

УДК 57.05;57.07

ДОСТИЖЕНИЯ И ПЕРСПЕКТИВЫ ЭПИГЕНЕТИЧЕСКИХ ИССЛЕДОВАНИЙ ДРЕВНЕЙ ДНК

Обзор

© 2021 К.В. Жур¹, В.А. Трифонов², Е.Б. Прохорчук^{1*}

¹ ФИЦ «Фундаментальные основы биотехнологии» РАН, 119071 Москва, Россия;
электронная почта: prokhortchouk@gmail.com

² Институт истории материальной культуры РАН, 191186 Санкт-Петербург, Россия

Поступила в редакцию 29.10.2021

После доработки 17.11.2021

Принята к публикации 17.11.2021

Развитие технологий высокопроизводительного полногеномного секвенирования и совершенствование методик пробоподготовки сделало возможным исследование ДНК из археологических образцов возрастом более миллиона лет. В процессе изучения древней ДНК (дДНК) учёными было сделано множество крупных открытий, касающихся миграций людей, замещений одних популяций другими, межвидовых скрещиваний кроманьонцев с неандертальцами и денисовцами, эволюции патогенов человека и др. Не менее революционным стало заявление о возможности исследовать эпигенетические модификации геномов древних людей, что позволило извлечь ранее недоступную информацию, в том числе об активности генов, позиционировании нуклеосом, метилировании ДНК. Анализируя статус метилирования определенных локусов в геноме, можно установить хронологический возраст человека на момент его смерти, реконструировать некоторые особенности фенотипа, как это было сделано для денисовца, и даже определить неблагоприятные факторы внешней среды, воздействию которых подвергался древний человек. Представленный обзор посвящён современным достижениям и перспективным направлениям изучения эпигенетических модификаций дДНК с анализом методических подходов, применяемых в данной области научных исследований.

КЛЮЧЕВЫЕ СЛОВА: древняя ДНК, палеоэпигенетика, эпигенетические часы, редактирование эпигенома, ожирение.

DOI: 10.31857/S0320972521120058

ВВЕДЕНИЕ

Стремительное развитие высокопроизводительных методов полногеномного секвенирования способствовало заметному увеличению количества и качества информации, получаемой из образцов древней ДНК (дДНК). Совершенствование методов пробоподготовки позволило преодолеть множество сложностей, характерных для работы с археологическими образцами, связанных с низким качеством и количеством ДНК, а также с контаминацией генетическим материалом других организмов. Долгое время самым древним образцом, из которого были получены полногеномные данные, считали останки лошади из среднего плейстоцена возрастом 560–780 тысяч лет [1], в связи с чем предполагали, что в образцах старше одного миллиона лет

генетический материал не сохраняется. Пересмотрели теоретический предел сохранности ДНК после выхода статьи в журнале *Nature*, в которой были проанализированы полногеномные данные, полученные из останков мамонтов раннего и среднего плейстоцена [2]. В соответствии с данным молекулярного датирования (байесовские молекулярные часы) [3], возраст животного, найденного на реке Крестовка, составил 1,65 миллиона лет [2].

Сохранность генетического материала свыше миллиона лет открывает огромные перспективы для более детального изучения эволюции человека. Анализируя последовательности ДНК, возможно отслеживать миграции людей, выявлять замещение одних популяций другими, устанавливать факт межвидового скрещивания, исследовать предрасположенность к различным заболеваниям, предсказывать физические и некоторые психофизиологические особенности древнего человека [4]. Исследуя генетический материал из зубного налёта, можно изучать диету архаичных людей и их микробиом [5, 6], в

Принятые сокращения: дДНК – древняя ДНК; CpG-локус – сайт, где цитозин (C) находится рядом с гуанином (G) в последовательности ДНК.

* Адресат для корреспонденции.

свою очередь, анализ ДНК древних патогенов даёт возможность проследить происхождение и эволюцию возбудителей различных заболеваний, например, как это было показано для бактерий *Yersinia pestis*, возбудителя чумы [7].

За последнее десятилетие учёные значительно продвинулись в изучении дДНК, в том числе вымерших видов людей. Многих интересовал вопрос: как сильно отличался архаичный человек от современного, и что позволило последнему выжить и успешно размножиться. Первый черновой вариант генома неандертальца был опубликован в 2010 г. [8], а уже в 2012–2013 гг. были прочитаны полные геномы денисовского человека [9] и неандертальца [10] с качеством, не уступающим образцам современной ДНК (с 30- и 52-кратным покрытием соответственно). Сравнение последовательностей геномов ныне живущих людей с геномами денисовского человека и неандертальца выявило более 30 000 однонуклеотидных позиций, отделяющих современных людей от архаичных [10]. Тем не менее большинство из этих замен были нейтральными, 3117 приходились на регуляторные элементы, 32 предположительно влияли на сайты сплайсинга и только 96 замен приводили к аминокислотным заменам. Учёные пришли к выводу, что выявленными отличиями в последовательностях геномов современных людей и архаичных нельзя объяснить весь спектр их фенотипических и адаптационных различий, а полученные результаты свидетельствуют в пользу значительного влияния эпигенетических модификаций на эволюцию человека [11, 12].

ПОДХОДЫ К ИЗУЧЕНИЮ ЭПИГЕНЕТИЧЕСКИХ МОДИФИКАЦИЙ ДРЕВНЕЙ ДНК

Существуют разнообразные механизмы эпигенетической регуляции активности генов, включая метилирование ДНК, модификации гистонов и негистоновых белков хроматина, воздействие на экспрессию генов посредством малых некодирующих молекул РНК и другие зачастую взаимосвязанные процессы. В случае с дДНК исследователи в основном фокусируются на изучении метилирования ДНК и составлении карт позиционирования нуклеосом [13–15], что в первую очередь связано с технической возможностью реконструкции данного вида эпигенетических модификаций в археологических образцах.

