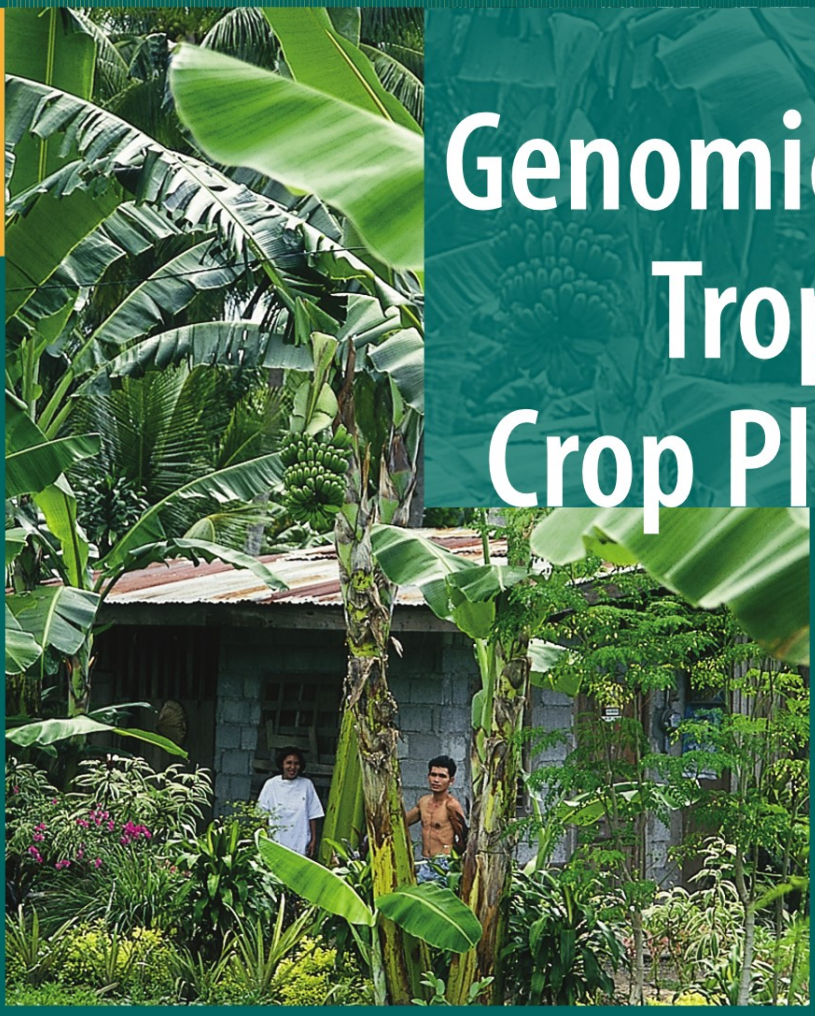
Paul H. Moore
Ray Ming
Editors

Plant Genetics / Genomics *Volume 1*



Genomics of Tropical Crop Plants

Springer

Paul H. Moore · Ray Ming
Editors

Genomics of Tropical Crop Plants

Foreword by Deborah Delmer

Springer

Editors

Paul H. Moore
USDA-ARS, PBARC
Hawaii Agriculture Research Center
99-193 Aiea Heights Drive
Aiea 96701
phmoore@hawaii.edu
p.moore@harc-hspa.com

Ray Ming
Department of Plant Biology
University of lllinois at Urbana-Champaign
Champaign, IL, USA
1201 W. Gregory Drive
Urbana 61801
288 ERML, MC-051
rming@life.uiuc.edu

ISBN: 978-0-387-71218-5 e-ISBN: 978-0-387-71219-2

Library of Congress Control Number: 2007942800

Cover illustration: Small holders in tropics: Photograph by Gisela Orjeda, Bioversity International. Most of bananas grown in the tropics are produced by smallholders for home consumption or selling in local markets.

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Chapter 9
Genomics of Coffee, One of the World's Largest Traded Commodities

Philippe Lashermes, Alan Carvalho Andrade, and Hervé Etienne

Abstract Coffee is one of the world's most valuable agricultural export commodities. In particular, coffee is a key export and cash crop in numerous tropical and subtropical countries having a generally favorable impact on the social and physical environment. While coffee species belong to the Rubiaceae family, one of the largest tropical angiosperm families, commercial production relies mainly on two species, *Coffea arabica* L. and *Coffea canephora* Pierre, known as Robusta. Although a considerable genetic diversity is potentially available, coffee breeding is still a long and difficult process. Nevertheless, genomic approaches offer feasible strategies to decipher the genetic and molecular bases of important biological traits in coffee tree species that are relevant to the growers, processors, and consumers. This knowledge is fundamental to allow efficient use and preservation of coffee genetic resources for the development of improved cultivars in terms of quality and reduced economic and environmental costs. This review focuses on the recent progress of coffee genomics in relation to crop improvement.

9.1 Introduction

Coffee species belong to the Rubiaceae family, one of the largest tropical angiosperm families. Variations in cpDNA classified the Coffeeae tribe into the Ixoroideae monophyletic subfamily, close to Gardenieae, Pavetteae and Vanguerieae (Bremer and Jansen 1991). Two genera, *Coffea* L. and *Psilanthus* Hook. f., were distinguished on the basis of flowering and flower criteria (Bridson 1982). Each genus was divided into two subgenera based on growth habit (monopodial *vs.* sympodial development) and type of inflorescence (axillary *vs.* terminal flowers). More than one hundred coffee species have been identified and new taxa are still being discovered (Bridson and Verdcourt 1988; Davis et al. 2006; Stoffelen et al. 2007). All species are perennial woody bushes or trees in inter-tropical forests of Africa

P. Lashermes
Institut de Recherche pour le Développement, UMR RPB, BP 64501,
34394 MontpellierCedex 5, France
e-mail: Philippe.Lashermes@mpl.ird.fr

P.H. Moore, R. Ming (eds.), *Genomics of Tropical Crop Plants*

and Madagascar for the *Coffea* genus, and Africa, Southeast Asia, and Oceania for the *Psilanthus* genus. Plants of the two genera differ greatly in morphology, size, and ecological adaptations. Some species, such as *C. canephora* and *C. liberica* Hiern, are widely distributed from Guinea to Uganda. Other species display specific adaptations, e.g., *C. congensis* Froehner to seasonally flooded areas in the Zaire basin and *C. racemosa* Lour. to very dry areas in the coastal region of Mozambique. Molecular phylogeny of *Coffea* species has been established based on DNA sequence data (Lashermes et al. 1997; Cros et al. 1998). The results suggest a radial mode of speciation and a recent origin in Africa for the genus *Coffea*. Several major clades were identified, which present a strong geographical correspondence (i.e., West Africa, Central Africa, East Africa, and Madagascar).

Commercial production relies mainly on two species, *C. arabica* L. and *C. canephora* Pierre, known as robusta. The cup quality (low caffeine content and fine aroma) of *C. arabica* makes it by far the most important species, representing 65% of the world production. Another species, *C. liberica* (liberica coffee) stands third with a share of less than one per cent of world coffee production.

C. arabica has its primary center of genetic diversity in the highlands of Southwest Ethiopia and the Boma Plateau of Sudan. Wild populations of *C. arabica* also have been reported in Mount Imatong (Sudan) and Mount Marsabit (Kenya) (Thomas 1942; Anthony et al. 1987). Cultivation of *C. arabica* started in Southwest Ethiopia about 1,500 years ago (Wellman 1961). Modern coffee cultivars are derived from two base populations of *C. arabica*, known as Typica and Bourbon, that were spread worldwide in the 18th century (Krug et al. 1939). Historical data indicate that these populations were composed of progenies of very few plants, i.e., only one for the Typica population (Chevalier and Dagron 1928) and the few plants that were introduced to the Bourbon Island (now Réunion) in 1715 and 1718 for the Bourbon population (Haarer 1956).

During the 18th and 19th centuries, only Arabica was produced. However, this species appeared to be very sensitive to parasitic threats, especially orange rust. That is why, in Africa, during the 19th century, the spontaneous forms of other species of coffee, especially *C. canephora*, were cultivated locally. In particular, coffee plants from local forest populations of the Belgian Congo (now the Democratic Republic of Congo) and Uganda were transferred to Java, a major breeding center from 1900 to 1930. At the same time, in Africa, the diversity of material cultivated was extended with the use of local spontaneous forms: Kouilou in Côte d'Ivoire, Niaouli in Togo and Benin, and Nana in the Central African Republic. The material selected in Java was reintroduced in the Belgian Congo around 1916 at INEAC (Institut National pour l'Etude Agronomique au Congo) research center which was the major breeding center of *C. canephora* from 1930 to 1960. Selected plant materials were largely distributed worldwide. Although the overall performance of cultivated trees has increased noticeably after a few breeding cycles, the cultivars nonetheless have remained genetically very close to individuals of the original natural populations (Dussert et al. 2003). Furthermore, the considerable genetic diversity observed in *C. canephora* is still largely unexploited.

