cocktail (Invivogen Cat# ANT-NR-1)
DMEM (Gibco 2196035)
FBS (Gibco 10270106)
GlutaMax (Glx) (Gibco 35050038)
Penicillin/Streptomycin (P/S) (Gibco 15140122)
Dead Cell Removal Kit (Miltenyi Biotec 130-090-101)
MACS MS columns (Miltenyi Biotec 130-042-201)
MACS Multistand (Miltenyi Biotec 130-042-303)
MACS separator (Miltenyi Biotec)
Myelin Removal kit (Milteny Biotec 130-096-731)
70 µm Flowmi Cell strainer (Sigma BAH136800070)

Methodology

NOTE: Samples need to be delivered at CNAG before 13:00h

To prepare:

- Dissociation media:
  160 U/ml Collagenase Type I
  0.8 U/ml Neutral Protease (Dispase grade I)
  1x Normocin cocktail
  DMEM 10%FBS +1xGlx +1xP/S
  Mix and filter using a sterile 0,22µm PES filter unit

  10 mL
  1 mL Collagenase (1600 U/mL stock)
  260 µL Dispase (30 U/mL stock)
  20µl normocin
  9,2 mL DMEM 10%FBS +1xGlx +1xP/S

- Wash buffer for Myelin Removal beads: PBS + 0,5% FBS (COLD)

- Binding buffer 1X from Dead Cell Removal Kit (in cell culture H_2O; prepare 3,5mL/sample + 0,5mL to calibrate the column)
    - 10 mL binding buffer: 500 µl 20x binding buffer + 9,5 mL H_2O
- Prepare 15 mL conical centrifuge tubs to collect life cells (negative fraction)
- 1,5ml Low Binding Eppendorf's to aliquot cells to take to CNAG

Protocol:

Dissociation

- Transfer a piece of peripheral nerve into a sterile tissue culture plate with one drop of dissociation media.
- Cut the nerve longitudinally and try to pull the fascicles out of the epineurium.
- Cut the fascicles into 2 mm x 2 mm fragments and transfer them into a 6-well plate with 2 mL of dissociation media per well.
- Incubate at 37ºC 16 h.
- Next day, mechanically dissociate the nerve tissue into suspension using a glass Pasteur pipette. Observe the sample under the microscope:
    - If there are aggregates, place sample at 37 ºC for an extra hour
- Repeat this procedure every 30 min until single cell suspension is obtained (around 17-18 h a 37 ºC with Pasteur pipette help).
- Transfer the sample suspension into a 15 mL conical centrifuge tube and add 8 mL complete DMEM to stop the activity of dissociation enzymes.
- Centrifuge at 1400 rpm 5 min.
- Recover supernatant (SN1) and repeat the centrifugation step. Keep the pellet of cells apart.
    - The reason of this step is because sometimes many cells remain in the supernatant.
- Centrifuge SN1 at 1400 rpm 5 min
    - If supernatant 2 (SN2) is still cloudy, you can centrifuge SN2
- Resuspend pellet 1 and pellet 2 in 0,5-1 mL (depending on the pellet) in complete DMEM.
- Filter sample using a 70 µm Flowmi Cell strainer
- Count cells using a hemocytometer
- **IMPORTANT:** If the cell pellet is very small, we DO NOT count the cells and go straight to perform Myelin Removal.
- Centrifuge 1400 rpm 5 min

Myelin Removal

- Add 180µl buffer (PBS + 0,5%FBS) COLD + 20µl myelin removal beads to the cell pellet.
- Incubate 15 min 4 ºC (protected from light).

Preparation of MS MACS columns MACS

- Mount columns in the MACS separator attached to the MACS Multistand (avoid position 1)
- Equilibrate columns with 500 µl Binding buffer 1x

Magnetic separation

- Add 2 mL cold buffer to the cells and beads.
- Centrifuge 1400 rpm 5 min.
- Aspirate supernatant.
- Resuspend in 1 mL of buffer.
- Apply cell suspension into the column.
- Wash the column with 4 x 500 µl Buffer

Dead Cell Removal (Milteny Biotec 130-090-101)

- Centrifuge at 1200 rpm 5 min
- Discard supernatant using a pipette tip.
- Add 100 µl Dead Cell Removal Microbeads (enough for 10M cells).
- Incubate 15 min at room temperature (light off).

Preparation of MS MACS columns MACS

- Mount columns in the MACS separator attached to the MACS Multistand (avoid position 1)
- Equilibrate columns with 500 µl Binding buffer 1x

Magnetic separation

- Add 500µl Binding buffer 1x to the sample containing the microbeads
- Apply cell suspension into the column

  IMPORTANT: if the column collapses, take the column and transfer it to a new equilibrated column.
- Collect alive cells (negative fraction) into a 15 mL conical centrifuge tube
- Wash the column with 4 x 500 µl Binding Buffer 1x
- Count alive cells: take 10 µl from the total volume of 2,5 mL without making any dilution
- Centrifuge alive cells at 1200 rpm 5 min
- Resuspend cells at concentration of 1.000 cells/uL(1.000.000 cells/mL) in **DMEM+10%FBS+Glx w/o antibiotic.**
    - Optional: Take column from the magnetic rack, add 1mL of 1x binding buffer and elute positive fraction of cells (death cells) into a 15 mL tube. Observe cells under a Neubauer chamber.

Sample preparation to take to CNAG

- Divide the sample for scRNAseq_ATAqseq (Multiome) (CNAG) and SMARTseq (sorting) into two low binding Eppendorf's:
    - Multiome: 0,5-1M cells
    - SMARTseq: Minimum 50.000 cells
- Place sample on ice and go:
    - Multiome: to CNAG straight away
    - SMARTseq: start de staining with DAPI to sort the cells (See SOP9).

Extended QC for stress detection and removal

Quality control (QC) is an integral part of any data analysis pipeline. In the case of scRNAseq data, there exist extensive QC options taking into account most frequent sources of noise and technical variability in the process from library preparation to sequencing, that is, from the moment the sample enters the Chromium Controller to the generation of sequencing data. This QC and normalization steps include adjusting for the total sequencing depth of each sample, detecting and removing empty GeMs, detecting and removing doublets, detecting and removing GeMs with dying cells or just cell death debris and many other factors affecting the quality of the scRNAseq data.

