Jul 01, 2019

# Targeted PCR-based deep sequencing of cfDNA with unique molecular indices by a customized QIAseq Targeted DNA Panel

Corinna Keup[1], Peter Hahn[2], Siegfried Hauch[2], Markus Sprenger-Haussels[2], Mitra Tewes[3], Pawel Mach[1], Ann-Kathrin Bittner[1], Rainer Kimmig[1], Sabine Kasimir-Bauer[1], Karim Benyaa[2]

[1]Department of Gynecology and Obstetrics, University Hospital of Essen, Germany, [2]QIAGEN GmbH, Hilden, Germany, [3]Department of Internal Medicine (Cancer Research), University Hospital of Essen, Germany

Working dx.doi.org/10.17504/protocols.io.trfem3n

Corinna Keup

ABSTRACT

This protocol describes in detail all steps of the library preparation from cfDNA with the QIAseq Targeted DNA Panel (QIAGEN GmbH, Hilden, Germany) and the subsequent data analysis by QIAGEN Biomedical Genomics Workbench and Ingenuity Variant Analysis (both QIAGEN GmbH, Hilden, Germany). Clinically relevant results obtained by this protocol will be published soon.

Since the input volume of the cfDNA eluate used in the beginning was 20 µl (and not maximal 16.75 µl as stated in the QIAseq Targeted DNA Panel handbook version Mai 2017), we adjusted the reagent volumes.

We used a customized QIAseq Targeted DNA Panel analyzing all exons of the 17 genes of interest (namely *AKT1, AR, BRCA1, BRCA2, EGFR, ERBB2, ERBB3, ERCC4, ESR1, KRAS, FGFR1MUC16, PIK3CA, PIK3R1, PTEN, PTGFR, TGFB1*).

In contrast to the QIAseq Targeted DNA Panel handbook, we used 20 cycles of amplification for the universal PCR.

For stringent data analysis, we defined criteria to exclude libraries with unsufficient sequencing qualities and we here also describe the bioinformatical filter used within the Ingenuity Variant Analysis software (QIAGEN GmbH, Hilden, Germany) for variant exclusion.

EXTERNAL LINK

https://www.qiagen.com/kr/resources/resourcedetail?id=8907edbe-a462-4883-ae1b-2759657e7fd0&lang=en;
https://www.qiagen.com/us/shop/sequencing/qiaseq-solutions/qiaseq-targeted-dna-panels/#orderinginformation;
https://www.qiagen.com/us/shop/sample-technologies/dna/qiaamp-minelute-ccfdna-kits/#orderinginformation

THIS PROTOCOL ACCOMPANIES THE FOLLOWING PUBLICATION

Publication: Keup, C.; Benyaa, K.; Hauch, S.; Sprenger-Haussels, M.; Tewes, M.; Mach, P.; Bittner, A.-K.; Kimmig, R.; Hahn, P.; Kasimir-Bauer, S. Targeted deep sequencing revealed variants in cell-free DNA of hormone receptor-positive metastatic breast cancer patients. Cellular and molecular life sciences : CMLS 2019. DOI: 10.1007/s00018-019-03189-z. AND Publication: Keup, C.; Storbeck, M.; Hauch, S.; Hahn, P.; Sprenger-Haussels, M.; Tewes, M.; Mach, P.; Hoffmann, O.; Kimmig, R.; Kasimir-Bauer, S. Cell-Free DNA Variant Sequencing Using CTC-Depleted Blood for Comprehensive Liquid Biopsy Testing in Metastatic Breast Cancer. Cancers 2019, 11, 238. DOI: 10.3390/cancers11020238.

DOCX

Targeted PCR-based deep sequencing of cfDNA with unique molecular indices by a customized QIAseq Targeted DNA Panel.docx

GUIDELINES

see QIAseq Targeted DNA Panel handbook (QIAGEN GmbH, Hilden, Germany)

SAFETY WARNINGS

see QIAseq Targeted DNA Panel handbook (QIAGEN GmbH, Hilden, Germany)

BEFORE STARTING

see QIAseq Targeted DNA Panel handbook (QIAGEN GmbH, Hilden, Germany)

## Starting material

1 cfDNA was isolated from preferably 4 ml plasma by affinity-based binding to magnetic beads according to the manufacturer's instructions (QIAamp MinElute ccfDNA Kit; QIAGEN GmbH, Hilden, Germany). cfDNA was eluted in 22 µl ultraclean water and stored at -20°C.cfDNA concentrations were analyzed by the Agilent Chip High Sensitivity DNA (Santa Clara). For the calculation of the cfDNA yield, only the concentration of fragments with a length between 100 and 700 bp were added up.

