

RESEARCH

Open Access



Mutational load in carotid body tumor

Anna V. Kudryavtseva^{1*}, Elena N. Lukyanova¹, Dmitry V. Kalinin², Andrew R. Zaretsky¹, Anatoly V. Pokrovsky², Alexander L. Golovyuk², Maria S. Fedorova¹, Elena A. Pudova¹, Sergey L. Kharitonov^{1,3}, Vladislav S. Pavlov¹, Anastasiya A. Kobelyatskaya¹, Nataliya V. Melnikova¹, Alexey A. Dmitriev¹, Andrey P. Polyakov³, Boris Y. Alekseev³, Marina V. Kiseleva³, Andrey D. Kaprin³, George S. Krasnov¹ and Anastasiya V. Snezhkina¹

From 11th International Multiconference "Bioinformatics of Genome Regulation and Structure\Systems Biology" - BGRS\SB-2018

Novosibirsk, Russia. 20-25 August 2018

Abstract

Background: Carotid body tumor (CBT) is a rare neoplasm arising from paraganglion located near the bifurcation of the carotid artery. There is great intra-tumor heterogeneity, and CBT development could be associated with both germline and somatic allelic variants. Studies on the molecular genetics of CBT are limited, and the molecular mechanisms of its pathogenesis are not fully understood. This work is focused on the estimation of mutational load (ML) in CBT.

Methods: Using the NextSeq 500 platform, we performed exome sequencing of tumors with matched lymph node tissues and peripheral blood obtained from six patients with CBT. To obtain reliable results in tumors with low ML, we developed and successfully applied a complex approach for the analysis of sequencing data. ML was evaluated as the number of somatic variants per megabase (Mb) of the target regions covered by the Illumina TruSeq Exome Library Prep Kit.

Results: The ML in CBT varied in the range of 0.09–0.28/Mb. Additionally, we identified several pathogenic/likely pathogenic somatic and germline allelic variants across six patients studied (including TP53 variants).

Conclusions: Using the developed approach, we estimated the ML in CBT, which is much lower than in common malignant tumors. Identified variants in known paraganglioma/pheochromocytoma-causative genes and novel genes could be associated with the pathogenesis of CBT. The obtained results expand our knowledge of the mutation process in CBT as well as the biology of tumor development.

Keywords: Carotid body tumor, Mutational load, Somatic variants, Germline variants, Exome, High-throughput sequencing

Background

Carotid body tumor (CBT) is the most frequent paraganglioma of the head and neck that arises from carotid glomus [1]. This tumor is highly vascularized and commonly involves carotid artery and cranial nerves. Surgery is a main method for CBT treatment, since radiation therapy and chemotherapy are not very effective.

Tumor development is closely associated with the accumulation of somatic mutations, which may be due to various processes such as endogenous and exogenous DNA damage, defective mechanisms of DNA replication, modification, and repair [2, 3]. These cause the changes in expression profiles of many genes, including activation of oncogenes and inactivation of tumor suppressor genes that lead to alterations in signaling pathways, cellular metabolism, and proliferation [4–14]. Distinct combinations of mutation types ("mutational signatures") depend on different mutation processes; multiple mutation processes generate jumbled composite signatures. In the study of

* Correspondence: rhizamoeba@mail.ru

¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia

Full list of author information is available at the end of the article



Alexandrov et al. (2013), more than 21 mutational signatures for 30 cancers were identified, and it has been shown that the prevalence of somatic mutations across human cancer types are different, ranging from approximately 0.04 to 20 somatic mutations per megabase (Mb) [15, 16].

In recent years, immunotherapy has been successfully used in cancer treatments. Notably, the blockade of immunosuppressive checkpoints, such as T-lymphocyte-associated antigen 4 (CTLA4), programmed cell death 1 (PD1) protein, and programmed cell death-ligand 1 (PD-L1), has demonstrated objective clinical responses in various cancers and other malignant neoplasms [17, 18]. Inhibition of both CTLA4 and PD1/PD-L1 reactivates lymphocytes against tumor-expressing neoantigens. They result from different mutations in tumor cells, and the potential of neoantigen formation is correlated with mutational load (ML) [19]. Melanoma and lung cancer, which are characterized by high MLs, were demonstrated to have clinical benefits from the immunotherapy with antibodies targeting CTLA4 and PD1 [20, 21]. Responses to immune checkpoint blockade therapy have been described in patients with colorectal cancer characterized by microsatellite instability (MSI), which is accompanied in most cases with high ML, resulting from defects in mismatch-repair pathways [22]. High ML, MSI (surrogate marker of high ML), and neoantigen production have been demonstrated to be promising markers of sensitivity to immune checkpoint blockade for several tumors [23–25]. However, these criteria also exhibited inconsistent patterns in patients with ovarian and urothelial cancer as well as those with glioblastoma; therefore, their potential use as prognostic factors requires further studies [26–28]. Additionally, intestinal microbiota, expression of PD-L1 in tumor cells, and tumor-infiltrating lymphocytes (TILs) were also determined to have a predictive role in immunotherapeutic responses [29–33].

In this study, we estimated the ML in CBT. We performed exome sequencing of tumors with matched lymph node tissues and peripheral blood derived from six patients with CBT. Additionally, a number of pathogenic/likely pathogenic somatic and germline variants were identified.

Methods

Patients and samples

Formalin-fixed paraffin-embedded (FFPE) tumor and lymph node tissues as well as peripheral blood from six patients with CBT were collected from Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation for exome sequencing. We also used a collection of 52 CBTs (exome sequencing data, available in the NCBI Sequence Read Archive [SRA] under accession number PRJNA411769) from a previous study [34]. All patients provided written informed consent. Clinicopathologic characteristics of the patients with CBT are presented in Table 1. The study was approved by the ethics committee from Vishnevsky Institute of Surgery and performed according to the Declaration of Helsinki (1964).

Exome sequencing

DNA was extracted from blood cells using a MagNA Pure Compact Nucleic Acid Isolation Kit I (Roche, Switzerland) on a MagNA Pure Compact Instrument (Roche); DNA from tumor and lymph node tissues was isolated with High Pure FFPET DNA Isolation Kit (Roche). DNA (100 ng per sample) was sheared to 150 bp using Covaris S 220 System (Thermo Fisher Scientific, USA) and was then subjected to library preparation with TruSeq Exome Library Prep Kit (Illumina, USA) according to manufacturer's instructions. The exome sequencing was performed on a NextSeq 500 System (Illumina) with paired-end reads. Read length was 76 bp for tumor and lymph node tissues and 151 bp – for blood. The obtained coverage was at least 300×. Raw sequencing data have been deposited at the NCBI SRA under accession number PRJNA476932.

Raw reads were trimmed and adapter sequences were removed with Trimmomatic [35]. We aligned 100 K randomly selected reads to bacterial genomes (NCBI, all bacterial genomes submitted up to 2014) with BWA [36] in order to evaluate contamination levels. All the samples demonstrated no greater than 0.1% bacterial DNA ratios. Next, reads were mapped to the reference human genome GRCh37.75 (Ensembl) with BWA. The derived BAM files

Table 1 Clinicopathologic characteristics of the patients with CBT

Patient	Gender	Age (years)	Family history of paragangliomas	Metastasis (lymph node/distance)	Multifocal growth	Comments
Pat100	Female	35	N/A	No	No	–
Pat101	Female	46	N/A	No	No	Tumor recurrence was diagnosed in a year after surgery*
Pat102	Female	31	N/A	No	No	–
Pat103	Male	57	N/A	No	No	–
Pat104	Female	58	N/A	No	Yes	Carotid body tumor and vagal paraganglioma were diagnosed*
Pat105	Female	67	N/A	No	No	–

* - In the study, only primary CBT from Pat101 and CBT (not vagal paraganglioma) from Pat104 were analyzed

were processed with picard-tools (reordered, supplied with group names, and duplicated reads were marked). Then, we performed base quality score recalibration (BQSR) using GATK4 (version 4.0.8.1) and dbSNP (common variants, 2015-06-05). To call somatic variants, two algorithms, VarScan [37] and Mutect2 [38], were used.

