Spring Webinar

Empowering Discovery through a Wide Variety of Data Types

May 11, 2022
2-4 pm ET

Follow @kidsfirstdrc
Visit kidsfirstdrc.org
Webinar Agenda

2:00 pm – Collaborate to Accelerate Discoveries in Pediatric Research
Host: Dr. Marcia Fournier, NIH Kids First Program Manager

2:15 pm – Keynote: Cancer Risk in Children with Birth Defects: How Kids First Can Move Observation into Understanding
Keynote Speaker: Dr. Philip Lupo, Kids First Investigator

2:45 pm – Kids First Variant Search Tool: Exploring Datasets to Jump Start Analyses
Guest Speakers: Dr. Adam Resnick & Dr. David Higgins, Kids First Data Resource Center

3:15 pm – Kids First Cloud Credits Pilot Announcement
Guest Speaker: Valerie Cotton, NIH NICHD

3:30 pm – Relevant NIH Mechanisms to Support Research Using Kids First Data & Resources
Guest Speaker: Dr. James Coulombe, NICHD

3:45 pm – Q&A
Collaborate to Accelerate Discoveries in Pediatric Research

Marcia Fournier, PhD
Program Manager
Gabriella Miller Kids First Pediatric Research Program
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Marcia Fournier  @MarciaFournier2
Introduction about Kids First Program
Pediatric Cancer Awareness and Scientific Hypothesis Brought the Kids First Program Together

Oct 2013
Gabriella Miller, childhood cancer advocate, died at 10 with Brain cancer

April 2014
Bipartisan bill Gabriella Miller Kids First Research Act signed by congress authorizes $12.6 million per year for ten years to NIH support pediatric research

Sept 2015
NIH Kids First Program first annual appropriation. The program is funded through 2024
Scientific Hypothesis: Intersection of Cancer and Pediatric Conditions

Sept 2012
First Birth Defect and Cancer Workshop by NIH

Shared mutations: BRAF, MAPK, ALK

The New York Times

Birth Defects Tied to Higher Cancer Risk

Major birth defects are associated with an increased risk for cancer in childhood. New research suggests the risk persists into adulthood.

By Nicholas Bakalar
Dec. 7, 2020

Original Investigation

June 20, 2019

Association Between Birth Defects and Cancer Risk Among Children and Adolescents in a Population-Based Assessment of 10 Million Live Births

Philip J. Lupo, PhD1,2, Jeremy M. Schraw, PhD3, Tania A. Desrosiers, PhD4, et al

Author Affiliations | Article Information

Cancer risk in individuals with major birth defects: large Nordic population based case-control study among children, adolescents, and adults

Dagrun Slettebø Daltveit,1 Kari Klungsøy,1,2 Anders Engeland,1,2 Anders Ekbom,1 Mika Gissler,1,5 Ingrid Glimelius,1,5 Tom Grotmol,1,5 Laura Madanat-Harjuoja,2,10 Anne Gulbekh Ording,1,11 Solbjørg Makalani Myrvig Sæther,1,12 Henrik Toft Sørensen,1,11 Rebecca Troisi,1,11 Tone Bjørge1,12
Who is Behind the Kids First Program

<table>
<thead>
<tr>
<th>Kids First Working Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Eunice Kennedy Shriver</em> National Institute of Child Health and Human Development (NICHD)</td>
</tr>
<tr>
<td>National Human Genome Research Institute (NHGRI)</td>
</tr>
<tr>
<td>National Heart, Lung, and Blood Institute (NHLBI)</td>
</tr>
<tr>
<td>National Cancer Institute (NCI)</td>
</tr>
</tbody>
</table>

Other Working Group Representation:

- NIDCR
- NIAAA
- NIDDK
- NEI
- NIAID
- ORIP
- NIDA
- NINDS
- NIEHS
- NIAMS
- NCATS
- CDC

The **Gabriella Miller Kids First Pediatric Research** is a trans-NIH initiative that enables researchers, clinicians, and patients to work together to accelerate collaborative research and promote new discoveries for children affected with cancer and structural birth defects.
Kids First Program Addresses Genomic Studies for Pediatric Cancer and Structural Birth Defects

- **Grow a data resource in the cloud** to accelerate pediatric cancer and structural birth defects research leading to better prevention, diagnosis, and treatments for patients and families
  - FAIR Data
  - Data sharing
  - Data analysis and visualization
  - Collaborative research
What Differentiates the Kids First Data & Resources

- Program offers opportunity of whole genome sequencing
- Public data available with phenotypes and clinical data
- Multiple pediatric conditions and childhood cancer
  - Commonalities/differences between phenotypes
  - Trios – patients and families
  - Variant call
How Kids First Program Supports Research
Research Program Empowered by Sequencing Centers and Data Resource Center

X01
PAR-22-054, Discovery of the Genetic Basis of Childhood Cancers and of Structural Birth Defects: Gabriella Miller Kids First Pediatric Research Program (X01 Clinical Trial Not Allowed)

1. **Selected investigators** send DNA/ RNA samples to Sequence Centers

2. **Sequencing Centers** generate genomic data with investigators’ samples

3. **Data Resource Center**
   - Compile sequencing with phenotypes and clinical data provided by investigators
   - Release data to investigators
   - Investigators have exclusivity of the data for 6-months
Kids First Accelerates Research Through Data Sharing

- Harmonized data, tools, and resources in the cloud
- Almost 3,000 registered users since 2018 launch

Quick Start

1. Join the Kids First Portal or Log In and fill out your user profile.
2. Go to your settings page to connect your account to GEN3 and CAVATICA to perform analysis.
3. Browse the data in the File Repository, performing queries based on your research interests.
Kids First Data Sharing Fast Growing

Approx. 20% growth in Kids First data access since Fall 2021

1. **56 sequencing projects**
2. **28 Publications**
3. **24 studies at dbGaP**
4. **442 data access requests approved to date**

Recent Public Data Releases

# Data Access Request Approvals (as of 5-5-22)

- phs002330.v1.p1
- phs002276.v1.p1
- phs002187.v1.p1
- phs002172.v1.p1
- phs002162.v1.p1
- phs002161.v1.p1
- phs002130.v1.p1
- phs001997.v1.p1
- phs001987.v1.p1
- phs001878.v2.p1
- phs001846.v1.p1
- phs001806.v1.p1
- phs001785.v1.p1
- phs001738.v1.p1
- phs001714.v1.p1
- phs001436.v1.p1
- phs001420.v1.p1
- phs001410.v1.p1
- phs001247.v1.p1
- phs001228.v1.p1
- phs001178.v1.p1
- phs001168.v2.p2
- phs001138.v4.p2
- phs001110.v3.p1
Kids First Data Release since Fall 2021

