

Il metodo EpiHRMAssay è rapido, economico e affidabile per l'analisi della stabilità genetica ed epigenetica delle piante propagate *in vitro*

Rosario Muleo^{1*}, Marco Cirilli^{1,2}, Emilia Caboni³, Ines Delfino¹, Fabiano Gattabria¹, Maurizio Iacona⁴

¹ Università degli studi della Tuscia, Viterbo

² Università degli studi di Milano

³ Consiglio per la ricerca in agricoltura e per l'analisi dell'economia agraria, Centro di ricerca per l'olivicoltura, la frutticoltura e l'agrumicoltura (CREA-OFA), sede di Roma

⁴ Università di Pisa

EpiHRMAssay is quick, economical and reliable tool for detecting genetic and epigenetic stability in agamic propagated plants

Abstract. Reliable and cost-effective assays with adequate sensitivity are required to detect the DNA methylation profile in plants for scientific and industrial purposes. To address this need, a novel approach was developed to quantify the overall methylation status at target loci, including CpG islands, and to enable high-throughput analyses in plants. The proposed assay, named EpiHRMAssay, is based on the combination of *in tube* High Resolution Melting Analysis (HRMA) on bisulfite-treated templates with the *in silico* prediction of the melting profile of virtual epialleles using uMELTSM software. The *in silico* simulation of predicted melting profiles provides the value of the predicted melting temperature (T_{m-s}) of a set of epialleles characterized by different numbers of methylated cytosines (#^mC) or the same number of #^mC, but located at different sites, generating different configurations. The T_{m-s} values and respective number of #^mC of the generated set of epialleles were used to build calibration models, enabling the quantification of methylation in unknown samples using only the *in tube*-observed melting temperature (T_{m-o}). The reliability of the calibration models was further improved using a few synthesized standard epialleles that have *ad hoc* configurations of methylated sites and extreme predicted T_{m-s} values. The EpiHRMAssay was validated by analysing the melting profile of the desired template of the promoter region of key genes (*CMT3*, *DDM1*, and *ROS1*) involved in the regulation of methylation/demethylation processes and chromatin remodelling within a clonal population of *in vitro*-propagated *Prunus persica* plants. The results demonstrate that EpiHRMAssay provides useful quantitative

information regarding the overall methylation status, representing a sensitive and reliable tool for locus-specific large-scale research and diagnostic contexts of the regulative regions of genes. Moreover, it is a rapid, affordable and cost-effective method, allowing a high-throughput screening of a large number of individuals. EpiHRMAssay also provides complementary information for the assessment of heterogeneous methylation in plants and can address an array of biological questions on epigenetic regulation for diversity studies, as well as for genetic identity and traceability of plant large-scale functional genomics.

Key words: DNA methylation, epigenetic regulation, High Resolution Melting Analysis (HRMA), *in silico* prediction of melting temperature, somaclonal variation.

Introduzione

La metilazione del DNA è un meccanismo epigenetico con un ruolo nella protezione del genoma da elementi mobili e retrovirus, nella determinazione dell'impronta genomica e del silenziamento indotta da eterocromatina (Zilberman *et al.*, 2007); inoltre, fornisce istruzioni per l'espressione genica, regolando dove, quando e quanto un gene debba essere espresso (Bell *et al.*, 2011). Nelle piante la metilazione del DNA avviene in contesti di sequenza simmetrica (CpG) e di sequenza asimmetrica, come CpHpG e CpHpH (H = A, C o T), diffuse tra regioni geniche e intergeniche (Becker *et al.*, 2011). Nelle piante sono presenti geni che controllano e regolano gli eventi epigenetici, come: *Methyltransferase1* (*MET1*), *Chromometilase1, 2, 3* (*CMT1, 2, 3*), *Domain Rearranged Methyltransferases* (*DRMs*), *Decreased DNA Methylation1* (*DDM1*), *Repressor Of Silencing1* (*ROS1*), e le proteine *Demeter-like* (*DML2* e *3*).

