Analytical and Clinical Validity Study of FirstStep[Dx] PLUS: A Chromosomal Microarray Optimized for Patients with Neurodevelopmental Conditions

Charles H. Hensel[1], Rena Vanzo[1], Megan Martin[1], Sean Dixon[1,] Christophe G. Lambert[2], Brynn Levy[3], Lesa Nelson[4], Andy Peiffer[1,5], Karen Ho[1,5], Moises Serrano[1], Sarah South[6], Kenneth Ward[4], E. Robert Wassman[1]

1) Lineagen, Inc. (Salt Lake City, UT); 2) University of New Mexico School of Medicine, Albuquerque, NM 3) Columbia University Medical Center, NY, NY 4) Affiliated Genetics Laboratory, Salt Lake City, UT; 5) University of Utah, Department of Pediatrics 6) ARUP, Salt Lake City, UT (Currently: 23&Me)

Abstract:

Introduction: Chromosomal microarray analysis (CMA) is recognized as the first-tier test in the genetic evaluation of children with developmental delays, intellectual disabilities, congenital anomalies and autism spectrum disorders of unknown etiology.

Array Design: To optimize detection of clinically relevant copy number variants associated with these conditions, we designed a whole-genome microarray, FirstStep[Dx] PLUS (FSDX). A set of 88,435 custom probes was added to the Affymetrix CytoScanHD platform targeting genomic regions strongly associated with these conditions. This combination of 2,784,985 total probes results in the highest probe coverage and clinical yield for these disorders.

Results and Discussion: Clinical testing of this patient population is validated on DNA from either non-invasive buccal swabs or traditional blood samples. In this report we provide data demonstrating the analytic and clinical validity of FSDX and provide an overview of results from the first 7,570 consecutive patients tested clinically. We further demonstrate that buccal sampling is an effective method of obtaining DNA samples, which may provide improved results compared to traditional blood sampling for patients with neurodevelopmental disorders who exhibit somatic mosaicism.

Clinical scenario: Neurodevelopmental disabilities, including developmental delays (DD), intellectual disabilities (ID), and autism spectrum disorders (ASD), affect up to 15% of children (1). In the majority of cases, a child's clinical presentation does not allow for a definitive etiological diagnosis. In such cases, CMA is recommended as the first-tier test that should be used to evaluate for a potential genetic etiology (2-7). A definitive genetic diagnosis allows patients to more often receive appropriate medical care tailored to their condition, as reflected by medical management changes and improved access to necessary support and educational services (8-13).

Test description: FirstStep[Dx] PLUS (FSDX) is an optimized clinical microarray test provided in the context of a comprehensive clinical service. Testing starts with either a non-invasive buccal swab sample or traditional blood sample from which DNA extraction using a Gentra Puregene® kit specific to the sample type (Qiagen, Inc., Valencia, CA) is performed in one of several contracted CLIA/CAP credentialed laboratories according to manufacturers' protocols. High quality genomic DNA is fragmented, labeled and hybridized to FSDX arrays using reagents, equipment and methodology as specified by the

manufacturer (Affymetrix, Inc., Santa Clara, CA)(14). Washed arrays are scanned and raw data files are processed to CYCHP files using a reference file comprising at least 100 samples with normal array findings. Data analysis is performed using Chromosome Analysis Suite software version 2.0.1 (Affymetrix). Hybridization of patient DNA to oligonucleotide and SNP probes is independently compared against a previously analyzed cohort of normal samples to call CNVs and allele genotypes. The percentage mosaicism of whole-chromosome aneuploidies is determined using the average log_2 ratio of the entire chromosome (14).

Microarray design: FSDX was optimized by the addition of 88,435 custom probes targeting genomic regions strongly associated with ID/DD/ASD (15-24). This was effected, under GMP by Affymetrix, to the CytoScanHD platform using their microarray design process specifications which have been previously described (14). This is consistent with the ACMG recommendation of "enrichment of probes targeting dosage-sensitive genes known to result in phenotypes consistent with common indications for a genomic screen" (25). Critical regions that did not meet a desired probe density >1 probe/1000 bp on the CytoScanHD were supplemented with additional probe content to allow for improved detection of smaller deletions and duplications in these critical regions. Finally, additional probes were added to improve detection of CNVs encompassing genes involved in other well-characterized neurodevelopmental disorders, for example *GAMT* (26) and *GATM* (27). All incremental probes were added in substitution for probes deemed sub-optimal by Affymetrix and previously masked, bringing FSDX to a grand total of 2,784,985 probes. Custom SNP probes (n =416) on FSDX are targeted by 12 oligonucleotides, three for each strand of each allele, which is approximately double the typical probe coverage for SNPs.

Test interpretation: CYCHP files are evaluated by ABMG certified cytogeneticists. Determination of CNVs is consistent with established cytogenetic standards. A minimum of 25-consecutive impacted probes is the baseline determinant for deletions and 50 probes for duplications independent of variant size. Rare CNVs are determined to be "pathogenic" if there is sufficient evidence published (at least two independent publications) to indicate that haploinsufficiency or triplosensitivity of the region or gene(s) involved is causative of clinical features or of sufficient overall size (28). If however, there is insufficient but at least preliminary evidence for a causative role for the region or gene(s) therein they are classified as variants of unknown significance (VOUS) independent of CNV size. Areas of absence of heterozygosity (AOH) are also classed as VOUS if of sufficient size and location to increase the risk for conditions with autosomal recessive inheritance or conditions with parent-of-origin/imprinting effects. Other CNVs are typically not reported after determination that they most likely represent normal common population variants and are contained in databases documenting presumptively benign CNVs, e.g., the Database of Genomic Variants (DGV) (29). These parameters were standard independent of the microarray used for analysis in comparative studies.