- Almost 4000 new subjects

**Kids First: Myeloid Malignancies** 408 subjects
- First Portal Release: September 23, 2021
- Data Types Available: Aligned Reads, Individual gVCFs, Family-Based VCFs
- Sequencing Center: HudsonAlpha Institute with additional harmonized data generated by the DRC
- About the Study: NIH X01 Project Abstract - Soheil Meshinchi, PI

**Kids First: T-Cell ALL** 1327 subjects
- First Portal Release: November 1, 2021
- Data Types Available: Aligned Reads, VCFs
- Sequencing Center: HudsonAlpha Institute with additional harmonized data generated by the DRC
- About the Study: NIH X01 Project Abstract - David Teachey, PI

**Kids First: Leukemia & Heart Defects in Down Syndrome** 2067 subjects
- First Portal Release: September 28, 2021
- Data Types Available: Aligned Reads, Individual gVCFs, Family-Based VCFs
- Sequencing Center: Broad Institute with additional harmonized data generated by the DRC
- About the Study: NIH X01 Project Abstract - Philip Lupo and Stephanie Sherman, PIs
NIH Support Mechanisms for Kids First Data Analysis

- **RO3**
  - RFA-RM-22-006, Expert-Driven Small Projects to Strengthen Gabriella Miller Kids First Discovery (R03 Clinical Trial Not Allowed) – Due date June 10!
  - PAR-19-375, Small Research Grants for Analyses of Gabriella Miller Kids First Pediatric Research Data (R03 Clinical Trial Not Allowed)
  - PAR-20-060, Small Research Grants for Establishing Basic Science-Clinical Collaborations to Understand Structural Birth Defects (R03 Clinical Trial Not Allowed)

- **R01**
  - PAR-20-137, In-Depth Phenotyping and Research Using IMPC-Generated Knockout Mouse Strains Exhibiting Embryonic or Perinatal Lethality or Sub viability (R01 Clinical Trial Not Allowed).
  - PAR-21-229, Screening and Functional Validation of Human Birth Defects Genomic Variants (R01 Clinical Trial Not Allowed)
Kids First Data is Enabling Research

- Over 70 NIH Research Grants Using Kids First Data & Resources

Number of Grants from Aug 2016 to May 2021

Cooperative agreements
Future with Multiple Data Types, Collaboration and Interoperability
Kids First Phase 2: $12.6M/year (FY22-24)

- Additional generation of childhood cancer and structural birth defects-related -omics data
  - Add multiple data types such as epigenomic and proteomic
  - Add longitudinal data, data enrichment with phenotypes, clinical data, and outcomes

- Continue development & improvement of the Data Resource
  - Enable discovery and data generation
  - Cross diseases collaborations
  - Interoperability

- Expert-driven activities to increase the value of Kids First data
  - Engage Kids First & community experts in activities such as integration, curation, and/or harmonization of rich clinical and phenotypic data
Kids First is Part of a Larger Data Ecosystem

Innovation across the Phenotypic Translational Divide Webinar Series

NIH Cloud-Based Platforms Interoperability (NCPI):
Empower end-user analyses across platforms through federation & interoperability
Keynote

Cancer Risk in Children with Birth Defects: How Kids First Can Move Observation into Understanding

Philip Lupo, PhD, MPH
Epidemiologist and Professor of Pediatrics
Baylor College of Medicine
Cancer Risk in Children with Birth Defects: How Kids First Can Move Observation into Understanding

Philip Lupo, PhD, MPH
Department of Pediatrics
Section of Hematology-Oncology
Baylor College of Medicine
● **Cancer risk in children with birth defects**: observations from registry-based studies

● **Building from Kids First to discover novel congenital anomaly-cancer syndromes**: Genetic Overlap Between and Anomalies and Cancer in Kids (GOBACK) Study

● **Leveraging Kids First to identify children with congenital anomalies who are more likely to develop cancer**: Down syndrome-acute lymphoblastic leukemia
Birth defects and cancer risk

- **Chromosomal anomalies**
  - Example: Trisomy 21 and acute lymphoblastic leukemia

- **Single gene defects**
  - Example: Costello syndrome and rhabdomyosarcoma

- **Non-syndromic birth defects**
  - Multifactorial; cancer risk largely unknown

March of Dimes; Taub JW, *J Pediatr Hematol Oncol*. 2001; Genetics Home Reference
(Some) research questions and the role of Kids First

1. Which birth defects are associated with which cancers?

2. Do specific birth defect-cancer associations represent undiscovered Mendelian syndromes?

3. Why do some children with birth defects develop cancer while others do not?
GOBACK (Genetic Overlap Between Anomalies and Cancer in Kids) Study

Texas Michigan Arkansas North Carolina

Birth Defect Registries Linking Cancer Registries

>10 million births
GOBACK registry linkage cohort

Unaffected children
N ≈ 9.6M

Children with cancer but no birth defect
N ≈ 15,000

Children with a birth defect who do not develop cancer
N ≈ 524,000

Children with both birth defects and cancer
N ≈ 2,100
## Risk of any cancer in children with birth defects

<table>
<thead>
<tr>
<th>Birth Defect</th>
<th>Cancer</th>
<th>N Comorbid</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any non-chromosomal birth defect</td>
<td>Any childhood cancer</td>
<td>1,740</td>
<td>2.6 (2.4-2.7)</td>
</tr>
<tr>
<td>Any chromosomal birth defect</td>
<td>Any childhood cancer</td>
<td>383</td>
<td>11.6 (10.4-12.9)</td>
</tr>
</tbody>
</table>
Specific “non-syndromic” birth defect-cancer associations

- Tested 72 birth defect-cancer associations (≥5 co-occurring cases)
- 40 birth defect-cancer associations were significant after correcting for multiple comparisons
## Non-syndromic birth defects and childhood cancer

<table>
<thead>
<tr>
<th>Birth Defect</th>
<th>Cancer</th>
<th>HR (95% CI)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ventricular septal defect</td>
<td>Hepatoblastoma</td>
<td>10.6 (5.8-19.2)</td>
</tr>
<tr>
<td>Pulmonary valve atresia</td>
<td>Hepatoblastoma</td>
<td>22.6 (9.1-55.7)</td>
</tr>
<tr>
<td>Pulmonary valve atresia</td>
<td>Neuroblastoma</td>
<td>7.6 (3.8-15.3)</td>
</tr>
<tr>
<td>Left ventricular outflow tract defects</td>
<td>Neuroblastoma</td>
<td>7.8 (3.5-17.3)</td>
</tr>
<tr>
<td>Biliary atresia</td>
<td>Non-Hodgkin lymphoma</td>
<td>164.2 (77.8-346.8)</td>
</tr>
<tr>
<td>Craniosynostosis</td>
<td>Hepatoblastoma</td>
<td>9.7 (4.3-22.2)</td>
</tr>
<tr>
<td><strong>Choanal atresia</strong></td>
<td><strong>Acute leukemia</strong></td>
<td><strong>9.2 (3.8-22.1)</strong></td>
</tr>
</tbody>
</table>

