Childhood Cancer Data Initiative (CCDI) Webinar Series

MCI Update: Pediatric Soft Tissue Sarcomas Jack Shern, M.D.



December 10, 2024

Agenda

- 1. Molecular Characterization Initiative (MCI): Creating a National Strategy
- 2. MCI Data Flow: Enabling Access to Data for Research and Analysis
- 3. MCI Soft Tissue Sarcoma Data: Clinical Impact
- **4**. Q&A

Today's Speaker

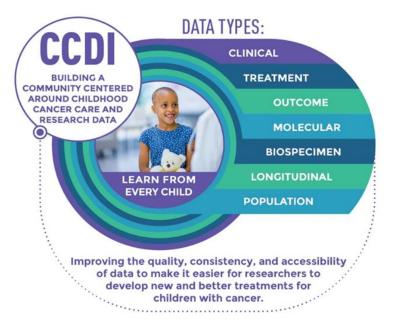


Jack Shern, M.D.

- Lasker Clinical Research Scholar
- Head of the Tumor Evolution and Genomics Section, National Cancer Institute

CCDI Goals

- Gather data from every child, adolescent, and young adult diagnosed with a childhood cancer
- Create a national strategy of molecular characterization to inform diagnosis and treatment
- Develop a platform and tools to bring together clinical and research data that will inform new insights in biology/etiology to improve preventive measures, treatment, quality of life, and survivorship for childhood cancers
- Engage the entire childhood cancer care and research community



*Flores-Toro JA et al., J Clin Oncol, 2023 (*PMID:37267580); *Jagu S et al., Pediatr Blood Cancer, 2024* (PMID: 37889049)

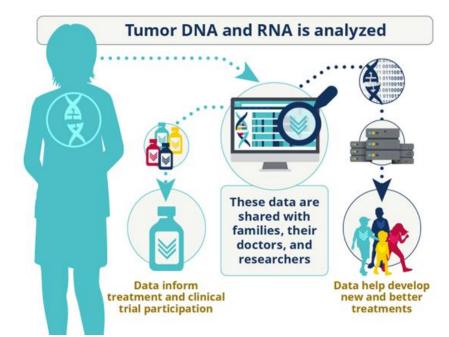
Molecular Characterization Initiative

Creating a National Strategy



CCDI Molecular Characterization Initiative (MCI)

- Launched in partnership with Children's Oncology Group's (COG) Project:EveryChild
- State-of-the-art molecular characterization at diagnosis (WES, fusions, methylation) in a CLIA-testing environment at no cost to participants
- Results returned to participants and treating physicians within 14-21 days
- Enrolled more than 4,500 participants from all 50 states, Canada, Australia, and New Zealand
- Learn more: <u>ccdi.cancer.gov/MCI</u>



MCI Enrollment Metrics

Enrollments by Diagnosis

Туре	Introduced	Number
CNS tumors	March 2022	3503
Soft tissue sarcomas	May 2022	1061
Rare tumors	September 2022	497
High-risk Neuroblastoma	February 2024	265

Enrollment Counts by State



Enrollment metrics as of November 12, 2024

Expanding MCI

- Plan to add Ewing sarcoma and relapsed or refractory tumors
- Expand to include AYAs outside of COG for the Coordinated Pediatric, Adolescent, and Young Adult Rare Cancer Initiative
- Prioritize diseases for research characterization and determine which assays (WGS, RNA Seq, single cell, epigenetics, proteomics/metabolomics) are appropriate to deepen our understanding of cancer biology
 - Partnering with COG disease-specific scientific committees and other subject matter experts