* muleo@unitus.it

La cultura dei tessuti *in vitro* è un'importante attività industriale e scientifica, con applicazioni nella propagazione, nel miglioramento genetico e biotecnologico, nella conservazione del germoplasma. Le condizioni *in vitro* sono stressanti (Cassells e Curry, 2001) e le piante coltivate talvolta mostrano fenomeni di variabilità genetica, epigenetica e fenotipica, con la comparsa di varianti somaclonali (Miguel e Marum, 2011). I cambiamenti genetici, epigenetici e fenotipici influenzano l'identità vegetale e l'efficienza economica della micropropagazione.

In questo lavoro si è considerato l'assunto che i prodotti PCR, generati da DNA trattati con bisolfito e con diversi contenuti di citosina metilata (^mC), hanno differenti temperature di melting (T_m), pertanto sono visualizzabili con coloranti intercalanti del DNA. Le differenze di T_m sono rilevabili in un termociclatore accoppiato ad un fluorometro (White *et al.*, 2007). Con il software uMELT (Dwight *et al.*, 2011) sono stati predetti i valori di T_m di una gamma di epialleli virtuali relativi a geni target, caratterizzati da numero e/o posizioni diverse delle ^mC. La temperatura di melting prevista *in silico* (T_m-s) di ogni epiallele e il rispettivo numero di ^mC sono stati utilizzati per costruire un modello di calibrazione per quantificare lo stato di metilazione complessivo di epialleli sconosciuti in campioni vegetali sulla base della temperatura di melting osservata sperimentalmente (T_m-o). EpiHRMAssay è stato validato analizzando lo stato di metilazione delle regioni del promotore dei geni *CMT3*, *DDMI* e *ROSI* all'interno di una popolazione di piante di pesco propagate *in vitro*.

Materiali e metodi

Gli espianti della cultivar di pesco Rich Lady sono stati clonati a partire da un'unica gemma ascellare, campionata da una pianta in campo presso il CREA-OFA. La coltura è stata effettuata in contenitori Magenta (Sigma, Italia) contenenti 50 ml di terreno di moltiplicazione (Gentile *et al.*, 2014). Il terreno è stato integrato da 1,11 μM BA, 16,3 μM di adenina solfato, 0,29 μM IBA e 0,19 μM GA₃. Il pH è stato titolato a 5,7. Le colture sono state mantenute a 24±2°C, a un fotoperiodo di 16 ore e all'irradianza di 40 μM m⁻² s⁻¹, generata da tubi fluorescenti (Fluora L58 vv/77, Osram, Italy). Gli espianti sono stati subcolturali ogni 21 giorni.

Il DNA è stato estratto da foglie campionate dalla pianta donatrice in campo e da piantine clonate coltivate *in vitro* alla 4^a, 18^a e 23^a subcoltura, con il protocollo CTAB modificato (Muleo *et al.*, 2009). Il trattamento con bisolfito del DNA è stato eseguito con il

kit EpiTect Bisulphite (Qiagen, Hilden, Germania) in un termociclatore, secondo il protocollo: 5 m. a 95°C, 25 m. a 60°C, 5 m. a 95°C, 85 m. a 60°C, 5 m. a 95°C, 175 m. a 60°C e a overnight a 20°C. Il DNA trattato, recuperato con la colonna spin EpiTect, è stato eluito in un volume finale di 50 μL, per una concentrazione teorica di 20 ng μL⁻¹.

Le sequenze nucleotidiche dei geni *DDMI*, *CMT3* e *ROSI* sono state identificate in Peach Genome V1.0 (Verde *et al.*, 2013) con tBLASTX e impiegate per il disegno di primer con Primer3. Gli epialleli standard non metilati (UMstd) dei geni sono stati ottenuti amplificando le regioni target da DNA non trattato, gli amplificati sono stati trattati con bisolfito utilizzando il kit EpiTect Bisulphite (Qiagen, Hilden, Germania) e utilizzati come controllo UMstd per il dosaggio HRMA. Epialleli standard di *CMT3* sono stati sintetizzati da GeneArt (Life Technologies, Carlsbad, CA) come frammenti di DNA a doppio filamento.

