Genetic mates
SNPedia write-ups
Final
Next-gen sequencing
Mike Snyder
QA: new technology and the future

Genetic Mates

Jonathan Mortensen/Francisco Gimenez

Will need class to upload data Tuesday night

- Analyze data

Class presentation on Thursday

SNPedia write-ups

- Due Monday May 27, 2013 midnight
- Upload to SNPedia.com
  - instructions at : http://www.stanford.edu/class/gene210/web/html/snpedia.html
- Review examples from last year:
  - Go to SNPedia.com
  - Sign in as gene210/stuart
  - Click “watchlist” at the top
  - Click “View and edit watchlist” at the top
  - Read a couple of SNPs from last year.
    - e.g. Rs7294919

SNPedia write-ups

Introduce the SNP and the phenotype

Summarize the data

- How many cases/controls?
- What was the p-value for association? Is this significant following multiple hypothesis testing?
- What was the odds ratio? Relative risk? Hazard ratio? % variance explained?
- Has the result been validated in a second study?
- Which is the risk allele?

Optional: Is the causal mutation known?

Is the causal gene known?

2012 SNP write-up example

SNPedia

Page | Discussion | Read | Edit with form | Edit | View history | Search | Go | Search

Rs7294919

rs7294919, showed a particularly strong link to a reduced hippocampus volume, suggesting that this gene is very important to hippocampus development or health. No associations for brain volume, but they did discover that intracranial volume was significantly associated with two loci: rs4273712, a known height locus on chromosome 6q22, and rs9915547, tagging the inversion on chromosome 17q21. The SNP is located between two genes, HRK [1] and FBXW8 [2], but evidence suggests that it influences the expression level of a gene 3' to FBXW8, TESC [3]. Each copy of the T allele was associated with a 107.8 mm^3 decrease in hippocampal volume [4]. In European populations, the effect allele (T) is found at frequency of 0.898 [5]. The minor allele (C) is found at a frequency of 0.102.

Background [edit]

The hippocampus is a critical brain structure involved in learning and memory. In particular, it is associated with the ability to form long-term memories of facts and events [6]. This is in contrast to short-term and working memory, which have been shown to be independent of the hippocampus [7]. Hippocampal size decreases with age and is diminished in several disorders including Alzheimer's Disease [8], Major Depressive Disorder [9], Post-traumatic Stress Disorder [10], and Schizophrenia [11]. Moreover, the size of the structure is heritable, with estimates of heritability ranging from 40-70% [12][13].

Studies [edit]

Two major studies were conducted which found an association between rs7294919 and hippocampal volume. They were published in the April 15th, 2012 issue of *Nature Genetics* [14][15]. Though the P-values and regression slopes differ, both studies, together comprising tens of thousands of individuals, agree that the T allele is negatively associated with hippocampal volume.

1. The first study uses the CHARGE consortium (Cohorts for Heart and Aging Research in Genomic Epidemiology) as its discovery cohort [16]. CHARGE comprises 8 sub-cohorts representing 9,232 people with an average age of 67.1

Orientation	plus
is a	snp
is	mentioned by
dbSNP	rs7294919
PheGenI	rs7294919
nextbio	rs7294919
hapmap	rs7294919
1000 genomes	rs7294919
hgdp	rs7294919
ensembl	rs7294919
gopubmed	rs7294919
geneview	rs7294919
scholar	rs7294919
google	rs7294919
pharmgkb	rs7294919
gwascentral	rs7294919
openSNP	rs7294919
23andMe	rs7294919
23andMe all	rs7294919
SNP Nexus	rs7294919

Navigation

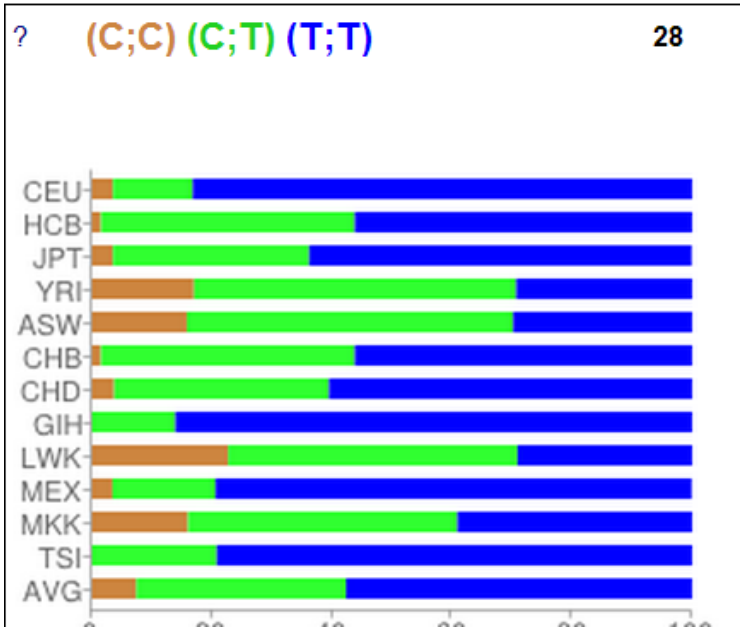
- SNPedia
- Promethease
- FAQ
- Blog
- Recent changes
- Random page

Toolbox

- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link
- Browse properties

Chromosome	12
Orientation	plus
GMAF	0.212
Position	117327592
Reference	GRCh37.p5 37.3/135
Max Magnitude	3

Geno	Mag	Summary
(C;C)	3	Enhanced hippocampal volume
(C;T)	3	Moderately enhanced hippocampal volume
(T;T)	0	Average hippocampal volume



SNPedia automatically puts in the chromosome, position and allele frequencies for different races.

