

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

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата биологических наук

Москва 2014

Работа выполнена в Лаборатории геномики адаптивного иммунитета в Федеральном государственном бюджетном учреждении науки Институте биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова (ИБХ РАН).

Научный руководитель:

Чудаков Дмитрий Михайлович, доктор биологических наук.

Официальные оппоненты:

Атауллаханов Равшан Иноятович, доктор медицинских наук, профессор, заведующий Лабораторией активации иммунитета (№25) Федерального государственного бюджетного учреждения “ГНЦ институт иммунологии” ФМБА России.

Логачёва Мария Дмитриевна, кандидат биологических наук, ведущий научный сотрудник Лаборатории эволюционной геномики Факультета биоинженерии и биоинформатики Московского государственного университета им. М.В. Ломоносова.

Ведущая организация:

Федеральное государственное бюджетное учреждение науки Институт общей генетики им. Н.И. Вавилова РАН.

Защита диссертации состоится «8» октября 2014 года в 10:00 часов на заседании совета Д 002.019.01 при ИБХ РАН по адресу: 117997, г. Москва, ул. Миклухо-Маклая, д. 16/10.

С диссертацией можно ознакомиться в библиотеке Института биоорганической химии им. академиков М.М. Шемякина и Ю.А. Овчинникова РАН и на сайте www.ibch.ru.

Автореферат разослан « »

2014 года.

Ученый секретарь диссертационного совета,

доктор физико-математических наук

В.А. Олейников



Общая характеристика работы

Актуальность проблемы

Эффективность и сбалансированность адаптивной иммунной защиты организма от инфекционных и онкологических заболеваний в значительной степени определяется функциональным и рецепторным разнообразием Т-лимфоцитов. Каждый функциональный клон Т-лимфоцитов несет уникальный Т-клеточный рецептор (TCR), распознающий чужеродный антиген в составе молекулы главного комплекса гистосовместимости (МНС) и определяющий специфичность адаптивного иммунного ответа. Индивидуальное разнообразие наивного репертуара TCR определяет вероятность того, что для новой инфекции или онкологического заболевания найдутся клоны Т-лимфоцитов, специфично распознающие характерные для данного патогена или патологических клеток антигены. После распознавания антигена антиген-специфичные Т-лимфоциты активно размножаются, при этом их концентрация в организме и периферической крови значительно возрастает.

Индивидуальный репертуар Т-клеточных рецепторов определяет эффективность иммунной защиты, диапазон распознаваемых антигенов, особенности патологических состояний иммунитета; а также содержит в себе информацию об инфекционных, онкологических и аутоиммунных заболеваниях данного пациента.

С развитием современных методов анализа ¹ исследования Т-клеточного иммунитета постепенно выходят на более системный уровень. Это приближает нас к созданию нового поколения методов диагностики и терапии заболеваний, методов вакцинации, а также методов восстановления и поддержания эффективной функции адаптивного иммунитета.

Настоящая работа посвящена анализу индивидуальных репертуаров TCR β человека методом глубокого массированного секвенирования.

¹ Например, методов массированного секвенирования репертуаров TCR, иммунофенотипирования Т-клеточных субпопуляций, транскриптомики, в том числе с единичных клеток, статистического анализа репертуаров TCR, и их потенциальных эпитопов в составе МНС, а также методов рентгеноструктурного анализа комплексов TCR/пептид/МНС и масс-спектрометрического анализа репертуаров презентруемых пептидов.

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 и воспроизводимого сравнения разнообразия репертуаров.

В третьей части работы с помощью разработанной методики исследуются изменения, происходящие в репертуаре TCR β человека при старении.

Цель работы

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 для массированного секвенирования библиотеки генов β -цепей Т-клеточных рецепторов из образцов периферической крови группы матерей и их детей;
2. Провести сравнительный анализ состава и структуры репертуаров генов 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 для массированного секвенирования молекулярно-баркодированные библиотеки генов β -цепей Т-клеточных рецепторов из периферической крови системно здоровых доноров разных возрастов;
5. Провести сравнительный анализ нормированного разнообразия репертуаров генов β -цепей Т-клеточных рецепторов периферической крови для доноров различного возраста.

Научная новизна и практическая ценность работы

Массированное секвенирование открыло новые возможности для исследования адаптивного иммунитета на уровне индивидуальных репертуаров антител и Т-клеточных рецепторов. В настоящей работе мы применили этот метод для анализа индивидуальных репертуаров CDR3 TCR β различных когорт системно здоровых доноров. Достигнутая глубина анализа (порядка 0,5-1 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 .

В данной работе мы изучали процесс формирования индивидуальных репертуаров TCR β и влияние наследственных факторов на этот процесс, исследовали структуру индивидуальных репертуаров и степень их пересечения между различными индивидуумами, а также оценивали изменения в разнообразии и общей структуре репертуара TCR β в ходе старения.

Мы уверены, что полученные нами данные интересны как с фундаментальной точки зрения, так и с практической. В частности, мы надеемся, что они будут полезны для разработки вакцин и адаптации их к иммунитету пожилых людей, в планировании трансплантаций, а также в разработке препаратов, поддерживающих иммунитет человека в старости. Кроме того, разработанный нами способ сравнительного анализа репертуаров TCR с помощью метода молекулярного баркодирования позволяет проводить точное и аккуратное сравнение образцов, и мы надеемся, что этот метод будет полезен в дальнейших исследованиях репертуаров TCR и антител.

Структура диссертации

Диссертационная работа изложена на 114 страницах и состоит из введения, обзора литературы, описания материалов и методов исследования, результатов и их обсуждения, выводов, заключения, списка сокращений, а также списка литературы, включающего 153 источника. Диссертация содержит 27 рисунков и 4 таблицы.

Апробация работы

Основные результаты диссертации были доложены на конференции «ZIBI summer symposium», 2013, Берлин, Германия.

Публикации

По результатам работы опубликовано 7 статей в рецензируемых зарубежных журналах.

Содержание работы

1. Сравнительный анализ репертуаров TCR β матерей и детей

Мы использовали массированное секвенирование для сравнения индивидуальных репертуаров TCR β матерей и их детей. Мы стремились выявить специфические характеристики, свойственные репертуарам TCR β родственных пар, чтобы определить степень влияния наследственности на формирование индивидуальных репертуаров. Кроме того, мы анализировали влияние отбора в тимусе на формирование индивидуального репертуара TCR β человека.

Взаимодействие иммунитетов матери и плода активно изучается в контексте осложнений в ходе беременности, аутоиммунных заболеваний, а также гаплоидентичных трансплантаций органов и гематопозитических стволовых клеток. Достигнутая глубина секвенирования позволила нам исследовать наличие зрелых микрохимерных Т-лимфоцитов матери или плода (то есть Т-клеток, мигрировавших через плаценту в качестве зрелых α/β Т-лимфоцитов 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 нашего программного обеспечения MiTCR было идентифицировано 500 000 – 2 000 000 различных клонотипов TCR β в образцах каждого из доноров. Полученные индивидуальные репертуары TCR β мы использовали для сравнительного анализа (Рисунок 1).

