

Validation of variants using cost effective high-resolution melting (HRM) analysis predicted from target re-sequencing in *Eucalyptus*

Abdul Bari Muneera Parveen, Divya Lakshmanan, Modhumita Ghosh Dasgupta*

Institute of Forest Genetics and Tree Breeding, R.S. Puram, Coimbatore–641002, Tamil Nadu, India

Abstract – The advent of next-generation sequencing has facilitated large-scale discovery and mapping of genomic variants for high-throughput genotyping. Several research groups working in tree species are presently employing next generation sequencing (NGS) platforms for marker discovery, since it is a cost effective and time saving strategy. However, most trees lack a chromosome level genome map and validation of variants for downstream application becomes obligatory. The cost associated with identifying potential variants from the enormous amount of sequence data is a major limitation. In the present study, high resolution melting (HRM) analysis was optimized for rapid validation of single nucleotide polymorphisms (SNPs), insertions or deletions (InDels) and simple sequence repeats (SSRs) predicted from exome sequencing of parents and hybrids of *Eucalyptus tereticornis* Sm. × *Eucalyptus grandis* Hill ex Maiden generated from controlled hybridization. The cost per data point was less than 0.5 USD, providing great flexibility in terms of cost and sensitivity, when compared to other validation methods. The sensitivity of this technology in variant detection can be extended to other applications including Bar-HRM for species authentication and TILLING for detection of mutants.

Key words: exome sequencing, genotyping, high resolution melting analysis, validation, variants

Introduction

High resolution melting (HRM) for DNA analysis is a closed-tube analysis system and post real-time PCR analytical technique based on the principle that the melting curves of DNA fragments vary depending on base composition (Montgomery et al. 2007). The melting kinetics from the amplification profile of HRM-designed primers facilitates scanning and cataloging of single nucleotide polymorphisms (SNPs), mutations, and methylation in the genomic signature of individual species (Wojdacz and Dobrovic 2007, Toi and Dwyer 2008, Wittwer 2009). The advantages of HRM analysis include reduced risk of cross contamination and decreased analytical time. It is considered an accurate and sensitive PCR-based genotyping technology for cost-effective screening of genetic markers without *a priori* knowledge of the variation in template DNA sequence (Wittwer et al. 2003). This fluorescence-based melting analysis allows greater discrimination between homozygote and sensitive detection of heterozygote (Montgomery et al. 2007).

HRM based genotyping has been used for SNP (Han et al. 2012, Marshall et al. 2015, Gomes et al. 2018), insertion/

deletion (InDel) (Zhou et al. 2015) and simple sequence repeat (SSR) genotyping (Mackay et al. 2008, Distefano et al. 2012) in several plant species like *Vitis vinifera* L., *Podocarpus* L'Hér. ex Pers. and *Medicago sativa* L. Further, the applications of HRM has also been extended to analysis of candidate DNA barcode marker as reported in leguminous forage and pasture species (Ganopoulos et al. 2012), *Crocus sativus* L. (Jiang et al. 2014), *Sideritis* L. (Kalivas et al. 2014), *Phyllanthus amarus* Schumach. et Thonn. (Buddhachat et al. 2015), *Calendula officinalis* L. (Schmiderer et al. 2015), *Artemisia* L. (Song et al. 2016) and *Senna alexandrina* Mill. (Mishra et al. 2018), where HRM was reported as an efficient molecular tool to authenticate species of traded medicinal plants.

Targeting induced local lesions in genomes (TILLING) is a high-throughput technique and reverse genetic strategy to identify induced mutants. It involves steps such as endonuclease digestion, cloning and electrophoresis, which are critical and time-consuming. In addition, lack of complete genome sequence information for many plant species has limited the development of suitable TILLING targets. Taheri et al. (2017) reported that TILLING-HRM is a mu-

* Corresponding author e-mail: gmodhumita@gmail.com

tant detecting method that is accurate, sensitive, fast and cost-effective when compared to conventional methods. In plants, TILLING-HRM is reported in wheat for the detection of EMS induced mutation (Botticella et al. 2011), *Solanum lycopersicum* L. (Gady et al. 2009) and *Brassica rapa* L. (Lochlainn et al. 2011) for detection of point mutation, and *Oryza sativa* L. (Shan et al. 2018) for screening γ ray-induced mutations.

However, in woody perennials, use of this technology is limited to gene based SNP marker assay in sweet cherry (Ganopoulos et al. 2013), validation of sex-linked markers in *Pistacia vera* L. (Kafkas et al. 2015) and identification of genetically distinct species in *Podocarpus* sp. (Marshall et al. 2015).

The genus *Eucalyptus* L'Hér. is a widely planted hardwood species because of its superior growth, adaptability and wood properties and occupies 19.61 million hectares globally. This genus is targeted world-wide for genetic improvement programs due to its high commercial value as raw material for the paper and pulp industries. In eucalypts, application of NGS technology has led to the identification of genetic markers for linkage map construction (Grattapaglia et al. 2011, Neves et al. 2011, Bartholomé et al. 2015), genetic diversity analyses (Novaes et al. 2008, Dillon et al. 2014), genome wide association studies (Cappa et al. 2013, Silva-Junior et al. 2015) and marker-assisted selection (Resende et al. 2012). However, marker prediction using deep sequencing has been predominantly validated using Sanger sequencing (Novaes et al. 2008) and GoldenGate assay (Grattapaglia et al. 2011). To our knowledge, no reports on the use of HRM analysis for confirmation of variants (SNPs/ InDels/ SSRs) are reported in forest tree species, including *Eucalyptus*.

In our previous study, we reported the use of target capture and deep sequencing of xylogenesis-related genes for identification of high-throughput genetic markers in three *Eucalyptus* species for family-based QTL and Association analysis (Dasgupta et al. 2015). The study was further extended and 7 genes were selected for sequencing in parents, *Eucalyptus tereticornis* Sm. (clone Et86) and *Eucalyptus grandis* Hill ex Maiden (clone Eg9) and their hybrids generated from controlled hybridization for variant identification. The present study was undertaken to demonstrate the use of HRM-based genotyping for rapid validation of predicted variants (SNPs/InDels/SSRs) in both heterozygous and homozygous conditions in parents and hybrids. The use of HRM as a rapid, homogenous, highly specific, sensitive and cost effective methodology for high-throughput validation of markers in *Eucalyptus* is reported.

