Published in final edited form as:

Methods Mol Biol. 2012 ; 882: . doi:10.1007/978-1-61779-842-9_10.

Genotyping of Single Nucleotide Polymorphisms by 5' Nuclease Allelic Discrimination

Mari Malkki and Effie W. Petersdorf

Abstract

Real-time quantitative PCR is an efficient method for high-throughput genotyping of single nucleotide polymorphisms (SNPs). In this chapter, we describe the 5' nuclease allelic discrimination assay for genotyping biallelic SNPs.

Keywords

Single nucleotide polymorphism; Real-time polymerase chain reaction; TaqMan® chemistry; Reporter dye; Fluorescent label

1. Introduction

Single nucleotide polymorphisms (SNPs) are the most common form of human genetic variation (1–8). SNPs are biallelic and occur approximately every 1,000 base pairs (bp) throughout the human genome (1–8). SNPs are powerful markers for mapping genes that cause disease (9–21). Association of SNPs residing within the major histocompatibility complex (MHC) have recently been described for diseases with well-known associations to HLA phenotypes and haplotypes, including rheumatoid arthritis, diabetes, and multiple sclerosis (12–21).

Because of their biallelic nature, SNPs can be readily genotyped using techniques that discriminate any two-way combination of adenine, guanine, cytosine, and thymine nucleotide bases. Among the most robust assays for SNP genotyping, real-time quantitative assays rely on sensitive and specific quantification of amplified DNA or cDNA using polymerase chain reaction (PCR) (22–24). Real-time PCR methods offer a variety of different labels for the detection of genetic variation, including DNA-intercalating agents, such as SYBR® Green (25), fluorogenic probes (26–28), and the widely used TaqMan® chemistry (29, 30).

In this chapter, we describe the real-time 5'-nuclease genotyping assay known as TaqMan® for discriminating between two alleles of a specific SNP. The method may be used for genotyping individual SNPs that reside throughout the human genome. In brief, a wild-type SNP Allele "A" is amplified separately from the alternative Allele "B" using region specific forward and reverse primers and two allele-specific TaqMan® probes designed to target the polymorphism. TaqMan® probes have a fluorescent reporter dye (VIC® specific for allele "A" and 6-carboxyfluorescein [FAM™] specific for allele "B") attached to its 5' end and a quencher dye (minor groove binder [MGB] or 6-carboxy-tetramethyl-rhodamine [TAMRA]) at its 3' end. The amplification is performed using a thermal cycler or a real-time PCR system. During amplification each uniquely labeled probe binds preferentially to one of the two alleles of the SNP of interest with different affinity. As amplification proceeds, the Taq

polymerase enzyme cleaves the bound probe, and a fluorescent signal is generated. Fluorescent signals are interpreted automatically using sequence detection software dedicated to real-time PCR instrumentation. A fluorescent signal from only the VIC dye indicates homozygosity for Allele “A”; the presence of only FAM dye fluorescence indicates homozygosity for Allele “B,” and the presence of both fluorescent signals indicates Allele “A”/Allele “B” heterozygosity.

2. Materials

2.1. TaqMan Assay

Supplies needed for performing the TaqMan[®] assay are as follows:

1. 96-Well optical plates and adhesive optical covers (Applied Biosystems, [ABI]).
2. Splash-free support base for 96-well plates.
3. Plate centrifuge.
4. 10, 200 and 1,000- μ L sterile tips.
5. 1–10, 20–200 and 100–1,000- μ L single or multiple tip electronic pipets or 10, 20, 200 and 1,000- μ L manual pipets.
6. Eppendorf tubes.
7. TaqMan[®] Universal Master Mix.

2.2. Real-Time PCR Instrumentation

ABI 7500 Real-Time PCR system (see Note 1).