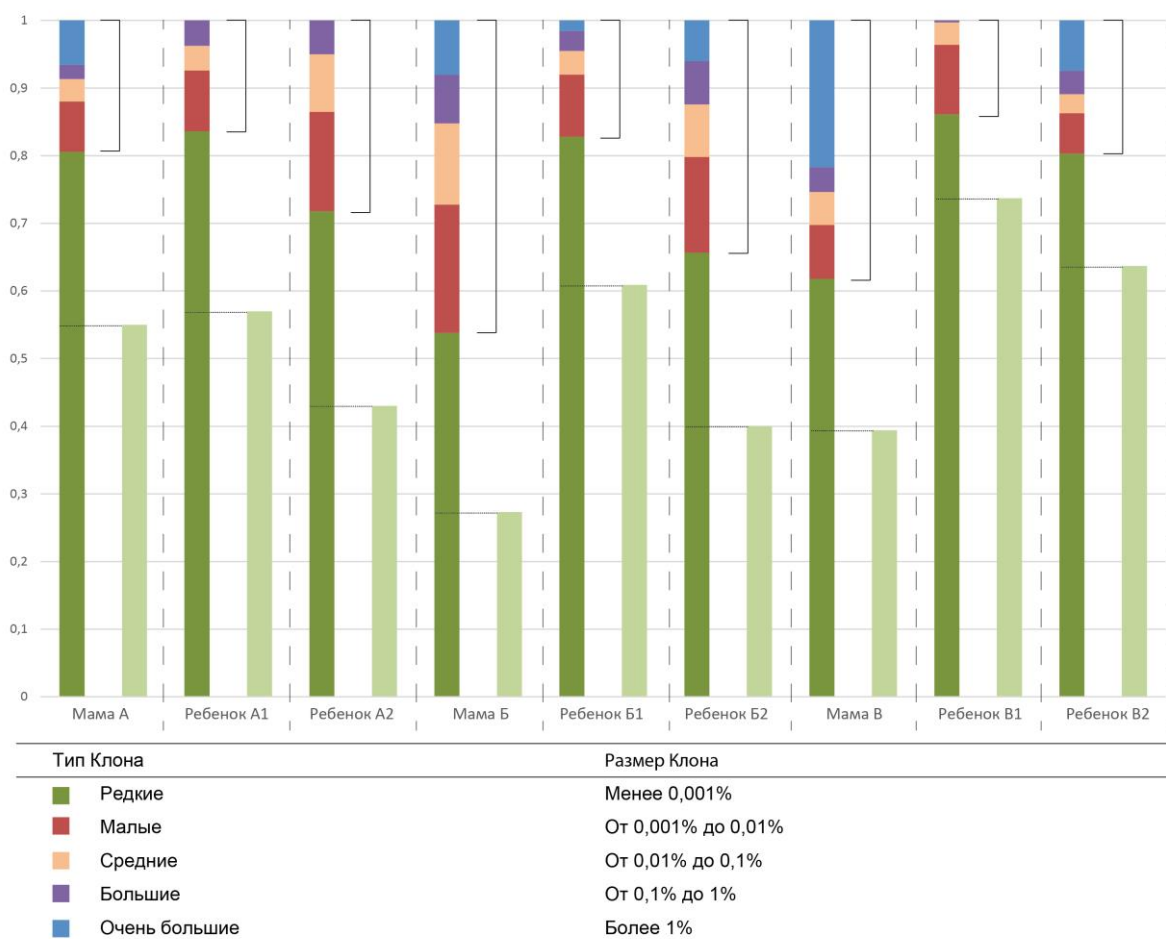


Рисунок 1. Представленность клонов Т-лимфоцитов разной частоты встречаемости в индивидуальных репертуарах TCRβ.

Разноцветные столбцы представляют собой Т-клеточное гомеостатическое пространство каждого из девяти доноров. В них цветами показана частота встречаемости клонов Т-лимфоцитов.

Светло-зеленые столбцы соответствуют доле наивных CD27^{high}CD45RA^{high} Т-клеток, измеренную с помощью проточной цитофлуориметрии.

Темные горизонтальные линии отсекают долю низкочастотных TCRβ клонотипов, эквивалентную содержанию наивных Т-лимфоцитов в периферической крови каждого донора. Квадратные скобки показывают долю высокочастотных клонотипов (более 0.001% от общего числа прочтений секвенирования).

1.1. Распределение частот использования Vβ сегментов

Существует 54 варианта генов Vβ-сегмента у человека, вносящих большой вклад в общее индивидуальное разнообразие TCRβ. Гены этого сегмента кодируют CDR1 и CDR2 участки TCRβ, взаимодействующие с молекулой MHC в составе комплекса пептид-MHC. Мы проанализировали относительное распределение использования Vβ сегментов в

репертуарах TCR β пар мама-ребенок на трех уровнях: на уровне репертуаров нефункциональных TCR β , репертуаров TCR β низкочастотных клонотипов и репертуаров TCR β высокочастотных клонотипов.

1) *Распределение частот использования V β сегментов в репертуарах нефункциональных TCR β*

Нефункциональные варианты TCR β - гены со сбитой рамкой считывания или со стоп-кодом в области CDR3 - находятся на гомологичной хромосоме примерно в 40% Т-лимфоцитов. Такие варианты присутствуют наряду с функциональным вариантом Т-клеточного рецептора в качестве пассивного балласта, не экспрессируются, и, соответственно, *не подвергаются отбору в тимусе*. Поэтому по их репертуару можно судить об особенностях исходной сборки TCR в ходе рекомбинационных событий.

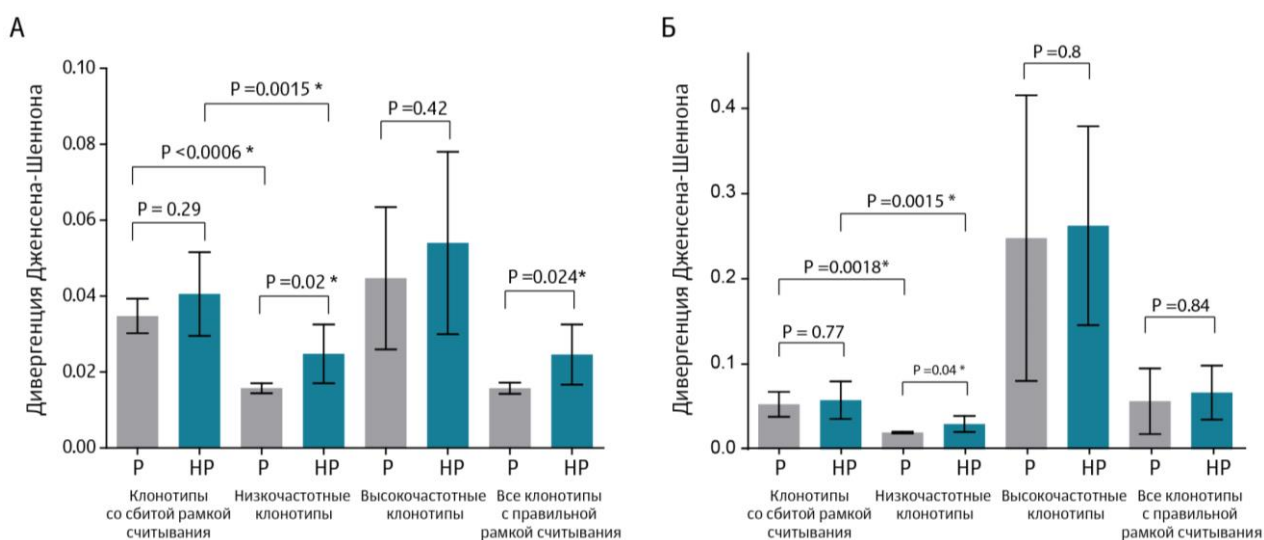


Рисунок 2. Средние значения дивергенции Дженсена-Шеннона представленности V β -сегментов в репертуарах TCR родственных (P, серые столбцы) и неродственных (HP, синие столбцы) пар мама-ребенок (с указанием стандартного отклонения). P-значения с поправкой Бонферрони представлены с учетом множественных выборок.

(А) Сравнение представленности V β -сегментов среди общего числа TCR β клонотипов (но не ридов – то есть влияние относительных размеров исследуемых клонотипов TCR β не учитывается).

(Б) Сравнение представленности V β -сегментов среди общего числа ридов – 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.

Сравнительный анализ репертуаров TCR β мам и детей на этом уровне показал, что распределение использования V β -сегментов сопоставимо для родственных и неродственных пар мама-ребенок. У родственных пар мама-ребенок наблюдается некоторое (статистически недостоверное) уменьшение значения дивергенции Дженсена-Шеннона² (Рисунок 2, два первых столбца), что может свидетельствовать о большем сходстве распределения частот использования V β -сегментов в ходе рекомбинации.

2) Распределение частот использования V β сегментов в репертуарах TCR β низкочастотных клонотипов

Мы предположили, что влияние селекции в тимусе на распределение представленности V β -сегментов можно оценить, сравнив репертуар нефункциональных TCR β с репертуаром TCR β наивных Т-клеток.

Долю наивных Т-лимфоцитов (CD3⁺CD27^{high}CD45RA^{high}) в образцах периферической крови каждого донора мы определяли с помощью метода проточной цитофлуориметрии.

Мы предположили, что фракция низкочастотных клонотипов, составляющая ту же долю общего количества ридов (полученного для каждого из доноров), что и определенная с помощью метода проточной цитофлуориметрии доля наивных Т-клеток в периферической крови этого донора, включает в себя преимущественно наивные Т-клетки (Рисунок 1).

