

УДК 575.113.1:575.21:577.24

ОПРЕДЕЛЕНИЕ ХРОНОЛОГИЧЕСКОГО ВОЗРАСТА ЧЕЛОВЕКА ПО БИОЛОГИЧЕСКИМ ОБРАЗЦАМ НА ОСНОВАНИИ АНАЛИЗА МЕТИЛИРОВАНИЯ CpG-ДИНУКЛЕОТИДОВ

© 2021 г. В. А. Лемеш¹ *, В. Н. Кипень¹ **, М. В. Богданова¹, А. А. Буракова¹, А. Г. Булгак², А. В. Байда³, С. А. Брускин⁴, О. В. Зотова², О. И. Добыш¹

¹Институт генетики и цитологии Национальной академии наук Беларуси, Минск, 220072 Республика Беларусь

²Республиканский научно-практический центр “Кардиология”, Минск, 220089 Республика Беларусь

³Белорусская медицинская академия последипломного образования, Минск, 220013 Республика Беларусь

⁴Институт общей генетики им. Н.И. Вавилова Российской академии наук, Москва, 119991 Россия

*e-mail: v.lemesh@igc.by

**e-mail: v.kipen@igc.by

Поступила в редакцию 27.05.2021 г.

После доработки 17.06.2021 г.

Принята к публикации 22.06.2021 г.

На основании биоинформатического и статистического анализов GEO-проектов по определению полногеномного профиля метилирования ДНК человека сформирован перечень из 41 CpG-динуклеотида с высоким прогностическим потенциалом для создания моделей предсказания возраста человека по биологическим образцам. Определен уровень метилирования для 1208 образцов индивидов из Республики Беларусь (275 – кровь, 466 – буккальный эпителий, 467 – сперма), рассчитаны коэффициенты корреляции R и построены математические модели определения возраста человека. Среднее значение точности предсказания возраста по образцам крови с использованием 12 CpG-динуклеотидов составило 3.4 года (для мужчин – 3.3, для женщин – 3.5); по образцам буккального эпителия с использованием 6 CpG-динуклеотидов – 4.6 года (для мужчин – 4.5, для женщин – 4.7); по образцам спермы с использованием 5 CpG-динуклеотидов – 3.0 года. Полученные результаты будут положены в основу при разработке калькуляторов предсказания возраста индивида по биологическим следам для экспертов-криминалистов.

Ключевые слова: криминалистика, определение возраста, кровь, буккальный эпителий, сперма, метилирование ДНК, human methylation 450K BeadChip.

DOI: 10.31857/S0016675821120092

Определение возраста человека по образцам биологических жидкостей и фрагментам тканей играет важную роль в криминалистической практике: способствует ограничению поисков при идентификации останков, сужению круга подозреваемых, тем самым позволяя сэкономить время, которое зачастую является лимитирующим фактором в процессе расследования. В связи со значимостью проблемы обеспечения безопасности граждан Союзного государства (Республика Беларусь и Российская Федерация) разработка методик определения возраста несомненно актуальна и востребована.

В настоящее время в судебно-медицинской экспертизе для оценки возраста человека применяются морфологический (характеристика внешних особенностей зубов или частей скелета) и биохимический (возрастание рацемизации аспарагиновой кислоты в дентине при старении) ме-

тоды. Однако ошибка при использовании этих методов довольно высока (3–10 лет). Кроме того, при минимальном количестве биологического материала или высокой степени его деградации их применение является затруднительным или невозможным. В таких случаях незаменимыми становятся методы, в которых матрицей для исследования выступает наследственный материал клетки.

Сравнение характеристик основных молекулярно-генетических подходов для определения возраста индивида представлено в обзоре А.Д. Золотаренко и др. [1]. Для решения криминалистической задачи по определению хронологического возраста человека наиболее воспроизводимым, линейно-зависимым, чувствительным и специфичным является подход, основанный на выявлении уровня метилирования ДНК в конкретных CpG-динуклеотидах [2–4].

Самыми распространенными биологическими образцами криминалистического характера являются следы крови, буккального эпителия и спермы. В мире предложено значительное количество методик определения возраста индивида по уровню прижизненной химической модификации ДНК [5–19]. Однако эффективность применения указанных методик для определения возраста представителей населения регионов Союзного государства не установлена и требует проработки, поскольку было продемонстрировано, что уровень метилирования CpG-динуклеотидов может различаться в зависимости от этногеографического происхождения индивидов и факторов окружающей среды [20, 21].

Цель данной работы – оценить точность предсказания возраста человека с использованием математических моделей на основании анализа метилирования CpG-динуклеотидов в образцах крови, буккального эпителия и спермы индивидов из Республики Беларусь.

МАТЕРИАЛЫ И МЕТОДЫ

Биологические образцы. Материалом для исследования служили образцы ДНК индивидов из Республики Беларусь. Забор венозной крови в пробирки BD Vacutainer K2E осуществлен у 275 индивидов (мужчины – 119, женщины – 156). Средний возраст добровольцев составил 57.3 ± 16.9 лет (для мужчин – 58.4 ± 16.8 лет, для женщин – 56.4 ± 17.0 лет), диапазон – от 18 до 93 полных лет. Забор буккального эпителия на ватный тампон-зонд осуществлен у 466 индивидов (мужчины – 262, женщины – 204). Средний возраст добровольцев составил 45.9 ± 18.1 лет (для мужчин – 46.7 ± 18.1 лет, для женщин – 44.8 ± 18.2 лет), диапазон – от 9 до 93 полных лет. Образцы спермы от 382 мужчин были получены на основании запроса в Республиканский банк ДНК (Институт генетики и цитологии НАН Беларуси). Дополнительно 85 образцов спермы получены от добровольцев, средний возраст которых составил 34.2 ± 5.2 лет, диапазон – от 22 до 53 полных лет.