Le amplificazioni PCR sono state eseguite in Light Cycler® 1.5 (Roche, Germania), utilizzando kit SensiMix (Quantace, USA) e LC-Green II plus colorante (Idaho Technology, USA), in un volume totale di 20 μL contenenti: 4 μL Mix Buffer, opportuna concentrazione di MgCl₂ (0,75 mM per *CMT3* e *ROSI* e 1 mM per *DDMI*) e 100 nM di ciascun primer disegnato per il DNA non trattato e per quello trattato con bisolfito, 1 μL di LC-Green colorante, 1,5 μL di Taq polimerasi e 1 μL (20 ng μL⁻¹) di DNA non trattato e di quello trattato con bisolfito. Il profilo di amplificazione è stato: 10 min a 95°C, 50 cicli di 10 s a 95°C, 10 s alla temperatura di annealing del primer e 10 s a 72°C. Le reazioni HRMA sono state eseguite separatamente nello strumento HR1 (Idaho Technology, Idaho USA) con rampa di temperatura da 70° a 90°C, con valori crescenti di 0,10°C s⁻¹. Due repliche biologiche, con due repliche tecniche ciascuna, sono state eseguite per ogni analisi. Le curve di fusione sono state normalizzate con il software fornito con lo strumento HR1 e visualizzate con gli algoritmi Derivative e Difference Plot. La fusione ad alta risoluzione, sensibile alla metilazione digitale (dMS-HRM), è stata eseguita usando una serie di diluizione da 1:50 a 1:100, a partire da 1 ng μL⁻¹ di DNA. Il profilo di amplificazione HRMA e PCR è stato lo stesso sopra descritto, ad eccezione del numero di cicli, impostato a 60. Ulteriori particolari sono riportati in Cirilli *et al.* (2017).

Risultati

Le mutazioni influenzano lo stato di metilazione con conseguente modifica del valore di T_m. Pertanto,

la prima analisi HRMA è stata effettuata su DNA non trattato con bisolfito per verificare la presenza di mutazioni nelle regioni del promotore dei geni. L'analisi degli ampliconi (profilo Tm-o) ha escluso la presenza di variazioni genetiche (fig. 1A). L'analisi condotta su DNA trattato con bisolfito degli stessi campioni ha mostrato differenze nei profili della curva (fig. 1B), rivelando la presenza di epialleli metilati. Nella maggior parte dei campioni è presente una riduzione del livello di metilazione rispetto a quella del donatore, come evidenziato dal valore Tm-o inferiore, mentre un livello della metilazione maggiore è stato rilevato in alcuni campioni (fig. 1B).

Per la quantificazione dello stato globale di metilazione dei campioni, sono stati sviluppati modelli di calibrazione per ciascun gene con le procedure:

a) la generazione di un insieme di epialleli virtuali delle regioni target è avvenuta simulando un trattamento *in silico* del DNA con bisolfito, convertendo ogni citosina, delle CpG island (denominate E), CpHpG e CpHpH (denominate L), individuate con Methyl Primer Express®, in uracile. Per ciascun gene, in base al numero potenziale di siti metilabili, è stato generato un insieme di classi virtuali di epialleli, tramite il tool uMELT. La modifica *in silico* dello stato di metilazione di ogni citosina ha permesso di ottenere per *DDMI*, *CMT3* e *ROS1*, rispettivamente, 22, 23 e 40 classi di epialleli virtuali. Ogni classe può assumere un insieme di configurazioni, con numero fisso di °C ma con pattern di distribuzione degli °C diverso tra i siti disponibili; pertanto ogni configurazione è caratterizzata da un proprio profilo di melting e da un valore specifico di Tm-s.

b) perché il modello di calibrazione sia robusto deve avere una variabilità dei valori di Tm-s all'interno di una classe minore a quello della variabilità tra le classi di epialleli. A tal fine è stata stimata una gamma ragionevole di valori Tm-s per ogni classe di epialleli, considerando un sottoinsieme di configurazioni in base alla frequenza della metilazione nei siti E (CpG) o nei siti L (CpHpG / CpHpH), consentendo di stimare la gamma di variabilità (Cirilli *et al.*, 2017). Pertanto una serie di configurazioni, da 3 a 7, sono state considerate per ogni classe di epialleli e i relativi valori Tm-s sono stati utilizzati nel modello di calibrazione.