Arabica production is constrained by numerous diseases and pests like leaf rust (*Hemileia vastatrix* Berk & Br.), coffee berry disease (*Colletotrichum kahawae*), coffee berry borer (*Hypothenemus hampei*), stem borer (*Xylotrechus quadripes* Chevr.), and nematodes (*Meloidogyne spp.* and *Pratylenchus spp.*). In contrast, Robusta is more tolerant to these diseases and pests. Hence, transfer of desirable genes, in particular for disease resistance, from coffee species into 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 (*H. vastatrix*), coffee berry disease (*C. kahawae*), and root-knot nematode (*Meloidogyne* spp.). Likewise, other coffee species are of considerable interest in this respect. For instance, *C. liberica* has been used as source of resistance to leaf rust (Srinivasan and Narasimhaswamy 1975), while *C. racemosa* constitutes a promising source of resistance to the coffee leaf miner (Guerreiro Filho et al. 1999). Exploitation of such 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. Undesirable genes from the donor parent are gradually eliminated by selection. In so doing, conventional coffee breeding methodology faces considerable difficulties. In particular, strong limitations are due to the long generation time of a coffee tree (5 years), the high cost of field trials, 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 thereby ensure good quality of the improved variety.

9.1.1 Economic and Societal Importance of Coffee

Coffee is one of the world's most valuable agricultural export commodities. In particular, coffee represents one of the key export and cash crops in tropical and subtropical countries with generally a favorable impact on the social and physical environment (International Coffee Organization, http://www.ico.org/). About 125 million people depend on coffee for their livelihoods in Latin America, Africa, and Asia. Coffee is produced in more than 68 countries between 22° N and 24° S latitudes, on a total of about 10.6 million ha. Total world production for the year 2005/06 was 6.6 million t green coffee beans, with an estimated market value of $12.2 billion. World coffee production over the last 5 years has fluctuated from 6.3 to 7.4 million t green beans.

Arabica coffee is cultivated in relatively cool mountain climates, at 1,000-2,000 m altitudes in equatorial regions and at 400 -1,200 m altitudes further from the equator. Robusta coffee requires the warm and humid climates of tropical lowlands and foothills. About 63% of the world coffees in 2004/05 were produced in Latin America, 24% in Asia, and 13% in Africa. Brazil (80% arabica) alone produced 35%, Vietnam (97% robusta) 12%, Colombia (100% arabica) 10%, and Ethiopia (100% arabica) 4% of all coffee. Smallholders (< 5 ha) account for about 70% of

world coffee production. However, medium to large coffee plantations (30-3,000 ha per estate) can be found in countries like Brazil, El Salvador, Guatemala, India, Indonesia, Vietnam, Kenya, Tanzania, and Ivory Coast.

World coffee consumption was 6.9 million t for 2004/05 (6.7 million t the previous year), of which about 1.8 million t (26%) is consumed domestically in coffee-exporting countries. Brazil, Mexico, and Ethiopia are large coffee exporters with notably high domestic coffee consumption (>40% of production). The European Community, Norway, Switzerland, the United States, and Japan consume about 80% of the coffee exported from producing countries. The demand for coffee has been growing steadily at 1.2% per year, but world production regularly exceeds demand by 0.3-0.6 million t per year. All these factors contribute to a high volatility of price on the main world coffee markets.

9.1.2 Coffea *as an Experimental Organism*

The coffee plant is an evergreen shrub or small tree, which, under free growth, may become 4-6 m tall for *C. arabica* and 8-12 m for *C. canephora*. In cultivation, both species are pruned to manageable heights of less than 2-3 m with one or more stems. The growth of the coffee plant is dimorphic. The main stems (orthotropic axes) grow vertically and the branches (plagiotropic axes) grow horizontally. Horticultural propagation is relatively easy. The plant may flower once or twice a year after a rainfall of at least 10 mm, which follows a period of water stress. Fruits mature in 7-9 (*C. arabica*) to 9-11 (*C. canephora*) months from date of flowering depending on the variety and environment. The coffee seeds do not behave in an orthodox manner when dehydrated or stored at low temperature. For instance, the viability of seeds of *C. arabica* decreases rapidly after 4-6 months at ambient temperature. However, short term storage (up to 3 years) is possible using adapted and controlled storage conditions (Eira et al. 2006).

All coffee species are diploid (2n=2x=22) and generally self-incompatible, except for *C. arabica* Linne which is tetraploid (2n=4x=44) and self-fertile. Molecular analyses (Lashermes et al. 1999) have indicated that *C. arabica* is an amphidiploid formed from the hybridization between two closely related diploid species (i.e., *C. canephora* and *C. eugenioides*). The evidence suggests recent speciation and low divergence between the two constitutive genomes of *C. arabica* and those of its progenitor species. The analysis of segregating molecular markers has confirmed earlier genetic and cytogenetic evidence that *C. arabica* is a functional diploid. Furthermore, homoeologous 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).

The nuclear DNA content of several coffee species has been estimated by flow cytometry (Cros et al. 1995). The DNA amount (2C values) varies between diploid coffee species from 0.95 to 1.8 pg. In comparison to other angiosperms (Bennett and Leitch 1995), the genomes of coffee species appear to be of rather low size (i.e.,

810 Mb for *C. canephora* and 1,300 Mb for *C. arabica*). These variations in DNA amount, other than variation due to 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. Furthermore, reduced fertility of certain interspecific F_1 hybrids appears to be associated with significant differences in nuclear content of parental species (Barre et al. 1998).

9.2 Genetic Mapping and Tagging in Coffee

9.2.1 Marker Diversity

In coffee, a whole range of techniques has been used to detect polymorphism at the DNA level, including randomly amplified polymorphic DNA (RAPD) (Orozco-Castilho et al. 1994; Lashermes et al. 1996a; Anthony et al. 2001; Aga et al. 2003), cleaved amplified polymorphisms (CAP) (Lashermes et al. 1996b; Orozco-Castilho et al. 1996), restriction fragment length polymorphisms (RFLP) (Paillard et al. 1996; Dussert et al. 2003; Lashermes et al. 1999), amplified fragment length polymorphism (AFLP) (Lashermes et al. 2000b; Anthony et al. 2002; Prakash et al. 2005), inverse sequence-tagged repeat (ISR) (Aga and Bryngelsson 2006), and simple sequence repeats or microsatellites (SSR) (Mettulio et al. 1999; Combes et al. 2000; Baruah et al. 2003; Moncada and McCouch 2004; Poncet et al. 2004). During the last few years, the number of co-dominant markers has been considerably increased by SSR mining in coffee expressed sequence tag (EST) databases (Bhat et al. 2005; Poncet et al. 2006; Aggarwal et al. 2007) offering new possibilities for genetic analysis.

9.2.2 Linkage Mapping

Several genetic maps have been constructed. The low polymorphism has been a major drawback for developing genetic maps of the *C. arabica* genome. Hence, the works reported so far are often restricted to alien DNA introgressed fragments into *C. arabica* (Prakash et al. 2004). Nevertheless, Pearl et al. (2004) recently obtained a genetic map from a cross between Catimor and Mokka cultivars of *C. arabica*. Furthermore, to overcome the limitation of low polymorphism, efforts were directed to the development of genetic maps in *C. canephora* or interspecific crosses.

The earliest attempt to develop a linkage map was based on Canephora doubled haploid (DH) segregating populations (Paillard et al. 1996; Lashermes et al. 2001). *C. canephora* is a strictly allogamous species consisting of polymorphic populations and of strongly heterozygous individuals. Conventional segregating populations are therefore somewhat difficult to generate and analyze. However, the ability to produce DH populations in *C. canephora* offers an attractive alternative approach. The method of DH production is based on the rescue of haploid embryos of maternal

origin occurring spontaneously in association with polyembryony (Couturon 1982). Two complementary segregating plant populations of *C. canephora* were produced from the same genotype. One population comprised 92 doubled haploids derived from female gametes, while the other population was a test-cross consisting of 44 individuals derived from male gametes. A genetic linkage map of *C. canephora* was constructed spanning 1,041 cM of the genome (Lashermes et al. 2001). This genetic linkage map comprised more than 40 specific STS markers, either single-copy RFLP probes or SSRs that are distributed on the eleven linkage groups. These markers constituted an initial set of standard landmarks of the coffee genome which have been used as anchor points for map comparison (Herrera et al. 2002) and coverage analysis of bacterial artificial chromosome (BAC) libraries (Leroy et al. 2005, Noir et al. 2004). Furthermore, the recombination frequencies in both populations were found to be almost indistinguishable. These results offer evidence in favor of the lack of significant sex differences in recombination of *C. canephora*.

More recently, Crouzillat et al. (2004) reported the development of a Canephora consensus genetic map based on a segregating population of 93 individuals from the cross of two highly heterozygous parents. Backcross genetic maps were established for each parent (i.e., elite clones BP409 and Q121) and then a consensus map was elaborated. More than 453 molecular markers such as RFLP and SSRs were mapped covering a genome of 1,258 cM. Recently, this map was used to map COS (i.e. Conserved Orthologous sequence) markers and perform comparative mapping between coffee and tomato (Wu et al. 2006).

In parallel several diploid interspecific maps were built. Those maps are based on either an F_1 hybrid population resulting from a cross between coffee diploid species (López and Moncada 2006) or progenies obtained by backcrossing of hybrid plants to one of the parental species (Ky et al. 2000; Coulibaly et al. 2003).

