



A rapid CRISPR/Cas9-mediated, amplification-free target enrichment method for native-strand sequencing*

Sequence selection by Cas9 provides a low-cost method of enriching native DNA from user-specified genomic loci, including regions which cannot be amplified by PCR

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

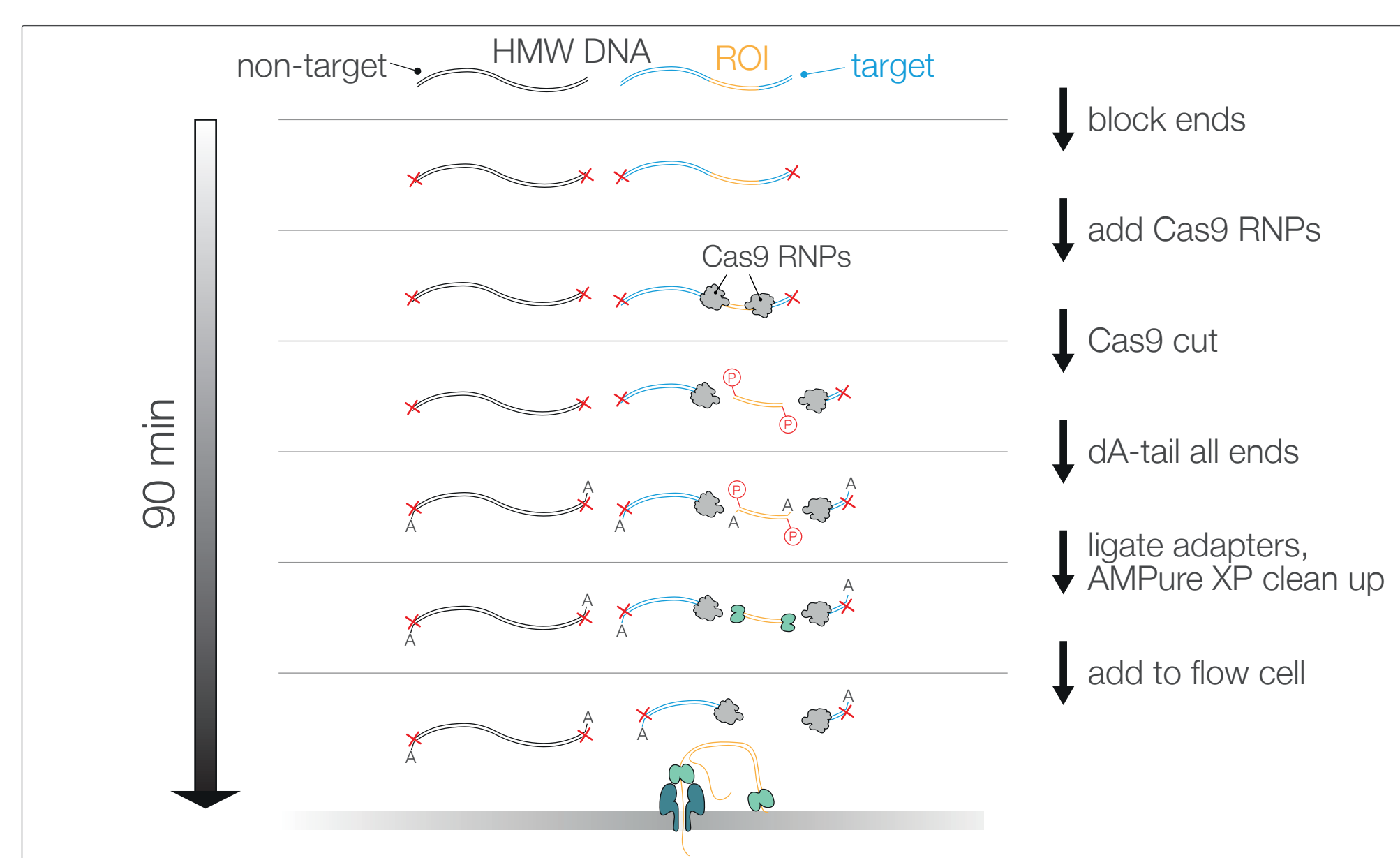


Fig. 1 Overview of library preparation for CRISPR/Cas9 target enrichment

90-minute Cas9 library preparation for PCR-free enrichment of target loci

It is an advantage in many situations to enrich for regions of interest (ROI) prior to sequencing. Here, we introduce a PCR-free enrichment method for nanopore sequencing, using Cas9 (Fig. 1). Because native strands are sequenced, fragment length and modifications are preserved. In the method, sample DNA is dephosphorylated, to prevent ligation. Cas9 is then used to cleave the DNA at predetermined sites, exposing ligatable ends. All 3' ends are dA-tailed and sequencing adapters are ligated only to the cleaved ends and the entire library is then added to the flow cell. In this way the fraction of reads corresponding to the ROI is enriched several thousand-fold, enabling many samples to be run on the same flow cell, or a lower-cost flow cell to be used.

