5X WGS assay as a more sensitive and cost-effective method to replace microarray in a diagnostic setting: Experience from the first 100 cases

Alka Chaubey[1], C. Alexander Valencia[1], Leah Stansberry[1], Zeqiang Ma[1], and Madhuri Hegde[1,]
[1]PerkinElmer Genomics, Duluth, GA

ABSTRACT

Introduction: Whole Genome Sequencing (WGS), is slowly picking up pace as a standard of care for individuals that need genetic/genomic testing. Numerous studies have reported the use of low resolution whole genome sequencing at an average of depth of 0.25X- 5X to detect genomic CNVs. 5X WGS (CNGnome™ test) developed by PerkinElmer Genomics, leverages the high throughput nature of NGS technologies along with bioinformatics tools to create a new standard for detecting large and intragenic copy number changes (CNV) and absence of heterozygosity throughout the genome without any probe bias. The goal of this 5X WGS methodology was to offer a more cost-effective option while maintaining or exceeding the diagnostic yield of a traditional microarray. Also, this test was validated on genomic DNA isolated from a variety of clinical specimen types, including saliva swabs, dried blood spots, and whole blood.
Methods: 100 clinical specimens from whole blood, dried blood spots, and saliva samples were used for performing the 5X WGS (CNGnome) test. The 5X WGS test methodology involves an automated DNA extraction using the Chemagic 360 workstation, followed by an automated library prep using the KAPA HyperPlus PCR-free library construction kit before direct sequencing using 2X150 bp reads on Illumina's NovaSeq 6000 system at a mean coverage of 5X in the target region. Alignment to the human reference genome (hg19) was performed and copy number variant (CNV) and AOH calls made using the NxClinical software v4.3 (BioDiscovery Inc., El Segundo, CA).
Results: The average global genome metrics for these samples was 5X depth with ~40% of the genome coverage distribution between 5-10X. The Multi-scale reference algorithm used in the NxClinical software utilized gender matched reference sets to assess ~250-600 million BAM reads for generation of virtual copy number (CN) probes and >15 million reads to generate B-allele frequency to detect AOH calls. Sequencing reads with a read base quality < 20 were rejected for the CNV analysis. The validation data set accuracy has been established at 100% for clinically relevant CNVs. Sixty-two cases had a negative result. Thirty-eight (38%) cases were identified with a clinically relevant CNV or reportable AOH or both. Fourteen cases (14%) were classified with known pathogenic CNVs and remaining abnormal cases either has variants of uncertain clinical significance or reportable AOH findings. For example, pathogenic heterozygous intragenic deletions involving the DMD (119 kb), NOTCH2 (27.8 kb), SHANK3 (183.1 kb) were identified. Interestingly, an intragenic homozygous deletion involving the TMCO1 gene (50.9 kb) consistent with the craniofacial dysmorphism, skeletal anomalies, and mental retardation syndrome was detected. All cytogenetically identifiable pathogenic CNVs such as, a 700 kb deletion involving the SQSTM1 gene, 1.9 Mb chr 5q35.2q35.3 microdeletion syndrome (involving the NSD1 gene), and 4.8 Mb chromosome 15q11q13 duplication syndrome was identified with base pair sequencing breakpoint accuracy. Due to the highest base pair resolution being awarded by this assay- a unique case of constitutional chromothripsis involving multiple pathogenic deletions of 7q was also identified by this test. The % AOH calculation data generated by this test matched the cases with known family history of consanguinity.
Conclusions: The performance parameters of this 5X test has been well documented and established as equal or better than those by a standard microarray test. This 5X WGS test offers a faster, cost-effective and a high-resolution method for the diagnosis of unbalanced translocations (CNVs), microdeletion/ microduplication syndromes and intragenic CNVs of clinical relevance. Our data on the first 100 cases paves the way for 5x WGS test to be the new standard of care in postnatal constitutional diagnostic testing and has the potential to replace the microarray technology.