9.3 Molecular Cytogenetics and BAC Cloning

9.3.1 Fluorescence in situ Hybridization

The basic chromosome number for the genus *Coffea* is considered to be $n = 11$, which is typical for most genera of the family Rubiaceae. Coffee somatic chromosomes are relatively small (1.5 to 3 μm) and morphologically similar to each other (Krug and Mendes 1940; Bouharmont 1959). Observations at the meiotic pachytene phase provided a significantly better chromosomal characterization and allowed the identification of most bivalents of *C. arabica* (Pinto-Maglio and Cruz 1998). Characterization of the longitudinal differentiation of the mitotic chromosomes of coffee has progressed using other techniques such as fluorescent banding (Lombello and Pinto-Maglio 2004). In addition, the fluorescence in situ hybridization (FISH) method that opened up new perspectives. In particular, genomic in situ hybridization (GISH) was successfully applied for characterization of genomes and chromosomes in polyploid, hybrid plants, and recombinant breeding lines. 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 (Raina et al. 1998; Lashermes et al. 1999). Furthermore, FISH was used to study the presence of alien chromatin in interspecific hybrids and plants derived from interspecific hybrids between coffee species (Barre et al. 1998; Herrera et al. 2007). More recently, the BAC-FISH procedure was used to rapidly localize a given introgression on a specific chromosome (Herrera et al. 2007).

9.3.2 Development of BAC Libraries

BAC libraries have been reported for both cultivated coffee species, *C. arabica* and *C. canephora*.

An Arabica BAC library using the cultivar IAPAR 59 was successfully constructed and validated (Noir et al. 2004). This introgressed variety, derived from the Timor Hybrid, was selected for resistance to leaf rust and root-knot nematodes, and it is widely distributed in Latin America. The library contains 88,813 clones with an average insert size of 130 kb and represents approximately 8 *C. arabica* dihaploid genome equivalents.

A Canephora BAC library was developed on a relatively good cup quality genotype (i.e., clone 126). The library contains 55,296 clones, with an average insert size of 135 kb, representing almost nine haploid genome equivalents. This resource was used to analyze the genome organization (copy number) of sucrose-metabolizing enzymes (mainly sucrose synthase and invertases) in the *C. canephora* genome (Leroy et al. 2005).

9.3.3 BAC Physical Mapping

Although the integration of physical and genetic mapping information would be particularly useful in coffee genome research, only few activities have been reported so far. Regional physical maps based on BAC contigs corresponding to agronomical important disease resistance genes were developed (Lashermes et al. 2004). Recently, a BAC contig linked to the S_H3 leaf rust resistance gene was used to assess microsynteny between coffee (*C. arabica* L.) and *Arabidopsis thaliana* (Mahé et al. 2007). Microsynteny was revealed and the matching counterparts to *C. arabica* contigs were seen to be scattered throughout four different syntenic segments of *Arabidopsis* on chromosomes (*Ath*) I, III, IV, and V.

9.4 EST Resources and Transcriptome Analysis

Large-scale sequencing of cDNAs to produce ESTs followed by comparisons of the resulting sequences with public databases has become the method of choice for the

rapid and cost-effective generation of data, becoming the fastest growing segment of public DNA databases. Hence, ESTs, DNA arrays, large-scale gene expression (transcriptome) profiling and associated bioinformatics are becoming routine in the plant sciences. However, only recently have efforts been dedicated to coffee.

9.4.1 Coffee EST Resources

Currently, there are nearly 43 million ESTs (dbEST release 031607, March 2007) in the NCBI public collection (http://www.ncbi.nlm.nih.gov/dbEST/) and the largest set of plant ESTs comes from the model species *Arabidopsis thaliana* and *Oryza sativa.* Nevertheless, a large variety of EST sequences from other plant species including coffee have been deposited in the dbEST database. Recently, a large set of ESTs (46,914) with a special focus on developing seeds of *C. canephora* has also been released (Lin et al. 2005). In addition, two other *C. canephora* EST sequence sets were developed from mRNA isolated from leaves and fruits at different development and maturation stages with 8,778 valid EST sequences (Poncet et al. 2006). All together, a significant set of 55,692 ESTs, mainly from fruits, is already publicly available for *C. canephora* (Table 9.1). Regarding *C. arabica,* public accessibility to EST collections is limited to 1,226 ESTs, comprising a suppression subtractive hybridization library of 618 EST sequences selected upon infection by *H. vastatrix* (Fernandez et al. 2004), 139 EST sequences identified as a differential response of *C. arabica* leaves and roots to chemically induced systemic acquired resistance (De Nardi et al. 2006), and 469 other ESTs from leaves (Cristancho et al. 2006; Joet et al. 2006). In total, there are fewer than 60 thousand entries of coffee ESTs, publicly available, deposited at the dbEST database.

However, other research groups involved in molecular genetics and genomics of *Coffea* sp. have also generated EST data that will become available to the coffee scientific community in the near future (Table 9.1). The Brazilian Coffee EST Project is such an example, which generated single-pass sequences of a total of 214,964 randomly picked clones from 37 cDNA libraries of *C. arabica*, *C. canephora,* and *C. racemosa*, representing specific stages of cells and plant development that after trimming resulted in 130,792, 12,805, and 10,510 good quality sequences for each species, respectively (Vieira et al. 2006). The ESTs assembled into 17,982 clusters and 32,155 singletons. Blast analysis of these sequences revealed that 22% had no significant matches to sequences in the NCBI database. Manually annotated sequencing results have been stored in two online databases (Vieira et al. 2006). Coffee EST resources have also been developed by the Cenicafe research group in Colombia, which have in their database, to date, 32,000 coffee EST sequences from 22 libraries organized in 9,257 *C. arabica* and 1,239 *C. liberica* unigenes (Cristancho et al. 2006). In addition, the Cenicafe database contains 6,000 *Beauveria bassiana* and 4,000 *H. hampei* (coffee berry borer) EST sequences. Aiming at the development of EST-SSR markers for coffee, an Indian research group reported an interim set of 2,092 ESTs of coffee generated at CCMB (Center for Cellular & Molecular

Biology), Hyderabad, India (Aggarwal et al. 2007). EST sequences of two cDNA libraries from leaves and embryonic roots of *C. arabica* were also produced by an Italian group to develop a cDNA microarray based on 1,587 non-redundant sequences (De Nardi et al. 2006).

With these efforts in progress on EST sequencing, the number of coffee ESTs in the public domain will continue to increase as can be seen from data presented in Table 9.1. In fact, the reports compiled here, from different research groups working worldwide, indicate that around 250,000 good quality ESTs from at least four different coffee species, have already been produced. The novelty and complementary nature of these upcoming data can be estimated from a clustering performed with the available coffee EST data at the dbEST database (NCBI) and the ESTs produced in Brazil (Fig. 9.1). The results indicate that only 13% of the unigenes is redundant. These results are expected in view of the different coffee species and tissues sampled to generate the ESTs.

Table 9.1 Available EST collections of *Coffea* species

Coffea species	Tissue/developmental stage	Number of valid ESTs	Reference
C. canephora	Leaves, young	8, 942	Lin et al. 2005
	Pericarp, all developmental stages	8, 956	Lin et al. 2005
	Whole cherries, 18 and 22 week after pollination	9, 843	Lin et al. 2005
	Endosperm and perisperm of seeds, 30 week after pollination	10, 077	Lin et al. 2005
	Endosperm and perisperm of seeds, 42 and 46 week after pollination	9, 096	Lin et al. 2005
	Embryogenic calli	7, 062	Vieira et al. 2006
	Leaves from water-deficit stressed plants	5, 743	Vieira et al. 2006
	Whole cherries of different developmental stages	5, 086	Poncet et al. 2006
	Leaves, young	3, 692	Poncet et al. 2006
C. arabica	Plantlets and leaves treated with araquidonic acid	4, 098	Vieira et al. 2006
	Suspension cells treated with acibenzolar-S-methyl	5, 605	Vieira et al. 2006
	Suspension cells treated with acibenzolar-S-methyl and brassinoesteroids	8, 252	Vieira et al. 2006
	Hypocotyls treated with acibenzolar-S-methyl	9, 882	Vieira et al. 2006
	Suspension cells treated with NaCl	7, 656	Vieira et al. 2006
	Embryogenic calli	7, 599	Vieira et al. 2006
	Zygotic embryo (immature fruits)	106	Vieira et al. 2006
	Germinating seeds (whole seeds and zygotic embryos)	8, 001	Vieira et al. 2006
	Flower buds in different developmental stages	15, 833	Vieira et al. 2006
	Flower buds + pinhead fruits + fruits at different stages	9, 451	Vieira et al. 2006
	Non embryogenic calli with and without 2,4 D	8, 558	Vieira et al. 2006
	Young leaves from orthotropic branch	9, 939	Vieira et al. 2006
	Mature leaves from plagiotropic branches	10, 319	Vieira et al. 2006
	Roots infected with nematodes	302	Vieira et al. 2006
	Primary embryogenic calli	2, 042	Vieira et al. 2006

Table 9.1 (continued)

