lald@CCF.org
@LalDennis

SLC6A1 – Estimated incidence & Variant interpretation

December 05, 2019

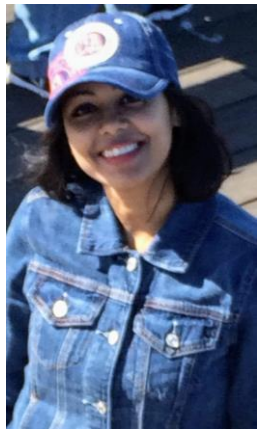
Cleveland Clinic

Dennis Lal
Genomic Medicine Institute, Lerner Research Institute ***Cleveland Clinic, Cleveland, Ohio, USA***
Epilepsy Center, Neurological Institute, ***Cleveland Clinic, Cleveland, Ohio, USA***
Stanley Center for Psychiatric Research, ***Broad Institute of Harvard and M.I.T, Cambridge, USA***
Analytic and Translational Genetic Unit, *Massachusetts General Hospital,* ***Harvard University, Boston, US***
Cologne Center for Genomics, ***University of Cologne, Cologne, Germany***

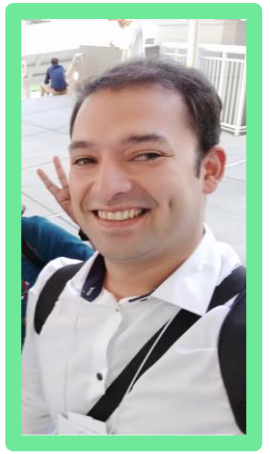
- All information/data (in more detail) are also included in the manuscript Amber sent to everyone

Team members who generated results which will be presented today

Javier Lopez
PhD student



Dr. Sumaiya Iqbal
Scientist



Dr. Eduardo Perez-Palma
Scientist

A lot of additional work for *SLC6A1, e.g. drafted the NORD report*

Estimating the birth incidence of de novo variant associated neurodevelopmental disorders

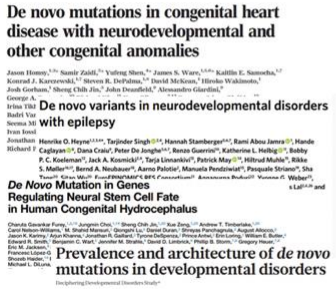
- Hundreds of monogenic neurodevelopmental disorders (NDDs) have been identified.
- Epidemiological gene-disorder specific incidence rates have been reported for only a few genes.
- **Goal: Predict incidence for all sporadic *de novo* variant associated NDDs using a genetic data-informed statistical model.**

Estimating the birth incidence of de novo variant associated neurodevelopmental disorders

Methods

Selecting genes of interest

101 *de novo* variant associated NDD genes identified from the literature


De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies

De novo variants in neurodevelopmental disorders with epilepsy

De Novo Mutation in Genes Regulating Neural Stem Cell Fate in Human Congenital Hydrocephalus

Prevalence and architecture of *de novo* mutations in developmental disorders

3,106 genes intolerant to variation identified using gnomAD constraint metrics

gnomAD

genome aggregation database

Estimation of incidence

Adjusted mutational model*

- Mutation rate
- Gene size
- Gene variant tolerance
- Pathogenic CNV rate

*Samocha et al., 2014

Validation of incidence estimates

1) Correlation of incidence estimates with reported pathogenic patient variants

ClinVar

HGMD®

2) Corroboration with reported estimates identified through systematic literature review

Lopez at al., (in review)