Public health importance: A definitive genetic diagnosis facilitates patient access to appropriate and necessary medical and support services. Defining the underlying genetic cause of DD/ID/ASD and/or multiple congenital anomalies (MCA) in each unique patient is vital to understanding etiology, prognosis, and course. It informs physicians of potential comorbid conditions for which a patient should be evaluated and treated proactively and optimally. Improved understandings of the appropriate

therapeutic and behavioral approaches to that patient are also enabled. Genetic testing is best provided in the context of an integrated service (30), so FSDX provides comprehensive, clear, readable, and personalized reports for the healthcare provider and a family-friendly section to facilitate understanding of the often-complex results. The report is complemented by availability of pre- and post-test genetic counseling and technical support to providers. Moreover, FSDX includes personalized insurance pre-authorization and appeals assistance to overcome barriers encountered by both providers and families that, in many circumstances, prevent access to crucial genetic testing services (10-11).

Published reviews, recommendations and guidelines: The American College of Medical Genetics (ACMG) (2,3), the American Academy of Child and Adolescent Psychiatry (4,7), the American Academy of Pediatrics (5), and the American Academy of Neurology/Child Neurology Society (6) recommend CMA as the first-tier test in the genetic evaluation of children with unexplained DD, ID, or ASD. Considerable data supporting these guidelines are documented in numerous reviews and publications (31-36). The ACMG has also published guidelines on both array design (25) and the validation of arrays, including validation of a new version of a platform in use by the laboratory from the same manufacturer, and of additional sample types (28).

EVIDENCE OVERVIEW

Validation of novel (Blood vs. Buccal) sample type on the established and optimized platforms

It is highly desirable to avail clinicians and families a less invasive sampling method for individuals with ID/DD/ASD due to the potential implications of venipuncture in many such individuals. We validated a buccal sampling methodology in conjunction with two independent CLIA-certified laboratories, ARUP (Salt Lake City, UT) and Fullerton Genetics Center (Asheville, NC) first on the CytoScanHD array and then the FSDX array. After preliminary consideration of multiple saliva and buccal DNA collection kits, we selected the ORAcollect-100® (DNA Genotek, Ottawa)(now cleared as a class II IVD medical device: ORAcollect-DX®) based upon ease of use for the intended patient population, ease of shipping processes, DNA stability, and post-extraction quality studies for further validation. Buccal swabs and blood samples from twenty-two individuals underwent parallel microarray analysis for concordance by each protocol. Twenty-two individuals' buccal and blood samples were analyzed in terms of array quality control metrics and no significant difference between the two sample sources was observed. In addition, there was 100% concordance of CNV calls between the two sample types. Finally, twenty-three individuals' blood and buccal samples underwent microarray analysis on the FSDX platform. Again no significant difference between the two sample sources in terms of quality metrics and 100% clinical concordance of CNV calls was observed.

Clinical Validation of a new version of a previously established platform

FSDX was validated independently in four CLIA-certified laboratories (Asuragen – Asuragen, Inc., Austin, TX; AGI – Affiliated Genetics, Inc., Salt Lake City, UT; Fullerton - Mission Hospital/Fullerton Genetics Center, Asheville, NC; CUMC – Columbia University Medical Center, Dept. of Pathology and Cell Biology, New York, NY.) all previously familiar with performing CMA on CytoScanHD (or its predecessor the Affymetrix 2.7M Cytogenetics array) and cross-validated between these laboratories as well. Data demonstrating the concordance of FSDX with these alternative arrays and across laboratories are shown

in Tables I-II. A total of six samples from patients having clinically significant CNVs findings on prior clinical analysis were re-analyzed using FSDX. This study mirrors ACMG guidance on clinical validation of a new version of a previously established platform when the total probe content change is less than 5% (here 3.3% increase) (28). There was complete clinical concordance between the initial clinical results and the results generated with FSDX in two independent laboratories (AGI and CUMC). Although the cross-platform and cross-laboratory results are unequivocally clinically concordant, minor differences in breakpoint determinations were observed in the majority of analyses (8 of 11), which is expectable and reasonable within the limits of the technology and interpretation overall. The differences in breakpoints between FSDX and CytoScanHD resulted in CNVs that were smaller, but only on average by $\leq$ 0.25% of the total CNV size. It was previously observed that arrays with increasing increased probe density result in smaller estimates of total CNV size presumably through the higher resolving power with increased probe density (14). In contrast, only a single inter-laboratory analysis of FSDX differed in breakpoint calls by the different cytogeneticists, and in this case the change was only 0.08% of the overall CNV size.

Gender	Clinical Laboratory	Clinical Array	Clinical Result	Validation Laboratory 1	Validation Result 1	Custom Probes	Validation Laboratory 2	Validation Result 2
M	ARUP	2.7M Array	arr 15q11.2q14(18,446,423-32,204,913)x2hmz, 15q25.1qter(77,139,194-100,201,798)x2hmz (hg18)	AGI	arr 15q11.2q14(22,752,398-34,652,683) hmz, 15q25.1q26.3(79,465,893-102,429,049) hmz(hg19)	1191	CUMC	arr 15 q11.2q14(22,752,398-34,652,683)x2hmz, 15q15.1q21.1(42,484,889-46,012,198)x2hmz, 15q25.1q26.3(79,465,893-102,429,049)x2hmz
M	ARUP	CytoScanHD	arr 16p11.2(29,580,020-30,178,406)x3 (hg19)	AGI	arr 16p11.2(29,421,694-30,177,916)x3(hg19)	0	CUMC	arr 16p11.2(29,580,020-30,178,406 x3)(hg19)
M	Asuragen	CytoScanHD	arr (1-22)x2,(XY)x1 (hg19) - LCSH on CHR1	AGI	arr (1-22)x2,(XY)x1 (hg19) - LCSH on CHR1 (10.5 Mb 1p)		CUMC	arr (1-22)x2,(XY)x1 (hg19)- LCSH on CHR1
F	Fullerton	CytoScanHD	arr 16p12.2(22,086,845-22,442,007)x1 (hg19)	AGI	arr 16p12.2(21,946,522-22,442,007)x1 (hg19)	0	CUMC	arr 16p12.2(22,086,845-22,442,007)x1 (hg19)
F	Fullerton	CytoScanHD	arr 22q13.33(49,423,616-51,197,838)x1 (hg19)	AGI	arr 22q13.33(49,424,409-51,197,838)x1(hg19)	56	CUMC	arr 22q13.33(49,423,616-51,197,838)x1(hg19)
M	Asuragen	CytoScanHD	arr 15q11.2q13.1(23,620,191-28,522,838)x1 (hg19)	AGI	arr 15q11.2q13.1(23,642,700-28,561,671)x1(hg19)	0		

