Effect of cryopreservation and post-cryopreservation somatic embryogenesis on the epigenetic fidelity of cocoa (*Theobroma cacao* L.)

Adu-Gyamfi, Raphael[1], Wetten, Andy[2] and Rodríguez López, Carlos Marcelino[3*]

[1]Faculty of Agriculture, University for Development Studies, P.O. Box TL 1882, Tamale, Ghana.

[2]School of Agriculture, Policy and Development, University of Reading, Whiteknights, Reading, Berks, UK RG6 6AS. [3]Plant Research Centre, School of Agriculture, Food and Wine, Faculty of Sciences, University of Adelaide, Waite Campus, PMB1 Glen Osmond.SA 5064. Australia.

*Corresponding author:

Email: carlos.rodriguezlopez@adelaide.edu.au

Phone: +61 450183372

1

Abstract

While cocoa plants regenerated from cryopreserved somatic embryos can demonstrate high levels of phenotypic variability, little is known about the sources of the observed variability. Previous studies have shown that the encapsulation-dehydration cryopreservation methodology imposes no significant extra mutational load since embryos carrying high levels of genetic variability are selected against during protracted culture. Also, the use of secondary rather than primary somatic embryos has been shown to further reduce the incidence of genetic somaclonal variation. Here, the effect of *in vitro* conservation, cryopreservation and post-cryopreservation generation of somatic embryos on the appearance of epigenetic somaclonal variation were comparatively assessed. To achieve this we compared the epigenetic profiles, generated using Methylation Sensitive Amplified Polymorphisms, of leaves collected from the ortet tree and from cocoa somatic embryos derived from three *in vitro* conditions: somatic embryos, somatic embryos cryopreserved in liquid nitrogen and somatic embryos generated from cryoproserved somatic embryos. Somatic embryos accumulated epigenetic changes but these were less extensive than in those regenerated after storage in LN. Furthermore, the passage of cryopreserved embryos through another embryogenic stage led to further increase in variation. Interestingly, this detected variability appears to be in some measure reversible. The outcome of this study indicates that the cryopreservation induced phenotypic variability could be, at least partially, due to DNA methylation changes.

Key words: somatic embryos, epigenetic, somaclonal variation, Methylation Sensitive Amplified Polymorphisms, MSAP, cryopreservation, *Theobroma cacao*, cocoa

Key message: Phenotypic variability observed in cryostored cocoa somatic-embryos is epigenetic in nature. This variability is partially reversible, not stochastic in nature but a directed response to the *in-vitro* culture and cryopreservation.

Introduction

The propagation of plant material through *in vitro* culture can lead to marked increases in the frequency of variants (Phillips et al.. 1994). These enhanced frequencies can be higher than those associated with the use of mutagens (Burg et al.. 2007) and have been termed somaclonal variation (Côte et al. 2001). The nature of such variation can be genetic (altering the DNA sequence of the ramets) and/or epigenetic (which does not affect the DNA sequence but affects its chemistry and structure, gene expression and may ultimately induce some form of phenotypic abnormality (Joyce and Cassells 2002). The addition of a methyl group to cytosine residues (cytosine methylation) is probably the most studied feature of epigenetic regulation in Eukaryotes. In plants, it occurs in three sequence contexts: CG, CNG, or CNN (in order of relative abundance) (N = any nucleotide other than G (Rodríguez López and Wilkinson 2015)). Cytosine methylation occurring within promoters or coding regions typically acts to repress gene transcription by changing local chromatin structure (Davey et al. 1997), thereby preventing the binding of DNA–binding proteins to the promoter regions (Inamdar et al. 1991) or as a binding cue for transcriptional repressions (Kass et al. 1997).

Different factors have been reported to affect the level of both genetic and epigenetic somaclonal variation induced during *in vitro* culture, including: **1.** medium composition (Schmitt et al. 1997; Bardini et al. 2003). **2.** The propagation technique used, with methods involving a dedifferentiated callus phase expected to show higher levels of variability (Miguel and Marum 2011). **3.** Origin of the donor tissue, with regenerants maintaining epigenetic features of the explant tissue used for propagation (Rodríguez López et al. 2010b). **4.** Time in culture and number of regeneration events (Rodríguez López et al. 2010b). Somaclonal variation generally increases with prolonged callus phase and with the number of multiplication cycles (Côte et al. 2001; Peredo et al. 2009). However, Rodriguez López *et al.*. (2010b) observed that older cocoa calli, as well as exhibiting reduced embryogenic potential, yielded somatic embryos (SE) containing less genetic and epigenetic aberrations. They hypothesised that totipotent cell lineages with few or no mutations are selected

3

during protracted cultures. In cocoa, primary somatic embryogenesis has previously been shown to induce high levels of chimeric (Rodríguez López et al. 2004) and homogenous (Rodríguez López et al. 2010a; Rodríguez López et al. 2010b) genetic and epigenetic mutants. Nevertheless, in the same studies fewer genetic variants were detected among secondary somatic embryos (SSEs) than among primary somatic embryos (PSEs). These findings are in line with Fang *et al.*., (2009) who reported an absence of mutations in the SSE population of cocoa studied probably due to SSEs being derived directly from epidermal cells rather than from dedifferentiated callus.

Due to the difficulty in generating large quantities of cocoa plants via cuttings and to the recalcitrant nature of cocoa seed with regard to low temperature storage, cryopreservation of tissue-cultured germplasm represents the most attractive backup for vulnerable field collections of the species. The lengthy culture periods associated with the establishment of embryogenic cocoa callus lines also mean that it will be a wise precaution to back up large *in vitro* germplasm collections with cryopreserved cell lines. Different groups have shown that the encapsulation-dehydration (Fang et al. 2009) and cryopreservation of plant material (Wilkinson et al. 2003) do not impose a significant extra genetic mutational load to the plants regenerated *in vitro*. However, little is known about how the cryopreservation methodology and subsequent propagation cycles affect the fidelity of the regenerant plants from an epigenetic perspective.

Methylation-sensitive amplified polymorphism (MSAP) is a modification of Amplified Fragment Length Polymorphism that makes use of differential sensitivity of certain restriction endonucleases to cytosine methylation to study the level and global patterns of methylation across the target genomes (Reyna-López et al. 1997). MSAP employs a pair of isoschizomer enzymes (*Hpa*II and *Msp*I) with differential sensitivity to methylation in their recognition sequence. In short, *Hpa*II is

4

inactive (does not cut its recognition site, 5'-CCGG-3') if one or both cytosines are methylated on both DNA strands, but cleaves when all cytosines are unmethylated or if one or both cytosines are methylated in only one strand. *Msp*I is, in contrast, considered by convention methylation to be insensitive and cleaves if all cytosines are unmethylated, and if the internal cytosine is methylated but not when the external cytosine is methylated. MSAP analysis has been previously used for the detection of tissue culture induced epigenetic changes in a large number of plant taxa (Alonso et al. 2015), including important crops such as: tobacco (Schmitt et al. 1997), rice (Xiong et al. 1999), oil palm (Matthes et al. 2001), barley (Bednarek et al. 2007), pea (Smýkal et al. 2007), potato (Dann and Wilson 2011), cocoa (Rodríguez López et al. 2010b), grapevine (Baránek et al. 2015), cassava (Kitimu et al. 2015), and garlic (Gimenez et al. 2015). In this study we used the MSAP technology to determine the effect of cryopreservation and post-cryopreservation regeneration of tertiary SEs (TSEs) on the appearance of somaclonal variation compared to equivalent variation in long term *in vitro* conservation of cocoa SSEs.

Materials and Methods

Plant Material

All the SEs used in this study were generated from a single AMAZ 15 cocoa tree held in the International Cocoa Quarantine Centre, Reading, UK, with somatic embryogenesis initiated from staminode cultures as described in Maximova et al. (2002). Three groups of propagules plus a set of reference samples from the ortet tree were compared: **1** SSEs maintained in ED medium (Maximova et al. 2002) culture for the 357 d duration of the experiment ('*in vitro*' hereafter); **2** SSEs recovered after 1 h storage in liquid nitrogen (LN) using vitrification-based cryopreservation (Adu-Gyamfi and Wetten 2012) ('1 h LN' hereafter); **3** TSEs regenerated from 1 h LN cotyledon explants following (Maximova et al. 2002) ('post 1 h LN' hereafter) and **4** the ortet tree reference set were 6 mm discs

5

obtained from 24 randomly selected recently expanded leaves from the donor tree. All samples were kept at -20°C until DNA extraction (see Table 1 and Figure 1).

DNA extraction

Genomic DNA was extracted from leaves of somatic embryos and donor plant using the DNeasy 96 plant kit (Qiagen, UK) following manufacturer's instructions and eluted in 50 µl Buffer AE (Qiagen). Two replicate extractions were performed for each sample. The DNA was quantified on a NanoDrop^{TM} 2000c (ThermoLifescience). Extracted DNAs were diluted with nanopure water to produce working stocks of 10 ng/µl.

