PG-Seq™ NGS Kit for Preimplantation Genetic Screening

Validation using Multi (5-10) Cells and Single Cells

from euploid and aneuploid cell lines

Introduction

Advances in next generation sequencing (NGS) technologies have gained considerable interest in genomic analysis, including in IVF embryo testing. Compared with array-based aneuploidy screening methodologies, NGS workflows offer potential improvements in accuracy, sensitivity and resolution for the detection of aneuploidy and copy number variants in embryos. RHS's PG-Seq™ kit has been developed to accurately detect whole chromosome aneuploidy along with structural rearrangements such as translocations and segmental errors using an efficient, streamlined workflow.

The PG-Seq™ kit contains

- DOPlify™, RHS's whole genome amplification (WGA) system which includes cell lysis and WGA reagents,
- library preparation reagents, indexes and
- data analysis software.

Each PG-Seq™ kit can process up to 96 samples in total, with 48 samples being multiplexed in each of 2 Illumina MiSeq runs. This is double the capacity of the current Illumina VeriSeq kit, significantly increasing throughput.

While several methods of whole genome amplification are commercially available, RHS' DOPlify™ has been independently validated to reliably amplify copy number variants down to 3Mb in size from 1, 3 or 5 cell templates, the only commercially available kit with equivalent accuracy to PicoPlex[1].

The PG-Seq™ library preparation technology prepares DOPlify™ amplified samples ready for clustering on an Illumina MiSeq flow cell through enzymatic fragmentation, end repair and A-tailing and adaptor ligation. Using an increased read length of 75bp, PG-Seq™ can multiplex 48 samples in a single NGS run, thereby doubling throughput and providing labour and cost efficiencies[2].

Recent data suggest an association between mitochondrial genome load in euploid embryos and implantation potential[3]. DOPlify™ and PG-Seq™ provide 100% mitochondrial genome coverage, which not only provides confidence in mitochondrial load determination but also uniquely allows unique embryo identification through the use of informative Single Nucleotide Variants (SNVs) to categorize embryos based on maternal origin (RHS, unpublished). The data analysis software is user-friendly and has been designed with advanced calling algorithms for whole chromosome aneuploidy and CNVs. The software is easy to navigate and provides the option to compare multiple samples at once. The software also allows customised setting changes for advanced experienced users to undertake higher resolution and more detailed analysis.

The aim of this study was to fully validate the performance of PG-Seq™ on single cells and 5 cell aliquots using cell lines with known ploidy, including euploid, single, and double trisomies. Cell line aberrations of a 31Mb gain on chromosome 3 and a 7Mb gain on 21 (GM09552), 7Mb loss and 31Mb gain on chromosome 8 (GM14485) and a 16Mb loss on 13 (GM07312) were also included. The use of single cell aliquots is a model of blastomere biopsy and the 5 cell aliquots are a model of trophectoderm biopsy.

Materials and Methods

Isolation and preparation of samples

Whole chromosome aneuploidy cell lines of the following karyotypes; 47,XY,+9 (GM09286), 47,XY,+13 (GM02948), 47,XY,+15 (GM07189), 47,XX,+18 (GM00143), 47,XX,+21 (GM04592), 48,XXY,+21 (GM04965), 48,XY,+16,+21 (GM04435) were purchased from Coriell Biorepository (USA). An additional fibroblast culture established from

chorionic villi containing a trisomy 22 was obtained from the Cytogenetics Department of SA Pathology (Adelaide, Australia). This culture is referenced as WCH+22 (47,XY,+22).

To test the resolution of the PG-Seq™ workflow, cell lines containing segmental errors ranging from 7Mb to 31Mb (GM09552, GM14485, GM07312) were also purchased from Coriell Biorepository. All cells lines were cultured to semi-confluence for 2 days prior to being harvested using trypsin-EDTA. Peripheral lymphocytes from a male and female of proven fertility were also isolated from peripheral blood mononuclear cells (PBMCs) as previously described[2]. These lymphocytes were designated 46,XY and 46,XX euploid cells, respectively.

Single and 5 cell aliquots of cultured fibroblasts or lymphocytes were manually isolated as described previously[4]. Briefly, cells were washed in droplets on a microscope slide and individually picked or manually grouped together in 5 cell groups, placed into 0.5ml PCR tubes in a volume of not more than 2µl and stored at –20°C until required.

Karyotyping by G-banding was performed on several cell lines from the same passage used in this study by the Women's and Children's Hospital (Australia) to ensure clonality. The results of karyotyping are provided in Table 5.

Aneuploidy screening using the PG-Seq™ kit

Cell lysis and whole genome amplification of the cell samples was performed according to the method described in the RHS PG-Seq™ TDS (version 1.0). Following WGA, the amplicons were run on an agarose gel (Figure 1) to confirm amplification success. The amplicons were then purified using 0.8x Ampure XP beads (Beckman Coulter, USA) according to standard protocol. Samples were quantified using a Qubit and 200ng was seeded into the library preparation, which was performed according to the standard PG-Seq™ protocol. Briefly, each sample was enzymatically fragmented followed by End Repair and A-tailing, Adapter Ligation and PCR. Final libraries were again quantified using a Qubit and analysed for size distribution using a 2200 TapeStation (Agilent, USA) (Figure 2). Equal nM concentrations of all 48 libraries were combined to create a 4nM pool. This pool was denatured and diluted to a final concentration of 15pM according to the standard MiSeq Denature and Dilute libraries guide version 01 (Illumina, USA). The final 15pM pool was run on an Illumina MiSeq with single index reads of 1x75 using the MiSeq 150 Cycle Reagent Kit v3 (Illumina, USA).

Fastq files generated from the sequencer were aligned to hg19 using the application BWA on Basespace (Illumina, USA). The resultant BAM files were analysed using the PG-Seq™ software. The reference was created from the combined PG-Seq™ data of 35 samples from various cell lines sequenced during this validation study. Each result that passed all quality control (QC) measures (Table 1) was then compared to the expected karyotype of the cell line.

Agarose gel electrophoresis

Whole genome amplified samples were visualised after electrophoresis in 1% agarose gels in 0.5X TBE (Tris, borate, EDTA) pre-stained with GelRed (Biotium, USA) using an ultra-violet light box.

