


DOI 10.18699/vjgb-24-40

Аллель-специфичная ПЦР с флуоресцентно-мечеными зондами: критерии подбора праймеров для генотипирования

В.А. Девяткин, А.А. Шкляр, А.Ж. Фурсова , Ю.В. Румянцева, О.С. Кожевникова  

Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия
 oidopova@bionet.nsc.ru

Аннотация. Однонуклеотидные полиморфизмы (SNP) могут служить надежными маркерами в генной инженерии, селекции, скрининговых обследованиях и других областях науки, медицины и производства. Полногеномное секвенирование и генотипирование при помощи секвенирования могут высокоспецифично детектировать SNP и выявлять новые аллели. Однако в ситуациях, когда интерес исследователей направлен на отдельные конкретные локусы, эти методы становятся избыточными, а их цена, доля ложноположительных и ложноотрицательных результатов и трудозатраты на пробоподготовку и анализ не оправдывают их применения. Поэтому точные и быстрые методы генотипирования отдельных аллелей все еще остаются востребованными, особенно при проверке кандидатных полиморфизмов в анализах ассоциации с определенным фенотипом. Один из таких методов – генотипирование с использованием аллель-специфичных зондов TaqMan (TaqMan dual labeled probes). Метод заключается в реакции ПЦР в реальном времени с использованием пары праймеров и двух олигонуклеотидных зондов, комплементарных последовательности вблизи данного локуса таким образом, что один зонд комплементарен аллелю дикого типа, а другой – мутантному аллелю. Преимущество метода заключается в его специфичности, чувствительности, невысокой стоимости и скорости получения результатов. Он позволяет с высокой точностью различать аллели в геноме в одностадийной ПЦР без дополнительного этапа разделения продуктов реакции, что делает его востребованным в исследованиях генетических ассоциаций в молекулярной генетике и медицине. Благодаря развитию технологий синтеза олигонуклеотидов и совершенствованию методов подбора праймеров и зондов можно ожидать расширения возможностей применения этого подхода в диагностике наследственных заболеваний. В настоящей статье мы разобрали основные принципы метода, процессы, влияющие на результат генотипирования, критерии подбора оптимальных праймеров и зондов, использование LNA-модификаций в олигонуклеотидах, а также привели протокол подбора праймеров, зондов и ПЦР на примере SNP rs11121704. Мы надеемся, что представленный протокол позволит исследовательским группам самостоятельно подбирать собственные эффективные тест-системы для проверки интересующих полиморфизмов.

Ключевые слова: генотипирование; однонуклеотидные полиморфизмы; зонды TaqMan; LNA-модификации; аллель-специфичная ПЦР.

Для цитирования: Девяткин В.А., Шкляр А.А., Фурсова А.Ж., Румянцева Ю.В., Кожевникова О.С. Аллель-специфичная ПЦР с флуоресцентно-мечеными зондами: критерии подбора праймеров для генотипирования. *Вавиловский журнал генетики и селекции*. 2024;28(3):351-359. DOI 10.18699/vjgb-24-40

Финансирование. Работа выполнена при финансовой поддержке Российского научного фонда, проект № 21-15-00047.

Allele-specific PCR with fluorescently labeled probes: criteria for selecting primers for genotyping

V.A. Devyatkin, A.A. Shklyar, A.Zh. Fursova , Yu.V. Rumyantseva, O.S. Kozhevnikova  

Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
 oidopova@bionet.nsc.ru

Abstract. Single-nucleotide polymorphisms (SNPs) can serve as reliable markers in genetic engineering, selection, screening examinations, and other fields of science, medicine, and manufacturing. Whole-genome sequencing and genotyping by sequencing can detect SNPs with high specificity and identify novel variants. Nonetheless, in situations where the interest of researchers is individual specific loci, these methods become redundant, and their cost, the proportion of false positive and false negative results, and labor costs for sample preparation and analysis do not justify their use. Accordingly, accurate and rapid methods for genotyping individual alleles are still in demand, especially for verification of candidate polymorphisms in analyses of association with a given phenotype. One of these techniques is genotyping using TaqMan allele-specific probes (TaqMan dual labeled probes). The method consists of real-time PCR with a pair of primers and two oligonucleotide probes that are complementary to a sequence near a given locus

in such a way that one probe is complementary to the wild-type allele, and the other to a mutant one. Advantages of this approach are its specificity, sensitivity, low cost, and quick results. It makes it possible to distinguish alleles in a genome with high accuracy without additional manipulations with DNA samples or PCR products; hence the popularity of this method in genetic association studies in molecular genetics and medicine. Due to advancements in technologies for the synthesis of oligonucleotides and improvements in techniques for designing primers and probes, we can expect expansion of the possibilities of this approach in terms of the diagnosis of hereditary diseases. In this article, we discuss in detail basic principles of the method, the processes that influence the result of genotyping, criteria for selecting optimal primers and probes, and the use of locked nucleic acid modifications in oligonucleotides as well as provide a protocol for the selection of primers and probes and for PCR by means of rs11121704 as an example. We hope that the presented protocol will allow research groups to independently design their own effective assays for testing for polymorphisms of interest.

Key words: genotyping; single-nucleotide polymorphisms; TaqMan probes; LNA modifications; allele-specific PCR.

