

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
НАУЧНОЕ УЧРЕЖДЕНИЕ
ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ЦЕНТР
ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ
СИБИРСКОГО ОТДЕЛЕНИЯ РАН

На правах рукописи

ЮНУСОВА АНАСТАСИЯ МАРАТОВНА

**Наследование способности фибробластов к
репрограммированию в индуцированные
плюрипотентные стволовые клетки**

03.03.04 – клеточная биология, цитология, гистология

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель:
кандидат биологических наук,
Баттулин Нариман Рашитович

Новосибирск 2017

Оглавление

Список использованных сокращений.....	6
Введение.....	8
Актуальность.....	8
Цели и задачи исследования.....	9
Научная новизна работы.....	10
Теоретическая и практическая значимость исследования.....	10
Положение, выносимое на защиту.....	10
Вклад автора.....	10
Апробация работы.....	11
Структура и объем работы.....	11
Глава I. Обзор литературы.....	12
I.1 Плюрипотентные стволовые клетки.....	12
I.1.1 Репрограммирование соматических клеток к плюрипотентному состоянию с помощью транскрипционных факторов.....	13
I.1.2 Гетерогенность исходной популяции как причина дифференциальной способности клеток к репрограммированию.....	14
I.1.2.1 Снижение гетерогенности репрограммируемых клеток: Полицистронные кассеты.....	15
I.1.2.1 Снижение гетерогенности репрограммируемых клеток: Вторичные репрограммируемые системы.....	16
I.1.3 Модель seesaw: репрограммирующие факторы как регуляторы дифференцировки.....	18
I.2 Основные этапы репрограммирования соматических клеток в ИПСК.....	20
I.2.1 Ранняя стадия репрограммирования.....	21
I.2.1.1 Увеличение скорости пролиферации клеток и подавление программы апоптоза.....	22
I.2.1.2 Мезенхимально-эпителиальный переход.....	24
I.2.2 Промежуточная стадия репрограммирования.....	26

I.2.3	Финальная стадия репрограммирования/Фаза стабилизации	28
I.3	«Project Grandiose». Альтернативные пути репрограммирования	30
I.4	Модели репрограммирования	43
I.4.1	Элитарные модели	43
I.4.1.1	Индукцированная элитарная модель	43
I.4.1.2	Элитарная модель	44
I.4.2	Стохастическая модель	45
I.4.3	Детерминистическая модель	46
I.4.4	Синтез стохастической и детерминистической модели	48
Глава II	Материалы и методы	50
II.1	Материалы	50
II.1.1	Клеточные линии	50
II.1.2	Вычислительные ресурсы	50
II.2	Методы	50
II.2.1	Молекулярно-биологические методы	50
II.2.1.1	Электрофорез в агарозном геле	50
II.2.1.2	Выделение ДНК из агарозного геля	50
II.2.1.3	Осаждение ДНК этанолом	50
II.2.1.4	Амплификация ДНК при помощи полимеразной цепной реакции (ПЦР)	51
II.2.1.5	Определение первичной структуры ДНК	51
II.2.1.6	Очистка ДНК с использованием Sephadex G50	51
II.2.1.7	Выделение плазмидной ДНК	51
II.2.1.8	Выделение геномной ДНК	52
II.2.1.9	Гидролиз ДНК эндонуклеазами рестрикции	52
II.2.1.10	Лигазная реакция	52
II.2.1.11	Трансформация электрокомпетентных клеток E.coli	52
II.2.2	Лентивирусные векторы	53
II.2.3	Получение библиотеки баркодированных лентивирусных векторов	53

II.2.3 Амплификация и секвенирование ДНК-баркодов репрограммированных клеток.....	54
II.3 Компьютерный анализ	56
II.3.1 Биоинформатическая обработка результатов секвенирования	56
II.3.2 Компьютерное моделирование и статистический анализ	56
II.4 Методы работы с клеточными культурами	58
II.4.1 Культуральные среды.....	58
II.4.2 Получение культуры первичных эмбриональных фибробластов.....	58
II.4.3 Культивирование фибробластов	59
II.4.4 Размораживание клеток.....	59
II.4.5 Замораживание клеток	59
II.4.6 Получение вирусных частиц.....	59
II.4.7 Сортировка Thy1.2-позитивных клеток с помощью проточной цитофлуорометрии	60
II.4.8 Репрограммирование эмбриональных фибробластов мыши к плюрипотентному состоянию	60
II.4.9 Определение времени удвоения клеточной популяции и количества потомков трансдуцированных клеток	61
II.4.10 Получение контрольных клеток.....	61
Глава III. Результаты	62
III.1 Схема эксперимента	62
III.2 Пилотный эксперимент по репрограммированию баркодированных фибробластов мыши к плюрипотентному состоянию	64
III.3 Параметры, определяющие эффективность баркодирования клеток.....	67
III.3.1 Разнообразие библиотеки баркодированных вирусов	67
III.3.2 Количество интеграций баркодов на одну клетку.....	68
III.3.3 Количество потомков исходных клеток	70
III. 4 Репрограммирование баркодированных фибробластов мыши к плюрипотентному состоянию: Эксперименты 1 и 2 (Несортированные ЭФМ).....	71

III.5 Секвенирование контрольных клеток с известными ДНК-баркодами	72
III.6 Анализ потери материала в ходе получения библиотеки (Секвенирование 3x5000 баркодированных клеток)	74
III.7 Оценка размера библиотеки баркодированных лентивирусных векторов	75
III.8 Описание компьютерной модели, симулирующей экспериментальные стадии	76
III.9 Компьютерное моделирование Экспериментов 1 и 2	78
III.10 Репрограммирование баркодированных фибробластов мыши к плюрипотентному состоянию: Эксперимент 3 (ЭФМ, сортированные по Thy-1) и Контрольный эксперимент.....	79
III.11 Анализ вклада быстроделющихся клеток в общее число двойных пересечений.....	82
III.12 Компьютерное моделирование двойных, тройных и четверных пересечений баркодов при дифференциальной способности клеток к репрограммированию	82
Глава IV. Обсуждение.....	84
IV.1 Оценка разнообразия библиотеки баркодированных лентивирусов.....	85
IV.2 Оценка наследуемости способности к репрограммированию через моделирование экспериментальных данных.....	86
Выводы	90
Благодарности	91
Список литературы.....	92
Приложения	111

Список использованных сокращений

АТФ – аденозинтрифосфат

ДНК – дезоксирибонуклеиновая кислота

ИПСК - индуцированные плюрипотентные стволовые клетки

МСК - мезенхимальные стволовые клетки

МЭП – мезенхимально-эпителиальный переход

п.н. – пар нуклеотидов

ПЦР – полимеразная цепная реакция

РНК - рибонуклеиновая кислота

ТФ – транскрипционный фактор

ЭДТА – этилендиаминтетраацетат

ЭМП – эпителио-мезенхимальный переход

ЭСК - эмбриональные стволовые клетки

ЭФМ – эмбриональные фибробласты мыши

ВМР – костный морфогенетический белок

DMEM (Dullbecco's Modified Eagle Medium) – среда Игла в модификации Дульбекко

DMR (differentially methylated regions)- дифференциально-метилованные регионы

Dox – доксициклин

FACS (fluorescence-activated cell sorting) – проточная цитофлуориметрия

FBS (fetal bovine serum) – эмбриональная бычья сыворотка

GFP (green fluorescent protein) – зеленый флуоресцентный белок

H3K27 – лизин в двадцать седьмом положении гистона H3

H3K4 – лизин в четвертом положении гистона H3

H3K36 – лизин в тридцать шестом положении гистона H3

H3K79 – лизин в семьдесят девятом положении гистона H3

IRES (internal ribosome entry site) - участок внутренней посадки рибосомы

KSR (Knockout serum replacement) - синтетический заменитель сыворотки

LeGO (Lentiviral Gene Ontology) – лентивирусный вектор третьего поколения

LIF (Leukemia inhibitory factor) - интерлейкин из семейства IL-6

MAPK (mitogen-activated protein kinases) – каскад митоген-активируемой протенкиназы

miRNA – малые некодирующие РНК

MOI (Multiplicity of infection) – среднее количество провирусных интеграций на геном клетки

Muse-клетки (Multilineage-differentiating stress enduring) – мультипотентные стрессоустойчивые клетки

NuRD (nucleosome remodeling and deacetylation) - хроматин-ремоделирующий комплекс

OG2 (Oct4-Gfp) – трансгенная линия мышей, клетки которых несут ген Gfp под контролем промотора Oct4.

OSKM (Oct4, Sox2, Klf4, c-Myc) – репрограммирующие факторы Яманаки

PBS (phosphate-buffered saline) – фосфатно-солевой буфер

PRC (Polycomb Repressive Complex) – хроматин-ремоделирующий комплекс (ингибиторный)

Tgf β – трансформирующий фактор роста β

Thy1 (Thy-1 cell surface antigen) – поверхностный антиген, маркер зрелых фибробластов

TRE (tetracycline response element) – последовательность, с которой связывается комплекс трансаиватор-доксидиклин

Введение

Актуальность

Репрограммирование генома открывает широкие возможности перед исследователями, позволяя изучать такие фундаментальные проблемы биологии, как молекулярные основы плюрипотентности, регуляцию дифференциальной активности генов, а также процессы, происходящие в раннем эмбриогенезе.

Принципиальная возможность репрограммирования генома соматической клетки была показана в опытах по клонированию сначала на амфибиях, а затем и на млекопитающих. Оказалось, что ядро соматической клетки взрослого организма, будучи пересаженным в энуклеированный ооцит, обеспечивает полное развитие организма [Gurdon, Uehlinger, 1966; Wilmut et al., 2002]. Позже появился альтернативный экспериментальный подход репрограммирования генома, основанный на слиянии эмбриональных стволовых клеток с соматическими клетками [Matveeva et al., 1998; Tada et al., 2001]. Такие клетки получили название гибридных. Для них характерно сохранение плюрипотентных свойств стволовых клеток на достаточно высоком уровне.

В 2006 году биология стволовых клеток получила мощный импульс к развитию, после того как было открыто репрограммирование с помощью транскрипционных факторов [Takahashi, Yamanaka, 2006]. Этот способ репрограммирования генома дифференцированных клеток представляет особый интерес, поскольку модуляция набора транскрипционных факторов и условий культивирования клеток, позволяют упростить исследования механизмов, лежащих в основе индукции и поддержании плюрипотентности [Buganim, Faddah, Jaenisch, 2013].

Однако эффективность получения индуцированных плюрипотентных клеток крайне низка — только малая часть клеток исходной популяции способна образовать плюрипотентного потомка. Так, например, в классической системе Яманаки (при использовании факторов Oct4, Sox2, Klf4 и c-Мус) репрограммируется не более 1% клеток. Анализ изменений в эпигенетическом профиле и профиле экспрессии генов на протяжении всего процесса репрограммирования показал, что большая часть клеток способна вступить в фазу инициации, то есть начать процесс репрограммирования. Однако только редкие клетки способны запустить экспрессию ключевых генов, необходимых для поддержания плюрипотентности. До сих пор неизвестно, что отличает эту малую часть клеточной популяции от остальных клеток.

Для объяснения механизмов репрограммирования были предложены две альтернативные модели: элитарная и стохастическая. Элитарная модель предполагает, что репрограммированию подвергается только особая субпопуляция клеток, которая может представлять собой недифференцированные тканевые стволовые клетки. Согласно же стохастической модели, способность к репрограммированию – случайное явление, независимо возникающее в каждой клетке после деления. [Yamanaka, 2009]. На данный момент существуют работы, подтверждающие как первую, так и вторую модель, поэтому невозможно однозначно ответить на вопрос, по какому механизму происходит репрограммирование клеток к плюрипотентному состоянию. Обладают ли клетки еще до индукции плюрипотентности некой предрасположенностью к репрограммированию? И может ли эта предрасположенность наследоваться дочерними клетками в ходе делений?

Для того, чтобы исследовать этот вопрос на уровне отдельных клеток, мы решили использовать метод баркодирования. Баркодирование клеток, один из методов маркирования, представляет собой внесение в геном клетки уникальной метки, ДНК-баркода, представляющего собой последовательность случайных нуклеотидов. ДНК-баркоды интегрируются в геном клетки с помощью лентивирусов, и в виде провирусов наследуются в ходе делений [Nguyen et al., 2014; Cheung et al., 2013]. Таким образом, потомки одной клетки несут один и тот же ДНК-баркод и могут быть легко определены с помощью секвенирования. Данный метод позволяет проследить за судьбой потомков тысяч клеток в течение процесса репрограммирования.

Цели и задачи исследования

Целью данной работы является изучение закономерностей распределения способности к репрограммированию в плюрипотентные стволовые клетки в популяции эмбриональных фибробластов мыши.

Для выполнения данной цели были поставлены следующие задачи:

1. Создать и охарактеризовать библиотеку ДНК-баркодированных лентивирусных векторов, подходящую для маркирования большой популяции клеток.
2. Провести эксперименты по репрограммированию баркодированных эмбриональных фибробластов мыши в плюрипотентные стволовые клетки.
3. Разработать компьютерную модель, позволяющую на основе частоты синхронного репрограммирования родственных клеток определить закономерности наследования способности к репрограммированию.

Научная новизна работы

Для изучения распределения способности к репрограммированию в популяции клеток был использован метод ДНК-баркодирования. В отличие от других методов исследования судьбы отдельных клеток, описываемый метод позволяет маркировать и отслеживать тысячи клеток в ходе эксперимента. Важно отметить, что в данной работе впервые с помощью метода ДНК-баркодирования была маркирована большая клеточная популяция (до 200 000 клеток). Кроме того, была разработана компьютерная модель, позволяющая на основе экспериментальных данных и с учетом параметров баркодирования, определить уровень наследования способности к репрограммированию.

Теоретическая и практическая значимость исследования

С теоретической точки зрения, изучение закономерностей распределения способности к репрограммированию способствует расширению наших знаний в области клеточной биологии и регуляции клеточной судьбы. Выявление характеристик успешно репрограммирующихся клеток позволит сделать поиск факторов, увеличивающих эффективность репрограммирования, более осмысленным и продуктивным. Плюрипотентные стволовые клетки являются уникальным объектом для исследования фундаментальных проблем биологии, связанных с основами плюрипотентности и регуляцией дифференциальной активности генов в процессе раннего развития млекопитающих. Заложенный в плюрипотентных стволовых клетках потенциал к дифференцировке позволяет использовать данные клетки в тканезаместительной терапии и скрининге лекарственных препаратов.

Положение, выносимое на защиту

На защиту выносится следующее положение:

В популяции эмбриональных фибробластов мыши присутствуют клетки с имманентной предрасположенностью к репрограммированию, передающуюся дочерним клеткам в ходе делений.

Вклад автора

Автором самостоятельно получены все основные результаты: создана и охарактеризована библиотека ДНК-баркодированных лентивирусных векторов, проведены эксперименты по репрограммированию маркированных ДНК-баркодами эмбриональных фибробластов мыши, проведен биоинформатический анализ данных высокопроизводительного секвенирования. Компьютерная модель, позволяющая на основе

экспериментальных параметров определять уровень наследуемости способности фибробластов к репрограммированию, была разработана совместно с В.С.Фишманом.

Апробация работы

Работа была доложена на следующих научных конференциях:

1. **Юнусова А.М.**, Фишман В.С., Баттулин Н.Р. Наследование способности фибробластов к репрограммированию в индуцированные плюрипотентные стволовые клетки //Материалы 51-й международной научной студенческой конференции "Студент и научно-технический прогресс" 12-18 апреля 2013 г. Биология, с. 175.
2. **Юнусова А.М.**, Фишман В.С., Баттулин Н.Р. Создание системы ДНК-баркодированных вирусов для маркирования клеток //Материалы VI Международной школы молодых ученых по молекулярной генетике на тему «ГЕНОМИКА И СИСТЕМНАЯ БИОЛОГИЯ» 16-21 ноября 2014г. Звенигород, с 51.
3. **A.Yunusova**, V.Fishman, N.Battulin, O.Serov. Susceptibility of embryonic fibroblasts to reprogramming inherited through cell division. Abstracts for Cell Technologies at The Edge: Research & Practice (CTERP) 6-8 April 2016 St. Petersburg, p 53.
4. **A.Yunusova**, V.Fishman, N.Battulin. Symmetric reprogramming fate in somatic cell lineages. Abstract Book for EMBO Workshop «Nuclear function and cell fate choice». 18-22 September 2016 Kyllini, Greece, p 116.
5. **A.Yunusova**, V.Fishman, N.Battulin. Reprogramming potential is an inheritable trait. Cell Symposia: 10 years of iPSCs. 25-27 September 2016, Berkley, USA, p 5.

По материалам работы опубликованы следующие статьи:

1. **Юнусова А.М.**, Баттулин Н.Р. Методы маркирования клеток для изучения судьбы клеточных поколений. //Вавиловский журнал генетики и селекции. 2016; 20(6) с 909-917. doi: 10.18699/VJ16.211.
2. **A.Yunusova**, V.Fishman, G.Vasiliev, N.Battulin. Deterministic versus stochastic model of reprogramming: new evidence from cellular barcoding technique. //Open Biology – 2017 (Apr); 7(4): 160311. doi:10.1098/rsob.160311.

Структура и объем работы

Диссертация состоит из оглавления, списка сокращений, введения, обзора литературы, описания используемых материалов и методов, результатов, обсуждения, выводов и списка литературы. Работа изложена на 118 страницах, содержит 17 рисунков, 3 таблицы и 7 приложений.

Глава I. Обзор литературы

I.1 Плюрипотентные стволовые клетки

Плюрипотентные стволовые клетки являются уникальным объектом для исследования фундаментальных проблем биологии, связанных с основами плюрипотентности и регуляцией дифференциальной активности генов в процессе раннего развития млекопитающих. Кроме того, заложенный в плюрипотентных стволовых клетках потенциал к дифференцировке позволяет использовать данные клетки в тканезаместительной терапии и скрининге лекарственных препаратов. Уникальность плюрипотентных стволовых клеток заключается в неограниченной возможности самовоспроизведения в культуре, а также способности дифференцироваться в производные всех трех зародышевых листков.

Впервые плюрипотентные стволовые клетки были получены из клеток внутренней клеточной массы бластоцист мыши в 1981 [Evans, Kaufman, 1981; Martin, 1981]. Полученные клетки были названы эмбриональными стволовыми клетками (ЭСК) и демонстрировали способность формировать все ткани и органы взрослого организма, то есть обладали плюрипотентностью. Только в 1998 году исследователям удалось получить ЭСК человека из 5-6 дневных бластоцист [Thomson, 1998].

Принципиальная возможность репрограммирования генома соматической клетки была доказана в многочисленных экспериментах по клонированию [Wilmut et al., 2002] и формированию гибридов между соматическими и эмбриональными стволовыми клетками [Matveeva et al., 1998; Tada et al., 2001]. Благодаря этим экспериментам, стало очевидно, что неоплодотворенные яйцеклетки и ЭСК содержат определенные факторы, играющие важную роль в индукции плюрипотентности в дифференцированных клетках. Для того чтобы подобрать набор факторов, необходимых и достаточных для репрограммирования соматических клеток, японскими исследователями Яманака и Такахаши были протестированы сочетания 24 транскрипционных факторов, вовлеченных в самоподдержание плюрипотентных клеток. Было установлено, что для индукции плюрипотентного состояния в мышинных эмбриональных фибробластах необходим набор всего из четырех факторов – *Oct4*, *Sox2*, *Klf4* и *c-Мyc* («коктейль Яманаки»). Полученные клетки были названы индуцированными плюрипотентными стволовыми клетками (ИПСК). ИПСК были получены из соматических клеток различных видов животных, включая человека [Park et al., 2008; Takahashi et al., 2007; Yu et al., 2007], крысу [Li et al., 2009b], макаку-резус [Liu et al., 2008], свиней [Esteban et al., 2009]. Это демонстрирует, что

фундаментальные свойства генных сетей, лежащих в основе индукции и поддержании плюрипотентности, остаются консервативными в процессе эволюции. Аналогичным образом были получены ИПСК из соматических клеток различного происхождения, таких как кератиноциты [Aasen et al., 2008], нейральные клетки [Eminli et al., 2008; Kim et al., 2008], клетки желудка и печени [Aoi et al., 2008], меланоциты [Utikal et al., 2009], панкреатические β -клетки [Stadtfield, Brennand, Hochedlinger, 2008] и терминально дифференцированные лимфоциты [Eminli et al., 2009; Hanna et al., 2008].

I.1.1 Репрограммирование соматических клеток к плюрипотентному состоянию с помощью транскрипционных факторов

В настоящее время разработано большое количество методов доставки репрограммирующих факторов в соматические клетки. В целом, эти методы могут быть разделены на две группы. К первой группе относятся интеграционные методы, где доставка факторов осуществляется с помощью ленти- и ретровирусных конструкций, встраивающихся в геном клеток [Aasen et al., 2008; Dimos et al., 2008; Park et al., 2008]. Именно с помощью ретровирусов были получены первые ИПСК (К. Takahashi & Yamanaka, 2006). Вторая группа включает методы, в ходе которых геном клеток либо не модифицируется вовсе, либо модифицируется минимально. Это репрограммирование клеток с помощью плазмидных, эписомных и аденовирусных векторов, которые не интегрируются в геном [Yu et al., 2009; Zhou, Freed, 2009]. А также введение в клетки рекомбинантных белков либо синтетических модифицированных РНК репрограммирующих факторов [Kim et al., 2009; Warren et al., 2010]. К этой же группе относятся методы, предусматривающие удаление генетических конструкций из генома с помощью рекомбиназы либо транспозазы [Kaji et al., 2009; Woltjen et al., 2009]. Методы разные, однако объединяет их то, что эффективность репрограммирования во всех случаях очень низкая. Как правило, значение лежит в диапазоне от 0.0001 до 1% [Fusaki et al., 2009; Naase et al., 2009; Warren et al., 2010]. Использование химических веществ и сверхэкспрессия дополнительных факторов позволяет лишь незначительно увеличить эффективность репрограммирования [Esteban et al., 2010; Feng et al., 2009; Huangfu et al., 2008; Ichida et al., 2009; Mali et al., 2010; Zhao et al., 2008]. В связи с этим большой интерес представляет вопрос, почему только редкие клетки из исходной популяции преуспевают в обретении плюрипотентности. Как происходит и чем обусловлено распределение способности к репрограммированию в популяции соматических клеток до сих пор неизвестно.

В последующих главах будут кратко представлены возможные причины низкой эффективности репрограммирования клеток к плюрипотентному состоянию с помощью транскрипционных факторов. Будут описаны изменения на транскрипционном, эпигенетическом и морфологическом уровнях, происходящие в клетках на разных этапах репрограммирования и обсуждены события, препятствующие репрограммированию. Обзор литературы будет завершён описанием всех предложенных на сегодняшний день моделей репрограммирования, позволяющих продемонстрировать как происходит распределение способности к репрограммированию в клеточной популяции.

I.1.2 Гетерогенность исходной популяции как причина дифференциальной способности клеток к репрограммированию

При трансдукции репрограммирующими факторами клетки демонстрируют целый спектр совершенно различных вариантов поведения. Это и остановка клеточного цикла, сопровождающаяся запуском апоптоза в ответ на пролиферативный стресс, и неполное репрограммирование в силу либо недостаточного подавления экспрессии тканеспецифичных генов, либо недостаточной активации генов, поддерживающих плюрипотентность [Mikkelsen et al., 2008]. Только малая часть клеток экспериментальной популяции преодолевает все препятствия, а образованные ими колонии плюрипотентных клеток демонстрируют все свойства эмбриональных стволовых клеток. В чём же причина такого разного ответа клеток на репрограммирующие факторы? Разумным выглядит предположение, что эта причина кроется в гетерогенности, присущей исходной клеточной популяции. Наиболее часто используемые в экспериментах по репрограммированию клетки, это первичная популяция эмбриональных фибробластов мыши. Эмбриональные фибробласты мыши представляют собой совокупность нескольких субпопуляций, значительно различающихся между собой и по молекулярным характеристикам, и по морфологическим [Singhal et al., 2016]. Кроме того, эти субпопуляции имеют различную предрасположенность к иммортализации и различный ответ на стресс, вызванный экспрессией экзогенных репрограммирующих факторов, что несомненно сказывается на выживании клеток в ходе процесса репрограммирования [Hanna, Saha, Jaenisch, 2010].

Лентивирусные конструкции, зачастую используемые в качестве доставщиков репрограммирующих факторов, встраиваются в геном случайным образом, преимущественно в интроны [Ustek et al., 2012]. Их транскрипционная активность, как известно, зависит от хромосомного контекста сайта интеграции: близости гетерохроматинизированных районов и различного рода регуляторных элементов. После трансдукции четырьмя вирусами с репрограммирующими факторами каждая клетка

характеризуется уникальным сайтом интеграции каждого провируса. А, следовательно, и уровень экспрессии репрограммирующих факторов значительно различается между клетками. Часть клеток при этом будет иметь большую вероятность репрограммирования, просто вследствие того, что имеет «благоприятные» сайты интеграции факторов и оптимальную стехиометрию репрограммирующих факторов.

I.1.2.1 Снижение гетерогенности репрограммируемых клеток: Полицистронные кассеты

Один из возможных вариантов снижения гетерогенности экспериментальной популяции клеток заключается в использовании полицистронных векторов, кодирующих гены Oct4, Sox2, Klf4 и c-Мyc в едином транскрипте, либо на основе лентивирусов, либо на основе транспозонов [Carey et al., 2011; Sommer et al., 2009; Woltjen et al., 2009]. Таким образом исследователям удастся добиться более равномерной представленности репрограммирующих факторов в клетках исходной популяции. Причем для получения колоний плюрипотентных клеток оказывается достаточно всего лишь по одной копии каждого из факторов Яманаки. Интересно, что в некоторых колониях ИПСК, полученных в пионерной работе Такахаши и Яманаки, количество копий репрограммирующих факторов превышало 15 [Takahashi, Yamanaka, 2006; Wernig et al., 2007]. Несомненным достоинством полицистронных кассет с репрограммирующими факторами является их большая эффективность по сравнению с отдельными векторами.

Интересно, что порядок расположения генов в полицистронной кассете влияет на уровень экспрессии белков, а, следовательно, и на эффективность репрограммирования. Рассмотрим этот случай на примере полицистронных кассет на основе транспозонов. В разных научных работах были использованы следующие конструкции, содержащие все 4 репрограммирующих фактора: MKOS [Kaji et al., 2009] [O'Malley et al., 2013], OSKM [Carey et al., 2010] [Yusa et al., 2009] [Kim et al., 2015], STEMCCA (OKSM) [Carey et al., 2011] и OKMS [Kim et al., 2015]. Порядок букв в аббревиатурах соответствует расположению генов репрограммирующих факторов в кассете. Причем в некоторых случаях кодирующие участки генов разделены 2A пептидами, в других – IRES последовательностями. Чантзура с коллегами сравнили эффективность репрограммирования эмбриональных фибробластов мыши при использовании указанных выше конструкций [Chantzoura et al., 2015]. При использовании кассеты OKMS наблюдалось наибольшее количество колоний с морфологией эмбриональных стволовых клеток. Однако подавляющее большинство из них не экспрессировали ген-маркер плюрипотентных клеток *Nanog*, что свидетельствует о неполном репрограммировании. Высокое содержание *Nanog*-негативных колоний было

обнаружено и при использовании STEMCCA (OKSM) кассеты. Тогда как при репрограммировании клеток с помощью MKOS или OSKM почти все колонии экспрессировали Nanog к 15 дню. Стоит отметить, однако, что общее количество колоний плюрипотентных клеток было значительно ниже чем в первых двух случаях.

В целом, все доступные на сегодняшний день полицистронные кассеты, кодирующие факторы Яманаки, могут быть разделены на две группы (по экспрессии гена *Klf4* [Kim et al., 2015]. Первая группа включает кассеты, несущие полноразмерный ген *Klf4L*. Ко второй группе относятся кассеты, кодирующие укороченную форму *Klf4S*, в которой отсутствуют первые 9 аминокислот. Экспериментально было показано, что с *Klf4S* продуцируется меньше белкового продукта. А кроме того, эффективность репрограммирования при использовании кассет второй группы значительно ниже. Важно отметить, что кассеты OKMS и STEMCCA(OKSM), использованные в описанной выше работе [Chantzoura et al., 2015] и показавшие наименьшую эффективность, принадлежат именно второй группе. При замене изоформы *Klf4S* на *Klf4L* в этих кассетах (OK+9MS and STEMCCA+9), эффективность репрограммирования повышается [Chantzoura et al., 2015; Kim et al., 2015].

Таким образом, уровень экспрессии и стехиометрия репрограммирующих факторов, оказывает значительное влияние на процесс репрограммирования. Для того чтобы индукция плюрипотентности прошла успешно, белковых продуктов генов Oct4 и Klf4 должно быть больше чем белковых продуктов генов Sox2 и C-Мус. В этом случае колонии ИПСК проявляют все свойства эмбриональных стволовых клеток, в обратном же – клетки репрограммируются не полностью и имеют aberrantное метилирование некоторых геномных регионов [Carey et al., 2010; Papapetrou et al., 2009].

I.1.2.1 Снижение гетерогенности репрограммируемых клеток: Вторичные репрограммируемые системы

Другой вариант снижения гетерогенности экспериментальной популяции клеток заключается в использовании, так называемых «вторичных» репрограммируемых систем. На первом этапе, из соматических клеток получают плюрипотентные клетки при помощи различных векторов, в том числе и полицистронных, описанных выше. Для возможности контроля экспрессии репрограммирующих факторов, промоторы снабжают операторной последовательностью, с которой связывается белок-трансактиватор rtTA при взаимодействии с доксициклином (как правило, используются клеточные линии, которые уже несут ген rtTA). Таким образом, экспрессия репрограммирующих факторов запускается при добавлении в культуральную среду доксициклина. Полученные линии ИПСК

используют для создания химерных, либо полностью трансгенных животных [Markoulaki et al., 2009; Smith et al., 2010; Wernig et al., 2008a; Woltjen et al., 2009]. И уже дифференцированные клетки этих животных репрограммируют, индуцировав экспрессию репрограммирующих факторов доксициклином. Второй раунд репрограммирования и дал название таким системам – «вторичные» репрограммируемые системы (secondary reprogramming system). Поскольку для клеток человека такой подход не применим, первичные колонии ИПСК человека дифференцируют *in vitro*, и уже эти производные репрограммируют вторично [Hockemeyer et al., 2008; Maherali et al., 2008].

Отличительная особенность таких «вторичных» систем в том, что все клетки являются потомками одной клетки, успешно репрограммировавшейся и образовавшей колонию ИПСК. А, следовательно, все они имеют одинаковые сайты интеграции репрограммирующих факторов. Логично предположить, что такие клетки должны репрограммироваться с гораздо большей эффективностью чем обычные фибробласты, трансдуцированные вирусами с репрограммирующими факторами. Теоретически, эффективность может достигать и 100%, так как все клетки являются клонами *той самой* клетки, которая имела «правильную» интеграцию репрограммирующих факторов, позволившую ей стать плюрипотентной. Вопреки ожиданиям, это оказалось не так. Только 4% «вторичных» фибробластов образовывали колонии плюрипотентного клеток и были способны длительно поддерживать состояние плюрипотентности в отсутствие экспрессии экзогенных репрограммирующих факторов. Стало очевидно, что даже генетически-идентичные клетки сильно различаются по способности к репрограммированию [Wernig et al., 2008b].

Есть и другой способ создания вторичных репрограммируемых систем, представленный сайт-направленным трансгенезом эмбриональных стволовых клеток. Индуцибельную полицистронную кассету с репрограммирующими факторами с помощью гомологичной рекомбинации встраивают в *Col1a1* локус [Carey et al., 2010; Stadtfeld et al., 2010]. Ген *Col1a1* кодирует коллаген первого типа и активно экспрессируется во многих типах клеток мыши, включая фибробласты, кератиноциты и нейральные предшественники. Таким образом достигается стабильная экспрессия трансгенной конструкции [Beard et al., 2006; Hochedlinger et al., 2005]. В качестве исходной линии ЭСК используют линию, которая уже несет ген трансактиватора (rtTA) в другом излюбленном генными инженерами локусе – *Rosa26*, обеспечивающем конститутивную экспрессию трансгена во всех тканях. Колонии ЭСК с правильной интеграцией используют для создания либо химерных, либо полностью трансгенных животных, клетки которых репрограммируют добавлением доксициклина. Эффективность репрограммирования таких вторичных систем значительно

превышает эффективность репрограммирования для систем со случайной интеграцией репрограммирующих факторов [Stadtfield et al., 2010], достигая 40% (в сравнении с 4% у [Wernig et al., 2008b]). Интересно также, что при интеграции репрограммирующих факторов и трансактиватора в один локус *Rosa26*, эффективность репрограммирования оказывается ниже, чем в *Col1a1* системах [Haenebalcke et al., 2013]. Причиной этого является более низкий уровень экспрессии репрограммирующих факторов в локусе *Rosa26*, что еще раз показывает важность «правильной» интеграции репрограммирующих факторов для индукции плюрипотентности.

При рассмотрении влияния уровня экспрессии и стехиометрии репрограммирующих факторов на эффективность репрограммирования представляется важным упомянуть о механизмах, лежащих в основе индукции плюрипотентности.

I.1.3 Модель seesaw: репрограммирующие факторы как регуляторы дифференцировки

Считается, что поддержание клетками плюрипотентного состояния достигается за счет совместного действия транскрипционных факторов – регуляторов плюрипотентности, которые препятствуют любой дифференцировке клеток [Jaenisch, Young, 2008; Silva, Smith, 2008; Young, 2011]. Хорошей иллюстрацией этому служит известное изображение эпигенетического ландшафта Конрадом Уоддингтоном. Здесь и дифференцированное состояние и плюрипотентность - относительно стабильные (устойчивые) и находятся на плато. Отсюда логично предположение, что дополнительная сверхэкспрессия в плюрипотентных же клетках факторов плюрипотентности, основная функция которых запрещать специализацию, не должна изменить судьбу клеток. Однако это оказалось не так. Например, повышенная экспрессия гена *Oct4*, как и гена *Nanog*, ключевых регуляторов плюрипотентности, приводит к дифференцировке ЭСК в клетки мезодермального ряда [Тео et al., 2011; Yu et al., 2011]. Гены *Sox2* и *Sip1* способствуют дифференцировке в нейроэктодерму, а гены *Esrrb*, *Sall4*, и *Tbx3* индуцируют развитие энтодермальных предшественников [Chng et al., 2010; Ivanova et al., 2006; Kopp et al., 2008; Lu, Yang, Jin, 2011; Zhang et al., 2006]. Это подтолкнуло некоторых исследователей к гипотезе о действии факторов плюрипотентности как классических индукторов дифференцировки. Причем работа этих факторов не кооперативная, а зачастую «противоборствующая»: многие из них запускают дифференцировку во взаимоисключающие типы клеток. Состояние плюрипотентности же поддерживается точным соотношением этих факторов, небольшой сдвиг которого приводит к дифференцировке в тот или иной тип специализированных клеток [Loh, Lim, 2011]. Этим может объясняться частая спонтанная дифференцировка ЭСК

при культивировании и необходимость поддерживать недифференцированное состояние клеток с помощью модуляции сигнальных путей химическими молекулами такими как LIF и CHIR 99021 + PD 0325901 (условия культивирования 2i).

Позднее эта гипотеза была расширена и на состояние индуцированной плюрипотентности [Shu et al., 2013]. Суть её в том, что факторы Яманаки при репрограммировании действуют как классические индукторы линейной дифференцировки. Экспериментальная проверка показала, что такие регуляторы развития мезодермальных линий как *Gata3*, *Gata6*, *Sox7* или *Pax1*, способны заменить *Oct4* в репрограммирующем коктейле. Достигается это, по-видимому, за счет подавления «эктодермальных» генов, экспрессия которых запускается другими репрограммирующими факторами: SOX2, KLF4 и C-MYC (SKM). В свою очередь, гены, связанные с развитием клеток эктодермального происхождения, такие как *Sox1*, *Sox3*, *Rcor2* и *Gmnn*, могут заменить *Sox2* в ходе репрограммирования. Аналогично предыдущему случаю, SOX2 и заменяющие его факторы подавляют экспрессию специфичных для мезодермы генов, которые активируются факторами OCT4, KLF4 и C-MYC (ОКМ). Совместная экспрессия *Gata6* и *Gmnn* позволяет репрограммировать фибробласты мыши только в присутствии экзогенных *Klf4* и *c-Myc* [Shu et al., 2013]. Позднее была показана возможность репрограммирования и фибробластов человека с помощью индукторов линейной дифференцировки: *Gata6* как замена *Oct4* и гены *Znf521*, *Otx2* и *Pax6* как замена *Sox2* [Montserrat et al., 2013]. Принципиально важно отметить, что все указанные гены (за исключением геминина *Gmnn*) не участвуют в поддержании плюрипотентности в ЭСК, но важны для линейной дифференцировки клеток.

