applied biosystems

Exon-Level Detection of Human Copy Number Variation Using High-Density DNA Oligonucleotide Arrays

Michael Shapero, Stephanie Nguyen, Benjamin Bolstad, Jeanette Schmidt, and Alan Roter
Thermo Fisher Scientific, 3450 Central Expressway, Santa Clara, CA, USA, 95051

ABSTRACT: Poster #775

There is increasing evidence that exon-level copy number changes are involved in a number of pathologies including intellectual disability and developmental delay, autism spectrum disorder (ASD), and congenital abnormalities [1]. Additionally, intragenic copy number variation (CNV) spanning only one or two exons are more prevalent in Mendelian disorders than previously suspected. Thus, there is an increasing need for the development of new genomic technologies that maximize CN coverage within genes and can be used in a clinical research setting with DNA samples that are also tested with whole-exome sequencing (WES).

The Applied Biosystems™ CytoScan XON™ Suite includes arrays, reagents, and software for cost-effective and streamlined analysis of exon-level CNVs and is primarily used to support germline samples (blood) that are also being analyzed with WES. This solution has been designed to cover the whole genome with increased coverage in 7,000 clinically relevant genes based on ClinVar classification and serves as a strong complement to mutation analysis performed by next-generation sequencing (NGS). The array uses whole-genome amplified DNA in conjunction with 6.85M empirically selected probes interrogating both polymorphic (0.3M) and nonpolymorphic (6.55M) regions of the genome. This allows for the combination of single-nucleotide polymorphism (SNP) genotypes and CN calls. The user-friendly Applied Biosystems™ Chromosome Analysis Suite (ChAS) software displays results by gene-level tiers based on relevant variants. This unique functionality allows visualization of CNVs for a targeted number of clinically relevant genes (based on ClinVar classification) or can expand the gene-level tiers for a whole-genome view.

System performance was evaluated using ~150 DNA samples composed of normal HapMap samples that include trios as well as unique aberration samples containing deletions or duplications spanning single or multiple exons in addition to regions of LOH. Performance was assessed for sample failure rate, sensitivity of CN calls, total call count in normal samples, LOH detection, and genotyping call rate and concordance. qPCR was used as an independent method to verify array-based CNV calls. Results suggest that the CytoScan XON Array can detect both simple and complex CNVs along with more precise boundary information, thereby enabling a more accurate analysis of their relationship to other genomic features. The addition of reliable exon-level deletion and duplication analysis via DNA microarrays should lead to an increase in disease-associated findings.

INTRODUCTION

Why is an exon-level CNV array an important research tool?

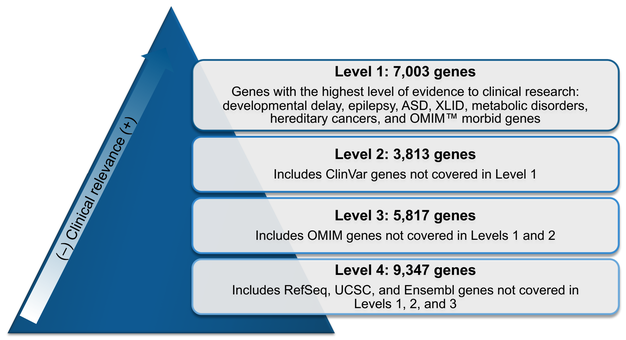
- Detection of deletions and duplications (CNVs) in conditions such as developmental delay
- Exon-level CNVs are important in many congenital pathologies
- Limitations of whole-exome sequencing (WES) in detecting small CNVs (<3 exons) due to poor coverage or highly variable regions
- An exome array is a complementary tool to confirm findings by WES
- An exome array can be helpful when a mutation is found by WES and a deletion/duplication is suspected on the other allele of the gene
- An exome array's SNP content allows detection of loss of heterozygosity/absence of heterozygosity (LOH/AOH)

With CytoScan XON Suite, you can:

- Comprehensively detect single-exon deletions and duplications in a cost-effective manner
- Complement NGS mutation analysis with reliable exon-level deletion and duplication detection
- Confirm CNV findings from alternative technologies
- Simplify and streamline sequence variant analysis

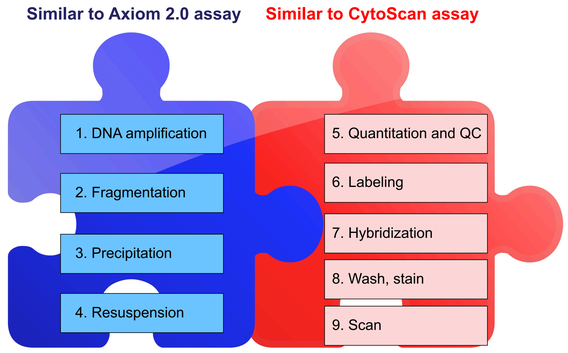
RESULTS

Figure 1. Description of gene-level tiers.



