**Palacký University Olomouc**

**Faculty of Science**

**Department of Botany**

**and**

**Centre of Plant Structural and Functional Genomics**

**Institute of Experimental Botany AS CR**

**Centre of the Region Haná for Biotechnological and Agricultural Research**

**Olomouc**

# Karyotype evolution in bananas (*Musa* spp.)

**Ph.D. Thesis**

**Denisa Šimoníková**

Olomouc 2021

Supervisor: Mgr. Eva Hřibová, Ph.D.

**Acknowledgements**

I would like to express my gratitude to my supervisor Mgr. Eva Hřibová, Ph.D., for her professional guidance, advice and support. My thanks also belong to the head of the laboratory, prof. Ing. Jaroslav Doležel, Dr.Sc., for the opportunity to be a member of the Centre of Plant Structural and Functional Genomics. I would like to thank the rest of my colleagues for their help and a friendly atmosphere during my Ph.D. studies.

**Declaration**

I hereby declare that I have written the Ph.D. thesis independently under the supervision of Mgr. Eva Hřibová, Ph.D., using the sources listed in references with no conflict of interest.

.....

This work was supported by the Czech Science Foundation (grant award 19-20303S).

# Bibliographical identification

**Author's name:** Mgr. Denisa Šimoníková

**Title:** Karyotype evolution in bananas (*Musa* spp.)

**Type of Thesis:** Ph.D. thesis

**Department:** Department of Botany

**Supervisor:** Mgr. Eva Hřibová, Ph.D.

**The year of presentation:** 2021

**Abstract:**

Until recently, cytogenetic studies in many plant species were hindered by limited availability of robust DNA probes. As the only possibility for the analysis of whole genomes or individual chromosomes, chromosome painting method based on labeling of pools of chromosome-specific BAC (bacterial artificial chromosome) clones was used. Unfortunately, this technique was applicable only in plants with small genomes and low number of repetitive sequences, and whose whole genome sequence was obtained by BAC by BAC sequencing strategy. However, the progress in next generation sequencing technologies enabled the development of oligo painting technique, which is based on the identification of short (45-nt long), chromosome-specific oligos from a reference genome sequence of selected species. Fluorescence *in situ* hybridization (FISH) with these fluorescently labeled oligos allowed the identification of individual chromosomes and large chromosomal rearrangements - translocations, comparative karyotype analysis and evolutionary studies in various plant species.

The Ph.D. thesis aims to provide comparative analysis of chromosome structure in selected banana species and edible banana cultivars from genus *Musa*, and their closely related genera *Musella* and *Ensete* using recently developed oligo painting FISH technique. Chromosome/chromosome-arm specific oligo painting probes developed using the reference genome sequence of *M. acuminata* ssp.

*malaccensis* ‘DH Pahang’ (A genome) were used for unambiguous identification of individual chromosomes and anchoring of the pseudomolecules to chromosomes *in situ*. Further, the cross-hybridization of the oligo painting probes specific to *M. acuminata* ssp. *malaccensis* to closely related genomes of *M. balbisiana* (B genome) and *M. schizocarpa* (S genome) proved the use of the probes for comparative karyotyping in banana (*Musa* spp.).

In the second part of the work, available chromosome/chromosome-arm specific oligo painting probes were used for the identification of large structural chromosome changes (translocations) in a set of twenty economically important diploid and triploid edible banana cultivars from the genus *Musa* and in species and subspecies, which were probably involved in their origin. Large differences in chromosome structure between individual species and subspecies of *Musa* were detected. Specific translocations, which occurred during the evolution of *Musa*, enabled the identification of putative progenitors of edible banana clones. Observed structural chromosome heterozygosity supported the hybrid origin of selected accessions and pointed to possible reason for their low fertility. These findings will be a valuable asset for banana breeders in selecting parents for further crosses.

FISH technique with previously developed chromosome-arm specific oligo painting probes was also applied to study chromosome structure of phylogenetically distinct species to *M. acuminata* within whole Musaceae family, differing in genome size and basic chromosome number. This part of the work aims to elucidate the genome organization at chromosomal level during the evolution and origin of species of this family. Finally, oligo painting probes were designed for the analysis of chromosome structures in other plant species, e. g. in fonio millet (*Digitaria exilis*), and genera *Lupinus* spp. and *Silene* spp.

**Keywords:** *Musa*, oligo painting FISH, chromosomal translocations, karyotype evolution

**Number of Pages/Appendices:** 110/VII

**Language:** English

# Bibliografická identifikace

**Jméno autora:** Mgr. Denisa Šimoníková

**Název práce:** Evoluce karyotypu banánovníku (*Musa* spp.)

**Typ práce:** Disertační práce

**Katedra:** Katedra botaniky

**Školitel:** Mgr. Eva Hřibová, Ph.D.

**Rok obhajoby:** 2021

**Abstrakt:**

Ještě nedávno byly cytogenetické studie u mnoha rostlinných druhů omezeny dostupností vhodných DNA sond. Pro analýzu struktury genomů/vybraných chromozomů se využívala metoda zvaná malování chromozomů (chromosome painting), založená na značení BAC klonů specifických pro jednotlivé chromozomy. Tuto metodu malování chromozomů bylo ale možné použít pouze u rostlin s malými genomy a nízkým obsahem repetitivních sekvencí, jejichž celogenomová sekvence byla získána sekvenováním jednotlivých BAC klonů. Nicméně pokrok v sekvenování nové generace umožnil u rostlin rozvoj tzv. "oligo painting" metody využívající celogenomové sekvence pro identifikaci krátkých (45 bází), chromozomově specifických oligomerů. Fluorescenční *in situ* hybridizace (FISH) s těmito fluorescenčně značenými oligomery unikátními pro jednotlivé chromozomy/oblasti chromozomů, tak umožnila mimo jiné identifikaci jednotlivých chromozomů a velkých strukturních změn - translokací, srovnávací analýzu a studium evoluce karyotypů u mnoha dalších rostlinných druhů.

Disertační práce si klade za cíl přispět k objasnění struktury chromozomů u vybraných druhů a jedlých kultivarů banánovníku z rodu *Musa,* a také jejich blízce příbuzných rodů *Musella* a *Ensete* s využitím nedávno vyvinuté metodiky pro malování chromozomů ("oligo painting FISH"). Poprvé tak bylo pomocí sond specifických pro jednotlivé chromozomy nebo jejich ramena identifikováno všech 11

chromozomů, ke kterým byly ukotveny DNA pseudomolekuly referenční genomové sekvence druhu *M. acuminata* ssp. *malaccensis* 'DH Pahang' (A genom). Oligo malovací sondy navržené pro *M. acuminata* ssp. *malaccensis* byly úspěšně hybridizovány i na blízce příbuzné druhy *M. balbisiana* (B genom) a *M. schizocarpa* (S genom), což poukázalo na jejich možné využití pro analýzu struktury chromozomů i u dalších zástupců banánovníku (*Musa* spp.).

Druhá část práce byla zaměřena na využití těchto oligo malovacích sond pro identifikaci velkých chromozomálních přestaveb (translokací) u dvaceti zástupců ekonomicky významných diploidních a triploidních jedlých typů banánovníku z rodu *Musa* a druhů, resp. poddruhů, které se pravděpodobně podílely na jejich vzniku. Mezi jednotlivými druhy, resp. poddruhy i odrůdami byly zjištěny velké rozdíly v genomové struktuře na úrovni chromozomů. Specifické chromozomové translokace, ke kterým došlo v průběhu evoluce, umožnily identifikaci pravděpodobných předků jedlých typů banánovníku. U vybraných testovaných položek byl potvrzen jejich hybridní charakter, což poukazuje na možnou příčinu jejich snížené fertility. Tyto poznatky o detailní struktuře jednotlivých chromozomů jsou cennou informací pro šlechtitele při výběru klonů vhodných pro další křížení.

Sondy specifické pro jednotlivé chromozomy nebo chromozomální ramena byly následně využity pro studium struktury chromozomů i u vybraných zástupců reprezentujících čeleď banánovníkovité (Musaceae), lišících se kromě velikosti genomu i základním chromozomovým číslem. Cílem bylo objasnění organizace genomu na úrovni chromozomů v průběhu evoluce a vzniku druhů v čeledi Musaceae. V neposlední řadě byly oligo malovací sondy navrženy pro studium struktury chromozomů i u dalších rostlinných druhů, např. rosičky útlé (*Digitaria exilis*), a rodů lupina (*Lupinus* spp.) a silenka (*Silene* spp.).

**Klíčová slova:** *Musa*, oligo painting FISH, chromozomové translokace, evoluce karyotypu

**Počet stran/příloh:** 110/VII

**Jazyk:** Anglický

# CONTENT

# 1 INTRODUCTION

Bananas and plantains are one of the major staple food crops for millions of people living in tropical and subtropical developing countries and represent important world trade commodity. In 2019, about 159 million tons of bananas and plantains were produced worldwide (FAOSTAT, 2019). However, majority of the bananas is consumed in a place of their origin, only about 10-20 % is exported. In 2019, around 26 million tons of bananas were distributed around the world (International trade statistics, 2019). Two types of bananas are available at the market, sweet and cooking bananas. While sweet bananas serve as a food supplement, cooking bananas are starchier and have to be processed before consuming.

Unfortunately, bananas are susceptible to many diseases and pests, mainly in plantations consisting of asexually propagated banana clones, which have a low genetic diversity. In the last century, Panama disease (Fusarium wilt of banana), caused by fungal plant pathogen *Fusarium oxysporum* f.sp. *cubense*, destroyed entire plantations of the most important export cultivar of sweet banana, Gros Michel. More resistant Cavendish cultivar was then selected to replace Gros Michel, becoming the world dominant export banana type. However, new highly virulent race of Panama disease (Tropical Race 4, TR4) has emerged in 1990s in Southeast Asia and spread though the world, namely to Africa, Western Asia, Australia and Latin America, where it infects also Cavendish cultivars (Ploetz, 2015). Moreover, abiotic stress associated with climate change has a negative effect on banana production (Van Asten *et al*., 2011). The use of pesticides is not a permanent solution, because this type of chemical control has negative economic and environmental impact. For many years, banana breeders make huge efforts to breed new resistant cultivars with high yield and nutritional value at the same time. Nevertheless, breeding is hampered by low fertility or even sterility associated with the polyploid nature of cultivated bananas and limited knowledge of genome structure of wild banana genotypes and improved diploids that are used in breeding programs (Ortiz and Swennen, 2014; Brown *et al*., 2017).

Despite huge socio-economic significance of bananas, a little is known about their genome structure and organization at chromosomal level, which is important for further genetic improvement of edible clones in breeding programs. Moreover, the evolution of whole Musaceae family is still unclear. However, the availability of banana reference genome sequence enabled the implementation of relatively new cytogenetic method called oligo painting FISH in banana studies and the identification of all chromosomes within a karyotype. The aim of this work is to shed light on genome organization and karyotype evolution in Musaceae family, especially in edible banana clones, using fluorescence *in situ* hybridization with whole chromosome oligo painting probes.

# 2 LITERATURE OVERVIEW

## 2.1 Characteristics of banana (*Musa* spp.)

### 2.1.1 Plant morphology

Bananas are monocotyledonous, tree-like, evergreen perennial herbs (Robinson and Saúco, 2010). Cultivated banana plants are about 2-9 meters high, some wild species can reach even 10-15 meters (Karamura *et al*., 2011).

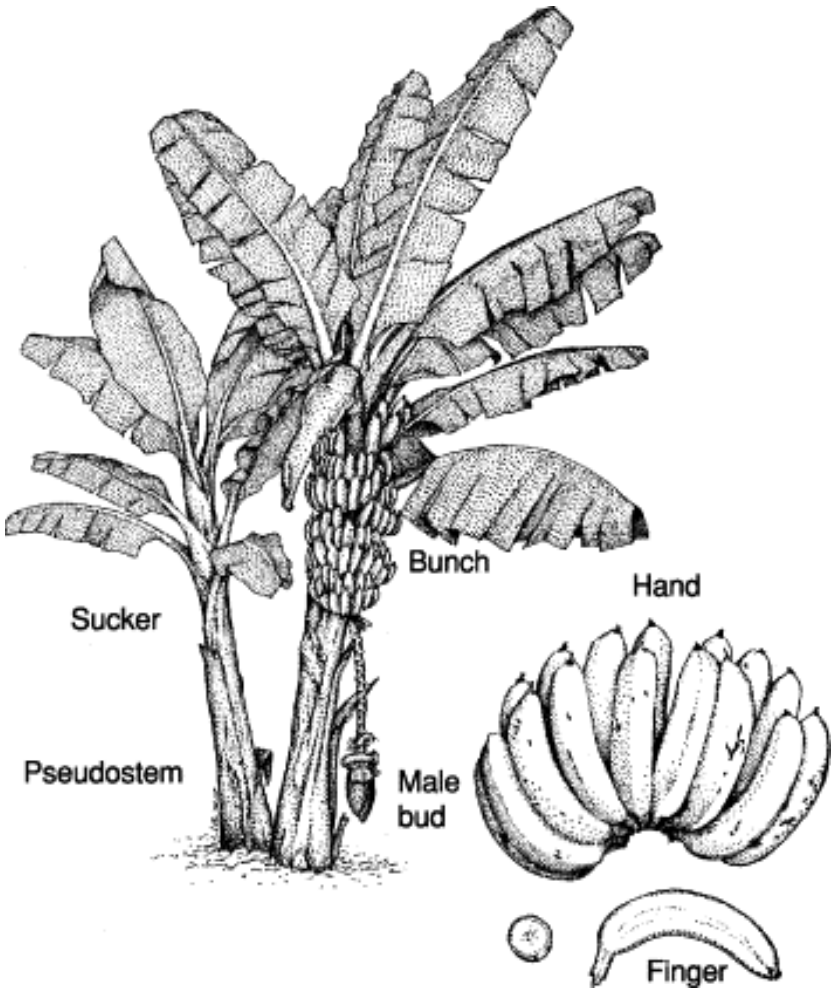



**Figure 1.** Structure of the banana plant (Macrae *et al.*, 1993).

The banana plant is composed of the basal rhizome, pseudostem, terminal crown of large leaves and inflorescence (Figure 1). Leaves are developed from apical

meristem, located on the top of the rhizome. New leaves, called cigar leaves, are enrolled and always grow in the middle of the pseudostem, which is composed of tightly packed leaf sheaths (Stover and Simmonds, 1987; Daniells, 2003). Mechanism of flower induction is widely discussed (e. g. Fahn *et al.*, 1963; Stover and Simmonds, 1987). The time of flowering and fruit maturity depends on the cultivar and environmental conditions. In general, floral initiation begins with suppression of leaf development, when about 30-40 leaves are already developed. Apical meristem starts to produce flower bracts instead of leaves, followed by the emergence of the inflorescence bud (Purseglove, 1972; Karamura *et al*., 2011).

The inflorescence is complex and comprises the peduncle with basal female flower clusters, which develop into the fruit bunch, followed by hermaphrodite flowers, and distal male (staminate) flower clusters with no fruit production. Pollen, produced by male flowers, is reduced or completely absent in cultivated bananas. The fruit (berry) is developed without the stimulus of pollination by vegetative parthenocarpy. Ovules shrivel inside the fruit, resulting in sterility or low fertility of cultivated seedless bananas. The fruit bunch consists of fruit clusters, called hands, while the individual fruit is called a finger (Figure 1). Despite the fact, that the pseudostem is very fleshy, full of water and with no woody parts, it can hold the bunch of fruits even heavier than 50 kilograms. After the growing season, fruit-bearing pseudostem dies and new plants are regenerated from suckers. Suckers emerge from vegetative buds, which are located opposite to the base of leaves. However, wild species of bananas, which have fruit full of dark seeds, reproduce also sexually (Robinson, 1996; Daniells, 2003; Karamura *et al*., 2011).

### 2.1.2 Taxonomy and geographical distribution

Banana family (Musaceae) originated and diversified during the early Eocene in tropical Southeast Asia, especially in Northern Indo-Burma (Southwestern China) or Malesian Region (Malaysia and Indonesia), the region which preserves most of the banana biodiversity. From the centre of their origin, bananas expanded also to Africa and South America (Simmonds, 1962; D'Hont *et al*., 2012; Häkkinen, 2013; Janssens *et al*., 2016).

Musaceae is a small, monophyletic group, which comprises one of the most important tropical fruit crops (Figure 2; Li *et al*., 2010; Liu *et al*., 2010). As a member of Zingiberales order, containing eight families, the family Musaceae is closely related to Heliconiaceae, Lowiaceae and Strelitziaceae (Kress *et al*., 2001; Kress and Specht, 2005). Musaceae comprises three genera, *Musa* L. (Linnaeus, 1753), *Ensete* Bruce ex Horan. (Horaninow, 1862) and *Musella* (Fr.) C.Y. Wu (Li, 1978). The largest of them, genus *Musa*, with about 75 species with a great economic importance mainly in Asia, Africa, Latin-America and Pacific islands, covers majority of Musaceae species (Häkkinen, 2013; Christelová *et al*., 2017). On the contrary, two closely related genera *Ensete* and *Musella* include just a few species (Li *et al*., 2010; Liu *et al*., 2010).

The large size of banana plants, long lasting life cycle and the lack of resolution of morpho-taxonomic markers made the taxonomic classification in Musaceae more difficult. However, according to basic chromosome number (x) and morphological diversity (IPGRI-INIBAP/CIRAD, 1996), Cheesman (1947) distinguished genus *Ensete* from *Musa* and established four sections within *Musa*, namely Eumusa (2n=2x=22), Rhodochlamys (2n=2x=22), Callimusa (2n=2x=20) and Australimusa (2n=2x=20). Additionally, a few species with ambiguous origin were observed, *Musa ingens* (2n=2x=14), *Musa beccarii* (2n=2x=18) and *Musa lasiocarpa* (2n=2x=18) (Simmonds, 1960). Lately, Simmonds (1962) classified *M. beccarii* into Callimusa section (2n=2x=20 or 2n=2x=18). More recently, Argent (1976) created another independent section Ingentimusa containing a single species *M. ingens* (2n=2x=14), which occurs in Papua New Guinea and is considered the largest herbaceous plant in the world. Based on plant morphology, taxonomic classification of *Musa lasiocarpa* has been studied over many years. This species was originally placed under *Musa* and *Ensete* (Franchet 1889; Baker 1893; Cheesman, 1947; Simmonds, 1960). Finally in 1978, Li classified this species as a representative of a new genus *Musella*. However, phylogenetic position of *Musella lasiocarpa* (previously named *Musa lasiocarpa*) still needs more investigation.

