EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR Affymetrix® CytoScan® Dx Assay
DECISION SUMMARY

A. 510(k) Number:

k130313

B. Purpose for Submission:

De novo request for evaluation of automatic class III designation of the Affymetrix® CytoScan® Dx Assay

C. Measurand:

Genome-wide chromosomal copy number variations

D. Type of Test:

Chromosomal Microarray

E. Applicant:

Affymetrix, Inc.

F. Proprietary and Established Names:

Affymetrix® CytoScan® Dx Assay

G. Regulatory Information:

1. Regulation section:

   21 CFR 866.5920

2. Classification:

   Class II (special controls)

3. Product code:

   PFX -- System, Microarray-based, genome-wide, postnatal chromosomal abnormality detection

4. Panel:

   Immunology

H. Intended Use:

1. Intended use(s):

   CytoScan® Dx Assay is a qualitative assay intended for the postnatal detection of copy number variations (CNV) in genomic DNA obtained from peripheral whole blood in patients referred for chromosomal testing based on clinical presentation. CytoScan® Dx Assay is intended for the detection of CNVs associated with developmental delay, intellectual disability, congenital anomalies, or dysmorphic features. Assay results are intended to be used in conjunction with other clinical and diagnostic findings, consistent with professional standards of practice, including confirmation by alternative methods, parental evaluation, clinical genetic evaluation, and counseling, as appropriate. Interpretation of assay results is intended to be performed only by healthcare professionals, board certified in clinical cytogenetics or molecular genetics. The assay is intended to be used on the GeneChip® System 3000Dx and analyzed by Chromosome Analysis Suite Dx Software (ChAS Dx Software).
   This device is not intended to be used for standalone diagnostic purposes, pre-implantation or prenatal testing or screening, population screening, or for the detection of, or screening for, acquired or somatic genetic aberrations.

2. Indication(s) for use:

   Same as Intended use above.

3. Special conditions for use statement(s):

   For prescription use.

4. Special instrument requirements:
   GeneChip® System 3000Dx v.2 and with Chromosome Analysis Suite Dx Software v.1.0 (ChAS Dx Software).

I. Device Description:

The CytoScan Dx consists of five reagent modules, a microarray kit, and analysis software. The five reagent modules are:

1. MOD R L A, CytoScan® Dx Pre-PCR contains buffers, nucleotides, enzyme, and primers and adaptors for amplification;
2. MOD T E W, CytoScan® Dx Pre-PCR contains buffer and nuclease free water for amplification;
3. MOD F L H, CytoScan® Dx Post-PCR contains buffers, nucleotides, and enzyme for fragmentation, labeling and hybridization;
4. MOD S AH W PB, CytoScan® Dx Post-PCR contains buffers, nuclease free water, and purification beads for stain and array hold;
5. MOD E PW CytoScan® Dx Post-PCR contains buffer for elution and purification wash.

The CytoScan® Dx Post-PCR CytoScan® Dx Array kit, 6-pack is designed for 6 runs. The microarray contains approximately 2,696,550 functional markers, each of which is approximately 25 bases long.

ChAS Dx Analysis Software and Browser v1.0.0 analyzes CEL file microarray data.

J. Substantial Equivalence Information:

1. Predicate device name(s) and 510(k) number(s):

   Not applicable.

2. Comparison with predicate:

   Not applicable.

K. Standard/Guidance Document Referenced (if applicable):

Not applicable.

L. Test Principle:

CytoScan Dx Assay provides genome-wide coverage for the detection of chromosomal imbalances. The CytoScan Dx array contains approximately 2.7 million markers which are representative of DNA sequences distributed throughout the genome with spacing, on average, approximately 880 bases apart in genic regions, and approximately 1700 bases apart in non-genic regions. The majority of the markers (1.9 million) are non-polymorphic markers, which provide overall genomic coverage of relevant cytogenetic regions and are used for assessing copy number. Approximately 750,000 SNP markers on the array are included to maximize genomic coverage and to enable detection of homozygosity. Both the SNP and non-polymorphic markers are approximately 25 bp long.

CytoScan Dx Assay consists of the following steps: (1) gDNA is isolated from peripheral blood and the isolated gDNA is digested with the restriction enzyme Nsp1 to reduce genomic complexity; (2) The digested gDNA is ligated to Nsp1 adapters and amplified in a multiplex PCR reaction to produce optimized amplicons in the 200 1100 bp size range; (3) The amplified PCR products are purified and then randomly fragmented using DNAse I to generate species of 25 125 bp, which are optimal for hybridization to 25-mer markers; (4) Reaction intermediates are visualized by gel electrophoresis after the PCR and fragmentation steps to confirm proper size distribution; (5) The final DNA product is end-labeled by the addition of a modified biotinylated base and hybridized to CytoScan Dx Arrays; (6) The arrays are sequentially washed and stained with a combination of a streptavidin-coupled dye and a biotinylated anti-streptavidin antibody in GeneChip® Fluidics Station 450Dx v.2; (7) The washed arrays are scanned using GeneChip® Scanner 3000Dx v.2 to acquire the signal intensity from each marker.

Chromosome Analysis Suite (ChAS Dx) software is used to analyze and visualize microarray data. The signal intensity of the hybridized DNA from the patient sample is compared to a reference DNA, which is based on an average of over 400 samples. The ratio of patient sample to reference intensity is expressed as a log2 ratio, and represents the relative intensity for each marker. A discrete copy number value is computed from the relative intensity data, and is displayed as the marker copy number state. The noninteger copy number states are calculated and displayed as the smoothed signal track, which can used to support an interpretation of a mosaic gain or loss. The SNP marker A- and B-allele intensities are also visualized in the Allele Track, which can be used to confirm copy number variation regions. The allele tracks show 3 bands (AA, AB, BB) in normal diploid regions, 4 bands (AAA, AAB, ABB, BBB) in triploid regions, and 2 bands (A, B) in haploid regions. The SNP markers are also analyzed for long contiguous stretches of homozygosity, which are visualized in the loss of heterozygosity (LOH) track. The absence or loss of heterozygosity (AOH / LOH) is calculated as a region significantly devoid of heterozygous genotype calls.

CytoScan Dx reports the copy number state (loss, gain), copy number (i.e., 0, 1, 2, 3, or 4 or greater), and position/location of chromosomal segment copy number changes across the queried genome.

M. Performance Characteristics (if/when applicable):

1. Analytical performance:

   a. Precision/Reproducibility:

   Two reproducibility studies were conducted. The first was a site to site reproducibility study, and the second was a between-lot reproducibility study.

   In each study, the reproducibility of CytoScan Dx results was evaluated for both copy number state determination (gain or loss), and localization of each CNV based on overlapping size of the CNVs or overlapping marker number, for two criteria of agreement: 50% overlap and 80% overlap. For every sample replicate tested, every CNV detected in each replicate was analyzed for reproducibility with the other replicates by pairwise agreement (i.e., CNV 1 in replicate 1 was compared to CNV 1 in replicate 2, then CNV 1 in replicate 1 was compared to CNV 1 in replicate 3, and so on). A pair of replicates was considered to agree if the two CNVs in each pair compared, overlapped by at least 50% or 80% of the CNV length at a given location on the chromosome, or marker number, provided the copy number state (gain/loss) was the same. Because pairwise analysis considers agreement between two CNVs for all combined pairs of replicates (e.g., a CNV detected in 9 replicates will have a total of 36 paired comparisons), in the scenario where two of 9 replicates had no CNV result, the 2 no calls were considered to be agreed in the pairwise replicate agreement calculation. Positive percent agreement (PPA) measures the pairwise agreement conditional on a replicate being gain or loss. Call rate is calculated as the average of the percent of replicates that call each CNV. Reproducibility was also assessed for agreement of copy number (0, 1, 2, 3, 4) without regard to size or marker number, and assessed for localization based on endpoint agreement between each CNV

replicate. To determine the precision estimate for localizing a CNV, % coefficient of variance of CNV length, median % absolute endpoint deviation, and standard deviation of the endpoints (both left and right endpoint) were calculated. The results in the tables below are grouped by size range and marker number. Results are shown including, and excluding Affymetrix-defined hypervariable regions. A more detailed description of the statistical methods for assessing the reproducibility variables is shown below:

Description of variables:

(1) Percent (%) overlap: For the criteria shown (50% overlap or 80% overlap between CNVs), the overlap data is the average of results for all pairwise replicates for the range of CNVs listed (row).
(2) Pairwise replicate agreement was determined by examining all pairs of replicates for overlap at 50% or 80% CNV length, summarized (for 9 replicates) as follows:

$$\sum_{i=1}^{9}\sum_{j>i}^{9} Pr\{rep_i = rep_j\}/36$$

Two replicates are considered to be agreed (or equal) if the CNVs overlap at least 50% (or 80%) and the copy number states are identical. Two replicates of no calls are considered to be agreed.
(3) Positive percent agreement (PPA) measures the pairwise agreement conditional on a replicate having the gain or loss, summarized (for 9 replicates) as follows:

$$\sum_{i=1}^{9}\sum_{j=1\&i\neq j}^{9} Pr\{rep_i = rep_j | rep_i = gain\ or\ loss\}/\ \ 8 * \sum_{i=1}^{9} \Pr\{rep_{i=}\ gain\ or\ loss\}$$

Two replicates are considered to be agreed (or equal) if the CNVs overlap at least 50% (or 80%) and the copy number states are identical.
(4) Call rate is calculated as the average of the percentage of replicates that call each CNV.
(5) Length, % Coefficient of Variance (%CV): The %CV of each CNV, calculated for both size in terms of kilobases (kb) and number of markers. For clarity, CV is calculated as the standard deviation divided by the mean length across the replicates (%CV = SD/Mean). The mean, minimum, median, and maximum for all CNVs, or stratified groups of CNVs, are presented.
(6) Median % Absolute Endpoint Deviation: For each CNV, the median left and right endpoint was determined. For each replicate CNV, the distance from the median endpoint was calculated for the left and right endpoints (DL and DR, respectively). The combined endpoint distance, DL+DR, was calculated. The fractional endpoint error is the combined endpoint distance divided by the CNV length, expressed as a percentage. For a CNV with replicate measurements, the CNV endpoints are e_L and e_R, the median left and right endpoints are defined as $\tilde{e_L}$ and $\tilde{e_R}$, respectively. For each replicate measurement, the absolute distance of the CNV endpoints from the medians, $D_L = |e_L - \tilde{e_L}|$ and $D_R = |e_R - \tilde{e_R}|$.

$$\%\ Endpoint\ Error = \frac{D_L + D_R}{e_R - e_L + 1} \times 100$$

$$Median\ \%\ Absolute\ endpoint\ Deviation = \%\ \widetilde{Endpoint\ Error}$$

(7) Standard deviation of identifying each endpoint is the standard deviation of the endpoint, separately for left and right endpoint.

Study 1. Site-to-site reproducibility:

To assess the reproducibility metrics described above across multiple sites, a study was performed on 93 genomic DNA samples (48 purified from blood, 44 purified from cell lines obtained from Coriell and the World Health Organization, 1 control sample from American Tissue Culture Collection) representing different CNV sizes and gains and losses. These samples were run by 2 operators at each of 3 sites across 3 non-consecutive days. Samples contained either gains or losses which covered a total of 50.8% of the genome with at least 1 CNV region on every chromosome with 42.3% of the CNVs detected as gains and 20.5% detected as losses.

A total of 9 chromosomal regions were defined by the sponsor as hypervariable regions because they contain genetic components of the immune system or members of gene families that have been shown to be associated with extensive copy number variations and/or rearrangement (e.g., the olfactory receptor family genes at 1q44 and 11q11), segmental duplication (17q21), and nonfunctional pseudogenes (ADAM3A at 8p11.2). None of these regions overlap with the CNV target regions for constitutional genomic disorders identified by the International Standards for Cytogenomic Arrays Consortium (ISCA). These excluded regions add up to approximately 3.6 Mb in size, or approximately 0.1% of the entire human genome (~3×10^9 bp). Only those CNV regions that were contained completely within the boundaries of the regions defined in the table above were removed. CNV regions containing any marker outside the boundaries of defined regions were still included in the analysis. Performance of the device including and excluding these 9 hypervariable regions were both evaluated. These 9 hypervariable regions are described as the limitations of the device and listed below.

Table 1. List of Affymetrix-defined Hypervariable Regions of the Human Genome (hg19 build).

Chromosomal Region	Boundaries	Size of Region (in bases)
1q44	248,681,754 – 248,835,053	153,300 bp
5q35.3	180,376,952 – 180,432,918	55,967 bp
7p14.1	38,273,345 - 38,419,181	145,837 bp
8p11.22	39,226,075 - 39,390,890	164,816 bp
11q11	55,347,529 - 55,481,854	134,326 bp
14q11.2	22,329,745 - 23,005,312	675,568 bp
14q32.33	106,035,612 - 107,297,169	1,261,558 bp
17q21.31	44,107,114 - 44,854,730	747,617 bp
22q11.22	22,992,312 - 23,260,235	267,924 bp

The reproducibility results demonstrated that when agreement across replicates was considered for CNVs that overlapped by 50% or more of the CNV length, the overall pairwise replicate agreement was 79.4% for all CNVs; range 65.7 to 100%. The agreement was 75.5% for all gains and 82.4% for all losses). For marker number, the overall pairwise replicate agreement is 82.2%; range 70.2% to 82.2%. The agreement was 79.8% for all gains and 83.9% for all losses. The results are shown in Tables 2B and 2A below, respectively.

The following tables 2A-2D show the results of the study to evaluate reproducibility of the device.

Table 2A. Reproducibility of CNVs Grouped by State (Gains and Losses) and Marker Number Based on Call Rate, Pairwise Agreement between Replicates and Positive Percent Agreement (PPA) for Two Criteria (50% and 80% Overlap) in All Regions in the Site-to-Site Study.

				Pairwise Replicate Agreement*		**PPA***	
State	**CNV Range (Markers)**	**# CNVs**	**Call Rate***	**Overlap**			
				50%	**80%**	**50%**	**80%**
Gain	50-75	230	49.7	70.8	68.7	72.7	68.7
	75-100	150	68.3	70.2	54.7	77.4	54.3
	100-150	131	70.2	81.5	71.6	86.5	73.0
	150-200	68	82.5	87.8	81.7	91.0	83.8
	200-300	72	92.4	91.0	81.1	92.4	81.9
	300-400	8	100.0	100.0	98.9	100.0	98.9
	400-1000	22	100.0	97.5	95.9	97.5	95.9
	1000-3000	22	100.0	100.0	100.0	100.0	100.0
	3000-5000	7	100.0	100.0	100.0	100.0	100.0
	5000+	60	100.0	100.0	100.0	100.0	100.0
	Total	770	71.4	79.8	72.5	85.3	75.3
Loss	25-50	430	48.8	79.5	76.4	75.4	68.0
	50-75	204	74.8	77.7	69.0	84.2	72.9
	75-100	107	81.9	87.5	80.7	92.1	83.8
	100-150	58	82.4	85.2	78.8	87.6	79.9
	150-200	24	86.4	82.4	69.4	85.1	70.2
	200-300	26	96.1	92.5	85.4	93.9	86.7
	300-400	12	100.0	94.2	83.0	94.2	83.0
	400-1000	37	100.0	95.2	77.4	95.2	77.4
	1000-3000	26	100.0	100.0	99.9	100.0	99.9
	3000-5000	12	100.0	100.0	100.0	100.0	100.0
	5000+	41	100.0	100.0	100.0	100.0	100.0
	Total	977	68.9	83.9	78.7	87.0	79.3
All	Total	1747	70.0	82.2	76.1	86.3	77.5

*refer to Description of variables section above for detail

Table 2B. Reproducibility of CNVs Grouped by State (Gains and Losses) and Sizes (in kb) Based on Call Rate, Pairwise Agreement between Replicates and Positive Percent Agreement (PPA) for Two Criteria (50% and 80% Overlap) in All Regions in the Site-to-Site Study.