For citation: Devyatkin V.A., Shklyar A.A., Fursova A.Zh., Romyantseva Yu.V., Kozhevnikova O.S. Allele-specific PCR with fluorescently labeled probes: criteria for selecting primers for genotyping. *Vavilovskii Zhurnal Genetiki i Selektii = Vavilov Journal of Genetics and Breeding*. 2024;28(3):351-359. DOI 10.18699/vjgb-24-40

Введение

Однонуклеотидные полиморфизмы (SNP) активно используются как надежные маркеры в геномной инженерии, селекции, скрининговых обследованиях и других областях науки, медицины и производства. Очевидно, что полногеномное секвенирование и генотипирование при помощи секвенирования могут высокоспецифично детектировать SNP и выявлять новые варианты. Однако в ситуациях, когда интерес исследователей направлен на отдельные конкретные локусы, эти методы становятся избыточными, а их цена, доля ложноположительных и ложноотрицательных результатов и трудозатраты на пробоподготовку и анализ не оправдывают их применения. Поэтому точные и быстрые методы генотипирования отдельных аллелей все еще остаются востребованными, особенно при проверке кандидатных полиморфизмов в анализах ассоциации с определенным фенотипом (Kalendar et al., 2022).

На сегодняшний день методы, основанные на аллель-специфичной ПЦР, позволяющие получать наиболее точные результаты при их небольшой стоимости, при этом не требуют высокой квалификации персонала и дорогого лабораторного оборудования. В настоящей работе мы анализируем принципы работы с одним из таких методов – генотипированием с использованием аллель-специфичных зондов, основанным на методе TaqMan (TaqMan dual labeled probes). Впервые он был описан 15 лет назад (Hui et al., 2008) и по-прежнему остается одним из наиболее часто используемых для детекции однонуклеотидных полиморфизмов. Точность метода обеспечивает ошибку генотипирования менее чем в одном случае из 2000 (Ranade et al., 2001). Правильный подбор праймеров и зондов возможен для большинства последовательностей генома и позволяет более чем в 90 % случаев получать достаточно точные результаты генотипирования для ДНК высокого качества без дальнейшей оптимизации.

Метод аллель-специфичной ПЦР с TaqMan-зондами способен разделить генотипы даже при небольших количествах исходного образца, не требует пост-ПЦР-обработки и хорошо коррелирует с другими методами (Broscanello et al., 2018). Однако коммерчески доступные и разрабатываемые под конкретный SNP наборы не отличаются выгодной ценой, а рекомендации по созданию и оптимизации собственных тест-систем в большинстве литературных источников описаны поверхностно.

В данной работе мы приводим подробный анализ метода с описанием процессов, способных влиять на результаты генотипирования, а также рекомендации по подбору собственных тест-систем.

Описание метода

Метод заключается в реакции ПЦР в реальном времени с использованием пары праймеров (прямого и обратного, между которыми расположен интересующий нас полиморфный локус) и двух олигонуклеотидных зондов, комплементарных последовательности вблизи данного локуса таким образом, что один зонд комплементарен аллелю дикого типа, а другой – мутантному аллелю. На 5'-конце каждый зонд имеет свой флуоресцентный краситель, на 3'-конце – гаситель флуоресценции и фосфатную группу. Фосфатная группа не позволяет зондам служить праймерами в реакции ПЦР. У интактного зонда из-за близости гасителя и красителя сигнала нет, так как происходит Фёрстеровский перенос энергии (Förster resonance energy transfer, FRET) и тушение флуоресценции. На стадии элонгации Taq ДНК-полимераза, дошедшая до полностью комплементарного спаренного с матрицей зонда, за счет 5'-3'-экзонуклеазной активности гидролизует его, разобщая гаситель и краситель, флуоресценцию детектирует прибор.

Гибридизация зонда с матрицей более эффективна при полной комплементарности, к тому же в случае неспаренного основания (мисматча), когда с матрицей, содержащей один аллель, связывается зонд, соответствующий другому аллелю, полимеразы преимущественно вытесняет его полностью, не разобщая хромофоры. Поэтому накопление сигнала будет гораздо эффективнее происходить в случае полной комплементарности зонда и матрицы. Таким образом, соотношение уровня флуоресценции разных красителей зависит от соотношения в исходной матрице аллелей, соответствующих зондам, меченым этими красителями (Hui et al., 2008).

В приближенной к идеальной ситуации график аллельной дискриминации, где значения осей X и Y обозначают уровни флуоресценции первого и второго красителя для каждого образца, представлен на рис. 1. Образцы с одинаковым генотипом формируют облако точек, удаленных от других кластеров. Уровень флуоресценции каждого красителя равен нулю для гомозигот, не имеющих алле-

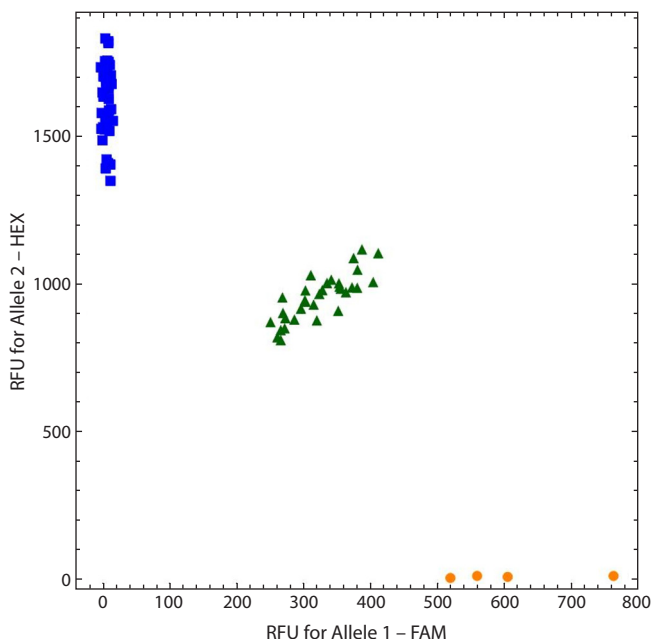


Рис. 1. График качественной аллельной дискриминации.

