

# Genomic Methods to Understand the Genetic Origins of Pediatric Cancers

Jinghui Zhang, PhD

Chair, Member

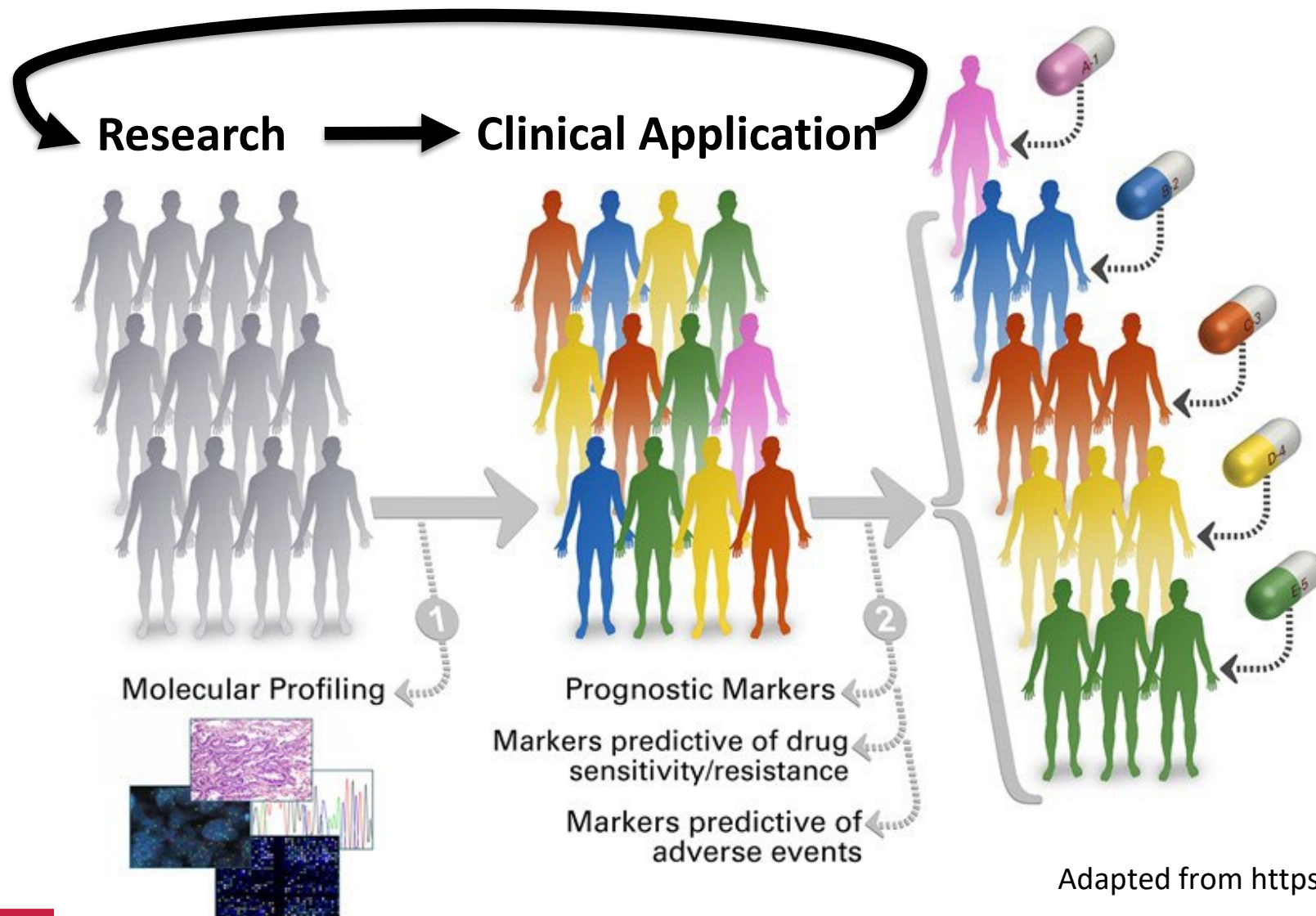
Endowed Chair of Bioinformatics

Department of Computational Biology

St. Jude Children's Research Hospital

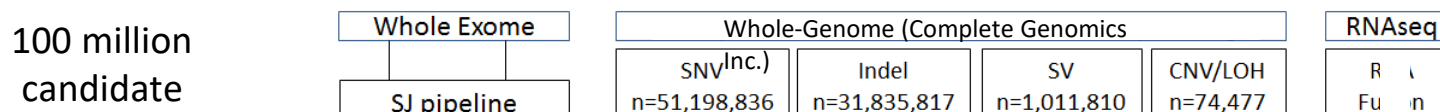
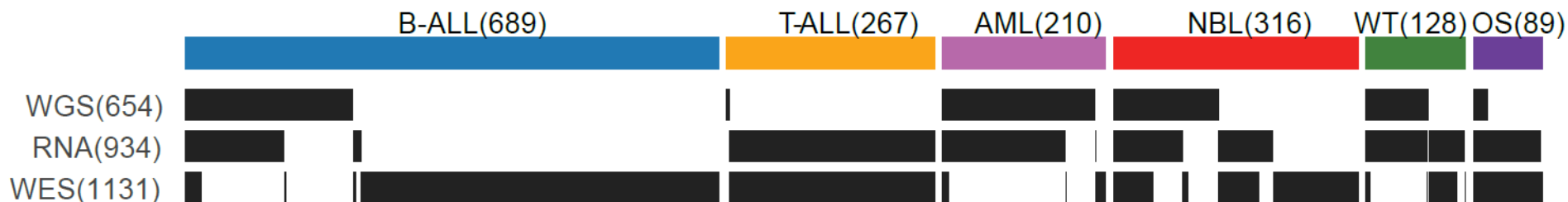
Memphis, TN; USA

# Genomic Analysis in Research & Clinic



Adapted from <https://pct.mdanderson.org/#/>

