Novel potential causative genes in carotid paragangliomas

Anastasiya V. Snezhkina¹, Elena N. Lukyanova¹, Andrew R. Zaretsky¹, Dmitry V. Kalinin², Anatoly V. Pokrovsky², Alexander L. Golovyuk², George S. Krasnov¹, Maria S. Fedorova¹, Elena A. Pudova¹, Sergey L. Kharitonov¹, Nataliya V. Melnikova¹, Boris Y. Alekseev³, Marina V. Kiseleva³, Andrey D. Kaprin³, Alexey A. Dmitriev¹ and Anna V. Kudryavtseva¹∗

From 11th International Multiconference “Bioinformatics of Genome Regulation and Structure\Systems Biology” - BGRSSB-2018
Novosibirsk, Russia. 20-25 August 2018

Abstract

Background: Carotid paragangliomas (CPGLs) are rare neuroendocrine tumors that arise from the paraganglion at the bifurcation of the carotid artery and are responsible for approximately 65% of all head and neck paragangliomas. CPGLs can occur sporadically or along with different hereditary tumor syndromes. Approximately 30 genes are known to be associated with CPGLs. However, the genetic basis behind the development of these tumors is not fully elucidated, and the molecular mechanisms underlying CPGL pathogenesis remain unclear.

Methods: Whole exome and transcriptome high-throughput sequencing of CPGLs was performed on an Illumina platform. Exome libraries were prepared using a Nextera Rapid Capture Exome Kit (Illumina) and were sequenced under 75 bp paired-end model. For cDNA library preparation, a TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Gold (Illumina) was used; transcriptome sequencing was carried out with 100 bp paired-end read length. Obtained data were analyzed using xseq which estimates the influence of mutations on gene expression profiles allowing to identify potential causative genes.

Results: We identified a total of 16 candidate genes (MYH15, CSP1, MYH3, PTGES3L, CSGALNACT2, NMD3, IFI44, GMCL1, LSP1, PPFIBP2, RBL2, MAGED1, CNIH3, STRA6, SLC6A13, and ATM) whose variants potentially influence their expression (cis-effect). The strongest cis-effect of loss-of-function variants was found in MYH15, CSP1, and MYH3, and several likely pathogenic variants in these genes associated with CPGLs were predicted.

Conclusions: Using the xseq probabilistic model, three novel potential causative genes, namely MYH15, CSP1, and MYH3, were identified in carotid paragangliomas.

Keywords: Carotid paragangliomas, Tumor-associated genes, Pathogenic variants, High-throughput sequencing, Exome, Transcriptome, Xseq

* 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

© The Author(s). 2019 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
Background
Paragangliomas are rare neuroendocrine neoplasms derived from paraganglionic tissue. In the head and neck, paragangliomas most frequently arise from carotid glomus at the bifurcation of the carotid artery [1]. Carotid paragangliomas (CPGLs) are highly vascularized tumors and are anatomically classified on three groups based on the size and involvement of the carotid artery [2]. The majority of carotid paragangliomas belong to groups 2 and 3 (primarily large size CPGLs with moderate or high arterial attachment), which are surgically challenging. CPGLs are typically characterized by slow-growing and non-aggressive tumors. Up to 10% of CPGLs form regional or distant metastases (malignant CPGLs) [3]. Approximately 10% of CPGLs are hereditary [4]. The new WHO classification describes paragangliomas as tumors with variable metastatic potential [5].

Almost all tumors, including paragangliomas, are characterized by altered energy metabolism [6–8]. Several genes involved in energy metabolism were found to be critical for CPGL pathogenesis. Germline variants in the SDHB [9], SDHC [10], SDHD [11], RET [12], and VHL [13, 14] genes were associated with carotid paragangliomas. Somatic variants in these genes, as well as ones in SDHA, SDHAF2, IDH1, NF1, MEN1, and KIF1B have also been described in sporadic CPGLs [10, 15, 16].

Germline and somatic variants in SDHx genes are often found in head and neck paragangliomas [17, 18]. These genes encode the four subunits (SDHA, SDHB, SDHC, and SDHD) of succinate dehydrogenase (SDH), also known as the mitochondrial complex II. SDH is a component of both the oxidative phosphorylation (OXPHOS) system and the tricarboxylic acid (TCA) cycle, which are key metabolic pathways in the mitochondria. Dysfunctional SDH resulting from genetic or epigenetic alterations can contribute to various pathologies, including cancer [19]. The molecular mechanisms underlying tumorigenesis include the increased generation of reactive oxygen species (ROS) and the disruption of mitochondrial function during apoptosis [20]. On the other hand, dysfunctional SDH leads to succinate accumulation, which in turn affects the activity of the prolyl hydroxylase (PHD) enzyme that catalyzes the hydroxylation of hypoxia inducible factor (HIF). Impaired PHD activity disrupts the binding between HIF and von Hippel-Lindau tumor suppressor protein (pVHL) and subsequently leads to HIF degradation via the ubiquitin-proteasome pathway [21]. The development of paragangliomas is often associated with the pseudohypoxia state caused by the stabilization and activation of HIF [22, 23]. Variants in SDHD occur in most carotid paragangliomas, whereas rarer SDHB variants result in aggressive disease and metastasis [24–26].