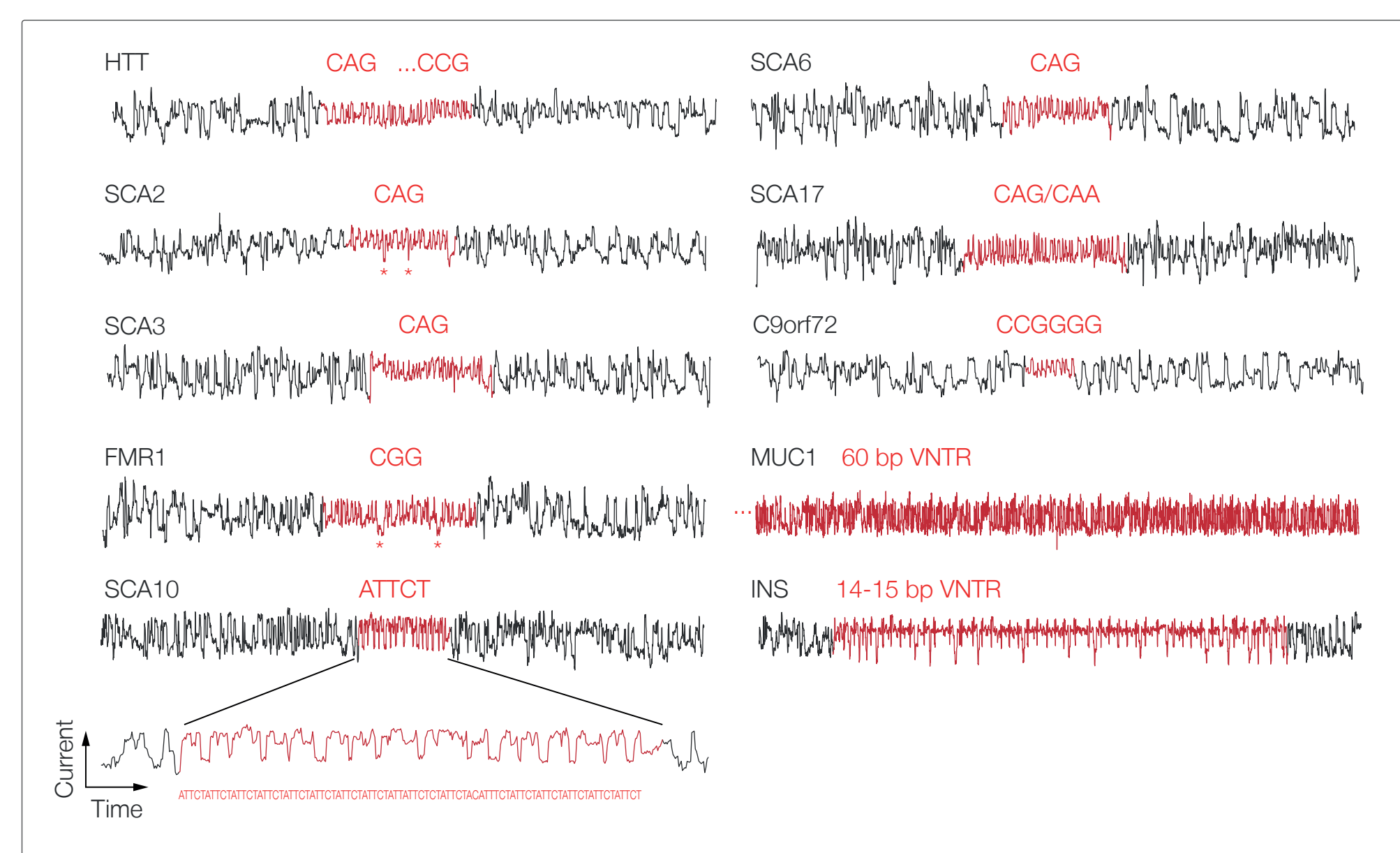


Fig. 3 Portions of raw data ("squiggles") from ROI in the panel; repetitive elements in red

Accurate counting of repeat elements in regions of interest not amenable to PCR

Ascertaining the repeat number in structured and variable number tandem repeats is important diagnostically. The present clinical approach is to perform repeat-primed PCR and Southern blot, and to count the repeats manually. However, such regions tend to be refractory to accurate PCR because their sequences are highly repetitive and GC-rich, which leads to imprecise estimates. Here a PCR-free enrichment and sequencing method is invaluable as it allows these difficult loci to be investigated without the need to sequence an entire genome. Fig. 3 shows raw MinION data ('squiggles') from a wide variety of STRs and VNTRs. Repeats can be counted accurately, either from basecalls or from raw data and clinically-relevant interruptions (asterisks), which may insulate against expansion of the repeats, can also be identified.

