

КЛОНАЛЬНЫЙ СОСТАВ МУЛЬТИПОТЕНТНЫХ МЕЗЕНХИМНЫХ СТРОМАЛЬНЫХ КЛЕТОК ЧЕЛОВЕКА: ПРИМЕНЕНИЕ ГЕНЕТИЧЕСКИХ ШТРИХ-КОДОВ ДЛЯ ИССЛЕДОВАНИЯ*

© 2019 А.Е. Бигильдеев^{1**}, А.М. Пилунов², Н.В. Сац¹, В.Л. Сурин¹,
И.Н. Шипунова¹, Н.А. Петинати¹, М.Д. Логачева³, А.В. Федотова³,
А.С. Касьянов⁴, А.С. Артюхов⁵, Э.Б. Дашинимаев⁶, Н.И. Дризе¹

¹ Национальный медицинский исследовательский центр
гематологии Минздрава России, 125167 Москва, Россия;
электронная почта: bigildeev.ae@gmail.com

² Московский государственный университет им. М.В. Ломоносова,
119991 Москва, Россия

³ НИИ физико-химической биологии МГУ им. М.В. Ломоносова,
119992 Москва, Россия

⁴ Институт общей генетики им. Н.И. Вавилова РАН,
119991 Москва, Россия

⁵ РНИМУ им. Н.И. Пирогова Минздрава России,
117997 Москва, Россия

⁶ Институт биологии развития им. Н.К. Кольцова РАН,
119334 Москва, Россия

Поступила в редакцию 12.09.2018

После доработки 27.11.2018

Принята к публикации 27.11.2018

В работе изучали клональный состав мультипотентных мезенхимных стромальных клеток (ММСК) человека, маркированных с помощью лентивирусных векторов, несущих генетические штрих-коды. ММСК были инфицированы клонированной библиотекой самоинактивирующихся лентивирусных векторов, несущих 667 неповторяющихся штрих-кодов. При каждом пассировании культуры 120 клеток клонировали в 96-луночных планшетах по одной клетке на лунку. Определяли эффективность клонирования и маркирования клоногенных клеток. Из клеточных колоний-клонов выделяли ДНК и идентифицировали штрих-коды с помощью секвенирования по методу Сэнгера. Из общей популяции ММСК каждого пассажа также выделяли ДНК и определяли разнообразие, а также представленность штрих-кодов с помощью метода глубокого секвенирования на платформе Illumina. Было показано, что доля маркированных лентивирусными векторами ММСК стабильна в пассажах. Процедура маркирования при высоких значениях множественности заражения может приводить к снижению пролиферации ММСК. Анализ штрих-кодов в отдельных клеточных клонах подтвердил поликлональность популяции ММСК. При пассировании культуры ее клональный состав значительно менялся вследствие истощения пролиферативного потенциала большинства клеток. Большие клоны были обнаружены на первом пассаже ММСК, далее в популяции выявлялось множество небольших клонов с ограниченным пролиферативным потенциалом. Результаты глубокого секвенирования подтвердили данные об изменении клонального состава ММСК. В поликлональной популяции ММСК содержится лишь небольшое количество клеток с высоким пролиферативным потенциалом, среди которых могут быть и стволовые. ММСК с высоким пролиферативным потенциалом чаще выявляются на самых ранних пассажах. В связи с этим для использования ММСК в подходах регенеративной медицины, опирающихся на пролиферацию клеток, рекомендуется применять ММСК ранних пассажей.

КЛЮЧЕВЫЕ СЛОВА: мультипотентные мезенхимные стромальные клетки, генетическое маркирование, штрих-код, лентивирусный вектор.

DOI: 10.1134/S0320972519030072

Принятые сокращения: СКК – стволовая кроветворная клетка; МСК – мезенхимная стволовая клетка; ММСК – мультипотентные мезенхимные стромальные клетки; eGFP – зеленый флуоресцентный белок (enhanced green fluorescent protein); NGS – высокопроизводительное параллельное секвенирование (next generation sequencing); ППС – полная питательная среда; ПЦР – полимеразная цепная реакция; CNV – количество провирусов в расчете на один диплоидный геном (copy number variation); MOI – множественность заражения (multiplicity of infection).

* Приложение к статье на английском языке опубликовано на сайте журнала «Biochemistry» (Moscow) и на сайте издательства Springer (Link.springer.com), том 84, вып. 3, 2019.

** Адресат для корреспонденции.

В костном мозге присутствуют два типа стволовых клеток: стволовые кроветворные клетки (СКК) и мезенхимные стволовые клетки (МСК). Если СКК интенсивно изучают почти 100 лет, то активные исследования МСК начались в последней четверти XX века. Впервые наличие МСК было описано А.Я. Фриденштейном [1], который называл их стромальными механоцитами. Термин МСК был введен А. Капланом для обозначения прилипающих к пластику стромальных клеток человека, способных к дифференцировке *in vitro* в костную, хрящевую и жировую ткани [2]. В 2005 г. Общество по клеточной терапии сделало заявление, что эти клетки не подходят под определение стволовых, т.к. обладают ограниченным пролиферативным потенциалом [3]. Было предложено называть их мультипотентными мезенхимными стромальными клетками (ММСК), а также был разработан четкий критерий клеток, подходящих под это определение [4]. Необходимость такого критерия была связана с тем, что эти клетки стали активно использовать для лечения различных заболеваний, а также в регенеративной медицине [5, 6]. Применение этих клеток часто опирается на их способность к пролиферации и дифференцировке. Оказалось, что популяция ММСК гетерогенна и представлена множеством клеточных клонов, различающихся по своим свойствам [7–9]. Результаты этих работ имеют общий недостаток: исследованные клетки не были маркированы, что не позволяло сделать вывод о клональном составе популяции ММСК. Нашей группой была предпринята попытка индивидуально маркировать ММСК с помощью лентивирусного вектора, содержащего зеленый флуоресцентный белок (eGFP). Была установлена гетерогенность популяции по пролиферативному потенциалу и ее поликлональность [10]. Однако эта работа также имела недостатки: проанализировано небольшое число клонов, а также не было проведено исследование общей популяции ММСК.

Несколько лет назад было предложено маркировать СКК с помощью библиотеки ретровирусных векторов, содержащей большое разнообразие коротких нуклеотидных последовательностей, так называемых «генетических штрих-кодов» [11]. После интеграции провируса каждая клетка исследуемой популяции приобретает уникальный, стабильный и наследуемый маркер, позволяющий проследить за судьбой потомства маркированной клетки и отличать его от потомства других меченых стволовых клеток. Важный аспект такого маркирования – разнообразие штрих-кодов в библиотеке должно многократно превосходить количество стволовых

клеток, которые предполагается маркировать [12]. Определить маркеры можно как в отдельных клонах, так и в общей популяции маркированных клеток с помощью современных методов высокопроизводительного параллельного секвенирования (next generation sequencing, NGS) [13]. Сочетание методов генетического штрих-кодирования и NGS позволило решить многие исследовательские задачи, в частности, определить клональный состав и его динамику в кроветворной системе [11, 14–18], в популяции стволовых клеток молочной железы [19] и других клеточных популяциях [20]. Однако маркирование мезенхимных стволовых клеток и ММСК костного мозга с помощью штрих-кодов до настоящего времени не проводилось. Целью данной работы являлось установление клонального состава популяции ММСК человека с помощью маркирования клеток библиотекой лентивирусных векторов с генетическими штрих-кодами.

МЕТОДЫ ИССЛЕДОВАНИЯ

Получение клонированной библиотеки генетических штрих-кодов. Библиотека плазмид LeGO-nude-BC32 была любезно предоставлена проф. Борисом Фезе и Керстин Корнильс (University Medical Center Hamburg-Eppendorf, Гамбург, Германия). Библиотека была получена на основе плазмиды системы лентивирусных векторов 3-го поколения LeGO [21], в которую вместо гена, кодирующего флуоресцентный белок, были клонированы последовательности генетических штрих-кодов вида NNNACTNCGANNCTTNCGANNCTTNNGGANNCTANNACTNNCGANNCTTNCGANNCTTNNGGANNCTANNACTNNCGANNCTC, где N – вырожденные нуклеотиды. Штрих-коды содержат 35 вырожденных нуклеотидов и 48 константных. Теоретическое разнообразие комбинаций вырожденных нуклеотидов в таких штрих-кодах составляет 4^{35} , или $\approx 10^{19}$. Нуклеотидный состав и структура штрих-кодов подробно охарактеризованы в работе Agnyosy et al. [22]. Компетентные клетки *E. coli* штамма XL1 трансформировали библиотекой плазмид методом теплового шока в стандартных условиях согласно руководству Maniatis et al. [23]. Отдельные колонии трансформантов получали на чашках Петри с 1,5%-ным LB-агаром и ампициллином («Синтез», Россия) в концентрации 100 мкг/мл после инкубации при 37 °C в течение 17 ч. Из индивидуальных колоний выделяли ДНК и определяли трансформанты, несущие вставку. Из ночных культур трансформантов объемом 2 мл выделяли плазмидную ДНК и определяли нуклеотидную последовательность штрих-кода

с помощью секвенирования по Сэнгеру (праймер для секвенирования – BC-PCR-Seq, табл. 1). Для создания библиотеки использовали штрих-коды, соответствующие следующим критериям: 1) штрих-коды, не содержащие мутаций в константной области; 2) штрих-коды, имеющие в константной части только замены; 3) штрих-коды, содержащие не более одной инсерции в константную или переменную часть, а также имеющие делецию, включающую не более четырех переменных нуклеотидов подряд; 4) штрих-коды, переменная часть которых отличается по меньшей мере по 10 позициям. Отобранные плазмиды с уникальными штрих-кодами смешивали в эквимольном соотношении, что позволило получить полностью охарактеризованную библиотеку штрих-кодов. Далее производили очистку смешанной библиотеки от бактериального эндотоксина на колонке EndoFree DNA Purification Kit («Qiagen», США) согласно инструкциям фирмы-производителя. Целевую ДНК после пересадки в изопропанол растворяли в стерильном TE-буфере без эндотоксина из набора и использовали для получения вирусного супернатанта.