Забор биологического материала был осуществлен с использованием процедур, одобренных Комитетом по биоэтике при Институте генетики и цитологии НАН Беларуси. Биологический материал хранили при температуре -20°C в течение 1–2 мес., либо при -80°C при долговременном хранении. ДНК экстрагировали с использованием набора MagMAX™ DNA Multi-Sample Kit (ThermoFisher, США) в соответствии с инструкциями производителя. Выделенную ДНК количественно оценивали с использованием NanoPhotometer N50 (IMPLEN, США).

Для того чтобы целенаправленно подобрать CpG-динуклеотиды для последующего молекулярно-генетического исследования образцов ДНК индивидов из Республики Беларусь, проведены биоинформатический и статистический анализы данных 18 GEO-проектов. Информация об уровне метилирования ДНК с использованием ИМ 450K BeadChip (Illumina) представлена в открытом доступе на платформе Gene Expression Omnibus NCBI: для образцов крови – GSE40279, GSE42861, GSE51032, GSE50660, GSE55763, GSE77696, GSE106648, GSE125105 (5935 человек, диапазон возраста – от 17 до 93 лет); для образцов буккального эпителия – GSE39560, GSE59505, GSE72120, GSE72556, GSE78874, GSE92767, GSE99029, GSE110128, GSE111223 (690 человек, диапазон возраста – от 1 мес. до 91 года); для образцов спермы – GSE114753 (141 человек, диапазон возраста – от 19 до 57 лет).

На основании проведенного анализа отобран 41 CpG-динуклеотид, уровень метилирования которого коррелировал с хронологическим возрастом человека: для образцов крови 27 CpG-динуклеотидов – cg00481951 (*SST*), cg02872426 (*DDO*), cg05213896 (*IL411*), cg06784991 (*ZYG11A*), cg06874016 (*NKIRAS2*), cg07547549 (*SLC12A5*), cg07553761 (*TRIM59*), cg08128734 (*RASSF5*), cg08468401, cg08928145 (*TSSK6*), cg10501210 (*MIR29B2CHG*), cg11807280 (*MEIS1*), cg12757011 (*SLC4A10*), cg14361627 (*LOC105375508*), cg16054275 (*F5*), cg16867657 (*ELOVL2*), cg18384097 (*PTPN7*), cg18473521 (*HOXC4*), cg19283806 (*CCDC102B*), cg19671120 (*CNGA3*), cg22454769 (*FHL2*), cg24079702 (*FHL2*), cg25410668 (*RPA2*), Chr.3:g.51707136 (*GRM2*), Chr.4:g.16573797 (*LDB2*), Chr.7:g.130734357 (*KLF14*) и Chr.20:g.63980491 (*SAMD10*) [6, 22]; для образцов буккального эпителия 7 CpG-динуклеотидов – cg00481951 (*SST*), cg07547549 (*SLC12A5*), cg08928145 (*TSSK6*), cg12757011 (*SLC4A10*), cg14361627 (*LOC105375508*), cg18384097 (*PTPN7*), cg19671120 (*CNGA3*) [23, 24]; для образцов спермы 14 CpG-динуклеотидов – cg00881487 (*MXRA5*), cg03762081, cg04123357 (*COL18A1*), cg06304190 (*TTC7B*), cg06979108 (*NOX4*), cg10836509, cg12837463, cg13014709 (*CDK4*), cg13030797, cg13837679 (*ALX4*), cg14750551 (*PARP14*), cg21843517 (*SNRNP35*), cg23488376 (*LINC01595*), cg24958325 (*ANKMY1*) [15].

Бисульфитное секвенирование с использованием SNaPshot. Праймеры для анализа CpG-динуклеотидов cg02872426, cg06784991, cg06874016, cg07553761, cg11807280, cg16054275, cg16867657, cg18473521, cg25410668, Chr.3:g.51707136, Chr.4:g.16573797, Chr.7:g.130734357, Chr.20:g.63980491 приведены в работе J. Naue с соавт. [6]. Праймеры для анализа CpG-динуклеотидов cg00481951, cg07547549, cg08928145, cg12757011, cg14361627, cg18384097, cg19671120 представлены в работе

S.R. Hong et al. [24]. Праймеры для анализа CpG-динуклеотидов cg06304190, cg06979108, cg12837463, cg00881487, cg13030797, cg14750551, cg23488376 представлены в работе H.Y. Lee et al. [15]. Праймеры и SBE-олигонуклеотиды для анализа CpG-динуклеотидов cg03762081, cg04123357, cg05213896, cg08128734, cg08468401, cg10501210, cg10836509, cg13014709, cg13837679, cg19283806, cg21843517, cg22454769, cg24079702 и cg24958325 смоделированы с использованием программы BiSearch (<http://bisearch.enzim.hu/>) – табл. 1.