c) un profilo di melting con un singolo picco è stato previsto da uMELT per quasi tutte le configurazioni degli epialleli *DDMI* e *ROS1*. Nella sequenza di *CMT3*, un dominio ricco di GC, con quattro siti CpG nidificati, influenza il profilo di fusione previsto in molte configurazioni, determinando una forma curva complessa (fig. 2A). La frequenza dei profili complessi è più elevata per le configurazioni con metilazione in tutti i quattro siti CpG, indipendentemente dallo stato di metilazione delle altre citosine che circondano il dominio e per le classi di epialleli con 5 - 15 °C. Ad esempio, il profilo di melting delle configurazioni 5E (classe 5°C) e 14E1L (classe 15°C) ha due picchi di melting, mentre la configurazione 10E (classe 10°C) ha un profilo asimmetrico con un'inclinazione verso temperature elevate (fig. 2B). Nelle classi con più di 15 °C, la frequenza di profili complessi diminuisce. La complessità del dominio di melting è risolvibile o attenuabile in uMELTSM, modificando i valori del parametro σ ,

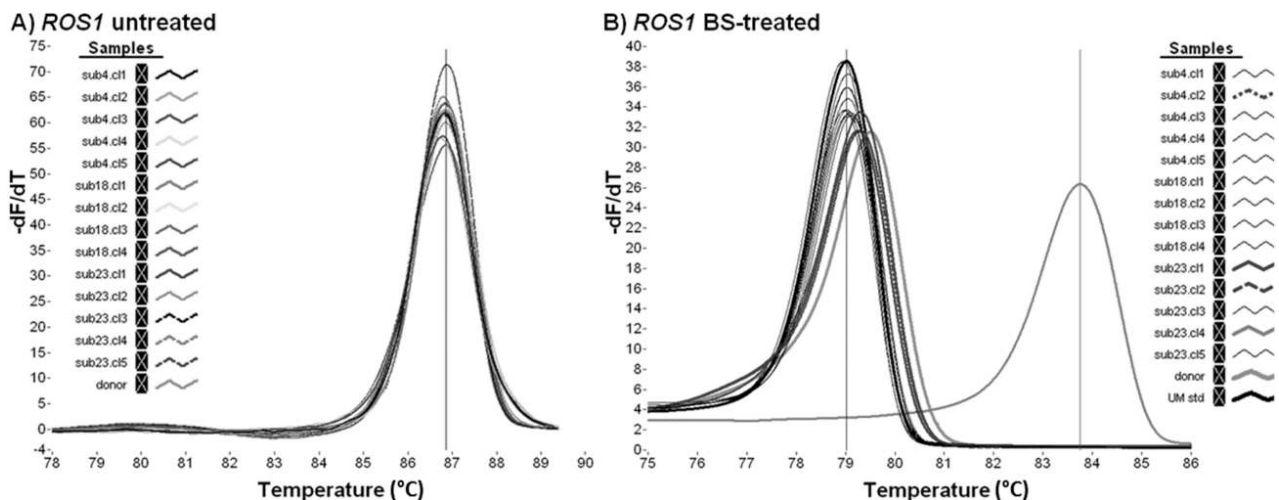


Fig. 1 - Derivate plot della curva di melting del test HRMA del promotore del gene *ROS1*, degli ampliconi ottenuti da DNA non trattato (A) e da DNA trattato con BS (B). I cursori verticali in (B) indicano rispettivamente il valore Tm (cursore sinistro) dei picchi di fusione del frammento standard (*UMstd*) e del clone 23sub.cl4 (cursore destro).

Fig. 1 - Derivative plot from the HRMA assay of the *ROS1* gene promoter, resulting from the amplification of untreated (A) and BS-treated (B) DNA templates. The vertical cursors in (B) indicate the T_m value of the melting peaks (left cursor) of the standard fragment (*UMstd*) and clone 23sub.cl4 samples (right cursor).