Table I: Clinical validity of the FirstStep^Dx PLUS array. Samples from patients analyzed clinically on commercially available Affymetrix microarrays were run independently on FSDX. Laboratory designations are as follows: Asuragen – Asuragen, Inc., Austin, TX; AGI – Affiliated Genetics, Inc., Salt Lake City, UT; Fullerton - Mission Hospital/Fullerton Genetics Center, Asheville, NC; CUMC – Columbia

University Medical Center, Dept. of Pathology and Cell Biology, New York, NY. All arrays used clinically were purchased from Affymetrix, Inc., Santa Clara, CA.

As further demonstration of the clinical validity of this array, 16 samples from an earlier research study using an alternative technology platform (Illumina, San Diego, CA) (15) were analyzed on FSDX in two participating CLIA-laboratories: ARUP & Asuragen (Table II). Samples all had significant CNVs present, which had been validated by quantitative PCR in the research study, and spanned both custom and standard probes. All results were concordant both across platforms and between laboratories.

Lineagen Index	CNV Type	Chromosome Band	Genes	Agree (Y/N)
CNV-1	Duplication	7p22.1	PMS2CL, RSPH10B, CCZ1B	Y
CNV-2	Duplication	5q23.1	DMXL1, TNFAIP8	Y
CNV-3	Duplication	4q13.3	COX18, ANKRD17	Y
CNV-4	Deletion	2p16.3	NRXN1	Y
CNV-4	Deletion	6p24.3	RIOK1	Y
CNV-5	Duplication	4q13.3	COX18, ANKRD17	Y
CNV-6	Duplication	5q23.1	DMXL1, TNFAIP8	Y
CNV-7	Duplication	7p22.1	PMS2CL, RSPH10B, CCZ1B	Y
CNV-8	Deletion	2p16.3	NRXN1	Y
CNV-8	Deletion	6p24.3	Intergenic	Y
CNV-9	Duplication	12q23.2-q24.32	>50	Y
CNV-10	Deletion	2p16.3	NRXN1	Y
CNV-11	Duplication	5q23.1	DTWD2, DMXL1, TNFAIP8, HSD17B4	Y
CNV-13	Duplication	12q23.2	CHPT1	Y
CNV-14	Duplication	7p22.1	PMS2CL, RSPH10B, CCZ1B	Y
CNV-15	Deletion	2p16.3	NRXN1 5'-flanking region	Y
CNV-16	Duplication	5q23.1	DTWD2, DMXL1, TNFAIP8, HSD17B4	Y

Table II: Research findings validated on FirstStep^Dx PLUS. Samples derived from research studies with significant CNVs present (15) which spanned both custom and standard probes were evaluated independently on FSDX in two laboratories and evaluated for agreement with the research findings as well their inter-laboratory concordance.

Inter-laboratory clinical performance validation

Further evidence of the inter-laboratory performance (Table III) is shown on twelve additional patients with clinically significant CNVs detected by clinical testing with FSDX at Asuragen, and then re-analyzed by both AGI and CUMC, again with completely concordant results. In addition, two patients run clinically

at CUMC were concordant with results subsequently generated by AGI, and six patients run clinically at AGI were concordant with results generated by Fullerton Genetics Laboratory Center.