MSAP procedure

A modification of the original MSAP method, as described by Rodríguez López et al. (2012), was used on all DNA extractions. In short, the method consists of the parallel digestion of genomic DNA with two methylation-sensitive isoschizomers (*Msp*I and *Hpa*II) as frequent cutters, each in combination with the same rare cutter (*Eco*RI), adaptor ligation, followed by two selective PCR amplifications with primers complementary to the adaptors but with unique 3' overhangs (Table 2). Thirty-two primer combinations consisting of eight +3*Eco*RI and four +2*Hpa*II/*Msp*I combinations were evaluated with a subset of 8 randomly selected pre-selective amplifications containing two samples from each group described in Table 1. The evaluation was done to assess the level of intra-treatment variation for each primer combination and their ability to generate informative and consistent MSAP profiles. Based on these results the best two primer combinations (*Hp1/Eco1* and *Hp3/Eco3*) were designated as primers for the selective amplification of 96 samples from the four experimental groups. Resultant PCR products labelled with FAM fluorescent dye were diluted 1/10 in sterile nano water and 1 µl was combined with 1 µl ROX/HiDi mix (50 µl ROX plus 1 ml HiDi formamide). Samples were then denatured by heating at 95 °C for 5 min, snap-cooled on ice for 2

min and run on an ABI PRISM 3100 Genetic Analyzer Capillary Sequencer (16 capillary array model) at 3 kV for 22 s and 15 kV for 45 min at 60 $^{\circ}$C.

Data analysis

GeneMapper® Software Version 4.0 was used to generate a binary data matrix of bands present (1) or absent (0) on the MSAP profiles from samples restricted using *Eco*RI and *Hpa*II or *Msp*I and amplified using the different primer combinations shown in Table 1. Only epiloci ranging from 100 to 500bp in size were considered for analysis. Fragments present/absent in all but one individual were considered uninformative and removed from all data sets. GenAlex (Peakall and Smouse 2012) 6.5 software was used to analyse the binary data matrix as described in Rodriguez Lopez et al. (2012). In brief, first the frequency of all epiloci was calculated for each group. Epiloci were considered unique to *in vitro* culture samples (*in vitro*, 1 h LN and/or Post 1 h LN) when the their frequency was higher than 75% in the donor plant samples and lower than 25% in the *in vitro* samples and *vice versa*. Equally, epiloci were considered unique to cryopreserved samples (1 h LN and/or Post 1 h LN) when their frequency was higher than 75% in the 1h LN samples and lower than 25% in the *in vitro* samples and *vice versa*. Finally, epiloci were considered as induced by *in vitro* culture or cryopreservation when their frequencies where higher than 60% in the donor plant/1h LN samples and lower than 30% in the *in vitro* samples and *vice versa*.

Next using GenAlex (v.6.4), the epigenetic variability between ortets (samples from donor plant) and ramet groups (samples generated under different *in vitro* conditions, Table 1) was visualized by Principal Coordinate Analysis (PCoA) based on the amalgamation of the MSAP profiles obtained from samples restricted with *Hpa*II or *Msp*I and amplified with primer combinations Hp1/Eco1 and

Hp3/Eco3Analysis of Molecular Variance (AMOVA) was then used to estimate and test the significance of the epigenetic diversity within and between the different groups. Pairwise PhiPT comparisons between each group based on 10000 permutations was used to infer the overall level of divergence in DNA methylation between groups. Finally, the level of epigenetic variability introduced by each of the treatments was estimated using two values from the AMOVA: **1** calculated PhiPT between ortet tree samples and samples from each treatment (*in vitro*, 1h LN, Post 1h LN) (i.e., the lower the PhiPT, the lower the level of DNA methylation variability introduced by that treatment), **2** The mean sum of squares within population (SSWP) was used to infer epigenetic variation within treatments (the higher the mean SSWP, the higher the level of epigenetic variability induced by that treatment) (Amoah et al. 2012).

Results

Effect of conservation methods on somatic embryo methylation patterns

To analyse the effect of the different *in vitro* conservation methods on the DNA methylation patterns of cocoa SE we compared the MSAP profiles of samples from the donor plant (AMAZ 15) to those obtained from SE conserved as described in Table 1. A total of 220 fragments ranging from 100bp to 500bp were generated by primer combinations Hp1/Eco1 and Hp3/Eco3 (108 and 112 fragments respectively). Of these, 3.6% (8 epiloci) and 8.6% (19 epiloci) were deemed unique or induced by cryopreservation and *in vitro* culture respectively (Table 3). In total 27 *in vitro* culture and cryopreservation markers were obtained, 13 using *Hpa*II and 14 using *Msp*I (Table 3). Of these markers, the frequency of epialleles *Hpa*II 1.1 (232), *Hpa*II 1.1 (233), *Hpa*II 3.3 (168), *Msp*I 3.3 (303) and *Msp*I 3.3 (407) showed a correlation with the accumulation of *in vitro* procedures (i.e. number of somatic embryogenesis and cryopreservation events) (Table 3, Supplementary Figure 1A-E). In contrast, epialleles *Hpa*II 1.1 (367), *Hpa*II 3.3 (181), *Hpa*II 3.3 (142), *Msp*I 3.3 (202) and *Msp*I 1.1 (278) showed a change in frequency associated with somatic embryogenesis/cryopreservation

8

followed by a reversion to donor tree allele frequency levels after cryopreservation or after tertiary embryogenesis (Table 3 and Supplementary Figure 1F-J).

AMOVA of profiles generated using *Hpa*II showed that 19% of the variation was explained by differences between groups while 81% was due to individual differences. Similarly, 21% of the variability detected by *Msp*I was explained by differences between groups and 79% was due to individual differences. Principal Coordinate Euclidean Analysis (PCoA) was used to provide an overview of the epigenetic variability introduced by somatic embryogenesis and cryopreservation. Overall, profiles generated using both *Msp*I and *Hpa*II provided a clear separation between donor plant samples and all SE samples (Figure 2A). However, the separation observed between the different ramet groups was slightly clearer when using profiles from samples restricted with *Msp*I (Figure 2B and C). PhiPT values confirmed the epigenetic variation observed between groups using PCoA analysis, with groups presenting higher levels of epigenetic divergence when *Msp*I profiles were used (Table 4). The largest differences between groups were detected between donor plant samples and cryopreserved samples (i.e. 1h LN and Post-1h LN) whilst the lowest levels of epigenetic divergence between groups were detected between both cryopreserved samples (i.e. 1h LN and Post-1h LN) restricted using *Hpa*II. AMOVA analysis using 10000 permutations showed that all pairwise distances were significant (P>0.0002). When PhiPT values generated from *Hpa*II and *Msp*I MSAP profiles were taken together, Post-1h LN samples were the most epigenetically divergent from the donor samples (Figure 3A). Analysis of epigenetic variance within groups estimated using the mean SSWP showed similar levels of variability amongst donor, *in vitro* and Post-1h LN. However, higher levels of variability were detected between 1h LN samples (Figure 3B).

Discussion

Somaclonal variation among conserved somatic embryos

Conventional propagation of elite cocoa clones via cuttings has been hampered by costs and undesired phenotypes associated with this type of propagule. Moreover, the long term conservation of cocoa is hindered by high levels of variability in agronomic performance and the poor low temperature storage ability of cocoa seeds (Li et al. 1998). For these reasons considerable effort has been put into the development of *in vitro* propagation and cryopreservation systems (for an extensive review see Guiltinan 2007). However, the passage of cells through *in vitro* culture (Miguel and Marum 2011) and cryopreservation (Kaity et al., 2008) may lead to undesirable changes (somaclonal variation) in the regenerants. Although the effect of *in vitro* culture, somatic embryogenesis and cryopreservation of somatic embryos on the morphological and genetic fidelity of cocoa plants has been extensively studied (Rodríguez López et al. 2004; Fang et al. 2009; Wetten et al. 2009; Rodríguez López et al. 2010a; Rodríguez López et al. 2010b), very little is known about the epigenetic modifications introduced by the cryopreservation of cocoa somatic embryos.

Not surprisingly, our results showed an increase in variability in all *in vitro* cultured samples (*in vitro*, 1h LN, Post 1h LN) (Figures 2 and 3A). As we have shown previously, this variability is both genetic and epigenetic in nature (Rodríguez López et al. 2004; Fang et al. 2009; Wetten et al. 2009; Rodríguez López et al. 2010a; Rodríguez López et al. 2010b). More interestingly, SSE recultured after one hour in liquid nitrogen (1h LN) showed: **a)** an increase in genetic/epigenetic distance from the donor samples when compared to the PhiPT between donor samples and non-cryopreserved SSE (*in vitro*) (Figure 3A), and **b)** higher levels of within group genetic/epigenetic variability detected using *Msp*I generated profiles (mean $SSWP_{MspI} = 278$) compared to both donor plant samples (mean $SSWP_{MspI} = 237$) and to non-cryopreserved somatic embryos (mean $SSWP_{MspI} = 229$) (Figure 3B). High frequencies of morphological abnormalities in cryostored and non-cryostored cocoa somatic

10

embryos have been reported (Alemanno et al. 1996; Alemanno et al. 1997; Tan and Furtek 2003) suggesting that both somatic embryogenesis and cryopreservation can induce somaclonal variability. Evidence from other plant species such as papaya (Kaity et al. 2008) and chrysanthemum (Martín et al. 2011) suggests that such variation is both genetic and epigenetic (i.e. DNA methylation) in nature. However, our previous work has shown that the genetic variability attributable to the encapsulation-dehydration and vitrification cryopreservation of cocoa SE methodologies used here (Fang et al. 2004; Adu-Gyamfi and Wetten 2012) is negligible (Fang et al. 2009; Wetten et al. 2009). We, therefore, speculate that this observed increased within group variability is mainly epigenetic in nature, as previously shown in somatic embryo-derived oil palm (Jaligot et al. 2000) and cryopreserved hop plants (Peredo et al. 2009). Moreover, this higher level of within group variability observed in cryopreserved samples (1 h LN) supports the hypothesis that a large fraction of stress (cold, dehydration, etc) induced epigenetic variation occurs randomly across the genome (Vogt 2015; Soen et al. 2015).

