UNIVERSITÉ MONTPELLIER II

SCIENCES ET TECHNIQUES DU LANGUEDOC

THÈSE

Présentée pour l'obtention du titre de

Docteur en Sciences de l'Université Montpellier II

Formation Doctorale : Biologie Intégrative des Plantes

École Doctorale : SIBAGHE – Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydro sciences et Environnement

par

Roberto Bobadilla LANDEY

Influence of micropropagation through somatic embryogenesis on somaclonal variation in coffee (*Coffea arabica*): assessment of variations at the phenotypical, cytological, genetic and epigenetic level

Soutenue publiquement devant le jury composé de,

Rapporteur	Marie-Anne Lelu-Walter, Directrice de Recherche, INRA Orléans
Rapporteur	Stéphane Maury, Professeur, Université Orléans
Examinateur	Michel Nicole, Directeur de Recherche, IRD Montpellier
Examinateur	Alberto Cenci, Chercheur, Bioversity International Montpellier
Directeur de thèse	Hervé Etienne, Chercheur, CIRAD Montpellier

A mis padres y hermanos…
…Por qué siempre están conmigo y siempre han creído en mí.

Agradezco a

Mi familia, que siempre me ha apoyado y está pendiente de mí.

Hervé Etienne, por aceptar trabajar conmigo y por todos sus consejos, suerte con tus nuevos proyectos.

June Simspon por creer en mí y recomendarme a este proyecto.

Michel Nicole, chef gracias por todo.

Benoît Bertrand, Alberto Cenci, Philippe Lashermes, Romain Guyot, Juan Carlos Herrera, Frédéric Georget y demás personas de mi equipo, verdaderos especialistas que contribuyeron en mi tesis, gracias por todos sus consejos para mejorar nuestros artículos y por sus enseñanzas.

Sylvain Santoni, boss gracias por todo, la siguiente vez secuenciación a fondo.

Eveline Dechamp, Isabelle Jourdan, Muriel Latreille, Morgane Ardisson, Audrey Webber, Katia Gil Vega por su asistencia técnica, sin ustedes no hubiera hecho ni una PCR.

A todos mis amigos y compañeros, gracias a ustedes nunca me sentí lejos de casa.

ABBREVIATIONS

2, 4-D 2, 4-dichlorophenoxyacetic acid
6-BA 6-benzylaminopurine
6-γ-γ-Dimethylallylaminopurine (2iP)
ABA acid absisic
AFLP Amplified Fragments Length Polymorphisms
AS agroforestry systems
BAC Bacterial Artificial Chromosome
CGA chlorogenic acids
DAPI 4',6-diamidino-2-phenylindole,
EG Embryo germination medium
ESP embryogenic suspensions
FISH Fluorescence
FCM flow cytometry
GA Gibberellins
GISH genomic *in situ* hybridization
HFSE High frequency somatic embryogenesis
IAA indole-3-acetic acid
IBA indole-3-butyric acid
LFSE Low frequency somatic embryogenesis
LTR Long Terminal Repeat
MSAP Methylation sensitive Amplified Polymorphisms
NAA naphthaleneacetic acid
NGS Next Generation Sequencing
PCR Polymerase Chain Reaction
PEMs Proembryogenic masses
PTC Plant tissue culture
RFLP Restriction fragment length polymorphism
RT-PCR Reverse transcriptase polymerase chain reaction
SDS Sodium Dodecyl Sulfate
SE Somatic Embryogenesis
SNP single nucleotide polymorphism
SSAP Sequence Specific Amplified Polymorphisms
SCE secondary embryogenesis
SV Somaclonal variaiton
TE Transposable element
TDZ Thidiazuron

SUMMARY

LIST OF FIGURES LITERATURE REVIEW

LIST OF TABLES LITERATURE REVIEW

Introduction générale

L'amélioration génétique du caféier Arabica (allopolyploïde, autogame) est rendue très difficile en raison d'une base génétique très étroite et du caractère pérenne de la plante (durée des cycles biologiques supérieure à cinq ans). Traditionnellement, l'amélioration génétique est basée sur la sélection généalogique qui aboutit à des lignées (homozygotes) propagées par semences. Les gains génétiques avec de tels programmes sont lents et peu importants. Au cours des années 90 en Amérique Centrale, le CIRAD avec ses partenaires a mis au point des clones d'hybrides F1 d'Arabica présentant une forte productivité, des résistances à la rouille orangée et à des nématodes à galles ainsi qu'une excellente qualité organoleptique. Etant donné la nature hétérozygote du matériel sélectionné, il a fallu développer une technique de propagation végétative pour permettre une diffusion rapide et à grande échelle de ces hybrides. Le choix s'est porté sur l'embryogenèse somatique (ES) en absence de techniques horticoles efficientes. Le défi était risqué car cette technique n'était alors utilisée au niveau industriel que chez le palmier à huile. Grâce aux recherches faites entre 1995 et 2005, un procédé d'ES relativement fiable a été mis au point. Fonctionnant de façon reproductible sur l'ensemble des hybrides, il permit la production expérimentale de plantes et aussi la mise en place du réseau d'essais avec les clones d'hybrides dans toute l'Amérique centrale pour l'évaluation multi-locale et enfin, la sélection définitive de nouvelles variétés résultant du croisement de variétés américaines et de souches 'sauvages' d'Ethiopie. Un accord de partenariat signé en 2002 entre une entreprise privée et le CIRAD a permis d'initier le transfert industriel de la méthode d'ES et de construire deux laboratoires de production au Nicaragua et au Mexique.

Avec comme objectif de répondre à un marché évalué à 10-15 millions de plantes/an, la maitrise des variations somaclonales (VS) et la propagation d'individus génétiquement conformes à la plante mère sont devenus des enjeux prioritaires pour les partenaires. Un ensemble de précautions a été pris au niveau de la production, en laboratoire comme en pépinière, pour limiter l'apparition et la diffusion d'arbres variants. Cependant, avec près de 4 millions d'hybrides actuellement au champ, il est apparu indispensable d'évaluer précisément les risques liés à la VS. La littérature sur l'ES rapporte régulièrement la production de VS mais peu d'études ont été faites à une échelle massive et la plupart du temps elles s'appuient sur des protocoles non optimisés mettant en jeu des concentrations excessives de régulateurs de croissance.

Générer régulièrement une information massive et précise sur l'impact des VS dans les conditions industrielles utilisées à moment donné est important pour un projet qui s'inscrit dans un changement d'échelle durable. Les risques augmentant probablement avec le changement d'échelle, une information aussi complète soit elle sur l'impact des variations somaclonales à une échelle donnée ne peut protèger complètement d'un accident à une échelle supérieure. Dans l'objectif d'établir un procédé industriel le plus sûr et solide possible il est donc apparu nécessaire de compléter ces études massives par des dispositifs expérimentaux destinés d'une part à estimer les limites du système en appliquant des conditions de culture que l'on suppose mutagènes (dans ce travail, la durée de prolifération cellulaire en présence de régulateurs de croissance), et destinés d'autre part à appréhender les mécanismes conduisant aux VS.

La présente étude s'appuie donc sur des données industrielles massives mais aussi sur un grand nombre de vitroplants passés plusieurs fois au crible des marqueurs moléculaires. Nous avons par ailleurs appliqué une approche multidisciplinaire combinant agronomie, cytologie et biologie moléculaire afin de mettre en évidence une large gamme de changements, qu'ils soient de nature chromosomique, génétique ou épigénétique, intervenant au cours du procédé d'ES chez le caféier.

CHAPTER I

LITERATURE REVIEW

PART I

Generalities about coffee

Coffee consumption, cultivation and economical trade

Coffee is the most popular beverage. It's consumed either hot or cold by about one-third of the people in the world. Its popularity can be attributed to its stimulant effect, which is produced by caffeine, an alkaloid present in green coffee in amounts between 0.8 - 1.4% for the Arabica varieties and 1.7 - 4.0% for Robusta (*Coffea canephora*) (www.ico.org). According to International Coffee Organization (www.ico.org) about 1.4 billion cups of coffee is consumed every day worldwide. The biggest consumers of coffee are the Nordic countries. The leader in consumption is Finland with 12 kg per capita in 2007 (**Table I.1**). Coffee is cultivated between the Tropics of Cancer and Capricorn forming a "bean belt" (**Figure I.1**). Growing regions typically offer moderate sunshine and rain, steady temperatures around 20 ºC (70 ºF), and rich, porous soil. The coffee beans are an economic mainstay for many countries and about 25 million people. It is the second natural commodity; having a monetary value surpassed only by petroleum. The total production for crop year 2009/10 totaled 120.6 million bags (www.ico.org). Brazil is responsible for about a third of all coffee production (39.4 million bags of 60 kg), making it by far the most important coffee-producing country, followed by Vietnam, Indonesia, Colombia with 18, 10.6 and 9 million bags of 60 kg, respectively in the crop year 2009/10.

Table I. 1 Top ten countries consumption of coffee ordered by annual *per capita* in 2007.

Rank	Country	Coffee consumption (kg per capita per year)
1	Finland	12
2	Norway	9.9
3	Iceland	9.0 (2006)*
4	Denmark	8.7
5	Netherlands	8.4
6	Sweden	8.2
7	Switzerland	7.9
8	Belgium/Luxembourg	6.8
9	Aruba	6.8 (2006)*
10	Canada	6.5

Where 2007 data is not available, 2006 data is provided. Available online at: http://www.ico.org/historical.asp.

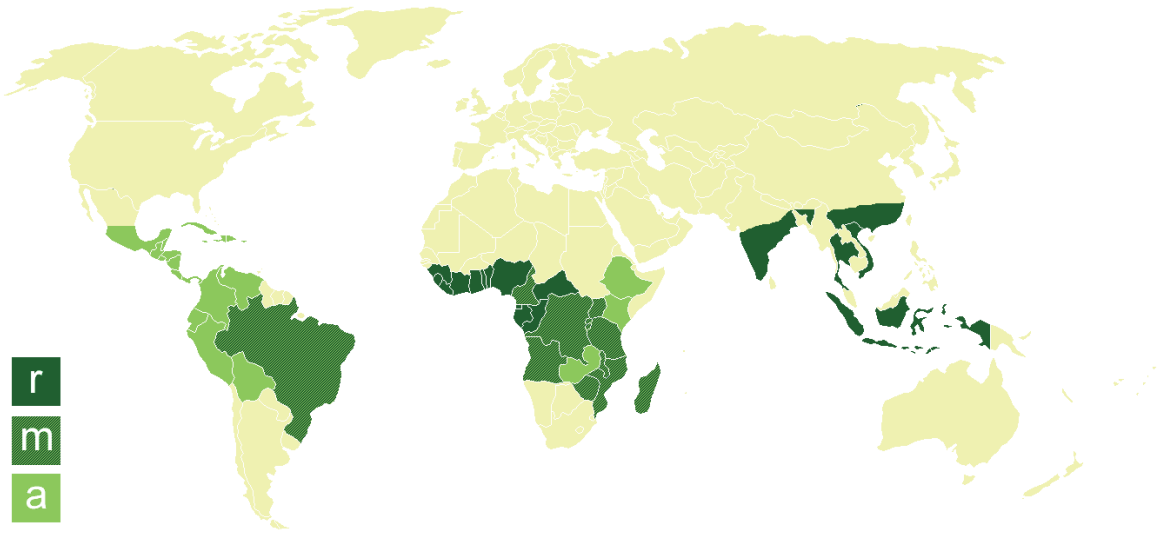


Figure 1.1 Map showing areas of coffee cultivation in the bean belt. r) *Coffea canephora*, m) *Coffea canephora* and *Coffea Arabica*, a) *Coffea Arabica* (source wikipedia).

Coffee taxonomy

Coffee belongs to Rubiaceae family, one of the five largest flowering plants families, with ~ 13000 species classified in 620 genera, more than 40 tribes and three subfamilies (Goevarts *et al.* 2006). Rubiaceae are found on all continents but most taxa are in tropical or subtropical areas (Goevarts *et al.* 2006). Most of them are woody trees and shrubs. Plants of economical importance in this family include: Quinine (*Cinchona*) which possess medicinal properties, *Ixora* and *Gardenia* cultivated as ornamentals plants, madder (*Rubia* sp) used as dyes and *Coffea* by far the most agronomical important species of Rubiaceae family. The Rubiaceae family is divided into three subfamilies: Cinchonoideae, Ixoroideae and Rubioideae (Bremer and Eriksson 2009). Ixoroideae subfamily that comprises about one-fifth of all Rubiaceae genera has 15 recognized tribes (Bremer and Eriksson 2009), among them the tribu Coffeeae DC which comprises 11 genera including the genus *Coffea* and *Psilanthus* (Davis *et al.* 2007).

The coffee species share the typical coffee bean morphology, i.e. a groove on the flat side of the seed. The 'husk' or 'parchment' (horny/ crustaceous endocarp) of the pyrene also has a deep ventral groove, which follows the invagination of the outer layer of the seed (exotesta) (Davis *et al.*2006) (**Figure I.2**). They have been classified in two genera, *Coffea* L. and *Psilanthus* Hook. f. which differ in their flower morphology (Leroy 1980; Bridson, 1988; Davis *et al.*2005). The flowers in *Coffea* genus have long style, medium corolla tube, exserted anthers while in *Psilanthus* genus the flowers have short style, long corolla tube and encased anthers (**Figure I.3**). The genus *Coffea* is classically divided in two subgenera: *Coffea* subgenus *Coffea* and *Coffea* subgenus *Baracoffea* (Bridson 2003; Davis 2003; Davis *et al.*2005, 2006).

However a sister relationship between subgenus *Coffea* and *Baracoffea* was demonstrated to be highly unlikely (Maurin *et al.*2007). Nowadays the subgenus *Baracoffea* is referred as the 'baracoffea alliance' as a mean of recognition of nine *Coffea* species that are not yet clearly positioned on *Coffea* phylogeny but represent a strongly and morphological distinct lineage (Davis and Rakotonasolo, 2008). The genus *Psilanthus* has also been divided in two subgenus: *Psilanthus* subgenus *Psilanthus* (two species) and *Psilanthus* subgenus *Afrocoffea* (Moens) (20 species) (Bridson 1988; Davis *et al.*2005, 2006; Davis and Rakotonasolo 2008). New taxa of *Coffea* are still being discovered (Stoffelen *et al.*2009). A recent phylogeographic analysis of *Coffea* subgenus *Coffea* species using data on plastid sequences suggested that Lower Guinea could be the centre of origin of *Coffea* subgenus *Coffea* (Anthony *et al.*2010). Species included into the genera *Coffea* are self-incompatible with exception of the allotetraploid species *C. arabica* (Carvalho *et al.*1991) and the diploid species *C. heterocalyx* Stoff (Coulibaly et al.2002) and *C. anthonyi* (formerly *C.* sp 'Moloundou') found out in Cameroon and in Republic of Congo (Stoffelen *et al.*2009), respectively. Only two species are of economical interest: *C. canephora* and *C. arabica*. The superior quality beverage is produced from *C. arabica* which is the most important traded species.

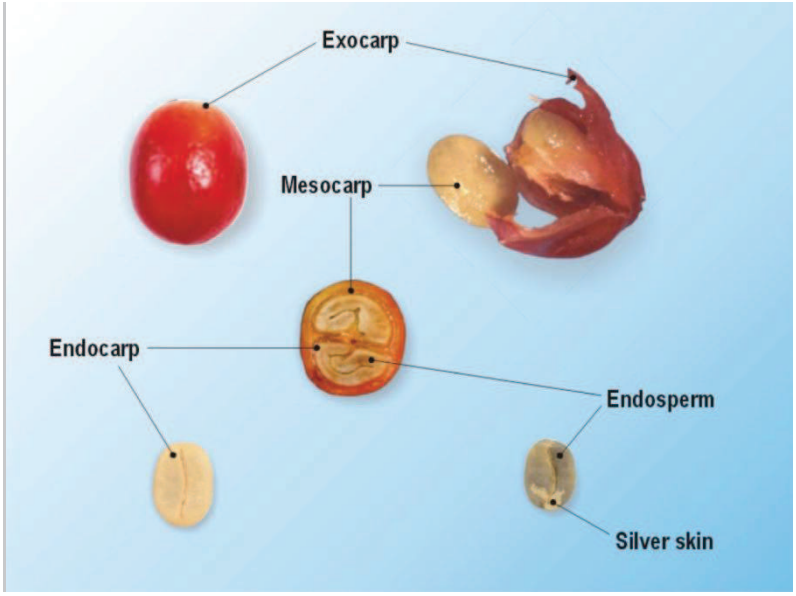


Figure I.2 Coffee bean morphology

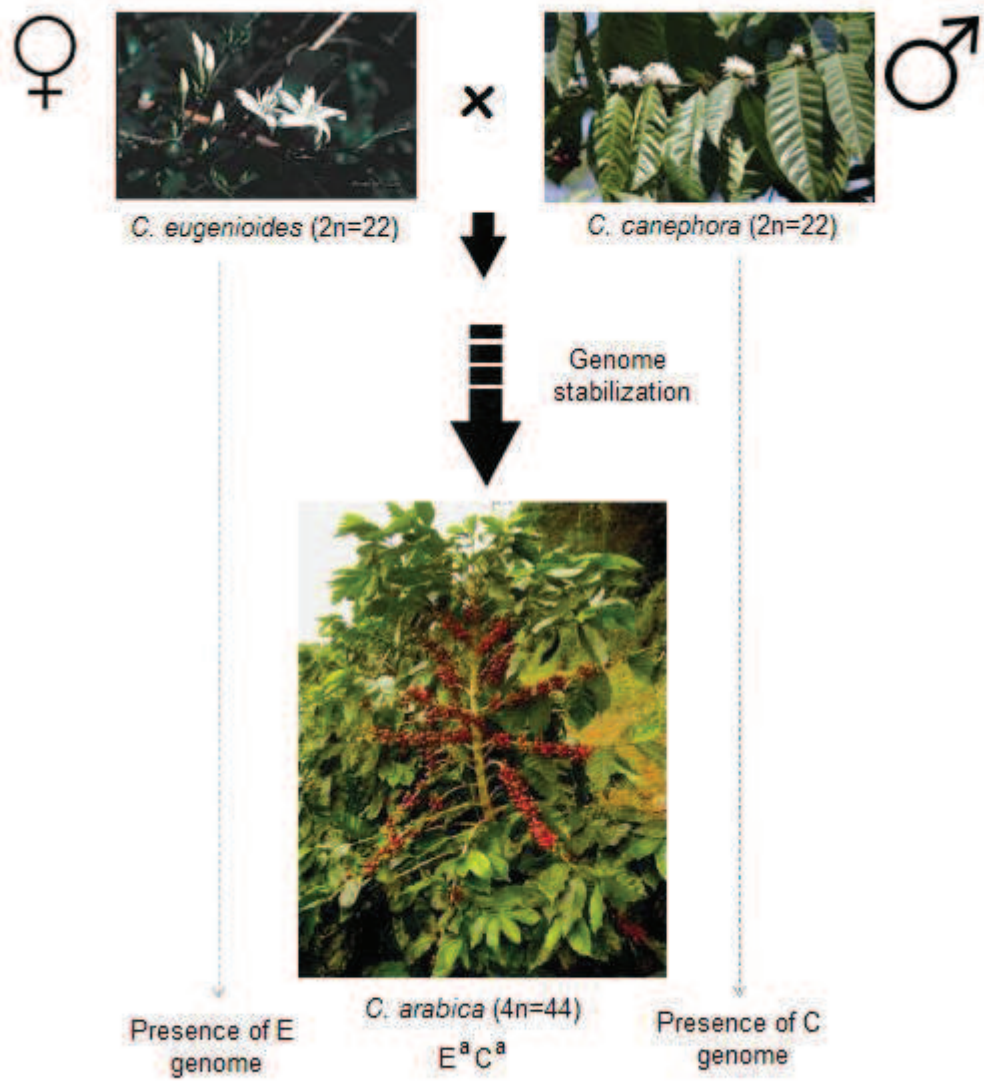
Figure I. 3 Flower morphology of *Coffea arabica*

Speciation and genome features of *C. arabica*

The genome constitution and mode of speciation of *C. arabica* have been subjected to several investigations. Based on cytological observations and fertility of interspecific hybrids, *C. canephora* and *C. congensis* were suggested to have a common ancestor with *C. arabica* (Carvalho 1952; Cramer 1957). Furthermore, *C. congensis* can be considered as an ecotype of *C. canephora* in the light of the fertility of interspecific hybrids (Louarn 1993) and the genetic diversity detected by molecular markers (Prakash *et al.* 2005). These phylogenetic assumptions were consolidated by a high homology found in the ITS2 region sequences of *C. arabica*, *C. canephora* and *C. congensis* (Lashermes *et al.* 1997). Regarding cpDNA that exhibited a maternal inheritance in coffee (Lashermes et al. 1996), *C. arabica* appeared to be similar to two species, *C. eugenioides* Moore and *Coffea* sp. "Moloundou" (Cros *et al.* 1998).

Restriction fragment length polymorphism (RFLP) markers in combination with genomic *in situ* hybridization (GISH) were used to investigate the origin of *C. arabica.* By comparing the RFLP patterns of potential diploid progenitor species with those of *C. arabica*, the source of the two sets of chromosomes, or genomes, combined in *C. arabica* was specified. The genome organization of *C. arabica* was confirmed by GISH using simultaneously labeled total genomic DNA from the two putative genome donor species as probes (Lashermes *et al.* 1999). These results clearly suggested that *C. arabica* is an amphidiploid (i.e. C^a E^a genomes) resulting from the hybridization between *C. eugenioides* (E genome) and *C. canephora* (C genome) or ecotypes related to those diploid species (**Figure I.4**) Results also indicated low divergence between the two constitutive genomes of *C. arabica* and those of its progenitor species, suggesting that the speciation of *C. arabica* took place very recently.

The precise localization in Central Africa of the speciation process of *C. arabica* based on the present distribution of the coffee species appeared difficult since the constitution and extent of tropical forest varied considerably during the late quaternary period. Furthermore, in spite of the close relationship among the two constitutive genomes, bivalents are observed in meiosis and *C. arabica* is considered to display a diploid-like meiotic behavior (Krug and Mendes 1940). Investigations suggest that homeologous chromosomes do not pair in *C. arabica*, not as a consequence of structural differentiation, but because of the functioning of pairing regulating factors (Lashermes *et al.* 2000a).



FigureI. 4 Genetic origin of allotetraploid *C. arabica*. Natural hybridization between *C. eugenioides* as female parent and *C. canephora* as male parent originated the *C. arabica* species.

Characteristics of coffee cultivated species

The coffee plant is a woody shrub, and it grows in the wild as high as 12 meters, but cultivated trees are pruned to 2 meters. It has a main vertical trunk (orthotropic) and primary, secondary and tertiary horizontal branches (plagiotropic). Small, white flowers give rise to a red or yellow fleshy fruit, the "coffee cherry" (**Figure I.5**) which contains a pair of beans (http://www.ico.org). A brief comparison of botanical aspects between *C. arabica* and *C. canephora* is presented in **Table I.2.**

Table I. 2. Comparison of some botanical aspects between *C. arabica* and *C. canephora* species

Botanical aspect	***C. arabica***	***C. canephora***
Date species described	1753	1895
Chromosomes (2n)	44	22
Time from flower to ripe cherry	9 months	10-11 months
Flowering	after rain	after rain
Ripe cherries	fall	stay
Yield (kg beans/ha)	1500-3000	2300-4000
Root system	deep	shallow
Optimum temperature (yearly average)	15-24° C	24-30° C
Optimal rainfall	1500-2000 mm	2000-3000 mm
Optimum altitude	1000-2000 m	0-700 m
Hemileia vastatrix	susceptible	resistant
Koleroga	susceptible	tolerant
Nematodes	susceptible	resistant
Tracheomycosis	resistant	susceptible
Coffee berry disease	susceptible	resistant
Caffeine content of beans	0.8-1.4%	1.7-4.0%
Shape of bean	flat	oval
Typical brew characteristics	acidity	bitterness, full
Body	average 1.2%	average 2.0%

Available in: http://www.ico.org/botanical.asp accessed in 23 July 2010 00:06

Coffea arabica is the only tetraploid from the genera *Coffea*. There are two distinct botanical varieties of arabica coffee: typica and bourbon. Historically, typica was cultivated in Latin America and Asia, whereas bourbon arrived in South America and, later, East Africa via the French colony of Bourbon Island (Reunion). Typica coffee plants have a conical shape with a main vertical trunk and secondary verticals that grow at a slight slant. Typica is a tall plant reaching 3.5-4 m in height. The lateral branches form 50-70° angles with the vertical stem. Because *C. arabica* is mainly self-pollinating, these varieties tend to remain genetically stable. However, spontaneous mutants showing desirable characteristics have been cultivated in their own right, as well as being exploited for cross-breeding purposes. Some of these mutants and cultivars are described next: (summarized from Clifford and Willson 1985). Caturra, a compact form of bourbon found near the town of Caturra, Brazil in the 1930s. A relatively recently selected botanical variety of the *C. arabica* species that generally matures more quickly, produces more coffee, and is more disease resistant than older, traditional arabica varieties. Caturra was largely exploited to industrialize the coffee culture thanks to the use of dwarf varieties.

Other cultivars have been developed to give the maximum economic return under specific regional conditions such as climate, soil, methods of cultivation and the prevalence of pests and diseases. Some of the better known cultivars are: Blue Mountain grown in Jamaica and Kenya; Mundo Novo a cross between typica and bourbon, originally grown in Brazil; Kent a variety originally developed in India, showing some disease resistance; Catuai an artificial hybrid between Mundo Novo and Caturra, characterized by either yellow or red cherries (Catuai-amarelo and Catuai-vermelho respectively). One special case is Timor normally called "Hibrido de Timor" is a natural hybrid of Arabica x Robusta which resembles Arabica coffee and has 44 chromosomes. It was found on the island of Timor (Indonesia) around the 1940s and it was cultivated because of its resistance to leaf rust.

Figure I.5 *Coffea arabica* tree with cherry beans.

Cytological aspects of genus *Coffea*

Cytological studies in coffee and other plants of Rubiaceae family demonstrated that *C. arabica* possess 2x=2n=44 chromosomes while diploid species possess 2x=2n=22 chromosomes (Krug 1934). They also determined n=11 as a basic number of chromosomes for the genus *Coffea*. These results were confirmed later by Fagerlind (1937). DNA content in plants and others eucariots are usually expressed in picograms (pg) or in megabase pairs of nucleotides (Mb) where, 1 Mb = 10^6 nucleotide base pairs and 1 pg = 965 Mb (Bennett and Leitch 1995). The chromosomal DNA 2C content (i.e a nucleus in G_1 phase of the cell cycle, with two copies of unreplicated genome) of *Coffea* species have been measured by flow cytometry (FCM) and slight variation has been reported by different authors depending on the methodology, standards, buffers or fluorochromes used. The mean content of chromosomal DNA (2C) estimated in different *C. arabica* and C. canephora cultivars ranged from 2.38 to 2.84 pg. and 1.32 to 1.76 pg. respectively (Cros *et al.*1995; Noirot *et al.* 2003; Clarindo and Carvalho 2009). The variations in DNA content among coffee species, besides of ploidy level (e.g. *C. arabica*), are probably due almost entirely to variation in the copy number of repeated DNA sequences. Differences may correspond to genomic evolution correlated with an ecological adaptation process (Lashermes and Anthony 2005).

PART II

Coffee breeding

Plant breeding can be conveniently separated into two activities: manipulating genetic variability (i.e. introducing new genetic material into the cultivated lines of interest) and plant evaluation (Brown and Thorpe 1995). The transfer of desirable genes in particular for disease resistance from diploid species like *C. canephora* and *C. liberica* into tetraploid Arabica cultivars without affecting quality traits has been the main objective of Arabica breeding (Carvalho 1988; Van der Vossen 2001). To date, *C. canephora* provides the main source of disease and pest resistance traits not found in *C. arabica*, including coffee leaf rust (*Hemileia vastatrix*), Coffee Berry Disease (*Colletotrichum kahawae*) and root-knot nematode (*Meloidogyne spp.*). Likewise, other diploid species present considerable interests in this respect. For instance, *C. liberica* has already been used as source of resistance to leaf rust (Srinivasan and Narasimhaswamy 1975) while *C. racemosa* constitutes a promising source of resistance againts the coffee leaf miner (Guerreiro Filho *et al.* 1999). Exploitation of coffee genetic resources has so far relied on conventional procedures in which a hybrid is produced between an outstanding variety and a donor genotype carrying the trait of interest, and the progeny is backcrossed to the recurrent parent (Bertrand *et al.* 2005).

Undesirable genes from the donor parent are gradually eliminated by selection. In doing so, conventional coffee breeding methodology faces considerable difficulties. In particular, strong limitations are due to the long generation time of coffee-tree (5 years), the high cost of field trial, and the lack of accuracy of current strategy. A minimum of 25 years after hybridization is required to restore the genetic background of the recipient cultivar and there by ensure good quality of the improved variety.

Development of F1 hybrid varieties

Since the 1980s, several researchers have proposed the creation of hybrid varieties to help in increasing genetic diversity, notably by using wild Sudan-Ethiopian origins (Charrier 1978) and to exploit heterosis between genetic groups (Walyaro 1983; Van der Vossen 1985). Positively, the introgression via "Timor Hybrid confers rust and nematodes resistance. Consequently, different cultivars with resistance to several leaf rust races and a better productivity of aproximately 15% were obtained in Central America, Brazil and Colombia (e.g. 'Costa Rica 95', 'Obata', 'IAPAR59' and 'Colombia' among others). Nevertheless, the gene introgression was usually accompanied with a substantial decrease in cup quality (Bertrand *et al.* 2003). Hence, most national coffee organizations in charge of disseminating varieties in Central America have now suspended the distribution of these

introgressed varieties. In the early 1990s, a regional breeding programme in the context of a regional collaborative project between PROMECAFE, CIRAD and CATIE was undertaken based on the use of hybrid vigor in crosses between traditional and wild genotypes from Ethiopia, Kenia, Sudan and Yemen from CATIE collection and american cultivar like "Caturra" and Catuai and cultivars derived from "Timor Hybrid" like "Catimor" (Caturra x Timor Hybrid) as well as "Sarchimor" (Villa Sarchi x Timor Hybrid). Traditional american cultivar present the inconvenient of a narrow genetic base resulting in a lack of resistance to main diseases like coffee leaf rust and nematodes but with the advantage to be adaptable to full-sun cultivation (unshaded) (Bertrand *et al.* 2012). On the other hand, the Sudan–Ethiopian materials provide resistance to nematodes (Anzueto *et al.* 2001), partial resistance to leaf rust (Gil *et al.* 1990) and possibly similar or new organoleptic qualities compared to traditional varieties. As a result, the intraespecific hybridization has succesfully broaden the narrow genetic base with new F1 *C. arabica* hybrids (i.e. F1 first filial generation from different parental lines) and introgressed resistant genes (Bertrand *et al.*2011). In relation to heterosis (i.e. hybrid vigour) in the species, Carvalho & Monaco, (1969), Walyaro, (1983), Ameha (1990), Bellachew (1997), then Cilas *et al*. (1998) demonstrated its existence by intercrossing (Figure of F1 hybrid creation).

Description of the hybrid vigour in *Coffea arabica*

In *C. arabica*, heterosis calculated on the basis of the best parent was evaluated from crosses between different genetic pools. The heterosis observed by different authors varies from 10% to 40% (Ameha 1990; Carvalho and Monaco 1969; Fazuoli *et al.* 1993; Walyaro 1983; Netto *et al.* 1993). The heterosis found in Central America (22.0 to 47.0%) was globally around the same magnitude as that observed by the majority of authors. For the vegetative variables, significant differences appeared between the two populations (hybrids and lines) for the length of primaries and for stem girth (Bellachew 2001; Bertrand 2005b). Otherwise, the hybrids were not significantly different from the lines for plant height, number of primary branches, internode's length of primaries, leaf area. In Central America or in Ethiopia, the yield differences between the parental lines and the hybrids were not explained by the yield components, such as the number of fruits per node or by the weight of 100 beans which were identical for both populations. Finally, the F1 hybrid population showed lower fertility than the population of lines. In Central America conditions, the difference in fertility rate was from 1.2 to 6.3% for floating fruits i.e. due to the presence of seedless fruits. In coffee, the number of seeds per fruit depends on ovule fertility (Louarn, 1992). Neither could heterosis be explained by better fertility, since hybrid fertility was even lower than that of the lines. Heterosis seemed to be permanently reflected in longer primary branches. This difference in length, which could not be attributed to a difference in internode's length, came from a larger number of internodes formed over the same time lapse.

The hybrid vigour in the Arabica species needs to be considered in relation to those obtained with other self–fertilizing species. In rapeseed, Lefort-Buson & Dattée (1985) found heterosis of 12.0 to 18.0% between populations of different origin. In barley, heterosis under suitable growing conditions was found to be between 10.0 and 25.0% (Scholz & Kunzel 1987). In wheat, heterosis was 17% to 40% depending on density (Oury *et al.*1990). The results obtained with *C. arabica* therefore showed heterosis that was comparable with that of most self-fertilizing species. The increase in biomass has led many authors to study heterosis in relation to density (Orozco 1975). The hypothesis is that competition between plants might lessen heterosis. However, the first results obtained from a network of trials set up in Central America since 1998, under commercial planting density conditions (5,000 to 10,000 plants/ha), do not seem to reveal such a tendency. Though more vigorous, hybrids seem to withstand strong competition very well. In one hybrid trial, at a density of 7,000 trees/ha, the difference between the best hybrid and the best line amounted to 36% (over 4 harvests). Heterosis would seem to go hand in hand with greater homeostasis, which enables hybrids to withstand stresses due to strong competition.

Performance of selected hybrids

In Central America, multi-trait selection in a hybrid population led to substantial expected genetic progress for yield and low progress for the weight of 100 beans (Bertrand *et al.* 2005). On the other hand, selection for productivity and fertility were opposite. This means that it is difficult to select for yield without, in return, reducing fertility. Conversely, by selecting for better fertility, no genetic progress will be made for yield within the population of hybrids. We put forward the hypothesis that heterozygosity leads to a drop in fertility and an increase in vigour. Indeed, it has been found that pedigree selection always goes hand in hand with progress in fertility (without directed selection). Conversely, "Arabusta" trees, which are interspecific hybrids obtained between *C. arabica* and *C. canephora* display a very high level of hybrid vigour combined with very high sterility (Reffye, 1973). The most productive hybrids came from introgressed maternal lines derived from the Timor Hybrid crossed by wild origins. Coffee growers who have adopted Green Revolution principles (i.e. involving the development of high-yielding varieties, modernization of management techniques, distribution of hybridized seeds, synthetic fertilizers, and the use of pesticides among others) are aware of the need to crop their coffee in agroforestry systems (AS) as they did in the past, but this can result in an estimated 20–40% drop in productivity (Bertrand *et al.*2005). For this economic reason, AS have been progressively abandoned by producers that prefered artificialized full sun conditions. Coffee variety selection must be adapted to ancient agronomic practices like agroforestry-based ecologically intensive cropping systems suitable for F1 hybrids. In the AS, the mean yield of hybrids was 58%

higher than that of the American cultivars (Caturra, Pacas, Catuai, Bourbon and Catimor cultivars) while the mean yield of hybrids in the full-sun system was 34% higher (Bertrand *et al.*2011).

F1 hybrids and coffee beverage quality

The organoleptic evaluations under various edapho-climatic conditions and elevations, by comparing F1 hybrids with traditional cultivars ('Bourbon') for bean chemical contents and cup quality did not show any clear difference (Bertrand et al. 2006). F1 hybrids appeared in turn to be inferior, similar, or superior to traditional cultivars for certain attributes, such as acidity, or aroma. Regarding the overall standard, F1 hybrids were equivalent or superior to traditional cultivars. For caffeine, as for trigonelline, the hybrids did not differ from the traditional varieties. The hybrids showed a tendency to be slightly richer in chlorogenic acids (CGA) than the traditional varieties. On the other hand, elevation did not seem to influence fat contents for the F1 hybrids. Higher vigour resulted in better nutrient supply to the fruits, disregarding elevation of plantation areas. The use of F1 hybrids should thus contribute to reducing variation in the fat content of coffee beans, and at the same time reduce variations in beverage quality.

PART III

Plant tissue culture and micropropagation

Methods of sexual and asexual of propagation in plants and tissue culture

Most of the horticultural plants are conventionally propagated by seeds (sexual propagation), nevertheless, there are certain cases for which the use of sexual propagation is not appropriate. Perennial crops (plants that persist for many growing seasons) often pass through a long juvenile phase of vegetative development before they are capable of seed production and propagation Also, cross pollinitation is difficult to control often resulting in highly heterozygous and phenotypic diverse plants (Sujatha, 2011). Aditionally, some perenial species present recalcitrant seed (i.e. not supporting dessication) and need to be stored under special conditions; some of the species include mango (*Mangifera indica*), rubber tree (*Hevea brasiliensis*) and avocado (*Persea americana*). These kind of species need to be asexually regenerated and in acceptable numbers to fulfill the demand of producers and eventually markets. Vegetative or asexual propagation method produce planting materials using plant vegetative parts (i.e. stem, leaves among others) instead of using seeds. However, some of these objectives can not be overcome with traditional vegetative techniques like cutting, grafting and buddings (see **table I.3** for advantages and disadvantages). There exist more advanced vegetative techniques based on tissue culture: axillary shoot propagation (a.k.a. enhancing axillary bud breaking), adventious shoot propagation (also presented in **table I.3**) and somatic embryogenesis.

Table I.3 Summary of vegetative propagation techniques, advantages and disadvantages

Method	Procedure	Type of plants propagated	Advantages	Disadvantages
Ground and air layering	Stems attached to the plants are covered by soil to induce rooting	Ornamentals, fruit trees, forest trees	Simple, low technology requirements	Expensive, involve many step of manual operation
Grafting	Tissues from donor plants (scion) are inserted in the receptor (rootstock) and the vascular tissues join together	Ornamentals, fruit trees	Low technology requirements, avoid rooting problems, scion profits from rootstock strength	Requires specialization, labour intensive, expensive, low multiplication rates
Cutting	Stems, roots or leaf sections are separated (cut) and induced to form shoots or roots	Ornamentals (succulents among others), fruit trees, forest trees	Moderate technology requirements	Needs rooting (usually in greenhouse), low multiplication rates, large plant stock needed to obtain cuttings, ageing produce rooting problems
Axillary shoot propagation	Prexisting meristems are used as explant to induce shoots spontaneously or after hormonal induction (cytokinins)	Ornamentals, fruit trees, forest trees	Moderate multiplication rates, moderate technology requirements	Need rooting, involves multiple manual operations, risk of losign genetic conformity related to phytohormones use, large plant stock needed to obtain explants
Adventitious shoot propagation	This method used *de novo* meristems produced directly or indirectly (callus) from embryos, stems, leaves	Ornamentals, fruit trees, forest trees	High multiplication rates, high technology required	Needs rooting, risk of of loosing genetic conformity related to phytohormones use, large plant stock needed to obtain explants

Plant tissue culture (PTC) is a form of vegetative propagation used for the large-scale production of plants known as micropropagation (Ahloowalia *et al*. 2004). A PTC system involves the establishment of cell, tissue or explants (meristems, leaf explants, roots or shoots) under defined media and tissue culture conditions, established proliferation cycle for a number of cell generations and the subsequent regeneration and aclimatization of plants (Larkin and Scowcroft 1981). Plant tissue culture technology offers the potential for efficiently propagating disease-free, genetically uniform and massive amounts of plants since its conception (Honda *et al*. 2001). The establishment of stable, efficient *in vitro* regeneration systems is not only important as a method for the clonally and mass propagation of economically important crops it is also a prerequisite for biotechnology and molecular breeding applications (Neelakandan and Wang 2012).

Totipotency and the developmental pathways of plant regeneration

Somatic cells within the plant contain all the genetic information necessary to create a complete and functional plant (Von Arnold 2008). This cell property is known as totipotency (i.e. the potential or inherent capacity to regenerate *in vitro* an entire plant). Plant cells are unique in that they retain totipotency and its developmental plasticity in the differentiated state (i.e. specialized plant cells) and have the ability to dedifferentiate, proliferate, and regenerate into mature plants under appropriate culture conditions (Neelakandan and Wang 2012). Plant tissue culture provides a good model system to study the mechanisms of cell totipotency (Komamine *et al.* 1992). Tissue culture techniques depend on plant cell totipotency. Micropropagated plants could be regenerated by two different morphogenetic pathways: **organogenesis** (the *in vitro* culture induction of plant organs or plants from different plant explants, cells or group of cells) or **somatic embryogenesis** (embryo formation from somatic cells and the subsequent regeneration of plants) (Peschke and Phillips 1992; Honda *et al.*2001). These plant regeneration pathways normally involve three analogous phases: **dedifferentiation** in which tissue become competent to responds to the organogenic or embryogenic stimulus, **induction** during which cells become determined to specifically form either a root, a shoot or embryos and **realization** meaning the outgrowth of an organ or embryos (De Klerk *et al.* 1996). In the regeneration pathway one particular aspect has been taken as evidence of cell totipotency, the stage of "callus formation" (Sugimoto *et al.* 2011).

Callus is defined as a mass of proliferating cells found after plant wounding and is stimulated in PTC for plant regeneration purposes. In this particular stage cells adquire similar features to meristematic cells (i.e. small cells with dense cytoplasm and small vacuoles) and the consequent loss of differentiated cell morphology leading to the assumption that cell plants deddiferentiate (Sugimoto *et al.* 2011). Recent researchs in *Arabidopsis* based on transcriptome study (RT-PCR, microarray) and microscopy analysis (FISH) have shown that callus transcriptome is more similar to the one found in root meristems than in shoot meristem or embryonic tissues. Therefore, callus formation is not a simple reprogramming process back to a dedifferentiated or embryonic state and might also be a transdifferentiation process (i.e. the switch to another cell type) instead of dedifferentiation (Sugimoto *et al.* 2010).

Plant growth regulators (PGRs) and their implication in the developmental pathways of organogenesis and somatic embryogenesis

De novo organogenesis and somatic embryogenesis (SE) are dependent on PGRs (phytohormones) for cell division and dedifferentiation; to acquire organogenetic competence; initiate organ formation and plant development (De Klerk *et al.* 1997). PGRs are synthetic or naturally produced signal molecules regulating cellular processes and widely used in tissue culture; the main PGRs are auxins, cytokinins, gibberellins and acid absisic (ABA), among other compounds. Auxins in plants are the agents responsible for the establishment of cell polarity, cell division, cell elongation, apical dominance and adventitious root formation. The original endogenous hormone is the indole-3-acetic acid (IAA) (Deo *et al.* 2010; Gutierrez- Mora *et al.* 2012). In PTC auxins are essential for the induction of SE, maintenance of defifferenciated state (proliferation of embryogenic cells) and rooting of somatic embryos at the cotyledonary stage (last developmental stage). Some of the most used synthetic auxin are indole-3-butyric acid (IBA), naphthaleneacetic acid (NAA), 2,4-dichlorophenoxyacetic acid (2,4-D). Cytokinins in plants promote cell division and differentiation, the natural cytokinins are zeatin and zeatin riboside (Dodeman *et al.* 1997; Gutierrez- Mora *et al.* 2012). Cytokinins in PTC induce adventitious shoot formation when used at high concentrations (1 to 10 mg/l), but in such conditions root formation is generally inhibited. Also, they promote axillary shoot formation by decreasing apical dominance. Some of the most used synthetic cytokinins are kinetin, Thidiazuron (TDZ), benzyladenine (6-BA) and 6-γ-γ-Dimethylallylaminopurine (2iP) (Gutierrez-Mora *et al*. 2012). Gibberellins (GA) regulate different processes of plant development (e.g. stem development and flowering). Plants produce more than 110 different kinds of GA, however in PTC the most used compounds are GA_1, GA_3, GA_4 and GA_7. In general, gibberellins induce elongation of internodes and the growth of meristems or buds *in vitro.* Therefore, GA3 is essential for meristem culture of some species (e.g. potato).

The developmental fate of the regenerating tissue (organogenesis or SE) can clasically be directed by the ratio of two plant hormones, auxin and cytokinin (Skoog and Miller 1957). In organogenesis, a high cytokinin to auxin ratio induces shoot morphogenesis (or high cytokinin with no auxin), whereas opposite low ratio results in root development. A high auxin signal (usually 2,4-D) is important to induce embryogenic cells, capable of plant regeneration in the process of SE (Komamine *et al.* 1992; Phillips 2004). This chemical regeneration pattern has been demonstrated over a wide range of plant species (Duclercq *et al*. 2011). Several other growth regulators (natural and synthetic) like oligosaccharides, jasmonate, polyamines, thidiazuron and brassinosteroids have also been reported to induce the embryogenic response (Feher *et al.* 2003; Yang and Zhang 2010). Particulary, cytokinins are also involved in the formation of embryogenic cells from the epidermis of immature zygotic embryos (Dodeman *et al.* 1997).

Somatic embryogenesis involves differentiated somatic cells acquiring embryogenic competence and proliferating as embryogenic cells (Dodeman *et al*. 1997; Fehér 2008; Yang and Zang 2010). Cells able to undergo embryo development generally appear as proembryogenic masses (PEMs) (Zimmerman *et al.* 1997). Proembryogenic masses are composed of somatic cells that completed their transition as embryogenic cells and in which no exogenous stimuli is necessary to produce somatic embryos (Komamine *et al*. 1992). Auxins are used in somatic embryogenesis for the proliferation of PEMs (i.e. for augmenting the yield of plants regenerated) and also to inhibit the development of somatic embryos (Nomura and Komamine 1995; Filonova *et al.* 2000; Yang and Zhang 2010). For these purposes both cell or embryo cultures are maintained or proliferated in a medium with similar composition of the medium used for callus induction (Dodeman *et al.* 1997; Von Arnold *et al.* 2002). Morphogenetic changes leading to the development of somatic embryos can be observed upon transferring callus to an auxin-free medium (Komamine *et al.* 1992; Zimmerman 1993). Also, auxin is rapidly taken up from the medium so the depletion of auxin in the culture medium starts after a few days. Consequently, if cultures are not regulary transferred to fresh medium the development of somatic embryos will start (Von Arnold *et al.* 2002).

De novo organogenesis and its applications

De novo organogenesis is the regeneration pathway in which shoots and roots are induced directly from plant explants (e.g. leaf, roots, buds) or indirectly from callus, (Duclerq *et al.* 2011). The organogenesis also requires the subsequent acclimatization of regenerated plants to the greenhouse before planting (Honda 2001). Nowdays *de novo* organogenesis is consider the most applied methodology for the *in vitro* plant regeneration (e.g. adventitious shoot proliferation) (Honda *et al.* 2001; Duclercq *et al.* 2011). Different organogenesis pathways induced by auxin and cytokinins can be observed in **figure I.6** Organogenesis is more often used in biotechnological breeding methods like in vitro micropropagation, haploid production and genetic engineering (Duclercq *et al.* 2011).

However, organogenesis methods are often labor-intensive and require many steps of manipulations resulting in high labour costs (Brown and Thorpe 1995). For example, adventitious shoot bud cultures (monopolar structures which in turn give rise to leaf primordial and the apical meristem) requires the selection, manual separation of shoots and induction of rooting in separated stages of the process. Organogenesis often involves more than one cell for the regeneration of one plant making this technique not apropiate to some biotechnological techniques like plant genetic transformation. Furthermore, the main disadvanatge for implement organogenesis in mass propagation programs is related to the low multiplication rates offered by this technique (Honda *et al.* 2001).

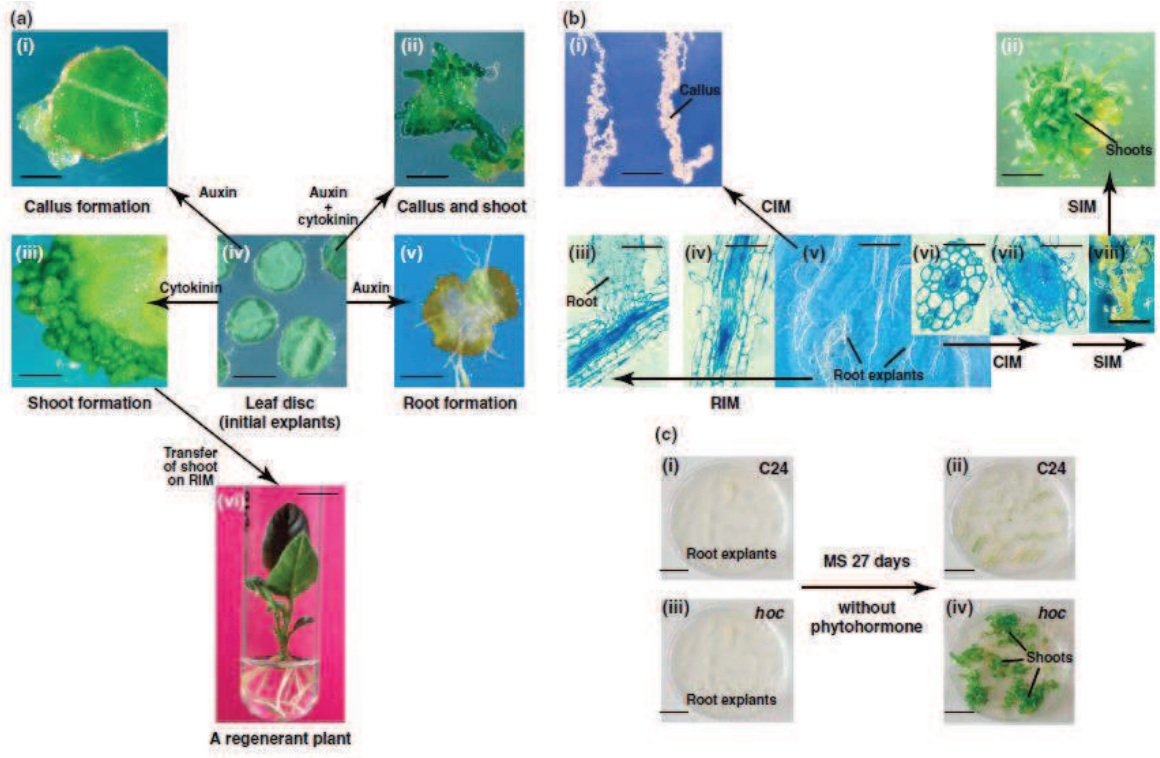


Figure I.6 Phytohormone-based de novo shoot and plant regeneration systems (Adapted from Duclercq *et al.*2011). Chemical regulation of callus, root and shoot formation from cultured leaf discs of (a) Datura and Petunia and (b) explants of *Arabidopsis*, grown on media containing different phytohormone combinations and concentrations, (c) illustration of the high shoot regeneration capacity of the hoc mutant compared with the wild-type *Arabidopsis thaliana* C24. In (a) Images i, ii, iv–vi show Datura explants and Image iii shows Petunia leaf explant. In Datura, callus developed on 2,4-D (1 mg/l) containing medium, roots on IAA (0.5 mg/l) medium, callus and shoots on 2,4-D (0.5 mg/l) and kinetin (1 mg/l) medium. In Petunia, shoots were developed on BA (1 mg/l) medium. Well-developed shoots were isolated and transferred to rooting medium (containing 0.5 mg/l IAA) to induce roots. A complete plant with well-developed roots and shoots was obtained, after a month of culture. In (b), Images i–viii show root explants of wild-type *A. thaliana* C24. In Arabidopsis, two-step media sequences (CIM and SIM) were used to obtain callus and shoots. Callus without shoot formation was obtained with 2,4-D (1 mg/l) and roots were obtained on IAA (0.5–1 mg/l) medium. In (c), Images i and ii show root explants of wild-type *A. thaliana* C24 and Images iii and iv are root explants of the hoc mutant. Note the formation of shoots on hoc roots without exogenous phytohormones, after 27 days of culture, whereas no shoots were formed on wild-type roots. CIM callus induction medium; RIM root induction medium; SIM shoot induction medium.

Somatic embryogenesis

Somatic embryogenesis is considered as the developmental process by which somatic cells (i.e. not involved in the sexual propagation process), under suitable induction conditions undergo restructuring through the embryogenic pathway to generate embryogenic cells (Yang and Zhang 2010). This developmental pathway was first described in carrot (*Daucus carota*) cell suspensions by the teams of Steward *et al*. (1958) and Reinert (1958). Plant regeneration via SE includes five common steps after explant selection: (1) induction of embryogenic cultures by culturing the primary explant on medium supplemented with PGRs, mainly auxins (Von Arnold *et al.* 2002); this step also involves the selection and physical separation of callus (Sujatha 2011), (2) the proliferation of embryogenic cultures (multiplication) on solidified or liquid medium supplemented with PGRs (Dodeman *et al*. 1997; Von Arnold *et al*. 2002, Sujatha 2011), (3) somatic embryo development in medium lacking PGRs; this inhibits proliferation and stimulates embryo formation and early development (Dodeman *et al*. 1997; Von Arnold *et al*. 2002), (4) embryo maturation by culture on medium supplemented with ABA or with reduced osmotic potential and (5) the development of plants on medium lacking PGRs, until plants are rooted and able to be transfered to sterilized soil for hardening in greenhouse. A conventional process of direct SE is presented in **Figure I.7.**

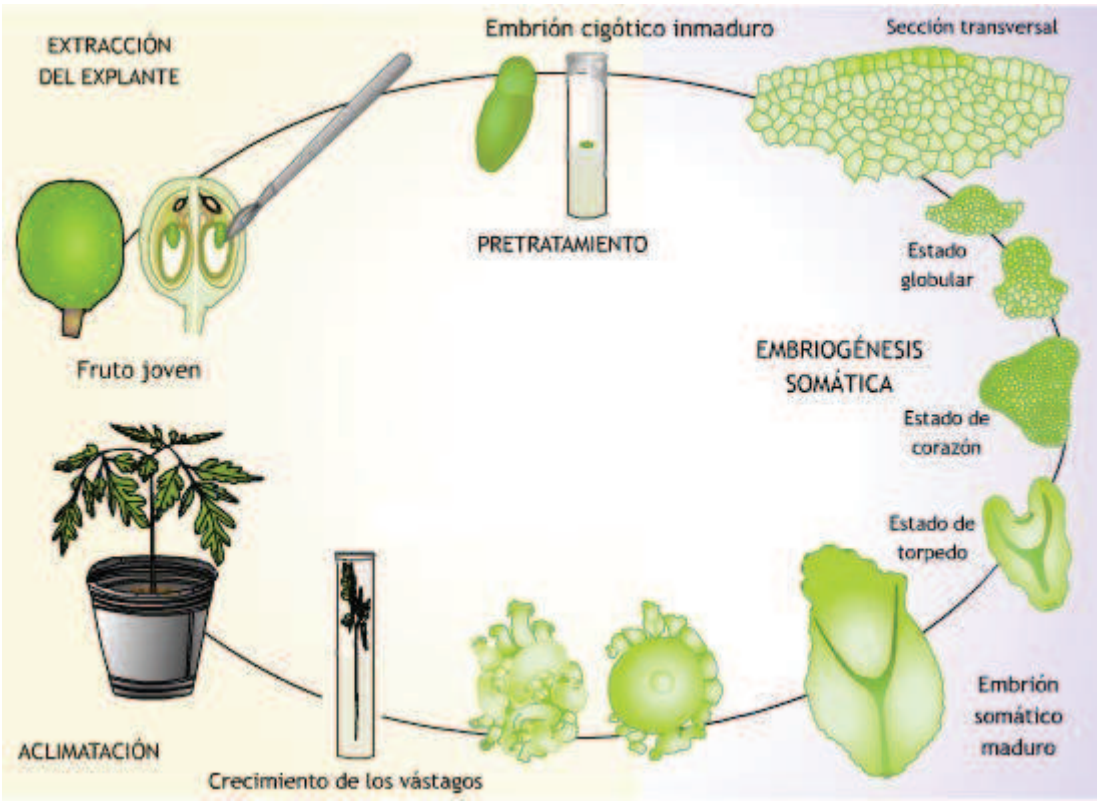


Figure I.7 Somatic embryogenesis of *Melia azedarach L*. (Available at: http://www.biologia.edu.ar).

Somatic embryos are derived either from unicelular or multicelular origin (Von Arnold *et al*. 2002; Quiroz-Figueroa *et al*. 2006; Yang and Zhang 2010). Nevertheless, not all somatic cells are capable of embryogenesis (Dodeman *et al*. 1997). The factors influencing the *in vitro* adaptability and regeneration are varied, ranging from genotype, origin of explant and culture conditions (Neelakandan and Wang 2012). Theoretically, all plants' cells are totipotent; however, it seems that the initiation of the embryogenic pathway is restricted to certain responsive cells owning the potential to activate genes involved in the generation of an embryogenic response (Feher 2008). The cells which are more competent for SE are generally those coming from juvenile tissues found in young plants and immature zygotic embryos (i.e. used in the propagation of forest trees by SE) among others (Gutierrez-Mora *et al*. 2012). It is believed that the induction of SE consists of the termination of a current gene expression pattern in the explant tissue and its replacement with embryogenic cell expression program (Quiroz-Figueroa *et al*. 2006; Von Arnold, 2008). In this respect, PGRs and abiotic stress (i.e. osmotic stress *per se* can induce SE without PGRs) could play a central role mediating the signal transduction cascade leading to the reprogramming of gene expression (Verdeil *et al*. 2007; Von Arnold 2008).

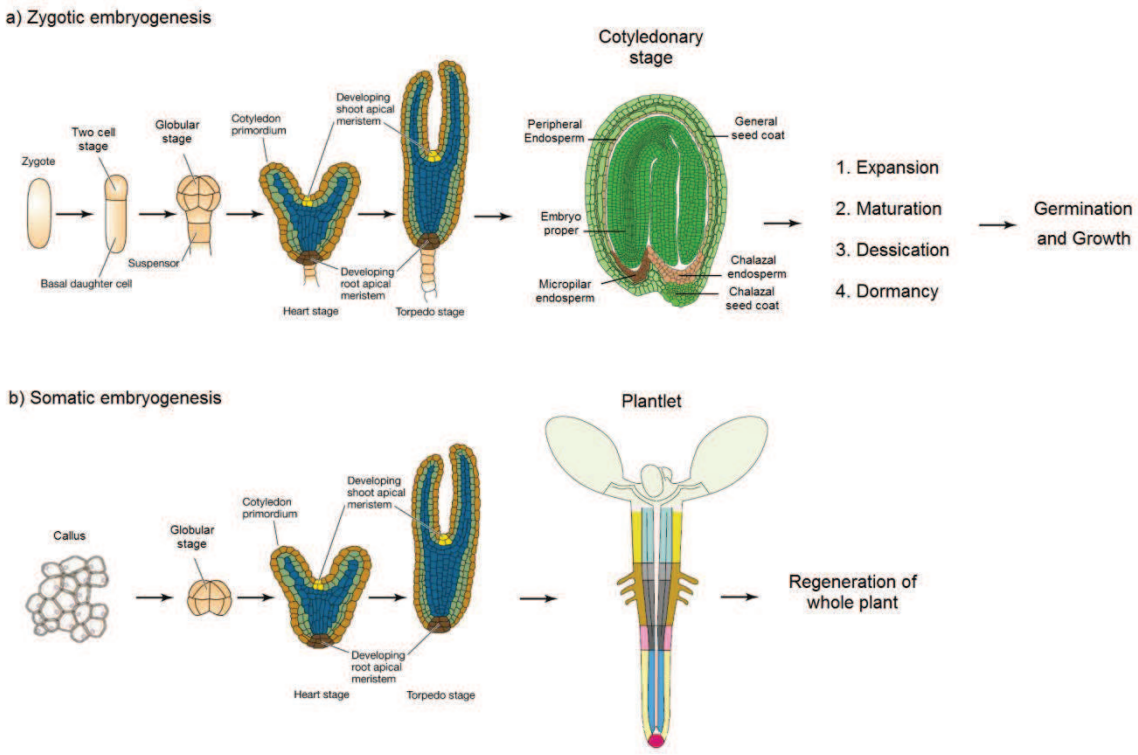
Direct and indirect somatic embryogenesis

Similarly to *de novo* organogenesis pathway, somatic embryos can differentiate either directly from the explant (direct SE) without an intervening callus phase or indirectly (indirect SE) after a callus phase involving the proliferation of embryogenic callus (Dodeman *et al*. 1997; Von Arnold *et al*. 2002; Quiroz-Figueroa *et al*. 2006; Namasivayam 2007). Direct SE is normally faster than indirect methods and could represent an advantag for the rapid release of the cloned material (e.g. 70 days in coffee). However, embryo production is less numerous than in indirect SE (Sujatha, 2011). Secondary embryogenesis (a.k.a repetitive embryogenesis) is a form of direct embryogenesis in which secondary embryo develop directly from epidermal or subepidermal cells instead of develop into a new plant. Secondary embryogenesis is of great importance in mass propagation to increase the production of plant regenerants. In indirect SE a friable embryogenic tissue is formed either in semi solid (as calli) or in liquid media (cell suspensions) from which the somatic embryos arise (Sujatha 2011). The indirect based methods use two different media: an induction media for primary callogenesis (normally auxin enriched) and a secondary regeneration medium (with a lesser auxin concentration) to produce the friable embryogenic callus.