First, we applied VarScan to reveal somatic variants in paired ('tumor *versus* normal') mode. We merged BAMs for blood and lymph nodes per each patient and submit pileups (bcftools) from these BAMs to VarScan. Reads with mapping quality lower than 20 and the bases with base calling quality lower than 20 were filtered out. Only regions with 20x or higher coverage (for both tumor and 'merged' norm) were included in the analysis.

Second, we used Mutect2 (from GATK 4.0.8.1) to identify somatic variants, SNVs and indels. Before calling somatic variants, we ran Mutect2 in 'tumor-only' mode with all the 12 normal samples (lymph nodes and blood) to create a panel of norms (PoN). Next, we merged BAMs for blood and lymph nodes per each patient and used these BAMs along with PoN to call somatic variants with Mutect2 in 'tumor *versus* normal' mode. The derived VCFs were analyzed with GATK FilterMutectCalls, and only passed somatic variants were included in the further analysis. We have decided to also include in the analysis clustered events, which are filtered out with FilterMutectCalls by default. Additionally, we called variants in artificial comparisons 'lymph node (FFPE) *versus* blood', 'tumor (FFPE) *versus* blood', 'blood *versus* tumor (FFPE)', etc.

The derived list of somatic variants was annotated using Annovar [39]. We included allele population frequency databases (gnomAD, 1000 Genomes Project, Kaviar, ESP 6500, and ExAC), public variant databases (dbSNP, ClinVar, and COSMIC), phastCons containing conservation data for vertebrates, primates, and placental mammals [40], and InterPro to analyze the localizations of variants in protein domains [41]. Additionally, prediction tools such as SIFT [42], PolyPhen2 [43], MutationTaster [44], LRT [45], InterVar [46], PROVEAN [47], M-CAP [48], MetaSVM, and MetaLR [49] were used to assess the pathogenicity of the variants. Variants were considered to be likely pathogenic if they were predicted as deleterious by at least three algorithms. However, in most cases the majority of algorithms gave consistent results.

We excluded variants with population frequency greater than 1%. Worth noting, the overall number of such variants comprised only 0–7% of all the exonic somatic variants (variants in gene coding regions) passed after analysis with FilterMutectCalls. Additionally, the list of somatic variants was filtered according to the minimal read coverage threshold (min 20 reads for a merged norm and min 10 reads for the tumor sample).

Results

Mutational load in CBT

First, we should mention once more that we used blood and FFPE samples taken from tumor and lymph node tissues. The accurate detection of variants in FFPE samples is often problematic because of DNA fragmentation and the occurrence of sequence artifacts resulted from fixation of tissues in formaldehyde [50]. To evaluate the effect of FFPE artifacts on the results, we compared 'lymph node (FFPE) *versus* blood' and revealed a great number of variants (hundreds) that was almost equal to the number of somatic variants found in the 'tumor (FFPE) *versus* blood' comparison (including variant with low alternative allele (AF) frequency). In contrast, when comparing either 'blood *versus* tumor (FFPE)' or 'blood *versus* lymph node (FFPE)', a very moderate number of variants (dozens) was revealed. This pronounced trend was observed for all six patients. It suggests that the most of variants identified in FFPE samples may be formalin-induced DNA artifacts. However, there were a variety of SNVs (A > T, C > A, G > A, A > G, etc.); and only a moderate bias towards the typical FFPE-induced transition, C > T, was observed. About 25% of all SNVs were represented with C > * substitutions, and approximately 40–60% of them were C > T transitions (10–15% of all SNVs).

When 'tumor *versus* lymph node' was compared, e.g. two FFPE samples, we derived about 1.3–1.5-fold lower amount of somatic variants relatively 'tumor (FFPE) *versus* blood' comparison, because read coverage for FFPE lymph node samples was 2-fold greater (on the average) than the coverage for blood samples. Additionally, this may suggest slight co-occurrence of formalin-induced variants, which are partially 'subtracted' when comparing two FFPE samples.

Among Mutect results, there are several false-positive somatic variants with low AF (5–10%) but high coverage (10–30 reads for alternative allele; the total coverage was at least 300x) observed in tumor, lymph node and blood samples. These variants obviously are neither germinal (too low AF), nor somatic (they are present in norms and tumors), nor FFPE artifacts (they are present in blood samples). Remarkably, 20–40% of reads that support the alternative alleles (AA) have low base quality score (before and after BQSR). This strongly suggests the presence of context-dependent sequencing errors that are not eliminated with BQSR procedure and are not sufficiently addressed by a variant calling algorithm (either by Mutect2 or VarScan). Indeed, we observed many of these variants in polyN-tracts (especially polyG), e.g. GGGT>GGGG, CCCC>CCCCC, susceptible for Illumina NextSeq-specific sequencing errors.

To eliminate such artifacts, we applied three approaches. First, we used strand bias filter (StrandOddsRatio annotation provided with GATK Mutect2 or

HaplotypeCaller). The presence of reads bearing a variant on only one strand indicates a false-positive. However, this option alone does not allow eliminating artifacts when they come from both strands (e.g. GGGTGGG>GGGG GGG). Second, we filtered out variants with abnormal distribution of base quality scores across the reads (GATK BaseQualityRankSumTest annotation). In other words, we excluded variants that were abnormally supported with too many reads with low base quality score at a current position, even if there were also many reads with high base quality score at this position. Third, we manually filtered out variants that were observed after/before four or more identical nucleotides (mainly, C/G). In most cases, when we are calling somatic variants on various types of cancer, their number may significantly exceed the ratio of the described false-positives. In contrast, when analyzing tumors with a low ML, this issue becomes especially important.

Totally, we found 70–130 potentially somatic variants in each patient comparing ‘tumor (FFPE) versus matched lymph node (FFPE) and blood + PoN’. Most of these variants have low AF (Fig. 1). Only 5–20% of variants passed threshold of AF > 15%.

When calculating the total mutational load, it is incorrect to consider a somatic variant that is observed only in a fraction of the tumor cells (such variants are featured with low AF) as a “whole one”. Otherwise, the higher sequencing coverage, the more we can find variants with very low AF values, and the higher calculated mutational load will be. The weight of such somatic variants should be adjusted for AF. Hence, we re-estimated the number of variants “in terms of heterozygous ones” as the sum of all AFs multiplied by 2 (Fig. 2, weighted somatic variants count).

As can be seen from Fig. 2, two of six patients demonstrated an elevated number of potentially somatic variants. These differences are much more pronounced when an elevated AF threshold is set (AF > 0.2, e.g. > 40% of cells are for heterozygous variant). Nevertheless, it is worth noting that the initial number of tumor cells in the tumor samples did not reach 100%, and it could vary from 70 to 90%.

To finally assess the weighted mutational load (wML) in CBT, we should reasonably pick up the AF threshold. For most patients, a significant reduce in the number of variants occurs in the region of AF = 10–15%. Therefore, wML may be estimated as 4–12 variants per genome or 0.09–0.28 variants per megabase taking into account the fact that we have used TruSeq Exome Library Prep Kit (the total length of target regions is 43 Mb). It should be noted that the evaluation of ML in such cases is close to the limit of sensitivity/specificity of the method.

Considering the structure of a list of potential somatic variants, 25–73 variants (SNVs and indels) are located either in coding regions or splice sites and are supported with at least three reads corresponding to the alternative allele (Fig. 3). Remarkably, 2–18% of these ones have already been annotated in COSMIC databases, and only 2–7% of variants have maximal population frequency (across multiple databases) greater than 1% (before filtering). The total number of all variants (including UTRs, intronic, and intergenic) was 2-fold greater on the average than the number of variants in the coding regions.

