

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА
БИОЛОГИЧЕСКИЙ ФАКУЛЬТЕТ
Кафедра молекулярной биологии

На правах рукописи

Ломов Николай Андреевич

**Изучение механизмов образования транслокаций,
ассоциированных со вторичными лейкозами, вызванными
терапией ингибиторами ДНК-топоизомераз II**

Специальность 1.5.3. — «Молекулярная биология»

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата биологических наук

Научный руководитель:
Михаил Александрович Рубцов
кандидат биологических наук

Москва — 2023

Оглавление

СПИСОК ИСПОЛЬЗУЕМЫХ СОКРАЩЕНИЙ	5
1. ВВЕДЕНИЕ.....	8
АКТУАЛЬНОСТЬ РАБОТЫ И СТЕПЕНЬ РАЗРАБОТАННОСТИ ТЕМЫ	8
ЦЕЛИ И ЗАДАЧИ	10
НОВИЗНА И НАУЧНАЯ ЗНАЧИМОСТЬ.....	11
ЛИЧНЫЙ ВКЛАД АВТОРА.....	11
МЕТОДОЛОГИЯ И МЕТОДЫ ИССЛЕДОВАНИЯ	11
ПОЛОЖЕНИЯ, ВЫНОСИМЫЕ НА ЗАЩИТУ	12
СТЕПЕНЬ ДОСТОВЕРНОСТИ И АПРОБАЦИЯ РЕЗУЛЬТАТОВ.....	12
2. ОБЗОР ЛИТЕРАТУРЫ.....	15
2.1 ОСТРЫЕ МИЕЛОИДНЫЕ ЛЕЙКОЗЫ, ВЫЗВАННЫЕ ТЕРАПИЕЙ.....	16
<i>Лейкозы и хромосомные транслокации</i>	<i>16</i>
<i>Причины развития лейкозов после терапии опухолей.....</i>	<i>17</i>
2.2. ТРАНСЛОКАЦИИ, РЕКУРРЕНТНЫЕ ДЛЯ ВТОРИЧНЫХ ОМЛ, ВЫЗВАННЫХ ТЕРАПИЕЙ ИНГИБИТОРАМИ ДНК-ТОПОИЗОМЕРАЗЫ II.....	18
<i>Ген AML1</i>	<i>18</i>
<i>Ген MLL</i>	<i>21</i>
<i>Развитие лейкозов при перестройках AML1 и MLL</i>	<i>26</i>
2.3. ТОПОИЗОМЕРАЗНЫЕ ЯДЫ – ЭФФЕКТИВНОЕ СРЕДСТВО ПРОТИВООПУХОЛЕВОЙ ТЕРАПИИ .26	
<i>Развитие противоопухолевой химиотерапии</i>	<i>26</i>
<i>ДНК-топоизомеразы и их активность в клетках</i>	<i>28</i>
<i>Топоизомеразы II в клетках человека</i>	<i>30</i>
<i>Топоизомеразные яды</i>	<i>33</i>
<i>Пути удаления топоизомеразы с ДНК</i>	<i>34</i>
2.4. РЕПАРАЦИЯ ДВУЦЕПОЧЕЧНЫХ РАЗРЫВОВ.....	35
<i>Репарация ДЦР, направляемая гомологией</i>	<i>36</i>
<i>Белки HDR.....</i>	<i>38</i>
<i>Негомологичное соединение концов разрыва.....</i>	<i>39</i>
<i>MMEJ и SSA.....</i>	<i>40</i>
<i>Выбор пути репарации</i>	<i>41</i>
<i>Вклад различных путей репарации ДЦР в образование транслокаций</i>	<i>41</i>
2.5. ИЗУЧЕНИЕ ФАКТОРОВ, ВЛИЯЮЩИХ НА ОБРАЗОВАНИЕ ТРАНСЛОКАЦИЙ	44
<i>Хромосомные территории.....</i>	<i>44</i>

<i>Изучение механизмов образования хромосомных транслокаций</i>	45
2.6. СОЗДАНИЕ КЛЕТОЧНЫХ МОДЕЛЕЙ.....	50
<i>Технологии редактирования генома</i>	50
<i>Индукция хромосомных перестроек</i>	53
3. МАТЕРИАЛЫ И МЕТОДЫ	54
3.1. РАБОТА С БАКТЕРИЯМИ	54
3.2. КУЛЬТУРЫ КЛЕТОК ЭУКАРИОТ.....	55
3.3. ТРЕХМЕРНАЯ ФЛУОРЕСЦЕНТНАЯ <i>IN SITU</i> ГИБРИДИЗАЦИЯ (3D FISH)	55
<i>Обработка клеток ингибитором топоизомераз</i>	55
<i>Облучение клеток рентгеновским излучением</i>	56
<i>Приготовление препаратов клеток</i>	56
<i>Конфокальная микроскопия препаратов</i>	57
<i>Компьютерная обработка и анализ полученных изображений</i>	57
<i>Статистическая обработка результатов</i>	59
3.4. МЕТОД 4С	60
<i>Приготовление 3С-библиотек</i>	60
<i>Приготовление 4С-библиотек</i>	61
<i>Анализ данных секвенирования 4С-библиотек</i>	62
3.5. СОЗДАНИЕ КЛЕТОЧНЫХ МОДЕЛЕЙ ОБРАЗОВАНИЯ ТРАНСЛОКАЦИЙ	63
<i>Общая стратегия</i>	63
<i>Подбор РНК-гидов</i>	64
<i>Сборка плазмид с генами гидовых РНК</i>	64
<i>Метод ENIT для проверки гидовых РНК</i>	66
<i>Вставка генов гидовых РНК в плазмиду Puro_Cas9 donor</i>	68
<i>Трансфекция клеток и селекция на антибиотиках</i>	69
3.6. ПРОВЕРКА ЭКСПРЕССИИ ТРАНСГЕНОВ.....	71
<i>Выделение РНК из клеток эукариот с помощью TRIzol reagent</i>	71
<i>Выделение ДНК из клеток с помощью TRIzol reagent</i>	72
<i>Обратная транскрипция</i>	72
<i>ПЦР в реальном времени (количественная ПЦР)</i>	73
3.7. АМПЛИФИКАЦИЯ ОБЛАСТИ ТРАНСЛОКАЦИИ AML1-ETO.....	75
3.8. ОБРАБОТКА КУЛЬТУРЫ КЛЕТОК IAML/ETO ИНГИБИТОРАМИ БЕЛКОВ РЕПАРАЦИИ И ХИМИОТЕРАПЕВТИЧЕСКИМИ ПРЕПАРАТАМИ.....	77
3.9. ИСПОЛЬЗУЕМОЕ ЛАБОРАТОРНОЕ ОБОРУДОВАНИЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ	80

4. РЕЗУЛЬТАТЫ	81
4.1. ИЗУЧЕНИЕ ВЛИЯНИЯ ЭТОПОЗИДА НА АЛЛЕЛИ <i>AML1</i> МЕТОДОМ 3D FISH.....	81
<i>Обработка этопозидом приводит к разрывам гена AML1 и расхождению концов разрыва, в отличие от высоких доз γ-излучения.</i>	<i>84</i>
<i>Разорванные под действием этопозиды аллели чаще оказываются вне хромосомной территории по сравнению с интактными аллелями</i>	<i>85</i>
4.2. 4С-АНАЛИЗ ПРЕДПОЧТИТЕЛЬНЫХ КОНТАКТОВ <i>AML1</i> И <i>MLL</i>	86
4.3. СОЗДАНИЕ КЛЕТОЧНОЙ МОДЕЛИ ОБРАЗОВАНИЯ ТРАНСЛОКАЦИЙ	90
<i>Общая схема получения клеточной культуры.....</i>	<i>92</i>
<i>Получение и проверка генов гидовых РНК.....</i>	<i>94</i>
4.4. ПОЛУЧЕНИЕ КУЛЬТУРЫ КЛЕТОК iAML/ETO	99
4.5. ИНДУКЦИЯ ТРАНСЛОКАЦИИ <i>AML1-ETO</i> В КЛЕТКАХ iAML/ETO	100
<i>Кинетика экспрессии трансгенов в клетках культуры iAML/ETO</i>	<i>101</i>
<i>Кинетика индукции транслокаций в клетках культуры iAML/ETO.....</i>	<i>102</i>
<i>Количественная детекция транслокаций</i>	<i>103</i>
<i>Анализ методом FISH генов AML1 и ETO после индукции экспрессии Cas9</i>	<i>105</i>
4.6. ВЛИЯНИЕ РАЗЛИЧНЫХ ВЕЩЕСТВ НА ЧИСЛО ТРАНСЛОКАЦИЙ	106
5. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ	108
ЗАКЛЮЧЕНИЕ	114
ВЫВОДЫ.....	115
СПИСОК ЛИТЕРАТУРЫ	116
БЛАГОДАРНОСТИ.....	142

Список используемых сокращений

Англоязычные сокращения

- 3C — chromosome conformation capture, метод захвата конформации хроматина
- 4C — chromosome conformation capture followed by sequencing, метод захвата конформации хроматина с последующим секвенированием
- BCR — breakpoint cluster region
- Cas9 — CRISPR-associated protein 9
- CRISPR — clustered regularly interspaced short palindromic repeats
- CSR — переключение класса антител, от class switch recombination
- CTD — C-terminal domain
- DNA-PKcs — DNA-dependent protein kinase
- DSBR — путь гомологической рекомбинации, от double-strand break repair
- E. coli* — *Escherichia coli*
- ENIT — engineered nuclease-induced translocations
- FISH — fluorescent in situ hybridization
- GFP — green fluorescent protein
- H3K4 — лизин-4 гистона H3
- H3K9 — лизин-9 гистона H3
- H3K27 — лизин-27 гистона H3
- HDR — homology-directed repair
- Hi-C — high-throughput chromosome conformation capture
- HR — homologous recombination
- iAML1-ETO — клеточная культура, от inducible *AML1-ETO* translocation
- M2rtTA — reverse tetracycline-controlled transactivator M2
- MLL — mixed-lineage leukemia
- MMEJ — microhomology-mediated end-joining
- mQ — деионизированная вода
- MRN — комплекс белков MRE11-RAD50-NBS1
- NHEJ — non-homologous end joining
- NHR — участки, гомологичные по последовательности белку Nervy дрозофилы
- NTD — N-terminal domain
- P-TEFb — positive transcription elongation factor b
- PAM — protospacer adjacent motive
- PBS — phosphate buffer saline
- PHD — plant homeodomain, белковый домен
- PIK-киназы — фосфатидилинозитол-3-киназа-подобные киназы

RHD — runt-homology domain
RPA — replication protein A
RUNX1 — runt-related transcription factor 1, AML1
SDSA — synthesis-dependent strand annealing, путь гомологической рекомбинации
SET — белковый домен Su(var)3-9, enhancer-of-zeste and trithorax
SHM — соматический гипермутагенез
SSA — single-strand annealing
SSC — saline sodium citrate
SUMO — small ubiquitin-like modifier
t-MN — therapy-related myeloid neoplasms
TAD — трансактивирующий белковый домен
TAE — буфер трис-ацетат-ЭДТА
TALEN — transcription activator- like effector nuclease
TDP2 — tyrosyl DNA phosphodiesterase 2
TMEJ — theta-mediated end joining
Top1 — ДНК-топоизомераза I
Top2 — ДНК-топоизомераза II
Top2cc — topoisomerase II cleaved complex
TOP2 α — топоизомераза 2 α
TOP2 β — топоизомераза 2 β
Ub — убиквитин
v/v — объем, мл / объем раствора, мл
w/v — масса, г / объем раствора, мл
WCP — whole chromosome painting, флуоресцентный гибридизационный зонд
на хромосомную территорию
WHD — winged helix домен
XLF — XRCC4-Like Factor, Cernunnos
XRCC1-6 — X-ray repair cross-complementing proteins 1-6
ZFN — нуклеазы на основе доменов цинковых пальцев

Русскоязычные сокращения

Гр — грей

ДМСО — диметилсульфоксид

дНТФ — дезоксирибонуклеозидтрифосфат

ДЦР — двуцепочечный разрыв ДНК

Зв — зиверт

мкл — микролитр

мкМ — микромоль/л

млн — миллион

млн/мл — миллион клеток в миллилитре

мМ — миллимоль/л

мс — миллисекунда

об/мин — оборотов в минуту

ОЛЛ — острый лимфоидный лейкоз

ОМЛ — острый миелоидный лейкоз

п.н. — пара (пары) нуклеотидов

ПЦР — полимеразная цепная реакция

ТА — транскрипционный активатор

ТАД — топологически-ассоциированный домен хроматина

Топо2 — ДНК-топоизомераза II

Трис — трис(гидроксиметил)аминометан

тыс.п.н. — тысяч пар нуклеотидов

ЭДТА — этилендиаминтетраацетат

1. Введение

Актуальность работы и степень разработанности темы

Ингибиторы ДНК-топоизомераз II получили широкое распространение в химиотерапии опухолей. Их эффективность связана с важной ролью ДНК-топоизомеразы II в делящихся клетках — с помощью этих ферментов клетка разделяет хромосомы после репликации для распределения по дочерним клеткам. Разделение топологически сцепленных между собой хромосом невозможно без образования двуцепочечных разрывов ДНК, через которые должны проходить фрагменты другой хромосомы. ДНК-топоизомераза II (топо2) контролирует этот процесс — она вносит временный двуцепочечный разрыв в один фрагмент хромосомы, пропускает через него другой фрагмент, после чего сшивает разрыв обратно. Ингибиторы, нарушающие возможность топо2 религиговать разрыв ДНК, называются топоизомеразными ядами. В результате действия топо2-ядов в делящихся клетках накапливаются двуцепочечные разрывы (ДЦР), которые приводят к апоптозу таких клеток. Наиболее чувствительны к этим ядам именно опухолевые клетки, и поэтому топоизомеразные яды, например, этопозид, эффективно применяют при лечении практически всех типов опухолей.

Однако терапия с применением топоизомеразных ядов имеет побочный эффект: у 1–6 % пациентов, прошедших такую терапию, в течение 1–3 лет развивается так называемый вторичный лейкоз. Этот лейкоз характеризуется острым течением заболевания, а также крайне неблагоприятным прогнозом. Цитогенетической особенностью при вторичном лейкозе являются определенные рекуррентные хромосомные перестройки, которые встречаются от пациента к пациенту. Считается, что к этим перестройкам приводят ошибки репарации ДЦР, возникающих под действием топо2-ядов в гемопоэтических клетках. Набор этих транслокаций ограничен — перестраивается либо ген *AML1* (*RUNX1*), либо ген *MLL* (*KMT2A*), каждый с одним из нескольких генов-партнеров. В норме гены *AML1* и *MLL* регулируют в гемопоэтических клетках транскрипцию ключевых для пролиферации генов. Образующиеся в результате транслокации слитые *AML1*- и *MLL*-белки приводят к нарушению баланса активирующего и репрессирующего действий *AML1* и *MLL*. Это приводит к изменению эпигенетического паттерна хроматина у делящихся клеток с перестройкой, и среди них может возникнуть опухолевый клон.

Следует отметить, что для других опухолей характерен гораздо более широкий набор хромосомных aberrаций, чем для лейкозов, ассоциированных с терапией топо2-ингибиторами. На данный момент не известно, какие факторы приводят к распространению именно таких транслокаций у пациентов с этими лейкозами.

Изучение механизмов образования транслокаций поначалу привело к выявлению двух основных факторов, определяющих результат перестройки. Первый фактор — колокализация перестраивающихся генов до момента образования в них разрыва. Методами высокопроизводительного секвенирования была показана прямая корреляция между частотой формирования транслокаций и пространственной близостью перестраивающихся локусов. Второй фактор — повышенная подвижность генов, в которых возник разрыв. Показано, что разрывы имеют свойство кластеризоваться в пространстве ядра. Далее было выявлено, что к факторам, определяющим формирование перестройки, относятся также способ и вероятность образования разрывов в локусах перестройки.

Однако было неясно, какие факторы вносят определяющий вклад при формировании транслокаций при терапии топоизомеразными ядами. Ранее в нашей лаборатории уже было показано, что для гена *MLL* характерна повышенная подвижность после обработки клеток этопозидом: в них ген *MLL* оказывался вне хромосомной территории чаще, чем в необработанных клетках. Поэтому в рамках данной диссертационной работы было изучено, как обработка этопозидом влияет уже на ген *AML1* и его локализацию относительно хромосомной территории. Также был проведен анализ пространственных контактов генов *MLL* и *AML1* со всеми участками генома (4C-анализ) в клетках до и после обработки этопозидом.

Изучение механизмов, задействованных в образовании хромосомных транслокаций, ассоциированных с терапией топо2-ядами, затруднено по нескольким причинам. Во-первых, формирование транслокаций — крайне редкое событие, и обнаружение транслокации у пациентов возможно, только когда клетки с транслокацией уже много раз поделились, и сформировался лейкозный клон. Во-вторых, точки разрывов, вызванных ингибированием топо2, могут находиться в довольно широких областях внутри перестраивающихся генов, что ограничивает набор методов, которыми эти транслокации могут быть детектированы. Чаще всего методом обнаружения таких транслокаций выступает флуоресцентная *in situ* гибридизация (FISH), требующая трудоемкого приготовления препаратов с использованием специализированных FISH-проб и конфокальной микроскопии с последующим анализом изображений. Сложность изучения механизмов образования транслокаций как на животных, так и на клетках, подвергнутых воздействию этопозидов, обуславливает необходимость моделирования транслокаций, характерных для вторичных лейкозов.

Стоит отметить, что культуры клеток, содержащие изучаемые транслокации, уже существуют. Это либо клетки, полученные от пациентов, либо клеточные линии, искусственно созданные с помощью технологий редактирования генома. Однако такие клеточные модели позволяют изучать последствия уже сформировавшейся транслокации, но не процесс ее образования.

Моделью для изучения процесса образования транслокаций может выступать клеточная культура, где исследуемые транслокации формировались бы только после воздействия на клетки определенным стимулом. Чтобы детектировать полученные транслокации наиболее простым методом — ПЦР — каждая транслокация должна происходить между локусами хромосом, заданными с точностью до нуклеотида. С помощью такой модели, применяя количественную ПЦР, можно изучать влияние различных соединений на вероятность образования заданной транслокации. Это важно как для исследования молекулярных механизмов, так и для проверки препаратов, используемых в химиотерапии, на предмет их влияния на возникновение транслокаций. Также клеточная модель может выступать тест-системой для поиска соединений, снижающих вероятность развития вторичного лейкоза после терапии с применением ингибиторов топо2.

Цели и задачи

Цель работы — выявить факторы, определяющие набор рекуррентных хромосомных транслокаций, ассоциированных с вторичными лейкозами, вызванными терапией топоизомеразными ядами.

Для достижения цели были поставлены следующие задачи:

1. Подтвердить, что обработка клеток этопозидом влияет на целостность гена *AML1*. Сравнить по этому параметру обработку клеток этопозидом с действием ионизирующего излучения.
2. Изучить распределение аллелей *AML1* относительно границ их хромосомных территорий при действии этопозида.
3. Получить профили пространственных контактов генов *AML1* и *MLL* с остальными участками генома в клетках, обработанных этопозидом, и в необработанных клетках.
4. Создать клеточную культуру, позволяющую изучать механизмы образования индуцируемой транслокации *AML1-ETO*.
5. С помощью полученной клеточной культуры изучить влияние ингибирования некоторых белков системы репарации на образование транслокации *AML1-ETO*.
6. С помощью полученной клеточной культуры *iAML-ETO* провести анализ химиотерапевтических препаратов, используемых в комбинации с ингибиторами ДНК-топоизомеразы II, на предмет их влияния на образование транслокации *AML1-ETO*.

Новизна и научная значимость

Работа посвящена актуальной теме — изучению механизмов возникновения хромосомных транслокаций, ассоциированных с лейкозами. В работе показано, что набор транслокаций, повторяющихся для лейкозов, вызванных терапией топоизомеразными ядами, во многом обусловлен особенностью двуцепочечных разрывов ДНК, которые формируются при ингибировании ДНК-топоизомеразы II. Предложен способ изучения механизмов образования онкогенных перестроек, заключающийся в создании специальной клеточной культуры, где транслокации индуцируются путем запуска экспрессии интегрированного гена нуклеазы Cas9, нацеленной на перестраиваемые локусы. При этом число транслокаций оценивается количественной ПЦР, что позволяет сравнивать частоту образования хромосомных перестроек в разных условиях. С помощью такой клеточной культуры iAML-ETO, моделирующей образование транслокации *AML1-ETO*, показано, что NU7026 — ингибитор белка репарации DNA-PKcs, а также метотрексат — использующийся в химиотерапии ингибитор дигидрофолатредуктазы — повышают число транслокаций. Клеточная культура iAML-ETO может быть использована как для дальнейшего изучения механизмов образования транслокаций, так и для поиска веществ, способных снизить вероятность развития вторичного лейкоза при совместной терапии с применением ингибиторов ДНК-топоизомераз II. В процессе создания клеточной модели был разработан метод проверки инструментов редактирования генома — Direct ENIT.

Личный вклад автора

Вклад соискателя в постановке задач, планировании и выполнении экспериментов, обработке данных, интерпретации результатов был решающим, за исключением разработки программы анализа конфокальных изображений, процедуры секвенирования 4C-библиотек и первичного анализа данных секвенирования. Соискатель писал текст всех статей, в которых указан первым автором, и занимался их подачей в журналы. Также соискатель непосредственно участвовал в закупке реактивов, расходных материалов и оборудования для исследования и занимался организацией работы лаборатории.

Методология и методы исследования

Работа выполнена с применением современных молекулярно-биологических методов. Изучение образования разрывов гена *AML1* и его расположения относительно хромосомной территории проводилось на клетках культуры Jurkat методом 3D FISH с последующим анализом изображений компьютерными программами. Анализ контактов генов-партнеров по транслокациям проводился с помощью метода 4C (chromosome conformation capture

followed by sequencing), включающего приготовление 4С-библиотек, высокопроизводительное секвенирование и биоинформатический анализ данных. Созданная клеточная культура iAML-ETO содержит в геноме элементы системы CRISPR/Cas под контролем системы Tet-ON, элементы этой системы были интегрированы с помощью технологий редактирования генома на основе нуклеаз TALEN. Эффективность гидовых РНК, направляющих Cas9 на локусы *AML1* и *ETO*, проверялась методом Direct ENIT. Сборка плазмидных конструкций осуществлялась методами молекулярного клонирования, в том числе бесшовного. Для оценки числа транслокаций была отработана система количественной ПЦР с TaqMan-зондами.

Положения, выносимые на защиту

1. Обработка клеток Jurkat этопозидом приводит к возникновению двуцепочечных разрывов ДНК внутри гена *AML1* и расхождению концов разрыва, детектируемому с помощью конфокальной микроскопии.
2. Под действием этопозиды аллели *AML1* локализуются снаружи границ своей хромосомной территории чаще, чем в необработанных клетках Jurkat.
3. Гены-партнеры по транслокациям, рекуррентным для лейкозов, вызванных топоизомеразными ядами, не демонстрируют пространственной близости по сравнению с другими генами ни в необработанных клетках Jurkat, ни после их обработки этопозидом.
4. Интегрированная в геном клеток индуцируемая система CRISPR/Cas позволяет воспроизводимо моделировать образование хромосомных транслокаций в культуре клеток и количественно детектировать их методом ПЦР.
5. С помощью полученной клеточной культуры iAML-ETO возможен поиск веществ, влияющих на возникновение транслокаций.

Степень достоверности и апробация результатов

Результаты работы были получены с использованием современных методик и на современном оборудовании. Результаты статистически достоверны и воспроизводимы. Основные положения и выводы исследования изложены в семи статьях в рецензируемых научных изданиях, индексируемых в базах данных WoS, Scopus, RSCI и рекомендованных для защиты в диссертационном совете МГУ, а также в патенте. Основные результаты работы были представлены на 10 конференциях.

Публикации

1. *Shmakova A. and ***Lomov N.**, Viushkov V., Tsfasman T., Kozhevnikova Y., Sokolova D., Pokrovsky V., Syrkina M., Germini D., Rubtsov M., Vassetzky Y. Cell models with inducible oncogenic translocations allow to evaluate the potential of drugs to favor secondary translocations // *Cancer communications*. — 2023. — 154-158. — 43(1). IF = 10,4 (Web of Science).
*авторы внесли равный вклад
2. **Lomov N.A.**, Viushkov V.S., Ulianov S.V., Gavrilov A.A., Alexeyevsky D.A., Artemov A.A., Razin S.V., Rubtsov M.A. Recurrent Translocations in Topoisomerase Inhibitor-Related Leukemia Are Determined by the Features of DNA Breaks Rather Than by the Proximity of the Translocating Genes // *Int J Mol Sci*. — 2022, — 23(17). — 9824. IF = 5,9 (Web of Science).
3. **Lomov N.**, Zerkalenkova E., Lebedeva S., Viushkov V., Rubtsov M. Cytogenetic and molecular genetic methods of chromosomal translocation detection with reference to the *KMT2A/MLL* gene // *Critical Reviews in Clinical Laboratory Sciences*. — 2021. — 180-206. — 58(3). IF = 8,5 (Web of Science).
4. **Lomov N.A.**, Viushkov V.S., Zamalutdinov A.V., Sboeva M.D., Rubtsov M.A. Direct ENIT: An easy and reliable tool for gRNA efficacy verification by tracking induced chromosomal translocation // *METHODSX*. — 2020. — Vol. 7. — P. 101-104. IF 2020 = 2,2 (Web of Science).
5. **Lomov N.A.**, Viushkov V.S., Petrenko A.P., Syrkina M.S., Rubtsov M.A. Methods for the evaluation of efficiency of CRISPR/Cas genome editing // *Molecular Biology*. — 2019. — Vol. 53, no. 6. — P. 982-997. IF = 1,4 (Web of Science). Статья доступна на русском языке в журнале Молекулярная биология. — 2019. — Т. 53, № 6. — С. 982-997.
6. Germini D., Bou Saada Y., Tsfasman T., Osina K., Robin C., **Lomov N.**, Rubtsov M., Sjakste N., Lipinski M., Vassetzky Y. A one-step PCR-based assay to evaluate the efficiency and precision of genomic DNA-editing tools // *Molecular Therapy - Methods & Clinical Development*. — 2017. — Vol. 5. — P. 43-50. IF = 6,7 (Web of Science).
7. **Lomov N.A.**, Borunova V.V., Rubtsov M.A. CRISPR/Cas9 technology for targeted genome editing // *Biopolymers and Cell*. — 2015. — Vol. 31, no. 4. — P. 243-248. IF = 0,34 (Scopus).

Патент

Ломов Н.А., Вьюшков В.С., Васецкий Е.С., Рубцов М.А. Тест-система для поиска препаратов, снижающих риск возникновения вторичных лейкозов // Патент номер #2018132409 от 30 апреля 2020.

Выступления на конференциях

Ломов Н.А., Вьюшков В.С. «Клеточная тест-система для сравнения химиотерапевтических препаратов на предмет риска лейкозогенных транслокаций». Объединенный VI конгресс гематологов России и III конгресс трансфузиологов России, Москва, 21–23 апреля 2022.

Lomov N., Rubtsov M., Viushkov V. “Cell model of inducible AML1-ETO translocation”. 26th Wilhelm Bernhard Workshop on the Cell Nucleus, Дижон, Франция, 20–24 мая 2019.

Canoy R., Germini D., Vassetzky Y., Andre F., **Lomov N.** “Factors that Affect Chromosomal Translocations in Cells”. Scientific Meetings “Mechanisms and Consequences of Chromosomal Translocations in Cancer”. Институт Пастера, Париж, Франция, 20–21 мая 2019

Ломов Н.А., Вьюшков В.С., Рубцов М.А. «Клеточная модель индуцируемой транслокации AML1-ETO». II Объединенный научный форум, включающий VI Съезд физиологов СНГ, VI Съезд биохимиков России и IX Российский симпозиум «Белки и пептиды», Дагомыс, 1–6 октября 2019

Ломов Н.А., Вьюшков В.С., Рубцов М.А. «Создание клеточных моделей для изучения лейкозогенных транслокаций». Всероссийский научно-практический конгресс с международным участием «Орфанные болезни», Москва, 31 мая — 2 июня 2018

Lomov N., Rubtsov M., Viushkov V. “Wide-scale observation of chromatin dynamics by computer processing of 3D-FISH images allows to study the mechanisms of leukemogenic translocations”. 25th Wilhelm Bernhard Workshop on the Cell Nucleus, Нижний Новгород, 19–22 июня 2017.

Ломов Н.А., Вьюшков В.С. «Пространственная мобильность концов двуцепочечных разрывов ДНК в условиях ингибирования белков-участников репарации двунитевых разрывов». XXIV Международная научная конференция студентов, аспирантов и молодых ученых «Ломоносов – 2017», МГУ имени М.В. Ломоносова, 20 апреля 2017

Lomov N., Rubtsov M., Viushkov V., Alexeevsky D. “Wide-scale observation of chromatin dynamics by computer processing of 3D-FISH images allows to study the mechanisms of leukemogenic translocations”. 1-ая Школа ADFLIM, ФИЦ Биотехнологии РАН, г. Москва, 12–14 декабря 2016.

Ломов Н.А., Рубцов М.А., Разин С.В., Яровая О.В. «Пространственная мобильность поврежденных генов AML1 и MLL как предпосылка возникновения вторичных лейкозов», III Конгресс Гематологов России, Москва, 14–16 апреля 2016

Lomov N., Rubtsov M., Alexeevsky D. “Broken proto-oncogenes AML1 and MLL leave the inherent chromosome territories in human lymphoid cells treated with DNA topoisomerase II poison etoposide”. VIII International meeting “From Molecular to Cellular Events in Human Pathologies”, Тбилиси, Грузия, 18–20 октября 2015.

2. Обзор литературы

Первый раздел обзора литературы содержит общую информацию о лейкозах и их цитогенетических маркерах — хромосомных транслокациях, а также описывает феномен вторичных лейкозов, вызванных предшествующей химиотерапией.

Второй раздел посвящен механизмам развития лейкоза при перестройках генов *AML1* и *MLL*. Приведена таблица с этими транслокациями, которые ассоциированы с лейкозами, вызванными терапией топоизомеразными ядами. Описаны нормальные функции генов *AML1* и *MLL* в организме, а также строение и принципы работы нормальных белков *AML1* и *MLL*. Данное описание объясняет молекулярные механизмы развития лейкоза при экспрессии *AML1*- и *MLL*-слитых белков в клетках.

В третьем разделе обзора обсуждаются топоизомеразные яды и механизм их действия. Описана история развития химиотерапии опухолей, и в частности, открытие ингибиторов ДНК-топоизомеразы II как одного из эффективнейших средств противоопухолевой терапии. Рассмотрены принципы функционирования ДНК-топоизомераз II типа и их значение для делящихся клеток, а также опасность, заключенная в работе этих ферментов для клеток. Описываемые факторы определяют эффективность ингибиторов ДНК-топоизомеразы II в химиотерапии. Один из классов этих ингибиторов, к которому принадлежат применяемые в химиотерапии соединения, — топоизомеразные яды. Описаны молекулярные механизмы действия этих соединений и пути процессинга ковалентных комплексов топоизомеразы с концами ДНК. Специальные пути процессинга концов разрывов необходимы для их превращения в стандартные двуцепочечные разрывы ДНК (ДЦР), которые уже могут быть зашиты клеточными системами репарации.

Механизмы репарации двуцепочечных разрывов ДНК в клетке рассматриваются в четвертом разделе. Это репарация двуцепочечных разрывов, направляемая гомологией, канонический и альтернативные пути негомологичного соединения концов разрыва. Рассматриваются механизмы, влияющие на выбор того или иного пути репарации ДЦР.

В пятом разделе рассматриваются работы, посвященные механизмам, определяющим формирование хромосомных транслокаций.

Шестой раздел посвящен методам, облегчающим получение новых клеточных моделей — технологиям редактирования генома.

2.1 Острые миелоидные лейкозы, вызванные терапией

Лейкозы и хромосомные транслокации

Гемобластозы, или онкогематологические заболевания, — это опухолевые заболевания кроветворной или лимфатической ткани. Среди них выделяют 2 основные группы — лейкозы и лимфомы, соответственно [1].

Лейкозы, или лейкемии, — опухолевые заболевания кроветворной ткани, которые развиваются, когда недифференцированные (бластные) клетки теряют способность к образованию зрелых форм и накапливаются в костном мозге. Считается, что лейкозы возникают в результате трансформации лишь одной исходной клетки — стволовой лейкозной клетки [2]. В результате её деления получается неопластический клон бластных клеток, который вытесняет нормальные элементы кроветворения из костного мозга, а также способен к накоплению во внутренних органах. Возникает недостаток зрелых клеток в периферической крови, а инфильтрация лейкозными клетками внутренних органов нарушает их работу. Симптомы такого заболевания — увеличение печени, селезенки и периферических лимфоузлов, геморрагический и анемический синдромы, иммунодефицит.

Лейкозы делятся на острые и хронические в зависимости от доли бластных клеток в костном мозге: при значении от 20% и выше лейкоз определяют как острый. В зависимости от типа бластных клеток, выделяют миелоидный (миелобластный) и лимфобластный острые лейкозы (ОМЛ, в английской литературе — acute myeloid leukemia, AML; и ОЛЛ, в английской литературе — acute lymphoblastic leukemia, ALL).

Считается, что к онкогенной трансформации бластных клеток приводят определенные хромосомные транслокации. Когда на границе такой перестройки соединяются фрагменты разных генов — эти гены называют генами-партнерами по транслокации — образуются слитые гены, продукты которых называют слитыми белками [3].

Хромосомная перестройка сама по себе может быть достаточным событием для развития лейкоза. В пользу этого говорят данные высокопроизводительного секвенирования: лейкозы несут очень мало соматических мутаций по сравнению с другими видами злокачественных новообразований [4], [5]. Конкретная транслокация, которой характеризуются клетки лейкоза, — в частности, какие гены являются партнерами по транслокации, — влияет на развитие заболевания, прогноз и выбор лечения [6]–[8]. Важность гена, участвующего в перестройке, отражена в последних изменениях в медицинской классификации лейкозов. Так, название «острый миелоидный лейкоз с перестройкой *KMT2A*» заменяет прежнее «острый миелоидный лейкоз с перестройкой *t(9;11)(p22;q23)*» [9]. Определение, какая именно хромосомная транслокация ассоциирована с данным лейкозом, является первостепенной задачей при уточнении диагноза у пациентов [10].

Причины развития лейкозов после терапии опухолей

Хромосомные транслокации возникают при ошибках репарации двуцепочечных разрывов ДНК (ДЦР), когда сшиваются вместе концы разных разрывов. Поэтому причиной хромосомных транслокаций могут быть любые процессы, приводящие к образованию разрывов в ДНК. Среди естественных причин возникновения большого количества ДЦР в клетке — работа ферментов во время таких процессов, как V(D)J-рекомбинация, переключение класса антител (CSR) и соматический гипермутагенез (SHM) при созревании лимфоцитов. С активностью этих ферментов связывают развитие многих лимфом, например, лимфомы Беркитта, которая сопровождается транслокацией t(8:14) между локусом *IGH* и геном *MYC* [11]–[13].

Другой причиной образования ДЦР в клетках может быть действие цитотоксических агентов или ионизирующего излучения, что происходит при радио- и химиотерапии опухолей и аутоиммунных заболеваний. Накопление ДЦР в раковых клетках приводит к остановке деления и развитию в них апоптоза [14]. Однако побочным эффектом терапии является развитие вторичных опухолей. Таким образом, с появлением различных видов терапии опухолей возник и особый класс заболеваний — обусловленные терапией миелоидные неоплазии (therapy-related myeloid neoplasms, t-MN, или myeloid neoplasms post cytotoxic therapy, MN-pCT), также называемые вторичными неоплазиями.

Обусловленные терапией миелоидные неоплазии составляют 10–20% от всех случаев миелоидных лейкозов. Они развиваются в 0,8–6,3% случаях в течение 20 лет после предшествующей терапии, с медианным временем 3–5 лет [15]. Выживаемость при таких вторичных заболеваниях измеряется месяцами, а не годами, что делает их одними из самых агрессивных видов рака — они считаются даже более опасными, чем аналогичные лейкозы, возникающие *de novo* [16]–[20].

Вторичные неоплазии можно разделить на группы в зависимости от терапии, которая им предшествовала — терапия алкилирующими агентами, радиотерапия или терапия ингибиторами топоизомеразы II [21]. Обусловленные алкилирующими агентами и/или радиотерапией вторичные лейкозы обычно возникают через 4–7 лет после терапии и чаще всего относятся к хроническому лейкозу, тогда как обусловленные ингибиторами топоизомеразы II вторичные лейкозы развиваются через 1–3 года и относятся к острому миелоидному лейкозу (topoisomerase inhibitor-related AML, TI-related AML). Примечательно, что обусловленные разной терапией неоплазии характеризуются разными хромосомными aberrациями. Так, для первой группы характерны крупные делеции участков 5 и 7 хромосом, тогда как для второй группы — сбалансированные (реципрокные) транслокации с участием хромосом 11 или 21 [15], [22]–[25].

2.2. Транслокации, рекуррентные для вторичных ОМЛ, вызванных терапией ингибиторами ДНК-топоизомеразы II

Вторичные лейкозы, вызванные терапией ингибиторами ДНК-топоизомеразы II, характеризуются транслокациями с участием хромосом 11 или 21. Причем список перестроек, которые раз за разом встречаются при этом типе лейкоза, включает лишь 8 транслокаций (таблица 2.1). Все эти рекуррентные транслокации являются перестройками с участием генов *AML1* и *MLL*. Поэтому мы рассмотрим нормальные функции генов *AML1* и *MLL* в клетке и последствия транслокаций с их участием, которые приводят к развитию лейкоза.

Таблица 2.1. Рекуррентные транслокации, обнаруживаемые при лейкозах, вызванных терапией ингибиторами топоизомераз II. Указаны различные названия, выделены те названия, которые приняты за официальные в современной классификации генов HUGO Gene Nomenclature Committee.

Ген	Транслокация	Ген-партнер	Ссылки
AML1 (RUNX1)	t(8;21)(q22;q22)	ETO (RUNX1T1)	[26]–[30]
	t(3;21)(q26.2;q22)	MDS1-EVI1 (MECOM , PRDM3)	[31], [32]
	t(1;21)(p36;q21)	PRDM16	[33], [34]
MLL (KMT2A)	t(9;11)(p22;q23)	AF9 (MLLT3)	[26], [28]
	t(4;11)(q21q23)	AF4 (AFF1 , MLLT2)	[35], [36]
	t(19;11)(q13q23)	ELL	[28], [37]
	t(11;19)(q23;p13.3)	ENL (MLLT1)	[28], [37], [38]
	t(11;16)(q23;p13)	CREBBP	[39]–[41]

Ген *AML1*

В 1991 году был идентифицирован точный локус хромосомной перестройки t(8;21) у пациентов с острым миелоидным лейкозом (Acute myeloid leukemia). Перестраиваемый ген был, соответственно, назван *AML1* [42]. Затем ученые выяснили, что ген *AML1* гомологичен гену дрозофилы *run1*, открытому еще в 1980-е годы, и поэтому *AML1* получил второе название *RUNX1*, от Runt-related transcription factor [31]. Вскоре было обнаружено, что этот ген часто перестраивается при лейкозах, вызванных терапией ингибиторами топоизомераз II [43], [44]. В настоящее время известно, что ген *AML1* участвует в перестройках не только при ОМЛ, но и при некоторых других видах лейкозов [31], [45].

Изоформы мРНК и белка AML1

Экспрессия *AML1* может осуществляться с двух промоторов: проксимального и дистального. Выделяют три основные изоформы белка: AML1A, AML1B, AML1C. Все они содержат вблизи от N-конца Runt-домен (runt-homology domain, RHD), необходимый для связывания с ДНК и для гетеродимеризации AML1 с белком CBF β (core binding factor β). Гетеродимеризация с CBF β повышает сродство AML1 к ДНК [46]. Более длинные изоформы B и C содержат трансактивирующий домен (transactivation domain, TAD) и консервативный участок VWRPY на C-конце белка, участвующий во взаимодействии с некоторыми белками-репрессорами (рисунок 2.1). Основная изоформа — AML1C [47].

Нормальная функция AML1

Ген *AML1* экспрессируется в кроветворных клетках на протяжении всего онтогенеза [48]. Изучение его функции началось с нокаута *Aml1* у мышей. В этих работах было показано, AML1 важен при созревании гемопоэтических стволовых клеток в раннем эмбриогенезе [49], [50]. При врожденном нокауте *Aml1* эмбрионы гибнут из-за серьезных нарушений миелопоэза и эритропоэза [51]. Кондиционный нокаут *Aml1* у взрослых мышей не приводит к утрате гемопоэза, однако у таких мышей нарушается процентное соотношение различных лейкоцитов в крови, а также клеток в костном мозге [52], [53].

Дальнейшие опыты, проведенные *in vitro* на клетках гомогенного эндотелия мыши, продемонстрировали направляющую функцию AML1 в гемопоэзе: при дифференцировке этих клеток до гемопоэтических стволовых клеток AML1 связывается с промоторами множества генов, привлекая коактиваторы или корепрессоры. AML1 активирует гены, участвующие в дифференцировке миелоидных клеток и эритроцитов, а к группе репрессируемых им принадлежат гены, играющие роль в развитии нервной системы, скелетных мышц и печени [54]. Таким образом, AML1 регулирует транскрипцию генов при дифференцировке гемопоэтических клеток.

Для активации генов комплекс AML1/CBF β способен рекрутировать такие белки, как гистонацетилтрансферазы p300 и CREBBP (cyclic AMP response element binding protein), коактиватор транскрипции C/EBP (CCAAT-enhancer-binding protein), транскрипционные факторы YAP (Yes[-киназа] associated protein), GATA-1 (названный по узнаваемой им последовательности), ETS1 (Erythroblast Transformation Specific) и другие [55]–[62]. А в роли рекрутируемых AML1 репрессоров выступают транскрипционные факторы TLE1 (Transducin like enhancer of split 1) и SIN3A (гомолог SIN3 дрожжей) [63], [64]. Связывание различных белков зависит от фосфорилирования AML1 и отличается у разных изоформ: так, связывание репрессора SIN3A регулируется фосфорилированием остатков серина 273 и треонина 276 AML1 киназой ERK (extracellular signal-regulated kinase), а TLE1

связывается с мотивом VWRPY, имеющимся на С-конце длинных изоформ AML1 [64]–[66].

Развитие лейкозов при перестройках гена AML1

Среди партнеров AML1 при рекуррентных транслокациях, вызванных ингибиторами топоизомераз, самый частый — ген ETO (Eight twenty one), или RUNX1T1 [67]. Оба названия гена отсылают к его транслокации t(8;21) с AML1 (RUNX1). Эта хромосомная транслокация встречается также при других лейкозах, и в достаточной мере изучена [68]. При транслокации AML1-ETO в получающемся слитом белке присутствует N-конец от белка AML1, содержащий домен RHD. С-конец слитого белка представлен почти полноразмерным белком ETO (рисунок 2.1) [46].

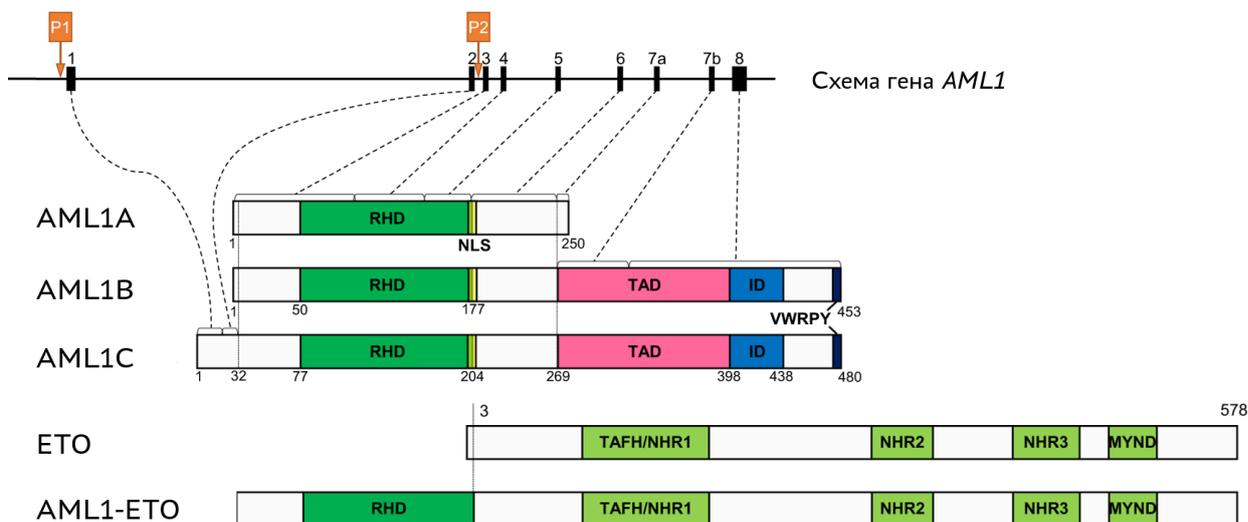


Рисунок 2.1. Схема гена *AML1* с экзонами и структуры изоформ AML1, ETO и слитого белка AML1-ETO. У AML1: RHD — runt-homology domain, TAD — трансактивирующий домен, ID — ингибиторный домен, на С-конце отмечен пентапептид VWRPY. У ETO: NHR1–3 — nervy homology regions, MYND — (myeloid, Nervy, and DEAF-1)-домен. По [47].

Белок ETO представляет собой транскрипционный репрессор, в структуре которого присутствуют четыре консервативных участка NHR (nervy homology regions — гомологичны по последовательности белку Nervy дрозофилы) [69]. ETO сам по себе не связывается с ДНК — к целевому локусу он привлекается через транскрипционные факторы. Репрессия также осуществляется с помощью других белков — корепрессоров, взаимодействующих с NHR-участками: NCoR1 (nuclear receptor corepressor), SMRT (silencing mediator of retinoid and thyroid receptors), SIN3A и деацетилаз гистонов [47], [70].

Ген *ETO* в норме не экспрессируется в кроветворных и лимфоидных органах [71] (www.proteinatlas.org/ENSG00000079102-RUNX1T1/tissue). Можно заключить, что образование слитого гена *AML1-ETO* приводит, во-первых, к появлению практически полноразмерного белка ETO в клетках, где он не должен присутствовать; а во-вторых, ETO

за счет участка AML1 привлекается к генам, к которым он не должен привлекаться. В итоге репрессируется транскрипция многих генов, экспрессия которых в норме регулируется транскрипционным фактором AML1 [72]–[74].

В реальности ситуация сложнее — экспрессия слитого белка приводит не только к репрессии многих генов, но и к активации некоторых других генов с помощью привлекаемых белков-активаторов. Так, было показано, что ETO способен привлекать гистонацетилтрансферазы с помощью первого участка NHR [75]. Вероятно, репрессирующее или активирующее действие AML1-ETO зависит от сигнальных каскадов, активных в данный момент в клетке, и особенностей хроматина в том месте, где находится регулируемый ген [76].

Важной особенностью клеток с t(8:21), полученных от пациентов, является наличие интактного аллеля *AML1*, что говорит о необходимости нормального AML1 для клетки [77]. По-видимому, интактный AML1 и слитый AML1-ETO конкурируют за связывание с промоторами генов — нокдаун AML1-ETO в этих клетках меняет распределение сайтов связывания нормального AML1 [78].

Также стоит отметить, что для развития лейкоза кроме перестройки AML1-ETO необходимы дополнительные мутации, по крайней мере в мышинных моделях [48], [79], [80]. Существуют данные о том, что на развитие лейкоза влияет и то, какие изоформы белка получают в результате альтернативного сплайсинга слитой мРНК [76], [81], [82]. Можно заключить, что экспрессия слитого белка AML1-ETO — первый шаг, который создает условия для дальнейших преобразований клетки на пути к развитию лейкоза.

Ген *MLL*

В 1980-х годах среди пациентов с острым лейкозом начали выделять подгруппу с крайне неблагоприятным прогнозом. С появлением проточной цитофлуориметрии было обнаружено, что бластные клетки этих агрессивных лейкозов часто экспрессируют поверхностные маркеры, свойственные как лимфоидной, так и миелоидной линиям. Иногда во время лечения наблюдалась даже смена набора этих поверхностных маркеров, и лейкоз, первоначально диагностированный как ОЛЛ, мог рецидивировать в ОМЛ. Соответственно, был введен термин «лейкоз смешанного фенотипа» [83]–[85]. Вскоре стало ясно, что для этого лейкоза характерна транслокация локуса 11q23. В 1991 году был идентифицирован ген, который находится в точке перестройки в клетках таких пациентов. Соответственно, ген был назван *MLL*, от mixed lineage leukemia (лейкоз смешанного фенотипа). Выявленный ген оказался гомологом гена *trithorax* дрозофилы, кодирующего лизинметилтрансферазу, которая активирует Нох-гены [86], [87]. Ген *MLL* кодирует лизинметилтрансферазу 2А, и отсюда его другое название — *KMT2A* (lysine (K)-specific methyltransferase 2A).

Хромосомные транслокации с участием гена *MLL* наблюдаются примерно в 10% случаев острого миелоидного и острого лимфоидного лейкозов [88], причем среди лейкозов, вызванных терапией топоизомеразными ядами, такие транслокации составляют большую часть [25], [89], [90]. Особенно часто — до 85% случаев — перестройки *MLL* встречаются при ОЛ у детей первого года жизни. Также мутации этого гена обнаруживаются и при многих других видах рака, в частности, рака толстой кишки, легких, мочевого пузыря, эндометрия и при раке груди [91]–[94].

Экспрессия перестроившегося *MLL*, слившегося с одним из своих многочисленных генов-партнеров, достаточна для запуска событий, приводящих к развитию лейкоза, что напрямую показано на мышах [88]. Это объясняется глобальной функцией *MLL* в клетках, где он экспрессируется: комплекс белков, формирующийся на основе *MLL*, осуществляет эпигенетическую маркировку промоторов тканеспецифичных генов. Такая маркировка необходима для поддержания паттерна экспрессии этих генов в процессе пролиферации и дифференциации клеток — другими словами, *MLL* отвечает за «транскрипционную память» клеток при их делении [95].

Ген *MLL* играет важную роль в эмбриональном развитии, особенно на ранних стадиях гемопоэза, хотя и во взрослом возрасте активно экспрессируется в некоторых тканях. Особая роль *MLL* в эмбриональном развитии состоит в активации транскрипции НОХ-генов и некоторых других генов, связанных с пролиферацией и регуляцией клеточного цикла [96]–[100]. Нокаут по *Mll* приводит к гибели мышей на ранних стадиях эмбриогенеза из-за нарушений гемопоэза и дифференциации гемопоэтических стволовых клеток [101]–[103]. Кондиционный нокаут *Mll* у мышей показывает, что во взрослом организме экспрессия *Mll* необходима клеткам костного мозга [104].