1. The first study uses the CHARGE consortium (Cohorts for Heart and Aging Research in Genomic Epidemiology) as its discovery cohort [16]. CHARGE comprises 8 sub-cohorts representing 9,232 people with an average age of 67.1 years. Most of the sub-cohorts were from Europe and hippocampal volume was determined using MRI. Second stage verification/replication was performed on two cohorts comprising 2,318 subjects. This study found that rs7294919 was associated with hippocampal volume with a meta-P value (combining the discovery and replication cohorts) of 2.9×10^{-11} and a regression slope (β) of -107.8 mm^3 hippocampal volume.

   The authors provide some speculation about the SNP's mechanism of action by reviewing the neighboring two genes. HRK is involved in apoptosis of neurons and is thought to play a role in ischemia-induced apoptosis. FBXW8 targets an E3 ubiquitin ligase to protein aggregates and has been shown to be involved in hippocampal neuron dendrite growth. Potential pitfalls of this study include the older age of the participants, the mixture of both computerized and manual hippocampus tracing in MRIs, and the predominantly European ethnic makeup of the cohorts.

Size of cohort?

P value?

Size of effect?

Replication in second set?

Final

Thursday May 23, 2013

Take home

No help

10% extra credit

Due Monday May 27 2013 at midnight

Submit to gene210.stanford@gmail.com

Discovery of new Mendelian disorders by Next Gen sequencing

Table 2: Number of genes common in sequential analysis of phenotypically ranked individuals

Sequential analysis	1	+2	+3	+4	+5	+6	+7	+8	+9	+10
NS/SS/I	5,282	3,850	3,250	2,354	2,028	1,899	1,772	1,686	1,600	1,459
Not in dbSNP129 or 1000 Genomes	687	214	145	84	63	54	42	40	39	34
Not in control exomes	675	134	50	26	13	13	8	5	4	2
Not in either	467	89	34	18	9	8	4	4	3	1
Is loss-of-function (nonsense/frameshift indel)	25	1	1	1	0	0	0	0	0	0

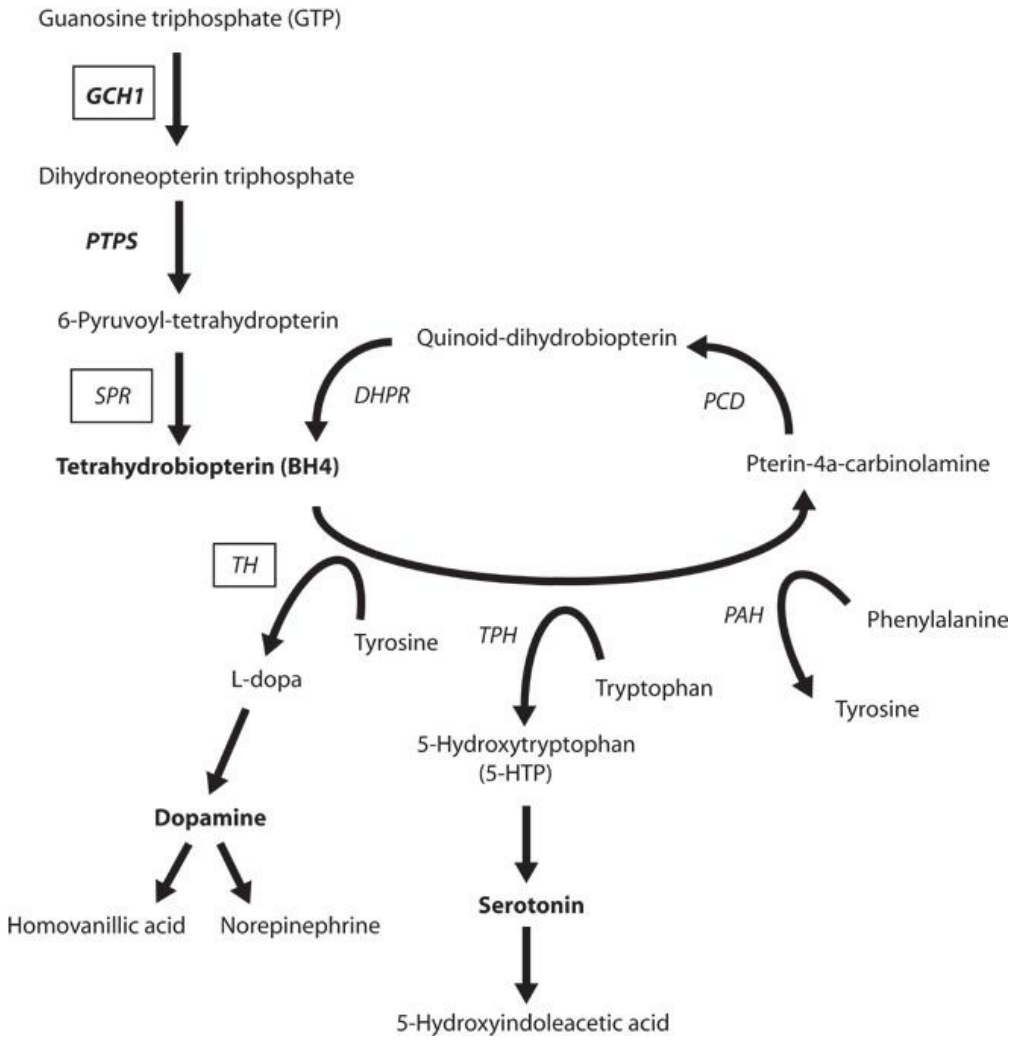
Variants were filtered as in **Table 1**. Exomes were added sequentially to the analysis by ranked phenotype; for example, column "+3" shows the number of genes at the intersection of the three top ranked cases (**Supplementary Fig. 1**). The gene with at least one NS/SS/I in all individuals is *MUC16*, which is very likely to be a false positive due to its extreme length (14,507 amino acids).