Pathogenic and likely pathogenic somatic variants

Across six patients, we revealed 50 likely pathogenic variants, and among them several potential driver variants were observed (Fig. 4). In two patients (Pat103

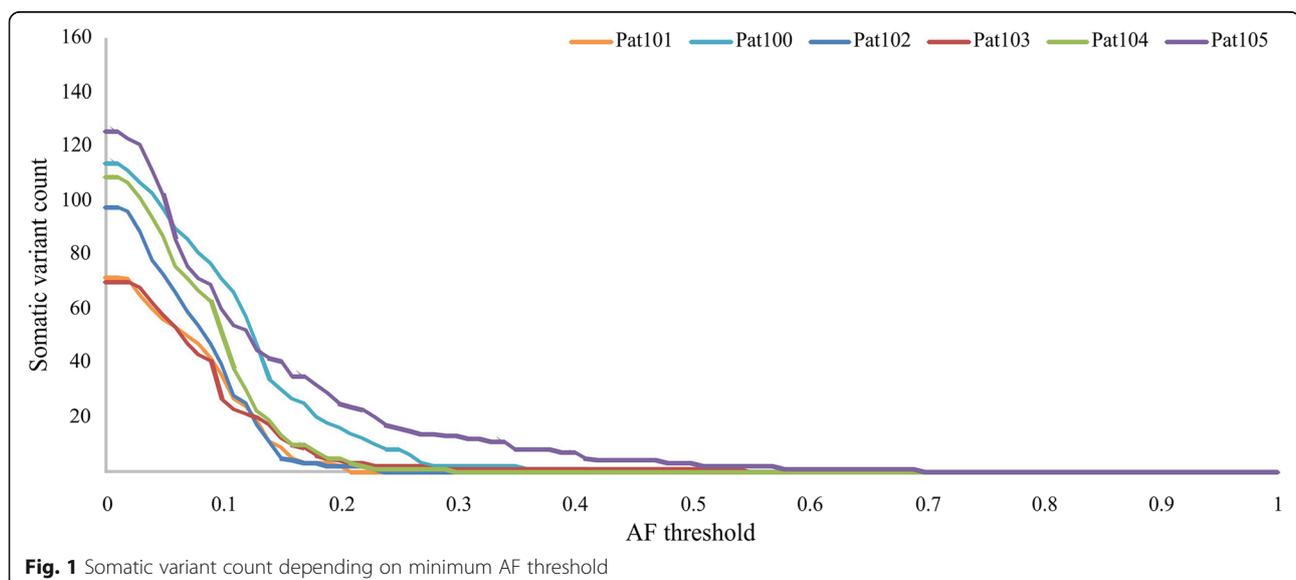
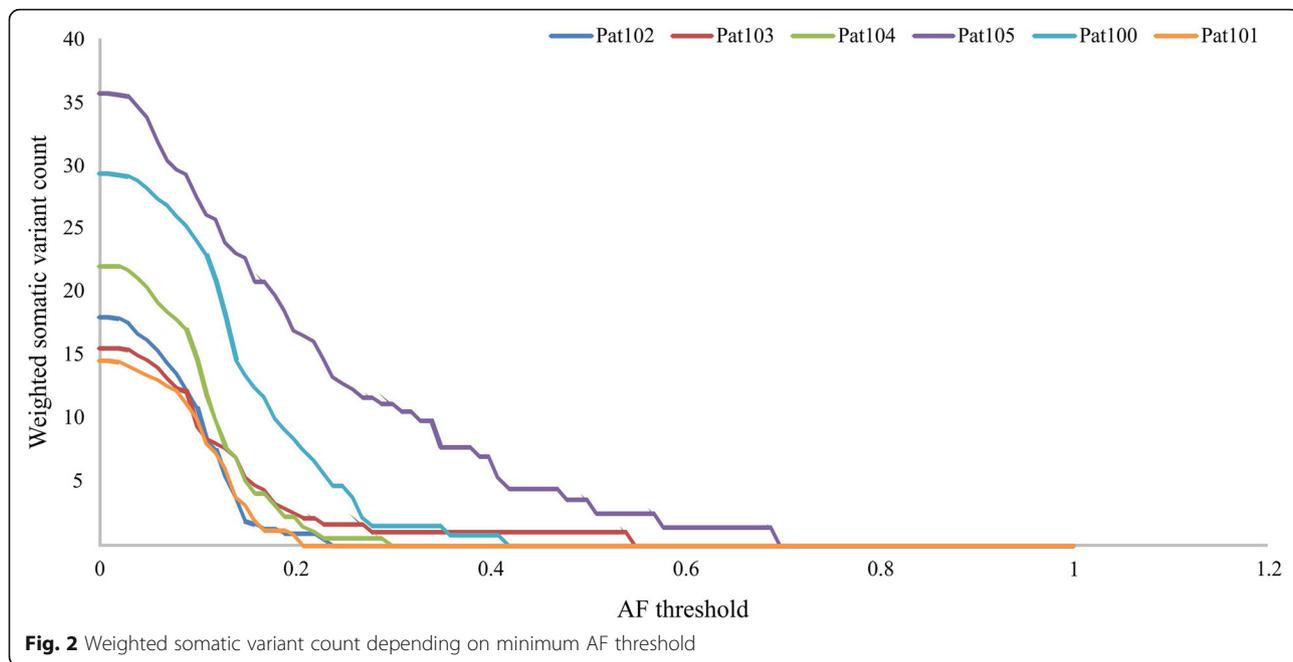


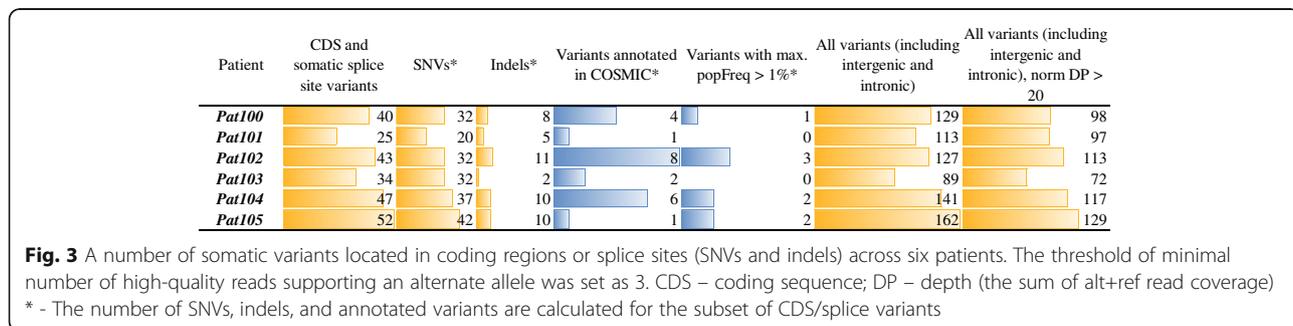
Fig. 1 Somatic variant count depending on minimum AF threshold



and Pat104) we found two co-occurred variants in *TP53* gene, NM_000546.5: c.842A > T, p.Asp281Val (chr17: 7,577,096, rs587781525) and NM_000546.5: c.A170A > G, p. Asp57Gly (chr17: 7,579,517). The first variant was described in dbSNP as both germline and somatic one, and has a pathogenic clinical significance according to the ClinVar database. The germline variant was associated with hereditary cancer-predisposing syndrome; the somatic variant has been found in many neoplasms, including neuroendocrine tumors (neuroblastoma and glioblastoma). Variant NM_000546.5: c.A170A > G, p. Asp57Gly (chr17: 7579517) was not previously described in databases and in the literature. In our study, this variant was found with high AF (0.55 and 0.23 for Pat103 and Pat104, respectively) and was classified to be likely pathogenic by most prediction tools. These two variants can be associated with biallelic inactivation of *TP53* gene and are involved in the pathogenesis of CBT.

Potential driver variant NM_003002: c.A1A > T, p. Met1Leu (chr11: 111,957,632, rs104894307) in *SDHD* gene was revealed for Pat103. It was annotated in dbSNP and ClinVar databases as germline pathogenic variant associated with hereditary cancer-predisposing syndrome and pheochromocytomas/paragangliomas. This somatic variant was also described as pathogenic in sporadic paraganglioma [51].

We also observed somatic variants in several known cancer-associated genes (for example, *JAG1*, *PRDM2*, *PRDM8*, *SETD2*, *ASPM*, *ZIC*, *GRIK1*, etc.), which may be important for cell growth and proliferation. They did not overlap between the patients. These variants have low AF values (5–10%) occurring in low fraction of cells (10–20%), but they may represent driver events. This demonstrates genetic intra-tumor heterogeneity of CBT. Apparently, they occurred after tumor onset and initial progression. Having received these variants, such cells could gain an advantage in their clonal expansion.



