

Федеральное государственное бюджетное учреждение науки
ИНСТИТУТ БИООРГАНИЧЕСКОЙ ХИМИИ
имени академиков М.М. Шемякина и Ю.А. Овчинникова
Российской академии наук

На правах рукописи

Путинцева Екатерина Викторовна
**Разнообразие репертуаров Т-клеточных рецепторов человека
и его изменения в ходе старения**

Специальность – 03.01.03 – молекулярная биология

Диссертация на соискание ученой степени
кандидата биологических наук

Научный руководитель:
Доктор биологических наук
Чудаков Дмитрий Михайлович

Москва 2014

Оглавление

| | |
|--|----|
| 1. Введение | 4 |
| 2. Обзор литературы..... | 6 |
| 2.1. Т-клеточный иммунитет и репертуар TCR..... | 6 |
| 2.1.1. Развитие Т-лимфоцитов в тимусе..... | 6 |
| 2.1.2. Генерация TCR..... | 9 |
| 2.1.3. Участие $\alpha\beta$ Т-лимфоцитов в адаптивном иммунном ответе | 13 |
| 2.1.4. Взаимодействие TCR и MHC..... | 15 |
| 2.1.5. Разнообразие репертуаров TCR | 16 |
| 2.1.6. Общественные TCR клонотипы..... | 17 |
| 2.1.7. Старение Т-клеточного иммунитета человека..... | 20 |
| 2.2. Методы исследования разнообразия индивидуальных репертуаров Т-клеточных рецепторов..... | 23 |
| 2.2.1. Проточная цитометрия..... | 23 |
| 2.2.2. Спектротипирование | 23 |
| 2.2.3. Гибридизация..... | 26 |
| 2.2.4. Секвенирование нуклеотидных последовательностей Т-клеточных рецепторов | 26 |
| 2.2.5. Ошибки ПЦР и секвенирования – искусственное разнообразие | 32 |
| 3. Цели и задачи | 35 |
| 4. Материалы и методы | 36 |
| 4.1. Оборудование и расходные материалы..... | 36 |
| 4.2. Реактивы..... | 36 |
| 4.3. Методы..... | 38 |
| 4.3.1. Сбор образцов..... | 38 |
| 4.3.2. Выделение РВМС..... | 38 |
| 4.3.3. Выделение тотальной РНК из клеток | 39 |
| 4.3.4. Синтез кДНК..... | 39 |
| 4.3.5. Первый раунд ПЦР-амплификации..... | 40 |
| 4.3.6. Второй раунд ПЦР-амплификации | 41 |
| 4.3.7. Секвенирование..... | 42 |
| 4.3.8. Анализ результатов массивного секвенирования..... | 43 |
| 4.3.9. Расчет дивергенции Дженсена-Шеннона между распределениями V β -сегментов | 43 |

| | |
|--|-----------|
| 4.3.10. Оценка нижней границы разнообразия репертуара CDR3 TCR β периферической крови..... | 44 |
| 4.3.11. Проточная цитометрия | 44 |
| 4.3.12. HLA-типирование..... | 45 |
| 5. Результаты и их обсуждение..... | 46 |
| 5.1. Сравнительный анализ репертуаров TCR β мам и детей | 46 |
| 5.1.1. Распределение частот использования V β сегментов..... | 47 |
| 5.1.2. Пересечения репертуаров TCR β CDR3 в родственных и неродственных парах мама - ребенок | 56 |
| 5.1.3. Одинаковые аминокислотные последовательности CDR3 чаще сочетаются с одинаковым V β -сегментом в родных парах доноров..... | 58 |
| 5.1.4. Селекция в тимусе снижает среднюю длину CDR3..... | 60 |
| 5.1.5. В поисках зрелых микрохимерных клонов Т-клеток..... | 61 |
| 5.1.6. Оценка пересечения полных индивидуальных TCR β репертуаров | 66 |
| 5.2. Разработка метода нормированного анализа репертуаров TCR с использованием молекулярного баркодирования..... | 70 |
| 5.3. Исследование изменений репертуара TCR β в ходе старения человека..... | 75 |
| 5.3.1. Индивидуальное разнообразие TCR β CDR3 падает в течение всей жизни человека..... | 75 |
| 5.3.2. Определение нижней границы индивидуального разнообразия TCR β CDR3..... | 79 |
| 5.3.3. Заполнение гомеостатического пространства активированными клонами Т-клеток..... | 81 |
| 5.3.4. Корреляция между разнообразием TCR β и долей наивных Т-клеток в периферической крови..... | 82 |
| 5.3.5. Образцы периферической крови долгоживущих доноров характеризуются высоким содержанием CD4 ⁺ наивных Т-клеток..... | 85 |
| 5.3.6. Возрастные изменения относительной представленности V β - и J β -сегментов | 86 |
| 5.3.7. Общественные клонотипы и старение | 88 |
| 6. Выводы | 91 |
| 7. Заключение | 93 |
| 8. Список сокращений..... | 94 |
| 9. Список литературы | 96 |

1. Введение

Эффективность и сбалансированность адаптивной иммунной защиты организма от инфекционных и онкологических заболеваний в значительной степени определяется функциональным и рецепторным разнообразием Т-лимфоцитов [1], [2]. Каждый клон Т-лимфоцитов кодирует уникальный Т-клеточный рецептор (TCR), потенциально способный распознать чужеродный антиген в составе молекулы главного комплекса гистосовместимости (МНС), и определяющий специфичность адаптивного иммунного ответа. Высокое индивидуальное разнообразие наивного репертуара TCR является залогом того, что для новой инфекции или онкологического заболевания найдутся клоны Т-лимфоцитов, специфично распознающие антигены, характерные для данного патогена или патологических клеток. Антиген-специфичные Т-лимфоциты активно размножаются и присутствуют в организме и периферической крови в повышенных концентрациях, причем клетки памяти могут сохраняться в течение десятков лет [3], [4]. Таким образом, индивидуальный репертуар Т-клеточных рецепторов определяет эффективность иммунной защиты, диапазон распознаваемых антигенов, особенности патологических состояний иммунитета и, более того, содержит в себе потенциально читаемую информацию о многих инфекционных, онкологических, и аутоиммунных заболеваниях данного пациента [1], [5].

С развитием современных методов массированного секвенирования репертуаров TCR [6], [7], иммунофенотипирования Т-клеточных субпопуляций [8], транскриптомики [9], в том числе с единичных клеток [10], статистического анализа репертуаров TCR [11], и их потенциальных эпитопов в составе МНС [12], а также методов рентгеноструктурного анализа комплексов TCR/пептид/МНС [13] и масс-спектрометрического анализа репертуаров презентруемых пептидов [14], исследования Т-клеточного иммунитета постепенно выходят на более системный уровень [15] [16], приближающий нас к созданию нового поколения методов диагностики и терапии ряда заболеваний, методов вакцинации, а также методов восстановления и поддержания эффективной функции адаптивного иммунитета.

Настоящая работа посвящена анализу индивидуальных репертуаров TCR β методом глубокого массированного секвенирования. В первой части работы

проводится сравнительный анализ состава и структуры индивидуальных репертуаров TCR β родственных и неродственных пар мама-ребенок. Вторая часть работы посвящена разработке методики нормализации библиотек генов TCR для корректного воспроизводимого межиндивидуального сравнения разнообразия репертуаров. В третьей части работы с помощью методики, разработанной во второй части, исследуются изменения, происходящие в репертуаре TCR β человека, связанные со старением.

2. Обзор литературы

Данный раздел состоит из двух частей. В первой части рассматривается общее строение и формирование индивидуального репертуара Т-клеточных рецепторов. Мы начинаем с освещения основных этапов развития Т-лимфоцитов в тимусе и генерации Т-клеточных рецепторов в ходе соматической рекомбинации. Затем кратко затрагивается роль Т-лимфоцитов в адаптивном иммунном ответе и взаимодействие Т-клеточного рецептора (TCR, англ. T-cell receptor) с комплексом пептид/МНС. После этого приводятся примерные оценки разнообразия индивидуального репертуара человека, существующие на сегодняшний день. В заключительной части этого раздела описываются такие характеристики Т-клеточного иммунитета, как общественные клонотипы и старение репертуара TCR.

Вторая часть обзора литературы посвящена основным существующим методам исследования репертуара Т-клеточных рецепторов. В перечисленные методы входят проточная цитометрия, гибридизация, спектротипирование, а также секвенирование нуклеотидных последовательностей Т-клеточных рецепторов. Основной упор делается на методы массивированного секвенирования (или секвенирования нового поколения – NGS, англ. next generation sequencing), позволившие вывести изучение репертуаров Т-клеточных рецепторов на новый уровень. Далее описываются проблемы и ограничения, связанные с использованием в качестве стартовой матрицы ДНК или РНК, а также проблема возникновения искусственного разнообразия и исправления ошибок полимеразной цепной реакции (ПЦР) и секвенирования.

2.1. Т-клеточный иммунитет и репертуар TCR

2.1.1. Развитие Т-лимфоцитов в тимусе

Предшественниками Т-клеток являются плюрипотентные гемопоэтические стволовые клетки, мигрировавшие из костного мозга в тимус, где происходит их дальнейшее созревание. На момент прибытия в тимус у клеток-предшественников отсутствуют поверхностные маркеры зрелых Т-клеток, и гены их рецепторов еще не рекомбинировали. Такие тимоциты называются *двойными отрицательными* (англ.

DN, double negative), поскольку они не экспрессируют на своей поверхности молекулы CD4 и CD8, определяющие специфичность их взаимодействия с другими клетками. DN-стадия развития тимоцитов состоит из четырех фаз (DN1-DN4), отличающихся друг от друга паттерном экспрессии поверхностных маркеров (Рисунок 2.1). После фазы DN4 тимоциты становятся *двойными позитивными* (англ. DP, double positive), т.е. экспрессируют на своей поверхности обе молекулы CD4 и CD8, и подвергаются дальнейшему отбору в тимусе [17].

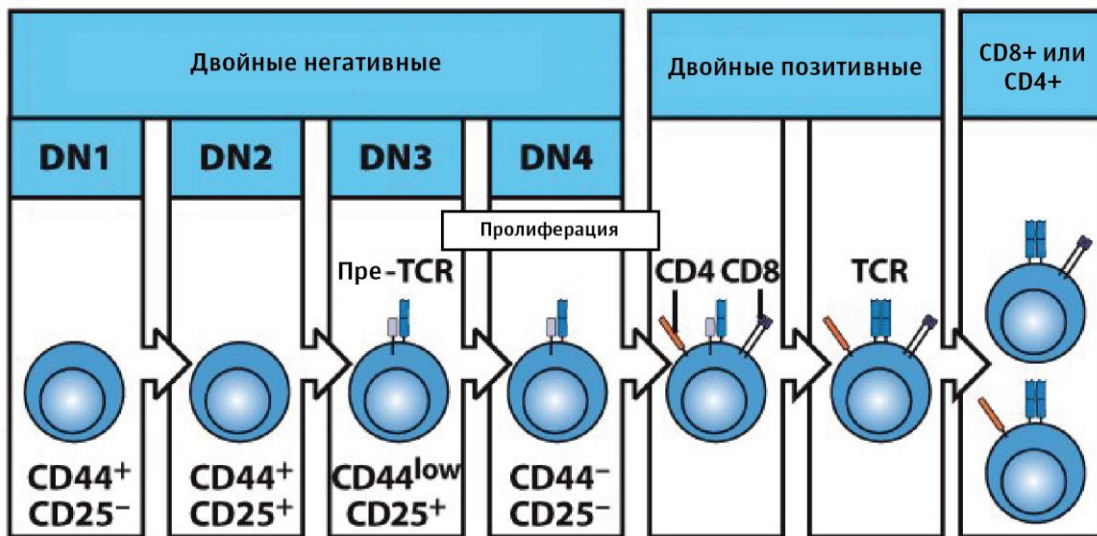


Рисунок 2.1 [17]. Стадии развития α : β Т-лимфоцитов в тимусе.

Ключевой характеристикой Т-лимфоцита является его рецептор к антигену – *Т-клеточный рецептор* (англ. TCR, T-cell receptor), образующийся в ходе созревания лимфоцита в тимусе. Он представляет собой трансмембранный белок, находящийся на поверхности Т-лимфоцитов. Структурно TCR состоит из двух цепей и в зависимости от структуры этих цепей бывает двух видов: α : β TCR в случае α и β цепей или γ : δ TCR в случае γ и δ цепей. Т-клетки, TCR которых состоит из соответствующих цепей, называются α : β или γ : δ Т-лимфоциты.

Каждая цепь TCR состоит из вариабельного и константного доменов, за которыми следует трансмембранный домен и короткий цитоплазматический хвост (Рисунок 2.2). Вариабельный домен β -цепи состоит из V- (англ. variable), D- (англ. diversity), и J- (англ. joining) участков; α -цепи – из V-, и J- участков.

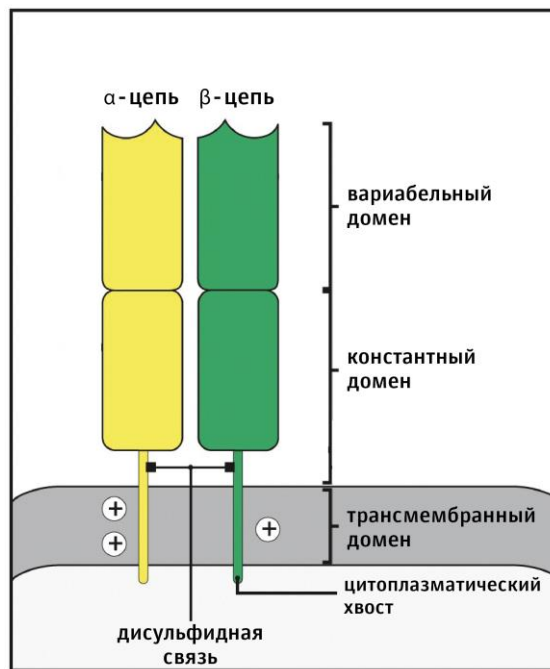


Рисунок 2.2 [17]. Структура $\alpha:\beta$ TCR.

Большая часть DN тимоцитов в результате развития становится $\alpha:\beta$ T-клетками, меньшая – $\gamma:\delta$ T-клетками. В процессе развития DN тимоциты начинают рекомбинацию генов γ , δ и β цепей одновременно, поэтому существует два сценария развития событий (Рисунок 2.3) [17]:

1) В том случае, если комплекс γ и δ цепей сформирован до образования удачной β -цепи, спаривающейся на этом этапе с замещающей α -цепью (см. ниже), тимоцит получает сигнал через $\gamma:\delta$ -рецептор, приводящий к остановке реорганизации генов β -цепи и превращению его в $\gamma:\delta$ T-лимфоцит;

2) В том случае, если готовая функциональная β -цепь сформирована до удачной комбинации γ и δ цепей, она спаривается с *замещающей* α -цепью предшественника TCR, что приводит к его полной сборке и формированию его комплекса с молекулой CD3. Через сформированный комплекс тимоцит получает сигнал, приводящий к остановке каких-либо дальнейших рекомбинаций β , γ и δ цепей, а также запуску фазы пролиферации. Эта фаза совпадает с последней стадией DN4 тимоцитов (Рисунок 2.1). После пролиферации начинается экспрессия молекул CD4 и CD8 на поверхности тимоцита, а также рекомбинация генов α -цепи. Перестройка в локусе α -цепи приводит к удалению генов δ -цепи, препятствуя таким образом экспрессии молекул γ - и δ - цепей на поверхности той же клетки. Когда вариант α -цепи,

подходящий для β -цепи тимоцита, готов, $CD3^{low}CD4^{+}CD8^{+}$ тимоцит переходит к стадии позитивной селекции в тимусе.

Процесс *позитивной* селекции в тимусе удаляет Т-лимфоциты, TCR которых не способны узнавать комплекс собственный пептид-МНС. Процесс *негативной* селекции удаляет Т-лимфоциты, TCR которых слишком сильно взаимодействует с комплексом собственный пептид-МНС, таким образом, являясь потенциальным источником развития аутоиммунного ответа [18].

2.1.2. Генерация TCR

Репертуар TCR настолько велик, что позволяет Т-лимфоцитам распознавать огромное разнообразие антигенов, не встречавшихся иммунной системе хозяина ранее. Однако закодировать достаточное количество вариантов TCR в геноме в готовом виде невозможно, поэтому для генерации достаточного разнообразия используется соматическая рекомбинация V(D)J участков.

В своей исходной конфигурации последовательность генов V, D, J и C сегментов не кодирует функциональных белков. Для их превращения в функциональные гены необходима сайт-специфическая ДНК-рекомбинация. Перестройки ДНК осуществляются по консервативным некодирующим последовательностям ДНК, расположенным в местах предполагаемой рекомбинации. Эти последовательности называются сигнальными последовательностями рекомбинации (или RSS последовательностями, англ. recombination signal sequence) и состоят из консервативного блока из семи нуклеотидов (септамера) – 5'-CACAGTG-3', – всегда следующего за кодирующей последовательностью, а также неконсервативного участка (так называемого *разделителя*), состоящего из 12 или 23 пар нуклеотидов. После разделителя находится еще один консервативный участок из девяти нуклеотидов (нонамер) – 5'-ACAАААСС-3'.

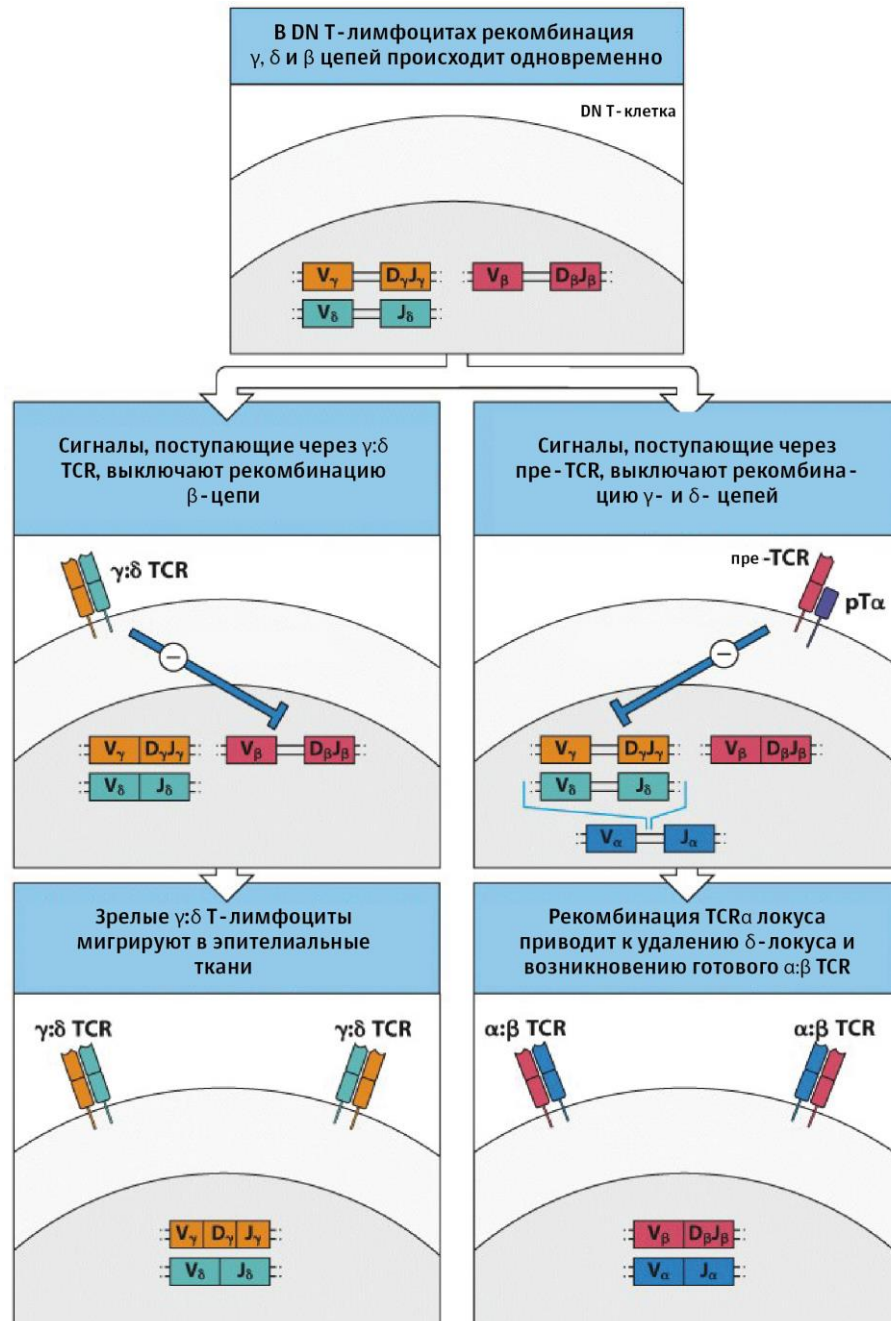


Рисунок 2.3 [17]. Определение судьбы T-лимфоцита сигналом, поступающим через γ : δ TCR или через пре- α : β TCR.

Последовательность разделителей различается, однако количество нуклеотидов постоянно и примерно соответствует либо одной (12 нуклеотидов), либо двум (23 нуклеотида) спиральям ДНК. Таким образом, септамер и нонамер всегда оказываются по одну сторону ДНК-спирали, где они могут быть связаны комплексами белков, осуществляющих катализ рекомбинации. Участок гена, фланкированный разделителем, состоящим из 23 пн, как правило, может быть присоединен только к последовательности, фланкированной разделителем из 12 пн. Сборка генов

происходит в несколько шагов с использованием ферментов, кодируемых генами, активирующих рекомбинацию (гены RAG-1 и RAG-2, англ. recombination activation genes). Ферменты вносят разрывы в ДНК в местах сигнальных последовательностей рекомбинации (RSS), далее происходит вырезание лишних фрагментов.

На следующей стадии формирования последовательности TCR происходит самолигирование двуцепочечных концов ДНК в месте разрыва. В случайном месте образовавшейся шпильки (однако недалеко от места лигирования) затем вносится одноцепочечный разрыв. Если внесение разрыва не совпадает с точкой лигирования, образуется две одноцепочечных палиндромных последовательности [19].

После этого происходит добавление случайных нуклеотидов терминальной дезоксирибонуклеотид трансферазой. Склонность этого фермента к добавлению G-нуклеотидов приводит к частому возникновению глицина (в кодонах которого преобладает гуанидин) [20]. Было выдвинуто предположение о том, что встраивание такой маленькой аминокислоты, как глицин, на стыке сегментов увеличивает пластичность CDR3-петли [21] (см. ниже).

Помимо добавления случайных нуклеотидов, существует также экзонуклеазная активность, направленная на концы разрывов. Механизм удаления нуклеотидов изучен хуже остальных этапов рекомбинации. Однако было замечено, что удаление нуклеотидов на стыках VD и VJ генов β -цепи происходит независимо (то есть большее количество удаленных букв в районе одного из стыков не приводит к меньшему количеству удаленных букв в районе другого) [11]. Кроме того, концевые последовательности, богатые G/C-нуклеотидами, оказываются более устойчивыми к экзонуклеазной активности, чем последовательности, богатые A/T-нуклеотидами [22].

Последняя стадия соматической рекомбинации TCR заключается в достраивании комплементарных цепей и лигировании разрывов [17].

В каждой клетке есть две копии TCR β локуса – по одной на каждой из двух аллелей. Однако за счет механизма аллельного исключения на одну клетку образуется лишь одна функциональная TCR β цепь [23]. В случае β -цепи сначала происходит рекомбинация генов D- и J-сегментов, а затем генов V-сегмента (Рисунок 2.4). При этом рекомбинация D β -J β происходит параллельно на обеих хромосомах, а V β -D β – сначала на одной и затем (в случае неудачной рекомбинации на первой) на другой

[23]. Теоретически рекомбинация β -цепи в каждой клетке может происходить вплоть до четырех раз, поскольку в каждом из двух локусов генов TCR β существует 2 сета генов D- и J- сегментов.

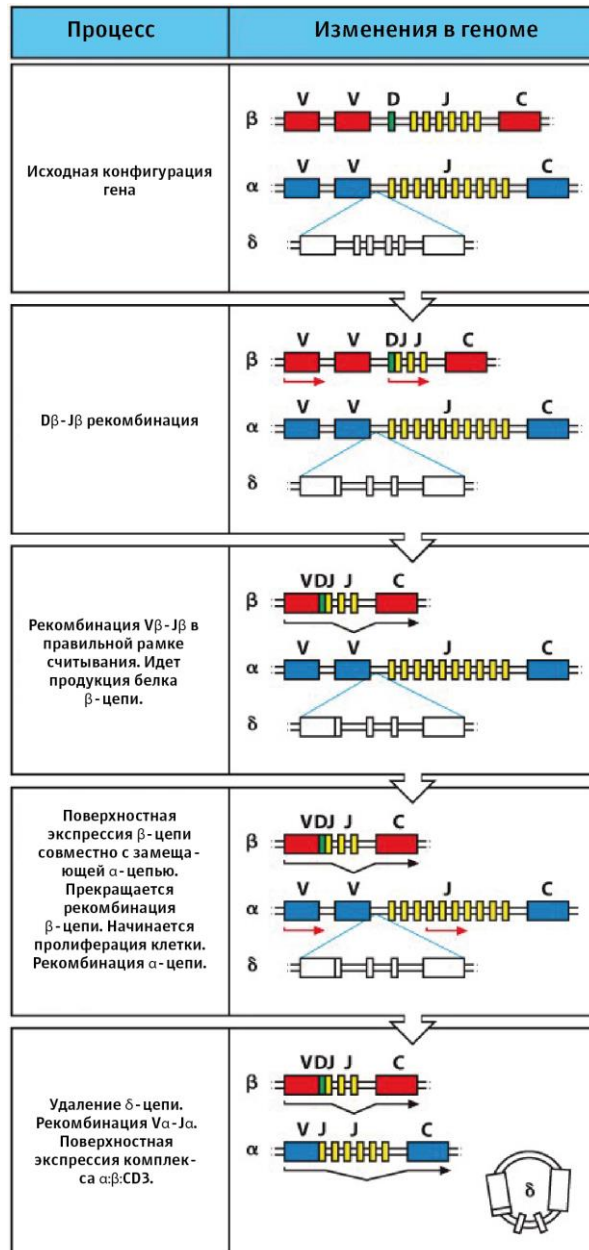


Рисунок 2.4 [17]. Стадии соматической рекомбинации генов α : β Т-лимфоцитов.

В случае α -цепи, благодаря большому количеству генов сегментов V α и J α (см. ниже), рекомбинация может происходить много раз, поэтому удачный вариант цепи образуется почти всегда. Неудачно рекомбинировавшие гены (то есть гены со сбитой рамкой считывания) продолжают экспрессироваться в выживших клетках, однако их

мРНК уничтожается с помощью механизма нонсенс-опосредованного разрушения мРНК (NMD, англ. nonsense-mediated decay). Этот механизм выполняет функцию контроля качества транскриптов: мРНК с преждевременным стоп-кодоном быстро разрушается в процессе трансляции, чтобы предотвратить синтез нефункциональных белков [24], [25].

В геноме человека локус генов α цепи TCR (*TRA*) находится на 14 хромосоме и включает в себя 47 генов $V\alpha$ -сегмента, 57 генов $J\alpha$ -сегмента и один $C\alpha$ -сегмент. Локус генов β цепи TCR (*TRB*) находится на 7 хромосоме человека и состоит из 54 генов V -сегмента, 2 генов D -сегмента, 13 генов J -сегмента и 2 генов C -сегмента. Количество возможных комбинаций, которое можно получить при разных сочетаниях этих генов составляет 7 522 632 [26]. Однако это лишь малая толика всего разнообразия вариантов Т-клеточных рецепторов, формирующегося в ходе случайной делеции и достройки нуклеотидов в процессе рекомбинации (см. выше).

2.1.3. Участие $\alpha\beta$ Т-лимфоцитов в адаптивном иммунном ответе

Как только Т-лимфоциты завершают свое развитие в тимусе, они покидают его и попадают в кровяное русло. Достигая периферических лимфатических узлов, они мигрируют из кровяных сосудов, затем снова возвращаясь в них по лимфатическим сосудам. Таким образом происходит постоянная рециркуляция зрелых Т-клеток между периферическими лимфатическими органами и кровяным руслом.

Зрелые Т-лимфоциты, еще не встретившие своего специфического антигена, называются *наивными* Т-лимфоцитами. Для участия в адаптивном иммунном ответе такие лимфоциты должны встретиться с узнаваемым ими пептидом в составе молекулы главного комплекса гистосовместимости (англ. МНС, major histocompatibility complex), расположенной на поверхности антиген-презентирующей клетки. При таком взаимодействии наивный Т-лимфоцит получает сигнал к пролиферации и дифференциации в клетки памяти и так называемые *эффекторные* Т-клетки, обладающие способностью прямого или опосредованного удаления антигена из организма хозяина.

Постоянная рециркуляция наивных Т-клеток через кровяное русло в лимфатические узлы, селезенку и лимфатические ткани слизистой – и затем обратно в кровь – позволяет им взаимодействовать со множеством дендритных антиген-

презентирующих клеток в лимфатических тканях каждый день и сканировать комплексы пептид-МНС на их поверхности. Таким образом, повышается вероятность встречи любой Т-клетки с антигеном патогена, начавшего воспалительный процесс в любой части организма. Наивные Т-лимфоциты, не повстречавшие специфический антиген, покидают лимфоидную ткань через эфферентные лимфатические сосуды, вновь поступая в кровяное русло и продолжая рециркуляцию. Однако при встрече со специфическим антигеном на поверхности взрослой дендритной клетки наивный Т-лимфоцит прекращает миграцию. В течение нескольких дней происходит его пролиферация, приводящая к клональной экспансии и дифференциации. В результате образуются эффекторный клоны Т-клеток, обладающие идентичной специфичностью к антигену. После окончания периода деления эффекторные Т-клетки выходят через лимфатические сосуды в кровяное русло и направляются к очагу воспаления. Исключением является рециркуляция лимфоцитов в селезенке, которая не связана с лимфатической системой – все клетки попадают в селезенку через кровь и выходят из нее непосредственно в кровь.

После узнавания антигена наивные Т-лимфоциты дифференцируются в разные функциональные классы эффекторных Т-клеток, после чего начинается их активная пролиферация. $CD8^+$ наивные Т-лимфоциты, распознающие антигены в комплексе с МНСI, дифференцируются в цитотоксические Т-клетки, способные распознавать и убивать зараженные клетки. $CD4^+$ Т-лимфоциты, распознающие антигены в составе МНСII, могут дифференцироваться в различные типы эффекторных клеток. Основными выделяемыми субпопуляциями $CD4^+$ Т-лимфоцитов на данный момент являются Th1 клетки (активирующие зараженные макрофаги, а также стимулирующие продукцию антител В-клетками), Th2 клетки (стимулирующие продукцию антител В-клетками и переключение классов антител), Th17 клетки (способствующие поступлению нейтрофилов в очаги заражения на ранних этапах адаптивного иммунного ответа), а также несколько субпопуляций регуляторных Т-лимфоцитов (наоборот, подавляющих иммунный ответ) [27].

Первая встреча Т-лимфоцита со специфичным антигеном, в ходе которой происходит его дифференциация и пролиферация, называется *первичным иммунным ответом*. В результате первичного иммунного ответа генерируется иммунологическая память, обеспечивающая намного более быструю индукцию

специфических эффекторных Т-лимфоцитов при повторной встрече с данным антигеном. Иммунный ответ, развивающийся при повторной встрече с антигеном, называется *вторичным иммунным ответом*. Иммунная память Т-клеточного ответа обеспечивается Т-лимфоцитами памяти, генерация которых происходит в момент первичного иммунного ответа. Время жизни таких клеток сопоставимо с продолжительностью жизни как у человек [3], так и у мыши [28].

2.1.4. Взаимодействие TCR и МНС

Лигандами TCR являются короткие пептиды в составе МНС, кодируемого генами человеческого лейкоцитарного антигена (англ. HLA, Human Leucocyte Antigens) в человеке. Эффективное взаимодействие комплекса пептид-МНС с TCR приводит к запуску трансмембранной TCR-опосредованной передачи сигнала и последующих внутриклеточных биохимических каскадов, приводящих к пролиферации, дифференциации, активации эффекторных функций или запуску механизмов клеточной смерти.

После трансляции подходящие друг другу α и β цепи формируют гетеродимер с образованием дисульфидных связей между константными участками этих цепей (Рисунок 2.2). Полученный функциональный гетеродимер затем гликозилируется, доставляется к клеточной мембране и встраивается в нее. Экспонированный конец α - β TCR гетеродимера имеет шесть петель, соответствующих шести высоко варьируемым CDR (англ. complementarity determining region) участкам, отвечающим за специфичное взаимодействие с комплексом пептид-МНС. CDR1 и CDR2 петли кодируются генами V_α / V_β . CDR3 петли, соответствующие гипервариабельному участку соматической рекомбинации, находятся на стыке V(D)J сегментов. Структурные исследования показали, что CDR1 и CDR2 петли преимущественно взаимодействуют с поверхностью комплекса МНС, а CDR3 петли взаимодействуют с экспонированным участком пептида в составе молекулы МНС [29]. TCR связывается с МНС, как правило, под углом к углеводородному остову пептида, зажато между двумя параллельными α -спиралями МНС (Рисунок 2.5). Угол связывания при этом варьируется от 20 до 70 градусов, в зависимости от структуры взаимодействующих молекул. Расположение α - и β - цепей TCR относительно связанного МНС пептида, судя

по структурам уже изученных TCR, высоко консервативно – CDR3 α петля взаимодействует с N-концом пептида, а CDR3 β петля – с его C-концом.

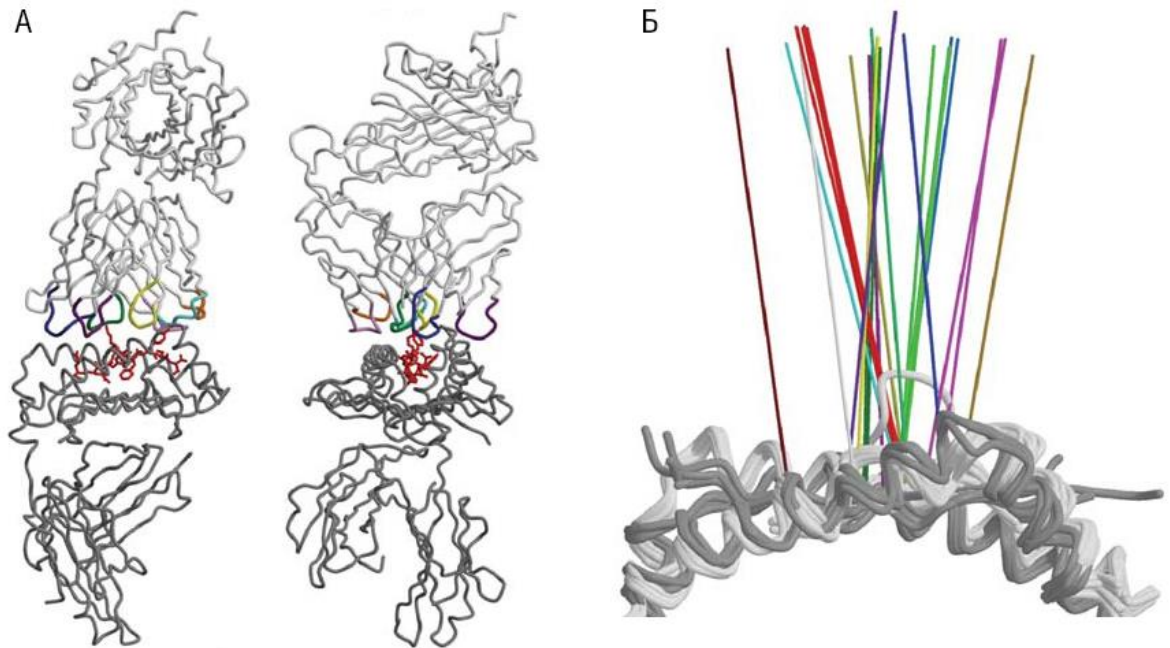


Рисунок 2.5 [29]. **Взаимодействие комплекса пептид-МНС и $\alpha:\beta$ TCR.** (А) Взаимодействие комплекса пептид(красный)-МНС(темно-серый) и $\alpha:\beta$ TCR (константный участок – светло-серый). Петли TCR изображены разными цветами: CDR1 α – темно-синий, CDR2 α – малиновый, CDR3 α – зеленый, CDR1 β – голубой, CDR2 β – розовый, CDR3 β – желтый. (Б) Ориентация 15 разных $\alpha:\beta$ TCR относительно МНС (серый) показана сдвоенными осями (соответствующими V α и V β) разных цветов.

