Next Generation Sequencing (NGS) metrics following DOP-PCR whole genome amplification (WGA) of single and multi-cell samples for PGS

Jasper, MJ.[1]; Hardy, T.[3]; Pinese, M.[2]; McCabe, M.[2]; Ledger, W.[3] and Fraser, M.[1]

[1]Reproductive Health Science Ltd, SA, Australia
[2]Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, NSW, Australia
[3]School of Women's & Children's Health, University of NSW, NSW, Australia

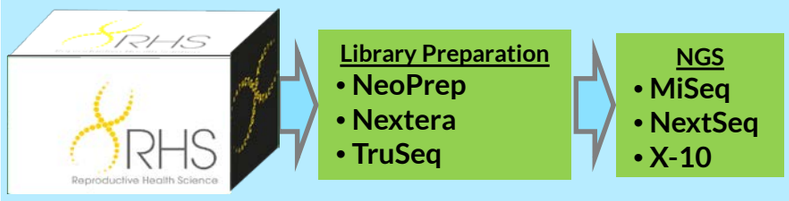
Background Information

Whole genome amplification (WGA) is often used to generate sufficient DNA for downstream analysis.
The aim of this study was to compare Next Generation Sequencing (NGS) workflows and aligned read data metrics from a range of NGS platforms as models for Pre-implantation Genetic Screening (PGS) and Preimplantation Genetic Diagnosis (PGD).
Comparisons were made using single cell and 5-cell aliquots amplified utilizing the Reproductive Health Science Ltd proprietary DOP-PCR based WGA as described in the DOPlify™ and EmbryoCellect™ kits.
Cell aneuploidy and mitochondrial DNA load was determined in addition to the evaluation of allele-drop out and the presence of genes of clinical interest.

Materials & Methods

Single cell and 5-cell aliquots sorted from euploid and aneuploid cell lines (Coriell Institute) were subjected to WGA using DOPlify™ (FIG 1; RHS; Australia). NeoPrep, Nextera or TruSeq libraries were prepared from the WGA products and unamplified gDNA samples from the same source and subsequently sequenced (paired-end) on either a MiSeq, NextSeq or X-10 platform according to standard protocol (Illumina). The sequencing data was bioinformatically aligned to hg19.

DOPlify™ WORKFLOW



WGA & NGS Metrics

MiSeq sequenced NeoPrep DNA libraries generated an average of 0.3 million reads/sample (n=12; 75bp), whereas Nextera DNA libraries generated an average of 1.3 million reads/samples (n=49; 300bp). NextSeq sequenced Nextera libraries generated an average of 5.8 million reads/sample (n=23; 150bp). However, TruSeq DNA libraries sequenced using X-10 technology generated substantially more reads, approximately 400 million reads/sample (n=6; 75bp, FIG 2). For MiSeq and NextSeq NGS runs, reads mapping to individual chromosomes were binned into 1Mb bins. The mean ± STDEV of the number of reads per bin for 3 samples sequenced simultaneously for NeoPrep-MiSeq prepared samples was 132 ± 44, Nextera-MiSeq prepared data was 445 ± 167, and for Nextera-NextSeq samples was 2597 ± 489 reads.

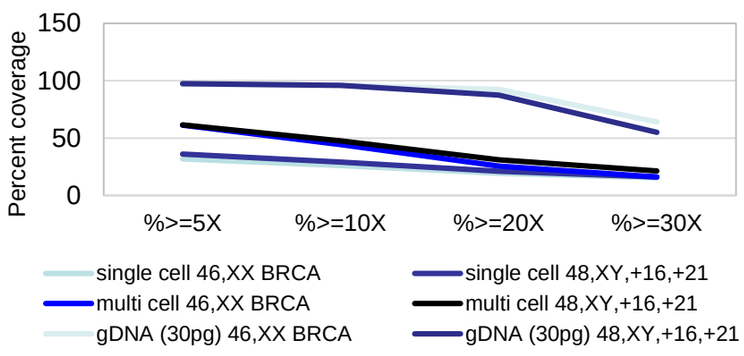


FIG 2. The percentage of the genome covered at >5x, >10x, >20x and >30x for 5-cell and 1 cell WGA DNA samples following X-10 sequencing.

