Dynamic allele-specific hybridization

A new method for scoring single nucleotide polymorphisms.

W. Mathias Howell*, Magnus Jobs, Ulf Gyllensten, and Anthony J. Brookes

The Human Genome Project is expected to yield thousands of single nucleotide polymorphisms (SNPs), which can be used to identify polygenic contributors to disease and ultimately to design individualized prognostic strategies and therapies. Exploiting these, however, will require powerful and automated scoring methods. Current techniques have their roots in allele-specific oligonucleotide hybridization (ASOH). In its basic form (i.e., hybridization, stringent washing, and signal detection), ASOH is limited by the difficult challenge of defining discriminatory assay conditions.

Therefore, newer methods include additional steps furnishing more robust allele scoring. Such procedures include the ligation chain reaction (ASOH plus selective ligation and amplification), mini-sequencing (ASOH plus a single base extension), and DNA "chips" (miniaturized ASOH with multiple oligonucleotide probe arrays). Alternatively, ASOH with single- or dual-labeled probes has been merged with PCR, as in the 5´ exonuclease assay, and with molecular beacons. While effective, these methods also entail considerable optimization efforts and/or costly enzymatic or oligonucleotide labeling steps. We discuss here a new SNP scoring method, dynamic allele-specific hybridization (DASH), that does not suffer these drawbacks.

The key to DASH is dynamic heating and coincident monitoring of DNA denaturation. No additional enzymes or reaction steps are involved. DASH therefore retains the simplicity of ASOH, but achieves unambiguous discrimination of all SNP variations using standardized reaction conditions. The assay is conducted in a microtiter plate format compatible with automation, and uses convenient fluorescence signal detection.

Method principle
An overview of the DASH procedure is shown in Figure 1. A target sequence is amplified by PCR in which one primer is biotinylated. The biotinylated product strand is then hybridized to the target at low temperature. This forms a duplex DNA region that interacts with a double-strand–specific intercalating dye.

Upon excitation, the dye emits fluorescence proportional to the amount of double-stranded DNA (probe-target duplex) present. The sample is then heated, and fluorescence is continually monitored. A rapid fall in fluorescence indicates the denaturing (or "melting") temperature of the probe-target duplex. When performed under appropriate buffer and dye conditions, a single-base mismatch between the probe and the target results in a dramatic lowering of melting temperature (T_m) that can be easily detected.

A model experiment
Synthetic 5´-biotinylated oligonucleotide targets were made, representing each of the two alleles of a SNP in the PSEN2 gene with the variant base located in the center. These mock target DNAs and a 50/50 mixture of the two (to mimic the heterozygous state) were bound to individual microtiter plate wells. DASH was then performed on these samples using a probe complementary to one of the alleles. Resulting fluorescence profiles are shown in Figure 2A. To most readily interpret DASH results, we also show the negative first derivatives (downward slopes) of the fluorescence curves. This provides peak values directly related to the probe-target T_mS. As shown in Figure 2B, the sample matched to the probe produces a single high-temperature peak. The sample mismatched to the probe by one base pair produces a single lower temperature peak. The heterozygous sample undergoes two phases of denaturation (one for each allele present) and therefore produces two peaks in the negative first derivative. The DASH assay thus reveals whether the target DNAs contain sequences matched or single base mismatched to the hybridized probe.

Using a series of model systems similar to the one described above, we tested the effects of a number of dyes, hybridization buffers, pH values, sodium and potassium concentra-
ties, and additives. We found that an additive free buffer comprising 0.1 M Hepes, 50 mM NaCl, 1 mM EDTA, pH 8.0 with Syber Green I dye at 1:10,000 dilution was most effective in two important ways. First, it raised probe-target temperature values to a practical range of 60–70°C (for 15 mer probes), this being approximately 20°C higher than predicted for normal hybridization conditions. Second, it rendered large temperature differences (6–12°C) for target matched and target mismatched 15–21 mer probes. This buffer was therefore used in all the experiments presented below.

**Probe design**

Having defined basic assay conditions on model DNAs, further experiments were conducted using short DNA fragments (60–90 bp). These were PCR amplified from human genomic DNA and used directly for DASH without purification.