Структура белка MLL

Структура нормального белка *MLL* позволяет понять механизмы его работы в клетке и последствия образования слитых с *MLL* белков. *MLL* — крупный многодоменный белок, состоящий из 3969 аминокислот, его молекулярная масса — 500 кДа (рисунок 2.2). Несколько доменов *MLL* отвечают за взаимодействие с определенными участками хроматина, а также с белками, работающими в комплексе с *MLL* — это домены AT-hook, CxxC, PHD. Каталитический домен осуществляет ди- и триметилирование лизина, находящегося в 4-м положении полипептидной цепи гистона H3 (H3K4me_{2/3}), что сопряжено с активацией транскрипции соответствующих генов [91]. После синтеза на рибосоме *MLL* расщепляется протеазой таспазой 1 на два полипептида — N-*MLL* и C-*MLL*, которые образуют единый комплекс, взаимодействуя доменами FYRN и FYRC (FY-rich участки N- и C-доменов) [105].

На N-фрагменте MLL находятся все домены, отвечающие за связывание транскрибируемой ДНК. AT-hook — участок неспецифического взаимодействия белка с малой бороздкой ДНК. Домен CxxC обеспечивает уже специфическое связывание MLL с неметилированными CpG-динуклеотидами, которые характерны для активных промоторов [106]–[108]. Окружающие CxxC-домен участки — pre-CxxC и post-CxxC — важны для взаимодействия MLL с элонгационными факторами РНК-полимеразы II PAFc (polymerase associated factor complex) [109]. Ближе к середине белка находятся домен PHD (plant homeodomain), распознающий H3K4me_{2/3} — метку промоторов активных генов [110]–[112]. Также на N-фрагменте MLL находится мотив для связывания адаптерных белков Menin1 и PSIP1 (PC4 [positive coactivator 4] And SFRS1 [Serine And Arginine Rich Splicing Factor 1] Interacting Protein 1; другое название белка — LEDGF/p75), которые дополнительно обеспечивают привлечение MLL к участкам хроматина с H3K36me_{2/3} [113], [114]. Также они привлекают к MLL транскрипционные факторы, такие как SMAD (название отсылает к гомологам из *C. elegans* SMA [«отвечающий за фенотип "small"»] и белкам дрозофилы MAD [Mothers Against Decapentaplegic]) и NFκB (Nuclear factor kappa-light-chain-enhancer of activated B cells) [115].

C-концевой фрагмент MLL содержит трансактивирующий домен TAD, привлекающий гистонацетилаз H3K27, H3K9 и H4K16 — p300/CREBBP, MOZ [monocytic leukemic zinc-finger protein] и MOF [назван по гомологу из дрозофилы MOF], соответственно [116]. На C-конце MLL находится домен SET (Su(var)3-9, Enhancer-of-zeste and Trithorax), который катализирует ди- и триметилирование H3K4 [117]. Вдобавок C-MLL привлекает белки WDR5 (WD repeat-containing protein 5), RBBP5 (Retinoblastoma-binding protein 5), ASH2L (Absent, small, or homeotic discs protein 2 like) и DPY30, и в составе полученного комплекса метилирующая активность SET-домена MLL многократно усиливается [91], [118]–[120].

Важным свойством MLL-комплекса является возможность его переключения между состояниями транскрипционного активатора и транскрипционного репрессора. Переключения происходит за счет белка CYP33 (cyclophilin 33) — пролин-изомеразы, при связывании с которой происходит *цис-транс* изомеризация пролина в составе PHD [121]. В результате MLL меняет свою структуру так, что вместо узнавания H3K4me_{2/3} домен PHD начинает связывать деацетилазы гистонов, репрессирующие транскрипцию [122]–[124]. Важность этого механизма подчеркивается тем, что удаление из MLL участка PHD приводит к опухолевой трансформации клетки [125].

Таким образом, белок MLL является регулятором транскрипции генов, которые несут метки активного хроматина — определенные гистоновые модификации и неметилированные CpG. В зависимости от контекста и активных сигнальных путей в клетке MLL записывает активирующие модификации с помощью SET-домена и привлекаемых

гистоновых ацетилтрансфераз либо репрессирующие — за счет привлекаемых гистоновых деацетилаз [95]. Можно заключить, что MLL обеспечивает транскрипционную память при линейной дифференровке клеток в ходе эмбрионального развития и в гемопоэзе.

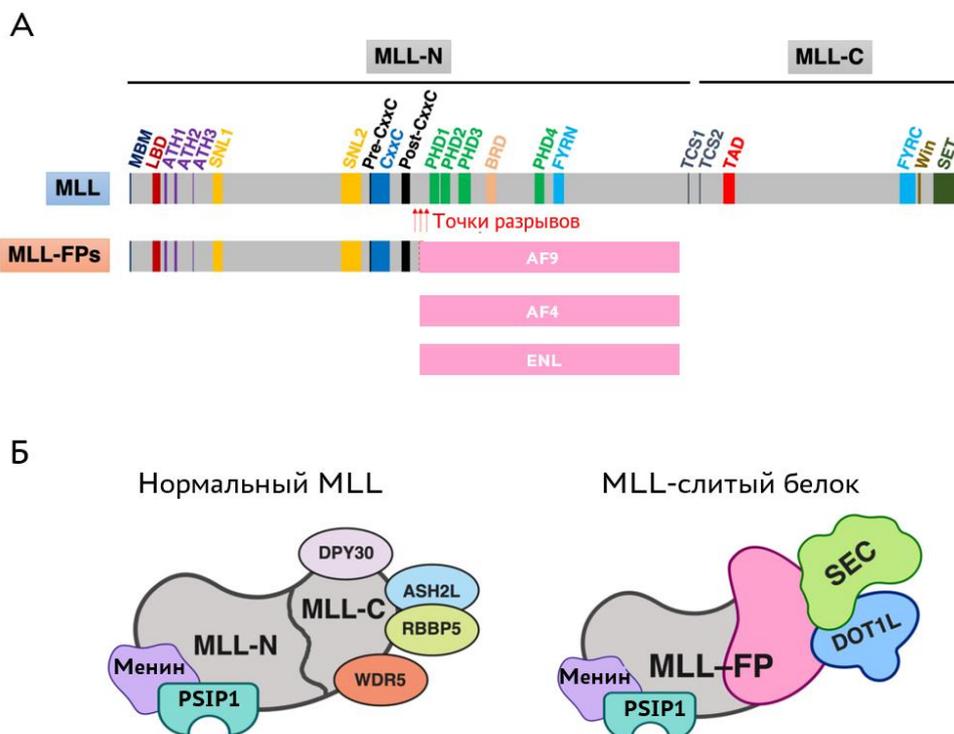


Рисунок 2.2. А. Строение MLL и слитого с MLL белка, по [115]. MLL-N и MLL-C — фрагменты, образующиеся при расщеплении таспазой; MBM и LBD — менин-связывающий мотив и LEDGF(PSIP1)-связывающий домен, ATH1–3 — участки AT-hook, SNL1–2 — сигналы ядерной локализации, pre-CxxC, CxxC и post-CxxC — CxxC-домен и окружающие его участки, PHD1–3 — участки plant homeodomain, вместе с BRD бромодоменом, FYRN и FYRC — FY-rich участки N- и C-доменов, TCS1–2 — сайты расщепления пре-MLL таспазой, TAD — трансактивирующий домен, Win — участок WDR5 interaction, SET — SET-домен. **Б** — комплексы, формируемые нормальным MLL и MLL-слитым белком (MLL-FP). Менин, PSIP1, DPY30, ASH2L, RBBP5, WDR5, DOT1L — белки, взаимодействующие с MLL и MLL-FP, SEC — суперэлонгационный комплекс. Изображение составлено на основе статьи [91].

Развитие лейкозов при перестройках гена MLL

Хромосомные транслокации гена MLL чаще всего происходят с разрывом в так называемом кластере точек разрыва, BCR (breakpoint cluster region), протяженностью около 10 т.п.н. Этот BCR захватывает интроны с 9 по 11. В структуре белка это соответствует области домена PHD. В слитых белках присутствует N-концевая часть белка MLL, а C-конец происходит от партнера по транслокации [115], [126], [127]. Исходя из структуры MLL и функции его доменов понятно, что образование таких слитых белков принципиально нарушает функционирование комплекса на основе MLL.

Необычной особенностью транслокаций с участием гена *MLL* является то, что существует более 100 генов-партнеров *MLL* по транслокации [8]. И несмотря на такое разнообразие, большинство случаев лейкозов с транслокациями *MLL* схожи по своим клиническим проявлениям [128], [129]. Это можно объяснить, во-первых, сходными точками разрыва в слитых генах, и во-вторых, тем, что большинство пациентов с перестройкой *MLL* несут транслокации с одним из генов *ENL*, *AF9*, *AF4*, *ELL*, и *AF10*, тогда как остальные транслокации встречаются крайне редко. В частности, все транслокации *MLL*, вызванные действием ингибиторов топоизомеразы II, как раз происходят с этими партнерами, за исключением *MLL-CREBBP* (таблица 2.1).

Схожесть транслокаций с участием генов *ENL*, *AF9*, *AF4*, *ELL*, и *AF10* объясняется тем, что они кодируют белки, образующие вместе с фактором Р-TEFb супер-элонгационный комплекс, необходимый для осуществления эффективной элонгации транскрипции РНК-полимеразой II [130], [131]. И когда *MLL* оказывается слит с одним из компонентов этого комплекса, то он привлекает к ДНК весь комплекс целиком, а также фактор DOT1L [132], [133]. Р-TEFb, входящий в данный комплекс, осуществляет фосфорилирование С-концевого домена РНК-полимеразы II, что стимулирует элонгацию транскрипции, а DOT1L осуществляет ди- и триметилирование H3K79 — эта метка ассоциирована с активным хроматином [133]–[135]. Следует также отметить, что при транслокации *MLL* теряется участок гена, который кодирует PHD-домен, отвечающий за переключение *MLL* в репрессирующее состояние. В итоге остаток *MLL* полностью теряет способность к репрессии генов, а слитый с ним супер-элонгационный комплекс и фактор DOT1L напрямую активируют транскрипцию тех генов, с которыми связывается *MLL*.

Механизм развития лейкоза при транслокации *MLL* включает активацию транскрипции генов НохА-кластера, а также генов, кодирующих белки сигнальных путей, например генов малой ГТФазы RAC1 и рецепторной тирозинкиназы FLT3. Это приводит к опухолевой трансформации кроветворных клеток. Показано, что слитые *MLL*-белки стимулируют экспрессию гомеобоксных генов *HOXA*, *MEIS1* и *PBX*, генов-транскрипционных регуляторов *AML1* и *MYC* и других генов. Промоторы активированных генов содержат метки активного хроматина H3K4me3 и H3K27ac, а также имеют повышенный уровень H3K79me2 [96], [98], [136].

В случае транслокации *MLL-CREBBP* механизм влияния слитого белка на клетку схож — CREBBP является гистонацетилтрансферазой, которая навешивает активирующие метки на гистоны, а сам *MLL* разорван в том же месте, что и при описанных выше его транслокациях [83].

Также в последнее время накапливается все больше данных, что для развития лейкоза необходима экспрессия слитого гена, полученного в результате реципрокной транслокации. Например, выдвигается гипотеза, что продукт реципрокной транслокации

AF4-MLL является глобальным активатором транскрипции, так как C-MLL вносит активирующие метки неспецифично, в отличие от интактного MLL [137]. Такие модификации могут привести к реактивации участков генома, которые уже были инактивированы в процессе дифференцировки клеток. В числе реактивированных могут быть гены, работающие в стволовых клетках, например, *NANOG* и *OCT4*, что было показано для клеток с транслокацией t(4;11) [138], [139].

Таким образом, при транслокациях с участием MLL происходит потеря контроля над транскрипцией генов, работающих в первую очередь в незрелых клетках. Активируются гены, отвечающие за пролиферацию, и теряется контроль над делением клеток, что приводит к развитию лейкоза.

Развитие лейкозов при перестройках *AML1* и *MLL*

Можно подытожить, что гены *AML1* и *MLL*, перестраиваемые в результате терапии ингибиторами топоизомераз II, оба в норме регулируют транскрипцию в гемопоэтических клетках. А образующиеся в результате транслокации слитые *AML1*- и *MLL*-белки приводят к разбалансировке транскрипции в этих клетках. При этом, видимо, не существует единственного варианта такого изменения паттерна экспрессии генов. Скорее, в клоне клеток с транслокацией появляются клетки с разным паттерном экспрессии. И из этих эпигенетически разнообразных клеток постепенно выделяется наиболее активно делящаяся, которые становятся лейкозными клетками [140].

Развитие опухоли — процесс, похожий на дарвиновскую эволюцию [141]. В клетках происходят соматические мутации, которые могут повышать приспособленность опухолевых клонов: их устойчивость к иммунной системе и скорость деления. Однако при делении клеток передаются не только мутации, но и эпигенетические метки, причем накопление эпигенетических изменений происходит гораздо быстрее, чем мутаций [142]. И эффект от изменения уровня экспрессии гена часто более масштабный, чем от точечной мутации. Поэтому клетки с транслокациями *AML1* и *MLL*, в результате которых наблюдаются масштабные эпигенетические эффекты, быстро дают клоны, развивающиеся в опасные лейкозы.

2.3. Топоизомеразные яды — эффективное средство противоопухолевой терапии

Развитие противоопухолевой химиотерапии

Противоопухолевая химиотерапия как конвенциональный метод лечения оформилась во второй половине 20 века, до этого момента опухоли лечили хирургически или с помощью радиотерапии. Разработка химиотерапии рака во многом связана с появлением моделей для

изучения противоопухолевого действия веществ — в начале XX века ученые научились подсаживать опухоли грызунам, что позволило экспериментировать в поиске веществ с противоопухолевым действием [143].

Однако обнаружение первого вещества, которое бы действовало против опухолей, произошло в результате трагичных событий и случайных совпадений. В 1943 году в городе Бари, занятом к тому времени союзниками, произошел налет немецкой авиации. Было потоплено много военно-транспортных судов. Пострадавшие моряки испытывали необычные симптомы — низкое давление, тошноту, апатию и повреждения кожи. На место был вызван подполковник Александр Стюарт, специалист по на химическому оружию. Он выяснил, что среди потопленных судов был американский транспорт, доставивший «на всякий случай» артиллерийские снаряды с ипритом — алкилирующим агентом, боевым отравляющим веществом. Один из обнаруженных симптомов отравления — резкое уменьшение содержания лейкоцитов в крови — напомнил Стюарту его предыдущие работы на грызунах по изучению иприта, в которых наблюдалось такое же явление. Это навело его на мысль о возможном применении иприта для борьбы с лейкозами и лимфомами, симптомом которых является высокое содержание лейкоцитов в крови. Поэтому после окончания войны были проведены клинические испытания иприта как противолейкозного препарата [144]. Однако это была не первая идея применения производных иприта для борьбы с опухолями — изучавшие его во время Второй мировой войны американские ученые из Йеля также заметили его действие на содержание лейкоцитов в крови и предположили возможность применения для терапии лейкоза, хотя их работа была опубликована только после окончания войны [145]. Так или иначе, химиотерапия опухолей с помощью алкилирующих агентов получила бурное развитие в послевоенные годы.

Изучение в конце 1940-х годов фолиевой кислоты показало, что ее нехватка напоминает азотистый иприт по действию на костный мозг. Были получены и проверены антагонисты фолиевой кислоты. Выяснилось, что их использование действительно приводит к положительным результатам у пациентов с лейкозом [146], однако высокая токсичность этих веществ ограничивала их применение. Развитием этих работ стало применение в химиотерапии метотрексата.

Экстракт из тканей Подофилла щитовидного (*Podophyllum peltatum*), или подофиллин, эпизодически применялся в медицине против кожных опухолей и бородавок еще в 19 веке [147]. В 1946 году было показано, что подофиллин обладает сходным с колхицином действием — препятствует делению клеток [148]. Однако попытки его применения против опухолей были ограничены его высокой токсичностью [149]. Поэтому в 1950-х начались поиски действующей молекулы из подофиллинового экстракта. Наилучшее противоопухолевое действие показал бензилиден 4'-деметилэпиподофиллин глюкозид, и в 1966 году был синтезирован его синтетический аналог — этопозид (VP-16),

годом позже — тенипозид (VM-26). В 1974 году завершились доклинические испытания, доказавшие их высокую эффективность против самых разных видов опухолей, а с 1983 было получено разрешение FDA (Food and Drug Administration, агентство Министерства здравоохранения и социальных служб США, контролирующее качество пищевых продуктов и применение лекарств) на применение этопозида в клинической практике [147]. На сегодня этопозид и аналогичные по механизму работы вещества находятся на переднем крае терапии самых разных типов опухолей: лейкозов, лимфом, саркомы и рака молочной железы, рака легких и других [150], [151].

Интересно, что механизм действия этопозида и тенипозида на момент их одобрения еще не был установлен. Было показано, что этопозид приводит к фрагментации ДНК *in vivo*, однако он не способен вносить разрывы в очищенную ДНК [152]. Только в 1984 году выяснили, что мишенью этопозида является ДНК-топоизомераза II типа — фермент, особенно востребованный в делящихся клетках, к которым относятся клетки опухоли. Его работа, с одной стороны, крайне необходима клетке, с другой — включает в себе смертельную опасность для клетки. Ниже пойдет речь о роли топоизомераз в клетке и механизме работы ДНК-топоизомераз II.

ДНК-топоизомеразы и их активность в клетках

К ДНК, представляющей из себя двойную спираль, применимы физические законы топологии. В двойной спирали одна из цепочек многократно пересекает плоскость, образуемой другой цепочкой. Данный параметр, называемый числом зацеплений, остается постоянным, пока ДНК замкнута или ее концы зафиксированы. В этом случае увеличение или уменьшение числа витков ДНК компенсируется добавлением супервитков, отрицательных или положительных, соответственно. Так как упаковка ДНК в ядре ограничивает ее во вращении, супервитки неизбежно возникают при плавлении ДНК в процессе транскрипции и репликации [153], [154]. Другой проблемой, связанной с топологией ДНК, является расхождение хромосом после репликации по дочерним клеткам. Так как репликация идет полуконсервативно — к каждой цепочке ДНК достраивается комплементарная ей цепь — то получившиеся хромосомы оказываются закручены относительно друг друга так же, как были закручены цепочки исходной двойной спирали ДНК. Эта ключевая проблема полуконсервативной модели репликации была сформулирована еще Фрэнсисом Криком и Джейсоном Уотсоном, а также Максом Дельбрюком, который для ее разрешения даже предложил альтернативную модель репликации — дисперсную модель [155], [156].

Проблемы, связанные с топологией ДНК, решаются ферментами-топоизомеразами с помощью временного разрыва цепочки ДНК. Топоизомеразы I типа вносят одноцепочечный разрыв, позволяя одной цепочке вращаться вокруг второй, тогда как

топоизомеразы II типа вносят двуцепочечный разрыв, через который может пройти другой фрагмент ДНК [154], [157]. И если проблемы сверхспирализации ДНК при транскрипции и репликации решаются топоизомеразами любого типа, то для расплетения хромосом может использоваться только ДНК-топоизомераза II типа, что было показано для многих организмов [158]–[161]. Поэтому любым делящимся клеткам критически необходима работа именно ДНК-топоизомераз типа II.

Задача ДНК-топоизомераз II типа заключается в контролируемом внесении двуцепочечных разрывов в ДНК, которые необходимы для расплетения хромосом, но в то же время представляют собой опасность для клетки. Почти все ДНК-топоизомеразы типа II из разных организмов гомологичны друг другу (их относят к подтипу IIА) и работают по одному механизму (рисунок 2.3). Топоизомераза IIА представляет собой гомодимер, мономеры которого соединяются друг с другом тремя участками, формируя трое «ворот» — N-, ДНК- и С-ворота. Ворота последовательно открываются и закрываются, пропуская фрагмент ДНК, называемый Т-фрагментом, или Т-сегментом (от «transport»). ДНК-ворота, расположенные в середине, связывают другой фрагмент ДНК (G-фрагмент, от «gate»), и открываются, временно разрывая его. Когда Т-фрагмент ДНК проходит ДНК-ворота и они закрываются, разрыв в G-фрагменте лигируется обратно. Образование двуцепочечного разрыва в G-фрагменте происходит за счет двух реакций переэтерификации: остатки тирозина каждой из субъединиц топоизомеразы атакуют фосфодиэфирные связи на разных цепях ДНК на расстоянии четырех нуклеотидов друг от друга, образуя ковалентные связи с 5'-фосфатами нуклеотидов [162]–[164]. Когда Т-фрагмент высвобождается через С-ворота, топоизомераза возвращается в исходное состояние и цикл ее работы завершается.

Хотя сама реакция переэтерификации не требует затрат энергии, в ходе каталитического цикла топоизомеразы типа IIА гидролизует две молекулы АТФ. В данном случае АТФ выступает как молекулярный переключатель: в комплексе с АТФ N-ворота закрыты и крепко удерживают вместе субъединицы фермента, в то время как ДНК-ворота открыты. Когда АТФ гидролизует до АДФ, субъединицы фермента в области N-ворот теряют сродство друг к другу, позволяя N-воротам открыться для следующего цикла работы; при этом уже ДНК-ворота стабилизируются в закрытой конформации. Согласованность в работе ворот крайне важна: если все ворота окажутся одновременно открытыми, это значит, что субъединицы топоизомеразы отсоединились друг от друга, и временный разрыв ДНК превратится в постоянный. Таким образом, молекулы АТФ аллостерически координируют открывание и закрывание ворот во время каталитического цикла топоизомеразы [165]–[167].

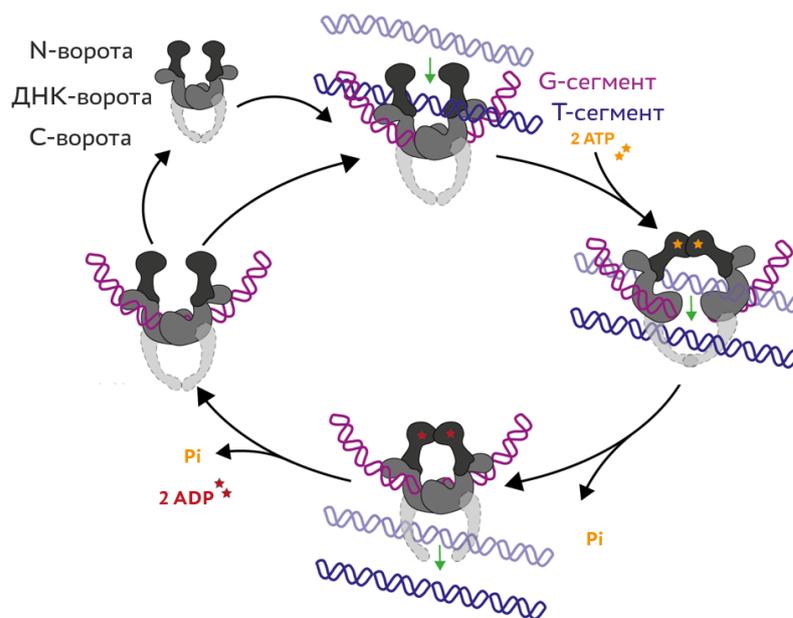


Рисунок 2.3. Механизм работы топоизомеразы типа IIА. На схеме показано согласованное изменение конформации фермента, что приводит к последовательному перемещению Т-сегмента ДНК через G-сегмент, при этом субъединицы фермента всегда удерживаются вместе. Рисунок на основе [163].

Топоизомеразы II в клетках человека

В геноме человека закодировано шесть ДНК-топоизомераз: TOP3 α , TOP3 β , TOP1, TOP1mt, TOP2 α и TOP2 β . Из них к топоизомеразам типа II относятся TOP2 α и TOP2 β [168]–[170] [171]. Аминокислотные последовательности TOP2 α и TOP2 β человека идентичны на 68%, имеют молекулярную массу 170 и 180 кДа, соответственно [172], [173].

Большая часть аминокислотной последовательности топоизомераз II консервативна. В каждой субъединице топоизомеразы II можно выделить N-концевой домен, он несет сайт связывания АТФ и образует N-ворота. Далее следует трансдуцирующий домен, в его состав входит К-петля (K-loop), на конце которой несколько остатков лизина взаимодействуют с ДНК (на рисунке 2.4 отмечено рамкой), что важно для активации АТФазной активности фермента [163]. В центральной части полипептида выделяют несколько доменов: TOPRIM-домен (topoisomerase-primase), Winged helix домен (WHD), где находится каталитический остаток тирозина, и TOWER-домен [174]–[176]. Центральные части мономеров образуют ДНК-ворота [177]. С-концевой домен (CTD) имеет сайты взаимодействия с другими белками, а также образует С-ворота [178]. Разные ортологи и разные изоформы фермента отличаются в первую очередь последовательностью аминокислот С-концевого домена.

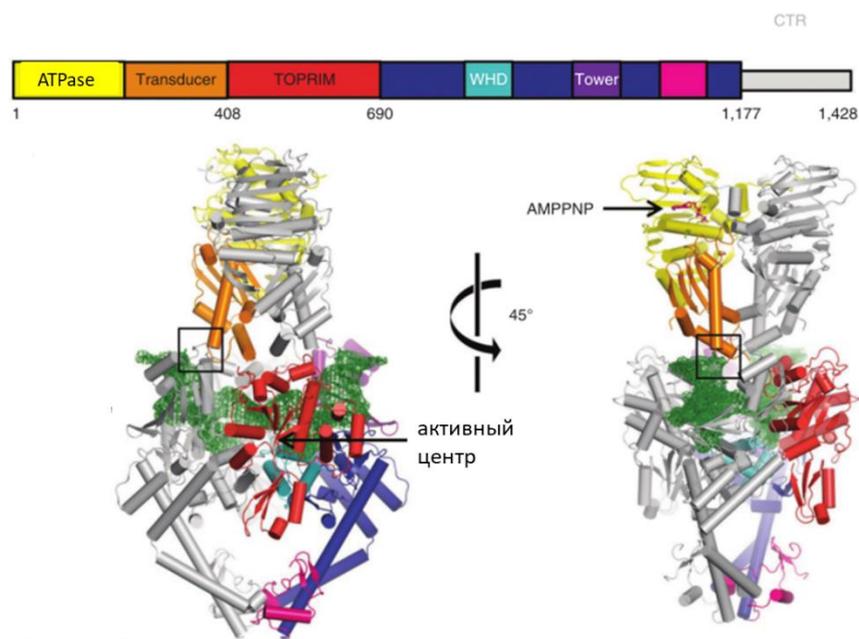


Рисунок 2.4. Структура топоизомеразы II *Saccharomyces cerevisiae*. В верхней части рисунка показана схема расположения доменов в субъединице фермента. Внизу показан димер топоизомеразы II, связанный с молекулой ДНК (показана зеленым цветом). АМППНР — негидролизуемый аналог АТФ, связанный с АТФазным доменом. Розовым цветом выделен С-концевой домен. Рисунок из статьи [163].

Интересно, что TOP2 α релаксирует положительно суперскрученную ДНК в 10 раз активнее, чем отрицательно суперскрученную, в то время как активность TOP2 β не зависит от знака суперспирализации ДНК. Такая же ситуация наблюдается и у *E. coli*, где есть 2 топоизомеразы II: гирана связывает положительный супервиток и превращает его в отрицательный, а топоизомераза IV релаксирует любые супервитки. Специфичность связывания обеспечивается доменом STD гираны, который устроен специальным образом для связывания именно положительного супервитка ДНК [179]. Аналогично и у изоформ топоизомеразы млекопитающих: при замене STD TOP2 β на STD TOP2 α , получившийся химерный белок будет предпочтительно релаксировать положительно суперскрученную ДНК [180], [181].

Топоизомеразы TOP2 α и TOP2 β выполняют в клетке разные функции. Топоизомераза TOP2 α выполняет свою функцию при делении клеток. Она экспрессируется в активно делящихся клетках в S-, G2- и M-фазах клеточного цикла и является мажорным белком метафазной хромосомы, в то время как экспрессия TOP2 β не зависит от фазы клеточного цикла [182]–[184]. Кроме того, топоизомераза TOP2 α , но не TOP2 β , колокализуется в ядре с фокусами инкорпорации бромдезоксиуридина, что предполагает функцию TOP2 α в разрешении топологических проблем при репликации [184]. Высокое сродство топоизомеразы TOP2 α к положительным супервиткам позволяет предположить, что она снимает супервитки, возникающие при работе ДНК-хеликазы [180]. Считается, что

TOP2 α принимает участие в компактизации хромосом при делении клетки [185]. Также TOP2 α необходима для сегрегации хромосом при делении. Нокаут гена топоизомеразы TOP2 α у мышей приводит к остановке развития эмбрионов на стадии 4–8 бластомеров из-за нарушений расхождения дочерних ядер [186], [187].

За локализацию TOP2 α в реплицирующемся хроматине ответственен ее STD, который *in vitro* способен взаимодействовать с эукариотическим скользящим зажимом PCNA (Proliferating cell nuclear antigen) [184]. Также известно, что участок STD TOP2 α взаимодействует с гистоном H3, причем различные модификации гистона усиливают или ослабляют это взаимодействие. Гистоны с этими модификациями распределены неравномерно по геному, что может влиять на неравномерность привлечения топоизомеразы к разным локусам [188].

TOP2 β выполняет довольно узкие функции, связанные с транскрипцией тканеспецифичных генов. Мыши с нокаутом по гену топоизомеразы TOP2 β демонстрируют тонкие дефекты в развитии нервных клеток, что ведет к гибели таких мышей вскоре после рождения [189]. Кондиционный нокаут TOP2 β в развивающемся головном мозге мыши приводит к нарушению нормального формирования слоев коры и также ведет к гибели животного [190]. Таким образом, эта изоформа не требуется для протекания глобальной транскрипции и репликации [191], [192]. Примечательно, что TOP2 β более активно экспрессируется в уже дифференцированных клетках. TOP2 β связывается с активными участками хроматина, в частности, с регуляторными участками генов и границами топологически ассоциированных доменов (ТАДов), что демонстрирует ее роль в регуляции экспрессии тканеспецифичных генов [193], [194]. В частности, показана важная роль топоизомераз в транскрипции длинных — более 200 тыс.п.н. — тканеспецифичных генов [195].

Как уже обсуждалось, субстратная специфичность топоизомераз TOP2 α и TOP2 β определяется их STD. Он же определяет различие изоформ в клеточной локализации и функциональной значимости. В норме TOP2 α при митозе входит в состав хромосом, в то время как TOP2 β равномерно распределена по цитозолю, однако при замене STD топоизомеразы TOP2 β на STD топоизомеразы TOP2 α первая обнаруживалась в составе метафазных хромосом. Также было показано, что в отсутствие активной TOP2 α , TOP2 β с STD от TOP2 α способна поддерживать пролиферацию клеток линии NT1080 [184], [196].

Топоизомеразные яды

На сегодняшний день изучены различные ингибиторы работы ДНК-топоизомеразы II. По механизму действия их принято делить на 2 группы — ингибиторы каталитического цикла и топоизомеразные яды (catalytic inhibitors и poisons, соответственно). Первые действуют на любых стадиях каталитического цикла, включая связывание и гидролиз АТФ, связывание и разрезание ДНК. Вторая группа — так называемые топоизомеразные яды, действие которых характеризуется повышением в клетке количества ковалентно связанных комплексов ДНК-фермент (Top2cc). Все используемые на сегодня в терапии опухолей ингибиторы топоизомеразы II относятся к группе топоизомеразных ядов [164], [197]. Чаще всего среди них используется этопозид, его механизм действия подробно изучен.

Этопозид связывается с ковалентным комплексом топоизомеразы с ДНК и стабилизирует его, препятствуя религированию ДНК [198]. Когда в G-сегмент ДНК внесен разрыв, молекула этопозида нековалентно связывается с ферментом, препятствуя полному закрытию ДНК-ворот и сближению концов разрыва ДНК (рисунок 2.5). Существует различие в связывании этопозида с изоформами топоизомеразы: в формировании связи с молекулой этопозида в TOP2 α участвует остаток метионина (Met762) вместо остатка глутамина (Gln778) в случае TOP2 β [199]. Одна молекула ингибитора взаимодействует с мономером фермента, т. е. стабилизирует одноцепочечный разрыв. Обычно обработка клеток этопозидом приводит к образованию 7–20 ОЦР на каждый ДЦР [200], [201].

Ряд экспериментов подтверждает ключевую роль ингибирования ДНК-топоизомеразы II в образовании транслокаций, вызванных соответствующей терапией. Так, этопозид индуцирует образование двуцепочечных разрывов в генах *AML1*, *MLL* и их партнеров, а к месту разрыва привлекаются белки репарации [126], [202]–[205]. Эти разрывы локализуются в кластерах точек разрывов, обнаруженных у пациентов с вторичным лейкозом.

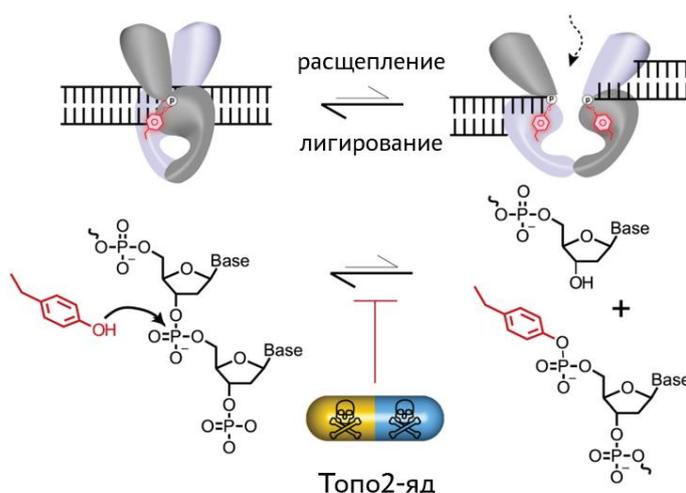


Рисунок 2.5. Ингибирование ДНК-топоизомеразы II топоизомеразными ядами. На основе [206].

Пути удаления топоизомеразы с ДНК

Топоизомераза, остановленная в результате действия топоизомерного яда и ковалентно сшитая с ДНК, препятствует транскрипции и репликации. Взаимодействие с транскрибирующей РНК-полимеразой запускает процесс удаления с ДНК топоизомеразы [207], [208]. В клетках человека существует несколько путей удаления топоизомеразы, ковалентно сшитой с ДНК. Их можно разделить на протеолитические и нуклеолитический пути (рисунок 2.6.). Протеолитические пути предполагают деградацию белка с помощью протеасомы или специальной протеазы SPRTN (Protein with SprT-like domain at the N terminus) с последующим отщеплением остатка полипептида ДНК-фосфодиэстеразой TDP2 (tyrosyl DNA phosphodiesterase 2). Нуклеолитический путь предполагает отщепление участка цепочки ДНК, на конце которого ковалентно пришит белок, с помощью эндонуклеаз систем репарации [206], [209].

Протеолитические пути удаления предполагают гидролиз связи между нуклеотидом и остатком тирозина специальным ферментом — TDP2. Показано, однако, что хотя этот фермент необходим для устойчивости клеток к этопозиду, его одного недостаточно для удаления топоизомеразы с ДНК, и требуется предварительная денатурация или частичный протеолиз белка [210]–[212]. Первичный протеолиз может осуществляться с помощью протеасомы — об этом говорят опыты с ингибитором протеасомы MG132 и с нокдауном гена, отвечающего за сборку протеасом [213]. Деградация топоизомеразы с помощью протеасом осуществляется после убиквитинилирования Top2cc [214]. Другая протеаза, которая может быть задействована в деградации Top2cc — SPRTN, металлопротеаза, которая специализируется на деградации пришитых к ДНК полипептидов [215]–[217]. Наконец, альтернативный фактор, обеспечивающий для TDP2 доступность Top2cc, — SUMO-лигаза ZAT1. Сумоилирование топоизомеразы приводит к изменению конформации, что позволяет TDP2 гидролизовать связь между нуклеотидом и остатком тирозина [206], [218], [219].

Нуклеолитический путь подразумевает удаление Top2cc вместе с небольшим участком ДНК с помощью эндонуклеазы системы репарации Mre11 (Meiotic Recombination 11 Homolog A). Доказательством того, что эта эндонуклеаза может удалить ковалентно связанный 5'-аддукт с ДНК, является изучение процессов рекомбинации, осуществляющейся с помощью белка Spo11. Этот белок гомологичен топоизомеразам типа IIВ, встречающимся у некоторых архей и растений [220],[221]. При рекомбинации Spo11 вносит разрывы в ДНК, образуя фосфотирозиновую связь, как и топоизомераза, однако не зашивает их. Показано, что белок Mre11 необходим для удаления Spo11-аддукта у дрожжей [222]–[224]. Исследования показывают, что в клетках с деплецией Mre11 накапливаются Top2cc, причем как в случае обработки топоизомеразными ингибиторами, так и без них, что подчеркивает возможность эндогенных нарушений работы топоизомераз

[174], [225], [226]. Кроме нуклеазы Mre11, в удалении Top2cc задействованы также белки BRCA1 (Breast cancer type 1) и CtIP (CtBP-interacting protein, CtBP — C-terminal-binding protein 1) [225], [227]. Эти белки вместе участвуют в одном пути репарации, который будет рассмотрен в соответствующем разделе. Пока не ясно, требуется ли процессинг топоизомеразы перед ее удалением по Mre11/CtIP/BRCA1-зависимому пути в клетке. По крайней мере, в экспериментах *in vitro* этого не требовалось [174].

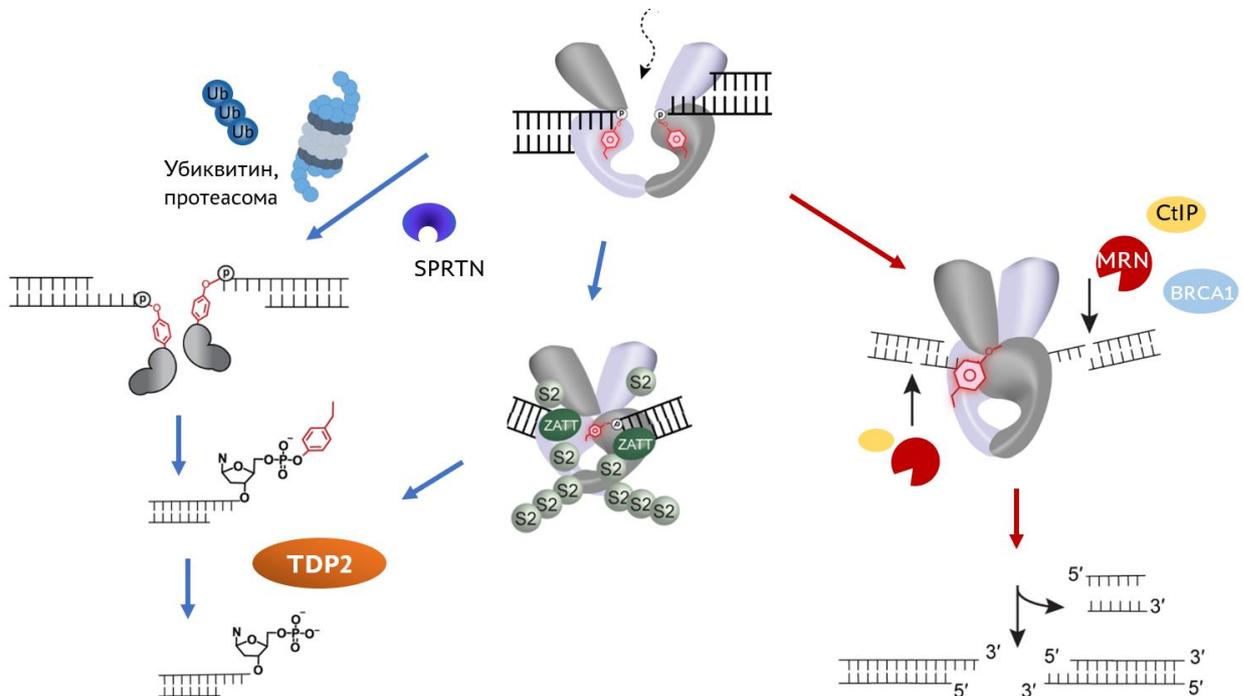


Рисунок 2.6. Протеолитический (синие стрелки) и нуклеолитический (красные стрелки) пути удаления Top2cc с ДНК. Показаны основные белки, участвующие в процессе (описание в тексте). S2 — SUMO, Ub — убиквитин.

В любом случае, удаление топоизомеразы приводит к образованию ДЦР ДНК. Множественные ДЦР приводят к запуску клеточного ответа на повреждение ДНК, что включает в себя фосфорилирование остатка S139 гистонов H2AX, репарацию ДНК или индукцию апоптоза [228]. Действие этопозида можно зарегистрировать по появлению фосфорилированного гистона H2AX (γ H2AX) уже через несколько минут после обработки им клеток [202], [229].

2.4. Репарация двуцепочечных разрывов

Двуцепочечные разрывы — самый опасный тип повреждений ДНК. Разрывы хромосом приводят к масштабным нарушениям распределения генов при делении клеток: фрагмент хромосомы, у которого отсутствует центромера, не сможет участвовать в распределении хромосом по дочерним клеткам. Поэтому существует несколько подстраховывающих друг

друга систем репарации двуцепочечных разрывов (ДЦР), а появление таких разрывов останавливает прохождение клеточного цикла [230].

Два основных пути репарации ДЦР — направляемая гомологией репарация (Homology-directed repair, HDR) и негомологичное соединение концов разрыва (Non-homologous end joining, NHEJ). Направляемая гомологией репарация требует присутствия гомологичной матрицы, а негомологичное соединение концов осуществляется путем прямого лигирования концов разрыва [231]. Также есть дополнительные пути, которые задействованы гораздо реже — это обусловленное микрогомогией соединение концов разрыва (microhomology-mediated end joining, MMEJ) и отжиг одноцепочечных концов (single-strand annealing, SSA) [232], [233].

Пути репарации ДЦР различаются по их месту в клеточном цикле, условиям, в которых они работают, по задействованным белкам и по устойчивости к ошибкам — например, появлению небольших инсерций или делеций или образованию транслокаций. Для понимания механизмов, задействованных при хромосомных перестройках, мы рассмотрим пути репарации ДЦР, работающие в клетках млекопитающих.

Репарация ДЦР, направляемая гомологией

Репарация ДЦР, направляемая гомологией, предполагает использование участков хромосом, гомологичных последовательностям, в которых случается разрыв. Этот путь репарации еще иногда называют гомологической рекомбинацией, так как именно такая репарация ДЦР обеспечивает образование кроссоверов при мейотической рекомбинации. В диплоидных клетках дрожжей донором гомологии выступает гомологичная хромосома, и для них этот путь репарации ДЦР является основным. Однако существенно больший размер генома млекопитающих усложняет поиск гомолога в интерфазе, и для них направляемая гомологией репарация ДЦР характерна только после удвоения хромосом — в этом случае донором гомологии выступает сестринская хроматида [234].

Репарация ДЦР, направляемая гомологией, осуществляется по следующей схеме (рисунок 2.7). Сначала 5'-концы ДНК в месте разрыва укорачиваются в направлении 5'→3'. Таким образом, получают длинные одноцепочечные выступающие 3'-концы ДНК. Эти участки ДНК покрываются белком RAD51, образуя нуклеофиламент, который взаимодействует с участком двуцепочечной гомологичной ДНК. В результате осуществляется инвазия нуклеофиламента в этот участок ДНК с образованием так называемой D-петли (от англ. displacement) и участка гетеродуплекса, в котором комплементарные цепи ДНК исходно принадлежали разным хромосомам. Затем следует удлинение 3'-конца по матрице ДНК. Удлиненный 3'-конец поврежденной хромосомы может покинуть D-петлю, вернуться в свою хромосому, и там по нему достраивается вторая цепь ДНК. Такой исход называется зависимым от синтеза отжигом цепи (synthesis-

dependent strand annealing pathway, SDSA) и наблюдается в большинстве случаев HDR в соматических клетках [235], [236]. Наличие в клетке хеликазы BLM семейства RecQ способствует разрушению D-петли, направляя HDR по пути SDSA [237]–[239].

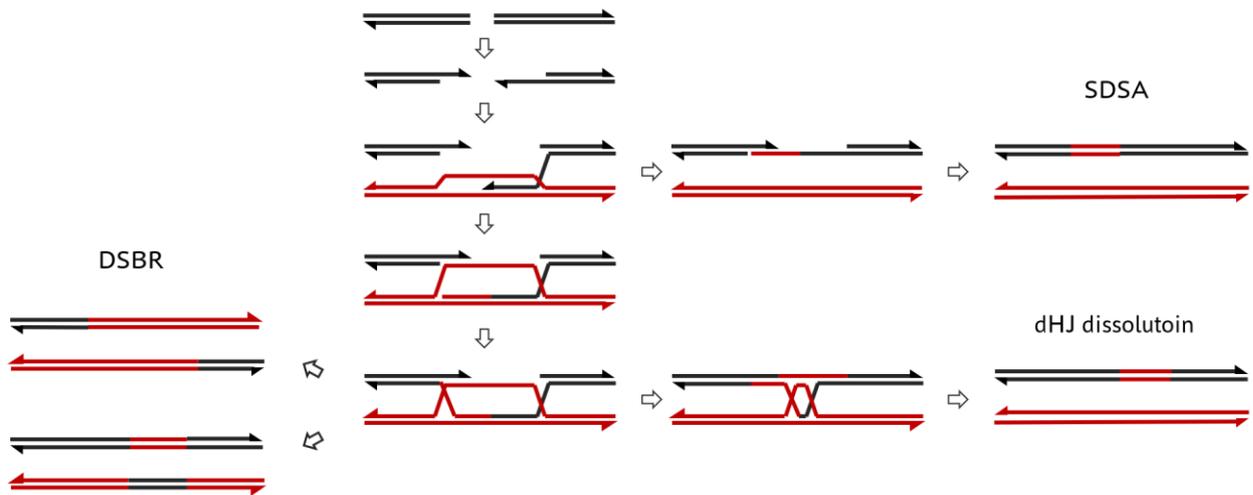


Рисунок 2.7 Схема путей репарации, направляемой гомологией (гомологической рекомбинации). Красным цветом показаны цепи ДНК донора гомологии, которым обычно выступает сестринская хроматида или гомологичная хромосома.

Если второй длинный 3'-конец разрыва захватывается D-петлей, то образуется двойная структура Холлидея (dHJ, double Holliday junction), которая должна будет потом разрезаться. Этот путь называется DSBR, от double-strand break repair — название связано с тем, что такой путь был предложен в 1983 году для объяснения молекулярного механизма кроссинговера, в отличие от объяснения, предполагавшего одноцепочечный разрыв ДНК [240]. Двойная структура Холлидея может быть разрезана с образованием кроссоверов — то есть с обменом плечами хромосом — и без образования кроссоверов. В последнем случае происходит лишь обмен фрагментами цепей ДНК в пределах небольшого участка, который был гетеродуплексом. Разрешение структур Холлидея в мейозе по пути образования кроссоверов осуществляется нуклеазами MSH4–MSH5 (MutSγ) и MLH1–MLH3 (MutLγ), тогда как в соматических клетках структуры «растворяются» без образования кроссоверов с помощью комплекса белков, включающего хеликазу BLM и топоизомеразу TOP3A [239], [241].

Запасной путь разрешения двойной структуры Холлидея, который осуществляется только в конце деления клеток и при наличии неразрешенных структур, предполагает простое разрешение с помощью белка GEN1, аналога резольваз бактерий или вирусов [242]. При использовании GEN1-зависимого механизма, могут получаться как некроссоверные продукты рекомбинации, так и кроссоверы. Если HDR проходит между сестринскими хроматидами, то образование кроссоверов не приводит к перераспределению генетической

информации, так как сестринские хроматиды идентичны. Однако образование кроссоверов может привести к хромосомным перестройкам, если HDR пройдет по эктопическим сайтам гомологии, находящимся на разных хромосомах. Такими сайтами могут быть повторяющиеся последовательности генома, например последовательности гомологичных мобильных элементов. Поэтому активность GEN1 в клетках человека регулируется ее доступом к ДНК — фермент имеет цитоплазматическую локализацию и может работать только в конце митоза при разрушенной ядерной оболочке [243]. Стоит отметить, что к образованию кроссоверов никогда не приводит SDSA, и поэтому именно этот путь характерен для HDR в соматических клетках [241][244].

Белки HDR

В начальных стадиях HDR задействованы следующие белки. Комплекс MRE11-RAD50-NBS1 (MRN) узнает разрыв ДНК и осуществляет процессинг концов разрыва. CtIP активирует у MRE11 эндонуклеазную активность, с помощью которой вносится ник в 20-40 нуклеотидах от конца ДЦР, а также 3'→5' экзонуклеазную активность, с помощью которой происходит удаление участка ДНК от ника до конца разрыва [245], [246]. Выступающий 3'-одноцепочечный участок ДНК покрывается белком RPA (Replication protein A). Затем белки BRCA2, BRCA1 (от Breast Cancer) и PALB2 (Partner and localizer of BRCA2) катализируют замещение RPA на белок RAD51 [247], [248]. RAD51 является ключевым белком, обеспечивающим поиск гомологичного партнера и инвазию в него цепи с образованием D-петли.

Важную роль в развитии ответа на ДЦР играет киназа ATM (Ataxia telangiectasia mutated), относящаяся к семейству фосфатидилинозитол-3-киназа-подобных киназ (PI3-киназ). Она привлекается субъединицей NBS1 комплекса MRN и фосфорилирует участвующие в репарации белки, что необходимо для их активности, а также запускает клеточные сигнальные пути ответа на повреждения ДНК, приводящие к аресту клеточного цикла. Так, ATM фосфорилирует p53 и чекпоинт-киназу CHK2 [249]. Одной из мишеней для ATM выступает вариантный гистон H2AX, который отличается от других форм гистона H2A несколькими аминокислотными остатками на С-конце, включающими серин 139, который и фосфорилируется. H2AX составляет несколько процентов от всех гистонов H2A и распределен по геному [250]. Волна фосфорилирования H2AX может распространяться по хроматину до 1 миллиона п.н. от точки ДЦР. Фосфорилированный H2AX (γ H2AX) привлекает к месту разрыва белки репарации, а также запускает сигнальные каскады в ответ на двуцепочечный разрыв [251].

Негомологичное соединение концов разрыва

Негомологичное соединение концов разрыва — путь репарации ДЦР, который преобладает у позвоночных в G1, хотя он активно задействуется и в остальных фазах клеточного цикла для репарации разрывов, не связанных с [252]–[254]. Этот путь используется при V(D)J-рекомбинации для сшивания фрагментов генома при формировании генов иммуноглобулинов и T-клеточных рецепторов. Данный путь является очень простым и состоит из следующих этапов: обнаружение разрыва, удержание концов, их процессинг при необходимости и их лигирование (рисунок 2.8). Обнаружение разрыва осуществляется комплексом KU70/80 (XRCC6/XRCC5, от X-ray repair cross-complementing protein 6/5, [255]). Это гетеродимер, мономеры которого имеют молекулярную массу 70 и 80 кДа. Гетеродимер образует кольцо, внутри которого находятся положительно заряженные остатки аминокислот. KU70/80 связывается неспецифически с концами ДНК, как тупыми, так и имеющими выступающие одноцепочечные участки ДНК («липкие концы») [256], [257]. KU70/80 привлекает к каждому концу разрыва субъединицу DNA-PKcs (DNA-dependent protein kinase, catalytic subunit). Две такие субъединицы димеризуются, удерживая концы разрыва вместе и формируя так называемый удаленный синаптический комплекс (long-range synaptic complex) [258]–[260]. DNA-PKcs также привлекает другие белки NHEJ, необходимые для процессинга концов разрыва и для их лигирования.