Somatic embryogenesis *vs.* zygotic embryogenesis

Somatic embryos are defined as bipolar structures containing both shoot and root apices that are anatomical and physiological comparable to zygotic embryos (Zimmerman *et al*. 1993; Phillips 2004; Verdeil *et al*. 2007). Thus, SE systems can be used as models for studying molecular, regulatory and morphogenetic events in plant embryogenesis (Quiroz-Figueroa *et al*. 2006; Deo *et al.* 2010). Conversely some aspects of zygotic embryogenesis can be explored with the aim of optimizing the regeneration of somatic embryos. Plant development can be divided in two main steps: embryogenesis *sensu stricto* beginning with the sporophitic generation of the zygote and finishing in the embryo at the cotyledonary stage and the maturation of a seed followed by germination (Dodeman *et al*. 1997). The sporophytic generation is initiated by a double fertilization event, a process unique in angiosperms in which both male gamets (spermatic cells) participate in the act of fusion, one unites with the egg cell (female gamete) to form the diploid zygote from wich the embryo develops, while the other gamete fuses with the central cell of the embryo sac and then develops into the triploid endosperm that provides nutrients for embryo development and later for germination (Quiroz-Figueroa *et al*. 2006; Sujatha, 2011). Embryogenesis in plants can also take place from somatic cells without the involvement of gamete fusion (e.g. apomixis process derived from unfertilized eggs or maternal tissue).

The embryogenesis *sensus stricto* can be divided into the following stages: globular shaped, heart-shaped, torpedo-shaped, and cotyledonal stages in dicots; globular, scutellar and coleoptilar stages in monocots; and globular, early cotyledonary and late cotyledonary embryos in conifers (Yang and Zhang 2010). Somatic embryos pass through characteristic morphological stages of zygotic embryogenesis, which are: globular shaped, heart-shaped, torpedo-shaped and cotyledonal (Zimmerman 1993; Dodeman *et al*. 1997; Phillips 2004; Quiroz-Figueroa *et al*. 2006). At the cotyledon stage, instead of entering desiccation or dormancy as observed in zygotic embryos, the somatic embryos continue to grow into fully differentiated plantlets (Zimmerman 1993; Yang and Zhang 2010). The completion of all the stages of embryogenesis results in the regeneration of a new plant (Zimmerman 1993; Quiroz-Figueroa *et al*. 2006). In **figure I.8** shows the characteristic developmental stages in (a) the zygotic and (b) the somatic embryogenesis and their main differences.



FigureI. 8 Comparison of somatic and zygotic embryogenesis (adapted from Zimmerman, 1993).

Applications of somatic embryogenesis

SE provides a valuable alternative to other vegetative propagation systems. This technology constitutes an important and powerful tool in the process of plant genetic transformation (i.e. intended for gene functional analysis or for genetically modified organisms developement), plant breeding programs and mass propagation of selected plants. Mass propagation is of particular interest for all breeding programs in view of the rapid and massive diffusion of novell and morphoagronomic attractive cultivars. For example, the application of SE in woody plants is a promising technology and particularly waited for the programs of reforestation. Although SE has been demonstrated in a very large number of plants and trees, this technology has been restricted for the large scale commercial production to only a few plant species like carrot and date palm (Ahloowalia *et al.* 2004). Similarly to other vegetative propagation methods SE offers advantages but also disadvantages that must be overcome before being a widely accepted and applied mass propagation technique.

The advantages of somatic embryogenesis over other vegetative propagation systems are presented as follows: (1) Higher regeneration rate than organogenesis providing an alternative approach to conventional micropropagation for large-scale clonal propagation of elite cultivars (interspecific or intraspecific hybrids) (Deo *et al*. 2010). (2) Distribution of diseasee free vitroplants (e.g.virus, bacteria and fungi) by SE systems coupled with thermotherapy and chemoterapy for pathogens elimination. (3)

Smaller areas needed for stock donor plants than in organogenesis methods (used for explant recolection). For example, large areas are needed to stock plants used for axillary bud culture of forest trees. (4) As in zygotic embryos, dormancy can also be induced in somatic embryos, hence long-term storage or even the creation of artificial seeds is possible (Deo *et al*. 2010; Sujatha 2011). (5) Cryopreservation (i.e. callus is progessively frozen at -35°C and then stored in liquid nitrogen at -196 °C) can cope the lost of embryogenic potential observed in callus ageing by storing young and fully functional callus. (6) Automation of the process through the use of liquid media (Etienne and Berthouly 2002). Its advantages are presented next.

Somatic embryogenesis in bioreactors

SE in liquid media is mainly applied along with bioreactors technology. The term "bioreactor" is generally used to describe a vessel carrying out a biological reaction for the production of microbial, plant or animal secondary metabolites (e.g. alkoids, terpenoids, phenylpropanoids) (Takayama and Akita, 2006). Two main approaches are commonly implemented when using this technology: batch and continous cultures. In batch methods culture media, cells (i.e. cell suspensions) or explants (e.g. shoots, buds, among others) are kept in the culture flask and mantained by subculturing. On the contrary, in continuous cultures the media is partially or completely replaced allowing using different media formulation, for example: media for the primary induction, cell proliferation or germination. Besides mass propagation by SE the use of bioreactors offer great potential in the production of secondary metabolites using transformed hairy roots and other vegetative techniques of mass propagation like shoot organogenesis (Honda *et al.*2001; Albarran *et al.*2005). Several examples of bioreactor applications for plant propagation are listed as follows. Shoots: *Musa acuminata, Pinus radiata, Citrus, Hevea brasiliensis, Saccharum spp*, *Atropa belldona, Begonia x hiemalis, Chrysanthemum morifolium, Dianthus caryophyllus, Fragaria ananassa, Nicotiana tabacum, Petuniahybrida, Primula obconica, Zoysia japonica, Scopolia japonica, Spathiphyllum, Stevia rebaudiana, ananas* Bulbs: *Fritillaria tunbergii, Hippeastrum hybridum, Hyacinthus orientalis, Lilium.* Corms (bulbo-tuber): *Caladium sp., Colocasia esculenta, Pinellia térnate*. Tubers: *Solanum tuberosum*. Embryos or adventitious buds: *Atropa belladona, C. arabica and C. Canephora* (Albarran *et al.*2005; Takayama and Akita, 2006).

The use of bioreactor present several advantages over solid media SE systems such as (1) higher multiplication rates of cell or embryos (i.e. secondary embryogenesis) (2) During regeneration in liquid media, root and shoot formation is simultaneous, thus, eliminating the need of root induction phase as with conventional micropropagation methods (e.g. solid media SE or organogenesis) (Deo *et al.* 2010). (3) Cultures can be manipulated such that embryo formation and germination can be synchronized maximizing plant output while minimizing labour cost (Honda *et al.* 2001; Von Arnold *et al.* 2002; Sujatha 2011). (4) Liquid cultures are easier to scale-up and subculturing than solid media

based systems. (5) Cell suspension offers the possibility of automation, improving the efficiency of embryo production in a short period of time and reducing labor costs (Ibaraki and Kurata, 2001).

As previously seen the optimization of PTC processes for mass propagation are principally intended to increase the production of regenerated plants and improve the quality of the plant regenerants. Nevertheless, the major concerns of any program made to clonally propagate morphoagronomic atractive cultivars is to **reduced the cost of vitro plants** and avoid or **limit the levels of phenotypic variation (off types)**. This phenomenon is commonly known as "somaclonal variation" and their aspects will be deeply discussed in section 4.

Coffee propagation and micropropagation

Sexual and traditional vegetative techniques for coffee propagation

Coffee varieties in both cultivated species are traditionally propagated by seeds (Juma *et al*., 1994; Etienne and Bertrand, 2003). The main constraint for coffee seed propagation is associated with the uncontrolled genetic variation found in heterozygous cultivars (e.g. F1 hybrids), the short span of seed viability and the slow rates of seed production i.e. after 4 years of plant development (Juma *et al*. 1994). Arabica varieties are sold in seed form after a relatively lengthy pedigree selection process of at least 20 years (Etienne *et al*. 2005).

Coffee clonal vegetative propagation is possible by traditional (e.g. layering, grafting and cuttings) and more advanced vegetative techniques (e.g. axillary and adventious shoot propagation). Many problems have limited the use of these techniques (1) The risk of disease propagation which limits their interchange across country' borders (Etienne *et al.,* 2005). (2) The limited numbers of orthotropic cuttings that a coffee tree can produce for shoot or bud propagation (4 or 5 per plant). (3) The need to establish large stock gardens for shoot production (Sujatha et al., 2011). (4) Rooting induction on vegetative cuttings can be very difficult, particularly in Arabica, normally requiring the application of hormones (Juma *et al*., 1994). (5) Finally, the most important limiting factor is associated to the low multiplication rates resulting from shoot organogenesis, budding and microcuttings (Juma *et al*., 1994; Etienne, 2005; Sujatha, 2011). For example, Juma *et al*. (1994) reported some thousand of plants could be produced *per annum* using microcutting propagation methods. A typical Arabica plantation normally requires a high plant density (4-6.000 plants/ha) meaning that those production rates are economically insufficient to satisfy the market demand.

Somatic embryogenesis

Consequent to the absence of efficient clonal propagation technology little of the genetic progress (e.g. F1 hybrid cultivars) is currently disseminated to growers (Bertrand *et al*., 2012). As seen in

chapter 2 the recent development of highly heterozygous F1 intraspecific hybrids in *C. arabica* made necessary to implement an efficient vegetative system for their mass propagation (Etienne and Bertrand, 2003; Bertrand *et al*., 2012). Reducing the time laps needed between the creation of valuable high yielding cultivar and its rapid dissemination is possible with efficient micropropagation systems (Sujatha, 2011). In coffee, several teams are currently working in the scaling-up of coffee SE with the goal of producing several millions of *in vitro* plants per year (Sujatha, 2011). SE techniques allow the production of relatively uniform plants on a massive scale, in a reduced time and in a greater scale than other vegetative techniques. For example, in coffee SE one leaf segment can produce hundreds of plantlets in a single one month-culture cycle, whereas the microcutting method requires several sub-culturing cycles requiring a lot of handling in order to get a comparable number of plantlets (Juma *et al*., 1994). This way, segregating hybrids and transgenic plants could be efficiently scale-up for commercial purposes (Sujatha, 2011).

SE in several *Coffea* species and genotypes is well documented (Starisky, 1970; Söndhal and Sharp, 1977; Pierson *et al*., 1983; Dublin, 1984; Yasuda *et al*., 1985; Berthouly and Michaux-Ferrière, 1996; Berthouly and Etienne, 1999). Based on the yield of plant regenerants obtained, these processes are classified in low frequency SE (producing from single digits to one hundred embryos per leaf explant within a period of 2 months) and high frequency SE (HFSE) processes (producing from one hundred to thousands embryos per leaf explants) (Sondahl and Sharp, 1977). A two step protocol using two different mediums (i.e. conditioning and induction of SE) succesfully resulted in the first description of a HFSE induction and subsequent plantlets development in Arabica cultivars (Söndhal and Sharp, 1977). For Arabica, tissue culture conditions for SE implemented the use of a primary culture on auxin-containing medium for callus induction (2.26 µM of 2, 4-D, 8.69 µM of 2iP and 4.92 µM of IBA) followed by a second culture medium for embryo induction with 2, 4-D and BA (4.52 µM and 17.75 µM respectively). Regeneration of coffee plants was achieved in a medium additoned only with cytokinin (4.4 or 17.6 µM BA) (Van Boxtel and Berthouly, 1996). The embryogenic potential of *C. canephora* can be maintained for 2 years using 5 g/L 6-BA and without any auxin supply. From this embryogenic culture 1 g of Robusta or Arabusta callus produced 1.2-1.9 $x10^{-5}$ embryos after 8-10 weeks in liquid medium regeneration.

Normally, HFSE methods take about 7-8 months from leaf explants to somatic embryos in Robusta cultivars and Arabusta and 9-10 months for Arabica cultivars. The use of lower concentration of 2, 4-D and 2iP in the induction medium was reported to increase the rate of embryos production and reduce the generation time of embryo formation at least 3 month in both *C. arabica* and *C. canephora* (Samson *et al*., 2006). In *C. canephora*, the use of the synthetic auxin 2, 4-D strongly increases HFSE (over 60%) when used in combination with kinetin or 6-BA in primary cultures. Coffee embryogenic calli used in indirect SE protocols are cytologically heterogeneous, composed of small isodiametric

cells (10–15 μm in diameter) that frequently divide during early stages of SE (Santana *et al*., 2007). These developmental stages (presented in **Fgure I.9)** include the classical preglobular, globular, heart-shaped, torpedo shaped, and cotyledonary of zygotic embryo development.

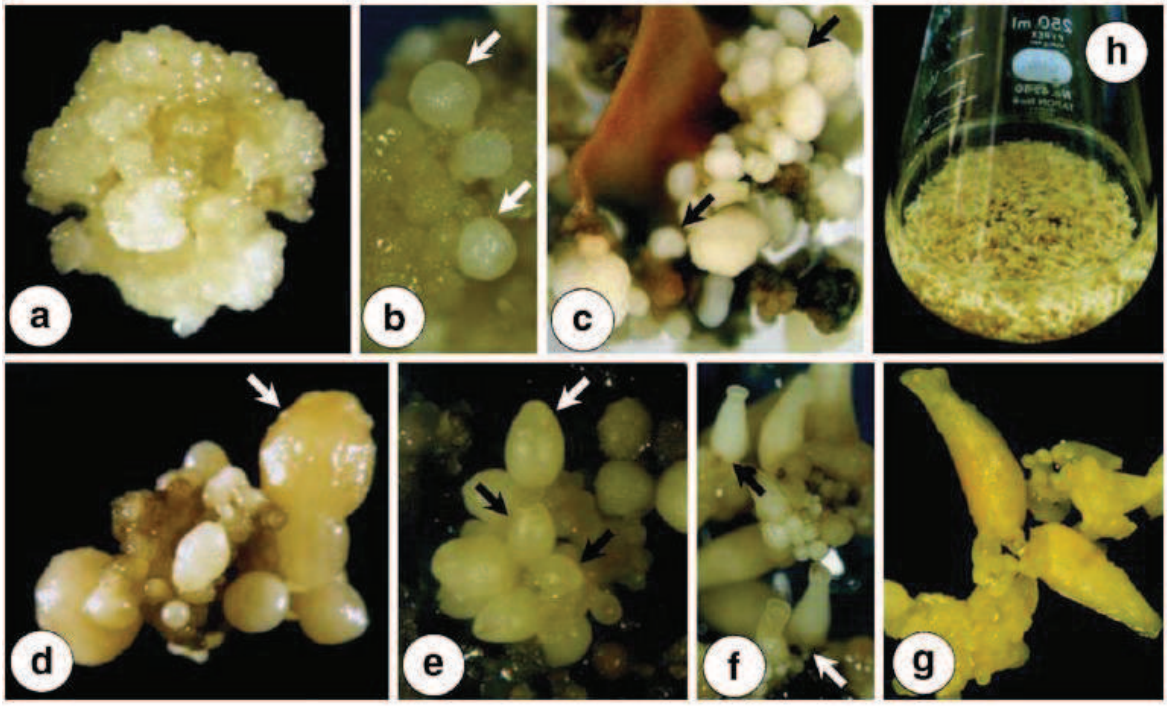


Figure I.9 Developmental stages of coffee somatic embryos obtained by indirect embryogenesis. (a) Embryogenic callus; (b) globular embryos, with arrows showing globular embryos emerging from callus; (c) globular embryos, with arrows showing globular embryos emerging directly from leaf explants; (d) oblong embryos (see arrow); (e) a cluster of heart-shaped embryos (black arrows) and elongated embryos (white arrows); (f) torpedo stage embryos, with arrows indicating suspensor-like structure; (g) cotyledonary stage embryos; (h) somatic embryogenesis from suspension cultures (Quiroz-Figueroa et al. 2002).

Cell suspension cultures have been established in coffee for a number of different purposes, particularly as an interphase for obtaining somatic embryos, for protoplast isolation and to increase the yield of plant regenerants in mass propagation protocols as a proliferation step. Using the media described by Yasuda *et al.* (1985) embryogenic calli can be induced and multiplied in most of the Robusta genotypes using 6-benzylaminopurine (6-BA) as the only PGR (Ducos *et al*., 2007). On the contrary, auxin (i.e. mostly 2,4-D) is generally required in *C. arabica* for blocking the embryogenic tissues at the undifferentiated state and inhibiting precocious embryo development during the multiplication step. As in other species coffee somatic embryos are regenerated directly from the explant or indirectly from the embryogenic callus. The depletion and/or decrease of auxin supply is usually essential for further somatic embryo development. In coffee propagation the use of direct SE has been considered as more suitable for *C. canephora* whereas indirect is more adapted to *C. arabica*. The use of direct methods of SE reduces the time for complete plant regeneration but the number of

plants produced is much lower. A maximum of 144.1±7.3 and 68.7±3.3 embryos per explant were produced rapidly through direct SE within a period of 2 months in *C. canephora* and *C. arabica,* respectively (Giridhar *et a*l., 2004a,b). For both cultivated species direct SE can be obtained with low auxin levels or without any auxin supply. The HFSE process enables the use of a liquid medium for both embryogenic tissue proliferation and embryo regeneration in Erlenmeyer flasks or bioreactors (Etienne et al. 2006). The recent development of temporary immersion bioreactors allowed for the first time in C. arabica to obtain the mass production of fully germinated and acclimatizable somatic embryos in liquid medium (Barry-Etienne et al. 1999). The indirect process was consequently preferred for scaling-up and development of industrial propagation procedures using bioreactors technology.

Scaling up of coffee micropropagation

Coffee SE is now considered as a model system for perennial species (Giridhar *et al.,* 2004). Pionner works were realized in coffee since twenty years on the development of new bioreactors and strategies for large-scale production of somatic embryos (Barry-Etienne et al. 1999, Ducos *et al*., 2007, Etienne and Berthouly 2002). Mass somatic embryo production has been successfully carried out in liquid medium using bioreactors, Erlenmeyer flask and even special disposable bags for the two main cultivated species (Starisky and Van Hasselt, 1980; Zamarripa *et al.,*1991; Ducos *et al*., 1993; Van Boxtel and Berthouly, 1996; Barry-Ettiene et al., 1999). The different types of bioreactors for coffee micropropagation initially comprised stirred bioreactors (Ducos et *al*. 1993b; Noriega and Söndhal, 1993; Ducos *et al.,* 1999) and more recently temporary immersion bioreactors like 1 L RITA [R] "Récipient à Immersion Temporaire Automatique" (Berthouly *et al., 1*995; Etienne *et al*., 1997; Etienne-Barry *et al.,* 1999), 10-L glass bottles (Ducos *et al*., 2007), 10-L horizontal disposable plastic bags (Ducos *et al*., 2007). Very high embryo yields could be obtained with these liquid culture systems. For example, using Erlenmeyer flasks in *C. canephora* Zamarripa *et al.* (1991) obtained 200,000 somatic embryos from 1 g of embryogenic tissue Ducos *et al*. (1993) obtained 600,000 embryos/l medium in a 3 L stirred bioreactor. With *C. arabica*, lower yields of embryos were obtained under such conditions compared with *C. canephora* and Arabusta. Zamarripa et al. (1991) and Zamarripa (1993) reported yields of about 46,000 embryos/3L Erlenmeyer flask or 20,000 embryos/L in a bioreactor. Mass production of *C. arabica* cv 'Catuai' somatic embryos in a 5-L stirred bioreactor were reported by Noriega and Söndhal (1993). A total yield of 45,000 embryos were recovered within 3 months. More recently, De Feria *et al.* (2003) produced *C. arabica* cv 'Catimor' embryos in 2-L bioreactors with a rate of 70,000 embryo/g fresh weight callus (Ducos *et al*., 2007). The current production capacity of the Neslté research Center in Tours is 2.5–3.0 million embryos per year. The multiplication of embryogenic cells is a key step because it greatly and rapidly scales up the number of potential embryos to be produced (Ducos *et al*., 2007). Bioreactors immersion techniques

with specially adapted culture vessels RITA^{R} allow the direct regeneration in the same container and without subculture of plantlets from cell suspensions. This small bioreactor has been used to grow embryogenic calli (Berthouly et *al.* 1995) and torpedo-stage embryos of ten *C. arabica* F1 hybrids (Etienne *et al.* 1997). Depending on genotype, yields ranging from 15 to 50,000 acclimatizable pre-germinated somatic embryos per g of embryogenic cells were recorded starting with a cell density of 1 g L-1 (i.e. 200 mg per system). This way, SE methods using temporary immersion technology offered the best potential to multiply F1 hybrids developed in breeding programs (Etienne 2005). This innovation also enables the mass and virtually synchronous production of germinated somatic embryos, minimizing the need for embryo selection before the acclimatization step (Etienne *et al*. 1997). Plant conversion rates for germinated somatic embryos (i.e. having a pair of open cotyledons, along with a well-developed chlorophyllous embryonic axis) frequently reach 60% in the nursery. It was shown that 86% of embryos in the same 1l-RITA® bioreactor reached the "germinated" stage and could be directly sown (Barry-Etienne *et al*., 2002). A major trend in commercial laboratories is the bulk cultivation of small propagules in photoautotrophic conditions, i.e. in sugar-free medium with enriched CO2 and high light intensity (Ducos *et al*., 2007) or photomixotropic conditions (additioned with very low concentrations of sugar). This method has been succesfully used for *in vitro* acclimatization in special vessel (VITRO^{TM}) in plants like strawberry, sweet potato, eucalyptus and papaya among others (Teixeira Da Silva, 2007). An improved process has been developed using a 10-l disposable bioreactor consisting in a bag containing a rigid plastic box (''Box-in-Bag'' bioreactor), insuring, amongst other advantages, a higher light transmittance to the biomass due to its horizontal design (Ducos *et al*., 2008). Recently, the use of a new bigger 5 l-MATIS® bioreactor, designed horizontally (area 355 cm^2) especially to favour embryo dispersion and light transmittance to SE, markedly improved the embryo-to-plantlet conversion rate (91%) (Etienne *et al,.* 2013).

PART IV

Somaclonal variation

Definition and extent of the problem

Most tissue culture studies were focused on optimization of culture conditions in view of enhanced propagation. Since the begining micropropagation methods based on tissue culture were conceived as a novell and atractive technologies for the clonal and rapid disemination of elite cultivars. However, the conditions that are optimal for plant multiplication may not be optimum for maintaining the genetic integrity of the cloned genotype (Ziauddin and Kasha 1990). The phenotypic variation observed among the plant regenerants derived from any form of tissue culture is defined as somaclonal variation (Larkin and Scowcroft 1981). Similarly, the term somaclonal variants is used to referr to all the off-types or phenotypic variants derived from any type of tissue culture process. Somaclonal variation (SV) was first described by the Hawaiian Sugar Planters' Association in their annual reports from 1967 to 1970 and later by Heinz in (1969, 1971 and 1973) as the morphogenetic, cytological and isozyme affectation of Sugarcane cultivars. Although first perceived as artifacts of tissue culture, further studies considered these offtypes as promising and useful variation profitable in breeding programs (Larkin and Scowcroft 1981). Nevertheless, in commercial micropropagation wherein the regenerant population is expected to be exact copies of the required genotype, the loss of genetic fidelity is a major concern (Larkin and Scowcroft 1981; Kaeppler *et al*. 2000; Miguel and Marun 2011; Wang and Wang 2012; Neelakandan and Wang 2012).

Qualitative and quantitative trait affectation by somaclonal variation

Somaclonal variants present a broad range of phenotypic changes, sometimes even affecting important agronomic traits (i.e. a distinguish feature). The affectation of traits has been described for both qualitative and quantitative features (Kaeppler *et al*. 2000). The variation of qualitative traits are widespread and not restricted to any culture system; some of them not seriously affecting the agronomic performance of plants (e.g. variegation or leaf abnormalities) (Nehra *et al.*1992; Kaeppler and Phillips 1993; Etienne and Bertrand 2003; Sato *et al*. 2011). However, some phenotypic changes can be detrimental or lethal for plants, especially when they affect the normal plant architecture. These qualitative variants are unable to adapt and develop in greenhouse or field; some examples include the multistem *C. arabica* variant reported by Etienne and Bertrand (2003) or the hooked stem phenotype found in white spruce pinus (*Picea glauca*), reported to be dead after 6 months in greenhouse (Tremblay et al. 1999). On the other hand, quantitative trait variations are frequently presented in somaclonal variants, in a general way positively or negatively altering the agronomic performance of crop species.

These alterations could affect plant height, plant biomass, grain yield and weight, content of solids in fruit and essential oils. Although this type of variation is more difficult to measure, in phenotypic aberrant regenerants the tendency is towards poorer agronomic performance (Nehra *et al*. 1992; Karp 1995; Kaeppler and Phillips 1993; Etienne and Bertrand 2003; Biswas *et al*. 2009). As seen, SV can have both negative (most frequent) and positive effects in plants. The generation of useful somaclonal variants has made tissue culture techniques atractive for plant breeding programs. Somaclonal variants showing better or novell agronomic traits have been found for both qualitative (e.g. the ormamentals Geranium cv. Velvet rose) (Duncan 1997) and quantitative/qualitative cultivars (e.g. Grand Naine and Cavendish banana cultivars) (Sahijram *et al*. 2003). Many somaclonal variants commercial cultivars were presented in the reviews made by Skirvin *et al*. (1994), Duncan (1997) and Jain (2001) and are summarized in **Table I.4.** Nevertheless, the use of SV in breeding programs has presented certain disadvantages: (1) most of the time somaclonal variants are not novell or useful (i.e. aberrant phenotypes), (2) the variation generated could be unstable or not reproducible (Karp 1995; Duncan 1997, Jain *et al*. 2001), (3) although some variants show positive changes other traits could be altered in a negative way (Karp 1994).

Table I.4 Somaclone variants released as new cultivars (Adapted from Duncan 1997 and Jain 2001).

Selected somaclones
Disease resistant
Banana, *Fusarium* wilt resistance, commercial cultivar
Celery, *Fusarium* yellows resistance *(F. oxysporum)* line MSU-SHK5
Celery, insect resistance (*Spodoptera exigua*), Fusarium yellow R somaclones K-26, K-108, & K-128 were also insect resistant, being integrated in breeding
Tomato, *Fusarium* R, variety DNAP-17
Rice, Picularia resistance, improved cooking quality, released cultivar DAMA
Rice, *Rhizoctonia* resistance, germplasm released LSBR-33 LSBR-5
Cyanodon dactylon, fall armyworm (Spodoptera) resistance P1572566 = Brazos-R3
Celery, *Fusarium* R, released cv UC- TC
Quantitative (high yield)
Wheat, high yield, released as new cv 'He Zu No.8
Lathyrus sativus, reduced neurotoxin in feed grain, high yield, early maturing, new cultivar released as P-24
Tomato, tomato high solid contents, var DNAP9
Cymbopogon (aromatic grass, 50–60% increased oil yield released cv CIMAP/Bio-13
Brassica juncea, high yield, shattering resistant, cv released Pusa Jai Kisan
Qualitative varieties
Blackberry, thornless, released in NZ as cv Lincoln Logan
Potato, non-browning, released cv 'White Baron'
Capsicum, yellow fruit, cv Bell sweet
Flax, salt & heat tolerance in the same somaclone, released as cv ANDRO
Sweet potato, cv Scarlet
Ornamentals
Torenia foumieri, white flowers, compact plant released cv Uconn White
Paulownia tomentosa, cv Snowstorm
Abiotic stress tolerance

Rice, submergence tolerance, excellent tol, line FR13A
Sorghum bicolor cv GAC, acid soil tolerance
Wheat heat and drought stress tolerance
Brassica tolerance to salt
Male sterile
Haemerocallis spp., Dwarf, short flowers, male sterile, released cv Yellow Tinkerbell

Somaclonal variation rates

The rates of SV in regenerated plants arise at greater frequencies than attributable to spontaneous mutations (Karp 1991; Skirvin *et al*. 1994; Duncan 1997). SV that specifically affect phenotypic traits vary in their frequency depending on the species, cultivars and micropropagation systems. For example the rate of SV found in bananas cv. Grande Naine was higher in shoot-tip derived cultures (5.3%) as compared to plant regenerants derived from SE (0.5 and 3.6%) (Shchukin *et al*. 1997). In any micropropagation program a normal SV frequency could be stablished between 1 and 5% per produccion cycle (Skirvin *et al.*1994; Côte *et al.* 2000; Sahijram et al. 2003). At this matter, in any micropropagation program 3-5% of phenotypic variation could be considered acceptable with the exception of banana where up to 10% is considered aceptable i.e. as practiced by commercial outfits (Sahijram *et al*. 2003). Unfortunally, sometimes the rates of variation could be strinkingly high or produce severe abnormal phenotypes not easily detected in early stages of plant development. These types of SV are unacceptable in any propagation system and result detrimental to any PTC method. For example, in 1990 Smith and Drew reported that in Australia the 90% of the *in vitro*-derived banana were off-types plants presenting the condition dwarf "choke-throat". Once in field this abnormality prevents the bunch from emerging from the banana shoot producing a closely packed and undersized fruit which affects plant productivity.

Genetic mechanisms associated to SV

Genetic changes were frequently associated with *in vitro* regenerated plants (Neelakandan and Wang 2012). Mutations are defined as heritable changes in the DNA (meiotically stable) not derived from genetic segregation or recombination (Van Harten 1998). Mutations affect the primary sequence of DNA and include numerical and structural chromosome changes, somatic recombination, point mutations, deletion and transpositions also ocurring in mitochondrial and chloroplast genomes (Orton 1983; Karp, 1991; Duncan 1997; Kaeppler *et al*. 2000; Neelakandan and Wang 2012). All of these classes of mutations would be expected to give rise to stable, sexually heritable variation with the possible exception of TEs (Orton 1983). These mechanisms are also implicated with allele inactivation (Larkin and Scowcroft 1981; D'amato 1985; Duncan, 1997; Kaeppler *et al*. 2000; Neelakandan and Wang, 2012). DNA sequence variations such as single base substitutions (SBS) and small InDels (i.e. base insertions or deletions) are predominant in progenies generated by tissue culture (Jiang *et al*.

2011; Carrier *et al*. 2012). Single base mutation can alter genes if they correspond to nonsynonymous mutations (i.e. the change of a nucleotide base in a given codon not specific of the required aminoacid) therefore altering the final protein.Therefore, genetic mutations can produce a great diversity of abnormal phenotypic traits. For example, in *Arabidopsis* somatic embryos-derived plants SBS and InDels resulted in different array of phenotypes including bleached and long hypocotyls, dwarfish plants, late flowering and large flower plants (Jiang *et al*. 2011).

In tissue culture, some phenotypic variants usually present similar phenotypes to previously described seed derived-mutants of their corresponding species (Phillips *et al* 1994). These observation could lead to the presumption that similar mutational mechanisms could be involved. Nevertheless, the term "**somaclonal variant**" is normally used to define all the phenotypic variants derived from plant tissue culture instead of the word "**mutants**" (Kaeppler *et al*. 2000). This differentiation is made because most of the time the nature of mechanisms producing phenotypic variation (i.e. stable or unstable) or the segregation of altered traits remains unknown (Kaeppler *et al*. 2000; Wang and Wang 2012). Several factor complicate the determination of the heritability (segregation) of altered traits after continued selfing and crossing: (1) Altered traits could arise in the form of homozygous and heterozygous mutations or in the form of dominant and recesive forms making complicated their observation (Karp 1991). (2) Sexual and ploidy imcompatibilities block the generation of progeny (Skirvin et al. 1994). (3) Finally not all the forms of SV are meiotically heritable or stable, specially the variation of "**epigenetic**" origin (Wang and Wang 2012). The epigenetic aspects of SV will be deeply discused in following sections.

Heritable SV involves either single or multiple genes from changes in DNA sequence including point mutations, polyploidy, aneuploidy, chromosomal rearrangements and the movement of transposable elements (Orton 1983; Kaeppler *et al*. 2000). This type of variation stable through the sexual cycle could be mantained with repeated asexual propagation (Skirvin *et al*. 1994). These somaclonal variants could be unequivocally referred as mutants instead of somaclonal variants. The analysis of the stability of the tissue culture-induced changes in selfed progenies of regenerated plants has demonstrated that many of the changes are stably inherited and therefore caused by genetic mutations (Wang and Wang 2012). Studies made by Jiang et al. (2011) reported that in SE Arabidopsis regenerants 6 out 28 selfed R1 families showed segregation of the abnormal traits, such as late flowering and long hypocotyle (presented in **Figure I.10).** Furthermore and as previosly mentioned, these somaclonal variants were correlated to the outcome of SBS and InDels (Jiamg *et al*.2011).

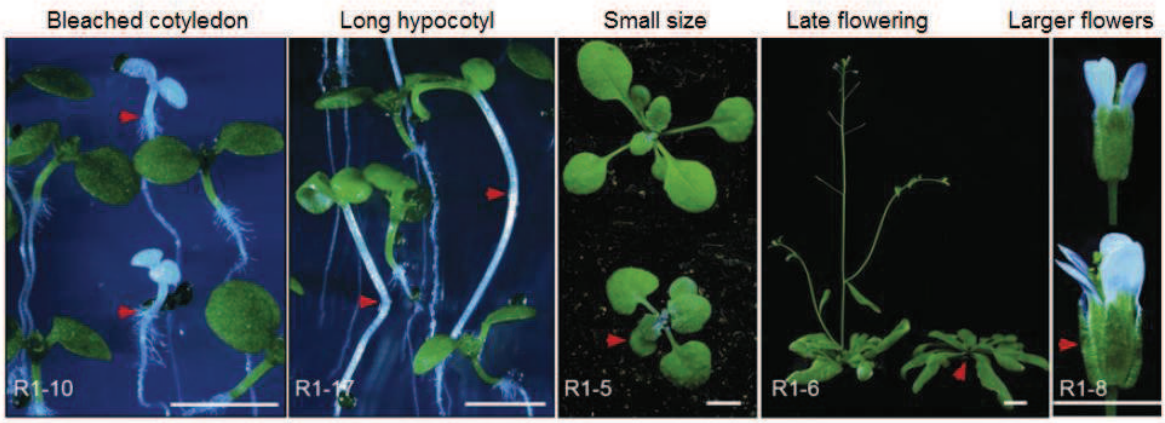


Figure I.10 Selected mutants (R1 plants) from Arabidopsis R0 lineages. Segregating variant plants or organs are highlighted with red arrows (Adapted from Jiang et al. 2011).

Cytological mechanisms associated with SV

The balance in gene dosage is essential for normal function of most eukaryotic genomes. An imbalance can cause severe phenotypic syndromes in both plants and animals (Birchler and Veitia 2007). Among the heritable types of variation, chromosome deletions, translocations and changes in ploidy have been encountered (i.e. ploidy refers to the normal number of chromosomal sets in a given cell) (D'Amato 1985). Cytogenetic changes such as variation in ploidy level, number of chromosomes, and structural changes represent major alterations to the genome and they are often generated during *in vitro* proliferation and differentiation (D'amato 1985; Lee and Phillips 1988; Kaeppler *et al.*2000; Neelakandan and Wang 2012). In PTC the production of cytogenetic abnormalities is specially found in long term callus cultures (Hao and Deng 2001; Smýkal *et al*. 2007; Rodriguez-López *et al*. 2010) in plant regenerants of a variety of species including strawberry, wheat, citrus, black spruce and banana (Nehra et al. 1992; Henry *et al*. 1996; Tremblay *et al*. 1999; Hao and Deng; 2001; Jambhale *et al*. 2001).

Karyotypic instability observed during *in vitro* culture includes chromosome breakage, deletions, translocations and inversions (Jain 2001). Ploidy level variations also occur in PTC with polyploidy generally occurring more frequently than aneuploidy (Kaeppler *et al.*2000). The production of mitotic irregularities in PTC underlies the occurrence of chromosomal variations (Larkin and Scowcroft 1981; Peschke and Phillips 1992). The use of high concentrations of PGRs can enhance the frequency of ploidy changes (Skirvin et al. 1994). Polyploidy in tissue culture is generally produced by endoreduplication or nuclear fusion (Bayliss 1973; Sunderland 1977; Kaeppler *et al*. 2000). Endoreduplication is produced when the nuclear genome continues to replicate without the normally succeeding cell division, leading to elevated nuclear gene content and polyploidy (Weber *et al*. 2008).

Unequal chromosome distribution, involving the distribution of replicated chromosomes into only one daughter cell also results in polyploidy (Lee and Phillips 1988).

Aberrant cell divisions during mitosis are relatively frequent in natural population leading to the gain or lost of chromosomes during meiosis (Harrison and Schwarzacher 2011). A change in chromosome number that is not the exact multiple of the haploid karyotype is known as aneuploidy (Torres et al. 2008). Aneuploidy involves the loss (monosomy) or gain (trisomy) of one or more specific chromosomes or large chromosomal segments (segmental aneuploidy) and results in a dosage imbalance of genes on the affected chromosome (Makarevitch and Harris 2009). Aneuploidy is found *in vitro*, not infrequently during the first phases of callus induction and supension cultures (Gözükirmizi et al. 1990; Hao and Deng 2001; Kumar and Mathur 2004; Giorgetti et al. 2011). It is produced through nuclear fragmentation (amitosis) followed by mitosis or by defective chromosome bahavior during mitosis (D'Amato 1985). These alterations found in natural and *in vitro* produced plant populations include multipolar anaphases spindles, non-disjunction (i.e. the failure of sister chromatids to separate during Anaphase I and II at meiosis or Anaphase at mitosis), aberrant spindles, lagging chromosomes, non-congression and chromosomes breakage (i.e. producing dicentric or acenctric chromosomes) (Sunderland 1977; D'Amato 1985; Kaeppler et al. 2000). Chromosome breakages may generate mutations directly through positional effects or through alterations in gene expression due to chromosomal rearrangement (Orton 1983; Peschke and Phillips 1992; Duncan 1997). A comparative analysis in oat and maize showed that chromosome breakage occurred more frequently than ploidy changes (Kaeppler et al. 1998). Translocations were the most frequent chromosomal abnormality observed with inversions and insertions/deletions also occurring (Peschke and Phillips 1992; Kaeppler et al. 2000). Different types of chromosome aberrations in regenerants from different PTC techniques are presented in **Table I.5**. As previously discussed, the most common abnormalities were chromosomal rearrangements whereas the occurrence of aneuploidy and polyploidy plants were less common.

Table I.5 Summary of chromosome aberrations observed in regenerant plants. Adapted from Lee and Phillips, (1988).

Species	Number of regenerated plants	Analysis	Aberration per plant					
			Polyploidy	Aneuploidy	Translocation	Deficiency	Duplication	Other *
Avena sativa	799	Meiotic	0	43	48	180	0	2
Zea mays	370	Meiotic	6	0	23	13	0	1
	267	Meiotic	2	1	45	59	0	1
	257	Meiotic	37	14	32	36	0	0
	142	Meiotic	1	1	7	0	0	0
	110	Meiotic	1	1	0	0	0	0
Triticosecale	51	Meiotic	0	12	3	29	0	0
Wheat-rye hybrid	10	Meiotic	0	0	4	10	6	0

Other include inversions, centric fusions and more complex unclassified rearrengements.

Aneuploid individuals frequently appear spontaneously within polyploid plant populations (e.g. maize); presumably due to a failure to partition equally the multiple chromosomes set at meiosis (Henry et al. 2010). Many plant aneuploids grow to generate adult plants, not least because plant genomes are often polyploid and have higher plasticity and mechanisms for gene dosage compensation (Harrison and Schwarzacher 2011). Aneuploidy results in gene dosage imbalance and often causes severe phenotypic alterations in plants and animals (Makarevitch and Harris 2009). Aneuploid plants exhibit a a variety of phenotypic sindromes, including developmental delays, partial sterility and alteration in plant architecture i.e in maize segmental aneuploids display leaf knotting, partial tassel sterility, late flowering time and overall changes in the plant architecture (Makarevitch *et al.* 2008). The effect of the dosage of specific chromosome types on traits is additive and could be used to predict the observed phenotype (Henry *et al*. 2010). **Table I.6** presents a description of the developmental abnormalities observed in *Arabidopsis* seedlings and the type of chromosomal abnormalities related to them. The morphological characteristics of these abnormalities can be seen in Figure **I.11.**

Table I.6 Description of phenotypic abnormalities observed in *Arabidopsis* aneuploids and the relation with different types of aneuploidy. Adapted from Henry et al. (2010).

Code	Trait	Description of abnormal phenotype	Chromosome abnormality
A	Rosette shape	Abnormal rossetes	3x + chromosomes 3 and 5
B	-	-	3x + chromosomes, 3 – chr. 5
C	-	-	2x + chromosome 2
D	-	-	2x + chromosomes 3
E	None	Whole rossetes wild type (E)	2n=2x=10*
F	Abnormal architecture	Aerial rossetes	3x – chromosome 5
G	-	-	2x + chromosome 5
H	-	-	2x + chromosomes 1, 2, and 3
I	-	-	2x + chromosome 3 and 4
J	Flower in axil	Direct conversion of an axillary meristem to a floral meristem.	2x + chromosomes 1 and 2
K	-	-	2x + chromosome 3
L	Irregular spacing	Periods of failed elongation resulting in disorganized and compacted nodes followed by longer than normal internodes.	2x + chromosomes 1, 2, and 3
M	-	-	2x + chromosome 5
N	Curly leaves	Rolled blades of cauline leaves.	2x + chromosome 5
O	Fasciation	Gross morphological evidence for radial stem growth resulting in divergence of the vasculature and bifurcation of the meristem.	2x + chromosome 3
P	Nubbin	Angular projection/bend in stem often with light irregular growth at position. Frequently found at the base of a secondary stem or immediately following or preceding a node	4x + chromosome 4
Q	Empty axils	Apparent lack of axillary buds at the basis of a cauline leaf.	2x + chromosome 3

R	None	Diploid cauline leaf showing wild-type stem and leaf trichomes.	2n=2x=10*
S	Hairy	Presence of higher density of trichomes on both the adaxial and abaxial leaf surfaces as well as on the stem.	2x + chromosome 4.

* Normal wild type

Figure I.11 Phenotypic alteration in *Arabidopsis thaliana* aneuploid seedlings.

Epigenetic mechanisms associated with SV

Epigenetic changes involve heritable, but potentially reversible, chemical modifications in genes without changing the primary DNA sequence (Kaeppler et al. 2000; Feschote et al. 2002; Saze 2008; Ahmad et al. 2010; Law and Jacobsen 2010). The first observation of somaclonal variation was reported in 1981 by Larkin and Scowcroft. Since than, it is considered to be caused either by genetic or epigenetic mechanisms. The discrimination between both genetic or epigenetic mechanisms is not easy to determinate (Karp 1991). The stability of qualitative traits mutations through sexual generations could indicate their molecular basis (Kaeppler and Phillips 1993). Unfortunately not all variation is stably inherited and some traits are lost on passage through sexual cycle or could be of epigenetic origin (i.e. reverted transposable elements insertions) (Kaeppler and Phillips 1993). Epigenetic modifications are mostly found in DNA (methylation) and histones (acetylation, methylation) and are associated with changes in the gene expression (Zhang et al. 2010; Vanyushin and Ashapkin 2011). Patterns of cytosine methylation are inherited from cell to cell through the action of maintenance methyltransferase enzymes on the symmetric CG and CHG and the asymmetric CHH contexts (Vaughn *et al*. 2007). Epigenetic regulation of gene expression is accomplished by DNA methylation, histone modifications, histone variants, chromatin remodeling, and may also involve small RNAs (Ahmad *et al*. 2010). The epigenetic modifications can mediate both a short-term (mitotic) or long-term (meiotic) transmission of an active or silent gene (Saze 2008). Changes in the DNA methylation patterns are thought to be a major source of epigenetic variation leading to SV (Kaeppler and Phillips 1993; Phillips *et al*. 1994; Kaeppler *et a*l. 2000). Hypomethylation of plant DNA can lead to abnormal development (e.g. *Arabidopsis*), finally affecting plant phenotype (Saze 2008; Mirouze and Paszkowski 2011). Another important fact is that DNA methylation is a key mechanism for silencing transposons; most transposable elements in plants and animals are methylated (Law and Jacobsen 2010). For this reason cytosine methylation is one of the most studied epigenetic mechanism (Kaeppler *et al*. 2000; Miguel and Marum 2011).

Cytosine methylation

The pattern of DNA methylation and the chromatin state are known to be modulated during plant and organ development or tissue differentiation (Furner and Matkze 2010; Law and Jacobsen 2010; Miguel and Marum 2011). Cytosine methylation is a major epigenetic mark involved in gene expression also playing an important role in plant regulation and development (Zhang et al. 2010; Vanyushin and Ashapkin 2011). Epigenetic variation in regenerated plants (i.e. referring to cytosine methylation) has been reported to be more frequent and variable than genetic SV (Miguel and Marum 2011). A possible explanation for the higher frequency of epigenetic changes found in PTC could be related to an imprecise resetting of the cytosine methylation patterns in plant regeneration, different to the normal patterns established in meiosis and fertilization (Kaeppler et al. 2000). During *in vitro*

culture, a significant alteration or resetting in the global methylation level of genomic DNA could be associated with the culture age, growth regulator type and concentration (LoSchiavo et al. 1989). For example, these authors showed that global methylation levels decreased with increasing concentration of **kinetin**, but increased with increasing amounts of the auxin **2,4-D**.

DNA methylation in plants commonly occurs at cytosine bases in the form of 5-Methylcytosine in all sequence contexts, the symmetric CG and CHG, and the asymmetric CHH contexts in which H denotes A, T or C (Cokus *et al*. 2008; Zhang *et al*. 2010; Law and Jacobsen 2010; Vanuyushin and Ashapkin 2011; Neelakandan and Wang 2012). Methylation at CG dinucleotides is catalysed by DNA methyltransferasease1 (MET1). The CG site is symmetrical on the opposite strand and MET1 can potentially maintain a pattern of methylation during DNA replication by a semiconservative mechanism (see **Figure 4.4**). Methylation in the context CHG is mediated by Chromomethylase3 (CMT3). Methylation in the non-symmetric CHH context results largely from the activities of the Domains Rearranged Methyltransferases (DRM1 and DRM2) with some contribution from CMT3 (Law and Jacobsen 2010; Furner and Matzke 2010). The alteration of the normal level of cytosine methylation could be detrimental for plant development resulting in phenotypic variation (Kaeppler and Phillips, 1993; Smulders and Klerk; 2010). That is the case of ***Arabidopsis*** *ddm1* mutants, characterized by abnormally low methylation levels (70% of reduction). These *Arabidopsis* mutants present abnormal developmental patterns resulting in a wide range of abnormalities which include leaf structure defects, altered flowering time and abnormal flower structures (Kakutani et al. 1996). Methylation can enhance both qualitative and quantitative trait variation because several genes can be affected simultaneously (Duncan 1997). Maintenance of methylation activity is necessary to preserve DNA methylation after every cellular DNA replication cycle. DNA modification by methylation are mediated by DNA cytosine methyltransferase enzymes belonging to chromomethylase (CMT), domains rearranged methyltransferase (DRM), and methyltransferase (MET) families (Law and Jacobsen 2010; Zhang et al. 2010; Vanuyushin and Ashapkin 2011). The proteins involved in DNA de novo methylation, maintenance of methylation and demethylation of the model plants *A. thaliana* are presented in **figure I.12.**

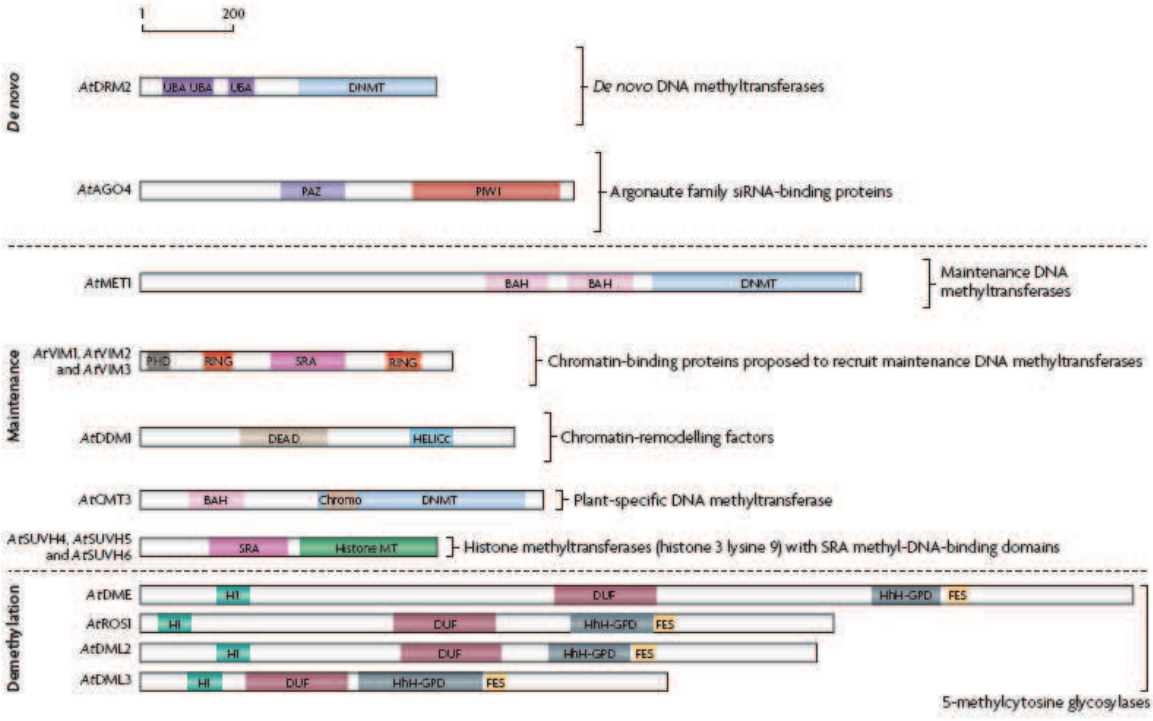


Figure I.12 Proteins involved in de novo DNA methylation, maintenance methylation and demethylation of A. thaliana. Adapted from Law and Jacobsen (2010).

A model describing the mechanisms involved in the maintenance of CG methylation during replication and the maintenance of CHG methylation in plants is presented in **Figure I.13**. The DNA methyltransferase 1 (DNMT1) is proposed to be recruited to replication foci through interactions with ubiquitin-like plant homeodomain and RING finger domain 1 (UHRF1) a SET or RING-associated (SRA) domain protein that specifically interacts with hemimethylated DNA and with proliferating cell nuclear antigen (PCNA) (Saze 2008). After being recruited, DNMT1 functions to maintain methylation patterns by restoring the hemimethylated DNA to a fully methylated state **(Figure I.13** section a). In plants, DNA METHYLTRANSFERASE 1 (MET1, also known as DMT1) and the VARIANT IN METHYLATION (VIM, also known as ORTHRUS) family of SRA domain proteins, which are homologues of DNMT1 and UHRF1, respectively, are likely to function in a similar manner to maintain CG methylation patterns. Black and white circles represent methylated and unmethylated cytosines respectively. A reinforcing loop of DNA and histone methylation is proposed to maintain CHG methylation in plants **(Figure I.13** section b). The CHROMOMETHYLASE 3 (CMT3) DNA methyltransferase maintains methylation in the CHG context, which is recognized by the SRA domain of the SUPPRESSOR OF VARIEGATION 3-9 HOMOLOGUE 4 (SUVH4, also known as KYP) histone methyltransferase (histone MT). SUVH4 catalyses histone 3 lysine 9 dimethylation (H3K9me2), a modification that is required for the maintenance of CHG methylation, and the

chromodomain of CMT3 binds methylated H3 tails (Henderson and Jacobsen 2007; Law and Jacobsen 2010).

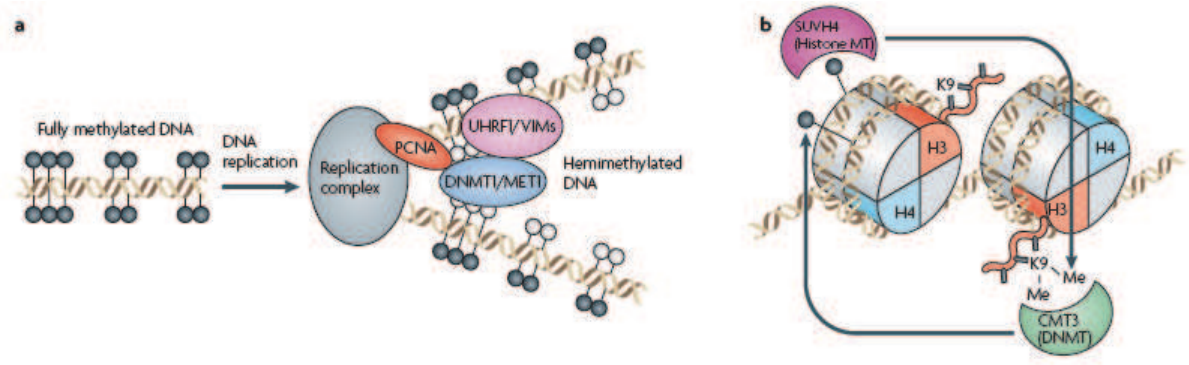


Figure I.13 Maintenance of DNA methylation in plants and mammals. a) Model depicting the maintenance of CG methylation during replication. b) Model depicting the maintenance of CHG methylation in plants (Law and Jacobsen 2010).

Other forms of DNA methylation

5-Methylcytosine remains as the best described form of DNA methylation; however, other genetic contexts have been described with this chemical modification including N6-methyladenosine (m^6A) and N4-methylcytosine (m^4C). N6-methyladenosine is a ubiquitous base modification found internally in the mRNA of most eukaryotes (Bodi *et al*. 2012). Similarly to cytosine methylation the alteration of the levels of adenine methylation during *Arabidopsis* development (90% reduction) gives rise to plants with altered growth patterns including reduced apical dominance, abnormal organ definition, and increased number of trichome branches (Bodi *et al*. 2012). Other modified bases, such as 5-hydroxymethylcytosine (hm^5C) and 5-hydroxymethyluracil (hm^5U), have also been described but their implications in epigenetic regulation is yet to be investigated.

Small interfering RNAs (siRNAs) and RNA-Directed DNA methylation (RdDM)

Chromatin structure is critically affected by the interplay among epigenetic mechanisms such as DNA methylation, histone modifications, and RNA interference (RNAi) (Henderson and Jacobsen 2007; Huettel *et al*. 2007). Regulation of these mechanisms influences gene expression by modifying the access to the underlying genetic information, ultimately affecting phenotypes (Miguel and Marum 2010). Histones are the main protein components of chromatin. The four corehistones, H2A, H2B, H3 and H4 forming a globular octameric complex called nucleosome upon which DNA is wrapped. Histones usually present chemical modifications like methylation or acetylation (Law and Jacobsen 2010). The loss of histone H3 methylation on Lys9 and, in some cases, gain of H3 trimethylation at Lys4 or multiacetylation of H3.1, H3.2 and H4 has been noted in cell suspension cultures of *Arabidopsis* (Tanurdzic *et al.* 2008) and *Solanum tuberosum* (Law and Shuttle 2005). An independent

class of small interfering RNAs (siRNAs), 24–26 nucleotides long, usually originates from TEs and tandem repeats. They are capable of transcriptional gene silencing by targeting specific DNA and histone sequences for methylation and heterochromatinisation via a specialized RNA-dependent DNA methylation (RdDM) pathway. These RNA silencing pathways are vital to the negative regulation of several transcription factor genes, repetitive elements, mobile elements and viruses, which are essential for maintenance of genome stability and survival (Almeida and Allshire 2005). The activation of RdDM promotes histone deacetylation and methylation changes that lead to the establishment of chromatin structures that repress transcription (Law and Jacobsen 2010; Federoff 2012). Small, non-protein-coding, regulatory RNAs are emerging as key players governing epigenetic processes in plants. MicroRNAs and trans-acting small RNAs, are short, usually 21–22 nucleotides in length, and typically mediate post-transcriptional gene silencing by mRNA degradation owing to near-perfect complementarity and cleavage, or by repression of translation. Insertion of a transposable element (TE) in the reverse orientation in the genome is believed to result in the production of an antisense transcript, followed by doublestranded RNA formation, thereby triggering the RNA interference (RNAi) phenomenon (Neelakandan and Wang 2012). Furthermore, mutations in both argonaute and dicer-family proteins cause the reactivation of TEs in many eukaryotic species. TEs give rise to numerous siRNAs in most species, including the recently reported TE siRNAs in humans, and TE-siRNA levels are correlated with element activity (Slotkin and Martienssen 2007). Transcriptional repression of TEs, such as class 2 transposons, is associated with the production of 24 nt sRNAs and the DNA methylation of the corresponding genomic sequence and the histone H3K9me2 (Tanurdzic *et al.* 2008).

Activation of transposable elements (TEs)

Transposable elements (TEs) are fragments of DNA that can insert into new chromosomal locations often making duplicate copies of themselves in the process (Feschote *et al*. 2002). Transposable elements can be transcriptionally (Hirochika 1996; Komatsu *et al*. 2003; Mirouze *et al*. 2009) and developmentally active (Komatsu *et al*. 2003). Despite the fact that they are widespread and abundant (they constitute the 76% and 80 % of maize and barley genome respectively) only a few retrotransposons have been found active in plants (Feschotte *et al*. 2002; Todorovska 2007). Some examples of active TEs include *Tnt1* (Grandbastien *et al*. 1989), *Tto1* (Hirochika *et al*. 2000), *Tos17* (Hirochika 1997), and *LORE1* and *LORE2* (Fukai *et al*. 2008). Active TEs are highly mutagenic, often targeting protein coding genes for insertion, causing chromosome breakage, illegitimate recombination and genome rearrangement (Hirochika *et al*. 1996; Slotkin and Martienssen 2007). TEs insertions can affect single genes, either by disrupting their function or influencing their regulation (Peschke *et al*. 1987). The transposition of TEs near or within genic regions can alter gene expression leading to altered phenotypes by both genetic and epigenetic mechanisms (Kashkush *et al*. 2003). Only a small

part of spontaneous mutations in plants has been shown to be caused by retrotransposons (Hirochika *et al*. 2000).

The structural features and classification of TEs are presented in **figure I.14.** For both Class I (RNA) and class II (DNA) is the element-encoded transcript (mRNA), and not the element itself that forms the transposition intermediate (Feschote et al. 2002). Each group of TEs contains autonomous and non-autonomous elements depending on their ability to produce the proteins necessary for transposition (Feschote *et al*. 2002; Hua-Van *et al*. 2005; Slotkin and Martienssen 2007; Todorovska, 2007). Autonomous elements have open reading frames (ORFs) that encode the products required for transposition (red boxes). Non-autonomous elements that are able to transpose have no significant coding capacity but retain the *cis*-sequences necessary for transposition found in the element Long Tandem Repeat (5' LTR) whereas 3' LTR provides terminator and polyadenilation signals (Feschote *et al*. 2002; Todorovska 2007).