However, in our datasets we identified a different source of sample variability that could greatly affect the analysis and conclusions: cellular stress. We detected that some of the samples presented a subset of cells with very high stress levels, in our case Endoplasmatic Reticulum stress. This stress had such an impact on the the global transcriptional profile of these cells that dominated their positioning in the dimensionality reduction plots.

For example when analyzing the first three PNF samples, we can see some regions with only cells from PNF19 (Figure 1). This has the effect of creating additional clusters that could be interpreted as different cell types or the tumor having cells in a special state.

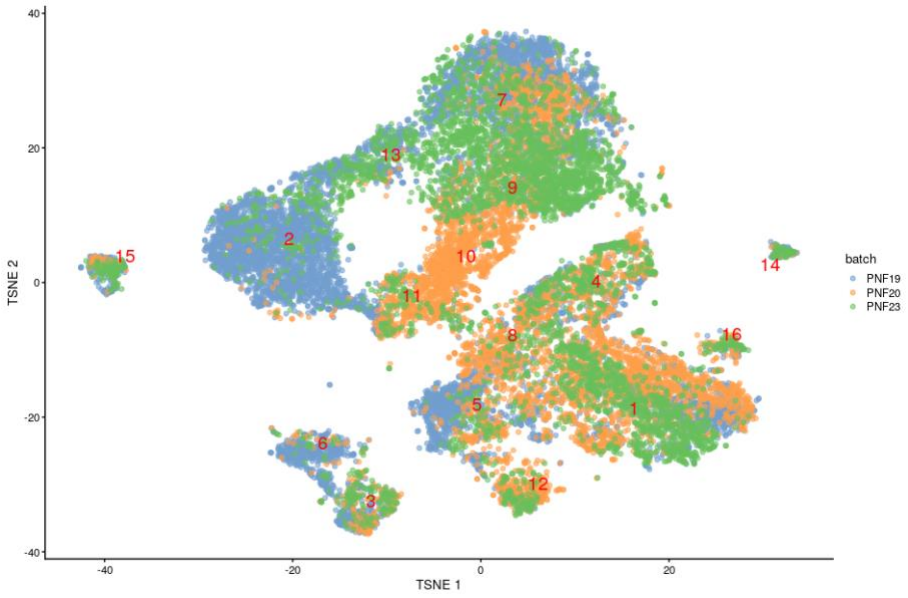


Figure 1 – Stress affects the positioning of cells in the dimensionality reduction plots. *We can see how some parts of the TSNE plot show cells from the PNF19 (blue) and create new additional cell clusters (red numbers).*

We developed a system leveraging stress-related gene-sets from gene ontology with the possibility of computing a score for each cell describing the expression level of a set of genes. With this, we can identify the stress levels of each cell and plot them in the dimensionality reduction plots.

In the example above, we can see how each of the regions containing only PNF19 cells is highly enriched in stressed cells (Figure 2).

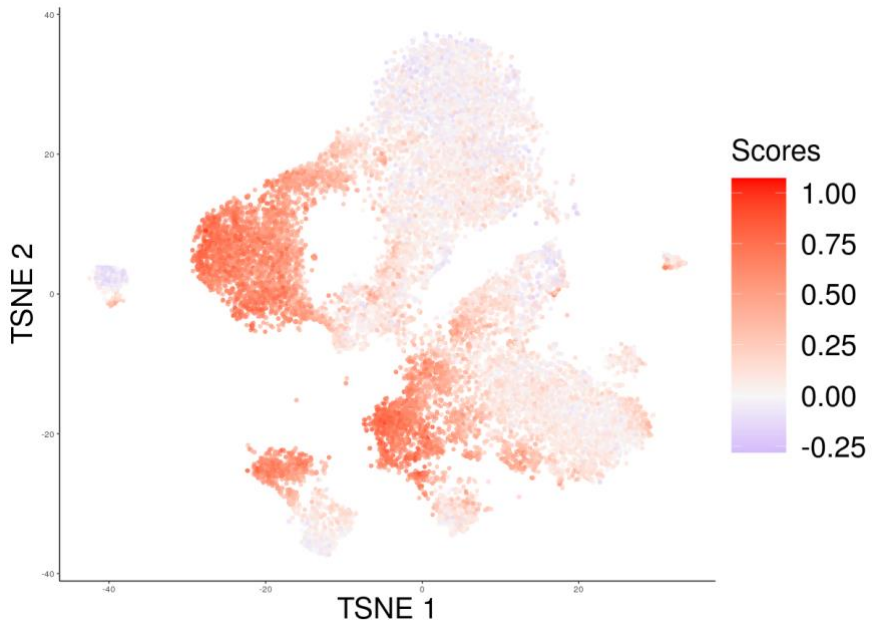


Figure 2 – Stress levels at single cell resolution. *The expression score of the set of genes associated to endoplasmatic reticulum stress shows a concentration of high scoring (stressed, in red) cells in the same regions that contained only PNF19 cells.*

We studied the distribution of stress scoring for each sample and saw how PNF19 had a significant accumulation of cells with high stress scores. (Figure 3)

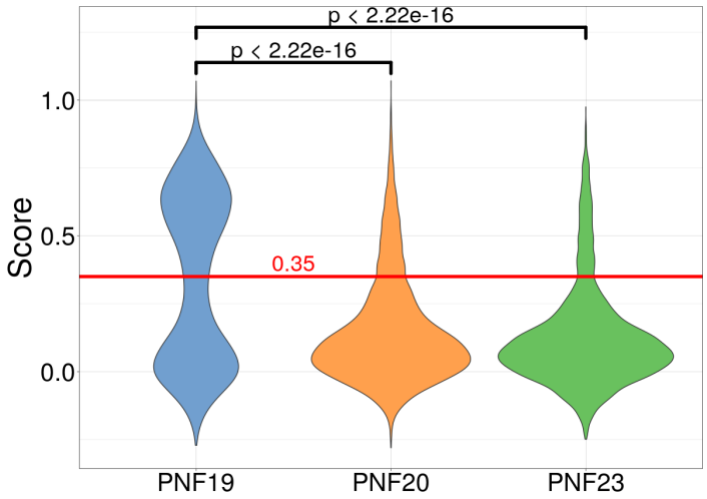


Figure 3 – Stress levels per sample. *The distribution of the scoring values for each sample show an accumulation of high scoring cells in PNF19 with respect to the other two samples.*

We think the stress levels are most likely not due to some intrinsic property of the tumor but to variations in the sample processing once extracted by the surgeons. Maybe the time elapsed between the tumor resection and the deposition of the specimen in the tube with media was different, or maybe there were differences in temperatures along the process, or any other environmental factor affecting the tumor.