Получение вирусного супернатанта. На клетках линии Phoenix-GP проводили трансфекцию стандартным методом преципитации с хлоридом кальция [24]. К клеткам добавляли смесь плазмид pHCMVC-VSV-G (R861), pMDLg/pRRE, pRSV-Rev, кодирующих структурные компоненты лентивируса, а также полученную библиотеку со штрих-кодами. Через 24 ч среду отбирали, центрифугировали 10 мин при 1500 об/мин и пропускали через фильтр с размером ячеек 0,45 мкм. Затем его центрифугировали в течение 3,5 ч при 18 000 об/мин. Осадок, содержащий вирусные частицы, растворяли в бессывороточ-

ной среде с 1% бычьего сывороточного альбумина (БСА) пятой фракции («Sigma», США), получая вирусный супернатант. Параллельно, в тех же условиях и с помощью тех же вспомогательных плазмид, получали супернатант лентивирусов, несущих ген *eGFP* вместо генетических штрих-кодов.

Определение вирусного титра. Часть объема вирусного супернатанта, содержащего вирусные с библиотекой штрих-кодов, отбирали и титровали, разбавляя полной питательной средой (ППС), содержащей αMEM с 10% эмбриональной телячьей сыворотки Hyclone («GE Life Sciences», США), 2 mM L-глутамин («ICN», Швейцария), 100 ед/мл пенициллина («Ферейн», Россия), 50 мкг/мл стрептомицина («Ферейн», Россия) с полибренном в концентрации 8 мкг/мл («Sigma», США). Было сделано шесть разведений с шагом 1 : 10. Титрованный вирусный супернатант добавляли к клеткам линии Phoenix-GP, из которых выделяли ДНК. Эту ДНК использовали для определения вирусного титра методом ПЦР в реальном времени. Титр определяли по калибровочной кривой, которую строили по шести разведениям библиотеки плазмид со штрих-кодами в диапазоне концентраций (5×10^7)...(5×10^2) копий/мкл. В каждую реакцию добавляли геномную ДНК неинфицированных клеток линии Phoenix-GP для того, чтобы уравнять эффективность реакции амплификации с плазмидами и ДНК, выделенной из инфицированных эукариотических клеток. Праймеры WPRE-F и WPRE-R и флуоресцентный зонд WPRE-probe были комплементарны участку WPRE, который содержали плазмиды, несущие штрих-коды (табл. 1). Реакцию проводили в приборе Applied Biosystems ABI PRISM 7500 («Thermo Fischer Scientific», США). По калибровочной кривой определяли количество прови-

Таблица 1. Праймеры и зонды, использованные в работе

Название	Последовательности праймеров (5'–3')
BC-PCR-Seq	CTGAATGATACGGCGACCAC
WPRE-F	CCTTGATAAATCCTGGTTGCTGTCT
WPRE-R	GGAAAGGAGCTGACAGGTGGT
WPRE-probe	R6G-TCAGGCAACGTGGCGTGGTGTG-BHQ1
ddPCR-BC-F	CATCGATACCGTCGACCTC
ddPCR-BC-R	GCGTCGTGTAGGGAAAGAG
ddPCR-BC-probe	R6G-TGCTGATTGTGCCTGGCTAGAAG-BHQ1
ddPCR-ACTB-F	GGCTTCTTGTCTTCTCTC
ddPCR-ACTB-R	CATGTCGTCCCAGTTGGTG
ddPCR-ACTB-probe	FAM-CCTGAAGTACCCCATCGAGCAC-RTQ1
BC-PCR-R	TCGTTGGGAGTGAATTAGCC
BC-PCR-F-nested	CACGACGCTCTCCGATCTN
BC-PCR-R-nested	GCTCTCCGATCTCTCGAGN

русов в зараженных клетках, вычисляли вирусный титр в инфекционных частицах на 1 мл вирусного супернатанта. Вирусный титр супернатанта с eGFP определяли с помощью проточной цитофлуориметрии, заражая клетки Phoenix-GP лимитирующими разведениями супернатанта.

Получение и маркирование ММСК. ММСК были получены от шести здоровых доноров (4 мужчины и 2 женщины в возрасте 20–47 лет, медиана 22 года). Костный мозг забирали во время эксфузии в отделении высокодозной химиотерапии и трансплантации костного мозга ФГБУ «НМИЦ гематологии» МЗ РФ после подписания донорами информированного согласия. ММСК получали из 8–10 мл костного мозга, как описано в предыдущих работах [25]. По достижении ММСК конfluence их пассировали по 10^5 клеток во флакон T25. Через сутки ППС заменяли на 5 мл свежей ППС, содержащей 100 мкл концентрата вирусного супернатанта и 4 мкг/мл полибрена («Sigma», США). Инкубировали 6 ч, после чего полностью меняли среду на свежую ППС. Эффективность маркирования, определенная в параллельном эксперименте с помощью лентивируса, содержавшего eGFP, составила $(27 \pm 6)\%$.

Клонирование и пассирование ММСК. Маркированные ММСК растили до достижения конfluence монослоя. Клетки промывали раствором Версена, а затем обрабатывали 0,025%-ным раствором трипсина («MP Biomedicals», США). Концентрацию клеток в одноклеточной суспензии определяли в камере Горяева. На следующий пассаж переносили по 10^5 ММСК. Уровень удвоения популяции ММСК определяли по методике, описанной ранее [26]. На каждом пассаже с помощью последовательных разведений получали суспензию, содержащую ~1 клетку на 200 мкл среды. Суспензию разносили на 60 ячеек двух 96-луночных планшетов из расчета 1 клетка (200 мкл) на лунку (клонирование). Клетки культивировали в среде α -MEM с 20% эмбриональной телячьей сыворотки с основным фактором роста фибробластов (любезно предоставлен к.б.н. М.Э. Гаспарян, лаборатория инженерии белка, ИБХ им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН [27]) в концентрации 5 нг/мл в течение 14 дней. Эффективность клонирования определяли с помощью распределения Пуассона [28]. Из оставшихся после пассирования и клонирования клеток выделяли ДНК. Ее использовали для определения среднего количества провирусов в расчете на один диплоидный геном (copy number variation, CNV) с помощью цифровой эмульсионной ПЦР («Bio-Rad», США), а также для определения штрих-кодов в общей клеточной популяции с помощью NGS.

Определение среднего количества провирусов на один диплоидный геном. CNV определяли с помощью цифровой эмульсионной ПЦР. Для этого собирали реакционную смесь, содержащую 3 мкл ДНК, разведенной до концентрации 2–20 мкг/мл, 11 мкл 2× смеси реагентов ddPCR™ Supermix for Probes («Bio-Rad», США), 450 нМ каждого из праймеров, 250 нМ каждого из зондов («Синтол», Россия), деионизированной воды до 22 мкл. Каждая реакционная смесь содержала две пары праймеров и два зонда. Первая пара праймеров (ddPCR-BC-F и ddPCR-BC-R) и соответствующий ей зонд (ddPCR-BC-probe) были комплементарны последовательности вектора (табл. 1). Вторая пара праймеров (ddPCR-ACTB-F и ddPCR-ACTB-R) и зонд (ddPCR-ACTB-probe) были комплементарны последовательности гена β -актина человека (табл. 1). Водно-капельная эмульсия была сгенерирована из каждой реакционной смеси с помощью Automated Droplet Generator («Bio-Rad», США). После этого следовала ПЦР на приборе C1000 Touch™ Thermal Cycler («Bio-Rad», США) со следующими параметрами: 95 °C – 5 мин; 40 циклов: 95 °C – 30 с, 60 °C – 1 мин; охлаждение и инкубация при 4 °C – 5 мин; запекание капель при 90 °C – 5 мин. Количество положительных капель в каждом образце после ПЦР было подсчитано с помощью ридера капель QX200 AutoDG Droplet Digital PCR System («Bio-Rad», США). CNV и абсолютную концентрацию мишеней определяли с помощью программного обеспечения QuantaSoft v.1.7.4 («Bio-Rad», США).

Анализ штрих-кодов в клонах ММСК. Из лунки, содержащих клеточные колонии, выделяли ДНК, которую использовали как матрицу для ПЦР. В качестве праймеров применяли олигонуклеотиды ddPCR-BC-F и BC-PCR-R, комплементарные областям провируса, фланкирующим штрих-код (табл. 1). Результаты анализировали с помощью гель-электрофореза в 1,5%-ном агарозном геле. Полоску геля с ПЦР-продуктом расчетной длины вырезали скальпелем и очищали на колонке Wizard SV Gel and PCR Clean-Up System («Promega», США) согласно инструкции фирмы-производителя. Нуклеотидную последовательность очищенного ПЦР-продукта определяли с помощью секвенирования по Сэнгеру. Если полученная последовательность содержала несколько штрих-кодов, которые невозможно было различить, то исходную ДНК клонов повторно анализировали с помощью двухстадийной ПЦР. В этом случае на первой стадии использовали те же праймеры, что и при одностадийной ПЦР, а на второй стадии один из праймеров был сдвинут внутрь вплотную к последо-

вательности штрих-кода («полугнездовая ПЦР»). Для второй стадии реакции продукт первой стадии разводили в 250 раз деионизованной водой и брали 2 мкл на реакционную смесь. Готовили четыре варианта реакционной смеси, содержавшие прямой праймер BC-PCR-F-nested (табл. 1), где N – один из четырех нуклеотидов, комплементарный первому нуклеотиду штрих-кода, а обратный праймер BC-PCR-R – тот же, что и в первой стадии ПЦР (или наоборот: прямой праймер ddPCR-BC-F – тот же, что в первой стадии, а обратный праймер BC-PCR-R-nested сдвинут в сторону штрих-кода, последний нуклеотид на 3'-конце комплементарен последнему вариабельному нуклеотиду штрих-кода). Результаты анализировали аналогично одностадийной ПЦР.