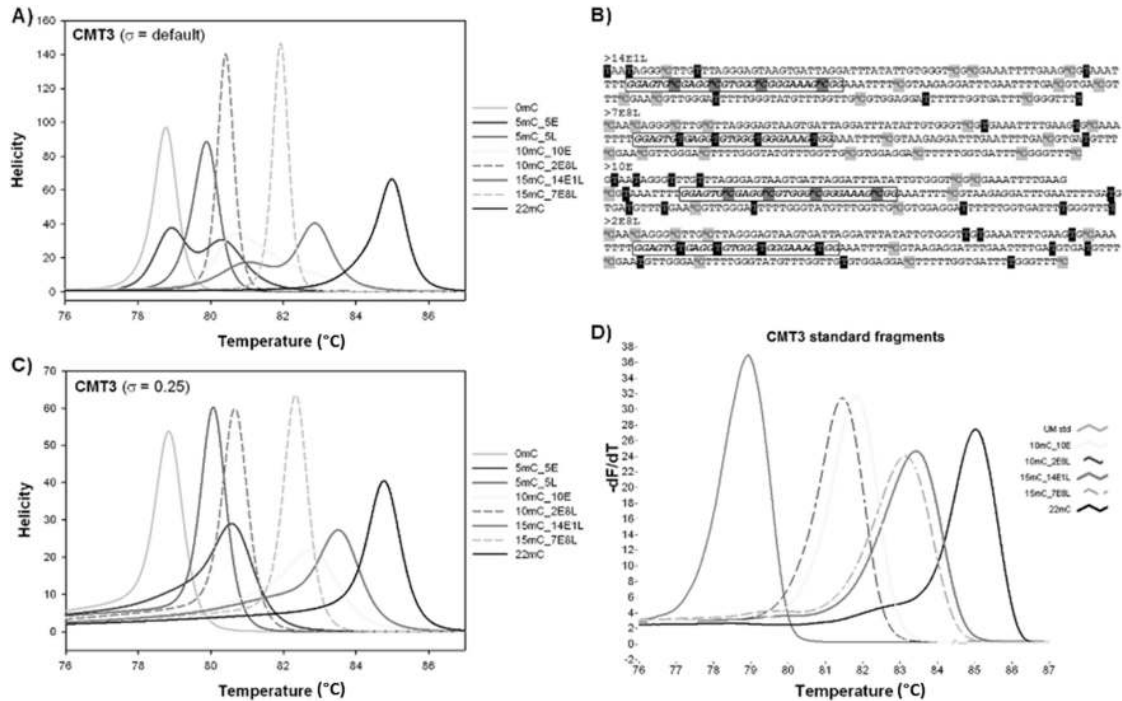


Fig. 2 - Il dominio ricco di GC è racchiuso all'interno della casella nera (A). L'allargamento della curva di melting derivata è regolabile dalla variazione del valore del parametro σ . Derivate plot della curva di melting di CMT3 predette *in silico* per le configurazioni estreme delle classi di epialleli con 0, 5, 10 e 15 citosine metilate, utilizzando il valore di default (0.00021) del parametro σ (B) e il valore σ di 0.25 (C). La regolazione dei parametri σ ha ridotto la complessità della curva di melting nelle configurazioni 14E1L e 10E, la quale è stata confermata sperimentalmente dei rispettivi ampliconi sintetici in test HRMA (D).

Fig. 2 - The GC-rich domain is enclosed within the empty black box (A). The broadening of the derivative melting curve can be modulated through the value of σ parameter. Derivative plot of *in silico* predicted melting curve of CMT3 amplicon for extreme configurations of the epiallele classes with 0, 5, 10 and 15 methylated cytosine, using default value (0.00021) of σ parameter (B), and σ value of 0.25 (C). The adjustment of σ parameters reduced the complexity of melting curve in the configurations 14E1L and 10E, as also experimental validated by the *in tube* HRMA of the respective synthetic amplicons (D).