Chr	Position	Ref.	Alt.	Gene	dbSNP	Tumor			Merged norm			Effect on protein function											ClinVar	COSMIC		
						Alt. cov	Ref. cov	Alt. ratio	Alt. cov	Ref. cov	Alt. ratio	max Pop Freq	SIFT	Polyphen2 HDIV	Polyphen2 HVAR	PROVEAN	MetaSVM	MetaLR	M-CAP	phastCons 100-way vertebrate*	phastCons 20-way mammalian*					
Pat100																										
10	118687318	G	A	<i>SHTN1</i>		3	21	0.13	1	40	0.02	0	D	D	D	D	T	D	D			1	0.785			
14	54417376	C	G	<i>BMP4</i>		5	150	0.03	0	225	0.00	0	D	D	D	D	D	D	D			1	1			
2	131520986	T	G	<i>AMER3</i>		8	46	0.15	1	92	0.01	0	D	D	D	N	T	T	T			0.312	0.289			
20	36769704	G	A	<i>TGM2</i>	rs368689999	5	84	0.06	0	145	0.00	0	D	D	D	D	D	D	D			1	0.556			
3	50339586	C	T	<i>HYAL1</i>	rs104893743	5	180	0.03	0	162	0.00	0	D	D	D	D	D	T	D			1	0.305	<< Deficiency of hyaluronoglu-	<< COSM1229101; large intestine	
4	38799482	G	A	<i>TLL1</i>	rs143619350	4	85	0.04	0	111	0.00	0	D	P	B	D	T	T	T			1	0.313	<< COSM1055140; cosaminidase	<< COSM1055140; large intestine, endometrium	
4	166021831	A	T	<i>TMEM192</i>		3	77	0.04	0	97	0.00	0	D	D	P	D	T	T	T			1	0.243			
6	39283003	A	C	<i>KCNK16</i>		18	135	0.12	0	245	0.00	0	D	P	B	N	T	T	T			0.088	0.107			
Pat101																										
1	114945479	C	G	<i>TRIM33</i>		3	64	0.04	0	115	0.00	0	T	D	D	D	D	D	D			1	1			
11	116744223	G	A	<i>SIK3</i>		4	34	0.11	0	108	0.00	0	D	D	P	N	T	T	D			1	1			
12	110656018	G	A	<i>IFT81</i>	rs756693673	4	17	0.19	0	36	0.00	0	D	D	P	N	T	T	T			1	0.661			
19	35993779	G	A	<i>DMKN</i>		3	20	0.13	0	76	0.00	0	D	D	P	D	T	T	D			0.362	0.17			
2	25964912	G	T	<i>ASXL2</i>		3	15	0.17	0	35	0.00	0	D	D	D	N	D	D	D			1	0.735			
Pat102																										
21	31062171	C	T	<i>GRIK1</i>		8	140	0.05	0	250	0.00	0	D	P	P	N	D	D	D			1	0.785			
3	52559100	C	T	<i>NTSDC2</i>		4	28	0.13	0	70	0.00	0	D	D	D	D	T	T	D			1	0.482			
Pat103																										
1	152185767	C	T	<i>HRNR</i>		3	43	0.07	0	107	0.00	0	D	D	P	N	T	T	T			0.088	0.018			<< COSM2156985; central nervous
11	111957632	A	T	<i>SDHD</i>	rs104894307	20	189	0.10	0	337	0.00	0	D	D	D	N	D	D	D			0.538	0.211	<< Hereditary cancer-predisposing syndrome		
12	103237494	A	C	<i>PAH</i>		6	101	0.06	0	255	0.00	0	D	D	D	D	D	D	D			1	0.252			
13	49075950	G	A	<i>RCBTB2</i>	rs376690025	4	71	0.05	0	127	0.00	0	D	B	B	D	T	T	D			1	0.343			
17	7577096	T	A	<i>TP53</i>	rs587781525	52	219	0.19	0	333	0.00	0	D	D	D	D	D	D	D			1	0.845	<< Neuroblastoma; Chronic lymphocytic leukemia; Renal cell carcinoma, papillary; Pancreatic adenocarcinoma; Skin squamous cell carcinoma; ... (total 19 tumors)	<< COSM562341, COSM45729, COSM1649342; lung, large intestine, breast	
17	7579517	T	C	<i>TP53</i>		71	59	0.55	0	123	0.00	0	D	P	P	N	D	D	D			0.088	0.212			
17	39881161	G	A	<i>HAP1</i>		4	59	0.06	0	168	0.00	0	D	P	P	D	T	T	T			0.362	0.083			
2	128459433	C	A	<i>SFT2D3</i>		5	158	0.03	0	283	0.00	0	D	P	D	N	T	T	D			1	0.785			
3	8667975	G	A	<i>SFUH2</i>		6	32	0.16	0	79	0.00	0	D	D	P	D	T	T	T			1	0.345			
Pat104																										
1	200956292	G	A	<i>KIF21B</i>		4	59	0.06	0	154	0.00	0	T	D	D	N	T	T	D			1	0.354			<< COSM902032; endometrium
11	94731996	A	T	<i>KDM4D</i>		2	1	0.67	0	22	0.00	0	D	P	B	D	T	T	T			0.192	0.048			
13	74420426	A	T	<i>KLF12</i>		3	137	0.02	0	252	0.00	0	T	P	P	N	T	T	T			0.487	0.434			
17	7577096	T	A	<i>TP53</i>	rs587781525	22	241	0.08	1	454	0.00	0	D	D	D	D	D	D	D			1	0.845	<< Neuroblastoma; Chronic lymphocytic leukemia; Renal cell carcinoma, papillary; Pancreatic adenocarcinoma; Skin squamous cell carcinoma; ... (total 19 tumors)	<< COSM562341, COSM45729, COSM1649342; lung, large intestine, breast	
17	7579517	T	C	<i>TP53</i>		16	55	0.23	1	169	0.01	0	D	P	P	N	D	D	D			0.088	0.212			
2	46605215	A	T	<i>EPAS1</i>		4	69	0.05	1	165	0.01	0	T	D	P	N	T	T	D			0.538	0.845			
2	230652330	G	C	<i>TRIP12</i>		10	72	0.12	0	134	0.00	0	D	P	P	D	T	T	D			1	1			
20	10639239	A	G	<i>JAG1</i>		11	171	0.06	0	261	0.00	0	D	P	D	D	D	D	D			1	0.845			
20	31017790	C	T	<i>ASXL1</i>	rs564841799	4	71	0.05	0	176	0.00	0	D	D	D	D	T	T	T			1	0.785			
6	70410745	T	C	<i>LMBRD1</i>		8	46	0.15	0	176	0.00	0	D	D	D	D	T	T	D			1	1			
6	160232781	T	C	<i>PNLDC1</i>		13	96	0.12	0	173	0.00	0	D	D	P	D	T	T	T			1	0.488			
Pat105																										
1	14106977	A	G	<i>PRDM2</i>	rs756985448	50	22	0.69	0	128	0.00	0	D	D	D	D	T	T	T			1	0.845			
1	157491079	G	A	<i>FCRL5</i>	rs142943534	4	48	0.08	0	64	0.00	0	T	P	B	D	T	T	D			1	0.222			
1	167849838	A	T	<i>ADCY10</i>		10	283	0.03	0	298	0.00	0	D	D	D	D	T	T	D			1	1			
1	197097759	A	T	<i>ASP</i>		64	134	0.32	0	194	0.00	0	D	D	D	D	T	T	D			1	1			
1	201832596	C	T	<i>IPO9</i>	rs763876238	10	28	0.26	0	53	0.00	0	D	D	D	D	D	D	D			1	0.785			
14	102469248	A	G	<i>DYNC1HI</i>		38	60	0.39	0	193	0.00	0	D	D	D	D	T	T	D			1	1			
2	43986107	G	A	<i>PLEKHH2</i>	rs372835795	8	99	0.07	0	132	0.00	0	D	D	D	D	D	D	D			1	0.39			
20	31473396	G	A	<i>EFCAB8</i>		3	66	0.04	0	116	0.00	0	D	D	D	D						0.593	0.488			
20	43108887	G	A	<i>TTPAL</i>	rs368005756	6	118	0.05	1	177	0.01	0	D	D	D	D	D	D	D			1	1			
3	47061257	C	T	<i>SETD2</i>		3	56	0.05	0	112	0.00	0	T	B	B	N	D	D	D			1	1			
3	47061259	G	T	<i>SETD2</i>		3	55	0.05	0	111	0.00	0	D	B	D	D	D	D	D			1	1			
3	142542410	C	T	<i>PCOLCE2</i>		34	64	0.35	0	117	0.00	0	D	D	D	D	T	T	D			1	0.525			COSM1693554; skin
4	94376840	A	C	<i>GRID2</i>		4	27	0.13	1	79	0.01	0	D	D	D	D	T	T	D			1	1			
9	74970993	A	C	<i>ZFAND5</i>		39	58	0.40	0	146	0.00	0	D	D	D	D	T	T	D			1	1			
9	95263292	C	T	<i>ECM2</i>	rs781440673	14	63	0.18	0	107	0.00	0	D	D	D	N	T	T	T			1	0.521			

