# ОПРЕДЕЛЕНИЕ ДОЛИ ПЛОДОВОЙ ДНК В ПЛАЗМЕ КРОВИ БЕРЕМЕННОЙ ЖЕНЩИНЫ С ПОМОЩЬЮ ВЫСОКОПРОИЗВОДИТЕЛЬНОГО СЕКВЕНИРОВАНИЯ НАБОРА ЧАСТОТНЫХ ОДНОНУКЛЕОТИДНЫХ ПОЛИМОРФИЗМОВ

Е. Шубина<sup>1,2</sup>  $\boxtimes$ , Т. Янкевич<sup>2</sup>, А. Ю. Гольцов<sup>1,2</sup>, И. С. Мукосей<sup>1</sup>, Т. О. Кочеткова<sup>1</sup>, А. А. Быстрицкий<sup>1</sup>, И. Ю. Барков<sup>1</sup>, Н. К. Тетруашвили<sup>1</sup>, Л. В. Ким<sup>1</sup>, Д. Ю. Трофимов<sup>1,2</sup>

Неинвазивный пренатальный ДНК-скрининг (НИПС) анеуплоидий по крови матери применяется для выявления хромосомных анеуплоидий (ХА) с 2011 г. Многочисленные клинические исследования показали, что важным параметром при проведении НИПС является доля плодовой ДНК. Целью работы была разработка тест-системы для оценки доли плодовой ДНК с помощью таргетного секвенирования однонуклеотидных полиморфизмов (SNP). По данным исследований международного проекта НАРМАР были отобраны полиморфизмы с высокой частотой встречаемости гетерозиготного генотипа. Для оценки частоты встречаемости отобранных полиморфизмов в российской популяции использовали 827 образцов ДНК доноров. С целью определения доли плодовой ДНК исследовали 87 образцов плазмы крови беременных женщин. Секвенирование проводили на приборах Ion Proton и Ion S5. В ходе работы были определены частоты встречаемости по данным секвенирования пулированных образцов. Проведено сравнение данных о 53 SNP в 87 отдельных образцах. Медиана разницы, полученой различными способами, составила 3,4%. Результаты определения доли плодовой ДНК с помощью SNP сравнивали с данными по Y-хромосоме, корреляция составила 0,7. Таким образом, разработанную тест-систему можно применять для определения доли плодовой ДНК с помощью SNP вне зависимости от пола плода.

**Ключевые слова:** неинвазивный пренатальный ДНК-скрининг, доля плодовой ДНК, однонуклеотидные полиморфизмы, хромосомные анеуплоидии

**Финансирование:** работа поддержана Министерством образования и науки Российской Федерации (соглашение № 14.607.21.0136, идентификатор проекта RFMEFI60715X0136).

🔀 Для корреспонденции: Шубина Екатерина

ул. Академика Опарина, д. 4, г. Москва, 117997; jekaterina.shubina@gmail.com

Статья получена: 15.04.2018 Статья принята к печати: 18.04.2018

**DOI:** 10.24075/vrgmu.2018.031

## QUANTIFICATION OF FETAL DNA IN THE PLASMA OF PREGNANT WOMEN USING NEXT GENERATION SEQUENCING OF FREQUENT SINGLE NUCLEOTIDE POLYMORPHISMS

Shubina  $J^{1,2} \boxtimes$ , Jankevic  $T^2$ , Goltsov AYu<sup>1,2</sup>, Mukosey IS<sup>1</sup>, Kochetkova TO<sup>1</sup>, Bystritsky AA<sup>1</sup>, Barkov IYu<sup>1</sup>, Tetruashvili NK<sup>1</sup>, Kim LV<sup>1</sup>, Trofimov DYu<sup>1,2</sup>

 $^{\mathrm{1}}$  Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology, Moscow