Class 2 elements DNA transposons (**Figure I.14 section a**) have terminal inverted-repeat (TIRs) (black triangles). Transposons do not require a reverse-transcription step to integrate into the genome. The transposon-encoded protein (transposase) recognizes the TIRs that flank the TE, excises the TE out of the donor position, and then integrates the transposon into the new genome location in cut and paste transposition (Feschote *et al.* 2002; Hua-Van *et al.* 2005; Slotkin and Martienssen 2007; Todorovska 2007). Class I elements can be divided into two types based on the transposition mechanism and structure (Feschote *et al.* 2002; Slotkin and Martienssen 2007). Retrotransposons undergo duplicative transposition, as their total number increases after each transposition with the potential to expand genomes (Slotkin and Martienssen 2007). These types of elements are predominant in plants (Wicker *et al.* 2007). LTR retrotransposons (**Figure I.14 section b**) have long terminal repeats (LTRs) in direct orientation (black triangles). Autonomous elements contain at least two genes closely related to retriviral proteins called gag (capsid-like) and pol which encodes several proteins and enzymes (protease, reverse transcriptase, RNase H and integrase) (Feschote *et al.* 2002; Slotkin and Martienssen 2007; Todorovska 2007). Non-LTR retrotransposons (**Figure I.14 group c**) are divided into long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINEs). This elements present different coding regions including ORF1, a gag-like protein; EN, endonuclease; and RT, reversetranscriptase but do not present LTR. Both LINEs and SINEs terminate by a simple sequence repeat, usually poly (A). All SINEs described so far are characterized by an internal RNA pol III promoter near the 5′ end (black stripes).

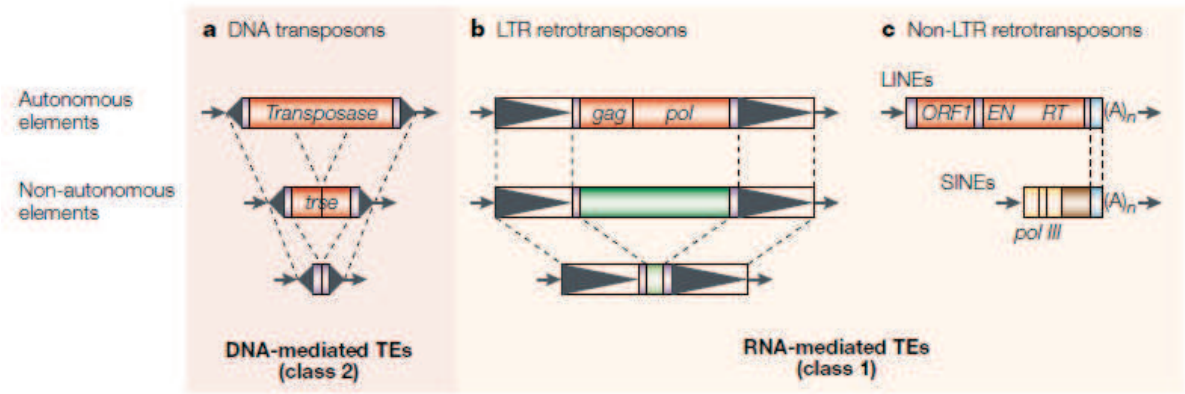


Figure I.14 Structural features and classification of plants transposable elements (Feschote *et al*. 2002).

Eukaryotic TEs transpose by different mechanisms. Class I retrotransposons (a.k.a. retroelements) transpose through an RNA intermediate in a mechanism replicative. Class II DNA transposons, transpose through DNA in a "cut and paste" mechanims in which no copy of the element remains at the excision site (Feschote *et al*. 2002; Hua-Van *et al.* 2005; Slotkin and Martienssen 2007; Todorovska 2007). Non LTR retroptransposon (Class II) transpose by the mechanism known as target-site primed reverse transcription (Levin and Moran 20011). The mechanisms of transposon mobilization are presented in **figure I.15**.

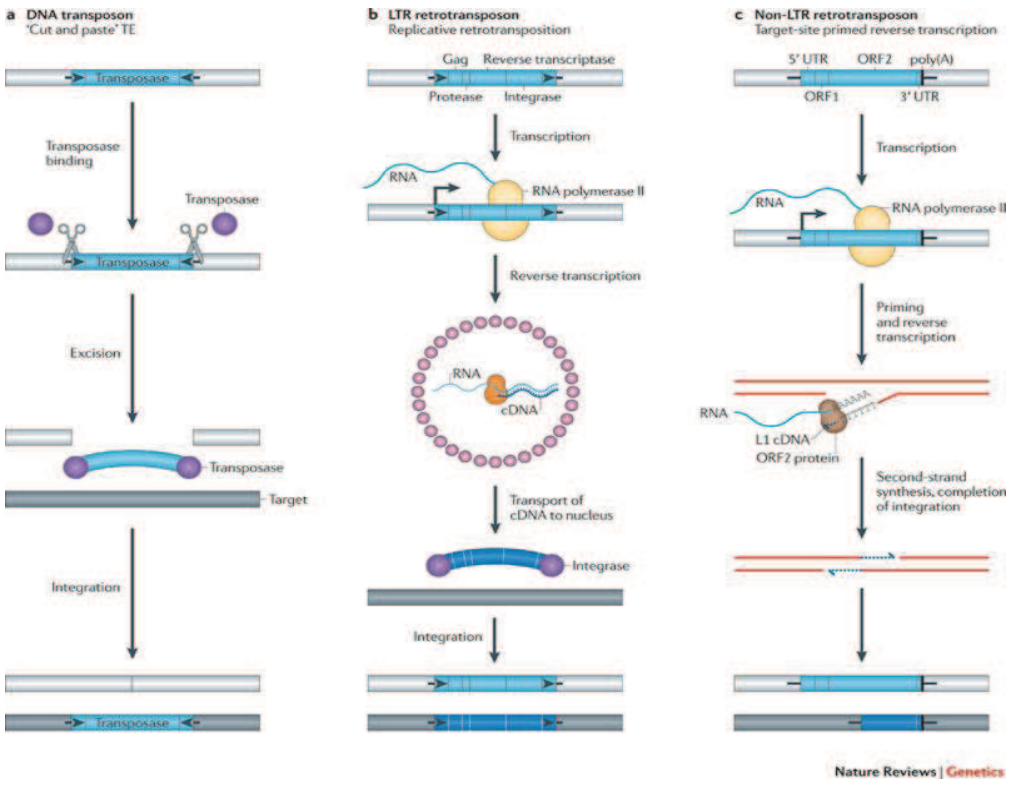


Figure I.15 Different mechanisms of transposon mobilization. Different mechanisms of transposon mobilization (Levin and Moran 2011).

Most TEs are dormant during normal growth and plant development (Feschotte *et al.* 2002). TEs are mobilized under specific conditions such as biotic (Grandbastien *et al.* 1998; Kashkush *et al.* 2003) and abiotic stress (Hirochika 1993; Kalendar *et al.* 2000). Although mutations and deletions abolish TEs transposition, some full-length autonomous TEs remain intact but silent (cryptic elements) in host genomes (Slotkin and Martiensen 2007). Eukariotic genomes have developed different epigenetic mechanisms to suppress their activity. TEs are targeted by the concerted silencing action of all DNA methylation mechanims including the novel RNAi and RdDM (Slotkin and Martienssen 2007; Law and Jacobson 2010). Transcription is the first step in the transposition of retrotransposon. However, most plant TEs are highly methylated and almost not transcribed (Hirochika *et al.* 2000). TEs reactivation is mostly associated with cytosine hypomethylation (Hirochika 2000; Miura *et al.* 2001; Barret *et al.* 2006; Smith *et al.* 2012). In *Arabidopsis* the *ddm1*mutation causes a 70% reduction in the genomic levels of 5-methylcytosine (Hirochika *et al.* 2000). Dormant endogenous *Tto1* retrotransposon (Hirochika 1996) and CACTA family elements (Miura *et al.* 2001; Kato *et al.* 2004) become transciptionally active and produce a new spectrum of insertions in *Arabidopsis ddm1* mutants (Miura *et al.* 2001). Methylation of the promoter region of the transposase gene in maize Activator (Ac), Suppressor-mutator (Spm) or MuDR reduces both the expression of the transposase and the frequency of transposition (Neelakandan and Wang 2012). Transcription reduction of *Tto1, Ttnt1*, and *Tos17* was correlated to their methylation status in *Arabidopsis* and rice (Hirochika *et al.* 2000; Liu *et al.* 2004; Perez-Hormaeche *et al.* 2008). However, transcriptional activation does not always imply new TE insertions in the genome. Copia-type retrotransposon EVD "évadé" is transcriptionally supressed by CG methylation and post-transcriptionally supported by RdDM in *met1 Arabidopsis* mutants (Mirouze and Paszkowski 2009). Chromatin modifications can also suppress TEs transcriptional activity including histone tails modifications, alterations in chromatin packing and chromatin condensation (Slotkin and Martienssen 2007). Silent TEs in plants are inherited from one generation to the next keeping the same methylation status (Miura *et al.* 2001; Vaughn *et al.* 2007).

Factors influencing SV

Many factors are intrinsically related to SV: the nature of explant, the genotype, the type and concentration of growth regulators, the time of culture i.e. the number of subculture cycles (Karp 1991, Duncan, 1997; Sahijram *et al.* 2003; Bairu *et al.* 2011).

Explant source

The use of explants from highly differentiated tissues such as roots, leaves and stems resulted in a higher frequency of variants when compared to meristem culture (Duncan, 1997). It is believed that the induction of a dedifferentiated state leads to major rates of variation (Peschke and Phillips, 1992). Hence variation is less likely to be observed from preformed shoots (axillary buds, shoot tips, and meristems) than from explants that have no preformed shoot meristems, such as leaves, roots, or

protoplasts (Karp, 1991; Duncan 1997). Explant tissue is generally assumed to be in a state of genetic uniformity, leading to the assumption that spontaneous genetic instability arise after *in vitro* culture (Orton 1983; D'Amato 1985). However, supporting data indicates that SV appears to result from both pre-existing genetic variation within the explants (Larkin and Scowcroft 1981; Evans et al.1984; Peschke and phillips, 1992; Skirvin et al. 1994) and variation induced de novo during the tissue culture (Duncan, 1997; Sato *et al.* 2011). Pre-existing genetic variation within the explants is commonly found in some explants and can be related to different states of ploidy including polysomaty (both polyploid and diploid cells that exist in the same tissue), chimerisms (organisms composed of two genetically different types of tissue) among other cytogenetic variations (D'Amato 1985; Duncan 1997). For example, sectored (chimeric) and non sectored chromosomal abnormalities in maize explants, taken from the same plant and used for organogenesis, were efficiently traced in their derived regenerants (Peschke and Phillips 1992). Other example include the outcome of variegation in plant tissue culture, mostly related to the preexisting abnormality in the chloroplast or with active transposon like in the case of the ornamental *Saintpaulia* (Sato *et al.* 2011).

Genotype

The rates of SV differ even when the same explant, genotype and conditions are used. However, certain genotypes are more susceptible to produce SV (Karp 1991; Kaeppler and Phillips 1993). For example, in modern bananas (*Musa* spp.) plants derived from ancestors *Musa acuminata* (the 'A' genome) show higher variability than plants derived from the ancestor *M. balbisiona* (known as the 'B' genome). Furthermore tissue culture in banana plants with the ABB genome affected different traits; the frequency of SV producing fruits malformation (hands per bunch and atrophied fruits) was 10.4% and 19.6% in plant height variation , whereas plants derived from the Americani (group AAA, subgroup Cavendish) showed only variation in plant height (Sahijram *et al.* 2003).

Growth regulators

As previously seen in section 3, the use of PGRs are essential in many stages of *in vitro* culture like callus induction or undifferentiated cell proliferation for augmenting plant yield. Unfortunately, the use of PGRs, particularly 2,4-D and 6-BA also favor the occurrence of SV (Evans 1988; Karp 1991; Duncan 1997; Kaeppler *et al.* 2000; Bairu *et al.* 2011; Neelakandan and Wang 2012). The specific mechanisms of PGRs on variation remains unclear, however, most reviews agree that the overall concentrations and the time of culture on PGRs enriched media increase the rate of variation (Karp 1991; Skirvin *et al.* 1994; Duncan 1997; Kaeppler *et al*. 2000; Bairu *et al.* 2011; Neelakandan and Wang 2012).

The induction of callus using high 2, 4-D concentrations has been implicated as a cause of SV in strawberry (Nehra *et al.* 1992). Auxins can produce a rapid and desorganized growth during callus

induction that may lead to genetic instability through asynchronous cell division affecting chromosomes structure or even ploidy levels (Lee and phillips 1988; Karp 1991; Neelakandan and Wang, 2012). Auxins can also cause cytological abnormalities by disrupting the cell cycle control leading to DNA synthesis and endoreduplication (Peschke and Phillips 1992). The synthetic auxin 2,4-D is not transportable out of the cells and is believed to facilitate a meristematic state by altering the endogenous auxin gradient (Morris 2000). Several other adversal effects related to the use of 2,4-D have been described in callus and plant tissue culture : (1) 2, 4-D causes an increased DNA, RNA and protein synthesis in plants, especially in meristematic tissues (Chinalia *et al*. 2007), (2) the induction of chromosome aberrations in meristematic mitotic cells of *Allium cepa* (Ateeq *et al.* 2002), (3) some degree of genotoxicity (mutagenic damage in DNA) at chromosomal level has been reported under high concentrations of 2, 4-D (Mohandas and Grant 1972). Other synthetic auxins were also implicated in SV occurrence including indolebutyric acid (IBA) and naphthaleneacetic acid (NAA). For instance, the use of IBA and GA3 (0.1 and 0.5 mg/L respectively) in meristem culture successfully produced useful SV for cultivar generation purposes in strawberry (Biswas *et al.* 2009). Also, long term cell cultures of Barley (12 months) under 0.1 mg/L of NAA produced higher percentages of polyploid cells than cultures using 2, 4-D and the same culture conditions (8.7 vs 2.7% respectively) (Ziauddin and Kasha 1990). Nevertheless, in the same study high concentrations of 2, 4-D (4 mg/L) produced higher percentage of polyploid cells in comparison to elevated concentrations of NAA (10 mg/L) (36 vs 15.3% respectively).

Cytokinins are also implicated in the production of SV and cytological abnormalities. The most used cytokinin in micropropagation is 6-benzylaminopurine (6-BA). In banana commercial micropropagation high concentrations of this cytokinin (5, 10 and 15 mg/L) were shown to induce genetic variability. Also the use of 6-BA and adenine induced variation in the chromosome number of banana (Sahijram *et al.* 2003). High concentrations of 6-BA (2 mg/L) and kinetin (0.5 mg/L) have been succesfully used to induce SV in strawberry breeding programs using shoot organogenesis (Biswas *et al*. 2009).

Culture time and long-term cultures

The length of time that a *in vitro* culture has been maintained is among the most important factors involved in inducing SV (Skirvin et al. 1994). The number and duration of subcultures especially in cell suspension and callus cultures also increased the rate of variation (Bairu et al. 2006). For example, an increased percentage of phenotypic variants were noted between the second and twelfh shoot organogenesis subculture of strawberry resulting in 3.8% variants from 943 plants and 10.4% variant from 1154 plants, respectively (Biswas *et al.* 2009). Time of culture presumably enhance SV due to an

accumulative effect of mutations over time (Peschke and Phillips 1992; Lee and Phillips 1998; Duncan, 1997). Cultures proliferating at excessive rates show more variation than those grown at moderate rates (Skirvin *et al.* 1994). Two posibilities have been hypothesized to explain an increased rate of mutation ocurring in long term cultures: a sequential accumulation of mutations over time and an augmented rate of mutation over time (Duncan 1997). Nevertheless, studies made in rice and maize regenerants have shown that a sequential accumulation of mutation is produced rather than an increased mutation rate over time (Kaeppler *et al.* 1998).

Plants regenerant derived from-long term cultures often developed several types of chromosome structural changes and sometimes major chromosomal abnormalities like aneuploidy (D'Amato 1991; Duncan 1997). In banana it has been reported that the number of somaclonal variants increased with a simultaneous decrease in the multiplication rate of propagules (Sahijram *et al.* 2003). The reduction or completely loss of regeneration ability is also commonly observed in strawberry long-term callus cultures or cell lines after 6 months culture (Nehra *et al.* 1992). The loss of regeneration ability is normally associated to major chromosomal events like aneuploidy (Orzechowska *et al.* 2012). Phenotypic abnormalities increase in frequency or/and severity in long-term cultures. This phenomenon was well characterized with coffee suspension cultures (Etienne and Bertrand 2003). Another example is the flower malformation known as mantled appearing in a frequency between 3 to 67% depending on culture time and cultivar (Duval *et al*. 1995).

Techniques used for the detection of cytological, genetic and epigenetic changes in tissue culture derived plants

Molecular markers to uncover genetic and epigenetic changes

A lack of any phenotypic variation in plant regenerants after tissue culture does not necessarily imply genetic or epigenetic conformity (Bednarek et al. 2007). The confirmation of trueness to type is a major concern in large scale clonal propagation (Miguel and Marum 2011). Plant regenerants are normally evaluated at genetic and epigenetic levels by DNA molecular markers showing different array of characteristics. In **Table I.7** are summarized the characteristics, advantages and disadvantages of different molecular markers techniques**.** Furthermore the use of molecular markers offers the opportunity to perform different types of anlysis like linkage mapping, population genetics and phylogenetics among others (Meudt and Clarke, 2007). The most common molecular markers techniques used for the evaluation of plant regenerants include restriction fragment length polymorphism (RFLPs), simple sequence repeats (SSRs) and amplified fragment length polymorphisms (AFLPs) (Vos *et al.*1995). The detection of polymorphisms associated to

retrotransposons activity or other sequences of interest is also possible with some AFLP adaptation like SSAP (Waugh *et al*.1997; Knox *et al.*2009) and differential display markers.

Table I.7 Classification of molecular markers systems (adapted from Jones et al. 2010).

Marker system	Advantages	Disadvantages
First-generation markers based on restriction fragment detection		
Restriction Fragment Length Polymorphism (RFLP)	Co-dominant; **highly reproducible**	**Low multiplex ratio***; **high on time/labour**.
Second-generation markers based on PCR		
Random Amplified Polymorphic DNA (RAPD)	Low on time/labour; medium multiplex ratio*.	Dominant; **low reproducibility.**
Amplified Fragment Length Polymorphism (AFLP).	Moderate to high reproducibility; commonly used; **high muliplex ratio***	Dominant; moderate time/labour
Methylation Sensitive Amplified Polymorhisms (MSAP)	**High reproducibility**; high muliplex ratio; detect some methylated states of the symetric sequence 5'-C$^{\downarrow}$CGG-3'	Dominant; moderate time/labour/some methylated states from CCGG sequence are not detected (see Reyna-Lopez et al. 1997).
Methylation AFLP (Met-AFLP)	**High reproducibility**; high muliplex ratio; detect some methylated states of the symetric sequence 5'-GGTAC$^{\downarrow}$C-3' (Bednarek et al. 2007).	Dominant; moderate time/labour/**not commonly used**
Sequence Specific Amplification Polymorphism (SSAP)	Applicable for targeting any gene, TE (Transposable Element) or sequence of interest (Waugh et al. 1997).	Sequence must be known to enable design of TE specific PCR primers; **moderate reproducibility**; dominant; moderate time/labour
Retrotransposon Microsatellite Amplified Polymorhisms (REMAP)	Co-dominant; Applicable for any TE of interest; low time/labour	**Low multiplex ratio**; sequence must be known to enable design of TE specific and µSAT PCR primers; **moderate reproducibility**;
Simple Sequence Repeat (SSR or Microsatellite)	Co-dominant; highly reproducible; commonly used; low time and labour	High cost of development; sequence must be known to enable design of specific µSAT PCR primers; low multiplex ratio
Variable Number Tandem Repeat (minisatellite) (VNTR)	Numerous multiallelic loci.	Low-resolution fingerprints in plants.
Third-generation markers based on DNA sequencing		
Single Nucleotide Polymorphism (SNP)	Common; evenly distributed; detection easily automated; high throughput; low assay cost; useful for association studies; commonly used; **high multiplex ratio***	Usually only two alleles present; moderate time/labour if not automated.
Genome scanning for expressed genes		
Expressed Sequence Tag (EST)	Easy to collect and sequence; reveals novel transcripts; good representation of transcripts.	Error-prone; isolation of mRNA may be difficult.
Sequence-Related Amplified Polymorphism (SRAP)	Simplicity; high throughput; numerous co-dominant markers; high reproducibility; targets coding sequences; detects multiple loci without previous knowledge of sequence information; PCR products directly sequenced.	Detects co-dominant and dominant markers, which can lead to complexity; null alleles detected directly.

Markers using array technology		
Microarrays (arrangements of small spots of DNA fixed to glass slides)	Whole-genome scanning; high-throughput technology; genotype–phenotype relationship; expression analysis of large numbers of genes.	Expensive; needs gene sequence data; technically demanding.
Diversity array technology (DArT to glass slides)	No sequence data required; high throughput; detects single base changes and indels; rapid germplasm characterization.	Dominant markers; technically demanding.

*The multiplex ratio is the number of independent loci detected in the assay.

The AFLP method combines the use of restriction enzymes with PCR amplification of fragments, and detects fragment length polymorphisms (Meudt and Clarke, 2007). AFLP molecular markers overcome problems related to low reproducibility and limited polymorphisms/mutations per analysis (Meudt and Clarke, 2007). In the development of PTC methods for plant propagation the use of molecular markers is mostly intended to evaluate the impact of tissue culture (genetic or epigenetic changes), the mechanism of variation and to establish the best true-to-type tissue culture conditions. As presented in Table I.8 AFLP has proved its efficacy in the evaluation of clonal fidelity in a range of tissue culture derived plants like *Freesia hybrida*, Rye (*Secale cereale* L.) and blue agave (*Agave tequilana* w.) among others finding variable levels of polymorphisms (De la Puente *et al.*2008; Gao *et al*. 2010, Diaz *et al.*2010).

The diversity of terms employed to report the level of genetic polymorphism within a population make difficult to compare the rates of variation. To avoid this problem and for comparative analysis data were recalculated in total polymorphisms in **Table I.8** using the formula: ***Total polymorphism*** = *[No. of polymorphic fragments/ (No. of fragments x No. of plant regenerants)] x 100* (Newbury et al. 2000).

Table I.8 Tissue culture induced genetic changes detected with AFLP techniques reported in the literrature. The variation rate is expressed in total polymorphisms, herein calculated for comparative analysis.

Species	Tissue culture system and PGRs conditions	Cause of genetic variation	Variation rate (Total polymorphisms %) *	Phenotype affectation	Reference
Arabidopsis	Indirect organogenesis, 2 mg/L 2, 4 D and 0.5 mg/L Kinetin callus induction	Tissue culture conditions	2.78	Not reported	Polanco and Ruiz. (2002)
Coffea arabica	Direct SE, 1.1 mg/L 6-BA and Indirect SE 1mg/L 2,4-D callus induction	Tissue culture conditions	2.58	Plants in nursery showed normal phenotype	Sanchez-Teyer et al. (2003)
Pea	Direct organogenesis (Shoot multiplication) 4.5 mg/L 6-BA and 0.01 mg/L NAA; long term cultures 24 years	Tissue culture conditions	0.15 in control plants vs 0.12 in regenerants, no statistical different	Not reported	Smýkal et al. (2007)
Grapevine	Indirect SE (Secondary embryogenesis) callus induction 0.5 mg/L 2,4-D and 0.1 mg/L 6-BA	Tissue culture conditions	0.79	Not reported, however author consider phenotype and development studies in older regenerants	Schellenbaum et al. (2008)

Echinacea purpurea	Direct organogenesis; shoot induction at 0.1 mg/L NAA and 1mg/L 6-BA	Tissue culture conditions, possibly diversity in plant material	9.40	Not reported	Chuang et al. (2009)
Freesia hybrida	Direct SE (2 mg/L IAA and 1-4 mg/L 6-BA)	Tissue culture conditions	0.07	Not reported	Gao et al. (2010)
	Indirect SE (2.5-10 mg/L 6-BA and 0.5-1 mg/L 2, 4-D)	-	0.02	Not reported	-

*Total polymorphism = [No. of polymorphic fragments/(No. of fragments x No. of plant regenerants)] x 100

Molecular markers are commonly used to evaluate the epigenetic status of tissue culture derived plants. For this purpose, two techniques are mainly used: Methylation Sensitive Amplified Polymorphisms (MSAP) developed by Reyna-López *et al*. (1997) and Methylation AFLP (Met-AFLP) developed by Bednarek *et al*. (2007). MSAP is an AFLP adaptation using a pair of methylation-sensitive restriction isoschizomers *Hpa*II and *Msp*I, that recognize the same tetranucleotide symetrical sequence 5'-C↓CGG-3' but have differential sensitivity to methylation at the inner or outer cytosine (Reyna-López et al. 1997). Another less implemented technique known as Met-AFLP uses a different pair of isoschizomers *Acc*651 and *Kpn*I which recognize another symmetric sequence 5'-GGTAC↓C-3'; the activity of the first one is blocked by *dcm* (methylase encoded by the dcm gene methylates the C^5-position of the internal cytosine residue) and CpG methylation (first cytosine) but is insensitive to *dam* (S-adenosylmethionine to the N^6 position of the adenine residues in the sequence), while the other one is insensitive to all form of methylation (Bednarek *et al*. 2007; Fiuk *et al*. 2010). The methylation sensitive molecular markers reavealed unexpectedly high frequencies of epigenetic variation among *in vitro* culture derived plants (Miguel and Marum 2010). **Table I.9** resumes different studies with methylation sensitive PCR techniques aimed at evaluating the impact of tissue culture on methylation changes.

Table I.9 Tissue culture induced methylation changes in propagated plants evaluated with methylation sensitive PCR techniques. The variation rate is expressed in total polymorphisms, herein calculated for comparative analysis.

Species	Tissue culture system and PGRs condiitons	Cause of methylation variation	Detection method	Total polymorphisms (%)	Phenotype affectation	Reference
Banana	Indirect organogenesis; callus induction with 5 mg/L 6-BA, 1 mg/L IBA	Tissue culture conditions	MSAP	1.72	Yes, leaf malformations were considered as epigenetic events by the authors	Peraza-Echeverria et al. (2001)
Barley	Indirect SE, hormonal conditions not specified	Tissue culture conditions	Met-AFLP	1.19	Not reported, plants developed and produced normally	Bednarek et al. (2007)
Codoponis lanceolata	Indirect organogenesis; Callus induction with 2.0 mg/L 6-BA 0.1 mg/L NAA	Tissue culture conditions	MSAP	6.49	Not reported	Guo et al. (2007)
Wild barley	Indirect organogenesis, callus induction 1- 6 mg/L	Tissue culture	MSAP	4.69	Not reported	Li et al. (2007)

	of 2, 4-D; callus maintenance 3 mg/L of 2, 4-D, regeneration 1 mg/L 6-BA.	conditions				
Pea	Direct organogenesis (shoot multiplication) 4.5 mg/L 6-BA and 0.02 mg/L NAA; long term cultures for 24 years	Tissue culture conditions	MSAP and HPCE	0.20 control plants vs 0.47% regenerants; statistical different**	No reported	Smýkal et al. (2007)
Hop plants	Direct organogenesis (Shoot multiplication) Conditions not specified ; long term culture from 3 to 12 months	Tissue culture conditions,	MSAP	0.10	No reported. Also they noted that polymorphisms accumulated during culture time	Peredo et al. (2008)
Grapevine	Indirect SE (Secondary embryogenesis) callus induction 0.05 mg/L 2,4-D and 0.01 mg/L 6-BA	Tissue culture conditions	MSAP	1.67	Not reported, however author planned to verifiy phenotype in older plants	Schellenbaum et al. (2008)
Gentiana pannonica	Shoot multiplication 2 mg/L 6-BA, 0.2 mg/L NAA then Indirect SE was induced with 1.0 mg/L NAA and 0.5 mg/L of kinetin	Tissue culture conditions	Met-AFLP	3.28	Not reported, plants developed and produced normally	Fiuk et al. (2010)
Freesia hybrida	Direct SE (2 mg/L IAA and 1-4 mg/L 6-BA)	Tissue culture conditions	MSAP	0.07	Not reported	Gao et al. (2010)
	Indirect SE (2.5-10 mg/L 6-BA and 0.5-1 mg/L 2, 4-D)	-	-	0.08	Not reported	-

*(SDE) secondary embryogenesis (ESP) embryogenic suspensions and define the type of multiplication step for augmenting plant yield in these industrial SE protocols. ** Based in their confidence intervals from the binomial distribution. Total polymorphism = [No. of polymorphic fragments/(No. of fragments x No. of plant regenerants)] x 100.

The activation of TEs by tissue culture conditions and their implication as a possible cause of SV was very early proposed by Larkin and Scowcroft (1981). The insertional activity of both RNA retrotransposons and DNA II transposons has been succesfully verified with different techniques including DNA and RNA hibridization techniques, PCR related techniques (e.g. transposon display and SSAP) and by phenotype evaluation. Different studies intended to uncover the factors and mechanisms implicated in the transposition of RNA class I retrotransposons and DNA class II transposons are sumarized in **Table I.10**. Interestingly, we can observed that the plant phenotype was only affected by DNA transposons (Class I); these elements were involved in the complete or partial color change of different plant structures (e.g. spotted coloration or variegation) and besides color alteration no other abnormalities were associated to TEs mobilization (e.g. structure or developmental abnormalities). Some of the given examples include spotted maize kernels (Planckaert and Walbot, 1989; Peschke and Phillips, 1990; Rhee et al.2010) variegation in carrot leaves (Ozeki et al.1997) and petal flowers (e.g. the ornamental African flower *Saintpaulia*) (Sato *et al.*2011) and the complete pigmentation of potato tubers found in red skinned potato (Momose et al.2010).

Table I.10 Tissue culture induced transposition in plant regenerants.

Type of TE and name	Species	Tissue culture system and PGRs conditions	Activity detection method	Cause of activation	Phenotype affectation	Reference
RNA Class I retrotransposon						
Tnt1, Tto1, and *Tto2* (Copia)	Tobacco	Indirect organogenesis, callus induction, 2mg/L NAA, 0.1 mg/L 6-BA. Shoot induction, 0.1 mg/L NAA, 2 mg/L 6-BA	Northern and Southern blott	Tissue culture conditions; higher TE activity in callus cultures vs plants regenerants	Not reported	Hirochika (1993)
Tos1- Tos 20	Rice	Indirect organogenesis,2 mg/L 2, 4 D	Southern blot	Tissue culture conditions; *Tos 17* and *Tos 19* active in short and long term cell cultures; *Tos 19* active in regenerants from 24 months cell cultures	Not reported	Hirochika (1996)
Tto1 (Copia)	Tobacco	Indirect organogenesis, callus induction, 2 mg/L NAA, 0.25 mg/L 6-BA	Northern blot	Tissue culture, wounding, methyl jasmonate, and some fungal elicitors	Not reported	Takeda et al. (1999)
Tnt1 (Copia)	Tobacco	Indirect organogenesis; conditions not specified	SSAP	Activated by fungal factors; Higher transposition in leaf derived protoplast vs leaf explants	Not reported	Melayah et al. (2001)
Karma (LINE – type)	Rice	Indirect organogenesis, conditions not specified	Southern blot	Inactive in cultured cells (5 months to 5 years) or R0 plants ; Activated in R1 plants (segregation)	Not reported	Komatsu et al. (2003)
Ty1- copia	Sweet potato	Indirect organogenesis (meristem cultures) callus induction, 0.2 mg/L NAA, 2 mg/L 6-BA	SSAP	Tissue culture conditions	Yes, abnormal leaves and roots possibly caused by TE activity	Tahara et al. (2004)
Lib (LINE-type)	Sweet potato	Indirect organogenesis (meristem cultures) callus induction (0.2 mg/L NAA, 2 mg/L 6-BA)	SSAP, Dot blot, Southern blot	Tissue culture conditions	Not reported	Yamashita and Tahara (2006)
Lullaby	Rice	Indirect organogenesis (transformed plants) conditions not specified	Microarrys, Southern blot and RBIP	Tissue culture conditions	Not reported	Picault et al. (2009)
MERE1- 1 (Copia family)	*Medicago truncatula*	Indirect organogenesis, conditions not specified	Southern blot, Transposon display	Tissue culture conditions; transcription related to methylation status*	Not reported	Rakocevic et al. (2009)
LORE-1 (*Gypsy*	*Lotus*	Indirect organogenesis (1-3	Southern blot,	Active in both transformed plants and	Not reported	Fukai et al.

superfamily)		mg/L 2, 4 D)	SSAP	plants regenerants; chromatin affinity; positive correlation with methylation variation		(2010)
DNA Class II transposon						
Mu (*Mutator*)	Maize	Indirect organogenesis, callus induction, 1 mg/L 2,4-D. Long term cultures from 4 to 14 months	Southern blot, phenotype evaluation	Methylation; Inactive lines remained inactive, active cell lines showed TE stability while other lines lost activity in long term cultures*	Yes, bronze spotted maize kernels	Planckaert and Walbot (1989)
En/Spm (*Supressor mutator*)	Maize	Indirect organogenesis, callus induction (1 mg/L 2,4-D), long term cultures from 4 to 22 months	Phenotype evaluation of kernels	Tissue culture conditions, active in two R1 from a single R0 from 8 months old callus culture	Yes, spotted kernels	Peschke and Phillips (1990)
Tdc1 (*En/Spm-like*)	Carrot	Indirect organogenesis 1 mg/L 2,4-D, long term cultures from 4 to 12 years	Southern blot, phenotype evaluation	Tissue culture conditions, active in long term cell cultures (12 years)	Yes, Leaf pigmentation in plant regenerants	Ozeki et al. (1997)
mPing (Tourist-like MITE)	Rice	Indirect organogenesis (anther culture)	Transposon display	Tissue culture conditions	Not reported	Kikuchi et al. (2003)
Pong (Tourist-like MITE)	Rice	Indirect organogenesis, 2 mg/L 2, 4 D	Southern blot	Tissue culture conditions	Not reported	Jiang et al. (2003)
RYS1	Rye	Indirect organogenesis ; conditions not specified	Southern blot	Tissue culture conditions	Not reported	Alves et al. (2005)
ZmTPAPong like-1 and 3 (MITE)	Maize	Indirect organogenesis, callus induction, 2 mg/L 2,4-D, 1 mg/L 6-BA	Southern blot	Tissue culture conditions	Not reported	Barret et al. (2006)
mPing (MITE)	Rice	Indirect organogenesis; callus induction, 2mg/L 2,4-D; 1 mg/L NAA	Transposon display	Tisue culture conditions, positive correlation with methylation alteration	Not reported	Ngezahayo et al. (2009)
nDaiZ (*h*AT superfamily)	Rice	Indirect organogenesis (anther cultures) conditions not specified	Southern blot	Tissue culture conditions, hypomethylation	Not reported	Huang et al. (2009)
dTstu-1 and 2, dTstu1-like (*MITE Stowaway family*)	Potato	Indirect SE (meristem cultures) conditions not specified	Phenotype evaluation, Transposon display	Tissue culture conditions; reversion of flavonoid 3',5'-hydroxylase function by removal of *dTstu1* TE from the gene	Yes, purple tuber pigmentation	Momose et al. (2010)
MyB (*Mutator*)	Maize	Indirect organogenesis ,	Methylation sensitive	Hypermethylation induced lack of activity	Yes, Kernels	Rhee et al.

		conditions not specified	southern blot; phenotypic evaluation	of *pericarp color* (*p1*) gene of maize*	pigmentation	(2010)
VGs1 (*h*AT superfamily)	Saintpaulia	Direct organogenesis, 1 mg/L NAA, 2 mg/L 6-BA	Phenotype evaluation; quantification of mutated cells by RT-PCR	Tissue culture conditions	Yes, flower petals variegation	Sato et al. (2011 a and b)
TCUP (hAT family*)*	Maize	Indirect organogenesis, callus induction , 1.5 mg/L 2,4-D. Cell suspensions treated with 5-aza-2-deoxycytidine	Southern blot; Transposon display	Tissue culture conditions + 5-aza-2-deoxycytidine treatment	Not reported	Smith et al. (2012)

* Preexisting activity of certain elements.

The mechanism of color variegation in maize kernels controled by TEs reactivation is shown in **Figure I.16** No major alterations like structural or developmental abnormalities were noted in these studies. Other studies made in *Arabidopsis* vitro plants by Jiang *et al.* (2011) found little contribution of TEs related to the generation of SV. This observation is also supported by Yu *et al*. (2011) studies in maize where the extent of TEs polymorphisms in *in vitro* generated plants was less prominent, in spite of their abundance in the genome.

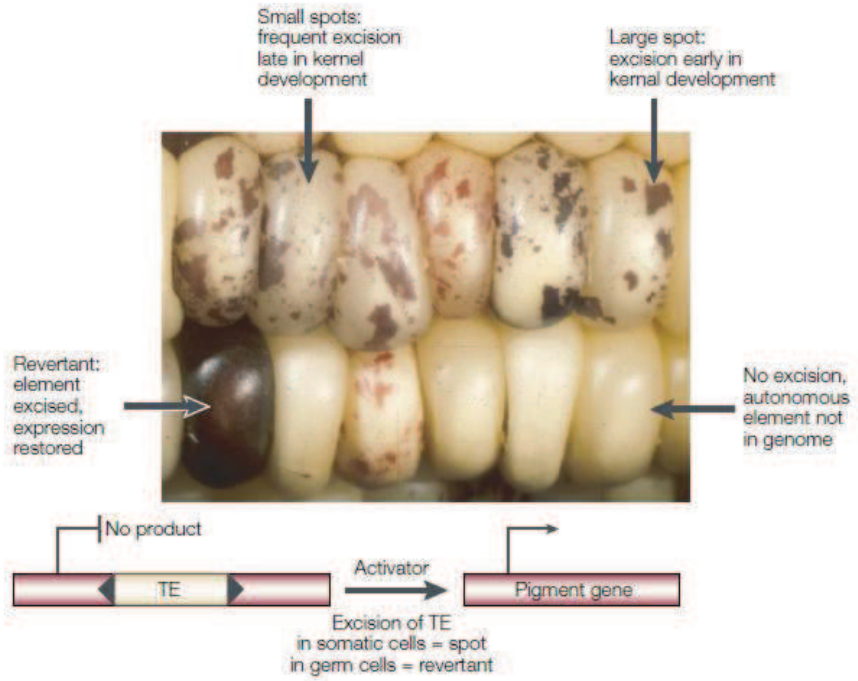


Figure I.16 Kernels on a maize ear showing unstable phenotypes due to the interplay between a TE (transposon) and a gene that encodes an enzyme in the anthocyanin (pigment) biosynthetic pathway. Sectors of revertant (pigmented) aleurone tissue result from the excision of the TE in a single cell. The size of the sector reflects the time in kernel development at which excision occurred (Feschote *et al.* 2002).

Cytological techniques

The techniques applied for their detection of both chromosomal structural changes and ploidy variations are technical demanding and require specialized training (Orton 1983). **Table I.11** presents the characteristics and uses of different cytogenetic techniques intended for the analysis of different kinds of abnormalities generated by tissue culture.

Table I.11 Summary of the strengths and limitations of different cytogenetic techniques for the detection of genetic variation in cultured somatic tissues. Adapted from Orton (1983).

	Type of variation							
	Quantitative			Qualitative				
Technique	Repetitive/ unique fractions	Inverted repeats	Aneuploidy /Polyploidy	Deletions/ Inversions/ Tranaslocations	Somatic recombination	Bases changes	Insertions	Technical aspects
Flow cytometry	-	-	+,-	+,-	-	-	-	High variability
Chromosome counts	-	-	+	+,-	-	-	-	Samples only mitotic cells
Kariotype analysis	-	-	+	+,-	-	-	-	Samples only mitotic cells
Southern blot	+,-	+,-	+	+,-	+,-	+	+,-	Technical demanding
In situ Hybridisation	-	+,-	+	+	-	-	+,-	Technical demanding
Cytogenetic analysis on progenie	-	-	+,-	+,-	+,-	+,-	+,-	Some variation may not be transmited

(+) Technique yields direct conclusive information (+,-) Technique yields indirect conclusive information or special assumption or conditions are necessary to yield conclusive information; -) Technique yields little or non pertinen information.

As seen in **Table I.12** cytological abnormalities are frequently produced under high concentrations of growth regulators and/or prolonged culture times. More severe abnormalities have been noted in PTC derived plants presenting major cytological abnormalities (e.g. aneuploidy) if compared with those showing methylation alterations or active TEs. For example, the presence of dwarf or undersized plants is an altered trait frequently observed in somaclonal variants with major chromosomal changes like aneuploidy (Nehra *et al*. 1992; Etienne and Bertrand, 2003; Bairu *et al*. 2006).

Table I.12 Tissue cultre induced chromosomal variations in plant regenerants

Species and ploidy level	Tissue culture system and PGRs	Type of chromosomal abnormality	Detection method	Phenotype affectation and cause of variation	Reference
Potato (2n=4x=48)	Indirect organogenesis; PGRs conditions not specified	Polyploidy (4x and 8x plants) mixoploidy and aneuploidy	Chloroplast counts in stomatal guard cells	Yes, wide array of abnormalities in leaf; yellow sectors and size variation, mostly related to mixoploid plants	Sree-Ramulu et al. (1983)
Maize (2n=2x=20)	Indirect organogenesis; callus induction and maintainance, 0.75 mg/L of 2, 4-D. Callus culture from 3 to 4 months and 8 to 9 months	Chromosomal abnormalities in 8 to 9 months callus derived plants; 3 to 4 showed no abnormalities.	Cytological analysis	Yes, variants phenotypes were not described	Lee and Phillips (1987)
Barley (2n=2x=14)	Indirect organogenesis; callus induction, 2.0 mg/L 2, 4-D. Callus maintenance from 2 to 12 months ,0.1, 1, 2 and 3 mg/L of 2, 4 –D.	Cell culture showed a rise in abnormal polyploid mitosis in long term cultures under high levels of 2, 4-D levels	Centromere banding stain	Yes, albino plants derived from high levels of 2, 4- D and long term cultures	Ziauddin and Kasha. (1990)
Barley (2n=2x=14)	Indirect organogenesis/Anter culture 1 mg/L 6-BA and 2, 4-D. suspension cultures, 2 mg/L 2, 4-D	Loss of regeneration capability in long term cultures indicates aneuploidy	Not verified	Yes, albino plants in 1 year old callus derived plants	Jänhe et al. (1990)
Strawberry cv Redcoat (2n=8x=56)	Indirect organogenesis; callus induction, 0.2,1 , 2 and 4.5 mg/L of 6-BA and 2, 4-D.	Aneuploidy related to leaf malformation in plants, chimerism described in yellow leaves plants.	Flow cytometry	Yes, high auxin levels(20 µM of 6-BA and 2, 4-D) produced Dwarf plants; culture ageing produce malformed or yellow leaves plants (4 to 6 months at 5µM of 6-BA and 2, 4-D)	Nehra et al. (1992)
Orange (2n=2x=18)	Indirect SE, PGRs conditions not specified. Long term cultures.	Low polyploidy (4x) and aneuploidy levels in somatic embryos and plant regenerants. Abnormal mitosis in callus cultures (lagging chromosomes, chromosomes bridges).	Chromosome counts and chromosome examination	Not reported	Hao and Deng (2002)
Coffea arabica (2n=4x=44)	Indirect SE , callus induction 0.5mg/ml 2, 4 D; suspension cultures, 1 mg/L 2, 4-D. Long term cultures from 2 to 12 months	No abnormality detected	Flow cytometry	Yes, high frequency of variants phenotypes in long term cultures (9-12 months). Variants: 1) Angustifolia. 2) Dwarf. 3) Mutistem. 4) Bullata	Etienne and Bertrand (2003)
Rice (2n=2x=24)	Indirect organogenesis, 2.5 mg/L 2, 4-D	Haploidy, polyploidy (3x, 4x) Monosomy (2n-1), trisomy (2n+1), chromosomal rearrengment s	Karyotype analysis	Not verified	Kharabian and Darabi (2005)
Papaya (2n=2x=18)	Indirect SE, embryo induction, 2 mg/L of 2, 4-D	Polyploidy (3x and 4x plants) mixoploidy and aneuploidy	Flow cytometry	Not verified	Clarindo et al. (2008)

Arabidopsis (2n=2x=10) ecotype columbia	Indirect organogenesis; callus induction, 0.5 mg/L of 2, 4-D and 0.05 mg/L of kinetin; 2 and 6 weeks old calli used for regeneration.	Flow cytometry: 2x, 4x and 8x plants in 2 weeks old calli plants; 2x, 4x in 6 weeks old calli plants. Chromosome count: mixoploidy and high frequency of aneuploid cells detected in 2-week-old calli. FISH: chromosomal rearrangements	Flow cytometry, chromosome count and FISH (Fluorescence in situ Hybridization)	Yes, correlation with ploidy level, 4x and 8x plants were shorter than 2x and had sturdy stems and leaves. Aneuploid cells detected in 2-week-old calli were correlated with plant malformation and fertility loss.	Orzechowska et al. (2012)

State of the art of SV in coffee plants propagated *in vitro*

Frequency of somaclonal variation in coffee

SV in coffee has been described in both of the most important species of the genera, *C. canephora* and *C. arabica*. In the case of *C. canephora* first studies reported the absence of somaclonal variants. Subsequent studies performed by Ducos (2003) described the presence of less vigorous plants in different field trials reporting 2.3% of overall variation from a total population of 1900 plants in Philippines trials and 3.8% in Thailand trials from a total population of 3156 plants (Ducos *et al.* 2003). Only one type of phenotypic variant was described characterized by less vigorous plants with yellowish leaves and low agronomic performance. Although authors consider less vigorous plants as the result of planting problems, the same study does not discard this kind of phenotype as a resut of *in vitro* passage. This altered trait has been also reported in maize somaclonal variants and could be related to physiological disorders (Peschke and Phillips 1992).

Different studies reported variable levels of SV in *C.arabica* plants produced through SE. Söndhal and Lauritis (1992) estimated an overall rate of 10% in a total population of 12,000 SE derived plants. Etienne and Bertrand (2001) estimated an 2.1% of overall variation in a total population of 644 somatic seedlings derived from 6 month-suspension cultures. In coffee the use of auxin (2, 4-D) is needed for the production of embryogenic cells (Santana *et al.* 2007) and the maintenance of proliferating embryogenic cells in suspension (Etienne and Bertrand, 2003). A detailed evaluation intended to establish the better culture conditions with embryogenic suspensions reported 1.3% of variation for all genotypes analyzed. This late in both plants derived from embryogenic callus and 3 months old- cell suspensions (Etienne and Bertrand 2003). Interestingly, the variation rate drastically increased when embryogenic suspensions proliferate beyond the sixth month reaching 10% of variation with 494 plants analyzed. After 12 months, this percentage increases to 25% (826 plants analyzed) sometimes producing the more abnormal phenotypes like multistem variants that were not able to survive in field conditions. Plant size, leaf morphology and plant arhitecture were the most reported altered traits in coffee (Söndhal and Lauritis, 1992; Etienne and Bertrand, 2001, 2003). Many of the variants found in these studies has significant similarities with mutants previously described in

seed progenies by Krug and Carvalho (1951). Some of the most common *C. arabica* mutants presenting aberrant phenotypes will be presented next.

Coffee mutants in seed progenies

Relatively elevated frequencies of mutants (1-8%) were observed in seed-derived coffee plant progenies. A greater number of mutants are known in the tetraploid *C. arabica* than in the diploid *C. canephora* (Krug and Carvalho, 1951). Some of the coffee mutants described at the time were recognized as resulting from chromosomal reduction or duplication (further data suporting this affirmation was not given in this review). The following descriptions correspond to some of the most common abnormal types of mutants (with the exception of *Anormalis*) affecting the normal morphoagronomic characteristics of *C.arabica*. One of the most common mutants seen in field correspond to plants presenting narrow leaves commonly known as Angustifolia (Krug *et al.* 1939). The shape of Angustifolia leaves is their main difference. They are oblanceolate, elongated, narrow at the base and thicker than the leaves of typica plants. This mutants seems to be originated by two recessive independent genes, ag_1, ag_1 and ag_2 ag_2 (Krug et al., 1939; Krug and Carvalho, 1951). Variegata plants normally presents abnormal plastids in leaf producing white, yellow and grenish yellowing sectoring resulting in different patterns of color known as variegation. Anormalis mutant is a very rare mutation severely affecting the whole plant architecture including fruits and seeds (Krug and Carvalho, 1951). Anormalis was first described in the coffee nursery of the central experimental central of Campinas Brazil (Krug and Carvalho, 1951). This mutation affects plant size and branching habits, their leaves are abnormal, subdivided almost palmate and at various sizes, the fruits are rather large and have a disc of variable size. These trees produced less abundant primary branches than typica, their leaves present ovaleliptic shape, their color is dark green, with a particular bronze coloration when fully matured. Bullata leaves are also more cariaceous than typica leaves with thicker leaf blades and midribs (Krug et al. 1939). The presence of cytological abnormalities has been described in sexual propagated plants. First reports made by Bragantia (1947) described *C. arabica* Angustifolia mutants as an aneuploid plant presenting 2n-1=43 chromosomes monosomy. Another plant presumably cytological abnormalities corresponded to *C. arabica* variety Bullata; this cultivar is usually referred as an hexaploid (2n=6x=66) or octoploid (2n=8x=88) plant (Krug, 1937).

Somaclonal variants obtained by tissue culture

Coffee somaclonal variants are possible counterparts of previous mutants described in literature. Coffee somaclonal variants present leaf abnormalities (e.g. Angustifolia and Bullata) and variegation (Variegata) abnormal plant architecture (Multistem and Anormalis) or altered size (Dwarf and Giant). Most somaclonal variants present flower and fruit abnormalities and consequently have a lower productivity (Etienne and Bertrand, 2003). According to these authors, the majority of SV apart from Dwarf variant, can be detected and eliminated in the nursery. In coffee as in many other cultured plant

it was found a relation between the genotype and the susceptibility to generate SV (Etienne and Bertrand, 2003). The effect of genotype and culture time can be seen at figure **figure I.17**, where the rate of SV rises only in some F1 hybrids clones after six months of culture with a clear genotypic effect (Etienne and Bertrand 2003).

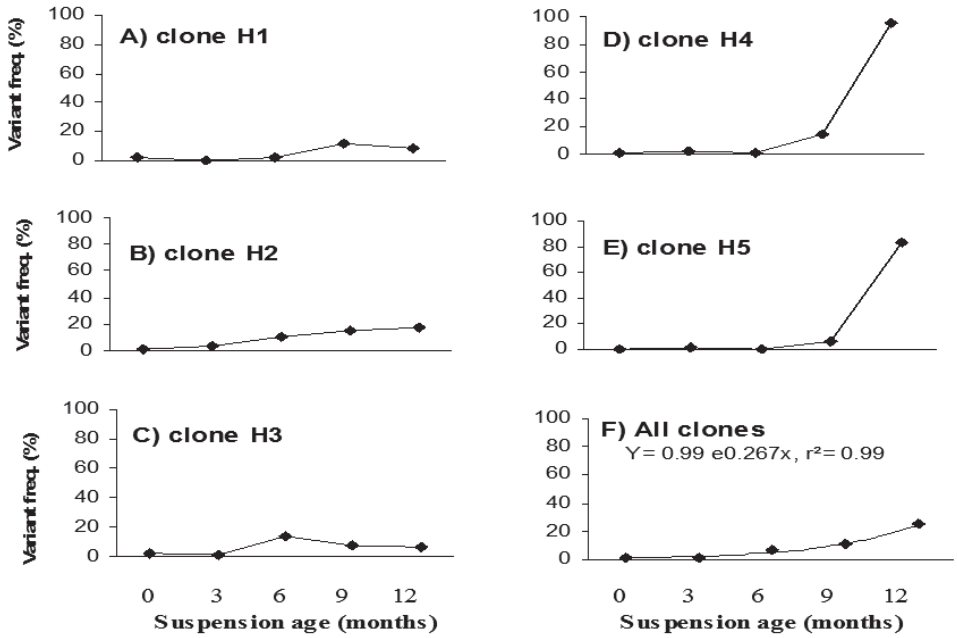


Figure I.17 Frequency of SV depending on proliferation duration in cell suspension cultures for five *C. arabica* F1 hybrids (Etienne et al. 2003).

As in other species e.g. oil palm (Giorgetti *et al.* 2011) there is evidence of pre-existing cytological abnormalities in coffee. Non disjunction, double prophases and lagging chromosomes were described in coffee leaf explants (pre-existing) and callus cultures (*de novo*) , those abnormalities included the presence of polyploids and aneuploids cells (Menéndez-Yuffá *et al.,* 2000). Additionally, interphase cells with micronuclei or binucleated were also observed. A few other abnormalities were described in the same study in coffee calllus during the induction of SE including c-mitosis, chained chromosomes, multipolar metaphases and chromosome bridges. Following the possible involvement of major chromosomal events causing abnormalities in coffee SE derived plants Etienne and Bertrand (2003) performed a FCM analysis in different types of somaclonal variants (e.g. Dwarf, Angustifolia, Mutistem and Bullata). Normal ploidy levels were found in this variants. However, a more detailed study of kariotype was proposed by the authors to verify the presence or absence of chromosomal abnormalities.

Molecular markers were used to evaluate the genetic and epigenetic changes in plant regenerants of coffee after both direct and indirect somatic embryogenesis. However, these studies were done using very reduced plant populations and also produced with non reliable protocols. Their results (i.e.

herein calculated in total polymorphisms for comparative analysis) are presented next. Rani et al. (2000) evaluate the impact of tissue culture in 27 *C. arabica* plants derived from indirect SE using different types of molecular markers (RFLP, RAPDs, SSRs). Their results indicated higher levels of changes in mitocondrial DNA (2.82%) than in nuclear DNA (0.07% RAPDs and 0.2% SSRs). In other study, Sanchez-Teyer *et al.* (2003) evaluate the genetic conformity of *C. arabica* regenerants from direct and indirect SE using AFLP. In this study, two samples of bulked DNA produced a total of 1,446 AFLP bands with 11.4% being polymorphic. However, the data from bulked DNA is not easily comparable with those obtained from other SE studies. In the same study they measured the intraclonal variation of 24 plants derived from direct and indirect SE (12 plants each) founding 2.5% of polymorphic bands. Nevertheless, the study did not detailed the percentage corresponding to each system of SE.

THESIS GENERAL OBJECTIVES

- Evaluate the somaclonal variation rate at the industrial level and the frequency of variant phenotypes generated by two somatic embryogenesis systems using secondary embryogenesis or cell suspension proliferation as proliferation methods.

- Identify the changes occurring at genetic and epigenetic level (methylation) during SE.

- Identify the mechanisms underlying somaclonal variation in coffee and if possible relate a mechanism to an abnormal phenotype.

THESIS SPECIFIC OBJECTIVES

In chapter II

- Large scale phenotypic evaluation (800.000 plants from two *C. arabica* F1 hybrids) at nursery and field level: frequency and phenotypic characteristics of somaclonal variants derived from secondary embryogenesis and cell suspensions.

- Verify the induction of genetic and epigenetic changes in batches of plants derived from both processes and two F1 hybrids using AFLP and MSAP molecular markers.

- Verify the induction of chromosomal aberrations (abnormal chromosome numbers) in a some plants obtained from both processes showing normal and variant phenotypes.

In chapter III

- Evaluate the effect of culture age on somaclonal variation using somatic seedlings (180 plants) derived from three different established cell lines (A, B and C) used as biological replicates and from variable culture periods (4, 11 and 27 months of cell proliferation) on auxin enriched media (1 mg/L 2, 4-D).

- Verify the induction of genetic and epigenetic changes in plants derived from the different conditions of this experimental design using AFLP and MSAP molecular markers.

- Determinate the random or not random character of somaclonal variation by comparing the data obtained from the plant batches regenerated from the different cell lines depending the culture age.

- Verify the activation of TEs under the *in vitro* propagation conditions by using SSAP markers.

- Verify the presence/absence of chromosomal aberrations by performing chromosome counting on plants derived from cell cultures of variable culture ages.

CHAPTER II

High Genetic and Epigenetic Stability in *Coffea arabica* Plants Derived from Embryogenic Suspensions and Secondary Embryogenesis as Revealed by AFLP, MSAP and the Phenotypic Variation Rate

This chapter was published in PLoS ONE 8(2): e56372. doi:10.1371/journal.pone.0056372

High genetic and epigenetic stability in *Coffea arabica* plants derived from embryogenic suspensions and secondary embryogenesis as revealed by AFLP, MSAP and the phenotypic variation rate

Roberto Bobadilla Landey[1], Alberto Cenci[2], Frédéric Georget[1], Benoît Bertrand[1], Gloria Camayo[3], Eveline Déchamp[1], Juan Carlos Herrera[3], Sylvain Santoni[4], Philippe Lashermes[2], June Simpson[5], Hervé Etienne[1,*]

1 *Unité Mixte de Recherche Résistance des Plantes aux Bioagresseurs, Centre de Coopération Internationale en Recherche Agronomique pour le Développement, Montpellier, France*

2 *Unité Mixte de Recherche Résistance des Plantes aux Bioagresseurs, Institut de Recherche pour le Développement, Montpellier, France*

3 *Centro Nacional de Investigaciones de Café, Manizales, Colombia*

4 *Unité Mixte de Recherche Amélioration Génétique et Adaptation des Plantes tropicales et méditerranéennes, Institut National de la Recherche Agronomique, Montpellier, France*

5 *Department of Plant Genetic Engineering, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, Guanajuato, Mexico*

* Corresponding author: E-mail: herve.etienne@cirad.fr, Tel.+33-467416227, Fax:+33-467416181

Email addresses:

RBL: ravenlandey@hotmail.com

AC: a.cenci@cgiar.org

FG: frederic.georget@cirad.fr

BB: benoit.bertrand@cirad.fr

GC: gloria.camayo@cafedecolombia.com

ED: eveline.dechamp@ird.fr

JCH: juanC.herrera@cafedecolombia.com

SS: santoni@supagro.inra.fr

PL: philippe.lashermes@ird.fr

JS: jsimpson@ira.cinvestav.mx

HE: herve.etienne@cirad.fr

Abstract

Embryogenic suspensions that involve extensive cell division are risky in respect to genome and epigenome instability. Elevated frequencies of somaclonal variation in embryogenic suspension-derived plants were reported in many species, including coffee. This problem could be overcome by using culture conditions that allow moderate cell proliferation. In view of true-to-type large-scale propagation of *C. arabica* hybrids, suspension protocols based on low 2,4-D concentrations and short proliferation periods were developed. As mechanisms leading to somaclonal variation are often complex, the phenotypic, genetic and epigenetic changes were jointly assessed so as to accurately evaluate the conformity of suspension-derived plants. The effects of embryogenic suspensions and secondary embryogenesis, used as proliferation systems, on the genetic conformity of somatic embryogenesis-derived plants (emblings) were assessed in two hybrids. When applied over a 6 month period, both systems ensured very low somaclonal variation rates, as observed through massive phenotypic observations in field plots (0.74% from 200 000 plant). Molecular AFLP and MSAP analyses performed on 145 three year-old emblings showed that polymorphism between mother plants and emblings was extremely low, i.e. ranges of 0-0.003% and 0.07-0.18% respectively, with no significant difference between the proliferation systems for the two hybrids. No embling was found to cumulate more than three methylation polymorphisms. No relation was established between the variant phenotype (27 variants studied) and a particular MSAP pattern. Chromosome counting showed that 7 of the 11 variant emblings analyzed were characterized by the loss of 1-3 chromosomes. This work showed that both embryogenic suspensions and secondary embryogenesis are reliable for true-to-type propagation of elite material. Molecular analyses revealed that genetic and epigenetic alterations are particularly limited during coffee somatic embryogenesis. The main change in most of the rare phenotypic variants was aneuploidy, indicating that mitotic aberrations play a major role in somaclonal variation in coffee.