На основе полученных данных, исследователи выдвинули новую балансовую модель индукции плюрипотентности, получившую название seesaw model (модель типа качели) [Shu, Deng, 2013]. Здесь плюрипотентное состояние находится в динамическом равновесии, которое устанавливается за счет взаимодействия между «конкурирующими» факторами линейной дифференцировки. Согласно данной модели, добиться установления плюрипотентности в дифференцированных клетках можно только при условии, что все действующие силы уравниваются друг друга. Несмотря на привлекательность такой модели, необходимо отметить то, что не все исследованные индукторы дифференцировки оказались способны заменить главные репрограммирующие факторы – Oct4 и Sox2. Другая сложность заключается в необходимости доказательства действия этих индукторов дифференцировки независимо от «коровой генной сети» плюрипотентности [Ben-David, Nissenbaum, Benvenisty, 2013]. Так, например, геминин экспрессируется в плюрипотентных клетках и через хроматин-ремоделирующий фактор Brg1 стимулирует экспрессию факторов Oct4, Sox2, и Nanog [Yang et al., 2011].

1.2 Основные этапы репрограммирования соматических клеток в ИПСК

Переход клетки из дифференцированного состояния в плюрипотентное требует масштабных эпигенетических преобразований, приводящих к изменению транскрипционного профиля клетки, изменению метаболизма, фенотипических и других характеристик. Первые исследования механизмов, лежащих в основе этих изменений, были проведены на гетерогенных клеточных популяциях в разные временные точки репрограммирования [Li et al., 2010; Maherali et al., 2008; Mikkelsen et al., 2008; Samavarchi-Tehrani et al., 2010; Stadtfeld et al., 2008]. Под гетерогенностью здесь подразумевается то, что анализируются все репрограммируемые клетки, независимо от их возможной судьбы: апоптоз, клеточное старение или неполное репрограммирование. Такие исследования позволяют определить глобальные процессы, запускаемые репрограммирующими факторами. Однако, точную последовательность событий, приводящих конкретную клетку к плюрипотентности, установить достаточно сложно. Это, в первую очередь, связано с низкой эффективностью репрограммирования. Есть несколько стратегий для решения этой задачи. Во-первых, прижизненное наблюдение за отдельными клетками с помощью микроскопии [Araki et al., 2010; Smith et al., 2010]. К сожалению, это позволяет детектировать изменения лишь на морфологическом уровне. Альтернативный способ – это обогащение популяции клетками, которые с большей вероятностью образуют плюрипотентного потомка [Hansson et al., 2012; O'Malley et al., 2013; Polo et al., 2012]. Зная маркеры ранней стадии репрограммирования, можно разделить клеточную популяцию на «предрасположенные» и «непредрасположенные» или рефрактерные к репрограммированию клетки. Сравнение событий, происходящих в этих группах клеток, позволяет выявить природу барьеров, стоящих на пути индукции плюрипотентности. И, наконец, наиболее чувствительный метод – это анализ эпигенетических модификаций и экспрессии генов в единичных клетках [Buganim et al., 2012]. Кроме того, использование вторичных репрограммируемых систем и клональный анализ позволяют заметно снизить гетерогенность клеток в исследуемых популяциях [Buganim et al., 2012; Golipour et al., 2012; Hanna et al., 2009].

Независимо от типа клеток и времени, необходимого для репрограммирования, последовательность событий, происходящих в ходе индукции плюрипотентности в общих чертах одинакова. Процесс репрограммирования принято делить на три стадии: раннюю (стадия инициации = *initiation*), промежуточную (стадия созревания = *maturation*) и позднюю (стадия стабилизации = *stabilization*) [Apostolou, Hochedlinger, 2013; Li et al., 2010; Samavarchi-Tehrani et al., 2010].

I.2.1 Ранняя стадия репрограммирования

Ранняя стадия репрограммирования является наиболее хорошо охарактеризованной. Возможно, это является отражением одного простого факта: в раннюю стадию вступает *большинство* исходных клеток, а, следовательно, фиксировать происходящие изменения значительно легче. Согласно микроскопическому анализу клетки уменьшаются в размере и приобретают округлую форму, их способность к миграции снижается. При этом скорость деления увеличивается, а между дочерними клетками образуются плотные адгезионные контакты [Smith et al., 2010], свойственные эпителиальным клеткам. Переход от мезенхимального типа клеток к эпителиальному получил название мезенхимально-эпителиального перехода [Li et al., 2010; Samavarchi-Tehrani et al., 2010].

Маркерная точка раннего этапа – потеря соматическими клетками поверхностного антигена THY-1 (Thymus cell antigen-1) [Brambrink et al., 2008; Polo et al., 2012; Stadtfeld et al., 2008]. Клетки, неспособные инактивировать Thy1 в течение первых трех дней репрограммирования, не образуют колоний плюрипотентных клеток, то есть рефрактерны к репрограммированию. Аналогичная ситуация наблюдается и для клеток, которые хоть и снижают экспрессию Thy1, но не активируют поверхностный антиген SSEA-1 (Stage-specific embryonic antigen-1) через шесть дней после начала репрограммирования.

Детальный анализ транскриптома и протеома двух клеточных субпопуляций (восприимчивой и рефрактерной) показал, что в ходе репрограммирования можно выделить две волны наибольших изменений в экспрессии генов [Hansson et al., 2012; Polo et al., 2012]. И первая волна приходится именно на раннюю стадию репрограммирования (первые три дня) [David, Polo, 2014]. В клетках увеличивается экспрессия генов, связанных с пролиферацией, регуляцией репликации ДНК, процессингом РНК, метаболизмом и организацией цитоскелета [Hansson et al., 2012; Li et al., 2010; Mikkelsen et al., 2008; Polo et al., 2012; Samavarchi-Tehrani et al., 2010]. В то время как тканеспецифичные гены дифференцированных клеток репрессируются. Интересно, что изменения в экспрессии генов, характерные для ранней фазы, наблюдаются в большинстве клеток исходной популяции, в том числе и в рефрактерных клетках. Однако в процессе репрограммирования количество генов, чья активность различается в восприимчивых и рефрактерных клетках, все более увеличивается, к концу процесса достигая отметки более чем 1500 генов. Этот список включает гены, кодирующие компоненты внеклеточного матрикса и комплексы связывания с ретиноевой кислотой, а также гены иммунного ответа (Mmp12, Rarres2, Fgf2, Fndc1, Il1f10 и Hsd11b1) [Polo et al., 2012].

Относительно анализа экспрессии генов в единичных клетках данные разнятся. В одной из работ исследователи провели микрочиповой анализ 26 генов и показали синхронное изменение их экспрессии в первые дни репрограммирования [Polo et al., 2012]. В другой работе при анализе 48 генов было обнаружено, что в первые шесть дней репрограммирования даже среди сестринских клеток наблюдается большая вариабельность [Buganim et al., 2012].

Вероятно, такие различия в экспрессии генов могут объясняться беспорядочным связыванием репрограммирующих факторов с сайтами-мишенями. Эта беспорядочность (promiscuous binding) обусловлена избытком экзогенных репрограммирующих факторов. Было показано, что при репрограммировании транскрипционные факторы OSKM связываются с промоторными регионами и энхансерными последовательностями гораздо большего числа генов чем в ЭСК (35% против 3%) [Soufi, Donahue, Zaret, 2012]. Большинство изменений в профиле экспрессии генов на ранней стадии связывают с транскрипционным фактором C-MYC. C-MYC связывается с промоторными районами открытого хроматина и способствует привлечению РНК-полимеразы и позитивного фактора элонгации P-TEF-b, действуя таким образом, как усилитель транскрипции [Nie et al., 2012; Rahl et al., 2010]. Кроме того, C-MYC облегчает связывание других репрограммирующих факторов с хроматином. В отличие от C-MYC, факторы OSK действуют как транскрипционные пионер-факторы, способные связываться с целевыми последовательностями в контексте закрытого хроматина [Soufi, Donahue, Zaret, 2012]. Стоит отметить, что среди генов, которые активируются OSKM факторами на ранней стадии репрограммирования не только гены, связанные с плюрипотентностью и мезенхимально-эпителиальным переходом, но также гены регуляторы апоптоза, что объясняет повышенный уровень гибели клеток.

Подводя итог, можно выделить следующие критические события, которые имеют место на ранней стадии репрограммирования: это ускорение клеточной пролиферации, избежание апоптоза и мезенхимально-эпителиальный переход. Более подробно речь о них пойдет ниже.

1.2.1.1 Увеличение скорости пролиферации клеток и подавление программы апоптоза

С помощью прижизненной микроскопической съемки на протяжении двух недель, исследователи из группы Александра Мейсснера провели ретроспективный анализ, проследив путь от колонии ИПСК до исходной клетки, из которой эта колония образовалась. Оказалось, что отличительной характеристикой успешно репрограммированных клеток является увеличение скорости пролиферации в первые 24

часа после начала экспрессии репрограммирующих факторов [Smith et al., 2010]. Анализ профиля экспрессии генов после индукции плюрипотентности показал, что в «восприимчивых» клетках в отличие от рефрактерных происходит увеличение экспрессии генов циклинов (*Ccnd1*, *Ccnd2*), циклин-зависимых киназ (*Cdk1*, *Cdk2*) и других генов, связанных с пролиферацией (*Plk1*, *Npat*, *Mki67*) [Hanna et al., 2009; Kim et al., 2008; Polo et al., 2012]. Это, в первую очередь, опосредовано действием репрограммирующих факторов. Например, С-МЫС активирует клеточную пролиферацию. Другой фактор, LIN 28, использующийся для репрограммирования клеток человека, ингибирует активность микро РНК let-7, которая, в свою очередь, репрессирует с-Мус (Viswanathan & Daley, 2010). О важности усиления пролиферации свидетельствует тот факт, что белки-ингибиторы циклин-зависимых киназ, таких как CIP1, INK4A и ARF, значительно снижают эффективность репрограммирования [Banito et al., 2009; Li et al., 2010; Utikal et al., 2009]. И напротив, коэкспрессия репрограммирующих факторов вместе с факторами, позитивно-регулирующими пролиферацию (REM2 или циклин D1), позволяет увеличить количество колоний ИПСК [Edel et al., 2010; Tanabe et al., 2013].

Подавляющее большинство клеток оказываются неспособными запустить благоприятствующие репрограммированию изменения в клеточном цикле. Такие клетки, сохраняя морфологию соответствующего дифференцированного типа, либо подвергаются апоптозу, либо клеточному старению и остановке клеточного цикла [Banito et al., 2009; Smith et al., 2010]. Более дифференцированные клетки и клетки на поздних пассажах репрограммируются с гораздо меньшей эффективностью [Eminli et al., 2009]. Однако, если дополнительно экспрессировать в них микро РНК из семейств miR-290 и miR-302, ингибирующие клеточное старение, эффективность репрограммирования можно увеличить [Judson et al., 2009; Wang et al., 2008]. Важно отметить, что экспрессия указанных микро РНК запускается с-Мус, показывая, тем самым, его положительное влияние на процесс репрограммирования достигается не только за счет увеличения скорости пролиферации, но и за счет подавления клеточного старения [Judson et al., 2009]. Клеточное старение вызывается увеличением экспрессии репрессоров клеточного цикла, таких как p16INK4a, ARF, p21 и p53 [Li et al., 2009a]. О том, что перечисленные факторы являются барьерами на пути к плюрипотентности свидетельствует тот факт, что иммортализованные посредством делеции гена p53 или локуса INK4a/ARF, вторичные меланоциты репрограммируются с эффективностью 65%, а вторичные фибробласты с эффективностью – 40% [Utikal et al., 2009].

Экспрессия экзогенных репрограммирующих факторов сама по себе представляет стресс для клетки: повреждения ДНК при интеграции вирусов вызывают активацию p53,

что приводит либо к остановке клеточного цикла, либо к запуску программы апоптоза. И хотя в норме эта реакция крайне важна для предотвращения неконтролируемого клеточного деления, при репрограммировании, такая система защиты клеток совсем не играет на руку исследователям. В добавок ко всему, среди факторов репрограммирующего коктейля – два онкогена, c-Myc и Klf4. Только те клетки, которые, в конечном счете, могут справиться со стрессом, способны стать плюрипотентными. Многими группами исследователей было показано, что репрессия p53 (*Trp53*), приводит к увеличению эффективности репрограммирования [Banito et al., 2009; Hong et al., 2009; Kawamura et al., 2009; Zhao et al., 2008]. Однако полного понимания задействованных механизмов еще нет. Согласно работе Джейкоба Ханны, клетки, нокаутные по p53, демонстрируют повышенную скорость пролиферации [Hanna et al., 2009], а, следовательно, имеют большую вероятность быть репрограммированными: либо за счет увеличения общего количества клеток, либо за счет эпигенетических модификаций, происходящих во время деления клеток. Другие авторы придерживаются альтернативной гипотезы, не связывающей позитивное влияние репрессии p53 на репрограммирование с ускорением клеточного деления [Utikal et al., 2009]. Так, оказалось, что при культивировании дефицитных по p53 клеток мыши в среде с низким содержанием сыворотки, их скорость деления соответствует таковой для клеток дикого типа, однако эффективность репрограммирования выше.

1.2.1.2 Мезенхимально-эпителиальный переход

На молекулярном уровне мезенхимально-эпителиальный переход (МЭП) характеризуется подавлением экспрессии генов *Snai1/2*, *Zeb1/2*, за счет действия экзогенных факторов OCT4 и SOX2 [Mikkelsen et al., 2008; Sridharan et al., 2009; Stadtfeld et al., 2008; Wang et al., 2013a]. Гены *Snai1/2*, *Zeb1/2* кодируют транскрипционные факторы, поддерживающие фенотип мезенхимальных клеток путем подавления программы «эпителиальных» генов [Thiery et al., 2009]. В свою очередь, гены-маркеры эпителиальных клеток, такие как *Cdh1* и *Epcam*, участвующие в образовании плотных контактов эпителиальных клеток, увеличивают свою активность при МЭП. Нокаут гена *Cdh1* также, как и сверхэкспрессия гена *Snai1* препятствуют репрограммированию [Li et al., 2010; Samavarchi-Tehrani et al., 2010].

Неожиданно, среди активаторов МЭП оказался белок NUP210, принимающий участие в формировании ядерных пор, причем преимущественно в эпителиальных клетках [Eriksson, Rustum, Hallberg, 2004]. В отсутствии его экспрессии формирование колоний ИПСК не происходит: клетки сохраняют мезенхимальные характеристики [Hansson et al., 2012].

Интересен факт, что форма клеток сама по себе может оказывать влияние на способность клеток к репрограммированию [Downing et al., 2013]. В данной работе исследователи репрограммировали фибробласты, культивирующиеся на специальных полимерных подложках с параллельными бороздками шириной 10 микрометров. Форма клеток, растущих в бороздках была более вытянутая. Оказалось, что эффективность репрограммирования клеток в бороздках в 4 раза выше по сравнению с контрольными клетками, растущими на обычной плоской поверхности. Проанализировав экспрессию генов, исследователи сделали вывод, что увеличение эффективности репрограммирования связано именно с ускорением мезенхимально-эпителиального перехода; в клетках с более вытянутой формой значительно увеличивалась экспрессия эпителиальных генов (*E-cad*, *Epcam*, *Krt8*, *Ocln* и *Cldn3*), в то время как экспрессия генов-маркеров мезенхимальных клеток (*Tgfbr1*, *Snai1/2*, *Vim* и *Itgb1*) уменьшалась [Downing et al., 2013].

Помимо описанного выше механического воздействия на клетки, модуляция активности сигнальных путей также играет важную роль в запуске мезенхимально-эпителиального перехода. Так, например, сигнальный путь TGF β (трансформирующий ростовой фактор бета) контролирует дифференцировку мезенхимальных клеток и играет важную роль в эпителио-мезенхимальном переходе (процесс, обратный МЭП) [Maherali, Hochedlinger, 2009; Xu, Lamouille, Derynck, 2009]. Поэтому использование ингибиторов этого сигнального пути, таких как E-616452, E-616451 и SB-431542, позволяет увеличить эффективность репрограммирования, способствуя МЭП даже при исключении из репрограммирующего коктейля транскрипционного фактора Sox2 [Ichida et al., 2009; Woltjen, Stanford, 2009]. Компоненты сигнального пути TGF β являются мишенями для кластера микро РНК 302-367. Дополнительная экспрессия этих микро РНК в ходе репрограммирования значительно увеличивает эффективность получения колоний ИПСК [Anokye-Danso et al., 2011; Lin et al., 2009b; Miyoshi et al., 2011]. И достигается это именно за счет увеличения экспрессии *E-cad* и благоприятствования МЭП. Важно отметить, что в последних двух работах исследователям удалось получить колонии плюрипотентных клеток без использования транскрипционных факторов.

И если сигнальный путь TGF β препятствует МЭП, то сигнальный путь костного морфогенетического белка BMP, наоборот, способствует ему, а, следовательно, и репрограммированию клеток в целом. Достигается это, по-видимому, за счет синергетического действия компонентов сигнального пути BMP и репрограммирующих факторов OKSM, приводящего к активации нескольких семейств микро РНК – miR-205 miR-200. Эти микро РНК, в свою очередь, и способствуют мезенхимально-эпителиальному переходу, активируя экспрессию *Cdh1*, *Epcam* и *Ocln*, и подавляя экспрессию *Zeb1/2*, *Snai1*

и *Slug* [Korpal, Kang, 2008; Samavarchi-Tehrani et al., 2010]. Важно отметить также, что микро РНК из этого семейства, в частности miR-141, miR-200 и miR-429, способствуют поддержанию плюрипотентности эмбриональных стволовых клеток мыши, даже если они культивируются в среде для дифференцировки [Lin et al., 2009a]. Кроме того, BMP может заменить Klf4 в репрограммирующем коктейле и индуцировать плюрипотентность в фибробластах при использовании всего лишь двух репрограммирующих факторов *Oct4* и *Sox2* [Chen et al., 2011].

Важность мезенхимально-эпителиального перехода подчеркивает тот факт, что транскрипционный фактор OCT4, один из главных факторов в репрограммирующем коктейле, может быть успешно заменен Е-кадгерином [Redmer et al., 2011]. Механизмы действия Е-кадгерина как репрограммирующего фактора еще не вполне ясны. Однако, есть предположение, что помимо влияния на МЭП, сверхэкспрессия Е-кадгерина предотвращает накопление β -катенина в ядре, сопровождающее старение клеток [Ho et al., 2013].

Важно отметить, что не всем клеткам в ходе репрограммирования необходимо проходить мезенхимально-эпителиальный переход. Например, кератиноцитам и гепатоцитам, принадлежащим к эпителиальным тканям. Ожидается, эти клетки репрограммируются до 100 раз эффективнее, а кроме того, быстрее, чем фибробласты – представители клеток мезенхимального типа [Aasen et al., 2008; Aoi et al., 2008]. Тем не менее, большинство клеток эпителиального типа, также оказываются неспособными репрограммироваться и образовать плюрипотентного потомка. Таким образом, очевидно, что мезенхимально-эпителиальный переход не единственное критическое событие в процессе репрограммирования.

I.2.2 Промежуточная стадия репрограммирования

На промежуточную стадию приходится вторая волна наибольших изменений в транскрипционном профиле [Hansson et al., 2012; Polo et al., 2012; Samavarchi-Tehrani et al., 2010]. Постепенно увеличивается экспрессия генов, контролирующих раннее развитие, включая гены *Fbxo15*, *Nanog*, *Sall4*, *Oct4* et al, тогда как тканеспецифичные гены дифференцированных клеток репрессируются [Buganim et al., 2012; Golipour et al., 2012; Polo et al., 2012].

Маркерами начального периода этой стадии для клеток мыши и человека служат поверхностные антигены SSEA-1 (Stage-specific embryonic antigen-1) и TRA-1-16 (tumor related antigen) соответственно [Brambrink et al., 2008; Polo et al., 2012; Stadtfeld et al., 2008]. Около 20% исходных инфицированных фибробластов человека экспрессируют TRA-1-16 на 7-ой день репрограммирования [Tanabe et al., 2013]. Однако впоследствии большинство

TRA-1-16-позитивных клеток перестают экспрессировать этот маркер, а скорость их пролиферации уменьшается. Это свидетельствует о том, что клетки, начав репрограммироваться и вступив в стадию инициации, затем столкнулись с неким барьером. Один из классических способов повышения эффективности репрограммирования – дополнительная сверхэкспрессия фактора LIN28, позволяет увеличить количество колоний ИПСК именно за счет предотвращения реверсии экспрессии TRA-1-60, а не за счет увеличения общего количества клеток, которые экспрессируют этот маркер на стадии инициации [Tanabe et al., 2013].

На промежуточной стадии происходит завершение мезенхимально-эпителиального перехода, и клетки формируют небольшие колонии [Smith et al., 2010]. Кроме того, происходит переключение пути синтеза АТФ: с окислительного фосфорилирования в митохондриях на гликолиз, что, стоит отметить, характерно для быстроделющихся плюрипотентных клеток [Heiden Vander, Cantley, Thompson, 2009; Prigione et al., 2010]. В ходе репрограммирования такое переключение клеточного метаболизма достигается путем снижения экспрессии генов, кодирующих компоненты комплекса переноса электронов вкупе с увеличением экспрессии генов, позитивно регулирующих гликолиз (*Pfkf1*, *Gpi*, *Pgk1*) [Hansson et al., 2012; Zhang et al., 2012]. Увеличение гликолитического потока способствует увеличению эффективности репрограммирования [Folmes et al., 2011].

Изменение в ходе репрограммирования метаболических и морфологических характеристик на характеристики, свойственные плюрипотентным клеткам, выглядит вполне логичным и закономерным процессом: клетки проходят путь, обратный тому, что прошли в ходе дифференцировки. Однако, оказалось, что промежуточная стадия несколько сложнее этого представления и даже сравнивается некоторыми исследователями с черным ящиком [Cyranoski, 2014]. Было показано, что при репрограммировании эмбриональных фибробластов мыши в промежуточную фазу активируются гены, характерные для эктодермальных клеток (*Krt6a*, *Krt17*, *Ehf*, *Ovo1*) [O'Malley et al., 2013]. Такой же феномен засвидетельствован и для неполностью репрограммированных В-лимфоцитов мыши [Mikkelsen et al., 2008]. При этом ни исходные клетки, ни полученные из них ИПСК, не экспрессируют указанные гены, что свидетельствует об их специфичности для промежуточной стадии. Аналогичная ситуация наблюдается и при репрограммировании клеток человека. Исследователи под руководством Синьи Яманаки обнаружили, что в промежуточную стадию TRA-1-16-позитивные клетки экспрессируют гены примитивной мезодермы (*Brachyury*, *Eomes*, *Cer1*, *Lhx1*, *Fgf4* et al) [Takahashi et al., 2014]. Закономерен вопрос, насколько экспрессия этих генов важна для репрограммирования? Возможно ли, что она имеет стохастическую природу и не играет никакой роли в индукции

плюрипотентности? Оказалось, при коэкспрессии мезодермальных факторов вместе с репрограммирующими факторами эффективность репрограммирования повышается. Наибольший эффект показан для гена *Foxh1* – он сравним с репрессией p53. Достигается это за счет активации МЭП путем усиления экспрессии *Ersat* и подавления CD13 (маркер фибробластов). Кроме того, FOXH1 активирует экспрессию поздних маркеров плюрипотентности, таких как DPPA4, DNMT3B, LIN28A, ZFP42 и SOX2. Авторы даже делают вывод, что для успешного репрограммирования клетки *должны* пройти состояние, характерное для клеток первичной полоски в ходе гаструляции [Takahashi et al., 2014].

Другим важным событием, необходимым для перехода к стабильному плюрипотентному состоянию, является инактивация трансгенных конструкций с репрограммирующими факторами. В этом участвует механизм, задействованный в инактивации ретровирусных последовательностей в ЭСК [Stadtfield et al., 2008]. Клетки, которые сохраняют высокий уровень экспрессии экзогенных факторов, не активируют ряд ключевых генов плюрипотентности, таких как *Sox2* и *Dppa2*, а их способность к дифференцировке значительно снижена [Brambrink et al., 2008; Golipour et al., 2012; Sommer et al., 2009]. Интересно, что большинство генов, задействованных в переходе в стадию стабилизации, несущественны для поддержания плюрипотентности и связаны с регуляцией организации цитоскелета и сегрегацией хромосом.

Таким образом, пролонгированная экспрессия экзогенных факторов является еще одним барьером для репрограммирования. В пользу этого свидетельствует также тот факт, что сверхэкспрессия *Oct4*, *Sox2* в ЭСК негативно влияет на плюрипотентный статус, вызывая дифференцировку клеток [Niwa, Miyazaki, Smith, 2000; Wang et al., 2012].

1.2.3 Финальная стадия репрограммирования/Фаза стабилизации

К финальной стадии репрограммирования относят изменения, связанные не с индукцией, но с поддержанием плюрипотентности в клетках [Stadtfield, Hochedlinger, 2010]. Полностью репрограммированные клетки характеризуются экспрессией генов *Sox2*, *Utf1*, *Esrrb*, *Dppa2* и *Lin28*. Указанные гены являются более строгими маркерами полноты репрограммирования, чем *Oct4* и *Nanog*, однако даже их активация не является стопроцентной гарантией успешного репрограммирования [Buganim et al., 2012; Golipour et al., 2012; Polo et al., 2012]. В дифференцированных соматических клетках локусы этих генов несут метки неактивного хроматина и, вероятно, напрямую не взаимодействуют с репрограммирующими факторами [Soufi, Donahue, Zaret, 2012]. Таким образом, для того чтобы стать плюрипотентной, клетке необходимо устранить эпигенетические барьеры в виде меток неактивного хроматина с ключевых генов плюрипотентности.

В полностью репрограммированных клетках происходит удлинение теломер до уровня, характерного эмбриональным стволовым клеткам [Marion et al., 2009; Stadtfeld et al., 2008]. Кроме того, в ИПСК мыши происходит реактивация транскрипционно-неактивной X хромосомы [Maherali et al., 2008; Stadtfeld et al., 2008]. Интересно, что в ИПСК человека неактивная X-хромосома сохраняется, что свидетельствует о некоторых различиях в ходе процесса репрограммирования в клетках человека и мыши [Tchieu et al., 2010]. Масштабные эпигенетические изменения приводят к установлению бивалентных доменов хроматина, свойственных эмбриональным стволовым клеткам [Huang et al., 2015]. И также, как и эмбриональные, индуцированные плюрипотентные стволовые клетки приобретают способность дифференцироваться в производные всех трех зародышевых листков как *in vitro*, так и *in vivo*, давать вклад в формировании химер, а также способность развиваться в целый организм при инъекции в тетраплоидные бластоцисты [Kang et al., 2009; Zhao et al., 2009]. Тем не менее, существуют работы, сообщающие о значительных различиях индуцированных плюрипотентных и эмбриональных стволовых клеток по эпигенетическому профилю и уровню экспрессии генов [Deng et al., 2009; Lister et al., 2011]. Этот факт вынудил некоторых исследователей выделить ИПСК в особый тип плюрипотентных клеток, обладающий своим, уникальным паттерном экспрессии [Chin et al., 2009]. Однако стоит отметить, что исследователи, сообщающие о значимых различиях между ИПСК и ЭСК, рассматривали в своих работах относительно небольшое количество линий (в среднем не больше 10) [Kim et al., 2011; Ohi et al., 2011]. Тогда как авторы, анализирующие большее количество линий, говорят о почти полной неразличимости ИПСК и ЭСК [Bock et al., 2011; Guenther et al., 2010]. Из этого следует, что различия ИПСК и ЭСК, возникающие вследствие неполноты репрограммирования, не являются фундаментальной проблемой этого процесса и в большей степени обусловлены техническими причинами, а именно различными способами индукции плюрипотентности и условиями культивирования в каждой лаборатории [Newman, Cooper, 2010; Yamanaka, 2012].

Описанные выше исследования, посвященные всестороннему изучению процесса репрограммирования, были сделаны на разных моделях и с использованием разных репрограммирующих факторов. Однако, даже стехиометрия факторов оказывает влияние на динамику репрограммирования. Поэтому для того, чтобы получить панорамный вид на процесс репрограммирования представляется полезным провести унифицированный анализ одной культуры репрограммируемых клеток. В конце 2014 года в журналах *Nature* и *Nature Communications* вышло сразу 5 статей, посвященных проекту с громким названием

Project Grandiose. Данный проект был запущен в 2010 году большой коллаборацией ученых из 8 институтов Канады, Австралии, Нидерландов и Кореи.

1.3 «Project Grandiose». Альтернативные пути репрограммирования

В рамках проекта был проведен многомерный анализ клеточного репрограммирования к плюрипотентному состоянию на разных этапах, включая анализы метилома, транскриптома, протеома, эпигенетических меток хроматина, а также анализ малых и длинных некодирующих РНК.

В качестве исходных клеток исследователи использовали одну из вторичных репрограммируемых систем – эмбриональные фибробласты мыши с факторами Яманаки под контролем доксициклин-зависимого промотора [Woltjen et al., 2009]. В течение 8 дней клетки культивировались в среде с высоким содержанием доксициклина (1500 нг/мл), следовательно, уровень экспрессии репрограммирующих факторов также был высок. Затем часть клеток переводили на среду с низким содержанием доксициклина (50 нг/мл). Интересно, что только в этом случае исследователи наблюдали образование колоний ИПСК. Тогда как постоянное культивирование клеток в среде с высоким уровнем доксициклина приводило к образованию другого класса плюрипотентных клеток, названного авторами F-классом, где F означает “fuzzy” т.е. нечеткий, неясный.

В отличие от колоний ИПСК, представляющих собой плотные колонии С-класса (С от “compact”), колонии клеток F-класса рыхлые и с нечеткими границами. F-клетки имеют высокий пролиферативный потенциал, дифференцируются в производные всех трех зародышевых листков и формируют тератомы. Однако, они не проходят самый строгий тест на плюрипотентность - не дают вклад в образование химер. Кроме того, для поддержания плюрипотентности F-клеткам необходима экспрессия репрограммирующих факторов. В противном случае клетки дифференцируются либо в разные типы нейронов при бессывороточном культивировании, либо в мезодермальные и эндодермальные предшественники при культивировании в среде с сывороткой [Tonge et al., 2014]. Здесь очевиден вопрос: возможно F-клетки просто неполностью репрограммированы? Во многих более ранних работах уже было показано существование/получение частично или неполностью репрограммированных клеток [Mikkelsen et al., 2008; Papp, Plath, 2013]. Однако, авторы Project Grandiose однозначны: F-клетки имеют отличный от ИПСК профиль экспрессии генов, который поддерживается при длительном культивировании (т.е. стабильный). Следовательно, F-клетки представляют собой особый тип плюрипотентных клеток, не описанный ранее. Среди свойств этого типа клеток – независимость самообновления популяции от сигнальных путей LIF и JAK. А кроме того, чрезвычайно

высокая скорость пролиферации даже по сравнению с ЭСК мыши: F-клетки будучи смешанными с ЭСК в соотношении 1:99 уже через 3 пассажа являли собой большинство. Данный факт вкупе с ослабленной межклеточной адгезией, делает F-клетки чрезвычайно перспективными для использования в биореакторах.

Авторы проекта детально проанализировали изменения в репрограммируемых клетках, приводящие к установлению фенотипа плюрипотентных клеток C- и F-классов.

Транскриптом и протеом [Benevento et al., 2014]. Сравнительный анализ показал, что в F-клетках снижена экспрессия ряда маркеров плюрипотентности, таких как SALL4, UTF1, DPAP4, LEFTY and ESRRB. При этом, один из главных регуляторов плюрипотентности, NANOG, экспрессируется на уровне, сравнимом с таковым в ЭСК и клетках C-класса. Это значит, что даже частичной активации транскрипционной сети, поддерживающей плюрипотентность в ЭСК, достаточно для способности дифференцироваться в производные трех зародышевых листков и давать тератомы. Как и в ранее опубликованных работах, посвященных изучению изменений транскриптома и протеома в ходе репрограммирования [Hansson et al., 2012; Polo et al., 2012], авторы Project Grandiose выделили два пика наибольших изменений. Независимо от траектории репрограммирования в первые 48 часов увеличивается экспрессия генов, связанных с пролиферацией и метаболизмом клеток, с регуляцией экспрессии и процессингом РНК. В то же время, уменьшается экспрессия мезенхимальных маркеров, и белков, участвующих в везикулярном транспорте, гликозилировании и энергетическом метаболизме. При этом, большинство активированных генов, являются «мишенью» транскрипционного фактора c-Myc, что говорит о его ключевой роли на первых этапах репрограммирования. На промежуточном этапе репрограммирования (с пятого по 11 день) в клетках увеличивается уровень эпителиальных маркеров и компонентов сигнального пути ретиноевой кислоты. Затем траектории репрограммирования разделяются: клетки C-класса характеризуются увеличением экспрессии модификаторов хроматина и транскрипционных факторов, поддерживающих плюрипотентность; в F-клетках же наблюдается еще большее увеличение «акселераторов клеточного цикла» и снижение экспрессии эпителиальных маркеров, что и обуславливает их мезенхимально-подобную морфологию. Мастер-регуляторами программы C-клеток на поздних этапах репрограммирования являются факторы – OCT4, SOX2, NANOG и CHD1, а программы F-клеток – C-MYC, JARID1A и KDM5B.

Микро РНК и длинные некодирующие РНК [Clancy et al., 2014]: Известно, что микро РНК и длинные некодирующие РНК играют большую роль при репрограммировании клеток. Они участвуют в мезенхимально-эпителиальном переходе, блокируют программу

поддержания клетками дифференцированного состояния, а также необходимы для поддержания плюрипотентности [Li et al., 2011; Lüningschrör et al., 2013; Papp, Plath, 2013]. Исследователи из группы Project Grandiose показали, что изменения в экспрессии микроРНК также имеют два характерных пика, в первые 48 часов после старта экспрессии факторов Яманаки и спустя 8 дней. И несмотря на то, что набор экспрессирующихся микроРНК значительно различается между клетками С- и F-классов, процессы, регулируемые этими микроРНК, одинаковы. Это апоптоз, пролиферация, клеточная морфология, дифференцировка и регуляция экспрессии генов посредством изменения структуры хроматина. Возможно, это свидетельствует о разных траекториях, по которым клетки достигают плюрипотентного состояния. Особое внимание авторы уделили кластерам микроРНК, регулирующим МЭП и ЭМП (мезенхимально-эпителиальный и эпителио-мезенхимальный переходы соответственно). Вполне ожидаемо, на 8-ой день репрограммирования пик экспрессии про-эпителиальных miR-192/194 и miR-200, способствующих МЭП [Li et al., 2011; Polo et al., 2012], совпадает с пиком экспрессии E-кадгерина [Benevento et al., 2014]. Однако затем их экспрессия снижается, при увеличении экспрессии мезенхимальных генов, что согласуется с ранее опубликованными результатами [Hansson et al., 2012]. При этом, в F-клетках наблюдается повышенный уровень экспрессии про-мезенхимальных микроРНК, таких как miR-10b-5p, а морфология скорее напоминает морфологию мезенхимальных клеток [Tonge et al., 2014]. В целом, профиль экспрессии микроРНК клеток С-класса (т.е. привычные нам ИПСК) схож с таковым для ЭСК мыши, золотым стандартом плюрипотентности. В то время как клетки F-класса отличаются отсутствием таких специфичных для ЭСК микроРНК как miR 106b, miR-302/367 и miR-290/295 [Leonardo et al., 2012]. Таким образом, те свойства плюрипотентных клеток, которыми обладают F-клетки, обеспечиваются набором микроРНК, отличным от того, что поддерживает плюрипотентность в ЭСК.

Эпигенетические модификации хроматина [Hussein et al., 2015]. Вскоре после начала экспрессии репрограммирующих факторов происходит глобальное снижение уровня H3K27me3 (метка неактивного хроматина), достигающее минимума на восьмой день. Это обусловлено повышенной экспрессией гистоновых деметилаз *Kdm6a (Utx)*, *Jhdm1d* и *Phf8* вкупе со снижением активности комплекса PRC2 (Polycomb Repressive Complex 2), способствующего триметилированию H3K27. Таким образом, в С- и F-клетках начинаются масштабные эпигенетические изменения, приводящие к формированию открытого, транскрипционно-активного хроматина. Среди активирующихся при этом генов – гены *Cldn4*, *Lefty2* и *Lin28b*, чей высокий экспрессионный статус характерен для плюрипотентных клеток. После восьмого дня репрограммирования происходит восстановление уровня

H3K27me3 до изначального уровня в эмбриональных фибробластах. В клетках, репрограммирующихся по F-траектории, часть ново-активированных генов сохраняет свою экспрессию, в то время как другая часть инактивируется либо за счет вновь приобретенной репрессивной метки H3K37me3, либо за счет метилирования ДНК [Lee et al., 2014]. И наоборот, некоторые гены, в частности *Nkx2-3*, *Isl1* и *Cxcr4*, отвечающие за нейральную дифференцировку и клеточную миграцию, инактивируются в клетках C-класса, но остаются активными в клетках F-класса.