Interestingly, TSE regenerated from cryopreserved SSE (Post 1h LN) showed intermediate within group genetic/epigenetic variability (mean $SSWP_{MspI} = 264$). This would indicate that the cryopreservation induced variability observed in 1h LN samples is partially reversible, supporting the hypothesis of it being epigenetic. Conversely, tertiary somatic embryos in the Post 1h LN group showed a significantly higher PhiPT distance from donor samples than that observed for SSE (*in vitro* and 1h LN samples), suggesting that this extra somatic embryogenesis event does introduce additional genomic variability. Analysis of scanning electron microscopy images have shown that TSE after cryopreservation generally regenerate directly from SSE epidermal cells (Fang *et al.,* 2009) without the intermediate callus phase which is normally associated with higher levels of somaclonal variation. Nonetheless, epigenetic somaclonal variation is not limited to the passage

11

through a dedifferentiated phase as shown in micropropagated cassava (Kitimu et al. 2015). Furthermore, this increase in variation following another step of somatic embryogenesis is in line with a number of studies on *in vitro* culture variation which report that variation increases with the number of multiplication cycles (Côte et al. 2001; Peredo et al. 2009). However, we have also shown that successive rounds of somatic embryogenesis in cocoa reduce the detected levels of genetic variability (Fang *et al.*., 2009). It is therefore tempting to speculate that this increase in PhiPT observed using MSAPs is also mainly epigenetic in nature. A factor possibly responsible for this elevated epigenetic variability might be the presence of growth regulators (2,4-D and 6-BA at 1 mg/l and 50 µl /l respectively) used during the induction of TSE, which have been reported to contribute to epi-mutations in tissue culture derived materials (Chakrabarty et al. 2003).

Finally, it is widely accepted that in general the majority of the developmental epigenetic variability occurs randomly across the genome (Vogt 2015; Soen et al. 2015). It would therefore be expected that the accumulation of stochastic variation during successive somatic embryogenesis events would increase the observed differences between samples of the same group. If this was the case, then non cryopreserved SSEs should present higher levels of within group variability than donor samples but lower than TSE. However, we observed similar levels of within group variability, measured here as mean SSWP, among all compared groups (donor, in vitro, and post 1h LN) suggesting that the somatic embryogenesis-induced variability detected here is not stochastic but probably resulted from a directed adaptation to the successive growing environments to which SEs are exposed.

The results shown here support previous evidence (Fang et al. 2009) that the high levels of phenotypic variability observed in cryostored SE may be symptomatic of epigenetic change. More importantly, our results suggest that this observed variability is not necessarily stochastic in nature,

but might be partially a response to the environmental stresses that plant cells are exposed to during *in vitro* culture and cryopreservation. One would expect that the main contributor to this observed variability would be low temperature. In fact RNA directed DNA methylation is upregulated by low temperature (Tuttle et al. 2008). However previous work has shown that low temperature alone does not explain all the variability observed during cryopreservation. For example, the use of the cryoprotectant DMSO, which chelates to nucleic acids, as a cryoprotectant might not only be introducing genetic changes, as shown previously (Aronen et al. 1999) but may be potentially causing numerous DNA methylation changes. What is more, (Harding 1994) reported that changes on methylation of DNA sequences may be an adaptive response to conditions of high osmotic stress induced by the used of high concentration of sucrose during vitrification. Understanding how these factors affect the fidelity of the regenerant plants will ultimately help in the development of new cryopreservation methodologies with lower levels of phenotypic, genetic and epigenetic variation.

Author contribution statement

A.W. designed and supervised the project; R.A.-G. generated the somatic embryos. R.A.-G. and C.M.R.L. carried out the MSAPs and analysed their results; All authors discussed the results, wrote the manuscript and read and approved its final version.

Conflict of interest

The authors declare no conflict of interests.

Bibliography

Adu-Gyamfi R, Wetten A (2012) Cryopreservation of cocoa (Theobroma cacao L.) somatic embryos by vitrification. Cryo Letters 33:494–505.

Alemanno L, Berthouly M, Michaux-Ferrière N (1996) Histology of somatic embryogenesis from floral tissues cocoa. Plant Cell, Tissue and Organ Culture (PCTOC) 46:187–194. doi: 10.1007/BF02307094

Alemanno L, Berthouly M, Michaux-Ferriere N (1997) A comparison between *Theobroma cacao* L. zygotic embryogenesis and somatic embryogenesis from floral explants. In Vitro CellDevBiol-Plant 33:163–172. doi: 10.1007/s11627-997-0016-8

Alonso C, Pérez R, Bazaga P, et al. (2015) MSAP markers and global cytosine methylation in plants: a literature survey and comparative analysis for a wild-growing species. Mol Ecol Resour. doi: 10.1111/1755-0998.12426

Amoah S, Kurup S, Rodriguez Lopez CM, et al. (2012) A hypomethylated population of *Brassica rapa* for forward and reverse epi-genetics. BMC Plant Biol 12:193. doi: 10.1186/1471-2229-12-193

Aronen TS, Krajnakova J, Häggman HM, Ryynänen LA (1999) Genetic fidelity of cryopreserved embryogenic cultures of open-pollinated *Abies cephalonica*. Plant Sci 142:163–172. doi: 10.1016/S0168-9452(98)00244-1

Baránek M, Čechová J, Raddová J, et al. (2015) Dynamics and reversibility of the DNA methylation landscape of grapevine plants (*Vitis vinifera*) stressed by in vitro cultivation and thermotherapy. PLoS ONE 10:e0126638. doi: 10.1371/journal.pone.0126638

Bardini M, Labra M, Winfield M, Sala F (2003) Antibiotic-induced DNA methylation changes in calluses of *Arabidopsis thaliana*. Plant Cell, Tissue and Organ Culture 72:157-162

Bednarek PT, Orłowska R, Koebner RM, Zimny J (2007) Quantification of the tissue-culture induced variation in barley (*Hordeum vulgare* L.). BMC Plant Biol 7:10. doi: 10.1186/1471-2229-7-10

Burg K, Helmersson A, Bozhkov P, von Arnold S (2007) Developmental and genetic variation in nuclear microsatellite stability during somatic embryogenesis in pine. J Exp Bot 58:687–698. doi: 10.1093/jxb/erl241

Chakrabarty D, Yu KW, Paek KY (2003) Detection of DNA methylation changes during somatic embryogenesis of Siberian ginseng (Eleuterococcus senticosus). Plant Sci 165:61–68. doi: 10.1016/S0168-9452(03)00127-4

Côte FX, Teisson C, Perrier X (2001) Somaclonal variation rate evolution in plant tissue culture: Contribution to understanding through a statistical approach. In Vitro CellDevBiol-Plant 37:539–542. doi: 10.1007/s11627-001-0093-z

Dann AL, Wilson CR (2011) Comparative assessment of genetic and epigenetic variation among regenerants of potato (*Solanum tuberosum*) derived from long-term nodal tissue-culture and cell selection. Plant Cell Rep 30:631–639. doi: 10.1007/s00299-010-0983-9

Davey C, Pennings S, Allan J (1997) CpG methylation remodels chromatin structure in vitro. J Mol Biol 267:276–288. doi: 10.1006/jmbi.1997.0899

Fang J, Wetten A, Hadley P (2004) Cryopreservation of cocoa (*Theobroma cacao* L.) somatic embryos for long-term germplasm storage. Plant Sci 166:669–675. doi: 10.1016/j.plantsci.2003.11.002

Fang J-Y, Wetten A, Adu-Gyamfi R, Rodriguez-Lopez C (2009) Use of secondary somatic embryos promotes genetic fidelity in cryopreservation of cocoa (*Theobroma cacao* L.). Agric and Food Science 18:

Gimenez MD, Yañez-Santos AM, Paz RC, et al. (2015) Assessment of genetic and epigenetic changes in virus-free garlic (*Allium sativum* L.) plants obtained by meristem culture followed by in vitro propagation. Plant Cell Rep. doi: 10.1007/s00299-015-1874-x

Guiltinan MJ (2007) Cacao. In: Pua E, Davey MR (eds) Transgenic Crops V. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 497–518

Harding K (1994) The methylation status of DNA derived from potato plants recovered from slow growth. Plant Cell, Tissue and Organ Culture (PCTOC) 37:31–38. doi: 10.1007/BF00048114

Inamdar NM, Ehrlich KC, Ehrlich M (1991) CpG methylation inhibits binding of several sequence-specific DNA-binding proteins from pea, wheat, soybean and cauliflower. Plant Mol Biol 17:111–123.

