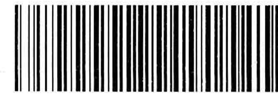

D0000743614
Befund CeGaT
M-0081273 1978-01-27
Banat Mohammed Hasan Mousa

CeGaT

CeGaT GmbH | Paul-Ehrlich-Str. 23 | D-72076 Tübingen | Germany

Dr. Bartolome Oliver Abadal
Neurospine Center Sl
Calle Vilana 12
08022 BARCELONA
SPAIN

Name	Banat, Mohammed Hasan Mousa (*27.01.1978)
Sex	Male
Patient-ID	81273
Sample receipt	13.01.2021 (EDTA blood) 13.01.2021 (Tumor-FFPE)
Report date	09.02.2021

cc: Dr. med. Dr. rer. nat. Saskia Biskup, FÄ für Humangenetik, Praxis für Humangenetik Tübingen, Paul-Ehrlich-Str. 23, 72076 Tübingen

Neoepitope prediction and peptide selection – Banat, Mohammed Hasan Mousa (*27.01.1978)

Indication **Left temporal astrocytoma, WHO grade II (ID 02/2009)**

Material **Tumor tissue: Brain biopsy with neoplastic glial proliferation**
Sample collection 01/2021
DNA and RNA isolation from tumor in FFPE (FFPE-ID TB 2021 53 A1) after macrodissection with estimated tumor content of 90% (HE staining)

Normal tissue: EDTA blood

Order

1. Somatic molecular genetic analysis of a tumor tissue sample:
   Tumor exome-analysis TUM02, evaluation of somatic variants of potential clinical relevance in 766 tumor-related genes (TUM01)
2. **Tumor transcriptome analysis, neoepitope prediction and peptide selection**

Dear Dr. Abadal,

Thank you for your request for molecular genetic analysis, neoepitope prediction and peptide selection.

RESULTS

19 peptides were selected based on neoepitope prediction results of your patient's tumor sequencing data.

The selected peptides are predicted to activate not only cytotoxic T cells but also T helper cells. Therefore, in addition to short peptides (8-12 amino acids) potentially binding to HLA class I molecules also long peptides (~17 amino acids) potentially binding to HLA class II molecules were included.

Individual peptide selection based on neoepitope prediction:

No	Peptide	Gene and Coding_info	NAF DNA	NAF RNA	FPKM RNA	HLA allele
1	PIIIGHHAY	IDH1:NM_005896.3:c.395G>A:p.R132H	0.38	0.35	12.85	A*29:02
2	FIYLSSNCF	CPE:NM_001873.4:c.989A>T:p.N330I	0.40	0.48	200.96	A*29:02, B*15:03, C*12:03
3	ALDPANGYMY	LRP5:NM_002335.4:c.484C>A:p.H162N	0.21	0.26	3.55	A*01:01, A*29:02
4	AGVLNAGSY	CLASP2:NM_015097.3:c.1898C>T:p.A633V	0.17	0.23	42.84	B*15:03
5	SGYVRPIPV	SMCHD1:NM_015295.3:c.664G>A:p.V222I	0.16	0.20	20.86	C*12:03
6	VAVETMHKM	TOGARAM1:NM_015091.4:c.4651C>G:p.L1551V	0.15	0.43	7.87	B*15:03, C*12:03
7	EKAWNVYPY	PITPNB:NM_012399.5:c.269C>T:p.A90V	0.08	0.03	21.47	B*15:03
8	PEHLKDESA	MAP6:NM_033063.1:c.1940C>T:p.P647L	0.22	0.20	5.60	B*45:01
9	WVKPIIIGHHAYGDQYR	IDH1:NM_005896.3:c.395G>A:p.R132H	0.38	0.35	12.85	class II
10	VPMVLVGNNCDLPSRTV	KRAS:NM_004985.5:c.351A>T:p.K117N	0.12	0.13	8.09	class II
11	YSVPGGMQDFIYLSSNCFE	CPE:NM_001873.4:c.989A>T:p.N330I	0.40	0.48	200.96	class II
12	RSGRLGAGVLNAGSYAS	CLASP2:NM_015097.3:c.1898C>T:p.A633V	0.17	0.23	42.84	class II
13	DHSGYVRPIPVPRSLNS	SMCHD1:NM_015295.3:c.664G>A:p.V222I	0.16	0.20	20.86	class II
14	AVNSDLSSSLEERMQSP	NOL4:NM_003787.4:c.722A>G:p.N241S	0.19	0.33	9.76	class II
15	DPMVPEHLKDESAMAT	MAP6:NM_033063.1:c.1940C>T:p.P647L	0.22	0.20	5.60	class II
16	RAIALDPANGYMYWTDW	LRP5:NM_002335.4:c.484C>A:p.H162N	0.21	0.26	3.55	class II
17	DVNQHGSDPESEETRKL	IWS1:NM_017969.3:c.358T>C:p.S120P	0.08	0.08	20.97	class II
18	VFHEKAWNVYPYCRTIV	PITPNB:NM_012399.5:c.269C>T:p.A90V	0.08	0.03	21.47	class II
19	NSKVNLVAVETMHKMIP	TOGARAM1:NM_015091.4:c.4651C>G:p.L1551V	0.15	0.43	7.87	class II