In any case, we decided that the best course of action was to completely remove the stressed cells instead of trying to correct the stress signature by statistical means, given how the stress affected in unknown ways a large number of genes.

Therefore, we set a scoring threshold (Figure 3) and removed all cells from all samples above that threshold (Figure 4).

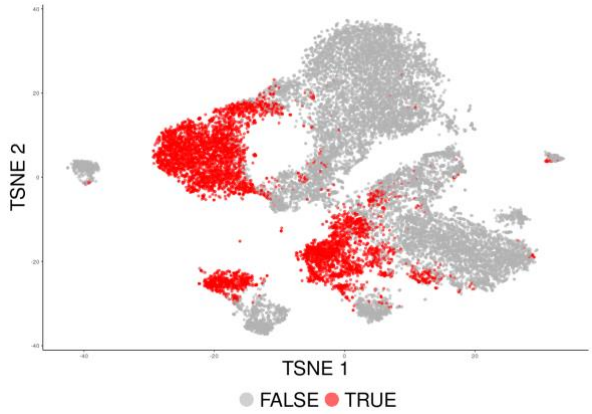


Figure 4 – Removed cells. *Cells with the stress scoring level above the threshold (red) were removed from the analysis.*

After removing the stressed cells we performed again the clustering and dimensionality reduction steps and obtained a much more informative image, which was used for all subsequent analysis. In this new image we can clearly see the two main components (Schwann cells and fibroblastic/mesenchymal) (Figure 5).

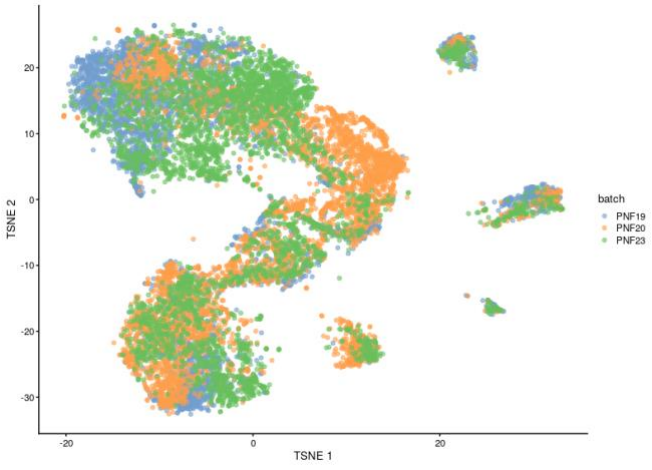


Figure 5 – Reprocessed non-stressed cells. *After the removal of the stressed cells, the TSNE plot changed producing a much more expected distribution.*

All in all, we think the identification of stressed cells is an important step in the processing and QC of scRNAseq data that has been largely missing in standard analysis so far. Our pipeline contains a step testing for this and other kinds of cellular stress and has the possibility of removing the stressed cells. We are working on a new software package to facilitate this kind of analysis and our intention is to publish an additional bioinformatics-oriented journal article describing this finding and the software package.

We performed a preliminary analysis of the 4 samples sequenced (3 ANFs and 1 PNF) together with the 3 PNF samples previously generated as preliminary data using the scRNAseq data analysis pipeline.

A few observations we deemed scientifically interesting follow:

Note: It is important to bear in mind that these are just mere observations and that a rigorous analysis and exploration is needed before any conclusion can be drawn.

A – There are more samples with a significant amount of stress.

When analyzing the cellular stress levels of all 7 samples we identified 2 samples with a significant amount of stressed cells, PNF19 (seen before in previous section) and 44ANFD. Both samples have a very similar stress profile, while the other 5 samples share a similar low stress profile (Figure 6).

NOTE: the actual values and the threshold levels are different than described in the previous section because this data is the scRNAseq part of the Multiome analysis.

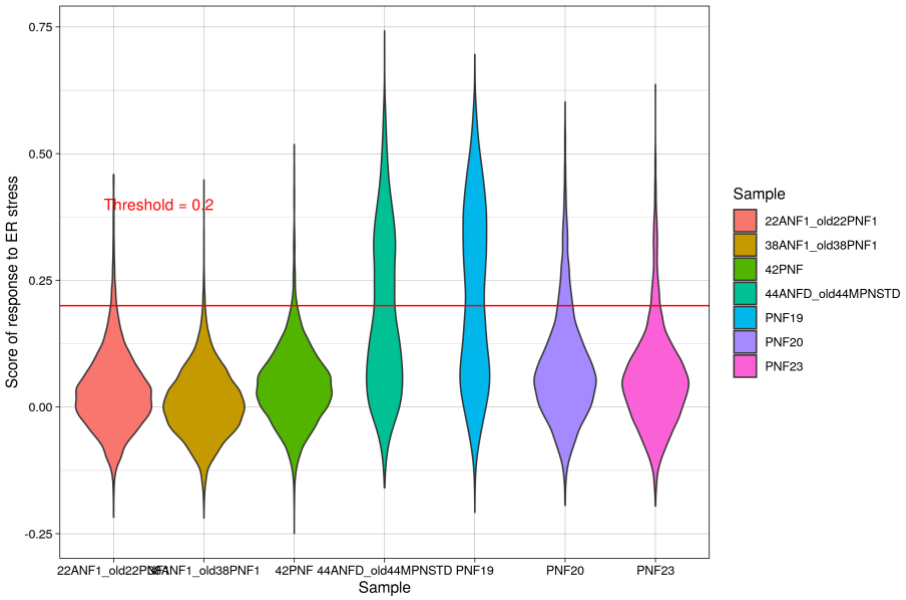


Figure 6 – Stress affects the positioning of cells in the dimensionality reduction plots. *We can see how some parts of the TSNE plot show cells from the PNF19 (blue) and create new additional cell clusters (red numbers).*

As in the analysis described in the previous section, the stressed cells clustered together and were placed together in the dimensionality reduction plots, showing the impact of the stress in the global transcriptomic profile (Figure 7).

