PG-Seq™ A novel complete NGS solution

# Tailoring resolution for PGT-A and PGT-SR by modifying NGS sample throughput

Unlike traditional DNA analysis platforms used for Preimplantation Genetic Testing (PGT), Next Generation Sequencing (NGS) technologies provide a highly flexible and customizable workflow for the detection of segmental copy number variations (CNVs) offering variable resolution. However, greater resolution is achieved at the expense of sample throughput.

**Aim** – We aim to define the level of resolution achievable using the PG-Seq™ kit with different NGS sample throughput.

**Materials** – PG-Seq™ kit NGS files reflecting the expected total mapped read count per sample for different throughput scenarios from 1-96 samples in a single MiSeq® instrument (Illumina®) run were created in silico from PG-Seq™ kit NGS files using random sampling. Files with total mapped read counts of 250,000 (96 samples), 500,000 (achieved using the standard PG-Seq™ kit workflow for 48 samples), 1 million (24 samples), 3 million (8 samples), 12 million (2 samples) and 24 million (1 sample) were created. Artificial CNVs, including duplications and deletions ranging from 100 Kbp to 10 Mbp were added in silico across the q arm of chromosomes 1, 5, 11, 15, 17, 19, and 22 (eg Figure 1). PG-Seq™ software was used to analyse CNVs with uniform settings across all samples.

**Results** – As expected, the CNV resolution increased with increasing read count per sample (Figure 2).

High resolution CNV calling, for the detection of germline or de novo 100 Kbp duplications or deletions, required > 12 million reads, or no more than 2 samples per MiSeq® instrument run.

A standard 48 sample PG-Seq™ kit run detected 5 Mbp segmental CNVs at a frequency of > 99.5% and should theoretically be suitable for detecting de novo CNVs of this size.

A higher throughput of 96 samples per run detected 10 Mbp CNVs with a frequency of 100% (Figure 3). While germline segmental CNVs could be called at this read count, there is a risk of false positive detection of de novo segmental CNVs.

The location of CNVs and their proximity to other CNVs, centromeres, and telomeres affects software CNV detection. CNVs distal to the centromere were detected at a higher frequency than those located proximal to the centromere. Additionally, at lower read counts per sample, two or more nearby duplications or deletions were sometimes detected as a single CNV.

S Myers, KA Warren, MJ Jasper

PerkinElmer Health Sciences (Australia) Pty Ltd, 40-46 West Thebarton Road, Thebarton, 5031, Australia

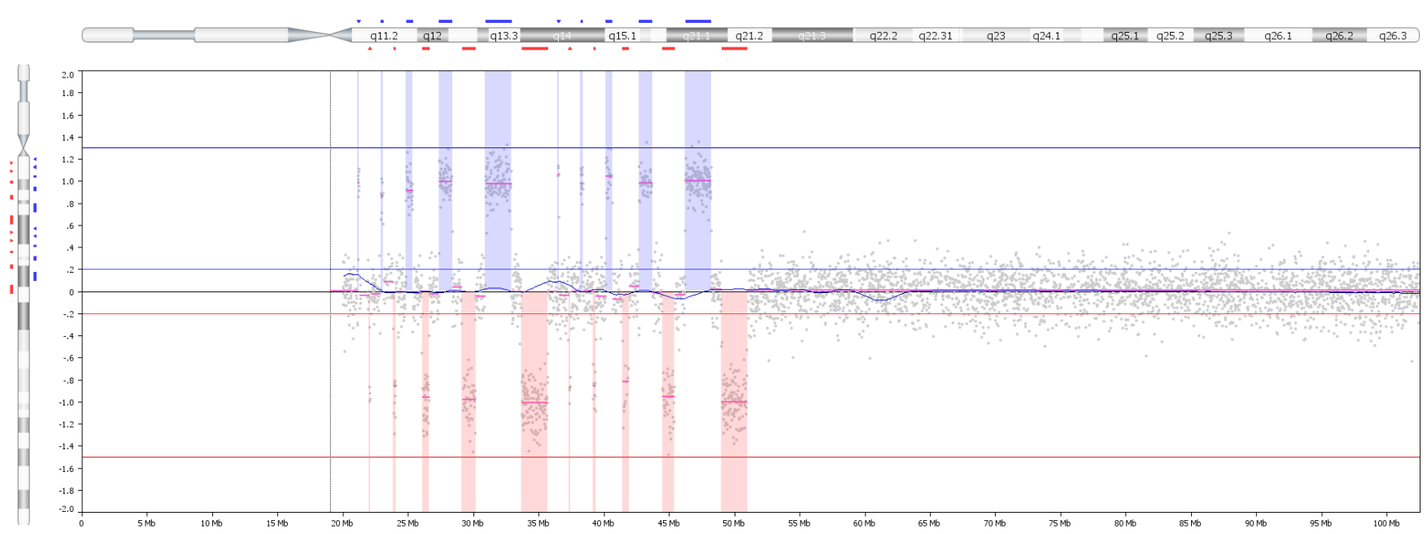

**Figure 1. –** Chromosome 15 CNV overview showing five duplications and five deletions ranging from 100 Kbp, 200 Kbp, 500 Kbp, 1Mbp to 2 Mbp replicated in duplicate for a sample with 24,000,000 reads, equivalent to a single sample throughput NGS run.