2.1.5. Разнообразие репертуаров TCR

Математически, суммарное потенциальное разнообразие репертуара $\alpha:\beta$ TCR, с учетом V(D)J рекомбинации и инсерции/делеции случайных нуклеотидов, составляет $>10^{15}$ уникальных последовательностей $\alpha\beta$ TCR для мышей и $>10^{18}$ – для людей [30].

Учитывая тот факт, что всего лишь около 3% Т-лимфоцитов переживает селекцию в тимусе, потенциальное разнообразие TCR в периферической крови можно оценить в порядка 10^{13} вариантов для мышей и 10^{16} – в человеческой популяции [31].

Потенциальное индивидуальное разнообразие ограничивается общим количеством Т-лимфоцитов в организме – 10^8 для мыши [32] и 10^{12} для человека [33].

Эти цифры, тем не менее, на много порядков превышают посчитанную до нашей работы нижнюю границу разнообразия TCR у одного организма – 10^6 у мыши [34] и 10^7 у человека [35].

Наблюдаемая огромная разница в потенциальном и реальном разнообразии репертуара TCR привела к появлению гипотезы о стохастической манере его генерации. Так, в иммунной системе каждого индивидуума генерируется некий случайный набор последовательностей из общего пула потенциального разнообразия. Однако в ранние 1990-е годы эта гипотеза была опровергнута исследованиями, в ходе которых в периферической крови человека были обнаружены так называемые *общественные* TCR [36], [37]. В этих исследованиях в крови разных индивидуумов были найдены идентичные или очень похожие по последовательности TCR популяции антиген-специфических Т-клеток памяти. С тех пор эти данные были подтверждены множеством работ.

2.1.6. Общественные TCR клонотипы

Поскольку выживание и селекция из огромного репертуара определенных TCR, идентичных в крови разных доноров, требует невероятного количества совпадающих этапов, существование общественных TCR в Т-клетках памяти считалось весьма маловероятным. Так, для их появления было бы необходимо, чтобы некоторые TCR:

- 1) появились в результате соматической рекомбинации из огромного числа возможных вариантов, создающихся в тимусе;
- 2) прошли отбор в тимусе;
- 3) выжили на периферии;
- 4) их наивные предшественники обладали достаточно высокой частотой встречаемости в крови, чтобы эффективно соревноваться с остальным доступным репертуаром TCR, специфичным к данному антигену.

Если бы появление всех вариантов TCR было равновероятным, можно было бы ожидать лишь очень редкие совпадения в репертуарах разных индивидуумов – и при этом, скорее всего, в ответ на определенный специфический антиген. Но общественные TCR, тем не менее, найдены в разнообразных по специфичности иммунных ответах, внутри разных видов [37]–[47]. Более того, высокий уровень

пересечения репертуаров TCR **наивных** Т-лимфоцитов был обнаружен как у мыши [48], [49], так и у человека [50], [51].

Для объяснения явления общественных клонотипов существует несколько гипотез:

1) Существует ряд объяснений наличия общественных TCR, основанных на особенностях их пространственной **структуры** [52]. Одно из них предполагает, что структура комплекса пептид-МНС может приводить к неравномерности репертуара TCR – как формируемого в ходе позитивной и негативной селекции в тимусе, так и в иммунном ответе Т-лимфоцитов, тем самым создавая благоприятные условия для экспансии общественных TCR [52]–[54]. Альтернативная версия, напротив, говорит о структуре самих общественных TCR (а не комплекса пептид-МНС), как о детерминанте увеличения их частоты встречаемости [55]. Однако закономерности в пространственной структуре всех известных общественных TCR или специфичных им комплексов пептид-МНС найдено не было. Кроме того, еще **до** возможности структурного отбора TCR, существует необходимость появления их в репертуаре многих индивидуумов.

2) Другая группа объяснений опирается на предположение, что определенные, наиболее вероятные, рекомбинационные события могут в точности воспроизводиться как у разных индивидуумов, так и в пределах одного организма. Например, более частое возникновение в ходе соматической рекомбинации определенных нуклеотидных последовательностей TCR может происходить за счет их более простой организации (меньшего количества добавленных нуклеотидов, менее укороченных генов V-, (D-) и J- сегментов) [50], [51] или за счет повторения каких-либо событий рекомбинации (например, добавления одного или нескольких остатков гуанидина – цепочки из которых можно часто наблюдать в общественных TCR [56]). Однако это не объясняет независимого возникновения в периферической крови разных индивидуумов Т-лимфоцитов с идентичной аминокислотной последовательностью TCR, но не совпадающей при этом нуклеотидной последовательностью [40], [45], [47].

3) Третьим объяснением является *конвергентная рекомбинация* [57] - то есть конвергенция многих различных по структуре событий V(D)J-рекомбинации, производящая идентичные нуклеотидные последовательности TCR, а также конвергенция нескольких нуклеотидных последовательностей TCR, создающая идентичные аминокислотные последовательности TCR. Была найдена сильная положительная корреляция между частотой встречаемости последовательностей аминокислот и количеством разных сочетаний нуклеотидов, кодирующих эти аминокислоты [57]. Кроме того, было показано, что конвергентная рекомбинация формирует как межиндивидуальный, так и индивидуальный репертуар TCR β , – частота встречаемости конвергентных последовательностей TCR β наивных CD8+ клеток выше как среди общественных клонотипов, так и внутри индивидуального репертуара мышей [51].

Существующие на сегодняшний день данные указывают на основной вклад в повторение одних и тех же последовательностей TCR в разных индивидуумах со стороны как воспроизведения идентичных по структуре событий соматической рекомбинации [50], [58], так и конвергентной рекомбинации [57]–[59]. Несмотря на то, что конвергентная рекомбинация дает статистически достоверные предсказания перекрытия индивидуальных репертуаров, менее половины последовательностей репертуаров TCR CD4⁺CD8⁺ клеток может быть отнесено к случайной конвергентной рекомбинации [60]. Кроме того, существуют такие последовательности TCR, вероятность воспроизведения которых в ходе конвергентной рекомбинации очень велика, но реальная частота встречаемости оказывается сильно ниже предсказанной [49], [57], [59], [61]. Такие случаи также указывают на вклад неравномерности частот событий в ходе соматической рекомбинации.

Несмотря на то, что в ходе селекции в тимусе происходит резкое снижение разнообразия TCR, считается, что она практически не вносит вклада в формирование репертуаров общественных TCR [62], [63].

2.1.7. Старение Т-клеточного иммунитета человека

Процессы, происходящие в стареющей иммунной системе, носят название иммунной недостаточности и поражают как врожденный, так и адаптивный иммунитет [64], [65].

В частности, снижается эффективность Т-клеточного иммунитета. Снижается способность адаптивного иммунитета генерировать антиген-специфичные ответы при инфекционных заболеваниях и вакцинации. Это приводит к повышенному уровню количества инфекций, опухолей и аутоиммунных заболеваний среди пожилых индивидуумов.

Одним из главных изменений в ходе старения иммунной системы является регрессия (или *инволюция*) тимуса [66]–[68], происходящая, по всей видимости, у всех позвоночных животных, и являющаяся древним и консервативным процессом [69]. При этом происходит уменьшение количества клеток и тканей тимуса, а также деградация внутренней организации, следствием чего в конце концов становится снижение количества производимых в нем наивных Т-лимфоцитов [66]–[68].

Процесс инволюции тимуса у человека имеет две фазы [70]: первая происходит в раннем возрасте – еще до пубертатного периода [71] – ее принято называть “инволюция тимуса, связанная с ростом”, поскольку она происходит в момент физического роста человека. Вторая фаза связана с возрастными изменениями в организме и происходит ближе к старости. Ее принято называть “инволюция тимуса, связанная со старением”.

Однако в то время, как многие исследования на животных показывают, что для поддержания стабильного количества наивных Т-лимфоцитов в периферической крови, действительно, необходима нормальная работа тимуса [72], связь между количеством этих клеток и активностью тимуса **у человека** до сих пор обсуждается [73], [74]. По всей видимости, в отличие от наивных Т-клеток мыши, основным источником которых – даже у пожилых особей – является функционирующий тимус, у взрослого человека большинство наивных Т-клеток образуется путем периферической пролиферации [72]. Значительный вклад последнего явления в количество периферических наивных Т-клеток был показан даже у детей [75], [76].

В отличие от генерации новых наивных Т-клеток в тимусе, периферическая пролиферация не приводит к увеличению разнообразия репертуара TCR взрослого

индивидуума. Тем не менее, минимальная продукция новых наивных клеток тимусом у человека все еще происходит в возрасте 55 лет [77], практически исчезая к 70 годам [78].

Иммунная система пожилого человека характеризуется также менее эффективным противостоянием даже уже знакомым патогенам, к которым за долгое время в организме были накоплены специфичные Т-лимфоциты памяти (например, к таким распространенным вирусам, как цитомегаловирус (CMV, англ. cytomegalovirus) или вирус Эпштейна-Барра).

Подобная неэффективность связана с появлением множества дефектов в передаче сигналов через TCR, что ухудшает скорость и эффективность иммунного ответа [79]. Кроме того, было показано, что большие олигоклональные популяции цитомегаловирус-специфичных Т-клеток препятствуют эффективному иммунному ответу на вирус Эпштейна-Барра и герпесвирусы [80]. Большие клоны Т-клеток памяти, по всей видимости, также препятствуют эффективной активации наивных Т-клеток при новой инфекции [81]. Цитомегаловирус же, по всей видимости, при старении не влияет на количество наивных CD8⁺ Т-лимфоцитов, однако приводит к достоверному уменьшению наивных CD4⁺ Т-лимфоцитов [82].

Способность клеток к пролиферации исчерпывается к старости. Это связано с общим уменьшением длины теломер Т-лимфоцитов пожилых людей по сравнению с молодыми [83]. Но – самое важное – происходит значительное укорочение длины теломер **наивных** Т-лимфоцитов [77]. Это наблюдение подтверждает идею о периферической пролиферации наивных Т-лимфоцитов, как способе поддержания количества Т-клеток в пожилом возрасте. Длина теломер для соматических клеток служит своеобразными митотическими часами [84]. Деление клеток приводит к быстрому уменьшению их длин, и при достижении критической длины митотическая активность клеток останавливается. Поскольку активная пролиферация Т-лимфоцитов является необходимым условием для TCR-обусловленного иммунного ответа, укорочение длины теломер, по-видимому, лежит в основе многих нарушений антиген-специфического ответа пожилых людей [85].

В гомеостатическом пространстве происходит постоянный процесс обращения наивных Т-клеток в Т-клетки памяти [86], связанный с постоянной конфронтацией их организма с окружающими антигенами. В результате оно постепенно заполняется, с

одной стороны, пролиферирующими наивными Т-клетками, с другой, – олигоклональными Т-клетками памяти [87], [88].

В работе, оценивающей репертуар TCR методом гибридизации [77], было показано, что >65% репертуара TCR наивных CD4⁺ Т-лимфоцитов молодых доноров (моложе 30 лет) состоит из очень редких клонотипов с частотой встречаемости порядка $1/2 \times 10^7$. Оставшиеся 35% занимают клоны с частотой $1/2 \times 10^5$, – наличие которых соответствует гипотезе о периферической пролиферации наивных Т-клеток. Этот уровень разнообразия – по данным исследования – держится примерно до возраста 65 лет и стремительно падает после. В возрасте 75 лет наивные CD4⁺ Т-клетки находятся в крови доноров с частотой более $1/2 \times 10^5$, то есть остаточный репертуар пожилых людей составляет менее 1%.

Другое исследование, использовавшее метод спектротипирования для анализа репертуара TCR β CD4⁺ Т-лимфоцитов людей разных возрастов – однако не превышающих 65 лет – показало отсутствие значительных возрастных изменений [89]. Падение репертуара CD4⁺ Т-клеток, тем не менее, не сопровождается достоверным уменьшением количества периферических наивных CD4⁺ клеток в крови пожилых (>65 лет) доноров [82]. Количество же наивных CD8⁺ Т-лимфоцитов, изменение репертуара которых не изучено на сегодняшний день, резко падает с возрастом, что, однако, не приводит к накоплению CD8⁺ Т-клеток памяти (количество которых либо не изменяется, либо снижается) [82].

Однако более подробного исследования изменения репертуара TCR человека при старении до нашей работы не было проведено [90]. В то же время, именно разнообразие репертуара TCR является залогом эффективного иммунного ответа на новые антигены, так как определяет вероятность нахождения нового специфического клона Т-клеток.

2.2. Методы исследования разнообразия индивидуальных репертуаров Т-клеточных рецепторов

Разработано множество методов для оценки разнообразия репертуаров TCR на уровне ДНК, РНК и белка. Проточная цитометрия позволяет сортировать клетки с определенным фенотипом и оценивать репертуар используемых V-сегментов генов TCR. Методы молекулярной биологии позволяют оценить репертуар TCR на уровне последовательностей геномной ДНК и мРНК.

2.2.1. Проточная цитометрия

Частота Т-лимфоцитов, экспрессирующих определенный V-сегмент одной из цепей TCR, может быть определена с помощью проточной цитометрии при условии доступности специфичных моноклональных антител. Этот метод дает возможность проводить совмещенный анализ TCR и других поверхностных маркеров в клетке. В зависимости от используемого прибора можно анализировать до 20 параметров одновременно (с помощью проточных цитометров) и до 70-100 параметров (с помощью масс-цитометров). С разработкой моноклональных антител, специфичных к различным V β -сегментам TCR человека, стало возможным подробное изучение репертуара V β сегментов. Проточная цитометрия позволяет проводить количественный и качественный анализ V-сегмента гетерогенных клеточных популяций, например, с целью определить общие тенденции изменения репертуара TCR при старении или применении лекарственных препаратов.

Однако этот метод позволяет оценить лишь разнообразие V-сегментов и не дает никакой информации о разнообразии гипервариабельных CDR3 участков.

2.2.2. Спектротипирование

Метод спектротипирования позволяет оценивать разнообразие и структуру репертуара TCR по соотношению участков CDR3 разной длины. Целевые последовательности TCR α и β локусов амплифицируют с помощью набора 5'-концевых праймеров, специфичных V-сегментам, и 3'-концевого C-специфического праймера. Полученную библиотеку используют в реакции праймер-достройки, где вместо

обычного 3'-концевого С-специфического праймера, применяется флуоресцентно, либо радиоактивно меченый праймер.

Далее ПЦР-продукты визуализируют при помощи нативного гель-электрофореза. Альтернативно, распределение длин CDR3 участков может быть оценено с помощью стандартного плащечного ДНК-секвенатора, используемого при секвенировании по Сэнгеру [91].

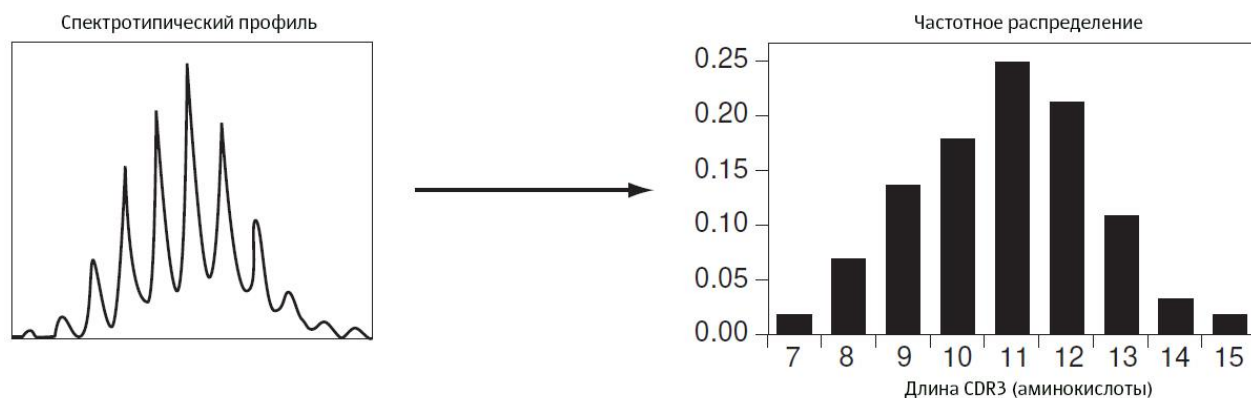


Рисунок 2.6 [87]. Пример спектротипического профиля.

Идея метода спекротипирования состоит в том, что в популяции Т-клеток основные различия TCR друг от друга сосредоточены в гипервариабельном участке β цепи CDR3. Причем эти различия заключаются не только в нуклеотидной последовательности разных CDR3 у разных TCR, но и в различии их длин (Рисунок 2.6). То есть степень равномерности распределения длин нуклеотидных последовательностей разных CDR3 фрагментов позволяет оценить степень поликлональности репертуара TCR в образце.

Так, спекротип наивных Т-лимфоцитов мыши представляет собой восьмипиковое колоколообразное распределение длин, в то время как на спекротипе иммунизированных животных доминируют один или несколько пиков, отражающих клональную экспансию лимфоцитов.

Наиболее точную оценку данный метод дает при работе с образцами, репертуар Т-клеток которых был сильно изменен в ходе каких-либо патологических процессов. Например, у пациентов, страдающих ревматоидным артритом, и здоровых людей спектротипический анализ периферических лимфоцитов крови позволил дать

качественную оценку репертуара TCR среди 22 V β семейств и показал существенное изменение репертуара T-клеток у больных по сравнению со здоровыми людьми [92].

Также метод спектrotипирования успешно применялся для мониторинга изменения разнообразия репертуара T-клеток у пациентов, с нарушениями гемопоэза до и после аллогенной трансплантации стволовых клеток в течение продолжительного времени [92]–[94]. В частности, было показано, что репертуар TCR β пациентов, страдающих заболеваниями гемопоэза, отличается меньшей вариабельностью по сравнению с репертуаром TCR здоровых людей. С течением времени после аллогенной трансплантации стволовых клеток у больных наблюдается постепенное восстановление репертуара T-клеточных рецепторов (Рисунок 2.7).

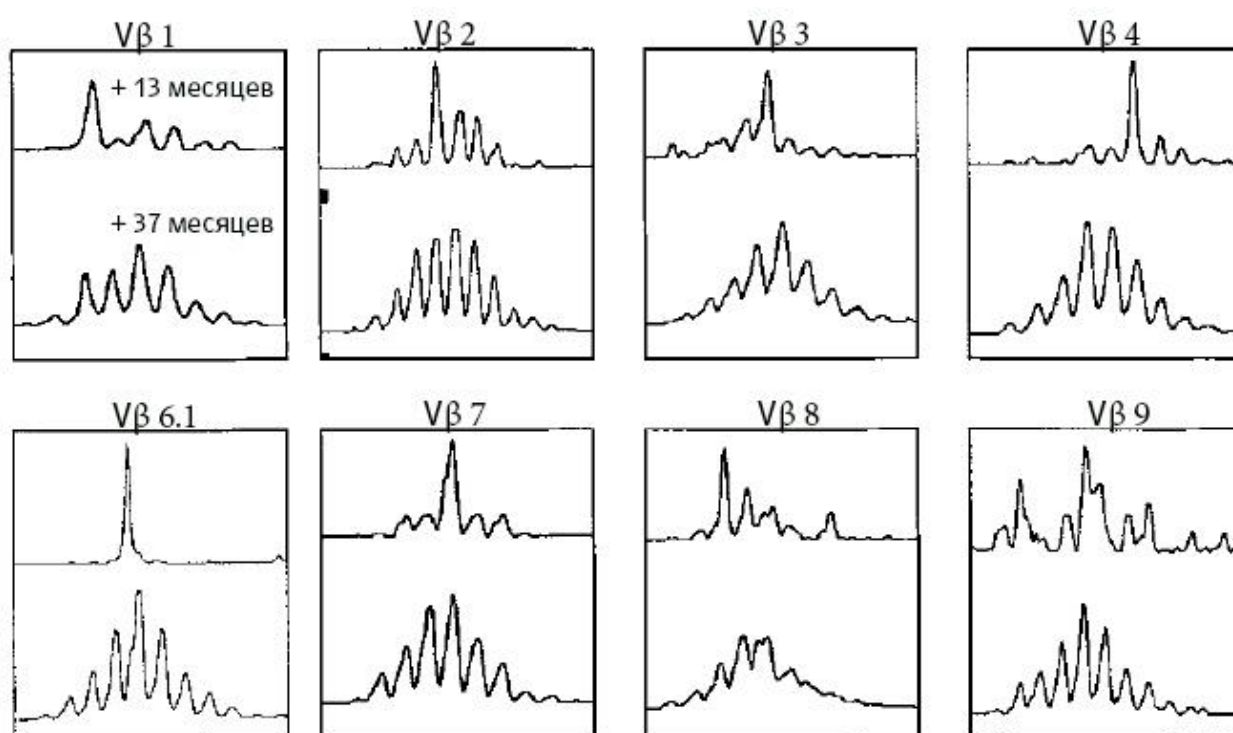


Рисунок 2.7 [92]. Спектротипические профили репертуаров CDR3 TCR β некоторых субпопуляций T-лимфоцитов пациента, страдающего хронической миелоидной лейкемией, через 13 и 37 месяцев после аллогенной трансплантации стволовых клеток.

Метод спектrotипирования позволяет качественно оценить разнообразие репертуара TCR образца (выявить наличие клональных экспансий). Тем не менее, этот метод неприменим для непосредственного анализа разнообразия TCR и определения клонального состава репертуара T-лимфоцитов.

2.2.3. Гибридизация

Альтернативным методом оценки разнообразия TCR является метод гибридизации. Разработанный для оценки количества повторяющихся последовательностей в геномной ДНК [95], этот метод может быть успешно применен и для оценки сложности смесей любых молекул нуклеиновых кислот, в том числе генов Т-клеточных рецепторов.

Кинетика ренатурации денатурированной ДНК зависит от её сложности (количества повторяющихся участков). Более сложные смеси нуклеиновых кислот ренатурируют медленнее, чем смеси, богатые повторяющимися последовательностями. Измеряя с помощью интеркалирующих красителей скорость ренатурации амплифицированных генов TCR, можно оценить их разнообразие, не прибегая к сложным и дорогостоящим методам, связанным с массивированным секвенированием или окраской антителами. Адаптация метода гибридизации для оценки разнообразия TCR называется AmpliCot [96].

Для использования AmpliCot требуется лишь прибор для ПЦР в реальном времени, а оценки разнообразия, полученные этим методом, относительно хорошо совпадают с данными, полученными с помощью секвенирования. В то же время, метод гибридизации дает лишь общую оценку разнообразия и не дает информации о структуре репертуара и последовательностях CDR3-участков TCR.

2.2.4. Секвенирование нуклеотидных последовательностей Т-клеточных рецепторов

Более полное понимание разнообразия TCR на уровне нуклеотидных последовательностей достигается с помощью секвенирования ДНК.

Ранее, секвенирование по Сэнгеру участков TCR, клонированных в бактериальные плазмиды, было единственным способом определения нуклеотидных последовательностей Т-клеточных рецепторов. Эти методы позволили впервые получить детальную информацию о репертуаре Т-клеточных рецепторов и привели к обнаружению некоторых антиген-специфичных последовательностей TCR. Однако, помимо высокой стоимости и трудоемкости такого анализа, количество анализируемых

последовательностей составляет 10^2 - 10^3 – в то время как полноценный анализ сложных репертуаров требует сотен тысяч и миллионов прочтений секвенирования.

Стремительное развитие за последнее десятилетие технологий массивированного секвенирования (так называемого *секвенирования нового поколения* – NGS, англ., *next generation sequencing*) позволяет получать информацию о миллионах и даже миллиардах последовательностей ДНК.

За последние годы было разработано несколько технологий массивированного секвенирования, отличающихся длиной, качеством, и количеством получаемых последовательностей (*ридов*), а также скоростью работы. Принципиально, все три современных платформы 454/Roche¹, IonTorrent и Illumina могут быть использованы для решения тех или иных задач анализа иммунных репертуаров с точки зрения длины ридов и глубины анализа. Все эти платформы используют метод “секвенирования путем синтеза”, который заключается в том, что для секвенирования одноцепочечной ДНК ферментативно синтезируют комплементарную цепочку. В Illumina используется *bridge*-амплификация (англ., мост), в результате которой на плоской подложке образуются образуются «молекулярные колонии» одноцепочечных молекул ДНК. В IonTorrent происходит амплификация в микролунках, в 454/Roche используется эмульсионная ПЦР на шариках. Принципы детекции сигнала отличаются у всех трех платформ.

Illumina использует терминирующие нуклеотиды с флуоресцентной меткой, сигнал от которой детектируется с помощью лазера и камеры. Roche использует принцип пиросеквенирования, при котором детектируется пирофосфат, отщепляющийся при включении в цепь комплементарного нуклеотида. В отличие от этих двух методов, IonTorrent не использует визуализацию молекулярных колоний – детекция сигнала осуществляется путем регистрации изменения pH при встраивании нуклеотида.

Секвенирование методом синтеза на матрице – Solexa, разработанное компанией Illumina, представляет собой одну из наиболее популярных современных платформ NGS [97]. Весь механизм секвенирования можно разбить на четыре последовательных этапа:

- 1) пробоподготовка;
- 2) построение кластеров;

¹ В 2013 году компания Roche анонсировала прекращение поддержки платформы 454.

- 3) секвенирование посредством синтеза;
- 4) биоинформатический анализ полученных данных.

На первом этапе к коротким целевым фрагментам ДНК с обеих сторон пришивают специальные адаптеры известной последовательности. Пришивание этих адаптеров проводится методом Vectorsite [98].

На следующем этапе подготовленную таким образом ДНК наносят на пластинку с большим количеством иммобилизованных праймеров двух типов, комплементарным пришитым к ДНК адаптерам, и проводят амплификацию (Рисунок 2.8, А). Связавшаяся с иммобилизованным на подложке праймером молекула ДНК амплифицируется (Рисунок 2.8, Б). Затем подложку прогревают, при этом происходит денатурация ДНК и смывание не связанной с подложкой комплементарной цепи (Рисунок 2.8, В). При этом на пластинке остается связанная одноцепочечная ДНК, комплементарная целевой последовательности и фланкированная двумя типами праймеров – одним свободным и одним связанным с подложкой.

Цепь ДНК может изгибаться и связываться с другим праймером на подложке (Рисунок 2.8, Г) – при этом происходит образование характерного одноцепочечного «моста» (отсюда bridge-амплификация). В ходе следующего раунда амплификации происходит достройка одноцепочечного «моста» до двухцепочечного (Рисунок 2.8, Д). После нагревания цепи вновь расходятся, но в этом случае обе из них остаются связанными с подложкой (Рисунок 2.8, Е). Так, из одной связанной с подложкой цепи получаются две (одна прямая и одна обратная), которые также могут изогнуться, связаться с другими праймерами и вступить в новые раунды амплификации (Рисунок 2.8, Ж-З).

После нескольких раундов амплификации на подложке образуются обособленные колонии – кластеры ДНК, – соответствующие единичным молекулам исходной ДНК (Рисунок 2.8, И-К). После завершения амплификации обратные цепи отрезаются и вымываются, на пластинке остаются лишь кластеры молекул, соответствующих прямой последовательности исходной ДНК (Рисунок 2.8, Л).

На третьем этапе происходит секвенирование (Рисунок 2.9). На подложку наносят смесь из четырех модифицированных оснований и праймер для синтеза, комплементарный свободному концу последовательности, иммобилизованной на пластинке (Рисунок 2.9, А). Модифицированные основания представляют собой 3'-O-

азидометил-2'-деоксинуклеозид-трифосфаты, каждый из которых имеет флуоресцентную метку своего цвета. 3'-O-азидометил не позволяет ДНК-полимеразе присоединять более одного нуклеотида за раз. После присоединения одного нуклеотида оставшиеся нуклеотиды смываются. Далее происходит освещение ячейки коротким импульсом лазера. Флуоресцентные метки каждого из присоединенных четырех нуклеотидов излучают свет определенной длины волны, что позволяет отличать их друг от друга. Поскольку после амплификации ДНК каждая молекула представлена множеством копий, сигнал одновременно фиксируется от большого количества одинаковых флуорофоров (Рисунок 2.9, В-Г).

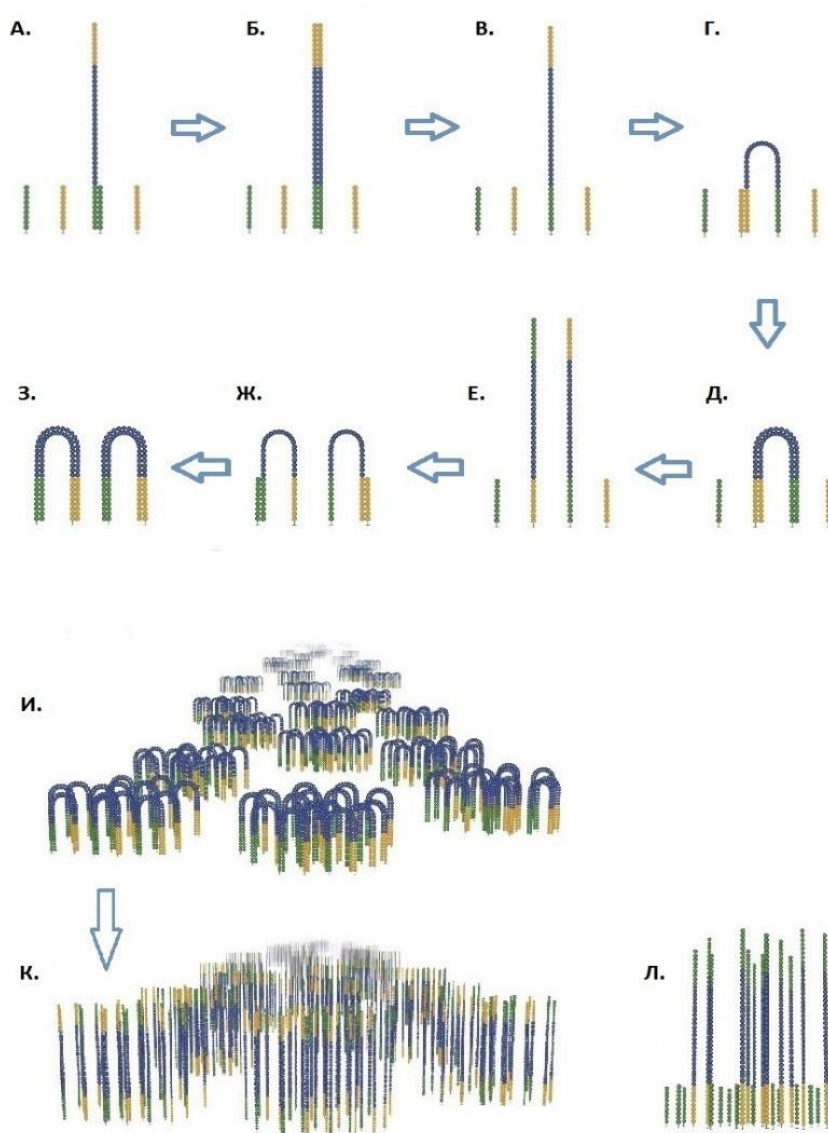


Рисунок 2.8 (©Illumina). **Механизм образования молекулярных колоний на подложке.**

Затем происходит добавление трис(2-карбокsetил)фосфина (ТСЕР), отделяющего флуорофор и азидометил от основания. При этом происходит освобождение 3'-гидроксильной группы для присоединения следующего нуклеотида.

Проблемой применения секвенирования нового поколения к анализу репертуара TCR является наличие ошибок ПЦР и секвенирования, создающих искусственное разнообразие TCR. Ошибки секвенирования отличаются для разных платформ. Для корректной оценки репертуара TCR необходимо проводить коррекцию ошибок, возникающих как на этапе секвенирования, так и на этапе пробоподготовки.

Существующие методы оценки разнообразия TCR на основе анализа последовательностей нуклеиновых кислот, как правило, применимы как для геномной ДНК, так и для мРНК. Однако использование геномной ДНК в качестве матрицы имеет ряд недостатков. Во-первых, при этом не элиминируется наличие второй гомологичной хромосомы, то есть в анализе будут участвовать как удачно, так и неудачно рекомбинировавшие гены. Во-вторых, при ПЦР-амплификации участков TCR с геномной ДНК используют специфические мультиплексы праймеров к семействам V- и J- генов, вносящие неравномерность в процесс амплификации. В-третьих, малое количество копий геномной ДНК (по сравнению с мРНК) существенно снижает эффективность анализа.