Introduced into clinical practice in 2011, non-invasive prenatal testing (NIPT) allows detection of chromosomal aneuploidies in the fetus using maternal blood samples. Multiple studies have shown that one of the key factors affecting the result of this test is the fetal DNA fraction. The aim of this work was to develop a method capable of measuring the fetal DNA fraction based on targeted SNP sequencing. We selected polymorphisms with high frequency of heterozygous genotype from the international HapMap database. To estimate the frequency of these polymorphisms in the Russian population, we used 827 DNA donor samples. Fetal DNA fraction was measured in 87 plasma samples of pregnant women. Sequencing was performed on Ion Proton and Ion S5. We determined the frequencies of the studied polymorphisms in the pooled samples and compared the data on 53 SNPs in the pooled and 87 individual samples. The median difference was 3.4%. The correlation between the results obtained by targeted SNP sequencing and Y chromosome read count was 0.7. Thus, the proposed method can be used to estimate the fetal DNA fraction using SNP genotyping regardless of the fetus's sex.

Keywords: non-invasive prenatal testing, fetal DNA fraction, single nucleotide polymorphisms, chromosome aneuploidy

Funding: this work was supported by the Ministry of Education and Science of the Russian Federation (Agreement 14.607.21.0136, Project ID RFMEFI60715X0136).

Correspondence should be addressed: Jekaterina Shubina
Akademika Oparina 4, Moscow, 117997; jekaterina.shubina@gmail.com

Recieved: 15.04.2018 Accepted: 18.04.2018

DOI: 10.24075/brsmu.2018.031

<sup>&</sup>lt;sup>1</sup> Национальный медицинский исследовательский центр акушерства, гинекологии и перинатологии имени В. И. Кулакова, Москва

<sup>&</sup>lt;sup>2</sup> ООО «НПФ ДНК-Технология», Москва

<sup>&</sup>lt;sup>2</sup> DNA-Technology LLC, Moscow

Неинвазивный пренатальный ДНК-скрининг (НИПС) анеуплоидий по крови матери применяется для обнаружения хромосомных анеуплоидий (ХА) с 2011 г. [1]. Многочисленные клинические исследования показали, что важным параметром при проведении НИПС является доля плодовой ДНК [2, 3]. При недостаточном количестве плодовой ДНК снижается чувствительность теста и возможны ложноотрицательные результаты [3].

Долю плодовой ДНК просто определить в случае с беременностью плодом мужского пола. Для этого используют сравнение покрытия У-хромосомы с покрытием остальных аутосом. Если женщина вынашивает плод женского пола, то оценка представленности плодовых фрагментов становится более сложной задачей.

Известные способы определения представленности плодовой ДНК среди всей свободной циркулирующей ДНК (сцДНК) основаны на выявлении и подсчете фрагментов, для которых можно определить их происхождение. Среди них использование фрагментов, специфичных для У-хромосомы, которое возможно только для плода мужского пола; не зависящий от пола анализ по-разному метилированных фрагментов сцДНК [4], однонуклеотидных полиморфизмов [5–7], разницы длин фргаментов плодовой и материнской ДНК [8], распределения фрагментов плодовой ДНК по геному [9–11].

С помощью таргетного секвенирования однонуклеотидных полиморфизмов (SNP) помимо определения доли плодовой ДНК становится возможной генетическая идентификация образца. Кроме того, этот же подход можно использовать для неинвазивного определения отцовства и пренатальной диагностики [12, 13].

Целью данной работы была разработка тест-системы, позволяющей определять долю плодовой ДНК вне зависимости от пола плода с помощью таргетного секвенирования SNP.

#### МАТЕРИАЛЫ И МЕТОДЫ

#### Выбор однонуклеотидных полиморфизмов

По данным широкомасштабных популяционных исследований НАРМАР [14] (http://hapmap.ncbi.nlm.nih.gov/) были отобраны 73 полиморфизма с частотой встречаемости гетерозиготного генотипа, равной 49–51% для представителей популяции СЕU (люди, произошедшие из западной и северной Европы), и 45–55% для представителей африканской (ASW), китайской (CHD, CHB) и японской популяций (JPT); отобранные полиморфизмы находились на расстоянии не менее 20 млн п.н. на 1–12 хромосомах. Для каждого из отобранных полиморфизмов подобрали специфичные