Gender	Clinical Laboratory	Clinical Array	Clinical Result	Validation Laboratory 1	Validation Result 1	Custom Probes	Validation Laboratory 2	Validation result 2
M	Asuragen	FSDx PLUS	arr[1-22]x2,(XY)x1 (hg19)	AGI	arr[1-22]x2,(XY)x1 (hg19)		CUMC	arr [1-22]x2,(XY)x1(hg19)
M	Asuragen	FSDx PLUS	arr 4p16.3p15.31(68,345-17,983,528)x1, 4p15.31(17,983,558-18,562,949)x3 (hg19)	AGI	arr 4p16.3p15.31(68,345-17,983,528)x1, 4p15.31(17,983,558-18,562,949)x3(hg19)	648	CUMC	arr 4p16.3p15.31(68,345- 17,983,528)x1, 4p15.31(17,983,558- 18,562,949)x3(hg19)
F	Asuragen	FSDx PLUS	arr 7p13(45,058,922-45,131,920)x1, 12p13.31p13.1(7,686,191-13,902,005)x1 (hg19)	AGI	arr 7p13(45,058,822-45,135,419)x1, 12p13.31p13.1(7,686,191-13,900,539)x1(hg19)	202	CUMC	arr 7p13(45,058,922-45,131,920)x1, 12p13.31p13.1(7,686,191-13,902,005)x1 (hg19)
F	Asuragen	FSDx PLUS	arr 2q24.1q24.3(158,429,675-166,350,700)x3 (hg19)	AGI	arr 2q24.1q24.3(158,421,800-166,359,851)x3(hg19)	163	CUMC	arr 2q24.1q24.3(158,429,675- 166,350,700)x3 (hg19)
M	Asuragen	FSDx PLUS	arr 4p16.3p16.1(68,345-7,442,049)x1, 8q24.3(146,082,484-146,295,771)x3 (hg19)	AGI	arr 4p16.3p16.1(69,195-7,452,141)x1, 8q24.3(146,086,086-146,295,771)x3(hg19)	329	CUMC	arr 4p16.3p16.1(68,345-7,442,049)x1,8q24.3(146,082,484-146,295,771)x3 (hg19)
M	Asuragen	FSDx PLUS	arr 17p12(14,085,623-15,484,630)x3 (hg19)	AGI	arr 17p12(14,087,918-15,513,053)x3(hg19)	0	CUMC	arr 17p12(14,085,623-15,484,630x3)(hg19)
F	Asuragen	FSDx PLUS	arr 4p16.3(68,345-1,740,152)x1 (hg19)	AGI	arr 4p16.3(68,345-1,768,917)x1(hg19)	127	CUMC	arr 4p16.3(68,345-1,740,152)x1 (hg19)
M	Asuragen	FSDx PLUS	arr[1-22]x2,(XY)x1 (hg19)	AGI	arr[1-22]x2,(XY)x1 (hg19)		CUMC	arr [1-22]x2,(XY)x1(hg19)
F	Asuragen	FSDx PLUS	arr 4p16.3(68,345-2,283,825)x1 (hg19)	AGI	arr 4p16.3(68,345-2,276,067)x1(hg19)	172	CUMC	arr 4p16.3(68,345-2,276,067)x1(hg19)
M	Asuragen	FSDx PLUS	arr[1-22]x2,(XY)x2 (hg19)	AGI	arr[1-22]x2 (XY)x2 (hg19)		CUMC	arr[1-22]x2 (XY)x2 (hg19)
M	Asuragen	FSDx PLUS	arr[1-22]x2,(XY)x1 (hg19)	AGI	arr[1-22]x2,(XY)x1 (hg19)		CUMC	arr [1-22]x2,(XY)x1(hg19)
F	Asuragen	FSDx PLUS	arr Xp11.3p11.23(46,362,229-46,413,283)x3 (hg19)	AGI	arr Xp11.3p11.23(46,362,229-46,413,283)x3(hg19) array defect	0	CUMC	arr Xp11.3p11.23(46,362,229-46,413,283)x3(hg19)
M	CUMC	FSDx PLUS	arr[1-22]x2,(XY)x1 (hg19)	AGI	arr[1-22]x2,(XY)x1 (hg19)			
M	CUMC	FSDx PLUS	arr[1-22]x2,(XY)x1 (hg19)	AGI	arr[1-22]x2,(XY)x1 (hg19)			
F	AGI	FSDx PLUS	11q23.3q25(120,877,867-134,938,470)x3, 17q25.3(80,1...	Fullerton	arr 11q23.3q25(120,877,614-134,938,470)x3, 17q25.3(80,198,434-81,041,938)x1 (hg19)	62		
M	AGI	FSDx PLUS	Xq13.3(74,449,166-74,654,323)x3 (hg19)	Fullerton	arr Xq13.3(74,449,166-74,654,323)x4 (hg19)	3		
M	AGI	FSDx PLUS	arr 10q11.22q11.23(48,121,431-51,780,596)x1, 16p12.2...	Fullerton	arr 10q11.22q11.23(46,966,534-51,780,596)x1, 16p12.2,2(21,596,299-21,734,618)x1 (hg19)	383		
F	AGI	FSDx PLUS	1q21.1(145,382,123-145,832,965)x3 (hg19)	Fullerton	arr 1q21.1(145,372,549-145,832,965)x3 (hg19)	42		
F	AGI	FSDx PLUS	19p13.3(3,968,733-4,092,266)x1 (hg19)	Fullerton	arr 19p13.3(3,974,964-4,091,216)x1 (hg19)	2		
M	AGI	FSDx PLUS	arr 5q23.1(118,462,828-118,622,926)x4 (hg19)	Fullerton	arr 5q23.1(118,456,925-118,622,926)x3~4 (hg19)	124		

Table III: Inter-laboratory validation of the FirstStep^{Dx} PLUS array. Independent samples from patients analyzed clinically at our participating laboratories were re-analyzed at other participating laboratories. Laboratory designations are as described in Table II.

These data clearly demonstrate the ability of FSDX to detect copy number variants on a consistent basis in independent laboratories, across a range of genomic locations with all samples yielding concordant results to those performed on three separate array platforms. The excellent performance in cases with CNVs previously detected in regions with custom probes supports the overall clinical consistency and appropriate performance of the custom probe content on the array.

Extended comparative clinical sensitivity studies in a in real-world clinical population

Data from 7570 consecutive patient samples tested clinically with FSDX from July 2012 (when the optimized FirstStep^{Dx} PLUS (FSDX) microarray was implemented into routine use) through May 2016 are shown in Figure 1. Overall there were 717 (9.5%) pathogenic abnormalities and 1534 (20.2%) VOUS observed, or a 29.7% overall CNV diagnostic yield for potentially abnormal findings. We also compared these results to patients who were tested by Lineagen through the same referral channels and in

comparable patient cohorts and general time windows using the 2.7M and CytoScanHD arrays. The 2.7M arrays (n=378) detected 7.4% pathogenic and 8.5% VOUS for an overall yield of 15.9%. The higher density CytoScanHD arrays (n=1194) detected 9.0% pathogenic and 14.2% VOUS for an overall yield of 23.2%. It is clear that these incremental probe additions translate into potentially clinically significant yields, likely with resultant improvements in patient medical management.