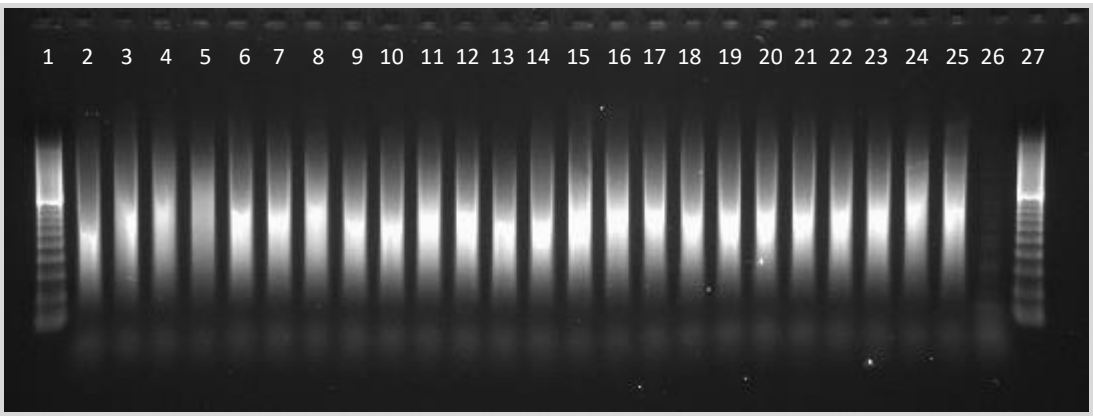

Figure 1. Example of WGA DNA. Lanes 1 and 27 contain a 100bp increment DNA size ladder (DMW-100M, GeneWorks, Australia), lanes 2-25 are WGA amplified 5 cell samples and lane 26 is a no template control (NTC) where no DNA was added to the reaction.

Table 1: Summary of Quality Control check points for PG-Seq™.

Protocol stage	Step	Quality control checks
WGA PCR	Agarose gel	Clearly visible smear of DNA amplification products in the range of 200 - 1000 bp.
Purification	Yield after 0.8x bead purification	> 5 ng/µL
Library Preparation	Final Library Fragment Size	200-500bp
	Final Library Yield	>20 ng/µL
Sequencing and Analysis	Cluster Density	1100-1600 K/mm
	% of clusters passing filter	>75%
	% of reads reaching Q30	>90%
	Mapped reads per sample	>300,000
	Nexus Quality Score	<0.0425

Quality Control Metrics

The QC metrics used in PG-Seq™ are based on RHS' experience amassed over the past 10+ years of product development and are used as indicators for confidence in the NGS result. The Nexus Quality Score is a measure of noise in the system, with lower quality scores reflecting smoother data sets. All steps after the WGA DNA purification can be repeated in case of QC failure by returning to the start of the failed step.

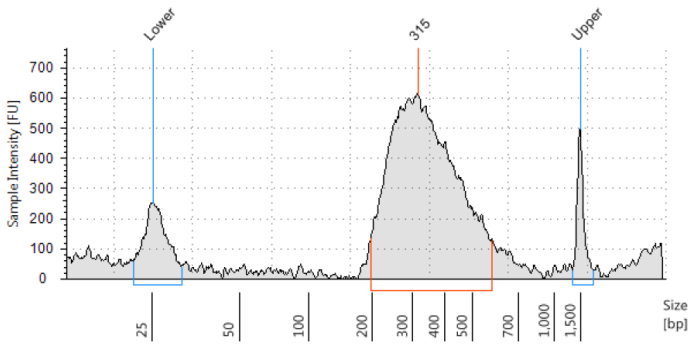

Figure 2. Agilent Tapestation tracing of fragmented and adapter ligated WGA products from PG-Seq™ confirming that the DNA has been fragmented to the target length of 200-600 bp in preparation for sequencing.

Whole Chromosome Aneuploidy Results

A total of 197 multi-cell samples and 226 single cell samples were subjected to WGA, library preparation and sequencing according to the PG-Seq™ kit TDS version 1.0. Where the logR ratio across the length of a chromosome was observed to be greater than 0.375 or less than -0.5, the sample was reported as aneuploid for that chromosome. A total of 9 5 cell samples and 34 single cell samples failed QC checks and were excluded from the analysis. The 5 cell aliquot and single cell results for each cell line are provided in Tables 2 and 3.

Table 2: 5 cell performance results

Cell Line ID – 5 cells	Karyotype	Number passed QC	Correct	Mosaic >60%
GM09286	47,XY,+9	20	19	1
GM02948B	47,XY,+13	21	21	0
GM07189	47,XY,+15	18	18	0
GM00143	47,XX,+18	20	20	0
WCH	47,XY,+22	18	17	1
GM04592	47,XX,+21	6	6	0
GM04965	48,XXY,+21	24	24	0
GM04435	48,XY,+16,+21	20	19	1
G00316	46,XX	21	21	0
G00317	46,XY	20	20	0
	Total	**188**	**185 (98.4%)**	**3 (1.60%)**

Table 3: Single Cell performance results

Cell Line ID – single cells	Karyotype	Number passed QC	Correct	False Positive	False Negative	False Negative + False Positive
GM09286	47,XY,+9	26	26	0	0	0
GM02948B	47,XY,+13	17	15	1	1	0
GM07189	47,XY,+15	21	21	0	0	0
GM00143	47,XX,+18	24	23	1	0	0
WCH	47,XY,+22	19	17	2	0	0
GM04592	47,XX,+21	4	4	0	0	0
GM04965	48,XXY,+21	21	20	1	0	0
GM04435	48,XY,+16,+21	19	17	1	0	1
G00316	46,XX	25	25	0	0	0
G00317	46,XY	16	16	0	0	0
	Total	**192**	**184 (95.8%)**	**6 (3.13%)**	**1 (0.52%)**	**1 (0.52%)**

One-hundred and eighty-five out of the 188 (98.4%) 5 cell samples that passed QC provided the correct results as determined by the PG-Seq™ protocol. The 3 incorrect results were mosaic for the copy number change that was expected from the cell line (samples were called mosaic when the logR ratio of a segment was 0.225 to 0.375 or 60% to 100% mosaic, see figure 3).