Таблица 1. Результаты секвенирования пулированных образцов

Номер пула	Количество образцов	Количество ридов	Количество ридов/полиморфизм med (q1-q3)		
1	96	2 681 517	12476 (5195–40260)		
2	114	2 002 697	13408 (4724–34127)		
3	96	2 711 707	17753 (7810–48959)		
4	84	3 037 177	20001 (6742–44910)		
5	78	3 884 900	28124 (10032–66108)		
6	96	1 677 467	9808 (2860–24624)		
7	92	1 592 401	8826 (2503–26345)		
8	63	1 759 487	11359 (3629–28146)		
9	57	2 340 385	14355 (3983–38147)		
10	51	2 403 795	16195 (4686–37680)		

праймеры для амплификации выбранного фрагмента. Праймеры отбирали таким образом, чтобы длина ПЦР-продукта была не более 110 п.н.

#### Образцы ДНК и плазмы

Оценку частоты встречаемости полиморфизмов в российской популяции проводили по результатам исследования 827 образцов ДНК, выделенных из крови доноров. Долю плодовой ДНК определяли в 87 образцах плазмы крови беременных женщин (в 45 образцах — у женщин с плодом мужского пола, в 42 — у женщин с плодом женского пола).

### Оценка частоты встречаемости полиморфизмов в российской популяции

Поскольку широкомасштабных популяционных данных о частоте встречаемости различных полиморфизмов в российской популяции нет и для отбора полиморфизмов использовали данные, полученные для других популяций, частота встречаемости выбранных полиморфизмов в исследуемой популяции может отличаться от опубликованных данных. В работе проводили оценку частоты встречаемости в российской популяции с помощью таргетного секвенирования смешанных (пулированных) в равных концентрациях образцов ДНК доноров.

Проводили секвенирование 10 пулов из 827 образцов (51–114 образцов/пул). Перед смешиванием ДНК оценивали ее концентрацию с помощью метода полимеразной цепной реакции (ПЦР) в реальном времени. Для оценки частоты встречаемости полиморфизмов в популяции суммировали частоты встречаемости, полученные для каждого пула, с учетом количества образцов в пуле. Результаты оценки частоты встречаемости полиморфизмов, полученные с помощью пулирования, сравнивали с результатами, полученными секвенированием 87 отдельных образцов.

#### Определение доли плодовой ДНК

Оценку доли плодовой ДНК проводили с помощью секвенирования 53 частотных полиморфизмов, которые были отобраны на основании предварительных данных секвенирования пулированных образцов.

Для определения доли плодовой ДНК использовали полиморфизмы, для которых представленность одного аллеля была более 80%, но менее 99,5%. В этом случае полагали, что мать имеет гомозиготный генотип, а плод гетерозиготный. Долю плодовой ДНК определяли по формуле  $ff = 2 \cdot B / (A + B)$ , где A — более представленный

#### ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ І МЕДИЦИНСКАЯ ГЕНЕТИКА

аллель, В — менее представленный аллель. В качестве значения доли плодовой ДНК использовали медиану значений для всех информативных полиморфизмов. Значение доли плодовой ДНК, определенные с помощью SNP, сравнивали со значением доли плодовой ДНК, определенным по покрытию Y-хромосомы.

#### Секвенирование

Библиотеки из ПЦР-продуктов подготавливали согласно рекомендациям производителя (Thermo Fisher Scientific Inc., США). Секвенирование проводили на приборах Ion Proton и Ion S5 (Thermo Fisher Scientific Inc., США) согласно протоколу производителя.

#### Анализ данных

Первичный анализ данных осуществляли с помощью программного обеспечения Torrent Server 4.4.3. Выравнивание на референсный геном версии GRCh37/hg19 проводили с помощью программного модуля ТМАР (Thermo Fisher Scientific Inc., США), после чего для позиций на геноме, соответствующих выбранным полиморфизмам, подсчитывали количество прочтений для каждого из аллелей (разработанный авторами скрипт с использованием программного пакета

рузат [15]), для анализа отбирали фрагменты только с качеством выравнивания не менее 30 и с длиной не менее 80% от ожидаемой длины фрагмента.

#### РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

#### Секвенирование пулированных образцов

Результаты секвенирования пулированных образцов представлены в табл. 1.

После оценки работы системы в целом и частоты встречаемости отдельных полиморфизмов для дальнейшей работы было отобрано 53 SNP. Исключены полиморфизмы с низкой представленностью. В табл. 2 представлены результаты сравнения частоты встречаемости для 53 SNP по данным секвенирования пулированных образцов и по данным анализа 87 отдельных образцов. Медиана разницы полученной двумя способами частоты встречаемости составила 3,4%.

#### Результаты определения доли плодовой ДНК

Среднее количество полиморфизмов, для которых у матери был гомозиготный генотип, составило 28 (25–32), из них информативных полиморфизмов было 14 (10–18). На

Таблица 2. Сравнение частот встречаемости SNP по данным секвенирования пулированных образцов и по данным для 87 отдельных образцов

номер SNP -	Пулированые образцы (827)		отдельные образцы (87)		номер SNP	Пулированые образцы (827)		отдельные образцы (87)	
	аллель1	аллель2	аллель1	аллель2	помер от	аллель1	аллель2	аллель1	аллель2
rs4846002	0,619	0,381	0,592	0,408	rs1265758	0,599	0,4	0,576	0,424
rs4926658	0,577	0,423	0,529	0,471	rs2143829	0,574	0,426	0,616	0,384
rs9434166	0,576	0,424	0,598	0,402	rs591356	0,438	0,562	0,453	0,547
rs10753750	0,564	0,436	0,586	0,414	rs9373116	0,579	0,421	0,494	0,506
rs1973943	0,532	0,467	0,494	0,506	rs7770051	0,479	0,521	0,407	0,593
rs7597744	0,49	0,51	0,494	0,506	rs16	0,516	0,484	0,506	0,494
rs2121304	0,56	0,44	0,558	0,442	rs12333726	0,564	0,435	0,552	0,448
rs1726025	0,517	0,483	0,517	0,483	rs6958027	0,593	0,407	0,523	0,477
rs11164111	0,513	0,487	0,494	0,506	rs314320	0,724	0,276	0,75	0,25
rs981841	0,49	0,509	0,558	0,442	rs625218	0,597	0,403	0,618	0,382
rs1978346	0,653	0,347	0,698	0,302	rs7005848	0,457	0,542	0,5	0,5
rs9843942	0,565	0,435	0,523	0,477	rs952559	0,547	0,453	0,612	0,388
rs6777416	0,587	0,413	0,616	0,384	rs827584	0,72	0,279	0,7	0,3
rs957303	0,61	0,39	0,593	0,407	rs9987271	0,577	0,422	0,541	0,459
rs1553212	0,514	0,486	0,448	0,552	rs6559467	0,583	0,417	0,612	0,388
rs751834	0,561	0,438	0,663	0,337	rs4132699	0,667	0,332	0,647	0,353
rs6771838	0,645	0,354	0,622	0,378	rs10980011	0,599	0,4	0,571	0,429
rs7696439	0,629	0,37	0,663	0,337	rs2583839	0,603	0,397	0,565	0,435
rs4864809	0,452	0,548	0,517	0,483	rs7904536	0,793	0,207	0,941	0,059
rs17002804	0,484	0,516	0,541	0,459	rs4917915	0,531	0,468	0,453	0,547
rs978373	0,497	0,502	0,459	0,541	rs845085	0,681	0,319	0,724	0,276
rs4621390	0,607	0,393	0,57	0,43	rs4333997	0,522	0,477	0,488	0,512
rs7703985	0,491	0,509	0,442	0,558	rs602991	0,696	0,304	0,747	0,253
rs2962799	0,542	0,456	0,541	0,459	rs2289300	0,716	0,283	0,647	0,353
rs902987	0,52	0,478	0,512	0,488	rs7973612	0,704	0,296	0,765	0,235
rs6859147	0,551	0,449	0,547	0,453	rs7971962	0,606	0,39	0,606	0,394
rs4921132	0,495	0,505	0,494	0,506					