Первые попытки изучения метилирования генома древних людей были предприняты ещё в 2010 г. группой С. Паабо; тогда учёные продемон-

стрировали, что ДНК неандертальца возрастом более 38 000 лет сохраняет картину метилирования [13]. Несколько позже, в 2014 г., была получена первая карта позиционирования нуклеосом для генома древнего человека [14]. Возможность установить, какие участки ДНК входят в состав нуклеосомного комплекса, а какие являются межнуклеосомными, т.е. открытыми, позволяет определить доступность генетического материала для взаимодействия с регуляторными белками, в том числе факторами, регулирующими транскрипцию, и, соответственно, предсказать экспрессию тех или иных генов у древних людей. В свою очередь, метилирование ДНК участвует в регуляции структуры хроматина, клеточной дифференцировки и многих других процессов, в том числе контролирует экспрессию генов на уровне транскрипции. Несмотря на то что метилирование генома и составление карты позиционирования нуклеосом не являются единственными механизмами эпигенетической регуляции, тем не менее изучение этих процессов даёт возможность извлечь из образцов дДНК ранее недоступную информацию об активности генов.

Основная сложность, возникающая при работе с дДНК – это существенная посмертная деградация генетического материала до фрагментов длиной в 25 пар нуклеотидов под действием эндогенных нуклеаз, а также в результате случайного гидролиза и окисления [16]. Наиболее частым гидролитическим повреждением дДНК является дезаминирование цитозина (С), т.е. отщепление аминогруппы от азотистого основания с образованием урацила (U). Дезаминированные цитозины в основном располагаются на концах фрагментов дДНК, что свидетельствует о наличии выступающих одноцепочечных концов, так как скорость дезаминирования цитозина в десятки раз выше в одноцепочечной, чем в двухцепочечной ДНК [17–19]. Особенность определения последовательности дДНК заключается в том, что при секвенировании методом синтеза на месте дезаминированных цитозинотимина исследователь наблюдает тимин (Т). Это связано с тем, что в процессе приготовления библиотек проводится обогащение с помощью полимеразной цепной реакции (ПЦР) с применением полимеразы, которая способна продолжать синтез комплементарной цепи в присутствии урацила. Во время синтеза комплементарной цепи урацил распознаётся полимеразой как тимин, в результате чего достраивается комплементарное тимину основание – аденин (А). В итоге исходная пара оснований С/Г (где Г – гуанин) в результате дезаминирования цитозина превращается в пару U/G,

которая в процессе ПЦР преобразуется в Т/А в одной из дочерних молекул. Соответственно, при секвенировании библиотек фрагментов дДНК исследователь наблюдает замены С → Т на 5'-конце молекулы ДНК или G → A — на 3'-конце в зависимости от особенностей выбранного протокола пробоподготовки. Важно отметить, что частота таких замен коррелирует с возрастом образца [20] и используется в качестве критерия, подтверждающего, что ДНК принадлежит именно древнему организму (эндогенная ДНК), а не является результатом контаминации современным генетическим материалом. В целом анализ фрагментов митохондриальной ДНК, выделенной из образцов возрастом от 18 до 60 000 лет, продемонстрировал сильную положительную корреляцию между возрастом и частотой замены С на Т. В группе образцов возрастом 117 лет и младше частота замен С → Т не превышала 5%, в то время как для образцов возрастом 500 лет и старше частота замен составляла не менее 11%. В свою очередь, для образцов ДНК, выделенных из тканей неандертальцев, данный показатель обычно составляет не менее 20%, а зачастую достигает 40–50% [16, 20].

Посмертная деградация дДНК, а конкретно процесс дезаминирования цитозина, усложняет исследование дДНК, но в то же время даёт возможность изучать метилирование геномов древних людей. Группа С. Паабо в процессе приго-

товления библиотек дДНК для секвенирования обрабатывала дДНК смесью ферментов урацил-ДНК-гликозилазы и эндонуклеазы VIII, что позволяет вырезать урацил с образованием одностороннего разрыва, оставляя неповреждённые части нетронутыми [13]. Такой способ значительно повышает точность секвенирования, а также даёт возможность определить были ли цитозин метилирован, так как в результате дезаминирования метилированного цитозина последний превращается в остаток тимина и, соответственно, не распознаётся смесью ферментов. Следовательно, участки генома, в которых цитозины были прижизненно метилированы, будут содержать больше чтений с тиминном, по сравнению с неметилированными участками, а доля замен С → Т будет отражать степень метилирования дДНК.

Метод подготовки библиотек для секвенирования с обработкой ДНК смесью ферментов урацил-ДНК-гликозилазы и эндонуклеазы VIII представляет собой непрямой способ анализа метилирования генома и предполагает применение специальных вычислительных алгоритмов для реконструкции метилирования. Существует и ещё один непрямой способ оценки степени метилирования дДНК, основанный на использовании двух различных полимераз для исходной матрицы: Taq-полимеразы, не чувствительной к урацилу, и Pfu-полимеразы, которая в присутствии урацила обрывает синтез цепи [14]. В случае дезаминирования метилированных цитозинов образуется тимин, следовательно, обе полимеразы смогут продолжить синтез цепи, и в результате секвенирования исследователь в обеих реакциях будет наблюдать замену С → Т в последовательности ДНК (рис. 1). В свою очередь, дезаминирование неметилированных цитозинов создаёт «непроходимый» для Pfu-полимеразы урацил, и синтез цепи обрывается. Таким образом, на основании соотношения замен С → Т, выявляемых при секвенировании библиотек фрагментов ДНК, приготовленных с помощью Taq- и Pfu-полимераз, становится возможным восстановление профиля метилирования дДНК.