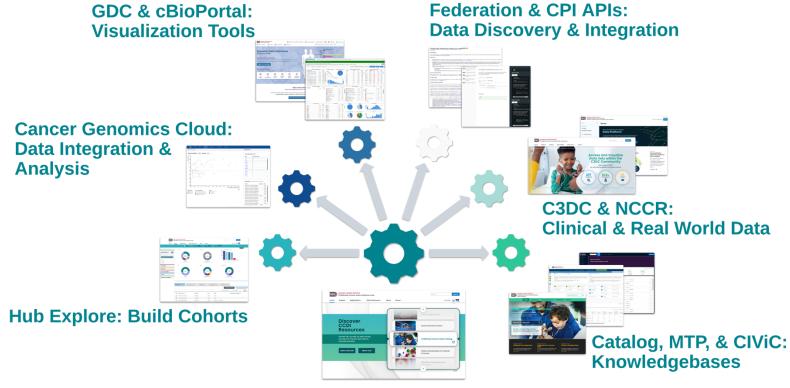


MCI Data Flow

Enabling Access to Data for Research and Analysis

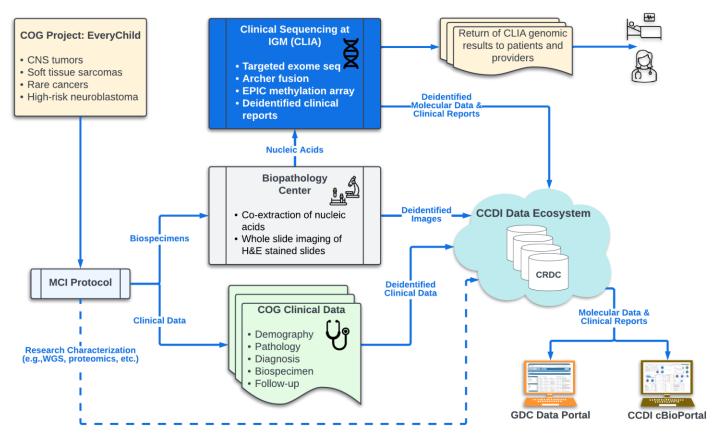


CCDI Data Ecosystem: A Connected Network of Resources



Hub: Entry Point

MCI Data Flow



NIH NATIONAL CANCER INSTITUTE

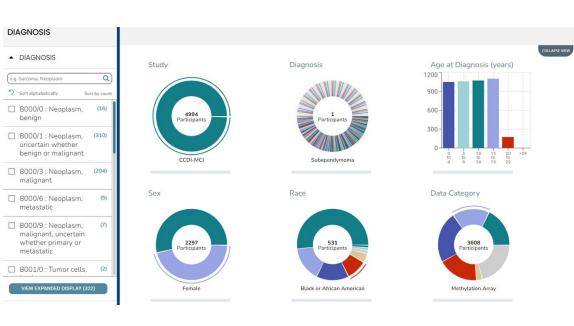
MCI: Assay Types & Data File Formats

Sample	Assay	Analytics	Result Type	File Description	File Format
	Enhanced WES	250x Churchill	Germline + somatic SNVs, INDELs, CNV, and LOH	Germline and Somatic Read Alignments	.FASTQ, .cram, .crai;
Tumor + normal DNA	Note: > 10-fold less input	alignment to GRCh38; IGMseq Pipeline	Note: Clinical report will	Germline and Somatic VCFs	.vcf.gz
	exome	Analysis in AWS	include tumor mutational burden (TMB)	PII Redacted Report	.pdf, .json
	Archer Dx Pan-Solid FusionPlex			Targeted RNA Read Alignment	.cram, .crai
Tumor RNA	note: minoeq min replace	Archer Analysis v. 6.0 in AWS	Fusion/ITD detection	Archer Fusion Results	.txt
	Archer panel for fusion calling			PII Redacted Report	.pdf, .json
				Illumina EPIC Array Intensity File	.idat
Tumor DNA	,	DKFZ Classifier v. 12.5	Disease classification	DKFZ Classification	.html
				PII Redacted Report	.pdf, .json
				Biospecimen and clinical data	.json

Institutional reimbursement for completion of data submission

CCDI Hub Explore Dashboard: Inventory of Data Sets

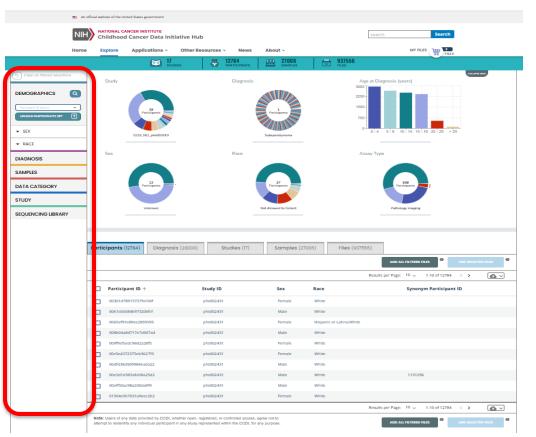
- An inventory of CCDI-managed childhood cancer data
- Provides:
 - Faceted search
 - Visualization of search results
 - Export results for further analysis
- Facilitates data discovery across studies, enabling the building of cohorts





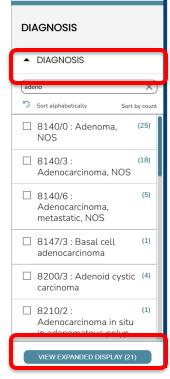
CCDI Hub Explore: Faceted Filtering

- Filter available data by various annotations of interest
 - Diagnosis
 - Demographics
 - Treatment
 - Treatment response
 - Survival
 - Samples
 - Data category
 - Study
- Facet selections dynamically reload visualization dashboard and tabular lists

