SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

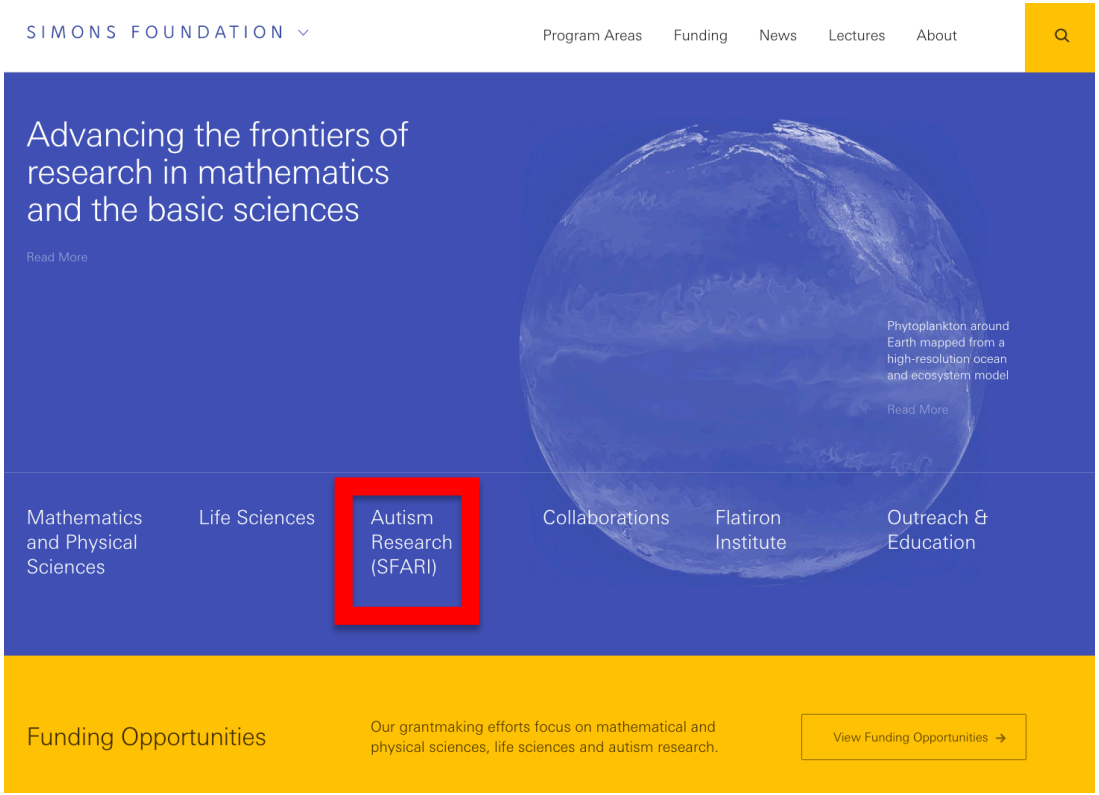
SFARI Resources and Data Sharing Practices

Natalia Volfovsky, PhD
Director, Data & Analytics
Informatics, Simons Foundation

March 25, 2021

Simons Foundation

Endowment: ~$3B
Budget in 2020
~$512M



SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Mission: improve the understanding, diagnosis and treatment of autism spectrum disorders (ASD) by funding innovative research of the highest quality and relevance

Identify risk factors whether genetic, environmental or epidemiological.