Figure 1 visually summarizes the genomic range of CNVs detected to date using FSDX. These include known microdeletion and microduplication syndrome regions as well as variants of unknown significance.
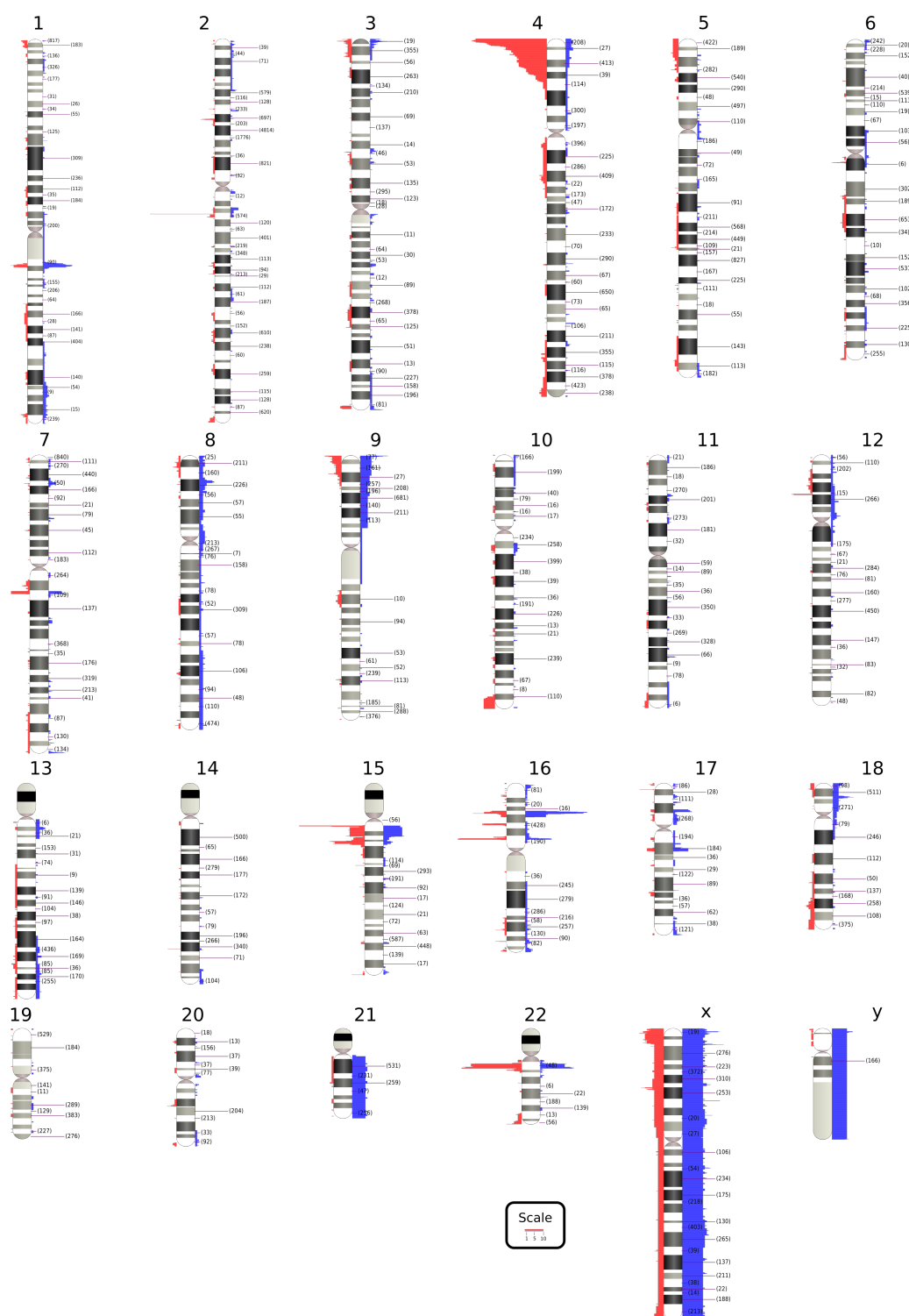
Figure 1: Summary of clinical CNV findings using FirstStep^{Dx} PLUS. Reported clinical findings are displayed next to chromosome ideograms. Deletions are shown on the left in red, and duplications are shown on the right in blue. Numbers in parentheses after chromosome band labels indicate the number of custom-designed probes in those bands. An excess of abnormalities is observed in the 4p region due to a research cohort; however this data is not reflected in the overall detection rates cited in the text.

Quality Control:

Similar to the CytoScanHD, the independent analysis of CNVs with oligonucleotide and SNP probes provides both detection and confirmation simultaneously (14). Three empirically determined quality control metrics are used that reflect overall data quality on an Affymetrix array: 1) waviness-SD, 2) median of the absolute values of all pairwise differences (MAPD), and 3) SNPQC (measure of how well genotype alleles are resolved). The criteria determined empirically for the CytoScanHD array extend to the FSDX array by virtue of design, with all three needed to meet minimum requirements for an array to be analyzed. FSDX arrays meeting these requirements can be interpreted over 99% of the time.

Analytic validity:

To rigorously demonstrate the comparable analytic validity of FSDX, we employed an independent tool, the Golden Helix, Inc. Copy Number Analysis Method (CNAM) (Golden Helix, Inc. Bozeman, MT)(36). First, Affymetrix Chromosome Analysis Suite (ChAS) version 2.1 (Affymetrix, Santa Clara, CA) was used to create an evaluation set of 205 samples with no reportable clinical finding from a single laboratory. To characterize the analytical function of custom probes added to the FSDX microarray, we used the sum of raw probe signal data across the array as a proxy for input DNA concentration. Across the 205 samples, there was roughly a 4.6 fold difference in total array signal. We then plotted individual probe signal intensity vs. total array signal for all probes on the array and used regression analysis to calculate slopes and y-intercepts for each probe. We divided total intensity by 933,696,453 (the lowest total intensity sum) to generate an X-axis that ranged from 1 to 4.6. Finally, we plotted y-intercept vs. slope for each probe and compared custom probes to CytoScanHD probes. Figure 2 shows a scatter-plot based on data from this subset of 205 samples. Data in red reflect custom probes only found on FSDX, while data in black designate standard CytoScanHD probes. The plot indicates that there is significant overlap between the custom probes and standard probes. However, at the extreme, a cluster of custom probes, with slope and intercept near zero, do not appear to respond well to increasing DNA input. Thus, these are less likely to be strongly copy number responsive, but are not necessarily non-informative. This "sub-par" population represents approximately 20% of the custom FSDX probes. This shift in performance characteristics of the single-pass design custom probes is still within the overall desired operating parameters; it suggests at most a 0.64% (% poorly functioning probes over total probe content) deviation in overall analytical sensitivity. However, given that the total probe content on FSDX is 3.27% greater, these data suggest a net gain in sensitivity of 2.63% (net increase in total optimally performing probes compared to CytoScanHD).