Анализ штрих-кодов в суммарной популяции ММСК. Из общей клеточной популяции ММСК каждого пассажа выделяли ДНК и секвенировали совокупность штрих-кодов с помощью метода глубокого секвенирования на платформе Illumina. ДНК выделяли стандартным способом с осаждением этанолом [23]. По 50 нг выделенной ДНК было использовано для определения последовательности штрих-кодов с помощью NGS на платформе Illumina MiSeq. Области ДНК, содержавшие штрих-код в составе интегрированного провируса, амплифицировали с помощью ПЦР. Для этого использовали готовую реакционную смесь NEBNext Ultra II («New England Biolabs, Inc.», США) с полимеразой Q5 в своем составе (12,5 мкл на реакцию), универсальный (2,5 мкл на реакцию) и одиночные индексные праймеры (2,5 мкл на реакцию) из наборов NEBNext Multiplex Oligos for Illumina® («New England Biolabs, Inc.», США) с добавлением раствора ДНК (7,5 мкл на реакцию). ПЦР проводили на приборе MJ Mini («Bio-Rad», США) по следующей программе: 98 °C – 30 с; 25 циклов: 98 °C – 10 с, 60 °C – 30 с, 72 °C – 30 с; 72 °C – 5 мин. Затем продукты ПЦР очищали с помощью магнитных частиц Ampure XP («Beckman-Coulter», США) в соотношении 1,2 : 1 (объем частиц к объему ДНК). Длину продуктов ПЦР определяли на приборе Bioanalyzer 2100 («Agilent», США) с помощью набора реагентов DNA High Sensitivity Kit («Agilent», США). После очистки концентрацию ДНК измеряли на флуориметре Qubit с использованием набора реагентов dsDNA HS Assay Kit («Thermo Fisher Scientific», США) и объединяли в равных количествах. После объединения концентрацию образца измеряли, доводили до 2 нМ и далее проводили подготовку образца к секвенированию согласно рекомендациям производителя прибора. Для секвенирования использовали реагенты Miseq

Reagent Kit v.3 (150 циклов) в режиме парных чтений. Для прочтения индексов использовали отдельное прочтение. Длина смыслового прочтения составляла 80 нуклеотидов. Перевод исходных данных в формат fastq проводили с помощью программы bcl2fastq2 («Illumina», США). Число прочтений на каждый образец (глубина секвенирования) приведено в табл. S1 (см. Приложение).

Скрипты, использованные в ходе дальнейшей биоинформатической обработки, доступны по адресу: <http://makarich.fbb.msu.ru/kasianov/Bigildeev/>.

На первом этапе производили соотнесение прочитанных индексов образцам ДНК. Далее проводили кластеризацию смысловых чтений по штрих-кодам. Для этих целей применяли скрипт Cluster.perl. На вход он принимал набор чтений в файле формата fastq. Перечень использованных при кластеризации референсных штрих-кодов приведен в табл. S2 (см. Приложение). Кластеризацию по штрих-кодам проводили в трех режимах: точное совпадение подпоследовательности чтения с референсным штрих-кодом, совпадение подпоследовательности чтения с референсным штрих-кодом с учетом одной замены, совпадение подпоследовательности чтения с референсным штрих-кодом с учетом шести замен. Количество прочтений всех кластеризованных последовательностей суммировали для каждого референсного штрих-кода. Результаты кластеризации собирали в одну таблицу для всех образцов с использованием скрипта ConcatBarco-desReadCountsFile.perl. Результаты объединения приведены в табл. S3, S4 и S5 (см. Приложение). Для анализа представленности штрих-кодов в суммарной клеточной популяции ММСК использовались результаты кластеризации с шестью заменами.

Зная количество клеток, ДНК которых была использована для секвенирования, и количество классифицированных чтений в соответствующем образце, вычисляли долю чтений (вклад), соответствующую одной клетке. Сравнивая с этим значением вклад каждого штрих-кода, определяли, какому количеству клеток он соответствует. Условно все штрих-коды были разделены на три категории: 1) вклад соответствует <8 клеткам; 2) вклад соответствует 8–32 клеткам; 3) вклад соответствует >32 клеткам.

Статистический анализ. Статистическую значимость различий оценивали по критерию Манна–Уитни. Для сравнения кумулятивной клеточной продукции маркированных и немаркированных ММСК от тех же доноров использовали тест Вилкоксона для сопряженных пар. Различия между группами считали статистически значимыми при $p < 0,05$.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Дизайн эксперимента. Общую схему эксперимента можно представить следующим образом (рис. 1). В культуральный флакон добавляли вирусную библиотеку на следующие сутки после посадки 10^5 ММСК нулевого пассажа. Через 6 ч среду полностью меняли и продолжали культивировать ММСК до конfluентности. При пассировании культуры 120 клеток клонировали по одной клетке на лунку 96-луночного планшета, 10^5 клеток возвращали во флакон, из оставшихся клеток выделяли ДНК. Эту процедуру проводили на первом, втором, третьем и четвертом пассажах культуры каждого донора костного мозга.

Библиотека штрих-кодов. Для построения библиотеки плазмид со штрих-кодами было проанализировано 1876 трансформантов *E. coli*. Было отмечено, что 70% трансформантов содержали плазмиду со вставкой нужной длины. Только такие колонии использовали для последующих этапов создания библиотеки. У 867 клонированных и выделенных плазмид была успешно установлена нуклеотидная последовательность штрих-кода. Таким образом, эффективность метода получения клонированных плазмид из общей смеси составила 46%. На основании критериев качества и разнообразия, описанных в разделе «Методы исследования», было отобрано 667 плазмид для создания библиотеки. Среднее расстояние Хэмминга между штрих-кодами полученной библиотеки было

равно 26 ± 3 при том, что общее количество вариативных нуклеотидов в штрих-коде составляло 35. Такой набор гарантировал уникальность каждого штрих-кода и простоту его отличия от других при анализе. Плазмиды, содержавшие отобранные штрих-коды, были смешаны в равной пропорции, так что каждый штрих-код в общей библиотеке имел одну и ту же частоту встречаемости.

Определение вирусного титра и множественности заражения. Титр вируса, оцененный с помощью ПЦР, составил $(6,0 \pm 2,7) \times 10^6$ инфекционных частиц/мл. Параллельно в тех же условиях ММСК инфицировали тем же вектором, содержащим ген eGFP вместо штрих-кодов. В этом случае титр определяли методом проточной цитофлуориметрии. Титр вируса, определенный этим способом, оказался равным $(55 \pm 11) \times 10^6$ инфекционных частиц/мл. Поскольку титр был определен двумя независимыми методами с различной чувствительностью, можно предполагать, что истинный титр находился между двумя этими значениями.

Опираясь на полученные значения вирусного титра, была вычислена множественность заражения (multiplicity of infection, MOI). Для заражения культуры ММСК использовали 100 мкл вирусного супернатанта с концентрацией $\sim(6 \times 10^6) \dots (55 \times 10^6)$ инфекционных частиц/мл; следовательно, суммарно в культуру добавили $(6-55) \times 10^5$ инфекционных частиц. Количество клеток в культуре составляло $\sim 10^5$.

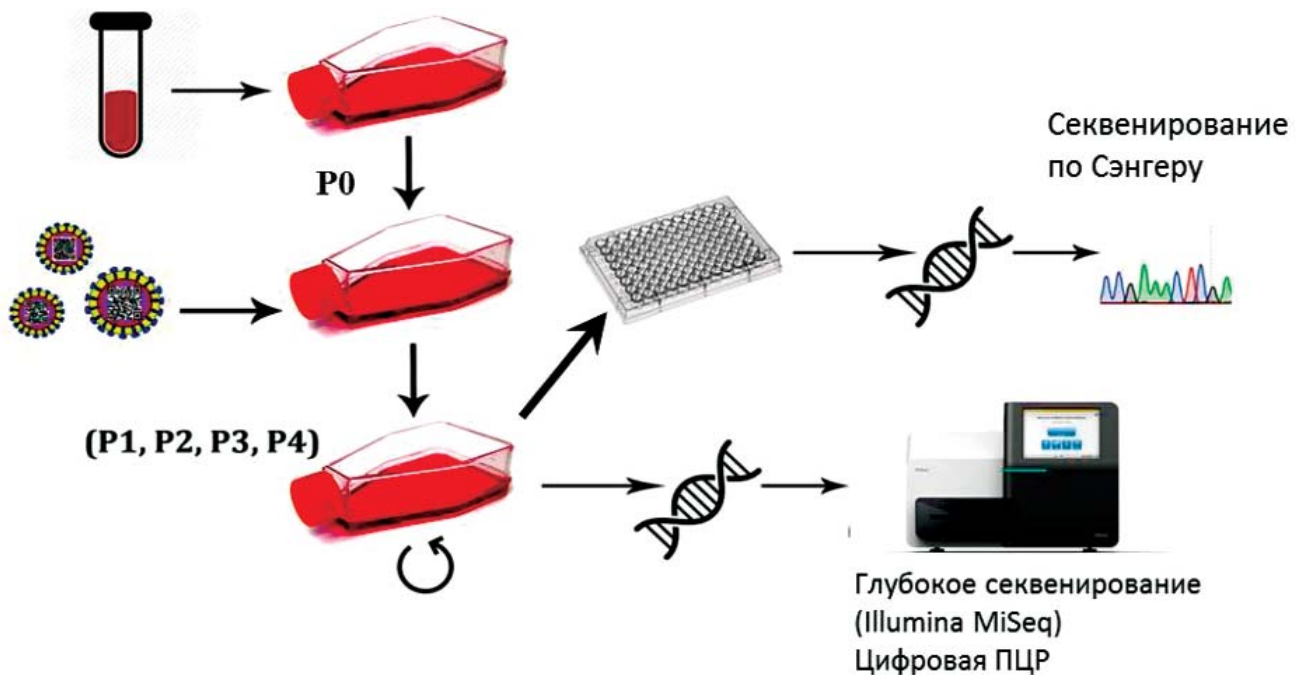


Рис. 1. Схема эксперимента. Обозначения: P0–P4 – номера пассажей.