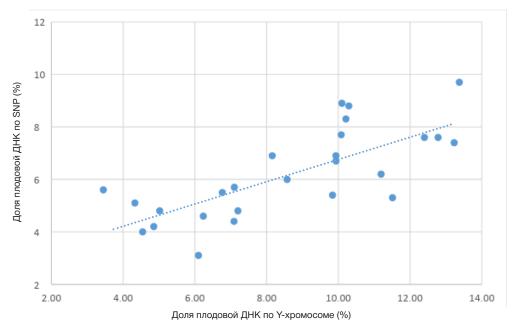


Рис. 1. Сравнение определения доли плодовой ДНК по Y-хромосоме и с помощью таргетного секвенирования частотных полиморфизмов

рис. 1 представлено сравнение результатов определения доли плодовой ДНК с помощью SNP и по данным о покрытии Y-хромосомы, корреляция составила 0,7.

#### ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Проведена оценка встречаемости выбранных полиморфизмов в исследуемой популяции с помощью секвенирования пулированных образцов. Продемонстрировано, что при пулировании и при определении генотипов отдельных образцов получаются сопоставимые результаты. Показана возможность оценки доли плодовой ДНК с помощью таргетного секвенирования небольшого количества частотных полиморфизмов вне зависимости от пола плода. Корреляция результатов определения доли плодовой ДНК с применением таргетного секвенирования и по

У-хромосоме в нашей работе ниже, чем по упоминающимся в литературе данным об использовании сопоставимого количества полиморфизмов для определения доли плодовой ДНК [7], что можно объяснить использованием в работе этих авторов молекулярных индексов и подсчета отдельных молекул.

#### выводы

Предложенный метод можно использовать для оценки частоты встречаемости аллелей в случае частотных полиморфизмов. Помимо определения доли плодовой ДНК, описанный подход может быть применен для идентификации образцов, неинвазивного определения отцовства и неинвазивной диагностики наследственных заболеваний в случае наследования от отца мутации, которой нет у матери.

#### Литература

- Agarwal A, Sayres LC, Cho MK, Cook-Deegan R, Chandrasekharan S. Commercial landscape of noninvasive prenatal testing in the United States. Prenat Diagn. 2013; 33 (6): 521–31.
- Canick JA, Palomaki GE, Kloza EM, Lambert-Messerlian GM, Haddow JE. The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies. Prenat Diagn. 2013; 33 (7): 667–74.
- Сухих Г. Т., Каретникова Н. А., Баранова Е. Е., Шубина Е. С., Коростин Д. О., Екимов А. Н. и др. Неинвазивная пренатальная диагностика анеуплоидий методом высокопроизводительносго секвенирования (NGS) в группе женщин высокого риска. Акушерство и гинекология. 2015; 4: 5–10.
- Nygren AO, Dean J, Jensen TJ, Kruse S, Kwong W, van den Boom D, et al. Quantification of fetal DNA by use of methylationbased DNA discrimination. Clin Chem. 2010 Oct; 56 (10): 1627–35.
- Sparks AB, Struble CA, Wang ET, Song K, Oliphant A. Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: evaluation for trisomy 21 and trisomy 18.
   Am J Obstet Gynecol. 2012 Apr; 206 (4): 319. e1–9.
- Nicolaides KH, Syngelaki A, Gil M, Atanasova V, Markova D. Validation of targeted sequencing of single-nucleotide polymorphisms