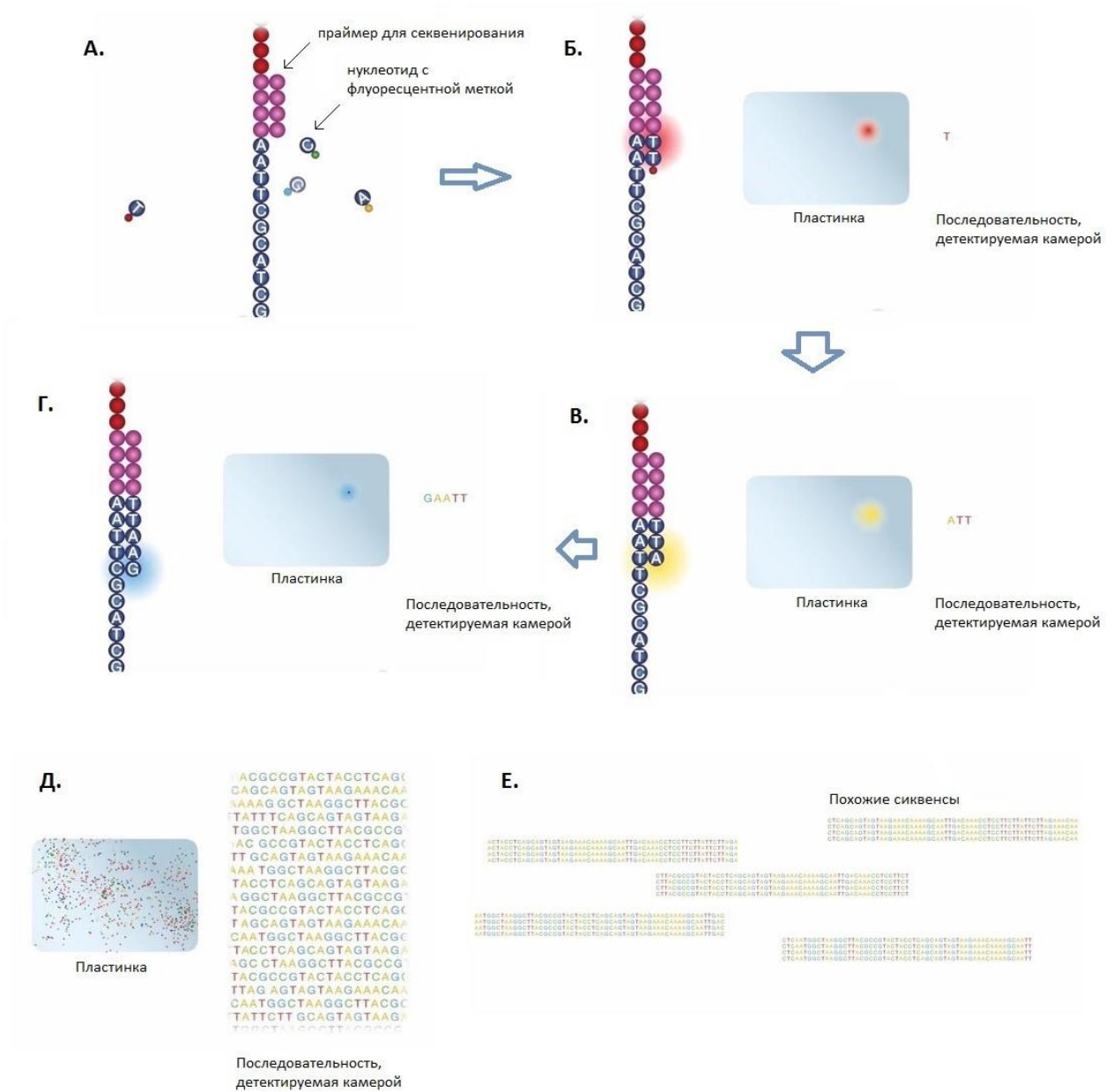


Рисунок 2.9 (©Illumina). Механизм секвенирования методом синтеза по матрице.

Библиотеки TCR, приготовленные с использованием мРНК в качестве стартового материала, менее искажены и обладают рядом преимуществ.

Во-первых, большое количество копий молекул РНК, кодирующей β и α цепи TCR в каждом Т-лимфоците, существенно увеличивает вероятность амплификации каждого из вариантов TCR.

Во-вторых, наличие механизма нонсенс-опосредованного разрушения мРНК (NMD, англ. nonsense mediated decay, см. раздел 2.1.2) позволяет избегать амплификацию нефункциональных вариантов TCR [99], [100].

В-третьих, последовательность TCR β может быть амплифицирована с кДНК с использованием универсальных праймеров, специфичных к константным участкам генов TCR β [101][102]. Кроме того, для синтеза кДНК может быть использован эффект смены матрицы ревертазой [103]. Использование этого эффекта обеспечивает ввод известной последовательности на 5'-конце мРНК, что позволяет использовать специфический к этому участку праймер в процессе амплификации. За счет использования одной пары олигонуклеотидов, удается значительно снизить погрешности, связанные с неравномерностью амплификации с помощью большого количества праймеров, специфичных к разным V β - и J- β сегментам, и провести более количественный анализ представленностей клонотипов TCR в репертуаре.

Тем не менее, и для таких библиотек существует ряд ограничений, которые следует учитывать при анализе результатов и планировании экспериментов.

Во-первых, уровень мРНК Т-клеточного рецептора может отличаться в различных клонах Т-лимфоцитов, что сказывается на представленности этих клонов в библиотеке TCR.

Во-вторых, необходимо учитывать состояние используемых клеток при приготовлении библиотеки: клетки из свежесыводенных образцов крови и клетки, прошедшие процедуру заморозки/разморозки, могут различаться по качеству и количеству РНК.

Важно также помнить, что обратная транскрипция и процесс смены матрицы ревертазой могут иметь разную эффективность в зависимости от используемого фермента и адаптера для смены матрицы. Предпочтительно использовать одну и ту же обратную транскриптазу и один и тот же адаптер для приготовления библиотек, разнообразие которых будет в дальнейшем сравниваться (например, при сравнении относительного разнообразия TCR наивных Т-клеток в образце РВМС).

2.2.5. Ошибки ПЦР и секвенирования – искусственное разнообразие

Основная проблема, отличающая анализ репертуаров Т-клеточных рецепторов от анализа заранее известного генома или транскриптома, вытекает из того факта, что с увеличением количества прочтений возрастает количество ошибочно прочитанных последовательностей, отличающихся на одну, две или три нуклеотидных замены.

Помимо ошибок секвенирования, в получаемых данных присутствуют ошибки, возникшие на других этапах приготовления библиотеки - ошибки обратной транскриптазы (в случае старта с мРНК) и ошибки, накопленные в ходе ПЦР амплификации библиотеки.

При анализе данных массивного секвенирования возникает проблема интерпретации ошибочных прочтений как реально существующих вариантов TCR. В совокупности, перечисленные типы ошибок могут привести к искусственному увеличению разнообразия TCR в несколько раз [104], [105].

Для решения этой проблемы в нашей лаборатории был разработан алгоритм избавления от ошибок, возникающих на этапе ПЦР и секвенирования [104]. Идентичные прочтения секвенирования (*риды*) хорошего качества в каждой нуклеотидной позиции формируют "сердцевинные" клонотипы. Затем происходит поиск соответствия между риды, имеющими плохое качество секвенирования в менее, чем в четырех нуклеотидных позициях, и сформированными "сердцевинными" клонотипами. Риды, имеющие плохое качество в более, чем трех нуклеотидных позициях, а также риды, для которых не обнаружилось соответствия с сформированными "сердцевинными" клонотипами, элиминируются. Поскольку Т-лимфоциты не имеют стадии гипермутации TCR, исправление ошибок происходит внутри V-, D- и J- сегментов и допускает лишь определенное количество ошибок в пределах каждого из них [104]. При этом учитывается количество ридов, соответствующее каждой последовательности (слияние клонотипов происходит лишь в случае, если количество ридов, соответствующее "правильному" клонотипу, существенно (в 3, 5, 10 раз – параметр задается в зависимости от глубины секвенирования и поставленной задачи) превосходит количество ридов, соответствующее клонотипу с ошибкой, Рисунок 2.10).

Данный алгоритм позволяет элиминировать искусственное разнообразие с учетом качества секвенирования каждого нуклеотида и повысить точность определения относительной концентрации клонотипов. В отличие от слепого удаления 4% прочтений секвенирования, описанного в работе [106], такой способ делает возможным рациональное избавление от ошибок ПЦР и секвенирования, максимально сохраняя при этом естественное разнообразие репертуара TCR.

| Количество ридов | Нуклеотидная последовательность CDR3 | Длина CDR3, нт | Аминокислотная последовательность CDR3 | V-сегмент | J-сегмент | D-сегмент |
|---------------------|---|----------------------|--|--------------|-------------|-----------|
| 18732 | tgtgcca caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKT TSGIVDEQFF | V10-3 | J2-1 | D2 |
| 74 | tgtgcc g tcaagacgactagcgggattgtggatgagcagttcttc | 45 | CAVKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 2 | tgtgcc g tcaagacgactagc g ggattgtggatgagcagttcttc | 45 | CAVKTTSRIVDEQFF | V10-3 | J2-1 | D2 |
| 1 | tgtg g ccatcaagacgactagcgggattgtggatgagcagttcttc | 45 | CVVKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 74 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDGQFF | V10-3 | J2-1 | D2 |
| 1 | tgtgcca g caagacgactagc g gattgtggatgagcagttcttc | 45 | CAIKTTSEIVDGQFF | V10-3 | J2-1 | D2 |
| 1 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CATKTTSGIVDGQFF | V10-3 | J2-1 | D2 |
| 63 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTGGIVDEQFF | V10-3 | J2-1 | D2 |
| 58 | tgtgcca g caagacgactag g gggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 1 | tgtgcca g caagacgactag g gggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDE*FF | V10-3 | J2-1 | D2 |
| 55 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CATKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 52 | tgtgcca g caagacgactagc g ggattgtggatgagcagttcttc | 45 | CAIKTTSRIVDEQFF | V10-3 | J2-1 | D2 |
| 45 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 1 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CANKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 28 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 15 | tgtg g ccatcaagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 8 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 4 | tgtg g caatcaagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDEQFF | V10-3 | J2-1 | D2 |
| 2 | tgtgcca g caagacgactagcgggattgtggatgagcagttcttc | 45 | CAIKTTSGIVDDQFF | V10-3 | J2-1 | D2 |
| 1 | tgtgcca g caagacgactagc g gattgtggatgagcagttcttc | 45 | CAIKTTSVIVDEQFF | V10-3 | J2-1 | D2 |

Рисунок 2.10 [104]. Типичный пример слияния прочтений секвенирования, содержащих ошибки, с доминирующим "сердцевинным" клонотипом.

Доминантный клонотип показан жирным шрифтом. Показаны V- (рыжий), D- (синий) и J- (зеленый) сегменты, а также исправленные ошибки (красный) внутри CDR3-участка. Серым фоном выделены производные ошибки от первичного ошибочного варианта последовательности.

3. Цели и задачи

Цель настоящей работы заключалась в изучении индивидуального репертуара β -цепей Т-клеточных рецепторов периферической крови человека и влияния на него наследственных факторов и старения. Были поставлены следующие задачи:

- 1) Подготовить библиотеки и методом массированного секвенирования получить данные по индивидуальным репертуарам β -цепей Т-клеточных рецепторов периферической крови для группы матерей и их детей;
- 2) Провести сравнительный анализ состава и структуры репертуаров β -цепей Т-клеточных рецепторов родственных и неродственных пар мама-ребенок;
- 3) Разработать метод на основе молекулярного баркодирования, позволяющий проводить корректное сравнение разнообразия репертуаров CDR3 участков библиотек генов TCR, полученных от различных доноров, в условиях разного количества клеток на старте, разного качества РНК, и разной глубины массированного секвенирования;
- 4) Подготовить молекулярно-баркодированные библиотеки и методом массированного секвенирования получить данные по индивидуальным репертуарам β -цепей Т-клеточных рецепторов периферической крови для когорты системно здоровых доноров разных возрастов;
- 5) Провести сравнительный анализ нормированного разнообразия репертуаров β -цепей Т-клеточных рецепторов периферической крови для доноров различного возраста.

4. Материалы и методы

4.1. Оборудование и расходные материалы

Покрытые ЭДТА пробирки Vacutainer (BD Biosciences, США); камера Горяева; автоматические микропипетки (Eppendorf, Германия; Gilson, Франция); пластиковые наконечники для микропипеток (SSI, США); центрифуги Eppendorf 5415R, 5804R (Eppendorf, Германия); вортекс Bio Vortex V1 (Biosan, Латвия); термостат Термит (ДНК-технология, Россия); пластиковые пробирки объемом 15 мл (Corning, США) и 50 мл (Greiner, Австрия); 0.2 мл пробирки для ПЦР (SSI, США); 1.5 мл пробирки (Sarstedt, Германия); автоматический ПЦР-амплификатор ABI 9700 Thermal Cycler (Applied Biosystems, США); камера для электрофореза (Хеликон, Россия); источник питания Эльф-4 (ДНК-технология, Россия); Nanodrop (Thermo Scientific, США); Qubit флуорометр (Invitrogen, США); набор для выделения ДНК из ПЦР-смеси Qiaquick (Qiagen, Голландия); секвенатор Illumina HiSeq2500; NEBNext DNA Library Prep Master Mix Set для Illumina (New England Biolabs, США); киты для типирования HLA: SSP AllSet Gold HLA-ABC Low Res и SSP AllSet Gold HLA-DRDQ Low Res (Invitrogen, США); проточный цитофлуориметр Cytomics FC 500 (Beckman Coulter, США).

4.2. Реактивы

Фиколл ($\rho = 1,077 \text{ г/см}^3$; ПанЭко, Россия); буфер PBS (137 mM NaCl, 2.7 mM KCl, 1.47 mM KH_2PO_4 , 4.29 mM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, pH 7.3-7.5) (ПанЭко, Россия); Trizol (Invitrogen, США); хлороформ (Sigma, США); изопропиловый спирт (Sigma, США); этиловый спирт (Sigma, США); буфер для синтеза первых цепей ревертазой Mint (5x) (Евроген, Россия); обратная транскриптаза Mint (Евроген, Россия); смесь dNTP (Евроген, Россия); дитиотрейтол (ДТТ) (Евроген, Россия); IP-раствор (Евроген, Россия); урацил-ДНК-гликозилаза (NEB, США); буфер для ПЦР реакции (10x) для Encyclo-полимеразы (Евроген, Россия); смесь ДНК-полимераз Encyclo (Евроген, Россия); агароза для электрофореза (Promega, США); глицерин (Sigma, США); буфер для нанесения образцов ДНК на агарозный гель (0.1% бромфеноловый синий, 0.1% ксиленцианол, 30% глицерин); антитела: CD3-PC7 (клон UCNT1; eBioscience, США), CD27-PC5 (клон 1A4CD27; Invitrogen, США), CD4-PE (клон 13B8.2; Beckman Coulter, США), CD45RA FITC (клон JS-B3; eBioscience, США).

Использованные олигонуклеотиды:

| | |
|-------------|---|
| BC_R4_short | GTATCTGGAGTCATTGA |
| SmartNNNNa | AAGCAGUGGTAUCAACGCAGAGUNNNNNUNNNNNUNNNNUCTTrGrGrGrG |
| BC_uni_R | TGCTTCTGATGGCTCAAACAC |
| M1SS | AAGCAGTGGTATCAACGCA |
| M1s | (N) ₂₋₄ (XXXXX) CAGTGGTATCAACGCAGAG |
| VJ1-1 | GTGAGTCTGGTGCCTTGCCAA |
| VJ1-2 | GTTAACCTGGTCCCCGAACCGA |
| VJ1-3 | GTGAGCCAACTCCCTCTCCAA |
| VJ1-4 | GAGAGCTGGGTCCACTGCCAA |
| VJ1-5 | GAGAGTCGAGTCCCATCACCAA |
| VJ1-6 | GTGAGCCTGGTCCCRTTCCCAA |
| VJ2-1 | GTGAGCCGTGTCCCTGGCCCGA |
| VJ2-2 | GTCAGCCTAGAGCCTTCTCCAA |
| VJ2-3 | GTCAGCCGGGTGCCTGGGCCAA |
| VJ2-4 | GAGAGCCGGGTCCCGGCGCCGA |
| VJ2-5 | AGGAGCCGCGTGCCTGGCCCGA |
| VJ2-6 | GTCAGCCTGCTGCCGGCCCCGA |
| VJ2-7 | GTGAGCCTGGTGCCCGGCCCGA |

Все олигонуклеотиды, за исключением SMART-адаптера, были синтезированы компанией Евроген (Россия). SMART-адаптер был синтезирован компанией ДНК-синтез (Россия).

4.3. Методы

4.3.1. Сбор образцов

Эта работа была одобрена комиссией по биоэтике и проведена, согласно с хельсинской декларацией. Все доноры были проинформированы об окончательной цели использования образцов их крови и дали письменное согласие.

В работе принимало участие 48 системно здоровых европеоидных донора. Из них в первой части работы (в исследовании репертуаров TCR пар мама-ребенок) принимало участие 3 мамы (средний возраст 40 ± 4) и их 6 детей (средний возраст 11 ± 4 ; по 2 ребенка на каждую мать). Во второй части работы (в исследовании изменений репертуаров TCR при старении) принимало участие 39 человек. В зависимости от возраста, они были разделены на 4 группы:

Группа 1 – молодые доноры (в ср. – 16 лет),

Группа 2 – доноры среднего возраста (в ср. – 39 лет),

Группа 3 – пожилые доноры (в ср. – 62 года),

Группа 4 – долгоживущие доноры (в ср. – 82 года).

10 мл венозной крови каждого из 48 доноров собирали в покрытые ЭДТА пробирки Vacutainer (BD Biosciences, США).

4.3.2. Выделение РВМС

Выделение *периферических мононуклеарных клеток крови* (РВМС, англ. peripheral blood mononuclear cells) из цельной крови проводили с помощью центрифугирования в градиенте фикола. Все операции проводили при комнатной температуре. К образцу периферической крови добавляли 2 объема раствора PBS (ПанЭко) и перемешивали. Затем аккуратно наслаивали разведенную кровь на 20 мл фикола (ПанЭко, $\rho = 1,077$ г/см³) в две пробирки объемом 50 мл (25 мл разведенной крови + 20 мл фикола). Центрифугировали образцы 30 минут с ускорением 600 g. Во время центрифугирования в пробирках происходило расслоение: верхний прозрачный слой, мутная интерфаза, нижний красный слой, красный осадок. Пипеткой аккуратно отбирали интерфазу, представляющую собой суспензию РВМС, переносили в новую пробирку объемом 15 мл. Отобранную интерфазу разбавляли раствором PBS до 15 мл. Центрифугировали 15 минут с ускорением 400 g. Большую часть супернатанта отбирали, клетки

ресуспендировали в оставшемся объеме. Объем доводили раствором PBS до 15 мл. Центрифугировали 15 минут при 400 g. Супернатант отбирали, клетки ресуспендировали в оставшемся объеме раствора PBS. Количество клеток в образце подсчитывали с помощью камеры Горяева.

4.3.3. Выделение тотальной РНК из клеток

К осадку клеток добавляли 1 мл реактива Trizol (Invitrogen, США), тщательно перемешивали, инкубировали при комнатной температуре 5 мин. Затем добавляли 0.2 мл хлороформа, встряхивали и инкубировали при комнатной температуре 2-3 мин, после чего центрифугировали с ускорением 12 000g при 4°C 15 мин. Водную фазу отбирали в новые пробирки и добавляли 0.5 мл изопропилового спирта. Тщательно перемешивали и центрифугировали с ускорением 12 000g при 4°C 15 мин. Осадок РНК промывали 95% этанолом и затем 80% этанолом. Супернатант отбирали и сушили при комнатной температуре до полного испарения этанола. Сухой осадок перерастворяли в 30 мкл mQ, не содержащей РНКаз (Qiagen, Голландия), и прогревали 10 мин. при 50°C.

Качество полученного раствора РНК и количество РНК в нем анализировали с помощью гель-электрофореза и прибора Nanodrop (Thermo Scientific, США).

4.3.4. Синтез кДНК

Синтез первых цепей кДНК производили с помощью набора для синтеза кДНК Mint (Евроген, Россия) по разработанному в нашей лаборатории протоколу [107]. Для синтеза кДНК использовали всю полученную из каждого образца РНК (как минимум, по 6 мкг/образец). В одну реакцию синтеза кДНК объемом 15 мкл добавляли 1.5 мкг РНК, то есть для каждого образца замешивали от четырех до шести реакций.

В качестве праймера для синтеза использовали олигонуклеотид BC_R4_short, специфичный к обоим вариантам константного домена TCR β . Для денатурации РНК и отжига праймера смесь РНК и олигонуклеотида инкубировали при температуре 70°C 2 минуты и затем при 40°C 2 минуты. Затем к реакции добавляли буфер для синтеза первой цепи, dNTP, ДТТ и SMART-адаптер SmartNNN_{Na}. Ревертаза воспринимает SMART-адаптер, как продолжение РНК матрицы и продолжает синтез первой цепи кДНК. Это явление называется эффектом смены матрицы ревертазой (*свитч*). Благодаря этому эффекту таргетная последовательность оказывается фланкирована известной

последовательностью как с 5', так и с 3'-концов. В состав используемого нами SMART-адаптера входит молекулярный идентификатор (т.е. последовательность из 12 случайных "N"-нуклеотидов, см. Разд. 5.2), а также 6 дезоксиурацилов и 4 рибогуанидина. Дезоксиурацилы необходимы для избавления от лишних молекул SMART-адаптера после реакции обратной транскрипции. Рибогуанидины повышают эффективность свитча [107].

Полученную смесь инкубировали при температуре 42°C 40 минут. Затем к реакции добавляли 5 мкл IP-раствора (Евроген, Россия) для повышения эффективности свитча. После этого реакцию инкубировали при температуре 42°C 130 минут.

Конечная реакционная смесь включала в себя:

- 2 мкл 5x буфера для синтеза первой цепи (Евроген, Россия);
- 1 мкл dNTP (10 мМ) (Евроген, Россия)
- 1 мкл ДТТ (20 мМ) (Евроген, Россия)
- 1 мкл обратной транскриптазы Mint (Евроген, Россия)
- 1 мкл РНК (1.5 мкг)
- 1 мкл олигонуклеотида BC_R4_short (10 мкМ)
- 1 мкл 5'-адаптера SmartNNNNa (15 мкМ)
- 1 мкл H₂O (Евроген, Россия)
- 5 мкл IP-раствора (Евроген, Россия)

Для удаления оставшихся молекул свитча SmartNNNNa реакцию синтеза завершали добавлением 1 мкл УДГ (урацил-ДНК-гликозилазы NEB, BioLabs) с последующей инкубацией образца при 37°C в течение 40 минут и последующей дезактивацией фермента при 95°C в течение 10 минут. До непосредственной амплификации полученный раствор кДНК хранили при 4°C.

4.3.5. Первый раунд ПЦР-амплификации

Амплификацию кДНК проводили с использованием смеси полимераз Encyclo (Евроген, Россия). В первом раунде ПЦР-амплификации использовалась вся кДНК каждого из образцов. В каждую реакцию объемом 50 мкл добавляли 3 мкл кДНК (до 32 реакций первого раунда ПЦР-амплификации на образец).

Для амплификации переменного участка TCR β использовались два олигонуклеотида: M1SS праймер, комплементарный инвариантной последовательности SmartNNNNa, и BC_uni_R праймер, комплементарный обоим вариантам константного участка TCR β .

Состав реакционной смеси:

- 5 мкл ПЦР-буфера Encyclo (Евроген, Россия)
- 1 мкл dNTP (10 мМ) (Евроген, Россия)
- 1 мкл смеси полимераз Encyclo (Евроген, Россия)
- 1 мкл праймера M1ss (10 мкМ)
- 1 мкл праймера BC_uni_R (10 мкМ)
- 3 мкл кДНК
- 38 мкл H₂O (Евроген, Россия)

Реакцию проводили в ПЦР-амплификаторе ABI 9700 Thermal Cycler (Applied Biosystems, США) в пробирках объемом 0.2 мл. Амплификация проводилась в течение 18 циклов с предварительным прогревом до 95°C в течение 2 мин. Использовалась следующая программа амплификации:

95°C 10 сек -> 62°C 20 сек -> 72°C 30 сек

После амплификации все реакции, относящиеся к одному образцу, смешивали в одной пробирке и тщательно перемешивали. От полученной смеси отбирали аликвоту объемом 100 мкл и очищали ДНК из реакционной смеси с помощью колонок QIAquick (Qiagen, Голландия) по протоколу производителя. Элюцию ДНК проводили в 20 мкл буфера EB (Qiagen, Голландия).

4.3.6. Второй раунд ПЦР-амплификации

Второй раунд ПЦР-амплификации также проводили с использованием смеси полимераз Encyclo (Евроген, Россия). В реакцию объемом 50 мкл добавляли 1 мкл очищенного продукта первого раунда ПЦР-амплификации. Всего для каждого образца было поставлено по 4 реакции второго раунда ПЦР-амплификации.

В качестве специфического 5'-праймера мы использовали олигонуклеотид M1s, комплементарный M1ss и имеющий в своем составе на 5'-конце 2-4 случайных нуклеотида ((N)₂₋₄) для лучшей кластеризации на секвенаторе Illumina, а также идентификатор образца (XXXXX), то есть известная последовательность из 5

нуклеотидов, уникальная для каждого образца. Со стороны 3'-конца мы использовали мультиплекс праймеров, комплементарных J β -сегментам [102].

Состав реакционной смеси:

- 5 мкл ПЦР-буфера Encyclo (Евроген, Россия)
- 1 мкл dNTP (10 мМ) (Евроген, Россия)
- 1 мкл смеси полимераз Encyclo (Евроген, Россия)
- 1 мкл праймера M1s (10 мкМ)
- 1 мкл смеси праймеров J β (10 мкМ)
- 1 мкл кДНК
- 40 мкл H₂O (Евроген, Россия)

Реакцию проводили в ПЦР-амплификаторе ABI 9700 Thermal Cycler (Applied Biosystems, США) в пробирках объемом 0.2 мл. Амплификация проводилась в течение 7-11 циклов (в зависимости от образца) с предварительным прогревом до 95°C в течение 2 мин. Использовалась следующая программа амплификации:

95°C 10 сек -> 62°C 20 сек -> 72°C 30 сек

4.3.7. Секвенирование

Концентрация продуктов ПЦР была определена с помощью прибора Qubit флуорометра (Invitrogen, США). ПЦР-продукты, включающие в себя баркоды образцов, смешивались эквимолярно. Адаптеры Illumina были лигированы к образцам по протоколу производителя с использованием NEBNext DNA Library Prep Master Mix Set для Illumina (New England Biolabs, США).

Для первой части исследования было использовано 2 дорожки Illumina – на одной из которых анализировали образцы мам, на другой – детей. Такое распределение было связано с элиминированием вероятности контаминации между сравниваемыми образцами на стадии секвенирования.

Для второй части исследования, с целью избежания возникновения перекосов в ходе секвенирования, доноры разных возрастов были случайно распределены по трем дорожкам секвенирования.

Все исследуемые библиотеки были проанализированы с помощью Illumina HiSeq2500, с использованием двустороннего прочтения секвенирования 100 + 100 нуклеотидов и собственных праймеров Illumina.

4.3.8. Анализ результатов массивного секвенирования

В обеих частях работы идентификацию V β сегмента (с использованием номенклатуры IMGT [108]), идентификацию CDR3 (основанную на распознавании последовательности, располагающейся между консервативными Cys-104 и Phe-118, включительно), кластеризацию клонотипов, а также коррекцию ошибок обратной транскрипции, ПЦР и секвенирования проводили с использованием разработанного в нашей лаборатории программного обеспечения MiTCR [109], в котором используется выше описанный алгоритм избавления от ошибок на стадии ПЦР и секвенирования [104] (см. разд. 2.2.5). Значение минимального удовлетворяющего качества каждого нуклеотида внутри CDR3 региона соответствовало Phred>25. При этом происходило вытягивание последовательностей плохого качества путем установления соответствий между ними и сформированными клонотипами из последовательностей хорошего качества. Для коррекции ошибок использовался наиболее строгий алгоритм “eliminate these errors”, исправляющий максимальное количество ошибок ПЦР и секвенирования. Этот алгоритм элиминирует 98% искусственного разнообразия CDR3 клонотипов. В то же время во время такого исправления ошибок теряется 2% истинного разнообразия TCR β (показано нами с помощью контрольных экспериментов и *in silico* моделирования).

4.3.9. Расчет дивергенции Джессена-Шеннона между распределениями V β -сегментов

Чтобы определить степень сходства распределения использования V β -сегментов родных и неродных пар мама-ребенок, мы использовали метод дивергенции Джессена-Шеннона (JS). Этот метод основан на дивергенции Кульбака-Лейблера (KL), однако имеет ряд существенных изменений. В частности, он симметризован и позволяет получить конечное значение. Дивергенция Кульбака-Лейблера (KL) и дивергенция Джессена-Шеннона (JS) определяются следующим образом [110]:

$$JS(P, Q) = \frac{1}{2} (KL(P, \frac{P+Q}{2}) + KL(Q, \frac{P+Q}{2})),$$

$$KL(P, Q) = \sum_i p_i \log_2 \frac{p_i}{q_i}.$$

Где P и Q – распределения частот встречаемости V β -сегментов в двух анализируемых индивидуумах, а p_i и q_i – частота определенного V-сегмента в первом и втором анализируемом индивидууме, соответственно.

Для сравнения значений дивергенции Дженсена-Шеннона родных и неродных пар мама-ребенок мы использовали двусторонний t-тест для независимых выборок, при этом различие считалось достоверным при P-значении < 0.05. Для контроля групповой вероятности ошибок использовались р-значения с поправкой Бонферрони.

4.3.10. Оценка нижней границы разнообразия репертуара CDR3 TCR β периферической крови

Оценка полного числа уникальных клонотипов TCR β было оценено так же, как в [6], используя модель ненаблюдаемых видов, описанную в [111]. Вкратце, для описания зависимости ожидаемого числа событий x_s для данного клонотипа от глубины прочтения t была выбрана модель Пуассоновского процесса. В соответствии с [111] это приводит к оценке полного числа уникальных клонотипов в виде $S(t) = \sum_{x=1}^{\infty} h_x n_x$, где n_x – число клонотипов обнаруженных в точности x раз. Суммы этого ряда, однако, расходятся с увеличением глубины прочтения, что делает оценку полного разнообразия невозможной. Эту проблему можно решить используя преобразование Эйлера, которое улучшает сходимость получившегося ряда. Таким образом, мы суммируем первые x_0 членов, где x_0 выбран как наибольшее из чисел, при котором коэффициент вариации для оценки $S(t)$ меньше 0.1. Это позволяет выбрать оптимальную оценку с учетом минимального смещения и разброса (см параграф 5 в [111]).

4.3.11. Проточная цитометрия

Небольшая аликвота РВМС каждого из образцов была мечена антителами для анализа с помощью проточной цитометрии. Использовались следующие антитела к человеческим антигенам: CD3-PC7 (клон UCHT1; eBioscience, США), CD27-PC5 (клон 1A4CD27; Invitrogen, США), CD4-PE (клон 13B8.2; Beckman Coulter, США), CD45RA FITC (клон JS-B3; eBioscience, США).

Клетки в PBS инкубировали с антителами в темном месте при комнатной температуре в течение 20 минут. Затем дважды отмывали от избыточных антител раствором PBS и хранили при температуре +4°C до непосредственного анализа методом проточной цитофлуориметрии.

Для анализа использовали проточный цитофлуориметр Cytomics FC 500 (Beckman Coulter, США). Анализ полученных данных осуществляли с помощью программного обеспечения Cytomics RXP Analysis (Beckman Coulter, США).

4.3.12. HLA-типирование

Образцы мам и детей были HLA-типированы с помощью китов SSP AllSet Gold HLA-ABC Low Res и SSP AllSet Gold HLA-DRDQ Low Res (Invitrogen, США) по протоколу производителя. Результаты типирования были проанализированы с помощью программного обеспечения UniMatch (Life Technologies, США).

5. Результаты и их обсуждение

Данная работа состоит из двух частей. Целью первой части исследования было выявление характерных черт репертуаров TCR β родственных пар мама-ребенок. Во второй части работы мы исследовали влияние старения человека на разнообразие репертуара TCR β . Кроме того, часть работы была посвящена разработке метода нормированного анализа репертуаров TCR с использованием молекулярного баркодирования.

5.1. Сравнительный анализ репертуаров TCR β матерей и детей

Взаимодействие иммунитетов матери и плода активно изучается в контексте осложнений в ходе беременности [112], аутоиммунных заболеваний [113]–[115], а также гаплоидентичных трансплантаций органов [116], [117] и гематопозитических стволовых клеток [118].

В нашем исследовании мы использовали метод массированного секвенирования для сравнения индивидуальных репертуаров TCR β матерей и их детей. Нами было получено 500 000 – 2 000 000 уникальных TCR β CDR3 клонотипов для каждого донора. Целью сравнительного анализа было выявление специфических характеристик, свойственных репертуарам TCR родственных пар, а также выяснение степени влияния наследственного фактора на формирование индивидуальных репертуаров. Кроме того, сравнивая клонотипы со сбитой рамкой считывания (то есть не подвергавшиеся селекции в тимусе) с клонотипами с нормальной рамкой считывания мы анализируем влияние отбора в тимусе на формирование индивидуального репертуара TCR β человека.

Помимо этого, достигнутая глубина секвенирования позволила нам исследовать наличие зрелых микрохимерных Т-лимфоцитов матери или плода, мигрировавших через плаценту в качестве зрелых α/β Т-клеток и сохранившихся в периферической крови обоих доноров.

В исследовании разнообразия репертуаров TCR β принимали участие три мамы и их дети (по два ребенка каждой из мам). Мы получили более 1×10^7 CDR3-содержащих

прочтений секвенирования (*ридов*) для каждой мамы и около 3×10^6 ридов – для каждого ребенка. Из полученных массивов последовательностей с помощью разработанного нами программного обеспечения MiTCR [109] было идентифицировано 500 000 – 2 000 000 разных клонотипов TCR β в образцах каждого из доноров (Таблица 1). Полученные индивидуальные репертуары TCR β (Рисунок 5.1) мы использовали в сравнительном анализе для выявления характерных черт родственных репертуаров.

5.1.1. Распределение частот использования V β сегментов

У человека существует 54 варианта генов V β -сегмента, вносящих большой вклад в общее индивидуальное разнообразие TCR β . Гены этого сегмента кодируют CDR1 и CDR2 участки TCR, взаимодействующие с молекулой МНС в составе комплекса пептид-МНС. Мы проанализировали относительное распределение использования семейств V β сегмента в репертуарах TCR β для пар мама-ребенок на трех уровнях:

1) Репертуары TCR β клонотипов со сбитой рамкой считывания

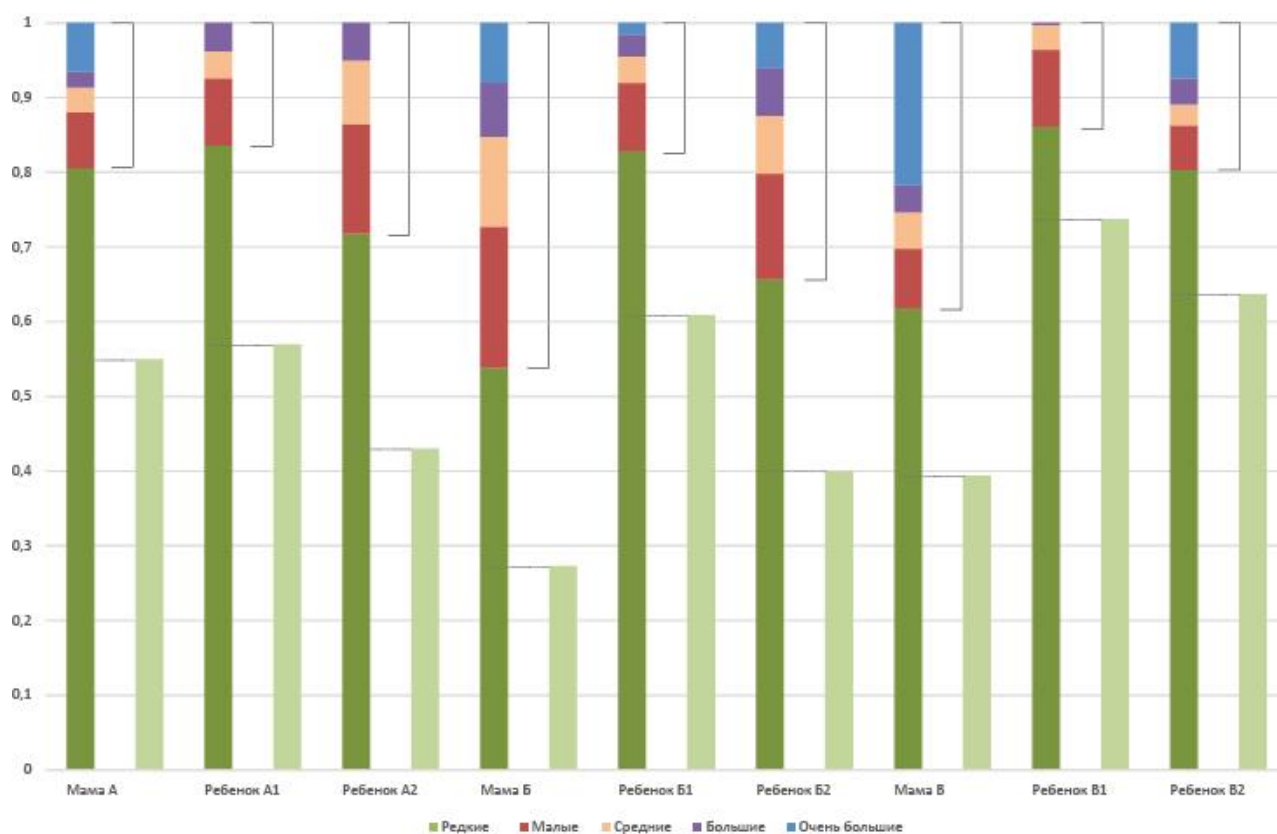
Поскольку варианты TCR β со сбитой рамкой считывания, присутствующие на гомологичной хромосоме примерно в 40% Т-лимфоцитов, не подвергаются селекции в тимусе, по их репертуару можно судить об особенностях исходной сборки TCR в ходе рекомбинационных событий [119]. Из-за действия механизма нонсенс-опосредованного разрушения мРНК (NMD, англ. nonsense mediated decay, см. раздел 2.1.2), в библиотеках TCR, созданных на основе РНК, варианты TCR β со сбитой рамкой считывания составляют очень низкий процент [99], [100], [102], [104]. Тем не менее, в проанализированных нами библиотеках такие TCR составили порядка 2,5%, то есть 16 000 – 45 300 клонотипов в каждом из образцов (Таблица 1). Этого количества было вполне достаточно для проведения статистического анализа частот использования различных V β сегментов на уровне первичной сборки TCR до селекции в тимусе.

