Annals of Botany **100**: 1073–1084, 2007
doi:10.1093/aob/mcm191, available online at www.aob.oxfordjournals.org

OXFORD JOURNALS
OXFORD UNIVERSITY PRESS

REVIEW

Domestication, Genomics and the Future for Banana

J. S. HESLOP-HARRISON* and TRUDE SCHWARZACHER
Department of Biology, University of Leicester, Leicester LE1 7RH, UK

Received: 16 July 2007 Returned for revision: 22 July 2007 Accepted: 25 July 2007 Published electronically: 31 August 2007

• *Background* Cultivated bananas and plantains are giant herbaceous plants within the genus *Musa*. They are both sterile and parthenocarpic so the fruit develops without seed. The cultivated hybrids and species are mostly triploid ($2n = 3x = 33$; a few are diploid or tetraploid), and most have been propagated from mutants found in the wild. With a production of 100 million tons annually, banana is a staple food across the Asian, African and American tropics, with the 15 % that is exported being important to many economies.
• *Scope* There are well over a thousand domesticated *Musa* cultivars and their genetic diversity is high, indicating multiple origins from different wild hybrids between two principle ancestral species. However, the difficulty of genetics and sterility of the crop has meant that the development of new varieties through hybridization, mutation or transformation was not very successful in the 20th century. Knowledge of structural and functional genomics and genes, reproductive physiology, cytogenetics, and comparative genomics with rice, *Arabidopsis* and other model species has increased our understanding of *Musa* and its diversity enormously.
• *Conclusions* There are major challenges to banana production from virulent diseases, abiotic stresses and new demands for sustainability, quality, transport and yield. Within the genepool of cultivars and wild species there are genetic resistances to many stresses. Genomic approaches are now rapidly advancing in *Musa* and have the prospect of helping enable banana to maintain and increase its importance as a staple food and cash crop through integration of genetical, evolutionary and structural data, allowing targeted breeding, transformation and efficient use of *Musa* biodiversity in the future.

Key words: *Musa*, banana, plantain, genome evolution, biodiversity, selection, plant breeding.

THE BOTANY OF BANANA (GENUS *MUSA*)

Bananas and plantains are monocotyledonous plants in the genus *Musa* (Musaceae, Zingiberales). They are giant herbs, commonly up to 3 m in height, with no lignification or secondary thickening of stems that is characteristic of trees (Tomlinson, 1969; see Fig. 2A). The centre of origin of the group is in South-East Asia, where they occur from India to Polynesia (Simmonds, 1962). The centre of diversity has been placed in Malaysia or Indonesia (Daniells *et al.*, 2001), although considerable diversity is known throughout the range. The plants are distributed mainly on margins of tropical rainforests (Wong *et al.*, 2002).

The taxonomy of the approximately 50 species within the genus *Musa* remains poorly resolved, not least because of the widespread vegetative reproduction and natural occurrence of many hybrids. Most frequently, the genus is divided into four (sometimes five) sections, Eumusa and Rhodochlamys with a basic chromosome number of $x = 11$, Australimusa ($x = 10$), and Callimusa ($x = 10$ or $x = 9$) (after Cheesman, 1947; Simmonds and Weatherup, 1990; Dolezel and Bartos, 2005). Various minor and major regroupings have been suggested (Wong *et al.*, 2002). At the species level, the number of species and the status of subspecies has been debated (Taxonomic Advisory Group for *Musa*, 2007). However, from the point of view of the taxa related to crops, the morphological features are well defined (*Musa* Germplasm Information System, MGIS; IPGRI–INIBAP/CIRAD, 1996). In the last decade, in conjunction with molecular studies of *Musa* accessions at the DNA level (see Bartos *et al.*, 2005), aspects of the taxonomy have been clarified but a careful treatment of the complementary and contrasting data, along with judicious filling of gaps in the data, is required to resolve the relationships and phylogeny in the genus.

In a broader phylogenetic view based on groupings that can be regarded as strictly monophyletic, Angiosperm Phylogeny Group (2003) puts *Musa* within the Zingiberales, one of four sister Orders in the monophyletic grouping Commelinids, along with the Poales (grasses), Commelinales and Aracales. This puts *Musa* in an important taxonomic position from the point of view of comparative genetics, since it is a sister group to the well-studied grasses. Apart from the cereals, bananas are the major crop species within the Commelinids. The Musaceae family includes a second genus, *Ensete*, with the Ethiopian Banana, used occasionally as a food in East Africa. Ginger, where the root is eaten, is the other significant crop in the Zingiberales. There are a number of horticultural species in the genus *Musa*, and *Strelitzia reginae* (bird-of-paradise) lies in the sister familiy Strelitziacea. The leaves of *Musa* are used for their fibre content: when fresh as plates for eating or wrapping food parcels for steaming, or when dry as strips for weaving into various articles and for roofing shelters. Specific names such as *M. ornata* and *M. textilis* reflect these uses.

* For correspondence. E-mail phh4@le.ac.uk

© 2007 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

THE BANANA CROP AND DOMESTICATION

Bananas and plantains are the fourth most important crop in developing countries, with a worldwide production of about 100 Mt. Bananas provide a starch staple across some of the poorest parts of the world in Africa (with consumption up to 400 kg per person per year) and Asia, while dessert bananas are a major cash crop in many countries (FAOStat, 2007). Bananas and plantains are cultivated throughout the humid tropics and sub-tropics in the Americas, Africa and Asia (each representing about a third of production), extending into Europe (Canary Islands) and Australia (Queensland). Worldwide, well over a thousand banana cultivars or landraces are recognized. The vast majority of the cultivated bananas (Pollefeys *et al.*, 2004) are derived from inter- and intraspecific crosses between two diploid ($2n = 2x = 22$) wild species, *Musa acuminata* and *Musa balbisiana* (Simmonds and Shepherd, 1955). In terms of the chromosome sets, these are designated as having the genome constitution AA (*M. acuminata*) or BB (*M. balbisiana*). These diploid *Musa* species have seeded fruit with little starch and only a small amount of fleshy pith, and are of no value as a crop.

The cultivated bananas and plantains differ from their wild relatives by being seedless and parthenocarpic – the fruit develops without seed development or pollination and fertilization. The genetic basis of the mutation (or mutations) in the A genome that gives rise to parthenocarpy has not been characterized, and no parthenocarpy has been identified in B genome diploids, although hybrids of A and B show the character. Most of the cultivars are wild collections made by farmers of spontaneously occurring mutants with parthenocarpic fruit production, which were brought into cultivation and then multiplied and distributed by vegetative propagation. There is no straightforward botanical distinction between bananas and plantains but, in general, bananas refer to the sweeter forms that are eaten uncooked, while starchy fruits that are peeled with a knife when unripe and then cooked are referred to as plantains and cooking bananas, while some cultivars are 'beer bananas' for fermentation of the juice, or used for deep frying as banana chips.

Many of the domesticated bananas (Figs 1 and 2) have proved to be triploid, $2n = 3x = 33$, with genome constitutions of AAA (mainly the sweet dessert bananas), AAB or ABB (mainly but not exclusively starchy plantains eaten after cooking). There are also seedless cultivated AA and AB diploids, and tetraploids ($2n = 4x = 44$) with genome constitutions of AAAA, AAAB, AABB and ABBB. These various plants have been collected from multiple, independent sources in the wild, so the hybridization events and mutations giving rise to the seedless and parthenocarpic characters have occurred many hundreds of times. Where fertile plants occur together, hybridization continues to produce new diversity (Pollefeys *et al.*, 2004) and parental combinations. Simmonds (1962) considered five plant characteristics that lead to farmers picking plants for cultivation: plant vigour, yield, seedlessness, hardiness and fruit quality, the first four of which are related to polyploidy (triploidy).