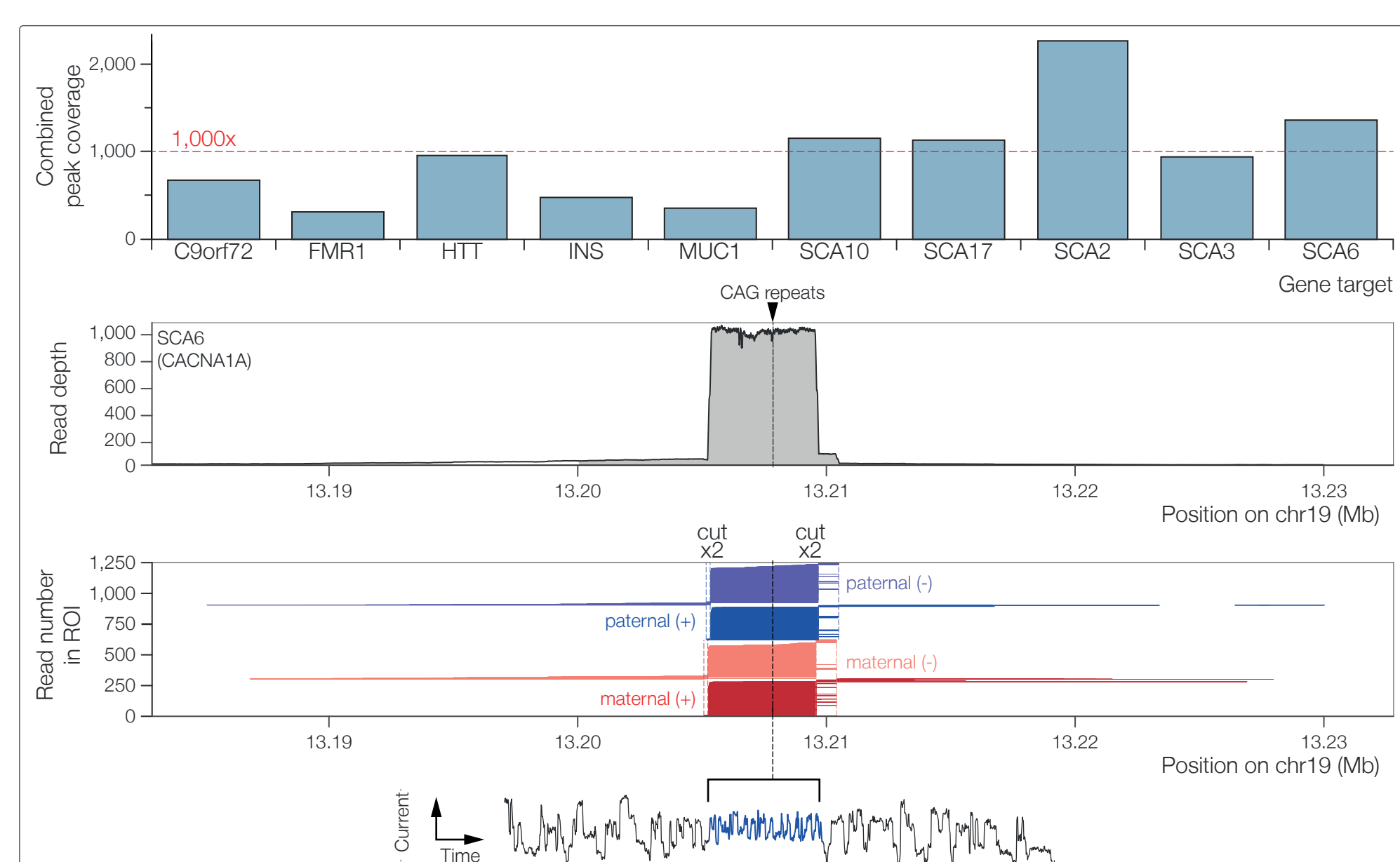


Fig. 2 PCR-free enrichment of ten human gene regions using a panel of 40 CRISPR RNAs

High coverage of ten human repeat expansion genes on one MinION flow cell

We designed probes to enrich ten human gene regions ranging from 2 kb - 20 kb, using a panel of 40 Cas9 CRISPR RNAs (crRNAs) in such a way that targeted regions would be captured as single fragments in both forward and reverse directions. An equimolar mix of the crRNAs was used with 5 µg high molecular weight NA12878 genomic DNA using the protocol outlined in Fig. 1 and the resulting library was run on a single MinION flow cell for 24 hours. The aligned coverage for each of the captured regions was 200-2,000x (Fig. 2a). Taking the CACNA1A gene as an example, >80% of the reads of the 5 kb ROI were single-pass, full-length, native molecules.