Jaligot E, Rival A, Beulé T, et al. (2000) Somaclonal variation in oil palm (*Elaeis guineensis* Jacq.): the DNA methylation hypothesis. Plant Cell Reports 19:684–690. doi: 10.1007/s002999900177

Joyce SM, Cassells AC (2002) Variation in potato microplant morphology in vitro and DNA methylation. Plant Cell, Tissue and Organ Culture 70:125–137.

Kaity A, Ashmore SE, Drew RA, Dulloo ME (2008) Assessment of genetic and epigenetic changes following cryopreservation in papaya. Plant Cell Rep 27:1529–1539. doi: 10.1007/s00299-008-0558-1

Kass SU, Pruss D, Wolffe AP (1997) How does DNA methylation repress transcription? Trends Genet 13:444–449. doi: 10.1016/S0168-9525(97)01268-7

Kitimu SR, Taylor J, March TJ, et al. (2015) Meristem micropropagation of cassava (Manihot esculenta) evokes genome-wide changes in DNA methylation. Frontiers in Plant Science. doi:10.3389/fpls.2015.00590

Li Z, Traore A, Maximova S, Guiltinan MJ (1998) Somatic embryogenesis and plant regeneration from floral explants of cacao (*Theobroma cacao* L.) using thidiazuron. In Vitro CellDevBiol-Plant 34:293–299. doi: 10.1007/BF02822737

Martín C, Cervera MT, González-Benito ME (2011) Genetic stability analysis of chrysanthemum (*Chrysanthemum x morifolium* Ramat) after different stages of an encapsulation-dehydration cryopreservation protocol. J Plant Physiol 168:158–166. doi: 10.1016/j.jplph.2010.06.025

Matthes M, Singh R, Cheah S-C, Karp A (2001) Variation in oil palm (*Elaeis guineensis* Jacq.) tissue culture-derived regenerants revealed by AFLPs with methylation-sensitive enzymes. TAG Theoretical and Applied Genetics 102:971–979. doi: 10.1007/s001220000491

Maximova SN, Alemanno L, Young A, et al. (2002) Efficiency, genotypic variability, and cellular origin of primary and secondary somatic embryogenesis of *Theobroma cacao* L. In Vitro CellDevBiol-Plant 38:252–259. doi: 10.1079/IVP2001257

Miguel C, Marum L (2011) An epigenetic view of plant cells cultured in vitro: somaclonal variation and beyond. J Exp Bot 62:3713–3725. doi: 10.1093/jxb/err155

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research--an update. Bioinformatics 28:2537–2539. doi: 10.1093/bioinformatics/bts460

Peredo EL, Arroyo-García R, Revilla MA (2009) Epigenetic changes detected in micropropagated hop plants. J Plant Physiol 166:1101–1111. doi: 10.1016/j.jplph.2008.12.015

Phillips RL, Kaeppler SM, Olhoft P (1994) Genetic instability of plant tissue cultures: breakdown of normal controls. Proc Natl Acad Sci U S A 91:5222–5226.

Reyna-López GE, Simpson J, Ruiz-Herrera J (1997) Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. Mol Gen Genet 253:703–710.

Rodríguez López CM, Wetten AC, Wilkinson MJ (2004) Detection and quantification of in vitro-culture induced chimerism using simple sequence repeat (SSR) analysis in *Theobroma cacao* (L.). Theor Appl Genet 110:157–166. doi: 10.1007/s00122-004-1823-5

Rodríguez López C, Bravo H, Wetten A, Wilkinson M (2010a) Detection of somaclonal variation during cocoa somatic embryogenesis characterised using cleaved amplified polymorphic sequence and the new freeware Artbio. Mol Breeding 25:501–516. doi: 10.1007/s11032-009-9348-x

Rodríguez López CM, Wetten AC, Wilkinson MJ (2010b) Progressive erosion of genetic and epigenetic variation in callus-derived cocoa (*Theobroma cacao*) plants. New Phytol 186:856–868. doi: 10.1111/j.1469-8137.2010.03242.x

Rodríguez López C, Morán P, Lago F, et al. (2012) Detection and quantification of tissue of origin in salmon and veal products using methylation sensitive AFLPs. Food Chem 131:1493–1498. doi: 10.1016/j.foodchem.2011.09.120

Rodríguez López CM, Wilkinson MJ (2015) Epi-fingerprinting and epi-interventions for improved crop production and food quality. Front Plant Sci 6:397. doi: 10.3389/fpls.2015.00397

Schmitt F, Oakeley EJ, Jost JP (1997) Antibiotics induce genome-wide hypermethylation in cultured Nicotiana tabacum plants. J Biol Chem 272:1534–1540.

Smýkal P, Valledor L, Rodríguez R, Griga M (2007) Assessment of genetic and epigenetic stability in long-term in vitro shoot culture of pea (*Pisum sativum* L.). Plant Cell Rep 26:1985–1998. doi: 10.1007/s00299-007-0413-9

18

Soen Y, Knafo M, Elgart M (2015) A principle of organization which facilitates broad Lamarckian-like adaptations by improvisation. Biol Direct 10:68. doi: 10.1186/s13062-015-0097-y

Tan CL, Furtek DB (2003) Development of an in vitro regeneration system for *Theobroma cacao* from mature tissues. Plant Sci 164:407–412. doi: 10.1016/S0168-9452(02)00428-4

Tuttle JR, Idris AM, Brown JK, et al. (2008) Geminivirus-mediated gene silencing from Cotton leaf crumple virus is enhanced by low temperature in cotton. Plant Physiol 148:41–50. doi: 10.1104/pp.108.123869

Vogt G (2015) Stochastic developmental variation, an epigenetic source of phenotypic diversity with far-reaching biological consequences. J Biosci 40:159–204. doi: 10.1007/s12038-015-9506-8

Wetten A, Adu-Gyamfi R, Fang J-Y, Rodriguez-Lopez C (2009) Cryopreservation of cocoa (*Theobroma cacao* L.) somatic embryos - is somaclonal variation an issue? Cryoletters 30:84–85.

Wilkinson T, Wetten A, Prychid C, Fay MF (2003) Suitability of cryopreservation for the long-term storage of rare and endangered plant species: a case history for *Cosmos atrosanguineus*. Ann Bot 91:65–74.

Xiong LZ, Xu CG, Saghai Maroof MA, Zhang Q (1999) Patterns of cytosine methylation in an elite rice hybrid and its parental lines, detected by a methylation-sensitive amplification polymorphism technique. Mol Gen Genet 261:439–446.

Figures and Tables

Table 1: **Description of cocoa samples used for MSAP analysis.**

Conservation method	Treatment description	Duration of treatment (days)			
		Cryo	P-cryo *in vitro* time	P-TSE *in vitro* time	Tissue sampled for MSAP
In vitro	24 SSEs maintained in ED medium	NA	NA	NA	357 d after culture initiation
1 h LN	24 SSEs recovered after 1hour storage in LN	175	182	NA	182 d after cryopreservation and 357 d after culture initiation
Post 1 h LN	24 TSEs obtained from 1 h LN SSE	175	87	95	95 d after TSE initiation, 182 d after cryopreservation and 357 d after culture initiation
Donor	24 Ortet plant leaves	NA	NA	NA	NA

NA: not applicable

Cryo: Number of days since the initiation of somatic embryos from staminodes, prior to samples being subjected to cryopreservation.

P-cryo in vitro time: Number of days that secondary somatic embryos (SSEs) spent in ED medium (Maximova et al. 2002) since they were removed from cryopreservation.

P-TSE in vitro time: Number of days that tertiary somatic embryos (TSEs) spent in ED medium

(Maximova et al. 2002) since the tertiary somatic embryogenesis event.