# Map the Genetic Drivers in Tumor Genome by Variant Enrichment Analysis



Of the 142 statistically significant driver genes:

- 78 (55%) absent in three adult pan-cancer studies
  - Kandoth et al Nature (2013)
  - Lawrence et al Nature (2013)
  - Zack et al Nature Genetics (2013)
- 62% are Structural/Copy Number Variations (SVs/CNVs)

142 driver  
genes

Statistical  
analysis  
(MutSigCV, GRIN)

Pathogenicity  
classification

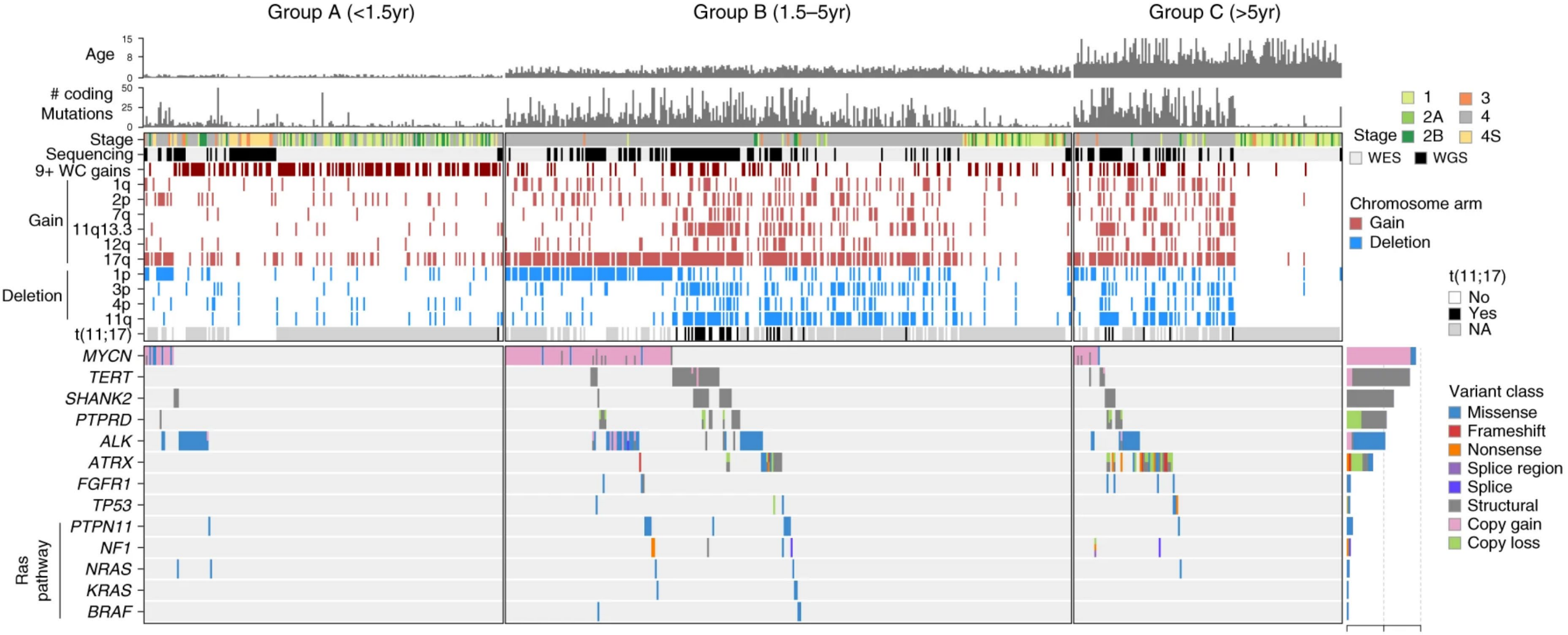
62 additional  
driver genes  
with P/LP  
variants