tenendo conto della cooperatività dei domini di melting. Aumentando il valore σ , fino a 0,25, le curve di fusione si risolvono in un singolo picco di melting (fig. 2C). L'effetto della metilazione CpG all'interno del dominio ricco di GC è stato sperimentalmente verificato utilizzando ampliconi sintetici per configurazioni rappresentative della classe 10^{mC} (10E e 2E8L) e classe 15^{mC} (14E1L e 7E8L), inclusa anche quella completamente metilata 22^{mC} (14E8L). Il profilo di fusione derivante dall'analisi HRMA delle configurazioni 14E1L e 10E ha mostrato un picco singolo leggermente allargato (fig. 2D), confermando l'ipotesi che un'adeguata impostazione del parametro σ migliori la simulazione e risolva la complessità. Il valore del parametro σ che meglio approssima le curve sperimentali è stato di 0,25.

d) il modello di calibrazione è stato costruito ipotizzando la dipendenza lineare tra il numero di citosine metilate ($\#mC$) e l'aumento della temperatura di melting del corrispondente frammento (T_m), rispetto alla temperatura di fusione del frammento standard (UMstd T_m). I valori di T_m-s ottenuti dall'insieme degli epialleli selezionati sono stati utilizzati per costruire un modello che preveda il numero di ^{mC} di

una regione, in base all'ipotesi di una correlazione lineare tra $\#mC$ e T_m , $\#mC = a + b(T_m-s - UMstd T_m-s)$ (fig. 3). I valori osservati del coefficiente di correlazione ($R = 0,993, 0,978$ e $0,991$ per *DDMI*, *ROSI* e *CMT3*) hanno confermato l'ipotesi della dipendenza lineare tra le variabili considerate. Oltre al set termodinamico MELTSIM, possono essere impiegati i set di predizione dei profili della curva di melting descritti da Huguet *et al.* (2010) e Weber (2015), entrambi disponibili in uMELTSM.

I modelli di calibrazione sono stati applicati anche agli ampliconi di ciascun gene. Nel caso di *ROSI* (fig. 4), i cloni hanno mostrato bassi livelli di metilazione, ad eccezione del campione sub23.c14 ($T_m-o = 83,76 \pm 0,05^\circ C$) e un rispettivo numero previsto di 26,65 ^{mC}. L'affidabilità delle previsioni è stata valutata verificando il $\#mC$ determinato con EpiHRMssay con il sequenziamento del DNA degli stessi campioni trattato con bisolfito. I risultati del sequenziamento del DNA trattato di tutti gli amplificati sperimentati hanno confermato la fiducia dei modelli di previsione. Infatti, il numero previsto di $\#mC$ differisce, entro gli errori, da quelli osservati inferiori a ± 1 , che è la minima probabile incertezza di una quantità discreta, con-

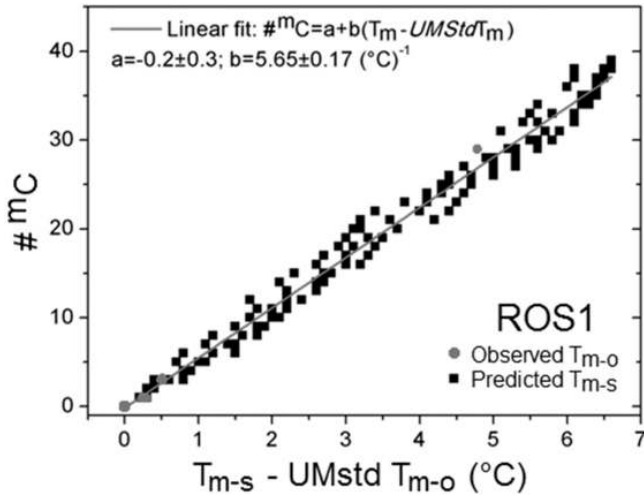


Figura 3. Valori T_{m-s} del gene *ROS1* ottenuti in silico (quadrati neri) e valori T_{m-o} ottenuti con reazioni HRMA (cerchi grigi) impiegati per il calcolo di $\#mC$ in funzione di $(T_m - UMstd T_m)$. La linea grigia è il risultato della procedura di correlazione lineare che ha generato l'equazione $[\#mC = a + b(T_m - UMstd T_m)]$ (vedi testo). I valori dei parametri della relazione sono: $a = -0,2 \pm 0,3$; $b = 5,65 \pm 0,17(^\circ C)^{-1}$; con $R = 0,991$.

