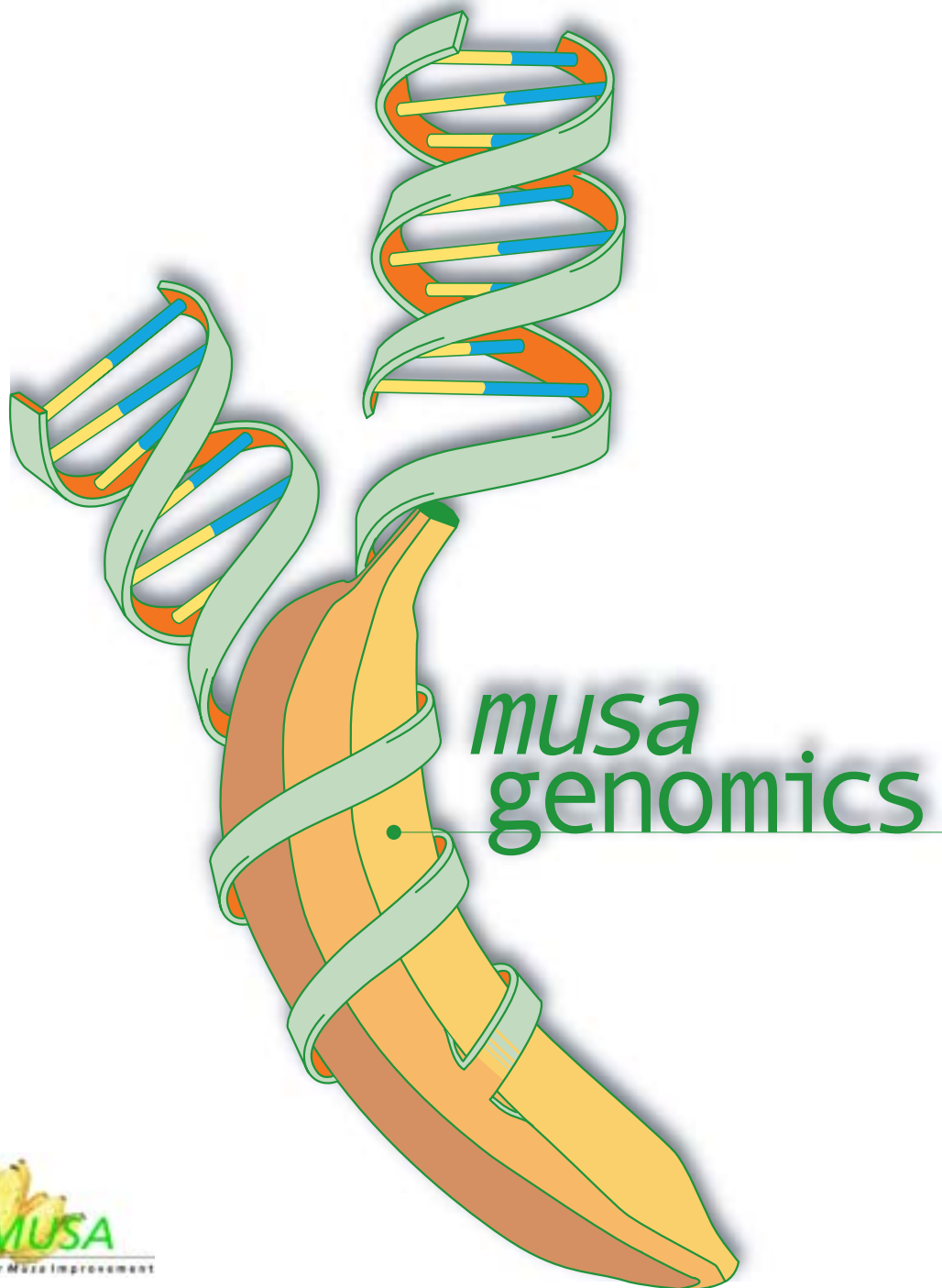




Strategy for the Global *Musa* Genomics Consortium

Report of a meeting held in Arlington, USA
17-20 July 2001
The Global *Musa* Genomics Consortium



The mission of the **International Network for the Improvement of Banana and Plantain** is to sustainably increase the productivity of banana and plantain grown on smallholdings for domestic consumption and for local and export markets.

The Programme has four specific objectives:

- To organize and coordinate a global research effort on banana and plantain, aimed at the development, evaluation and dissemination of improved cultivars and at the conservation and use of *Musa* diversity
- To promote and strengthen collaboration and partnerships in banana-related research activities at the national, regional and global levels
- To strengthen the ability of NARS to conduct research and development activities on bananas and plantains
- To coordinate, facilitate and support the production, collection and exchange of information and documentation related to banana and plantain.

INIBAP is a programme of the International Plant Genetic Resources Institute (IPGRI), a Future Harvest center.

The **International Plant Genetic Resources Institute** (IPGRI) is an autonomous international scientific organization, supported by the Consultative Group on International Agricultural Research (CGIAR). IPGRI's mandate is to advance the conservation and use of genetic diversity for the well-being of present and future generations. IPGRI's headquarters is based in Rome, Italy, with offices in another 19 countries worldwide. It operates through three programmes: (1) the Plant Genetic Resources Programme, (2) the CGIAR Genetic Resources Support Programme, and (3) the International Network for the Improvement of Banana and Plantain (INIBAP).

The international status of IPGRI is conferred under an Establishment Agreement which, by January 2000, had been signed and ratified by the Governments of Algeria, Australia, Belgium, Benin, Bolivia, Brazil, Burkina Faso, Cameroon, Chile, China, Congo, Costa Rica, Côte d'Ivoire, Cyprus, Czech Republic, Denmark, Ecuador, Egypt, Greece, Guinea, Hungary, India, Indonesia, Iran, Israel, Italy, Jordan, Kenya, Malaysia, Mauritania, Morocco, Norway, Pakistan, Panama, Peru, Poland, Portugal, Romania, Russia, Senegal, Slovakia, Sudan, Switzerland, Syria, Tunisia, Turkey, Uganda and Ukraine.

Citation:

The Global *Musa* Genomics Consortium. 2002. A strategy for the Global *Musa* Genomics Consortium. Report of a meeting held in Arlington, USA, 17-20 July 2001. The International Network for the Improvement of Banana and Plantain, Montpellier, France.

INIBAP ISBN: 2-910810-48-08

© International Plant Genetic Resources Institute, 2002

INIBAP
Parc Scientifique Agropolis 2
34397 Montpellier Cedex 5
France

IPGRI
Via dei Tre Denari 472/a
00057 Maccarese (Fiumicino)
Italy



Strategy for the Global *Musa* Genomics Consortium

Report of a meeting held in Arlington, USA

17-20 July 2001

The Global *Musa* Genomics Consortium

Acknowledgements

INIBAP would like to thank all the participants for their dynamic contribution to laying the foundations of the Global *Musa* Genomic Consortium at the meeting in Arlington.

The financial support from the National Science Foundation is gratefully acknowledged and we thank Chris Cullis, Jane Silverthorne, Machi Dilworth and Mary Clutter for their interest in this initiative. Sincere thanks are extended to Andrew Paterson for his efforts in mobilizing support and for administering the NSF grant. We would also like to thank the rapporteurs, Laura Silva, Diego González de León and Chris Town for capturing the essence of the discussions and Jean-Vincent Escalant for producing the initial draft report. Special thanks also go to Pat Heslop-Harisson for transforming the initial draft into an articulated strategy, to Suzanne Sharrock for her valuable inputs to the document and to Claudine Picq for supervising its publication. Finally, a word of appreciation for Emile Frison for his initiative in organizing the meeting and for facilitating the discussions and also for Charlotte Lusty, Future Harvest and Burness Communications for their successful contribution to raising the awareness of the general public about the launching of the Consortium.

Table of Contents

Introduction	5
Background	6
Aims of the Global <i>Musa</i> Genomics Consortium	7
Review of the status of <i>Musa</i> genomic research	8
A strategy for the <i>Musa</i> genomics programme	9
Deliverables	10
Costs and value of a <i>Musa</i> genomics programme	10
How genomics can be applied for <i>Musa</i> improvement	11
Resources for <i>Musa</i> genomics	12
Components of the incremental strategy for genomics	14
How to implement the strategy	15
Operation of the Consortium	17
Rules of the Consortium	17
Functioning	17
Composition of the Management Committee	18
Annex 1: Available resources	19
Segregating populations	19
Transcriptome	20
BACs	21
Transformation	22
Mutants	22
Annex 2: Background document:	
Beyond <i>Arabidopsis</i> and rice: <i>Musa</i> as a model for genomics	25
Annex 3: Programme of the Arlington meeting	31
Annex 4: Press release: Future Harvest	33
Global Consortium announces plans to sequence banana genome	33
Annex 5: Names and addresses of the participants	39
Acronyms and abbreviations	43

Introduction

The Global *Musa* Genomics Consortium was launched at a meeting held in Arlington, Virginia, USA from 17–20 July, 2001. This meeting was attended by scientists from 12 countries worldwide.