С цветным вариантом рис. 1 можно ознакомиться в электронной версии статьи на сайте: www.elibrary.ru

Следовательно, МОИ находилась в диапазоне 6–55. Высокое значение МОИ было подтверждено цифровой капельной ПЦР, которую использовали для определения среднего количества провирусов в расчете на диплоидный геном (CNV) в популяции маркированных ММСК на каждом пассаже культуры (рис. 2).

Было показано, что все исследованные в работе культуры содержат маркер. CNV уменьшалось от первого пассажа (медиана 17,8) к четвертому (медиана 2,2). По нашему предположению это обусловлено тем, что те клетки, в которых произошли множественные интеграции провирусов, погибают или утрачивают способность к пролиферации и поэтому достаточно быстро элиминируются из культуры, в отличие от тех клеток, в которых количество интеграций было небольшим.

Ростовые характеристики ММСК, эффективность клонирования и маркирования. Процедура маркирования ММСК оказывает влияние на их способность к пролиферации, что выражается в достоверном ($p = 0,04$) уменьшении суммарной клеточной продукции к третьему пассажу по сравнению с контрольными ММСК (рис. 3, а). Результат согласуется с ранее полученными данными, указывающими на то, что стрессовое воздействие в виде вирусной трансдукции может негативно сказываться на росте ММСК [29].

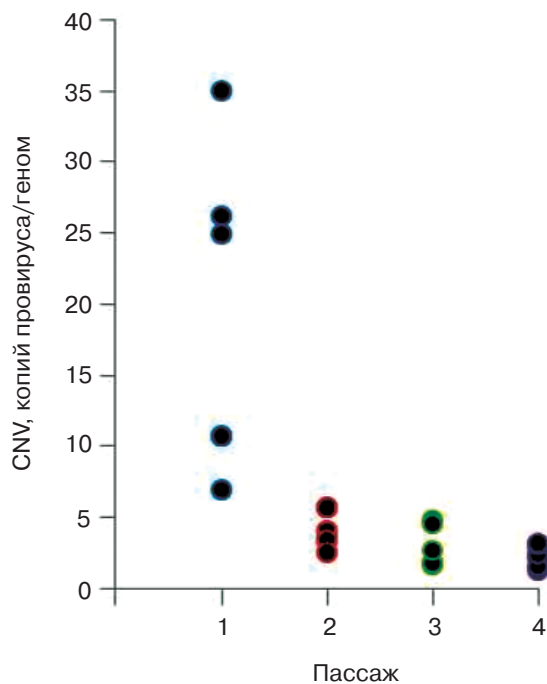


Рис. 2. Динамика изменения среднего количества провирусов в расчете на диплоидный геном (CNV). С цветным вариантом рис. 2 можно ознакомиться в электронной версии статьи на сайте: www.elibrary.ru

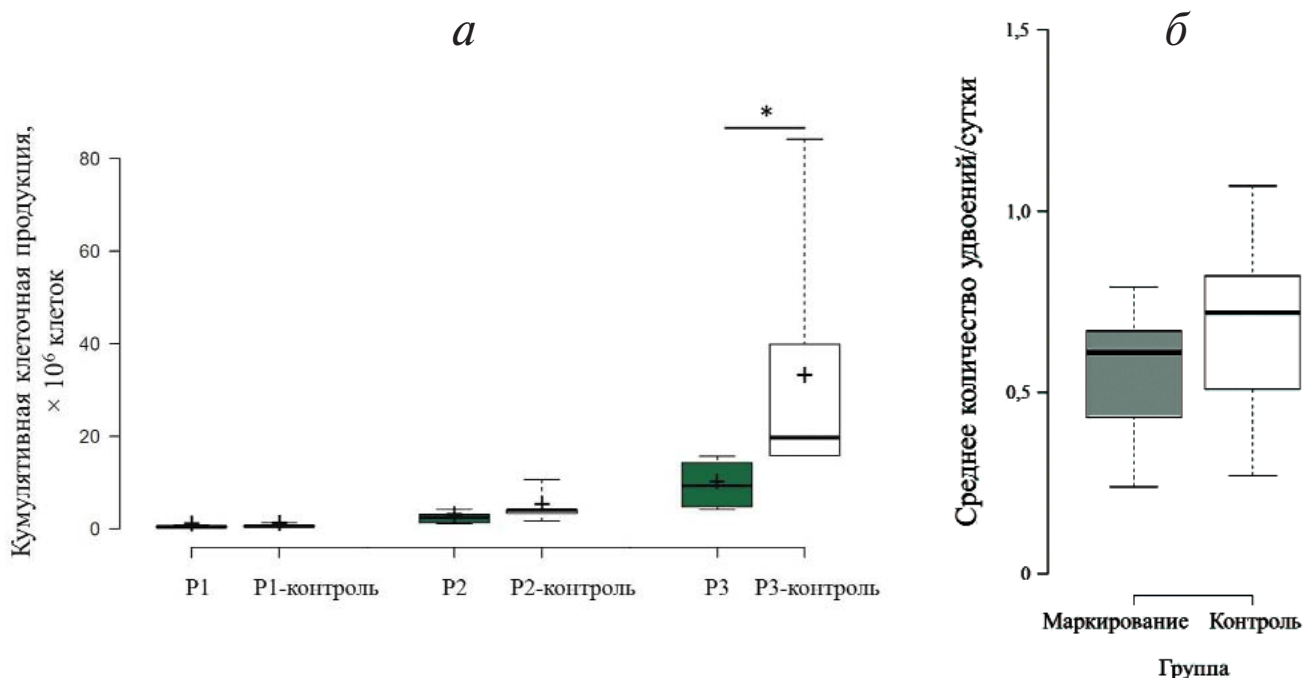


Рис. 3. Сравнение кумулятивной клеточной продукции (а) и среднего уровня пролиферации (б) маркированных и контрольных ММСК на 1–3-ем пассажах. Указан разброс значений, медианы и межквартильный диапазон Q3–Q1. Плюсы обозначают среднее значение. Звездочка указывает на статистически значимые различия по парному тесту Вилкоксона. С цветным вариантом рис. 3 можно ознакомиться в электронной версии статьи на сайте: www.elibrary.ru

При анализе среднего количества клеточных удвоений отмечалась тенденция к снижению этого параметра у маркированных ММСК, однако различия были статистически незначимыми в среднем для трех пассажей (рис. 3, б). Третий пассаж был проанализирован отдельно, но здесь также не наблюдали статистически значимых различий, вероятно, в силу небольшой выборки (данные не представлены). По мере пассирования наблюдалось снижение эффективности клонирования (рис. 4) маркированных ММСК аналогично тому, как это происходило для немаркированных ММСК в предыдущих экспериментах [29]. При этом пропорция маркированных клеток практически не изменялась по мере пассирования (табл. 2).

Учитывая результаты, полученные по среднему количеству провирусов в расчете на один диплоидный геном, можно заключить, что процедура переноса генетического маркера может негативно сказываться на пролиферации клеточной культуры, но это влияние ограничивается теми клетками, в ДНК которых интегрировалось значительное количество провирусов (более четырех).

Анализ штрих-кодов в клонах ММСК. Проанализировано 458 маркированных клонов, в которых удалось выявить 381 уникальный штрих-код из 667 штрих-кодов (57%), исходно присутствовавших в библиотеке. Полученные результаты свидетельствуют о том, что разнообразие штрих-кодов в вирусном супернатанте меньше, чем в исходной смеси плазмид. Вероятно, это объясняется стохастической природой процесса трансфекции клеток пакующей линии, процессов сборки вирионов и последующей трансдукции ММСК. Более того, анализ частоты встречаемости штрих-кодов в клеточных колониях-клонах ММСК от всех доноров показал, что некоторые штрих-коды встречаются чаще, чем другие (рис. 5, а). Несмотря на то что большинство штрих-кодов было обнаружено 1–5 раз, некоторые из них встречались >10 раз. Подобную зави-

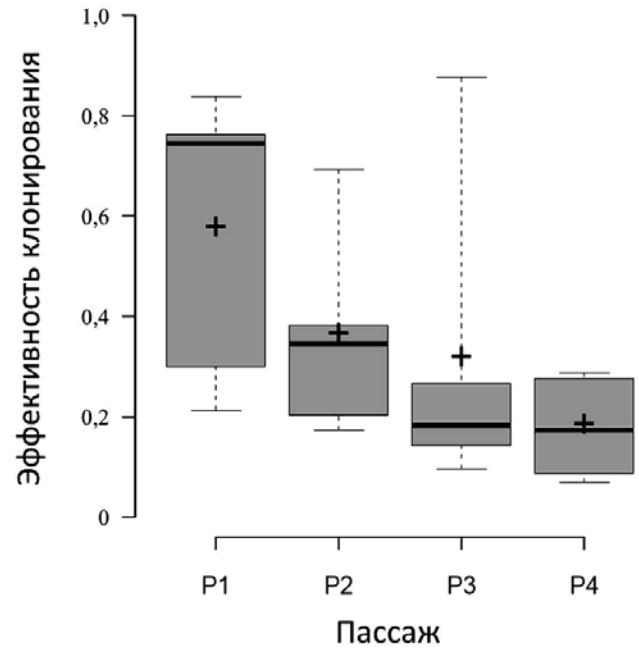


Рис. 4. Эффективность клонирования маркированных ММСК. Указан разброс значений, медианы и межквартильный диапазон Q3–Q1. Плюсы обозначают среднее значение

симость наблюдали при анализе штрих-кодов, обнаруженных в ММСК каждого донора по отдельности (данные не приведены).