Ma et al, Nature 2018

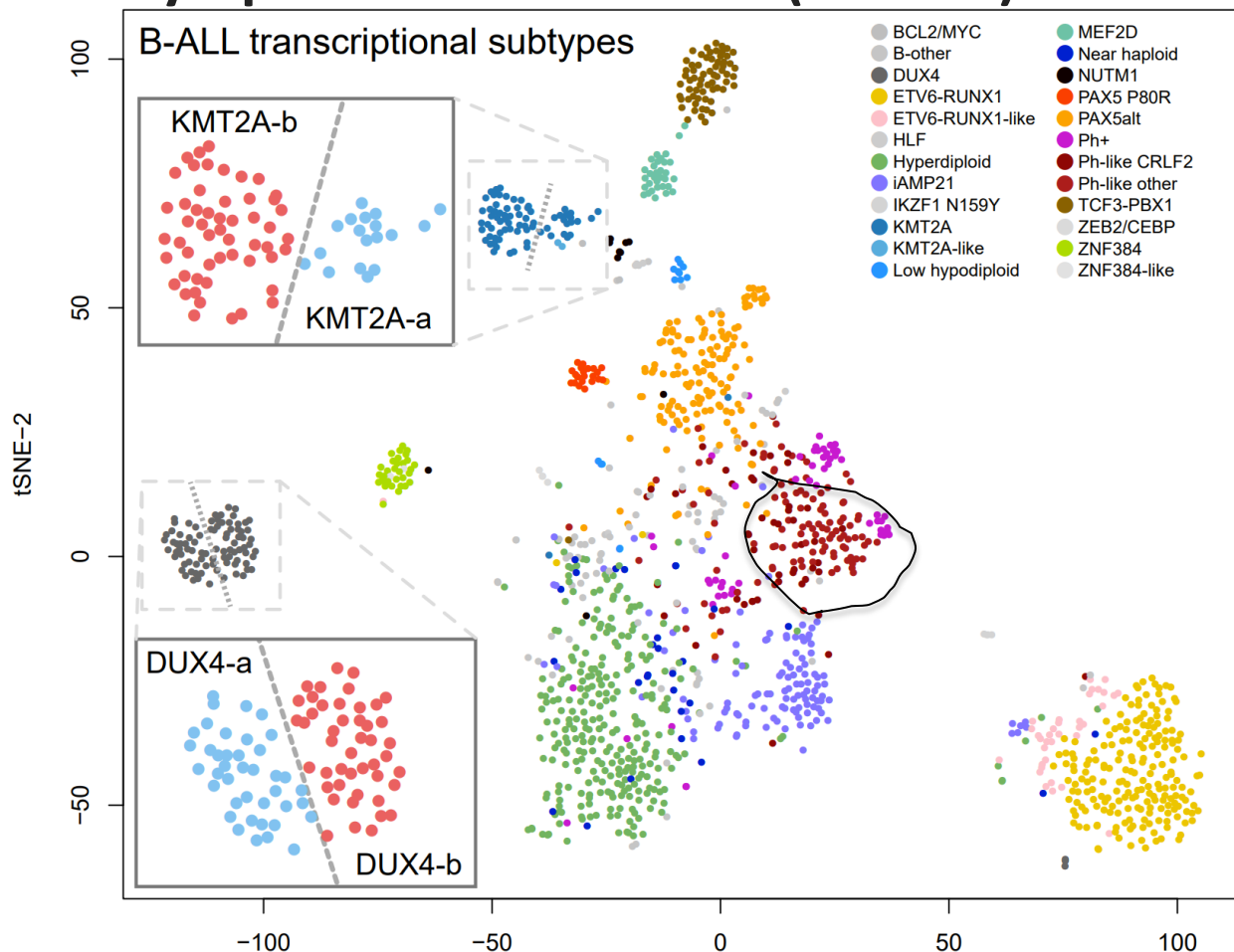
# Define Genetic Basis by Age of Onset

## Pan-neuroblastoma Study of ~700 Patients



# Define Genetic Basis by Following the Cluster Pattern of Expression/Methylation Data

## The genomic landscape of pediatric acute lymphoblastic leukemia (n=2754)

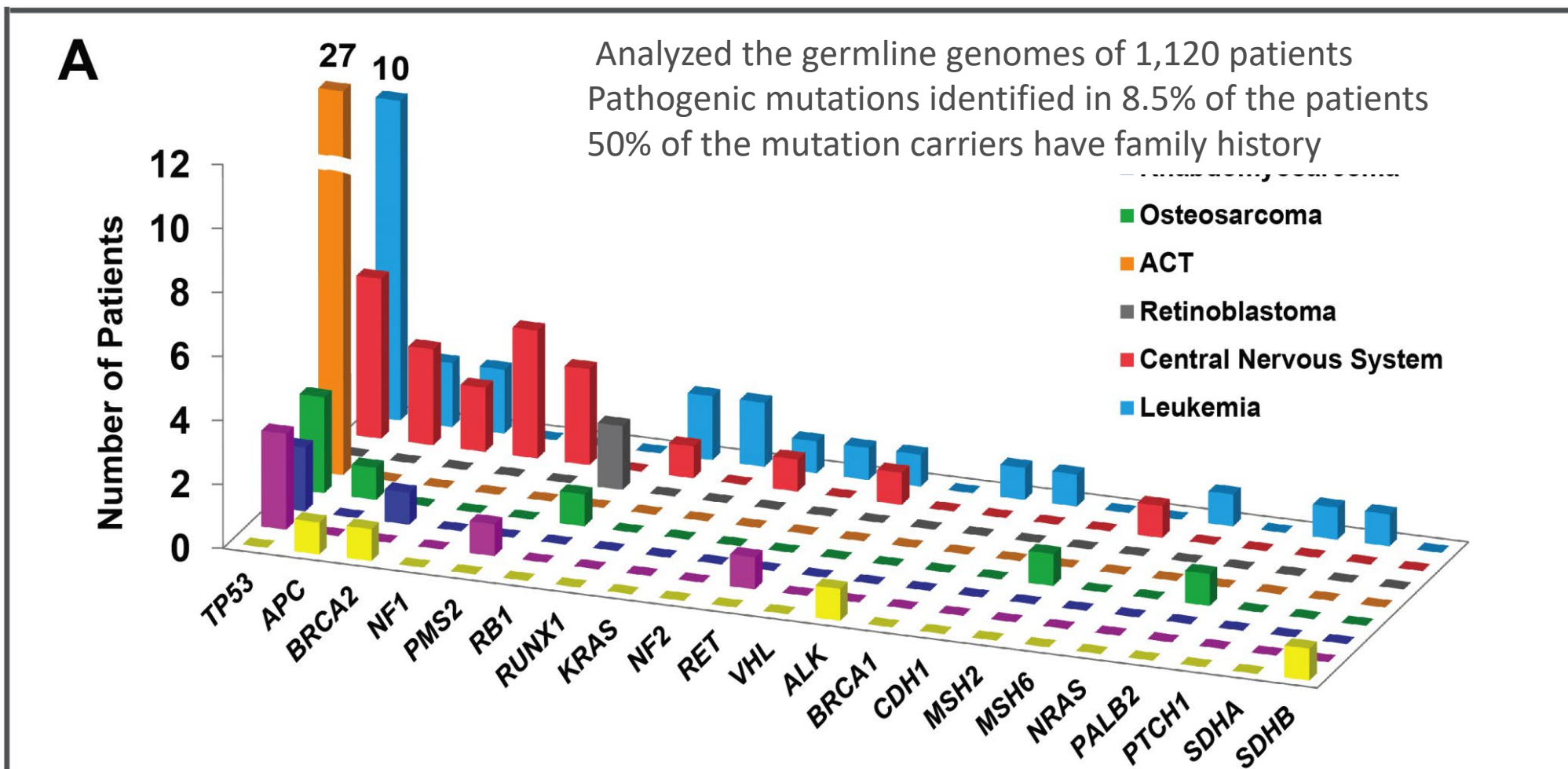


## Approaches

1. Perform clustering using RNA-seq expression of methylation data.
  - Samples are organized by genetic alterations or cell of origin or both
2. Analysis of genetic variants for samples within a cluster can identify cluster-specific recurrent drivers (e.g. NUTM1 fusion)
3. Heterogeneous drivers affecting the same program can be identified (e.g. Ph-like ALL)
4. Additional collaborative mutations led to substructure within a cluster