Распределение использования V β -сегментов на этом уровне рассмотрения было более сходным для всех доноров, чем на уровне нефункциональных клонотипов (Рисунок 2). По всей вероятности, это является следствием дальнейшего формирования репертуара наивных Т-лимфоцитов под действием их отбора в тимусе. В таком случае 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.

Представленность V β -сегментов в репертуарах функциональных TCR β на уровне низкочастотных Т-лимфоцитов (то есть после позитивной и негативной селекции в тимусе) оказалась достоверно ($P = 0.02$) более сходной у родственных, чем у неродственных пар (Рисунок 2, столбцы 3 и 4).

В соответствии с анализом Дженсена-Шеннона, в репертуарах TCR β родственных групп доноров заметна тенденция к выравниванию распределения использования V β -

² Дивергенция Дженсена-Шеннона – статистический метод для определения близости двух распределений вероятностей.

сегментов в низкочастотных клонотипах по сравнению с нефункциональными клонотипами. Интересно, что наблюдаемые изменения в целом схожи и для неродственных доноров. Зависимость степени сближения частоты использования V β -сегментов от количества общих аллелей HLA оказалась недостоверной ($R = 0.12$, $P = 0.63$). То есть, по всей вероятности, тимус оказывает значительное и схожее для разных индивидуумов давление на формирование наивного репертуара TCR β в отношении предпочтительно отбираемых V-сегментов.

3) *Распределение частот использования V β сегментов в репертуарах TCR β высокочастотных клонотипов*

Мы предположили, что влияние встречи Т-лимфоцитов со специфичным антигеном на разнообразие используемых V-сегментов может быть отслежено при сравнении репертуаров TCR наивных и активированных Т-лимфоцитов. В популяцию высокочастотных клонотипов (преимущественно состоящую из активированных Т-лимфоцитов) мы включили те, доля прочтений секвенирования которых превышает 0.001% от всех полученных CDR3-содержащих сиквенсов. Таким образом, нижняя граница этой группы приблизительно на порядок превосходит верхнюю границу группы низкочастотных клонотипов каждого из доноров (Рисунок 1). Это позволяет снизить количество “засоров” наивными Т-клонотипами группы высокочастотных клонотипов.

Мы показали, что высокочастотные клонотипы характеризуется повышенной гетерогенностью используемых V β -сегментов. При этом достоверно большего *сходства* распределения частот использования V β -сегментов в репертуарах родственных пар по сравнению с неродственными обнаружено не было.

Повышенная гетерогенность используемых V β -сегментов может являться следствием увеличения разброса за счет сравнительно небольшого количества высокочастотных клонотипов, входящих в анализ. Также его причиной может быть случайный характер встречи специфичных Т-клонотипов и антигенов, приводящий к экспансии разных клонотипов в разных индивидуумах – даже при иммунном ответе против одних и тех же антигенов.

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

Исследования последних лет выявили значительные межиндивидуальные пересечения репертуаров TCR β . Тем не менее, до сих пор неизвестно, влияет ли гаплоидентичность доноров (то есть идентичность половины аллелей HLA) на степень пересечения их репертуаров TCR.

Чтобы ответить на этот вопрос, мы провели сравнительный анализ репертуаров TCR β родственных и неродственных 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. Для корректного сравнения степени перекрытия репертуаров TCR β в разных субпопуляциях, мы нормализовали количество совпадающих CDR3 вариантов на объемы сравниваемых выборок по формуле:

$$[\text{нормализованное количество общих TCR}\beta \text{ CDR3 клонотипов выборок А и Б}] = \frac{[\text{количество общих TCR}\beta \text{ CDR3 клонотипов выборок А и Б}]}{([\text{количество TCR}\beta \text{ CDR3 клонотипов в выборке А}] \times [\text{количество TCR}\beta \text{ CDR3 клонотипов в выборке Б}]}$$

Нормализованные результаты пересечений представлены на Рисунке 3.

На всех уровнях сравнения пересечений репертуаров CDR3 относительное количество общих клонотипов было выше для родственных пар, по сравнению с неродственными. Тем не менее, эти различия не достигали статистически значимых величин. Вероятно, для достижения статистически значимых различий необходим анализ большего количества пар доноров.

Наибольшая степень пересечения репертуаров TCR β CDR3 у всех рассматриваемых пар доноров наблюдается в субпопуляции высокочастотных клонотипов. Это может быть следствием сочетания двух причин.

Во-первых, наивные клонотипы с упрощенной структурой CDR3 продуцируются в тимусе с повышенной частотой – как внутри одного индивидуума, так и в рамках всей

популяции. Это повышает вероятность встречи таких клонотипов со специфичными антигенами и, как следствие, их экспансии.

Во-вторых, поскольку все изучаемые нами доноры живут в схожей 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 Т-лимфоцитов разных доноров.

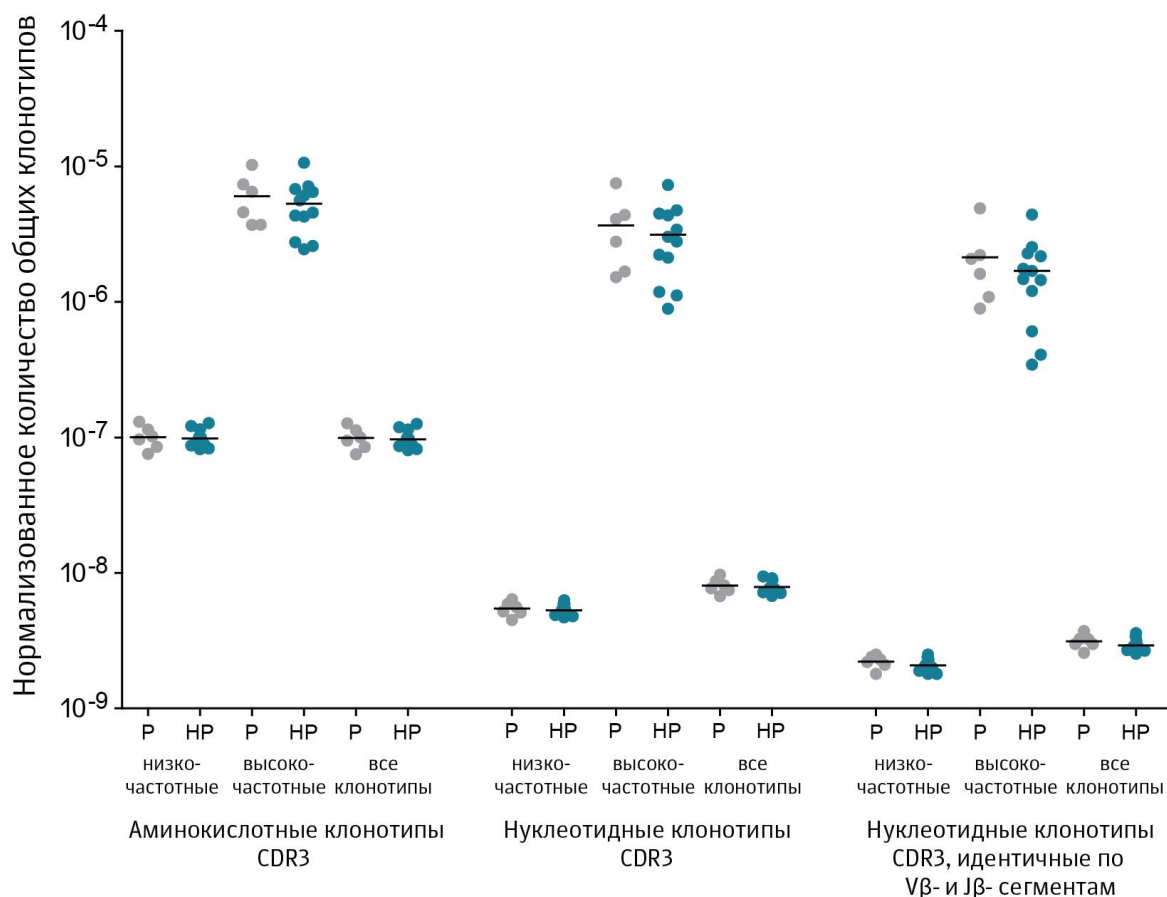


Рисунок 3. Относительное количество общих TCRβ CDR3 клонотипов пар родственных и неродственных доноров.