				Pairwise Replicate Agreement*		PPA*	
State	CNV Range (kb)	# CNVs	Call Rate*	Overlap			
				50%	80%	50%	80%
Gain	50-75	99	50.4	71.4	67.2	70.4	62.1
	75-100	98	63.7	70.7	62.6	74.4	63.1
	100-150	149	64.9	70.8	63.6	76.3	65.2
	150-200	101	62.3	65.7	47.1	64.9	36.4
	200-300	100	78.2	73.3	67.5	78.1	70.8
	300-400	27	68.6	70.3	65.3	71.5	64.2
	400-1000	94	88.8	86.8	71.3	89.5	70.7
	1000-3000	39	86.6	87.4	78.5	90.5	80.2
	3000-5000	3	100.0	100.0	100.0	100.0	100.0
	5000+	60	100.0	99.1	98.4	99.1	98.4
	Total	770	71.4	75.5	66.5	79.4	67.0
Loss	25-50	351	53.7	78.8	73.9	78.5	68.8
	50-75	177	65.6	77.9	70.1	76.7	63.2
	75-100	88	66.9	77.4	68.5	72.4	57.1
	100-150	133	80.9	82.4	76.3	88.0	80.7
	150-200	30	73.6	82.5	80.1	85.5	82.2
	200-300	29	86.4	81.8	62.9	84.8	63.0
	300-400	35	88.3	89.8	69.5	91.0	68.1
	400-1000	43	86.9	89.1	74.4	91.9	74.9
	1000-3000	35	98.5	96.3	92.8	96.8	93.3
	3000-5000	16	60.0	84.2	84.2	86.9	86.9
	5000+	40	100.0	100.0	99.9	100.0	99.9
	Total	977	68.9	82.4	75.9	84.7	75.1
All	Total	1747	70.0	79.4	71.9	82.4	71.5

*refer to Description of variables section above for detail

Table 2C. Reproducibility of CNVs Grouped by State (Gains and Losses) and Marker Number Based on Call Rate, Pairwise Agreement between Replicates and Positive Percent Agreement (PPA) for Two Criteria (50% and 80% Overlap) in Regions Excluding Affymetrix-defined Hypervariable Regions in the Site-to-Site Study.

				Pairwise Replicate Agreement*		PPA*	
State	CNV Range (Markers)	# CNVs	Call Rate*	Overlap			
				50%	80%	50%	80%
Gain	50-75	166	45.4	75.1	73.6	74.7	71.6
	75-100	34	57.0	77.9	73.7	75.5	66.5
	100-150	75	63.6	85.1	76.9	88.3	76.4
	150-200	48	78.7	90.8	88.9	94.0	91.8
	200-300	36	85.6	91.3	87.3	94.0	89.3
	300-400	8	100.0	100.0	98.9	100.0	98.9
	400-1000	22	100.0	97.5	95.9	97.5	95.9
	1000-3000	22	100.0	100.0	100.0	100.0	100.0
	3000-5000	7	100.0	100.0	100.0	100.0	100.0
	5000+	60	100.0	100.0	100.0	100.0	100.0
	Total	478	68.9	85.4	82.5	89.1	85.0
Loss	25-50	387	49.8	81.0	78.1	77.4	70.6
	50-75	74	71.5	84.8	82.8	89.0	86.4
	75-100	67	92.1	95.5	89.0	97.2	90.1
	100-150	42	82.7	87.0	82.8	88.5	83.4
	150-200	19	86.9	84.9	74.9	87.3	75.8
	200-300	16	99.5	99.4	96.7	99.6	96.9
	300-400	8	100.0	100.0	94.9	100.0	94.9
	400-1000	17	100.0	100.0	99.8	100.0	99.8
	1000-3000	26	100.0	100.0	99.9	100.0	99.9
	3000-5000	12	100.0	100.0	100.0	100.0	100.0
	5000+	41	100.0	100.0	100.0	100.0	100.0
	Total	709	67.9	87.3	84.5	89.8	85.6
All	Total	1187	68.3	86.6	83.8	89.5	85.4

*refer to Description of variables section above for detail

Table 2D. Reproducibility of CNVs Grouped by State (Gains and Losses) and Sizes (in kb) Based on Call Rate, Pairwise Agreement between Replicates and Positive Percent Agreement (PPA) for Two Criteria (50% and 80% Overlap) in Regions Excluding Affymetrix-defined Hypervariable Regions in the Site-to-Site Study.

				Pairwise Replicate Agreement*		PPA*	
State	CNV Range (kb)	# CNVs	Call Rate*	Overlap			
				50%	80%	50%	80%
Gain	50-75	79	49.5	75.4	73.4	74.2	70.1
	75-100	55	58.3	75.9	73.0	75.0	70.9
	100-150	89	66.8	79.8	74.0	84.2	75.7
	150-200	48	43.5	77.6	74.0	71.5	63.0
	200-300	56	80.3	90.9	89.4	93.9	92.1
	300-400	17	71.1	80.0	74.8	81.2	73.9
	400-1000	32	73.8	78.7	76.2	78.6	74.4
	1000-3000	39	86.6	87.4	78.5	90.5	80.2
	3000-5000	3	100.0	100.0	100.0	100.0	100.0
	5000+	60	100.0	99.1	98.4	99.1	98.4
	Total	478	68.9	82.6	79.2	85.1	80.4
Loss	25-50	286	54.3	81.0	77.2	80.9	73.4
	50-75	115	65.4	85.6	81.8	84.8	77.8
	75-100	54	59.6	81.2	77.2	74.0	65.8
	100-150	68	82.2	93.1	89.3	95.3	91.0
	150-200	26	75.7	85.3	84.2	87.8	86.4
	200-300	21	85.6	87.4	73.6	89.9	73.7
	300-400	17	75.8	83.5	76.4	84.4	75.1
	400-1000	31	82.0	87.6	78.8	91.3	80.5
	1000-3000	35	98.5	96.3	92.8	96.8	93.3
	3000-5000	16	60.0	84.2	84.2	86.9	86.9
	5000+	40	100.0	100.0	99.9	100.0	99.9
	Total	709	67.9	86.4	82.7	88.4	82.8
All	Total	1187	68.3	85.0	81.4	87.2	81.8

*refer to Description of variables section above for detail

CytoScan Dx Assay determines the copy number of each identified CNV as being 0, 1, 2, 3 or 4. For copy number agreement, reproducibility of the numerical copy number was evaluated by performing all pairwise comparisons for replicate measurements of each CNV. Each pairwise comparison with exactly matching copy number was scored as concordant. The summary table 2E presents the fraction of the pairwise comparisons scored as percent agreement, for each size range shown. Agreement = Number of Matching Pairwise Comparisons * 100 / Total Number of Pairwise comparisons, where matching is defined as having identical copy numbers for both of the pair. Copy number agreement was assessed without regard to location, size or marker number agreement. Note that this is a different calculation of agreement because it relies on identical copy number, not just copy number state.

Table 2E. Copy Number Reproducibility in the Site-to-Site Reproducibility Study.

		All Regions (Markers)		All Regions (kb)		Regions Excluding Affymetrix-Defined Hypervariable Regions (Markers)		Regions Excluding Affymetrix-Defined Hypervariable Regions (kb)	
		CNV (N)	% Agreement	CNV (N)	% Agreement	CNV (N)	% Agreement	CNV (N)	% Agreement
Gain	50-75	230	69.6%	99	71.1%	166	73.5%	79	74.3%
	75-100	150	70.7%	98	78.0%	34	78.6%	55	87.5%
	100-150	131	81.9%	149	71.2%	75	86.3%	89	78.7%
	150-200	68	90.1%	101	77.5%	48	91.6%	48	79.7%
	200-300	72	93.5%	100	80.4%	36	90.5%	56	90.5%
	300-400	8	100.0%	27	79.7%	8	100.0%	17	86.7%
	400-1000	22	98.1%	94	88.3%	22	98.1%	32	82.0%
	1000-3000	22	97.4%	39	89.7%	22	97.4%	39	89.7%
	3000-5000	7	100.0%	3	100.0%	7	100.0%	3	100.0%
	5000+	60	93.6%	60	93.6%	60	93.6%	60	93.6%
	All	770	79.6%	770	79.6%	478	84.3%	478	84.3%
Loss	25-50	430	79.3%	351	78.8%	387	81.0%	286	81.1%
	50-75	204	79.5%	177	82.7%	74	86.3%	115	87.7%
	75-100	107	88.0%	88	87.0%	67	96.2%	54	87.4%
	100-150	58	90.7%	133	85.2%	42	93.0%	68	94.2%
	150-200	24	90.5%	30	86.3%	19	91.8%	26	89.0%
	200-300	26	97.0%	29	89.2%	16	99.6%	21	92.2%
	300-400	12	100.0%	35	95.5%	8	100.0%	17	90.5%
	400-1000	37	100.0%	43	92.4%	17	100.0%	31	89.5%
	1000-3000	26	100.0%	35	98.9%	26	100.0%	35	98.9%
	3000-5000	12	100.0%	16	84.2%	12	100.0%	16	84.2%
	5000+	41	100.0%	40	100.0%	41	100.0%	40	100.0%
	All	977	85.1%	977	85.1%	709	87.9%	709	87.9%
All		1747	82.7%	1747	82.7%	1187	86.6%	1187	86.6%

*Agreement = Number of Matching Pairwise Comparisons * 100 / Total Number of Pairwise comparisons, where matching is defined as having identical copy numbers for both of the pair. Copy number agreement was assessed without regard to location, size or marker number agreement.

For endpoint analysis, only those CNVs detected were assessed for endpoint agreement (i.e.,no calls in replicates could not be included). The CNVs need to have same copy number state (gain, loss) in order to be included in the endpoint agreement calculation. Endpoint agreement is assessed by median % absolute endpoint deviation, standard deviation of left endpoint, and standard deviation of right endpoint, as indicated in the Description of variables section. The results are shown in Tables 2F-2I.

Table 2F. Reproducibility of CNV Length and Endpoints (in Markers) in All Regions in the Site-to-Site Study.*

Gain/Loss	CNV Range (Markers)	N	% CV CNV Length Mean (Min, Median, Max)	Average % Overlap	Median % Absolute Endpoint Deviation Mean (Min, Median, Max)	SD Left Endpoint (M) Mean (Min, Median, Max)	SD Right Endpoint (M) Mean (Min, Median, Max)
Gain	50-75	157	5.1 (0.0, 4.1, 30.1)	68.5	0.03 (0.00, 0.01, 0.22)	1.8 (0.0, 1.2, 12.5)	2.2 (0.0, 1.6, 11.3)
	75-100	136	11.6 (0.0, 11.4, 29.9)	63.8	0.06 (0.00, 0.04, 0.30)	8.1 (0.0, 8.4, 27.9)	4.6 (0.0, 2.0, 21.6)
	100-150	111	11.7 (0.0, 7.0, 66.2)	76.9	0.06 (0.00, 0.02, 0.63)	10.1 (0.0, 3.0, 78.1)	6.2 (0.0, 2.1, 48.8)
	150-200	59	11.7 (0.3, 4.5, 54.2)	85.0	0.05 (0.00, 0.01, 0.47)	15.1 (0.0, 3.3, 107.5)	9.8 (0.0, 2.8, 93.4)
	200-300	68	9.7 (0.0, 5.9, 76.5)	86.1	0.05 (0.00, 0.02, 0.77)	12.5 (0.0, 5.4, 79.3)	13.2 (0.0, 9.6, 166.9)
	300-400	8	1.9 (0.0, 1.0, 7.9)	98.1	0.01 (0.00, 0.01, 0.02)	1.4 (0.0, 1.3, 3.1)	5.7 (0.0, 2.6, 24.1)
	400-1000	22	2.7 (0.0, 0.5, 38.9)	95.8	0.02 (0.00, 0.00, 0.28)	12.5 (0.0, 1.4, 210.1)	10.3 (0.0, 1.9, 152.7)
	1000-3000	22	0.3 (0.0, 0.2, 1.4)	99.3	0.00 (0.00, 0.00, 0.01)	3.0 (0.0, 1.4, 15.6)	3.2 (0.0, 1.5, 34.1)
	3000-5000	7	0.3 (0.0, 0.1, 1.1)	99.5	0.00 (0.00, 0.00, 0.00)	4.3 (0.0, 3.1, 11.3)	7.7 (0.0, 0.5, 41.8)
	5000+	60	0.0 (0.0, 0.0, 0.5)	99.7	0.00 (0.00, 0.00, 0.00)	3.1 (0.0, 0.9, 55.0)	2.0 (0.0, 0.6, 25.5)
	Total	650	7.8 (0, 4.3, 76.5)	76.1	0.04 (0.00, 0.01, 0.77)	7.4 (0.0, 2.3, 210.1)	5.6 (0.0, 2.0, 166.9)
Loss	25-50	282	6.9 (0.0, 5.4, 63.7)	77.3	0.04 (0.00, 0.02, 0.48)	1.4 (0.0, 1.0, 18.4)	1.8 (0.0, 0.9, 26.8)
	50-75	186	13.1 (0.0, 8.2, 73.1)	73.3	0.08 (0.00, 0.03, 0.71)	3.5 (0.0, 1.7, 22.4)	6.2 (0.0, 3.5, 42.4)
	75-100	101	8.2 (0.0, 5.2, 33.1)	82.9	0.04 (0.00, 0.01, 0.30)	3.3 (0.0, 1.7, 22.4)	4.7 (0.0, 2.8, 22.6)
	100-150	51	10.8 (0.0, 1.9, 66.0)	82.9	0.06 (0.00, 0.01, 0.53)	8.6 (0.0, 1.4, 71.6)	6.5 (0.0, 1.6, 70.5)
	150-200	22	14.2 (0.0, 5.4, 57.7)	78.6	0.04 (0.00, 0.01, 0.23)	17.3 (0.0, 3.0, 109.7)	10.5 (0.0, 3.5, 78.9)
	200-300	26	14.7 (0.0, 6.8, 68.1)	90.5	0.07 (0.00, 0.02, 0.33)	27.8 (0.0, 5.8, 136.0)	8.4 (0.0, 1.7, 46.9)
	300-400	12	8.3 (0.3, 1.1, 62.8)	90.0	0.04 (0.00, 0.00, 0.41)	19.0 (0.0, 1.4, 204.9)	10.7 (0.0, 1.8, 63.9)
	400-1000	37	8.2 (0.0, 3.2, 50.0)	87.6	0.04 (0.00, 0.02, 0.26)	41.5 (0.0, 12.2, 216.5)	6.5 (0.0, 5.2, 21.4)
	1000-3000	26	0.5 (0.0, 0.2, 6.1)	99.1	0.00 (0.00, 0.00, 0.01)	5.1 (0.0, 0.9, 42.5)	5.1 (0.0, 2.0, 65.9)
	3000-5000	12	0.1 (0.0, 0.0, 0.1)	99.3	0.00 (0.00, 0.00, 0.00)	1.1 (0.0, 0.7, 3.2)	1.4 (0.0, 1.1, 5.3)
	5000+	41	0.0 (0.0, 0.0, 0.3)	99.8	0.00 (0.00, 0.00, 0.00)	1.5 (0.0, 1.1, 6.2)	1.8 (0.0, 0.5, 21.9)
	Total	796	8.6 (0.0, 4.8, 73.1)	81.1	0.05 (0.00, 0.01, 0.71)	6.2 (0.0, 1.4, 216.5)	4.4 (0.0, 1.5, 78.9)
All	Total	1446	8.3 (0.0, 4.6, 76.5)	79.0	0.04 (0.00, 0.01, 0.77)	6.7 (0.0, 1.7, 216.5)	4.9 (0.0, 1.8, 166.9)