Previous studies using genotyping with several types of DNA markers, such as microsatellite-based (SSR) genotyping, analysis of nuclear ribosomal ITS

sequences and chloroplast DNA or cytogenetic studies, questioned the classification system in *Musa* proposed by Cheesman (1947). Phylogenetic relationships provided by molecular markers showed close evolutionary relationships between Eumusa and Rhodochlamys species (both x=11), and between Ingentimusa, Australimusa and Callimusa representatives, which vary in basic chromosome number (x =7, 9, 10) (e. g. Wong *et al*., 2002; Bartoš *et al*., 2005; Risterucci *et al*., 2009; Li *et al*., 2010; Liu *et al*., 2010; Christelová *et al*., 2011; Hřibová *et al*, 2011; Janssens *et al*., 2016). Based on these findings, merging of section Eumusa and Rhodochlamys into one section called Musa, and sections Callimusa, Ingentimusa and Australimusa into another section called Callimusa have been suggested by Häkkinen (2013). However, this idea was not accepted within banana research community and traditional taxonomy based on four sections proposed by Cheesman (1947) is still widely used.

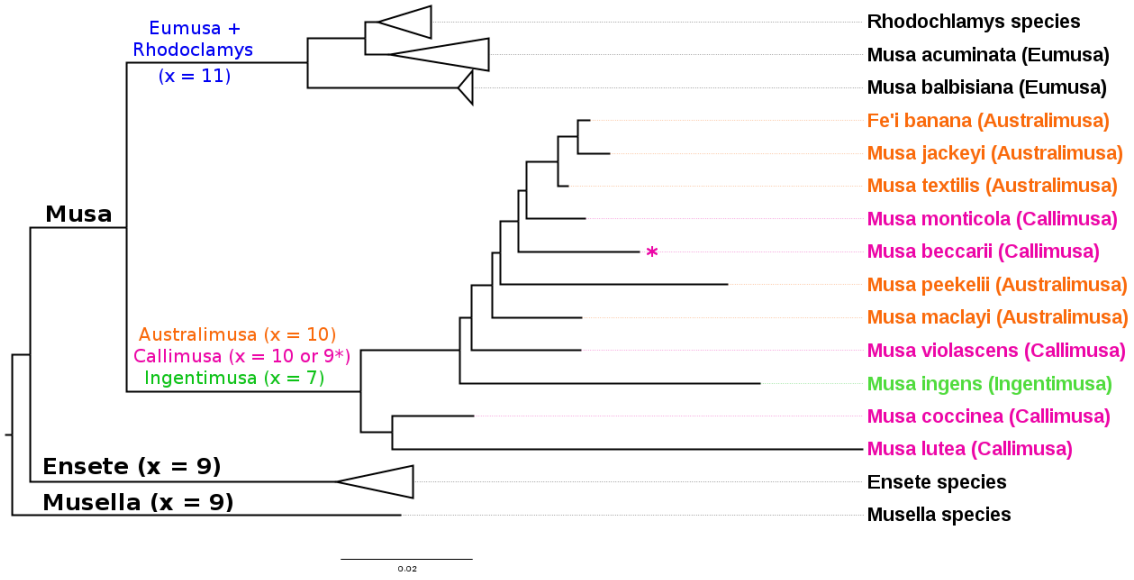



**Figure 2.** Phylogenetic tree of Musaceae family constructed using Neighbor-joining method of ITS1-ITS2 sequence regions of ribosomal DNA (retrieved from GenBank), rooted on genus *Musella* (Šimoníková *et al*., unpublished).

#### 2.1.2.1 Genus *Musa*

Genus *Musa* is distributed throughout Southeast Asia (from India to Papua New Guinea), Africa and South America (Häkkinen, 2013; Janssens *et al*., 2016).

There are two important species among the representatives of Eumusa section, *Musa acuminata* and *Musa balbisiana*, which contributed to the origin of majority of edible banana clones. *M. acuminata* grows throughout the whole Southeast Asia, whereas *M. balbisiana* is limited to the region from East India to South China (Figure 3; Simmonds and Shepherd, 1955; De Langhe *et al*., 2009). Despite genetic variability observed in *M. balbisiana* (e. g. Sotto and Rabara, 2000; Ude *et al*., 2002), no subspecies have been distinguished to date. *M. balbisiana* has valuable agronomic traits, such as an ability to grow well in drier contidions and resistance to diseases and pests as compared to *M. acuminata* (Sotto and Rabara, 2000). Moreover, large and waterproof leaves of *M. balbisiana* are used in wide variety of applications, mainly as traditional food wrappers or disposable plates for serving meals especially in South India and Southeast Asia (Jacob, 1952; Subbaraya, 2006; Maidin and Latiff, 2015).

On the contrary, nine subspecies (namely *banksii, burmannica, burmannicoides, errans, malaccensis, microcarpa, siamea, truncata, and zebrina*) and three varieties (namely chinensis, sumatrana, and tomentosa) have been recognized within *M. acuminata* based on geographical distribution (Simmonds, 1962; Perrier *et al*., 2009; 2011; Martin *et al*., 2017; WCSP, 2018). Regions, where individual subspecies of *M. acuminata* occur are almost non-overlapping (Figure 3; Perrier *et al*., 2011). Individual subspecies of *M. acuminata* cannot be clearly distinguished based on plant morphology, only a few morphological differences have been detected. However, large phenotypic variation has been observed between species *M. acuminata* and *M. balbisiana* (Simmonds, 1956).

While species from Eumusa section are characterized by large (3 meters or more) plants with pendent inflorescences with many flowers, plants from section Rhodochlamys, which is phylogenetically closely related to section Eumusa, are smaller (less than 3 meters), bearing erect inflorescences with few flowers (Wong *et al*., 2002). Rhodochlamys species are widely used as ornamental plants because of their typically brightly coloured bracts. Species from this section are resistant to seasonal droughts, which are frequent in the monsoonal areas mainly in Northeast India, Bangladesh, Myanmar and Thailand (Häkkinen and Sharrock, 2002).

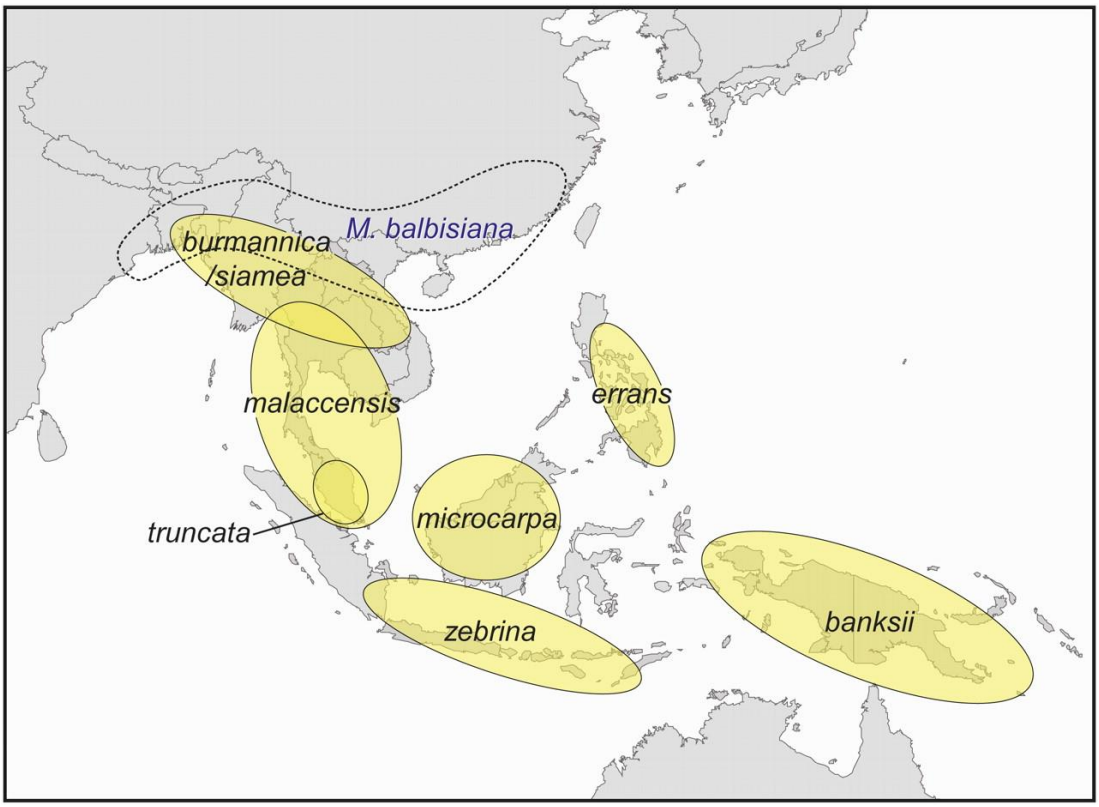



**Figure 3.** Geographical distribution of *M. balbisiana* and subspecies of *M. acuminata* (Perrier *et al*., 2011).

Section Australimusa comprises important banana species, which are grown for fibers, e. g. *Musa textilis*, and also edible cultivars, known as polynesian Fe'i bananas (*Musa* × *troglodytarum* L.; 2n=2x=20). Fe'i are parthenocarpic and vegetatively propagated clones (MacDaniels, 1947; Bartoš *et al*., 2005; Häkkinen, 2013). Compare to edible cultivars derived from *M. acuminata* and *M. balbisiana*, Fe'i bananas are grown only on Pacific islands (particularly French Polynesia), with the center of origin in Papua New Guinea. Fe'i cultivars were domesticated independently from edible cultivars of Eumusa section, and their ancestors are not known. Fe'i bananas are believed to contain T genome from *Musa textilis*, which belongs to the same section (Daniells *et al*., 2001). Compared to edible clones from Eumusa section, Fe'i bananas are characterized by distinctive erect bunch consisting of fruit clusters with typical orange pulp and high carotenoid content (Englberger *et al*., 2003; 2006).

The most variable section Callimusa contains species differing in basic chromosome number (2n=2x=18 and 20; Cheesman, 1947; Simmonds, 1962; Čížková *et al*., 2015). Species from this section are characterized by unique seed morphology. Seeds are cylindrical, barrel-shaped and possessing a large apical chamber. On the contrary, species from Eumusa, Rhodochlamys and Australimusa sections have dorsiventrally compressed seeds, possessing a small apical chamber (Wong *et al*., 2002). Recent molecular studies showed, that this section comprises also *M. ingens* (2n=2x=14; Burgos-Hernández *et al*., 2017; Šimoníková *et al*., unpublished).

#### 2.1.2.2 Genus *Ensete*

Genus *Ensete* (2n=2x=18), which comprises seven species (namely *E. ventricosum, E. livingstonianum, E. glaucum, E. homblei, E. perrieri, E. superbum, E. lecongkietii*), colonized sub-Saharan Africa, Madagaskar and Asia (Cheesman, 1947; Simmonds, 1962; Matheka *et al*., 2019). *Ensete* is well-adapted to cooler and drier environments than most *Musa* species (Cheesman, 1947; Baker and Simmonds, 1953). *Ensete ventricosum*, also called Ethiopian banana, represents an important food security crop, whose cultivation dominates in highlands of Ethiopia (Borrell *et al*., 2019). The fruit is seldom edible (enset is known as a false banana), but the pseudostem and underground corm are carbohydrate-rich (Brandt *et al*., 1997).

High levels of genetic diversity and overlapping geographical distribution of wild and domesticated enset have been observed in Ethiopia, indicating multiple domestications events and introgressions from wild populations (Tobiaw and Bekele, 2011; Olango *et al*., 2015). However, all domesticated enset landraces arose probably from a single species, *E. ventricosum*, in contrast to cultivars from genus *Musa* (Borrell *et al*., 2019).

#### 2.1.2.3 Genus *Musella*

*Musella* (2n=2x=18) is a monospecific genus native to Southern Sichuan and Northern Yunnan, growing only in Southwestern China (Li, 1978; Wu and Kress,

2001; Liu *et al*., 2002; Ma *et al*., 2011). Because of a narrow distribution, *Musella lasiocarpa* populations have limited genetic variability (Pan *et al*., 2007). Interestingly, *Musella* tolerates much drier and colder environment compare to genera *Musa* and *Ensete* (Liu *et al*., 2002; Janssens *et al*., 2016). The current analyses confirmed that *Musella* is phylogenetically closely affiliated with *Ensete* (e. g. Li *et al*., 2010; Liu *et al*., 2010).

## 2.2 Origin of edible banana clones

Edible banana cultivars, which belong to Eumusa section of genus *Musa*, originated after natural inter- or intra-specific cross hybridization between two wild species *Musa acuminata* Colla (2n=2x=22, AA genome) and *Musa balbisiana* Colla (2n=2x=22, BB genome). Cultivated bananas differ in ploidy level and genomic constitution. Crosses resulted mostly in seed sterile diploid (AA, AB), triploid (AAA, AAB, ABB) or even tetraploid hybrids (AAAB, AABB), obtained in the breeding programs (Simmonds and Shepherd, 1955; Simmonds, 1956). Moreover, molecular studies revealed the contribution of *Musa schizocarpa* (Eumusa section, 2n=2x=22, SS genome), which is very close to *M. acuminata*, and diploid *Musa textilis* (Australimusa section, 2n=2x=20, TT genome) to the origin of edible banana clones after crosses with *M. acuminata* (Carrell *et al*., 1994; Čížková *et al*., 2013; Němečková *et al.*, 2018).

The most common edible cultivars are triploid bananas, that emerged after the fusion of unreduced gamete from edible diploid, almost sterile cultivar with haploid gamete from fertile diploid (Simmonds, 1962; Carreel *et al.*, 1994; Raboin *et al*., 2005). Two types of triploid bananas are known, autotriploids derived from *M. acuminata* (AAA group) and allotriploids (with AAB and ABB constitution), in which *M. balbisiana* was involved in the crosses (Price, 1955).

Very strong drop in sea levels during glaciation periods facilitated human migration from Southeast Asia to Bismark Archipelago, New Guinea and close islands, which were already occupied around 30 000 years ago (Sand, 1989; Kirch, 1997; Kagy *et al*., 2016). Horticulture has been developed around 7000 years ago. At the time, tuber plants and bananas were grown in New Guinea Highlands (Denham *et*

*al*., 2004). From New Guinea, which is considered the most active centre of diversity for *Musa* (Perrier *et al.*, 2009), people migrated eastwards within centuries, reaching Polynesia, and thereby introducing traditional banana cultivars and other domestic plants (Horrocks *et al.*, 2009; Horrocks and Rechtman, 2009).

*M. balbisiana* was transferred southward, enabling hybridization with *M. acuminata* (Perrier *et al*., 2011). *M. balbisiana* provides valuable traits to the resulting inter-specific hybrids, because of its resistance to biotic and abiotic stresses and a strong root system (Bakry *et al*., 2009). On the other hand, several studies indicated that the genome of *M. balbisiana* 'Pisang Klutuk Wulung', which is often used in breeding programs, contains infectious segments of endogenous Banana streak virus (eBSV) sequences. Upon activation by biotic and abiotic stresses, integrated eBSVs lead to infections by several species of BSV harbouring both *M. acuminata* and *M. balbisiana* genomes in inter-specific hybrids (Gayral *et al*., 2008; Chabannes *et al*., 2013; Umber *et al*., 2016).

### 2.2.1 Contribution of wild diploid *M. balbisiana* and subspecies of *M. acuminata* to evolution of edible banana clones

As mentioned above, due to climatic changes during glaciation periods, different *M. acuminata* subspecies got in close proximity, resulting in cross hybridization and the origin of intra-specific hybrids, which were further selected and propagated by humans (Perrier *et al*., 2011; Martin *et al*., 2017).

Three particular subspecies of *M. acuminata* from different geographic regions have been proposed as the main contributors to the A genome of edible banana cultivars. *Musa* cultivars have diverse and more complex mosaic genome structure than it was previously assumed. Recent studies support multiple hybridization events and contribution of more ancestral genotypes in their origin (Rouard *et al*., 2018; Baurens *et al*., 2019; Martin *et al*., 2020a). *M. acuminata* ssp. *banksii*, which is endemic in New Guinea, played a key role in the origin of cultivated clones (Sharrock, 1990; Perrier *et al*., 2009; Němečková *et al*., 2018), followed by *M. acuminata* ssp. *zebrina*, which originated in Indonesia (Java island) (Rouard *et al*., 2018), and *M. acuminata* ssp. *malaccensis* with the center of diversity

in Malay peninsula (De Langhe *et al*., 2009; Perrier *et al*., 2011). Several studies indicate a minor contribution of *M. acuminata* ssp. *burmannica*, which originated in Burma (nowadays called Myanmar) (Cheesman, 1948) and *M. acuminata* ssp. *burmannicoides* and *M. acuminata* ssp. *siamea* from the same *burmannica* genetic group (Carreel *et al*., 1994; 2002; Perrier *et al*., 2009; 2011; Christelová *et al*., 2017; Martin *et al*., 2020b).

Clarifying the evolution of individual subspecies of *M. acuminata*, which played important role as genetic contributors to modern cultivars, is crucial for understanding the domestication and diversification of banana. Until recently, several studies have attempted to resolve relationships among subspecies of *M. acuminata*, however only few subspecies have been analysed and limited number of markers was used (Jarret *et al.*, 1992; Christelová *et al*., 2011; Janssens *et al*., 2016; Sardos *et al*., 2016; Christelová *et al.*, 2017). Lately, analysis of *de novo* draft genome assemblies of three *M. acuminata* subspecies (namely *banksii*, *zebrina* and *burmannica*) (Rouard *et al*., 2018), combined with reference genome sequences of *M. acuminata* ssp. *malaccensis* (D'Hont *et al*., 2012; Martin *et al*., 2016) and *M. balbisiana* (Davey *et al*., 2013), revealed a close relationship between subspecies. Individual subspecies are relatively homogenous, suggesting the rapid radiation within subspecies of *M. acuminata* after the divergence from *M. balbisiana*. Moreover, an introgression between *M. acuminata* ssp. *burmannica* and *M. acuminata* ssp. *malaccensis* was confirmed and pointed out to the geographical overlap of these two subspecies (Perrier *et al*., 2011; Rouard *et al*., 2018).