Три наиболее часто встречающихся штрих-кода (штрих-коды № 332, 611 и 968) были суммарно обнаружены по 13 раз среди всех исследованных ММСК. При анализе ММСК отдельных доноров именно эти штрих-коды были наиболее часто встречающимися в ММСК каждого донора. Такое распределение частоты встречаемости может быть обусловлено несколькими причинами. Например, оно может свидетельствовать о том, что в вирусном супернатанте нарушена равномерность представленности штрих-кодов (вирионы с определенными штрих-кодами встречаются чаще, чем другие). Вероятно, это связано со стохастической природой процесса

Таблица 2. Доля маркированных клонов в ММСК

Донор	Пассаж 1	Пассаж 2	Пассаж 3	Пассаж 4	Всего для донора
250	12/22 (55%)	12/22 (55%)	15/28 (54%)	14/22 (64%)	53/94 (56%)
251	29/39 (74%)	28/46 (61%)	4/10 (40%)	11/19 (58%)	72/114 (63%)
252	45/61 (74%)	18/19 (95%)	8/15 (53%)	3/7 (43%)	74/102 (73%)
253	54/59 (92%)	18/21 (86%)	17/22 (77%)	8/9 (89%)	97/111 (87%)
254	49/62 (79%)	43/53 (81%)	52/65 (80%)	18/29 (62%)	162/209 (78%)
256*	N/A	N/A	N/A	N/A	N/A

* ММСК донора 256 не были клонированы. Для каждого донора на каждом пассаже культуры указано количество маркированных клонов/общее количество клонов (процент маркированных клонов).

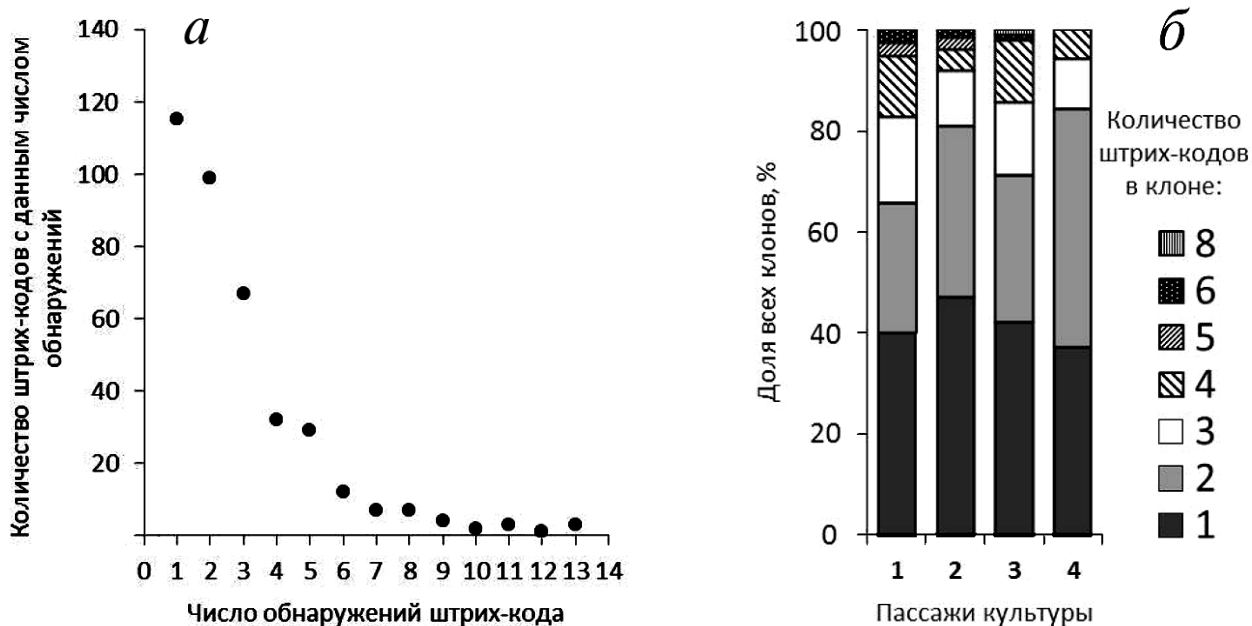


Рис. 5. а – Распределение штрих-кодов по частоте встречаемости. По горизонтали отложено, сколько раз штрих-код был встречен в клонах ММСК (суммарно по всем культурам и пассажам). По вертикали отложено число штрих-кодов, которые были обнаружены соответствующее число раз; б – распределение клеточных клонов по количеству штрих-кодов (интеграций провирусов)

трансфекции клеток-продуцентов и продукции ими вирусных частиц. Сужение разнообразия штрих-кодов и искажение равномерности их представленности в вирусном супернатанте по сравнению с библиотекой плазмид является неизбежным недостатком применения лентивирусных векторов. Степень сужения репертуара штрих-кодов и искажения представленности каждого из них остается открытым вопросом и требует детального изучения, поскольку может влиять на выводы исследователей, применяющих подобные методики. В этом исследовании 90% штрих-кодов было детектировано не более 5 раз. Данные результаты позволили заключить, что большинство штрих-кодов представлено равномерно, т.к. частота их встречаемости невелика, а увеличенная частота встречаемости некоторых из них может объясняться наличием в ММСК каждого донора нескольких клонов – потомков одной маркированной клетки, обладающей высоким пролиферативным потенциалом.

Интересно, что распределение клеточных клонов по количеству интеграций провирусов было практически одинаковым на всех исследованных пассажах (рис. 5, б). В подавляющем большинстве клонов было обнаружено от одного до четырех штрих-кодов. Доля клонов с большим количеством штрих-кодов была мала и снижалась к поздним пассажам до 0, что еще раз указывает на то, что те клетки, в ДНК которых

встроилось более четырех провирусов, утрачивают способность к пролиферации и элиминируются из культуры.

Оценка уникальности маркирования. Степень уникальности маркирования клеточной популяции штрих-кодами может быть предсказана с помощью биномиального распределения [12]:

$$P(k, n) = \frac{n!}{(n-k)! k!} p^k (1-p)^{n-k}, \quad (1)$$

где P – вероятность того, что все клетки будут маркированы уникально, k – количество маркированных клоногенных клеток в тестируемой популяции, n – количество уникальных штрих-кодов в библиотеке, p – вероятность «успеха», т.е. того, что клетка будет маркирована конкретным штрих-кодом.

Оценим вероятность того, что каждая клетка будет уникально маркирована для донора № 250 (репрезентативный образец). Пусть в культуре на момент заражения было 100 000 клеток. Эффективность клонирования ММСК донора № 250 на первом пассаже (наиболее близком к моменту маркирования) равнялась 0,21 (табл. S6, Приложение), а средняя эффективность маркирования ММСК этого донора составляла 56% (табл. 2). Таким образом, исходная культура должна содержать ~11 700 маркированных клоногенных клеток (k) в популяции ($100\,000 \times 0,21 \times 0,56$). Разнообразие штрих-кодов в библиотеке $n = 667$,

а вероятность «успеха» равна $1/n = 1/667$. Подставив значения n и k в формулу (1), получим, что вероятность того, что все клоны маркированы уникально при интеграции одного штрих-кода, стремится к нулю.

Следовательно, если были встречены два клона, несущие каждый по одному одинаковому штрих-коду, то их считали потомками разных маркированных клеток. Истинное разнообразие библиотеки увеличивается за счет того, что одна клетка может быть маркирована более чем одним штрих-кодом. Предположим, что клетка маркирована двумя штрих-кодами. Число уникальных комбинаций можно оценить с помощью формулы комбинаций m элементов из множества n элементов с повторениями:

$$\bar{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m! \cdot (n-1)!} \quad (2)$$

Приняв $n = 667$ и $m = 2$, получим, что количество комбинаций составит >225 тыс. Оценив уникальность маркирования по формуле биномиального распределения (1), получим, что вероятность того, что все клоны маркированы уникально, стремится к 1. Таким образом, приобретение клеткой двух штрих-кодов позволяет обеспечить уникальность маркирования. Следовательно, если в эксперименте выявлялись два клеточных клона, совпадающих по двум (и более) штрих-кодам, такие клеточные клоны считали произошедшими из одной исходно трансдуцированной клетки. Всего в работе было обнаружено пять таких клонов (рис. 6).

