No-Amp Targeted SMRT Sequencing Using a CRISPR-Cas9 Enrichment Method

Ian McLaughlin, Yu-Chih Tsai, Brett Bowman, Janet Ziegle
PacBio, 1305 O’Brien Drive, Menlo Park, CA 94025

Introduction

While PCR amplification techniques can be used successfully with genomic regions of moderate to high complexity, low-complexity regions and regions containing repetitive elements can lead to poor or failed PCR-based enrichment. Many genetic diseases caused by repeat element expansions have proven to be challenging enrichment targets with PCR amplification, used either alone or in combination with a hybridization capture method.

We have developed an enrichment method for targeted SMRT Sequencing on the PacBio Sequel System using the CRISPR-Cas9 system that requires no PCR amplification. This method, paired with SMRT Sequencing’s long reads, high consensus accuracy, and uniform coverage, allows sequencing of genomic regions regardless of challenging sequence context that cannot be investigated with other technologies. The method is amenable to analyzing multiple samples and/or targets in a single reaction. In addition, this method also preserves epigenetic modifications allowing for the detection and characterization of DNA methylation which has been shown to be a key factor in the disease mechanism for some repeat expansion diseases.

Here we present results of our latest No-Amp Targeted Sequencing procedure applied to the characterization of CAG triplet repeat expansions in the HTT gene responsible for Huntington’s Disease.

No-Amp Targeted SMRT Sequencing

![Workflow](image)

Fig. 1. Workflow for No-Amp Targeted SMRT Sequencing. If multiplexing samples, barcoded EcoRI target adapters are used to form a uniquely barcoded library for each sample. Barcoded libraries are pooled prior to the Complexity Reduction step where background gDNA not associated with SMRTbell templates containing the targeted region is digested with a mixture of restriction enzymes. Target multiplexing is facilitated by the use of multiple guide RNAs in the CRISPR-Cas9 Digestion step.

![Methodology](image)

Fig. 2. CRISPR-Cas9 enrichment methodology. The CRISPR-Cas9 system is used to target a region of interest in a background-reduced collection of SMRTbell templates followed by capture adapter ligation and magnetic bead enrichment of the resulting asymmetric SMRTbell library. Enriched SMRTbell templates are subsequently prepared for SMRT Sequencing on a Sequel System.

Results – Non-Expanded HTT Alleles

![Plot](image)

Fig. 4. Representative genomic coverage by ZMW plot for all 5 samples in Cas9 Digest Reaction Set 2. While HTT is the focus of the current protocol, other targets are occasionally tested such as FMR1 in the plot above. Other relatively small peaks such as in chromosomes 1, 7, 17, and Y may represent SMRTbell templates containing sequence homologous to the guide RNA sequence. These off-target molecules are a common feature of any CRISPR-Cas9 system and do not impact the ability to obtain on-target reads in our optimized protocol.

![Table](image)

Table 1. HTT results for a multiplexed sample set testing two different Cas9 digestion input amounts. HTT target counts are consistent within Cas9 Digest sets and scale proportionally to Cas9 input amount. The %On-Target values are also consistent between Cas9 Digest sets revealing that with only 368 ng input, approximately 1,000 target reads per sample can be achieved. This suggests that a greater multiplexing level is possible with the protocol as currently only 50 reads per allele are required to provide an accurate repeat call estimate.

![Histogram](image)

Fig. 3. Representative repeat region size histogram for Mb4-5 of Cas9 Digest Reaction Set 2. There are 17 contiguous CAG repeats for the vast majority of CCS reads. Other repeat contexts are included in the plot to study additional sequence elements associated with the CAG repeat.

Results – Expanded HTT Alleles

![Plot](image)

Fig. 5. CAG repeat count estimate plots for expanded HTT target alleles. These results come from a separate experiment testing the same No-Amp protocol with a multiplex of samples containing expansion alleles. Protocol input was 10 μg per sample. Expanded alleles were detected and matched the expected values within a few repeat units. The count variation in each size class observed with the No-Amp protocol could reflect biological variation present in different cell populations (mosaicism).

Conclusions

- No-Amp Targeted SMRT Sequencing is a powerful tool for capturing and assessing sequence variation in repeat expansion elements.
- Thousands of target reads can be obtained without the need for amplification enabling an accurate assessment of the natural variation present in a population of repeat elements which is often impossible to do with a PCR-based approach.
- With less than 3 μg of input DNA, greater than 1,000 HTT target molecules can be sequenced. We estimate that an accurate repeat count can be achieved with as little as 50 reads per allele. These results indicate that higher sample multiplexing than demonstrated here is possible.

References


No-Amp Targeted Sequencing

- PacBio Webpage: www.pacb.com/applications/targeted-sequencing/no-amp-targeted-sequencing

Acknowledgments

The authors would like to thank everyone who helped generate data for the poster.