Estimated incidence matches reported estimates

2) Comparison with reported incidence estimates

Gene	Disease	Incidence in 100,000 births		Significant Difference
		Calculated	Reported	
SCN1A	Dravet Syndrome (OMIM: 607208)	6.69-7.62	4.78	no
			4.54	no
			5.90	no
			4.10	no
SLC2A1	GLUT1 Deficiency Syndrome (OMIM: 606777)	1.65-2.22	1.20	no
			2.95	no
TBX5	Holt-Oram Syndrome (OMIM: 142900)	0.39-0.45	0.56	no
STXBP1	STXBP1 Encephalopathy (OMIM: 612164)	3.30-3.81	1.09	yes
			0.82	yes
SALL1	Townes-Brocks Syndrome (OMIM: 107480)	0.30-0.36	0.42	no
KCNQ2	KCNQ2 Encephalopathy (OMIM: 613720)	2.93-3.59	1.18	no
			1.23	no
CDKL5	CDKL5 Deficiency Disorder (OMIM: 300672)	1.81-2.49	1.77	no
			0.96	no

Lopez at al., (in review)

Estimating the birth incidence of de novo variant associated neurodevelopmental disorders

2) Comparison with reported incidence estimates

Gene	Disease	Incidence in 100,000 births		Significant
		Calculated	Reported	Difference
SCN1A	Dravet Syndrome (OMIM: 607208)	6.69-7.62	4.78	no
			4.54	no
			5.90	no
			4.10	no
SLC2A1	GLUT1 Deficiency Syndrome (OMIM: 606777)	1.65-2.22	1.20	no
			2.95	no
TBX5	Holt-Oram Syndrome (OMIM: 142900)	0.39-0.45	0.56	no
STXBP1	STXBP1 Encephalopathy (OMIM: 612164)	3.30-3.81	1.09	yes
			0.82	yes
SALL1	Townes-Brocks Syndrome (OMIM: 107480)	0.30-0.36	0.42	no
KCNQ2	KCNQ2 Encephalopathy (OMIM: 613720)	2.93-3.59	1.18	no
			1.23	no
CDKL5	CDKL5 Deficiency Disorder (OMIM: 300672)	1.81-2.49	1.77	no
			0.96	no

SLC6A1

2.65 (90%CI: 2.38-2.86) per 100,000 births

Lopez at al., (in review)

Estimating the birth incidence of de novo variant associated neurodevelopmental disorders

2) Comparison with reported incidence estimates

Gene	Disease	Incidence in 100,000 births		Significant
		Calculated	Reported	Difference
SCN1A	Dravet Syndrome (OMIM: 607208)	6.69-7.62	4.78	no
			4.54	no
			5.90	no
			4.10	no
SLC2A1	GLUT1 Deficiency Syndrome (OMIM: 606777)	1.65-2.22	1.20	no
			2.95	no
TBX5	Holt-Oram Syndrome (OMIM: 142900)	0.39-0.45	0.56	no
STXBP1	STXBP1 Encephalopathy (OMIM: 612164)	3.30-3.81	1.09	yes
			0.82	yes
SALL1	Townes-Brocks Syndrome (OMIM: 107480)	0.30-0.36	0.42	no
KCNQ2	KCNQ2 Encephalopathy (OMIM: 613720)	2.93-3.59	1.18	no
			1.23	no
CDKL5	CDKL5 Deficiency Disorder (OMIM: 300672)	1.81-2.49	1.77	no
			0.96	no

SLC6A1

2.65 (90%CI: 2.38-2.86) per 100,000 births

Number of births for the United States in 2018 was 3,788,235 (cdc.gov)
->
98 new patients every year in the US alone

Lopez at al., (in review)

In the most comprehensive sequencing screens *SLC6A1* is among the most frequent mutated genes

- Three large genomic screens with 8565, 9170 and 9769 **epilepsy** patients have been reported. In one study Pathogenic variants of *SLC6A1* are consistently identified in these large datasets, falling in the **top twenty** in one study and **top ten in the other two studies**.

- Further, in the largest **autism** sequencing study to date (N=11,986 patients), *SLC6A1* was also among the **top ten genes**, with the most significant enrichment in autism patients compared to 23,598 controls.