Coffea species	Tissue/developmental stage	Number of valid ESTs	Reference
	Leaves infected with leaf miner and coffee leaf rust	3, 072	Vieira et al. 2006
	Roots	149	Vieira et al. 2006
	Roots with acibenzolar-S-methyl	1, 051	Vieira et al. 2006
	Suspension cells stressed with aluminum	4, 981	Vieira et al. 2006
	Stems infected with *Xylella spp*.	8, 045	Vieira et al. 2006
	Water-deficit stressed field plants (pool of tissues)	5, 743	Vieira et al. 2006
	Well-watered field plants (pool of tissues)	798	Vieira et al. 2006
	SSH - young coffee leaves	618	Fernandez et al. 2004
	Young coffee leaves	448	Cristiancho et al. 2006
	SSH-young coffee leaves	16	Joet et al. 2006
	Embryonic roots and leaves	2, 016	De Nardi et al. 2006
	17 cDNA libraries (not described)	30, 000	Cristiancho et al. 2006
C. racemosa	Fruits (*Coffea racemosa*)	5, 041	Vieira et al. 2006
	Fruits, stages 1,2 and 3 (*Coffea racemosa*)	5, 469	Vieira et al. 2006
C. liberica	4 cDNA libraries (not described)	2, 000	Cristiancho et al. 2006
Coffea sp.	cDNA libraries not described	2, 092	Aggarwal et al. 2007
	Total *Coffea sp*. ESTs	246,541	

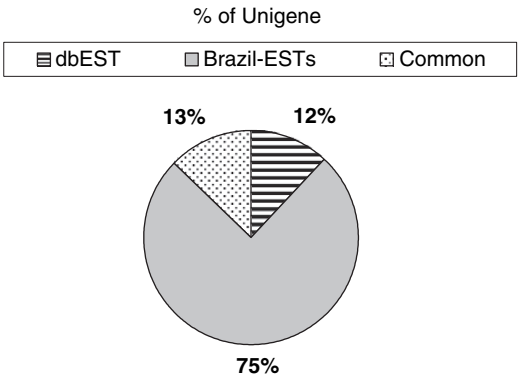


Fig. 9.1 Origin distribution of unigenes from ESTs present on two databases (Brazil-ESTs - 75% of total unigenes are only represented on this database; dbEST – 13 % of total unigenes are only represented on this database; Common- 12 % of total unigenes are represented on both databases)

9.4.2 Coffee EST Exploitation

As sequence and annotation data continue to accumulate, computer-based tools for systematic collection, organization, storage, access, analysis, and visualization of these data will become increasingly valuable to the plant science community. Advances in computational molecular biology and biostatistics will make it possible to analyze large EST datasets more precisely and efficiently producing more reliable digital expression profiling data, taking into account the methodological limitations of constructing cDNA libraries (Alba et al. 2004).

Table 9.2 presents a list of several databases containing coffee data with a diversity of bioinformatics resources. The Solanaceae Genomics Network (SGN; http://sgn.cornell.edu/), serves well as an example of these on-line resources. SGN is dedicated to the biology of Solanaceae species, including tomato, potato, tobacco, eggplant, pepper, petunia, and coffee. This database contains sequence data derived from nearly 300,000 ESTs coming from different plant species, extensive mapping data for the tomato genome, in addition to mapping data for the genomes of potato and eggplant, which can facilitate cross-species homology relationships (via comparative genome analysis) among solanaceus and related species (Wu et al. 2006).

The inventory of the available resources on coffee ESTs provided in this section indicates that considerable information has been generated the last few years. These data are valuable tools for discovering genes (a large number of unknown genes already identified), developing EST-SSR markers, providing the basis for mapping and the establishment of breeding programs based on marker-assisted selection (MAS), and providing the foundation for the development of novel tools such as microarrays. Initiatives for the generation of these arrays have already begun and some results have started to appear (De Nardi et al. 2006). Microarrays and real-time reverse transcription polymerase chain reaction (qRT-PCR) will allow comprehensive transcriptome analysis making it possible to identify and dissect complex genetic networks that underlie important biological processes critical to physiology,

Table 9.2 Bioinformatics resources and databases containing coffee ESTs

Databases	URL	Valid ESTs	Reference
SGN-USA	http://sgn.cornell.edu/	46, 914	Lin et al. 2005
IRD-France	http://www.mpl.ird.fr/bioinfo/	9, 412	Fernandez et al. 2004 Poncet et al. 2006 Joet et al. 2006
CoffeeDNA-Trieste	http://www.coffeedna.net/	5, 391	De Nardi et al. 2006
CCMB-India	http://www.ccmb.res.in/	2, 092	Aggarwal et al. 2007
CBP&D-Café Brazil	http://www.lge.ibi.unicamp.br/cafe/ http://www.cenargen.embrapa.br/biotec/genomacafe/	154, 107	Vieira et al. 2006
Cenicafe-Colombia	http://bioinformatics.cenicafe.org/	32, 000	Cristancho et al. 2006
NCBI	http://www.ncbi.nlm.nih.gov/dbEST/	56, 918	Fernandez et al. 2004 Lin et al. 2005 Poncet et al. 2006 De Nardi et al. 2006

development, quality, and responses to abiotic and biotic stresses in coffee. Recently, the early molecular resistance responses of coffee (*C. arabica* L.) to the rust fungus (*H. vastatrix*) have been monitored using real-time quantitative qRT-PCR (Ganesh et al. 2006). Similarly, several studies of particular biosynthetic pathways in relation with the ripening and development of coffee fruit were reported (Geromel et al. 2006; Hinniger et al. 2006; Simkin et al. 2006; Bustamante et al. 2007; Lepelley et al. 2007). Further characterization of gene networks in coffee plants will help us to identify new targets for manipulation of physiological, biochemical, and developmental processes of this very important crop species.

Moreover, modern biology is facing a new momentum. Recent advances in high-throughput methodologies and equipment allowed researchers from different disciplines to attempt to combine large-scale DNA sequence, gene expression, protein, metabolite, genotype, and/or phenotype data to develop resourceful and integrated databases for a better understanding of biological processes (Alba et al. 2005). These combined tools through integrated efforts of genomics research and breeding will certainly be essential to quickly overcome practical problems faced by the coffee agribusiness such as control of pre- and post-harvest physiological factors involved in quality, disease, and pest control and management of plant responses to environmental changes (e.g., limited water availability and adverse temperatures). Finally, integrated coffee genomic research may result in increased beverage consumption through new value-added products derived from coffee (e.g., nutraceuticals, oils, and flavors), ensuring the sustainability of the coffee production chain.

9.5 Genetic Transformation

Coffee genetic engineering emerged during the last decade as a tool to elucidate the function, regulation, and interaction of agronomically interesting genes through functional genomics approach. Genetic transformation became available only after the establishment of protocols for in vitro regeneration through somatic embryogenesis for the two principal commercial species (Berthouly and Etienne 2000; Etienne 2005). Nowadays, genetic transformation of coffee plants has been successfully achieved by several research groups. However, despite significant advances over the last 15 years, coffee transformation is still very laborious, with bottlenecks in the methodology that make it far from a routine laboratory technique. Up to now, only a few genes have been transferred into coffee genotypes. However, the recent development of coffee genomics has led to the identification of numerous genes involved in agronomically important biological processes and which have potential for transformation.

9.5.1 Direct Gene Transfer

Barton et al. (1991) reported a transformation method of coffee embryos by electroporation using the *nptII* (i.e. kanamycin resistance) gene. This method remained

little noticed until the recent works of Fernandez-Da Silva and Menéndez-Yuffá (2003) that described improved conditions to regenerate transformed *C. arabica* somatic embryos expressing the GUS and *bar* genes. The biolistic delivery method has been improved considerably since the first report of GUS transient expression in coffee using a gunpowder driven device by van Boxtel et al. (1995). Using a helium gun device, Rosillo et al. (2003) determined that a short endosperm pre-treatment with two osmotic preconditioning agents (i.e., mannitol and sorbitol) increased the number of cells expressing GUS. Ribas et al. (2005a) described a protocol for transformation of *C. canephora* embryogenic callus and somatic embryos using a helium gun, associated with sub-culturing onto medium containing mannitol before and after bombardment. Their protocol allowed 12.5% of transformed callus expressing GUS-positive reaction to histochemical assay.

9.5.2 Indirect Gene Transfer

Hatanaka et al. (1999) achieved the first successful *A. tumefaciens*-mediated transformation of *C. canephora* plants exhibiting strong GUS stable expression. Leroy et al. (2000) also reported efficient regeneration of *A. tumefaciens*-transformed coffee plants of both *Coffea* sp. expressing the *uidA* and *cry1Ac* genes. The use of embryogenic callus is often preferred to somatic embryos and becomes the most common way for coffee transformation (Fig. 9.2). From this callus, only 30% (*C. canephora*) and 10% (*C. arabica*) developed somatic embryos, and from these, only 50% regenerated into plantlets. Ribas et al. (2006) transformed *C. canephora* explants submitted to sonification during immersion in a suspension of an *A. tumefaciens* strain encoding *uidA* and *bar* genes and regenerated transformed plants. Canche-Moo et al. (2004) transformed leaf explants through

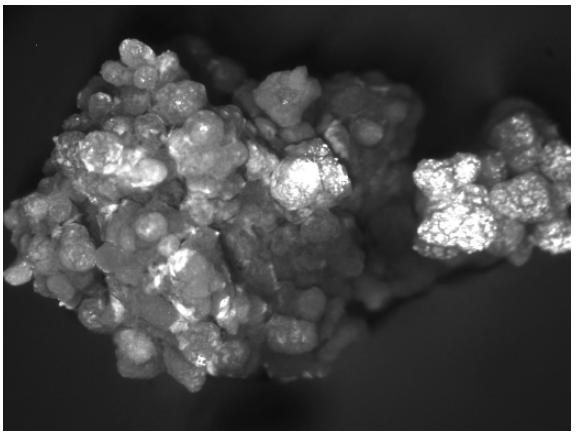


Fig. 9.2 Expression of the 35S-GFP reporter gene on *C. arabica* transformed embryogenic callus three weeks after *Agrobacterium tumefaciens*-mediated transformation

A. tumefaciens-mediated transformation involving a vacuum infiltration protocol in a bacterial suspension. *A. tumefaciens*-mediated transformation has also served to induce stable gene silencing through RNA interference (RNAi) technology of genes encoding theobromine synthase (Ogita et al. 2004) and N-methyl transferase, respectively (Kumar et al. 2004), both genes involved in caffeine biosynthesis, in both cultivated species. Ogita et al. (2003) showed that leaves of 1-year-old transformed trees exhibited reduced theobromine and caffeine content (30 to 50% compared with the control). Ribas et al. (2005b) achieved inhibition of ethylene burst in *C. arabica* by introducing the ACC-oxidase gene in antisense orientation.