2.3. Design of Probes and Primers

Currently, over 160,000 inventoried, ready-to-use SNP assays and over 4.5 million predesigned, made-to-order SNPs assays are available from ABI (<https://products.appliedbiosystems.com>). The particular SNP of interest can be searched by its rs number (<http://www.ncbi.nlm.nih.gov/SNP/>) from the ABI Web site. ABI 40 $\times$ or 80 $\times$ TaqMan[®] SNP genotyping assays can be ordered for different sample sizes: small scale for 1,500 reactions, medium for 5,000 reactions, and large for 12,000 reactions calculated from 5 μ L reaction/sample. The ready-to-use assay mixture is preloaded with (1) the forward and reverse primers for amplification of the polymorphic sequence, and (2) the two allele-specific TaqMan[®] probes descriptive of the SNP of interest, one labeled with VIC[®] dye and the other with 6FAM[™] dye. Every assay comes with an information file containing details on the chromosomal location of the SNP, allele frequency (only for validated assays) and the nucleotide sequence to which the VIC/FAM reporter dye is linked. Upon delivery, all assays should be protected from light by wrapping the tube in foil. The 40 $\times$ and 80 $\times$ assays should be diluted with 1 $\times$ TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) to 20 $\times$ working mixture for best stability. This assay mixture should be aliquoted in batches to avoid freeze–thaw cycling. Assay should be stored at –15 to –25°C.

For SNPs which do not have ready-made assays available, there are several approaches for the user to design probe and primer sets: (1) use ABI Primer Express Software (see below); (2) use ABI Custom TaqMan[®] design service; and (3) search of the literature or available

¹When the Taqman assay for a particular SNP is performed for the first time, it is recommended that a real-time PCR system (such as the ABI 7500 Real-time PCR system) is used to optimize the assay. After optimization, a regular PCR thermocycler, such as the ABI 9700 PCR system, can be used for the amplification step and the real-time PCR machine for the post-PCR data analysis.

databases, such as SNP500Cancer (<http://variantgps.nci.nih.gov/cgfsnp/pages/snp500.do>) for TaqMan[®] genotyping reagents for the SNP of interest.

Should the user require probes and primers to be designed, the ABI Primer Express software v2.0 provides assay design guidelines developed specifically for quantification assays. The general guidelines are as follows:

1. Import the DNA sequence containing the SNP of interest directly from NCBI dbSNP database (<http://www.ncbi.nlm.nih.gov/SNP/>).
2. Design forward and reverse amplification primers each with optimal lengths of 20 nucleotides and with matching T_m (59°C optimal) within 1°C of each other. Strive for an amplicon length of 75–150 bp. Avoid a guanine residue at the 5' end of the probe because a G residue adjacent to reporter dye will quench the reporter fluorescent even after cleavage. Target a T_m of 65–67°C for each probe and a minimum length of 13 nucleotides. Avoid runs of an identical nucleotide in the probe sequence (especially runs of four or more consecutive Gs should be avoided). Position the SNP site in the central third of the probe; if this is not possible, then shift the position of SNP site towards the 3' end of the probe but do not place it in the last two 3' nucleotides.

3. Methods

3.1. Preparation of Genomic DNA

Prepare 5–25 ng/μL of DNA in TE buffer (10 mM Tris pH 8.0, 0.1 mM EDTA) or in nuclease-free water. All samples should preferably have the same DNA concentration for genotyping to avoid a wide range of fluorescence activity within a genotype group (see Notes 2 and ³).

3.2. Preparation of Probes and SNP Assay Mixture

Dissolve self-designed MGB and TAMRA quenched probes in TE (10 mM Tris–HCl, 1 mM EDTA, pH8.0) or nuclease-free water at 100 μM concentration and the nonlabeled reverse and forward primers in TE (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) at 1,000 μM concentration. To make 20× SNP genotyping assay cocktail containing MGB quencher, use 4 μM for each of the VIC and FAM probe and 18 μM for each of the nonlabeled primer. To make 20× SNP genotyping assay containing TAMRA quencher, use 2 μM for each of the VIC and FAM probe and 14 μM for each of the nonlabeled primer. For example, to make 500 μL of 20× SNP assay mix with MGB quencher use 20 μL each of the VIC and FAM probes, 9 μL each of the nonlabeled primers and 442 μL nuclease-free water.

SNP assays purchased from ABI are ready to use.