Здесь и на рис. 2: построено с помощью программы Bio-Rad CFX Manager. RFU (relative fluorescence unit) – относительная единица флуоресценции.

ля, меченного данным красителем, и практически вдвое выше для гомозигот по этому аллелю по сравнению с гетерозиготами.

На рис. 2 показан результат менее специфичной реакции, когда отжиг и последующая рестрикция зондов происходят также на матрице, соответствующей другому аллелю. В этом случае интенсивность свечения обоих красителей отлична от нуля для всех образцов. При этом генотипы образцов также хорошо дискриминируются с высокой точностью. Следует отметить, что в ходе реакции концентрация специфичного зонда снижается, концентрация матрицы увеличивается, а зонд с мисматчем остается интактным, что смещает равновесие в сторону образования дуплекса с мисматчем. В таком случае может помочь использование в реакции меньшего количества исходной матрицы либо считывание аллелограммы на более ранних циклах.

Исходя только из соотношения уровней сигналов красителей в отдельно взятом образце нельзя определить, какой зонд лучше связывается и, соответственно, каким был генотип, поскольку красители имеют разную интенсивность свечения, зонды отличаются по эффективности связывания, и на данный процесс могут влиять другие стохастические факторы. Поэтому для анализа необходимо использовать несколько образцов, среди которых встречаются разные генотипы. Рекомендуется анализировать не менее 20 образцов за один запуск прибора для достоверной дискриминации. По тому, насколько один сигнал нарастает интенсивнее, чем другой (т.е. по углу радиус-вектора точки на графике для каждого образца), можно разделить все образцы на группы, и в случае обнаружения трех групп быть уверенными в точном определении генотипа каждой из них. При отсутствии какого-то генотипа в выборке при оценке результатов следует опираться на

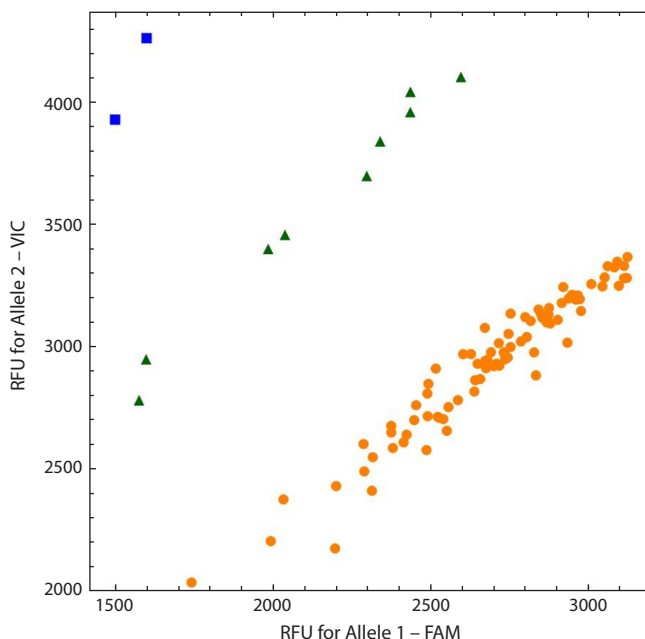


Рис. 2. График аллельной дискриминации с низким отношением целевой флуоресценции к фоновому сигналу.

соответствие реальной частоты генотипов и ожидаемой, по данным литературы или по равновесию Харди–Вайнберга. Чтобы убедиться в правильности аллельной дискриминации, рекомендуется проверить секвенированием по Сэнгеру образцы каждого генотипа.

Критерии подбора оптимальных праймеров и зондов для генотипирования

Подбор праймеров

- GC-состав в пределах 30–80 % (в идеале 40–60 %).
- Следует избегать повторов одинаковых нуклеотидов, особенно четырех и больше подряд идущих G.
- Температура плавления (T_m) в диапазоне 58–60 °C. Разница T_m между прямым и обратным праймером не больше 2 °C.
- Среди пяти нуклеотидов на 3'-конце не рекомендуется больше двух G и/или C. Следует избегать T на 3'-конце; G или C в последней позиции на 3'-конце является более подходящим сайтом связывания для ДНК полимераз.
- Длина 18–30 нуклеотидов.
- Праймеры и зонды не должны перекрываться.
- Рекомендуемая длина ампликона, по данным литературы, 80–120 нуклеотидов.

Увеличение длины продукта снижает эффективность реакции и нуклеазную активность Taq-полимеразы (Debo-de et al., 2017). В отдельных случаях, в зависимости от нуклеотидного состава, для выполнения остальных критериев можно довести длину ампликона до 1000. Минимальная длина обуславливается суммарной длиной праймеров и зондов. Расстояние между зондами и праймером, отжигающимися на ту же цепь, рекомендуют делать как можно меньше, по возможности не больше 20 нуклеотидов. Как показывает наш опыт, длина ампликона от 98 до 469 п. н.

и расстояние от зонда до праймера, комплементарного той же цепи, что и зонд, от 22 до 348 п. н. не оказывали видимого эффекта на точность дискриминации.