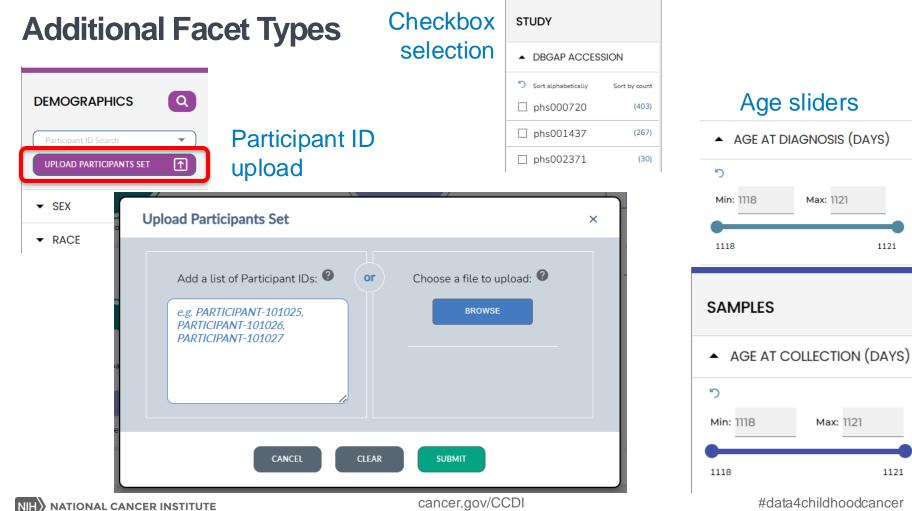


CCDI Hub Explore: Diagnosis Search

- By default, first several diagnoses are listed sorted by numeric code
- Entering a search term will reduce the space of listed diagnoses to the first several matches
- Clicking the "View Expanded Display" button will show all possible matches
- Checking the associated checkbox will update the visual dashboard and table lists accordingly



7	Diagnosis Facet Search	×					
Search Diagnosis (adeno Q) 🤊							
3 selections of 21 search results 5 Sort alphabetically Sort by							
8140/3 : Adenocarcinoma, (18) NOS	8140/6 : Adenocarcinoma, (metastatic, NOS	adenocarcinoma, NOS (4)					
□ 8140/0 : Adenoma, NOS (25)	B147/3 : Basal cell adenocarcinoma) 🗌 8200/3 : Adenoid cystic (4) carcinoma					
8210/2 : Adenocarcinoma (1) in situ in adenomatous polyp	8220/3 : Adenocarcinoma in adenomatous polyposis coli) 28255/3 : Adenocarcinoma (1) with mixed subtypes					
8260/3 : Papillary (9) adenocarcinoma, NOS	□ 8271/0 : Lactotroph adenoma) 8272/0 : Pituitary adenoma, (3 NOS					
8330/0 : Follicular (1) adenoma, NOS	8331/3 : Follicular adenocarcinoma, well differentiated)					
8380/3 : Endometrioid adenocarcinoma, NOS (1)	B480/3 : Mucinous denocarcinoma	 8574/3 : Adenocarcinoma (1) with neuroendocrine differentiation 					



[#]data4childhoodcancer

CCDI Hub Explore: File-based Access

- Files can be added to the My Files cart from any of the tab lists
- Options differ by table
 - **Example 1:** add all the files for a Participant from the "Participants" tab
 - Example 2: add single files from "Files" tab
- Download study metadata from Studies tab

Partio	cipants (4615) Diagnosis (124	416)	Studies (1)	Samples (67	20)	Files (3657	72)			
							ADD ALL	FILTERED FILES	ADD SELEC	
							Results per Page:	10 ~ 1-1	0 of 36572 < ゝ	
3 row(s) selected									
•										•
8	File Name	File Category	File Description	File Type ↑	File Size	Study ID	Participant ID	Sample ID	GUID	MD5sum
	IGM_PBCNBA-0DVSBN_20240329.clinical_ report.json	Clinical data	IGM Clinical JSON	json	517 Bytes	phs002790	PBCNBA		dg.4DFC/a48632eb- fc7a-4527-b62a- 1d067c9e2679	b3f693931ecb2856 21f5e7bd7a55558
	IGM_PBCNBA-0DVSBW_20240402.clinical _report.json	Clinical data	IGM Clinical JSON	json	11.81 KB	phs002790	PBCNBA		dg.4DFC/dca0a1a1- adc6-44e1-b5a2- 9e1e552c1a00	a8098e1a93069046 22c333e3cb1cc3de
	IGM_PBCNBB-0DVSBO_20240329.clinical_ report.json	Clinical data	IGM Clinical JSON	json	517 Bytes	phs002790	PBCNBB		dg.4DFC/b40996b8- d607-46a9-9b75- 2926113fe90e	78e0fbde284cae72 7981ed8909d8bca
ATIO				cance	r aov/(CDI			do /DEC/895186a1- #	data4childho
ΙΑΙΙΟΙ	NAL CANCER INSTITUTE			Cance	1.90%				TT	

Obtain Controlled-access Files

- 1. The database of Genotypes and Phenotypes (dbGaP) access is given using eRA Commons accounts.
 - Go to the <u>eRA Commons site</u> and create an account under your organization or institution.