Метилирование ДНК [Lee et al., 2014]. С помощью полногеномного бисульфитного секвенирования (MethylC-seq) исследователи проанализировали дифференциально метилированные районы (DMR – differentially methylated regions) и выявили интересную закономерность: наличие как активных (H3K4me3), так и неактивных модификаций хроматина (H3K27me3) коррелирует с низким уровнем метилирования DMR (96,9%). И наоборот, отсутствие модификаций гистонов, как правило, сопряжено с высоким уровнем метилирования (79,7%). Это говорит о возможном наличии интерферирующих механизмов эпигенетической регуляции. В соответствии с этим, в ходе репрограммирования гены, содержащие неметилированные промоторы, приобретают H3K27me3 и активируются на ранних этапах (*Ersat* и *Cgh1*). А гены, имеющие в дифференцированном состоянии сильно метилированные промоторы, активируются только после деметилирования ДНК, происходящего на поздних этапах репрограммирования (*Essrb*, *Nr5a2*, *Pecam1* et al.). Таким образом, метилирование промоторов ключевых генов плюрипотентности является одним из главных барьеров при репрограммировании.

Важно отметить то, что авторы данной работы не побоялись заявить о так называемом «слоне в комнате», - проблеме которую предпочитают не обсуждать. И заключается она в том, что независимо от метода репрограммирования многие получаемые колонии не являются ИПСК и представлены разными типами клеток с разным уровнем плюрипотентности [Baumann, 2015]. Кроме того, в зависимости от целей исследования, клетки с альтернативными траекториями репрограммирования могут обладать некоторым преимуществом перед ИПСК. Так, например, F-клетки быстрее делятся, их дешевле культивировать (без дорогостоящего соединения LIF), а кроме того, они хорошо дифференцируются в нейроны без специальных индукторов.

На сегодняшний день ИПСК являются наиболее перспективным источником клеток для тканезаместительной терапии. Однако на пути применения ИПСК в клинической практике стоят важнейшие препятствия. Во-первых – для доставки репрограммирующих

факторов в клетки, используют конструкции, модифицирующие геном клеток. Во-вторых, экспрессия генов *Oct4*, *Sox2*, *Klf4* и *c-Мус*, использующихся для индукции плюрипотентности, ассоциирована с развитием множества видов опухолей. Уменьшение числа факторов, необходимых для репрограммирования, и использование неинтегративных методов их доставки критически снижает и без того низкую эффективность получения ИПСК. Именно поэтому так важно исследовать, какие барьеры на пути к плюрипотентности существуют в клетке, и как их можно преодолеть. Ниже в таблице просуммированы сведения о некоторых факторах, позитивно или негативно влияющих на индукцию плюрипотентности в дифференцированных клетках. Факторы, сверхэкспрессия которых приводит к увеличению, а репрессия – к уменьшению эффективности репрограммирования, обозначены как энхансеры репрограммирования. Их активация на разных этапах репрограммирования необходима для полного репрограммирования соматических клеток. И наоборот, ингибиторами репрограммирования являются факторы, подавление активности которых способствует индукции плюрипотентности.

Таблица 1. Энхансеры и ингибиторы индукции плюрипотентности. Зеленым выделены транскрипционные факторы; желтым – факторы, осуществляющие эпигенетические модификации, изменяющие структуру хроматина; голубым – сигнальные и метаболические пути; красным – кластеры микро-РНК; серым - другое. Красным шрифтом выделены факторы, роль которых в репрограммировании противоречива и не выяснена до конца.

Роль в индукции плюрипотентности	Ключевой участник	Механизм действия	Примечания	Источник
Энхансеры репрограммирования	C/EBPa (CCAAT/enhancer-binding protein-α) Транскрипционный фактор	Активирует экспрессию TET2 (гидроксиметилтрансфераза), тем самым способствует деметилированию ДНК и формированию открытой структуры хроматина. Облегчает связывание OSKM факторов с сайтами-мишенями.	Эффективность репрограммирования достигает 95% для В-лимфоцитов мыши (нет эффекта для МЭФ)	[Stefano Di et al., 2013]
	UTF1 (Undifferentiated Embryonic Cell Transcription Factor 1)	Точный механизм не известен. Возможно через деградацию мРНК генов Arf и Ink4a способствует активной пролиферации [Jia et al., 2012].	Эффективность репрограммирования в 100 раз выше по сравнению с конвенциональн	[Yang, Chang, Rana, 2014; Zhao et al., 2008]

Транскрипционный фактор		ым методом (для фибробластов человека)	
FOX1 (forkhead box H1) Транскрипционный фактор	Способствует MET и активирует экспрессию поздних маркеров плюрипотентности.	Эффект показан на клетках человека.	[Takahashi et al., 2014]
GLIS1 (Gli-like transcription factor Glis1) Транскрипционный фактор	Напрямую взаимодействует с OSKM факторами. Увеличивает экспрессию генов, задействованных в поддержании плюрипотентности.	Эффективность репрограммирования в 2-30 раз выше по сравнению с конвенциональным методом	[Maekawa et al., 2011]
SNAIL Транскрипционный фактор	Подавляет экспрессию miR let-7 за счет связывания с промотором.	Парадоксально, но является ключевым фактором, поддерживающим мезенхимальный фенотип клеток, т.е. репрессирует МЭП.	[Unternaehrer et al., 2014]
ZFP217 (zinc finger protein 217) Транскрипционный фактор	Активирует экспрессию генов-регуляторов плюрипотентности, включая <i>Nanog</i> , <i>Sox2</i> , <i>Klf4</i> , и <i>c-Мyc</i> . Взаимодействует с METTL3 (Methyltransferase Like 3) и предотвращает METTL3-опосредованное метилирование (m ⁶ A) и деградацию мРНК факторов плюрипотентности.	Эффект показан на ЭФМ	[Aguilo et al., 2015]
Kdm3a, Kdm3b, Kdm4c, Kdm4b/2b И Kdm6a лизин-К-специфические деметилазы	Деметилируют H3K9me2/me3 в локусах генов плюрипотентности, способствуют формированию активного хроматина.		[Chen et al., 2012a; Mansour et al., 2012]
Lhdn1a/1b Лизин-К-специфические деметилазы	Увеличивают скорость пролиферации, репрессируют локус <i>Ink4/Arf</i> , тем самым ингибируя клеточное старение. Взаимодействуют с OCT4 и активируют экспрессию кластера <i>mir 302/367</i> .		[Chen et al., 2012a; Wang et al., 2011]

MOF (males absent on the first_ гистон ацетилтрансфераза	Взаимодействует с WDR5 (H3K4 метилтрансфераза) и активирует экспрессию эндогенного <i>Oct4</i>	Увеличение эффективности репрограммирования для ЭФМ	[Mu et al., 2015]
TET1/TET2 (Ten-eleven translocation methylcytosine dioxygenase 1) - деоксигеназа TGD (Thymine DNA glycosylase) – тимин-ДНК гликозилаза	Способствуют деметилированию промоторов и активации ключевых генов плюрипотентности. TET-TGD необходимы на раннем этапе репрограммирования: активируют miR-200 и способствуют МЭП.	Увеличение эффективности репрограммирования для фибробластов человека и мыши	[Costa et al., 2013; Doege et al., 2012; Gao et al., 2013; Hu et al., 2014; Kohli, Zhang, 2013]
Brg1 Baf155 из комплекса BAF (Brg1 Associated Factors) – хроматин-ремоделирующий комплекс	Способствуют переходу хроматина в активное состояние. Усиливают связывание репрессирующих факторов с промоторами активируемых генов.		[Singhal et al., 2010]
Комплекс PRC2 (Polycomb repressive complex 2) + субъединицы: JARID2, MTF2 и PRC2p48 Хроматин-ремоделирующий комплекс	Репрессируют линейно-специфичные гены (<i>TGFβR2</i> , <i>Snail1</i> , and <i>TGFβ1</i> для фибробластов) за счет триметилирования H3K27 промоторных регионов.		[Zhang et al., 2011]
RYBP (RING1 and YY1-binding protein) – компонент комплекса PRC1	На раннем этапе репрограммирования способствует репрессии линейно-специфичных генов (<i>Egr4</i> , <i>Jun</i> , <i>Tril</i> , <i>Klf15</i> и <i>Rasd1</i>) за счет установления неактивной метки H2AR119Ub1. На позднем этапе репрограммирования действует как коактиватор транскрипции: взаимодействует с OCT4 и способствует экспрессии генов, участвующих в индукции плюрипотентности (<i>Kdm2b</i>).	Эффект показан на ЭФМ	[Li et al., 2017; Nefzger, Polo, 2017]

Mbd3 (хроматин- ремоделирующий комплекс NuRD)	Точный механизм неизвестен. Возможно действует как медиатор транскрипции, взаимодействуя с ТФ, поддерживающими плюрипотентность (с NANOG, например). Может супрессировать mTOR, активируя аутофагию, необходимую на раннем этапе репрограммирования.	Увеличение эффективности репрограммирования нейральных стволовых клеток	[Santos Dos et al., 2014; Wang et al., 2013b]
BCL-2 (B-cell lymphoma 2) Регулятор апоптоза	Подавление апоптоза.		[Kawamura et al., 2009]
Сигнальный путь белков морфогенеза костной ткани (Bmp)	На раннем этапе репрограммирования способствует МЭП за счет увеличения экспрессии E-кадгерина и окклюдина. Подавляет p16/INK4A опосредованное клеточное старение. На поздних этапах снижает эффективность репрограммирования.		[Chen et al., 2011; Hamasaki et al., 2012; Hayashi et al., 2016; Samavarchi-Tehrani et al., 2010]
Сигнальный путь Wnt/ β -катенина	На раннем этапе препятствует репрограммированию. Возможно за счет конкурирования TCF3 (один из компонентов пути) с SOX2 за связывание с сайтами-мишенями [Zhang et al., 2013]. Необходим на позднем этапе репрограммирования – поддерживает активную транскрипцию <i>Oct4</i> .		[Aulicino et al., 2014; Ho et al., 2013; Zhang et al., 2014]
Сигнальный путь PI3/АКТ PI3К – фосфотидилинози тол-3-ОН протеинкиназа АКТ – протеинкиназа В	Способствует переключению синтеза АТФ: с окислительного фосфорилирования на гликолиз за счет увеличения экспрессии гликолитических генов <i>Glut1</i> , <i>Hk2</i> и <i>Pfk1</i> . Сигнальный каскад ERas-Akt способствует пролиферации и за счет репрессии FoxO1 подавляет апоптоз.	Тиреоидный гормон Т3, увеличивающий эффективность репрограммирования соматических клеток человека, напрямую активирует экспрессию генов,	(Zhu et al., 2010; M. Chen et al., 2012; Yu et al., 2014)

			вовлеченных в метаболизм глюкозы, а также активирует PI3/АКТ-сигнальный путь.	
FGF2 фактор роста фибробластов	На раннем этапе способствует подавлению экспрессии генов <i>Slug</i> , <i>Snail</i> , <i>Zeb1</i> , <i>Zeb2</i> , и <i>Cdh2</i> , а также <i>Col1a1</i> , благоприятствуя МЭП. Ингибирует p16, в результате чего увеличивается скорость пролиферации. На поздних этапах снижает эффективность репрограммирования.			[Jiao et al., 2013]
сигнальный путь JAK/STAT JAK – тирозинкиназа Janus STAT (signal transducer and activator of transcription) - ТФ	Блокирует de novo метилтрансферазу DNMT1 и гистоновые деацетилазы HDAC2, HDAC3 и HDAC8. Способствует деметилированию промоторов ключевых генов плюрипотентности. Крайне важен на поздних этапах репрограммирования – для активации эндогенного <i>Oct4</i> . Участвует в репрессии ретровирусных трансгенов.	Увеличение эффективности репрограммирования для клеток мышей		[Tang et al., 2012; Tang, Tian, 2013; Yang et al., 2010]
NAD ⁺ зависимые метаболические пути NAD – никотинамидадениндинуклеотид, кофермент	НАМ (никотинамид) - предшественник NAD ⁺ снижает уровень активных форм кислорода и оксидативный стресс, способствуя увеличению выживаемости клеток. Подавляет апоптоз и клеточное старение.	Эффект показан на фибробластах человека		[Son et al., 2013]
miR-290-295 miR-302-367 miR-106a/363	Активируются репрограммирующими факторами. Способствуют МЭП и активации генов плюрипотентности. Подавляют апоптоз и увеличивают скорость	Позволяют репрограммировать клетки без экзогенных транскрипционных факторов [Anokye-Danso et al., 2011]		[Greer Card et al., 2008; Judson et al., 2009; Li et al., 2011; Liao et al., 2011; Lin et al.,

		пролиферации. Ингибируют TGFβ -сигнальный путь.		2008; Subramanyam et al., 2011]
	miR-205 miR-200	Репрессируют мезенхимальные факторы Zeb1/Zeb2. Через активацию сигнального пути BMP стимулируют МЭП.		[Chen et al., 2011; Gregory et al., 2008; Samavarchi-Tehrani et al., 2010; Wiklund et al., 2011]
	miR-130 miR-301 miR-721	Подавляют экспрессию гена <i>Meox2</i> (гомеобокс-содержащий ТФ) (см. ниже).		[Pfaff et al., 2011]
	miR-106b miR-93	Посттранскрипционно репрессирует <i>Tgfb2</i> (рецептор TGFβ) и p21, способствуя МЭП и переход в S-фазу клеточного цикла.		[Li et al., 2011]
	MDM2 (Mouse double minute 2 homolog) убиквитин-лигаза	p53-зависимый механизм: способствует деградации белка p53. Предполагаемый p53-независимый механизм – через связывание с комплексом PRC2 подавление экспрессии линейно-специфичных генов [Wienken et al., 2016; Wienken, Moll, Dobbstein, 2017].		[Hong et al., 2009; Kawamura et al., 2009]
	CCL2 ((C-C motif) ligand 2) β-хемокин	Активирует комплекс гипоксия-индуцибельного фактора (HIF-1). Приводит к увеличению экспрессии генов <i>Klf4</i> , <i>Nanog</i> и <i>Tbx3</i> .	Увеличение эффективности репрограммирования для фибробластов мыши и человека	[Hasegawa et al., 2011; Hasegawa et al., 2014]
	Netrin-1	Блокирует DCC (Deleted in Colorectal Carcinoma) зависимый апоптоз (не связан с p53).		[Ozmadenci et al., 2015]
	Dmbx1, Hnf4g, Nobox и Asb4		Определены с помощью полногеномного скрининга на основе РНК-интерференции	[Yang, Chang, Rana, 2014]

Ингибиторы репрограммирования	Zfp281 (репрессор транскрипции)	Связывается с комплексом NuRD и репрессирует экспрессию гена <i>Nanog</i>		[Fidalgo et al., 2012]
	MEOX2 Транскрипционный фактор	Точный механизм неизвестен. Возможно действует как эффектор TGFβ – сигнального пути (Valcourt et al., 2007).		[Pfaff et al., 2011]
	Bright/Arid3A Транскрипционный фактор	Способствует клеточному старению через активацию сигнальных путей p53-p21 и p16INK4a –pRb. Напрямую связывается с промоторными/энхансерными последовательностями генов <i>Oct4</i> , <i>Sox2</i> и <i>Nanog</i> , препятствуя их активации.		[Popowski et al., 2014]
	DOT1L (H3K79 гистон метилтрансфераза)	Поддерживает метилирование H3K79 и активный транскрипционный статус линейно-специфичных генов (<i>Twist1</i> , <i>Snail1</i> , <i>Zeb1</i>). Подавляет МЭП.	Увеличение эффективности репрограммирования фибробластов и клеток крови	[Onder et al., 2012]
	Jmjd3 (гистон H3K27 деметилаза)	Способствует клеточному старению путем увеличения экспрессии p21 и Ink4a. Через убиквитин-лигазу Trim26 вызывает деградацию белка PHF20 (plant homeodomain finger protein 20), который необходим для реактивации эндогенного <i>Oct4</i> .		[Zhao et al., 2013]
	Mbd3 (хроматин-ремоделирующий комплекс NuRD)	Mbd3, связываясь с репрограммирующими факторами, привлекается к их генам-мишеням и формирует закрытую структуру хроматина. Блокирует действие репрограммирующих факторов.	Эффективность репрограммирования различных типов дифференцированных клеток мыши, нокаутных по Mbd3, достигает 100%. Синхронная траектория для всех клеток.	[Brumbaugh, Hochedlinger, 2013; Luo et al., 2013; Rais et al., 2013]
	p53 и p21	Активируют программу апоптоза, вызванную повреждениями ДНК при интеграции	Увеличение эффективности репрограммирования от 10 до	[Banito et al., 2009; Brosh et al., 2013;

		репрограммирующих факторов, а также сверхэкспрессией протоонкогенов c-Мус и Klf4. Снижают скорость пролиферации. Подавляют МЭП.	100 раз по сравнению с конвенциональным методом. До 80% [Utikal et al., 2009] для вторичных меланоцитов	Guo et al., 2014; Hanna et al., 2009; Kawamura et al., 2009; Li et al., 2009a; Marión et al., 2009; Okita et al., 2008; Rasmussen et al., 2014; Utikal et al., 2009; Zhao, Xu, 2010; Zhao et al., 2008]
	p16(Ink4a) p19(Arf) ингибиторы циклин-зависимых киназ	Активируют p53-p21 сигнальный путь, индуцируют арест пролиферации клеток на стадиях G1/G2 клеточного цикла		[Kawamura et al., 2009; Li et al., 2009a; Utikal et al., 2009]
	МАРК (p38) - сигнальный путь	Препятствует МЭП, индуцирует клеточное старение	Ингибитор PD0325901	[Li, Rana, 2012; Lin et al., 2009b; Silva et al., 2008]
	TGFβ-сигнальный путь	Препятствует МЭП. Активирует МАРК-сигнальный путь, увеличивает экспрессию мезодермальных генов, препятствуя МЭП.	Ингибиторы A-83-01 E616452 SB431542	[Hou et al., 2013; Li et al., 2010; Lin et al., 2009b; Maherali, Hochedlinger, 2009; Thiery, Sleeman, 2006; Zhu et al., 2011]
	Сигнальный путь Hippo LATS2 (Large Tumor Suppressor Kinase 2)	Препятствует репрограммированию за счет подавления активности TAZ (транскрипционного коактиватор). Точный механизм неизвестен.	Эффект показан на фибробластах человека	[Qin et al., 2012]

	серин/треонин-протеинкиназа			
	let-7	Способствует инаktivации генов, вовлеченных в поддержании плюрипотентности (<i>c-Myc</i> , <i>Lin28</i> , <i>Sall4</i>)		[Marson et al., 2008; Melton, Judson, Blueloch, 2010; Worringer et al., 2014]
	Кластер miR-34	Посттранскрипционно репрессирует <i>Nanog</i> , <i>Sox2</i> и <i>N-Myc</i> .		
	miR-145	Препятствует МЭП (за счет снижения экспрессии <i>E-cad</i>) и увеличению экспрессии генов плюрипотентности <i>Sox2</i> , <i>Klf4</i> и <i>c-Myc</i> . Активирует <i>let-7</i> .	Эффект показан на фибробластах человека	[BARTA et al., 2016]
	TESK1 LIMK2 серин/треонинкиназы	Фосфорилируют кофилин, что приводит к полимеризации актина. Препятствуют МЭП.		[Sakurai et al., 2014]
	Fbxw7 (E3-убиквитин лигаза)	Способствует деградации <i>C-MYC</i>		[Buckley et al., 2012]
	DDX5 (DEAD-Box Helicase 5) РНК-связывающий белок. Компонент комплекса Droscha	Через miR-125b посттранскрипционно репрессирует RYBP (RING1 and YY1-binding protein), компонент комплекса PRC1. Препятствует МЭП.	Роль RYBP в индукции плюрипотентности описана выше	[Li et al., 2017]
	956 генов. Среди них: 1. Система убиквитинирования: UBE2D3, UBE2E3, RNF40 2. Клеточная адгезия, подвижность ADAM7, ADAM21, ADAM29 3. Эндоцитоз и везикулярный транспорт	1. Способствуют деградации репрограммирующих факторов 2. Точный механизм неизвестен. Возможно препятствуют переключению класса интегринов ($\alpha 5$ на $\alpha 6$). 3. Способствуют активации сигнального пути TGF β , ингибирующего МЭП.	Определены с помощью полногеномный скрининга на основе РНК-интерференции	[Qin et al., 2014; Yang, Chang, Rana, 2014]

	ARSD, SLC17A5, DRAM1 4. Система дефосфорили рования PTPRJ, PTPRK, PTPN11 5. Регуляция структуры хроматина ATF7IP, ARID4A, CENPB, MED19 6. Регуляция транскрипции TTF1, TTF2, TMF1			
--	--	--	--	--

I.4 Модели репрограммирования

На сегодняшний день предложено несколько типов моделей репрограммирования соматических клеток к плюрипотентному состоянию. Это элитарные модели, а также детерминистическая и стохастическая модели (Рисунок 1).

I.4.1 Элитарные модели

I.4.1.1 Индуцированная элитарная модель

Согласно индуцированной элитарной модели, помимо репрограммирующих факторов на эффективность получения ИПСК влияют факторы, которые могут быть активированы или инактивированы путем интеграции конструкций с репрограммирующими факторами в геном. То есть только клетки со специфическим сайтом интеграции подвержены репрограммированию. Опровергает эту модель тот факт, что колонии ИПСК были получены без использования конструкций, встраивающихся в геном [Okita et al., 2008; Zhou, Freed, 2009]. Кроме того, вторичные репрограммируемые системы (описанные выше) показали, что даже те клетки, которые имеют одинаковые сайты интеграции факторов, сильно различаются по способности к репрограммированию [Stadtfield et al., 2010].

1.4.1.2 Элитарная модель

Элитарная модель предполагает, что полностью репрограммированные ИПСК могут быть получены только из особой субпопуляции клеток (Рисунок 1А). Возможными кандидатами на эту роль являются мезенхимальные стволовые клетки (МСК), всегда присутствующее в первичной культуре. В соответствующих условиях МСК способны дифференцироваться *in vitro* в разные типы клеток, включая эндотелиальные клетки [Oswald et al., 2004], кардиомиоциты [Makino et al., 1999], гепатоциты [Oyagi et al., 2006], et al. Однако, несмотря на это, МСК не способны вносить вклад в образование химерных животных, то есть не обладают плюрипотентностью как таковой. Вполне вероятно, что по причине менее дифференцированного статуса, МСК более эффективно подвергаются репрограммированию, в отличие от других, более дифференцированных клеток, которые либо совсем не вступают в процесс репрограммирования, либо проходят его лишь частично. На данный момент в большинстве работ по получению ИПСК используются гетерогенные клеточные популяции (первичные культуры), содержащие клетки разного происхождения и разной степени дифференцировки, в том числе и МСК. Контаминация первичных культур разными типами клеток неизбежна при выделении клеточного материала и может быть исключена только при использовании клеточного сортирования.

Ранее было показано, что степень дифференцировки репрограммируемых клеток значительно влияет на эффективность репрограммирования. Например, гемопоэтические стволовые клетки с гораздо большей эффективностью дают начало колониям ИПСК, по сравнению с терминально дифференцированными В-лимфоцитами [Eminli et al., 2009; Polo et al., 2010]. Доказательства справедливости элитарной модели, по крайней мере в случае репрограммирования фибробластов человека, были получены Вакао с коллегами. Авторам удалось показать, что в культуре фибробластов человека присутствуют особые мультипотентные стволовые клетки, названные Muse-клетки (Multilineage-differentiating stressenduring – мультипотентные стрессоустойчивые клетки) [Wakao et al., 2011]. Культивированные в суспензионной культуре Muse-клетки дифференцируются в кластеры, напоминающие эмбрионидные тельца, характерные для ЭСК, культивируемых при тех же условиях. Кластеры Muse клеток экспрессируют маркеры плюрипотентности, такие как *SSEA-1*, *Nanog*, *Oct4* и *Sox2*, и образуют производные всех трех зародышевых листков. Однако по сравнению с ЭСК и полностью репрограммированными ИПСК, Muse-клетки неспособны формировать тератомы и не дают вклад в формирование химер при инъекции в бластоцисту. Помимо этого, уровень экспрессии генов, связанных с регуляцией клеточного цикла, в Muse-клетках сравним с таковым в фибробластах и значительно

снижен по сравнению с плюрипотентными клетками. По оценкам исследователей в первичной культуре фибробластов содержится около 2-5% Muse клеток. Исследователи по поверхностным маркерам отсортировали популяцию фибробластов человека на Muse клетки и non-Muse клетки, и подвергли их репрограммированию путем сверхэкспрессии генов-индукторов плюрипотентности. В конечном итоге только Muse клетки дали начало колониям полностью репрограммированных ИПСК. Производные non-Muse клеток также образовывали колонии, но их морфология существенно отличалась от морфологии, характерной для ЭСК человека. К тому же, колонии, произошедшие от non-Muse клеток не экспрессировали многие гены-маркеры плюрипотентности, такие как *Sox2*, *Nanog*, *Rex1*. Согласно авторам, non-Muse колонии подвергаются аресту на ранних этапах репрограммирования и не способны развиваться в ИПСК, и только Muse клетки под действием репрограммирующих факторов подвергаются последовательным эпигенетическим изменениям, приводящим к установлению плюрипотентного состояния [Kitada, Wakao, Dezawa, 2012]. Однако даже в этом случае эффективность репрограммирования не превышает 0,03%. Возможно, в субпопуляции Muse клеток можно выделить еще несколько субпопуляций, которые с более высокой вероятностью способны дать начало колониям полностью репрограммированных ИПСК.

Как бы ни была привлекательна элитарная модель, в рамках нее невозможно объяснить тот факт, что колонии ИПСК были получены из терминально дифференцированных клеток [Aoi et al., 2008; Stadtfeld, Brennand, Hochedlinger, 2008]. Кроме того в некоторых экспериментах с использованием классических факторов Яманаки эффективность репрограммирования достигала 90% [Hanna et al., 2009] (описан ниже). Маловероятно, что ткани содержат такое большое количество соматических стволовых клеток и клеток-предшественников.

I.4.2 Стохастическая модель

Стохастическая модель предполагает, что каждая дифференцированная клетка, после обработки репрограммирующими факторами способна дать начало колониям ИПСК. При этом действия репрограммирующих факторов случайно приводят к активации генов-индукторов плюрипотентности, поэтому предсказать, когда произойдет и произойдет ли вообще полное репрограммирование клетки невозможно (Рисунок 1Б). Убедительное доказательство справедливости стохастической модели было получено группой исследователей под руководством Рудольфа Яниша [Hanna et al., 2009]. Исследователи выделили популяцию вторичных В-лимфоцитов, рассадили по одной клетке в индивидуальные лунки и запустили экспрессию репрограммирующих факторов. В

результате 93% клеточных клон образовали колонии ИПСК, хотя и требовали разное время для репрограммирования (от 2 до 18 недель). Полученные колонии ИПСК имели специфическую реорганизацию генов иммуноглобулинов, что однозначно определяет их происхождение от зрелых В-лимфоцитов, исключая возможность контаминации мезенхимальными стволовыми клетками. Авторы обнаружили, что скорость репрограммирования прямо коррелирует с количеством клеточных делений. Это вполне согласуется со стохастической моделью: если способность к репрограммированию возникает независимо в каждой клетке, то с увеличением числа клеток будет увеличиваться и число событий репрограммирования. Есть и альтернативная гипотеза, заключающаяся в том, что с делением клеток связаны некоторые лимитирующие события, необходимые для достижения плюрипотентности. Так, например, изменения структуры хроматина, происходящие в S-фазе, могут приводить к приобретению клеткой эпигенетических модификаций, способствующих установлению плюрипотентного состояния. Тогда чем больше делений пройдет клетка, тем больше вероятность того, что она будет репрограммирована [Hanna, Saha, Jaenisch, 2010].

Таким образом, было доказано, что любая терминально дифференцированная клетка может превратиться в плюрипотентную с достаточно высокой вероятностью. При этом ключевым фактором репрограммирования является количество клеточных делений, которые клетка претерпевает на пути к плюрипотентности.

I.4.3 Детерминистическая модель

Согласно детерминистической модели, изменения в экспрессии генов, приводящие к установлению плюрипотентности, происходят в строгом порядке, и колонии ИПСК могут быть получены с постоянной латентностью. Латентность в этом случае определяется абсолютным временем или количеством клеточных делений, необходимых для того, чтобы дифференцированная клетка дала плюрипотентного потомка (Рисунок 1Б).

Другие способы репрограммирования генома клеток – перенос ядра соматической клетки в энуклеированный ооцит (клонирование) или формирование гибридов между соматическими и эмбриональными стволовыми клетками являются детерминистическими. Активация *Oct4* происходит уже через два клеточных деления в случае клонирования и без деления в случае гибридизации и этом, не наблюдается случаев неполного репрограммирования, частого явления при получении ИПСК с помощью транскрипционных факторов [Boiani et al., 2002; Hasegawa et al., 2010; Jaenisch, Young, 2008; Yamanaka, Blau, 2010].

В 2013 году была опубликована еще одна статья из лаборатории Джейкоба Ханна, в которой было показано, что репрессия MBD3 (methyl-CpG binding domain protein 3), белка, принадлежащего хроматин-ремоделирующему комплексу NURD (nucleosome remodeling and deacetylation), приводит к максимальной эффективности репрограммирования (близкой к 100%) [Rais et al., 2013]. В этом случае каждая клетка образует плюрипотентного потомка, причем с фиксированной латентностью – за 7 дней, что подразумевает детерминистическую траекторию. Комплекс Mbd3/NURD и его роль в репрограммировании заслуживают отдельного внимания. Ранее в этом же 2013 году другим коллективом авторов также была показана положительная роль ингибирования Mbd3 при репрограммировании [Luo et al., 2013]. Однако в этом исследовании ученым не удалось добиться 100% эффективности репрограммирования. Отличает эти две работы лишь то, что в первой из них авторы использовали «вторичную» репрограммируемую систему (либо гомозиготную, либо гетерозиготную по делеции Mbd3), а во второй – репрессировали Mbd3 с помощью РНК-интерференции. В 2014 году коллективом авторов под руководством Жозе Сильвы была опубликована работа, результаты которой значительно отличаются, если не сказать опровергают выводы Джейкоба Ханна [Santos Dos et al., 2014]. Исследователи индуцировали плюрипотентность в нейральных столовых клетках, и обнаружили, что нокаутные по Mbd3 клетки репрограммируются значительно хуже. Согласно их выводам, экспрессия Mbd3 необходима на стадии инициации репрограммирования, а, следовательно, Mbd3 не препятствует индукции плюрипотентности, а наоборот, способствует ей. Интересно, что целый ряд работ с использованием полногеномных скринингов на основе малых интерферирующих РНК, не выявил Mbd3 среди генов, препятствующих репрограммированию. При том, что наличие других известных барьеров, не подвергающихся сомнению, таких как ген p53, было показано [Qin et al., 2014; Yang, Chang, Rana, 2014]. Исследователи из лаборатории Мариуса Вернига, еще одного признанного исследователя стволовых клеток, также работали на вторичных гетерозиготных Mbd3^{+/flox} клетках, предоставленных им Джейкобом Ханной. Согласно их данным, клетки с нарушенным Mbd3 действительно демонстрируют более равномерную траекторию достижения плюрипотентности по сравнению с другими системами. Тем не менее из работы не очевидно, что эффективность репрограммирования достигает 100% [Lujan et al., 2015]. Более года назад еще несколько групп исследователей из других лабораторий получили клеточные линии, с которыми работали Райз с соавторами, для того, чтобы воспроизвести результаты в своей лаборатории. Возможно, в скором времени, роль Mbd3 в репрограммировании станет более ясна. Данный пример еще раз наглядно показывает, что использование разных конструкций, условий культивирования и разные способы

получения трансгенных клеточных линий могут привести к совершенно противоположным результатам.

I.4.4 Синтез стохастической и детерминистической модели

Следующая модель, предложенная на основе анализа экспрессии генов в единичных клетках представляет собой синтез стохастической и детерминистической модели [Buganim et al., 2012] (Рисунок 1В). Так, на примере 48 исследованных генов было показано, что в первые 6 дней после индукции репрограммирования наблюдается наибольший разброс в уровне экспрессии даже между сестринскими клетками. На более поздних стадиях, после старта экспрессии генов *Sox2* и *Nanog*, вариабельность между клетками значительно уменьшается, что свидетельствует о синхронизации транскрипционных процессов. Авторы статьи сделали предположение, что OSKM факторы, взаимодействуя с доступными районами по всему геному вызывают стохастические, случайные изменения в экспрессии генов, в редких случаях приводя к активации генов *Sox2* или *Nanog*, после чего запускается экспрессия генов *Utf1*, *Esrrb*, *Dppa2* и *Lin28*; клетки становятся полностью плюрипотентными. Таким образом, данная модель соединяет в себе раннюю стохастическую фазу и следующую за ней позднюю детерминистическую.

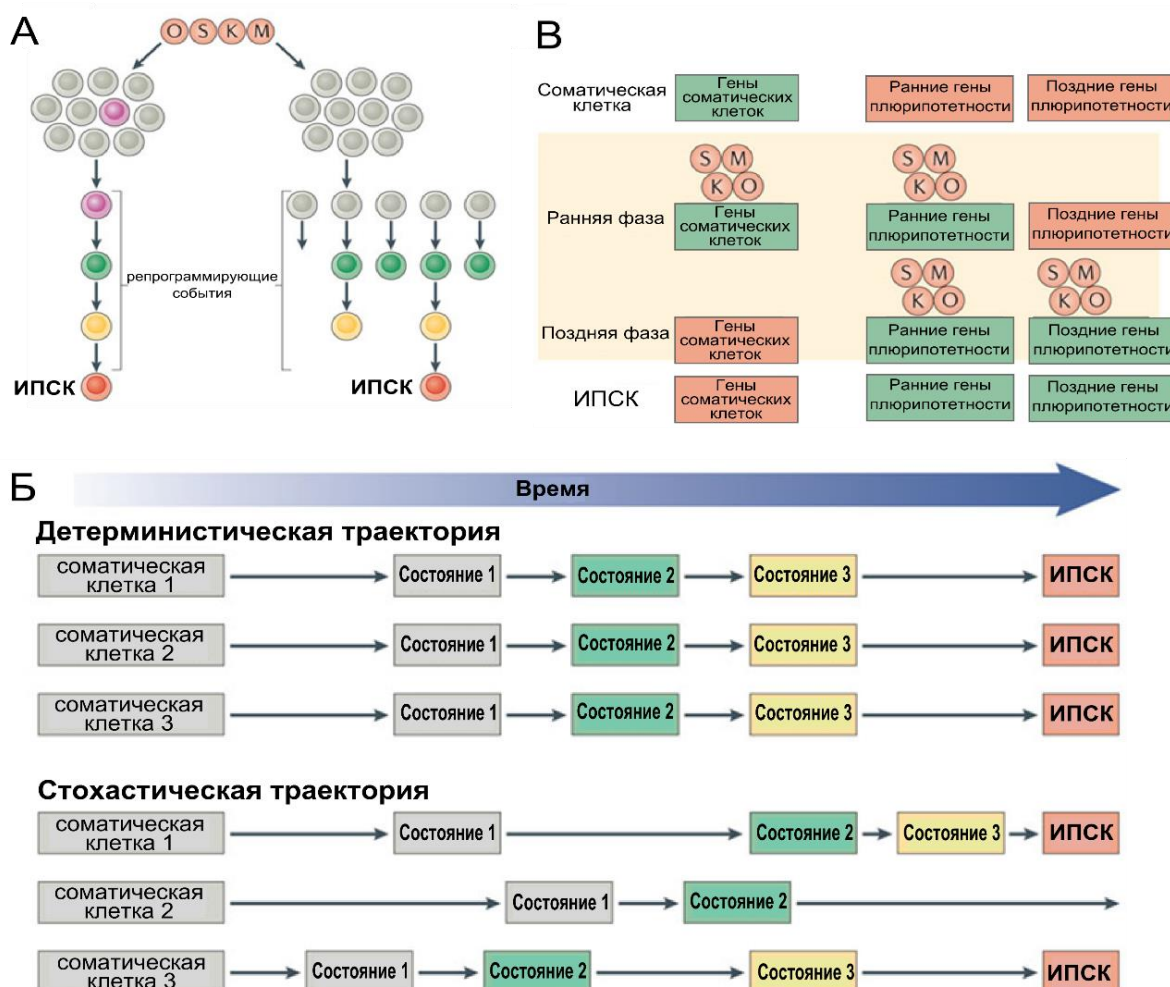


Рисунок 1. Модели репрограммирования из [Takahashi, Yamanaka, 2016].

А. Согласно элитарной модели, только «элитарные» клетки из исходной популяции (выделена розовым цветом) обладают способностью к репрограммированию (слева). Либо репрограммироваться может каждая клетка, но вероятность полного репрограммирования как случайного явления, очень мала (справа).

Б. Репрограммирование проходит либо по детерминистической траектории (сверху), либо по стохастической (снизу). Двигаясь по детерминистической траектории, клетки проходят одинаковую цепочку событий из одного состояния в другое с фиксированной латентностью. Тогда как стохастическая траектория подразумевает варьирование времени переходов для каждой клетки.

В. Репрограммирование как двухфазовый процесс. В соматических клетках активны гены, специфические для данного типа дифференцированных клеток (обозначены зеленым), тогда как гены плюрипотентности репрессированы (обозначены красным). На ранней стадии репрограммирования OSKM-факторы беспорядочно связываются с промоторными и энхансерными последовательностями как линейно-специфичных генов, так и ранних генов плюрипотентности (стохастическая фаза). На поздней стадии репрограммирования OSKM-факторы оказываются способны связываться с поздними генами плюрипотентности, в результате чего полностью активируется «коровая» сеть поддержания плюрипотентности (детерминистическая фаза).