¹ Adjusted for maternal age, child’s sex and state of birth. Models including hepatoblastoma are adjusted for birthweight. Models including ventricular septal defect are adjusted for birthweight and gestational age.
Cancer risk increased for children with multiple non-syndromic birth defects

<table>
<thead>
<tr>
<th>Number of Major Birth Defects</th>
<th>Adjusted HR (95% CI)</th>
<th>Number of cancer cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>No defect</td>
<td>1.00</td>
<td>13,111</td>
</tr>
<tr>
<td>1 defect</td>
<td>1.32 (1.20-1.45)</td>
<td>477</td>
</tr>
<tr>
<td>2 defects</td>
<td>3.51 (3.19-3.56)</td>
<td>446</td>
</tr>
<tr>
<td>3 defects</td>
<td>4.62 (4.08-5.22)</td>
<td>261</td>
</tr>
<tr>
<td>≥4 defects</td>
<td>5.85 (5.31-6.44)</td>
<td>432</td>
</tr>
</tbody>
</table>

JAMA Oncol. 2019 Jun 20;5(8):1150-8
GOBACK family cohort

**GOBACK**

<table>
<thead>
<tr>
<th>Texas</th>
<th>Michigan</th>
<th>North Carolina</th>
<th>Arkansas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth Defect Registries</td>
<td>Linking</td>
<td>Cancer Registries</td>
<td></td>
</tr>
</tbody>
</table>

>10 million births

Identify novel associations

Recruit cohort of families

Whole genome sequencing

Variant analysis

SNVs/indels

*SVs

Sharon Plon, MD, PhD

Aniko Sabo, PhD

**GOBACK TO THE BASES**

Genetic Overlap Between Anomalies and Cancer in Kids
De novo ~5kb one exon deletion in USP9X

Genomic location: chrX:41066285-41071603

PCR validation of the heterozygous deletion

Saumya Sisoudiya
Phenotype of the proband

Female
Birth defects
  ○ Coloboma
  ○ Heart defects
  ○ **Choanal atresia**
  ○ Ear anomalies
  ○ Genitourinary anomalies
Indicative of CHARGE syndrome but did not have a CHD7 pathogenic variant

Cancer: **Precursor cell lymphoblastic leukemia**

**Choanal atresia-acute leukemia HR=9.2, 95% CI: 3.8-22.1**
**USP9X** involved in several developmental and cancer pathways

Adapted from Murtaza et. al., *Cell and Mol Life Sci.*, 2015

Created with BioRender.com
Hypothesis: **USP9X** is a novel ALL susceptibility gene associated with a CHARGE-like syndrome.

Ma et al, Nature, 2018
Leverage Children’s Oncology Group Project: EveryChild (PEC)

1. Determine the frequency of known cancer predisposition variants among children with congenital anomalies and cancer.

2. Identify variants that underlie novel anomaly-cancer predisposition syndromes and describe the landscape of somatic alterations in these children.

Sharon Plon
Logan Spector
Children’s Oncology Group (COG)

~ 200 research sites throughout United States
~ 90% of children diagnosed with cancer in the US cared for at COG member sites

As of 2/2021
APEC14B1, Project EveryChild Registry

Future Contact: optional information

**Patient Contact Information**

- **Patient's Address:**
  - Street Address:
  - City: [ ]
  - State/Province: [ ]
  - Zip/Postal Code: [ ]
  - Country: [ ]

- **Phone Number:**
  - City: [ ]
  - State/Province: [ ]
  - Zip/Postal Code: [ ]
  - Country: [ ]

- **Driver's license number:**
  - Driver's license issued state: [ ]

**Parent/Guardian Future Contact Information**

- **First Parent/Guardian:**
  - First Name: [ ]
  - Middle Name: [ ]
  - Last Name: [ ]

- **Address:**
  - Street Address:
  - City: [ ]
  - State/Province: [ ]
  - Zip/Postal Code: [ ]
  - Country: [ ]

- **First Parent/Guardian date of birth:**
  - Date: [ ]
  - Email address:

- **Second Parent/Guardian:**
  - First Name: [ ]
  - Middle Name: [ ]
  - Last Name: [ ]

- **Address:**
  - Street Address:
  - City: [ ]
  - State/Province: [ ]
  - Zip/Postal Code: [ ]
  - Country: [ ]

- **Second Parent/Guardian date of birth:**
  - Date: [ ]
  - Email address:

Please indicate the language spoken in the home, circle all that apply:
- English
- French
- Spanish
- Other, specify:

**Other Key Contact Information**

- **Key Contact:**
  - First Name: [ ]
  - Middle Name: [ ]
  - Last Name: [ ]

- **Address:**
  - Street Address:
  - City: [ ]
  - State/Province: [ ]
  - Zip/Postal Code: [ ]
  - Country: [ ]

- **Key contact date of birth, if known:**
  - Date: [ ]
  - Email address:

Please indicate the language spoken in the home, circle all that apply:
- English
- French
- Spanish
- Other, specify:

**APEC14B1, Project EveryChild Registry**

1. **Where was the baby born?**
   - City: [ ]
   - State/Province: [ ]
   - Zip/Postal Code: [ ]
   - Country: [ ]

2. **Was this patient a single or multiple birth?**
   - [ ] Single
   - [ ] Twins
   - [ ] Triples or more
   - 2a. If twin, specify: [ ] Identical
     [ ] Fraternal
     [ ] Unknown

3. **Was patient conceived through use of in vitro fertilization?**
   - [ ] Yes
   - [ ] No
   - [ ] Not sure

4. **Was cord blood banked at birth?**
   - [ ] Yes
   - [ ] No
   - [ ] Not sure

5. **Has anyone in the patient's immediate family (biological mother, father, brothers, sisters) ever had cancer?**
   - [ ] Yes
   - [ ] No
   - [ ] Not sure

6. **My child's...**
   - [ ] Mother
   - [ ] Father
   - [ ] Full brother
   - [ ] Full sister
   - [ ] Son
   - [ ] Daughter
   - [ ] Other
   - [ ] Other gender

7. **Does the patient have any structural birth defects known at this time?**
   - [ ] Yes
   - [ ] No
   - [ ] Not sure
   - Cleft lip: [ ]
   - Cleft palate: [ ]
   - Clubfoot: [ ]
   - Gastroschisis: [ ]
   - Heart defect: [ ]
   - Other: [ ]

8. **Does the patient have any known genetic disorder?**
   - [ ] Yes
   - [ ] No
   - [ ] Not sure
   - Down Syndrome: [ ]
   - Turner Syndrome: [ ]
   - Neurofibromatosis Type 1: [ ]
   - Other: [ ]

