The Journal of Molecular Diagnostics, Vol. ■, No. ■, ■ 2015

ELSEVIER

the Journal of Molecular Diagnostics
jmd.amjpathol.org

# Multiplexed High-Resolution Melting Assay for Versatile Sample Tracking in a Diagnostic and Research Setting

Q44 Céline Helsmoortel, R. Frank Kooy, and Geert Vandeweyer

*From the Department of Medical Genetics, University of Antwerp, Antwerp, Belgium*

Accepted for publication June 25, 2015.

Q3 Address correspondence to Geert Vandeweyer, Department of Medical Genetics, University of Antwerp, Prins Boudewijnlaan 43, 2650 Edegem, Belgium E-mail: geert.vandeweyer@uantwerpen.be.

Modern experimental procedures in molecular genetics, such as next-generation sequencing experiments, require that samples are taken along a whole series of wet- and dry-laboratory steps. It generally is Q4 accepted that by increasing the complexity and number of steps in the experimental pipeline, the risk of sample swaps increases. It therefore is recommended to confirm the identity of each individual sample at the end of any pipeline. Here, we present a versatile assay to determine the identity of samples rapidly and efficiently by genotyping 21 single-nucleotide polymorphisms (SNPs) using multiplex high-resolution melting. The selected SNPs also are present in whole-exome sequencing data, and comparison of the differentially obtained genotypes allows reliable identification of individual samples. In this assay, we combined primers interrogating two to three SNPs per high-resolution melting reaction, enabling the generation of the SNP genotype profile in only eight reactions per sample, limiting the hands-on time and Q5 minimizing the amount of reagents. This SNP profiling approach also can be used to track samples in custom next-generation sequencing enrichment panels by including these 21 SNPs in the target region, allowing for the often-required independent validation of sample identity in both clinical and research settings. *(J Mol Diagn 2015, ■: 1–7; http://dx.doi.org/10.1016/j.jmoldx.2015.06.011)*

Q6 The experimental set-up of next-generation sequencing (NGS) applications is complex because it includes DNA preparation, library preparation, sequencing, bioinformatics, and a final interpretation of the results. Each of these steps may be performed by different researchers and even at different facilities. The variety of analyses this technology offers in combination with the reduction in experimental costs has contributed to the establishment of NGS pipelines in virtually every genetics laboratory. A drawback of the often-complex NGS workflow is an increased risk for sample swaps. Hence, independent end point confirmation of sample identity is recommended. In fact, according to the guidelines from the American College of Medical Genetics, every diagnostic laboratory is urged to establish a system to prevent sample mix-up, ensuring the identity of tested Q7 samples (ISO 15189:2012).[1]

Different approaches are in use to assess the identity of DNA samples, depending on the setting. For instance, in forensic medicine or paternity testing, short tandem repeat polymorphisms are evaluated routinely. Because short tandem repeats are mostly outside coding regions, these repeats are hardly covered by the NGS technology, which focuses primarily on the coding regions of the genome.[2] Furthermore, accurate genotyping of repeats is a technical challenge. As a consequence, single-nucleotide polymorphisms (SNPs) located in coding regions are the preferred genetic markers for sample identification, especially because these can be interrogated with relative ease through independent techniques such as Sanger sequencing, high-resolution melting (HRM), SNP arrays, or the MassARRAY in combination Q8 with the iPlex kits (Agena). Q9

As elegantly argued by Pengelly et al,[3] a suitable SNP set to uniquely identify a sample should consist of bi-allelic markers that can be genotyped accurately using whole-exome

Supported by grants from the Belgian National Fund for Scientific Research-Flanders (R.F.K. and G.V.).

Q1 Q2 Disclosures: None declared.


http://dx.doi.org/10.1016/j.jmoldx.2015.06.011

sequencing (WES) and Sanger sequencing, have allele frequencies in HapMap populations ranging from 0.2 to 0.8, and do not have a deleterious effect. Furthermore, the SNPs should be positioned within the target regions of each commercially available WES enrichment technology (eg, Nextera Rapid Capture Exome; Illumina, San Diego, CA; SureSelect All Exon v4; Agilent, Santa Clara, CA; and SeqCap EZ Human Exome v3.0; NimbleGen, Madison, WI, kits). Based on this framework, we developed a fast and cost-efficient multiplexed HRM assay to generate genotype profiles containing 21 SNPs from the theoretically derived panel.[3] We show that genotyping of these 21 SNPs generates discriminative sample profiles, enabling identification and tracking in routine applications. Multiple approaches for multiplex HRM have been described.[4,5] Here, we did not include primer tails to adjust melting temperatures, but applied an *in silico* amplicon melting temperature prediction in combination with stringent primer design to achieve nonoverlapping melting temperatures. The sensitivity and specificity of this approach was shown in previous studies and allows for the combination of up to four assays per reaction.[4,5]

## Materials and Methods

### NGS Data

Validation was performed on three patient cohorts included in WES or targeted sequencing experiments.[6–8] Genomic DNA was extracted from peripheral blood using standard methods. DNA quality was confirmed by Qubit measurements, $A_{260}/A_{280}$ ratios, and agarose gelectrophoresis.

