Gisela Mir, Tim Semple, Jason Li, Richard Tothill, Andrew Fellowes

Library preparation, targeted-hybridisation capture and sequencing of limiting amounts of FFPE samples.

Abstract

Massively-parallel (next-generation) sequencing is now being rapidly adopted for the detection of pathogenic mutations in cancer patient samples. One of the great challenges faced in this process is the often limiting amounts and poor quality DNA that can be obtained from formalin-fixed and paraffin-embedded (FFPE) tissues. With the steady improvements to DNA sequencing and the chemistry used for library preparation, there are now a number of commercial options for processing and sequencing of FFPE samples. However, these methods still require thorough testing before they can be used routinely in a clinical setting under strict quality assurance guidelines. Targeted-hybridisation capture is routinely used in research and has been established as one of the most cost-effective methods to explore large genomic regions for diverse germ-line and somatic variations. We therefore aimed to explore the sensitivity, accuracy and reliability of different library preparation methods combined with hybridisation capture for sequencing of FFPE samples from pathology archives.

In this technical work we selected a good quality reference sample (NA12878, HapMap) and several tumor FFPE clinical samples to initially compare four library preparation methods (SureSelect XT, Ovation SP Ultralow, Kapa and ThruPLEX-FD) and two capture methods (SureSelect XT and NimbleGen SeqCap EZ) at different starting amounts of DNA (20 ng to 500 ng). We used a 3 Mb custom panel designed to capture 600 genes frequently mutated in cancer, some known fusions and one virus. Captured libraries were sequenced with an Illumina HiSeq2500 at 100 paired-end reads aiming a mean coverage >500x.

Although successful libraries were generated from 20 ng of DNA, using 50 ng or less DNA resulted in low complexity libraries and inefficient captures that produced high duplication levels and/or low on-target when aligned to the targeted regions. Underperformance at these conditions was also true for the good quality reference sample. When starting with 100 ng of DNA duplication rates were still higher than those expected in optimal situations but coverage was good for most of the targeted regions, with results very similar for the two capture platforms tested; at 300 ng we reached standard performance QC metrics. The highest variability in library efficiency between FFPE samples was irrespective of the library preparation method, emphasizing the importance of a good sample QC to estimate capture and sequencing success at an early stage.

Variant calling and copy-number variation analyses are being run to determine the sensitivity of the method. We aim to increase the number of samples and run validation tests to establish the combination of wet-lab and bioinformatics that best suits the use of targeted-capture and high-throughput sequencing for the study of FFPE samples in a clinical setting.