Dopa (3,4-dihydroxyphenylalanine)–responsive dystonia (DRD).

- hereditary dystonia with marked diurnal variation.
- begins in childhood.
- associated with mutations in genes encoding guanosine 5’-triphosphate (GTP), cyclohydrolase (*GCH1*), tyrosine hydroxylase (*TH*), and sepiapterin reductase (*SPR*).

Bainbridge, M.N. et al. Whole-genome sequencing for optimized patient management. Sci. Transl. Med. 3, 87re3 (2011).

Fig. 1 Metabolic pathways of neurotransmitter production.



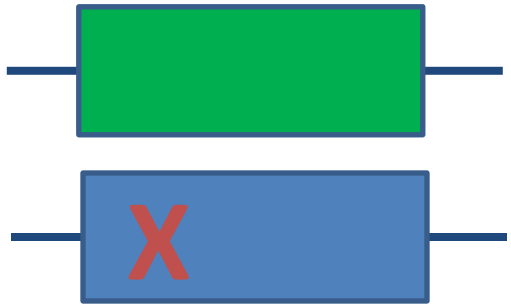
Bainbridge M N et al. Sci Transl Med 2011;3:87re3-87re3

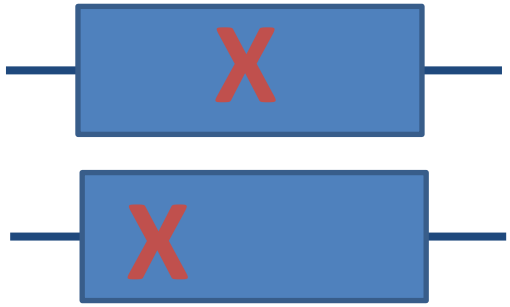
Clinical history

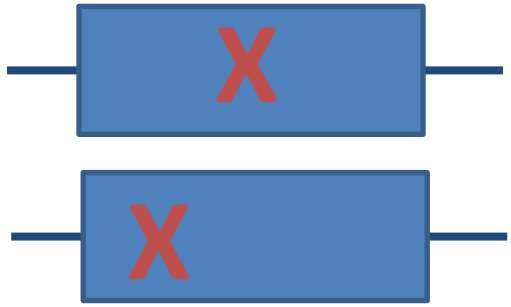
- fraternal twin pair.
- diagnosed with DRD at age 5.
- L-dopa was found to alleviate the clinical symptoms of dystonia in one twin.
- no identified deleterious variants in the *TH* or *GCH1* genes.
- Sequencing of the *SPR* gene was not available.
- high-throughput sequencing used to interrogate the whole genomes of the male and female twins.

