

DOI: 10.25205/978-5-4437-1691-6-22

**ГЕНЕТИЧЕСКИЕ ДЕТЕРМИНАНТЫ ЛЕГОЧНОЙ АРТЕРИАЛЬНОЙ ГИПЕРТЕНЗИИ:
РЕЗУЛЬТАТЫ ПОЛНОГЕНОМНОГО СЕКВЕНИРОВАНИЯ И АНАЛИЗА SNP У 74 ПАЦИЕНТОВ****GENETIC DETERMINANTS OF PULMONARY ARTERIAL HYPERTENSION:
RESULTS OF WHOLE GENOME SEQUENCING AND SNP ANALYSIS IN 74 PATIENTS**

В. И. Замятин¹, Г. С. Охрименко¹, И. И. Боровикова², Е. В. Зобова³,
Д. А. Никулин³, Е. Н. Данковцева^{3,4}, Д. А. Затеишчиков^{3,4}, М. С. Попцова¹

¹Национальный исследовательский университет «Высшая школа экономики», Москва

²Московский государственный университет им. М. В. Ломоносова

³Центральная государственная медицинская академия УД Президента РФ, Москва

⁴Городская клиническая больница № 29 им. Н. Э. Баумана, Москва

V. I. Zamyatin¹, G. S. Okhrimenko¹, I. I. Borovikova², E. V. Zobova³,
D. A. Nikulin³, E. N. Dankovtseva^{3,4}, D. A. Zateyshchikov^{3,4}, M. S. Poptsova¹

¹National Research University Higher School of Economics, Moscow

²Lomonosov Moscow State University

³Central State Medical Academy of the Presidential Administration of the Russian Federation, Moscow

⁴City Clinical Hospital No. 29 named after N. E. Bauman, Moscow

✉ vzamyatin@hse.ru

Аннотация

В исследовании проведено WGS у 74 пациентов с ЛАГ с использованием платформы MGI DNBSEQ. Для анализа использовали инструменты bwa-mem, MarkDuplicates, GATK HaplotypeCaller и VEP с dbNSFP и SpliceAI. Выявлены 6 известных (5 патогенных, 1 вероятно патогенный) и 1 *de novo* (вероятно патогенный) SNP. Выявленные SNP создают стоп-сигналы или изменяют сплайсинг гена BМPR2, что связано с патогенезом ЛАГ.

Abstract

The study performed WGS in 74 patients with PAH using the MGI DNBSEQ platform. The bwa-mem, MarkDuplicates, GATK HaplotypeCaller and VEP tools with dbNSFP and SpliceAI were used for analysis. Six known (5 pathogenic, 1 likely pathogenic) and 1 *de novo* (likely pathogenic) SNPs were identified. The identified SNPs create stop signals or alter splicing of the BМPR2 gene, which is associated with the pathogenesis of PAH.

Введение и цель. Легочная артериальная гипертензия (ЛАГ) представляет собой серьезное и прогрессирующее заболевание, характеризующееся повышением давления в легочных артериях, что приводит к правожелудочковой недостаточности и высокой смертности. Несмотря на значительные достижения в понимании патофизиологических механизмов ЛАГ, остаются вопросы о генетических факторах, влияющих на развитие и прогрессирование данного заболевания. Современные методы молекулярной генетики, такие как полногеномное секвенирование (WGS) и анализ однонуклеотидных полиморфизмов (SNP calling), предоставляют уникальные возможности для выявления генетических вариантов, ассоциированных с ЛАГ, и оценки их патогенности.

Целью данного исследования является проведение полногеномного секвенирования и анализа однонуклеотидных полиморфизмов у пациентов с легочной артериальной гипертензией для выявления потенциально патогенных генетических вариантов и оценки их роли в патогенезе заболевания.

Методы. Популяция исследования. В исследование включено 74 пациента с диагнозом «легочная артериальная гипертензия» (ЛАГ). Все пациенты дали информированное согласие на участие в исследовании и предоставление образцов для генетического анализа.