The ploidy and genome constitution of banana accessions has been determined by study of plant and fruit morphology since the 1940s; now, methods of numerical taxonomy have refined the approach (Simmonds and Weatherup, 1990; Ortiz, 1997*a*; Ortiz *et al.*, 1998; Pollefeys *et al.*, 2004). Flow cytometric analysis (Dolezel and Bartos, 2005) is proving valuable for accurate and rapid surveys without growing mature plants. *In situ* hybridization to chromosome preparations using DNA probes that label the A and B ancestral genomes separately have shown that the full sets of $x = 11$ chromosomes are present (Fig. 3; Osuji *et al.*, 1997), and most cultivars have complete genomes of 11 chromosomes. However, d'Hont *et al.* (2000) used *in situ* hybridization to show that the variety 'Pelipita' ($2n = 3x = 33$) included eight A chromosomes and 25 B chromosomes rather than the 11 A and 22 B of the normal ABB

FIG. 1. The diversity of banana and plantains on sale in a shop in south India (Varkala, Kerala State) with various genome compositions. Cultivars are indicated by letters above bunches: a, cultivar 'Red' (AAA genome constitution), a prized sweet dessert banana cultivar. Differences between bunches are mostly from water and nitrogen conditions in the field, and not genetic. b, 'Palayam Codan' (AAB). c, 'Njalipoovan' AB (unripe and ripe, green and yellow) sweet dessert banana with small fingers, thin skin and delicate flavour but poor keeping quality and the fruits fall off bunches. d, 'Robusta' ('Cavendish' group, AAA); 'Cavendish' cultivars ripen without changing to yellow (green ripe) when above 22 °C. e, 'Nendran' (AAB), used for cooking and for making chips. f, 'Peyan' (ABB) used as a vegetable for curries and for cooked snacks. g, 'Poovan' (AAB). (A light tube has been edited out of the picture in the top left.)

FIG. 2. A, A banana plant with ripening fruit bunch. A sucker is growing from the base of the stem which will form a replacement plant after the fruit is harvested and the mother plant cut down. B, The dessert banana 'Gros Michel' (AAA, 2n=3x=33) killed by Panama disease or *Fusarium* wilt. 'Gros Michel' was the major export banana before spread of the disease led to its replacement by the variety 'Cavendish' which accounts for nearly all the export trade in banana.

type, while two other AAB plantains with 33 chromosomes had more than 11 B genome chromosomes. Analysis of IRAP markers (see below) indicate that 'Pelipita' lies in an anomalous phylogenetic position away from other species (Desai and Heslop-Harrison, unpubl. res. in work from 2004). Thus backcrossing or chromosome elimination has occurred during the derivation of some varieties. Molecular analyses have suggested the presence of chromosome markers from *Musa* species other than the A and B genome, and d'Hont *et al.* (2000) used *in situ* hybridization to confirm the presence of complete S and T genomes, from *Musa schizocarpa* and *M. textilis*, respectively, in some varieties. The diploid variety 'Wompa' was AS, while other genotypes were established as AAT and ABBT.

Polyploids can originate by doubling of somatic chromosome numbers from diploids, but the most common origin in the wild is likely to be from 2*n* gametes, where gametogenesis has been modified, sometimes through failure of one division of meiosis and generation of a restitution gamete with the same genetic constitution as the parent. The importance of 2*n* gametes has been considered in wild diploid species and cultivated polyploid crops (Ortiz and Peloquin, 1992). Raboin *et al.* (2005) used RFLP markers to show that 'Cavendish' and 'Gros Michel' AAA bananas had a common diploid ancestor from the Indian Ocean Islands that was included in the triploid cultivars through 2*n* gametes, with the two cultivars having a different genotype contributing the third genome through a haploid gamete.

As well as the nuclear genome, the origin and fate of the plastid and mitochondrial genomes are important in hybrid crops, contributing to critical agronomic traits. In the cereals, the organellar or plasmon genomes have significant consequences for fertility, disease resistance and yield (Tsunewaki and Ogihara, 1983), as well as indicating the nature and directions of hybrids that have given rise to new species. In *Musa*, Faure *et al.* (1994) made controlled and reciprocal crosses that demonstrated strong bias towards maternal transmission of chloroplast DNA, but showed the unusual phenomenon of paternal transmission of mitochondrial DNA in *Musa acuminata.* The study was extended by Carreel *et al.* (2002) to analyse the origins of more

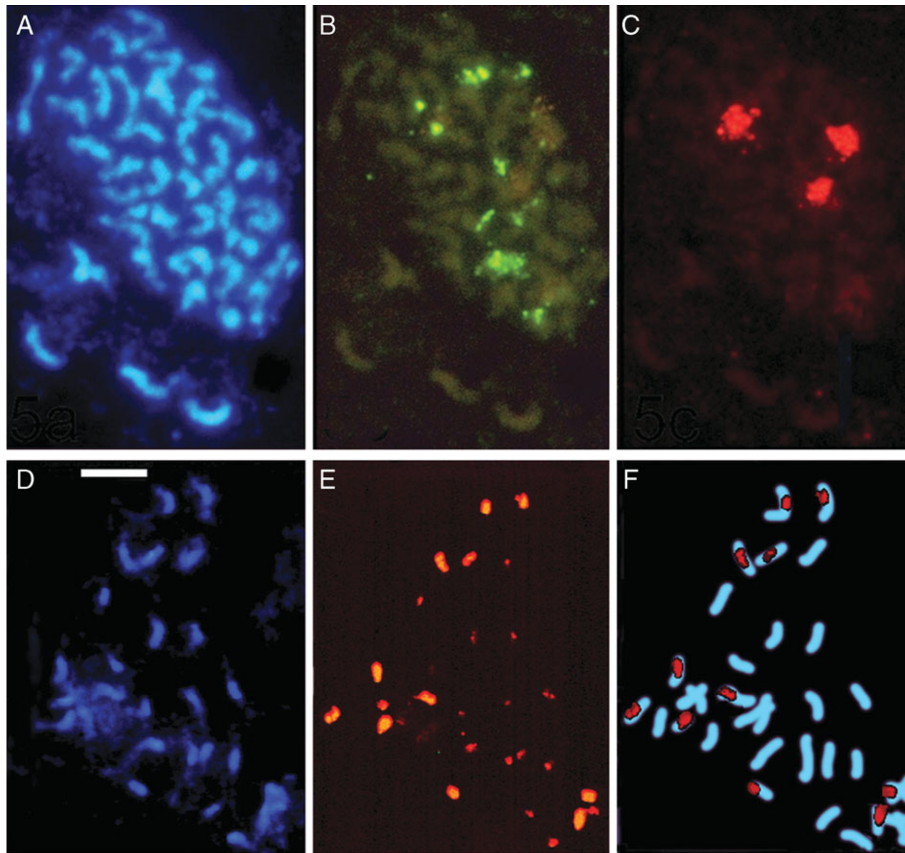

FIG. 3. *In situ* hybridization to banana chromosomes ($2n = 3x = 33$) stained blue (A, D) with the fluorochrome DAPI. (A–C) hybridization of 5S (labelled green, B) and 45S (red, C) rDNA to chromosomes from the ABB cooking banana 'Fougamou' showing three major 45S rDNA loci (one chromosome in each genome carries locus), while multiple chromosomes have 5S loci. (D–F) A metaphase from the ABB ($2n = 3x = 33$) cultivar 'Bluggoe' labelled with total genomic DNA from the diploid genome donor *Musa acuminata* (A genome; red in E). The DNA hybridizes predominantly to the centromeres of A genome origin chromosome and identifies these 11 chromosomes, shown in the drawing in F (Osuji *et al.*, 1997, 1998). Scale bar in (D) = 5 μm.

than 300 *Musa* genotypes, leading to the conclusion that most cultivars are linked to two subspecies of *M. acuminata*, *M. acuminata banksii* and *M. acuminata errans*, through their mitochondrial genomes.

BANANA PROPAGATION

The banana plant readily produces vegetative suckers next to the mother pseudostem at the base of plant, with strong vascular connection to the mother (Fig. 2A). These can be removed from the parent and planted separately, where they rapidly develop new leaves and root systems, allowing rapid vegetative propagation and multiplication. In cultivation, unwanted suckers are removed to avoid weakening the parent plant. The suckers are the major source of planting material and normally remain true-to-type. After planting, at a typical density of 1500 to 2500 plants ha^{-1}, each plant produces a single pseudostem with one fruit bunch of 20–40 kg harvested 9–14 months after planting. The plant is then cut to ground level, the leaves removed and destroyed to control disease, and a side sucker allowed to grow up to produce the next crop. In intensively managed plantations, the plants are replaced with new, disease-free planting material after three-to-eight of these ratooning cycles. Where plants are not replaced, a gradual and continuous yield decline is usually observed, attributed to disease build-up.

During propagation, some somatic clonal variants have been observed and selected, in particular for infloresecnce, fruit and height characteristics (Krikorian *et al.*, 1993; Szymkowiak and Sussex, 1996). Good examples come from the 'Cavendish' group of dessert bananas, where there are several height variants such as (in approximate descending order) 'Lacatan', 'Robusta', 'Valery', 'Giant Cavendish', 'Grand Naine', 'Dwarf Cavendish', 'Petit Naine' and 'Dwarf Parfitt', and other variants, such as 'Williams' and 'Zelig'. The changes giving rise to these independently named varieties are considered to be genetic mutations, although without the possibility of carrying out genetic segregation tests and without cloning and sequencing the relevant genes this is not proven and they may be epigenetic variants. However, there are some detectable changes between the 'Cavendish' groups at the DNA level. The diversity of new forms derived through a combination of accumulation of somatic mutations and human selection has led Ortiz (1997*a*) to consider sub-Saharan Africa as a secondary centre of banana diversity.