Materials and methods

Plant material and DNA isolation

A bi-parental mapping population of *Eucalyptus tereticornis* (clone Et86) \times *Eucalyptus grandis* (clone Eg9) was developed for construction of linkage map and QTL tagging

for wood property traits (Rajasugunasekar et al. 2015). The leaf tissues from the two parents and eight hybrids (H13, H48, H137, H160, H190, H218, H265, and H275) were harvested and immediately frozen at -80°C . Genomic DNA was isolated from the leaf tissues using the ArborEasy DNA isolation kit (Institute of Forest Genetics and Tree Breeding, India) and quantified using a NanoDrop ND1000 spectrophotometer (Thermo Scientific, USA).

Target capture and deep sequencing

Seven genes presumed to be involved in wood formation (On-line Suppl. Tab. 1) were mined from literature (Dasgupta et al. 2015) for in-solution target capture and deep sequencing in parents and hybrids. The regions targeted for capture included the exon, 3' and 5' UTRs and 500-1000 bp upstream regions.

Approximately 200 ng of genomic DNA from each sample was made up to 50 μL with Tris EDTA buffer and sonicated for ~ 210 s to fragment DNA into a size range of 200 to 400 bp. The size distribution was checked on the Agilent High Sensitivity Tape Station. Subsequently, libraries for each sample were constructed using SureSelect^{XT} Target Enrichment System for Illumina Paired-End Sequencing (Agilent Technologies, Santa Clara, California, USA) using the manufacturer's protocol. The library was hybridized at 65°C for 65 hours and the captured library was isolated using Magnetic Streptavidin-coated beads (Thermo Fisher Scientific, USA). The captured library was PCR amplified (12 cycles) and purified using the HighPrep PCR clean-up system (MagBio Genomics, MD, USA). The enriched amplified library was quantified using Qubit fluorometer (Thermo Scientific, USA) and 2×150 bp paired end sequencing was conducted using NextSeq 550 System (Illumina Inc., San Diego, CA, USA). Raw reads were quality checked using FastQC (Andrews, 2010). The quality passed raw reads were processed using Trim Galore (Krueger, 2015) for adapter clipping and low quality base trimming with filtering criteria of minimum read length 20 bp and minimum base quality Q30.

The processed reads were aligned to the reference *E. grandis* genome database hosted by the Phytozome portal (Goodstein et al. 2012) using Bowtie 2-2.0.5 (Langmead and Salzberg 2012) allowing one mismatch in a seed length of 22. The consensus sequences of all ten files were generated to identify the variants with SAMtools v.1.23 (Li 2011) and VarScan v. 2.34 (Koboldt et al. 2009). The genes were sequenced with an average read depth of 268.62 X to 561.72 X. The regions that were monomorphic between the parent were filtered out and positions that had polymorphic alleles between the parents were further explored in all the eight progenies and polymorphic variants (SNPs and InDels) were cataloged. The presence of polymorphic SSRs across parents and hybrids were mined from the consensus sequences using MISA computational tool (Beier et al. 2017). In the present study, HRM analysis was conducted in the parents (Et86 and Eg9) and eight hybrids (H13, H48, H137, H160, H190, H218, H265, and H275) (Tab. 1).

Tab. 1. Genotypes of parents (*Eucalyptus tereticornis* clone Et86 and *Eucalyptus grandis* clone Eg9) and eight hybrids (H13, H48, H137, H160, H190, H218, H265, and H275) of *Eucalyptus* mapping population obtained from target capture and exome sequencing data. SNP – Single Nucleotide Polymorphism, InDel – Insertion Deletion, SSR – Simple Sequence Repeat.

Gene name	Eg9	Et86	H13	H 48	H 137	H 160	H 190	H 218	H 265	H 275
SNPs										
<i>BP</i>	C/C	T/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T
<i>XTH</i>	C/C	T/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T
<i>PAAPA</i>	T/T	C/C	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T
<i>LIM1</i>	C/T	T/T	C/T	C/T	T/T	T/T	T/T	C/T	T/T	C/T
InDel										
<i>CesA4</i>	G/AC	G/-AC	G/-AC	G/-AC	G/-AC	G/-AC	G/-AC	G/-AC	G/-AC	G/-AC
SSRs										
<i>MUR3</i>	(TC)7	(TC)5	(TC)7	(TC)7	(TC)7	(TC)7	(TC)7	(TC)7	(TC)7	(TC)7
<i>ARF4</i>	(TGG)6	(TGG)4	(TGG)6	(TGG)6	(TGG)4	(TGG)4	(TGG)6	(TGG)4	(TGG)3	(TGG)6

Primer design and HRM analysis

Primer pairs targeting specific SNPs, InDels and SSRs were designed using Beacon designer version 7.90 (On-line Suppl. Tab. 2). Primers were designed following the criteria: predicted annealing temperature (T_m) of 58–61 °C, limited self-complementarity and amplicon lengths of 80–160 bp. PCR optimization was done to ascertain amplification of single product. After testing of different combinations of primers and template DNA concentrations, the reaction was optimized at 30 ng of DNA, 100 nmol L⁻¹ of each primer, 0.8 mg mL⁻¹ BSA, 2.5 mM MgCl₂ and 5 µL of Kappa Fast HRM master mix (Sigma, St. Louis, MO, USA) in a total volume of 10 µL. As suggested by the manufacturer, the amplification cycle was initially performed in two steps and the PCR program was as follows: 95 °C for 10 minutes; 40 cycles of 95 °C for 15 seconds, 55–60 °C for 1 minute (On-line Suppl. Tab. 2) and a dissociation cycle of 95 °C for 10 seconds, 60 °C for 1 minute, and 95 °C for 15 seconds (ramp rate, 0.3%) was used. Quantitative real time PCR was performed in the ABI PRISM 7500 Step one plus Sequence Detection System and HRM curves were analyzed using HRM software version 2.3 (Applied Biosystems, USA). PCR amplification was analyzed through the assessment of the CT value, end point fluorescence level, and the amplification efficiency for data QC. Data from low quality amplification including runs with CT value of over 30; outliers having end point fluorescence less than 50% of average fluorescence were removed from the analysis. HRM curve for each genotype was visually scored and variants were identified by examining normalized, difference and derivative melt plots. Melting temperatures of parents and hybrids for the genes used in HRM analysis are presented in On-line Suppl. Tab. 3. HRM assays for all samples and variants were conducted in duplicate and in independent replicates to ascertain the consistency of the results.