С точки зрения простоты и дешевизны эксперимента мы рекомендуем подбирать праймеры с учетом возможности секвенирования по Сэнгеру с этих же праймеров. В таком случае желательно, чтобы от праймера, с которого будет осуществляться чтение, до полиморфного сайта и от полиморфного сайта до максимальной длины прочтения (т. е. последовательности слева и справа от полиморфного сайта) было не менее 50 нуклеотидов.

Подбор зондов

- Оба зонда должны отжигаться на одну и ту же цепь и не быть комплементарными между собой.
- T_m должна быть примерно на 6–8 °C выше, чем у праймеров (в отличие от количественного ПЦР с использованием только одного зонда, T_m для которого должна быть выше на 10 °C, чем у праймеров). Это связано с тем, что по мере остывания реакционной смеси олигонуклеотидные зонды должны связаться с ДНК матрицей раньше праймеров.
- Чем больше разница в T_m для полностью спаренного с матрицей зонда и для зонда с неспаренным основанием, тем эффективнее разделяются аллели. Минимальная разница, позволявшая нам разделять аллели, была 3 °C, но мы рекомендуем по возможности подбирать зонды так, чтобы было минимум 4–5 °C. Чем больше окно температур, тем проще подобрать температуру отжига, при которой вероятность отжига зонда с мисматчем ничтожно мала по сравнению с таковой для полностью комплементарного зонда.
- Не помещать G на 5'-конце, так как он будет тушить присоединенный к нему флуорофор после разрезания зонда. Taq-полимераза также хуже разрезает такие зонды (Huang, Li, 2009). Разрезание олигонуклеотида начинается с появления 1–2 неспаренных нуклеотидов на 5'-конце, узнаваемых нуклеазным доменом. Неспаренный G на конце резко нарушает комплементарность, иногда приводя к полному разделению цепей быстрее, чем Taq-полимераза начнет разрезать зонд, и, таким образом, зонд вытесняется целиком, как если бы изначально был не полностью комплементарен.
- Из двух цепей выбирать такую, чтобы зонды содержали больше C, чем G, так как эмпирические данные показывают, что такие зонды чаще дают высокий сигнал (<https://www.thermofisher.com/order/catalog/product/450025>).
- Полиморфный сайт должен находиться примерно в средней трети зонда.
- GC состав в пределах 20–80 % (в идеале 30–70 %).
- Желательно подобрать позиции начала и конца каждого зонда так, чтобы T_m у обоих зондов стала примерно одинакова.
- Длина зондов 18–30 нуклеотидов, оптимальная длина составляет 20 нуклеотидов. Эти ограничения обусловлены тем, что зонд должен связываться специфично только с одним участком внутри амплифицируемого фрагмента и удовлетворять требованиям по T_m . Чем длиннее весь зонд целиком, тем меньший вклад в тем-

пературу плавления составляет полиморфный сайт, тем меньше в процентном соотношении будет различаться T_m для случаев полной комплементарности и мисматча, тем хуже дискриминируются аллели. Длина больше 30 нуклеотидов допустима, но в таких случаях гаситель следует располагать не на 3'-конце, а внутри зонда, примерно на расстоянии 18–25 нуклеотидов от 5'-конца. Это связано с тем, что при расстояниях между красителем и гасителем больше, чем 100 Å, соответствующих примерно 30 п. о. в структуре В-ДНК, нарушается FRET, и интактные пробы могут давать свечение, ухудшая отношение уровня сигнала к фоновому уровню.

- Репортерные флуорофоры (красители) должны иметь разные спектры эмиссии. Следует выбирать флуорофоры для разных каналов из списка совместимых с амплификатором. Пробы, меченные FAM и HEX, разрезаются более эффективно, чем пробы, меченные ROX или CY5.
- Рекомендуется метить более ярким красителем зонд с меньшими значениями T_m /GC-состава или с хуже связывающимся с матрицей аллелем А/Т (например, FAM дает более сильный сигнал, чем HEX).

Использование LNA-модификаций в олигонуклеотидах

Коммерчески доступные зонды TaqMan могут быть конъюгированы с мотивом, связывающим малую бороздку (minor groove binder, MGB), дигидроциклопирролиндо-трипептидом (DPI3), чтобы повысить аффинность связывания зонда с целевой последовательностью. Это позволяет увеличить температуру плавления зонда без увеличения его длины, что улучшает дискриминацию между комплементарным и некомплементарным зондом.

Коммерчески доступная альтернатива этой технологии – замкнутые нуклеиновые кислоты (locked nucleic acid, LNA или bridged nucleic acid, BNA), являющиеся аналогами РНК с рибозой, заблокированной в 3'-эндоконформации за счет 2'-O, 4'-C метиленового мостика. Включение данных модифицированных оснований в олигонуклеотид усиливает термостабильность и специфичность гибридизации (Owczarzy et al., 2011). Обозначаются такие нуклеотиды обычно как [+X] или +X (где X = A, T, G, C). За счет замены отдельных нуклеотидов в зонде на их LNA-аналоги можно сделать сам зонд короче, а вклад, который вносит полиморфный сайт в общую T_m , – больше, что облегчит разделение аллелей. Как правило, делают модификацию одного нуклеотида в позиции SNP, но для каждой последовательности различные варианты могут давать разный вклад (You et al., 2006).

Выбор инструментов для подсчета T_m олигонуклеотидов с LNA меньше, чем в случае с немодифицированными основаниями; можно использовать сервисы Oligo Evaluator (<http://www.oligoevaluator.com/LoginServlet>) или OligoAnalyzer Tool (<https://www.itdna.com/calc/analyzer>).

