

ГЕНЕТИКА / GENETICS

DOI: <https://doi.org/10.18454/jbg.2023.22.7>**ЭКСПРЕССИОННАЯ МОДЕЛЬ ПРЕДСКАЗАНИЯ МУТАЦИЙ В ГЕНАХ *SDHx* ПРИ КАРОТИДНЫХ ПАРААНГЛИОМАХ**

Научная статья

Снежкина А.В.^{1,*}, Кобеляцкая А.², Аюпова А.Ф.³, Маркова Д.⁴, Федорова М.С.⁵, Павлов В.С.⁶, Головюк А.Л.⁷, Калинин Д.В.⁸, Кудрявцева А.В.⁹¹ ORCID : 0000-0002-4421-4364;⁸ ORCID : 0000-0001-6247-9481;^{1, 2, 3, 4, 5, 6, 9} Институт молекулярной биологии имени В.А. Энгельгардта, Москва, Российская Федерация^{7, 8} Национальный медицинский исследовательский центр хирургии имени А.В. Вишневского, Москва, Российская Федерация

* Корреспондирующий автор (leftger[at]rambler.ru)

Аннотация

Каротидная параганглиома (КПГ) – это редкое злокачественное новообразование, которое формируется из клеток парасимпатического параганглия, располагающегося в области бифуркации сонной артерии. КПГ преимущественно характеризуются медленным ростом и бессимптомным течением, однако их сложная анатомическая локализация и высокая степень васкуляризации представляет серьезные проблемы для терапии. Около 40% всех параганглиом головы и шеи ассоциированы с мутациями в генах *SDHx*. Мутации в этих генах могут быть связаны с наследственными синдромами, а также риском развития метастатических и мультифокальных форм опухолей. Диагностика *SDHx* мутаций имеет важное значение для ведения пациентов с параганглиомами. В этом исследовании, на основе анализа транскриптомных профилей и данных о *SDHx* мутациях в образцах КПГ мы провели поиск модели предсказания мутационного статуса этих генов. Поиск модели проводился с использованием полносвязной нейронной сети. В результате создана модель на базе экспрессии пяти генов – *GFPT2*, *GABRG2*, *KCNA6*, *PCDH9* и *CACNG3* со статистическим показателем AUC 0,92. Таким образом, полученная модель может быть перспективным инструментом для определения мутаций в генах *SDHx* при КПГ.

Ключевые слова: параганглиомы головы и шеи, каротидная параганглиома, *SDHx* мутации, экспрессия генов, предсказательная модель.

AN EXPRESSION MODEL OF PREDICTION OF MUTATIONS IN *SDHx* GENES IN CAROTID PARAGANGLIOMAS

Research article

Snezhkina A.V.^{1,*}, Kobelyatskaya A.², Ayupova A.F.³, Markova D.⁴, Fedorova M.S.⁵, Pavlov V.S.⁶, Golovyuk A.L.⁷, Kalinin D.V.⁸, Kudryavtseva A.V.⁹¹ ORCID : 0000-0002-4421-4364;⁸ ORCID : 0000-0001-6247-9481;^{1, 2, 3, 4, 5, 6, 9} Institute of Molecular Biology named after V.A. Engelhardt, Moscow, Russian Federation^{7, 8} National Medical Research Center of Surgery named after A.V. Vishnevsky, Moscow, Russian Federation

* Corresponding author (leftger[at]rambler.ru)

Abstract

Carotid paraganglioma (CPG) is a rare malignant neoplasm that forms from the cells of the parasympathetic paraganglion located at the carotid bifurcation. CPGs are predominantly characterized by slow growth and an asymptomatic course, but their complex anatomical localization and high degree of vascularization present serious problems for therapy. About 40% of all head and neck paragangliomas are associated with mutations in *SDHx* genes. Mutations in these genes may be associated with inherited syndromes as well as the risk of developing metastatic and multifocal forms of tumours. Diagnosis of *SDHx* mutations is important for the management of patients with paragangliomas. In this study, we searched for a model to predict the mutational status of these genes based on analyses of transcriptomic profiles and *SDHx* mutation data in CPG samples. The model search was conducted using a full-link neural network. As a result, a model based on the expression of five genes – *GFPT2*, *GABRG2*, *KCNA6*, *PCDH9* and *CACNG3* with a statistical AUC of 0.92 was created. Thus, the obtained model can be a promising tool for determining mutations in *SDHx* genes in CPG.