Results

Selection of candidate genes for in-solution target capture

In the present study, seven xylogenesis-related genes involved in different stages of wood formation including bio-

synthesis of lignin, cellulose, xyloglucan, cell wall formation and phytohormone signaling were selected (On-line Suppl. Tab. 1) based on our earlier study (Dasgupta et al. 2015). The length of the gene, position of the variants and their biological functions are presented in On-line Suppl. Tab. 1. The presence of polymorphic SNPs, InDels and SSRs across parents and eight hybrids was predicted using different computational pipelines (Tab. 1).

The raw sequence data was deposited in NCBI Short Read Archive with the accession number SRP152786 including *Eucalyptus tereticornis* (SRX4367138), *E. grandis* (SRX4367137), *Eucalyptus* hybrids H13 (SRX4367139), H48 (SRX4367141), H137 (SRX4367144), H160 (SRX4367136), H190 (SRX4367130), H218 (SRX4367129), H265 (SRX4367132) and H275 (SRX4367131).

DNA isolation and Primer design

HRM analysis was conducted for two *Eucalyptus* species (*E. tereticornis* and *E. grandis*) and eight hybrids generated by controlled hybridization. Total DNA isolated from leaves of two parents and hybrids was quantified and the concentration ranged from 173.5 ng µL⁻¹ to 868.8 ng µL⁻¹ with OD 260/280 = 1.69 to 1.98 across all samples. Seven primer pairs amplifying fragments less than 160 bp were chosen for the study.

HRM analysis for variant determination

HRM assay targeted four SNP regions of four different genes to determine if the HRM curves were in concordance with the exome sequencing results as shown in Tab. 1. Distinct HRM melting curves were observed for each SNP genotype in the parents and hybrids, each of which is represented in different color (Tab. 1, Figs. 1–4, On-line Suppl. Figs.1–4). In all samples evaluated, the replicates resulted in similar curves and were assigned to the same group, indicating the consistency and reproducibility of the HRM assay. All the four SNP markers produced polymorphic melting curves among the parents and hybrids investigated through HRM analysis.

All the three SNPs (*BP*, *PAAPA* and *XTH*) showed polymorphic melting curves differentiating between the homo-

zygous parents and heterozygous hybrids (Tab. 1, Figs. 1–3, On-line Suppl. Figs. 1–3). In *XTH*, the melting curve of hybrids (blue) was above that of the parents (red) (Fig. 3, On-line Suppl. Fig. 3) since the melting temperatures of hybrids were higher than those of the parents (Fig. 3, On-line Suppl. Fig. 3). In *BP* (Fig. 1, On-line Suppl. Fig. 1) and *PAAPA* (Fig. 2, On-line Suppl. Fig. 2) the melting curves of the hybrids were intermediate between the parents. The exome sequencing data for the parent Et86 showed a presence of nucleotide change in the position (718 bp) around the target SNP for the HRM analysis (Fig. 3). The change in melting curve profile of *XTH* in the parents when compared to *BP* and *PAAPA* could be attributed to the nucleotide change in the parent (Et86) at this locus.

Melting curve of *Lim1* indicated a heterozygous (C/T) and homozygous (T/T) allelic composition at the target locus for Eg9 (green) and Et86 (red) respectively (Fig. 4, On-line Suppl. Fig. 4). The melting curve of the four hybrid genotypes (H13, H48, H218, and H275) clustered in the same group as Eg9 (green), indicating similar allelic status. The remaining hybrids (H137, H160, H190 and H265) (red) grouped with Et86 and represented a homozygous condition. However, in H13 and H275, the melting curves did not show a decrease in fluorescence to the same baseline as other hybrids (H48 and H218) (Fig. 4, On-line Suppl. Fig. 4). The lack of decrease in fluorescence to the baseline in the two hybrids may indicate the presence of a second nucleo-

tide change in the PCR product or presence of secondary structures, resulting in inconsistent melting of PCR product. The presence of rare or unknown SNPs near the region selected for study may also interfere with the analysis and interpretation.

Targeted exome sequencing data for *CesA4* predicted the presence of deletion in Et86 (G/-AC) and insertion in Eg9 (G/G) with the amplicon size differing by 2 bp between the two parents (Tab. 1). All the eight hybrids selected for HRM analysis had the presence of InDel and so they grouped with the Et86 (red), where as Eg9 had a distinct melting profile (green) in the analysis (Tab. 1, Fig. 5, On-line Suppl. Fig. 5). The HRM curve results were consistent with sequencing results (Tab. 1).

Two SSR markers chosen for this study produced polymorphic melting curves differentiating the parents and hybrids when investigated using HRM analysis. For the SSR loci of *Mur3*, the parents Eg9 had the (TC)₇ repeat while Et86 had the (TC)₅ repeat and all the hybrids sequenced had the (TC)₇ repeat, similar to Eg9. The derivative melting curve and difference plot of SSR loci of *Mur3* showed a narrow curve for Eg9 (TC)₇ and a broader curve for Et86 (TC)₅ (Tab. 1, Fig. 6, On-line Suppl. Fig.6). All the hybrids grouped with parent Eg9 but a slight deviation in the shape of the melt curves for H218, H275 and H265 (green) was observed. Et86 was identified as different variant (red) (Fig. 6, On-line Suppl. Fig.6).