Пересечения показаны на уровне аминокислотных последовательностей CDR3, нуклеотидных последовательностей CDR3 и нуклеотидных последовательностей CDR3 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. На графике представлены нормализованные значения пересечений (см. в тексте). P – родственные пары, NP – неродственные пары.

1.3. Одинаковые аминокислотные последовательности CDR3 чаще сочетаются с совпадающим V β -сегментом в высокочастотных клонотипах родственных пар 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 МНС. 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 анализ вошло по 10 000 наиболее часто встречающихся аминокислотных CDR3 клонотипов каждого из доноров. Мы вычисляли пересечения репертуаров этих клонотипов в родственных и неродственных парах, а также их пересечения при условии совпадения V β -сегментов (то есть при условии совпадения CDR1, CDR2, и CDR3 участков TCR β).

Мы определяли соотношение количества аминокислотных клонотипов с общими CDR1, CDR2, CDR3 участками к количеству аминокислотных клонотипов с общим CDR3 участком (то есть соотношение V β -CDR3/CDR3). Оно всегда было выше в родственных парах.

Более того, мы наблюдали достоверную положительную корреляцию этого соотношения и количества общих аллелей МНСI у анализируемых доноров. Несмотря на то, что сравнение проводилось на несортированных Т-лимфоцитах, вклад CD8⁺ клеток был достаточным для обеспечения корреляции.

Это наблюдение, по всей видимости, показывает, что для узнавания определенного комплекса пептид-МНС CD8⁺ клетками часто необходима полная конвергенция TCR β цепи.

1.4. Селекция в тимусе снижает среднюю длину CDR3

Сравнение репертуаров функциональных и нефункциональных³ CDR3 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). То есть репертуар TCR β CDR3, возникающий в результате соматической рекомбинации (параметры которого отражены в репертуаре нефункциональных CDR3), в среднем характеризуется большей длиной. Последующая же

³ Гены TCR β со сбитой рамкой считывания или стоп-кодоном в области CDR3.

селекция в тимусе приводит к уменьшению его средней длины и количества добавленных нуклеотидов.

1.5. Поиск зрелых микрохимерных клонов Т-клеток

На сегодняшний день считается, что в ходе беременности материнские клетки и клетки плода (в частности, предшественники Т-лимфоцитов и зрелые Т-лимфоциты) активно переходят через плаценту и могут сохраняться в организме нового хозяина в течение десятков лет после рождения. Тем не менее, до сих пор неизвестно, могут ли Т-лимфоциты, перешедшие через плаценту в зрелом состоянии, сохраняться в крови нового индивидуума и работать, как функциональные клоны.

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.

В качестве маркеров таких клеток мы выбрали:

- 1) Полное совпадение V β - и J β - сегментов в сочетании с идентичными нуклеотидными последовательностями CDR3;
- 2) Количество добавленных нуклеотидов как характеристику вероятности независимого возникновения данной последовательности CDR3 в разных людях⁴.

Тем не менее, чувствительность выбранного нами метода оказалась ограничена естественной дисперсией кривых распределения количества добавленных в CDR3 нуклеотидов, характеризующих совпадение части репертуара TCR β любых двух людей.

В ходе проведенного нами сравнительного анализа мы не наблюдали присутствия зрелых микрохимерных клонов Т-лимфоцитов в периферической крови родственных доноров. Однако 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.

⁴ Чем больше добавленных нуклеотидов, тем ниже вероятность независимого возникновения данной последовательности и, соответственно, выше вероятность действительно микрохимерной природы наблюдаемого общего клонотипа.

1.6. Оценка полного пересечения индивидуальных TCR β репертуаров

Мы предприняли попытку оценить полное пересечение репертуаров TCR β двух индивидуумов. Для этого мы экстраполировали данные по пересечениям репертуаров 12 неродственных 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. При увеличении объемов выборок в арифметической прогрессии количество общих клонотипов в каждой паре росло в геометрической прогрессии (Рисунок 4 А-В).

При максимальной глубине анализа (порядка одного миллиона клонотипов для каждого из доноров) 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.

Экстраполируя эти данные на оцениваемое общее индивидуальное разнообразие TCR β , мы провели приблизительную оценку полного перекрытия CDR3 репертуаров TCR β двух индивидуумов, составившую порядка 2 200 000 аминокислотных последовательностей CDR3; 2 060 000 аминокислотных последовательностей CDR3 с несовпадающими нуклеотидными последовательностями и 180 000 нуклеотидных последовательностей CDR3 (что представляет собой 44.1%, 41.3% и 3.6% индивидуальных репертуаров, соответственно).

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.

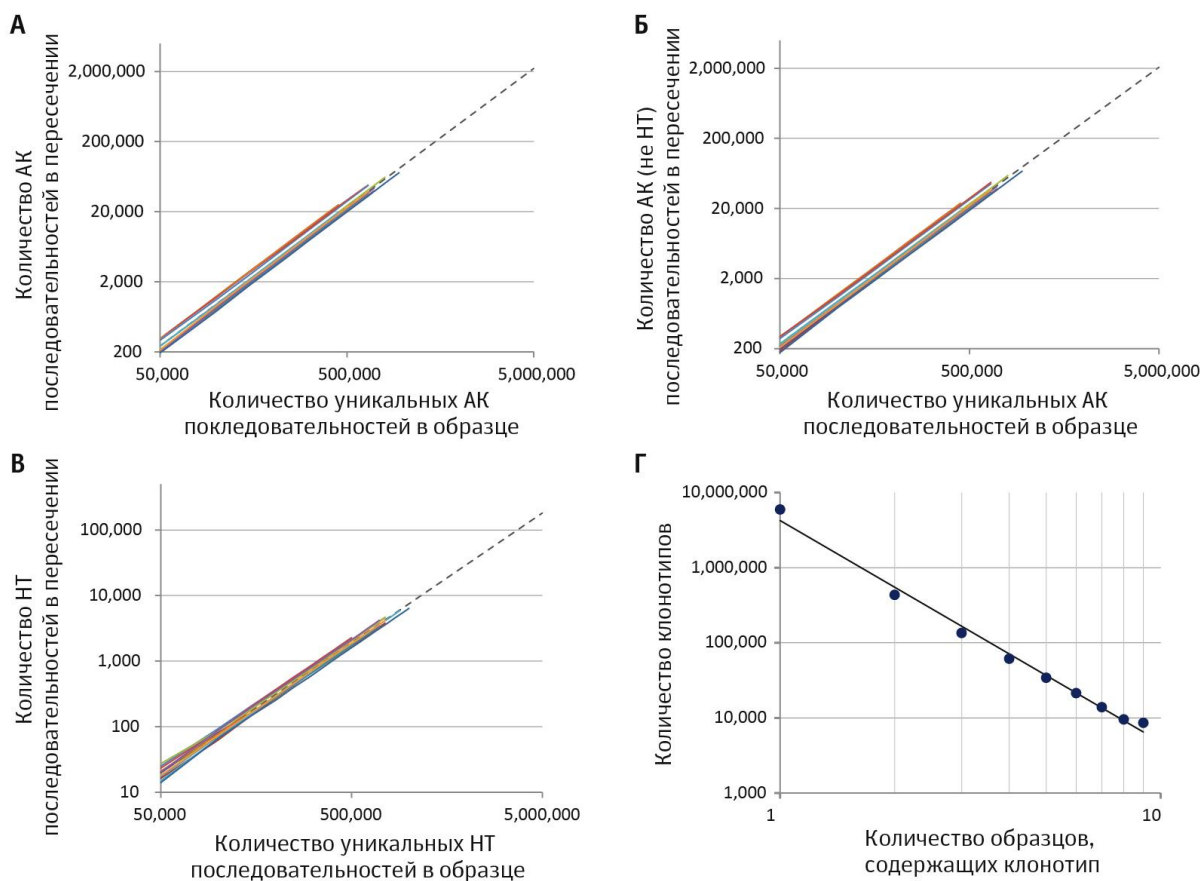


Рисунок 4. Перекрывание индивидуальных TCRβ CDR3 репертуаров растет геометрически при увеличении объема анализируемых массивов данных.