Introduction

Among micropropagation methods, somatic embryogenesis has the best potential for rapid and large-scale multiplication of selected varieties in a wide range of economically important species. Schematically, the initial step of dedifferentiation leading to the acquisition of embryogenic competence is common to most plant species, whereas for the following step of proliferation of embryogenic material, efficient procedures can be classified under two main strategies. The first is the proliferation through secondary embryogenesis (SCE) which involves first differentiating the somatic embryos before enhancing their proliferation by adventitious budding (Figure 1). The second consists of establishing embryogenic suspensions (ESP) to favor large-scale embryogenic cell proliferation before the subsequent embryo differentiation step. In order to come up with an industrial procedure, the development of ESP represents the best option to ensure synchronous and massive somatic embryo production [1]. In addition, ESP allows the production of large numbers of embryogenic-competent cells and this process can be easily scaled up. Nevertheless, tissue culture systems such as somatic embryogenesis that involve the acquisition of competence for pluripotentiality and extensive cell division are more risky with respect to genome and epigenome instability [2]. The use of ESP has frequently been associated with an increased risk of genetic instability and somaclonal variation in the regenerated plants [3-5]. Although ESP has been developed for some important crops, it has therefore not been widely applied for commercial purposes. Somaclonal variation in ESP-derived plants is probably related to the presence of 2,4-dichlorophenoxyacetic acid (2,4-D), which is often essential for maintaining proliferating cells in an embryogenic, undifferentiated state [6,7]. This auxin could enhance somaclonal variation through the stimulation of rapid disorganized growth that can influence the mitotic process, resulting in chromosomal aberrations [8,9].

The term ‘somaclonal variation’ (SV) describes the tissue culture-induced stable genetic, epigenetic or phenotypic variation in clonally propagated plant populations [10]. Somaclonal variation is considered to be one of the main bottlenecks in the development of micropropagation procedures, especially in view of large-scale commercial operations, for which the strict maintenance of genetic and agronomic traits from selected individuals is required. An analysis of the progeny of phenotypic variants showed that some of the variations produced by somatic embryogenesis can occur in the form of stable and heritable mutations [8,11]. In maize, Kaeppler and Phillips [12] also reported stable segregation of somaclonal variant phenotypic qualities in several seed generations. It has also been well documented that somaclonal variants commonly present cytological aberrations such as chromosomal rearrangements (deletions, duplications, inversions and translocations), and sometimes more severe alterations like aneuploidy or polyploidy [12-16].

Although most mutants segregate in a Mendelian fashion upon selfing and outcrossing [12], SV is sometimes present in the form of transient mutations, suggesting the involvement of epigenetic events [11]. Epigenetic traits are heritable changes associated with chemical modification of DNA without alteration of the primary DNA sequence [17]. Cytosine methylation has been proposed as a possible cause of SV [11,12]. Epigenetic modifications (methylation) can mediate the transmission of an active or silent gene in the short-term (mitotic cell division) or long-term (meiotic divisions leading to transmission across generations) [18]. DNA methylation in plants commonly occurs at cytosine (5-methylcytosine, m^5C) bases in all sequence contexts: the symmetric CG and CHG (in which H could be A, T or C) and the asymmetric CHH contexts [17,18]. Molecular marker approaches like methylation-sensitive amplified polymorphism (MSAP) and Met-AFLP have proved efficient in the analysis of methylation patterns [19,20]. The existence of zones susceptible to methylation variations was recently shown in somatic embryogenesis-derived plants (emblings) in grapevine [21] and barley [20]. SV was also associated with the activity of mobile DNA elements or retroelements [22,23]. Novel mechanisms such as RNAi directed demethylation have recently been proposed to explain retrotransposon activation [2,24].

Coffea arabica is an allotetraploid tree species (2n=4X=44) characterized by low molecular polymorphism [25]. Somatic embryogenesis is currently applied industrially for large-scale and rapid dissemination of selected F1 hybrids that provide a highly significant increase in the yield of high quality coffee [26,27]. Regarding industrial-scale micropropagation, upgrading production to several million vitroplants per production unit would undoubtedly boost economic profitability. This would require switching from an SCE- to an ESP-based protocol. However, former field observations revealed that SV occurs at relatively high rates in ESP-derived *C. arabica* plants [28,29]. Apart from different phenotypic variants easily identifiable through morphological characteristics, we did not discover in trees showing a normal phenotype any variations involving agronomically important quantitative and physiological characteristics [30]. In view of true-to-type propagation of selected *C. arabica* hybrid varieties, we previously developed improved ESP protocols based on the use of low exogenous 2,4-D concentrations and short proliferation periods, allowing reliable somatic embryo mass regeneration [27]. For potential commercial applications, here using two *C. arabica* hybrids we verified the conformity of suspension-derived plants with that of plants obtained by secondary embryogenesis, i.e. the industrial process currently in use. The objectives were to assess large-scale phenotype conformity in commercial field plots, to quantify genetic and epigenetic modifications in the regenerated plants through AFLP (Amplified fragment length polymorphism) and MSAP molecular markers, and to cytologically characterize the karyotype of different phenotypic variants detected in the study.

Materials and Methods

Plant material and somatic embryogenesis

Selected F1 hybrids of *C. arabica* [26] obtained by crossing traditional dwarf American varieties (Caturra, Sarchimor T5296) and wild accessions originating from Ethiopia and Sudan are disseminated in Central America through somatic embryogenesis. In the present study, emblings derived from the two hybrids Sarchimor T5296 x Rume Sudan and Caturra x ET531, named respectively H1 and H3, were analyzed to assess the SV level. Large-scale phenotypic observations were performed in Nicaragua both at the nursery (more than 600 000 young emblings) and field level (more than 200 000 three year-old emblings) on 11 coffee plots belonging to the 'La Cumplida' coffee research experimental sites in the Matagalpa region (Nicaragua). Nursery and field phenotypic observations were done on balanced amounts of plants from hybrids H1 and H3 (Table 1). Field observations were performed for all trees by visual evaluation of growth and morphology, flowering and fruit yield during the first and second production years. No specific permits were required for the described field studies that were performed with the authorization of the Coffee Research Department of 'La Cumplida', owner of the experimental sites. These sites are not protected and the studies did not involve endangered or protected species. Molecular marker analysis was applied on F1 hybrid mother plants propagated by rooted horticultural cuttings (four for each hybrid) and used as source material for *in vitro* propagation, as well as on adult emblings (3 years after planting) for which plants exhibiting an abnormal phenotype (phenotypic variant) were distinguished from plants with a normal phenotype and productivity (**Table 8**). Molecular marker patterns obtained for emblings were systematically compared with those obtained with mother plants (four plants per hybrid propagated by horticultural cutting) used as reference.

Table 8: Plant material used in molecular marker analyses. Total number of plants analyzed from two F1 *Coffea arabica* hybrid lines (H1 and H3) corresponding to somatic-embryo derived 3 year-old plants (emblings) with normal or variant phenotypes along with their respective mother plants as reference. The numbers of emblings for each variant phenotype are also given.

Somatic embryogenesis proliferation step	Type of plant material	No. of plants per *C. arabica* hybrid		Total no. of plants
		H1	H3	
Secondary embryogenesis	Emblings normal phenotype	28	31	59
	Emblings variant phenotype	5	14	19
	Angustifolia (A)	0	4	
	Bullata (B)	2	4	

	Dwarf (D)	0	2	
	Variegata (V)	3	4	
	Total	33	45	78
Embryogenic suspension	Emblings normal phenotype	25	34	59
	Emblings variant phenotype	1	7	8
	Angustifolia (A)	1	2	
	Dwarf (D)	0	1	
	Variegata (V)	0	4	
	Total	26	41	67

The somatic embryogenesis process (Figure 1) involved the following stages:

1) Production of embryogenic callus: pieces of young leaves were surface-sterilized and used as explants. The explants were cultured for 1 month on a 1/2 strength MS [74] "C" callogenesis medium [75] containing 2.3 µM 2,4-D, 4.9 µM indole-3-butyric acid (IBA) and 9.8 µM iso-pentenyladenine (iP), and then transferred for 6 months to MS/2 "ECP" embryogenic callus production medium [75] containing 4.5 µM 2,4-D and 17.7 µM benzylaminopurine (6-BA). All media were solidified using 2.4 g/l Phytagel (Sigma, Steinheim, Germany). These steps were carried out at 26-27°C in the dark.

2) 6-month multiplication step and embryo regeneration. Fully developed somatic embryos were mass regenerated via two distinct multiplication processes, i.e. either secondary (or repetitive) embryogenesis or embryogenic suspensions (Figure 1).

Secondary embryogenesis (SCE). Two hundred mg of embryogenic aggregates were placed in a 1 liter-RITA® temporary immersion bioreactor (CIRAD, Montpellier, France; [76]) along with 200 ml of "R" MS/2 regeneration medium [75] containing 17.76 µM 6-BA, in darkness for 6 weeks. An immersion frequency of 1 min every 12 h was applied. Proliferating embryo masses were then placed in 1/2 strength MS [74] regeneration medium containing 5.6 µM 6-BA and subcultured once every 6 weeks for three proliferation cycles. Secondary embryos were produced with an immersion frequency of 1 min every 12 h and a high culture density (approx. 10 000 embryos). The cultures were kept at 27°C, with a 12 h/12 h photoperiod and 50 µmol m^{-2} s^{-1} photosynthetic photon flux density.

Embryogenic cell suspensions (ESP). Embryogenic calli were transferred to 100 ml Erlenmeyer flasks at a density of 1 g/L in 1/4 MS strength Yasuda liquid proliferation medium [77] with 1.36 µM 2,4-D and 4.4 µM 6-BA. Suspension cultures were maintained by the monthly transfer of 1 g/L of embryogenic aggregates into fresh medium. Six month-old suspensions were used for somatic embryo regeneration. Embryo differentiation was initiated by transferring embryogenic aggregates at a density

of 1 g/L in 250 ml Erlenmeyer flasks in a full strength MS medium containing 1.35 µM 6-BA. Fully developed torpedo-shaped embryos were obtained after two 4 week subcultures in such conditions. All suspension cultures were shaken at 110 rpm at 27 °C under 50 µmol m^{-2} s^{-1} photosynthetic photon flux density.

3) Pre-germination in a bioreactor. Germination was triggered by applying a low culture density of around 800-900 embryos per 1 l-RITA® bioreactor. An "EG" embryo germination medium [75] containing 1.33 µM BA was used for 2 months and finally, for 2 weeks, the "EG" culture medium was supplemented with 234 mM sucrose. By the end of the in vitro culture stage, each bioreactor contained around 700 pre-germinated cotyledonary somatic embryos with an elongated embryonic axis and a pair of open, chlorophyllous cotyledons. Pre-germination was conducted in the light (12/12 h, 50 µmol m^{-2} s^{-1}).

4) Plantlet conversion was obtained after direct sowing of mature somatic embryos in the nursery. Mature embryos were sown vertically on top of the substrate (two parts soil, one part sand, one part coffee pulp) sterilized by chemical treatment (Dazomet (DMTT), Union Carbide). The somatic embryo culture density in the plastic boxes (l.w.h = 30/21/10 cm) was approximately 3600 m^{-2}. The cultures were placed under a transparent roof that provided 50% shade, and were watered for 2 min twice daily. Conversion of somatic embryos into plants was generally observed 12 weeks after sowing, and characterized by the emergence of a stem bearing at least two pairs of true leaves.

5) Growth and hardening in the nursery (21 weeks). Plantlets grown from somatic embryos were transferred to 0.3 L plugs on a substrate comprising peat-based growing medium (Pro-mix, Premier Tech Ltd, Canada) and coffee pulp (3/1, v/v) under conventional nursery conditions until they reached the required size for planting in the field (approx. 30 cm). During this stage, the shade (50% light interception) and relative humidity (80%) were gradually reduced over 4 weeks to 0% light interception, with natural RH ranging from 65 to 90%.

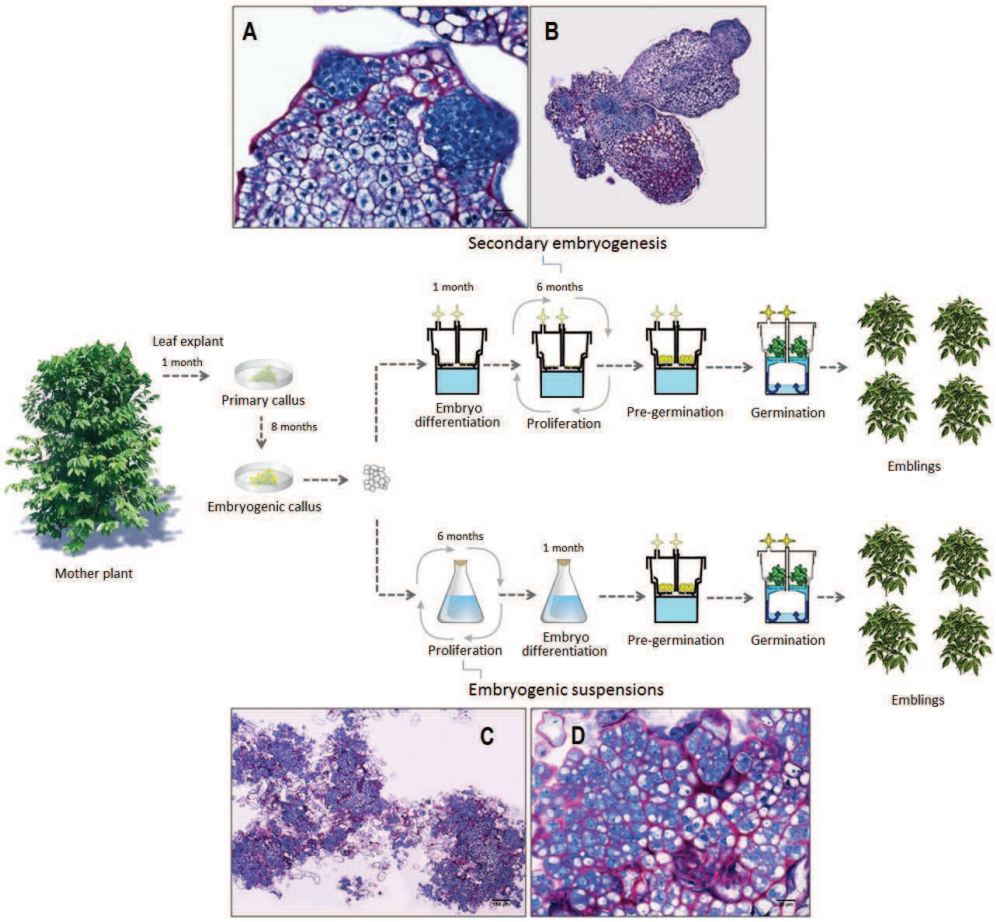


Figure 1. Schematic representation of two somatic embryogenesis processes applied at the industrial level. The first somatic embryogenesis process (upper section of the flow diagram) involved a proliferation step based on secondary embryogenesis in RITA® temporary immersion bioreactors (photos 1A, 1B). The second process (lower section of flow diagram) included a proliferation step based on embryogenic suspensions (photos 1C, 1D). 1A, initial developmental stages of secondary embryos at the root pole of primary somatic embryos; 1B, clusters of primary and secondary embryos; 1C, clusters of embryogenic cells in suspension; 1D, embryogenic cells in suspension.

Molecular analysis

DNA extraction

Young fully expanded leaves were selected on three year-old plants for molecular analysis. Genomic DNA was isolated from 100 mg of lyophilized leaves using Dellaporta buffer [78] containing sodium dodecyl sulfate (SDS) detergent and sodium bisulfite 1% w/v to avoid leaf oxidation. DNA was purified in spin-column plates as described in the DNeasy plant kit protocol from QIAGEN.

AFLP markers

AFLP analysis was carried out as described by Vos *et al*. [79], with a total of four primer combinations (Table 2), using 5-FAM or 5-HEX fluorescently labeled *Eco*R1 (+3) and unlabeled *Mse*I (+3) primers. A touchdown PCR program for selective amplification was performed in an Eppendorf thermocycler under the following conditions: 3 min at 94°C, 12 cycles of 45 s at 94°C, 12 cycles of 45 s at 65°C and 1 min at 72°C; the annealing temperature was decreased by 0.7°C per cycle from a starting point of 65°C during this stage, with a final round of 25 cycles of 94°C for 45 s, 56°C for 45 s, 72°C for 1 min and a final elongation step of 72°C for 1 min. The same PCR conditions were found to be appropriate for MSAP in preliminary tests.

Table 2: Primer combinations used for AFLP and MSAP analyses.

AFLP primer combinations (*Eco*+3 ^{labeled}/*Mse*+3)	MSAP primer combinations (*Eco*-3/*Hpa*+2 ^{labeled})
Eco-ACT/*Mse*-AGT	C1 *Eco*-AAC/*HPA*-AA
Eco-AGG/*Mse*-AGT	C2 *Eco*-AAC/*HPA*-AT
Eco-CGC/*Mse*-CCA	C3 *Eco*-AGG/*HPA*-AA
Eco-CAC/*Mse*-CCA	C4 *Eco*-AGG/*HPA*-AT
	C5 *Eco*-ACT/*HPA*-CA
	C6 *Eco*-ACT/*HPA*-CT
	C7 *Eco*-AGA/*HPA*-CA
	C8 *Eco*-AGA/*HPA*-CT

Fluorescent dyes for marked primers correspond to 5'-FAM ^{TM} and 5'-HEX ^{TM}

MSAP markers

MSAP analysis was carried out as described by Reyna-López *et al*. [19] with minor adaptations for capillary electrophoresis. The MSAP protocol is an adaptation of the AFLP method for the evaluation of different states of methylation in the symmetric sequence CCGG. In the MSAP protocol, the frequent cutting endonuclease (*Mse*I) was replaced by the two isoschizomeric restriction enzymes *Hpa*II and *Msp*I with different sensitivity to the methylation state of the symmetric sequence CCGG (Table 4). Specifically, *Hpa*II is able to recognize and cut only when the CCGG sequence is unmethylated or hemi-methylated on the external cytosine. *Msp*I is able to cut when CCGG sequence is unmethylated or if the internal cytosine is fully or hemi-methylated. Both *Hpa*II and *Msp*I are unable to cut if the external cytosine presents full methylation. DNA methylation in plants commonly occurs at cytosine bases in all sequence contexts: the symmetric CG and CHG, in which (H = A, T or C) and the asymmetric CHH contexts [17]. Selective amplification included a total of eight primer combinations per isoschizomer (Table 2). *Hpa*II (+2) primers were fluorescently labeled with 5-FAM or 5-HEX while *Eco*R1 (+3) remained unlabeled. In order to reduce the possibility of technical

artifacts, two repetitions using different DNA extractions were performed for each primer combination.

Capillary electrophoresis and data analysis

PCR products were separated by capillary electrophoresis with Pop 7™ polymer in a 16 capillary 3130 XL Genetic Analyzer from Applied Biosystems using an internally manufactured 524 ROX fluorophore as sizing standard. The fragments used for fingerprinting were visualized as electropherograms in applied Biosystems software GeneMapper® version 3.7. Informative fragments were mostly found in the 100-450 bp range. All amplified fragments were classified based on the primer combination used and their size. The sample fingerprint data was converted to binary code, with "1" denoting the presence of the fragment and "0" the absence. Different binary matrices were constructed for comparative analysis depending on the kind of molecular marker. As shown in Table 4, the MSAP patterns were classified as follows: Pattern 1 when a comigrating amplified fragment was obtained from the DNA template digested by both restriction enzymes *Hpa*II and *Msp*I; Pattern 2 and 3 when an amplification fragment was obtained only from the DNA template digested by *Hpa*II or *Msp*I, respectively.

Table 4: MSAP patterns corresponding to different methylation states of the symmetric sequence CCGG, as revealed by the specificity of the restriction enzymes *Hpa* II and *Msp* I.

Restriction enzymes	MSAP patterns after enzymatic digestion			
	Pattern 1	Pattern 2	Pattern 3	Pattern 4
***Hpa* II**	1	1	0	0
***Msp* I**	1	0	1	0
Methylation state	Unmethylated	Hemi-methylated	Fully-methylated	Fully-methylated
Methylation position	None	External cytosine	Internal cytosine	External cytosine
Schematic representation	CCGG GGCC	CH3 CCGG GGCC	CH3 CCGG GGCC CH3	CH3 CCGG GGCC CH3

Before analysis of the embling versus mother plant population, we successfully verified, on a set of plants from the Caturra variety, that the same MSAP patterns were systematically generated whatever the plant age and the leaf developmental stage (data not shown). Hence, a possible developmental variability in the studied plant material does not seem to introduce any additional source of variation in

the methylation state. Nevertheless, in all experiments, only leaves from the same developmental stage were chosen.

Slide preparation and karyotyping

Root tips were harvested from individual adult emblings and placed in an aqueous solution of 8-hydroxyquinoline (2.5 mM) used as pre-treatment, for 4 h in darkness (2 h at 4°C plus 2 h at room temperature). A solution of Carnoy (absolute ethanol and glacial acetic acid, 3:1 v/v) was used to fix the tissues for at least 24 h at -20°C. Fixed material was then stored in 70% ethanol at 4°C until use for slide preparation. The stored root tips were used for slide preparations by employing the technique for cell dissociation of enzymatically macerated roots, as described previously by Herrera *et al.* [80]. Preparations were frozen in liquid nitrogen in order to remove the coverslips, stained with 4',6-diamidino-2-phenylindole, DAPI (1μg/mL), and mounted in Vectashield (Vector Laboratories, Peterborough, UK).

In order to determine the occurrence of chromosome modifications, individual plants of *C. arabica* regenerated by somatic embryogenesis showing a normal (2 plants) or variant (11 plants) phenotype were submitted to karyotype analysis. The Angustifolia, Bullata, Dwarf, Giant and Variegata phenotypic variants were analyzed. During slide examination, mitotic cells at metaphase or prometaphase stages were used for chromosome counting. Between 4 and 8 mitotic cells from each individual were analyzed to determine the chromosome number. The best examples were photomicrographed at metaphase to document the chromosome number and morphology. A Nikon Eclipse 90i epi-fluorescence microscope equipped with a digital, cooled B/W CCD camera (VDS 1300B Vosskühler ®) was used with the appropriate filter (UV-2E/C excitation wavelength 340-380).

Results

Frequency of phenotypic variants

Embling batches of hybrids HI and H3 obtained from both SCE and ESP were checked for phenotype variation at both nursery and field levels. The frequency of phenotypic variants assessed among more than 600 000 plants in the nursery was very low (approx. 0.1%) and not significantly affected by the proliferation system nor the hybrid variety (**Table 1**). Observation of around 200 000 emblings in the field two years after planting revealed roughly an additional 0.74% of abnormal phenotypes, still without any significant difference between the two proliferation systems and hybrids. Apart from these phenotypic variants, all the other studied trees flowered, grew and produced normally. In conclusion, the overall phenotypic variation rate obtained by pooling the data obtained both in the nursery and in the field was less than 1% and no significant differences were noted between the proliferation systems or between the hybrids used.

Table 1: Frequency of coffee phenotypic variants detected in the nursery and field on three year-old plants from two *C. arabica* hybrids depending on the type of proliferation system used in the industrial somatic embryogenesis process.

Proliferation system in the somatic embryogenesis process	Hybrid	Observations after 10 months in nursery				Observations after 36 months in field			
		No. of observed emblings	No. of variants	Somaclonal variation frequency (%)	3σ confidence interval*	No. of emblings	No. of variants	Somaclonal variation frequency (%)	3σ confidence interval*
Secondary embryogenesis	H1	117.115	148	0.13	[0.09-0.16]	51.131	373	**0.73**	[0.62-0.84]
	H3	121.894	159	0.13	[0.10-0.16]	49.126	390	**0.79**	[0.67-0.91]
	Total	239.009	307	0.13	[0.10-0.15]	100.257	763	**0.76**	[0.67-0.84]
Embryogenic suspension	H1	204.871	206	0.10	[0.08-0.12]	54.218	402	**0.74**	[0.63-0.85]
	H3	197.705	183	0.09	[0.07-0.11]	54.566	394	**0.72**	[0.61-0.83]
	Total	402.576	389	0.09	[0.08-0.11]	108.784	796	**0.73**	[0.65-0.80]

The variable analyzed was the proportion (p) of variant (p = X/n), where X was the number of variant and n the number of plants observed. A 3σ confidence limit for binomial distribution was calculated using the formula $p \pm 3\left(\sqrt{p(1-p/n)}\right)$ with a level of confidence of 99%.

Both proliferation systems generated the same kind of phenotypic variants (**Figures 2A, B**), with the Dwarf and Angustifolia types (**Figures 3E, B**) being the most frequent. Note that the secondary embryogenesis proliferation system specifically enhanced the occurrence of Dwarf variants, whereas the embryogenic suspensions favored the occurrence of the Angustifolia type. This latter phenotype can easily be detected and eliminated at the nursery level along with the Variegata variant (**Figure 3C**). A comparison of **Figures 2A** and **2B** clearly shows that elimination in the nursery is not efficient for the Dwarf type. This somaclonal variation is more easily observable 2-3 years after planting in the field thanks to the characteristic grouped canopy morphology and low yield. Similarly, the Giant and Bullata (**Figure 3F**) phenotypic variants could only be detected in the field on well-developed trees.

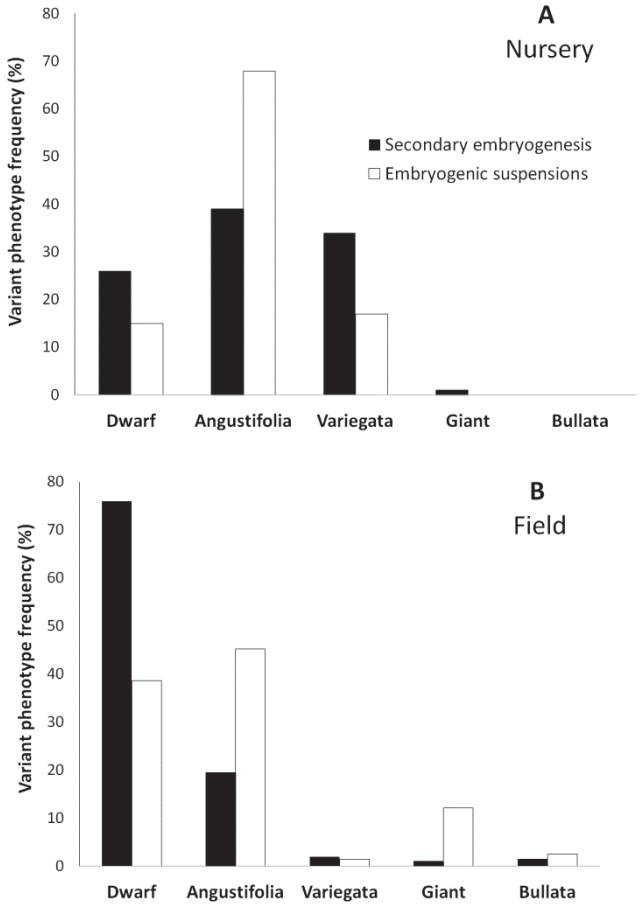


Figure 2. Proportions (%) of the different types of phenotypic variants in comparison to the total number of variants. Variants representing less than 1% of somatic embryogenesis-derived plants were observed in *C. arabica* embling batches at nursery (A) and field (B) levels, depending on the proliferation system used, i.e. secondary embryogenesis (SCE) or embryogenic suspension (ESP). In the nursery, the data were obtained from 239 009 emblings derived from SCE and 402 576 emblings from ESP. In the field, the data were obtained through the observation of 100 257 emblings derived from secondary embryogenesis and 108 784 from embryogenic suspensions.

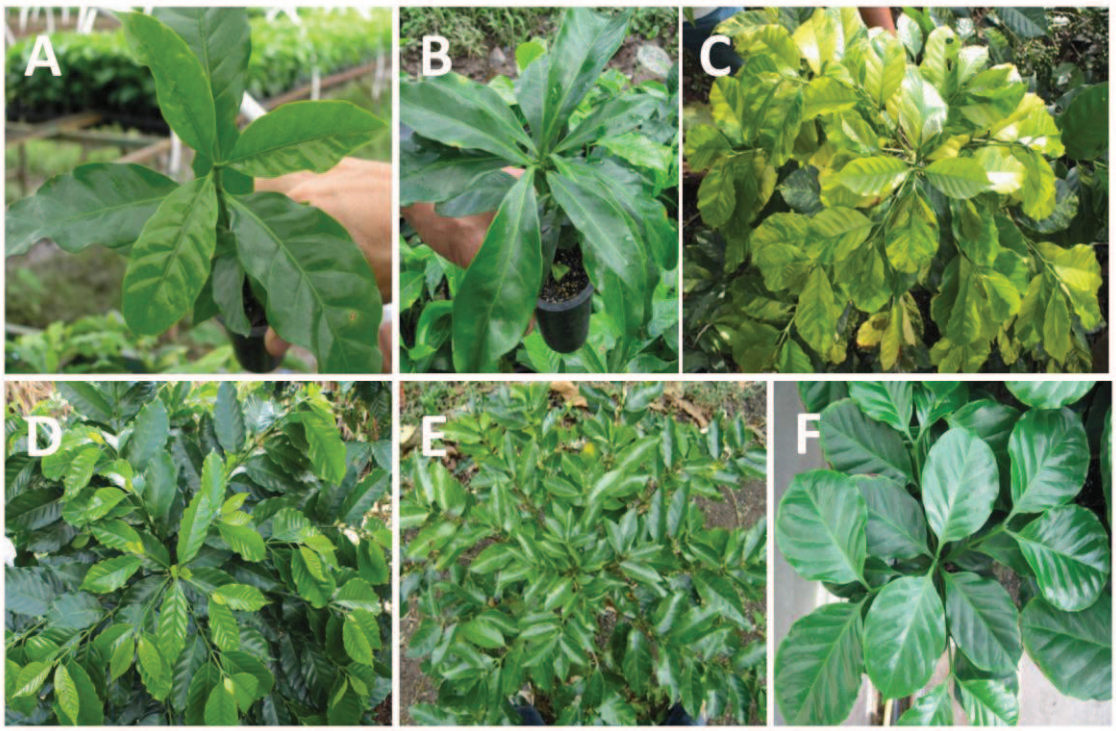


Figure 3. Examples of different *C. arabica* phenotypic variants in plants mass propagated through somatic embryogenesis. A, plant exhibiting a normal phenotype in the nursery; B, Angustifolia variant in the nursery; C, Variegata variant in the field; D, plant showing a normal phenotype in the field; E, Dwarf variant in the field; F, Bullata variant in nursery.

Locus specific polymorphisms revealed by AFLP

In order to verify the induction of molecular polymorphism by the somatic embryogenesis process, AFLP analysis (four primer combinations, **Table 2**) was carried out on mother plants and their derived emblings. From a total of 204 bands obtained, only one polymorphic fragment of 173 bp in size (*Eco*-ACT/*Mse*-AGT) shared by two emblings of hybrid H1 and exhibiting a normal phenotype, was found (**Table 3**). From a total of 198 bands obtained, no polymorphism was found in emblings of hybrid H3. All variants had the same AFLP pattern as the mother plants. For both hybrids, no significant quantitative effect on AFLP was detected when comparing SCE and ESP.

Table 3: Summary of AFLP data and observed polymorphisms among mother plants and emblings derived from secondary embryogenesis or embryogenic suspensions. Data were obtained for two *C. arabica* hybrids (H1 and H3) and compared with the patterns of the mother plants.

Proliferation system in the somatic embryogenesis process	Hybrid	No. of analyzed emblings	No. of fragments	Polymorphic fragments		Emblings showing polymorphisms		Total polymorphism **
				No.	(%)	No.	(%)	(%)
Secondary embryogenesis	H1	33	204	1 *	0.5	2	6	**0.03**
	H3	45	198	0	0	0	0	0

Embryogenic suspensions	H1	26	204	0	0	0	0	0
	H3	41	198	0	0	0	0	0

*Found in 2 emblings with normal phenotype (N°210 and 232) showing a new AFLP band Eco-ACT/Mse-AGT 173 bp

**Total polymorphism = [No. of polymorphic fragments/(No. of fragments x No. emblings)] x 100

Methylation changes revealed by MSAP

In order to evaluate the occurrence of possible epigenetic modifications in the micropropagated plants, a study on the alteration of methylation patterns was performed by MSAP analysis using eight primer combinations (**Table 2**). Only clear and reproducible bands were selected for the analysis. More than 395 fragments were considered. First, MSAP patterns were obtained from DNA digested by the two isoschizomers (*Hpa*II and *Msp*I), as illustrated in **Figure 4**. They were further compared with those of mother plants to classify the amplified fragments according to the methylation pattern, as shown in **Table 4**. The percentages of monomorphic fragments (pattern 1) were elevated and similar for both hybrids at nearly 91%. The remaining 9% of fragments (8.5% for H1 and 9.1% for H3) almost exclusively corresponded to pattern 3.

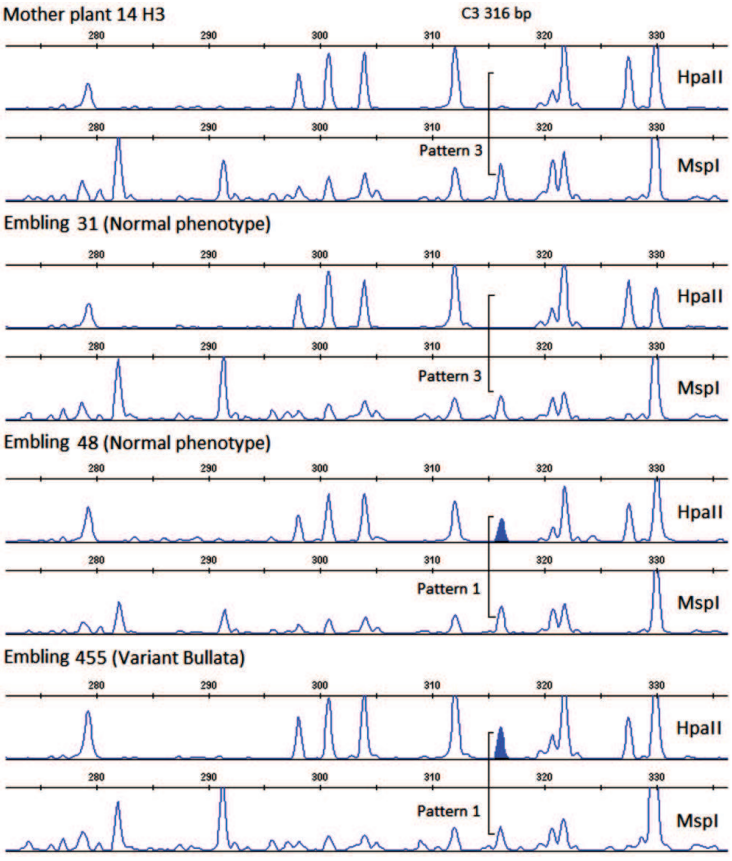


Figure 4. Representation of MSAP electropherograms observed for coffee mother plants and embling progeny using the isoschizomers *Hpa*II and *Msp*I. Illustration of the pattern variation obtained for normal and variant phenotypes within the embling progeny.

A comparison of amplification patterns in mother plants and their respective progeny are reported in **Table 5**. All differences between mother and derived plants were switches between patterns 1 and 3 and vice versa, i.e. likely modifications in the restriction ability of *Hpa*II. Among the polymorphic bands, eight bands corresponded to a change from pattern 3 in mother plants towards the unmethylated pattern 1 in emblings, suggesting demethylation. Changes associated with certain polymorphic MSAP bands were more frequent than others (**Table 5**). Seventy percent of the changes were linked to only five polymorphic bands. The detected polymorphism was very low and similar for both hybrids (**Table 6**) but slightly higher (0.07-0.18%) when compared to AFLP molecular markers. Similar to the AFLP results, the total MSAP polymorphism was not significantly different between the two proliferation systems nor between the two hybrid varieties.

Table 5: MSAP methylation patterns in mother plants and modified patterns in emblings. Relation with the type of *C. arabica* hybrid, type of proliferation system [secondary embryogenesis (SCE) and embryogenic suspension (ESP)] and regenerant phenotype.

Polymorphic MSAP fragments (size in bp)	MSAP methylation patterns		Proliferation system affected by the methylation change		Presence of the methylation change depending on the plant phenotype		No. of methylation changes for each fragment	
	Mother plants	Emblings	Hybrid H1	Hybrid H3	Normal	Variant	No.	(%)
C3- 107 bp	Pattern 3	Pattern 1	SCE, ESP	SCE, ESP	**+****	-	8	9.8
C2 -127 bp	Pattern 3	Pattern 1	0	ESP	**+**	-	1	1.2
C4 -134 bp	Pattern 3	Pattern 1	0	SCE	**+**	-	4	4.8
C1- 251 bp	Pattern1 *	Pattern 3	SCE, ESP	0	**+**	-	12	14.6
C3- 253 bp	Pattern 3	Pattern 1	0	SCE	**+**	-	2	2.4
C2- 302 bp	Pattern 3	Pattern 1	0	SCE, ESP	**+**	+	16	19.5
C3- 316 bp	Pattern 3	Pattern 1	0	SCE	**+**	+	4	4.8
C6 -370 bp	Pattern 3	Pattern 1	SCE, ESP	ESP	**+**	+	10	12.2
C8- 370 bp	Pattern 3	Pattern 1	SCE, ESP	SCE, ESP	**+**	+	12	14.6
C5- 387 bp	Pattern 1	Pattern 3	SCE	SCE	**+**	-	13	15.8
No. changes							82	

*Pattern 1: Fragment present in both HpaII and MspI restriction digests (1:1); Pattern 3: Fragment absent in HpaII digests and present in MspI digests (0:1). ** Relationship with a particular phenotype is indicated with (+) for presence and (-) for absence.

Most emblings showing changes in MSAP pattern had only one or two methylation polymorphisms (Figure 5). We did not find any emblings with more than three methylation polymorphisms. The same polymorphic bands were shared by plants from both proliferation systems and/or both hybrids in approximately half of the cases (Table 5). No relation was established between the variant phenotype and a particular MSAP pattern (Table 6).

Table 6: Overall MSAP data and methylation polymorphism among mother plants and emblings of *C. arabica* hybrids derived from secondary embryogenesis or embryogenic suspensions.

Proliferation in the somatic embryogenesis process	Hybrid	No. of emblings analyzed	No. of fragments	Methylation polymorphic fragments		Total polymorphism *(%)	3σ confidence intervals**
				No.	Percentage (%)		
Secondary embryogenesis	H1	33	399	5	1.2	**0.18**	[0.071-0.294]
	H3	45	396	7	1.7	**0.16**	[0.068-0.246]
		78		12	1.5	**0.17**	[0.098-0.238]
Embryogenic suspensions	H1	26	399	4	1.0	**0.18**	[0.057-0.309]
	H3	41	396	5	1.0	**0.07**	[0.006-0.129]
		67		9	1.1	**0.11**	[0.051-0.174]

*Total polymorphism = [No. of methylation polymorphic fragments/ (No. of fragments x No. emblings)] x 100.

**The variable analyzed was the proportion (p) of methylation polymorphisms (p=X/n), where X was the number of methylation polymorphisms and n the total number of fragments. A 3σ confidence limit for binomial distribution was calculated using the formula $p \pm 3\left(\sqrt{p(1-p/n)}\right)$ with a level of confidence of 99%.

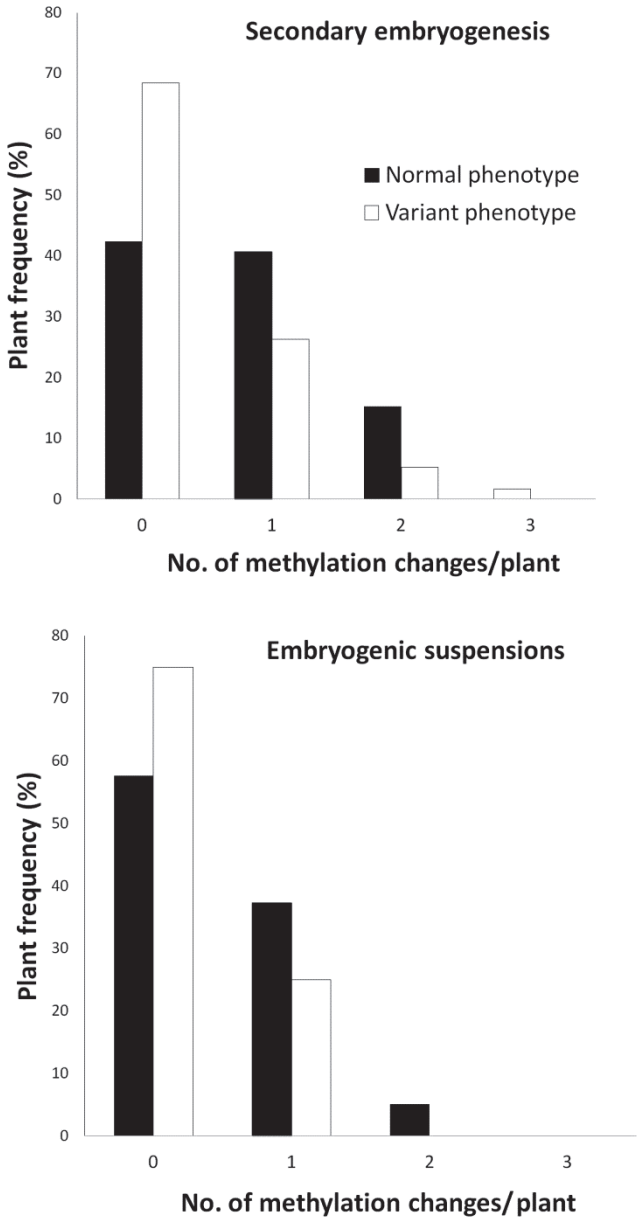


Figure 5. Methylation polymorphism accumulation in coffee emblings showing a normal *vs*. variant phenotype depending on the somatic embryogenesis process used. For the secondary embryogenesis process, data were derived from the analysis of 59 phenotypically normal and 19 variant emblings. For the embryogenic suspension process, 59 phenotypically normal and 8 variant emblings were studied.

Chromosome counting of somaclonal variants

Chromosome numbers were assessed in 2 phenotypically normal emblings and 11 somaclonal variants (**Fig. 6**). Table 7 shows that the two normal regenerants exhibited the expected chromosome number for the allotetraploid *C. arabica* species (2n=4x=44) whereas 7 of the 11 variant emblings showed a different chromosome number (aneuploids). In almost all cases, the aneuploid karyotypes were characterized by the loss of 1-3 chromosomes. One Angustifolia variant had an extra chromosome.

Moreover, the results showed that different chromosome numbers - including the normal number - could be observed for the same variant phenotype and that abnormal chromosome numbers were obtained for most of the variant phenotypes.

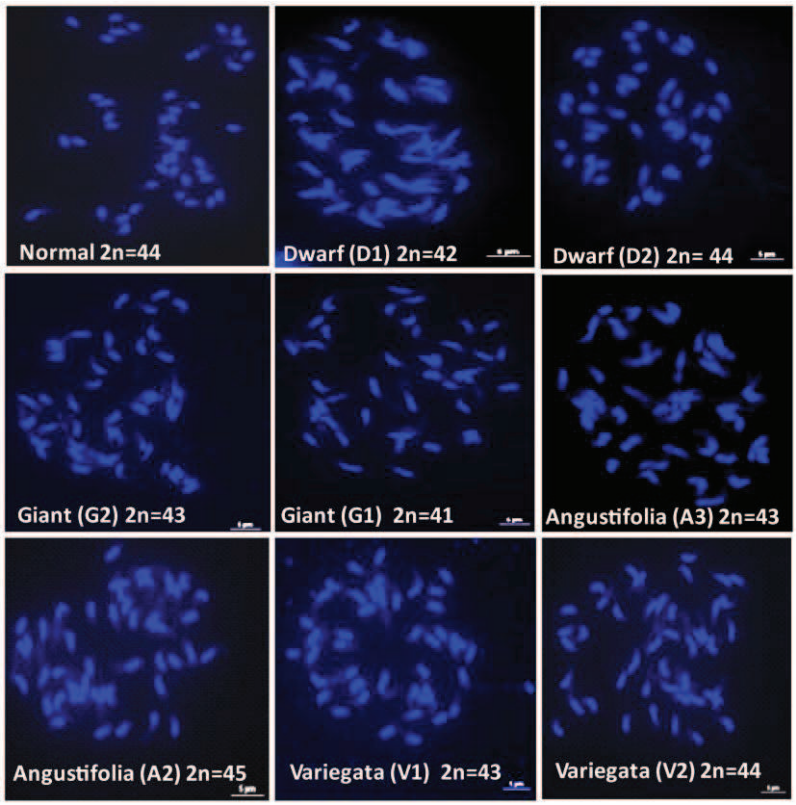


Figure 6. Mitotic cells at metaphase or prometaphase stages and observed ploidy levels of some normal and variant emblings from the allotetraploid *C. arabica* species (2n=44). Karyotype analyses were performed by counting chromosomes on four to eight clear metaphase spreads obtained from root tips of three year-old plants.

Table 7: Summary of chromosome counting in some normal *versus* variant *C. arabica* hybrids derived from somatic embryogenesis. The chromosome numbers obtained from root tips are indicated for 3 year-old emblings derived from embryogenic suspensions (ESP) or secondary embryogenesis (SCE) showing normal or abnormal phenotypes.

Embling phenotype (normal or variant)	Somatic embryogenesis process	Code	No. of metaphases analyzed	No. of chromosomes
Normal	ESP	N1	8	44
Normal	SCE	N2	7	44
Angustifolia	ESP	A1	4	44
Angustifolia	SCE	A2	8	**45**
Angustifolia	ESP	A3	8	**43**
Bullata	ESP	B1	4	44

Dwarf	SCE	D1	5	**42**
Dwarf	ESP	D2	4	44
Giant	ESP	G1	6	**41**
Giant	ESP	G2	6	**43**
Giant	SCE	G3	6	**43**
Variegata	ESP	V1	6	**43**
Variegata	SCE	V2	4	44

Discussion

Until now, relatively high somaclonal variation rates have been reported in *C. arabica* emblings, particularly with embryogenic suspensions [28,29,30]. The variant frequency was found to strongly increase in embryogenic suspensions after 6 months proliferation (25% after 12 months) in the presence of 4.52 µM 2,4-D [29]. The presence of disorganized rapid growth phases in tissue culture, such as callus and cell suspension cultures, has often been considered as one of the factors that cause SV [8,31]. In view of true-to-type propagation, we further established processes with limited disorganized rapid growth phases. The first was based on SCE proliferation in temporary immersion bioreactors enhanced by the addition of cytokinin, similar to methods described in other woody plants such as rubber [32], oak [33] and tea [34] and is currently used on a commercial scale for coffee. The second involves the proliferation of ESP in the presence of both auxin and cytokinin albeit at a very low level of 2,4-D (1.36 µM) with a short proliferation time (6 months). These conditions allow sufficient amplification of embryogenic material to ensure the cost-effectiveness of the industrial process. A large-scale phenotypic evaluation for each process: 230 000-400 000 emblings in the nursery and 100 000 in the field was performed to obtain valuable and accurate information on genetic stability. The present study showed for the first time that, for both somatic embryogenesis processes, the variant phenotype frequencies were extremely low (less than 1%) and not statistically different. Moreover no statistical difference could be noted between the two studied hybrids. This clearly demonstrates that it is possible to control SV by optimizing the ESP culture conditions. In another cultivated *Coffea* species, i.e. the diploid *C. canephora*, Ducos *et al*. [35] found 2-4% of a low-vigor phenotype off-type among 5 067 emblings derived from 5-7-month-old ESP. In contrast to the *C. arabica* protocol, all somatic embryogenesis steps for *C. canephora* were performed with cytokinins as sole source of growth regulator. Hence, in coffee, both auxin and cytokinin could probably be involved in the generation of SV. The Angustifolia, Variegata and Dwarf variants were the most frequent types, irrespective of the proliferation strategy employed, confirming previous studies conducted on a lesser scale [29, 30]. Interestingly, except for the Dwarf variant which seems to be

tissue culture specific, all the phenotypic variants observed in the somatic embryogenesis progenies were also observed among *C. arabica* seed-derived progeny. Indeed, phenotypic mutations are frequent in coffee nurseries, and Krug and Carvalho [36] previously characterized the numerous different morphologies in detail.

In several cases, AFLP markers have proved useful in the detection of genetic variation in tissue culture derived plants [20,37-40]. In coffee, the genetic stability of coffee emblings has been poorly evaluated by molecular markers and limited to plants derived from experimental protocols. The present study revealed no or extremely limited mutations at the DNA level in large-scale somatic embryogenesis-propagated plants**.** From a total sample of 145 plants analyzed belonging to two hybrid varieties, only 1 out of 204 bands was polymorphic in only two SCE-produced plants. ESP-derived plants showed no AFLP polymorphism when compared with mother plants. Our results significantly differed from those previously reported in *C. arabica*. In a first approach on only 27 plants, Rani *et al.* [41] used different DNA markers (RAPD, random amplified polymorphic DNA and SSR, simple sequence repeat) to assess the genetic integrity of *C. arabica* emblings obtained from embryogenic calli, and they found a higher polymorphism level (4%) in the nuclear genome. By performing RAPD analyses on Norway spruce emblings, Heinze and Schmidt [42] concluded that gross somaclonal variation was absent in their plant regeneration system. In contrast, RAPD and SSR markers allowed the detection of high genetic variation in cotton emblings regenerated in the presence of 2,4-D [39]. AFLP analysis of 24 rye emblings led to the scoring of 887 AFLP markers, among which 8.8% identified the same polymorphism in plants obtained independently, revealing putative mutational hot spots [38].

DNA cytosine methylation plays an important role in plant regulation and development [43,44]. Since the pioneer studies on maize emblings [12], substantial evidence has been obtained which indicates that demethylation can occur at a high frequency during somatic embryogenesis and can be an important cause of tissue culture induced variation [45]. DNA methylation has also been implicated in gene silencing and transposable element reactivation [24,46]. To our knowledge, epigenetic deregulation during coffee micropropagation has not yet been studied. The very low total methylation polymorphism values obtained for both somatic embryogenesis processes and both hybrids (0.07-0.18% range) indicated that the tissue culture procedures employed in coffee weakly affected DNA methylation of the regenerated plants. This finding is in accordance with the 0.87% total polymorphism recently found in *Freesia hybrid* emblings by Gao *et al.* [40]. Moreover, the low number of methylation polymorphisms per embling (range 1-3) confirmed that neither SCE nor ESP induced additional stress at the methylation level during embryogenic material proliferation. In contrast, a significantly higher accumulation of methylation changes in some emblings has been regularly observed in other species [20, 21, 47, 48]. For example, in grapevine, most emblings showed

between zero and three changes, similar to our findings, but a few accumulated up to 18 [21]. It has also been well demonstrated that auxin levels strongly alter DNA methylation of embryogenic cell cultures [49]. However, similar to our results, examples of stable MSAP patterns have already been reported using bamboo tissues at different developmental stages of somatic embryogenesis [50]. The timing of plant regeneration from proliferating callus cultures could be crucial for the appearance of variation. In callus-derived hop plants, an increase in the variation was detected by MSAP in prolonged callus cultures [20]. Our results demonstrated that very few changes are possible by limiting both the auxin level and culture duration.

MSAP markers have already been successfully used to demonstrate epigenetic instabilities (methylation alteration) induced by somatic embryogenesis in a great variety of plants, such as the ornamental flower *Freesia*, banana, barley, grapevine and maize [20,21,40,51-53], also indicating that demethylation events were generally the most frequent. Although occurring at low frequency, our results also indicated demethylation events and mainly the loss of methylation in the internal cytosine of the 5'-CCGG-3´ sequence to produce a new *Hpa*II band not detected in mother plants but present in the amplification pattern of the isoschizomer *Msp*I. The detection of the same MSAP polymorphic fragments in independent plant samples from different hybrids and somatic embryogenesis processes suggests the existence of hotspots of DNA methylation changes in the genome. The existence of non-randomly behaving methylation polymorphic fragments in micropropagated plants has already been described using Met-AFLP [20,48] and MSAP [21,53-55].

Gross changes, such as variation in ploidy level, number of chromosomes and structural changes, are mitotic aberrations that represent major genomic alterations of *in vitro* plants often generated during *in vitro* proliferation and differentiation [12,56-58]. Variations in chromosome number and structure have been described among emblings for several species [59-62]. We demonstrated that gross changes occur during coffee somatic embryogenesis and are related to SV, whereas genetic and epigenetic (methylation) alterations are very weak. Until now, by using flow cytometry analysis, normal ploidy levels were reported in coffee emblings [29,63] but chromosome counting was not performed in these studies. The sensitivity of flow cytometry analysis was probably not sufficient to identify aneuploid plants. Nevertheless, the presence of mitotic aberrations, including double prophase, lagging chromosomes, aneuploids and polyploid cells, has previously been described in leaves [64,65] and embryogenic calli [64] of *C. arabica* but not in the later steps of somatic embryogenesis, and without establishing any relation with SV. The presence of aneuploidy has also been well documented in embryogenic calli of *Hordeum vulgare* [66] and sweet orange [60].

The mechanisms underlying SV remain largely theoretical and unclear [16]. Thus it is often difficult to correlate a well-described genetic or epigenetic mechanism to a variant phenotype. For example,

although DNA methylation has often been suggested as a possible cause of SV, a number of studies have reported high levels of methylation variation with no effect on the plant phenotype [20,21,52]. Another example is given by oil palm emblings, approximately 5% of which exhibit the 'mantled' phenotype affecting the formation of floral organs in both male and female flowers. Although a decrease in DNA methylation was observed, it was not possible to determine the nature of the epigenetic deregulation [67]. In the present study, it was possible to reveal a large proportion of aneuploid karyotypes in different variant phenotypes, hence showing that chromosomal rearrangements could be directly involved in the occurrence of phenotypic variation. The addition or subtraction of a single chromosome has a greater impact on phenotype than whole genome duplication, i.e. polyploidy [68]. It was clearly demonstrated in *Arabidopsis thaliana* that certain phenotypic traits are strongly associated with the dosage of specific chromosomes and that chromosomal effects can be additive [69]. Similarly, in maize seedlings, trisomic plants showed characteristic features such as reduced stature, tassel morphology changes and the presence of knots on the leaves, suggesting a phenotypic effect caused by the altered copy of specific chromosome related genes [70]. A similar mechanism could explain most of the variant phenotypes observed in *C. arabica*. The observation of a variant phenotype in plants with the expected chromosome number could be explained by the coexistence of monosomic and trisomic chromosomes in the same genome or by other chromosomal-like structural changes associated with undetected deletions, duplications, inversions or translocations of specific chromosomal segments [10,71]. The karyotype analyses performed in the present study were limited to chromosome counting and did not enable observation of such chromosomal alterations.

Conclusions

This report shows that somatic embryogenesis is reliable for true-to-type and large-scale propagation of elite varieties in the *C. arabica* species. Both ESP and SCE ensured high proliferation rates along with very low SV rates, as observed through massive phenotypic observations in a commercial nursery and field plots. Molecular analysis (AFLP and MSAP) performed on 145 emblings derived from two proliferation processes and two different F1 hybrids showed that polymorphism between mother plants and emblings was extremely low. Consequently, it can be concluded that genetic and epigenetic alterations were also particularly limited during somatic embryogenesis in our controlled culture conditions. *C. arabica* is a young allopolyploid still having the most of its genes in duplicated copies [72]. It could be hypothesized that the impact of genetic or epigenetic variations on phenotype was restricted because of the buffer effect due to polyploidy. The main change in most of the rare phenotypic variants was aneuploidy. Although further studies are necessary for an accurate

understanding of the chromosome anomalies involved in the acquisition of a particular phenotype, it is now obvious that mitotic aberrations play a major role in SV in coffee. The identification and use of molecular markers at the heterozygous state in mother plants (i.e. polymorphic between the two parental lines) would be required to investigate further this type of chromosome abnormality. Current studies based on the use of long-term embryogenic cultures [73] are aimed at establishing the full range of cytological, genetic and epigenetic (with a special focus on transposable elements) mechanisms underlying SV.

Acknowledgements

The authors thank Clément Poncon, owner of the Coffee Research Center 'La Cumplida', and José Martin Hidalgo for their interest and invaluable assistance during this work.

References

1. Etienne H, Dechamp E, Barry Etienne D, Bertrand B (2006) Bioreactors in coffee micropropagation (Review). Braz J Plant Physiol 18: 45-54.

2. Miguel C, Marum L (2011) An epigenetic view of plant cells cultured *in vitro*: somaclonal variation and beyond. J Exp Bot 62: 3713-3725.

3. Jähne A, Lazzeri PA, Jäger Gussen M, Lörz H (1991) Plant regeneration from embryogenic cell suspensions derived from anther cultures of barley (*Hordeum vulgare* L.). Theor Appl Genet 82: 74-80.

4. Rival A, Beulé T, Barre P, Hamon S, Duval Y, et al. (1997) Comparative flow cytometric estimation of nuclear DNA content in oil palm (*Elaeis guineensis* Jacq.) tissue cultures and seed-derived plants. Plant Cell Rep 16: 884-887.

5. Lu S, Wang Z, Peng X, Guo Z, Zhang G, et al. (2006) An efficient callus suspension culture system for triploid bermudagrass (*Cynodon transvaalensis* x *C. dactylon*) and somaclonal variations. Plant Cell Tiss Org Cult 87: 77-84.

6. Lambé P, Mutambel HSN, Fouche JG, Deltour R, Foidart JM, et al. (1997) DNA methylation as a key process in regulation of organogenic totipotency and plant neoplastic progression. In Vitro Cell Dev Biol Plant 33: 155–162.

7. Von Aderkas P, Bonga J (2000) Influencing micropropagation and somatic embryogenesis in mature trees by manipulation of phase change, stress and culture environment. Tree Physiol 20: 921–928.

8. Karp A (1994) Origins, causes and uses of variation in plant tissue cultures. In: Vasil IK, Thorpe TA, editors. Plant cell and tissue culture. Dordrecht: Kluwer Academic Publishers. pp. 139-152.

9. Bukowska B (2006) Toxicity of 2,4-Dichlorophenoxyacetic acid – Molecular mechanisms. Polish J of Environ Stud 15: 365-374.

10. Larkin PJ, Scowcroft WR (1981) Somaclonal variation: a novel source of variability from cell cultures for plant improvement. Theor Appl Genet 60: 197–214.

11. Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 43: 179-188.

12. Kaeppler SM, Phillips RL (1993) DNA methylation and tissue culture-induced DNA methylation variation in plants. In Vitro Cell Dev Bio Plant 29: 125-130.

13. Peschke VM, Phillips RL (1992) Genetic implications of somaclonal variation in plants. Adv Genet 30: 41–75.

14. Duncan RR (1997) Tissue culture-induced variation and crop improvement. Adv Agron 58: 201-240.

15. Sahijram L, Soneji J, Bollamma K (2003) Analyzing somaclonal variation in micropropagated bananas (*Musa* spp.). In Vitro Cell Dev Bio Plant 39: 551-556.

16. Bairu MW, Aremu OA, Van Staden J (2011) Somaclonal variation in plants: causes and detection methods. Plant Growth Regul 63:147-173.

17. Law J, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 11: 204-220.

18. Saze H (2008) Epigenetic memory transmission through mitosis and meiosis in plants. Sem Cell Dev Biol 19: 527-536.

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

20. Bednarek PT, Orłowska R, Koebner RMD, Zimny J (2007) Quantification of tissue-culture induced variation in barley (*Hordeum vulgare* L.). BMC Plant Biol 7: 10.

21. Schellenbaum P, Mohler V, Wenzel G, Walter B (2008) Variation in DNA methylation patterns of grapevine somaclones (*Vitis vinifera* L.). BMC Plant Biol 8: 78.

22. Peschke VM, Phillips RL, Gegenbach BG (1987) Discovery of transposable element activity among progeny of tissue culture-derived maize plants. Science 238: 804-807.

23. Mckenzie NL, Wen LY, Dale J (2002) Tissue-culture enhanced transposition of the maize transposable element Dissociation in *Brassica oleracea* var. 'Italica'. Theor Appl Genet 105: 23-33.

24. Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. Nat Rev Genet 8: 272-285.

25. Anthony F, Bertrand B, Quiros O, Lashermes P, Berthaud J, et al. (2001) Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. Euphytica 118: 53-65.

26. Bertrand B, Alpizar E, Llara L, Santacreo R, Hidalgo M, et al. (2011) Performance of Arabica F1 hybrids in agroforestry and full-sun cropping systems in comparison with pure lines varieties. Euphytica 181: 147-158.