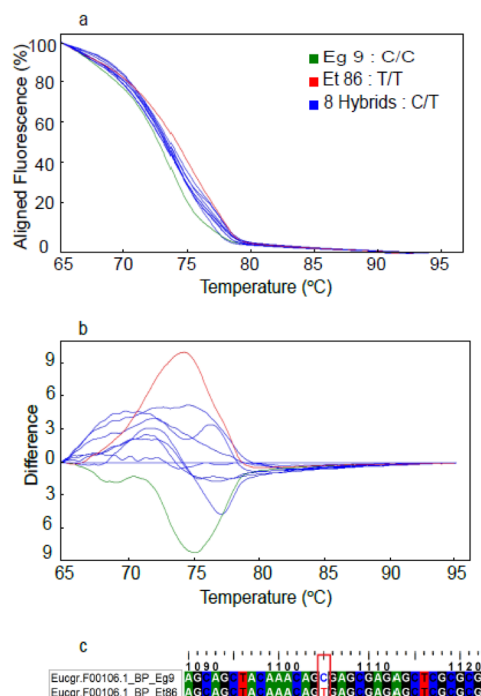


Fig. 1. Melting curve analysis of the single nucleotide polymorphic (SNP) marker in *BP*. Polymorphic high resolution melting (HRM) curves differentiated the genotypes into three groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *BP* gene sequence for the parents. SNPs are shown and highlighted in the red box.

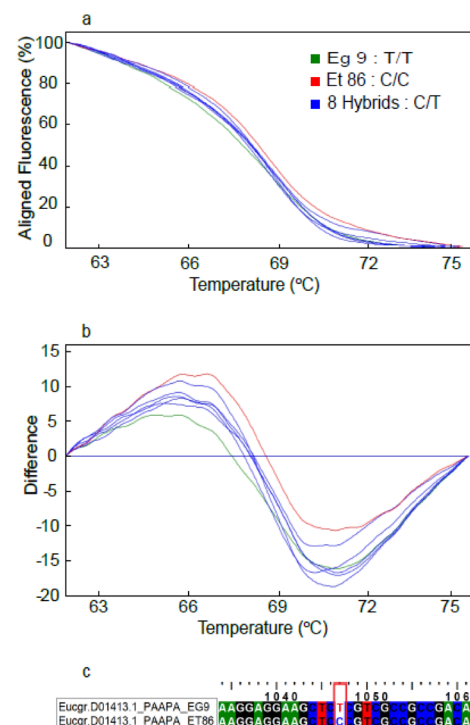


Fig. 2. Melting curve analysis of the single nucleotide polymorphic (SNP) marker in *PAAPA*. Polymorphic high resolution melting (HRM) curves differentiated the genotypes into three groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *PAAPA* gene sequence for the parents. SNPs are shown and highlighted in the red box.

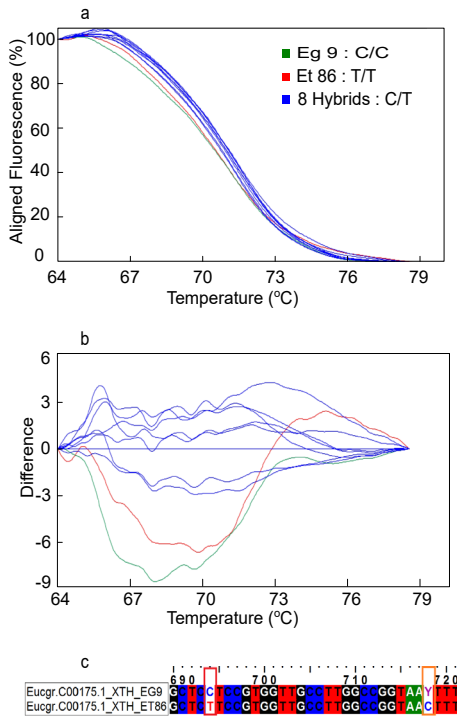


Fig. 3. Melting curve analysis of the single nucleotide polymorphic (SNP) marker in *XTH*. Polymorphic high resolution melting (HRM) curves differentiated the genotypes into three groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *XTH* gene sequence for the parents. SNPs are shown and highlighted in the red box. The change in the second nucleotide is highlighted with orange box.

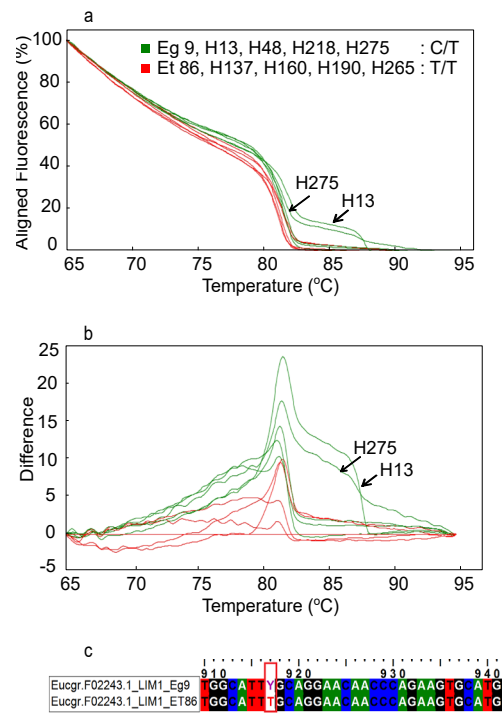


Fig. 4. Melting curve analysis of the single nucleotide polymorphic (SNP) marker in *Lim1*. Polymorphic high resolution melting (HRM) curves differentiated the genotypes into two groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *Lim1* gene sequence for the parents. SNPs are shown and highlighted in the red box.