Variants in other known paragangioma/pheochromocytoma-causative genes (e.g., RET, NF1, and IDH1) are less common in CPGLs and are primarily observed in sporadic ones. RET and NF1 are a proto-oncogene and a tumor suppressor gene respectively; their protein products participate in the PI3K/AKT and RAS/MAPK pathways and mTOR signaling [27, 28]. Variants in RET and NF1 lead to deregulation of these pathways and are believed to be critical drivers of tumorigenesis. Several variants of IDH1 but not those of IDH2 were reported in CPGLs [10, 16]. Inactivation mutations in IDH1 are known to result in HIF accumulation and tumor development under pseudohypoxia conditions [29]. Two potential pathogenic variants in KIF1B were identified in CPGLs [16]. KIF1B encodes a motor protein responsible for organelle transport and cell division [30, 31]. Several studies have demonstrated the involvement of KIF1B in tumorigenesis and its tumor suppressor role [32–34]. A study using rat pheochromocytoma cells indicated that the effects of KIF1B in apoptosis are mediated by the Egln3 prolyl hydroxylase pathway [35]. Variants associated with CPGLs were detected in a group of TP53-related genes [TP53BP1, TP53BP2 (ASPP2), and TP53I13 (DSCP1)] that encode proteins involved in the regulation of cell proliferation, DNA damage response, and apoptosis. These proteins can act as tumor suppressors by modulating p53 function and promoting apoptosis [36–38]. The mutated ARNT (HIF1B) gene, which encodes an aryl hydrocarbon nuclear translocator, was identified in CPGLs [16]. ARNT functions as a co-factor for various transcription factors, including HIF1α [39]. Aberrant ARNT expression can contribute to tumor growth via interactions with specific transcription factors that regulate the expression of multiple genes [40]. In addition, likely pathogenic variants have been identified in MEN1, BAP1, BRAF, BRCa1, BRCa2, CDKN2A, CSDE1, and FGFR3 in individual CPGLs [16]. The above findings indicated high genetic inter-heterogeneity of these tumors.

Although recent studies identified several potential mechanisms involved in paraganglioma pathogenesis, further investigations are required to elucidate the molecular genetic basis of the paragangliomas with different localizations and predispositions to be aggressive. Moreover, most reports focused on mutational status of the genes associated with CPGLs. The functional effects of germline or somatic mutations on gene expression in paragangliomas remain poorly understood. Therefore, the present study aimed to identify the potential causative genes associated with carotid paragangliomas. Using exome data of CPGLs from previous study [16] and RNA-Seq data, we evaluated the effects of somatic variants on gene expression profiles using the xseq probabilistic model. This approach allowed the identification
of novel candidate genes and likely pathogenic variants involved in the pathogenesis of CPGLs.

Methods
Patients and samples
Carotid paraganglioma samples were collected after obtaining written informed consent from the patients in the Vishnevsky Institute of Surgery. Pathological evaluation was performed for all samples, and the minimum percentage of tumor cells was 80%. The study was approved by the ethics committee from Vishnevsky Institute of Surgery and was conducted in accordance with the principles outlined in the Declaration of Helsinki (1964). Clinicopathologic characteristics of the CPGL patients are presented in Table 1.

RNA extraction
Total RNA was extracted from 25 formalin-fixed paraffin-embedded (FFPE) tissues of carotid paragangliomas using a RNeasy FFPE Kit (Qiagen, Germany). RNA concentration was measured on a Qubit 2.0 fluorometer (Thermo Fisher Scientific, USA), and RNA quality was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, USA).

Library preparation and transcriptome sequencing
The cDNA libraries were prepared using a TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Gold (Illumina, Germany) according to the manufacturer’s instructions. Paired-end (2 × 100 bp) sequencing was performed on an Illumina HiSeq 2500 Sequencing System (Illumina). RNA-Seq data have been deposited to the NCBI Sequence Read Archive (SRA) under accession number PRJNA476709.

Raw reads were subjected to quality control with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter sequences and low-quality reads were removed using Trimmomatic [41]. Kallisto [42] was used for pseudoalignment and quantifying abundances of transcripts. Gene expression levels were normalized using the TPM (transcripts per million) method.

Whole-exome sequencing data
We analyzed exome data of carotid paragangliomas generated from a previous study [16]. In that study, exome libraries were prepared using a Nextera Rapid Capture Exome Kit (Illumina). High-throughput sequencing of 75 bp paired-end reads was carried out on a NextSeq500 System (Illumina) at a median sequencing depth of 300×. The raw sequence reads are available at the SRA under the accession number PRJNA411769.

The quality of raw sequencing data was examined with FastQC. Trimmomatic was used to trim low-quality reads and adapter sequences. Reads were then mapped to the reference human genome (GRCh37/hg19) using BWA-MEM [43]. SAMTools software [44, 45] was applied for removing of duplicate reads, sorting, merging, etc. Freebayes [46] was used for variant calling; VCF files were processed with submodules from vcflib (https://github.com/vcflib/vcflib). SnpSift from the SnpEff toolbox [47] was used for annotation of genomic variants; the HGMD, OMIM, ClinVar, dbSNP, dbNSFP, COSMIC, ConsensusPathDB, 1000 Genomes Project, and ExAC databases were used as additional resources. Pathogenicity of all variants was assessed in silico using the SIFT [48], PolyPhen2 [49], Mutation-Taster [50], and LRT [51] algorithms. PhastCons [52] and PhyloP [53] methods were used to analyze the evolutionary conservation of regions harboring the target variants.