Fig. 3 *In silico* T_{m-s} values (black squares) and *in tube* T_{m-o} values (grey circles) for $\#mC$ as a function of $(T_m - UMstd T_m)$ for *ROS1*. The grey line is the result of the linear fitting procedure of *in silico*-predicted data $[\#mC = a + b(T_m - UMstd T_m)]$ (see text). Fitting parameters: $a = -0.2 \pm 0.3$, $b = 5.65 \pm 0.17(^\circ C)^{-1}$ with $R = 0.991$.

dei siti mC , ma offre l'opportunità di stimare lo stato di metilazione complessivo della regione di un gene attraverso una semplice e veloce analisi, rilevando e risolvendo campioni in cui è presente un'eterogeneità di metilazione, spesso difficile da gestire con approcci di sequenziamento del DNA trattato con bisolfito.

L'epigenetica descrive i cambiamenti ereditabili nella funzione genica che si verificano indipendentemente dalla sequenza del DNA. La facilità e la rapidità di utilizzo di EpiHRMAssay, rispetto ad altri metodi, lo rendono uno strumento rapido e facile per rilevare il grado di metilazione in sequenze sconosciute di campioni di piante. EpiHRMAssay è adatto per l'analisi della metilazione della regione di DNA associabili ad un determinato gene, per identificare regioni sensibili alla metilazione sconosciute, indipendentemente dalla conoscenza iniziale del loro stato di metilazione. EpiHRMAssay trova impiego nella diagnosi di metilazione delle piante propagate, contribuendo a definire l'identità genetica e la tracciabilità.

Riassunto

EpiHRMAssay combina l'analisi High Resolution Melting Analysis, su DNA trattato con bisolfito, con un sistema previsionale *in silico* delle temperature di melting del DNA di epialleli virtuali, ottenute con il software uMELTSM, e propone un nuovo approccio di quantificazione della metilazione e delle mutazioni

fermando così l'affidabilità delle predizioni di modelli (Cirilli *et al.*, 2017).

Discussione e conclusioni

La procedura non individua la posizione specifica

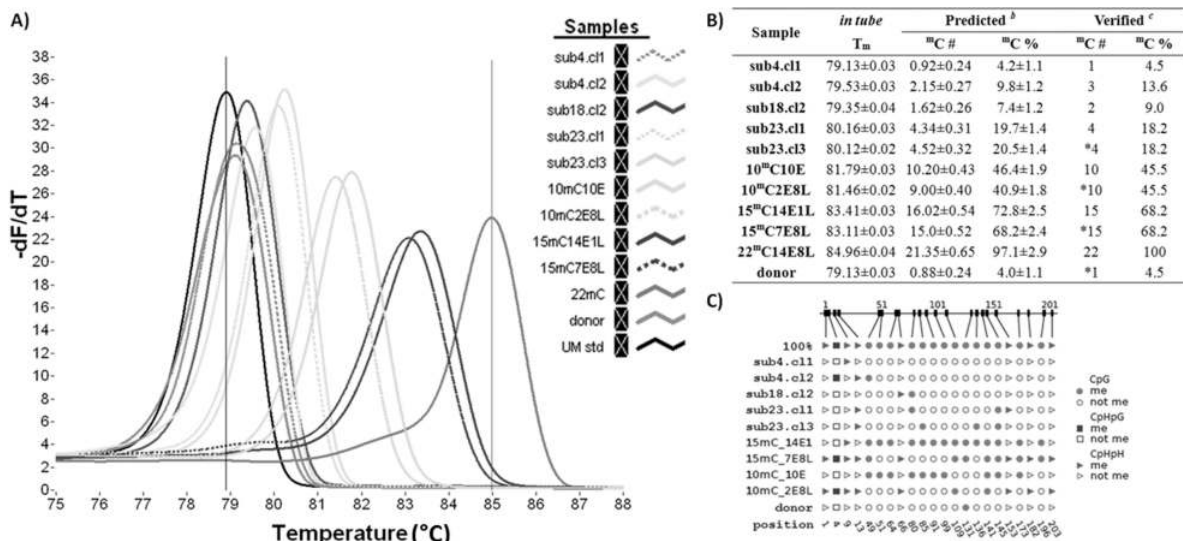


Fig. 4 - Derivative melting curve che evidenziano il valore T_{m-o} , corrispondente al picco della curva, di ciascun amplicone del gene *ROS1* (A). Nella tabella (B) sono riportati i dati generati dal confronto tra i numeri e le percentuali di mC previsti e quelli osservati negli ampliconi sequenziati.