Бисульфитная конверсия ДНК проводилась для 200–500 нг с использованием MethylEdge® Bisulfite Conversion System Kit (Promega, США). ПЦР проводили в объеме 20 мкл, содержащем 10–15 нг конвертированной ДНК, 1 ед. ArtStart ДНК-полимеразы (ArtBioTech, Беларусь), 2 мкл 10× буфера для ПЦР (в состав которого входил хлорид магния в концентрации 3.0 мМ), 200 нМ каждого дНТФ и 0.4–1.0 мкМ каждого праймера. Протокол амплификации на термоциклере ProFlex PCR System (ThermoFisher, США) включал следующие этапы: 95°C – 4 мин, далее 34 цикла (94°C – 20 с, 56°C – 30 с, 72°C – 45 с), финальная элонгация при 72°C в течение 7 мин. В дальнейшем 5 мкл амплификата очищали с использованием Exo-CIP™ Rapid PCR Cleanup Kit (NEB, США). Реакцию микросеквенирования (SBE, Single-base extension) проводили с использованием SNaPshot Kit (Applied Biosystems, США) согласно протоколу производителя, каждая реакция включала также 3 мкл очищенного на предыдущей стадии амплификата и 0.2–0.4 мкМ соответствующего SBE-олигонуклеотида. В дальнейшем 10 мкл амплификата очищали с использованием 1 ед. FastAP Thermosensitive Alkaline Phosphatase (ThermoFisher).

Электрофоретическая детекция продуктов реакции микросеквенирования проводилась с использованием генетического анализатора ABI PRISM 3500 и программного обеспечения GeneMapper® 5.0 (Applied Biosystems).

Процент метилирования (0–100%) для каждого CpG-динуклеотида рассчитывали согласно формуле: интенсивность флуоресцентного сигнала для G (метилированные, т.е. неконвертированные цитозины; согласно правилу комплементарности C-G) разделить на сумму флуоресцентных сигналов G и A (неметилированные, т.е. конвертированные в T-цитозины; согласно правилу комплементарности T-A).

Статистический анализ данных. В ходе биоинформатического и статистического анализов были рассчитаны: коэффициенты корреляции R по Спирману; коэффициенты детерминации (R^2), а также скорректированные R^2 , равные $1 - (1 - R^2)(n - 1)/(n - k)$, где k – число независимых переменных и n – объем выборки; значения среднего абсолютного отклонения (Mean Absolute De-

viation, MAD) от истинного возраста и средне-квадратическая ошибка модели (Root Mean Square Error, RMSE). Данные уровня метилирования CpG-динуклеотидов были нормализованы с использованием алгоритма Batch normalization ($n = 30$). Статистический анализ проводили с использованием программного обеспечения SPSS v20.0.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Коэффициенты корреляции R

Коэффициенты R для CpG-динуклеотидов, рассчитанные по результатам молекулярно-генетического анализа для образцов крови и буккального эпителия с указанием половой принадлежности, представлены в табл. 2, для спермы – в табл. 3. Наибольшие расчетные значения R показаны для следующих CpG-динуклеотидов (в порядке уменьшения абсолютного значения): для образцов крови – cg19283806, cg14361627, cg08128734, cg16867657, cg22454769, cg07553761, cg08468401 и cg11807280; для образцов буккального эпителия – cg07547549, cg14361627, cg00481951, cg12757011 и cg08928145; для образцов спермы – cg13014709, cg12837463, cg06304190, cg00881487 и cg23488376. Из-за слабой или статистически незначимой ассоциации с хронологическим возрастом из дальнейшего моделирования исключены CpG-динуклеотиды cg08928145, cg18384097 (табл. 2) и cg13837679 (табл. 3).

Регрессионная модель предсказания возраста

В модели предсказания возраста по биологическим образцам отобраны CpG-динуклеотиды с использованием процедуры Stepwise, реализованной в SPSS, оказывающие достоверное и значительное влияние на изменение процента объясненной дисперсии R^2 : 12 CpG-динуклеотидов для крови – cg19283806, cg14361627, cg22454769, cg11807280, cg08468401, cg05213896, cg16867657, cg07547549, cg25410668, cg08128734, cg10501210 и Chr.4:g.16573797; 6 CpG-динуклеотидов для буккального эпителия – cg00481951, cg07547549, cg08928145, cg12757011, cg14361627 и cg19671120; 5 CpG-динуклеотидов для спермы – cg13014709, cg06304190, cg04123357, cg12837463, cg06979108.

При использовании множественной линейной регрессии в качестве модели предсказания возраста по образцам крови в общей группе индивидов (без разделения по полу) значение MAD составило 3.4 года, RSME – 2.4 года. При разделении по полу для мужчин MAD – 3.3 года, RSME – 2.3 года; для женщин MAD – 3.5 года, RSME – 2.4 года. Скорректированный процент объяснен-

Таблица 1. Праймеры и SBE-олигонуклеотиды, смоделированные с использованием программы BiSearch