На графиках представлено количество общих TCRβ CDR3 клонотипов 12 неродственных пар доноров в зависимости от объема анализируемых образцов. Анализ пересечения репертуаров был произведен для: **(А)** аминокислотных последовательностей (АК), **(Б)** аминокислотных последовательностей с несовпадающими нуклеотидными последовательностями (АК, не НТ) и **(В)** нуклеотидных последовательностей (НТ).

Каждая из 12 линий разных цветов представляет собой наблюдаемое количество общих TCRβ CDR3 клонотипов каждой пары неродственных доноров. Пунктирной прямой показана степенная зависимость ($Y = aX^b$), которая использовалась для экстраполяции величины перекрытия индивидуальных репертуаров TCRβ при анализе полных индивидуальных репертуаров.

(Г) На данном графике представлено распределение количества общих клонотипов для 9 анализируемых людей. То есть показано, сколько общих клонотипов встречается одновременно у разного количества доноров. Из графика видно, что частота, с которой TCRβ клонотипы встречаются в репертуарах человека, распределена по степенному закону.

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 количественной информации о реальном репертуаре.

В нашей работе мы разработали метод для решения этой проблемы. Наш метод объединил в себе принципы разработанной нами ранее технологии подготовки библиотек генов TCR и технологии введения молекулярных баркодов (Рисунок 5). Молекулярный баркод – это последовательность из случайных нуклеотидов, уникальная для каждой молекулы кДНК.

В методе используется эффект смены матрицы ревертазой. В синтезируемую в процессе обратной транскрипции кДНК происходит встраивание 5'-адаптера, содержащего участок из 12 случайных нуклеотидов (молекулярный баркод, Рисунок 5). Таким образом, каждая синтезированная молекула кДНК оказывается специфически мечена одним из 4^{12} (>16.7 миллионов) вариантов уникальных молекулярных баркодов.

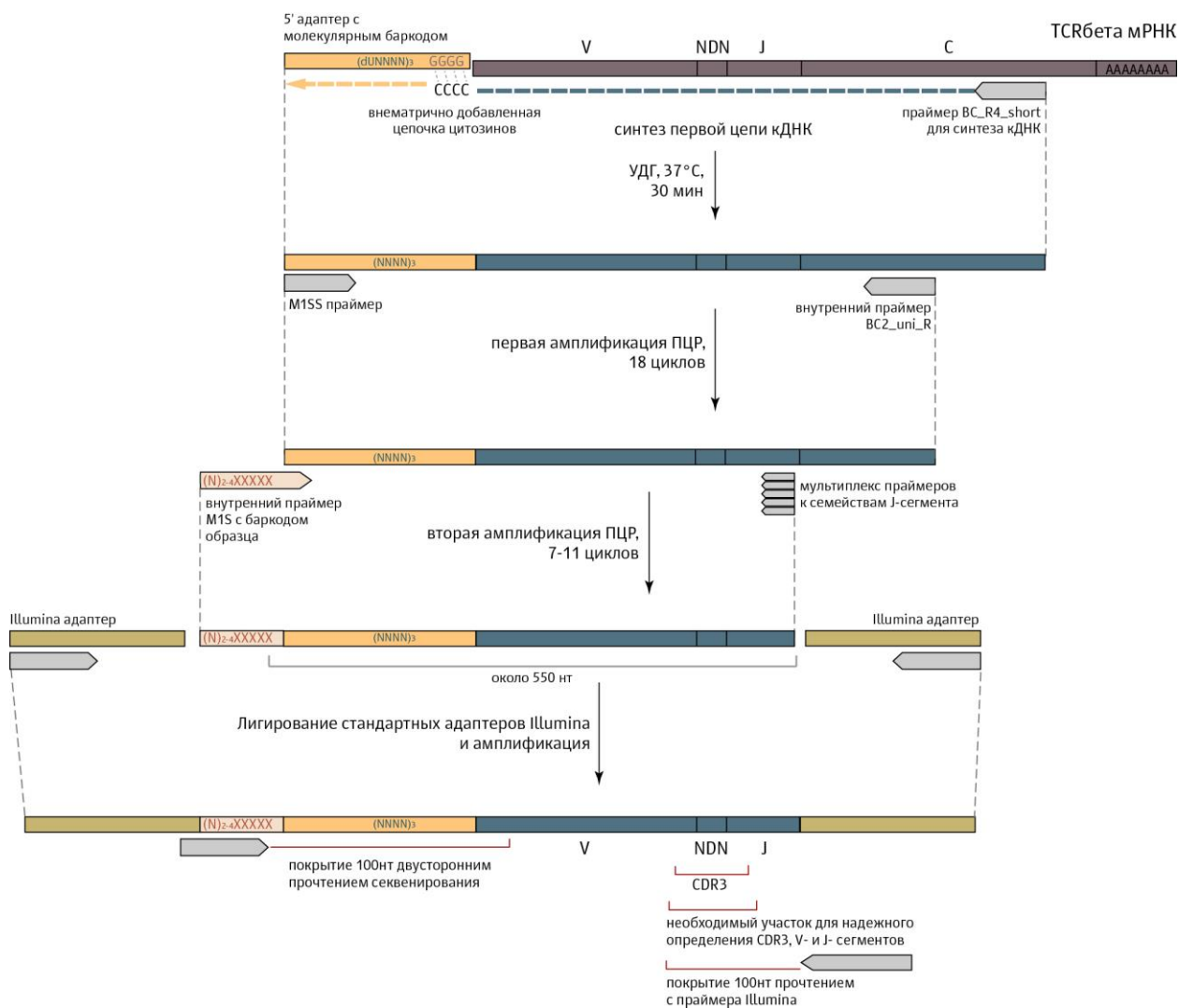


Рисунок 5. Схема подготовки библиотеки, созданной на основе мРНК, для массированного секвенирования.

(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.

При дальнейшем анализе данных массированного секвенирования библиотеки такие молекулярные баркоды позволяют провести точную оценку стартового количества молекул кДНК в образце, а также значительно 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 при фиксированном количестве анализируемых прочтенных последовательностей (Рисунок 6). Кроме того, применение данного метода нормализации позволяет снизить количественные искажения представленности вариантов внутри каждого конкретного массива.

Разработанный метод мы применили в следующей части работы для нормированного сравнительного анализа разнообразия репертуаров TCR β в периферической крови здоровых доноров различного возраста.

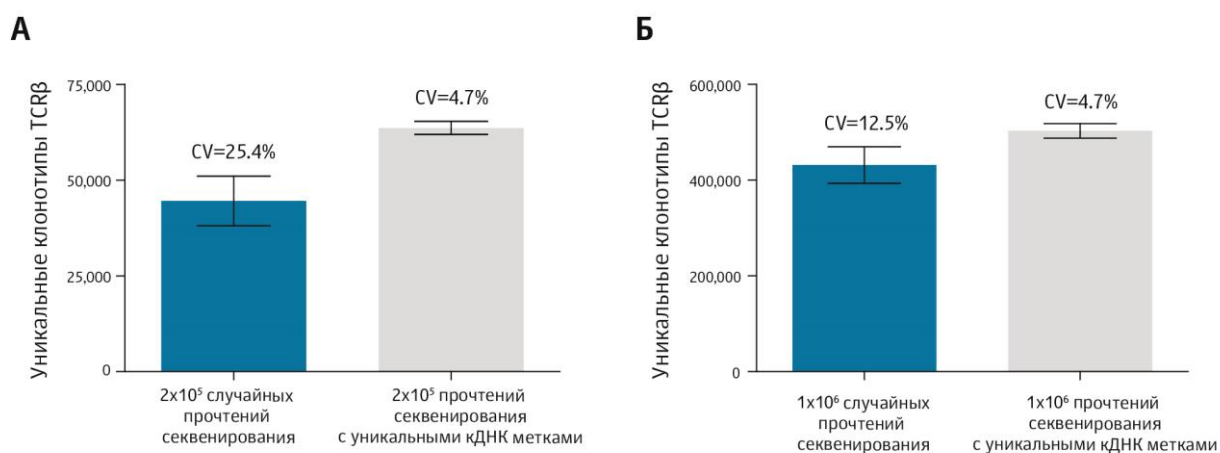


Рисунок 6. Нормализация с использованием молекулярного баркодирования увеличивает разнообразие TCR β внутри образца и уменьшает разброс между образцами. CV – коэффициент вариации.