Martin *et al*. (2020a) used transcriptomic data for the identification of specific single-nucleotide polymorphisms (SNPs), and confirmed that *M. balbisiana*, *M. acuminata* ssp. *zebrina*, and *M. acuminata* ssp. *malaccensis* are represented by distinct ancestries. Moreover, a close relationship between *M. acuminata* ssp. *banksii* and *M. acuminata* ssp. *microcarpa* 'Borneo', which was previously suggested by Carreel *et al.* (1994), was approved by having one common ancestor. In addition, transcriptomic data indicated that *M. acuminata* ssp. *burmannica*, *M. acuminata* ssp. *burmannicoides* and *M. acuminata* ssp. *siamea* shared common ancestry (Martin *et al*., 2020a), supporting the idea of a single genetic group comprising these closely

related subspecies (Perrier *et al*., 2011; Dupouy *et al*., 2019). In contrast to previous studies (Carreel *et al*., 1994; Perrier *et al*., 2009; Christelová *et al*., 2017), which indicated a close relationship between *M. acuminata* ssp. *errans* and *M. acuminata* ssp. *banksii*, Martin *et al*. (2020a) surprisingly showed that *M. acuminata* ssp. *errans* is closely associated with *M. acuminata* ssp. *malaccensis*, *M. acuminata* ssp. *zebrina* and *M. acuminata* ssp. *burmannica/siamea*. However, *M. acuminata* ssp. *errans*, together with *M. acuminata* ssp. *truncata* and *M. acuminata* ssp./var. sumatrana have not been predicted to contribute to the origin of edible cultivars (Martin *et al*., 2020a). More recently, Martin *et al*. (2020b) precisely characterized six large reciprocal translocations, which originated in different subspecies of *M. acuminata* and helped to better understand the subspecies evolution in bananas.

However, there are some limitations in the study of Martin *et al*. (2020a), in which only transcriptomic data specific to *Musa* hybrids have been analysed. In this case, one parental genome could be possibly silenced or completely substituted by other parental genome, which was already observed in other plant species, e. g. in *Tragoporon mirus* (Buggs *et al*., 2010) or cotton (Yoo *et al*., 2013). Following this, contribution of various *M. accuminata* subspecies in the origin of *Musa* hybrids detected in the study of Martin *et al*. (2020a) might have been inaccurate or incomplete.

### 2.2.2 Diploid edible banana cultivars

Several groups of diploid edible banana cultivars are known. Ney Poovan group of bananas (also known as Elakki Bale) is represented by diploid inter-specific cultivars with AB genomic constitution. These edible diploids are one of the most significant varieties grown in India. The fruit is highly fragrant and have unique taste. The Elakki cultivars are drought-tolerant, have a resistance to leaf spot, but they are susceptible to fusarium wilt and banana bract mosaic virus (Bohra *et al*., 2014; Selvakumar and Parasurama, 2020).

The most important diploid, phenotypically diversified edible AA cultivars, known as Mchare, are grown in East Africa, especially in highlands of Kenya, Tanzania and Malawi. These varieties subsequently spread also onto Indian Ocean

islands and Madagaskar. Except of the name Mchare (widely used in Tanzania), AA cultivars are also called Mlali in the Comoro islands and Muraru in Kenya. Studies indicate that these cultivars share some alleles with *M. acuminata* ssp. *zebrina* and *M. acuminata* ssp. *banksii*, with possible origin in Borneo, Java or Sumatra (Perrier *et al*., 2009; Hippolyte *et al*., 2012). However, no similar forms are grown nowadays in that region and hypothesis about a single ancestor is the most probable one. AA edible diploids are genetically homogeneous, but well adapted to highly diverse ecological zones. Most importantly, these Mchare cultivars have a huge potential for breeding and genetic improvement of edible triploid Gros Michel/Cavendish AAA subgroups of bananas (Perrier *et al.*, 2019). Recently, Martin *et al*. (2020a) proposed the complex hybridization scheme of Mchare bananas. Transcriptomic data indicated that *M. acuminata* ssp. *banksii/microcarpa* and *M. acuminata* ssp. *zebrina* could be involved in their origin, which was previously proposed by Perrier *et al*. (2009). Moreover, Martin *et al*. (2020a) also suggested participation of *M. acuminata* ssp. *malaccensis* in the emergence of Mchare cultivar 'Chicame'.

### 2.2.3 Intra-specific triploid edible banana cultivars

Several groups of triploid intra-specific edible cultivars are well-described, including Ibota, Rio, Red and the most important groups of dessert banana clones - Cavendish and Gros Michel, as well as a group of cooking bananas known as Mutika-Lujugira.

Dessert bananas represent an important source of income for people from tropical developing countries. The most popular dessert type of banana is represented by one particular genotype - Cavendish, which dominates at the international market. In 2018, Cavendish cultivars accounted for 57 % of the global banana production. However, most of that production is consumed at the place of origin, only about 16 % of Cavendish was exported in 2018 (Lescot, 2020). However, despite a huge economic importance of Cavendish cultivars, as well as Gros Michel, export industry relies on monoculture of these genetically closely related intra-specific sterile triploid clones with monospecific *M. acuminata* origin (AAA) (Simmonds, 1962; Carreel, 1994; Jeridi *et al*., 2011). In the last century, Gros Michel represented the main cultivar of sweet banana used for the export. Unfortunately, fungal pathogen

*Fusarium oxysporum* f.sp. *cubense*, causing Panama disease, destroyed entire plantations of these clonally propagated bananas. Thus, Gros Michel was replaced by Cavendish cultivar, which is more resistant (Stover, 1962; Ploetz, 2015).

Only a small genetic difference between Cavendish and Gros Michel cultivars has been observed (Raboin *et al*., 2005; Christelová *et al*., 2017). These cultivars most probably originated after the hybridization of a partially sterile diploid Mchare banana (*zebrina/microcarpa* and *banksii* ascestry) with fertile Khai subgroup (*malaccensis* ascestry), which comprises the 'Pisang Madu' and 'Pisang Pipit' accessions. Mchare bananas (formerly Mlali subgroup) provided unreduced diploid (2n) restitution gamete, whereas Khai served as the donor of normal haploid gamete (Simmonds, 1966; Raboin *et al*., 2005; Perrier *et al*., 2009; Hippolyte *et al*., 2012; Martin *et al*., 2017; Perrier *et al*., 2019; Martin *et al*., 2020a). Transcriptomic data indicated that 'Pisang Madu' genotypes contributed to the origin of Cavendish cultivar 'Grande Naine' (Martin *et al*., 2020a). 'Pisang Madu' genotypes originated on Borneo and are represented by highly heterozygous genotypes with complex and not completely known origin (Rosales *et al*., 1999; Perrier *et al*., 2009; Martin *et al*., 2020a)

Triploid East African Highland bananas, also known as matooke, are representatives of Mutika/Lujugira group of bananas (Shepherd, 1957). The starchy Matooke clones represent a nutritionally very important banana group for over 80 millions of people from Great Lakes region of East Africa, namely Burundi, Democratic Republic of Congo, Kenya, Rwanda, Tanzania and Uganda. Matooke cultivars are sterile, vegetatively propagated bananas with AAA genomic constitution. Nowadays, about 120 cultivars, which were selected by farmers, are grown (Němečková *et al*., 2018). These bananas are usually cooked before consuming or fermented in the form of banana beer. In Uganda, the term 'matooke' is equal to food and these edible clones serve as a staple food and are traditionally prepared as steamed bananas. Around 250 kg of fruit per inhabitant per year is consumed in Uganda, Burundi and Rwanda (Perrier *et al*., 2019).

Minimal genetic variation has been observed in Mutika/Lujugira group of bananas (Kitavi *et al*., 2016; Christelová *et al*., 2017). However, due to accumulation

of somatic mutations in the absence of sexual reproduction and/or epigenetic changes, different phenotypes exist (e. g. Shepherd, 1957; De Langhe, 1961; Kitavi *et al.*, 2016; Němečková *et al*., 2018). Diploid *M. acuminata* ssp. *banksii* together with *M. acuminata* spp. *zebrina* are considered to be the putative parents of these triploid clones (Carreel *et al*., 2002; Perrier *et al.*, 2009; Hippolyte *et al*., 2012; Li *et al*., 2013; Kitavi *et al.*, 2016; Christelová *et al*., 2017; Němečková *et al*., 2018; Martin *et al*., 2020b). Surprisingly, participation of *M. schizocarpa* (genome S) in their origin was proposed by several authors (Carreel *et al.*, 2002; Heslop-Harrison and Schwarzacher, 2007; Němečková *et al*., 2018). Moreover, previous studies suggested, that matooke did not evolve from East African diploids, known as Mchare cultivars (Hippolyte *et al*., 2012; Sardos *et al*., 2016; Němečková *et al*., 2018).

### 2.2.4 Inter-specific triploid edible banana cultivars

One of the most important group of edible inter-specific hybrids with AAB genomic constitution – the Plantains, cover about 18 % of total banana production and are mainly grown in West Africa and Central/South America. Together with cultivars characterized by ABB genomic composition, these triploid edible starchy bananas make up almost 40 % of global production (Baurens *et al*., 2019).

Allotriploids with ABB genome composition represent starchy bananas commonly used for cooking, dessert and beer production (Karamura *et al*., 1998). Moreover, resistance to weevils, nematodes and black leaf streak has been observed (Karamura *et al*., 1998). These ABB triploids are also drought-tolerant (van Wesemael *et al*., 2019). ABB banana cultivars are morphologically diverse, nine subgroups have been recognized, namely Bluggoe, Monthan, Ney Mannan, Klue Teparod, Kalapua, Peyan, Pisang Awak, Pelipita and Saba (Daniells *et al*., 2001). ABB bananas arose in two geographical regions, Southeast Asia and India (De Langhe *et al*., 2009; Perrier *et al*., 2011).

Next to the previously mentioned Plantains, another allotriploid subgroups with AAB genome composition have been described, e. g. Mysore, Pome or Silk. Plantains, which are the most important AAB triploids, have been divided into two subgroups according to their secondary centers of diversification (De Langhe *et al*.,

2009). African plantains, restricted to the African continent (growing mainly in Central and West Africa), represent the most morphologically diverse subgroup of triploid bananas (Swennen, 1990), which could be explained by somaclonal variation (Vuylsteke *et al.* 1988; 1991). However, these bananas exhibit a low genetic diversity. Around 100 kg of fruit per inhabitant per year is consumed in Africa (Perrier *et al*., 2019). Second group is known as the Pacific plantains (also called Maoli/Popoulu plantains), which have been observed in South India (Simmonds, 1966), Northern Philippines (De Langhe and Valmayor, 1979) and spread to all Pacific Islands from Southeast Asia due to human migration (e. g. Sand, 1989; Kirch, 1997; Donohue and Denham, 2009). Low or almost no genetic diversity has been noticed in this subgroup of plantains (De Langhe *et al*., 2015).

Based on variation of the inflorescence, fruit morphology and size, three morphotypes have been recognized in plantains (Figure 4; De Langhe *et al*., 2005). The 'Horn' types are characterized by the absence of male flowers, high number of neutral flowers and the presence of horn-like fruit. The 'French' (common) plantains have male and female flowers and smaller fruit (Kervégant, 1935). The intermediate type known as the 'False Horn' plantains, have no male bud and low number of neutral flowers (De Langhe, 1961). No molecular markers characteristic for individual morphological subgroups have been identified so far (De Langhe *et al*., 2005). Epigenetic factors, such as DNA methylation, could be involved in the phenotypic diversification, which was previously studied in mangroves or oil palm (Lira-Medeiros *et al*., 2010; Ong-Abdullah *et al*., 2015).

Inter-specific triploid cultivars with AAB or ABB genome composition are not simple allopolyploids, which would arise from a few crosses. Their evolution most probably has passed through an inter-specific AB hybrid (with unreduced gamete), which hybridized with diploid *M. acuminata* ssp. *banksii* or *M. balbisiana*, donors of a haploid gamete, followed by several backcrosses to one of the parents (De Langhe *et al*., 2010; Perrier *et al.,* 2011; Hippolyte *et al.,* 2012; Baurens *et al.,* 2019; Cenci *et al*., 2020). AB hybrids and ABB allotriploids originated in India and surrounding regions and in Southeast Asia, where *M. balbisiana* is abundant (De Langhe *et al*., 2009; Perrier *et al*., 2011).

Recent studies showed that genomes of inter-specific hybrids contain different proportion of A and B chromosomes and/or recombinant chromosomes. Unbalanced proportion of A and B genome alleles complicates banana breeding. Moreover, morphology of cultivated inter-specific bananas does not correspond to the simple genome formulas suggested by Simmonds and Shepherd (1955), but a bias towards the A or B phenotype was detected (De Langhe *et al*., 2010).

Subsequently, Baurens *et al*. (2019) sequenced selected A/B inter-specific cultivars through whole genome sequencing (WGS) or RadSeq to characterize the A and B genome composition in their genomes. It has been confirmed that inter-specific recombination occurs very often between A and B chromosomes of these cultivars. This phenomenon could facilitate recombination of traits between two species during the breeding of improved cultivars. For example, unfavourable endogenous virures, which are integrated in B genome, could be eliminated (Chabannes *et al*., 2013; Noumbissié *et al*., 2016). However, a reciprocal translocations and inversions in the genomes of inter-specific banana clones have been detected. These chromosomal rearrangements affect recombination, followed by segregation distortion and aneuploidy in triploids, which hampers the breeding efforts. Nevertheless, the presence of large structural variations and inter-specific recombination elucidated the cause of mosaic genome structure of inter-specific edible cultivars (Baurens *et al*., 2019). Recently, Cenci *et al*. (2020) used SNP markers called from RadSeq data and confirmed the frequent occurrence of homologous chromatin exchanges between A and B subgenomes in ABB cultivars. Thus, *M. balbisiana* could serve as a source of useful variability in breeding of new improved cultivars. It was found out that ABB cultivars are more drought tolerant (van Wesemael *et al*., 2019; Cenci *et al*., 2020).

In plantains (AAB), two A genomes are highly similar, indicating that *M. acuminata* ssp. *banksii* was most probably involved in the origin of AB inter-specific hybrid. Folowing this, resulted AB hybrid hybridized with haploid gamete also from *M. acuminata* ssp. *banksii* (Horry, 1989; Lebot *et al*., 1993; Carreel *et al*., 2002; Perrier *et al*., 2011; Hippolyte *et al*., 2012; Kagy *et al*., 2016; Baurens *et al*., 2019, Martin *et al*., 2020b). Perrier *et al*. (2009) pointed to *M. acuminata* ssp.

*malaccensis* as the most probable progenitor in inter-specific cultivars with ABB genome composition. However, ABB clones from Saba and Bluggoe-Monthan subgroups most likely contain A genome from *M. acuminata* ssp. *banksii* similarly to plantains (AAB), which was later proposed by several authors (e. g. Christelová *et al*. 2017; Martin *et al*., 2020b).

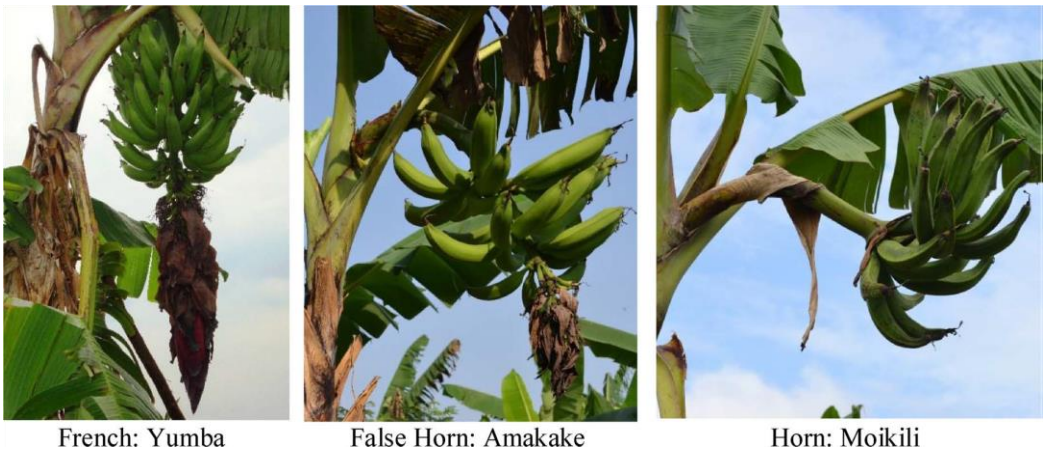



**Figure 4.** Three morphotypes of plantains (Adheka *et al*., 2018).

## 2.3 Chromosomal rearrangements have accompanied the evolution of banana genome

Several cytogenetic studies on chromosome pairing during meiosis in *Musa* indicated that evolution of edible cultivated banana clones and some wild species was accompanied by chromosomal rearrangements, particularly translocations, which lead to differences in chromosome structure (Dodds, 1943; Wilson, 1945; Simmonds, 1962; Dessauw, 1987; Fauré *et al*., 1993; Shepherd, 1999; Therdsak *et al*., 2010). Structural chromosome heterozygosity causes irregularities in meiosis (such as irregular chromosomal pairing) and subsequent development of low fertility or complete sterility in hybrids and aberrant chromosome numbers in the progeny (Jáuregui *et al*., 2001; Rieseberg, 2001; Ostberg *et al*., 2013).

Based on chromosome pairing during meiosis in inter-subspecific hybrids of *M. acuminata*, seven (to eight) translocation groups of structurally homogenous accessions have been distinguished among *M. acuminata* subspecies by Shepherd

(1999). One Standard group (ST) and six groups differing from Standard group by 1 to 4 translocations only partially corresponded to the classification of individual subspecies. These findings were supported by several authors, who observed segregation distortion during genetic mapping in inter-subspecific hybrids (Fauré *et al*., 1993; Hippolyte *et al*., 2010; Mbanjo *et al*., 2012; Noumbissié *et al*., 2016).

The Standard group (ST) is the largest one, comprising *M. acuminata* ssp. *banksii*, *microcarpa* and *malaccensis*, whereas other six groups are based on their geographic origin. The Northern Malayan group (NM) consists of some *M. acuminata* ssp. *malaccensis* accessions, the Northern 1 group comprises *burmannicoide*s and some *siamea* accessions, the Northern 2 group includes *burmannica* and some *siamea* accessions, the Malayan Highland group consists of one *truncata* accession, the Javanese group includes two *zebrina* accessions and finally the East African group comprises one unclassified accession (Shepherd, 1999; Martin *et al*., 2017).

Recently, the production of whole genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang' (D'Hont *et al*., 2012; Martin *et al*., 2016) allowed the implementation of next generation sequencing (NGS) technologies and comparative genomics to reveal chromosomal rearrangements in *Musa*. Until now, the presence of several translocations, inversions and duplications in *Musa* genome was confirmed. Characterization of large structural variations is necessary for the easier utilization of genetic resources in *Musa* breeding strategies, especially for improvement of disease resistance (Dupouy *et al*., 2019).