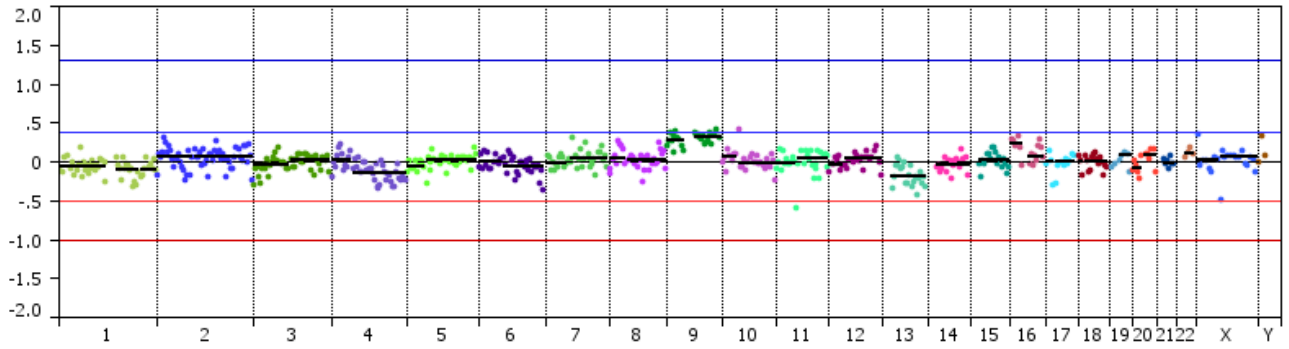

Figure 3: A mosaic result for GM09286 5 cell sample reported as false negative for aneuploidy for chromosome 9 due to the average segment LogR value being 0.293 (78% mosaic).

One-hundred and eighty-four out of 192 (95.8%) single cell samples that passed QC were found to produce the expected whole chromosome aneuploidy result. Of the 8 incorrect results, 6 detected the correct chromosome aneuploidy, however there was a further aneuploidy detected on another chromosome (false positive). There was 1 false negative result where the expected aneuploidy was not detected, however karyotyping has confirmed that this cell line (GM02948) is mosaic with only 88% of cells containing the expected trisomy 13. One result was false negative for the expected trisomy and false positive for other chromosomal aneuploidies. All 5 cells and single cells samples of the female and male euploid cells (G00316 and G00317 respectively) achieved 100% accuracy with no false positives.

The total number of chromosomes analysed in the 5 cell samples was 8,839; a total of 3 expected aneuploid chromosomes appeared to be mosaic. If the mosaic results are counted as false negatives, the results were 99.96% correct. Including the mosaic results as correct due to observed mosaicism in the cell lines, the results are 100% correct. The total number of chromosomes analysed in the single cell samples was 9,023; including the 6 false positive, 1 false negative and 1 false positive + false negative result, the chromosome results were 99.91% correct.

Example graphs for each cell line are provided in figures 4 to 13.