Таблица 1. Результаты секвенирования: количество клонотипов и прочтений секвенирования TCRβ.

| Донор | Пол | Возраст | Общее количество ридов | CDR3 - содержащие риды TCRβ (%) | TCR бета клонотипы до коррекции ошибок | Конечные TCR бета клонотипы | Клонотипы со сбитой рамкой считывания, % от общего количества клонотипов | Наивные Т-клетки, % от всех Т-клеток (FACS анализ) | Клонотипы, занимающие >0.001% ридов, % от количества ридов |
|------------|-----|---------|------------------------|---------------------------------|--|-----------------------------|--|--|--|
| Мама А | Ж | 36 | 13,656,054 | 12,513,543 (91.6) | 2,044,290 | 1,708,037 | 2.7 | 55.0 | 19.4 |
| Мама Б | Ж | 43 | 13,872,805 | 12,901,795 (93.0) | 1,213,738 | 918,557 | 2.6 | 27.3 | 46.2 |
| Мама В | Ж | 43 | 11,167,059 | 10,038,463 (89.9) | 2,180,886 | 1,978,745 | 3.0 | 39.4 | 38.2 |
| Ребенок А1 | М | 11 | 4,687,578 | 2,889,352 (61.6) | 756,772 | 729,800 | 2.6 | 57.2 | 16.4 |
| Ребенок А2 | М | 9 | 4,202,419 | 2,467,388 (58.7) | 558,173 | 535,283 | 2.4 | 43.0 | 28.2 |
| Ребенок Б1 | М | 16 | 4,830,536 | 3,173,376 (65.7) | 821,908 | 790,592 | 3.2 | 60.9 | 17.2 |
| Ребенок Б2 | М | 10 | 4,615,093 | 3,009,961 (65.2) | 545,730 | 517,410 | 2.3 | 40.1 | 34.3 |
| Ребенок В1 | Ж | 6 | 6,081,365 | 3,940,367 (64.8) | 1,104,982 | 1,060,854 | 2.3 | 73.7 | 13.9 |
| Ребенок В2 | М | 16 | 4,564,646 | 3,008,748 (65.9) | 785,164 | 760,572 | 2.6 | 63.7 | 19.7 |

Таблица 2. Средние количества общих TCRβ CDR3 клонотипов родных и неродных пар мама-ребенок.

| Пары | Аминокислотные последовательности | | | Нуклеотидные последовательности | | | Нуклеотидные последовательности (Vβ- и Jβ- сегменты идентичны) | | |
|----------|-----------------------------------|----------------------------|---------------|---------------------------------|----------------------------|---------------|--|----------------------------|---------------|
| | Низко-частотные клонотипы | Высоко-частотные клонотипы | Все клонотипы | Низко-частотные клонотипы | Высоко-частотные клонотипы | Все клонотипы | Низко-частотные клонотипы | Высоко-частотные клонотипы | Все клонотипы |
| Родные | 71,714±25,890 | 117±34 | 83,257±26,056 | 4,938±2,237 | 69±7 | 8,407±3,396 | 2,015±893 | 40±10 | 3,456±1,359 |
| Неродные | 68,979±18,822 | 99±15 | 80,304±18,626 | 4,640±1,587 | 59±12 | 7,955±2,227 | 1,823±624 | 31±11 | 3,135±897 |



| Тип Клона | Размер Клона |
|---------------|--------------------|
| Редкий | Менее 0,001% |
| Малый | От 0,001% до 0,01% |
| Средний | От 0,01% до 0,1% |
| Большой | От 0,1% до 1% |
| Очень большой | Более 1% |

Рисунок 5.1. Представленность клонов Т-лимфоцитов разной частоты встречаемости в индивидуальных репертуарах TCRβ. Цветные столбцы соответствуют доле клонального пространства, занимаемой клонами определенной частоты встречаемости у каждого из девяти доноров. Светло-зеленые столбцы соответствуют доле наивных CD27^{high}CD45RA^{high} Т-клеток, измеренную в ходе анализа методом проточной цитофлуориметрии. Пунктирные линии отсекают долю низкочастотных TCRβ клонотипов, эквивалентную содержанию наивных Т-лимфоцитов в периферической крови каждого донора, использованную в анализе низкочастотных клонотипов. Квадратные скобки показывают долю высокочастотных клонотипов в образце каждого донора, соответствующую доле, занимаемой клонами Т-клеток, каждый из которых составляет более 0.001% от общего числа Т-клеток.

Сравнительный анализ представленности сегментов Vβ в репертуарах TCRβ со сбитой рамкой считывания для мам и детей показал, что распределение использования семейств Vβ-сегментов сопоставимо для родственных и неродственных пар мама-ребенок. У родственных пар мама-ребенок наблюдалось некоторое уменьшение

значения дивергенции Дженсена-Шеннона² (Рисунок 5.2 А, 5.3 А, Б, два первых столбца), что могло бы свидетельствовать о большем сходстве частот используемых сегментов, однако оно носило статистически недостоверный характер. То есть, по всей видимости, наследуемых в паре мама-ребенок характеристик рекомбинаторного механизма недостаточно для выраженного сходства в выборе V-сегментов при сборке TCR, либо требуется большая выборка для подтверждения этого факта. *ы*

2) Репертуары TCR β низкочастотных клонотипов

Влияние селекции в тимусе на распределение представленности V β -сегментов можно оценить, сравнив репертуар TCR β со сбитой рамкой считывания с репертуаром TCR β наивных Т-клеток. Это связано с тем, что TCR с неправильной рамкой считывания не проходят селекцию в тимусе – в отличие от зрелых наивных Т-лимфоцитов (см. разд. Развитие Т-лимфоцитов в тимусе).

Поскольку мы стремились к максимальной глубине анализа, в данной работе мы не использовали метод проточной цитофлуориметрии (FACS, англ., fluorescence activated cell sorting) для сортировки наивных Т-лимфоцитов периферической крови доноров (с тем чтобы избежать потери клеток или РНК и возможных количественных искажений). Пул клонотипов, большинство из которых является наивными Т-лимфоцитами, мы определяли следующим образом.

С помощью FACS-анализа мы определяли долю наивных Т-лимфоцитов (CD3⁺CD27^{high}CD45RA^{high}) в образцах периферической крови каждого донора. Такой способ определения наивных Т-лимфоцитов был описан в работе [121]. В образцах детей наивные Т-лимфоциты составляли 40–73% от общего количества Т-клеток, в образцах матерей – 27–55% (Рисунок 5.1). В крови человека наивные Т-клоны, как правило, представлены минимальным количеством идентичных по TCR Т-клеток [122].

² Дивергенция Дженсена-Шеннона – статистический метод для определения близости двух распределений вероятностей (см. разд. Методы)

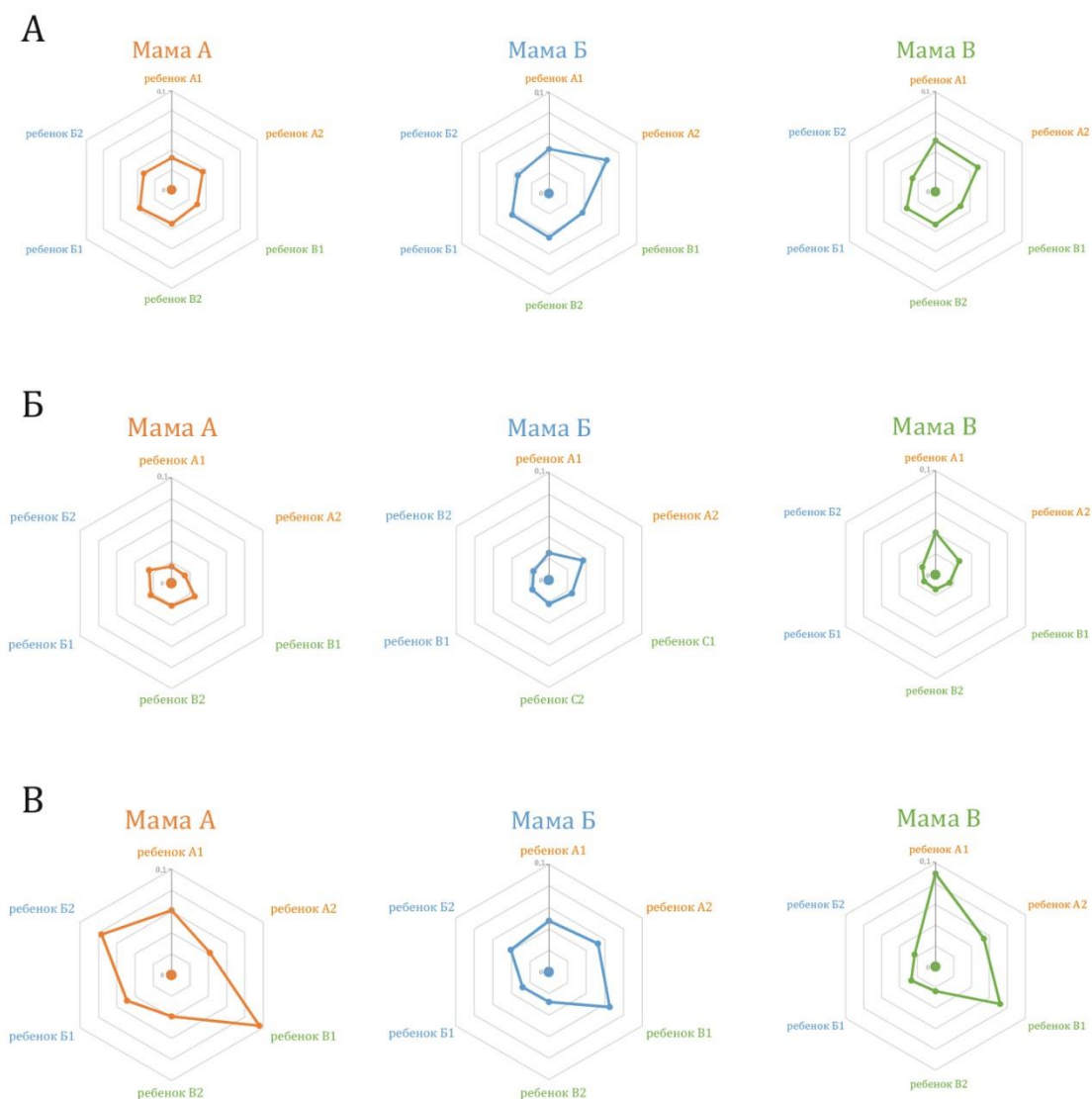


Рисунок 5.2. Сравнительный анализ распределения использования V β -сегментов. Для анализа используются значения дивергенции Дженсена-Шеннона. (А) TCR β клонотипы со сбитой рамкой считывания; (Б) Низкочастотные TCR β клонотипы с правильной рамкой считывания; (В) Высокочастотные TCR β клонотипы с правильной рамкой считывания. Центральная точка каждой из диаграмм соответствует одной из мам (оранжевая – маме А, голубая – маме Б и зеленая – маме В). Остальные точки на каждой из диаграмм соответствуют шести детям. Чем ближе точка ребенка находится к точке мамы, тем ниже значение дивергенции Дженсена-Шеннона (и, соответственно, тем более схожа представленность V β -сегментов).

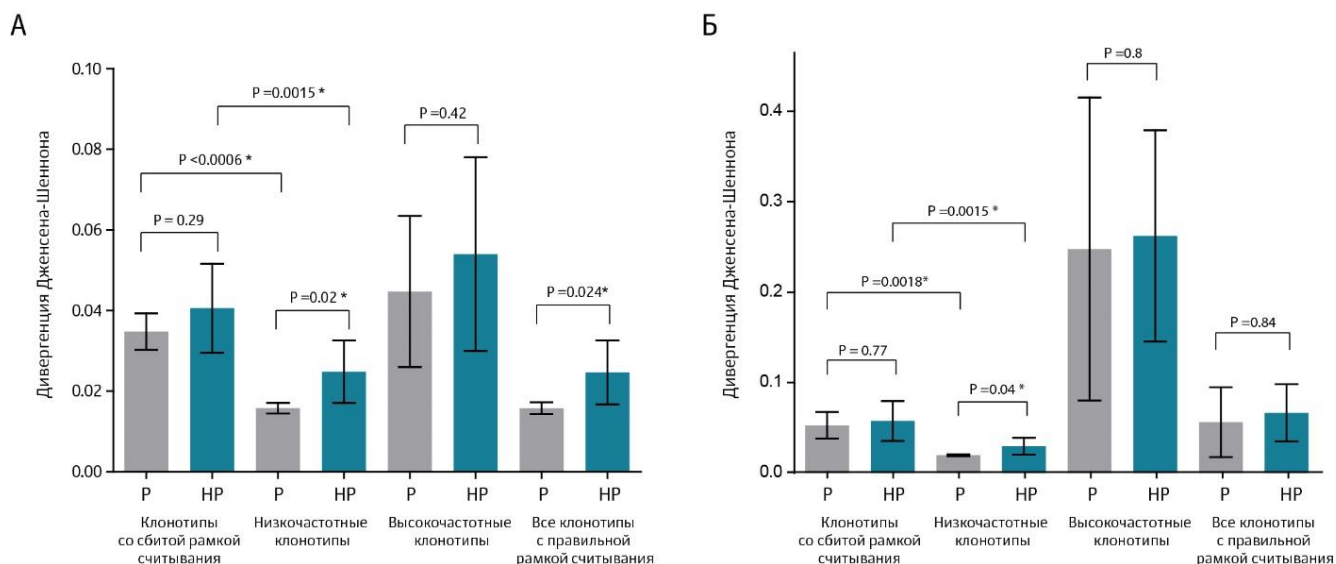


Рисунок 5.3. Средние значения дивергенции Дженсена-Шеннона представленности V β -сегментов в репертуарах TCR родственных (P, серые столбцы) и неродственных (NP, синие столбцы) пар мама-ребенок (с указанием стандартного отклонения). P-значения с поправкой Бонферрони представлены с учетом множественных выборок. (А) Сравнение представленности V β -сегментов, основанное на количестве TCR β клонотипов (но не сиквенсов – то есть влияние относительных размеров исследуемых клонотипов TCR β не учитывается). (Б) Сравнение представленности V β -сегментов, основанное на количестве сиквенсов – то есть с учетом относительных размеров исследуемых клонотипов TCR β . Чем меньше значение дивергенции Дженсена-Шеннона, тем ближе распределение использования V β -сегментов в репертуарах TCR исследуемых пар. * – различия статистически достоверны.

Соответственно, мы предположили, что фракция низкочастотных клонотипов, составляющая ту же долю гомеостатического пространства, что и определенная с помощью FACS-анализа доля наивных Т-клеток в крови этого донора, включает в себя преимущественно наивные Т-клетки (433 293 – 1 797 650 клонотипов в каждом доноре, Рисунок 5.1). При этом учитывались только клонотипы с правильной рамкой считывания TCR.

Распределение использования V β -сегментов на этом уровне рассмотрения было значительно более сходным у всех доноров, чем на уровне клонотипов со сбитой рамкой считывания (Рисунок 5.2 Б, 5.3 А, Б). Это происходит, вероятнее всего, вследствие дальнейшего формирования репертуара наивных Т-лимфоцитов под действием их отбора в тимусе. Его влияние приводит к выраженной конвергенции частот использования различных V β -сегментов как в родственных (P = 0.0006), так и в неродственных (P = 0.0015) парах. То есть селекция в тимусе формирует индивидуальные

репертуары TCR на основе полученных в ходе рекомбинации вариантов, действуя схоже в разных людях.

Родственные пары мама-ребенок гаплоидентичны (то есть имеют 50% одинаковых аллелей HLA). Поскольку в V-сегментах закодированы участки TCR, взаимодействующие с молекулой MHC [29], [123], можно предположить, что представленность V β -сегментов в репертуарах функциональных TCR на уровне наивных Т-лимфоцитов (после позитивной и негативной селекции в тимусе) более сходна в родственных парах мама-ребенок. Она, действительно, оказалась достоверно ($P = 0.02$) более сходной у родственных, чем у неродственных пар (Рисунок 5.3 А, Б – столбцы 3 и 4).

В соответствии с анализом Дженсена-Шеннона, в репертуарах TCR β родственных групп доноров заметна тенденция к выравниванию частот использования определенных V β -сегментов в низкочастотных клонотипах по сравнению с клонотипами со сбитой рамкой считывания (Рисунок 5.4). То есть, например, в каждой группе (семьях А, Б и В) мы наблюдали тенденцию к сближению уровня представленности V β -сегментов 12-3, 12-4, 20-1, 21-1 и 23-1 в пуле низкочастотных клонотипов. Мы, кроме того, наблюдали одновременное падение представленности V β -сегмента 7-3 в репертуаре низкочастотных клонотипов каждого донора семей А и В, а также одновременное увеличение представленности V β -сегмента 28 в семье Б.

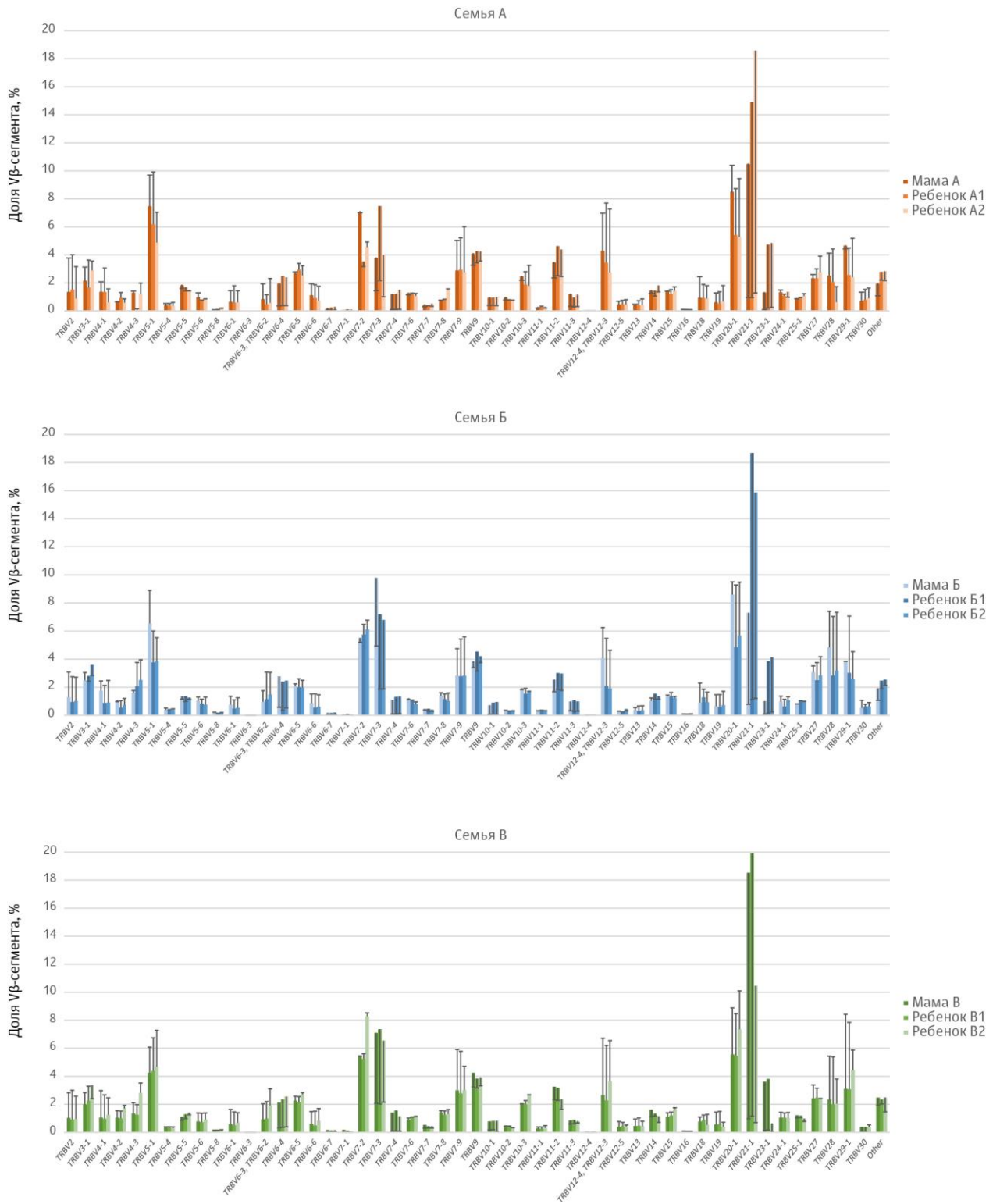


Рисунок 5.4. Распределение частот использования Vβ-сегментов в субпопуляции функциональных низкочастотных TCRβ клонотипов в сравнении с вариантами TCRβ со сбитой рамкой считывания. Цветные столбцы соответствуют представленности определенного сегмента Vβ в репертуаре TCR со сбитой рамкой считывания для каждого донора. Линии показывают изменения представленности каждого из сегментов при переходе к репертуару функциональных низкочастотных TCRβ клонотипов. **(А)**, **(Б)** и **(В)** отображают распределение частот использования Vβ сегментов родственных доноров семьи А, Б и В, соответственно.

Интересно, что наблюдаемые изменения были в целом схожи и для неродственных доноров (сравните Рисунок 5.4 А-В). Зависимость степени сближения частоты использования V β -сегментов от количества общих аллелей HLA оказалась недостоверной ($R = 0.12$, $P = 0.63$). Значит, по всей вероятности, тимус оказывает значительное и схожее для разных индивидуумов давление на формирование наивного репертуара TCR β .

3) Репертуары TCR β высокочастотных клонотипов

Влияние встречи Т-лимфоцитов со специфичным антигеном на разнообразие используемых V-сегментов может быть отслежено при сравнении репертуаров TCR наивных и активированных Т-лимфоцитов.

Следуя вышеприведенной логике, мы предположили, что наиболее часто встречающиеся клоны в периферической крови донора состоят преимущественно из активированных Т-клеток. В эту популяцию мы включили клонотипы, количество сиквенсов которых составляет более 0.001% от всех CDR3-содержащих сиквенсов. Таким образом, нижняя граница этой группы приблизительно на порядок превосходит верхнюю границу группы низкочастотных клонотипов каждого из доноров (Рисунок 5.1). Такое разграничение позволяет в значительной степени избежать “засоров” наивными Т-клонотипами группы высокочастотных клонотипов. Тем не менее, следует оговориться, что последняя группа может содержать некоторое количество наивных клонотипов с CDR3 низкой сложности, часто продуцируемых в тимусе за счет конвергентной рекомбинации (общественных клонотипов) и следовательно высоко представленных в периферической крови [51].

Определенная таким образом субпопуляция высокочастотных клонотипов составляет 13.9 – 46.2% гомеостатического Т-клеточного пространства для образца каждого из доноров (2 803 – 8 285 клонотипов на образец). Эта субпопуляция характеризуется повышенным разнообразием частот использования V β -сегментов, представленных в TCR разных доноров. Достоверно большего сходства распределения частот использования V β -сегментов в репертуарах родственных пар по сравнению с неродственными обнаружено не было.

Такая гетерогенность представленности V β -сегментов в репертуарах TCR разных людей, вероятно, связана с тем, что основным фактором, формирующим этот репертуар,

является специфичность к конкретным антигенам (а не взаимодействие V β -сегментов с МНС). Относительная случайность встречи того или иного специфического клона с антигеном и последующая экспансия таких клонов размывает сходство использования семейств V-сегментов в репертуарах TCR разных людей, наблюдаемое в наивных Т-лимфоцитах.

Необходимо еще раз отметить, что приведенный выше анализ относится к низко- и высоко- частотным клонотипам, которые не совпадают полностью с наивными и эффекторными Т-лимфоцитами/Т-клетками памяти периферической крови доноров. Было показано, что вследствие рекомбинантной конвергенции и общей неравномерности частот событий рекомбинации некоторые наивные клонотипы формируются чаще других [50], [51], [106]. В то же время, некоторые Т-клетки памяти иногда присутствуют в периферической крови в низких концентрациях. Кроме того, в этих же исследованиях было показано, что репертуары наивных Т-лимфоцитов и Т-клеток памяти в некоторой степени перекрываются, то есть один и тот же TCR β -клонотип может оказаться как в составе наивного Т-лимфоцита, так и в составе Т-клетки памяти. Тем не менее, такие клонотипы относительно редки и не вносят существенного вклада в оценку основных тенденций представленности V-сегментов в репертуарах TCR наивных и эффекторных Т-лимфоцитов.

5.1.2. Пересечения репертуаров TCR β CDR3 в родственных и неродственных парах мама - ребенок

Многие исследования последних лет выявили значительные межиндивидуальные пересечения репертуаров TCR β [50], [51], [58], [60]–[62]. Тем не менее, до сих пор неизвестно, имеет ли место более существенное перекрытие репертуаров TCR гаплоидентичных доноров (половина аллелей HLA которых идентична).

Чтобы ответить на этот вопрос, мы провели сравнительный анализ репертуаров TCR β родных и неродных пар мама-ребенок на трех уровнях:

- 1) на уровне аминокислотной последовательности CDR3,
- 2) на уровне нуклеотидной последовательности CDR3,
- 3) на уровне нуклеотидной последовательности CDR3 при совпадающих V- и J-сегментах (т.е. полностью идентичных TCR β цепей).

Мы оценили такие пересечения отдельно для низкочастотных и высокочастотных клонотипов, а также для всех клонотипов с правильной рамкой считывания. В Таблице 2 приведены ненормализованные средние количества общих вариантов CDR3 для родных и неродных пар мама-ребенок.

Для корректного сравнения степени перекрытия репертуаров TCR в разных субпопуляциях, мы нормализовали количество совпадающих CDR3 вариантов на объемы сравниваемых выборок по формуле: [нормализованное количество общих TCRβ клонотипов выборок А и Б] = [количество общих TCRβ клонотипов выборок А и Б]/([количество клонотипов в выборке А]×[количество клонотипов в выборке Б]). Нормализованные результаты пересечений представлены на Рисунке 5.5.

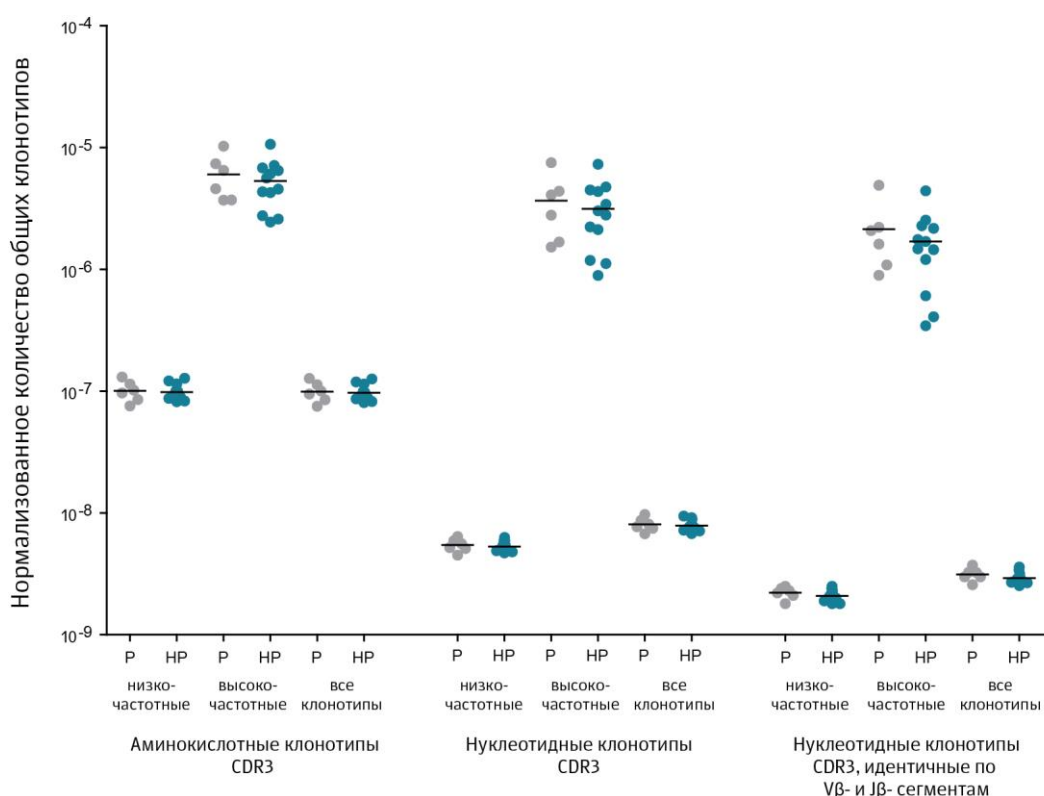


Рисунок 5.5. Нормализованное количество общих клонотипов индивидуальных TCRβ репертуаров. Пересечения показаны на уровне аминокислотных последовательностей CDR3, нуклеотидных последовательностей и нуклеотидных последовательностей с идентичными Vβ-и Jβ-сегментами – для родственных (серые круги) и неродственных (синие круги) пар мама-ребенок, для низкочастотных, высокочастотных и всех клонотипов с правильной рамкой считывания. Количество пересечений было нормализовано (см. в тексте). Р – родственные пары, НР – неродственные пары.

Наибольшая степень пересечения репертуаров TCR β CDR3 у всех рассматриваемых пар доноров наблюдается в субпопуляции высокочастотных клонотипов. Это наблюдение согласуется с предыдущими данными [51], и может быть следствием сочетания двух параметров. Во-первых, наивные клонотипы с упрощенной структурой CDR3 продуцируются в тимусе с повышенной частотой – как внутри одного индивидуума, так и в рамках всей популяции (см. раздел 2.1.6). Во-вторых, поскольку все изучаемые нами доноры живут в одной и той же среде, их иммунная система сталкивается в основном с одними и теми же антигенами. Отбор на уровне специфичности к этим антигенам повышает вероятность клональной экспансии схожих клонов в гомеостатическом пространстве разных доноров.

На всех уровнях сравнения пересечений репертуаров CDR3 нормализованное количество общих клонотипов было выше для родных пар, по сравнению с неродными, однако различия не достигали статистически значимых величин. Это наблюдение согласуется с результатами, полученными ранее в работе [50], выполненной на CD8+ клетках периферической крови трех доноров. В этой работе не было обнаружено влияния количества общих аллелей HLA на степень перекрывания репертуаров CDR3 TCR β наивных CD8+ клеток. В нашей работе нам удалось провести более аккуратное сравнение большей когорты родственных доноров. Однако, несмотря на повышенную точность использованного нами метода и значительно более глубокое секвенирование каждого из доноров, мы наблюдали лишь небольшую тенденцию к сближению родственных репертуаров TCR β .

5.1.3. Одинаковые аминокислотные последовательности CDR3 чаще сочетаются с одинаковым V β -сегментом в родных парах доноров

Известно, что участок CDR3 преимущественно взаимодействует с пептидом антигена, в то время как участки CDR1 и CDR2, последовательность которых закодирована в V-сегменте, преимущественно взаимодействуют с молекулой MHC [29], [123]. Участки некоторых V β -сегментов, участвующие в формировании CDR3, очень похожи друг на друга. Поэтому разные V β -сегменты могут формировать идентичные аминокислотные последовательности CDR3.

В донорах с одинаковыми или похожими аллелями HLA могут быть избирательно активированы клоны с одинаковым V β -сегментом и аминокислотной

последовательностью CDR3, узнающие один и тот же комплекс пептид-МНС [26]. Так как родственные пары мама-ребенок имеют как минимум 50% общих HLA-аллелей, можно ожидать, что активированные клоны с одинаковой аминокислотной последовательностью CDR3 (то есть узнающие один и тот же антиген) будут чаще иметь общий V β -сегмент, кодирующий CDR1 и CDR2 участки, отвечающие за узнавание МНС.

Для проверки этой гипотезы мы проанализировали репертуары, состоящие из 10 000 наиболее часто встречающихся аминокислотных CDR3 клонотипов каждого из доноров. Затем мы вычислили их пересечения в родственных и неродственных парах, а также пересечения аминокислотных CDR3 клонотипов, имеющих одинаковый V β -сегмент (т.е. идентичные CDR1, CDR2, и CDR3 участки TCR β).

Мы определяли соотношение количества аминокислотных клонотипов с общими CDR1, CDR2, CDR3 участками к количеству аминокислотных клонотипов с общим CDR3 участком (соотношение V β -CDR3/CDR3). Это соотношение было всегда выше в родственных парах (в 1,3 раза, ± 0.16 , Рисунок 5.6 А). Кроме того, мы наблюдали достоверную положительную корреляцию этого соотношения и количества общих аллелей МНС-I у анализируемых доноров ($R = 0.62$, $P < 0.006$, Рисунок 5.6 Б, Таблица 3).