# Germline Predisposition in Pediatric Cancer

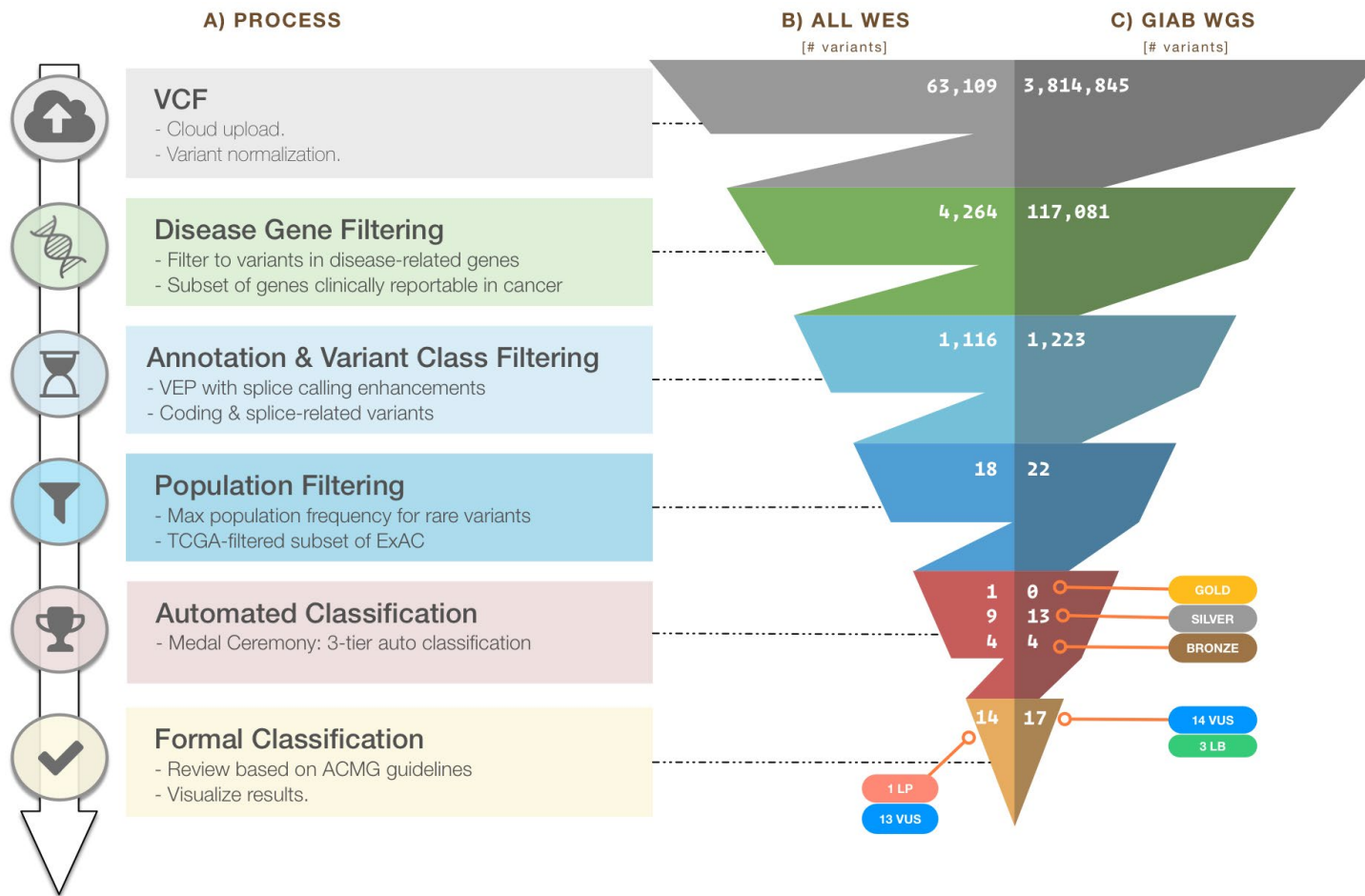


J. Zhang et al, NEJM 2015

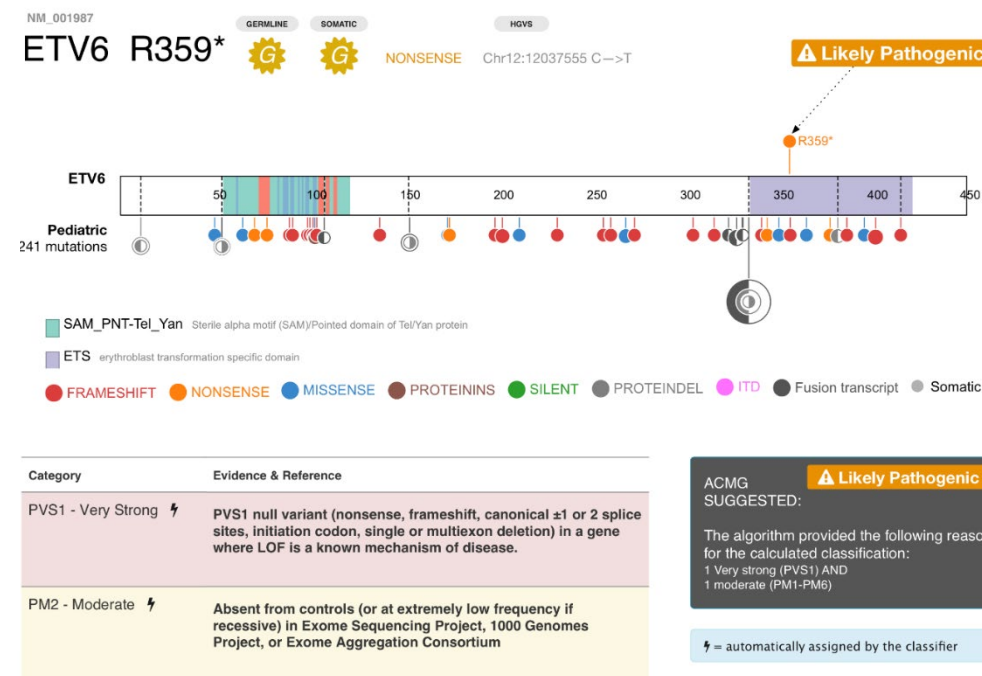
# Germline Variant Classification by PeCanPIE

## Pediatric Cancer Variant Pathogenicity Information