In vitro tissue culture propagation systems are very efficient in *Musa*. These can give high-quality, uniform plants free of disease and nematodes, and much of the planting material used in commercial plantations, and increasingly in smallholder production, comes from mass micropropagation. Shoot tip cultures have been most widely used (Strosse *et al.*, 2004), but suspension cultures are also being developed (Roux *et al.*, 2001). In some tissue culture systems, high levels of chimerism are found, where chromosome number and genotype vary (Roux *et al.*, 2001) in the resulting plants. The valued South Indian 'Red' sweet banana (see Fig. 1) shows regular reversion of the colour character to green, particularly in tissue-culture propagated plants but also in the field (Stover and Simmonds, 1987), although the basis of this has not be confirmed. A programme checking varietal characteristics of material grown up after a decade of storage *in vitro* is showing that very few morphological or ploidy variants have been induced (van den Houwe *et al.*, 1995). Applications of molecular markers (see below) do show some DNA changes (Ray *et al.*, 2006) arising following tissue culture. Notably, Oh *et al.* (2007) demonstrated that some genomic regions of *Musa* show higher rearrangements and mutation rates than others under the stresses imposed during tissue culture.

PLANT DISEASES AND ABIOTIC STRESS CHALLENGES

The *Musa* crop suffers from several devastating diseases (Robinson, 1996), and there is a continuous challenge to existing varieties by new diseases and newly virulent disease strains, which are met with agronomic practices to control disease spread, the development and application of chemical controls, and the search for genetically resistant cultivars. In commercial plantations, twice-weekly agrochemical applications may account for 30 % of the production costs, while steady yield declines in many smallholder systems may be accounted for by disease. Historically, Panama disease or *Fusarium* wilt, caused by the fungus *Fusarium oxysporum* f. sp. *cubense* has devastated banana production and 'is widely regarded as one of the most destructive plant diseases in recorded history' (Moore *et al.*, 1995). The disease was first reported in Australia in 1874 and spread to nearly all banana-growing regions, eventually destroying the export trade based on the variety 'Gros Michel' (Fig. 2B) and its leading to its replacement in the 1950s and 60s by the resistant 'Cavendish' (AAA) group of cultivars that are now the major export variety. More recently, an extremely virulent form of the pathogen, 'Tropical Race 4', has been spreading and is causing substantial losses to both subsistence farmers and commercial growers since it attacks 'Cavendish' (Hwang and Ko, 2004). Once established in an area, *Fusarium* cannot be controlled chemically by fungicides or soil fumigants, or by cultural practices such as rotations or soil amendment, so the only long-term option for continuing banana production is replacement of a susceptible variety with a resistant variety (Hwang and Ko, 2004; Daly and Walduck, 2006). However, most commercial varieties are susceptible to 'Tropical Race 4' (e.g. all those in Taiwan tested by Su *et al.*, 1986). Although a number of varieties have been identified with resistance genes that may be useful in breeding or gene-transfer programmes, these varieties have weaknesses and are not suitable as replacements for 'Cavendish' (Daly and Walduck, 2006).

Another fungal disease, Black sigatoka leaf spot or black leaf streak disease (BLSD, *Mycosphaerella fijiensis*) has been serious in recent years (Ferreira *et al.*, 2004), with infection commonly leading to 50 % crop losses and the need for control with environmentally undesirable and expensive fungicides. There is some genetic resistance in *Musa* with potential for exploitation (Ortiz and Vuylsteke, 1995), and genomic studies of the pathogen, including complete sequencing, are underway (Conde-Ferráez *et al.*, 2007). A bacterial wilt caused by *Xanthomonas* is spreading rapidly in East Africa; although control of spread by cultural practices is being attempted, a long-term solution may again come through genetic resistance. Other viral and fungal diseases include banana bunchy top virus (BBTV, controlled by eradication of infected plants), banana streak virus (BSV), freckle (*Cladosporium musae*), *Phyllostictina musarium*, *Cordana* and Moko disease (*Ralstonia solanacearum*; Jain and Swennen, 2004). Burrowing nematodes (*Radopholus similis* and *Pratylenchus* spp.) and weevil (*Cosmopolites sordidus*) pests also constrain banana production, with little genetic resistance in widely grown cultivars and infection often leading to plantations becoming uneconomic and being abandoned.

An important group of viruses, the badnaviruses including BSV (Hull, 2002), are related by sequence to the retroelements. Harper *et al.* (1999) showed that the BSV-related sequences are integrated within the nuclear genome, and since then it has become clear that the elements can be expressed and give rise to a viral infection (Harper *et al.*, 2002), although integration is not an essential part of the viral life cycle. Hull *et al.* (2000) and others have speculated that the presence of integrated copies may confer virus resistance through induction of transcriptional or post-transcriptional gene silencing of homologous sequences.

BANANA GERMPLASM

Within collections of *Musa* germplasm worldwide, there are between 1500 and 3000 accessions, representing a wide range of morphological variation and genome constitutions. Many of the varieties of banana have been maintained locally, although in recent years the distribution of selections with improved characteristics, as well as consumer preference for defined varieties, has led to a reduction in the number of varieties grown. This is evident in South India where three or four clones are most widely available. Johannessen (1970) pointed out how farmers in Central America collect different varieties, with typically 12–21 being present on each farm, which he considered 'as a form of insurance against possible disaster'. It is not clear whether this is as widely practiced now, and anecdotal observations might suggest that the practice is declining.

Some 1000 *Musa* cultivars and 180 wild species, representing much of the diversity of the banana crop, are maintained in tissue culture at the Bioversity International Transit Centre (ITC) in Belgium, and these provide a valuable reference collection that is mostly in the public domain and freely accessible for research and breeding, and distributed as tissue-culture plantlets. Elsewhere in the world, many researchers, particularly in Asia, have developed field-based germplasm collections, and well-curated internet databases are now disseminating information about these collections (Pollefeys *et al.*, 2004). While morphology and flow cytometry are measures of diversity, the analyses have limitations and there remain questions about the presence of multiple genotypes with a single name or a single genotype with multiple names. Where plants have been studied in different regions, how much diversity is environmentally induced and how much genotypic? What diploid genotypes are represented in the triploids, and how diverse are the triploid selections? Answers to these questions, now coming from DNA-based molecular diversity studies (see below) will help to direct plant breeders towards appropriate germplasm to test and select, and to focus germplasm collections towards representing the full range of diversity present in the genus at all ploidy levels.

BANANA AND PLANTAIN IN PLANT BREEDING

Plant breeders face similar challenges in all species: the need for higher quality and productivity in the face of changing pressures from both biotic (disease-caused) and abiotic (environmental) stresses. As with any crop, a combination of intensive cultivation and extensive growth of single varieties leads to the emergence of virulent pathogens, either through new mutations in the pathogens or through spread of existing virulent strains. Abiotic stress, whether from water, salinity, wind or temperature, coming from increasing the range where the crop is grown, changing water availability or climate change, is affecting the crop and can lead to large yield instability and fluctuations in production and price. *Musa* germplasm includes disease resistances, abiotic stress resistances, and altered agronomic performance within various accessions. The genepool, as in other domesticated species, provides a valuable resource for improvement of the crop. However, conventional cross-breeding programmes using elite cultivars are not practical in banana because cultivated bananas are sterile: the diversity in cultivars is derived from wild collections that are vegetatively propagated. In *Musa*, the challenges are to identify and characterize the relevant genes and genetic diversity, and then to utilize this variation in a largely sterile crop propagated by tissue culture. Biotechnology and gene technology, together with conventional methods, can assist in overcoming the problems in developing new banana cultivars.

Banana breeding – the deliberate creation and selection by breeders of synthetic hybrids – has been carried out since the 1920s (Rowe, 1984; Ortiz and Vuylsteke, 1995) but by early in the 21st century, none of these synthetic cultivars and radiation mutants were widely grown (Jain and Swennen, 2004). Rowe (1984) has summarized sexual methods for the breeding of bananas at that time. Vuylsteke *et al.* (1997) and Ortiz (1997*c*) reviewed the wide array of breeding schemes being explored, combining conventional and innovative approaches, and producing potential cultivars from primary tetraploids, secondary triploids and other populations. There are a number of new cultivars that are now released or in field tests arising from these programmes, and it is anticipated that they will soon become more widely grown and accepted.