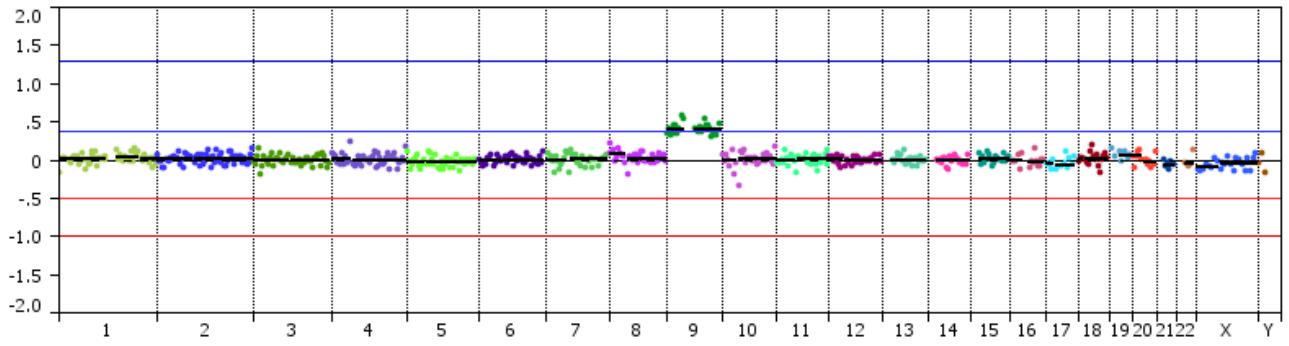

Figure 4: GM09286 47,XY,+9 single cell

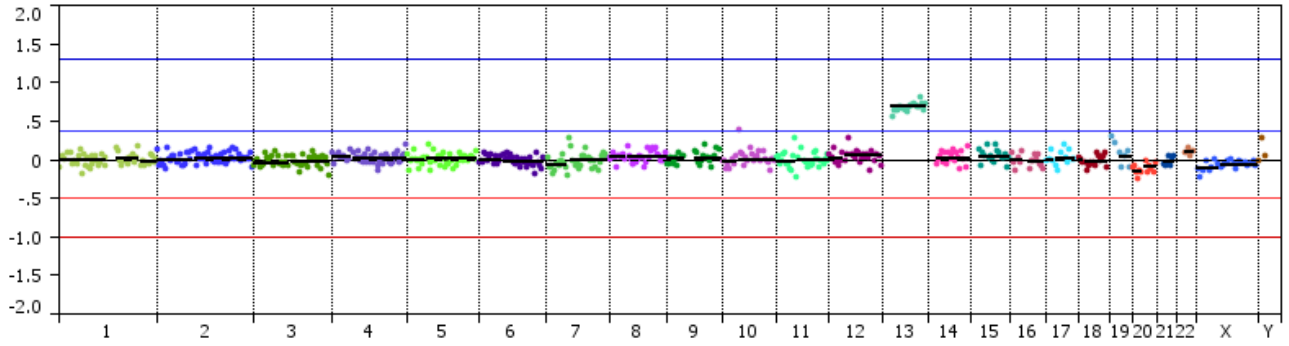

Figure 5: GM02948 47,XY,+13 5 cell

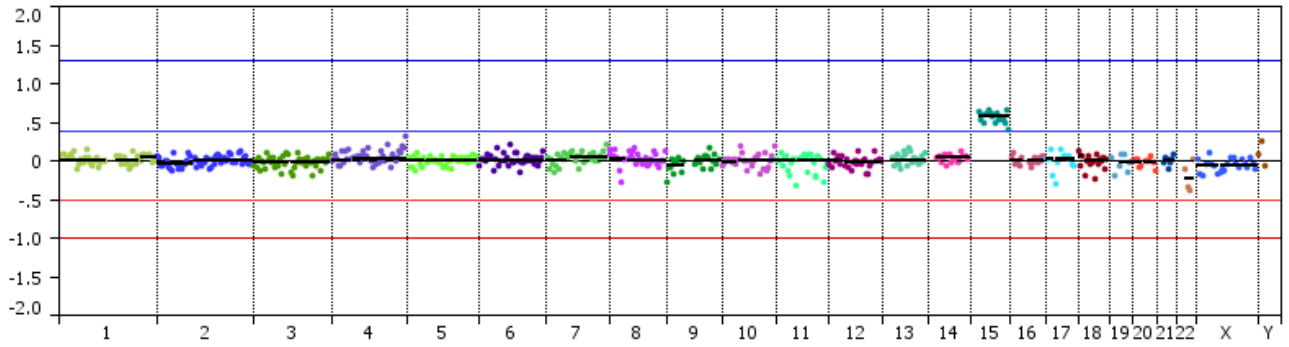

Figure 6: GM07189 47,XY,+15 single cell

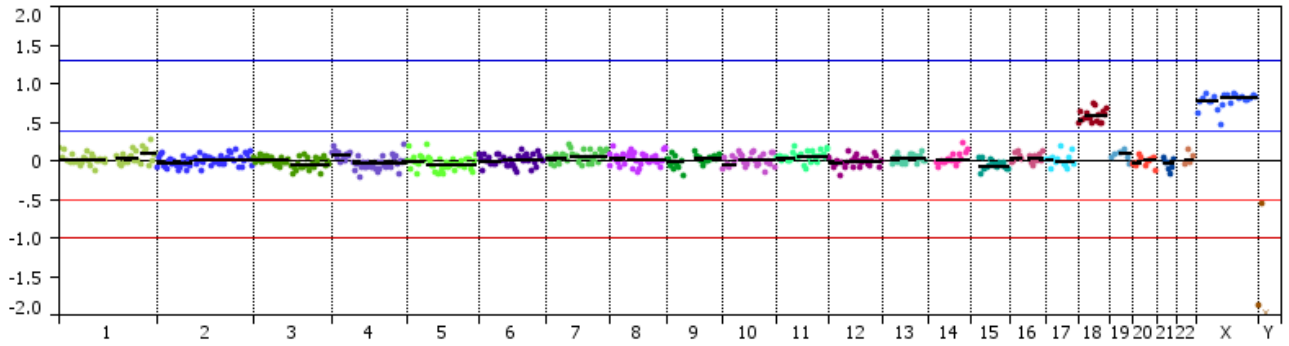

Figure 7: GM00143 47,XX,+18 single cell

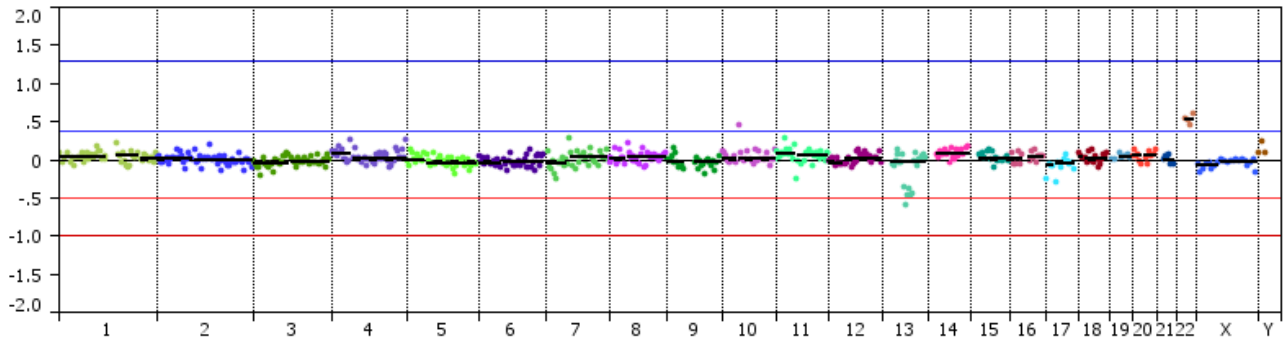

Figure 8: Clinical WCH 47,XY,+22 5 cell

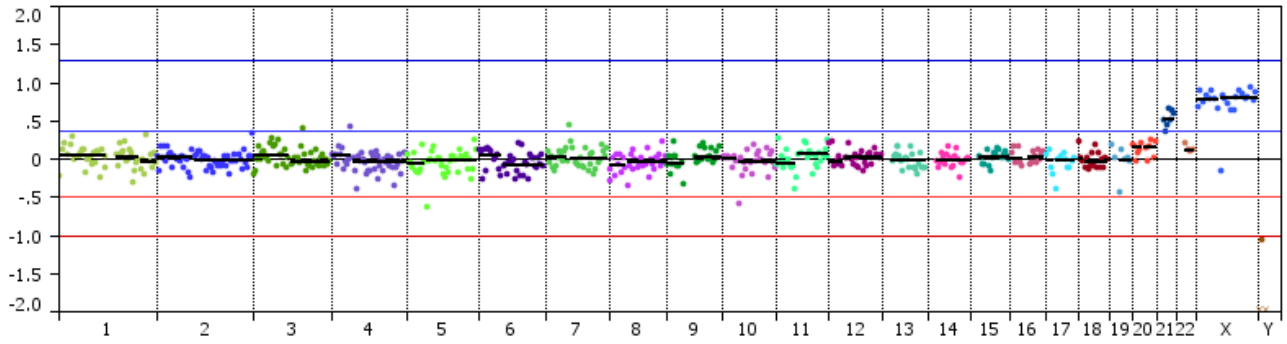

Figure 9: GM04592 47,XX,+21 single cell

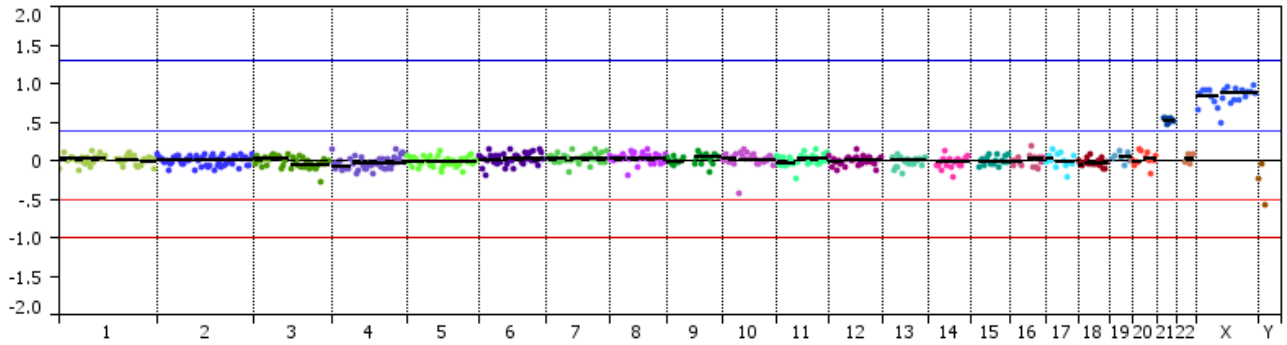

Figure 10: GM04965 48,XXY,+21 5 cell

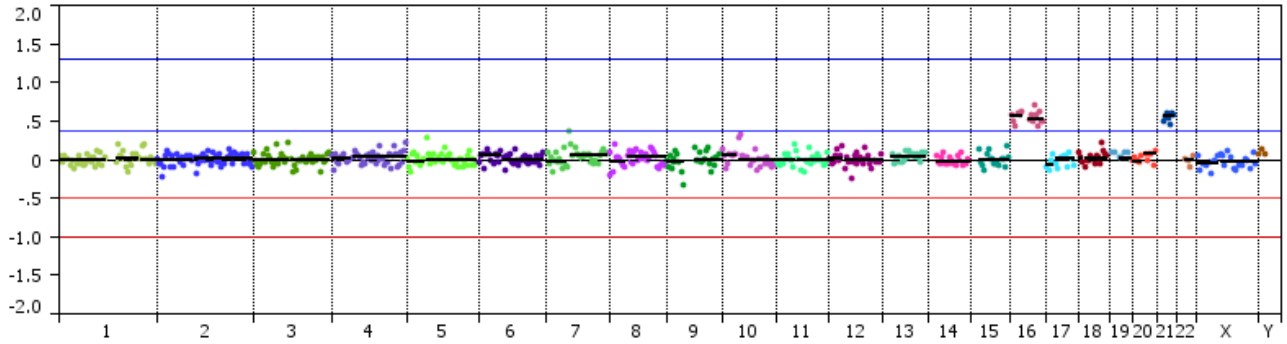

Figure 11: GM04435 48,XY,+16,+21 single cell

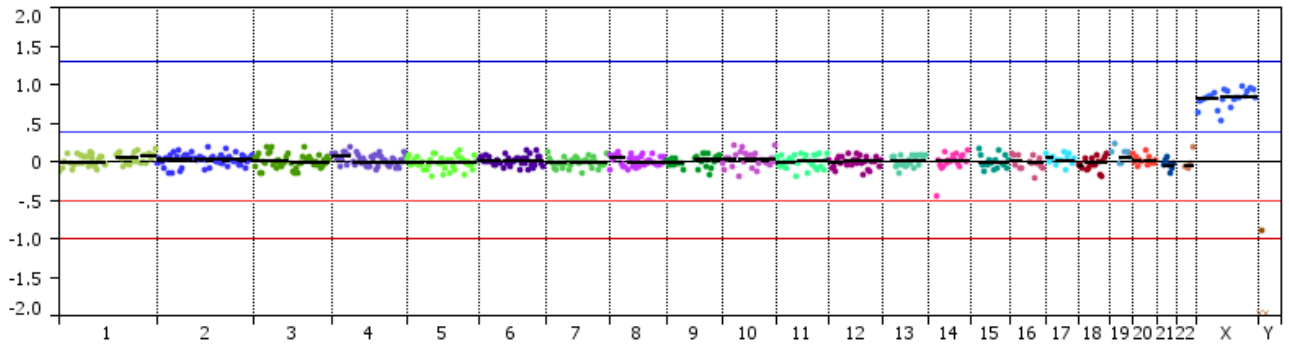

Figure 12: 46,XX single cell

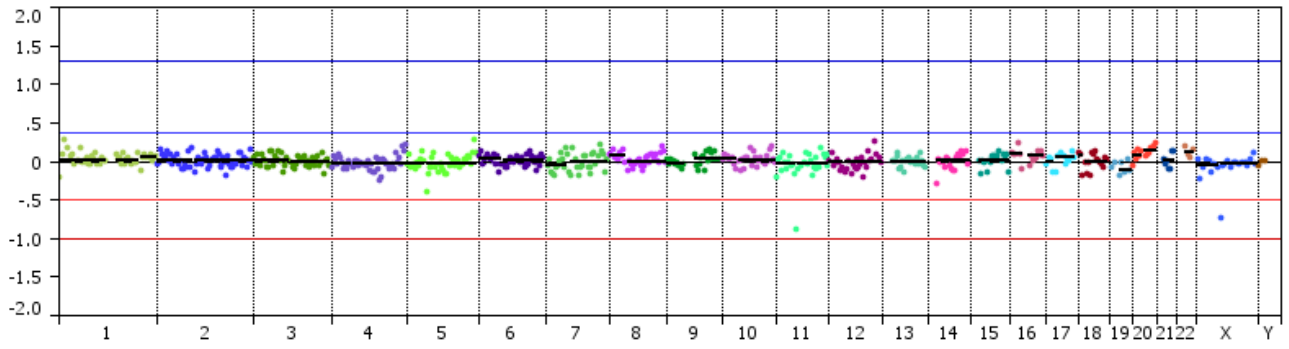

Figure 13: 46,XY single cell

Structural rearrangement detection

In order to confirm the theoretical resolution of PG-Seq™, 5 cell samples from cell lines known to harbor segmental errors were sourced from Coriell Biorepository and subjected to the PG-Seq™ workflow. These cell lines contained aberrations ranging in size from 7-31mb (GM09552, GM14485 and GM07312). Adjusted software parameters were used to increase the detection resolution when analysing these samples in order to increase sensitivity and accurately determine the breakpoints of the segmental aberrations.

A total of sixty-three 5 cell samples were subjected to the PG-Seq™ protocol with 1 sample failing QC parameters. The results for each segmental cell line is detailed in Table 4 below.

Table 4: 5 Cell results for cell lines with structural rearrangements

Cell line ID – 5 cells	Karyotype	Number Passed QC	Correct	False Negative Mosaic	False Positive Segmental	False Negative Segmental
GM09552	47,XY,+der(21)t(3;21)(p24.1;q21)	22	20	0	1	1
GM14485	46,XY,der(8)del(8)(p23.3)dup(8)(:p23.1->p11.2::p23.1->qter)	21	19	1	1	0
GM07312	46,XX,del(13)(pter>q14.1::q21.2>qter)	19	18	0	1	0
	Total	**62**	**57 (91.93%)**	**1 (1.61%)**	**3 (4.84%)**	**1 (1.61%)**

Segmental aberrations of 7-31Mb were detected with 98.3% sensitivity and specificity in 5 cell samples. Sixty-one of 62 samples showed the expected rearrangement. One result that appeared to be high level mosaic was included as correct. Examples of the results can be seen in figures 14 to 16.