Exchange

[https://platform.stjude.cloud/tools/pecan\\_pie](https://platform.stjude.cloud/tools/pecan_pie)



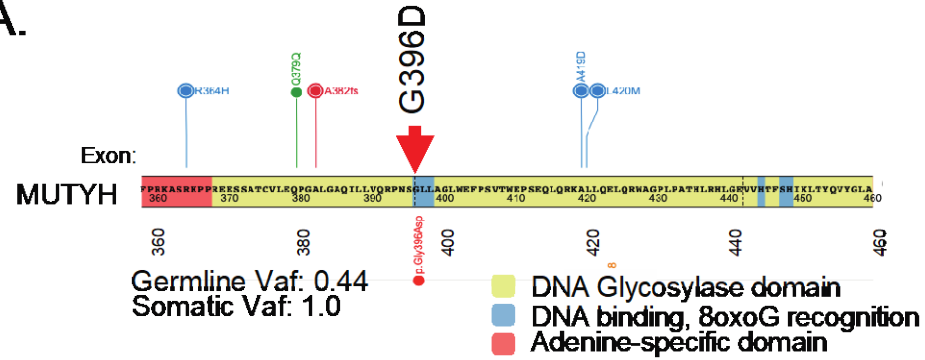
- ❑ Used for JCO 2018, Lancet Oncology 2015, Cancer Discovery 2021
- ❑ Used for clinical and research