Keywords: head and neck paragangliomas, carotid paraganglioma, *SDHx* mutations, gene expression, prediction model.

Введение

Параганглиомы (ПГ) головы и шеи образуются из параганглиев парасимпатической нервной системы с частотой 1 на 100 000 [1]. Наиболее частые места локализации этих опухолей – область бифуркации сонной артерии (каротидные ПГ), область вдоль блуждающего нерва (вагальные ПГ) и среднее ухо (югулотимпанальные ПГ) [2]. Каротидные параганглиомы (КПГ) составляют 60% всех ПГ головы и шеи. Развитие КПГ на начальных этапах часто протекает бессимптомно и проявляется в виде медленно растущего новообразования. Диагностика этих опухолей осуществляется преимущественно с использованием таких инструментальных методов как компьютерная томография, магнитно-резонансная томография и ангиография, однако окончательный диагноз устанавливается только после

операции гистопатологическими и иммуногистохимическими методами. Хирургическое удаление признано эффективным и практически единственным способом лечения КПП, но связано с высоким риском осложнений из-за длительного времени операции, значительной кровопотери и паралича черепно-мозговых нервов [3].

До 25% КПП являются мультифокальными (развиваются с обеих сторон шеи и/или одновременно с другими ПГ), в 4-6% случаев наблюдается метастазирование [4]. Важно отметить, что все ПГ относятся к злокачественным новообразованиям с переменным потенциалом метастазирования, который сложно предсказать [2]. На сегодняшний день эффективные биомаркеры, предсказывающие агрессивный фенотип ПГ, не известны. Тем не менее генетические исследования последних лет показали, что наследственные мутации в генах *SDHB* с высоким риском развития метастатических ПГ головы и шеи, в то время как мутации в генах *SDHD* ассоциированы с мультифокальными формами опухолей [4]. Мутации в генах *SDHx* (*SDHD*, *SDHB*, *SDHC* и *SDHA*) являются самыми частыми (около 40%) среди наследственных ПГ головы и шеи и могут быть связаны с развитием синдромов параганглиом [5]. Эти гены кодируют четыре субъединицы сукцинатдегидрогеназы, которая является важным ферментом цикла Кребса и окислительного фосфорилирования. Предполагают, что одним из ключевых механизмов развития *SDHx*-ассоциированных ПГ является накопление онкометаболита сукцината, приводящее к активации пути псевдогипоксии, компоненты которого могут быть потенциальными мишенями для противоопухолевой терапии [6]. Высокая частота мутаций в генах *SDHx*, связь с развитием наследственных синдромов параганглиом и агрессивным поведением опухоли обуславливают необходимость генетического консультирования и тестирование всех пациентов с ПГ, которое также рекомендовано Международным Обществом Эндокринологов [7].

В данной работе, на основе ранее полученных данных секвенирования транскриптома репрезентативной выборки КПП, мы разработали модель для предсказания мутаций в генах *SDHx*. Это исследование важно для развития методов диагностики и оценки прогноза течения заболевания.

Методы и материалы исследования

2.1. Набор данных

Для разработки модели использовались данные транскриптомного секвенирования 77 образцов опухолевых тканей КПП (данные не опубликованы). Подготовка транскриптомных библиотек проводилась с помощью набора TruSeq Stranded Total RNA Library Prep Kit with Ribo-Zero Human/Mouse/Rat фирмы Illumina (США). Секвенирование выполнено на приборе NextSeq 500 System (Illumina) в режиме одноконцевых прочтений длиной 76 нуклеотидов. Первичный биоинформатический анализ выполнялся по стандартной схеме, включающей такие этапы как контроль качества прочтений, обрезка и фильтрация ридов, удаление адаптеров, картирование последовательностей на референсный геном человека (GRCh38.p12), агрегация отчётов, подсчёт числа ридов на ген и импорт данных в среду R. Анализ дифференциальной экспрессии выполнялся с использованием пакета edgeR.

2.2. Поиск модели

Поиск модели проводился с использованием полносвязной нейронной сети (FCNN) и библиотеки Keras. Исследуемая выборка образцов КПП случайным образом была разделена на обучающий и тестовый наборы данных в соотношении 2:1 с сохранением пропорций сравниваемых групп. В качестве входных данных использовали значение CPM, отражающее нормализованный уровень экспрессии генов. Для построения FCNN были задействованы 5 предикторов в качестве входного слоя, затем построены 7 скрытых слоев и выходной слой, состоящий из двух нейронов. Шаг обучения алгоритма оптимизации составил 0,003. Метриками качества модели во время обучения выбраны чувствительность, специфичность, точность и площадь под ROC-кривой (AUC). Для предотвращения переобучения, модель после прохождения каждой эпохи подлежала сохранению. Количество эпох, дающее максимальное качество модели, считалось оптимальным (но не более 200).