Although seed set in triploid banana accessions is very low, some triploid cultivars have residual fertility, and hand-pollination with diploid parents gives some seed – typically averaging less than one to 20 or more seeds per fruit bunch (30 kg; Ortiz and Vuylsteke, 1995). These authors were able to produce useful hybrids following *in vitro* germination of seed from triploid × diploid crosses. Ortiz (1997*b*) has reported the production of $2n$ pollen in both male-fertile diploid and polyploid accessions of *Musa*, but suggested that $2n$ pollen production may be a character that evolved late in the domestication of cultivated *Musa*. The occurrence of $2n$ pollen allowed Ortiz to suggest it will be useful for transfer of desired attributes to the polyploid level through sexual polyploidization, rather than alternative analytical breeding schemes. Particularly where the parental stocks are not well characterized, a good knowledge of relationships between diploid (wild and cultivated) and triploid cultivars is essential to help choose diploid parents for use in producing agriculturally useful hybrids (Carreel *et al.*, 2002).

The banana breeding programmes at FHIA in Honduras (Rowe and Rosales, 1993) has been based on selection of improved diploids that were crossed to eventually generate triploids, a number of which have been released as varieties. Other programmes involve crossing AA diploids with AAB triploids to generate diploid and tetraploid (AAAB) hybrids with improved disease resistance and agronomic characters, a strategy used by the International Institute of Tropical Agriculture (IITA) in Nigeria and Kenya (Crouch *et al.*, 1999) and in Brazil by EMBRAPA (Ferreira *et al.*, 2004). CARBAP in Africa and CIRAD-FLHOR in the West Indies also have active crossing and selection projects (Escalant *et al.*, 2002).

Breeding through mutation is an important approach for *Musa*, and as with many mutation programmes is particularly valuable to compensate for defined weaknesses in existing cultivars. Roux (2004) has summarized the history of the two commercially released banana lines derived from gamma-ray-induced mutations. Both have useful agronomic traits: ‘Novaria’, released in 1995, flowers about 10 weeks earlier than the original parental clone (‘Grande Naine’ in the ‘Cavendish’ group), while the Thai variety ‘Klue Hom Thong’ KU1 has large bunch size and a cylindrical shape with larger banana fingers. A number of other traits derived from mutation treatments or from somaclonal variants, including dwarfism, *Fusarium* wilt and Black Sigatoka resistance, are currently being evaluated (Hwang and Ko, 2004; Roux, 2004).

A number of alternative ways to generate transgenic bananas have been developed. Transformation with alien genes and selectable markers is usually low-efficiency and dependent on the genotype and state of material, although procedures are improving. Both particle bombardment (Sági *et al.*, 1995) and *Agrobacterium*-mediated (May *et al.*, 1995; Khanna *et al.*, 2004; Rodríguez-Zapata *et al.*, 2005) strategies have been applied. As well as marker genes, anti-fungal, anti-nematode (Atkinson *et al.*, 2004), virus-resistance genes and some for control of ripening have been the earliest targets for transformation, and clearly both the precision and speed of this technology (e.g. using cell suspensions and BIBAC vectors, Ortiz-Vazquez *et al.*, 2005) is increasing as plants are trialled in the field. The use of banana transformed with fruit promoters and appropriate protein genes has also been considered to produce and deliver edible vaccines (Mason *et al.*, 2002).

BANANA GENETICS

Because of the sterility, genetics based on segregating populations has not been possible in the triploid bananas, and efforts at mapping diploid *Musa* species have not been particularly successful. The first genetic linkage map was published by Faure *et al.* (1992) based on 90 loci in a population of 92 F_2 individuals from an F_1 cross between two divergent *M. acuminata* accessions. Finding suitable polymorphisms for mapping was difficult because of the heterozygosity of the grandparents, and 15 linkage groups (rather than 11) were detected, accounted for by chromosome structural rearrangements between the two grandparents. There are several programmes aiming to make new F_1 mapping populations, but none is yet available for researchers. Where diploid segregating populations have been established, it seems they are difficult to maintain: maybe viability of hybrid populations is low because of chromosomal translocations and segregation of inviable gene alleles. A combination of cost, space and disease/biosecurity constraints, with potential restrictions because they are intellectual property, also limit generation of *Musa* mapping populations.

Because of the difficulty in making crosses for linkage mapping, alternative strategies are being explored for *Musa* genetics. Firstly, now that thousands of PCR-based markers can be created from *Musa* DNA sequences in the databases, the HAPPY mapping strategy could be applied (Thangavelu *et al.*, 2003). Genomic DNA is diluted to a fraction of a genome equivalent per tube (effectively segregating markers to different tubes) and amplified non-selectively, before aliquots are tested for amplification by pairs of PCR markers to assess linkage. Another approach to mapping could use *Musa* double-haploids (as generated by Assani *et al.*, 2003) and, if large numbers could be produced from a single heterozygous recombining pollen source, then linkage mapping could be established in the progeny, perhaps in tissue culture. In the absence of a structured segregating population, gene isolation, characterization and functional genomics can be carried out using collections of mutations – a mutation grid – (Ahloowalia *et al.*, 2004), providing an alternative strategy that might be applicable, particularly in sterile cultivars.

The lack of mapping populations in *Musa* means there is no prospect of mapping quantitative trait loci (QTLs) using the conventional approach with large segregating populations of F_2, backcross or doubled-haploid plants. The use of linkage disequilibrium (LD) mapping (also called association genetics), as in human populations (with its related approach, genome-wide – single nucleotide polymorphism typing and association with phenotype, Gibbs and Singleton, 2006), may be an important development for discovery of important trait genes and valuable alleles in *Musa*. LD mapping requires the ability to genotype hundreds of accessions with thousands of markers, and the diversity of *Musa* accessions available in germplasm banks from unstructured populations is appropriate for this approach.

MUSA GENOMICS

The genome size of *Musa* was determined by Dolezel *et al.* (1994): the unreplicated haploid genome size is 550 Mbp in *M. balbisiana* and 600 Mbp in *M. acuminata* (Lysak *et al.*, 1999; extended to a range of species in all sections of *Musa* and in *Ensete* by Bartos *et al.*, 2005). This is larger than the genomes found in species such as rice or *Arabidopsis thaliana* (between 150 and 500 Mbp), but much smaller than the Triticeae cereals (5500 Mbp in barley, 9000 Mbp in rye, and 17 000 Mbp in hexaploid wheat; Bennett and Leitch, 2004). Divided among 11 chromosomes in the haploid set, this means that the chromosomes are relatively small, and all are similar sizes with about 50 Mbp of DNA. Osuji *et al.* (1998) showed that there was one major 45S rDNA site on each chromosome set, but variation in the number of 5S rDNA sites between two and six (Dolezelova *et al.*, 1998; Osuji *et al.*, 1998; Bartos *et al.*, 2005). Figure 3 shows metaphase chromosomes of a triploid *Musa* cultivar.

In contrast to the genetic mapping of *Musa*, the study of genomics advanced rapidly in the first years of the 21st century. Several BAC clone libraries were developed from both A and B genome diploid *Musa* species (Vilarinhos *et al.*, 2003; Safár *et al.*, 2004; Ortiz-Vazquez *et al.*, 2005; Musagenomics, 2007) with a total genomic coverage of more than 30 times. The Musa Genomics Resource Centre was established in the Czech Republic to distribute these resources, and by 2007, 42 BACs totalling 3·3 Mbp (starting with Aert *et al.*, 2004) and 4·5 Mbp of BAC end sequences (BES; Cheung and Town, 2007) were published. These allow an overview of the *Musa* genome showing that it has a 47 % GC content, the development of large numbers of PCR-based markers, and a comparison of genes and genomic structures with other species.

The gene sequences of banana have been analysed from genomic DNA and from RNA. Analysis of the BAC and BES sequences show that a coding gene occurs every 6·4 to 6·9 kb in genomic DNA (Aert *et al.*, 2004; Cheung and Town, 2007) except in regions with large numbers of transposable elements where there are fewer genes. Several thousand ESTs (expressed sequence tags, gene sequences

analysed by isolation of mRNA), important for examination of gene expression, responses and differentiation of the plants, and examination of diversity, have been published (Santos *et al.*, 2005) and many tens of thousands more are becoming available. Comparisons of EST libraries will be very valuable for identification of genes that are differentially expressed under stress conditions – Santos *et al.* (2005) made libraries from plants grown in cold (5 °C) and hot (45 °C) conditions, and found that about 30 % of the genes in their library had been identified in other species as being involved in responses to environmental stress, and that there were substantial differences in the expression between the two libraries. Coemans *et al.* (2005) used an alternative method for gene-expression profiling (transcriptome mapping), SuperSAGE, which they suggest will be very useful for screening and identifying genes that are differentially expressed under biotic or abiotic stresses.