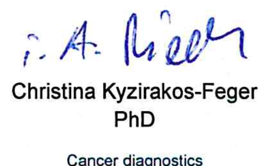
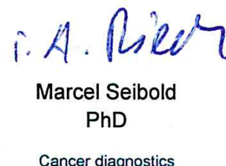
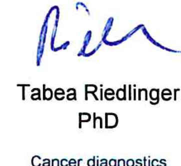
NAF: Novel allele frequency, the frequency with which the mutated allele occurs in the sequencing (1 is 100%). The observed frequencies are influenced by the tumor content and do not correlate directly to the variant frequency in the tumor. **FPKM:** Fragments Per Kilobase Million, expression level metric.

Based on the DNA sequencing analysis of the blood sample the HLA genotype was determined to be:

HLA-A*01:01, HLA-A*29:02, HLA-B*15:03, HLA-B*45:01, HLA-C*06:02, HLA-C*12:03.

No peptides that met our selection criteria were predicted to bind to HLA-C*06:02.

Yours sincerely,

Saskia Biskup MD, PhD Consultant for Human Genetics	Christina Kyzirakos-Feger PhD Cancer diagnostics	Marcel Seibold PhD Cancer diagnostics	Tabea Riedlinger PhD Cancer diagnostics

ADDITIONAL INFORMATION

Methods

DNA/RNA isolation: DNA and RNA from tumor tissue was isolated after macrodissection at CeGaT Tübingen after evaluation by a pathologist.

The pathological services (confirmation of the histological diagnosis and determination of the tumor content) were carried out on our behalf by a specialist in pathology. Pathology services are not within the scope of the ISO 15189 accreditation.

DNA sequencing: The coding and flanking intronic regions were enriched using in solution hybridization technology and were sequenced using the Illumina HiSeq/NovaSeq system.

Computational analysis: Illumina bcl2fastq2 was used to demultiplex sequencing reads. Adapter removal was performed with Skewer. The trimmed reads were mapped to the human reference genome (hg19) using the Burrows Wheeler Aligner. Reads mapping to more than one location with identical mapping score were discarded. Read duplicates that likely result from PCR amplification were removed. The remaining high-quality sequences were used to determine sequence variants (single nucleotide changes and small insertions/deletions). The variants were annotated based on several internal as well as external databases.

Genetic data evaluation: Only variants (SNVs/small indels) with a novel allele frequency (NAF) of ≥ 5% in the tumor sample within the coding regions and their adjacent intronic regions (-/+ 8 base pairs) were evaluated. The sensitivity of the test is dependent on the tumor content of the analyzed material, the sample quality, and the sequencing depth.