INTRODUCTIONS AND METHODS

- **APPROACH:** 5X WGS (CNGnome, low resolution WGS, low pass WGS) was performed by using the KAPA HyperPlus PCR-free library construction kit and sequenced on Illumina NovaSeq™ 6000 (2 x 150 bp mode).
- **CLINICAL SAMPLES: USA, South America, Malaysia and India.**
- **AUTOMATION:** Unified automated workflow for high throughput and cost effective method for the 5X WGS assay
- **ANALYSIS and INTERPRETATION:** NxClinical 5.0 software (BioDiscovery, El Segundo, CA) was utilized for analysis, interpretations and reporting of CNVs and AOH from 5X WGS data. Multi Scale Reference algorithm is utilized by NxClinical 5.0.
- **5X WGS Data Quality in NxClinical 5.0 software:**
  - ~40% of the human genome had coverage distribution between 5x-10X depth
  - >200-250 million BAM reads were utilized for generation of log2 ratio and virtual copy number probes.
  - >15 million BAM reads were utilized to generate B allele frequency for AOH calls

RESULTS

- 100 clinical/diagnostic cases were subjected to 5x WGS or CNGNome test (aligned to human reference genome hg19).
- Sixty two (62%) cases had a negative result for copy number aberrations.
- Thirty eight (38%) cases were identified with a positive result including reportable CNVs or absence of heterozygosity or both.
- Fourteen (14%) cases were classified with known pathogenic CNVs of clinical significance.
- All cytogenetically identifiable microdeletions/microduplications (5q microdeletion, 15q11q13 duplication) were detected at approximate base pair sequence accuracy.
- Intragenic heterozygous deletions were identified in the *NOTCH2*, and *SHANK3* and intragenic homozygous deletions were identified in the *TMCO1* gene.
- **Dual Mendelian Diagnoses:** 2 cases with dual clinical diagnoses resulting from 2 pathogenic CNVs were identified.

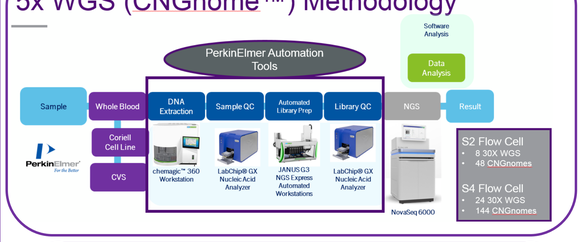
5X WGS detection of clinically relevant CNVs (hg19)

Sample ID	Result	Nomenclature	Classification
DS182258	DUAL DIAGNOSIS: Kleefstra syndrome and Duchenne Muscular Dystrophy	seq [hg19] Xp21.1(32688830_32808729)x0, 9q34.3(140198815_140598814)x1	Pathogenic
DS187239	Intragenic *NOTCH2* deletion- Hajdu-Cheney syndrome	seq [hg19] 1p12(120553275_120581101)x1	Likely Pathogenic
DS187003	Nephronophthisis (NPHP1) deletion syndrome-phenotype related	seq [hg19] 2q13(110841711_111028510)x1	Likely Pathogenic
DS183970	Xp22.31 microduplication (1.6 Mb)	seq [hg19] Xp22.31(6450180_8134280)x3	Likely Pathogenic
DS185913	Homozygous intragenic deletion of *TMCO1* (exons 5-7)	seq [hg19] 1q24.1(165671310_165722209)x0	Likely Pathogenic
DS183397	de novo intragenic SHANK3 deletion	seq [hg19] 22q13.33(51122267_51304566)x1 dn	Pathogenic
DS187606	Multiple CNVs in chr 9p- complex rearrangement	seq [hg19] 9p24.3p24.1(42115_8214109)x1, 9p24.1p21.1(8214110_30253840)x3, 9p21.1p13.1(30402040_38858214)x3	Pathogenic
DS181038_CNGNOM4	Chr 19p13 microdeletion	seq [hg19] 19p13.2(10140165_13379514)x3 dn	Pathogenic
DS181740_CNGNOM7	Constitutional chromothripsis (deletions) of Chr7q35q36.3	seq [hg19] 7q35q36.1(145448731_148600230)x1, 7q35q36.1(145448731_148600230)x1, 7q36.1(148843231_151338630)x1, 7q36.1q36.3(152440831_158744830)x1	Pathogenic
DS183730_CNGNOM15	Chr 5q35 microdeletion (Sotos) syndrome	seq [hg19] 5q35.2q35.3(175515930_177464729)x1	Pathogenic
DS183979	Chr 2q27 microdeletion syndrome	seq [hg19] 2q37.1q37.3(233643757_243199373)x1	Pathogenic
LPGNOM9: DS180649	Chr 15q11-q13 duplication syndrome	seq [hg19]15q11.2q13.1(23681146_28567945)x3	Pathogenic
DS180675	de novo Chr 5q35 microdeletion and 13q AOH	seq [hg19] 5q35.3(178,930,830_179,633,929)x1 dn, 13q31.1q31.3(86,587,765_92,648,539)x2 hmz	Pathogenic and Reportable AOH
19CT021947	DUAL DIAGNOSIS: Klinefelter syndrome and DiGeorge syndrome	seq[hg19] Xp22.22q28(16,227,930- 155,270,560)x2, 22q11.21(18884133_20307762)x1	Pathogenic