Lindy et al 2018, Epi25 Collaborative 2019, Truty et al 2019, Satterstrom et al 2019

Patients captured by routine clinical sequencing – challenges and opportunities

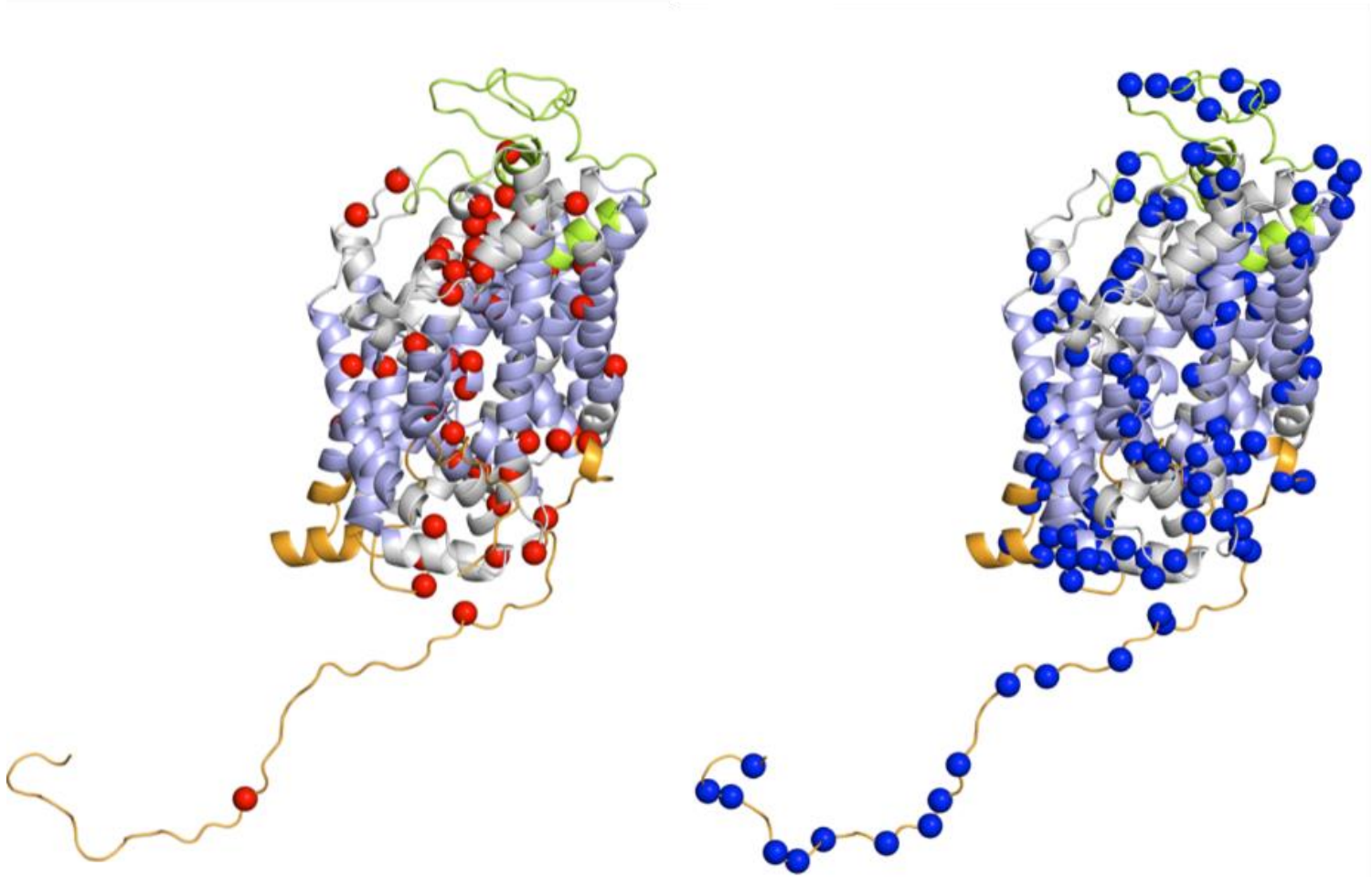
- http://simple-clinvar.broadinstitute.org