The meeting was critical for laying a solid foundation for future collaboration in *Musa* genomics and allowed the first important steps to be taken towards the development of a coherent strategy for *Musa* genomics. This included initiating the definition of tasks and their distribution among members of the Consortium.

The meeting also provided the opportunity to discuss the functioning of the Consortium and to agree on the terms of collaboration. This will facilitate the initiation of new collaborations, and the creation of synergies based on the complementarity of the skills available among the members and the diversity of interests of the participating public sector research institutions.

The creation of the Consortium has already raised considerable interest for *Musa* genomics and several partners have decided to increase their investment in this area, while others have decided to start investing. As a consequence of this meeting, the Consortium is now in a strong position to raise additional resources to substantially increase investment in research on this important crop.

This document provides further background information about the establishment of the Consortium and its aims and objectives. It also provides a review of the current status of *Musa* genomics research and provides details of the nature and scale of the work to be carried out by the Consortium members. Information is provided on an incremental strategy developed by the Consortium to achieve its goals and the proposed *modus operandi*, as agreed during the Arlington meeting. Further details are provided in the Annexes:

- The resources available to the Consortium (Annex 1),
- A background document 'Beyond *Arabidopsis*' (Annex 2),
- The programme of the Arlington meeting (Annex 3),
- The Press statement released at the time of the meeting (Annex 4),
- The list of participants at the meeting (Annex 5).

Background

Bananas (including plantains and various types of cooking bananas) are of major global importance in terms of food and income security to millions of smallholder farmers throughout the developing countries of the tropics and subtropics. Productivity increases are essential in the face of the growing populations in many of the countries where bananas provide a vital food source. Such increases in production must however, be brought about in the face of growing pest and disease pressure and ever-changing environmental and economic conditions.

The International Network for the Improvement of Banana and Plantain (INIBAP) was created in 1985 in response to the threat posed by the spread of a devastating fungal disease, black Sigatoka, to Africa. This disease, together with a range of other fungi, bacteria, insects, nematodes and viruses cause heavy losses to banana farmers worldwide, the majority of whom cannot afford chemical pesticides to control these parasites. INIBAP has always recognised the need for improved, pest and disease resistant varieties as the most effective and sustainable way of increasing yields, but is also aware of the difficulty of breeding such varieties in a crop where most of the cultivated varieties are highly sterile. Indeed, despite the efforts of INIBAP and its partners, progress in the development of improved varieties suitable for small-scale farmers has been slow. Recognizing this, the *Musa* research community agreed that a global-level initiative was required in order to accelerate the impact of improvement efforts. This culminated in 1997 with the formation of PROMUSA, the Global Programme for *Musa* Improvement. The major thrust of PROMUSA is to develop a wide range of improved *Musa* varieties, bringing together conventional breeding and biotechnology supported by research that is carried out on pests and diseases within the various working groups.

The formation of PROMUSA means that a mechanism for collaboration and information exchange between researchers involved in *Musa* genetic improvement is now in place. This has allowed the global prioritization of research needs and the acceleration of progress through the formation of synergistic partnerships. More than 100 researchers worldwide participate in the Global Programme.

Given the importance of genetic information for the sustainable improvement of *Musa*, a meeting of researchers wishing to apply the new and rapidly developing genomics technologies to *Musa* was held in France in 2000. At this meeting, it was agreed that deciphering the banana genome would be an enormous task that would require participation and collaboration of scientists around the world. In order to bring together and enhance the combined

expertise of the different participating laboratories, it was therefore agreed to form a “*Musa* Genomics Consortium” in the framework of PROMUSA.

The launch of the *Musa* Genomics Consortium took place in Arlington, Virginia in July 2001.

Aims of the Global *Musa* Genomics Consortium

The Consortium aims to apply genomics to the sustainable improvement of *Musa* (banana and plantain), a crop of world importance. The Consortium believes that newly available genomics technologies, which cover the analysis and sequencing of all the DNA, its genes, their expression, recombination and diversity, can now be applied directly to the sustainable improvement of this major crop. The Consortium aims to develop freely-accessible resources for *Musa* genomics, and use the new knowledge and tools to enable both targeted conventional breeding and transgenic strategies. The genomics strategy will also allow better utilization and maintenance of *Musa* biodiversity. Furthermore, knowledge of *Musa* genomics, a monocotyledon cultivated as a polyploid, will provide a model resource for the exploitation of the genomes of other important species, increasing the utility of all genomic information. The Consortium will benefit from the successes of current genome and genomics programmes and enabling technologies in delivering new genetic knowledge and tools, towards the development of new varieties. Information will be leveraged from the complete genomic sequences of *Arabidopsis* and rice, as well as the extensive sequence tags of other species, and these will be applied to *Musa* improvement. High throughput technologies, developed primarily for human genome analysis but available widely, will be accessed to put in place rapidly the resources and tools needed.

The overall aims of the Global *Musa* Genomics Consortium lie within the context of improvement of the world’s prosperity through fighting poverty and food insecurity in developing countries. The Consortium will achieve this aim by using the tools and expertise at its disposal to increase the productivity of the developing world’s fourth most important crop, a staple food and key cash crop for nearly a billion people. As a result of genomics research, banana and plantain productivity can be increased in ways which will remain sustainable, particularly in the face of changing economic, social and environmental conditions. The Consortium believes that such increases in productivity gained through fundamental knowledge and application of genomics will help to ensure future food and income security for millions of men, women and children in the developing world.

Review of the status of *Musa* genomics research

There is a very good knowledge of the structure of the *Musa* species complex based on morphological descriptors as well as molecular markers for the chloroplastic, mitochondrial and nuclear genome [work of the *Centre de coopération internationale en recherche agronomique pour le développement* (CIRAD) International Institute of Tropical Agriculture (IITA) and others]. Studies have revealed a great diversity in *Musa*, providing a good model for the study of gene regulation.

Genetic and physical maps on banana are being developed at CIRAD, France and the Institute of Experimental Botany (IEB), Czech Republic, respectively, which should allow the identification of specific genes of interest. Various ongoing research projects have allowed genetic molecular markers to be developed for *Musa*. The first genetic map on banana (shared within PROMUSA), based on a wild segregating population, is currently being constructed at CIRAD using molecular markers: RFLP, AFLP and microsatellites (STMS). Complementary work is ongoing at the Queensland Department of Primary Industries (QDPI) in Australia, where the banana segregating population from CIRAD is currently used to identify genes of resistance to *Fusarium* wilt. Segregating populations are being studied at IITA, Nigeria and *Centre africain de recherches sur bananiers et plantains* (CARBAP), Cameroon. Activities are also being developed at IEB to establish a physical map on banana using FISH, fibre-FISH, PRINS and chromosome painting.

BAC and EST libraries are already available in *M. acuminata*. A new BAC library for *M. acuminata* is being developed in the framework of the Agropolis Advanced Research Platform at Montpellier, France, where a researcher from EMBRAPA, Brazil is working at CIRAD on the mapping of *Musa acuminata* translocation break points implementing on banana, molecular cytogenetic methods such as *in situ* hybridization using BACS as probes.

A BAC library of *M. balbisiana* is also being developed at IEB in Czech Republic. The project aims at producing an ordered large-insert banana B-genome library, which will be highly complementary with the BAC library on A-genome.

Genetic transformation of bananas is now a routine at a number of institutes around the world including *Katholieke Universiteit Leuven* (KULeuven), Belgium; Queensland University of Technology (QUT), Australia; DNA Plant Technology Corporation (DNAP), USA; and Syngenta (Zeneca), UK. Two main methodologies are being used: particle bombardment and *Agrobacterium*-mediated transformation. The latter technique is considered to hold the

greatest potential as it allows the insertion of larger pieces of DNA and results in fewer insertion sites and fewer gene copies. Transient and stable expression of introduced genes has been observed in a range of cultivars. The sterile nature of many cultivars means that concerns over genetic manipulation in field-grown material are small. There is, in contrast to brassicas and cereals, very little chance of the spreading of introduced genes into wild species.

KULeuven in Belgium, in collaboration with the University of Queensland, Australia, has isolated promoters from various strains of banana streak badnavirus. These promoters have allowed a very high constitutive expression of reporter genes in banana as well as other monocot, and dicot species. A promoter from sugarcane bacilliform badnavirus, cloned and shown to infect banana at the University of Minnesota, was also found by the Minnesota, Belgian and Australian groups to be highly active in both monocots and dicots. The Belgian and Australian groups have also shown that this promoter is active in banana. At QUT, Australia, promoter regions derived from banana bunchy top nanovirus satellite components (S1 and S2) and banana actin genes have been isolated and characterized in transgenic banana plants.

The existence of several somatic mutants in *Musa*, a result of vegetative propagation, but mainly as mutants obtained after irradiation or using chemical mutagenic agents, creates a resource of potential value for functional genomics (relating gene sequence to function), and also for the study of evolutionary mechanisms.

Haploid plants of *M. acuminata* have been produced through anther culture and perfect homozygous diploid plants have been obtained through colchicine treatment. Tetraploid plants can be easily obtained through doubling of diploids.

