# Diagnostic Exome Sequencing Identifies a Homozygous Whole-Gene Deletion of *DPY19L2* that Was Not Detected by a High-Density Single Nucleotide Polymorphism (SNP) Array

Samin A. Sajan[1], Sheila Saliganan[1], Katherine L. Helbig[1], Brady Barrows[2], Rocio Martinez[2], Zöe Powis[1], Rebecca Burns[3], Luis Rohena[3], Jamie A. Massie[4], Elizabeth Spiteri[1], Wendy A. Alcaraz[1]

[1]Department of Clinical Genomics, [2]Clinical Laboratory Department, Ambry Genetics, Aliso Viejo, CA
[3]Departments of Pediatrics and Obstetrics and Gynecology, San Antonio Military Medical Center, San Antonio, TX
[4]Department of Obstetrics and Gynecology, Madigan Army Medical Center, Tacoma, WA

## BACKGROUND

- While both exome sequencing and high-density single nucleotide polymorphism (SNP) arrays are capable of detecting copy number variants (CNVs), neither methodology is comprehensive and each can miss certain regions of the genome.
- In a clinical setting, SNP array testing usually precedes exome sequencing primarily due to its lower cost and ability to reliably detect large CNVs that are at least 100 kilobases (kb) in length.
- It is therefore important for SNP arrays to be able to detect, at the very least, known clinically relevant CNVs that meet the length cutoff so as to minimize additional unnecessary investigations towards a genetic diagnosis.

## CASE REPORT

- We present a case example of a 38 years old male with globozoospermia, a spermatogenic defect that can cause male infertility due to acrosome-less sperm, who pursued diagnostic exome sequencing at our clinical genetic testing laboratory.
- The proband had no other presentations nor a family history of infertility, and was sent for exome sequencing as a proband-only case.
- The proband has a 7 month old daughter born via *in-vitro* fertilization (IVF) using his sperm.
- Prior to exome sequencing the proband had a negative result from the Affymetrix CytoscanHD Array from LabCorp, an array that contains 743,000 SNPs and 1,953,000 non-polymorphic copy number probes.
- The ordering clinician suspected several genes implicated in spermatogenic defects including *DPY19L2, SPATA16, AURKC, CATSPER1, KLHL10, HR5A1, SEPT12, SYCP3, USP9Y*.