Perez-Palma 2019., Nucleic Acid Research

Variants on GAT-1 the *SLC6A1* protein

- No experimentally solved structure
- Structure computationally determined using an existing solved structure of 41% sequence identity
  - Using RaptorX software
  - P-value of the predicted structured = $2.79e^{-14}$ ($< e^{-14}$ is a good indicator)
- Domains: 12 Helical domains, 1 Extracellular and 2 cytoplasmic domains, and rest are the hinge regions in between helical segments

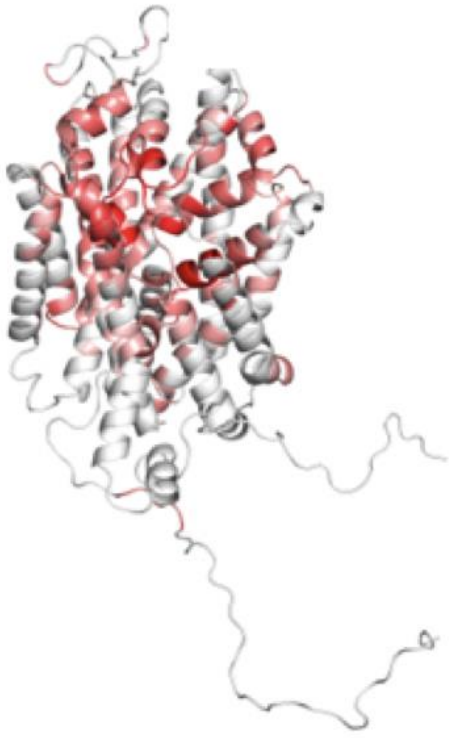
Latest data set: Genetic variants on *SLC6A1* and the GAT1 protein



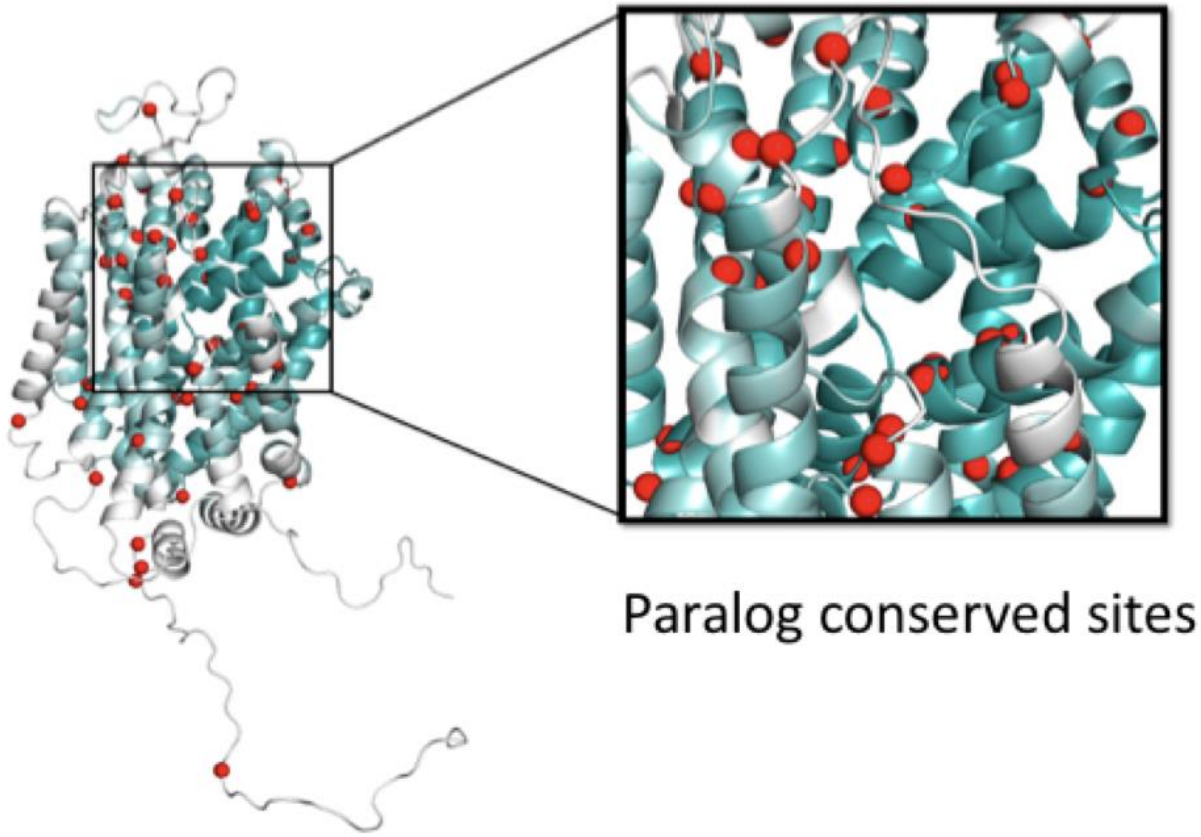
gnomAD variant positions in SLC6A1 (control)

