Integrated genomic analysis of NF1-associated peripheral nerve sheath tumors: an updated biorepository dataset

Stavriani Makri[1], Jineta Banerjee[2], Yang Lyu[3], Alexandra Scott[2], Lindy Zhang[1], Ana Calizo[1], Kai Pollard[4], Kuangying Yang[3], Lisa Fridman[1], John Gross[5], Calixto-Hope Lucas[5], Jiawan Wang[1], Adam Levin[6], Allan Belzberg[7], Carlos Romo[8], Robert Allaway[2], Angela Hirbe[3] , Jaishri O. Blakeley[8], Christine A. Pratilas[1]

[1] Division of Pediatric Oncology, Johns Hopkins University School of Medicine; [2]Sage Bionetworks; [3]Division of Oncology, Washington University St. Louis; [4]Personal Genome Diagnostics; [5]Department of Pathology, Johns Hopkins University School of Medicine; [6]Department of Orthopedic Surgery, Johns Hopkins University School of Medicine; [7]Department of Neurosurgery, Johns Hopkins University School of Medicine; [8]Department of Neurology, Johns Hopkins University School of Medicine

JOHNS HOPKINS MEDICINE
THE SIDNEY KIMMEL COMPREHENSIVE CANCER CENTER

We Are HOPKINS KIMMEL

Introduction

Neurofibromatosis type 1 (NF1) is an inherited neurogenetic condition associated with increased risk of developing a variety of tumors, including cutaneous neurofibromas (cNF), plexiform neurofibromas (pNF), atypical neurofibromatous neoplasms of uncertain biologic potential (ANNUBP), and malignant peripheral nerve sheath tumors (MPNST). The Johns Hopkins NF1 Biospecimen Repository was established in 2016 to enhance research for NF1 tumors via access to primary human tissues and preclinical models. It has since become a vital resource for NF1 research worldwide and continues to expand the scope of available biospecimens to address growing requests from the NF research community. In addition to the growing collection of banked tumor tissue, the biorepository has broadened its offerings to include single-cell suspensions from tumors, tissue microarrays (TMA), genetically diverse cell lines, and patient-derived xenografts (PDX).

Methods

- Patients with clinically or genetically confirmed NF1 undergoing surgical resection or biopsy of NF1-associated tumors are invited to participate in this IRB-approved study for specimen collection.
  - Tissue fragments are frozen, paraffin-embedded, and digested into single-cell suspensions.
  - Cell lines and PDX are attempted with malignant tumors.
- For new cell lines, IncuCyte live-cell imaging is used to assess growth and response to MEK and SHP2 inhibitors.
- Banked specimens are genomically characterized via whole exome sequencing (WES), whole genome sequencing (WGS), and RNA sequencing (RNAseq), with data accessible through the NF Data Portal.
- Clinical annotations and outcomes data are made available to investigators upon scientific review and IRB approval.

Tumor sample preparation for DNA and RNA sequencing

- Genomic DNA from blood and tumor was extracted. RNA was isolated from tumor by using TRIZOL.

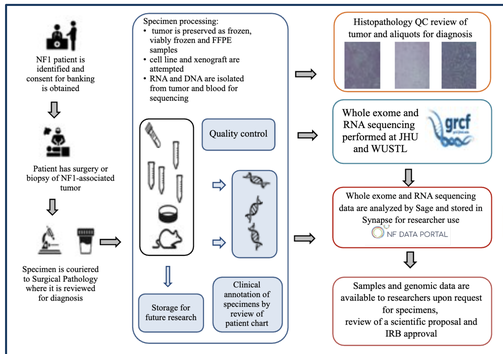
WES data processing for somatic variant calling and analysis

- Raw fastq data files were quality checked using FastQC v0.11.9 and a report was generated using MultiQC v1.8. Fastq files were aligned to GRCh38 using BWA 0.7.17-r1188.
- Duplicates were marked using GATK MarkDuplicates, and bases recalibrated using GATK BaseRecalibrator and GATK ApplyBQSR (GATK v4.3.0.0). Somatic single nucleotide variants (SNVs) were then called using Strelka2 software (Strelka v2.9.10) and Mutect2 (GATK v4.4.0.0).
- To avoid detection of false positive variant calls in the samples, we identified and reported the variant calls that had consensus between Strelka2 and Mutect2 callers, using a previously benchmarked method called SomaticSeq.