Белки, осуществляющие процессинг концов разрыва — полимеразы или нуклеазы. Такой нуклеазой выступает белок ARTEMIS, обладающий как эндонуклеазной, так и 5'→3' экзонуклеазной активностями [261], [262]. Именно эта нуклеаза разрезает шпильки, образующиеся в процессе разрезания ДНК при V(D)J-рекомбинации, и мутации в ее гене приводит к тяжелому комбинированному иммунодефициту [263],[264]. В качестве полимераз в NHEJ используются полимеразы семейства Pol X: полимеразы μ (POLM) и полимеразы λ (POLL), способные в дополнение к обычной полимеразной активности добавлять нуклеотиды нематрично [265]. Процессинг концов разрыва необходим, если ДЦР имеет липкие концы. Так, при выступающих 5'-липких концах ARTEMIS отрезает их для образования тупых концов, а при выступающих 3'-липких концах используется POLM. Липкие концы разрыва могут удерживаться друг за друга несколькими комплементарными нуклеотидами — так называемыми участками микрогомологии в 1–4 нуклеотида, что облегчает последующее лигирование. Процессинг концов разрывов приводит к тому, что после репарации в данном локусе возникают делеции и небольшие инсерции [266],[267].

Для завершения NHEJ DNA-PKcs привлекает дополнительные коровые факторы, которые позволяют прочнее связаться концам разрыва, формируя так называемый ближний синаптический комплекс (short-range synaptic complex). В данной конформации комплекса концы ДНК уже выравниваются для лигирования [268]. К необходимым для этого факторам

относятся XRCC4 (X-ray repair cross-complementing protein 4), XLF (XRCC4-like factor/Cernunos), и ДНК-лигаза IV [266],[269].

DNA-РКcs — белок, относящийся к семейству РlК-киназ. Связывание с KU70/80 активирует ее серин/треониновую киназную активность — DNA-РКcs автофосфорилируется и затем начинает фосфорилировать различные субстраты: KU70, KU80, XRCC4, XLF, ARTEMIS, ДНК-лигазу IV и гистон H2AX. Эта киназная активность DNA-РКcs очень важна для запуска работы перечисленных выше белков [270],[249]. Таким образом, автофосфорилирование является одним из регуляторов для завершения репарации путем NHEJ. В случае ингибирования автофосфорилирования DNA-РКcs, например, белком BRCA1, возможна сборка комплекса для репарации по пути HDR [245].

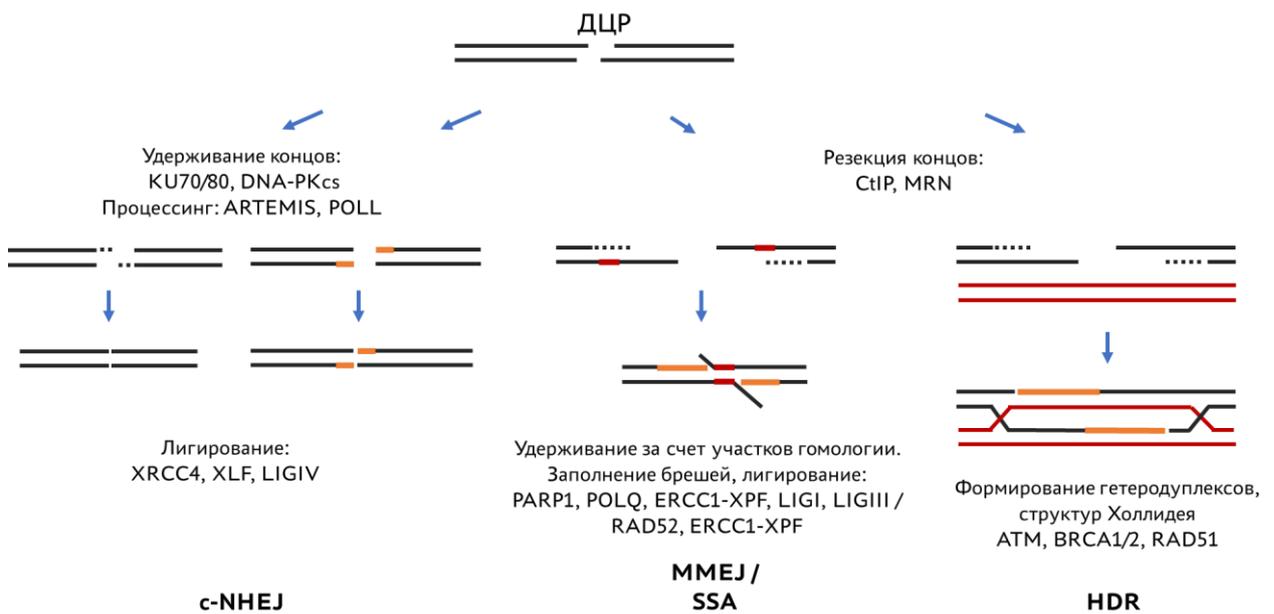


Рисунок 2.8. Общая схема путей репарации ДЦР в клетках. Указаны основные этапы и перечислены белки, участвующие в них. Оранжевым цветом показаны достраиваемые участки ДНК, красным — гомологичные участки ДНК, пунктиром — удаление нуклеотидов.

MMEJ и SSA

Помимо описанных путей репарации HDR и NHEJ в клетках существуют альтернативные пути репарации двуцепочечных разрывов. Так, было обнаружено, что даже при нокауте генов белков NHEJ все еще происходит репарация путем соединения концов разрыва [271]. Этот путь репарации назвали альтернативным NHEJ, или alt-NHEJ. В дальнейшем были расшифрованы детали этого механизма: первичное соединение концов разрыва осуществляется за счет микрогомологии в несколько нуклеотидов, для чего требуется резекция концов разрыва. Поэтому такой путь называют опосредованной микрогомологией соединением концов — MMEJ (microhomology-mediated end joining) [272]. Похожим образом происходит репарация по механизму отжига одноцепочечных

участков — SSA (Single strand annealing, SSA), только в нем используется гораздо большая по размеру область гомологии. Таким образом, MMEJ и SSA используют резекцию концов, аналогично HDR, хотя и не требуют последовательности-матрицы гомологии, что напоминает NHEJ. Очевидно, что в результате соединения концов этими способами в месте, где был ДЦР, происходят делеции (рисунок 2.8). Выбор между MMEJ и SSA, по-видимому, осуществляется в зависимости от интенсивности резекции концов разрыва. Обычно интенсивная резекция концов разрывов осуществляется в S- и G2-фазах клеточного цикла, где SSA может выступать запасным для HDR путем репарации разрывов.

MMEJ проходит без участия KU, DNA-PKcs, LIG4 и XRCC4, зато в нем участвуют поли(АДФ-рибоза)-полимераза PARP1, эндонуклеаза системы эксцизии нуклеотидов ERCC1/XPF (excision repair cross-complementation group 1 и xeroderma pigmentosum group F), полимеразы θ (POLQ) и лигазы I и III (LIG1 и LIG3) [273]–[275]. Участие в этом процессе особой полимеразы привело к тому, что данный путь еще называют тета-опосредованным соединением концов разрыва — TMEJ (theta-mediated end joining)[273]. Путь SSA задействует RAD52, обеспечивающий отжиг участков ДНК, и эндонуклеазу ERCC1/XPF [276].

Выбор пути репарации

Разветвление путей репарации ДЦР происходит на ранних этапах. С разрывом либо связывается Ku70-80, и запускается c-NHEJ, либо MRN, и запускается резекция концов. Однако переключение между путями репарации может происходить и уже после узнавания концов разрыва [245]. Белок 53BP1 вместе с белками RIF1 и PTIP тормозит резекцию концов, направляя репарацию по пути NHEJ. Белок BRCA1 является антагонистом 53BP1 и направляет репарацию по пути HDR или другим путям, связанным с резекцией (MMEJ, SSA). Показано, что инициация резекции эндонуклеазой MRE11 также направляет репарацию по путям, связанным с резекцией [277]–[279]. Поэтому путь репарации в случае разрывов, индуцированных ингибированием топоизомеразы, может зависеть и от пути удаления Top2cc. Важную роль в выборе пути играет хроматиновый контекст: показано, что в клетках, находящихся в G2, ДЦР, лежащие в неактивном хроматине репарируются преимущественно путем NHEJ, в то время как ДЦР в области активных генов — преимущественно HDR [280].

Вклад различных путей репарации ДЦР в образование транслокаций

HDR может привести к транслокациям только при эктопической рекомбинации между гомологичными участками на негомологических хромосомах. Такими участками могут быть многочисленные повторы — последовательности транспозонов. Однако вероятность встречи разрыва и гомологичного участка на другой хромосоме довольно низка вследствие

пространственной организации хромосом в хромосомные территории и другие пространственные структуры. Но даже в случае такой встречи и инвазии 3'-конца разрыва в такой участок, миграция ветвей структуры Холлидея по короткому участку гомологии будет невозможна, и рекомбинация нарушится [273]. Поэтому такие транслокации крайне редки и имеют значение лишь в эволюционном масштабе [281].

SSA может привести к межхромосомной транслокации, если концы разрывов потеряются и будут диффундировать в ядре. Но вероятность задействования именно этого механизма репарации гораздо ниже, чем других механизмов соединения концов разрывов, так как в случае SSA оба перестраивающихся локуса должны иметь протяженные участки гомологии, например Alu-элементы, хотя такое встречается [282].

Наиболее часто транслокации образуются в результате репарации по механизму NHEJ или MMEJ [283]. Репарация вызванных топоизомеразными ядами разрывов осуществляется этими путями: нокаун белков, участвующих в этих путях репарации, повышает чувствительность клеток к [284]–[286]. При V(D)-рекомбинации разрывы, вносимые комплексом RAG1/RAG2, в норме репарируются исключительно путем негомологичного соединения концов [287]. В настоящее время ведется дискуссия о том, какой из путей — канонический NHEJ или альтернативный — чаще приводят к транслокациям.

При изучении клеточных экстрактов из лейкозных клеток было показано, что соединение фрагментов ДНК происходит как классическим путем NHEJ, так и с использованием микрогомологии. И выбор того или иного пути определяется во-многом свойствами концов молекул ДНК, выступающих субстратами для репарации [288].

В другой работе в мышинных В-клетках с нокаутом *Ku70* и *Lig4* методом FISH наблюдали повышенное количество транслокаций локуса *IgH* после активации процесса переключения класса антител, а также повышенное число разрывов хромосомы, содержащей локус *IgH* [289]. Секвенирование нескольких областей перестроек показало, что соединение разрывов часто проходило по участкам микрогомологии, что характерно для MMEJ. Авторы предполагают, что раз в таких клетках выключен канонический путь NHEJ и при этом число транслокаций выше, то значит именно альтернативные пути соединения концов разрывов отвечают за формирование транслокаций.

Другие две работы были сделаны на специально созданной клеточной модели. Авторы оценивали число транслокаций, образующихся в результате внесения разрывов мегануклеазой I-SceI в два локуса. В геноме эмбриональных стволовых клеток мыши были встроены конструкции, несущие сайты узнавания I-SceI и фрагменты генов устойчивости к антибиотику. В случае образования транслокации клетки приобретали устойчивость, и количество транслокаций в культуре оценивалось по числу выросших колоний на чашке с антибиотиком. В случае нокаута с таких клеток генов *Xrcc4* (компонент лигазного

комплекса при NHEJ) или *Ku70*, количество транслокаций было выше [290]. В случае нокадауна CtIP уже наблюдалось уменьшение числа транслокаций по сравнению с контролями. Секвенирование точек перестройки в выросших клонах показало, что среди них реже встречается микрогомология и в целом более короткие участки делеций.

Однако в работах других авторов, выполненных на лимфоцитах мышей с помощью оценки числа микроядер, а также по анализу последовательностей сшивок генов иммуноглобулинов, было показано, что клетки-мутанты по *POLQ* более чувствительны к действию радиации и этопозиду, что указывает на роль MMEJ в репарации разрывов [291],[292].

В работе [293], где использовали уже клетки человека и вносили разрывы с помощью программируемых нуклеаз Cas9 (описание технологии ниже), напротив, показали снижение числа транслокаций при нокаутах генов белков c-NHEJ — *XRCC4* и *LIG4*. А области транслокаций не содержали участков микрогомологии и длинных делеций. При этом нокадаун *CTIP* и *LIG3*, необходимых для MMEJ, не приводил к снижению числа транслокаций.

Тем не менее, в работе с индукции разрывов ДНК радиацией в клетках HCT116 цитогенетическими методами было показано, что в G2 при ингибировании белков альтернативного пути число транслокаций уменьшается [294]. Однако авторы отмечают, что именно в G2, когда активнее происходит резекция концов, альтернативный путь соединения концов значительно чаще используется, чем в G1.

Наконец, в работе [295] авторы индуцировали в клетках человека два ДЦР на разных хромосомах с помощью двух программируемых нуклеаз ZFN (описание технологии ниже) и секвенировали как область стыка у транслокаций, так и неперестроившиеся целевые локусы. Авторы наблюдали многочисленные делеции и инсерции в области стыка, что связано с особенностями ZFN-индуцированных разрывов — разрывы формировались с липкими концами в 5 нуклеотидов — и согласуется с механизмом канонического NHEJ. Однако в 27% случаев в области транслокаций наблюдались участки микрогомологии длиной 2–4 нуклеотида, которые могут свидетельствовать о соединении путем MMEJ.

Можно подытожить, что и канонический, и альтернативный пути соединения концов разрыва оба задействованы в нормальных клетках человека при репарации ДЦР и могут являться причиной образования транслокаций.

2.5. Изучение факторов, влияющих на образование транслокаций

Хромосомные территории

Генетический материал клетки огромен, но при этом он должен транскрибироваться, поддерживаться в целостности и реплицироваться, что подразумевает его сложную организацию внутри ядра. За несколько десятилетий было получено множество свидетельств сложной организации ядра и наличия в нем субкомпарментов [296]. Выяснилось, что каждая хромосома в ядре занимает определенный объем и имеет свою радиальную позицию относительно центра ядра [297], [298]. С использованием многоцветной FISH с флуоресцентными зондами к целым хромосомам или спектрального кариотипирования (методы описаны в [10]) была четко показана организация хромосом в ядре в виде хромосомных территорий [299], [300], хотя границы между ними могут быть несколько размыты [301]. При этом гомологичные хромосомы как правило оказываются на противоположных полюсах ядра.

Биохимические методы изучения контактов участков генома между собой также демонстрируют предпочтительные *цис*-контакты участков хромосом. В частности, применяемые для изучения топологии укладки хромосом полногеномные C-методы (англ. Chromosome Conformation Capture), наиболее широко используемый из них — Hi-C [302], [303]. В методе Hi-C фиксированный хроматин сначала фрагментируют эндонуклеазами, а затем обрабатывают лигазой. В результате предпочтительно лигируются участки генома, сближенные в пространстве ядра, и частоту лигирования любых двух фрагментов генома можно детектировать тем или иным способом. Чем выше эта частота, тем чаще данные участки генома оказывались поблизости друг от друга в ядрах клеток или образовывали прямой контакт. В Hi-C используется глубокое секвенирование, что позволяет в одном эксперименте получить информацию о частотах лигирования в масштабе всего генома или внутри выбранного участка хромосомы, если предварительно фрагменты из этого участка были отобраны посредством гибридизации со специфическими зондами. Данные Hi-C визуализируют в виде тепловых карт, на которых области высоких значений сигнала соответствуют пространственным взаимодействиям, часто встречающимся в клеточной популяции. Анализ таких тепловых карт показал несколько уровней иерархии генома млекопитающих, в частности, что в клетках эукариот имеет место пространственная сегрегация хромосомных территорий [304]. Многочисленные *цис*-контакты участков хромосом хорошо заметны на Hi-C картах даже при невысоком разрешении. При увеличении разрешения на Hi-C картах заметны более низкие уровни организации, вплоть до отдельных хроматиновых петель. Детали организации хроматина рассмотрены в обзорах [305], [306].

Изучение механизмов образования хромосомных транслокаций

Чтобы произошла межхромосомная транслокация, концы разрывов разных хромосом очевидно должны оказаться рядом, хотя обычно они располагаются в своих хромосомных территориях. Были выдвинуты две гипотезы насчет того, как гены, находящиеся в разных хромосомных территориях, встречаются при транслокации [307], [308]. Согласно первой гипотезе, хромосомные перестройки возникают в ходе репарации ДЦР при ошибочном объединении изначально близко расположенных фрагментов разных хромосом. Эта модель образования хромосомных перестроек известна как гипотеза «первенства контакта» (contact first). Согласно второй гипотезе, «первенства разрыва» (breakage first), внесение разрывов является первоначальным событием, приводящим к повышенной мобильности концов разрывов и сближению генов-партнеров по транслокациям.

Пространственная предрасположенность локусов к транслокациям

В пользу влияния исходной пространственной близости локусов на то, какая транслокация произойдет, свидетельствуют результаты многих работ. Так, методом 3D FISH было показано, что в клетках щитовидной железы гены *H4* и *NCOA4* сближены с геном *PET* сильнее, чем можно было бы ожидать, отталкиваясь от их взаимной локализации на хромосоме и сравнивая с другими генами на этой хромосоме. Изучаемые гены формируют между собой перестройки (в данном случае крупные инверсии), которые приводят к раку щитовидной железы у пациентов, получивших высокую дозу радиации, что наводит авторов на мысль об их пространственной предрасположенности к перестройке [309], [310]. В другой работе исследователи попытались обнаружить закономерности уже на статистической выборке: методом крио-FISH с высоким разрешением были определены степени смешивания хромосомных территорий между собой. Полученные значения сравнивали с числом транслокаций, возникающих между соответствующими хромосомами после облучения лимфоцитов дозой 3–6 Гр [311], и увидели положительную корреляцию [301]. И наоборот, в работе [312] методом 3D FISH сравнивали, насколько глубоко в своих хромосомных территориях находятся гены, участвующие в межхромосомных и внутрихромосомных перестройках при раке щитовидной железы. Выяснили, что участвующие во внутрихромосомных перестройках гены *RET*, *NTRK* и *BRAF* располагаются в глубине территории, а участвующие в межхромосомных транслокациях гены *PAX8*, *PPARG* и *THADA* находятся у поверхности их хромосомных территорий. Эти результаты, полученные на основе микроскопии, могут привести к выводу, что чем ближе гены друг к другу, тем скорее между ними будут происходить хромосомные перестройки после возникновения разрывов.

Однако в исследовании [313], также проведенным с помощью 3D FISH на линиях лимфоцитов и миелоидных клеток человека, напрямую было показано, что не всегда

частота контактов коррелирует с частотой транслокаций (по крайней мере, наблюдаемой в клинической практике). Так, во всех четырех клеточных линиях среднее расстояние между *MLL* и *ENL* было меньше, чем между *MLL* и *AF4*, при этом транслокация *MLL-AF4* значительно более распространена, чем *MLL-ENL*. Напомним, что молекулярный механизм онкогенеза у обеих этих транслокаций одинаков — и *AF4*, и *ENL* относятся к компонентам суперэлонгационного комплекса — что не позволяет объяснить повышенную частоту одной перестройки относительно другой клональным преимуществом.

Работы, выполненные с помощью C-методов, также показывают важность пространственной близости для образования перестроек. Так, в работе [314] использовали метод высокопроизводительного скрининга транслокаций (HTGTS — High-Throughput Genomic Translocation Sequencing) и Hi-C для анализа контактов генов. На основе мышиных про-B-клеток были созданы клеточные линии, в геном которых была введена кассета, содержащая многочисленные сайты для *I-SceI*. Причем клетки были нокаутами по гену киназы *ATM*. Индуцируя экспрессию *I-SceI*, провели анализ спектра транслокаций и обнаружили, что они происходят предпочтительно с сайтами *RAG*-нуклеазы. Чтобы исключить это влияние, на клетки воздействовали ионизирующим излучением. В итоге было показано, что существует корреляция между частотой перестроек между локусами и частотой их контактов, особенно четко это проявлялось для *цис*-перестроек и *цис*-контактов.

В другой работе, основанной на анализе данных Hi-C клеток лимфобластоидной линии и линии K562, сравнивали пространственную близость генов, которые известны уже из клинической практики своими перестройками [315]. Были проанализированы контакты полутора тысяч генов-партнеров по транслокациям. Показано, что многие пары генов-партнеров по транслокациям действительно имеют повышенную частоту контактов. К ним относятся такие известные пары, как *BCR-ABL* и *MYC-IGH*. Однако авторы лишь подмечают тенденцию, но не говорят, что это абсолютное правило — так, для некоторых транслокаций, например *AML-ETO*, повышенной частоты контактов замечено не было.

Мобильность разрывов ДНК

Наблюдение за двуцепочечными разрывами ДНК часто демонстрировало их подвижность в пространстве ядра. Так, в работе [316] наблюдали за распределением фосфорилированной формы гистона H2AX в ядрах клеток HeLa после облучения линейным пучком альфа-частиц. Спустя 60 минут после облучения фокусы гамма-H2AX отклонялись от своего линейного расположения, и в значительной доле клеток (до 50%) образовывали кластеры. В другой работе, где визуализировали в живых клетках слитый с RFP белок 53BP1, связывающийся с ДЦР, после обработки фибробластов и клеток карциномы γ -лучами наблюдали кластеризацию фокусов 53BP1-RFP, а также вытеснение кластеров в менее

плотные участки хроматина (хроматин визуализировали гистоном H2B, слитым с GFP) [317]. Кластеры часто колокализировались с комплексом MRN, который участвует в репарации ДЦР. Примечательно, что некоторые кластеры оставались в клетках в течение даже нескольких дней, и этот эффект был сильнее при большей дозе облучения. Кластеризацию фокусов 53BP1, возникающих после воздействия на клетки ионизирующим излучением, наблюдали также и в работе [318].

Однако изучение пространственной динамики двуцепочечных разрывов в живой клетке не всегда демонстрирует высокую подвижность разрывов. В работе [319] в геном мышинных фибробластов NIH3T3 были встроены кассета операторов тетрациклинового оперона и кассета операторов лактозного оперона, между которыми помещался сайт нуклеазы I-SceI. Экспрессия CFP-меченного Lac-репрессора и YFP-меченного Tet-репрессора позволяла наблюдать за подвижностью локусов в реальном времени, однако при индукции экспрессии нуклеазы I-SceI метки CFP- и YFP- удерживались вместе, хотя 80–90% локусов и содержали сигнал γ H2AX. Разделения концов разрыва удалось добиться лишь при нокдауне *KU80*. Развитием данной работы стало наблюдение за разрывами уже на разных хромосомах [320]. В одну хромосому 7 встроили сайт I-SceI и рядом с ним кассету LacO. В две хромосомы 10 были встроены сайты I-SceI, фланкированные двумя кассетами TetO. Использование высокопроизводительной микроскопии позволяло вести наблюдение за фокусами в тысячах клеток. До трансфекции I-SceI доля сближенных фокусов составляла ~2%. Спустя 36 часов после трансфекции I-SceI флуоресцентные фокусы сближались в 12% клеток, и в 0,3% клеток возникали хромосомные транслокации, что было определено с помощью количественной ПЦР.

Далее авторы проанализировали динамику фокусов в течение 24 часов в клетках, в которых в итоге образовывались транслокации. Критерием образования транслокации было сближение фокусов, после которого фокусы двигались как единое целое до конца наблюдения. Более 80% локусов, которые в итоге дали транслокацию, находились ближе в пределах 1,5 мкм за 4 часа до образования транслокации. Однако в 10% случаев транслокации формировали локусы, расположенные более чем за 3 мкм друг от друга. Эти наблюдения позволяют заключить, что вероятность образования транслокации между разрывами снижается с увеличением расстояния между ними.

В этой работе также не удалось обнаружить повышенной мобильности двуцепочечных разрывов по сравнению с интактным хроматином. Как и в предыдущей работе [319], авторы не видели расхождения концов разрыва ни в одной из тысяч клеток. Ингибирование DNA-ПКcs увеличивало частоту транслокаций, однако не влияло на подвижность разрывов.

В другой работе, где визуализировали с помощью 53BP1-GFP в живой клетке разрывы, вызванные индуцируемой нуклеазой, наблюдали кластеризацию разрывов [321].

В работе использовали линию клеток человека D1vA с индуцируемой нуклеазой AsiSI, которая вызывает порядка сотни разрывов при индукции системы. Авторы ингибировали киназы ATM и DNA-PKcs, и в итоге делают вывод, что ATM может быть белком, отвечающим за кластеризацию, а DNA-PKcs более необходима для репарации разрывов, так как при ее ингибировании существенно повышается чувствительность клеток к разрывам.

Противоречивые результаты наблюдения за динамикой разрывов могут объясняться разным хроматиновым контекстом разрезаемых локусов. Авторы предыдущей работы в новом исследовании [322] попытались проверить данное предположение. В работе вновь использовали линию клеток D1vA. Был проведен анализ контактов разрезаемых локусов с помощью метода capture Hi-C (модификация протокола Hi-C, направленная на анализ контактов отдельных областей генома со всеми остальными локусами). Было показано, что после индукции нуклеазы разрезаемые локусы начинают чаще контактировать с другими разрезаемыми локусами. При этом локусы, которые показывали высокий уровень кластеризации, показывают наиболее высокий уровень транскрипционной активности, который определяли иммунопреципитацией хроматина по присутствию РНК-полимеразы и триметилированного H3K36.

Динамика разрывов может отличаться и вследствие разных способов их индукции. Так, в работе [323] изучали мобильность различных участков хроматина, в том числе фокусов разрывов, наблюдая за 53BP1-GFP в живых клетках. Разрывы индуцировались либо ионизирующим излучением дозой 5–8 Гр, либо этопозидом (15 мин 15 мкг/мл). По сравнению с интактными локусами разрывы показали более высокую динамику, особенно те, что индуцировались этопозидом. Интересно, что авторы смогли уменьшить подвижность разрывов, обрабатывая клетки ингибитором гистонацетилтрансфераз циркумином, тем самым подтверждая важность хроматинового контекста для определения свойств разрыва.

Изучение методом FISH клеток, обработанных этопозидом, также свидетельствует в пользу повышения подвижности перестраивающихся генов. Так, обработка клеток линии Nalm-6 этопозидом приводит к увеличению частоты колокализации генов *AML1* и *ETO* [203]. Также об увеличении частоты колокализации генов *AML1* и *ETO* в результате обработки этопозидом клеток линии Jurkat говорят результаты работы [324]. В другом исследовании, проведенном в нашей лаборатории, методом 3D FISH показали расхождение концов разрывов гена *MLL*, индуцированных этопозидом [325]. При этом ген *MLL* после обработки клеток этопозидом чаще локализовался за границей территории 11 хромосомы, что говорит о возможной подвижности концов разрывов этого гена, индуцированных этопозидом. Напротив, локализация гена *CCND1*, который также расположен на 11 хромосоме, но не замечен в транслокациях, не изменялась после обработки этопозидом.

Предрасположенность локусов к разрывам

Очевидным фактором, который определяет вероятность возникновения транслокаций, является предрасположенность перестраивающихся локусов к разрыву. Так, в работе [326] убедились, что если вносить только один разрыв в *MLL* — для этого использовали нуклеазу с цинковыми пальцами — то транслокации не детектируются.

В двух очень похожих работах [327], [328] использовали высокопроизводительное секвенирование транслокаций, которые индуцировались в мышинных В-лимфоцитах с помощью *I-SceI*. Обычно партнерами по транслокациям выступали сайты узнавания дезаминазы *AID*, которая в лимфоцитах вносит разрывы при переключении класса антител. В случае же нокаута *AID* транслокации происходили с локусами, содержащими сайты, напоминающие сайты *I-SceI*, и наблюдалось перераспределение спектра перестроек в пользу *транс*-перестроек. Также авторы обеих работ замечают, что транслокации происходят предпочтительнее между активно транскрибируемыми участками. Однако это может быть объяснено тем, что *AID* работает только в транскрибируемых участках. В любом случае, авторы показывают, что транслокации случаются именно между локусами, которые подвержены образованию разрывов.

Авторы одного из описываемых выше исследований дополнили предыдущую работу 4C-анализом (один из C-методов, позволяющий проанализировать в популяции клеток контакты интересующих локусов со всеми участками генома) и сравнили его результаты с данными по транслокациям [329]. Было показано, что в клетках, в которых экспрессируется *AID*, транслокации происходят между локусами, в которых *AID* вносит разрыв, а для нокаутных по *AID* клеток спектр транслокаций начинает коррелировать с профилем контактов. Авторы делают вывод, что если вносятся разрывы, не направленные на отдельные локусы, то перестройки определяются скорее взаимным расположением генов-партнеров, в то время как при направленном внесении разрывов в локусы вероятность перестроек будет определяться в первую очередь вероятностью разрыва в этих локусах.

Обработка клеток этопозидом, с одной стороны, вызывает множественные разрывы в геноме, с другой стороны, результаты работ демонстрируют, что сайты, чувствительные к разрывам, вызванным ингибированием топоизомеразы, распределены по геному не случайно. Так, *END-seq* (метод, направленный на определение точек разрывов в популяции клеток, основанный на высокопроизводительном секвенировании) показал, что разрывы под действием этопозида возникают часто в основаниях петель, на границах топологически ассоциированных доменов (ТАДов) хроматина [330], при этом такие «хрупкие» сайты совпадают с точками разрывов, наблюдаемыми у пациентов с лейкозами. Более того, в следующей работе авторы показали, что сайты ДЦР, индуцируемых этопозидом, можно предсказывать по местам связывания когезина и участкам активной транскрипции [207].

2.6. Создание клеточных моделей

Для изучения различных процессов, протекающих в клетке, иногда создают клеточные модели с заданными свойствами. Несколько таких клеточных моделей, применяемых для изучения транслокаций, было описано выше. Кроме того, клеточные модели можно использовать для высокопроизводительного скрининга потенциальных лекарств. Благодаря четко измеряемому свойству модели и возможности автоматизации процесса проводят скрининг целых библиотек химических веществ с целью обнаружения молекулы с необходимым свойством [331]. Часто для создания клеточной модели необходимо генетически модифицировать существующую клеточную культуру. Эта процедура стала доступной с развитием простых технологий редактирования генома.

Технологии редактирования генома

В настоящее время в лабораторных исследованиях используются технологии редактирования генома, основанные на трех системах: CRISPR/Cas (Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein), TALEN (Transcription activator-like effector nuclease) и ZFN (Zinc-finger nuclease) (рисунок 2.9). Эти системы позволяют направленно вносить ДЦР в заданное место генома, что является основой соответствующих технологии редактирования генома [332]. Репарация данного разрыва по пути негомологичного соединения концов часто приводит к инсерциям или делециям, что может использоваться для получения генного нокаута. С другой стороны, введение в клетку матрицы для гомологичной рекомбинации одновременно с осуществлением разрыва позволяет редактировать целевую последовательность или интегрировать в геном новые конструкции.

Нуклеазы ZFN

Из рассматриваемых технологий первыми были открыты ZF-нуклеазы. ZFN — химерные белки: С-концевая часть представляет собой нуклеазный домен эндонуклеазы *FokI*, а на N-конце содержится несколько доменов цинковых пальцев C2H2-типа, каждый из которых имеет размер около 30 аминокислот и распознает последовательность из трех нуклеотидов [333], [334]. Для эффективного внесения ДЦР необходима димеризация эндонуклеазного домена *FokI*, поэтому ZFN конструируются так, чтобы узнавать два полусайта, которые фланкируют желаемую точку разрыва [335]. Но несмотря на то, что почти для каждой комбинации из трех пар нуклеотидов существует распознающий ее ZF-модуль, ZFN не всегда работают эффективно — объединение хорошо характеризованных модулей часто не приводит к созданию белка, надежно распознающего заданную последовательность [336]. Поэтому технологии ZFN в настоящий момент полностью заменяются технологиями на основе TALE и CRISPR/Cas.

Нуклеазы TALEN

Следующий прорыв в редактировании генома был связан с расшифровкой кода TALE-белков. Эти белки являются факторами вирулентности бактерий рода *Xanthomonas sp.*, активирующими транскрипцию ряда генов в клетках инфицированного растения. Белки TALE построены из повторяющихся блоков по 34 аминокислоты, и каждый такой блок узнает один нуклеотид в молекуле ДНК за счет двух переменных аминокислот [337]. Путем слияния нескольких мотивов TALE с нуклеазным доменом FokI удалось получить эффективные сайт-специфические нуклеазы [338]. Как и ZFN, TALEN работают в составе димера, узнавая два полусайта, расположенных на расстоянии 12–21 нуклеотидов друг от друга [339]. По сравнению с ZFN получение TALE-нуклеаз заданной специфичности является более простым процессом, поскольку один домен TALE распознает один нуклеотид, а не три. Кроме того, в отличие от ZFN, можно получить работающие TALE для распознавания любых комбинаций нуклеотидов.

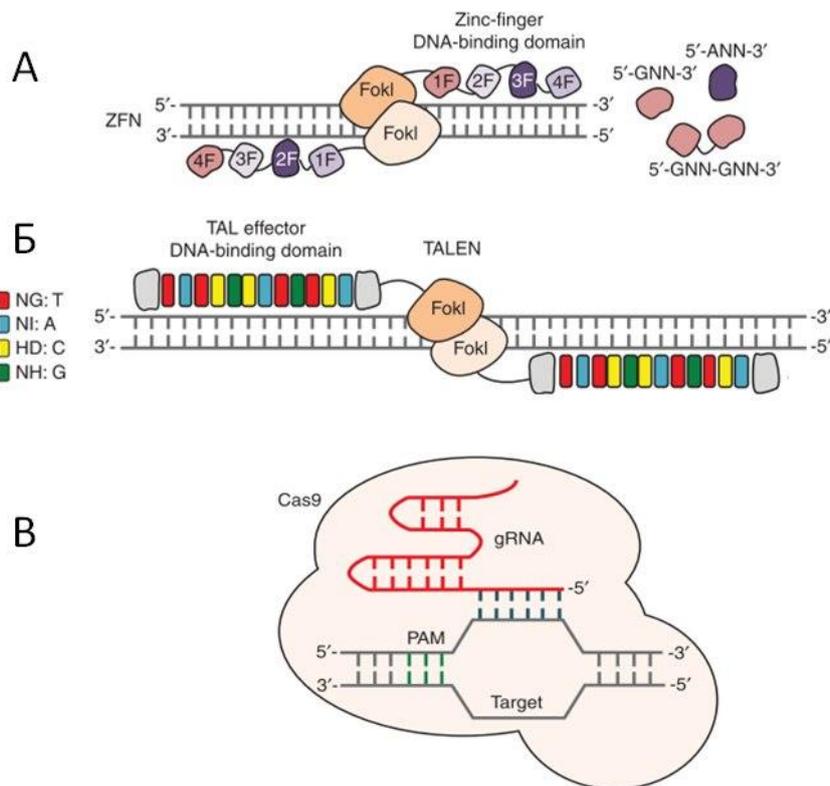


Рисунок 2.9. Нуклеазы, на которых основаны технологии редактирования генома. **А** — внесение разрыва в ДНК с помощью ZFN. Показаны отдельные домены цинковых пальцев и узнаваемые ими последовательности ДНК. FokI — нуклеазный домен в ZFN и TALEN. **Б** — TALEN. Разными цветами отмечены TALE-домены, каждый из которых узнает один нуклеотид в ДНК. **В** — нуклеаза Cas9 в комплексе с химерной гидовой РНК (gRNA+tracrRNA) и ДНК-мишенью. Рисунок из [332].

Система CRISPR/Cas

В настоящее время среди технологий редактирования генома наиболее широко используется система CRISPR/Cas, получившее свое развитие относительно недавно. В 1987 году были обнаружены особые последовательности в бактериальных геномах, состоящие из уникальных последовательностей (спейсеров), перемежающихся повторами. Их назвали CRISPR (clustered regularly interspaced short palindromic repeats) — короткие палиндромные повторы, регулярно расположенные группами [340]. В 2005 году выяснилось, что спейсеры соответствуют последовательностям вирусной ДНК. Было также замечено, что с CRISPR ассоциированы гены, кодирующие нуклеазы. Тогда было сделано предположение, что бактерии используют эти нуклеазы для уничтожения чужеродной ДНК, при этом последовательность спейсеров используется в качестве примера того, какую ДНК необходимо разрезать [341], [342], что вскоре было доказано экспериментально [343]. Таким образом, система CRISPR/Cas представляет собой прокариотический аналог адаптивного иммунитета, который направлен на борьбу с бактериофагами и другими эгоистичными молекулами ДНК [344]. Сам механизм созревания комплекса нуклеаза-РНК и направленного расщепления ДНК был подробно изучен к 2011 году [345]. Было выяснено, что фермент образует тройственный комплекс с двумя небольшими РНК (crRNA и tracrRNA), закодированными в кластере. Этот комплекс расплетает ДНК и, если crRNA комплементарно связывается с одной из цепей ДНК, нуклеаза вносит двуцепочечных разрыв. Также было показано, что crRNA и tracrRNA могут быть объединены в одну так называемую гидовую (направляющую) РНК, и что данную конструкцию можно использовать для направленного внесения разрывов, а следовательно, и для редактирования генома эукариот [346]–[349].

Эффекторным элементом системы CRISPR/Cas II типа является нуклеаза Cas9. Чтобы Cas9 смогла внести разрыв, в ДНК-мишени должна присутствовать последовательность PAM (Protospacer adjacent motive). PAM должна примыкать к 3'-концу последовательности, распознаваемой гидовой РНК. Cas9 вносит двуцепочечный разрыв, отступив 3–4 нуклеотида от PAM в направлении 5'-конца цепи ДНК, в которой находится PAM [350]. По сравнению с TALE-нуклеазами система CRISPR/Cas9 является более простой в применении: программирование специфичности нуклеазы достигается подбором подходящей 20 нуклеотидной узнающей последовательности гидовой РНК и не требует белковой инженерии. Для облегчения подбора последовательностей РНК-гидов существуют онлайн-инструменты, а для сборки генов — плазмидные конструкции, в которые необходимо лишь вставить олигонуклеотид, соответствующий узнающему участку РНК-гида.

Индукция хромосомных перестроек

Для формирования хромосомных перестроек можно индуцировать разрывы искусственно с помощью мегануклеаз или программируемых нуклеаз. Помимо описанных выше работ, упомянем еще несколько примеров клеточных моделей с перестройками, индуцированными с помощью нуклеаз. Так, с помощью системы CRISPR/Cas9 в клетках мышинных фибробластов NIH3T3 удалось получить инверсию участка 17 хромосомы, которая приводила к образованию слитого онкогена *Eml4-Alk* [351]. Данная перестройка часто обнаруживается при немелкоклеточном раке легкого. Более того, инфекция мышей аденовирусом, несущим слитый ген, приводила к развитию опухолей легких, похожих по фенотипу на немелкоклеточный рак легкого человека. В другой работе с помощью системы CRISPR/Cas9 удалось получить линии клеток HEK293A, несущих транслокации *FLI1-EWSR1* или *AML1-ETO* в геноме [352]. Полученные линии клеток могут быть использованы как модели саркомы Юинга и острого миелоидного лейкоза. Тем не менее, являясь удобными моделями для тестирования терапевтических подходов к лечению, такие клеточные модели не позволяют исследовать ранние события, происходящие при возникновении хромосомных перестроек.

3. Материалы и методы

3.1. Работа с бактериями

Приготовление компетентных клеток

Для проведения трансформации готовили компетентные клетки *E. coli* штамма Top10. Штамм Top10 хранился в замороженном виде при $-80\text{ }^{\circ}\text{C}$. Вначале проводили посев штрихом замороженных клеток на чашку с агаризованной средой SOB (2% w/v триптона, 0,5% w/v дрожжевого экстракта, 8,5 mM NaCl, 2,5 mM KCl, 10 mM MgCl₂, 1,5% w/v агара). После появления колоний на чашке проводился смыв культуры в 250 мл жидкой среды SOB. Культура инкубировалась в шейкер-инкубаторе Innova 40 (New Brunswick Scientific) при 200 об/мин при $20\text{ }^{\circ}\text{C}$ до достижения оптической плотности в 0,3 единицы. Определение оптической плотности проводилось с помощью спектрофотометра «BioPhotometer Plus» (Eppendorf) при длине волны 600 нм. Затем культура осаждалась центрифугированием при 2500 об/мин при $4\text{ }^{\circ}\text{C}$ на центрифуге «5804 R» (Eppendorf). Осадок промывался стерильным холодным буфером ССМВ80 (10 mM ацетат калия, 80 mM CaCl₂, 20 mM MnCl₂, 10 mM MgCl₂, 10% v/v глицерин, pH 6,4) и повторно центрифугировался в тех же условиях. Осадок снова ресуспендировался в холодном буфере ССМВ80 и полученная суспензия замораживалась на $-80\text{ }^{\circ}\text{C}$ аликвотами по 50 мкл. Компетентность оценивали по результату трансформации 10 пг плазмиды pUC-18 одной аликвоты клеток. Достаточной принимали компетентность 10^{-6} .

Проведение трансформации компетентных клеток

Аликвота 50 мкл компетентных клеток *E. coli* штамма Top10 размораживалась на льду, после чего к ней добавлялось необходимое количество плазмиды (от 1 до 10 нг). Клетки инкубировались на льду в течение 30 минут, и затем подвергались тепловому шоку: 1 минута при $42\text{ }^{\circ}\text{C}$ на водяной бане (Biosan). После этого к клеткам добавлялось 5 объемов среды SOC, которая представляет собой среду SOB с добавлением глюкозы до 20 mM. Суспензия инкубировалась 1 час при $37\text{ }^{\circ}\text{C}$ в термомиксере Thermomixer Compact (Eppendorf) при 550 об/мин. 20 мкл суспензии засеивалось на чашку со средой LB (1% w/v триптона, 0,5% w/v дрожжевого экстракта, 170 mM NaCl, 1,5% агара) с добавлением соответствующего антибиотика: ампициллин («Синтез») в конечной концентрации 100 мкг/мл или канамицин («Панэко») в конечной концентрации 25 мкг/мл. Колонии бактерий выращивались в течение ночи в инкубаторе при $37\text{ }^{\circ}\text{C}$.

3.2. Культуры клеток эукариот

В работе использовались перевиваемые линии клеток человека: прикрепленная культура HeLa (карцинома шейки матки), суспензионная культура Jurkat (Т-лимфоциты) и суспензионная культура LCL (другое название — RPMI-8866, клетки лимфобластоидной линии). Клетки HeLa культивировали в пластиковых культуральных чашках в среде DMEM («ПанЭко») с добавлением до 10% объема эмбриональной телячьей сыворотки (Hyclone), а также антибиотиков пенициллина в конечной концентрации 50 ед/мл («ПанЭко») и стрептомицина в конечной концентрации 50 мкг/мл («ПанЭко»). Культуры Jurkat и LCL культивировали в культуральных матрасах в среде RPMI-1640 («ПанЭко», или Gibco) с добавлением до 10% объема эмбриональной телячьей сыворотки (Hyclone или FBS certified, Gibco), а также антибиотиков пенициллина в конечной концентрации 50 ед/мл и стрептомицина в конечной концентрации 50 мкг/мл. Среда с добавленными антибиотиками и сывороткой далее в тексте будет называться полной.

Культивировали клетки в инкубаторе (Binder) при 37°C во влажной атмосфере с содержанием 5% CO₂. Плотность клеточной суспензии поддерживали в пределах 0,2–1 млн/мл. В экспериментах использовали клетки, находящиеся в экспоненциальной стадии роста, с плотностью 0,5–1 млн/мл. Плотность клеточной суспензии оценивали под микроскопом с помощью камеры Горяева.

3.3. Трехмерная флуоресцентная *in situ* гибридизация (3D FISH)

Обработка клеток ингибитором топоизомераз

Суспензию клеток, находящихся на логарифмической стадии роста (0,7–0,9 млн/мл), обрабатывали ингибитором топоизомераз: добавляли этопозид (раствор 20 мг/мл в DMSO) до конечной молярной объемной концентрации 0,17 мкМ (100 мкг/мл) и инкубировали клетки в течение 1 часа. Затем отмывали от среды с ингибитором (осаждали клетки при 700 g 4 мин, удаляли среду, ресуспендировали осадок клеток в свежей среде) и оставляли в свежей полной среде на 10–60 мин. Затем клеткам меняли среду на бессывороточную из расчета 1,5 млн/мл среды, наносили по 0,5 мл на покровные стекла, размещенные в пустом 6-луночном планшете в клеточном инкубаторе и предварительно обработанные согласно инструкции производителя моллюсковым клеем CellTak (Corning). Количество клеток на покровном стекле было подобрано так, чтобы они равномерно в один слой покрывали покровное стекло. Клетки, используемые в качестве контроля, подвергали всем тем же процедурам, но вместо ингибитора топоизомераз к ним добавляли соответствующее количество DMSO.

Облучение клеток рентгеновским излучением

Для облучения клеток использовали прибор Radgil (Gilardoni) — специальную медицинскую рентгеновскую установку для облучения компонентов и препаратов крови. Использовалась программа №10, клетки находились в полной среде в концентрации 0,8 млн/мл и облучались в установке 13 мин — 25 Зв (около 7 Гр). Спустя час клеткам меняли среду на бессывороточную и сажали полчаса в клеточном инкубаторе на покровные стекла, обработанные моллюсковым клеем CellTak по описанному выше протоколу.

Приготовление препаратов клеток

Все замены растворов производили аккуратно, чтобы не смыть клетки с покровных стекол. Растворы удаляли водоструйным насосом, и сразу добавляли следующий раствор, чтобы не высушить клетки. Чтобы смыть не прикрепившиеся клетки, добавляли к стеклам с клетками 0,3X PBS. Заменяли 0,3X PBS на раствор 4 % формальдегида в 0,3X PBS и инкубировали в течение 10 мин при комнатной температуре. Дважды промывали по 5 мин в 1X PBS. Далее стекла обрабатывали раствором 0,5 % Тритон X-100 в 1X PBS в течение 10 мин, дважды промывали по 5 мин в 1X PBS и выдерживали в 20 % растворе глицерина в 1X PBS в течение ночи при 4 °С. Затем стекла с клетками подвергали 4-х кратному замораживанию в жидком азоте и оттаиванию в 20 % растворе глицерина в 1X PBS. Инкубировали 20 мин в 0,1 М HCl при комнатной температуре, дважды промывали по 5 мин в 1X PBS. Обработывали препараты раствором РНКазы А (200 мкг/мл в 2X SSC) в течение 30 мин при 37 °С, дважды промывали по 5 мин в 2X SSC. Затем уравнивали в 50 % растворе деионизованного формамида в 2X SSC минимум в течение суток при 4 °С. До момента гибридизации клетки хранили в этом же растворе при 4 °С.

Гибридизация с ДНК-зондами

Начало и конец гена *AML1* визуализировали с помощью флуоресцентно меченых бактериальных искусственных хромосом (BlueGnome). ВАС RP11-299D9 перекрывала 140 тыс.п.н. и была помечена с помощью AlexaFluor 488, а ВАС RP11-177L11 перекрывала 160 тыс.п.н. и была помечена с помощью Cy3 оранжевого флуорофора. Хромосомная территория 21 визуализировались с помощью зонда ID Labs, помеченную флуорофором IDYE™ 647 (дальний красный спектр флуоресценции). Для денатурации ДНК покровное стекло с клетками помещали в раствор 70 % деионизованного формамида в 2X SSC, нагретого до 75 °С на термоблоке амплификатора Perkin Elmer и выдерживали в течение 15 мин. Готовили гибридизационную смесь из расчета на 2 покровных стекла: 1,5 мкл флуоресцентного зонда (пробы) на хромосомную территорию, 1 мкл двуцветной пробы на исследуемый ген, 5 мкл буфера из комплекта к зондам на хромосомные территории, mQ до общего объема 10 мкл. Денатурировали смесь зондов термоциклере «Bio-Rad» при 75 °С в

течение 5–7 мин. На предварительно разогретое на термостате до 60 °С предметное стекло наносили 5 мкл гибридизационной смеси и помещали сверху покровное стекло клетками вниз. Герметизировали препарат резиновым клеем, переносили в темную влажную камеру и инкубировали в термостате при 37 °С в течение 36–48 часов.

Для визуализации гена *ETO* использовались флуоресцентно меченые бактериальные искусственные хромосомы (BlueGnome) ВАС RP11-643O11 (начало гена) и ВАС RP11-118O8 (конец гена).

Постгибридизационные отмывки и монтирование препаратов

Покровные стекла с клетками отлепляли от предметных стекол и отмывали в растворах 0,3 % NP-40 в 0,4X SSC в течение 2 мин при 70° С, затем в 0,1 % NP-40 в 2X SSC 1 мин при комнатной температуре. Один раз промывали стекла 2X SSC в течение 5 мин и один раз 5 мин в 1X PBS. Монтировали покровные стекла с клетками на предметные, используя среду для заключения препаратов Dako fluorescent mounting medium (ДАКО). Хранили препараты в темноте при 4 °С.

Конфокальная микроскопия препаратов

Для получения многослойных изображений препаратов использовали метод лазерной сканирующей конфокальной микроскопии. Для возбуждения флуорофоров использовались: 488 нм аргоновый лазер для возбуждения флуорофоров AlexaFluor 488 и SpectrumGreen, 543 нм гелиево/неоновый лазер для возбуждения флуорофоров Cy3 и SpectrumOrange (гибридизационные зонды на *AML1* и *MLL*, соответственно), 633 нм гелиево/неоновый лазер для возбуждения флуорофора IDYE™ 647, маркирующего хромосомные территории. Использовали 63-кратный объектив. Сканирование велось при разрешении 1024, каждая точка сканировалась дважды, при этом значение сигнала усреднялось, скорость сканирования — 1,6 мкс на точку. Динамический диапазон сканирования — 12 бит. Шаг между слоями изображений составлял 300–500 нм, одно изображение состояло из 20–30 оптических слоев, высота стопки слоев 10–12 мкм. Нижний и верхние слои задавались вручную для каждого поля сканирования так, чтобы между ними находилась большая часть ядра.