3.3. TaqMan[®] Universal PCR Master Mix

This can be purchased from ABI. To avoid PCR contamination from previously amplified material, “carry-over,” inclusion of AmpErase[®] uracil-*N*-glycosylase (UNG) in the buffer is recommended (see Note 4). A single lot of master mix should be used in the same

²If specific controls for unique SNP genotypes are desired, the user has several resources for verifying whether the SNP of interest has been characterized for a reference cell from the International HapMap Project (<http://hapmap.ncbi.nlm.nih.gov/>) or the 1000 Human Genomes Project (www.1000genomes.org/). The user may determine whether the cells of interest can be purchased from Coriell Institute for Medical Research (<http://ccr.coriell.org/sections/Collections/NHGRI/?SsId=11>).

³The most common cause for trailing of genotype clusters, as illustrated in Fig. 1, is variation in the DNA concentration of the test samples. Trailing may be minimized by accurate measurement of DNA concentration using the picogreen method (following the manufacturer's instructions or by using a standard spectrometer) and use of concentrations of 5–25 ng/mL. Trailing of genotype clustering can also be caused by pipetting error/s, use of expired reagents, the presence of PCR inhibitor in DNA or degraded DNA.

experiment, since lot-to-lot variability can contribute to variation in fluorescence values. The PCR master mix includes a passive reference dye ROXTM, which is used to normalize fluorescent signals in individual wells caused by small differences in PCR reaction mixture volume (see Note 5).

3.4. PCR Mixture

For one well, a basic cocktail volume is 3.75 μ L containing 2.5 μ L 2 $\times$ Taqman universal PCR master mix, 0.25 μ L 20 $\times$ SNP assay mix, and 1 μ L of nuclease-free water. For full 96-well plate this cocktail is made as a 105 $\times$ mixture to allow for dead volume and pipetting errors as follows: 262.5 μ L 2 $\times$ TaqMan[®] universal PCR master mix, 26.25 μ L 20 $\times$ SNP assay mix, and 105 μ L of nuclease-free water. This cocktail can be scaled up for as many plates as needed.

3.5. Amplification of Target Sequence

1. Clean up the working surface with ethanol and water.
2. Label the 96-well optical plate with assay name and plate identifier and put the plate into a PCR splash-free support base (see Note 6).
3. Prepare the PCR mixture for as many plates as needed.
4. Pipet 1.3 μ L of individual DNAs into the bottom of wells, changing the tip between every DNA sample. Leave two wells in each plate devoid of template DNA as negative controls; for these wells, add 1.3 μ L of water in place of DNA.
5. Add 3.75 μ L of PCR master mix onto wall of each well. Change tips if necessary to prevent carry over from well to well.
6. Cover the plate with adhesive optical cover and centrifuge the plates at 1,500 rpm for 1 min to concentrate reagents at the bottom of the wells, and to remove any air bubbles.
7. Place the plate/s in ABI 9700 PCR thermal cycler or 7500 Real-time PCR system (see Note 1). If a real-time machine is used for the amplification step, please follow the manufacturer's instructions for plate setups and use of the machine. For MGB quencher probes, program the following cycling conditions: 2 min 50°C (this step only if UNG containing buffer is used), 10 min 92°C followed by 45 cycles of 30 s 92°C and 1 min 30 s 60°C. For TAMRA quencher probes, recommended conditions are: 2 min 50°C (perform this step only if UNG containing buffer is used), 10 min 95°C followed by 45 cycles of 30 s 95°C and 1 min 30 s 60°C. After cycling, keep plates at 4°C.
8. Read the plates in 7500 Real-Time PCR system using allelic discrimination with automated allele calling settings for the SDS 2.1 software (please refer to manufacturer's instructions) (see Note 7).

⁴TaqMan[®] Universal PCR Master Mix can be purchased with or without AmpErase[®] UNG. AmpErase[®] UNG containing Taqman PCR master mix is recommended for high-throughput PCR applications. AmpErase[®] UNG can prevent the reamplification of carryover PCR products by removing any uracil incorporated to single- or double-stranded DNA.