Epilepsy patient variants positions in SLC6A1

Patient variants cluster in a hot spot, which is conserved across related transporters

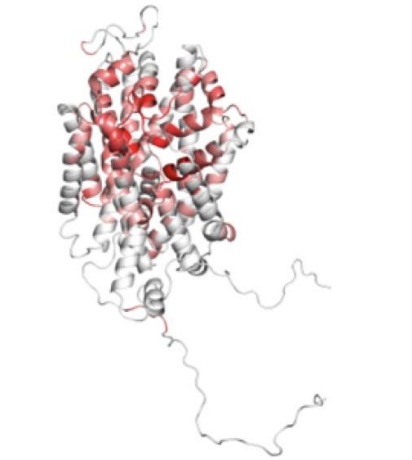


Hot-spots: enriched with pathogenic and depleted of control

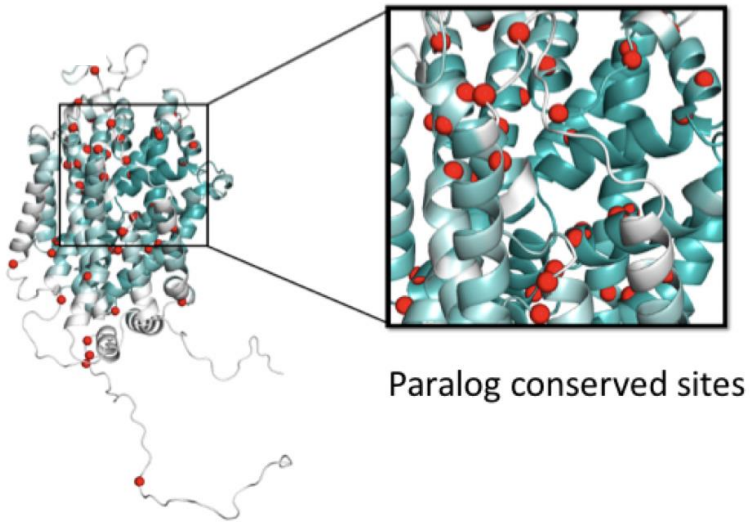


Paralog conserved sites

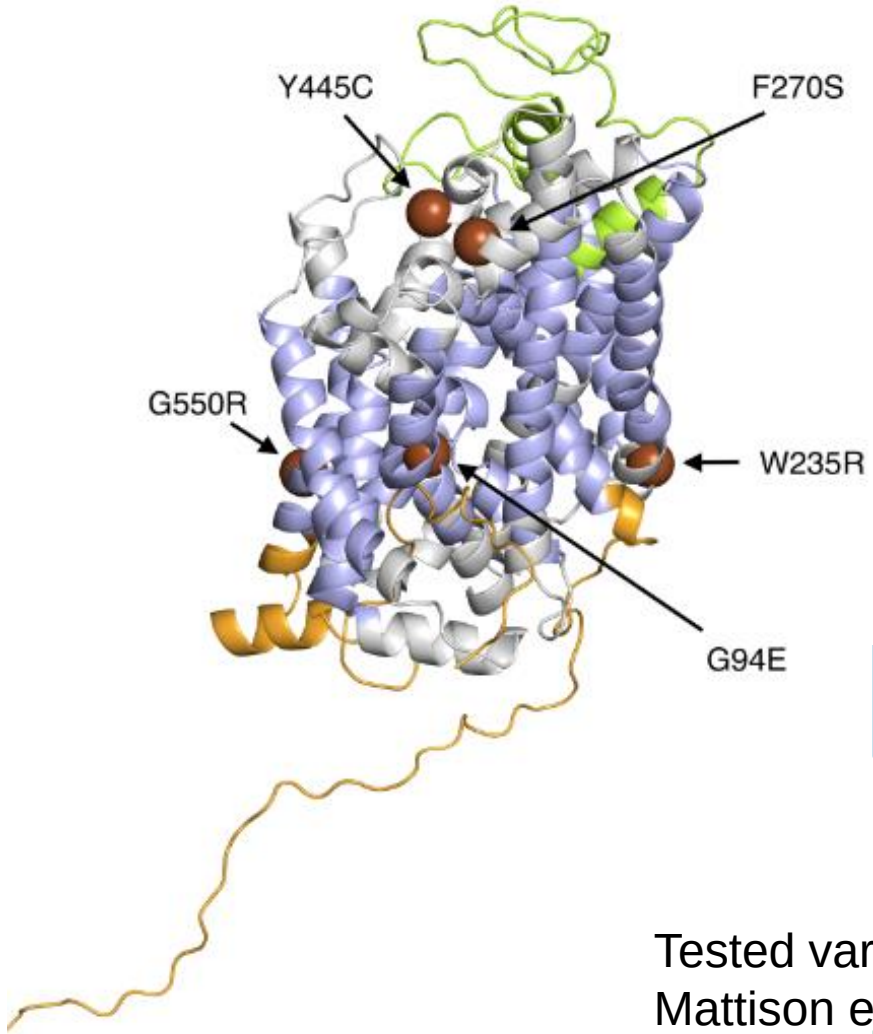
All functionally tested variants where loss-of-function missense variants and are located across the hotspot



Hot-spots: enriched with pathogenic and depleted of control



Paralog conserved sites



Tested variants from Mattison et al. 2018

