Oxford NANOPORE Technologies™

APPLICATIONS

Comprehensive analysis of hereditary cancers with PCR-free targeted sequencing using adaptive sampling

Simple targeted sequencing of the human genome on a MinION™ Flow Cell allows SNP-calling and phasing, methylation analysis and genome-wide, low-pass copy-number-variant analysis from a single dataset

Contact: **apps_posters@nanoporetech.com** More information at: **www.nanoporetech.com** and **publications.nanoporetech.com**

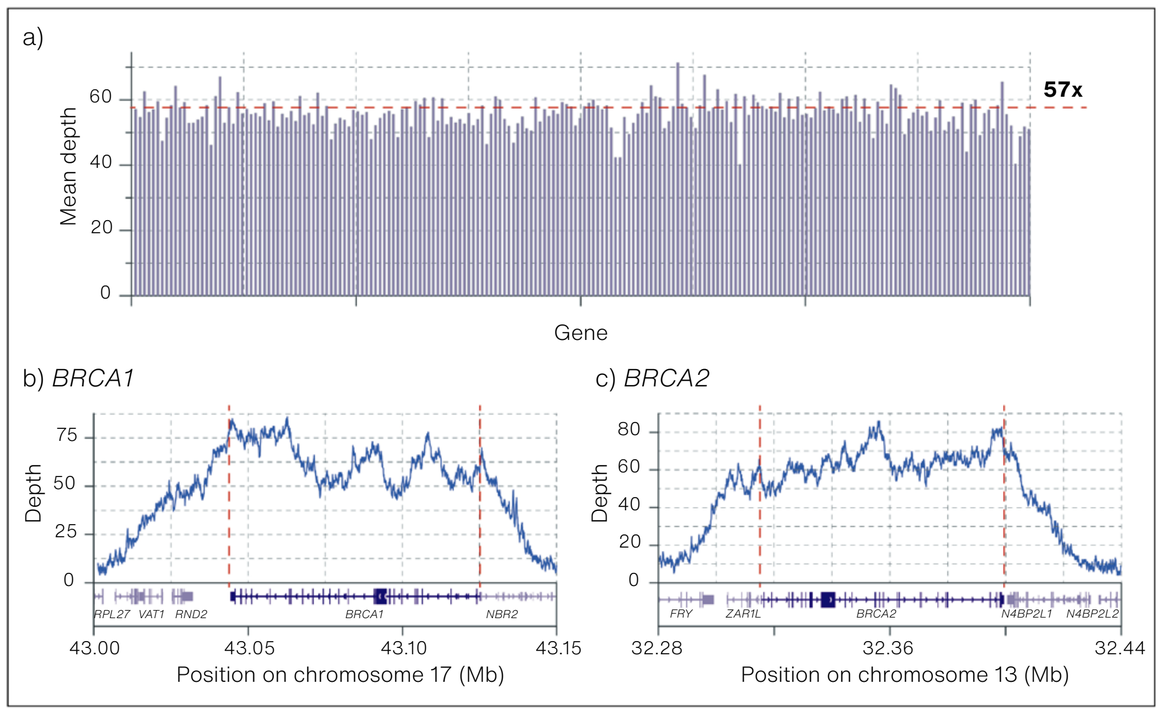

Fig. 1 Adaptive sampling coverage across a) >200 cancer-related genes b) *BRCA1* c) *BRCA2*

Adaptive sampling allows genomic regions to be targeted during sequencing run

In adaptive sampling, strands of DNA are base-called as they pass through the pore, and if the first 400-500 nucleotides of the read aligns within a predefined target region, that read is allowed to progress. Otherwise, the electrical potential across the individual pore is reversed, which rejects the strand and opens the pore to further template molecules. Here we designed a panel to target 201 hereditary cancer-related genes and > 15,000 promoter regions across the genome. Mean coverage of the cancer genes was 57x after 72 hours of sequencing (Fig. 1a). The end-to-end coverage profile was reasonably uniform along the length of particularly long target genes such as *BRCA1* and *BRCA2* (Figs. 1b and 1c).