Таблица 3. Результаты HLA-типирования доноров.

| | HLA-A | HLA-B | HLA-C | DRB1 | DRB3 | DRB4 | DQB1 |
|-------------------|--------|--------|-----------|--------|-------|-------|--------|
| Мать А | 02, 24 | 15, 57 | 03, 06 | 07, 14 | 01-03 | 01 | 03, 05 |
| Ребенок А1 | 02, 25 | 15, 18 | 03, 12 | 04, 14 | 01-03 | 01-02 | 03, 05 |
| Ребенок А2 | 02, 02 | 15, 44 | 03, 05 | 11, 14 | 01-03 | - | 03, 05 |
| Мать Б | 02, 23 | 44, 51 | 02, 04 | 07, 11 | 01-03 | 01-02 | 02, 03 |
| Ребенок Б1 | 02, 23 | 38, 44 | 04, 12 | 07, 13 | 01-03 | 01-02 | 02, 06 |
| Ребенок Б2 | 02, 23 | 27, 44 | 02, 04 | 07, 11 | 01-03 | 01-02 | 02, 03 |
| Мать В | 01, 11 | 08, 35 | 04/08, 07 | 01, 03 | 01-03 | - | 02, 05 |
| Ребенок В1 | 02, 11 | 35, 38 | 04, 12 | 01, 13 | 01-03 | - | 05, 06 |
| Ребенок В2 | 02, 11 | 35, 38 | 04, 12 | 01, 13 | 01-03 | - | 05, 06 |

Для анализа зависимости между соотношением пересечений V β -CDR3/CDR3 и количеством общих аллелей главного комплекса гистосовместимости I (МНСI) мы использовали модель линейной регрессии. Линейная модель (V β -CDR3/CDR3 соотношение = $b_0 + b_1 \times$ [количество общих аллелей МНСI]) была фиттирована методом наименьших квадратов. Для оценки корреляции был использован коэффициент Пирсона.

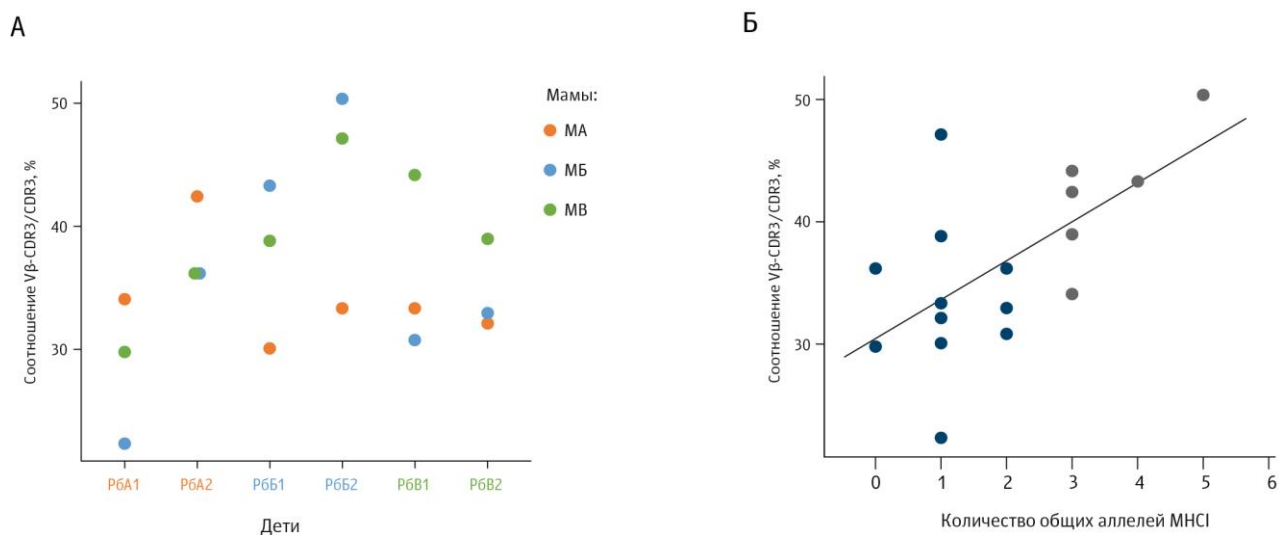


Рисунок 5.6. Соотношение пересечений аминокислотных последовательностей Vβ-CDR3/CDR3. (А) Соотношение пересечений аминокислотных последовательностей Vβ-CDR3/CDR3 для каждой пары мама-ребенок, основанное на сравнении 10 000 наиболее часто встречающихся клонотипов в каждом из образцов. **(Б)** Корреляция между количеством общих аллелей МНС1 и соотношением пересечений аминокислотных последовательностей Vβ-CDR3/CDR3 для 10 000 наиболее часто встречающихся клонотипов в каждом из доноров. Сплошная линия представляет собой модель линейной регрессии, линейный коэффициент корреляции составляет 0.62 ($P < 0.006$). Родственные и неродственные пары показаны серым или черным цветом, соответственно.

Это наблюдение показывает, что для узнавания определенного комплекса пептид-МНС часто необходима полная конвергенция TCRβ цепи, следствием которой является повышенная доля совпадающих по Vβ-сегменту идентичных последовательностей CDR3 гаплоидентичных доноров. Несмотря на то, что сравнение проводилось на несортированных Т-лимфоцитах, вклад CD8⁺ клеток был достаточным для обеспечения корреляции. Вероятно, наблюдаемая корреляция была бы выше при анализе сортированной субпопуляции CD8⁺ Т-клеток.

5.1.4. Селекция в тимусе снижает среднюю длину CDR3

Сравнение репертуаров CDR3 с правильной и со сбитой рамкой считывания показало, что у всех анализируемых доноров последние характеризуются большей средней длиной CDR3 участка ($45,6 \pm 0,4$ против $43,3 \pm 0,2$) и большим количеством добавленных нуклеотидов ($8,6 \pm 0,2$ против $7,4 \pm 0,1$; Рисунок 5.7 А).

Таким образом, репертуар TCRβ CDR3, возникающий в результате соматической рекомбинации (параметры которого отражены в репертуаре клонотипов со сбитой

рамкой считывания), в среднем характеризуется большей длиной. Последующая же селекция в тимусе приводит к уменьшению его средней длины и количества добавленных нуклеотидов.

5.1.5. В поисках зрелых микрохимерных клонов Т-клеток

На сегодняшний день считается, что в ходе беременности материнские клетки и клетки плода активно переходят через плаценту и сохраняются в организме нового хозяина в течение десятков лет после рождения [124], [125].

Явление *микрохимеризма* (присутствие в организме одного индивидуума небольшого числа клеток другого индивидуума) обнаружено в популяциях разных клеток крови плода и матери, включающих Т- и В-лимфоциты, гранулоциты, дендритные клетки, моноциты и макрофаги [126]–[129].

Считается, что, как правило, предшественники лимфоцитов переходят через плаценту и затем их потомки частично заселяют гомеостатическое пространство нового организма [127], [129]–[131]. Однако некоторые исследования показывают, что зрелые Т-лимфоциты также могут мигрировать через плаценту [128]–[130], [132].

Однако до сих пор неизвестно, могут ли зрелые микрохимерные Т-лимфоциты сохраняться в крови нового индивидуума и работать, как функциональные клоны. Большинство исследований микрохимеризма проводилось с помощью отслеживания мужской геномной ДНК в материнских тканях, или ненаследуемой аллели HLA в тканях плода [127], [133]. С помощью этих методов, тем не менее, невозможно отследить судьбу определенного материнского клона Т-клеток в крови ребенка, или клона Т-клеток плода в крови матери.

Мы предположили, что проводимый нами анализ методом глубокого секвенирования значительной порции репертуаров TCR β матери и ребенка может выявить присутствие клонов Т-клеток, перешедших через плаценту в зрелом состоянии и давших начало функциональной популяции Т-клеток в новом хозяине.

С помощью нашего метода, тем не менее, невозможно определить направление передачи клонов (от матери к ребенку или наоборот), поскольку микрохимерные клоны, перешедшие через плаценту в зрелом состоянии, характеризовались бы идентичными последовательностями TCR.

Потенциальным маркером таких клеток могли бы быть полностью совпадающие V- и J- сегменты в сочетании с идентичными нуклеотидными последовательностями CDR3. Однако около 40% общих клонотипов любых двух анализируемых доноров имеют одинаковые V- и J- сегменты. В анализируемых нами парах было обнаружено от 1 766 до 5 410 таких клонотипов (Рисунок 5.5). Такое большое количество полностью совпадающих клонотипов делает невозможным использование V β -CDR3-J β участка в качестве ключевого параметра для выделения зрелых микрохимерных клонов среди общественных клонов, возникших независимо в разных людях в результате конвергентной рекомбинации ([51], см. разд. “Общественные клонотипы”). Таким образом, даже если зрелые микрохимерные клоны присутствуют в периферической крови исследуемых доноров, они теряются среди намного большего количества общественных клонотипов. Значит, необходимы дополнительные характеристики таких клонотипов.

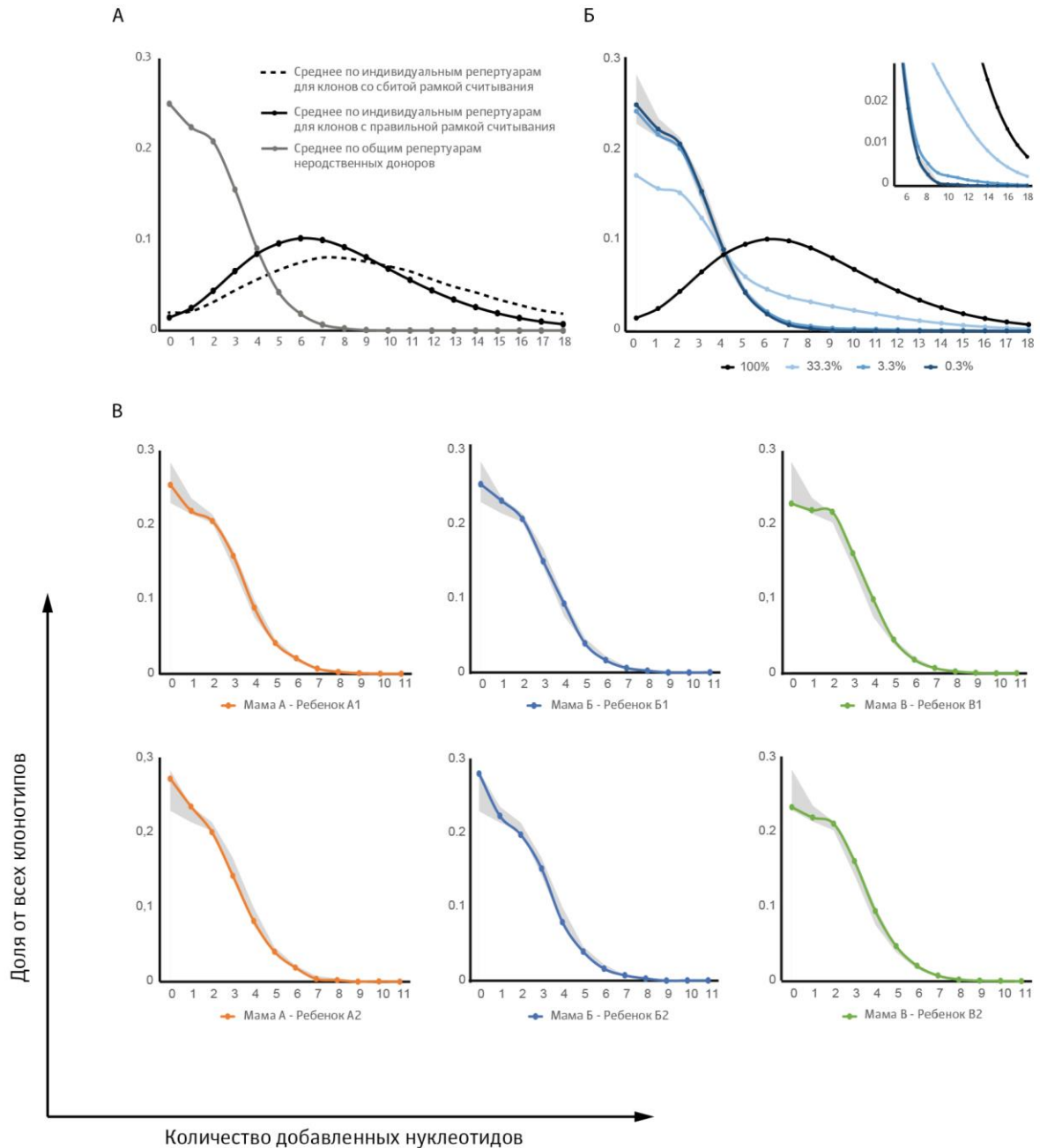


Рисунок 5.7. Кривые распределения добавленных нуклеотидов. (А) Усредненные кривые распределения добавленных в CDR3 нуклеотидов индивидуальных репертуаров TCR β (с правильной рамкой считывания – черная сплошная линия, со сбитой рамкой считывания – пунктирная линия) и для TCR β клонотипов, общих для пар неродственных доноров (серая сплошная линия). (Б) Моделирование кривых распределения добавленных в CDR3 нуклеотидов TCR β клонотипов, совпадающих по V β -CDR3-J β сегментам в родных парах мама-ребенок, – в зависимости от доли зрелых микрохимерных клеток. Эти модели были получены на основе серой (эквивалентной близкому к нулевому вкладу зрелых микрохимерных клеток в общий репертуар) и черной (эквивалентной 100%-ному вкладу зрелых микрохимерных клеток в общий репертуар) сплошных кривых (А), сочетаемых в разных пропорциях. Линии представляют собой модельный репертуар TCR β , вклад зрелых микрохимерных клонов в который составляет 100%, 33%, 3.3% или 0.3% от общего количества клонотипов. Затемненное поле отображает разброс кривых общих репертуаров неродственных доноров. В правом верхнем углу находится увеличенное изображение. (В) Кривые распределения добавленных в CDR3 нуклеотидов TCR β клонотипов, идентичных по V β -CDR3-J β нуклеотидным последовательностям и общих для родных пар мама-ребенок. Затемненные поля отображают разброс аналогичных кривых общих репертуаров неродных пар.

В ряде работ было показано, что большинство общественных клонотипов характеризуется малым количеством добавленных нуклеотидов в CDR3 участке [50], [51], [62]. Поэтому в качестве дополнительного параметра мы выбрали количество добавленных нуклеотидов, как характеристику вероятности независимого возникновения данной последовательности CDR3 в разных людях [62]. Сравнение по этому параметру индивидуальных репертуаров TCR β и репертуаров TCR β , общих для клонотипов любых двух доноров, действительно, показало значительно меньшее количество добавленных букв у последних (Рисунок 5.7 А).

Перенос зрелых Т-клеток через плаценту очень маловероятно зависит от длины CDR3 и количества добавленных нуклеотидов. Кроме того, было показано, что у человека среднее количество добавленных букв в CDR3 репертуаре TCR β периферической крови взрослого человека и пуповинной крови достоверно не различается [134]. Поэтому эта характеристика должна быть одинакова как для индивидуальных репертуаров TCR, так и для репертуаров микрохимерных Т-клонов (в отличие от репертуаров общественных TCR). То есть, теоретически, содержание зрелых микрохимерных клонов Т-клеток в общих репертуарах TCR β родственных пар мама-ребенок будет отражено в кривой распределения количества добавленных нуклеотидов по клонотипам – пропорционально количеству микрохимерных клонов в образце (Рисунок 5.7 Б).

Чувствительность этого метода ограничена естественной дисперсией кривых распределения количества добавленных в CDR3 нуклеотидов, характеризующих перекрывание репертуаров TCR β двух людей. Например, если зрелые микрохимерные клоны Т-клеток составляют ~0.3% от общих Т-клонов, совпадающих по V β -CDR3-J β сегментам, пары мама-ребенок (то есть ~10 из 3 000 общих клонотипов, входящих в ~1x10⁶ индивидуальных клонотипов, обнаруженных в крови каждого донора в результате секвенирования), то по форме кривой распределения количества добавленных нуклеотидов невозможно было бы определить наличие микрохимерных клонов. Т.е. такая доля зрелых микрохимерных клонов периферической крови находится за пределами чувствительности выбранного нами метода.

Если же в одной паре родственных доноров зрелые микрохимерные клоны составляют 100 из 3 000 общих клонотипов (~3.3%), кривая распределения количества добавленных нуклеотидов отчетливо отражала бы присутствие таких клонов (Рисунок

5.7 Б). Такая доля зрелых микрохимерных клонотипов является примерной границей чувствительности выбранного нами метода.

Анализ кривых распределения количества добавленных нуклеотидов для всех шести исследуемых родственных пар не показал наличия зрелых микрохимерных клонов Т-клеток в периферической крови доноров (Рисунок 5.7 В). Соответственно, среднее количество добавленных нуклеотидов общих клонов Т-клеток в родных и неродных парах было одинаковым.

Описанное выше сравнение кривых распределения количества добавленных нуклеотидов производилось на уровне *клонотипов* TCR β – а не *количества сиквенсов*. Таким образом, представленность каждого отдельного клона в крови не учитывалась.

Тем не менее, при сравнении кривых распределения количества добавленных нуклеотидов на уровне количества прочтений секвенирования (то есть при анализе с учетом относительного размера каждого клона) были получены похожие, хотя и более шумные, результаты (данные не приводятся).

В ходе проведенного нами сравнительного анализа репертуаров TCR β на достигнутом уровне чувствительности метода (порядка 100 зрелых микрохимерных клонов на 10^6 анализируемых клонов TCR β) мы не наблюдали присутствия зрелых микрохимерных клонов в крови родных доноров. Тем не менее, это не исключает присутствия таких клеток в количествах ниже, чем чувствительность нашего метода, в небольшом количестве индивидуумов или при таких патологических состояниях, как вызванные микрохимеризмом аутоиммунные заболевания.

Стоит отметить, что в связи с высокой долей общественных клонов в периферической крови человека, методы глубокого секвенирования индивидуальных репертуаров TCR β оказываются недостаточными для обнаружения зрелых микрохимерных клонов. Однако дальнейшее развитие этих методов – особенно в комбинации с использованием разработанного нами метода эмульсионного ПЦР с живых клеток для определения пар TCR α и β цепей [135], в связке с HLA – будет способствовать изучению этого явления.

Таким образом, мы не обнаружили значительных различий в общих нуклеотидных вариантах V β -CDR3-J β родственных и неродственных пар мама-ребенок, позволивших бы нам говорить о субпопуляции зрелых микрохимерных клонов Т-клеток, перешедших

в ходе беременности и присутствующих в периферической крови нового хозяина в течение многих лет.

5.1.6. Оценка пересечения полных индивидуальных TCR β репертуаров

Нижняя граница пересечения разнообразия индивидуальных репертуаров TCR β CD8+ наивной субпопуляции Т-лимфоцитов любых двух доноров по существующим на сегодняшний день примерным оценкам составляет порядка 14 000 идентичных аминокислотных последовательностей CDR3 [50]. Эта оценка основана на данных по 200 000 – 600 000 индивидуальных TCR β клонотипов. Обладая данными об 1-2 миллионах индивидуальных клонотипах, мы решили провести собственную оценку степени пересечения полных репертуаров TCR β двух индивидуумов.

Мы сравнивали репертуары TCR β 12 неродных пар мама-ребенок. Увеличивая количество входящих в анализ уникальных клонотипов, выбираемых случайным образом из каждого образца, мы анализировали число идентичных вариантов CDR3, встречающихся в этих выборках (Рисунок 5.8). При увеличении объемов выборок в арифметической прогрессии количество общих клонотипов в каждой паре росло геометрически (Рисунок 5.8 А–В). При максимальной глубине анализа (порядка миллиона клонотипов на человека) было обнаружено 72 000 клонотипов, идентичных на аминокислотном уровне, 68 000 клонотипов, идентичных на аминокислотном уровне при условии несовпадения нуклеотидных последовательностей, и 6 000 клонотипов, идентичных на нуклеотидном уровне. Такое количество общих клонотипов в несколько раз превосходит предыдущие оценки пересечения разнообразия репертуаров человека [50]. Самая большая степень перекрытия репертуаров была показана для образцов, количество идентифицированных клонотипов в которых составляло 1×10^6 и 1.7×10^6 . Количество одинаковых аминокислотных клонотипов в этих образцах составляло 113 000; аминокислотных клонотипов с несовпадающими нуклеотидными последовательностями – 108 000; идентичных нуклеотидных клонотипов - 11 000.

Нижняя граница разнообразия репертуара TCR β человека в предыдущей работе была оценена как 5×10^6 уникальных клонотипов [6]. Учитывая это, мы экстраполировали полученные нами кривые пересечения, аппроксимируя их с помощью степенной функции: $Y = aX^b$, аналогично [50]. Коэффициент b при этом был близок к 2.0, а $R^2 > 0.999$ во всех проанализированных случаях (Рисунок 5.8 А–В, пунктирные кривые). Наша

приблизительная оценка полного перекрытия CDR3 репертуаров TCR β двух индивидуумов составляет порядка 2 200 000 аминокислотных последовательностей CDR3, 2 060 000 аминокислотных последовательностей CDR3 с несовпадающими нуклеотидными последовательностями и 180 000 нуклеотидных последовательностей CDR3, что составляет 44.1%, 41.3% и 3.6% индивидуальных репертуаров, соответственно.

Таким образом, несмотря на то, что общее потенциальное разнообразие составляет порядка 5×10^{11} аминокислотных вариантов TCR β [50], пересечение репертуаров любых двух здоровых доноров составляет практически половину репертуара каждого из них.

В приготовлении библиотек TCR β периферической крови доноров мы использовали большое количество мер предосторожности для защиты от контаминаций между образцами, включая секвенирование сравниваемых образцов в ходе двух независимых запусков Illumina. Однако даже в присутствии некоторого количества засоров, они бы не могли повлиять на количество общих аминокислотных вариантов CDR3, последовательность которых не совпадает на нуклеотидном уровне (Рисунок 5.8 Б). Кроме того, мы проводили экстракцию CDR3 и коррекцию ошибок с помощью программы MiTCR [109] с использованием наиболее строгих алгоритмов избавления от ошибок, элиминирующих 98% ошибок в ходе ПЦР и процесса секвенирования с минимальной потерей естественного разнообразия TCR β [104], [109] (см. раздел 2.2.5). Таким образом, полученные нами результаты достаточно однозначно свидетельствуют о необычайно высоком пересечении индивидуальных репертуаров TCR β .

Такая высокая степень пересечения свидетельствует о существовании довольно ограниченного набора наиболее часто продуцируемых тимусом аминокислотных последовательностей TCR β CDR3.

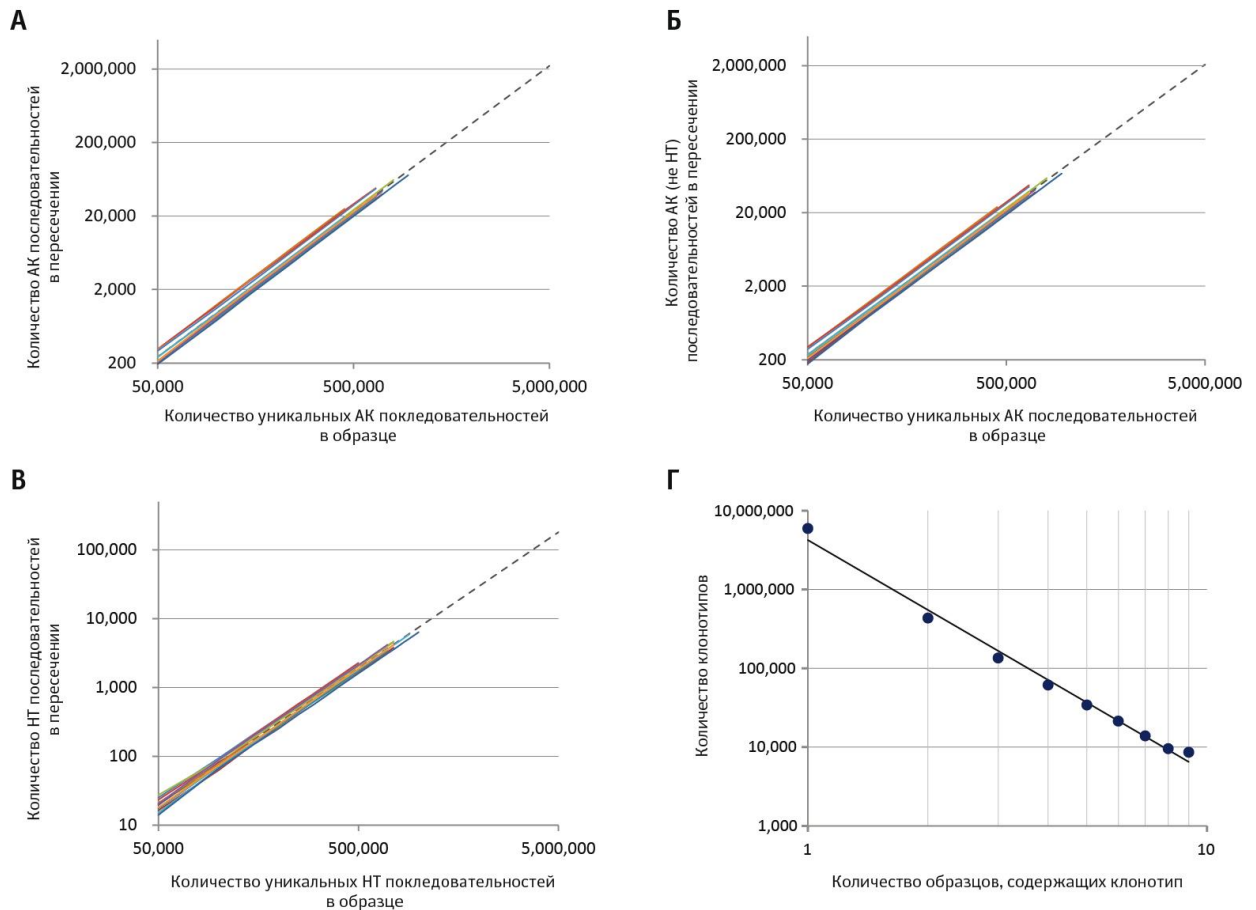


Рисунок 5.8. Перекрывание индивидуальных TCR β CDR3 репертуаров растет геометрически при увеличении объема анализируемых массивов данных. На графиках представлено количество общих клонотипов 12 неродственных пар доноров в зависимости от размера анализируемого образца на уровне **(А)** аминокислотных последовательностей (АК), **(Б)** аминокислотных последовательностей с несовпадающими нуклеотидными последовательностями (АК не НТ) и **(В)** нуклеотидных последовательностей (НТ). Каждая из 12 представленных цветных линий представляет собой наблюдаемое количество общих клонотипов внутри пула определенного количества случайно взятых клонотипов каждой пары неродственных доноров. Пунктирной прямой дана степенная зависимость ($Y = aX^b$), которая использовалась для экстраполяции величины перекрытия индивидуальных репертуаров TCR β на случай анализа полных индивидуальных репертуаров. **(Г)** На данном графике представлено распределение количества общих клонотипов для 9 анализируемых людей. Из графика видно, что частота, с которой TCR β клонотипы встречаются в репертуарах человека распределена по степенному закону.

Мы также вычислили верхнюю и нижнюю границу их разнообразия, используя оценку видового разнообразия по *Chao*, описанную в [136]. Это простая оценка использует количество клонотипов, наблюдаемое в одном и двух донорах. Доверительный интервал этой оценки составил от 1.2×10^7 до 5.4×10^7 уникальных аминокислотных последовательностей CDR3, с уровнем достоверности $\alpha = 0.001$.

Эти наблюдения вносят существенный вклад в понимание адаптивной иммунной системы человека. По всей видимости, частотные предпочтения в ходе рекомбинации и селекция в тимусе настолько сильно влияют на формирование репертуара человека, что большинство CDR3 вариантов TCR β , экспрессируемых в наивных Т-лимфоцитах, покидающих тимус, выбираются из некоторого “шорт-листа”, состоящего из менее, чем 10^8 вариантов аминокислотных последовательностей.

Тем не менее, индивидуальный репертуар TCR β имеет сложную структуру, в которой клонотипы с упрощенной последовательностью CDR3 включают большую долю общественных клонотипов (Рисунок 5.7). Интересно, что при анализе пересечений репертуаров всех девяти исследуемых доноров, оказалось, что количество доноров, в крови которых встречается определенный клонотип, распределено в соответствии со степенной функцией (степень -2.95 , $R^2 = 0.99$) (Рисунок 5.8 Г). Это показывает в целом фрактальную структуру репертуара TCR β CDR3 человека, определяющую ландшафт пула общественных клонотипов, которая усложняется при увеличении глубины анализа репертуара.

5.2. Разработка метода нормированного анализа репертуаров TCR с использованием молекулярного баркодирования

Любая из стадий подготовки библиотеки генов антител или Т-клеточных рецепторов (TCR) к анализу (забор образца клеток, выделение нуклеиновых кислот, ПЦР-амплификация, секвенирование) может приводить к искажению информации о реальном репертуаре. Так, например, число лимфоцитов в образцах крови, полученных от разных доноров, или от одного и того же донора, но в разное время (например, до или после проведения той или иной терапии), может существенно варьировать от образца к образцу. Кроме того, независимо от выбранной технологии создания библиотеки (например, создание на основе геномной ДНК [6], [137] или на основе кДНК [102], [104], [138]), невозможно обеспечить равномерное и равноэффективное течение всех этапов процессов пробоподготовки и секвенирования. Описанная проблема, во-первых, существенно осложняет точную оценку исходного количества Т-клеток и реального разнообразия TCR, присутствующего в каждом конкретном образце. А, во-вторых, в значительной мере ограничивает возможную глубину сравнительного анализа разнообразия репертуаров TCR для двух и более образцов.

Для решения указанной проблемы, во второй части нашего исследования мы применили метод молекулярного баркодирования каждой анализируемой молекулы кДНК, объединивший в себе принципы разработанной нами ранее технологии подготовки библиотек генов TCR [102], [107] и технологии введения молекулярных баркодов, описанной в [139]. Этот метод подразумевает использование эффекта смены матрицы [103] для встраивания в синтезируемую в процессе обратной транскрипции кДНК 5'-адаптера, содержащего участок из 12 вырожденных (случайных) нуклеотидов. Таким образом, каждая синтезированная молекула кДНК оказывается специфически помечена одним из 4^{12} (>16.7 миллионов) вариантов уникальных молекулярных баркодов. При дальнейшем анализе данных массивного секвенирования библиотеки, подготовленной описанным способом, такой молекулярный баркод позволяет провести точную оценку стартового количества молекул кДНК в образце, что, как уже было отмечено, невозможно сделать, используя только данные о прочтенных последовательностях генов TCR [140]. Последующая амплификация кДНК библиотеки происходит с использованием универсальных праймеров [102], [104]. Двустороннее

секвенирование по технологии Illumina позволяет получить информацию как о нуклеотидной последовательности CDR3 участка, так и о последовательности уникальной молекулярной метки каждой исходной молекулы кДНК (Рисунок 5.9).

С помощью данной технологии можно нормировать друг относительно друга любые две и более библиотеки, даже если они наработаны с разного количества Т-клеток и отсекуены с разной глубиной покрытия. Для этого необходимо произвольно выбрать из каждого анализируемого массива данных (образца) равное число анализируемых молекул кДНК, характеризующихся индивидуальным, отличным от других молекул баркодом, и уже затем проводить сравнительный анализ. Поскольку каждая прочтенная последовательность с уникальным баркодом соответствует отдельной молекуле кДНК, такая нормализация данных секвенирования позволяет кардинально снизить уровень искажения количественной информации в ходе подготовки библиотеки генов TCR. В результате снижается разброс данных для независимо полученных реплик и повышается детектируемое разнообразие репертуара β цепей TCR в пересчете на фиксированное число прочтенных последовательностей (так как каждое прочтение соответствует новой молекуле кДНК) (Рисунок 5.10). Кроме того, применение данного метода нормализации позволяет снизить количественные искажения представленности вариантов внутри каждого конкретного массива.

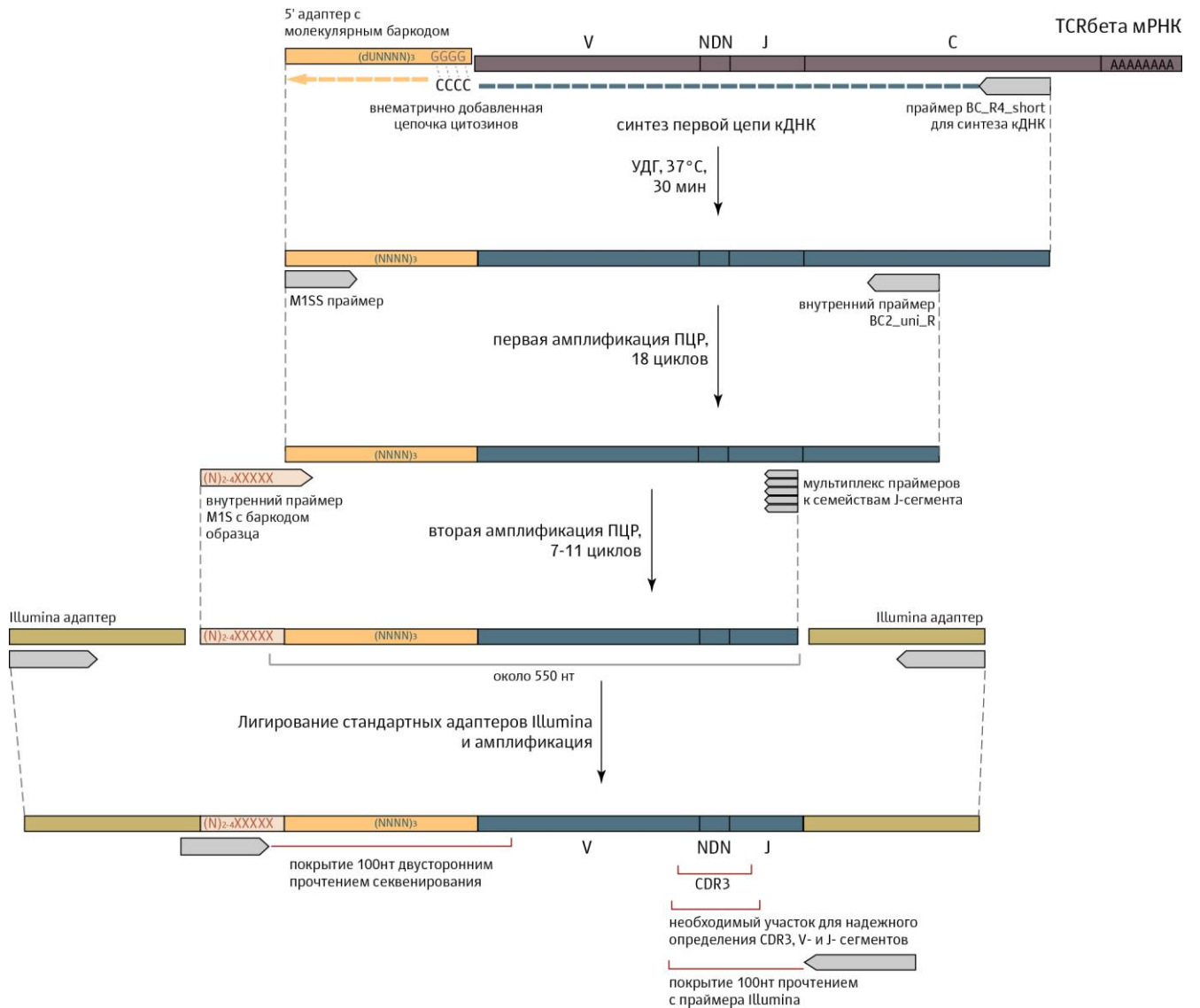


Рисунок 5.9. Схема подготовки библиотеки, созданной на основе мРНК, для секвенирования нового поколения. (dUNNNN)₃ = молекулярный баркод; XXXXX = баркод образца; (N)₂₋₄ – случайные нуклеотиды, добавленные к 5' концу библиотеки с целью создания разнообразия для лучшего распознавания кластеров на секвенаторе Illumina.