Mother

Father







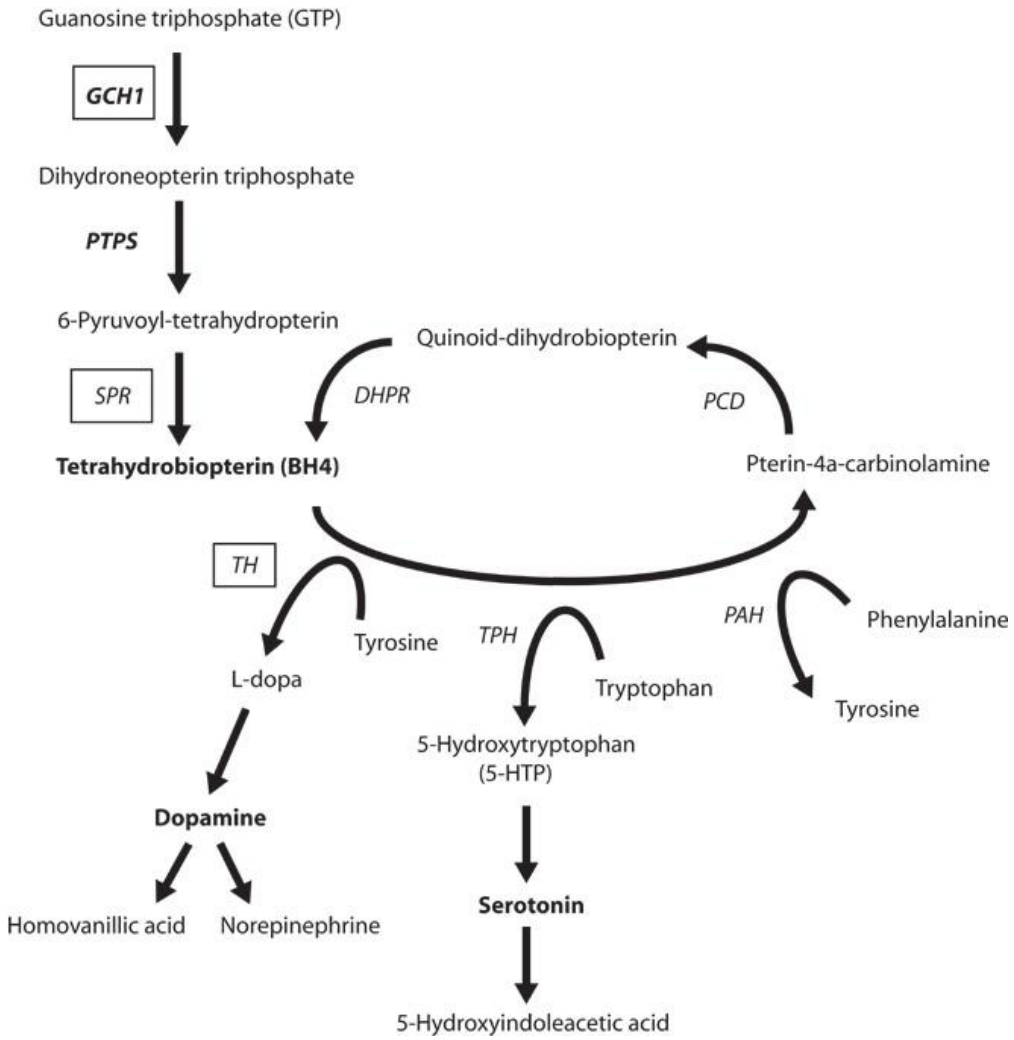
Affected

Affected

Results

- Three genes found with rare compound heterozygous mutations.
  - *ZNF544* encodes a computationally predicted zinc finger protein with no known function or targets
  - *C2orf16.*
  - *SPR* encoding sepiapterin reductase.
- *SPR* mutations seen in two previous families with DRD.

Fig. 1 Metabolic pathways of neurotransmitter production.

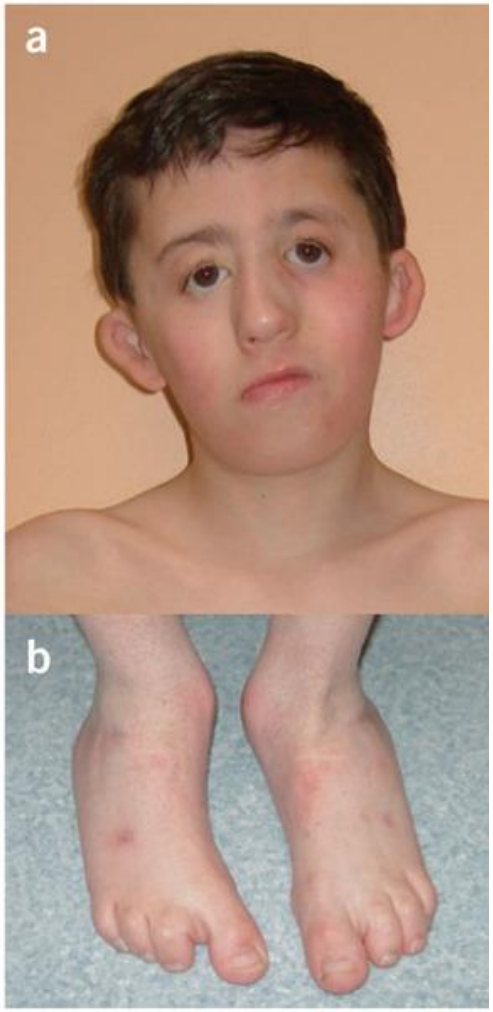


Bainbridge M N et al. Sci Transl Med 2011;3:87re3-87re3

Results

- recommended treatment of DRD caused by SPR mutations is with both the dopamine precursor L-dopa, which the twins were already prescribed, and the serotonin precursor 5-hydroxytryptophan (5-HTP), which the twins were not receiving.
- The male DRD patient reported improved focus in school, as well as improved coordination in athletics. Further, the male showed reduced drooling and hand tremor, and objective evidence for the latter was provided by serial handwriting samples.
- The female twin reported reduced frequency of laryngeal spasms, improved sleep and focus, and improved tolerance for exercise and was able to resume participation in sports after a 14-month absence. In the female DRD patient, there were also reduced choreiform movements of the tongue by objective physical examination.

Miller Syndrome

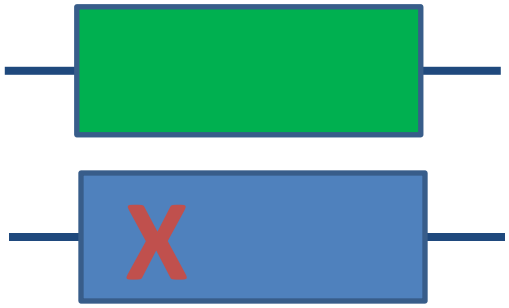


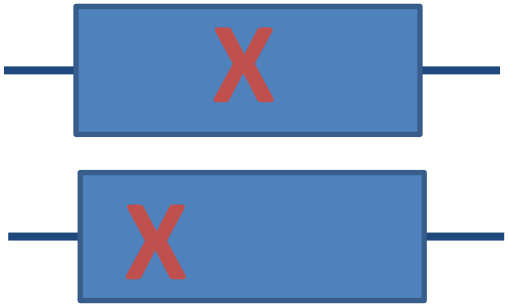
- Cupped ears
- coloboma of the lower eyelids,
- prominent nose,
- micrognathia
- absence of the fifth digits of the feet

Ng, S.B. et al. Exome sequencing identifies the cause of a Mendelian disorder. Nat. Genet. 42, 30–35 (2010).