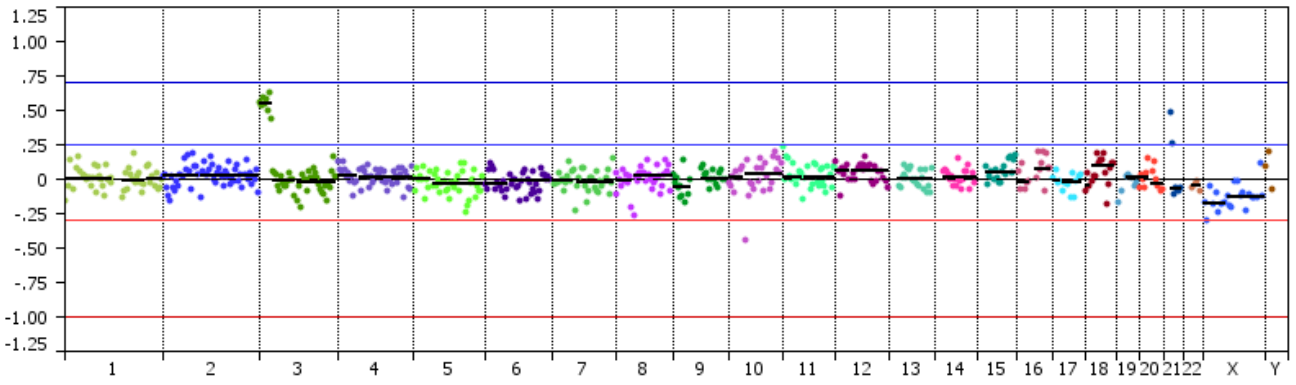

Figure 14a: GM09552 ,47,XY,+der(21)t(3;21)(p24.1;q21)

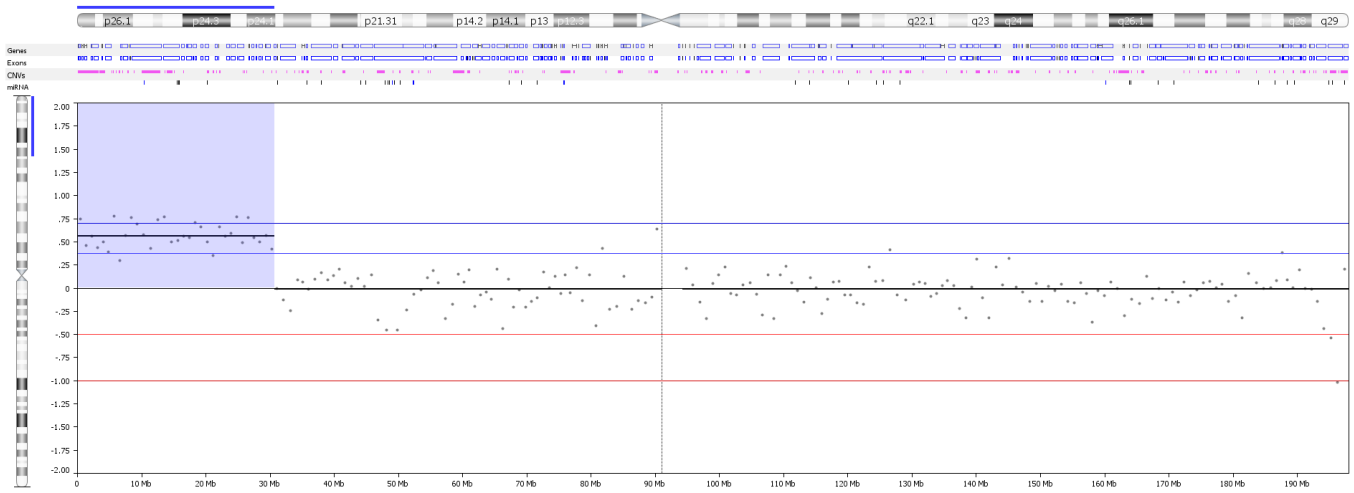

Figure 14b: GM09552 ,47,XY,+der(21)t(3;21)(p24.1;q21) closer view of the gain on chromosome 3

Figure 14c: GM09552 ,47,XY,+der(21)t(3;21)(p24.1;q21) closer view of the gain on chromosome 21

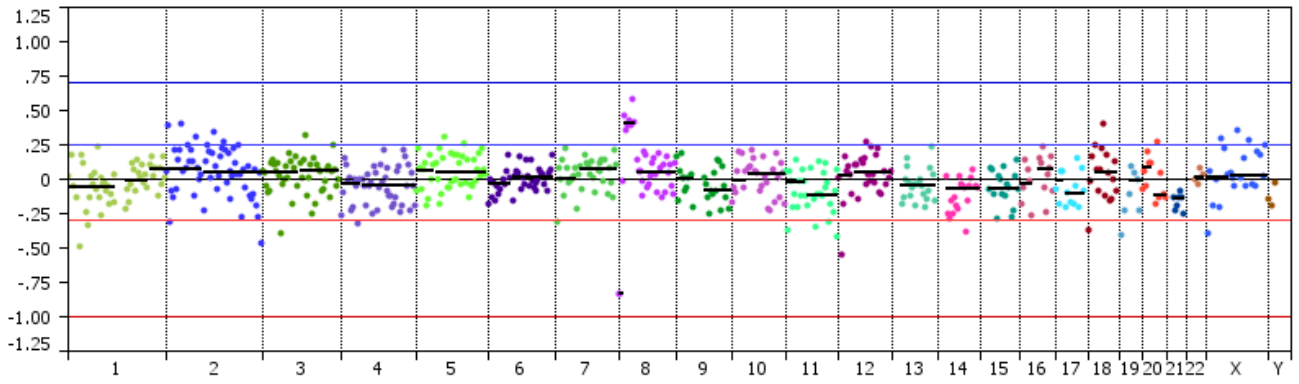

Figure 15a: GM14485 , 46,XY,der(8)del(8)(p23.3)dup(8)(:p23.1->p11.2::p23.1->qter)