Существует вероятность того, что в одну ячейку планшета при клонировании попадает более одной клоногенной клетки. В результате в ячейке образуется популяция из нескольких маркированных клонов, что при анализе будет выглядеть как один клон, содержащий несколько штрих-кодов. В силу этого анализ ДНК, выделенной из двух ячеек, содержащих одинаково маркированные клоны плюс дополнительные клеточные колонии, будет давать частично перекрывающиеся наборы штрих-кодов. Если эти наборы перекрывались по двум и более штрих-кодам, то считали, что в той ячейке, где обнаруживались дополнительные штрих-коды, присутствует несколько клеточных клонов, один из которых несет общие штрих-коды с клоном из первой ячейки, а остальные несут любые другие комбинации.

Клональный состав ММСК. В клеточных клонах, полученных на 1–4-м пассажах ММСК каждого донора, было проведено сравнение штрих-кодов. С учетом соображений, описанных в предыдущем разделе, были выявлены клоны — потомки одних и тех же маркированных

клеток. Результаты анализа представлены на рис. 6. Большинство клонов встречается один раз на протяжении всех четырех пассажей культуры. Полученные данные свидетельствуют о том, что ММСК поддерживаются, в основном, за счет клонов с небольшим пролиферативным потенциалом, который исчерпывается в течение одного пассажа культуры. В результате на каждом следующем пассаже клональный состав ММСК существенно изменяется, в пролиферацию вступают те клоны, которые на предыдущем пассаже не делились и не вносили большого вклада в общую клеточную популяцию. В ММСК доноров 250 и 254 было обнаружено по одной паре одинаковых клонов на разных пассажах культуры (отмечены красным). Ранее было показано, что ММСК могут содержать клоны, сохраняющиеся в культуре на протяжении нескольких пассажей [10]. Динамика их вклада в общую клеточную популяцию может быть различной. Этим можно объяснить то, что, например, у донора № 250 одинаковые клоны (такowymi считали клоны, совпадающие по двум и более штрих-кодам) были обнаружены только на первом и четвертом пассажах, а на втором и третьем их выявить не удалось.

В ММСК донора 254 было обнаружено три пары одинаково маркированных клонов на первом пассаже культуры. Они соответствуют таким клонам, которые вносят существенный вклад в общую клеточную популяцию. Только в этом случае вероятность попадания нескольких их представителей в 120 клонированных клеток могла быть высокой. Ранее было показано, что в ММСК встречаются клоны, вклад которых в общую клеточную популяцию составляет 10–40% [10]. Тот факт, что такие клоны были обнаружены только у одного из пяти исследованных доноров, свидетельствует о том, что количество клонов с высоким пролиферативным потенциалом невелико по сравнению с общим количеством клонов на каждом пассаже культуры. Частота встречаемости таких клонов (например, для донора № 250) можно оценить как отношение маркированных клоногенных клеток к общему количеству клоногенных клеток в пассаже, которое было оценено ранее, или $60/11\ 000 = 0,005$. Наличие и количество клонов с высоким пролиферативным потенциалом может быть связано с концентрацией МСК в исследуемом костном мозге. Полученные данные еще раз подтверждают то, что ММСК не идентичны МСК.

Анализ штрих-кодов в общей клеточной популяции ММСК. На каждом пассаже культуры из суммарной популяции клеток, оставшихся после пассирования и клонирования, выделяли ДНК и анализировали вклад штрих-кодов. Репрезента-

тивное распределение вклада штрих-кодов для ММСК донора 254 на первом пассаже приведено на рис. 7.

Из рис. 7 видно, что NGS выявило подавляющее большинство штрих-кодов исходной библиотеки. В каждом образце выявлялось >600 из исходных 667 штрих-кодов, при этом >550 штрих-кодов имели ≥5 прочтений. Эти данные в сово-

купности с данными цифровой капельной ПЦР по оценке количества провирусов в расчете на один диплоидный геном позволяют сделать однозначный вывод о поликлональности культуры. Действительно, цифровая ПЦР продемонстрировала, что в среднем по всем пассажам и всем донорам на один геном приходится 9 провирусов. Одновременно с этим, NGS показало,

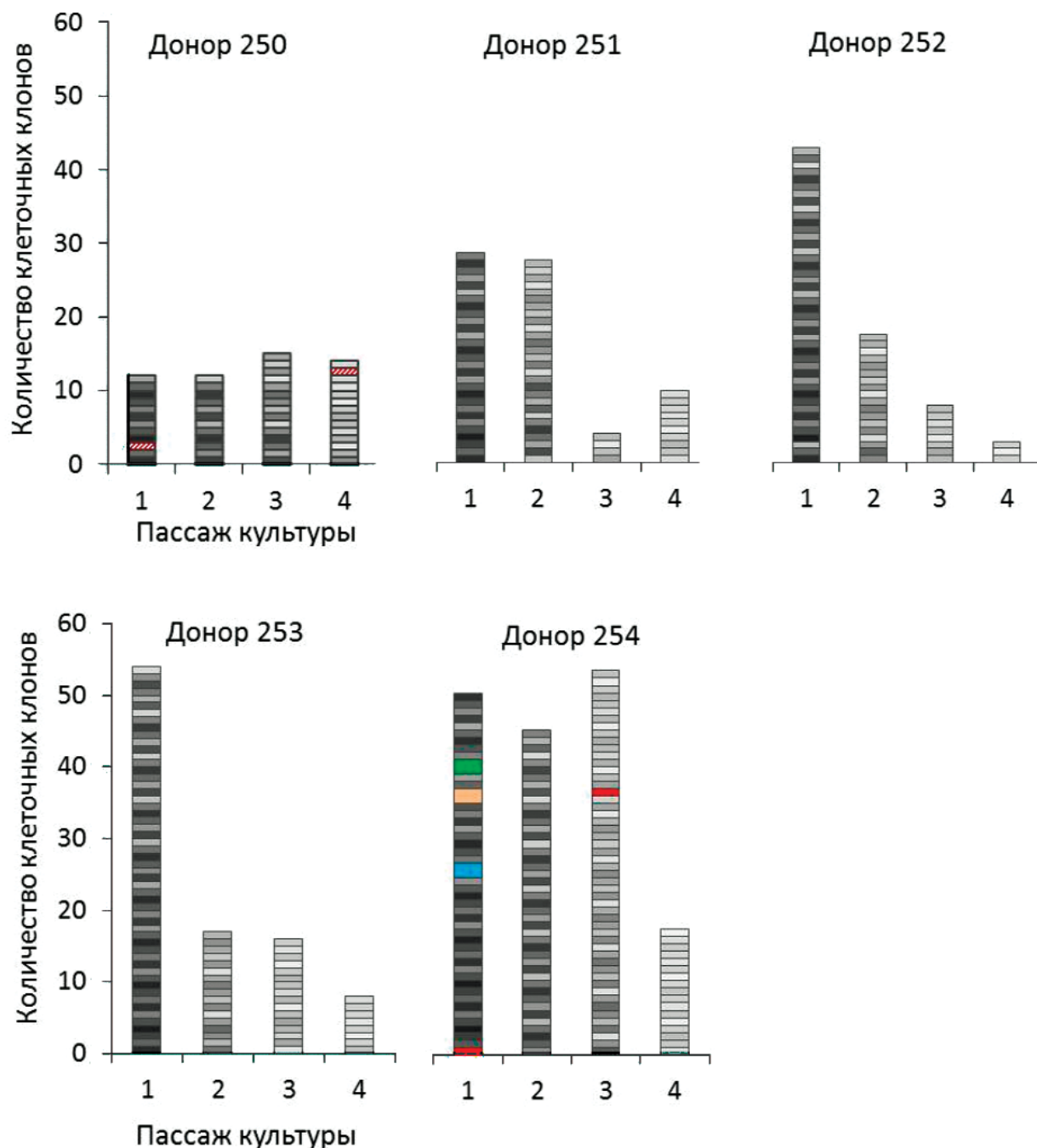


Рис. 6. Клональный состав ММСК. Каждый прямоугольник – клеточный клон. Серые прямоугольники обозначают такие клеточные клоны, которые у данного донора встретились всего один раз (независимо от пассажа). Цветные прямоугольники обозначают клоны, встретившиеся более одного раза

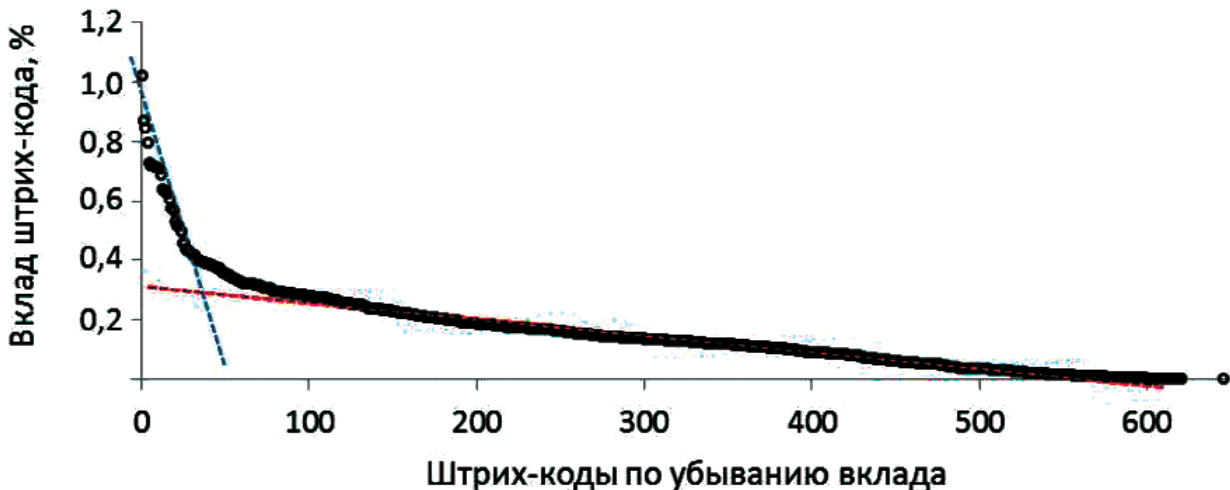


Рис. 7. Распределение вклада штрих-кодов в общую клеточную популяцию ММСК донора 254 на первом пассаже культуры. По вертикали – вклад штрих-кода, представленный как отношение количества прочтений данного штрих-кода к количеству классифицированных прочтений, выраженное в процентах. По горизонтали – штрих-коды, расположенные по вкладу от большего к меньшему.