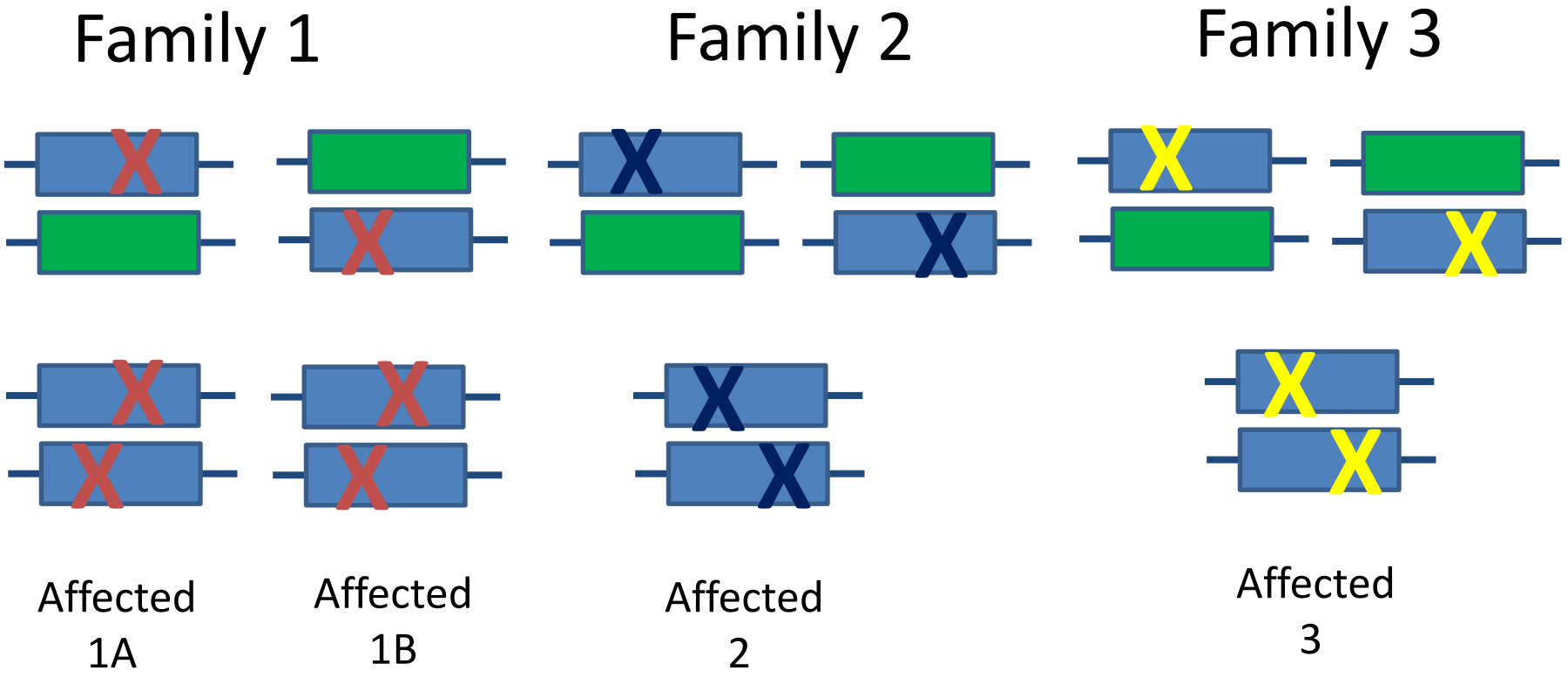
Mother

Father



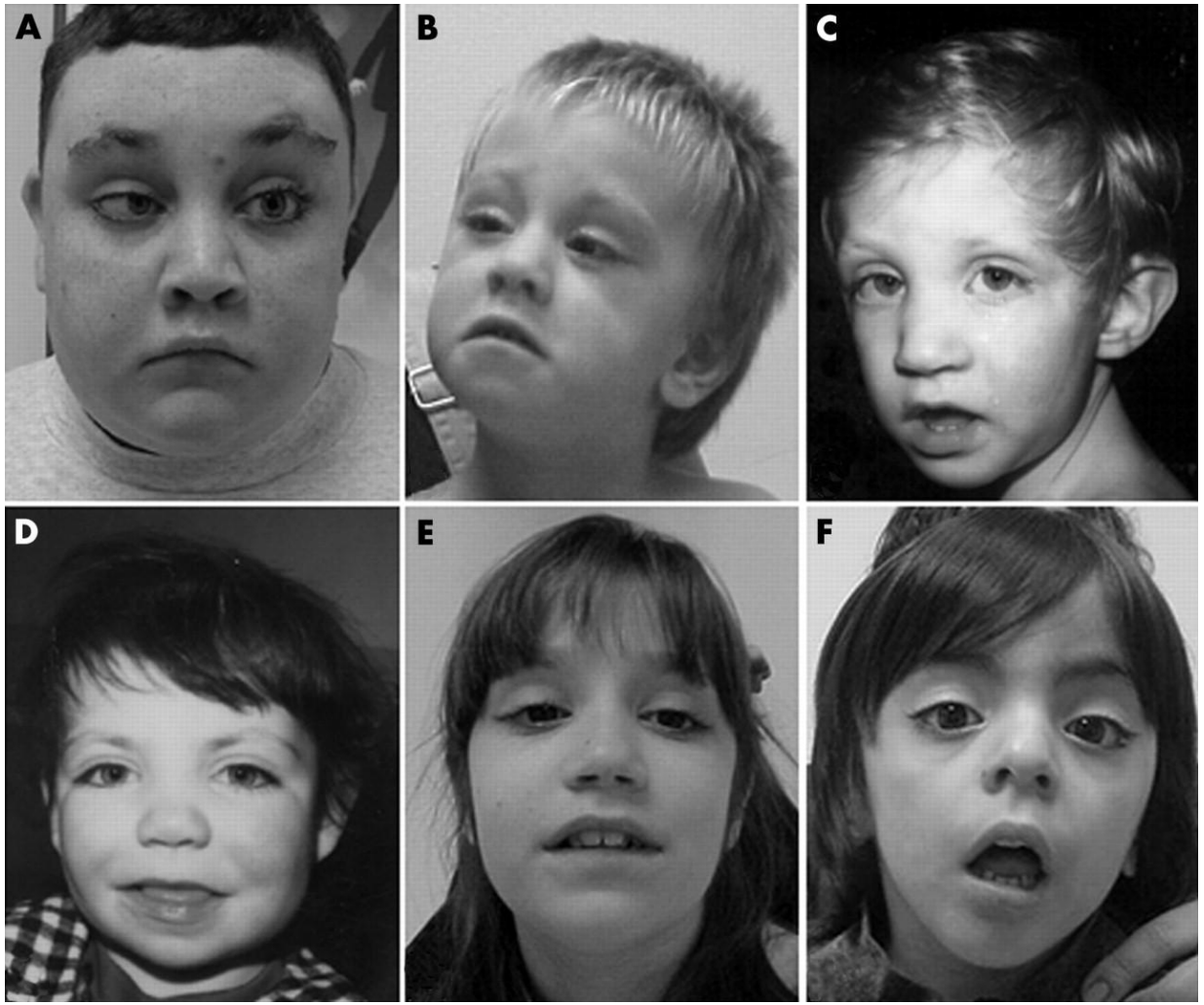


Affected


Family 1
Family 2
Family 3
Affected 1A
Affected 1B
Affected 2
Affected 3

- Non-synonymous and Rare (~450/person)
- Two variants in the same gene (30/person)
- Same gene in 1A and 1B = 9 genes
- Same gene in 1A, 1B, 2 and 3 = *DHODH*
- dihydroorotate dehydrogenase
- Validate by sequencing the gene in 3 more cases.

Kabuki Syndrome



Ng, S.B. *et al*. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* **42**, 790–793 (2010).

Kabuki syndrome

- Kabuki syndrome is a rare, multiple malformation disorder characterized by a distinctive facial appearance, cardiac