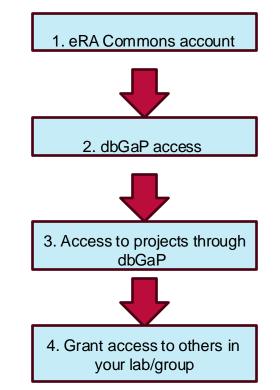
2. Go to the dbGaP Controlled Access Data section and select Authorized Access. Log in with your eRA Commons account.

3. Create a Research Project.

- Select the projects you would like controlled access to.
- Create a Research Use Statement explaining the need for the projects.
- Confirm project structure and send off for review to the Data Access Committee.

4. Go to the My Requests tab to see all current access that is linked to your eRA Commons Account.

 Go to the "Downloaders" tab and search for other members in your lab/group and add them to the selected Research Projects



How to Apply for Controlled Access on dbGaP: <u>https://www.youtube.com/watch?v=m0xp_cCO7kA</u>

How Do I Access MCI FASTQ Files for Rhabdomyosarcoma?

Go to ccdi.cancer.gov/explore

- Select "Molecular Characterization Initiative" under Study Name or choose "phs002790" under Study dbGaP Accession
- Download the Study Manifest
- Find the tabs ending with "_file" (look for FASTQ, JSON, etc.)
- Create a Data Repository Service Manifest for Cloud Analysis
- Export the manifest to Cancer Genomics Cloud (CGC):
 - In CGC, go to your Project → Files → Add Files → GA4GH Data Repository Service (DRS)
 - Files will load into your project for analysis

Participants (4994) Diag	nosis (16238)	Studies (1)	Samples (10912)	Files (147422)	ADD ALL FILTERED FI	LES	ADD SELECTED FILES
4						Results per Page: 10	✓ 1-1 of 1	
Study ID	Access	Diagnosis (Top 5)		Diagnosis Anatomic Site (Number of Fop 5) Participants	Number of Samples	Number of Files	File Type (Top
phs002790	٥	see diagnosis_comme 9380/3 : Glioma. mali 9421/1 : Pilocytic astr 9470/3 : Medulloblast Low-Grade Glioma (3 Basd More	gnant (517) ocytoma (461) toma, NOS (427)	C72.9 : Central nervous system C71.9 : Brain. NOS (3206) C42.0 : Blood (2282) Invalid value (1885) Not Reported (1215) Read More	(4339) 4994	10912	147422	fastq (47408) json (32278) pdf (9605) vcf (7326) cram (7324) Bead More

WY FILES

To access and analyze files, select or deselect any files as needed, and then choose an option from the 'Available Export Options' dropdown by clicking either 'Export to Cancer Genomics Cloud' or 'Download File Manifest'.				USER	GUIDE	AVAILABLE EXPORT OPTIONS	
File Name 个	Study Name	Study Accession	Participant ID	Sample ID	File Type		DAD MANIFEST
207678140100_R01 C01_Grn.idat	Molecular Characterization Initiative	phs002790	PBCXUF	0E2LK3	idat	13.05 MB	Ō
207678140100_R01 C01_Red.idat	Molecular Characterization Initiative	phs002790	PBCXUF	0E2LK3	idat	13.05 MB	Ō

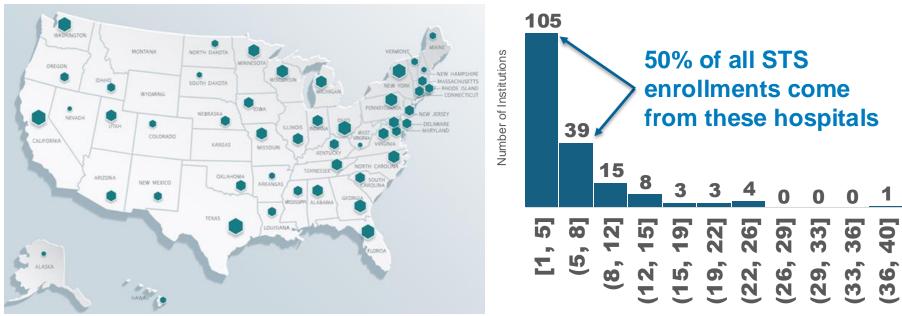
MCI Data Access: FAQs

- Does the lab data exist in a structured format, or is it expected that we use NLP on the PDF documents to make it computable?
 - Deidentified clinical sample reports are available in both PDF and JSON formats.
- Are there institutional IDs attached to the data so we can organize it by institutions in the platform?
 - Institutional IDs are not attached to the data to reduce the risk of re-identification, especially for extremely rare pediatric cancers.
- Is it possible to download the de-identified summary reports directly without going through the dbGaP approval process?
 - The summary reports were not provided in this manner due to potential inclusion of sensitive information.