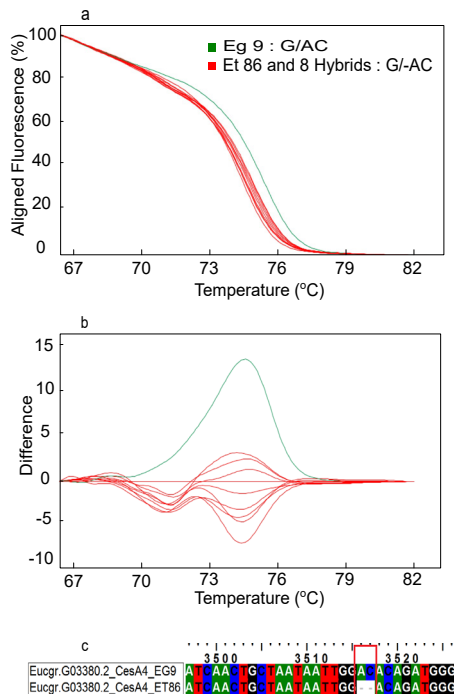


Fig. 5. Melting curve analysis of the insertion or deletion (InDel) marker in *Cesa4*. Polymorphic high resolution melting (HRM) curves differentiated the genotypes into two groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *Cesa4* gene sequence for the parents. Position of InDel is shown and highlighted in the red box.

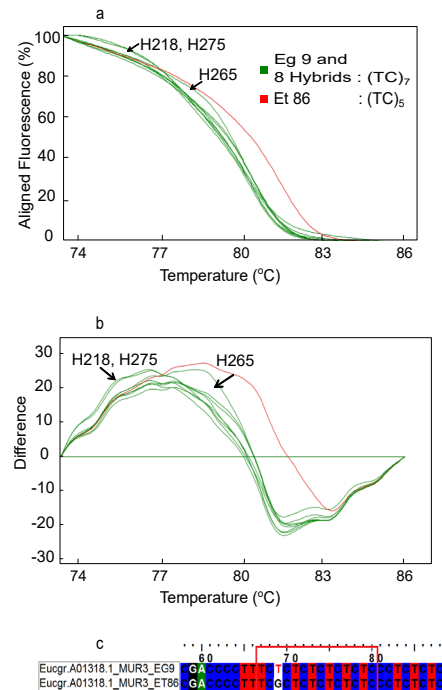


Fig. 6. Melting curve analysis of the simple sequence repeat (SSR) marker in *Mur3*. Polymorphic high resolution melting (HRM) differentiated the genotypes into two groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *Mur3* gene sequence for the parents. Position of SSR is shown and highlighted in the red box.

The sequencing results for the SSR loci in *ARF4* showed the (TGG)₆ repeat in Eg9 and (TGG)₄ in Et86. Three distinct variants were identified in the HRM analysis with melt curve difference between parents (Et86 and Eg9). The hybrids (H13, H48, H190 and H275) grouped with Eg9 (green) and hybrids (H137, H160 and H218) clustered with Et86 (red). The hybrid H265 showed distinct melting curve (blue) with repeat SSR motif of (TGG)₃. The sensitivity of HRM analysis allowed the detection of a complex situation for *ARF4*, which is the presence of (TGG)₆ repeat for parent Eg9, H13, H48, H190, H275, (TGG)₃ repeat for H265 and (TGG)₄ repeat in Et86, H137, H160 and H218 (Tab. 1). The variants identified using the assay is consistent with the sequencing results, confirming the discriminating power of the analysis (Fig. 7, On-line Suppl. Fig. 7).

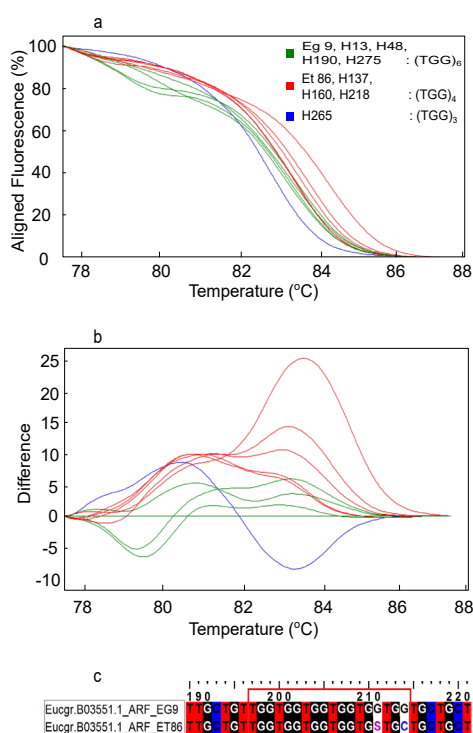


Fig. 7. Melting curve analysis of the simple sequence repeat (SSR) marker in *Arf4*. Polymorphic high resolution melting (HRM) differentiated the genotypes into three groups. a – normalized melting plot, b – difference melting curve. The curves with different color represent different genotypes. c – sequence alignment of the *Arf4* gene sequence for the parents. Position of SSR is shown and highlighted in the red box.

Discussion

Next generation sequencing has revolutionized genome research since it enables cost effective sequencing of millions of loci in a single run and has replaced the conventional approaches for genotyping like polyacrylamide gel electrophoresis system and denaturing capillary electrophoresis (Wenz et al. 1998, Wang et al. 2003). It is presently an indispensable approach for identifying and cataloging the full spectrum of genetic variation across the genome at a scale unattainable by traditional techniques such as Sanger sequencing (Ku

et al. 2012). This technology has facilitated and accelerated large-scale discovery, validation and assessment of genetic markers for high-throughput genotyping (Harismendy et al. 2009, Paszkiewicz and Studholme 2012). However, data generated using NGS platforms may suffer from high error rates mainly due to base-calling and alignment errors, limiting accurate variant prediction. Determining the rate of false positives and false negatives is one of the critical areas in genome analysis (Shigemizu et al. 2013), since it drastically influences downstream analyses including the identification of rare mutation and estimation of allele frequencies (Nielsen et al. 2011).

Variants predicted using NGS platforms are validated with secondary, orthogonal methods such as Sanger sequencing, Kompetitive allele specific PCR (KASP), TaqMan system (Applied Biosystems, Foster City, CA) and Illumina GoldenGate assay (Myakishev et al. 2001, Fan et al. 2003, De la Vega et al. 2005, Mu et al. 2016). While the cost of marker discovery using NGS platforms has drastically reduced over the past decade, inclusion of additional confirmation of variants has added additional costs to genetic testing. Hence, the need for alternative assays that are rapid, simple, homogenous, highly specific, sensitive and cost effective has increased for high throughput validation of markers.