The tools to create mono- or interspecific diploids and tetraploids are available, allowing a comparison of plants with the same genetic composition, but with different ploidy levels.

Several additional segregating populations are being established in support of the *Musa* genomics initiative. These efforts will allow the identification of specific genes of interest (pest and disease resistance etc.).

A strategy for the *Musa* genomics programme

In collective discussions, current resources available in *Musa* genomics, the types of resources and information available in other plant species and examples of the utilization of these resources were reviewed. With a focus on *Musa* an optimized strategy was defined to develop and utilize genomic tools

for rapid improvement of the crop, including the discovery of genes of agronomic and economic importance, and the measurement and utilization of the biodiversity in the genus.

Deliverables

The deliverables of the *Musa* genomics programme will be:

- The complete sequence of the *Musa* genome,
- The identification of each gene in the genome,
- The definition of the function of each gene in the genome,
- The complete map of gene expression – the transcriptome – during development and under various biotic and abiotic stresses,
- The definition of all the alleles of each gene present in all genotypes of the species,
- Development of usable genetic markers for major traits,
- The measurement of variation and diversity between all accessions.

These deliverables will enable the directed breeding of improved varieties of banana and plantains allowing environmentally friendly and sustainable production. The technology will allow transfer of genes freely between different accessions, maintaining their regulation, and also give the ability to transfer genes between other species and banana. The ability to manipulate bananas at the genome level has considerable potential for the improvement of the sterile, mostly triploid crop. A secondary aim, one which is in line with the goals of some funding agencies, would be provision of both genomics information and technology which will assist exploitation of genomes of other agronomically and medically important species; it is clear that *Musa* provides a well-justified platform for gaining an insight into the genomes of all species, increasing the generality and value of all genomic information. A separate document on this subject, “Beyond *Arabidopsis* and rice: *Musa* as a model for genomics” has been prepared by PROMUSA, the Global Programme for *Musa* Improvement (Annex 2).

PROMUSA offers an outstanding framework to assume worldwide leadership in the initiative on *Musa* genomics, whose activities will be developed within the *Musa* Genomics Consortium.

Costs and value of a *Musa* genomics programme

The coordinated programme to put the deliverables above within reach, including complete sequencing of the genome and transcriptome development,

was estimated to cost upwards of USD\$50 million. About half of this is the cost of sequencing the 600 Mbp *Musa* genome, with the rest divided among the development of mutants, transcriptome analysis, diversity measurement and the provision of accessible bioinformatic tools. Compared to the expenditure on *Arabidopsis*, *Caenorhabditis* and human, this figure is relatively small, reflecting the leveraging of existing information and the enormous efficiency and technology improvements made in these pioneering programmes. We have no doubt that the completion of a *Musa* genomics programme will make a major contribution to developing country agriculture and the quality of life of the billion people reliant on banana; hence it represents a most worthwhile and efficient expenditure of resources.

As noted in the press release (Annex 4), the Consortium will press for the initiation of a comprehensive genomics programme, to be completed over a five-year timeframe, by a coordinated multi-national Consortium.

Genomics is already a key to plant breeding and biodiversity utilization strategies, and there is little doubt that the goals will be reached in due course. However, members of the Consortium recognise that funding on the scale required might not be available over the next five years, so an incremental strategy was developed. The coordinated programme will enable access and application of markers and genes in the short term, while driving towards, in the longer term, the complete genomic information required, with minimal duplication of work.

How genomics can be applied for *Musa* improvement

In most crops, improvement and breeding methods are based on crossing elites and selecting the best performing progeny, a strategy impossible in the sterile *Musa* cultivars. Towards the end of the 20th century, a number of strategies to improve *Musa* using diploid and tetraploid lines are beginning to produce improved cultivars. Genomics involves different tools whose value is well proven in model species. New visions, resources and methodologies were discussed with relevance to their application in *Musa*.

In the short term, the development of molecular markers and marker-assisted selection (MAS) methods will without doubt improve the efficiency of selection of improved cultivars for defined traits such as pest and disease resistance, abiotic stress tolerance, quality and post-harvest fruit characteristics. The genetic maps now under construction will improve selectability of quantitative traits such as yield, and also allow better selection of parents for breeding programmes. Participants in Arlington agreed that more DNA markers are needed to analyse more individuals in less time.

Because we do not know how the environment affects the pattern of the expression of the genome, the development of new and more markers should be a way to predict how a particular genotype could respond and therefore to predict if a particular genotype is more or less adapted to a specific situation.

In the medium term (5 to 10 years), breeding should provide new varieties. Gene isolation and the analysis of mutants will allow characterization of critical alleles of genes of agronomic value and the relevant regulatory sequences. Transgenic approaches will offer the potential for direct transfer of these genes into current triploid cultivars, avoiding the difficulty of reconstructing already satisfactory varieties from crosses between unimproved and minimally selected lines.

High-throughput technologies are critical to major genomics programmes, enabling identification of all genes and gene functions in a species. While relatively expensive to set-up, because all genes are characterized in parallel the per-gene costs are substantially lower than the cost for isolation of particular targeted genes. Furthermore, there is no cost of 'failure' to isolate a particular gene, errors are less likely, work in particular genome regions is not duplicated between laboratories, and the inherent systematic approach means genes are not missed in the assay.

Resources for *Musa* genomics

All genetic and genomic work relies on access to genotype collections and intercrossed families. Key resources for genomics include large-insert recombinant DNA libraries, of which the most appropriate for *Musa* is made in BAC vectors. Complementing this, full sets of expressed gene libraries, cDNA libraries, are required to examine gene expression profiles in different tissues and conditions. Finally, mutants and transgenic plants (including of other species with *Musa* genes) are essential to define the function of many genes.

The creation of these resources, many of which are underway now, and the mechanisms to handle them involving robotics, are essential pre-requisites for a complete genomics programme. Once created, they become tools for the global Consortium to gain data.

Many resources are already available for *Musa*, and public access is assured by the functioning of INIBAP. These were reviewed by the participants and are listed in Annex 1 in detail. All the participants agreed in identifying additional resources needed and the ways to develop them, trying also to establish a timeframe. In contrast to other species, we start from a very strong position

with an effective international coordination in place and a good availability of the resources already developed.

The wide diversity in the genus *Musa* is the basic element that sustains production of this important crop, a situation unchanged but made efficient and accessible by the tools of genomics. Breeders use diversity to produce improved varieties, allowing the crop to be grown in a wide range of environments and to meet the varied needs of the millions of people who depend on it for food and income. INIBAP, entrusted by FAO, aims to conserve and make available this diversity for the benefit of present and future generations. As part of this effort, a major international germplasm collection is available. This resource is almost unparalleled in any other species. Some 1400 accessions of *Musa* are available from the International *Musa* Germplasm Collection, maintained by INIBAP at KULeuven, Belgium. The need for continued, open access, and fully indexed (pathogen-free) status was highlighted by the Consortium, and the continuing expansion of the collection with new accessions, particularly of wild and diploid material, was strongly supported.

There are already framework genetic maps based on genetic markers, large-insert DNA libraries (BACs) are becoming available, and populations of hybrids are being made for mapping and trait evaluation. Regarding resistance genes to pests and diseases, existing maps should allow isolation of genes involved in black Sigatoka resistance. Currently, a QTL for Sigatoka resistance has been anchored on a map developed at CIRAD and could be mapped allowing isolation of genes (or cluster of genes) involved in the resistance to the disease.

With respect to pathogens, two problems were identified: firstly, the presence of the banana streak pararetrovirus (BSV) within the genome itself. If disease can be spread through this mechanism, there will be potential for problems in distribution of germplasm. Secondly, wild *Musa* collections are made without their context of pathogens; hence devastating but unrecognised disease pathogens may be present in the location of origin, which could spread rapidly if a cultivar became widespread.

Research is required into the efficient strategies to utilize information from *Arabidopsis* and Rice genomic research. A full-normalized cDNA from different banana tissue could be isolated and made as a microarray for high-throughput genomic screening. CHIP could be created (one chip may contain 15 000 elements) and used on *Arabidopsis* or / and rice maps to find any similarity and therefore to identify new QTLs.

Components of the incremental strategy for genomics

Key resources are already becoming available, including the BAC libraries. A series of parallel projects are proposed, involving genomic DNA and cDNA in parallel, with additional work leading to the application of these results by breeding teams.

Physical maps

Following creation of the BAC library resources, physical maps require three different types of data: fingerprints, hybridization of data using anchored probes and cytogenetic data. The physical organization data can be obtained by individual groups working in partnership, and leads on to the sequencing programme.