Fig. 4 Pathogenic/likely pathogenic somatic variants in six samples of carotid body tumor

Pathogenic and likely pathogenic germline variants in causative genes

We analyzed germline variants in 42 genes (*VHL*, *SDHA*, *SDHB*, *SDHC*, *SDHD*, *NF1*, *RET*, *HRAS*, *KRAS*, *EPAS1* (*HIF2A*), *ATRX*, *CSDE1*, *BRAF*, *FGFR1*, *FGFR2*, *FGFR3*, *FGFR4*, *FGFRL1*, *SETD2*, *ARNT*, *TP53*, *TP53BP1*, *TP53BP2*, *TP53I13*, *KMT2D*, *BAP1*, *IDH1*, *IDH2*, *SDHAF1*, *SDHAP2*,

FH, *EGLN1*, *MDH2*, *TMEM127*, *MAX*, *KIF1B*, *MEN1*, *GNDF*, *GNAS*, *CDKN2A*, *BRCA1*, and *BRCA2*) reported previously to be involved in the development of paragangliomas/pheochromocytomas [34, 52]. Three pathogenic and two likely pathogenic germline variants were found across six patients with CBT according to the predicted algorithms and public databases (Table 2, Fig. 5). These

variants were characterized by high conservation scores (PhastCons) and had an allele frequency less than 0.01% in 1000 Genomes Project, ESP 6500, and ExAC databases.

Discussion

The frequency of somatic mutations and neoantigen production correlates with responses to immunotherapy. In high mutated cancer, the ML was considered as a prognostic factor of prolonged survival with immune checkpoint inhibitors [20, 21]. However, this association has not been observed for tumors with low ML. In the present work, we estimated the ML in CBT. CBT belongs to rare neoplasms, and it is difficult to collect a representative set of CBT samples. Moreover, matched morphological normal tissues (“conventional norm”) are unavailable due to the tumor localization. We collected and performed exome sequencing of tumor samples with matched lymph node tissues and peripheral blood from six patients with CBT. This revealed actual somatic variants and allowed estimating the ML in the six samples studied. CBT carried low ML (0.09–0.28/Mb) compared to other common cancers [15]. Even though high ML is a factor for immunotherapy, it also generally indicates high aggressiveness of the tumors and correlates with increased genetic instability and poor prognosis [53–56]. CBT is usually a slow-growing tumor, and in 10% of cases, it can become aggressive and metastatic [57]. Thus, low ML in CBT is possibly associated with its non-aggressive behavior and probably indicates inefficiency of immune checkpoint blockade therapy.

Paragangliomas/pheochromocytomas can be caused by germline and somatic variants in at least 42 known genes [52]. Potential driver somatic variants were found in *SDHD* (Pat103) and *TP53* (Pat103 and Pat104) genes. Likely pathogenic variants were revealed in *EPAS1* (*HIF2A*) (Pat104) and *SETD2* (Pat105) genes, and no somatic variants were observed in other known genes. However, several pathogenic and likely pathogenic germline variants in *SDHB*, *SDHC*, *SDHD*, and *RET* genes were detected. A majority of them were identified in *SDHx* genes that are often associated with hereditary pheochromocytomas/paragangliomas [58]. Two variants were observed in *SDHB* gene in different patients (Pat101 and Pat102): a novel missense variant NM_003000.2: c.463C > A, p.Pro155Thr (chr1: 17,354,321) and a nonsense high-impact variant NM_003000.2: c.136C > T, p.Arg46* (chr1: 17,371,320, rs74315370). The latter was described in the dbSNP and ClinVar databases as a pathogenic germline variant associated with hereditary cancer predisposition syndrome, paragangliomas/pheochromocytomas, and gastrointestinal stromal tumor [59–62]. Notably, this variant has been reported in patients with aggressive extra-adrenal paraganglioma in the chest and CBT and has been considered as a high-risk factor for malignancy or recurrence of

paragangliomas/pheochromocytomas [63–66]. Indeed, the patient tested (Pat101) was characterized by the tumor recurrence that is one of the features indicating aggressive phenotype of CBT. The variant NM_003001.3: c.149G > A, p.Arg50His (chr1: 161,298,257, rs769177037) in *SDHC* was also found in Pat102. It was described in dbSNP as a variant of uncertain clinical significance. Germline variants in *SDHC* are more rarely associated with the development of paragangliomas/pheochromocytomas than variants in *SDHB* or *SDHD* [58]. It should be noted that in Pat102 we observed two somatic likely pathogenic variants in *GRIK1* and *NT5DC2* genes. Therefore, according to our previous data, the formation of CBT can be probably caused by the cumulative effect of several highly or not highly pathogenic variants [34]. In this particular case, it seems that the main driver is the pathogenic germline variant in *SDHB* gene.

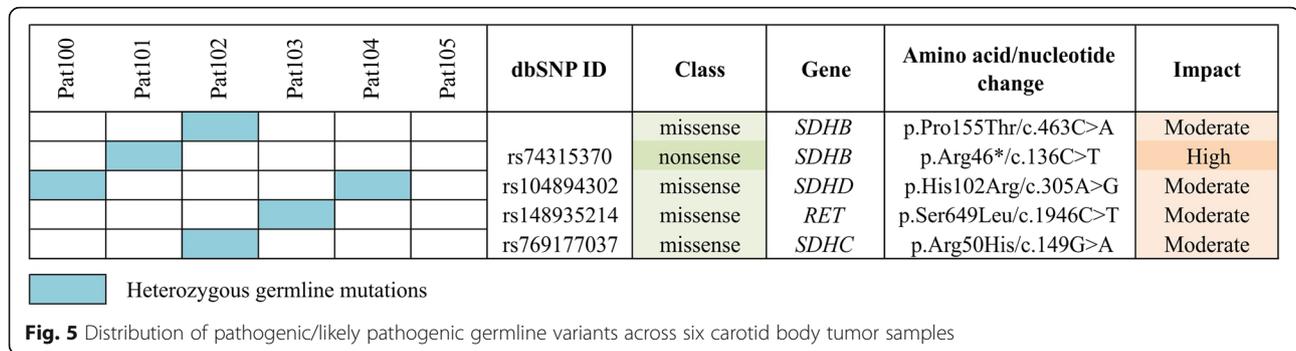
The germline variant NM_003002.3: c.305A > G, p.His102Arg (chr11: 111,959,726, rs104894302) in *SDHD* was identified in two patients – Pat100 and Pat104. This variant is found in dbSNP and ClinVar databases as a pathogenic germline variant associated with hereditary cancer-predisposing syndrome, paragangliomas/pheochromocytomas, gastric stromal sarcoma, and Cowden syndrome 3 [67, 68]. This variant has been detected in malignant CBT [69]. Data on the aggressive behavior of the tumor in tested patients have not been reported; one patient (Pat104) was characterized by multiple tumors (vagal paraganglioma and CBT) with multifocal growth. In this patient (Pat104), we also found pathogenic and likely pathogenic somatic variants in *TP53* gene.