## Library construction

2 The input amount preferred for library preparation was in the range of 30-60 ng, but cfDNA samples with lower input were also included for the library preparation. The library construction was done using the QIAGEN QIAseq Targeted DNA Panel protocol, which is separated in four steps all performed on ice. For end-repair and a-addition, 20 µl of the template cfDNA was mixed with 3.13 µl of 10x QIAGEN Fragmentation Buffer, 0.94 µl QIAGEN FERA Solution, 6.25 µl 5x Fragmentation Enzyme Mix and 1.6 µl QIAGEN FG Solution, which inhibits the fragmentation enzyme but has no effect on the end-repair and a-addition reaction. The reaction mix was incubated for 14 min at 32°C and 30 min at 72°C. The incubated reaction mix was transferred in an adapter ligation master mix, which contained 12.75 µl 5x QIAGEN Ligation Buffer, 6.4 µl QIAGEN DNA Ligase, 9.2 µl QIAGEN Ligation Solution and added 0.64 µl QIAGEN IL-N7## UMI adapter per reaction (same IL-N7## adapter for eight samples) [1,2]. Since building of a consensus sequence of amplified fragments with the same UMI, added in this step of the library preparation, excludes PCR artefacts. The mix was incubated in a thermocycler for 15 min at 20°C without heated lid. After the adapter ligation, a two stepped magnetic bead cleanup was performed, using 136.3 µl (1.4x) QIAGEN QIAseq Beads and 36.5 µl RNase-free water for the first cleanup step, washed twice with 80% ethanol and dried the beads for 10 min, before cfDNA was eluted in 52 µl RNase-free water. For the second clean up, 70 µl of the mentioned magnetic beads were used, washed twice, dried and cfDNA was eluted in 12 µl RNase-free water. Afterwards, targeted enrichment of the cfDNA library was performed by transferring the eluted library to a targeted enrichment PCR master mix, which contained 4 µl 5x QIAGEN TEPCR Buffer, 0.8 µl QIAGEN IL-Forward primer, 0.8 µl HotStarTaq Polymerase and 5 µl customized QIAGEN QIAseq Targeted DNA Panel primer. The customized panel was designed to amplify all coding regions of the following genes: AKT1, AR, BRCA1, BRCA2, EGFR, ERBB2, ERBB3, ERCC4, ESR1, KRAS, FGFR1, MUC16, PIK3CA, PIK3R1, PTEN, PTGFR, TGFB1 and showed high specificity and uniformity (Supplementary Table S3). The reaction mix was incubated with following conditions: 13 min at 95°C, 2 min at 98°C, 8 cycles of 15 seconds at 98°C and 10 min at 68°C, 5 min at 72°C and 5 min at 4°C. The cycling was followed by a magnetic bead cleanup with 108 µl (1.2x) QIAGEN QIAseq beads and 70 µl RNase-free water, using 16 µl RNase-free water for the elution of the library. The last step of the library preparation is the universal PCR amplification, in which the cfDNA library was added to a master mix, which contained 4 µl 5x QIAGEN UPCR Buffer, 1 µl HotStartTaq Polymerase and 1.6 µl RNase-free water and each reaction mix with the same IL-7## adapter was pipetted into separate wells of an IL-5## adapter plate to add sample specific tags [2]. The following conditions for the cycling were used: 13 min at 95°C, 2 min at 98°C, 20 cycles of 15 seconds at 98°C and 2 min of 60°C, 5 min at 72°C and 5 min at 4°C. At the end, a magnetic bead cleanup with 108 µl (1.2x) QIAGEN QIAseq beads and 70 µl RNase-free water was performed. The final targeted enriched cfDNA library was eluted in 30 µl RNase-free water and stored at -20°C.