*refer to Description of variables section above for detail

Table 2G. Reproducibility of CNV Length and Endpoints (in kb) in All Regions in the Site-to-Site Study.*

			% CV CNV length	Average % Overlap	Median % Absolute Endpoint Deviation	SD Left Endpoint (kb)	SD Right Endpoint (kb)
Gain/ Loss	CNV Range (kb)	N	Mean (Min, Median, Max)		Mean (Min, Median, Max)		
Gain	50-75	67	9.1 (0.0, 3.9, 54.8)	69.6	0.04 (0.00, 0.00, 0.37)	5.1 (0.0, 0.9, 40.2)	2.9 (0.0, 1.1, 14.0)
	75-100	77	11.7 (0.0, 8.9, 55.6)	68.6	0.06 (0.00, 0.02, 0.56)	7.6 (0.0, 4.4, 46.8)	7.3 (0.0, 5.7, 52.7)
	100-150	130	13.8 (0.0, 6.4, 65.8)	67.4	0.07 (0.00, 0.00, 0.59)	13.6 (0.0, 5.5, 63.8)	7.0 (0.0, 2.6, 61.4)
	150-200	76	35.9 (0.0, 29.7, 137.4)	62.6	0.13 (0.00, 0.09, 0.62)	31.6 (0.0, 31.5, 91.7)	37.2 (0.0, 16.6, 253.5)
	200-300	89	25.5 (0.0, 6.4, 118.6)	70.8	0.07 (0.00, 0.01, 0.68)	22.8 (0.0, 9.1, 238.5)	43.5 (0.0, 5.5, 269.0)
	300-400	23	41.9 (0.0, 14.1, 144.4)	69.2	0.42 (0.00, 0.02, 2.23)	24.8 (0.0, 5.1, 298.4)	123.8 (0.0, 15.2, 488.7)
	400-1000	90	16.2 (0.0, 9.8, 100.6)	79.3	0.08 (0.00, 0.04, 1.21)	36.7 (0.0, 16.8, 358.2)	71.8 (0.0, 44.7, 695.3)
	1000-3000	35	12.1 (0.0, 1.4, 61.7)	83.8	0.06 (0.00, 0.00, 0.44)	106.2 (0.0, 4.5, 538.4)	83.8 (0.0, 3.2, 1011.9)
	3000-5000	3	0.8 (0.1, 0.2, 2.1)	98.9	0.00 (0.00, 0.00, 0.00)	20.1 (2.2, 6.9, 51.1)	20.4 (0.0, 4.8, 56.5)
	5000+	60	3.0 (0.0, 0.0, 146.8)	98.6	0.00 (0.00, 0.00, 0.03)	183.1 (0.0, 0.3, 10583.8)	352.8 (0.0, 0.0, 11297.1)
	Total	650	17.4 (0.0, 7.2, 146.8)	72.4	0.08 (0.00, 0.01, 2.23)	39.6 (0.0, 5.4, 10583.8)	64.4 (0.0, 4.7, 11297.1)
Loss	25-50	252	9.5 (0.0, 6.6, 60.3)	76.3	0.04 (0.00, 0.00, 0.66)	2.0 (0.0, 0.1, 22.9)	2.8 (0.0, 1.3, 24.3)
	50-75	145	20.3 (0.0, 12.7, 71.4)	75.4	0.08 (0.00, 0.02, 0.68)	8.4 (0.0, 4.0, 43.4)	6.1 (0.0, 3.6, 33.8)
	75-100	69	25.7 (0.0, 12.5, 115.9)	76.2	0.12 (0.00, 0.03, 0.50)	18.2 (0.0, 8.2, 116.5)	7.2 (0.0, 3.3, 47.0)
	100-150	121	11.7 (0.0, 8.1, 57.3)	79.5	0.03 (0.00, 0.00, 0.38)	8.0 (0.0, 1.6, 78.5)	8.0 (0.0, 3.8, 40.3)
	150-200	24	10.6 (0.2, 2.8, 104.1)	80.6	0.03 (0.00, 0.00, 0.36)	14.6 (0.0, 3.5, 163.9)	3.7 (0.0, 1.8, 43.9)
	200-300	27	16.2 (0.0, 14.3, 53.4)	76.2	0.07 (0.00, 0.02, 0.42)	29.4 (0.0, 13.1, 128.7)	15.0 (0.0, 3.5, 79.3)
	300-400	32	12.9 (0.0, 7.7, 94.8)	82.2	0.03 (0.00, 0.02, 0.17)	35.3 (0.0, 22.4, 293.4)	13.0 (0.0, 3.9, 184.3)
	400-1000	40	17.2 (0.0, 5.8, 167.4)	82.4	0.10 (0.00, 0.01, 1.26)	55.2 (0.0, 8.9, 466.3)	54.7 (0.0, 8.3, 1442.1)
	1000-3000	35	3.2 (0.0, 0.4, 33.4)	94.5	0.01 (0.00, 0.00, 0.10)	21.3 (0.0, 1.1, 342.6)	28.5 (0.0, 4.7, 202.4)
	3000-5000	11	0.3 (0.0, 0.0, 2.6)	83.9	0.00 (0.00, 0.00, 0.00)	11.4 (0.0, 0.0, 117.9)	4.0 (0.0, 0.0, 24.6)
	5000+	40	0.5 (0.0, 0.0, 9.4)	99.6	0.00 (0.00, 0.00, 0.06)	8.1 (0.0, 0.6, 164.7)	50.9 (0.0, 0.1, 1484.5)
	Total	796	13.1 (0.0, 7.0, 167.4)	79.8	0.05 (0.00, 0.00, 1.26)	12.1 (0.0, 1.6, 466.3)	11.6 (0.0, 2.3, 1484.5)
All	Total	1446	15.1 (0.0, 7.1, 167.4)	76.6	0.07 (0.00, 0.00, 2.23)	24.5 (0.0, 2.9, 10583.8)	35.3 (0.0, 3.0, 11297.1)

*refer to Description of variables section above for detail

Table 2H. Reproducibility of CNV Length and Endpoints (in Markers) in Regions Excluding Affymetrix-defined Hypervariable Regions in the Site-to-Site Study.*

			% CV CNV Length	Average % Overlap	Median % Absolute Endpoint Deviation	SD Left Endpoint (M)	SD Right Endpoint (M)
Gain/ Loss	**CNV Range (Markers)**	**N**	**Mean (Min, Median, Max)**	**Mean (Min, Median, Max)**			
Gain	50-75	97	4.3 (0.0, 2.6, 30.1)	73.1	0.03 (0.00, 0.02, 0.22)	1.9 (0.0, 1.4, 12.5)	1.5 (0.0, 1.0, 11.3)
	75-100	23	7.5 (0.0, 5.4, 29.7)	75.6	0.05 (0.00, 0.02, 0.26)	3.9 (0.0, 2.0, 13.5)	4.2 (0.0, 2.1, 21.6)
	100-150	56	11.7 (0.0, 3.4, 66.2)	81.3	0.08 (0.00, 0.02, 0.63)	10.8 (0.0, 2.4, 78.1)	6.8 (0.0, 2.6, 48.8)
	150-200	39	6.9 (0.3, 1.7, 54.2)	89.0	0.03 (0.00, 0.01, 0.23)	8.4 (0.0, 2.0, 65.1)	5.6 (0.0, 1.5, 93.4)
	200-300	32	7.9 (0.0, 2.4, 76.5)	88.1	0.05 (0.00, 0.01, 0.77)	9.4 (0.0, 2.0, 79.3)	11.4 (0.0, 2.4, 166.9)
	300-400	8	1.9 (0.0, 1.0, 7.9)	98.1	0.01 (0.00, 0.01, 0.02)	1.4 (0.0, 1.3, 3.1)	5.7 (0.0, 2.6, 24.1)
	400-1000	22	2.7 (0.0, 0.5, 38.9)	95.8	0.02 (0.00, 0.00, 0.28)	12.5 (0.0, 1.4, 210.1)	10.3 (0.0, 1.9, 152.7)
	1000-3000	22	0.3 (0.0, 0.2, 1.4)	99.3	0.00 (0.00, 0.00, 0.01)	3.0 (0.0, 1.4, 15.6)	3.2 (0.0, 1.5, 34.1)
	3000-5000	7	0.3 (0.0, 0.1, 1.1)	99.5	0.00 (0.00, 0.00, 0.00)	4.3 (0.0, 3.1, 11.3)	7.7 (0.0, 0.5, 41.8)
	5000+	60	0.0 (0.0, 0.0, 0.5)	99.7	0.00 (0.00, 0.00, 0.00)	3.1 (0.0, 0.9, 55.0)	2.0 (0.0, 0.6, 25.5)
	Total	366	5.1 (0.0, 1.6, 76.5)	83.2	0.03 (0.00, 0.00, 0.77)	5.7 (0.0, 1.7, 210.1)	4.7 (0.0, 1.4, 166.9)
Loss	25-50	250	6.3 (0.0, 4.9, 63.7)	78.8	0.03 (0.00, 0.02, 0.48)	1.5 (0.0, 1.1, 18.4)	1.5 (0.0, 0.8, 26.8)
	50-75	61	7.5 (0.0, 2.9, 73.1)	83.0	0.05 (0.00, 0.02, 0.71)	1.5 (0.0, 0.7, 11.2)	3.7 (0.0, 1.4, 42.4)
	75-100	64	6.4 (0.0, 3.1, 33.1)	91.1	0.02 (0.00, 0.00, 0.30)	2.7 (0.0, 1.3, 22.4)	3.6 (0.0, 1.5, 20.9)
	100-150	36	8.3 (0.0, 1.6, 55.1)	85.6	0.04 (0.00, 0.00, 0.46)	4.5 (0.0, 1.3, 65.1)	6.9 (0.0, 0.9, 70.5)
	150-200	17	9.5 (0.0, 2.1, 50.7)	81.7	0.02 (0.00, 0.01, 0.12)	7.7 (0.0, 1.9, 40.7)	10.7 (0.0, 1.7, 78.9)
	200-300	16	4.6 (0.0, 1.0, 19.2)	97.7	0.03 (0.00, 0.01, 0.21)	3.9 (0.0, 0.8, 19.7)	7.2 (0.0, 1.4, 43.5)
	300-400	8	3.0 (0.3, 0.5, 19.1)	96.5	0.01 (0.00, 0.00, 0.09)	1.1 (0.0, 0.7, 4.1)	9.5 (0.9, 1.8, 63.9)
	400-1000	17	0.7 (0.0, 0.3, 2.9)	97.8	0.00 (0.00, 0.00, 0.02)	2.0 (0.0, 0.0, 11.9)	2.5 (0.0, 1.4, 21.4)
	1000-3000	26	0.5 (0.0, 0.2, 6.1)	99.1	0.00 (0.00, 0.00, 0.01)	5.1 (0.0, 0.9, 42.5)	5.1 (0.0, 2.0, 65.9)
	3000-5000	12	0.1 (0.0, 0.0, 0.1)	99.3	0.00 (0.00, 0.00, 0.00)	1.1 (0.0, 0.7, 3.2)	1.4 (0.0, 1.1, 5.3)
	5000+	41	0.0 (0.0, 0.0, 0.3)	99.8	0.00 (0.00, 0.00, 0.00)	1.5 (0.0, 1.1, 6.2)	1.8 (0.0, 0.5, 21.9)
	Total	548	5.5 (0.0, 2.7, 73.1)	85.2	0.03 (0.00, 0.00, 0.71)	2.3 (0.0, 1.1, 65.1)	3.2 (0.0, 1.1, 78.9)
All	Total	914	5.3 (0.0, 2.2, 76.5)	84.5	0.03 (0.00, 0.00, 0.77)	3.6 (0.0, 1.3, 210.1)	3.8 (0.0, 1.3, 166.9)

*refer to Description of variables section above for detail

Table 2I. Reproducibility of CNV Length and Endpoints (in kb) in Regions Excluding Affymetrix-defined Hypervariable Regions in the Site-to-Site Study.