К сожалению, доступных сервисов для расчета разницы T_m между полностью спаренным дуплексом и гетеродуплексом с неспаренным основанием для олигонуклеотидов с LNA на данный момент нет. Ранее проводить такой расчет позволял OligoAnalyzer Tool, но из-за низкой точности эту опцию убрали. Для грубой оценки влияния LNA на мисматчи можно использовать результаты работ

по изучению термодинамики олигонуклеотидов с LNA. В некоторых случаях LNA модификация даже уменьшает разницу с мисматчем по сравнению с немодифицированным олигонуклеотидом, так что использовать эти модификации следует с осторожностью (You et al., 2006).

Пример подбора

Изменяя длину ампликона, температуру плавления, GC-состав, позиции и длину праймеров и зондов, можно получить комбинации, удовлетворяющие описанным критериям, и выбрать из всех лучшую. Не всегда анализируемые последовательности с SNP позволяют подобрать праймеры и зонды, удовлетворяющие всем описанным критериям, но это не означает, что подобранная тест-система не будет работать на практике.

Рассмотрим алгоритм подбора праймеров и зондов для анализа полиморфизма rs11121704.

1. В базе данных dbSNP (<https://www.ncbi.nlm.nih.gov/snp/>) находим полиморфизм, переходим на страницу с детальным описанием (<https://www.ncbi.nlm.nih.gov/snp/rs11121704>).

2. Находим нужную замену, обращаем внимание на сборку генома, для которой указана позиция. Например,

в нашем случае – последняя на текущий момент GRCh38. На следующем этапе понадобится информация о хромосоме (NC_000001.11), позиции на ней (11233902) и замене (C>T). Одному rsID может соответствовать несколько вариантов в одном локусе, все они перечислены на странице (C>A/C>T). Обычно у большей части популяции присутствует или референсный аллель, или наиболее распространенный из альтернативных. Нужную информацию можно найти во вкладке “Frequency” (рис. 3).

Используя данный метод генотипирования, рассматриваем все полиморфизмы как биаллельные, и нас обычно интересует наиболее распространенная в данной позиции замена, поскольку варианты с частотой, близкой к нулю, могут проявиться только на крупномасштабных выборках.

3. Для подбора праймеров используем открытый онлайн-ресурс Primer-Blast (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>).

В поле “Enter accession” вводим хромосому с интересующим полиморфизмом (NC_000001.11).

В полях “Range” указываем границы, в которых должны лежать прямой и обратный праймеры около позиции SNP (11233902). Устанавливаем границы участка по-

ALFA Allele Frequency

The ALFA project provide aggregate allele frequency from dbGaP. More information is available on the project [page](#) including descriptions, data access, and terms of use.

Release Version: 20230706150541

Search:

Population	Group	Sample Size	Ref Allele	Alt Allele
Total	Global	75060	C=0.28937	A=0.00000, T=0.71063
European	Sub	59104	C=0.27827	A=0.00000, T=0.72173
African	Sub	4888	C=0.6279	A=0.00000, T=0.3721
African Others	Sub	174	C=0.718	A=0.000, T=0.282
African American	Sub	4714	C=0.6245	A=0.00000, T=0.3755
Asian	Sub	238	C=0.046	A=0.000, T=0.954
East Asian	Sub	164	C=0.049	A=0.000, T=0.951
Other Asian	Sub	74	C=0.04	A=0.00, T=0.96
Latin American 1	Sub	400	C=0.305	A=0.000, T=0.695
Latin American 2	Sub	3384	C=0.1690	A=0.00000, T=0.8310
South Asian	Sub	4968	C=0.1842	A=0.00000, T=0.8158

Рис. 3. Основная информация о варианте rs11121704: хромосома, позиция, сборка генома, частота встречаемости нуклеотидных замен.

садки праймеров не ближе 15 нуклеотидов от позиции SNP, потому что иначе посадке праймера может мешать зонд, и не дальше 200, чтобы ампликон не был слишком длинным и эффективность реакции была выше: прямой праймер от 11233702 (11233902-200) до 11233887 (11233902-15), обратный – от 11233917 (11233902+15) до 11234102 (11233902+200). “PCR product size” устанавливаем 100–250.

В поле “Database Refseq” указываем “Refseq representative genomes” или “Genomes for selected organisms (primary reference assembly only)”. Обе базы данных содержат первичные сборки хромосомных последовательностей с минимумом избыточности, в “Representative genomes” также указаны альтернативные локусы и митохондриальные геномы при наличии таковых.

Пункт “Advanced parameters” дает доступ к дополнительным параметрам, из которых нас интересует параметр “Primer GC content (%)”, для него устанавливаем диапазон 40–60 %.

Остальные параметры оставляем без изменений. Перед нажатием кнопки “Get primers” выбираем опцию “Show results in a new window”, чтобы после выдачи результатов было проще изменить отдельные параметры запуска для повторного поиска (рис. 4).

4. Переходим на страницу с результатами поиска праймеров. Выбираем праймеры, подходящие по позиции, Tm и специфичности. С неспецифичными праймерами можно работать, однако необходимо потом проверить, что зонд связывается только со специфичным ампликоном, но не с побочными.