MCI-Soft Tissue Sarcoma Data Clinical Impact



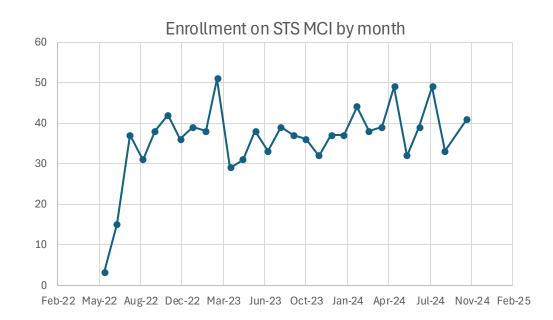
Distribution of Enrollment of Soft Tissue Sarcoma (STS) Patients



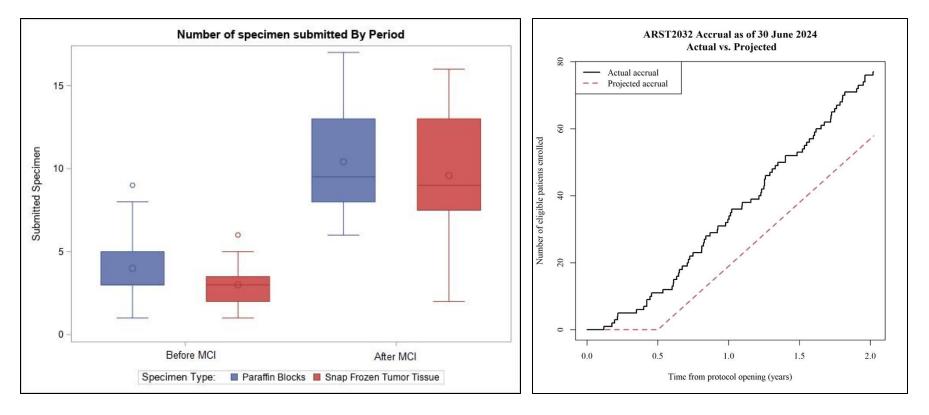
Number of submitted samples

MCI - STS Study: As of 10/29/2024

- 1,043 enrolled patients
- 851 (82%) of enrollees have a submitted paired blood/tumor specimen
- 777 (75%) pass QC and have completed profiling
- 2,170 completed assays



Increased Submission of Samples with Opening of MCI

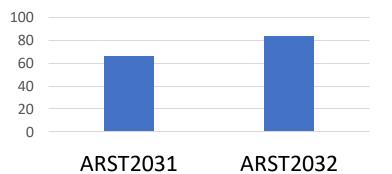


Sapna Oberoi; Wei Xu COG

Co- enrollment on Soft Tissue Sarcoma Clinical Trials

- MCI is required for eligibility on ARST2032 for low-risk rhabdomyosarcoma
- Transition to "Regimen M" for patients with MYOD1 or TP53 mutant tumors allowed at Week 3 or 6

Patients co-enrolled on APEC MCI and clinical trials



2. SOMATIC CANCER-ASSOCIATED SEQUENCE VARIATION IN THE TUMOR

Variant Information	Genomic Change (GRCh38)	Etiology	Germline Variant Allele Fraction (%)	Tumor Variant Allele Fraction (%)	Variant Classification (AMP/ASCO/CAP)
NRAS (NM_002524.5) c.35G>C p.Gly12Ala	chr1:114716126 C>G	Somatic	not detected	66%	Tier I (Level B)
TP53 (NM_000546.6) c.404G>T p.Cys135Phe	chr17:7675208 C>A	Somatic	not detected	2%**	Tier I (Level B)
NF1 (NM_001042492.3) c.6420_6427+18del p.?	chr17:31336907 _31336932del	Somatic	not detected	24%*	Tier I (Level B)

*The variant allele frequency/fraction for this variant may be underestimated due to its nature as a deletion event.

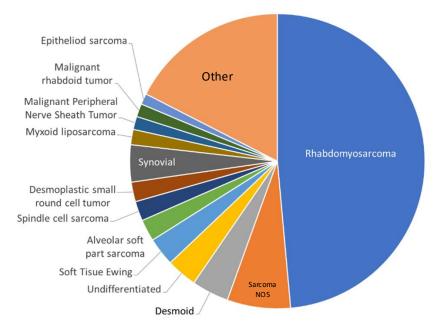
3. PERTINENT NEGATIVES

TP53 Germline Sequence Variants - No pathogenic or likely pathogenic germline single nucleotide variants or small insertion-deletion events were detected.