Using exome and transcriptome data from the same carotid paragangliomas (25 samples), we evaluated the impact of somatic loss-of-function variants on gene expression profiles using xseq [54]. The xseq model evaluates the posterior probability of the mutation influence in a certain gene on the expression profiles of a number of genes, including itself (cis-effect) and associated genes (trans-effect). The output of xseq consists of \( P(F_{g,m}) \) and \( P(D_g) \) values, where \( P(F_{g,m}) \) is the probability that an individual mutation in gene \( g \) in an individual patient \( m \) influences expression within that patient, and \( P(D_g) \) is the probability that a recurrently mutated gene \( g \) influences gene expression across the population of patients [54].

Validation of variants by Sanger sequencing
Potential pathogenic variants identified using xseq were selected and validated by Sanger sequencing. Amplification was carried out in 20 μL final reaction volume using a Tersus PCR Kit (Evrogen, Russia). Nested-PCR was performed to amplify the two variants (c.2706-7855G > A and c.2706-7798delA) in the CSP1 gene. For the first

| Characteristic                        | Total, n |
|--------------------------------------|----------|
| Gender                               |          |
| Male                                 | 5        |
| Female                               | 20       |
| Age (years)                          |          |
| ≤ 50                                 | 13       |
| > 50                                 | 12       |
| Family history of paragangliomas     |          |
| Yes                                  | 1        |
| No                                   | 1        |
| N/A                                  | 23       |

Table 1 Clinicopathologic characteristics of the CPGL patients
round of amplification, the outer primers (F: 5′-GGAGGCTGAGGCGGGCAGATC-3′ and R1: 5′-CACCATTCTCCTGCGCTTCAGCCCTCT-3′) producing a 181 bp product were used. PCR products from the first step were diluted at 1:1000 and subsequently amplified using the inner primers (F: 5′-GGAGGCTGAGGCGGGCAGATC-3′ and R2: 5′-GTAGCTGGGACTACAGGCGCCT-3′) to produce a 154 bp fragment. PCR and nested-PCR were performed using the following amplification profile: initial denaturation step at 95 °C for 3 min; 35 cycles of denaturation at 95 °C for 30 s, annealing at 60–65 °C for 30 s, and extension at 72 °C for 60 s; and a final extension of 72 °C for 5 min. All PCR products were subjected to 2% agarose gel electrophoresis and visualized using a Gel Doc XR+ Gel Documentation System (BioRad, USA) with GelRed DNA stain (Biotium, USA). Primer sequences are listed in Table 2.

**Results**

Using the xseq model, we estimated the probability that potential loss-of-function somatic variants (frameshift, nonsense, and splice-site variants) impact gene expression in individual patients [P(F) value] and across the population of patients [P(D) value]. The essential first step was the identification of somatic variants. In the study, we analyzed previously collected data from carotid paraganglioma patients. Tumor samples were collected over a span of more than 20 years in the Vishnevsky Institute of Surgery; no blood samples were obtained from the same patients to date. Thus, data from the 1000 Genomes Project and ExAC databases were used to exclude germline variants. Variants with more than 1% frequency in these databases were excluded from the analysis. We identified 16 genes (MYH15, CSP1, MYH3, PTGES3L, CSGALNACT2, NMD3, IFI44, GMCL1, LPSI, PPFIBP2, RBL2, MAGED1, CNH3, STRA6, SLC6A13, and ATM) whose somatic variants can influence their gene expression. These genes had maximum P(F) and P(D) values ranging from 0.51 to 0.99 and from 0.32 to 0.99 respectively. Based on the criteria, MYH15, CSP1, and MYH3 were considered to have significant cis-effect of somatic variants on gene expression and were proposed as candidate genes associated with carotid paragangliomas.

To analyze all potential somatic loss-of-function variants in MYH15, CSP1, and MYH3, novel filter conditions were set for variants derived from the 1000 Genomes Project and ExAC databases to exclude germline variants. Variants with more than 1% frequency in these databases were excluded from the analysis. We identified 16 genes (MYH15, CSP1, MYH3, PTGES3L, CSGALNACT2, NMD3, IFI44, GMCL1, LPSI, PPFIBP2, RBL2, MAGED1, CNH3, STRA6, SLC6A13, and ATM) whose somatic variants can influence their gene expression. These genes had maximum P(F) and P(D) values ranging from 0.51 to 0.99 and from 0.32 to 0.99 respectively. Based on the criteria, MYH15, CSP1, and MYH3 were considered to have significant cis-effect of somatic variants on gene expression and were proposed as candidate genes associated with carotid paragangliomas.