Б

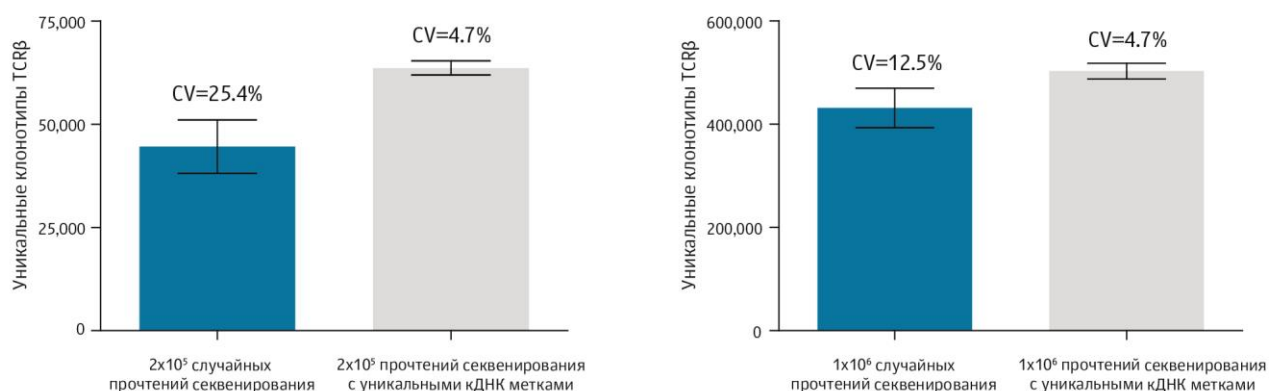


Рисунок 5.10. Нормализация с использованием молекулярного баркодирования увеличивает разнообразие TCRβ внутри образца и уменьшает разброс между образцами. (А) Библиотеки TCRβ, полученные из трех независимо взятых 2 мл образцов крови одного и того же донора. Правый столбец соответствует количеству уникальных TCRβ клонотипов, обнаруженных в 2×10^5 случайно выбранных, CDR3-содержащих прочтений секвенирования, без учета молекулярных баркодов. Правый столбец соответствует аналогичному количеству прочтений секвенирования, характеризующихся уникальными молекулярными баркодами. **(Б)** Библиотеки TCRβ, полученные из образцов с разным количеством РВМС – 5×10^6 и 1×10^7 клеток. Оба образца были получены в результате одного забора крови у донора. Левый столбец соответствует количеству уникальных TCRβ клонотипов, обнаруженных среди 1×10^6 случайно выбранных, CDR3-содержащих прочтений, без учета молекулярных баркодов. Правый столбец соответствует аналогичному количеству прочтений, характеризующихся уникальными молекулярными баркодами. CV – коэффициент вариации.

Ситуации, в которых количества прочтений разных вариантов кДНК, равнопредставленных в исходном образце, отличаются в несколько раз, объясняются стохастичной и неравномерной природой процессов ПЦР-амплификации и секвенирования. Анализ таких образцов приводит к появлению искаженного репертуара, в котором ряд клонов будет представлен избыточно, информация же о других клонах, особенно редких, может быть утеряна.

В модельных экспериментах для библиотек генов TCR нами показано, что средняя эффективность использованной нами технологии смены матрицы ревертазой равна примерно одному молекулярному событию на две Т-клетки. Например, в данных массивированного секвенирования библиотек, приготовленных с использованием в качестве стартового материала примерно 2×10^4 или 1×10^6 Т-клеток, содержится

информация о приблизительно 1×10^4 или 5×10^5 уникальных вариантов баркодов, соответственно. Это означает, что, например, каждая анализируемая молекула кДНК β -цепи TCR, меченая отдельным баркодом, соответствует единичной случайной Т-клетке, а один миллион таких молекулярных событий – одному миллиону Т-клеток, случайно выбранных из более чем 10 миллионов Т-лимфоцитов образца периферической крови объемом 10 мл.

Реализация предложенной нами технологии нормирования данных массивированного секвенирования делает возможным корректное сравнение разнообразия TCR β CDR3 цепей для образцов, полученных от различных пациентов, даже в условиях разного количества клеток на старте, разного качества РНК, и разной глубины секвенирования образцов. Кроме того, применение данного подхода позволяет минимизировать влияние на результаты эксперимента таких случайных и неизбежных факторов, как индивидуальные особенности каждого конкретного образца донорской крови и неравномерность течения процессов на всех стадиях подготовки образца к анализу, начиная от забора крови, и заканчивая секвенированием.

5.3. Исследование изменений репертуара TCR β в ходе старения человека

5.3.1. Индивидуальное разнообразие TCR β CDR3 падает в течение всей жизни человека

В нашем исследовании влияния старения на индивидуальный репертуар TCR β человека принимало участие 39 системно здоровых донора. В зависимости от возраста, они были разделены на 4 группы (Таблица 4):

Группа 1 – молодые доноры (в среднем – 16 лет),

Группа 2 – доноры среднего возраста (в среднем – 39 лет),

Группа 3 – пожилые доноры (в среднем – 62 года),

Группа 4 – долгоживущие доноры (в среднем – 82 года).

На основе РВМС, выделенных из образцов периферической крови каждого из доноров, были приготовлены библиотеки TCR β с применением метода баркодирования молекул кДНК (см. предыдущий раздел). Полученные библиотеки были проанализированы двусторонним секвенированием Illumina. Нормализацию полученных массивов последовательностей проводили с помощью уникальных молекулярных баркодов кДНК. Последующая кластеризация клонотипов и исправление ошибок проводились с помощью программного обеспечения MiTCR, разработанного в нашей лаборатории [109]. Полученные нормализованные массивы последовательностей содержали полную количественную информацию о клональном составе и разнообразии CDR3 репертуара TCR β , соответствующем 10^6 Т-клеток образцов периферической крови каждого донора. Сравнение этого разнообразия у доноров разных возрастов показало, что с увеличением возраста оно падает практически линейно ($R = -0.75$, $P = 4 \times 10^{-8}$), изменяясь от 7.8×10^5 клонотипов у самого молодого донора до 1.3×10^5 у одного из самых пожилых доноров. Примечательно, что достоверное падение разнообразия репертуара TCR β происходит уже к возрасту 40 лет (Рисунок 5.11 А), что не было показано ни в одной из предшествующих работ [77] [89].

Таблица 4. Характеристика возрастных групп.

| Возрастная группа | Средний возраст | Количество доноров (Ж/М - пол) | Наблюдаемое разнообразие на 10^6 Т-клеток, $\times 10^5$ | Нижняя граница общего разнообразия, $\times 10^6$ | Среднее содержание наивных клеток в CD3+ субпопуляции | Среднее содержание наивных клеток в CD4+ субпопуляции | Среднее содержание наивных клеток в CD8+ субпопуляции | Среднее соотношение доли CD4+/CD8+ Т-клеток |
|------------------------------|-----------------|--------------------------------|--|---|---|---|---|---|
| Группа 1 (молодые) | 16 (6-25) | 11 (4/7) | 6.4 ± 1.0 | 5.0 ± 1.3 | 54.3 ± 10.3 | 49.8 ± 9.6 | 59.7 ± 14.0 | 1.38 ± 0.49 |
| Группа 2 (среднего возраста) | 39 (34-43) | 10 (6/4) | 5.0 ± 1.3 | 4.1 ± 1.0 | 35.2 ± 9.4 | 33.7 ± 12.4 | 37.5 ± 8.1 | 1.97 ± 0.56 |
| Группа 3 (пожилые) | 62 (61-66) | 7 (4/3) | 3.1 ± 1.0 | 2.4 ± 6.5 | 17.3 ± 8.1 | 19.9 ± 7.9 | 18.8 ± 12.6 | 2.45 ± 0.95 |
| Группа 4 (долгоживущие) | 82 (71-90) | 11 (5/6) | 3.6 ± 0.9 | 2.6 ± 6.2 | 27.5 ± 14.8 | 34.6 ± 20.9 | 11.2 ± 6.1 | 2.54 ± 1.08 |

Мы наблюдали значительное уменьшение количества уникальных клонотипов TCR β в группе доноров среднего возраста по сравнению с группой молодых доноров ($P < 0.01$, падение в 1.3 раза), а также в группе пожилых доноров относительно группы доноров среднего возраста ($P < 0.01$, падение в 1.6 раз). Однако в группе долгоживущих доноров – по сравнению с группой пожилых доноров – напротив - наблюдалось небольшое *увеличение* разнообразия TCR β , которое, впрочем, не было статистически достоверным (Рисунок 5.11 А).

Примечательно, что анализ 10^6 случайно выбранных CDR3-содержащих прочтений секвенирования *без* учета молекулярных баркодов обнаруживал слабые различия между возрастными группами ($P > 0.05$, двусторонний t-тест для зависимых выборок для Групп 1 и 2). Вероятнее всего, это связано с тем, что Группа 1, например, характеризуется наиболее разнообразным репертуаром, значительная часть которого вытесняется из анализа более представленными клонотипами (Рисунок 5.10, Рисунок 5.12 А).

Анализ абсолютного количества Т-клеток в крови выборочных доноров не выявил возрастных изменений в их численности. Это наблюдение согласуется с исследованиями, проведенными ранее [141]–[143], в которых либо не было найдено никаких возрастных изменений, либо было замечено незначительное возрастное уменьшение абсолютного количества Т-клеток или их доли от общего количества РВМС. Таким образом, наблюдаемые изменения разнообразия репертуара TCR β , по всей видимости, относятся как к относительному разнообразию на 1 миллиону Т-клеток, так и к относительному разнообразию Т-клеток в пересчете на взятый объем периферической крови.

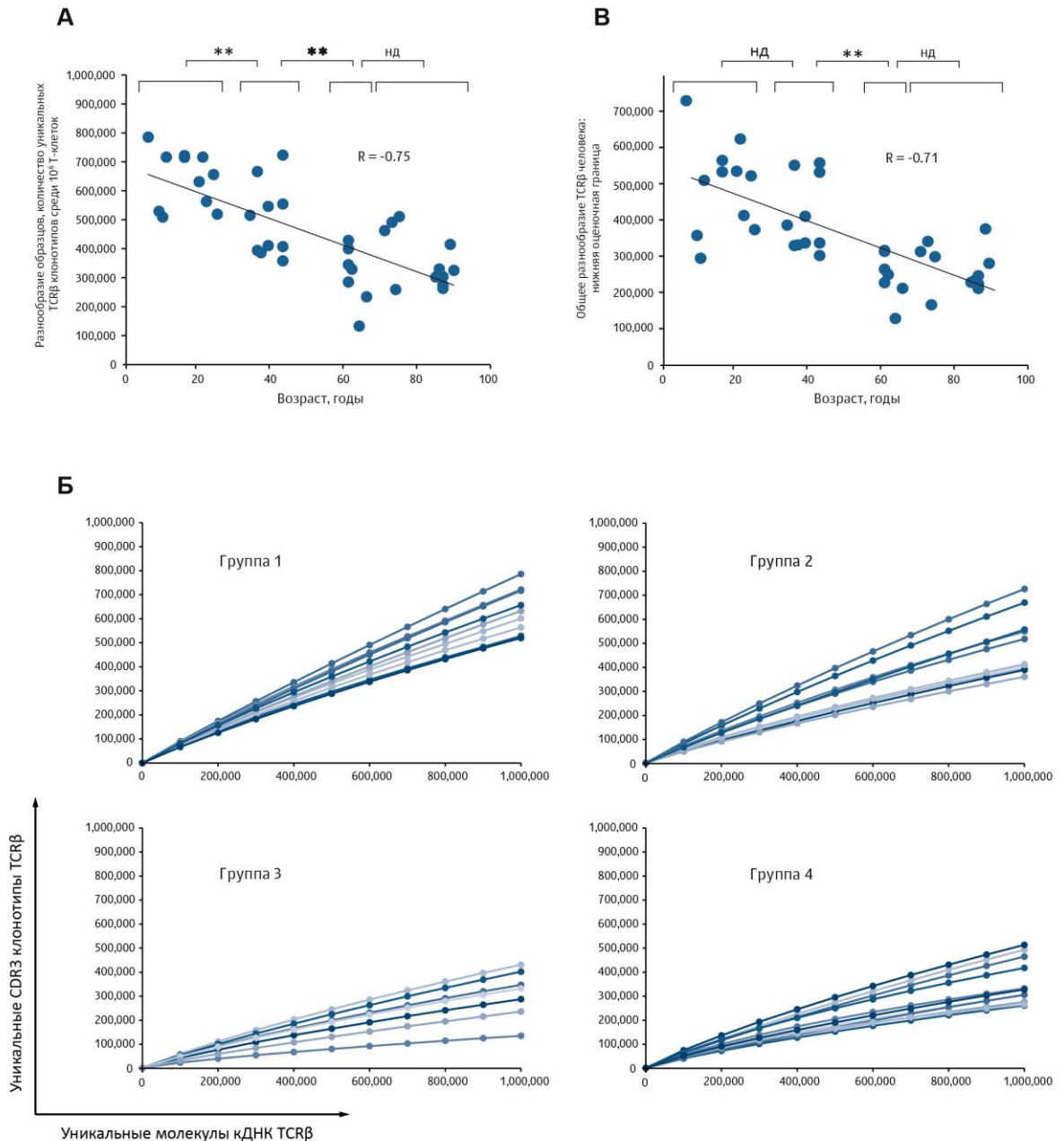


Рисунок 5.11. Возрастные изменения разнообразия репертуаров TCRβ. (А) На графике изображена зависимость количества уникальных TCRβ CDR3 клонотипов на 10^6 Т-клеток, от возраста донора. Наблюдаемое разнообразие падает с увеличением возраста практически линейно и достоверно ($R = -0.75$; $P < 0.0001$, критерий Краскела-Уоллиса). ** – $P < 0.01$, двусторонний t-тест для зависимых выборок; нд – недостоверно. **(Б)** Кривые видового накопления TCRβ клонотипов для четырех возрастных групп, с приращением 1×10^5 . В анализе принимали участие 10^6 молекул кДНК каждого образца. Графики показывают среднее количество уникальных клонотипов, полученных методом случайной выборки ($n=5$) молекул кДНК из образца каждого донора. **(В)** Оценка нижней границы суммарного разнообразия TCRβ доноров.

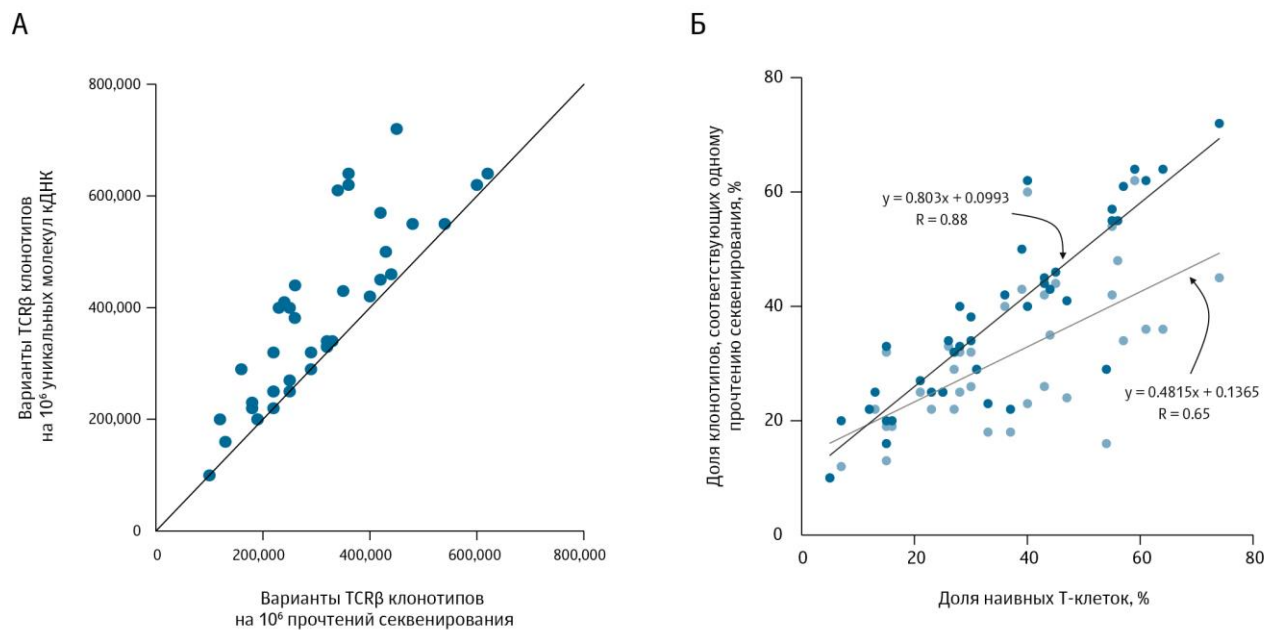


Рисунок 5.12. Сравнение данных, анализируемых по количеству прочтений секвенирования и по количеству молекул кДНК на старте. (А) В данных, анализируемых исходя из числа прочтений секвенирования (10^6 прочтений), разнообразие образца, как правило, занижено – в среднем в 1.15 раз (варьируясь от 1.01 до 1.76 – для разных массивов данных, $P = 2 \times 10^{-7}$, двусторонний t-тест для зависимых выборок) – относительно данных, анализируемых исходя из числа уникальных молекулярных баркодов (10^6 баркодов). **(Б)** Количество редких клонотипов, каждый из которых представлен в массиве данных одной уникальной молекулой кДНК хорошо коррелирует с процентным содержанием наивных Т-клеток в периферической крови (темные круги). Ненормированные массивы данных демонстрируют явную неравномерность и, как следствие, худшую корреляцию с процентным содержанием наивных Т-клеток (светлые круги).

5.3.2. Определение нижней границы индивидуального разнообразия TCRβ CDR3

Анализ кривых видового накопления (Рисунок 5.11 Б) показал, что количество TCRβ клонотипов, обнаруженное в каждом из доноров, возрастает пропорционально количеству молекул кДНК, находящихся в анализе, – практически в линейной зависимости. Кривые видового накопления схожи для доноров в пределах одной группы и различаются для доноров из разных возрастных групп. Большинство этих кривых далеко от насыщения, что затрудняет оценку общего индивидуального разнообразия TCRβ.

Для оценки нижней границы общего разнообразия TCR β в периферической крови изучаемых доноров, нами была применена модель ненаблюдаемых видов (см. Материалы и Методы). Так, нами были получены следующие, последовательно снижающиеся, значения оценочной нижней границы разнообразия для разных возрастных групп (Таблица 4):

Группы 1 – $\sim 7 \times 10^6$ клонотипов;

Группа 2 – $\sim 4 \times 10^6$ клонотипов;

Группа 3 – $\sim 2,4 \times 10^6$ клонотипов (происходит достоверное снижение относительно Группы 2 – в 1.6 раз, $P = 0.001$) (Рисунок 5.11 В).

Стоит отметить, что это лишь нижняя оценочная граница разнообразия, которая будет возрастать при увеличении глубины анализа. Чтобы подтвердить это, мы использовали разбитые на 2 реплики (на уровне PBMC) образцы двух доноров возрастом 25 и 87 лет, в которых наивные Т-клетки (CD45RA^{high}/CD27^{high}) составляли 44% и 15%, соответственно. Каждая реплика состояла из 10^6 уникальных стартовых молекул кДНК. Добавление таких реплик к анализу увеличило оценку разнообразия в 1.91 и в 1.74 раза, а оценку нижней границы разнообразия – в 2.3 ± 0.4 и в 2.4 ± 0.1 раза для молодого и долгоживущего доноров, соответственно. Таким образом, оценка нижней границы общего индивидуального разнообразия TCR β , основанная на 1×10^6 Т-клеток на старте, действительно, ниже реального разнообразия TCR β человека. Скорее всего, с увеличением глубины секвенирования эта оценка будет расти.

В целом, чрезвычайно сложно оценить или измерить общее разнообразие TCR β CDR3 ввиду высокой гетерогенности частот наивных Т-лимфоцитов. В ходе пролиферации на периферии некоторые клоны наивных Т-лимфоцитов делятся более эффективно, в то время как некоторые постепенно исчезают. Соответственно, в определенный момент времени некоторые наивные Т-клоны могут быть представлены в крови всего одной клеткой. Формально такие клоны по-прежнему вносят вклад в общее разнообразие индивидуального репертуара, хотя, вероятно, уже не несут физиологической значимости. В данном исследовании мы достигли рекордной глубины секвенирования репертуара TCR β человека, охватив при этом, тем не менее, лишь до двух миллионов Т-клеток из общего пула Т-лимфоцитов человека, составляющего порядка 10^{12} Т-клеток. Поэтому экстраполировать полученные данные таким образом, чтобы аккуратно и точно предсказать полное разнообразие индивидуального

репертуара TCR β , математически невозможно. В данной работе мы лишь примерно определяем возможную нижнюю границу разнообразия с помощью метода, предложенного [111].

Однако степень увеличения разнообразия TCR β при добавлении миллиона событий сравнима для молодого и старого доноров. Таким образом этот подход в целом применим для относительной оценки общего разнообразия TCR β доноров разных возрастов. Тем не менее, лишь очень глубокое секвенирование сортированных наивных Т-лимфоцитов сможет точно ответить на вопрос об общем индивидуальном разнообразии наивных TCR β клонотипов и его изменении в ходе старения, поскольку в ходе общего анализа разное количество наивных Т-клеток попадает в образцы доноров разных возрастов, что существенно влияет на наблюдаемое разнообразие.

5.3.3. Заполнение гомеостатического пространства активированными клонами Т-клеток

Объем занимаемого гомеостатического пространства размножившимися TCR β клонами имеет четкую возрастную зависимость, прослеживаемую при сравнительном анализе репертуаров TCR β молодой, среднего возраста и пожилой возрастных групп (Группы 1, 2 и 3). Процентное же содержание единично представленных клонотипов (1 Т-лимфоцит = 1 молекула кДНК = один молекулярный баркод), наоборот, падает с возрастом, составляя 55% в среднем в Группе 1 и 24% – в Группе 3. Интересно, что эти закономерности не распространяются на состоящую из долгоживущих доноров Группу 4. В последней, напротив, наблюдается падение количества размножившихся клонов и увеличение количества редко встречающихся клонов, по сравнению с Группой 3 (Рисунок 5.13).

Наблюдаемое в нашем исследовании заполнение гомеостатического пространства активированными Т-лимфоцитами с возрастом соответствует результатам предыдущих исследований [81]. Учитывая относительно стабильное общее количество Т-лимфоцитов, разрастание субпопуляции активированных Т-клеток неизбежно приводит к снижению количества наивных Т-клеток. При этом, общее разнообразие репертуаров формируется в основном редкими клонотипами, а в разнообразии часто встречающихся клонотипов (> 1 копии на миллион) нами не было найдено никаких возрастных изменений ($R = 0.22$, $p = 0.2$).

В совокупности, вышеперечисленные факты могут означать, что с возрастом происходит потеря наивных Т-клонов, при этом активированные клоны к новым антигенам возникают в сравнительно небольших количествах, а размер клонов уже существующих антиген-специфических активированных Т-лимфоцитов постоянно растет.

Тем не менее, без анализа отсортированных субпопуляций Т-лимфоцитов, невозможно окончательно судить о гомеостатических изменениях как внутри субпопуляций, так и между ними. Стоит также отметить, что CMV-статус исследуемых нами доноров нам неизвестен. Поскольку этот вирус оказывает увеличивающее влияние на количество Т-клеток памяти в крови [82], [144], [145], эта информация могла бы быть очень важна. Но так как антитела к CMV есть у абсолютного большинства людей в возрасте более 40 лет, скорее всего, большинство взрослых доноров данного исследования были позитивны к CMV. Для определения чистого эффекта старения на репертуар TCR необходимо проведение дальнейших исследований на отсортированных наивных и активированных Т-лимфоцитах с учетом CMV-статуса выбранных доноров.

5.3.4. Корреляция между разнообразием TCR β и долей наивных Т-клеток в периферической крови

Чтобы оценить корреляцию между количеством наивных Т-клеток в периферической крови, возрастом донора и наблюдаемым разнообразием TCR β , мы измерили процентное содержание наивных Т-клеток (CD45RA^{high}/CD27^{high}) в образцах периферической крови каждого из исследуемых доноров методом проточной цитофлуориметрии (Рисунок 5.14 А; Таблица 4).

Как и предполагалось, была обнаружена обратная зависимость доли наивных Т-клеток в периферической крови донора от его возраста. Коэффициент корреляции R, посчитанный методом линейной регрессии для всех четырех возрастных групп, составил -0.68. Исключение группы долгоживущих доноров (Группа 4) привело к увеличению корреляции (R = -0.85). Это объясняется тем, что доля наивных Т-клеток в этой группе больше, по сравнению с Группой 3, и не отличается достоверно от Группы 2 (P=0.13) (Рисунок 5.14 Б).

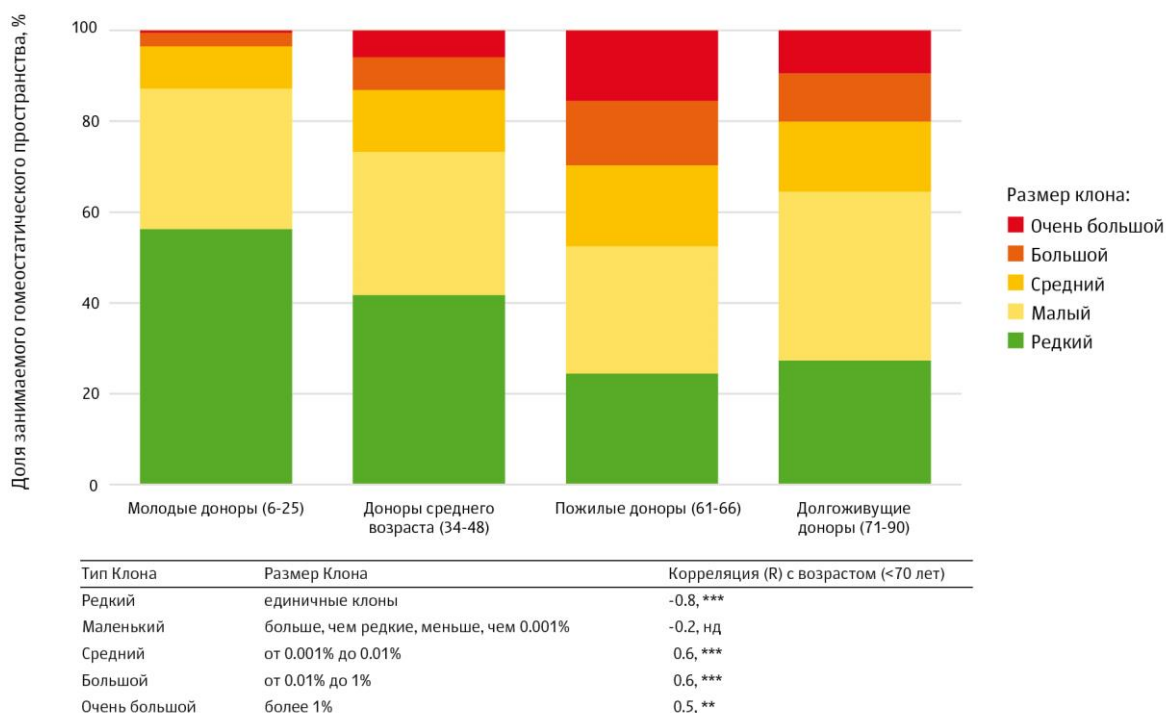


Рисунок 5.13. Распределение гомеостатического пространства доноров разных возрастов между клонами разного размера. На графике представлена доля гомеостатического пространства, занимаемого клонами определенного размера для четырех возрастных групп. Указана корреляция между возрастом донора и занимаемой долей гомеостатического пространства клонами определенного размера. Размер клонов указан в процентах от общего количества анализируемых Т-клеток (10^6). *** – $P < 0.0001$, ** – $P = 0.001$, нд – недостоверно.

Обнаруженная выраженная корреляция между долей наивных Т-клеток и наблюдаемым разнообразием $TCR\beta$ ($R = 0.90$, $p = 1.4 \times 10^{-14}$; Рисунок 5.15) подтверждает как высокую точность методики молекулярного баркодирования для анализа репертуаров $TCR\beta$, так и правильность выделения $CD45RA^{high}/CD27^{high}$ субпопуляции Т-лимфоцитов, как наивных клеток. Как было сказано выше, данные, нормализованные на стартовое количество молекул кДНК, оказались намного более надежными, чем ненормализованные данные, основанные на количестве прочтений секвенирования. Корреляция последних с экспериментальными измерениями доли наивных Т-лимфоцитов была значительно снижена (Рисунок 5.12 Б).

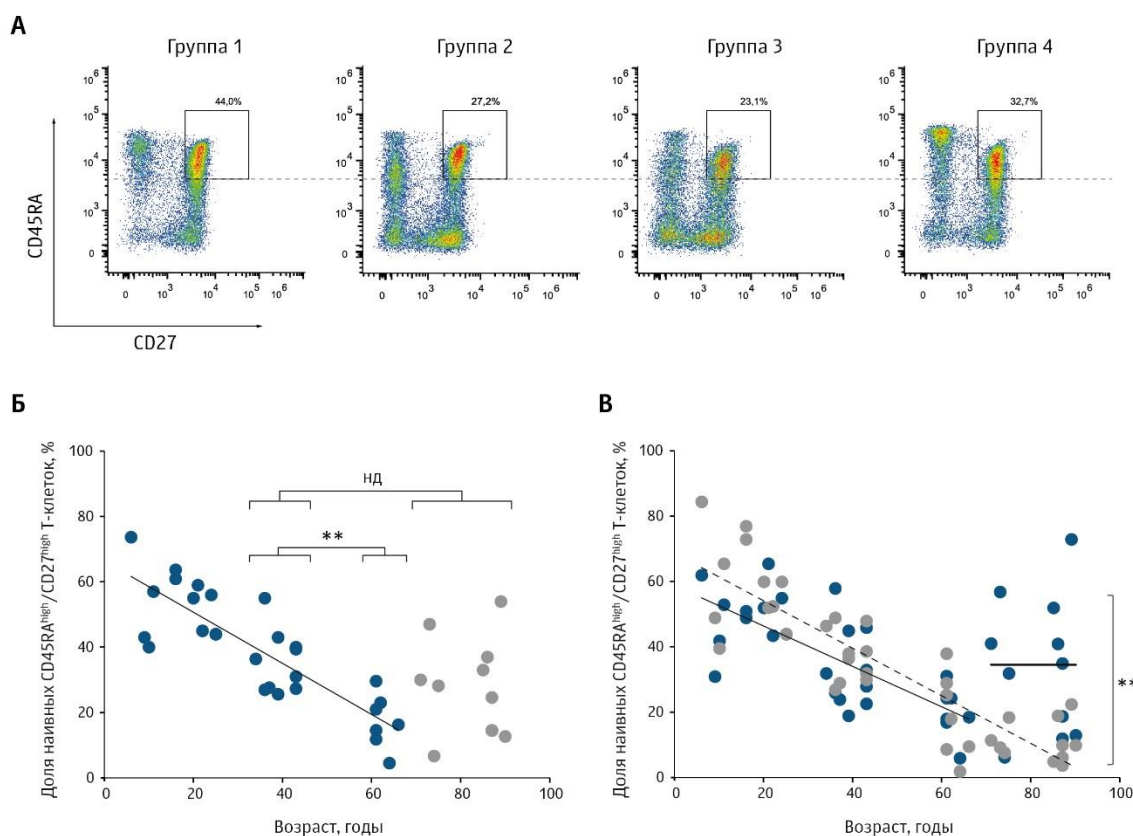


Рисунок 5.14. Возрастное изменение содержания наивных Т-клеток в периферической крови. (А) Способ определения наивных Т-клеток методом проточной цитофлуориметрии (FACS). Периферические мононуклеарные клетки крови были проанализированы методом проточной цитофлуориметрии с использованием меток различных цветов. Т-клетки определяли с помощью антител к CD3. Среди CD3⁺ клеток в дальнейшем выделяли наивные Т-клетки, используя антитела к CD45RA и CD27. На графиках представлены типичные данные по анализу процента наивных Т-клеток из всех CD3⁺ клеток для Групп 1-4. Указанные значения соответствуют процентному содержанию наивных Т-клеток (CD45RA^{high}/CD27^{high}) в пуле CD3⁺ Т-клеток. **(Б, В)** Возрастные изменения численности наивных Т-клеток. **(Б)** Процентное содержание наивных Т-клеток (CD45RA^{high}/CD27^{high}) в пуле CD3⁺ клеток показано в зависимости от возраста донора. Наблюдаемая зависимость линейна и статистически достоверна ($R = -0.85$, $P = 1 \times 10^{-8}$) для Групп 1-3 (синие круги), но не для Группы 4 (серые круги). ** – $P = 0.0012$, нд – недостоверно. **(В)** Процентное содержание наивных Т-клеток в пуле CD8⁺ клеток (серые круги, пунктирная линия) и в пуле CD4⁺ клеток (синие круги, сплошная линия) для Групп 1-3 и для среднего Группы 4). Процентное содержание наивных Т-клеток в CD8⁺ субпопуляции линейно убывает с возрастом ($R = -0.88$). Однако в CD4⁺ субпопуляции такое убывание характерно только для Групп 1-3 ($R = -0.75$). В Группе 4 в CD4⁺ субпопуляции доля наивных клеток статистически достоверно превосходит их долю в CD8⁺ субпопуляции. ** – $P = 0.004$, двусторонний t-тест для зависимых выборок.

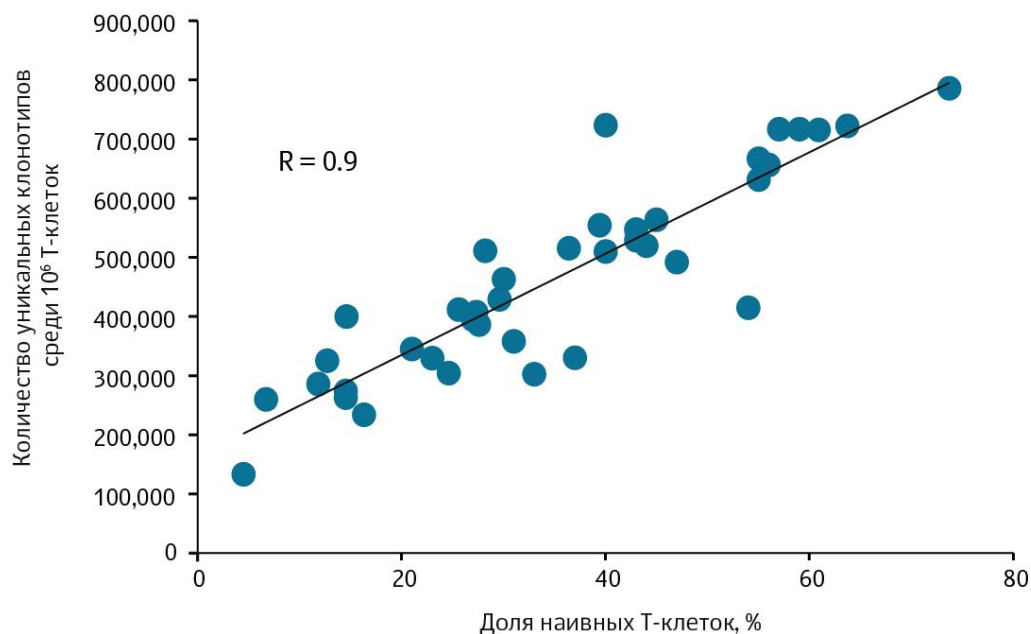


Рисунок 5.15. Наблюдаемое разнообразие TCRβ пропорционально доле наивных Т-клеток в периферической крови. Показана зависимость разнообразия TCRβ, наблюдаемого в 10^6 Т-клеток периферической крови, и процентного содержания наивных Т-клеток ($CD45RA^{high}/CD27^{high}$) в пуле $CD3^+$ клеток ($R = 0.90$. $P = 1.4 \times 10^{-14}$).