27. Etienne H, Bertrand B, Montagnon C, Bobadilla Landey R, Dechamp E, et al. (2012) Un exemple de transfert technologique réussi en micropropagation : la multiplication de *Coffea arabica* par embryogenèse somatique. Cah Agric 21: 115-24.

28. Söndahl MR, Lauritis JA (1992) Coffee. In: Hammerschlag FA, Litz RE, editors. Biotechnology of Perennial Fruit Crops. Wallingford: CAB International. pp. 401-420.

29. Etienne H, Bertrand B (2003) Somaclonal variation in *Coffea arabica*: effects of genotype and embryogenic cell suspension age on frequency and phenotype of variants. Tree Physiol 23: 419-426.

30. Etienne H, Bertrand B (2001) The effect of the embryogenic cell suspension micropropagation technique on the trueness to type, field performance, bean biochemical content and cup quality of *Coffea arabica* trees. Tree Physiol 21: 1031-1038.

31. Rani V, Raina S (2000) Genetic fidelity of organized meristem-derived micropropagated plants: a critical reappraisal. In Vitro Cell Dev Bio Plant 36: 319-330.

32. Hua HW, Huang TD, Huang HS (2010) Micropropagation of self-rooting juvenile clones by secondary somatic embryogenesis in *Hevea brasiliensis*. Plant Breeding 129: 202-207.

33. Mallón R, Covelo P, Vieitez AM (2012) Improving secondary embryogenesis in *Quercus robur*: Application of temporary immersion for mass propagation. Trees Structure and Function 26: 731-741.

34. Akula A, Becker D, Bateson M (2000) High-yielding repetitive somatic embryogenesis and plant recovery in a selected tea clone, 'TRI-2025', by temporary immersion. Plant Cell Rep 19: 1140-1145.

35. Ducos JP, Alenton R, Reano JF, Kanchanomai C, Deshayes A, et al. (2003) Agronomic performance of *Coffea canephora* P. trees derived from large-scale somatic embryo production in liquid medium. Euphytica 131: 215-223.

36. Krug CA, Carvalho A (1951) The genetics of *Coffea*. Advances in Genetics 4: 127-158.

37. Hornero J, Martínez I, Celestino C, Gallego FJ, Torres V, et al. (2001) Early checking of genetic stability of cork oak somatic embryos by AFLP analysis. Int J Plant Sci 162: 827-833.

38. De la Puente R, Gonzalez AI, Ruiz ML, Polanco C (2008) Somaclonal variation in rye (*Secale cereale* L.) analyzed using polymorphic and sequenced AFLP markers. In Vitro Cell Dev Biol Plant 44: 419-426.

39. Jin S, Mushke R, Zhu H, Tu L, Lin Z, et al. (2008) Detection of somaclonal variation of cotton (*Gossypium hirsutum*) using cytogenetics, flow cytometry and molecular markers. Plant Cell Rep 27: 1303-1316.

40. Gao X, Yang D, Cao D, Ao M, Sui X, et al. (2010) *In vitro* micropropagation of *Freesia hybrida* and the assessment of genetic and epigenetic stability in regenerated plantlets. J Plant Growth Regul 29: 257–267.

41. Rani V, Singh KP, Shiran B, Nandy S, Goel S, et al. (2000) Evidence for new nuclear and mitochondrial genome organizations among high-frequency somatic embryogenesis-derived plants of allotetraploid *Coffea arabica* L. (Rubiaceae). Plant Cell Rep 19: 1013-1020.

42. Heinze B, Schmidt J (1995) Monitoring genetic fidelity *vs* somaclonal variation in Norway spruce (*Picea abies*) somatic embryogenesis by RAPD analysis. Euphytica 85: 341-345.

43. Zhang M, Kimatu JN, Xu K, Liu B (2010) DNA cytosine methylation in plant development. Review. J Genet Genomics 37: 1-12.

44. Vanyushin BF, Ashapkin VV (2011**)** DNA methylation in higher plants: Past, present and future. *Biochim Biophys Acta* 1809**:** 360-368**.**

45. Smulders MJM, de Klerk GJ (2011) Epigenetics in plant tissue culture. Plant Growth Reg 63: 137-146.

46. Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: when genetics meets genomics. Nat Rev Genet 3: 329-341.

47. Li X, Yu X, Wang N, Feng Q, Dong Z, et al. (2007) Genetic and epigenetic instabilities induced by tissue culture in wild barley (*Hordeum brevisubulatum* (Trin.) Link). Plant Cell Tiss Org Cult 90: 153–168.

48. Fiuk A, Bednarek PT, Rybczyński JJ (2010) Flow cytometry, HPLC-RP and met AFLP analyses to assess genetic variability in somatic embryo-derived plantlets of *Gentiana pannonica* scop. Plant Mol Biol Rep 28: 413-420.

49. LoSchiavo F, Pitto L, Giuliano G, Torti G, Nuti Ronchi V, et al. (1989) DNA methylation of embryogenic carrot cell cultures and its variations as caused by mutation, differentiation, hormones, and hypomethylating drugs. Theor Appl Genet 77: 325-331.

50. Gillis K, Gielis J, Peters H, Dhooghe E, Orpins J (2007) Somatic embryogenesis from mature *Bambusa balcooa* Roxburgh as basis for mass production of elite forestry bamboos. Plant Cell Tiss Org Cult 91: 115-123.

51. Peraza Echeverria S, Herrera Valencia VA, Kay AJ (2001) Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplification polymorphism (MSAP). Plant Sci 161: 359–367.

52. Yu X, Li X, Zhao X, Jiang L, Miao G, et al. (2011) Tissue culture-induced genomic alteration in maize (*Zea mays*) inbred lines and F1 hybrids. Ann Appl Biol 158: 237-247.

53. Baránek M, Křižan B, Ondružíková E, Pidra M (2010) DNA-methylation changes in grapevine somaclones following *in vitro* culture and thermotherapy. Plant Cell Tiss Org Cult 101: 11-22.

54. Guo WL, Wu R, Zhang YF, Liu XM, Wang HY, et al. (2007) Tissue culture-induced locus-specific alteration in DNA methylation and its correlation with genetic variation in *Codonopsis lanceolata* Benth et Hook. Plant Cell Rep 26: 1297-1307.

55. Díaz Martínez M, Nava Cedillo A, Guzmán López JA, Escobar Guzmán R, Simpson J (2012) Polymorphism and methylation patterns in *Agave tequilana* Weber var. 'Azul' plants propagated asexually by three different methods. Plant Sci 185-186: 321-330.

56. Karp A (1991) On the current understanding of somaclonal variation. In: Miflin BJ, editor. Oxford surveys of plant molecular and cell biology. Vol 17. Oxford: Oxford University Press. pp. 1-58.

57. Larkin PJ (1987) Somaclonal variation: history, method and meaning. Iowa State 1 Res 61: 393-43.

58. Neelakandan AK, Wang K (2012) Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. Plant Cell Rep 31: 597– 620.

59. Al Zahim MA, Ford Loyd BV, Newbury HJ (1999) Detection of somaclonal variation in garlic (*Allium sativum* L.) using RAPD and cytological analysis. Plant Cell Rep 18: 473-477.

60. Hao YI, Deng XX (2002) Occurrence of chromosomal variations and plant regeneration from long-term-cultured citrus callus. In Vitro Cell Dev Biol Plant 38: 472-476.

61. Mujib A, Banerjee S, Dev Ghosh P (2007) Callus induction, somatic embryogenesis and chromosomal instability in tissue culture-raised hippeastrum (*Hippeastrum hybridum* cv. United Nations). Propag Ornam Plants 7: 169-174.

62. Leal F, Loureiro J, Rodriguez E, Pais MS, Santos C, et al. (2006) Nuclear DNA content of *Vitis vinifera* cultivars and ploidy level analyses of somatic embryo-derived plants obtained from anther culture. Plant Cell Rep 25: 978-985.

63. Sanchez Teyer LF, Quiroz Figueroa F, Loyola Vargas V, Infante D (2003) Culture-induced variation in plants of *Coffea arabica* cv. Caturra rojo, regenerated by direct and indirect somatic embryogenesis. Mol Biotechnol 23: 107–115.

64. Ménendez Yuffá A, Da Silva R, Rios L, Xena de Enrech N (2000) Mitotic aberrations in coffee (*Coffea arabica* cv. 'Catimor') leaf explants and their derived embryogenic calli. Electron J Biotechnol 3: 0-5. ISSN 0717-3458.

65. Zoriniants SE, Nosov AV, Monforte Gonzalez M, Mendes Zeel M, Loyolas Vargas VM (2003) Variation of nuclear DNA content during somatic embryogenesis and plant regeneration of *Coffea arabica* L. using cytophotometry. Plant Sci 164: 141-146.

66. Gözükirmizi N, Ari S, Oraler G, Okatan Y, Ünsal N (1990) Callus induction, plant regeneration and chromosomal variations in barley. Acta Botanica Neerlandica 39: 379-387.

67. Jaligot E, Beulé T, Baurens FC, Billote N, Rival A (2004) Search for methylation-sensitive amplification polymorphism associated with the 'mantled' variant phenotype in oil palm (*Elaeis guineensis* Jacq.). Genome 47: 224-228.

68. Birchler JA, Veitia RA (2007) The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell 19: 395-402.

69. Henry IM, Dilkes BP, Miller ES, Burkart Waco D, Comai L (2010) Phenotypic consequences of aneuploidy in *Arabidopsis thaliana*. Genetics 186: 1231-45.

70. Makarevitch I, Harris C (2010) Aneuploidy causes tissue specific qualitative changes in global gene expression patterns in maize. Plant Physiol 152: 927–938.

71. Fukuoka H, Kawata M, Tkaiwa M (1994) Molecular changes of organelle DNA sequences in rice through dedifferentiation, long-term culture or the morphogenesis process. Plant Mol Biol 26: 899-907.

72. Cenci A, Combes MC, Lashermes P (2012) Genome evolution in diploid and tetraploid *Coffea* species as revealed by comparative analysis of orthologous genome segments. Plant Mol Biol 78: 135-145.

73. Ribas A, Dechamp E, Bertrand B, Champion A, Verdeil JL, et al. (2011) *Agrobacterium tumefaciens*-mediated genetic transformation of *Coffea arabica* (L.) is highly enhanced by using long-term maintained embryogenic callus. BMC Plant Biol **11:** 92.

74. Murashige T, Skoog FA (1962) Revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15: 473-497.

75. Etienne H (2005) Protocol of somatic embryogenesis : Coffee (*Coffea arabica* L. and *C. canephora* P.). In: Jain SM, Gupta PK, editors. Protocols for somatic embryogenesis in woody plants. Series: Forestry Sciences. Vol 77. The Netherlands: Springer edn. pp. 167-179.

76. Teisson C, Alvard D (1995) A new concept of plant *in vitro* cultivation liquid medium : temporary immersion. In: Terzi M et al., editors. Current Issues in Plant Molecular and Cellular Biology. Dordrecht: Kluwer Academic Publishers. pp. 105-110.

77. Yasuda T, Fujii Y, Yamaguchi T (1985) Embryogenic callus induction from C*offea arabica* leaf explants by benzyladenine. Plant Cell Physiol 26: 595-597.

78. Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA miniprepration: version II. Plant Mol Biol Rep 1: 19-21.

79. Vos P, Hogers R, Bleeker M, Rijans M, Van de Lee T, et al. (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414.

80. Herrera JC, D'hont A, Lashermes P (2007) Use of fluorescence *in situ* hybridization as a tool for introgression analysis and chromosome identification in coffee (*Coffea arabica* L.). Genome 50: 619-626.

CHAPTER III

Long-term cell cultures for understanding the basis of somaclonal variations in coffee: assessment of phenotypic, genetic, epigenetic and chromosomal changes in *Coffea arabica* (L.) regenerated plants

Chapter written as an article to be submitted to *New Phytologist.*

Long-term cell cultures for understanding the basis of somaclonal variations in coffee: assessment of phenotypic, genetic, epigenetic and chromosomal changes in *Coffea arabica* (L.) regenerated plants

Bobadilla Landey R[1], Cenci A[2], Guyot R[4], Bertrand B[1], Déchamp E[1], Herrera JC[5], Lashermes P[3], Etienne H[1,*]

[1] *Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) –Département des Systèmes Biologiques. UMR-RPB, 911 Avenue Agropolis, BP 64501, 34394 Montpellier, France*

[2] *Bioversity, Montpellier*

[3] *IRD, UMR RPB, 911 Avenue Agropolis, BP 64501, 34394 Montpellier, France*

[4] *IRD, UMR DIADE 911 Avenue Agropolis, BP 64501, 34394 Montpellier France*

[5] *Centro Nacional de Investigaciones de Café, CENICAFE (www.cenicafe.org). PO BOX. 2427 Manizales, Caldas, Colombia*

* Corresponding author herve.etienne@cirad.fr

Introduction

Since long time *in vitro* culture has been associated with the outcome of variant phenotypes ranging from moderate to strikingly high rates of variation (Karp 1991, Skirvin et al. 1994). The term Somaclonal Variation (SV) defines the genetic or phenotypic variation observed in clonally propagated plants derived from tissue culture (Larkin and Scowcroft, 1981). Somatic embryogenesis (SE) processes intended to mass propagation classically include an embryogenic material proliferation step *via* cell suspension cultures or secondary (cycling) embryogenesis. However, the duration of the proliferation step has often been associated with increased SV rates (Larkin and Scowcroft 1981, Karp 1991, Duncan 1997, Lee and Phillips 1998). Growth regulators like auxin and/or cytokinin are generally indispensable to sustain the cell and embryo multiplication, unfortunately their involvement in the occurrence of SV has been reported (Nehra et al. 1992, Duncan 1997, Bairu et al. 2006). Although some somaclonal variants have been positively evaluated at agronomic level (Skirvin et al. 1994, Jain 2001, Biswas et al. 2009), the loss of genetic fidelity is detrimental for commercial purposes when the main objective is strictly the clonal propagation of the elite plant material to ensure the maintenance of its agronomic characteristics.

Little information about SV related mechanisms is still available in studies performed either on cell cultures (Tanurdzic et al., 2008) or regenerated plants (Hao and Deng 2002, Smýkal et al. 2007, Rodriguez-Lopez et al. 2010). Early attempts to explain SV have evidenced some implication of chromosomal abnormalities (Lee and Phillips 1988, Karp 1991, Duncan 1997). It has been demonstrated that aneuploid cell lines increased in frequency and severity in prolonged culture times (reviewed by D'Amato 1985). Chromosomal instabilities and ploidy alterations were observed in cell cultures of some species including barley, orange, pea and oil palm (Gözükirmizi et al. 1990, Hao and Deng 2002, Kumar and Mathur 2004, Giorgetti et al. 2011). The involvement of epigenetic events in SV has been proposed because some variants reverted back to the normal phenotype after some generations of vegetative propagation or sometime after field planting e.g. unstable abnormal phenotypes of strawberry reported by Biswas et al. (2009). Genetic or supposedly epigenetic in origin, a wide range of mechanisms has been recently described as possible causes of SV including TEs (Transposable Elements) transposition and small RNA directed DNA methylation (Miguel and Marum 2011, Smulders and de Klerk 2011, Anjanasree et al. 2012). DNA methylation in the form of 5-Methylcytosine (5mC) is a major epigenetic mark involved in gene expression and playing an important role in plant regulation and development (Law and Jacobsen 2010, Zhang et al. 2010, Vanyushin and Ashapkin 2011). DNA methylation in plants commonly occurs at cytosine bases in all sequence contexts the symmetric CG and CHG (where H indicates A, T or C) and the asymmetric CHH contexts. DNA hypomethylation can produce a severe affectation on plant phenotype as noted in *ddm1* Arabidopsis mutants (Kakutani et al. 1996). Tissue culture induces sequence specific

methylation variation that can be detected by DNA molecular markers like Met-AFLP (Bednarek et al. 2007, Fiuk et al. 2010) or MSAP (Peraza-Echeverria et al. 2001, Diaz-Martínez et al. 2012). The loss of DNA methylation in cell cultures can be maintained upon plant regeneration, and be stably inherited over subsequent generations (Stroud et al. 2012). However, the effect of DNA methylation deregulation associated with SV remains unclear.

Strong evidences have been provided that abiotic and biotic stresses play an important role in the TE reactivation (Grandbastien et al. 1998; Kalendar et al. 2000; Jiang et al. 2003). An elevated transposition of TEs (class I retrotransposons and class II transposons) has been positively correlated to demethylating events induced by tissue culture in Arabidopsis and maize (Hirochika 2000, Miura et al. 2001, Barret et al. 2006, Smith et al. 2011). TEs could potentially alter gene expression when insertions occur inside or near to encoding protein genes finally affecting phenotype (Kumar and Bennetzen 1999, Sato et al. 2011).

Coffea arabica is an allotetraploid (2n=4x=44) with a basic chromosome number of 11 chromosomes. Previous studies on various cultivars have shown that SE can generate variable and relatively high levels (5-90%) of phenotypic variants among the regenerated plants depending on *in vitro* culture conditions (Söndahl and Bragin 1991, Etienne and Bertrand 2003). The effect of culture ageing on the induction of phenotypic variation has been studied in coffee (Etienne and Bertrand 2003). Whatever the genotype analyzed, the degree of SV was low (1.3%) in plants directly produced from embryogenic callus and increased progressively in line with the culture age reaching 25% in average in plants produced from 12 month-aged suspensions in presence of 2,4-D (2,4-dichlorophenoxyacetic acid). However the nature of the mechanisms occurring during culture ageing and underlying the SV has not yet been investigated.

The objective of the present work was to identify the cytological and molecular mechanisms associated with SV occurrence as well as the stochastic character of the phenomenon. In this aim, by establishing three independent embryogenic callus lines from a same variety, we investigated the effects of culture ageing on the derived plants by regenerating different plant batches after more or less extended cell proliferation periods (up to 27 months) in an auxin enriched medium. Phenotypic, genetic, epigenetic and chromosomal changes were systematically assessed on somatic seedlings. Molecular analyses were performed with AFLP and two related techniques MSAP and SSAP in order to investigate changes in the DNA methylation patterns and the activity of different TEs.

Material and Methods

Plant material and somatic embryogenesis

Plant material corresponds to somatic embryo-derived plants (somatic seedlings) of *C. arabica* var. Caturra regenerated from 3 independent embryogenic cell lines (A, B and C) established with distinct embryogenic calli and regenerated after 3 different culture periods i.e. 4, 11, 27 months (**Fig. 1**). The dispositive included different cell lines with the aim of identifying the stochastic or time dependant character of SV occurrence. Leaf explants for SE were collected from mother trees maintained in a greenhouse located at IRD (Montpellier, France).

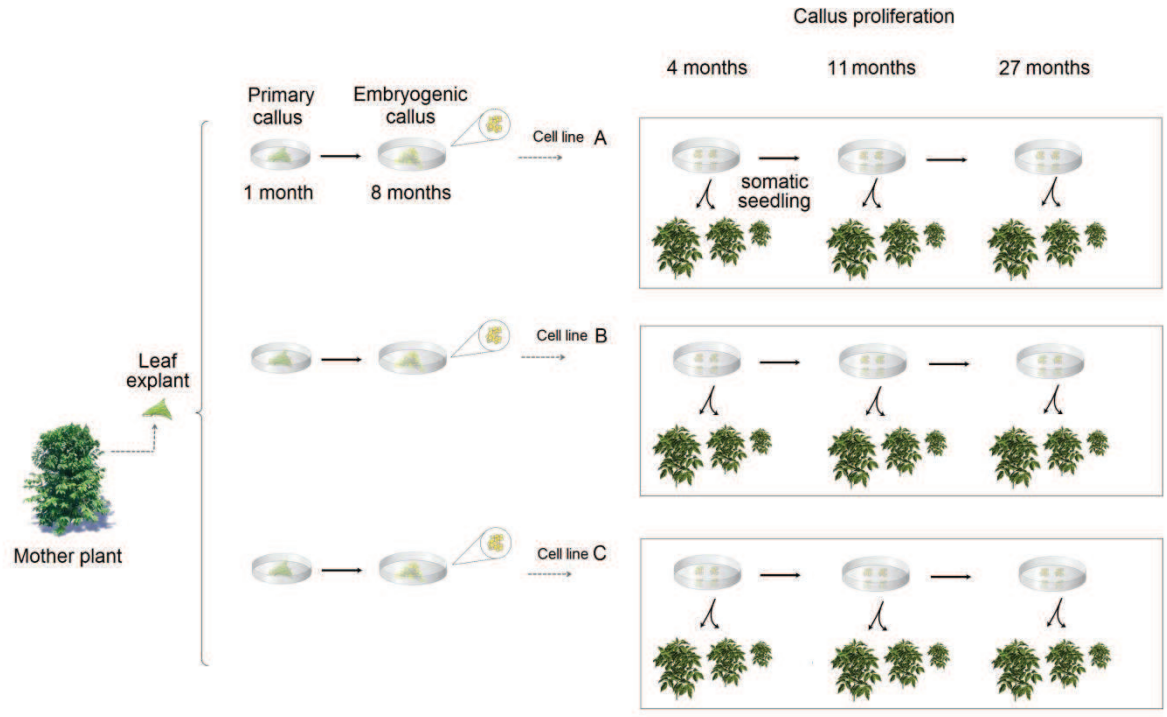


Figure 1. Schematic representation of the experimental design based on the establishment and long-term maintenance of independent cell lines and further regeneration of somatic embryos-derived plantlets at different culture ages (4, 11 and 27 months) for each of them.

Embryogenic callus induction

Embryogenic calli were induced as previously described by Etienne (2005). Briefly, young leaves' pieces of 1 cm^2 from mother-tree were surface-sterilized and used as explants. Leaves were disinfected by immersion in 10% calcium hypochlorite solution containing 1% Tween 80 for 20 min followed by 8% calcium hypochlorite solution for 10 min and rinsed 4 times with sterile water. The explants were cultured for 1 month on MS/2 (Murashige and Skoog 1962) callogenesis medium

containing 4.52 µM 2,4-D, 4.92 µM indole-3-butyric acid (IBA) and 9.84 µM iso-pentenyladenine (iP) until production of primary calli. The explants were then transferred for 6-8 months to MS/2 embryogenic callus production medium containing 4.52 µM 2,4-D and 17.76 µM 6-benzylaminopurine (6-BA). Yellow embryogenic callus appeared on the initial oxidized callus. All the media used in this work were supplemented with 30 g/L sucrose and had the pH adjusted to 5.7 prior to the addition of 2.8 g/L phytagel. The media were autoclaved at 120°C and 1.1 kg/cm2 for 20 min, and then 25 ml of medium was dispensed into each baby food jar. These steps were performed at 27°C in the dark.

Establishment of cell cultures and plant regeneration

Three long-term cell cultures were successfully established on semi-solid medium by isolating initially three distinct primary embryogenic calluses that proliferated independently on MS/2 embryogenic callus production medium (Etienne 2005) enriched with 4.5 µM 2, 4 D and 12 µM 6-BA and solidified with 2.8 g/L phytagel. Long-term maintenance of the cell cultures was achieved during 3 years by one-month proliferation cycles, i.e. by monthly transfers of the most active part of the embryogenic callus (located on the upper part) onto fresh medium. Plants could be regenerated after 27 months proliferation but all cell cultures from 36 months lost the ability to regenerate somatic embryos.

Cell cultures were submitted to embryo regeneration conditions by subculturing approximately 200 mg callus every four weeks twice on 'R' regeneration medium (Etienne 2005) containing 17.76 µM 6-BA and twice on 'M' maturation medium containing 1.35 µM 6-BA until plantlets developed. Plantlets exhibiting one leaf pair were transferred to the same maturation medium devoid of plant growth regulators and enriched with 1 g/l active charcoal until. During the entire regeneration process up to acclimatization, the cultures were maintained under a 14 h photoperiod (20 $\mu mol.m^{-2}s^{-1}$ light intensity) at 26 °C. Three to 4 cm tall plantlets were acclimatized in the greenhouse. Plantlets were transferred in pots to a horticultural substrate, placed in a partially shaded acclimatization tunnel (80-90% relative humidity) for 1 month and then cultivated at 24-26°C at 60% of relative humidity. Afterwards they were moved to a nursery (2000–4000 lux) until phenotypic observations.

Phenotype characterization

The detection of variants was performed on 180 total somatic seedlings from the 3 embryogenic cell lines (50-65 plants per embryogenic line) after 12 and 24 months in nursery. For each culture age (4, 11, 27 months), 40-75 somatic seedlings plants were observed. The phenotypic characterization consisted in assessing plant height, internodes number and length, leaf length, leaf width and collar diameter. Other characteristics like presence or absence of domatia, leaf shape, canopy shape and orthotropic or plagiotropic development were also observed and compared to those

of normal plants. All observations were used to characterize the phenotype and eventually to relate it to an already described variant or mutant phenotype.

Molecular analyses

One hundred twenty four *C. arabica* cv. Caturra plants derived from the three cell cultures A, B and C and from the different culture ages were subjected to molecular markers to estimate their genetic and epigenetic conformity in comparison with mother plants (**Table 1**). In this aim, AFLP, MSAP and SSAP analyzes were performed on both the mother plants and on the SE progeny to evidence polymorphisms in the molecular markers patterns. DNA was extracted from fully developed young leaves for molecular analysis. Genomic DNA was isolated from 100 mg of lyophilized leaves using Dellaporta buffer (Dellaporta 1983) containing sodium dodecyl sulfate (SDS) detergent and sodium bisulfite 1% w/v to avoid leaf oxidation. DNA was purified in spin-column plates as described in the DNeasy plant kit protocol from QIAGEN.

Table 1. Plant material used in molecular marker analyses. Somatic embryogenesis derived plants from *C. arabica* var. caturra were regenerated from cell cultures after variable proliferation periods. Plants were grouped and analyzed based on the embryogenic cell line (A, B, C) and cell culture age (4, 11, 27).

Cell culture age (months)	Cellular lines	No. of plants analyzed
4	A	10
	B	20
	C	15
11	A	ND
	B	16
	C	17
27	A	15
	B	20
	C	11
Total		**124**

AFLP markers

AFLP analysis was carried out as described by Vos et al. (1995). Selective amplification include a total of five EcoR1/MspI primer combinations using 5-FAM or 5-HEX fluorescently labeled *Eco*R1 (+3) and unlabeled *Mse*I (+3) primers (**Table 2**). PCR program for selective amplification was performed under the following conditions: 3 min at 94°C, following of 12 cycles of 45 s at 94°C, 45 s at 65°C and 1 min at 72°C; the annealing temperature was decreased by 0.7°C per cycle from a starting point of 65°C during this stage (Touchdown), with a final round of 25 cycles of 94°C for 45 s, 56°C for 45 s, 72°C for 1 min and a final extension of 72°C for 1 min. MSAP and S-SAP molecular markers were performed under the same conditions. Two repetitions using different DNA extractions were performed for each primer combination and for all molecular markers to reduce the possibility of technical artifacts.

Table 2. Primer combinations used for AFLP and MSAP analyses**.**

AFLP primer combinations (Eco+3 ^{marked}/Mse+3)	MSAP primer combinations (Eco-3/Hpa+2 ^{marked})
Eco-ACT/Mse-AGT	C1 Eco-AAC/HPA-AA
Eco-AGG/Mse-AGT	C2 Eco-AAC/HPA-AT
Eco-CGC/Mse-CCA	C3 Eco-AGG/HPA-AA
Eco-CAC/Mse-CCA	C4 Eco-AGG/HPA-AT
Eco-CTC/Mse-CCA	C5 Eco-ACT/HPA-CA
	C6 Eco-ACT/HPA-CT
	C7 Eco-AGA/HPA-CA
	C8 Eco-AGA/HPA-CT

Fluorescent dyes for marked primers correspond to 5'-FAM ™ and 5'-HEX ™

MSAP molecular markers

MSAP analysis was performed as described by Reyna-López et al. (1997) with minor adaptations for capillary electrophoresis. The MSAP protocol is an adaptation of the AFLP method for the evaluation of different states of methylation in the symmetric sequence CCGG. In the MSAP protocol, the frequent cutting endonuclease (*Mse*I) is replaced by the two restriction enzymes isoschizomers *Hpa*II and *Msp*I with different sensitivities to the methylation state of CCGG sequence (**Table 3**). Specifically, *Hpa*II is able to recognize and cut when the CCGG sequence is unmethylated or hemi-methylated on the external cytosine. *Msp*I is able to cut when CCGG sequence is

unmethylated or if the internal cytosine is fully or hemi-methylated. Both *Hpa*II and *Msp*I are unable to cut if the external cytosines are fully methylated. Selective amplification included a total of 8 primer combinations per isoschizomer (Table 2). *Hpa*II (+2) primers were fluorescently labeled with 5-FAM or 5-HEX while *Eco*R1 (+3) remained unlabeled. The ensemble of fragments from both digestions was grouped by primer combination, fragment size and sample to determinate all individual profiles. After fragment classification three possible EcoRI/HpaII and EcoRI/MspI MSAP patterns were found by presence (1) or absence (0).

Table 3. MSAP DNA electrophoretic pattern as revealed by the specificity of the restriction enzymes *Hpa*II and *Msp*I to different methylation states of the symmetric sequence CCGG.

Restriction enzymes	**MSAP DNA electrophoretic patterns after enzymatic digestion**			
	Pattern 1	**Pattern 2**	**Pattern 3**	**Pattern 4**
HpaII	1	1	0	0
MspI	1	0	1	0
Methylation state	Unmethylated	Hemi-methylated	Fully-methylated	Fully-methylated
Methylation position	None	External cytosine	Internal cytosines	External cytosines
Schematic representation	CCGG GGCC	CH_3 CCGG GGCC	CH_3 CCGG GGCC CH_3	CH_3 CCGG GGCC CH_3

SSAP molecular markers

SSAP was performed as described by Waugh *et al* (1997) with minor modifications to capillary electrophoresis. SSAP protocol is an adaptation of AFLPs markers to discover polymorphisms associated to a sequence of interest like TEs or genes. Transposable elements sequences used in the present study were identified in genomic data available for coffee trees: *C. canephora* and *C. arabica*. Seventeen coffee BACs (Bacterial Artificial clone Chromosomes) sequences were retrieved from databases using SRS@EBI (http://srs.ebi.ac.uk/, accessions GU123896, GU123899, GU123898, GU123894, GU123897, GU123895, HQ696508, HQ696510, HQ696509, HQ696511, HQ834787 and HQ832564 for *C. arabica* and EU164537, HQ696512, HQ696507, HQ696513 and HM635075 for *C. canephora*) and 131,412 BAC-end sequences (Dereeper A, Guyot R et al., accepted manuscript in PMB). Using these resources, transposable elements were identified based on sequence similarities (BLAST) against Repbase (Jurka et al. 2005) (http://www.girinst.org/),

fine structure analysis using dotter alignments (Sonnhammer and Durbin 1995) or identified through assembly using AAARF algorithm (DeBarry et al., 2008). Finally a sequence database of transposable elements comprising Class I retrotransposons and Class II transposons were constructed and used for SSAP-based primer design. Different elements were used for primer design including 13 different *Copia* and *Gypsy* LTR retrotransposon families, one Sine and 2 different transposons families (MuLEs and Mariner families; **Table 4**). Primers were designed within a maximum distance of 200 bp at both extremities of each type of elements to find nearby polymorphic TEs insertions associated to *Eco*RI restriction sites. Specific TEs primers were fluorescently labeled either with 5-FAM, 5-HEX or 5-Tamra while *Eco*R1 (+1) remained unlabeled (Table 4). SSAP molecular marker was then performed to indentify polymorphisms possibly linked with transposable elements mobility. A total of 35 primers were designed and used in a previous screening. Only clear and repetitive profiles were used in the final analysis.

Table 4. List of SSAP primers used in the study.

Selective retrotransposon primers	**Sequence 3' to 5'**
Gypsy 13a$^{5'\text{-FAM}}$	CCGTAGGGTTCGGAGTCTTC
Gypsy 13b$^{5'\text{-HEX}}$	TTCCTTACAGGGACGACAGC
Copia 28a$^{5'\text{-FAM}}$	TTGTAGTATGCCAACCCACCT
Copia 28b$^{5'\text{-HEX}}$	AGAGTTGGGGTTTGGGAATC
Copia 9a$^{5'\text{-FAM}}$	TGCTGTCTCCAAAATTAAGGG
Copia 9b$^{5'\text{-HEX}}$	CAGGGTGAATGTAGGAATCTTGA
Gypsy 30a$^{5'\text{-FAM}}$	CCTAAGAGGGAATCCGACCT
Gypsy 30b$^{5'\text{-HEX}}$	CGTACCTTGTACCGTGGAGG
Copia 57a$^{5'\text{-FAM}}$	TCTCAACAAAACATAATCTCCTTCT
Copia 57b$^{5'\text{-HEX}}$	CACTTTTGCTGCAGATGTTGA
1696Rom37a$^{5'\text{-FAM}}$	TCAACCTTGGGATCCTTGAA
1696Rom37b$^{5'\text{-HEX}}$	TTCAAGGATCCCAAGGTTGA
Gypsy 5a$^{5'\text{-TAMRA}}$	ATTGCTGAATTTGGGAACCA
Gypsy 18a$^{5'\text{-FAM}}$	CCTTGGAGTTCCTCCCTCAT
Gypsy 18b$^{5'\text{-HEX}}$	CCTCTTGTCGATCCTGAAGC
Copia 30a$^{5'\text{-FAM}}$	CAAGCCATGCTCCTACCATT
Copia 30b$^{5'\text{-HEX}}$	CAAGCCATGCTCCTACCATT

Gypsy 39a$^{5'\text{-FAM}}$	TTGCATCCAAAAGTCCCTCT
Gypsy 39b$^{5'\text{-HEX}}$	TATTATTGCGCAGGCTGATG
Copia 22a$^{5'\text{-HEX}}$	CAGGTATCTTGGGAGGAAAACA
Copia 22b$^{5'\text{-FAM}}$	TGTTTTCCTCCCAAGATACCTG
Gypsy 15b$^{5'\text{-FAM}}$	AACTCTTCGGGAACATGGTG
Gypsy 15c$^{5'\text{-HEX}}$	TTATTGCACAGGCTGACGAC
Copia 33a$^{5'\text{-FAM}}$	ACGTCCGATGCAGTCTTTTT
Copia 33b$^{5'\text{-HEX}}$	TAAAACGTCCTGGCTTGACC
SINE-FAM1a$^{5'\text{-FAM}}$	AATGGCTCAGTGGCCACCA
SINE-FAM 1b$^{5'\text{-Hex}}$	TGGTGGCCACTGAGCCATT
SINE-FAM 2a$^{5'\text{-FAM}}$	CGTAAGAATTGACCCGGA
SINE-FAM 2b$^{5'\text{-Hex}}$	TCCGGGTCAATTCTTACG
Selective transposons primers	
MuLE-Deless$^{5'\text{-FAM}}$	AGGGATAATATCAGAAACCTC
MuLE -Altilio$^{5'\text{-FAM}}$	GGGAAAATCGCCAATTTAGTC
MuLE-Adriano$^{5'\text{-FAM}}$	GCAAAATGTTCCAATGATC
Mutor$^{5'\text{-FAM}}$	CTAATTTCTCAGATTATTTGC
MuLE-Michel$^{5'\text{-FAM}}$	CGAAGTCTCGGCCGAGA
Mariner Cafe$^{5'\text{-FAM}}$	TACTCCCTCCGTCCCACTTTA
***Fluorescent dyes 5'-FAM™, 5'-HEX™, 5'-Tamra™**	

Capillary electrophoresis and data analysis

PCR products were separated by capillary electrophoresis with Pop 7™ polymer in a 16 capillary 3130 XL Genetic Analyzer from Applied Biosystems. Fragment size was estimated using an internally manufactured 524 ROX fluorophore as sizing standard. The fragments used for fingerprinting were visualized as electropherograms using applied Biosystems software GeneMapper® version 3.7. Informative fragments were mostly found in the 100-450 bp range. All amplified fragments were classified based on the primer combination and their size. The sample fingerprint data was converted to binary code, with (1) denoting the presence of the fragment and (0) its absence. Different binary matrices were constructed for comparative analysis depending on the kind of molecular marker. Global genetic and epigenetic levels were estimated as total polymorphisms. Total polymorphism = [no. of polymorphic fragments/(no. of fragments x no. emblings)] x 100. A 3σ

confidence limit for binomial distribution was calculated to analyze statistical difference between the global genetic and epigenetic levels derived from different molecular markers using the formula $p \pm 3\sqrt{p(1-p)/n}$ with a confidence level of 99%. The variable analyzed was the proportion (p) of polymorphisms ($p = x/n$), where "x" was the number of polymorphic fragments and "n" the total number of fragments observed.

Slide preparation and chromosome counting

Root tips were collected from individual somatic seedling and placed in an aqueous solution of 8-hydroxyquinoline (2.5 mM) used as pre-treatment, for 4 h in darkness (2 h at 4°C plus 2 h at 21°C room temperature). A solution of Carnoy (absolute ethanol and glacial acetic acid, 31 v/v) was used to fix the tissues for at least 24 h at -20°C. Fixed material was then stored in 70% ethanol at 4°C until slide preparation. The stored root tips were used for slide preparations by employing the technique for cell dissociation of enzymatically macerated roots, as described previously by Herrera et al. (2007). Preparations were frozen in liquid nitrogen to remove the coverslips, stained with 4', 6-Diamidino-2-Phenylindole, Dihydrochloride DAPI (1µg/mL), and mounted in Vectashield (Vector Laboratories, Peterborough, UK). During slide examination, mitotic cells at metaphase or prometaphase stages were used for chromosome counting. Between 4 and 8 mitotic cells from each individual were analyzed to determine the chromosome number. The best examples were photomicrographed at metaphase to document the chromosome number and morphology. A Nikon Eclipse 90i epi-fluorescence microscope equipped with a digital, cooled B/W CCD camera (VDS 1300B Vosskühler ®) was used with the appropriate filter (UV-2E/C excitation wavelength 340-380).

Results

Effect of culture age on somaclonal variation

Different sets somatic seedlings from 3 different culture ages (4, 11 and 27 months old cell cultures) and regenerated from 3 different cell lines (A, B and C) were phenotyped (Fig. 1). From the different cell cultures only those from 4, 11 and 27 months old were able to regenerate plants. After 36 months of proliferation, all the embryogenic lines lost the capability to regenerate plants. Based on plant architecture and leaf morphology we identified a clear pattern of phenotypic abnormalities among somatic seedlings after 12 months of development in nursery. We found a predominant aberrant phenotype characterized by lower size, oval-elliptic leaves and close canopy (Fig. 2D), corresponding to "Bullata" coffee mutants described by Krug (1939).

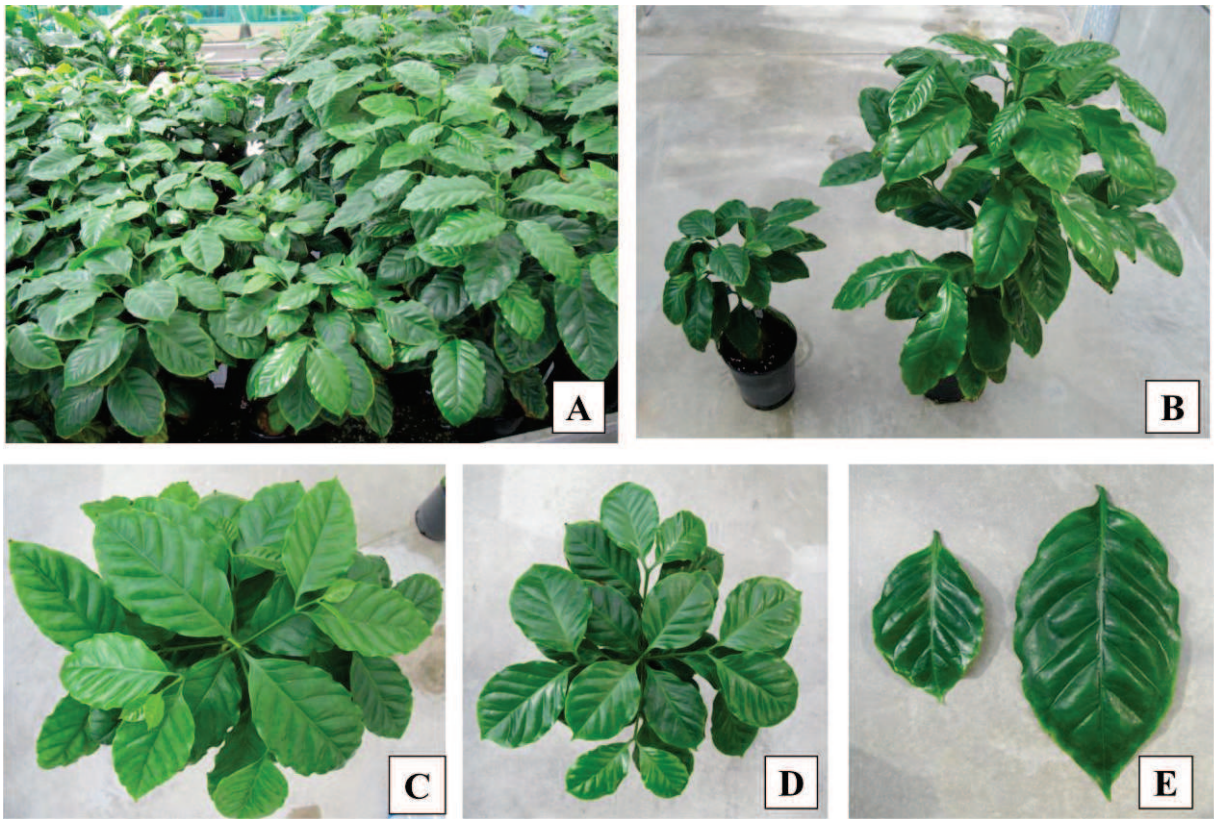


Figure 2. Morphological differences between normal and *C. arabica* variant phenotypes observed in somatic seedlings regenerated after two different cell culture times (4 and 27 months). (A and B) Differences in plant size between somatic seedlings derived from 4 and 27 months old cell cultures. Size after 8 months of parallel development in nursery. (C) Phenotypically normal somatic seedling from 4 months old cell culture. (D) Bullata variant phenotype from 27 months old culture. (E) Comparison between leaves from normal and Bullata phenotype. Bullata leaves (left) are rounder and smaller than those from true-to-type coffee somatic seedlings (right).

Somaclonal variants were observed in somatic seedlings from long-term cell cultures i.e. 11 and 27 months calluses affecting in the later almost all the regenerated plants, albeit all plants derived from younger cell cultures (4 months) exhibited normal phenotypes (**Fig. 2C**, **Fig. 3**). Thirty percents of somatic seedlings from 11 months cell cultures corresponded to abnormal phenotypes (20 and 40% for cell lines B and C, respectively). From a total of nine variant plants over 30 plants observed, eight were Dwarf and one Bullata. Among 75 somatic seedlings from 27 months old cultures, ninety four percents of presented abnormal phenotype (100, 92 and 92% for cell lines A, B and C, respectively). Only four plants presented normal phenotypes. The mutants corresponded to 4 Dwarf phenotype, one Angustifolia and the rest of the plants were Bullata phenotype.

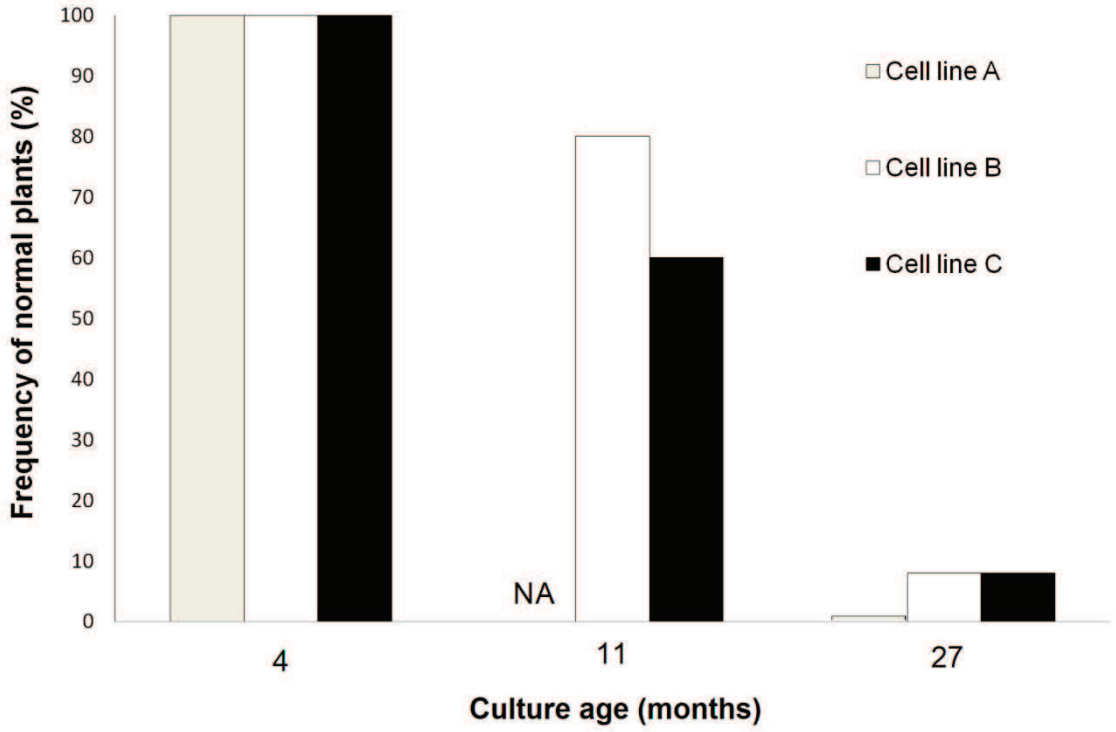


Figure 3. Frequency of normal plants found somatic seedlings derived from cell cultures of 4, 11 and 27 months. Data were obtained from 180 observed plants regenerated from 3 independent cellular lines (A, B, C). NA means not analyzed.

No genetic changes revealed by AFLP in the somatic seedlings

AFLP were conducted to evidence genetic polymorphisms associated with somatic embryogenesis and more particularly with the age of the embryogenic culture. Five primer combinations produced 182 total fragments with an average of 36 fragments per primer combination (**Table 5**). No differences were found between the electrophoretic profiles of mother plants and the 124 somatic seedlings studied. SSAP molecular markers were performed to uncover polymorphisms associated with TEs' activity possibly activated by tissue culture (i.e. as novell insertions). A total of 1003 SSAP fragments were generated after 45 selective SSAP amplification resulting in 22 fragments per primer combination (**Table 6**). No polymorphic fragments were observed between the electrophoretic profiles of mother plants and those of the *in vitro* progeny (**Fig. 6**). The number of fragments produced after selective amplification was in a range of 25-133 depending on the element tested. Both SSAP and AFLP analyses on somatic seedlings indicated that significant levels of genetic variations were not enhanced neither by short nor by long-term multiplications periods.

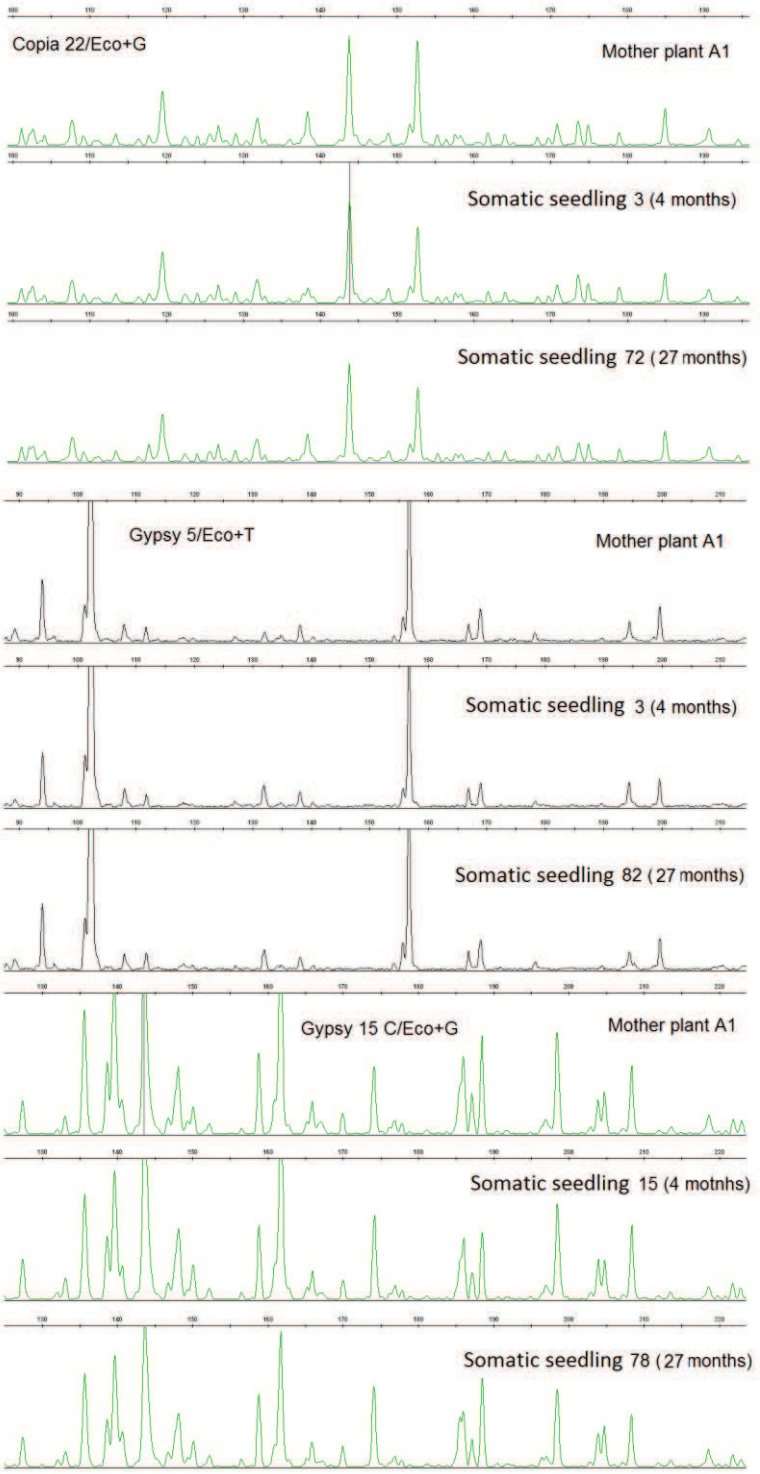


Figure 6. Examples of SSAP electropherograms observed for coffee mother plants and somatic seedlings derived from 4 months and 27 months cell cultures. Illustration of monomorphic electrophoretic profiles.

Table 5. Overall AFLP data and observed polymorphisms among mother plants and somatic seedlings regenerated from cell cultures with different ages (4, 11 and 27 months). All plants corresponded to *C. arabica* cv. Caturra.

Cell culture age (months)	No. of analyzed somatic seedlings	No. of fragments	Polymorphic fragments		Somatic seedlings showing polymorphisms		Total polymorphisms (%)*
			No.	(%)	No.	(%)	
4	45	182	0	0	0	0	0
11	33	182	0	0	0	0	0
27	46	182	0	0	0	0	0

*Total polymorphism = [No. of polymorphic fragments/ (No. of fragments x No. somatic seedlings)] x 100.

Table 6. Overall SSAP data and observed polymorphisms among *C. arabica* mother plants cv. caturra and somatic seedlings derived from 4 and 27 months cell cultures.

TE primer	Culture age (months)	No. of analyzed somatic seedlings	No. of primer couples	No. of SSAP fragments	No. of polymorphic fragments
Copia 22 a*	4	45	4	72	0
	27	45	4	72	0
Copia 22 b*	4	45	4	63	0
	27	45	4	63	0
Copia 28 a	4	45	4	96	0
	27	45	4	96	0
Copia 30 b	4	45	4	133	0
	27	45	4	133	0
Copia 33 a	4	45	4	87	0
	27	45	4	87	0
Copia 57 b	4	45	1	25	0
	27	45	1	25	0
Gypsy 5 a	4	45	4	73	0

	27	45	4	73	0
Gypsy 13 a	4	45	4	87	0
	27	45	4	87	0
Gypsy 15 c	4	45	4	93	0
	27	45	4	93	0
Gypsy 18 a	4	45	4	82	0
	27	45	4	82	0
Gypsy 18 b	4	45	4	77	0
	27	45	4	77	0
Mariner Cafe	4	45	1	31	0
	27	45	1	31	0
Mut-Adriano	4	45	1	27	0
	27	45	1	27	0
Mutor	4	45	1	27	0
	27	45	1	27	0
Sine-Fam. 1 a.	4	45	1	30	0
	27	45	1	30	0

* Elements marked with “a” or “b” correspond to forward or reverse primers designed from a particular element.

Low induction of methylation changes in short and long-term cell cultures

MSAPs markers were performed in coffee somatic seedlings to investigate the methylation changes induced by tissue culture and more particularly by the embryogenic culture age. Eight primer combinations from the separate EcoRI/HpaII and EcoRI/MspI digestions produced a total of 402 MSAP fragments (**Table 7**). The pattern distribution resulted in 348 non methylated sites with pattern 1 (1/1), 52 fragments with internally methylated cytosines with pattern 3 (0/1) and only 2 externally hemi-methylated cytosines with pattern 2 (1/0). The interpretation of the possible methylation state of these patterns indicated that 86.6% of the fragments corresponded to non methylated sites (pattern 1) whereas 13.4% were methylated (patterns 2 and 3).

Table 7. Overall MSAP data and methylation polymorphisms among *C. arabica* mother plants cv. caturra and somatic seedlings derived from three different cell lines of different ages (4, 11 and 27 months).

Culture age (months)	No. of analyzed somatic seedlings	No. of fragments	Polymorphic MSAP fragments		Total no. of changes	Total polymorphism (%)*	3σ Confidence intervals**
			No.	(%)			
4	45	402	9	2.2	27	**0.149**	[0.063-0.235]
11	33	402	6	1.5	15	**0.113**	[0.026-0.201]
27	46	402	9	2.2	16	**0.087**	[0.022-0.151]

*Total polymorphism = [No. of polymorphic fragments/ (No. of fragments x No. somatic seedlings)] x 100.

** A 3σ confidence limit for binomial distribution was calculated using the formula $P \pm 3\sqrt{P(1-P)/n}$ with a confidence level of 99%. The variable analyzed was the proportion (P) of methylation polymorphisms (P=X/n), where X was the number of methylation polymorphisms and n the total number of fragments.

The mother plants' profiles and those from somatic seedlings were then compared to find polymorphic methylation fragments (**Fig. 4**). A total of 58 methylation changes associated to 14 polymorphic fragments were detected in the whole plant population analyzed (**Table 8**). Somatic seedlings derived from 4 months-old cell cultures presented 27 methylation changes representing 46.5% of the overall changes (Table 7). Plant batches from 11 and 27 months old cell cultures presented similar amount of changes, i.e. 15 and 16 changes representing 26 and 28% respectively. The rate of epigenetic variation in somatic seedlings from 4, 11 and 27 months old cell cultures was calculated as a total methylation polymorphism for global estimation. Total polymorphism ranged between 0.087 and 0.149% disregarding the culture age (Table 7). These rates were not statistically different for short and long-term cell cultures-derived somatic seedlings indicating no significant effect of culture age at methylation level. We also noted that methylation changes arose from both methylated and unmethylated sites. The methylation patterns of the 14 polymorphic fragments are indicated in Table 8. Eight fragments methylated in mother plants were found unmethylated in somatic seedlings (the internal methylation of $CC^{Met}GG$ changing to the unmethylated state of CCGG sequence). The remaining 6 fragments were found unmethylated in mother plants and methylated in somatic seedlings evidencing methylation events (the unmethylated state of CCGG changing to the methylated state of $CC^{Met}GG$ sequence).

Table 8. MSAP methylation patterns in *C. arabica* mother plants cv. Caturra and modified patterns in regenerated somatic seedlings. Relation with a specific culture age (4, 11 and 27 months) and phenotype. Presence of a methylation change in a specific cell line was indicated by the cellular line code (A, B or C.).

Polymorphic MSAP fragments (size in bp)	**MSAP Methylation patterns***		**Methylation changes relation with a particular culture age (month) and cell line**			**Presence of the methylation change depending on the plant phenotype****		**Number of methylation changes per fragment**	
	Mother plants	Somatic seedlings	4	11	27	Normal	Variant	No.	(%)
C3 107	Met.	Unmet.	B	C		+	-	10	17.2
C8 115	Met.	Unmet.			C	-	+	1	1.7
C3 122	Met.	Unmet.	B			+	-	1	1.7
C2 127	Unmet.	Met.	B		A	+	+	4	6.9
C1 133	Unmet.	Met.	B		A	+	+	5	8.6
C7 182	Unmet.	Met.	B		A	+	+	2	3.4
C3 202	Unmet.	Met.	C	B, C	C	+	+	9	15.5
C3 229	Met.	Unmet.		B, C		+	-	4	6.9
C1 243	Met.	Unmet.		B		+	-	2	3.4
C4 264	Met.	Unmet.	B		C	+	+	6	10.3
C7 305	Met.	Unmet.			C	-	+	3	5.2
C6 323	Unmet.	Met.	B		A	+	+	2	3.4
C7 386	Met.	Unmet.		C		+	-	1	1.7
C5 454	Unmet.	Met.	C	A, B, C	C	+	+	8	13.8
No. changes								**58**	**100**

*Unmethylated (Unmet.) DNA site restricted by both HpaII and MspI (1, 1); Methylated (Met.) DNA site not restricted by HpaII and restricted by MspI (0, 1).

**Relation with a particular plant phenotype is indicated with (+) for presence and (-) for absence.

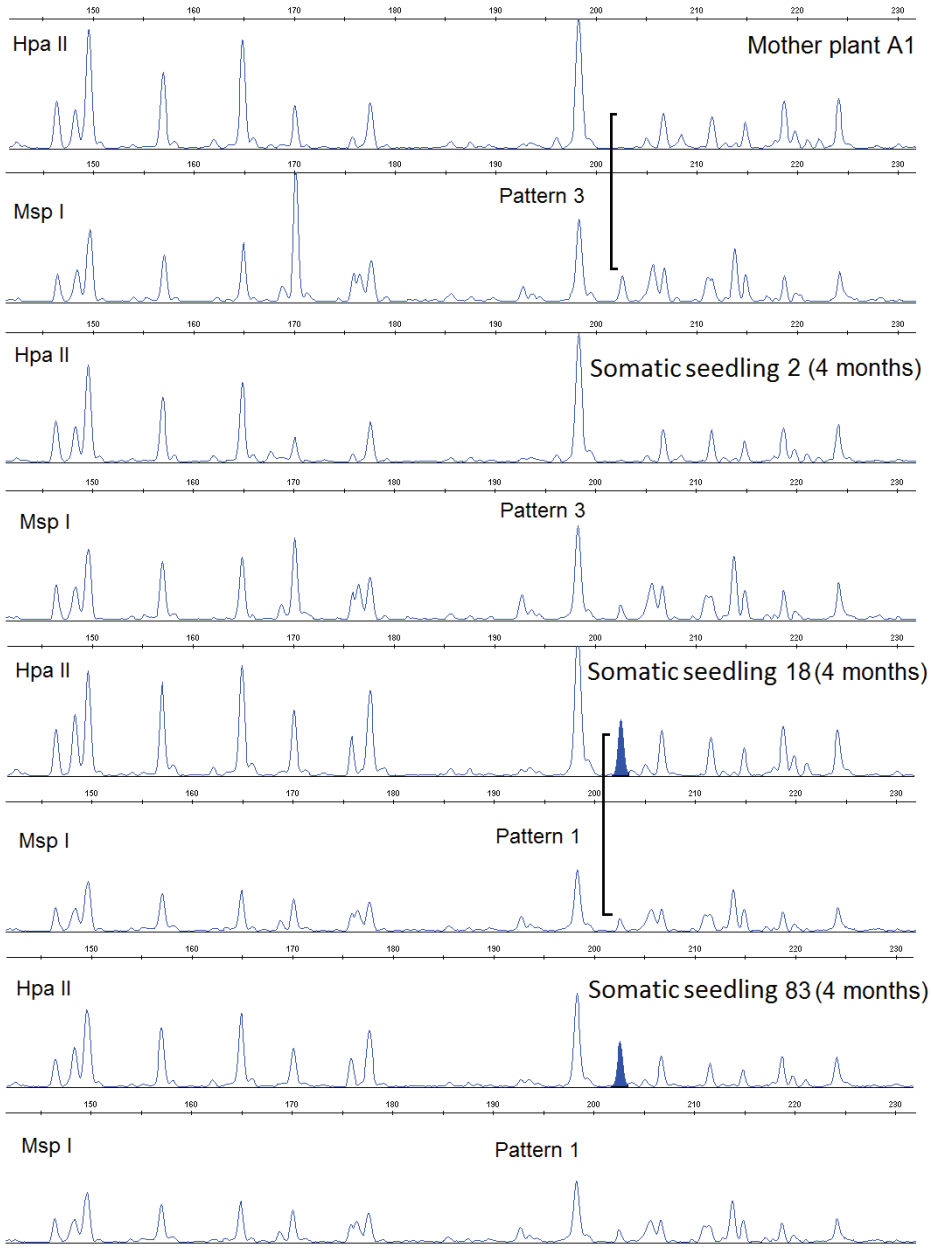


Figure 4. Examples of MSAP electropherograms observed for coffee mother plants and somatic seedlings progeny derived from 4 month cell cultures using the isoschizomers *Hpa*II and *Msp*I. Illustration of the variation patterns obtained for the phenotypically normal plants.