Компьютерная обработка и анализ полученных изображений

Для анализа полученных снимков использовалась программа, разработанная Д.А. Алексеевским (Отдел математических методов в биологии, НИИ ФХБ имени А.Н. Белозерского) специально для наших целей. Программа получает на вход изображение, состоящие из конфокальных слоев и набор настроек, далее программа выполняет следующие шаги: предобработка снимков, детектирование сигналов от пробы на *AML1* (зелёный и оранжевый цвета), детектирование хромосомных территорий (дальняя

красная область спектра) и ядер, нормирование окрестностей сигналов, выдача статистики. В качестве результата программа выдает идентификаторы всех детектированных сигналов, территорий, ядер, их размер и координаты их центров, расстояния между сигналами, расстояния до ближайшей границы территории. Пример детектирования — на рисунке 4.1.

Предобработка снимков

Предобработка сводится к формированию в памяти компьютера изображения в виде трёхмерной матрицы значений яркости пикселей для каждого из каналов. Для снимков, в которых один или несколько каналов существенно зашумлены, применялся медианный фильтр: значение в каждой точке заменяется на медианное значение среди трех значений: 1) в точке, находящейся слоем ниже, 2) в самой точке, 3) в точке, находящейся слоем выше.

Детектирование сигналов

Распознавание сигналов подразумевает применение фильтра для выделения небольших пятен и нормировки сигнала. Для каждой точки со значением a фильтр считает значение гауссова размытия снимка $b=g(a)$ и значение общего фона в окрестности $c=m(g(a))$, равное максимуму среди значений b в большой окрестности точки. В качестве результата фильтр выдаёт значение $(a-b)/c$. В полученном кадре выделяются точки, имеющие значение выше заданного в настройках порога, и совокупность находящихся рядом (в трёхмерном пространстве) выделенных точек называется сигналом. Значения порогов выставляются вручную для серии снимков одного эксперимента. Для отделения шума и артефактов слишком маленькие и слишком большие сигналы отбрасываются. Для определения точной координаты сигнала использовалась координата его центра масс.

Детектирование границ ядер и хромосомных территорий

Детектирование границ ядер и хромосомных территорий производится по фиксированному порогу после применённых фильтров. Цвет сигнала, по которому определяются границы ядра, является одним из настраиваемых параметров. Обычно любая из флуоресцентных проб дает фоновое окрашивание всего ядра. Детектирование границ ядра осуществляется путем расположения цилиндра, обозначающего границы клетки, оптимальным образом — чтобы минимальное число пикселей находилось вне границы. Используется цилиндр, а не шар, так как верхний и нижний «полюсы» ядер не снимались.

Для определения границ хромосомной территории в окрестности каждого сигнала строится распределение наблюдаемых значений яркости флуорофора территории, определяются значения квантилей 0,3 и 0,99, и все значения между ними линейно интерполируются на значения от 0 до 1. Значения, меньшие, чем значения квантили 0,3, заменяются на 0 (точка лежит вне хромосомной территории), а большие — на 1 (точка является частью хромосомной территории).

Определение расстояния от сигнала до территории

Расстояние до хромосомной территории определялось, как расстояние от центра масс сигнала до плоскости, проведённой через центры трех ближайших вокселей территории; Положительное расстояние — если сигнал исходно находился вне территории и отрицательное — если сигнал лежал внутри территории.

Отбрасывание возможных артефактов, определение пар сигналов от одного аллеля

Чтобы исключить возможные артефакты, на полученные результаты накладываются различные фильтры. Из используемых в анализе данных удалялись случаи, когда на изображении видно, что снимок не удался – невозможно визуально выделить отдельные точечные сигналы или хромосомные территории. Алгоритм работы программы следующий: 1) выбираются и учитываются только те ядра, в которых есть ровно два зелёных и два оранжевых сигнала, 2) из этих сигналов формируются такие пары зелёного и оранжевого, чтобы они были самыми близкими в ядре (распределение сигналов на пары, представляющие начало и конец одного аллеля), 3) между сигналами в этих парах считаются расстояния. Если сигналы слишком далеко друг от друга или от хромосомной территории, такие клетки также не учитываются в статистике, как ошибки детектирования территории или ошибки разбиения сигналов по парам. Все детектированные программой разрывы были подтверждены визуально.

Статистическая обработка результатов

Для определения, является ли «хвост» распределения расстояний между сигналами в клетках, обработанных этопозидом, неслучайным, сформулировали гипотезу: истинная мера хвоста p (доля выбросов) больше нуля. Исходя из желаемого уровня достоверности 0,997, по правилу трех сигм, и согласно теореме об асимптотике эмпирического распределения [353] (Боровков А.А., 1997), истинная мера p отличается от экспериментальной p^* не более, чем на величину $x = \text{три сигма делить на корень из числа точек}$, где сигма есть корень из числа $p \times (1-p)$. Если $p^* - x > 0$, то с достоверностью $> 0,997$ хвост распределения есть. Если p примерно равно значению $p^* = 0,029$, вычисляемому через гистограмму, то для нашего случая хвост с достоверностью $> 0,997$ не случаен.

Для проверки гипотезы, что распределение аллелей вне и внутри хромосомной территории отличается у обработанных и не обработанных клеток, использовали критерий Хи-квадрат.

3.4. Метод 4С

Приготовление 3С-библиотек

Подготовка ядер клеток

Обработанные этопозидом либо контрольные клетки осаждали, по 10 млн клеток. Меняли среду на 10 мл полной среды с добавлением формальдегида до 2,5%. Клетки фиксировали 10 мин при комнатной температуре. Останавливали реакцию, добавляя до 0,124 М глицин, инкубировали 5 мин во льду. Клетки осаждали при 300 g 10 мин на 4 °С, промывали холодным PBS с добавлением до 10% телячей сыворотки (FBS) и повторяли центрифугирование. Ресуспендировали осадок в 10 мл холодного лизирующего буфера (10 mM Трис-HCl pH=8, 10 mM NaCl, 0,2% (по объему) NP40, 1X ингибитор протеаз [Thermo Scientific, #78430]). Инкубировали 15 мин во льду. Осаждали при 600 g 5 мин на 4 °С, ресуспендировали осадок в PBS/10%FBS, удаляли супернатант и замораживали осадок в жидком азоте.

Рестрикция HindIII и лигирование

Замороженные осадки растворяли в 200 мкл 1,1X буфера NEB2, осаждали при 1000 g 5 мин при 4 °С. Осадок ресуспендировали в 200 мкл 1,1X буфера NEB2 с добавлением SDS до 0,3%. Инкубировали 60 мин в термошейкере на 1400 об/мин при 37 °С. Затем добавляли 330 мкл 1,1X буфера NEB2 и тритон-X100 до 1,8%. Инкубировали еще 60 мин в термошейкере на 1400 об/мин при 37 °С. Добавляли 600 единиц рестриктазы HindIII (NEB #R0104M), оставляли на ночь в термошейкере на 1400 об/мин при 37 °С. Затем добавляли еще 200 ед. HindIII на 2 часа. Инактивировали рестриктазу, добавляя SDS до 1,3% и инкубируя смесь 20 мин в термошейкере на 1400 об/мин при 65 °С.

Для лигирования суспензию переносили в 50-мл пробирки, где в каждой предварительно готовили 0,7 мл 10X буфера для лигазы T4, 375 мкл 20% тритон-X100 и 5,3 мл воды. Инкубировали 1 час на качалке (220 об/мин) при 37 °С. Добавляли 100 ед. (20 мкл) T4 DNA-лигазы (Fermentas #EL0011). Инкубировали 5 часов в водяной бане при 16 °С с помешиванием и затем еще 30 мин при комнатной температуре.

Очистка ДНК 3С-библиотек

Добавляли 600 мкг (30 мкл) Протеиназы К, инкубировали ночь при 65 °С в водяной бане. Охлаждали, добавляли к смеси равный объем фенол-хлороформа (1:1). Перемешивали, осаждали при 3220 g в течение 10 мин при комнатной температуре. Водную фазу отбирали, добавляли равный объем хлороформа. Перемешивали, осаждали при 3220 g 10 мин при комнатной температуре. Отбирали водную фазу в фальконы, пригодные для центрифугирования на высоких оборотах (Corning Центристар #430829), измеряли объем (~7 мл), добавляли 7 мкл гликогена (Fermentas #R0561, 20 mg/ml) 5,6 мл воды mQ, 1,4 мл

3М NaOAc до 0,3 М, перемешивали, добавляли 96% этанол до 50 мл, выдерживали 3 часа на -80°C .

Осаждали при 7500 g 60 мин при 4°C . Растворяли осадок в 200 мкл 10 мМ Трис-НСI pH=8,0. Добавляли 5 мкл РНКазы А (Fermentas #EN0531), инкубировали 45 мин при 37°C . Добавляли 360 мкл магнитных шариков AMPure XP beads, инкубировали при комнатной температуре 10 мин. Промывали ДНК дважды 75% спиртом, элюировали в 200 мкл 10 мМ трис-НСI pH=8,0, измеряли концентрацию ДНК. Проводили аналитический электрофорез для контроля лигирования образцов (в том числе отобранных до лигирования).

Приготовление 4С-библиотек

Рестрикция DpnII и лигирование

К 20–50 мкг 3С-библиотеки, добавляли 50 мкл 10X DpnII-Buffer, 200 ед (4 мкл), рестриктазу DpnII (NEB #R0543M), воды до 500 мкл. Инкубировали ночь при 37°C . Инактивировали 20 мин в термошейкере на 1400 об/мин при 65°C . Чистили ДНК с помощью фенол-хлороформной очистки, осаждали с гликогеном. Растворяли ДНК в 100 мкл 10 мМ Трис-НСI pH=8,0.

Для лигирования образцы переносили в 50 мл фальконы с 10X лигазным буфером (1,4 мл) и 12,5 мл воды, добавляли 200 ед. лигазы Т4, инкубировали 5 часов в водяной бане при 16°C с помешиванием. Чистили ДНК с помощью фенол-хлороформной очистки и магнитных шариков, Растворяли ДНК в 200 мкл 10 мМ Трис-НСI pH=8,0. Измеряли концентрацию ДНК. Проводили аналитический электрофорез для контроля качества вторых рестрикции и лигирования, сравнивая образцы до и после.

Аmplификация фрагментов ДНК, содержащих якорные участки

Лигированные молекулы представляют из себя кольца, часть из которых содержит якорный участок. Для амплификации фрагментов ДНК, содержащих якорный участок, проводили инвертированную ПЦР (последовательности праймеров для 4С в таблице 3.1). Использовали набор Expand Long Template PCR System (Roche #11 681 842 001). Сначала проводили аналитическую ПЦР, чтобы подобрать количество циклов, оптимальное для приготовления 4С-библиотек. Ставили с каждой парой праймеров 4 реакции (4 варианта количества циклов ПЦР) в 20 мкл и один отрицательный контроль. В каждой реакции замешивали (на льду) 50 нг DNA, 0,3 мМ праймеров, 0,35 мМ дНТФ, 1,5 ед. полимеразы. Программа ПЦР: 94° 2 мин [94° 15 сек, Та 1 мин, 68° 3 мин] \times 27–36. Та: для AML1 $59,7^{\circ}\text{C}$, MLL $59,1^{\circ}\text{C}$, CCND1 $62,0^{\circ}\text{C}$. Доставали соответствующие образцы после 27, 30, 33, 36 цикла. Анализировали образцы на электрофорезе в 1% агарозном геле по полосе мажорного продукта, который образуется в случае, если рестриктаза не разрешила сайты на

границах якорной последовательности, или произошло их религирирование (размеры продуктов указаны в таблице 3.1.)

Таблица 3.1. Последовательности праймеров для приготовления 4С-библиотек с якорями на BCR генов *AML1* и *MLL*, а также на контрольный ген *CCND1*. Указаны температуры отжига праймеров и размеры мажорного продукта ПЦР.

Название праймера	Последовательность	Ta, °C	Продукт, п.н.
4C-AML1-H	CAGGTGAGTGTGGAGGTAGAGAG	59,3	308
4C-AML1-D	CTTGGTCCCCAGCTGAGAT	60,2	
4C-MLL-H v2	ATAGGCTCCATGTTGGCTCA	59,1	764
4C-MLL-D v2	CACAGGATACAAAGCAGAАCTACTC	58,8	
4C-CCND1-H	CCTGCCAACTTCGGTGTCC	62,1	950
4C-CCND1-D	AAGTTACCCGAGGCGGAGTC	61,9	

Для препаративной ПЦР с каждой парой праймеров ставили по 4 одинаковые отдельные реакции в 50 мкл. Программа 94° 2 мин [94° 15 сек, Ta 1 мин, 68° 3 мин] × 30 циклов (определено по результатам аналитического ПЦР). Полученные 4С-библиотеки объединяли и чистили с помощью набора Qiagen PCR Purification Kit согласно инструкции производителя, ДНК элюировали в 50 мкл Трис-НСl 10 мМ рН=8,0. Препаративную ПЦР ставили с соблюдением всех мер предосторожности против контаминации ПЦР-продуктами: в отдельной комнате и не в той одежде, что была при работе с продуктами аналитической ПЦР. Смешивали по 1,3 мкг ПЦР-продуктов от трех пар праймеров для приготовления каждой объединенной 4С-библиотеки и отдавали на секвенирование. Библиотеки были подготовлены к секвенированию М.Д. Логачёвой (Лаборатория эволюционной геномики, факультет Биоинженерии и Биоинформатики) и секвенированы парными прочтениями на платформе Illumina HiSeq 4000.

Анализ данных секвенирования 4С-библиотек

Первичный анализ данных был проведен А.В. Артемовым (Лаборатория №10 Интеллектуального анализа данных и предсказательного моделирования ИППИ РАН) и описан в работах [354], [355]. Вкратце, прочтения, или риды, были демультимплексированы, обрезаны, и в дальнейшем использовались из пары прочтений те, что содержали сайт узнавания HindIII. Последовательности, располагающиеся в прочтениях после сайта узнавания HindIII, были картированы на геном человека (сборка hg19), используя программный пакет Bowtie.

«Сырой» 4С-сигнал представлял из себя количество прочтений, картированных возле каждого сайта узнавания HindIII в геноме. Однако для удобства анализа данных

частоты контактов якорных локусов определялись для геномных отрезков (бинов) размером в 500 тыс.п.н.

Для всех экспериментов 4С была построена матрица корреляций. Критерий Пирсона показал высокую корреляцию для повторов экспериментов. Критерий Мана-Уитни использовался для сравнения числа контактов якорей с генами-партнерами по транслокациям и остальными генами.

3.5. Создание клеточных моделей образования транслокаций

Общая стратегия

Для создания клеточной культуры, в которой заданная транслокация индуцируется доксициклином, в геном клеток линии LCL интегрировались элементы системы CRISPR/Cas под контролем системы Tet-ON. Элементами CRISPR/Cas были: ген нуклеазы Cas9 и гены РНК-гидов, направляющих Cas9 в локусы, между которыми возникает перестройка. Экспрессия Cas9 контролировалась TRE-промотором, активным только при наличии гетерологически экспрессируемого белка-активатора, связывающего доксициклин. Кроме перечисленных генетических элементов, в геном клеток встраивались гены устойчивости к селектирующим антибиотикам. Плазмида AAVS1-M2rtTA, содержащая ген белка-активатора, была получена из Addgene (#60843). Плазмида Puro_Cas9_gRNA1+2, содержащая элементы системы CRISPR/Cas, собиралась на основе плазмиды Puro_Cas9 donor (номер #58409 в Addgene).

Для сборки плазмиды Puro_Cas9_gRNA1+2 сначала были подобраны РНК-гиды и собраны плазмиды на основе плазмиды phU6-gRNA (номер в Addgene #53188). Каждая такая плазмида содержала ген одной из гидовых РНК (далее иногда просто гидов). Эти плазмиды затем использовали для проверки работоспособности гидов. Затем участок, содержащий ген гидовой РНК (вместе с промотором и терминатором), амплифицировали с помощью ПЦР, чтобы затем вставить его в плазмиду Puro_Cas9 donor. Сборка осуществлялась методом Gibson Assembly. Полученная плазмида использовалась при трансфекции клеток. Дальнейшее описание процесса создания клеточных моделей образования транслокаций будет сделано на примере клеточной культуры iAML1-ETO (от «inducible *AML1-ETO* translocation»), моделирующей образование транслокации между генами *AML1* и *ETO*.

Подбор РНК-гидов

РНК-гиды для индукции транслокации подбирали следующим способом. В первую очередь определяли участок генома, внутри которого необходимо индуцировать разрыв. Затем в нем проводили поиск последовательностей, которые могут быть узнаны комплексом Cas9–РНК-гид с помощью онлайн-сервиса crispr.mit.edu. В качестве исходного материала в этот сервис загружали последовательность размером до 250 нуклеотидов из выбранного участка генома, включавшую в себя максимальное число точек разрыва, обнаруженных у пациентов [203]. Указанный сервис автоматически распознавал PAM в пределах региона поиска и выдавал список возможных РНК-гидов (пример на рисунке 3.1), а также проводил оценку числа их потенциальных «оффтаргетов» по референсному геному человека (сборка hg19). Для каждого локуса на основании сформированного сервисом списка мы выбрали по три гидовые РНК. Среди узнающих участков гидов старались брать такие, которые имеют температуру отжига около 60 °С с учетом добавленных нуклеотидов CACC, необходимых для дальнейшего встраивания в плазмиду phU6-gRNA. В этом случае в дальнейшем удобнее проводить скрининг клонов с плазмидой со вставкой. Начинаться гид должен с G, в противном случае добавляли G на 5'-конец.

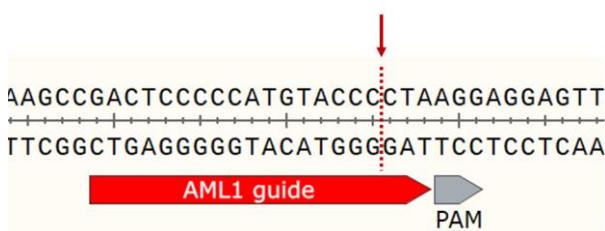


Рисунок 3.1. Последовательность участка ДНК, который комплементарен узнаваемой гидовой РНК (по сути, последовательность РНК-гида), PAM и точка двуцепочечного разрыва, который вносит Cas9, на примере участка гена *AML1*.

Сборка плазмид с генами гидовых РНК

Сборка экспрессионных векторов, несущих гены гидовых РНК, осуществлялась на основе плазмиды phU6-gRNA (номер #53188, Addgene). Этот вектор содержит промотор гена малой ядерной РНК U6, последовательность конститутивной части гидовой РНК (скаффолд), и небольшой спейсер между ними (рисунок 3.2). Спейсер содержит сайты рестрикции для эндонуклеазы *BstV2I*, вырезающей этот фрагмент из плазмиды. Вместо него в плазмиду можно встроить олигонуклеотид, представляющий из себя последовательность узнающего участка гидовой РНК (рисунок 3.3). Таким образом, плаزمида phU6-gRNA использовалась для создания целого семейства плазмид, несущих гены разных гидовых РНК.

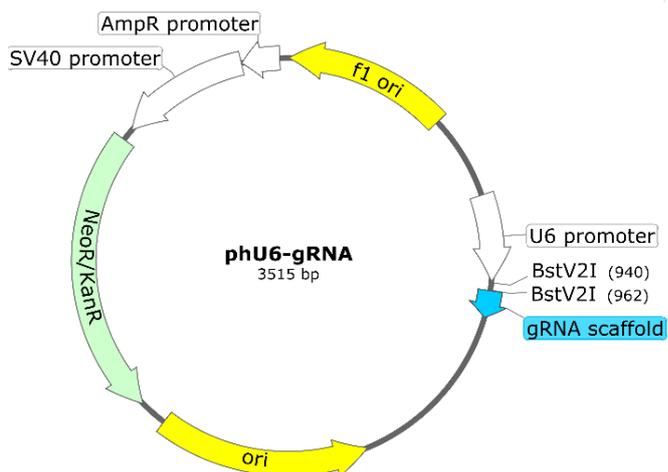


Рисунок 3.2. Карта плазмиды phU6-gRNA. На карте отмечены элементы: KanR — ген устойчивости к канамицину под контролем бактериального промотора AmpR, Ori — ориджин репликации в клетках *E. coli*, U6 — промотор РНК полимеразы III человека, gRNA скаффолд — последовательность константной части гидовой РНК. Отмечены сайты узнавания рестриктазы *BstV2I*.

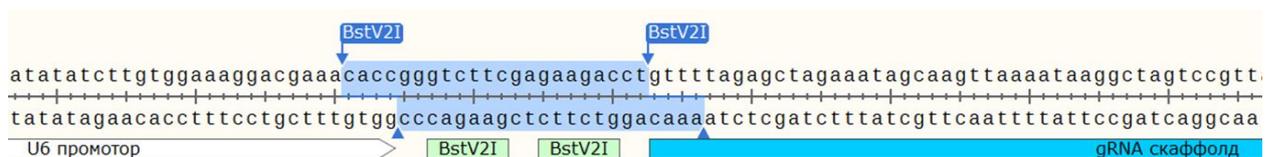


Рисунок 3.3. Участок плазмиды phU6-gRNA. Фрагмент, вырезаемый рестриктазой *BstV2I*, выделен синим цветом. Зеленым цветом отмечены сайты узнавания рестриктазы *BstV2I*. Остальные обозначения — как на предыдущем рисунке.

Получение олигонуклеотидов-вставок

Последовательность узнающего участка гидовой РНК далее называется прямым гидом, а комплементарная ей — обратным гидом. К прямому гиду добавлялся на 5'-конец САСС, к обратному — АААС. Заказывали в компании «Евроген» олигонуклеотиды сразу 5'-фосфорелированными.

Олигонуклеотиды разводили до 10 мМ, брали 31 пмоль прямого и столько же соответствующего ему обратного (3,1 мкл), смешивали, нагревали до 95° и инкубировали 5 мин, затем давали остыть 1 час до комнатной температуры. Получившиеся вставки с липкими концами лигировали в плазмиду phU6-gRNA, разрезанную *BstV2I*, дефосфорилированную с помощью щелочной фосфатазы FastAP (Thermo Fisher Scientific) и очищенную набором для очистки ПЦП-продукта NucleoSpin® Gel and PCR Clean-up (MACHEREY-NAGEL) согласно протоколам производителей.

Лигирование плазмиды со вставкой и трансформация бактерий

Реакцию лигирования 100 нг плазмиды и 5 пмоль вставки проводили в 15 мкл ДНК-лигазой T4 (Thermo Fisher Scientific) 20 минут при комнатной температуре согласно протоколу производителя. Затем проводили трансформацию: 3-5 мкл лигазной смеси добавляли к 50 мкл компетентных клеток, подвергали клетки тепловому шоку 45 с при 42 °С. Охлаждали 2 мин на льду, добавляли 750 мкл среды SOC, которая представляет собой среду SOB с добавлением глюкозы до 20 мМ. Инкубировали 1 час при 37 °С в термомиксере. По 200 мкл размазывали на чашку с агаризованной средой LB (1% w/v триптона, 0,5% w/v дрожжевого экстракта, 170 мМ NaCl, 1,5% агара) с добавлением в качестве селективирующего антибиотика канамицина («Панэко») 25 мкг/мл. Колонии бактерий выращивались в течение ночи в инкубаторе при 37 °С.

Скрининг клонов

На следующий день после трансформации ставили ПЦР-скрининг: скалывали колонию в 200 мкл воды, брали этой суспензии 1 мкл и замешивали ПЦР с использованием обратного праймера M13r и прямого олигонуклеотида (тот, который с добавлением CACC) в качестве прямого праймера. Результаты ПЦР визуализировали на электрофореze. Из колоний, несущих плазмиду со вставкой, затем наращивали ночную культуру (в 5–10 мл среды LB с канамицином 25 мкг/мл) и выделяли плазмиду с помощью набора NucleoSpin® Plasmid (MACHEREY-NAGEL) или GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) по протоколу производителя. Плазмиду затем секвенировали по Сэнгеру на предмет корректности вставки, используя праймер M13r.

Метод ENIT для проверки гидовых РНК

Метод назван ENIT (engineered nuclease-induced translocations) предполагает котрансфекцию клеток плазмидами, кодирующими нуклеазы и гидовые РНК (в случае проверки работы элементов системы CRISPR/Cas). В случае эффективной работы нуклеаз между точками разрывов с некоторой вероятностью образуются хромосомные перестройки. Спустя 48 часов после трансфекции с ДНК клеток ставится ПЦР на предмет наличия ожидаемых транслокаций, которые происходят между разрезаемыми локусами.

Вместе с подбором РНК-гидов сразу подбирали и заказывали праймеры для ПЦР, детектирующей транслокацию, подбор проводили с помощью сервиса Primer-blast (ncbi.nlm.nih.gov/tools/primer-blast). Расстояние до точки разрыва выбирали 30–100 нуклеотидов, чтобы использовать праймеры впоследствии и для ПЦР в реальном времени. Заказывали по паре вариантов каждого праймера. Старались допускать не больше 2 GC в последних 5 нуклеотидах на 3'-конце праймера, всего длина 18–30 нуклеотидов. Синтез праймеров осуществлялся в компании ООО «ДНК-синтез» (Россия) или в Eurofins Genomics.

Трансфекция клеток HeLa

Для проверки гидов методом ENIT использовали клетки HeLa. За день до трансфекции клетки засеивали в шестилуночный планшет на 30 % конфлюентности. Трансфицировали клетки реагентом TurboFect (Thermo Fisher Scientific) согласно инструкции производителя: 3–4 мкг плазмид суммарно из расчета на одну лунку и реагент TurboFect — 6–8 мкл. На следующий день после трансфекции меняли клеткам среду на свежую полную DMEM.

Прямая вложенная ПЦР

Транслокации лучше всего детектируются, если в методе ENIT использовать стратегию вложенной ПЦР (nested PCR). Также вложенная ПЦР менее требовательна к подготовке образцов, поэтому возможно использование «прямой» ПЦР без выделения ДНК с помощью специальных наборов (рисунок 3.4). Условия прямой ПЦР были отработаны отдельно [373].

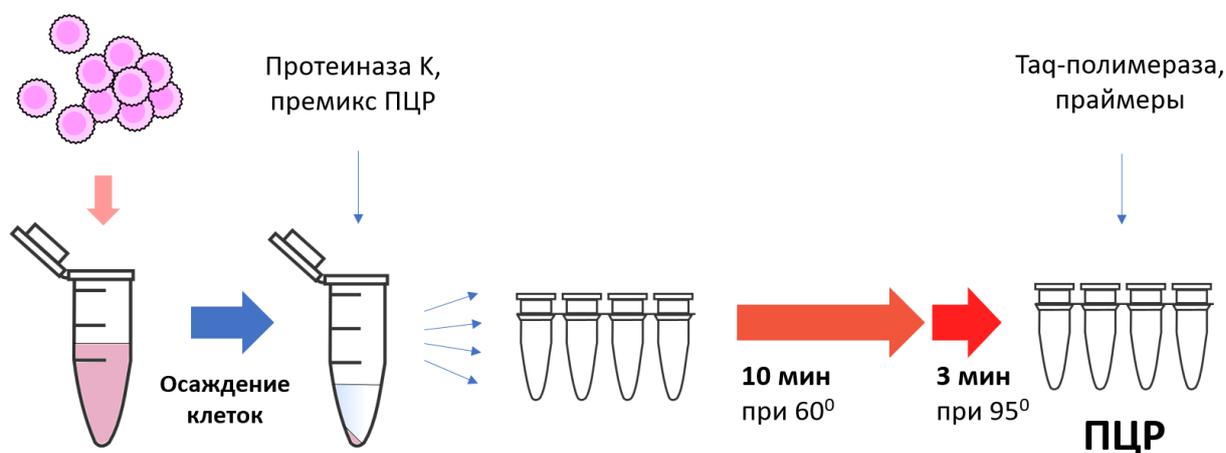


Рисунок 3.4. Последовательность действий при проведении прямой ПЦР с культуры клеток. Лизис клеток и ПЦР проводятся в одних и тех же ПЦР-пробирках.

Спустя 48 часов после трансфекции клетки снимали с чашки с помощью раствора 0,05% трипсина — 0,53 мМ ЭДТА («ПанЭко») — 300 мкл на лунку, затем добавляли 2,7 мл PBS. На одну реакцию ПЦР брали 5–100 тыс клеток. Осаждали 3 мин 5 000 об/мин на настольной центрифуге Eppendorf Minispin. К осадку клеток добавляли 2X премикс для ПЦР (буфер, нуклеотиды, ионы магния — согласно инструкции производителя к ДНК-полимеразе) и протеиназу К до 0,1 мг/мл так, чтобы получилось по 10 мкл на каждую предстоящую реакцию. Распределяли по 10 мкл в ПЦР-пробирки. Ставили в амплификатор, программа: 10 мин 60 °С и 3 мин 95 °С.

После лизиса в пробирки добавляли HotStart Taq-полимеразу, праймеры для первой стадии вложенной ПЦР и воду до 20 мкл. Ставился ПЦР1 с внешними праймерами с программой: 95 °С 5 мин, (95 °С 20 сек, 60 °С 1 мин, 72 °С 45 сек) ×20, 72 °С 5 мин.

Затем 1 мкл из реакции разбавлялся в 100 мкл воды и 1 мкл этого разбавленного раствора использовался для ПЦР2 (в 20 мкл) с внутренними праймерами: 95 °С 5 мин, (95 °С 20 сек, 60 °С 1 мин, 72 °С 30 сек) ×25, 72 °С 5 мин. Результаты ПЦР2 анализировались с помощью электрофореза, длины полос сравнивались с теоретически ожидаемыми.

Вставка генов гидовых РНК в плазмиду Puro_Cas9 donor

Метод Gibson Assembly

Метод основан на образовании длинных одноцепочечных концов на соединяемых фрагментах ДНК и отжиге таких комплементарных концов друг на друга. Комплементарные концы формируются в результате наработки соединяемых фрагментов методом ПЦР с длинными праймерами, содержащими необходимые участки на неотжигающихся 5'-концах. Нам необходимо было вставить 2 фрагмента, в качестве места встраивания выбрали сайт узнавания рестриктазой SalI. Использовался набор Gibson Assembly Master Mix (NEB), выполняемый протокол основан на протоколе производителя набора.

Праймеры со «свешивающимися» 5'-концами подбирались согласно протоколу Gibson Assembly от NEB, исходя из последовательности плазмиды вокруг сайта узнавания рестриктазой SalI. Фрагментами для вставки являлись гены гидовых РНК, в качестве матрицы для их наработки использовались проверенные плазмиды pHU6-gRNA со вставкой узнающего участка гидовой РНК.

Сначала проводили оптимизацию условий ПЦР: ставили градиент по температурам отжига праймеров. В качестве точки отсчета брали температуру отжига, рассчитанную в онлайн-сервисах NEB Tm Calculator. Качество ПЦР и концентрацию фрагментов оценивали по яркости полос на электрофорезе с помощью программы ImageJ. Затем ставили препаративный ПЦР в 50 мкл. ПЦР-продукт очищали от компонентов ПЦР специальным набором NucleoSpin® Gel and PCR Clean-up по инструкции производителя.

Рестриксию 1–2 мкг плазмиды Puro_Cas9 donor проводили рестриктазой SalI в течение 3 часов при 37 °С. Очищали продукт рестрикции с помощью набора для очистки ПЦР-продукта NucleoSpin® Gel and PCR Clean-up.

Каждую реакцию Гибсон-сборки ставили в 4–8 мкл. В каждую реакцию брали плазмиды 110 нг, 10 нг фрагментов (или воды для контроля) и воды до 4 мкл, добавляли 4 мкл Gibson Assembly Master Mix. Помещали в амплификатор на 30 мин на 50 °С, затем оставляли на 4 °С.

Брали по 2 мкл этой смеси к 50 мкл компетентных клеток, затем проводили трансформацию по обычному протоколу. Высеивали клетки на агаровые чашки с ампициллином в качестве селективного антибиотика.

Скрининг колоний на наличие вставки

После трансформации на чашках в опыте выросло на порядок больше колоний, чем в негативном контроле. В этом случае ставили ПЦР-скрининг колоний — скальвали штук 20 колоний в 200 мкл воды, брали 1 мкл этой суспензии и замешивали ПЦР с праймерами PuroCas9_F_Sal и PuroCas9_R_Sal, 30 циклов, 60 °С отжиг. На электрофорезе выявлялись полосы: 200 п.н. — если вставки нет, 500 п.н. — если вставился только один фрагмент, 800 п.н. — если вставились оба фрагмента. В последнем случае такую колонию скальвали, наращивали ночную культуру, выделяли плазмиду с помощью набора NucleoSpin® Plasmid и отправляли на секвенирование по Сэнгеру, используя праймеры PuroCas9_F_Sal и PuroCas9_R_Sal. Если секвенирование показывало корректность вставки, то наращивали большой объем культуры (150 мл) с этой плазмидой и выделяли плазмиду с помощью набора NucleoBond® Xtra Midi (MACHEREY-NAGEL) по протоколу производителя для дальнейшей трансфекции плазмидой клеток.

Трансфекция клеток и селекция на антибиотиках

Электропорация клеток культуры LCL

Трансфекцию культуры LCL осуществляли с помощью электропорации с использованием прибора Gene Pulser Xcell (Bio-Rad Laboratories). Сначала подобрали оптимальные условия электропорации из числа нескольких предустановленных протоколов (таблица 3.2). В качестве тестовой плазмиды использовали 2,5 мкг плазмиды, кодирующей EGFP. Помещали 3 млн клеток в 200 мкл полной среды в кювету для электропорации шириной 0,2 см. После электрического пульса быстро переносили клетки в лунки шестилуночного планшета в 2 мл полной среды.

Таблица 3.2. Параметры протоколов электропорации клеток LCL.

Название протокола	Тип электрического разряда	Напряжение, В	Длительность разряда *, мс	Емкость конденсатора *, микрофарад
«HuT78»	Квадратная волна	130	25	-
«CV1»	Квадратная волна	100	25	-
«A549»	Квадратная волна	150	10	-
«Jurkat»	Экспоненциальный спад	140	-	1000
«K562»	Экспоненциальный спад	155	-	1000

* — длительность указана только для разрядов квадратной формы. Разряд по типу экспоненциального спада происходит в результате разрядки конденсатора, емкость которого указана в последнем столбце таблицы.

Через сутки после электропорации определяли долю живых флуоресцирующих клеток на проточном цитофлуориметре FACS Aria SORP (BD Biosciences) с участием Д.М. Поташниковой (кафедра клеточной биологии и гистологии). Наибольший процент живых флуоресцирующих клеток (2,6% от общего числа клеток) был получен по протоколу «Jurkat», который далее использовался в работе.

Для получения клеточной культуры iAML-ETO клетки культуры LCL трансфицировали одновременно четырьмя плазмидами: Puro_Cas9 _AML1_ETO_puro (рисунок 3.5), AAVS1-M2rtTA (рисунок 3.6) и двумя плазмидами, содержащими гены субъединиц нуклеазы TALEN, которая вносит разрыв по сайту AAVS1: TALEN_L (Addgene #59025) и TALEN_R (Addgene #59026). Трансфекция всех плазмид осуществлялась путем электропорации по протоколу «Jurkat», описанному выше. Для электропорации использовалось по 2,5 мкг плазмид TALEN_L и TALEN_R, и по 2,3 мкг плазмид AAVS1-M2rtTA и Cas9+gRNA _AML1_ETO. На 2–3-й день после электропорации визуально под микроскопом оценивалось состояние культур: большая часть клеток должна была погибнуть, однако отдельные живые клетки начинали образовывать кластеры. В этом случае к клеткам добавляли селективирующий антибиотик.

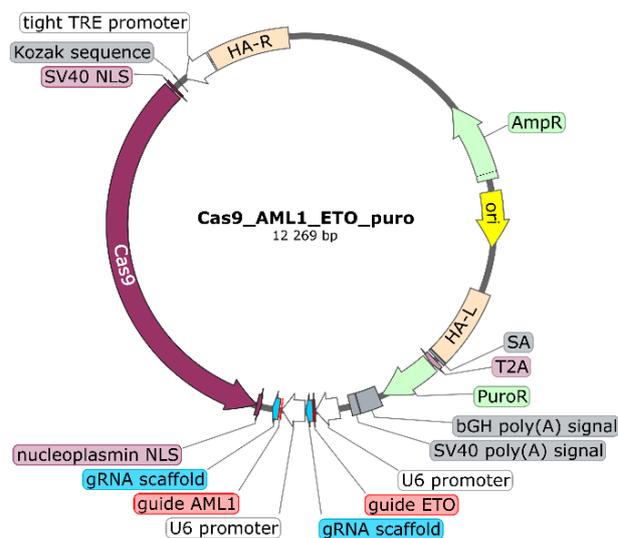


Рисунок 3.6.

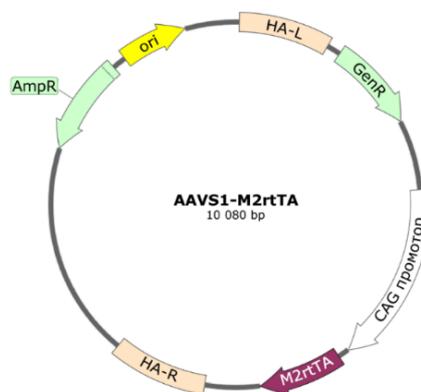


Рисунок 3.5. Плазмидный вектор Cas9 _AML1_ETO_puro, который содержит гены гидовых РНК gRNA_AML1 и gRNA_ETO (отмечены переменные участки генов и конститутивный скэффолдный), а также ген Cas9 под контролем промотора TRE. В состав TRE входит 7 сайтов посадки активатора транскрипции M2rtTA. AmpR — ген устойчивости к ампициллину, PuroR — ген устойчивости к пурамицину, SA — акцептор сплайсинга, T2A — последовательность T2A-пептида. Ori — ориджин репликации в клетках *E. coli*, U6 — промотор РНК полимеразы III человека. HA-R и HA-L — плечи гомологии, необходимые для интеграции кассеты генов в геном.

Рисунок 3.6. Плазмидный вектор AAVS1-M2rtTA, содержащий ген активатора транскрипции M2rtTA и ген устойчивости генетицину (GenR). Промотор CAG необходим для поддержания высокого уровня экспрессии M2rtTA в клетках человека.

Селекция клеток iAML-ЕТО на антибиотиках

Для выявления рабочих концентраций антибиотиков проводились предварительные эксперименты по определению чувствительности клеток LCL к антибиотикам. Клетки распределяли по по 200–250 тыс. клеток по 12-луночному планшету и в среду к ним добавляли 10 разных концентраций пурамицина от 0 до 1,1 мкг/мкл. Спустя сутки, двое, трое и т.д. оценивали число живых клеток в лунках с помощью микроскопии суспензии клеток с красителем трипановым синим. Минимальная концентрация антибиотика, при которой клетки погибают на 2–3 день (0,3 мкг/мл), была взята как селективирующая для дальнейших экспериментов. Максимальная концентрация, при которой клетки не успевают за 7 дней погибнуть, но и не делятся, была взята как поддерживающая. Аналогичным способом была определена селективная и поддерживающая концентрации генетицина (гентамицин, G418) — 750 мкг/мл и 300 мкг/мл, но на селекцию в этом случае закладывали 7–10 дней ввиду довольно больших значений концентраций.

Спустя два-три дня после трансфекции клеток им добавляли антибиотик генетицин G-418 (Gibco) до концентрации 800 мкг/мл. Трансфицированные клетки росли в среде с генетицином в течение 10 дней. Затем клетки пересеивали на среду с пурамицином (Sigma-Aldrich) в концентрации 0,3 мкг/мл и культивировали 4 дня. После этого клетки пересеивались на среду с «поддерживающей» концентрацией антибиотиков: 0,1 мкг/мл пурамицина и 300 мкг/мл генетицина. Трансфицированные клетки далее культивировали при такой концентрации антибиотиков на протяжении всех экспериментов по определению свойств полученной культуры. Спустя один месяц клетки повторно проверяли на устойчивость к антибиотикам. В отличие от контрольных клеток LCL, трансфицированные клетки выдерживали 4 дня инкубации при концентрации генетицина в 1000 мкг/мл или при концентрации пурамицина в 0,5 мкг/мкл.

3.6. Проверка экспрессии трансгенов

Выделение РНК из клеток эукариот с помощью TRIzol reagent

Клетки (от 3 до 5 млн) центрифугировали 4,5 минуты при 600 g на центрифуге 5417 R (Eppendorf). Осадок клеток ресуспендировали в DPBS («ПанЭко»), переносили в пробирку на 1,7 мл и центрифугировали 3 минуты при 2,1 тысяч об/мин на центрифуге Eppendorf MiniSpin (ротор 6 см). Осадок клеток лизировали в 750 мкл TRIzol reagent (Invitrogen). Далее к лизату добавляли 150 мкл хлороформа, и раствор перемешивали с помощью вортекса Reax Top (Heidolph). Лизат центрифугировали при 12000 g 10–15 минут при 4 °C. Верхнюю (водную) фазу отбирали для выделения РНК. Нижнюю фазу и интерфазу использовали далее для выделения ДНК.

Для осаждения РНК к отобранной водной фазе добавляли 350 мкл изопропилового спирта. После 10 минут инкубации при комнатной температуре центрифугировали при 12 000 g 10 минут при 4 °С. Осадок РНК промывали 75% этиловым спиртом, центрифугировали при 12 000 g 5 минут при 4 °С. Осадок РНК растворяли в 15 мкл буфера 10 mM трис-НСl pH=7,5. Выделенная РНК хранилась при -80 °С.

Качество выделенной РНК оценивали с помощью электрофореза в агарозном геле. РНК считалась пригодной для дальнейшего применения, если на электрофореграмме детектировались четкие полосы 28S и 18S рибосомальной РНК.

Выделение ДНК из клеток с помощью TRIzol reagent

Параллельно с выделением РНК выделяли ДНК. После отбора верхней фазы к оставшимся в пробирке фенольной фазе и интерфазе добавляли 250 мкл 96% этилового спирта. После тщательного перемешивания осадок ДНК осаждали при 2 000 g 5 мин при 4 °С. Фенольно-спиртовой супернатант удаляли и ресуспендировали осадок в растворе 0,1 М цитрата натрия на 10% этиловом спирте. После 30 минут инкубации при комнатной температуре центрифугировали при 2 000 g при 4 °С в течение 5 минут. Супернатант удаляли, и осадок ДНК повторно ресуспендировали в растворе 0,1 М цитрата натрия на 10% этиловом спирте, инкубировали и осаждали. После удаления супернатанта к осадку добавляли 75% этиловый спирт на ночь (4 °С). На следующий день осадок ДНК осаждали при 2000 g при 4 °С в течение 5 минут. Супернатант удаляли и осадок подсушивали с открытой крышкой в течение 10–15 минут. Растворяли осадок в 100–200 8 mM NaOH в термомиксере при 56 °С при 550 об/мин в течение 1 часа. Не растворившийся осадок удаляли, центрифугируя смесь 10 минут при 12 000 g при 4 °С.

Обратная транскрипция

Перед проведением обратной транскрипции образец РНК обрабатывали ДНКазой для удаления примесей геномной ДНК. В реакционную смесь добавляли 5 мкг РНК, 1 единицу активности ДНКазы I (Thermo Fisher Scientific) и соответствующий реакционный буфер. Обрабатывали ДНКазой 30 минут при 37 °С, затем останавливали реакцию, добавляя ЭДТА до концентрации 5 mM и прогревая реакционную смесь при 65 °С в течение 10 минут.

Для синтеза кДНК в реакционную смесь добавляли 1 мкг РНК, обработанной ДНКазой, дАТФ, дТТФ, дГТФ и дЦТФ до 1 mM каждого, 100 пмоль вырожденных праймеров (смесь случайных гексамеров), 20 единиц активности ингибитора РНКазы RiboLock (Thermo Fisher Scientific) и 200 единиц активности обратной транскриптазы RevertAid H minus (Thermo Fisher Scientific) в поставляемом буфере. Объем реакции — 20 мкл. Реакционную смесь инкубировали 10 минут при 25 °С, затем еще 1 час при 42 °С. Реакцию останавливали, прогревая смесь в течение 10 минут при 70 °С.

ПЦР в реальном времени (количественная ПЦР)

Для проведения ПЦР в реальном времени использовали набор реактивов Maxima SYBR Green qPCR Master mix (Thermo Fisher Scientific), позволяющий проводить реакцию с использованием «горячего старта». Объем одной реакции — 25 мкл. Концентрация праймеров составляла 250 нМ. Последовательности праймеров приведены в таблице 3.3. Синтез праймеров осуществлялся в компании ООО «ДНК-синтез» (Россия). По окончании циклов ПЦР прибор снимал кривую плавления для подтверждения специфичности амплификации. В качестве дополнительной проверки специфичности продукты некоторых реакций анализировали с помощью электрофореза в агарозном геле.

Таблица 3.3. Последовательности праймеров, использовавшихся для проведения ПЦР.

Название праймера	Последовательность нуклеотидов
Cas9_f	CCGAAGAGGTCGTGAAGAAG
Cas9_r	GCCTTATCCAGTTCGCTCAG
TA_f	AAATCAGCTCGCGTTCCTGT
TA_r	CGCTTTCGCACTTTAGCTGT
GAPDH_f	CAAGGTCATCCATGACAACCTTG
GAPDH_r	GTCCACCACCCTGTTGCTGTAG
AML_f1	TGGGGAAGCTCACCAGATAG
AML_r1	CAGGCTTTTCACAAGCATTCCT
AML_f2	CTCCAGTGGGCTATTCTGC
AML_r2	TCCTTCTTCCTGAAGCCGA
ETO_f1	TTGGGACACCTAGGAGTGGT
ETO_r1	TGCATGAACATAAACAGGCACTC
ETO_f2	GGCCTGAGGAGCTTGCTATC
ETO_r2	GAACATAAACAGGCACTCTTTGC
Myc_f	CCAGTAACTCCTCTTTCTTCGGAC
Myc_r	CGCTATGCTGGATTTTGCTGCA
Myc_f_out	GGCCGTTTTAGGGTTTGTTGG
M13f	CCCAGTCACGACGTTGTAACACG
M13r	AGCGGATAACAATTCACACAGG

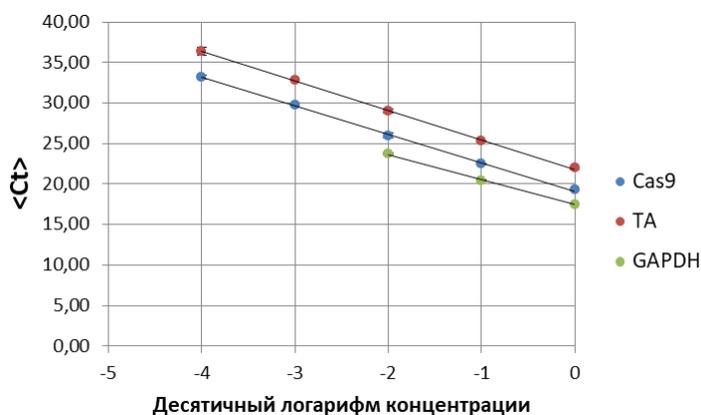
Для измерения экспрессии генов Cas9 и M2rtTA в реакционные смеси вносили по 1 мкл соответствующей кДНК, полученной в ходе обратной транскрипции. Cas9 и M2rtTA амплифицировали с помощью праймеров Cas9_f и Cas9_r, TA_f и TA_r, соответственно. ПЦР проводили по протоколу: 95°C 10 мин, (95 °C 15 сек, 60 °C 20 сек, 72 °C 30 сек) x 40. Каждую реакцию ставили в трех повторностях. Нормировку экспрессии указанных генов проводили по экспрессии гена *GAPDH* (праймеры GAPDH_f и GAPDH_r) определенной в том же образце кДНК. Все реакции ставили одновременно в 96-луночном планшете. Анализ экспрессии проводили в соответствии с методом, описанным в [356].

В предварительном эксперименте определяли эффективность амплификации с используемыми парами праймеров. Для этого приготовили серию из 5 последовательных десятикратных разведений одного из образцов кДНК и с 1 мкл смеси проводили ПЦР в реальном времени с каждой из трех пар праймеров и в трех повторностях, по их результатам вычисляли среднее значение порогового цикла ($\langle Ct \rangle$) для каждой пары праймеров для каждого разведения. Строили график зависимости $\langle Ct \rangle$ от десятичного логарифма концентрации кДНК (рисунок 3.7). Исходная концентрация кДНК принималась равной единице.

Эффективность амплификации с пары праймеров вычисляли по формуле:

$$10^{\frac{1}{-tg\alpha}},$$

где $tg\alpha$ — тангенс угла наклона соответствующей прямой (см. рисунок 3.7).



Пара праймеров	Эффективность амплификации
Cas9_f и Cas9_r	1,93
TA_f и TA_r	1,89
GAPDH_f и GAPDH_r	2,10

Рисунок 3.7. Графики зависимости среднего значения порогового цикла ($\langle Ct \rangle$) для каждой пары праймеров от десятичного логарифма концентрации кДНК. Легенда: Cas9 — пара праймеров Cas9_f и Cas9_r, TA — пара праймеров TA_f и TA_r, GAPDH — пара праймеров GAPDH_f и GAPDH_r. Эффективность амплификации кДНК с указанными парами праймеров указана справа.

Оценку уровня экспрессии генов Cas9 и M2rtTA в экспериментальном образце проводили по формуле:

$$R_i = \frac{E_{GAPDH}^{\langle Ct \rangle_{GAPDH}}}{E_i^{\langle Ct \rangle_i}}, \text{ где } R_i \text{ — относительный уровень экспрессии гена } i \text{ (Cas9 или M2rtTA), нормированный на экспрессию гена } GAPDH. E_i \text{ — эффективность амплификации с соответствующей пары праймеров, } \langle Ct \rangle_i \text{ — среднее по трем точкам значение порогового цикла амплификации образца кДНК с соответствующей парой праймеров.}$$

Перед проведением количественной ПЦР ставили контрольную реакцию на отсутствие примесей геномной ДНК: в качестве матрицы для проведения ПЦР с указанными праймерами использовалась РНК, обработанная ДНКазой I, но до синтеза кДНК.

3.7. Амплификация области транслокации AML1-ETO

Для детекции транслокации использовали либо вложенную ПЦР, либо ПЦР в реальном времени с TaqMan-зондом. Транслокация осуществлялась между локусами, в которые вносила разрыв нуклеаза Cas9, направляемая гидовыми РНК со следующими узнающими последовательностями: gRNA_AML1: GACTCCCCCATGTACCCСТА, gRNA_ETO: GATGTAAGAGGAAGCAGСТТ.

Вложенная ПЦР

Вложенную ПЦР проводили с полимеразой TrueStart Hot Start Taq (Thermo Fisher Scientific), позволяющей проводить реакцию с использованием «горячего старта», либо полимеразой DreamTaq Hot Start (Thermo Fisher Scientific). Количества компонентов реакции подбирались согласно инструкции производителя. Концентрация каждого из праймеров в реакционной смеси была 250 нМ. Объем реакции — 20 мкл. Количество матрицы варьировало в зависимости от эксперимента. Подбор праймеров осуществлялся с помощью онлайн-сервиса «Primer Blast». Последовательности праймеров, использовавшихся в работе, приведены в таблице 3.3.