Agrobacterium rhizogenes-mediated transformation in both *C. canephora* and *C. arabica* species was first reported by Spiral et al. (1993) and Sugiyama et al. (1995), respectively. However in these studies the regeneration protocol was laborious and plants frequently showed a "hairy" phenotype with short internodes and stunted growth. Such abnormal phenotype is stable as demonstrated by Perthuis et al. (2005), who showed that four out of the nine independently transformed *C. canephora* clones obtained with *A. rhizogenes* were still displaying this phenotype in field conditions; furthermore all these plants died within few months after planting. Kumar et al. (2006) described an adapted method for *A. rhizogenes* sonification-assisted embryos transformation. The percentage of plantlets with aberrant phenotypes was significantly low compared with the results previously described by Sugiyama et al. (1995). Alpizar et al. (2006a) developed an *A. rhizogenes*-mediated transformation protocol (Fig. 9.3) that enables efficient and rapid regeneration of transformed roots from hypocotyls of zygotic embryos and subsequent production of composite plants (i.e., transformed roots induced on non-transformed shoots). This methodology was specifically developed for functional analysis of genes involved

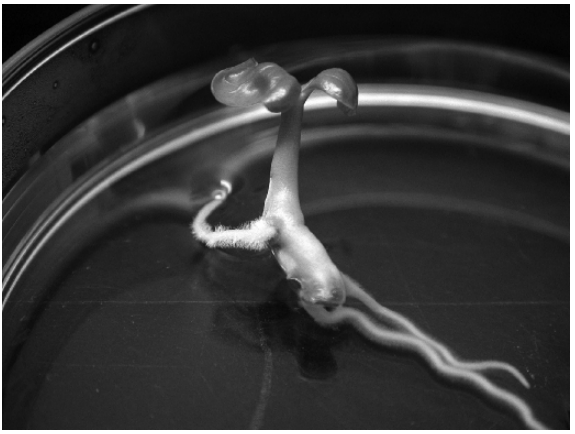


Fig. 9.3 Regeneration of transformed roots on the hypocotyl of zygotic embryo following *Agrobacterium rhizogenes*-mediated transformation protocol. Emergence of a transformed root at the wound site 4-8 weeks after the end of the co-cultivation with A4 RS strain. The transformed zygotic embryo is subcultured every 4 weeks on MS germination medium containing decreasing cefotaxime concentrations (Alpizar et al. 2006a)

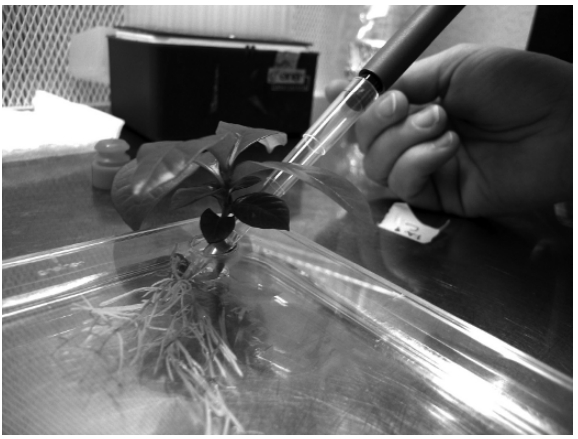


Fig. 9.4 Inoculation of nematodes on transformed roots of *Coffea arabica* cv. Caturra composite plants to evaluate the resistance:susceptibility to the root-knot nematode *Meloidogyne exigua*. The Caturra cv. is susceptible to *M. exigua*. Composite plants were generated after 12 weeks by inducing transformed roots on non-transformed shoots. One month after acclimatization to *ex vitro* conditions, they were ready to be inoculated with *M. exigua* nematode individuals. Gall symptoms caused by *M. exigua* on transgenic roots will be visible three months after nematode inoculation (Alpizar et al. 2006a)

in resistance to root specific pathogens like nematodes and/or in plant morphology and development. The authors demonstrated that the *Meloidogyne exigua* root-knot nematode could normally proliferate in transgenic roots and that consequently this transformation methodology could efficiently be applied for the functional analysis of the *Mex-1* resistance gene (Fig. 9.4).

9.5.3 Selection of Transformed Tissue

The first studies in coffee transformation were done using antibiotics and particularly kanamycin. However, kanamycin has exhibited contradictory results as a potential selection agent for transformed coffee embryos. Barton et al. (1991) and Spiral et al. (1993) concluded that kanamycin possesses poor selective capacity since non-transformed somatic embryos could develop even at high concentration (400 mg/L). Nevertheless, kanamycin was recently used successfully at 400 mg/L by Cunha et al. (2004) and at 100 mg/L by Canche-Moo et al. (2006). Hatanaka et al. (1999), Naveen et al. (2002), and Ogita et al. (2004) observed that hygromycin at 50-100 mg/L allowed an acceptable regeneration of *A. tumefaciens*-mediated transformed somatic embryos. Because of the low selection efficiency of antibiotics in coffee, along with biosafety issues, other types of selection markers including herbicide selection or positive selection were used.

Van Boxtel et al. (1997) and Fernandez-Da Silva and Menéndez-Yuffá (2004) showed that low concentration (6 mg/L) of the herbicide ammonium glufosinate

was sufficient to inhibit non-transformed callus growth. Subsequently, research by Ribas et al. (2005a, 2006) confirmed the reliability of the *bar* gene as a selection marker in both *C. canephora* and *C. arabica.* Leroy et al. (2000) successfully used selectable marker gene *csr1-1* from *A. thaliana* that allowed selection with the herbicide chlorsulfuron at a concentration of 80 μg/L in the two cultivated species. Cruz et al. (2004) also succeeded in regenerating transformed plantlets containing the *ppt* gene, which confers resistance to phosphinothricine, on selective medium containing 10 μM of the herbicide. Samson et al. (2004) demonstrated the potential use of xylose isomerase (*XylA*) gene as a positive selection marker in coffee transformation.

Visual markers have been used to replace those based on herbicide selection. Ogita et al. (2004) and Canche-Moo et al. (2006) used *gfp*, the gene encoding for green fluorescent protein (Chalfie et al. 1994), and *DsRFP*, the gene encoding for red fluorescent protein (Gallie et al. 1989) as reporter genes for visual selection of somatic embryos of *C. canephora* following *A. tumefaciens*-mediated transformation. The first complete selection of coffee transformed tissues without using any marker gene was achieved by Alpizar et al. (2006a, 2006b), after selection of putative *A. rhizogenes* transformed roots using *uidA* and *gfp* genes through histochemical GUS assay or GFP epifluorescence.

9.5.4 Testing Transgenic Coffee Plants

Ribas et al. (2006) regenerated coffee somatic embryos transformed with the *bar* gene on selective medium containing ammonium glufosinate. Regenerated plants supported up to eight times the herbicide doses recommended for field applications. Leroy et al. (2000) described *C. canephora* plants transformed with *cry1Ac* gene from *Bacillus thuringiensis* expressing resistance against the coffee leaf miner (*Perileucoptera coffeella*) in green-house conditions. Perthuis et al. (2005) reported that this resistance was stable and effective after six releases of a natural population of *P. coffeella* during four years of field assessments. In addition, production of transformed coffee plants with resistance to coffee berry borer (*Hypothenemus hampei*) was attempted by Cruz et al. (2004) using the *α-Al1*gene from common bean encoding for an inhibitor of alfa-amylase (Powers and Culbertson, 1983). The authors achieved transformation with this gene and bioassays with the insect are under way to confirm functional activity in coffee. Following recent achievements in other crops, gene pyramiding can be envisaged in coffee to introduce a larger number of genes for resistance to diverse races of a particular pathogen or a combination of different pathogens.

Ribas et al. (2005b) achieved inhibition of the ethylene burst by introducing the *ACC oxidase* gene in antisense orientation; this technique would permit the understanding of genes involved in fruit maturation and ethylene production.