anomalies, skeletal abnormalities, immunological defects and mild to moderate mental retardation.
- Six cases of parent-child transmission.
- Autosomal dominant disorder.

Strategy

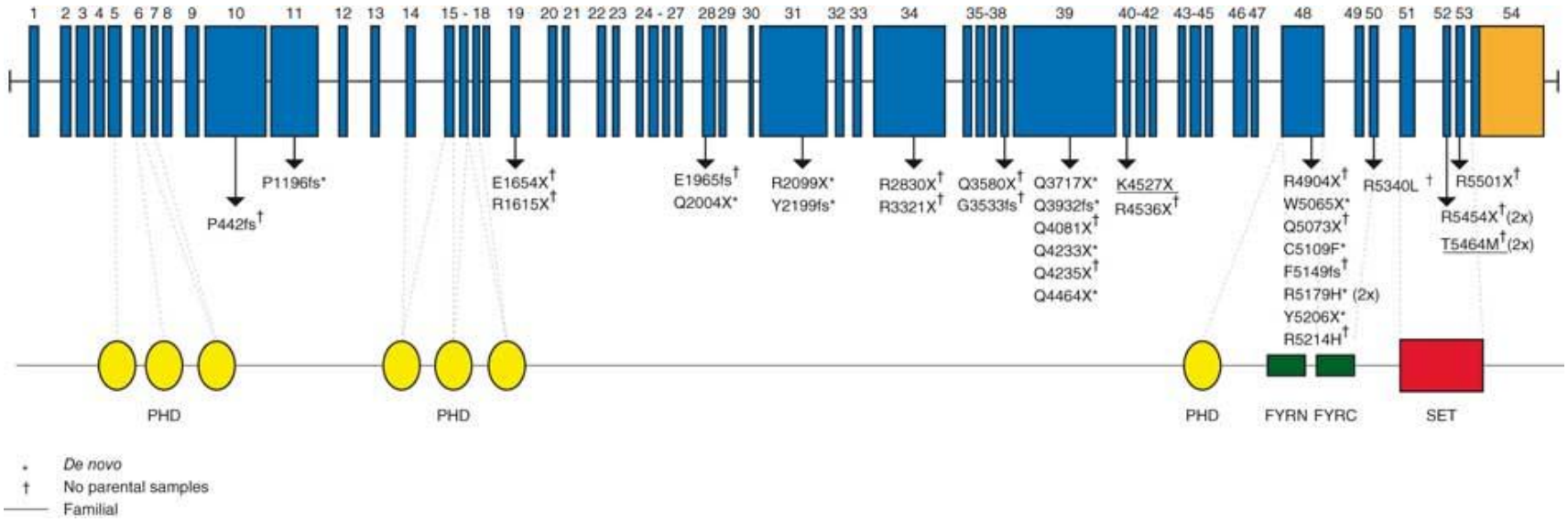
- Look for rare, loss-of-function protein coding changes in strongest patient
- See if find mutations in same gene in the second strongest patient, then third, then fourth, etc.
- After 4 patients, only one gene is common. MLL2
- Retrospectively, see that MLL2 has suggestive mutations in 5 more patients. Ultimately, 9/10 patients had mutations in MLL2.
- For validation, found 26/43 patients had mutations in MLL2. 0/190 controls had mutations.