Непрямые способы анализа метилирования генома являются предпочтительными при работе с дДНК в отличие от прямых методов, таких как бисульфитная конверсия [21] или обогащение молекул ДНК, содержащих метилированные CpG, с помощью белков, содержащих метил-связывающие домены (methyl-CpG-binding domain, MBD) [22]. Например, для бисульфитной конверсии необходимо большое количество генетического материала, что уже делает этот подход неприменимым для древних образцов, а обогащение с помощью белков, содержащих ме-

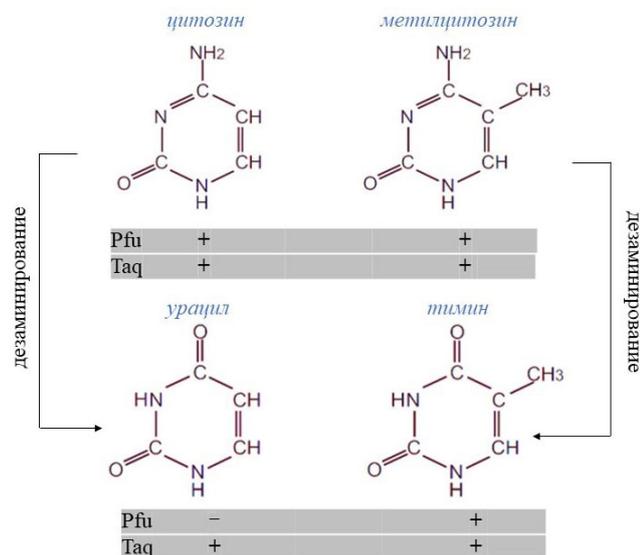


Рис. 1. Схематическое изображение метода оценки метилирования дДНК, основанного на использовании двух различных полимераз для исходной матрицы. Способность Pfu-полимеразы и Taq-полимеразы продолжать синтез цепи в присутствии различных типов нуклеотидов во время ПЦР-амплификации обозначена знаками «+» (способна) и «-» (неспособна)

тил-связывающие домены, может вносить существенный сдвиг в сторону анализа более длинных фрагментов ДНК, игнорируя короткие, действительно эндогенные, молекулы.

РЕКОНСТРУКЦИЯ ЭПИГЕНОМОВ АРХАИЧНЫХ ЛЮДЕЙ

В 2014 г. научная группа Д. Гохмана (David Gokhman) впервые опубликовала полные реконструированные карты метилирования ДНК неандертальца и денисовца [23]. Исследователи сравнили картины метилирования ДНК, выделенной из костной ткани древнего и современного человека, и выявили порядка 2000 дифференциально метилированных участков [23]. На тот момент особое внимание привлекли отличия в метилировании кластера генов *HOXD*, регулирующих развитие разных частей тела у многоклеточных животных. В промоторе гена *HOXD9* и в теле гена *HOXD10* костной ткани древнего человека было установлено гиперметилирование, в то время как у всех исследованных образцов костной ткани современных людей для этих участков характерно гипометилирование. Ранее, в исследовании на мышинных модельных системах, было показано, что изменения в экспрессии генов данного кластера приводят к морфологическим преобразованиям, напоминающим различия в строении конечностей неандертальца и современного человека [24]. Таким образом, эпигенетические модификации данного региона могут играть ключевую роль в эволюции конечностей современного человека, что в очередной раз свидетельствует о том, что при изучении микроэволюционных и макроэволюционных процессов необходимо учитывать как генетический, так и эпигенетический фактор.

В 2019 г. команда Д. Гохмана опубликовала работу, в которой представила новый подход в изучении функциональной роли дифференциально метилированных участков генома архаичных людей [25]. Предложенный способ позволял предсказывать фенотипические особенности древних людей, основываясь на допущении, что изменения, вызываемые сильным метилированием промотора, похожи на изменения, возникающие в результате мутаций в гене, которые либо существенно снижают уровень его экспрессии, либо полностью выключают ген. При разработке метода исследователи использовали базу данных «Онтология человеческого фенотипа» (The Human Phenotype Ontology, HPO) [26], где содержится информация о том, как мутации, нарушающие работу того или ино-

го гена, сказываются на фенотипе человека, в данном случае на строении скелета. Метод был предварительно проверен на видах с известной морфологией, таких как неандертальцы и шимпанзе. Точность определения расходящихся черт составила более 85%, что позволило применить данный подход для реконструкции фенотипа денисовца (девочки из Денисовой пещеры) [25]. На текущий момент не найдено полных скелетов или черепов архаичных людей данного вида, и восстановление их внешности на основании эпигенетических модификаций генома является единственно возможным способом. Исследователям удалось предсказать 56 отличий денисовцев от неандертальцев и/или современных людей. Денисовцы имели ряд общих черт с неандертальцами: удлинённое лицо, низкий лоб, выступающие челюсти, широкий таз и крупную грудную клетку, также были выявлены и уникальные особенности, например, более длинная зубная дуга и более выраженное латеральное расширение черепа [25]. Важно отметить, что полученные результаты не противоречат морфологии обнаруженных на сегодняшний день фрагментов скелета денисовцев.

Несмотря на впечатляющие результаты, полученные командой Д. Гохмана, у исследователей возник ряд трудностей, связанных с отсутствием полных референсных карт метилирования костной ткани современного человека, малым количеством образцов, необходимостью сравнивать данные метилирования, полученные с помощью различных технологий, а также с отсутствием таких данных для организмов, которые можно было бы использовать в качестве внешней группы (аутгруппы) при определении эволюционных взаимоотношений между исследуемыми популяциями людей разных видов [25].

Для преодоления перечисленных сложностей исследователи приняли решение дополнительно реконструировать профили метилирования ещё нескольких древних геномов, в том числе эпигеном древних людей современного анатомического облика, а также целенаправленно реконструировать полные карты метилирования в костной ткани современных людей и нескольких шимпанзе с применением различных технологий. Такой подход позволил максимально учесть факторы, которые могут влиять на результат поиска дифференциально метилированных регионов: технологию секвенирования, различия в покрытии генома и в степени дезаминирования, возраст и пол человека (например, некоторые регионы генома подвержены гиперметилированию только у мужчин), часть скелета (бедренная кость, височная кость черепа, фаланга пальца, зуб), из которой выде-

лялась ДНК, а также потенциальный шум, который возникает во время реконструкции карт метилирования.