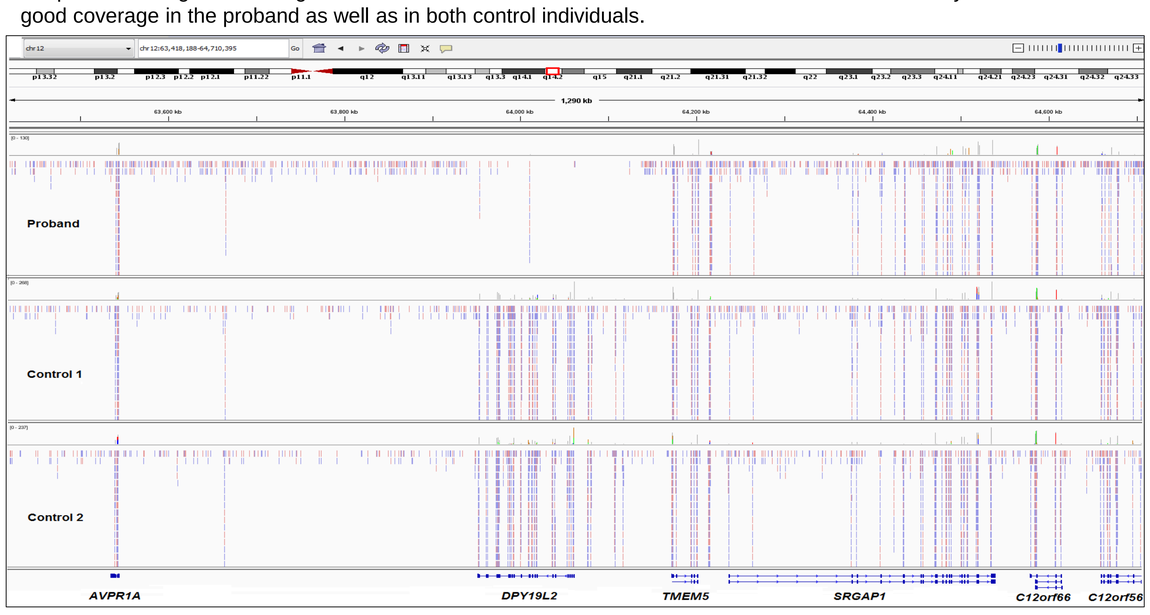
## METHODS

- Diagnostic whole exome sequencing was performed on the proband as previously described (Farwell KD, *et al*. 2014) with the exception of the capture reagent used which was IDT xGen Exome Research Panel V1.0. Approximately 97% of the proband's exome was covered at 20x or higher.
- Population frequency cutoffs for variant filtering included 0.1% for dominant, 1% for recessive, and 0.2% for incomplete penetrance models. The latter model included variants that are found in the Human Gene Mutation Database (HGMD), those that cause truncations such as frameshift, nonsense, and canonical splice, and those found in known imprinted disorder-causing genes. Population databases used were 1000 Genomes Project, Exome Sequencing Project (ESP), and Exome and Genome Aggregation Consortia (ExAC and gnomAD, respectively).

## RESULTS

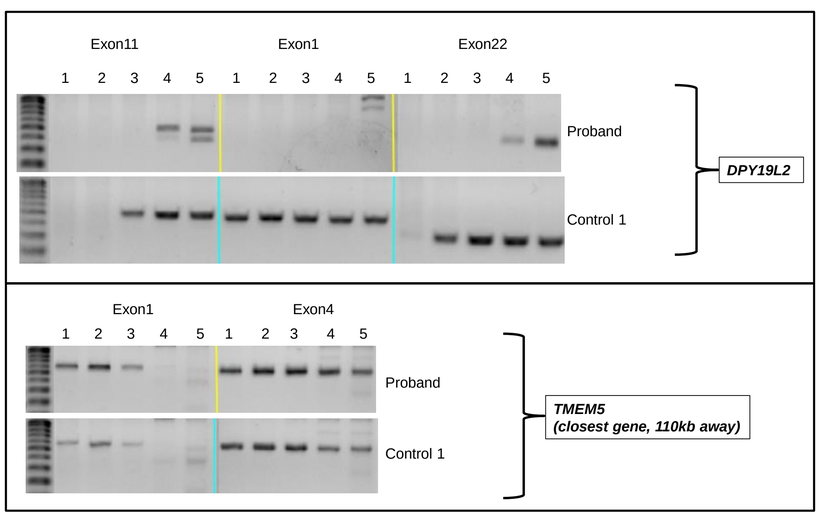
- Exome data processing and filtering resulted in 51 individual variants in 31 genes, none of which were clinically relevant for the proband's presentation.
- Manual inspection of coverage of genes of interest showed an almost complete lack of coverage in *DPY19L2* indicative of a homozygous whole gene deletion (Figure 1).

**Figure 1.** Integrated Genome Browser (IGV) view showing an almost complete absence of coverage across *DPY19L2* in the proband but good coverage in two unrelated control individuals. Genes in the immediate vicinity of *DPY19L2* show good coverage in the proband as well as in both control individuals.




- Polymerase chain reaction (PCR) of select exons of *DPY19L2* as well as the nearby *TMEM5* gene in the proband and an unrelated control individual confirmed that the proband had a homozygous deletion of *DPY19L2* (Figure 2).

**Figure 2.** PCR showing homozygous deletion of *DPY19L2* in the proband. Exons 11, 1, and 22 of *DPY19L2* (upper panel) as well as exons 1 and 4 of the nearby *TMEM5* gene (lower panel) were amplified by PCR using genomic DNA of the proband and an unrelated control individual using standard methods. *DPY19L2* failed to amplify in the proband but amplified well in the control whereas *TMEM5* amplified in both, suggesting that the lack of amplification of the former in the proband was not due to DNA quality issues. *DPY19L2* has four pseudogenes which most likely account for the bands seen in some of the proband's lanes in the upper panel. The lane numbers represent different primer annealing temperatures: Lane 1=65ºC; Lane 2=63.4ºC; Lane 3=60.9ºC; Lane 4=57.1ºC; Lane 5=52.26ºC.

