Whole Exome Sequencing Reveals the Genetic Basis of a Case of Idiopathic Hemolytic Anemia and Suggests Candidate Rare Variants for ADHD in a Utah Pedigree

Reid Robison,[1,10] Gholson J. Lyon,[1,2,8,10] Tao Jiang,[3,4,10] Richard Van Wijk,[6] Lin Yang,[3] Peng Zhang,[3] Renhua Wu,[3] Ying Liu,[3,4] Xu Yang,[3] Jinchuan Xing,[5] Barry Moore,[5] Joseph Glessner,[8] Fred Reimherr,[1] Wouter van Solinge,[6] Josephine Elia,[8] Mark Yandell,[5] William McMahon,[1] Evan Johnson,[9] Mark Clement,[9] Hakon Hakonarson,[8] Jun Wang,[3,4,7] Kai Wang,[8,11]

Introduction

ADHD is a common disorder affecting more than 1 in 20 children in the U.S., with as many as 50% remaining symptomatic into adulthood. Genetic factors are thought to play a large role in the etiology of the disorder, but studies thus far have been inconsistent, accounting for a small portion of the risk for ADHD. It is hypothesized that rare, family-specific genetic variants may account for the remaining missing heritability of ADHD, and advances in next-generation sequencing have made whole exome sequencing a potentially viable option for gene variant finding in complex disease pedigrees. Exome sequencing has identified the causes of nearly a dozen Mendelian disorders. Whether exome sequencing can be used to find risk variants for complex traits and has clinical utility as a diagnostic tool remain to be explored. Here we examined the utility of exome sequencing to identify candidate genes for attention deficit/hyperactivity disorder (ADHD) in a Utah pedigree in which ADHD clustered in an apparently Mendelian pattern.

Methods

Whole exome sequencing was conducted on a father and two sons with severe ADHD, combined hyperactive and inattentive subtype, all of whom improved with treatment with stimulant medication in an adult ADHD clinical trial. Exome capture was done using a commercially available Agilent SureSelect in solution method. Paired end sequencing was performed using the Illumina Genome Analyzer IIx platform with read lengths of 76 base pairs, providing at least 20x-coverage. Sequence reads were aligned to human reference genome builds hg18 and hg19 using both Novoalign 2.07 and SOAPaligner 2.20. Single nucleotide polymorphism (SNP) calling was conducted using both Maq 0.5.0 and SOAPsnp 1.03. SNP results were filtered as followed: Base quality > 20, depth from 4-200, copy number estimate < 2, and distance between two adjacent SNPs no less than 5. Nonsynonymous SNPs (nsSNPs) were identified using ANNOVAR and USeq Alleler, and insertions/deletions (indels) were detected using SOAPindel. Illumina 610k SNP microarrays were also run on each of the three samples, with copy number variant (CNV) detection using PennCNV.

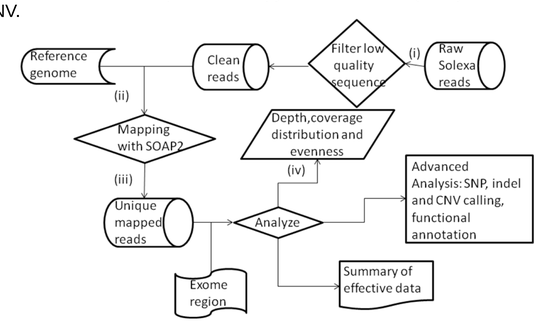
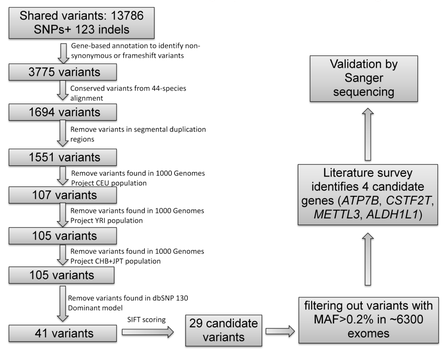


Figure 1. Identification of family-specific ADHD candidate genes by ANNOVAR stepwise reduction procedure, based on shared variants among 3 patients and assuming a dominant disease model.



