RHS
Unique Genomic Solutions

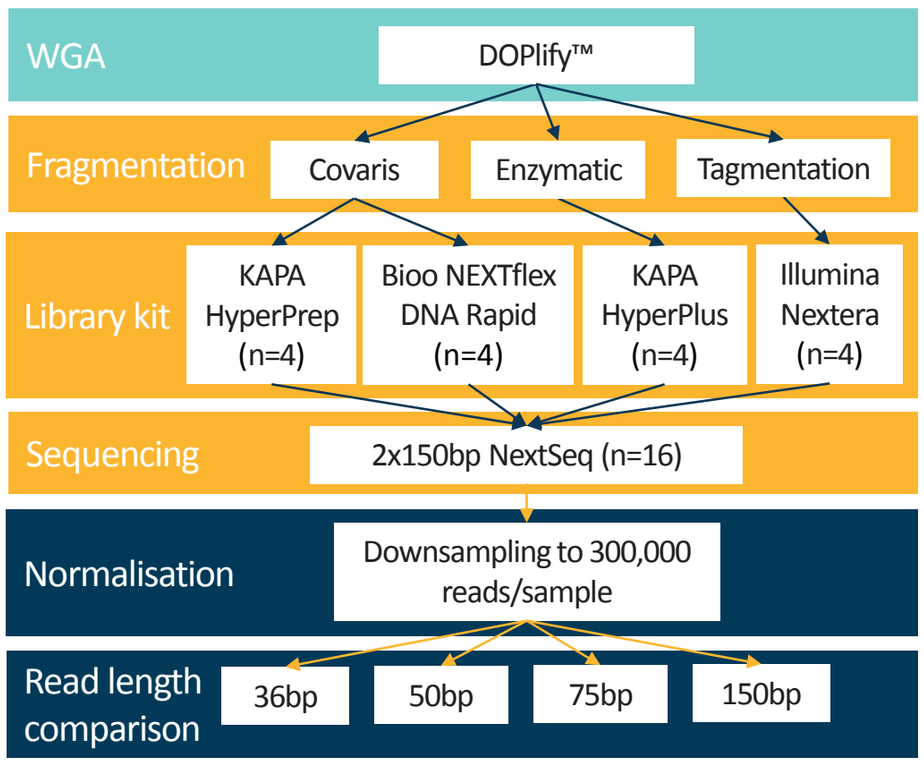
rhsc.com.au

Validation of a high throughput, low cost NGS PGS assay: impact of library preparation and read length on resolution

Limited Whole Genome Amplification (WGA) technologies suitable for copy number detection are commercially available. RHS DOPlify™ has been specifically developed for Next Generation Sequencing (NGS). The time and financial efficiencies offered by NGS of clinical samples is mediated by multiplexing, with the cost per sample decreasing as more samples are multiplexed. However, this can also negatively impact the test resolution and quality of results. When limited source material for genome wide evaluation is available; effective fragmentation, highly efficient adapter ligation and optimal read length is imperative to maximise the reads per sample available for interrogation.

Aim – This study evaluated the impact of library preparation kits and read length on PGS results using whole genome amplified single cell and 5-cell inputs.

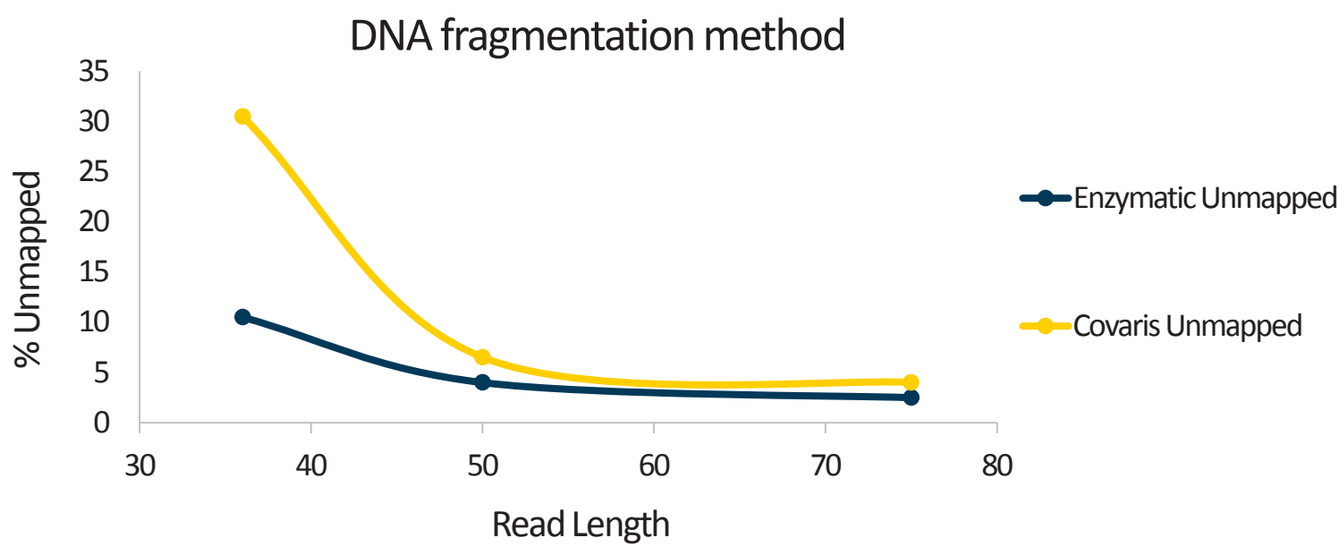
Methods - Individually sorted single cell and 5-cell aliquots from four aneuploid cell lines (Coriell Institute) underwent WGA using DOPlify™. PCR products from single cells and 5 cells from each cell line were pooled to make a consistent template. This template was then split over the different kits tested. Libraries were pooled and sequenced in a single NGS run. Bioinformatics was used to normalise the number of reads per sample and to trim the reads for each of the read lengths being compared.