Table 2: Number of genes common in sequential analysis of phenotypically ranked individuals

Sequential analysis	1	+2	+3	+4	+5	+6	+7	+8	+9	+10
NS/SS/I	5,282	3,850	3,250	2,354	2,028	1,899	1,772	1,686	1,600	1,459
Not in dbSNP129 or 1000 Genomes	687	214	145	84	63	54	42	40	39	34
Not in control exomes	675	134	50	26	13	13	8	5	4	2
Not in either	467	89	34	18	9	8	4	4	3	1
Is loss-of-function (nonsense/frameshift indel)	25	1	1	1	0	0	0	0	0	0

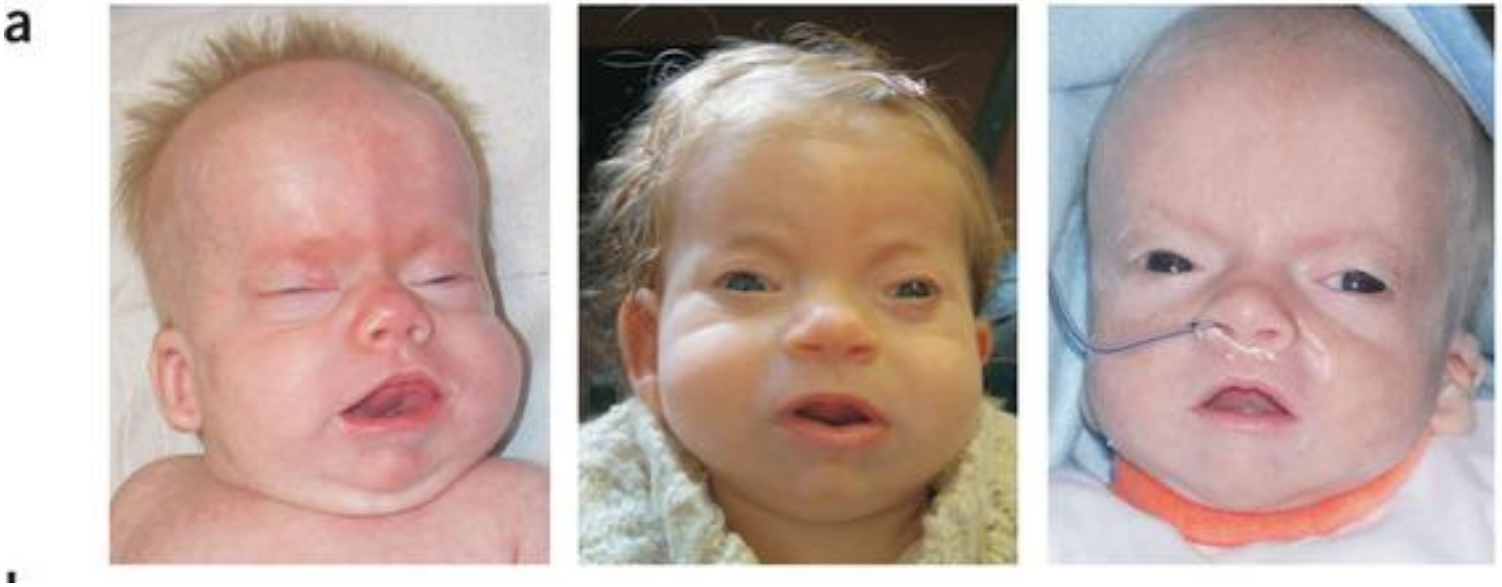
Variants were filtered as in **Table 1**. Exomes were added sequentially to the analysis by ranked phenotype; for example, column "+3" shows the number of genes at the intersection of the three top ranked cases (**Supplementary Fig. 1**). The gene with at least one NS/SS/I in all individuals is *MUC16*, which is very likely to be a false positive due to its extreme length (14,507 amino acids).

Genomic structure and allelic spectrum of *MLL2* mutations that cause Kabuki syndrome



The SET domain of MLL2 confers strong histone 3 lysine 4 methyltransferase activity and is important in the epigenetic control of active chromatin states.

Schinzel-Giedon syndrome



- Severe mental retardation, distinctive facial features, multiple congenital malformations (including skeletal abnormalities, genitourinary and renal malformations, and cardiac defects) and a higher-than-normal prevalence of tumors, notably neuroepithelial neoplasia.

Hoischen, A. et al. De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. Nat. Genet. 42, 483–485 (2010).