Sequencing

The Consortium strongly recommended that a start be made as soon as possible on sequencing complete BACs (80-125 kb each) so that some information about the structure of large genomic tracts becomes available. Already, considerable information about other genomes is coming from such a strategy, and the results will feed into the complete and systematic *Musa* genomics project.

cDNA characterization

EST sequencing – the sequencing of the ends of expressed genes cloned as cDNA – has been used in many species, but new information is showing these data are not as valuable as full-length cDNA sequences. The data is however useful to search for genes, so it was concluded that a framework with relatively low coverage of the genome would be useful, which would be integrated with the BAC sequencing data. It was felt that data from both A or B genomes would be essential, although noting that in terms of breeding and sources of resistance genes to pests and diseases, the A genome is used much more by the different breeding programmes than the B genome. According to work carried out at IITA, AFLP analysis of the B genomes shows much more diversity than expected, while BB diversity, analysed using SSR derived from the A genome, is shown to be as variable as that observed in A. There are several B materials in Asia, and the uncertainty on the diversity of the B genome could be resolved by ongoing work in the Philippines and India throughout the next year.

Functional genomics / transcriptomics

Leading on from the cDNA characterization, the Consortium felt that production of a cDNA microarray would be very valuable and should be given high priority, so that the characterization of gene expression profiles could begin rapidly. Once available, the microarray would be exploited to reveal

genes involved in disease resistance in the first instance. These would then be sequenced and genomic clones obtained from the BAC libraries.

Mutagenesis

The Consortium recognised that approaches to gene discovery and functional analysis through mutation analysis (both insertion and deletion) are important components of the Consortium. A detailed proposal for the numbers of plants and types of neutron bombardment felt to be optimum was also prepared.

Bioinformatics

To ensure early availability of data from genomics, and exploitation of comparative data from other species, the Consortium aimed to set up a common central resource as entry point for the data. The PROMUSA web site (www.promusa.org) could host a data repository. Specific modules are needed to host specific tasks that could be combined within and between others; several participants have major international roles in genomic databases, including the Munich Information Center for Protein Sequences (MIPS) and The Institute for Genomic Research (TIGR). To ensure maximum exploitation of sequence data, state-of-the-art bioinformatics techniques will be applied and value added rapidly to all generated sequence by knowledge transfer from other species. To this end, major plant bioinformatics centers, e.g. TIGR and MIPS, are involved in the Consortium from the earliest stage, and will also contribute their experience in genomic database development. It was suggested that early work should consider database development and its nomenclature, including Diego González de León, Mexico; Andrew Paterson, USA; Jaroslav Dolezel, Czech Republic; Michael Pillay, IITA; Elizabeth Arnaud, INIBAP; Françoise Carreel, CIRAD.

How to implement the strategy

Different institutions have already expressed their interest in participating in the development of some of the activities defined in the strategy. For much of the genomic work, expansion and direction of the work is dependent on funding becoming available.

IITA is already developing segregating populations (Borneo x SF247; 6142 x 8075) and will communicate to the Consortium about the pedigree of the populations. CIRAD and INIBAP have also started with the development of different populations, aiming the segregation on three different characters: resistance to black leaf streak disease (black Sigatoka); resistance to nematodes (*R. similis*); and quality (parthenocarpy). The *Empresa Brasileira de Pesquisa Agropecuária* (EMBRAPA) in Brazil has provided to PROMUSA, through INIBAP and the *Corporación Bananera Nacional* (CORBANA), a population that

could show the segregation on nematode resistance. The F1 population is now being evaluated at CORBANA, Costa Rica. The group recommends contacting the National Research Center on Banana (NRCB) in India and the *Centro Nacional de Pesquisa de Recursos Geneticos y Biotecnologia* (CENARGEN)/EMBRAPA in Cruz das Almas, Brazil, for further involvement.

CIRAD, Leicester University, the *Centro de Investigación y de Estudios avanzados del IPN* (CINVESTAV) and the Indian Institute for Horticultural Research (IIHR) are interested to participate with the development of EST and cDNA libraries, fully microarray printed and normalized. Leicester University, UK would be happy to coordinate normalizing cDNA and printing of microarrays, with reading and hybridization facilities if required by Consortium members. However, funds will be required.

Two BAC libraries should be available within a year on both B and A genomes. These libraries should be ordered and characterized using fingerprinting and sequencing. To complement the development of 'B' library at IEB, the University of Minnesota offered to work on some of the 'BB' DNA to develop a 5x coverage BAC library on B genome. The group agreed that in the current situation, coverage of 6x would be sufficient. An ideal coverage of 10x will require additional funds.

TIGR offered to sequence a few BACs to see if there are any sequence features evident and start the programme of gene discovery. KULeuven will also start to sequence some BACs with their own resources. Other institutions expressed their interest in participating to the BAC sequencing. CENARGEN/*Universidade Catolica de Brasilia* (UCB) could make 30 000 reads/year with trained people and additional funds. Brazilian teams needs to be trained on micro-arrays. CINVESTAV is recognised by FAO (UNESCO) as a training centre and could facilitate the training course. FAO/International Atomic Energy Agency (IAEA) Agriculture and Biotechnology Laboratory is well known for arranging several interregional training courses. Recently, in October 2001, a training course on DNA markers was organized.

Cytogenetics and physical maps are an important part of the BAC strategy. The University of Georgia, Leicester University and IEB would continue with the development and implementation of different techniques especially related to physical mapping: chromosome and fibre stretching, flow cytometry.

An EU project proposal on comparative genomics on banana will be submitted to the next call for proposals of the 5th Framework programme, aiming to meet the externally defined goals of developing model species for genomics. This project involves seven different European institutions: KULeuven, CIRAD,

University of Leicester, MIPS/National Research Center on Environment and Health (GSF), IEB, FAO/IAEA and INIBAP.

Operation of the Consortium

Rules of the Consortium

Membership: Membership is on individual basis, with the possibility that one individual represents his institution. Membership will be submitted to the Management Committee, which will consult with the group. New membership will formally be accepted through a letter of invitation.

Members of the Consortium shall agree to:

- Share materials and resources that are relevant to the development of enabling technologies, where appropriate through the use of a Material Transfer Agreement (MTA) and / or after agreement about co-authorship to secure joint ownership of technology.
- To share within the group all the information obtained from a project funded in the framework of the Consortium. This information will be freely available to all the members.
- To facilitate access to infrastructure within the Consortium
- Immediately share among members of the Consortium all sequencing data which will be published and placed in the public domain.
- To avoid as much as possible, the submittal of competing project proposals.

Results that have a possible commercial application and for which the inventors want to take a patent protection, will be made available through a royalty-free non-exclusive licence for applications in developing countries by smallholder farmers. The members of the Consortium will also be allowed to use these results for research purposes.

The Consortium will negotiate access to enabling technologies with the private sector.

Functioning

To facilitate the functioning of the Consortium, a Management Committee and a Secretariat were nominated. An external Advisory Committee may be necessary as the Consortium grows and access to funding increases: this would provide guidance and expertise, facilitating the Consortium with new opportunities of funding, new alliances, and also advise the Consortium on

IPR and related aspects. Currently this role will be carried out by the Management Committee and seconded members of the Consortium.

The Management Committee will consist of up to seven members and will act as an internal body to evaluate new candidatures, ideas and project proposals. The Committee will elect a chair and each member will represent the Consortium in scientific meetings or contacts with donors. The Committee calls the meetings. They will advise and act as a sounding vote with the Secretariat. The members of the Committee also agreed to write up the rules and will elect their own chair.

Composition of the Management Committee

Jaroslav Dolezel	(Chair of the Genetic improvement working group of PROMUSA), IEB, Czech Republic
Charles Arntzen	University of Arizona, USA
Françoise Carreel	CIRAD, France
James Dale	QUT, Australia
Michael Pillay	IITA, Uganda
Laszlo Sagi	KULeuven, Belgium
Jean-Philippe Vielle-Calzada	CINVESTAV, Mexico.

INIBAP has been nominated and agreed to act as the Secretariat to the Consortium. INIBAP will assist the Consortium, ensuring a regular exchange of information within the group. INIBAP will facilitate the development of an Internet web site-platform to link the different databases. The secretariat will also assume responsibility with regard to press releases and will facilitate the development of a listserver for the Consortium. INIBAP as a secretariat, will assist the Management Committee in organizing future meetings, and gathering information on potential donors and partners.

Annex 1: Available resources

Segregating populations

Different segregating populations already exist (or are in the process of) with different levels of accessibility. However, all the existing populations are from the A genome resulting from crosses between 2x X 2x; 2x X 3x. There is no cross that involves the B genome. The diversity of B genome will be completed and available within a year. CIRAD (Station of Neufchâteau in Guadeloupe) could

Issue	Resources	Contributor(s) Contact(s) and notes	Concerns	Solutions recommendations comments
Segregating populations	Eight populations will become available over the next 12-18 months.		Indexing takes one year and is very costly. Most currently at F1 level.	DNA and mapping can be performed while indexing goes on. Indexing would be limited to the mother F2 plants from which sets of complete populations would be generated through micropropagation.
Population to study BLS resistance Autohybridization F1 hybrids <i>M. acuminata truncata</i> (HR) x <i>M.a. madang</i> (S) – TRUMA.		CIRAD INIBAP	Ideally would need multiple sites to score polygenic traits in addition to Costa Rica (India, Brazil, Mexico, ...).	
Population to study nematodes resistance (R. <i>similis</i>) <i>M.a. pahang</i> (AA) x Pisang Jari Buaya (AA). <i>M.a. pahang</i> (AA) x Yangambi Km5 (AAA).				
Population to study parthenocarpy Calcutta 4 (AA) x Pisang lilin (AA). Calcutta 4 (AA) x IDN 110 (AA).				
AFCAM Characterized for BLS (CARBAP, CIRAD).		CARBAP K. Tomekpe		Use of bioassays <i>ex-situ</i> for disease scoring.
IITA segregating populations Two segregating populations at IITA: SF247 (Onne) x Borneo and 6142 x 8075-diploid x diploid cross.		IITA M. Pillay	Check pedigree and mating scheme.	Start parental screenings to verify quality of crosses.

start with crosses using B genome. Most of the participants agreed on the interest/importance to have AB and BB segregating populations. IITA-Uganda has 4 F1 plants from an AA X BB cross. IIHR will also contribute to the development of STMS markers for the B genome.