			% CV CNV Length	Average % Overlap	Median % Absolute Endpoint Deviation	SD Left Endpoint (kb)	SD Right Endpoint (kb)
Gain/ Loss	**CNV Range (kb)**	**N**	**Mean (Min, Median, Max)**	**Mean (Min, Median, Max)**			
Gain	50-75	48	6.2 (0.0, 2.0, 54.8)	74.2	0.01 (0.00, 0.00, 0.18)	3.2 (0.0, 0.4, 40.2)	1.6 (0.0, 0.0, 8.4)
	75-100	36	10.8 (0.0, 6.2, 55.6)	76.9	0.06 (0.00, 0.00, 0.51)	7.4 (0.0, 3.1, 46.8)	6.8 (0.0, 3.0, 52.7)
	100-150	72	11.8 (0.0, 5.5, 55.1)	76.3	0.05 (0.00, 0.00, 0.46)	12.5 (0.0, 4.0, 63.8)	4.3 (0.0, 1.6, 56.4)
	150-200	24	10.5 (0.0, 3.2, 61.7)	76.2	0.05 (0.00, 0.00, 0.36)	10.9 (0.0, 4.2, 51.8)	9.1 (0.0, 0.7, 114.7)
	200-300	47	8.0 (0.0, 2.9, 89.3)	88.5	0.02 (0.00, 0.00, 0.40)	15.0 (0.0, 3.7, 238.5)	7.6 (0.0, 4.1, 99.3)
	300-400	13	24.1 (0.0, 5.5, 144.4)	78.9	0.05 (0.00, 0.01, 0.40)	34.8 (0.0, 8.0, 298.4)	55.2 (0.0, 3.2, 488.7)
	400-1000	28	16.1 (0.0, 1.8, 100.6)	77.8	0.08 (0.00, 0.01, 1.21)	41.2 (0.0, 3.6, 358.2)	78.0 (0.0, 5.4, 695.3)
	1000-3000	35	12.1 (0.0, 1.4, 61.7)	83.8	0.06 (0.00, 0.00, 0.44)	106.2 (0.0, 4.5, 538.4)	83.8 (0.0, 3.2, 1011.9)
	3000-5000	3	0.8 (0.1, 0.2, 2.1)	98.9	0.00 (0.00, 0.00, 0.00)	20.1 (2.2, 6.9, 51.1)	20.4 (0.0, 4.8, 56.5)
	5000+	60	3.0 (0.0, 0.0, 146.8)	98.6	0.00 (0.00, 0.00, 0.03)	183.1 (0.0, 0.3, 10583.8)	352.8 (0.0, 0.0, 11297.1)
	Total	366	9.7 (0.0, 2.5, 146.8)	81.1	0.04 (0.00, 0.00, 1.21)	51.0 (0.0, 2.7, 10583.8)	77.2 (0.0, 1.6, 11297.1)
Loss	25-50	201	7.3 (0.0, 5.0, 60.3)	78.7	0.03 (0.00, 0.00, 0.66)	1.6 (0.0, 0.5, 22.6)	1.8 (0.0, 0.4, 23.9)
	50-75	84	11.3 (0.0, 6.5, 71.4)	83.5	0.04 (0.00, 0.00, 0.68)	4.0 (0.0, 2.2, 43.4)	4.0 (0.0, 0.6, 33.8)
	75-100	38	22.5 (0.0, 6.3, 115.9)	80.5	0.09 (0.00, 0.01, 0.50)	15.6 (0.0, 2.3, 116.5)	6.5 (0.0, 2.4, 47.0)
	100-150	57	6.8 (0.0, 3.8, 41.5)	91.0	0.01 (0.00, 0.00, 0.20)	6.5 (0.0, 2.0, 55.6)	2.9 (0.0, 1.5, 16.4)
	150-200	21	8.5 (0.2, 2.1, 104.1)	83.8	0.02 (0.00, 0.00, 0.17)	11.0 (0.0, 3.0, 163.9)	3.6 (0.0, 0.6, 43.9)
	200-300	19	11.2 (0.4, 7.1, 43.4)	83.2	0.04 (0.00, 0.00, 0.42)	18.0 (0.0, 4.4, 121.4)	12.3 (0.0, 1.6, 79.3)
	300-400	14	15.4 (0.0, 0.9, 94.8)	82.1	0.01 (0.00, 0.00, 0.05)	35.5 (0.0, 0.3, 293.4)	21.8 (0.0, 2.4, 184.3)
	400-1000	28	19.7 (0.0, 4.6, 167.4)	82.9	0.12 (0.00, 0.01, 1.26)	57.9 (0.0, 5.3, 466.3)	75.1 (0.0, 12.0, 1442.1)
	1000-3000	35	3.2 (0.0, 0.4, 33.4)	94.5	0.01 (0.00, 0.00, 0.10)	21.3 (0.0, 1.1, 342.6)	28.5 (0.0, 4.7, 202.4)
	3000-5000	11	0.3 (0.0, 0.0, 2.6)	83.9	0.00 (0.00, 0.00, 0.00)	11.4 (0.0, 0.0, 117.9)	4.0 (0.0, 0.0, 24.6)
	5000+	40	0.5 (0.0, 0.0, 9.4)	99.6	0.00 (0.00, 0.00, 0.06)	8.1 (0.0, 0.6, 164.7)	50.9 (0.0, 0.1, 1484.5)
	Total	548	9.0 (0.0, 3.7, 167.4)	84.4	0.03 (0.00, 0.00, 1.26)	10.1 (0.0, 1.2, 466.3)	12.6 (0.0, 1.3, 1484.5)
All	Total	914	9.3 (0.0, 3.2, 167.4)	83.2	0.03 (0.00, 0.00, 1.26)	26.4 (0.0, 1.5, 10583.8)	38.5 (0.0, 1.4, 11297.1)

*refer to Description of variables section above for detail

Reproducibility of Loss of Heterozygosity (LOH) regions:

Reproducibility of LOH calls was calculated in the site-to-site study and is shown in Table 3.

Table 3. Reproducibility of LOH calls in the site-to-site study.

Size Range (Mb)	# LOH regions	Call Rate	Pairwise Replicate Agreement		PPA	
			Overlap			
			50%	80%	50%	80%
3-4	228	84.6	90.0	82.4	93.5	83.9
4-5	38	97.9	95.4	81.0	96.3	81.6
5-10	68	96.7	91.5	84.9	93.4	86.6
>10	147	99.7	96.0	93.4	96.3	93.6
Total	481	91.9	92.4	86.0	94.6	87.4

Study 2. Lot-to-Lot reproducibility:

A between-lot reproducibility study was conducted to test the impact of array lot, reagent lot, and operator on the reproducibility of the CytoScan Dx Assay. Forty-seven (47) genomic DNA (gDNA) samples (including 1 control) were randomized across 2 plates with each plate including 12 blood gDNA, 11 cell-line gDNA, and 1 cell-line control gDNA. Samples in each replicate plate were independently randomized for plate location. All 47 gDNA samples were evaluated with the 3 lots of arrays, 3 lots of reagents, and 6 operators at 1 investigative site – Affymetrix, Santa Clara, CA.

The samples were selected to maximize the variation across the genome with consideration to gain and loss segments, chromosomal representation, CNV regions in genic and non-genic regions, and in telomeric and centromeric regions. The samples had aberrations that collectively covered 31.3% of the genome. Each sample was run 10 times across the study, except for one control sample which was run in every batch and therefore run 24 times. Data analysis methods are same as those used in the site-to-site reproducibility study. Reproducibility results observed in the lot-to-lot reproducibility study were similar to the results in the site-to-site study.

b. Linearity/assay reportable range:

N/A

c. *Traceability, Stability, Expected values (controls, calibrators, or methods):*

Stability:

Stability studies were performed which included packaging/shipping stability, shelf life, and freeze/thaw and inversion stability. The stability of gDNA samples under expected storage conditions was also evaluated. In process QC standard; visual inspection for discoloration, leakage, volume loss or precipitation of reagents; and array QC metrics were assessed.

Reagent packaging/shipping stability was evaluated under ambient, refrigerated and frozen shipping conditions, and array stability was evaluated under ambient conditions (summer or winter) over 120 hour periods in a kit box of 6 arrays. One gDNA sample was used for both shelf life testing and freeze/thaw and inversion stability studies. CytoScan Dx microarrays and reagents were each randomly selected from 3 design validation lots and 12 replicates of the gDNA samples were tested with CytoScan Dx reagents and array at storage time = 3 months. A variety of storage, freeze/thaw, in use and inversion conditions were tested.

Real-time stability studies currently support a three month shelf life. CytoScan® Dx reagents can undergo a total of 10 freeze/thaw cycles (-20°C/Wet Ice Freeze/Thaw) total, including 6 in-use cycles and 4 cycles prior to use in the assay. Enzymes can withstand 5 dry ice/-20°C freeze/thaw cycles. The recommended number of freeze/thaw cycles for components stored at -15 and -25°C are 4 in-use cycles within 30 days, 3 cycles prior to use in the assay, and 4 dry ice/-20°C freeze/thaw cycles for the enzymes. Purified gDNA samples are stable when stored for up to 6 weeks following DNA extraction, at either 15°C to 30°C (ambient), 2°C to 8°C (refrigerated), or −15°C to −25°C (frozen, with up to 4 freeze/cycles).

Assay controls:

The package insert recommends including a positive control in every batch of samples processed.

d. *Detection limit:*

DNA input:

The amount of genomic DNA recommended for testing per sample with the CytoScan Dx is 250 ng. To determine the performance of the CytoScan Dx across a range of genomic DNA input concentrations, the amount of genomic DNA used in the initial fragmentation step of the assay was tested at levels ranging from 1 ng to 500 ng. Eight genomic DNA samples containing 31 CNV regions (representing 317 kb- 83.1 Mb of sizes of CNV region gains and losses) were tested in the study. These include 7 cell line derived samples from Coriell or ATCC that have known chromosomal aberrations and 1 blood derived sample from a healthy donor. A set of input DNA

levels serially diluted were tested at concentrations of 1, 2, 10, 50, 250, 300, 375, and 500 ng. The parameters measuring CytoScan performance, copy number state within copy number variation region (PMC), the endpoints of the copy number region (PMCB), and the number of non-copy number 2 segments (AASC) were compared with predefined acceptance criteria. The Limit of Detection is defined as the DNA input level where 100% of CNV markers can be detected accurately 95% of the time. DNA input of 250 ng was set as the control level. The results demonstrated that 10 ng of genomic DNA is a conservative lower limit for the CytoScan Dx assay based on the evaluated criteria. Array QC metrics including median absolute pair-wise difference (MAPD) and single nucleotide polymorphism quality control (SNPQC) also showed concordant results, supporting that array data meeting prespecified acceptance criteria were achievable for input DNA levels ≥ 10 ng. The assay functions appropriately with genomic DNA input of up to 500 ng.

Mosaicism:

To determine the level of mosaicism reliably detected by the CytoScan Dx, a total of 8 gDNA samples purified from Coriell cell lines were tested in the study. Each sample contained 1 gain and 1 loss on autosomal chromosomes, for a total of 16 CNV regions. The study was designed to simulate different percentages of mosaicism by mixing different proportions of 2 different gDNA samples. Four pairs of samples were mixed at 12 different mixture levels (0, 5, 10, 20, 30, 40, 60, 70, 80, 90, 95, and 100%) to generate samples representative of specified levels of mosaicism. This design generated 48 different DNA mixtures that were analyzed on 48 CytoScan Dx microarrays.

Receiver operating characteristics (ROC) curve analysis was used as a metric of how well the mosaic copy number markers are separated from the adjacent normal copy number markers. An ROC curve was generated for each mosaic region in each sample mixture. The mosaic region was scored as detected if the area under the curve (AUC) of the ROC curve met or exceeded the predefined threshold of 0.90. An AUC of 0.90 means that each marker considered has a 90% probability of detecting a mosaic segment correctly. Results with the 8 gDNA samples showed that the CytoScan Dx can detect mosaic segments in samples with greater than 15% mosaicism.

Another 2 gDNA samples at 8 different mixture levels were run as 9 replicates on the CytoScan Dx system to establish repeatability of mosaicism detection, which generated data from 72 CytoScan Dx microarrays. Samples NA13330 and NA07216 were mixed in 11 different proportions scored for the ability to distinguish a mosaic segment in 5 genomic regions (chromosomes 1, 3, 4, 7 and X). Each condition was run on 9 replicate tests in CytoScan Dx to assess repeatability. Mosaic regions were scored by calculating the AUC for the markers in the region and an equal number of adjacent markers. The results show that at a 15% mosaicism level, the mean AUC was 0.96 with 50% of the data falling within the inter quartile range of ± 1.9%. The CV at the 15% mosaicism level was 1.6%. Overall, the claimed detection level is that the CytoScan Dx reliably detects mosaicism greater than 20%, and that mosaicism less than 20% may not be reliably detected.

e. Analytical specificity:

Interfering Substances:
To assess the impact of interfering substances on the CytoScan Dx Assay, simulated aberrant or normal blood samples were either spiked with conjugated and unconjugated bilirubin (60 mg/dL), triglycerides (triolein, 3000 mg/dL), and hemoglobin (>0.5g/dL). Twenty-four (24) normal blood samples with no known chromosomal abnormalities and 9 spike-in cell lines with 19 known aberrant regions were utilized in the study. Each cell line was cultured, harvested, counted and re-suspended in leukocyte-depleted blood to a final concentration of $8x10^6$ cells/mL. DNA was extracted from each sample to be tested using CytoScan assay and the assay performance was measured against critical limits of in-process QC, array-based QC and performance metrics. For each interferent test condition, array QC and assay performance results met the sponsor's acceptance criteria and no interference was observed with any of the tested substances.

Carryover:
To determine the effect of potential gDNA carry-over from 1 array to the next when processing multiple arrays on the same Fluidics workstation 450Dx v.2, 12 gDNA samples that represent a variety of copy number gains and losses on over 15 chromosomes were tested in the study. For a given fluidic station module, a sequence of 1 normal sample followed by 3 successive aberrant samples (with the same aberrations) then followed by another run of the same normal sample was used to determine any potential carry-over effect due to genomic DNA. These 5 sequential runs were performed in a predefined order on 12 modules of 3 fluidics stations. Prespecified criteria for establishing the absence of carryover was to compare the array performance metric values and copy number state determination on a set of predefined aberrations between the first run and the fifth run. There was no significant difference in copy number state determination (Wilcoxon signed rank statistic (S) = 0.5, p-value = 1.00) or array QC metrics (Wilcoxon signed rank statistic, $p > 0.05$) between the first run and the fifth run, suggesting no carry-over from the aberrant gDNA samples by potential sources such as the fluidic station.

Cross-Contamination:
Sensitivity to cross-contamination was evaluated under simulated cross-contaminating conditions using 8 gDNA samples, mixed into 4 pairs at 12 different ratios before the hybridization step. A total of 48 arrays were analyzed for the assay performance and QC metrics. Prespecified acceptance criteria were that the single-nucleotide polymorphism quality control (SNPQC) should not fall below the standard acceptance limit (SNPQC ≥12) at approximately 20% contamination. SNPQC values ≥12 indicate the absence of substantial cross-contamination.

f. Assay cut-off:

N/A

2. Comparison studies:

a. Accuracy:

Accuracy of the CytoScan Dx Assay results was assessed by comparing the CNVs identified by CytoScan Dx Assay to the results obtained using alternative methods. A total of 1515 CNVs were identified by the CytoScan Dx Assay in a total of 137 gDNA samples (48 purified from blood, 86 from Coriell Cell Repository cell lines, 3 from ATCC cell lines). The samples were selected to maximize the variation across the genome with consideration to gain and loss segments of various sizes/number of probes, chromosomal representation, CNV regions in genic and non-genic regions, and in telomeric and centromeric regions. Of the 137 samples in the sequencing study, 5 samples were excluded for a total of 132 evaluable samples (3 of the 5 omitted samples were incorrectly annotated samples and 2 samples were hyper-segmented i.e., >40 segments, indicating poor quality DNA). These 132 samples had 1280 eligible CNVs for inclusion in the analysis. The CNVs covered 63.5% of the genome and were more prevalent in non-telomeric/non-centromeric regions than in telomeric/centromeric (62.3% vs 37.7%) and in genic (79.1%) than non-genic regions. A total of 28.91% of the CNVs had high (>45%) GC content.