9. **Does the patient have any known autoimmune disease?**
   - [ ] Yes
   - [ ] No
   - [ ] Not sure
   - Juvenile Idiopathic Arthritis: [ ]
   - Cellulose disease: [ ]
   - Diabetes mellitus (Type I): [ ]
   - Inflammatory bowel diseases (Crohn's or ulcerative colitis): [ ]
   - Other: [ ]

10. **Which describe the patient?**
    - [ ] White
    - [ ] African American
    - [ ] American Indian, Alaskan, Eskimo
    - [ ] Asian
    - [ ] Other
    - [ ] Other Spanish/Hispanic origin includes Europeans

Please return to hospital or clinic staff. Phone: [ ] FAX: [ ] Thank you for your information!
<table>
<thead>
<tr>
<th>Anomaly</th>
<th>Hematologic Malignancies</th>
<th>CNS Tumors</th>
<th>Non-CNS Solid Tumors</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spina Bifida</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Central Nervous System</td>
<td>11</td>
<td>11</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>Eye</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Ear, Face, and Neck</td>
<td>9</td>
<td>2</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>Heart</td>
<td>88</td>
<td>18</td>
<td>53</td>
<td>159</td>
</tr>
<tr>
<td>Circulatory</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Respiratory</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Clefts</td>
<td>14</td>
<td>4</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>16</td>
<td>5</td>
<td>15</td>
<td>36</td>
</tr>
<tr>
<td>Genitourinary</td>
<td>28</td>
<td>6</td>
<td>27</td>
<td>61</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>60</td>
<td>18</td>
<td>54</td>
<td>132</td>
</tr>
<tr>
<td>Integument</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Other/ unspecified</td>
<td>2</td>
<td>0</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>TOTAL</td>
<td>243</td>
<td>73</td>
<td>184</td>
<td>500</td>
</tr>
</tbody>
</table>
Kids First GOBACK Sequencing

- 700 blood and 500 tumor DNA samples for whole genome sequencing at 30X coverage
- 500 tumor DNA samples for whole exome sequencing at 100X coverage
- 120 tumor RNA samples for transcriptome sequencing
7. Does the patient have any structural birth defects known at this time?

<table>
<thead>
<tr>
<th>Birth Defects</th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cleft lip</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cleft palate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clubfoot</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastrochisis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart defect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other specify</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

8. Does the patient have any known genetic disorder?

<table>
<thead>
<tr>
<th>Genetic Disorder</th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down Syndrome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uj Frumeni Syndrome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurofibromatosis Type I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other specify</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. Which of these describe the patient? Check all that apply.

- [ ] White
- [ ] African American
- [ ] American Indian, Eskimo, or Aleut
- [ ] Asian
- [ ] Black or African American
- [ ] American Indian, Eskimo, or Aleut, (except for enrolled American Indians)
- [ ] Other (specify)
- [ ] Unknown
- [ ] Other Spanish/Hispanic origin includes Europeans

Please return to hospital or clinic staff.

Phone: ___________________ FAX: _______________ Thank you for your information!
1. Collect extensive phenotypic and clinical data from children with congenital anomalies and cancer enrolled in Project:EveryChild

2. Integrate phenotypic and clinical data from Project:EveryChild into the Gabriella Miller Kids First Pediatric Data Resource Center

Allison Heath
Adam Resnick
Population-based registries can be leveraged to inform genomic analyses

Birth defects could account for ~10% of childhood cancers

Non-syndromic birth defects are associated with an increased risk for childhood cancer, especially embryonal tumors and germ cell tumors

Reasons for these associations are complex and multifactorial: in some children, inherited and de novo genetic variants likely explain both phenotypes
Down syndrome-associated leukemia
Down syndrome (DS) and leukemia

- First reported 1930
- First systematic study in 1957
- ~20-fold increased risk
- Cumulative risk of 2% by age 5
- Comprises ~2% of childhood acute lymphoblastic leukemia (ALL) and 10% of childhood acute myeloid leukemia (AML)

Hasle et al, Lancet 200
Clinical features of DS - ALL

- Similar age range, except rare in infants
- Distinctive immunophenotype - T-ALL and Burkitt very rare
- Distinctive spectrum of genetic alterations
  - Typical ALL alterations (hyperdiploidy, ETV6-RUNX1) less common
  - CRLF2-R and JAK alterations more common
- Poorer outcomes
  - Increased risk of relapse
  - Increased treatment-related mortality, primarily infections

Li et al, Leukemia 2016
Buitenkamp et al, Blood 2014
Why does ALL arise more often in children with DS?

- Are there germline genetic variants associated with the ALL susceptibility in children with DS?

- How do germline and somatic variants interact within the context of trisomy 21?
## Genome-wide Association Studies (GWASs) of non-DS ALL

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>SNP</th>
<th>OR (95% CI)</th>
<th>P-value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARID5B</td>
<td>10</td>
<td>rs10821936</td>
<td>1.86 (1.71-2.03)</td>
<td>5.9×10^{-46}</td>
<td>Xu et al. 2013</td>
</tr>
<tr>
<td>ARID5B</td>
<td>10</td>
<td>rs7089424</td>
<td>1.65 (1.54-1.76)</td>
<td>6.7×10^{-19}</td>
<td>Papaemmanuil et al. 2009</td>
</tr>
<tr>
<td>IKZF1</td>
<td>7</td>
<td>rs11978267</td>
<td>1.59 (1.45-1.74)</td>
<td>5.3×10^{-24}</td>
<td>Xu et al. 2013</td>
</tr>
<tr>
<td>IKZF1</td>
<td>7</td>
<td>rs6944602</td>
<td>1.64 (1.37-2.07)</td>
<td>3.4×10^{-15}</td>
<td>Papaemmanuil et al. 2009</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>9</td>
<td>rs3731249</td>
<td>2.23 (1.90-2.61)</td>
<td>9.0×10^{-23}</td>
<td>Xu et al. 2015</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>9</td>
<td>rs17756311</td>
<td>1.36 (1.18-1.56)</td>
<td>1.4×10^{-5}</td>
<td>Xu et al. 2013</td>
</tr>
<tr>
<td>CEBPE</td>
<td>14</td>
<td>rs4982731</td>
<td>1.36 (1.24-1.48)</td>
<td>9.0×10^{-12}</td>
<td>Xu et al. 2013</td>
</tr>
<tr>
<td>PIP4K2A</td>
<td>10</td>
<td>rs7088318</td>
<td>1.40 (1.28-1.53)</td>
<td>1.1×10^{-11}</td>
<td>Xu et al. 2013</td>
</tr>
<tr>
<td>GATA3</td>
<td>10</td>
<td>rs3824662</td>
<td>3.85 (2.71-5.47)</td>
<td>2.2×10^{-14}</td>
<td>Perez-Andreu et al. 2013</td>
</tr>
</tbody>
</table>
GWAS of ALL susceptibility in children with DS