Like all plant genomes, the *Musa* genome consists of repetitive DNA and single copy sequences, and understanding the composition and organization of the genome at the large-scale level is helpful to allow gene isolation and to understand long-term and short-term evolutionary processes (Kubis *et al.*, 1998; Schmidt and Heslop-Harrison, 1998). Baurens *et al.* (1996) used a competitive PCR to show that some repetitive DNA sequence families represented more than 1 % of the *Musa* genomes. By using labelled total genomic DNA from A and B genome diploid species to hybridize to chromosomes from hybrid cultivars, Osuji *et al.* (1997) showed that the repetitive DNA components of the two genomes were substantially different since the two or three chromosome sets in the hybrids were labelled differentially (see Fig. 3). These results indicated that there were large arrays of repetitive sequence at the centromeres of the chromosomes, since these regions were both stained brightly as chromocentres using the DNA stain DAPI, and the *in situ* hybridization showed that the major genomic repeats represented in the labelled genomic DNA were from the centromeres. Valarik *et al.* (2002) cloned and characterized many repetitive DNA sequences and located those that were not related to rDNA or retroelements in the centromeric region of the chromosomes.

Several pairs of BACs containing homoeologous regions from the *M. acuminata* (A) and *M. balbisiana* (B) genomes have been sequenced. Figure 4 shows a dot-plot of two clones of *Musa*. The continuous diagonal line shows regions where a high proportion of nucleotides are identical between the two species. Gaps may be horizontal or vertical (regions where there are additional sequences present in one genome) or diagonal (regions where the sequences differ between the two genomes). At the scale of less than 100 000 bp, these plots show that the genomes are similar and that the gene order is largely conserved. Cheung and Town (2007) compared pairs of BAC end sequences from single BACs and found that a small number also mapped to adjacent regions of the rice genome, indicating conserved microsynteny over a larger taxonomic range, and showing how part of the *Musa* genome can be anchored to rice. As the authors point out, this provides a cost-effective and efficient way to understand *Musa* genes and the genome by informatics and conserved synteny with the model reference species rice and *Arabidopsis thaliana*.

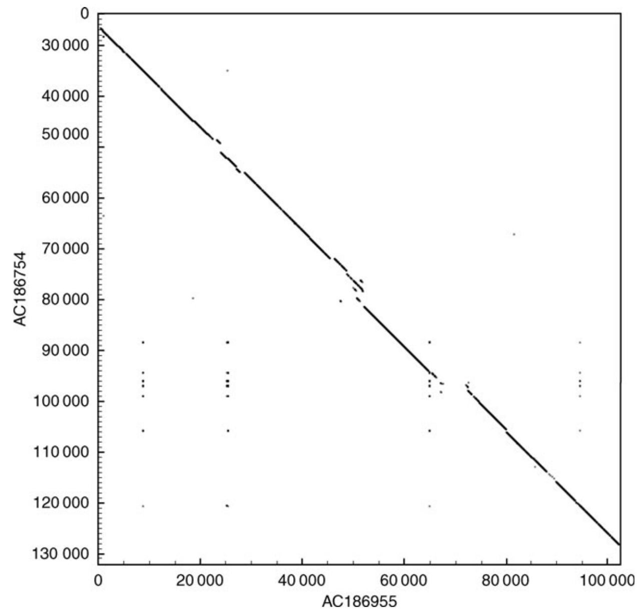

FIG. 4. A dot-plot comparing about 100 000 bp (*x* and *y* axes) of the genomic DNA sequence in homoeologous BAC clones from the A and B genome. Horizontal: Genbank accession AC186955 from *Musa acuminata*; vertical AC186754 from *M. balbisiana* (submitted to database by Chris Town *et al.*, TIGR, 2006; dot-plot made using program by Sonnhammer and Durbin, 1995). Similar DNA sequences are indicated by lines made up of dots, while gaps indicate regions of different sequence or insertions and deletions.

Retroelements, class I transposable elements or transposons, are abundant in the *Musa* genome (Baurens *et al.*, 1997; Balint-Kurti *et al.*, 2000; Teo *et al.*, 2002), as in other species (Heslop-Harrison *et al.*, 1997). Automated annotation of the BAC libraries shows that more than a third of the open reading frames are related to retroelements (Musagenomics, 2007). The analysis of two BACs by Aert *et al.* (2004) revealed that one BAC consisted of 45 kb of gene-rich sequence without retroelements, followed by 28 kb containing mostly transposon-like sequences and repetitive DNA. BAC-end sequencing, allowing a survey of the whole genome, showed that 36 % of the BESs contained sequences homologous to transposable elements (Cheung and Town, 2007). The evidence suggests that *Musa* has repeat-rich regions in the centromeres and perhaps elsewhere, and there may be gene-rich regions, as suggested in other species (see Heslop-Harrison, 1991).

BANANA DIVERSITY

Morphological data has suggested that *Musa* is diverse with well-defined characters (MGIS, Pollefeys *et al.*, 2004) giving accurate indicators of genome constitution and ploidy. However, phenotyping for many physiological characters, including biotic and abiotic stress tolerance, particularly under controlled, contained, reproducible conditions is difficult because of the size of the plants, and long life

cycle. International exchange of disease-free germplasm (e.g. from the ITC), allows identical genotypes to be trialled under contrasting conditions, and is particularly important for testing field disease resistance; the 'Gros Michel' plant shown in Fig. 2B was grown in a *Fusarium*-infected 'hot-spot'. However, phenotyping for drought tolerance has been less systematic; an ideal study would require alteration of the single variable of water under field conditions, with effects of chronic and acute drought onset being tested independently. This could be achieved with controlled irrigation in a dry area, since large plants with deep roots would probably be too difficult to grow under shelters. Greenhouse-grown plants and *in vitro* tests where, for example, plantlets or cultured cells are challenged with fungal toxins or osmotic shock, are being used to evaluate important characters, but correlation to field behaviour remains essential.

Molecular marker methods using polymorphic-anonymous (that is, not of any known functional importance to the plant) markers are widely used for germplasm characterization. In the 1980s, isozyme and anthocyanin analysis confirmed that *Musa* germplasm was indeed genetically diverse (Jarret and Litz, 1986; Horry and Jay, 1988). Soon after, DNA markers became available. Although little attention has since been paid to non-DNA based biochemical markers, including antibodies, it is worth noting that they can be very high throughput, fast and accurate, with little requirement for skilled labour or advanced laboratory equipment compared to DNA markers. They may prove to be valuable in some plant diversity, selection and breeding contexts, allowing larger samples and complementing use of DNA markers.

Projects aiming to characterize and sample the diversity of *Musa* germplasm including wild species and cultivars with world-wide (Tenkouano *et al.*, 1999) and regional (Pillay *et al.*, 2001) focus have used most of the different classes of DNA markers including RFLP, AFLP (Ude *et al.*, 2002*a*, *b*), PCR–RFLP (Nwakanma *et al.*, 2003*a*, *b*), microsatellites and retroelement (Teo *et al.*, 2005) markers, as well as the unreliable (Jones *et al.*, 1997) RAPD method. These studies are proving valuable in national breeding programmes and fundamental studies of biodiversity, while it is also important that genetic diversity is evaluated within the genebanks. In the decade to 2007, some 30 refereed papers have been published covering diversity in national or continental collections, wild species, cultivars and landraces, and presenting different techniques and primers.

The data have been used to infer relationships between *Musa* accessions, including investigation of the representation of wild species' diversity in cultivars, to build phylogenetic trees suggesting relationships and groupings of diverse germplasm. Methods using PCR-based techniques are particularly appropriate because of the high levels of polymorphism that can be detected efficiently. Primers flanking polymorphic simple sequence repeats (SSRs) have been used in a number of studies in *Musa* where they provide co-dominant, chromosome-specific markers; SSRs have been successfully used for germplasm analysis and estimation of genetic relationships between accessions. As pointed out by Buhariwalla *et al.* (2005), SSR markers are routinely used in numerous studies in *Musa* (Crouch *et al.*, 1999, 2000 and references therein; Creste *et al.*, 2004), and a large number of markers have been published from both the A and B genome; many of the markers amplify sequences from both genomes.