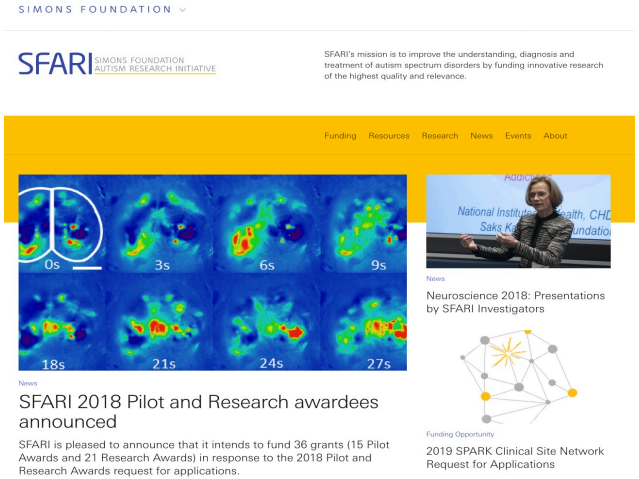
Use non-human organisms to understand how these risk factors alter brain function and animal behavior.

Promote preclinical and clinical investigations to improve autism diagnosis & therapy.

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Simons Foundation Autism Research Initiative (SFARI)

- 2020 budget: **$92 million**
- Since launch in 2003: **>$525 million** in external research support
- >**550 investigators**



@SFARIorg

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

anthonycz/Shutterstock.com

SFARI Cohorts

SSC

Simons Simplex Collection
~10,000 individuals

SIMONS SEARCHLIGHT

Driven by science. United by hope.

~2,0000 individuals

SPARK

Simons Powering Autism Research

More than 300,000 individuals

AIC Autism Inpatient Collection

~1,000 individuals

autism BRAIN NET

~250 individuals

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Mission: improve the understanding, diagnosis and treatment of autism spectrum disorders (ASD) by funding innovative research of the highest quality and relevance

- Current Funding Opportunities: *sfari.org/funding-opportunities*
- Policies and Procedure: *https://www.sfari.org/funding-opportunities/policies-and-procedures/*
- Resources : *sfari.org/resources*
- Data & biospecimen requests, recruitment : *base.sfari.org*

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Data Sharing Policy

SFARI aims to support reproducible scientific research of the highest quality. This can be greatly facilitated by making all raw data, analysis methods, and, when applicable, computer code, available to the research community as quickly and transparently as possible.

Data generated from collections is considered "community resources" and, as such, all these data, including, but not limited to, alignment files and variant calls, will be made available to the entire research community, pre-publication, in real-time, as they are produced.

Besides being subject to any limitations by consent and IRB protocol, a publication embargo will also be enforced for 6 months after the defined project is complete while foundation-sponsored analyses are underway.

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Intellectual Property

The Simons Foundation has no wish to participate in the commercialization of any intellectual property that may result from a grant. Notwithstanding this commercialization disclaimer, the Simons Foundation has the right to release a summary of findings of the research within a reasonable period of time, typically within six (6) months of the date of the grant expiration.

Some Current Funding Opportunities

Funding Opportunity

2021 Human Cognitive and Behavioral Science – Request for Applications

To produce foundational knowledge about neurobehavioral differences associated with ASD, which will directly inform the development of refinement of tools needed for translational efforts, such as biomarkers and outcome measures.

Quantitative and objective measures of human behavior

- Psychophysics
- Computer vision and machine learning
- Wearables and app-based data collection

Human neuroscience

- EEG, fMRI, MEG

Encourage use of SPARK and Simons Searchlight Research Match

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

SFARI Base

SFARI Base is a clearinghouse for autism and autism-related research data and biospecimens supported by the SFARI funding as well as an online portal for the submission of research recruitment requests

The use of SFARI cohorts(SSC, Simons Searchlight and AIC) data and biospecimens is limited to projects related to advancing the field of autism and related neurodevelopmental disorder research. SPARK phenotypic and genomic data does not have this limitation. Applications will be reviewed to ensure the use of data and biospecimens is in compliance with the Institutional Review Board (IRB)-approved consent language.

All investigators using SFARI resources must adhere to the guidelines outlined in our Acknowledgement Policy.

.