One patient (Pat103) carried the germline variant NM_020975.4: c.1946C > T, p.Ser649Leu (chr10: 43,609,994, rs148935214) in the proto-oncogene *RET*. It was deposited to dbSNP from the gnomAD database as a germline variant. In ClinVar, another allele was reported with conflicting interpretations of pathogenicity found in hereditary cancer-predisposing syndrome (uncertain significance) and multiple endocrine neoplasia (MEN) type 2 (uncertain significance/likely benign) characterized by medullary thyroid carcinoma, pheochromocytomas, and hyperparathyroidism [70–72]. In this patient, we also identified pathogenic and likely pathogenic somatic variants in *TP53* gene (NM_000546.5: c.842A > T, p.Asp281Val (chr17: 7,577,096, rs587781525) and NM_000546.5: c.A170A > G, p. Asp57Gly (chr17: 7,579,517) that are the same in Pat104. These variants can be potential driver ones.

Interestingly, we did not reveal any pathogenic germline variants in known paraganglioma/pheochromocytoma-causative genes in the patient Pat105, which is characterized with the greatest number of somatic variants and the highest ML. However, this patient was characterized by at least two somatic variants in CBT in one known CBT-causative gene – *SETD2*. This gene

Table 2 Pathogenic/likely pathogenic germline variants in six samples of carotid body tumor

Gene	dbSNP	GenBank	Coordinate	Nucleotide change	Amino acid change	Genotype	Predictions		MutationTaster	LRT	Clinical significance (dbSNP/ClinVar)
							SIFT	PolyPhen2			
<i>SDHB</i>	-	NM_003000.2	chr1: 17,354,321	c.463C > A	p.Pro155Thr	Het	Deleterious (0)	Probably damaging (1)	Disease-causing (p-values = 1)	Neutral (0)	N/A
<i>SDHB</i>	rs74315370	NM_003000.2	chr1: 17,371,320	c.136C > T	p.Arg46*	Het	Deleterious (0)	Benign (0)	Disease-causing (p-values = 1)	Neutral (0)	Pathogenic
<i>SDHD</i>	rs104894302	NM_003002.3	chr11: 111,959,726	c.305A > G	p.His102Arg	Het	Deleterious (0)	Probably damaging (0.99)	Disease-causing (p-values = 1)	Neutral (0)	Pathogenic
<i>RET</i>	rs148935214	NM_020975.4	chr1043609,994	c.1946C > T	p.Ser649Leu	Het	Deleterious (0)	Probably damaging (1)	Disease-causing (p-values = 1)	Neutral (0)	Pathogenic
<i>SDHC</i>	rs769177037	NM_003001.3	chr1: 161,298,257	c.149G > A	p.Arg50His	Het	Deleterious (0)	Probably damaging (1)	Not disease-causing (p-values = 0)	Neutral (0)	Uncertain significance



encodes for histone methyltransferase, an epigenetic modifier with tumor suppressor functionality [73]. Mutations in *SETD2* are found in many tumors, including neoplasms of the central nervous system [74].

In the previous work, we performed exome sequencing of 52 archival FFPE samples of CBT [34]. Peripheral blood or other normal tissues were unavailable; therefore, germline variants were excluded with strong filtering using the 1000 Genomes Project and ExAC databases. We estimated the ML as the number of potentially somatic deleterious variants per megabase of coding regions. However, we derived an obviously elevated ML: the average ML was 6–8 variants per Mb. Therefore, such approach does not allow efficient elimination of germline variants. Indeed, this method excludes 96–98% or more germline variants, but the remaining 2–3% of the germline variants may significantly outnumber somatic ones. In this study, we filtered the pool of somatic variant candidates from a previous work (52 patients) using exome sequencing data on blood and lymph nodes derived in the present work (6 patients). This resulted in at least 2-fold reduction of the estimated ML, but this value was still excessively high. Thus, the use of matched normal tissues is necessary to be able to accurately estimate ML.

Recently, Roche (Switzerland) announced the AVE-NIO ctDNA Analysis Kits for personalized oncology assays [75, 76]. The AVENIO ctDNA Surveillance Kit targets frequently mutated regions across 197 genes and has been optimized for monitoring of ML in lung and colorectal cancers. This kit contains the main genes that are associated with lung, colorectal, breast, gastric, prostate, ovarian, thyroid, and pancreatic cancers, as well as glioma and melanoma according to the U.S. National Comprehensive Cancer Network (NCCN) Guidelines (<https://www.nccn.org/>). The kit did not include genes that have been shown to be involved in the pathogenesis of paragangliomas and pheochromocytomas, except *TP53*, *BRCA1*, and *BRCA2*. A panel of genes accurately reflecting the ML in CBT is also unknown. Moreover, the kit and the appropriate analysis software focus on quantitating ML basing on ctDNA

sequencing, and this approach is more acceptable for malignant tumors with a high frequency of metastases, while CBT is primarily a slow-growing tumor with indeterminate potential of malignancy. Thus, whole exome sequencing, which was used in the study, is currently the only method for estimating the ML in CBT.

Conclusion

The ML varied in the range of 0.09–0.28/Mb in the analyzed cohort of patients with CBT (six individuals). Several pathogenic/likely pathogenic somatic and germline allelic variants in both known paraganglioma/pheochromocytoma-causative genes and novel ones were identified. These results improve the understanding of CBT pathogenesis.

Abbreviations

AA: Alternative allele; AF: Allele frequency; CBT: Carotid body tumor; ML: Mutational load

Acknowledgements

Authors thank Vishnevsky Institute of Surgery for tissue samples and National Medical Research Radiological Center for assistance in the data analysis. This work was performed using the equipment of EIMB RAS “Genome” center (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

Funding

This work and publication costs were financially supported by the Russian Foundation for Basic Research (grant 16–04-01521a), the Russian Science Foundation (grant 17–75-20105), and ICGEB project CRP/RUS15–01. Collection of six CBT samples with matched lymph node tissues and blood, exome sequencing of these samples, and bioinformatic analysis of the obtained data and a part of publication costs were financially supported by the Russian Foundation for Basic Research (grant 16–04-01521a) and ICGEB project CRP/RUS15–01. Reanalysis of exome sequencing data from 52 CBTs and a part of publication costs were financially supported by the Russian Science Foundation, grant 17–75-20105.

Availability of data and materials

All data generated or analyzed in this study are included in the published article. The sequence data are available in the NCBI Sequence Read Archive (SRA) under the accession numbers PRJNA411769 and PRJNA476932.

About this supplement

This article has been published as part of BMC Medical Genomics Volume 12 Supplement 2, 2019: Selected articles from BGR/SB-2018: medical genomics. The full contents of the supplement are available online at <https://bmcmmedgenomics.biomedcentral.com/articles/supplements/volume-12-supplement-2>.

Authors' contributions

AVK, AVS, and ENL conceived and designed the work; MSF, AVK, EAP, SLK, DVK, ARZ, ALG, NVM, AVP, AAK, and AVS performed the experiments; ENL, GSK, APP, MVK, VSP, and AAD analyzed the data; AVS, AVK, BYA, ADK, and ENL wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The study was approved by the ethics committee from Vishnevsky Institute of Surgery and performed according to the Declaration of Helsinki (1964).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia. ²Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation, Moscow, Russia. ³National Medical Research Radiological Center, Ministry of Health of the Russian Federation, Moscow, Russia.