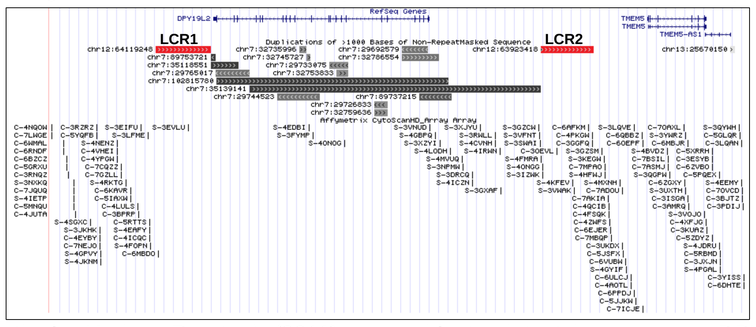



## INTERPRETATION

- Homozygous whole gene deletions of *DPY19L2* are common in men with globozoospermia (Elinati E, *et al*. 2012).
- These deletions are due to 28kb low-copy repeats (LCR1 and LCR2) with ~97% sequence identity on either side of *DPY19L2* that can cause non-allelic homologous recombination resulting in a deletion of approximately 200kb (Figure 3).
- Our proband's deletion was missed by the Affymetrix CytoscanHD Array due to poor probe coverage between LCR1 and LCR2. This region contains only 3 intragenic SNPs in *DPY19L2* i.e. 1 marker/36,666 base-pairs and a total of 27 probes in the region between and including LCR1 and LCR2 i.e. 1 marker/7,407 base-pairs (Figure 3). This is in contrast with the average coverage of OMIM genes on this array which is 1 marker/723 base-pairs.

### Detecting DPY19L2 deletion is influenced by the type of array, probe coverage, and analysis methods

**Figure 3.** UCSC Genome Browser view showing segmental duplications (grey/red bars) as well as individual probes from the Affymetrix CytoscanHD Array in the chr12:63,870,000-64,223,568 region (hg19). Segmental duplications in red are LCR1 and LCR2 which are responsible for approximately 200kb homozygous whole gene deletion of *DPY19L2* in globozoospermia. Overall probe coverage between LCR1 and LCR2 is poor, thus this deletion was missed in our proband.




- Interestingly, Harbuz R, *et al*. (2011) were able to detect homozygous *DPY19L2* deletions in patients with globozoospermia by aray comparative genomic hybridization (aCGH) using the Agilent 180K Human Genome Microarray (Agilent Technologies, Santa Clara, CA, USA) even though there were only three probes in the region (Figure 4 of their paper). They also used the Affymetrix 250K StyI SNP array which contains four SNPs in the region between LCR1 and LCR2 but they had to manually inspect the genotypes because only one SNP gave the expected results. Additionally, quantitative analysis of this region using Illumina Infinium II HumanHap 550 BeadChip SNP arrays allowed the detection of this deletion.
- This suggests that arrays can detect this deletion in-spite of low coverage, and that some manual examination of the data may also be warranted.

## TAKE-HOME POINTS

- SNP array testing costs less than, and is usually done prior to exome sequencing to detect pathogenic copy number variations (CNVs). However, SNP arrays can sometimes miss large CNVs due to poor probe coverage.
- An adult male with globozoospermia (male infertility) had previous negative SNP array test results and subsequently underwent exome sequencing which identified a homozygous whole-gene deletion of *DPY19L2* which is the cause of his infertility.
- This demonstrates that exome sequencing can detect large CNVs missed by SNP arrays and that some regions of the genome known to harbor pathogenic CNVs may need to be better represented on SNP arrays.
- In some cases, such as the proband described here, aCGH may be a better option for detecting pathogenic CNVs. Additionally, based on the patients' presentation, clinicians should be encouraged to provide genes/regions of interest to clinical laboratories that provide SNP/aCGH tests so that they can be examined manually for any CNVs that do not pass the standard cutoff thresholds (such as those in regions with segmental duplications). It will also be beneficial if the clinicians ensured that most, if not all, genes of interest have adequate probe coverage on a given array platform before ordering the test.

## REFERENCES

- Elinati E, *et al*. (2012) *Hum Mol Genet* **21**(16):3695-702.
- Farwell KD, *et al*. (2014) *Genet Med* **17** (7):578-86.
- Harbuz R, *et al*. (2011) *Am J Hum Genet* **88**(3):351-361.