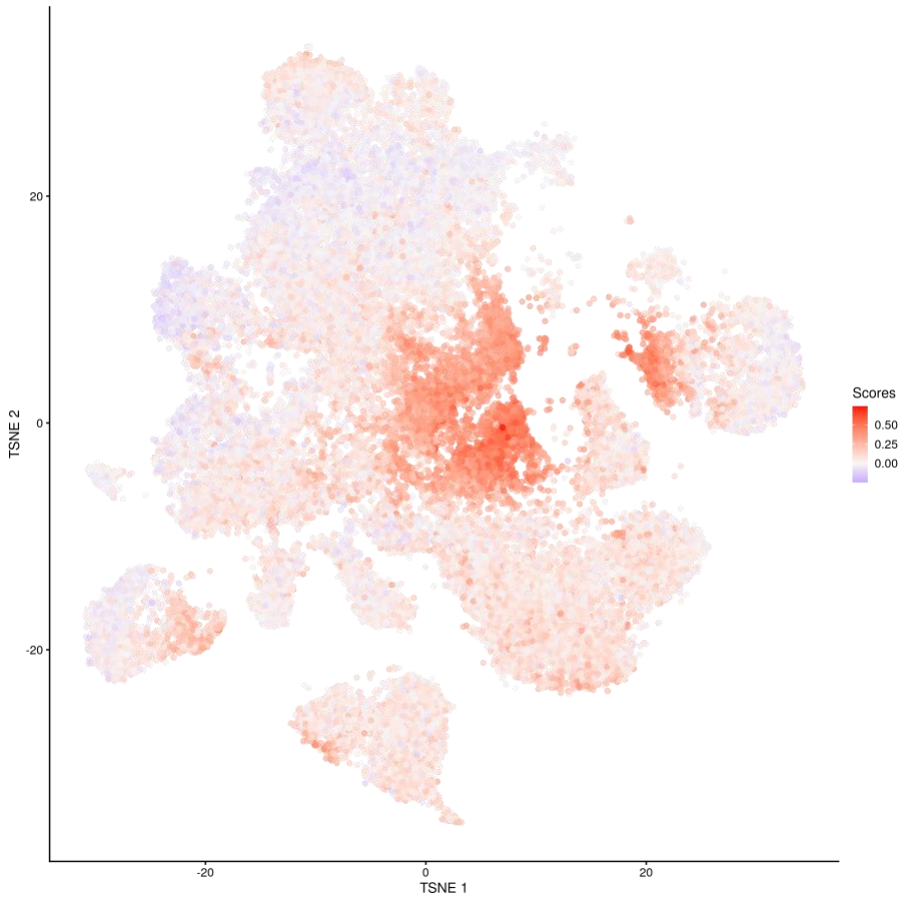


Figure 7 – Global impact of stress in the transcriptomic profile. *Cells with high stress score cluster together in this TSNE plot, almost regardless of their actual cell-type. This shows the impact of the high stress levels in the transcriptional profile.*

B – The observed cell components are the expected ones, but there are important differences between samples

The main cell components identified by marker gene expression (Figure 8) and single-cell level cell type assignment (Figure 9) showed the expected main cellular components and cell types: Schwann-cell component (the tumor cells), mesenchymal component (the main part of the microenvironment), as well as the endothelial cells and the immune infiltrate.

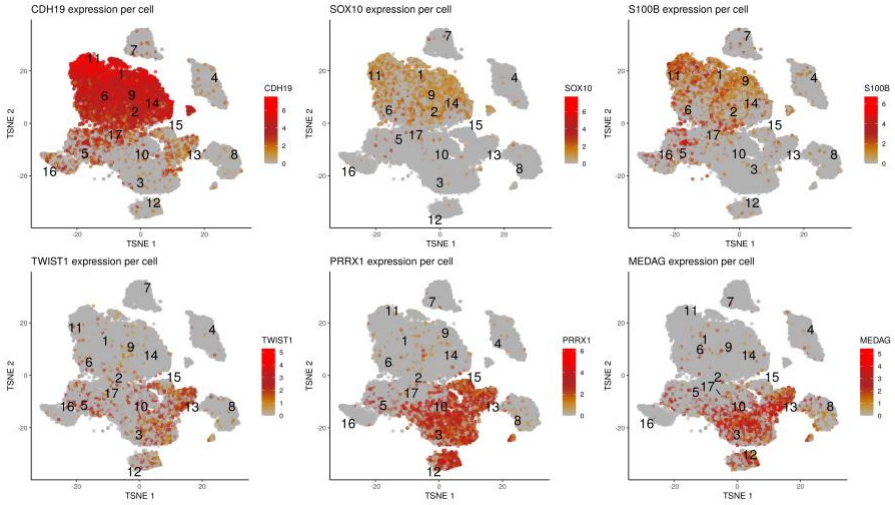


Figure 8 – Main cell component by marker genes. *The expression of marker genes for the Schwann cell component (top row) and mesenchymal component (bottom row) clearly identifies them. Markers for the other minor components no shown.*

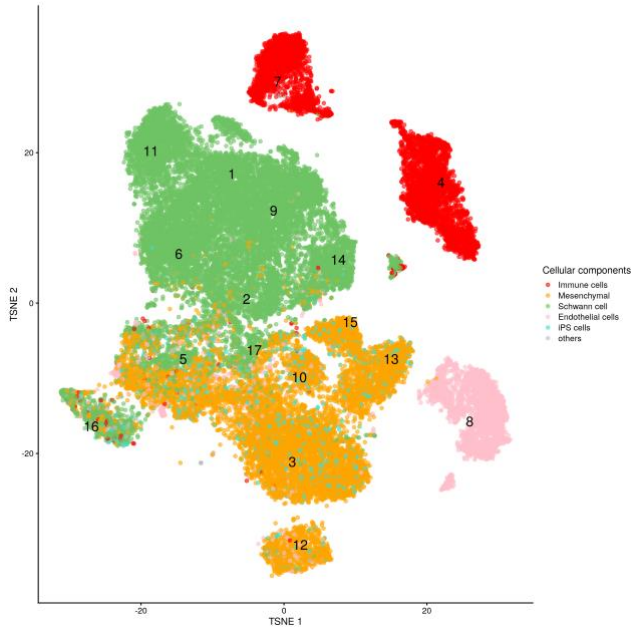


Figure 9 – Main cell component by cell type assignment. *The components are also clearly recovered by cell-type assignment based on a reference dataset (Human Primary Cell Atlas).*

However, although the whole dataset perfectly recovers the expected components in the expected proportions we can see quite striking differences in cellular composition between the different samples (Figure 10 and Figure 11). We can clearly see how ANFs (22ANF, 38ANF and 44ANF) have a lower proportion of mesenchymal cells when compared to PNFs (42PNF, PNF19, PNF20 and PNF23), with the most extreme

example in 22ANF, which is almost completely depleted from mesenchymal cells and endothelial cells and has a very small amount of immune cells. The other two ANFs show a reduced presence of mesenchymal cells but much higher immune infiltrate, which is missing (totally or partially) in the different PNFs. It is important to note that some of these differences may be related to the exact part of the tumor sampled, but the difference in the presence of mesenchymal cells between ANFs and PNFs seems consistent so far and is a biologically plausible result of the expansion of Schwann cells with no CDKN2A in ANFs.