В результате командой Д. Гохмана была создана уникальная платформа для изучения динамики метилирования ДНК в процессе эволюции человека, что позволило обнаружить дифференциально метилированные участки в 588 генах [25]. В том числе было установлено гиперметилирование в участках генов *SOX9*, *ACAN*, *COL2A1*, *NFIX* и *XYLTI*, связанных с анатомией лица и голосового тракта, характерное только для современных людей. Известно, что *SOX9*, *ACAN* и *COL2A1* представляют собой группу скоординированно функционирующих генов, взаимодействующих друг с другом и регулирующих рост и формирование скелета, включая каркас гортани и кости лица [27]. Предположительно гиперметилирование этих генов могло сыграть ключевую роль в формировании лица и голосового тракта современного человека [15].

Важно отметить, что в тканях, не относящихся к скелетным (мозг, кровь и др.), профиль метилирования данных генов практически не различается у шимпанзе и современного человека, что свидетельствует о том, что различия являются специфическими именно для скелетных тканей. Такой вывод позволяет сделать наличие профиля метилирования для шимпанзе, организма, выступающего в качестве внешней эволюционной группы. Если исследователю известен статус метилирования определенного региона у шимпанзе в костной и других тканях, то, имея эту информацию, становится возможным предсказать статус метилирования исследуемого локуса в тканях архаичных людей, не сохранившихся для исследования, в том числе в тканях мозга [28], в то время как в костной ткани уровень метилирования можно определить экспериментально.

На рис. 2 представлена схема предсказания статуса метилирования геномного локуса в тка-

нях мозга архаичных людей. Например, если исследуемый локус неметилирован в костной ткани шимпанзе и у неандертальского или денисовского человека (сценарий 1), но метилирован в костной ткани современного человека и тканях мозга современных людей и шимпанзе, то данная эпимутация (изменением метилирования ДНК без изменения последовательности) является специфичной для костной ткани, которая возникла у современных людей, изменив состояние локуса с неметилированного на метилированный. Следовательно, у архаичных видов людей этот локус в тканях мозга будет скорее всего метилирован. В том случае, если исследуемый локус неметилирован в костной ткани шимпанзе и у неандертальского или денисовского человека (сценарий 2), но метилирован в костной ткани современного человека, и при этом в тканях мозга современных людей локус метилирован, а в тканях мозга шимпанзе неметилирован, то данная эпимутация затронула все ткани, изменив статус локуса на метилированный. Такое изменение могло произойти, например, если эпигенетическое состояние локуса определяется до тканевой дифференцировки. В данном случае у архаичных видов людей этот локус в тканях мозга будет скорее всего неметилирован. Стоит также различать такие тканеспецифические эпимутации с эпимутациями, которые приводят к противоположному состоянию метилирования гена во всех органах и тканях у людей (современных и архаичных) и шимпанзе. Последние являются отражением генетических изменений в процессе эволюции от обезьяны к человеку.

Следует также отметить, что кроме генов, регулирующих рост и формирование скелета, дифференциально метилированные участки в геномах архаичных и современных людей выявлены в генах, ассоциированных с развитием нервной, мышечной, сердечно-сосудистой систем и другими важными процессами в организме, что несомненно требует дальнейшего детального изу-

Организм		Шимпанзе		Современный человек		Архаичный человек	
		костная ткань	мозг	костная ткань	мозг	костная ткань	мозг (предсказание)
Статус метилирования исследуемого локуса	сценарий 1	—	+	+	+	—	? (+)
	сценарий 2	—	—	+	+	—	? (-)



Рис. 2. Схема предсказания статуса метилирования исследуемого геномного локуса в тканях мозга архаичных людей. «+» — Метилированное состояние геномного локуса; «—» — неметилированное состояние геномного локуса

чения. Таким образом, исследование эпигенетических модификаций древних геномов открывает новые возможности для изучения микро- и макроэволюционных процессов живых организмов, а также предоставляет уникальную возможность реконструировать фенотипические признаки архаичных людей, в том числе тех, для которых ещё не обнаружено полных скелетов.

ЭПИГЕНЕТИЧЕСКИЕ «ЧАСЫ» ДЛЯ ОПРЕДЕЛЕНИЯ ВОЗРАСТА ДРЕВНЕГО ЧЕЛОВЕКА

Ещё одним интересным направлением исследований является возможность определения хронологического возраста древнего человека на основе оценки статуса метилирования определенных CpG-локусов (сайтов, где цитозин (C) находится рядом с гуанином (G) в последовательности ДНК), степень метилирования которых изменяется по мере взросления и старения организма. Повышенный интерес к такому подходу связан с тем, что установление возраста древнего человека на момент его смерти как правило вызывает трудности, так как останки древних людей зачастую представлены в виде отдельных фрагментов скелета. Например, в случае денисовского человека были найдены всего несколько мелких костей и зубов [9, 29], что не позволяет достоверно определить возраст классическими подходами, изучая морфологию костей черепа или других частей скелета. Кроме того, традиционные методы основаны на визуальной, зачастую довольно субъективной, оценке морфологических особенностей скелета, что может приводить к ошибкам.

Практическая возможность определения возраста древнего человека на момент его смерти на основании данных метилирования ДНК была впервые показана на примере 4000-летнего палеоэскимоса, принадлежащего к культуре Saqqaq [14]. ДНК палеоэскимоса была выделена из стержня волос, хорошо сохранившихся в вечной мерзлоте, что позволило произвести широкомасштабное секвенирование образца с 20-кратным покрытием. Для определения возраста исследователи применили линейную модель Koch и Wagner 2011 года [30], описывающую связь возраста и статуса метилирования пяти CpG-локусов, из которых четыре, ассоциированные с генами *NPTX2* (cg12799895), *TRIM58* (cg07533148), *GRIA2* (cg25148589) и *KCNQ1DN* (cg01530101), подвергаются гиперметилированию с возрастом и один CpG-локус в гене *BIRC4BP* (cg23571857) по мере взросления организма претерпевает гипометилирование. Для

определения возраста палеоэскимоса исследователи использовали только два локуса (cg07533148 и cg01530101) из пяти, которые предварительно показали наибольшую информативность для ДНК, выделенной из волос современного человека, с различиями между прогнозируемым и реальным возрастом в пределах 1,7–12,4 года. Согласно статусу метилирования регионов вблизи данных локусов, возраст палеоэскимоса на момент смерти составил приблизительно 40 лет.