TP53 Copy Number - No biallelic loss detected. However, a single copy loss of TP53 was detected on 17p in the tumor specimen.

Soft Tissue Sarcoma patients enrolled in the first year of MCI

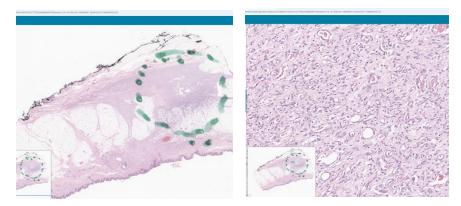
Total Number of patients is 425		
Median age at diagnosis	11.2 (0.0	07, 25.86)
Variable	Frequency	Percentage
Age at diagnosis	382	89.9
≤ 18 years	302 43	89.9 10.1
> 18 years	43	10.1
Sex		
Female	196	46.1
Male	229	53.9
Race		
American Indian or Alaska Native	5	1.2
Asian	18	4.2
Black or African American	61	14.4
Multiple Races	7	1.7
Native Hawaiian or other Pacific Islander	6	1.4
White	270	63.5
Not Reported or UKN	58	13.7
Ethnicity		
Hispanic or Latino	79	18.6
Not Hispanic or Latino	301	70.8
Not Reported or UKN	45	10.6



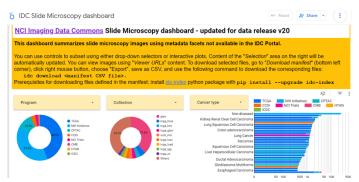
MCI H&E Images Now Available in Imaging Data Commons

- Hematoxylin and eosin (H&E) stained images in Digital Imaging and Communications in Medicine (DICOM) format are now accessible as openaccess through the Imaging Data Commons (IDC)
- Additional images associated will be released in the coming months, new images will be periodically added as participants are enrolled
- Links to images on IDC are accessible through CCDI Hub

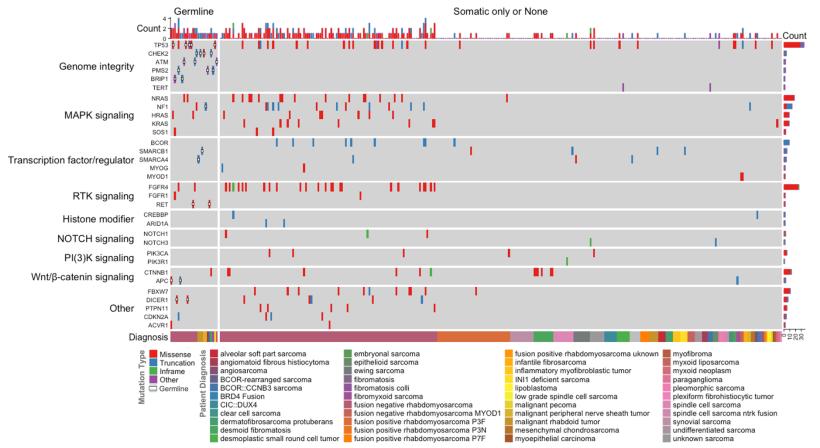
Imaging Data Commons Collections



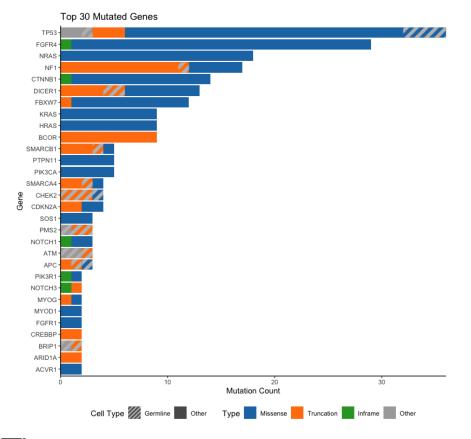
IDC Slide Microscopy Dashboard



Soft Tissue Sarcoma Mutational Summary

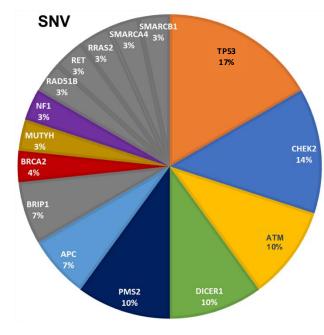


Recurrently Mutated Genes in the STS cohort



- Recurrent gene rank list is largely driven by fusion negative RMS
- Mutation of *FGFR4* is common in fusion negative RMS
- MYOD1 L122R mutations are uncommon (at least at the time of diagnosis)