С цветным вариантом рис. 7 можно ознакомиться в электронной версии статьи на сайте: www.elibrary.ru

что в любом образце идентифицируется несколько сотен штрих-кодов. Такое соотношение неизбежно означает, что культура состоит из множества одновременно функционирующих клонов.

Медиана вклада каждого отдельного штрих-кода составила 0,013%. Из рис. 7 видно, что среди всего множества выделяется группа из нескольких десятков штрих-кодов, вклад которых значительно выбивается из общего распределения и превышает 0,4%. Наличие такого излома можно объяснить несколькими способами. Эти штрих-коды могут соответствовать совокупности клеточных клонов, которые давали относительно большой вклад в общую клеточную популяцию ММСК на данном пассаже. В пользу этой версии говорит то, что все одинаковые клоны, обнаруженные на первом пассаже культуры этого донора, содержали штрих-коды, входящие в топ-35 по своему вкладу. Таким образом, результаты, полученные при секвенировании штрих-кодов в отдельных клонах, и данные NGS коррелируют между собой.

Возможно альтернативное объяснение. Наиболее часто встречающимися могут быть штрих-коды, чья представленность в исходном вирусном супернатанте была больше. Следовательно, они интегрировались в сравнительно большее количество клеток при маркировании культуры, если предположить, что вероятность интеграции провирусной ДНК не зависит от нуклеотидной последовательности штрих-кода. В пользу этой версии говорит то, что одни и те же штрих-коды входят в топ-10 штрих-кодов у нескольких доноров одновременно. Из 10 штрих-кодов, об-

ладающих максимальным вкладом в ММСК донора № 250 на первом пассаже, один штрих-код (№ 967) входил в 10 наиболее представленных на первом пассаже штрих-кодов у всех проанализированных доноров, еще пять штрих-кодов (№ 31, 42, 463, 669, 699) входили в топ-10 у четырех из пяти доноров, а остальные четыре штрих-кода также встречались среди топ-10 у более чем одного донора. Из этого следует, что в исходном общем вирусном супернатанте, которым заражали ММСК всех доноров, вирусов с данными штрих-кодами было относительно больше, чем всех остальных. По мере пассирования вклад каждого отдельного штрих-кода меняется. Было показано, что количество штрих-кодов, вносящих существенный вклад в общую популяцию, невелико и уже ко второму пассажу определяется как десятые доли процента (рис. 8).

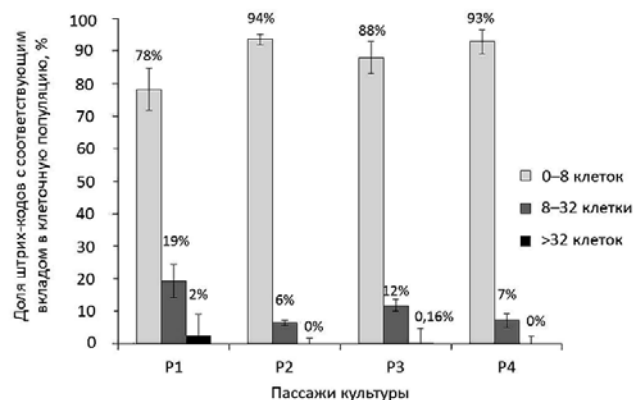


Рис. 8. Динамика распределения штрих-кодов, обладающих вкладом, соответствующим различному количеству клеток популяции ММСК

Максимальный вклад штрих-кодов соответствовал ~100 клеткам, такие штрих-коды были выявлены у половины проанализированных доноров. У одного донора (Д250) их было 90, у оставшихся — шесть (Д254) и один (Д251). Учитывая, что для секвенирования было взято количество ДНК, соответствующее ~8000 клеток, можно оценить максимальная представленность клеточных клонов в культуре. В этом эксперименте оно составило $100/8000 = 1,2\%$. Один штрих-код может присутствовать в нескольких клонах, поэтому реальная представленность клонов в культуре ММСК может быть ниже, что подтверждается полученными ранее данными [10]. Подавляющее большинство штрих-кодов на всех пассажах выявлялось в ≤ 8 клетках.

Описанные в работе сложности, связанные с уникальностью маркирования и искажением представленности штрих-кодов в библиотеке, указывают на то, что необходимо продолжать поиск новых стратегий по маркированию ММСК. В их основе должны лежать плазмидные библиотеки, в которых достигается баланс между разнообразием и точностью характеристики библиотеки. Более того, желательна анализировать лентивирусную библиотеку по разнообразию и представленности штрих-кодов непосредственно перед ее использованием.

Применение техники маркирования ММСК с помощью генетических штрих-кодов и последующего глубокого секвенирования маркеров подтвердило наличие поликлональности в ММСК человека, которая была установлена ранее с помощью ПЦР, опосредованной лигированием [10]. При этом в данной работе было проанализировано существенно большее количество клонов и выявлено пропорционально большее их разнообразие. Полученные результаты означают, что в популяции ММСК реальное количество кло-

нов может быть еще больше. Значительные изменения клонального состава популяции от пассажа к пассажи подтверждают ограниченный пролиферативный потенциал большинства представителей ММСК. По-прежнему нельзя исключить, что в состав этой популяции не входят клетки, близкие по своим характеристикам к стволовым. ММСК в основном представляют собой гетерогенную популяцию стромальных предшественников с ограниченной способностью к пролиферации, но сохраняющих плюрипотентность.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Соблюдение этических норм

Все процедуры, выполненные в исследовании с участием людей, соответствуют этическим стандартам институционального комитета по исследовательской этике и Хельсинской декларации 1964 года и ее последующим изменениям и сопоставимым нормам этики. От каждого из включенных в исследование участников было получено информированное добровольное согласие.

Благодарности

Авторы благодарят проф. Б. Фезе и К. Корнилс (University Medical Center Hamburg-Eppendorf, Гамбург, Германия) за предоставленную библиотеку плазмид, содержащих штрих-коды, и линию-продуцент для получения лентивирусного вектора.

СПИСОК ЛИТЕРАТУРЫ

1. Friedenstein, A.J., Chailakhjan, R.K., and Lalykina, K.S. (1970) The development of fibroblast colonies in monolayer cultures of guinea-pig bone marrow and spleen cells, *Cell Tissue Kinet.*, **3**, 393–403, doi: 10.1111/j.1365-2184.1970.tb00347.
2. Caplan, A.I. (1991) Mesenchymal stem cells, *J. Orthop. Res.*, **9**, 641–650, doi: 10.1002/jor.1100090504.
3. Horwitz, E.M., Le Blanc, K., Dominici, M., Mueller, I., Slaper-Cortenbach, I., Marini, F.C., Deans, R.J., Krause, D.S., Keating, A., and International Society for Cellular Therapy. (2005) Clarification of the nomenclature for MSC: The International Society for Cellular Therapy position statement, *Cytotherapy*, **7**, 393–395, doi: 10.1080/14653240500319234.
4. Dominici, M., Le Blanc, K., Mueller, I., Slaper-Cortenbach, I., Marini, F., Krause, D., Deans, R., Keating, A., Prockop, D.J., and Horwitz, E. (2006) Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement, *Cytotherapy*, **8**, 315–317, doi: 10.1080/14653240600855905.
5. Jones, B.J. and McTaggart, S.J. (2008) Immunosuppression by mesenchymal stromal cells: from culture to clinic, *Exp. Hematol.*, **36**, 733–741, doi: 10.1016/j.exphem.2008.03.006.
6. Spees, J.L., Lee, R.H., and Gregory, C.A. (2016) Mechanisms of mesenchymal stem/stromal cell function, *Stem Cell Res. Ther.*, **7**, 125, doi: 10.1186/s13287-016-0363-7.
7. Pittenger, M.F., Mackay, A.M., Beck, S.C., Jaiswal, R.K., Douglas, R., Mosca, J.D., Moorman, M.A., Simonetti, D.W., Craig, S., and Marshak, D.R. (1999) Multilineage potential of adult human mesenchymal stem cells, *Science*, **284**, 143–147, doi: 10.1126/science.284.5411.143.