Основная цель разработчиков моделей для прогнозирования хронологического возраста состоит в том, чтобы научиться определять биологический возраст человека, который не всегда совпадает с хронологическим, что может отражать интенсивность процессов старения для каждого организма индивидуально. В свою очередь, более детальное изучение эпигенетических механизмов старения поможет разрабатывать стратегии по замедлению этого процесса и увеличению продолжительности и качества жизни. На сегодняшний момент разработано не менее 15 эпигенетических «часов» старения [31], в одних моделях для оценки возраста достаточно 3 CpG-локуса [32], в других – необходимо проанализировать несколько сотен [33, 34]. В свою очередь, каждая модель характеризуется определенной погрешностью и типом ткани, для которой прогнозирование работает наилучшим образом. Например, модель Horvath использует 353 CpG-локуса, работает с большим количеством тканей человека и имеет среднюю ошибку предсказания 3–5 лет [33], что является весьма неплохим показателем. Тем не менее подавляющее большинство имеющихся сегодня эпигенетических калькуляторов прогнозируют возраст по профилю метилирования ДНК, выделенной из цельной крови (наиболее доступного материала), что абсолютно неприменимо для образцов дДНК. Следовательно, разработка новых моделей или изучение возможности применения существующих моделей для прогнозирования хронологического возраста древнего человека по данным метилирования ДНК, выделенной из костной ткани, костей черепа и зубов, является актуальной задачей и требует дополнительных исследований в этом направлении.

ИССЛЕДОВАНИЕ НЕДОСТУПНЫХ РАННЕ АСПЕКТОВ ЖИЗНИ ДРЕВНИХ ЛЮДЕЙ

Отдельным перспективным направлением в палеоэпигенетике является изучение локусов генома, характеризующихся чувствительностью к условиям окружающей среды [35]. Такие локусы, выступая в качестве «посредников» между

внешней средой и ДНК, модифицируют свой паттерн метилирования в ответ на определенные воздействия, такие как изменения в питании, смена климатических условий, взаимодействие с химическими веществами и др. На сегодняшний день опубликовано большое количество работ, подтверждающих влияние различных внешних факторов на эпигенетические модификации генома, в том числе на эпигеном человека. Например, в исследовании под руководством Т. Пауса (T. Paus) было показано, что воздействие сигаретного дыма на плод во время внутриутробного развития изменяет статус метилирования целого ряда генов [36]. В работе М. Росарио (M. Rosario) с соавторами установлено, что у детей, рожденных от матерей с диабетом второго типа, выявляется более 4000 дифференциально метилированных регионов в геноме при сравнении с детьми, рожденными от женщин без данного заболевания [37]. Интересные данные были получены Б. Хейманс (B. Heijmans) в ходе исследования детей, выношенных в период массового голода, поразившего население Нидерландов под конец Второй мировой войны, известного как «Голодная зима 1944 года». Исследователями было показано, что в группе людей, подвергшихся воздействию голода на самых ранних этапах своего развития, наблюдается снижение уровня метилирования гена *IGF2* [38]. Данный ген представляет особый интерес, поскольку кодирует инсулиноподобный фактор роста 2 и играет важную роль в регуляции процессов роста и развития организма. Хорошо известно, что активность данного гена регулируется по механизму геномного импринтинга, в том числе посредством метилирования ДНК [39], соответственно, снижение уровня метилирования гена *IGF2* может приводить к изменению активности инсулиноподобного фактора роста 2, оказывая существенное влияние на процессы роста и развития организма, а также на развитие патологических состояний во взрослом возрасте [40, 41]. Таким образом, существует множество доказательств того, что воздействие неблагоприятных факторов на ранних стадиях внутриутробного развития может изменять эпигенетический статус определенных локусов генома и сохраняться на протяжении жизни человека, оказывая влияние на его здоровье.

Исследование локусов в геноме древнего человека, профиль метилирования которых ассоциирован с тем или иным воздействием внешней среды, может служить новым подходом в изучении условий проживания архаичных людей, их диеты, заболеваний, а также токсических веществ, действию которых они подвергались в первую очередь в наиболее чувствитель-

ный эмбриональный период развития. В связи с тем, что изменение в метилировании, возникшее на ранней стадии развития, ещё до дифференцировки, будет перенесено на все дочерние клетки организма, это позволяет изучать эпигенетические локусы, чувствительные к воздействиям внешней среды, в тканях, доступных для исследования в палеоэпигенетике (зубы, волосы и кости скелета) [35].

Особый интерес для учёных представляет изучение эпигенетических модификаций геномов древних людей для исследования механизмов таких заболеваний, как ожирение и ассоциированные с ним патологии. Различают моногенную тяжёлую форму ожирения с ранней манифестацией, вызванную мутацией в одном гене, и полигенную форму. При моногенной форме внешние факторы не оказывают существенного влияния на развитие заболевания. В то время как полигенная форма относится к мультифакторным гетерогенным заболеваниям и развивается в результате дисбаланса между потребляемой и расходуемой энергией, в основе которого лежат нарушения метаболизма, генетическая предрасположенность, нарушение поведенческих реакций и влияние внешних факторов. Полигенная форма ожирения является наиболее распространённой в современном обществе [42].

С развитием технологий секвенирования в ходе исследований методом полногеномного поиска ассоциаций (genome-wide association studies, GWAS) выявлено более 1000 локусов, ассоциированных с признаками ожирения, тем не менее совместно они могут объяснить только 8,4% вариативности индекса массы тела человека [43]. Большинство обнаруженных вариантов, выявленных в результате полногеномного анализа ассоциаций, по отдельности имеют слабую связь с ожирением и связанными с ним патологиями и зачастую расположены в некодирующей части генов или в межгенном пространстве [44–46].