11% of profiled STS patients have a reportable germline finding



*Single nucleotide variants reported in 1 or more patients

1. GERMLINE CANCER-ASSOCIATED SEQUENCE VARIATION

Gene (Transcript ID)	Genomic Change (GRCh38)	Nucleotide Change	Etiology/Zygosity	Predicted Protein Change	Associated Disease/Condition	Variant Interpretation (ACMG/AMP Evidence)
RET (NM_020975.6)	chr10:43114598 G>C	c.1998G>C	Het	p.Lys666Asn	 (AD) Medullary thyroid carcinoma (OMIN: 155240) (AD) Multiple endocrine neoplasia IIA (OMIN: 171400) (AD) Multiple endocrine neoplasia IIB (OMIN: 12300) (AD) Pheochromocytoma (OMIN: 171300) 	Pathogenic (PS1, PS3_Moderate, PS4_Moderate, PM2, PM5)

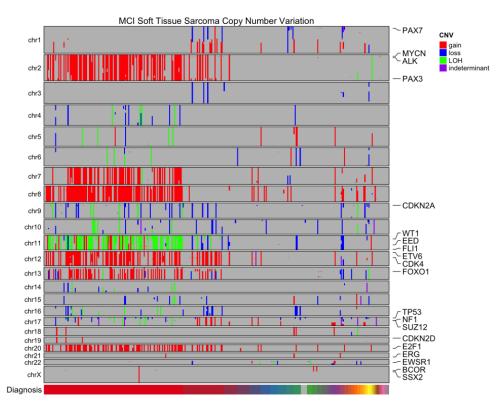
NM_020975.6(RET):c.1998G>C p.(Lys666Asn)

This individual's comparator peripheral blood specimen harbors a heterozygous missense variant in the *RET* gene, which is one of the receptor tyrosine kinases that plays a key role in cell proliferation and differentiation (PMID: 16979782). Pathogenic variants in this gene are associated with autosomal dominant disorders; medullary thyroid carcinoma (OMIM: 155240), multiple endocrine neoplasia IIA (OMIM: 171400), multiple endocrine neoplasia IIB (OMIM: 162300) and pheochromocytoma (OMIM: 171300). A germline *RET* variant has also been reported in the setting of relapsed leukemia (PMID: 30936199).

This missense variant (p.Lys666Asn) is a well-described pathogenic variant as it has been reported in the setting of medullary thyroid carcinoma in multiple unrelated individuals in the literature. *In-vitro* studies showed that this variant results in increase phosphorylation of ERK and high kinase activity (ACMG/AMP: PS4_Moderate, PS3_Moderate; PMIDs: 27673361, 15858153, 20103606). This variant is present at an extremely low frequency in gnomAD, a large-scale control population database (ACMG/AMP: PM2) and this same amino acid change resulting from a different nucleotide change (c.1998G>T) has been reported as pathogenic in literature and in ClinVar by multiple clinical laboratories (ACMG/AMP: PS1; ClinVar ID: 24932). Additionally, a different amino acid change at this same residue (Lys666Glu) has been reported as pathogenic in literature and in ClinVar by accordance with ACMG/AMP: guidelines, this variant is classified as pathogenic.

"Rhabdomyosarcoma NOS"

Soft Tissue Sarcoma Copy Number Summary



Diagnosis fusion negative rhabdomyosarcoma fusion positive rhabdomvosarcoma P3F synovial sarcoma desmoid fibromatosis spindle cell sarcoma ewing sarcoma malignant rhabdoid tumor unknown sarcoma CIC::DUX4 desmoplastic small round cell tumor INI1 deficient sarcoma alveolar soft part sarcoma fusion positive rhabdomvosarcoma P7F myxoid liposarcoma BCOR::CCNB3 sarcoma dermatofibrosarcoma protuberans inflammatory myofibroblastic tumor malignant peripheral nerve sheath tumor undifferentiated sarcoma angiomatoid fibrous histiocytoma epithelioid sarcoma fibromyxoid sarcoma pleomorphic sarcoma spindle cell sarcoma ntrk fusion clear cell sarcoma embryonal sarcoma fusion negative rhabdomyosarcoma MYOD1 infantile fibrosarcoma mesenchymal_chondrosarcoma BCOR-rearranged sarcoma BRD4 Fusion angiosarcoma fibromatosis fibromatosis colli fusion positive rhabdomyosarcoma P3N fusion positive rhabdomyosarcoma uknown lipoblastoma low grade spindle cell sarcoma malignant pecoma myoepithelial carcinoma myofibroma myxoid neoplasm paraganglioma plexiform fibrohistiocytic tumor

- Fusion Negative RMS has a very distinct CN profile
- Focal deletion and amplification of *CDKN2A/CDK4* occurs in ~20% of the population
- Multiple tumors especially patients with Li Fraumeni have evidence of chromothripsis