8. Palau, P., Dominguez, E., Lopez, L., Heredia, R., Gonzalez, J., Ramon, J.M., Serra, P., Santas, E., Bodi, V., Sanchis, J., Chorro, F.J., and Nunez, J. (2016) Inspiratory muscle training and functional electrical stimulation for treatment of heart failure with preserved ejection fraction: rationale and study design of a prospective randomized controlled trial, *Clin. Cardiol.*, **39**, 433–439, doi: 10.1002/clc.22555.
9. Russell, K.C., Phinney, D.G., Lacey, M.R., Barrilleaux, B.L., Meyertholen, K.E., and O'Connor, K.C. (2010) *In vitro* high-capacity assay to quantify the clonal heterogeneity in trilineage potential of mesenchymal stem cells reveals a complex hierarchy of lineage commitment, *Stem Cells*, **28**, 788–798, doi: 10.1002/stem.312.
10. Bigildeev, A.E., Zhironkina, O.A., Shipounova, I.N., Sats, N.V., Kotyashova, S.Y., and Drize, N.J. (2012) Clonal composition of human multipotent mesenchymal stromal cells, *Exp. Hematol.*, **40**, 847–856, doi: 10.1016/j.exphem.2012.06.006.
11. Gerrits, A., Dykstra, B., Kalmykova, O.J., Klauke, K., Verovskaya, E., Broekhuis, M.J.C., de Haan, G., and Bystrykh, L.V. (2010) Cellular barcoding tool for clonal analysis in the hematopoietic system, *Blood*, **115**, 2610–2618, doi: 10.1182/blood-2009-06-229757.
12. Bystrykh, L.V., and Belderbos, M.E. (2016) Clonal analysis of cells with cellular barcoding: when numbers and sizes matter, *Methods Mol. Biol.*, **1516**, 57–89, doi: 10.1007/7651_2016_343.
13. Ребриков Д.В., Коростин Д.О., Шубина Е.С., Ильинский В.В. (2015) *NGS: высокопроизводительное секвенирование*, Москва, Лаборатория знаний.
14. Glimm, H., Ball, C.R., and von Kalle, C. (2011) You can count on this: barcoded hematopoietic stem cells, *Cell Stem Cell*, **9**, 390–392, doi: <https://doi.org/10.1016/j.stem.2011.10.013>.
15. Maetzig, T., Brugman, M.H., Bartels, S., Heinz, N., Kustikova, O.S., Modlich, U., Li, Z., Galla, M., Schiedlmeier, B., Schambach, A., and Baum, C. (2011) Polyclonal fluctuation of lentiviral vector-transduced and expanded murine hematopoietic stem cells, *Blood*, **117**, 3053–3064, doi: 10.1182/blood-2010-08-303222.
16. Lu, R., Neff, N.F., Quake, S.R., and Weissman, I.L. (2011) Tracking single hematopoietic stem cells *in vivo* using high-throughput sequencing in conjunction with viral genetic barcoding, *Nat. Biotechnol.*, **29**, 928–933, doi: 10.1038/nbt.1977.
17. Verovskaya, E., Broekhuis, M.J.C., Zwart, E., Ritsema, M., van Os, R., de Haan, G., and Bystrykh, L.V. (2013) Heterogeneity of young and aged murine hematopoietic stem cells revealed by quantitative clonal analysis using cellular barcoding, *Blood*, **122**, 523–532, doi: 10.1182/blood-2013-01-481135.
18. Biasco, L., Pellin, D., Scala, S., Dionisio, F., Basso-Ricci, L., Leonardelli, L., Scaramuzza, S., Baricordi, C., Ferrua, F., Cicalese, M.P., Giannelli, S., Neduva, V., Dow, D.J., Schmidt, M., Von Kalle, C., Roncarolo, M.G., Ciceri, F., Vicard, P., Wit, E., Di Serio, C., Naldini, L., and Aiuti, A. (2016) *In vivo* tracking of human hematopoiesis reveals patterns of clonal dynamics during early and steady-state reconstitution phases, *Cell Stem Cell*, **19**, 107–119, doi: 10.1016/j.stem.2016.04.016.
19. Nguyen, L.V., Pellacani, D., Lefort, S., Kannan, N., Osako, T., Makarem, M., Cox, C.L., Kennedy, W., Beer, P., Carles, A., Moksa, M., Bilenky, M., Balani, S., Babovic, S., Sun, I., Rosin, M., Aparicio, S., Hirst, M., and Eaves, C.J. (2015) Barcoding reveals complex clonal dynamics of *de novo* transformed human mammary cells, *Nature*, **528**, 267–271, doi: 10.1038/nature15742.
20. Cornils, K., Thielecke, L., Huser, S., Forgber, M., Thomaschewski, M., Kleist, N., Hussein, K., Riecken, K., Volz, T., Gerdes, S., Glauche, I., Dahl, A., Dandri, M., Roeder, I., and Fehse, B. (2014) Multiplexing clonality: combining RGB marking and genetic barcoding, *Nucleic Acids Res.*, **42**, e56, doi: 10.1093/nar/gku081.
21. Weber, K., Bartsch, U., Stocking, C., and Fehse, B. (2008) A multicolor panel of novel lentiviral «gene ontology» (LeGO) vectors for functional gene analysis, *Mol. Ther.*, **16**, 698–706, doi: 10.1038/mt.2008.6.
22. Aranyossy, T., Thielecke, L., Glauche, I., Fehse, B., and Cornils, K. (2017) Genetic barcodes facilitate competitive clonal analyses *in vivo*, *Hum. Gene Ther.*, **28**, 926–937, doi: 10.1089/hum.2017.124.
23. Sambrook, J., Fritsch, E., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor, Cold Spring Harbor Laboratory Press, N.Y.
24. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J., Smith, J.A., and Struhl, K. (1991) *Current protocols in molecular biology*, New York, John Wiley & Sons, Inc.
25. Kuzmina, L.A., Petinati, N.A., Parovichnikova, E.N., Lubimova, L.S., Gribova, E.O., Gaponova, T.V., Shipounova, I.N., Zhironkina, O.A., Bigildeev, A.E., Svinareva, D.A., Drize, N.J., and Savchenko, V.G. (2012) Multipotent mesenchymal stromal cells for the prophylaxis of acute graft-versus-host disease—a phase II study, *Stem Cells Int.*, **2012**, 1–8, doi: 10.1155/2012/968213.
26. Todoerti, K., Lisignoli, G., Storti, P., Agnelli, L., Novara, F., Manferdini, C., Codeluppi, K., Colla, S., Crugnola, M., Abeltino, M., Bolzoni, M., Sgobba, V., Facchini, A., Lambertenghi-Deliliers, G., Zuffardi, O., Rizzoli, V., Neri, A., and Giuliani, N. (2010) Distinct transcriptional profiles characterize bone microenvironment mesenchymal cells rather than osteoblasts in relationship with multiple myeloma bone disease, *Exp. Hematol.*, **38**, 141–153, doi: 10.1016/j.exphem.2009.11.009.
27. Gasparian, M.E., Elistratov, P.A., Drize, N.I., Nifontova, I.N., Dolgikh, D.A., and Kirpichnikov, M.P. (2009) Overexpression in *Escherichia coli* and purification of human fibroblast growth factor (FGF-2), *Biochemistry (Moscow)*, **74**, 221–225, doi: 10.1134/S000629790902014X.
28. Bailey, N.T.J. (1995) *Statistical methods in biology*, 3rd edition, Cambridge, Cambridge University Press.
29. Жиронкина О.А., Шипунова И.Н., Бигильдеев А.Е., Сац Н.В., Петинати Н.А., Дризе Н.И. (2011) Пролиферативный потенциал мультипотентных мезенхимальных стромальных клеток, выделенных из костного мозга человека, *Клеточные технол. биол. мед.*, 230–235.

**CLONAL COMPOSITION OF HUMAN MULTIPOTENT
MESENCHYMAL STROMAL CELLS: APPLICATION
OF GENETIC BARCODES FOR RESEARCH**

**A. E. Bigildeev^{1*}, A. M. Pilunov², N. V. Sats¹, V. L. Surin¹,
I. N. Shipounova¹, N. A. Petinati¹, M. D. Logacheva³, A. V. Fedotova³,
A. S. Kasyanov⁴, A. S. Artyukhov⁵, E. B. Dashinimaev⁶, and N. J. Drize¹**

¹ *National Research Center for Hematology, Ministry of Health
of the Russian Federation, 125167 Moscow, Russia;
E-mail: bigildeev.ae@gmail.com*

² *Lomonosov Moscow State University, 119991 Moscow, Russia*

³ *Belozersky Institute of Physico-Chemical Biology, Lomonosov
Moscow State University, 119992 Moscow, Russia*

⁴ *Vavilov Institute of General Genetics, Russian Academy
of Sciences, 119991 Moscow, Russia*

⁵ *Pirogov Russian National Research Medical University (RNRMU),
Ministry of Health of the Russian Federation, 117997 Moscow, Russia*

⁶ *Koltzov Institute of Developmental Biology, Russian
Academy of Sciences, 119334 Moscow, Russia*

Received September 12, 2018

Revised November 27, 2018

Accepted November 27, 2018

Clonal composition of human multipotent mesenchymal stromal cells (MMSCs), marked with lentiviral vectors carrying genetic barcodes, was studied. MMSCs were infected with a cloned library of self-inactivating lentiviral vectors carrying 667 non-repeating barcodes. At each passage of the culture, 120 cells were cloned one cell per well in 96-well plates. The effectiveness of cloning and labeling of clonogenic cells was determined. From the one cell-derived colonies DNA was extracted, and barcodes were identified by Sanger sequencing. DNA was also extracted from the total MMSC population at each passage, and the diversity and representation of barcodes was determined by the Illumina-based deep sequencing. It was shown that the proportion of MMSCs marked with lentiviral vectors is stable in passages. Given high multiplicity of infection, the labeling procedure can decrease MMSCs proliferative potential. Barcode identification in individual cell clones confirmed the polyclonal character of the MMSC population. Clonal composition of MMSCs changes with the passages significantly due to depletion of the proliferative potential of most cells. Large clones were found at the first passage of MMSCs, and then many small clones with a limited proliferative potential were detected in the population. The results of deep sequencing confirmed the changes in MMSCs clonal composition. The polyclonal population of MMSC contains only a small number of cells with a high proliferative potential, among which there may be stem cells. MMSCs with high proliferative potential are more often detected in the earliest passages. In this regard, the MMSCs of early passages are recommended for use in regenerative medicine based on cell proliferation.

Keywords: multipotent mesenchymal stromal cells, genetic marking, barcode, lentiviral vector