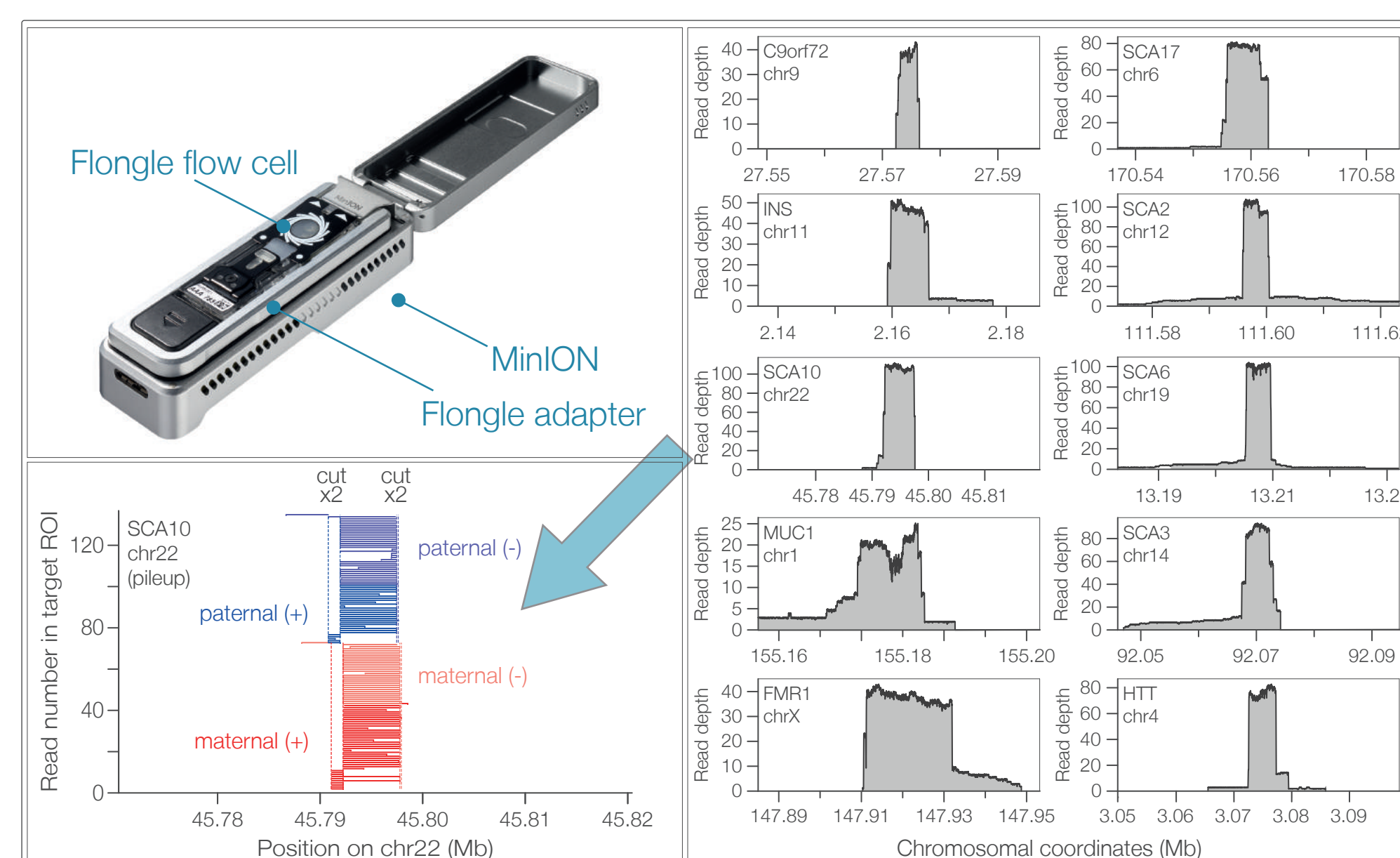


Fig. 4 Sequencing of the Cas9 targeted repeat panel using the Flongle adapter

Low-cost, rapid, amplification-free targeted sequencing on Flongle

When looking at a relatively small number of target loci, a single full-length MinION run may generate more data than required for the subsequent analysis. The Flongle is an adapter for the MinION and GridION that allows small and inexpensive single-use flow cells to be run. As with all of our sequencing devices, data is available for analysis within seconds of the run starting. The speed and simplicity of the Cas9-mediated enrichment workflow, when combined with the Flongle, creates a quick, accessible and cost-efficient targeted sequencing system for smaller tests and experiments. Fig. 4 shows data from the same 40-crRNA panel used in Fig. 2, but run on a single Flongle flow cell for 24 hours; coverage of up to 100x was obtained, equivalent to 50x coverage per allele.