(А) Библиотека TCR β , полученная из трех независимо полученных образцов крови одного и того же донора объемом 2 мл. Левый столбец соответствует количеству уникальных TCR β 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) Библиотека TCR β , полученная из двух 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 TCR β клонотипов, обнаруженных среди 1×10^6 случайно выбранных, CDR3-содержащих прочтений (без учета молекулярных баркодов). Правый столбец соответствует аналогичному количеству прочтений с уникальными молекулярными баркодами.

3. Исследование изменений репертуара TCR β в ходе старения человека

В нашем исследовании влияния старения на индивидуальный репертуар TCR β человека принимало участие 39 системно здоровых донора, разделенных на 4 возрастные группы:

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).

3.1. Наблюдаемое индивидуальное разнообразие TCR β CDR3 падает в течение всей жизни человека

Мы показали практически линейное убывание индивидуального разнообразия TCR β (наблюдаемого на 1 миллион взятых Т лимфоцитов) с возрастом (Рисунок 7) – начиная от примерно 600 000 клонотипов в детском возрасте, с кинетикой снижения разнообразия примерно 5 000 клонотипов в год. Наблюдается статистически значимое уменьшение числа уникальных клонотипов 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, наблюдается небольшое *увеличение* разнообразия TCR β , которое, впрочем, не является статистически достоверным.

3.2. Определение нижней границы общего индивидуального разнообразия TCR β CDR3

Для оценки нижней границы общего индивидуального разнообразия TCR β в периферической крови изучаемых доноров, нами была применена модель ненаблюдаемых видов⁵. Были получены следующие – последовательно снижающиеся – значения оценочной нижней границы общего разнообразия TCR β для разных возрастных групп:

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

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

⁵ Модель, позволяющая на основании доступной части видов оценить их полное разнообразие.

Группа 3 – $\sim 2.4 \times 10^6$ клонотипов.

Нижняя оценочная граница для доноров Группы 4 оказалась чуть выше, чем у доноров Группы 3, и составила $\sim 2.6 \times 10^6$ клонотипов.

Следует отметить, что мы получили лишь нижнюю оценку разнообразия, которая, вероятно, будет возрастать при увеличении глубины анализа. Чтобы проверить это предположение, мы использовали разделенные на 2 реплики (на уровне периферических мононуклеарных клеток крови) образцы двух доноров возрастом 25 и 87 лет. Доля наивных Т-клеток (CD45RA^{high}/CD27^{high}) в образцах периферической крови этих доноров составляла 44% и 15%, соответственно.

В анализе использовалось 10^6 уникальных стартовых молекул кДНК каждой из реплик. Сначала мы оценивали детектируемое разнообразие и нижнюю границу разнообразия CDR3 TCR β клонотипов, основываясь на одной из двух реплик. Затем в анализ добавляли вторые реплики. Их добавление увеличивало напрямую детектируемое разнообразие клонотипов 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 одном миллионе Т-клеток, действительно, ниже реального разнообразия TCR β человека. Скорее всего, с дальнейшим увеличением глубины секвенирования эта оценка будет и далее увеличиваться. Однако, учитывая клональность выходящих из тимуса наивных Т-лимфоцитов, воспроизводимость наиболее частых вариантов рекомбинации, а также конечный размер общего пула Т-лимфоцитов человека (до 10^{12} клеток), общее индивидуальное разнообразие TCR β не должно превышать 10^8 - 10^9 вариантов.

3.3. Заполнение гомеостатического пространства активированными клонами Т-клеток

При сравнительном анализе репертуаров TCR β групп молодого, среднего и пожилого возраста (Группы 1, 2 и 3) прослеживается возрастная зависимость объема Т-клеточного гомеостатического пространства, занимаемого высоко-представленными клонами. Напротив, процентное содержание единично представленных клонотипов (то есть клонотипов, кДНК которых представлена 1 молекулой в анализируемом миллионе молекул кДНК образца), падает с возрастом, составляя в среднем 55% в Группе 1 и 24% – в Группе 3 (Рисунок 8).

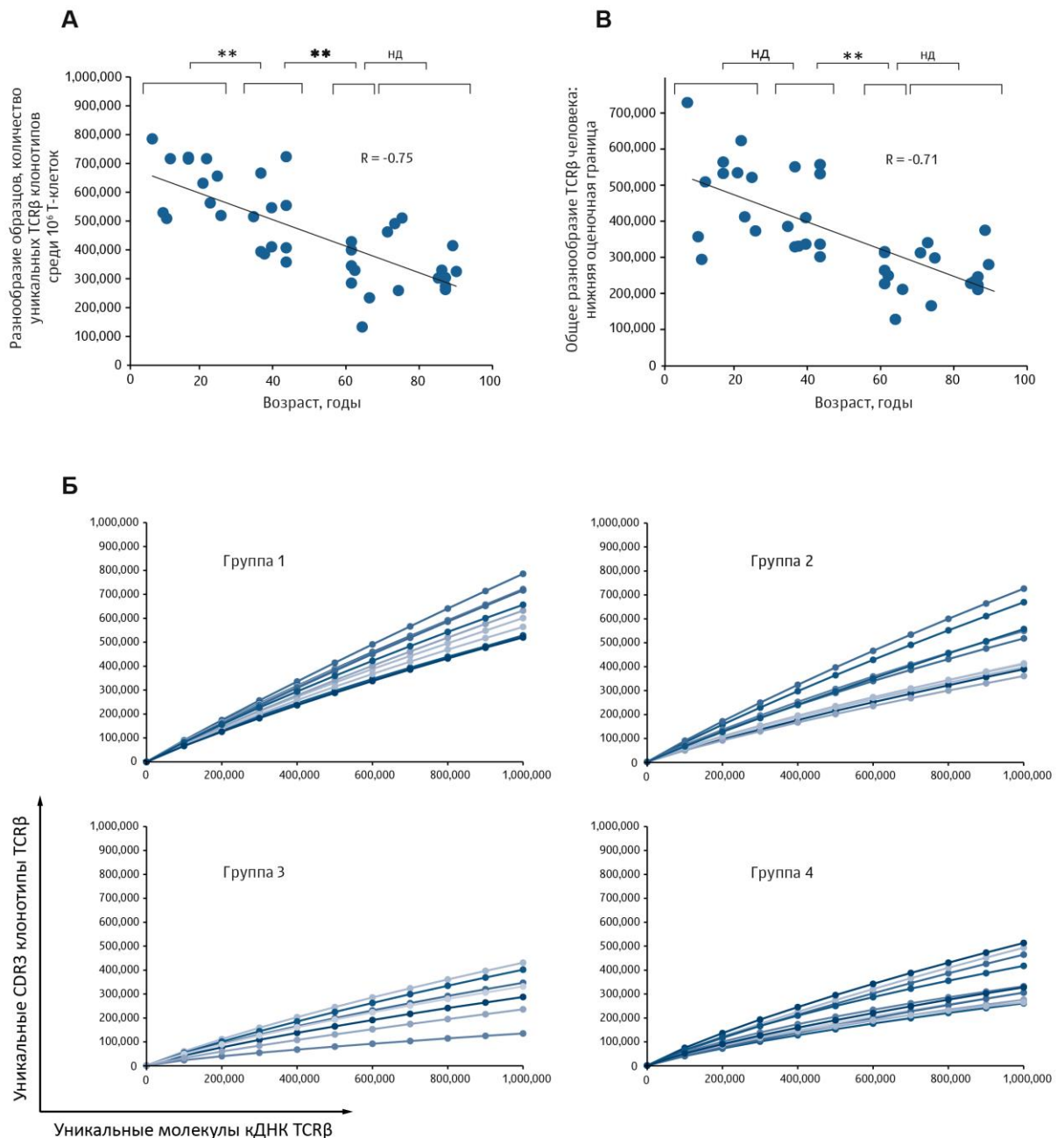


Рисунок 7. Возрастные изменения разнообразия репертуаров TCRβ. R – коэффициент корреляции.

(А) На графике изображена зависимость количества уникальных TCRβ CDR3 клонотипов на 10^6 T-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 анализ включены 10^6 молекул кДНК каждого образца. Графики показывают среднее количество уникальных TCRβ CDR3 клонотипов, полученных методом случайной выборки ($n=5$) молекул кДНК из образца каждого донора.

(В) Оценка нижней границы суммарного разнообразия TCRβ доноров в зависимости от возраста.