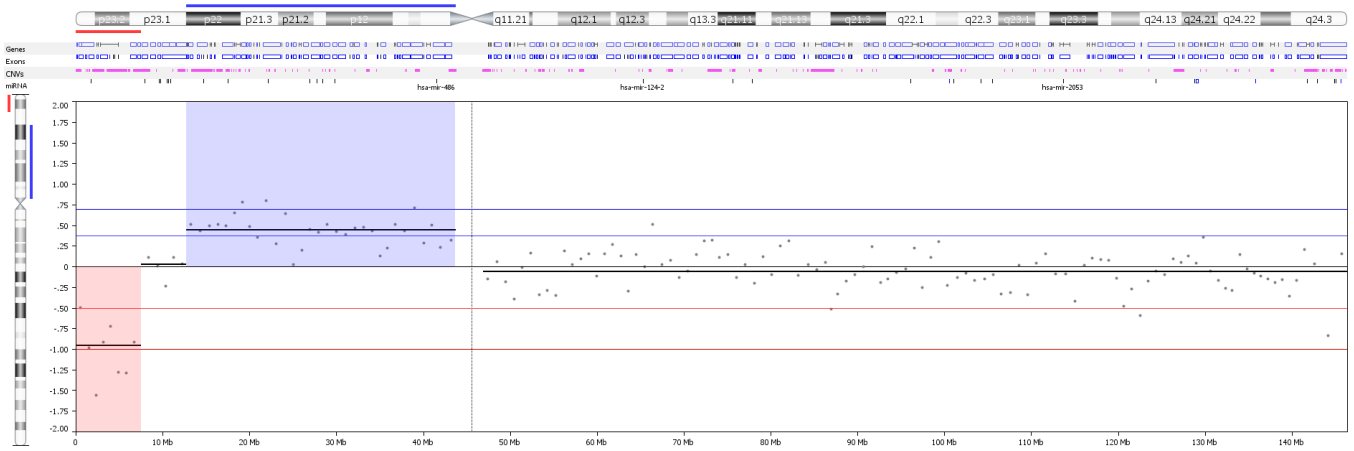

Figure 15b: GM14485 , 46,XY,der(8)del(8)(p23.3)dup(8)(:p23.1->p11.2::p23.1->qter); closer view of the loss and gain on chromosome 8

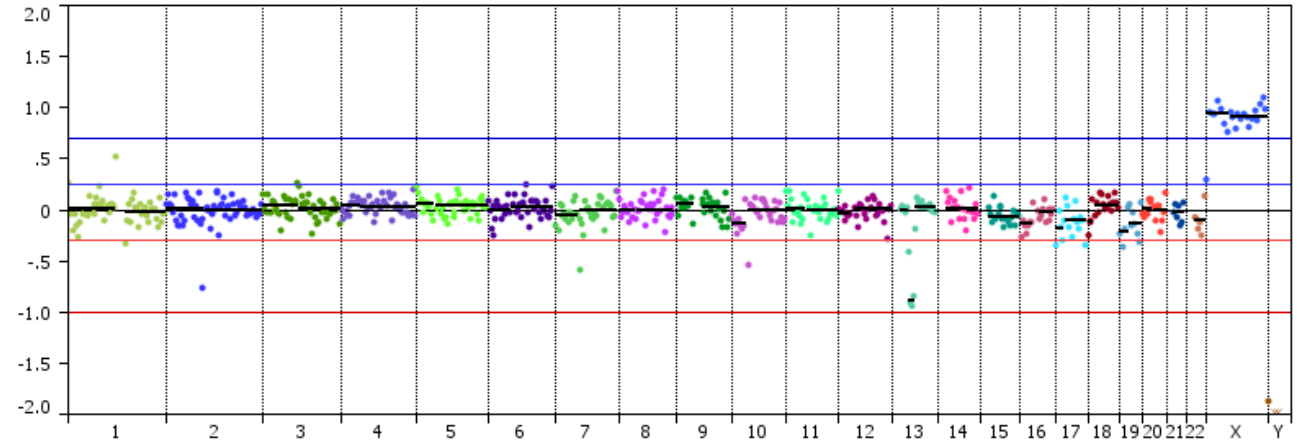

Figure 16a: GM70312 , 46,XX,del(13)(pter>q14.1::q21.2>qter)

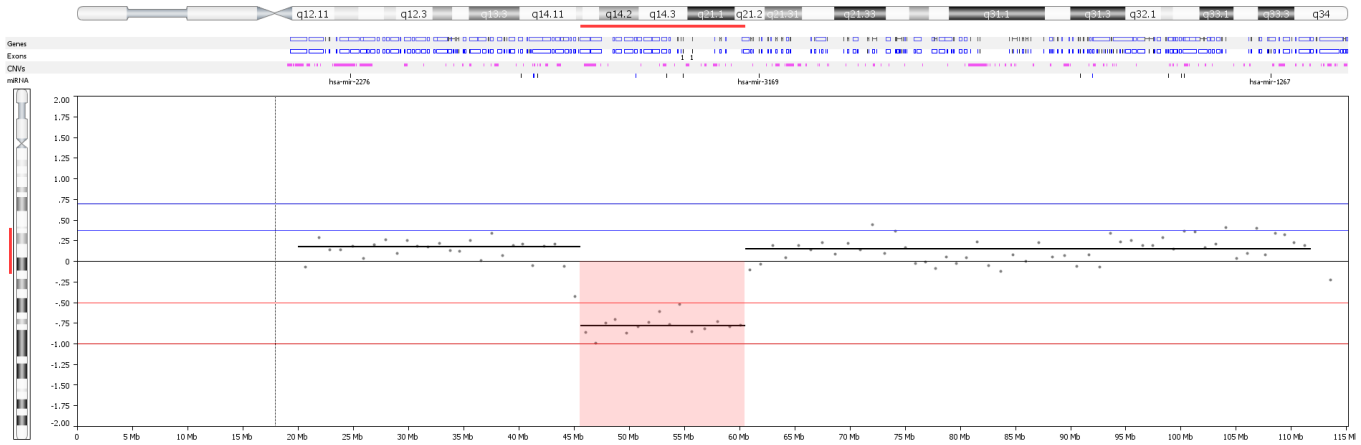

Figure 16b: GM70312 , 46,XX,del(13)(pter>q14.1::q21.2>qter), closer view of the loss on chromosome 13

