Complex rearrangements and extrachromosomal circular DNA in Colorectal Cancer

Master's Thesis in Bioinformatics (30 ECTS)

Elías Calandra Jiménez, Studienr: 202303411

Supervisors:

Claus Lindbjerg Andersen, Professor Department of Molecular Medicine, Aarhus University

Nicolai Juul Birkbak, Professor Department of Molecular Medicine, Aarhus University

ABSTRACT

Genomic instability in cancer often manifests as structural variants (SVs) and extrachromosomal circular DNAs (ecDNAs), which can drive oncogene amplification and therapeutic resistance. In this study three SV callers (SvABA, Manta, Delly), the genome-graph tool JaBbA, and AmpliconArchitect (AA) were benchmarked on the NCI-H2009 cell line sequenced by two library preparations. Caller concordance, karyotype support, and ecDNA detection were assessed. The tools were then applied to ten colorectal cancer (CRC) tumors. A difference in the number of structural variants called was detected between the two sequencing strategies, while AmpliconArchitect and JaBbA were not highly affected by the sequencing. In CRC, unstable tumors harbored more SVs (mean 432 vs. 198, p=0.06), and one contained a ecDNA on 6p21 amplifying several oncogenes. Recurrent deletions were found disrupting the gene *PRKN* in 5 of 10 tumors. Overall, integrating SV calling with JaBbA and AA provides a great framework for complex rearrangements and ecDNA detection for short-read WGS, although long-read sequencing and empirical validations would help in the ecDNA reconstruction.

ABBREVIATIONS

AA: AmpliconArchitect

CRC: Colorectal cancer

ecDNA: Extrachromosomal DNA/Extrachromosomal circular DNA

dm: double minute/extrachromosomal DNA

BFB: Breakage-fusion-bridge

SVs: Structural Variants

SGA: String Graph Assembly

LOESS: locally weighted regression regression

INDEX