Objectives
- Identify inherited genetic variants associated with ALL susceptibility in DS
- Explore association between inherited variation and common somatic alterations
- Compare frequency of established risk alleles in ALL cases with vs without DS

Meta-analysis of 4 cohorts
- Cases – children with DS ALL (n=542)
- Controls – children with DS, no ALL (n=1,192)

Brown et al, Blood 2019
GWAS findings: ALL susceptibility in children with DS

Brown et al, Blood 2019
Table 3. Results from case-case analysis of association between ALL risk alleles and DS status

<table>
<thead>
<tr>
<th>SNP</th>
<th>Pos</th>
<th>Gene</th>
<th>DS-ALL COG molecular subgroup adjusted comparison‡ (255 DS-ALL, 2387 non-DS ALL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11978267</td>
<td>Chr7:50466304</td>
<td>IKZF1</td>
<td>0.97 (0.77-1.23)</td>
</tr>
<tr>
<td>rs3731249</td>
<td>Chr9:21970916</td>
<td>CDKN2A</td>
<td>1.72 (1.10-2.69)</td>
</tr>
<tr>
<td>rs3824662</td>
<td>Chr10:8104208</td>
<td>GATA3</td>
<td>0.81 (0.63-1.06)</td>
</tr>
<tr>
<td>rs12769953</td>
<td>Chr10:22407656</td>
<td>BMI1</td>
<td>1.12 (0.83-1.50)</td>
</tr>
<tr>
<td>rs10741006</td>
<td>Chr10:22856019</td>
<td>PIP4K2A</td>
<td>0.96 (0.75-1.23)</td>
</tr>
<tr>
<td>rs7089424</td>
<td>Chr10:63752159</td>
<td>ARID5B</td>
<td>0.80 (0.64-1.01)</td>
</tr>
<tr>
<td>rs2239633</td>
<td>Chr14:23589057</td>
<td>CEBPE</td>
<td>1.17 (0.92-1.47)</td>
</tr>
</tbody>
</table>

‡Analysis adjusted for top 5 principal components and molecular subgroups (CRLF2 high, high hyperdiploidy, ETV6-RUNX1, and B other).
<table>
<thead>
<tr>
<th>SNP</th>
<th>Position</th>
<th>Gene</th>
<th>CRLF2-high (151 DS-ALL, 55 non-DS ALL)</th>
<th>High hyperdiploid (19 DS-ALL, 888 non-DS ALL)</th>
<th>ETV6-RUNX1 (45 DS-ALL, 547 non-DS ALL)</th>
<th>B-other (40 DS-ALL, 859 non-DS ALL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs11978267</td>
<td>Chr7:50466304</td>
<td>IKZF1</td>
<td>0.71 (0.48-1.05)</td>
<td>1.26 (0.67-2.36)</td>
<td>0.89 (0.55-1.45)</td>
<td>1.37 (0.87-2.17)</td>
</tr>
<tr>
<td>rs3731249</td>
<td>Chr9:21970916</td>
<td>CDKN2A</td>
<td>2.16 (0.96-4.89)</td>
<td>1.62 (0.49-5.43)</td>
<td>1.68 (0.64-4.39)</td>
<td>1.49 (0.60-3.68)</td>
</tr>
<tr>
<td>rs3824662</td>
<td>Chr10:8104208</td>
<td>GATA3</td>
<td>0.73 (0.50-1.06)</td>
<td>1.02 (0.48-2.20)</td>
<td>1.02 (0.59-1.78)</td>
<td>0.74 (0.42-1.29)</td>
</tr>
<tr>
<td>rs12769953</td>
<td>Chr10:22407656</td>
<td>BMI1</td>
<td>1.28 (0.78-2.11)</td>
<td>1.52 (0.58-3.95)</td>
<td>0.75 (0.45-1.22)</td>
<td>1.45 (0.76-2.77)</td>
</tr>
<tr>
<td>rs10741006</td>
<td>Chr10:22856019</td>
<td>PIP4K2A</td>
<td>0.85 (0.56-1.31)</td>
<td>0.75 (0.38-1.48)</td>
<td>0.78 (0.49-1.23)</td>
<td>1.73 (1.00-2.97)</td>
</tr>
<tr>
<td>rs7089424</td>
<td>Chr10:63752159</td>
<td>ARID5B</td>
<td>0.45 (0.30-0.67)</td>
<td>1.38 (0.70-2.74)</td>
<td>1.02 (0.65-1.61)</td>
<td>1.06 (0.68-1.66)</td>
</tr>
<tr>
<td>rs2239633</td>
<td>Chr14:23589057</td>
<td>CEBPA</td>
<td>1.35 (0.93-1.98)</td>
<td>0.92 (0.47-1.80)</td>
<td>1.19 (0.76-1.88)</td>
<td>1.01 (0.62-1.63)</td>
</tr>
</tbody>
</table>

*Analyses comparing DS-ALL and non-DS ALL cases enrolled on COG P9900 or AALL0232 trials. P values and ORs calculated using logistic regression tests assuming additive allelic effects, adjusting for the top 5 principal components.*
GWAS conclusions

- Established non-DS ALL susceptibility loci also contribute to ALL risk in children with DS
  - Genome-wide significance loci at *IKZF1*, *CDKN2A*, *ARID5B*, *GATA3*

- Rather than exhibiting unique susceptibility loci, trisomy 21 appears to modify penetrance of inherited ALL susceptibility
  - Greater magnitude of effect, particularly for *CDKN2A*
Next steps

Whole genome sequencing (WGS) for comprehensive analysis of heritable variation associated with ALL in children with DS

(INvestigation of Co-occurring conditions across the Lifespan to Understand Down syndrome)
Collaboration: WGS in DS - ALL and AVSD

- Children with DS have 2,000-fold increased risk of atrioventricular septal defect (AVSD) and 10-fold increased risk of ALL.
- Genetic factors underlying these associations unknown.

- **Central hypothesis:** risk-associated genetic variants in the background of trisomy 21 lead to higher penetrance of AVSD and ALL.
- **Secondary hypothesis:** rare variants explain a significant proportion of this increased risk.