RNAseq data processing and analysis

- Raw fastq files were processed using nf-core/rnaseq (v3.11.2) and quantified using salmon: v1.10.1. ComBat from sva R package (v3.42.0) was used for batch correction. DESeq2 (v1.34.0) was applied to call the differentially expressed genes (|fold change| > 1.2, adjust $P < 0.05$), and ggplot2 R package (v3.3.6) was applied to draw volcano plots.
- Principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) were applied for the visualization of samples.
- Gene set enrichment analysis (GSEA) was conducted by using R-package fgsea (v1.20.0). ggplot2 (v3.3.6) was used for the gene volcano plot and pathway enrichment plot. Pheatmap (v1.0.12) was applied to plot gene expression heatmap.

Results



Overview of the Johns Hopkins NF1 Biospecimen Repository

- 440 samples, 225 unique patients (May 2025)
- RNAseq 104 primary tumor samples
- Whole exome 114 primary tumor samples
- Whole genome 47 primary tumor samples
- 5 cell lines and 12 MPNST PDX
- More than 80 shares to investigators in US and Europe
- New tissue microarray (2024)

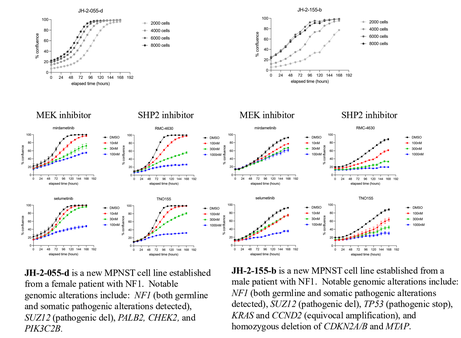
Figure 1. NF1 patients are identified at the Johns Hopkins Comprehensive Neurofibromatosis Center (JHCNC). Samples are collected when patients undergo biopsy or surgical excision. In addition to routine frozen samples, FFPE, and RNA and DNA isolation, MPNST specimens are processed by mouse implantation, cell line generation, and dissociation for single-cell suspension (modified from *Pollard, et. al, Scientific Data, 2020*).

Johns Hopkins NF1 biospecimen repository

New MPNST cell lines and PDX

Specimen ID	Cell line	PDX	Male/ female	SUZ12/EED status (JH NGS)	Other clinical genomic sequencing (JH NGS)
JH-2-002	Y	Y	M	SUZ12 mutation	NF1, CDKN2A/B, MTAP del
JH-2-009-d		Y	M	SUZ12 mutation	ATM, BLM, RECQL4, DICER1, EGFR
JH-2-023		Y	M	NA	NTRK1, ATR, MAP2K2, MAP3K1
JH-2-031		Y	M	SUZ12 mutation	PIK3CA Q546K
JH-2-055-b		Y	F	SUZ12 mutation	NF1, AGAP9, AGAP3, ARID1A, ARID1B
JH-2-055-d	Y		F	SUZ12 mutation	PALB2, CHEK2, PIK3C2B
JH-2-060-d		Y	M	SUZ12 mutation EED WT	NF1, BRCA1 amp, MDM2 amp, CDKN2A/B homo del (MTAP intact)
JH-2-079-c	Y	Y	F	EED mutation	NF1, EED, CDKN2A, PIK3C2B, PDGFRB, RECQL4
JH-2-079-d		Y	F	EED mutation	NF1, EED, CDKN2A, PIK3C2B
JH-2-103	Y	Y	M	SUZ12 mutation	NF1, MAP2K2, EGFR (A289V), TP53
JH-2-120		Y	M	SUZ12/EED WT	NF1, TP53
JH-2-155-b	Y		M	SUZ12 mutation EED WT	NF1, SUZ12, CCND2 equiv amp, KRAS equiv amp, TP53, CDKN2A/B homo del, MTAP del
JH-2-169		Y	M	SUZ12/EED WT	NF1, MTAP del, CDKN2A/B del, AKT2 equiv amp
JH-2-242-a		Y	F	NA	NA