Вложенная ПЦР осуществлялась в два этапа. На первом этапе ставили реакцию с внешними праймерами по протоколу: 95 °С 5 мин, (95 °С 20 сек, 60 °С 1 мин, 72 °С 45 сек) × 20. Далее к 20 мкл реакционной смеси добавляли 80 мкл деионизированной воды. 1 мкл получившегося раствора использовали в качестве материала матрицы во втором раунде ПЦР с внутренними праймерами (протокол 95 °С 5 мин, (95 °С 20 сек, 60 °С 1 мин, 72 °С 30 сек) × 25). Таким образом, на втором этапе использовалась 1/100 часть продукта первой реакции. Пары праймеров, использовавшиеся для проведения вложенной ПЦР, можно найти на рисунке 3.8.

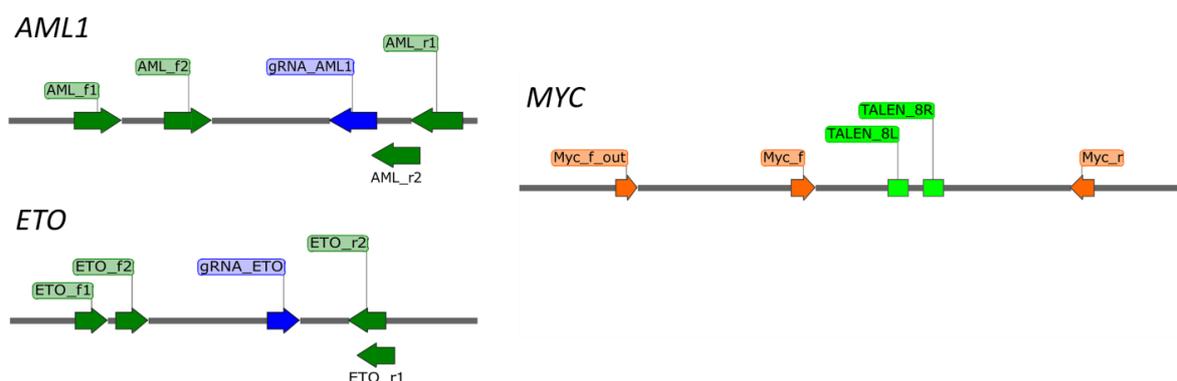


Рисунок 3.8. Взаимное расположение праймеров для проведения вложенной ПЦР на участках генов *AML1*, *ETO* и *MYC*. Синим отмечены последовательности, узнаваемые гидовыми РНК в генах *AML1* и *ETO*, светло-зеленым — субъединицами TALEN в гене *MYC*. Темным зеленым и оранжевым цветом отмечены праймеры. В некоторых случаях для проведения второго раунда вложенной ПЦР использовался тот же праймер, что и для проведения первого раунда. Последовательности праймеров показаны в таблице 3.3.

Количественная ПЦР с TaqMan-зондом

Для определения количества транслокаций проводили количественную ПЦР с TaqMan-зондом. Праймеры и TaqMan-зонд подбирали с помощью онлайн-сервиса PrimerQuest™ Tool на сайте www.idtdna.com. Руководствовались следующими критериями: температура отжига зонда выше на 5–10 °С, чем у праймеров; зонд не более 30 нуклеотидов длиной; содержание GC в пределах 30–80%; не имеет G на 5'-конце, где находится флуорофор; зонд комплементарен цепи, в которой больше G, чем C. Был подобран и заказан в компании «ДНК-синтез» (Россия) зонд со следующей последовательностью: CCCAGCAGAATAGGCCCACTGGAGCCCA, зонд содержал флуорофор FAM и гаситель флуоресценции BHQ1 по концам. Для детекции флуоресценции на амплификаторе CFX96 выставляли фильтр возбуждения 450–490 нм и фильтр детекции 515–530 нм. Реакцию ПЦР проводили с помощью полимеразы DreamTaq Hot Start (Thermo Fisher Scientific) в объеме 25 мкл, использовали 1 мкг ДНК, праймеров до 900 нМ каждого (последовательность праймеров: ETO_F2: TGCATGAACATAAACAGGCACTC, AML_fus_R: GAAGSTCACCAGATAGGCTGTA), 250 нМ TaqMan-зонда, дНТФ до 0,2 мМ, 10X буфер из набора и воду для ПЦР (PanReac Applichem). Также в реакцию в конце добавляли 0,15 мкл эндонуклеазы рестрикции FastDigest HindIII (Thermo Fisher Scientific), что повышает качество количественной ПЦР с TaqMan-зондом (Акишев, 2021). Все экспериментальные точки ставились в трех повторностях. Предварительно смешивали премикс на все реакции из компонентов ПЦР, за исключением раствора ДНК, и раскапывали по 20 мкл премикса во все лунки планшета с помощью степпера Eppendorf Multipette. ДНК образцов предварительно разбавляли до концентрации 200 нг/мкл, и в каждую реакцию добавляли 5 мкл раствора ДНК. Объемы премикса и раствора ДНК были подобраны, исходя из оптимальных объемов для раскапывания степпером (фиолетовый наконечник) и автоматической пипеткой с наименьшей ошибкой. Планшет заклеивали оптически-прозрачной пленкой для ПЦР (SSI-3622-00) с помощью специального валика (BioRad), капли сбрасывали на центрифуге с планшетным ротором при 1000 g в течение 2 мин. Реакцию проводили на амплификаторе CFX96 по протоколу: 95 °С 5 мин, (95 °С 20 сек, 60 °С 1 мин, 72 °С 30 сек) x 35.

Во всех реакциях ставили отрицательный контроль без матрицы. ПЦР замешивалась в отдельной комнате в специальном ламинарном шкафу со строгим соблюдением всех мер предосторожности против контаминации ПЦР-продуктами.

Для использования в качестве положительного контроля и для определения абсолютного количества транслокаций в образцах создали плазмиду, содержащую участок транслокации, амплифицируемый при ПЦР с праймерами AML_F9 ACCAAACTAGAATGCCAGGT и AML_R9 GTGGAGAAGTGCACAATGCTT. Плазида была создана на базе вектора pUC-18 (Addgene #50004) путем вставки участка

транслокации по сайтам рестрикции HindIII и BamHI. Использовали рестриктазы FastDigest (Thermo Fisher Scientific), после рестрикции вектор дефосфорилировали щелочной фосфатазой FastAP. Вставка нарабатывалась с помощью вложенной ПЦР с ДНК активированных клеток iAML-ЕТО (48 часов после активации) по протоколу, описанному выше. Внутренние праймеры отличались от AML_fus_R и ETO_F2 наличием неотжигающихся 5'-участков, содержащих сайты рестрикции HindIII и BamHI, и 4 дополнительных нуклеотидов. После очистки ПЦР-продукта набором QIAquick PCR purification kit (Qiagen) по инструкции производителя, проводилась его рестрикция HindIII и BamHI и повторная очистка. Лигирование вектора и вставки проводилось в соотношении 1:5 по молям, трансформацию компетентных клеток проводили, по протоколу, описанному выше. Выросшие колонии скринировали с помощью ПЦР с использованием праймеров M13F и M13R, фланкирующих область вставки. По размеру ПЦР-продукта на электрофореze в агарозном геле определяли наличие вставки, 2 такие колонии скальвались в ночную культуру. Выделение плазмиды из ночной культуры проводили набором реактивов GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) по инструкции производителя. Последовательность вставки подтвердили секвенированием по Сэнгеру.

3.8. Обработка культуры клеток iAML/ЕТО ингибиторами белков репарации и химиотерапевтическими препаратами

Тесты на цитотоксичность для подбора рабочих концентраций ингибиторов

Тест на цитотоксичность ингибитора NU7026 был поставлен с использованием клеток культуры iAML-ЕТО и с помощью набора ХТТ согласно инструкции производителя. В разных работах использовалась концентрация от 0,2 мкМ ([357]) до 10 мкМ (из [358] следует, что такая концентрация не токсична для клеток сама по себе, но повышает чувствительность к радиации). Клетки были распределены по в 96-луночный планшет в количестве 90 тысяч в 200 мкл среды на лунку. Добавляли NU7026 в концентрациях от 0,3 до 20 мкМ с одновременным добавлением к клеткам до 0,1 мкг/мл доксицилина, также были клетки без добавления веществ (контроль). Все экспериментальные точки были представлены в трех повторностях. Спустя 3 дня с помощью ХТТ-теста по поглощению при оценивали выживаемость клеток в разных образцах, вычитая значение среды без клеток и нормировав на контроль люминисценции (детектировалась на планшетном ридере) с контролем. Выживаемость клеток при 20 мкМ оказалась на уровне 80–85%, поэтому данная концентрация была использована в дальнейших экспериментах.

Тесты на цитотоксичность для подбора рабочих концентраций препаратов химиотерапии

Данная работа была выполнена совместно с коллегами из Лаборатории комбинированной терапии опухолей «НМИЦ онкологии им. Н.Н. Блохина» — В.С. Покровским и Д.В. Соколовой. Сначала были проведены предварительные эксперименты со значениями веществ, основанными на литературных данных. Затем значения выживаемости были уточнены после второго теста, концентрации которого выбирали уже на основе результатов первого теста так, чтобы крайние точки были ориентировочно концентрациями 1% и 99% выживаемости. Для проведения теста клетки линии iAML-ЕТО разбавляли до 22 тысяч/мл и распределяли по лункам 96-луночного планшета. Вещества тестировали в пяти концентрациях. После инкубации клеток с различными концентрациями веществ в течение 4 дней проводили МТТ-тест по протоколу производителя. Поглощение при 540 нм измеряли на планшетном ридере Tirerdeck Multiscan MCC/340 (Flow Lab). Выживаемость считали по значению, нормированному на не обработанные клетки. Все эксперименты проводили в трех повторностях. Полученные данные обрабатывали с помощью программы Prism Graphpad version 9.1.0, находя IC10, IC25, IC50.

Таблица 3.4. Значения цитотоксичности химиотерапевтических препаратов для клеток культуры iAML-ЕТО. MNTD — минимальная нетоксичная доза, IC10, IC25, IC50 — значения, при которых жизнеспособность культуры снижается на 10, 25 и 50%.

Препарат	MNTD	IC10	IC25	IC50	Производитель
Карбоплатин, мкг/мл	0,1	0,1	0,3	0,9	Accord
Цисплатин, нг/мл	7,2	11,5	20,4	62,0	Accord
Цитарабин, нМ	0,01	0,02	0,05	0,21	Sigma
Этопозид, нг/мл	1,17	1,93	3,57	11,48	Teva
Ифосфамид, мкг/мл	9,3	16,7	33,6	103,3	Veropharm
Иринотекан, нг/мл	9,1	15,9	30,9	94,8	Veropharm
Метотрексат, нг/мл	1,3	2,0	3,2	7,9	EBEWE PHARMA
Оксалиплатин, нг/мл	2,8	6,3	15,8	55,7	Sanofi
Паклитаксел, нг/мл	0,004	0,005	0,007	0,014	Bristol-Myers Squibb

Обработка клеток культуры iAML-ЕТО, выделение ДНК и подготовка к ПЦР-анализу на транслокации

Для проведения опытов по влиянию различных веществ на количество транслокаций при активации клеток культуры iAML-ЕТО, клетки рассеивались в концентрации, позволяющей им не перерасти концентрацию 1 млн/мл на момент окончания эксперимента (3–4 день). Клеткам добавляли доксициклин в концентрации 0,1 мкг/мл и тестируемые вещества. Клетки осаждались при 10 000 g 5 минут, затем супернатант удаляли и из осадка выделяли клетки с помощью коммерческого набора Биолабмикс (Россия) по протоколу производителя. При этом на одно выделение брали 3–5 млн клеток, и ключевым моментом выделения было разбивание агрегата лизированных клеток с помощью ресуспендирования пипеткой — иногда на это уходило несколько минут на образец. Элюцию проводили 60–70 мкл элюирующего буфера из набора, предварительно нагретого до 60 °С. Концентрацию образцов измеряли спектрофотометрически прибором NanoPhotometer P330 (Implen) по поглощению при длине волны 260 нм и флуориметрически с помощью прибора Qubit и набора реактивов к нему (Invitrogen). Если концентрации, полученные разными способами измерения, не совпадали, то брали среднее значение. Затем образцы разбавляли водой для ПЦР до концентрации ДНК 100 нг/мкл.

Статистическая обработка результатов количественной ПЦР

Для определения статистической значимости различий в экспериментах использовался критерий Даннета [359]. Использование этого критерия оправдано, поскольку мы не знаем, распределены ли наши значения нормально, а также проводим множественное попарное сравнение нескольких выборок с одной контрольной. Формула для расчета критерия Даннета:

$$q = \frac{R_{cont} - R_i}{\sqrt{\frac{nl(l+1)}{6}}}, \text{ где } R_{cont} \text{ и } R_i \text{ — суммарные ранги контрольной и опытной}$$

выборок, соответственно. n — число элементов в каждой выборке, l — число сравниваемых выборок. Полученное значение q сравнивали с табличным [359] для определения статистической значимости. Поскольку для каждой обработки число биологических повторностей равнялось 2, в формулу подставляли $n=2$. Соответственно, ранги определяли на основании двух значений F для каждой обработки.

3.9. Используемое лабораторное оборудование и программное обеспечение

Клеточный инкубатор BINDER CB 210 (Binder), Центрифуги Biosan microspin (Biosan) и Eppendorf: 5702, minispin (Eppendorf), суховоздушный термостат TC-1/80 СПУ (Смоленское СКТБ СПУ), термостат Термит (ДНК-технология), амплификаторы: PCR system (Perkin Elmer), Bio-Rad C1000 ThermalCycler и Bio-Rad CFX96 (Bio-Rad Laboratories), водяные бани (Elmi), вакуумный насос (Millipore), автоматические пипетки HTL (PZ HTL S.A.), инвертированный микроскоп Nikon Eclipse TS100 (Nikon), оптико-электронный модуль в составе комплекта DIC оптики и сканирующего блока для конфокального микроскопа LSM-510 Meta на базе инвертированного флуоресцентного микроскопа Axiovert 200M (Carl Zeiss). Для получения изображений на конфокальном микроскопе использовали программное обеспечение Laser Scanning Microscope LSM 510 Version 3.2 (Carl Zeiss) и NIS-Elements 4.13.01 (Nikon), для работы с изображениями — Carl Zeiss ZEN 2011, ImageJ 1.52b, KolorPaint 4.8.3, для построения карт плазмид и геномных локусов использовали SnapGene Viewer (GSL Biotech LLS). Для работы с изображениями, таблицами и текстами использовали программы пакета Microsoft Office 365. Для фотографирования результатов электрофореза использовали VisionWorks LS (UVP). Для анализа данных ПЦР в реальном времени использовали Bio-Rad CFX Manager версии 3.1 (Bio-Rad Laboratories).

4. Результаты

4.1. Изучение влияния этопозиды на аллели *AML1* методом 3D FISH

Ранее в нашей лаборатории методом трехмерной флуоресцентной *in situ* гибридизации (3D FISH) было изучено влияние обработки клеток этопозидом на ген *MLL*. Метод 3D FISH включает в себя способ приготовления препаратов клеток, когда фиксируется и сохраняется трехмерная структура ядер, затем препараты гибридизуются с флуоресцентно меченым зондом на определенные локусы ДНК, и далее проводится конфокальная микроскопия препаратов, позволяющая по серии конфокальных снимков восстановить трехмерное изображение. Изучение целостности гена *MLL* проводилось с помощью двуцветного флуоресцентного зонда, который визуализировал начало и конец гена, а также флуоресцентного зонда на хромосомную территорию (WCP, whole chromosome painting), в которой находится ген *MLL*. В качестве клеточной модели использовали клетки линии Jurkat – суспензионные перевиваемые клетки лимфоидной природы, имеющие крупное ядро правильной шаровидной формы, идеально подходящие для приготовления трехмерных препаратов. Клетки обрабатывали этопозидом в концентрации 0,17 мкМ (100 мкг/мл) в течение часа. В этих условиях клетки еще не входят в апоптоз, но имеют детектируемое число разрывов в *MLL*. Затем клеткам меняли среду на свежую без ингибитора и инкубировали в такой среде 60–90 минут для того, чтобы эффект от обработки этопозидом успел накопиться. Пробовали и другие условия обработки клеток: например, анализировали клетки, обработанные ингибитором, затем отмытые и проинкубированные в нормальных условиях в течение 0, 90 или 180 минут, однако в итоге пришли к одному варианту обработки — 90 минут. В этих условиях изучаемые эффекты от обработки этопозидом были ярко выражены и воспроизводились раз за разом, но при этом эксперимент не занимал слишком много времени.

Препараты исследовали с помощью конфокального микроскопа, получая серии конфокальных оптических срезов для каждого поля микроскопирования. Для детекции и статистически значимого подсчета редких событий, которые происходят далеко не в каждой клетке, и выявления каких-либо закономерностей необходимо было провести исследование большого количества клеток и проанализировать результаты. Для этого использовали программу, позволяющую обработать значительный массив данных — так, в нашей работе были проанализированы десятки тысяч клеток. Программа была написана для данного исследования Д.А. Алексеевским. Она получала на вход стопки конфокальных изображений и реконструировала трехмерные изображения. Затем определяла сигналы от флуоресцентных зондов: границы сигналов на начало и конец гена, границы хромосомных

территорий, и подсчитывала их перекрытие. Сигналы, которые перекрывались с территорией менее, чем на 20% по своей площади, принимались за вышедшие. В отдельных экспериментах без визуализации хромосомной территории проводился подсчет аллелей *MLL*, в которых сигналы от начала и конца разошлись, с помощью программы Nemo [360].

В результате опытов удалось детектировать разрывы ДНК в гене *MLL*, возникающие в условиях ингибирования ДНК-топоизомеразы II. В опытах с визуализацией хромосомной территорией было показано, что 3'-конец *MLL* чаще локализуется вне хромосомной территории в обработанных этопозидом клетках по сравнению с необработанными. Для контрольного гена *CCND1*, не замеченного в хромосомных перестройках, но активно экспрессируемого в этих клетках и лежащего в той же 11 хромосоме, такого не наблюдалось [325].

Нашей задачей было исследовать, как обработка клеток этопозидом влияет на второй ген, перестраивающийся при соответствующих вторичных лейкозах — ген *AML1*. По сравнению с предыдущей работой мы усовершенствовали протокол исследования. В работе по наблюдению за *MLL* использовали двуцветный ДНК-зонд с зеленым и оранжевым флуорофорами, а территорию хромосомы 11 визуализировали WCP-зондом, конъюгированной с флуорофором того же зеленого цвета. Это не позволяло визуально отличить 5'-конец *MLL* от хромосомной территории, соответственно, была возможность детектировать выходы за пределы хромосомной территории только 3'-конца гена *MLL*. В нашей работе мы использовали WCP-зонды, конъюгированные с флуорофором IDYE™ 647, который испускает свет в дальней красной области спектра, и двуцветные гибридизационные зонды на целевой ген, флуорофоры которых испускают свет в зеленой и оранжевой областях спектра.

Также вместе с Д.А. Алексеевским была существенно доработана программа для анализа изображений с микроскопа. В новой версии программа определяла границы сигналов трех цветов: от двух зондов на ген и на хромосомную территорию (рисунок 4.1), и высчитывала расстояния между центрами сигналов (за центр сигнала принимался центр тяжести сигнала, состоящего из нескольких вокселей). Программа распределяла сигналы по парам, состоящим из начала и конца одного аллеля, определяла расстояния между центром сигнала и до границы хромосомной территории, определяемой как плоскость между ближайшими центрами трех вокселей территории (см. Материалы и методы). Детектировались и границы ядер. В процессе анализа данных применялись различные фильтры: отбирались только такие случаи, когда внутри детектированного ядра была минимум одна хромосомная территория (обычно их две, но иногда они соприкасаются и сливаются в одну), и внутри ядра было по 2 пары сигналов (в интерфазных ядрах должно быть по 2 аллеля гена, зонд на целевой ген — двуцветный).

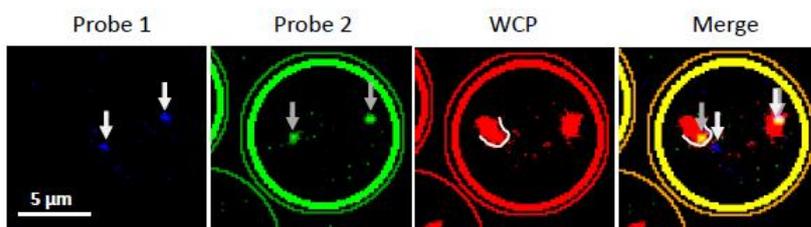
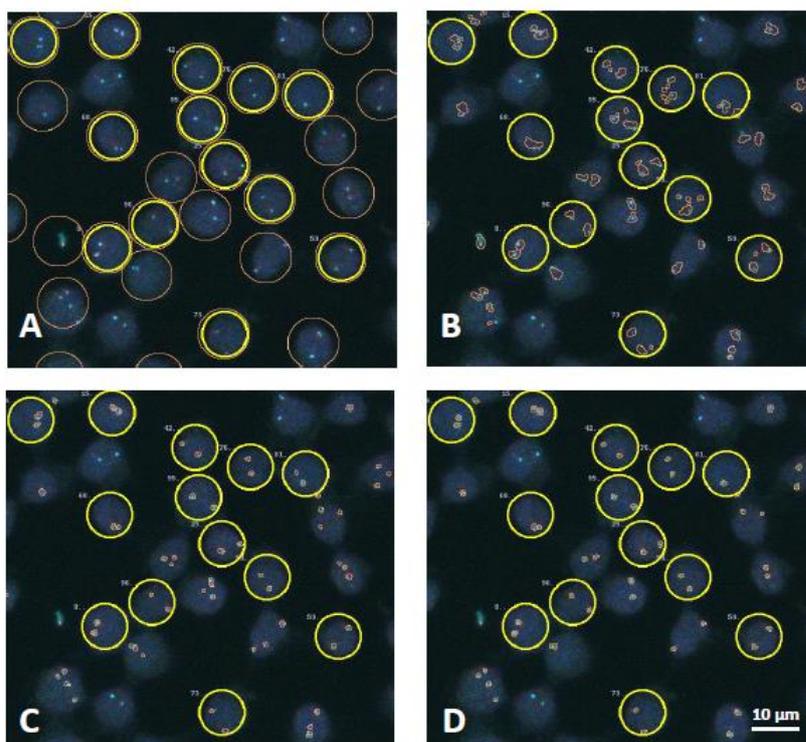


Рисунок 4.1. Распознавание компьютерной программой границ ядер, хромосомных территорий и сигналов от начала и конца гена *AML1* на снимке препарата клеток, полученном с помощью конфокального микроскопа. Изображен один конфокальный слой, но при анализе программа реконструирует трехмерную структуру, используя серию конфокальных снимков.

В результате изменения протокола и доработки программы для анализа изображений мы могли одновременно детектировать разрывы гена *AML1* и выходы концов разрыва этого гена за его хромосомную территорию. Как и в предыдущей работе, клетки обрабатывали этопозидом в концентрации 0,17 мкМ в течение часа с последующей инкубацией в среде без этопозидом в течение часа — полутора часов. Затем проводили 3D FISH, конфокальную микроскопию — для каждого препарата снимали 12–15 полей, содержащих порядка ста клеток, — и анализ данных. Эксперимент был проведен несколько раз, данные объединялись, так как распределения были идентичными. Кроме обработанных этопозидом клеток, 3D FISH проводили с контрольными необработанными клетками, а также с клетками, подвергнутыми воздействию высокой дозы γ -излучения (7 Гр). Доза излучения была подобрана, исходя из литературных данных [361].

Обработка этопозидом приводит к разрывам гена *AML1* и расхождению концов разрыва, в отличие от высоких доз γ -излучения.

Сначала мы построили распределение расстояний между центрами сигналов начала и конца для каждого аллеля *AML1* (мы будем называть их парными расстояниями) в необработанных этопозидом клетках (рисунок 4.2). Квантиль 99% (99-й процентиль) мы приняли, как максимальное расстояние между началом и концом интактного аллеля. Если расстояние между началом и концом аллеля будет больше, то такие аллели мы далее будем называть разорванными. Анализ распределения значений парных расстояний в клетках, обработанных этопозидом, показывает, что 4% сигналов — разорванные. Отметим, что это не просто аллели, внутри которых возник ДЦР, а такие аллели с ДЦР, концы разрывов в которых физически разошлись на расстояние в несколько десятков — сотни нанометров. Также отметим, что квантиль 99% — условное значение для границы разрыва, которое можно увеличить, например, до 99,5%, и тогда конкретные результаты будут иметь другие значения, но общие закономерности останутся теми же: в опытах с этопозидом тогда будет 3,6% разрывов. Для подтверждения неслучайности «хвоста» в распределении в случае обработки этопозидом по сравнению с контролем был проведен статистический анализ согласно теореме об асимптотике эмпирического распределения [353], подтвердивший неслучайность с достоверностью $> 0,997$ (см. Материалы и методы).

При воздействии на клетки высокой дозой радиации, мы не наблюдали такого эффекта, как при воздействии этопозидом — число разорванных аллелей *AML1* не выросло по сравнению с контролем (рисунок 4.2).

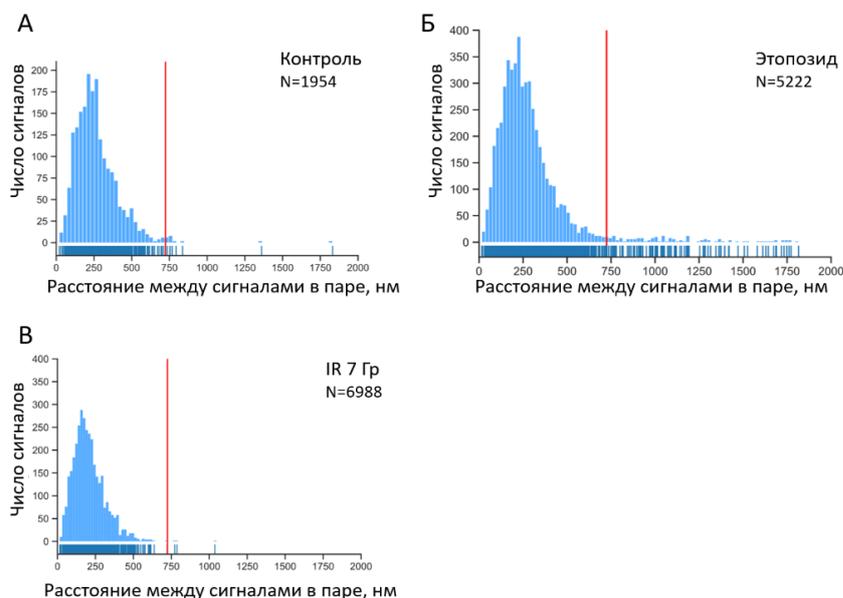


Рисунок 4.2. Гистограммы распределения расстояния между сигналами, визуализирующими концы одного аллеля *AML1* (использовали центры тяжести сигналов). А — контроль, Б — клетки, обработанные этопозидом, В — клетки, подвергнутые ионизирующему излучению. N — общее число сигналов на гистограмме. Вертикальная линия отсекает значение, равное 99-й перцентили в контроле.

Было обнаружено, что доля аллелей *AML1*, находящихся вне хромосомной территории, у клеток после обработки этопозидом была выше, чем в контрольных клетках (более чем в 2 раза, $p < 0,05$ по критерию хи-квадрат). Причем за хромосомную территорию чаще выходили разорванные аллели. Так, у обработанных этопозидом клеток вне хромосомной территории находилось 2,4% разорванных аллелей, а среди «неразорванных» — 0,8% (разница значима, $p < 0,05$ по критерию хи-квадрат). Однако мы не можем считать все «неразорванные» аллели интактными, так как клетки все же подвергались действию этопозиды, а часть разрывов скорее всего не детектируется с помощью микроскопии. Поэтому мы сравнили разорванные аллели из эксперимента с аллелями из контрольных клеток. В этом случае разница была более заметна (в 5,8 раз, $p < 0,001$ по критерию хи-квадрат). Другими словами, разорванные под действием этопозиды аллели *AML1* чаще оказываются вне хромосомной территории по сравнению с интактными. Отсюда следует, что разошедшиеся концы разрывов, вызванных обработкой этопозидом, имеют повышенную подвижность в пространстве ядра по сравнению с интактными аллелями. Это повышает вероятность образования транслокации с партнером из другой хромосомы.

Разорванные под действием этопозиды аллели чаще оказываются вне хромосомной территории по сравнению с интактными аллелями

Далее на основе полученных данных были построены двумерные графики, где для каждого сигнала (точка на графике) координата X означала расстояние до парного ему сигнала (вторая часть того же аллеля), а по координате Y откладывалось расстояние до границы его хромосомной территории (рисунок 4.3).

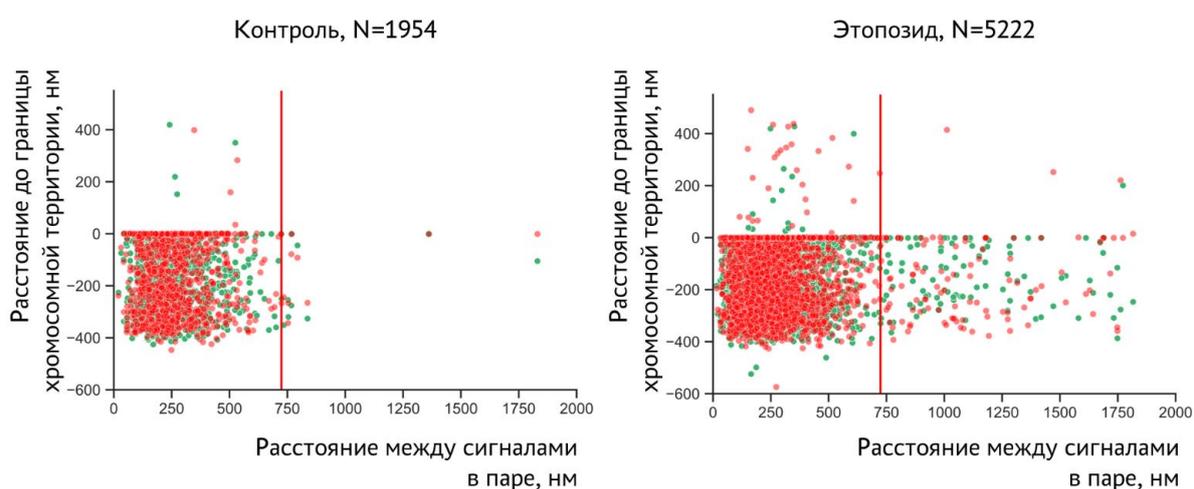


Рисунок 4.3. Графики, отражающие расстояния между началом и концом *AML1* (ось X), и между сигналом (началом или концом *AML1*) до границы хромосомной территории (Ось Y). Отрицательные значения Y означают, что сигнал находится внутри хромосомной территории. Вертикальная линия отсекает значения расстояний между сигналами от начала и конца одного аллеля, принятые за «разорванный аллель». Красные и зеленые точки — 5'- и 3'-концы *AML1*, соответственно.

4.2. 4С-анализ предпочтительных контактов *AML1* и *MLL*

Следующая задача заключалась в том, чтобы определить, существует ли пространственная предрасположенность к транслокации между генами-партнерами по изучаемым транслокациям, и возникает ли она при обработке клеток этопозидом. Другими словами, сближены ли гены-партнеры исходно в ядре и сближаются ли они друг с другом более предпочтительно, чем с любыми другими локусами, после воздействия этопозидом. Для решения задач, связанных с определением вероятности контактов между различными локусами в пространстве ядра, применяют 3С-методы (от chromosome conformation capture — захват конформации хромосом). Это группа методов для изучения пространственной организации хромосом, основанных на протоколе 3С [362]. Метод 3С предполагает фиксацию всех контактирующих участков ДНК формальдегидом, затем фрагментацию ДНК рестриктазой и лигирование сближенных участков. Получившиеся химерные молекулы представляют из себя 3С-библиотеку, которая затем может анализироваться различными способами.

Метод 4С представляет из себя анализ 3С-библиотеки с помощью инвертированной ПЦР с последующим массовым параллельным секвенированием продуктов ПЦР (рисунок 4.4) [363]. Для этого фрагменты 3С-библиотеки дополнительно дробятся рестриктазой с большой распространенностью сайтов в геноме, чем для первой рестриктазы, и лигируются с образованием колец ДНК. Таким образом, эти кольца представляют собой химерные молекулы, некоторые из которых («целевые кольца») содержат фрагменты интересующих нас генов. Чтобы провести анализ контактов интересующих нас генов, необходимо отсеку секвенировать неизвестные участки целевых колец. Для этого с целевых колец проводят инвертированную ПЦР — когда прямой и обратный праймеры, подобранные на известную нам последовательность гена интереса, направлены не друг к другу, а наоборот. Для генов *MLL*, *AML1* и контрольного гена *CCND1*, контакты которых мы изучали, мы выбрали участок между сайтами рестрикции *HindIII* (первая рестриктаза, сайт узнавания — 6 пар нуклеотидов) и *DpnI* (вторая рестриктаза, сайт узнавания — 4 пары нуклеотидов), внутри которого подобрали праймеры для инвертированной ПЦР. Такой участок называют якорем, и мы для *MLL* и *AML1* выбирали якорь возле локуса, который оказывается на границе перестроек у пациентов с лейкозом [364], [36].

Полученная 4С-библиотека секвенируется с помощью массового параллельного секвенирования (также называемого секвенированием нового поколения, или NGS), и проводится анализ прочтений: обрезание концевых участков, принадлежащих якорю, наложение оставшихся участков на геном и подсчет их количества. Результат является числом выявленных контактов между якорем и другими участками генома [365].

Для удобства визуального анализа данных 4C весь геном разбивают на равные фрагменты (бины) размером в несколько сотен тысяч пар нуклеотидов и строят 4C-профили контактов якоря со всеми биномы каждой хромосомы. 4C-профиль отражает количество обнаруженных контактов с каждым бином.

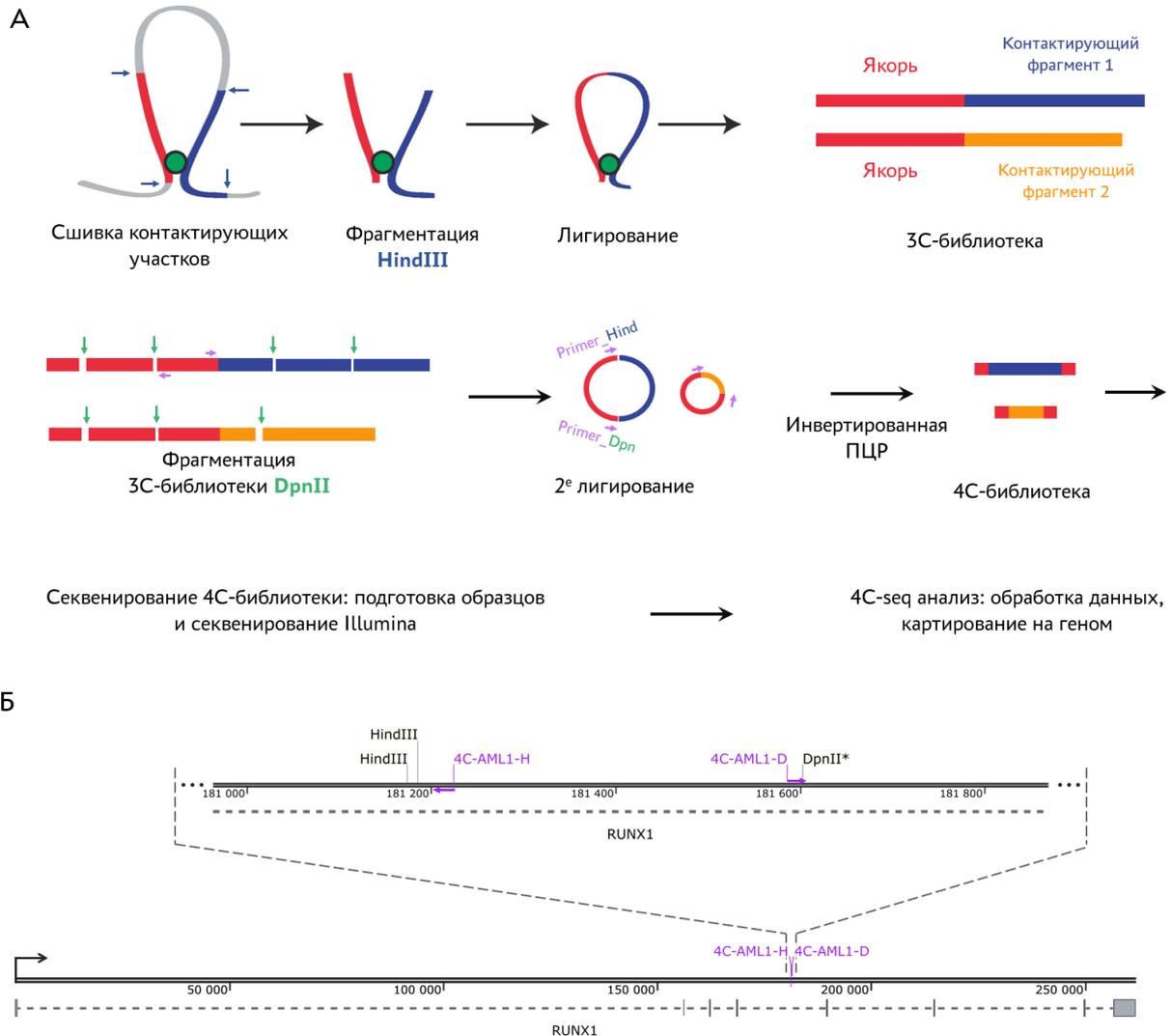


Рисунок 4.4. А. Общая схема протокола метода 4C. Б. Праймеры для инвертированной ПЦР для получения 4C-библиотеки на примере *AML1*. Якорем выступает область гена, в которой возникают разрывы, предшествующие перестройкам гена при лейкозах.

Мы провели анализ пространственных контактов генов *MLL* и *AML1*, а также контрольного гена *CCND1*, со всеми участками генома до и после обработки этопозидом. Эксперимент был проведен дважды. Для каждой из библиотек мы получили 7–20 уникальных прочтений, которые аннотировались на геном. Данные между двумя повторностями эксперимента продемонстрировали высокую корреляцию (критерий Пирсона $r > 0,999$), поэтому мы объединили значения 4C-сигнала в повторностях, предварительно нормализовав значения

в бинах (по 500 тыс.п.н.) на общее число прочтений в данной библиотеке. Данные визуализировались в UCSC-браузере. Для генов *AML1* и *MLL* данные 4C на якорной хромосоме и на хромосомах, где лежат их гены-партнеры по изучаемым транслокациям, представлены на рисунках 4.5 и 4.6.



Рисунок 4.5. 4C-профили контактов *AML1* необработанных клеток (синие) и обработанных этопозидом (сиреневые). Показаны профили *cis*-контактов *AML1* (с хромосомой 21) и профили контактов с хромосомами 1, 3 и 8, в которых лежат гены-партнеры *AML1* по перестройкам, ассоциированным с лейкозами, которые вызваны терапией топо2-ядами. Локусы генов-партнеров отмечены стрелками. Разрешение по оси X — 500 тыс.п.н. Профили построены в геномном браузере UCSC.



Рисунок 4.6. 4C-профили контактов *MLL* необработанных клеток (оранжевые) и обработанных этопозидом (зеленые). Показаны профили *cis*-контактов *MLL* (с хромосомой 11) и профили контактов с хромосомами 4, 9 и 19, в которых лежат гены-партнеры *MLL* по перестройкам, ассоциированным с лейкозами, которые вызваны терапией топо2-ядами. Лocusы генов-партнеров отмечены стрелками. Разрешение по оси X — 500 тыс.п.н. Профили построены в геномном браузере UCSC.

Анализ профилей 4С-сигнала показал следующие результаты. Во-первых, профили 4С-сигнала на хромосоме, в которой находится якорный ген, ожидаемо имели высокий пик в области якоря, что является внутренним контролем адекватности результатов 4С, так как чаще всего в пространстве сближены участки, которые недалеко друг от друга по геному. Во-вторых, визуальный анализ не выявил, чтобы *AML1* и *MLL* контактировали со своими генами-партнерами по изучаемым транслокациям чаще, чем с другими областями генома, как в контрольных клетках, так и в обработанных этопозидом. Также общий профиль контактов в опыте не поменялся по сравнению с контролем, как можно было бы ожидать, если бы гены-партнеры по транслокации направлялись бы после разрывов в какие-то определенные компартменты ядра. Эти выводы были подтверждены статистическим анализом сырых данных: количество контактов генов-партнеров по транслокациям не выбивалось из значений количества контактов якорей с любыми другими генами на других хромосомах. Также статистический анализ показал полную корреляцию профилей 4С-сигналов в обработанных этопозидом и не обработанных клетках. В-третьих, количество контактов якорей с другими хромосомами (транс-контактов) не увеличивается в клетках после обработки этопозидом. Это означает, что большинство аллелей все-таки остаются в своих хромосомных территориях, что совпадает с наблюдениями, полученными методом FISH, при которых доля вышедших аллелей в контроле была 0,4%, а в опыте — 0,9%.

4.3. Создание клеточной модели образования транслокаций

Более детальное изучение механизмов репарации, задействованных при формировании транслокаций, а также поиск с помощью метода FISH веществ, способных влиять на образование перестроек, малоэффективно. Лейкозогенные хромосомные транслокации легко обнаруживаются в клетках крови больных, потому что приводят к клональному преимуществу раковых клеток при делении. Тем не менее, первичное событие, запускающее раковое перерождение клетки, — формирование лейкозогенной транслокации — случается в популяции клеток крайне редко, и «поймать» момент возникновения транслокации методом FISH почти невозможно, так как этот метод основан на микроскопировании отдельных клеток.

Поэтому мы провели анализ методов, которыми возможно детектировать транслокации [10]. Наиболее удобным методом видится ПЦР, позволяющая количественно детектировать транслокации, однако данный метод применим только в случае, когда известны места разрывов, между которыми происходит транслокация, с точностью до нескольких нуклеотидов. Но если разрывы индуцируются путем ингибирования ДНК-топоизомеразы, то они возникают практически в любой точке на протяжении больших участков: при образовании лейкозогенных транслокаций эти участки (кластеры точек

разрыва, BCR) составляют десятки тысяч пар нуклеотидов, что не позволяет подобрать праймеры для детекции методом ПЦР такого широкого спектра транслокаций между генами-партнерами. Выходом не будет и ПЦР с обратной транскрипцией, несмотря на то что сплайсированные мРНК уже гораздо короче, и несколько пар праймеров покрывают большинство вариантов точек разрыва при транслокации между двумя генами. Это связано с тем, что этопозид индуцирует разрывы в большом количестве генов, и теоретически любая пара разрывов может привести к транслокациям. Поэтому отдельно взятая транслокация будет случаться слишком редко для того, чтобы можно было набирать какую-то статистику.

Выходом видится использование для индукции разрывов программируемых нуклеаз. Это нуклеазы, специфичность которых можно задавать для направления их в целевые локусы. К ним относятся нуклеазы с цинковыми пальцами, TALE-нуклеазы и нуклеазы систем CRISPR/Cas. Так как они позволяют направленно вносить ДЦР в заданное место генома, то являются основой соответствующих технологий редактирования генома [332]. Среди этих систем представляется самой удобной CRISPR/Cas, так как она требует использование генетических конструкций наименьшего размера: ген нуклеазы Cas9, а также гены двух гидовых РНК, суммарная длина конструкций — 5 тыс.п.н. Все необходимые элементы такой системы могут быть собраны в одну векторную плазмидную конструкцию. Для сравнения: если бы мы использовали систему нуклеаз TALEN, то для внесения двух разрывов нам бы потребовалось 4 гена по 4 тыс.п.н. (каждая TALE-нуклеаза состоит из двух субъединиц). Для доставки таких генов в клетку понадобилось бы 4 отдельные плазмиды, которые должны были бы попасть в клетку все вместе одновременно. Кроме того, создание нуклеаз TALEN, специфичных к определенному локусу, является отдельной сложной биоинженерной задачей.

Для индукции транслокаций во время проведения эксперимента можно использовать трансфекцию клеток плазмидами, кодирующими необходимые элементы системы CRISPR/Cas. Однако такой способ не обеспечит достаточного количества транслокаций, так как не во все клетки будут попадать плазмиды, кодирующие нуклеазы. Но главное, что трансфекция представляет из себя слишком жесткую процедуру для клеток, в результате которой многие погибают, а биологические эффекты будут во многом обусловлены состоянием выживших клеток. Поэтому мы решили получить культуру клеток с интегрированными в геном элементами системы CRISPR/Cas. А так как мы изучаем не влияние уже возникших транслокаций, как во многих других работах, а процесс их формирования, то транслокации должны запускаться по сигналу. Для этого ген нуклеазы Cas9 должен находиться под индуцируемым промотором. Для реализации такой схемы экспрессии мы выбрали систему Tet-ON.

Система Tet-ON основана на бактериальной системе регуляции экспрессии Tet-оперона, отвечающего за устойчивость бактерии к антибиотикам тетрациклинового ряда. Система Tet-On состоит двух компонентов: промотора перед регулируемым геном и белка транс-активатора. В состав промотора входит минимальный CMV-промотор и несколько операторных последовательностей (TRE — tetracycline response element), а транс-активатор содержит домен, связывающий эти операторные последовательности только в присутствии тетрациклина или его аналогов, например доксициклина. Также в состав транс-активатора входит сильный транскрипционный домен вирусного белка VP16, который запускает сборку инициаторного комплекса РНК-полимеразы II (рисунок 4.7) [366].

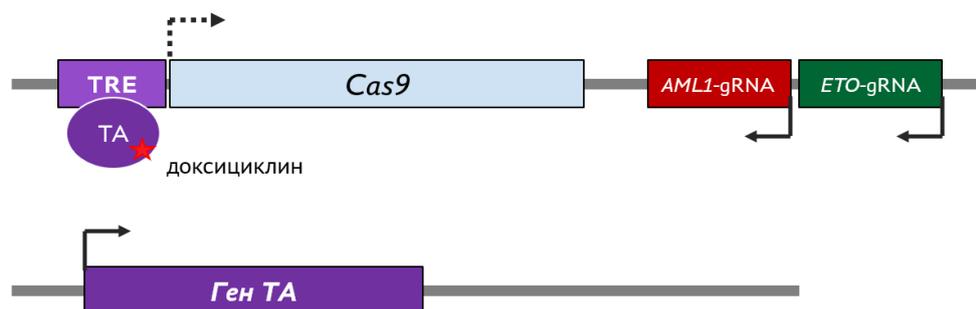


Рисунок 4.7. Механизм активации экспрессии Cas9 в полученной клеточной культуре iAML-EТО. Активатор транскрипции (ТА) экспрессируется конститутивно. Белок-активатор связывается с добавляемым к клеткам доксициклином и активирует экспрессию Cas9. TRE — промотор, содержащий сайты посадки ТА.

Общая схема получения клеточной культуры

Таким образом, в геном клеток необходимо было интегрировать ген нуклеазы Cas9, находящийся под промотором TRE, гены гидовых РНК под своими индивидуальными промоторами для экспрессирующей их РНК-полимеразы III, и ген транс-активатора под конститутивным промотором. Кроме того, для отбора клеток, содержащих в геноме необходимые конструкции, необходимы гены селективных маркеров. Наконец, для одинаковой и надежной экспрессии всех конструкций, они должны быть предсказуемо интегрированы в safe-harbor локус (от «спокойная гавань»), который обеспечивает стабильную экспрессию трансгенов в нем. Поэтому все необходимые элементы были собраны в две кассеты. В состав одной кассеты входили ген Cas9 под промотором TRE, гены двух гидовых РНК под промоторами U6 (промотор гена малой ядерной РНК U6) и ген устойчивости к селективному антибиотику пурамицину. В состав второй кассеты входил ген транс-активатора под конститутивным искусственным промотором CAG [367] и ген устойчивости к антибиотику неомицину (геницидину).

Чтобы интегрировать кассеты в определенное место в геноме, мы выбрали стратегию интеграции трансгенов по механизму гомологичной рекомбинации. Встраиваемые кассеты находились в составе плазмидных векторов и были окружены плечами гомологии к определенному участку генома AAVS1, который является safe-harbor локусом. AAVS1 находится на 19 хромосоме внутри первого интрона гена *PPP1R12C*, кодирующего регуляторную субъединицу фосфатазы миозина. Назвали этот локус AAVS1, потому что в него *in vitro* интегрировался аденоассоциированный вирус AAV, хотя *in vivo* это и не является местом предпочтительной интеграции этого вируса. Так или иначе, было показано, что трансгены, встроенные в этот локус, демонстрируют стабильную экспрессию, отчего его часто используют, как место для интеграции трансгенов [368]. Также мы использовали стратегию генной ловушки: в плазмидных векторах перед генами устойчивости к антибиотикам находились последовательность 3'-сайта сплайсинга и последовательность T2A-пептида из вируса бабочки-мокрицы *Thosea asigna*, позволяющая в процессе трансляции отщепиться следующему за ним участку пептида [369]. Так как эти гены не имели своего собственного промотора, то экспрессировались только при встраивании кассеты с ними в интрон гена, так как в этом случае могли образоваться химерные мРНК, с которой транслировался их белковый продукт. Проскок терминатора и альтернативный сплайсинг мРНК гена *PPP1R12C* позволяет синтезироваться в клетке и его нормальным белковым продуктам.

Для индукции разрывов в сайте AAVS1 также были необходимы программируемые нуклеазы, направленные на последовательность AAVS1. Однако в этом случае уже не могла быть использована система CRISPR/Cas9, так как гены гидовых РНК находятся под конститутивным промотором, и при внедрении в клетки плазмиды, постоянно и на высоком уровне экспрессирующей Cas9, целевые транслокации формировались бы еще на стадии получения клеточной культуры. Поэтому для внесения разрывов в AAVS1 необходимо использовать либо ортологичную систему CRISPR/Cas, нуклеаза которой не будет направляться РНК-гидами для Cas9, либо TALE-нуклеазы. В нашем случае выбор конкретных решений был обусловлен существованием набора плазмид, на основе которого мы могли легко получить все необходимые нам конструкции. Так, мы использовали готовые плазмиды, кодирующие субъединицы TALE-нуклеазы на локус AAVS1, и плазмиду AAVS1-M2rtTA, содержащую кассету с геном белка-активатора, геном устойчивости к неомицину и плечи гомологии к AAVS1. Еще одна плаزمида была собрана на базе существующей Puro_Cas9 donor, в которой была фланкированная плечами гомологии к AAVS1 кассета, содержащая Cas9 под промотором TRE, и ген устойчивости к пурамицину. Нашей задачей было вставить в нее гены гидовых РНК, направляющих Cas9 на локусы, между которыми будет формироваться транслокация.