The 132 samples were analyzed by a validated high throughput method to confirm the accuracy of the CytoScan Dx CNV results. The criterion for accuracy was agreement between the CNVs identified by the CytoScan Dx Assay and sequencing method based on a ≥50% overlap in markers or size, and the same copy number state (gain or loss) between the two methods. The results are summarized and stratified by copy number state, size or marker range, and genomic region in Tables 4A-E. Due to the number of CNVs that did not show gain or loss (i.e., were copy number neutral state) using high throughput sequencing, another analytically validated molecular method was performed on a statistically appropriate number of CNVs, and CytoScan Dx Assay results were then compared to this composite analytical method (Tables 5A-C). This proportion analysis demonstrated a modest improvement in results from 78.8% (95%CI, 76.4-80.9%) to 88.7% (95% CI, 84.2-92.2%). The results in Tables 4A-E and 5A-C summarize the accuracy (% agreement between CytoScan Dx Assay and comparison method) and the false positive rate (FPR) (i.e., the alternate method(s) did not confirm the CNV detected by the Cytoscan Dx; 1-agreement; see footnote to tables for explanation).

Table 4A. CytoScan Dx Assay Accuracy for All CNV Regions Stratified by Gain/Loss and Size (kb) when Compared to Sequencing Method.

Gain/ Loss	CNV Range (kb)	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR** (95% CI)* Based on Sequencing
Gain	50-75	49	39	79.6% (66.4%, 88.5%)	20.4% (11.5%, 33.6%)
	75-100	48	21	43.8% (30.7%, 57.7%)	56.3% (42.3%, 69.3%)
	100-150	133	95	71.4% (63.2%, 78.4%)	28.6% (21.6%, 36.8%)
	150-200	39	20	51.3% (36.2%, 66.1%)	48.7% (33.9%, 63.8%)
	200-300	56	38	67.9% (54.8%, 78.6%)	32.1% (21.4%, 45.2%)

Gain/ Loss	CNV Range (kb)	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR** (95% CI)* Based on Sequencing
	300-400	42	31	73.8% (58.9%, 84.7%)	26.2% (15.3%, 41.1%)
	400-1000	123	67	54.5% (45.7%, 63.0%)	45.5% (37.0%, 54.3%)
	1000+	83	82	98.8% (93.5%, 99.8%)	1.2% (0.2%, 6.5%)
	Total	573	393	68.6% (64.7%, 72.3%)	31.4% (27.7%, 35.3%)
Loss	25-50	157	117	74.5% (67.2%, 80.7%)	25.5% (19.3%, 32.8%)
	50-75	92	80	87.0% (78.6%, 92.4%)	13.0% (7.6%, 21.4%)
	75-100	42	36	85.7% (72.2%, 93.3%)	14.3% (6.7%, 27.8%)
	100-150	168	149	88.7% (83.0%, 92.6%)	11.3% (7.4%, 17.0%)
	150-200	26	24	92.3% (75.9%, 97.9%)	7.7% (2.1%, 24.1%)
	200-300	35	31	88.6% (74.0%, 95.5%)	11.4% (4.5%, 26.0%)
	300-400	26	22	84.6% (66.5%, 93.9%)	15.4% (6.1%, 33.5%)
	400-1000	51	49	96.1% (86.8%, 98.9%)	3.9% (1.1%, 13.2%)
	1000+	110	107	97.3% (92.3%, 99.1%)	2.7% (0.9%, 7.7%)
	Total	707	615	87.0% (84.3%, 89.3%)	13.0% (10.7%, 15.7%)
Total		1280	1008	78.8% (76.4%, 80.9%)	21.3% (19.1%, 23.6%)

*95%CI calculated using the Wilson score method.

** FPR=Pr(Sequencing≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

Table 4B. CytoScan Dx Assay Accuracy for All CNV Regions Stratified by Gain/Loss and Size (Markers) when Compared to Sequencing Method.

Gain/ Loss	CNV Range (Markers)	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR** (95% CI)* Based on Sequencing
Gain	50-75	121	90	74.4% (65.9%, 81.3%)	25.6% (18.7%, 34.1%)
	75-100	70	36	51.4% (40.0%, 62.8%)	48.6% (37.2%, 60.0%)
	100-150	131	83	63.4% (54.8%, 71.1%)	36.6% (28.9%, 45.2%)
	150-200	69	43	62.3% (50.5%, 72.8%)	37.7% (27.2%, 49.5%)
	200-300	78	43	55.1% (44.1%, 65.7%)	44.9% (34.3%, 55.9%)
	300-400	16	13	81.3% (57.0%, 93.4%)	18.8% (6.6%, 43.0%)
	400-1000	21	20	95.2% (77.3%, 99.2%)	4.8% (0.8%, 22.7%)
	1000+	67	65	97.0% (89.8%, 99.2%)	3.0% (0.8%, 10.2%)
	Total	573	393	68.6% (64.7%, 72.3%)	31.4% (27.7%, 35.3%)
Loss	25-50	158	121	76.6% (69.4%, 82.5%)	23.4% (17.5%, 30.6%)
	50-75	115	104	90.4% (83.7%, 94.6%)	9.6% (5.4%, 16.3%)
	75-100	154	137	89.0% (83.0%, 93.0%)	11.0% (7.0%, 17.0%)
	100-150	58	51	87.9% (77.1%, 94.0%)	12.1% (6.0%, 22.9%)
	150-200	31	24	77.4% (60.2%, 88.6%)	22.6% (11.4%, 39.8%)
	200-300	31	22	71.0% (53.4%, 83.9%)	29.0% (16.1%, 46.6%)
	300-400	17	16	94.1% (73.0%, 99.0%)	5.9% (1.0%, 27.0%)
	400-1000	39	36	92.3% (79.7%, 97.3%)	7.7% (2.7%, 20.3%)
	1000+	104	104	100.0% (96.4%, 100.0%)	0.0% (0.0%, 3.6%)
	Total	707	615	87.0% (84.3%, 89.3%)	13.0% (10.7%, 15.7%)
Total		1280	1008	78.8% (76.4%, 80.9%)	21.3% (19.1%, 23.6%)

*95%CI calculated using the Wilson score method.

** FPR=Pr(Sequencing≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

Table 4C. CytoScan Dx Assay Accuracy for CNVs in Regions Excluding Affymetrix-defined Hypervariable Regions Stratified by Gain/Loss and Size (kb) when Compared to Sequencing Method.

Gain/ Loss	CNV Range (kb)	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR** (95% CI)* Based on Sequencing
Gain	50-75	31	28	90.3% (75.1%, 96.7%)	9.7% (3.3%, 24.9%)
	75-100	24	17	70.8% (50.8%, 85.1%)	29.2% (14.9%, 49.2%)
	100-150	85	72	84.7% (75.6%, 90.8%)	15.3% (9.2%, 24.4%)
	150-200	28	20	71.4% (52.9%, 84.7%)	28.6% (15.3%, 47.1%)
	200-300	44	38	86.4% (73.3%, 93.6%)	13.6% (6.4%, 26.7%)
	300-400	35	31	88.6% (74.0%, 95.5%)	11.4% (4.5%, 26.0%)
	400-1000	37	32	86.5% (72.0%, 94.1%)	13.5% (5.9%, 28.0%)
	1000+	83	82	98.8% (93.5%, 99.8%)	1.2% (0.2%, 6.5%)
	Total	367	320	87.2% (83.4%, 90.2%)	12.8% (9.8%, 16.6%)
Loss	25-50	104	64	61.5% (51.9%, 70.3%)	38.5% (29.7%, 48.1%)
	50-75	34	23	67.6% (50.8%, 80.9%)	32.4% (19.1%, 49.2%)
	75-100	20	14	70.0% (48.1%, 85.5%)	30.0% (14.5%, 51.9%)
	100-150	67	55	82.1% (71.3%, 89.4%)	17.9% (10.6%, 28.7%)
	150-200	23	21	91.3% (73.2%, 97.6%)	8.7% (2.4%, 26.8%)
	200-300	25	23	92.0% (75.0%, 97.8%)	8.0% (2.2%, 25.0%)
	300-400	11	8	72.7% (43.4%, 90.3%)	27.3% (9.7%, 56.6%)
	400-1000	35	33	94.3% (81.4%, 98.4%)	5.7% (1.6%, 18.6%)
	1000+	110	107	97.3% (92.3%, 99.1%)	2.7% (0.9%, 7.7%)
	Total	429	348	81.1% (77.1%, 84.5%)	18.9% (15.5%, 22.9%)
Total		796	668	83.9% (81.2%, 86.3%)	16.1% (13.7%, 18.8%)

*95%CI calculated using the Wilson score method.

** FPR=Pr(Sequencing≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

Table 4D. CytoScan Dx Assay Accuracy for CNVs in Regions Excluding Affymetrix-defined Hypervariable Regions Stratified by Gain/Loss and Size (Marker) when Compared to Sequencing Method.

Gain/ Loss	CNV Range (Markers)	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR** (95% CI)* Based on Sequencing
Gain	50-75	76	57	75.0% (64.2%, 83.4%)	25.0% (16.6%, 35.8%)
	75-100	38	35	92.1% (79.2%, 97.3%)	7.9% (2.7%, 20.8%)
	100-150	56	46	82.1% (70.2%, 90.0%)	17.9% (10.0%, 29.8%)
	150-200	49	43	87.8% (75.8%, 94.3%)	12.2% (5.7%, 24.2%)
	200-300	46	41	89.1% (77.0%, 95.3%)	10.9% (4.7%, 23.0%)
	300-400	14	13	92.9% (68.5%, 98.7%)	7.1% (1.3%, 31.5%)

Gain/ Loss	CNV Range (Markers)	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR** (95% CI)* Based on Sequencing
	400-1000	21	20	95.2% (77.3%, 99.2%)	4.8% (0.8%, 22.7%)
	1000+	67	65	97.0% (89.8%, 99.2%)	3.0% (0.8%, 10.2%)
	Total	367	320	87.2% (83.4%, 90.2%)	12.8% (9.8%, 16.6%)
Loss	25-50	104	68	65.4% (55.8%, 73.8%)	34.6% (26.2%, 44.2%)
	50-75	47	36	76.6% (62.8%, 86.4%)	23.4% (13.6%, 37.2%)
	75-100	50	34	68.0% (54.2%, 79.2%)	32.0% (20.8%, 45.8%)
	100-150	41	36	87.8% (74.5%, 94.7%)	12.2% (5.3%, 25.5%)
	150-200	24	21	87.5% (69.0%, 95.7%)	12.5% (4.3%, 31.0%)
	200-300	25	18	72.0% (52.4%, 85.7%)	28.0% (14.3%, 47.6%)
	300-400	13	12	92.3% (66.7%, 98.6%)	7.7% (1.4%, 33.3%)
	400-1000	21	19	90.5% (71.1%, 97.3%)	9.5% (2.7%, 28.9%)
	1000+	104	104	100.0% (96.4%, 100.0%)	0.0% (0.0%, 3.6%)
	Total	429	348	81.1% (77.1%, 84.5%)	18.9% (15.5%, 22.9%)
Total		796	668	83.9% (81.2%, 86.3%)	16.1% (13.7%, 18.8%)

*95%CI calculated using the Wilson score method.

** FPR=Pr(Sequencing ≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

Table 4E. CytoScan Dx Assay Accuracy for CNVs in Affymetrix-defined Hypervariable Regions when Compared to Sequencing Method.

Region	Sample Size (N)	Agree	% Agreement (95% CI)* Based on Sequencing	FPR (95% CI)* Based on Sequencing
1q44	43	43	100.0% (91.8%, 100.0%)	0.0% (0.0%, 8.2%)
5q35.3	48	48	100.0% (92.6%, 100.0%)	0.0% (0.0%, 7.4%)
7p14.1	3	3	100% (43.8%, 100.0%)	0.0% (0.0%, 56.2%)
8p11.22	75	72	96.0% (88.9%, 98.6%)	4.0% (1.4%, 11.1%)
11q11	85	58	68.2 (57.7%, 77.2%)	31.8% (22.8%, 42.3%)
14q11.2	27	19	70.4% (51.5%, 84.1%)	29.6% (15.9%, 48.5%)
14q32.33	118	24	20.3% (14.1%, 28.5%)	79.7% (71.5%, 85.9%)
17q21.31	71	70	98.6% (92.4%, 99.8%)	1.4% (0.2%, 7.6%)
22q11.22	14	3	21.4% (7.6%, 47.6%)	78.6% (52.4%, 92.4%)
Total	484	340	70.2% (66.0%, 74.1%)	29.8% (25.9%, 34.0%

*95%CI calculated using the Wilson score method.

** FPR=Pr(Sequencing ≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

Table 5A. CytoScan Dx Assay Accuracy for CNV Regions Excluding Affymetrix-defined Hypervariable Regions Stratified by Gain/Loss and Size (kb) when Compared to a Composite Method.

Gain/ Loss	CNV Range (kb)	Sample Size (N)	% Agreement (95% CI)* Based on Composite Methods	FPR** (95% CI)* Based on Composite Methods
Gain	50-75	31	93.5% (78.2%, 94.6%)	6.5% (5.4%, 21.8%)
	75-100	24	83.3% (68.1%, 86.4%)	16.7% (13.6%, 31.9%)
	100-150	85	55.5% (29.1%, 80.4%)	44.5% (19.6%, 70.9%)
	150-200	28	76.0% (62.3%, 80.6%)	24.0% (19.4%, 37.7%)
	200-300	45	88.9% (72.0%, 91.0%)	11.1% (9.0%, 28.0%)
	300-400	35	71.4% (31.2%, 100.0%)	28.6% (0.0%, 68.8%)
	400-1000	37	94.6% (71.5%, 95.9%)	5.4% (4.1%, 28.5%)
	1000+	83	88.6% (62.5%, 96.7%)	11.4% (3.3%, 37.5%)
	Total	368	79.1% (69.7%, 86.7%)	20.9% (13.3%, 30.3%)
Loss	25-50	103	91.6% (88.3%, 92.4%)	8.4% (7.6%, 11.7%)
	50-75	34	100.0% (89.8%, 100.0%)	0.0% (0.0%, 10.2%)
	75-100	20	95.0% (79.7%, 95.8%)	5.0% (4.2%, 20.3%)
	100-150	67	95.1% (83.7%, 95.8%)	4.9% (4.2%, 16.3%)
	150-200	23	100.0% (67.9%, 100.0%)	0.0% (0.0%, 32.1%)
	200-300	25	100.0% (67.3%, 100.0%)	0.0% (0.0%, 32.7%)
	300-400	11	100.0% (77.5%, 100.0%)	0.0% (0.0%, 22.5%)
	400-1000	35	100.0% (65.3%, 100.0%)	0.0% (0.0%, 34.7%)
	1000+	110	100.0% (76.1%, 100.0%)	0.0% (0.0%, 23.9%)
	Total	428	97.0% (95.1%, 97.1%)	3.0% (2.9%, 4.9%)
All		796	88.7% (84.2%, 92.2%)	11.3% (7.8%, 15.8%)

*95%CI calculated using the Wilson score method.