Karen Rabin
Jun Yang
Stephanie Sherman
1. Identify genetic variants underlying AVSD in children with DS

2. Identify genetic variants underlying ALL in children with DS
   - Particular attention to rare, structural, and chromosome 21 variants
   - Evaluation of relationship between germline and somatic features (WGS of paired leukemia-germline samples)
DS-ALL sequencing: normal and tumor samples

- Normal: N=167
- Paired: N=362
- Tumor: N=6
Plans underway

- Perform a comprehensive analysis of heritable variation associated with risk of ALL in children with DS
  - Assess structural, rare, and chr21 variants
  - Evaluate relationship between the inherited genome and somatic features

- Conduct deep phenotyping of children with DS-ALL to identify congenital risk factors for DS-ALL and their impact on leukemia clinical features and outcomes
  - Determine co-occurring conditions among children in COG biology and/or registry protocols
  - Link to other data resources
DS-ALL conclusions

- Children with DS and ALL have a distinctive clinical presentation, which reflects underlying differences in somatic alterations and germline risk factors

- Known non-DS ALL susceptibility loci also contribute to ALL risk in children with DS, but with differing effect sizes

- Rare, structural, and chr21 variants may also explain a proportion of ALL risk
Overall conclusions

- Kids First provides a unique opportunity to explore the overlap between birth defects and cancer
- Some associations may be driven by uncharacterized syndromes – Kids First data can be leveraged to explore this possibility
- Insights into factors influencing cancer among children with birth defects may guide improved genetic counseling, surveillance, and treatment interventions
Acknowledgements

Baylor/TXCH
- Sharon Plon
- Karen Rabin
- Jeremy Schraw
- Austin Brown
- Olga Taylor
- Dani Mitchell
- Lauren Sanclemente
- Michael Scheurer
- Jacob Junco
- Saumya Sisoudiya
- Tiffany Chambers

COG
- Logan Spector
- Mignon Loh
- Stephen Hunger
- Meenakshi Devidas
- Yunfeng Dai
- Michael Borowitz
- Brent Wood
- Nyla Heerema
- Andrew Carroll

Crnic Institute
- Joaquin Espinosa

Kids First
- Valerie Cotton
- James Coulombe
- Marcia Fournier

St Jude
- Jun Yang
- Gang Wu
- Ti-Cheng Chang
- Wentao Yang
- Zhongshan Cheng
- Dale Hedges
- Jeremy Hunt

The patients and families who participated in this research

Emory
- Stephanie Sherman
- Elizabeth Leslie
- David Cutler
- Mike Zwick
- Tracie Rosser

CPRIT

NIH National Cancer Institute

INCLUDE PROJECT

Emory University School of Medicine

HGSC Human Genome Sequencing Center

Baylor College of Medicine

Texas Children’s Cancer Center
Email: GOBACK_Studystudy@bcm.edu | T 1-855-474-4520
Kids First Variant Search Tool: Exploring Datasets to Jump Start Analyses

Adam Resnick, PhD
Co Director
Center for Data Driven Discovery in Biomedicine
Children’s Hospital of Philadelphia

David Higgins, PhD
Informatics Program Manager
Center for Data Driven Discovery in Biomedicine
Children’s Hospital of Philadelphia

Adam Resnick
@adamcresnick

David Higgins

LinkedIn
LinkedIn
Kids First Variant Search Tool: Exploring Datasets to Jump Start Analyses

Kids First Spring Public Webinar
May 13th, 2022

Follow @kidsfirstdrc
Visit kidsfirstdrc.org
Gabriella Miller Kids First Data Resource Center

- Kids First: Orofacial Cleft: African and Asian Ancestry
- Kids First: Neuroblastoma
- Kids First: Myeloid Malignancies
- Kids First: Esophageal Atresia & Tracheoesophageal Fistulas
- Kids First: Disorders of Sex Development
- Kids First: Leukemia & Heart Defects in Down Syndrome
- Kids First: Novel Cancer Susceptibility in Families (from BASIC3)
- Kids First: Congenital Diaphragmatic Hernia
- Kids First: Enchondromatosis
- Kids First: Orofacial Cleft - European Ancestry
- Kids First: T Cell ALL
- Kids First: Microtia - Hispanic
- Kids First: Intersections of Cancer & SBD
- Kids First: Orofacial Cleft - Latin American
- Kids First: Familial Leukemia
- Kids First: Craniofacial Microsomia
- Kids First: Syndromic Cranial Dysinnervation
- Kids First: Kidney and Urinary Tract Defects
- Kids First: Adolescent Idiopathic Scoliosis
- Kids First: Hemangiomas (PHACE)
- Kids First: Ewing Sarcoma
- Kids First: Nonsyndromic Craniosynostosis
- Kids First: Congenital Heart Defects
- Kids First: Osteosarcoma

24 Studies Released on the Portal

- Participants: 20,267
- Families: 8,462
- Files: 81,927
- Data: 1.09 PB
Across disease
Across modalities
Across disciplines
Across data resources
Across disease
Across modalities
Across disciplines
Across data resources
Across disease
Across modalities
Across disciplines
Across data resources
Gabriella Miller Kids First Data Resource Center

Across disease
Across modalities
Across disciplines
Across data resources
Gabriella Miller Kids First Data Resource Center

Across disease
Across modalities
Across disciplines
Across data resources

Platform Empowered
Across disease
Across modalities
Across disciplines
Across data resources
KFDRC: Connectivity on behalf of discovery

https://anvilproject.org/ncpi

https://includedcc.org/

https://www.nih-cfde.org/
KFDRC: Connectivity on behalf of discovery

https://anvilproject.org/ncpi

https://includedcc.org/

https://www.nih-cfde.org/
Variant Search Tool
Acknowledgements - Variant Search Tool

Dr. Vincent Ferretti
Jeremy Costanza
Adrian Paul
Christophe Botek
Denis Beauregard
Evans Girard
Francis Iavoie
Lucas Lemmonier

Kids First DRC
Partner Institutions
Variant Search allows users to query and search Kids First genetic data to **identify new samples** and **build cross-disease cohorts** to **jump start discoveries** in pediatric cancer and structural birth defects.
## Variant Queries