Для вставки генов гидовых РНК Puro_Cas9 donor в мы использовали метод сборки Гибсона. Метод основан на образовании длинных одноцепочечных концов на соединяемых фрагментах ДНК и отжиге таких комплементарных концов друг на друга. Он позволяет вставлять в одной реакции сразу несколько фрагментов, причем в любое место плазмиды, где мы можем внести разрыв. Сборка плазмиды Puro_Cas9_gRNA1+2 описана в соответствующем разделе Материалов и методов.

Получение и проверка генов гидовых РНК

Сначала нам требовалось подобрать последовательности гидовых РНК, которые бы направляли нуклеазу Cas9 в нужные локусы. В качестве этих локусов мы выбрали области внутри *AML1* и *ETO*, где по литературным данным происходят разрывы, приводящие к лейкозогенной перестройке *AML-ETO*, — это области интрона 5 гена *AML1* и интрона 2 гена *ETO* [203]. Внутри этих областей мы подобрали по три последовательности, которые бы узнавались РНК-гидами. Для подбора использовали онлайн-инструмент crispr.mit.edu. Было подобрано сразу несколько вариантов, так как не все подобранные даже с помощью специальных сервисов последовательности будут достаточно эффективно разрезаться соответствующим комплексом Cas9–РНК-гид [370].

Чтобы сконструировать гены гидовых РНК, мы использовали плазмиду phU6-gRNA, в которой уже есть промотор, терминатор и ген гидовой РНК, за исключением участка, кодирующего переменную узнающую часть РНК. Этот двадцатинуклеотидный участок мы заказали в виде одноцепочечных комплементарных олигонуклеотидов, а затем встроили его в плазмиду phU6-gRNA с помощью рестриктазы IIS-типа, сайты узнавания которой находятся внутри спейсера, выщепляемого ей из последовательности гена РНК-гида.

Для выбора эффективно работающих РНК-гидов можно использовать различные способы [371]. Однако в процессе работы совместно с коллегами института Гюстава Русси (Вильжуиф, Франция) мы разработали наиболее удобный, на наш взгляд, метод проверки эффективности нуклеаз — инструментов редактирования генома. Метод назван «Engineered Nuclease-induced Translocations» (ENIT) [372]. Он предполагает котрансфекцию клеток плазмидами, кодирующими нуклеазы и гидовые РНК (в случае проверки работы элементов системы CRISPR/Cas). Если нуклеазы эффективно вносят разрывы в необходимые локусы ДНК, то при репарации с некоторой вероятностью образуются хромосомные перестройки между точками разрывов. Спустя 48 часов после трансфекции с ДНК клеток ставится ПЦР на предмет наличия ожидаемых транслокаций, которые происходят между разрезаемыми локусами. В этом методе можно использовать любые легко трансфицируемые клетки, например, HeLa.

Искомая перестройка случается далеко не во всех клетках трансфицируемой культуры. Во-первых, при трансфекции не каждая клетка получает и начинает

экспрессировать все плазмиды, которыми ее трансфицируют. Во-вторых, чтобы искомая транслокация случилась, сделанные нуклеазами разрывы должны сшиться, что происходит только при ошибках репарации. Для надежного обнаружения транслокаций необходимо сначала оптимизировать ПЦР — подобрать наилучшие пары праймеров и условия реакции, чтобы редкий продукт транслокации специфически амплифицировался на фоне большого количества копий геномной ДНК.

Чтобы не заниматься предварительной оптимизацией ПЦР, для метода проверки гидов мы выбрали другой подход — вложенную (nested) ПЦР, которая проводится в две стадии [373]. Сначала ставится небольшое количество циклов ПЦР1 со внешними праймерами. Затем реакционная смесь разбавляется в 100–1000 раз и с ней ставится вторая стадия ПЦР2 уже со внутренними праймерами. Неспецифические продукты амплификации посторонних последовательностей, которые неизбежно появляются во время ПЦР с геномной ДНК, не будут амплифицироваться при ПЦР со внутренней парой праймеров. Метод отличается высокой чувствительностью, так как за две стадии проводится большее число циклов реакции, чем при одностадийной ПЦР. Эффективность вложенной ПЦР позволяет существенно ускорить и упростить процесс подготовки образцов для реакции — вместо выделения ДНК из клеток с помощью специальных коммерческих наборов мы используем «прямую» ПЦР (Direct PCR). В этом случае небольшое количество клеток суспендируется в ПЦР-премиксе, обрабатывается протеиназой К, после инактивации протеиназы сразу ставится первая реакция ПЦР1, для чего в раствор добавляются ДНК-полимераза и соответствующие праймеры (см. Материалы и методы).

Мы проверяли наши гиды по-отдельности: котрансфицировали клетки линии HeLa одной из собранных плазмид с геном гидовой РНК, плазмидой с геном нуклеазы Cas9 (#41815, Addgene) и двумя плазмидами, которые кодируют субъединицы нуклеазы TALEN (TALEN_8F и TALEN_8R), вносящей двуцепочечный разрыв в районе промотора гена *MYC* на восьмой хромосоме. Плазмиды с этой TALE-нуклеазой уже были в нашей лаборатории, и было показано, что она достоверно вносит двуцепочечный разрыв в указанном локусе. Если Cas9 в комплексе с проверяемой гидовой РНК вносит двуцепочечный разрыв в узнаваемом ею локусе, между этим локусом и геном *MYC* с определенной вероятностью возникает хромосомная транслокация (рисунок 4.8).

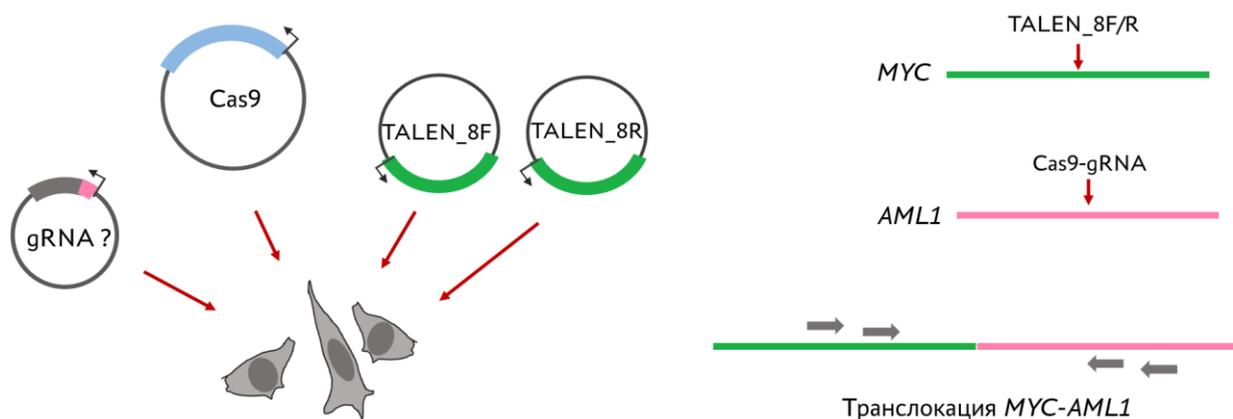


Рисунок 4.8. Принцип метода ENIT, используемого для проверки внесения двуцепочечных разрывов системой CRISPR/Cas9. В качестве примера — гидовая РНК gRNA_AML1. Слева — трансфекция клеток плазмидами, справа — образование транслокации, детектируемой вложенной ПЦР (горизонтальные стрелки — внешние и внутренние праймеры). Вертикальными стрелками отмечены места внесения разрывов в гены *AML1* и *MYC*, соответственно. «?» обозначает проверяемую гидовую РНК.

Всего возможны 4 варианта сшивки между концами двух разных разрывов (рисунок 4.9). Соответственно, можно обнаружить 4 транслокации, если использовать соответствующие пары праймеров. Длины продуктов ПЦР будут отличаться по размеру, что хорошо заметно на электрофорезе.

Результаты проверки гидовых РНК к локусам *AML1* и *ETO*

Спустя двое суток после трансфекции клеток HeLa мы собрали клетки и провели несколько реакций вложенной ПЦР, используя комбинации праймеров для каждой из 4-х возможных перестроек (рисунок 4.9). В качестве отрицательного контроля использовалась геномная ДНК из нетрансфицированных клеток HeLa. Результаты электрофореза продуктов второй стадии вложенной ПЦР — на рисунке 4.10.

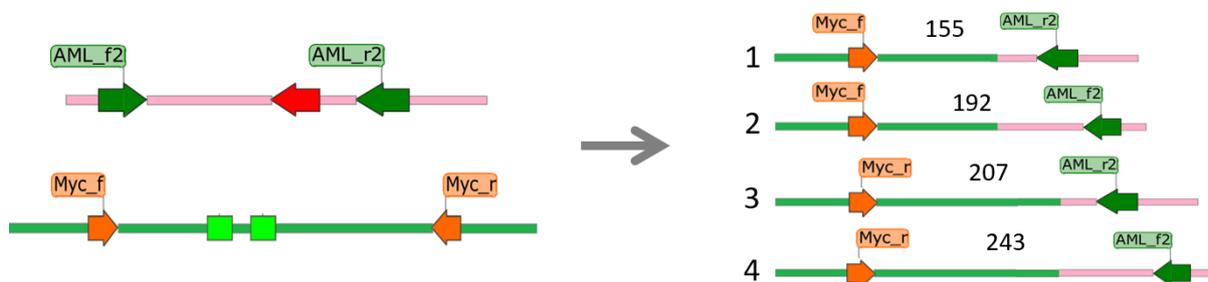


Рисунок 4.9. 4 варианта транслокаций, возникающих при внесении разрывов в области генов *AML1* и *C-MYC*. Последовательность, узнаваемая РНК-гидом *AML1* отмечена красной стрелкой, нуклеазами TALEN_8F/R — зелеными прямоугольниками. Различные варианты перестроек дают разные длины продуктов ПЦР, что хорошо различается на электрофорезе. На схеме не показаны внешние праймеры, используемые на первой стадии ПЦР.

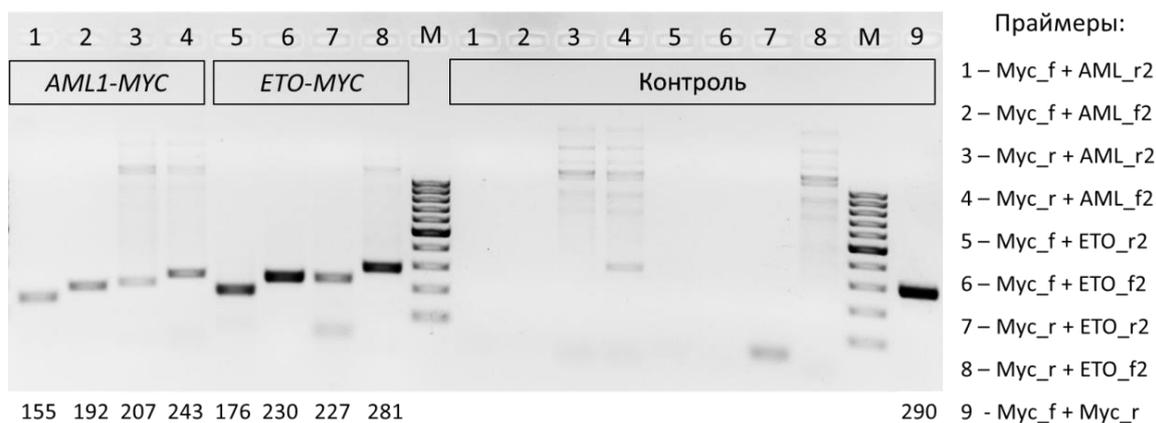


Рисунок 4.10. Детектирование транслокаций *AML1-MYC* и *ETO-MYC* в клетках HeLa. На рисунке — результат электрофореза продуктов второй стадии вложенной ПЦР. В качестве матрицы в ПЦР использовалась: дорожки «*AML1-MYC*» — ДНК клеток, трансфицированных плазмидами p_gRNA_AML1, Cas9, TALEN_8F и TALEN_8R; дорожки «*ETO-MYC*» — ДНК клеток, трансфицированных плазмидами p_gRNA_ETO, Cas9, TALEN_8F и TALEN_8R; дорожки «Контроль» — ДНК нетрансфицированных клеток HeLa. Над каждой дорожкой цифрой указана пара праймеров, использовавшаяся при проведении второго раунда вложенной ПЦР (название праймеров справа на рисунке). Под дорожками — ожидаемый размер продукта транслокации. М — маркер молекулярных масс с шагом в 100 п.н. Пара праймеров Myc_f/Myc_r использовалась в качестве положительного контроля.

При электрофорезе ПЦР-продуктов на дорожках, соответствующих трансфицированным клеткам (рисунок 4.10, левая половина), видны полосы, которых нет после ПЦР с ДНК контрольных клеток. При этом размер таких ПЦР-продуктов во всех случаях соответствует ожидаемому, который легко вычисляется из расстояния между предполагаемыми точками разрыва и сайтами посадки соответствующих праймеров. Как в случае транслокации *AML1-MYC*, так и в случае транслокации *ETO-MYC*, мы наблюдали все четыре возможных способа соединения концов разрывов (рисунок 4.9). Для подтверждения того, что обнаруживаемые при электрофорезе полосы действительно соответствуют транслокациям, некоторые ПЦР-продукты были секвенированы по Сэнгеру (рисунок 4.11). Можно видеть, что в последовательности каждого ПЦР-продукта четко локализуются участки, произошедшие от генов-партнеров по перестройке. Как и ожидалось, в гене *MYC* разрыв происходил между сайтами посадки субъединиц TALEN-нуклеазы. В генах *AML1* и *ETO* разрывы происходят на расстоянии нескольких нуклеотидов от 3'-конца гидовой РНК, что также соответствует данным о механизме работы Cas9 [374],[375].

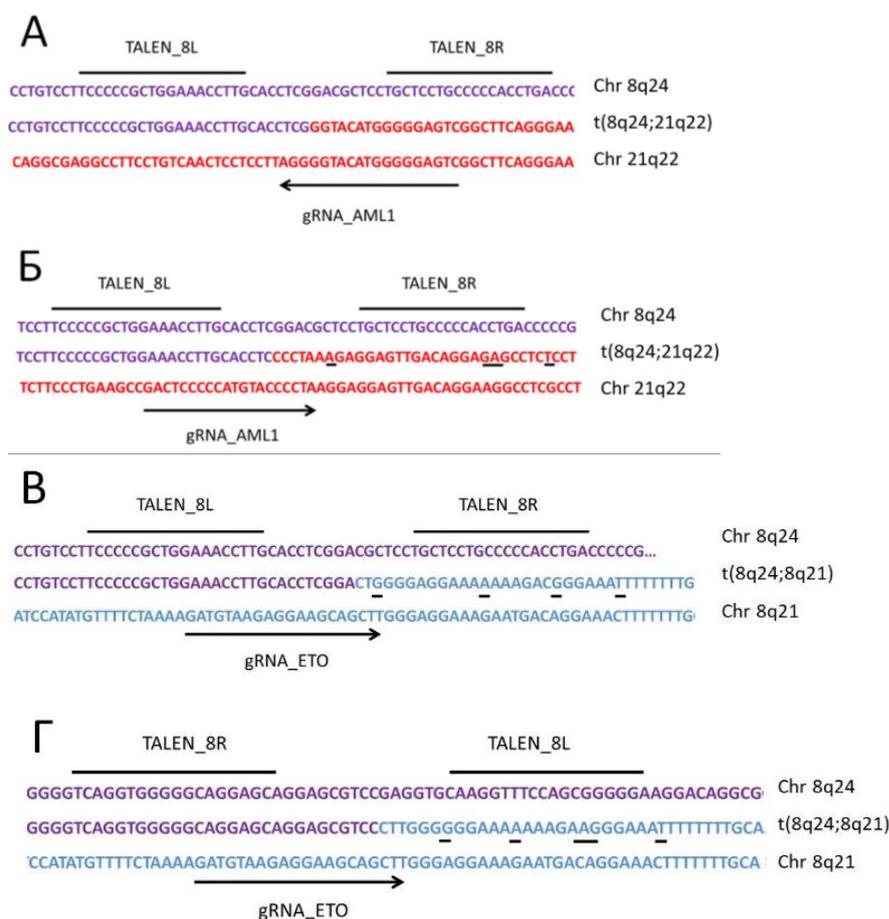


Рисунок 4.11. Последовательности транслокаций *MYC-AML1* и *MYC-ETO* (последовательность в середине), выравненные относительно генов-партнеров по перестройкам (последовательности сверху и снизу). Последовательности получены при секвенировании следующих ПЦР-продуктов: **А** — *Мyc_f+AML_r2*, **Б** — *Мyc_f+AML_f2*, **В** — *Мyc_f+ETO_r2*, **Г** — *Мyc_r+ETO_r2*. Во всех случаях последовательность гена *MYC* показана фиолетовым цветом, гена *AML1* — красным, а гена *ETO* — синим. Отмечены сайты, распознаваемые субъединицами TALEN_8F и TALEN_8R в гене *MYC*, а также участки, узнаваемые гидовыми РНК в генах *AML1* и *ETO*. Черными линиями подчеркнуты различия между последовательностями транслокаций и последовательностями геномных локусов.

Отметим, что каждый ПЦР-продукт в действительности может представлять собой гетерогенную смесь продуктов амплификации транслокаций, произошедших в разных клетках. В пользу данного предположения свидетельствуют первичные данные, полученные в ходе секвенирования (рисунок 4.12). Можно видеть, что качество прочтения резко падает в районе предполагаемой точки разрыва. При секвенировании использовался праймер со стороны гена *MYC*, поэтому со стороны гена *MYC* до точки разрыва последовательность всех ПЦР-продуктов одинакова. Однако после точки разрыва из-за отдельных инсерций или делеций, которые часто происходят при негомولوجном соединении концов разрывов, последовательности оказываются «рассинхронизованы». Тем не менее, наиболее вероятная последовательность, которая определяется секвенированием, соответствует лигированию концов разрыва стык в стык.

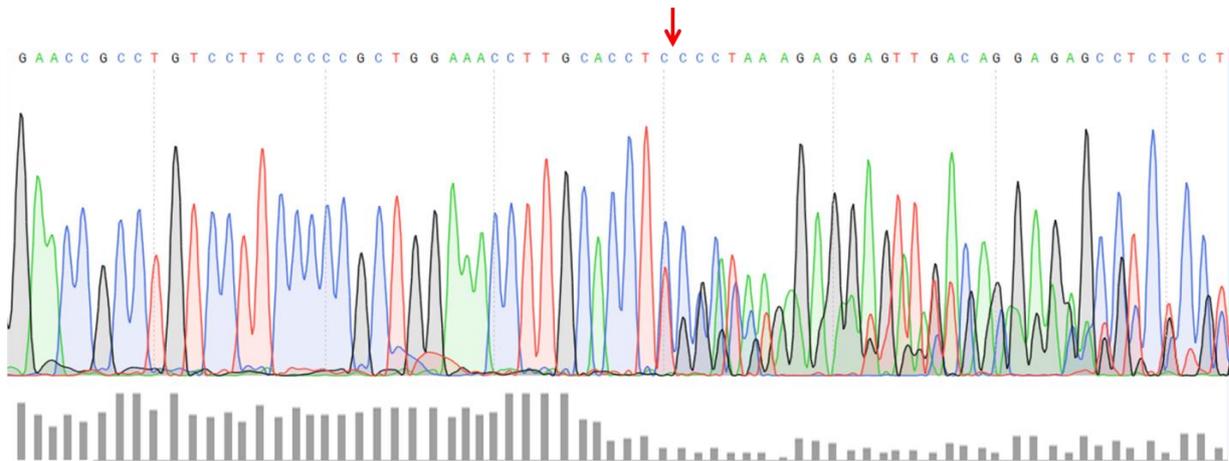


Рисунок 4.12. Хроматограмма, полученная в ходе секвенирования ПЦР-продукта Мус_f+AML_f2. Для секвенирования использовался праймер Мус_f. Сверху показана последовательность, автоматически определяемая прибором как наиболее вероятная. Красной стрелкой отмечена точка разрыва. Столбчатая гистограмма снизу отражает качество определения соответствующего нуклеотида.

Результаты проверки гидов демонстрируют, что Cas9 в комплексе с гидовыми РНК gRNA_AML1 и gRNA_ETO эффективно вносит разрывы в требуемые локусы *AML1* и *ETO*, и что для индукции хромосомной транслокации в геноме клеток человека достаточно внести двуцепочечные разрывы в целевые локусы. Ошибки, возникающие в ходе репарации этих разрывов, приводят к появлению хромосомных перестроек.

4.4. Получение культуры клеток iAML/ETO

Для получения нашей клеточной модели мы использовали культуру LCL, поскольку она является удобной в работе линией лейкоцитов человека. Итак, мы трансфицировали клетки этой культуры одновременно четырьмя плазмидами: Puro_Cas9_AML1_ETO, AAVS1-M2rtTA и двумя плазмидами, содержащими гены субъединиц нуклеазы TALEN, которая вносит разрыв по сайту AAVS1 (рисунок 4.13).

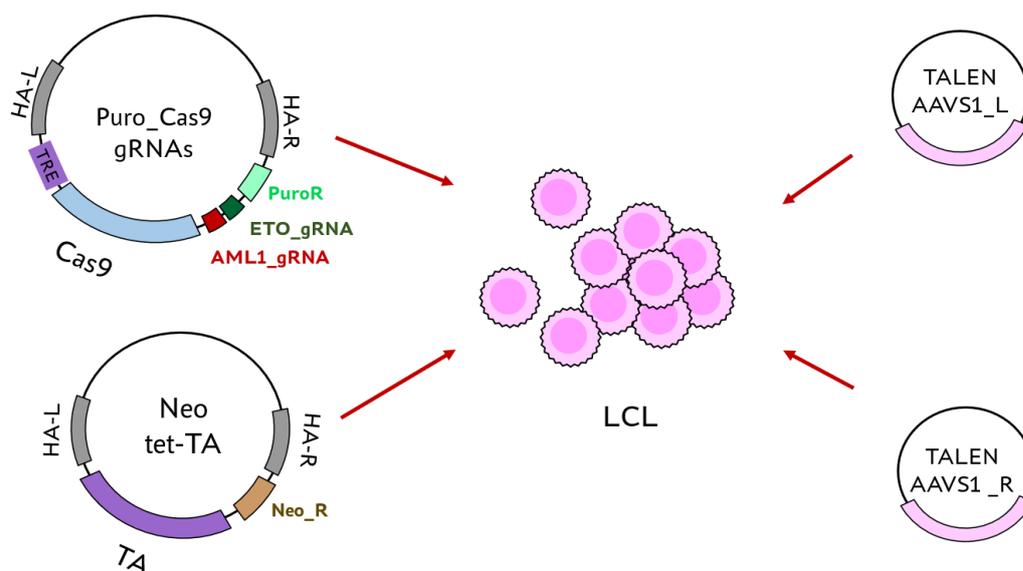


Рисунок 4.13. Схема трансфекции клеток LCL плазмидами с интегрируемыми генными кассетами, содержащими элементы индуцируемой системы CRISPR/Cas, и плазмидами, кодирующими субъединицы TALEN, индуцирующей разрыв в локусе AAVS1. HA-L/R — плечи гомологии к областям слева и справа от данного разрыва, соответственно. TA — транскрипционный активатор. PuroR и NeoR — гены устойчивости к пурамицину и неомицину, соответственно, TRE — индуцируемый промотор.

После попадания в клетку плазмид начинается транзистентная экспрессия, и начинает нарабатываться TALEN-нуклеаза, которая вносит разрыв в локусы AAVS1 на обеих гомологичных хромосомах. В результате репарации по механизму гомологичной рекомбинации в место разрывов будут встраиваться кассеты, окруженные плечами гомологии. В ряде случаев в одну гомологичную хромосому клетки встроится кассета из плазмиды Puro_Cas9+gRNA_AML1_ETO, содержащая гены гидовых РНК gRNA_AML1 и gRNA_ETO, ген нуклеазы Cas9 и ген устойчивости к пурамицину, а в другую гомологичную хромосому встроится кассета из плазмиды AAVS1-M2rtTA, содержащая ген активатора транскрипции TA и ген устойчивости к генетицину (G-418). Селекция трансфицированных клеток на генетицине и пурамицине позволяет отобрать клетки, содержащие в геноме обе конструкции. Для краткости в тексте работы она будет называться просто культурой iAML/ETO. Подробнее процесс получения этой клеточной культуры описан в разделе «Материалы и методы».

4.5. Индукция транслокации *AML1-ETO* в клетках iAML/ETO

Согласно литературным данным [376], [377], механизм активации экспрессии Cas9 в культуре iAML/ETO следующий: конститутивно экспрессирующийся активатор транскрипции (TA) связывается в присутствии доксициклина с промотором гена Cas9 и запускает его транскрипцию (рисунок 4.7). Внесение двуцепочечных разрывов запускает

процесс их репарации, приводящий в некоторых случаях к образованию хромосомных перестроек. Для дальнейшей работы с клетками iAML/ETO важно было понять, как изменяется уровень экспрессии Cas9 после активации доксициклином и через какой промежуток времени мы можем достоверно детектировать хромосомные транслокации.

Для активации экспрессии Cas9 к клеткам культуры iAML/ETO добавляли доксициклин в концентрации 1 мкг/мл. Затем, спустя различные промежутки времени (4 часа, 8 часов, 24 часа, 32 часа, 48 часов, 60 часов и 72 часа) отбирали по 4 млн клеток для выделения РНК и ДНК.

Кинетика экспрессии трансгенов в клетках культуры iAML/ETO

С выделенной РНК синтезировали кДНК, с которой далее ставили ПЦР в реальном времени. Перед проведением ПЦР в реальном времени мы поставили ПЦР с детекцией в конечной точке для подтверждения специфичности праймеров и соответствия размера получаемого продукта ожидаемому. Результаты трех биологических повторностей эксперимента по анализу экспрессии Cas9 и TA с помощью количественной ПЦР показаны на рисунке 4.14. Под биологическими повторностями в данном случае подразумеваются эксперименты, проведенные на полученных клетках iAML-ETO в разные дни. Можно видеть, что в каждой индивидуальной повторности экспрессия Cas9 быстро увеличивается после добавления доксициклина, достигая максимума примерно к 24 часам (рисунок 4.14 А и Б). Далее уровень экспрессии колеблется вокруг максимального значения. Необходимо заметить, что небольшая экспрессия Cas9 обнаруживается и до активации ее экспрессии: значения R равны 0,02, 0,03 и 0,01 для первой, второй и третьей повторности, соответственно. Тем не менее, после активации экспрессия Cas9 увеличивается более чем на два порядка во всех повторностях (рисунок 4.14 Б). С другой стороны, экспрессия активатора транскрипции значительно не изменяется после добавления доксициклина, что соответствует представлениям о конститутивном характере его экспрессии (рисунок 4.14 В и Г). Таким образом, характер экспрессии Cas9 и TA в активированных клетках полностью соответствовал ожидаемому.

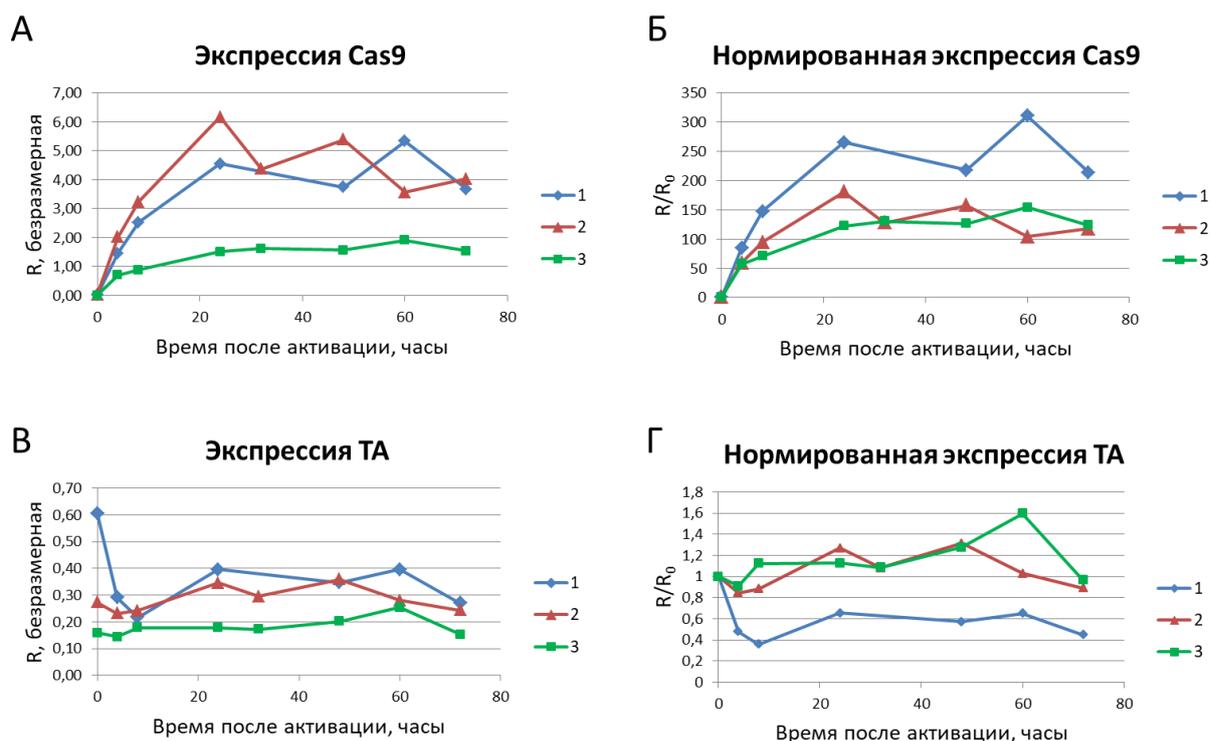


Рисунок 4.14. Анализ экспрессии Cas9 и TA в клетках iAML/ETO после добавления к клеткам доксициклина. На каждом графике точка с абсциссой «0» соответствует клеткам, отобраннным перед добавлением доксициклина. Последняя временная точка — 72 часа. Кривые «1», «2» и «3» на каждом графике соответствуют биологическим повторностям эксперимента (нумерация сквозная через все графики). Значение R представляет собой уровень экспрессии указанного гена относительно экспрессии гена GAPDH (см. Материалы и методы). На графиках А и В отложена зависимость значения R от времени для генов Cas9 и TA, соответственно. Для построения графиков Б и Г значение R в каждой точке делилось на значение R_0 (уровень экспрессии в клетках до активации).

Кинетика индукции транслокаций в клетках культуры iAML/ETO

С выделенной из активированных клеток геномной ДНК проводили вложенную ПЦР на транслокацию *AML1-ETO* (рисунок 4.15). Этот подход характеризуется более высокой чувствительностью и сниженным риском амплификации посторонних последовательностей [372]. Можно видеть, что полоса ожидаемого размера появляется спустя 48 часов после активации экспрессии Cas9. Отсутствие этой полосы в клетках до активации (точка «0 часов») позволяет предположить, что данный фрагмент ДНК соответствует продукту транслокации.

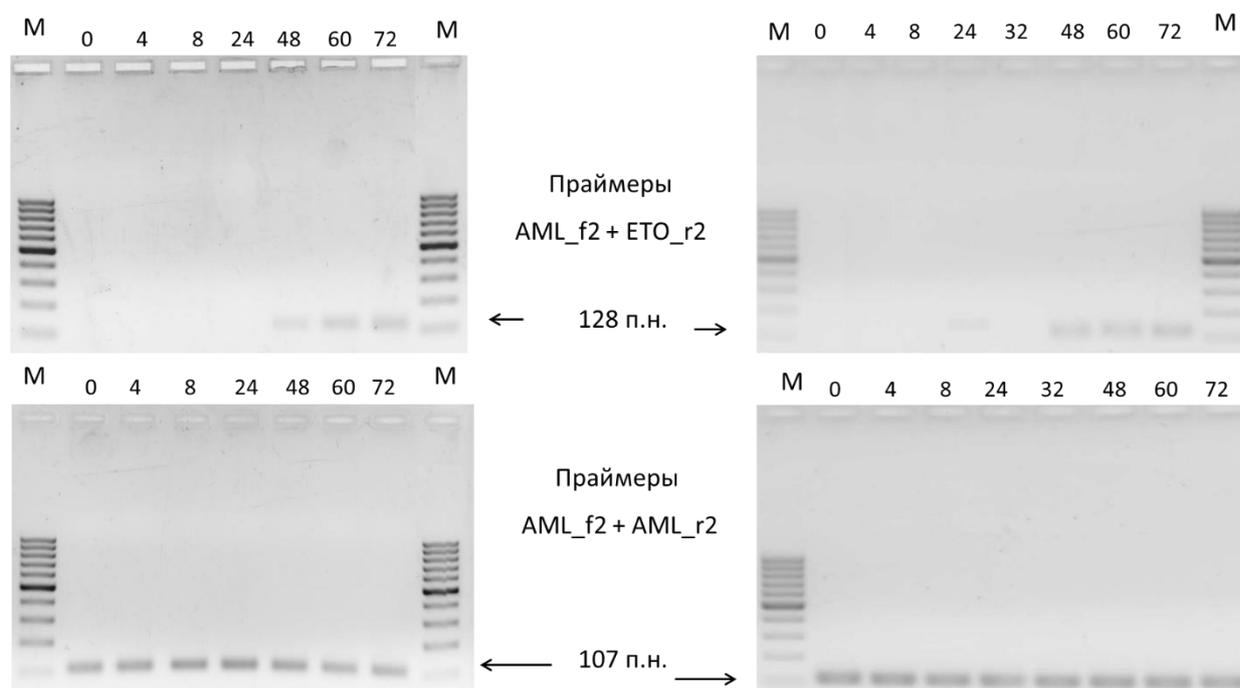


Рисунок 4.15. Кинетика индукции транслокаций в активированных клетках iAML/ETO. Показаны результаты второго раунда вложенной ПЦР с ДНК, выделенной из активированных клеток двух повторностей эксперимента. Цифры над дорожками показывают время выделения ДНК из клеток с момента активации. Точка «0» соответствует клеткам, отобраным перед добавлением доксициклина. На рисунке отмечены ожидаемые размеры полос. Пара праймеров AML_f2 + AML_r2 (электрофорезы снизу) использовалась в качестве положительного контроля ПЦР.

Секвенирование по Сэнгеру продукта ПЦР из точки 48 часов подтвердило, что обнаруживаемая полоса действительно соответствует продукту транслокации. В отличие от транслокаций, индуцированных при проверке гидов, сиквенс транслокаций в клетках iAML1-ETO полностью совпадал с ожидаемым.

Проведенные эксперименты подтверждают индукцию транслокации *AML1-ETO* в клетках iAML/ETO после активации экспрессии Cas9 доксициклином. Транслокация впервые детектируется спустя 48 часов после добавления к клеткам доксициклина. Важно подчеркнуть, что транслокации могут возникать и раньше, однако только к 48 часам накапливается необходимое количество перестроек, детектируемое в нашем эксперименте.

Количественная детекция транслокаций

Важное свойство созданной клеточной модели — возможность изучать на ней факторы, влияющие на образование транслокаций. Для этой цели необходимо иметь возможность определять количество транслокаций, во-первых, точно, а во-вторых, воспроизводимо, чтобы сравнивать между собой число транслокаций в разных условиях. Поэтому мы уделили особое внимание разработке методики экспериментов на этой культуре клеток по подсчету числа транслокаций.

Основной проблемой детекции транслокаций методом ПЦР является низкая представленность амплифицируемого участка (мишени) среди прочей геномной ДНК. В связи с этим резко повышается требование к специфичности применяемой ПЦР. Поэтому в качестве метода детекции продукта в количественной ПЦР был выбран метод с TaqMan-зондом, как более защищенный от нецелевого сигнала, чем метод детекции продукта по накоплению сигнала от интеркалирующего красителя. Для ПЦР были подобраны праймеры и TaqMan-зонды. Для подбора оптимальных условий ПЦР, а также в качестве стандарта для сравнения, была создана плазмидная конструкция pUC_EA, содержащая участок транслокации. С помощью плазмиды была определена оптимальная температура отжига праймеров, а также было проверено, что эффективность наших праймеров не меняется при добавлении 1 мкг «балластной» геномной ДНК: поставили ПЦР с тремя разведениями чистой плазмиды и тремя теми же разведениями, но с добавлением геномной ДНК клеток LCL, и Ct были одинаковыми (рисунок 4.16). Наконец, была отработана процедура постановки ПЦР с учетом необходимости максимально точно отбирать объемы реактивов и соблюдения всех мер предосторожности против контаминации ПЦР-продуктами. Отметим, что все условия ПЦР, флуорофоры TaqMan-зонда и праймеры подбирались, исходя из потенциального использования цифровой капельной ПЦР для еще более точного определения числа транслокаций в нашей экспериментальной системе. В дальнейшем полученную плазмиду использовали для построения калибровочных кривых из нескольких концентраций при ПЦР.

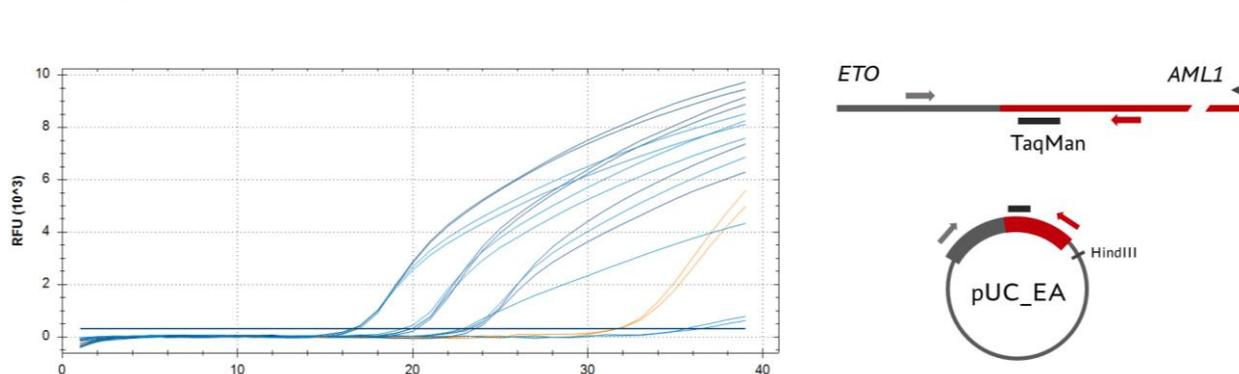


Рисунок 4.16. Кривые амплификации трех образцов с разным содержанием линейаризованной плазмиды pUC_EA (фиолетовые и синие кривые), образца ДНК клеток спустя 48 часов после индукции транслокаций (оранжевые кривые) и образца ДНК без индукции (линии с Ct \approx 36). Каждый образец представлен в двух повторностях, каждое из разведений плазмиды представлено в двух вариантах: чистом и с «балластной» ДНК. Справа схема области транслокации и плазмиды pUC_EA. Цветные стрелки обозначают праймеры.

Для экспериментов по определению числа транслокаций клетки обрабатывали доксициклином, через 3–4 дня из клеток выделяли ДНК и ставили количественную ПЦР по отработанному протоколу. Так, с помощью количественной ПЦР была выбрана наиболее

эффективная концентрация доксициклина для индукции транслокаций — 0,1 мкг/мл. Затем была изучена кинетика накопления транслокаций в культуре клеток после индукции доксициклином: от популяции активированных клеток спустя определенные промежутки отбирали часть, выделяли ДНК и ставили количественную ПЦР на транслокацию (рисунок 4.17). Было определено, что число транслокаций варьирует в зависимости от условий культивирования клеток — при низких концентрациях клеток в суспензии число транслокаций меньше. В достаточно плотной культуре клеток (0,7–1 млн/мл) число транслокаций на 3–4 сутки после индукции доксициклином может достигать 10–12 транслокаций на 1000 клеток. Для подсчета количества транслокаций в 1 мкг ДНК использовали калибровочную кривую, построенную по результатам ПЦР с разными количествами плазмиды — от 1 фг до 10 пг (в пересчете на количество плазмид — от ≈ 340 штук). Количество ДНК, которое добавляли в реакцию, определялось с помощью флуориметрического измерения концентрации ДНК и подтверждалось с помощью внутреннего контроля — количественной ПЦР на интактный аллель *AML1* (оба способа определения количества ДНК давали практически идентичные результаты).

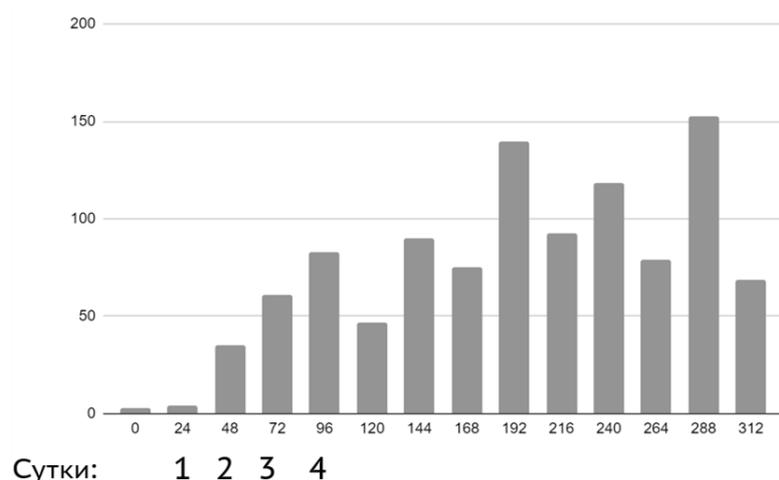


Рисунок 4.17. Кинетика накопления транслокаций в культуре клеток iAML-ETO после активации экспрессии гена Cas9 доксициклином. Показаны результаты одной из повторностей экспериментов, другие повторности давали аналогичные результаты.

Анализ методом FISH генов *AML1* и *ETO* после индукции экспрессии Cas9

Мы также изучили с помощью FISH с двуцветной пробой к генам *AML1* или *ETO* клетки iAML-ETO, была индуцирована экспрессия Cas9. Через 24 часа после добавления к клеткам доксициклина были приготовлены препараты клеток, с которыми затем провели FISH. Однако анализ снимков не показал наличие разрывов этих генов, как в случае обработки этопозидом.

4.6. Влияние различных веществ на число транслокаций

Влияние ингибиторов комплексов репарации, на число транслокаций AML1-ETO

В качестве проверки, что на клеточной культуре iAML-ETO можно наблюдать эффект от обработки клеток различными веществами, заключающийся в изменении числа транслокаций, были поставлены следующие эксперименты. Индуцировали целевую перестройку и одновременно обрабатывали клетки некоторыми ингибиторами белков репарации: NU7026 — ингибитором белка репарации DNA-PKcs — и мирином — ингибитором комплекса репарации MRN. Выбор веществ был обусловлен результатами работы по изучению влияния различных ингибиторов репарации на образование другой транслокации — IGH-MYC [378]. Для изучения формирования этой транслокации мы получили клеточную культуру по той же схеме, что и культуру iAML-ETO. Среди проанализированных веществ выраженный эффект — увеличение числа транслокаций — наблюдался в результате ингибирования DNA-PKcs. Мирин же, напротив, несколько снижал число транслокаций IGH-MYC.

Мы добавляли в среду к клеткам культуры iAML-ETO NU7026 или мирин одновременно с добавлением доксициклина. Также был контроль с клетками, в которых добавляли только доксициклин. Затем на 3–4 день выделяли из клеток ДНК и проводили ПЦР по отработанному ранее протоколу. Данные вещества показали некоторое увеличение частоты транслокаций в системе относительно базового уровня, то есть числа транслокаций в контрольном эксперименте (рисунок 4.18).

Влияние этопозида на число транслокаций AML1-ETO

Аналогичный эксперимент на линии iAML-ETO был поставлен с этопозидом. При одновременной индукции перестроек и обработке клеток этопозидом (в концентрации IC10) частота транслокации AML-ETO было несколько ниже, чем в контроле — в клетках, обработанных только доксициклином (рисунок 4.18).

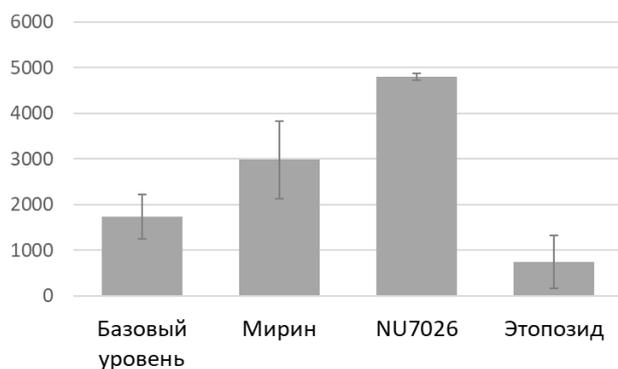


Рисунок 4.18. Число транслокаций (на миллион клеток), образующихся при ко-обработке клеток линии iAML-ETO ингибиторами белков репарации — NU7026 и мирина, а также этопозида.

Влияние препаратов, используемых в химиотерапии, на число транслокаций *AML1-ETO*

В современных протоколах химиотерапии опухолей, как правило, используются комбинации различных веществ. Поэтому нашей задачей было сравнить различные используемые в терапии соединения с точки зрения их опасности для пациента — риска индукции лейкозогенной транслокации. На полученной клеточной линии *iAML-ETO* был проведен скрининг веществ, применяющихся в химиотерапии опухолей. Подбор препаратов и их рабочих концентраций осуществлялся вместе с коллегами из лаборатории комбинированной терапии РОНЦ им.Блохина В.С. Покровским и Д.В. Соколовой. Были протестированы следующие препараты: антиметаболиты (цитарабин, метотрексат), препараты на основе платины (карбоплатин, цисплатин, оксалиплатин), алкилирующий агент (ифосфамид), ингибитор ДНК-топоизомеразы I (иринотекан), ингибитор сборки микротрубочек (паклитаксел). Сначала было проведено исследование цитотоксичности данных препаратов на клетки культуры *iAML-ETO*. Затем уже проводился эксперимент, когда препараты добавлялись к клеткам одновременно с доксициклином в нелетальных концентрациях $\approx IC_{10}$. Было проведено 2 независимых эксперимента, в каждом из которых были контрольные точки («базовый уровень»). В результате можно наблюдать, что метотрексат приводит к повышению частоты индуцированных транслокаций *AML1-ETO* относительно базового уровня, соответствующего частоте транслокаций в клетках, обработанных только доксициклином (рисунок 4.19). Данный результат для метотрексата затем был отдельно подтвержден в двух независимых экспериментах.

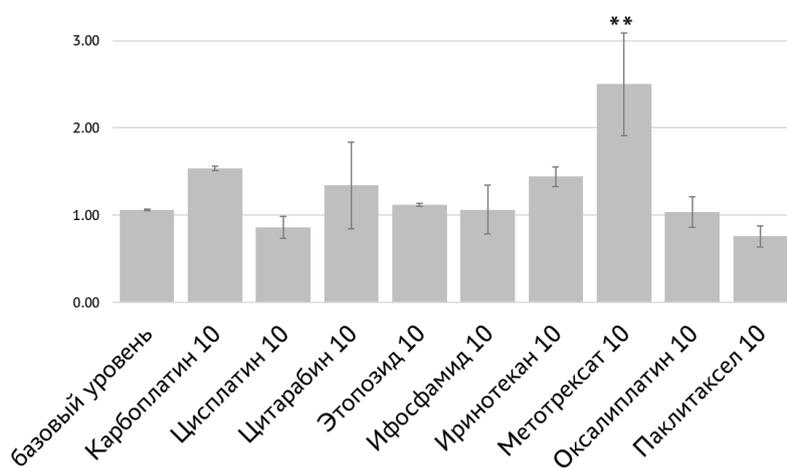


Рисунок 4.19. Результаты скрининга с помощью культуры *iAML-ETO* соединений, применяющихся в химиотерапии опухолей совместно с ингибиторами ДНК-топоизомеразы II. Оценивалось число транслокаций при ко-обработке клеток этими соединениями дозой IC_{10} вместе с доксициклином относительно базового уровня (обработка только доксициклином).

5. Обсуждение результатов

Факторы, определяющие хромосомные перестройки

Очевидно, что в клетке могут формироваться хромосомные перестройки между теоретически между любыми локусами. Однако только физиологически значимые транслокации приводят к получению клеткой преимущества в пролиферации и формированию лейкозного клона. Но нельзя сказать, что лишь пролиферативное преимущество является фактором, определяющим набор транслокаций: известны сотни лейкозогенных транслокаций, в том числе с участием генов *AML1* и *MLL*. Набор их генов-партнеров при *de novo* лейкозах гораздо шире, чем при лейкозах, вызванных терапией ингибиторами ДНК-топоизомераз II. Кроме того, для вторичных лейкозов, вызванных терапией алкилирующими агентами или радиацией, набор хромосомных aberrаций будет отличаться [15], [27]. Поэтому в данной работе исследовалось, какие факторы влияют на образование транслокаций под действием ингибиторов ДНК-топоизомераз II типа, в частности, этопозида.

Работы по изучению механизмов, связанных с образованием хромосомных транслокаций, демонстрируют важность разных факторов для определения того, какая перестройка вероятнее произойдет. Некоторые исследования показывают взаимосвязь между частотой контактов перестраивающихся генов и частотой транслокаций между ними [314], [315], [319], [379]. В других работах подчеркивается повышенная мобильность разрывов по сравнению с интактными локусами, которая повышает вероятность транслокации между не сближенными изначально локусами [317], [322], [380]. Эти результаты обычно интерпретировались в контексте двух гипотез: гипотеза о первенстве контакта (“contact-first”) подразумевает, что транслокация происходит вследствие ошибочной репарации разрывов, которые произошли в находящихся рядом локусах. Гипотеза о первенстве разрыва (“breakage-first”) предполагает, что ДЦР могут перемещаться на значительные расстояния в ядре перед тем, как встретиться с другим разрывом для транслокации, что отменяет необходимость в исходной пространственной близости перестраивающихся локусов [316], [381]. Однако накопленные данные не позволяют выявить единый механизм, однозначно определяющий возникающие транслокации. Скорее, все факторы влияют на образование транслокаций, но в разной степени для разных ситуаций [382]. Так, в работе по визуализации процесса образования транслокаций в живых клетках было показано, что и изначально близкие, и далекие локусы формировали транслокации после внесения в них разрывов, однако чаще происходили транслокации между близкими локусами [320].

Вдобавок к факторам исходной близости локусов и мобильности разрывов, на образование определенных транслокаций, очевидно, влияет вероятность возникновения

разрывов в том или ином локусе [314], [327]–[329], а также способ, которым был индуцирован разрыв. Так, *in vivo* наблюдение за фокусами белка 53BP1, который связывается с разрывами ДНК, показывает более высокую подвижность разрывов, индуцированных этопозидом, по сравнению с разрывами, вызванными радиацией [323]. А работы по секвенированию точек разрыва, вызванных обработкой этопозидом, показывают, что к нему в большей степени чувствительны участки в основаниях хроматиновых петель [207], [383]. На самом деле, факторы связаны друг с другом — так, способ внесения ДЦР влияет на то, какие локусы более подвержены разрыву, как много разрывов будет в клетке, по какому пути будет идти репарация разрыва, как долго он будет оставаться не репарированным.