Martin *et al.* (2017) revealed a heterozygous reciprocal translocation using mate-pair sequencing, BAC-FISH, targeted PCR and DArTseq. The translocation involved 3 Mb segment of chromosome 1 and 10 Mb segment of chromosome 4 in wild diploid *M. acuminata* ssp. *malaccensis*, compare to a reference genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang', which is derived from Standard translocation group (ST; Shepherd, 1999; D'Hont *et al*., 2012; Martin *et al*., 2016; 2017). The 1/4 translocation, corresponding to Northern Malayan translocation group (NM), proposed previously by Shepherd (1999), was detected only in some of the *M. acuminata* ssp. *malaccensis* accessions, thus probably originated within this

subspecies (Martin *et al*., 2017). This heterozygous reciprocal translocation leaded to a high segregation distortion and lower recombination rate. Interestingly, it was preferentially transmitted to the progeny. The 1/4 translocation was confirmed in one copy of chromosomes e g. in clone 'Pisang Lilin', Mchare banana 'Akondro Mainty' (ST x NM hybrid) and in triploid cultivars Cavendish and Gros Michel (Martin *et al*., 2017). The study indicated the close genetic relationship between Mchare and Cavendish subgroup of dessert bananas, which was already proposed by several authors (Raboin *et al*., 2005; Perrier *et al*., 2009). Lately, Martin *et al*. (2020b) verified the presence of 1/4 translocation in heterozygous state in 'Pisang Lilin'.

In *Musa schizocarpa*, which is considered to be the contributor to the origin of some edible banana clones (Carrell *et al*., 1994; Čížková *et al*., 2013; Němečková *et al*., 2018), no large structural variation was detected as compared to the reference genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang' (Belser *et al*., 2018).

Baurens *et al*. (2019) detected a paracentric inversion on chromosome 5 and a large reciprocal translocation involving chromosomes 1 and 3 in *M. balbisiana* 'Pisang Klutuk Wulung' after anchoring a dense genetic map to the reference genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang' (ST group). The inversion comprised a large 9 Mb segment, and a reciprocal translocation involved a large 8,5 Mb segment of chromosome 3, which was translocated to chromosome 1 and a small 0,65 Mb segment of chromosome 1, which was translocated to chromosome 3. In *M. balbisiana*, which is less genetically variable than *M. acuminata*, no subspecies have been identified (Swangpol *et al*., 2007; Gayral *et al*., 2010), thus detected translocations could be present in other *M. balbisiana* acccessions involved in the origin of inter-specific cultivars (Baurens *et al*., 2019).

Moreover, the aneuploidy involving chromosomes 1, 3 and 5 was detected in the progeny of tetraploid possessing one copy of *M. balbisiana* genome. Additionally, high segregation distortion and lower recombination rate for these chromosomes were confirmed (Baurens *et al*., 2019). In inter-specific cultivars with AAB and ABB genomic constitution, no translocated chromosomes from *M. acuminata* have been detected (Martin *et al*., 2020b). These findings are in line

with the study from Hippolyte *et al*. (2012), in which *M. acuminata* ssp. *banksii*, belonging to Standard group (ST), is considered as a donor of A genome to AAB and ABB cultivars.

Dupouy *et al*. (2019) detected two large homozygous reciprocal translocations in seedy *M. acuminata* ssp. *burmannicoides* 'Calcutta 4', compared with the *M. acuminata* ssp. *malaccensis* reference sequence. Thus, the Northern 1 group was suggested for 'Calcutta 4', as it differs from 'DH Pahang' (ST group) by two translocations (Shepherd, 1999). The first one involved an exchange of a 240 kb distal region of chromosome 2 with a 7,2 Mb distal region of chromosome 8, the second one consisted of an exchange of a 20,8 Mb distal region of chromosome 1 with a 11,6 Mb distal region of chromosome 9. Both translocations were identified in selected wild accessions of *M. acuminata* ssp. *burmannicoides, M. acuminata* ssp. *burmannica* and *M. acuminata* ssp. *siamea* after sequencing mate-pair libraries using Illumina technology. However, only 2 of 87 analysed cultivars were heterozygous for 2/8 translocation, whereas 1/9 translocation was not detected. These findings indicate that these subspecies were not important contributors to the origin of present cultivars (Carreel *et al*., 1994).

Additionally, 1/9 and 2/8 translocations (Dupouy *et al*., 2019) were confirmed in homozygous state in *M. acuminata* ssp. *burmannicoides* 'Calcutta 4' and *M. acuminata* ssp. *siamea* 'Pa Rayong' (Martin *et al*., 2020b). Using BAC-FISH method, the presence of 1/9 translocation (Dupouy *et al.*, 2019) at homozygous state was confirmed in *M. acuminata* ssp. *siamea* 'Khae Phrae' and ssp. *burmannica* 'Long Tavoy' (Martin *et al*., 2020b).

*M. acuminata* ssp. *burmannicoides, M. acuminata* ssp. *burmannica* and *M. acuminata* ssp. *siamea* are phylogenetically closely related, which was already proposed by several authors (Shepherd, 1999; Carreel *et al*., 2002; Perrier *et al*., 2009; Hřibová *et al*., 2011; Christelová *et al*., 2017; Němečková *et al*., 2018; Martin *et al*., 2020a) and have similar geographical distribution and morphological characteristics (Simmonds, 1962; Perrier *et al*., 2009). However, recent studies did not support separation into three distinct subspecies (Sardos *et al*., 2016; Christelová *et al*., 2017). Based on analysis of chromosome pairing during meiosis in inter-

subspecific hybrids of *M. acuminata*, Shepherd (1999) placed *burmannica* and some *siamea* accessions into Northern 2 group, which vary by one additional translocation from Northern 1 group containing *burmannicoides* and some *siamea* accessions. Thus, the idea of unified *burmannica* genetic group needs more investigation and the hypothesis about the presence of a third translocation proposed by Shepherd needs to be confirmed. Nevertheless, Dupouy *et al*. (2019) suggested, that detected translocations emerged in the *burmannica* genetic group. Moreover, several accessions from this *burmannica* genetic group are disease resistant-rich, which is an important attribute in breeding of new cultivars. For example, clone 'Calcutta 4' (*M. acuminata* ssp. *burmannicoides*) is resistant to nematodes or black leaf streak disease, clones 'Long Tavoy' (*M. acuminata* ssp. *burmannica*) and 'Khae Phrae' (*M. acuminata* ssp. *siamea*) showed resistance to nematodes and to F*usarium oxysporum* f.sp. *cubense* (Vuylsteke *et al*., 1993; Jones, 2000; Quénéhervé *et al*., 2009).

Another three large reciprocal translocations were revealed after sequencing of 155 *Musa* accessions (containing banana cultivars and representatives of *Musa* diversity) using Illumina Hiseq 4000 platform and genotyping by sequencing of 1059 individuals from 11 progenies (Martin *et al*., 2020b). These three newly observed translocations, together with three previously reported reciprocal translocations (Martin *et al*., 2017; Dupouy *et al*., 2019), which emerged in different (sub)species of *M. acuminata*, were precisely characterized (Figure 5; Martin *et al*., 2020b). The majority of diploid and triploid cultivars were structurally heterozygous for 1 to 4 translocations, indicating their hybrid origin. Detected translocations induced a reduction of recombination and translocated chromosomes were preferentially transmitted to the progeny. Moreover, out of the 1059 progeny individuals studied, 96 aneuploids have been observed (Martin *et al*., 2020b). The geographical distribution of individual translocations corresponded to the habitat of individual subspecies (Figure 3), which was previously suggested by Perrier *et al*. (2009).

The presence of a homozygous reciprocal translocation involving one arm of chromosome 3 and one arm of chromosome 8 was observed in *M. acuminata* ssp. *zebrina* (Martin *et al*., 2020b). This translocation corresponds to Javanese group

proposed by Shepherd (1999). The presence of 3/8 translocation was also detected in two copies in matooke types of East African Highland bananas (EAHBs) (Martin *et al*., 2020b), which is in line with previously suggested *M. acuminata* ssp. *zebrina* contribution to the origin of this group of edible bananas (Hippolyte *et al*., 2012; Němečková *et al*., 2018). Martin *et al*. (2020b) proposed that 3/8 translocation emerged in *M. acuminata* ssp. *zebrina*.

In cultivar 'Pisang Madu', which is considered to be the n gamete donor in the origin of Cavendish and Gros Michel dessert bananas, the presence of a heterozygous reciprocal translocation involving chromosomes 1 and 7 was proposed (Martin *et al*., 2020b). The origin of 1/7 translocation could not be classified, because it was detected only in heterozygous state in selected hybrids and cultivars. Martin *et al*. (2020b) hypothesize that the 1/7 translocation corresponds to the cryptic wild gene pool, which needs more investigations. The 1/7 translocation could be characteristic for the eighth translocation group, proposed by Shepherd (1999). This translocation occurred frequently in cultivated clones, e. g. in 'Pisang Madu', Cavendish and Gros Michel cultivars (Martin *et al*., 2020b).

A homozygous reciprocal translocation involving one arm of chromosome 7 and one arm of chromosome 8 was detected in 'Khae Phrae' (*M. acuminata* ssp. *siamea*) and 'Long Tavoy' (*M. acuminata* ssp. *burmannica*). In these two accessions, also a homozygous 2/8 translocation was detected previously (Dupouy *et al*., 2019). The 7/8 translocation corresponds to the Northern 2 group previously proposed by Shepherd (1999). Martin *et al*. (2020b) suggested that 7/8 translocation originated in *burmannica* group after the emergence of 1/9 and 2/8 translocations.

It was found out that *M. acuminata* ssp. *microcarpa* 'Borneo' is structurally homozygous with the reference chromosome structure of *M. acuminata* ssp. *malaccensis* 'DH Pahang'. In *M. acuminata* ssp. *banksii* and in most accessions of *M. acuminata* ssp. *malaccensis*, no large translocations have been detected (Martin *et al*., 2020b). These findings indicate the presence of ancestral chromosome structure corresponding to the Standard group (ST; Shepherd, 1999), as in *M. acuminata* ssp. *malaccensis* 'DH Pahang', which was proposed previously by Dupouy *et al*. (2019).

Martin *et al*. (2020b) suggested that the East African group proposed by Shepherd (1999) comprises a translocation involving chromosome 9 and one of chromosomes 5, 6, 10 or 11. Finally, the last Malayan Highland group could contain a reciprocal translocation involving chromosome 1 and/or chromosome 4 and another chromosome.

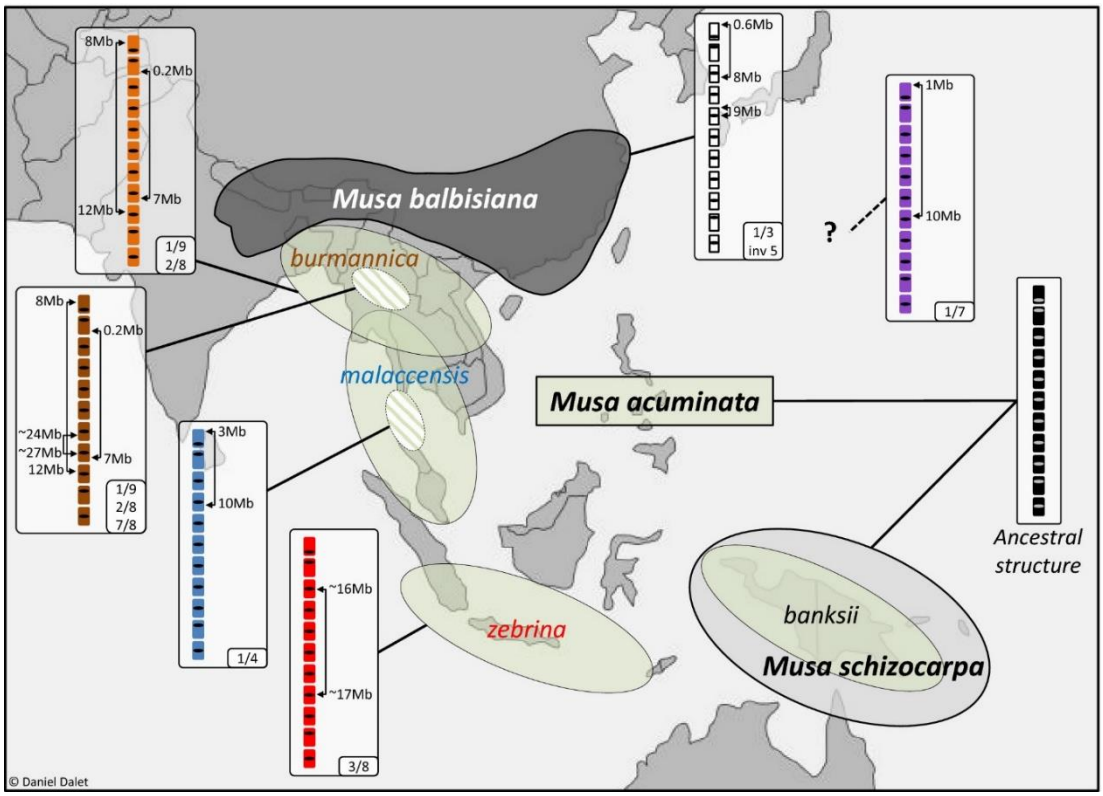



**Figure 5.** Geographical distribution and associated translocated chromosome structures of *Musa* species and subspecies, which contributed to the origin of cultivated edible clones (Perrier *et al*., 2011; Martin *et al*., 2020b).

Translocations 1/4 (from *M. acuminata* ssp. *malaccensis*) and 3/8 (from *M. acuminata* ssp. *zebrina*) were detected frequently among analysed accessions, indicating their huge contribution to the origin of cultivated bananas (Carreel *et al*., 2002; Perrier *et al*., 2009; 2011; Christelová *et al*., 2017). On the contrary, translocations 1/9, 2/8 and 7/8, which are typical for *burmannica* group, were rarely observed in edible *Musa* cultivars. These findings confirmed quite low contribution of *burmannica* group to the emergence of cultivated banana clones (Carreel *et al*., 1994; Perrier *et al*., 2011). In dessert-type banana cultivars, e. g. Cavendish and Gros

Michel, the presence of 1/4, 1/7 and 3/8 translocations was confirmed (Martin *et al*., 2020b). These data approved their complex origin and pointed to their most probable progenitors as previously suggested (Raboin *et al*., 2005; Hippolyte *et al*., 2012; Martin *et al*., 2020a).

## 2.4 Breeding of new banana cultivars

Sterility of cultivated banana clones, which is the result of inter-(sub)specifc hybridization with the contribution of unreduced gametes in the origin of triploid clones, leads to parthenocarpy. Parthenocarpy in plants causes the production of edible fruits without seeds. This characteristic was desirable for the early farmers, who selected clones producing seedless fruits during *Musa* domestication process in the Southeast Asia (D'Hont *et al*., 2000). Thereafter, cultivated bananas were maintained by vegetative propagation (Simmonds, 1962; De Langhe, *et al*., 2010).

Classical breeding of edible diploids is based on direct crosses with wild resistant genotypes or improved diploids (e. g. Brown *et al*., 2017; Batte *et al*., 2019; Brown *et al*., 2020). Conventional banana breeding of triploid banana clones comprises the development of tetraploids (4x), which are obtained by crossing of inferior parthenocarpic landrace varieties of triploids with wild seeded diploids (3x × 2x). Subsequently, these tetraploids are crossed with improved diploids (4x × 2x), resulting in secondary triploid hybrids (3x), which are evaluated as possible improved varieties. This breeding strategy, which is labour and time-consuming, lasts up to 10-15 years (Figure 6; Bakry and Horry, 1992; Tomepke *et al*., 2004; Ortiz, 2013; Nyine *et al*., 2017). After crossing of cultivated varieties with resistant wild diploids, resulting hybrids contain also unfavorable genes from the wild diploids. Nevertheless, generated tetraploids are both female and male fertile, thus further improvement is possible. On the other hand, maintaining of desirable features in existing varieties, while incorporating a resistance, is challenging for banana breeders (Nyine *et al*., 2017).

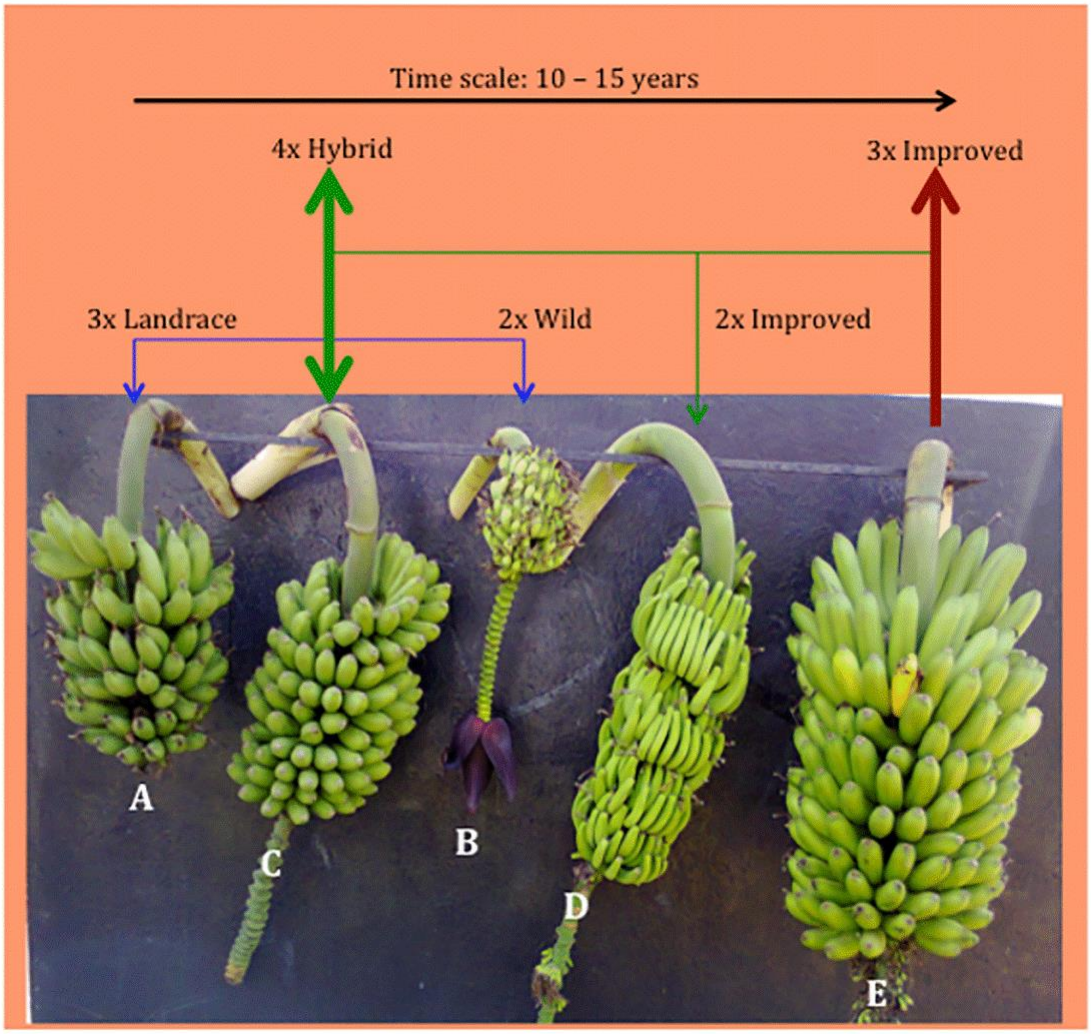



**Figure 6**. Conventional banana breeding involves crossing of triploid inferior and parthenocarpic landrace varieties (A) with wild diploid banana (B). Resulting tetraploid (C) is crossed with improved diploid hybrid banana (D). Resulting secondary triploid (E) is evaluated as a potential improved cultivar (Nyine *et al*., 2017).