**Table 2** List of primers used for validation of the variants

| Gene   | Variant          | Primer sequence (5′ → 3′)                  | Product size (bp) |
|--------|------------------|--------------------------------------------|-------------------|
| CPS1   | c.2706-7855G > A | F: GGAGGCTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 181               |
|        | c.2706-7798delA  | F1: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 154               |
|        | c.2706-7745C > G | F1: GAG ACC ATC CTG GCT AAC ACG GTG AAA C<br>R1: GTG CCG CTC TCT CTC CCA GGC TG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 195               |
|        | p.Gly1382Ser/c.4144G > A | F: CTAAAATCTCCTACACCATCCTGAGAAG<br>R: AGATTCCTGAGAAGAAAGAAAC<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 276               |
|        | p.Arg1748Gln/c.5243G > A | F: TCTCTTCTGAGTAGAAGACCATCTGCT<br>R: TTTTTCATTGAGAGCGCAAC<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 113               |
| MYH15  | p.Glu1811Lys/c.5431G > A | F: ACTGCTCAAGGCTCCTGATTACC<br>R: AAGAAAGCAGAAGGAGGAGAAGAC<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 190               |
|        | p.Ala1003Val/c.3008C > T | F: TCTCTACTGAGTTAAGACCCCTTACTGAG<br>R: GTTTTCTCTGAGTTTGTGTCAAC<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 167               |
| MYH3   | p.Ala1198Thr/c.3592G > A | F: TGACCCGGCACTCGAGGTTG<br>R: CTGGTACCCCTCAGGAGATAG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 176               |
|        | p.Ala1604Thr/c.4810G > A | F: TCAGCTCCTCCAGGAGGTTG<br>R: CTGAGACGGAGGAGGAGGAGGAGG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 118               |
|        | p.Ala1752Thr/c.5254G > A | F: CTCATGCAGCTCGAGGAGG<br>R: GGGCTGTAGAAGAGGAGGAGG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 141               |
|        | p.Ile264Thr/c.791T > C | F: CAAAAAGAGGCACTGCTTTCTTACA<br>R: GCCCTGTTATGCAATGCCAATCG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 182               |
|        | p.Met798Val/c.2392A > G | F: GAGCTCTGACGTTGAGGAGG<br>R: GCCCTGTTATGCAATGCCAATCG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 142               |
|        | p.Ser1240Leu/c.3719C > T | F: CAGGTGTCAATGACGGAGGAGG<br>R: GGAGAAGGAGAAGGAGGAGG<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 142               |
|        | p.Thr442Ile/c.1325C > T | F: GATCTTGAAACACACACGAGGACT<br>R: GGAGGCTGAGGCGGGCAGATC<br>R2: GTACCGTGAGGCGGGCAGATC<br>R1: CACCATTCTCCTGCGCTTCAGCCCTCT<br>R2: GTAGCTGGGACTACAGGCGCCT | 195               |

*Amplification of fragments carrying two variants (c.2706-7855G > A and c.2706-7798delA) in the CSP1 gene was performed using nested-PCR*
Project and ExAC; in particular, variants with frequencies of 2% or less were used in the analysis. Predictions using the SIFT, PolyPhen2, MutationTaster, or LRT algorithms indicated that the MYH3 gene was predominantly enriched in likely pathogenic variants (Table 3). The PhastCons and PhyloP values for these variants were also high, indicating strong evolutionary conservation. However, several variants were frequently identified in the carotid paraganglioma samples and appear to be germline ones. These variants were excluded from the sample set, and the cis-effects were measured again using xseq. In this case, only MYH15 was characterized by variants that can result in gene expression alterations [average P(F) value – 0.66; and P(D) value – 0.68].

All identified variants were validated by Sanger sequencing.

### Discussion

Using integrative analysis of whole exome and RNA-Seq data from a representative sample set, we identified novel potential causative genes in carotid paragangliomas. We employed methods based on the calculation of the posterior probability of a mutation to influence gene expression. From the set of 16 candidate genes, somatic loss-of-function variants in three genes, namely CPS1, MYH15, and MYH3, were found to significantly alter their gene expression levels.

*CPS1* encodes carbamoyl-phosphate synthase 1, a mitochondrial enzyme that mediates the first step of the

---

**Table 3** Likely pathogenic variants in MYH15, CPS1, and MYH3

| Gene | rs ID number | GenBank Coordinate | Nucleotide change | Amino acid change | Genotype/number of samples | Predictions |
|------|-------------|-------------------|-------------------|-------------------|----------------------------|-------------|
| **CPS1** | – | Chr2: 211494571 | c.2706-7855G > A | – | Het/2 | SIFT PolyPhen2 MutationTaster LRT |
| | – | Chr2: 211494681 | c.2706-7745C > G | – | Het/1 | N/A N/A N/A N/A |
| | – | Chr2: 211494618 | c.2706-7798delA | – | Het/1 | N/A N/A N/A N/A |
| | – | NM_001122633.2 Chr2: 211539650 | c.4144G > A | p.Gly1382Ser | Het/2 | Tolerated Benign Disease-causing Deleterious |
| **MYH15** | rs56118396 | NM_014981.1 Chr3: 108112954 | c.5243G > A | p.Arg1748Gln | Het/4 | Tolerated Benign Disease-causing N/A |
| | – | Chr3: 108110666 | c.5431G > A | p.Glu1811Lys | Het/1 | Deleterious Possibly damaging Disease-causing N/A |
| **MYH3** | rs34088014 | NM_002470.3 Chr17: 10542709 | c.3008C > T | p.Ala1003Val | Het/2 | Tolerated Benign Disease-causing N/A |
| | – | Chr17: 10541997 | c.3592G > A | p.Ala1198Thr | Het/1 | Tolerated Possibly damaging Disease-causing N/A |
| | – | Chr17: 10539939 | c.4810G > A | p.Ala1604Thr | Het/1 | Deleterious Possibly damaging Disease-causing N/A |
| | – | Chr17: 10534960 | c.5254G > A | p.Ala1752Thr | Het/3 | Deleterious Possibly damaging Disease-causing N/A |
| | – | Chr17: 10550688 | c.791 T > C | p.Ile264Thr | Het/1 | Deleterious Possibly damaging Disease-causing N/A |
| | – | Chr17: 10543684 | c.2392A > G | p.Met798Val | Het/1 | Deleterious Benign Possibly disease-causing N/A |
| | – | Chr17: 10541370 | c.3719C > T | p.Ser1240Leu | Het/1 | Deleterious Benign Disease-causing N/A |
| | – | Chr17: 10547753 | c.1325C > T | p.Thr442Ile | Het/1 | Tolerated Benign Possibly disease-causing N/A |

SIFT scores range from 0 to 1, and variants with SIFT scores less than 0.05 are considered deleterious. SIFT scores ranging from 0.05 to 1 are predicted to be tolerated variants.