Особенности формирования перестроек под действием этопозидов

В нашей работе анализ FISH-препаратов показал, что этопозид приводит к разрыву в гене *AML1*, и после обработки клеток этопозидом аллели *AML1* чаще локализуются за пределами хромосомной территории по сравнению с контрольными клетками. Причем доля аллелей, находящихся вне территории, для разорванных аллелей выше, чем для интактных аллелей. Отсюда можно предположить, что концы разрывов, вызванных обработкой этопозидом, приобретают повышенную подвижность в пространстве ядра, покидая свою хромосомную территорию.

Образование разрывов внутри генов *AML1* и *MLL* при воздействии этопозидом было показано и ранее [30], [36], [202], [204], [205]. Однако наши результаты подчеркивают, что этопозид приводит не просто к ДЦР внутри гена *AML1*, а к расхождению концов разрывов в пространстве ядра, которое можно детектировать с помощью микроскопии, — этого не наблюдается при действии γ -излучения на клетки Jurkat или при индукции экспрессии Cas9 в клетках культуры iAML-ETO. В норме разрывы быстро узнаются системами репарации и концы разрыва удерживаются вместе с помощью белковых комплексов или гомологии между одноцепочечными участками [232], [258]–[260], [384]. Вероятно, это связано с большей сложностью репарации разрывов, вызванных ингибированием ДНК-топоизомеразы II, по сравнению с разрывами иного происхождения. Разрывы, вызванные ингибированием ДНК-топоизомеразы, могут быть репарированы только после удаления с их концов ковалентно связанной субъединицы топоизомеразы [206]. Сложность репарации приводит к тому, что такие разрывы дольше существуют и диффундируют в ядре. Отсутствие детектируемых методом FISH разрывов после действия γ -излучения или после индукции нуклеазы Cas9 говорит о том, что они, напротив, сразу репарируются соответствующими системами клетки и не успевают диффундировать в пространстве ядра на расстояние, детектируемое световым микроскопом. Это подтверждается и другими работами, в которых разрывы индуцировали нуклеазами [319], [320].

Для изучения влияния обработки клеток этопозидом на контакты генов *AML1* и *MLL* с их партнерами по транслокациям был проведен 4С-анализ. Результаты показывают, что в профилях контактов генов *AML1* и *MLL* с остальными локусами генома после обработки этопозидом изменений не наблюдается. Можно предположить, что если бы концы разрывов в генах-партнерах сосредотачивались в некоторых ядерных компартментах, например, кластеризовались в одни и те же «фабрики репарации» или, например, в области ядрышка, как предполагалось ранее [324], то это отразилось бы на 4С-профилях. Скорее всего, такого направленного движения генов-партнеров друг к другу или в один компартмент не происходит.

Несмотря на данные литературы, которые связывают частоту образования транслокаций с пространственной близостью перестраивающихся локусов, 4С-анализ не выявил в клетках Jurkat такой пространственной близости для генов-партнеров по транслокациям, изучаемым в данной работе. С другой стороны, клетки Jurkat относятся к уже дифференцированным клеткам крови, в то время как лейкозы развиваются из недифференцированных бластных клеток, и пространственная организация их генома может несколько отличаться. И все же есть основания предполагать, что так как *AML1* и *MLL* остаются активно экспрессируемыми в клетках Jurkat, то и профили их контактов принципиально не отличаются от таковых в бластных клетках. Тогда можно заключить, что гены-партнеры по изучаемым транслокациям не имеют пространственной предрасположенности для перестроек. Эти результаты подтверждают данные [36], полученные на клетках линии KG1 (линия на основе макрофагов из костного мозга): авторы не выявили методом 3С повышенной частоты контактов между *MLL* и его партнерами по транслокациям *AF9* и *AF4*.

Изучение механизмов образования транслокаций с помощью клеточной модели

В нашей работе мы разработали способ изучения механизмов образования онкогенных перестроек, заключающийся в создании специальной клеточной культуры, где транслокации индуцируются путем запуска экспрессии интегрированного гена нуклеазы Cas9, нацеленной на перестраивающиеся локусы. Главное свойство такой модели заключается в возможности точно оценить частоту образования транслокаций. По разработанному протоколу можно создавать клеточные модели и для других транслокаций. Так, мы создали клеточную модель для изучения транслокации *IGH-MYC*, которая используется для изучения причин образования лимфомы Беркитта [378],[385].

Результаты экспериментов с полученной в данной работе клеточной культурой iAML-ETO подтвердили, что она является удобным инструментом для моделирования транслокации *AML1-ETO*. Образование транслокаций происходит с постоянной частотой в рамках одного эксперимента, а выявленные закономерности воспроизводятся при повторе

экспериментов. Это говорит о возможности применения модели для исследования механизмов формирования хромосомных транслокаций. Общее число транслокаций при индукции двух ДЦР — одна на 100–300 клеток — сопоставимо с результатами, полученными в других работах методами *in vivo* микроскопии [320] и высокопроизводительного секвенирования [386]. К недостаткам такого моделирования транслокаций, вызванных терапией топоизомеразными ядами, можно отнести неестественность способа внесения ДЦР в этой модели и отсутствие ковалентно сшитого с ДНК белкового аддукта в месте разрыва. Однако применение программируемой нуклеазы обусловлено необходимостью в простом и точном способе детекции получаемых перестроек с помощью ПЦР.

Изучение кинетики экспрессии Cas9 и кинетики накопления транслокаций показывает существенную задержку между индукцией разрывов и формированием транслокаций: Cas9 выходит на пик экспрессии спустя 24 часа, а резкое накопление транслокаций начинается спустя 36–48 часов после индукции системы доксициклином. Это является еще одним аргументом в пользу того, что транслокации происходят между «потерявшимися» концами разрывов, которые долгое время диффундируют в ядре, пока не встретят партнера для репарации. В некоторых работах также высказываются предположения, что транслокации происходят между концами разрывов, репарация которых характеризуется медленной кинетикой [387].

Разработанный во время выполнения диссертационной работы метод ENIT [372], [373] демонстрирует, что если внести разрывы в два локуса на разных хромосомах, то между ними произойдет перестройка хотя бы в некоторых клетках исследуемой популяции. Данный метод широко применяется в разнообразных экспериментах по редактированию генома, и какие бы пары локусов ни разрезались, практически всегда удавалось детектировать транслокации между ними. Это доказывает, что внесение ДЦР в перестраивающиеся локусы является достаточным фактором для образования транслокаций, а остальные факторы уже просто способствуют тем или иным транслокациям.

Исследование кинетики накопления транслокаций в культуре клеток iAML-ЕТО показало, что через несколько суток число транслокаций выходит на плато. Остановка образования новых транслокаций может быть объяснена тем, что индуцируемые нуклеазой разрывы в первую очередь репарируются правильным способом — когда сшиваются концы одного разрыва, то есть без формирования транслокации. При этом репарация происходит по механизму негомологичного соединения концов разрыва (NHEJ), который допускает вставку в место разрыва дополнительных нуклеотидов или же образование небольших делеций [374]. В результате на месте разрыва появляются инделы, которые в дальнейшем не позволяют РНК-гиду направлять Cas9 в этот локус, и возможность образования

транслокации в такой клетке исчезает. С учетом неизбежного подтекания промотора гена Cas9, можно предположить, что культура iAML-ЕТО со временем снижает свой потенциал к образованию транслокаций, и для экспериментов стоит использовать аликвоты, замороженные на ранних стадиях работы с культурой.

Меньшее число транслокаций в экспериментах, где клеточная культура находится в фазе быстрого роста, может объясняться большим вкладом гомологической рекомбинации в репарацию ДЦР. Для делящихся клеток характерно переключение путей репарации с NHEJ на направляемую гомологией рекомбинацию [252]. Считается, что именно первый путь может приводить к транслокации, тогда как второй путь защищен от таких ошибок [283]. В плотной культуре доля делящихся клеток снижается, и в среднем по популяции более предпочтительным становится путь репарации NHEJ.

Анализ влияния ингибиторов белков репарации на число транслокаций в культуре клеток iAML-ЕТО показал, что при ингибировании DNA-РКс число транслокаций возрастает. Этот результат совпадает с результатами, полученными при визуализации с помощью системы флуоресцентный репрессор-оператор (fluorescent repressor-operator systems, FROS, описание метода: [388]) процесса образования транслокаций, образуемых мегануклеазой I-SceI [320]. В этом контексте также необходимо упомянуть работу, в которой в клетках линии K562 с помощью нуклеазы ZFN индуцировали разрыв в гене *MLL* и анализировали методом FISH число разрывов, а число транслокаций — методом инвертированной ПЦР (описание метода: [10]). В отличие от нашей работы, разрывы индуцировали путем транзientной экспрессии плазмиды, кодирующей ZFN. В этих условиях авторы смогли детектировать разрывы *MLL*, причем при ингибировании DNA-РКс с помощью NU7026 число разрывов было выше. Так как второй разрыв не индуцировался, то при трансфекции ZFN транслокаций авторы не обнаружили, однако добавление NU7026 привело к обнаружению одной транслокации [326]. Наши результаты, вместе с результатами описанных работ, могут объясняться следующим образом. В процессе репарации ДЦР по механизму NHEJ, комплекс с участием DNA-РКс удерживает концы разрыва вместе перед тем, как они будут лигированы [258]–[260]. При ингибировании этого белка вырастает шанс «потерять» концы одного разрыва, а значит, повышается шанс в дальнейшем лигироваться концам разных разрывов с образованием транслокаций.

Репарацию разрывов в условиях ингибирования DNA-РКс можно сравнить с репарацией разрывов, вызванных ингибированием ДНК-топоизомераз: они не могут быть сразу репарированы из-за ковалентно связанного с ДНК фермента и тоже часто «теряют» концы. Аналогично может быть объяснено влияние мирина — ингибирование комплекса MRN, задействованного в начальных этапах репарации разрывов [279], также повышает шансы «потерять» концы разрыва. Таким образом, результаты по ингибированию

DNA-РКcs и MRN подтверждают, что сложность процессинга концов и их потеря повышает шансы на формирование хромосомных перестроек.

Некоторое снижение числа транслокаций *AML-ETO* в экспериментах с линией *iAML-ETO* и этопозидом может объясняться тем, что у разрывов, индуцированных нуклеазой, появляется больше других вариантов для репарации — разрывов в других локусах, индуцированных этопозидом [389]–[391].

Полученную клеточную модель *iAML-ETO* можно использовать для поиска лекарств и для тестирования веществ на предмет влияния на образование транслокаций. Так, был протестирован набор препаратов, применяемых в терапии опухолей совместно с топоизомеразными ядами. Было обнаружено, что один из препаратов — метотрексат — приводил к повышению числа транслокаций. Эти данные согласуются с описанными в литературе примерами развития АМЛ с транслокациями $t(8;21)$ и $t(3;21)$ у пациентов с ревматоидным артритом, которых лечили низкими дозами метотрексата [32]. Метотрексат — ингибитор дигидрофолатредуктазы, и возможные механизмы его влияния на образование транслокации пока сложно предположить, для этого необходимы дополнительные исследования.

Другим применением клеточной модели является ее использование в качестве тест-системы для различных соединений на предмет снижения частоты образования транслокаций. Широкий скрининг библиотек химических соединений на данной клеточной модели может привести к обнаружению веществ, которые способны снизить вероятность образования лейкозогенных перестроек вследствие химиотерапии опухолей.

Заключение

Химиотерапия опухолей с применением ингибиторов ДНК-топоизомераз II типа широко распространена, однако может приводить к развитию вторичных лейкозов, характеризующихся определенными хромосомными транслокациями. На основании данных литературы и результатов нашего исследования, можно выделить три основных фактора, определяющих набор этих транслокаций. Во-первых, транслокации определяются локусами, которые подвержены действию таких ядов вследствие активности на этих участках ДНК-топоизомераз. Во-вторых, при обработке клеток этопозидом характерна «потеря» концов разрывов, что приводит к их расхождению и диффузии по ядру. Сложность репарации таких разрывов способствует этому расхождению и нивелирует фактор пространственной близости перестраивающихся локусов. В-третьих, среди образующихся транслокаций только некоторые дают клеткам пролиферативное преимущество, и поэтому набор перестроек, наблюдаемых при лейкозах, ограничен.

Во время проведения исследования были разработаны методики, значимые для работ, связанных с редактированием генома — метод проверки эффективности инструментов редактирования генома Direct ENIT, методики получения клеточной линии для моделирования образования транслокации. Полученная клеточная культура может применяться как тест-система для анализа используемых в терапии препаратов на предмет риска возникновения лейкозогенных транслокаций. Также эта клеточная модель может стать тест-системой для обнаружения соединений, способных снизить вероятность развития вторичного лейкоза при терапии с применением ингибиторов ДНК-топоизомераз II. Наконец, такая модель может быть использована для изучения механизмов репарации двуцепочечных разрывов ДНК.

Выводы

1. Обработка клеток культуры Jurkat этопозидом приводит к возникновению двуцепочечных разрывов ДНК внутри гена *AML1* и расхождению образовавшихся концов ДНК в пространстве ядра.
2. Воздействие высоких доз γ -излучения не приводит к увеличению числа разошедшихся концов разрывов гена *AML1*.
3. Под действием этопозида аллели *AML1* чаще, чем в необработанных клетках, локализуются вне хромосомной территории.
4. Ни в контрольных клетках, ни в клетках, обработанных этопозидом, не выявлено повышенного числа контактов генов *AML1* и *MLL* с их генами-партнерами по транслокациям по сравнению с другими генами.
5. Создана клеточная культура iAML-ETO, содержащая в геноме индуцируемую доксициклином систему CRISPR/Cas, которая формирует транслокацию *AML1-ETO*, количественно определяемую методом ПЦР.
6. С помощью полученной клеточной культуры iAML-ETO показано, что присутствие метотрексата или ингибитора DNA-PKcs NU7026 приводит к повышению частоты формирования транслокаций *AML1-ETO*.

Список литературы

- 1 Schürch CM, Riether C, Ochsenbein AF. Dendritic cell-based immunotherapy for myeloid leukemias. *Front. Immunol.* 2013; **4**. doi:10.3389/fimmu.2013.00496.
- 2 Becker MW, Jordan CT. Leukemia stem cells in 2010: Current understanding and future directions. *Blood Rev* 2011; **25**: 75–81.
- 3 Wang JH. Mechanisms and impacts of chromosomal translocations in cancers. *Front. Med. China.* 2012; **6**. doi:10.1007/s11684-012-0215-5.
- 4 Andersson AK, Ma J, Wang J, Chen X, Gedman AL, Dang J *et al.* The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nat Genet* 2015; **47**. doi:10.1038/ng.3230.
- 5 Greaves M. When one mutation is all it takes. *Cancer Cell.* 2015; **27**. doi:10.1016/j.ccell.2015.03.016.
- 6 Balgobind B V., Raimondi SC, Harbott J, Zimmermann M, Alonzo TA, Auvrignon A *et al.* Novel prognostic subgroups in childhood 11q23/MLL-rearranged acute myeloid leukemia: Results of an international retrospective study. *Blood* 2009; **114**. doi:10.1182/blood-2009-04-215152.
- 7 Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T *et al.* Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood.* 2017; **129**. doi:10.1182/blood-2016-08-733196.
- 8 Meyer C, Burmeister T, Gröger D, Tsaour G, Fechina L, Renneville A *et al.* The MLL recombinome of acute leukemias in 2017. *Leukemia* 2018; **32**. doi:10.1038/leu.2017.213.
- 9 Khoury JD, Solary E, Abal O, Akkari Y, Alaggio R, Apperley JF *et al.* The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia* 2022; **36**: 1703–1719.
- 10 Lomov N, Zerkalenkova E, Lebedeva S, Viushkov V, Rubtsov MA. Cytogenetic and molecular genetic methods for chromosomal translocations detection with reference to the *KMT2A/MLL* gene. *Crit Rev Clin Lab Sci* 2021; **58**: 180–206.
- 11 Germini D, Tsfasman T, Klibi M, El-Amine R, Pichugin A, Iarovaia O V *et al.* HIV Tat induces a prolonged MYC relocalization next to IGH in circulating B-cells. *Leukemia* 2017; : 1–8.
- 12 Grande BM, Gerhard DS, Jiang A, Griner NB, Abramson JS, Alexander TB *et al.* Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. *Blood* 2019; **133**. doi:10.1182/blood-2018-09-871418.
- 13 Lieber MR. Mechanisms of human lymphoid chromosomal translocations. *Nat. Rev. Cancer.* 2016; **16**: 387–398.
- 14 Roos WP, Kaina B. DNA damage-induced cell death by apoptosis. *Trends Mol. Med.* 2006. doi:10.1016/j.molmed.2006.07.007.
- 15 Bhatia S. Therapy-related myelodysplasia and acute myeloid leukemia. *Semin Oncol* 2013; **40**: 666–675.

- 16 Schoch C, Kern W, Schnittger S, Hiddemann W, Haferlach T. Karyotype is an independent prognostic parameter in therapy-related acute myeloid leukemia (t-AML): An analysis of 93 patients with t-AML in comparison to 1091 patients with de novo AML. *Leukemia* 2004; **18**. doi:10.1038/sj.leu.2403187.
- 17 Borthakur G, Estey EE. Therapy-related acute myelogenous leukemia and myelodysplastic syndrome. *Curr Oncol Rep* 2007; **9**: 373–377.
- 18 Godley LA, Larson RA. Therapy-Related Myeloid Leukemia. *Semin Oncol* 2008; **35**: 418–429.
- 19 Østgård LSG, Medeiros BC, Sengeløv H, Nørgaard M, Andersen MK, Dufva I *et al*. Epidemiology and clinical significance of secondary and therapy-related acute myeloid leukemia: A national population-based cohort study. *J Clin Oncol* 2015; **33**: 3641–3649.
- 20 Tiruneh T, Enawgaw B, Shiferaw E. Genetic Pathway in the Pathogenesis of Therapy-Related Myeloid Neoplasms: A Literature Review. *Oncol Ther* 2020; **8**: 45–57.
- 21 Pedersen-Bjergaard J, Andersen MK, Christiansen DH, Nerlov C. Genetic pathways in therapy-related myelodysplasia and acute myeloid leukemia. *Blood*. 2002; **99**. doi:10.1182/blood.V99.6.1909.
- 22 Kayser S, Döhner K, Krauter J, Köhne C-H, Horst HA, Held G *et al*. The impact of therapy-related acute myeloid leukemia (AML) on outcome in 2853 adult patients with newly diagnosed AML. *Blood* 2011; **117**: 2137–45.
- 23 McNerney ME, Godley LA, Le Beau MM. Therapy-related myeloid neoplasms: When genetics and environment collide. *Nat Rev Cancer* 2017; **17**: 513–527.
- 24 Ratain MJ, Rowley JD. Therapy-related acute myeloid leukemia secondary to inhibitors of topoisomerase II: From the bedside to the target genes. *Ann Oncol* 1992; **3**: 107–111.
- 25 Mauritzson N, Albin M, Rylander L, Billström R, Ahlgren T, Mikoczy Z *et al*. Pooled analysis of clinical and cytogenetic features in treatment-related and de novo adult acute myeloid leukemia and myelodysplastic syndromes based on a consecutive series of 761 patients analyzed 1976–1993 and on 5098 unselected cases reported in the . *Leukemia* 2002; **16**: 2366–2378.
- 26 Rowley JD, Olney HJ. International Workshop on the relationship of prior therapy to balanced chromosome aberrations in therapy-related myelodysplastic syndromes and acute leukemia: Overview report. *Genes Chromosom Cancer* 2002; **33**: 331–345.
- 27 Pedersen-Bjergaard J, Andersen MK, Andersen MT, Christiansen DH. Genetics of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia* 2008; **22**: 240–248.
- 28 Imamura T, Taga T, Takagi M, Kawasaki H, Koh K, Taki T *et al*. Nationwide survey of therapy-related leukemia in childhood in Japan. *Int J Hematol* 2018; **108**: 91–97.
- 29 Gustafson SA, Lin P, Chen SS, Chen L, Abruzzo L V, Luthra R *et al*. Therapy-related acute myeloid leukemia with t(8;21) (q22;q22) shares many features with de novo acute myeloid leukemia with t(8;21)(q22;q22) but does not have a favorable outcome. *Am J Clin Pathol* 2009; **131**: 647–655.
- 30 Smith KA, Cowell IG, Zhang Y, Sondka Z, Austin CA. The role of topoisomerase II beta on breakage and proximity of RUNX1 to partner alleles RUNX1T1 and EVI1. *Genes, Chromosom Cancer* 2014; **53**: 117–128.

- 31 Sood R, Kamikubo Y, Liu P. Role of RUNX1 in hematological malignancies. *Blood*. 2017; **129**. doi:10.1182/blood-2016-10-687830.
- 32 Tanaka K, Oshikawa G, Akiyama H, Ishida S, Nagao T, Yamamoto M *et al*. Acute myeloid leukemia with t(3;21)(q26.2;q22) developing following low-dose methotrexate therapy for rheumatoid arthritis and expressing two AML1/MDS1/EVI1 fusion proteins: A case report. *Oncol Lett* 2017; **14**: 97–102.
- 33 Sato Y, Izumi T, Kanamori H, Davis EM, Miura Y, Larson RA *et al*. t(1;3)(p36;p21) is a recurring therapy-related translocation. *Genes Chromosom Cancer* 2002; **34**. doi:10.1002/gcc.10055.
- 34 Duhoux FP, Ameye G, Montano-Almendras CP, Bahloula K, Mozziconacci MJ, Laibe S *et al*. PRDM16 (1p36) translocations define a distinct entity of myeloid malignancies with poor prognosis but may also occur in lymphoid malignancies. *Br J Haematol* 2012; **156**: 76–88.
- 35 Andersen MK, Christiansen DH, Jensen BA, Ernst P, Hauge G, Pedersen-Bjergaard J. Therapy-related acute lymphoblastic leukaemia with MLL rearrangements following DNA topoisomerase II inhibitors, an increasing problem: Report on two new cases and review of the literature since 1992. *Br J Haematol* 2001; **114**: 539–543.
- 36 Cowell IG, Austin CA. Mechanism of Generation of Therapy Related Leukemia in Response to Anti-Topoisomerase II Agents. ... *J Environ Res Public* ... 2012; **9**: 2075–91.
- 37 Moorman a V, Hagemeijer A, Charrin C, Rieder H, Secker-walker LM. Clinical Profile of 53 Patients. *Leukemia* 1998; **12**: 805–810.
- 38 Meyer C, Burmeister T, Strehl S, Schneider B, Hubert D, Zach O *et al*. Spliced MLL fusions: a novel mechanism to generate functional chimeric MLL-MLLT1 transcripts in t(11;19)(q23;p13.3) leukemia. *Leuk Off J Leuk Soc Am Leuk Res Fund, UK* 2007; **21**: 588–590.
- 39 Lavau C, Du C, Thirman M, Zeleznik-Le N. Chromatin-related properties of CBP fused to MLL generate a myelodysplastic-like syndrome that evolves into myeloid leukemia. *EMBO J* 2000; **19**. doi:10.1093/emboj/19.17.4655.
- 40 Yoo BJ, Nam MH, Sung HJ, Lim CS, Lee CK, Cho YJ *et al*. A case of therapy-related acute lymphoblastic leukemia with t(11;19) (q23;p13.3) and MLL/MLLT1 gene rearrangement. *Korean J Lab Med* 2011; **31**. doi:10.3343/kjlm.2011.31.1.13.
- 41 Xie W, Tang G, Wang E, Kim Y, Cloe A, Shen Q *et al*. t(11;16)(q23;p13)/KMT2A-CREBBP in hematologic malignancies: presumptive evidence of myelodysplasia or therapy-related neoplasm? *Ann Hematol* 2020; **99**: 487–500.
- 42 Miyoshi H, Shimizu K, Kozu T, Maseki N, Kaneko Y, Ohki M. (8;21) breakpoints on chromosome 21 in acute myeloid leukemia are clustered within a limited region of a single gene, AML1. *Proc Natl Acad Sci U S A* 1991; **88**. doi:10.1073/pnas.88.23.10431.
- 43 Roulston D, Espinosa R, Nucifora G, Larson RA, Le Beau MM, Rowley JD. CBFA2(AML1) Translocations With Novel Partner Chromosomes in Myeloid Leukemias: Association With Prior Therapy. *Blood* 1998; **92**: 2879–2885.

- 44 Pedersen-Bjergaard J, Rowley JD. The balanced and the unbalanced chromosome aberrations of acute myeloid leukemia may develop in different ways and may contribute differently to malignant transformation. *Blood*. 1994; **83**. doi:10.1182/blood.v83.10.2780.2780.
- 45 De Braekeleer E, Ferec C, De Braekeleer M. RUNX1 translocations in malignant hemopathies. *Anticancer Res* 2009; **29**: 1031–1038.
- 46 Lam K, Zhang D-E. RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis. *Front Biosci (Landmark Ed)* 2012; **17**: 1120–39.
- 47 Hayashi Y, Harada Y, Harada H. Myeloid neoplasms and clonal hematopoiesis from the RUNX1 perspective. *Leukemia* 2022; **36**: 1203–1214.
- 48 Ichikawa M, Yoshimi A, Nakagawa M, Nishimoto N, Watanabe-Okochi N, Kurokawa M. A role for RUNX1 in hematopoiesis and myeloid leukemia. *Int. J. Hematol.* 2013; **97**: 726–734.
- 49 Chen MJ, Yokomizo T, Zeigler BM, Dzierzak E, Speck NA. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* 2009; **457**. doi:10.1038/nature07619.
- 50 Ottersbach K. Endothelial-to-haematopoietic transition: An update on the process of making blood. *Biochem. Soc. Trans.* 2019; **47**. doi:10.1042/BST20180320.
- 51 Wang Q, Stacy T, Binder M, Marin-Padilla M, Sharpe AH, Speck NA. Disruption of the *Cbfa2* gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci* 1996; **93**: 3444–3449.
- 52 Growney JD, Shigematsu H, Li Z, Lee BH, Adelsperger J, Rowan R *et al.* Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood* 2005; **106**: 494–504.
- 53 Ichikawa M, Goyama S, Asai T, Kawazu M, Nakagawa M, Takeshita M *et al.* AML1/Runx1 Negatively Regulates Quiescent Hematopoietic Stem Cells in Adult Hematopoiesis. *J Immunol* 2008; **180**: 4402–4408.
- 54 Tanaka Y, Joshi A, Wilson NK, Kinston S, Nishikawa S, Göttgens B. The transcriptional programme controlled by Runx1 during early embryonic blood development. *Dev Biol* 2012; **366**: 404–419.
- 55 Yamagata T, Maki K, Mitani K. Runx1/AML1 in Normal and Abnormal Hematopoiesis. *Int J Hematol* 2005; **82**: 1–8.
- 56 Elagib KE, Racke FK, Mogass M, Khetawat R, Delehanty LL, Goldfarb AN. RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation. *Blood* 2003; **101**. doi:10.1182/blood-2002-09-2708.
- 57 Kim WY, Sieweke M, Ogawa E, Wee HJ, Englmeier U, Graf T *et al.* Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J* 1999; **18**. doi:10.1093/emboj/18.6.1609.
- 58 Zhang DE, Hetherington CJ, Meyers S, Rhoades KL, Larson CJ, Chen HM *et al.* CCAAT enhancer-binding protein (C/EBP) and AML1 (CBF alpha2) synergistically activate the macrophage colony-stimulating factor receptor promoter. *Mol Cell Biol* 1996; **16**: 1231–1240.

- 59 Kitabayashi I, Yokoyama A, Shimizu K, Ohki M. Interaction and functional cooperation of the leukemia-associated factors AML1 and p300 in myeloid cell differentiation. *EMBO J* 1998; **17**: 2994–3004.
- 60 Yagi R, Chen LF, Shigesada K, Murakami Y, Ito Y. A WW domain-containing Yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO J* 1999; **18**: 2551–2562.
- 61 Segrelles C, Paramio JM, Lorz C. The transcriptional co-activator YAP: A new player in head and neck cancer. *Oral Oncol.* 2018; **86**. doi:10.1016/j.oraloncology.2018.08.020.
- 62 Riddell A, McBride M, Braun T, Nicklin SA, Cameron E, Loughrey CM *et al.* RUNX1: An emerging therapeutic target for cardiovascular disease. *Cardiovasc Res* 2020; **116**: 1410–1423.
- 63 Lutterbach B, Westendorf JJ, Linggi B, Isaac S, Seto E, Hiebert SW. A mechanism of repression by acute myeloid leukemia-1, the target of multiple chromosomal translocations in acute leukemia. *J Biol Chem* 2000; **275**: 651–656.
- 64 Imai Y, Kurokawa M, Yamaguchi Y, Izutsu K, Nitta E, Mitani K *et al.* The Corepressor mSin3A Regulates Phosphorylation-Induced Activation, Intranuclear Location, and Stability of AML1. *Mol Cell Biol* 2004; **24**. doi:10.1128/mcb.24.3.1033-1043.2004.
- 65 Seo W, Tanaka H, Miyamoto C, Levanon D, Groner Y, Taniuchi I. Roles of VWRPY motif-mediated gene repression by Runx proteins during T-cell development. *Immunol Cell Biol* 2012; **90**. doi:10.1038/icb.2012.6.
- 66 Alkadi H, McKellar D, Zhen T, Karpova T, Garrett LJ, Gao Y *et al.* The VWRPY Domain Is Essential for RUNX1 Function in Hematopoietic Progenitor Cell Maturation and Megakaryocyte Differentiation. *Blood* 2018; **132**. doi:10.1182/blood-2018-99-113400.
- 67 Huret JL, Ahmad M, Arsaban M, Bernheim A, Cigna J, Desangles F *et al.* Atlas of genetics and cytogenetics in oncology and haematology in 2013. *Nucleic Acids Res* 2013; **41**.
- 68 Peterson LF, Zhang DE. The 8;21 translocation in leukemogenesis. *Oncogene* 2004; **23**: 4255–4262.
- 69 Davis JN, McGhee L, Meyers S. The ETO (MTG8) gene family. *Gene* 2003; **303**: 1–10.
- 70 Hiebert SW, Reed-Inderbitzin EF, Amann J, Irvin B, Durst K, Linggi B. The t(8;21) fusion protein contacts co-repressors and histone deacetylases to repress the transcription of the p14ARF tumor suppressor. *Blood Cells, Mol Dis* 2003; **30**. doi:10.1016/S1079-9796(03)00021-4.
- 71 Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A *et al.* Tissue-based map of the human proteome. *Science (80-)* 2015; **347**. doi:10.1126/science.1260419.
- 72 Perry C, Eldor A, Soreq H. Runx1/AML1 in leukemia: disrupted association with diverse protein partners. *Leuk Res* 2002; **26**: 221–228.
- 73 Gelmetti V, Zhang J, Fanelli M, Minucci S, Pelicci PG, Lazar MA. Aberrant Recruitment of the Nuclear Receptor Corepressor-Histone Deacetylase Complex by the Acute Myeloid Leukemia Fusion Partner ETO. *Mol Cell Biol* 1998; **18**: 7185–7191.

- 74 Lutterbach B, Westendorf JJ, Linggi B, Patten A, Moniwa M, Davie JR *et al.* ETO, a target of t(8;21) in acute leukemia, interacts with the N-CoR and mSin3 corepressors. *Mol Cell Biol* 1998; **18**: 7176–84.
- 75 Wang L, Gural A, Sun XJ, Zhao X, Perna F, Huang G *et al.* The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science (80-)* 2011; **333**. doi:10.1126/science.1201662.
- 76 Rejeski K, Duque-Afonso J, Lübbert M. AML1/ETO and its function as a regulator of gene transcription via epigenetic mechanisms. *Oncogene*. 2021; **40**: 5665–5676.
- 77 Ben-Ami O, Friedman D, Leshkowitz D, Goldenberg D, Orlovsky K, Pencovich N *et al.* Addiction of t(8;21) and inv(16) acute myeloid leukemia to native RUNX1. *Cell Rep* 2013; **4**: 1131–43.
- 78 Ptasinska A, Assi SA, Mannari D, James SR, Williamson D, Dunne J *et al.* Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding. *Leukemia* 2012; **26**: 1829–1841.
- 79 Yuan Y, Zhou L, Miyamoto T, Iwasaki H, Harakawa N, Hetherington CJ *et al.* AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations. *Proc Natl Acad Sci* 2001; **98**: 10398–10403.
- 80 Higuchi M, O’Brien D, Kumaravelu P, Lenny N, Yeoh EJ, Downing JR. Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia. *Cancer Cell* 2002; **1**. doi:10.1016/S1535-6108(02)00016-8.
- 81 Koza T, Fukuyama T, Yamami T, Akagi K, Kaneko Y. MYND-less splice variants of AML1-MTG8 (RUNX1=CBFA2T1) are expressed in leukemia with t(8;21). *Genes Chromosom Cancer* 2005; **43**. doi:10.1002/gcc.20165.
- 82 Yan M, Kanbe E, Peterson LF, Boyapati A, Miao Y, Wang Y *et al.* A previously unidentified alternatively spliced isoform of t(8;21) transcript promotes leukemogenesis. *Nat Med* 2006; **12**: 945–949.
- 83 Slany RK. The molecular biology of mixed lineage leukemia. *Haematologica* 2009; **94**: 984–993.
- 84 Mirro J, Zipf TF, Pui CH, Kitchingman G, Williams D, Melvin S *et al.* Acute mixed lineage leukemia: Clinicopathologic correlations and prognostic significance. *Blood* 1985; **66**. doi:10.1182/blood.v66.5.1115.bloodjournal6651115.
- 85 Mirro J, Kitchingman GR, Williams DL, Murphy SB, Zipf TF, Stass SA. Mixed lineage leukemia: the implications for hematopoietic differentiation. *Blood*. 1986; **68**. doi:10.1182/blood.v68.2.597.597.
- 86 Ziemin-Van Der Poel S, McCabe NR, Gill HJ, Espinosa R, Patel Y, Harden A *et al.* Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc Natl Acad Sci U S A* 1991; **88**. doi:10.1073/pnas.88.23.10735.
- 87 Djabali M, Selleri L, Parry P, Bower M, Young BD, Evans GA. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. *Nat Genet* 1992; **2**. doi:10.1038/ng1092-113.

- 88 Marschalek R. Mechanisms of leukemogenesis by MLL fusion proteins. *Br J Haematol* 2011; **152**: 141–154.
- 89 Super HJG, McCabe NR, Thirman MJ, Larson RA, Le Beau MM, Pedersen- Bjergaard J *et al.* Rearrangements of the MLL gene in therapy-related acute myeloid leukemia in patients previously treated with agents targeting DNA-topoisomerase II. *Blood* 1993; **82**. doi:10.1182/blood.v82.12.3705.bloodjournal82123705.
- 90 Felix CA. Secondary leukemias induced by topoisomerase-targeted drugs. *Biochim. Biophys. Acta - Gene Struct. Expr.* 1998; **1400**. doi:10.1016/S0167-4781(98)00139-0.
- 91 Rao RC, Dou Y. Hijacked in cancer: The KMT2 (MLL) family of methyltransferases. *Nat Rev Cancer* 2015; **15**: 334–346.
- 92 Borkhardt A, Wuchter C, Viehmann S, Pils S, Teigler-Schlegel A, Stanulla M *et al.* Infant acute lymphoblastic leukemia - Combined cytogenetic, immunophenotypic and molecular analysis of 77 cases. *Leukemia* 2002; **16**. doi:10.1038/sj.leu.2402595.
- 93 Pieters R, Schrappe M, De Lorenzo P, Hann I, De Rossi G, Felice M *et al.* A treatment protocol for infants younger than 1 year with acute lymphoblastic leukaemia (Interfant-99): an observational study and a multicentre randomised trial. *Lancet* 2007; **370**. doi:10.1016/S0140-6736(07)61126-X.
- 94 Creutzig U, Zimmermann M, Dworzak M, Bourquin J-P, von Neuhoff C, Sander A *et al.* Additional Prognostic Impact of Specific Sites of Extramedullary Leukemia in Pediatric AML with MLL-Rearrangements but Not in Core-Binding Factor (CBF) AML. *Blood* 2012; **120**. doi:10.1182/blood.v120.21.2621.2621.
- 95 Marschalek R. Systematic Classification of Mixed-Lineage Leukemia Fusion Partners Predicts Additional Cancer Pathways. *Ann Lab Med* 2016; **36**: 85–100.
- 96 Liu H, Cheng EHY, Hsieh JJD. MLL fusions: Pathways to leukemia. *Cancer Biol Ther* 2009; **8**: 1204–1211.
- 97 Guenther MG, Jenner RG, Chevalier B, Nakamura T, Croce CM, Canaani E *et al.* Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci* 2005; **102**: 8603–8608.
- 98 Ayton PM, Cleary ML. Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. *Genes Dev* 2003; **17**. doi:10.1101/gad.1111603.
- 99 Wang P, Lin C, Smith ER, Guo H, Sanderson BW, Wu M *et al.* Global Analysis of H3K4 Methylation Defines MLL Family Member Targets and Points to a Role for MLL1-Mediated H3K4 Methylation in the Regulation of Transcriptional Initiation by RNA Polymerase II. *Mol Cell Biol* 2009; **29**. doi:10.1128/mcb.00924-09.
- 100 Artinger EL, Mishra BP, Zaffuto KM, Li BE, Chung EKY, Moore AW *et al.* An MLL-dependent network sustains hematopoiesis. *Proc Natl Acad Sci U S A* 2013; **110**. doi:10.1073/pnas.1301278110.
- 101 Yu BD, Hess JL, Horning SE, Brown GAJ, Korsmeyer SJ. Altered Hox expression and segmental identity in Mll-mutant mice. *Nature* 1995; **378**. doi:10.1038/378505a0.
- 102 Yagi H, Deguchi K, Aono A, Tani Y, Kishimoto T, Komori T. Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. *Blood* 1998; **92**. doi:10.1182/blood.v92.1.108.413k11_108_117.

- 103 Ernst P, Fisher JK, Avery W, Wade S, Foy D, Korsmeyer SJ. Definitive hematopoiesis requires the mixed-lineage leukemia gene. *Dev Cell* 2004; **6**: 437–443.
- 104 McMahon KA, Hiew SYL, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ *et al.* Mll Has a Critical Role in Fetal and Adult Hematopoietic Stem Cell Self-Renewal. *Cell Stem Cell* 2007; **1**: 338–345.
- 105 García-Alai MM, Allen MD, Joerger AC, Bycroft M. The structure of the FYR domain of transforming growth factor beta regulator 1. *Protein Sci* 2010; **19**. doi:10.1002/pro.404.
- 106 Birke M, Schreiner S, García-Cuellar M-P, Mahr K, Titgemeyer F, Slany RK. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res* 2002; **30**: 958–65.
- 107 Allen MD, Grummitt CG, Hilcenko C, Min SY, Tonkin LM, Johnson CM *et al.* Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. *EMBO J* 2006; **25**. doi:10.1038/sj.emboj.7601340.
- 108 Risner LE, Kuntimaddi A, Lokken AA, Achille NJ, Birch NW, Schoenfelt K *et al.* Functional specificity of CpG DNA-binding CXXC domains in mixed lineage leukemia. *J Biol Chem* 2013; **288**. doi:10.1074/jbc.M113.474858.
- 109 Muntean AG, Tan J, Sitwala K, Huang Y, Bronstein J, Connelly JA *et al.* The PAF Complex Synergizes with MLL Fusion Proteins at HOX Loci to Promote Leukemogenesis. *Cancer Cell* 2010; **17**. doi:10.1016/j.ccr.2010.04.012.
- 110 Wang Z, Song J, Milne TA, Wang GG, Li H, Allis CD *et al.* Pro Isomerization in MLL1 PHD3-Bromo Cassette Connects H3K4me Readout to CyP33 and HDAC-Mediated Repression. *Cell* 2010; **141**: 1183–1194.
- 111 Chang PY, Hom RA, Musselman CA, Zhu L, Kuo A, Gozani O *et al.* Binding of the MLL PHD3 Finger to Histone H3K4me3 Is Required for MLL-Dependent Gene Transcription. *J Mol Biol* 2010; **400**. doi:10.1016/j.jmb.2010.05.005.
- 112 Roloff TC, Ropers HH, Nuber UA. Comparative study of methyl-CpG-binding domain proteins. *BMC Genomics* 2003; **4**. doi:10.1186/1471-2164-4-1.
- 113 Yokoyama A, Cleary ML. Menin Critically Links MLL Proteins with LEDGF on Cancer-Associated Target Genes. *Cancer Cell* 2008; **14**. doi:10.1016/j.ccr.2008.05.003.
- 114 Hughes CM, Rozenblatt-Rosen O, Milne TA, Copeland TD, Levine SS, Lee JC *et al.* Menin associates with a trithorax family histone methyltransferase complex and with the Hoxc8 locus. *Mol Cell* 2004; **13**. doi:10.1016/S1097-2765(04)00081-4.
- 115 Chan AKN, Chen CW. Rewiring the epigenetic networks in MLL-rearranged leukemias: Epigenetic dysregulation and pharmacological interventions. *Front Cell Dev Biol* 2019; **7**: 1–15.
- 116 Slany RK. The molecular mechanics of mixed lineage leukemia. *Oncogene*. 2016; **35**. doi:10.1038/onc.2016.30.
- 117 del Rizzo PA, Trievel RC. Substrate and product specificities of SET domain methyltransferases. *Epigenetics*. 2011; **6**. doi:10.4161/epi.6.9.16069.

- 118 Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL *et al.* Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* 2006; **13**. doi:10.1038/nsmb1128.
- 119 Patel A, Dharmarajan V, Vought VE, Cosgrove MS. On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J Biol Chem* 2009; **284**. doi:10.1074/jbc.M109.014498.
- 120 Li Y, Han J, Zhang Y, Cao F, Liu Z, Li S *et al.* Structural basis for activity regulation of MLL family methyltransferases. *Nature* 2016; **530**. doi:10.1038/nature16952.
- 121 Lloyd NR, Wuttke DS. Cyp33 binds AU-rich RNA motifs via an extended interface that competitively disrupts the gene repressive Cyp33-MLL1 interaction in vitro. *PLoS One* 2021; **16**: 1–18.
- 122 Fair K, Anderson M, Bulanova E, Mi H, Tropschug M, Diaz MO. Protein Interactions of the MLL PHD Fingers Modulate MLL Target Gene Regulation in Human Cells. *Mol Cell Biol* 2001; **21**. doi:10.1128/mcb.21.10.3589-3597.2001.
- 123 Xia ZB, Anderson M, Diaz MO, Zeleznik-Le NJ. MLL repression domain interacts with histone deacetylases, the polycomb group proteins HPC2 and BMI-1, and the corepressor C-terminal-binding protein. *Proc Natl Acad Sci U S A* 2003; **100**. doi:10.1073/pnas.1436338100.
- 124 Grow EJ, Wysocka J. Flipping MLL1's switch one proline at a time. *Cell*. 2010; **141**. doi:10.1016/j.cell.2010.06.013.
- 125 Chen J, Santillan DA, Koonce M, Wei W, Luo R, Thirman MJ *et al.* Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. *Cancer Res* 2008; **68**. doi:10.1158/0008-5472.CAN-07-6514.
- 126 Cowell IG, Sondka Z, Smith K, Lee KC, Manville CM, Sidorczuk-Lesthuruge M *et al.* Model for MLL translocations in therapy-related leukemia involving topoisomerase II - mediated DNA strand breaks and gene proximity. *Proc Natl Acad Sci* 2012; **109**: 8989–8994.
- 127 Wright RL, Vaughan ATM. A systematic description of MLL fusion gene formation. *Crit Rev Oncol Hematol* 2014; **91**: 283–291.
- 128 Ross ME, Zhou X, Song G, Shurtleff SA, Girtman K, Williams WK *et al.* Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. *Blood* 2003; **102**. doi:10.1182/blood-2003-01-0338.
- 129 Ferrando AA, Armstrong SA, Neuberg DS, Sallan SE, Silverman LB, Korsmeyer SJ *et al.* Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: Dominance of HOX dysregulation. *Blood* 2003; **102**. doi:10.1182/blood-2002-10-3221.
- 130 Lin C, Smith ER, Takahashi H, Lai KC, Martin-Brown S, Florens L *et al.* AFF4, a Component of the ELL/P-TEFb Elongation Complex and a Shared Subunit of MLL Chimeras, Can Link Transcription Elongation to Leukemia. *Mol Cell* 2010; **37**. doi:10.1016/j.molcel.2010.01.026.

- 131 Yokoyama A, Lin M, Naresh A, Kitabayashi I, Cleary ML. A Higher-Order Complex Containing AF4 and ENL Family Proteins with P-TEFb Facilitates Oncogenic and Physiologic MLL-Dependent Transcription. *Cancer Cell* 2010; **17**. doi:10.1016/j.ccr.2009.12.040.
- 132 Bitoun E, Oliver PL, Davies KE. The mixed-lineage leukemia fusion partner AF4 stimulates RNA polymerase II transcriptional elongation and mediates coordinated chromatin remodeling. *Hum Mol Genet* 2007; **16**: 92–106.
- 133 Mueller D, García-Cuellar MP, Bach C, Buhl S, Maethner E, Slany RK. Misguided transcriptional elongation causes mixed lineage leukemia. *PLoS Biol* 2009; **7**. doi:10.1371/journal.pbio.1000249.
- 134 Nguyen AT, Zhang Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev* 2011; **25**: 1345–1358.
- 135 Godfrey L, Crump NT, Thorne R, Lau IJ, Repapi E, Dimou D *et al*. DOT1L inhibition reveals a distinct subset of enhancers dependent on H3K79 methylation. *Nat Commun* 2019; **10**. doi:10.1038/s41467-019-10844-3.
- 136 Prange KHM, Mandoli A, Kuznetsova T, Wang SY, Sotoca AM, Marneth AE *et al*. MLL-AF9 and MLL-AF4 oncofusion proteins bind a distinct enhancer repertoire and target the RUNX1 program in 11q23 acute myeloid leukemia. *Oncogene* 2017; **36**: 3346–3356.
- 137 Marschalek R. The reciprocal world of MLL fusions: A personal view. *Biochim Biophys Acta - Gene Regul Mech* 2020; **1863**. doi:10.1016/j.bbagr.2020.194547.
- 138 Gaussmann A, Wenger T, Eberle I, Bursen A, Bracharz S, Herr I *et al*. Combined effects of the two reciprocal t(4;11) fusion proteins MLL·AF4 and AF4·MLL confer resistance to apoptosis, cell cycling capacity and growth transformation. *Oncogene* 2007; **26**. doi:10.1038/sj.onc.1210125.
- 139 Eberle I, Pless B, Braun M, Dingermann T, Marschalek R. Transcriptional properties of human NANOG1 and NANOG2 in acute leukemic cells. *Nucleic Acids Res* 2010; **38**. doi:10.1093/nar/gkq307.
- 140 Ferrando AA, López-Otín C. Clonal evolution in leukemia. *Nat Med* 2017; **23**: 1135–1145.
- 141 Nowell PC. The clonal evolution of tumor cell populations. *Science (80-)* 1976; **194**. doi:10.1126/science.959840.
- 142 Greaves M, Maley CC. Clonal evolution in cancer. *Nature*. 2012; **481**. doi:10.1038/nature10762.
- 143 DeVita VT, Chu E. A history of cancer chemotherapy. *Cancer Res* 2008; **68**: 8643–8653.
- 144 Ledford H. The poisonous history of chemotherapy. *Nature* 2020; **585**. doi:10.1038/d41586-020-02605-w.
- 145 Goodman LS, Wintrobe MM, Dameshek W, Goodman MJ, Gilman A, McLennan MT. Nitrogen mustard therapy: Use of Methyl-Bis(Beta-Chloroethyl)amine Hydrochloride and Трис(Beta-Chloroethyl)amine Hydrochloride for Hodgkin's Disease, Lymphosarcoma, Leukemia and Certain Allied and Miscellaneous Disorders. *J Am Med Assoc* 1946; **132**. doi:10.1001/jama.1946.02870380008004.

- 146 Farber S. Some observations on the effect of folic acid antagonists on acute. *Blood* 1949; **4**: 160–167.
- 147 Hande KR. Etoposide: Four Decades of Development of a Topoisomerase II Inhibitor. *Eur J Cancer* 1998; **34**: 1514–1521.
- 148 King LS, Sullivan M. The similarity of the effect of podophyllin and colchicine and their use in the treatment of condylomata acuminata. *Science (80-)* 1946; **104**. doi:10.1126/science.104.2698.244.
- 149 Greenspan EM. Effect of alpha-peltatin, beta-peltatin, and podophyllotoxin on lymphomas and other transplanted tumors. *J Natl Cancer Inst* 1950; **10**. doi:10.1093/jnci/10.6.1295.
- 150 Baldwin EL, Osheroff N. Etoposide, topoisomerase II and cancer. *Curr Med Chem - Anti-Cancer Agents* 2005; **5**. doi:10.2174/1568011054222364.
- 151 Burden DA, Osheroff N. Mechanism of action of eukaryotic topoisomerase II and drugs targeted to the enzyme. *Biochim. Biophys. Acta - Gene Struct. Expr.* 1998; **1400**. doi:10.1016/S0167-4781(98)00132-8.
- 152 Wozniak AJ, Ross WE. DNA damage as a basis for 4'-demethylepipodophyllotoxin-9-(4,6-O-ethylidene-beta-D-glucopyranoside) (etoposide) cytotoxicity. *Cancer Res* 1983; **43**: 120–4.
- 153 Deweese JE, Osheroff N. The DNA cleavage reaction of topoisomerase II: Wolf in sheep's clothing. *Nucleic Acids Res* 2009; **37**: 738–748.
- 154 McKie SJ, Neuman KC, Maxwell A. DNA topoisomerases: Advances in understanding of cellular roles and multi-protein complexes via structure-function analysis. *BioEssays*. 2021; **43**: 2000286.
- 155 Watson JD, Crick FHC. Genetical implications of the structure of deoxyribonucleic acid. *Nature* 1953; **171**: 964–967.
- 156 Delbrück M. ON THE REPLICATION OF DESOXYRIBONUCLEIC ACID (DNA). *Proc Natl Acad Sci* 1954; **40**. doi:10.1073/pnas.40.9.783.
- 157 Champoux JJ. DNA TOPOISOMERASES: Structure, Function, and Mechanism. 2001; : 369–413.
- 158 Bliska JB, Cozzarelli NR. Use of site-specific recombination as a probe of DNA structure and metabolism in vivo. *J Mol Biol* 1987; **194**. doi:10.1016/0022-2836(87)90369-X.
- 159 Zechiedrich EL, Cozzarelli NR. Roles of topoisomerase IV and DNA gyrase in DNA unlinking during replication in Escherichia coli. *Genes Dev* 1995; **9**. doi:10.1101/gad.9.22.2859.
- 160 Yang L, Wold MS, Li JJ, Kelly TJ, Liu LF. Roles of DNA topoisomerases in simian virus 40 DNA replication in vitro. *Proc Natl Acad Sci U S A* 1987; **84**. doi:10.1073/pnas.84.4.950.
- 161 Buchenau P, Saumweber H, Arndt-Jovin DJ. Consequences of topoisomerase II inhibition in early embryogenesis of Drosophila revealed by in vivo confocal laser scanning microscopy. *J Cell Sci* 1993; **104**. doi:10.1242/jcs.104.4.1175.