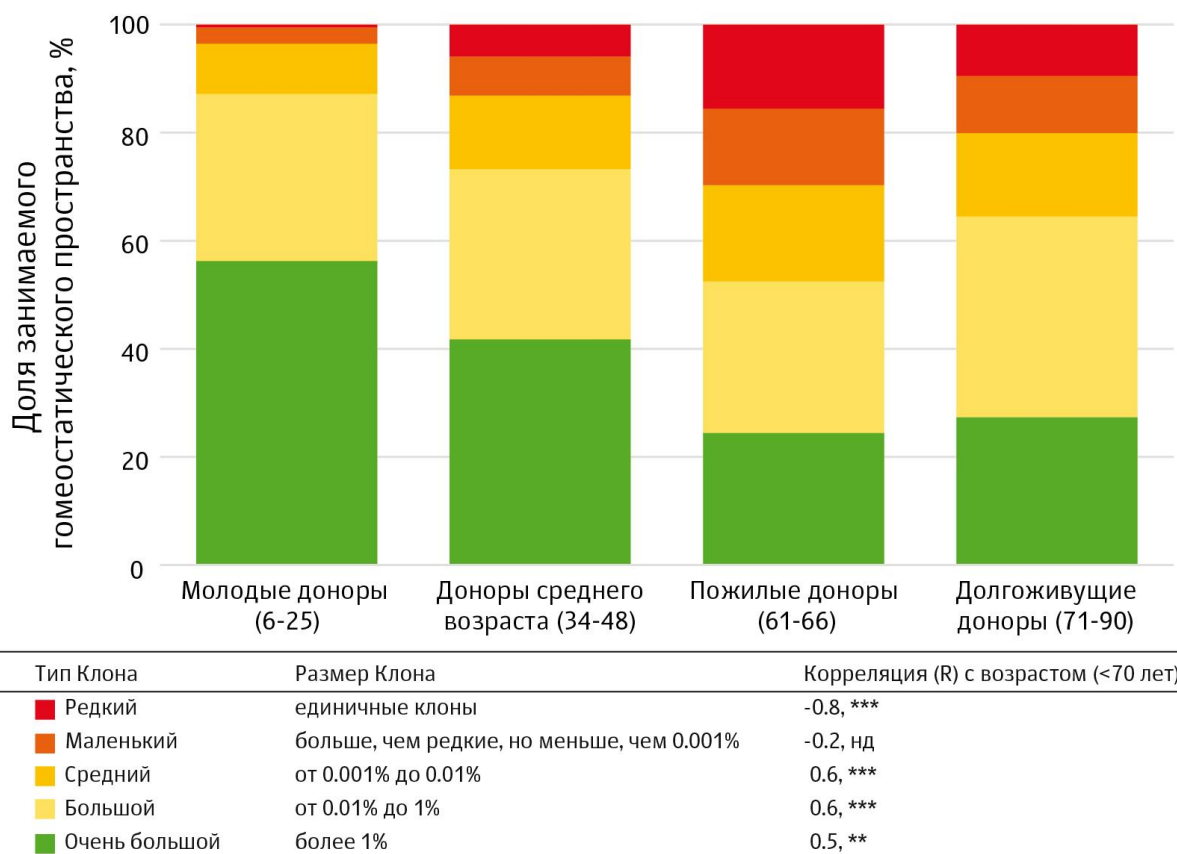


Рисунок 8. Распределение Т-клеточного гомеостатического пространства доноров разных возрастов между клонами разной представленности.

На графике показана доля гомеостатического пространства, занимаемого клонами разной представленности для четырех возрастных групп. Представленность клонов указана в процентах от общего количества анализируемых Т-клеток (10^6). *** – $P < 0.0001$, ** – $P = 0.001$, нд – недостоверно. R – коэффициент корреляции.

Учитывая относительно стабильное общее количество Т-лимфоцитов (практически не изменяющееся с возрастом), разрастание субпопуляции активированных Т-клеток неизбежно приводит к снижению количества наивных Т-клеток. Это приводит к снижению вероятности эффективного иммунного ответа против новых патогенов или патогенных клеток.

Интересно, что эти закономерности не распространяются на состоящую из относительно долгоживущих доноров Группу 4. В последней, напротив, наблюдается снижение объема, занимаемого высоко-представленными клонами и увеличение количества редко встречающихся клонов, по сравнению с Группой 3 (Рисунок 8).

3.4. Корреляция между разнообразием CDR3 TCR β и долей наивных Т-клеток в периферической крови

Чтобы оценить корреляцию между процентом наивных Т-клеток в периферической крови, возрастом донора и наблюдаемым разнообразием TCR β , мы измерили долю наивных Т-клеток (CD45RA^{high}/CD27^{high}) каждого из исследуемых образцов методом проточной цитофлуориметрии.

Как и предполагалось, была обнаружена обратная зависимость доли наивных Т-клеток в периферической крови донора от его возраста (Рисунок 9). Мы также показали наличие выраженной корреляции между долей наивных Т-клеток и наблюдаемым разнообразием TCR β (Рисунок 10).

Эта корреляция подтверждает как высокую точность методики молекулярного баркодирования для анализа репертуаров TCR β , так и верность определения CD45RA^{high}/CD27^{high} субпопуляции Т-лимфоцитов человека, как наивных клеток.

3.5. Образцы периферической крови долгоживущих доноров характеризуются высокой долей 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 ($R = -0.88$; $P < 1 \times 10^{-8}$; Рисунок 9). 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 процента 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. Более высокая доля наивных CD4⁺ Т-клеток сочетается с большим количеством низкочастотных клонотипов и сравнительно высоким разнообразием TCR β в Группе 4 (см. раздел 3.3).

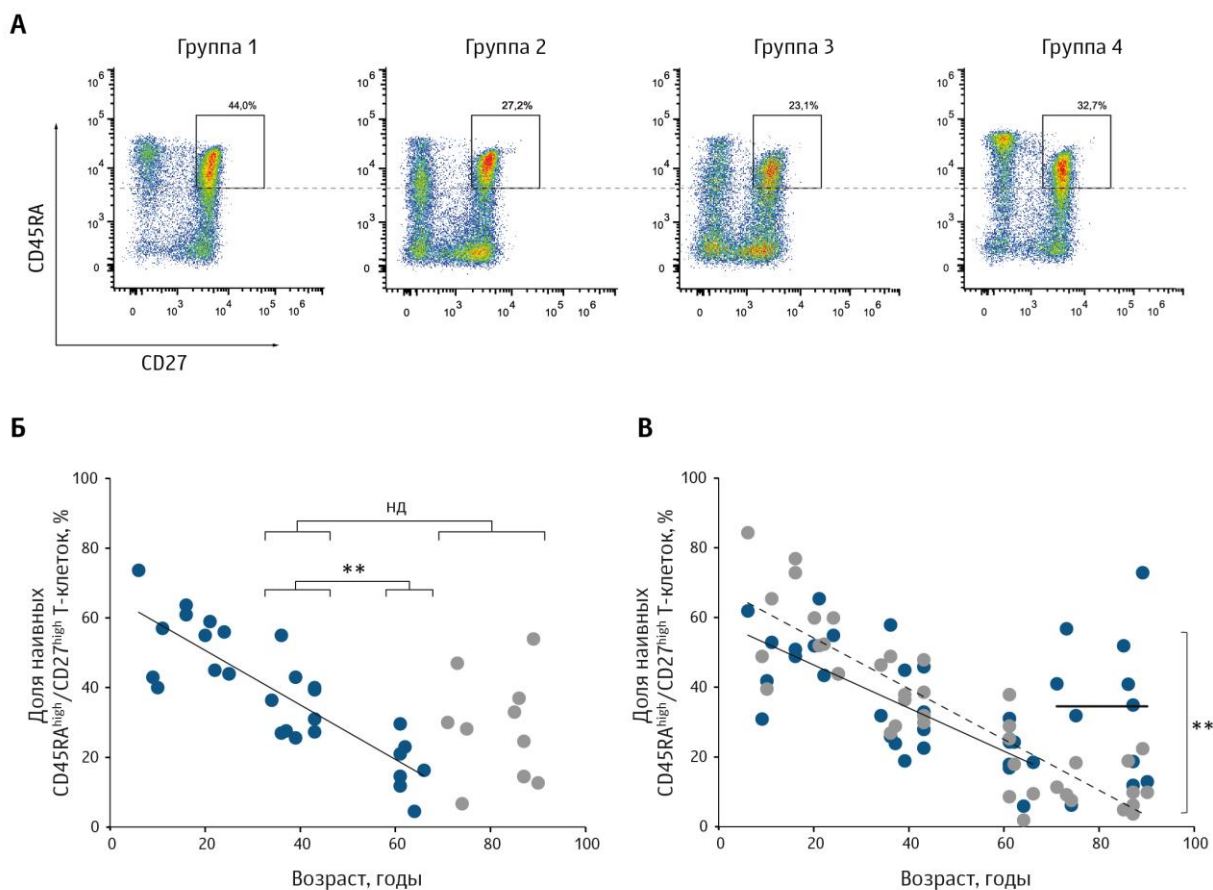


Рисунок 9. Возрастное изменение содержания наивных Т-клеток в периферической крови.