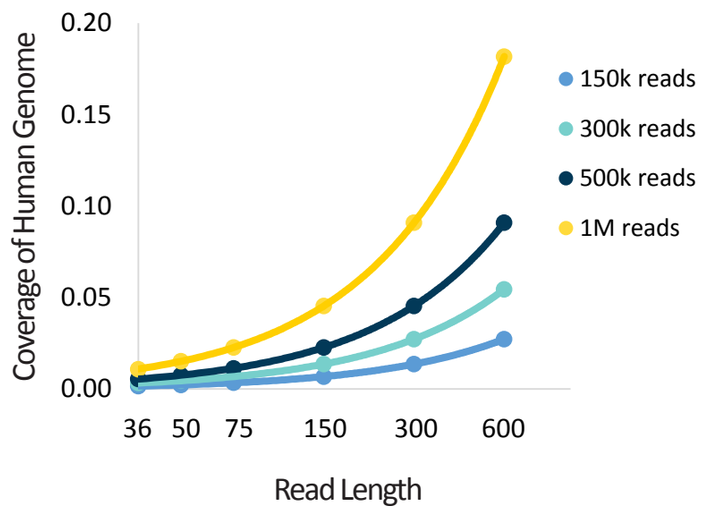
K. Warren[1], P. Wang[2], W. Parker[2], B. Hodgson[1], M. Jasper[1].
[1]RHS Ltd, Thebarton, Australia. [2]ACRF Cancer Genomics Facility, Centre for Cancer Biology, Adelaide, Australia.

RHS Ltd, 40-46 West Thebarton Road, Thebarton
South Australia 5031, Australia | **rhsc.com.au**

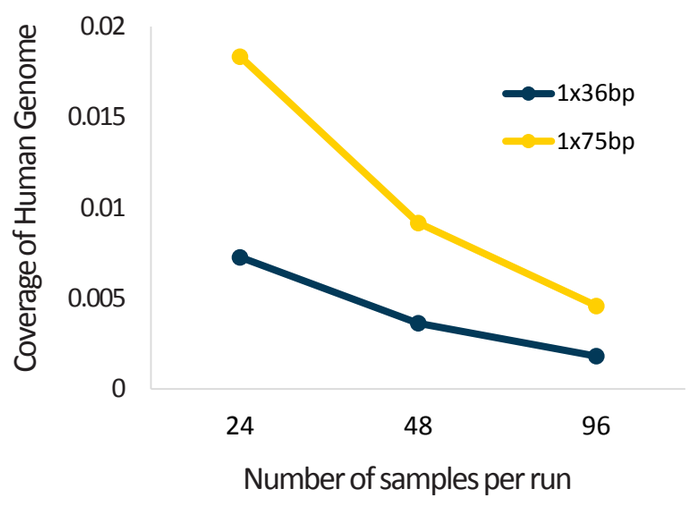
Results - Enzymatic fragmentation is the most effective and time efficient method for NGS library preparation, providing higher mapping rates and less data loss. This is particularly evident at the shorter read length of 1x36bp.



Genome coverage is greater using longer read lengths, allowing increased sample multiplexing. Modelling demonstrates that by increasing the read length from 36bp to 75bp, greater genomic coverage is achieved even when the number of samples in the run is doubled from 24 to 48.



Genome coverage is dependent on read length and total reads sequenced per sample.



Genome coverage is equivalent even when larger numbers of samples are multiplexed, if longer reads lengths are sequenced.

The resolution of PG-Seq is determined by the number of samples multiplexed. By multiplexing 48 samples with an average 600k reads per sample, it is possible to detect 2Mbp CNVs.

Samples	Average reads per sample	Resolution (5x minimum window size)
96	300k	5Mb
48	600k	2Mb
24	1.5M	1Mb

Conclusions

- The quantity and quality of sequencing data is highly dependent on multiplexing, library preparation method and read length.
- RHS PG-Seq with DOPlify™ has been purposefully designed to provide optimal read length to multiplex 48 samples per NGS run with high resolution.

DOPlify™ A New Generation of Whole Genome Amplification