The PolyPhen-2 scores range from 0 to 1. Variants with scores ranging from 0 to 0.15 in this range are predicted to be benign. Variants with scores ranging from 0.15 to 0.85 are potential damaging, whereas values above 0.85 have higher probability of being damaging variants.

MutationTaster p-values (probability) range from 0 to 1. P-values close to 1 indicate pathogenicity (disease-causing).

LRT is a likelihood ratio test. LRT scores range from 0 (neutral) to 1 (deleterious).

Het heterozygous genotype.
urea cycle [55]. Germline variants in CPS1 cause rare autosomal recessive disorders called carbamoyl-phosphate synthetase 1 deficiency (CPSID) [56]. Recent studies have indicated that CPS1 is primarily expressed in liver tissues and participates in the conversion of ammonia to urea in hepatocytes in humans. However, several studies have demonstrated the involvement of CPS1 in carcinogenesis. Alterations in CPS1 were identified in malignant glioma [57]. CPS1 expression has been associated with the initiation and progression of colorectal cancer [58]. Hypermethylation of CPS1 accompanied by its downregulated expression was detected in hepatocellular carcinoma [59]. On the other hand, CPS1 overexpression was observed in rectal, lung, and cervical cancers and was correlated with poor prognosis [60, 61]. In lung adenocarcinoma cells, inhibition of CPS1 expression was found to reduce tumor growth [61, 62]. Our current findings revealed that CPS1 is a potential causative gene in carotid paragangliomas. The variants in CPS1 can influence its expression level. Both genetic and epigenetic factors influence the expression of CPS1, consistent with previous studies and our current findings. Similar to the development of inherited CPSID, variants identified in CPS1 can be assumed to cause CPS1 inactivation, thereby leading to CPGLs. In turn, CPS1 variants could lead to alterations in the urea cycle and accumulation of toxic ammonia, which is normally converted to urea. The brain is known to be highly sensitive to the effects of excess ammonia. Ammonia can be transported through the cell membranes in many cell types, including cells of the nervous tissue. In vitro studies have demonstrated that the presence of glutamate and ammonium is necessary for lactate release in astrocytes and glial cells [63]. In turn, extracellular lactate is associated with malignant transformation and can regulate cancer-related signaling pathways (e.g., NF-kB/IL-8, HIF1, and PI3K/Akt/mTOR/Bcl-2), angiogenesis, and ATP production, which in turn influence cell proliferation and migration [64–66]. Such molecular mechanism may be responsible for the development of CPGLs in some cases.

Variants in MYH3 and MYH15 were additionally demonstrated to exert cis-effects on gene expression in carotid paragangliomas. MYH3 and MYH15 belong to the highly conserved myosin heavy chain (MYH) family that encode heavy chains of myosins [67]. Myosins are motor proteins consisting of heavy and light chains and are involved in various types of cellular movements resulting from actin-myosin interactions. In muscle cells, myosin and actin filaments form myofibrils; in non-muscle cells, myosins were demonstrated to participate in various biological processes, such as transport of cellular organelles, chromosomal segregation, actin organization, plasma membrane tension, endocytosis, cell division, and cell motility [68, 69]. Expression of MYH genes varies in human tissues. MYH3 is expressed during the early stages of development in skeletal muscles, and its expression decreases after birth, whereas MYH15 expression is exclusively detected in adult extraocular muscles [70, 71]. Data from the Human Proteome Atlas (HPA, https://www.proteinatlas.org) indicate that the peak expression of the MYH3 protein is observed in the heart and skeletal muscles, esophagus, prostate, and seminal vesicle. MYH15 is expressed in the brain, lungs, small intestines, and testes. The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/) project, which contains RNA-Seq data from various types of cancer, indicate that MYH3 and MYH15 are expressed in almost all tumors at the lowest levels. Slightly elevated mRNA level of MYH3 gene was found in head and neck cancer with average fragments per kilobase million (FPKM) 3.3 in a set of 499 samples. However, MYH15 proteins are expressed in glioma, urothelial and colorectal cancers (HPA data). In addition, proteome studies revealed deferential expression of MYH15 in early onset dementia, an atypical frontotemporal lobar degeneration (aFTLD) [72].

Variants in these genes were observed in different diseases. Variants in MYH3 were detected as benzo(a)-pyrene exposure-genomic alterations in lung cancer patients [73]. Germline variants in MYH3 are associated with distal arthrogryposis syndromes [distal arthrogryposis type 1, Freeman-Sheldon syndrome (DA2A), and Sheldon-Hall syndrome (DA2B)], multiple pterygium syndrome (MPS), and spondylocarpotarsal synostosis syndrome (SCT) [74–77]. These disorders are characterized by skeletal anomalies and multiple congenital contractures in the limbs, thereby indicating an essential role of MYH3 in bone development. Variants in MYH15 were reported to be associated with higher risk for non-cardioembolic stroke and ischemic stroke [78, 79]. In the present study, we identified several likely pathogenic variants in MYH3 and MYH15 in carotid paragangliomas, and most of these variants were found to be located in MYH3. Currently, the direct roles of MYH3 and MYH15 in normal cellular processes and the development of various pathologies, including cancer, remain unclear. Moreover, only few studies in literature have reported the involvement of MYH3 and MYH15 in tumorigenesis. Therefore, variants identified in MYH3 and MYH15 are likely to affect the function of MYH proteins and are potentially associated with carotid paragangliomas. However, further studies are required to elucidate the mechanisms by which MYH3 and MYH15 can promote tumor pathogenesis.

