Validation and Implementation of a 62-Gene Next Generation Sequencing Panel for Syndromic Autism: Experience from a Clinical Diagnostic Laboratory

Julie R Jones[1], Monica J Basehore[1], Stephen McGee[1], Amanda Lindy[1], Kristen J Champion[1], Christin D Collins[2], Madhuri Hegde[2], Michael J Friez[1]

[1]Greenwood Genetic Center, Greenwood, SC [2]Emory Genetics Laboratory, Decatur, GA

Greenwood Genetic Center
Where Compassion Inspires Progress

www.GGC.org

Abstract

Autism spectrum disorders (ASD) are complex developmental disabilities affecting as many as one in 150 individuals. Coexisting genetic disorders are found in a subgroup of patients with autism (estimated to be between 11 and 37%). In an effort to provide comprehensive testing for this subgroup, the Molecular Diagnostic Laboratory at the Greenwood Genetic Center (GGC) collaborated with investigators at Emory Genetics Laboratory and RainDance Technologies™ to design a Next Generation Sequencing (NGS) panel for 62 genes associated with ASD (available commercially from RainDance Technologies™ as the ASDSeq™ Research Screening Panel). These genes were selected to represent the most common single gene etiologies associated with a syndrome that includes autism as a significant clinical feature. The panel utilizes microdroplet-based multiplex PCR technology to enrich for 1164 exons for NGS with greater than 98% coverage of the targeted regions. The coverage of the panel includes all exons for each gene, at least 50 nucleotides upstream and downstream of each exon and 1 kb of both the 5' promoter regions and the 3' UTRs. The panel was recently validated at GGC as a diagnostic test on both the SOLiD™ 3 Plus and 5500xl systems using 16 positive controls. In addition, 40 patients with autism of unknown etiology have been tested to date. Here we describe our validation process and the implementation of diagnostic testing using the NGS Syndromic Autism 62-Gene Panel. This panel provides comprehensive testing for patients with syndromic autism and may also serve as a useful 2nd tier test for patients with features overlapping those of Rett and Angelman syndromes.

Syndromic Autism 62-Gene Panel

- *AP1S2*
- *ARX*
- *ATRX*
- *AVPR1A*
- *BDNF*
- *BRAF*
- *CACNA1C*
- *CASK*
- *CDKL5*
- *CHD7*
- *CNTNAP2*
- *CREBBP*
- *DHCR7*
- *DMD*
- *EHMT1*
- *FGD1*
- *FMR1*
- *FOLR1*
- *FOXG1*
- *FOXP1*
- *FOXP2*
- *GABRB3*
- *GLUT1*
- *HOXA1*
- *HPRT1*
- *KDM5C*
- *L1CAM*
- *MBD5*
- *MECP2*
- *MED12*
- *MEF2C*
- *MET*
- *MID1*
- *NHS*
- *NIPBL*
- *NLGN3*
- *NLGN4X*
- *NRXN1*
- *NSD1*
- *OPHN1*
- *PAFAH1B1*
- *PCDH19*
- *PHF6*
- *PQBP1*
- *PTCHD1*
- *PTEN*
- *PTPN11*
- *RAB39B*
- *RAI1*
- *RELN*
- *SCN1A*
- *SHANK3*
- *SLC6A4*
- *SLC9A6*
- *SMC1A*
- *TCF4*
- *TSC1*
- *TSC2*
- *UBE3A*
- *VPS13B*
- *ZEB2*

Upon reviewing the current literature and utilizing the Autism Genetic Database (http://wren.bcf.ku.edu/), 62 genes were selected for inclusion in the ASDSeq™ Research Screening Panel. These genes fall into 3 general categories: 1) those known to harbor mutations associated with ASD; 2) genes associated with syndromic autism (genes highlighted in red are associated with disorders having phenotypic overlap with Rett and Angelman syndromes); and 3) candidate genes for autism based on association studies.

Panel design:

❖1164 exons (2314 amplicons; 220,140 bp)

❖50 nucleotides upstream and downstream of each exon*

❖1 kb of both the 5' promoter regions and the 3' UTRs*

*GGC is analyzing the coding regions and 25 nucleotides upstream and downstream of each exon

Methods

5 µg of each genomic DNA sample was fragmented to 4-6 kb using a Covaris. RainDance enrichment for the 62 genes was performed on the RDT1000 instrument (RainDance Technologies, Lexington, MA). Standard fragment libraries for analysis by SOLiD sequencing were prepared for each sample. Library amplification was performed using emulsion PCR, and the products of the emulsion PCR were purified (AMPure) and then deposited onto a glass slide for analysis by either the SOLiD™ 3 Plus or 5500xl system (Applied Biosystems, Foster City, CA). After bioinformatically processing the data, a tertiary analysis was performed using NextGENe software (SoftGenetics, State College, PA). The data was then reviewed with an initial emphasis on novel alterations and those reported in the Human Gene Mutation Database (HGMD: http://www.hgmd.cf.ac.uk/ac/index.php). The alterations identified were cross-referenced to the other samples within the same run as well as to a cumulative database that archives all the sample data from previous runs. After filtering the known SNPs, this comparison allows us to determine which changes are most likely novel and warrant further consideration. All changes deemed of potential clinical relevance were confirmed by Sanger sequencing. Sixteen abnormal controls that had been previously identified by Sanger sequencing were used in the validation of the panel. Eight controls were run on the SOLiD™ 3 Plus instrument, and 8 were run on the 5500xl. In addition, 40 samples from patients with ASD of unknown etiology were analyzed.