№	СрG-динуклеотид	Ген	Хромосомная позиция*	Нуклеотидная последовательность
1	cg03762081	—	Chr.19:g.51020309	ACTCTGTGGC[CG]TGAGTGAGCT
				F-праймер (5' > 3'): GAGGTТАAGAGAAATAGAAATAAG R-праймер (5' > 3'): AATATAAAACAAAACCACAACCC SBE-олигонуклеотид (5' > 3'): (T) _n -AACACACACAAAАСТСАСТСАС
2	cg04123357	COL18A1	Chr.21:g.45509586	ACAGCCACCG[CG]ACTTCCAGCC
				F-праймер (5' > 3'): AGATGATATTTTGTTAGTT R-праймер (5' > 3'): AAАСТАААСССТСТТСТАСТА SBE-олигонуклеотид (5' > 3'): (T) _n -САСТСАССАССРАСТАААААТС
3	cg05213896	IL4I1	Chr.19:g.49890396	TGCGCTTCAC[CG]CCGGTCCGCT
				F-праймер (5' > 3'): ААСТТАААААСАССТТААТА R-праймер (5' > 3'): ААТТТГААГГТГТТГААГГТ SBE-олигонуклеотид (5' > 3'): (T) _n -АСААСААСААСАААААААТААТАС
4	cg08128734	RASSF5	Chr.1:g.206512090	GATTTAGCCC[CG]TGTTTGGCTT
				F-праймер (5' > 3'): ТТТАТТТТГГГТТТТГГААГГТ R-праймер (5' > 3'): ССАТТТСТАТАСТТСАТТТТСТ SBE-олигонуклеотид (5' > 3'): (T) _n -СТТАААСТСТААААССАААСАС
5	cg08468401	—	Chr.3:g.14261631	TGAGATGGTT[CG]AGGTGCCAGG
				F-праймер (5' > 3'): ТТТТАГТТГТТТТГТГГТГТТ R-праймер (5' > 3'): ТТТТСТТТАААСТТААТСТСАТТС SBE-олигонуклеотид (5' > 3'): (T) _n -АТААСАТААСТССТААСАССТС
6	cg10501210	MIR29B2CHG	Chr.1:g.207823675	CACCCATCGC[CG]CCTGCGCAGA
				F-праймер (5' > 3'): GAAAAGAAATТАААГТГГГГАА R-праймер (5' > 3'): АССААТТСТААААСАТТСРАААС SBE-олигонуклеотид (5' > 3'): (T) _n -ССАААТТТАААТСТАССРАААС
7	cg10836509	—	Chr.1:g.22611902	CTCCCATGGC[CG]CAGGCAGCAG
				F-праймер (5' > 3'): GTTAGGTTTGGTGATGTAGT R-праймер (5' > 3'): СААААСТТАААСТССААА SBE-олигонуклеотид (5' > 3'): (T) _n -АААСРСРТСТССТАСТАССТАС
8	cg13014709	CDK4	Chr.12:g.57754606	TCAGACAATC[CG]ATTCCATCC
				F-праймер (5' > 3'): ТТАТАТАТТААТТАТАГГ R-праймер (5' > 3'): АСТСАААСААТАТАСАСТТТАС SBE-олигонуклеотид (5' > 3'): (T) _n -САААТТТСААААТТТСС
9	cg13837679	ALX4	Chr.11:g.44308365	AGGCCTCCGC[CG]TGТССАГСС
				F-праймер (5' > 3'): GTTTGTGAAAGAAGGAATTT R-праймер (5' > 3'): ААТССАААСТААТАААААААСАС SBE-олигонуклеотид (5' > 3'): (T) _n -АССТАААААСАААААСТААААСАС
10	cg19283806	CCDC102B	Chr.18:g.68722183	TGTTTGCACATCC[CG]CATCCTTTTTTGGCT
				F-праймер (5' > 3'): GGGGAATGTTTGTATTTATTT R-праймер (5' > 3'): ТААААААТТАСССТТСТСТ SBE-олигонуклеотид (5' > 3'): (T) _n -АСААААТАААСАССААААААААТАС

Таблица 2. Коэффициенты корреляции (R) с возрастом для CpG-динуклеотидов (кровь и буккальный эпителий)