** FPR=Pr(Composite≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

***Agreement with composite methods may be negatively adversely impacted by variable marker density on the CytoScan Dx Array.

Table 5B. CytoScan Dx Assay Accuracy for CNVs Regions Excluding Affymetrix-defined Hypervariable Regions Stratified by Gain/Loss and Size (Markers) when Compared to a Composite Method.

Gain/ Loss	CNV Range (Markers)	Sample Size (N)	% Agreement (95% CI)* Based on Composite Methods	FPR** (95% CI)* Based on Composite Methods
Gain	50-75	76	73.6% (67.4%, 75.8%)	26.4% (24.2%, 32.6%)
	75-100	38	78.3% (61.0%, 83.1%)	21.7% (16.9%, 39.0%)
	100-150	57	72.4% (60.9%, 76.8%)	27.6% (23.2%, 39.1%)
	150-200	49	72.4% (57.0%, 78.3%)	27.6% (21.7%, 43.0%)
	200-300	46	74.7% (61.6%, 79.1%)	25.3% (20.9%, 38.4%)
	300-400	14	79.5% (17.4%, 95.5%)	20.5% (4.5%, 82.6%)
	400-1000	21	95.2% (60.1%, 97.0%)	4.8% (3.0%, 39.9%)
	1000+	67	94.0% (68.7%, 95.6%)	6.0% (4.4%, 31.3%)
	Total	368	79.1% (76.6%, 79.7%)	20.9% (20.3%, 23.4%)
Loss	25-50	104	90.9% (87.8%, 91.2%)	9.1% (8.8%, 12.2%)
	50-75	47	95.7% (83.7%, 96.3%)	4.3% (3.7%, 16.3%)
	75-100	50	97.6% (90.6%, 97.7%)	2.4% (2.3%, 9.4%)
	100-150	41	100.0% (72.1%, 100.0%)	0.0% (0.0%, 27.9%)
	150-200	23	100.0% (76.2%, 100.0%)	0.0% (0.0%, 23.8%)
	200-300	25	98.9% (77.7%, 99.1%)	1.1% (0.9%, 22.3%)
	300-400	13	100.0% (84.4%, 100.0%)	0.0% (0.0%, 15.6%)
	400-1000	21	100.0% (36.5%, 100.0%)	0.0% (0.0%, 63.5%)
	1000+	104	100.0% (69.3%, 100.0%)	0.0% (0.0%, 30.7%)
	Total	428	97.0% (95.2%, 97.0%)	3.0% (3.0%, 4.8%)
All		796	88.7% (87.7%, 88.8%)	11.3% (11.2%, 12.3%)

*95%CI calculated using the Wilson score method.

** FPR=Pr(Composite≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

***Agreement with composite methods may be negatively adversely impacted by variable marker density on the CytoScan Dx Array.

Table 5C. CytoScan Dx Assay Accuracy for CNVs in Affymetrix-defined Hypervariable Regions when Compared to a Composite Method.

Region	Sample Size (N)	% Agreement (95% CI)* Based on Composite Methods	FPR** (95% CI)* Based on Composite Methods
1q44	43	100% (59.6%, 100%)	0.0% (0.0%, 40.4%)
5q35.3	48	100% (59.2%, 100%)	0.0% (0.0%, 40.8%)
7p14.1	3	100% (61%, 100%)	0.0% (0.0%, 39.0%)
8p11.22	75	88% (59.6%, 99%)	12.0% (1.0%, 40.4%)
11q11	85	100% (72.4%, 100%)	0.0% (0.0%, 27.6%)
14q11.2	27	90.1% (58.2%, 95.4%)	9.9% (4.6%, 41.8%)
14q32.33	118	14.7% (4.9%, 39.3%)	85.3% (60.7%, 95.1%)
17q21.31	71	100.0% (70.1%, 100.0%)	0.0% (0.0%, 29.9%)
22q11.22	14	21.4% (15.7%, 42.5%)	78.6% (57.5%, 84.3%)

Region	Sample Size (N)	% Agreement (95% CI)* Based on Composite Methods	FPR** (95% CI)* Based on Composite Methods
All	484	74.5% (67.7%, 79.0%)	25.5% (21.0%, 32.3%)

*95%CI calculated using the Wilson score method for kb. A modified Wilson's method with adjustment for the effects of weighted and stratified sampling and sampling from finite populations was used to calculate the 95% C for markers.

** FPR=Pr(Composite≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with qPCR defined as truth. Agreement can also be referred to as PPV.

***Agreement with composite methods may be negatively adversely impacted by variable marker density on the CytoScan Dx Array.

Positive percent agreement for larger aberrations when compared to historical karyotype and FISH results:

Historical testing data were available for the 132 samples, which had previously identified 161 CNVs detected with either karyotype or FISH. As would be expected, these CNVs represented mostly large size CNVs (mean size 25.1 MB, median size 11.2 MB, range 1.8-175.9 MB). Either the ISCN karyotype or the historical diagnosis was manually parsed to identify copy number direction and starting and ending cytobands for the aberration(s). The genomic map locations for each cytoband were obtained from the UCSC Genome Browser using hg19 build. Due to the variability in human interpretation and the similarity of appearance of cytobands from different chromosomes in cases of translocations, cytoband locations were not very accurate. The regions identified by karyotype and FISH were expanded by 1 cytoband at each end, and these cytobands were converted to genomic base pair coordinates and the CytoScan Dx Assay CNV locations were compared to these historical karyotype or FISH results. Agreement was defined as the same copy number state, gain, or loss, with any amount of overlap. Positive percent agreement is defined as the proportion of karyotype or FISH identified CNVs (Gain/Loss) which have the same CytoScan CNV state call (Gain/Loss). The positive percent agreement between CytoScan Dx Assay and RPC was 91.4% (149/163; Wilson method 95% CI 86.1%-94.8%). Of the 14 missed aberrations, four were balanced translocations that are not detected by CytoScan Dx Assay, 3 CNVs were outside of the CytoScan Dx Assay marker regions (two at the Y-ter and one in the acrocentric p-arm of chromosome 22), and 1 low level mosaic that is below the stated detection limit for CytoScan Dx Assay.

Endpoint accuracy:

The endpoint distance for CNVs detected by both sequencing and CytoScan Dx assay was evaluated. Endpoint agreement between CytoScan Dx Assay and sequencing was assessed for those CNVs determined to have the same copy number state (gain or loss) by both CytoScan Dx and Sequencing. For the CNVs that have the same copy state, the endpoints were considered to agree if they were within ≤12 markers for losses and ≤25 markers for gains for each end (the endpoint agreement was not considered relative to total CNV size or probe number). In this analysis, Affymetrix-defined hypervariable regions of the genome were excluded. A total of 1367

endpoints were included in the analysis.

For the left (start) endpoint, the distance was calculated as the difference between the CytoScan Dx Assay start marker position and the sequencing start marker position using the following formula: Distance=CytoScan Dx Assay_start_marker_position - sequencing_start_marker_position. For the right (end) endpoint, the distance was calculated as the difference between the sequencing end marker position and the CytoScan Dx Assay end marker position using the following formula: Distance= sequencing_end_marker_position - CytoScan Dx Assay_end_marker_position. Negative values indicated that the CNV endpoint is further from the center of the CNV for CytoScan Dx Assay compared to sequencing, while positive values indicated that the CNV endpoint is to closer to the center of the CNV for CytoScan Dx Assay compared to sequencing. The overall endpoint agreement based on specified criteria (≤12 markers for loss segments and ≤25 markers for gain segments) for CNVs with state agreement is 93.4% (95% CI: 93.2-93.4%), with 94.8% and 92.4% for gains and losses, respectively. The data is presented in Table 6 below.

Table 6. Endpoint Agreement for Endpoint Criteria ≤12 Markers for Loss Segments, and ≤ 25 Markers for Gain Segments (Excluding CNVs within the Hypervariable Regions of the Genome).

Gain/ Loss	CNV Size (Markers)	Endpoints, N	Endpoint Agreement, N	Endpoint Agreement, % (95%CI)
Gain	50-75	104	104	100.0% (96.4%, 100.0%)
	75-100	60	60	100.0% (94.0%, 100,0%)
	100-150	73	71	97.3% (92.4%, 97.4%)
	150-200	77	71	92.2% (87.7%, 97.7%)
	200-300	86	76	88.4% (84.5%, 89.0%)
	300-400	22	20	90.9% (77.1%, 92.6%)
	400-1000	42	38	90.5% (82.7%, 91.5%)
	1000+	128	121	94.5% (91.7%, 94.7%)
	Total	592	561	94.8% (94.1%, 94.8%)
Loss	25-50	206	206	100.0% (98.2%, 100.0%)
	50-75	84	79	94.0% (89.9%, 94.4%)
	75-100	80	73	91.3% (87.0%, 91.7%)
	100-150	68	68	100.0% (94.7%, 100.0%)
	150-200	34	33	97.1% (87.1%, 97.4%)
	200-300	38	33	86.8% (78.6%, 88.3%)
	300-400	28	25	89.3% (78.2%, 90.9%)
	400-1000	34	24	70.6% (62.9%, 74.1%)
	1000+	203	175	86.2% (84.5%, 86.5%)
	Total	775	716	92.4% (91.9%, 92.4%)
All		1367	1277	93.4% (93.2%, 93.4%)

For those endpoints that met the criteria of demonstrating agreement between sequencing and CytoScan Dx Assay within 12 markers for loss segments and within 25 markers for gains, the distribution of the difference in markers between the

sequencing endpoint and the Cytoscan Dx Assay endpoint is shown below in Figure 1. The Y-axis represents number of CNV endpoints with the endpoint difference indicated on the X-axis.

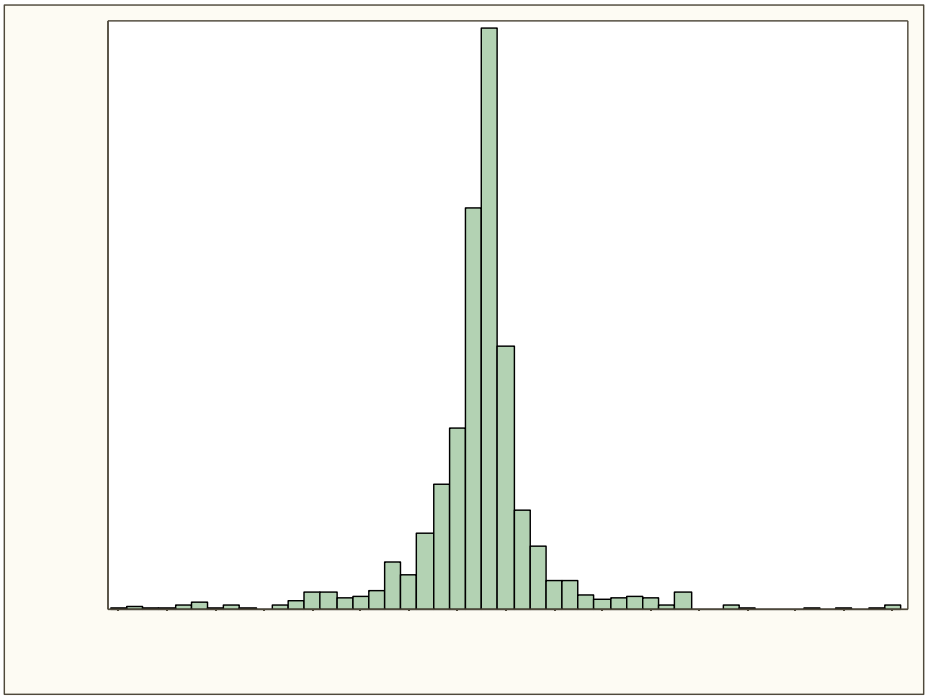


Figure 1. Endpoint Difference in Markers between Sequencing and CytoScan Dx Assay for CNVs that met criteria.

Negative values indicate that the endpoint determined by CytoScan Dx Assay is further from the center of the CNV than the endpoint determined by sequencing; positive values indicate that the endpoint determined by CytoScan Dx Assay is closer to the center of the CNV than the endpoint determined by sequencing. Count indicates the number of CNV endpoints with the indicated endpoint distance on x-axis.

LOH Accuracy:

Accuracy for LOH regions was determined by comparing genotypes derived from CytoScan Dx Assay to LOH calls from orthogonal high throughput sequence data. Each CytoScan array LOH segment was compared to the sequence data LOH segmentation. A CytoScan LOH segment was scored as matched if it overlapped with the high throughput sequencing LOH segment. The PPA between CytoScan Dx Assay and sequencing was 159/159 (100%, 95% CI 97.5%-100.0%) for LOH regions 3 MB and greater.

Additionally, CytoScan Dx Assay LOH calls were compared to a published table of LOH regions derived from HapMap sample genotypes published by Gibson et al [11]. The manuscript reported the presence of 1393 LOH regions >1 MB in length in 209 HapMap individuals and provided details on 20 of these regions, ranging in size from 5 MB and above. CytoScan Dx Assay identified all 20 (100.0%) of the regions

specifically described by Gibson et al. In total, CytoScan identified 42 LOH regions 5 MB and above and 79 LOH regions 2 MB and above in these samples.

b. Matrix comparison:

To verify consistent performance of CytoScan Dx, regardless of whether gDNA was extracted from peripheral blood collected with heparin or with EDTA, and regardless of whether the collection was made in full (~3 mL) or partial (~1.5 mL) blood collection tube volumes, 24 normal blood samples with no known chromosomal abnormalities and 10 spike-in cell lines with 21 known aberrant regions were utilized in the study. Each cell line was cultured, harvested, counted and re-suspended in leukocyte-depleted blood to a final concentration of $8x10^6$ cells/mL to create simulated aberrants. 3mL and 1.5 mL of each spike-in blood sample was dispensed into 3 mL BD Vacutainer EDTA and heparin tubes. A total of 144 samples (48 simulated CNV blood samples, including 8 no cell spike controls and 96 normal blood samples) were generated for this study. The gDNA was extracted and tested by CytoScan Dx for the evaluation of the potential impacts on assay performance due to different anticoagulants and fill volumes. The assay results were evaluated using standard in-process, array-based QC criteria, and performance criteria. This study met the sponsor's acceptance criteria and demonstrated that CytoScan performed consistently, regardless of whether genomic DNA was extracted from peripheral blood collected with heparin or with EDTA, and regardless of whether the collection was made in full (~3 mL) or partial (~1.5 mL) blood collection tube volumes.

3. Clinical studies:

a. Clinical Sensitivity:

A retrospective clinical study was performed to characterize the clinical performance characteristics of CytoScan Dx Assay for the purpose of reporting the pathogenic detection rate (potential diagnostic yield) of the assay. A total of 960 gDNA samples from previously tested DD/ID patients referred to three clinical laboratories for routine chromosomal testing were collected and analyzed with the CytoScan Dx. Each sample had a historical clinical laboratory interpretation of benign, pathogenic, or variant of unknown clinical significance (VOUS) made by a cytogeneticist based on CNVs detected by one or more methods used by the clinical site including karyotype, FISH, microarray (non-Affymetrix), or other technique (collectively referred to as routine patient care, RPC).