В последние годы появляется всё больше работ, свидетельствующих о том, что ключевая роль в патогенезе ожирения может принадлежать эпигенетическим механизмам [47–50]. Широкое распространение ожирения за последние десятилетия в странах с низким и средним уровнем доходов, в первую очередь в районах с большим количеством ресторанов быстрого питания, может быть следствием изменений в профиле метилирования ДНК в результате малоподвижного современного образа жизни и несбалансированного питания. Тем не менее молекулярный механизм патогенеза ожирения до конца неясен, и непонятно: являются ли выявленные эпигенетические изменения причиной

развития заболевания или его следствием. Зачастую данные об уровне и спектре дифференциально метилированных генов противоречивы, а число таких работ ограничено.

Исследование дДНК может улучшить наше понимание эпигенетической составляющей молекулярной природы ожирения и ассоциированных с ним патологий. Отличный от современного образ жизни древних людей позволяет использовать их профиль метилирования генома в качестве эпигенетической «нормы». Предполагая, что древние люди современного типа в меньшей степени были подвержены таким неблагоприятным факторам, как малоподвижный образ жизни и высококалорийное питание с высокой долей жиров и простых углеводов, следует ожидать выявления различий в профиле метилирования геномов древних людей и ныне живущих. Логично предположить, что обнаруженная эпигенетическая вариабельность определённых участков генома может играть важную роль в развитии ожирения и ассоциированных с ним заболеваний. С этой точки зрения интересно сравнить профили метилирования современных людей и, например, людей бронзового века, живших порядка 5000 лет назад. Учитывая, что такой промежуток времени довольно короткий с точки зрения эволюции (приблизительно 200 поколений), можно предположить, что изменения в метилировании вследствие изменения самой нуклеотидной последовательности ДНК маловероятны. Следовательно, адаптация на малых временных масштабах возможна преимущественно за счёт эпигенетических изменений.

На сегодняшний момент исследования доступны способы проверки функциональной значимости выявленных различий в профиле метилирования современных и древних людей: например, на клеточных линиях путём моделирования статуса метилирования исследуемого локуса с помощью лентивирусных конструкций, экспрессирующих каталитически неактивный белок dCas9 слитый с каталитическим доменом TET1CD (вызывает таргетное деметилирование) или Dnmt3a (вызывает таргетное метилирование) [51, 52]. Такие конструкции позволяют изменять метилирование регионов генома, заранее заданных исследователем с помощью РНК-гидов, каждый из которых направит dCas9 к собственной мишени. В случае с ожирением моделирование статуса метилирования изучаемых локусов в индуцированных плюрипотентных стволовых клетках с последующей направленной их дифференцировкой в адипогенном направлении позволит сравнить транскрипционные профили адипоцитов с модифицированным метилированием заданных локусов и без

него и выявить, какие гены изменяют свою активность, в том числе с учётом их функциональных связей между собой.

Работы такого рода довольно трудоёмкие, но в то же время могут способствовать получению новых сведений об эпигенетической составляющей молекулярной природы ожирения и связанных с ним болезней. В свою очередь, выявленные с использованием такого оригинального подхода эпигенетические локусы, ассоциированные с исследуемым заболеванием, могут быть использованы в качестве перспективных мишеней для поиска новых лекарственных препаратов.

ЗАКЛЮЧЕНИЕ

Исследование последовательностей и эпигенетических модификаций геномов древних людей является довольно дорогим и трудозатратным направлением, требующим специальных условий для работы, большого объёма секвенирования, вычислительных мощностей и владения специфическими инструментами анализа данных. Тем не менее поддержание и развитие таких исследований является задачей государственной важности не только с точки зрения биологических наук, например, для изучения эволюции человека и его заболеваний, но и для истории и политологии. Зачастую это единственный метод, способный объективно и наглядно продемонстрировать и обосновать бездоказательность экстремистских теорий о превосходстве одних народов над другими. Ярким историческим примером служат идеи немецкого археолога Kossinna [53], которые позже были подхвачены нацистами, что привело к известным событиям и огромному числу жертв.

На сегодняшний день исследования дДНК перестали быть исключительно фундаментальными и приобрели очевидный прикладной аспект. Например, изучение ДНК древних людей способно улучшить наше понимание генетической и эпигенетической составляющей молекулярной природы различных заболеваний, в том числе таких как ожирение и ассоциированные с ним патологии. Исследование эволюции мозга архаичного и современного человека, а конкретно генетических и эпигенетических механизмов регуляции нервной системы, может пролить свет на механизмы патогенеза психических и нейродегенеративных заболеваний. Анализ ДНК патогенов древних людей даёт возможность изучать происхождение и эволюцию возбудителей заболеваний современных людей, что позволит оценить шансы появления ещё более

опасных патогенов и поможет избежать новых вспышек эпидемий в будущем.

Таким образом, палеогенетика и недавно выделившееся её направление, палеоэпигенетика, являются важными приоритетными направлениями для страны, требующими поддержки со стороны государства. До недавнего времени полногеномное секвенирование уникальных образцов из местных археологических памятников мирового значения в большинстве случаев проводилось в зарубежных лабораториях, что не способствовало развитию исследований дДНК в России и не позволяло местным археологам на равных участвовать в интерпретации полученных результатов. Полноценное становление и развитие данного направления предоставит возможность учёным самостоятельно проводить

исследования ценных археологических образцов, обнаруженных на территории страны, повысит конкурентоспособность российской науки на мировом уровне, а, следовательно, позволит на равных участвовать в интерпретации исторических событий прошлого.

Финансирование. Работа выполнена при финансовой поддержке Минобрнауки России, грант № 13.1902.21.0023 (согл. № 075-10-2020-116).

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм. Настоящая статья не содержит описания каких-либо исследований с участием людей или животных в качестве объектов.