Variants are named according to the HGVS recommendations. The sample fulfilled our quality criteria upon arrival and during/after each processing step in the laboratory. The procedure described above was developed and validated in-house (Laboratory developed test; LDT). A minimal tumor content of 20 % was taken as a basis.

RNA sequencing: Library preparation was performed using the TruSeq Total RNA (RiboZero rRNA removal Kit) or-SMARTer Stranded Total RNA Library Kit and subsequently analyzed using high-throughput sequencing on the HiSeq/NovaSeq system (Illumina). Mapping of sequencing reads to the hg19 reference genome was performed with STAR (Version 2.5.2b). Gene expression analysis (counting of aligned reads per gene, calculation of normalized read counts and calculation of FPKM values) was done with DESeq2 (Love et al., 2014, PMID: 25516281) in R (R Core Team 2015).

Prediction of neoepitopes: Somatic single nucleotide variants and small InDels (in frame or frameshift insertions and/or deletions) only present in the tumor but absent in the normal tissue of the patient were identified by sequencing analysis. Typing of HLA class I is performed on sequencing results from patient's normal tissue using OptiType (Szolek et al., 2014, PMID: 25143287). Identified somatic variants and in-phase germline variants are translated into peptide sequences and sent to the Center for Bioinformatics Tübingen, Dept. of Computer Science, University Tübingen for epitope prediction. MHC class I epitopes are predicted using SYFPEITHI, netMHC-4.0 and netMHCpan-3.0 (Rammensee et al., 1999, PMID: 10602881; Andreatta et al., 2016, PMID: 26515819; Nielsen et al., 2016, PMID:27029192). Peptides containing somatic variants that are classified as binder by at least one prediction method are further evaluated. The respective thresholds for classification as binder are defined as <500 nM for netMHC and netMHCpan as well as >50% of maximal score for SYFPEITHI. Peptides resembling a wildtype sequence in the human proteome (based on UniProtKB/Swiss-Prot, human, 9/7/14) are excluded.

Selection of peptides: Peptides derived from genes most probably not expressed in the patient's tumor entity were excluded. For this purpose, expression data for the respective variant were analyzed using RNA sequencing data of the tumor sample. Putative HLA class I epitopes with a high HLA class I binding prediction score derived from variants with high allele frequencies were selected. Peptides predicted to bind to different HLA class I molecules of the patient were prioritized. Peptides which are predicted to bind to several HLA types were further prioritized. Putative HLA class II epitopes with a length of +/-17 amino acids were designed to contain variants with high allele frequencies. Peptides spanning variants in possible tumor drivers were prioritized. Peptides with a high percentage of hydrophobic amino acids, peptides with a high probability for gelation or dimerization were excluded to avoid solubility problems in an aqueous solution and problems during synthesis. However, since it is impossible to reliably predict these characteristics, it cannot be guaranteed that all selected peptides can be synthesized and solubilized.

The bioinformatically identified somatic variants corresponding to all selected peptides were manually reviewed in the sequencing data and filtered for false positives.

A file containing all peptides predicted to bind to your patient's HLA class I molecules can be sent to you by E-mail upon request.

Befund CeGaT
M-0081273 1978-01-27
Banat Mohammed Hasan Mousa

CeGaT GmbH | Paul-Ehrlich-Str. 23 | D-72076 Tübingen | Germany
Tel: +49 7071 565 44 55 | Fax: +49 7071 565 44 56 | info@cegat.com | www.cegat.com
Court District Stuttgart - HRB 729958 | VAT No: DE265504070
Managing Directors: Dr. Dr. Saskia Biskup, Dr. Dirk Biskup, Christian Kraft
VR Bank Tübingen eG | IBAN: DE20 6406 1854 0021 2890 00 | SWIFT / BIC: GENODES1STW

CAP ACCREDITED
COLLEGE of AMERICAN PATHOLOGISTS
CLIA CERTIFIED ID: 99D2130225
Accredited by the College of American Pathologists

DAkkS Deutsche Akkreditierungsstelle
D-ML-13206-01-00
Accredited according to DIN EN ISO 15189:2014