HRM analysis has been reported as a convenient and cost effective PCR based method for genotyping SNPs, SSRs and InDels in plant species (Simko 2016). It was used in gene scanning, species identification (Słomka et al. 2017), as an alternate to SNaPshot analysis (Mehta et al. 2017), barcoding (Song et al. 2016, Mishra et al. 2018) and mutant screening (Taheri et al. 2017).

In the present study HRM was used to predict variants (SNPs/InDels/SSRs) in both heterozygous and homozygous conditions in *Eucalyptus* parents and hybrids. The sensitivity and utility of this technology for SNP genotyping was demonstrated in tetraploid *Medicago sativa* L., where parental genotypes and segregating progenies were characterized (Han et al. 2012). Mutation scanning for homozygous or heterozygous substitution and genotyping with the use of HRM assays was reported in *Olea europaea* L. germplasm (Muleo et al. 2009). Furthermore, HRM assay was used for species discrimination and phylogenetic analysis in *Podocarpus* sp. (Marshall et al. 2015), varietal discrimination in *Vitis vinifera* L. (Gomes et al. 2018), cultivar discrimination in sweet cherry (Ganopoulos et al. 2012) and validation of sex linked markers in *Pistacia vera* L (Kafkas et al. 2015).

The discriminating power of HRM melt curves was proven in *CesA4*, where the analysis could differentiate parents with and without the predicted InDel. This is in agreement with the earlier report in barley, where InDel markers were scored using electrophoresis and the HRM method (Zhou et al. 2015). They reported the discriminating ability of the HRM technique to distinctly differentiate samples with 2 bp difference.

The two SSR markers (*Mur3* and *Arf4*) targeted for this study produced polymorphic melting curves differentiating

the parents and hybrids when investigated through HRM analysis. However, mild changes in the shape of the melt curves for the same variants could be probably due to the minor systematic errors like pipetting of reagents. There are a number of potential problems and challenges that could affect the sensitivity of HRM assays. The quality and quantity of the DNA, the presence of salt from DNA isolation procedure, presence of primer dimer or other agents that interfere with PCR amplification can affect the sensitivity and reproducibility of HRM analysis, leading to incorrect genotyping or variant calling (Twist et al. 2013).

Results on SSR genotyping using HRM assay in trees is limited to grapevine (Mackay et al. 2008), citrus (Distefano et al. 2012) and sweet cherry (Ganopoulos et al. 2013). The studies have reported HRM as a flexible, cost-effective and closed-tube microsatellite genotyping method well suited for varietal certification and as an effective alternate to gel based platforms. In the present study, for the first time HRM analysis was optimized for the validation of variants predicted using deep sequencing methods in *Eucalyptus*. It could detect the presence of variants with high resolution and grouped

genotypes based on their predicted allelic composition. This enabled rapid scanning for the presence of SNPs/InDels/SSRs among heterozygous mapping population. The use of small amplicons (shorter than 150 bp) for genotyping allowed differentiation between homozygous genotypes and easy identification of heterozygous genotypes. A cost per data point analysis revealed that HRM assay per reaction was less than 0.5 USD and hence provided a greater flexibility for marker validation in terms of cost than sanger sequencing and other genotyping platforms.

Acknowledgements

The authors acknowledge Department of Biotechnology, Government of India for funding the research work (Project no. BT/PR10539/PBD/16/1064/2013). The funding support as research fellowship was provided to AMP by Department of Biotechnology, Government of India. The technical support extended by Dr. Neeta S. Madan, Senior Scientist, Reliance Research and Development Centre, Mumbai, India in conducting HRM assay is gratefully acknowledged.