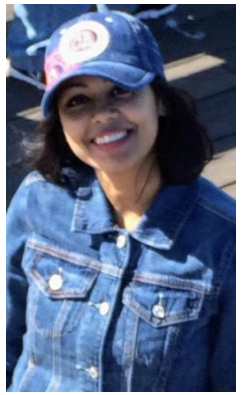
Summary

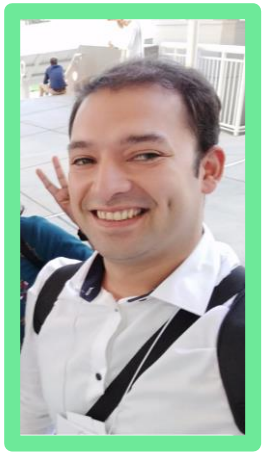
- *SLC6A1* deficiency disorder is a frequent rare neurodevelopmental disorder
- The number of patients are increasing
- Variant analyses support that GAT-1 loss-of-function is the patho-mechanism also for missense variants

Acknowledgement

SLC6A1 Connect
SLC6A1 Research & Support


Javier Lopez
PhD student


Sumaiya Iqbal
Postdoc


Eduardo Perez-Palma
Postdoc

lald@CCF.org
dlal@Broadinstitute.org

@LalDennis

Cleveland Clinic

STANLEY CENTER
FOR PSYCHIATRIC RESEARCH
AT BROAD INSTITUTE

University of Cologne