An alternative marker system to SSRs exploits the abundance and polymorphic nature of retrotransposon insertions into the nuclear genome. By using PCR primers facing outwards from conserved regions of retroelements, DNA fragments can be amplified and inter-retroelement amplified polymorphisms (IRAPs) detected (Waugh *et al.*, 1997; Kalendar *et al.*, 1999); as with SSR markers, polymorphisms can be used to infer relationships and phylogenies as the retroelements are a relatively rapidly evolving genomic component. IRAP methods can reveal high levels of polymorphism, and have the advantages that no DNA digestion, ligations or probe hybridization are needed to generate the marker data (Nair *et al.*, 2005; Teo *et al.*, 2005).

SUPER-DOMESTICATION AND BANANAS

As with all other crop species, banana production faces major challenges from biotic and abiotic stresses. As well as the critical need for better genetic disease resistance, there is the need for banana to contribute to better food security through greater yields and more efficient production, greater yield stability, better use of water, less and less toxic chemical input, and higher quality. About 85 % of banana production is eaten as a staple food in the country of production, but this market is changing rapidly with urbanization of the population and a smaller proportion of the crop being produced for homestead use, meaning transport and distribution chains. In the export market, there is a huge investment in the shipping chain for the single important export cultivar group 'Cavendish', and even legislation defining fruit sizes and prohibiting abnormal curvature (European Commission, 1994). Such strict standards make changes in variety difficult (Raboin *et al.*, 2005), but banana export remains a significant part of the economy of many developing countries.

Compared to other species discussed in this volume, banana has an unusual triploid genetic background with parthenocarpy, and the process of domestication has been largely through collection of individual varieties from spontaneous mutations in the wild. Hence conventional genetics and plant breeding are relatively poorly developed. However, genomic approaches are now rapidly advancing in *Musa*: the Global *Musa* Genomics Consortium was established in 2001 to assure the sustainability of banana as a staple food crop by developing an integrated genetic and genomic understanding, allowing targeted breeding, transformation and more efficient use of *Musa* biodiversity.

The gene-pool within wild *Musa* species and represented in the cultivars has the diversity to meet the challenges faced by banana breeders and its characterization is a key to making better use of the biodiversity. Effective plant breeding requires knowledge of *Musa* genome evolution and structure (Dolezel and Bartos, 2005), complemented

with functional genomics, allowing bioinformatics and gene discovery. Understanding the evolution of the banana genome through the DNA changes, hybridization and polyploidy in the past will allow prediction of how genomes will change in the future, whether under natural evolutionary processes, or the accelerated processes of plant breeding. New data are allowing us to exploit comparative genomics to gain an understanding of *Musa* genetics, and the complete DNA sequence of a *Musa* species would be valuable for both fundamental research and for application in future improvement of the crop, enabling access to all the genes and their control sequences. The genetic understanding and development of tools means there are prospects for 'new generations' of super-domestication through breeding technology involving sexual hybridization, mutation breeding and targeted transformation approaches.

ACKNOWLEDGEMENTS

We thank the Generation Challenge Programme for support, and some work was in collaborations under an IAEA/FAO Coordinated Research Project. We are extremely grateful to our many collaborators on the banana projects. Funding to pay the Open Access publication charges for this article was provided by the OECD.

LITERATURE CITED

Aert R, Sági L, Volckaert G. 2004. Gene content and density in banana (*Musa acuminata*) as revealed by genomic sequencing of BAC clones. *Theoretical and Applied Genetics* **109**: 129–139.

Ahloowalia BS, Maluszynski M, Nichterlein K. 2004. Global impact of mutation-derived varieties. *Euphytica* **135**: 187–204.

Angiosperm Phylogeny Group. 2003. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG II. *Botanical Journal of the Linnean Society* **141**: 399–436.

Assani A, Bakry F, Kerbellec F, Haicour R, Wenzel G, Foroughi-Wehr B. 2003. Production of haploids from anther culture of banana [*Musa balbisiana* (BB)]. *Plant Cell Reports* **21**: 511–516.

Atkinson HJ, Grimwood S, Johnston K, Green J. 2004. Prototype demonstration of transgenic resistance to the nematode *Radopholus similes* conferred on banana by a cystatin. *Transgenic Research* **13**: 135–142.

Balint-Kurti PJ, Clendennen SK, Dolezelovca M, Valarik M, Dolezel J, Beetham PR, May GD. 2000. Identification and chromosomal localization of the monkey retrotransposon in *Musa* sp.. *Molecular and General Genetics* **263**: 908–915.

Bartos J, Alkhimova O, Dolezelova M, De Langhe E, Dolezel J. 2005. Nuclear genome size and genomic distribution of ribosomal DNA in *Musa* and *Ensete* (Musaceae): taxonomic implications. *Cytogenetics and Genome Research* **109**: 50–57.

Baurens FC, Noyer JL, Lanaud C, Lagoda P. 1996. Use of competitive PCR to assay copy number of repetitive elements in banana. *Molecular and General Genetics* **253**: 57–64.

Baurens FC, Noyer JL, Lanaud C, Lagoda P. 1997. Copia-like elements in banana. *Journal of Genetics and Breeding* **51**: 135–142.

Bennett MD, Leitch IJ. 2004. *Plant DNA C-values database* (release 3·0, Dec. 2004). http://www.rbgkew.org.uk/cval/homepage.html. Accessed July 2007

Buhariwalla HK, Jarret RL, Jayashree B, Crouch JH, Ortiz R. 2005. Isolation and characterization of microsatellite markers from *Musa balbisiana*. *Molecular Ecology Notes* **5**: 327–330.

Carreel F, Gonzalez de Leon D, Lagoda P, Lanaud C, Jenny C, Horry J, Tezenas du Montcel H. 2002. Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome* **45**: 679–692.

Cheesman EE. 1947. Classification of the bananas. II. The genus *Musa* L. *Kew Bulletin* **2**: 106–117.

Cheung F, Town CD. 2007. A BAC end view of the *Musa acuminata* genome. *BMC Plant Biology* **7**:29. doi:10.1186/1471–2229–7–29.

Coemans B, H. Matsumura H, Terauchi R, Remy S, Swennen R, Sági L. 2005. SuperSAGE combined with PCR walking allows global gene expression profiling of banana (*Musa acuminata*), a non-model organism. *Theoretical and Applied Genetics* **111**: 1118–1126.

Conde-Ferráez L, Waalwijk C, Canto-Canché BB, Kema GHJ, Crous PW, James AC, Abeln ECA. 2007. Isolation and characterization of the mating type locus of *Mycosphaerella fijiensis*, the causal agent of black leaf streak disease of banana. *Molecular Plant Pathology* **8**: 111–120.

Creste S, Neto AT, Vencovsky R, Silva SO, Figueira A. 2004. Genetic diversity of *Musa* diploid and triploid accessions from the Brazilian banana breeding program estimated by microsatellite markers. *Genetic Resources and Crop Evolution* **51**: 723–733.

Crouch JH, Crouch HK, Tenkouano A, Ortiz R. 1999. VNTR-based diversity analysis of 2*x* and 4*x* full-sib *Musa* hybrids. *Electronic Journal of Biotechnology* **2**: 130–139.

Crouch HK, Crouch JH, Madsen S, Vuylsteke DR, Ortiz R. 2000. Comparative analysis of phenotypic and genotypic diversity among plantain landraces (*Musa* spp., AAB group). *Theoretical and Applied Genetics* **101**: 1056–1065.

Daly A, Walduck G. 2006. *Fusarium* wilt of bananas (Panama disease) Agnote 151. Australia: Northern Territory Government. https://transact.nt.gov.au/ebiz/dbird/TechPublications.nsf/1717DB3D7CD580C06925723C004EA6DD/$file/786.pdf?OpenElement. Accessed July 2007

Daniells JW, Jenny C, Karamura DA, Tomekpe K, Arnaud E, Sharrock S. (eds). 2001. *Musalogue: A catalogue of* Musa *germplasm. Diversity in the genus* Musa. Montpellier, France: INIBAP.

Dolezel J, Bartos J. 2005. Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany* **95**: 99–110.

Dolezel J, Doelzelova M, Novak FJ. 1994. Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *M. balbisiana*). *Biologia Plantarum* **36**: 351–357.

Dolezelova M, Valarik M, Swennen R, Horry J-P, Dolezel J. 1998. Physical mapping of the 18S–25S and 5S ribosomal RNA genes in diploid bananas. *Biologia Plantarum* **41**: 497–505.

Escalant JV, Sharrock S, Frison E. 2002. *The genetic improvement of Musa using conventional breeding, and modern tools of molecular and cellular biology*. Montpellier, France: INIBAP. http://www.promusa.org/research/conventional-modern.pdf. Accessed July 2007

European Commission. 1994. Commission regulation (EC) No 2257/94 of 16 September 1994 laying down quality standards for bananas. Brussels: European Commission, 1–7.