Наблюдаемое разнообразие TCRβ на 1×10^6 Т-клеток периферической крови, лучше отражает возрастные изменения, чем доля наивных Т-лимфоцитов в периферической крови. Частная корреляция возраста донора и наблюдаемого разнообразия образца за вычетом эффекта, вызванного возрастным изменением доли наивных Т-клеток, достоверна, $R = -0.47$ ($P = 0.003$). Напротив, частная корреляция возраста донора и доли наивных Т-лимфоцитов за вычетом тренда, связанного с изменением наблюдаемого разнообразия образца, недостоверна, $R = -0.03$ ($P = 0.86$).

5.3.5. Образцы периферической крови долгоживущих доноров характеризуются высоким содержанием $CD4^+$ наивных Т-клеток

При анализе доли наивных Т-лимфоцитов ($CD45RA^{high}/CD27^{high}$) в $CD8^+$ субпопуляции периферической крови доноров оказалось, что она падает линейно с возрастом донора ($R = -0.88$; $P < 1 \times 10^{-8}$). Доля наивных Т-клеток в $CD4^+$ субпопуляции также падает линейно ($R = -0.75$) и со схожей кинетикой, но лишь у доноров моложе 70 лет. В группе же долгоживущих доноров происходит увеличение содержания наивных Т-лимфоцитов в $CD4^+$ субпопуляции в 1.7 раз по сравнению с донорами Группы 3 – различие близкое к

достоверному ($P = 0.054$). Кроме того, в отличие от других рассматриваемых в исследовании возрастных групп, в Группе 4 доля наивных Т-клеток в $CD4^+$ субпопуляции в 3,6 раз превышает их долю в $CD8^+$ субпопуляции (Рисунок 5.14 В). Более высокая доля наивных $CD4^+$ Т-клеток сочетается с большим количеством низкочастотных клонотипов (Рисунок 5.13) и сравнительно высоким разнообразием TCR β (Рисунок 5.11 А) в Группе 4.

Вероятным объяснением этих наблюдений является тот факт, что доноры Группы 4, дожив до глубокой старости, тем самым уже прошли некоторый отбор. То есть их долголетие является следствием (а не причиной) отклонений, наблюдаемых в их периферической крови.

В этом случае можно предположить связь повышенного содержания наивных Т-клеток в $CD4^+$ субпопуляции с увеличенной продолжительностью жизни. Такая зависимость может объясняться тем, что более разнообразная популяция $CD4^+$ Т-лимфоцитов в целом лучше регулирует работу иммунной системы, ограничивая развитие аутоиммунных реакций, частых среди пожилых людей, и обеспечивая более сбалансированный выбор типа и интенсивности иммунного ответа.

Наше объяснение, однако, отличается от такового в недавней работе других авторов [146]. В этой работе предполагается, что остаточное функционирование тимуса пожилых людей может быть сдвинуто в сторону выработки $CD4^+$ наивных Т-лимфоцитов. Однако на Рисунке 5.14 В видно, что до 70 лет содержание $CD4^+$ и $CD8^+$ наивных Т-клеток падает со схожей кинетикой. Таким образом, мы считаем, что именно возрастной отбор, а не общие тенденции иммунной системы, определяют повышенное содержание наивных $CD4^+$ Т-лимфоцитов, наблюдаемое нами в периферической крови долгоживущих доноров.

5.3.6. Возрастные изменения относительной представленности V β - и J β - сегментов

Исходя из полученных нами данных, представленность большинства генов V β - и J β -сегментов в репертуаре TCR практически не меняется с возрастом. Однако в использовании некоторых из них все же наблюдается возрастная корреляция (Рисунок 5.16). Изменение соотношения доли $CD4^+$ и $CD8^+$ Т-клеток, равномерно увеличивающееся с возрастом (наблюдаемое нами и в работе: [147]), может быть частичным объяснением этого явления. Доля семейств V β -7-9 и V β -13, характерных для субпопуляции $CD8^+$ Т-

клеток [148], предположительно должна уменьшаться со снижением относительного количества $CD8^+$ Т-клеток. И это действительно происходит (Рисунок 5.16). Наоборот, представленность семейства $V\beta$ -18, более характерного для субпопуляции $CD4^+$ Т-клеток, возрастает с возрастом доноров (Рисунок 5.16).

При сравнении других параметров между изучаемыми группами – таких как средняя длина CDR3 или количества добавленных нуклеотидов – не было найдено никаких возрастных изменений.

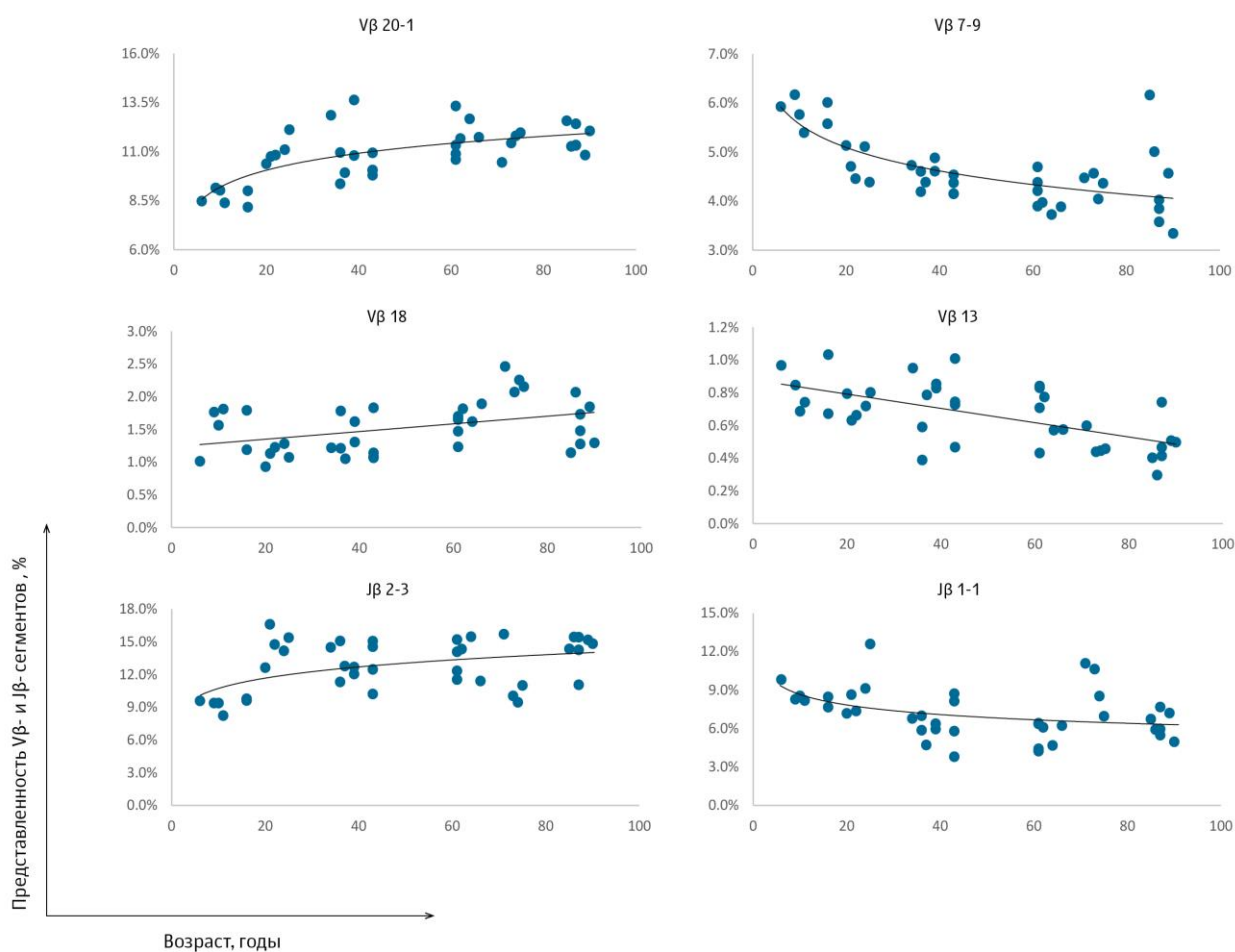


Рисунок 5.16. Обобщенные возрастные изменения представленности $V\beta$ - и $J\beta$ -сегментов TCR. Процентное содержание клонотипов, имеющих в составе TCR определенное семейство $V\beta$ - или $J\beta$ - сегментов, в 10^6 анализируемых Т-лимфоцитов. Для расчета корреляции и аппроксимации данных были использованы степенная ($V\beta$ -20-1, $V\beta$ -7-9, $J\beta$ -2-3 и $J\beta$ -1-1) и линейная ($V\beta$ -13, $V\beta$ -18) функции. Все коэффициенты корреляции статистически достоверны ($P < 0.01$).

5.3.7. Общественные клонотипы и старение

В соответствии с ранее полученными данными [26], [50], [51], [58], [60], [149], [150], сравнительный анализ репертуаров TCR β (нормализованных на количество молекул кДНК на старте) для каждого двух исследуемых нами доноров показал значительное пересечение этих репертуаров, состоящее из 8 000 – 55 000 идентичных *аминокислотных* вариантов CDR3 и 500 – 7 000 идентичных *нуклеотидных* вариантов CDR3.

Чтобы получить список наиболее распространенных общественных клонотипов (repPC, англ. representative public clonotypes) TCR β , был проведен сравнительный анализ репертуаров исследуемых доноров. В анализ вошли по 100 000 наиболее часто встречающихся клонотипов каждого из образцов. Затем были выбраны аминокислотные варианты CDR3, найденные как минимум в 6 из 39 доноров (т.е. в 15% всех доноров). Так, был получен список 10 000 самых часто встречающихся общественных клонотипов (герPC, этот список полностью приведен в нашей статье [151]). Как и ожидалось исходя из предыдущих работ [50], [51], клонотипы, входящие в этот список, характеризуются короткой длиной CDR3 с медианой в 39 нт. Медиана же длин CDR3 всех проанализированных клонотипов из всех доноров составляет 45 нт. Среднее количество добавленных к CDR3 букв также ниже у герPC клонотипов (Рисунок 5.17 А, Б).

В представленности J β -сегментов для герPC не было найдено никаких особенностей. Однако представленность семейств V β -12-4, V β -5-1, V β -7-2 и V β -7-9 была выше, чем в репертуаре TCR β 100 000 самых представленных клонотипов в каждом из 39 доноров, а семейств V β -20-1 и V β -29-1, наоборот, ниже (Рисунок 5.17 В).

Интересно, что количество герPC клонотипов, наблюдаемых в составе 100 000 самых представленных клонотипов, уменьшается от Группы 1 к Группе 3 ($R = 0.58$, $P = 1 \times 10^{-3}$, Рисунок 5.17 Г) и коррелирует с наблюдаемым разнообразием репертуара TCR β (Рисунок 5.17 Д).

Наличие обширных пересечений индивидуальных репертуаров TCR по всей вероятности является важным феноменом, играющим большую роль в эффективности антиген-специфических ответов [60], [62], [149], [150]. RepPC клонотипы, найденные в нашем исследовании, обладают короткой длиной CDR3 и небольшим количеством добавленных к CDR3 нуклеотидов. Интересно, что количество герPC клонотипов в репертуаре периферической крови человека падает с возрастом, повторяя тенденцию

количества наивных Т-клеток и общего разнообразия TCR. Возможно это связано с тем, что значительное количество герПС клонотипов представлено наивными Т-клетками, последовательность CDR3 TCR β которых имеет упрощенную структуру (то есть низкое количество добавленных нуклеотидов), многократно продуцируемую тимусом и вследствие конвергентной рекомбинации составляющую значительную часть репертуара TCR β общественных клонотипов.

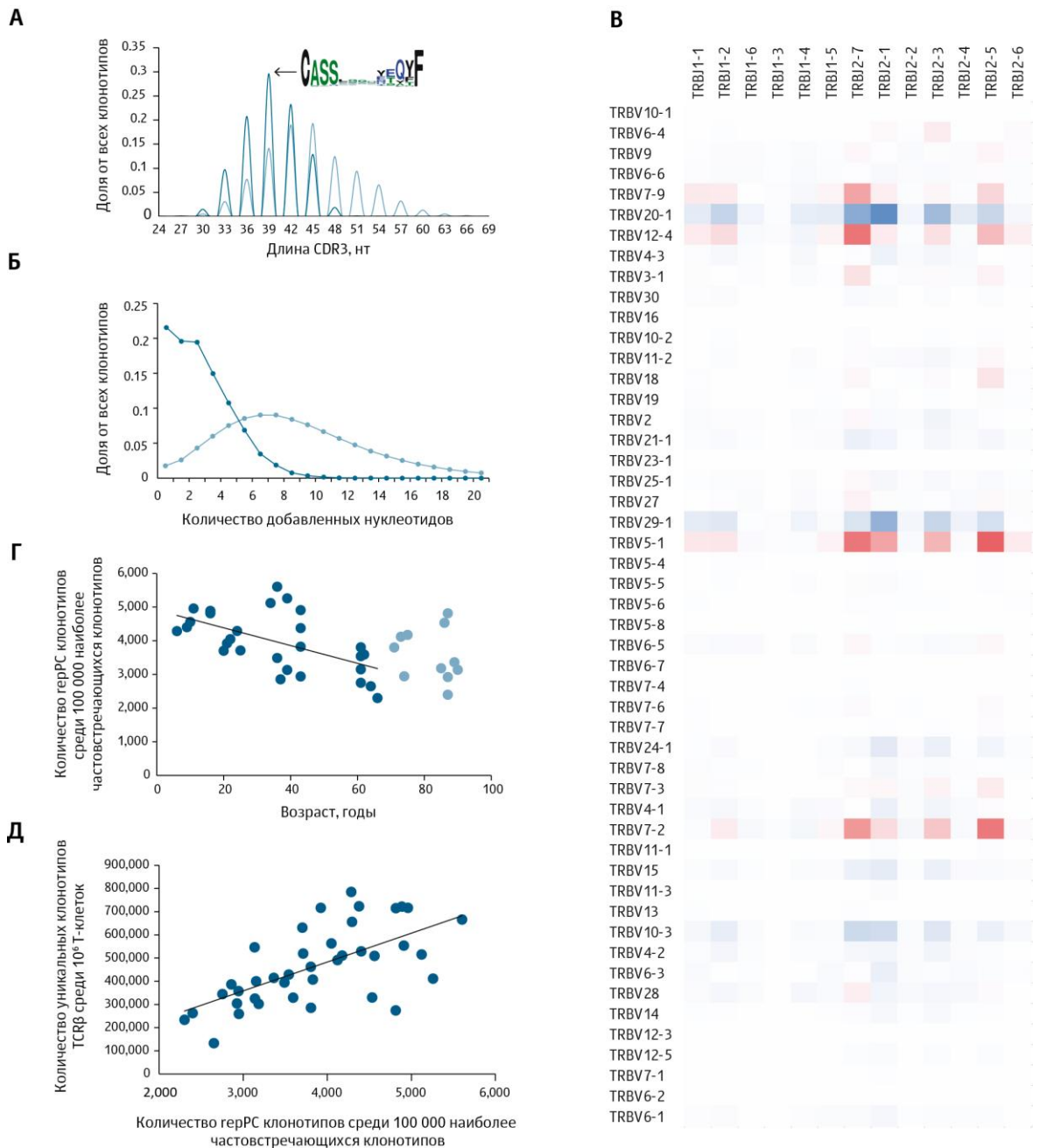


Рисунок 5.17. Характеристика наиболее распространенных общественных клонотипов (герРС). (А, Б) Распределение длины CDR3 участка (А) и добавленных нуклеотидов (Б) для 100 000 наиболее часто встречающихся индивидуальных клонотипов в каждом из 39 доноров (светлая линия) и для 10 692 герРС клонотипов (темная линия). (В) Неравномерность средних частот парной встречаемости семейств V β - и J β - сегментов в герРС клонотипах ([средняя частота парной встречаемости в 100 000 наиболее представленных клонотипов] – [средняя частота парной встречаемости в герРС клонотипах]). (Г) Возрастные изменения количества герРС клонотипов, найденных в составе 100 000 наиболее представленных клонотипов в каждом из исследуемых доноров ($R = 0.58$). (Д) Корреляция количества герРС клонотипов и наблюдаемым разнообразием TCR β . $R = 0.64$, $P < 0.001$; частная корреляция R разнообразия и количества герРС клонотипов, вычисленная при фиксированном значении возраста, составляет 0.48, $P < 0.01$.

6. Выводы

1. Был проведен сравнительный анализ структуры репертуаров TCR β по частотам использования V β -сегментов для родных и неродных пар мама-ребенок. Было показано, что:

- в исходно генерируемом в ходе рекомбинации репертуаре TCR β относительная частота использования V β -сегментов сопоставима для родных и неродных пар;
- давление селекции в тимусе значительно и сонаправленно для всех доноров изменяет частоты использования V β -сегментов при формировании наивного репертуара TCR β ;
- в репертуаре наивных Т-лимфоцитов, после селекции в тимусе, частоты использования V β -сегментов достоверно ближе для родных пар.

2. Впервые был проведен глубокий сравнительный анализ репертуаров последовательностей CDR3 TCR β для родных и неродных пар мама-ребенок на аминокислотном и нуклеотидном уровне. Было показано, что:

- достоверных различий в пересечении репертуаров для родных и неродных пар не наблюдается;
- наибольшая степень пересечения репертуаров TCR β CDR3 для любых двух индивидуумов наблюдается на уровне высокочастотных клонотипов;
- на уровне высокочастотных клонотипов идентичные аминокислотные последовательности CDR3 в родственных парах мама-ребенок чаще сочетаются с идентичными V β -сегментами, и эта величина хорошо коррелирует с числом общих аллелей MHC I;
- селекция в тимусе снижает среднюю длину CDR3 участка репертуара TCR β человека;
- микрохимерных клонов Т-лимфоцитов, перешедших через плаценту в зрелом состоянии, не наблюдается;
- степень пересечения аминокислотных репертуаров TCR β CDR3 для любых двух людей может составлять до 40% от общего индивидуального разнообразия.

3. Был разработан метод на основе молекулярного баркодирования библиотек, который позволяет корректно и глубоко сравнивать разнообразие репертуаров TCR β CDR3 образцов, полученных от различных пациентов, в условиях разного количества клеток на старте, разного качества РНК, и разной глубины секвенирования.

4. С использованием разработанного метода было показано, что относительное разнообразие репертуара CDR3 TCR β периферической крови на 1 миллион Т-лимфоцитов практически линейно убывает с возрастом, начиная от $\sim 6 \times 10^5$ клонотипов в детском возрасте, с кинетикой снижения разнообразия $\sim 5 \times 10^3$ клонотипов в год. Снижение разнообразия хорошо коррелирует с долей наивных Т-лимфоцитов, которая также линейно убывает с возрастом.

5. На основе данных по разнообразию последовательностей CDR3 TCR β на 1 миллион Т-лимфоцитов была проведена оценка нижней границы общего разнообразия TCR β в периферической крови человека в зависимости от возраста. Эта величина составила $\sim 7.0 \times 10^6$ вариантов у детей и $\sim 2.5 \times 10^6$ у пожилых людей.

6. Было показано, что образцы периферической крови доноров в возрасте >70 лет характеризуются:

- повышенным содержанием наивных клеток в популяции CD4+ Т-лимфоцитов;
- повышенным содержанием низкочастотных клонотипов;
- относительно высоким разнообразием CDR3 TCR β .

Эти наблюдения указывают на влияние возрастного отбора и ассоциацию перечисленных параметров с долгожительством.

7. Были идентифицированы последовательности около 10 000 наиболее часто встречающихся в популяции клонотипов TCR β . Было показано, что:

- такие клонотипы характеризуются малой длиной CDR3 и малым количеством добавленных нуклеотидов;
- содержание таких клонотипов в периферической крови снижается с возрастом.

7. Заключение

Массированное секвенирование открывает новые возможности для исследования адаптивного иммунитета на уровне индивидуальных репертуаров антител и Т-клеточных рецепторов.

В настоящей работе мы применили этот подход для анализа индивидуальных репертуаров CDR3 TCR β различных когорт системно здоровых доноров. Достигнутая глубина анализа (порядка 1 миллиона индивидуальных клонотипов для каждого донора), а также применение разработанной нами технологии молекулярного баркодирования для количественной нормализации сравниваемых библиотек позволило нам на качественно новом уровне охарактеризовать индивидуальные репертуары TCR β . Мы изучили процесс формирования индивидуальных репертуаров TCR β и влияние наследственных факторов на этот процесс, исследовали структуру индивидуальных репертуаров и степень их пересечения между различными индивидуумами [152], [153], а также оценили изменения в разнообразии и общей структуре репертуара TCR β в ходе старения [151].

В данной работе мы не проводили физического разделения исследуемых образцов на функциональные субпопуляции Т-лимфоцитов, такие как CD4 $^+$ и CD8 $^+$; наивные, эффекторные и Т-клетки памяти; различные функциональные субпопуляции CD4 $^+$ лимфоцитов. Дальнейшие глубокие исследования репертуаров TCR для различных сортированных функциональных субпопуляций Т-лимфоцитов представляют безусловный интерес.

Мы надеемся, что разработанные нами технологии и программное обеспечение [7], [107], [109], [135] окажутся востребованными для нормированного и безошибочного сравнительного анализа репертуаров Т-клеточных рецепторов и антител как в фундаментальных, так и в прикладных исследованиях адаптивного иммунитета.

8. Список сокращений

CDR – регион, определяющий комплементарность (англ., complementarity determining region)

CMV – цитомегаловирус (англ., cytomegalovirus)

CV – коэффициент вариации (англ., coefficient variance)

DN – двойные негативные тимоциты (англ., double negative)

dNTP – дезоксирибонуклеозидтрифосфат (англ., deoxinucleotide triphosphate)

DP – двойные позитивные тимоциты (англ., double positive)

FACS – метод проточной цитофлуориметрии (англ., fluorescence activated cell sorting)

FITC – флуоресцеин изотиоцианат (англ., fluorescein isothiocyanate)

HLA – человеческий лейкоцитарный антиген (англ., human leucocyte antigen)

MHC – главный комплекс гистосовместимости (англ., major histocompatibility complex)

NGS – массивированное секвенирование или секвенирование нового поколения (англ., next generation sequencing)

NMD – нонсенс-опосредованное разрушение мРНК (англ., nonsense-mediated decay)

PBMC – периферические мононуклеарные клетки крови (англ., peripheral blood mononuclear cells)

PC7 – фикоэритрин цианин 7 (англ., phycoerythrin cyanin 7)

PC5 – фикоэритрин цианин 5.1 (англ., phycoerythrin cyanin 5.1)

PE – фикоэритрин (англ., phycoerythrin)

RAG – гены, активирующие рекомбинацию (англ., recombination activation genes)

repPC – наиболее распространенные общественные клонотипы (англ., representative public clonotypes)

RSS – сигнальные последовательности рекомбинации (англ., recombination signal sequence)

ТСЕР - трис(2-карбоксиэтил)фосфин (англ., tris(2-carboxyethyl)phosphine)

ТСR – Т-клеточный рецептор (англ., T-cell receptor)

Th – Т-хелперы (англ., T-helper cells)

ДТТ - дитиотрейтол

ПЦР – полимеразная цепная реакция

ЭДТА – этилендиаминтетрауксусная кислота

УДГ – урацил-ДНК гликозилаза

9. Список литературы

- [1] D. J. Woodsworth, M. Castellarin, and R. A. Holt, "Sequence analysis of T-cell repertoires in health and disease.," *Genome Med.*, vol. 5, no. 10, p. 98, Jan. 2013.
- [2] H. Robins, "Immunosequencing: applications of immune repertoire deep sequencing.," *Curr. Opin. Immunol.*, vol. 25, no. 5, pp. 646–52, Oct. 2013.
- [3] E. Hammarlund, M. W. Lewis, S. G. Hansen, L. I. Strelow, J. A. Nelson, G. J. Sexton, J. M. Hanifin, and M. K. Slifka, "Duration of antiviral immunity after smallpox vaccination.," *Nat. Med.*, vol. 9, no. 9, pp. 1131–7, Sep. 2003.
- [4] O. V Britanova, a G. Bochkova, D. B. Staroverov, D. a Fedorenko, D. a Bolotin, I. Z. Mamedov, M. a Turchaninova, E. V Putintseva, a a Kotlobay, S. Lukyanov, a a Novik, Y. B. Lebedev, and D. M. Chudakov, "First autologous hematopoietic SCT for ankylosing spondylitis: a case report and clues to understanding the therapy.," *Bone Marrow Transplant.*, vol. 47, no. 11, pp. 1479–81, Nov. 2012.
- [5] V. Venturi, B. D. Rudd, and M. P. Davenport, "Specificity, promiscuity, and precursor frequency in immunoreceptors.," *Curr. Opin. Immunol.*, vol. 25, no. 5, pp. 639–45, Oct. 2013.
- [6] H. S. Robins, P. V Campregher, S. K. Srivastava, A. Wacher, C. J. Turtle, O. Kahsai, S. R. Riddell, E. H. Warren, and C. S. Carlson, "Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells.," *Blood*, vol. 114, no. 19, pp. 4099–107, Nov. 2009.
- [7] M. Shugay, O. V Britanova, E. M. Merzlyak, M. A. Turchaninova, I. Z. Mamedov, T. R. Tuganbaev, D. A. Bolotin, D. B. Staroverov, E. V Putintseva, K. Plevova, C. Linnemann, D. Shagin, S. Pospisilova, S. Lukyanov, T. N. Schumacher, and D. M. Chudakov, "Towards error-free profiling of immune repertoires.," *Nat. Methods*, vol. 11, no. 6, pp. 653–5, May 2014.

- [8] E. W. Newell, N. Sigal, S. C. Bendall, G. P. Nolan, and M. M. Davis, "Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes.," *Immunity*, vol. 36, no. 1, pp. 142–52, Jan. 2012.
- [9] G. Vahedi, A. C Poholek, T. W. Hand, A. Laurence, Y. Kanno, J. J. O'Shea, and K. Hirahara, "Helper T-cell identity and evolution of differential transcriptomes and epigenomes.," *Immunol. Rev.*, vol. 252, no. 1, pp. 24–40, Mar. 2013.
- [10] A. K. Shalek, R. Satija, X. Adiconis, R. S. Gertner, J. T. Gaublomme, R. Raychowdhury, S. Schwartz, N. Yosef, C. Malboeuf, D. Lu, J. J. Trombetta, D. Gennert, A. Gnirke, A. Goren, N. Hacohen, J. Z. Levin, H. Park, and A. Regev, "Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells.," *Nature*, vol. 498, no. 7453, pp. 236–40, Jun. 2013.
- [11] A. Murugan, T. Mora, A. M. Walczak, and C. G. Callan, "Statistical inference of the generation probability of T-cell receptors from sequence repertoires.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 40, pp. 16161–6, Oct. 2012.
- [12] Y. Kim, A. Sette, and B. Peters, "Applications for T-cell epitope queries and tools in the Immune Epitope Database and Analysis Resource.," *J. Immunol. Methods*, vol. 374, no. 1–2, pp. 62–9, Nov. 2011.
- [13] X.-L. Li, M.-K. Teng, E. L. Reinherz, and J.-H. Wang, "Strict Major Histocompatibility Complex Molecule Class-Specific Binding by Co-Receptors Enforces MHC-Restricted $\alpha\beta$ TCR Recognition during T Lineage Subset Commitment.," *Front. Immunol.*, vol. 4, p. 383, Jan. 2013.
- [14] K. Kasuga, "Comprehensive analysis of MHC ligands in clinical material by immunoaffinity-mass spectrometry.," *Methods Mol. Biol.*, vol. 1023, pp. 203–18, Jan. 2013.
- [15] E. W. Newell and M. M. Davis, "Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells.," *Nat. Biotechnol.*, vol. 32, no. 2, pp. 149–57, Feb. 2014.

- [16] B. A. Kidd, L. A. Peters, E. E. Schadt, and J. T. Dudley, "Unifying immunology with informatics and multiscale biology.," *Nat. Immunol.*, vol. 15, no. 2, pp. 118–27, Feb. 2014.
- [17] K. Murphy, *Janeway's Immunobiology, Eighth Edition*. 2011.
- [18] A. J. Yates, "Theories and Quantification of Thymic Selection.," *Front. Immunol.*, vol. 5, p. 13, Jan. 2014.
- [19] D. B. Roth, J. P. Menetski, P. B. Nakajima, M. J. Bosma, and M. Gellert, "V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes.," *Cell*, vol. 70, no. 6, pp. 983–91, Sep. 1992.
- [20] K. J. L. Jackson, M. J. Kidd, Y. Wang, and A. M. Collins, "The Shape of the Lymphocyte Receptor Repertoire: Lessons from the B Cell Receptor.," *Front. Immunol.*, vol. 4, p. 263, Jan. 2013.
- [21] M. Hofle, D. S. Linthicum, and T. Ioerger, "Analysis of diversity of nucleotide and amino acid distributions in the VD and DJ joining regions in Ig heavy chains.," *Mol. Immunol.*, vol. 37, no. 14, pp. 827–35, Oct. 2000.
- [22] B. Nadel and A. J. Feeney, "Influence of coding-end sequence on coding-end processing in V(D)J recombination.," *J. Immunol.*, vol. 155, no. 9, pp. 4322–9, Nov. 1995.
- [23] J. L. Abbey and H. C. O'Neill, "Expression of T-cell receptor genes during early T-cell development.," *Immunol. Cell Biol.*, vol. 86, no. 2, pp. 166–74, Feb. 2008.
- [24] K. E. Baker and R. Parker, "Nonsense-mediated mRNA decay: terminating erroneous gene expression.," *Curr. Opin. Cell Biol.*, vol. 16, no. 3, pp. 293–9, Jun. 2004.
- [25] I. Behm-Ansmant and E. Izaurralde, "Quality control of gene expression: a stepwise assembly pathway for the surveillance complex that triggers nonsense-mediated mRNA decay.," *Genes Dev.*, vol. 20, no. 4, pp. 391–8, Feb. 2006.

- [26] J. J. Miles, D. C. Douek, and D. a Price, "Bias in the $\alpha\beta$ T-cell repertoire: implications for disease pathogenesis and vaccination.," *Immunol. Cell Biol.*, vol. 89, no. 3, pp. 375–87, Mar. 2011.
- [27] C. T. Weaver, L. E. Harrington, P. R. Mangan, M. Gavrieli, and K. M. Murphy, "Th17: an effector CD4 T cell lineage with regulatory T cell ties.," *Immunity*, vol. 24, no. 6, pp. 677–88, Jun. 2006.
- [28] S. A. Valkenburg, V. Venturi, T. H. Y. Dang, N. L. Bird, P. C. Doherty, S. J. Turner, M. P. Davenport, and K. Kedzierska, "Early priming minimizes the age-related immune compromise of CD8⁺ T cell diversity and function.," *PLoS Pathog.*, vol. 8, no. 2, p. e1002544, Feb. 2012.
- [29] M. G. Rudolph, R. L. Stanfield, and I. A. Wilson, "How TCRs bind MHCs, peptides, and coreceptors.," *Annu. Rev. Immunol.*, vol. 24, pp. 419–66, Jan. 2006.
- [30] M. M. Davis and P. J. Bjorkman, "T-cell antigen receptor genes and T-cell recognition.," *Nature*, vol. 334, no. 6181, pp. 395–402, Aug. 1988.
- [31] K. Shortman, M. Egerton, G. J. Spangrude, and R. Scollay, "The generation and fate of thymocytes.," *Semin. Immunol.*, vol. 2, no. 1, pp. 3–12, Jan. 1990.
- [32] P. Ferrier, *V(D)J Recombination*. Springer, 2009, p. 211.
- [33] E. Bianconi, A. Piovesan, F. Facchin, A. Beraudi, R. Casadei, F. Frabetti, L. Vitale, M. C. Pelleri, S. Tassani, F. Piva, S. Perez-Amodio, P. Strippoli, and S. Canaider, "An estimation of the number of cells in the human body.," *Ann. Hum. Biol.*, vol. 40, no. 6, pp. 463–71.
- [34] A. Casrouge, E. Beaudoin, S. Dalle, C. Pannetier, J. Kanellopoulos, and P. Kourilsky, "Size estimate of the alpha beta TCR repertoire of naive mouse splenocytes.," *J. Immunol.*, vol. 164, no. 11, pp. 5782–7, Jun. 2000.
- [35] T. P. Arstila, A. Casrouge, V. Baron, J. Even, J. Kanellopoulos, and P. Kourilsky, "A direct estimate of the human alphabeta T cell receptor diversity.," *Science*, vol. 286, no. 5441, pp. 958–61, Oct. 1999.

- [36] P. A. Moss, R. J. Moots, W. M. Rosenberg, S. J. Rowland-Jones, H. C. Bodmer, A. J. McMichael, and J. I. Bell, "Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 88, no. 20, pp. 8987–90, Oct. 1991.
- [37] V. P. Argat, C. W. Schmidt, S. R. Burrows, S. L. Silins, M. G. Kurilla, D. L. Doolan, A. Suhrbier, D. J. Moss, E. Kieff, T. B. Sculley, and I. S. Misko, "Dominant selection of an invariant T cell antigen receptor in response to persistent infection by Epstein-Barr virus.," *J. Exp. Med.*, vol. 180, no. 6, pp. 2335–40, Dec. 1994.
- [38] R. Cibotti, J. P. Cabaniols, C. Pannetier, C. Delarbre, I. Vergnon, J. M. Kanellopoulos, and P. Kourilsky, "Public and private V beta T cell receptor repertoires against hen egg white lysozyme (HEL) in nontransgenic versus HEL transgenic mice.," *J. Exp. Med.*, vol. 180, no. 3, pp. 861–72, Sep. 1994.
- [39] N. E. Annels, M. F. Callan, L. Tan, and A. B. Rickinson, "Changing patterns of dominant TCR usage with maturation of an EBV-specific cytotoxic T cell response.," *J. Immunol.*, vol. 165, no. 9, pp. 4831–41, Nov. 2000.
- [40] A. Lim, L. Trautmann, M. A. Peyrat, C. Couedel, F. Davodeau, F. Romagné, P. Kourilsky, and M. Bonneville, "Frequent contribution of T cell clonotypes with public TCR features to the chronic response against a dominant EBV-derived epitope: application to direct detection of their molecular imprint on the human peripheral T cell repertoire.," *J. Immunol.*, vol. 165, no. 4, pp. 2001–11, Aug. 2000.
- [41] N. Khan, M. Cobbold, R. Keenan, and P. A. H. Moss, "Comparative analysis of CD8+ T cell responses against human cytomegalovirus proteins pp65 and immediate early 1 shows similarities in precursor frequency, oligoclonality, and phenotype.," *J. Infect. Dis.*, vol. 185, no. 8, pp. 1025–34, Apr. 2002.
- [42] N. Khan, N. Shariff, M. Cobbold, R. Bruton, J. A. Ainsworth, A. J. Sinclair, L. Nayak, and P. A. H. Moss, "Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals.," *J. Immunol.*, vol. 169, no. 4, pp. 1984–92, Aug. 2002.