Another type of breeding strategy is commonly used in CIRAD (The French Agricultural Research Centre for International Development). In this case, triploids are generated after crosses between diploids and autotetraploids, which have been obtained through chromosome doubling using colchicine treatment (Bakry and Horry, 1994; Bakry *et al*., 2001; 2009).

Triploidy is the most favourable ploidy level in banana breeding. Triploid banana plants are more vigorous, having larger fruits and higher sterility, with the

absence of seeds (Bakry *et al*., 2009). Hovewer, diploids are important in the breeding strategy for the introduction of genetic variability (Tenkouano *et al*., 2003; Amorim *et al*., 2011; 2013). Identification of fertile cultivars, which produce seeds under specific conditions, is crucial for the breeding process (Ortiz, 2013).

Nevertheless, breeding of improved banana cultivars is greatly hampered by their polyploid nature, reduced fertility, low breeding efficiency and long breeding and selection cycle (e. g. Burke and Arnold, 2001; Ortiz and Swennen, 2014; Martin *et al*., 2017; Batte *et al*., 2019; Baurens *et al*., 2019). The knowledge of large-scale structural chromosome rearrangements could reveal possible causes of their low fertility and help breeders in selecting wild seeded diploid parents with valuable traits, which could be used for breeding of the new banana cultivars.

## 2.5 Molecular cytogenetics of banana

Nowadays we have a large range of molecular tools, including next generation sequencing technologies, which have been already applied for characterization of family Musaceae. However, the application of cytological methods, such as analysis of nuclear genome size, ploidy level (or chromosome number), genomic distribution of repetitive sequences (such as rDNA), and analysis of large chromosomal changes significantly shed light on genome structure and evolution of Musaceae, which is also valuable in breeding of new *Musa* cultivars.

### 2.5.1 Genome size and ploidy level of banana

Flow cytometry is a rapid and convenient method, which enables accurate measurement of nuclear DNA content (Fox and Galbraith, 1990; Doležel, 1991). Using this technique, the genome size of 600 Mb/1C was determined in *M. acuminata* and 550 Mb/1C in M. *balbisiana* (Doležel *et al.*, 1994). These data indicated significantly lower DNA content in *M. balbisiana* compared to *M. acuminate* and facilitated the estimation of nuclear DNA content and genome composition also in triploid hybrids (Lysák *et al*., 1999; Čížková *et al*., 2013).

Bartoš *et al*. (2005) and Čížková *et al*. (2015) enlarged the knowledge about nuclear genome size by analysing wild diploid species of *Musa*, including

*M. schizocarpa* and *M. textilis*, and species from genus *Ensete*. Genome sizes of species from sections Eumusa (x=11) and Rhodochlamys (x=11) were overlapping, ranging from 578 to 704 Mb/1C and 609 to 664 Mb/1C, respectively. Representatives of the section Australimusa (x=10) had higher genome size, between 734-791 Mb/1C. However, the highest genome size (798 Mb/1C) was estimated in *M. beccarii* (x=9) from Callimusa section (Bartoš *et al*., 2005). Later on, Čížková *et al*. (2015) measured even higher nuclear DNA content in several wild banana species, with the highest in *M. borneensis* (x=9) from Callimusa section (867 Mb/1C). *E. gilleti* (x=9) from genus *Ensete* had genome size 619 Mb/1C, similar to sections Eumusa and Rhodochlamys (Bartoš *et al*., 2005). Within all analysed accessions, the genome size of *M. balbisiana* was determined as the smallest one, followed by *M. acuminata*. Moreover, variation in genome size within individual subspecies of *M. acuminata* was detected. This phenomen could be caused by diversification within *M. acuminata* since the separation from a common ancestor (Moore *et al*., 1993) and/ or by different selective pressure on *M. acuminata* accessions, which occurred in different geographical areas exhibiting various abiotic stresses (Lysák *et al*., 1999).

Except the estimation of the genome size, flow cytometry is also an effective method for rapid ploidy determination in *Musa* (Doležel *et al*., 1997; Doleželová *et al*., 2005; Christelová *et al*., 2017). In this case, the suspension of DAPI-stained nuclei isolated from fresh leaf tissue of *Musa* is mixed with the chicken red blood cell nuclei (CRBC), which are used as an internal reference standard (Galbraith *et al*., 1998). The gain of flow cytometer is adjusted so that the G1 peak of CRBC is positioned approximately at channel 100 and relative nuclear DNA content of *Musa* is estimated by comparing peak positions (corresponding to the relative fluorescence intensity) of CRBC nuclei and nuclei of a tested sample. For example, peak appearing on channel 50 corresponds to a diploid *Musa* accession. Alternatively, instead of CRBC nuclei used as an internal reference standard, *Musa* accession with known ploidy level can serve as a reference for ploidy determination of unknown *Musa* sample (Christelová *et al*., 2017).

Using flow cytometry, a ploidy level of 1150 accessions of *Musa* from ITC collection (the International *Musa* Germplasm Transit Centre, Leuven, Belgium) was determined in 2004 (Doleželová *et al*., 2005). Thus, the flow cytometry has a significant potential in an effective characterization of banana accessions, especially rapid screening of progenies obtained by breeding process.

### 2.5.2 Structure and organization of banana genome at chromosomal level

Banana genome, which is relatively small, is divided into morphologically similar, usually 1-2 μm long, highly condensed mitotic chromosomes. Average chromosome contains about 50 Mb (Doleželová *et al*., 1998; Osuji *et al*., 1998; D'Hont *et al*., 2000). Chromosomes seem to be metacentric or submetacentric. Due to a tight chromosome condensation, determination of centromere position and identification of chromosome arms is very difficult, also measuring of short and long chromosome arms is nearly impossible. However, secondary constriction (nucleolar organizing region, NOR) is usually clearly visible at the end of chromosome arms and delimits satellite, which is well separated. Thus, the microscopic identification of all *Musa* chromosomes based on morphological characteristics is not possible, only chromosomes bearing secondary constriction can be identified within karyotype.

Number of chromosomes bearing secondary constriction corresponds to ploidy level in genus *Musa*. However, higher number of chromosomes with NOR have been observed in the genera *Ensete* and *Musella*, with four and two chromosome pairs bearing a secondary contriction, respectively (Doleželová *et al*., 1998; Osuji *et al*., 1998; Bartoš *et al*., 2005; Čížková *et al*., 2013; Šimoníková *et al*., unpublished). Satellite can be completely separated from the rest of corresponding chromosome during the preparation of chromosome spreads, which can result in errors in chromosome counting. Aneuploidy, frequently detected in bananas, could be in some cases related to presence of these 'small chromosomes' (Cheesman and Larter, 1935; Sandoval *et al*., 1996; Shepherd and Da Silva, 1996). Chromosome counting is the only method in bananas for reliable detection of aneuploidy (Sandoval *et al*., 1996; Shepherd and Da Silva, 1996; Bartoš *et al*., 2005; Čížková *et*

*al*., 2013; 2015; Němečková *et al*., 2018). In this case, the usage of flow cytometry is not convenient because of differences in genome size between individual *Musa* species, subspecies or their hybrids (e. g. Čížková *et al*., 2013; Christelová *et al*., 2017; Němečková *et al*., 2018). Despite that, Roux *et al*. (2003) successfully detected aneuploidy in mutant triploid plants derived from gamma-irradiated shoot tips using DNA flow cytometry. However, this method is laborious and inconvenient for large-scale screening.

The ability to identify all chromosomes within karyotype and correlate them with chromosome-specific landmarks is crucial for characterization of banana genome at chromosomal level. However, relatively high number of *Musa* chromosomes (9-11 in haploid state), their small size and high degree of chromatin condensation make the application of classical molecular cytogenetic methods even more complicated (e. g. Valárik *et al*., 2002; Bartoš *et al*., 2005; Čížková *et al*., 2013; 2015).

Chromosome banding, introduced at the turn of 1970s (e. g. Caspersson *et al*., 1968; Pardue and Gall, 1970), was widely used for chromosome identification in plant species with large, repeat-rich genomes, comprising wheat or rye from Poaceae family (Gill and Kimber, 1977; Gill *et al*., 1991; Song *et al*., 1994) or species from genus *Lilium* (Lim *et al*., 2001). However, chromosomes of many plant species, including bananas, have uniform banding pattern, thus the application of chromosome banding is not possible (Greilhuber, 1977; Schubert *et al*., 2001).

#### 2.5.2.1 Fluorescence *in situ* hybridization (FISH)

Development of DNA *in situ* hybridization, which was based on the usage of radioactively labeled DNA sequences (probes), transferred the classical cytogenetic methods to modern molecular cytogenetics (John *et al*., 1969; Pardue and Gall, 1969). Few years later, radioactively labeled probes were substituted by the application of non-radioactive, biotin-labeled probes, based on the strong affinity of biotin to streptavidin (or avidin). Streptavidin was conjugated with the enzyme (biotinylated horseradish peroxidase), which reacted with the enzyme substrate, resulting in colored precipitate (Rayburn and Gill, 1985). The establishment of

fluorescence *in situ* hybridization (FISH) provided new possibilities for chromosome identification and analysis of chromosome structural changes by mapping of fluorescently labeled probes on mitotic and meiotic chromosomes or even in interphase nuclei (Langer-Safer *et al*., 1982).

In plants, FISH was introduced by Schwarzacher *et al*. (1989) and Leitch *et al*. (1991) and become the most important technique in plant molecular cytogenetics research, helping to generate physical cytogenetic maps of various plant species. This technique was first used for mapping of repetitive sequences, which produce unique FISH pattern on individual chromosomes (e. g. Mukai *et al*., 1993; Kato *et al*., 2004; Jiang and Gill, 2006). Except of repeats, including genes for ribosomal DNA, also BAC (bacterial artificial chromosome) clones comprising large, single-copy DNA sequences (100-200 kb) were used as FISH probes (e. g. Jiang *et al*., 1995; Lysák *et al*., 2001). Recently, even short single copy sequences (1,5-3 kb) were successfully used as probes for FISH in plants (e. g. Lamb *et al*., 2007; Danilova *et al*., 2012; Karafiátová *et al*., 2016). However, FISH-based chromosome identification was developed only in a few plant species, because non-model species have usually limited genomic resources (Braz *et al*., 2020a).

FISH is the only method in Musaceae, which allows chromosome identification, reconstruction of molecular karyotypes or comparative analysis among related species. However, only several DNA probes are available in bananas, including ribosomal DNA, dispersed and tandemly organized repeats or BAC clones (Doleželová *et al*., 1998; Osuji *et al*., 1998; Valárik *et al*., 2002; Hřibová *et al*, 2008; de Capdeville *et al*., 2009; Čížková *et al*., 2013; 2015).

#### 2.5.2.2 Ribosomal DNA sequences (rDNA)

Ribosomal 45S and 5S RNA genes (rRNA) represent the most conserved, highly frequent and tandemly organized repetitive units in all eukaryotes (Appels *et al*. 1980, Long and Dawid, 1980, Ellis *et al*. 1988). These ribosomal DNA sequences provided useful FISH markers for chromosome identification and karyotyping in various plant species, e. g. in *Aegilops* spp. (Badaeva *et al*., 1996), *Trifolium* spp. (Ansari *et al*., 1999) or *Brassica* spp. (Hasterok *et al*., 2001).

Genes coding 45S rRNA and 5S rRNA were mapped as the first probes on banana chromosomes. Doleželová *et al*. (1998) used as a probe for 45S rDNA plasmid VER 17 (Yakura and Tanifuji, 1983) containing genes from *Vicia faba*, the probe for 5S rDNA was prepared by PCR using specific primers amplifying 303 bp region in rice and some other plants. Osuji *et al*. (1998) used specific DNA sequences from *Triticum aestivum* as probes for rDNA (Gerlach and Bedbrook, 1979; Gerlach and Dyer, 1980). Later on, Valárik *et al*. (2002) isolated several DNA clones suitable as probes for rDNA by screening of partial genomic DNA libraries of *M. acuminata* and *M. balbisiana*. Radka 1 DNA clone, containing part of 26S rRNA gene, was used as a FISH probe for localization of 45S rDNA, and Radka 2 DNA clone, containing 400 bp insert of a part of the 5S rDNA, was utilized for mapping of 5S rRNA gene sites (Valárik *et al*., 2002).

*Musa* representatives of Eumusa and Australimusa sections contained 45S rDNA loci exclusively at the end of short arms of chromosomes, in NOR, thus their number corresponds to the ploidy level. However, number of 45S rDNA sites differs from two to four chromosomes in diploid species of Rhodochlamys section. In Callimusa, which is the most variable section (2n=2x=18 or 20), number of 45S rDNA loci varies from two to six signals in diploids. For example, 45S rDNA loci are localized, not only on the two chromosomes with secondary constriction, but also interstitially on another four chromosomes in *M. beccarri* (2n=2x=18). This phenomenon could be caused by chromosome rearrangements during the evolution. In *Ensete gilleti* from genus *Ensete*, even eight chromosomes bearing 45S rDNA loci have been detected.

The number of 5S rDNA sites, which is significantly more variable among individual accessions compared to the 45S rDNA loci, differs between three (in selected species from Rhodochlamys and Callimusa) to twelve (in *M. schizocarpa* from Eumusa) hybridization signals in genera *Musa* and *Ensete*. 5S rDNA loci were detected mostly in distal chromosome regions (Doleželová *et al*., 1998; Osuji *et al*., 1998; Bartoš *et al*., 2005; Čížková *et al*., 2013; 2015).

FISH with simultaneous localization of 45S and 5S rDNA increased the number of identified chromosomes. Odd number of 45S and 5S rDNA loci in some

accessions revealed structural chromosome heterozygosity, which could be caused by hybridization between genotypes with different number of rDNA loci. These findings supported proposed hybrid origin of selected accessions. Another possibility could be due to the variation in intensity of FISH signals caused by lower copy number of rDNA sites, which have been detected. In this case, the signal could be under detection limit of FISH (Doleželová *et al*., 1998; Osuji *et al.*, 1998; Bartoš *et al*., 2005; Čížková *et al*., 2013; 2015).

#### 2.5.2.3 Repetitive DNA sequences

Repetitive DNA sequences are ubiquitous in higher plant genomes. Based on genomic organization, two types of repeats are known. Disperse repeats comprise various transposable elements, which contain coding sequences necessary for their proliferation (Kubis *et al*., 1998; Zhao *et al*., 1998; Chang *et al*., 2008). On the contrary, the majority of tandem repeats have usually no function or coding capacity, but have a strong hybridization signal after FISH on plant chromosomes (Pedersen *et al*., 1996; Tsujimoto *et al*., 1997; Navrátilová *et al*., 2003). Tandemly organized repeats used as FISH probes helped with the identification of individual chromosomes in many plant species, e. g. in barley (Tsujimoto *et al*., 1997), wheat (Pedersen and Langridge, 1997; Badaeva *et al*., 2015), *Vicia* spp. (Navrátilová *et al*., 2003), *Aegilops* spp. (Liu *et al*., 2011), *Festuca pratensis* Huds. (Křivánková *et al*., 2017) or *Agropyron cristatum* (Said *et al*., 2018). Satellite DNA sequences, consisting of thousands of tandemly organized repetitive units, are usually localized in heterochromatin, typical for centromeric or telomeric regions (Charlesworth *et al*., 1994; Schmidt and Heslop-Harrison, 1998; Macas *et al*., 2000; Zatloukalová *et al*., 2011). Satellites displayed specific banding pattern in several species and were used as cytogenetic markers for identification of individual chromosomes (e. g. Sharma and Raina, 2005; Macas *et al*., 2006).

About 55 % of *M. acuminata* and *M. balbisina* genome (Wang *et al*., 2019; Belser *et al*., 2021) and about 67 % of *M. schizocarpa* genome (Belser *et al*., 2018) is represented by various repeats from which LTR retrotransposons are the most abundant. However, bananas are poor in satellites, which comprise only around 3 %

of the genome (Hřibová *et al*., 2010). So far, only few repeats suitable as FISH probes were identified in bananas. From disperse repeats, *gypsy*-like LTR retroelement *monkey*, showing significant homology to gypsy-like LTR from other species, was identified from partial genomic libraries of species *M. acuminata* and *M. balbisiana* and localized by FISH on their chromosomes. Several copies of this element colocalized with 45S rDNA loci and thus did not provide additional cytogenetic marker, while other copies were dispersed throughout the genome (Balint-Kurti *et al*., 2000). Valárik *et al.* (2002) described another repetitive DNA sequences, but FISH signals were visible along all chromosomes or in their (peri)centromeric regions. Hřibová *et al.* (2007) used DNA reassociation kinetics (Cot analysis; Britten and Kohne, 1968; Britten and Davison, 1976) for isolation of the highly repeated fraction of the banana genome. After sequencing of 614 highly repetitive Cot-clones, tandem repeats and various retrotransposons were identified and localized to (peri)centromeric regions and secondary constrictions in *M. acuminata* ssp. *burmannicoides* 'Calcutta 4'. These results were consistent with data obtained by Balint-Kurti *et al.* (2000).

Later on, 454/Roche sequencing platform was used for the characterization of repetitive DNA sequences from *M. acuminata* ssp. *burmannicoides* 'Calcutta 4' and for the analysis of repeat diversity across the whole banana family (Hřibová *et al*., 2010; Novák *et al*., 2014). Majority of retroelements has been dispersed along all banana chromosomes, only a LINE element and a Ty3/Gypsy element MusA1 from Chromoviridae family were localized to (peri)centromeric regions in *M. acuminata* and *M. balbisiana* (Hřibová *et al*., 2010; Neumann *et al*., 2011). A LINE element proved to be a suitable centromere-specific cytogenetic marker in species and clones from Eumusa section of *Musa* (Čížková *et al*., 2013).