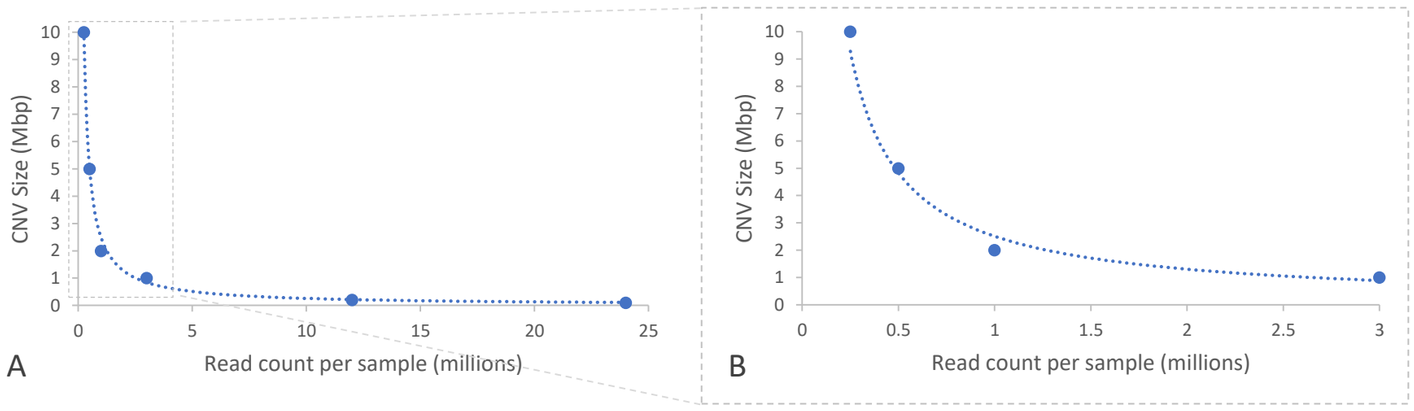


**Figure 2. –** Resolution attained in silico using the PG-Seq™ kit for different NGS sample throughput scenarios (n ≥180 replicates per CNV size) A) 96 samples down to 1 sample (0.25-24 million reads) and B) highlighted section of A replicated, 96 down to 8 samples (0.25-3 million reads) per NGS run.

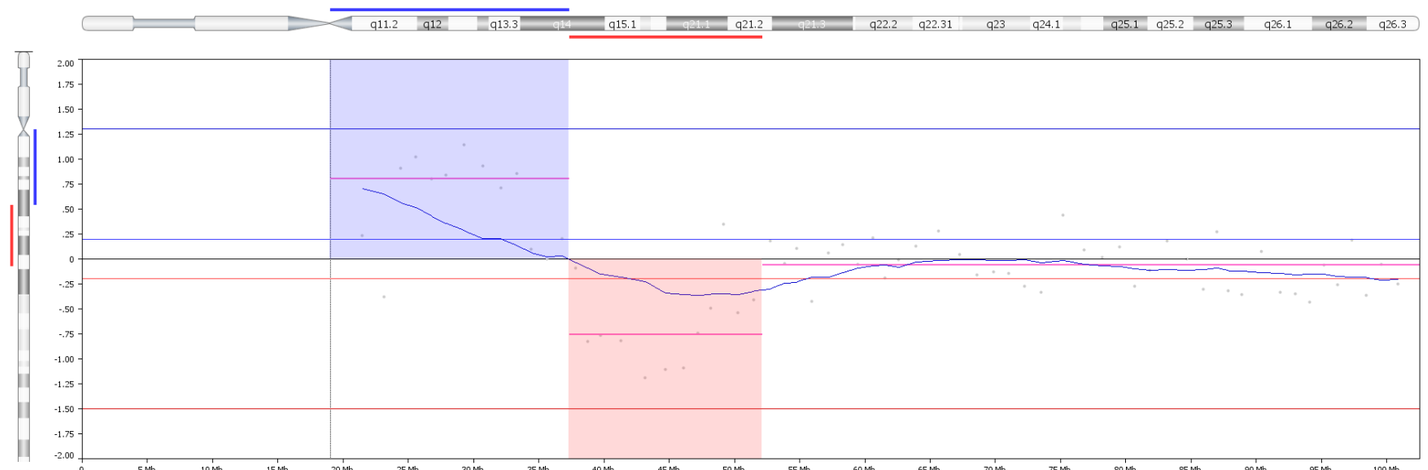

**Figure 3. –** Detection of a single 10 Mbp duplication and 10 Mbp deletion on Chromosome 15 for a sample with 250,000 reads, equivalent to a 96 sample throughput NGS run.

**Conclusions**

*In silico* analysis of Next Generation Sequencing data highlights:

- technical limitations and financial implications for the detection of small segmental CNVs,
- the PG-Seq™ kit offers scalable and economical sequencing for PGT-A and PGT-SR, and
- a standard 48 sample PG-Seq™ kit NGS run can detect *de novo* CNVs as small as 5 Mbp.

These *in silico* observations should be validated using clinical samples with known CNVs, particularly if considering the detection of small *de novo* aberrations.