**Genes Symbol** = MYC and **VEP** = HIGH, MODERATE

**+ New query**

Showing 1 - 20 out of 134

<table>
<thead>
<tr>
<th>Variant</th>
<th>Type</th>
<th>dbsNP</th>
<th>Consequences</th>
<th>CLINVAR</th>
<th>Studies</th>
<th>Part.</th>
<th>Freq.</th>
<th>ALT</th>
<th>Homo</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr8:g.127735492&gt;T&gt;A</td>
<td>SNV</td>
<td>rs560576306</td>
<td>▲ stop_gained MYC L7*</td>
<td>-</td>
<td>1</td>
<td>1/11030</td>
<td>9.07e-5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ upstream_gene_variant CASC11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127735730&gt;del</td>
<td>deletion</td>
<td>--</td>
<td>▲ upstream_gene_variant CASC11</td>
<td></td>
<td>1</td>
<td>2/11030</td>
<td>1.81e-4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ frameshift_variant MYC A86X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127736259&gt;C&gt;T</td>
<td>SNV</td>
<td>rs1399116825</td>
<td>▲ upstream_gene_variant CASC11</td>
<td></td>
<td>1</td>
<td>2/11030</td>
<td>1.81e-4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ stop_gained MYC Q72*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127738438&gt;del</td>
<td>deletion</td>
<td>--</td>
<td>▲ frameshift_variant MYC T73X</td>
<td></td>
<td>1</td>
<td>1/11030</td>
<td>9.07e-5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ upstream_gene_variant CASC11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127735477&gt;G&gt;A</td>
<td>SNV</td>
<td>rs773815056</td>
<td>▲ missense_variant MYC R2H</td>
<td></td>
<td>1</td>
<td>1/11030</td>
<td>9.07e-5</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Navigate to Variant Search using the top menu bar.
Apply filters to the database using the categories on the left.
Filters will appear in the query at the top of the tool.
Results will be shown below, including dbSNP IDs and protein consequences.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Type</th>
<th>dbSNP</th>
<th>Consequences</th>
<th>CLINVAR</th>
<th>Studies</th>
<th>Part.</th>
<th>Freq.</th>
<th>ALT</th>
<th>Homo</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr8:g.127735492T&gt;A</td>
<td>SNV</td>
<td>rs560576306</td>
<td>▲ stop.gained MYC L7*</td>
<td>--</td>
<td>1</td>
<td>1/11030</td>
<td>9.07e-5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ upstream_gene_variant CASC11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ frameshift_variant MYC A86X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127735730del</td>
<td>deletion</td>
<td>--</td>
<td>▲ upstream_gene_variant CASC11</td>
<td>--</td>
<td>1</td>
<td>2/11030</td>
<td>1.81e-4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ frameshift_variant MYC Q72*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127736259C&gt;T</td>
<td>SNV</td>
<td>rs1399116825</td>
<td>▲ upstream_gene_variant CASC11</td>
<td>--</td>
<td>1</td>
<td>2/11030</td>
<td>1.81e-4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ stop.gained MYC Q72*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127738438del</td>
<td>deletion</td>
<td>--</td>
<td>▲ frameshift_variant MYC T73X</td>
<td>--</td>
<td>1</td>
<td>1/11030</td>
<td>9.07e-5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>▲ upstream_gene_variant CASC11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr8:g.127735477G&gt;A</td>
<td>SNV</td>
<td>rs377815066</td>
<td>▲ missense_variant MYC P2H</td>
<td>--</td>
<td>1</td>
<td>1/11030</td>
<td>9.07e-5</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Using the Kids First Variant Search Tool

Search an individual variant
... building on findings from Dr. Mary Marazita's group

Search a chromosomal region
○ ... building on findings from Dr. Sharon Diskin's group

Search based on disease-associated genes
○ ... novel discovery leveraging the power of Kids First datasets.
Orofacial Clefts

• Cleft lip and cleft palate are among the most common birth defects worldwide.

• Genome-wide association studies (GWAS) can be used to identify variations more likely to be present in individuals with an orofacial cleft.

A child with a unilateral, complete cleft lip.

(Image: chop.edu)
Orofacial Clefts

- A 2020 study using Gabriella Miller Kids First sequencing data identified a region on chromosome 21 from a population of patients from Colombia associated with orofacial cleft.

Is this variant present in participants in other Kids First studies?

(Mukhopadhyay et al., 2020)
Neuroblastoma

- **Neuroblasts** are a type of cell found in unborn children which mature into neurons.
- In some children, neuroblasts continue to grow uncontrolled, forming tumors. This cancer type is called **neuroblastoma**.
- Neuroblastomas most frequently arise in adrenal glands but can be found in other tissues as well.

Clinical presentation of neuroblastoma.

(Image: John M. Maris, 2010)
A small deletion on chromosome 16 was found to be more frequently observed in patients with neuroblastoma.

Kids First sequencing data identified these deletions frequently occurred spontaneously.

What variants from other Kids First studies are in this region?

(Egolf et al., 2019)
An atrial septal defect allows blood to move from the left atrium to the right atrium.

A diaphragmatic hernia allows organs such as the liver to develop out of place.

(Images: CHOP)
An atrial septal defect allows blood to move from the left atrium to the right atrium.

A diaphragmatic hernia allows organs such as the liver to develop out of place.

Is there overlap in genes underlying these birth defects?

Are any variants in these genes present in Kids First participants?
Want to learn more about Kids First platforms and data?

Monthly User Support Office Hours

Next Session:
RESCHEDULED
Tuesday, May 17
3:00pm to 4:00pm/ET

To Join, Follow the Link:
tinyurl.com/KidsFirstOfficeHours

Questions? Contact:
higginsd@chop.edu
Vision

Alleviate suffering from childhood cancer and structural birth defects by fostering collaborative research to uncover the etiology of these diseases and supporting data sharing within the pediatric research community.
The Kids First Community is Growing!

X01 Childhood Cancer & Structural Birth Defects Cohorts

Kids First Sequencing Centers

Kids First
PEDIATRIC RESEARCH PROGRAM
Data Resource Center
Key Principle:

Invite a diversity of researchers, developers, and other community members to contribute tools, data, resources, and knowledge to the Kids First ecosystem.

You are part of Kids First!
Expert-Driven Small Projects to Strengthen Gabriella Miller Kids First Discovery (R03)

1. Collecting, extracting, submitting deeper data or new data types associated with Kids First datasets
2. Harmonizing or processing data to promote cross-disease or cross-species (or cross-dataset analysis)
3. Port analysis workflows to deploy within the Kids First Data Resource (e.g., CAVATiCA)
4. Creating or integrating, a new or separate tool to federate with the Kids First Data Resource
5. Consenting for broader data sharing
The Kids First Data Resource for Collaborative Discovery

**CAVATICA**
Pull data from multiple sources into one workspace. Use notebooks, bring-your-own or use available workflows.

**Data Resource Portal**
Enter point, query/search/discover & build synthetic cohorts

**Knowledge Base Integrations (PedcBioPortal)**
Integrations with existing curated/published data visualizations

**Data Services**
Model clinical data in FHIR-based data services for semantic interoperability and coordination

**Framework Services**
Index and point to files in the cloud across NIH (for approved users)
Researcher challenges to data and compute

**Current State**

- Laptop or On Premise
- Shell Based Workflow
- Downloading Data
- Not sustainable

**Future State**

- Cloud Environment
  - CAVATICA
    - (Platform as a Service)
- Community based
  - Standard Workflow Language
- Computational Analysis on a platform
  where the data lives
- Sustainable
Lessons Learned from Early Cloud Pilots

• Early pilots: Open only to Kids First X01 investigators
• Lessons learned:
  o Benefits (positive experiences)
  o Challenges & solutions

Nara Sobreira
Johns Hopkins University
Research Focus: Cartilage tumors and vascular anomalies

Sharon Plon
Baylor College of Medicine
Research Focus: Identifying novel cancer susceptibility mutations

Mary Marazita
University of Pittsburgh
Research Focus: Human Genetics of Complex Traits