Table 2: **Sequence of oligonucleotides used for MSAP analysis**. All sequences are shown on the 5' to 3' orientation. 3'selective bases are highlighted in bold. * indicates primers chosen for selective amplification

Oligo name	Function	Sequence
Rv.Ad.HpaII/MspI	HpaII/MspI Reverse Adaptor	GACGATGAGTCTAGAA
Fw.Ad.HpaII/MspI	HpaII/MspI Forward Adaptor	CGTTCTAGACTCATC
Rv.Ad. EcoRI	EcoRI Reverse Adaptor	AATTGGTACGCAGTCTAC
Fw. Ad. EcoRI	EcoRI Forward Adaptor	CTCGTAGACTGCGTACC
Pre. HpaII/MspI	HpaII/MspI preselective primer	GATGAGTCTAGAACGG**T**
Pre.EcoRI	EcoRI preselective primer	GACTGCGTACCAATTC**A**
Hp 1*	HpaII/MspI selective primer	GATGAGTCTAGAACGG**TA**
Hp 2	HpaII/MspI selective primer	GATGAGTCTAGAACGG**TC**
Hp 3*	HpaII/MspI selective primer	GATGAGTCTAGAACGG**TG**
Hp 4	HpaII/MspI selective primer	GATGAGTCTAGAACGG**TT**
Eco 1*	EcoRI selective primer	GACTGCGTACCAATTC**AAA**
Eco 2	EcoRI selective primer	GACTGCGTACCAATTC**AAC**
Eco 3*	EcoRI selective primer	GACTGCGTACCAATTC**AAG**
Eco 4	EcoRI selective primer	GACTGCGTACCAATTC**AAT**
Eco 5	EcoRI selective primer	GACTGCGTACCAATTC**ACA**
Eco6	EcoRI selective primer	GACTGCGTACCAATTC**ACC**
Eco 7	EcoRI selective primer	GACTGCGTACCAATTC**ACG**
Eco 8	EcoRI selective primer	GACTGCGTACCAATTC**ACT**

		Primer (epilocus)	Donor	Invitro	1h LN	Post 1 h LN
Cryopreservation induced	HpaII	1.1 (232)	0%	29%	80%	100%
		1.1 (233)	100%	52%	5%	0%
		1.1 (235)	89%	52%	20%	25%
		1.1 (416)	37%	14%	75%	85%
		3.3 (168)	86%	29%	15%	0%
	MspI	1.1 (232)	0%	0%	75%	79%
		3.3 (157)	91%	48%	25%	25%
		3.3 (202)	19%	29%	70%	48%
In vitro induced	HpaII	1.1 (367)	84%	29%	65%	100%
		1.1 (446)	11%	76%	65%	75%
		3.3 (133)	0%	76%	45%	57%
		3.3 (142)	5%	76%	60%	48%
		3.3 (168)	86%	29%	15%	0%
		3.3 (181)	0%	76%	80%	48%
		3.3 (187)	29%	95%	100%	95%
		3.3 (371)	95%	19%	10%	0%
	MspI	1.1 (278)	15%	71%	65%	0%
		1.1 (319)	90%	24%	25%	0%
		1.1 (447)	15%	86%	80%	89%
		3.3 (188)	70%	10%	20%	5%
		3.3 (223)	78%	29%	70%	15%
		3.3 (226)	87%	62%	50%	25%
		3.3 (288)	74%	24%	45%	5%
		3.3 (303)	4%	52%	55%	75%
		3.3 (371)	91%	0%	5%	0%
		3.3 (407)	100%	71%	55%	30%
		3.3 (472)	70%	14%	45%	5%

| Allele frequency | 0 | 25 | 50 | 75 | 100 |

Table 3. Somatic embryogenesis/cryoprerservation induced epigenetic polymorphisms: MSAP generated epiloci unique to *in vitro* culture samples (*in* vitro, 1 h LN and/or Post 1 h LN) and to cryopreserved samples (1 h LN and/or Post 1 h LN). MSAP profiles were obtained using methylation sensitive isoschizomers *Hpa*II and *Msp*I and primer combinations and *Hp*1/*Eco*1 and *Hp*3/*Eco*3. Column Primer (epilocus) shows primer combination used to generate the specific MSAP product (i.e. 1.1 = *Hp*1/*Eco*1 and 3.3 = *Hp*3/*Eco*3). Columns Donor, Invitro, 1 h LN and Post 1 h LN show the frequency of each epiloci in each of the compared groups (i.e. **Donor:** AMAZ 15 cocoa ortet tree used to regenerate all somatic embryos; **Invitro**: *In vitro* maintained secondary somatic embryos; **1 h LN**: Secondary somatic embryos recovered from after 1 hour storage in liquid nitrogen and **Post 1 h LN**: tertiary somatic embryos generated from 1 h LN samples).

		Donor	Invitro	1 h LN	Post 1h LN
HpaII	Donor	-	0.0001	0.0001	0.0001
	Invitro	0.2232	-	0.0001	0.0002
	1 h LN	0.2798	0.1401	-	0.0001
	Post 1h LN	0.2707	0.0658	0.0778	-
MspI	Donor	-	0.0001	0.0001	0.0001
	Invitro	0.2559	-	0.0001	0.0001
	1 h LN	0.2340	0.0896	-	0.0001
	Post 1h LN	0.3764	0.1206	0.1536	-

Table 4. Somatic embryogenesis and cryopreservation induced epigenetic differences. Epigenetic distances (PhiPT) between groups (i.e. **Donor:** AMAZ 15 cocoa ortet tree used to regenerate all somatic embryos; **Invitro**: *In vitro* maintained secondary somatic embryos; **1 h LN**: Secondary somatic embryos recovered from after 1 h storage in liquid nitrogen and **Post 1 h LN**: tertiary somatic embryos generated from 1 h LN samples) were calculated using Analysis of Molecular Variance (AMOVA) inferred from the analysis of methylation-sensitive amplified polymorphism (MSAP) assays using methylation sensitive isoschizomers *Hpa*II and *Msp*I and primer combinations *Hp*1/*Eco*1 and *Hp*3/*Eco*3. PhiPT Values are shown below diagonal (-). Probability of having a more extreme PhiPT than the observed values by chance alone based on 10,000 permutations is shown above diagonal.
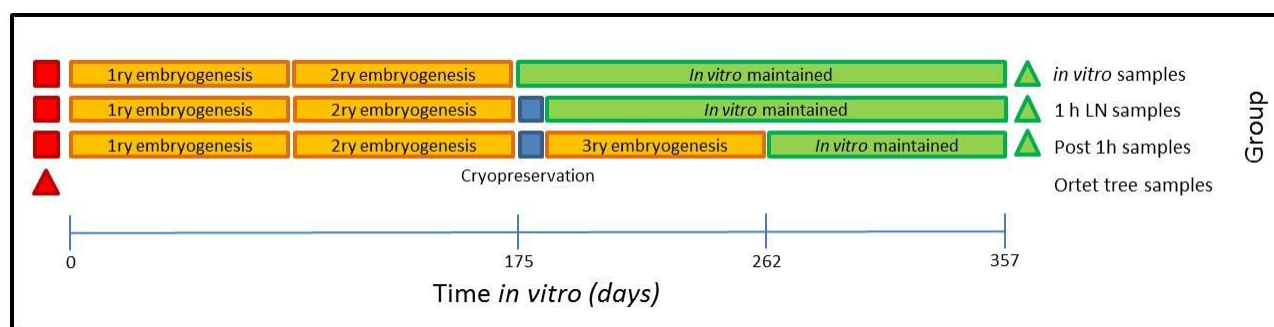
Figure 1: Schematic representation of experimental design. Somatic embryogenesis was initiated from staminodes (red squares) obtained from a single AMAZ 15 cocoa tree from the International Cocoa Quarantine Centre, Reading, UK. Triangles represent 24 samples collected from individual newly expanded leaves for DNA extraction from each group (Donor plant (Red) and somatic embryo derived plants (Green)). Orange boxes represent successive somatic embryogenic events (as described in Maximova et al. 2002). Blue boxes represent a cryopreservation event (1 h in liquid nitrogen) of somatic embryos (as described in Adu-Gyamfi and Wetten, 2012). Green boxes represent maintenance of somatic embryos *in vitro* on ED medium (Maximova et al. 2002). Horizontal bar represents accumulated time spend under *in vitro* conditions by the samples used in this study.