The number of methylation changes accumulated by somatic seedlings was between 1 and 4 (Fig. 5). We did not find any somatic seedling accumulating a great quantity of methylation changes or any accumulative effect related to culture age. The number of accumulated methylation changes was similar for somatic seedlings regenerated from all proliferation durations (4, 11, 27 months) (Fig. 5). Most of the methylation changes appeared and disappeared with no clear relation with a particular cell line or culture age. Two polymorphic molecular markers (C3 202 bp and C5 454 bp) were observed in some emblings derived from cellular line "C" and for the 3 culture ages analyzed (Table 8). However,

changes from these two fragments C3 were also shared with other plants derived from different cellular lines and without any relation with the culture age.

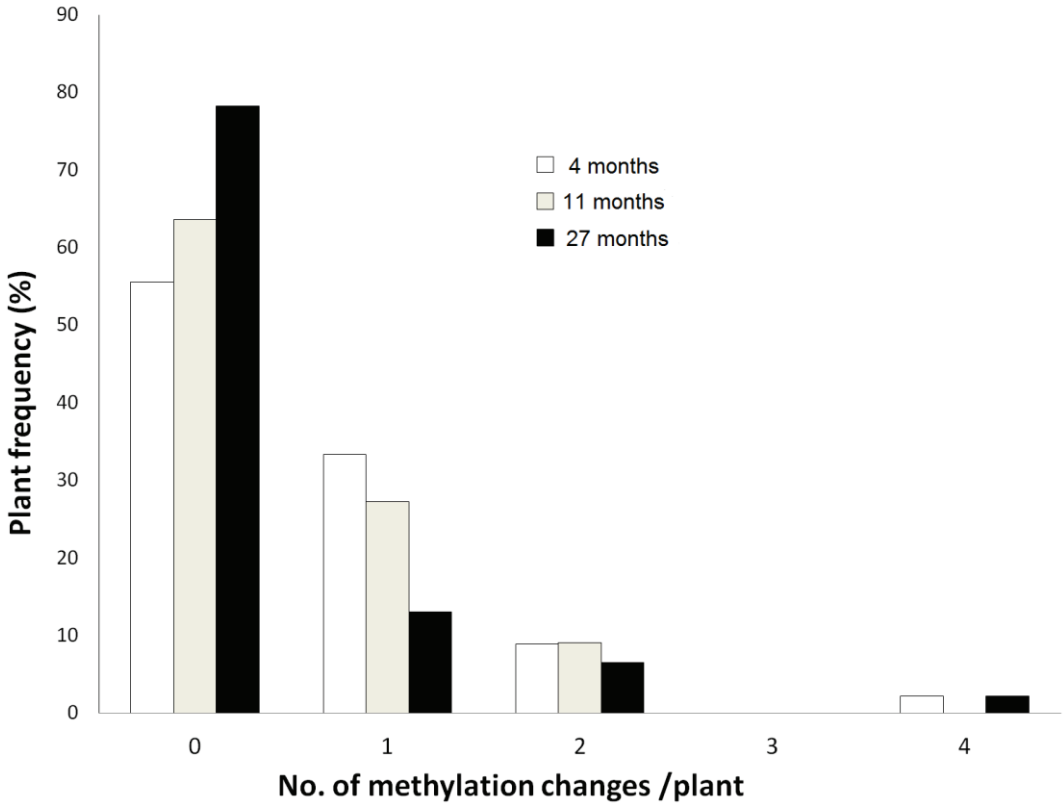


Figure 5. Accumulation of methylation changes in 124 coffee somatic seedlings derived from three different cell lines (A, B and C) depending on the time of culture (4, 11 and 27 months).

Abnormal chromosome numbers limited to phenotypic variants

Chromosome counts were performed on 15 somatic seedlings obtained from all culture age (Fig. 7, Table 9). All plants showing a normal phenotype systematically presented normal chromosome numbers disregarding the culture age. This behavior was observed even in the rare case of one phenotypically normal plant regenerated from a 27 month-old culture. Conversely, the chromosome counts performed in eight abnormal phenotypes plants derived from the two longer culture periods evidenced aneuploidy. All phenotypic variants presented monosomies i.e. the lack of one or several chromosomes. The Dwarf variants presented 2n-1, 2n-2 and 2n-3 monosomies while the three Bullata phenotypic variants showed 2n-1 monosomy (43 chromosomes).

Table 9. Summary of chromosome countings in some normal *versus* variant *C. arabica* somatic seedlings derived from cell cultures. The chromosome numbers obtained from root tips are indicated for the specific culture age of cell cultures (4, 11 and 27 months) from which somatic seedlings derived.

Cell culture age (months)	**Somatic seedling phenotype**	**Chromosome number**
4 C	Normal	44
4 A	Normal	44
11 B	Normal	44
11 C	Normal	44
11 B	Normal	44
11 B	Normal	44
27 B	Normal	44
11 B	Dwarf variant	**43**
27 A	Dwarf variant	**42**
27 A	Dwarf variant	**41**
27 A	Dwarf variant	**43**
27 A	Bullata variant	**43**
27 C	Bullata variant	**43**
27 B	Bullata variant	**43**
27 C	Dwarf variant	**42**

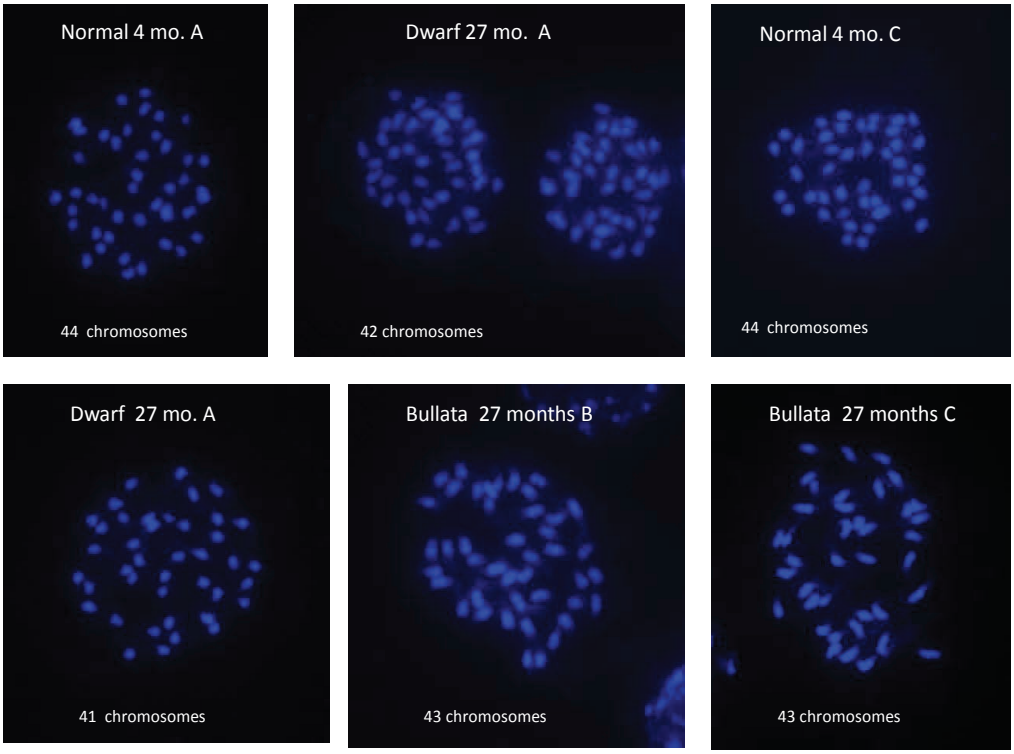


Figure 7. Mitotic cells at metaphase or prometaphase stages and observed ploidy levels of some normal and variant somatic seedlings from the allotetraploid *C. arabica* species (2n=4x=44). Plants were regenerated from 4 and 27 month old cell cultures. Karyotype analyses were performed by counting chromosomes on four to eight clear metaphase spreads obtained from root tips of 12 month-old nursery plants.

Discussion

Elevated levels of SV affecting phenotype have been frequently reported with long-term cell cultures (Nehra et al. 1992, Tremblay et al. 1999, Jambhale et al. 2001, Biswas et al. 2009). In the present work we showed that the occurrence of SV is closely related with cell's culture time of proliferation. The phenotypic evaluation of regenerated plants derived from 3 independent cell lines showed that most of them were variants (94%) when the regeneration was induced from two years' cultures, whereas no SV or intermediate levels (20-40%) were found with short and medium-term cultures, respectively. The use of different cell lines established for a same genotype allowed to strongly support the existence of a non random mechanism for SV. This is in accordance with previous observations on plants derived from cell suspensions of different *C. arabica* genotypes (Etienne and Bertrand 2003). Low levels of SV (1.3% in average) were found in plants regenerated from recently obtained callus or from 6 months old cell cultures. In that work, the SV rate reached

25% in batches of plants produced from suspensions aged 12 months. The frequency of variants increasing exponentially with the age of the cell culture.

Different types of molecular markers were performed in this study to explore a genetic or epigenetic origin of SV and evidence a possible accumulation of polymorphisms over time. AFLP markers did not show any polymorphism in the whole plant population analyzed indicating low levels of genetic variation. Similar approaches with long-term cultures using different kinds of molecular markers found contrasting results. Polymorphisms were not found using RAPDs markers in 2 years old-cultures of hop (Peredo et al. 2006). Using SSR markers Smýkal et al. (2007) reported no genetic variation in 24 years old in multiple shoots cultures of pea. However, in the same work AFLP molecular markers successfully evidenced 22 polymorphic and singleton fragments among 436 total fragments.

It has been long proposed the involvement of TEs on phenotypic affectation (Larkin and Scowcroft 1981). The most visible effect of TEs in somaclonal variants is related to color affectation. This alteration is produced by the transposition of DNA Class II transposon and was found in *in vitro* regenerants from different species. Some examples of color affectation are the variegation produced in maize kernels from indirect organogenesis (Planckaert and Walbot 1989, Peschke and Phillips 1991, Rhee et al. 2010), carrot leaves (Ozeki et al. 1997), the petals of the ornamental flower Saintpaulia from direct organogenesis (Sato et al. 2011) or the total color pigmentation from red to purple of potato tubers derived from indirect SE (Momose et al. 2010). Clear evidences of activity of *Tos 17* and *Tos 19* retrotransposons were provided in short and long-term cell cultures respectively (Hirochika 1996). *Tos 19* was found active in plants from 24 months cell cultures (Hirochika 1996). Other studies associated the activity of TEs (*Athila* element) in long-term cell suspensions cultures of *Arabidopsis* to specific hypomethylation in heterochromatic regions (Tanurdzic et al. 2008). Additionally, in cell cultures *Tnt1* tobacco retrotransposon could remain active after *in vitro* plant regeneration (Hirochika 1993). In coffee, no information was available about the potential role played by the activity of TEs on cell cultures. Extended cell proliferation periods did not allow evidencing TEs activity during the cell culture. These finding are in agreements with AFLPs results indicating no genetic variation. Some scenarios are possible to explain the lack of polymorphisms. Evidence has been found that some TEs (e.g. Mutator) are not activated or become hypermethylated and silenced after tissue culture (Planckaert and Walbot 1989). Transcriptional (methylation) and post-transcriptionally mechanisms (siRNA) could also repress the activity of TEs like Arabidopsis 'Évadé' retrotransposon and prevent their insertion (Mirouze et al. 2009).

The epigenetic modifications induced by tissue culture were increasingly studied during the last ten years. Many studies have reported high and unexpected levels of epigenetic variation related to cytosine methylation (Miguel and Marum 2011). In spite of this, the epigenetic affectation in long-term cultures and the derived plants has been poorly evaluated. Two approaches were followed to study the epigenetic changes. The first one was aimed at exploring specific sequence contexts with methylation-sensitive molecular markers (e.g. CCGG for MSAP) and the other at measuring the global methylation levels in all three cytosine contexts CG, CHG and CHH with HPLC methods. In our work, the number of methylation changes was low (i.e. from 15 to 27) and the resulting total polymorphism was less than 1% independently of the age of the cell culture from which they derived indicating little perturbation of the epigenetic status. Similarly, low levels of methylation variation were found in fresia somatic seedling derived from from short-term cultures (Gao et al. 2009), long term multiple shoots cultures of pea (Smýkal et al. 2007) and hop-plants axillary buds (Peredo et al. 2006). The involvement of methylation events on SV could be a factor overestimated. Many studies with methylation-sensitive molecular markers which reported high level of epigenetic variation have described null affectation on phenotype or morpho-agronomic traits (Bednarek et al. 2007; Li et al. 2007; Schellenbaum et al. 2008; Fiuk et al. 2010). The methylation changes contribute poorly to global estimations when they are not produced in high frequencies (Smulders and Klerk 2011). The individual accumulation of methylation changes could be a more trusted indicator of tissue culture stress or plant epigenetic instability. In the epigenetic evaluation made by molecular markers (e.g. MSAP) most of the *in vitro* plants generally do not tend to accumulate great quantities of methylation changes (Peraza-Echeverría et al. 2001, Li et al. 2007, Bednarek et al. 2007, Schellenbaum et al. 2008, Fiuk et al. 2010). However, it was noted in these works that a few plants can individually accumulate many methylation changes indicating a higher epigenetic instability. Coffee somatic seedlings from short and long-term cell cultures accumulated comparable and weak numbers of methylation changes (1-4 by individual) although in the last situation most of them were variants. The same tendency was previously observed in both normal and variant coffee somatic seedlings obtained in two industrially applied SE processes using embryogenic suspensions and secondary embryogenesis proliferation systems (Bobadilla-Landey et al. 2013). These findings indicate that the epigenetic status of plants derived from medium and long-term cultures remained stable.

First studies in tissue-culture derived plants identified many chromosomal abnormalities as one of the main factors involved in SV (D'Amato 1985, Kaeppler et al. 2000). Many altered mechanisms observed during meiosis and mitosis can produce chromosomal alterations, such as non-disjunction (i.e. the failure of sister chromatids to separate during meiosis or mitosis), multipolar mitotic spindles, lagging anaphase chromosomes and chromosome missegregation into micronuclei (D'Amato 1985, Kaeppler et al. 2000, Holland and Cleveland 2009). Such mechanisms can produce different extent

and severity of chromosomal abnormalities in *in vitro* plants including segmental or complete aneuploidy (Nehra et al. 1992, Kaepler and Phillips 1993, Tremblay et al. 1999, Jambhale et al. 2001). Aneuploidy was frequently reported in tissue culture during the phases of callus induction and cell suspension cultures (Gözükirmizi et al. 1990, Hao and Deng 2002, Kumar and Mathur 2004, Giorgetti et al. 2011). Furthermore, chromosomal instabilities in callus cultures can be enhanced when prolonged culture times are applied (Henry et al. 1996, Hao and Deng 2002). Henry et al. (1996) reported that the levels of chromosomal abnormalities obtained from 4 months callus cultures of wheat were 12% whereas they reached approximately 80% in 14 months cultures. Such anomalies included both polyploidy and aneuploidy. Chromosomal abnormalities were also reported in *in vitro* regenerants derived from long-term cultures in a variety of species (Nehra et al. 1992, Tremblay et al. 1999, Hao and Deng 2002, Jambhale et al. 2001).

The cytological examination of coffee somatic seedlings derived from cell cultures revealed different types of aneuploidy exclusively in the monosomic form. Such abnormalities were systematically found in the analyzed somaclonal variants derived from 11 and 27 months old cultures. On the contrary normal ploidy levels were always observed in plants showing normal phenotype. Hence, most of somaclonal variants derived from the two longest proliferation times could presumably present some degree of aneuploidy (i.e. segmental or total). In coffee, many cytological abnormalities were described in leaf explants and cell cultures by Menéndez-Yuffá et al. (2000) like non disjunction, double prophases cells, lagging chromosomes, micronuclei and binucleated cells, also including polyploid and aneuploid cells. Our results showed that cytological abnormalities could be enhanced by tissue culture specially in long-term cultures. Aneuploidy plays a major role in *C. arabica* SV. In this regard, the mechanisms of non disjunction of chromosomes during anaphase could be the best explanation for the production of aneuploid cells and plants during tissue culture (Holland and Cleveland 2009). Despite of this, neither non disjunction nor other mechanisms could explain the predominance of monosomic plants. Phenotypic variation is observed in many forms in tissue culture-derived plants. The *in vitro* plants usually present abnormalities in plant size, leaf shape, plant architecture and reproductive structures. The gene imbalance associated to specific altered chromosomes results in well defined and sometimes predictable variant phenotypes (Birchler and Veitia 2007, Makarevitch et al. 2008, Henry et al. 2010, Makarevitch and Harris 2010). The presence of dwarf plants was frequently reported in somaclonal variants of many species including coffee (Tremblay et al. 1999, Etienne and Bertrand 2003, Sajihram et al. 2003, Bairu et al. 2006). The association of a predominant phenotype with a particular culture age was already observed in coffee (Etienne and Bertrand 2003).

The elevated rates of SV frequently reported with long-term cell cultures offer a unique opportunity to understand the underlying molecular and physiological mechanisms. The use of long-term cultures was aimed at establishing the full range of cytological, genetic and epigenetic mechanisms leading to SV. Although the molecular markers approach used cannot generate exhaustive information, it can be concluded that genetic polymorphisms and epigenetic changes are particularly limited during the cell culture and that aneuploidy is directly linked to SV. *C. arabica* is a young allopolyploid still having the most of its genes in duplicated copies (Cenci et al. 2012). It could be hypothesized that the impact of genetic or epigenetic polymorphisms was restricted because of the buffer effect due to polyploidy. However the allopolyploid structure also allowed aneuploid cells to survive and regenerate viable plants what would not certainly be the case with non polyploid species. The absence of one or more chromosome was probably buffered by the other homeologous chromosomes. Hence, *C. arabica* is an interesting species to study the occurrence and consequences of chromosomal abnormalities in relation with SV. Although further studies are necessary for an accurate understanding of the chromosome anomalies involved in the acquisition of a particular phenotype, the present results obviously showed that mitotic aberrations play a major role in SV in coffee.

References

Holland AJ, Cleveland DW (2009) Boveri revisited: Chromosomal instability, aneuploidy and tumorigenesis. Nat Rev Mol Cell Biol. 2009 July; 10(7): 478–487. doi: 10.1038/nrm2718.

Anjanasree K, Neelakandan, Kan Wang (2012) Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. *Plant Cell Rep*. 31: 597– 620.

Bairu MW, Fennell CW, Van Staden J (2006) The effect of plant growth regulators on somaclonal variation in. Cavendish banana (Musa AAA cv. 'Zelig'). SCI. HORT. 108(4): 347-351.

Barret P, Brinkman M, Beckert M (2006) A sequence related to rice Pong transposable element displays transcriptional activation by *in vitro* culture and reveals somaclonal variations in maize. Genome 49:1399–1407

Bednarek PT, Orłowska R, Koebner RMD, Zimny J (2007) Quantification of tissue-culture induced variation in barley (*Hordeum vulgare* L.). *BMC Plant Biol*, 7:10.

Biswas MK, Dutt M, Roy UK, Islam R, Hossain M (2009) Development and evaluation of *in vitro* somaclonal variation in strawberry for improved horticultural traits. *Scientia Horticulturae*, 122 (3): 409- 416.

Birchler JA, Veitia RA (2007) The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell 19: 395–402.

Cenci A, Combes M-C, Lashermes P (2012) Genome evolution in diploid and tetraploid *Coffea* species as revealed by comparative analysis of orthologous genome segments. Plant Mol Biol 78: 135-145.

D'Amato, F. (1985) Cytogenetics of plant cells and tissue cultures and their regenerates. *Plant Sci. 3*: 73-112.

DeBarry JD, Liu R, Bennetzen JL (2008) Discovery and assembly of repeat family pseudomolecules from sparse genomics sequence data using the assisted automated assembler of repeat families (AAARF) algorithm. BMC Bioinformatics. 13(9):235. doi:10.1186/1471-2105-9-235.

Díaz-Martínez M, Nava-Cedillo A, Guzmán-López JA, Escobar-Guzmán R, Simpson J (2012) Polymorphism and methylation patterns in *Agave tequilana* Weber var. 'Azul' plants propagated asexually by three different methods. *Plant Sci*, 185-186:321-330.

Duncan RR (1997) Tissue culture-induced variation and crop improvement. *Adv Agron* 58:201-240.

Etienne H, Bertrand B (2001) The effect of the embryogenic cell suspension micropropagation technique on the trueness to type, field performance, bean biochemical content and cup quality of *Coffea arabica* trees. Tree Physiol 21: 1031-1038.

Etienne H, Bertrand B (2003) Somaclonal variation in *Coffea arabica*: effects of genotype and embryogenic cell suspension age on frequency and phenotype of variants. *Tree Physiol*, 23:419-426.

Etienne H (2005) Protocol of somatic embryogenesis: *Coffee* (*Coffea arabica* L. and *C. canephora* P.). In Protocols for somatic embryogenesis in woody plants (the Netherlands, Jain SM and Gupta PK), pp. 167-179.

Fiuk A, Bednarek PT, Rybczýnski J J (2010) Flow cytometry, HPLC-RP and met AFLP analyses to assess genetic variability in somatic embryo-derived plantlets of *Gentiana pannonica* scop. *Plant Mol Biol Rep*, 28:413-420.

Furner IJ, Matzke M (2010) Methylation and demethylation of the *Arabidopsis* genome. Curr Opin Plant Biol 14:1-5.

Gao X, Yang D, Cao D, Ao M, Sui X, Wang Q, et al. (2010) *In vitro* micropropagation of *Freesia hybrida* and the assessment of genetic and epigenetic stability in regenerated plantlets. *J Plant Growth Regul*, 29:257–267.

Giorgetti L, Castiglione M, Turrini A, Ronchi V, Geri C (2011) Cytogenetic and histological approach for early detection of "mantled" somaclonal variants of oil palm regenerated by somatic embryogenesis: first results on the characterization of regeneration system. *Caryologia* 64: 223-234.

Gözükirmizi N, An S, Oraler G, Okatan Y, Onsal N (1990) Callus induction, plant regeneration and chromosomal variations in barley. *Acta Bot. Neerl*. 39(4): 379-387.

Grandbastien MA *(*1998*).* Activation of plant retrotransposons under stress conditions. *Trends Plant Sci.* 3*,* 181–187.

Hao Y I, Deng XX (2002) Occurrence of chromosomal variations and plant regeneration from long-term-cultured *citrus* callus. *In vitro Cellular and Developmental Biology-Plant* 38: 472-476.

Henry IM, Dilkes BP, Miller ES, Burkart-Waco D, Comai L (2010) Phenotypic consequences of aneuploidy in *Arabidopsis thaliana*. *Genetics*, 186:1231-45.

Henry Y, Marcotte JL, de Buyser J (1996) The effects of aneuploidy on karyotype abnormalities in wheat plants regenerated from short- and long-term somatic embryogenesis. *Plant Sci*. 114: 101–109.

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. EMBO J 12:2521–2528

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M (1996). Retrotransposons of rice involved in mutations induced by tissue culture. Proc. Natl. Acad. Sci. USA 93, 7783–7788.

Hirochika H, Okamoto H, Kakutani T. (2000). Silencing of retrotransposons in Arabidopsis and reactivation by the ddm1 mutation. *Plant Cell*. 12: 357-369.

Jain SM (2001) Tissue culture-derived variation in crop improvement. *Euphytica* 118: 153–166.

Jambhale, N.D., S.C. Patil, A.S. Jadhav, S.V. Pawar and B.D. Waghmode (2001) Effect of number of subcultures on In vitro multiplication of four banana clones. Infomusa, 10: 38-39.

Jiang N, Bao ZR, Zhang XY, Hirochika H, Eddy SR, McCouch SR et al. (2003). An active DNA transposon family in rice. Nature 421 , 163–167.

Jiang C, Mithani A, Gan X, Belfield EJ, Harberd P et al. (2011) Regenerant Arabidopsis lineages display a distinct genome-wide spectrum of mutations conferring variant phenotypes. Curr Biol 21:1385–1390.

Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J (2000) Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110, 462–467.

Kaeppler SM, Phillips RL (1993). DNA methylation and tissue culture-induced DNA methylation variation in plants. *In Vitro Cell Dev Bio-Plant*, 29:125-130.

Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 43: 179-188.

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci U S A. 1996;93(22):12406.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000). Genome evolution of wild barley (*Hordeumspontaneum*) by BARE-1 retrotransposon dynamics in response to Sharp microclimatic divergence. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6603–6607.

Karp A, (1991) Origins, causes and uses of variation in plant tissue cultures. In *Plant cell and tissue culture*. Edited by: Vasil IK, Thorpe TA. Dordrecht: Kluwer Academic Publishers; 1994:139-152.

Kumar PS, Mathur VL. (2004). Chromosomal instability in callus culture of Pisum sativum. *Plant Cell Tissue Organ Cult* 78:267–271.

Kumar A, Bennetzen JL (1999) Plant retrotransposons. *Annu. Rev. Genet*. 33:479–532.

Larkin PJ, Scowcroft WR (1981). Somaclonal variation: a novel source of variability from cell cultures for plant improvement. *Theor Appl Genet*, 60:197–214.

Law J, Jacobsen SE (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*, 11: 204-220.

Lee ML, Phillips RL (1988) The chromosomal basis of somaclonal variation. *Annu. Rev. Plant Physiol. Plant Mol.* Biol. 39: 413–437.

Li X, Yu X, Wang N, Feng Q, Liu B, et al. (2007) Genetic and epigenetic instabilities induced by tissue culture in wild barley (*Hordeum brevisubulatum* (Trin.) Link). *Plant Cell Tiss Org Cult*, 90:153–168.

Makarevitch I, Phillips RL, Springer NM (2008) Profiling expression changes caused by segmental aneuploidy in maize. *BMC Genomics* 9: 7.

Makarevitch I, Harris C (2010) Aneuploidy causes tissue specific qualitative changes in global gene expression patterns in maize. *Plant Physiol*, 152:927–938.

Menéndez-Yuffá A, Fernandez R, Rios L, Xena N, (2000) Mitotic aberrations in coffee (*Coffeea arabica* cv. 'Catimor') leaf explants and their derived embryogenic calli. *Electron. J. Biotechnol.* 3, 1–6.

Miguel C, Marum L (2011) An epigenetic view of plant cells cultured *in vitro*: somaclonal variation and beyond. J Exp Bot 62: 3713-3725.

Mirouze M, Reinders J, Bucher E, Nishimura T, Mathieu O, et al. (2009) Selective epigenetic control of retro transposition in Arabidopsis. *Nature*, 461:427-430.

Miura A, Yonebayashi S, Watanabe K, Toyama T, Kakutan T, et al. (2001) Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. *Nature* 411:212–214.

Momose M, Abe Y, Ozeki Y (2010) Miniature inverted-repeat transposable elements of Stowaway are active in potato. Genetics 186:59–66.

Nehra NS, Kartha KK, Stushnoff C, Giles KL (1992) The influence of plant growth regulator concentrations and callus age on somaclonal variation in callus culture regenerants of strawberry *Plant Cell, Tissue and Organ Culture* 29: 257-268.

Ozeki Y, Davies E, Takeda J (1997) Somatic variation during long term subculturing of plant cells caused by insertion of a transposable element in a phenylalanine ammonia-lyase (PAL) gene. Mol Gen Genet 254:407–416.

Peraza-Echeverria S, Herrera-Valencia VA, Kay A-J (2001) Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplification polymorphism (MSAP). Plant Sci 161:359–367.

Peredo EL, Angeles Revilla M, Arroyo-Garcia R. 2006. Assessment of genetic and epigenetic variation in hop plants regenerated from sequential subcultures of organogenic calli. Journal of Plant Physiology 163, 1071–1079.

Peschke VM, Phillips RL (1991) Activation of the maize transposable element Suppressor-mutator (Spm) in tissue culture. Theor Appl Genet 81:90–97

Planckaert F, Walbot V, (1989) Molecular and genetic characterization of *Mu* transposable elements in *Zea mays*: behavior in callus culture and regenerated plants. *Genetics,* 123, 567–578.

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

Rhee Y, Lin H, Buell R, Childs K, Kaeppler S (2009) Ac2 allele of maize identified in regenerant-derived progeny from tissue culture results from insertion of a novel transposon. Maydica 54:429–437.

Rodríguez López, C. M., Wetten, A. C. and Wilkinson, M. J. (2010), *Progressive erosion of genetic* and epigenetic variation in callus-derived New Phytol. Jun;186(4):856-68. doi: 10.1111/j.1469-8137.2010.03242.x.

Sato M, Hosokawa M, Doi M (2011) Somaclonal Variation Is Induced De Novo via the Tissue Culture Process: A Study Quantifying Mutated Cells in Saintpaulia. *PLoS ONE* 6(8): e23541. doi:10.1371/journal. Pone. 0023541.

Schellenbaum P, Mohler V Wenzel G, Walter B (2008) Variation in DNA methylation patterns of grapevine somaclones (*Vitis vinifera* L.). *BMC Plant Biol*, 8:78.

Skirvin RM, McPheeters KD, Norton M (1994) Sources and frequency of somaclonal variation. *HortScience* 29:1232-1237.

Smith AM, Hansey CN, Kaeppler SM (2012) TCUP a novel *h*AT transposon active in maize tissue culture. Frontier in Plant Science, 3: 6.

Smulders MJM, de Klerk GJ (2011) Epigenetics in plant tissue culture. *Plant Growth Reg*, 63:137-146.

Smýkal P, Valledor L, Rodriguez 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.

Söndahl, M.R. and A. Bragin. 1991. Somaclonal variation as a breeding tool for coffee improvement. In Proc. 14th Colloquium International Coffee Science Association, San Francisco, USA. ASIC Publishers, Vevey, Switzerland, pp 701–710.

Sonnhammer EL, Durbin R (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene. 167(3):GC1–G10.

Stroud H, Ding B, Simon SA, Feng S, Bellizzi M, Pellegrini M, Wang GL, Meyers BC, Jacobsen SE (2012) Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife*. 2013;2:e00354. doi: 10.7554/eLife.00354. Epub 2013 Mar 19.

Tanurdzic M, Vaughn MW, Jiang H, Tae-Jin L, Martienssen RA, et al. (2008) Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biol* 6, 2880-2895.

Tremblay L, Levasseur C, Tremblay FM (1999) Frequency of somaclonal variation in black spruce (*Picea mariana,* Pinaceae) and White spruce (*P. glauca,* Pinaceae) derived from somatic embryogenesis and identification of some factors involved in genetic instability. *Am J Bot*, 86:1373-1381.

Vanyushin BF, Ashapkin VV (2011) DNA methylation in higher plants: Past, present and future. Biochim Biophys Acta 1809: 360-368.

Vos P, Hogers R, Bleeker M, Rijans M, Zabeau M et al. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23:4407-4414.

Waugh R, McLean K, Flavell A J, Pearce SR, Powell W, et al. (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence- specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet*. 253, 687-694.

Zhang M, Kimatu JN, Xu K, Liu B (2010) DNA cytosine methylation in plant development. Review. *J Genet Genomics*, 37:1-12.

CHAPTER IV

GENERAL DISCUSSION AND PERSPECTIVES

Somatic embryogenesis is an efficient and promising technology that has been long-expected for large-scale plant micropropagation since the end of the seventies. Unfortunately, it was quickly perceived the complexity to develop this technique. In this sense, SE involves a lot of developmental steps and a wide range of poorly understood mechanisms complicating the establishment of the regeneration pathway. The first steps of SE involve the mechanism of differentiation/redifferentiation leading to the production of cells having in theory individually the potential to regenerate complete plants. At this matter, the production of the embryogenic tissues presents the first obstacle to overcome, and is even more difficult to achieve starting with aged or mature selected plants. For example, the recalcitrance to produce embryogenic tissues is a phenomenon commonly observed in woody plants (e.g. forest trees). A rejuvenation step is often a prerequisite to solve this problem. Embryo production is also difficult as it depends on the optimization of many factors (i.e. genotype, PGRs, physical environment among others). The conversion of somatic embryos into plantlets, step of crucial importance for the cost effectiveness of the whole process, is very difficult to optimize. The embryo/plant conversion rate intrinsically depends on the quality of embryos obtained and hence on the efficiency of the previous maturation/desiccation steps. Another difficulty frequently reported during the last decades is the production of plants genetically uniform to the mother plant. The obtention of genetically uniform plants i.e. true-to-type micropropagation implies a rigorous optimization of culture conditions such as i.e. type and concentration of growth regulators, number of regeneration cycles/subcultures, type of initial explant and genotype used among others. Additionally, once plant regeneration is achieved a very detailed evaluation of the morpho-agronomic characteristics is needed. However, the incidence of SV is sometimes difficult to measure, particularly in plants with long developmental cycles. That is the case of conifers because the agronomic trait evaluation cannot be made before 20 years of field development. This greatly delays and complicates the experiments performed for test the genetic conformity in function of variable *in vitro* culture conditions. Although an accurate and complete agronomic evaluation is not an easy task, it is even more difficult to associate biological mechanisms or candidate genes to a particular variant phenotype. The last barrier limiting the extensive use of SE processes at industrial scale is to succeed in reaching their economic profitability. All SE systems developed frequently compete directly with other multiplication systems based on vegetative and/or sexual horticultural propagation techniques than can represent simpler and cheaper ways to propagate the elite varieties.

This thesis work contributed in generating responses and information on the trueness-to-type of the SE processes established for coffee (*Coffea arabica* L.), and particularly that of a system based on the use of embryogenic cell suspensions. The approach used was also aimed at better understanding the mechanisms leading to SV in coffee. More precisely, the main expectations of the present thesis were: 1) to know if it's possible to apply at industrial level a SE process based on embryogenic

suspensions to obtain high levels of genetically uniform plants. The deployment of such protocols in already existing coffee micropropagation units directly depends on the responses generated by this work. During the last years, the practical experience acquired in these units showed that embryogenic suspension protocols appeared as the more adapted for the large scale and cost effective production of coffee elite cultivars; 2) To better identify the culture conditions responsible for SV, specifically the time of proliferation and growth regulators known to be among the most critical factors. 3) To evaluate the levels of cytological, genetic and epigenetic changes after a complete cycle of SE but for different technical itineraries, and if possible to relate them to a well characterized phenotypic variation.

Two complementary approaches to uncover the impact and origin of SV

In a first approach we evaluated the plant regenerants produced in two types of SE processes: a first one *a priori* considered as low mutagenic (short proliferation times, low auxin concentration but enough efficient to ensure the production of sufficient quantity of embryogenic material). The second approach was more adapted to research purposes and based on previous experiences of the team (Etienne and Bertrand 2001, 2003) using a much more mutagenic process with long periods of proliferation and high auxin and cytokinin concentrations. The work has been planned to study at the same time genetic, epigenetic and cytological mechanisms rather than more deeply focus on one particular current tendency mechanisms like epigenetic or the activation of TEs. The reason of this choice was the total lack of knowledge about the mechanisms leading to SV in coffee. Hence, it seemed more efficient to explore simultaneously some of the candidate mechanisms in view of the quick identification of the responsible(s) one(s). It is evident that each mechanism *per se* could require several years of research. Another objective in the two approaches was to relate any identified cytological, genetic or epigenetic change to the observed phenotypic variation. In many papers, although significative and clearly consecutive to *in vitro* propagation, the changes observed could not be related to SV. For this reason, an important effort has been done on phenotypic characterization and the abnormal phenotypes identified in the SE progenies were systematically compared to both mother plants and phenotypically normal micropropagated plants in all kinds of analysis. Additionally, whatever the species considered, to our knowledge no evaluation for genetic conformity has been made using industrially applied SE protocols.

A micropropagation system ensuring trueness-to-type

We have shown that is possible to limit the occurrence of SV by optimizing culture conditions. Micropropagation using cell suspensions is now well mastered for *C. arabica* which belongs to the 3 or 4 plant species for which such level of liability has been achieved. In SE the use of cell suspensions offers great advantages in terms of work simplification, rate of proliferation, reduction of production

costs, synchronization of embryo development and the possibility of long-term storage in liquid nitrogen (-196°C). Thanks to this technology it will be possible to produce 5 to 10 million vitroplants per production unit and by this way to disseminate quickly and massively the new varieties. Furthermore, with the specific objective of ensuring true-to-type propagation by using secure SE protocols, the team developed several strategies during the last 7 years to limit the production and the diffusion to producers of somaclonal variants (Etienne et al. 2012). The strategy named "risk dilution" is based on several rules which are presented next:

1. To maintain a low auxin concentration during the embryogenic suspension proliferation step,
2. To limit the number of proliferation cycles during the embryogenic suspension step (max. 6).
3. To incorporate a total traceability system per production lot and cellular line to identify better the affected plant batches, understand the origin of the problem and quickly minimize the risk of SV.
4. To apply a phenotypic screening to detect somaclonal variants in nursery with the goal of minimizing the risk of delivering them to producers. For this reason, we can see the importance that the control of nursery be directly done under the strict control of the micropropagation company which will be always responsible for the presence of undesired plants at field level.

These simple measures are based on common sense and made possible the true-to-type multiplication of *C. arabica* hybrids.

In coffee, SE technology is directly competing with seed propagation to distribute cultivars. Although abnormal phenotypes arise from both propagation methodologies (Krug *et al.* 1939; Krug and Carvalho 1951), it is clear that the rates of variation arising from *in vitro* methodologies must stay in acceptable rates, approximately less than 5%. In the future, with the continuous and necessary scaling-up of the process, it will be of primary importance of controlling continuously the SV rates. Further optimizations of culture conditions with the goal of improving the *cost-effectiveness* along with the quantity of plants produced are susceptible to enhance their occurrence. Nevertheless, both objectives could be compatible and be achieved either by intervening on the two callogenesis steps: 1) the initial dedifferentiation/redifferentiation step in the leaf explants leading to the production of embryogenic tissues (first month of the process), 2) the proliferation of embryogenic suspensions that permits to drastically increase the *availability* of embryogenic material, hence the number of regenerated plants and productivity of the whole SE process.

It has been previously shown in coffee that the initial step of embryogenic callus production systematically produced a low percentage of SV, estimated to be around 1% in the conditions used by Etienne and Bertrand (2003) that were identical to those used in the thesis work. Then, it seems that the first mutational events can undergo very soon in the process and that the embryogenic suspension step can increase their frequency depending on the culture conditions applied. It could be useful to lower the SV occurrence during the embryogenic callus production by decreasing the hormonal supply

(i.e. actually a mix of IAA, IBA, 2, 4-D, 6-BA) which is possibly too high. It seems possible to reduce it without decreasing the quantity or the quality of the embryogenic tissues produced. To our knowledge, the hormonal balance during this initial step has never been optimized to reduce the impact of SV.

The thesis experiment design did not allow identifying the respective responsibilities of growth regulators and culture duration in the production of SV during the proliferation of embryogenic cultures (see chapter 3). As previously discussed, the establishment of embryogenic suspensions is difficult because it needs at the beginning an accurate selection of the embryogenic tissues (different types of tissues are simultaneously regenerated on the leaf explants) and monthly subcultures also requiring an important technical know-how to avoid stressing the suspensions. Nevertheless at industrial scale, suspension cultures are eliminated after 6 months of proliferation to prevent the apparition of SV. Due to both economical and practical reasons it would be interesting to extend the period of proliferation of already well established and highly embryogenic suspension cultures. The most logical way to do this is to extend the proliferation duration maintaining the 2, 4-D and 6-BA supply at a low level and testing the batches of plants regenerated from older cultures for SV. A complementary objective is to establish a cytological technique of chromosome counting for the early detection of aneuploidy in cell suspension cultures. This way, it would be not necessary to induce plant regeneration and wait approximately 18 months for well developed nursery plants to perform chromosome counting on root tips or stems meristems. This would represent an efficient tool to early check for chromosomal aberrations directly on the embryogenic cultures.

Effectiveness of molecular markers to evidence changes at genetic level

DNA molecular markers are commonly applied to measure the impact of PTC in micropropagated plant genome and to elucidate the mechanisms of genetic variation (Miguel and Marum 2010). Polymorphisms are caused by the gain or loss of a restriction site caused by single base insertion/deletions a.k.a InDels, a length polymorphism between the restriction sites (e.g. transposable elements insertions/deletions or a variable microsatellite) or a change in the selective primer binding site like Single Nucleotide Polymorphisms or SNPs (Meudt and Clarke 2007), also the change of the methylation status of a specific sequence (CCGG differentially digested by Hpa II or MspI in MSAP). In our studies, very low levels of polymorphisms were found in somatic seedlings with AFLP markers in both normal and abnormal phenotypes, i.e. 0–0.003% Total Polymorphism for plants derived from secondary embryogenesis or embryogenic suspensions. However other forms of genetic variation could be involved in coffee SV and are not easily detectable through AFLP analysis. Genetic variation could be produced as single nucleotide polymorphisms in DNA, in molecular biology these changes are known as single nucleotide polymorphisms (SNPs). SNPs arise as mutations involving base pair

substitutions, from errors in DNA replication and from mutagenic agents (Jones et al. 2009). Commonly studied with Microarray and Next Generation Sequencing (NGS) technologies, SNPs are part of the third generation of molecular markers. In recent studies, variant phenotypes of Arabidopsis derived from SE were positively correlated with SNPs and InDels (Jiang *et al.* 2011). Additionally, some of the abnormalities reported in Arabidopsis by these authors presented different levels of segregation, such as small size plants and late flowering, indicating stable mutations. For this reason, the development and use of SNPs molecular markers represents the next step for the evaluation of genetic variation associated to SV in coffee somatic seedlings. This perspective is complementary to the study of genetic mechanisms done in this work. Furthermore, a bioinformatics tool especially designed for the high throughput detection of SNPs in plant genomes and named SNiPloid (http://sniplay.cirad.fr/cgi-bin/sniploid.cgi) was developed in GALAXY by a member of our team. It was adapted to facilitate the study of these variations in polyploid plants like *C. arabica* (Dereeper et al. 2011).

Compared efficiencies of different molecular marker methods to uncover the TEs transposition

In this thesis a molecular marker analysis was performed to evidence a possible relationship between the time of culture in a SE procedure and the activation of different TEs. The plant material for this analysis was regenerated in stressful and mutagenic conditions differing from the industrial ones to increase the probability of TEs activation. Our experiment design took into account several aspects: (1) the use of cloned plants allowed discarding polymorphism related to genotypic diversity. (2) The choice of TEs elements with intact long tandems repeats (LTRs) (100% identity) implying possible transcriptional activity. (3) The use of different TEs including two abundant families in genomes (i.e. Copia, Gypsy) for which several bibliographical references reported their possible activation through tissue culture (see table I.10 from literature review). Finally, to uncover the presence of new copies of transposable elements we choose SSAP technique which has successfully evidenced TEs insertions in different species like tobacco, sweet potato and lotus (Melayah et al. 2001, Tahara et al. 2004, Fukai et al. 2010). SSAP protocols were not available in coffee at the beginning of this work and were successfully adapted to this species. After a delicate standardization step of SSAP technique we proceed to perform the analysis. This analysis consisted in a total of 45 primer combinations using 15 different primers (i.e. selected among 35 designed) applied in a total population of 94 plants. Despite of the precautions taken (i.e. candidate element approach and the relatively important scale of the work), we did not find any active TEs for the entire set of element tested in plant derived from short and long-term cell cultures. In conclusion we did not evidence any polymorphism associated to new TEs copies. The first hypothesis is that these elements were not reactivated under stressful SE conditions or are inactive. Although TEs constitute a large percentage of

plant genome - for example 85% in maize and 90% in wheat (Schnable et al. 2009, Devos et al. 2005), only few elements were reported active in the literature.

A second possible explanation is that the inactivity of the chosen elements resulted from their transcriptional inactivation through methylation or epigenetic post-transcriptional mechanisms preventing their insertion. The TEs activity is known to be transcriptionally silenced by CG methylation supported by RNA-directed DNA methylation and post-transcriptionally silenced by siRNA (Mirouze et al. 2009).

Further studies aimed at studying TEs activation in coffee micropropagation can rely on different or new methodologies. PCR techniques could present several limitations related to electrophoresis (lack of reproducibility, technical artifacts, low resolution profiles) and primer design (unspecific amplification or null amplification). Even in the hypothetical case of finding TEs insertion in the form of polymorphisms, all of them need to be validated before declaring the activity of a given element. For example, in the study of wild barley (*Hordeum brevisubulatum*) micro shoots made by Li et al. (2007), all SSAP polymorphisms found in rice vitro plants were related to a phenomenon different from transposition mechanisms and identified as changes at the restriction sites or the selective bases. There are several alternatives to study the activity of TEs without performing PCR analysis. One alternative is the use of hybridization based techniques (e.g. southern blot, northern blot and microarrays). Most of the current knowledge about TEs activity was obtained using hybridization-based techniques like Southern and Northern blot (see Table I.10 from literature review). Although technical specialization is needed to perform such techniques, they present the advantage to be less susceptible to the outcome of technical artifacts and are equally or more effective than molecular markers. For example, Restriction Fragment Length Polymorphism (RFLP) is in fact a Southern blot used as molecular marker. Southern blot/RFLPs have been efficiently used for the detection of new inserted TEs copies in the genome of plants propagated by indirect organogenesis in crop species like rice (Hirochika 1993, 1996; Komatsu et al. 2003; Jiang et al. 2003) and maize (Planckaert and Walbot 1989; Peschke and Phillips 1990; Smith et al. 2012). However, hybridization techniques using Southern or Northern blot principles are not easy and require many steps for their realization and high degree of technical specialization. Also, hybridization based techniques like Southern blot often imply the use of radioactive marked probes to obtain better resolution in electrophoretic profiles.

Most of studies on TEs activation mechanisms in tissue culture derived plants include TEs expression analysis and structural genetics (Hirochika 1993; Takeda et al. 1999). Recent advances in coffee sequencing allow the design of sequence specific probes for high-throughput hybridization techniques using fluorescent marks. That is the case of microarrays (a.k.a. CHIPs) capable of simultaneously analyze thousands of genes in a simple CHIP and particularly useful for expression

studies in plants (Adomas et al. 2008). Recently, the drop in the price of most of the sequencing technologies has opened the possibility to perform transcriptional analyses and whole genome DNA sequencing, both using NGS (Next Generation Sequencing) technologies like 454 Sequencing, Illumina or SOLiD. Through these approaches the transcriptional activity or the structural changes caused by TEs transposition can be deeply study at different levels of genome (i.e. shallow or deep sequencing) leading to exhaustive information about the effects on coffee *in vitro* culture.

Transposable elements and their role in SV, an overstatement?

The role of TEs in generating genetic and phenotypic diversity has been largely demonstrated in seedlings. Some examples are the sexual determination in Melon (Martin et al. 2009), the elongated form in tomato (Xiao et al. 2008), the wrinkled seed malformation in Pea (Bhattacharyya et al. 1990) or the skin change color of grapes (Kobayashi et al. 2004). However, this is not the case for tissue culture derived plants in which the only phenotypic effect of TEs was related to color variation (see Table I.10 from literature review). For example, the total pigmentation of potato tubers from red to purple produced by Mite element *dTstu* in plants obtained from indirect SE (Momose et al. 2010) or the classical example of mutator elements producing spotted kernels in maize plants obtained from indirect organogenesis (Planckaert and Walbot 1989; Peschke and Phillips 1990). In a general way, somaclonal variants seem to be more related to major chromosomal abnormalities, like in the case of the dwarf strawberry derived from micro shoots described by Nehra et al (1992). Although the presence of cytological abnormalities was often associated to SV, such mechanisms do not exclude the simultaneous involvement of genetic or epigenetic events. Little is still known about TEs despite of the abundant bibliography on their activation/insertion mechanisms and the associated consequences on plant genome and gene expression. For biologists, one important remaining question is whether or not TEs are involved in the alteration of important agronomic traits after plant tissue culture. In coffee as in other plant species we can not eliminate this possibility, even when the major responsibility of chromosomal aberrations in SV has been demonstrated here. The results obtained with SSAP did not show polymorphism produced by TEs activation during coffee SE and long-term cell cultures. However, this kind of molecular marker is neither exhaustive nor can exclude the possibility of unseen insertions of TE belonging to families specifically activated by tissue culture. Thanks to the advances in *C. canephora* sequencing, a deeper study of genome structural changes using sequencing technologies like NGS could now be performed in both normal and variant phenotypes of *C. arabica* for comparative analysis.

The use of methylation-sensitive molecular markers in the evaluation of the somatic seedlings

The analysis of tissue culture-derived plants with methylation-sensitive molecular markers (MSAP, Met-AFLP or Southern blot) follows a recent tendency aimed at evaluating the impact of tissue culture at epigenetic level. As seen in previous section there is a wide array of methylation marks in both symmetric and asymmetric sequences (Cokus et al. 2010). Different sequences susceptible to methylation can be analyzed to determinate the frequency of methylation and demethylation generated by *in vitro* propagation. Nevertheless, most methylation sensitive markers (e.g. MSAP, Met-AFLP) evaluate only symmetric sequences meaning that other forms of cytosine methylation are not taken into account and consequently global levels are underestimated. A literature review indicates that epigenetic markers always allowed to show some degree of variation in the propagated plants, ranging from low to high, unlike genetic molecular markers (e.g. AFLPs or SSRs among others), (see Table I.9 from literature review).

Molecular markers evaluations are not intended to obtain an exhaustive situation of the new methylation landscape after tissue culture. They represent an efficient and reliable tool to measure the presence or absence of changes in the methylation patterns of micropropagated plants and to give an idea of their extent. In our MSAPs study, the levels of polymorphisms evaluated were low (i.e. ranged from 0.07–0.18%) for both plant batches obtained from secondary embryogenesis and cell suspensions with no significant difference between the proliferations systems. In the plants regenerated from more or less aged cell cultures, the levels of polymorphisms were similar (0.049-0.087%). Moreover, none of the studied coffee somatic seedlings from both studies including the phenotypic variants accumulated more than 4 methylation changes, showing that epigenetic status was weakly affected even after long-term *in vitro* culture.

Other approaches have been intended in other species to measure the global levels of methylation variation in micropropagated plants with HPLC methods (Baurens et al. 2004), microarrays (Tanurdzic et al. 2008) or sequencing techniques like NGS Illumina Solexa sequencing (Cokus et al. 2010). These studies contributed to a better understanding of the dynamics of methylation. Some examples included the differential levels of methylation found in acacia micro shoots between young and aged leaf developmental stages (Baurens et al. 2004), the transcriptional activation of TEs by the hypomethylation produced in heterochromatic region of Arabidopsis suspension cultures (Tanurdzic et al. 2008) and the differential distribution of epigenetic marks according to the sequence context evaluated in Arabidopsis seedlings (Cokus et al. 2010). Molecular markers are also adaptable to the analysis of global levels but at a lower scale and limited to the sequence context evaluated i.e. the specific sequence context recognized by the restriction enzyme. In view of studying global methylation level variations, molecular marker approaches must use many

primer combinations and evaluate more sequences susceptible to methylation. Other techniques like bisulfite sequencing -i.e. the use of bisulfite treatment of DNA to determine the pattern of cytosine methylation - offer the possibility to work in massive estimations as recently published in Arabidopsis (Cokus et al. 2010). In coffee this could be possible thanks to the advances in *C. canephora* genome sequencing (Coffee.genome.org).

Plant methylation is not static, meaning that methylation changes are also permanently generated through physiological and developmental processes (Furner and Matzke 2010, Law and Jacobsen 2010). Consequently, other sources of variation might be simultaneously evaluated during the analysis of vitro plants. We see the importance in view of further epigenetic change evaluation to take precautions. Of course, all the analyzed plants must grow in the same environment in nursery and field where the culture conditions are more optimized and uniform. On the other side, it is important to study plants and organs showing similar developmental stages to suppress a possible variation related to the physiological variability within the plant batches. All these precautions have been taken into account in the present thesis. Moreover, preliminary studies on the *C. arabica* cv. Caturra seedlings clearly showed that the use of plants and leaves with different developmental stages had no effect on the MSAP patterns obtained. The observation of great quantities of epigenetic changes could lead researchers to the wrong assumption that tissue culture produces a strong stress probably affecting plant phenotype. Nevertheless, no phenotype affectation was reported in all recent studies **presented** in Table I.9 from literature review in spite of the high levels of epigenetic changes reported in some of them. Two hypotheses can be done: (1) The micropropagated plants can somehow tolerate and buffer the changes generated during tissue culture, (2) These methylation changes are related to the normal physiology and development of plants. Using bisulfite sequencing and RNA Seq the next step of epigenetic evaluation in coffee could be to identify molecular marks closely associated with biotic and abiotic stresses including that specifically induced by SE. This approach can also be performed in specific sequence contexts to know their methylaiton status i.e. candidates genes or TEs. The implication of microRNAs in the generation of SV have been recently proposed (Rodriguez-Enriquez et al. 2011). Plant phenotypic plasticity (i.e. changes in plant phenotype in response to the environment) could also be intrinsically related with methylation.

Are all molecular markers' variations indicative of SV?

Molecular markers have efficiently proved the production of polymorphisms during tissue culture. When applied, it is well known that the deeper the analysis (i.e. add more markers or primers combinations into analysis) the greater the possibility to find polymorphisms. However as seen above, even when they are found in high numbers, polymorphisms do not often correlate with phenotypic affectation. Most of polymorphisms are largely found in non coding genome fragments (Meudt and Clarke 2007). This in part could explain that null phenotype affectation be frequently reported. Some

examples included vitro plants from *Arabidopsis thaliana*, *C. arabica* and *Echinacea purpurea* (Polanco and Ruiz 2002, Sanchez-Teyer et al. 2003, Chuang et al. 2009). DNA molecular markers frequently reveal neutral sites of variation at the sequence level that do not reflect in the phenotype (Jones et al 2009). However, most of polymorphisms are indicative of structural DNA changes induced by tissue culture. For this reason this kind of evaluation needs to be performed to evaluate the level and the nature of the variations induced.

Knowing that it will be easier in the future to show molecular variations in the micropropagated plants, either by increasing the depth of the molecular analysis or because of the continuous development of more powerful genomic tools, one wonders: (1) if the information obtained from genomics studies will be useful for the biotechnologists in charge of the development of micropropagation processes, and (2) if any molecular variation found does really correspond to a somaclonal variation. Caricaturing the positions of each other we can think that the point of view of researchers in genomics ("all plants showing an heritable variation at DNA level must be considered as somaclonal variant") will probably strongly differ from that of the biotechnologists involved in micropropagation ('all plants showing a normal phenotype and producing normally are acceptable, disregarding if they show some genetic or epigenetic variation'). In the strict sense, the production of a stable DNA polymorphism "normally heritable" represents a mutational event. However, polymorphisms/mutations as previously stated can be produced in non coding regions or in the case of SNPs can generate synonymous mutations i.e. a base substitution in a given codon that also codes for the required amino acid. Both could have null effect on phenotype. In our point of view, it is preferable to respect the first SV definition given by Larkin and Scowcroft in 1985: "**somaclonal variation is the morphologic variation observed in tissue culture-derived plants and their progeny**". The term SV must be limited to the heritable phenotypic variations. Then, it is obvious that a complete morpho-agronomic evaluation of somatic seedlings must be performed before declaring the genetic conformity. Elevated rates of the corresponding abnormal phenotypes were observed in the seed progenies of Angustifolia, Dwarf and Variegated *C. arabica* somaclonal variants (F. Georget, personal comm.). The heritability of the variant phenotype was checked several times. Hence the observed abnormal phenotypes in SE-derived coffee plants correspond to true somaclonal variants.

The role of cytological mechanisms in SV

How the alteration of specific chromosome segments or loss of whole chromosomes contribute to produce a specific abnormal phenotype is far to be understood. The missing chromosome alters the gene balance in the chromosome set. In Arabidopsis seedlings, aneuploid trisomies (2n+1) produced different abnormal phenotypes due to an additive effect of the extra chromosome(s) (Henry et al. 2010). The monosomic forms of aneuploidy imply the lost of specific chromosome (s) (i.e. 2n-1 chromosomes/plants) and other mechanisms are probably involved in the alteration of phenotype. This

type of aneuploidy could involve two mechanisms both implying: (1) the deregulation of a given group of genes or (2) the expression of recessive alleles in the monosomic chromosome (Griffiths et al. 2000). Coffee abnormal phenotypes were systematically related to abnormal chromosome numbers, mostly in the form of monosomies. However, some of these somaclonal variants presenting the same phenotype (e.g. Dwarf and Angustifolia, see Plos one article) have also exhibited normal chromosomal counts which could contradict this observation. There are two possible explanations. The involvement of chromosomal mosaicism or segmental aneuploidy (extra or missing chromosome segments a.k.a. partial aneuploidy). Chromosomal mosaicism produced by non disjunction generates normal and aneuploid cells in the same tissue i.e. monosomic and trisomic (Griffiths et al. 2000, Holland and Cleveland 2009). Chromosomal mosaicism could pass undetected in chromosomal counts if the numbers of aneuploid cells are low also producing less severe phenotype affectation. However, this seems to be not the case of coffee aneuploidy induced by SE in which the chromosomal counts consistently showed the production of completely aneuploid plants.

Segmental aneuploidy (i.e. the situation in which only one part of the chromosome(s) is lost or gained) seems to be a better explanation. However, this type of aneuploidy is difficult to prove using classical techniques of karyotyping like Centromeres or Giemsa banding. The *C. arabica* chromosomes are quite short and difficult to stain; consequently several attempts to produce quality karyotypes with distinguishable banding patterns have failed (Clarindo and Carvalho 2008). Flow cytometry does not possess the sensitivity and precision needed to detect coffee aneuploidy. This technique can only be used to evidence polyploidy. Other more advances techniques like fluorescence *in situ* hybridization (FISH) are attractive but require technical specialization and the construction of molecular probes to identify efficiently the lacking or extra chromosome segments or aneuploidy. Nevertheless, the advances in the *C. canephora* genome sequencing open the possibility to identify molecular markers specific of each chromosome and more interestingly, to have a set of genes specific of each chromosome and located throughout the segments. In NGS sequencing, the pair-ending mapping property of Illumina allows the identification of many structural differences including the lost of short and long sequence segments. Consequently the partial and total gain/loss of chromosomes could be studied with this technique (Scouarnec and Gribble 2012). This strategy, along with gene expression studies like RT-PCR or more precisely by NGS sequencing (RNA Seq or genome sequencing) represent a good option to study segmental or total aneuploidy in *C. arabica* and other species of the genera. It could be interesting to know if the lacking or additional chromosomes are always the same.