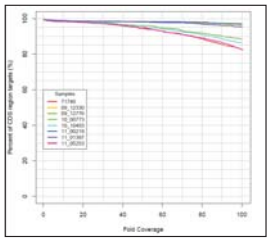
Validation of the Syndromic Autism 62-Gene Panel

Gene	Alteration	Detected by NGS?
MED12	c.4832G>A hemi (p.R1611H)	Yes (3 Plus)
FGD1	c.912delG hemi	Yes (3 Plus)
UBE3A	c.1814_1816delCTT het	Yes (3 Plus)
MECP2	c.763C>T het (p.R255X)	No (3 Plus)
PTPN11	c.922A>G het (p.N308D)	Yes (3 Plus)
TCF4	c.1418_1421delCGGT het	Yes (3 Plus)
NSD1	c.5990A>G het (p.Y1997C)	Yes (3 Plus)
PTEN	c.235G>A het (p.A79T)	Yes* (3 Plus)
ATRX	c.7153_7154insG hemi	Yes (5500)
CDKL5	c.1178C>T het (p.T393I)	Yes (5500)
FGD1	c.1643_1644delTG hemi	Yes (5500)
L1CAM	c.523+12C>T het	Yes (5500)
UBE3A	c.911_912insT	Yes (5500)
PTEN	c.235G>A het (p.A79T)	Yes (5500)
CHD7	c.559C>T het (p.Q187X)	Yes (5500)
TCF4	c.559C>T het (p.Q187X)	Yes (5500)

All of the known alterations for the 16 abnormal controls were detected with the exception of the *MECP2* nonsense mutation. We knew from our experience with the 92-Gene XLID Panel that this region of the gene was problematic for NGS due to poor coverage. This region of *MECP2* is, therefore, included in a recurrent dropout panel that is analyzed by Sanger sequencing for every sample submitted for the Syndromic Autism 62-Gene Panel. Other targeted regions with routinely poor coverage as well as those regions that consistently have no coverage by NGS are also included in the Sanger recurrent dropout panel.

*The *PTEN* c.235G>A alteration was detected by the SOLiD™ 3 Plus system; however, our initial analysis did not identify the change since the abnormal allele percent (18%) was below our established threshold of 20%. When we lowered the threshold, the analysis identified the alteration. This abnormal control was also subsequently run on the 5500xl, and the alteration was detected in our initial analysis (abnormal allele = 55%).