In addition to dispersed repeats, two satellites CL18 (~ 2 kb monomer unit) and CL33 (~ 130 bp monomer unit) have been identified *in silico* from 454 sequencing data by Hřibová *et al*. (2010) on specific chromosome loci. These satellites were mapped on mitotic chromosomes of diploid *M. acuminata* ssp. *burmannicoides* 'Calcutta 4' and proved to be feasible cytogenetic markers in further studies (Čížková *et al*., 2013). Between individual banana species, a high sequence

conservation and homology of these satellites have been detected, thus it was not possible to discriminate A and B genomes (Čížková *et al*., 2013).

Karyotype reconstruction in *Musa* based only on dispersed and tandem repetitive DNA sequences used as cytogenetic markers is unfeasible, because most of these sequences do not provide a chromosome-specific pattern.

#### 2.5.2.4 Single copy BAC (bacterial artificial chromosome) clones

BAC clones have been widely used as another type of molecular cytogenetic markers for chromosome identification in plant research, e. g. in cotton (Hanson *et al*., 1995), rice (Jiang *et al*., 1995), barley (Lapitan *et al*., 1997), sorghum (Kim *et al*., 2002), wheat (Zhang *et al*., 2004) or *Brachypodium* (Idziak *et al*., 2014). BAC libraries usually comprise 'single copy' BAC clones carrying inserts about 70 - 200 kb (Vilarinhos *et al*., 2003; Šafář *et al*., 2004; Ortiz-Vazquez *et al*., 2005). However, in some studies long inserts contained repetitive sequences, resulting in dispersed FISH signals. To suppress the non-specific hybridization, blocking DNA or subcloning have been used (Zhang *et al*., 2004; Janda *et al*., 2006).

To identify chromosome-specific BAC clones with low number of repetitive sequences, which could be used as probes for FISH in bananas, the genomic BAC library of *M. acuminata* ssp. *burmannicoides* 'Calcutta 4' (Vilarinhos *et al*., 2003) have been analysed by Hřibová *et al*. (2008). Out of set of eighty selected BAC clones, only one chromosome-specific BAC clone 2G17, which was localized to subtelomeric regions on two chromosomes of diploid *M. acuminata* ssp. *burmannicoides* 'Calcutta 4', was proved to be a suitable cytogenetic marker, giving a strong FISH signal. After FISH with another available cytogenetic markers used in *Musa*, the clone colocalized with genes for 5S rRNA on the same chromosome pair (Hřibová *et al*., 2008). Another tested BAC clones produced dispersed signals along chromosomes or even no signal was detected. Following this, nineteen BAC clones were additionally subcloned, from which low-copy subclones (containing only 5-10 kb inserts) were selected for FISH. One subclone was successfully mapped to a secondary constriction, whereas another three subclones were localized on (peri)centromeric regions of all chromosomes (Hřibová *et al*., 2008).

Thus, only few BAC clones have been successfully mapped on mitotic banana chromosomes (Vilarinhos *et al*., 2006; Hřibová *et al*., 2008). Sequencing of gene-rich BAC clones confirmed the presence of various dispersed repeats, resulting in non-specific FISH signals (Valárik *et al*., 2002; Aert *et al*., 2004; Hřibová *et al*., 2008). Following this, the identification of BAC clones containing only few copies of repeats was under detection limit of Southern hybridization, which was used for their selection (Hřibová *et al*., 2008). Additionally, no FISH signal could be caused by presence of mitochondrial or chloroplast DNA (Vilarinhos *et al*., 2003) and/or due to a high compactness of chromatin of mitotic chromosomes (Hřibová *et al*., 2008). FISH on pachytene chromosomes could solve this problem, which was demonstrated on other species (Lysák *et al*., 2001). Later, de Capdeville *et al*. (2009) successfully mapped few more BAC clones even on banana pachytene chromosomes.

The availability of reference genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang' (D'Hont *et al*., 2012) allowed the selection of BAC clones from a BamH1 and HindIII BAC libraries of 'DH Pahang' (D'Hont *et al*., 2012; http://banana-genome.cirad.fr) and the implementation of BAC-FISH for validation of translocations detected previously *in silico* (Martin *et al*., 2017; Martin *et al*., 2020b).

#### 2.5.2.5 Chromosome painting

Chromosome painting, which was first used by Pinkel *et al*. (1988), is defined as labeling of entire homologous chromosomes or their regions using chromosome-specific DNA probes. Since the identification of individual chromosomes is crucial for successful cytogenetic studies, this powerful technique became popular in human and animal molecular cytogenetics. For instance, using combinatorial multi-fluor FISH with chromosome-specific painting probes, inter-chromosomal rearrangements have been identified in human karyotypes (Speicher *et al.*, 1996). Arrangement of chromosome territories in chicken or mammalian nuclei (Cremer and Cremer, 2001; Habermann *et al*., 2001) or mammalian karyotype evolution have been also studied using chromosome-specific painting (Ferguson-Smith and Trifonov, 2007).

In human research, chromosome-specific painting probes can be obtained from DNA of flow sorted (Cremer *et al.*, 1988) or microdissected chromosomes (Meltzer *et al*., 1992). To prevent the cross-hybridization of repetitive DNA, an excess of unlabeled sheared total genomic DNA or a fraction enriched for highly repetitive sequences (Cot-1 DNA) is usually used. This technique was named chromosomal *in situ* suppression (CISS) (Lichter *et al*., 1988).

However, despite the great efforts, all attempts to provide chromosome-specific (unique) probes in plants have failed due to high complexity of plant genomes (Fuchs *et al*., 1996a). Many plant species contain large amount of repetitive DNA sequences, which are dispersed across the genome. In many studies, blocking of these repeats was not successful and thus resulted in their cross-hybridization (Schubert *et al*., 2001).

2.5.2.5.1 Alternative approaches of chromosome painting used in plants

As mentioned above, the application of classical chromosome painting techniques, which are frequently used in human and animal research, is not possible in plants. However, painting of the terminal heterochromatic segment of microdissected B chromosomes of *Secale cereale* (Houben *et al*., 1996) and microdissected Y chromosome of *Rumex acetosa* (Shibata *et al*., 1999) was succesfully performed in plants. In this case, chromosome painting was based on chromosome-specific dispersed repetitive sequences, which were not present on A chromosomes or autosomes, respectively (Lysák *et al*., 2001; Schubert *et al*., 2001). Schubert *et al*. (2001) also indicated possible occurrence of repetitive spacer elements of the 45S rRNA genes, which could be transposed from NOR to other parts of the satellite chromosome (e. g. Fuchs *et al*., 1996b).

Another possibility, how to paint plant chromosomes, is a modification of FISH called genome *in situ* hybridization (GISH). This method, which was established in plants by Schwarzacher *et al.* (1989), enables discrimination and thus the identification of parental chromosomes in inter-specific hybrids and their progeny (Lysák *et al*., 2001). Painting of plant chromosomes using GISH is achieved by labeling of total genomic DNA of one parent, which is used as a

probe. Alien chromosomes are distinguished on the basis of divergent dispersed repeats (Schubert *et al*., 2001). The origin of natural amphiploids, the introgression of alien chromosomes or homeologous pairing and thus an exchange between genomes in hybrids have been studied (e. g. Jiang and Gill, 1994; Gill and Friebe, 1998; Silva and Souza, 2013; Kopecký *et al.,* 2008; 2017).

GISH have been also used for discrimination of parental chromosomes in inter-specific cultivated banana clones (Osuji *et al*., 1997; D'Hont *et al*., 2000). Unfortunately, only (peri)centromeric regions were painted, whereas distal parts of the chromosomes remained unlabeled. This could be caused by uneven repeat sequence content, because repeats are usually more abundant in (peri)centromeric regions of chromosomes (D'Hont, 2005).

For the first time, A and B genomes in cultivated triploid (AAB, ABB) clones and artificial *Musa* hybrids were distinguished using GISH with labeled total genomic DNA from A or B genome (Osuji *et al*., 1997). Data obtained were in agreement with the classification of plantains (AAB) and cooking bananas (ABB) by phenotypic descriptors (Simmonds and Shepherd, 1955). D'Hont *et al*. (2000) determined chromosomes belonging to four genomes (A, B, S, T) of studied banana cultivars. Results corresponded with presumable genetic constitution of selected hybrids based on their ploidy level, morphological and molecular data. One exception was Pelipita clone with ABB genome composition, in which 8 A chromosomes and 25 B chromosomes were detected, instead of presumed 11 A and 22 B chromosomes. Despite the labeling of (peri)centromeric regions of the chromosomes, this observation can indicate presence of some irregularities during meiosis (D'Hont *et al*., 2000).

However, this method has limitations due to non-uniform chromosome labeling and cross-hybridization of the probe between different genomes, thus the identification of chromosomal rearrangements using GISH in *Musa* would not be possible. The cross-hybridization occurred mostly between A and S genomes, followed by A and B genomes. Much smaller cross-hybridization was detected between more distantly related species *M. acuminata* from Eumusa section (A genome) and *M. textilis* (T genome) from the Australimusa section. This

phenomenon could be connected to their variable genome size and putative presence of different DNA repeats. D'Hont *et al*. (2000) pointed to the critical genome size and repeat content, which could cause problems during GISH labeling. The level of cross-hybridization indicates that two species have similar compositions of repeats, which give most of the GISH signals, and thus there is a very close relationship between these genomes (Osuji *et al*., 1997).

Using GISH, Jeridi *et al.* (2011) also analysed chromosome pairing in meiosis in triploid inter-specific clones (AAB, ABB), and confirmed frequent inter-specific recombinations between A and B genomes. Again, hybridization signals were limited to (peri)centromeric regions on meiotic chromosomes. The findings were consistent with the results obtained on mitotic chromosomes (Osuji *et al.*, 1997; D'Hont *et al*., 2000). Based on variation in labeling intensity, Jeridi *et al*. (2011) also confirmed a higher content of species-specific repeats in *M. acuminata* (having larger genome), compared to *M. balbisiana*, by the usage of probes derived from A and B genomes, which was already proposed by Baurens (1997).

2.5.2.5.2 Chromosome-specific painting using bulked BAC clones

For the first time, Lysák *et al*. (2001) demonstrated chromosome-specific painting technique in plants, which was based on pooling of BAC clones covering specific chromosome of *Arabidopsis thaliana*. Single- or low-copy BAC clones, representing minimum tiling path (a minimal number of BAC clones covering selected chromosome), were fluorescently labeled and used as a probe for painting the entire chromosome. One chromosome arm of one of the relatively small chromosomes of *A. thaliana* was painted by labeling of more than hundred BACs (Lysák *et al*., 2001).

This technique requires whole genome sequence obtained by BAC by BAC (clone by clone) sequencing approach. Ordered BAC contigs, covering entire genome, were used for the identification of chromosome-specific BAC clones and their further localization *in situ*. However, the method is applicable only in species with small genomes and low number of repetitive DNA sequences, whose genomes were assembled by clone-by-clone approach - *Arabidopsis thaliana* (125 Mb) and

*Brachypodium distachyon* (~ 300 Mb) (The *Arabidopsis* Genome Initiative, 2000; The International *Brachypodium* Initiative, 2010). Genome of *A. thaliana* is largely euchromatic, only centromeric regions are heterochromatic, thus BACs containing these sequences were excluded from the probe preparation (Lysák *et al*., 2001; Jiang and Gill, 2006).

Chromosome painting in *A. thaliana* was initially used for tracking of individual chromosomes in the interphase nuclei. Later on, spatial arrangement and functional properties of chromatin domains were elucidated (Fransz *et al*., 2002; Pečinka *et al*., 2004). Moreover, chromosome-specific painting probes obtained from one species can be hybridized on chromosomes of other closely related species, allowing comparative cytogenetic studies across whole families. Genome multiplications, chromosomal rearrangements and evolution in Brassicaceae family were studied using BAC-based probes developed for *A. thaliana* (Lysák *et al*., 2005; 2006; Mandáková and Lysák, 2008; Mandáková *et al*., 2010). Few years later, BAC-painting FISH was also successfully applied in *B. distachyon* (Betekhtin *et al*., 2014; Idziak *et al*., 2014).

Despite the small size of nuclear genome and quite low proportion of repetitive DNA sequences in *Musa*, FISH with selected BAC clones resulted mostly in dispersed or no signals on chromosomes. Moreover, different approaches used for creation of whole genome sequences (D'Hont *et al*., 2012; Belser *et al*., 2018; Wang *et al*., 2019), which were not based on BAC clones sequencing, were not applicable for chromosome-specific painting using a bulked BAC clones method.

#### 2.5.2.5.3 Chromosome-specific oligo painting

For many years, the application of FISH techniques has been limited by the lack of robust DNA probes, especially in non-model plant species (Jiang, 2019). Until recently, BAC-FISH was the only method for chromosome painting in a few plant species. However, progress in next generation sequencing technologies (NGS) allowed the production of reference genome assemblies in many plants, even in non-model species.

Technical advances in DNA synthesis enabled recent development of the method called oligo painting FISH using synthetic oligo probes designed from single-copy DNA sequences. This method is applicable in any species with a sequenced genome and allows painting of entire chromosomes (Jiang, 2019). FISH with these oligo panting probes was previously applied in mammalian and *Drosophila* species (Boyle *et al*., 2011; Yamada *et al*., 2011; Beliveau *et al*., 2012). In 2015, Han *et al.* introduced this method for the first time in plants.

The method requires the availability of reference genome sequence, from which thousands of short (~ 45-nt long), single-copy oligonucleotides (oligos) are computationally identified. Oligos specific to a chromosome region, a whole chromosome or multiple chromosomes are then synthesized in parallel as a pool, fluorescently labeled and used as a probe for FISH. Probes designed from one species can be applied among related species, enabling comparative karyotype analysis and the identification of chromosomal rearrangements during their evolution and speciation. These oligonucleotide-based probes, which are highly versatile, give new possibilities of the FISH application especially in non-model plant species, thus represent the most significant advance of FISH (Jiang, 2019).

The bioinformatic pipeline called Chorus (https://github.com/forrestzhang/Chorus) was developed by Han *et al*. (2015) to design chromosome-specific probes using bulked oligos. Chorus software allowed the identification of non-overlapping and unique oligos, which are specific to entire chromosome or chromosomal region. Oligos containing repetitive DNA sequences or sequences located on other chromosomes are excluded. Oligo length, sequence similarity and melting temperature (Tm) of the oligos can be edited in the pipeline to control the stringency of oligo selection. Set of unique oligos is then selected according to their density, which represents the number of oligos per kb, and location of oligos along individual chromosome. All selected oligos, comprising genomic sequence (usually 45 – 48-nt long), forward primer with T7 RNA promotor sequence and reverse primer are synthesized as a pool by the MYcroarray (Ann Arbor, Michigan, USA; Figure 7A). Oligo library is amplified by emulsion PCR and the PCR product is then used for T7 *in vitro* transcription. The resulting RNA is reverse-transcribed using a biotin- or

digoxigenin-labeled reverse primer, producing RNA:DNA hybrids. In the last step, RNA is hydrolyzed and single-stranded labeled oligomers are directly used as a FISH probe (Figure 7B). After FISH, hybridization signals are clearly visible on somatic metaphase chromosomes and even on pachytene chromosomes, but with less signal intensity. As expected, centromeric regions are not labeled (Han *et al*., 2015).

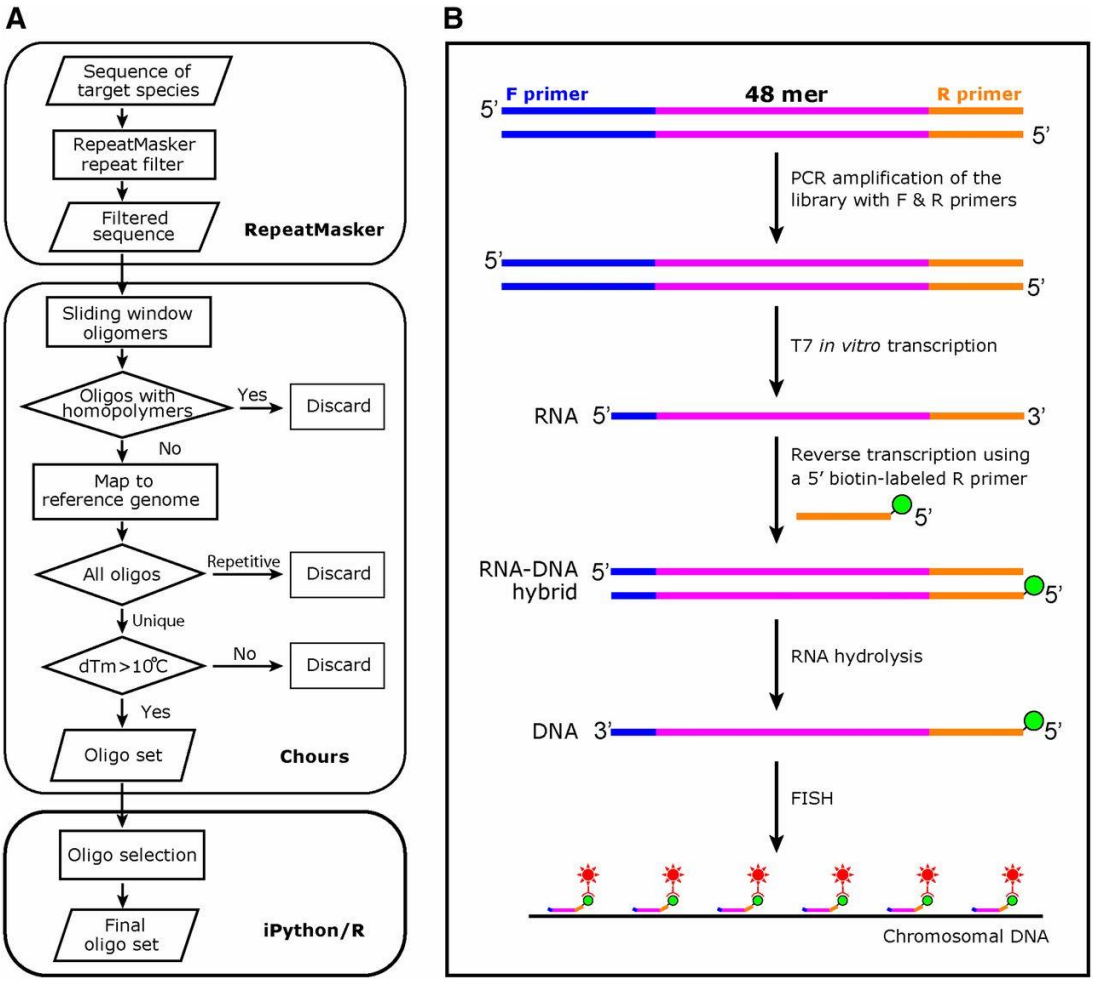



**Figure 7.** Development of oligo painting FISH probes using pools of chromosome-specific oligos. (A) The pipeline for oligo selection (B) The amplification and labeling of oligo pools for oligo painting FISH. (Han *et al*., 2015).