The image shows a screenshot of the Primer-Blast web interface. It is divided into several sections:

- PCR Template:** Includes fields for "Enter accession, gi, or FASTA sequence" (with NC_000001.11 entered), "Forward primer" (233702), "Reverse primer" (233917), and "Range" (From 233702 To 233887).
- Primer Parameters:** Includes "PCR product size" (100-250), "Primer GC content (%)" (40-60), "Primer melting temperatures (Tm)", and "Secondary Structure Alignment Methods".
- Exon/Intron selection:** Includes "Exon junction span" (No preference), "Exon junction match", and "Intron inclusion".
- Primer Pair Specificity Checking Parameters:** Includes "Specificity check" (checked), "Search mode" (Automatic), "Database" (Refseq representative genomes), "Organism" (Homo sapiens), and "Primer specificity stringency".
- Advanced parameters:** Includes "Primer Pair Specificity Checking Parameters" with values like "Max number of sequences returned by Blast" (50000), "Blast word size" (7), and "Max primer pairs to screen" (500).
- Internal hybridization oligo parameters:** Includes "Hybridization oligo" and "Hyb Oligo GC%".

Buttons for "Get Primers" and checkboxes for "Show results in a new window" and "Use new graphic view" are visible at the bottom of the main sections.

Рис. 4. Пример настроек для подбора праймеров в программе Primer-Blast.

Primer pair 1

	Sequence (5'->3')	Template strand
Forward primer	TTTTTCCTCATTTTGGGCGA	Plus
Reverse primer	TATCAGTTGCAGGAAAGTGC	Minus
Product length	130	

Products on intended targets
>NC_000001.11 Homo sapiens chromosome 1, GRCh38.p14 Primary Assembly

product length = 130
Features associated with this product:
serine/threonine-protein kinase mtor isoform 1
serine/threonine-protein kinase mtor isoform x3

Forward primer	1	TTTTTCCTCATTTTGGGCGA	20
Template	11233836	11233855
Reverse primer	1	TATCAGTTGCAGGAAAGTGC	20
Template	11233965	11233946

Products on potentially unintended templates
>NC_000013.11 Homo sapiens chromosome 13, GRCh38.p14 Primary Assembly

product length = 1186
Features associated with this product:
fibroblast growth factor 14 isoform 1a
fibroblast growth factor 14 isoform 3

Forward primer	1	TTTTTCCTCATTTTGGGCGA	20
Template	101877194	..A.G.....T..	101877213
Reverse primer	1	TATCAGTTGCAGGAAAGTGC	20
Template	101878379	CTGA....T.....	101878360

Рис. 5. Выбранные праймеры в программе Primer-Blast дают специфичный целевой ПЦР-продукт длиной 130 нуклеотидов, содержащий SNP rs11121704, и один потенциальный ПЦР-продукт длиной 1186 нуклеотидов, который не будет образовываться при нормальных условиях.

Мы выбрали для прямого праймера последовательность 5'-TTTTTCCTCATTTTGGGCGA-3', для обратного – 5'-TATCAGTTGCAGGAAAGTGC-3'. На странице с результатами показано, что выбранные праймеры образуют целевой специфичный продукт длиной 130 нуклеотидов (рис. 5). А также представлен один потенциальный неспецифичный ПЦР-продукт длиной 1186 нуклеотидов, который не будет образовываться вследствие неполной комплементарности сайтов посадки последовательностям праймеров (см. рис. 5).

5. Далее в пункте “Tracks” выбираем подпункт “Configure tracks”, находим и отмечаем галочками пункты “Common variations (MAF>=0.01)”, “Cited variations” и “ClinVar variants with precise endpoints”, добавляем их в отображение кнопкой “Configure”. Следует избегать пересечения зонда с другими полиморфизмами, кроме интересующего нас.

6. В результатах поиска выделяем область последовательности около SNP (-20...+20 нуклеотидов), нажимаем “Copy sequence (selection)” (рис. 6).

Строим также комплементарную этой последовательности цепь. Это можно сделать вручную или использовать любой доступный сервис, например (https://www.bioinformatics.org/sms/rev_comp.html).

Последовательность прямой цепи:

5'-TTCTCCTTTCCAAACATCTG(C)GATGATGTGCC TGAAGCAT-3'

Последовательность обратной цепи:

5'-AATGCTTCAGGCACATCATC(G)CAGATGTTTGG AAAGGAGAA-3'

В скобках обозначена позиция интересующего SNP.

Рис. 6. Последовательность нуклеотидов, фланкирующая SNP и не пересекающаяся с другими известными полиморфными сайтами.

Основные характеристики праймеров и зондов для rs11121704

Олигонуклеотид	Tm OligoEvaluator	Tm OligoCalc	Tm Oligo- Analyzer Tool	Tm среднее	Tm мисматч	delta Tm
5'-TTTTTCCTCATTTGGGCGA-3'	66.3	54.3	63.0	61.2		
5'-TATCAGTTGCAGGAAAGTGC-3'	60.5	56.4	62.4	59.8		
5'-FAM-CAGGCACATCATCGCAGATGTTT-BHQ1-3'	70.2	62.9	66.4	66.5	62.4	4.0
5'-VIC-CAGGCACATCATCACAGATGTTT-BHQ1-3'	68.6	63.6	64.9	65.7	60.3	4.6

Примечание. Перечислены температуры плавления (Tm), предсказанные несколькими сервисами, их среднее. Tm мисматч – температура плавления дуплекса зонда с некомплементарной матрицей. delta Tm – разница между Tm (Oligo-Analyzer Tool) и Tm мисматч.

7. Внутри последовательности около SNP выбираем подходящий по длине и составу фрагмент. Судя по GC-составу, нужно взять последовательность обратной цепи, потому что так в зонде будет больше C, чем G:

5'-CAGGCACATCATC(G)CAGATGTTT-3'

Поскольку мы берем не ту цепь, для которой указана SNP, следует помнить, что для нашей последовательности замена C>T в обратной цепи соответствует замене G>A.