⁵The fluorescence signal from the passive reference dye ROXTM should be consistent across all wells and, therefore, it serves as an internal fluorescence signal reference to which the reporter dye (FAM and VIC) signal is normalized during data analysis. Signal normalization is necessary to correct for well-to-well non-PCR-related fluorescence signal fluctuations caused by minor differences in concentrations or volume from well-to-well. In the data analysis, the normalized reporter (Rn) is the ratio of the fluorescence intensity of the reporter dye (FAM or VIC) signal to the fluorescence intensity of the passive reference dye signal.

⁶Keeping the plate off the work surface prevents dirt getting onto outside of wells. Dirt can get into the real-time machine plate block and cause carryover fluorescence for all plates. Routine background fluorescence check of real-time machine plate block is recommended (please refer to manufacturer's instructions).

References

1. Kruglyak L. The use of a genetic map of biallelic markers in linkage studies. *Nat Genet.* 1997; 17:21–24. [PubMed: 9288093]
2. Wang DG, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science.* 1998; 280:1077–1082. [PubMed: 9582121]
3. Kruglyak L. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. *Nat Genet.* 1999; 22:139–144. [PubMed: 10369254]
4. The International SNP Map Working Group. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature.* 2001; 409:928–933. [PubMed: 11237013]
5. Kruglyak L, Nickerson DA. Variation is the spice of life. *Nat Genet.* 2001; 27:234–236. [PubMed: 11242096]
6. Sherry ST, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001; 29:308–311. [PubMed: 11125122]
7. International HapMap 3 Consortium et al. Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010; 467:52–58. [PubMed: 20811451]
8. 1000 Genomes Project Consortium et al. A map of human genome variation from population-scale sequencing. *Nature.* 2010; 467:1061–1073. [PubMed: 20981092]
9. Johnson GC, et al. Haplotype tagging for the identification of common disease genes. *Nat Genet.* 2001; 29:233–237. [PubMed: 11586306]
10. Carlson CS, et al. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet.* 2004; 74:106–120. [PubMed: 14681826]
11. Johnson AD, O'Donnell CJ. An open access database of genome-wide association studies. *BMC Med Genet.* 2009; 10:6–23. [PubMed: 19161620]
12. The MHC Sequencing Consortium. Complete sequence and gene map of a human major histocompatibility complex. *Nature.* 1999; 401:921–923. [PubMed: 10553908]
13. Walsh EC, et al. An integrated haplotype map of the human major histocompatibility complex. *Am J Hum Genet.* 2001; 73:580–590. [PubMed: 12920676]
14. Horton R, et al. Gene map of the extended human MHC. *Nat Rev Genet.* 2004; 5:889–899. [PubMed: 15573121]
15. Miretti MM, et al. A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am J Hum Genet.* 2005; 76:634–646. [PubMed: 15747258]
16. de Bakker PI, et al. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet.* 2006; 38:1166–1172. [PubMed: 16998491]

⁷The vast majority of TaqMan[®] assays perform well using standard assay and PCR cycling conditions (Fig. 2). However, should modifications be required, it is recommended that for every new assay, the first assay is run on a real-time PCR machine to observe PCR amplification plots (please refer to manufacturer's instructions). An Rn (fluorescence of the reporter dye divided by the fluorescence of a passive reference dye ROX[™]) vs. PCR amplification cycle number plot and delta Rn (Rn minus baseline) vs. PCR amplification cycle number plot are useful to identify and examine irregularities in amplification; a Ct (threshold cycle) vs. individual well plot is useful to locate outliers in detector data sets and to identify dirty wells in the instrument block.

If assay clustering for AA, AB, and BB genotypes can be discriminated as shown in Fig. 2, the rest of the amplifications for that SNP can be performed using a standard PCR machine followed by real-time PCR machine analysis for post-PCR end-point allelic discrimination plate reads. After amplification, the assay plates can be covered in foil and left at 4°C for few days before doing the end-point reads.