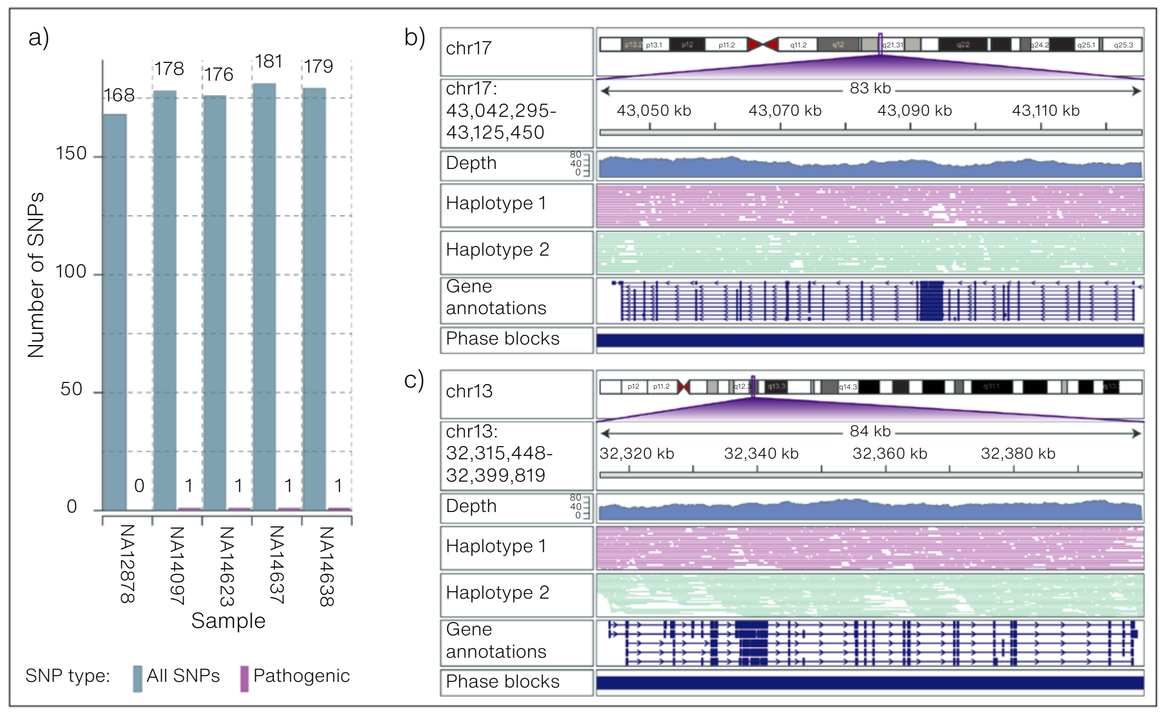

Fig. 2 SNP-calling a) in breast- and ovarian-cancer samples b) phasing in *BRCA1* c) *BRCA2*

SNP-calling on enriched targeted regions to find pathogenic variants in cancer samples

We called SNPs using *clair3* and intersected the calls with ClinVar (Fig. 2a). We identified a T > G SNP in NA14097 *BRCA1* exon 5 leading to a missense mutation which results in a non-functional protein, a nonsense mutation in NA14637 *BRCA1* exon 13, and a cryptic splice site in *BRCA1* intron 5 in NA14638. Although NA14623 is a BRCA negative breast cancer sample, we found a pathogenic variant in PALB2, a critical protein in the formation of the *BRCA1- PALB2-BRAC2* repair complex. Phasing these complex genes separates them fully into individual haplotypes (Figs. 2b and 2c) which is important in the study of compound heterozygosity, and also allows the target genes to be distinguished from their pseudogenes.

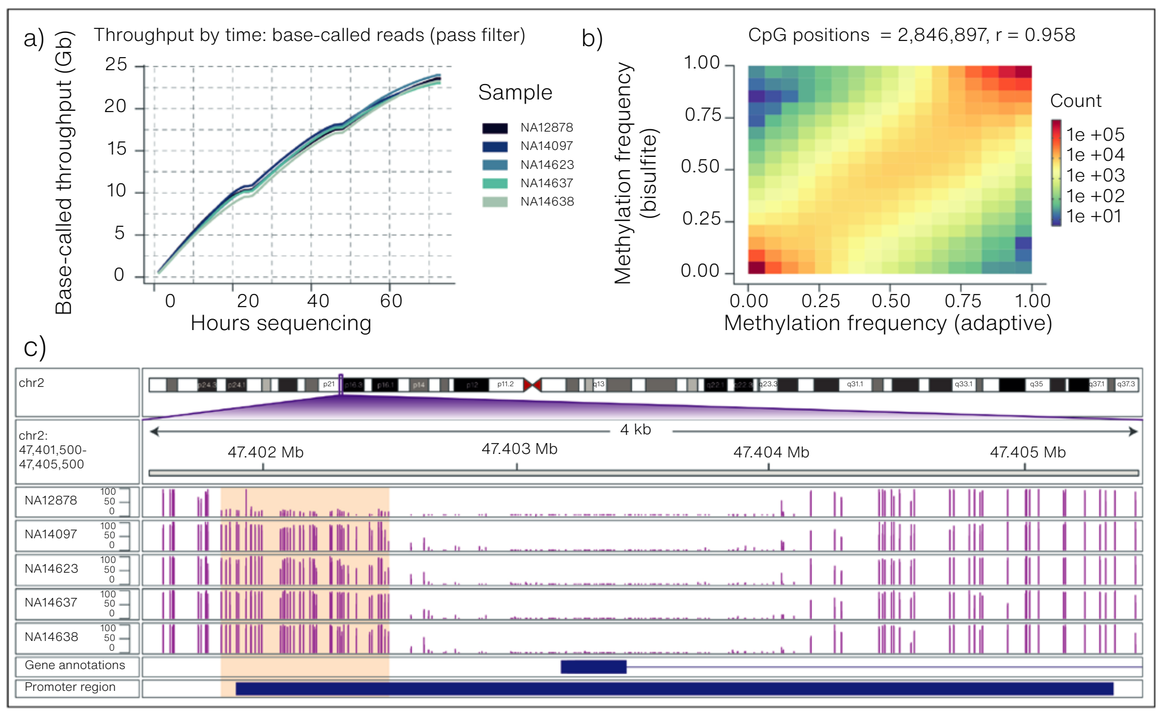

Fig. 3 Adaptive methylation analysis a) throughput b) correlation with bisulfite c) *MSH2* gene