Results

Table 1. Summary of SNPs for exome capture samples

ExomeCapture	84060	84615	92157
Number of genomic positions for calling SNPs	52,105,343	52,105,343	52,105,343
Number of high-confidence genotypes*	46,651,026	45,555,039	46,393,954
Number of high-confidence genotypes in target regions	35,571,754	34,856,090	35,742,301
Number of known SNPs in target region	192,415	192,415	192,415
Coverage of known SNPs	148,795 (77.33%)	142,933 (74.28%)	148,396 (77.12%)
Total number of novel SNPs	25,523	24,470	25,321
Total number of novel indels	335	325	335

Table 2. Biochemical assays of enzyme activities in the patient affected with idiopathic hemolytic anemia confirmed PKLR deficiency.

	Patient 84060	Control	Reference values
PK (U/gHb)	3.3 L	8.6	6.1 – 12.3
HK (U/gHb)	3.2 H	1.1	0.8 – 1.5
G6PD (U/gHb)	15.8 H	9.2	6.4 – 10.5

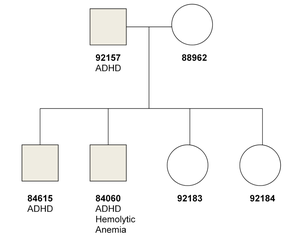
PK, pyruvate kinase; HK, hexokinase; G6PD, glucose-6-phosphate dehydrogenase.

Analysis and filtering of shared novel variants among three family members with ADHD identified 29 candidate variants. Comparison to variant data from ~6,300 exomes revealed that 12 of these variants have minor allele frequency >0.2%. In addition, one family member, who had undergone a splenectomy for idiopathic hemolytic anemia, was found to have two rare non-synonymous mutations in PKLR, consistent with a diagnosis of red blood cell pyruvate kinase deficiency. This finding was confirmed by functional biochemical testing the majority of the SNPs were validated by Sanger sequencing, and validation of the indels is underway.

Table 3. 29 candidate variants for ADHD and their population frequencies in other cohorts

# chrom	Position in HG18	Reference allele	Mutant allele	Gene	Type of Mutation	Amino acid change	# variants in BGI exomes[2]	% in BGI exomes	# variants in ~600 Baylor exomes	% in Baylor exomes
chr17	64384287	A	G	ABCA8	Nonsynonymous	C1387R	0	0.00%	0	0.00%
chr11	68323378	G	A	CPT1A	Nonsynonymous	L193F	0	0.00%	0	0.00%
chr8	101063450	A	G	RGS22	Nonsynonymous	I1084T	0	0.00%	0	0.00%
chr18	59805227	G	T	SERPINB8	Nonsynonymous	G287V	0	0.00%	0	0.00%
chr1	205267500	-	T	C1orf116	frameshift insertion		34	1.40%	0	0.00%
chr17	42571111	-	TGGC	CDC27	frameshift insertion		5	0.20%	0	0.00%
chr17	42589406	A	-	CDC27	frameshift deletion		31	1.30%	0	0.00%
chr1	109266693	TTC	-	GPSM2	amino acid deletion		9	0.40%	0	0.00%
chr9	8577604	AT	-	HNRNPK	frameshift deletion		37	1.60%	0	0.00%
chr11	5432010	-	CA	OR51I2	frameshift insertion		15	0.60%	0	0.00%
chr17	36038691	C	-	SMARCE1	frameshift deletion		19	0.80%	0	0.00%
chr7	98839005	C	-	PDAP1	frameshift deletion		37	1.60%	1	0.20%
chr18	27355154	T	G	DSG2	Nonsynonymous	V158G	1	~0.0%	1	0.20%
chr3	**127359980**	**G**	**A**	**ALDH1L1**	**Nonsynonymous**	**P107L**	**2**	**~0.0%**	**0**	**0.00%**
chr13	**51440681**	**A**	**G**	**ATP7B**	**Nonsynonymous**	**V536A**	**1**	**~0.0%**	**1**	**0.20%**
chr10	**53128652**	**A**	**C**	**CSTF2T**	**Nonsynonymous**	**C222G**	**4**	**0.10%**	**1**	**0.20%**
chr14	**21041859**	**G**	**A**	**METTL3**	**Nonsynonymous**	**R36W**	**9**	**0.20%**	**1**	**0.20%**
chr11	76632438	-	A	GDPD4	frameshift insertion		36	1.50%	6	1.00%
chr7	86998554	A	T	ABCB1	Nonsynonymous	S893T	815	14.3%[1]	9	1.50%
chr11	133634133	C	G	ACAD8	Nonsynonymous	S171C	112	2.00%	20	3.30%
chr20	17904347	C	T	C20orf72	Nonsynonymous	R178W	23	0.40%	8	1.30%
chr8	33438433	T	C	FUT10	Nonsynonymous	Q27R	15	0.30%	3	0.50%
chr13	19695025	A	T	GJB6	Nonsynonymous	S199T	68	1.20%	4	0.70%
chr16	69572830	G	T	HYDIN	Nonsynonymous	P1491H	77	1.40%	dozens	>5.0%
chr10	22059861	G	A	MLLT10	Nonsynonymous	R713H	15	0.30%	6	1.00%
chr17	10355994	A	G	MYH1	Nonsynonymous	Y435H	99	1.70%	14	2.30%
chr1	143727234	G	T	PDE4DIP	Nonsynonymous	L142I	1256	22.10%	hundreds	>30.0%
chr2	98175864	T	C	VWA3B	Nonsynonymous	I513T	15	0.30%	16	2.70%
chr5	115230317	AAGA	-	AP3S1	frameshift deletion		185	7.80%	19	3.20%