Gene-level tier analysis is incorporated into the user-friendly ChAS software, providing a novel and flexible approach, where the software displays results by gene-level tiers based on clinical research. This unique functionality allows visualization of CNVs for a targeted number of relevant genes (Level 1), or expands the gene-level tiers for a whole-genome view (Figure 1). Thus, the user determines the gene-level tiers of interest for analysis and reporting.

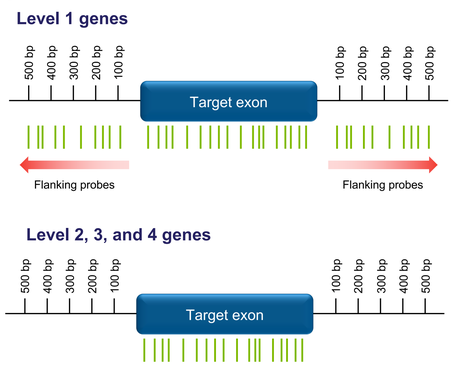
Figure 2. Assay workflow.



The workflow is a combination of two well-characterized and robust target preparation chemistries that leverage whole-genome amplification.

From QC to scanning of the arrays, the process is nearly identical to the CytoScan HD workflow. Thus there is no need to purchase new equipment for current users of CytoScan assays, which leads to easier adoption for labs.

Figure 3. Probe selection.



There are 6.85M probes that were empirically selected for whole-genome coverage including:

- 6.5M copy number probes
- 300K SNP probes for LOH/AOH analysis as well as duo/trio assessment and sample tracking

The sensitivity is 95% for the detection of exon-level CNVs. This was calculated from Level 1 genes.

There are 25,980 total genes that are covered:

- Full coverage: 21,844 genes
- Partial coverage: 4,136
- Genes for medical research (including cancer genes): 7,003
- Exon-level CNV detection with an average of 15 probes per call

Table 1. Performance estimates (median) by gene levels.

Level	Sensitivity (gains + losses)*	Reproducibility (gains + losses)*
1	96.1%	97.0%
2	92.7%	94.1%
3	92.6%	93.9%
4	86.0%	88.4%

* Computed at a region level.

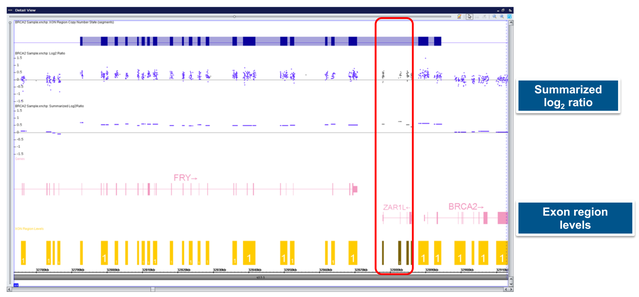
Table 2. qPCR verification of CytoScan XON Suite calls.

		Verified by SYBR™ Green Assay			Verified by TaqMan® Assay		
Call type	Total regions	Available data*	Yes	Percent validated	Available data*	Yes	Percent validated
Agilent Array	52	36	3	8%	34	9	26%
CytoScan XON Array	18	17	14	82%	11	8	72%

qPCR was run on 70 regions from 2 samples that were run on both arrays:

- A minimum of 1 primer was required to record a fold change (FC)
- If >1 primer, a FC where the majority of the primers agreed was used (if no agreement, region was excluded from analysis)
- 72–82% of calls on the CytoScan XON Array were confirmed, which was a much higher rate than the Agilent Array
- 1 call on the CytoScan XON Array set was not attempted; if this call did not confirm, the validation rate would drop to 67–77%—still well above the Agilent Array
- These results indicate that the CytoScan XON Array shows a lower false positive rate than an alternative technology

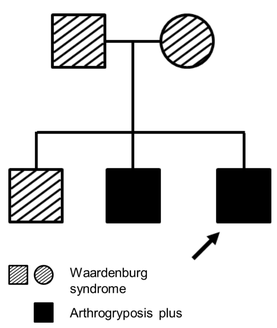
Figure 4. Visualization of regions with exon-level CNVs.