STS Fusion Oncogene Summary – Archer Results

Detected Fusion Oncogenes in MCI STS

COL1A1	PDGFB
HEY1	
SRF	
	NCOA2
NSD3	
РАХЗ	FOX01
PAX7	NCOA1
SS18	SSX2
	SSX1
EWSR1	WT1
ETV6	FLI1
- RBPMS	ATF1
- MYO5A	CREB1 -
	ERG
ASPSCR1	ETV1
NONO	
	BEND2
BCOR	KLF5 —
	NTRK3
KIAA1549 BCL6	TEE3
EML4	CCNB3
-FGFR1	BRAF
SORT1	DDIT3
FUS	
CLTC	RAB11FIP1
- TRPM1	NOTCH2
CARS	ALK
	0050010
	CREB3L2
CIC	FXR2
SEC16A	DUX4
TFG	NOTCH1 -
GOPC	ROS1 -
	USP6;TP53 -
-TBL1XR1	NTRK1 -
-PRKD3	
	RAF1 —
COL1A2	
- MEG3	PLAG1 -
MGA	
BRD4	NUTM1 -
	CITED1
CCDC6	RET
NCOA4	
-CRTC1	MAML2
CAPZA2	MET
CCDC127	TERT

"Soft tissue sarcoma NOS"

Gene Fusion	5' Fusion Partner	3' Fusion Partner	Classification (AMP/ASCO/CA
EWSR1::KLF5	EWSR1:NM_001163287.2 Exon: 9 (GRCh37) chr22:29686457	<i>KLF5</i> :NM_001730.4 exon: 2 (GRCh37) chr13;73636660	Tier I (Level B)

"Spindle cell neoplasm"

Gene Fusion	5' Fusion Partner	3' Fusion Partner	Classification (AMP/ASCO/CAF
EML4::NTRK3	<i>EML4:NM_019063.4</i> exon: 2 (GRCh37) chr2:42472827	NTRK3:NM_001012338.2 exon: 14 (GRCh37) chr15:88576276	Tier I (Level A)

NIH



Access CCDI resources through the CCDI Hub



CCDI Data Access:

- Access guide: <u>https://ccdi.cancer.gov/static/media/CCDI_Usage_Instructions_Nov2024_v2.5.0.69ea3c_d5.pdf</u>
- Information about CCDI study data: <u>datacatalog.ccdi.cancer.gov/resource/CCDI</u>
- Cancer Data Standards Registry and Repository: <u>cadsr.cancer.gov/onedata/Home.jsp</u>

Data Models and Federation API:

- CCDI: github.com/CBIIT/ccdi-model
- CCDI template: <u>github.com/CBIIT/Model_to_Submission</u>
- Childhood Cancer Clinical Data Commons: github.com/CBIIT/c3dc-model
- CCDI Federation API: <u>cbiit.github.io/ccdi-federation-api</u>

IH NATIONAL CANCER INSTITUTE

Team Members

- Patients and Families
- Advocacy Community
- Local treatment, pathology and surgical teams
- Children's Oncology Group Project EveryChild Team
- Children's Oncology Group Soft Tissue Sarcoma Committee
- Nationwide Biopathology Center Team
- Institute for Genomic Medicine Team
- NCI Childhood Cancer Data Initiative Team

Yassmine Akkari, Institute for Genomic Medicine Michael A. Arnold, University of Colorado Natalie Bir, Biopathology Center Catherine E. Cottrell. Institute for Genomic Medicine Avery Funkhouser, National Cancer Institute Doug Hawkins, Seattle Childrens Hospital Subhashini Jagu, National Cancer Institute Javed Khan, National Cancer Institute Corinne M. Linardic, Duke School of Medicine Elaine R. Mardis. Institute for Genomic Medicine Mariam Mathew. Institute for Genomic Medicine Yvonne Moyer, Biopathology Center Sapna Oberoi, University of Manitoba Nilsa C. Ramirez, Biopathology Center Greg Reaman, National Cancer Institute Erin R. Rudzinski, Seattle Childrens Hospital Kathleen Schieffer, Institute for Genomic Medicine Malcolm Smith, National Cancer Institute Shountea Stover, Biopathology Center Diana Thomas, Biopathology Center Rajkumar Venkatramani, Texas Children's Hospital Wei Xue, University of Florida Aaron Weiss, Maine Medical Center Greg Wheeler, Institute for Genomic Medicine





How You Can Engage with CCDI



Learn about CCDI and subscribe to our monthly newsletter: cancer.gov/CCDI



Access CCDI data and resources: ccdi.cancer.gov



Questions? Email us at: NCIChildhoodCancerDataInitiative@mail.nih.gov



Thank you for attending!



cancer.gov/espanol

cancer.gov