Samples were tested by CytoScan Dx and the results were provided to one of two independent cytogeneticists who then generated an overall clinical laboratory interpretation for each sample based on CytoScan result, including diagnosing a syndrome, as appropriate. The cytogeneticist also assessed each of the CNVs identified in each sample and classified them as benign, pathogenic, or VOUS. The cytogeneticist was permitted to request the data from parental samples that were run as part of routine care at the clinical laboratory. If the cytogeneticist requested parental results, and if the results were available, then the cytogeneticist was

permitted to use them to complete the interpretation. Neither the original clinical laboratory diagnosis, nor data from testing conducted during routine patient care such as karyotyping, FISH, microarray, PCR, MLPA, or any other type of copy number data were available to the cytogeneticist interpreting the CytoScan Dx results.

A total of 61.7% of all samples were male with an average age of 7.8 ± 11.3. Overall, 86% of the subjects were assessed using a microarray (excluding an Affymetrix array) as part of routine patient care. 425 CNVs were reported by the investigative sites.

Potential diagnostic yield was calculated for RPC methods as well as for CytoScan Dx Assay (Table 7) as the probability of a pathogenic interpretation by the RPC or CytoScan Dx Assay, stratified by RPC methods used in the sample collection sites. The diagnostic yield using CytoScan Dx was 13.8% vs. 13.3% for routine patient care. The diagnostic yield stratified by various methods is confounded by differences in technology utilization patterns in routine patient care, as well as interpretation, and therefore comparisons should be interpreted with caution and presented for interest only.

Table 7. Potential diagnostic yield for RPC and CytoScan Dx Assay.

		RPC		CytoScan Dx Assay	
RPC Method	**N**	**Pathogenic Calls (N)**	**Diagnostic Yield % (95% CI)***	**Pathogenic Calls (N)**	**Diagnostic Yield % (95% CI)***
Karyotype Only	72	10	13.9% (7.7%, 23.7%)	19	26.4% (17.6%, 37.6%)
Karyotype + FISH	11	6	54.6 % (28.0%, 78.7%)	6	54.6 % (28.0%, 78.7%)
Karyotype + FISH + Other**	10	2	20.0 % (5.7%, 51.0%)	2	20.0 % (5.7%, 51.0%)
Karyotype + Other	45	2	4.4 % (1.2%, 14.8%)	6	13.3 % (6.3%, 26.2%)
Microarray* Only**	351	10	2.9 % (1.6%, 5.2%)	9	2.6 % (1.4%, 4.8%)
Microarray + Karyotype	74	11	14.9 % (8.5%, 24.7%)	7	9.5 % (4.7%, 18.3%)
Microarray + FISH	77	54	70.1 % (59.2%, 79.2%)	48	62.3 % (51.2%, 72.3%)
Microarray + FISH + Other	2	0	0.0 % (0.0%, 65.8%)	0	0.0 % (0.0%, 65.8%)
Microarray + Karyotype + FISH + Other	67	6	9.0 % (4.2%, 18.2%)	8	11.9 % (6.2%, 21.8%)
Microarray + Karyotype + FISH	17	4	23.5 % (9.6%, 47.3%)	4	23.5 % (9.6%, 47.3%)
Microarray + Karyotype + Other	187	18	9.6 % (6.2%, 14.7%)	18	9.6 % (6.2%, 14.7%)
Microarray + Other	47	5	10.6 % (4.6%, 22.6%)	5	10.6 % (4.6%, 22.6%)
Total	960	128	13.3 % (11.3%, 15.6%)	132	13.8 % (11.7%, 16.1%)

At the sample level, clinical Positive Percent Agreement (PPA), defined as proportion of samples with routine patient care classified as pathogenic which are classified as pathogenic based on CytoScan Dx Assay results [i.e., Probability (CytoScan Dx Assay = pathogenic |Routine patient care classification = pathogenic)], is 105/128= 82.0% (95%CI, 74.5-87.7%) and shown in Table 8. Out of 23 samples that were pathogenic by RPC but deemed non-pathogenic based on CytoScan Dx Assay results,

22 samples were called VOUS by the cytogeneticist using CytoScan, and 1 sample was called benign. In these 23 samples, 49 CNVs were identified by RPC, of which 45 CNVs were identified by CytoScan Dx Assay; therefore, 45/49 CNVs (91.8%) that did not agree on the clinical call level agreed at the analytical level.

Table 8. Comparison of Sample Classification between RPC and CytoScan Dx Assay in the Prospective Study.

Cytogeneticist Interpretation of CytoScan Dx Assay Result		Diagnosis at Original Site				
		Pathogenic	Non-Pathogenic			Total
			VOUS	Benign	No CNVs	
Pathogenic		105	7	0	20	132
Non-Pathogenic	**VOUS**	22	81	69	199	371
	Benign	1	21	44	391	457
Total		128	109	113	610	960
PPA*		105/128 = 82.0% (95%CI 74.5-87.7%)				
NPA**		805/832 = 96.8% (95%CI 95.3-97.8%)				

*Percent Positive Agreement: Pr(CytoScan Dx Assay = pathogenic |Routine patient care classification = pathogenic)
**Percent Negative Agreement: Pr(CytoScan Dx Assay = non-pathogenic |Routine patient care classification = non-pathogenic)

At the CNV level, analytical accuracy was calculated as the percentage of CNVs identified by routine patient care (RPC) that was identified by CytoScan Dx Assay. In the 960 samples, RPC identified 680 CNVs. Out of these 680 CNVs, 639 CNVs were identified by CytoScan Dx and RPC (analytical agreement, 639/680 = 94.0% [95%CI 91.9%, 95.5%]). Of the 41 disagreements, 34 of the CNVs identified by RPC were outside of CytoScan Dx Assay reportable categories (2 CNVs were low-level mosaics, 1 mosaic ring chromosome CNV, 5 CNVs on Y in PAR regions, 14 balanced translocations/inversions CNVs, and 12 CNVs below the reported resolution of CytoScan Dx Assay) and 7 were not identified by CytoScan Dx Assay.

A total of 43 different clinical syndromes were represented in the prospective clinical study. Syndrome agreement was compared between clinical interpretation based upon RPC and CytoScan Dx Assay, with an overall interpretation positive agreement of 80.7%. Of the 18 disagreements, only 1 of the disagreements included a sample which disagreed analytically; 17/18 disagreements were due to differences in clinical laboratory interpretation. The one analytical disagreement was a sample called Nebulette syndrome by RPC. The syndrome types and positive percent agreement of clinical interpretations are itemized in Table 9.

Table 9. List of Syndromes Observed in the Prospective Study.

Syndrome	Number of Patients Based on RPC Reports	Number of Cases for which CytoScan & RPC Interpretation Agree	Positive Percent Agreement (%)
15q13.3 microdeletion	1	1	100
16p11.2 microdeletion	5	4	80
16p11.2 microduplication	3	2	67
16p13.11 microdeletion	2	2	100
16p13.11 microduplication neurocognitive disorder susceptibility locus	2	2	100
17p11.2 duplication / Potocki-Lupski	1	1	100
17q21.3 microdeletion / Koolen-de Vries	1	1	100
18p deletion	1	1	100
1p36 deletion	1	1	100
1q21.1 microdeletion	2	1	50
1q21.1 microdeletion (susceptibility locus to neurodevelopmental disorders)	2	1	50
1q21.1 susceptibility locus to thrombocytopenia-absent radius (TAR) syndrome	2	2	100
22q11 microduplication	5	4	80
22q13.3 deletion / Phelan-McDermid	1	1	100
2q37 monosomy	2	2	100
3q29 microdeletion	1	0	0
9q subtelomeric deletion / Kleefstra	1	1	100
Alagille	1	0	0
Angelman OR Prader-Willi	4	3	75
Bannayan-Riley-Ruvalcaba (BRRS) / PTEN hamartoma tumor	1	1	100
Cardiofaciocutaneous	1	0	0
Cri-du-chat	3	3	100
Dentinogenesis imperfecta	1	0	0
DiGeorge / Velocardiofacial / 22q11.21 Microdeletion	3	2	67
Down / Trisomy 21	11	10	91
Edwards / Trisomy 18	2	2	100
Francois-Neetens fleck corneal dystrophy	1	0	0
Jacobsen / 11q deletion	1	0	0

Syndrome	Number of Patients Based on RPC Reports	Number of Cases for which CytoScan & RPC Interpretation Agree	Positive Percent Agreement (%)
Klinefelter / XXY	4	2	50
Nebulette Syndrome[1]	1	0	0
Neurofibromatosis 1 with intellectual disability	1	1	100
Pallister-Killian	1	1	100
Patau / Trisomy 13	5	5	100
Smith-Magenis	1	1	100
Steroid sulfatase deficiency, X-linked / Ichthyosis	2	2	100
Tetrasomy 9p	1	1	100
Trisomy 8	1	1	100
Von Hippel-Lindau	1	1	100
Williams-Beuren	3	3	100
Wolf-Hirschhorn	2	2	100
Xq28 (MECP2) duplication	1	1	100
Other	2	2	100
Total	89	71	80.7

[1] Result was not analytically detected by CytoScan Dx.

b. Clinical specificity:

See above

c. Other clinical supportive data (when a. and b. are not applicable):

N/A

4. Clinical cut-off:

N/A

5. Expected values/Reference range:

The prevalence of CNVs in patient populations depends on risk factors such as age, gender, presence of symptoms, and family history. A blinded study was conducted to assess the potential impact of the CytoScan Dx Assay CNV results on interpretation using two sample sets. One set of 149 samples represented known syndromes (selected for breadth of representation of syndromes than expected prevalence in the intended use population). The second set of specimens were from 108 phenotypically normal subjects. The syndromic and normal samples were collected at 8 and 1 site, respectively. The syndrome classification was designated from the compilation of routine patient care testing, with the exception that any Affymetrix SNP6.0 microarray results were excluded. For the syndromic samples, the original clinical laboratory diagnosis accompanying the

samples was the diagnosed clinical syndrome determined by the external laboratories following routine patient care in patients referred for chromosomal testing due to developmental delay, intellectual disability, congenital anomalies, and/or dysmorphic features. The samples collected from phenotypically normal individuals, are assumed to be absent of pathogenic CNVs. CytoScan Dx Assay results were interpreted by a single trained cytogeneticist. In this retrospective study of subjects with syndromes and phenotypically normal subjects, on average CytoScan Dx Assay identified 15.4±11.5 (mean±SD) CNVs per syndromic sample and 10.4±3.40 (mean±SD) CNVs per phenotypically normal sample.

CytoScan Dx Assay supported the correct interpretation for 145 of 149 syndromic specimens, while 3 samples from 108 phenotypically normal individuals were interpreted as pathogenic. Cytogeneticist interpretations of CytoScan Dx Assay results for this study are summarized in Table 10.

Table 10. Interpretation of CytoScan Dx Assay Results in a Syndromic and Phenotypically Normal Sample Set.

		Syndromic	Phenotypically Normal
Pathogenic		145	3
Non-Pathogenic	**VOUS**	4	47
	Benign	0	58
Total		149	108

In both groups, CNVs were more common in genic than in non-genic regions (77.6% of the CNVs were in genic regions) and in non-telomeric or non-centromeric regions than in telomeric or centromeric regions (72.7% of the CNVs were in non-telomeric or non-centromeric regions).

In a multi-center prospective study of 960 subjects (see Clinical performance for details), CytoScan Dx Assay identified 11.2±4.1 (mean±SD) CNVs per subject.

N. Instrument Name:

Affymetrix GeneChip Microarray Instrumentation System

O. System Descriptions:

1. Modes of Operation:

   Batch - The GeneChip® System 3000Dx Scanner has an autoloader that enables the automated scanning of up to 48 CytoScan Dx Arrays.

2. Software:

   FDA has reviewed applicant's Hazard Analysis and software development processes for

this line of product types:

Yes __x____ or No _____

3. Specimen Identification:

   Each CytoScan Dx Array has a unique barcode. Operators register one Array barcode with one specimen ID in the Worklist. The Array barcode and the Worklist are used by the device software to identify the specimens.

4. Specimen Sampling and Handling:

   Specimens are processed according to CytoScan Dx® assay instructions.

5. Calibration:

   Installation and calibration are performed by the manufacturer. No user calibration required.

6. Quality Control:

   The CytoScan Dx Assay employs both in-process QC checks and array QC metrics to assist in identifying problems in the assay and instances in which the assay has failed. The in-process QC includes PCR, fragmentation gels, and purified PCR DNA yield. The array QC metrics used in CytoScan Dx Assay are the median absolute pair-wise difference (MAPD), and single nucleotide polymorphism quality control (SNPQC). MAPD compares the log2 ratios of each adjacent pair of markers along the chromosome and measures the local variability of the log2 ratios. SNPQC measures the distance between peaks representing the a/a, a/b, b/b genotypes in signal contrast space (minimum of inter-peak distances). ChAS Dx Software checks the DXCHP files for array QC values. The software issues a notice if the array QC parameters do not meet the thresholds.

P. Other Supportive Instrument Performance Characteristics Data Not Covered In The "Performance Characteristics" Section above:

Overall, a total of 1435 unique samples were utilized to determine the CytoScan Dx performance characteristics. Both cell lines and blood sample were used in the studies to provide adequate coverage of the genome. A total of 76% genome were tested in the studies, with 66% for gains and 39.7% for losses. During the assay development and validation stages, 77.5% of genome was tested. With all studies combined, 89.6% of the genome was evaluated with 84.2% for gains and 46.6% for losses. Studies with performance characteristics data not covered in the "performance characteristics" section were listed below.

1. Performance of Copy Number Variations Reported at the Lower Range of the Assay:

   The performance of CNVs reported at the lower range of the assay was assessed. The minimum size and number of markers of the CNVs reported is dictated by the filter setting, which is set at 25 kb and 25 markers for CN losses and 50 kb and 50 markers for CN gains. CNVs included in this study included 50-75 markers or kb for gains and 25-50 markers or kb for losses. Reproducibility at the filter setting is summarized in Tables 11A-B and analytical accuracy is summarized in Table 12.

   In the reproducibility study, pairwise replicate agreement for copy number gains 50-75 kb was 78.8% and for copy number losses 25-50 kb was 81.0% when the 50% overlapping criteria cutoff was used.

   In the analytical accuracy study, at the filter setting, accuracy for copy number gains 50-75 kb was 93.5% (95% CI: 78.2%, 94.6%) and for copy number losses 25-50 kb was 91.6% (95% CI: 88.3%, 92.4%) when using composite analytical method as a comparator.

   The filter setting for regions of loss of heterozygosity (LOH) is 3 MB. Reproducibility of calling LOH 3-4 MB in length is 91.9% as measured by call rate, and % agreement with high throughput sequencing genotype is 100.0% for LOH regions 3-5 MB in length.