○ Msp Post 1 h LN   ○ Msp 1 h LN   ◓ Msp Invitro   ● Msp Donor   ◇ Hpa Post 1 h LN   ◈ Hpa 1 h LN   ◆ Hpa Invitro   ◆ Hpa Donor
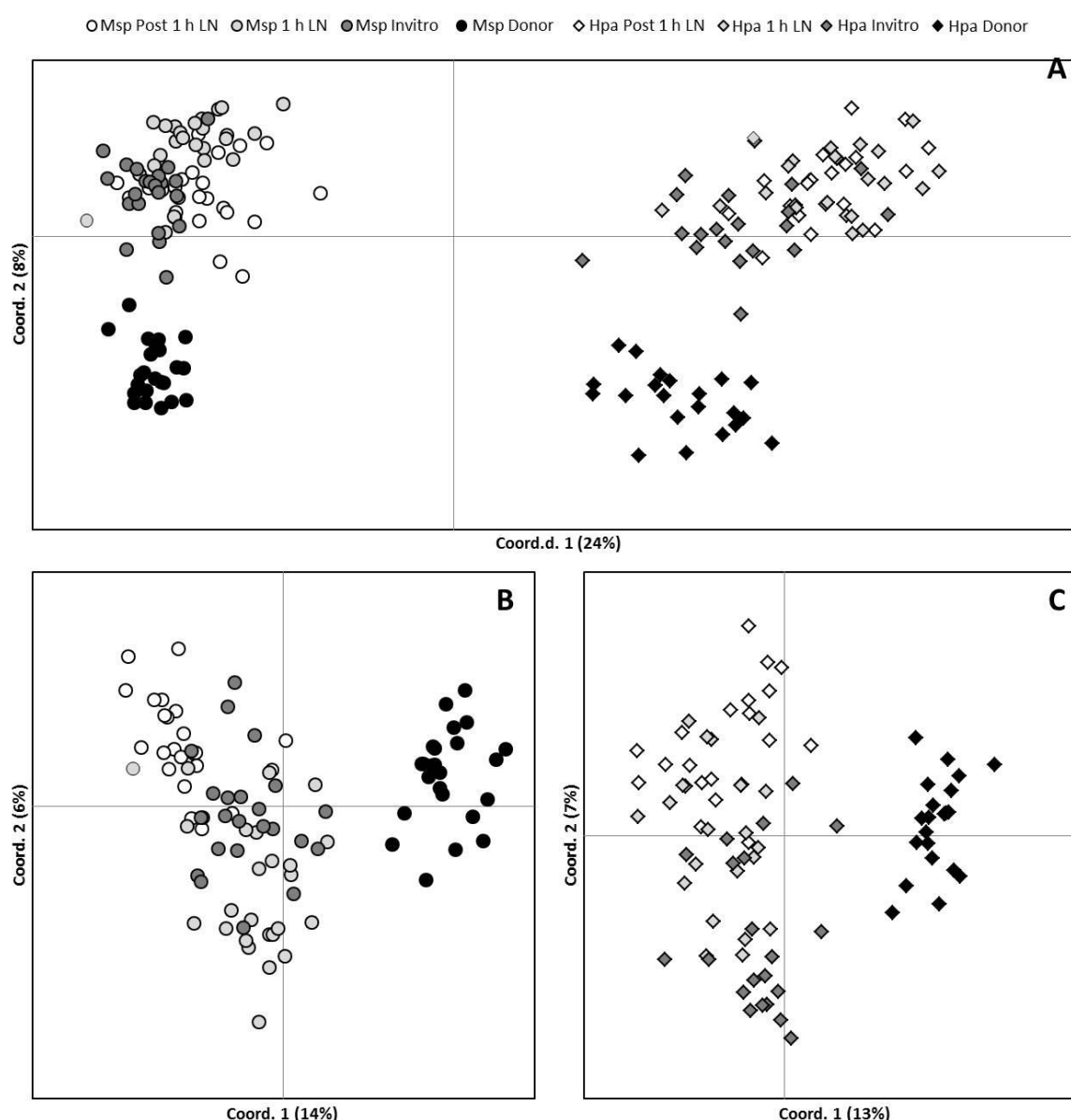
Figure 2: **Effect of conservation method on somaclonal variation pattern in cocoa somatic embryos.** Principal coordinate analysis based on Euclidean Analysis of MSAP distances between 72 somatic embryos grouped by method of conservation and 24 leaf samples of the donor tree. MSAP profiles were obtained using methylation sensitive isoschizimers *Hpa*II (rhomboids) and *Msp*I (circles) and primer combinations and *Hp*1/*Eco*1 and *Hp*3/*Eco*3. 1 h: SEs recovered from secondary SE after 1 hour storage in liquid nitrogen; Post 1 h: tertiary SEs generated from 1 h samples; Invitro: *In vitro* maintained secondary SEs, and Donor: AMAZ 15 cocoa ortet tree used to