Christine Seidman
President and Fellows of Harvard College
Research Focus: Genetic Basis of Structural Heart and other birth defects

Jonathan Rios
UT Southwestern Medical Center
Research Focus: Genomics of orthopaedic disease program
Benefits: Lowers the barrier to using the cloud

- Speed of approval - did not have to wait full grant cycle
  - Request approved by NIH within 4 weeks
  - Quick set up of Cavatica billing group – connected to funds provided through STRIDES

- Safety net for getting started
  - Able to get comfortable with the Cavatica environment without consuming other resources

- Generate preliminary data for NIH grant applications
  - Helps with planning for more robust analysis
Benefit: Scalability

- "the work supported by her credits that took her weeks on CAVATICA would have taken her months to do at her institution"
- CAVATICA enables researchers to activate dozens of machines rapidly to support multiple tasks/runs at a time
Benefit: Ability to use, build & share workflows

Structural Variant Calling

- Run multiple pipelines already available, tune them, and build and share a new pipeline
- Expand this work for long-read structural variant analyses
Benefit: Collaboration

CAVATICA continues to check permissions (e.g., dbGaP approval)
Benefit: Taking advantage of Interoperability

NIH Cloud Platforms Interoperability (NCPI): Empower end-user analyses across platforms through federation & interoperability
Interoperability: Kids First, TOPMed, TARGET, GTEx data all in one workspace
What if a child represented in the Undiagnosed Disease Network had phenotypes similar to children represented in Kids First? Could you use the Kids First Data Resource to better understand the condition?

- Compare phenotypes between UDN and Kids First
- Look for overlapping genetic pathways
Analyzing "data in place"

4,000+ genomes

Up to 24,000 genomes
Lessons learned: Challenges & Solutions

Challenges:
- Workspace storage costs from *output files* or *uploaded data* led to over-spending
- Didn't know how to get started, concerned about costs
- Favorite tools are not in the cloud
- Not all users fully appreciated the value/scalability (especially if currently analyzing only their own dataset)

Solutions:
- New rules on closing billing groups & deleting workspaces that are overspending
- Office hours and trainings
- New [Kids First Cloud Credits Program GitHub Page](#)!
Cloud Credit Requests

- Requests will be reviewed by the NIH Kids First Working Group & chaired by **Huiqing Li, NHBLI**
- Initially can request $1000
  - An additional $1000 if you describe cross-dataset analyses (Kids First + other interoperable data)
- After initial credits are spent; report on progress and request up to $5000 to continue/expand your work
  - Can apply for more if/after you share workflows/pipelines
- Inactive billing groups will be closed after 6 months; Cavatica workspaces will be deleted if they accrue costs beyond the allocation
- Researchers can continue by committing their own funds to a billing group
Cloud Credits: How to Apply

- If you would like to participate in the Kids First 2022 Cloud Credits Pilot program, please send an email to KidsFirst@od.nih.gov with "Kids First 2022 Cloud Credits Pilot" in the subject line, and the Working Group will follow up with additional information.

- In your email, please describe:
  - Which Kids First datasets you are interested in analyzing and whether you plan to cross-analyze Kids First data with other data.
  - Whether you have previously participated in earlier Kids First cloud credit pilots, either as an X01 or R03 investigator. In that case, please indicate whether this is a request for an additional allocation.
  - What analyses you intend to do on the CAVATICA platform
Thank you!
Relevant NIH Mechanisms to Support Research Using Kids First Data & Resources

James Coulombe, PhD
Working Group Coordinator,
Gabriella Miller Kids First Pediatric Research Program
Chief, Developmental Biology and Structural Variation Branch,
Eunice Kennedy Shriver National Institute of Child Health and Human Development
Scientific Vision

Alleviate suffering from childhood cancer and structural birth defects by fostering collaborative research to uncover the etiology of these diseases and supporting data sharing within the pediatric research community.
Kids First Program Addresses Genomic Studies for Pediatric Cancer and Structural Birth Defects

- **Grow a data resource in the cloud** to accelerate pediatric cancer and structural birth defects research leading to better prevention, diagnosis, and treatments for patients and families
  - FAIR Data
  - Data sharing
  - Data analysis and visualization
  - Collaborative research
Public Law 113–94 113th Congress

SEC. 1 This Act may be cited as the “Gabriella Miller Kids First Research Act”.

SEC. 3. 10–YEAR PEDIATRIC RESEARCH INITIATIVE. SUPPLEMENT, NOT SUPPLANT; PROHIBITION AGAINST TRANSFER.

Funds appropriated pursuant to section 402A(a)(2) of the Public Health Service Act, as added by subsection (b)— (1) shall be used to supplement, not supplant, the funds otherwise allocated by the National Institutes of Health for pediatric research; and (2) notwithstanding any transfer authority in any appropriation Act, shall not be used for any purpose other than allocating funds for making grants as described in section 402(b)(7)(B)(ii) of the Public Health Service Act, as added by subsection (a).

Approved April 3, 2
Small Research Grants for Analyses of Gabriella Miller Kids First Pediatric Research Data

- Program Announcement with Referral (PAR)
- R03 small grant mechanism
  - Awards for up to 2 years
  - Budget expanded up to $100,000/year in direct costs
  - Preliminary data is not required
  - 97 applications received
  - 24 received funding
  - PAR-19-375 Application dates until October/November 16th, 2022
Expert-Driven Small Projects to Strengthen Gabriella Miller Kids First Discovery (R03 Clinical Trial Not Allowed)

Awards for up to 2 years
Budget expanded up to $100,000/year in direct costs
Preliminary data is not required

- RFA-RM-21-011
  - Eleven applications received
  - 6 Funded (one through the INCLUDE program)

- RFA-RM-22-006
  - Applications due June 10th, 2022
Investigator Initiated “Parent” FOAs

R01:
PA-20-185 NIH Small Research Grant Program (Parent R03 Clinical Trial Not Allowed)

R03:
PA-20-200 NIH Small Research Grant Program (Parent R03 Clinical Trial Not Allowed)

R21:
PA-20-195 NIH Exploratory/Developmental Research Grant Program (Parent R21 Clinical Trial Not Allowed)
NIH Funding Opportunities

Go shopping: https://grants.nih.gov/funding/searchguide/index.html#

- **R01: PAR-21-229**: Screening and Functional Validation of Human Birth Defects Genomic Variants (R01 Clinical Trial Not Allowed)

- **R21: RFA-CA-22-02**: Development of Innovative Informatics Methods and Algorithms for Cancer Research and Management (R21 Clinical Trial Optional)

- **R21: PAR-20-078**: Secondary Analysis of Existing Datasets in Heart, Lung, and Blood Diseases and Sleep Disorders (R21 Clinical Trial Not Allowed)
Questions?

Thank you!