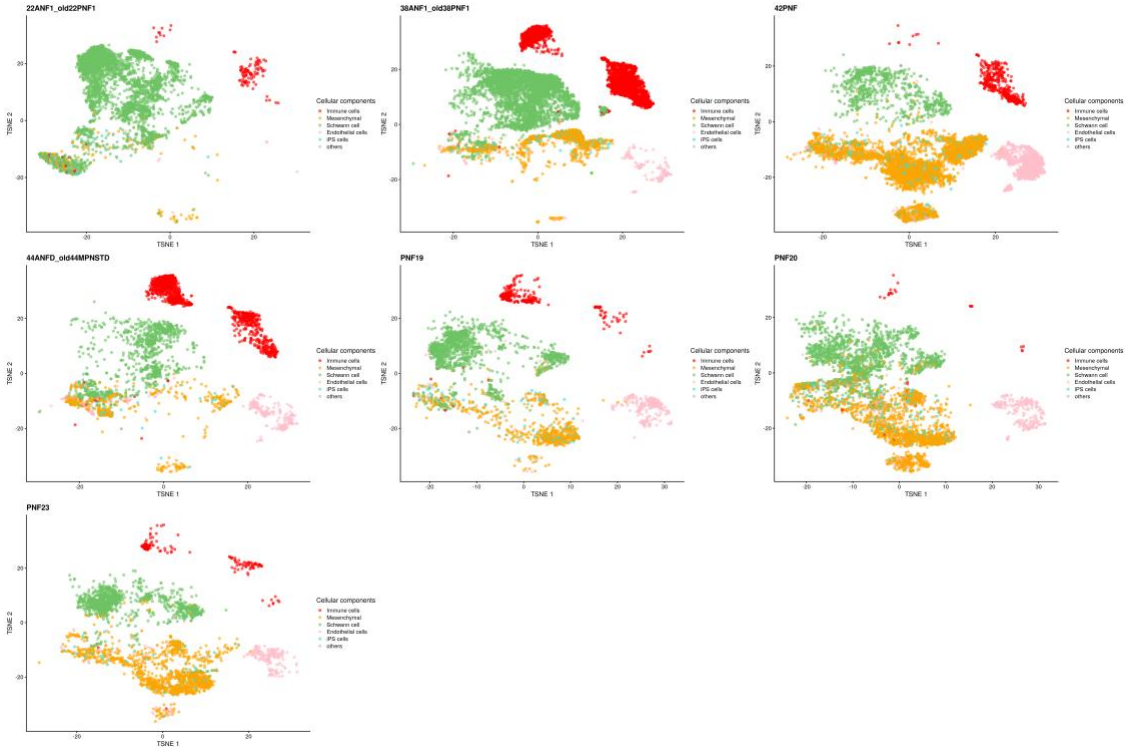


Figure 10 – Main cell components per sample. *Different samples present different cell-type composition. These plots show the TSNE representation of each sample with each cell colored according to its assigned cell-type.*

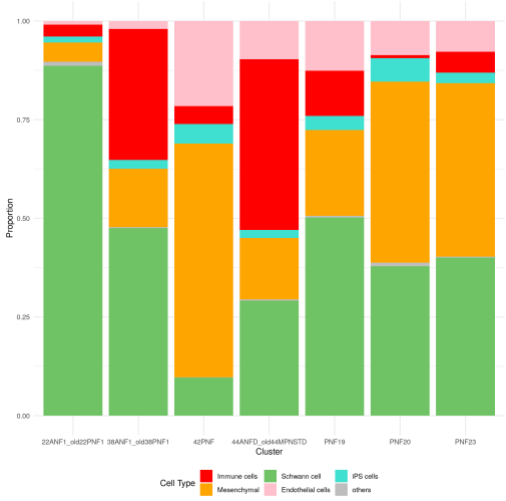


Figure 11– Main cell components per sample (barplots). *Barplots showing the proportion of each cell type in each sample.* ***Note:*** *While iPS cells can not exist in a primary tissue like the ones studied here, they were identified probably due to the incompleteness of the reference datasets regarding our tissues of interest, which we expect to partially improve with the execution of this project.*

C – Differences in gene expression in the different samples

In addition to differences in the cellular composition of the different samples, we can observe differences in the expression of some important genes.
One example is *CDKN2A*. Its inactivation defines at the molecular level the transition from PNF to ANF and so we expect its expression to be decreased in ANFs (Figure 12). While the differences are not clear cut, we can see how in 38ANF and 44ANF the expression of *CDKN2A* in Schwann cells is almost null as expected. In the case of 22ANF, which presents a very high expression level of *CDKN2A,* the inactivation of this gene is caused by a translocation event. We hypothesize that it might be expressed in a non-functional form and its absence activates a feedback loop that increases its expression to levels even higher than those seen in PNFs.