Кровь								
CpG-динуклеотид	мужчины ($n = 119$)				женщины ($n = 156$)			
	НГ 95%	R	ВГ 95%	p -уровень	НГ 95%	R	ВГ 95%	p -уровень
cg00481951	-0.010	0.182	0.344	4.80E-02	0.013	0.161	0.316	4.51E-02
cg02872426	-0.287	-0.423	-0.558	1.42E-06	-0.318	-0.485	-0.596	1.12E-06
cg05213896	0.355	0.515	0.636	2.12E-09	0.452	0.572	0.673	6.29E-15
cg06784991	0.157	0.334	0.495	2.03E-04	0.312	0.454	0.583	2.60E-09
cg06874016	-0.551	-0.383	-0.192	1.74E-05	-0.510	-0.375	-0.223	1.40E-06
cg07547549	0.267	0.441	0.579	5.33E-07	0.384	0.508	0.632	1.35E-11
cg07553761	0.488	0.629	0.732	1.91E-14	0.571	0.675	0.764	4.28E-22
cg08128734	-0.775	-0.682	-0.555	1.27E-17	-0.799	-0.730	-0.642	2.90E-27
cg08468401	-0.663	-0.530	-0.349	5.51E-10	-0.795	-0.713	-0.610	1.48E-25
cg08928145*	-0.111	0.078	0.251	0.398	-0.127	0.046	0.216	0.568
cg10501210	-0.568	-0.430	-0.269	1.09E-06	-0.608	-0.496	-0.366	4.64E-11
cg11807280	-0.715	-0.605	-0.458	3.15E-13	-0.718	-0.619	-0.501	7.51E-18
cg12757011	0.149	0.336	0.486	1.88E-04	0.203	0.363	0.506	3.25E-06
cg14361627	0.679	0.776	0.845	3.55E-25	0.665	0.751	0.813	1.56E-29
cg16054275	-0.574	-0.417	-0.245	2.37E-06	-0.496	-0.370	-0.225	2.05E-06
cg16867657	0.456	0.610	0.724	1.72E-13	0.628	0.716	0.791	7.70E-26
cg18384097	-0.486	-0.308	-0.114	6.47E-04	-0.426	-0.300	-0.143	1.42E-04
cg18473521	0.344	0.503	0.626	5.62E-09	0.300	0.450	0.574	3.88E-09
cg19283806	-0.858	-0.789	-0.676	1.65E-26	-0.871	-0.820	-0.751	3.23E-39
cg19671120	0.275	0.431	0.576	1.01E-06	0.076	0.243	0.388	2.21E-03
cg22454769	0.410	0.571	0.690	1.19E-11	0.653	0.734	0.794	1.29E-27
cg24079702	-0.087	0.106	0.299	0.252	0.078	0.238	0.386	2.73E-03
cg25410668	0.314	0.484	0.620	2.40E-08	0.579	0.686	0.767	4.88E-23
Chr.3:g.51707136	0.031	0.234	0.422	1.04E-02	0.337	0.465	0.590	9.58E-10
Chr.4:g.16573797	-0.627	-0.508	-0.354	3.55E-09	-0.600	-0.482	-0.334	1.92E-10
Chr.7:g.130734357	0.300	0.470	0.608	6.81E-08	0.322	0.466	0.584	8.98E-10
Chr.20:g.63980491	-0.599	-0.466	-0.308	9.48E-08	-0.703	-0.608	-0.493	3.99E-17

Буккальный эпителий								
CpG-динуклеотид	мужчины ($n = 262$)				женщины ($n = 204$)			
	НГ 95%	R	ВГ 95%	p -уровень	НГ 95%	R	ВГ 95%	p -уровень
cg00481951	0.654	0.722	0.779	1.89E-43	0.638	0.717	0.778	1.58E-33
cg07547549	0.714	0.770	0.819	1.23E-52	0.710	0.780	0.833	5.43E-43
cg08928145	0.467	0.555	0.634	1.56E-22	0.579	0.671	0.743	5.04E-28
cg12757011	0.551	0.633	0.702	8.56E-31	0.519	0.613	0.687	1.89E-22
cg14361627	0.694	0.753	0.796	3.68E-49	0.700	0.770	0.826	2.30E-41
cg18384097*	-0.230	-0.112	-0.001	0.071	-0.310	-0.184	-0.058	0.009
cg19671120	0.139	0.254	0.359	3.13E-05	0.080	0.210	0.328	0.003

Примечание. НГ 95% – нижняя граница 95%-ного доверительного интервала; ВГ 95% – верхняя граница 95%-ного доверительного интервала; * исключенные из моделирования CpG-динуклеотиды; p -уровень – уровень статистической значимости.

Например, при разбиении всего массива образцов на две категории (без разделения по полу) и использовании логистической регрессии на основании молекулярно-генетического анализа 12 CpG-динуклеотидов в образцах крови с вероятностью 92.0% имеется возможность сделать заключение, что возраст неизвестного индивида составляет меньше 40 лет, с вероятностью 98.2% – больше 40 лет, с вероятностью 89.7% – меньше 50 лет, с вероятностью 99.0% – больше 50 лет; с вероятностью 88.2% – меньше 60 лет, с вероятностью 89.9% – больше 60 лет. Также при анализе уровня метилирования 6 CpG-динуклеотидов в образцах буккального эпителия с вероятностью 93.8% имеется возможность сделать заключение, что возраст неизвестного индивида составляет меньше 40 лет, с вероятностью 89.8% – больше 40 лет. На основании анализа 5 CpG-динуклеотидов в образцах спермы с вероятностью 84.4% можно сделать заключение, что возраст неизвестного индивида составляет меньше 35 лет.

Таким образом, для полученных нами результатов характерен баланс точности предсказания возраста с использованием небольшого количества CpG-динуклеотидов – от 5 для спермы до 12 для крови. Мультиплексирование при молекулярно-генетическом тестировании позволяет проводить анализ в генетическом анализаторе за 1–2 реакции. Для технологии секвенирования SNaPshot характерны высокие значения точности и воспроизводимости. Тем не менее экспертам-криминалистам при использовании подобного подхода для предсказания возраста по биологическим следам необходимо будет учитывать ряд факторов, которые позволят получить достоверный результат: получать на этапе конверсии ДНК достаточное количество материала для ПЦР, так как повторная конверсия может оказать существенное влияние на воспроизводимость; проводить исследование метилирования CpG-динуклеотидов не менее чем в двух проворностях, чтобы усреднить результат. Выполнение данных рекомендаций позволит максимально нивелировать методические ошибки.

Полученные нами данные будут положены в основу при разработке калькуляторов предсказания возраста по биологическим следам для экспертов-криминалистов.

