CENTRE FOR INTEGRATIVE SEQUENCING

Development of methods for NGS based RNA-Seq and DNA-Seq on FFPE materials

Jakob Hedegaard[1], Kasper Thorsen[1], Mette Katrine Lund[2], Anne-Mette K. Hein[3], Stephen Jacques Hamilton-Dutoit[4], Henrik Hager[4], Søren Vang[1], Iver Nordentoft[1], Karin Birkenkamp-Demtröder[1], Liselotte Kahns[3], Mogens Kruhøffer[2], Brian Sørensen[2], Bjarne Knudsen[3], Claus Lindbjerg Andersen[1], Karina Dalsgaard Sørensen[1], Jakob Skou Pedersen[1], Torben Falck Ørntoft[1], Lars Dyrskjøt[1]

1 Department of Molecular Medicine (MOMA), Molecular Diagnostic Laboratory, Aarhus University Hospital, Skejby, Brendstrupgaardsvej 100, DK-8200 Aarhus N, Denmark
2 AROS Applied Biotechnology A/S, Science Park Skejby, Brendstrupgaardsvej 102, 8200 Aarhus N, Denmark
3 CLC bio, Finlandsgade 10-12, Katrinebjerg, 8200 Aarhus N, Denmark
4 Institute of Pathology, Aarhus University Hospital, Nørrebrogade 44, DK-8000 Aarhus C, Denmark

Formalin-fixed, paraffin-embedded (FFPE) patient tissue samples stored globally in pathology archives represent an invaluable resource for clinical research. Unfortunately, nucleic acids isolated from FFPE tissues tend to be fragmented and chemically modified, interfering with many classical molecular analyses. Since Next Generation Sequencing (NGS) technologies rely on randomly fragmented nucleic acids, we have focused a systematic study on the potential use of FFPE samples in NGS-based clinical studies. Available FFPE extraction kits were evaluated with respect to purification of RNA and DNA from different tissue types. Although the quality and quantity of the extracted nucleic acids was found to be highly dependent on the purification method used, it was possible to isolate high molecular DNA and RNA from recent FFPE specimen (fixed within one year). From older FFPE specimen, the quality and quantity was though found to be reduced. Extracted DNA and RNA from matching cryopreserved and FFPE specimens were used for the preparation of targeted genomic (Exome-Seq) and whole transcriptome (RNA-Seq) sequencing libraries. Illumina's TruSeq exome preparation protocol was used for the preparation of Exome-Seq libraries and was found to provide functional libraries for some samples, while others failed during amplification of the libraries, most likely due to the modifications of the DNA. In contrast, using depletion of rRNA followed by preparation of strand-specific RNA-Seq libraries (Ribo-Zero and ScriptSeq, Eppicentre/Illumina) succeeded in all samples, even from tissues sampled and fixed two decades ago. Analysis of the obtained paired-end sequencing data (2x101, multiplex, HiSeq 2000) revealed a slightly lower fraction of reads mapping in FFPE compared to FF libraries and a slightly higher fraction of non-perfect matches in FFPE compared to FF libraries. We hence recommend sequencing FFPE libraries deeper than FF and use long reads to compensate for the higher error rate. We found good concordance in variant calls between exome data from FF and FFPE libraries, but observed some systematic as well as random biases. The expression profiles obtained from the paired FF-FFPE samples were also found to be in concordance – even from old (12 years) samples. This work demonstrates the potential application of NGS to study FFPE samples. The results are promising regarding the possible use of the concealed information of archived FFPE specimens for large scale retrospective NGS-based genomic studies.