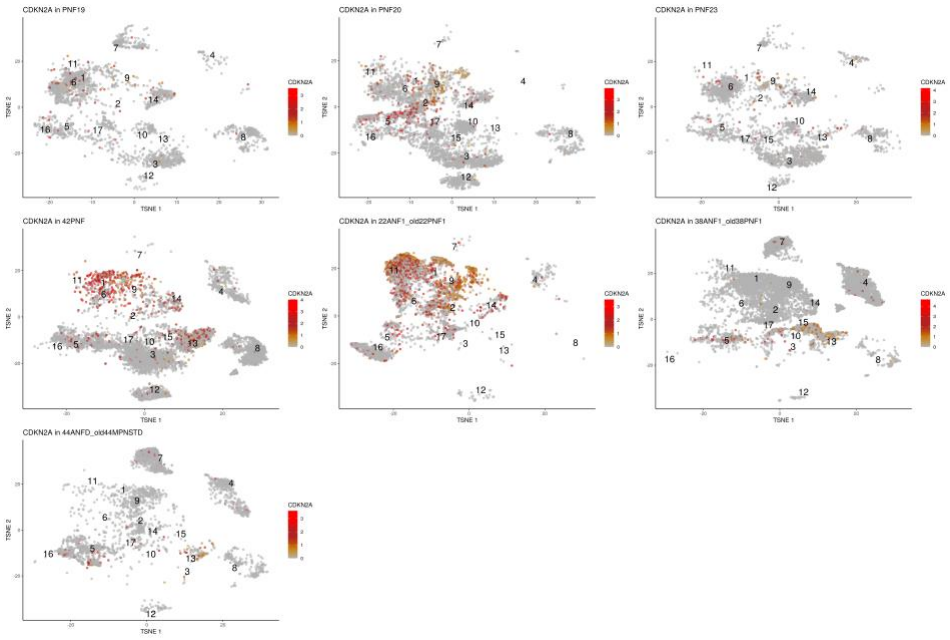


Figure 12 – CDKN2A expression per sample. *The expression level of* CDKN2A *in the different samples show slight differences between samples and tumor types.*

Similar differences can be seen in other genes. For example, while the number of cell expressing *CDH19* does not change between PNFs and ANFs, the actual expression level is somewhat higher in ANFs (Figure 13).

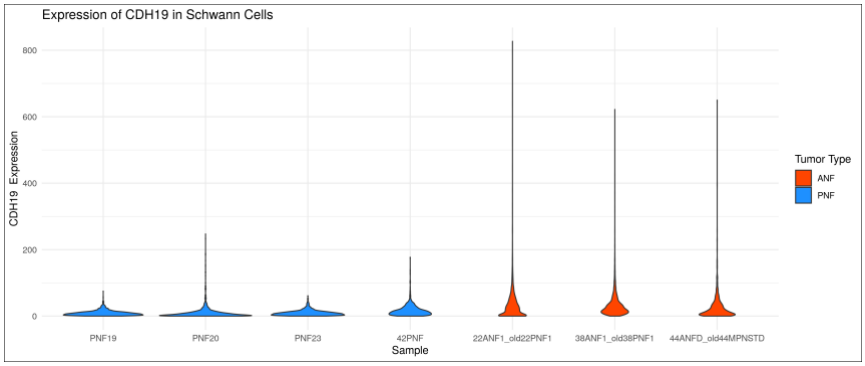


Figure 13*– CDH19 *expression in Schwann cells. *Violin plots representing the expression levels of* CDH19 *in all cells identified as Schwann cells in the different samples. A clear difference between PNFs (blue) and ANFs (red) can be seen.*

What opportunities for training and professional development has the project provided

Training: One-on-one work with a mentor. Pere Pericot was an undergraduate student in our lab for 6 months under the supervision of Dr. Bernat Gel. He worked in the area of single cell data analysis of NF1 related tumors. He presented a poster at the Children's Tumor Foundation Conference "NF conference 2022" (see **Appendices**).

How were the results disseminated to communities of interest?

Nothing to Report

4. IMPACT

Nothing to Report

5. CHANGES/PROBLEMS

There have been two changes in the execution of the present project:

1) Small delay in obtaining the RNA-seq data from all proposed samples.

During these first 12 month of the project, we established obtaining the RNA-seq data by 10x Genomics and SMART-seq analyses of the 16 samples proposed. However, given the value of the samples and the different conditions needed to set up for each of the different groups of samples (nerves, neurofibromas, MPNSTs), we prioritized first to set up the whole single cell analysis workflow (sample dissociation; single cell preparation; library constructions; sequencing; bioinformatic analysis) of few samples of the different tumor/tissue types, instead of performing sample preparation and sequencing for all 16 samples, before checking that the whole workflow was performing as expected. Since we know now that the workflow has been set up for all different types of tissue/tumors, now we can proceed with the analysis of all samples. This change in the order of stablished priorities (that is, finish the whole workflow up to sequencing analysis for a few samples of each type instead of performing the sequencing of the 16 samples before performing a bioinformatic quality control of the data) delayed just a bit the completion of the RNA-seq analysis proposed. However, this is just a small delay, since at the end of the summer, all 16 samples will be already completed and we will catch up time, since the bioinformatic pipelines are already ready to be executed for all of them.

2) Adding non-planned single cell data to the dataset

One of the aims of the present project is to generate a shared resource consisting in a comprehensive cell diversity map of the cellular composition, spatial distribution and genomic content of NF1-associated peripheral nerve sheath tumors. In order to make this single cell data set even more complete, we decided to include single cell epigenetic information, initially not planned to be part of this data set, that has been generated in the context of another previous project. In this regard, with no additional cost for the present project, there will be another layer of information (ATAC-seq data) from the same samples, in the final repository.

6. PRODUCTS

Publications, conferences, papers and presentations:

Poster: "Single-cell analysis of Plexiform Neurofibromas: heterogeneity and tumor cell-of-origin"

Pere Pericot, Miriam Magallon-Lorenz, Helena Mazuelas, Itziar Uriarte, Holger Heyn, Meritxell Carrió, Eduard Serra & Bernat Gel

Children's Tumor Foundation **2022 NF Conference**, Philadelphia June 18-21 2022

Poster: "A genomic portrait of MPNST cells: governed by genomic structural and copy-number alterations and a mesenchymal identity"

Miriam Magallon-Lorenz, Ernest Terribes, Helena Mazuelas, Sara Ortega, Edgar Creus, Ignacio Blanco, Héctor Salvador, Nancy Ratner, Margaret Wallace, Juani Fernández-Rodríguez, Conxi Lázaro, Meritxell Carrió, Eduard Serra & Bernat Gel

Children's Tumor Foundation **2022 NF Conference**, Philadelphia June 18-21 2022

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked in the project?