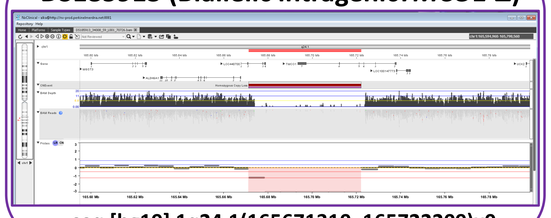
AUTOMATION & CASE EXAMPLES

PKIG Automation

5x WGS (CNGnome™) Methodology

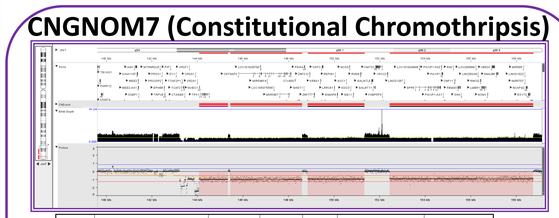


DS185913 (Biallelic intragenic*TMCO1* Δ)



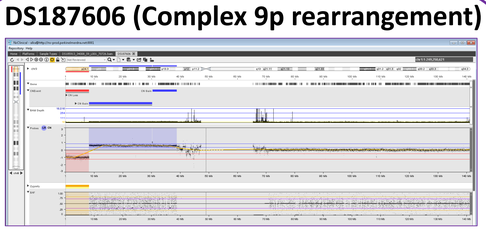
seq [hg19] 1q24.1(165671310_165722209)x0

CNGNOM7 (Constitutional Chromothripsis)



Event	Chromosome Region	Cytoband	Length (bp)	OMIM Genes Count	OMIM Morbid Map	Classification
CN Loss	chr7:144,080,031-145,348,130	7q35	1268100	2	*NOBOX, TPK1*	Pathogenic
CN Loss	chr7:145,448,731-148,600,230	7q35q36.1	3151500	3	*CNTNAP2, EZH2*	Pathogenic
CN Loss	chr7:148,843,231-151,338,630	7q36.1	2495400	38	*KCNH2, NOS3, CDK5, ASB10, PPKAG2*	Pathogenic
CN Loss	chr7:152,440,831-158,744,830	7q36.1q36.3	6304000	19	*DPP6, SHH, LMBR1, MNX1, DNAJB6, WDR60*	Pathogenic

DS187606 (Complex 9p rearrangement)



Event	Chromosome region	Cytoband	Length (bp)	Classification
CN Loss	chr9:42,115-8,214,109	9p24.3 - p24.1	8171995	Pathogenic
CN Gain	chr9:8,214,110-30,253,840	9p24.1 - p21.1	22039731	Pathogenic
CN Gain	chr9:30,402,040-38,858,214	9p21.1 - p13.1	8456175	Pathogenic

DISCUSSION/CONCLUSIONS

- **Diagnostic performance and clinical significance:** 5X WGS has been established as equal to or better than a standard microarray test.
- **Highest possible accurate breakpoint determination:** Approximate base pair resolution of deletions and duplications for microdeletions and microduplications and intragenic deletions/ duplications.
- **Fast TAT and reduced cost:** Simple automated workflow reduces the TAT and multiplexing on the NovaSeq is extremely cost effective.
- **AOH detection**: AOH has been validated for various degrees
- **Dual Mendelian Diagnosis:** Two cases with dual CNVs consistent with Mendelian diagnoses consistent with clinical phenotype were detected highlighting the power and potential of this 5X WGS assay.
- **WGS as standard of care:** 5X WGS is an effective method for the diagnosis of chromosomal diseases or microdeletion/microduplication syndromes and has the potential to replace a standard microarray test.

Disclosure: All the authors are employees of Perkin Elmer Genomics.