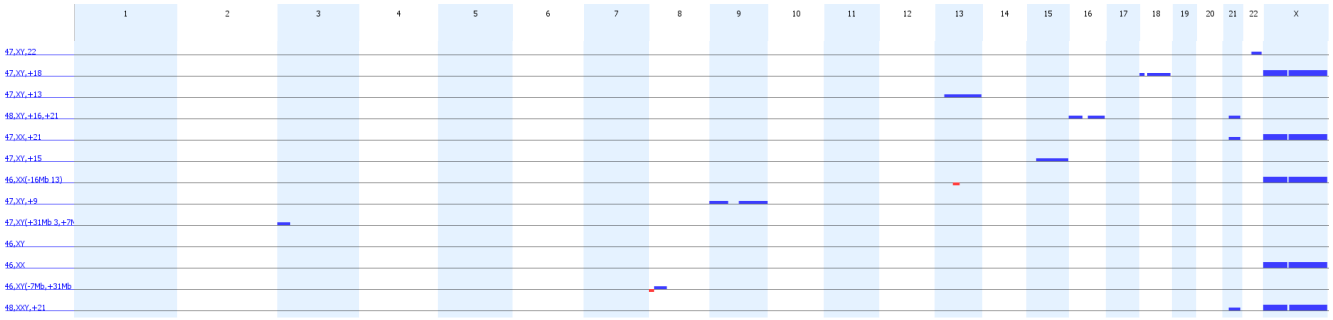

Figure 17: Summary of all aneuploidy and structural rearrangement aberrations validated using PG-Seq™

Inherent mosaicism in the cell lines used

A limitation of performing comprehensive chromosome screening on small numbers of cells is the potential mosaicism associated with a cell line being grown in culture. Limited data is available that documents mosaicism in commercially available cell lines (Table 5).

Table 5: Reported karyotype of the cell lines used in validating PG-Seq™ and the year of assessment (Source; Coriell Cell Repository pers. comm. and RHS data).

Cell line	Karyotype	Karyotype by Coriell (No. of cells; Year)	Karyotype by RHS (No. of cells; Year)
GM07189	47,XY,+15	100% 47,XY,+15; 10; 1984	100% 47,XY,+15; 39; 2015
GM00143	47,XX,+18		100% 47,XX,+18; 50; 2015
GM04435	48,XY,+16,+21	10% 47,XY,+21, 90% 48,XY,+16,+21; 10; 1981	
GM02948	47,XY,+13		88% 47,XY,+13, 12% 46,XY; 42; 2009
GM04965	48,XXY,+21	100% 48,XXY,+21; 10; 1981	100% 48,XXY,+21; 51; 2015
GM04592	47,XX,+21	100% 47,XX,+21; 10; 1981	

Based on karyotyping performed by Coriell and RHS, it was found that there is mosaicism present in some cell lines which may have increased in prevalence during subsequent passage. Considering these findings, it would not be unexpected that other cell lines used in this study and having been through numerous passages, may similarly be mosaic by now. During karyotyping of 4 cell lines from 2009-2015, in addition to the expected trisomy for each cell line having been confirmed, a partial chromosome loss was reported for 1-2 cells from three of the cell lines. This observation suggests that as well as loss of trisomy, there may be the introduction of sub-chromosomal errors in cell lines during sequential culture.

Additional considerations

Failed WGA is caused by failure to transfer the biopsy into the PCR tube or placement of the cell above the meniscus of the lysis and WGA reagents. RHS recommends visually inspecting the transfer pipette to ensure that the cells have been transferred. RHS also recommends marking the location of the transferred biopsy on the outside of the PCR tube while the transfer buffer is still visible, particularly if the transferred cell is not at the very bottom of the tube. If the latter occurs, it may be necessary to add the lysis buffer above the marked cell position to ensure that the biopsy is lysed.

Clinical data on the implantation rates of mosaic embryo biopsies suggests that the transferred mosaic embryo has an equivalent chance of implantation success and live birth as an untested embryo. Clinical decisions on mosaic PGS results will depend on the individual circumstances (ie number of available embryos, affected chromosome, patient consent, confidence in the PGS result and likelihood of technical causes). Whilst RHS has validated the detection of mosaicism down to 20% using cell lines, it is much more difficult in our experience to call low level mosaicism in an unknown clinical sample. It has been our recommendation to clinics to call results suggesting 80% or greater aneuploidy as aneuploidy, results suggesting 50-80% as mosaic and results below 50% aneuploidy as euploid. PGS should always be followed up with prenatal screening as it is widely accepted that embryo mosaicism can occur and the ploidy of the inner cell mass is not tested during PGS.

Conclusion

The choice of Preimplantation Genetic Screening kit should be determined to suit several factors that impact on the patient, laboratory and clinician. The kit of choice should be tailored to address its clinical intent and the degree of resolution required for the indication. The PG-Seq™ kit for analysis of chromosomes in a multi-cell or single cell low-DNA template sample provides the sensitivity, flexibility in its sample format and accuracy for ready introduction into a routine IVF laboratory. The data herein generated using cell lines of known karyotype and lymphocytes manually sorted illustrate the accuracy of PG-Seq™ and confirm the test performance for small numbers of cells is equivalent to other PGS solutions.

PG-Seq™ has been specifically designed and validated for analysis of whole chromosomal aneuploidy and structural rearrangements down to 7Mb in size with correct results obtained in over 95% of single cells and 5 cell aliquots tested. This performance has been achieved multiplexing 48 samples using a standard Illumina MiSeq chemistry. RHS has also developed a proprietary approach to combining PGS with Pre-implantation Genetic Diagnosis using target sequence enrichment. Combined PGS and PGD results are achievable using the same workflow as has been validated herein other than a specific modification to the whole genome amplification protocol. PG-Seq™ offers a unique solution to meet the varied testing requirements of IVF embryos using a single high throughput, adaptable workflow.

References

1. Deleye L, et al. Performance of four modern whole genome amplification methods for copy number variant detection in single cells. Sci Rep, 2017, 7:3422 doi:10.1038/s41598-017-03711-y
2. Warren K, et al Validation of high throughput low cost NGS PGS assay: impact of library preparation and read length on resolution. ESHRE 2017, Geneva
3. Fragouli E, et al. Altered levels of mitochondrial DNA are associated with female age, aneuploidy, and provide an independent measure of embryonic implantation potential. PLoS Genet. 2015, 11(6):e1005241
4. Hussey ND, Donggui H, Froiland DA, et al. Analysis of five Duchenne muscular dystrophy exons and gender determination using conventional duplex polymerase chain reaction on single cells. Mol Hum Reprod 1999, 5:1089-94

PG-Seq™ is for Research Use Only and is not to be used for clinical diagnosis.

RHS Ltd

40-46 West Thebarton Road

Thebarton South Australia 5031

AUSTRALIA

ASX:RHS

info@rhsc.com.au | www.rhsc.com.au