- 162 Liu LF, Rowe TC, Yang L, Tewey KM, Chen GL. Cleavage of DNA by mammalian DNA topoisomerase II. *J Biol Chem* 1983; **258**: 15365–15370.
- 163 Schmidt BH, Osheroff N, Berger JM. Structure of a topoisomerase II–DNA–nucleotide complex reveals a new control mechanism for ATPase activity. *Nat Struct Mol Biol* 2012; **19**: 1147–1154.
- 164 Nitiss JL. DNA topoisomerase II and its growing repertoire of biological functions. *Nat. Rev. Cancer*. 2009; **9**: 327–337.
- 165 Bates AD, Berger JM, Maxwell A. The ancestral role of ATP hydrolysis in type II topoisomerases: prevention of DNA double-strand breaks. *Nucleic Acids Res* 2011; **39**: 6327–39.
- 166 Bush NG, Evans-roberts K, Maxwell A. Macromolecules DNA Topoisomerases. *EcoSal Plus* 2015; **6**: 1–34.
- 167 Bax BD, Murshudov G, Maxwell A, Germe T. DNA Topoisomerase Inhibitors: Trapping a DNA-Cleaving Machine in Motion. *J Mol Biol* 2019; **431**: 3427–3449.
- 168 Wang JC. Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol* 2002; **3**: 430–440.
- 169 Pommier Y, Sun Y, Huang SYN, Nitiss JL. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat. Rev. Mol. Cell Biol.* 2016; **17**: 703–721.
- 170 Chen SH, Chan N-L, Hsieh T. New Mechanistic and Functional Insights into DNA Topoisomerases. *Annu Rev Biochem* 2013; **82**: 139–170.
- 171 Pommier Y, Nussenzweig A, Takeda S, Austin C. Human topoisomerases and their roles in genome stability and organization. *Nat. Rev. Mol. Cell Biol.* 2022; **23**: 407–427.
- 172 Austin CA, Sng J-H, Patel S, Fisher LM. Novel HeLa topoisomerase II is the II β isoform: complete coding sequence and homology with other type II topoisomerases. *Biochim Biophys Acta - Gene Struct Expr* 1993; **1172**: 283–291.
- 173 Austin CA, Marsh KL. Eukaryotic DNA topoisomerase II beta. *Bioessays* 1998; **20**: 215–26.
- 174 Lee S, Jung S-R, Heo K, Byl JAW, Dewese JE, Osheroff N *et al.* DNA cleavage and opening reactions of human topoisomerase II α are regulated via Mg²⁺-mediated dynamic bending of gate-DNA. *Proc Natl Acad Sci U S A* 2012; **109**: 2925–30.
- 175 Schmidt BH, Burgin AB, Dewese JE, Osheroff N, Berger JM. A novel and unified two-metal mechanism for DNA cleavage by type II and IA topoisomerases. *Nature* 2010; **465**: 641–4.
- 176 Vanden Broeck A, Lotz C, Drillien R, Haas L, Bedez C, Lamour V. Structural basis for allosteric regulation of Human Topoisomerase II α . *Nat Commun* 2021; **12**. doi:10.1038/s41467-021-23136-6.
- 177 Chen SF, Huang NL, Lin JH, Wu CC, Wang YR, Yu YJ *et al.* Structural insights into the gating of DNA passage by the topoisomerase II DNA-gate. *Nat Commun* 2018; **9**. doi:10.1038/s41467-018-05406-y.

- 178 Adachi N, Miyaike M, Kato S, Kanamaru R, Koyama H, Kikuchi A. Cellular distribution of mammalian DNA topoisomerase II is determined by its catalytically dispensable C-terminal domain. *Nucleic Acids Res* 1997; **25**: 3135–3142.
- 179 Corbett KD, Shultzaberger RK, Berger JM. The C-terminal domain of DNA gyrase A adopts a DNA-bending beta-pinwheel fold. *Proc Natl Acad Sci U S A* 2004; **101**: 7293–7298.
- 180 McClendon AK, Rodriguez AC, Osheroff N. Human Topoisomerase II α Rapidly Relaxes Positively Supercoiled DNA. *J Biol Chem* 2005; **280**: 39337–39345.
- 181 McClendon AK, Gentry AC, Dickey JS, Brinch M, Bendsen S, Andersen AH *et al.* Bimodal Recognition of DNA Geometry by Human Topoisomerase II α : Preferential Relaxation of Positively Supercoiled DNA Requires Elements in the C-Terminal Domain. *Biochemistry* 2008; **47**: 13169–13178.
- 182 Taagepera S, Rao PN, Drake FH, Gorbsky GJ. DNA topoisomerase II alpha is the major chromosome protein recognized by the mitotic phosphoprotein antibody MPM-2. *Proc Natl Acad Sci* 1993; **90**: 8407–8411.
- 183 Woessner RD, Mattern MR, Mirabelli CK, Johnson RK, Drake FH. Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells. *Cell Growth Differ* 1991; **2**: 209–214.
- 184 Niimi A, Suka N, Harata M, Kikuchi A, Mizuno S. Co-localization of chicken DNA topoisomerase II α , but not β , with sites of DNA replication and possible involvement of a C-terminal region of α through its binding to PCNA. *Chromosoma* 2001; **110**: 102–114.
- 185 Farr CJ, Antoniou-Kourounioti M, Mimmack ML, Volkov A, Porter ACG. The α isoform of topoisomerase II is required for hypercompaction of mitotic chromosomes in human cells. *Nucleic Acids Res* 2014; **42**: 4414–4426.
- 186 Carpenter AJ, Porter ACG. Construction, Characterization, and Complementation of a Conditional-Lethal DNA Topoisomerase II Mutant Human Cell Line. *Mol Biol Cell* 2004; **15**: 5700–5711.
- 187 Akimitsu N, Adachi N, Hirai H, Hossain MS, Hamamoto H, Kobayashi M *et al.* Enforced cytokinesis without complete nuclear division in embryonic cells depleting the activity of DNA topoisomerase II α . *Genes to Cells* 2003; **8**: 393–402.
- 188 Lane AB, Giménez-Abián JF, Clarke DJ. A novel chromatin tether domain controls topoisomerase II α dynamics and mitotic chromosome formation. *J Cell Biol* 2013; **203**. doi:10.1083/jcb.201303045.
- 189 Yang X, Li W, Prescott ED, Burden SJ, Wang JC. DNA topoisomerase IIbeta and neural development. *Science* 2000; **287**: 131–4.
- 190 Lyu YL, Wang JC. Aberrant lamination in the cerebral cortex of mouse embryos lacking DNA topoisomerase IIbeta. *Proc Natl Acad Sci U S A* 2003; **100**: 7123–8.
- 191 Austin CA, Lee KC, Swan RL, Khazeem MM, Manville CM, Cridland P *et al.* TOP2B: The first thirty years. *Int. J. Mol. Sci.* 2018; **19**. doi:10.3390/ijms19092765.
- 192 Lyu YL, Lin C-P, Azarova AM, Cai L, Wang JC, Liu LF. Role of Topoisomerase II β in the Expression of Developmentally Regulated Genes. *Mol Cell Biol* 2006; **26**: 7929–7941.

- 193 Uusküla-Reimand L, Hou H, Samavarchi-Tehrani P, Rudan MV, Liang M, Medina-Rivera A *et al.* Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biol* 2016; **17**. doi:10.1186/s13059-016-1043-8.
- 194 Neha S, Dholaniya PS. The Prevailing Role of Topoisomerase 2 Beta and its Associated Genes in Neurons. *Mol. Neurobiol.* 2021; **58**. doi:10.1007/s12035-021-02561-0.
- 195 King IF, Yandava CN, Mabb AM, Hsiao JS, Huang HS, Pearson BL *et al.* Topoisomerases facilitate transcription of long genes linked to autism. *Nature* 2013; **501**. doi:10.1038/nature12504.
- 196 Linka RM, Porter ACG, Volkov A, Mielke C, Boege F, Christensen MO. C-Terminal regions of topoisomerase II α and II β determine isoform-specific functioning of the enzymes in vivo. *Nucleic Acids Res* 2007; **35**: 3810–3822.
- 197 Vann KR, Oviatt AA, Osheroff N. Topoisomerase II Poisons: Converting Essential Enzymes into Molecular Scissors. *Biochemistry* 2021; **60**: 1630–1641.
- 198 Robinson MJ, Osheroff N. Effects of antineoplastic drugs on the post-strand-passage DNA cleavage/religation equilibrium of topoisomerase II. *Biochemistry* 1991; **30**: 1807–1813.
- 199 Pommier Y, Marchand C. Interfacial inhibitors: targeting macromolecular complexes. *Nat Rev Drug Discov* 2012; **11**: 25–36.
- 200 Long BH, Musial ST, Brattain MG. Single- and double-strand DNA breakage and repair in human lung adenocarcinoma cells exposed to etoposide and teniposide. *Cancer Res* 1985; **45**: 3106–3112.
- 201 Pommier Y, Schwartz RE, Kohn KW, Zwelling LA. Formation and rejoining of deoxyribonucleic acid double-strand breaks induced in isolated cell nuclei by antineoplastic intercalating agents. *Biochemistry* 1984; **23**: 3194–3201.
- 202 Kantidze OL, Razin S V. Chemotherapy-related secondary leukemias: A role for DNA repair by error-prone non-homologous end joining in topoisomerase II — Induced chromosomal rearrangements. *Gene* 2007; **391**: 76–79.
- 203 Smith KA, Cowell IG, Zhang Y, Sondka Z, Austin CA. The role of topoisomerase II beta on breakage and proximity of RUNX1 to partner alleles RUNX1T1 and EVI1. *Genes Chromosom Cancer* 2014; **53**: 117–128.
- 204 Zhang Y, Stricteel P, Strick R, Chen J, Nucifora G, Beau MM Le *et al.* Genomic DNA breakpoints in AML1/RUNX1 and ETO cluster with topoisomerase II DNA cleavage and DNase I hypersensitive sites in t(8;21) leukemia. *Proc Natl Acad Sci U S A* 2002; **99**: 3070–3075.
- 205 Stricteel PL, Strick R, Tomek RJ, Roe B a, Rowley JD, Zeleznik-Le NJ. DNA structural properties of AF9 are similar to MLL and could act as recombination hot spots resulting in MLL/AF9 translocations and leukemogenesis. *Hum Mol Genet* 2000; **9**: 1671–1679.
- 206 Riccio AA, Schellenberg MJ, Williams RS. Molecular mechanisms of topoisomerase 2 DNA–protein crosslink resolution. *Cell. Mol. Life Sci.* 2020; **77**: 81–91.
- 207 Canela A, Maman Y, Huang S yin N, Wutz G, Tang W, Zagnoli-Vieira G *et al.* Topoisomerase II-Induced Chromosome Breakage and Translocation Is Determined by Chromosome Architecture and Transcriptional Activity. *Mol Cell* 2019; **75**: 252–266.e8.

- 208 Yan H, Tammaro M, Liao S. Collision of trapped topoisomerase 2 with transcription and replication: Generation and repair of DNA double-strand breaks with 5' adducts. *Genes (Basel)* 2016; **7**. doi:10.3390/genes7070032.
- 209 Swan RL, Cowell IG, Austin CA. Mechanisms to Repair Stalled Topoisomerase II-DNA Covalent Complexes. *Mol Pharmacol* 2022; **101**: 24–32.
- 210 Zeng Z, Cortés-Ledesma F, El Khamisy SF, Caldecott KW. TDP2/TTRAP is the major 5'-tyrosyl DNA phosphodiesterase activity in vertebrate cells and is critical for cellular resistance to topoisomerase II-induced DNA damage. *J Biol Chem* 2011; **286**. doi:10.1074/jbc.M110.181016.
- 211 Ledesma FC, El Khamisy SF, Zuma MC, Osborn K, Caldecott KW. A human 5'-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage. *Nature* 2009; **461**. doi:10.1038/nature08444.
- 212 Gao R, Schellenberg MJ, Huang SYN, Abdelmalak M, Marchand C, Nitiss KC *et al*. Proteolytic degradation of topoisomerase II (Top2) enables the processing of Top2·DNA and Top2·RNA covalent complexes by tyrosyl-DNA-phosphodiesterase 2 (TDP2). *J Biol Chem* 2014; **289**. doi:10.1074/jbc.M114.565374.
- 213 Ban Y, Ho C-W, Lin R-K, Lyu YL, Liu LF. Activation of a Novel Ubiquitin-Independent Proteasome Pathway when RNA Polymerase II Encounters a Protein Roadblock. *Mol Cell Biol* 2013; **33**. doi:10.1128/mcb.00403-13.
- 214 Swan RL, Poh LLK, Cowell IG, Austin CA. Small molecule inhibitors confirm ubiquitin-dependent removal of TOP2-DNA covalent complexes. *Mol Pharmacol* 2020; **98**. doi:10.1124/mol.119.118893.
- 215 Lopez-Mosqueda J, Maddi K, Prgomet S, Kalayil S, Marinovic-Terzic I, Terzic J *et al*. SPRTN is a mammalian DNA-binding metalloprotease that resolves DNA-protein crosslinks. *Elife* 2016; **5**. doi:10.7554/eLife.21491.
- 216 Vaz B, Popovic M, Newman JA, Fielden J, Aitkenhead H, Halder S *et al*. Metalloprotease SPRTN/DVC1 Orchestrates Replication-Coupled DNA-Protein Crosslink Repair. *Mol Cell* 2016; **64**. doi:10.1016/j.molcel.2016.09.032.
- 217 Ruggiano A, Vaz B, Kilgas S, Popović M, Rodriguez-Berriguete G, Singh AN *et al*. The protease SPRTN and SUMOylation coordinate DNA-protein crosslink repair to prevent genome instability. *Cell Rep* 2021; **37**. doi:10.1016/j.celrep.2021.110080.
- 218 Zagnoli-Vieira G, Caldecott KW. TDP2, TOP2, and SUMO: What is ZATT about? *Cell Res*. 2017; **27**. doi:10.1038/cr.2017.147.
- 219 Schellenberg MJ, Lieberman JA, Herrero-Ruiz A, Butler LR, Williams JG, Muñoz-Cabello AM *et al*. ZATT (ZNF451)-mediated resolution of topoisomerase 2 DNA-protein crosslinks. *Science (80-)* 2017; **357**. doi:10.1126/science.aam6468.
- 220 Buhler C, Gadelle D, Forterre P, Wang JC, Bergerat a. Reconstitution of DNA topoisomerase VI of the thermophilic archaeon *Sulfolobus shibatae* from subunits separately overexpressed in *Escherichia coli*. *Nucleic Acids Res* 1998; **26**: 5157–62.
- 221 Bergerat A, de Massy B, Gadelle D, Varoutas PC, Nicolas A, Forterre P. An atypical topoisomerase II from Archaea with implications for meiotic recombination. *Nature* 1997; **386**: 414–417.

- 222 Neale MJ, Pan J, Keeney S. Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* 2005; **436**. doi:10.1038/nature03872.
- 223 Milman N, Higuchi E, Smith GR. Meiotic DNA Double-Strand Break Repair Requires Two Nucleases, MRN and Ctp1, To Produce a Single Size Class of Rec12 (Spo11)-Oligonucleotide Complexes. *Mol Cell Biol* 2009; **29**. doi:10.1128/mcb.01127-09.
- 224 Rothenberg M, Kohli J, Ludin K. Ctp1 and the MRN-complex are required for endonucleolytic Rec12 removal with release of a single class of oligonucleotides in fission yeast. *PLoS Genet* 2009; **5**. doi:10.1371/journal.pgen.1000722.
- 225 Aparicio T, Baer R, Gottesman M, Gautier J. MRN, CtIP, and BRCA1 mediate repair of topoisomerase II-DNA adducts. *J Cell Biol* 2016; **212**: 399–408.
- 226 Hoa NN, Shimizu T, Zhou ZW, Wang Z-Q, Deshpande RA, Paull TT *et al.* Mre11 Is Essential for the Removal of Lethal Topoisomerase 2 Covalent Cleavage Complexes. *Mol Cell* 2016; **64**: 580–592.
- 227 Nakamura K, Kogame T, Oshiumi H, Shinohara A, Sumitomo Y, Agama K *et al.* Collaborative action of Brca1 and CtIP in elimination of covalent modifications from double-strand breaks to facilitate subsequent break repair. *PLoS Genet* 2010; **6**. doi:10.1371/journal.pgen.1000828.
- 228 Tomicic MT, Kaina B. Topoisomerase degradation, DSB repair, p53 and IAPs in cancer cell resistance to camptothecin-like topoisomerase I inhibitors. *Biochim. Biophys. Acta - Rev. Cancer*. 2013; **1835**: 11–27.
- 229 Montecucco A, Biamonti G. Cellular response to etoposide treatment. *Cancer Lett.* 2007; **252**: 9–18.
- 230 Callegari AJ, Kelly TJ. Shedding light on the DNA damage checkpoint. *Cell Cycle*. 2007; **6**: 660–666.
- 231 Hartlerode AJ, Scully R. Mechanisms of double-strand break repair in somatic mammalian cells. *Biochem J* 2009; **423**: 157–168.
- 232 Hanscom T, McVey M. Regulation of Error-Prone DNA Double-Strand Break Repair and Its Impact on Genome Evolution. *Cells*. 2020; **9**. doi:10.3390/cells9071657.
- 233 Han J, Huang J. DNA double-strand break repair pathway choice: the fork in the road. *Genome Instab Dis* 2020; **1**: 10–19.
- 234 Burma S, Chen BPC, Chen DJ. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. *DNA Repair (Amst)* 2006; **5**: 1042–8.
- 235 Zapotoczny G, Sekelsky J. Human cell assays for synthesis-dependent strand annealing and crossing over during double-strand break repair. *G3 Genes, Genomes, Genet* 2017; **7**: 1191–1199.
- 236 Mitchel K, Lehner K, Jinks-Robertson S. Heteroduplex DNA Position Defines the Roles of the Sgs1, Srs2, and Mph1 Helicases in Promoting Distinct Recombination Outcomes. *PLoS Genet* 2013; **9**. doi:10.1371/journal.pgen.1003340.
- 237 Sung P, Klein H. Mechanism of homologous recombination: Mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* 2006; **7**: 739–750.

- 238 Larsen NB, Hickson ID. RecQ helicases: Conserved guardians of genomic integrity. *Adv. Exp. Med. Biol.* 2013; **767**. doi:10.1007/978-1-4614-5037-5_8.
- 239 Harami GM, Pálinkás J, Seol Y, Kovács ZJ, Gyimesi M, Harami-Papp H *et al.* The topoisomerase IIIalpha-RMI1-RMI2 complex orients human Bloom's syndrome helicase for efficient disruption of D-loops. *Nat Commun* 2022; **13**. doi:10.1038/s41467-022-28208-9.
- 240 Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW. The double-strand-break repair model for recombination. *Cell.* 1983; **33**. doi:10.1016/0092-8674(83)90331-8.
- 241 Hunter N. Meiotic recombination: The essence of heredity. *Cold Spring Harb Perspect Biol* 2015; **7**. doi:10.1101/cshperspect.a016618.
- 242 Ip SCY, Rass U, Blanco MG, Flynn HR, Skehel JM, West SC. Identification of Holliday junction resolvases from humans and yeast. *Nature* 2008; **456**. doi:10.1038/nature07470.
- 243 Chan YW, West SC. Spatial control of the GEN1 Holliday junction resolvase ensures genome stability. *Nat Commun* 2014; **5**. doi:10.1038/ncomms5844.
- 244 Sanchez A, Reginato G, Cejka P. Crossover or non-crossover outcomes: tailored processing of homologous recombination intermediates. *Curr Opin Genet Dev* 2021; **71**: 39–47.
- 245 Ackerson SM, Romney C, Schuck PL, Stewart JA. To Join or Not to Join: Decision Points Along the Pathway to Double-Strand Break Repair vs. Chromosome End Protection. *Front Cell Dev Biol* 2021; **9**: 1–19.
- 246 Tisi R, Vertemara J, Zampella G, Longhese MP. Functional and structural insights into the MRX/MRN complex, a key player in recognition and repair of DNA double-strand breaks. *Comput Struct Biotechnol J* 2020; **18**: 1137–1152.
- 247 Takaoka M, Miki Y. BRCA1 gene: function and deficiency. *Int. J. Clin. Oncol.* 2018; **23**: 36–44.
- 248 Godin SK, Sullivan MR, Bernstein KA. Novel insights into RAD51 activity and regulation during homologous recombination and DNA replication. *Biochem Cell Biol* 2016; **94**: 407–418.
- 249 Blackford AN, Jackson SP. ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol Cell* 2017; **66**: 801–817.
- 250 Redon C, Pilch D, Rogakou E, Sedelnikova O, Newrock K, Bonner W. Histone H2A variants H2AX and H2AZ. *Curr Opin Genet Dev* 2002; **12**: 162–169.
- 251 Valdiglesias V, Giunta S, Fenech M, Neri M, Bonassi S. γ H2AX as a marker of DNA double strand breaks and genomic instability in human population studies. *Mutat. Res. - Rev. Mutat. Res.* 2013; **753**. doi:10.1016/j.mrrev.2013.02.001.
- 252 Karanam K, Kafri R, Loewer A, Lahav G. Quantitative Live Cell Imaging Reveals a Gradual Shift between DNA Repair Mechanisms and a Maximal Use of HR in Mid S Phase. *Mol Cell* 2012; **47**: 320–329.
- 253 Shibata A. Regulation of repair pathway choice at two-ended DNA double-strand breaks. *Mutat Res Mol Mech Mutagen* 2017; **803–805**: 51–55.

- 254 Shahar OD, Ram EVSR, Shimshoni E, Hareli S, Meshorer E, Goldberg M. Live imaging of induced and controlled DNA double-strand break formation reveals extremely low repair by homologous recombination in human cells. *Oncogene* 2012; **31**: 3495–3504.
- 255 Thacker J, Zdzienicka MZ. The mammalian XRCC genes: Their roles in DNA repair and genetic stability. *DNA Repair (Amst)*. 2003; **2**: 655–672.
- 256 Grundy GJ, Moulding HA, Caldecott KW, Rulten SL. One ring to bring them all-The role of Ku in mammalian non-homologous end joining. *DNA Repair (Amst)* 2014; **17**: 30–38.
- 257 Frit P, Ropars V, Modesti M, Charbonnier JB, Calsou P. Plugged into the Ku-DNA hub: The NHEJ network. *Prog Biophys Mol Biol* 2019; **147**: 62–76.
- 258 Graham TGW, Walter JC, Loparo JJ. Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Mol Cell* 2016; **61**: 850–858.
- 259 Chaplin AK, Hardwick SW, Stavridi AK, Buehl CJ, Goff NJ, Ropars V *et al*. Cryo-EM of NHEJ supercomplexes provides insights into DNA repair. *Mol Cell* 2021; **81**: 3400–3409.e3.
- 260 DeFazio LG, Stansel RM, Griffith JD, Chu G. Synapsis of DNA ends by DNA-dependent protein kinase. *EMBO J* 2002; **21**: 3192–3200.
- 261 Jiang W, Crowe JL, Liu X, Nakajima S, Wang Y, Li C *et al*. Differential Phosphorylation of DNA-PKcs Regulates the Interplay between End-Processing and End-Ligation during Nonhomologous End-Joining. *Mol Cell* 2015; **58**: 172–185.
- 262 Menon V, Povirk LF. End-processing nucleases and phosphodiesterases: An elite supporting cast for the non-homologous end joining pathway of DNA double-strand break repair. *DNA Repair (Amst)*. 2016; **43**: 57–68.
- 263 Volk T, Pannicke U, Reisli I, Bulashevskaya A, Ritter J, Björkman A *et al*. DCLRE1C (ARTEMIS) mutations causing phenotypes ranging from atypical severe combined immunodeficiency to mere antibody deficiency. *Hum Mol Genet* 2015; **24**. doi:10.1093/hmg/ddv437.
- 264 Ma Y, Pannicke U, Schwarz K, Lieber MR. Hairpin Opening and Overhang Processing by an Artemis/DNA-Dependent Protein Kinase Complex in Nonhomologous End Joining and V(D)J Recombination. *Cell* 2002; **108**: 781–794.
- 265 Moon AF, Garcia-Diaz M, Bebenek K, Davis BJ, Zhong X, Ramsden DA *et al*. Structural insight into the substrate specificity of DNA Polymerase μ . *Nat Struct Mol Biol* 2007; **14**. doi:10.1038/nsmb1180.
- 266 Chang HHY, Watanabe G, Gerodimos CA, Ochi T, Blundell TL, Jackson SP *et al*. Different DNA End Configurations Dictate Which NHEJ Components Are Most Important for Joining Efficiency. *J Biol Chem* 2016; **291**: 24377–24389.
- 267 Pannunzio NR, Watanabe G, Lieber MR. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J Biol Chem* 2018; **293**: 10512–10523.
- 268 Chen S, Lee L, Naila T, Fishbain S, Wang A, Tomkinson AE *et al*. Structural basis of long-range to short-range synaptic transition in NHEJ. *Nature* 2021; **593**: 294–298.
- 269 Menon V, Povirk LF. XLF/Cernunnos: An important but puzzling participant in the nonhomologous end joining DNA repair pathway. *DNA Repair (Amst)*. 2017; **58**: 29–37.

- 270 Menolfi D, Zha S. DNA-PKcs kinase activity orchestrates both end-processing and end-ligation. *Trends Cell Biol* 2022; **32**: 91–93.
- 271 Kabotyanski EB, Gomelsky L, Han JO, Stamato TD, Roth DB. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res* 1998; **26**: 5333–5342.
- 272 Wang HH, Perrault AR, Takeda Y, Qin W, Iliakis G. Biochemical evidence for Ku-independent backup pathways of NHEJ. *Nucleic Acids Res* 2003; **31**: 5377–5388.
- 273 Schimmel J, van Schendel R, den Dunnen JT, Tijsterman M. Templated Insertions: A Smoking Gun for Polymerase Theta-Mediated End Joining. *Trends Genet* 2019; **35**: 632–644.
- 274 Gu J, Lu H, Tippin B, Shimazaki N, Goodman MF, Lieber MR. XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO J* 2007; **26**. doi:10.1038/sj.emboj.7601559.
- 275 Liang L, Deng L, Nguyen SC, Zhao X, Maulion CD, Shao C *et al*. Human DNA ligases I and III, but not ligase IV, are required for microhomology-mediated end joining of DNA double-strand breaks. *Nucleic Acids Res* 2008; **36**: 3297–3310.
- 276 Bhargava R, Onyango DO, Stark JM. Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends Genet* 2016; **32**: 566–575.
- 277 Li J, Xu X. DNA double-strand break repair: A tale of pathway choices. *Acta Biochim Biophys Sin (Shanghai)* 2016; **48**: 641–646.
- 278 Shibata A. Regulation of repair pathway choice at two-ended DNA double-strand breaks. 2017.
- 279 Shibata A, Moiani D, Arvai AS, Perry J, Harding SM, Genois MM *et al*. DNA Double-Strand Break Repair Pathway Choice Is Directed by Distinct MRE11 Nuclease Activities. *Mol Cell* 2014; **53**: 7–18.
- 280 Aymard F, Bugler B, Schmidt CK, Guillou E, Caron P, Briois S *et al*. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat Struct Mol Biol* 2014; **21**. doi:10.1038/nsmb.2796.
- 281 Montgomery EA, Huang SM, Langley CH, Judd BH. Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: Genome structure and evolution. *Genetics* 1991; **129**. doi:10.1093/genetics/129.4.1085.
- 282 Stoppa-Lyonnet D, Carter PE, Meo T, Tosi M. Clusters of intragenic Alu repeats predispose the human C1 inhibitor locus to deleterious rearrangements. *Proc Natl Acad Sci U S A* 1990; **87**. doi:10.1073/pnas.87.4.1551.
- 283 Bunting SF, Nussenzweig A. End-joining, translocations and cancer. *Nat Rev Cancer* 2013; **13**: 443–454.
- 284 Maede Y, Shimizu H, Fukushima T, Kogame T, Nakamura T, Miki T *et al*. Differential and common DNA repair pathways for topoisomerase I- and II-targeted drugs in a genetic DT40 repair cell screen panel. *Mol Cancer Ther* 2013; **13**. doi:10.1158/1535-7163.MCT-13-0551.

- 285 Sasanuma H, Tsuda M, Morimoto S, Saha LK, Rahman MM, Kiyooka Y *et al.* BRCA1 ensures genome integrity by eliminating estrogen-induced pathological topoisomerase II-DNA complexes. *Proc Natl Acad Sci U S A* 2018; **115**: E10642–E10651.
- 286 Zhang B, Tang Z, Li L, Lu LY. NBS1 is required for SPO11-linked DNA double-strand break repair in male meiosis. *Cell Death Differ* 2020; **27**: 2176–2190.
- 287 Roth DB. V(D)J Recombination: Mechanism, Errors, and Fidelity. *Microbiol Spectr* 2014; **2**: 1–11.
- 288 Bentley J, Diggle CP, Harnden P, Knowles MA, Kiltie AE. DNA double strand break repair in human bladder cancer is error prone and involves microhomology-associated end-joining. *Nucleic Acids Res* 2004; **32**: 5249–5259.
- 289 Boboila C, Jankovic M, Yan CT, Wang JH, Wesemann DR, Zhang T *et al.* Alternative end-joining catalyzes robust IgH locus deletions and translocations in the combined absence of ligase 4 and Ku70. *Proc Natl Acad Sci U S A* 2010; **107**: 3034–3039.
- 290 Simsek D, Jasin M. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4–ligase IV during chromosomal translocation formation. *Nat Struct Mol Biol* 2010; **17**: 410–416.
- 291 Goff JP, Shields DS, Seki M, Choi S, Epperly MW, Dixon T *et al.* Lack of DNA Polymerase θ (POLQ) radiosensitizes bone marrow stromal cells in vitro and increases reticulocyte micronuclei after total-body irradiation. *Radiat Res* 2009; **172**: 165–174.
- 292 Yousefzadeh MJ, Wyatt DW, Takata K ichi, Mu Y, Hensley SC, Tomida J *et al.* Mechanism of Suppression of Chromosomal Instability by DNA Polymerase POLQ. *PLoS Genet* 2014; **10**. doi:10.1371/journal.pgen.1004654.
- 293 Ghezraoui H, Piganeau M, Renouf B, Renaud JB, Sallmyr A, Ruis B *et al.* Chromosomal Translocations in Human Cells Are Generated by Canonical Nonhomologous End-Joining. *Mol Cell* 2014; **55**: 829–842.
- 294 Soni A, Siemann M, Pantelias GE, Iliakis G. Marked contribution of alternative end-joining to chromosome-translocation-formation by stochastically induced DNA double-strand-breaks in G2-phase human cells. *Mutat Res - Genet Toxicol Environ Mutagen* 2015; **793**: 2–8.
- 295 Brunet E, Simsek D, Tomishima M, DeKelver R, Choi VM, Gregory P *et al.* Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl Acad Sci U S A* 2009; **106**: 10620–10625.
- 296 Bouvet P, Mongelard F. *Epigenetics: Development and Disease*. Springer Netherlands: Dordrecht, 2013 doi:10.1007/978-94-007-4525-4.
- 297 Cremer M, von Hase J, Volm T, Brero A, Kreth G, Walter J *et al.* Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res* 2001; **9**: 541–567.
- 298 Cremer T, Cremer M. Chromosome territories. *Cold Spring Harb Perspect Biol* 2010; **2**: a003889.
- 299 Speicher MR, Ballard SG, Ward DC. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet* 1996; **12**. doi:10.1038/ng0496-368.

- 300 Schröck E, Du Manoir S, Veldman T, Schoell B, Wienberg J, Ferguson-Smith MA *et al.* Multicolor spectral karyotyping of human chromosomes. *Science (80-)* 1996; **273**. doi:10.1126/science.273.5274.494.
- 301 Branco MR, Pombo A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biol* 2006; **4**: 780–788.
- 302 Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014; **159**. doi:10.1016/j.cell.2014.11.021.
- 303 Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A *et al.* Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science (80-)* 2009; **326**: 289–293.
- 304 Cremer T, Cremer M, Cremer C. The 4D Nucleome: Genome Compartmentalization in an Evolutionary Context. *Biochem.* 2018; **83**. doi:10.1134/S000629791804003X.
- 305 Razin S V., Ulianov S V., Gavrilov AA. 3D Genomics. *Mol Biol* 2019; **53**: 802–812.
- 306 Mirny LA, Imakaev M, Abdennur N. Two major mechanisms of chromosome organization. *Curr. Opin. Cell Biol.* 2019; **58**. doi:10.1016/j.ceb.2019.05.001.
- 307 Falk M, Lukasova E, Kozubek S. Higher-order chromatin structure in DSB induction, repair and misrepair. *Mutat Res* 2010; **704**: 88–100.
- 308 Meaburn KJ, Misteli T, Soutoglou E. Spatial genome organization in the formation of chromosomal translocations. *Semin. Cancer Biol.* 2007; **17**. doi:10.1016/j.semcancer.2006.10.008.
- 309 Nikiforova MN, Stringer JR, Blough R, Medvedovic M, Fagin JA, Nikiforov YE. Proximity of chromosomal loci that participate in radiation-induced rearrangements in human cells. *Science (80-)* 2000; **290**: 138–141.
- 310 Gandhi M, Medvedovic M, Stringer JR, Nikiforov YE. Interphase chromosome folding determines spatial proximity of genes participating in carcinogenic RET/PTC rearrangements. *Oncogene* 2006; **25**: 2360–6.
- 311 Arsuaga J, Greulich-Bode KM, Vazquez M, Bruckner M, Hahnfeldt P, Brenner DJ *et al.* Chromosome spatial clustering inferred from radiogenic aberrations. *Int J Radiat Biol* 2004; **80**. doi:10.1080/09553000410001723857.
- 312 Gandhi MS, Stringer JR. Gene position within chromosome territories correlates with their involvement in distinct rearrangement types in thyroid cancer cells. ..., *Chromosom ...* 2009; **228**: 222–228.
- 313 Gué M, Sun JS, Boudier T. Simultaneous localization of MLL, AF4 and ENL genes in interphase nuclei by 3D-FISH: MLL translocation revisited. *BMC Cancer* 2006; **6**: 3–7.
- 314 Zhang Y, McCord RP, Ho Y-J, Lajoie BR, Hildebrand DG, Simon AC *et al.* Chromosomal translocations are guided by the spatial organization of the genome. *Cell* 2012; **148**: 908–21.
- 315 Engreitz JM, Agarwala V, Mirny LA. Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease. *PLoS One* 2012; **7**: e44196.

- 316 Aten J a, Stap J, Krawczyk PM, van Oven CH, Hoebe R a, Essers J *et al.* Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* 2004; **303**: 92–5.
- 317 Falk M, Lukasova E, Gabrielova B, Ondrej V, Kozubek S. Chromatin dynamics during DSB repair. *Biochim Biophys Acta - Mol Cell Res* 2007; **1773**: 1534–1545.
- 318 Neumaier T, Swenson J, Pham C, Polyzos A, Lo AT, Yang PA *et al.* Evidence for formation of DNA repair centers and dose-response nonlinearity in human cells. *Proc Natl Acad Sci U S A* 2012; **109**: 443–448.
- 319 Soutoglou E, Dorn JF, Sengupta K, Jasin M, Nussenzweig A, Ried T *et al.* Positional stability of single double-strand breaks in mammalian cells. *Nat Cell Biol* 2007; **9**: 675–82.
- 320 Roukos V, Voss TC, Schmidt CK, Lee S, Wangsa D, Misteli T. Spatial Dynamics of Chromosome Translocations in Living Cells. *Science (80-)* 2013; **341**: 660–664.
- 321 Caron P, Choudjaye J, Clouaire T, Bugler B, Daburon V, Aguirrebengoa M *et al.* Non-redundant Functions of ATM and DNA-PKcs in Response to DNA Double-Strand Breaks. *Cell Rep* 2015; **13**: 1598–609.
- 322 Aymard F, Aguirrebengoa M, Guillou E, Javierre BM, Bugler B, Arnould C *et al.* Genome-wide mapping of long-range contacts unveils clustering of DNA double-strand breaks at damaged active genes. *Nat Struct Mol Biol* 2017; **24**: 353–361.
- 323 Krawczyk PM, Borovski T, Stap J, Cijssouw T, ten Cate R, Medema JP *et al.* Chromatin mobility is increased at sites of DNA double-strand breaks. *J Cell Sci* 2012; **125**: 2127–2133.
- 324 Rubtsov MA, Glukhov SI, Allinne J, Pichugin A, Vassetzky YS, Razin S V *et al.* Treatment of lymphoid cells with the topoisomerase II poison etoposide leads to an increased juxtaposition of. *Biopolym Cell* 2011; **27**: 398–403.
- 325 Glukhov SI, Rubtsov M a, Alexeyevsky D a, Alexeevski A V, Razin S V, Iarovaia O V. The Broken MLL Gene Is Frequently Located Outside the Inherent Chromosome Territory in Human Lymphoid Cells Treated with DNA Topoisomerase II Poison Etoposide. *PLoS One* 2013; **8**: e75871.
- 326 Do TU, Ho B, Shih S-J, Vaughan A. Zinc Finger Nuclease induced DNA double stranded breaks and rearrangements in MLL. *Mutat Res* 2012; **740**: 34–42.
- 327 Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H *et al.* Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. *Cell* 2011; **147**: 95–106.
- 328 Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ *et al.* Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. *Cell* 2011; **147**: 107–119.
- 329 Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon K-R, Jankovic M *et al.* DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. *Nature* 2012; **484**: 69–74.
- 330 Canela A, Maman Y, Jung S, Wong N, Callen E, Day A *et al.* Genome Organization Drives Chromosome Fragility. *Cell* 2017; **170**: 507–521.e18.

- 331 Torrance CJ, Agrawal V, Vogelstein B, Kinzler KW. Use of isogenic human cancer cells for high-throughput screening and drug discovery. *Nat Biotechnol* 2001; **19**. doi:10.1038/nbt1001-940.
- 332 Gaj T, Sirk SJ, Shui SL, Liu J. Genome-editing technologies: Principles and applications. *Cold Spring Harb Perspect Biol* 2016; **8**: a023754.
- 333 Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc Natl Acad Sci* 1996; **93**: 1156–1160.
- 334 Dutta S, Madan S, Sundar D. Exploiting the recognition code for elucidating the mechanism of zinc finger protein-DNA interactions. *BMC Genomics* 2016; **17**. doi:10.1186/s12864-016-3324-8.
- 335 Smith J, Bibikova M, G. Whitby F, Reddy AR, Chandrasegaran, Srinivasan Carroll D. Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res* 2000; **28**: 3361–3369.
- 336 Ramirez CL, Foley JE, Wright DA, Müller-Lerch F, Rahman SH, Cornu TI *et al*. Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat. Methods*. 2008; **5**. doi:10.1038/nmeth0508-374.
- 337 Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science (80-)* 2009; **326**: 1501.
- 338 Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A *et al*. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* 2010; **186**: 756–761.
- 339 Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF *et al*. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol* 2011; **29**: 143–150.
- 340 Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* 1987; **169**: 5429–5433.
- 341 Bolotin A, Quinquis B, Sorokin A, Dusko Ehrlich S. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 2005; **151**: 2551–2561.
- 342 Makarova KS, Grishin N V, Shabalina SA, Wolf YI, Koonin E V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* 2006; **1**: 7.
- 343 Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S *et al*. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 2007; **315**: 1709–1712.
- 344 Sorek R, Lawrence CM, Wiedenheft B. CRISPR-Mediated Adaptive Immune Systems in Bacteria and Archaea. *Annu Rev Biochem* 2013; **82**: 237–266.
- 345 Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA *et al*. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* 2011; **471**: 602–607.

- 346 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* (80-). 2012; **337**: 816–821.
- 347 Cong L, Ran F, Cox D, Lin S, Barretto R. Multiplex Genome Engineering Using CRISPR / Cas Systems. *Science* (80-) 2013; : 3–6.
- 348 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE *et al.* RNA-guided human genome engineering via Cas9. *Science* 2013; **339**: 823–6.
- 349 Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. *Elife* 2013; **2013**.
- 350 Menzorov AG, Lukyanchikova VA, Korablev AN, Serova IA, Fishman VS. Genome editing using CRISPR/ Cas9 system: a practical guide. *Vavilov J Genet Breed* 2016; **20**: 930–944.
- 351 Maddalo D, Manchado E, Concepcion CP, Bonetti C, Vidigal JA, Han YC *et al.* In vivo engineering of oncogenic chromosomal rearrangements with the CRISPR/Cas9 system. *Nature* 2014; **516**: 423–428.
- 352 Torres R, Martin MC, Garcia A, Cigudosa JC, Ramirez JC, Rodriguez-Perales S. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system. *Nat Commun* 2014; **5**: 1–8.
- 353 Borovkov AA. *Stochastic Processes in Queueing Theory*. 1976 doi:10.1007/978-1-4612-9866-3.
- 354 Gushchanskaya ES, Artemov A V., Ulyanov S V., Logacheva MD, Penin AA, Kotova ES *et al.* The clustering of CpG islands may constitute an important determinant of the 3D organization of interphase chromosomes. *Epigenetics* 2014; **9**: 951–963.
- 355 Lomov NA, Viushkov VS, Ulianov S V, Gavrilov AA, Alexeyevsky DA, Artemov A V *et al.* Recurrent Translocations in Topoisomerase Inhibitor-Related Leukemia Are Determined by the Features of DNA Breaks Rather Than by the Proximity of the Translocating Genes. *Int J Mol Sci* 2022; **23**: 9824.
- 356 Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001; **29**: 45e–45.
- 357 Baleriola J, Suárez T, De La Rosa EJ. DNA-PK promotes the survival of young neurons in the embryonic mouse retina. *Cell Death Differ* 2010; **17**. doi:10.1038/cdd.2010.46.
- 358 Nutley BP, Smith NF, Hayes A, Kelland LR, Brunton L, Golding BT *et al.* Preclinical pharmacokinetics and metabolism of a novel prototype DNA-PK inhibitor NU7026. *Br J Cancer* 2005; **93**. doi:10.1038/sj.bjc.6602823.
- 359 Glantz SA. Biostatistics: how to detect, correct and prevent errors in the medical literature. *Circulation* 1980; **61**. doi:10.1161/01.CIR.61.1.1.
- 360 Iannuccelli E, Mompert F, Gellin J, Lahbib-Mansais Y, Yerle M, Boudier T. NEMO: A tool for analyzing gene and chromosome territory distributions from 3D-FISH experiments. *Bioinformatics*. 2010; **26**. doi:10.1093/bioinformatics/btq013.
- 361 Syljuåsen RG, McBride WH. Radiation-induced apoptosis and cell cycle progression in Jurkat T cells. *Radiat Res* 1999; **152**: 328–331.

- 362 Баттулин НР, Фишман ВС, Орлов ЮЛ, Мензоров АГ, Афонников ДА, Серов ОЛ. 3С-Методы В Исследованиях Пространственной Организации Генома. *Вавиловский Журнал Генетики И Селекции* 2012; **16**: 872–878.
- 363 Splinter E, de Wit E, van de Werken HJG, Klous P, de Laat W. Determining long-range chromatin interactions for selected genomic sites using 4C-seq technology: From fixation to computation. *Methods* 2012; **58**: 221–230.
- 364 Peterson LF, Boyapati A, Ahn E-Y, Biggs JR, Okumura AJ, Lo M-C *et al.* Acute myeloid leukemia with the 8q22;21q22 translocation: secondary mutational events and alternative t(8;21) transcripts. *Blood* 2007; **110**: 799–805.
- 365 Zhao Z, Tavoosidana G, Sjölander M, Göndör A, Mariano P, Wang S *et al.* Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet* 2006; **38**: 1341–1347.
- 366 Gossen M, Freundlieb S, Bender G, Müller G, Hillen W, Bujard H. Transcriptional activation by tetracyclines in mammalian cells. *Science* (80-) 1995; **268**: 1766–1769.
- 367 Jun-ichi M, Satoshi T, Kimi A, Fumi T, Akira T, Kiyoshi T *et al.* Expression vector system based on the chicken β -actin promoter directs efficient production of interleukin-5. *Gene* 1989; **79**: 269–277.
- 368 Sadelain M, Papapetrou EP, Bushman FD. Safe harbours for the integration of new DNA in the human genome. *Nat Rev Cancer* 2011; **12**: 51–58.
- 369 Kim JH, Lee SR, Li LH, Park HJ, Park JH, Lee KY *et al.* High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* 2011; **6**. doi:10.1371/journal.pone.0018556.
- 370 Xiang X, Corsi GI, Anthon C, Qu K, Pan X, Liang X *et al.* Enhancing CRISPR-Cas9 gRNA efficiency prediction by data integration and deep learning. *Nat Commun* 2021; **12**. doi:10.1038/s41467-021-23576-0.
- 371 Lomov NA, Viushkov VS, Petrenko AP, Syrkina MS, Rubtsov MA. Methods of Evaluating the Efficiency of CRISPR/Cas Genome Editing. *Mol. Biol.* 2019; **53**: 862–875.
- 372 Germini D, Saada YB, Tsfasman T, Osina K, Robin C, Lomov N *et al.* A one-step PCR-based assay to evaluate efficiency and precision of genomic DNA-editing tools. *Mol Ther - Methods Clin Dev* 2017; **5**: 43–50.
- 373 Lomov NA, Viushkov VS, Zamalutdinov A V., Sboeva MD, Rubtsov MA. Direct ENIT: An easy and reliable tool for gRNA efficacy verification by tracking induced chromosomal translocation. *MethodsX* 2020; **7**: 101104.
- 374 Lomov NAA, Borunova VV V, Rubtsov MAA. CRISPR/Cas9 technology for targeted genome editing. *Biopolym Cell* 2015; **31**: 243–248.
- 375 Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* 2014; **513**: 569–573.
- 376 Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H, Hillen W. Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci* 2000; **97**: 7963–7968.

- 377 González F, Zhu Z, Shi ZD, Lelli K, Verma N, Li Q V. *et al.* An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. *Cell Stem Cell* 2014; **15**: 215–226.
- 378 Shmakova A, Lomov N, Viushkov V, Tsfasman T, Kozhevnikova Y, Sokolova D *et al.* Cell models with inducible oncogenic translocations allow to evaluate the potential of drugs to favor secondary translocations. *Cancer Commun* 2023; **43**: 154–158.
- 379 Sathitruangsak C, Righolt CH, Klewes L, Tung Chang D, Kotb R, Mai S. Distinct and shared three-dimensional chromosome organization patterns in lymphocytes, monoclonal gammopathy of undetermined significance and multiple myeloma. *Int J Cancer* 2017; **140**: 400–410.
- 380 Wang H, Nakamura M, Abbott TR, Zhao D, Luo K, Yu C *et al.* CRISPR-mediated live imaging of genome editing and transcription. *Science (80-)* 2019; **365**: 1301–1305.
- 381 Iarovaia O V., Rubtsov M, Ioudinkova E, Tsfasman T, Razin S V., Vassetzky YS. Dynamics of double strand breaks and chromosomal translocations. *Mol Cancer* 2014; **13**: 249.
- 382 Ramsden DA, Nussenzweig A. Mechanisms driving chromosomal translocations: lost in time and space. *Oncogene* 2021; **40**: 4263–4270.
- 383 Kantidze OL, Razin S V. Chromatin loops, illegitimate recombination, and genome evolution. *BioEssays*. 2009; **31**. doi:10.1002/bies.200800165.
- 384 Pannunzio NR, Watanabe G, Lieber MR. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J. Biol. Chem.* 2018; **293**: 10512–10523.
- 385 Sall FB, Germini D, Shmakova A, Diouf PMD, Wiels J, Ndour M *et al.* S-2. Role of epstein-barr virus zebra protein in induction of t(8;14) translocation. *Biopolym Cell* 2019; **35**. doi:10.7124/bc.0009F6.
- 386 Hu J, Meyers RM, Dong J, Panchakshari RA, Alt FW, Frock RL. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc* 2016; **11**. doi:10.1038/nprot.2016.043.
- 387 Löbrich M, Jeggo P. A Process of Resection-Dependent Nonhomologous End Joining Involving the Goddess Artemis. *Trends Biochem Sci* 2017; **42**: 690–701.
- 388 Viushkov VS, Lomov NA, Rubtsov MA, Vassetzky YS. Visualizing the Genome: Experimental Approaches for Live-Cell Chromatin Imaging. *Cells* 2022; **11**. doi:10.3390/cells11244086.
- 389 Luzhin A V., Velichko AK, Razin S V., Kantidze OL. Automated Analysis of Cell Cycle Phase-Specific DNA Damage Reveals Phase-Specific Differences in Cell Sensitivity to Etoposide. *J Cell Biochem* 2016. doi:10.1002/jcb.25615.
- 390 Ivashkevich AN, Martin OA, Smith AJ, Redon CE, Bonner WM, Martin RF *et al.* γ H2AX foci as a measure of DNA damage: A computational approach to automatic analysis. *Mutat Res - Fundam Mol Mech Mutagen* 2011; **711**. doi:10.1016/j.mrfmmm.2010.12.015.
- 391 Reddig A, Roggenbuck D, Reinhold D. Comparison of different immunoassays for γ H2AX quantification. *J Lab Precis Med* 2018; **3**. doi:10.21037/jlpm.2018.09.01.

Благодарности

В первую очередь я выражаю благодарность научному руководителю, Михаилу Александровичу Рубцову, за его роль наставника: за помощь в постановке задач и подборе материалов для исследования, за всестороннюю поддержку, а также за возможность работать на кафедре молекулярной биологии Биологического факультета МГУ. Также я благодарен Сергею Владимировичу Разину и Егору Сергеевичу Васецкому за возможность выполнения некоторых этапов исследования в Институте биологии гена РАН и в институте Gustave-Roussy во Франции. Отдельная благодарность за содействие при выполнении работы и за рабочую атмосферу моим коллегам из лаборатории молекулярной биологии МГУ, в особенности Владимиру Вьюшкову.

Я хочу поблагодарить Данила Алексеевского за его программу для анализа изображений с конфокального микроскопа, Сергея Ульянова и Алексея Гаврилова за помощь в подготовке 4С-библиотек, Артема Артемова за обработку данных 4С, Татьяну Цфасман за всестороннюю помощь во время получения клеточных линий, Вадима Покровского и Дарину Соколову за помощь в работе с химиотерапевтическими препаратами и других коллег, которые помогли мне во время исследования. Также я выражаю благодарность коллективу преподавателей кафедры молекулярной биологии за знания и навыки, которые я получил во время учебы.

Я благодарен своим родным, близким и друзьям, которые поддерживали меня во время работы над этим диссертационным исследованием.

Отдельно я хочу выразить благодарность Ольге Владимировне Яровой, Андрею Борисовичу Сударикову и Андрею Александровичу Замятнину за согласие выступить в роли официальных оппонентов этой диссертации.