The evaluation of the segregating populations in different locations around the world requires great effort and coordination. The variation in the comportment of the individuals in different environments could reflect a difference in the measurement of the field traits rather than a real variation due to biotic or abiotic factors. It was suggested to establish three different sites around the world: Africa, Latin America and Asia. However, because parthenocarpy is not affected by the environment, only one site should be established in Latin America to evaluate this characteristic (CORBANA, Costa Rica).

Transcriptome

cDNA libraries already exist, and may need to be normalized. CIRAD, Leicester University, IIHR, University of Queensland and CINVESTAV could be involved in the development of a full-length library and its sequencing.

Resources	Contributor(s) Contact(s) and notes	Concerns	Solutions recommendations comments
Buy?		SYNGENTA - 70 000 at \$1 M, but may be poor.	Probably not desirable at this stage (see below).
Need full length clones. Tissue types Developmental stages, etc. Normalized (subtractive) libraries.		Coordination of complementary efforts. Quality control. DB development and tracking of each clone.	Very desirable: full length, fully normalized, fully representative cDNA library. No need to sequence at first as tremendous amounts of data can be recovered by printing this onto microarrays and probing with various sources generated from projects studying specific traits.
ESTs / cDNAs	Univ. Queensland CIRAD Un. Leicester IIHR CINVESTAV		30 000

BACs

Different BAC libraries will be available in the framework of *PROMUSA*. CIRAD, in France, is developing a BAC library on the A genome while IEB in Czech Republic is working on the B genome. A 6x coverage would be enough. Scaling up to 10x will require additional funding. There is already a BAC library available at Texas & AM University (TAMU), USA. It is 4x coverage (111-120 kb). BAC libraries should be fingerprinted, ABC end sequenced and linked to a physical map. The study of the polymorphisms between two BAC libraries from Calcutta 4 should allow assessing the extent of the polymorphism.

Resources	Contributor(s) Contact(s) and notes	Concerns	Solutions recommendations comments
Two to be available within a year? One is available for characterization now (D. Kaemmer). Developed with two RE's. Fingerprinting and quality control urgent (as they are developed?)	CIRAD IEB	Quality control... Distribution.	Coverage? 6x? Scaling up would require additional funding, especially in the case of B.
Molecular cytogenetics.	IEB Univ. Leicester		Fibre/chromosome stretched molecules for local ordering.
Hybridization based data.			Scaffolding hybridization framework for monocots being developed by A. Paterson.
Physical maps End sequencing Other large inserts from triploid genotypes.			Interest from TIGR if funding is available.

Transformation

Genetic transformation can be used as tool for genomics. Different institutions already expressed their interest to participate in the transformation of diploid cell suspensions (QUT and the University of Queensland, Australia; KULeuven, Belgium; International Laboratory for Tropical Agricultural Biotechnology (ILTAB), USA. Diploid cell suspensions are already available at CIRAD (IDN 110 and Galeo) and new suspension using Calcutta 4 could be developed.

Resources	Contributor(s) Contact(s) and notes	Concerns	Solutions recommendations comments
Role of Leuven as a reference center.		Capacity to take on multiple requests. Only suspension cultures of triploid A.	1000 new transgenic lines per year could be scaled up to 10,000 given additional resources. Calcutta IV suspension culture to be established as soon as possible. Urgent need to develop more efficient techniques, preferably avoiding suspension cultures, on 2x materials.
Development of alternative techniques.		Coordination with insertional mutant production work.	C. Fauquet J. Dale

Mutants

Mutants are considered important in the study of genomics and identification of genes. However, mutants derived from radiation are not tagged. A suggestion has been made on the use of *in situ* hybridization. Both the array technology to clone the 'hole' in the DNA sequences and the use of fast neutron bombardment to create mutants to study 'repetitive DNA sequences' is very promising. Fast neutron irradiation induces a variety of phenotypes, but it includes a high proportion of aneuploids. It has been suggested to map all interested phenotypes. Any deletion due to fast neutron bombardment will be analysed on micro-arrays.

Tagging techniques can be useful once applied on fast neutron irradiated plant material to trap promoters and genes using T-DNA and to trap enhancers and genes through the introduction of transposable element systems (Ac/Ds).

The use of two different clones: one susceptible and one resistant to BLSA has been suggested (Calcutta 4 as resistant and Pisang madang as susceptible). Seeds could be irradiated and the progenies would be screened to detect any deviated phenotype.

The installation of a mutant germplasm repository (MGR) at the FAO/IAEA Laboratory will facilitate the exchange of plant mutants material between scientists.

Resources	Contributor(s) Contact(s) and notes	Concerns	Solutions recommendations comments
Fast neutron irradiation and/or Gamma ray irradiation.	FAO/IAEA and KULEuven.	IAEA support Seed production for irradiation. Size of M2 population.	No problem in seed production of Calcutta IV. Microarray based deletion mapping makes fast neutron irradiation more desirable than EMS mutagenesis.
Insertional.	Interest of KULEuven and QUT.	Confirm transposable element utility. Other approaches (interference RNA, ...).	

Annex 2: Background document: Beyond *Arabidopsis* and rice: *Musa* as a model for genomics

***Musa* as a model for genomics**

So far, *Arabidopsis* and rice have been used as model species in plant genomics because of their small genome size and ease of handling, but both species have their limitations. Complementing these two species, the *Musa* genome provides a powerful platform for gaining a fundamental insight into the genomes of other species. Both in structural genomics and in functional genomics, *Musa* can be a model for several fundamental aspects for which the other two model species cannot be used.

The *Musa* genome offers new perspectives in the monocotyledons, with its relatively small haploid genome of 500 to 600 Mb (only 25% larger than rice) divided among 11 chromosomes. Because of its relatively small size, the *Musa* genome is highly tractable to complete functional and sequence analysis and extensive characterization of the genes will be possible using realistic high-throughput technologies.

Musa offers an interesting model for genetic studies, as it is one of the few plant species with bi-parental cytoplasmic inheritance: paternal inheritance of mitochondria and maternal inheritance of chloroplasts.

In the centre of origin and diversity in Southeast Asia there are many sterile clones that have been genomically static or fixed for thousands of years of vegetative propagation in the same environment. There are also partially fertile and highly fertile wild diploid equivalents that have been actively evolving for the same period in the same environment, ranking *Musa* among one of the most perfect models to study plant evolution at a genomic level. Moreover, they have co-evolved for the same period and in the same environment with most of the *Musa* pathogens. Both of these co-evolutions reinforce the position of *Musa* as a perfect model to study plant and pathogen evolutions at a genomic level.

About 3000 years ago, a few plantain varieties were introduced from the centre of origin into Africa where they underwent secondary diversification exclusively through mutations, in the absence of most of their pathogens. This provides also a valuable tool for the study evolution in the presence and absence of pathogens.

As a vegetatively propagated crop, it offers a good model to study the role of somaclonal variation and phenomena such as “imprinting”.

The variability in the different ploidy levels in *Musa* offers a very special opportunity to gain insight to the greater-than-additive gains in crop productivity that often accompany polyploidy. Even the most studied polyploids like cotton and sugarcane contain only one 'type' of polyploidy within the taxon, whereas *Musa* includes a number of autopolyploids (AAA, AAAA, AAAAAA) and different types of allopolyploids (allotriploids AAB, ABB and allotetraploids AABB, AAAB) in addition to the diploid *M. acuminata* (AA) and *M. balbisiana* (BB) and AB hybrids. Bananas and plantains are attractive models to study the role of hybridization and polyploidy in the evolution of cultivated crops, and to analyze the interaction of parental genomes at chromosomal and sequence level.

The diverse arrays of polyploid and diploid genomes in *Musa* create opportunities not only to study the relationship of polyploid formation to phenotype, but also the causes and consequences of polyploidy for genome organization.

The combination of parthenocarpy and sterility that has led to the typical edible banana is of basic interest in a wide range of other fruit crops, and is especially unusual in that there are relatively few parthenocarpic monocots.

The small size of the *Musa* genome and the different ploidy levels of the existing crops and species within *Musa*, as well as the combination of parthenocarpy with sterility and vegetative propagation, place the banana among one of the few opportunities to study the specific gene expression in different chromosomal environments.