Figure 2
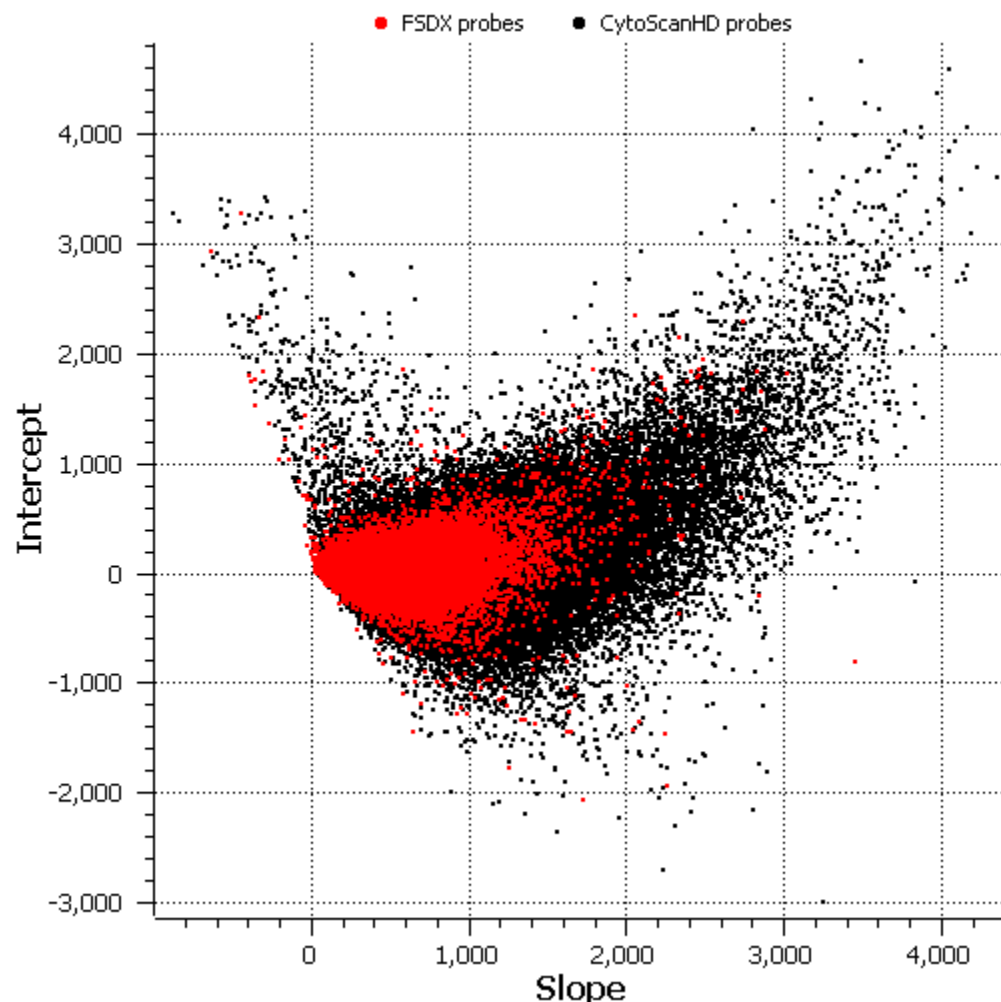
Figure 2: Functional Overlap of Custom Probes with CytoScanHD Probes. Histogram of slope and Y-intercept for each probe on the FirstStep[Dx]PLUS array. Probes shown in red are custom probes found only on FSDX, which overlie probes shown in black common to both FSDX and CytoScanHD. Probes with slope and intercept values near 0 are considered to be sub-optimal probes.

We next analyzed the relative impact of the custom probes compared to the standard array probes on overall detection rates of CNVs. The weighted $\log_2$ ratios from 184 arrays that each had at least one clinically reported finding previously analyzed using ChAS in a single laboratory were evaluated again with CNAM (36). This utilized the univariate option with no moving window; a maximum of 10 segments per 10,000 probes; a minimum segment size of 1 marker; and a stringent permutation p-value threshold of 0.001. After segmentation, we classified segments as losses if their mean was <0.20 with 25 or more probes and classified segments as gains if their mean was >0.20 with 50 or more probes (consistent with our clinical workflows). These calls were compared to the clinically reported findings generated with ChAS. The analysis with CNAM was performed in two ways. First, only probes present on the

CytoScanHD standard array were considered. Second, FSDX custom probes in addition to the standard probes were considered. All clinically relevant findings were observed using both analysis methods (Table IV). Each reported CNV was detectable independent of the presence of custom probes in the CNV and independent of the use of the custom probes in the analysis. These data show no evidence that the previously described sub-optimally responding subset of custom probes interfere with the function of responsive probes, whether standard or custom, or with the overall sensitivity for CNV detection.

The data in Table I also demonstrate the analytic validity of custom probes that are included in some of the clinically reported findings. Increased sensitivity and resolution has been demonstrated with increasing probe density (14), and since the custom probes were added specifically to regions important for ASD and other neurodevelopmental disorders, the resolution of FSDX for these conditions is predicted to be enhanced, since approximately 80% of the custom probes are fully analytically responsive to DNA input. Preliminary analysis (data not shown) suggests that the overall detection of CNVs may be increased by inclusion of the custom probes in this analysis, and further evaluation of this is under investigation.

Chr	# Findings Abnormal		VOUS		# findings with Custom Probes	Agreement +/- custom probes
	Deletions	Duplications	Deletions	Duplications		
chr1	1		2	4	1	Y
chr2	2	1	5	7	9	Y
chr3		2	4	2	1	Y
chr4	2			5	3	Y
chr5	1	1	4	2	3	Y
chr6			6	5	2	Y
chr7	2		4	5	0	Y
chr8	1		2	1	2	Y
chr9	1		6	7	5	Y
chr10			3	4	3	Y
chr11		2		3	1	Y
chr12		2	1	3	3	Y
chr13	3		1	1	3	Y
chr14			3	2	0	Y
chr15	5	1	9	2	1	Y
chr16	5	3	4	5	5	Y
chr17	1	4	3		1	Y
chr18	1		3	2	2	Y
chr19			1		0	Y
chr20			2	1	0	Y
chr21		1		1	1	Y
chr22	5	1	1		6	Y

Table IV: Analytic validity of FirstStep^{Dx}PLUS. Data from clinical samples were evaluated using CNAM on weighted $\log_2$ ratios from 184 arrays as described above. The data from CNVs with and without custom FSDX probe content were evaluated to determine any discrepancy in detection based on inclusion of the custom probes and no evidence of non-concordance was observed.