СПИСОК ЛИТЕРАТУРЫ

- Orlando, L., Ginolhac, A., Zhang, G., Froese, D., Albrechtsen, A., et al. (2013) Recalibrating Equus evolution using the genome sequence of an early Middle Pleistocene horse, *Nature*, **499**, 74-78.
- Van der Valk, T., Pečnerová, P., Díez-Del-Molino, D., Bergström, A., Oppenheimer, J., et al. (2021) Million-year-old DNA sheds light on the genomic history of mammoths, *Nature*, **591**, 265-269.
- Reis, M., Donoghue, P., and Yang, Z. (2016) Bayesian molecular clock dating of species divergences in the genomics era, *Nat. Rev. Genet.*, **17**, 71-80.
- Krause, J., and Pääbo, S. (2016) Genetic time travel, *Genetics*, **203**, 9-12.
- Weyrich, L. S., Duchene, S., Soubrier, J., Arriola, L., Llamas, B., et al. (2017) Neanderthal behaviour, diet, and disease inferred from ancient DNA in dental calculus, *Nature*, **544**, 357-361.
- Jensen, T. Z. T., Niemann, J., Iversen, K. H., Fotakis, A. K., Gopalakrishnan, S., et al. (2019) A 5700 year-old human genome and oral microbiome from chewed birch pitch, *Nat. Commun.*, **10**, 5520.
- Rasmussen, S., Allentoft, M. E., Nielsen, K., Orlando, L., Sikora, M., et al. (2015) Early divergent strains of *Yersinia pestis* in Eurasia 5,000 years ago, *Cell*, **163**, 571-582.
- Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., et al. (2010) A draft sequence of the Neandertal genome, *Science*, **328**, 710-722.
- Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., et al. (2012) A high-coverage genome sequence from an archaic Denisovan individual, *Science*, **338**, 222-226.
- Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., et al. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains, *Nature*, **505**, 43-49.
- Jablonka, E. (2013) Epigenetic inheritance and plasticity: the responsive germline, *Progr. Biophys. Mol. Biol.*, **111**, 99-107.
- Luo, X., Song, R., Moreno, D. F., Ryu, H.-Y., Hochstrasser, M., et al. (2020) Epigenetic mechanisms contribute to evolutionary adaptation of gene network activity under environmental selection, *Cell Rep.*, **33**, 108306.
- Briggs, A. W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., et al. (2010) Removal of deaminated cytosines and detection of *in vivo* methylation in ancient DNA, *Nucleic Acids Res.*, **38**, e87.
- Pedersen, J. S., Valen, E., Velazquez, A. M. V., Parker, B. J., Rasmussen, M., et al. (2014) Genome-wide nucleosome map and cytosine methylation levels of an ancient human genome, *Genome Res.*, **24**, 454-466.
- Gokhman, D., Nissim-Rafinia, M., Agranat-Tamir, L., Housman, G., García-Pérez, R., et al. (2020) Differential DNA methylation of vocal and facial anatomy genes in modern humans, *Nat. Commun.*, **11**, 1189.
- Rohland, N., Glocke, I., Aximu-Petri, A., and Meyer, M. (2018) Extraction of highly degraded DNA from ancient bones, teeth and sediments for high-throughput sequencing, *Nat. Protoc.*, **13**, 2447-2461.
- Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *Nature*, **362**, 709-715.
- Kavli, B., Otterlei, M., Slupphaug, G., and Krokan, H. (2007) Uracil in DNA – general mutagen, but normal intermediate in acquired immunity, *DNA Repair*, **6**, 505-516.
- Srivastava, S., and Singh, N. (2011) The probability analysis of opening of DNA, *J. Chem. Phys.*, **134**, 115102.
- Sawyer, S., Krause, J., Guschanski, K., Savolainen, V., and Pääbo, S. (2012) Temporal patterns of nucleotide misincorporations and DNA fragmentation in ancient DNA, *PLoS One*, **7**, e34131.
- Llamas, B., Holland, M. L., Chen, K., Cropley, J. E., Cooper, A., et al. (2012) High-resolution analysis of cytosine methylation in ancient DNA, *PLoS One*, **7**, e30226.
- Seguin-Orlando, A., Gamba, C., Sarkissian, C. D., Ermini, L., Louvel, G., et al. (2015) Pros and cons of methylation-based enrichment methods for ancient DNA, *Sci. Rep.*, **5**, 11826.
- Gokhman, D., Lavi, E., Prüfer, K., Fraga, M. F., Riancho, J. A., et al. (2014) Reconstructing the DNA methylation maps of the Neandertal and the Denisovan, *Science*, **344**, 523-527.
- Zakany, J., and Duboule, D. (2007) The role of Hox genes during vertebrate limb development, *Curr. Opin. Gen. Dev.*, **17**, 359-366.
- Gokhman, D., Mishol, N., de Manuel, M., de Juan, D., Shuqrun, J., et al. (2019) Reconstructing Denisovan anatomy using DNA methylation maps, *Cell*, **179**, 180-192.e10.
- Köhler, S., Gargano, M., Matentzoglou, N., Carmody, L. C., Lewis-Smith, D., et al. (2021) The human phenotype ontology in 2021, *Nucleic Acids Res.*, **49**, D1207-D1217.
- Lee, Y.-H., and Saint-Jeannet, J.-P. (2011) Sox9 function in craniofacial development and disease, *Genesis*, **49**, 200-208.
- Gokhman, D., Meshorer, E., and Carmel, L. (2016) Epigenetics: it's getting old. Past meets future in paleoepigenetics, *Trends Ecol. Evol.*, **31**, 290-300.
- Reich, D., Green, R. E., Kircher, M., Krause, J., Patterson, N., et al. (2010) Genetic history of an archaic