Table 11A. Pairwise Replicate Agreement and Positive Percent Agreement for CNVs Near the Filter Setting.*

					Pairwise Replicate agreement		**PPA**	
			N	**Call Rate**	**Overlap**			
Gain/Loss	**Region**	**CNV range**			**50%**	**80%**	**50%**	**80%**
Gain	All	Marker	230	49.7	70.8	68.7	72.7	68.7
		kb	99	50.4	71.4	67.2	70.4	62.1
	Non-HV**	Marker	430	48.8	79.5	76.4	75.4	68.0
		kb	351	53.7	78.8	73.9	78.5	68.8
Loss	All	Marker	166	45.4	75.1	73.6	74.7	71.6
		kb	79	49.5	75.4	73.4	74.2	70.1
	Non-HV	Marker	387	49.8	81.0	78.1	77.4	70.6
		kb	286	54.3	81.0	77.2	80.9	73.4

*refer to Description of variables section above for detail

** CNVs located outside Affymetrix-defined hypervariable regions

Table 11B. Reproducibility of Length and Endpoints for CNVs Near the Filter Setting.*

				% CV CNV Length	Average % Overlap	Median % Absolute Endpoint Deviation	SD Left Endpoint	SD Right Endpoint
Gain/ Loss	**Region**	**Size Category***	**N**	**Mean (Min, Median, Max)**		**Mean (Min, Median, Max)**		
Gain	All	Marker	157	5.1 (0.0, 4.1, 30.1)	68.5	0.03 (0.00, 0.01, 0.22)	1.8 (0.0, 1.2, 12.5)	2.2 (0.0, 1.6, 11.3)
		kb	67	9.1 (0.0, 3.9, 54.8)	69.6	0.04 (0.00, 0.00, 0.37)	5.1 (0.0, 0.9, 40.2)	2.9 (0.0, 1.1, 14.0)
	Non-HV**	Marker	97	4.3 (0.0, 2.6, 30.1)	73.1	0.03 (0.00, 0.02, 0.22)	1.9 (0.0, 1.4, 12.5)	1.5 (0.0, 1.0, 11.3)
		kb	48	6.2 (0.0, 2.0, 54.8)	74.2	0.01 (0.00, 0.00, 0.18)	3.2 (0.0, 0.4, 40.2)	1.6 (0.0, 0.0, 8.4)
Loss	All	Marker	282	6.9 (0.0, 5.4, 63.7)	77.3	0.04 (0.00, 0.02, 0.48)	1.4 (0.0, 1.0, 18.4)	1.8 (0.0, 0.9, 26.8)
		kb	252	9.5 (0.0, 6.6, 60.3)	76.3	0.04 (0.00, 0.00, 0.66)	2.0 (0.0, 0.1, 22.9)	2.8 (0.0, 1.3, 24.3)
	Non-HV	Marker	250	6.3 (0.0, 4.9, 63.7)	78.8	0.03 (0.00, 0.02, 0.48)	1.5 (0.0, 1.1, 18.4)	1.5 (0.0, 0.8, 26.8)
		kb	201	7.3 (0.0, 5.0, 60.3)	78.7	0.03 (0.00, 0.00, 0.66)	1.6 (0.0, 0.5, 22.6)	1.8 (0.0, 0.4, 23.9)

*refer to Description of variables section above for detail

** CNVs located outside Affymetrix-defined hypervariable regions

Table 12. Analytical Accuracy for CNVs Near the Filter Setting.

Gain/Loss	Region	Size Category	Comparator Method	N	% Agreement (95% CI)**	FPR*** (95% CI)**
Gain	All	Marker	Sequencing	121	74.4% (65.9%, 81.3%)	25.6% (18.7%, 34.1%)
		Kb	Sequencing	49	79.6% (66.4%, 88.5%)	20.4% (11.5%, 33.6%)
	Non-HV*	Marker	Sequencing	76	75.0% (64.2%, 83.4%)	25.0% (16.6%, 35.8%)
		Marker	Composite	76	73.6% (67.4%, 75.8%)	26.4% (24.2%, 32.6%)
		Kb	Sequencing	31	90.3% (75.1%, 96.7%)	9.7% (3.3%, 24.9%)
		Kb	Composite	31	93.5% (78.2%, 94.6%)	6.5% (5.4%, 21.8%)
Loss	All	Marker	Sequencing	158	76.6% (69.4%, 82.5%)	23.4% (17.5%, 30.6%)
		Kb	Sequencing	157	74.5% (67.2%, 80.7%)	25.5% (19.3%, 32.8%)
	Non-HV*	Marker	Sequencing	104	65.4% (55.8%, 73.8%)	34.6% (26.2%, 44.2%)
		Marker	Composite	103	90.9% (87.8%, 91.2%)	9.1% (8.8%, 12.2%)
		Kb	Sequencing	104	61.5% (51.9%, 70.3%)	38.5% (29.7%, 48.1%)
		Kb	Composite	103	91.6% (88.3%, 92.4%)	8.4% (7.6%, 11.7%)

* CNVs located outside Affymetrix-defined hypervariable regions

**95%CI calculated using the Wilson score method. A modified Wilson's method with adjustment for the effects of weighted and stratified sampling and sampling from finite populations was used to calculate the 95% CI for markers.

** FPR=Pr(comparator method≠Gain or loss | CytoScan Dx Assay=Gain or loss) in this context is 1-Agreement rather than the conventional 1-specificity, i.e. FPR = 1-Agreement, where agreement is defined as TP/(TP+FP), with sequencing defined as truth. Agreement can also be referred to as PPV.

2. DNA extraction study:

   Three common genomic DNA (gDNA) extraction methods were evaluated for the isolation of gDNA from peripheral blood for this test. Twenty-four (24) normal blood samples with no known chromosomal abnormalities and 8 spike-in cell lines with 16 known aberrant regions were utilized in the study. To establish simulated aberrant, each cell line was cultured, harvested, counted and re-suspended in leukocyte-depleted blood to a final concentration of $8x10^6$ cells/mL. The gDNA was extracted using 3 different methods and tested on CytoScan Dx Assay. DNA sample QC, Array QC metrics and performance metrics were evaluated against critical limits. No obvious quality issue was noticed by testing DNA concentration, $A_{260/280}$ purity ratio and DNA integrity on gel. All samples passed assay performance assessment criteria, as well as the in-process and array-based QC criteria. CytoScan performance was not affected by choice of sample DNA extraction method.

3. Thermal cycler study:

   To evaluate the effect of thermal cycler on CytoScan Dx assay performance, 4 different commercially available thermal cyclers were used to test 24 cell line gDNA samples obtained from Coriell and ATCC. 24 samples were run on each of 4 thermal cyclers and the ability to determine copy number state compared. Array QC (SNPQC $\geq$ 15 and MAPD $\leq$ 0.25) metrics, copy number gain or loss state were evaluated. Pairwise confidence intervals were used to compare 4 thermal cycler models and assess

differences in performance for each thermal cycler model. For all performance measures for all pairwise thermal cycler comparisons, confidence intervals were overlapping and contained 0. This study met the sponsor's acceptance criteria and demonstrated that any commercially available thermal cycler would be adequate for use with the CytoScan Dx Assay.

Q. Proposed Labeling:

Labeling satisfies the requirements of 21 CFR 809.10, 21 CFR 801.109, including an appropriate prescription statement as required by 21 CFR 801.109(b), and the special controls for this type of device.

R. Identified Potential Risks and Required Mitigation Measures:

Identified Risks to Health	Required Mitigations
Inaccurate test results that provide false positive and false negative results can lead to improper patient management.	Special controls (1) and (2)
Failure to correctly interpret test results can lead to false positive and false negative results and accordingly improper patient management.	Special controls (1)(iii) and (2)

S. Benefit/Risk Analysis:

Summary of the Benefit(s)	Patients with developmental delay, intellectual disability, congenital anomalies, or dysmorphic features may potentially benefit in the intended use population by use of the device with assay results interpreted by healthcare professionals, board certified in clinical cytogenetics or molecular genetics.
Summary of the Risk(s)	Erroneous device results could adversely influence clinical interpretation and consultation for patients with developmental delay, intellectual disability, congenital anomalies, or dysmorphic features due to false negative or false positive results.
Summary of Other Factors	In addition to potential diagnostic yield study using the device, analytical performance evaluation and labeling along with requirements of special controls supports the intended use. De novo regulatory approach leverages device use by healthcare professionals, board certified in clinical cytogenetics or molecular genetics in conjunction with other clinical and diagnostic findings, consistent with professional standards of practice including confirmation by alternative methods, parental evaluation, clinical genetic evaluation, and counseling as appropriate.
Conclusions Do the probable benefits outweigh the probable risks?	Yes. Based on the potential diagnostic yield study for the diagnostic device along with review of the analytical performance and labeling, the probable benefits outweigh the probable risks.

T. Conclusion:

The information provided in this de novo submission is sufficient to classify this device into class II under regulation 21 CFR 866.5920. FDA believes that special controls, along with the applicable general controls, provide reasonable assurance of the safety and effectiveness of the device type. This device is classified under the following:

Product Code:	PFX
Device Type:	Postnatal chromosomal copy number variation detection system
Class:	II (special controls)
Regulation:	21 CFR 866.5920

(a) *Identification*. A postnatal chromosomal copy number variation detection system is a qualitative assay intended for the detection of copy number variations (CNVs) in genomic DNA obtained from whole blood in patients referred for chromosomal testing based on clinical presentation. It is intended for the detection of CNVs associated with developmental delay, intellectual disability, congenital anomalies, or dysmorphic features. Assay results are intended to be used in conjunction with other clinical and diagnostic findings, consistent with professional standards of practice, including confirmation by alternative methods, parental evaluation, clinical genetic evaluation, and counseling, as appropriate. Interpretation of assay results is intended to be performed by qualified healthcare professionals such as clinical cytogeneticists or molecular geneticists. This device is not intended to be used for standalone diagnostic purposes, pre-implantation or prenatal testing or screening, population screening, or for the detection of, or screening for, acquired or somatic genetic aberrations.

(b) *Classification*. Class II (special controls). The special controls for this device are:

(1) Design verification and validation must include the following information:
  (i) A detailed description of all components in the test system that includes:
    (A) A description of the assay components, array composition and layout, all required reagents, instrumentation, and equipment, including illustrations or photographs of non-standard equipment or methods;
    (B) A description of the design of the array in terms of chromosomal coverage and probe density for different regions;
    (C) An identification of the number of probes and size of the copy number variations reported at the lower range of the assay;
    (D) Detailed documentation of the device software, including standalone software applications and hardware-based devices that incorporate software;
    (E) Methodology and protocols for detecting copy number and visualizing results;
    (F) A description of the result outputs along with sample reports, and a description of any links to external databases provided by the device to the user or accessed by the device;

(G) Specifications for the methods to be used in specimen collection, extraction, including DNA criteria for DNA quality and quantity to perform the assay, and storage; and

(H) A description of appropriate internal and external controls that are recommended or provided. The description must identify those control elements that are incorporated into the testing procedure.

(ii) Information that demonstrates the performance characteristics of the system, including:

(A) Device reproducibility data generated, at a minimum, using three sites, with two operators at each site, for three non-consecutive days using at least three instruments. A well characterized panel of samples that provide a wide range of copy number variations (i.e., gains, losses, adequate size coverage across the range of sizes claimed by the device, adequate chromosomal coverage, challenging regions in the genome, copy number variations reported at the lower range of the assay, interstitial, subtelomeric, and pericentromeric rearrangements, aneuploidy, unbalanced translocations, mosaicism, and known syndromic regions) must be used. The results must be itemized for all copy number variations detected in each sample across all replicates and summarized in a tabular format stratified by size range and range of probe numbers for gains and losses separately and calculated for overall. The results must be analyzed using pairwise replicate agreement, and summarized as overall pairwise replicate agreement as well as pairwise replicate agreement conditional on replicates having a positive copy number state call (gains or losses), call rate, copy number variation size variation, and endpoint agreement;

(B) Device accuracy data using cell lines and clinical samples representing a variety of copy number variations and syndromes. In this analytical study, accuracy must be determined for every copy number variation detected in a particular sample. The accuracy data provided must include the copy number state determination and endpoint accuracy. The accuracy samples must cover different genomic variations across the genome (i.e., gains, losses, adequate copy number variation size coverage across the range of sizes claimed by the device, adequate chromosomal coverage, challenging regions in the genome, copy number variations reported at the lower range of the assay, interstitial, subtelomeric, and pericentromeric rearrangements, aneuploidy, unbalanced translocations, mosaicism, and known syndromic regions). Copy number variations identified by the device must be compared to comparator method(s). Agreement between the copy number variations detected by the array and the comparator must be summarized in a tabular format that includes the positive percent agreement and false positive rate stratified by size range and range of probe numbers for gains and losses separately and calculated for overall;

(C) Assay performance data for copy number variations reported at the lower range of the assay for both gains and losses;

(D) Device analytical sensitivity data, including DNA input and limit of detection for mosaicism, if applicable;

(E) Device analytical specificity data, including interference, carryover, and cross-contamination data;
(F) Device stability data, including real-time stability under various storage times, temperatures, and freeze-thaw conditions;
(G) Specimen matrix comparison data if more than one specimen type or anticoagulant can be tested with the device;
(H) Data that demonstrates the clinical validity, including diagnostic yield, of the device using a minimum of 800 retrospective clinical samples that were collected prospectively and obtained from three or more clinical laboratories, with results interpretation equally divided between two or more qualified healthcare professionals (e.g., cytogeneticists). Patients must be representative of the intended use population and not limited to common syndromes. Diagnostic yield data must be summarized in tabular format and stratified by the comparison methodologies. Data must be summarized in tabular format comparing interpretation of results, with description of reasons for variability in calls between the device and the standard of care methods. Data to support the accuracy of calls for known syndromes must be included; and
(I) Data that demonstrates device results when a minimum of 100 apparently healthy, phenotypically normal individuals are tested and interpreted by one or more cytogeneticists blinded to the patient status.

(iii) Identification of risk mitigation elements used by the device, including a description of all additional procedures, methods, and practices incorporated into the directions for use that mitigate risks associated with testing.

(2) The labeling required under § 809.10 of this chapter must include:
(i) A warning statement that the device is not intended to be used for standalone diagnostic purposes, pre-implantation or prenatal testing or screening, population screening, or for the detection of, or screening for, acquired or somatic genetic aberrations.
(ii) Limitations regarding the assay's performance with respect to validated copy number; variations reported at the lower range of the assay, stratified by size range and range of probe numbers for gains and losses separately. Limitations regarding problematic (hypervariable) regions; loss of heterozygosity; mosaicism; and inability to detect balanced translocations, as appropriate;
(iii) A warning statement that interpretation of assay results is intended to be performed by qualified healthcare professionals such as clinical cytogeneticists or molecular geneticists; and
(iv) A description of the performance studies performed in accordance with paragraph (b)(1)(ii) of this section and a summary of the results.