26

regenerate SEs. Hp or Msp preceding a treatment means MSAP profiles were generated using restriction enzyme *Hpa*II and *Msp*I respectively.
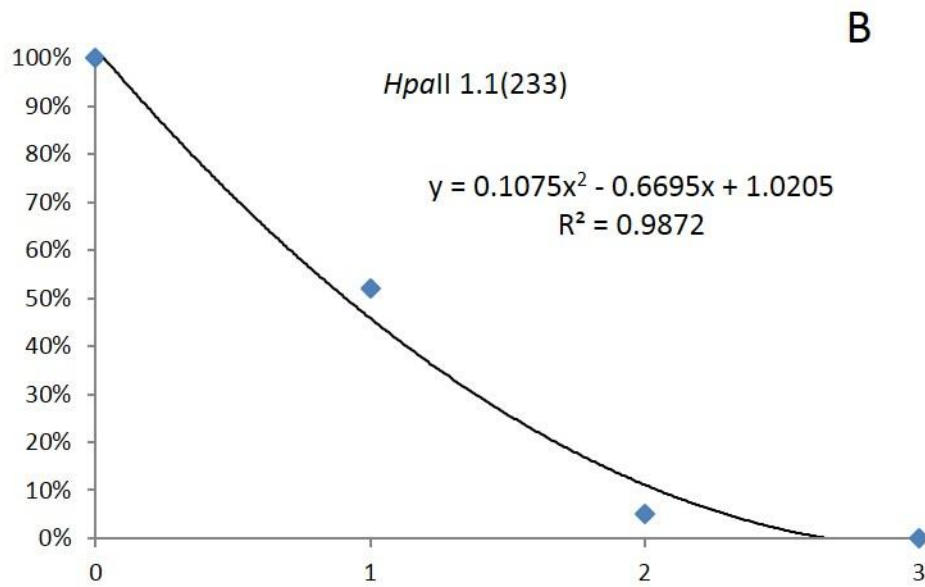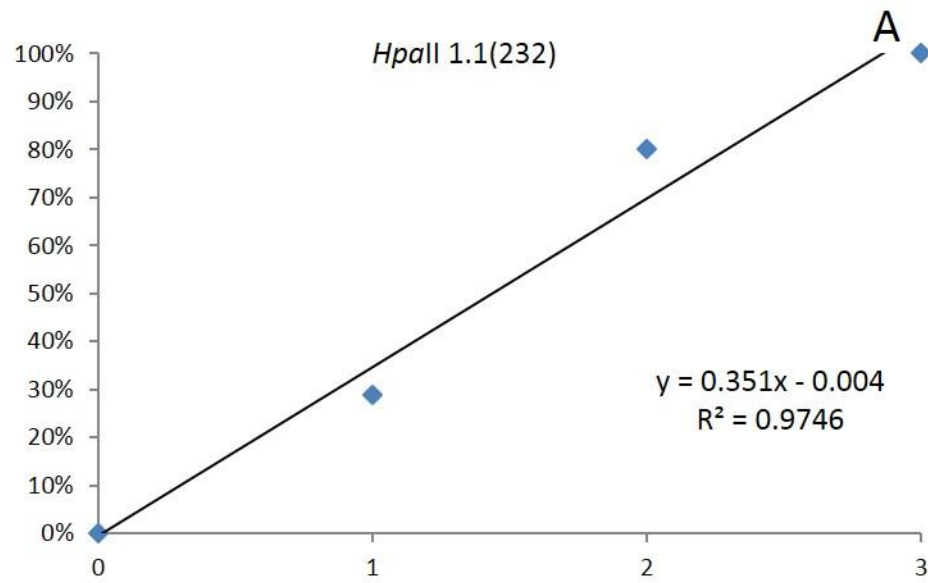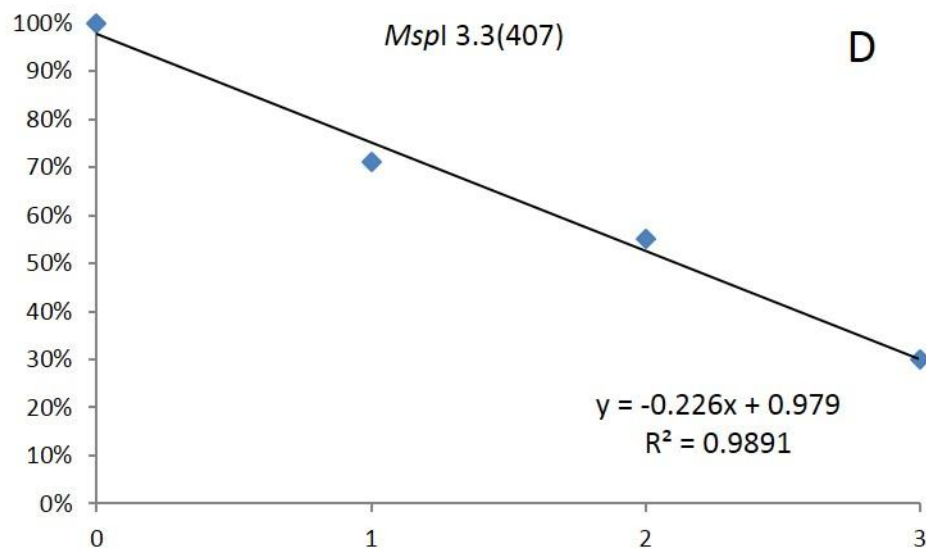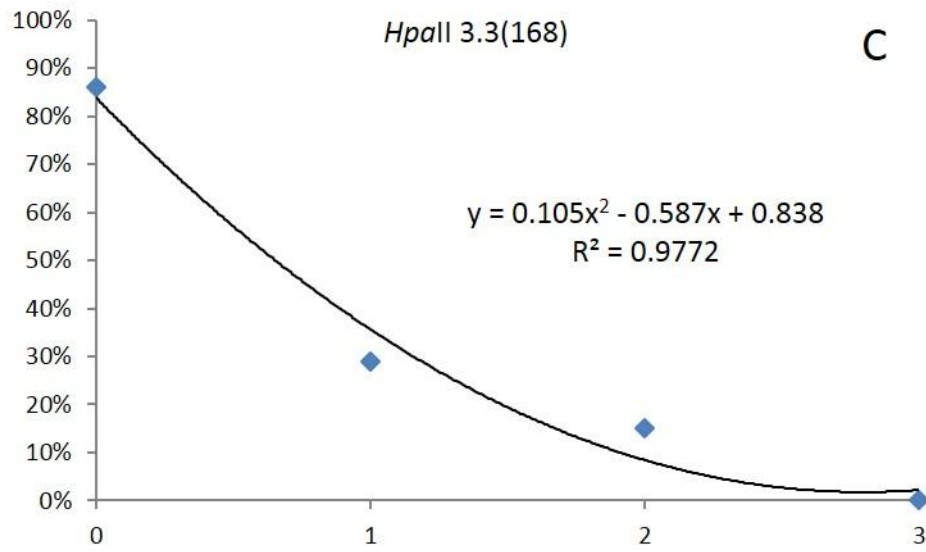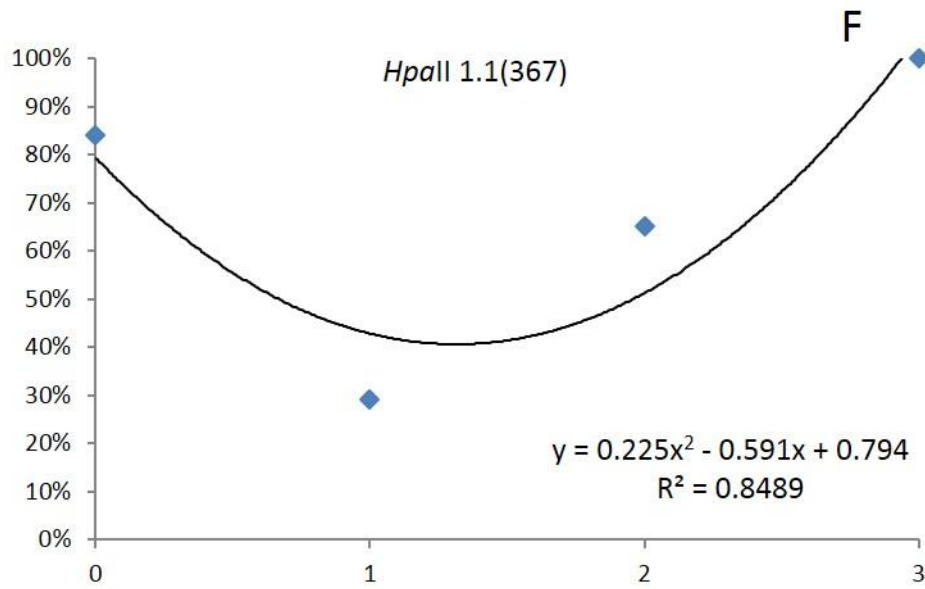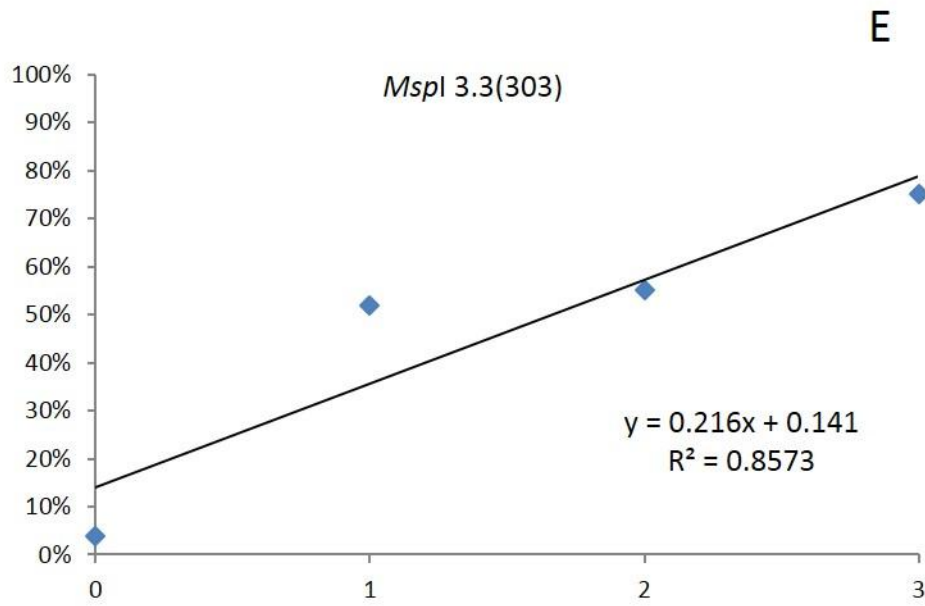
27
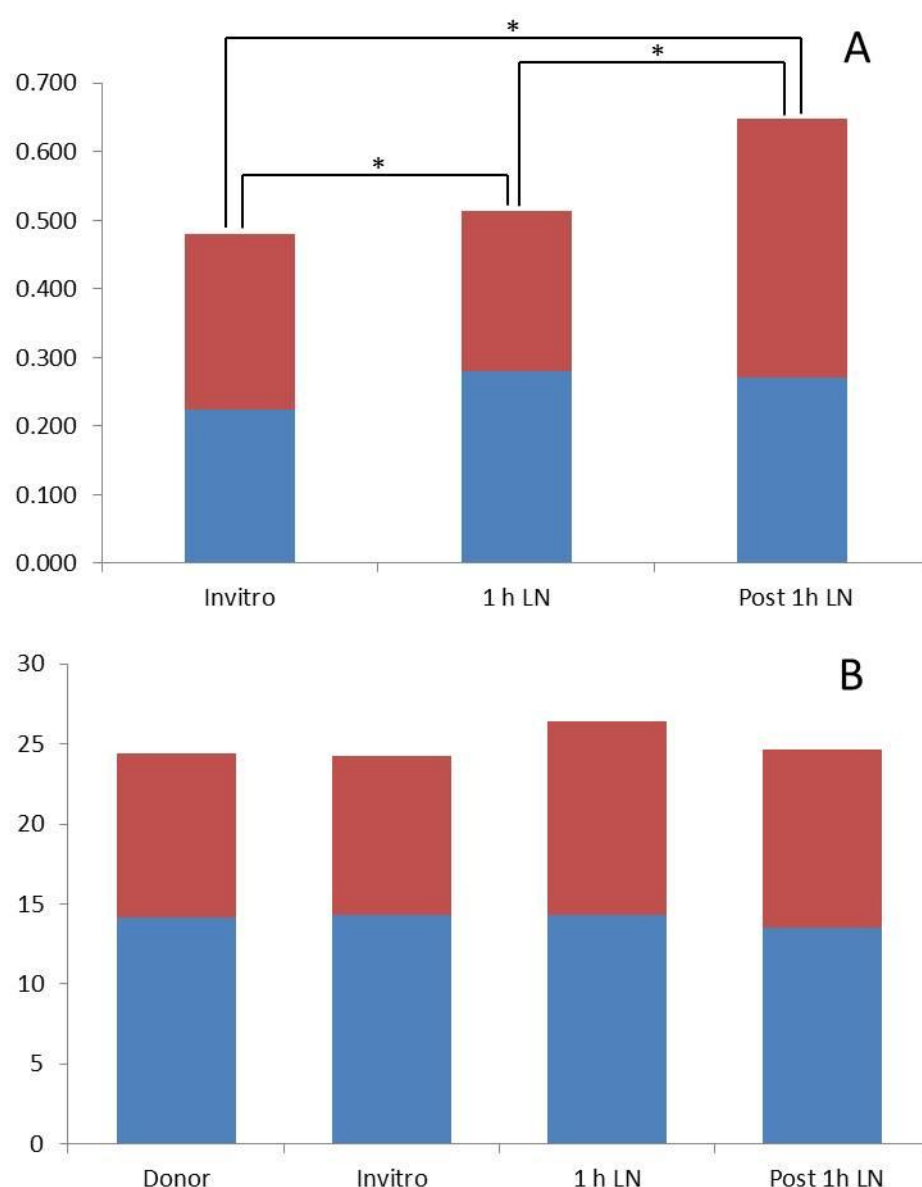
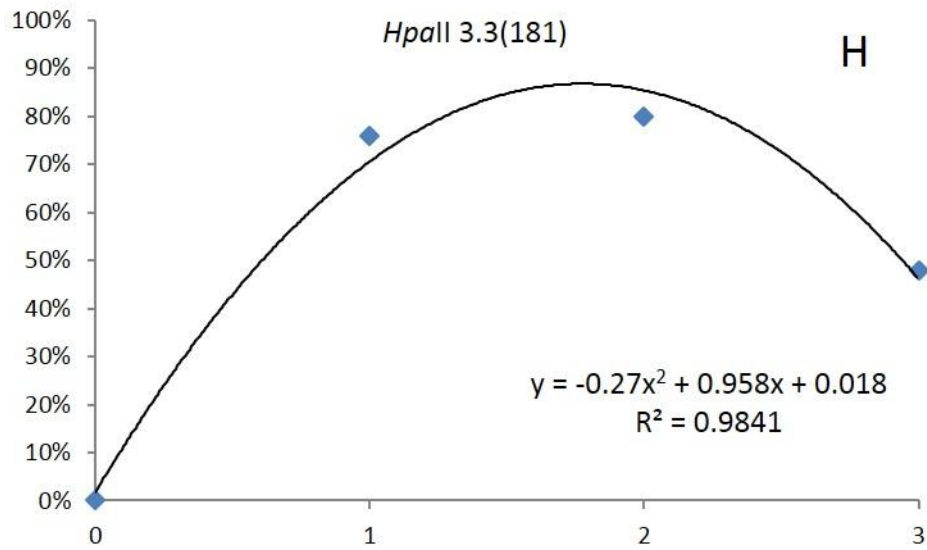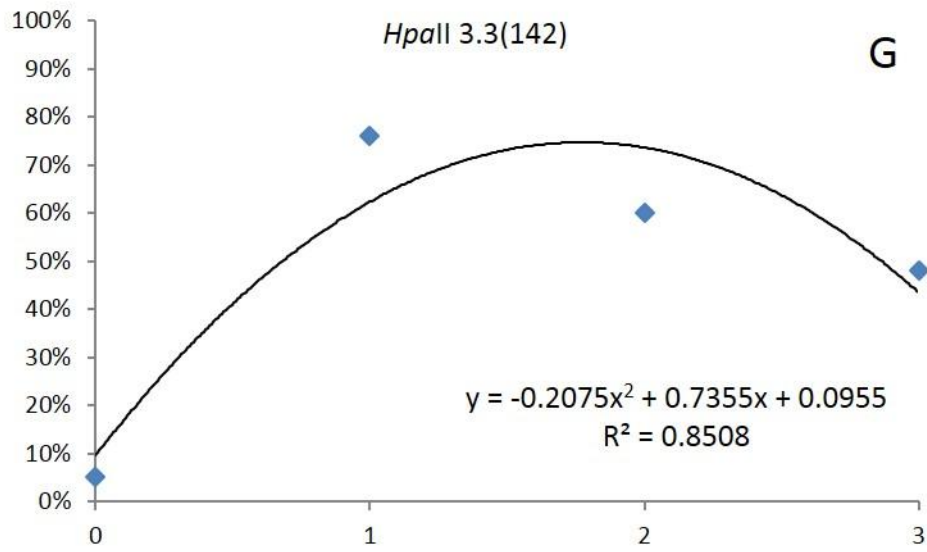**Figure 3: Somatic embryogenesis and cryopreservation induced between epigenetic variability.**