Figure 4. New cell lines and PDX have been developed from primary human MPNST tissue specimens. Representative cell growth curves are shown for untreated cells, plated at a range of cell densities, and in response to MEKi (mirdametinib or selumetinib) or SHP2i (RMC-4630 or TNO155).



JH-2-055-d is a new MPNST cell line established from a female patient with NF1. Notable genomic alterations include: *NF1* (both germline and somatic pathogenic alterations detected), *SUZ12* (pathogenic del), *PALB2, CHEK2,* and *PIK3C2B*.

JH-2-155-b is a new MPNST cell line established from a male patient with NF1. Notable genomic alterations include: *NF1* (both germline and somatic pathogenic alterations detected), *SUZ12* (pathogenic del), *TP53* (pathogenic stop), *KRAS* and *CCND2* (equivocal amplification), and homozygous deletion of *CDKN2A/B* and *MTAP*.

New tissue microarray (TMA)

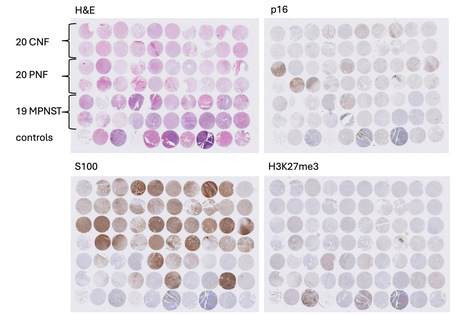


Figure 5. We constructed a 7x10 array of 1mm cores from formalin-fixed paraffin-embedded tissue blocks consisting of 20 cutaneous neurofibroma (CNF), 20 plexiform neurofibroma (PNF), 19 malignant peripheral nerve sheath tumors (MPNST), and 11 on-slide non-neoplastic control tissues. Immunohistochemical stains were also performed on the TMA, including for SOX10, S100, H3K27me3, p16, p53, and Ki-67. H&E and IHC whole slide digital images have been deposited into an online repository (Concentriq by Proscia) for sharing, upon request.

Relatedness between WES and RNAseq samples

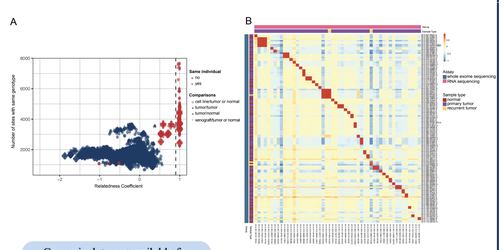


Figure 2. (A) A scatter plot of the relatedness coefficient versus the number of sites with the same genotype between two samples, as calculated by Somalier. Comparisons between two samples annotated as being from the same individual are shown in red and those from different individuals are colored blue. The shape of the point indicates the tissue of origin of the samples compared. The vertical line indicates the cutoff value (≥ 0.9) used to determine whether two samples are from the same individual based on the relatedness coefficient. **(B)** A heatmap of the relatedness coefficient of pairwise comparisons between WES and RNAseq data from primary human samples as shown in (A), excluding cell lines or xenografts derived from these samples. RNAseq samples are shown on the x-axis and WES samples are on the y-axis. The relatedness coefficient is displayed according to the color scale bar shown in the figure.