Более подробные объяснения приведены в тексте. Аббревиатура OSKM означает репрограммирующие факторы Яманаки.

Подводя итог, подчеркнем, что пути репрограммирования, по которым проходят дифференцированные клетки, достигая состояния плюрипотентности, представляют собой все еще неизведанную территорию. Часть клеток выбирает более короткий путь, и репрограммируется за несколько дней. Тогда как большинство других клеток, идут долгими запутанными путями, требуя длительной экспрессии репрограммирующих факторов. Несмотря на шквал работ, направленных на изучение механизмов индукции плюрипотентности, все еще невозможно однозначно ответить, каким образом происходит распределение способности к репрограммированию в клеточных популяциях. Существуют работы, поддерживающие как стохастическую, так и элитарную модель репрограммирования. Сложность заключается также в том, чтобы отличить предрасположенность к репрограммированию, присущую клеткам еще до индукции плюрипотентности, от симметричного ответа сестринских клеток, запущенного уже после начала экспрессии репрограммирующих факторов.

Глава II. Материалы и методы

II.1 Материалы

II.1.1 Клеточные линии

В работе были использованы эмбриональные фибробласты мышей (ЭФМ) линии OG2, любезно предоставленные чл.-корр. РАН, проф. А.Н. Томилиным (Санкт-Петербург). В качестве системы для наработки вирусов были использованы клетки линии Phoenix (клетки эмбриональной почки человека). Все клетки культивировались при 37°C в атмосфере 5% углекислого газа.

II.1.2 Вычислительные ресурсы

Для запуска скриптов использовался вычислительный кластер Новосибирского Государственного Университета (<http://www.nusc.ru/>) под управлением ОС SUSE Linux Enterprise Server 11. Все скрипты были написаны на языке Python (версии 2.6 и 2.7) сотрудником лаборатории генетики развития В.С. Фишманом.

II.2 Методы

II.2.1 Молекулярно-биологические методы

II.2.1.1 Электрофорез в агарозном геле

Фрагменты ДНК разделяли в 0,8-3% агарозном геле, приготовленном на 1xTAE буфере (Трис-ацетат-ЭДТА) с 10 мкг/мл бромистого этидия. Электрофорез проводили в 1xTAE при напряжении электрического поля 5В/см. Буфер для нанесения пробы, содержал 50% глицерола, 0,1 ЭДТА (рН 8.0) и ~0.025% ксиленианола. Визуализацию фрагментов ДНК осуществляли в ультрафиолетовом свете с помощью ChemiDoc XRS System (BioRad).

II.2.1.2 Выделение ДНК из агарозного геля

Выделение ДНК из агарозного геля проводилось с использованием наборов (Биосилика, QIAGEN). Все действия проводились строго в соответствии с протоколом, рекомендуемым фирмой производителем.

II.2.1.3 Осаждение ДНК этанолом

К образцу добавляли 25 мкг/мл соосадителя (LPA фирмы GenElute), 0,1 объем 3М NaOAc и 2 объема 96% этанола. Центрифугировали 10 мин при 14500 g, отбирали супернатант и дважды промывали 70% этанолом. Пробирки высушивали в термостате

(37°C), затем растворяли ДНК в ddH₂O или TE-буфере (10 mM Трис-HCl pH 7,4; 1 mM ЭДТА-Na₂, pH 8,0).

II.2.1.4 Амплификация ДНК при помощи полимеразной цепной реакции (ПЦР)

Реакционная смесь объемом 25 мкл включала: буфер для Taq-полимеразы (65 mM Tris-HCl (pH 8,8); 16 mM (NH₄)₂SO₄; 0,01% Tween; 2,5 mM MgCl₂), 10pM каждого дезоксинуклеозидтрифосфата (дАТФ, дЦТФ, дГТФ, дТТФ), 10mM прямого и обратного праймеров, 0,5 ед. Taq-полимеразы, а также матрицу - 50-500 нг ДНК.

Условия ПЦР: денатурация геномной ДНК 95°C 2 мин, с последующей амплификацией (35 циклов): 95°C – 15 сек., 55°C – 15 сек., 72°C – 40 сек. После циклической амплификации - элонгация при 72°C в течение 3 мин. ПЦР проводили на автоматическом амплификаторе PTC 100 (MJ Research Inc.)

II.2.1.5 Определение первичной структуры ДНК

Определение первичной структуры ДНК проводили прямым секвенированием с использованием набора BigDye Terminator v 3.1 cycle Sequencing Kit. Реакционная смесь объемом 20 мкл включала: 3,5 мкл BigDye Terminator buffer; 1 мкл BigDye Terminator v3.1; 5 пмоль прямого или обратного праймера и 50-500 нг матрицы (плазмидная или геномная ДНК). Условия реакции включали 40 циклов: 10 сек -96°C, 10 сек - 50°C, 4 мин - 60°C, и заключительный этап - 3 мин - 94°C. Полученный продукт реакции очищали на колонках с сорбентом Sephadex G50. Определение нуклеотидной последовательности проводилось в Межинститутском центре секвенирования ДНК (ИХБФМ СО РАН).

II.2.1.6 Очистка ДНК с использованием Sephadex G50

К 1г сорбента добавляли 20 мл ddH₂O.Полученную суспензию инкубировали при комнатной температуре как минимум 3 часа. Заполняли колонку (Amersham Pharmacia Biotech, AutoSeq) 1 мл суспензии сорбента и центрифугировали 5 (900 g). Затем промывали сорбент 200 мкл ddH₂O, и центрифугировали 5 мин (900 g). Реакцию секвенирования ДНК переносили на сорбент и центрифугировали 5 мин (900 g). Собранную жидкость упаривали до сухого состояния центрифугированием при 40°C в течение 30 мин.

II.2.1.7 Выделение плазмидной ДНК

Выделение плазмидной ДНК проводилось с использованием наборов (QIAGEN MiniPrep, QIAGEN MidiPrep и Биосилика) строго по протоколам, прилагаемым фирмами-производителями.

II.2.1.8 Выделение геномной ДНК

Клетки центрифугировали 2 мин при 14000 об/мин. Клеточный осадок ресуспендировали в 400 мкл экстрагирующего буфера (10 mM Трис-HCl pH 8,0; 25mM ЭДТА-Na₂; 0,1M NaCl, 0,5% SDS), и 400 мкг протеиназы К (NEB). Инкубировали 2 часа при 65°C. Очистку ДНК проводили фенольной экстракцией. В водную фазу добавляли равный объем смеси фенол-хлороформа, тщательно перемешивали, центрифугировали 5 мин. 14000 об/мин. Раствор ДНК без интерфазы переносили в чистую пробирку, добавляли равный объем хлороформа, перемешивали и центрифугировали (5 мин. 14000 об/мин), раствор ДНК переносили в чистую пробирку. ДНК осаждали этанолом по описанной выше методике. Осажденную ДНК растворяли в буфере TE.

II.2.1.9 Гидролиз ДНК эндонуклеазами рестрикции

К реакционной смеси, содержащей однократный рестрикционный буфер и плазмидную ДНК, добавляли 10-100 единиц активности эндонуклеазы рестрикции. Инкубировали при оптимальной для фермента температуре. При необходимости инактивировали фермент прогреванием 20 мин при 65°C или 80°C. Для электрофоретического разделения фрагментов использовали 0.8-3% агарозный гель.

II.2.1.10 Лигазная реакция

Расчет соотношения количества вектора и вставки проводили по следующей формуле: $\frac{X * Y}{Z} * k = N$, где X-количество вектора в нг, Y – размер вставки в kb (тысячах нуклеотидов), Z – размер вектора в kb, k - молярное соотношение вставки к вектору, N-количество вставки в нг. Лигирование проводили в реакционной смеси в 1x лигазном буфере (50 mM Tris-HCl pH 7,5; 10 mM MgCl₂; 10 mM DTT; 1 mM ATP) (NEB) и 0,5 мкл лигазы фага T4 (NEB) с концентрацией 400 у/мкл в течении 12 часов при температуре 15°C.

II.2.1.11 Трансформация электрокомпетентных клеток E.coli

Кювету Bio-Rad с зазором 0,1 см, используемую для электропорации, заранее остужали во льду. Электрокомпетентные клетки оттаивали во льду в течение 5 мин. ДНК добавляли в оттаявшие клетки. Переносили клетки в охлажденную кювету и подвергали электропорации на приборе MicroPulser (Bio-Rad) со следующими параметрами: 1,8 kV, 25 μF, 200 Ом, зазор кюветы 0,1 см. Немедленно вымывали клетки из кюветы 1 мл среды SOC (0,05% NaCl; 2% бактотриптон; 0,5% дрожжевой экстракт; 2,5 mM KCl; 10 mM MgCl₂; 10 mM глюкоза; pH 7,5), предварительно прогретой до 37°C и переносили в микроцентрифужные пробирки на 1,5 мл. Пробирки качали на шейкере при 37°C в течение

45 мин (220 об/мин). Полученную клеточную суспензию наносили на чашки Петри с селективной питательной средой и растирали шпателем до сухого состояния. Чашки оставляли в термостате при 37°C на 12-18 часов.

II.2.2 Лентивирусные векторы

В работе были использованы следующие лентивирусные векторы:

pLeGO-TRE-OSKM. Содержит полицистронную кассету, кодирующую репрограммирующие факторы Яманаки (Oct4, Sox2, Klf4 и c-Мyc) в едином транскрипте. Промотор содержит специфическую последовательность TRE (tetracycline response element), с которой связывается белок-трансактиватор M2rtTA при взаимодействии с доксициклином. Только после этого запускается экспрессия репрограммирующих факторов. Лентивирусы, собранные на основе этого вектора, использовались для репрограммирования эмбриональных фибробластов мыши.

pLeGO-M2rtTA. Кодировал вышеупомянутый белок-трансактиватор под конститутивным промотором. Лентивирусы, собранные на основе этого вектора, использовались для репрограммирования фибробластов, а также в дополнительных экспериментах по оценке MOI (количество вирусных интеграций на клетку). Кроме того, именно этот вектор был выбран в качестве носителя ДНК-баркода.

pLeGO-TRE-GFP. Кодировал ген GFP под контролем доксициклин-индуцибельного промотора. Лентивирусы, собранные на основе этого вектора, использовались в дополнительных экспериментах по оценке MOI. Кроме того, этот вектор использовался для оценки функциональности баркодированного pLeGO-M2rtTA.

В качестве основы для всех векторов послужил вектор pLeGO-G2 (Addgene #25917), представляющий собой лентивирусный вектор третьего поколения. Все вышеописанные вектора были сконструированы автором в рамках дипломной работы. Их функциональность была продемонстрирована в экспериментах по репрограммированию эмбриональных фибробластов мыши.

II.2.3 Получение библиотеки баркодированных лентивирусных векторов

Отжиг баркодированных олигонуклеотидов: Синтез вырожденных олигонуклеотидов, содержащих 30 вырожденных позиций в которых могут располагаться любые основания (т.е. ДНК-баркод), проводился фирмой Биоссет (Новосибирск). Один из одноцепочечных олигонуклеотидов содержал ДНК-баркод (30 bp), фланкированный последовательностью M13F справа, и последовательностью, комплиментарной сайту посадки праймера M13R, слева.

Другой одноцепочный олигонуклеотид содержал последовательность праймера M13R (5'-catgtatgaattcGTTGTAAACGACGCCAGTnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnGGTCATA-GCTGTTTCCTGctcgagta-3'/5'-catgtaacatgtCAGGAAACAGCTATGACC-3'). Кроме того, на 5'-концы праймеров были добавлены некомплементарные последовательности, содержащие сайты связывания эндонуклеаз рестрикции EcoRI и PciI (выделены красным). В буфере для отжига (0,1M NaCl, TE до конечного объема 50 мкл) в эквимольном соотношении (100 пмоль) разводили одноцепочечные олигонуклеотиды и инкубировали 2 мин при температуре 95°C, затем смесь охлаждали до температуры 25°C в течение 30 мин. Реакция проводилась на автоматическом амплификаторе PTC-100 (MJ Research Inc).

Достройка дуплекса: К 50 мкл смеси, полученной на предыдущем шаге, добавляли 8 мкл 10 mM дезоксинуклеотидтрифосфатов, 20 мкл 10x буфера для фрагмента Кленова и 15 ед. самого фермента (Сибэнзим). Водой ddH₂O доводили объем до 200 мкл. Инкубировали 10 мин при температуре 37°C, затем инактивировали фермент при 65°C в течение 15 мин.

Очистка олигонуклеотидов из реакционной смеси: Очистку двухцепочечных олигонуклеотидов проводили с помощью набора для очистки ДНК из реакционных смесей (Biosilica) строго по прилагаемой инструкции.

Клонирование в вектор pLeGO-M2rtTA. Очищенные баркодированные олигонуклеотиды встраивали в плазмидный вектор pLeGO-M2rtTA по сайтам рестрикции EcoRI (SibEnzyme) и PciI (SibEnzyme) как описано выше в разделах: «гидролиз ДНК эндонуклеазами рестрикции» и «лигазная реакция».

Наработка плазмиды в бактериях и выделение плазмидной ДНК. Полученный вектор был трансформирован в Top10 электрокомпетентные клетки. Трансформированные клетки культивировались в жидкой среде LB с добавлением ампициллина (50мкг/мл) в течение двух часов. Затем около 0.0015% объема клеточной суспензии было нанесено на чашку Петри с селективной питательной средой – для подсчета количества трансформантов. Из остальной части клеточной суспензии была выделена плазмидная ДНК (см. «выделение плазмидной ДНК»). Подсчет колоний на чашке Петри на следующий день выявил 727 колоний. Таким образом, в общей клеточной суспензии предполагается наличие около 50 000 000 бактерий, несущих баркодированные плазмиды.

II.2.3 Амплификация и секвенирование ДНК-баркодов репрограммированных клеток

Отсортированные репрограммированные клетки были ресуспендированы в 5-10 мкл буфера для выделения ДНК (0.005% SDS и 400 мкг/мл протеиназа К) и инкубированы на

55°C в течение 3 часов [Goldenberger et al., 1995]. После инактивации фермента (в течение 10 минут при 95°C) весь объем реакционной смеси был использован для амплификации ДНК-баркодов с помощью ПЦР. В реакции была использована одна из самых точных ДНК-полимераз - high-fidelity DNA-polymerase (New England Biolabs). Праймеры, используемые в реакции, содержали адапторы, необходимые для секвенирования ПЦР-продуктов на платформе Ion Torrent. Кроме того, каждый из forward праймеров, содержал уникальную последовательность из 10 нуклеотидов, персональный идентификатор (Таблица 2).

Название и последовательность праймера 5'-3'
Праймер Р CCTCTCTATGGGCAGTCGGTGATATGCTCCCCGGGTAACТАAG
Праймер А2 CCATCTCATCCCTGCGTGTCTCCGACTCAGTAAGGAGAACGATGTGCAGGAAACAGCTATGACC
Праймер А3 CCATCTCATCCCTGCGTGTCTCCGACTCAGAAGAGGATTCGATGTGCAGGAAACAGCTATGACC
Праймер А4 CCATCTCATCCCTGCGTGTCTCCGACTCAGTACCAAGATTCGATGTGCAGGAAACAGCTATGACC
Праймер А5 CCATCTCATCCCTGCGTGTCTCCGACTCAGCAGAAGGAACGATGTGCAGGAAACAGCTATGACC
Праймер А6 CCATCTCATCCCTGCGTGTCTCCGACTCAGCTGCAAGTTCGATGTGCAGGAAACAGCTATGACC
Праймер А7 CCATCTCATCCCTGCGTGTCTCCGACTCAGGTTACGGCTGATGTGCAGGAAACAGCTATGACC
Праймер А8 CCATCTCATCCCTGCGTGTCTCCGACTCAGCGATGGACTTGATGTGCAGGAAACAGCTATGACC
Праймер А9 CCATCTCATCCCTGCGTGTCTCCGACTCAGGCAGGTGTCTGATGTGCAGGAAACAGCTATGACC

Таблица 2. Список олигонуклеотидов, использованных для амплификации ДНК-баркодов репрограммированных клеток. Красным выделены последовательности персональных идентификаторов. В первом, втором и третьем экспериментах были использованы олигонуклеотиды А2-А5. В контрольном эксперименте – олигонуклеотиды А6-А9.

Амплификация ДНК-баркодов каждого образца осуществлялась с помощью праймера Р и одного из праймеров А. Секвенирование ПЦР-продуктов всех образцов проводилось в одной реакции, и принадлежность сиквенса определялось по уникальной последовательности в праймере А. ПЦР-продукты правильного размера были выделены из агарозного геля с помощью чипа PE Caliper LabChip XT with DNA 300. Анализ качества и количества ПЦР-продуктов был проведен на анализаторе Agilent Bioanalyzer 2100.

Секвенирование образцов было проведено на платформе Ion Torrent (PGM) в секторе геномных исследований ИЦИГ СО РАН согласно инструкциям.

II.3 Компьютерный анализ

II.3.1 Биоинформатическая обработка результатов секвенирования

Биоинформатический анализ проводился с помощью алгоритмов на языке Python. Риды были распределены по разным библиотекам согласно последовательности персонального идентификатора. С помощью программы cutadapt 1.3 были удалены константные фрагменты, фланкирующие ДНК-баркод. Риды, не содержащие таких фрагментов, были исключены из дальнейшего анализа. Качество ридов было проверено с помощью программы FastQC v0.10.1. Риды с низким качества прочитки (где 20% позиций имеют значение Phred ниже 20) были удалены. Кроме того, были удалены риды, длина которых более 33 и менее 26 нуклеотидов. Оставшиеся риды были переведены в FASTA формат и кластеризованы по гомологии с помощью программы DNAClust с пороговым значением гомологии 80%. Кластеры ДНК-баркодов, представленные только одним ридом, также были исключены из анализа. Сравнение ДНК-баркодов между библиотеками было проведено с помощью скрипта, написанного В.С. Фишманом.

II.3.2 Компьютерное моделирование и статистический анализ

Мы разработали компьютерную модель, симулирующую все стадии эксперимента и позволяющую учесть ряд экспериментальных параметров. Первый шаг симулирует рассадку клеток и трансдукцию их вирусами. Каждая клетка в модели представлена как независимый объект. Эффективное число клеток N_0 считается как:

$$N_0 = N_{seeded} * (1 - e^{-m})$$

где N_{seeded} это число клеток, посаженных на чашку в начале эксперимента, а m – MOI, или среднее количество вирусных интеграций на клетку. Этот параметр рассчитывается только для баркодированного вируса (LeGO-M2rtTA) из дополнительных экспериментов (см. результаты и методы ниже «Репрограммирование эмбриональных фибробластов»). Каждой клетке из N_0 присваивается ДНК-баркод. ДНК-баркод представляет собой случайное целое число, выбранное из количества вариантов, равного разнообразию библиотеки баркодов. В нашем случае это разнообразие составляет от 50 000 до 8 000 000 уникальных вариантов (см. результаты). Следующий шаг моделирования симулирует период культивирования клеток в отсутствие доксициклина и последующую рассадку клеток на 4 чашки. На этом этапе каждая N_0 клетка образует от 0 до 8 потомков. При этом, все потомки несут тот же

баркод, что и родительская клетка. Количество потомков определяется для каждой клетки случайно, исходя из анализа делений клеток с помощью прижизненной микроскопии. Затем клетки рассаживаются на 4 разные чашки – этот этап реализуется путем случайного присвоения каждому потомку номера в диапазоне от 1 до 4. Затем клетки на каждой чашке подвергаются случайному удалению с вероятностью от 0 до 0.9. Этот этап отражает потерю баркодов на разных стадиях культивирования клеток и приготовления библиотеки. Поскольку параметр потери материала не может быть вычислен однозначно, в модели он варьируется от 10 до 90%. Последний этап моделирования симулирует стадию репрограммирования. Из экспериментальных данных нам известно количество уникальных ДНК-баркодов, присутствующих на каждой чашке. Следовательно, мы можем рассчитать среднее количество (μ) и дисперсию (σ^2) этого параметра, предполагая нормальность распределения. На этом этапе моделирования создается четыре случайных числа $BA_1 \dots BA_4$ согласно нормальному распределению с параметрами μ и σ^2 . Затем с каждой чашки выбирается такое количество баркодов, чтобы, с одной стороны, эти числа соответствовали числам $BA_1 \dots BA_4$, а с другой стороны, чтобы общее число выбранных баркодов ($BA_1 + \dots + BA_4$) было приблизительно равно общему числу баркодов на всех чашках в эксперименте. Клетки с выбранными (т.е. репрограммированными) баркодами делятся на подгруппы так, чтобы сестринские клетки оказались в одной подгруппе. Этот шаг необходим, поскольку клетки могут нести одинаковый баркод и не быть при этом сестринскими (такие клетки будут отнесены в разные подгруппы). Затем, только одной клетке из каждой подгруппы присваивается возможность быть репрограммированной. В этом случае вероятность передать способность к репрограммированию по наследству или коэффициент наследуемости равен 0. Мы можем варьировать этот параметр от 0 до 1. В этом случае каждой из сестринских клеток присваивается заданная вероятность быть репрограммированной. Результатом моделирования является количество пересечений или общих баркодов между выбранными баркодами на разных уровнях наследуемости. Пересечения при этом могут быть двойными (между двумя любыми группами), тройными (между тремя группами), и наконец, четверными (пересечения между всеми четырьмя группами).

Алгоритм моделирования был написан на языке Python сотрудником лаборатории генетики развития В.С. Фишманом. Все вычисления были проведены на вычислительном кластере Новосибирского Государственного Университета (<http://www.nusc.ru/>). Мы проводили как минимум 100 запусков моделирования на каждое значение коэффициента наследуемости и потери материала (значения варьировались от 0 до 1). Результаты симуляции сравнивались с экспериментальными результатами с помощью

непараметрического дисперсионного анализа (критерий Краскела-Уоллиса) при уровне значимости $p < 0.05$. Все статистические тесты были выполнены в программе GraphPad Prism 6.00.

II.4 Методы работы с клеточными культурами

II.4.1 Культуральные среды

Среда для культивирования эмбриональных фибробластов мыши и клеток линии Phoenix:

88% среда DMEM (Dulbecco's Modified Eagle Medium (Sigma, Thermo Fisher Scientific))

10% FBS (фетальная бычья сыворотка) (Thermo Fisher Scientific)

1 mM L-глутамин (Invitrogen)

50 ед/мл пенициллин (Sigma)

50 ед/мл стрептомицин (Sigma)

Среда для культивирования репрограммирующихся ЭФМ:

80% среда DMEM (Sigma- Aldrich, Thermo Fisher Scientific)

7,5% FBS (бычья эмбриональная сыворотка) (Thermo Fisher Scientific)

7,5% KSR (синтетический заменитель сыворотки) (Thermo Fisher Scientific)

1 mM L-глутамин (Invitrogen)

1 mM заменимых аминокислот (Invitrogen)

0,1 mM β -меркаптоэтанол (Invitrogen)

50 ед/мл пенициллин (Sigma)

50 ед/мл стрептомицин (Sigma)

1000 ед/мл LIF (PolyGene)

50 мкг/мл аскорбиновая кислота (Wako)

3 мкМ ингибитор GSK3b (Sigma)

2 мкг/мл DOX (доксидиклин) (Sigma)

II.4.2 Получение культуры первичных эмбриональных фибробластов

Для получения эмбриональных фибробластов использовали 12-дневных эмбрионов мышей линий C57BL и OG2. Эмбрионы извлекали из матки и стерильно переносили в среду для культивирования эмбриональных фибробластов мыши. После двухминутной инкубации эмбрионы переносили в чашку Петри в каплю раствора трипсина (0,25%). После удаления головы и внутренних органов ткани эмбриона тщательно измельчали и переносили в пробирку, содержащую 3 мл трипсина. После инкубации в трипсине при температуре 37°C в течение 30 мин, трипсин инактивировали добавлением 8 мл среды для

культивирования. Центрифугировали в течение 5 мин при 900g (LMC-3000 Biosan). Супернатант сливали и клеточный осадок ресуспендировали в среде для фибробластов. Затем переносили для выращивания в культуральные матрасы площадью 25 см² (Corning).

II.4.3 Культивирование фибробластов

Фибробласты пересаживали в новую емкость после достижения ими монослоя. При пересеве из лунки отбирали всю ростовую среду и промывали клетки фосфатно-солевым буфером (PBS) (Медиген). Затем добавляли трипсин (раствор 0,25%) и инкубировали 2-5 минут в термостате при 37°C. Затем инактивировали трипсин культуральной средой с сывороткой, тщательно ресуспендировали клетки, центрифугировали 5 мин при 900g, сливали супернатант. Клеточный осадок ресуспендировали в среде для фибробластов и переносили в новую культуральную емкость.

II.4.4 Размораживание клеток

Ампулы с клетками размораживали на водяной бане при 37°C, затем помещали суспензию клеток в пробирку (5 мл) со средой DMEM+10%FBS, центрифугировали 5 мин при 900g, сливали супернатант, ресуспендировали клетки в ростовой среде и переносили в емкость для культивирования.

II.4.5 Замораживание клеток

Клетки замораживали, когда их плотность достигала 80-85%. Клетки снимали трипсином как при пересадке. После центрифугирования клеточный осадок ресуспендировали в среде для заморозки (50% FBS, 10% DMSO, 40% DMEM). Полученную суспензию по 1 мл переносили в криопробирки и помещали в криоконтейнер (Mr. Frosty) на -70°C. Через сутки криопробирки переносили в дьюары с жидким азотом для длительного хранения.

II.4.6 Получение вирусных частиц

День 0, рассев: клетки линии Phoenix рассевали на желатинизированные культуральные чашки (примерно по 7×10^5 клеток/см²).

День I, трансфекция: Клетки трансфицировали с помощью реагента Lipofectamine 2000 (Invitrogen). Состав смеси ДНК: pRSV/Rev, pMDLg/pRRE, pCMV-VSVG (содержат компоненты, необходимые для упаковки и интеграции вирусной последовательности) и векторная плазмида в соотношении 5:10:2:10. В качестве векторных использовались следующие плазмиды: pLEGO-TRE-OKSM, pLEGO-M2rtTA или pLEGO-TRE-GFP. Трансфекцию проводили в полном соответствии с протоколом производителя.

День 3, 4, сбор вирусов: Раз в сутки собирали с клеток среду, и меняли её на новую, содержащую DMEM (GIBCO), 10% FBS, инактивированную путем нагревания на 56 °C в течение 30 мин, 1% заменимых аминокислот (Invitrogen), 1% глутамина (Invitrogen). Вирусный супернатант фильтровали через 0,22 мкм фильтр (Millipore) и либо сразу использовали для заражения клеток, либо аликвотили и замораживали на -70°C.

II.4.7 Сортировка Thy1.2-позитивных клеток с помощью проточной цитофлуориметрии

ЭФМ 1-го пассажа снимали трипсином как при пересадке, центрифугировали 900g в течении 5 мин. Осадок промывали 1 раз в PBS. Клетки ресуспендировали в ростовой среде (100 мкл на 10^6 клеток) и добавляли антитела Thy1.2 (FITC) Biolegend в количестве 0.5 мкл на 10^6 клеток. Инкубировали клетки в темноте во льду в течение 30 мин, дважды промывали PBS. Затем ресуспендировали в ростовой среде и сортировали по 250000 клеток на 10 см² лунку с помощью проточного цитофлуориметра FACS Aria (BD Biosciences) в ЦКП микроскопического анализа биологических объектов СО РАН.

II.4.8 Репрограммирование эмбриональных фибробластов мыши к плюрипотентному состоянию

ЭФМ линии OG2 первого пассажа высаживали на 10 см² чашки с плотностью 20×10^3 на 1 см². Клетки инкубировали с супернатантом, содержащим избыток вирусов LEGO-TRE-OKSM и баркодированный вирус LEGO-M2rtTA. При этом, во избежание множественного баркодирования клеток, последний вирус был взят в недостатке. Среднее количество интеграций LEGO-M2rtTA, приходящееся на одну клетку (MOI) вычислялось из дополнительных экспериментов. Дополнительные эксперименты проводились параллельно экспериментам по репрограммированию и в аналогичных условиях, но вместо вируса с репрограммирующими факторами был использован вирус LeGO-TRE-GFP. Долю GFP-позитивных клеток определяли на проточном цитофлуориметре FACS Aria (BD Biosciences) в ЦКП микроскопического анализа биологических объектов СО РАН. Значение MOI определялось по количеству GFP-позитивных клеток согласно распределению Пуассона. Для улучшения эффективности трансдукции к вирусному супернатанту добавляли полибрен (5мкг/ мл). На следующий день меняли среду на среду для фибробластов. После удвоения клеточной популяции, клетки пересаживали на обработанные митомицином-С мышинные эмбриональные фибробласты (фидерные клетки). И уже с этого момента клетки культивировались в среде для ИПСК с добавлением доксициклина (см. состав культуральных сред). Смену среды проводили каждый день.

II.4.9 Определение времени удвоения клеточной популяции и количества потомков трансдуцированных клеток

Дополнительно, для определения времени удвоения клеточной популяции, ЭФМ линии OG2 были рассажены с той же плотностью и трансдуцированы при тех же условиях, что и в эксперименте по репрограммированию. Все клетки были из одной разморозки, а эксперименты проводились параллельно. Подсчет клеток проводился каждые 12 часов (на протяжении 48 часов) с помощью камеры Горяева. Каждая временная точка была подсчитана трижды. Кроме того, клетки в эксперименте по репрограммированию после трансдукции культивировались в приборе CellIQ в ЦКП Клеточных технологий ИЦиГ СО РАН в течение 24 -48 часов с детальной съемкой. На каждой лунке было выбрано 36 полей; частота фотосъемки каждого поля составила один снимок в 10 минут. Количество потомков клеток было установлено с помощью анализа фотографий (вручную).

II.4.10 Получение контрольных клеток

Клетки линии Phoenix были трансдуцированы двумя типами лентивирусов: LeGO-TRE-GFP, несущим ген GFP под контролем доксициклин-индуцибельного промотора и LeGO-M2rtTA, несущим ДНК-баркод. При этом, вирус LeGO-M2rtTA был взят в недостатке. Это необходимо для того, чтобы большинство клеток, которые получают этот вирус, получили только одну его копию. Трансдуцированные клетки культивировались в присутствии доксициклина (2 мг/мл). Затем отдельные GFP-позитивные клетки были субклонированы. Из нескольких клонов были получены стабильные клеточные линии. Из клеток двух линий, выбранных в качестве контрольных, была выделена ДНК, с последующей амплификацией ДНК-баркодов и секвенированием ПЦР-продуктов. Последовательность праймеров, используемых для амплификации: 5'-GTGGCCTGGAGAAACAGCTA-3' и 5'-CCACATAGCGTAAAAGGAGCA-3'. Секвенирование подтвердило присутствие только одного варианта ДНК-баркода на линию. Последовательности контрольных ДНК-баркодов были следующие:

Линия_1 - 5'-TGCTGACCGCAGGTACGACGCCGAAGGATG-3'

Линия_2 - 5'-GGACGCGCCTACTACTTCGCGAGCATCCTG-3'

Контрольные клетки были добавлены в каждый образец репрограммированных клеток (в каждом эксперименте было 4 образца клеток (см. схему эксперимента в главе результаты)). В первом эксперименте в каждый образец было добавлено 100 клеток Линии_1 и 400 клеток Линии_2. Для второго и третьего эксперимента числа равнялись 20/10 клеток для Линии_1 и 80/50 клеток для Линии_2 соответственно.

Глава III. Результаты

III.1 Схема эксперимента

В качестве источника эмбриональных фибробластов мы выбрали трансгенную линию мышей OG2, клетки которой несут ген *Gfp* под контролем промотора гена *Oct4* [Szabó et al., 2002]. При активации указанного промотора, запускается экспрессия *Gfp*, что позволяет легко отобрать популяцию репрограммированных клеток на клеточном сортере. Схема эксперимента представлена на Рисунке 2.

Эмбриональные фибробласты мыши были ко-трансдуцированы двумя лентивирусными конструкциями, одна из которых содержит репрограммирующие факторы Яманаки (*Oct4*, *Sox2*, *Klf4* и *c-Myc*) под контролем доксициклин-индуцибельного промотора, а другая – индивидуальный ДНК-баркод и белок-трансактиватор M2rtTA (reverse tetracycline transactivator). Именно этот трансактиватор при взаимодействии с доксициклином запускает экспрессию репрограммирующих факторов. Это позволяет нам контролировать экспрессию OSKM простым добавлением доксициклина в культуральную среду. Клетки культивировались около двух суток до удвоения клеточной популяции, а затем были рассажены на 4 чашки, для того чтобы потомки одной клетки могли попасть в разные чашки (для двух потомков одной клетки вероятность попасть на разные чашки равна 75%). Только после этого была запущена экспрессия репрограммирующих факторов, путем добавления в культуральную среду доксициклина. Таким образом, все клетки вступили в процесс репрограммирования одновременно. После 7 дней индукции GFP-позитивные клетки были отобраны, а их ДНК-баркоды амплифицированы и секвенированы. После биоинформационной обработки мы получили состав ДНК-баркодов на каждой чашке. Присутствие одинаковых ДНК-баркодов на разных чашках свидетельствует о том, что потомки одной клетки попали на разные чашки и успешно репрограммировались. Поэтому, зная количество пересечений, мы можем определить долю сестринских клеток, синхронно репрограммировавшихся в ходе эксперимента. Если способность к репрограммированию предетерминирована и может передаваться по наследству, доля баркодов, найденных на разных чашках, будет высока. В другом случае, если успешное репрограммирование является простой случайностью (стохастическим событием), доля общих баркодов будет незначительна.

Схемы получения лентивирусных векторов для индукции плюрипотентности, а также баркодированных лентивирусных векторов описаны в материалах и методах.

1. ТРАНСДУКЦИЯ

Контрансдукция ЭФМ доксициклин-индуцибельными лентивирусами, кодирующими репрограммирующие факторы и баркодированными лентивирусами. На этом этапе ДНК-баркод интегрируется в геномную ДНК. Экспрессии репрограммирующих факторов еще нет.

2. УДВОЕНИЕ КЛЕТОЧНОЙ ПОПУЛЯЦИИ

Трансдуцированные клетки культивируются до тех пор, пока не достигнут удвоения клеточной популяции (24-48 часов).

3. РАССАДКА

Клетки в равном количестве рассаживаются на 4 чашки с фидерными клетками. На этом этапе происходит распределение дочерних клеток на разные чашки.

4. СТАРТ РЕПРОГРАММИРОВАНИЯ

Запуск экспрессии репрограммирующих факторов путем добавления доксициклина в культуральную среду. Клетки культивируются в присутствии доксициклина в течение недели.

5. СОРТИРОВКА

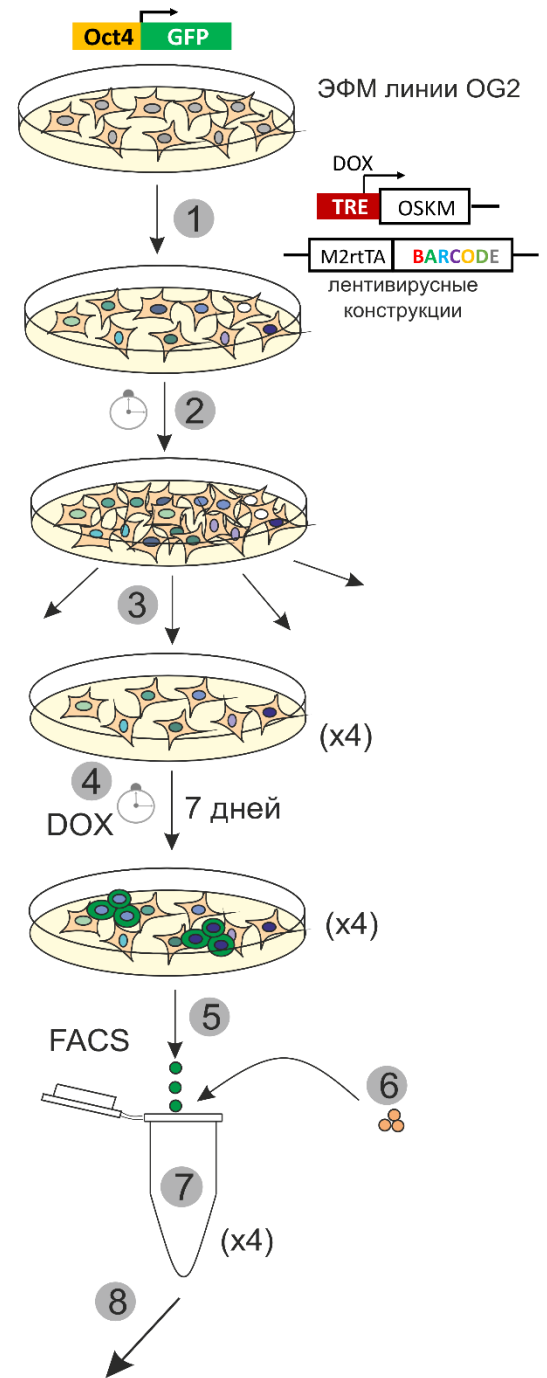
Репрограммированные клетки с каждой чашки отбираются на клеточном сортере (сортировка по GFP).