In coffee, most somaclonal variants present phenotypes very similar to previously described mutants observed in seed progenies. The studies of Krug and Carvalho (1951) proposed that most of the mutants found in seed progenies derived from chromosome duplications, especially in the case of

Bullata mutants described as hexaploid and octoploid. We expected a similar situation with the corresponding phenotypes derived from SE. However, we clearly demonstrated that most of these mutants correspond to aneuploid plants rather than polyploids, regardless the weakly or highly mutagenic culture conditions that led to their production. Otherwise, the studies made recently by JC Herrera (CENICAFE, Colombia, pers. comm.) reported that many abnormal phenotypes derived from sexual propagation (see photos below) identical to those described by Krug and Carvalho and to those described here, also correspond to aneuploids and not to polyploids plants. *Coffea arabica* is a recent allotetraploid still having most of its genes in duplicated copies (Cenci et al. 2012). This particularity allows a buffer effect that probably explains its surprising capacity to produce viable forms of aneuploidy. Similarly to our results, the Colombian team of CENICAFE reported higher percentages of zygotic variants presenting monosomy than trisomy (**see figure V.1 and V.2**). Furthermore, similar phenotypes presented monosomy, trisomy or normal chromosome number, as also observed for the Angustifolia and Dwarf variants in our studies. The first conclusion is that aneuploidy is a widespread mechanism in *C. arabica* that can occur during both meiosis (non disjunction in Anaphase I and II) and mitosis (Anaphase) affecting seedlings and somatic seedlings, respectively. It is difficult to explain how aneuploidy *per se* can produce similar phenotypes. As observed for other species, it is possible that most of the time the same chromosomes are concerned and that the lost of such chromosomes lead to a limited number of particular phenotypes. Some repressed genes would be only expressed in the absence of the missing counterparts (Griffiths et al. 2000). One of the hypotheses is that multiple chromosomal losses could create a new genetic landscape in which recessive alleles could become dominant and transcriptionally expressed.

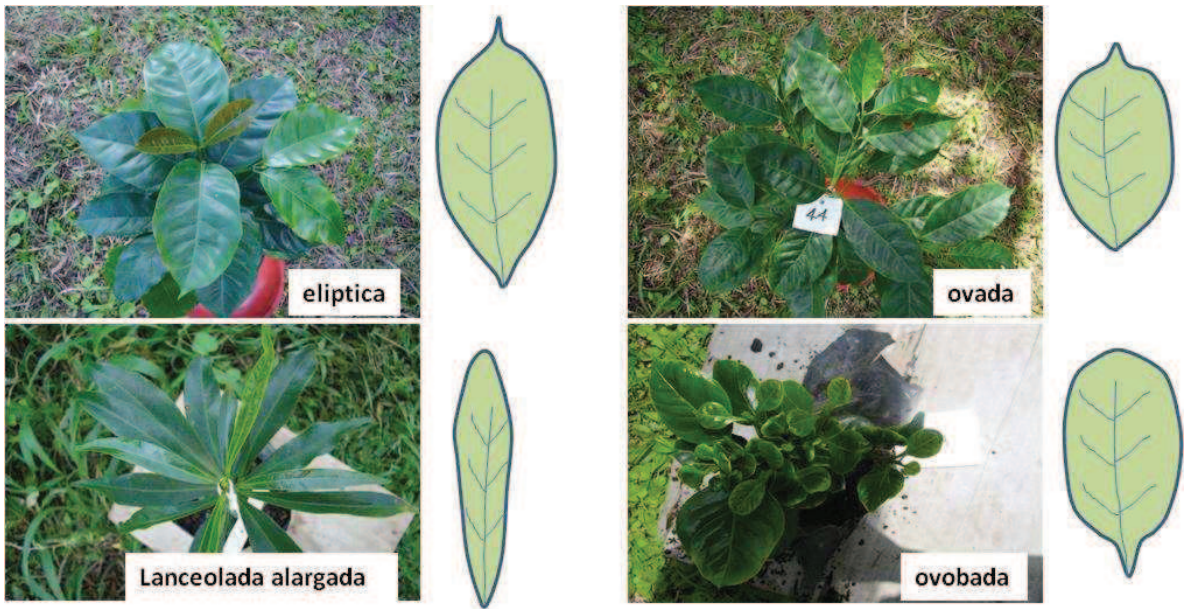


Figure V.1 Different types of seed-derived coffee mutants observed in different intra and interspecific hybrids populations (JC Herrera, CENICAFE unpublished data from CENICAFE, Colombia).

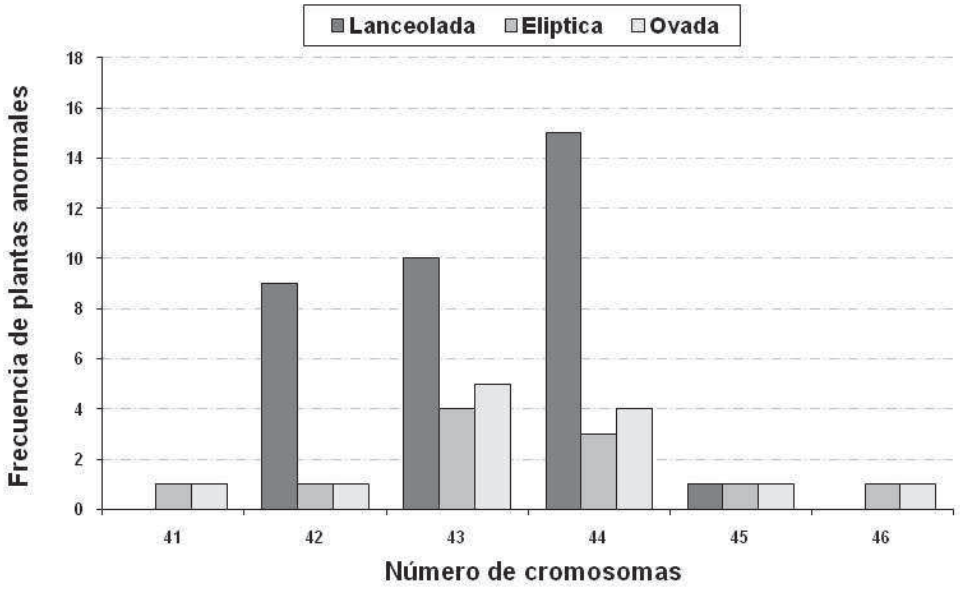


Figure V.2 Frequency of coffee zygotic phenotypic mutants and corresponding ploidy levels. Plants derived from different intra and interspecific hybrid populations (JC Herrera, CENICAFE).

The frequency of aneuploid plants increases in line with embryogenic cultures' age: proposed mechanisms

In eukaryotes, different mechanisms of anaphase non disjunction are related to the production of aneuploidy producing both monosomic and trisomic cells. Such mechanisms were described in recent reviews (Holland and Cleveland 2009) including: (a) weakened mitotic checkpoint that might allow cells to enter anaphase in the presence of unattached or misaligned chromosomes, (b) the premature loss of sister chromatid cohesion and, (c) persisting merotelic attachments (i.e. the attachment of one kinetochore to both mitotic spindle poles) (**see Figure V.3**). Aneuploidy can be generated by the previous mechanisms in proliferating cells. The presence of several types of chromosomal instabilities in coffee was documented in leaf explants and calli by Menendez-Yuffa et al. (2000). Our results indicated the almost total absence of normal plants and the prevalence of monosomic plants, some of them with more than one chromosome lost (i.e. 4n-2 or 4n-3), derived from long-term cell cultures. We wondered if such mechanisms could be sufficient to explain the strong and regular increase of the proportion of aneuploid plants with the ageing of embryogenic cultures till the almost total absence of normal plants and the prevalence of monosomic plants in the regenerated somatic seedlings.

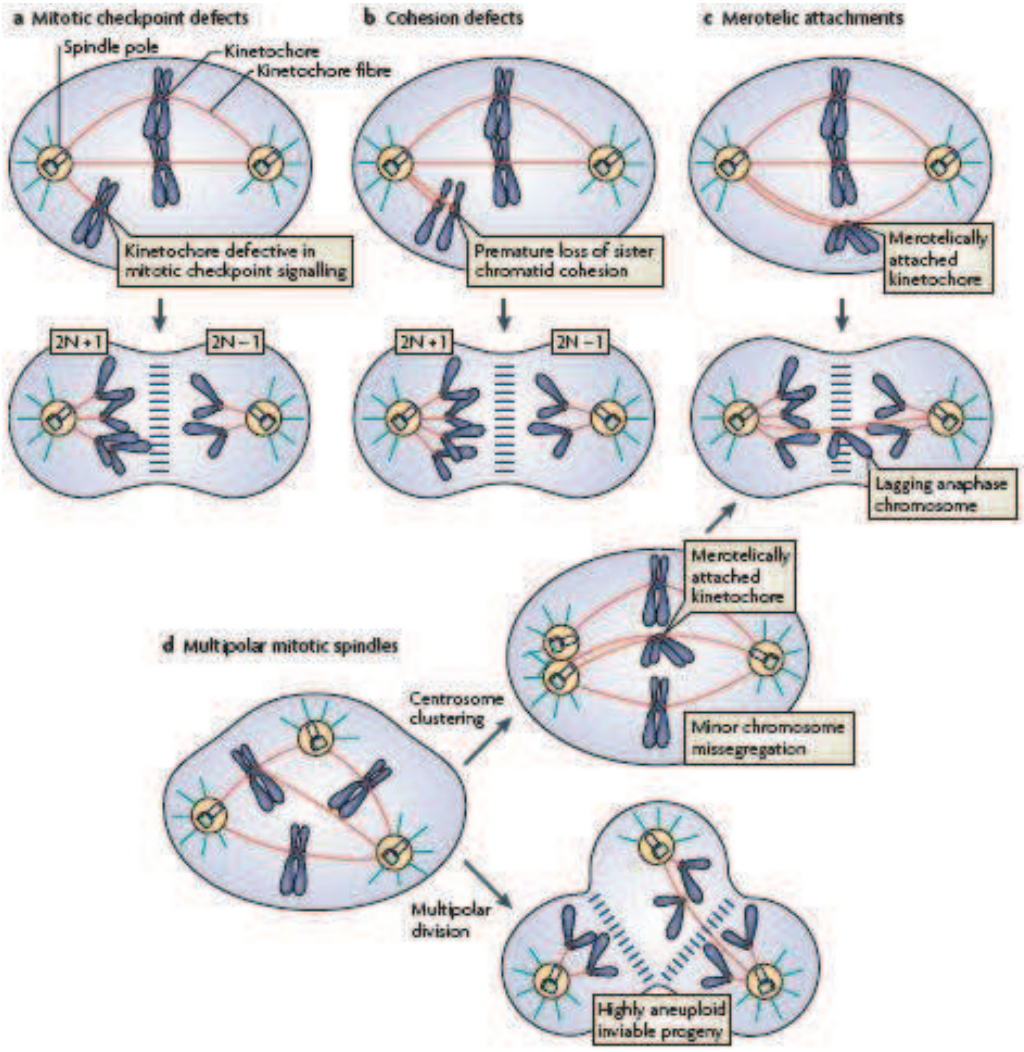


Figure V.3 Pathways in the generation of aneuploidy in eukaryotes (Holland and Cleveland 2009).

Another mechanism, the presence of multipolar mitotic spindles (**Figure V.3-d and Figure V.4-c**) seems to be a better explanation. Cells that possess more than two centrosomes might form multiple spindle poles during mitosis. Plants also present centrosomes but unlike animals they do not contain centrioles. Multipolar mitotic spindles cause high rates of chromosome missegregation and severe aneuploidy (Holland and Cleveland 2009, William et al. 2012). The observations made in the present study never documented this chromosomal instability because cells in prophase or telophase are rarely seen in coffee mitotic preparations (JC Herrera, pers. Comm.). However, this chromosomal instability was reported in coffee *in vitro* culture at the callus stage by Menendez-Yuffa et al. (2000) [**see figure V.4**]. The formation of multiple mitotic spindles implies the disruption of the normal mitotic controls.

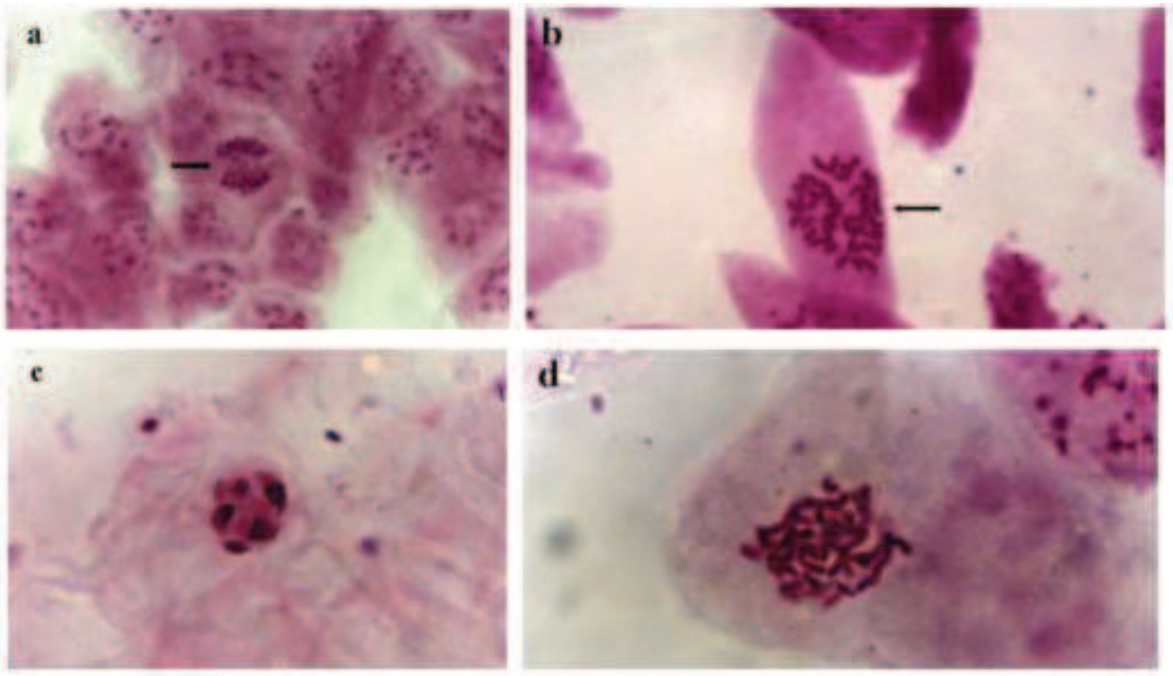


Figure V.4 Coffee callus cells with abnormal mitosis: (a) anaphase with chromosome bridges and lagging chromosomes, (b) C-mitosis failure to form the achromatic spindle, (c) cell with multipolar spindles, (d) metaphase with chained chromosomes (adapted from Menendez-Yuffa et al. 2000).

An abundant literature has described the effects of PGRs on the occurrence of SV and more precisely those associated to elevated concentrations. However there is not conclusive information indicating candidate mechanisms of their action. It is well known that benzyl and phenolic compounds are mutagenic and that some of the most used PGRs contain these molecules (e.g. 2,4-D and 6-BA, Thidiazuron). However, auxins like IAA, NAA and cytokinins known to be less physiologically active and that do not contain these molecules were also shown to produce SV (Biswas et al. 2009). Additionally, PGRs are rarely used alone, so the effect frequently attributed to only one auxin could in fact result from their interaction. A possible mechanism of action could be disruption of the mitotic checkpoint (Holland and Cleveland 2009). In coffee tissue culture 2, 4-D is used to maintain the undifferentiated embryogenic state in callus culture (Ribas et al. 2011). This plant hormone could be involved in the disruption of the mitotic check point. For example, differentiated and undifferentiated cell suspensions of two culture lines of carrot (*Daucus carota*) maintained for 18 days in presence of 2,4-D showed significant higher frequencies of cells with lagging chromosomes and multipolar spindles at anaphase and telophase in comparison to root meristem cells (Bayliss 1973). An elevated frequency of mitotic spindle abnormalities, chromosome bridges and micronuclei and also polyploidy were reported in long-term (12 months) cell cultures of diploid barley maintained under high 2,4-D concentrations but also in cell cultures in presence of NAA (Ziauddin and Kasha 1990). Furthermore, molecules presenting benzoic ring (6-BA) rather than phenolic (2, 4-D) have been recently linked to the production of chromosomal instabilities, especially multipolar spindles. For example, different

chromosomal instabilities including multiple mitotic spindles were produced in *Allium cepa* cells from root tips after 24 and 48 h exposure treatment with Trifluralin herbicide (0.42 to 3.34 ppm) (Fernandes et al. 2009). Other benzoic compound Paclitaxel, an anticancer drug produced multipolar spindles in mitotic cells in human HeLa cell lines (Bian et al. 2010). Consequently both PGRs used in coffee for embryogenic cell proliferation and especially 6-BA could potentially disrupt the mitotic checkpoint and finally produce multiple mitotic spindles. We propose that the increasing and elevated frequencies of aneuploid (i.e. mostly monosomic) plants in proliferating cell cultures be produced by the mechanism of multiple mitotic spindles as a consequence of the disruption of the normal mitotic check point by PGRs action. A detailed cytological evaluation of callus cells after different times of proliferation and under different concentrations of several PGRs could be performed in order to validate the proposed mechanism.

General conclusion

The main purpose of this thesis was to evaluate SV in SE derived *C. arabica* plants and to uncover the underlying mechanisms. The first objective was to evaluate the frequency of abnormal phenotypes in somatic seedlings produced in different SE processes applied in industrial conditions. The phenotyping was performed on a huge quantity of plants i.e. 800 000 somatic seedlings both at nursery and field level. The second objective was to contribute to a better understanding of the mechanisms involved in SV. In this aim different batches of somatic seedlings including normal and variant plants were evaluated at genetic and epigenetic level by using PCR based molecular markers technology. More precisely AFLP and derived techniques were chosen and adapted to coffee to detect polymorphisms affecting genetic mutation (ALFP), methylation (MSAP) and transposable elements (SSAP). The polymorphisms were detected through the systematic comparison with the mother plants that were never submitted to *in vitro* conditions. At the same time, some somatic seedlings were subjected to chromosome counting to evaluate the impact of chromosomal aberrations. The last objective was to produce information allowing to identify some culture parameters possibly involved in SV production and leading to further optimization of SE conditions in view of true-to-type propagation.

In the first chapter we presented a detailed bibliographical review which describes many possible mechanisms and factors associated to SV. Many mechanisms causing SV and proposed long-time ago in pioneer works were discussed. Recent reviews were analyzed with a special focus in SE and the tissue culture conditions applied. Of special interest were the epigenetic aspect related to methylation, the activation of transposable elements and mitotic aberrations. One difficulty found along this bibliographical review work was the lack of uniformity in the presentation of data from genetic or epigenetic evaluation of *in vitro* regenerants. For this reason, and to facilitate comparative

analysis we recalculated all genetic or epigenetic data and transformed them into Total Polymorphisms. This unit represents the global level of genetic/epigenetic variation within the studied population.

In the second chapter (Article Bobadilla et al. 2013 *PLoS ONE*), we showed the trueness-to-type of two SE systems using different proliferation strategies. Both embryogenic suspensions and secondary embryogenesis achieved high proliferation rates along with very low SV rates (less than 1%). SV was evaluated through massive phenotypic observations in a commercial nursery and field plots. Based on molecular markers analysis we also concluded that the genetic and epigenetic changes were limited under controlled culture conditions. The chromosome counting on seven somaclonal variants indicated that abnormal chromosome numbers, more precisely aneuploidy (4n-1, 4n-2, 4n-3 and 4n+1), in a majority of them.

In the Third chapter (Article Bobadilla et al. 2013 *in preparation*), somatic seedlings produced from short, medium and long-term cultures were evaluated at phenotypic, genetic and epigenetic level.The study with 3 independent cell lines showed that SV do not occur trough a random pathway but are directly and strongly affected by the culture duration. .The molecular marker analyses revealed that genetic and epigenetic changes are limited during cell culture. It could be hypothesized that the impact of genetic or epigenetic variations are compensated by the buffer effect of *C. arabica* polyploidy. The particular allopolyploid structure of Arabica (2n=4x=44) also allowed aneuploid cells to survive and regenerate viable plants. The absence of one or more chromosome was buffered by the other homeologous chromosomes. This could not be true for other non polyploid species in which aneuploidy probably occurs but aneuploid cells are non viable and cannot regenerate plants. Hence, *C. arabica* is an interesting species to investigate the occurrence of chromosomal abnormalities in relation with SV. Aneuploidy was detected in somatic seedlings derived from medium and long term cell cultures. This form of genetic variation showed increased number of aneuploid plants closely with time of culture. Aneuploidy was observed in all cellular lines evaluated. Consequently, we can conclude that the production of this form of variation is not random and is time dependent. Other conclusion is that cytological abnormalities are intrinsically related to the mechanisms producing SV in coffee.

In the fourth chapter we discussed the effectiveness of many technologies intended to evaluate plant genetic and epigenetic conformity. Although the technologies applied have shown to be effective or suitable, we concluded we must explore different form of mutation like SNPs. Based on the different analysis possible with NGS, we also concluded that this technology must be the next step in the analysis of coffee SV. This analysis can be useful to uncover other genetic mechanisms involved in SV like SNPs or InDels. On the other hand major cytological abnormalities have been strongly related

with coffee SV but the identity of lacking chromosomes remains unknown. NGS sequencing could not only help to identify major chromosomal lost but also partial aneuploidy.

In the Anexes (Article Etienne et al. 2012 *Cahiers de l'Agriculture*) presented the history of the technological transfer of a coffee SE process especially developed in partnership public/private CIRAD/ECOM group for the large-scale diffusion in Central America and Mexico of selected *C. arabica* F1 hybrids. It describes the difficulties encountered for the transfer of an embryogenic suspension-based SE process and the biological efficiencies obtained at each developmental switch. It showed the technical feasibility of such procedure. This was possible thanks to several technological achievements that made technology reliable and cost effective: the establishment and maintenance of highly embryogenic suspensions for all genotypes, the mass production of somatic embryos in temporary immersion bioreactors and the possibility of sowing them directly in the nursery for plant regeneration.

References

Adomas A, Heller G, Olson A, Osborne J, Karlsson M, Nahalkova J, Van Zyl L, Sederoff R, Stenlid J, Finlay R, Asiegbu FO (2008). Comparative analysis of transcript abundance in Pinus sylvestris after challenge with a saprotrophic, pathogenic or mutualistic fungus. Tree Physiol. 28 (6): 885–897. PMID 18381269.

Ahmad A, Zhang, Y, Cao XF (2010) Decoding the epigenetic language of plant development. Mol. Plant. 3, 719–728.

Ahloowalia BS, Prakash J, Savangikar VA, Savangikar C (2004) Plant tissue culture In: Low Cost Options for Tissue Culture Technology in Developing Countries, IAEA-TECDOC-1384 pp 3-9.

Alves E, Ballesteros I, Linacero R, Va´zquez AM (2005) RYS1, a foldback transposon, is activated by tissue culture and shows preferential insertion points into the rye genome. Theor Appl Genet 111:431–436

Anthony F, Bertrand B, Quiros O, Wilches A, Lashermes P, et al. (2001) Genetic diversity of wild coffee (*Coffea arabica* L.) using molecular markers. Euphytica 118, 53–65

Anthony, F., Combes, M.C., Astorga, C., Bertrand, B., Graziosi, G., and Lashermes, P. (2002). The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. Theor. Appl. Genet. 104, 894–900.

Anthony, F. & Lashermes, P. 2005. The coffee (*Coffea arabica* L.) genome: diversity and evolution. Pp. 207–228 in: Sharma, A.K. & Sharma, A. (eds.), Plant Genome: Biodiversity and Evolution, vol. 1B, Phanerogams. Science Publishers Inc., New Hampshire.

Anthony, F., Diniz, L.E.C., Combes, M.-C., and Lashermes, P. (2010). Adaptive radiation in Coffea subgenus Coffea L. (Rubiaceae) in Africa and Madagascar. Plant Syst Evol.

Anzueto F, Bertrand B, Sarah JL, Eskes AB and Decazy B. 2001. Resistance to *Meloidogyne incognita* in Ethiopian *Coffea arabica* origins: detection and study of resistance transmission. Euphytica 118: 1-8.

Ateeq B, Abdul FM, Niamat AM, Ahmad W (2002) Clastrogenicity of pentachlorophenol, 2,4-D and butachlor evaluated by Allium root tip test. Mut. Res., 514, 105.

Bairu MW, Fennell CW, Van Staden J (2006) The effect of plant growth regulators on somaclonal variation in. Cavendish banana (Musa AAA cv. 'Zelig'). SCI. HORT. 108(4): 347-351.

Bairu MW, Aremu OA, Van Staden J (2011) Somaclonal variation in plants: causes and detection methods. Plant Growth Regul 63:147-173.

Barret P, Brinkman M, Beckert M (2006) A sequence related to rice Pong transposable element displays transcriptional activation by in vitro culture and reveals somaclonal variations in maize. Genome 49:1399–1407

Bayliss MW (1973) Origin of chromosome number variation in cultured plant cells. Nature 246: 529-530.

Bednarek PT, Orłowska R, Koebner RMD, Zimny J (2007) Quantification of tissue-culture induced variation in barley (*Hordeum vulgare* L.). *BMC Plant Biol*, 7:10.

Bellachew, B., 1997. Arabica coffee breeding in Ethiopia: a review. ASIC, 17e coll., Nairobi, Kenya, pp 406-414.

Bennett, M. D., Leitch, I. J. (1995). Nuclear DNA Amounts in Angiosperms. Annals of Botany 76, 113-176.

Berthouly M, Etienne H (1999) Somatic embryogenesis of coffee. In: S.M. Jain, P.K. Gupta & R.J. Newton (Eds.), Somatic Embryogenesis in Woody Plants, pp. 259–288. Kluwer Academic Publishers, UK.

Bertrand B, Etienne H, Cilas C, Charrier A and Baradat B. 2005. *Coffea arabica* hybrid performance for yield, fertility and bean weight. Euphytica 141: 255–262.

Bertrand B, Vaast P, Alpizar E, Etienne H, Davrieux F and Charmetant P. 2006. Comparison of bean biochemical composition and beverage quality of Arabica hybrids involving Sudanese-Ethiopian origins with traditional varieties at various elevations in Central America. Tree Physiology, 26, 1239-1248.

Bertrand, B., Alpizar, E., Montagnon, C., Georget, F., and Etienne, H. (2010). Performance of Arabica F1 hybrids in agroforestry and full-sun cropping systems in comparison with pure lines varieties. Agroforestery Systems.

Bertrand B, Alpizar E, Lara L, SantaCreo R, Hidalgo M, Quijano JM, Montagnon C, Georget F, Etienne H, (2011) Performance of *Coffea arabica* F1 hybrids in agroforestry and full-sun cropping systems in comparison with American pure line cultivars. Euphytica 181: 147-58.

Bhattacharyya, M.K., Smith, A.M., Ellis, T.H.N., Hedley, Martin, C. (1990) The wrinkled-seed character of pea described by Mendel is caused by a trasnposon-like insertion in a gene encoding starch-branching enzyme. Cell 60: 115-122.

Bian M, Fu J, Yan Y, Chen Q, Yang C, Shi Q, Jiang Q, Zhang C (2010) Short exposure to paclitaxel induces multipolar spindle formation and aneuploidy through promotion of acentrosomal pole assembly. Sci China Life Sci. Nov; 53(11):1322-9. doi: 10.1007/s11427-010-4086-1. Epub 2010 Nov 3.

Birchler JA, Veitia RA (2007) The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell 19: 395-402.

Biswas MK, Dutt M, Roy UK, Islam R, Hossain M (2009) Development and evaluation of *in vitro* somaclonal variation in strawberry for improved horticulturel traits. *Scientia Horticulturee*, 122 (3): 409- 416.

Bremer, B. and Eriksson, T. (2009). Time tree of rubiaceae: phylogeny and dating the family, subfamilies, and tribes Int. J. Plant Sci. 170, 766–793.

Bridson, D. M., Verdcourt, B. (1988). Rubiaceae. ed. Flora of tropical East Africa. Balkema, Rotterdam. 415-747.

Bridson, D. M, Verdcourt, B. (2003). Rubiaceae. In: Pope GV, ed. Flora Zambesiaca, Vol. 5, Part 3, Rubiaceae. Kew: Royal Botanic Gardens, 379–720.

Brown, D. C. W. and Thorpe, 1995. Crop improvement through tissue culture, World Journal of Microbiology and Biotechnology, 11: 409-415.

Bodi Z, Zhong S, Mehra S, Song J, Graham N, Li H, May S, Fray RG (2012). "Adenosine Methylation in Arabidopsis mRNA is Associated with the 3' End and Reduced Levels Cause Developmental Defects". Front Plant Sci 3: 48. doi:10.3389/fpls.2012.00048. PMC 3355605. PMID 2263964.

Carrier G, Le Cunff L, Dereeper A, Legrand D, Sabot F, et al. (2012) Transposable Elements Are a Major Cause of Somatic Polymorphism in Vitis vinifera L. PLoS ONE 7(3): e32973. doi:10.1371/journal.pone.0032973.

Carvalho, A. (1952). Taxonomia de Coffea arabica L. Caracteres morfologicos dos haploides. Bragantia 12, 201–212.

Carvalho, A., & L.C. Monaco, 1969. The breeding of arabica coffee. In: F.P. Fewerda & F. Wit (Ed.), Outlines of perennial crop breeding in the tropics. Misc. Pap. Agric. Univ. Wageningen 4: 198-216.

Carvalho, A., Medina Filho, H. P., Fazuoli, L. C., Guerreiro Filho, O., Lima, M. M. A. (1991). Aspectos genéticos do cafeeiro. Revista Brasileira de Genética l4, 135–183.

Charrier A. 1978. Etude de la structure et de la variabilité génétique des caféiers. Bulletin IFCC 14. 100 p.

Clarindo, WR, Carvalho CR (2009). Comparison of the Coffea canephora and C. arabica karyotype based on chromosomal DNA content Plant Cell Rep (2009) 28:73–81.

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE. 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature 452, 215–219.

Coulibaly, I., Noirot, M., Lorieux, M., Charrier, A., Hamon, S., Louarn, J. (2002). Introgression of self-compatibility from Coffea heterocalyx to the cultivated species *Coffea canephora*. Theoretical and Applied Genetics 105, 994–999.

Côte F.X., M. Folliot, R. Domergue & C. Dubois. 2000a. Field performance of embryogenic cell suspension-derived banana plants (Musa AAA, cv. Grande naine). Euphytica 112:245-251.

Cramer, P.J.S. (1957) in F.L. Wellman, Review of Literature of Coffee Research in Indonesia. ASIC Editorial lnternational American Institute of Agriculturel Sciences, Turrialba, Costa Rica.

Cros, J., Combes, M. C., Chabrillange, N., Duperray, C., Monnot des Angles A., et al. (1995). Nuclear DNA content in the subgenus Coffea (Rubiaceae): inter- and intra-specific variation in African species. Can J Bot 73, 14–20.

Cros, J., Combes, M. C., Trouslot, P., Anthony. F., Hamon, S., Charrier, A., Lashermes, P. (1998). Phylogenetic relationships of *Coffea* species: new evidence based on the chloroplast DNA variation analysis. Mol Phylogenet Evol 9, 109–117.

Davis, A. P., Bridson, D., Rakotonasolo, F. (2005). A reexamination of *Coffea* subgenus *Baracoffea* and comments on the morphology and classification of Coffea and Psilanthus (Rubiaceae-Coffeeae). In: Keating RC, Hollowell VC, Croat T (eds) Festschrift for William G. D'Arcy: the legacy of a taxonomist (Monograph in Syst Bot 104). MBG Press, St. Louis, pp 398–420.

Davis, A., Govaerts, R., Bridson, D., and Stoffelen, P. (2006). An annotated taxonomic conspectus of the genus Coffea (Rubiaceae). Botanical Journal of the Linnean Society, 465-512.

Davis, A. P., Govaerts, R., Bridson, D. M., Stoffelen, P. (2006). An annotated taxonomic conspectus of the genus Coffea (Rubiaceae). Bot J Linn Soc 152, 465–512.

Davis, A. P., Chester, M., Maurin, O., Fay, M. F. (2007). Searching for the relatives of *Coffea* (Rubiaceae, Ixoroideae): the circumscription and phylogeny of Coffeeae based on plastid sequence data and morphology. Am J Bot 94, 313–329.

Davis, A. P., Rakotonasolo, F. (2008). A taxonomic revision of the baracoffea alliance: nine remarkable Coffea species from western Madagascar. Bot J Linn Soc 158, 355–390.

De la Puente R, Gonzalez AI., Ruiz ML, Polanco C (2008) Somaclonal variation in rye (*Secale cereale* L.) analyzed using polymorphic and sequenced AFLP markers. In Vitro Cell Dev Biol-Plant 44(5): 419-426.

Dereeper A., Nicolas S., Lecunff L., Bacilieri R., Doligez A., Peros JP., Ruiz M., This P. SNiPlay: a web-based tool for detection, management and analysis of SNPs. Application to grapevine diversity projects.. BMC Bioinformatics. 2011. May 5;12(1):134. (PubMed).

Deo, PC, Tyagi AP, Taylor M, Harding R, Becker, DK. (2010) Factors affecting somatic embryogenesis and transformation in modern plant breeding. The South Pacific Journal of Natural and Applied Sciences, 28(1), pp. 27-40.

Devos KM, Ma J, Pontaroli AC, Pratt LH, Bennetzen JL (2005) Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. Proc Natl Acad Sci USA, 102:19243-19248.

Díaz-Martínez M, Nava-Cedillo A, Guzmán-López JA, Escobar-Guzmán R, Simpson J. Polymorphism and methylation patterns in Agave tequilana Weber var. 'Azul' plants propagated asexually by three different methods. Plant Sci. 2012;185-186:321–330.

Dodeman VL, Ducreux G, Kreis M (1997) Zygotic embryogenesis versus somatic embryogenesis, *J. Exp. Bot.* 48: 1493–1509

Duclercq J, Sangwan-Norreel B, Catterou M, Sangwan RS (2011) De novo shoot organogenesis. Trends Plant Sci. 2011 Nov;16 (11):597-606. doi: 10.1016/j.tplants.2011.08.004. Epub 2011 Sep 8.

Ducos JP, Alenton R, Reano JF, Kanchanomai C, Deshayes A, et al. (2003) Agronomic performance of *Coffea canephora* P. trees derived from large-scale somatic embryo production in liquid medium. Euphytica 131: 215-223.

Duncan RR (1997) Tissue culture-induced variation and crop improvement. *Adv Agron* 58:201-240.

Etienne, H. and B. Bertrand. 2001. Trueness-to-type and agronomic characteristics of *Coffea arabica* trees micropropagated by the embryogenic cell suspension technique. Tree Physiol. 21:1031–1038.

Etienne, H., and Bertrand, B. (2003). Somaclonal variation in *Coffea arabica*: effects of genotype and embryogenic cell suspension age on frequency and phenotype of variants. Tree Physiology 23, 419-426.

Etienne, H. (2005). Protocol of somatic embryogenesis: *Coffee* (*Coffea arabica* L. and *C. canephora* P.). In Protocols for somatic embryogenesis in woody plants (the Netherlands, Jain SM and Gupta PK), pp. 167-179.

Fagerlind F (1937). Embryologische, Zytologische und Bestäubugsexperimentelle Studien in der Familie Rubiaceae nebst Bermerkungen über Einige Polyphoiditatsprobleme. Acta Horti-Bergiani 11, 199-460.

Fazuoli LC, Gonçalvez W, Toma-Braghini M and Silvarolla MB. 2006. Tupi RN 1669-13: A coffee cultivar resistant to *Hemileia vastatrix* and *Meloidogyne exigua* nematode. *In*: XXI International Conference of Coffee Science. ASIC. Montpellier, France. In Press.

Fehér, A., Pasternak, T. P., and Dudits, D. 2003. Transition of somatic plant cells to an embryogenic state. *Plant Cell Tiss Organ Cult.* 74: 201–228.

Fehér A (2008) The initiation phase of somatic embryogenesis: what we know and what we don't. *Acta Biologica Szegediensis*, vol. 52, no. 1, p. 53-56.

Fernandes TC, Mazzeo DE, Marin-Morales MA (2009) Origin of nuclear and chromosomal alterations derived from the action of an aneugenic agent Trifluralin herbicide Ecotoxicol Environ Saf. 2009 Sep;72(6):1680-6. doi: 10.1016/j.ecoenv.2009.03.014. Epub 2009 May 5.

Feschotte C, Jiang N, Wessler SR (2002) Plant transposable elements: when genetics meets genomics. Nat Rev Genet 3: 329-341.

Feschotte C, Pritham E (2007) DNA transposons and the evolution of eukaryotic genomes. AnnuRev Genet 41:331–368.

Filonova LH, Bozhkov PV, Von Arnold S (2000) Developmental pathway of somatic embryogenesis in Picea abies as revealed by time-lapse tracking. *J. Exp. Bot.* 51: 249–264.

Fiuk A, Bednarek PT, Rybczýnski J J (2010) Flow cytometry, HPLC-RP and met AFLP analyses to assess genetic variability in somatic embryo-derived plantlets of *Gentiana pannonica* scop. *Plant Mol Biol Rep*, 28:413-420.

Furner IJ, Matzke M (2010) Methylation and demethylation of the *Arabidopsis* genome. Curr Opin Plant Biol 14:1-5.

Gao X, Yang D, Cao D, Ao M, Sui X, Wang Q, et al. (2010) Invitro micropropagation of *Freesia hybrida* and the assessment of genetic and epigenetic stability in regenerated plantlets. *J Plant Growth Regul*, 29:257–267.

Gil S, Berry D and Bieysse D. 1990. Recherches sur la résistance à *Hemileia vastatrix* Berk. et Br. dans un groupe de génotypes de *Coffea arabica* L. d'origines Ethiopiennes. Café Cacao Thé 34: 105-144.

Giridhar P, Indu EP, Ravishankar GA et al (2004a) Influence of Triacontanol on somatic embryogenesis in *Coffea arabica* L. and *Coffea canephora* P. ex. Fr. In Vitro Cell Dev Biol Plant 40:200–203

Giridhar P, Indu EP, Vinod K et al (2004b) Direct somatic embryogenesis from *Coffea arabica* L. and Coffea canephora P. Ex. Fr. under the influence of ethylene action inhibitor-silver nitrate. Acta Physiol Plant 26:299–305Giridhar P, Vinod K, Indu EP et al (2004c) Thidiazuron induced somatic embryogenesis in *Coffea arabica* L. and Coffea canephora P ex Fr. Acta Bot Croat 63:25–33

Giorgetti L, Castiglione M, Turrini A, Ronchi V, Geri C (2011) Cytogenetic and histological approach for early detection of "mantled" somaclonal variants of oil palm regenerated by somatic embryogenesis: first results on the characterization of regeneration system. *Caryologia* 64: 223-234.

Goevarts, R. M. Ruhsam, L., Andersson, E., Robbrecht, D., Bridson, A., Davis, I., Schanzer, B., Sonké (2006). World checklist of Rubiaceae. Royal Botanic Gardens, Kew. http://www.kew.org/wcsp/rubiaceae/.

Gözükirmizi N, An S, Oraler G, Okatan Y, Onsal N (1990) Callus induction, plant regeneration and chromosomal variations in barley. *Acta Bot.Neerl*. 39(4): 379-387.

Grandbastien MA (1998) Activation of plant retrotransposons under stress conditions. Trends Plant Sci 3:181–187.

Griffiths AJF, Miller JH, Suzuki DT, et al. (2000) An Introduction to Genetic Analysis. 7th edition. New York: W. H. Freeman; Aneuploidy. Available from: http://www.ncbi.nlm.nih.gov/books/NBK21870/.

Guerrero Filho O, Silvarolla MB, Eskes AB (1999) Expression and mode of inheritance of resistance in coffee to leaf miner *Perileucoptera coffeella*. Euphytica 105:7-15

Guo WL, Wu R, Zhang YF, Liu XM, Wang HY, Gong L, Zhang ZH, Liu B (2007) Tissue culture-induced locus-specific alteration in DNA methylation and its correlation with genetic variation in Codonopsis lanceolata Benth. et Hook. f. Plant Cell Rep 26:1297–1307.

Gupta, P. K., Varshney, R. K., Sharma, P. C. & Ramesh, B. 1999 Molecular markers and their applications in wheat breeding. Plant Breed. 118, 369–390. (doi:10.1046/j.1439- 0523.1999.00401.x).

Gutiérrez-Mora A, González-Gutiérrez AG, Rodríguez-Garay B, Ascencio-Cabral, Li-Wei L (2012). Plant Somatic Embryogenesis: Some Useful Considerations, in: Embryogenesis, Dr. Ken-Ichi Sato (Ed.), ISBN: 978-953-51-0466-7.

Hao Y I, Deng XX (2002) Occurrence of chromosomal variations and plant regeneration from long-term-cultured *citrus* callus. *In vitro Cellular and Developmental Biology-Plant* 38: 472-476.

Henderson IR, Jacobsen SE. 2007. Epigenetic inheritance in plants. Nature 447, 418–424.

Henry IM, Dilkes BP, Young K, Watson B, Wu H, et al. (2005) Aneuploidy and genetic variation in the Arabidopsis thaliana triploid response. Genetics.;170:1979–1988.

Henry IM, Dilkes BP, Miller ES, Burkart-Waco D, Comai L (2010) Phenotypic consequences of aneuploidy in *Arabidopsis thaliana*. *Genetics*, 186:1231-45.

Henry Y, Marcotte JL, de Buyser J (1996) The effects of aneuploidy on karyotype abnormalities in wheat plants regenerated from short- and long-term somatic embryogenesis. *Plant Sci*. 114: 101–109.

Hirochika H (1993) Activation of tobacco retrotransposons during tissue culture. EMBO J 12:2521–2528

Hirochika H, Sugimoto K, Otsuki Y, Tsugawa H, Kanda M (1996). Retrotransposons of rice involved in mutations induced by tissue culture. Proc. Natl. Acad. Sci. USA 93, 7783–7788.

Hirochika H, Okamoto H, Kakutani T. (2000). Silencing of retrotransposons in Arabidopsis and reactivation by the ddm1 mutation. *Plant Cell*. 12: 357-369.

Honda H, Liu C, Kobayashi T. (2001) Large-Scale Plant Micropropagation. Adv Biochem Eng Biotechnol.; 72:157-82. Review.

Huang J, Zhang K, Shen Y, Huang Z, Li M, Tang D, Gu M, Cheng Z (2009) Identification of a high frequency transposon induced by tissue culture, nDaiZ, a member of the hAT family in rice.Genomics 93:274–281.

Hua-Van A, Le Rouzic A, (2005) Maisonhaute C, Capy P. Abundance, distribution and dynamics of retrotransposable elements and transposons: similarities and differences. Cytogenet Genome Res ;110: 426–40.

Iida S, Morita Y, Choi JD, Park K I, Hoshino A (2004) Genetics and epigenetics in flower pigmentation associated with transposable elements in morning glories. *Adv. Biophys.* 38, 141–159.

Jain SM (2001) Tissue culture-derived variation in crop improvement. *Euphytica* 118: 153–166.

Jiang N, Bao Z, Zhang X, Hirochika H, Eddy SR, McCouch SR, Wessler SR (2003) An active DNA transposon family in rice. Nature 421:163–167

Jiang C, Mithani A, Gan X, Belfield EJ, Harberd P et al. (2011) Regenerant Arabidopsis lineages display a distinct genome-wide spectrum of mutations conferring variant phenotypes. Curr Biol 21:1385–1390.

Jones RN, Ougham H, Thomas H. 1997. Markers and mapping: we are all geneticists now. New Phytologist 137: 165–177.

Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci U S A. 1996;93(22):12406.

Kaeppler SM, Phillips RL (1993). DNA methylation and tissue culture-induced DNA methylation variation in plants. *In Vitro Cell Dev Bio-Plant*, 29:125-130.

Kaeppler SM, Kaeppler HF, Rhee Y (2000) Epigenetic aspects of somaclonal variation in plants. Plant Mol Biol 43: 179-188.

Kalendar R, Tanskanen J, Immonen S, Nevo E, Schulman AH (2000). Genome evolution of wild barley (*Hordeumspontaneum*) by BARE-1 retrotransposon dynamics in response to Sharp microclimatic divergence. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6603–6607.

Karp A (1991) On the current understanding of somaclonal variation. In: Miflin BJ, editor. Oxford surveys of plant molecular and cell biology. Volume 17. Oxford: Oxford University Press. pp. 1-58.

Karp A (1995) Origins, causes and uses of variation in plant tissue cultures. In: Vasil IK, Thorpe TA (eds) Plant cell and tissue culture. Kluwer Academic Publishers, Dordrecht, pp 139–152

Kashkush, K., Feldman,M. and Levy, A. A.(2003).Transcriptional activation of retrotransposons alters the expression of adjacent genesin wheat. *Nat.Genet.* 33, 102–106.

Kazazian HH (2004) Mobile Elements: Drivers of Genome Evolution. *Science* 303: 1626–1632.

Kikuchi K, Terauchi K, Wada M, Hirano HY (2003) The plant MITE mPing is mobilized in anther culture. Nature 421:167–170

Komatsu M, Shimamoto K, Kyozuka J (2003) Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon Karma. Plant Cell 15:1934–1944

Kumar PS, Mathur VL. (2004). Chromosomal instability in callus culture of Pisum sativum. *Plant Cell Tissue Organ Cult* 78:267–271.

Kobayashi, S., Goto-Yamamoto, N. and Irochika, H. (2004). Retrotransposon-induced mutations in grape skin color. Science 304: 982.

Komamine A, Kawahara R, Matsumoto M, Sunabori S, Toya T, Fujiwara A, Tsukuhara M, Smith J, Ito M, Fukuda H, Nomura K, Fujimura T (1992) Mechanisms of somatic embryogenesis in cell cultures: physiology, biochemistry, and molecular biology. *In Vitro Cell Dev. Biol. Plant* **28**: 11–14.

Krug CA, Mendes JET, Carvalho A (1939) Taxonomia de *Coffea arabica L*. Campinas: Instituto Agronômico do Estado, Bolétim Técnico nº62

Krug CA, Carvalho A (1951) The genetics of *Coffea*. Advances in Genetics 4:127-158.

Lashermes, P., Cros, J., Combes, M. C., Trouslot, P., Anthony, F., Hamon, S., Charrier, A. (1996). Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus *Coffea* L. Theoretical and Applied Genetics 93, 626– 632.

Lashermes, P., Combes, M. C., Trouslot, P., Charrier, A. (1997) Phylogentic relationships of coffee-tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. Theoretical and Applied Genetics 94, 947–955.

Lashermes, P., Combes, M. C., Robert, J., Trouslot, P., D'hont, A., Anthony, F., Charrier, A. (1999). Molecular characterisation and origin of the *Coffea arabica* L. genome. Mol Gen Genet 261, 259-266.

Lashermes, P., Bertrand, B., Herve, E. (2009) Breeding Coffee (*Coffea arabica*) for Sustainable Production in S.M. Jain, P.M. Priyadarshan (eds.), Breeding Plantation Tree Crops: Tropical Species, Springer Science 2009 Chapter 14, 525-543.

Larkin PJ, Scowcroft WR (1981). Somaclonal variation: a novel source of variability from cell cultures for plant improvement. *Theor Appl Genet*, 60:197–214.

Law J, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*, 11: 204-220.

Law RD, Suttle JC. (2005) Chromatin remodeling in plant cell culture: patterns of DNA methylation and histone H3 and H4 acetylation vary during growth of asynchronous potato cell suspensions. Plant Physiology and Biochemistry 43, 527–534.

Lee M, Phillips ML (1987) Genetic variants in progeny of regenerated maize plants Genome, 29, pp. 834–838

Lee ML, Phillips RL (1988) The chromosomal basis of somaclonal variation. *Annu. Rev. Plant Physiol. Plant Mol.* Biol. 39: 413–437.

Lefort-Buson, M. & Y. Dattée (1985). Etude de l'hétérosis chez le colza oléagineux d'hiver (*Brassica napus* L.) – Comparaison de deux populations l'une homozygote et l'autre hétérozygote, Agronomie 5: 101-110.

Leroy JF. (1980). Evolution et taxogenèse chez les Caféiers (*Coffea* L., Psilanthus Hook.f. et Nostolachma Durand). Hypothèse sur leur origine. Compte Rendu Hedomadaire des Séances de l'Académie des Sciences Série D, Sciences Naturelles 29: 593–596.

Le Scouarnec S, Gribble SM. Characterising chromosome rearrangements: recent technical advances in molecular cytogenetics. Heredity. 2012;108:75–85.

Li X, Yu X, Wang N, Feng Q, Liu B, et al. (2007) Genetic and epigenetic instabilities induced by tissue culture in wild barley (*Hordeum brevisubulatum* (Trin.) Link). *Plant Cell Tiss Org Cult*, 90:153–168.

Liu ZL, Han FP, Tan M, Shan XH, Dong YZ, Wang XZ, Fedak G, Hao S, Liu B (2004) Activation of a rice endogenous retrotransposon Tos17 in tissue culture is accompanied by cytosine demethylation and causes heritable alteration in methylation pattern of flanking genomic regions. Theor Appl Genet 109:200–209

LoSchiavo F, Pitto L, Giuliano G, Torti G, Nuti-Ronchi V, et al. (1989) DNA methylation of embryogenic carrot cell cultures and its variations as caused by mutation, differentiation, hormones, and hypomethylating drugs. Theor Appl Genet 77: 325-331.

Louarn, J. (1992). La fertilité des hybrides interspécifiques et les relations génomiques entre caféiers diploïdes d'origine africaine (Genre *Coffea* L. sous-genre *Coffea*). Thèse de Doctorat, Université d' Orsay-Paris-Sud, France.

Makarevitch I, Phillips RL, Springer NM (2008) Profiling expression changes caused by segmental aneuploidy in maize. *BMC Genomics* 9: 7.

Makarevitch I, Harris C (2010) Aneuploidy causes tissue specific qualitative changes in global gene expression patterns in maize. *Plant Physiol*, 152:927–938.

Maurin, O., Davis, A. P., Chester, M., Mvungi, E. F., Jaufeerally-Fakim, Y., Fay, M. F. (2007). Towards a phylogeny for *Coffea* (Rubiaceae): identifying well-supported lineages based on nuclear and plastid DNA sequences. Ann Bot Lond 100, 1565–1583.

Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, et al. (2009) A transposon-induced epigenetic change leads to sex determination in melon. Nature 461: 1135–1138doi:10.1038/nature08498.

Menéndez-Yuffá A, Fernandez R, Rios L, Xena N, (2000) Mitotic aberrations in coffee (*Coffeea arabica* cv. 'Catimor') leaf explants and their derived embryogenic calli. *Electron. J. Biotechnol.* 3, 1–6.

Melayah D, Bonnivard E, Grandbastien M.A, et al. (2001) The mobility of the tobacco Tnt1 retrotransposon correlates with its transcriptional activation by fungal factors. Plant J. 28, 159–168.

Meyer, G.M., Banasiak, M., Ntoyi, T.T., Nicholson, T.L. and Snyman, S.J. 2009. Sugarcane plants from temporary immersion culture: acclimating for commercial production. Acta Hort. (ISHS) 812:323-328 http://www.actahort.org/books/812/812_45.html.

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

Mirouze M, Reinders J, Bucher E, Nishimura T, Mathieu O, et al. (2009) Selective epigenetic control of retro transposition in Arabidopsis. *Nature*, 461:427-430.

Miura A, Yonebayashi S, Watanabe K, Toyama T, Kakutan T, et al. (2001) Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. *Nature* 411:212–214.

Mohandas T, Grant WF (1972) Cytogenetic effects of 2,4-D and amitrole in relation to nuclear volume and DNA content in some higher plants Can. J. Genet. Cytol., 14, pp. 773–783.

Momose M, Abe Y, Ozeki Y (2010) Miniature inverted-repeat transposable elements of Stowaway are active in potato. Genetics 186:59–66.

Morris DA (2000) Transmembrane auxin carrier systems: dynamic regulators of polar auxin transport. Plant Growth Regul 32: 161–172.

Namasivayam P (2007) Acquisition of embryogenic competence during somatic embryogenesis. Plant Cell. Tissue Organ. Cult. 90: 1-8.

Neelakandan AK, Wang K (2012) Recent progress in the understanding of tissue culture-induced genome level changes in plants and potential applications. Plant Cell Rep 31: 597– 620.

Nehra NS, Kartha KK, Stushnoff C, Giles KL (1992) The influence of plant growth regulator concentrations and callus age on somaclonal variation in callus culture regenerants of strawberry *Plant Cell, Tissue and Organ Culture* 29: 257-268.

Netto, K. A., A.E. Miguel, A.R. Queiroz & J.B.D. Pereira, 1993. Estudios de híbridos de *C. Arabica* : Catimor versus Catuai, Catindu versus Catuai e outros. 19éme congresso Brasileiro de Pesquisas cafeeiras, IAC, Campinas, Brazil, pp. 38-40.

Newbury HJ, Howell EC, Crouch JH, Ford-Lloyd BV (2000) Natural and culture induced genetic variation in plantains (Musa spp. AAB group), Aust. J. Bot. 48 493–500.

Ngezahayo F, Xu C, Wang H, Jiang L, Liu B, et al. (2009) Tissue culture-induced transpositional activity of mPing is correlated with cytosine methylation in rice. BMC Plant Biol 9:91.

Noirot M, Poncet V, Barre P, Hamon P, Hamon S, Kochko A (2003). Genome size variations in diploid African *Coffea species*. Ann Bot 92, 709–714.

Nomura K., Komamine A (1995) Physiological and biochemical aspects of somatic embryogenesis. In: Thorpe, T. A., (eds), *In Vitro Embryogenesis in Plants*. Kluwer, Dordrecht, pp. 249–266.

Oury, F.X., P. Koenig, P. Berard & M. Rousset, 1990. Une comparaison entre des blés produits par voie chimique et leurs parents : niveaux d'hétérosis et élaboration du rendement. Agronomie 10: 735-748.

Orozco, F (1975) Heterosis and genotype environment interaction: theorical and experimental aspects. Bull Techn Dep Gen Anim 24: 43-56.

Orton TJ *(*1983) Experimental approaches to the study of somaclonal variation. Plant Mol. Biol. Rep. 1:67-76.

Orzechowska M, K. Stepien K, Kaminska T, Siwinska D (2012) Chromosome variations in regenerants of Arabidopsis thaliana derived from 2- and 6-week-old callus detected using flow cytometry and FISH analyses. Plant Cell Tiss Organ Cult. DOI 10.1007/s11240-012-0232-8.

Ozeki Y, Davies E, Takeda J (1997) Somatic variation during long term subculturing of plant cells caused by insertion of a transposable element in a phenylalanine ammonia-lyase (PAL) gene. Mol Gen Genet 254:407–416.

Peraza-Echeverria S, Herrera-Valencia VA, Kay A-J (2001) Detection of DNA methylation changes in micropropagated banana plants using methylation-sensitive amplification polymorphism (MSAP). Plant Sci 161:359–367.

Peredo EL, A ngeles Revilla M, Arroyo-Garcia R.(2006). Assessment of genetic and epigenetic variation in hop plants regenerated from sequential subcultures of organogenic calli. Journal of Plant Physiology 163, 1071–1079.

Peschke VM, Phillips RL (1991) Activation of the maize transposable element Suppressor-mutator (Spm) in tissue culture. Theor Appl Genet 81:90–97

Peschke VM, Phillips RL (1992) Genetic implications of somaclonal variation in plants. Adv Genet 30: 41–75.

Phillips RL, Kaeppler SM, Peschke VM (1990) Do we understand somaclonal variation? In: Nijkamp, H.J.J., L.H.W. van der Plast, and J. van Aartrijk. (eds.). Progress in plant cellular and molecular biology. Amsterdam: Kluwer Academic Publishers: 131-141.

Phillips GC (2004) In vitro morphogenesis in plants-recent advances. *In Vitro Cell Dev Biology - Plant*, 40(4):342-345. DOI: 10.1079/IVP2004555

Picault N, Chaparro C, Piegu B, Stenger W, Formey D, Llauro C, Descombin J, Sabot F, Lasserre E, Meynard D, Guiderdoni E, Panaud O (2009) Identification of an active LTR retrotransposon in rice. Plant J 58:754–765

Planckaert F, Walbot V, (1989) Molecular and genetic characterization of *Mu* transposable elements in *Zea mays*: behavior in callus culture and regenerated plants. *Genetics,* 123, 567–578.

Polanco C, Ruiz ML (2002) AFLP analysis of somaclonal variation in Arabidopsis thaliana regenerated plants. Plant Sci 162:817–824.

Prakash, N. S., Combes, M. C., Naveen, K. S., Lashermes, P. (2002). AFLP analysis of introgression in coffee cultivars (*Coffea arabica* L.) derived from a natural interspecific hybrid. Euphytica 124, 265-271.

Prakash, N. S., Combes, M. C., Dussert, S., Naveen, S., Lashermes, P. (2005a). Analysis of genetic diversity in Indian robusta coffee genepool (*Coffea canephora*) in comparison with a representative core collection using SSRs and AFLPs. Genetic Res Crop Evol 52, 333–343.

Quiroz-Figueroa FR, Rojas-Herrera R, Galaz-Avalos RM, Loyola-Vargas VM (2006) Embryo production through somatic embryogenesis can be used to study cell differentiation in plants. *Plant Cell Tiss. Org. Cult.* 86:285–301.

Rakocevic A, Mondy S, Tirichine L, Cosson V, Brocard L, Iantcheva A, Cayrel A, Devier B, Abu El-Heba GA, Ratet P (2009) MERE1, a low-copy-number copia-type retroelement in Medicago truncatula active during tissue culture. Plant Physiol 151:1250–1263.

Reffye (de), P., 1973. Le contrôle de la fructification et de ses anomalies chez les *Coffea arabica*, Robusta et leurs hybrides Arabusta. Café Cacao Thé 18: 237-254.

Reinert, J. (1959). Uber die Kontrolle der Morphogenese und die Induktion von Adventivembryonen an Gewebekulturen aus Karotten. Planta, 53, 318–333.

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

Rhee Y, Lin H, Buell R, Childs K, Kaeppler S (2009) Ac2 allele of maize identified in regenerant-derived progeny from tissue culture results from insertion of a novel transposon. Maydica 54:429–437.

Ribas A, Dechamp E, Bertrand B, Champion A, Verdeil JL, et al. (2011) *Agrobacterium tumefaciens*-mediated genetic transformation of *Coffea arabica* (L.) is highly enhanced by using long-term maintained embryogenic callus. BMC Plant Biol 11: 92.

Rodriguez-Enriquez, J, Dickinson, H.G, Grant-Downton, R.T. (2011) MicroRNA misregulation: an overlooked factor generating somaclonal variation? Trends in Plant Science. doi:10.1016/j.tplants.2011.03.002.

Sahijram L, Soneji J, Bollamma K (2003) Analyzing somaclonal variation in micropropagated bananas (*Musa* spp.). In Vitro Cell Dev Bio-Plant 39: 551-556.

Sanchez-Teyer LF, Quiroz-Figueroa F, Loyola-Vargas V, Infante D (2003) Culture-induced variation in plants of *Coffea arabica* cv. Caturra rojo, regenerated by direct and indirect somatic embryogenesis. Mol Biotechnol 23: 107–115.

Sato M, Kawabe T, Hosokawa M, Tatsuzawa F, Doi M (2011) Tissue culture-induced flower-color changes in *Saintpaulia* caused by excision of the transposon inserted in the *flavonoid 3′*, *5′ hydroxylase* (*F3′5′H*) promoter. Plant Cell Rep 30: 929–939.

Sato M, Hosokawa M, Doi M (2011) Somaclonal Variation Is Induced De Novo via the Tissue Culture Process: A Study Quantifying Mutated Cells in Saintpaulia. *PLoS ONE* 6(8): e23541. doi:10.1371/journal. Pone. 0023541.

Saze H, Kakutani T (2011) Differentiation of epigenetic modifications between transposons and genes. *Current Opinion in Plant Biology* 14:81–87.

Schnable PS, et al. (2009) The B73 maize genome: Complexity, diversity, and dynamics. Science 326(5956): 1112–1115.

Schellenbaum P, Mohler V Wenzel G, Walter B (2008) Variation in DNA methylation patterns of grapevine somaclones (*Vitis vinifera* L.). *BMC Plant Biol*, 8:78.

