**“A functional homologous recombination assay predicts primary chemotherapy response and long-term survival in ovarian cancer patients”**

Manuela Tumiati1\*, Sakari Hietanen2, Johanna Hynninen2, Elina Pietilä1, Anniina Färkkilä1,4, Katja Kaipio3, Pia Roering3, Kaisa Huhtinen3, Amjad Alkodsi1, Yilin Li1, Rainer Lehtonen1, Erdogan Pekcan Erkan1, Minna M. Tuominen1, Kaisa Lehti1,5, Sampsa K. Hautaniemi1, Anna Vähärautio1, Seija Grénman2, Olli Carpén1,3, and Liisa Kauppi1\*

**Supplementary materials and methods**

**Whole genome sequencing analysis**

Raw paired-end sequencing reads were subject to quality and adapter trimming using Trimomatic [1] (minimum quality: 20, minimum read length: 50, sliding window: '5:20', head crop: 5). Next, quality-controlled reads were mapped to human reference genome GRCh38.d1.vd1 using bwa-mem with default parameters. Resulted bam files were sorted and marked for duplicates using PICARD tools (<https://broadinstitute.github.io/picard/>) and then subjected to base quality score recalibration using the genome analysis toolkit GATK (<https://software.broadinstitute.org/gatk/>).

Somatic base substitutions and short insertions and deletions (indels) were detected from matched tumor-normal pairs using MuTect2 [2] (GATK v3.6). We used COSMIC v79 and dbSNP v144 variants as white and black lists for MuTect2 respectively. For MuTect2 normal panel input, we constructed a panel of normal variants using MuTect artifact detection mode from a set of 39 normal whole genome samples from TCGA ovarian cancer project. Subsequently, we kept variants marked as PASS by MuTect2. Since multiple tumor samples were sequenced from a single patient, we used forced-calling strategy to enhance sensitivity. For each variant detected by MuTect2 in one tumor sample, we counted the reads covering the reference and variant allele for that variant in all other tumor samples from the same patient, and HaplotypeCaller (GATK v3.6) in “genotype given alleles” mode was used for that purpose. Finally, we used Annovar [3] for annotation and functional prediction.

Somatic copy number alterations were detected from whole-genome sequencing following the AscatNgs workflow [4]. The ASCAT algorithm [5] allows estimation of absolute allelic copy number together with purity and average ploidy. Genes with absolute total copy number > 1.5 \* ploidy were considered amplified and those with total copy number < 0.5 \* ploidy were considered deleted. Loss of heterozygosity (LOH) was determined when minor copy number is 0.

**Whole exome sequencing analysis**

Raw sequencing reads were mapped to human reference genome GRCh38.d1.vd1 using bwa-mem with default parameters. Resulted bam files were sorted and marked for duplicates using PICARD tools and then subjected to base quality score recalibration using the genome analysis toolkit (GATK).

Base substitutions and indels were detected from tumor samples using MuTect2 in tumor-only mode (GATK v3.6). Cosmic, dbSNP and a panel of normal variants were used as inputs to MuTect2 as in whole genome analysis. Annovar was used for variant annotation and functional prediction.

We used the GATK4 implementation of RecapSeg (<https://github.com/broadinstitute/gatk-protected/blob/master/docs/CNVs/CNV-methods.pdf>) to obtain allelic copy number profiles from exome sequencing data. Read depths at informative capture targets in cancer samples were normalized against a panel of normals constructed from 10 unmatched normal samples. Unmatched normal samples have been sequenced and processed with the same protocols and pipelines as in the tumor samples. The copy-ratio profiles resulted from normalization were then segmented using the circular binary segmentation algorithm (CBS) [6]. To incorporate allelic information with the resulted relative copy number profiles, heterozygous SNPs in a matched normal are typically needed. Since matched normals were not available, we used GATK4 'GetBayesianHetCoverage' utility to predict heterozygous SNPs using only tumor samples. Subsequently, resulted allelic copy number profiles were used as inputs to the ASCAT algorithm allowing estimation of purity and ploidy. Amplifications, deletions and LOH were then determined in the same way as in whole-genome sequencing.

**Targeted exome sequencing**

For two samples, DNA was extracted from unstained FFPE sections and sequenced against a panel of 315 cancer-related genes and 28 genes often rearranged in cancers with FoundationOneTM (Foundation Medicine Inc., Cambridge, MA). Analysis of base substitutions, indels, copy number alterations and rearrangements was performed as previously described [7].

**LOH and mutational signature analysis**

The LOH analysis was performed using the algorithm already described by Abkevich *et al* [8], while mutational signatures were obtained from whole genome and whole exome sequencing data as previously described [9-11].

1. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-20.
2. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213-9.
3. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
4. Raine KM, Van Loo P, Wedge DC, et al. ascatNgs: Identifying Somatically Acquired Copy-Number Alterations from Whole-Genome Sequencing Data. *Curr Protoc Bioinforma*. 2016;56:15.9.1-15.9.17.
5. Van Loo P, Nordgard SH, Lingjærde OC, et al. Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci U S A*. 2010;107(39):16910-5.
6. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*. 2004;5(4):557-72.
7. Frampton GM, Fichtenholtz A, Otto GA, et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol*. 2013;31(11):1023-31.
8. Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer*. 2012;107(10):1776-1782.
9. Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational processes molding the genomes of 21 breast cancers. *Cell*. 2012;149(5):979-993.
10. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.
11. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering Signatures of Mutational Processes Operative in Human Cancer. *Cell Rep*. 2013;3(1):246-259.