Depth of Coverage



❖Average fold coverage of targeted regions: 168X

❖AVG exon span coverage: 97.6%

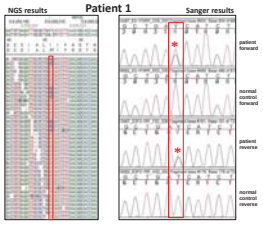
❖AVG number of dropouts: 18

sample 1	sample 2	sample 3	sample 4	sample 5	sample 6	sample 7	sample 8
BRAF_cds2	BRAF_cds2	BRAF_cds2	BRAF_cds2	BRAF_cds2	BRAF_cds2	BRAF_cds2	BRAF_cds2
CASK_cds8	CASK_cds8	CASK_cds8	CASK_cds8	CASK_cds8	CASK_cds8	CASK_cds8	CASK_cds8
DMD_cds43	DMD_cds43	DMD_cds43	DMD_cds43	DMD_cds43	DMD_cds43	DMD_cds43	DMD_cds43
EHMT1_cds1	EHMT1_cds1	EHMT1_cds1	EHMT1_cds1	EHMT1_cds1	EHMT1_cds1	EHMT1_cds1	EHMT1_cds1
FMR1_cds7	FMR1_cds7	FMR1_cds7	FMR1_cds7	FMR1_cds7	FMR1_cds7	FMR1_cds7	FMR1_cds7
KDM5C_cds24	KDM5C_cds24	KDM5C_cds24	KDM5C_cds24	KDM5C_cds24	KDM5C_cds24	KDM5C_cds24	KDM5C_cds24
L1CAM_cds2	L1CAM_cds2	L1CAM_cds2	L1CAM_cds2	L1CAM_cds2	L1CAM_cds2	L1CAM_cds2	L1CAM_cds2
MED12_cds36	MED12_cds36	MED12_cds36	MED12_cds36	MED12_cds36	MED12_cds36	MED12_cds36	MED12_cds36
NIPBL_cds14	NIPBL_cds14	NIPBL_cds14	NIPBL_cds14	NIPBL_cds14	NIPBL_cds14	NIPBL_cds14	NIPBL_cds14
NRXN1_cds5	NRXN1_cds5	NRXN1_cds5	NRXN1_cds5	NRXN1_cds5	NRXN1_cds5	NRXN1_cds5	NRXN1_cds5
PCDH19_cds2	PCDH19_cds2	PCDH19_cds2	PCDH19_cds2	PCDH19_cds2	PCDH19_cds2	PCDH19_cds2	PCDH19_cds2
PTEN_cds7	PTEN_cds7	PTEN_cds7	PTEN_cds7	PTEN_cds7	PTEN_cds7	PTEN_cds7	PTEN_cds7
RELN_cds51	RELN_cds51	RELN_cds51	RELN_cds51	RELN_cds51	RELN_cds51	RELN_cds51	RELN_cds51
RELN_cds64	RELN_cds64	RELN_cds64	RELN_cds64	RELN_cds64	RELN_cds64	RELN_cds64	RELN_cds64
SHANK3_cds12	SHANK3_cds12	SHANK3_cds12	SHANK3_cds12	SHANK3_cds12	SHANK3_cds12	SHANK3_cds12	SHANK3_cds12
SLC6A4_cds8	SLC6A4_cds8	SLC6A4_cds8	SLC6A4_cds8	SLC6A4_cds8	SLC6A4_cds8	SLC6A4_cds8	SLC6A4_cds8
TSC1_cds20	TSC1_cds20	TSC1_cds20	TSC1_cds20	TSC1_cds20	TSC1_cds20	TSC1_cds20	TSC1_cds20
HPRT1_cds1	HPRT1_cds1	HPRT1_cds1	HPRT1_cds1	HPRT1_cds1			RELN_cds62
	SHANK3_cds1	RELN_cds62					

The table above shows the regions that yielded no coverage from the last run of autism samples. Data from previous runs was very similar. The regions highlighted in dark blue showed zero coverage for all 8 samples. The regions highlighted in light blue showed zero coverage for two or more samples, and the region highlighted in yellow was a unique dropout for sample 2. All unique dropouts were Sanger sequenced in addition to the recurrent dropout panel described above.

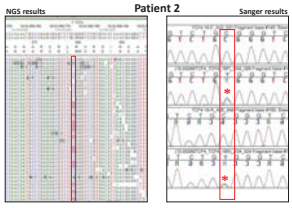
Patient Results

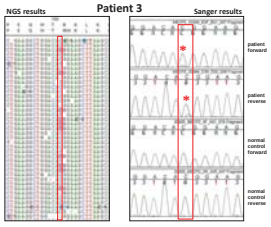
For each of the 40 patients tested thus far, an average of 5 to 10 variants of unknown clinical significance were detected. Family studies will be performed to help define the pathogenicity of these changes. In addition to these variants, two pathogenic mutations and one likely pathogenic alteration were detected. These three cases are described below:



Patient 1 is a 19 year old female with intellectual disability, autism, speech delay and seizures. A heterozygous missense change (c.137T>C; p.I46T) was detected in the *MEF2C* gene by NGS which is predicted by both SIFT and PolyPhen to be damaging. The change was confirmed by Sanger and was not maternally inherited. The paternal sample is not available for analysis.

Patient 2 is a non-verbal, non-ambulatory 2 1/2 year old male with hypotonia, global developmental delay and acquired microcephaly. A heterozygous missense mutation (c.1738C>T; p.R580W) was detected in the *TCF4* gene by NGS and confirmed by Sanger.





Patient 3 is an 11 year old female with acquired microcephaly and other features consistent with Rett syndrome. In 2002, a *MECP2* analysis was performed by Sanger sequencing which yielded a normal result. NGS using this panel detected a mosaic heterozygous mutation (c.316C>T; p.R106W) in *MECP2*. Sanger sequencing was repeated, and the abnormal allele was again not detected by the software but could be visualized at a very low level.

Summary

❖Our evaluation of the ASDSeq™ Research Screening Panel using both the SOLiD™ 3Plus and 5500xl systems yielded good results with an average coverage of 168X for the targeted regions and an average exon span coverage of 97.6%.

❖We found concordance of positive controls during the validation of the panel with the exception of a *MECP2* mutation located in an area with poor coverage. Sanger sequencing panels were designed for this region as well as other areas with poor coverage, recurrent dropouts and unique dropouts for each sample. The addition of these Sanger panels brings the average exon span coverage to 99.3%.

❖This 62-gene NGS panel provides a more comprehensive analysis for syndromic autism and may serve as a 2nd tier analysis for Rett/Angelman-like phenotypes.

Acknowledgments

We sincerely thank the families that have enrolled in the South Carolina Autism Project at GGC. We also thank the clinicians and counselors that have referred patients for clinical testing. Lastly, we recognize the contributions of our laboratory staff and the clinical team at GGC for their considerable efforts in providing comprehensive testing and evaluation of patients with autism.