## Library quality control and quantification

3 To check the quality of the libraries in the range of 200bp to 3000bp the Agilent Chip High Sensitivity DNA (Santa Clara) was used. A qPCR was performed for calculation of the library yield, using 2 µl 10 µM IL-Forward/IL-Reverse Primer Mix and 4 µl of the 1:25000 and 1:250000 diluted library in QuantiNova SYBR Green protocol (QIAGEN). 6 different concentrations (20 pM to 2e-4 pM) Illumina Library Quantification Standard were used for the standard curve. Cycling on the qPCR Cycler ABI 7900 was as follows: 2 min at 95°C, 40 cycles with 5 sec at 95°C and 10 sec at 60°C, plus melt curve.

## Sequencing

4 Libraries were diluted to 2 nM (MiSeq) or 4 nM (NextSeq) and libraries with a lower yield were excluded. All pooled libraries were analyzed on Illumina sequencers (Illumina MiSeq Sequencer with the MiSeq Reagent Kit v2, 2x150bp reads or Illumina NextSeq Sequencer with a NextSeq 550 System High-Output Kit, 2x150bp reads) with the QIAseq Customer Primer 1A.

## Data analysis/ Bioinformatical analysis

5 To check whether the sequencing data are reliable and comparable, it is important to control different sequencing quality parameters and specify exclusion criteria. All samples with less than 5 million read fragments were excluded, given that a low read amount is often caused

by bad library quality. In a statistical point of view, the comparability of data from a sample with a low and a sample with a high read fragment output are statistically insufficient [3,4]. Furthermore all samples with an UMI coverage lower than 400 were excluded, to ensure that the read fragments were as diverse as possible for a statistical powerful statement [1]. To guarantee a balanced distribution of the reads on all target regions, samples were excluded, if less than 95% of the target region was covered with at least 5% of the mean UMI coverage.The QIAGEN Biomedical Genomics Workbench was used to build alignments of mapped reads to the targeted regions (reference genome hg19), for the analysis of the UMIs and for first variant calling [1]. The Ingenuity Variant Analysis (IVA; QIAGEN) was used for the quality scoring, detailed variant calling with customized filter (described below) and data interpretation. The IVA used five different filters for exclusion of called variants. The confidence filter excluded all variants with a call quality below 20. Variants with a prevalence of >3% in the normal population (reference data bases: 1. Allel Frequency Community (gnomAD&CGI), 2. 1000 Genomes Project, 3. ExAC and 4. NHLBI ESP exomes) were excluded, unless the variant was already known to be a pathogenic common variant, to identity rare variants potentially associated with the evaluated condition of the tested cohort (common variant filter). The filter also only kept variants that were associated with gain of function or which inheritance pattern (homozygous, compound heterozygous, haplosufficient, hemizygous, het-ambiguous or heterozygous) occurred in at least 50% of all cases at gene level (genetic analysis filter). Only variants located no more than 20 bases into intron were included (part of the predicted deleterious filter). The cancer driver variant filter kept only variants that were found in 1.) Cancer-associated mouse knockout phenotypes, 2.) cancer-associated cellular processes with any directionality 3.) cancer-associated pathways with any directionality 4.) cancer therapeutic targets 5.) published cancer literature variant and gene level findings 6.) Known or predicted cancer subnetwork regulatory sites 7.) COSMIC at a frequency greater than or equal to 0.01% 8.) TCGA at a frequency greater than or equal to 0.01%. Moreover, the IVA classified the variants according to their clinical significance (by predicted or observed evidence) into five groups: pathogenic, likely pathogenic, uncertain significance, likely benign and benign (as defined by the American College of Medical Genetics and Genomics; part of the predicted deleterious filter) [5].

## References


6 [1] Xu C, Nezami Ranjbar MR, Wu Z, DiCarlo J, Wang Y. Detecting very low allele fraction variants using targeted DNA sequencing and a novel molecular barcode-aware variant caller. BMC genomics 2017;18:5.[2] Meyer M, Stenzel U, Myles S, Prüfer K, Hofreiter M. Targeted high-throughput sequencing of tagged nucleic acid samples. Nucleic acids research 2007;35:e97.[3] Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. Nature reviews. Genetics 2011;12:443–51.[4] Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing-Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. The Journal of molecular diagnostics : JMD 2017;19:341–65.[5] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in medicine : official journal of the American College of Medical Genetics 2015;17:405–24.