- ChAS software will highlight regions where there are more than one exon showing the same copy number change
- If analysis of only Level 1 genes has been selected, the $\log_2$ ratio of genes in other classifications will be grayed out

Figure 5. Waardenburg syndrome studied at Greenwood Genetic Center [2].

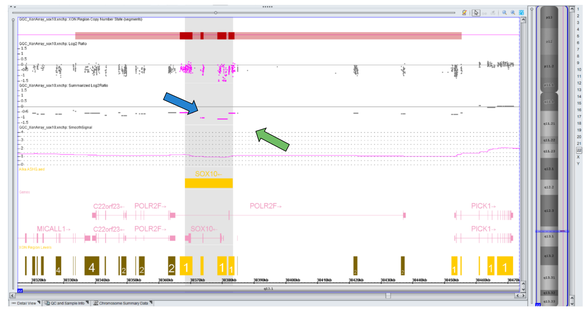
- US (stillborn fetus at 32 weeks): fixed limbs, scalp edema, micrognathia
- Exam (32 weeks):
  - White hair
  - Cleft palate
  - Absence of palmar and plantar creases
  - Dystopia canthorum
  - Four-limb pterygia
- Run on CytoScan HD Array
  - Large deletion seen across *SOX10* gene
- Waardenburg syndrome panel run:
  - No findings
- WES
  - No sequence variants
- Run on CytoScan XON Array
  - Additional *SOX10* findings



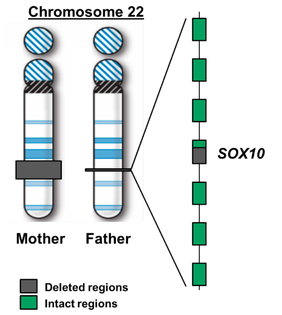
Waardenburg syndrome genes:

- **WS1:** *PAX3* (100%)
- **WS2:** *MITF* (~15%), *SOX10* (~15%), *EDN3/EDNRB* (~5%), *SNA12* (~5%)
- **WS3:** *PAX3*
- **WS4:** *SOX10* (50%), *EDN3/EDNRB* (~20%)

Figure 6. ChAS output of Waardenburg syndrome sample.



- Nested homozygous deletion in a heterozygous one
- qPCR confirmed deletion as well in **maternal** (larger deletion) and **paternal** (smaller deletion) samples
- Researchers anticipated that biallelic sequence variants of *SOX10* would result in similar pathogenic symptoms and observed that double heterozygosity for two different Waardenburg genes had not been reported
- Parents may not be aware of the risk of babies with multiple and potentially lethal anomalies
- Microarray analysis and sequencing of Waardenburg-associated genes may be required to identify couples at risk of complex birth defects



CONCLUSIONS

The CytoScan XON Suite is an exon-level copy number solution that provides the sensitivity and flexibility required to improve and complement the analysis of important variants for clinical research. It is designed to cover the whole genome, with an increased focus on 7,000 clinically relevant genes defined using ClinVar and ClinGen classifications. The CytoScan XON Suite is a full solution composed of the assay, array, and analysis software.

- Sensitive exon-level copy number and superior coverage across the whole genome
- Smart design improves resolution in key genes
- Reporting flexibility with gene panel or gene-level tier analysis
- Find out more at **thermofisher.com/microarrays**

REFERENCES

1. Miller DT, Adam MP, Aradhya S et al. (2010) Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86:749–764.
2. Stevenson RE, Vincent V, Spellicy CJ et al. (2018) Biallelic deletions of the Waardenburg II syndrome gene, *SOX10*, cause a recognizable arthrogryposis syndrome. *Am J Med Genet A* 176:1968–1971.

ACKNOWLEDGEMENTS

We would like to thank Chuan Chen, Vicky Huynh, Renee Zamora, and Gargi Mamtora for their contributions in generating lab data. We also would like to thank John Burrill, Patrick Weaver, Bina Shanbhag, and Yuan Lin for their contributions.

TRADEMARKS/LICENSING

ThermoFisher SCIENTIFIC