Clinical utility: The goal of any diagnostic test is improvement in medical management of patients. This is achieved first by reaching the correct diagnosis and second by following the appropriate management & surveillance procedures for that diagnosis. CMA has a clear and positive impact on medical management, as documented in several studies (8-13). Further, CMA testing often results in a correct,

additional, or modified diagnosis for those who are difficult to diagnose with clinical observation alone (37, 38). Even in patients with well-defined conditions, better determining CNV breakpoints with a higher resolution methodology can provide information beyond what is known or assumed from other tests (39). A test, like CMA, which allows us to identify multiple individually rare diagnoses is difficult to assess by traditional measures of clinical utility. This is due, in part, to the fact that each potential diagnosis has specific benefits in terms health and cost of care for a given patient. However, the literature has documented numerous improvements in care (8-13) as a result of reaching a genetic diagnosis for individuals with DD, ID, and/or ASD. Further this represents this a continuum as clinical understanding and care evolves stepwise as a direct result of improved diagnostic clarity. The ability of increased probe density in genomic regions of interest to improve diagnosis is becoming apparent, and each additional correct diagnosis allows incremental opportunities for increased composite clinical utility of such tools.

Limitations:

Given the rarity of individual CNVs and current limits of understanding, this is a relatively small clinical validation cohort. This could limit the certainty and range of conclusions from this evaluation, however it exceeds established guidelines for such validations (28). Analytical validity across millions of data points as in a CMA, can only be assessed by in silico methods such as applied here, however this is superior to mere presumptions of performance typical in the literature.

CMA has proven an important clinical diagnostic advance, however it is limited by the failure to detect a diagnosis in still a majority of affected individuals even with the increased performance with added probe content in FSDX. However, emerging evidence suggests that another significant portion of these cases, which is negative on CMA will have mutations detectable by massively parallel sequencing (NGS) in the future complementing CMA (41). In addition, roughly two-thirds of reported CNVs are VOUS, largely due to the limitations of our clinical experience with these rare conditions. The use of new informatics approaches and databases are emerging to better define the potential relevance and pathogenesis of such currently uncertain findings (42).

Conclusions: Analytic validation of FirstStep[Dx] PLUS coupled with over three years of clinical use have demonstrated the utility of FirstStep[Dx] PLUS in identifying genetic causes for neurodevelopmental conditions. The simple non-invasive sampling procedure, high clinical sensitivity and extensive support services make FirstStep[Dx] PLUS an ideal choice for the clinical genetic evaluation of patients with these disorders.

Competing Interest Statement: CH, RV, MM, SD, KH, MS, and ERW are employees of and hold stock option grants in Lineagen, Inc. AP is an employee of the University of Utah and a paid consultant to Lineagen. LN and KW are employees of Affiliated Genetics, Inc., a contract service laboratory performing FSDX tests. BL is an employee of Columbia University Medical Center, a contract service laboratory performing FSDX tests. Sarah South is a paid consultant to Lineagen, a current employee of 23&Me, and was previously an employee of ARUP, a contract service laboratory conducting tests for Lineagen at the

time of this study. CL is a faculty member at the University of New Mexico and a paid consultant to Lineagen.

Ethics Statement: All involvement with human subjects was done in accordance with the World Medical Association Declaration of Helsinki, and with patient informed consent for anonymous post-clinical testing use of specimens and data.

Acknowledgements:

References:

1. Boyle CA, Boulet S, Schieve LA, et al. Trends in the Prevalence of Developmental Disabilities in US Children, 1997–2008. *Pediatrics* 2011; 127(6):1034-1042.
2. Manning M, Hudgins L. Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. *Genet Med* 2010; 12(11):742-745.
3. Schaefer GB, Mendelsohn NJ. Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genet Med* 2013; 15(5):399-407.
4. Volkmar F, Siegel M, Woodbury-Smith M, et al. Practice Parameter for the Assessment and Treatment of Children and Adolescents with Autism Spectrum Disorder. *J Am Acad Child Adolesc Psychiatry* 2014; 53(2):237-257.
5. Moeschler JB, Shevell M. Comprehensive Evaluation of the Child with Intellectual Disability or Global Developmental Delays. *Pediatrics* 2014; 134(3):e903-e918.
6. Michelson DJ, Shevell MI, Sherr EH, et al. Evidence Report: Genetic and metabolic testing on children with global developmental delay. Report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. *Neurology* 2011; 77(17):1629-1635.
7. Heil KM, Schaaf CP. The Genetics of Autism Spectrum Disorders – A Guide for Clinicians. *Curr Psychiatry Rep* 2013; 15(1):334.
8. Saam J, Gudgeon J, Aston E, Brothman AR. How physicians use array comparative genomic hybridization results to guide patient management in children with developmental delay. *Genet Med* 2008; 10(3):181-186.
9. Coulter ME, Miller DT, Harris DJ, et al. Chromosomal microarray testing influences medical management. *Genet Med* 2011; 13(9):770-776.
10. Ellison JW, Ravnan JB, Rosenfeld JA, et al. Clinical Utility of Chromosomal Microarray Analysis. *Pediatrics* 2012; 130(5):e1085-e1095.
11. Riggs ER, Wain KE, Riethmaier D, et al. Chromosomal Microarray Impacts Clinical Management. *Clin Genet* 2013; 85(2):147-153.