Targeting with adaptive sampling is PCR-free, so methylation signals are retained

We also included over 15,000 promoter regions that contained > 4 CpGs in the adaptive sampling bed file, allowing for a form of reduced representation methylation calling to be performed. The sequencing depth across the majority of the 15,000 promoter regions was sufficient to provide robust, per CpG methylation calls (Fig. 3a). Comparing the outputs of methylation calling from the adaptive sampling with a bisulfite data set for the same sample gave a correlation coefficient of 0.958 over 2.8 million CpG sites (Fig. 3b). Many sites across the five genomes showed contrasting methylation patterns. However a stark contrast between hypermethylation in all cancer samples and the NA12878 non-cancer control was observed in the promoter/enhancer region of the known mismatch-repair gene *MSH2* (Fig. 3c).

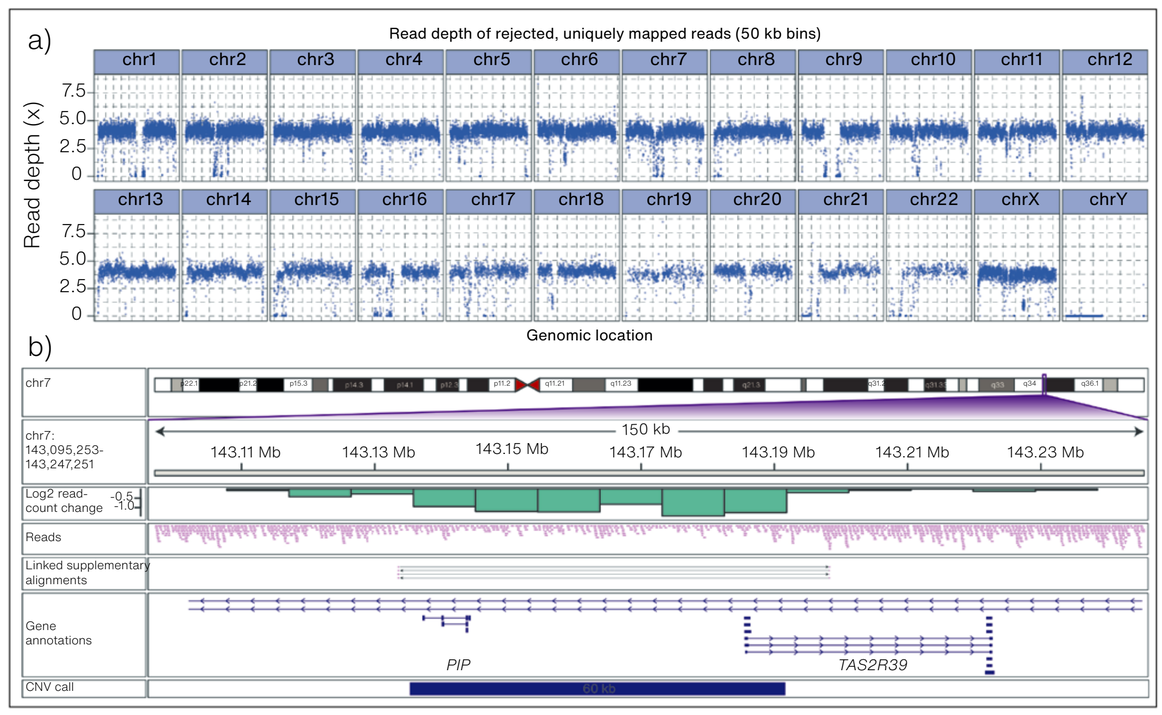

Fig. 4 Genome-wide CNV analysis a) read depth b) deletion in the *PIP* gene

Using rejected, off-target reads for the analysis of copy-number variants

As the first 400–500 bp of each read is base called prior to being either accepted or rejected, the millions of short, rejected reads can be used to generate low-pass whole-genome coverage of approximately 4– 5x depth (Fig. 4a) and exploited to detect CNV changes in a non-targeted manner, using *QDNAseq*. In this way, a ~60 kb deletion encompassing the PIP gene can be seen in the BRCA-negative breast cancer sample (Fig. 4b). When supplementary alignments of reads spanning this region are linked, the CNV breakpoint can be pinpointed. PIP gene is thought to be associated with anti-tumour immune response mechanisms and higher expression levels have been linked to better prognosis in triple negative breast cancer. However, the impact of germline CNVs involving this gene is poorly understood.