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Depositing Data to SFARI Base

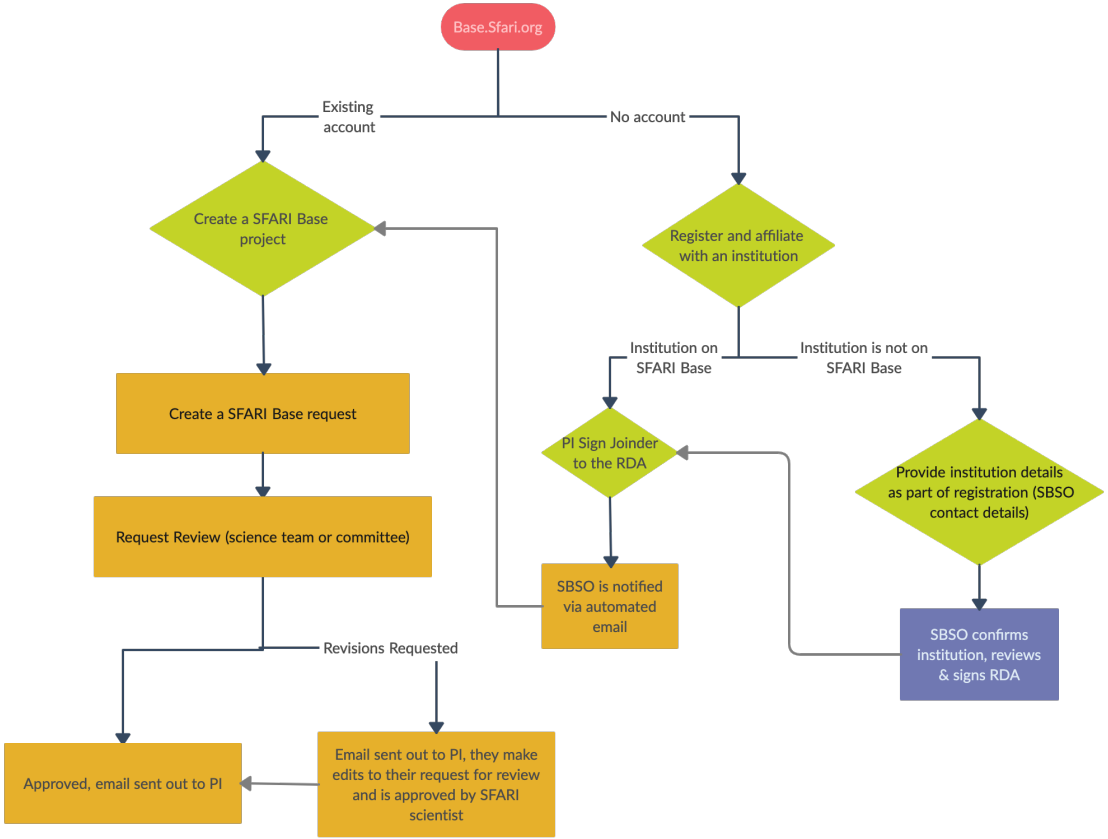
Information for submission:

1. Project Title/Abstract/PI/Data manager
2. Data Manifest: List of all files, Each file should be accompanied by the Md5 checksum.
3. Samples IDs and all types of files that will be available per sample.
4. Size of the dataset.
5. Description of the methods/analysis used for generation of the data/analysis files including versions of the software and parameters applied. Github link.
6. Embargo information
7. Manuscripts other references.

SFARIBase ID will be issued to the deposited data set and it can be used as a reference in further publications.

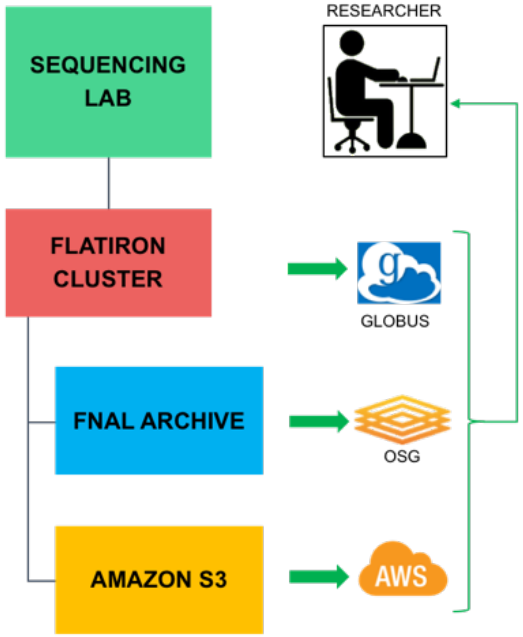
SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Requests Workflow



SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Genomic Data Sharing



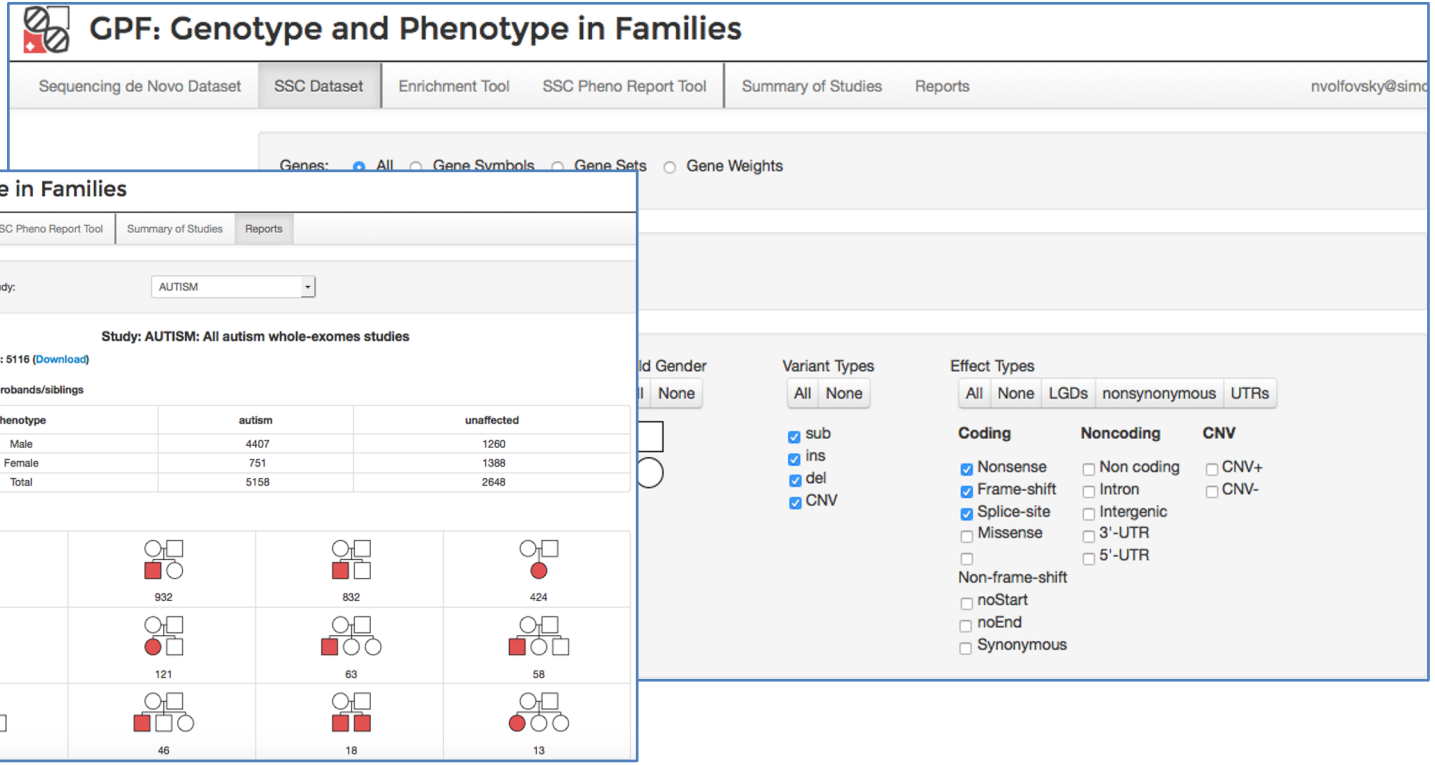
SFARI Base Collections
- Principal Investigator submits request for sets of genetic data (BAM, CRAM, SV, VCF, docs).