Edmonson et al, Genome Research 2019

# Tumor WGS Mutational Signature for Germline Variant Classification

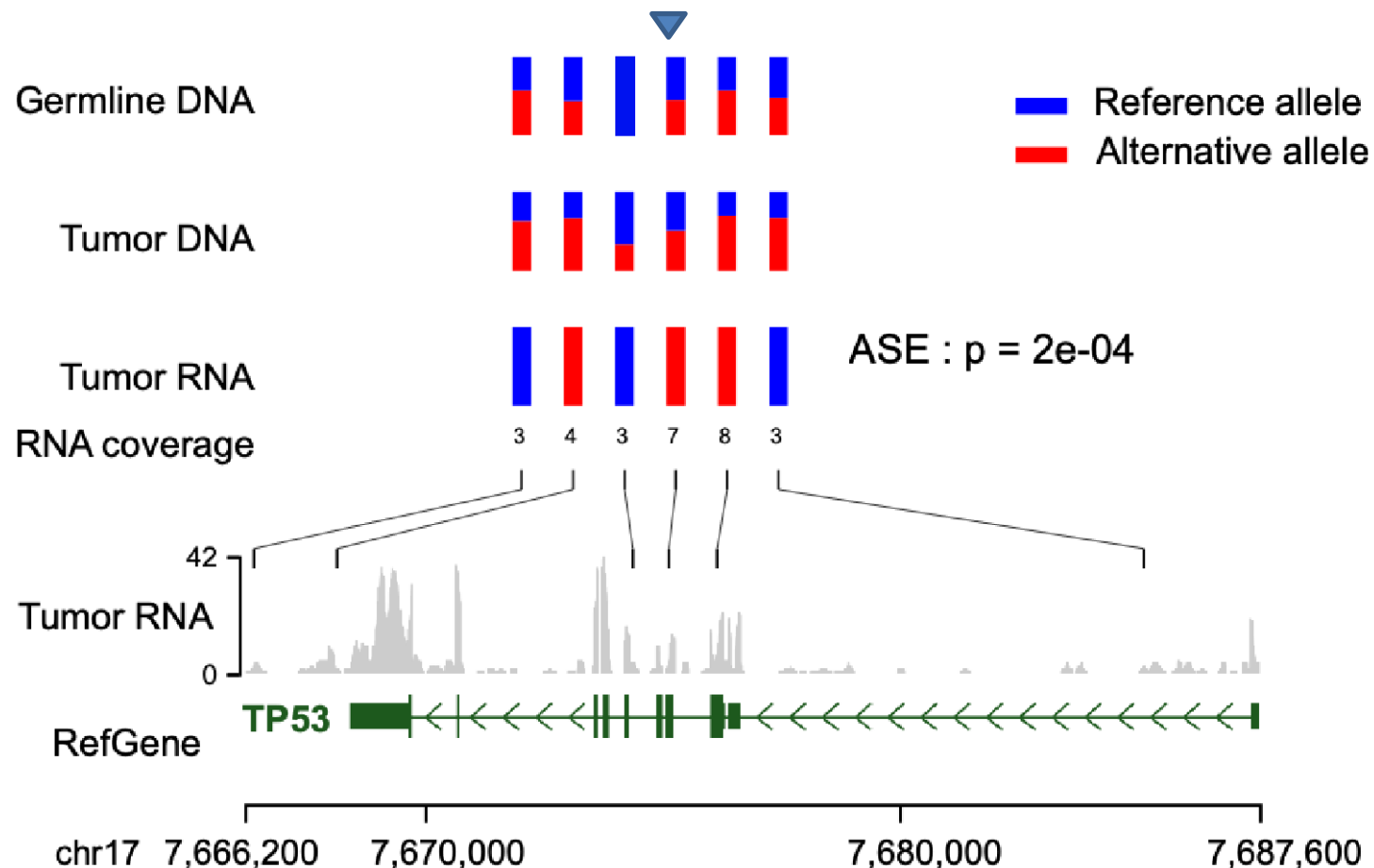
A.