Published: 13 March 2019

References

- Pellitteri PK, Rinaldo A, Myssiorek D, Gary Jackson C, Bradley PJ, Devaney KO, Shaha AR, Nettierville JL, Manni JJ, Ferlito A. Paragangliomas of the head and neck. *Oral Oncol.* 2004;40(6):563–75.
- Nakagawa H, Fujita M. Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer Sci.* 2018;109(3):513–22.
- Tchurikov NA, Kretova OV, Fedoseeva DM, Chechetkin VR, Gorbacheva MA, Snezhkina AV, Alembekov IR, Kravatskaya GI, Kravatsky YV. Genome-wide mapping of hot spots of DNA double-strand breaks in human cells as a tool for epigenetic studies and cancer genomics. *Genomics data.* 2015;5:89–93.
- Krasnov GS, Dmitriev AA, Snezhkina AV, Kudryavtseva AV. Deregulation of glycolysis in cancer: glyceraldehyde-3-phosphate dehydrogenase as a therapeutic target. *Expert Opin Ther Targets.* 2013;17(6):681–93.
- Pudova EA, Kudryavtseva AV, Fedorova MS, Zaretsky AR, Shcherbo DS, Lukyanova EN, Popov AY, Sadritdinova AF, Abramov IS, Kharitonov SL, et al. HK3 overexpression associated with epithelial-mesenchymal transition in colorectal cancer. *BMC Genomics.* 2018;19(Suppl 3):113.
- Fedorova MS, Snezhkina AV, Pudova EA, Abramov IS, Lipatova AV, Kharitonov SL, Sadritdinova AF, Nyushko KM, Klimina KM, Belyakov MM, et al. Upregulation of NETO2 gene in colorectal cancer. *BMC Genet.* 2017; 18(Suppl 1):117.
- Kudryavtseva AV, Fedorova MS, Zhavoronkov A, Moskalev AA, Zasedatelev AS, Dmitriev AA, Sadritdinova AF, Karpova IY, Nyushko KM, Kalinin DV, et al. Effect of lentivirus-mediated shRNA inactivation of HK1, HK2, and HK3 genes in colorectal cancer and melanoma cells. *BMC Genet.* 2016;17(Suppl 3):156.
- Snezhkina AV, Krasnov GS, Zaretsky AR, Zhavoronkov A, Nyushko KM, Moskalev AA, Karpova IY, Afremova AI, Lipatova AV, Kochetkov DV, et al. Differential expression of alternatively spliced transcripts related to energy metabolism in colorectal cancer. *BMC Genomics.* 2016;17(Suppl 14):1011.
- Snezhkina AV, Krasnov GS, Lipatova AV, Sadritdinova AF, Kardymon OL, Fedorova MS, Melnikova NV, Stepanov OA, Zaretsky AR, Kaprin AD, et al. The dysregulation of polyamine metabolism in colorectal Cancer is associated with overexpression of c-Myc and C/EBPbeta rather than Enterotoxigenic Bacteroides fragilis infection. *Oxidative Med Cell Longev.* 2016;2016:2353560.
- Kudryavtseva AV, Krasnov GS, Dmitriev AA, Alekseev BY, Kardymon OL, Sadritdinova AF, Fedorova MS, Pokrovsky AV, Melnikova NV, Kaprin AD, et al. Mitochondrial dysfunction and oxidative stress in aging and cancer. *Oncotarget.* 2016;7(29):44879–905.
- Krasnov GS, Dmitriev AA, Sadritdinova AF, Fedorova MS, Snezhkina AV, Melnikova NV, Poteryakhina AV, Nyushko KM, Belyakov MM, Kaprin AD, et al. evaluation of gene expression of hexokinases in colorectal Cancer with the use of bioinformatics methods. *Biofizika.* 2015;60(6):1050–6.
- Snezhkina AV, Krasnov GS, Zhikrivetskaya SO, Karpova IY, Fedorova MS, Nyushko KM, Belyakov MM, Gnuchev NV, Sidorov DV, Alekseev BY, et al. Overexpression of microRNAs miR-9, 98, and 199 Correlates with the Downregulation of HK2 Expression in Colorectal Cancer. *Mol Biol+.* 2018; 52(2):190–9.
- Fedorova MS, Kudryavtseva AV, Lakunina VA, Snezhkina AV, Volchenko NN, Slavnova EN, Danilova TV, Sadritdinova AF, Melnikova NV, Belova AA, et al. Downregulation of OGDHL expression is associated with promoter hypermethylation in colorectal cancer. *Mol Biol+.* 2015;49(4):608–17.
- Oparina NY, Snezhkina AV, Sadritdinova AF, Veselovskii VA, Dmitriev AA, Senchenko VN, Melnikova NV, Speranskaya AS, Darii MV, Stepanov OA, et al. Differential expression of genes that encode glycolysis enzymes in kidney and lung cancer in humans. *Russ J Genet+.* 2013;49(7):707–16.
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415–21.
- Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013; 499(7457):214–8.
- Liontos M, Anastasiou I, Bamias A, Dimopoulos MA. DNA damage, tumor mutational load and their impact on immune responses against cancer. *Annals of translational medicine.* 2016;4(14):264.
- Gong J, Le TQ, Massarelli E, Hendifar AE, Tuli R. Radiation therapy and PD-1/PD-L1 blockade: the clinical development of an evolving anticancer combination. *J Immunother Cancer.* 2018;6(1):46.
- Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. *Science.* 2015;348(6230):69–74.
- Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science.* 2015;350(6257):207–11.
- Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science.* 2015;348(6230):124–8.
- Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med.* 2015;372(26):2509–20.
- Saleh K, Eid R, Haddad FG, Khalife-Saleh N, Kourie HR. New developments in the management of head and neck cancer - impact of pembrolizumab. *Ther Clin Risk Manag.* 2018;14:295–303.
- Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science.* 2017;357(6349):409–13.
- Davarpanah NN, Yuno A, Trepel JB, Apolo AB. Immunotherapy: a new treatment paradigm in bladder cancer. *Curr Opin Oncol.* 2017.
- Matsushita H, Hasegawa K, Oda K, Yamamoto S, Nishijima A, Imai Y, Asada K, Ikeda Y, Karasaki T, Fujiwara K, et al. The frequency of neoantigens per somatic mutation rather than overall mutational load or number of predicted neoantigens per se is a prognostic factor in ovarian clear cell carcinoma. *Oncoimmunology.* 2017;6(8):e1338996.
- Snyder A, Nathanson T, Funt SA, Ahuja A, Buros Novik J, Hellmann MD, Chang E, Aksoy BA, Al-Ahmadie H, Yusko E, et al. Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: an exploratory multi-omic analysis. *PLoS Med.* 2017;14(5):e1002309.
- Garg AD, Vandenberk L, Van Woensel M, Belmans J, Schaaf M, Boon L, De Vleeschouwer S, Agostinis P. Preclinical efficacy of immune-checkpoint monotherapy does not recapitulate corresponding biomarkers-based clinical predictions in glioblastoma. *Oncoimmunology.* 2017;6(4):e1295903.
- Pitt JM, Vetzizou M, Waldschmitt N, Kroemer G, Chamillard M, Boneca IG, Zitvogel L. Fine-tuning Cancer immunotherapy: optimizing the gut microbiome. *Cancer Res.* 2016;76(16):4602–7.
- Zakharevich NV, Averina OV, Klimina KM, Kudryavtseva AV, Kasianov AS, Makeev VJ, Danilenko VN. Complete genome sequence of *Bifidobacterium longum* GT15: identification and characterization of unique and global regulatory genes. *Microb Ecol.* 2015;70(3):819–34.
- Zakharevich NV, Averina OV, Klimina KM, Kudryavtseva AV, Kasianov AS, Makeev VJ, Danilenko VN. Complete Genome Sequence of *Bifidobacterium longum* GT15: Unique Genes for Russian Strains. *Genome Announc.* 2014;2(6).
- Patel SP, Kurzrock R. PD-L1 expression as a predictive biomarker in Cancer immunotherapy. *Mol Cancer Ther.* 2015;14(4):847–56.