На основании данных, представленных в открытом доступе на платформе GEO NCBI для проектов по определению полногеномного профиля метилирования ДНК, сформирован перечень CpG-динуклеотидов с высокими значениями линейной зависимости между уровнем метилирования и хронологическим возрастом. Для молекулярно-генетического анализа отобран 41 CpG-динук-

Таблица 3. Коэффициенты корреляции (R) с возрастом для CpG-динуклеотидов (сперма)

CpG-динуклеотиды	НГ 95%	R	ВГ 95%	p -уровень
cg00881487	–0.301	–0.389	–0.222	2.93E-11
cg03762081	–0.261	–0.348	–0.181	9.84E-09
cg04123357	0.361	0.276	0.439	7.88E-16
cg06304190	–0.369	–0.446	–0.282	1.50E-16
cg06979108	0.326	0.242	0.402	5.07E-13
cg10836509	–0.266	–0.351	–0.175	5.16E-09
cg12837463	–0.373	–0.452	–0.289	6.81E-17
cg13014709	–0.422	–0.497	–0.331	1.28E-21
cg13030797	–0.199	–0.287	–0.113	1.48E-05
cg13837679*	–0.071	–0.156	0.021	0.124
cg14750551	–0.256	–0.334	–0.176	2.02E-08
cg21843517	0.244	0.15	0.331	9.20E-08
cg23488376	–0.294	–0.368	–0.216	9.71E-11
cg24958325	–0.181	–0.269	–0.091	8.17E-05

Примечание. НГ 95% – нижняя граница 95%-ного доверительного интервала, ВГ 95% – верхняя граница 95%-ного доверительного интервала, * – исключенные из моделирования CpG-динуклеотиды, p -уровень – уровень статистической значимости.

леотид с высоким прогностическим потенциалом для создания моделей предсказания возраста человека по биологическим образцам. Для данных CpG-динуклеотидов определен уровень метилирования для 1208 биологических образцов индивидов из Республики Беларусь, рассчитаны коэффициенты корреляции R с возрастом. С использованием множественной линейной регрессии созданы модели для предсказания возраста по образцам крови, буккального эпителия и спермы.

Предложенные нами модели предсказания возраста на основании профиля метилирования CpG-динуклеотидов крови, буккального эпителия и спермы отличаются относительной простотой, так как в анализе используется небольшое количество маркеров и методика, разработанная с их использованием, может применяться в криминалистических лабораториях молекулярно-генетической направленности. Ошибка предсказания возраста для рассчитанных нами моделей соответствует аналогичным исследованиям [6, 15, 24].

Разрабатываемая нами методика определения возраста по биологическим образцам на основе анализа профиля метилирования ДНК основывается на результатах отечественных научно-технических исследований. В результате данный под-