SFARI Base Approval
- Request goes through approval chain to confirm objective for research, adherence to HIPAA, key persons in lab.

Access Granted
- Access granted according to preferred method, chosen from Globus Online, AWS S3, or FermiLab

Access Confirm
- Check in with lab manager to confirm access to datasets, and address technical issues with access.

GPF: gpf.sfari.org



Ivan Iossifov Ph.D.
CSHL & NYGC

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

gnomAD-like SFARI Browser: genomes.sfari.org

CHD8 chromodomain helicase DNA binding protein 8

Ensembl gene ID	ENSG00000100888
Ensembl transcript ID	ENST00000399982 (canonical)
UCSC Browser	14:21385195-21456127
GeneCards	CHD8
OMIM	610528

Gene Constraint

Category	Exp. SNVs	Obs. SNVs	Constraint metrics
Synonymous	502.4	473	Z = 1.03 o/e = 0.94 (0.87 - 1.02)
Missense	1441.3	806	Z = 5.95 o/e = 0.56 (0.53 - 0.59)
LoF	129	5	pLI = 1 o/e = 0.04 (0.02 - 0.08)

Include: CDS UTR Non-coding transcript

Show transcripts

Bins All variant

ClinVar pathogenic and likely pathogenic variants (57)

SFARI Variants (2482)

Viewing in table

21,385,542 21,391,942 21,394,060 21,397,952 21,401,364 21,403,653 21,408,618 21,415,717 21,429,314 21,437,

LoF only | Missense only | Synonymous only | Other only

SNVs Indels

† denotes a consequence that is for a non-canonical transcript

Variant ID	Source	Cohort	Consequence	Annotation	Flags	Affected Allele Count	Total Allele Count
14-21393671-CAT-C	E	SPARK	p.Tyr2041Ter	frameshift	De Novo	1	1
14-21393692-G-A	E	SPARK	p.Arg2035Ter	stop gained	De Novo	1	1
14-21400060-G-A	E	SPARK	p.Arg1580Trp	missense	De Novo	1	2
14-21402335-C-T	E	SPARK	c.3882+1G>A	splice donor	De Novo	1	1
14-21405208-C-T	E	SPARK	c.3307+1G>A	splice donor	De Novo	1	1
14-21405800-A-T	E	SPARK	p.Leu991His	missense	De Novo	1	1
14-21406908-C-T	E	SPARK	p.Arg952Gln	missense	De Novo	1	1

Monkol Lek Ph.D.
Yale University

SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Statistics of SFARI Data Requests

SFARIbase

Monthly Report

Feb 2021

33 New users	106 Returning users	7 New institutions	4 Executed institutions	18 Manuscripts published
607 Active users		232 Active institutions		234 Active projects