Conclusions

In the present study, we identified several novel genes, namely CPS1, MYH15, and MYH3, associated with carotid paragangliomas. These results contribute to a better
understanding of the molecular mechanisms behind the pathogenesis of CPGLs.

Acknowledgments

The authors thank the Vishnevsky Institute of Surgery for providing the samples and their morphological characterization and the National Medical Research Radiological Center for the assistance in data analysis. This work was performed using the equipment of the EMB RAS "Genome“ center [http://www.emb.ru/rus/kdp/cccu_genome_c.php].

Funding

This work and publication costs were funded 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. Whole-exome sequence data are available at the NCBI Sequence Read Archive (SRA) under the accession number PRJNA411769. RNA-Seq data are available at the NCBI SRA under accession number PRJNA476709.

About this supplement

This work has been published as part of BMC Medical Genetics Volume 20 Supplement 1, 2019: Selected articles from BGSIS-2018: medical genetics. The full contents of the supplement are available online at https://bmcmedgenet.biomedcentral.com/articles/supplements/volume-20-supplement-1.

Authors’ contributions

AVK, AVS, and ENL conceived and designed the work; MSF, EAP, SLK, DVK, ARZ, ALG, NVM, AVP, and AVS performed the experiments; ENL, GSK, MVK, AVK, AVS, 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 of Vishnevsky Institute of Surgery, Ministry of Health of the Russian Federation. The study was conducted in accordance with the principles outlined in 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: 9 April 2019

References

1. Paris J, Facon F, Thomassin JM, Zanaret M. Cervical paragangliomas: neurovascular surgical risk and therapeutic management. Eur Arch Otorhinolaryngol. 2006;263(9):860–5.
2. Chamblin WR, ReMine WH, Shipps SG, Harrison EG Jr. Carotid body tumor (chemodectoma). Clinicopathologic analysis of ninety cases. Am J Surg. 1971;122(6):732–9.
3. 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.
4. Riehl A, Ravi A, Vijayaraghavan K. Paraganglioma neck - a neuroendocrine tumour revisited. Int J Angiol. 2008;17(3):162–5.
5. El-Naggar AK, Chan JKC, Jennifer RG, Takashi T, Pieter JS. Classification of Head and Neck Tumours. Fourth edition. World Health Organization. 2017:9, 348.
6. Pudova EA, Kudystavtseva AV, Fedorova MS, Zaretsky AR, Shcherbo DS, Lukyanova EN, Popov AV, Sadridtinnova AF, Abramov IS, Khantovorov SL, et al. HR3 overexpression associated with epithelial-mesenchymal transition in colorectal cancer. BMC Genomics. 2018;19(Suppl 3):113.
7. Snezhkina AV, Krasnov GS, Zaretsky AR, Zhavoronok A, Nyzhko KM, Moskalev AA, Karpova IV, Afremova AL, Lipatova AI, Kochevtov DV, et al. Differential expression of alternatively spliced transcripts related to energy metabolism in colorectal cancer. BMC Genomics. 2016;17(Suppl 1):1011.
8. Snezhkina AV, Krasnov GS, Zhukrivetskaya SO, Karpova IV, Fedorova MS, Nyzhko KM, Belyakov MM, Gruchiev 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.
9. Collins N, Dietzek A. Contiguous bilateral head and neck paragangliomas in a carrier of the SDHB germline mutation. J Vasc Surg. 2012;55(1):216–9.
10. Gaal J, Burnichon N, Kopershoek E, Roncelin I, Bertherat J, Plouin P, de Krijger RR, Gimenez-Roqueplo AP, Dinjens WN. Lactate dehydrogenase- mutations are rare in pheochromocytomas and paragangliomas. J Clin Endocrinol Metab. 2010;95(3):1274–8.
11. Liapis CD, Bellos JK, Halapas A, Lembessis P, Koutsieliers M, Kostakis A. Carotid body paraganglioma and SDHD mutation in a Greek family. Anticancer Res. 2005;25(3):2449–52.
12. Choi Jd, Tucker KM, Lee TT, Chong GC. Hereditary paraganglioma- pheochromocytoma syndromes associated with SDHD and RET mutations. Head & neck. 2014;36(10):E99–E102.
13. Majchrzak K, Cybulski C, Bobek-Billewicz B, Majchrzak H, Lubinski J. A case of carotid body paraganglioma and haemangioblastoma of the spinal cord in a patient with the N131K missense mutation in the VHL gene. Neurosci. 2011;32(2):491–6.
14. Gross DJ, Avishai N, Meiner V, Filon D, Zbar B, Abeliovich D. Familial pheochromocytoma associated with a novel mutation in the von Hippel-Lindau gene. J Clin Endocrinol Metab. 1996;81(1):147–9.
15. Smith JD, Harvey RN, Daar OA, Prince ME, Bradford CR, Wolf GT, Else T, Basura GJ. Head and neck paragangliomas: a two-decade institutional experience and algorithm for management. Laryngoscope Investigative otolaryngology. 2017;2(6):380–9.
16. Snezhkina AV, Lukyanova EN, Kalinin DV, Pokrovsky AV, Kudristova 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.
17. Fakhry N, Naccoli-Sire P, Parlier-Sett A, Giorgi R, Giovanni A, Zanaret M. Cervical paragangliomas: is SDH genetic analysis systematically required? Eur Arch Otorhinolaryngol. 2008;265(5):567–63.
18. Chen H, Zhu W, Li X, Xue L, Wang Z, Wu H. Genetic and epigenetic patterns in patients with the head-and-neck paragangliomas associate with differential clinical characteristics. J Cancer Res Clin Oncol. 2017;143(6):953–60.
19. Bezawork-Geleta A, Rohlena J, Dong L, Pacak K, Neuzil J. Mitochondrial complex II at the crossroads. Trends Biochem Sci. 2017;42(4):312–25.
20. Kudystavtseva AV, Krasnov GS, Dmitriev AA, Alekseyev BY, Kardymon OL, Sadridtinnova 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.
21. Mole DR, Maxwell PH, Pugh CW, Ratcliffe PJ. Regulation of HIF by the von Hippel-Lindau tumour suppressor: implications for cellular oxygen sensing. JUBMB Life. 2001;52(1):243–7.
22. Selak MA, Armour SM, Mackenzie ED, Boulahbel H, Watson DG, Mansfield KD, Pan Y, Simon MC, Thompson CB, Gottlieb E. Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-alpha prolyl hydroxylase. Cancer Cell. 2005;7(1):77–87.
23. van Hulsteijn LT, Dekkers OM, Hes FJ, Smit JW, Corssmit EP. Risk of malignant paraganglioma in SDHB-mutation and SDHD-mutation carriers: a systematic review and meta-analysis. J Med Genet. 2012;49(12):768–75.
26. Boederck CC, Neumann HP, Maier W, Bausch B, Schipper J, Ridder GJ. Malignant head and neck paragangliomas in SDHB mutation carriers. Otolaryngol Head Neck Surg. 2007;137(1):126–9.