Name	Helena Mazuelas
Project Role	Postdoctoral Fellow
Researcher Identifier	0000-0001-7212-2144
Nearest person month worked	12
Contribution to the project	Dr. Mazuelas worked in the area of setting up conditions and preparing samples for single cell analysis

Name	Meritxell Carrió
Project Role	Wet Lab coordinator
Researcher Identifier	0000-0002-1258-6593
Nearest person month worked	2.4
Contribution to the project	Dr. Carrió coordinated and supervised the area of setting up conditions and preparing samples for single cell analysis

Name	Bernat Gel
Project Role	Dry Lab coordinator
Researcher Identifier	0000-0001-8878-349X
Nearest person month worked	2.4
Contribution to the project	Dr. Gel set up bioinformatic tools for single cell data analysis

Name	Eduard Serra
Project Role	PI. Project coordinator
Researcher Identifier	0000-0003-2895-9857
Nearest person month worked	2.4
Contribution to the project	Dr. Serra supervised and directed the project

Have there been a change in the active other support of the PI?

Nothing to report

What other organizations where involved as partners?

Organization Name: Centro Nacional de Análisis Genómico (CNAG-CRG);

Location of organization: Barcelona, Spain

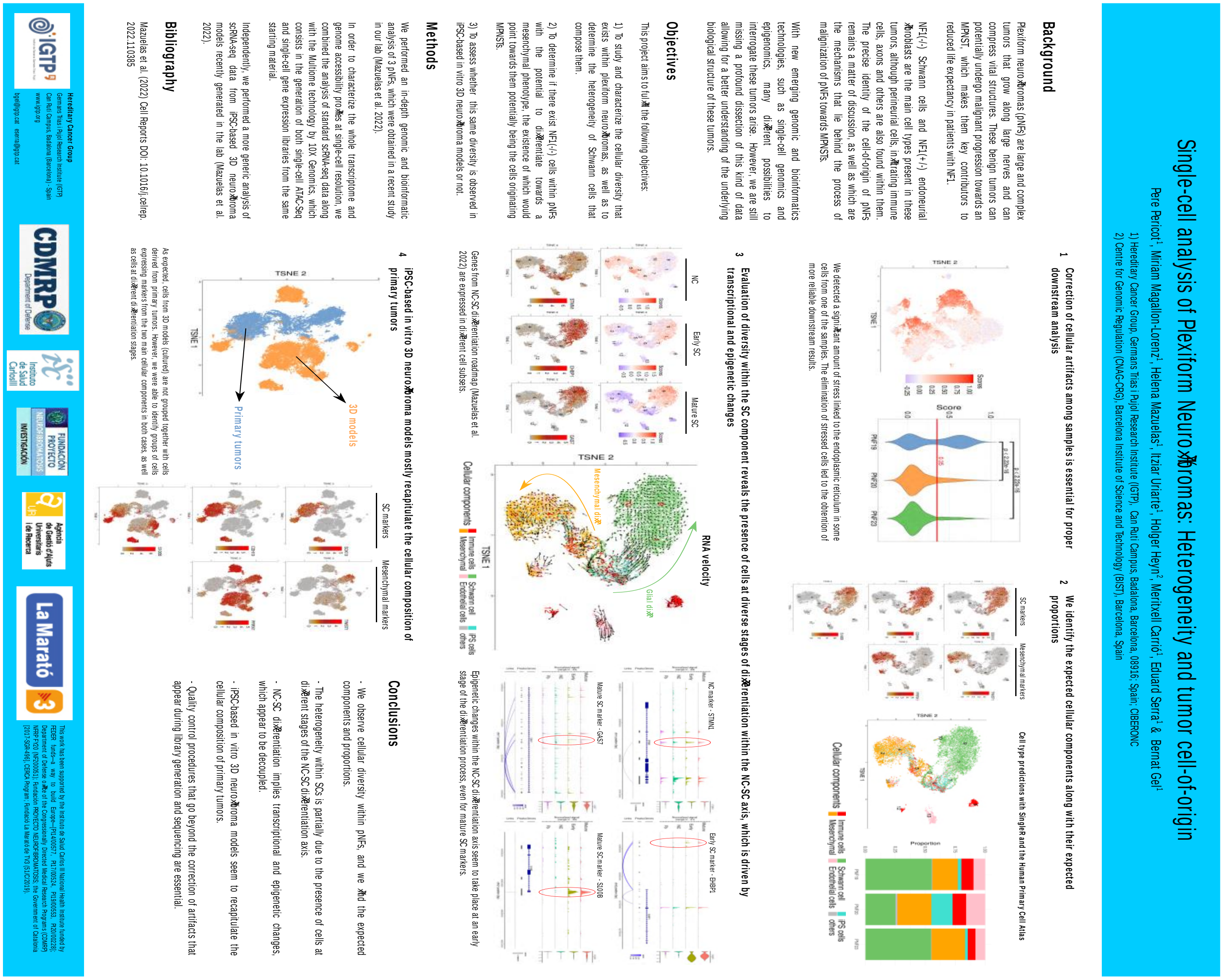
Partner's contribution to the project: Single Cell Genomic group at the CNAG-CRG; **Facility**: Singe Cell genomic facility; **Collaboration**: Dr. Holger Heyn Lab

8. SPECIAL REPORTING REQUIREMENTS

Nothing to report

9. APPENDICES

Poster 1 presented at CTF NF Congress 2022



A genomic portrait of MPNST cells:

governed by genomic structural and copy-number alterations and a mesenchymal identity

Miriam Magallon-Lorenz1, Ernest Terribas1, Itziar Uriarte1, Helena Mazuelas1, Sara Ortega2, Edgar Creus2,3, Ignacio Blanco4, Hector Salvador5, Nancy Ratner6, Margaret Wallace7, Juani Fernández-Rodríguez2,3, Conxi Lázaro2,3, Meritxell Carrió1, Eduard Serra1,3 & Bernat Gel1,8

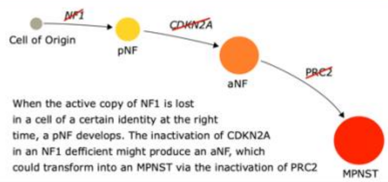
1 - Hereditary Cancer Group, Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain; 2 - Hereditary Cancer Program, Catalan Institute of Oncology (ICO-IDIBELL), L'Hospitalet de Llobregat, Spain; 3 - Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain; 4 - Clinical Genetics Service, Northern Metropolitan Clinical Laboratory, Germans Trias i Pujol University Hospital (HGTP), Badalona, Spain; 5 - Pediatric Oncology Department, Sant Joan de Déu Barcelona Children's Hospital, Barcelona, Spain; 6 - Division of Experimental Hematology and Cancer Biology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA; 7 - Department of Molecular Genetics & Microbiology, University of Florida College of Medicine, Gainesville, FL, USA; 8 - Departament de Fonaments Clínics, Universitat de Barcelona.