33. Lee N, Zakka LR, Mihm MC Jr, Schatton T. Tumour-infiltrating lymphocytes in melanoma prognosis and cancer immunotherapy. *Pathology*. 2016;48(2):177–87.
34. Snezhkina AV, Lukyanova EN, Kalinin DV, Pokrovsky AV, Dmitriev AA, Koroban NV, Pudova EA, Fedorova MS, Volchenko NN, Stepanov OA, et al. Exome analysis of carotid body tumor. *BMC Med Genet*. 2018;11(Suppl 1):17.
35. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
36. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25(14):1754–60.
37. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22(3):568–76.
38. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213–9.
39. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
40. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, Rosenbloom K, Clawson H, Spieth J, Hillier LW, Richards S, et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res*. 2005;15(8):1034–50.
41. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, Brown SD, Chang HY, El-Gebali S, Fraser MI, et al. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res*. 2018.
42. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nat Protoc* 2016, 11(1):1–9.
43. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248–9.
44. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 2014;11(4):361–2.
45. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. *Genome Res*. 2009;19(9):1553–61.
46. Li Q, Wang K. InterVar: clinical interpretation of genetic variants by the 2015 ACMG-AMP guidelines. *Am J Hum Genet*. 2017;100(2):267–80.
47. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. *PLoS One*. 2012;7(10):e46688.
48. Jagadeesh KA, Wenger AM, Berger MJ, Guturu H, Stenson PD, Cooper DN, Bernstein JA, Bejerano G. M-CAP eliminates a majority of variants of uncertain significance in clinical exomes at high sensitivity. *Nat Genet*. 2016;48(12):1581–6.
49. Dong C, Wei P, Jian X, Gibbs R, Boerwinkle E, Wang K, Liu X. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Hum Mol Genet*. 2015;24(8):2125–37.
50. Do H, Dobrovic A. Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin Chem*. 2015;61(1):64–71.
51. Riemann K, Sotlar K, Kupka S, Braun S, Zenner HP, Preyer S, Pfister M, Pusch CM, Blin N. Chromosome 11 monosomy in conjunction with a mutated SDHD initiation codon in nonfamilial paraganglioma cases. *Cancer Genet Cytogenet*. 2004;150(2):128–35.
52. Zhikrivetskaya SO, Snezhkina AV, Zaretsky AR, Alekseev BY, Pokrovsky AV, Golovyuk AL, Melnikova NV, Stepanov OA, Kalinin DV, Moskalev AA, et al. Molecular markers of paragangliomas/pheochromocytomas. *Oncotarget*. 2017;8(15):25756–82.
53. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature*. 1998;396(6712):643–9.
54. Palmieri G, Colombino M, Cossu A, Marchetti A, Botti G, Ascierto PA. Genetic instability and increased mutational load: which diagnostic tool best direct patients with cancer to immunotherapy? *J Transl Med*. 2017;15(1):17.
55. Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM, Proszek PZ, Melchor L, Pawlyn C, Kaiser MF, et al. APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nat Commun*. 2015;6:6997.
56. Miller A, Asmann Y, Cattaneo L, Braggio E, Keats J, Auclair D, Lonial S, Network MC, Russell SJ, Stewart AK. High somatic mutation and neoantigen burden are correlated with decreased progression-free survival in multiple myeloma. *Blood Cancer J*. 2017;7(9):e612.
57. Gardner P, Dalsing M, Weisberger E, Sawchuk A, Miyamoto R. Carotid body tumors, inheritance, and a high incidence of associated cervical paragangliomas. *Am J Surg*. 1996;172(2):196–9.
58. van Nederveen FH, Gaal J, Favier J, Korpershoek E, Oldenburg RA, de Bruyn EM, Sleddens HF, Derckx P, Riviere J, Dannenberg H, et al. An immunohistochemical procedure to detect patients with paraganglioma and pheochromocytoma with germline SDHB, SDHC, or SDHD gene mutations: a retrospective and prospective analysis. *Lancet Oncol*. 2009;10(8):764–71.
59. Benn DE, Croxson MS, Tucker K, Bambach CP, Richardson AL, Delbridge L, Pullan PT, Hammond J, Marsh DJ, Robinson BG. Novel succinate dehydrogenase subunit B (SDHB) mutations in familial pheochromocytomas and paragangliomas, but an absence of somatic SDHB mutations in sporadic pheochromocytomas. *Oncogene*. 2003;22(9):1358–64.
60. Benn DE, Gimenez-Roqueplo AP, Reilly JR, Bertherat J, Burgess J, Byth K, Croxson M, Dahia PL, Elston M, Gimm O, et al. Clinical presentation and penetrance of pheochromocytoma/paraganglioma syndromes. *J Clin Endocrinol Metab*. 2006;91(3):827–36.
61. Ricketts C, Woodward ER, Killick P, Morris MR, Astuti D, Latif F, Maher ER. Germline SDHB mutations and familial renal cell carcinoma. *J Natl Cancer Inst*. 2008;100(17):1260–2.
62. Mason EF, Sadow PM, Wagner AJ, Remillard SP, Flood TA, Belanger EC, Hornick JL, Barletta JA. Identification of succinate dehydrogenase-deficient bladder paragangliomas. *Am J Surg Pathol*. 2013;37(10):1612–8.
63. Ghayee HK, Havekes B, Corssmit EP, Eisenhofer G, Hammes SR, Ahmad Z, Tessnow A, Lazurova I, Adams KT, Fojo AT, et al. Mediastinal paragangliomas: association with mutations in the succinate dehydrogenase genes and aggressive behavior. *Endocr Relat Cancer*. 2009;16(1):291–9.
64. Bayley JP, van Minderhout I, Weiss MM, Jansen JC, Oomen PH, Menko FH, Pasini B, Ferrando B, Wong N, Alpert LC, et al. Mutation analysis of SDHB and SDHC: novel germline mutations in sporadic head and neck paraganglioma and familial paraganglioma and/or pheochromocytoma. *BMC Med Genet*. 2006;7:1.
65. Gimenez-Roqueplo AP, Favier J, Rustin P, Rieubland C, Crespín M, Nau V, Khau Van Kien P, Corvol P, Plouin PF, Jeunemaitre X, et al. Mutations in the SDHB gene are associated with extra-adrenal and/or malignant pheochromocytomas. *Cancer Res*. 2003;63(17):5615–21.
66. Amar L, Bertherat J, Baudin E, Ajzenberg C, Bressac-de Paillerets B, Chabre O, Chamontin B, Delemer B, Giraud S, Murat A, et al. Genetic testing in pheochromocytoma or functional paraganglioma. *J Clin Oncol*. 2005;23(34):8812–8.
67. Burnichon N, Rohmer V, Amar L, Herman P, Lebouleux S, Darrouzet V, Niccoli P, Gaillard D, Chabrier G, Chabolle F, et al. The succinate dehydrogenase genetic testing in a large prospective series of patients with paragangliomas. *J Clin Endocrinol Metab*. 2009;94(8):2817–27.
68. Baysal BE, Ferrell RE, Willett-Brozick JE, Lawrence EC, Myssiorek D, Bosch A, van der Mey A, Taschner PE, Rubinstein WS, Myers EN, et al. Mutations in SDHD, a mitochondrial complex II gene, in hereditary paraganglioma. *Science*. 2000;287(5454):848–51.
69. Poeppel TD, Yucee A, Boy C, Metz KA, Kaminsky E, Neumann HP, Rosenbaum SJ, Mann K, Moeller LC. Novel SDHD gene mutation (H102R) in a patient with metastatic cervical paraganglioma effectively treated by peptide receptor radionuclide therapy. *J Clin Oncol*. 2011;29(33):e812–5.
70. Colombo-Benkmann M, Li Z, Riemann B, Hengst K, Herbst H, Keuser R, Gross U, Rondot S, Raue F, Senninger N, et al. Characterization of the RET protooncogene transmembrane domain mutation S649L associated with nonaggressive medullary thyroid carcinoma. *Eur J Endocrinol*. 2008;158(6):811–6.
71. Prazeres H, Couto JP, Rodrigues F, Vinagre J, Torres J, Trovisco V, Martins TC, Sobrinho-Simoes M, Soares P. In vitro transforming potential, intracellular signaling properties, and sensitivity to a kinase inhibitor (sorafenib) of RET proto-oncogene variants Glu511Lys, Ser649Leu, and Arg886Trp. *Endocr Relat Cancer*. 2011;18(4):401–12.
72. Eric Z, Hoffmann MM, Sullivan M, Franke G, Peczkowska M, Harsch I, Schott M, Gabbert HE, Valimaki M, Preuss SF, et al. Pathogenicity of DNA variants and double mutations in multiple endocrine neoplasia type 2 and von Hippel-Lindau syndrome. *J Clin Endocrinol Metab*. 2010;95(1):308–13.
73. Li J, Duns G, Westers H, Sijmons R, van den Berg A, Kok K. SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget*. 2016;7(31):50719–34.

74. Viaene AN, Santi M, Rosenbaum J, Li MM, Surrey LF, Nasrallah MP. SETD2 mutations in primary central nervous system tumors. *Acta Neuropathol Commun.* 2018;6(1):123.
75. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med.* 2014;20(5):548–54.
76. Newman AM, Lovejoy AF, Klass DM, Kurtz DM, Chabon JJ, Scherer F, Stehr H, Liu CL, Bratman SV, Say C, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol.* 2016;34(5):547–55.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