- [43] L. Trautmann, M. Rimbert, K. Echasserieu, X. Saulquin, B. Neveu, J. Dechanet, V. Cerundolo, and M. Bonneville, "Selection of T cell clones expressing high-affinity public TCRs within Human cytomegalovirus-specific CD8 T cell responses.," *J. Immunol.*, vol. 175, no. 9, pp. 6123–32, Nov. 2005.
- [44] L. Trautmann, N. Labarrière, F. Jotereau, V. Karanikas, N. Gervois, T. Connerotte, P. Coulie, and M. Bonneville, "Dominant TCR V alpha usage by virus and tumor-reactive T cells with wide affinity ranges for their specific antigens.," *Eur. J. Immunol.*, vol. 32, no. 11, pp. 3181–90, Nov. 2002.
- [45] D. A. Price, "Avidity for antigen shapes clonal dominance in CD8+ T cell populations specific for persistent DNA viruses," *J. Exp. Med.*, vol. 202, no. 10, pp. 1349–1361, Nov. 2005.
- [46] P. Boudinot, S. Boubekour, and A. Benmansour, "Rhabdovirus infection induces public and private T cell responses in teleost fish.," *J. Immunol.*, vol. 167, no. 11, pp. 6202–9, Dec. 2001.
- [47] D. A. Price, S. M. West, M. R. Betts, L. E. Ruff, J. M. Brenchley, D. R. Ambrozak, Y. Edghill-Smith, M. J. Kuroda, D. Bogdan, K. Kunstman, N. L. Letvin, G. Franchini, S. M. Wolinsky, R. A. Koup, and D. C. Douek, "T cell receptor recognition motifs govern immune escape patterns in acute SIV infection.," *Immunity*, vol. 21, no. 6, pp. 793–803, Dec. 2004.
- [48] P. Bousso, A. Casrouge, J. D. Altman, M. Haury, J. Kanellopoulos, J. P. Abastado, and P. Kourilsky, "Individual variations in the murine T cell response to a specific peptide reflect variability in naive repertoires.," *Immunity*, vol. 9, no. 2, pp. 169–78, Aug. 1998.
- [49] M. F. Quigley, H. Y. Greenaway, V. Venturi, R. Lindsay, K. M. Quinn, R. A. Seder, D. C. Douek, M. P. Davenport, and D. A. Price, "Convergent recombination shapes the clonotypic landscape of the naive T-cell repertoire.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 107, no. 45, pp. 19414–9, Nov. 2010.
- [50] H. S. Robins, S. K. Srivastava, P. V Campregher, C. J. Turtle, J. Andriesen, S. R. Riddell, C. S. Carlson, and E. H. Warren, "Overlap and effective size of the human CD8+ T cell receptor repertoire.," *Sci. Transl. Med.*, vol. 2, no. 47, p. 47ra64, Sep. 2010.

- [51] V. Venturi, M. F. Quigley, H. Y. Greenaway, P. C. Ng, Z. S. Ende, T. McIntosh, T. E. Asher, J. R. Almeida, S. Levy, D. A. Price, M. P. Davenport, and D. C. Douek, "A mechanism for TCR sharing between T cell subsets and individuals revealed by pyrosequencing.," *J. Immunol.*, vol. 186, no. 7, pp. 4285–94, Apr. 2011.
- [52] S. J. Turner, P. C. Doherty, J. McCluskey, and J. Rossjohn, "Structural determinants of T-cell receptor bias in immunity.," *Nat. Rev. Immunol.*, vol. 6, no. 12, pp. 883–94, Dec. 2006.
- [53] G. B. E. Stewart-Jones, A. J. McMichael, J. I. Bell, D. I. Stuart, and E. Y. Jones, "A structural basis for immunodominant human T cell receptor recognition.," *Nat. Immunol.*, vol. 4, no. 7, pp. 657–63, Jul. 2003.
- [54] S. J. Turner, K. Kedzierska, H. Komodromou, N. L. La Gruta, M. A. Dunstone, A. I. Webb, R. Webby, H. Walden, W. Xie, J. McCluskey, A. W. Purcell, J. Rossjohn, and P. C. Doherty, "Lack of prominent peptide-major histocompatibility complex features limits repertoire diversity in virus-specific CD8+ T cell populations.," *Nat. Immunol.*, vol. 6, no. 4, pp. 382–9, Apr. 2005.
- [55] L. Kjer-Nielsen, C. S. Clements, A. W. Purcell, A. G. Brooks, J. C. Whisstock, S. R. Burrows, J. McCluskey, and J. Rossjohn, "A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity.," *Immunity*, vol. 18, no. 1, pp. 53–64, Jan. 2003.
- [56] G. H. Gauss and M. R. Lieber, "Mechanistic constraints on diversity in human V(D)J recombination.," *Mol. Cell. Biol.*, vol. 16, no. 1, pp. 258–69, Jan. 1996.
- [57] V. Venturi, K. Kedzierska, D. A. Price, P. C. Doherty, D. C. Douek, S. J. Turner, and M. P. Davenport, "Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 103, no. 49, pp. 18691–6, Dec. 2006.
- [58] V. Venturi, D. A. Price, D. C. Douek, and M. P. Davenport, "The molecular basis for public T-cell responses?," *Nat. Rev. Immunol.*, vol. 8, no. 3, pp. 231–8, Mar. 2008.

- [59] V. Venturi, H. Y. Chin, D. A. Price, D. C. Douek, and M. P. Davenport, "The role of production frequency in the sharing of simian immunodeficiency virus-specific CD8+ TCRs between macaques.," *J. Immunol.*, vol. 181, no. 4, pp. 2597–609, Aug. 2008.
- [60] H. Li, C. Ye, G. Ji, and J. Han, "Determinants of public T cell responses.," *Cell Res.*, vol. 22, no. 1, pp. 33–42, Jan. 2012.
- [61] V. Venturi, H. Y. Chin, T. E. Asher, K. Ladell, P. Scheinberg, E. Bornstein, D. van Bockel, A. D. Kelleher, D. C. Douek, D. A. Price, and M. P. Davenport, "TCR beta-chain sharing in human CD8+ T cell responses to cytomegalovirus and EBV.," *J. Immunol.*, vol. 181, no. 11, pp. 7853–62, Dec. 2008.
- [62] H. Li, C. Ye, G. Ji, X. Wu, Z. Xiang, Y. Li, Y. Cao, X. Liu, D. C. Douek, D. A. Price, and J. Han, "Recombinatorial biases and convergent recombination determine interindividual TCR β sharing in murine thymocytes.," *J. Immunol.*, vol. 189, no. 5, pp. 2404–13, Sep. 2012.
- [63] A. L. Furmanski, C. Ferreira, I. Bartok, S. Dimakou, J. Rice, F. K. Stevenson, M. M. Millrain, E. Simpson, and J. Dyson, "Public T cell receptor beta-chains are not advantaged during positive selection.," *J. Immunol.*, vol. 180, no. 2, pp. 1029–39, Jan. 2008.
- [64] D. Aw, A. B. Silva, and D. B. Palmer, "Immunosenescence: emerging challenges for an ageing population.," *Immunology*, vol. 120, no. 4, pp. 435–46, Apr. 2007.
- [65] A. C. Shaw, S. Joshi, H. Greenwood, A. Panda, and J. M. Lord, "Aging of the innate immune system.," *Curr. Opin. Immunol.*, vol. 22, no. 4, pp. 507–13, Aug. 2010.
- [66] A. J. George and M. A. Ritter, "Thymic involution with ageing: obsolescence or good housekeeping?," *Immunol. Today*, vol. 17, no. 6, pp. 267–72, Jun. 1996.
- [67] D. D. Taub and D. L. Longo, "Insights into thymic aging and regeneration.," *Immunol. Rev.*, vol. 205, pp. 72–93, Jun. 2005.
- [68] H. E. Lynch, G. L. Goldberg, A. Chidgey, M. R. M. Van den Brink, R. Boyd, and G. D. Sempowski, "Thymic involution and immune reconstitution.," *Trends Immunol.*, vol. 30, no. 7, pp. 366–73, Jul. 2009.

- [69] D. P. Shanley, D. Aw, N. R. Manley, and D. B. Palmer, "An evolutionary perspective on the mechanisms of immunosenescence.," *Trends Immunol.*, vol. 30, no. 7, pp. 374–81, Jul. 2009.
- [70] D. Aw and D. B. Palmer, "It's not all equal: a multiphasic theory of thymic involution.," *Biogerontology*, vol. 13, no. 1, pp. 77–81, Feb. 2012.
- [71] E. Montecino-Rodriguez, H. Min, and K. Dorshkind, "Reevaluating current models of thymic involution.," *Semin. Immunol.*, vol. 17, no. 5, pp. 356–61, Oct. 2005.
- [72] I. den Braber, T. Mugwagwa, N. Vrisekoop, L. Westera, R. Mögling, A. B. de Boer, N. Willems, E. H. R. Schrijver, G. Spierenburg, K. Gaiser, E. Mul, S. a Otto, A. F. C. Ruiter, M. T. Ackermans, F. Miedema, J. a M. Borghans, R. J. de Boer, and K. Tesselaar, "Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans.," *Immunity*, vol. 36, no. 2, pp. 288–97, Feb. 2012.
- [73] L. Cicin-Sain, I. Messaoudi, B. Park, N. Currier, S. Planer, M. Fischer, S. Tackitt, D. Nikolich-Zugich, A. Legasse, M. K. Axthelm, L. J. Picker, M. Mori, and J. Nikolich-Zugich, "Dramatic increase in naive T cell turnover is linked to loss of naive T cells from old primates.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 50, pp. 19960–5, Dec. 2007.
- [74] C. Bourgeois, Z. Hao, K. Rajewsky, A. J. Potocnik, and B. Stockinger, "Ablation of thymic export causes accelerated decay of naive CD4 T cells in the periphery because of activation by environmental antigen.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 25, pp. 8691–6, Jun. 2008.
- [75] M. D. Hazenberg, S. A. Otto, A. M. C. van Rossum, H. J. Scherpbier, R. de Groot, T. W. Kuijpers, J. M. A. Lange, D. Hamann, R. J. de Boer, J. A. M. Borghans, and F. Miedema, "Establishment of the CD4+ T-cell pool in healthy children and untreated children infected with HIV-1.," *Blood*, vol. 104, no. 12, pp. 3513–9, Dec. 2004.
- [76] I. Bains, R. Antia, R. Callard, and A. J. Yates, "Quantifying the development of the peripheral naive CD4+ T-cell pool in humans.," *Blood*, vol. 113, no. 22, pp. 5480–7, May 2009.

- [77] K. Naylor, G. Li, A. N. Vallejo, W.-W. Lee, K. Koetz, E. Bryl, J. Witkowski, J. Fulbright, C. M. Weyand, and J. J. Goronzy, "The influence of age on T cell generation and TCR diversity.," *J. Immunol.*, vol. 174, no. 11, pp. 7446–52, Jun. 2005.
- [78] D. C. Douek, R. D. McFarland, P. H. Keiser, E. A. Gage, J. M. Massey, B. F. Haynes, M. A. Polis, A. T. Haase, M. B. Feinberg, J. L. Sullivan, B. D. Jamieson, J. A. Zack, L. J. Picker, and R. A. Koup, "Changes in thymic function with age and during the treatment of HIV infection.," *Nature*, vol. 396, no. 6712, pp. 690–5, Dec. 1998.
- [79] A. A. Sadighi Akha and R. A. Miller, "Signal transduction in the aging immune system.," *Curr. Opin. Immunol.*, vol. 17, no. 5, pp. 486–91, Oct. 2005.
- [80] N. Khan, A. Hislop, N. Gudgeon, M. Cobbold, R. Khanna, L. Nayak, A. B. Rickinson, and P. A. H. Moss, "Herpesvirus-specific CD8 T cell immunity in old age: cytomegalovirus impairs the response to a coresident EBV infection.," *J. Immunol.*, vol. 173, no. 12, pp. 7481–9, Dec. 2004.
- [81] I. Messaoudi, J. Lemaoult, J. A. Guevara-Patino, B. M. Metzner, and J. Nikolich-Zugich, "Age-related CD8 T cell clonal expansions constrict CD8 T cell repertoire and have the potential to impair immune defense.," *J. Exp. Med.*, vol. 200, no. 10, pp. 1347–58, Nov. 2004.
- [82] A. M. Wertheimer, M. S. Bennett, B. Park, J. L. Uhrlaub, C. Martinez, V. Pulko, N. L. Currier, D. Nikolich-Žugich, J. Kaye, and J. Nikolich-Žugich, "Aging and cytomegalovirus infection differentially and jointly affect distinct circulating T cell subsets in humans.," *J. Immunol.*, vol. 192, no. 5, pp. 2143–55, Mar. 2014.
- [83] N. H. Son, S. Murray, J. Yanovski, R. J. Hodes, and N. Weng, "Lineage-specific telomere shortening and unaltered capacity for telomerase expression in human T and B lymphocytes with age.," *J. Immunol.*, vol. 165, no. 3, pp. 1191–6, Aug. 2000.
- [84] J. W. Shay and W. E. Wright, "Hayflick, his limit, and cellular ageing.," *Nat. Rev. Mol. Cell Biol.*, vol. 1, no. 1, pp. 72–6, Oct. 2000.
- [85] R. B. Effros, "The role of CD8 T cell replicative senescence in human aging.," *Discov. Med.*, vol. 5, no. 27, pp. 293–7, Jun. 2005.

- [86] F. F. Fagnoni, R. Vescovini, G. Passeri, G. Bologna, M. Pedrazzoni, G. Lavagetto, A. Casti, C. Franceschi, M. Passeri, and P. Sansoni, "Shortage of circulating naive CD8(+) T cells provides new insights on immunodeficiency in aging.," *Blood*, vol. 95, no. 9, pp. 2860–8, May 2000.
- [87] A. Wack, A. Cossarizza, S. Heltai, D. Barbieri, S. D'Addato, C. Franceschi, P. Dellabona, and G. Casorati, "Age-related modifications of the human alphabeta T cell repertoire due to different clonal expansions in the CD4+ and CD8+ subsets.," *Int. Immunol.*, vol. 10, no. 9, pp. 1281–8, Sep. 1998.
- [88] D. N. Posnett, R. Sinha, S. Kabak, and C. Russo, "Clonal populations of T cells in normal elderly humans: the T cell equivalent to 'benign monoclonal gammopathy'.," *J. Exp. Med.*, vol. 179, no. 2, pp. 609–18, Feb. 1994.
- [89] R. D. Kilpatrick, T. Rickabaugh, L. E. Hultin, P. Hultin, M. A. Hausner, R. Detels, J. Phair, and B. D. Jamieson, "Homeostasis of the naive CD4+ T cell compartment during aging.," *J. Immunol.*, vol. 180, no. 3, pp. 1499–507, Feb. 2008.
- [90] S. D. Boyd, Y. Liu, C. Wang, V. Martin, and D. K. Dunn-Walters, "Human lymphocyte repertoires in ageing.," *Curr. Opin. Immunol.*, vol. 25, no. 4, pp. 511–5, Aug. 2013.
- [91] J. R. Currier and M. A. Robinson, "Spectratype/immunoscope analysis of the expressed TCR repertoire.," *Curr. Protoc. Immunol.*, vol. Chapter 10, p. Unit 10.28, May 2001.
- [92] S. Verfuërth, K. Peggs, P. Vyas, L. Barnett, R. J. O'Reilly, and S. Mackinnon, "Longitudinal monitoring of immune reconstitution by CDR3 size spectratyping after T-cell-depleted allogeneic bone marrow transplant and the effect of donor lymphocyte infusions on T-cell repertoire.," *Blood*, vol. 95, no. 12, pp. 3990–5, Jun. 2000.
- [93] S. A. Memon, C. Sportès, F. A. Flomerfelt, R. E. Gress, and F. T. Hakim, "Quantitative analysis of T cell receptor diversity in clinical samples of human peripheral blood.," *J. Immunol. Methods*, vol. 375, no. 1–2, pp. 84–92, Jan. 2012.
- [94] C. J. Wu, A. Chillemi, E. P. Alyea, E. Orsini, D. Neuberg, R. J. Soiffer, and J. Ritz, "Reconstitution of T-cell receptor repertoire diversity following T-cell depleted

- allogeneic bone marrow transplantation is related to hematopoietic chimerism.," *Blood*, vol. 95, no. 1, pp. 352–9, Jan. 2000.
- [95] R. J. Britten and D. E. Kohne, "Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms.," *Science*, vol. 161, no. 3841, pp. 529–40, Aug. 1968.
- [96] P. D. Baum and J. M. McCune, "Direct measurement of T-cell receptor repertoire diversity with AmpliCot.," *Nat. Methods*, vol. 3, no. 11, pp. 895–901, Nov. 2006.
- [97] M. L. Metzker, "Sequencing technologies - the next generation.," *Nat. Rev. Genet.*, vol. 11, no. 1, pp. 31–46, Jan. 2010.
- [98] C. Arnold and I. J. Hodgson, "Vectorette PCR: a novel approach to genomic walking.," *PCR Methods Appl.*, vol. 1, no. 1, pp. 39–42, Aug. 1991.
- [99] A. D. Bhalla, J. P. Gudikote, J. Wang, W.-K. Chan, Y.-F. Chang, O. R. Olivas, and M. F. Wilkinson, "Nonsense codons trigger an RNA partitioning shift.," *J. Biol. Chem.*, vol. 284, no. 7, pp. 4062–72, Feb. 2009.
- [100] J. Wang, V. M. Vock, S. Li, O. R. Olivas, and M. F. Wilkinson, "A quality control pathway that down-regulates aberrant T-cell receptor (TCR) transcripts by a mechanism requiring UPF2 and translation.," *J. Biol. Chem.*, vol. 277, no. 21, pp. 18489–93, May 2002.
- [101] J. D. Freeman, R. L. Warren, J. R. Webb, B. H. Nelson, and R. a Holt, "Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing.," *Genome Res.*, vol. 19, no. 10, pp. 1817–24, Oct. 2009.
- [102] I. Z. Mamedov, O. V Britanova, D. A. Bolotin, A. V Chkalina, D. B. Staroverov, I. V Zvyagin, A. A. Kotlobay, M. A. Turchaninova, D. A. Fedorenko, A. A. Novik, G. V Sharonov, S. Lukyanov, D. M. Chudakov, and Y. B. Lebedev, "Quantitative tracking of T cell clones after haematopoietic stem cell transplantation.," *EMBO Mol. Med.*, vol. 3, no. 4, pp. 201–7, Apr. 2011.

- [103] M. Matz, D. Shagin, E. Bogdanova, O. Britanova, S. Lukyanov, L. Diatchenko, and A. Chenchik, "Amplification of cDNA ends based on template-switching effect and step-out PCR.," *Nucleic Acids Res.*, vol. 27, no. 6, pp. 1558–60, Mar. 1999.
- [104] D. A. Bolotin, I. Z. Mamedov, O. V Britanova, I. V Zvyagin, D. Shagin, S. V Ustyugova, M. A. Turchaninova, S. Lukyanov, Y. B. Lebedev, and D. M. Chudakov, "Next generation sequencing for TCR repertoire profiling: platform-specific features and correction algorithms.," *Eur. J. Immunol.*, vol. 42, no. 11, pp. 3073–83, Nov. 2012.
- [105] P. Nguyen, J. Ma, D. Pei, C. Obert, C. Cheng, and T. L. Geiger, "Identification of errors introduced during high throughput sequencing of the T cell receptor repertoire.," *BMC Genomics*, vol. 12, p. 106, Jan. 2011.
- [106] R. L. Warren, J. D. Freeman, T. Zeng, G. Choe, S. Munro, R. Moore, J. R. Webb, and R. a Holt, "Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes.," *Genome Res.*, vol. 21, no. 5, pp. 790–7, May 2011.
- [107] I. Z. Mamedov, O. V Britanova, I. V Zvyagin, M. a Turchaninova, D. a Bolotin, E. V Putintseva, Y. B. Lebedev, and D. M. Chudakov, "Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling.," *Front. Immunol.*, vol. 4, no. December, p. 456, Jan. 2013.
- [108] M. P. Lefranc, V. Giudicelli, C. Busin, A. Malik, I. Mougnot, P. Déhais, and D. Chaume, "LIGM-DB/IMGT: an integrated database of Ig and TcR, part of the immunogenetics database.," *Ann. N. Y. Acad. Sci.*, vol. 764, pp. 47–9, Sep. 1995.
- [109] D. a Bolotin, M. Shugay, I. Z. Mamedov, E. V Putintseva, M. a Turchaninova, I. V Zvyagin, O. V Britanova, and D. M. Chudakov, "MiTCR: software for T-cell receptor sequencing data analysis.," *Nat. Methods*, vol. 10, no. 9, pp. 813–4, Sep. 2013.
- [110] J. Lin, "Divergence Measures Based on the Shannon Entropy," *IEEE TRANSACTIONS ON INFORMATION THEORY.*, 1991.

- [111] B. EFRON and R. THISTED, “Estimating the number of unseen species: How many words did Shakespeare know?,” *Biometrika*, vol. 63, no. 3, pp. 435–447, Dec. 1976.
- [112] A. C. Zenclussen, “Adaptive immune responses during pregnancy.,” *Am. J. Reprod. Immunol.*, vol. 69, no. 4, pp. 291–303, Apr. 2013.
- [113] J. L. Nelson, D. E. Furst, S. Maloney, T. Gooley, P. C. Evans, A. Smith, M. A. Bean, C. Ober, and D. W. Bianchi, “Microchimerism and HLA-compatible relationships of pregnancy in scleroderma.,” *Lancet*, vol. 351, no. 9102, pp. 559–62, Feb. 1998.
- [114] K. Sarkar and F. W. Miller, “Possible roles and determinants of microchimerism in autoimmune and other disorders.,” *Autoimmun. Rev.*, vol. 3, no. 6, pp. 454–63, Aug. 2004.
- [115] T. Lepez, M. Vandewoestyne, S. Hussain, F. Van Nieuwerburgh, K. Poppe, B. Velkeniers, J.-M. Kaufman, and D. Deforce, “Fetal microchimeric cells in blood of women with an autoimmune thyroid disease.,” *PLoS One*, vol. 6, no. 12, p. e29646, Jan. 2011.
- [116] W. J. Burlingham and G. Benichou, “Bidirectional alloreactivity: A proposed microchimerism-based solution to the NIMA paradox.,” *Chimerism*, vol. 3, no. 2, pp. 29–36.
- [117] E. Jankowska-Gan, A. Sheka, H. W. Sollinger, J. D. Pirsch, R. M. Hofmann, L. D. Haynes, M. J. Armbrust, J. D. Mezrich, and W. J. Burlingham, “Pretransplant immune regulation predicts allograft outcome: bidirectional regulation correlates with excellent renal transplant function in living-related donor-recipient pairs.,” *Transplantation*, vol. 93, no. 3, pp. 283–90, Feb. 2012.
- [118] M. Stern, L. Ruggeri, A. Mancusi, M. E. Bernardo, C. de Angelis, C. Bucher, F. Locatelli, F. Aversa, and A. Velardi, “Survival after T cell-depleted haploidentical stem cell transplantation is improved using the mother as donor.,” *Blood*, vol. 112, no. 7, pp. 2990–5, Oct. 2008.
- [119] B. J. Manfras, D. Terjung, and B. O. Boehm, “Non-productive human TCR beta chain genes represent V-D-J diversity before selection upon function: insight into biased

- usage of TCRBD and TCRBJ genes and diversity of CDR3 region length.," *Hum. Immunol.*, vol. 60, no. 11, pp. 1090–100, Nov. 1999.
- [120] I. V Zvyagin, M. V Pogorelyy, M. E. Ivanova, E. A. Komech, M. Shugay, D. A. Bolotin, A. A. Shelenkov, A. A. Kurnosov, D. B. Staroverov, D. M. Chudakov, Y. B. Lebedev, and I. Z. Mamedov, "Distinctive properties of identical twins' TCR repertoires revealed by high-throughput sequencing.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 111, no. 16, pp. 5980–5, Apr. 2014.
- [121] D. Favre, C. A. Stoddart, B. Emu, R. Hoh, J. N. Martin, F. M. Hecht, S. G. Deeks, and J. M. McCune, "HIV disease progression correlates with the generation of dysfunctional naive CD8(low) T cells.," *Blood*, vol. 117, no. 7, pp. 2189–99, Feb. 2011.
- [122] T. P. Arstila, "A Direct Estimate of the Human T Cell Receptor Diversity," *Science (80)*, vol. 286, no. 5441, pp. 958–961, Oct. 1999.
- [123] K. C. Garcia, J. J. Adams, D. Feng, and L. K. Ely, "The molecular basis of TCR germline bias for MHC is surprisingly simple.," *Nat. Immunol.*, vol. 10, no. 2, pp. 143–7, Feb. 2009.
- [124] K. M. Adams and J. L. Nelson, "Microchimerism: an investigative frontier in autoimmunity and transplantation.," *JAMA*, vol. 291, no. 9, pp. 1127–31, Mar. 2004.
- [125] J. L. Nelson, "Your cells are my cells.," *Sci. Am.*, vol. 298, no. 2, pp. 64–71, Feb. 2008.
- [126] P. C. Evans, N. Lambert, S. Maloney, D. E. Furst, J. M. Moore, and J. L. Nelson, "Long-term fetal microchimerism in peripheral blood mononuclear cell subsets in healthy women and women with scleroderma.," *Blood*, vol. 93, no. 6, pp. 2033–7, Mar. 1999.
- [127] L. S. Loubière, N. C. Lambert, L. J. Flinn, T. D. Erickson, Z. Yan, K. a Guthrie, K. T. Vickers, and J. L. Nelson, "Maternal microchimerism in healthy adults in lymphocytes, monocyte/macrophages and NK cells.," *Lab. Invest.*, vol. 86, no. 11, pp. 1185–92, Nov. 2006.
- [128] A. M. Jonsson, M. Uzunel, C. Götherström, N. Papadogiannakis, and M. Westgren, "Maternal microchimerism in human fetal tissues.," *Am. J. Obstet. Gynecol.*, vol. 198, no. 3, pp. 325.e1–6, Mar. 2008.

- [129] J. E. Mold, J. Michaëlsson, T. D. Burt, M. O. Muench, K. P. Beckerman, M. P. Busch, T.-H. Lee, D. F. Nixon, and J. M. McCune, “Maternal alloantigens promote the development of tolerogenic fetal regulatory T cells in utero.,” *Science*, vol. 322, no. 5907, pp. 1562–5, Dec. 2008.
- [130] C. Götherstrom, A. M. Johnsson, J. Mattsson, N. Papadogiannakis, and M. Westgren, “Identification of maternal hematopoietic cells in a 2nd-trimester fetus.,” *Fetal Diagn. Ther.*, vol. 20, no. 5, pp. 355–8.
- [131] K. Khosrotehrani, M. Leduc, V. Bachy, S. Nguyen Huu, M. Oster, A. Abbas, S. Uzan, and S. Aractingi, “Pregnancy allows the transfer and differentiation of fetal lymphoid progenitors into functional T and B cells in mothers.,” *J. Immunol.*, vol. 180, no. 2, pp. 889–97, Jan. 2008.
- [132] W. J. Burlingham and J. L. Nelson, “Microchimerism in cord blood: mother as anticancer drug.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 109, no. 7, pp. 2190–1, Feb. 2012.
- [133] J. L. Nelson, “The otherness of self: microchimerism in health and disease.,” *Trends Immunol.*, vol. 33, no. 8, pp. 421–7, Aug. 2012.
- [134] M. A. Hall, J. L. Reid, and J. S. Lanchbury, “The distribution of human TCR junctional region lengths shifts with age in both CD4 and CD8 T cells.,” *Int. Immunol.*, vol. 10, no. 10, pp. 1407–19, Oct. 1998.
- [135] M. A. Turchaninova, O. V Britanova, D. A. Bolotin, M. Shugay, E. V Putintseva, D. B. Staroverov, G. Sharonov, D. Shcherbo, I. V Zvyagin, I. Z. Mamedov, C. Linnemann, T. N. Schumacher, and D. M. Chudakov, “Pairing of T-cell receptor chains via emulsion PCR.,” *Eur. J. Immunol.*, vol. 43, no. 9, pp. 2507–15, Sep. 2013.
- [136] M. I. Eren, A. Chao, W.-H. Hwang, and R. K. Colwell, “Estimating the richness of a population when the maximum number of classes is fixed: a nonparametric solution to an archaeological problem.,” *PLoS One*, vol. 7, no. 5, p. e34179, Jan. 2012.
- [137] C. Linnemann, B. Heemskerk, P. Kvistborg, R. J. C. Kluin, D. A. Bolotin, X. Chen, K. Bresser, M. Nieuwland, R. Schotte, S. Michels, R. Gomez-Eerland, L. Jahn, P. Hombrink, N. Legrand, C. J. Shu, I. Z. Mamedov, A. Velds, C. U. Blank, J. B. A. G. Haanen, M. A.

- Turchaninova, R. M. Kerkhoven, H. Spits, S. R. Hadrup, M. H. M. Heemskerk, T. Blankenstein, D. M. Chudakov, G. M. Bendle, and T. N. M. Schumacher, "High-throughput identification of antigen-specific TCRs by TCR gene capture.," *Nat. Med.*, vol. 19, no. 11, pp. 1534–41, Nov. 2013.
- [138] D. C. Douek, M. R. Betts, J. M. Brenchley, B. J. Hill, D. R. Ambrozak, K.-L. Ngai, N. J. Karandikar, J. P. Casazza, and R. A. Koup, "A novel approach to the analysis of specificity, clonality, and frequency of HIV-specific T cell responses reveals a potential mechanism for control of viral escape.," *J. Immunol.*, vol. 168, no. 6, pp. 3099–104, Mar. 2002.
- [139] T. Kivioja, A. Vähärautio, K. Karlsson, M. Bonke, M. Enge, S. Linnarsson, and J. Taipale, "Counting absolute numbers of molecules using unique molecular identifiers.," *Nat. Methods*, vol. 9, no. 1, pp. 72–4, Jan. 2012.
- [140] J. A. Casbon, R. J. Osborne, S. Brenner, and C. P. Lichtenstein, "A method for counting PCR template molecules with application to next-generation sequencing.," *Nucleic Acids Res.*, vol. 39, no. 12, p. e81, Jul. 2011.
- [141] I. M. Rea, S. E. McNerlan, and H. D. Alexander, "CD69, CD25, and HLA-DR activation antigen expression on CD3+ lymphocytes and relationship to serum TNF-alpha, IFN-gamma, and sIL-2R levels in aging.," *Exp. Gerontol.*, vol. 34, no. 1, pp. 79–93, Jan. 1999.
- [142] J. E. Nagel, F. J. Chrest, and W. H. Adler, "Enumeration of T lymphocyte subsets by monoclonal antibodies in young and aged humans.," *J. Immunol.*, vol. 127, no. 5, pp. 2086–8, Nov. 1981.
- [143] S. E. McNerlan, I. M. Rea, and H. D. Alexander, "A whole blood method for measurement of intracellular TNF-alpha, IFN-gamma and IL-2 expression in stimulated CD3+ lymphocytes: differences between young and elderly subjects.," *Exp. Gerontol.*, vol. 37, no. 2–3, pp. 227–34.
- [144] G. C. Wang, P. Dash, J. A. McCullers, P. C. Doherty, and P. G. Thomas, "T cell receptor $\alpha\beta$ diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection.," *Sci. Transl. Med.*, vol. 4, no. 128, p. 128ra42, Apr. 2012.

- [145] R. Solana, R. Tarazona, A. E. Aiello, A. N. Akbar, V. Appay, M. Beswick, J. A. Bosch, C. Campos, S. Cantisán, L. Cicin-Sain, E. Derhovanessian, S. Ferrando-Martínez, D. Frasca, T. Fulöp, S. Govind, B. Grubeck-Loebenstein, A. Hill, M. Hurme, F. Kern, A. Larbi, M. López-Botet, A. B. Maier, J. E. McElhane, P. Moss, E. Naumova, J. Nikolich-Zugich, A. Pera, J. L. Rector, N. Riddell, B. Sanchez-Correa, P. Sansoni, D. Sauce, R. van Lier, G. C. Wang, M. R. Wills, M. Zieliński, and G. Pawelec, “CMV and Immunosenescence: from basics to clinics.,” *Immun. Ageing*, vol. 9, no. 1, p. 23, Jan. 2012.
- [146] S. Ferrando-Martínez, E. Ruiz-Mateos, A. Hernández, E. Gutiérrez, M. del M. Rodríguez-Méndez, A. Ordoñez, and M. Leal, “Age-related deregulation of naive T cell homeostasis in elderly humans.,” *Age (Dordr)*, vol. 33, no. 2, pp. 197–207, Jun. 2011.
- [147] J. Yan, J. M. Greer, R. Hull, J. D. O’Sullivan, R. D. Henderson, S. J. Read, and P. A. McCombe, “The effect of ageing on human lymphocyte subsets: comparison of males and females.,” *Immun. Ageing*, vol. 7, p. 4, Jan. 2010.
- [148] R. Emerson, A. Sherwood, C. Desmarais, S. Malhotra, D. Phippard, and H. Robins, “Estimating the ratio of CD4+ to CD8+ T cells using high-throughput sequence data.,” *J. Immunol. Methods*, vol. 391, no. 1–2, pp. 14–21, May 2013.
- [149] D. A. Price, T. E. Asher, N. A. Wilson, M. C. Nason, J. M. Brenchley, I. S. Metzler, V. Venturi, E. Gostick, P. K. Chattopadhyay, M. Roederer, M. P. Davenport, D. I. Watkins, and D. C. Douek, “Public clonotype usage identifies protective Gag-specific CD8+ T cell responses in SIV infection.,” *J. Exp. Med.*, vol. 206, no. 4, pp. 923–36, Apr. 2009.
- [150] M. C. Iglesias, J. R. Almeida, S. Fastenackels, D. J. van Bockel, M. Hashimoto, V. Venturi, E. Gostick, A. Urrutia, L. Wooldridge, M. Clement, S. Gras, P. G. Wilmann, B. Autran, A. Moris, J. Rossjohn, M. P. Davenport, M. Takiguchi, C. Brander, D. C. Douek, A. D. Kelleher, D. A. Price, and V. Appay, “Escape from highly effective public CD8+ T-cell clonotypes by HIV.,” *Blood*, vol. 118, no. 8, pp. 2138–49, Aug. 2011.
- [151] O. V Britanova, E. V Putintseva, M. Shugay, E. M. Merzlyak, M. a Turchaninova, D. B. Staroverov, D. a Bolotin, S. Lukyanov, E. a Bogdanova, I. Z. Mamedov, Y. B. Lebedev, and D. M. Chudakov, “Age-Related Decrease in TCR Repertoire Diversity Measured with

Deep and Normalized Sequence Profiling,” *J. Immunol.*, vol. 192, no. 6, pp. 2689–98, Mar. 2014.

- [152] E. V Putintseva, O. V Britanova, D. B. Staroverov, E. M. Merzlyak, M. a Turchaninova, M. Shugay, D. a Bolotin, M. V Pogorelyy, I. Z. Mamedov, V. Bobrynina, M. Maschan, Y. B. Lebedev, and D. M. Chudakov, “Mother and child T cell receptor repertoires: deep profiling study,” *Front. Immunol.*, vol. 4, no. December, p. 463, Jan. 2013.
- [153] M. Shugay, D. a Bolotin, E. V Putintseva, M. V Pogorelyy, I. Z. Mamedov, and D. M. Chudakov, “Huge Overlap of Individual TCR Beta Repertoires,” *Front. Immunol.*, vol. 4, no. December, p. 466, Jan. 2013.