Introduction

Malignant peripheral nerve sheath tumors (MPNST) are aggressive soft-tissue sarcomas with a bad prognosis, mainly due to a high metastatic potential and lack of therapeutic options. MPNST diagnosis is challenging, especially in sporadic cases, since specific histological criteria have not been completely established.

We performed a deep genomic characterization of about 40 MPNSTs and 8 cell lines using whole-genome sequencing (WGS), whole-exome sequencing (WES), RNA-seq, SNP-array and methylation array. A different subset of techniques was applied to each MPNST, mainly depending on tissue availability. Single-cell RNA-seq (scRNA) and single-cell ATAC-seq (scATAC) were also performed in a small subset of MPNST and a set of plexiform neurofibromas (PNF).

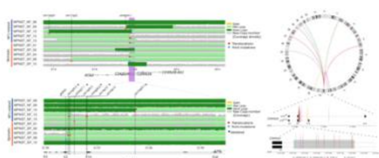
Working Model



When the active copy of NF1 is lost in a cell of a certain identity at the right time, a pNF develops. The inactivation of CDKN2A in an NF1 deficient might produce an aNF, which could transform into an MPNST via the inactivation of PRC2

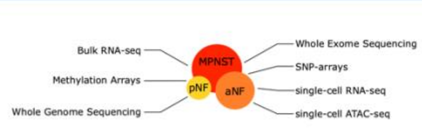
Inactivation of Tumor Suppressor Genes

For the generation of an MPNST 3 tumor supressor genes (TSG) must be inactivated: *NF1*, *CDKN2A* (both encoded proteins p16 and p14) and a component of the polycomb repressor complex 2 (PRC2) (either *SUZ12* or *EED*). Inactivations must be in that order and in many cases are caused by structural variants (SV). That is part of the reason these inactivations are not always found in all MPNSTs.

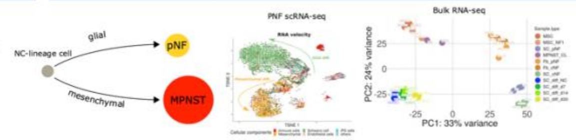


Adapted from Magallón-Lorenz, 2021

Integrative Analysis



Mesenchymal Identity

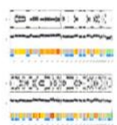


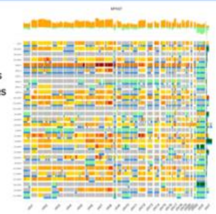
see poster by Pere Pericot *et. al.*

Genomic Profile

Copy Number changes

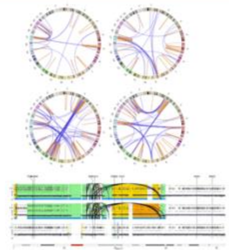
MPNSTs have a highly rearranged hyperploid genome, with recurrent gains in 7, 8, 1q and 17q and losses in 9p, but no recurrent focal gains or losses.





Structural Variants

MPNST genomes usually present a large number of translocations and inversions, but there lot pf variability between tumors. They present no recurrent SVs and they have no fusion genes, but SVs frequently inactivate the key tumor suppressor genes. Some MPNSTs show regions of the genome with accumulation of SVs, suggesting thy might the result of chromotripsis or similar processes.



Small variants

MPNST have a low mutation burden. They have a low number of small mutations (SNVs and small insertions and deletions). Mutational signature analysis point to clock-like signatures and identifies no additional mutattional process.

Most SNVs occur after the genome rearrangements, meaning they are not driving tumorogenesis. MPNSTs usually don't have activating mutations in oncogenes

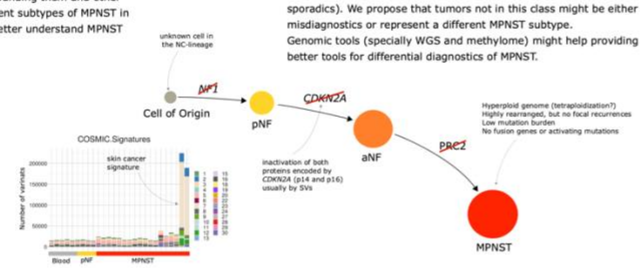
Potential Misdiagnoses

There exist a high potential of misdiagnosis of MPNSTs by counfounding them and other similar entities. The presence of misdiagnosed MPNSTs or different subtypes of MPNST in the literature impacts our capacity to establish generalities, to better understand MPNST biology and has an impact in precision medicine efforts.

In our hands, "classic MPNSTs" are defined by:

- the inactivation of *NF1*, *CDKN2A* and PRC2
- a hyperploid and highly rearranged genome, with recurrent alteration of 7, 8, 9p, 1q and 17q
- no expression of glial differeniation markers (e.g. *S100B*)
- low mutation burden and no overrepresentation of specific mutational signatures
- lack of activating mutations or fusion genes
- classified as MPNST by the methylome-based sarcoma classifier

For evidences of misdiagnosis of often used cell lines see poster by Miriam Magallón-Lorenz *et. al.*



Conclusions

We performed a portrait of the "classic MPNST" (either NF1-related or sporadics). We propose that tumors not in this class might be either misdiagnostics or represent a different MPNST subtype.

Genomic tools (specially WGS and methylome) might help providing better tools for differential diagnostics of MPNST.

Hereditary Cancer Lab
Germans Trias i Pujol Research Institute
Ctra de Can Ruti, camí de les escoles, s/n
Badalona, Spain

This work has been supported by the Instituto de Salud Carlos III National Health Institute funded by FEDER funds—a way to build Europe—[PI14/00577, PI17/00524, PI18/00053, PI20/00228]; Department of Defense office of the Congressionally Directed Medical Research Programs (CDMRP) NFRP FY20 (NF200051); Fundación PROYECTO NEUROFIBROMATOSIS (projects to CL and BG-ES); the Government of Catalonia [2017-SGR-496], CERCA Program; Fundació La Marató de TV3 (51/C/2019). MM-L is supported by Fundación PROYECTO NEUROFIBROMATOSIS.

bgel@igtp.cat
@bernatgel