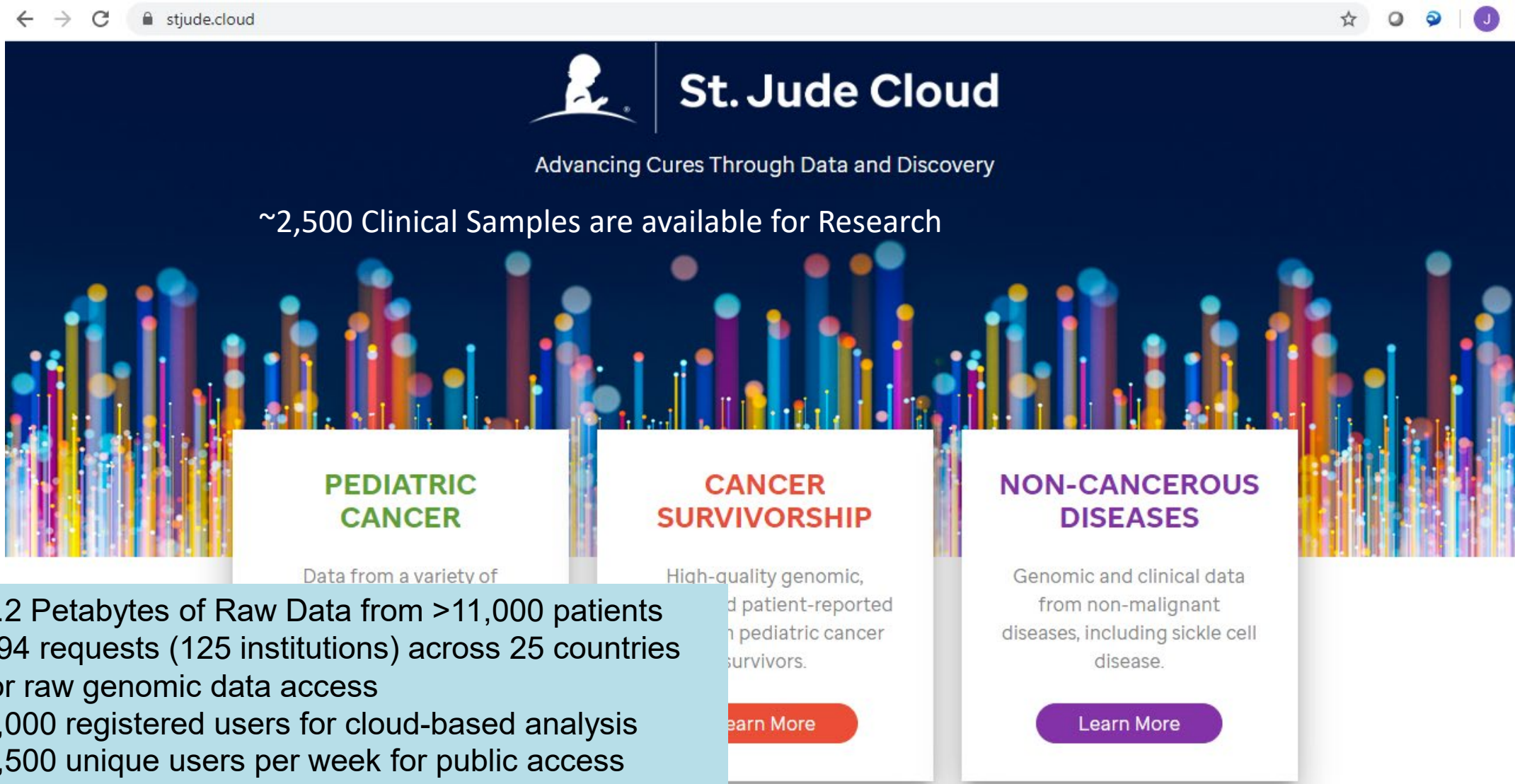
# Tumor Allelic-Specific Expression (ASE) for Germline Variant Classification

**TP53 A161T in a ganglioneuroblastoma patient.**



# Data Sharing & Visualization on St Jude Cloud

<https://www.stjude.cloud/>



The screenshot shows the St. Jude Cloud website. At the top, there is a navigation bar with the St. Jude logo and the text "St. Jude Cloud". Below this, a tagline reads "Advancing Cures Through Data and Discovery". A large banner features a background of colorful, vertical, glowing lines of varying heights and colors (blue, green, yellow, orange, red, purple). Overlaid on this banner are three white rectangular boxes, each representing a different data category. The first box is titled "PEDIATRIC CANCER" in green text and mentions "Data from a variety of". The second box is titled "CANCER SURVIVORSHIP" in red text and mentions "High-quality genomic, and patient-reported" and "n pediatric cancer survivors.". The third box is titled "NON-CANCEROUS DISEASES" in purple text and mentions "Genomic and clinical data from non-malignant diseases, including sickle cell disease.". Each box has a "Learn More" button at the bottom. The browser's address bar shows "stjude.cloud".

**PEDIATRIC CANCER**  
Data from a variety of

**CANCER SURVIVORSHIP**  
High-quality genomic, and patient-reported  
n pediatric cancer survivors.  
[Learn More](#)

**NON-CANCEROUS DISEASES**  
Genomic and clinical data from non-malignant diseases, including sickle cell disease.  
[Learn More](#)

- 1.2 Petabytes of Raw Data from >11,000 patients
- 394 requests (125 institutions) across 25 countries for raw genomic data access
- 3,000 registered users for cloud-based analysis
- 2,500 unique users per week for public access