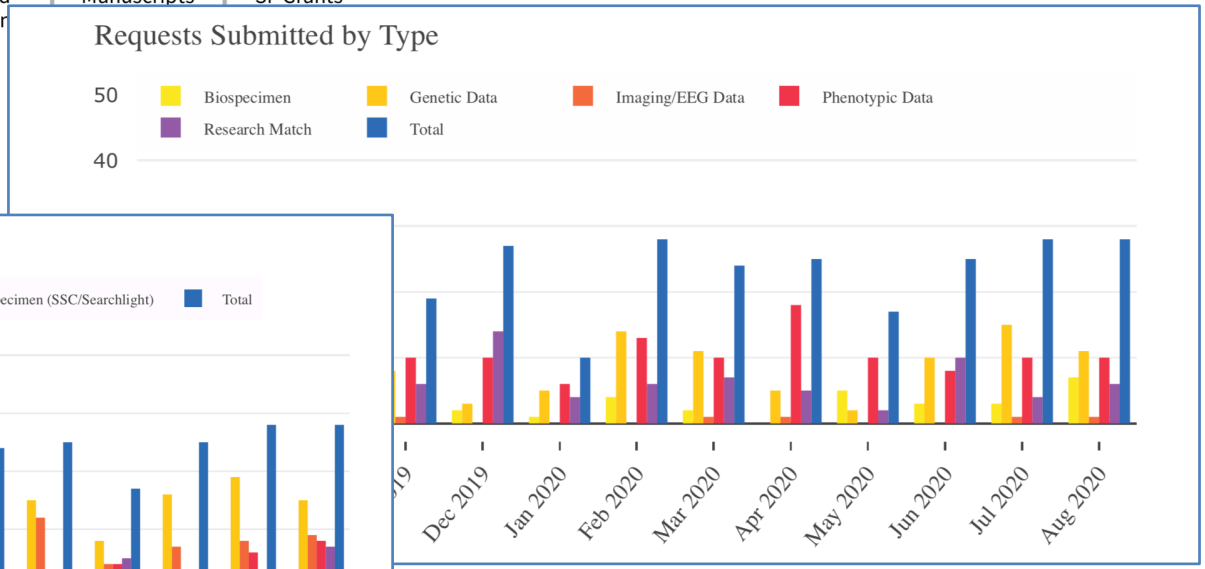
SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

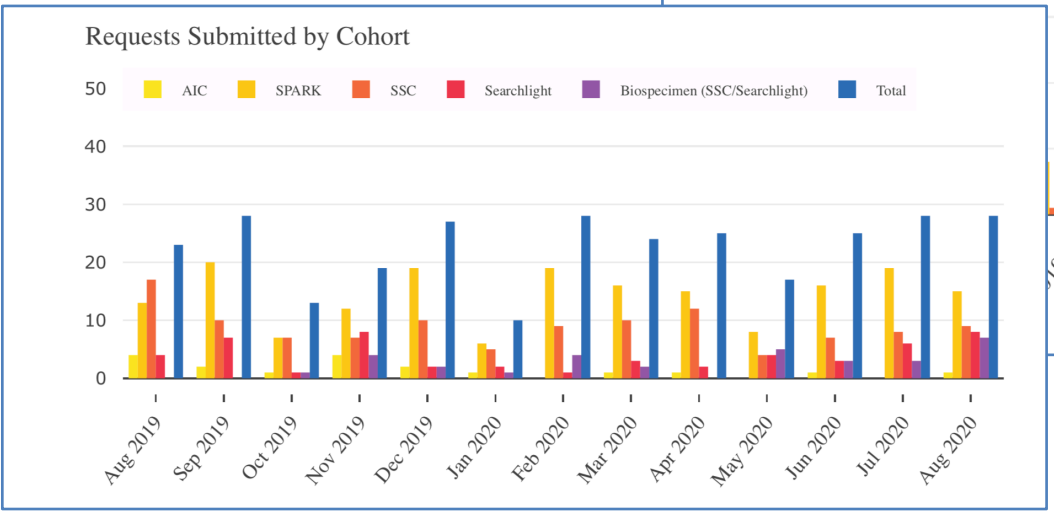
SFARIbase

Monthly Report

Aug 2020

33	108	5	3	5	0
New users	Returning users	New institutions	Executed institution	Manuscripts	SF Grants





SFARI SIMONS FOUNDATION AUTISM RESEARCH INITIATIVE

Conclusions

Functional Data Sharing Policy is important for facilitation of the reproducible biomedical research.

Major stakeholders - Researchers, Funders, Scientific Grant Managers, Informatics teams - should be included in the transparent communications and discussions