6. ДОБАВЛЕНИЕ КОНТРОЛЬНЫХ КЛЕТОК

Определенное количество контрольных клеток с известными ДНК-баркодами добавляется в каждый образец с сортированными клетками.

7. АМПЛИФИКАЦИЯ ДНК-БАРКОДОВ

Клетки лизируются, и все ДНК-баркоды амплифицируются при помощи ПЦР.

**8. Секвенирование ДНК-баркодов на платформе Ion Torrent****Биоинформационная обработка ридов**

Распределение ридов по принадлежности одной из 4-ех библиотек

Сравнение пересечений ДНК-баркодов между библиотеками

Рисунок 2. Схема эксперимента.

[illegible]

Как упоминалось выше, вирус на основе LeGO-TRE-OSKM был взят в избытке, чтобы максимизировать количество клеток, получивших интеграцию репрограммирующих факторов. Тогда как второй вирус был взят в недостатке во избежание множественного баркодирования одной клетки. Мы культивировали инфицированные клетки в течение двух суток, затем рассадили их в равной пропорции на 4 чашки с фидерными клетками и добавили в культуральную среду доксициклин. Уже на третий день культивирования мы наблюдали появление GFP-позитивных клеток, которые к седьмому дню сформировали колонии двух типов: небольшие округлые колонии с плотно контактирующими клетками, а также крупные колонии с нечеткими границами (Рисунок 4А).

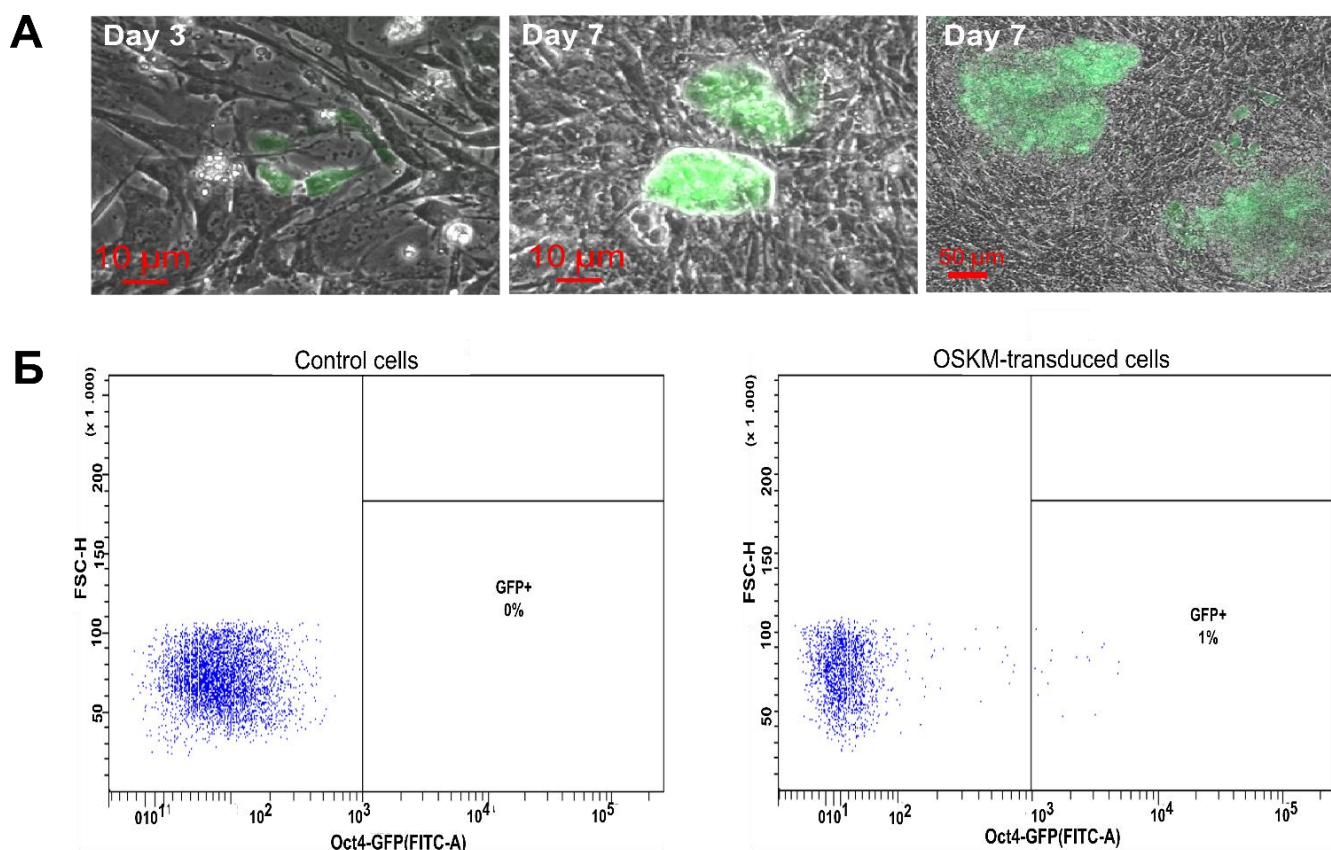


Рисунок 4. Получение GFP-положительных репрограммированных клеток, маркированных ДНК-баркодами.

(А). Последовательная активация экспрессии GFP в ходе репрограммирования клеток. Нумерация дней с начала добавления доксициклина в культуральную среду.

(Б). Анализ количества GFP-положительных клеток после 7 дней экспрессии репрограммирующих факторов, оцененное методом проточной цитофлуориметрии.

GFP-положительные клетки были отобраны на клеточном сортере (Рисунок 4Б), а их ДНК-баркоды амплифицированы и просеквенированы на платформе Ion Torrent в центре секвенирования ИЦИГ СО РАН. После биоинформатической в нашем распоряжении остались только 30-нуклеотидные последовательности ДНК-баркодов обработки (см. материалы и методы). Суммарная библиотека содержала 160000 качественно прочтенных последовательностей, что свидетельствует о невысокой глубине секвенирования (Приложение 1). Тем не менее, мы кластеризовали оставшиеся последовательности по гомологии, и получили 1675 уникальных последовательностей ДНК-баркодов. По чашкам эти баркоды были распределены следующим образом: 610, 596, 325 и 509 вариантов баркодов в 1-ой, 2-ой, 3-ей и 4-ой чашке соответственно. Из этих данных видно, что суммарное количество баркодов, найденных в каждой чашке ($610 + 596 + 325 + 509 = 2040$) превышает количество баркодов в общей библиотеке до распределения по четырем библиотекам, соответствующим четырем чашкам (1675). Это означает, что некоторые варианты ДНК-баркодов присутствуют более чем на одной чашке. Доля ДНК-баркодов,

общих и уникальных для каждой чашки, отражены на Рисунке 5А. Сравнив ДНК-баркоды каждой из четырех библиотек между собой, мы обнаружили большое количество пересечений между ними (Рисунок 5Б). Отсюда можно предположить следующую цепочку событий: материнская клетка получила ДНК-баркод, поделилась, при рассадке дочерние клетки попали в разные чашки, и каждая из них успешно репрограммировалась. Именно тогда на разных чашках мы обнаружим одинаковый ДНК-баркод. Таким образом, количество пересечений ДНК-баркодов с разных чашек отражает количество событий, когда обе дочерние клетки репрограммировались.

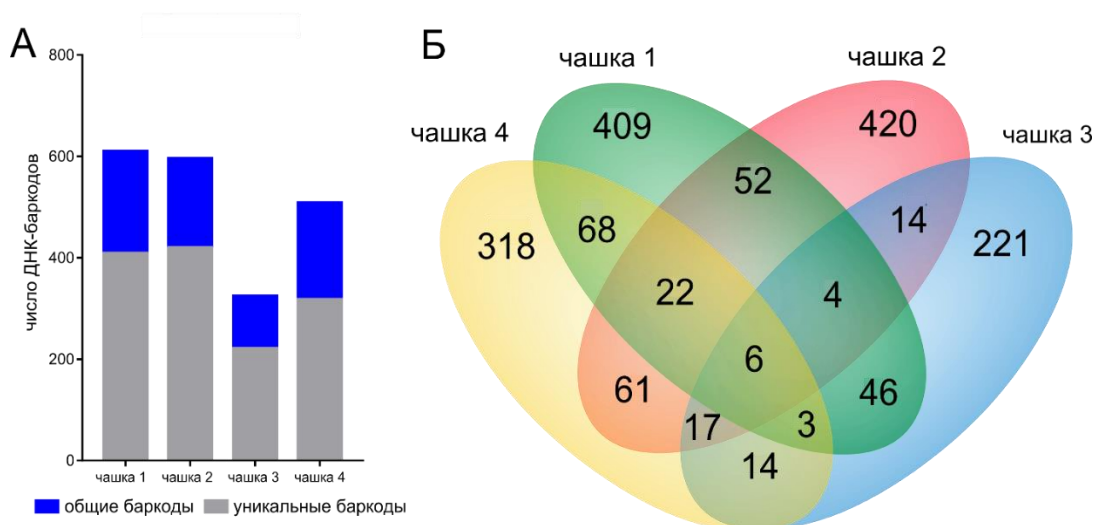


Рисунок 5. Сравнение ДНК-баркодов успешно репрограммированных клеток, собранных с четырех чашек в пилотном эксперименте.

(А). Столбики представляют собой количество ДНК-баркодов, полученных с каждой чашки. Доля уникальных вариантов ДНК-баркодов, присутствующих только на одной чашке, обозначена серым блоком. Доля баркодов, общих для двух и более чашек отмечена синим блоком.

(Б). Диаграмма Венна, отражающая число общих ДНК-баркодов для четырех библиотек/чашек.

Для подсчета количества наблюдаемых в эксперименте двойных пересечений ДНК-баркодов, мы использовали следующую формулу:

$$T_{\text{observed}} = \sum_{i=2..4} ((\sum P_i) * (i - 1) * (-1)^i) \quad (1)$$

где P_i отражает пересечение кластеров между двумя разными чашками. В нашем эксперименте $T_{\text{observed}} = 215$. Если же сестринские клетки репрограммировались за время эксперимента случайно, то ожидаемое количество пересечений кластеров T_{expected} будет рассчитываться по следующей формуле:

$$T_{\text{expected}} = N * C * E^2, \quad (2)$$

где N – количество трансдуцированных клеток, C – вероятность того, что дочерние клетки попадут при рассадке в разные чашки, и E – эффективность репрограммирования. В

пилотном эксперименте количество трансдуцированных клеток равнялось 170 000. Поскольку в формуле (1) рассчитывается суммарное количество *двойных* пересечений, то и параметр C в формуле (2) определяет вероятность распределения *двух* сестринских клеток на две разные чашки. Для четырех чашек эта вероятность равна 0.75. Эффективность репрограммирования (E) была оценена как отношение общего количества уникальных ДНК-баркодов к числу трансдуцированных клеток и составила примерно 0.0001. Вычисленное таким образом T_{expected} равняется 13, что намного меньше наблюдаемого пересечения T_{observed} (215). Этот результат говорит в пользу того, что способность к репрограммированию может иметь наследственную компоненту, и не является случайным процессом.

Таким образом, мы показали, что метод ДНК-баркодирования может быть использован для выявления наследственной компоненты способности к репрограммированию. Однако для того, чтобы дать более точную оценку, необходимо учесть ряд дополнительных параметров, непосредственно связанных с самой методикой баркодирования.

III.3 Параметры, определяющие эффективность баркодирования клеток

Эффективность метода ДНК баркодирования для изучения судьбы отдельных клеток в клеточных поколениях была показана в многочисленных экспериментах [Lu et al., 2011], [Cheung et al., 2013], [Nguyen et al., 2014]. Напомним, что основной принцип этого метода заключается во внесении вырожденной метки в геном клеток с помощью вирусов. Эта метка, последовательность из случайных нуклеотидов, наследуется в ходе поколений и, таким образом, по ней можно установить клональную принадлежность клеток. Ниже в тексте обсуждены некоторые параметры, оказывающие влияние на эффективность баркодирования клеточной популяции.

III.3.1 Разнообразие библиотеки баркодированных вирусов

Как упоминалось ранее, один из наиболее важных параметров – это размер или разнообразие библиотеки баркодированных лентивирусов. Под разнообразием подразумевается количество уникальных вариантов ДНК-баркодов. Согласно исчерпывающему обзору [Naik, Schumacher, Peri, 2014], посвященному баркодированию, количество клеток в эксперименте должно составлять не более 10% от разнообразия ДНК-баркодов. В противном случае, клетки с разным происхождением и, как правило, с разной судьбой, имеют большую вероятность получить одинаковый ДНК-баркод, и будут восприниматься исследователем как одна клетка (Рисунок 6А). Это осложняет

интерпретацию данных, поскольку в эксперименте мы будем иметь завышенный уровень пересечений ДНК-баркодов с разных чашек. Именно поэтому так важно оценить реальное разнообразие библиотеки ДНК-баркодов.

Теоретически, разнообразие библиотеки определено длиной ДНК-баркода. Для 30-нуклеотидного баркода это более чем 10^{17} уникальных вариантов. Однако, получение лентивирусной библиотеки включает клонирование баркода в векторную плазмиду, трансформацию бактериальных клеток, а также трансфекцию клеток млекопитающих (сборка вирусных частиц). Безусловно, эти стадии являются лимитирующими/неэффективными и не позволяют сохранить все изначальное разнообразие ДНК-баркодов.

Как правило, для определения разнообразия библиотеки рекомендуется проведение массового параллельного секвенирования ДНК-баркодов на разных этапах создания библиотеки: начиная от секвенирования плазмидных векторов, заканчивая секвенированием клеток, зараженных баркодированными лентивирусами. На основе полученных данных создаются референсные библиотеки, включающие в себя все «доступные» варианты баркодов. Тем не менее, для больших библиотек этот способ является весьма трудоемким. Чем больше библиотека, тем больше должна быть глубина секвенирования. Мы использовали альтернативный метод, позволяющий оценить, по крайней мере, порядок разнообразия библиотеки. Поскольку все эксперименты должны были быть проведены с использованием вирусных частиц, собранных при одной трансфекции (то есть из одного стока), то оценив долю баркодов, общих для нескольких экспериментов, мы можем предсказать примерное разнообразие общей библиотеки. Таким образом, ДНК-баркоды из разных экспериментов являются случайными выборками одной генеральной совокупности. Чем меньше пересечений между экспериментами, тем больше разнообразие общей библиотеки. Следовательно, чем больше экспериментов, включающих секвенирование баркодов из одной общей лентивирусной библиотеки, мы проведем, тем больше будет данных для оценки разнообразия этой общей библиотеки.

III.3.2 Количество интеграций баркодов на одну клетку

Другой важный параметр, который необходимо учесть при баркодировании, – это количество интеграций ДНК-баркодов (т.е. провирусных копий), приходящееся на одну клетку. Известно, что количество провирусных копий на геном клетки (MOI – Multiplicity of infection) подчиняется распределению Пуассона. Значение MOI, равное = 1, говорит о том, что в среднем на одну клетку приходится одна вирусная интеграция, при этом с вероятностью 36,7% клетка инфицирована одной вирусной частицей, с вероятностью

18,39% двумя, вероятность инфекции тремя частицами 6,13%, а вероятность того, что в клетку не попала ни одна вирусная частица – 36,7%. Таким образом, чем больше значение MOI, тем больше клеток несет одну и более вирусных интеграций. Для того, чтобы уменьшить вероятность множественного баркодирования одной клетки, необходимо проводить трансдукцию клеток баркодированными лентивирусами с низкими значениями MOI (меньше 1). В противном случае, клетка будет инфицирована несколькими вирусными частицами, и реальное пересечение ДНК-баркодов с разных чашек будет завышено. Так, например, одна клетка с двумя ДНК-баркодами будет интерпретироваться как две клетки с индивидуальными ДНК-баркодами (Рисунок 6Б). Ситуация с лентивирусами, несущими репрограммирующие факторы, строго противоположна: необходимо, чтобы каждая клетка имела хотя бы одну копию полицистронной кассеты с факторами Яманаки. Поэтому этот лентивирус был взят в избытке.

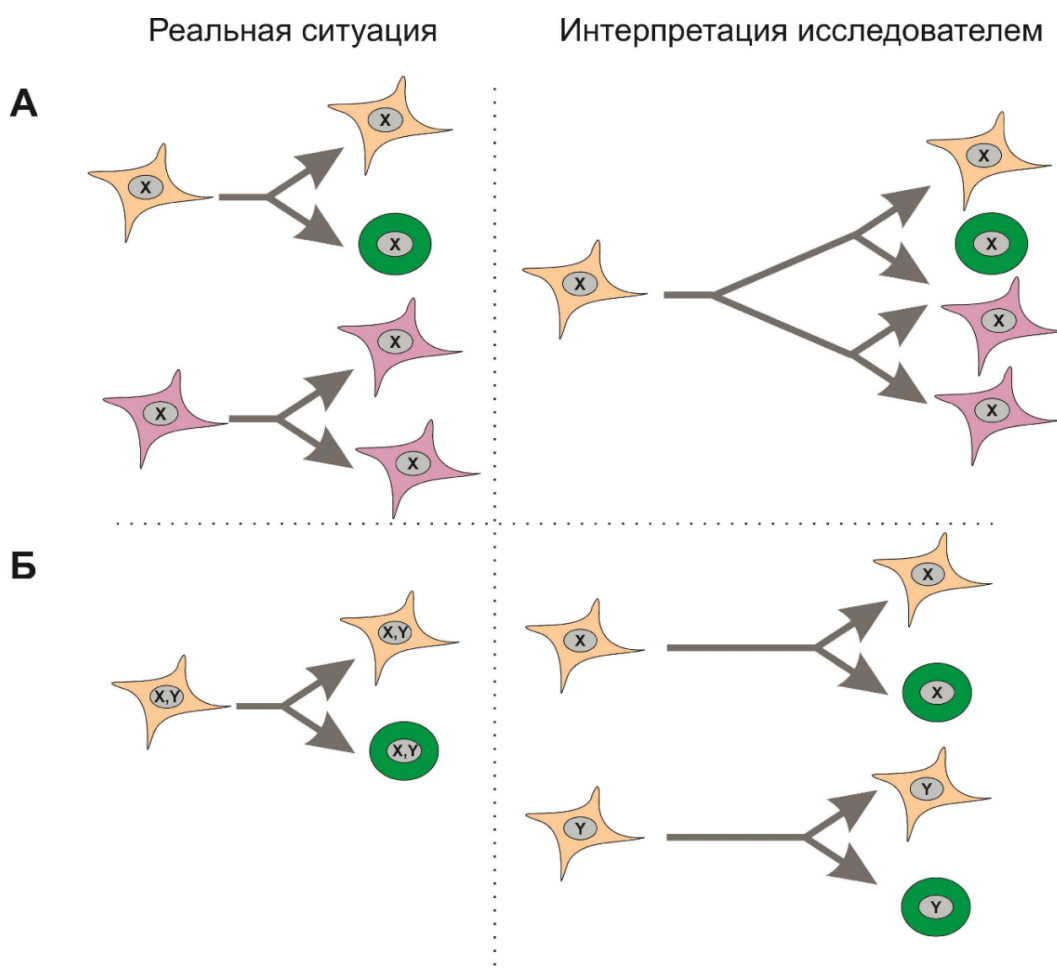


Рисунок 6. Вероятные случаи неправильной интерпретации судьбы клеток (адаптировано из [Naik, Schumacher, Peri, 2014])

(А). Вариант маркирования двух клеток одинаковым ДНК-баркодом. В этом случае клетки с разной судьбой будут интерпретироваться как одна клетка.

(Б). Вариант множественного баркодирования одной клетки. В этом случае одна клетка с несколькими баркодами будет восприниматься как несколько клеток с одинаковой судьбой.

Поскольку в прямом эксперименте по репрограммированию фибробластов установить MOI не представляется возможным, мы провели дополнительные эксперименты. Суть последних заключалась в трансдукции эмбриональных фибробластов мыши в условиях, аналогичных условиям в эксперименте по репрограммированию, включая: плотность рассадки клеток до трансдукции, время трансдукции, а также количество вируса с ДНК-баркодом. За одним исключением; вирусы с репрессирующими факторами были заменены на вирусы с *Gfp* под контролем доксициклин-индуцибельного промотора. Напомним, что баркодированные вирусы несут белок-трансактиватор, который при взаимодействии с доксициклином будет запускать экспрессию GFP. Таким образом, заражая клетки избытком вирусов с GFP и баркодированными вирусами, мы можем оценить долю клеток, получивших баркод, по наличию экспрессии GFP. Процент GFP-позитивных клеток был определен с помощью проточного цитофлуориметра (FACS), а MOI вычислено по формуле Пуассона. Все дополнительные эксперименты проводились параллельно экспериментам по репрограммированию.

III.3.3 Количество потомков исходных клеток

Схема эксперимента предполагает рассадку клеток на 4 чашки после удвоения клеточной популяции. Однако, в связи с асинхронностью клеточного цикла, некоторые клетки за отведенное время не успеют поделиться, тогда как другие клетки делятся более одного раза. В пользу последнего свидетельствует тот факт, что в пилотном эксперименте мы наблюдали присутствие одного и того же ДНК-баркода на всех 4-ех чашках (Рисунок 5Б). Это может иметь следующее объяснение: некоторые клетки образовали как минимум 4 потомка, которые попали на разные чашки и репрограммировались. Чем больше потомков имеет клетка, тем больше вероятность распределения этих потомков по разным чашкам. Стоит отметить, что в нашем дизайне эксперимента случаи совместного репрограммирования сестринских клеток мы можем детектировать только при их распределении по разным чашкам.

Таким образом, параметр *C* (вероятность распределения потомков на разные чашки) в формуле (2), основанный на предположении о единичном случае деления клеток, некорректен. И для того, чтобы более точно его оценить, нам необходима информация о количестве потомков каждой клетки, которое, следует отметить, может равняться даже 0 (гибель клетки до деления). Для мониторинга количества делений трансдуцированных фибробластов, мы проводили прижизненную микроскопию с детальной съемкой в течение 24 часов на приборе CellIQ в центре клеточных технологий ИЦиГ СО РАН (Рисунок 7).

III. 4 Репрограммирование баркодированных фибробластов мыши к плюрипотентному состоянию: Эксперименты 1 и 2 (Несортированные ЭФМ)

Следующие два эксперимента мы провели в соответствии с вышеуказанными изменениями. В течение 24 часов после трансдукции клетки культивировались в приборе CellIQ со съемкой каждые 15 минут. На Рисунке 7 представлен пример фотографий клеток на одном поле в разные промежутки времени.

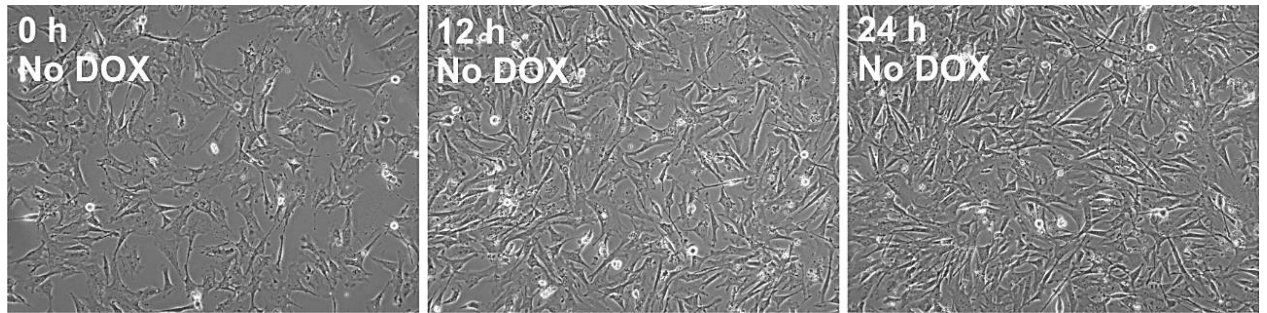


Рисунок 7. Эмбриональные фибробласты мыши после трансдукции лентивирусами с репрограммирующими факторами в указанные промежутки времени. Клетки культивировались в отсутствие доксициклина

Вручную проанализировав фотографии, мы проследили за судьбой около 300 клеток в каждом эксперименте и определили процентную долю клеток, оставивших от 0 до 8 потомков за отведенное время (Приложение 2). Значения MOI были вычислены из дополнительных экспериментов (замена вируса с репрограммирующими факторами на вирус с GFP). Мы полагаем, что эти значения верны и для экспериментов по репрограммированию. Согласно полученным данным, менее половины клеток (39% в первом эксперименте и 29% - во втором) получили в ходе вирусной трансдукции, как минимум один баркод, а вместе с ним белок-трансактиватор и принципиальную возможность репрограммироваться. Это соответствует требованиям низкого MOI для метода баркодирования с помощью лентивирусов. Все параметры этих и последующих экспериментов отражены в Приложении 3.

После 7-ми дней культивирования в среде с DOX, GFP-позитивные клетки были отобраны с помощью проточного цитофлуориметра, а их ДНК-баркоды амплифицированы и секвенированы.

Мы провели биоинформационную обработку данных массового параллельного секвенирования в соответствии с описанием, приведенным в материалах и методах. Количество ридов до обработки и после прохождения всех фильтров отражено в Приложении 1. Оставшиеся риды были кластеризованы по гомологии и распределены на четыре группы в соответствии с принадлежностью одной из четырех чашек. Мы проанализировали качественный состав ДНК-баркодов на каждой чашке и обнаружили, что

значительная часть баркодов представлена более чем на одной чашке, т.е. пересекается (Рисунок 8Б, Г). Также на рисунке 8 (А, В) представлено соотношение количества общих кластеров к уникальным. Видно, что значительную долю составляют именно общие баркоды. Это свидетельствует в пользу синхронного репрограммирования дочерних клеток

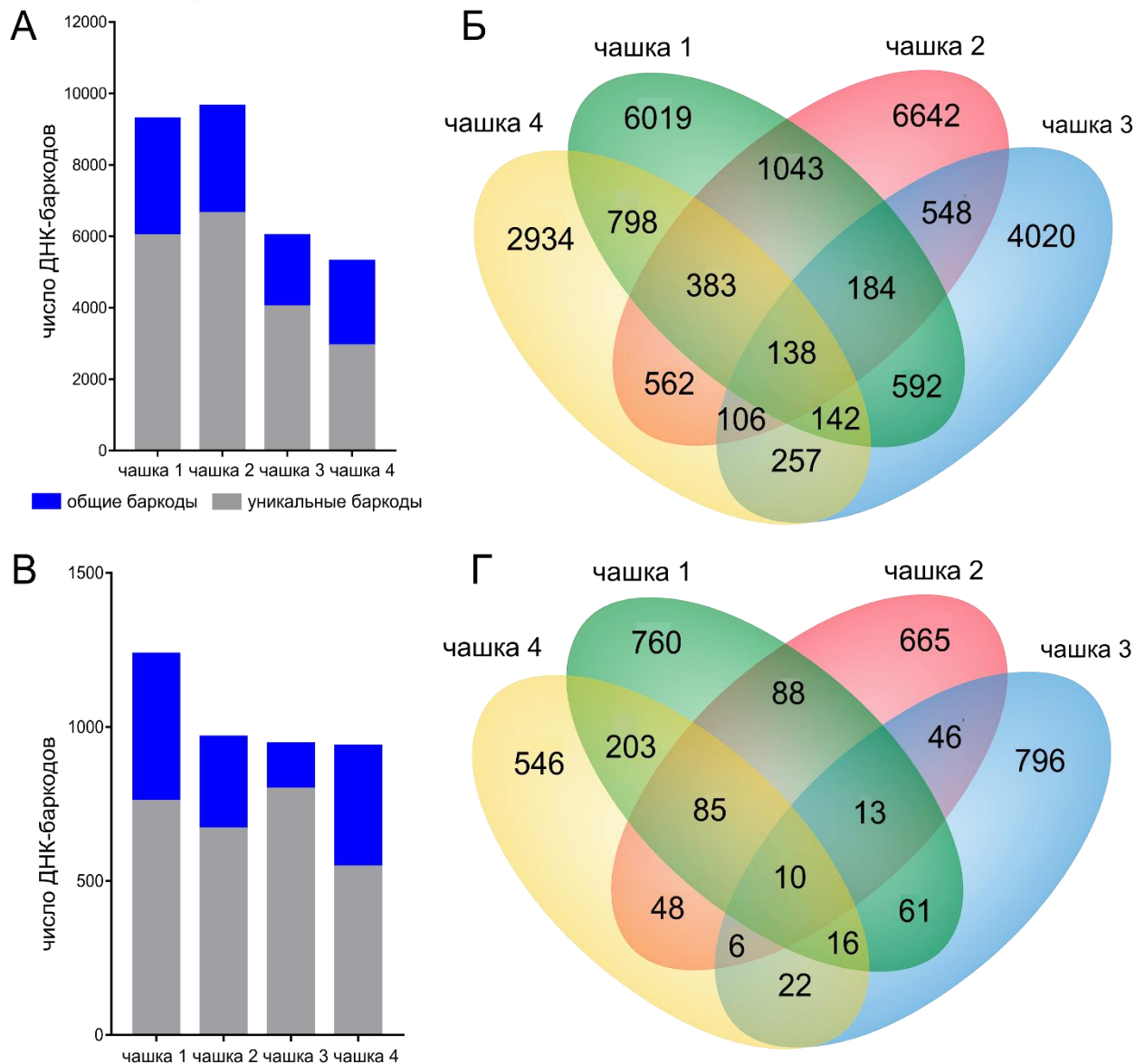


Рисунок 8. Сравнение ДНК-баркодов GFP-позитивных клеток в четырех чашках для первого (А, Б) и второго (В, Г) экспериментов.

(А) и (В) отражают соотношение общих ДНК-баркодов (синие блоки) и уникальных ДНК-баркодов (серые блоки) на каждой чашке. Уникальные баркоды – баркоды, присутствующие только на указанной чашке; общие баркоды – баркоды, обнаруженные на нескольких чашках.

(Б) и (Г) Диаграммы Венна для первого и второго эксперимента соответственно. Здесь показаны все пересечения ДНК-баркодов между четырьмя чашками.

III.5 Секвенирование контрольных клеток с известными ДНК-баркодами

Для проверки, вносят ли ПЦР и секвенирование смещение в представленность баркодов мы включили в каждый эксперимент набор контрольных клеток с известными

ДНК-баркодами (см. схема эксперимента Рисунок 2). В качестве контрольных клеток были использованы два субклона клеток линии Phoenix, каждый из которых нес один единственный, причем известный нам баркод, далее по тексту – контрольный баркод. Получение данных субклонов и последовательности нуклеотидов контрольных баркодов описаны в материалах и методах. В первом эксперименте контрольные клетки в количестве 400 (для первого субклона) и 100 (для второго субклона) были внесены в каждую из четырех пробирок с сортированными GFP-позитивными клетками. Для второго эксперимента соответствующие числа клеток равнялись 80 и 20. Контрольные клетки подвергались тем же воздействиям (проходили те же стадии), что и отсортированные в конце эксперимента репрограммированные GFP-позитивные клетки. Это лизис клеток, амплификация и секвенирование баркодов. После биоинформационного анализа данных секвенирования мы кластеризовали риды контрольных баркодов согласно процедуре, описанной в материалах и методах. Во-первых, каждый из контрольных баркодов был найден во всех библиотеках (Рисунок 9). Во-вторых, доля кластеров контрольных баркодов коррелировала с количеством контрольных клеток, которым этот баркод принадлежал: большему числу контрольных клеток соответствовало большее количество ридов контрольных баркодов, за исключением чашки 4 в эксперименте 2 (Рисунок 9Б).

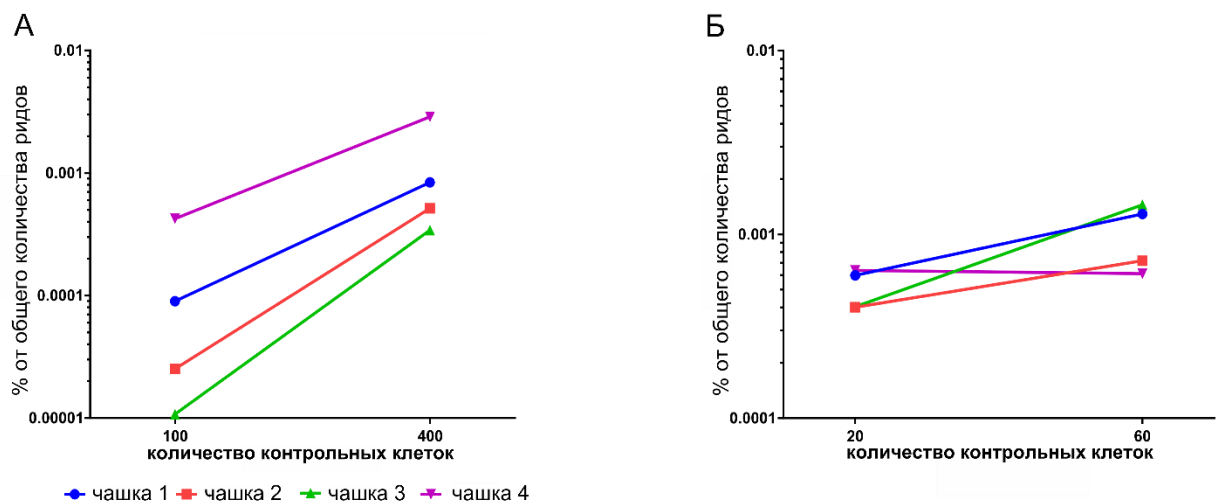


Рисунок 9. Зависимость процентного содержания ридов контрольных баркодов в библиотеке от количества соответствующих контрольных клеток для первого (А) и второго (Б) экспериментов. По оси Х отложено количество контрольных клеток, добавленных в каждый образец; по оси Y – доля ридов контрольных баркодов от общего количества ридов.

Важно отметить, что количество ридов в кластерах контрольных баркодов значительно отличалось для библиотек, полученных с разных чашек. Это хорошо видно на рисунке 9А (эксперимент 1): для чашек 3 и 4 разница в представленности ридов достигает почти два порядка. Это свидетельствует о том, что качество библиотеки не является постоянной величиной, и на разных стадиях эксперимента часть материала может теряться.

Кроме того, количество уникальных вариантов ДНК-баркодов варьирует для разных чашек в одном эксперименте (Приложение 4). Однако, нет оснований полагать, что исходные клетки, поделенные на четыре равные части и культивирующиеся при одинаковых условиях, будут демонстрировать значительно различающуюся эффективность репрограммирования.

III.6 Анализ потери материала в ходе получения библиотеки (Секвенирование 3x5000 баркодированных клеток)

Напомним, что параллельно каждому эксперименту по репрограммированию мы проводили дополнительный эксперимент по оценке MOI баркодированного вируса. В дополнительном эксперименте, параллельном первому эксперименту по репрограммированию, мы отобрали 3 группы GFP-позитивных клеток в количестве 5000 на клеточном сортере. С каждой из групп мы получили библиотеку и провели массовое параллельное секвенирование их ДНК-баркодов.

Полученные библиотеки являются хорошей тест-системой; мы знаем точное число клеток, из которых они были получены (5000), а также количество баркодов, приходящееся на одну клетку (ведь именно по этому эксперименту мы определяем MOI). Таким образом, мы получаем возможность предсказать количество кластеров баркодов после секвенирования. Однако, сравнив теоретические данные с экспериментальными, мы обнаружили значительные несоответствия. В эксперименте, после биоинформационной обработки и кластеризации ридов, мы получили 4251, 5176 и 3620 вариантов уникальных ДНК-баркодов. Тогда как согласно теории, с поправкой на MOI, количество кластеров должно быть около 6000. Этот результат в очередной раз свидетельствует о потере ДНК-баркодов в ходе получения библиотек. Мы вычислили потерю ДНК-баркодов исходя из разницы в наблюдаемом и ожидаемом количестве баркодов (Рисунок 10). Из рисунка видно, что этот параметр варьирует для трех библиотек, начиная от 18% для второй группы и до 42% для третьей группы.

Необходимо отметить, что потеря ДНК-баркодов возможна на любом этапе эксперимента, начиная от потерь клеток при рассадке на четыре чашки и сортировки репрограммированных клеток до неэффективной амплификации и недостаточной глубины секвенирования.