12. Henderson LB, Applegate CD, Wohler E, et al. The impact of chromosomal microarray on clinical management: a retrospective analysis. *Genet Med* 2014; 16(9):657-664.

13. Tao VQ, Chan KYK, Chu VWY, et al. The Clinical Impact of Chromosomal Microarray on Paediatric Care in Hong Kong. *PLoS ONE* 2014; 9(10):e109629.

14. Mason-Suares H, Kim W, Grimmett L, et al. Density matters: comparison of array platforms for detection of copy-number variation and copy-neutral abnormalities. 2013; 15(9):706-712.

15. Matsunami N, Hadley D, Hensel CH et al. Identification of rare recurrent copy number variants in high-risk autism families and their prevalence in a large ASD population. *PLoS One* 2013; 8(1):e52239.

16. Sebat J, Lakshmi B, Malhotra D, et al. Strong Association of De Novo Copy Number Mutations with Autism. *Science* 2007; 316(5823):445–449.

17. Marshall CR, Noor A, Vincent JB, et al. Structural Variation of Chromosomes in Autism Spectrum Disorder. *Am J Hum Genet* 2008; 82(2):477–488.

18. Christian SL, Brune CW, Sudi J, et al. Novel Submicroscopic Chromosomal Abnormalities Detected in Autism Spectrum Disorder. *Biol Psychiatry* 2008 63(12):1111–1117.

19. Glessner JT, Wang K, Cai G, et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 2009 459(7246):569–573.

20. Bucan M, Abrahams BS, Wang K, et al. Genome-Wide Analyses of Exonic Copy Number Variants in a Family-Based Study Point to Novel Autism Susceptibility Genes. *PLoS Genet* 2009 5(6):e1000536.

21. Pinto D, Pagnamenta AT, Klei L, et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* 2010; 466(7304):368–372.

22. Szatmari P, Paterson AD, Zwaigenbaum L, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007; 39(3):319–328.

23. Weiss LA, Shen Y, Korn JM, et al. Association between Microdeletion and Microduplication at 16p11.2 and Autism. *N Engl J Med* 2008; 358(7):667–675.

24. Jacquemont M-L, Sanlaville D, Redon R, et al. Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. *J Med Genet* 2006; 43(11):843–849.

25. Kearney HM, South ST, Wolff DJ, et al. American College of Medical Genetics recommendations for the design and performance expectations for clinical genomic copy number microarrays intended for use in the postnatal setting for detection of constitutional abnormalities. *Genet Med* 2011; 13(7):676-679.

26. Cerebral creatine deficiency syndrome 2. http://omim.org/entry/612736.

27. Cerebral creatine deficiency syndrome 3. http://omim.org/entry/612718.

28. South S. T., *et al.* ACMG Standards and Guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: revision 2013. *Genetics in Medicine* **15**(11):901–909 (2013).

29. MacDonald JR, Ziman R, Yuen RK, et al. The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res* 2014; 42(Database issue):D986-992.

30. Committee on Bioethics, Committee on Genetics, and, The American College of Medical Genetics and, Genomics Social, Ethical, and Legal Issues Committee: Ethical and Policy Issues in Genetic Testing and Screening of Children. Pediatrics Mar 2013, 131 (3) 620-622; **DOI:** 10.1542/peds.2012-3680

31. Bernardini L, Alesi V, Loddo S, et al. High-resolution SNP arrays in mental retardation diagnostics: how much do we gain? *Eur J Human Genet* 2010; 18(2):178-195.

32. McGrew SG, Peters BR, Crittendon JA, Veenstra-VanderWeele J. Diagnostic Yield of Chromosomal Microarray Analysis in an Autism Primary Care Practice: Which Guidelines to Implement? *J Autism Dev Disord* 2012; 42(8):1582-1591.

33. Howell KB, Kornberg AJ, Harvey AS et al. High resolution chromosomal microarray in undiagnosed neurological disorders. *J Paediatr Child Health* 2013; 49(9):716-724.

34. Battaglia A, Doccini V, Bernardini L et al. Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *Eur J Paediatr Neurol* 2013; 17(6):589-599.

35. Miller DT, Adam MP, Aradhya S, et al. Consensus Statement: Chromosomal Microarray Is a First-Tier Clinical Diagnostic Test for Individuals with Developmental Disabilities or Congenital Anomalies. *Am J Hum Genet* 2010; 86(5):749-764.

36. Shen Y, Dies KA, Holm IA, et al. Clinical Genetic Testing for Patients with Autism Spectrum Disorders. 2010; *Pediatr* 2010; 125(4):e727-e735.

37. SNP & Variation Suite ™ (Version 8.x) [Software]. Bozeman, MT: Golden Helix, Inc. Available from http://www.goldenhelix.com.

38. Sdano MR, Vanzo RJ, Martin MM, et al. Clinical Utility of Chromosomal Microarray Analysis of DNA from Buccal Cells: Detection of Mosaicism in Three Patients. *J Genet Counsel* 2014; 23(6):922-927.

39. Martin MM, Vanzo RJ, Sdano MR, et al. Mosaic Deletion of 20pter Due to Rescue by Somatic Recombination. Am J Med Genet A. 2016 Jan;170A(1):243-8. doi: 10.1002/ajmg.a.37407. Epub 2015 Oct 5.

40. Ho KS, South ST, Lortz A, et al. Chromosomal microarray testing identifies a 4p terminal region associated with seizures in Wolf-Hirschhorn syndrome. J Med Genet 2016;0:1–8. doi:10.1136/jmedgenet-2015-103626

41. Jiang Y, Yuen, RKC, Jin, X, et al., Detection of Clinically Relevant Genetic Variants in Autism Spectrum Disorder by Whole Genome Sequencing, The American Journal of Human Genetics (2013), http://dx.doi.org/10.1016/j.ajhg.2013.06.012

42. Uddin M, Tammimies K, Pellecchia G, et al., Brain-expressed exons under purifying selection are enriched for de novo mutations in autism spectrum disorder. Nat. Genetics. Vol 46(7):742-49; 2014.