FAOStat. 2007. *FAOSTAT*. http://faostat.fao.org/. Accessed July 2007

Faure S, Noyer JL, Horry JP, Bakry F, Lanaud C, Gonzalez de Leon D. 1992. A molecular marker based linkage map of diploid bananas (*Musa acuminata*). *Theoretical and Applied Genetics* **87**: 517–526.

Faure S, Noyer J-L, Carreel F, Horry J-P, Bakry F, Lanaud C. 1994. Maternal inheritance of chloroplast genome and paternal inheritance of mitochondrial genome in bananas (*Musa acuminata*). *Current Genetics* **25**: 265–269.

Ferreira CF, Silva SO, Sobrihno NP, Paz OP. 2004. Molecular characterization of banana (AA) diploids with contrasting levels of black and yellow sigatoka resistance. *American Journal of Applied Sciences* **1**: 276–278.

Gibbs JR, Singleton A. 2006. Application of genome-wide single nucleotide polymorphism typing: Simple association and beyond. *PLoS Genetics* **2**:10, e150. doi:10.1371/journal.pgen.0020150.

Harper G, Osuji JO, Heslop-Harrison JS, Hull R. 1999. Integration of banana streak badnavirus into the *Musa* genome: molecular and cytogenetic evidence. *Virology* **255**: 207–213.

Harper G, Hull R, Lockhart B, Olszewski N. 2002. Viral sequences integrated into plant genomes. *Annual Review of Phytopathology* **40**: 119–136.

Heslop-Harrison JS. 1991. The molecular cytogenetics of plants. *Journal of Cell Science* **100**: 15–21.

Heslop-Harrison JS, Brandes A, Taketa S, Schmidt T, Vershinin AV, Alkhimova EG, *et al* 1997. The chromosomal distributions of

Ty1-*copia* group retrotransposable elements in higher plants and their implications for genome evolution. *Genetica* **100**: 197–204.

d'Hont A, Paget-Goy A, Escoute J, Carreel F. 2000. The interspecific genome structure of cultivated banana, *Musa* spp, revealed by genomic DNA *in situ* hybridization. *Theoretical Applied Genetics* **100**: 177–183.

Horry JP, Jay M. 1988. Distribution of anthocyanins in wild and cultivated banana varieties. *Phytochemistry* **27**: 2667–2672.

van den Houwe I, De Smet K, Tezenas de Montcel H, Swennen R. 1995. Variability in storage potential of banana shoot cultures under medium term storage conditions. *Plant Cell, Tissue and Organ Culture* **42**: 269–274.

Hull R. 2002. *Matthews' plant virology*, 4th edn. San Diego: Academic Press.

Hull R, Harper G, Lockhart B. 2000. Viral sequences integrated into plant genomes. *Trends in Plant Science* **5**: 362–365.

Hwang S-C, Ko W-H. 2004. Cavendish banana cultivars resistant to *Fusarium* wilt acquired through somaclonal variation in Taiwan. *Plant Disease* **88**: 580–588.

IPGRI–INIBAP/CIRAD. 1996. *Descriptors for banana (Musa spp.)*. Rome: IPGRI; Montpellier, France: INIBAP; Montpellier, France: CIRAD.

Jain SM, **Swennen R.** (eds)**. 2004.** *Banana improvement: cellular, molecular biology, and induced mutations*. Enfield, NH: Science Publishers. Available online at http://www.fao.org/docrep/007/ae216e/ae216e00.htm. 30 chapters. Accessed July 2007

Jarret RL, Litz RE. 1986. Isozymes as genetic markers in bananas and plantains. *Euphytica* **35**: 539–549.

Johannessen CL. 1970. The dispersal of *Musa* in Central America: the domestication process in action. *Annals of the Association of American Geographers* **60**: 689–699.

Jones CJ, Edwards KJ, Castiglione S, Winfield MO, Sala F, Van De Wiel C, *et al*. 1997. Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* **3**: 381–390.

Kalendar R, Grab T, Regina M, Suoniemi A, Schulman AH. 1999. IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. *Theoretical and Applied Genetics* **98**: 704–711.

Khanna H, Becker D, Kleidon J, Dale J. 2004. Centrifugation Assisted *Agrobacterium tumefaciens*-mediated Transformation (CAAT) of embryogenic cell suspensions of banana (*Musa* spp. Cavendish AAA and Lady finger AAB). *Molecular Breeding* **14**: 239–252.

Krikorian AD, Irizarry H, Cronaur-Mitra SS, Rivera E. 1993. Clonal fidelity and variation in plantain (*Musa* AAB) regenerated from vegetative stem and floral axis tips *in vitro*. *Annals of Botany* **71**: 519–535.

Kubis S, Schmidt T, Heslop-Harrison JS. 1998. Repetitive DNA elements as a major component of plant genomes. *Annals of Botany* **82**(Suuplement): 45–55.

Lysak MA, Dolezelova M, Horry JP, Swennen R, Dolezel J. 1999. Flow cytometric analysis of nuclear DNA content in *Musa*. *Theoretical and Applied Genetics* **98**: 1344–1350.

Mason HS, Warzecha H, Mor T, Arntzen CJ. 2002. Edible plant vaccines: applications for prophylactic and therapeutic molecular medicine. *Trends in Molecular Medicine* **8**: 324–329.

May GD, Afza R, Mason HS, Wiecko A, Novak FJ, Arntzen CJ. 1995. Generation of transgenic banana (*Musa acuminata*) plants via *Agrobacterium*-mediated transformation. *Bio/Technology* **13**: 486–492.

Moore NY, Bentley S, Pegg KG, Jones DR. 1995. Musa *Disease Fact Sheet No 5.* Fusarium *wilt of banana*. Montpellier, France: INIBAP.

Musagenomics 2007. http://www.musagenomics.org. Accessed July 2007

Nair AS, Teo CH, Schwarzacher T, Heslop-Harrison JS. 2005. Genome classification of banana cultivars from South India using IRAP markers. *Euphytica* **144**: 285–290.

Nwakanma DC, Pillay M, Okoli BE, Tenkouano A. 2003*a*. PCR-RFLP of the ribosomal DNA internal transcribed spacers (ITS) provides markers for the A and B genomes in *Musa* L. *Theoretical and Applied Genetics* **108**: 154–159.

Nwakanma DC, Pillay M, Okoli BE, Tenkouano A. 2003*b*. Sectional relationships in the genus *Musa* L. inferred from the PCR-RFLP of organelle DNA sequences. *Theoretical and Applied Genetics* **107**: 850–856.

Oh TJ, Cullis MA, Kunert K, Engelborghs I, Swennen R, Cullis CA. 2007. Genomic changes associated with somaclonal variation in banana (*Musa* spp.). *Physiologia Plantarum* **129**: 766–774.

Ortiz R. 1997*a*. Morphological variation in *Musa* germplasm. *Genetic Resources and Crop Evolution* **44**: 393–404.

Ortiz R. 1997*b*. Occurrence and inheritance of 2*n* pollen in *Musa*. *Annals of Botany* **79**: 449–453.

Ortiz R. 1997*c*. Secondary polyploids, heterosis, and evolutionary crop breeding for further improvement of the plantain and banana (*Musa* spp. L) genome. *Theoretical and Applied Genetics* **94**: 1113–1120.

Ortiz R, Peloquin SJ 1992. Recurrent selection for 2*n* gamete production in 2*x* potatoes. *Journal of Genetics and Breeding* **46**: 383–390.

Ortiz R, Vuylsteke D. 1995. Factors influencing seed set in triploid *Musa* spp. L. and production of euploid hybrids. *Annals of Botany* **75**: 151–155.

Ortiz R, Madsen S, Vuylsteke D. 1998. Classification of African plantain landraces and banana cultivars using a phenotypic distance index of quantitative descriptors. *Theoretical and Applied Genetics* **96**: 904–911.

Ortiz-Vazquez E, Kaemmer D, Zhang HB, Muth J, Rodrıguez-Mendiola M, Arias-Castro C, James A. 2005. Construction and characterization of a plant transformation-competent BIBAC library of the black Sigatoka-resistant banana *Musa acuminata* cv. Tuu Gia (AA). *Theoretical and Applied Genetics* **110**: 706–713.

Osuji JO, Harrison G, Crouch J, Heslop-Harrison JS. 1997. Identification of the genomic constitution of *Musa* L. lines (bananas, plantains and hybrids) using molecular cytogenetics. *Annals of Botany* **80**: 787–793.