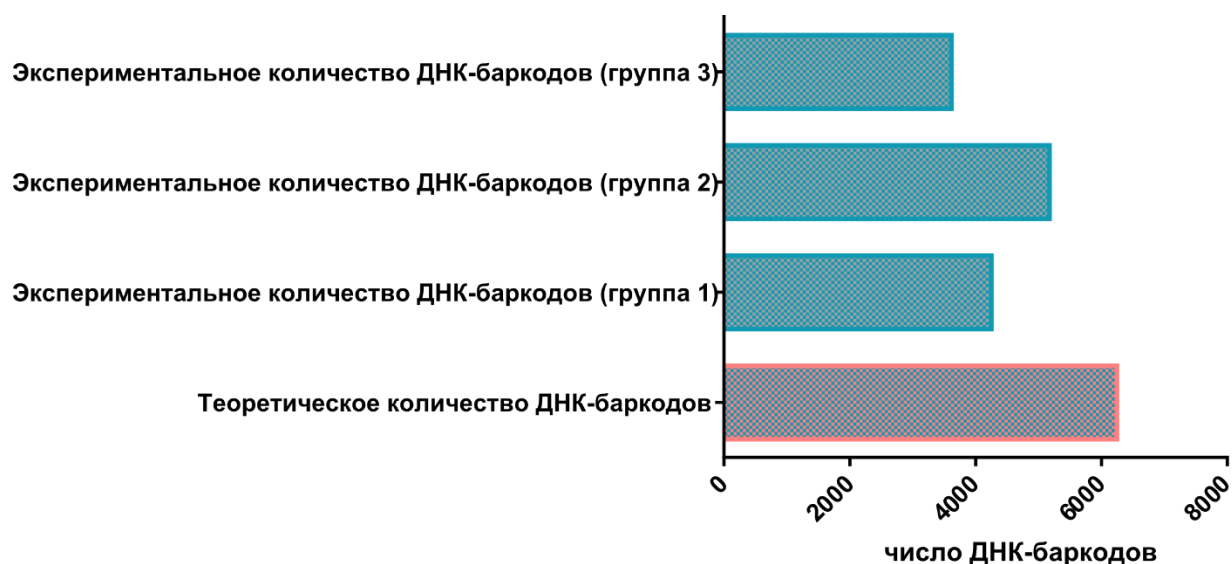


Рисунок 10. Сравнение теоретических и экспериментальных данных по количеству кластеров ДНК-баркодов в группах из 5000 клеток.

III.7 Оценка размера библиотеки баркодированных лентивирусных векторов

Всего на данный момент нами были просеквенированы три суб-библиотеки ДНК-баркодов, полученные из клеток, зараженных вирусами из одного стока. Это суб-библиотеки из репрограммированных клеток в первом и втором экспериментах, а также суб-библиотека из трех групп по 5000 клеток в дополнительном эксперименте. Количество уникальных ДНК-баркодов, принадлежащих каждой суб-библиотеке, а также число ДНК-баркодов, общих для двух любых суб-библиотек показаны на Рисунке 11.

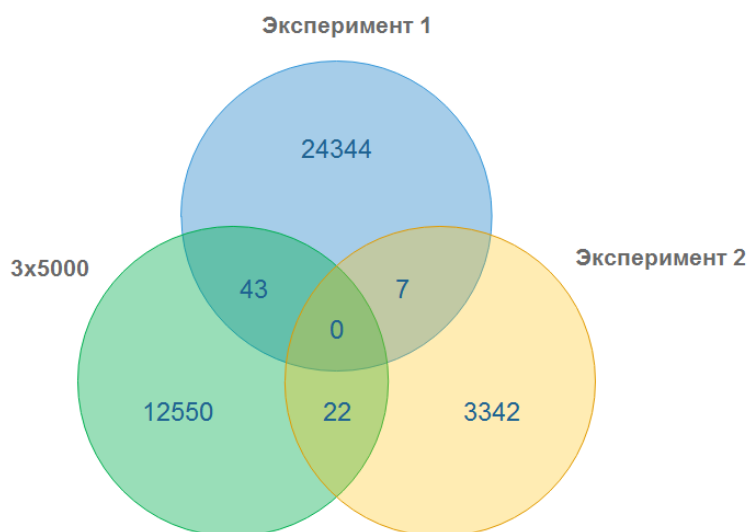


Рисунок 11. Пересечения ДНК-баркодов между тремя суб-библиотеками.

Как было описано выше, количество пересечений ДНК-баркодов между экспериментами позволяет судить о размере общей библиотеки (см раздел «Разнообразие библиотеки баркодированных вирусов»). С помощью компьютерного моделирования мы

подобрали наиболее подходящие размеры общей библиотеки для каждого из попарных сравнений трех суб-библиотек (Таблица 3 и Приложение 5).

	Эксперимент 1	Эксперимент 2	3x5000
Эксперимент 1		12×10^6	7×10^6
Эксперимент 2	12×10^6		2×10^6
3x5000	7×10^6	2×10^6	

Таблица 3. Наиболее подходящие размеры общей библиотеки баркодированных вирусов для каждого попарного сравнения суб-библиотек.

Из таблицы видно, что значения довольно сильно различаются между собой, однако, даже наименьшее из них, 2000000, по определению, подходит для маркирования нашей экспериментальной популяции клеток.

III.8 Описание компьютерной модели, симулирующей экспериментальные стадии

Как было описано ранее, мы обнаружили значительные пересечения ДНК-баркодов между чашками. Было бы интересно оценить, какая доля этих пересечений действительно обусловлена синхронным репрограммированием потомков одной клетки (и наследуемым!), а не особенностями методики баркодирования или случайным совпадением. Для этого необходимо учесть ряд описанных выше параметров: это размер библиотеки баркодированных вирусов, количество интеграций ДНК-баркода на клетку, различное число потомков трансдуцированных клеток, а, следовательно, и вероятность распределения на разные чашки, и, напоследок, потеря материала (как клеток, так и ДНК) на разных стадиях приготовления библиотеки. Поскольку с учетом этих параметров расчет по формуле не представляется возможным, мы разработали компьютерную модель, симулирующую все стадии эксперимента и позволяющую оценить вероятность репрограммирования сестринских клеток (т.е. ту самую «наследуемость» способности к репрограммированию). Схема функционирования модели отражена на Рисунке 12 и более подробно описана в главе «Материалы и методы». Кратко, алгоритм модели следующий. Сначала создается некоторое количество объектов, численно равное экспериментальному количеству клеток. Затем объектам присваиваются номера, что отражает мечение клеток ДНК-баркодом, причем номера присваиваются в соответствие со значением MOI, то есть какие-то клетки получают два и более номера, тогда как другие не получают свой номер-метку совсем и исключаются из дальнейшего моделирования. Далее объекты мультиплицируются, исходя из количества делений клеток после трансдукции (от 0 до 8). После этого, объектам присваивается еще один номер – номер чашки (от 1 до 4), таким

образом имитируется рассадка на 4 чашки. Затем часть каждой группы элиминируется (параметр «потеря материала», d). Поскольку этот параметр не может быть оценен экспериментально, в модели он варьируется от 10 до 90%. И, наконец, последний шаг в моделировании – это отбор количества объектов из каждой группы, равного экспериментально полученному количеству кластеров ДНК-баркодов на каждой чашке. Отбор объектов проводится согласно нормальному распределению, где n – среднее количество кластеров ДНК-баркодов на четырех чашках с соответствующей дисперсией s . Кроме того, в модель был введен коэффициент наследуемости способности к репрограммированию, h , который также варьируется от 0 до 90%. Этот коэффициент определяет вероятность синхронного репрограммирования потомков одной клетки. Для простоты, в дальнейшем этот коэффициент будет обозначаться как коэффициент или уровень наследуемости.



Рисунок 12. Схема компьютерной модели, симулирующей экспериментальные этапы. Параметр d обозначает потерю материала, параметр h – коэффициент наследуемости. Подробное описание в тексте.

Результатом моделирования является число (т.е. баркоды), общих для двух групп (двойные пересечения), трех групп (тройные пересечения) и, наконец, четырех групп (четверные пересечения). Цель моделирования заключается в подборе такого коэффициента наследуемости, при котором количество двойных пересечений между четырьмя группами в модели значительно не отличается от двойных пересечений баркодов

между чашками в эксперименте. Поскольку на каждую точку d приходится несколько подходящих значений h , наиболее подходящее выбирается по близости средних значений.

III.9 Компьютерное моделирование Экспериментов 1 и 2

Мы провели моделирование первого и второго эксперимента при соответствующих параметрах, которые отражены в Приложении 3. Для каждого эксперимента было сделано 100 запусков моделирования при разнообразии библиотеки в 2 000 000 уникальных вариантов ДНК-баркодов. Поскольку разнообразие библиотеки не было определено однозначно, мы решили использовать наименьшее значение данного параметра (см раздел оценка размера библиотеки баркодированных лентивирусов).

В результате моделирования мы получили диапазон подходящих значений коэффициентов наследуемости в зависимости от потери материала (Рисунок 13). Из рисунка видно, что чем больше ДНК-баркодов теряется в эксперименте, тем больше значение наследуемости мы должны задать в модели, чтобы двойные пересечения совпали. Согласно результатам моделирования, уровень наследуемости при минимальной потере материала варьирует от 0 до 10% в первом эксперименте, во втором эксперименте от 0 до 30% (Рисунок 13).

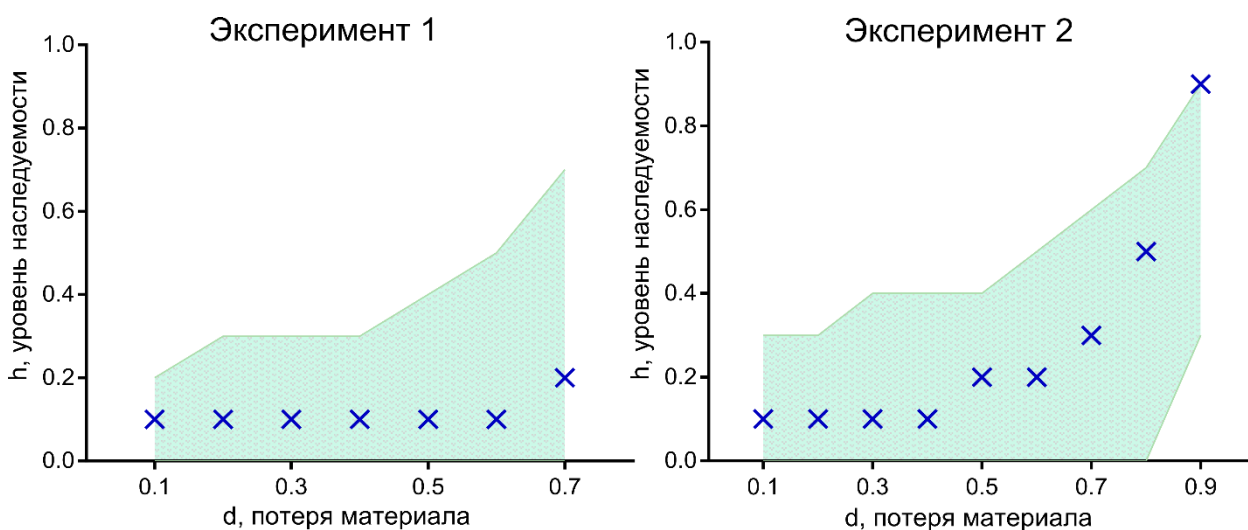


Рисунок 13. Результаты компьютерного моделирования первого и второго экспериментов. По оси X отмечена потеря материала; по оси Y –уровень наследуемости, при котором количества пересечений ДНК-баркодов между чашками значимо не отличаются в эксперименте и модели ($p < 0.05$). Заштрихованная область показывает подходящий уровень наследуемости при определенной потере материала. Крестиками выделены *наиболее* подходящие значения. Анализ проводился с помощью непараметрического дисперсионного анализа (критерий Краскела-Уоллиса) в программе GraphPad Prism 6. График, показывающий результаты моделирования первого эксперимента, не включает значения h при d , равной 0.8 и 0.9, по причине невозможности функционирования алгоритма при данных параметрах для этого эксперимента.

Для того, чтобы проверить достоверность полученного результата и определить число пересечений баркодов между чашками при условии, что успешное приобретение плюрипотентности определяется лишь случаем (то есть соответствие стохастической модели), мы разработали контрольный эксперимент, симулирующий отбор репрограммированных клеток случайным образом (т.е. при нулевой наследуемости).

III.10 Репрограммирование баркодированных фибробластов мыши к плюрипотентному состоянию: Эксперимент 3 (ЭФМ, сортированные по Thy-1) и Контрольный эксперимент

Схема контрольного эксперимента почти полностью повторяет схему эксперимента по репрограммированию. За исключением того, что вирус с репрограммирующими факторами был заменен на вирус с доксициклин-индуцибельным GFP. ДНК-баркодирование этих клеток было проведено при тех же условиях, включая начальную плотность рассадки клеток, количество вирусов с ДНК-баркодом и время трансдукции. Контрольный эксперимент проводился параллельно эксперименту по репрограммированию. После вирусной трансдукции клетки также культивировались в приборе CellIQ в течение суток для подсчета количества делений. Затем клетки были рассажены на 4 чашки и культивировались в присутствии доксициклина. Через неделю, мы *случайным образом* отбирали на клеточном сортере столько же GFP-позитивных клеток с каждой чашки, сколько в параллельном эксперименте по репрограммированию. Поскольку клетки были отобраны случайным образом, пересечения ДНК-баркодов между четырьмя чашками в этом эксперименте обусловлены случайными событиями. И сравнив эти пересечения с пересечениями в эксперименте по репрограммированию, мы можем выявить наличие наследственной компоненты у способности к репрограммированию.

Кроме того, мы решили ограничить исходную экспериментальную популяцию клеток только Thy1-позитивными клетками (Thy1 – маркер терминально дифференцированных фибробластов). Это позволяет снизить гетерогенность исходных клеток и исключить из эксперимента менее дифференцированные клетки, являющиеся одними из претендентов на роль «элитарной» субпопуляции. Поэтому и контрольный эксперимент, и эксперимент по репрограммированию проводились на сортированных Thy1-позитивных клетках.

Библиотеки ДНК-баркодов из контрольных и репрограммированных клеток были просеквенированы на платформе Ion Torrent. Результаты биоинформационной обработки полученных данных представлены в Приложении 1 (обозначены как общая библиотека для третьего эксперимента). Мы обнаружили, что количество двойных, тройных и четверных

пересечений ДНК-баркодов между чашками в контрольном эксперименте значительно меньше чем в эксперименте по репрограммированию (Рисунок 14). Данный результат хорошо согласуется с тем фактом, что баркоды в контрольном эксперименте были выбраны случайно и наследуемости не предполагается совсем.

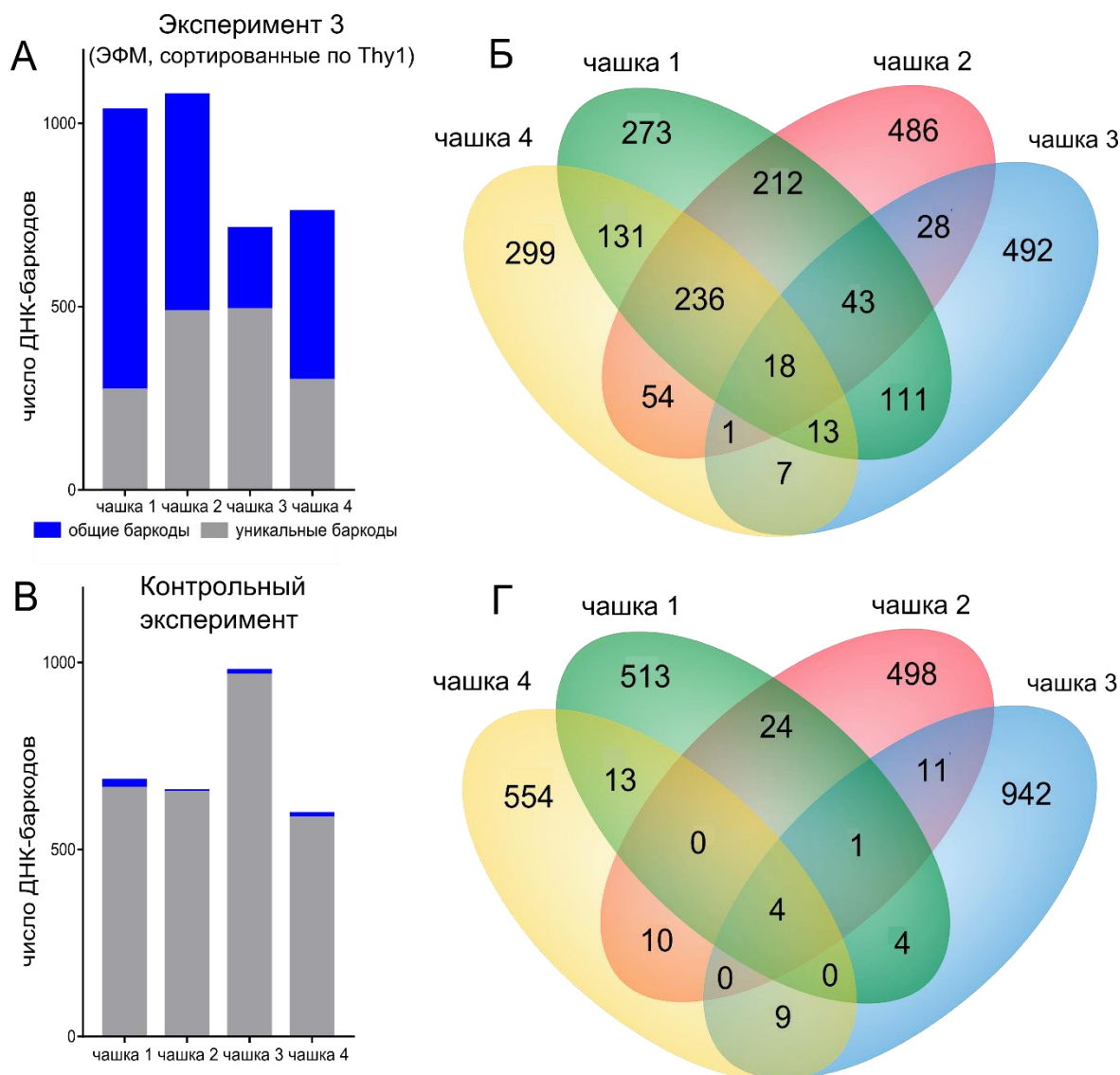


Рисунок 14. Сравнение ДНК-баркодов успешно репрограммированных клеток в четырех чашках для третьего эксперимента (репрограммирование Thy⁺ клеток) (А, Б) и для контрольного эксперимента со случайным отбором клеток (В, Г).

Аналогично рисунку 8, диаграммы Венна (Б) и (Г) иллюстрируют количество уникальных кластеров баркодов, а также двойные, тройные и четверные пересечения между чашками.

(А) и (В) показывают соотношение общих ДНК-баркодов (синие блоки) и уникальных ДНК-баркодов (серые блоки) на каждой чашке. Уникальные баркоды – баркоды, присутствующие только на указанной чашке; общие баркоды – баркоды, обнаруженные на нескольких чашках.

После секвенирования еще двух суб-библиотек баркодов мы получили новые данные, которые можно использовать для коррекции оценки размера общей библиотеки.

Согласно попарному сравнению суб-библиотек из первого и второго эксперимента с суб-библиотекой из третьего эксперимента (контрольный эксперимент + эксперимент по репрограммированию), размер библиотеки составляет 6 000 000 и 160 000 уникальных баркодов соответственно (Приложение 6). Когда же мы сравнили между собой ДНК-баркоды из контрольного эксперимента и эксперимента по репрограммированию, то размер общей библиотеки оказался еще меньше – около 50 000 вариантов баркодов. Поэтому при моделировании третьего эксперимента мы решили придерживаться минимального значения размера библиотеки.

Согласно компьютерному моделированию уровень наследуемости способности к репрограммированию получился даже больше, чем в предыдущих экспериментах – его минимальное значение составило 20% (Рисунок 15).

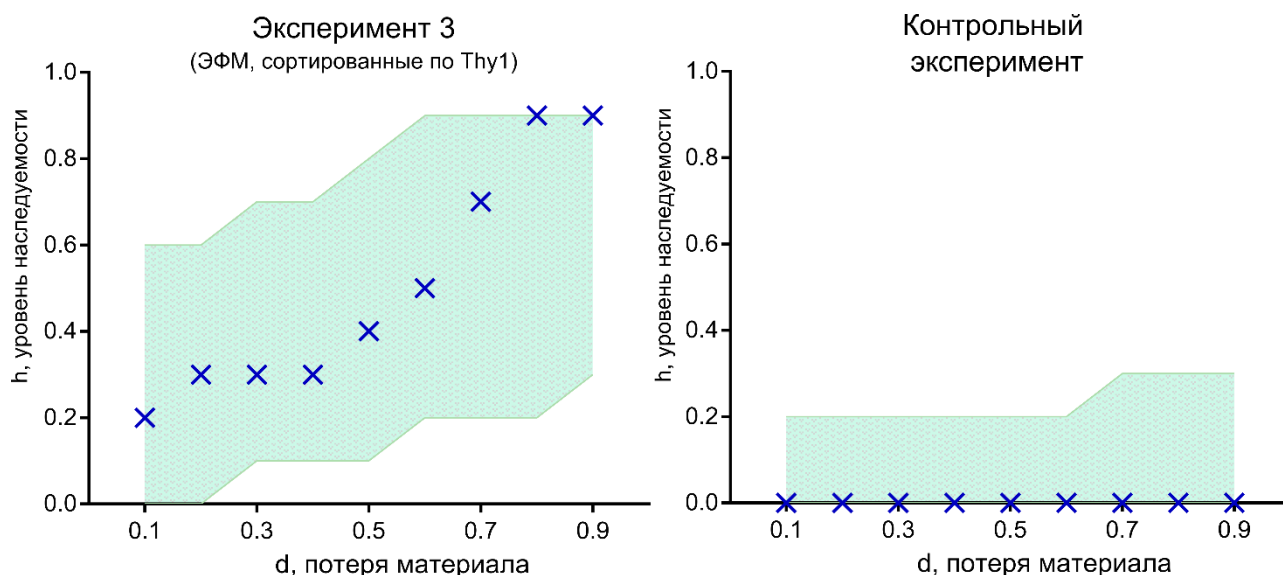


Рисунок 15. Результаты компьютерного моделирования Эксперимента 3 и контрольного эксперимента (сортированные по Thy1 фибробласты). Аналогично рисунку 13, по оси X отмечена потеря материала; по оси Y –уровень наследуемости, при котором количества пересечений ДНК-баркодов между чашками значимо не отличаются в эксперименте и модели ($p < 0.05$). Заштрихованная область показывает подходящий уровень наследуемости при определенной потере материала. Крестиками выделены *наиболее* подходящие значения. Анализ проводился с помощью непараметрического дисперсионного анализа (критерий Краскела-Уоллиса) в программе GraphPad Prism 6.

На рисунке хорошо видно, насколько сильно различаются результаты моделирования для контрольного эксперимента и эксперимента по репрограммированию. Таким образом, мы полагаем, что некоторые клетки способны передавать способность к репрограммированию в плюрипотентные клетки по наследству. Кроме того, исключение малодифференцированных соматических стволовых клеток из нашего эксперимента не привело к уменьшению доли синхронно репрограммирующихся дочерних клеток.

III.11 Анализ вклада быстроделющихся клеток в общее число двойных пересечений

Из результатов компьютерного моделирования можно узнать не только количество пересечений ДНК-баркодов между чашками, но то, какие клетки за них ответственны. Мы проанализировали вклад клеток с разным количеством потомков в общее число двойных пересечений (Рисунок 16). В Экспериментах 1 и 2 клетки, образовавшие 4 и больше потомков, отвечали более чем за половину случаев двойных пересечений, хотя их доля в исходной популяции не превышала 25% (Приложение 2). Для Эксперимента 3 результаты немного отличались. Так, максимальное количество оставленных потомков не превышало четырех. Наиболее правдоподобным выглядит следующее объяснение: при сортировке Thy1-позитивных клеток из анализа были исключены быстроделющиеся соматические стволовые клетки. И даже в этом случае, при пересчете на долю от исходной популяции, вклад клеток с четырьмя потомками больше чем клеток с двумя потомками.

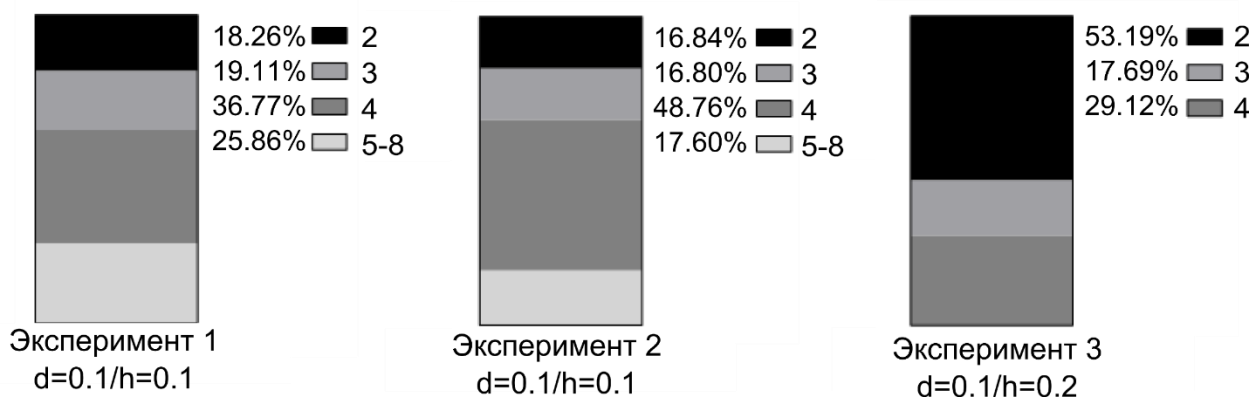


Рисунок 16. Клетки, образовавшие большее число потомков, дают больший вклад в количество репрограммированных клеток. На гистограммах отражено соотношение вкладов клеток с разным количеством потомков в общее количество двойных пересечений. Данные получены с помощью компьютерного моделирования на выбранных параметрах d и h (наиболее подходящий уровень наследуемости, h при минимальной потере материала, d).

III.12 Компьютерное моделирование двойных, тройных и четверных пересечений баркодов при дифференциальной способности клеток к репрограммированию

Необходимо упомянуть, что в нашей модели уровень наследуемости варьировался от 0 до 0.9 для каждого значения потери материала, но был равным для всех объектов-клеток. То есть модель предполагает однородное распределение передачи способности к репрограммированию в клеточной популяции. Тогда вклад клеток с большим числом потомков в пересечения баркодов между чашками обусловлен исключительно количеством потомков, а, следовательно, большей вероятностью быть выбранными. Между тем, известна гипотеза, что быстроделющиеся клетки могут иметь предрасположенность к репрограммированию за счет необходимых эпигенетических преобразований,

происходящих в ходе делений. Мы решили протестировать эту гипотезу и при компьютерном моделировании присвоили быстроделаящимся клеткам больший уровень наследуемости h . Кроме того, мы сравнили тройные и четверные пересечения в эксперименте и модели. Оказалось, что при однородном распределении уровня наследуемости количество тройных и четверных пересечений ДНК-баркодов в эксперименте было значительно больше чем в результате моделирования (Рисунок 17). Тогда как при увеличении уровня наследуемости для клеток с большим количеством потомков эта разница уменьшается.

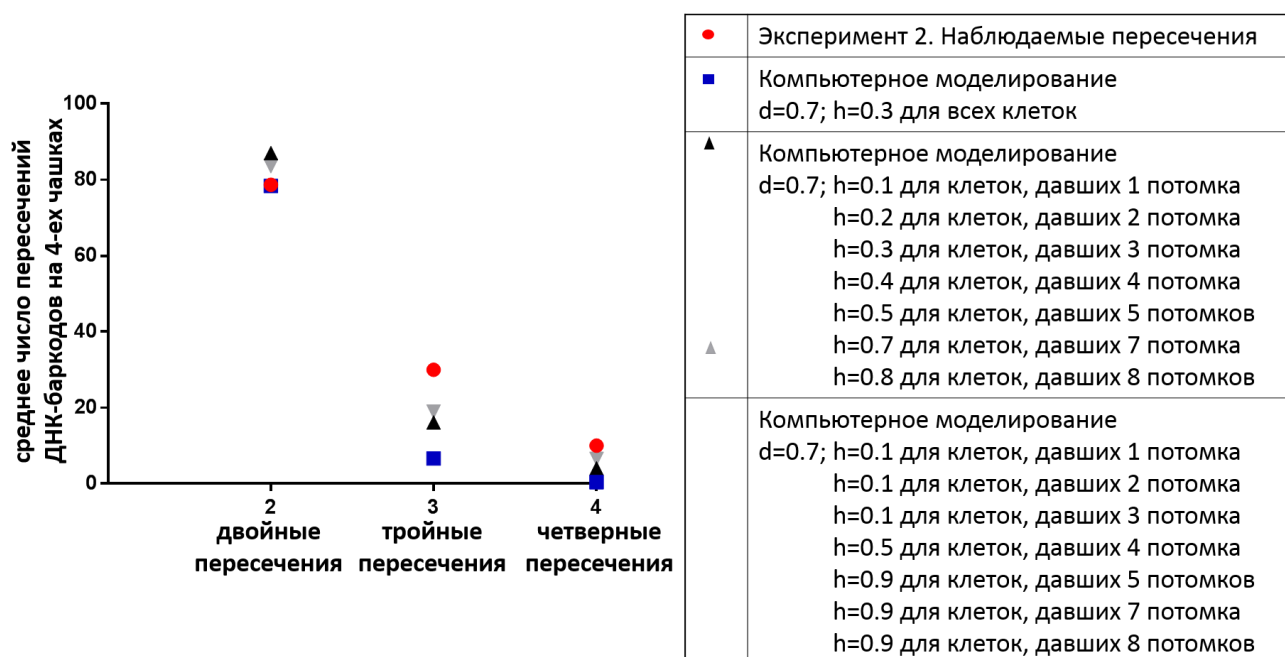


Рисунок 17. Вероятность синхронного репрограммирования сестринских клеток больше для быстроделаящихся клеток. На графике показано сравнение двойных, тройных и четверных пересечений ДНК-баркодов при разных распределениях способности к репрограммированию в клеточной популяции (однородное и дифференциальное распределения). В качестве модельных взяты пересечения, полученные при параметрах $d=0.7$ и $h=0.3$ (см рисунок 12). Результаты моделирования при других значениях h и d и других экспериментов не показаны в силу того, что они не отличаются от результатов, представленных выше.

Глава IV. Обсуждение

Вопрос о том, обладают ли клетки наследуемой предрасположенностью к репрограммированию, ставится исследователями уже давно. Тем не менее, определенного ответа на него все еще нет, а имеющиеся данные противоречивы. Так, например, Араки с коллегами при помощи прижизненной микроскопической съемки на протяжении нескольких недель исследовали судьбу клеток в ходе репрограммирования. Ретроспективный анализ (от плюрипотентной клетки к образовавшему ее «предковому» фибробласту) показал, что первое деление клеток после трансдукции репрограммирующими факторами является асимметричным. Эта асимметричность проявляется в том, что одна из сестринских клеток образует потомков, которые становятся плюрипотентными, тогда как другая сестринская клетка, как правило, погибает через несколько дней после начала эксперимента. Интересно, что второе деление оставшейся в живых клетки симметричное, и оба ее потомка дают начало плюрипотентным клеткам [Araki et al., 2010]. Таким образом, даже *после* индукции репрограммирующими факторами сестринские клетки имеют разную судьбу, а, следовательно, никакой наследуемости быть не может. Другая работа, опубликованная не так давно в журнале EMBO Reports свидетельствует об обратном [Pour et al., 2015]. В этой работе была также использована прижизненная микроскопия, однако наблюдать за клетками и их родственными связями исследователи стали еще до запуска экспрессии факторов. Все ретроспективно проанализированные пары сестринских клеток делили общую судьбу: либо обе клетки репрограммировались, либо обе оставались рефрактерными (не увеличивали скорость пролиферации и не уменьшались в размере – см в литобзоре). Таким образом, исследователи предположили, что способность к репрограммированию предопределена еще до индукции факторов, и является наследуемой характеристикой. Недостатком обеих работ является небольшое количество проанализированных клеток.

В нашей работе, для исследования того, как способность к репрограммированию распределяется между сестринскими клетками, был использован метод ДНК-баркодирования, который позволяет проследить за судьбой тысяч клеток. ДНК-баркод представляет собой вырожденную последовательность из 30 нуклеотидов, которая, будучи интегрированной в геном с помощью лентивирусов, наследуется в ходе делений. Таким образом, потомки одной клетки несут один и тот же баркод, и могут быть легко определены с помощью массового параллельного секвенирования.

Мы маркировали эмбриональные фибробласты мыши ДНК-баркодами и репрограммировали их к плюрипотентному состоянию. После чего отобрали успешно репрограммированные клетки (активировавшие ген *Oct4*) и провели массовое параллельное секвенирование их ДНК-баркодов. Оказалось, что довольно большую долю ДНК-баркодов составляют именно баркоды родственных клеток (Рисунок 5; Рисунок 8; Рисунок 14).

Для того, чтобы определить, какая часть ДНК-баркодов действительно обусловлена репрограммированием сестринских клеток, а не особенностями методики баркодирования, мы разработали компьютерную модель, симулирующую эксперимент и позволяющую учесть различные экспериментальные параметры. Среди этих параметров – исходное количество клеток, посаженных на чашку, разнообразие библиотеки баркодированных лентивирусов, MOI/количество ДНК-баркодов, приходящееся на одну клетку, а также количество потомков исходных клеток. С помощью моделирования мы получаем возможность количественно оценить вероятность синхронного репрограммирования родственных клеток или, другими словами, оценить уровень наследуемости способности к репрограммированию.

IV.1 Оценка разнообразия библиотеки баркодированных лентивирусов

Разнообразие библиотеки, один из важнейших параметров для моделирования, заслуживает отдельного внимания. Как обсуждалось выше, во всех экспериментах, включающих секвенирование баркодов, были использованы вирусные частицы из одного стока (общей библиотеки). Поэтому мы оценивали разнообразие библиотеки по количеству пересечений баркодов между двумя любыми экспериментами (подробно описано в результатах: оценка размера библиотеки). Для разных экспериментов, полученные оценки варьировались от 50 000 до 12 000 000 и больше. В обзоре [Naik, Schumacher, Peri, 2014], единственном на сегодняшний день обзоре, посвященном методике баркодирования, отмечается, что минимальный размер библиотеки должен в десять раз превышать количество маркируемых клеток. Тем не менее, это правило является «правилом большого пальца», и скорее, может быть использовано как рекомендация. В нашей схеме исходное количество фибробластов, посаженных на чашку, составляет приблизительно 200 000. Однако трансдукция вирусами проходит при низких значениях MOI, что сокращает «эффективную» численность клеток до 60 000-80 000. Кроме того, эффективность репрограммирования достаточно низка, – согласно нашим результатам, в каждом эксперименте репрограммировалось не более 4000 клеток. В силу этого условия, «эффективная» численность сокращается еще больше. Таким образом, даже при минимальной оценке разнообразие библиотеки является вполне приемлемым. Кроме того,

при моделировании экспериментов при разных значениях разнообразия библиотеки, мы не обнаружили значительных различий (Приложение 7).

IV.2 Оценка наследуемости способности к репрограммированию через моделирование экспериментальных данных

Согласно моделированию, значение уровня наследуемости (h) лежит в диапазоне 0-30%. Здесь стоит оговорить, что данная оценка получена при минимальной потере материала (d). Однако вероятность минимальной потери материала все же довольно мала (что хорошо продемонстрировано в дополнительном эксперименте с секвенированием баркодов групп из 5000 клеток (Рисунок 10)), поэтому реальный уровень наследуемости должен превышать указанные значения. Разумно предположить, что потеря материала в эксперименте составляет не более 50%, а, следовательно, уровень наследуемости лежит в диапазоне от 0 до 60%, при наиболее подходящем значении равном 10%. Это значит, что при условии репрограммирования любой отдельно взятой клетки, вероятность репрограммирования родственных ей клеток составляет 10%.

Для того, чтобы проверить достоверность полученного результата мы провели контрольный эксперимент, симулирующий отбор репрограммированных клеток при нулевой наследуемости. В этом случае все клетки имеют равную вероятность быть отобранными, и родственные клетки могут попасть в анализ только за счет случая, что соответствует стохастической модели репрограммирования. Оказалось, что доля ДНК-баркодов родственных клеток в контрольном эксперименте значительно меньше чем в эксперименте по репрограммированию (Рисунок 14). Компьютерное моделирование показало такие же результаты: уровень наследуемости в контрольном эксперименте оказался равен 0 (при всех значениях потери материала) (Рисунок 15). Таким образом, события, когда родственные клетки делят общую судьбу, не являются абсолютно случайными. Кроме того, контрольный эксперимент показывает, что наша компьютерная модель вполне адекватно отражает суть моделируемого процесса.

Более того, если бы способность к репрограммированию не передавалась по наследству (с какой-то долей вероятности), а возникала у сестринских независимо после деления, то вероятность их синхронного репрограммирования определялась бы общей эффективностью репрограммирования. Во всех проведенных экспериментах эффективность репрограммирования, в среднем, не превышала 4%. Это уступает оценке вероятности синхронного репрограммирования родственных клеток, вычисленной путем моделирования (10-60%). Исключение составляет первый эксперимент, в котором эффективность репрограммирования превысила 27%. Поскольку условия проведения всех

экспериментов были идентичны, причина такого разброса не ясна. Известно, тем не менее, что эффективность репрограммирования может широко варьировать для разных линий клеток, полученных от одной особи, и зависит от множества факторов. Высокая эффективность репрограммирования в первом эксперименте «маскирует» вклад наследственной компоненты способности к репрограммированию. Поэтому мы не учитывали данные этого эксперимента, поскольку по ним невозможно определить параметр наследуемости.

Подводя итог проделанной работе, мы заключаем, что в популяции эмбриональных фибробластов мыши присутствуют клетки, имеющие предрасположенность к репрограммированию, которая выражается в синхронном репрограммировании их потомков.