- hominin group from Denisova Cave in Siberia, *Nature*, **468**, 1053-1060.
30. Koch, C. M., and Wagner, W. (2011) Epigenetic-aging-signature to determine age in different tissues, *Aging*, **3**, 1018-1027.
 31. Bergsma, T., and Rogaeva, E. (2020) DNA methylation clocks and their predictive capacity for aging phenotypes and healthspan, *Neurosci. Insights*, **15**, 2633105520942221.
 32. Bocklandt, S., Lin, W., Sehl, M. E., Sánchez, F. J., Sinsheimer, J. S., et al. (2011) Epigenetic predictor of age, *PLoS One*, **6**, e14821.
 33. Horvath, S. (2013) DNA methylation age of human tissues and cell types, *Genome Biol.*, **14**, R115.
 34. Zhang, Q., Vallerga, C. L., Walker, R. M., Lin, T., Henders, A. K., et al. (2019) Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing, *Genome Med.*, **11**, 54.
 35. Gokhman, D., Malul, A., and Carmel, L. (2017) Inferring past environments from ancient epigenomes, *Mol. Biol. Evol.*, **34**, 2429-2438.
 36. Toledo-Rodriguez, M., Lotfipour, S., Leonard, G., Perron, M., Richer, L., et al. (2010) Maternal smoking during pregnancy is associated with epigenetic modifications of the brain-derived neurotrophic factor-6 exon in adolescent offspring, *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **7**, 1350-1354.
 37. Rosario, M. C., Ossowski, V., Knowler, W. C., Bogardus, C., Baier, L. J., et al. (2014) Potential epigenetic dysregulation of genes associated with MODY and type 2 diabetes in humans exposed to a diabetic intrauterine environment: an analysis of genome-wide DNA methylation, *Metabolism*, **63**, 654-660.
 38. Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., et al. (2008) Persistent epigenetic differences associated with prenatal exposure to famine in humans, *Proc. Natl. Acad. Sci. USA*, **105**, 17046-17049.
 39. Smith, F. M., Garfield, A. S., and Ward, A. (2006) Regulation of growth and metabolism by imprinted genes, *Cytogenet. Genome Res.*, **113**, 279-291.
 40. Cui, H., Cruz-Correa, M., Giardiello, F. M., Hutcheon, D. F., Kafonek, D. R., et al. (2003) Loss of IGF2 imprinting: a potential marker of colorectal cancer risk, *Science*, **299**, 1753-1755.
 41. Ravenel, J. D., Broman, K. W., Perlman, E. J., Niemitz, E. L., Jayawardena, T. M., et al. (2001) Loss of imprinting of insulin-like growth factor-II (IGF2) gene in distinguishing specific biologic subtypes of Wilms tumor, *J. Nat. Cancer Inst.*, **93**, 1698-1703.
 42. Loos, R. J. F., and Yeo, G. S. H. (2021) The genetics of obesity: from discovery to biology, *Nat. Rev. Genet.*, **23**, 1-14.
 43. Khera, A. V., Chaffin, M., Wade, K. H., Zahid, S., Brancale, J., et al. (2019) Polygenic prediction of weight and obesity trajectories from birth to adulthood, *Cell*, **177**, 587-596.e9. doi: 10.1016/j.cell.2019.03.028.
 44. Goodarzi, M. O. (2018) Genetics of obesity: what genetic association studies have taught us about the biology of obesity and its complications, *Lancet Diab. Endocrinol.*, **6**, 223-236.
 45. Müller, M. J., Geisler, C., Blundell, J., Dulloo, A., Schutz, Y., et al. (2018) The case of GWAS of obesity: does body weight control play by the rules? *Int. J. Obesity*, **42**, 1395-1405.
 46. Locke, A. E., Kahali, B., Berndt, S. I., Justice, A. E., Pers, T. H., et al. (2015) Genetic studies of body mass index yield new insights for obesity biology, *Nature*, **518**, 197-206.
 47. Cheng, Z., Zheng, L., and Almeida, F. A. (2018) Epigenetic reprogramming in metabolic disorders: nutritional factors and beyond, *J. Nutr. Biochem.*, **54**, 1-10.
 48. Pietiläinen, K. H., Ismail, K., Järvinen, E., Heinonen, S., Tummers, M., et al. (2016) DNA methylation and gene expression patterns in adipose tissue differ significantly within young adult monozygotic BMI-discordant twin pairs, *Int. J. Obesity*, **40**, 654-661.
 49. Muniandy, M., Heinonen, S., Yki-Järvinen, H., Hakkarainen, A., Lundbom, J., et al. (2017) Gene expression profile of subcutaneous adipose tissue in BMI-discordant monozygotic twin pairs unravels molecular and clinical changes associated with sub-types of obesity, *Int. J. Obesity*, **41**, 1176-1184.
 50. Perflyev, A., Dahlman, I., Gillberg, L., Rosqvist, F., Iggman, D., et al. (2017) Impact of polyunsaturated and saturated fat overfeeding on the DNA-methylation pattern in human adipose tissue: a randomized controlled trial, *Am. J. Clin. Nutr.*, **105**, 991-1000.
 51. Morita, S., Noguchi, H., Horii, T., Nakabayashi, K., Kimura, M., et al. (2016) Targeted DNA demethylation *in vivo* using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions, *Nat. Biotechnol.*, **34**, 1060-1065.
 52. Nuñez, J. K., Chen, J., Pommie, G. C., Cogan, J. Z., Replogle, J. M., et al. (2021) Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing, *Cell*, **184**, 2503-2519.e17.
 53. Kossinna, G. (1919) *Das Weichselland, ein Uralter Heimatboden der Germanen*, Kafermann, Leipzig.

ACHIEVEMENTS AND PROSPECTS FOR EPIGENETIC STUDIES OF ANCIENT DNA

Review

K. V. Zhur¹, V. A. Trifonov², and E. B. Prokhortchouk^{1*}

¹ Federal Research Centre "Fundamentals of Biotechnology", Russian Academy of Sciences, 119071 Moscow, Russia; E-mail: prokhortchouk@gmail.com

² Institute for the History of Material Culture Russian Academy of Sciences, 191186 St.-Petersburg, Russia

The development of high-throughput whole-genome sequencing technology and improvement of sample preparation techniques make it possible to study ancient DNA (aDNA) from archaeological samples that are over a million years old. Study of aDNA sequences shed the light on a history of human migrations, clarified a chain of populations changes, interbreeding of Cro-Magnons with Neanderthals and Denisovans, evolution of human pathogens, etc. Equally important was the announcement of the possibility to study epigenetic modifications of ancient genomes, which allows receiving previously inaccessible information, including gene expression data, nucleosome positioning and DNA methylation. Cytosine methylation maps of Denisovan genome opens a road for chronological age prediction, phenotype reconstruction and even elucidating a list of unfavorable environmental factors that have influenced the ancient person. In this review we discuss actual progress in epigenetics of ancient DNA including methodological approaches and promising directions of the studies.

Keywords: ancient DNA, paleoepigenetics, epigenetic clock, epigenome editing, obesity