9.5.5 Identification of Coffee Promoters

Availability of tissue-specific and inducible promoters would be helpful for successful development of coffee functional analysis and transgenically improved coffee. With few exceptions, the 35S promoter derived from the cauliflower mosaic virus, has been the most common component of transgenic constructs used. However, Rosillo et al. (2003) comparing pCaMV35S with two coffee promoters (α-tubulin and α-arabicin) showed that they all resulted in similar transient expression of the *uidA* gene. Marraccini et al. (1999) cloned a complete coffee 11S seed storage protein gene and carried out promoter analysis in transgenic tobacco plants. Acuña et al. (1999) also reported the cloning of a bean (endosperm)-specific promoter partially corresponding to the sequence of Marraccini et al. (1999). However, no functional analysis (test in a heterologous plant) of their sequence was performed. In another work, Marraccini et al. (2003) reported the cloning of the rubisco small subunit (*CaRBCS1)* promoter of *C. arabica* that was tested in transgenic tobacco and shown to be leaf-specific. Satyanarayana et al. (2005) reported the cloning of the promoter of N-methyl transferase (NMT) gene involved in caffeine biosynthesis pathway. The promoter region was tested in tobacco, exhibiting expression of the *GUS* reporter gene in leaves. The authors mentioned that current efforts are focused on the use of this promoter sequence for down-regulation of NMT gene through transcriptional gene silencing. This cloning of the first promoter of a gene involved in caffeine biosynthesis together with the near identification of genes involved in sucrose and drought tolerance metabolism opens up the possibility for coffee transformation to validate and study the molecular mechanisms that regulate the production of these important targets for *Coffea* sp. cultivation.

Even without functional analysis in transgenic plants, other coffee promoters were also cloned and described in several publications, some of them recently and directly deduced from the use of EST sequencing: promoter of a protein homologous to the yeast translation initiation factor SUI1 (Gaborit et al. 2003), promoter oleosin *CcOLE-1* (Simkin et al. 2006), promoter dehydrin *CcDH2a* (Hinninger et al. 2006). A number of highly expressed genes that showed high specificity for different stages of seed development, as well as for the pericarp tissue that surrounds the seed, have been identified (Lin et al. 2005). These genes could lead to promoters which can potentially be used to drive gene expression in specific stages/tissues of the coffee plant. Many of these genes are important for insect/pathogen resistance and determining the quality of the coffee bean. It is proposed to create a collection of promoters in the International Coffee Genomics Network (ICGN) as a resource for functional analysis of coffee genes.

9.6 Perspectives

In the last decade, overproduction has resulted in historically low coffee prices that have had devastating effects on coffee producers. Farmers have survived with great difficulties or have abandoned coffee culture and switched to alternative crops.

Consequently, bean quality may be lowered and supply become less stable to adversely affect the coffee market of consumer countries. Furthermore, in the light of climate changes and increasing awareness of the negative impacts of the non-sustainable use of natural resources, coffee production will have to evolve. This is particularly relevant for perennial plants such as coffee whose productive life is very long and for which rapid genetic gains are tedious to obtain.

In the twentieth century, coffee production benefited from the selection of spontaneous Arabica mutants (e.g., the compact cultivar Caturra), natural hybrids between species (e.g., the Timor hybrid), highly productive Canephora clones, and new cultural practices (e.g., open sun monoculture). Coffee germplasm was collected, but little has been done at the molecular level to exploit biodiversity in coffee species as well as their genomic resources.

The recent development of high-capacity methods for analyzing the structure and function of genes, which is collectively termed "genomics," represents a new paradigm with broad implications. The advent of large-scale molecular genomics for plant species such as *C. arabica and C. canephora* will provide access to previously inaccessible sources of genetic variation which could be exploited in breeding programs. Anticipated outcomes in current and future coffee breeding include (1) rapid characterization and managing of germplasm resources, (2) enhanced understanding of the genetic control of priority traits, (3) identification of candidate genes or tightly linked genomic regions underlying important traits, and (4) identification of accessions in genetic collections with variants of genomic regions or alleles of candidate genes having a favorable impact on priority traits. In this way, the recent efforts to set up an international commitment (ICGN, http://www.coffeegenome.org) to work jointly for the development of common sets of genomic tools, plant populations, and concepts would be extremely useful.

References

Acuña R, Bassüner R, Beillinson V, Cortina H, Cadena-Gómez G, et al. (1999) Coffee seeds contain 11S storage proteins. Physiol Plant 105:122–131

Aga E, Bryngelsson T (2006) Inverse sequence-tagged repeat (ISTR) analysis of genetic variability in forest coffee (*Coffea arabica* L.) from Ethiopia. Genetic Res Crop Evol 53:721–728

Aga E, Bryngelsson T, Bekele E, Salomon B (2003) Genetic diversity of forest arabica coffee (*Coffea arabica* L.) in Ethiopia revealed by random amplified polymorphic DNA (RAPD) analysis. Hereditas 138:36–46

Aggarwal RK, Hendre PS, Varshney RK, Bhat PR, Krishnakumar V, et al. (2007) Identification, characterization and utilization of EST-derived genetic microsatellite markers for genome analyses of coffee and related species. Theor Appl Genet 114:359–372

Alba R, Fei Z, Payton P, Liu Y, Moore SL, et al. (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. Plant J 39:697–714

Alba R, Payton P, Fei Z, McQuinn R, Debbie P, et al. (2005) Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. Plant Cell 17:2954–2965

Alpizar E, Dechamp E, Espeout S, Royer M, Lecouls A-C, et al. (2006a) Efficient production of *Agrobacterium rhizogenes*-transformed roots and composite plants for studying gene expression in coffee roots. Plant Cell Rep 25:959–967

Alpizar E, Dechamp E, Bertrand B, Lashermes P, Etienne H (2006b) Transgenic roots for functional genomics of coffee resistance genes to root-knot nematodes. Proc Intl Sci Colloquium on Coffee 21:653–659, ASIC (http://www.asic-cafe.org)

Anthony F, Berthaud J, Guillaumet JL, Lourd M (1987) Collecting wild *Coffea* species in Kenya and Tanzania. Plant Genet Res Newsl 69:23–29

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 MC, Astorga C, Bertrand B, Graziosi G, et al. (2002) The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. Theor Appl Genet 104:894–900

Barre P, Layssac M, D'Hont A, Louarn J, Charrier A, et al. (1998) Relationship between parental chromosomic contribution and nuclear DNA content in the coffee interspecific hybrid: *C. pseudozanguebariae* x *C. liberica* var. *dewevrei*. Theor Appl Genet 96:301–305

Barton CR, Adams TL, Zarowitz M (1991) Stable transformation of foreign DNA into *Coffea arabica* plants. Proc Intl Conf Coffee Sci 14:460–464

Baruah A, Naik V, Hendre PS, Rajkumar R, Rajendrakumar P, et al. (2003) Isolation and characterization of nine microsatellite markers from *Coffea arabica* L., showing widecross-species amplifications. Mol Ecol Notes 3:647–650

Bennett MD, Leitch IJ (1995) Nuclear DNA amounts in angiosperms. Ann Bot 76:113–176

Berthouly M, Etienne H (2000) Somatic embryogenesis of coffee. In: Coffee biotechnology and quality. Proc Intl Seminar Biotechnology Coffee Agro-industry 13:71–90

Bhat PR, Krishnakumar V, Hendre PS, Rajendrakumar P,Varshney RK, et al. (2005) Identification and characterization of gene-derived EST–SSR markers from robusta coffee variety 'CxR' (an interspecific hybrid of *Coffea canephora* and *Coffea congensis*). Mol Ecol Notes 5:80–83

Bouharmont J (1959) Recherches sur les affinités chromosomiques dans le genre *Coffea*. Publ. INEAC, Série Scientifique 77:94 p

Bremer B, Jansen RK (1991) Comparative restriction site mapping of chloroplast DNA implies new phylogenetic relationships within *Rubiaceae*. Am J Bot 78:198–213

Bridson D (1982) Studies in *Coffea* and *Psilanthus* (*Rubiaceae* subfam. *Cinchonoideae*) for Part 2 of 'Flora of Tropical East Africa': *Rubiaceae*. Kew Bull 36:817–859

Bridson D, Verdcourt B (1988) *Coffea*. In: Polhill RM (ed) Flora of Tropical East Africa. Rubiaceae (Part 2). A.A. Balkema, Rotterdam, pp 703–727

Bustamante J, Campa C, Poncet V, Noirot M, Leroy T, et al. (2007) Molecular characterization of an ethylene receptor gene (*CcETR1*) in coffee trees, its relationship with fruit development and caffein content. Mol Genet Genomics (in press), DOI: 10.1007/s00438-007-0219-z

Canche-Moo RLR, Ku-Gonzales A, Burgeff C, Loyola-Vargas VM, Rodríguez-Zapata LC, et al. (2006) Genetic transformation of *Coffea canephora* by vacuum infiltration. Plant Cell Tiss Org Cult 84:373–387

Carvalho A (1988) Principles and practice of coffee plant breeding for productivity and quality factors: *Coffea arabica*. In: Clarke RJ, Macrae R (eds) Coffee, volume 4: Agronomy. Elsevier Applied Science, London, pp 129–165

Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC (1994). Green florescent protein as a marker for gene expression. Science 263:663–664

Chevalier A, Dagron M (1928) Recherches historiques sur les débuts de la culture du caféier en Amérique. Communications et Actes de l'Académie des Sciences Coloniales (Paris) 5:1–38

Combes MC, Andrzejewski S, Anthony F, Bertrand B, Rovelli P, et al. (2000) Characterisation of microsatellite loci in *Coffea arabica* and related coffee species. Mol Ecol 9:1178–1180