Aneuploidy & mtDNA

Correct aneuploidy diagnoses were achieved for single cells carrying either trisomy 15 or 21 where a total of ~300,000 reads were obtained following MiSeq sequencing of Neoprep WGA DNA libraries. In addition, correct diagnoses of the sex chromosomes, including XXY were obtained (Fig 3A). Additionally, it was observed that the mean percent of total reads mapping to the mtDNA genome (Fig 3B) was significantly higher in aneuploid single cells than in euploid single cells (Student's T-test; $p<0.01$; euploid single cell (n=23) mean ± STDEV 0.13 ± 0.08 and aneuploid single cell (n=24), mean ± STDEV 1.34 ± 1.91 n=24).

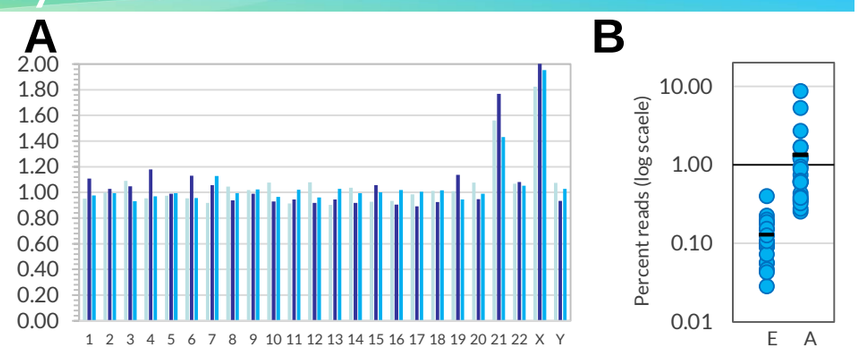


FIG 3. A) 48,XXY,+21 single cell ratios of mapped reads per chromosome normalized to an internal single cell reference. Each bar represents a single cell following MiSeq sequencing (n=3). B) Euploid (E) and aneuploidy (A) single cell mtDNA content determined using MiSeq NGS and Nextera libraries.

Allele Drop-Out

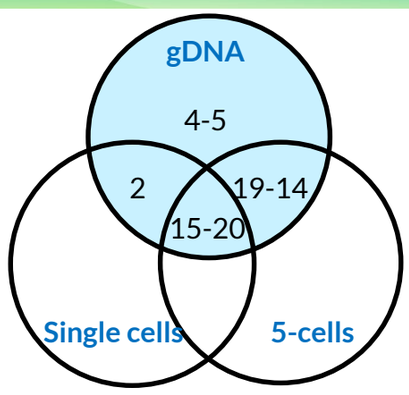


FIG 4. SNP's common to DNA template following X-10 sequencing of gDNA, single cell and 5-cell WGA DNA samples.

Allele drop out (ADO) is an important indicator of PCR efficiency in primer-specific PCR. WGA however is designed to amplify the genome and metrics to determine its robustness and efficiency are more complex to generate and assess. As NGS library preparation and sequencing parameters alone impact on sample complexity, targeted PCR of the WGA DNA would be a more accurate measure of ADO instead of NGS assessment. However, ADO was assessed using a 40 SNP panel described by Pakstis et al (2007). All 40 SNPs were sequenced in 2 gDNA samples to an average depth of 36 ± 6 reads (Euploid) and 33 ± 7 reads (48,XY,+16,+21; GM04435), equivalent to >23x using X-10 technology. The number of SNP's sequenced correlated to starting DNA template and for 5-cell samples 90%-83% SNP's were sequenced, for single cell WGA samples 40%-53% SNP's were sequenced respectively (FIG 4).

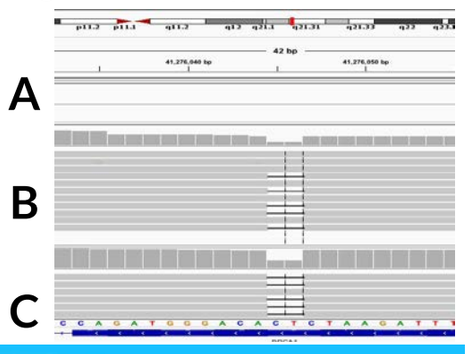
Genes of Clinical Interest

The single cell WGA sample reads mapping to genes of clinical significance contained in an ISCA Panel (377 genes; International Standards for Cytogenomic Arrays Consortium) and Newborn Screening Panel (2013) (72 genes; American College of Medical Genetics; Genetic Med 2006) were evaluated. As the total reads per sample increases due to both the increased capacity of the NGS platform and the reduction in the number of samples multiplexed in the sequencing run, so gene detection increased (Table 1).
X-10 sequencing was utilized to determine the suitability of the WGA for small variant changes in genes of clinical interest. A pathogenic variant (2bp deletion 185delAG in exon 3) was directly sequenced in the 5-cell and gDNA BRCA1-positive sample (GM14090B) confirming that the deletion was heterozygous, but was not sequenced in the single cell sample (FIG 5). Detection of genes of clinical interest requires either an enrichment approach or expensive deep sequencing workflows with limited multiplexing on large format sequencers.