Scholz, F. & G. Kunzel, 1987. Hybrid barley. Problems and advances especially in developing genetic systems. In: S. Yasuda et K. Konishi (Ed.), Barley Genetics,. V Proc IBGS, Okayama, Japon, pp. 937-943.

Shchukin, A.; Ben-Bassat, D.; Israeli, Y. Plant regeneration via somatic embryogenesis in Grand Naine banana and its effect on somaclonal variation. Acta Hort. 447:317–318; 1997.

Silvestrini, M., Maluf, M., Silvarolla, M., Guerreiro, O., Medina, H., Vanini, M., Oliveira, A., de Gaspari-Pezzopane, C., and Fazuoli, L. (2008). Genetic diversity of a Coffea Germplasm Collection assessed by RAPD markers. Genetic Resources and Crop Evolution, 901-910.

Skirvin RM, McPheeters KD, Norton M (1994) Sources and frequency of somaclonal variation. *HortScience* 29:1232-1237.

Skoog F, and Miller CO (1957). Chemical regulation of growth and organ formation in plant tissues cultured in vitro. Symp. Soc. Exp. Biol. 54, 118–130.

Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet*, 8:272-285.

Smith MK, Drew RA (1990) Current applications of tissue culture in plant propagation and improvement. Austral. J. Plant Physiol. 17:267–289.

Smith AM, Hansey CN, Kaeppler SM (2012) TCUP a novel *h*AT transposon active in maize tissue culture. Frontier in Plant Science, 3: 6.

Smulders MJM, de Klerk GJ (2011) Epigenetics in plant tissue culture. *Plant Growth Reg*, 63:137-146. Steward FC, Mapes MO, Kent AE, Holsten RD (1964) Growth and development of cultured plant cells. Science 143:20–27.

Smýkal P, Valledor L, Rodriguez 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.

Söndahl, M. R.; Sharp, W. R. (1977) High frequency induction of somatic embryos in cultured leaf explants of Coffea arabica L. Z. Pflanzenphysiol. 81:395–408.

Söndahl MR, Lauritis JA (1992) Coffee. In : Hammerschlag FA, Litz RE, editors. Biotechnology of Perennial Fruit Crops. Wallingford: CAB International. pp. 401-420.

Sree Ramulu K, Dijkhuis P, Roest S (1983) Phenotypic variation and ploidy level of plants regenerated from protoplasts of tetraploid potato (Solanum tuberosum L. cv. 'Bintje'). Theor Appl Genet 65:329–338.

Srinivasan KH, Narasimhaswamy RL (1975) A review of coffee breeding work done at the Government coffee experiment station, Balehonnur. Indian coffee 34:311-321

Steward, F. C., Mapes, M. O., & Mears, K. (1958). Growth and organized development of cultured cells. II. Organization in cultures grown from freely suspended cells. American Journal of Botany, 45, 705–708.

Stoffelen P, Noirot M, Couturon E, Bontems S, De Block P, Anthony F (2009) *Coffea anthonyi* Stoff. & F. Anthony, a new selfcompatible Central African coffee species, closely related to an ancestor of *C. arabica* L. Taxon 58, 133–140.

Sugimoto K. et al. (2010) Arabidopsis regeneration from multiple tissues occurs via a root development pathway. Dev. Cell 18, 463–471.

Sugimoto K, Gordon SP, Meyerowitz EM (2011) Regeneration in plants and animals: dedifferentiation, transdifferentiation or just differentiation? Trends Cell Biol;21(4):212–218.

Sunderland N (1977) Nuclear cytology. In: H.E. Street (Ed.), Plant Tissue and Cell Culture, pp. 177–205. Blackwell, Oxford.

Takeda S, Sugimoto K, Otsuki H, Hirochika H (1999) A 13-bp cisregulatory element in the LTR promoter of the tobacco retrotransposon Tto1 is involved in responsiveness to tissue culture, wounding, methyl jasmonate and fungal elicitors. Plant J 18:383–393

Tahara M, Aoki T, Suzuka, S., Yamashita H, Tanaka M, Matsunaga S, Kokumai S (2004). Isolation of an active element from a high-copy-number family of retrotransposons in the sweetpotato genome. Mol. Gen. Genomics 272: 116–127.

Takayama S and Akita M (2006) Bioengineering aspects of bioreactor application in plant propagation In: *S* Dutta Gupta, Y Ibaraki (eds.) *Plant tissue culture engineering*, series: focus on biotechnology, vol. 6. Dordrecht, the Netherlands Springer, pp. 82-98.

Tanurdznic M, Vaughn MW, Jiang H, Tae-Jin L, Martienssen RA, et al. (2008) Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biol* 6, 2880-2895.

Teixeira da Silva JA, Giang DTT, Chan M-T, Sanjaya, Norikane A, Chai M-L, Ruíz JC, Penna S, Granström T, Tanaka M (2007) The influence of different carbon sources, photohetero-, photoauto- and photomixotrophic conditions on protocorm-like body organogenesis and callus formation in thin cell layer culture of hybrid *Cymbidium* (Orchidaceae). *Orchid Science and Biotechnology* 1, 15-23.

Tremblay L, Levasseur C, Tremblay FM (1999) Frequency of somaclonal variation in black spruce (*Picea mariana,* Pinaceae) and White spruce (*P. glauca,* Pinaceae) derived from somatic embryogenesis and identification of some factors involved in genetic instability. *Am J Bot*, 86:1373-1381.

Torres, E.M., Williams, B.R., and Amon, A. (2008). Aneuploidy: cells losing their balance. Genetics 179, 737–746.

Van der Vossen DJA. (1985). Coffee selection and breeding. *In*: Coffee botany biochemistry and production of beans and beverage. Ed. by C. Clifford and J. Wilson. Westford, Connecticut, EE.UU. pp. 46–68.

Van Harten AM (1998) Mutation breeding: theory and practical applications. Cambridge University Press, Cambridge.

Van Boxtel J and Berthouly M (1996) High frequency somatic embryogenesis from coffee leaves: factors influencing embryogenesis, and subsequent proliferation and regeneration in liquid medium. *Plant Cell, Tissue Organ Culture*, January 1996, vol. 44, no. 1, p. 7-17.

Vanyushin BF, Ashapkin VV (2011) DNA methylation in higher plants: Past, present and future. Biochim Biophys Acta 1809: 360-368.

Verdeil JL, Alemanno L, Niemenak N, Tranbarger TJ (2007) Pluripotent versus totipotent plant stem cells: dependence versus autonomy? Trends in Plant Science 12, 245-52.

Von Arnold S, Sabala I, Bozhkov P, Dyachok J and Filonova L (2002) Developmental pathways of somatic embryogenesis. Plant Cell Tiss Org. Cult. 69: 233–249.

Vos P, Hogers R, Bleeker M, Rijans M, Zabeau M et al. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23:4407-4414.

Walyaro DJ (1983) Considerations in breeding for improved yield and quality in Arabica coffee (Coffea arabica L.). Doctoral Thesis, Agricultural University, Wageningen.

Waugh R, McLean K, Flavell A J, Pearce SR, Powell W, et al. (1997) Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence- specific amplification polymorphisms (S-SAP). *Mol. Gen. Genet*. 253, 687-694.

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P and Schulman AH (2007) A unified classification system for eukaryotic transposable elements. Nat Rev Genet 8: 973-982.

Xiao H, Jiang N, Schaffner EK, Stockinger EJ, van der Knaap E (2008) A retrotransposon mediated gene duplication underlies morphological variation of tomato fruit. Science 319:1527–1530.

Yamashita H, Tahara M (2006) A LINE-type retrotransposon active in meristem stem cells causes heritable transpositions in the sweet potato genome. Plant Mol Biol 61:79–94

Yang X, Zhang X (2010). Regulation of somatic embryogenesis in higher plants. Critical Reviews in Plant Sciences, 29: 36-57.

Yasuda, T., Y. Fujii, and T. Yamaguchi (1985) Embryogenic callus induction from *Coffea arabica* leaf explants by benzyladenine, Plant Cell Physiol. 26:595-597.

Yu X, Li X, Zhao X, Jiang L, Miao G, Pang J, Qi X, Liu B (2011) Tissue culture-induced genomic alteration in maize (Zea mays) inbred lines and F1 hybrids. Ann Appl Biol 158:237–247.

Zhang M, Wang H, Dong Z, Qi B, Xu K, Liu B (2010a) Tissue culture-induced variation at simple sequence repeats in sorghum (Sorghum bicolor L.) is genotype-dependent and associated with down-regulated expression of a mismatch repair gene, MLH3. Plant Cell Rep 29:51–59.

Zhang M, Kimatu JN, Xu K, Liu B (2010) DNA cytosine methylation in plant development. Review. *J Genet Genomics*, 37:1-12.

Zhang MS, Xu CM, Yan HY, Zhao N, Von Wettstein D, Liu B (2009). Limited tissue culture-induced mutations and linked epigenetic modifications in F1 hybrids of sorghum pure lines are accompanied by increased transcription of DNA methyltransferases and 5-methylcytosine glycosylases. *Plant Journal*, 57:666-679.

Zimmerman JL (1993) Somatic embryogenesis: A model for early development in higher plants. Plant Cell 5, 1411-1423.7:607–613. [73] DOI 10.1007/s11101-008-9089-1.

ANNEX

Un exemple de transfert de technologie réussi dans le domaine de la micropropagation: la multiplication de *Coffea arabica* par embryogenèse somatique

Etienne H, Bertrand B, Montagnon C, Bobadilla-Landey R, Déchamp E, Jourdan I, Alpizar E, Malo R and Georget F

This chapter was written in French and published in Cah Agric 21: 115-24. doi: 10.1684/agr.2012.0553

Résumé

Parmi les techniques possibles de micropropagation, la multiplication végétative par embryogenèse somatique est de loin la plus prometteuse pour la diffusion rapide et à grande échelle d'individus élites. Pourtant à ce jour, il n'existe que de très rares exemples de procédés de propagation par embryogenèse somatique appliqués au niveau commercial. Les blocages sont multiples et se situent le plus souvent au niveau d'un effet génotypique important en particulier pour l'obtention de tissus embryogènes, au niveau de la qualité des embryons somatiques régénérés, de l'incidence de variations somaclonales et plus généralement au niveau du manque de reproductibilité et d'efficacité de certaines étapes conduisant à des coûts de production jusqu'ici prohibitifs. Les recherches sur l'embryogenèse somatique du caféier ont débuté dans différents instituts, y compris le CIRAD, dès la fin des années 70. Entre 1995 et 2001, le CIRAD a fait évoluer la technique fonctionnant à l'échelle laboratoire de recherche vers un procédé pouvant permettre la diffusion industrielle d'hybrides F1 de *Coffea arabica* extrêmement prometteurs. Au cours de cette période, deux innovations technologiques ont rendu le transfert technologique économiquement envisageable : la production en masse d'embryons somatiques en bioréacteurs à immersion temporaire et la possibilité de les semer directement en pépinière. Parallèlement, des données rassurantes concernant la conformité génétique des plantes régénérées étaient obtenues (fréquence de variations somaclonales < 3%). En 2002, le CIRAD en partenariat avec le groupe ECOM a décidé de mettre en œuvre le transfert industriel de la méthode d'embryogenèse somatique pour diffuser en Amérique centrale quatre clones d'hybrides d'Arabica sélectionnés pour des systèmes de culture en agroforesterie. Le présent article décrit chez le caféier les différentes étapes et les difficultés qui ont du être surmontées jusqu'à un transfert technologique réussi en 2010 et à un des tout premiers exemples d'application commerciale de la technologie d'embryogenèse somatique.

Mots clés: Conformité, coûts de production, embryogenèse somatique, pépinière, transfert technologique, variations somaclonales.

Summary

Of all the possible micropropagation techniques, vegetative propagation by somatic embryogenesis is by far the most promising one for the rapid, large-scale dissemination of elite individuals. Yet, to date, examples of somatic embryogenesis processes applied on an industrial scale are very few and far between. There are many complications. They usually involve a major genotypic effect, particularly for obtaining embryogenic tissues, or are related to the quality of regenerated somatic embryos, the incidence of somaclonal variation and, more generally, a lack of reproducibility and efficiency at certain stages of the process, leading to production costs that are prohibitive. Research on coffee somatic embryogenesis began at the end of the 1970s at various institutes, including CIRAD. Between 1995 and 2001, CIRAD moved the technique forward from a research laboratory scale to a technique enabling industrial dissemination of extremely promising *Coffea arabica* F1 hybrids. Over that period, two technological innovations made technology transfer economically feasible: mass production of somatic embryos in temporary immersion bioreactors and the possibility of sowing them directly in the nursery. At the same time, reassuring data were obtained on the genetic conformity of regenerated plants (somaclonal variation frequency < 3%). In 2002, in partnership with the ECOM group, CIRAD decided to transfer the somatic embryogenesis method on an industrial scale to Central America so that four Arabica hybrid clones, that were selected for agroforestry-based farming systems could be disseminated throughout that part of the word. This article describes in coffee tree the different stages and the difficulties we had to overcome before successful technology transfer could occur in 2010. It describes one of the first examples of somatic embryogenesis technology applied on a commercial scale.

Key words: Fidelity, nursery, production costs, somaclonal variation, somatic embryogenesis, technology transfer

Introduction

L'embryogenèse somatique, une technologie très attendue!

Lorsqu'elle est initiée comme chez le caféier à partir de tissus (ou 'explants') maternels prélevés sur l'arbre adulte, l'embryogenèse somatique permet la propagation végétative rapide et massive de génotypes élites en s'affranchissant des longs et coûteux processus de sélection généalogique. A ce jour, il n'existe pourtant que de très rares exemples d'application commerciale de procédés d'embryogenèse somatique. On peut cependant citer comme exemples le pin à encens (*Pinus taeda*) [Gupta et Hartle, 2010], le palmier à huile (Khaw et al., 1999) ou encore le caféier, que ce soit l'espèce Arabica (présent article) ou Canephora (Sampote et al., 2000 ; Ducos et al., 2010), pour lesquels la production annuelle dépasse à présent un à plusieurs millions de plantes. Pourtant, de l'avis général, cette technique de multiplication végétative est de loin la plus prometteuse pour capturer au plus vite le gain génétique par la diffusion rapide et à grande échelle des individus élites. Ceci est d'autant plus vrai avec les espèces ligneuses pour lesquelles les cycles biologiques et, partant, les délais pour la sortie variétale, sont longs. Dans les années 80 l'engouement pour développer cette technologie est fort et les attentes immenses, expliquant que des recherches soient alors engagées chez un très grand nombre d'espèces, parfois sans justification agronomique immédiate dans certains cas. Les blocages pour développer cette technologie sont multiples et se situent le plus souvent au niveau d'un effet génotypique important en particulier pour l'obtention de tissus embryogènes, au niveau de la qualité des embryons somatiques régénérés, de l'incidence de variations somaclonales et plus généralement au niveau du manque de reproductibilité et d'efficacité de certaines étapes conduisant à des coûts de production jusqu'ici prohibitifs.

Diffuser au plus vite le progrès génétique chez l'espèce Arabica

Les recherches sur l'embryogenèse somatique du caféier ont débuté dans différents instituts, y compris au CIRAD, dès la fin des années 70 mais sans objectif très clair. Au début des années 90, le CIRAD en partenariat avec le réseau de recherche centro-américain PROMECAFE se lance dans la création d'hybrides F1 intraspécifiques chez *Coffea arabica* en croisant les variétés traditionnellement cultivées en Amérique Latine et des individus sauvages originaires d'Ethiopie et du Kenya. Les hybrides ainsi obtenus se révèlent extrêmement prometteurs car ils présentent un fort niveau d'hétérosis, avec des gains de rendement de l'ordre de 40% comparés aux meilleures variétés cultivées, et pour certains d'entre eux, produisent un café dont la qualité organoleptique s'avère supérieure à celle des variétés de référence (Bertrand et al., 2005 ; cf article dans ce même numéro). La nécessité de disposer d'un procédé d'embryogenèse somatique capable de propager massivement les meilleurs clones d'hybrides F1 d'Arabica s'impose rapidement auprès des co-obtenteurs de ces

nouvelles variétés. Mais il existe un pas important de la technique mise au point en laboratoire de recherche à un procédé industriel permettant la production annuelle de plusieurs millions de plants. Les co-obtenteurs décident de financer la recherche nécessaire pour réussir ce premier changement d'échelle. Il sera réalisé sous la direction du CIRAD au CATIE, un centre de recherche régional situé au Costa Rica.

Résultats

L'état des lieux avant le transfert technologique (1995-1996): repérage des points à améliorer

Plusieurs limitations sont identifiées qui s'opposent à un transfert technologique en l'état du procédé d'embryogenèse somatique développé jusqu'ici par le CIRAD. Tout d'abord, les coûts de production : en effet la mise au point d'un logiciel permettant d'estimer une fourchette des coûts de production dans différentes conditions de culture livre son verdict : 1,5-2 USD/plante alors qu'un semenceau conventionnel se vend 0,25-0,35 USD ! Handicap supplémentaire, en Amérique latine, les densités de plantation pratiquées chez l'Arabica se situent entre 6 et 8.000 arbres/hectare en raison du nanisme des variétés utilisées qui a permis l'intensification de la production dans les années 60. Le surcoût nécessaire pour planter des vitroplants doit être limité même si une plus-value significative est espérée avec du matériel hybride. Le logiciel s'est avéré également utile pour identifier précisément les étapes du procédé responsables de ce coût de production élevé ; il s'agit des étapes tardives incluant la germination et le développement en plantules sevrables, c'est-à-dire possédant au moins deux paires de feuilles pour supporter le choc de l'acclimatation aux conditions *ex vitro*. Cette période de croissance *in vitro* génère classiquement une main d'œuvre importante à cause de l'inévitable repiquage mensuel et de la fabrication des milieux nutritifs gélifiés mais aussi la mobilisation de surfaces importantes dans les salles de culture. La deuxième limitation est le risque non apprécié jusqu'alors que le procédé d'embryogenèse somatique puisse provoquer une fréquence élevée de variations somaclonales. Ces 'erreurs de photocopie' sont indésirables car les plants 'variants' ne présentent plus toutes les qualités agronomiques de la 'plante mère' sélectionnée. Les variations somaclonales représentent un problème récurrent dans les cultures *in vitro* et particulièrement avec les systèmes d'embryogenèse somatique qui font intervenir des concentrations relativement élevées d'auxines telles que le 2,4-D et l'AIA pour induire la formation et la multiplication de cellules embryogènes. L'implication de ces régulateurs de croissance dans l'induction des variations somaclonales est connue depuis longtemps (Karp, 1994).

Des innovations technologiques et des informations rassurantes sur la conformité génétique (1996-2001)

Au cours de cette période, deux innovations techniques rendent un transfert technologique économiquement envisageable : i) la production en masse d'embryons somatiques pré-germés en bioréacteurs à immersion temporaire (Etienne-Barry et al., 1999 ; Albarrán et al. 2005) et ii) la possibilité de les semer directement sur substrat horticole pour obtenir une régénération de plantules photo-autotrophes en pépinière (Etienne-Barry et al. 1999, 2002a, b). Ces deux sauts technologiques permettent de transférer la plus grande partie des étapes tardives (germination, conversion des embryons en plantes) du laboratoire vers la pépinière et ceci réduit considérablement les coûts de production. Le prix de revient d'une plantule est ainsi estimé à 0,5 USD. Le pari est fait qu'une réduction des coûts sera encore possible en passant aux conditions de production industrielle.

Parallèlement, des données rassurantes concernant la conformité génétique des plantes régénérées sont obtenues. D'une part, la fréquence de variations somaclonales au champ s'avère relativement faible (inférieure à 3%). D'autre part, les seules variations observées en cinq années sont qualitatives, c'est-à-dire aisément identifiables au niveau phénotypique, et non quantitatives. Par exemple, la quantité de café produite ou la teneur du grain en éléments biochimiques ne sont pas modifiées (Etienne et Bertrand 2001). Sept types de variants phénotypiques sont alors décrits ; les variants Angustifolia (feuilles étroites), Variegata (panachure des feuilles) et Nain sont les plus fréquents (Etienne et Bertrand 2003). Par ailleurs sont précisées des conditions de multiplication du matériel embryogène en suspensions cellulaires permettant de contrôler la régénération de variants somaclonaux. En 2001, le procédé est donc considéré comme transférable au niveau industriel, d'autant plus qu'il a fonctionné sur la totalité des 19 hybrides F1 testés.

Montage du partenariat (2003)

A partir de 1999-2000, le CIRAD est décidé à aller jusqu'au terme du processus de valorisation de l'embryogenèse somatique pour multiplier à grande échelle les hybrides F1 mais aussi pour se construire une expérience utile pour les autres espèces tropicales pour lesquelles l'application de cette technologie est envisagée. Il recherche un partenaire intéressé par ce transfert technologique chez *C. arabica*. Un contrat est signé avec le groupe ECOM en 2003. Le groupe suisse, négociant en café de qualité et très implanté en Amérique latine, particulièrement au Mexique et en Amérique Centrale se révèle très intéressé, car il est soucieux pour son approvisionnement en café haut de gamme dans cette zone. A ce moment là, les essais agronomiques à partir de clones d'hybrides F1 révèlent leur remarquable adaptation aux conditions d'agroforesterie et confirment l'excellence de certains clones au niveau organoleptique (Bertrand et al. 2011). Le groupe ECOM est logiquement très intéressé car la

majorité des caféières en Amérique latine est conduite en agroforesterie. L'adoption des hybrides F1 pourrait permettre d'augmenter les quantités et la qualité du café produit. Les partenaires choisissent le Nicaragua comme lieu de transfert technologique pour diffuser les hybrides d'Arabica dans toute l'Amérique Centrale.

Chacun des deux partenaires est cependant conscient de la difficulté d'un tel transfert technologique mais probablement pour des raisons différentes. Le CIRAD est concentré sur les difficultés d'ordre technique liées au transfert technologique proprement dit et au changement d'échelle important à réussir (passer d'une production annuelle de 50.000 plantes à plusieurs millions). De son côté, le principal souci du partenaire est de pouvoir vendre les vitroplants car il s'inquiète à juste titre de la double particularité de ce nouveau matériel végétal : hybride F1 et vitroplant. Effectivement, à cette date il n'existe pas d'exemple connu de programme de sélection ayant conduit à la diffusion commerciale d'hybrides F1 ou de vitroplants. Le marché est à créer de toutes pièces et les réticences de la part des caféiculteurs risquent d'être nombreuses.

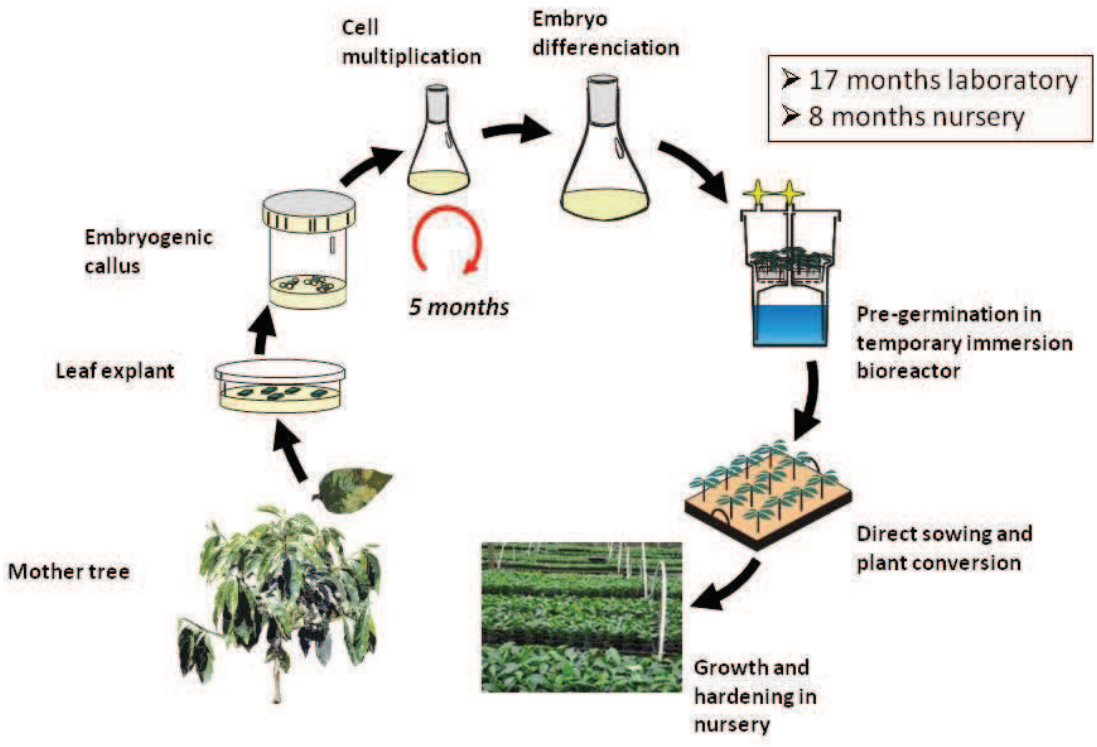


Figure 1. Représentation schématique du procédé industriel de propagation *in vitro* du caféier par embryogenèse somatique.

Construction des infrastructures et premiers ajustements (2004-2006)

Le choix a été fait de construire un laboratoire fonctionnel de 300 m^2 (**Fig. 2A**), une structure assez petite pour ne pas alourdir les coûts de production par des frais incompressibles (fluides, surfaces..). Il est également décidé de le localiser sur le même site qu'une grande usine de traitement de café ('Beneficio') à Sebaco, une petite ville située à 100 kms de la capitale Managua, l'objectif étant que les nombreux producteurs qui apportent leur café « pergamine » pour traitement à l'usine, puissent également découvrir le matériel hybride et se familiariser avec cette méthode innovante de micropropagation *in vitro*.

Une collection de 'pied-mères' (boutures horticoles ou greffes des hybrides sélectionnés) est installée à proximité du laboratoire. Elle fournira le matériel végétal de base nécessaire à la propagation *in vitro*. Six à huit copies clonales de chaque arbre sélectionné y sont maintenues à l'état végétatif (pas de fructification) et dans un excellent état sanitaire afin de favoriser la réactivité des explants après introduction *in vitro*. Les pépinières d'acclimatation des embryons somatiques, d'endurcissement et de grossissement sont mises en place à proximité de Matagalpa dans une ferme appartenant à l'un des producteurs du groupe (Finca « La Cumplida »), située à 30 kms du laboratoire, dans la zone de culture du caféier.

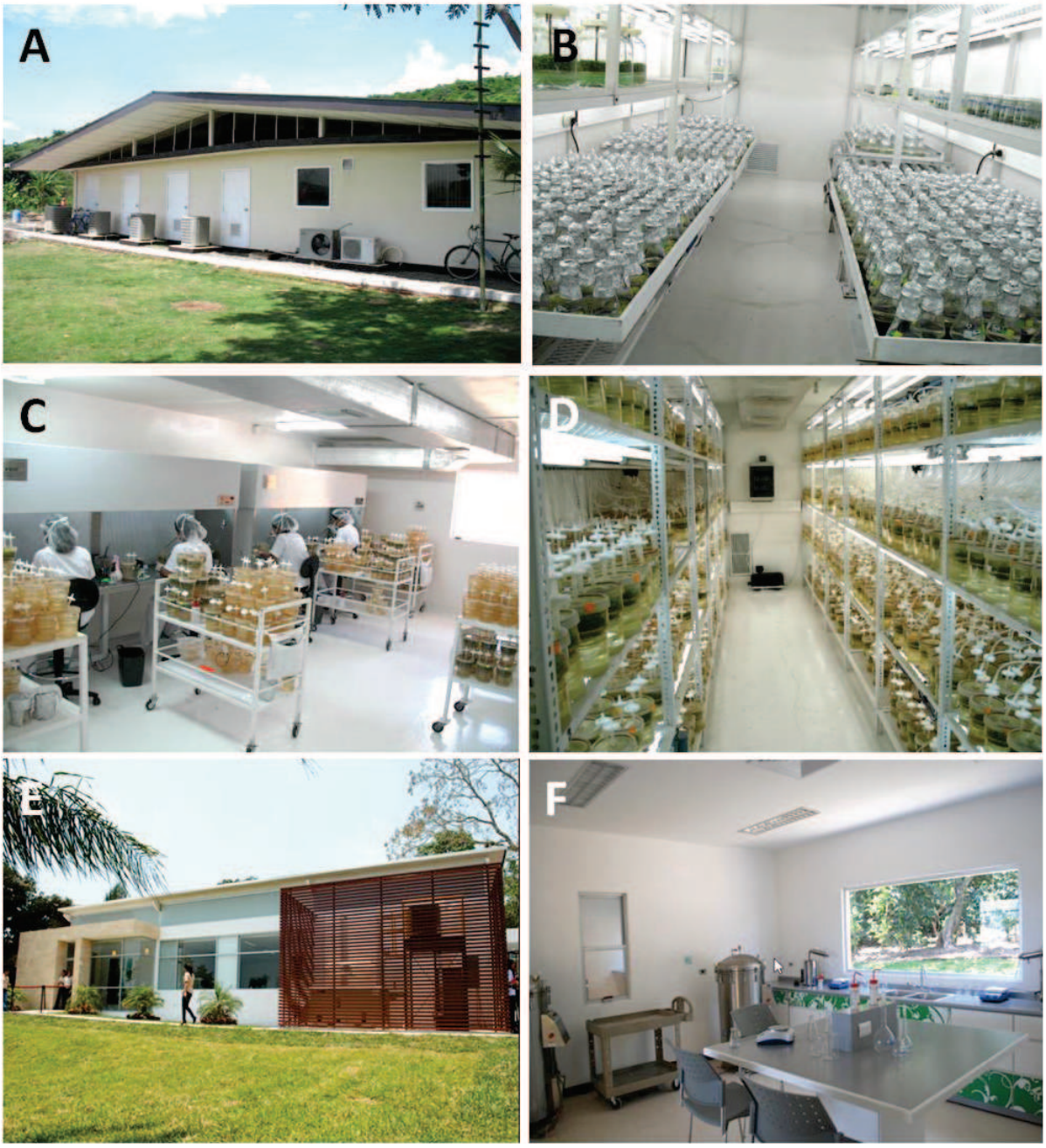


Figure 2. Aperçus des laboratoires de propagation *in vitro* du caféier de Sebacco au Nicaragua (A, vue extérieure ; B, salle des suspensions cellulaires ; C, salle de repiquage ; D, salle de culture en bioréacteurs) et de Jalapa au Mexique (E, vue extérieure ; F, salle de préparation des milieux).

De nombreuses difficultés sont rencontrées au cours de cette période qui empêchent la mise en place d'une production en routine. L'eau utilisée s'avère extrêmement dure (>350 ppm) et des dépôts de calcaire importants recouvrent les résistances des autoclaves et des distillateurs, les bioréacteurs mais également les feuilles des pieds-mères. Des pompes à injection d'acide seront ultérieurement mises en place pour neutraliser les ions calcium en régulant le pH des eaux d'arrosage. Les coupures d'électricité sont nombreuses et constantes au cours de l'année. Elles gênent considérablement

l'utilisation des tables orbitales en laboratoire sur lesquelles sont multipliées les différentes lignées cellulaires. Pour résoudre le problème, on équipera le laboratoire de deux groupes électrogènes (un petit groupe électrogène connecté exclusivement aux différentes tables orbitales permet de palier la déficience du groupe électrogène principal, lorsque le réseau de distribution d'électricité se trouve endommagé pendant plusieurs jours consécutifs). Sous ses latitudes, des niveaux de contaminations élevés sont enregistrés dans le laboratoire en période de saison des pluies. L'installation de filtres HEPA, couplée à des désinfections fréquentes et à des mesures de maintenance systématique des conduits d'aération, a permis de contrôler ce problème. Des autoclaves trop sophistiqués (nombreux composants électroniques) ont été achetées et chaque panne électrique ralentit l'activité du laboratoire car l'accès à des produits manufacturés - introuvables localement – retarde la réparation de ces appareils importés. L'impossibilité de trouver du personnel formé à la culture *in vitro* (cette formation n'existe pas dans les universités Nicaraguayennes) a imposé de former totalement le personnel aux différentes tâches nécessaires à l'activité de production. Enfin, les structures d'acclimatation des embryons somatiques, jusqu'alors placées à 30 kms du laboratoire, ont été finalement transférées à proximité pour un meilleur contrôle de cette étape très délicate qui nécessite à la fois une préparation technique importante, et un suivi rigoureux des procédures sanitaires, dans la continuité du travail réalisé *in vitro*.

Production industrielle et changement d'échelle (2007-2010)

Au cours de l'année 2007, la plupart des problèmes techniques cités ci-dessus sont réglés et une équipe de 25 personnes dont onze travaillant au laboratoire est formée et organisée. Pour chaque poste spécialisé (préparation des milieux, autoclavage, entretien des suspensions cellulaires, reporting, acclimatation..), plusieurs personnes sont formées pour palier à de possibles défections. L'activité de production démarre début 2007 et augmente régulièrement au cours des 3 années suivantes : 30.000 plantes sont vendues en 2007, 280.000 en 2008, 650.000 en 2009, 1.000.000 en 2010 et 1.500.000 plantes en 2011. A terme, l'objectif de production annuelle de ce laboratoire, sans modification ou supplément de moyens, est de 5 millions de plantes. Comme on va le voir par la suite, ce changement d'échelle peut être atteint par des optimisations techniques sur le procédé. Au cours de cette période, une dizaine de clones d'hybrides F1 ont été produits. A partir de cette production, un réseau de plusieurs centaines de milliers de plantes cultivées en conditions d'agroforesterie en Méso-Amérique et au Mexique a été installé. Les premières productions pré-industrielles sont l'occasion de tester chaque étape du procédé d'embryogenèse somatique, d'identifier des points de blocage et de réaliser des optimisations importantes. Nous détaillons par la suite cette expérience étape par étape.

Faisabilité au niveau industriel des différentes étapes du procédé

Les différentes étapes du procédé d'embryogenèse somatique sont schématisées dans la Figure 1. L'ensemble du clonage, allant de l'introduction des fragments de feuilles jusqu'à la production de plantes pouvant être transférées au champ, nécessite près de 2 ans dont 8 mois de pépinière. La production commerciale a commencé après vérification de la conformité génétique de tous les pieds-mères en utilisant des marqueurs moléculaires de type microsatellites (SSR). Trois arbres se sont avérés non conformes au génotype attendu et ont été écartés.

- **Production des tissus embryogènes.** Cette étape ne pose pas de problème chez l'Arabica sinon qu'elle est relativement longue (8-10 mois). La totalité des explants réagit en produisant un cal primaire cicatriciel (Fig. 3A) et entre 10 et 40% d'entre eux selon les génotypes produisent un cal secondaire embryogène (Fig. 3B). Ces fréquences sont suffisantes pour une production à grande échelle d'autant plus que les tissus embryogènes sont multipliés par la suite sous forme de suspensions d'agrégats cellulaires. Un effet génotypique existe mais il est facilement pris en compte en adaptant les quantités d'explants foliaires introduits.

- **Multiplication des tissus embryogènes et différenciation des embryons.** Ces deux étapes sont réalisées en milieu liquide en fioles d'Erlenmeyer sous agitation (Fig. 2B), ce qui permet de réduire drastiquement les besoins en main d'œuvre et en surface de laboratoire. A titre d'exemple, 4 millions d'embryons sont produits annuellement sur 4 m^2 de tables d'agitation. Elles ne posent aucun problème au niveau industriel mais requièrent un savoir-faire technique important par rapport aux autres étapes, en particulier pour l'établissement initial des suspensions. Ces étapes permettent également une synchronisation importante du développement du matériel végétal qui perdurera par la suite et permettra de réduire le travail associé au tri des embryons acclimatables. A chacune de ces étapes correspond ainsi un stade unique de développement, i.e. agrégats embryogènes puis embryons au stade torpille pleinement développés au terme de l'étape de différenciation (Fig. 3C).

- **Pré-germination des embryons somatiques en bioréacteurs.** Le changement d'échelle de cette étape a été réalisé avec succès au cours des 3 dernières années ; 4 millions d'embryons pré-germés ont été produits en 2010 en bioréacteurs à immersion temporaire d'un litre de type RITA® (Teisson et Alvard 1995) [Figs. 2C,D ; 3D]. La collecte de l'ensemble des embryons pré-germés (Fig. 3E) capables de poursuivre leur développement en plantules en pépinière nécessite 2 à 3 récoltes à partir de ces bioréacteurs. Des essais avec des bioréacteurs plus grands (3 litres) ont montré plusieurs avantages. Ainsi la réduction du nombre total de bioréacteurs permet un gain d'efficacité, c'est-à-dire de produire une plus grande quantité d'embryons acclimatables pour une même quantité de travail. Moins de bioréacteurs signifie

également moins d'investissement et moins de travail de nettoyage des différentes pièces constitutives. Par ailleurs, le brassage des embryons obtenu dans un volume plus grand est bien plus efficace et s'accompagne d'une meilleure synchronisation au cours des premières étapes de la germination. Cette optimisation se traduit directement par la possibilité de récolter en une seule fois la totalité des embryons transférables en pépinière.

- **Semis direct des embryons pré-germés en sol horticole et conversion en plantules.** Il s'agit de l'étape la plus délicate du procédé durant laquelle il faut permettre aux embryons de s'adapter aux conditions *ex vitro*, non stériles et plus sujettes à des variations en température et hygrométrie que ne le sont les conditions de laboratoire. Ce transfert est réalisé sous hygrométrie saturante sous tunnels plastique (Fig. 4A). Les embryons somatiques y sont cultivés à forte densité sur un substrat inerte à base de tourbe (Fig. 4B). Actuellement, cette étape constitue un goulot d'étranglement au niveau industriel puisque en moyenne, pour l'ensemble des 11 génotypes propagés, seulement 60% des embryons régénèrent des plantules (Fig. 4C). Un effet génotypique assez fort est observé entre les clones d'hybrides propagés. La durée moyenne nécessaire à la conversion en plantes après semis est relativement longue (22 à 24 semaines) en comparaison avec les plantules issues de semences (14 à 15 semaines). La conversion en plantes est asynchrone et deux ou trois récoltes successives sont nécessaires. Ces observations illustrent la marge de progression importante possible sur cette étape.

- **Grossissement en pépinière.** Les plantules présentant 2 à 3 paires de feuilles sont transférées dans des conditions de pépinière plus traditionnelles où elles sont 'endurcies' aux conditions extérieures en réduisant progressivement l'humidité relative et en augmentant l'intensité lumineuse (Figs. 4D, E, G, H). Cette étape est bien maîtrisée et ne pose aucune difficulté. Même si la croissance initiale des plants issus d'embryons somatiques des vitroplants est plus lente et hétérogène que celle de leurs homologues issus de graines, en fin de pépinière ils ont rattrapé leur retard et s'avèrent même plus vigoureux que ces derniers (Menéndez-Yuffa et al. 2010, Fig. 4H). Les pertes concernent environ 9% des plantes qui sont écartées au cours du contrôle qualité réalisé au terme de l'étape de pépinière. Pour l'essentiel, ces plantes présentent des défauts horticoles (manque de vigueur, tige courbée..) ; quelques plantes exhibant des modifications précoces du phénotype liées à des variations somaclonales [Angustifolia (Fig. 5A) et Variegata (Fig. 5B)] sont également écartées mais elles ne représentent que 0,3% de la totalité de la production. L'effet génotypique est faible ou inexistant au cours de l'étape de grossissement des plants. La principale difficulté pour le changement d'échelle de cette étape s'est située au niveau du choix du container. Des volumes trop grands imposaient l'utilisation de quantités de substrat horticole et des surfaces de pépinière très importantes, et compliquaient ultérieurement le transport des plants auprès des

producteurs, puis la plantation au champ, en particulier dans les zones montagneuses prépondérantes en Amérique centrale. Le choix d'un petit container de 200 ml appelé 'tubete' a été fait (Fig. 4F) étant donné la vocation du laboratoire de diffuser les vitroplants dans un substrat inerte, indemne de nématodes et de maladies du sol, sur l'ensemble de la zone Centroaméricaine (Fig. 5D).

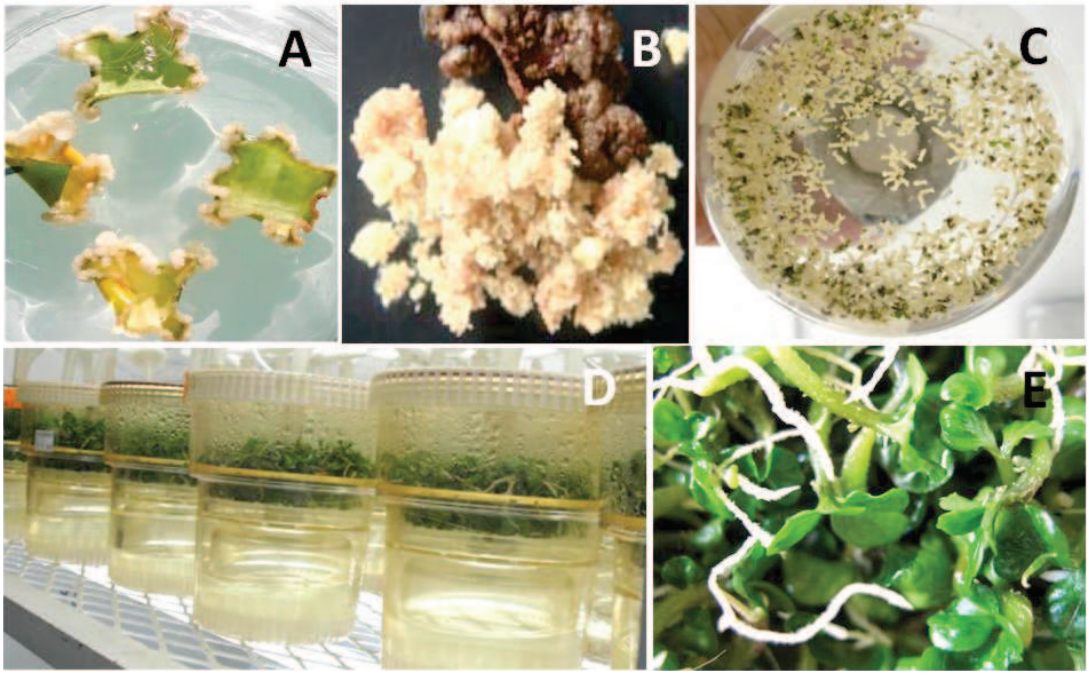


Figure 3. Différents stades de développement des tissus de caféier cultivés au cours des phases *in vitro* du procédé d'embryogenèse somatique. A, explant foliaire un mois après introduction *in vitro* ; B, cal embryogène 8 mois après introduction ; C, différenciation des embryons somatiques en fioles d'Erlenmeyer ; D, pré-germination des embryons en bioréacteurs à immersion temporaire de type RITA® ; E, embryons somatiques pré-germés prêts pour leur transfert *ex vitro*.

Conformité génétique des plantes issues d'embryogenèse somatique

L'obtention d'informations sur les variations somaclonales était l'une des attentes de ce transfert technologique. Nous n'avons jusqu'ici identifié chez *Coffea arabica* que des variations morphologiques (qualitatives) et avons montré que des plants phénotypiquement normaux poussaient et produisaient normalement. Une partie des variants peut ainsi être détectée et écartée précocement en pépinière ; c'est le cas des variants Angustifolia et Variegata. Comme on l'a vu, ces variants ne représentent que 0,3% de la production totale en pépinière. Les autres variants phénotypiques ne peuvent être détectés qu'au stade adulte, un ou deux ans après plantation. Les premières observations dans les parcelles commerciales du Nicaragua révèlent une fréquence d'environ 1% (comptage réalisé en champs sur environ 100 000 plants plantés en 2008), ce qui est tout à fait acceptable commercialement. Les variations observées concernent principalement des variants nain (84 % des variations observées), non détectables en pépinière d'acclimatation.

Afin de garantir à l'échelle industrielle la conformité génétique des plantes issues d'embryogenèse somatique, le laboratoire du Nicaragua a développé une véritable stratégie de « dilution du risque ». Outre le « screening » précoce des plants en pépinière au moment de l'expédition des plants, la stratégie en amont repose sur les points suivants : i) diminuer au maximum la présence de régulateurs de croissance (2,4D) dans les suspensions cellulaires, ii) mettre en place une traçabilité totale de la production par lot. Un lot de production représentant l'ensemble des plants issus de la descendance *in vitro* d'une même séquence d'introduction (même date d'introduction, même hybride, même pied mère), iii) mettre en place une traçabilité totale de la production par lignées au sein des lots de culture. Une lignée de production représentant l'ensemble des plants issus de la descendance *in vitro* d'une même lignée cellulaire (cal embryogène d'origine), iv) mélanger plusieurs lots et plusieurs lignées au moment de l'expédition des plants (une expédition de plants étant constituée d'au moins 5 lots de culture différents, lui-même constitué d'au moins 5 lignées cellulaires différentes, iii) et enfin, limiter le nombre de plants produits par lignée et par lot. Cette stratégie de 'dilution du risque' porte ses fruits puisque, on la vu précédemment, la fréquence de variations somaclonales dans les parcelles commerciales est faible (1% en moyenne) mais aussi parce qu'elle varie peu d'une parcelle à l'autre (entre 0,5 et 1,5% maximum).

Figure 5. A, variant panaché ('variegata') ; B, variant 'angustifolia' ; C, clone d'hybride F1 d'Arabica en fructification ; D, transport des vitroplants chez les producteurs ; E, clones d'hybrides F1 d'Arabica dans des systèmes agroforestiers.

Conclusions sur le transfert technologique

Nous avons acquis une expérience importante au cours de ce transfert technologique et observé qu'il était inévitable de progresser par essais et erreurs tant il est impossible d'anticiper la totalité des problèmes, particulièrement dans un pays en voie de développement. Cette expérience a montré la nécessité en cas de difficulté d'analyser et d'intervenir très rapidement afin d'éviter des pertes de production vite dramatiques, et pour ce faire la nécessité d'avoir établi une cohésion et une communication sans faille au sein de l'équipe constituée avec le partenaire. Le transfert de technologie a été le lieu d'une rencontre de partenaires venant de deux mondes différents, l'industrie/privé et la recherche/public. Pour arriver à un partenariat efficace basé sur des complémentarités équilibrées, il a fallu former une équipe mixte multidisciplinaire, ce qui a pris un peu de temps. Cette action en partenariat est en tout cas exemplaire car elle démontre que la synergie public/privé peut sous certaines conditions être efficacement au service d'une politique en faveur d'une agriculture durable. En effet, les hybrides F1 sont susceptibles de servir de catalyseur au retour des planteurs vers l'agroforesterie (Fig. 5E), qui pour des raisons de productivité se sont tournés dans les années 80 vers les pratiques peu écologiques de la révolution verte.

Figure 4. Aperçus des pépinières de vitroplants de caféier au Nicaragua. A et B, vue des tunnels d'acclimatation ; C, aspect des plantules obtenues par semis des embryons pré-germés ; D , transfert des plantules en 'tubetes' de 200 ml (F) et endurcissement (E) ; G et H, vues des pépinières de grossissement (ferme 'La Cumplida, Matagalpa, Nicaragua).

Le transfert technologique de l'embryogenèse somatique est achevé chez l'espèce Arabica et démontre la faisabilité d'une propagation de masse par embryogenèse somatique. Un changement d'échelle conséquent a été possible pour chaque étape du procédé rendant possible un flux de production continu du laboratoire vers la pépinière et par conséquent une identification précise des forces et faiblesses sur l'itinéraire technique choisi. L'activité commerciale a débuté à partir de 2008 et le cap du million de plantes vendues aux producteurs centroaméricains a été atteint en 2010. Par ailleurs, les hybrides F1 (Fig. 5C) ont confirmé leur supériorité sur les lignées traditionnelles et génèrent un tel engouement que la demande est maintenant supérieure à la capacité de production (2 millions de

plantes pour le seul Nicaragua). L'objectif du partenariat est donc de répondre au plus vite à la demande par un changement d'échelle important de l'unité de production du Nicaragua (5 à 6 millions de plantes d'ici 4 ans) et mais aussi par le montage d'autres unités dans la région. En effet, un prototype est à présent disponible pour la partie laboratoire comme pour la partie pépinière qu'il est possible de « photocopier » sur d'autres sites. Une autre unité de production est déjà opérationnelle dans l'état de Veracruz au Mexique pour propager d'autres hybrides (Figs. 2E, F); par ailleurs plusieurs pépinières industrielles ont été mises en place au Mexique, au Guatemala et au Salvador.

Enfin, la démonstration de la faisabilité d'une propagation à grande échelle par embryogenèse somatique chez *C. arabica* renouvelle le champ des possibles dans le domaine de la sélection génétique chez cette espèce. En effet, la réussite de ce transfert technologique permet dorénavant d'envisager l'introduction sur le marché de nouvelles variétés à partir de matériels hybrides ou mutants.

Références

Albarran J, Bertrand B, Lartaud M, Etienne H. Cycle characteristics in a temporary immersion bioreactor affect regeneration, morphology, water and mineral status of coffee (*Coffea arabica* L.) somatic embryos. *Plant Cell Tiss Org Cult* 2005 ; 81:27-36.

Barry-Etienne D, Bertrand B, Schlongvoïgt A, Etienne H. The morphological variability within a population of coffee somatic embryos produced in a bioreactor affects the regeneration and the development of plants in the nursery. *Plant Cell Tiss Org Cult* 2002a ; 68:153-62.

Barry-Etienne D, Bertrand B, Vásquez N, Etienne H. Direct sowing of *Coffea arabica* somatic embryos mass-produced in a bioreactor and regeneration of plants. *Plant Cell Rep* 1999 ; 19:111-17.

Barry-Etienne D, Bertrand B, Vásquez N, Etienne H. Comparison of somatic embryogenesis-derived coffee (*Coffea arabica* L.) plantlets regenerated *in vitro* or *ex vitro* conditions : morphological, mineral and water characteristics. *Ann Bot* 2002b ; 90:77-85.

Bertrand B, Alpizar E, Llara L *et al.* Performance of Arabica F1 hybrids in agroforestry and full-sun cropping systems in comparison with pure lines varieties. *Euphytica* 2011 sous presse, DOI 10.1007/s10681-011-0372-7.

Bertrand B, Etienne H, Cilas C, Charrier A, Baradat P (2005) *Coffea arabica* hybrid performance for yield, fertility and bean weight. *Euphytica*. 141:255-62.

Ducos JP, Gibault E, Lambot C (2010) Coffee propagation by somatic embryogenesis at Nestle R&D Center-Tours. In: Proceedings of the Conference on Advances in Somatic Embryogenesis of Trees and its application for the future forests and plantations, IUFRO-KFRI 2010 conference, August19-21, 2010, Suwon, Korea.

Etienne H, Bertrand B. The effect of the embryogenic cell suspension micropropagation technique on the trueness to type, field performance, bean biochemical content and cup quality of *Coffea arabica* trees. *Tree Physiol* 2001 ; 21:1031-38.

Etienne H, Bertrand B. Somaclonal variation in *Coffea arabica* : effects of genotype and embryogenic cell suspension age on frequency and phenotype of variants. *Tree Physiol* 2003 ; 23:419-26.

Gupta PK, Hartle JE (2010) Scale-up somatic embryogenesis of conifers for reforestation. In: Proceedings of the Conference on Advances in Somatic Embryogenesis of Trees and its application for the future forests and plantations, IUFRO-KFRI 2010 conference, August19-21, 2010, Suwon, Korea.

Karp A (1994) Origin, causes and uses of variation in plant tissue culture. In: I. Vasil and T. Thorpe, Editors, *Plant Cell and Tissue Culture*, Kluwer Academic Publishers, Dordrecht (1994), pp. 139–151.

Khaw CH, Ng SK, Thong KC (1999) Commercial production of clonal palms by tissue culture – prerequisites, constraints and issues. In: Proceedings of the 1999 PORIM International Palm Oil Congress – Agriculture. Emerging Technologies and Opportunities in the Next Millenium. Palm Oil Research Institute of Malaysia (PORIM), Kuala Lumpur, Malaysia, 1-6 February 1999.

Ménendez-Yuffá A, Barry-Etienne D, Bertrand B, Georget F, Etienne H. A comparative analysis of the development and quality of nursery plants derived from somatic embryogenesis and from seedlings for large-scale propagation of coffee (*Coffea arabica* L.). *Plant Cell Tiss Org Cult* 2010 ; 102:297-307.

Sampote S, Ducos J-P, Lambot C, Kunasol T, Kasinkasaempong Y, Chantanumat P, Pétiard V (2006) First industrial massive propagation of *Coffea canephora* through somatic embryogenesis organized in Thailand. In: Proceedings of the 21st International Conference on Coffee Science, Montpellier (France), 11th - 15th September 2006.

Teisson C, Alvard D. A new concept of plant *in vitro* cultivation liquid medium : temporary immersion. In : M Terzi et al. (eds) *Current Issues in Plant Molecular and Cellular Biology* (pp 105-110). Kluwer Academic Publishers, Netherlands, 1995.

Influence of micropropagation through somatic embryogenesis on somaclonal variation in coffee (*Coffea arabica*) : assessment of variations at the phenotypical, cytological, genetic and epigenetic level

Somaclonal variation (SV) is a major concern in all micropropagation systems. It is described as the phenotypic variation displayed in *in vitro*-derived regenerants and it is believed to be originated from a large array of genetic and epigenetic mechanisms. Highly productive *Coffea arabica* hybrids are clonally disseminated in Meso-American region through somatic embryogenesis (SE). The objective of the present work in coffee is to evaluate the trueness-to-type of SE, to understand better the mechanisms involved in SV and further optimize SE conditions. We assessed the variations in the propagated plants at the phenotypic, cytogenetic (chromosome counting), genetic (mutations/AFLP, activation of transposable elements/S-SAP) and epigenetic (methylation/MSAP) level by using two complementary approaches. First, with 2 hybrids we studied industrial culture conditions expected to be weakly mutagenic thanks to the combined use of short term proliferation period (6 months) and low auxin supply (0-1.4 µM 2,4-D). Two proliferation systems i.e. secondary embryogenesis and embryogenic suspensions were compared, the latter being more productive and economic. AFLP and MSAP molecular analyses on 145 somatic seedlings showed that genetic and epigenetic polymorphisms between mother plants and emblings were extremely low, i.e. ranges of 0–0.003% and 0.07–0.18% respectively, with no significant difference between the proliferation systems. No plant was found to cumulate more than 3 methylation polymorphisms. For the two hybrids tested, massive phenotypic observations in nursery and field plots showed very low levels of SV (0.9% from 800,000 plants). Cytological analysis showed abnormal chromosome numbers (41-43, 45) in most of coffee somaclonal variants and normal numbers (44) in phenotypically normal plants. Stressful experimental conditions were also applied by using extended proliferation periods (4, 12 and 27 months) for three independent embryogenic lines established for the Caturra var. in presence of high growth regulator concentrations (4.5 µM 2,4-D, 17.8 µM 6-BA) to understand the mechanisms of culture ageing on SV. The proliferation time strongly affected the SV frequency among the 180 regenerated plants and in a highly similar way with the three embryogenic lines. No variant was found after 4 months proliferation although 30% and 94% phenotypic variants were observed in plants derived from 12 and 27 month-old cultures, respectively. Regardless the culture age and the embryogenic line, no polymorphisms were found in the 124 plants analyzed neither with AFLP nor with S-SAP using 13 different transposable elements from several families, and very limited polymorphisms were found in the methylation patterns using MSAP markers (0.049-0.087%). However, similarly to plants derived from under industrial conditions, phenotypic variants systematically showed abnormal chromosome numbers and normal plants systematically showed normal numbers. This work showed that SE based on embryogenic suspensions is reliable for true-to-type propagation of selected *C. arabica* varieties. It also demonstrated the importance of culture age on SV and hence the non random nature of this phenomenon. The genetic and epigenetic alterations are particularly limited during SE. The main change in most of phenotypic variants was aneuploidy showing that mitotic aberrations play a major role in SV in coffee.

Key-words: aneuploidy, coffee, epigenetic, micropropagation, methylation, molecular markers, somaclonal variation, somatic embryogenesis, transposable elements

Influence de la micro-propagation par embryogenèse somatique sur la variation somaclonale chez le caféier (*Coffea arabica*) : évaluation des changements au niveau phénotypique, cytologique, génétique et épigénique

La variation somaclonale (VS) est une préoccupation majeure de tous les systèmes de micropropagation. Elle est décrite comme un changement phénotypique présent chez les vitroplants et pourrait être générée par une large gamme de mécanismes génétique et épigénétiques. Des hybrides de *Coffea arabica* hautement productifs sont distribuées sous forme clonale par embryogenèse somatique (ES) en Meso-Amérique. L'objectif de ce travail chez le caféier est d'évaluer la conformité génétique des plants multipliés par ES (somaplants), de mieux comprendre les mécanismes impliqués dans les SV et finalement d'optimiser les conditions de culture. Nous avons évalué les variations dans les plantes régénérées au niveau phénotypique, cytologique (comptage de chromosomes), génétique (mutations/marqueurs AFLPs, réactivation d'éléments transposables/ marqueurs S-SAP) et épigénétique (méthylation/marqueurs MSAP) en utilisant deux approches complémentaires. Tout d'abord, nous avons étudié chez deux hybrides des conditions de cultures industrielles supposées peu mutagènes grâce à l'utilisation combinée d'une courte période de prolifération (6 mois) et d'un faible apporte en auxine (0-1.4 µM 2,4-D). Deux systèmes de prolifération i.e. l'embryogenèse secondaire et les suspensions embryogènes seront comparés, le dernier étant plus productif et économique. Les analyses moléculaires AFLP et MSAP sur 145 somaplants montrent que les polymorphismes génétique et épigénétique entre plantes mères et somaplants sont extrêmement réduits, i.e. dans l'intervalle 0-0,003% et 0,07-0,18% respectivement, sans différence significative entre les systèmes de prolifération. Aucun somaplant ne cumule plus de 3 polymorphismes de méthylation. Pour les deux hybrides testés, des observations phénotypiques massives en pépinière et au champ ont révélé de très faibles niveaux de VS (0,9% pour 800.000 plantes). Des analyses cytologiques ont mis en évidence des nombres de chromosomes anormaux (41-43, 45) chez la plupart des variants et des nombres normaux (44) chez les plants ayant un phénotype normal. Des conditions expérimentales *a priori* mutagènes ont également été appliquées en utilisant des périodes de prolifération prolongées (4, 12 et 27 mois) chez trois lignées embryogènes indépendantes de la variété Caturra en présence de concentrations élevées en régulateurs de croissance (4.5 µM 2,4-D, 17.8 µM 6-BA), afin de comprendre les mécanismes liés vieillissement des cultures interviennent sur les VS. L'étude des 180 somaplants régénérés a montré que le temps de prolifération affecte fortement la fréquence de VS et d'une manière hautement similaire pour les 3 lignées embryogènes. Aucun variant n'a été trouvé après 4 mois de prolifération alors que 30% et 94% de variants phénotypiques ont été caractérisés chez les plants issus de cultures de 12 et 27 mois, respectivement. Quels que soient l'âge de culture et la lignée embryogène, aucun polymorphisme n'a été trouvé chez les 124 somaplants analysés que ce soit avec les marqueurs AFLP, ou même S-SAP en étudiant 13 éléments transposables appartenant à plusieurs familles, et des polymorphismes très faibles ont été trouvés dans les profils de méthylation avec les MSAP (0,049-0,087%). Cependant, de façon similaire aux somaplants produits en conditions industrielles, les variants phénotypiques montrent systématiquement des nombres de chromosomes anormaux (41-43) et les plants normaux le nombre de chromosomes attendu. Ce travail montre que l'ES s'appuyant sur des suspensions embryogènes peut garantir une propagation conforme des variétés sélectionnées de *C. arabica*. Il démontre également l'importance de l'âge des cultures sur l'apparition de VS et donc le caractère non aléatoire du phénomène. Les changements génétiques et épigénétiques sont particulièrement limités durant l'ES. Le principal changement chez la plupart des variants est l'aneuploidie, ce qui montre que les aberrations mitotiques jouent un rôle majeur dans les VS chez le caféier.

Key-words: aneuploïdie, caféier, éléments transposables, embryogenèse somatique, épigénétique, marqueurs moléculaires, méthylation, micropropagation, variation somaclonale