Being a monocotyledon but taxonomically very distantly related to rice, *Musa* is an ideal candidate for studying synteny between distantly related species.

Musa was the first species where a pararetrovirus was shown to be integrated in the plant genome with the capacity to give rise to episomal banana streak badnavirus. Understanding the mechanism behind this phenomenon may lead to important applications, such as gene targeting.

State of the art and available tools for *Musa* genomics

A genebank holding more than 90% of the existing diversity of cultivated bananas and a representative sample of the wild species of *Musa* is maintained by INIBAP at the International *Musa* genebank in Leuven, Belgium. The collection has been tested for bacteria and viruses and all varieties that tested negative are available for distribution.

There is a very good knowledge of the structure of the *Musa* species complex based on morphological descriptors as well as molecular markers for the

chloroplastic, mitochondrial and nuclear genome (work of CIRAD, IITA and others). Studies have revealed a great diversity in *Musa*, providing a good model for the study of gene regulation.

Genetic and physical maps on banana are being developed at CIRAD, France and the Institute of Experimental Botany (IEB), Czech Republic, respectively, which should allow the identification of specific genes of interest. Various ongoing research projects have allowed genetic molecular markers to be developed for *Musa*. The first genetic map on banana (shared within PROMUSA), based on a wild segregating population, is currently being constructed at CIRAD using molecular markers: RFLP, AFLP and microsatellites (STMS). Complementary work is ongoing at QDPI in Australia, where the banana segregating population from CIRAD is currently used to identify genes of resistance to Fusarium wilt. Segregating populations are being studied at IITA, Nigeria and CRBP, Cameroon. Activities are also being developed at IEB to establish a physical map on banana using FISH, fibre-FISH, PRINS and chromosome painting.

BAC and EST libraries are already available in *M. acuminata*. A new BAC library for *M. acuminata* is being developed in the framework of the Agropolis Advanced Research Platform at Montpellier, France, where a researcher from EMBRAPA, Brazil is working at CIRAD on the mapping of *Musa acuminata* translocation break points implementing on banana, molecular cytogenetic methods such as *in situ* hybridization using BACS as probes.

A BAC library of *M. balbisiana* is also being developed at IEB in Czech Republic. The project aims at producing an ordered large-insert banana B-genome library, which will be highly complementary with the BAC library on A-genome.

Genetic transformation of bananas is now a routine at a number of institutes around the world including *Katholieke Universiteit Leuven* (KULeuven), Belgium; Queensland University of Technology (QUT), Australia; DNA Plant Technology Corporation (DNAP), USA; and Syngenta (Zeneca), UK. Two main methodologies are being used: particle bombardment and *Agrobacterium*-mediated transformation. The latter technique is considered to hold the greatest potential as it allows the insertion of larger pieces of DNA and results in fewer insertion sites and fewer gene copies. Transient and stable expression of introduced genes has been observed in a range of cultivars. The sterile nature of many cultivars means that concerns over genetic manipulation in field-grown material are small. There is, in contrast to brassicas and cereals, very little chance of the spreading of genes into wild species.

KULeuven in Belgium, in collaboration with the University of Queensland, Australia, has isolated promoters from sugarcane bacilliform badnavirus that is able to infect bananas. The promoter has allowed a very high constitutive expression of reporter genes in banana as well as other monocot and dicot species. At QUT, Australia, promoter regions derived from banana bunchy top nanovirus satellite components (S1 and S2) and banana actin genes have been isolated and characterized in transgenic banana plants.

The existence of several somatic mutants in *Musa*, a result of vegetative propagation, but mainly as mutants obtained after irradiation or using chemical mutagenic agents, creates a resource of potential value for functional genomics (relating gene sequence to function), and also for the study of evolutionary mechanisms.

Haploid plants of *M. acuminata* have been produced through anther culture and perfect homozygous diploid plants have been obtained through colchicine treatment. Tetraploid plants can be easily obtained through doubling of diploids.

The tools to create mono- or interspecific diploids and tetraploids are available, allowing a comparison of plants with the same genetic composition, but with different ploidy levels.

Several additional segregating populations are being established in support of the *Musa* genomics initiative. These efforts will allow the identification of specific genes of interest (pest and disease resistance, etc.).

Institutional framework

A mechanism for collaboration and information exchange between researchers involved in *Musa* genetic improvement has been put in place in the form of PROMUSA, the Global Programme for *Musa* Improvement. This has allowed the global prioritization of research needs and the acceleration of progress through the formation of synergistic partnerships. More than 100 researchers worldwide participate in this Global Programme.

In the framework of PROMUSA, a “*Musa* Genomics Consortium” has been formed in order to increase collaboration in genomics research. Deciphering the banana genome is an enormous task that will require participation and collaboration of scientists around the world. The *Musa* Genomics Consortium will bring together and enhance combined expertise of the different participating laboratories.

Socioeconomic importance of bananas

Besides the fact that *Musa* is a powerful model for plant genomic studies, it is also one of the most important staple food crops in the developing world and an important source of income millions of poor in the humid tropics as illustrated by the following:

- Bananas are the developing world's fourth most important food crop (after rice, wheat and maize). The crop is grown in more than 100 countries throughout the tropics and sub-tropics.
- Bananas as a crop consist of a wide range of varieties, including both cooking and dessert types. Plantains, which are a specific type of cooking banana found mainly in West Africa and Latin America, are included within the banana group.
- Annual world production is around 88 million tonnes, of which around a third is produced in each of the African, Asia-Pacific and Latin American and Caribbean regions.
- Around 87% of all the bananas grown worldwide are produced by small-scale farmers for home consumption or for sale in local and regional markets.
- Bananas provide a staple food for millions of people, particularly in Africa, an area where the green revolution has had little influence.
- Bananas are an important food security crop, providing a cheap and easily-produced source of energy. In addition, they are rich in certain minerals and in vitamins A, C and B6.
- Growing urbanization in many developing countries means that the crop is becoming more and more important as a source of revenue, sometimes providing the main source of income for rural communities. Bananas thus play an important role in poverty alleviation.
- Bananas will grow in a range of environments and produce fruit year-round, thus providing a source of energy during the 'hungry-period' between the harvest of other crops.
- Bananas are particularly suited to intercropping systems and to mixed farming with livestock. Due to their suitability for production in backyard systems, bananas are also an important component of periurban agriculture.
- When grown in perennial production systems, bananas maintain cover throughout the year, thus protecting the soil from rain and wind erosion. Furthermore, if their biomass is used as mulch, soil fertility and organic matter remain stable.

- Approximately 13% of worldwide banana production is destined to the export market. For many countries, especially in Latin America and the Caribbean, bananas provide an essential source of foreign exchange. The value of banana exports well outranks those of other fruits, such as apples and oranges, as well as vegetables, such as tomatoes and potatoes. Banana exports are worth nearly \$1 billion annually for Ecuador, the world's largest banana exporter.
- The export banana industry is also the backbone of the economies of many Caribbean countries, and the crop plays a vital role in the social and political fabrics of the islands, including the French departments of Guadeloupe and Martinique. The banana crop is often the sole source of income for rural communities in the Caribbean. In this region, bananas are the only year-round crop that can be viably cultivated to produce a regular weekly income for small-scale farmers. It is the only crop that can produce again within months of damage or destruction by storm, floods or hurricanes, which are perennial Caribbean hazards.
- In many countries, bananas are more than just a food crop. They provide an important source of fibre (for example Abaca/Manila hemp in the Philippines), and among other uses, can be fermented to produce alcohol.
- Bananas have been considered as a useful tool to deliver edible vaccines. The fruit can be eaten uncooked; it is sterile before peeling and it is often the first solid food eaten by babies.

Annex 3. Programme of the Arlington meeting

Tuesday 17 July 2001

09.00	Introduction Emile Frison, Chair of <i>PROMUSA</i> NSF representatives
09.30 – 12.30	Presentations from the participants J. Dolezel, IEB, Czech Republic M. Pillay, IITA R. Miller UCB, Brazil P. Lagoda, CIRAD, France L. Silva, CINVESTAV, Mexico D. Kaemmer, CICY, Mexico P. Heslop-Harrison, Univ. of Leicester, UK L. Anand, IIHR, India L. Sagi, KULeuven, Belgium H. Schoof, MIPS, Germany N. Roux, IAEA, Austria A. Paterson, Univ. of Georgia, USA M. Souza, CENARGEN/EMBRAPA, Brazil N. Olszewski, Univ of Minnesota, USA C. Town, TIGR, USA
12.30	Lunch
PM	Forum discussion

Wednesday 18 July 2001

08.30 – 09.00	Presentations from the participants
	G. Volckaert, KULeuven, Belgium
AM	Strategy for <i>Musa</i> Genomics
PM	Discussion on functioning of the Consortium (rules)

Thursday 19 July 2001

AM	Workplan / Division of tasks
PM	Project proposals working groups meetings

Friday 20 July 2001

AM	Project proposals working groups meetings (continued)
----	--

End of the meeting

Annex 4. Press release: Future Harvest Global Consortium announces plans to sequence¹ banana genome

New data could benefit small-scale banana farmers worldwide and reduce reliance on chemicals in commercial production

Arlington, Virginia, 19 July 2001 – Scientists from 11 countries today announced the founding of an international Consortium to sequence the banana genome within five years. The scientists from governmental, university, and nonprofit organizations will use the new genetic data to enable developing-world farmers to grow bananas that are able to resist the fungus “black Sigatoka,” as well as other diseases and pests. Bananas are a staple food for nearly half a billion people worldwide, but their crops are increasingly lost to disease. The genome sequence will also benefit US and European consumers of the popular dessert banana, one of the world’s most chemically dependent crops.