27. Jiang SM. The RET proto-oncogene in human cancers. Oncogene. 2000; 19(49):5900–7.

28. Phlipott C, Towll H, Frayling IM, Cooper DN, Upathyaya M. The NF1 somatic mutational landscape in sporadic human cancers. Human genomics. 2017;1(1):13.

29. Zhao S, Lin Y, Xu W, Jiang W, Zha W, Wang P, Yu W, Li Z, Gong L, Peng Y, et al. Gloma-derived mutations in IDH1 dominantly inhibit IDH1 catalytic activity and induce HIF-1alpha. Science. 2009;324(5924):261–5.

30. Nagamuk A, Sato-Yoshitake R, Okada Y, Noda Y, Takemura R, Yamazaki H, Hirokawa N. KIF18, a novel microtubule plus-end directed monomeric motor protein for transport of mitochondria. Cell. 1994;79(7):1209–20.

31. Kong WX, Wang DH, Zhou CJ, Zhou BX, Liang CG. Loss of function of KIF1B impairs cnoyte meiotic maturation and early embryo development in mice. Mol Reprod Dev. 2016;83(11):1027–40.

32. Munirajan AK, Anand K, Mukai A, Takahashi M, Suena A, Ohira M, Koda T, Hirata T, Ozaki T, Nakagawa A. KIF1Bbeta functions as a haploinsufficient tumor suppressor gene mapped to chromosome 1p56.2 by inducing apoptotic cell death. J Biol Chem. 2008;283(36):24246–34.

33. Shi TY, Jiang Z, Jiang R, Yin S, Wang MY, Yu KD, Shao ZM, Sun MH, Zeng W, Qi O. Polymorphisms in the kinesin-like factor 1 gene and risk of epithelial ovarian cancer in eastern Chinese women. Tumour Biol. 2015;36(9):6999–727.

34. Yeh IT, Lenci RE, Qin Y, Buddavarapu K, Ligon AH, Leteune E, Do Cao C, Cardot-Bauters C, Pigny P, Dahia PL. A germline mutation of the KIF18B gene on 1p36 in a family with neural and nonneural tumors. Hum Genet. 1980;64(3):279–85.

35. Schlisio S, Kenchappa RS, Vredeveld LC, George RE, Stewart R, Greulich H, Shahriari K, Nguyen NV, Pigny P, Dahia PL, et al. The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. Gene Dev. 2008;22(7):984–99.

36. Lichenstein H, Xiong Z, Zhu Y, Hu C, Fawzi S, Wolf GM. KIF1Bbeta functions as a haploinsufficient tumor suppressor gene mapped to chromosome 1p56.2 by inducing apoptotic cell death. J Biol Chem. 2008;283(36):24246–34.

37. Tordella L, Koch S, Salter V, Pagotto A, Doonidee JB, Feller SM, Ratnayka I, Leung CM, Yuen H, et al. ASPP2 suppresses squamous cell carcinoma via RelA/p65-mediated repression of p63. Proc Natl Acad Sci U S A. 2004;101(5):1573–8.

38. Li X, Xu B, Mosan MS, Zhao Y, Su P, Hafty GB, Shao C, Yang Q. S38P functions as a tumor suppressor in breast cancer via the inhibition of NF-kappaB through miR-146a. Carcinogenesis. 2012;33(12):2593–600.