Ключевой вопрос заключается в том, какова природа этой предрасположенности. Согласно первому сценарию (стохастическому), она возникает спонтанно у какой-то доли клеток. Возможно ее возникновение обусловлено стохастичностью, присущей всем молекулярным процессам, происходящим в клетке, начиная от экспрессии генов, заканчивая процессами деградации белков [Elowitz et al., 2002; Ozbudak et al., 2002]. Случаи, когда стохастические процессы определяют судьбу клетки хорошо известны в литературе [Biggins et al., 2015; Jukam, Desplan, 2010; Ohnishi et al., 2013]. Наиболее ярким примером является моноаллельная экспрессия одного из более чем 1000 генов ольфакторных рецепторов мыши, закрепленная формированием специфической трехмерной структуры генома и обеспечивающая специфичность конкретного нейрона к определенному стимулу [Lomvardas et al., 2006; Tsuboi et al., 1999]. Стохастическая экспрессия гена *spineless* на стадии куколки определяет выбор светочувствительности омматидиев плодовой мушки *Drosophila* [Yau, Hardie, 2009], а вариабельность экспрессии гена *Sca-1* в гемопоэтических клетках млекопитающих, приводит к дифференцировке клональных клеток в разные классы клеток крови [Chang et al., 2008]. Долгое время считалось, что программа поддержания плюрипотентности строго репрессируется в дифференцированных клетках, поэтому обнаружить в них экспрессию факторов плюрипотентности невозможно. Однако недавно было показано, что один из главных транскрипционных факторов плюрипотентных клеток, *Oct4*, присутствует в гладкомышечных клетках кровеносных сосудов, выполняя атеропротективную функцию [Cherepanova et al., 2016]. Такая плеiotропность факторов плюрипотентности позволяет предположить их особую роль в регуляции клеточной пластичности. Возможно, в силу стохастичности/регуляции клеточной пластичности в

некоторых клетках уже экспрессируются факторы, способствующие репрограммированию, или наоборот, не экспрессируются факторы, препятствующие ему. И если такое состояние устойчиво, как минимум, несколько делений, мы будем наблюдать частые случаи синхронного репрограммирования родственных клеток. Таким образом наши результаты можно интерпретировать следующим образом: предрасположенность возникает спонтанно, но устойчива в ходе делений и с какой-то долей вероятности передается дочерним клеткам, которые репрограммируются по детерминистической траектории.

Согласно второму сценарию (элитарному), предрасположенность к репрограммированию – характеристика привилегированных клеток. Свойство «элитарности» подразумевает то, что обладающие им клетки имеют некий общий признак, отличающий их от других клеток, не предрасположенных к репрограммированию. Так, менее дифференцированный статус соматических стволовых клеток и большая эффективность репрограммирования сделали их главными кандидатами на роль элитарных клеток (Eminli et al., 2009). Мы проверили гипотезу об элитарности менее дифференцированных клеток и провели эксперимент по репрограммированию *Thy1*-позитивных терминально-дифференцированных фибробластов. К нашему удивлению, доля синхронно-репрограммирующихся дочерних клеток в эксперименте не уменьшилась, что заставляет отвергнуть предположение о преимущественном вкладе соматических стволовых клеток в общее число синхронно-репрограммированных потомков.

Среди других характеристик клеток, обладающих большей вероятностью образовать плюрипотентного потомка, – высокая скорость пролиферации [Hanna et al., 2009]. На это есть, как минимум, две причины. Во-первых, если способность к репрограммированию исключительно стохастическое явление, возникающее независимо в каждой клетке, то с увеличением числа клеток, будет увеличиваться и вероятность репрограммирования хоть одной из них. Во-вторых, быстроделаящиеся клетки могут иметь преимущество за счет необходимых для репрограммирования эпигенетических преобразований, происходящих в ходе деления.

С помощью моделирования мы оценили вклад клеток с разным количеством потомков в общее число двойных пересечений ДНК-баркодов между чашками. Несмотря на то, что доля быстроделаящихся клеток в исходной клеточной популяции была небольшая, их вклад в число двойных пересечений составил более 50% (Рисунок 16). Клетки с большим количеством потомков имеют больший шанс попасть на разные чашки, и вместе с тем большую вероятность быть выбранными в модели. Таким образом, модель уже отражает ситуацию, когда быстроделаящиеся клетки имеют больший шанс репрограммироваться только за счет количества потомков. И тем не менее, для того, чтобы наблюдать сходимость

в количестве пересечений баркодов между чашками в модели и эксперименте, необходимо задать ненулевой уровень наследуемости (как минимум 10%).

Для анализа альтернативной гипотезы, мы присвоили больший уровень наследуемости быстроделющимся клеткам, и получили даже большую сходимость результатов относительно тройных и четверных пересечений (Рисунок 17). Таким образом, предположение о том, что быстроделющиеся клетки являются элитарными, и их потомки репрограммируются синхронно с гораздо большей вероятностью чем потомки других клеток, хорошо согласуется с результатами моделирования. Стоит отметить, что существование привилегированных клеток с ультрабыстрым клеточным циклом уже было показано для популяции гемопоэтических клеток. [Guo et al., 2014]. Эти привилегированные клетки вносят подавляющий вклад в долю всех событий репрограммирования (до 99%). При этом, все потомки привилегированных клеток становятся плюрипотентными синхронно (после пятого деления), то есть по детерминистической траектории. Однако вопрос о присутствии подобных клеток в других клеточных популяциях все еще открыт.

Выводы

1. Создана и охарактеризована библиотека ДНК-баркодированных лентивирусных векторов, позволяющая эффективно маркировать клеточную популяцию, включающую до 200 000 эмбриональных фибробластов мыши.
2. Разработана компьютерная модель, позволяющая оценить вероятность синхронного репрограммирования дочерних клеток.
3. Согласно нашей модели, вероятность синхронного репрограммирования дочерних клеток составляет около 10%. Это значение превышает вероятность их случайного репрограммирования за время наблюдения, что свидетельствует в пользу передачи предрасположенности к репрограммированию при делении для первичных эмбриональных фибробластов мыши.
4. Показано, что исключение из эксперимента Thy1-негативных фибробластов (15-20%), представляющих собой менее дифференцированные клетки, не приводит к уменьшению частоты синхронного репрограммирования дочерних клеток.
5. С помощью компьютерного моделирования установлено, что быстроделяющиеся клетки (оценка по количеству дочерних клеток при 2-х суточном наблюдении), при меньшем содержании в исходной популяции фибробластов, являются основным источником репрограммированных клеток.
6. Быстроделяющиеся клетки являются возможными кандидатами на роль «элитарных» клеток. Присвоение им большего уровня наследуемости способности к репрограммированию в компьютерной модели удовлетворяет количеству пересечений ДНК-баркодов в эксперименте.

Благодарности

Автор выражает благодарность своим научным руководителям – О.Л. Серову и Н.Р. Баттулину, которые помогали автору ценными советами в ходе выполнения дипломной и диссертационной работы. Автор благодарит всех сотрудников лаборатории генетики развития ИЦиГ СО РАН за доброе отношение и помощь на всех этапах работы. Автор выражает особую признательность В.С. Фишману, который помогал с биоинформатической обработкой и компьютерным моделированием, а также участвовал в обсуждении полученных данных. Искреннюю благодарность автор приносит рецензентам С.В. Кулемзину и С.А. Лашину за ценные критические замечания и советы по написанию диссертации. Благодарность за неустанную поддержку автор выражает своим друзьям и семье.

Список литературы

1. Aasen T. и др. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. // Nat. Biotechnol. 2008. Т. 26. № 11. С. 1276–1284.
2. Aguilo F. и др. Coordination of m6A mRNA Methylation and Gene Transcription by ZFP217 Regulates Pluripotency and Reprogramming // Cell Stem Cell. 2015. Т. 17. № 6. С. 689–704.
3. Anokye-Danso F. и др. Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency // Cell Stem Cell. 2011. Т. 8. № 4. С. 376–388.
4. Aoi T. и др. Generation of pluripotent stem cells from adult mouse liver and stomach cells. // Science. 2008. Т. 321. № 5889. С. 699–702.
5. Apostolou E., Hochedlinger K. Chromatin dynamics during cellular reprogramming // Nature. 2013. Т. 502. № 7472. С. 462–471.
6. Araki R. и др. Conversion of ancestral fibroblasts to induced pluripotent stem cells // Stem Cells. 2010. Т. 28. № 2. С. 213–220.
7. Aulicino F. и др. Temporal perturbation of the Wnt signaling pathway in the control of cell reprogramming is modulated by TCF1 // Stem Cell Reports. 2014. Т. 2. № 5. С. 707–720.
8. Banito A. и др. Senescence impairs successful reprogramming to pluripotent stem cells // Genes Dev. 2009. Т. 23. № 18. С. 2134–2139.
9. BARTA T. и др. Inhibition of miR-145 Enhances Reprogramming of Human Dermal Fibroblasts to Induced Pluripotent Stem Cells // Stem Cells. 2016. Т. 34. С. 246–251.
10. Baumann K. Stem cells: Multiple routes to pluripotency // Nat. Rev. Genet. 2015. Т. 16. № FEBRUARY 2015. С. 67–67.
11. Beard C. и др. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells // Genesis. 2006. Т. 44. № 1. С. 23–28.
12. Ben-David U., Nissenbaum J., Benvenisty N. XNew balance in pluripotency: Reprogramming with lineage specifiers // Cell. 2013. Т. 153. № 5.
13. Benevento M. и др. Proteome adaptation in cell reprogramming proceeds via distinct transcriptional networks // Nat. Commun. 2014. Т. 5. С. 5613.
14. Biggins J.S. и др. Towards understanding the roles of position and geometry on cell fate

decisions during preimplantation development // *Semin. Cell Dev. Biol.* 2015. T. 47-48. C. 74–79.

15. Bock C. и др. Reference maps of human es and ips cell variation enable high-throughput characterization of pluripotent cell lines // *Cell*. 2011. T. 144. № 3. C. 439–452.

16. Boiani M. и др. Oct4 distribution and level in mouse clones: Consequences for pluripotency // *Genes Dev.* 2002. T. 16. № 10. C. 1209–1219.

17. Brambrink T. и др. Sequential Expression of Pluripotency Markers during Direct Reprogramming of Mouse Somatic Cells // *Cell Stem Cell*. 2008. T. 2. № 2. C. 151–159.

18. Brosh R. и др. p53 Counteracts reprogramming by inhibiting mesenchymal-to-epithelial transition // *Cell Death Differ.* 2013. T. 20. № 2. C. 312–320.

19. Brumbaugh J., Hochedlinger K. Removing reprogramming roadblocks: Mbd3 depletion allows deterministic iPSC generation // *Cell Stem Cell*. 2013. T. 13. № 4. C. 379–381.

20. Buckley S.M. и др. Regulation of pluripotency and cellular reprogramming by the ubiquitin-proteasome system // *Cell Stem Cell*. 2012. T. 11. № 6. C. 783–798.

21. Buganim Y. и др. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase // *Cell*. 2012. T. 150. № 6. C. 1209–1222.

22. Buganim Y., Faddah D.A., Jaenisch R. Mechanisms and models of somatic cell reprogramming. // *Nat. Rev. Genet.* 2013. T. 14. № 6. C. 427–39.

23. Carey B.W. и др. Single-gene transgenic mouse strains for reprogramming adult somatic cells. // *Nat. Methods*. 2010. T. 7. № 1. C. 56–9.

24. Carey B.W. и др. Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells // *Cell Stem Cell*. 2011. T. 9. № 6. C. 588–598.

25. Chang H.H. и др. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells // *Nature*. 2008. T. 453. № 7194. C. 544–547.

26. Chantzoura E. и др. Reprogramming Roadblocks Are System Dependent // *Stem Cell Reports*. 2015. T. 5. № 3. C. 350–364.

27. Chen J. и др. BMPs functionally replace Klf4 and support efficient reprogramming of mouse fibroblasts by Oct4 alone // *Cell Res.* 2011. T. 21. № 1. C. 205–212.

28. Chen J. и др. H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs // Nat. Genet. 2012a. Т. 45. № 1. С. 34–42.
29. Chen M. и др. Promotion of the induction of cell pluripotency through metabolic remodeling by thyroid hormone triiodothyronine-activated PI3K/AKT signal pathway // Biomaterials. 2012b. Т. 33. № 22. С. 5514–5523.
30. Cherepanova O.A. и др. Activation of the pluripotency factor OCT4 in smooth muscle cells is atheroprotective // Nat. Med. 2016. Т. 22. № 6. С. 657–665.
31. Cheung A.M.S. и др. Analysis of the clonal growth and differentiation dynamics of primitive barcoded human cord blood cells in NSG mice // Blood. 2013. Т. 122. № 18. С. 3129–3137.
32. Chin M.H. и др. Induced Pluripotent Stem Cells and Embryonic Stem Cells Are Distinguished by Gene Expression Signatures // Cell Stem Cell. 2009. Т. 5. № 1. С. 111–123.
33. Chng Z. и др. SIP1 Mediates Cell-Fate Decisions between Neuroectoderm and Mesendoderm in Human Pluripotent Stem Cells // Cell Stem Cell. 2010. Т. 6. № 1. С. 59–70.
34. Clancy J.L. и др. Small RNA changes en route to distinct cellular states of induced pluripotency // Nat. Commun. 2014. Т. 5. С. 5522.
35. Costa Y. и др. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency // Nature. 2013. Т. 495. № 7441. С. 370–374.
36. Cyranoski D. Stem cells: The black box of reprogramming // Nature. 2014. Т. 516. № 7530. С. 162–164.
37. David L., Polo J.M. Phases of reprogramming // Stem Cell Res. 2014. Т. 12. № 3. С. 754–761.
38. Deng J. и др. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. // Nat. Biotechnol. 2009. Т. 27. № 4. С. 353–60.
39. Dimos J.T. и др. Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. // Science. 2008. Т. 321. № 5893. С. 1218–1221.
40. Doege C.A. и др. Early-stage epigenetic modification during somatic cell reprogramming by Parp1 and Tet2 // Nature. 2012. Т. 488. № 7413. С. 652–655.
41. Downing T.L. и др. Biophysical regulation of epigenetic state and cell reprogramming // Nat. Mater. 2013. Т. 12. № 12. С. 1154–1162.

42. Edel M.J. и др. Rem2 GTPase maintains survival of human embryonic stem cells as well as enhancing reprogramming by regulating p53 and cyclin D1 // *Genes Dev.* 2010. Т. 24. № 6. С. 561–573.
43. Elowitz M.B. и др. Stochastic gene expression in a single cell // *Science* (80-.). 2002. Т. 297. № 5584. С. 1183–1186.
44. Eminli S. и др. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. // *Stem Cells.* 2008. Т. 26. № 10. С. 2467–2474.
45. Eminli S. и др. Differentiation stage determines potential of hematopoietic cells for reprogramming into induced pluripotent stem cells. // *Nat. Genet.* 2009. Т. 41. № 9. С. 968–76.
46. Eriksson C., Rustum C., Hallberg E. Dynamic properties of nuclear pore complex proteins in gp210 deficient cells // *FEBS Lett.* 2004. Т. 572. № 1-3. С. 261–265.
47. Esteban M. а и др. Generation of induced pluripotent stem cell lines from Tibetan miniature pig. // *J. Biol. Chem.* 2009. Т. 284. № 26. С. 17634–17640.
48. Esteban M.A. и др. Vitamin C Enhances the Generation of Mouse and Human Induced Pluripotent Stem Cells // *Cell Stem Cell.* 2010. Т. 6. № 1. С. 71–79.
49. Evans M.J., Kaufman M.H. Establishment in culture of pluripotential cells from mouse embryos. // *Nature.* 1981. Т. 292. С. 154–156.
50. Feng B. и др. Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. // *Nat. Cell Biol.* 2009. Т. 11. № 2. С. 197–203.
51. Fidalgo M. и др. Zfp281 mediates Nanog autorepression through recruitment of the NuRD complex and inhibits somatic cell reprogramming // *Proc. Natl. Acad. Sci.* 2012. Т. 109. № 40. С. 16202–16207.
52. Folmes C.D.L. и др. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming // *Cell Metab.* 2011. Т. 14. № 2. С. 264–271.
53. Fusaki N. и др. Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. // *Proc. Jpn. Acad. Ser. B. Phys. Biol. Sci.* 2009. Т. 85. № 8. С. 348–62.
54. Gao Y. и др. Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming // *Cell Stem Cell.* 2013. Т. 12. № 4. С. 453–469.

55. Goldenberger D. и др. A simple «universal» DNA extraction procedure using SDS and proteinase K is compatible with direct PCR amplification // *Genome Res.* 1995. Т. 4. № 6. С. 368–370.
56. Golipour A. и др. A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network // *Cell Stem Cell.* 2012. Т. 11. № 6. С. 769–782.
57. Greer Card D.A. и др. Oct4/Sox2-Regulated miR-302 Targets Cyclin D1 in Human Embryonic Stem Cells // *Mol. Cell. Biol.* 2008. Т. 28. № 20. С. 6426–6438.
58. Gregory P.A. и др. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1 // *Nat. Cell Biol.* 2008. Т. 10. № 5. С. 593–601.
59. Guenther M.G. и др. Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells // *Cell Stem Cell.* 2010. Т. 7. № 2. С. 249–257.
60. Guo S. и др. Nonstochastic reprogramming from a privileged somatic cell state // *Cell.* 2014. Т. 156. № 4. С. 649–662.
61. Gurdon J., Uehlinger V. “Fertile” Intestine Nuclei // *Nature.* 1966. Т. 210. С. 1240–1241.
62. Haase A. и др. Generation of Induced Pluripotent Stem Cells from Human Cord Blood // *Cell Stem Cell.* 2009. Т. 5. № 4. С. 434–441.
63. Haenebalcke L. и др. The ROSA26-iPSC Mouse: A Conditional, Inducible, and Exchangeable Resource for Studying Cellular (De)Differentiation // *Cell Rep.* 2013. Т. 3. № 2. С. 335–341.
64. Hamasaki M. и др. Pathogenic mutation of ALK2 inhibits induced pluripotent stem cell reprogramming and maintenance: Mechanisms of reprogramming and strategy for drug identification // *Stem Cells.* 2012. Т. 30. № 11. С. 2437–2449.
65. Hanna J. и др. Direct Reprogramming of Terminally Differentiated Mature B Lymphocytes to Pluripotency // *Cell.* 2008. Т. 133. № 2. С. 250–264.
66. Hanna J. и др. Direct cell reprogramming is a stochastic process amenable to acceleration. // *Nature.* 2009. Т. 462. № 7273. С. 595–601.
67. Hanna J.H., Saha K., Jaenisch R. Pluripotency and cellular reprogramming: Facts, hypotheses, unresolved issues // *Cell.* 2010. Т. 143. № 4. С. 508–525.
68. Hansson J. и др. Highly Coordinated Proteome Dynamics during Reprogramming of Somatic Cells to Pluripotency // *Cell Rep.* 2012. Т. 2. № 6. С. 1579–1592.

69. Hasegawa K. и др. Comparison of reprogramming efficiency between transduction of reprogramming factors, cell-cell fusion, and cytoplasm fusion // *Stem Cells*. 2010. Т. 28. № 8. С. 1338–1348.
70. Hasegawa Y. и др. CC chemokine ligand 2 and leukemia inhibitory factor cooperatively promote pluripotency in mouse induced pluripotent cells // *Stem Cells*. 2011. Т. 29. № 8. С. 1196–1205.
71. Hasegawa Y. и др. CCL2 enhances pluripotency of human induced pluripotent stem cells by activating hypoxia related genes. // *Sci. Rep.* 2014. Т. 4. С. 5228.
72. Hayashi Y. и др. BMP-SMAD-ID promotes reprogramming to pluripotency by inhibiting p16/INK4A-dependent senescence // *Proc. Natl. Acad. Sci.* 2016. Т. 113. № 46. С. 13057–13062.
73. Heiden M.G. Vander, Cantley L.C., Thompson C.B. Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation // *Science* (80-.). 2009. Т. 324. № 5930. С. 1029–1033.
74. Ho R. и др. Stage-Specific Regulation of Reprogramming to Induced Pluripotent Stem Cells by Wnt Signaling and T Cell Factor Proteins // *Cell Rep.* 2013. Т. 3. № 6. С. 2113–2126.
75. Hochedlinger K. и др. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues // *Cell*. 2005. Т. 121. № 3. С. 465–477.
76. Hockemeyer D. и др. A Drug-Inducible System for Direct Reprogramming of Human Somatic Cells to Pluripotency // *Cell Stem Cell*. 2008. Т. 3. № 3. С. 346–353.
77. Hong H. и др. Suppression of induced pluripotent stem cell generation by the p53–p21 pathway // *Nature*. 2009. Т. 460. № 7259. С. 1132–1135.
78. Hou P. и др. Pluripotent Stem Cells Induced from Mouse Somatic Cells by Small-Molecule Compounds // *Science* (80-.). 2013. Т. 341. № 6146. С. 651–654.
79. Hu X. и др. Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming // *Cell Stem Cell*. 2014. Т. 14. № 4. С. 512–522.
80. Huang K. и др. Dynamically reorganized chromatin is the key for the reprogramming of somatic cells to pluripotent cells // *Sci. Rep.* 2015. Т. 5. С. 17691.
81. Huangfu D. и др. Induction of pluripotent stem cells by defined factors is greatly improved

- by small-molecule compounds. // *Nat. Biotechnol.* 2008. T. 26. № 7. С. 795–7.
82. Hussein S.M.I. и др. Corrigendum: Genome-wide characterization of the routes to pluripotency // *Nature*. 2015. T. 523. № 7562. С. 626–626.
83. Ichida J.K. и др. A Small-Molecule Inhibitor of Tgf- β Signaling Replaces Sox2 in Reprogramming by Inducing Nanog // *Cell Stem Cell*. 2009. T. 5. № 5. С. 491–503.
84. Ivanova N. и др. Dissecting self-renewal in stem cells with RNA interference // *Nature*. 2006. T. 442. № 7102. С. 533–538.
85. Jaenisch R., Young R. Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming // *Cell*. 2008. T. 132. № 4. С. 567–582.
86. Jia J. и др. Regulation of pluripotency and self-renewal of ESCs through epigenetic-threshold modulation and mRNA pruning // *Cell*. 2012. T. 151. № 3. С. 576–589.
87. Jiao J. и др. Promoting reprogramming by FGF2 reveals that the extracellular matrix is a barrier for reprogramming fibroblasts to pluripotency // *Stem Cells*. 2013. T. 31. № 4. С. 729–740.
88. Judson R.L. и др. Embryonic stem cell-specific microRNAs promote induced pluripotency // *Nat. Biotechnol.* 2009. T. 27. № 5. С. 459–461.
89. Jukam D., Desplan C. Binary fate decisions in differentiating neurons // *Curr. Opin. Neurobiol.* 2010. T. 20. № 1. С. 6–13.
90. Kaji K. и др. Virus-free induction of pluripotency and subsequent excision of reprogramming factors. // *Nature*. 2009. T. 458. № 7239. С. 771–5.
91. Kang L. и др. iPS Cells Can Support Full-Term Development of Tetraploid Blastocyst-Complemented Embryos // *Cell Stem Cell*. 2009. T. 5. № 2. С. 135–138.
92. Kawamura T. и др. Linking the p53 tumour suppressor pathway to somatic cell reprogramming // *Nature*. 2009. T. 460. № 7259. С. 1140–1144.
93. Kim D. и др. Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. // *Cell Stem Cell*. 2009. T. 4. № 6. С. 472–476.
94. Kim J.B. и др. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. // *Nature*. 2008. T. 454. № 7204. С. 646–50.
95. Kim K. и др. Donor cell type can influence the epigenome and differentiation potential of

- human induced pluripotent stem cells // *Nat. Biotechnol.* 2011. T. 29. № 12. C. 1117–1119.
96. Kim S. и др. KLF4 N-terminal variance modulates induced reprogramming to pluripotency // *Stem Cell Reports*. 2015. T. 4. № 4. C. 727–743.
97. Kitada M., Wakao S., Dezawa M. Muse cells and induced pluripotent stem cell: Implication of the elite model // *Cell. Mol. Life Sci.* 2012. T. 69. № 22. C. 3739–3750.
98. Kohli R.M., Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation // *Nature*. 2013. T. 502. № 7472. C. 472–479.
99. Kopp J.L. и др. Small Increases in the Level of Sox2 Trigger the Differentiation of Mouse Embryonic Stem Cells // *Stem Cells*. 2008. T. 26. № 4. C. 903–911.
100. Korpai M., Kang Y. The emerging role of miR-200 family of MicroRNAs in epithelial-mesenchymal transition and cancer metastasis // *RNA Biol.* 2008. T. 5. № 3. C. 115–119.
101. Lee D.-S. и др. An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator // *Nat. Commun.* 2014. T. 5. C. 5619.
102. Leonardo T.R. и др. The functions of microRNAs in pluripotency and reprogramming // *Nat. Cell Biol.* 2012. T. 14. № 11. C. 1114–1121.
103. Li H. и др. The Ink4/Arf locus is a barrier for iPS cell reprogramming // *Nature*. 2009a. T. 460. № 7259. C. 1136–1139.
104. Li H. и др. RNA Helicase DDX5 Inhibits Reprogramming to Pluripotency by miRNA-Based Repression of RYBP and its PRC1-Dependent and -Independent Functions // *Cell Stem Cell*. 2017. T. 20. № 4. C. 462–477.e6.
105. Li R. и др. A mesenchymal-to-Epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts // *Cell Stem Cell*. 2010. T. 7. № 1. C. 51–63.
106. Li W. и др. Generation of Rat and Human Induced Pluripotent Stem Cells by Combining Genetic Reprogramming and Chemical Inhibitors (DOI:10.1016/j.stem.2008.11.014) // *Cell Stem Cell*. 2009b. T. 4. № 4. C. 370.
107. Li Z. и др. Small RNA-mediated regulation of iPS cell generation // *EMBO J.* 2011. T. 30. № 5. C. 823–834.
108. Li Z., Rana T.M. A kinase inhibitor screen identifies small-molecule enhancers of reprogramming and iPS cell generation // *Nat. Commun.* 2012. T. 3. C. 1085.

109. Liao B. и др. MicroRNA cluster 302-367 enhances somatic cell reprogramming by accelerating a mesenchymal-to-epithelial transition // J. Biol. Chem. 2011. Т. 286. № 19. С. 17359–17364.
110. Lin C.-H. и др. Myc-regulated microRNAs attenuate embryonic stem cell differentiation // EMBO J. 2009a. Т. 28. № 20. С. 3157–3170.
111. Lin S.-L. и др. Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state // RNA. 2008. Т. 14. № 10. С. 2115–2124.
112. Lin T. и др. A chemical platform for improved induction of human iPSCs // Nat. Methods. 2009b. Т. 6. № 11. С. 805–808.
113. Lister R. и др. Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. // Nature. 2011. Т. 471. № 7336. С. 68–73.
114. Liu H. и др. Generation of Induced Pluripotent Stem Cells from Adult Rhesus Monkey Fibroblasts // Cell Stem Cell. 2008. Т. 3. № 6. С. 587–590.
115. Loh K.M., Lim B. A precarious balance: Pluripotency factors as lineage specifiers // Cell Stem Cell. 2011. Т. 8. № 4. С. 363–369.
116. Lomvardas S. и др. Interchromosomal Interactions and Olfactory Receptor Choice // Cell. 2006. Т. 126. № 2. С. 403–413.
117. Lu R. и др. Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding // Nat Biotechnol. 2011. Т. 29. № 10. С. 928–933.
118. Lu R., Yang A., Jin Y. Dual functions of T-box 3 (Tbx3) in the control of self-renewal and extraembryonic endoderm differentiation in mouse embryonic stem cells // J. Biol. Chem. 2011. Т. 286. № 10. С. 8425–8436.
119. Lujan E. и др. Early reprogramming regulators identified by prospective isolation and mass cytometry // Nature. 2015. Т. 521. № 7552. С. 352–356.
120. Lüningschrör P. и др. MicroRNAs in pluripotency, reprogramming and cell fate induction // Biochim. Biophys. Acta - Mol. Cell Res. 2013. Т. 1833. № 8. С. 1894–1903.
121. Luo M. и др. NuRD blocks reprogramming of mouse somatic cells into Pluripotent stem cells // Stem Cells. 2013. Т. 31. № 7. С. 1278–1286.
122. Maekawa M. и др. Direct reprogramming of somatic cells is promoted by maternal

transcription factor Glis1 // *Nature*. 2011. T. 474. № 7350. C. 225–229.

123. Maherali N. и др. A high-efficiency system for the generation and study of human induced pluripotent stem cells // *Cell Stem Cell*. 2008. T. 3. № 3. C. 340–345.

124. Maherali N., Hochedlinger K. Tgf β Signal Inhibition Cooperates in the Induction of iPSCs and Replaces Sox2 and cMyc // *Curr. Biol*. 2009. T. 19. № 20. C. 1718–1723.

125. Makino S. и др. Cardiomyocytes can be generated from marrow stromal cells in vitro. // *J. Clin. Invest*. 1999. T. 103. № 5. C. 697–705.

126. Mali P. и др. Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes // *Stem Cells*. 2010. T. 28. № 4. C. 713–720.

127. Mansour A.A. и др. The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming // *Nature*. 2012. T. 488. № 7411. C. 409–413.

128. Marion R.M. и др. Telomeres Acquire Embryonic Stem Cell Characteristics in Induced Pluripotent Stem Cells // *Cell Stem Cell*. 2009. T. 4. № 2. C. 141–154.

129. Marión R.M. и др. A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity // *Nature*. 2009. T. 460. № 7259. C. 1149–1153.

130. Markoulaki S. и др. Transgenic mice with defined combinations of drug-inducible reprogramming factors. // *Nat. Biotechnol*. 2009. T. 27. № 2. C. 169–71.

131. Marson A. и др. Connecting microRNA Genes to the Core Transcriptional Regulatory Circuitry of Embryonic Stem Cells // *Cell*. 2008. T. 134. № 3. C. 521–533.

132. Martin G.R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. // *Proc. Natl. Acad. Sci. U. S. A*. 1981. T. 78. № 12. C. 7634–7638.

133. Matveeva N.M. и др. In vitro and in vivo study of pluripotency in intraspecific hybrid cells obtained by fusion of murine embryonic stem cells with splenocytes. // *Mol. Reprod. Dev*. 1998. T. 50. № 2. C. 128–38.

134. Melton C., Judson R.L., Blueloch R. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells // *Nature*. 2010. T. 463. № 7281. C. 621–626.

135. Mikkelsen T.S. и др. Dissecting direct reprogramming through integrative genomic analysis // *Nature*. 2008. T. 454. № 7200. C. 49–55.

136. Miyoshi N. и др. Reprogramming of mouse and human cells to pluripotency using mature microRNAs // *Cell Stem Cell*. 2011. Т. 8. № 6. С. 633–638.
137. Montserrat N. и др. Reprogramming of human fibroblasts to pluripotency with lineage specifiers // *Cell Stem Cell*. 2013. Т. 13. № 3. С. 341–350.
138. Mu X. и др. The Histone Acetyltransferase MOF Promotes Induces Generation of Pluripotent Stem Cells // *Cell. Reprogram.* 2015. Т. 17. № 4. С. 259–267.
139. Naik S.H., Schumacher T.N., Perle L. Cellular barcoding: A technical appraisal // *Exp. Hematol.* 2014. Т. 42. № 8. С. 598–608.
140. Nefzger C.M., Polo J.M. DEAD-Box RNA Binding Protein DDX5: Not a Black-Box during Reprogramming // *Cell Stem Cell*. 2017. Т. 20. № 4. С. 419–420.
141. Newman A.M., Cooper J.B. Lab-specific gene expression signatures in pluripotent stem cells // *Cell Stem Cell*. 2010. Т. 7. № 2. С. 258–262.
142. Nguyen L. V. и др. Clonal analysis via barcoding reveals diverse growth and differentiation of transplanted mouse and human mammary stem cells // *Cell Stem Cell*. 2014. Т. 14. № 2. С. 253–263.
143. Nie Z. и др. c-Myc Is a Universal Amplifier of Expressed Genes in Lymphocytes and Embryonic Stem Cells // *Cell*. 2012. Т. 151. № 1. С. 68–79.
144. Niwa H., Miyazaki J., Smith A.G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. // *Nat. Genet.* 2000. Т. 24. № 4. С. 372–376.
145. O'Malley J. и др. High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. // *Nature*. 2013. Т. 499. № 7456. С. 88–91.
146. Ohi Y. и др. Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells // *Nat Cell Biol.* 2011. Т. 13. № 5. С. 541–549.
147. Ohnishi Y. и др. Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages // *Nat. Cell Biol.* 2013. Т. 16. № 1. С. 27–37.
148. Okita K. и др. Generation of mouse induced pluripotent stem cells without viral vectors. // *Science*. 2008. Т. 322. № 5903. С. 949–53.
149. Onder T.T. и др. Chromatin-modifying enzymes as modulators of reprogramming // *Nature*. 2012. Т. 483. № 7391. С. 598–602.

150. Oswald J. и др. Mesenchymal stem cells can be differentiated into endothelial cells in vitro. // *Stem Cells*. 2004. Т. 22. № 3. С. 377–384.
151. Oyagi S. и др. Therapeutic effect of transplanting HGF-treated bone marrow mesenchymal cells into CCl₄-injured rats // *J. Hepatol.* 2006. Т. 44. № 4. С. 742–748.
152. Ozbudak E.M. и др. Regulation of noise in the expression of a single gene // *Nat. Genet.* 2002. Т. 31. № 1. С. 69–73.
153. Ozmadenci D. и др. Netrin-1 regulates somatic cell reprogramming and pluripotency maintenance // *Nat. Commun.* 2015. Т. 6. С. 7398.
154. Papapetrou E.P. и др. Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. // *Proc. Natl. Acad. Sci. U. S. A.* 2009. Т. 106. № 31. С. 12759–64.
155. Papp B., Plath K. Epigenetics of reprogramming to induced pluripotency // *Cell*. 2013. Т. 152. № 6. С. 1324–1343.
156. Park I.-H. и др. Reprogramming of human somatic cells to pluripotency with defined factors. // *Nature*. 2008. Т. 451. № 7175. С. 141–6.
157. Pfaff N. и др. miRNA screening reveals a new miRNA family stimulating iPS cell generation via regulation of Meox2 // *EMBO Rep.* 2011. Т. 12. № 11. С. 1153–1159.
158. Polo J.M. и др. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. // *Nat. Biotechnol.* 2010. Т. 28. № 8. С. 848–55.
159. Polo J.M. и др. A molecular roadmap of reprogramming somatic cells into iPS cells // *Cell*. 2012. Т. 151. № 7. С. 1617–1632.
160. Popowski M. и др. Bright/Arid3A acts as a barrier to somatic cell reprogramming through direct regulation of Oct4, Sox2, and Nanog // *Stem Cell Reports*. 2014. Т. 2. № 1. С. 26–35.
161. Pour M. и др. Epigenetic predisposition to reprogramming fates in somatic cells. // *EMBO Rep.* 2015. Т. 16. № 3. С. 370–8.
162. Prigione A. и др. The senescence-related mitochondrial/oxidative stress pathway is repressed in human induced pluripotent stem cells // *Stem Cells*. 2010. Т. 28. № 4. С. 721–733.
163. Qin H. и др. Transcriptional analysis of pluripotency reveals the hippo pathway as a barrier to reprogramming // *Hum. Mol. Genet.* 2012. Т. 21. № 9. С. 2054–2067.

164. Qin H. и др. Systematic identification of barriers to human Ipsc generation // Cell. 2014. Т. 158. № 2. С. 449–461.
165. Rahl P.B. и др. C-Myc regulates transcriptional pause release // Cell. 2010. Т. 141. № 3. С. 432–445.
166. Rais Y. и др. Deterministic direct reprogramming of somatic cells to pluripotency // Nature. 2013. Т. 502. № 7469. С. 65–70.
167. Rasmussen M.A. и др. Transient p53 suppression increases reprogramming of human fibroblasts without affecting apoptosis and DNA damage // Stem Cell Reports. 2014. Т. 3. № 3. С. 404–413.
168. Redmer T. и др. E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming // EMBO Rep. 2011. Т. 12. № 7. С. 720–726.
169. Sakurai K. и др. Kinome-wide functional analysis highlights the role of cytoskeletal remodeling in somatic cell reprogramming // Cell Stem Cell. 2014. Т. 14. № 4. С. 523–534.
170. Samavarchi-Tehrani P. и др. Functional genomics reveals a BMP-Driven mesenchymal-to-Epithelial transition in the initiation of somatic cell reprogramming // Cell Stem Cell. 2010. Т. 7. № 1. С. 64–77.
171. Santos R.L. Dos и др. MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner // Cell Stem Cell. 2014. Т. 15. № 1. С. 102–110.
172. Shu J. и др. XInduction of pluripotency in mouse somatic cells with lineage specifiers // Cell. 2013. Т. 153. № 5.
173. Shu J., Deng H. Lineage Specifiers: New Players in the Induction of Pluripotency // Genomics, Proteomics Bioinforma. 2013. Т. 11. № 5. С. 259–263.
174. Silva J. и др. Promotion of reprogramming to ground state pluripotency by signal inhibition // PLoS Biol. 2008. Т. 6. № 10. С. 2237–2247.
175. Silva J., Smith A. Capturing Pluripotency // Cell. 2008. Т. 132. № 4. С. 532–536.
176. Singhal N. и др. Chromatin-remodeling components of the baf complex facilitate reprogramming // Cell. 2010. Т. 141. № 6. С. 943–955.
177. Singhal P.K. и др. Mouse embryonic fibroblasts exhibit extensive developmental and phenotypic diversity. // Proc. Natl. Acad. Sci. U. S. A. 2016. Т. 113. № 1. С. 122–7.