“Ancient farmers selected banana strains that were seedless and thus sterile, and grew the fruit through vegetative sprouting,” said Emile Frison, PhD, director of the Montpellier, France-based International Network for the Improvement of Banana and Plantain. “Cultivated bananas have, therefore, been at a near evolutionary standstill for thousands of years and lack the genetic diversity needed to fight off disease. A coordinated effort by scientists worldwide is needed to unlock the diversity found in bananas that still grow and reproduce in the wild.”

The group sequencing effort is being launched at a meeting held 17-19 July at the U.S. National Science Foundation in Arlington, Virginia. The International Network for the Improvement of Banana and Plantain (INIBAP), a programme of the Rome-based International Plant Genetic Resources Institute (IPGRI), is leading the effort, which brings together organizations from Australia, Belgium, Brazil, the Czech Republic, France, Germany, India, Mexico, the United Kingdom, and the United States. The newly founded “Global *Musa* (Banana) Genomics Consortium” includes the International Institute for Tropical Agriculture (IITA) based in Nigeria. IPGRI and IITA are Future Harvest Centers. The Consortium also includes the Institute for Genomic

¹ The headline phrase ‘sequencing the banana genome’ was used as shorthand for application of all aspects of genomics, including genome analysis, transcriptomics, gene discovery and breeding technologies, described elsewhere in this document.

Research (TIGR), which previously collaborated in sequencing the genomes of rice and *Arabidopsis* (a plant in the mustard family), as well as sequencing the parasite that causes East Coast fever—a leading cause of death in African cattle. Scientists will map the banana genome using a sexually reproducing wild species of banana from Southeast Asia.

“Banana will be the first exclusively tropical crop to be sequenced,” said Frison. “More than a popular snack, bananas are a staple food that many African families eat for every meal. This is our chance to develop a crop that won’t fail for them and that may help lift them out of hunger and poverty.”

Cutting back on agricultural chemicals

Farmers in 120 countries grow bananas and plantains. Plantains are long, green bananas—one of six major groups of cooking bananas—found mostly in West Africa and Latin America. Of the 95 million metric tons of bananas grown annually, approximately one-third is produced in each of Latin America, Africa, and Asia. Some 85 percent of the global crop is produced for home consumption and local trade, largely without the use of pesticides, leaving them highly susceptible to disease. The 15 percent of the global banana crop grown for export relies heavily on chemical inputs.

Bananas and plantains together are the developing world’s fourth most important food crop, following rice, wheat, and corn. In parts of Africa, bananas provide more than one-quarter of all food calories. When ripe, most banana types are not sweet like the imported dessert Cavendish bananas eaten in Europe and North America, but starchy like a potato and eaten cooked. Banana varieties, all grouped under the scientific name of *Musa*, are rich in vitamins A, C, and B6 and contain high levels of calcium, potassium, and phosphorus, providing an essential source of nutrition in developing countries.

Bananas are threatened by the rapidly spreading fungus black Sigatoka that has been undermining banana production for the past three decades. It has reached almost every banana-growing region in the world and typically reduces yield by 30 to 50 percent. Other diseases and pests that cripple yields include a soil fungus, parasitic worms, weevils, and viruses such as the Banana Streak Virus, which lurks inside the banana genome itself.

Commercial growers can afford and rely extensively on chemical fungicides, often spraying their crops 50 times per year—the equivalent of spraying nearly once per week, which is about 10 times the average for intensive agriculture in industrialized countries. Chemical inputs account for 27 percent of the production cost of export bananas. Agricultural chemicals used on bananas for

diseases and pests have harmed the health of plantation workers and the environment.

“If we can devise resistant banana varieties, we could possibly do away with fungicides and pesticides all together,” said Frison. “In addition, resistant strains are essential for small-holder farmers, who cannot afford the expensive chemicals to begin with. When black Sigatoka strikes, farmers can do little more than watch their plants die. Increased hunger can swiftly follow.”

Banana genome to reveal secrets of plant evolution

Following rice and *Arabidopsis*, the banana will be only the third plant sequenced. Comprised of just 11 chromosomes with a total of 500 000 000 to 600 000 000 base pairs, the banana genome is among the smallest of all plants and researchers expect quick results.

“If we’ve learned anything from genomics, it is how little we know about biology,” said Claire Fraser, PhD, president of TIGR in Rockville, Maryland. “We expect that the banana genome sequencing will reveal surprising insights into the evolution of plants.” J. Craig Venter, PhD, head of the human-genome sequencing company Celera, is chair of the TIGR board.

“Bananas have unique characteristics that will provide researchers with a powerful model, capable of investigating fundamental questions with potentially widespread applications to agriculture,” Frison said. He notes several areas of scientific interest

Scientists will ultimately be able to compare the genome of wild bananas that reproduce sexually with those of asexual crop bananas. This should provide important insights in to how, and how quickly, plant genomes evolve.

Bananas originated in Asia, but several thousand years ago, humans introduced them to Africa. The wild bananas that remained in Asia continued to co-evolve with their pests, while the African arrivals left most of their pests behind. Comparing the genomes of wild Asian varieties with those of African cultivars will provide an uncommon look at the effects of disease agents on genome evolution.

The majority of cells of most organisms have two sets of chromosomes (one inherited from the female, the other from the male). In the laboratory, bananas can be grown with anywhere from one to six sets of chromosomes. Once the banana genome is known, scientists will be able to probe the effects of multiple chromosome sets on basic plant functions, such as how plants use and store carbon.

Bananas are the only known plant in which a virus (the Banana Streak Virus) imbeds pieces of itself into the banana's own DNA, only to pop out during times of stress, reassemble itself, and cause disease. The banana genome sequence should reveal just how this virus is able to strike when the plant is most vulnerable. It may provide a powerful new tool for targeted genetic transformation.

North-South collaboration key to success

During this week's meeting at the National Science Foundation, Consortium members from developed as well as developing countries are discussing how to divide up the banana genome, which sequencing methods to use, and how to fund the project.

"We have ensured North-South collaboration right from the beginning," Frison said. "And we have the opportunity to benefit from all the players agreeing on how to proceed from the outset."

Prior to the founding of the Consortium, individual members, including the governments of Australia, Belgium, France, India, Mexico, and the United States were already investing an estimated US\$1.2 million in banana genomics research. Following the July meeting, this financing will increase to at least US\$2 million, with an additional 10 full-time researchers. However, in order to achieve their goals, the Consortium will require US\$4 to \$5 million annually. The Consortium expects some of this additional support to come from the U.S. National Science Foundation and the European Union.

"By building a strong public and non-profit Consortium that unites experts in genomics and plant biology, we hope to be able to approach the private sector later from a position of strength," Frison said. "We want to guarantee that any results, such as disease-resistant banana strains, will be made available to small-holder farmers on a royalty-free basis, even if they have commercial applications."

The International Network for the Improvement of Banana and Plantain (INIBAP) (www.inibap.org) is a programme of the Rome-based International Plant Genetic Resources Institute (IPGRI) (www.cgiar.org/ipgri), a Future Harvest Center. INIBAP was established in 1985 as an international organization with a mission to sustainably increase the productivity of banana and plantain grown on smallholdings for domestic consumption and for local and export markets. Using networking as its *modus operandi*, INIBAP has a small headquarters staff in Montpellier, France and regional offices in the four major banana-growing areas of the world.

Future Harvest (www.futureharvest.org) is a nonprofit organization that builds awareness and support for food and environmental research for a world with less poverty, a healthier human family, well-nourished children, and a better environment. Future Harvest supports research, promotes partnerships, and sponsors projects that bring the results of research to rural communities, farmers, and families in Africa, Latin America, and Asia. Future Harvest is an initiative of 16 food and environmental research centers that receive funding from the Consultative Group on International Agricultural Research (www.cgiar.org).