So far, oligo painting probes were developed and successfully hybridized on somatic metaphase chromosomes, meiotic pachytene chromosomes or interphase nuclei of many plant species. Oligo panting FISH technique has already been applied for reconstruction of molecular karyotypes and identification of chromosomal

rearrangements, for comparative chromosome painting, analysis of chromosome pairing and transmission in meiosis or for examination of the quality of genome assemblies e. g. in *Cucumis* spp. (Han *et al*., 2015; Bi *et al*., 2020), *Fragaria* spp. (Qu *et al*., 2017), *Aquilegia coerulea* (Filiault *et al*., 2018), *Solanum* spp. (Braz *et al*., 2018; He *et al*., 2018), *Oryza* spp. (Hou *et al*., 2018; Liu *et al.*, 2020a), *Saccharum spontaneum* (Meng *et al*., 2018; 2020), *Populus* spp. (Xin *et al*., 2018; 2020), *Zea* spp. (Albert *et al*., 2019; do Vale Martins *et al*., 2019; Braz *et al*., 2020b), *Citrus* spp. (He *et al*., 2020), cotton (Liu *et al*., 2020b); family Triticeae (Song *et al.*,2020) or subtribe Phaseolineae (do Vale Martins *et al*., 2021).

Two classes of oligo probes have been demonstrated in plant cytogenetic studies (He *et al*., 2020). Oligo probes, which are designed for labeling of entire chromosome allow tracking and identification of only a single chromosome (e. g. Han *et al.*, 2015; He *et al.*, 2018; Hou *et al.*, 2018; Albert *et al.*, 2019). However, oligo probes can be designed also for multiple regions of multiple chromosomes. For example, Braz *et al*. (2018) did not label entire chromosomes with chromosome-specific oligo probes, but so-called oligo-FISH barcode system was used in the study. This system was based on selection of only two oligomer libraries, which generated distinct FISH signals in specific locations represented about 250 kb long regions on individual chromosome in potato genome, and which after FISH enable unambiguous identification of all chromosomes (Figure 8).

The same set of oligo probes was used for chromosome identification in related *Solanum* species, allowing comparative karyotype analysis. Later on, the barcoding system was used in other studies (e. g. Braz *et al*., 2018; 2020b; Meng *et al.*, 2020). Nevertheless, using this barcoding system, only larger chromosome translocations could be revealed (Braz *et al*., 2018), but detection of smaller chromosomal rearrangements is not possible, compare to oligo probes covering entire chromosomes.

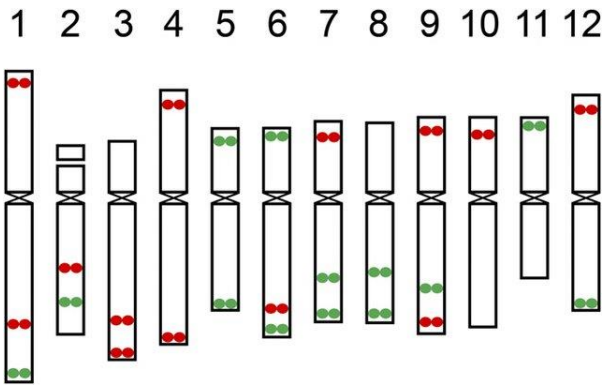



**Figure 8:** Predicted locations of the oligo-FISH signals on 12 potato chromosomes. Oligos were selected from 13 red and 13 green chromosomal regions. All chromosomes can be identified based on number and location of the red/green signals (Braz *et al*., 2018).

#### 2.5.2.6 Karyotype reconstruction in banana

Using multicolor FISH, probes for 45S rDNA, 5S rDNA, a BAC clone 2G17, a LINE element and two satellite repeats (CL18 and CL33) were successfully mapped on several *Musa* accessions including representatives of *M. acuminata*, *M. balbisiana*, *M. schizocarpa* and their inter-specific hybrids. A LINE element was used for the identification of primary constrictions, which are not always clearly visible on small, condensed chromosomes. In all tested *Musa* accessions, a probe for 45S rDNA was localized on chromosomes bearing secondary constriction and corresponded to their ploidy level. Idiograms of three diploid species, *M. acuminata* ssp. *malaccensis* 'DH Pahang' ITC 1511, *M. balbisiana* 'Pisang Klutuk Wulung' and *M. schizocarpa* 'Schizocarpa' ITC 0560, are depicted on Figure 9 (Čížková *et al*., 2013).

In diploid *M. acuminata*, satellite CL18 was detected in subtelomeric region on two chromosomes and co-localized with BAC clone 2G17. CL33 was localized in subtelomeric region on four chromosomes, except *M. acuminata* 'Maia Oa', in which only one chromosome pair was bearing CL33. Moreover, CL33 co-localized on two chromosomes with CL18, a BAC clone 2G17 and 5S rDNA in all diploid *M. acuminata* accessions. 5S rRNA genes, which are more diverse in number of

hybridization signals between individual accessions, were detected on another four chromosomes in *M. acuminata* ‘DH Pahang’ (Figure 9A; Čížková *et al*., 2013).

Compared to diploid *M. acuminata*, similar hybridization signals have been observed also in diploid *M. schizocarpa* accessions, one exception being the probe for 5S rRNA genes, which was localized on twelve chromosomes. Interestingly, two signals of 5S rDNA co-localized with 45S rDNA in *M. schizocarpa* ‘Schizocarpa’ (Figure 9B; Čížková *et al*., 2013).

In diploid *M. balbisiana*, CL33 was not detected in any tested accessions. CL18 was localized on four chromosomes with one exception being *M. balbisiana* ‘Cameroun’, in which CL18 was localized on only three chromosomes, indicating a structural chromosome heterozygosity. A BAC clone 2G17 co-localized with satellite CL18 on two chromosomes. However, *M. balbisiana* ‘Honduras’ contained a BAC clone 2G17 on even four chromosomes, which could be caused by the locus duplication. Another two chromosomes bearing CL18 co-localized with 5S rDNA. 5S rDNA was detected on additional four chromosomes in *M. balbisiana* ‘Pisang Klutuk Wulung’ (Figure 9C; Čížková *et al*., 2013).

However, a combination of these probes allows the identification of only few chromosomes (Figure 9). Although the quite high number of chromosomes carried probes for 5S rDNA in all species, these probes hybridized to similar chromosome regions, thus chromosomes could not be discriminated from each other and the number of identified chromosomes was not increased. Due to the lack of cytogenetic markers specific for different genomes (A, B, S or T), the identification of homologous chromosomes in inter-specific hybrids was not possible (Čížková *et al*., 2013).

Even the availability of reference genome sequence of *M. acuminata* ssp. *malaccensis* ‘DH Pahang’ (D’Hont *et al*., 2012; Martin *et al*., 2016) and resequencing of 123 wild and cultivated *Musa* accessions (Dupouy *et al*., 2019) did not facilitate the design of another cytogenetic markers for FISH, which would help the identification of all chromosomes in *Musa* and the anchoring of pseudomolecules of reference genome sequence to individual chromosomes *in situ*. So far, only few

chromosomes were identified in banana and complete karyotype is still lacking. The number of cytogenetic markers needs to be increased for the identification of all banana chromosomes within a karyotype, which is important for development of molecular karyotypes and mainly for the identification of structural chromosome changes in *Musa* spp.

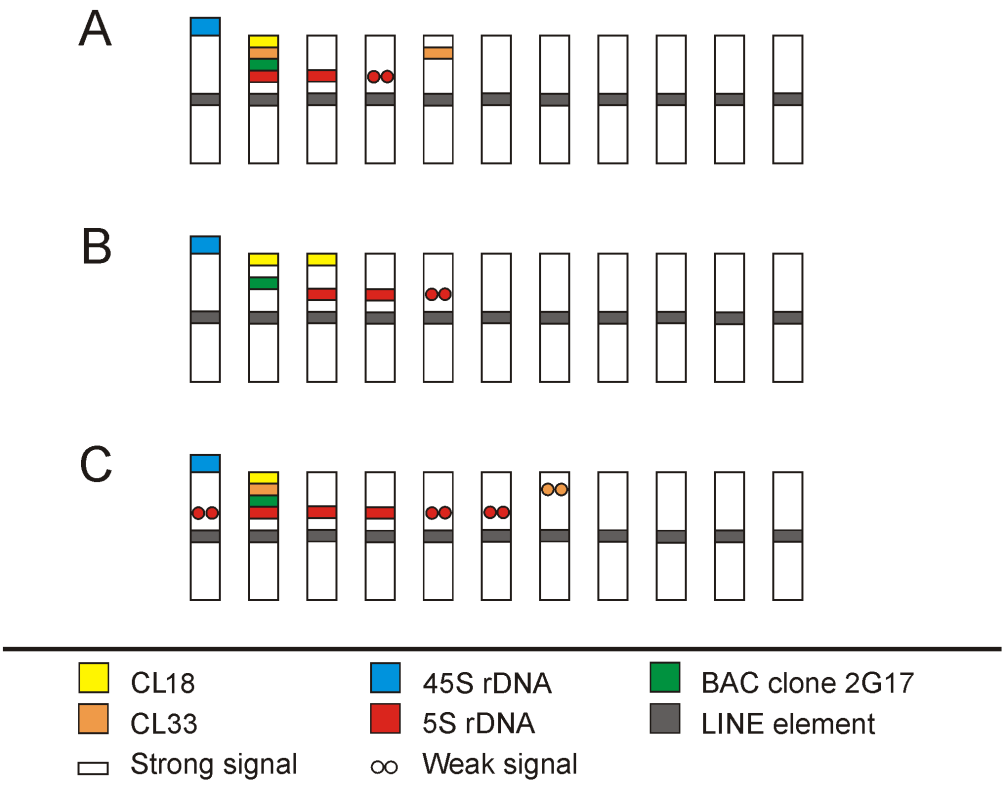



**Figure 9.** Idiograms of three diploid accessions of genus *Musa*. (A) *M. acuminata* ssp. *malaccensis* 'DH Pahang' ITC 1511; (B) *M. balbisiana* 'Pisang Klutuk Wulung'; (C) *M. schizocarpa* 'Schizocarpa' ITC 0560 (Čížková *et al*., 2013).

## 2.6 References


**Adheka, J.G., Dhed'a, D.B., Karamura, D., Blomme, G., Swennen, R. and De Langhe, E.** (2018). The morphological diversity of plantain in the Democratic Republic of Congo. *Sci. Hort*. **234**:126-133. doi: 10.1016/j.scienta.2018.02.034

**Aert, R., Sagi, L. and Volckaert, G.** (2004). Gene content and density in banana (*Musa acuminata*) as revealed by genomic sequencing of BAC clones. *Theor. Appl. Genet*. **109**:129-139. doi: 10.1007/s00122-004-1603-2

**Albert, P.S., Zhang, T., Semrau, K., Rouillard, J.M., Kao, Y.H., Wang, C.R., *et al*.** (2019). Whole-chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal relationships. *Proc. Natl. Acad. Sci. USA*. **116**:1679-1685. doi: 10.1073/pnas.1813957116

**Amorim, E.P., dos Santos-Serejo, J.A., Amorim, V.B.O., Ferreira and C.F., Silva, S.O.** (2013). Banana breeding at embrapa cassava and fruits. *Acta. Hortic*. **968**:171-176. doi: 10.17660/ActaHortic.2013.986.18

**Amorim, E.P., Silva, S.O., Amorim, V.B.O. and Pillay, M.** (2011). Quality improvement of cultivated *Musa*. In: Pillay, M., Tenkouano, A., eds. Banana Breeding: Progress and Challenges. New York: CRC Press, 252-280.

**Ansari, H.A., Ellison, N.W., Reader, S.M., Badaeva, E.D., Friebe, B., Miller, T.E., *et al*.** (1999). Molecular cytogenetic organization of 5S and 18S-26S rDNA loci in white clover (*Trifolium repens* L.) and related species. *Ann. Bot*. **83**:199-206. doi: 10.1006/anbo.1998.0806

**Appels, R., Gerlach, W.L., Dennis, E.S., Swift, H. and Peacock, W.J.** (1980). Molecular and chromosomal organization of DNA sequences coding for the ribosomal RNAs in cereals. *Chromosoma*. **78**:293-311. doi: 10.1007/BF00327389

**Argent, G.C.G.** (1976). The wild bananas of Papua New Guinea. Notes R. Bot. Gard. Edinburgh. **35**:77-114

**Badaeva, E.D., Amosova, A.V., Goncharov, N.P., Macas, J., Ruban, A.S., Grechishnikova, I.V., *et al*.** (2015). A Set of Cytogenetic Markers Allows the Precise Identification of All A-Genome Chromosomes in Diploid and Polyploid Wheat. *Cytogenet. Genome Res*. **146**:71-79. doi: 10.1159/000433458

**Badaeva, E.D., Friebe, B. and Gill, B.S.** (1996). Genome differentiation in *Aegilops*. 2. Physical mapping of 5S and 18S-26S ribosomal RNA gene families in diploid species. *Genome*. **39**:1150-1158. doi: 10.1139/g96-145

**Baker, J.G.** (1893). A synopsis of the genera and species of Museae. *Ann. Bot*. **7**:189-229. doi: 10.1093/aob/os-7.2.189

**Baker, R.E.D. and Simmonds, N.W.** (1953). The genus *Ensete* in Africa. Kew Bulletin. **8**:405-416, with a correction in Kew Bulletin **8**:574. doi: 10.2307/4115529

**Bakry, F., Carreel, F., Jenny, C. and Horry, J.P.** (2009). Genetic Improvement of Banana. In: Jain, S.M., Priyadarshan, P.M. (eds). Breeding Plantation Tree Crops: Tropical Species. Springer, New York, NY. doi: 10.1007/978-0-387-71201-7_1

**Bakry, F. and Horry, J.P.** (1992). Tetraploid hybrids from interploid 3x/2x crosses in cooking bananas. *Fruit*s. **47**:641-655.

**Bakry, F. and Horry, J.P.** (1994). *Musa* breeding at CIRAD-FLHOR, pp. 169-175. In: The Improvement and Testing of *Musa*: A Global Partnership (Jones, D.R. Ed.), INIBAP, Montpellier, France.

**Bakry, F., Carreel, F., Caruana, M.L., Cote, F., Jenny, C. and Tezenas du Montcel, H.** (2001). Banana, pp. 1-29. In: Tropical Plant Breeding (Charrier, A., Jacquot, M., Hamon, S., and Nicolas, D. Eds.), CIRAD Publisher, Collection Reperes, Montpellier, France.

**Balint-Kurti, P., Clendennen, S., Doleželová, M., Valárik, M., Doležel, J., Beetham, P.R., *et al.*** (2000). Identification and chromosomal localization of the *monkey* retrotransposon in *Musa* sp. *Mol. Gen. Genet*. **263**:908-915. doi: 10/1007/s004380000265

**Bartoš, J., Alkhimova, O., Doleželová, M., De Langhe, E. and Doležel, J.** (2005). Nuclear genome size and genomic distribution of ribosomal DNA in *Musa* and *Ensete* (Musaceae): taxonomic implications. *Cytogenet. Genome Res*. **109**:50-57. doi: 10.1159/000082381

**Batte, M., Swennen, R., Uwimana, B., Akech, V., Brown, A., Tumuhimbise, R., *et al.*** (2019). Crossbreeding East African Highland Bananas: Lessons learnt relevant to the botany of the crop after 21 years of genetic enhancement. *Front. Plant Sci.* **10**:81. doi: 10.3389/fpls.2019.00081

**Baurens, F.C.** (1997). Identification par PCR des espèces impliquées dans la composition génomique des cultivars de bananier, à l'aide de séquences répétées = PCR identification of the species involved in the genomic composition of the banana cultivar using repeated sequences. Thèse de doctorat, Université de Toulouse 3, Toulouse, France.

**Baurens, F.C., Martin, G., Hervouet, C., Salmon, F., Yohomé, D., Ricci, S., *et al.*** (2019). Recombination and large structural variations shape interspecific edible bananas genomes. *Mol. Biol. Evol*. **36**:97-111. doi: 10.1093/molbev/msy199

**Beliveau, B.J., Joyce, E.F, Apostolopoulos, N., Yilmaz, F., Fonseka, C.Y., McCole, R.B., *et al.*** (2012). Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. USA*. **109**:21301-21306. doi: 10.1073/pnas.1213818110.

**Belser, C., Baurens, F.C., Noel, B., Martin, G., Cruaud, C., Istace, B., *et al.*** (2021). Telomere-to-telomere gapless chromosomes of banana using nanopore sequencing. doi.org/10.1101/2021.04.16.440017

**Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F.C., Falentin, C., *et al.*** (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat. Plants*. **4**:879-887. doi: 10.1038/s41477-018-0289-4

**Betekhtin, A., Jenkins, G. and Hasterok, R.** (2014). Reconstructing the evolution of *Brachypodium* genomes using comparative chromosome painting. *PLoS ONE*. **9**:e115108. doi: 10.1371/journal.pone.0115108

**Bi, Y., Zhao, Q., Yan, W., Li, M., Liu, Y., Cheng, Ch., *et al.*** (2020). Flexible chromosome painting based on multiplex PCR of oligonucleotides and its application for comparative chromosome analyses in *Cucumis*. *Plant J.* **102**:178-186. doi: 10.1111/tpj.14600

**Bohra, P., Waman, A.A., Sathyanarayana, B.N., Umesha, K., Anu, S.R., Swetha, H.G., *et al.*** (2014). Aseptic culture establishment using antibiotics with reference to their efficiency and phytotoxicity in difficult-to-establish native Ney Poovan banana (*Musa*, AB). *Proc. Natl. Acad. Sci. India Sect. B: Biol. Sci*. **84**:257-261. doi: 10.1007/s40011-013-0220-8

**Borrell, J.S, Biswas, M.K., Goodwin, M., Blomme, G., Schwarzacher, T., Heslop-Harrison, J.S.P, *et al.*** (2019). Enset in Ethiopia: a poorly characterized but resilient starch staple. *Ann Bot*. **123**:747-766. doi: 10.1093/aob/mcy214

**Boyle, S., Rodesch, M.J., Halvensleben, H.A., Jeddeloh, J.A. and Bickmore, W.A.** (2011). Fluorescence *in situ* hybridization with high-complexity repeat-free oligonucleotide probes generated by massively parallel synthesis. *Chromosome Res*. **19**:901-909. doi: 10.1007/s10577-011-9245-0

**Brandt, S.A., Spring, A., Hiebsch, C., McCabe, J.T., Tabogie, E., Diro, M., *et al.* (1997).** The "Tree Against Hunger": enset-based agricultural system in Ethiopia. Washington, DC: American Association for the Advancement of Science.