#### Cohort 1

Whole-exome data were obtained after TruSeq DNA Sample Preparation (Illumina) and SeqCap EZ Human Exome Library v3.0 enrichment (Roche NimbleGen), followed by 2 × 100 bp sequencing on the HiSeq 2000 (Illumina).[6]

#### Cohort 2

Targeted enrichment data were obtained after Library Preparation (KAPA Biosystems, Wilmington, MA) and SeqCap EZ Choice custom enrichment (Roche NimbleGen), followed by 2 × 150 bp sequencing on the MiSeq (Illumina).[7]

#### Cohort 3

Target enrichment data were obtained after custom HaloPlex 1-500 kb, ILMFST, 96 custom design (Agilent, Santa Clara, CA) enrichment, followed by 2 × 150 bp sequencing on the MiSeq (Illumina).[8]

### NGS Data Analysis

Target regions of 75 bp upstream and downstream of each SNP from the profile were added to the custom SeqCap and Haloplex designs. Data analysis was performed using an in-house pipeline, according to the latest GATK best practices.[9]

Coverage reports including the genotypes of the 21 SNPs were generated using an in-house QC-reporting tool (version 1.0.2) and supporting files (available at *https://bitbucket.org/geertvandeweyer/multpilexed_hrm_assay*, last accessed May 27, 2015). The report includes pretrimming and post-trimming quality metrics taken from FastX[10] and CutAdapt.[11] Included alignment metrics from Picard are insert size distribution, on-target enrichment, and library complexity estimation (*http://broadinstitute.github.io/picard*, last accessed May 27, 2015). Target region coverage details were plotted using BedTools[12] and custom R-code. Finally, genotyping metrics were taken from VCFstats,[13] GATK CallableLoci, and GATK Variant Quality Score Recalibration analysis.[14] Genotypes for the presented SNP panel were generated by GATK Unified

**Table 1** Multiplex HRM Mastermixes: Primers Concentrations Were 100 μmol/L

| SNP/primer used | Volume (μL) | Concentration (μmol/L) |
|---|---|---|
| 4 — 6 — 7 | | |
| Primers SNP 4 | 0.07 | 0.23 |
| Primers SNP 6 | 0.05 | 0.17 |
| Primers SNP 7 | 0.05 | 0.17 |
| Water | 5.68 | |
| 2 — 5 — 15 | | |
| Primers SNP 2 | 0.05 | 0.17 |
| Primers SNP 5 | 0.05 | 0.17 |
| Primers SNP 15 | 0.05 | 0.17 |
| Water | 5.72 | |
| 3 — 13 — 17 | | |
| Primers SNP 3 | 0.08 | 0.27 |
| Primers SNP 13 | 0.12 | 0.40 |
| Primers SNP 17 | 0.05 | 0.17 |
| Water | 5.52 | |
| 1 — 16 | | |
| Primers SNP 1 | 0.07 | 0.23 |
| Primers SNP 16 | 0.08 | 0.27 |
| Water | 5.72 | |
| 8 — 9 — 18 | | |
| Primers SNP 8 | 0.05 | 0.17 |
| Primers SNP 9 | 0.05 | 0.17 |
| Primers SNP 18 | 0.05 | 0.17 |
| Water | 5.72 | |
| 12 — 14 — 19 | | |
| Primers SNP 12 | 0.15 | 0.50 |
| Primers SNP 14 | 0.05 | 0.17 |
| Primers SNP 19 | 0.08 | 0.27 |
| Water | 5.6 | |
| 20 — 21 — 22 | | |
| Primers SNP 20 | 0.07 | 0.23 |
| Primers SNP 21 | 0.2 | 0.67 |
| Primers SNP 22 | 0.05 | 0.17 |
| Water | 5.38 | |
| 10 — 11 | | |
| Primers SNP 10 | 0.05 | 0.17 |
| Primers SNP 11 | 0.05 | 0.17 |
| Water | 5.82 | |





FLA 5.4.0 DTD ■ JMDI455_proof ■ 21 November 2015 ■ 2:04 am ■ EO: JMD15_0114