39. Tordella L, Koch S, Salter V, Pagotto A, Doonidee JB, Feller SM, Ratnayka I, Leung CM, Yuen H, et al. ASPP2 suppresses squamous cell carcinoma via RelA/p65-mediated repression of p63. Proc Natl Acad Sci U S A. 2004;101(5):1573–8.

40. Schlisio S, Kenchappa RS, Vredeveld LC, George RE, Stewart R, Greulich H, Shahriari K, Nguyen NV, Pigny P, Dahia PL, et al. The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. Gene Dev. 2008;22(7):984–99.

41. Lichenstein H, Xiong Z, Zhu Y, Hu C, Fawzi S, Wolf GM. KIF1Bbeta functions as a haploinsufficient tumor suppressor gene mapped to chromosome 1p56.2 by inducing apoptotic cell death. J Biol Chem. 2008;283(36):24246–34.

42. Schlisio S, Kenchappa RS, Vredeveld LC, George RE, Stewart R, Greulich H, Shahriari K, Nguyen NV, Pigny P, Dahia PL, et al. The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. Gene Dev. 2008;22(7):984–99.

43. Tordella L, Koch S, Salter V, Pagotto A, Doonidee JB, Feller SM, Ratnayka I, Leung CM, Yuen H, et al. ASPP2 suppresses squamous cell carcinoma via RelA/p65-mediated repression of p63. Proc Natl Acad Sci U S A. 2004;101(5):1573–8.

44. Schlisio S, Kenchappa RS, Vredeveld LC, George RE, Stewart R, Greulich H, Shahriari K, Nguyen NV, Pigny P, Dahia PL, et al. The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. Gene Dev. 2008;22(7):984–99.

45. Tordella L, Koch S, Salter V, Pagotto A, Doonidee JB, Feller SM, Ratnayka I, Leung CM, Yuen H, et al. ASPP2 suppresses squamous cell carcinoma via RelA/p65-mediated repression of p63. Proc Natl Acad Sci U S A. 2004;101(5):1573–8.

46. Schlisio S, Kenchappa RS, Vredeveld LC, George RE, Stewart R, Greulich H, Shahriari K, Nguyen NV, Pigny P, Dahia PL, et al. The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. Gene Dev. 2008;22(7):984–99.

47. Tordella L, Koch S, Salter V, Pagotto A, Doonidee JB, Feller SM, Ratnayka I, Leung CM, Yuen H, et al. ASPP2 suppresses squamous cell carcinoma via RelA/p65-mediated repression of p63. Proc Natl Acad Sci U S A. 2004;101(5):1573–8.

48. Schlisio S, Kenchappa RS, Vredeveld LC, George RE, Stewart R, Greulich H, Shahriari K, Nguyen NV, Pigny P, Dahia PL, et al. The kinesin KIF1Bbeta acts downstream from EglN3 to induce apoptosis and is a potential 1p36 tumor suppressor. Gene Dev. 2008;22(7):984–99.

49. Tordella L, Koch S, Salter V, Pagotto A, Doonidee JB, Feller SM, Ratnayka I, Leung CM, Yuen H, et al. ASPP2 suppresses squamous cell carcinoma via RelA/p65-mediated repression of p63. Proc Natl Acad Sci U S A. 2004;101(5):1573–8.
expressed in extraocular muscles and muscle spindles. J Physiol. 2010;588(Pt 2):353–64.

72. Martins-de-Souza D, Guest PC, Mann DM, Roeber S, Rahmoune H, Bauder C, Kretzschmar H, Volk B, Baberie A, Bahn S. Proteomic analysis identifies dysfunction in cellular transport, energy, and protein metabolism in different brain regions of atypical frontotemporal lobar degeneration. J Proteome Res. 2012;11(4):2533–43.

73. Yu XJ, Yang MJ, Zhou B, Wang GZ, Huang YC, Wu LC, Cheng X, Wen ZS, Huang JY, Zhang YD, et al. Characterization of somatic mutations in air pollution-related lung cancer. EBioMedicine. 2015;2(6):583–90.

74. Alvarado DM, Buchan JC, Gurnett CA, Dobbs MB. Exome sequencing identifies an MYH3 mutation in a family with distal arthrogryposis type 1. J Bone Joint Surg Am. 2011;93(11):1045–50.

75. Toydemir RM, Rutherford A, Whitby FG, Jorde LB, Carey JC, Bamshad MJ. Mutations in embryonic myosin heavy chain (MYH3) cause freeman-Sheldon syndrome and Sheldon-Hall syndrome. Nat Genet. 2006;38(5):561–5.

76. Chong JX, Burrage LC, Beck AE, Marvin CT, McMillin MJ, Shively KM, Harrell TM, Buckingham KJ, Bacino CA, Jain M, et al. Autosomal-dominant multiple pterygium syndrome is caused by mutations in MYH3. Am J Hum Genet. 2015;96(5):841–9.

77. Carapito R, Goldenberg A, Paul N, Pichot A, David A, Hamel A, Dumant-Forest C, Leroux J, Ory B, Isidor B, et al. Protein-altering MYH3 variants are associated with a spectrum of phenotypes extending to spondylocarpotarsal synostosis syndrome. Eur J Hum Genet. 2016;24(12):1746–51.

78. Luke MW, O’Meara ES, Rowland CM, Shifman D, Bare LA, Arellano AR, Longstreth WT Jr, Lumley T, Rice K, Tracy RP, et al. Gene variants associated with ischemic stroke: the cardiovascular health study. Stroke. 2009;40(2):363–8.