Полногеномное секвенирование. Для проведения полногеномного секвенирования (WGS) использовалась платформа MGI DNBSEQ. Процедура включала экстракцию ДНК из периферической крови пациентов, подготовку библиотек и проведение секвенирования с покрытием, обеспечивающим высокую точность и полноту данных.

Анализ данных. Выравнивание и предобработка данных: сырые данные секвенирования выравнивались на референсный геном человека (GRCh38) с использованием инструмента bwa-mem; для удаления дубликатов прочтений применялся инструмент MarkDuplicates из пакета GATK. Поиск однонуклеотидных полиморфизмов (SNP calling) выполнялся с использованием GATK HaplotypeCaller; патогенность выявленных вариантов оценивалась с помощью аннотационного инструмента VEP (Variant Effect Predictor) с использованием баз данных dbNSFP и SpliceAI; классификация патогенности вариантов проводилась в соответствии с рекомендациями ACMG (American College of Medical Genetics and Genomics) с помощью InterVar. Анализ структурных вариаций (SV): выявление структурных вариаций проводилось с использованием инструмента Gridss; оценка патогенности структурных вариаций выполнялась с помощью CADD-SV (Combined Annotation-Dependent Depletion for Structural Variants).

Результаты. В результате биоинформатического варианта были найдены 6 (5 патогенных, 1 вероятно патогенный по ACMG) известных SNP и 1 (вероятно патогенный по ACMG) рассматриваемый нами *de novo* (табл. 1).

Таблица 1

Патогенные и вероятно патогенные варианты, обнаруженные у пациентов с ЛАГ

Chr	Pos	Ref	Alt	Consequence	Gene	Databases	Patient count	ACMG
2	202532663	C	T	stop_gained	BMP2	rs946132834&CM023609&CO SV65810324	1	Pathogenic
2	202530820	C	T	stop_gained	BMP2	rs137852751&CM010163&CO SV65808047	1	Pathogenic
2	202520195	C	T	stop_gained	BMP2	rs1060502581&CD1513986&CM 043449&COSV65807609	1	Pathogenic
2	202467648	A	G	Missense variant	BMP2	rs863223426&CM106873&CO SV101001618	2	Likely pathogenic
17	61482932	C	T	stop_gained	TBX4	rs1603256040&COSV53610400	1	Pathogenic
15	40032178	C	T	stop_gained	EIF2AK4	rs372637398&COSV99139556	1	Pathogenic
7	30922075	G	T	Missense variant	AQP1		1	Likely pathogenic

Варианты *rs946132834*, *rs137852751*, *rs1060502581* создают преждевременный трансляционный стоп-сигнал, они обнаружены у людей с семейной первичной легочной гипертензией (PMID: 12139571). Известно, что варианты с потерей функции в BMP2 являются патогенными (PMID: 16429395, 11015450, 30578397). *rs863223426* — это миссенс-вариант, наблюдался у лиц с легочной артериальной гипертензией (PMID: 19555857, 20534176, 21737554, 29843651, 31727138, 32581362). Он заменяет аспарагин, который является нейтральным и полярным, на аналогичный серин, в кодоне 126 белка BMP2 (p.Asn126Ser). В настоящее время данные указывают на то, что вариант является патогенным, но для окончательного доказательства этого необходимы дополнительные данные. Вариант *p.G132C (chr7:30922075 G>T)* соответствует критериям ACMG: PM1, PM2, PP2, PP3, вследствие чего обозначен как вероятно патогенный.

Вариант *rs75747031 (g.202533253A>G)* присутствует у 4 пациентов (табл. 2) и с высокой вероятностью изменяет процесс сплайсинга в гене BMP2 — создает новый акцепторный сайт, а значит, может являться патогенным.

Таблица 2

Сплайс-вариант, обнаруженный у 4 пациентов с ЛАГ

Chr	Pos	Ref	Alt	Databases	Gene	SpliceAI acceptor gain score	Patient count
chr2	202533253	A	G	rs75747031	BMP2	0.57	4