Границы второго зонда подберем так, чтобы сравнить их Tm:

5'-CAGGCACATCATC(A)CAGATGTTT-3'

8. Рекомендуется проверить Tm по нескольким сервисам и усреднить (см. таблицу).

Для примера мы использовали OligoCalc (<http://biotools.nubic.northwestern.edu/OligoCalc.html>), OligoEvaluator (<http://www.oligoevaluator.com/LoginServlet>) и OligoAnalyzer Tool (<https://www.idtdna.com/calc/analyzer>).

9. Для сравнения температуры плавления полностью комплементарно связанного зонда с зондом, образующим неспаренное основание, в OligoAnalyzer Tool выбираем опцию “Tm mismatch”.

Для первого зонда CAGGCACATCATC(G)CAGATGTTT мисматчем будет нуклеотид, комплементарный второму зонду (CAGGCACATCATC(A)CAGATGTTT), т.е. выбираем букву “T”, нажимаем “Use Exact Complement Tm”, “Calculate”. Чем больше разница между Tm полностью комплементарного олигонуклеотида и зонда с некомплементарным основанием (“deltaTm”) и, соответственно, меньше доля связанного зонда с мисматчем по сравнению с полностью комплементарным зондом на стадии отжига зондов, тем точнее будет разделение аллелей.

10. Для проверки специфичности подобранных зондов используем сервис Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome).

В поле “Enter accession number(s), gi(s), or FASTA sequence(s)” вставляем последовательность зонда, “Database” – “Refseq representative genomes”, “Organism” – “human (taxid:9606)”, выбираем опцию “Show results in a new window”, нажимаем кнопку “Blast”. Нам важно, чтобы зонд не связывался с неспецифичными ПЦР-продуктами (если такие есть) и садился в единственный участок целевого ампликона.

11. Праймеры и зонды рекомендуется проверить на само- и взаимокompлементарность и образование шпиклек в программе OligoAnalyzer Tool согласно их рекомендациям (<https://www.idtdna.com/pages/education/decoded/article/designing-pcr-primers-and-probes>). В программе про-

веряем параметр ΔG (изменение свободной энергии Гиббса) образования вторичных структур. При более положительных значениях ΔG , чем -9 ккал/моль, вторичные структуры не оказывают значимого эффекта на ПЦР, значения же больше нуля показывают, что при данных условиях вторичные структуры не формируются (https://www.gene-quantification.de/oligo_architect_glossary.pdf). Поэтому при проверке праймеров и зондов выбираем такие, у которых для потенциальных вторичных структур параметр $\Delta G \geq -9$ ккал/моль.

Проведение и подбор условий ПЦР

Меченные флюорофором зонды следует хранить в темноте, чтобы избежать фотообесцвечивания (https://assets.thermofisher.com/TFS-Assets/LSG/Application-Notes/cms_043004.pdf).

Смесь для ПЦР готовим на льду, на одну реакцию требуется:

- 10 мкл буфера (мы использовали BioMaster HS-qPCR (2x) («Биолабмикс», Россия), но его можно заменить на любой доступный аналог),
- 3.5 пмоль прямого праймера,
- 3.5 пмоль обратного праймера,
- 1.5 пмоль зонда с FAM,
- 1.5 пмоль зонда с VIC,
- 10 нг ДНК,
- ddH₂O до 20 мкл.

В пробирке смешиваем все перечисленные компоненты, кроме образца ДНК, с учетом количества образцов и с запасом примерно 10 %. Образцы ДНК раскапываем сразу в лунки, затем добавляем в каждую лунку по 18 мкл смеси, перемешиваем, сбрасываем капли.

Сначала на нескольких образцах проверяем работоспособность новых праймеров и зондов, подбираем оптимальные температуры отжига. Оптимальные концентрации праймеров и зондов также могут отличаться от приведенных выше, но финальные концентрации зондов обычно минимум в 2 раза ниже, чем праймеров (https://www.bioline.com/mwdownloads/download/link/id/3301/p/i/pi-50201_sensifast_probe_hi-rox_one-step_kit_v11.pdf).

Протокол ПЦР

1. Начальная денатурация, 95 °C 3 мин.
2. Амплификация и детекция (40 циклов):
денатурация, 95 °C 10 с,
отжиг праймеров и элонгация с детекцией сигнала, 60 °C 30 с.

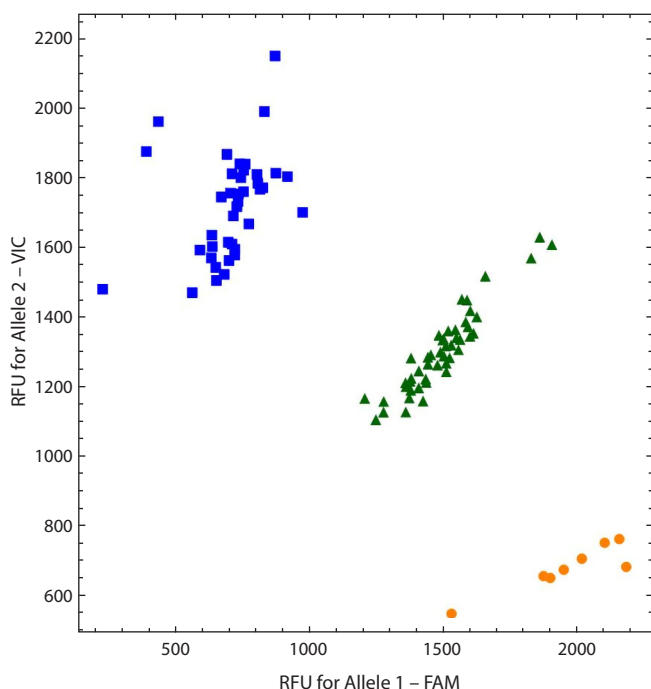


Рис. 7. Результат аллельной дискриминации с помощью зондов к SNP rs11121704.