References

- Andrews, S., 2010: FastQC: a quality control tool for high throughput sequence data, Available online. Retrieved May 17, 2018 from <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Bartholomé, J., Mandrou, E., Mabilia, A., Jenkins, J., Nabihoudine, I., Klopp, C., Schmutz, J., Plomion, C., Gion, J.M., 2015: High-resolution genetic maps of *Eucalyptus* improve *Eucalyptus grandis* genome assembly. *New Phytologist* 206, 1283–1296.
- Beier, S., Thiel, T., Münch, T., Scholz, U., Mascher, M., 2017: MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33, 2583–2585.
- Botticella, E., Sestili, F., Hernandez-Lopez, A., Phillips, A., Lafiandra, D., 2011: High resolution melting analysis for the detection of EMS induced mutations in wheat *SbElla* genes. *BMC Plant Biology* 11, 156.
- Buddhachat, K., Osathanunkul, M., Madesis, P., Chomdej, S., Ongchai, S., 2015: Authenticity analyses of *Phyllanthus amarus* using barcoding coupled with HRM analysis to control its quality for medicinal plant product. *Gene* 573, 84–90.
- Cappa, E.P., El-Kassaby, Y.A., Garcia, M.N., Acuña, C., Borralho, N.M.G., Grattapaglia, D., Poltri, S.N.M., 2013: Impacts of population structure and analytical models in genome-wide association studies of complex traits in forest trees: A case study in *Eucalyptus globulus*. *PLoS ONE* 8, e81267.
- Dasgupta, M.G., Dharanishanthi, V., Agarwal, I., Krutovsky, K.V., 2015: Development of genetic markers in *Eucalyptus* species by target enrichment and exome sequencing. *PLoS One* 10, e0116528.
- De la Vega, F.M., Lazaruk, K.D., Rhodes, M.D., Wenz, M.H., 2005: Assessment of two flexible and compatible SNP genotyping platforms: TaqMan SNP genotyping assays and the SNPlex genotyping system. *Mutation Research* 573, 111–135.
- Dillon, S., McEvoy, R., Baldwin, D.S., Rees, G.N., Parsons, Y., Southerton, S., 2014: Characterisation of adaptive genetic diversity in environmentally contrasted populations of *Eucalyptus camaldulensis* Dehnh. (River red gum). *PLoS One* 9, e103515.
- Distefano, G., Caruso, M., La Malfa, S., Gentile, A., Wu, S.B. 2012: High resolution melting analysis is a more sensitive and effective alternative to gel-based platforms in analysis of SSR – an example in citrus. *PLoS One* 7, e44202.
- Fan, J.B., Oliphant, A., Shen, R., Kermani, B.G., Garcia, F., Gunderson, K.L., Hansen, M., Steemers, F., Butler, S.L., Deloukas, P., Galver, L., Hunt, S., McBride, C., Bibikova, M., Rubano, T., Chen, J., Wickham, E., Doucet, D., Chang, W., Campbell, D., Zhang, B., Kruglyak, S., Bentley, D., Haas, J., Rigault, P., Zhou, L., Stuelpnagel, J., Chee, M. S., 2003: Highly parallel SNP genotyping. *Cold Spring Harbor Symposia on Quantitative Biology* 68, 69–78.
- Gady, A.L., Hermans, F.W., Van de Wal, M.H., van Loo, E.N., Visser, R.G., Bachem, C.W., 2009: Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations. *Plant Methods* 5, 13–14.
- Ganopoulos, I., Madesis, P., Tsafaris, A., 2012: Universal ITS2 barcoding DNA region coupled with High Resolution Melting (HRM) analysis for seed authentication and adulteration testing in leguminous forage and pasture species. *Plant Molecular Biology Reporter* 30, 1322–1328.
- Ganopoulos, I., Tsaballa, A., Xanthopoulou, A., Madesis, P., Tsafaris, A., 2013: Sweet cherry cultivar identification by High-Resolution-Melting (HRM) analysis using gene-based SNP markers. *Plant Molecular Biology Reporter* 31, 763–768.
- Gomes, S., Castro, C., Barrias, S., Pereira, L., Jorge, P., Fernandes, J.R., Martins-Lopes, P., 2018: Alternative SNP detection platforms, HRM and biosensors, for varietal identification in *Vitis vinifera* L. using F3H and LDOX genes. *Scientific Reports* 8, 5850.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., Rokhsar, D.S., 2012: Phytozome: a comparative platform for green plant genomics. *Nucleic acids research*, 40, D1178–D1186.
- Grattapaglia, D., Silva-Junior, O.B., Kirst, M., de Lima, B.M., Faria, D.A., Pappas, G.J.Jr., 2011: High-throughput SNP genotyping