- for non-invasive prenatal detection of aneuploidy of chromosomes 13, 18, 21, X, and Y. Prenat Diagn. 2013 Jun; 33 (6): 575–9.
- Song Y, et al. Quantitation of fetal DNA fraction in maternal plasma using circulating single molecule amplification and re-sequencing technology (cSMART). Clin Chim Acta. 2016; 456: 151–6.
- Yu SCY, et al. Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. Proc Natl Acad Sci USA. 2014 Jun; 111 (23): 8583–8.
- Kim SK, et al. Determination of Fetal DNA Fraction from the Plasma of Pregnant Women using Sequence Read Counts. Prenat Diagn. 2015: n/a-n/a.
- Straver R, Oudejans CBM, Sistermans EA, Reinders MJT. Calculating the fetal fraction for Non Invasive Prenatal Testing based on Genome-wide nucleosome profiles. Prenat Diagn. 2016: n/a-n/a.
- van Beek DM, et al. Comparing methods for fetal fraction determination and quality control of NIPT samples. Prenat Diagn. 2017; 37 (8): 769–73.
- 12. Lv W, et al. Noninvasive prenatal testing for Wilson disease by use of circulating single-molecule amplification and resequencing technology (cSMART). Clin Chem. 2015; 61 (1): 172–81.

#### ОРИГИНАЛЬНОЕ ИССЛЕДОВАНИЕ І МЕДИЦИНСКАЯ ГЕНЕТИКА

- Meng M, et al. Noninvasive prenatal testing for autosomal recessive conditions by maternal plasma sequencing in a case of congenital deafness. Genet Med. 2014; 16 (12): 972–6.
- 14. International HapMap Consortium. The International HapMap
- Project. Nature. 2003; 426 (6968): 789-96.
- Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug; 25 (16): 2078–9.

#### References

- Agarwal A, Sayres LC, Cho MK, Cook-Deegan R, Chandrasekharan S. Commercial landscape of noninvasive prenatal testing in the United States. Prenat Diagn. 2013; 33 (6): 521–31.
- Canick JA, Palomaki GE, Kloza EM, Lambert-Messerlian GM, Haddow JE. The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies. Prenat Diagn. 2013; 33 (7): 667–74.
- Sukhikh GT, et al. Noninvasive prenatal diagnosis of aneuploidies by next-generation sequencing (NGS) in a group of high-risk women."Obstetrics and Gynecology. 2015; 4: 5–10.
- Nygren AO, Dean J, Jensen TJ, Kruse S, Kwong W, van den Boom D, et al. Quantification of fetal DNA by use of methylationbased DNA discrimination. Clin Chem. 2010 Oct; 56 (10): 1627–35.
- Sparks AB, Struble CA, Wang ET, Song K, Oliphant A. Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: evaluation for trisomy 21 and trisomy 18.
   Am J Obstet Gynecol. 2012 Apr; 206 (4): 319. e1–9.
- Nicolaides KH, Syngelaki A, Gil M, Atanasova V, Markova D. Validation of targeted sequencing of single-nucleotide polymorphisms for non-invasive prenatal detection of aneuploidy of chromosomes 13, 18, 21, X, and Y. Prenat Diagn. 2013 Jun; 33 (6): 575–9.
- 7. Song Y, et al. Quantitation of fetal DNA fraction in maternal plasma using circulating single molecule amplification and re-sequencing technology (cSMART). Clin Chim Acta. 2016; 456: 151–6.

- Yu SCY, et al. Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. Proc Natl Acad Sci USA. 2014 Jun; 111 (23): 8583–8.
- 9. Kim SK, et al. Determination of Fetal DNA Fraction from the Plasma of Pregnant Women using Sequence Read Counts. Prenat Diagn. 2015: n/a-n/a.
- Straver R, Oudejans CBM, Sistermans EA, Reinders MJT. Calculating the fetal fraction for Non Invasive Prenatal Testing based on Genome-wide nucleosome profiles. Prenat Diagn. 2016: n/a-n/a.
- van Beek DM, et al. Comparing methods for fetal fraction determination and quality control of NIPT samples. Prenat Diagn. 2017; 37 (8): 769–73.
- Lv W, et al. Noninvasive prenatal testing for Wilson disease by use of circulating single-molecule amplification and resequencing technology (cSMART). Clin Chem. 2015; 61 (1): 172–81.
- Meng M, et al. Noninvasive prenatal testing for autosomal recessive conditions by maternal plasma sequencing in a case of congenital deafness. Genet Med. 2014; 16 (12): 972–6.
- 14. International HapMap Consortium. The International HapMap Project. Nature. 2003; 426 (6968): 789–96.
- Li H, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug; 25 (16): 2078–9.