Fig. 4 - Derivative melting curves that each show the T_{m-o} value corresponding to the curve peak for each *ROS1* amplicon (A). In the table shown in panel B are reported the data generated from the comparison between the numbers and percentages of the predicted mC and those observed in sequenced amplicons.

del DNA. Esso è un test affidabile, conveniente e di elevata sensibilità atto a rilevare variazioni nel DNA di piante propagate agamicamente (*in vitro* e non) per verificare la stabilità genetica ed epigenetica e garantire la tracciabilità dell'identità genetica e fenotipica, ed è applicabile anche ai fini della ricerca scientifica.

Parole chiave: High Resolution Melting Analysis (HRMA), metilazione del DNA, predizione *in silico* della temperatura di *melting*, regolazione epigenetica, variazione somaclonale.

Bibliografia

- BECKER C., HAGMANN J., MÜLLER J., KOENIG D., STEGLE O., BORWARDT K., WEIGEL D., 2011. *Spontaneous epigenetic variation in the Arabidopsis thaliana methylome*. Nature, 480: 245-249.
- BELL J.T., PAI A.A., PICKRELL J.K., GAFFNEY D.J., PIQUE-REGI R., DEGNER J.F., GILAD Y., PRITCHARD J.K., 2011. *DNA methylation patterns associate with genetic and gene expression variation in HapMap cell lines*. Genome Biol., 12:R10.
- CASSELLS A.C., CURRY R.F., 2001. *Oxidative stress and physiological, epigenetic and genetic variability in plant tissue culture: implications for micropropagators and genetic engineers*. Plant Cell Tiss. Organ Cult., 64: 145-157.
- CIRILLI M., DELFINO I., CABONI E., MULEO R., 2017. *EpiHRMAssay, an in tube and in silico combined approach for the scanning and epityping of heterogeneous DNA methylation*. Biol. Methods Protoc., 2: 1-11.
- DWIGHT Z., PALAIS R., WITTEWIT C.T., 2011. *uMELT: prediction of high-resolution melting curves and dynamic melting profiles of PCR products in a rich web application*. Bioinformatics, 27: 1019-1020.
- GENTILE A., JÁQUEZ GUTIÉRREZ M., MARTINEZ J., FRATTARELLI A., NOTA P., CABONI E., 2014. *Effect of meta-Topolin on micropropagation and adventitious shoot regeneration in Prunus rootstocks*. Plant Cell, Tiss Organ Cult., 118:373-381.
- HUGUET J.M., BIZARRO C.V., FORNS N., SMITH S.B., BUSTAMANTE C., RITORT F., 2010. *Single-molecule derivation of salt dependent base-pair free energies in DNA*. Proc. Natl. Acad. Sci. USA, 107: 15431-15436.
- MIGUEL C., MARUM L., 2011. *An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond*. J. Exp., Bot., 62: 3713-3725.
- 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.
- VERDE I., ABBOTT A.G., SCALABRIN S., JUNG S., SHU S., MARRONI, F., ET AL., *International Peach Genome Initiative*. 2013. *The high-quality draft genome of peach Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution*. Nat. Genet., 45 (5): 487-494.
- WEBER G., 2015. *Optimization method for obtaining nearest-neighbour DNA entropies and enthalpies directly from melting temperatures*. Bioinformatics, 3: 871-877.
- WHITE H.E., HALL V.J., CROSS N.C.P., 2007. *Methylation-sensitive high-resolution melting-curve analysis of the SNRPN gene as a diagnostic screen for Prader-Willi and Angelman syndromes*. Clin. Chem., 53(11): 1960-1962.
- ZILBERMAN D., GEHRING M., TRAN R.K., BALLINGER T., HENIKOFF S., 2007. *Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription*. Nat. Genet., 39: 61-69.