Mother

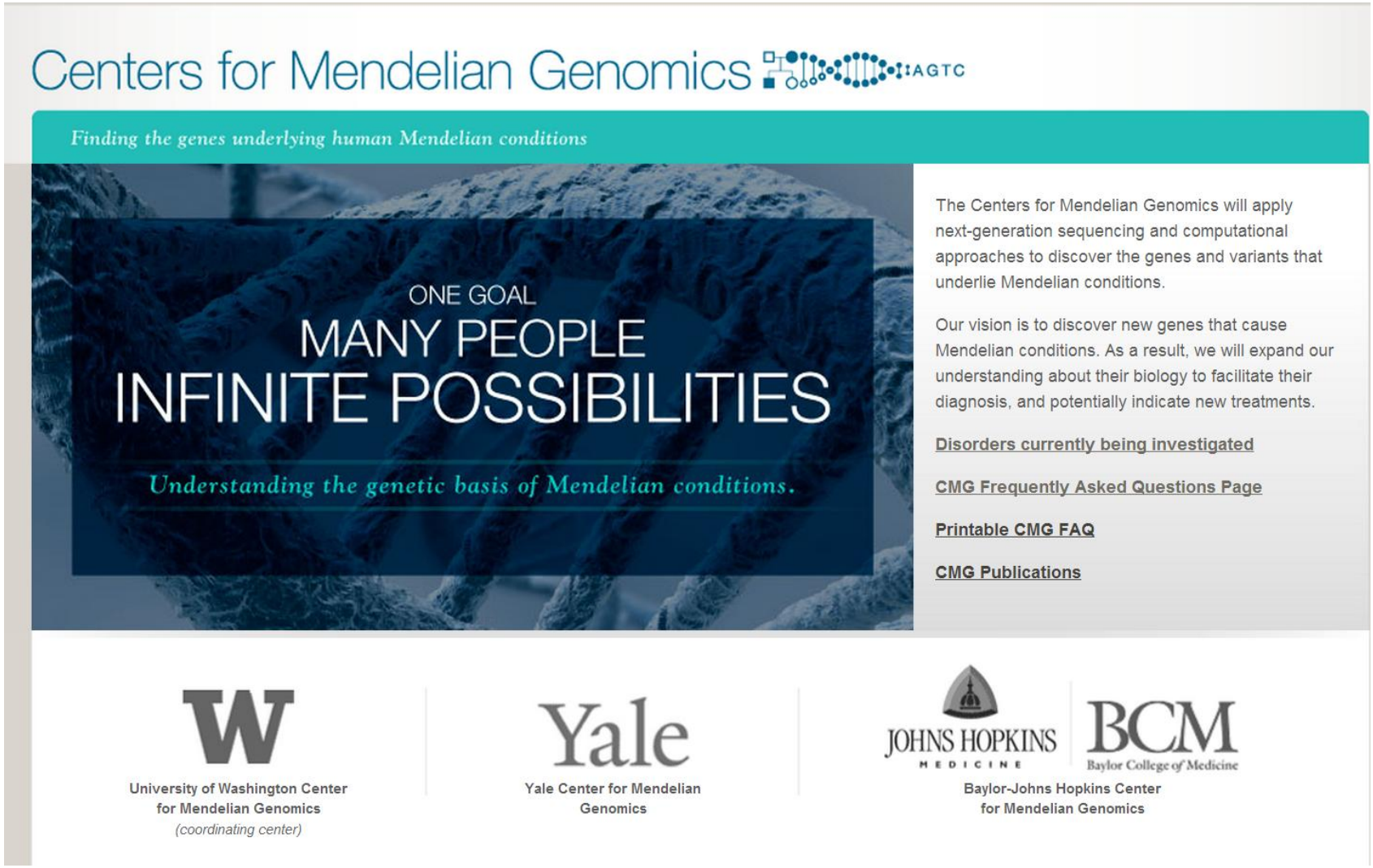
Father

Affected

Strategy

- Look for rare, loss-of-function protein coding changes
- Only 1 gene shared in four patients – SETB1.
- In all cases, mutations occurred de novo in patients and were absent from parents.
- Validated in 8/9 additional patients.
- Gain-of-function effect or dominant-negative effect

www.mendelian.org



www.mendelian.org

- Large scale public effort to identify Mendelian mutations
- UW, Yale/Hopkins/Baylor
- Doctors send in samples
- Centers sequence families, find mutations
- Joint publication

www.mendelian.org

	A	C
1	**Disorder**	**OMIM #**
2	Aarskog Syndrome, autosomal dominant	100050
3	Acromelic frontonasal dysostosis (AFND)	603671
4	Acute liver failure	
5	Adams-Oliver Syndrome (AOS)	100300
6	Agammaglobulinemia 1, autosomal recessive; AGM1	601495
7	Alopecia mental retardation	203650
8	Amyotrophic lateral sclerosis, juvenile, with dementia	205200
9	Aortic aneurysm and AV malformation	
10	Aortic aneurysm and hemangiomas and AV malformation	
11	Aortic aneurysm and hemangiomas and AV malformation	
12	Aortic aneurysm, autosomal dominant	
13	Aortic aneurysm, AVM	
14	Aortic aneurysm, familial abdominal, 1; AAA1	100070
15	Aortic aneurysm, familial thoracic 1; AAT1	607086
16	Aortic aneurysm, familial thoracic 2; AAT2	607087
17	Aortic valve disease	109730
18	Aplasia cutis congenita (ACC)	107600
19	Arrhythmogenic right ventricilar dysplasia, familial, 1; ARVD1	107970
20	Arthrogryposis, distal, type 1A; DA1A	108120
21	ARVD	610193
22	Ascending aortic dilation and branchial cysts	
23	Asphyxiation thoracic dystrophy	
24	Ataxia	
25	Atrial septal defect, autosomal dominant (ASD)	
26	Atrioventricular septal defect (AVSD)	606215
27	Autism	209850
28	Autism	
29	Autism, developmental delay, dysmorphic features	
30	Autoimmune Polyendocrine Syndrome, type I; APS1	240300
31	Bardet-Biedl Syndrome	209900

- Sequencing 285 Mendialian families
- Either exome or wgs