Основные результаты

На предварительном этапе перед поиском модели мы провели анализ данных дифференциальной экспрессии (ДЭ) генов между двумя группами образцов КПП – опухоли с мутацией в генах *SDHx* (n=45) и без мутаций в этих генах (n=32). В результате идентифицировано 79 ДЭ генов ($\text{LogFC} \geq 1/\leq -1$ при $p \leq 0,05$ по критерию Манна-Уитни и F-тесту квази-правдоподобия), из которых 26 генов характеризовались повышенной экспрессией, а 53 гена имели пониженную экспрессию. На основе идентифицированных ДЭ генов был проведен поиск экспрессионной модели, предсказывающей мутации в генах *SDHx* при КПП. Лучшими метриками качества характеризовалась модель экспрессии пяти генов: *GFPT2*, *GABRG2*, *KCNA6*, *PCDH9* и *CACNG3*. Чувствительность этой модели для обучающего набора данных составила 0,97, специфичность – 0,86, точность – 0,92, AUC – 0,915 (см. рис. 1). Метрики качества для тестового набора данных составили: чувствительность – 0,93, специфичность – 0,90, точность – 0,92, AUC – 0,92 (см. рис. 1).

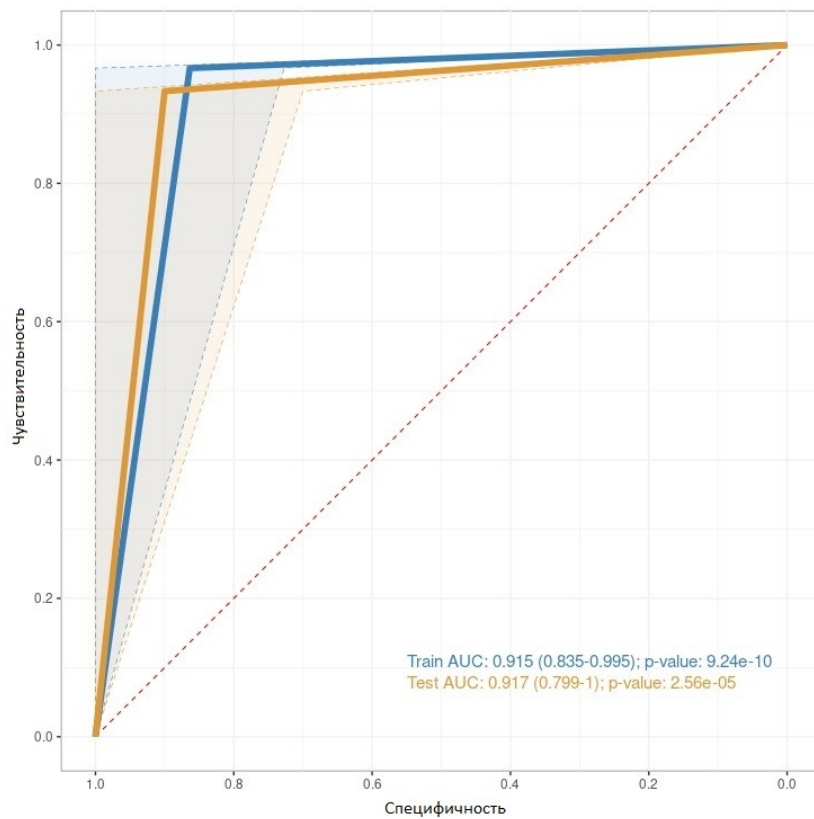


Рисунок 1 - ROC-анализ модели на основе экспрессии генов *GFPT2*, *GABRG2*, *KCNA6*, *PCDH9* и *CACNG3*, предсказывающей наличие *SDHx* мутаций при каротидных парагангиомах
 DOI: <https://doi.org/10.18454/jbg.2023.22.7.1>

Примечание: синяя кривая – обучающий набор данных, оранжевая кривая – тестовый набор данных

Уровень экспрессии генов, вошедших в модель, различался между *SDHx*-мутантными и не мутантными образцами не более 4 раз. Изменение экспрессии генов между исследуемыми группами образцов представлено в таблице 1.