Оранжевыми точками обозначены гомозиготы по референсному аллелю (C/C), зелеными треугольниками – гетерозиготы (C/T), синими квадратами – гомозиготы по альтернативному аллелю (T/T).

Результат ПЦР с подобранными праймерами и зондами к SNP rs11121704 представлен на рис. 7.

Поскольку T_m олигонуклеотида обозначает температуру, при которой половина молекул находится в расплавленном состоянии, а половина – в двуцепочечном, рекомендуемая температура отжига должна быть примерно на 5°C ниже наименьшего значения T_m для двух праймеров, так как при таких условиях оба праймера практически полностью свяжутся с комплементарными цепями. На практике из-за возможных неточностей при расчете T_m или несоответствия условий реакции условиям, для которых производился расчет, оптимальная температура подбирается эмпирически. Мы рекомендуем проверить интервал [$T_{m_{cp}} - 5^\circ\text{C} \dots T_{m_{cp}} + 5^\circ\text{C}$], где $T_{m_{cp}}$ – среднее значение T_m двух праймеров. Также рекомендуем подбор условий проводить для двух вариантов смеси: с зондами и с интеркалирующим красителем, например SYBR Green. График кривых плавления позволит выявить возможные неспецифичные продукты ПЦР.

Элонгация обычно занимает около 1 мин на 1000 п. о. Если температура отжига составляет больше 60°C , этот шаг часто объединяют с предыдущим, и элонгация происходит при температуре отжига. Хотя температурный

оптимум для большинства Taq-полимераз лежит в районе $75\text{--}80^\circ\text{C}$, элонгация не может протекать при температуре выше, чем температура плавления зондов.

Заключение

Генотипирование с помощью аллель-специфичной ПЦР является эффективным и точным инструментом для определения генетических вариаций. Преимущества данного метода заключаются в его специфичности, чувствительности, невысокой стоимости и скорости получения результатов. Он позволяет с высокой точностью различать разные аллели в геноме в одностадийной ПЦР без дополнительных этапов разделения продуктов реакции, что делает его особенно полезным для исследований генетических ассоциаций в молекулярной генетике и медицине.

Благодаря развитию технологий синтеза олигонуклеотидов и совершенствованию методов подбора праймеров и зондов можно ожидать расширения возможностей применения этого подхода в диагностике наследственных заболеваний. В настоящей статье мы детально разобрали критерии и условия оптимизации успешного подбора праймеров и олигонуклеотидных зондов для аллель-специфичной ПЦР. Мы надеемся, что представленный протокол позволит исследовательским группам самостоятельно подбирать собственные эффективные тест-системы для проверки интересующих их полиморфизмов.

Список литературы / References

- Broccanello C., Chiodi C., Funk A., McGrath J.M., Panella L., Stevanato P. Comparison of three PCR-based assays for SNP genotyping in plants. *Plant Methods*. 2018;14:28. DOI 10.1186/s13007-018-0295-6
- Debode F., Marien A., Janssen E., Bragard C., Berben G. The influence of amplicon length on real-time PCR results. *Biotechnol. Agron. Soc. Environ.* 2017;27(1):3-11. DOI 10.25518/1780-4507.13461
- Huang Q., Li Q. Characterization of the 5' to 3' nuclease activity of *Thermus aquaticus* DNA polymerase on fluorogenic double-stranded probes. *Mol. Cell. Probes*. 2009;23(3-4):188-194. DOI 10.1016/j.mcp.2009.04.002
- Hui L., DeMonte T., Ranade K. Genotyping using the TaqMan assay. *Curr. Protoc. Hum. Genet.* 2008;56(2):2.10.1-2.10.8. DOI 10.1002/0471142905.hg0210s56
- Kalendar R., Shustov A.V., Akhmetolayev I., Kairov U. Designing allele-specific competitive-extension PCR-based assays for high-throughput genotyping and gene characterization. *Front. Mol. Biosci.* 2022;9:773956. DOI 10.3389/fmolb.2022.773956
- Owczarzy R., You Y., Groth C.L., Tataurov A.V. Stability and mismatch discrimination of locked nucleic acid-DNA duplexes. *Biochemistry*. 2011;50(43):9352-9367. DOI 10.1021/bi200904e
- Ranade K., Chang M.S., Ting C.T., Pei D., Hsiao C.F., Olivier M., Pesich R., Hebert J., Chen Y.D., Dzaou V.J., Curb D., Olshen R., Risch N., Cox D.R., Botstein D. High-throughput genotyping with single nucleotide polymorphisms. *Genome Res*. 2001;11(7):1262-1268. DOI 10.1101/gr.157801
- You Y., Moreira B.G., Behlke M.A., Owczarzy R. Design of LNA probes that improve mismatch discrimination. *Nucleic Acids Res*. 2006;34(8):e60. DOI 10.1093/nar/gkl175

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 30.08.2023. После доработки 04.10.2023. Принята к публикации 12.10.2023.