A) Bars represent epigenetic distance (PhiPT) between ortet tree samples and each of the *in vitro* treatments (**Invitro**: *In vitro* maintained secondary somatic embryos; **1 h LN**: Secondary somatic embryos recovered from after 1 hour storage in liquid nitrogen and **Post 1 h LN**: tertiary somatic embryos generated from 1 h LN samples) calculated using 10.000 permutations and AMOVA analysis. **B)** Bars represent the average epigenetic variability (mean sum of squares within population (SSWP)) between samples with the same origin (i.e. Donor, Invitro, 1 h LN and Post 1 h LN). PhiPT
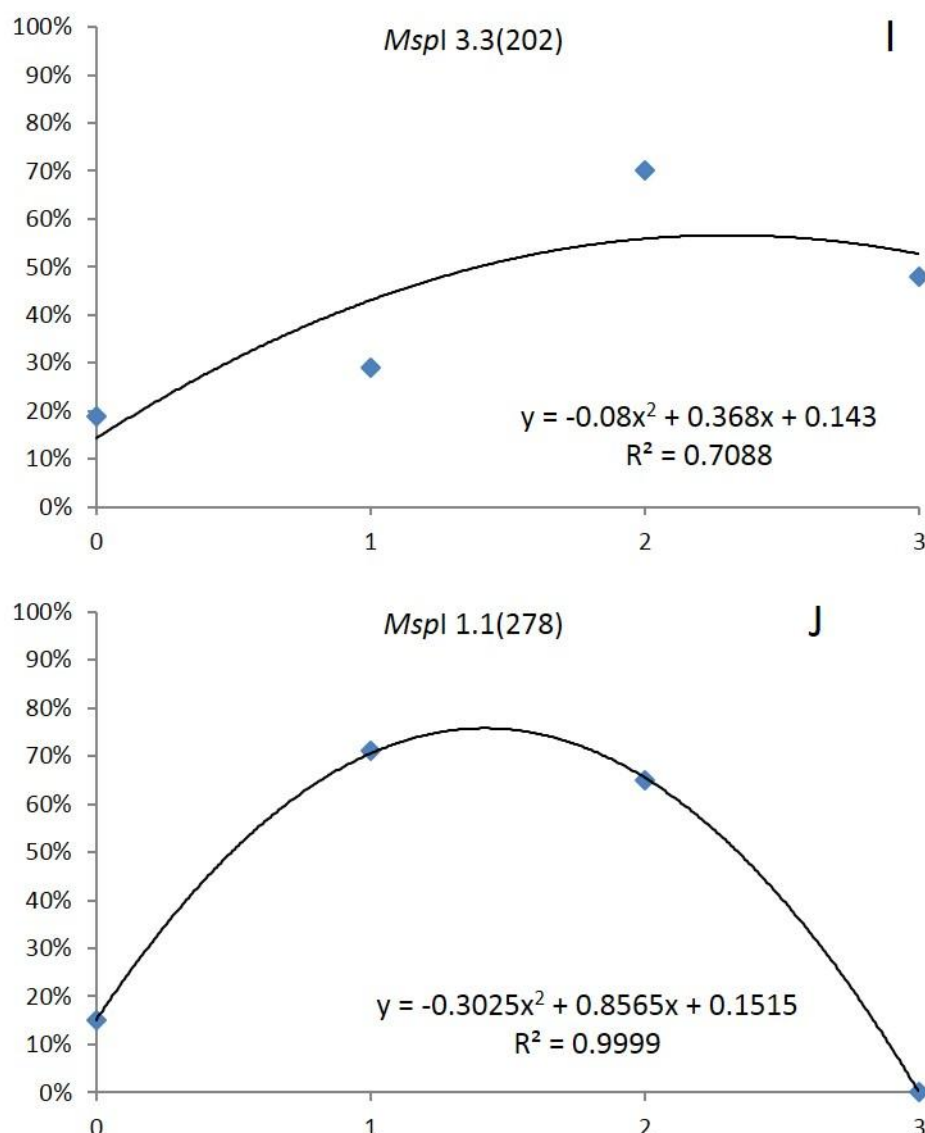
28

and SSWP values were calculated using GenAlex 6.1 software from MSAP profiles generated combining *Hp*3/*Eco*3 and *Hp*1/*Eco*1 selective primer combinations and restriction enzymes *Hap*II (Blue) and *Msp*I (red) as frequent cutters. All somatic embryos and ortet tree samples (24 per treatment) were initiated/collected from a single AMAZ 15 cocoa tree from the International Cocoa Quarantine Centre, Reading, UK. * Indicates significantly different (P>0.0002) PhiPT values between treatments.

A

*Hpa*II 1.1(232)

$y = 0.351x - 0.004$
$R^2 = 0.9746$

B

*Hpa*II 1.1(233)

$y = 0.1075x^2 - 0.6695x + 1.0205$
$R^2 = 0.9872$

Graph C: *Hpa*II 3.3(168)

$y = 0.105x^2 - 0.587x + 0.838$

$R^2 = 0.9772$

Graph D: *Msp*I 3.3(407)

$y = -0.226x + 0.979$

$R^2 = 0.9891$

31

E

*Msp*I 3.3(303)

$y = 0.216x + 0.141$
$R^2 = 0.8573$

F

*Hpa*II 1.1(367)

$y = 0.225x^2 - 0.591x + 0.794$
$R^2 = 0.8489$

*Hpa*II 3.3(142)

G

$$y = -0.2075x^2 + 0.7355x + 0.0955$$
$$R^2 = 0.8508$$

*Hpa*II 3.3(181)

H

$$y = -0.27x^2 + 0.958x + 0.018$$
$$R^2 = 0.9841$$

33

Supplementary Figure 1: Effect of somatic embryogenesis and cryopreservation on epilocus frequency. Values on the horizontal axis indicated the number of treatments each group has been subjected to (0= Donor Plant, 1= secondary somatic embryos maintained in ED medium, 2= secondary somatic embryos after cryopreservation for 1 h in liquid nitrogen (1h LN) and 3= tertiary somatic embryos generated from 1 h LN samples). Epilocus frequencies were calculated from presence/absence MSAP profiles generated combining $Hp3$/$Eco3$ and $Hp1$/$Eco1$ selective primer combinations and restriction enzymes HapII and MspI using GenAlex 6.1 software.