Osuji JO, Crouch J, Harrison G, Heslop-Harrison JS. 1998. Molecular cytogenetics of *Musa* species, cultivars and hybrids: location of 18S-5·8S-25S and 5S rDNA and telomere-like sequences. *Annals of Botany* **82**: 243–248.

Pillay M, Ogundiwin E, Nwakanma DC, Ude G, Tenkouano A. 2001. Analysis of genetic diversity and relationships in East African banana germplasm. *Theoretical and Applied Genetics* **102**: 965–970.

Pollefeys P, Sharrock S, Arnaud E. 2004. *Preliminary analysis of the literature on the distribution of wild Musa species using MGIS and DIVA-GIS*. Montpellier, France: INIBAP. http://bananas.bioversityinternational.org/files/files/pdf/publications/wildspecies_pollefeys.pdf. Accessed July 2007

Raboin L-M, Carreel F, Noyer JL, Baurens FC, Horry JP, Bakry F, *et al* 2005. Diploid ancestors of triploid export banana cultivars: molecular identification of 2*n* restitution gamete donors and n gamete donors. *Molecular Breeding* **16**: 333–341.

Ray T, Dutta I, Saha P, Das S, Roy SC. 2006. Genetic stability of three economically important micropropagated banana (*Musa* spp.) cultivars of lower Indo-Gangetic plains, as assessed by RAPD and ISSR markers. *Plant Cell, Tissue and Organ Culture* **85**: 11–21.

Robinson JC. 1996. *Bananas and plantains.* Oxford: CAB International.

Rodríguez-Zapata LC, Castaño E, Acereto-Escoffié POM, Chi-Manzanero BH, Echeverría-Echeverría S, Grijalva R, James-Kay A, González-Estrada T. 2005. *Agrobacterium*-mediated transformation of *Musa acuminata* cv. 'Grand Nain' scalps by vacuum infiltration. *Scientia Horticulturae* **105**: 359–371.

Roux NS. 2004. Mutation induction in *Musa* – review. In: Jain SM, Swennen R, eds. *Banana improvement: cellular, molecular biology, and induced mutations*. Enfield, NH: Science Publishers. http://www.fao.org/docrep/007/ae216e/ae216e00.htm. Chapter 3. Accessed July 2007

Roux N, Dolezel J, Swennen R, Zapata-Arias FJ. 2001. Effectiveness of three micropropagation techniques to dissociate cytochimeras in *Musa* spp. *Plant Cell, Tissue and Organ Culture* **66**: 189–197.

Rowe P. 1984. Breeding banana and plantains. *Plant Breeding Reviews* **2**: 135–155.

Rowe P, Rosales F. 1993. Diploid breeding at FHIA and the development of Goldfinger. *InfoMusa* **2**: 9–11.

Safár J, Noa-Carrazana JC, Vrána J, Bartos J, Alkhimova O, Sabau X, Simková H, Lheureux F, Caruana M-L, Dolezel J, Piffanelli P. 2004. Creation of a BAC resource to study the structure and evolution of the banana (*Musa balbisiana*) genome. *Genome* **47**: 1182–1191.

Sági L, Panis B, Remy S, Schoofs H, De Smet K, Swennen R, Cammue BPA. 1995. Genetic transformation of banana and plantain (*Musa* spp.) via particle bombardment. *Bio/Technology* **13**: 481–485.

Santos CMR, Martins NF, Horberg HM, de Almeida ERP, Coelho MCF, Togawa RC, *et al* 2005. Analysis of expressed sequence tags from *Musa acuminata* ssp. *burmannicoides*, var. Calcutta 4 (AA) leaves submitted to temperature stresses. *Theoretical and Applied Genetics* **110**: 1517–1522.

Schmidt T, Heslop-Harrison JS. 1998. Genomes, genes and junk: the large scale organization of plant chromosomes. *Trends in Plant Science* **3**: 195–199.

Simmonds NW. 1962. *Evolution of the bananas.* London: Longmans, Green & Co.

Simmonds NW, Shepherd K. 1955. Taxonomy and origins of cultivated bananas. *Botanical Journal of the Linnean Society* **55**: 302–312.

Simmonds NW, Weatherup STC. 1990. Numerical taxonomy of the wild bananas (*Musa*). *New Phytologist* **115**: 567–571.

Sonnhammer ELL, Durbin R. 1995. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167**: GC1–10.

Stover RH, Simmonds NW. 1987. *Bananas*, 3rd ed.. Harlow, UK: Longman.

Strosse H, Van den Houwe I, Panis B. 2004. Banana cell and tissue culture – a review. In: Jain SM, Swennen R, eds. *Banana improvement: cellular, molecular biology, and induced mutations*. Enfield, NH: Science Publishers. http://www.fao.org/docrep/007/ae216e/ae216e00.htm. Chapter 1. Accessed July 2007

Su HJ, Hwang S-C, Ko W-H. 1986. Fusarial wilt of Cavendish bananas in Taiwan. *Plant Disease* **70**: 814–818.

Szymkowiak EJ, Sussex IM. 1996. What chimeras can tell us about plant development. *Annual Review of Plant Physiology and Plant Molecular Biology* **47**: 351–376.

Taxonomic Advisory Group for *Musa*. 2007. *Launching the Taxonomic Advisory Group: developing a strategic approach to the conservation and use of Musa diversity*. Montpellier, France: INIBAP. http://bananas.bioversityinternational.org/files/files/pdf/publications/launching_tag.pdf. See also http://tag.inibap.org/. Accessed July 2007

Tenkouano AJH, Crouch HK, Crouch J, Vuylsteke D, Ortiz R. 1999. Comparison of DNA marker and pedigree-based methods of genetic analysis of plantain and banana (*Musa* spp.) clones. I. Estimation of genetic relationships, *Theoretical and Applied Genetics* **98**: 62–68.

Teo CH, Tan SH, Othman YR, Schwarzacher T. 2002. The cloning of Ty 1-copia-like retrotransposons from 10 varieties of banana (*Musa* sp). *Journal of Biochemistry, Molecular Biology and Biophysics* **6**: 193–201.

Teo CH, Tan SH, Ho CL, Faridah QZ, Othman YR, Heslop-Harrison JS, Kalendar R, Schulman AH. 2005. Genome constitution and classification using retrotransposon-based markers in the orphan crop banana. *Journal of Plant Biology* **48**: 96–105.

Thangavelu M, James AB, Bankier A, Bryan G, Dear PH, Waugh R. 2003. HAPPY mapping in a plant genome: reconstruction and analysis of a high-resolution physical map of a 1·9 Mbp region of *Arabidopsis thaliana* chromosome 4. *Plant Biotechnology Journal* **1**: 23–31.

Tomlinson P. 1969. *Anatomy of the monocotyledons. III. Commelinales–Zingiberales*. Oxford: Clarendon Press.

Tsunewaki K, Ogihara Y. 1983. The molecular basis of genetic diversity among cytoplasms of *Triticum* and *Aegilops* species. II. On the origin of polyploid wheat cytoplasms as suggested by chloroplast DNA restriction fragment patterns. *Genetics* **104**: 155–171.

Ude G, Pillay M, Nwakanma D, Tenkouano A. 2002*a*. Analysis of genetic diversity and sectional relationships in *Musa* using AFLP markers. *Theoretical and Applied Genetics* **104**: 1239–1245.

Ude G, Pillay M, Nwakanma D, Tenkouano A. 2002*b*. Genetic diversity in *Musa acuminata* Colla and *Musa balbisiana* Colla and some of their natural hybrids using AFLP markers. *Theoretical and Applied Genetics* **104**: 1246–1252.

Valarik M, Simkova H, Hribova E, Safár J, Dolezelova M, Dolezel J. 2002. Isolation, characterization and chromosome localization of repetitive DNA sequences in banana (*Musa* spp.). *Chromosome Research* **10**: 89–100.

Vilarinhos AD, Piffanelli P, Lagoda P, Thibivilliers S, Sabau X, Carreel F, D'Hont A. 2003. Construction and characterization of a bacterial artificial chromosome library of banana (*Musa acuminata* Colla) *Theoretical and Applied Genetics* **106**: 1102–1106.

Vuylsteke D, Ortiz R, Ferris RSB, Crouch JH. 1997. Plantain improvement. *Plant Breeding Reviews* **14**: 267–320.

Waugh R, McLean K, Flavell AJ, Pearce SR, Kumar A, Thomas BBT, Powell W. 1997. Genetic distribution of BARE-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). *Molecular and General Genetics* **253**: 687–694.

Wong C, Kiew R, Argent G, Set O, Lee SK, Gan YY. 2002. Assessment of the validity of the sections in *Musa* (Musaceae) using AFLP. *Annals of Botany* **90**: 231–238.