When the genotype clusters are not well separated (Figs. 3 and 4), adjustment of the annealing PCR temperature and/or reducing the number of PCR cycles may help. Start by increasing annealing temperature 1°C at a time from 60 to 61°C and then to 62° as needed (Fig. 3a, b) or decrease it by 1°C from 60 to 59°C (Fig. 4a, b). Although 5 µL reaction volumes are economical, some assays will require 10–25 µL for optimal results.

Failure to achieve clean separation of genotype clusters may indicate the presence of novel polymorphism in the probe sequence and requires sequencing of the target region.

17. Cree BA, et al. A major histocompatibility Class I locus contributes to multiple sclerosis susceptibility independently from HLA-DRB1* 15:01. *PLoS One*. 2010; 5:e11296. [PubMed: 20593013]
18. Baschal EE, et al. Defining multiple common “completely” conserved major histo-compatibility complex SNP haplotypes. *Clin Immunol*. 2009; 132:203–214. [PubMed: 19427271]
19. International MHC. Autoimmunity Genetics Network et al. Mapping of multiple susceptibility variants within the MHC region for 7 immune-mediated diseases. *Proc Natl Acad Sci USA*. 2009; 106(44):18680–18685. [PubMed: 19846760]
20. Valdes AM, Thomson G. Type 1 Diabetes Genetics Consortium. Several loci in the HLA class III region are associated with T1D risk after adjusting for DRB1-DQB1. *Diabetes Obes Metab*. 2009; 11(suppl 1):46–52. [PubMed: 19143814]
21. Shiina T, Inoko H, Kulski JK. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens*. 2004; 64:631–649. [PubMed: 15546336]
22. Saiki RK, et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*. 1985; 230:1350–1354. [PubMed: 2999980]
23. Mullis KB, Faloona FA. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol*. 1987; 155:335–350. [PubMed: 3431465]
24. Higuchi R, et al. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (NY)*. 1993; 11:1026–1030.
25. Morrison TB, Weis JJ, Wittwer CT. Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques*. 1998; 24:954–958. 960, 962. [PubMed: 9631186]
26. Piatek AS, et al. Molecular beacon sequence analysis for detecting drug resistance in *Mycobacterium tuberculosis*. *Nat Biotechnol*. 1998; 16:359–363. [PubMed: 9555727]
27. Marras SA, Kramer FR, Tyagi S. Multiplex detection of single-nucleotide variations using molecular beacons. *Genet Anal*. 1999; 14:151–156. [PubMed: 10084107]
28. Mhlanga MM, Malmberg L. Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. *Methods*. 2001; 25:463–471. [PubMed: 11846616]
29. Holland PM, et al. Detection of specific polymerase chain reaction product by utilizing the 5′—3′ exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proc Natl Acad Sci USA*. 1991; 88:7276–7280. [PubMed: 1871133]
30. Heid CA, et al. Real time quantitative PCR. *Genome Res*. 1996; 6:986–994. [PubMed: 8908518]

Fig. 1.

Example of “trailing” genotype clusters AA (*circle*), AB (*triangle*), and BB (*diamond*). Trailing describes the variation of fluorescence signals among a given genotype that has the appearance of being spread across an imaginary line that extends from the no template control (*square*). This phenomenon is often caused by variability in concentration of individual genomic DNAs.

Fig. 2.

Example of a good separation of genotype clusters AA (*circle*), AB (*triangle*), and BB (*diamond*). The no template controls (*square*) are distant from any clusters. Samples within a given cluster are close to the center of the cluster and the separation of the three clusters is very clear. Contrast this figure with that shown in Fig. 1.

Fig. 3. Example of SNPs assay requiring modification of PCR annealing temperature from 60 to 62°C. **(a)** At 60°C annealing, AB (*triangle*) and BB (*diamond*) clusters are not well separated, **(b)** the same assay with 62°C annealing temperature showing better separation of AB (*triangle*) and BB (*diamond*) clusters.

Fig. 4. Example of SNPs assay requiring modification of PCR annealing temperature from 60 to 59°C. **(a)** At 60°C annealing, AB (*triangle*) and AA (*circle*) clusters are not well separated, **(b)** the same assay with 59°C annealing temperature showing better separation of AB (*triangle*) and AA (*circle*) clusters.