We examined the effect of oligonucleotide probe length and the location of the polymorphic base. Two C-T (G-A) SNPs were tested, employing primers biotin:5’-GCCCTCATGCGGATGCTC-3’ plus 5’-GGAAAGCAACGGAGT- TATCCAG-3’ (PSEN2 gene), and primers biotin:5’-CTTCTAATATCCTTAGCTGTCA-3’ plus 5’-TAGAAGTTGGATGTTTAGG-3’ (NDUFV2P1 gene). Genomic DNAs either homozygous or heterozygous for each of the alleles of these SNPs were available from other studies. DASH probes of lengths 13, 15, 19, and 25 nucleotides were designed for one allele of the PSEN2 gene, centered on the variant base in the sequence 5’-GTGACATTAGG(C)GGGAGAACAGC-3’. An additional 15mer probe was made in which the polymorphic position was moved two bases to the 3’ side of the central position (base 10). For the NDUFV2P1 gene, single allele DASH probes of lengths 13, 15, 19, and 25 nucleotides were designed with the variant base positioned 2 bases 3’ of the central position in the sequence 5’-GACAATTACTATG(A)GATTTGACAG-3’ for the PSEN2 gene.

Performing DASH revealed that increasing probe length over the tested range leads to a twofold increase in fluorescence signal strength, but a greater than threefold decrease in the spacing of the temperature peaks used for allele discrimination. Consequently, probes of 15 to 21 nucleotides are recommended for DASH assays. Moving the polymorphic position two bases from the center results in weaker and less discriminatory signals. Designing effective DASH probes thus requires nothing more than selecting 15–21 long nucleotide allelic sequences centered on the polymorphic nucleotide. This simplicity will greatly assist the exploitation of DASH in high-throughput situations.

**Applicability to all SNP types**

There are six possibilities for SNP variations in one strand of a DNA sequence, namely G-A, C-A, T-A, C-G, T-G, and C-T. Two of these pairs (G-A and T-C, C-A and G-T) represent mirrored variants in complementary DNA strands, meaning that there are only four distinct SNPs types that an SNP assay needs to score. A study was therefore undertaken to assess whether DASH could reliably detect all of these mismatches these SNPs could generate.

Examples of the four possible SNP types in the human genome were ascertained by a combination of literature surveys and experimentation. Individuals either homozygous or heterozygous for each of the alleles at these loci were identified by PCR with restriction enzyme digestion or direct sequencing (data not shown). For each locus, two oligonucleotide probes matching the two naturally occurring alleles were constructed for use in DASH. Also for each locus, two additional probes for the nonexisting alleles (the other two bases at the SNP positions) were also constructed. This allowed us to extend our examination of all four SNP types to four different sequence contexts. Each target region was then PCR amplified from genomic DNAs representing each known genotype, and these were interrogated by their respective probes in a series of DASH experiments.

Graphs of the negative first derivatives produced by the naturally occurring sequence probes are shown in Figure 3, confirming that DASH achieves effective scoring of all SNP types. In each case, the peaks for the pairs of homozygous samples differ by between 5 and 9°C, enabling unambiguous allele scoring. Results for the other two single base mismatched probes (nonexistent alleles) gave equivalent levels of discrimination, proving the robustness of the assay over several sequence contexts (data not shown).

As expected, heterozygous samples produced curves that are the sum of the two homozygous curves, in most cases clearly resolved into two distinct peaks. Variations in peaks heights were minimal and are likely to represent differing efficiencies of PCR amplification. All the observed temperature values were highly reproducible in several iterations of this experiment using different microtiter plate regions.

In practice, we suggest that one should perform serial probing of an immobilized test DNA with allelic probes for both variants. By including an intervening alkali washing step, this strategy has been found to be effective with little loss of signal on the second probing. All genotypes are thus doubly confirmed, further enhancing the precision of SNP scoring by DASH. Full experimental details are available from the authors upon request.

**Acknowledgment**

We gratefully acknowledge the financial support and intellectual contributions made to this project by staff at Hybrid Ltd., UK. We are also grateful for gifts of trial plasticware provided by LabSystems in Finland, and Boehringer Mannheim in Sweden.