Таблица 1 - Гены экспрессионной модели, ассоциированные с *SDHx* мутациями при КПП

DOI: <https://doi.org/10.18454/jbg.2023.22.7.2>

Ген	Кодируемый белок	LogFC	Направление изменения	Значение <i>p</i> (тест Манна-Уитни)
<i>GFPT2</i>	Глутамин-фруктозо-6-фосфатаминотрансфераза	1,38	↑	0,0004
<i>CACNG3</i>	Субъединица гамма 3 регуляторного белка трансмембранного АМРА-рецептора I типа	1,21	↑	0,0002
<i>PCDH9</i>	Протокадгерин 9	1,04	↓	0,00002
<i>GABRG2</i>	Субъединица гамма 2 рецептора гамма-аминомасляной кислоты A	1,92	↓	0,0002

KCNA6	Белок Kv1.6 подсемейства калиевых потенциалзависимых каналов	1,03	↓	0,0002
-------	--	------	---	--------

Примечание: ↑ - повышение экспрессии, ↓ - понижение экспрессии

Обсуждение

Идентификация патогенных мутаций в генах, ассоциированных с канцерогенезом, является важным этапом в диагностике и лечении. Этот вопрос широко изучается как в теоретической, так и в прикладной биоимедине. В связи с наступившей эрой омиксных данных у исследователей появилась возможность проводить более глубокий и разносторонний анализ молекулярно-генетических характеристик опухолей. Эти данные также позволяют разрабатывать новые подходы с высокой предсказательной силой для обнаружения драйверных нарушений. В этой работе, для создания модели предсказания мутаций в генах *SDHx* мы использовали данные об экспрессии генов и данные о наличии мутаций в исследуемых образцах. В результате разработана модель на основе экспрессии генов *GFPT2*, *GABRG2*, *KCNA6*, *PCDH9* и *CACNG3*, демонстрирующая высокое значение AUC – 0,92, указывающее на качественный алгоритм классификации.

В этиологии ПГ выделяют спорадические и семейные случаи, при этом до 1/3 этих опухолей могут быть ассоциированы с семейными синдромами параганглиом [8]. *SDHx* мутации вызывают развитие наследственных синдромов параганглиом типов 1 (*SDHD*), 3 (*SDHC*), 4 (*SDHB*) и 5 (*SDHA*). Мутации в гене *SDHA* могут приводить к развитию синдрома Ли, характеризующегося подострой некротизирующей энцефаломиелопатией [9]. Соматические мутации в генах *SDHx* также выявляются в контексте спорадических ПГ [10]. Наследственные ПГ имеют более высокий риск развития метастатических и мультифокальных форм опухолей, тогда как спорадические ПГ характеризуются преимущественно не агрессивным фенотипом и односторонней локализацией [11]. Таким образом, *SDHx* мутации имеют высокую частоту среди пациентов с ПГ и являются первостепенной генетической характеристикой этих опухолей, определяющей наследуемость заболевания и его клиническое течение.

В настоящее время для генетического тестирования мутаций во всех четырех генах *SDHx* можно использовать метод полноэкзомного секвенирования или таргетную панель для секвенирования. Оба подхода имеют высокую стоимость и сроки их выполнения могут занимать длительное время, поскольку ячейка для запуска секвенатора имеет определенную емкость и, как правило, с использованием одной ячейки одновременно анализируют большое число образцов, которые собирают от разных пациентов. Поиск соматических мутаций с помощью таргетной панели на гены *SDHx* используют различные лаборатории при клиниках в США: Логан Хелт, Мэйо, Асанте и др. Компания Illumina предлагает коммерческую панель TruSight Oncology 500, в которую включены 500 различных генов, ассоциированных с опухолями, включая гены *SDHx*. Также можно проводить анализ индивидуальных генов *SDHx* с использованием таргетного высокопроизводительного секвенирования или секвенирования по Сэнгеру, пока не удастся выявить мутацию или не будут проверены все четыре гена. Данный подход может быть самым экономичным, если провести предварительный анализ опухолей на наличие мутаций в генах *SDHx* с использованием экспрессионной модели. Это позволит исключить около 60% случаев опухолей, не связанных с мутациями в этих генах. Кроме того, модель можно использовать для классификации КПП на *SDHx*-мутантные и не мутантные опухоли для изучения молекулярно-генетических особенностей развития заболевания.