(А) Способ определения наивных Т-клеток методом проточной цитофлуориметрии (FACS). Периферические мононуклеарные клетки крови были проанализированы с помощью проточной цитофлуориметрии. 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 анализа доли наивных Т-клеток в пуле CD3⁺ клеток для Групп 1-4.

(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). Доля 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.

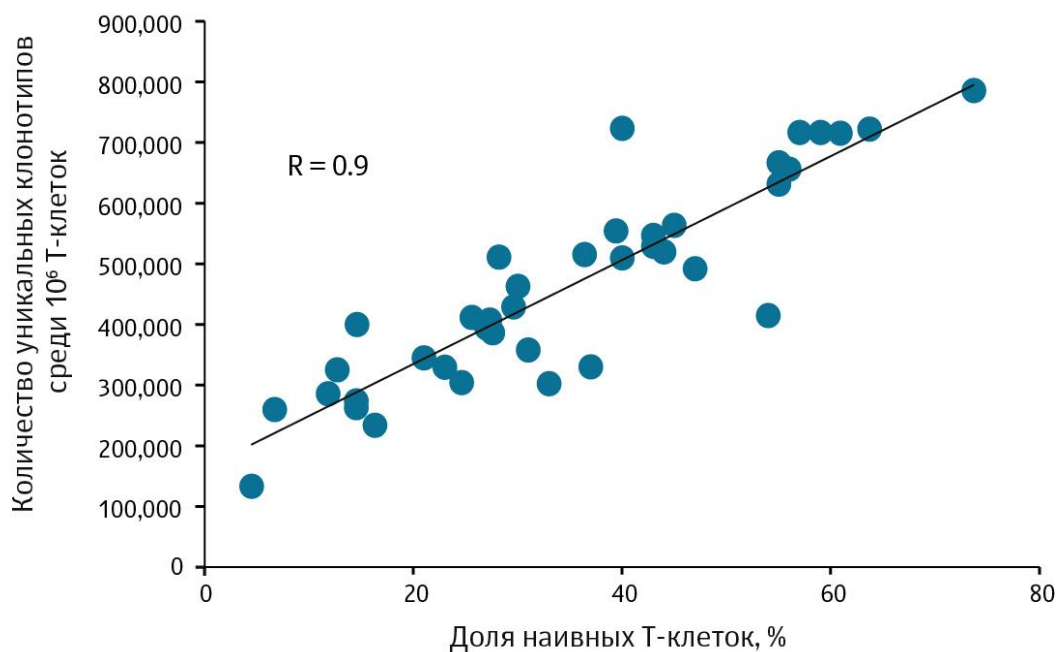


Рисунок 10. Наблюдаемое разнообразие TCRβ пропорционально доле наивных Т-клеток в периферической крови.

Показана зависимость разнообразия TCRβ, наблюдаемого в 10⁶ Т-клеток периферической крови, и доли наивных Т-клеток (CD45RA^{high}/CD27^{high}) в пуле CD3⁺ клеток ($R = 0.90$. $P = 1.4 \times 10^{-14}$). R – коэффициент корреляции.

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 повышенной доли 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.

3.6. **Общественные клонотипы⁶ и старение**

Сравнительный анализ репертуаров TCR β на 1 миллион Т лимфоцитов для каждого двух доноров оказал 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 клонотипов TCR β в популяции, был проведен сравнительный анализ репертуаров исследуемых доноров. В анализ вошли по 100 000 наиболее часто встречающихся клонотипов каждого из образцов. Из них были выбраны аминокислотные варианты CDR3, найденные как минимум в 6 из 39 доноров (т.е. в 15% всех доноров). Так, был получен список из примерно 10 000 самых часто встречающихся клонотипов в популяции.

Клонотипы, входящие в этот список, характеризуются короткой длиной CDR3 с медианой в 39 нт (в сравнение с медианой длин CDR3 всех проанализированных клонотипов из всех доноров, составляющей 45 нт). Кроме того, среднее количество добавленных к CDR3 букв также ниже у этой группы клонотипов.

Интересно, что количество таких наиболее распространенных общественных клонотипов в изучаемой популяции, наблюдаемых среди 100 000 самых представленных клонотипов, уменьшается от Группы 1 к Группе 3 и коррелирует с наблюдаемым разнообразием репертуара TCR β ($R = 0.48$, $P < 0.01$).

⁶ Клонотипы, встречающиеся в организмах многих индивидуумов.

Выводы

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

- в исходно генерируемом в ходе рекомбинации репертуаре TCR β относительная частота использования V β -сегментов сопоставима для родственных и не родственных 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 родственных пар.

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

- достоверных различий в пересечении репертуаров для родственных и неродственных 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 MHC1;
- 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 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 процентным 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.

Список работ, опубликованных по теме диссертации

Публикации

1. **Putintseva EV**, Britanova OV, Staroverov DB, Merzlyak EM, Turchaninova MA, Shugay M, Bolotin DA, Pogorelyy MV, Mamedov IZ, Bobrynina V, Maschan M, Lebedev YB, Chudakov DM. Mother and child T cell receptor repertoires: deep profiling study. *Front Immunol.* 2013; 4:463.
2. Britanova OV*, **Putintseva EV***, Shugay M*, Merzlyak EM, Turchaninova MA, Staroverov DB, Bolotin DA, Lukyanov S, Bogdanova EA, Mamedov IZ, Lebedev YB, Chudakov DM. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. *J Immunol.* 2014;192(6):2689-98.
3. Shugay M, Bolotin DA, **Putintseva EV**, Pogorelyy MV, Mamedov IZ, Chudakov DM. Huge Overlap of Individual TCR Beta Repertoires. *Front Immunol.* 2013;4:466.
4. Shugay M., Britanova O.V., Merzlyak E.M., Turchaninova M.A., Mamedov I.Z., Tuganbaev T.R., Bolotin D.A., Staroverov D.B., **Putintseva E.V.**, Plevova K., Linnemann C., Shagin D., Pospisilova S., Lukyanov S., Schumacher T.N., Chudakov D.M. Towards error-free profiling of immune repertoires. *Nat Methods.* 2014;11(6):653-5.
5. Bolotin DA, Shugay M, Mamedov IZ, **Putintseva EV**, Turchaninova MA, Zvyagin IV, Britanova OV, Chudakov DM. MiTCR: software for T-cell receptor sequencing data analysis. *Nat Methods.* 2013;10(9):813-4.
6. Mamedov IZ, Britanova OV, Zvyagin IV, Turchaninova MA, Bolotin DA, **Putintseva EV**, Lebedev YB, Chudakov DM. Preparing unbiased T-cell receptor and antibody cDNA libraries for the deep next generation sequencing profiling. *Front Immunol.* 2013;4:456.
7. Turchaninova MA, Britanova OV, Bolotin DA, Shugay M, **Putintseva EV**, Staroverov DB, Sharonov G, Shcherbo D, Zvyagin IV, Mamedov IZ, Linnemann C, Schumacher TN, Chudakov DM. Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol.* 2013;43(9):2507-15.

Тезисы докладов на конференциях

1. **Ekaterina Putintseva**, Dmitry Bolotin, Miklail Pogorely, Dmitry Staroverov, Mikhail Shugay, Olga Britanova, Dmitry Chudakov “Mother and child TCR repertoires: deep NGS-based profiling study”, летний симпозиум ZIBI (Центр инфекционной биологии и иммунологии), июнь 2013, Берлин, Германия.

* Вклад в работу данных авторов равнозначен.