178. Smith Z.D. и др. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. // *Nat. Biotechnol.* 2010. Т. 28. № 5. С. 521–526.
179. Sommer C.A. и др. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. // *Stem Cells.* 2009. Т. 27. № 3. С. 543–9.
180. Son M.J. и др. Nicotinamide overcomes pluripotency deficits and reprogramming barriers // *Stem Cells.* 2013. Т. 31. № 6. С. 1121–1135.
181. Soufi A., Donahue G., Zaret K.S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome // *Cell.* 2012. Т. 151. № 5. С. 994–1004.
182. Sridharan R. и др. Role of the Murine Reprogramming Factors in the Induction of Pluripotency // *Cell.* 2009. Т. 136. № 2. С. 364–377.
183. Stadtfeld M. и др. Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse // *Cell Stem Cell.* 2008. Т. 2. № 3. С. 230–240.
184. Stadtfeld M. и др. A reprogrammable mouse strain from gene-targeted embryonic stem cells. // *Nat. Methods.* 2010. Т. 7. № 1. С. 53–5.
185. Stadtfeld M., Brennand K., Hochedlinger K. Reprogramming of pancreatic beta cells into induced pluripotent stem cells. // *Curr. Biol.* 2008. Т. 18. № 12. С. 890–4.
186. Stadtfeld M., Hochedlinger K. Induced pluripotency: History, mechanisms, and applications // *Genes Dev.* 2010. Т. 24. № 20. С. 2239–2263.
187. Stefano B. Di и др. C/EBP α poises B cells for rapid reprogramming into induced pluripotent stem cells // *Nature.* 2013. Т. 506. № 7487. С. 235–239.
188. Subramanyam D. и др. Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells // *Nat Biotech.* 2011. Т. 29. № 5. С. 443–448.
189. Szabó P.E. и др. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells // *Mech. Dev.* 2002. Т. 115. № 1-2. С. 157–160.
190. Tada M. и др. Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells // *Curr. Biol.* 2001. Т. 11. № 19. С. 1553–1558.
191. Takahashi K. и др. Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors // *Cell.* 2007. Т. 131. № 5. С. 861–872.

192. Takahashi K. и др. Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm // *Nat. Commun.* 2014. Т. 5.
193. Takahashi K., Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors // *Cell.* 2006. Т. 126. № 4. С. 663–676.
194. Takahashi K., Yamanaka S. A decade of transcription factor-mediated reprogramming to pluripotency // *Nat. Rev. Mol. Cell Biol.* 2016. Т. 17. № 3. С. 183–193.
195. Tanabe K. и др. Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts // *Proc. Natl. Acad. Sci.* 2013. Т. 110. № 30. С. 12172–12179.
196. Tang Y. и др. Jak/Stat3 signaling promotes somatic cell reprogramming by epigenetic regulation // *Stem Cells.* 2012. Т. 30. № 12. С. 2645–2656.
197. Tang Y., Tian X. (Cindy). JAK-STAT3 and somatic cell reprogramming // *JAK-STAT.* 2013. Т. 2. № 4. С. e24935.
198. Tchieu J. и др. Female human iPSCs retain an inactive X chromosome // *Cell Stem Cell.* 2010. Т. 7. № 3. С. 329–342.
199. Teo A.K.K. и др. Pluripotency factors regulate definitive endoderm specification through eomesodermin // *Genes Dev.* 2011. Т. 25. № 3. С. 238–250.
200. Thiery J.P. и др. Epithelial-Mesenchymal Transitions in Development and Disease // *Cell.* 2009. Т. 139. № 5. С. 871–890.
201. Thiery J.P., Sleeman J.P. Complex networks orchestrate epithelial–mesenchymal transitions // *Nat. Rev. Mol. Cell Biol.* 2006. Т. 7. № 2. С. 131–142.
202. Thomson J.A. Embryonic Stem Cell Lines Derived from Human Blastocysts // *Science* (80-.). 1998. Т. 282. № 5391. С. 1145–1147.
203. Tonge P.D. и др. Divergent reprogramming routes lead to alternative stem-cell states // *Nature.* 2014. Т. 516. № 7530. С. 192–197.
204. Tsuboi A. и др. Olfactory neurons expressing closely linked and homologous odorant receptor genes tend to project their axons to neighboring glomeruli on the olfactory bulb // *J. Neurosci.* 1999. Т. 19. № 19. С. 8409–8418.
205. Unternaehrer J.J. и др. The epithelial-mesenchymal transition factor SNAIL paradoxically enhances reprogramming // *Stem Cell Reports.* 2014. Т. 3. № 5. С. 691–698.

206. Ustek D. и др. A genome-wide analysis of lentivector integration sites using targeted sequence capture and next generation sequencing technology // *Infect. Genet. Evol.* 2012. T. 12. № 7. С. 1349–1354.
207. Utikal J. и др. Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. // *Nature*. 2009. T. 460. № 7259. С. 1145–8.
208. Wakao S. и др. Multilineage-differentiating stress-enduring (Muse) cells are a primary source of induced pluripotent stem cells in human fibroblasts. // *Proc. Natl. Acad. Sci. U. S. A.* 2011. T. 108. № 24. С. 9875–9880.
209. Wang G. и др. Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation // *Proc. Natl. Acad. Sci.* 2013a. T. 110. № 8. С. 2858–2863.
210. Wang S. и др. Transient Activation of Autophagy via Sox2-Mediated Suppression of mTOR Is an Important Early Step in Reprogramming to Pluripotency // *Cell Stem Cell*. 2013b. T. 13. № 5. С. 617–625.
211. Wang T. и др. The histone demethylases Jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner // *Cell Stem Cell*. 2011. T. 9. № 6. С. 575–587.
212. Wang Y. и др. Embryonic stem cell–specific microRNAs regulate the G1-S transition and promote rapid proliferation // *Nat. Genet.* 2008. T. 40. № 12. С. 1478–1483.
213. Wang Z. и др. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells // *Cell Stem Cell*. 2012. T. 10. № 4. С. 440–454.
214. Warren L. и др. Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA // *Cell Stem Cell*. 2010. T. 7. № 5. С. 618–630.
215. Wernig M. и др. In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state // *Nature*. 2007. T. 448. № 7151. С. 318–324.
216. Wernig M. и др. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types // *Nat. Biotechnol.* 2008a. T. 26. № 8. С. 916–924.
217. Wernig M. и др. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. // *Nat. Biotechnol.* 2008b. T. 26. № 8. С. 916–24.
218. Wienken M. и др. MDM2 Associates with Polycomb Repressor Complex 2 and Enhances Stemness-Promoting Chromatin Modifications Independent of p53 // *Mol. Cell*. 2016. T. 61. №

1. С. 68–83.

219. Wienken M., Moll U.M., Dobbelstein M. Mdm2 as a chromatin modifier // 74 j J. Mol. Cell Biol. 2017. Т. 9. № 1. С. 74–80.

220. Wiklund E.D. и др. Coordinated epigenetic repression of the miR-200 family and miR-205 in invasive bladder cancer // Int. J. Cancer. 2011. Т. 128. № 6. С. 1327–1334.

221. Wilmut I. и др. Somatic cell nuclear transfer. // Nature. 2002. Т. 419. № October. С. 583–586.

222. Woltjen K. и др. piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. // Nature. 2009. Т. 458. № 7239. С. 766–70.

223. Woltjen K., Stanford W.L. Inhibition of Tgf- β Signaling Improves Mouse Fibroblast Reprogramming // Cell Stem Cell. 2009. Т. 5. № 5. С. 457–458.

224. Worringer K.A. и др. The let-7/LIN-41 pathway regulates reprogramming to human induced pluripotent stem cells by controlling expression of prodifferentiation genes // Cell Stem Cell. 2014. Т. 14. № 1. С. 40–52.

225. Xu J., Lamouille S., Derynck R. TGF-beta-induced epithelial to mesenchymal transition. // Cell Res. 2009. Т. 19. № 2. С. 156–72.

226. Yamanaka S. Elite and stochastic models for induced pluripotent stem cell generation // Nature. 2009. Т. 460. № 7251. С. 49–52.

227. Yamanaka S. Induced pluripotent stem cells: Past, present, and future // Cell Stem Cell. 2012. Т. 10. № 6. С. 678–684.

228. Yamanaka S., Blau H.M. Nuclear reprogramming to a pluripotent state by three approaches. // Nature. 2010. Т. 465. № 7299. С. 704–12.

229. Yang C.S., Chang K.Y., Rana T.M. Genome-wide Functional Analysis Reveals Factors Needed at the Transition Steps of Induced Reprogramming // Cell Rep. 2014. Т. 8. № 2. С. 327–337.

230. Yang J. и др. Stat3 activation is limiting for reprogramming to ground state pluripotency // Cell Stem Cell. 2010. Т. 7. № 3. С. 319–328.

231. Yang V.S. и др. Geminin Escapes degradation in G1 of mouse pluripotent cells and mediates the expression of Oct4, Sox2, and Nanog // Curr. Biol. 2011. Т. 21. № 8. С. 692–699.

232. Yau K.W., Hardie R.C. Phototransduction Motifs and Variations // *Cell*. 2009. T. 139. № 2. C. 246–264.
233. Young R.A. Control of the embryonic stem cell state // *Cell*. 2011. T. 144. № 6. C. 940–954.
234. Yu J. и др. Induced pluripotent stem cell lines derived from human somatic cells. // *Science*. 2007. T. 318. № 5858. C. 1917–20.
235. Yu J. и др. Human induced pluripotent stem cells free of vector and transgene sequences. // *Science*. 2009. T. 324. № 5928. C. 797–801.
236. Yu P. и др. FGF2 Sustains NANOG and Switches the Outcome of BMP4 Induced Human Embryonic Stem Cell Differentiation // *Cell Stem Cell*. 2011. T. 8. № 3. C. 326–334.
237. Yu Y. и др. Stimulation of somatic cell reprogramming by ERas-Akt-FoxO1 signaling axis // *Stem Cells*. 2014. T. 32. № 2. C. 349–363.
238. Yusa K. и др. Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. // *Nat. Methods*. 2009. T. 6. № 5. C. 363–369.
239. Zhang J. и др. Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1 // *Nat. Cell Biol.* 2006. T. 8. № 10. C. 1114–1123.
240. Zhang J. и др. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal // *Cell Stem Cell*. 2012. T. 11. № 5. C. 589–595.
241. Zhang P. и др. Regulation of induced pluripotent stem (iPS) cell induction by Wnt/b-catenin signaling // *J. Biol. Chem.* 2014. T. 289. № 13. C. 9221–9232.
242. Zhang X. и др. Gene regulatory networks mediating canonical wnt signal-directed control of pluripotency and differentiation in embryo stem cells // *Stem Cells*. 2013. T. 31. № 12. C. 2667–2679.
243. Zhang Z. и др. PRC2 complexes with JARID2, MTF2, and esPRC2p48 in ES cells to modulate ES cell pluripotency and somatic cell reprogramming // *Stem Cells*. 2011. T. 29. № 2. C. 229–240.
244. Zhao T., Xu Y. p53 and stem cells: new developments and new concerns // *Trends Cell Biol.* 2010. T. 20. № 3. C. 170–175.
245. Zhao W. и др. Jmjd3 inhibits reprogramming by upregulating expression of INK4a/Arf and

targeting PHF20 for ubiquitination // Cell. 2013. T. 152. № 5. С. 1037–1050.

246. Zhao X. и др. iPS cells produce viable mice through tetraploid complementation. // Nature. 2009. T. 461. № 7260. С. 86–90.

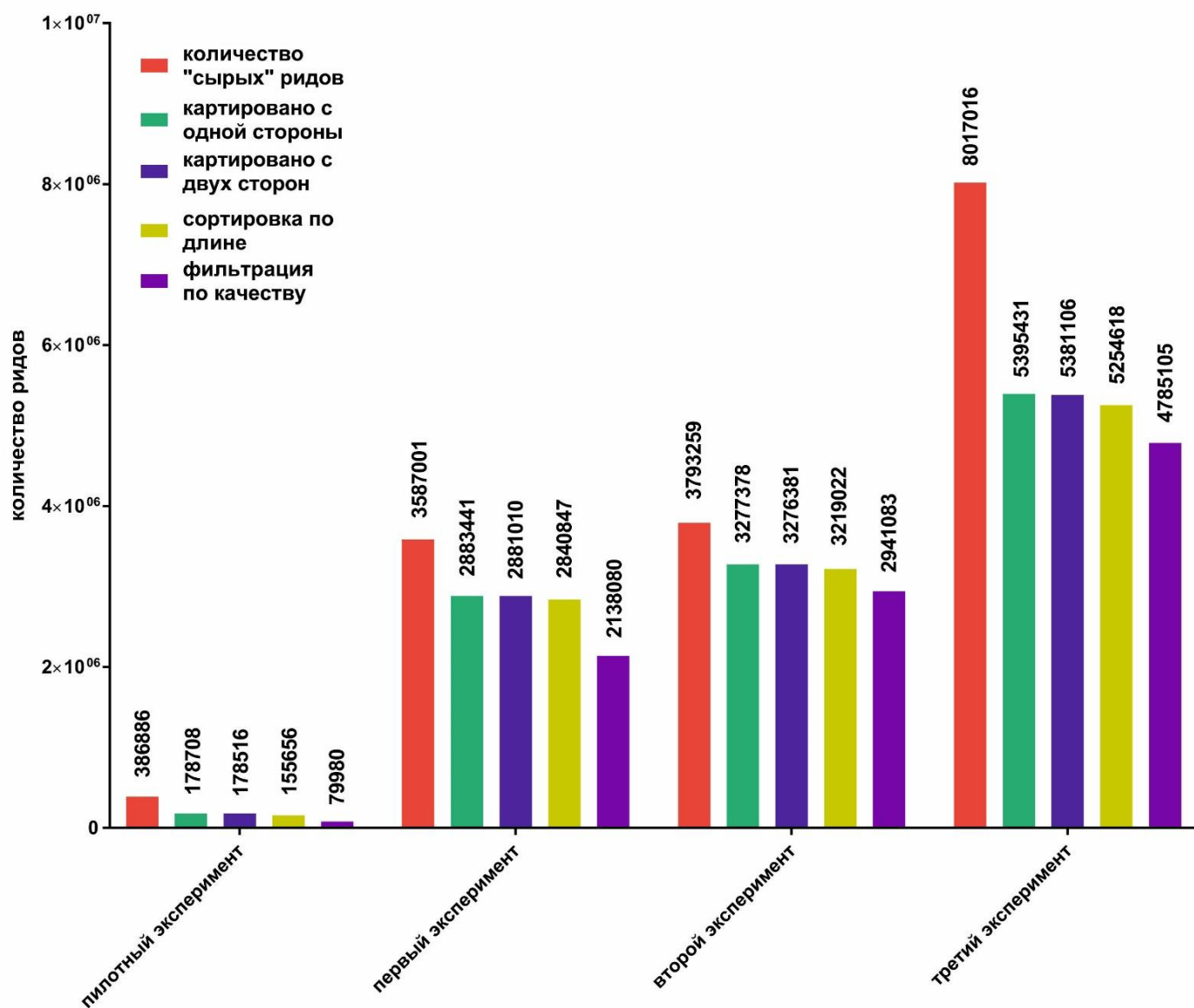
247. Zhao Y. и др. Two Supporting Factors Greatly Improve the Efficiency of Human iPSC Generation // Cell Stem Cell. 2008. T. 3. № 5. С. 475–479.

248. Zhou W., Freed C.R. Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells // Stem Cells. 2009. T. 27. № 11. С. 2667–2674.

249. Zhu S. и др. Recent advances in chemically induced reprogramming // Cell Cycle. 2011. T. 10. № 6. С. 871–872.

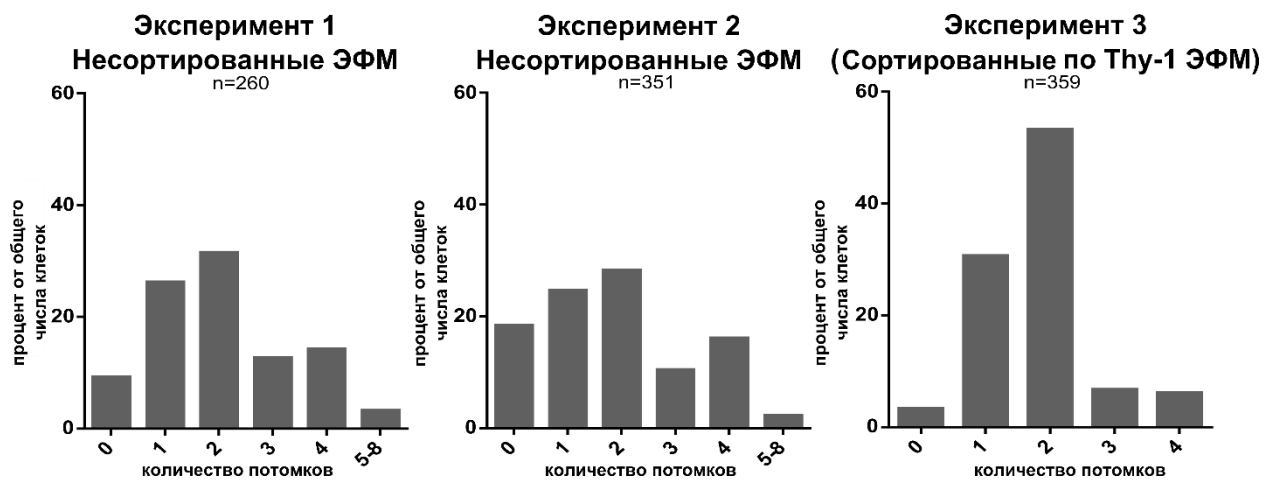
Приложения

Приложение 1



Анализ качества и количества данных массового параллельного секвенирования библиотек ДНК-баркодов на платформе Ion Torrent. «Сырые» риды – общее число ридов в библиотеке; картировано с одной стороны – число ридов, у которых картирован один конец; картировано с двух сторон – число ридов, у которых картированы оба конца; сортированные по длине – число ридов, длина которых лежит в пределах 26-33 нуклеотидов (длина ДНК-баркода – 30 нуклеотидов), отфильтрованные по качеству – количество ридов, в которых 80% позиций имеют значение баллов качества более 20.

Приложение 2



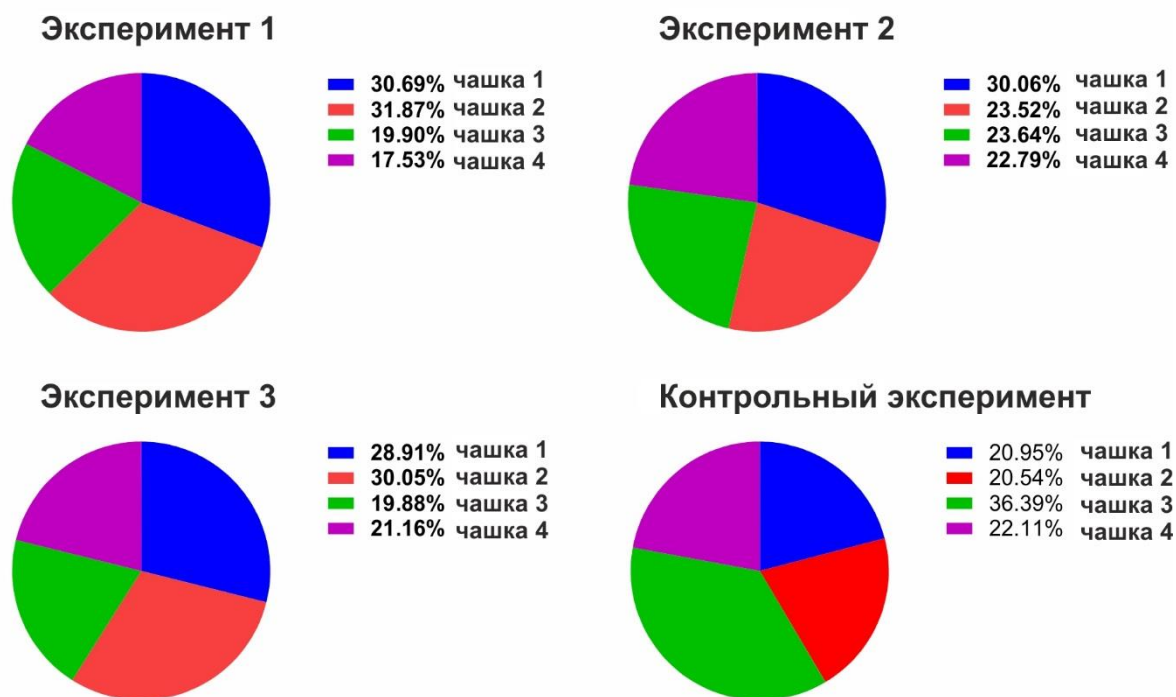
Количество потомков, образованных трансдуцированными клетками в течение двух суток культивирования в приборе CellIQ. Подсчет клеток проводился путем анализа фотографий, сделанных с частотой один кадр в 10 минут.

Приложение 3

Номер эксперимента	Пилотный эксперимент	Эксперимент 1	Эксперимент 2	Эксперимент 3	Контрольный эксперимент
Цель эксперимента	Апробация дизайна эксперимента	Анализ синхронного репрограммирования родственных клеток в гетерогенной (несортированной) популяции ЭФМ	Анализ синхронного репрограммирования родственных клеток в гетерогенной (несортированной) популяции ЭФМ	Анализ синхронного репрограммирования родственных клеток в популяции ЭФМ, позитивных по Thy1 (зрелые фибробласты)	Симуляция отбора репрограммированных клеток случайным образом (т.е. при нулевой наследуемости).
Количество клеток, посаженных под трансдукцию	170 000	230 000	200 000	190 000	190 000
Доля GFP-позитивных клеток в дополнительных экспериментах/MOI	59.1%/0.89	38.7%/0.49	29.1%/0.34	46.9%/0.63	46.9%/0.63
Культивирование в приборе CellIQ с фотографированием каждые 10 минут	Нет	50ч (Приложение 2)	30ч (Приложение 2)	28ч (Приложение 2)	28ч (Приложение 2)
Компьютерное моделирование	Нет	Да (Рисунок 13)	Да (Рисунок 13)	Да (Рисунок 15)	Да (Рисунок 15)

Характеристика проведенных в работе экспериментов.

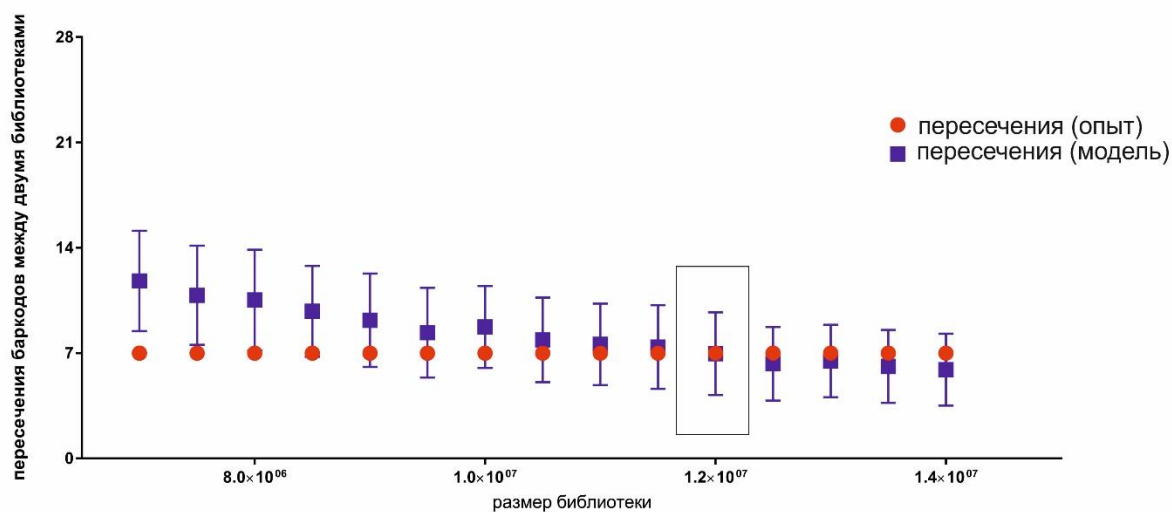
Приложение 4



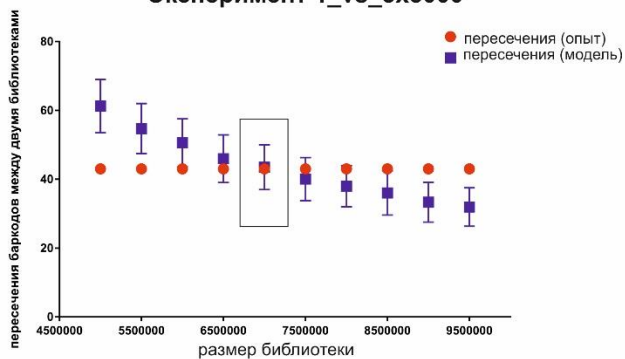
Представленность кластеров (уникальных вариантов) ДНК-баркодов на четырех чашках в каждом эксперименте.

Приложение 5

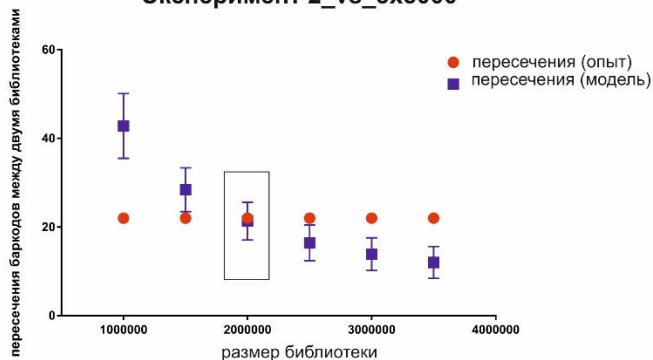
Эксперимент 1_vs_Эксперимент 2



Эксперимент 1_vs_3x5000

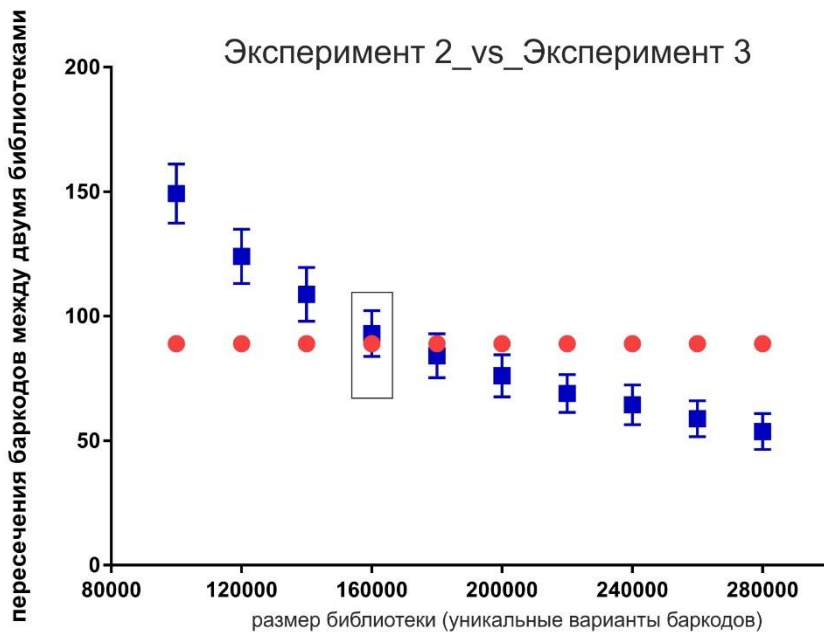
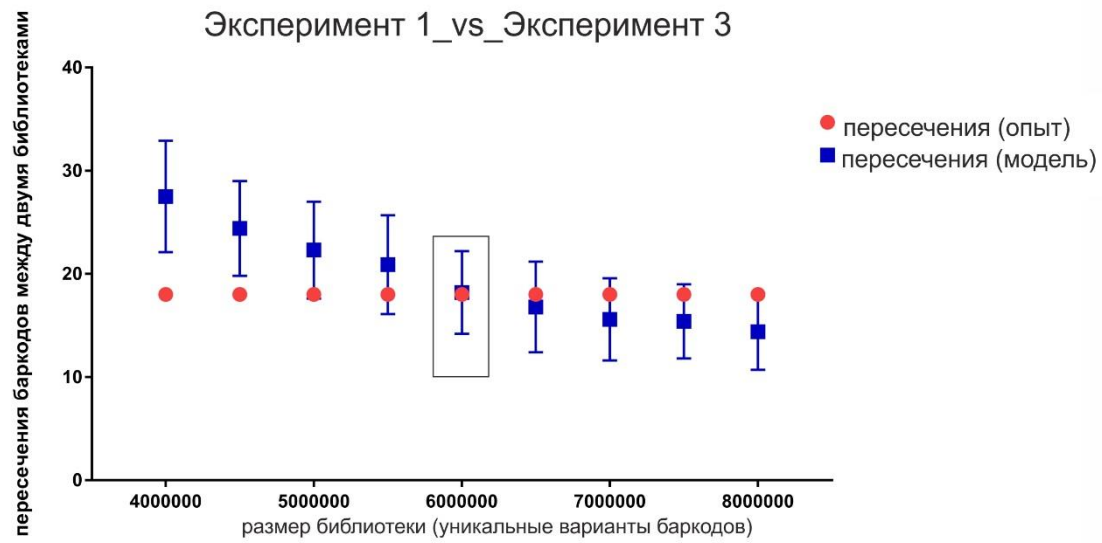


Эксперимент 2_vs_3x5000

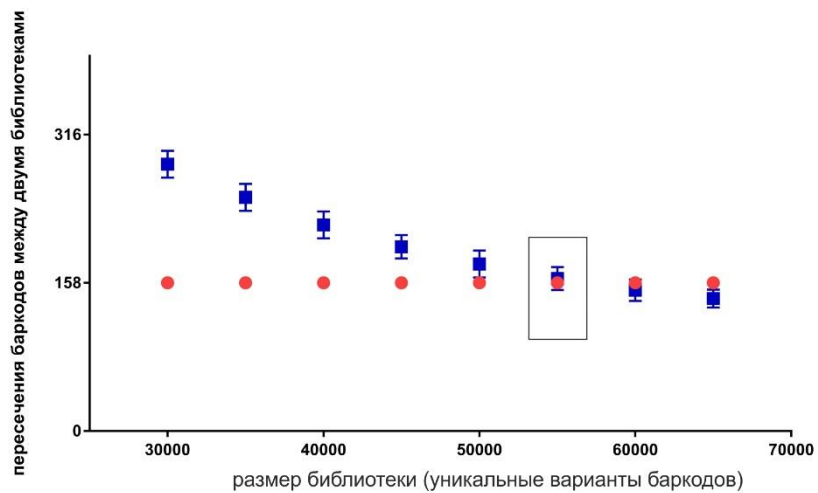


Анализ размера общей библиотеки баркодированных лентивирусов на основании результатов секвенирования трех суб-библиотек: эксперименты 1 и 2, а также дополнительного эксперимента 3x5000. Красными точками обозначены пересечения ДНК-баркодов между разными экспериментами; синими квадратами – ожидаемое среднее количество пересечений при различных размерах общей библиотеки. Черными прямоугольниками выделены наиболее подходящие значения размера библиотеки. Анализ проводился с помощью скриптов на языке Python, написанных Фишманом В.С. На каждое попарное сравнение суб-библиотек было проведено 200 запусков моделирования, «усы» иллюстрируют стандартное отклонение.

Приложение 6



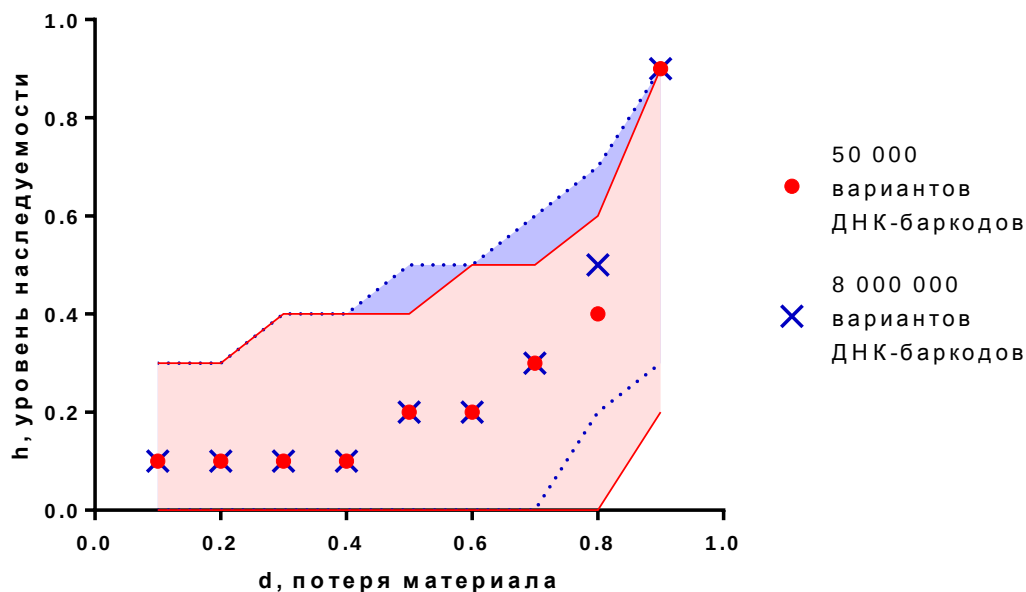
Эксперимент 3 (репрограммирование)_vs_Эксперимент 3 (контроль)



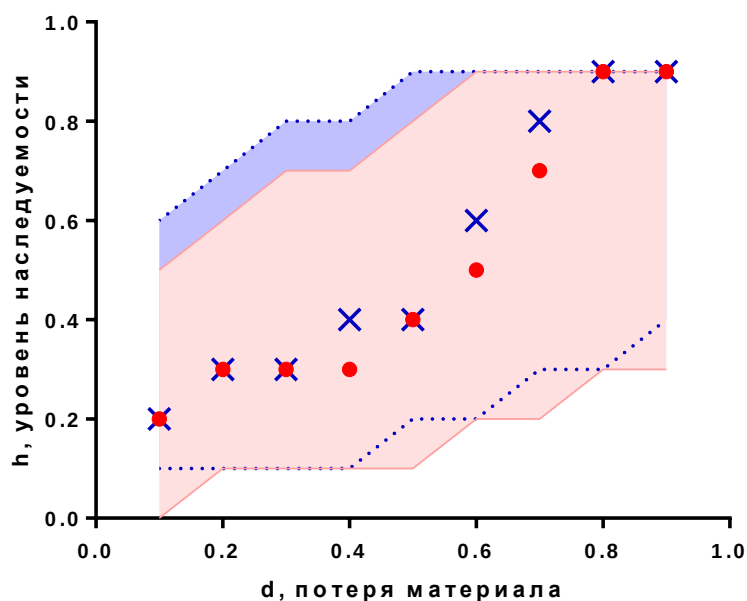
Подпись к приложению 6. Коррекция оценки размера общей библиотеки баркодированных лентивирусов на основании результатов третьего эксперимента, включающего секвенирование ДНК-баркодов репрограммированных Thy+клеток, а также ДНК-баркодов случайно отобранных клеток из контрольного эксперимента. Красными точками обозначены пересечения ДНК-баркодов между разными экспериментами; синими квадратами – ожидаемое среднее количество пересечений при различных размерах общей библиотеки. Черными прямоугольниками выделены наиболее подходящие значения размера библиотеки. Анализ проводился с помощью скриптов на языке Python, написанных Фишманом В.С. На каждое попарное сравнение суб-библиотек было проведено 200 запусков моделирования, «усы» иллюстрируют стандартное отклонение.

Приложение 7

Эксперимент 2



Эксперимент 3 (Thy+ клетки)



Размер библиотеки баркодированных вирусов не оказывает значительного влияния на подобранные с помощью моделирования уровни наследуемости. На графиках показаны результаты компьютерного моделирования второго и третьего экспериментов при разном размере библиотеки: 50000 и 80000000 уникальных вариантов ДНК-баркодов. По оси X отмечена потеря материала; по оси Y –уровень наследуемости, при котором количества пересечений ДНК-баркодов между чашками значимо не отличаются в эксперименте и модели ($p < 0.05$). Заштрихованная область показывает подходящий уровень наследуемости при определенной потере материала. Крестиками выделены *наиболее* подходящий значения. Анализ проводился с помощью непараметрического дисперсионного анализа (критерий Краскела-Уоллиса) в программе GraphPad Prism 6. На каждый эксперимент было проведено по 100 запусков моделирования.