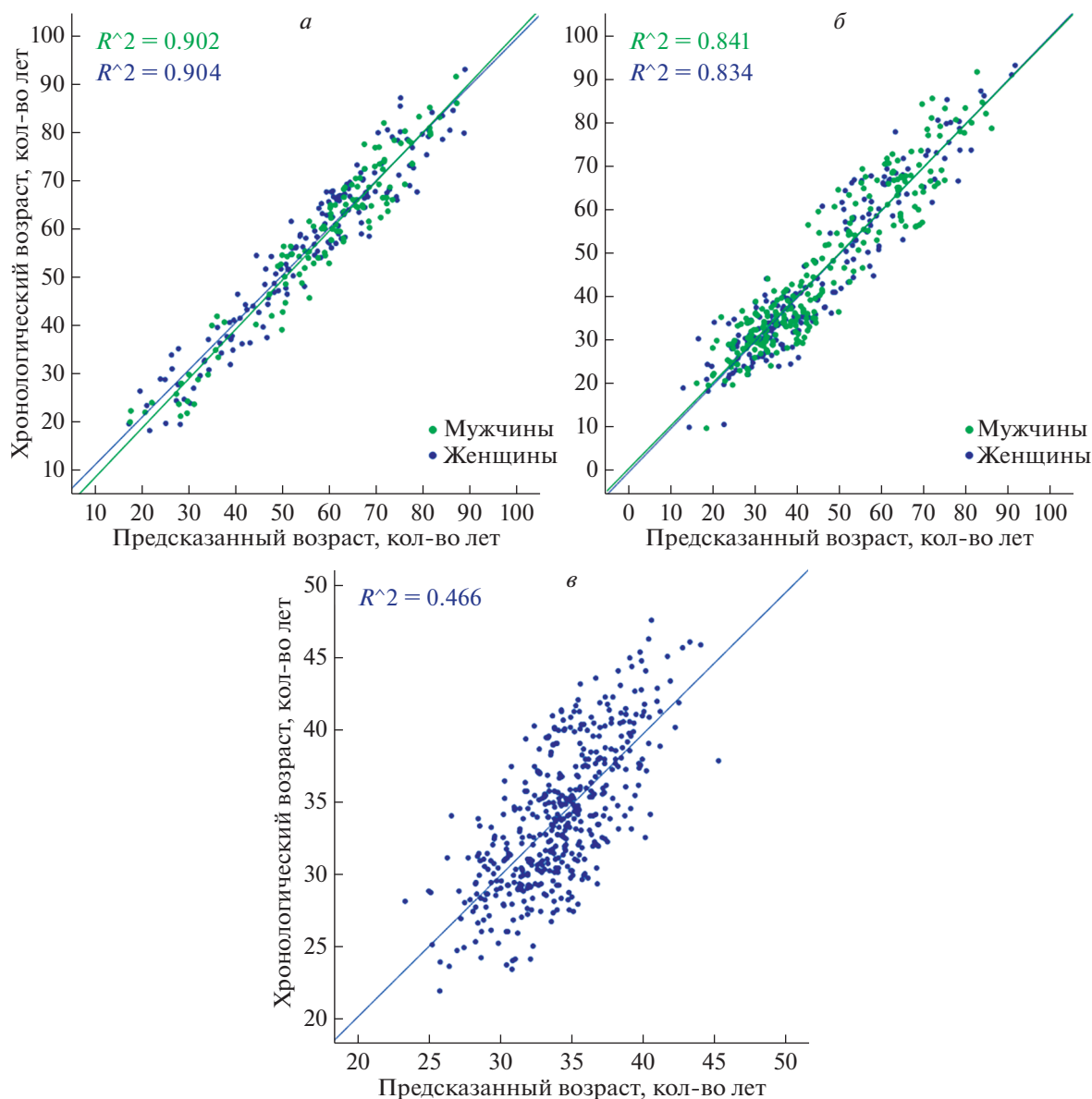


Рис. 1. Точность предсказания возраста индивида по образцам крови (*a*), буккального эпителия (*б*) и спермы (*в*).

ход обеспечит использование актуальных знаний в области прикладной генетики при решении криминалистических задач на территории Республики Беларусь и Российской Федерации.

Работа выполнена в рамках мероприятия 2 “Разработка методики определения вероятного возраста индивида по характеристике его ДНК” научно-технической программы Союзного государства “Разработка инновационных геногеографических и геномных технологий идентификации личности и индивидуальных особенностей человека на основе изучения генофондов регио-

нов Союзного государства” (ДНК-идентификация).

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального и/или национально-го комитета по исследовательской этике и Хельсинкской декларации 1964 г. и ее последующим изменениям или сопоставимым нормам этики.

От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Авторы заявляют об отсутствии конфликта интересов.