Table 1. Number of genes where at least one read mapped.

Sample	#	ISCA	Newborn Screen
SC MiSeq (n=4)	48	277 ± 29 (73%)	49 ± 3 (68%)
gDNA MiSeq	48	324 (86%)	70 (97%)
SC NextSeq (n=5)	24	338 ± 3.4 (90%)	72 (100%)
gDNA NextSeq	24	374 (99%)	72 (100%)

FIG 5. Direct detection of *BRCA1* mutation. The 2bp deletion can be visualized in the reference genome using IGV (Broad Institute), as well as in B) 5-cell WGA DNA sample and C) gDNA sample, but not in the A) single cell WGA DNA sample.



Combined PGD+PGS – Targeted Gene Enrichment

Reliable PGD requires greater breadth and depth of coverage than PGS, and evaluation of the specific gene of interest, in addition to adjacent SNP's if possible. These factors are a key indicator for PGS + PGD success. By multiplexing WGA primers and gene specific PCR primers, RHS have developed a method for aneuploidy and parallel assessment of the same WGA product for single gene disorder diagnosis and evaluation (Provisional Patent application filed). Enrichment of the targeted sequence in a gene specific PCR is evident (FIG 6A). Enrichment of the 3' and 5' ends of the desired PGD amplicon, corresponding to the 75bp fragment size sequenced is also evident in the NGS data (FIG 6B).

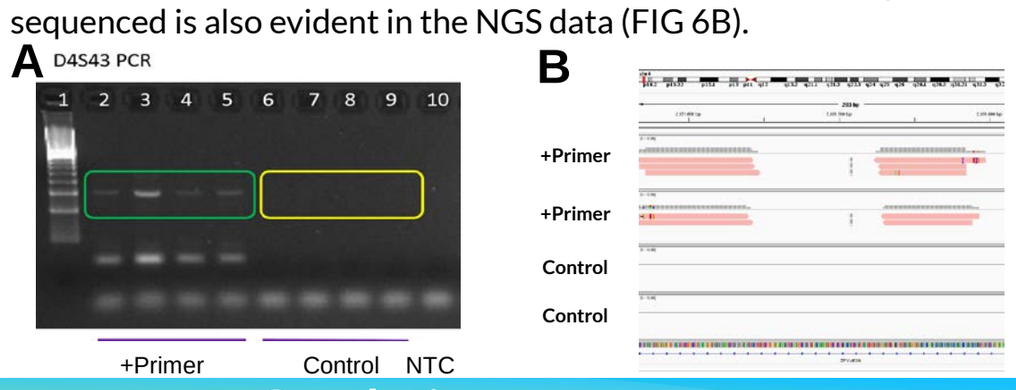


FIG 6. Combined PGD+PGS. A) Enrichment of targeted DNA during WGA following the addition of sequence specific primers. D4S43 amplicon size; 283bp. DNA ladder (DMW-100M; Geneworks), B) Visualization of D4S43 amplicon size; 283bp in data generated from MiSeq sequencing of NeoPrep DNA libraries constructed from single cell WGA DNA viewed in IGV.

Conclusions

- DOPlify™ WGA technology simultaneously provides cell ploidy status and mitochondrial DNA load using NGS platforms.
- The amount of sequencing data available for interrogation is highly dependent on the number of samples multiplexed in a library, the read length and the platform configuration. While specific gene sequence is present in single cell and 5-cell WGA NGS data, gene enrichment is required to acquire a level of resolution suitable for PGS+PGD.
- DOPlify™ WGA technology is compatible with targeted gene enrichment for PGS+PGD.

Acknowledgements: Bioinformatics and technical assistance performed by K Warren and technical assistance provided by C Robinson, S Protopsaltis and M Brockman. Funding for X-10 NGS was provided to TH & WL by Illumina Pilot the Possibility and Royal Hospital for Women Foundation. **Conflict of Interest:** MJ and MF are shareholders and employees of Reproductive Health Science Ltd.