Заключение

Полученная модель является перспективным инструментом для определения мутаций в генах *SDHx* при КПП для научно-исследовательских целей с использованием полнотранскриптомных данных. Кроме того, результаты этого исследования могут в дальнейшем получить развитие в клинике для более экономичного скрининга на носительство *SDHx* мутаций у пациентов с КПП методом количественной ПЦР.

Финансирование

Исследование выполнено при финансовой поддержке гранта РФФ № 21-14-00353. Работа выполнена с использованием оборудования ЦКП «Геном» ИМБ РАН (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование. Но рецензент или автор статьи предпочли не публиковать рецензию к этой статье в открытом доступе. Рецензия может быть предоставлена компетентным органам по запросу.

Funding

This study was funded by the Russian Science Foundation, grant number 21-14-00353. The work was performed using the equipment of the CCP "Genome" of the IMB RAS (http://www.eimb.ru/rus/ckp/ccu_genome_c.php).

Conflict of Interest

None declared.

Review

All articles are peer-reviewed. But the reviewer or the author of the article chose not to publish a review of this article in the public domain. The review can be provided to the competent authorities upon request.

Список литературы на английском языке / References in English

1. Zanoletti E. Vagal Paraganglioma / E. Zanoletti, A. Mazzoni // Skull Base. — 2006 — Vol. 16, no. 3 — P. 161-7.

2. El-Naggar A.K. WHO Classification of Head and Neck Tumours, 4th Edition / A.K. El-Naggar, J.K.C. Chan, J.R. Grandis [et al.] // International Agency for Research on Cancer. — 2017. — Vol. 9. — P. 347.
3. Shiga K. Challenges of Surgical Resection of Carotid Body Tumors — Multiple Feeding Arteries and Preoperative Embolization / K. Shiga, K. Katagiri, A. Ikeda, [et al.] // Anticancer Research. — 2022. — Vol. 42, No. 2. — P. 645-652.
4. Snezhkina A. Potential Biomarkers of Metastasizing Paragangliomas and Pheochromocytomas / A. Snezhkina, V. Pavlov, A. Dmitriev [et al.] // Life (Basel). — 2021. — Vol. 11, No. 11. — P. 1179.
5. Snezhkina A.V. Mutation Frequency in Main Susceptibility Genes Among Patients With Head and Neck Paragangliomas / A.V. Snezhkina, M.S. Fedorova, V.S. Pavlov [et al.] // Frontiers in Genetics. — 2020. — Vol. 11. — P. 614908.
6. Zhikrivetskaya S.O. Molecular Markers of Paragangliomas/Pheochromocytomas / S.O. Zhikrivetskaya, A.V. Snezhkina, A.R. Zaretsky [et al.] // Oncotarget. — 2017. — Vol. 8, No. 15. — P. 25756-25782.
7. Lenders J.W.M. Pheochromocytoma and Paraganglioma: an Endocrine Society Clinical Practice Guideline / J.W.M. Lenders, Q-Y. Duh, G. Eisenhofer [et al.] // Journal of Clinical Endocrinology & Metabolism. — 2014. — Vol. 99, No. 6. — P. 1915-42.
8. Joynt K.E. Paragangliomas: Etiology, Presentation, and Management / K.E. Joynt, J.J. Moslehi, K.L. Baughman // Cardiology in Review. — 2009. — Vol. 17, No. 4. — P. 159-64.
9. Horváth R. Leigh Syndrome Caused by Mutations in the Flavoprotein (Fp) Subunit of Succinate Dehydrogenase (SDHA) / R. Horváth, A. Abicht, E. Holinski-Feder [et al.] // Journal of Neurology, Neurosurgery & Psychiatry. — 2006. — Vol. 77, No. 1. — P. 74-6.
10. Savvateeva M. Somatic Mutation Profiling in Head and Neck Paragangliomas / M. Savvateeva, A. Kudryavtseva, E. Lukyanova [et al.] // Journal of Clinical Endocrinology & Metabolism. — 2022. — Vol. 107, No. 7. — P. 1833-1842.
11. Neumann H.P.H. Distinct Clinical Features of Paraganglioma Syndromes Associated with SDHB and SDHD Gene Mutations / H.P.H. Neumann, C. Pawlu, M. Peczkowska [et al.] // JAMA. — 2004. — Vol. 292, No. 8. — P. 943-51.