СПИСОК ЛИТЕРАТУРЫ

1. Золотаренко А.Д., Чекалин Е.В., Брускин С.А. Современные молекулярно-генетические методы определения возраста в криминалистике // Генетика. 2019. Т. 55. № 12. С. 1397–1409. <https://doi.org/10.1134/S0016675819120154>
2. Кильчевский А., Моссэ И., Шантуренко М., Буракова А. Генетика – судебной экспертизе Беларуси // Наука и инновации. 2020. № 10(212). С. 22–28.
3. Hannum G., Guinney J., Zhao L. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates // Mol. Cell. 2013. V. 24. P. 359–367. <https://doi.org/10.1016/j.molcel.2012.10.016>
4. Kurdyukov S., Bullock M. DNA methylation analysis: Choosing the right method // Biology (Basel). 2016. V. 5. № 1. P. e1–e21. <https://doi.org/10.3390/biology5010003>
5. Bocklandt S., Lin W., Sehl M.E. et al. Epigenetic predictor of age // PLoS One. 2011. V. 6. P. 1–6. <https://doi.org/10.1371/journal.pone.0014821>
6. Naue J., Hoefsloot H.C.J., Mook O.R.F. et al. Chronological age prediction based on DNA methylation: Massive parallel sequencing and random forest regression // Forensic Sci. Int. Genet. 2017. V. 31. P. 19–28. <https://doi.org/10.1016/j.fsigen.2017.07.015>
7. Park J.L., Kim J.H., Seo E. et al. Identification and evaluation of age-correlated DNA methylation markers for forensic use // Forensic Sci. Int. Genet. 2016. V. 23. P. 64–70. <https://doi.org/10.1016/j.fsigen.2016.03.005>
8. Vidaki A., Ballard D., Aliferi A. et al. Syndercombe Court D. DNA methylation-based forensic age prediction using artificial neural networks and next generation sequencing // Forensic Sci. Int. Genet. 2017. V. 28. P. 225–236. <https://doi.org/10.1016/j.fsigen.2017.02.009>
9. Zbiac-Piekarska R., Spolnicka M., Kupiec T. et al. Development of a forensically useful age prediction method based on DNA methylation analysis // Forensic Sci. Int. Genet. 2015. V. 17. P. 173–179. <https://doi.org/10.1016/j.fsigen.2015.05.001>
10. Eipel M., Mayer F., Arent T. et al. Epigenetic age predictions based on buccal swabs are more precise in combination with cell type-specific DNA methylation signatures // Aging (Albany NY). 2016. V. 8. № 5. P. 1034–1048. <https://doi.org/10.18632/aging.100972>
11. McEwen L.M., O'Donnell K.J., McGill M.G. et al. The PedBE clock accurately estimates DNA methylation age in pediatric buccal cells // PNAS. 2020. V. 117. № 38. P. 23329–23335. <https://doi.org/10.1073/pnas.1820843116>
12. Koop B.E., Mayer F., Gündüz T. et al. Postmortem age estimation via DNA methylation analysis in buccal swabs from corpses in different stages of decomposition—a “proof of principle” study // Int. J. Legal Med. 2021. V. 135. № 1. P. 167–173. <https://doi.org/10.1007/s00414-020-02360-7>
13. Dongen J., Ehli E.A., Jansen R. et al. Genome-wide analysis of DNA methylation in buccal cells: a study of monozygotic twins and mQTLs // Epigenetics Chromatin. 2018. V. 11. № 1. P. e1–e14. <https://doi.org/10.1186/s13072-018-0225-x>
14. Wozniak A., Heidegger A., Piniewska-Rog D. et al. Development of the VISAGE enhanced tool and statistical models for epigenetic age estimation in blood, buccal cells and bones // Aging (Albany NY). 2021. V. 13. № 5. P. 6459–6484. <https://doi.org/10.18632/aging.202783>
15. Lee H.Y., Jung S.E., Oh Y.N. et al. Epigenetic age signatures in the forensically relevant body fluid of semen: a preliminary study // Forensic Sci. Int. Genet. 2015. V. 19. P. 28–34. <https://doi.org/10.1016/j.fsigen.2015.05.014>
16. Alsaleh H., McCallum N.A., Halligan D.L., Hadrill P.R. A multi-tissue age prediction model based on DNA methylation analysis // Forensic Sci. Int. Genet. Suppl. Ser. 2017. V. 6. P. 62–64. <https://doi.org/10.1016/j.fsigs.2017.09.056>
17. Li L., Song F., Huang Y., Zhu H., Hou Y. Age-associated DNA methylation determination of semen by pyrosequencing in Chinese Han population // Forensic Sci. Int. Genet. 2017. V. 6. P. e99–e100. <https://doi.org/10.1016/j.fsigs.2017.09.042>
18. Jenkins T.G., James E.R., Alonso D.F. et al. Cigarette smoking significantly alters sperm DNA methylation patterns // Andrology. 2017. V. 5. № 6. P. 1089–1099. <https://doi.org/10.1111/andr.12416>
19. Lee J.W., Choung C.M., Jung J.Y. et al. A validation study of DNA methylation-based age prediction using semen in forensic casework samples // Leg. Med. (Tokyo). 2018. V. 31. P. 74–77. <https://doi.org/10.1016/j.legalmed.2018.01.005>
20. Fleckhaus J., Freire-Aradas A., Rothschild M.A., Schneider P.M. Impact of genetic ancestry on chronological age prediction using DNA methylation analysis // Forensic Sci. Int. Genet. Suppl. Ser. 2017. V. 6. P. e399–e400. <https://doi.org/10.1016/j.fsigs.2017.09.162>
21. Donkin I., Barres R. Sperm epigenetics and influence of environmental factors // Mol. Metab. 2018. V. 14. P. 1–11. <https://doi.org/10.1016/j.molmet.2018.02.00>
22. Купень В.Н., Богданова М.В., Буракова А.А. и др. Обоснование минимального объема выборки для предсказания хронологического возраста человека // Мол. и прикл. генетика. 2021. Т. 30. С. 39–48. <https://doi.org/10.47612/1999-9127-2021-30-39-48>
23. Купень В.Н., Богданова М.В., Буракова А.А. и др. Предсказательный потенциал CpG-маркеров для определения хронологического возраста человека // Мол. и прикл. генетика. 2020. Т. 28. С. 70–80.
24. Hong S.R., Jung S., Lee E.H. et al. DNA methylation-based age prediction from saliva: High age predictability by combination of 7 CpG markers // Forensic Sci. Int. Genet. 2017. V. 29. P. 118–125. <https://doi.org/10.1016/j.fsigen.2017.04.006>

Determination of Human Chronological Age from Biological Samples Based on the Analysis of CpG-Dinucleotides Methylation

V. A. Lemesh^{a, *}, V. N. Kipen^{a, **}, M. V. Bahdanava^a, A. A. Burakova^a, A. G. Bulgak^b,
A. V. Bayda^c, S. A. Bruskin^d, O. V. Zotova^b, and O. I. Dobysh^a

^a*Institute of Genetics and Cytology of the National Academy of Sciences of Belarus, Minsk, 220072 Belarus*

^b*State Institution "Republican Scientific and Practical Centre "Cardiology", Minsk, 220089 Belarus*

^c*Belarusian Medical Academy of Postgraduate Education, Minsk, 220013 Belarus*

^d*Vavilov Institute of General Genetics RAS, Moscow, 119991 Russia*

**e-mail: v.lemesh@igc.by*

***e-mail: v.kipen@igc.by*

Based on bioinformatics and statistical analysis of GEO-projects to determine the genome-wide profile of human DNA methylation, a list of 41 CpG dinucleotides with high predictive potential was formed to create models for predicting human age from biological samples. The methylation level was determined for 1208 samples of individuals from the Republic of Belarus (275 – blood, 466 – buccal epithelium, 467 – sperm). The correlation coefficients R were calculated, and mathematical models for determining the age of individual were constructed. The average value of the accuracy of age prediction from blood samples using 12 CpG-dinucleotides was 3.4 years (for men – 3.3, for women – 3.5); for buccal epithelium samples using 6 CpG-dinucleotides – 4.6 years (for men – 4.5, for women – 4.7); for sperm samples using 5 CpG-dinucleotides – 3.0 years. The results obtained will be used as a basis for the development of calculators for predicting the age of an individual based on traces of biological character for forensic experts.

Keywords: forensic, age determination, blood, buccal epithelium, semen, DNA methylation, human methylation 450K BeadChip.