Genomic data are available for access on the NF Data Portal. https://www.synapse.org/#!Synapse:syn4939902/wiki/235907

Summary of somatic variants detected in NF1-associated tumors

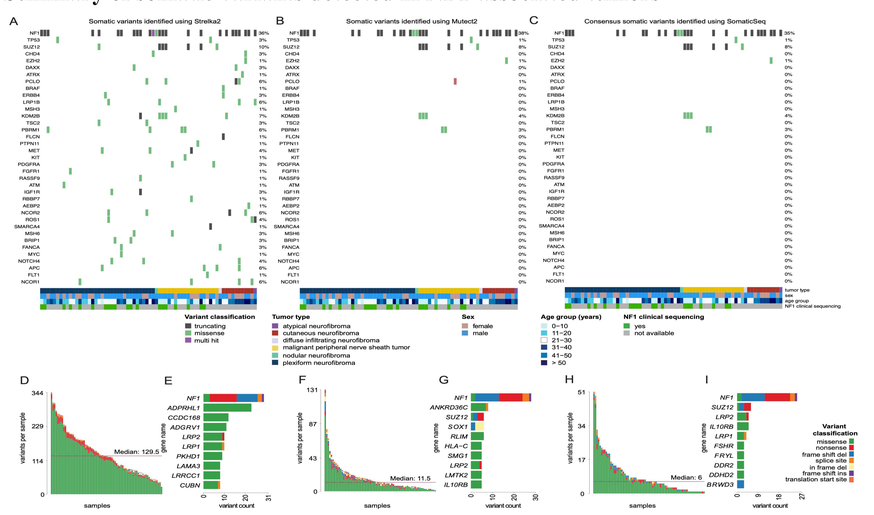


Figure 3. **(A)** Oncoplot of variants in selected genes of interest in the cohort of biorepository patients detected using Strelka2 of 10 CNF, 1 nodular NF, 1 diffuse NF, 36 PNF, 1 ANNUBP, and 19 MPNST. **(B)** Oncoplot of variants in selected genes of interest in the cohort of biorepository patients detected using Mutect2. **(C)** Oncoplot of the consensus of variants in selected genes of interest in the cohort of biorepository patients detected using SomaticSeq. **(D, F, H)** Bar plots showing the number of variants per sample, by SNV class. **(E, G, I)** Bar plots of the top 10 genes with variants of moderate or high impact in the cohort.

Conclusions

- The successful advancement of therapeutic developments for NF1-associated PNST necessitates ongoing efforts in the systematic acquisition and analysis of human tumor specimens and their corresponding model systems.
- The goal of this project is to bank blood and tumor samples with accompanying clinical data from patients with NF1-associated neoplasms. These tissues, and accompanying genomic data, represent a publicly available resource shared with researchers to promote ongoing collaboration for therapeutic discovery.
- A total of 187 samples (plus 71 corresponding normal samples) from banked tissues have been sequenced, including whole exome on a total of 114 tumors (33 MPNST, 4 ANNUBP, 57 PNF, 9 diffuse NF, and 11 CNF) and RNA sequencing on a total of 73 tumors (18 MPNST, 1 ANNUBP, 38 PNF, 9 diffuse NF, 6 CNF, and 1 nodular NF); data are available in the NF Data Portal.
- Our validation of the data confirms findings that are concordant with available literature regarding the genomics of NF1-associated tumors and validates the biospecimen repository genomic data as a useful resource for further downstream analysis by NF1 researchers. These genomic data are now publicly available to NF1 researchers to support an array of important lines of scientific inquiry.
- In response to evolving scientific needs, the biorepository has expanded to include single-cell suspensions, TMA blocks, and newly developed MPNST cell lines. These efforts further enhance its value for critical therapeutic advances for patients with NF1.

If you have received samples or data, please complete this survey:

Acknowledgements

Neurofibromatosis Therapeutic Acceleration Program
SageBionetworks
CDMRP CONGRESSIONALLY DIRECTED MEDICAL RESEARCH PROGRAMS
NF DATA PORTAL
Washington University in St. Louis

BioRxiv preprint: **doi:** https://doi.org/10.1101/2024.01.23.576977