Annex 5. Names and addresses of the participants

Mr Nicolas ROUX
FAO/IAEA Laboratories
A-2444 Seibersdorf
AUSTRIA
Tel.: +43 1 2600 28275
Fax: +43 1 2600 28222
Email: n.roux@aiea.org

Dr Laszlo SAGI
Laboratory of Tropical Crop
Improvement
Catholic University of Leuven
Kasteelpark Arenberg 13
B-3001 Leuven
BELGIUM
Tel.: +32 16 321421
Fax: +32 16 321993
Email: laszlo.sagi@agr.kuleuven.ac.be
Web site: www.agr.kuleuven.ac.be/
dtp/tro/home.htm

Prof. Guido VOLCKAERT
Katholieke Universiteit Leuven
Laboratory of Gene Technology (LoGT)
Kasteelpark Arenberg 21
B-3001 Leuven
BELGIUM
Tel.: +32 16 329667
Fax: +32 16 321965
Email: gvolckaert@agr.kuleuven.ac.be
Website: www.agr.kuleuven.ac.be/
dp/logt/logt.htm

Dr Rosana GARCIA COLLEVATTI
Universidade Catolica de Brasilia
Pro-Reitorato de Pos-Graduacao e
Pesquisa
SGAN 916, Mod. B, Asa Norte
70790-160 Brasilia ,DF
BRAZIL
Tel.: +55 61 340 5550
Fax : +55 61 356 1800
Email: rosanegc@pos.ucb.br
Web site: www.ucb.br

Dr Robert N.G. MILLER
Universidade Catolica de Brasilia
Pro-Reitorato de Pos-Graduacao e
Pesquisa
SGAN 916, Mod. B, Asa Norte
70790-160 Brasilia ,DF
BRAZIL
Tel.: +55 61 340 5550
Fax : +55 61 356 1800
E.mail: robert@unb.br or
bruna@tba.com.br
Web site: www.ucb.br

Dr Manoel Teixeira SOUZA
CENARGEN/EMBRAPA
Parque Estação Biológica
Final Avenida W5 Norte
Caixa Postal 02372
70.770-900 Brasilia-DF
BRAZIL
Tel.: +55 61 448 4700
Fax: +55 61 340 3624
Email: msouza@cenargen.embrapa.br

Dr Jaroslav DOLEZEL
Laboratory of Molecular Cytogenetics
and Cytometry
Institute of Experimental Botany
Sokolovska 6
CZ-77200 Olomouc
CZECH REPUBLIC
Tel.: +42 68 5228521
Fax: +42 68 5228523
Email: dolezel@ueb.cas.cz
Web site: www.ueb.cas.cz/olomouc1

Dr Francois COTE
CIRAD-FHLOR
Avenue Agropolis
34398 Montpellier Cedex 5
FRANCE
Tel.: +33 (0) 467 61 59 27
Fax: +33 (0) 467 61 56 05
Email: angelique.d'hont@cirad.fr

Dr Jean-Vincent ESCALANT
INIBAP
Parc Scientifique Agropolis II
F-34397 Montpellier Cedex 5
FRANCE
Tel.: +33 (0) 467 61 13 02
Fax: +33 (0) 467 61 03 34
Email: j.escalant@cgiar.org
Web site: www.inibap.org

Dr Emile FRISON
INIBAP
Parc Scientifique Agropolis II
F-34397 Montpellier Cedex 5
FRANCE
Tel.: +33 (0) 467 61 13 02
Fax: +33 (0) 467 61 03 34
Email: e.frison@cgiar.org
Web site: www.inibap.org

Dr Pierre LAGODA
CIRAD-BIOTROP
Avenue Agropolis
TA 40/03 (Bat. 3, Bur. 001)
34398 Montpellier Cedex 5
FRANCE
Tel.: +33 (0) 467 61 56 31
Fax: +33 (0) 467 61 56 05
Email: pierre.lagoda@cirad.fr

Dr Heiko SCHOOF
Inst. f. Bioinformatik Forschungs-zen-
trum f. Umwelt u. Gesundheit
85764 Neuherberg
GERMANY
Tel.: +49 89 3187 35 80
Fax: +49 89 3187 3585
Email: w.mewes@gsf.de
Web site: www.mips.gsf.de

Dr Lalitha ANAND
Division of Biotechnology
Indian Institute of Horticultural
Research
Hessaraghatta
Bangalore - 560089
INDIA
Tel.: +91 80 846 6420
Fax: +91 80 846 6291
Email: lalithaanand@yahoo.com

Dr Diego GONZALEZ de LEON
CINVESTAV
Km 9.6 Libramiento Norte
Carretera Irapuato-León
CP 36500
Irapuato, GTO
MÉXICO
Tel.: +52 462 396 00
Fax: +52 462 396 50
Email: mgomez@ira.cinvestav.mx

Dr Dieter KAEMMER
Centro de Investigación Científica de
Yucatán, A.C.
Unidad de Biotecnología
Calle 43, No. 130
Col. Chuburná de Hidalgo
97200 Mérida
Yucatán
MEXICO
Tel.: +52 998 139 23 Ext. 178
Fax: +52 998 139 00/813941
Email: DIETER@CICY.MX

Dr Laura SILVA
CINVESTAV
Department of Genetic Engineering
Laboratory of Reproductive
Development and Apomixis
Km 9.6 Libramiento Norte
Carretera Irapuato-León
CP 36500
Irapuato, GTO
MÉXICO
Tel.: +52 462 396 34
Fax: +52 462 458 49
Email: vielle@ira.cinvestav.mx

Dr Michael PILLAY
IITA Eastern and Southern Africa
Regional
Center (ESARC) - Namulonge
PO Box 7878
Kampala
UGANDA
Tel.: +256 41 22 34 60
Fax: +256 41 22 34 59
Email: m.pillay@cgiar.org

Prof. J.S. (Pat) HESLOP-HARRISON
Department of Biology
University of Leicester
LE1 7RH Leicester
UNITED KINGDOM
Tel.: +44 (0) 116 252 5079
(direct) or 3381 (office)
Fax: +44 (0) 116 252 2791
Email: PHH4@le.ac.uk
Web site: www.molcyt.com

Dr Christopher A. CULLIS
Francis Hobart Herrick Professor
of Biology
Department of Biology
Case Western Reserve University
Cleveland OH 44106-7080
USA
Tel.: +1 216 368 5362
Fax: +1 216 368 4672
Email: cac5@po.cwru.edu /
ccullis@nsf.gov

Dr Claire M. FRASER
The Institute for Genomic Research
(TIGR)
9712 Medical Center Drive
Rockville, MD 20850
USA
Tel.: +1 301 838 5812
Fax: +1 301 838 0208
Email: cmfraser@tigr.org
Web site: www.tigr.org

Prof. Andrew PATERSON
Plant Genome Mapping Laboratory
University of Georgia
Rm 162, Riverbend Research Center
110 Riverbend Road
Athens, GA 30602
USA
Tel.: +1 706 583 0162
Fax: +1 706 583 0160
Email: paterson@uga.edu

Dr Chris TOWN
The Institute for Genomic Research
(TIGR)
9712 Medical Center Drive
Rockville, MD 20850
USA
Tel.: +1 301 838 5812
Fax: +1 301 838 0208
Web site: www.tigr.org

Acronyms and abbreviations

AFLPs	amplified fragment length polymorphism
BAC	bacterial artificial chromosome
BLSD	black leaf streak disease
BSV	banana streak virus
CARBAP	Centre africain de recherches sur bananiers et plantains, Cameroon
cDNA	complementary deoxyribonucleic acid
CENARGEM	Centro Nacional de Pesquisa de Recursos Geneticos y Biotecnologia, Brazil
CINVESTAV	Centro de Investigación y de Estudios avanzados del IPN, Mexico
CIRAD	Centre de coopération internationale en recherche agronomique pour le développement, France
CORBANA	Corporación Bananera Nacional, Costa Rica
DNAP	DNA Plant Technology Corporation, USA
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuaria
EMS	ethylmethane sulphonate
EST	expressed sequence tag
EU	European Union
FAO	Food and Agriculture Organization of the United Nations, Italy
FISH	fluorescence <i>in situ</i> hybridization
GSF	National Research Center on Environment and Health, Germany
IAEA	International Atomic Energy Agency, Austria
IEB	Institute of Experimental Botany, Czech Republic
IIHR	Indian Institute for Horticultural Research
IITA	International Institute of Tropical Agriculture, Nigeria
ILTAB	International Laboratory for Tropical Agricultural Biotechnology, USA

INIBAP	International Network for the Improvement of Banana and Plantain, France
IPR	intellectual property rights
KULeuven	Katholieke Universiteit Leuven, Belgium
MAS	marker assisted selection
MGR	mutant germplasm repository
MIPS	Munich Information Center for Protein Sequences, Germany
MTA	Material Transfer Agreement
NRCB	National Research Center on Banana, India
NSF	National Science Foundation, USA
PRINS	primed <i>in situ</i>
PROMUSA	A Global Programme for <i>Musa</i> Improvement
QDPI	Queensland Department of Primary Industries, Australia
QTL	quantitative trait loci
RFLP	restriction fragment length polymorphism
SSR	simple sequence repeat
STMS	sequence-tagged microsatellite
TIGR	The Institute for Genomic Research, USA
TAMU	Texas & AM University, USA
UCB	Universidade Catolica de Brasilia
UNESCO	United Nations Educational, Scientific and Cultural Organization, France

Layout and cover design : Roberto Hamm - Crayon & cie
Printed in France by Arceaux 49, Montpellier

ISBN 2-910810-48-08

FUTURESM
HARVEST