Coulibaly I, Revol B, Noirot M, Poncet V, Lorieux M, et al. (2003) AFLP and SSR polymorphism in a Coffea interspecific backcross progeny [(*C. heterocalyx* x *C. canephora*) x *C. canephora*]. Theor Appl Genet 107:1148–1155

Couturon E (1982) Obtention d'haploïde spontanés de *Coffea canephora* Pierre par l'utilisation du greffage d'embryons. Café Cacao Thé 26(3):155–160

Cristancho MA, Rivera C, Orozco C, Chalarca A, Mueller L (2006) Development of a bioinformatics platform at the Colombia National Coffee Research Center Proc Intl Sci Colloquium Coffee 21: 638–643, ASIC (http://www.asic-cafe.org)

Cros J, Combes MC, 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 MC, Trouslot P, Anthony F, Hamon S, et al. (1998) Phylogenetic relationships of *Coffea* species: new evidence based on the chloroplast DNA variation analysis. Mol Phylo Evol 9:109–117

Crouzillat D, Rigoreau M, Bellanger L, Priyono S, Mawardi S, Syahrudi, McCarthy J, Tanksley S, Zaenudin I, Petiard V (2004) A Robusta consensus map using RFLP and microsatellites markers for the detection of QTL. Proc Intl Sci Colloquium Coffee 20: 546–553, ASIC (http://www.asic-cafe.org)

Cruz ARR, Paixao ALD, Machado FR, Barbosa MFF, Junqueira CS, et al. (2004) Metodologia para obtenção de plantas transformadas de *Coffea canephora* por co-cultivo e calos embriogênicos com *A. tumefaciens*. Boletim de Pesquisa e Desenvolvimento. Embrapa, Brasilia, Brasil, vol.58:15 p

Cunha WG, Machado FRB, Vianna GR, Texteira JB, Barros EVSA (2004) Obtenção de *Coffea arabica* geneticamente modificado por bombardeio de calos embriogênicos. Boletim de Pesquisa e Desenvolvimento. Embrapa, Brasilia, Brasil, vol. 73:15 p

Davis AP, Maurin O, Chester M, Mvungu EF, Fay MF (2006) Phylogenetic relationship in *Coffea* (Rubiaceae) inferred from sequence data and morphology. Proc Intl Sci Colloquium Coffee 21: 868-875, ASIC (http://www.asic-cafe.org)

De Nardi B, Dreos R, Del Terra L, Martellossi C, Asquini E, et al. (2006) Differential responses of *Coffea arabica* L. leaves and roots to chemically induced systemic acquired resistance. Genome 49:1594–1605

Dussert S, Lashermes P, Anthony F, Montagnon C, Trouslot P, et al. (2003) Coffee (*Coffea canephora*). In: Hamon P, Seguin M, Perrier X, Glaszmann JC (eds) Genetic diversity of cultivated tropical plants. Science Publishers Inc., Plymouth, pp 239–258

Eira MTS, Amaral da Silva EA, de Castro RD, Dussert S, Walters C, et al. (2006) Coffee seed physiology. Brazilian J Plant Physiol 18:149–163

Etienne H (2005) Protocol of somatic embryogenesis: Coffee (*Coffea arabica* L. and *C. canephora* P.). In: Protocols for somatic embryogenesis in woody plants. Series: Forestry Sci Vol. 77, Jain SM, Gupta PK (Eds). Springer, the Netherlands. ISBN: 1-4020-2984-5, pp. 167–179

Fernandez D, Santos P, Agostini C, Bon MC, Petitot AS, et al. (2004) Coffee (*Coffea arabica* L.) genes early expressed during infection by the rust fungus (*Hemileia vastatrix*). Mol Plant Pathol 5:527–536

Fernandez-Da Silva R, Menéndez-Yuffá A (2003) Transient gene expression in secondary somatic embryos from coffee tissues electroporated with genes *gus* and *bar*. Electronic J Biotech 6:29–35

Fernandez-Da Silva R, Menéndez-Yuffá A (2004) Efecto del herbicida glufosinate de amonio en diferentes explantes de *Coffea arabica* cv. Catimor. Acta Científica Venezolana 55:211–217

Gaborit C, Caillet V, Deshayes A, Marraccini P (2003) Molecular cloning of a full-length cDNA and gene from *Coffea arabica* encoding a protein homologous to the yeast translation initiation factor SUI1: expression analysis in plant organs. Braz J Plant Physiol 15: 55–58

Gallie DR, Lucas WJ, Walbot V (1989). Vizualing mRNA expression in plant protoplasts: factors influencing efficient mRNA uptake and translation. Plant Cell 1: 303–311

Ganesh D, Petitot A-S, Silva M, Alary R, Lecouls AC, et al. (2006). Monitoring of the early molecular resistance responses of coffee (*Coffea arabica* L.) to the rust fungus (*Hemileia vastatrix*) using real-time quantitative RT-PCR. Plant Sci 170:1045–1051

Geromel C, Ferreira LP, Guerreiro SMC, Cavalari AA, Pot D, et al. (2006) Biochemical and genomic analysis of sucrose metabolism during coffee (*Coffea arabica*) fruit development. J Exp Bot 57:3243–3258

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

Haarer AE (1956) Modern coffee production. Leonard Hill (books) Limited, London

Hatanaka T, Choi YE, Kusano T, Sano H (1999) Transgenic plants of *Coffea canephora* from embryogenic callus via *Agrobacterium tumefaciens*-mediated transformation. Plant Cell Rep 19:106–110

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

Herrera JC, Combes MC, Anthony F, Charrier A, Lashermes P (2002) Introgression into the allotetraploid coffee (*Coffea arabica* L.): segregation and recombination of the *C. canephora* genome in the tetraploid interspecific hybrid (*C. arabica* x *C. canephora*). Theor Appl Genet 104:661–668

Hinniger C, Caillet V, Michoux F, Ben Amor M, Tanksley S, et al. (2006) Isolation and characterization of cDNA encoding three dehydrins expressed during *Coffea canephora* (Robusta) grain development. Ann Bot 97:755–765

Joet T, Salmona J, Suzanne W, Descroix F, Dussert S, et al. (2006) Targeted transcriptome profiling during seed development in *C. arabica* cv. Laurina. Proc Intl Sci Colloquium Coffee 21: 687–694, ASIC (http://www.asic-cafe.org)

Krug CA, Mendes AJT (1940) Cytological observations in *Coffea* - IV. J Genet 39:189–203

Krug CA, Mendes JET, Carvalho A (1939) Taxonomia de *Coffea arabica L.* Campinas: Instituto Agronômico do Estado, Bolétim Técnico no 62

Kumar V, Sathyabarayana KV, Indu EP, Sarala Itty S, Giridhar P, et al. (2004) Post transcriptional gene silencing for down regulating caffeine biosynthesis in *Coffea canephora* P. Proc Intl Sci Colloquium Coffee 20:769–774

Kumar V, Satyanarayana KV, Sarala Itty S, Indu EP, Giridhar P, et al. (2006) Stable transformation and direct regeneration in *Coffea canephora* P ex. Fr. *Agrobacterium rhizogenes* mediated transformation without hairy-root phenotype. Plant Cell Rep 25:214–222

Ky CL, Barre P, Lorieux M, Trouslot P, Akaffou S, et al. (2000) Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee Coffea sp. Theor Appl Genet 101:669–676

Lashermes P, Trouslot P, Anthony F, Combes MC, Charrier A (1996a) Genetic diversity for RAPD markers between cultivated and wild accessions of *Coffea arabica*. Euphytica 87:59–64

Lashermes P, Cros J, Combes MC, Trouslot P, Anthony F, et al. (1996b) Inheritance and restriction fragment length polymorphism of chloroplast DNA in the genus *Coffea* L. Theor Appl Genet 93:626–632

Lashermes P, Combes MC, Trouslot P, Charrier A (1997) Phylogenetic relationships of coffee tree species (*Coffea* L.) as inferred from ITS sequences of nuclear ribosomal DNA. Theor Appl Genet 94:947–955

Lashermes P, Combes MC, Robert J, Trouslot P, D'Hont A, et al. (1999) Molecular characterisation and origin of the *Coffea arabica* L. genome. Mol Gen Genet 261:259–266

Lashermes P, Paczek V, Trouslot P, Combes MC, Couturon E, et al. (2000a). Single-locus inheritance in the allotetraploid *Coffea arabica* L. and interspecific hybrid *C. arabica* x *C. canephora*. J Heredity 91:81–85

Lashermes P, Andrzejewski S, Bertrand B, Combes MC, Dussert S, et al. (2000b) Molecular analysis of introgressive breeding in coffee (*Coffea arabica* L.). Theor Appl Genet 100:139–146

Lashermes P, Combes MC, Prakash NS, Trouslot P, Lorieux M, et al. (2001) Genetic linkage map of *Coffea canephora*: effect of segregation distortion and analysis of recombination rate in male and female meioses. Genome 44:589–595

Lashermes P, Noir S, Combes MC, Ansaldi C, Bertrand B, et al. (2004) Toward an Integrated Physical Map of the Coffee Genome. Proc 20th Intl Sci Colloquium Coffee 20: 554–559, ASIC (http://www.asic-cafe.org)

Lepelley M, Cheminade G, Tremillon N, Simkin A, Caillet V, et al. (2007) Chlorogenic acid synthesis in coffee**:** An analysis of CGA content and real-time RT-PCR expression of HCT, HQT,