Figure 2: Pedigree of a Utah family affected with idiopathic hemolytic anemia and ADHD.



Discussion

Preliminary results show four interesting candidate rare nonsynonymous variants in brain-expressed genes that are shared among 3 individuals (father and two sons) and thus may be associated with ADHD in this family. Validation of these findings is underway in each individual, along with expansion of the pedigree to include DNA sequence and phenotype data from the unaffected mother and other siblings, along with functional characterization of the nsSNPs and PKLR mutations. These preliminary results show that exome sequencing can identify genetic causes of mendelian disorders, as in the case of pyruvate kinase deficiency in one member of this family, and also has potential to discover rare variants associated with complex neuropsychiatric conditions such as ADHD.

Conclusion

Collectively, our results identify new candidate genes for ADHD, and suggest that exome sequencing may be an efficient strategy for solving diagnostic unknowns.

Author Affiliations

1. Department of Psychiatry, 30 North 1900 East, University of Utah, Salt Lake City, UT 84132
2. NYU Child Study Center, New York, NY 10016
3. BGI-Shenzhen 518083, China
4. Genome Research Institute, Shenzhen University Medical School, Shenzhen 518060, China
5. Eccles Institute of Human Genetics, University of Utah, Salt Lake City, UT 84132
6. Department of Clinical Chemistry and Haematology, University Medical Center Utrecht, Netherlands
7. Department of Biology, University of Copenhagen, Copenhagen, Denmark
8. Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA 19104
9. Department of Statistics, Brigham Young University, Provo, UT 84602
10. The first three authors should be regarded as joint first authors.
11. Present address: Zilkha Neurogenetic Institute, Department of Psychiatry and Preventive Medicine, University of Southern California, Los Angeles, CA 90089

Acknowledgements

We thank Matthew Bainbridge and Richard Gibbs at Baylor for access to unpublished exome data. This work was funded by University of Utah Department of Psychiatry funds to G.L., R.R. and F.R. B.M. & M.Y. were supported by NHGRI 1RC2HG005619. K.W. and H.H. were funded by a Pilot/Methodological Study Award from NIH/NCRR Grant UL1 RR025774.

THE UNIVERSITY OF UTAH