- in the highly heterozygous genome of *Eucalyptus*: assay success, polymorphism and transferability across species. *BMC Plant Biology* 11, 65.
- Han, Y., Khu, D.M., Monteros, M.J., 2012: High Resolution Melting analysis for SNP genotyping and mapping in tetraploid alfalfa (*Medicago sativa* L.). *Molecular Breeding* 29, 489–501.
- Harismendy, O., Ng, P.C., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., Schork, N.J., Murray, S.S., Topol, E.J., Levy, S., Frazer, K.A., 2009: Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biology* 10, R32.
- Jiang, C., Cao, L., Yuan, Y., Chen, M., Jin, Y., Huang, L., 2014: Barcoding melting curve analysis for rapid, sensitive, and discriminating authentication of Saffron (*Crocus sativus* L.) from its adulterants. *BioMed Research International* 2014, 809037.
- Kafkas, S., Khodaeiaminjan, M., Güney, M., Kafkas, E., 2015: Identification of sex-linked SNP markers using RAD sequencing suggests ZW/ZZ sex determination in *Pistacia vera* L. *BMC Genomics* 16, 98.
- Kalivas, A., Ganopoulos, I., Xanthopoulou, A., Chatzopoulou, P., Tsaftaris, A., Madesis, P., 2014: DNA barcode ITS2 coupled with high resolution melting (HRM) analysis for taxonomic identification of *Sideritis* species growing in Greece. *Molecular Biology Reports* 41, 5147–5155.
- Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., Ding, L., 2009: VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25, 2283–2285.
- Krueger, F., 2015: Trim Galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, Available online. Retrieved June 2, 2018 from http://www.bioinformatics.babraham.ac.uk/projects/trim_galore
- Ku, C.S., Cooper, D.N., Polychronakos, C., Naidoo, N., Wu, M., Soong, R., 2012: Exome sequencing: dual role as a discovery and diagnostic tool. *Annals of Neurology* 71, 5–14.
- Langmead, B., Salzberg, S.L., 2012: Fast gapped-read alignment with bowtie 2. *Nature Methods* 9, 357–359.
- Li, H., 2011: A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 27, 2987–2993.
- Lochlainn, S.Ó., Amoah, S., Graham, N.S., Alamer, K., Rios, J.J., Kurup, S., Stoute, A., Hammond, J.P., Østergaard, L., King, G.J., 2011: High resolution melt (HRM) analysis is an efficient tool to genotype EMS mutants in complex crop genomes. *Plant Methods* 7, 43–49.
- Mackay, J.F., Wright, C.D., Bonfiglioli, R.G., 2008: A new approach to varietal identification in plants by microsatellite high resolution melting analysis: Application to the verification of grapevine and olive cultivars. *Plant Methods* 4, 8.
- Marshall, C.W., Chagné, D., Deusch, O., Gruenheit, N., McCallum, J., Bergin, D., Lockhart, P.J., Wilcox, P.L., 2015: A DNA-based diagnostic for differentiating among New Zealand endemic *Podocarpus*. *Tree Genetics and Genomes* 11, 69.
- Mehta, B., Daniel, R., McNeven, D., 2017: HRM and SNaPshot as alternative forensic SNP genotyping methods. *Forensic Science, Medicine and Pathology* 13, 293–301.
- Mishra, P., Shukla, A.K., Sundaresan, V., 2018: Candidate DNA barcode tags combined with high resolution melting (Bar-HRM) curve analysis for authentication of *Senna alexandrina* Mill. with validation in crude drugs. *Frontiers in Plant Science* 9, 283.
- Montgomery, J., Wittwer, C.T., Palais, R., Zhou, L. 2007: Simultaneous mutation scanning and genotyping by high-resolution DNA melting analysis. *Nature Protocols* 2, 59–66.
- Mu, W., Lu, H.-M., Chen, J., Li, S., Elliott, A.M., 2016: Sanger confirmation is required to achieve optimal sensitivity and specificity in next-generation sequencing panel testing. *The Journal of Molecular Diagnostics* 18, 923–932.
- Muleo, R., Colao, M. C., Miano, D., Cirilli, M., Intrieri, M. C., Baldoni, L., Rugini E., 2009: Mutation scanning and genotyping by high resolution DNA melting analysis in olive germplasm. *Genome* 52, 252–260.
- Myakishev, M.V., Khripin, Y., Hu, S., Hamer, D.H., 2001: High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. *Genome Research* 11, 163–169.
- Neves, L.G., McMamani, E., Alfenas, A., Kirst, M., Grattapaglia, D., 2011: A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in *Eucalyptus*. *BMC Genomics* 12, 189.
- Nielsen, R., Paul, J.S., Albrechtsen, A., Song, Y.S., 2011: Genotype and SNP calling from next-generation sequencing data. *Nature Reviews Genetics* 12, 443–451.
- Novaes, E., Drost, D.R., Farmerie, W.G., Pappas, G.J.Jr., Grattapaglia, D., Sederoff, R.R., Kirst, M., 2008: High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics* 9, 312.
- Paszkievicz, K., Studholme, D.J., 2012: High-throughput sequencing data analysis software: current state and future developments. In: Rodriguez-Ezpeleta, N., Hackenberg, M., Aransay, A.M. (eds.), *Bioinformatics for high throughput sequencing*, 231–248. Springer Science, New York.
- Rajasugunasekar, D., Vellaichamy, P., Mayavel, A., Sivakumar, V., Modhumitha, D., Bachpai, V.K.W., Yasodha, R., Nagarajan, B., 2015: Interfacing classical and molecular breeding in red gums in India: Observed and expected gains. *Proceedings October 2015 IUFRO Eucalypt Conference, China*, 21–22.
- Resende, M.D., Resende, M.F.Jr., Sansaloni, C.P., Petroli, C.D., Missiaggia, A.A., Aguiar, A. M., Abad, J.M., Takahashi, E.K., Rosado, A.M., Faria, D.A., Pappas, G.J.Jr., Kilian, A., Grattapaglia, D., 2012: Genomic selection for growth and wood quality in *Eucalyptus*: capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytologist* 194, 116–128.
- Schmiderer, C., Lukas, B., Ruzicka, J., Novak, J., 2015: DNA-Based Identification of *Calendula officinalis* (Asteraceae). *Applications in Plant Sciences* 3, 1500069.
- Shigemizu, D., Fujimoto, A., Akiyama, S., Abe, T., Nakano, K., Boroevich, K. A., Yamamoto, Y., Furuta, M., Kubo, M., Nakagawa, H., Tsunoda, T., 2013: A practical method to detect SNVs and indels from whole genome and exome sequencing data. *Scientific Reports* 3, 2161.
- Silva-Junior, O.B., Faria, D.A., Grattapaglia, D., 2015: A flexible multi-species genome-wide 60K SNP chip developed from pooled resequencing of 240 *Eucalyptus* tree genomes across 12 species. *New Phytologist* 206, 1527–1540.
- Simko, I., 2016: High-Resolution DNA melting analysis in plant research. *Trends in Plant Science* 21, 528–537.
- Słomka, M., Sobalska-Kwapis, M., Wachulec, M., Bartosz, G., Strapagiel, D., 2017: High resolution melting (HRM) for high-throughput genotyping – limitations and caveats in practical case studies. *International Journal of Molecular Sciences* 18, 2316.
- Song, M., Li, J., Xiong, C., Liu, H., Liang, J., 2016: Applying High Resolution Melting (HRM) technology to identify five commonly used *Artemisia* species. *Scientific Reports* 6, 34133.
- Taheri, S., Abdullah, T.L., Jain, S.M., Sahebi, M., Azizi, P., 2017: TILLING, High Resolution Melting (HRM), and next-generation sequencing (NGS) techniques in plant mutation breeding. *Molecular Breeding* 37, 40.
- Toi, C.S., Dwyer, D.E., 2008: Differentiation between vaccine and wild type varicella-zoster virus genotypes by high-resolution

- melt analysis of single nucleotide polymorphisms. *Journal of Clinical Virology* 43, 18–24.
- Twist, G.P., Gaedigk, R., Leeder, J.S., Gaedigk, A., 2013: High-resolution melt analysis to detect sequence variations in highly homologous gene regions: application to CYP2B6. *Pharmacogenomics* 14, 913–922.
- Wang, D., Shi, J., Carlson, S.R., Cregan, P.B., Ward, R.W., Diers, B.W., 2003: A low-cost, high-throughput polyacrylamide gel electrophoresis system for genotyping with microsatellite DNA markers. *Crop Science* 43, 1828–1832.
- Wenz, H., Robertson, J.M., Menchen, S., Oaks, F., Demorest, D.M., Scheibler, D., Rosenblum, B.B., Wike, C., Gilbert, D.A., Efcavitch, J.W., 1998: High-precision genotyping by denaturing capillary electrophoresis. *Genome Research* 8, 69–80.
- Wittwer, C.T., 2009: High-resolution DNA melting analysis: advancements and limitations. *Human Mutation* 30, 857–859.
- Wittwer, C.T., Reed, G.H., Gundry, C.N., Vandersteen, J.G., Pryor, R.J., 2003: High-resolution genotyping by amplicon melting analysis using LCGreen. *Clinical Chemistry* 49, 853–860.
- Wojdacz, T.K., Dobrovic, A., 2007: Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Research* 35, e41.
- Zhou, G., Zhang, Q., Tan, C., Zhang, X.Q., Li, C., 2015: Development of genome-wide InDel markers and their integration with SSR, DArT and SNP markers in single barley map. *BMC Genomics* 16, 804.