C3H1, and CCoAOMT1 genes during grain development in *C. canephora*. Plant Sci (in press), DOI:10.1016/j.plantsci.2007.02.004

Leroy T, Henry A-M, Royer M, Altosaar I, Frutos R, et al. (2000) Genetically modified coffee plants expressing the *Bacillus thuringiensis cry1Ac* gene for resistance to leaf miner. Plant Cell Rep 19:382–389

Leroy T, Marraccini P, Dufour M, Montagnon C, Lashermes P, et al. (2005) Construction and characterization of a *Coffea canephora* BAC library to study the organization of sucrose biosynthesis genes. Theor Appl Genet 111:1032–1041

Lin C, Mueller LA, McCarthy J, Crouzillat D, Pétiard V, et al. (2005). Coffee and tomato share common gene repertoires as revealed by deep sequencing of seed and cherry transcripts. Theor Appl Genet 112:114–130

Lombello RA, Pinto-Maglio CAF (2004) Cytogenetic studies in *Coffea L.* and *Psilanthus* Hook. Using CMA/DAPI and FISH. Cytologia 69:85–91

López G, Moncada MDP (2006) Construction of an interspecific genetic Linkage map from a *Coffea liberica* x *C. eugenioides* F_1 Population. Proc Intl Sci Colloquium Coffee 21: 644–652, ASIC (http://www.asic-cafe.org)

Mahé L, Combes MC, Lashermes P (2007) Comparison between a coffee single copy chromosomal region and *Arabidopsis* duplicated counterparts evidenced high level synteny between the coffee genome and the ancestral *Arabidopsis* genome. Plant Mol Biol 64:699–711

Marraccini P, Deshayes A, Pétiard V, Rogers WJ (1999) Molecular cloning of the complete 11S seed storage protein gene of *Coffea arabica* and promoter analysis in transgenic tobacco plants. Plant Physiol Biochem 37:273–282

Marraccini P, Courjault C, Caillet V, Lausanne F, Lepage B, et al. (2003) Rubisco small subunit of *Coffea arabica*: cDNA sequence, gene cloning and promoter analysis in transgenic tobacco plants. Plant Physiol Biochem 41:17–25

Mettulio R, Rovelli P, Anthony F, Anzueto F, Lashermes P, et al. (1999) Polymorphic microsatellites in *Coffea arabica*. Proc Intl Sci Colloquium Coffee 18:344–347

Moncada P, McCouch S (2004) Simple sequence repeat diversity in diploid and tetraploid Coffea species. Genome 47:501–509

Naveen KS, Sreenath HL, Sreedevi G, Veluthambi K, Naidu R (2002) Transgenic coffee (*Coffea arabica*) plants with markers genes through *Agrobacterium tumefaciens*-mediated transformation. XV Plant Crops Symposium. Placrosym. Mysore, India, pp. 219–225

Noir S, Patheyron S, Combes MC, Lashermes P, Chalhoub B (2004) Construction and characterisation of a BAC library for genome analysis of the allotetraploid coffee species (*Coffea arabica* L.). Theor Appl Genet 109:225–230

Ogita S, Uefuji H, Yamaguchi Y, Sano H (2003) RNA interference: producing decaffeinated coffee plants. Nature 423:823

Ogita S, Uefuji H, Morimoto M, Sano H (2004) Application of RNAi to confirm theobromine as the major intermediate for caffeine biosynthesis in coffee plants with potential for construction of decaffeinated varieties. Plant Mol. Biol 54:931–941

Orozco-Castillo C, Chalmers KJ, Waugh R, Powell W (1994) Detection of genetic diversity and selective gene introgression in coffee using RAPD markers. Theor Appl Genet 87:934–940

Orozco-Castillo C, Chalmers KJ, Powell W, Waugh R (1996) RAPD and organelle specific PCR re-affirms taxonomic relationships within the genus *Coffea*. Plant Cell Rep 15:337–341

Paillard M, Lashermes P, Pétiard V (1996) Construction of a molecular linkage map in coffee. Theor Appl Genet 93:41–47

Pearl HM, Nagai C, Moore PH, Steiger DL, Osgood RV, et al. (2004) Construction of a genetic map for arabica coffee. Theor Appl Genet 108:829–835

Perthuis B, Pradon J, Montagnon C, Dufour M, Leroy T (2005) Stable resistance against the leaf miner *Leucoptera coffeella* expressed by genetically transformed *Coffea canephora* in a pluriannual field experiment in French Guiana. Euphytica 144:321–329

Pinto-Maglio CAF, Da Cruz ND 1998. Pachytene chromosome morphology in *Coffea* L. II. *C. arabica* L. complement. Caryologia, 51:19–35

Poncet V, Hamon P, Minier J, Carasco C, Hamon S, et al. (2004) SSR cross-amplification and variation within coffee trees (*Coffea* spp.). Genome 47:1071–1081

Poncet V, Rondeau M, Tranchant C, Cayrel A, Hamon S, et al. (2006) SSR mining in coffee tree EST databases: potential use of EST-SSRs as markers for the *Coffea* genus. Mol Genet Genomics 276:436–449

Powers J, Culbertson JD (1983). Interaction of a purified bean (*Phaseolus vulgaris*) glycoprotein with an insect amylase. Cereal Chem 60:107–117

Prakash NS, Marques DV, Varzea VMP, Silva MC, Combes MC, et al. (2004) Introgression molecular analysis of a leaf rust resistance gene from *Coffea liberica* into C*offea arabica* L. Theor Appl Genet 109:1311–1317

Prakash NS, Combes MC, Dussert S, Naveen S, Lashermes P (2005) 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

Raina SN, Mukai Y, Yamamoto M (1998) *In situ* hybridisation identifies the diploid progenitor of *Coffea arabica* (Rubiaceae). Theor Appl Genet 97:1204–1209

Ribas AF, Kobayashi AK, Pereira LFP, Vieira LGE (2005a) Genetic transformation of *Coffea canephora* by particle bombardment. Biol Plant 49:493–97

Ribas AF, Galvão RM, Pereira LFP, Vieira LGE (2005b) Transformação de *Coffea arabica* com o gene da ACC-oxidase em orientação antisenso. 50th Congreso Brasileiro de Genética p. 492

Ribas AF, Kobayashi AK, Pereira LFP, Vieira LGE (2006) Production of herbicide-resistance coffee plants (*Coffea canephora* L.) via *Agrobacterium tumefaciens*-mediated transformation. Braz Arc Biol Technol 49:11–19

Rosillo AG, Acuna JR, Gaitan AL, de Pena M (2003) Optimized DNA delivery into *Coffea arabica* suspension culture cells by particle bombardment. Plant Cell Tiss Org Cult 74:75–79

Samson NP, Campa C, Noirot M, De Kochko A (2004) Potential use of D-Xylose for coffee plant transformation. Proc Intl Conf Coffee Sci 20: 707-713 (ASIC) (http://www.asic-cafe.org)

Satyanarayana KV, Kumar V, Chandrashekar A, Ravishankar GA (2005) Isolation of promoter for *N*-methyltransferase gene associated with caffeine biosynthesis in *Coffea canephora*. J Biotechnol 119:20–25

Simkin AJ, Qian T, Caillet V, Michoux F, Ben Amor M, et al. (2006) Oleosin gene family of *Coffea canephora*: quantitative expression analysis of five oleosin genes in developing and germinating coffee grain. J Plant Physiol 163:691–708

Spiral J, Thierry C, Paillard M, Pétiard V (1993) Obtention de plantules de *Coffea canephora* Pierre (Robusta) transformées par *Agrobacterium rhizogenes*. C. R. Acad Sci Paris 3:1–6

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

Stoffelen P, Noirot M, Couturon E, Anthony F (2007) A new caffeine-free coffee species in the deep rain forest of Cameroon. Taxon (in press)

Sugiyama M, Matsuoka C, Takagi T (1995) Transformation of *Coffea* with *Agrobacterium rhizogenes*. Proc Intl Conf Coffee Sci 16:853–859

Thomas AS (1942) The wild arabica coffee on the Boma Plateau, Anglo-Egyptian Sudan. Empire J Expt Agric 10:207–212

van Boxtel J, Berthouly M, Carasco M, Dufour M, Eskes A (1995) Transient expression of ÿ-glucuronidase following biolistic delivery of foreign DNA into coffee tissue. Plant Cell Rep 14:748–752

van Boxtel J, Eskes A, Berthouly M (1997) Glufosinate as an efficient inhibitor of callus proliferation in coffee tissue. In Vitro Cell Dev Biol Plant 33:6–12

Van der Vossen HAM (2001) Agronomy I: Coffee Breeding Practices. In: Clarke RJ, Vitzthum OG (eds) Coffee: recent developments. Blackwell Science, United Kingdom, pp 184–201

Vieira LGE, Andrade AC, Colombo CA, Araujo AH, Metha A et al. (2006) Brazilian coffee genome project: an EST-based genomic resource. Braz J Plant Physiol 18:95–108

Wellman FL (1961) Coffee: botany, cultivation and utilization. London: Leonard Hill Books

Wu F, Mueller LA, Crouzillat D, Pétiard V, Tanksley SD (2006) Combining bioinformatics and phylogenetics to identify large sets of single-copy orthologous genes (COSII) for comparative, evolutionary and systematic studies: A test case in the euasterid plant clade. Genetics 174:1407–1420