**Braz, G.T., He, L., Zhao, H., Zhang, T., Semrau, K., Rouillard, J.M., *et al.*** (2018). Comparative oligo-FISH mapping: An efficient and powerful methodology to reveal karyotypic and chromosomal evolution. *Genetics*. **208**:513-523. doi: 10.1534/genetics.117.300344

**Braz, G.T., Martins, L.D., Zhang, T., Albert, P.S., Birchler, J.A. and Jiang, J.M**. (2020b). A universal chromosome identification system for maize and wild *Zea* species. *Chromosome Res*. **28**:183-194. doi: 10.1007/s10577-020-09630-5

**Braz, G.T., Yu, F., do Vale Martins, L. and Jiang, J.** (2020a). Fluorescent *In Situ* Hybridization Using Oligonucleotide-Based Probes. *Methods Mol. Biol*. **2148**:71-83. doi: 10.1007/978-1-0716-0623-0_4

**Britten, R.J. and Davidson, E.H.** (1976). Studies on nucleic acid reassociation kinetics: empirical equations describing DNA reassociation. *Proc. Natl. Acad. Sci. USA.* **73**:415-419. doi: 10.1073/pnas.73.2.415

**Britten, R.J. and Kohne, D.E.** (1968). Repeated sequences in DNA. *Science*. **161**:529-540. doi: 10.1126/science.161.3841.529

**Brown, A., Carpentier, S.C. and Swennen, R.** (2020). Breeding climate-resilient bananas. Chapter **4**:91-115. In: Kole, C. (eds) Genomic designing of climate-smart fruit crops. Springer, Cham. doi: 10.1007/978-3-319-97946-5_4

**Brown, A., Tumuhimbise, R., Amah, D., Uwimana, B., Nyine, M., Mduma, H., *et al*.** (2017). The genetic improvement of bananas and plantains (*Musa* spp.). In: Genetic Improvement of Tropical Crops (Campos, H., and Caligari, P.D.S. eds): Springer, Cham, pp. 219-240.

**Buggs, R.J., Elliott, N.M., Zhang, L., Koh, J., Viccini, L.F., Soltis, D.E., *et al*.** (2010). Tissue-specific silencing of homoeologs in natural populations of the recent allopolyploid *Tragopogon mirus*. *New Phytol*. **186**:175-83. doi: 10.1111/j.1469-8137.2010.03205.x

**Burgos-Hernández, M., González, D. and Castillo-Campos, G.** (2017). Phylogenetic position of the disjunct species *Musa ornata* (Musaceae): first approach to understand its distribution. *Genet. Resour. Crop Evol*. **64**:1889-1904. doi.org/10.1007/s10722-016-0479-8

**Burke, J.M. and Arnold, M.L.** (2001). Genetics and the fitness of hybrids. *Annu. Rev. Genet.* **35**:31-52. doi: 10.1146/annurev.genet.35.102401.085719

**Carreel, F.** (1994). Etude de la diversité génétique des bananiers (genre *Musa*) à l'aide des marqueurs RFLP, PhD Thesis, Institut National Agronomique, Paris-Grignon

**Carreel, F., Fauré, S., González de León, D., Lagoda, P.J.L., Perrier, X., Bakry, F., *et al*.** (1994). Evaluation of the genetic diversity in diploid bananas (*Musa* sp.*). Genet. Sel. Evol*. **26**:125-136. doi: 10.1051/gse:19940709

**Carreel, F., Gonzalez de Leon, D., Lagoda, P., Lanaud, C., Jenny, C., Horry, J., *et al*.** (2002). Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome.* **45**:679-692. doi: 10.1139/g02-033

**Caspersson, T., Farbar, S., Foley, G.E., Kudynowski, J., Modest, E.J., Simonsson, E., *et al.*** (1968). Chemical differentiation along metaphase chromosomes. *Exp. Cell Res*. **49**:219-222. doi: 10.1016/0014-4827(68)90538-7

**Cenci, A., Sardos, J., Hueber, Y., Martin, G., Breton, C., Roux, N., *et al.*** (2020). Unravelling the complex story of intergenomic recombination in ABB allotriploid bananas. *Ann. Bot*. **127**:7-20. doi: 10.1093/aob/mcaa032

**Chabannes, M., Baurens, F.C., Duroy, P.O., Bocs, S., Vernerey, M.S., Rodier-Goud, M., *et al.*** (2013). Three infectious viral species lying in wait in the banana genome. *J. Virol*. **87**:8624-8637. doi: 10.1128/JVI.00899-13

**Chang, S.B., Yang, T.J., Datema, E., van Vugt, J., Vosman, B., Kuipers, A., *et al.*** (2008). FISH mapping and molecular organization of the major repetitive sequences of tomato. *Chromosome Res*. **16**:919-933. doi:10.1007/s10577-008-1249-z

**Charlesworth, B., Sniegowski, P. and Stephan, W.** (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature*. **371**:215-220. doi: 10.1038/371215a0

**Cheesman, E.E.** (1947). Classification of the bananas. The genus *Ensete* Horan and the genus *Musa* L. Kew Bull. **2**:97-117.

**Cheesman, E.E.** (1948). Classification of the bananas. Kew Bull. **3**:145-153.

**Cheesman, E.E. and Larter, L.N.H.** (1935). Genetical and cytological studies of *Musa*. III. Chromosome numbers in the Musaceae. *J. Genet*. **30**:31-52. doi: 10.1007/BF02982204

**Christelová, P., De Langhe, E., Hřibová, E., Čížková, J., Sardos, J., Hušáková, M., *et al.*** (2017). Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. *Biodivers. Conserv.* **26**:801-824. doi: 10.1007/s10531-016-1273-9

**Christelová, P., Valárik, M., Hřibová, E., De Langhe, E. and Doležel, J.** (2011). A multi gene sequence-based phylogeny of the Musaceae (banana) family. *BMC Evol. Biol*. **11**:103. doi: 10.1186/1471-2148-11-103

**Čížková, J., Hřibová, E., Christelová, P., Van den Houwe, I., Häkkinen, M., Roux, N., *et al.*** (2015). Molecular and cytogenetic characterization of wild *Musa* species. *PLoS One*. **10**:e0134096. doi: 10.1371/journal.pone.0134096

**Čížková, J., Hřibová, E., Humplíková, L., Christelová, P., Suchánková, P. and Doležel J.** (2013). Molecular analysis and genomic organization of major DNA satellites in banana (*Musa* spp.). *PLoS One*. **8**:e54808. doi: 10.1371/journal.pone.0054808

**Cremer, T. and Cremer, C.** (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet*. **2**:292-301. doi: 10.1038/35066075

**Cremer, T., Lichter, P., Borden, J., Ward, D.C. and Manuelidis, L.** (1988). Detection of chromosome aberrations in metaphase and interphase tumor cells by *in situ* hybridization using chromosome-specific library probes. *Hum. Genet*. **80**:235-246. doi: 10.1007/BF01790091

**D'Hont, A.** (2005). Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenet. Genome Res*. **109**:27-33. doi: 10.1159/000082378

**D'Hont, A., Denoeud, F., Aury, J.M., Baurens, F.C., Carreel, F., Garsmeur, O., *et al.*** (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*. **488**:213-217. doi: 10.1038/nature11241

**D'Hont, A., Paget-Goy, A., Escoute, J. and Carreel, F.** (2000). The interspecific genome structure of cultivated banana, *Musa* spp. revealed by genomic DNA *in situ* hybridization. *Theor. Appl. Genet*. **100**:177-183. doi: 10.1007/s001220050024

**Daniells J., Jenny, C., Karamura, D. and Tomekpe, K**. (2001). Musalogue: a catalogue of *Musa* germplasm. Diversity in the genus *Musa* (Arnaud, E. and Sharrock, S. compil.). International Network for the Improvement of Banana and Plantain, Montpellier, France.

**Daniells, J.** (2003). Bananas and Plantains. Encyclopedia of Food Sciences and Nutrition, 2nd edition, 372-378. doi:10.1016/b0-12-227055-x/00080-8

**Danilova, T.V., Friebe, B. and Gill, B.S.** (2012): Single-copy gene Fluorescence *in situ* hybridization and genome analysis: Acc-2 loci mark evolutionary chromosomal rearrangements in wheat. *Chromosoma*. **121**:597-611. doi: 10.1007/s00412-012-0384-7

**Davey, M.W., Gudimella, R., Harikrishna, J.A., Sin, L.W., Khalid, N. and Keulemans, J.** (2013). A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. *BMC Genomics.* **14**:683. doi: 10.1186/1471-2164-14-683

**De Capdeville, G., Souza Junior, M.T., Szinay, D., Diniz, L.E.C., Wijnker, E. and Swennen, R., *et al.*** (2009). The potential of high-resolution BAC-FISH in banana breeding. *Euphytica*. **166**:431-443. doi: 10.1007/s10681-008-9830-2

**De Langhe E., Pillay, M., Tenkouano, A. and Swennen, R.** (2005). Integrating morphological and molecular taxonomy in *Musa*: the African plantains (*Musa* spp. AAB group). *Plant. Syst. Evol.* **255**:225-236. doi: 10.1007/s00606-005-0346-0

**De Langhe, E.** (1961). La taxonomie du bananier plantain en Afrique Équatoriale. *Journal d'agriculture tropicale et de botanique appliquée***. 8**:417-449.

**De Langhe, E. and Valmayor, R.V.** (1979). French plantains in Southeast Asia. IBPGRNewsletter for South-east Asia: 3-4.

**De Langhe, E., Hřibová, E., Carpentier, S., Doležel, J. and Swennen, R.** (2010). Did backcrossing contribute to the origin of hybrid edible bananas? *Ann. Bot.* **106**:849-857. doi: 10.1093/aob/mcq187

**De Langhe, E., Perrier, X., Donohue, M. and Denham, T.** (2015). The original Banana Split: Multi-disciplinary implications of the generation of African and Pacific Plantains in Island Southeast Asia. *Ethnobot. Res. Appl.* **14**:299-312. doi: 10.17348/era.14.0.299-312

**De Langhe, E., Vrydaghs, L., De Maret, P., Perrier, X. and Denham, T.P.** (2009). Why bananas matter: An introduction to the history of banana domestication. *Ethnobot. Res. Appl.* **7**:165-177. doi: 10.17348/era.7.0.165-177

**Denham, T., Haberle, S. and Lentfer, C.** (2004). New evidence and revised interpretations of early agriculture in Highland New Guinea. *Antiquity.* **78**:839-857. doi: 10.1017/S0003598X00113481

**Dessauw, D.** (1987). Etude des facteurs de la ste´rilite´ du bananier (*Musa* spp) et des relations cytotaxinomiques entre *M. acuminata* Colla et *M. balbisiana* Colla. PhD thesis, University of Paris-Sud Centre d'Orsay.

**D'Hont, A.** (2005). Unraveling the genome structure of polyploids using FISH and GISH; examples of sugarcane and banana. *Cytogenet. Genome Res*. **109**:27-33. doi: 10.1159/000082378

**do Vale Martins, L., de Oliveira Bustamante, F., da Silva Oliveira, A.R., da Costa, A.F., de Lima Feitoza, L., Liang, Q., *et al*.** (2021). BAC- and oligo-FISH mapping reveals chromosome evolution among *Vigna angularis*, *V. unguiculata*, and *Phaseolus vulgaris*. *Chromosoma*. doi: 10.1007/s00412-021-00758-9

**do Vale Martins, L., Yu, F., Zhao, H., Dennison, T., Lauter, N., Wang, H., *et al.* (2019).** Meiotic crossovers characterized by haplotype-specific chromosome painting in maize. *Nat. Commun*. **10**:4604. doi: 10.1038/s41467-019-12646-z

**Dodds, K.S.** (1943). Genetical and cytological studies of *Musa*. V. Certain edible diploids. *J. Genet*. **45**:113-138. doi: 10.1007/BF02982931

**Doležel, J.** (1991). Flow cytometric analysis of nuclear DNA content in higher plants. *Phytochem. Anal*. **2**:143-154. doi: 10.1002/pca.2800020402

**Doležel, J., Doleželová, M. and Novák, F.J.** (1994). Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *M. balbisiana*). *Biol. Plant.* **36**:351-357. doi: 10.1007/BF02920930

**Doležel, J., Lysák, M.A., Doleželová, M. and Roux, N.** (1997). Use of flow cytometry for rapid ploidy determination in *Musa* species. *InfoMusa.* **6**:6-9

**Doleželová, M., Doležel, J., Van den Houwe, I., Roux, N. and Swennen, R.** (2005). Ploidy levels revealed. *InfoMusa*. **14**:34-36.

**Doleželová, M., Valárik, M., Swennen, R., Horry, J.P. and Doležel, J.** (1998). Physical mapping of the 18S-25S and 5S ribosomal RNA genes in diploid bananas. *Biol. Plant*. **41**:497-505. doi: 10.1023/A:1001880030275

**Donohue, M. and Denham, T.P.** (2009). Banana (*Musa* spp.) domestication in the Asia-Pacific region: Linguistic and archaeobotanical perspectives. *Ethnobot. Res. Appl.* **7**:293-332. doi: 10.17348/era.7.0.293-332

**Dupouy, M., Baurens, F.C., Derouault, P., Hervouet, C., Cardi, C., Cruaud, C., *et al.*** (2019). Two large reciprocal translocations characterized in the disease resistance-rich *burmannica* genetic group of *Musa acuminata*. *Ann. Bot.* **124**:319329. doi: 10.1093/aob/mcz078

**Ellis, T.H., Lee, D., Thomas, C.M., Simpson, P.R., Cleary, W.G., Newman, M.A., *et al.*** (1988). 5S rRNA genes in *Pisum*: Sequence, long range and chromosomal organization. *Mol. Gen. Genet*. **214**:333-342. doi: 10.1007/BF00337732

**Englberger L., Wills R.B.H., Blades B., Dufficy L., Daniells J.W. and Coyne T.** (2006). Carotenoid content and flesh colour of selected banana cultivars growing in Australia. *Food Nutr. Bull*. **27**:281291. doi: 10.1177/156482650602700401

**Englberger, L., Aalbersberg, W., Ravi, P., Bonnin, E., Marks, G.C., Maureen, H.F., Fitzgerald, *et al.*** (2003). Further analyses on Micronesian banana, taro, breadfruit and other foods for provitamin A carotenoids and minerals. *J. Food Comp. Anal.* **16**:219-236. doi: 10.1016/S0889-1575(02)00171-0

**Fahn, A., Stoler, S. and Fist, T.** (1963). Vegetative shoot apex in banana and zonal changes as it becomes reproductive. *Bot. Gaz.* **124**:246-250

**FAOSTAT,** Agriculture Organization of the United Nations, FAO. 2019. http://www.fao.org/home/en/. Accessed June 2021.

**Fauré, S., Noyer, J.L., Horry, J.P., Bakry, F., Lanaud, C. and de León, D.G.** (1993). A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). *Theor. Appl. Genet.* **87**:517-526. doi: 10.1007/BF00215098

**Ferguson-Smith, M.A. and Trifonov, V.** (2007). Mammalian karyotype evolution. *Nat. Rev. Genet*. **8**:950-962. doi: 10.1038/nrg2199

**Filiault, D.L., Ballerini, E.S., Mandáková, T., Aköz, G., Derieg, N.J., Schmutz, J., *et al.*** (2018). The *Aquilegia* genome provides insight into adaptive radiation and

reveals an extraordinarily polymorphic chromosome with a unique history. *Elife*. **7**:e36426. doi: 10.7554/eLife.36426

**Fox, M.H. and Galbraith, D.W**. (1990). Application of flow cytometry and sorting to higher plant systems. - In: Melamed, M.R., Lindmo, T., Mendelsohn, M.L. (ed.): Flow Cytometry and Sorting. 2nd Edition, pp. 633-650. Wiley-Liss, New York.

**Franchet, A.R.** (1889). *Musa lasiocarpa* in Morot. *J. Bot*. **3**:329

**Fransz, P., De Jong, J.H., Lysák, M., Castiglione, M.R. and Schubert, I**. (2002). Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. *Proc. Natl. Acad. Sci. USA*. **99**:14584-14589. doi: 10.1073/pnas.212325299

**Fuchs, J., Houben, A., Brandes, A. and Schubert, I.** (1996a). Chromosome 'painting' in plants – a feasible technique? *Chromosoma*. **104**:315-320. doi: 10.1007/BF00337219

**Fuchs, J., Kloos, D.U., Ganal, M.W. and Schubert, I.** (1996b). *In situ* localization of yeast artificial chromosome sequences on tomato and potato metaphase chromosomes. *Chromosome Res.* **4**: 277-281. doi: 10.1007/BF02263677

**Galbraith, D.W., Doležel, J., Lambert, G. and Macas, J.** (1998). Nuclear DNA content and ploidy analyses in higher plants. In: Robinson, J.P. (ed) Current protocols in cytometry. Wiley, New York, pp 7.6.1-7.6.22

**Gayral, P., Blondin, L., Guidolin, O., Carreel, F., Hippolyte, I., Perrier, X., *et al*.** (2010). Evolution of endogenous sequences of banana streak virus: what can we learn from banana (*Musa* sp.) evolution? *J. Virol*. **84**:7346-7359. doi: 10.1128/JVI.00401-10

**Gayral, P., Noa, C., Lescot, M., Lheureux, F., Lockhart, B.E.L., Matsumoto, T., *et al*.** (2008). A single Banana streak virus integration event in the banana genome as the origin of infectious endogenous pararetrovirus. *J. Virol*. **82**:6697–6710. doi: 10.1128/JVI.00212-08

**Gerlach, W.L. and Bedbrook, J.R.** (1979). Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res.* **7**:1869-1885. doi: 10.1093/nar/7.7.1869

**Gerlach, W.L. and Dyer, T.A.** (1980). Sequence organization of the repeating units in the nucleus of wheat which contain 5S rRNA genes. *Nucleic Acids Res.* **8**:4851-4865. doi: 10.1093/nar/8.21.4851

**Gill, B.S. and Friebe, B.** (1998). Plant cytogenetics at the dawn of the 21st century. *Curr. Opin. Plant. Biol*. **1**:109-115. doi: 10.1016/s1369-5266(98)80011-3

**Gill, B.S. and Kimber, G.** (1977). Recognition of translocations and alien chromosome transfers in wheat by the giemsa C-banding technique. *Crop Sci*. **17**:264-266. doi**:** 10.2135/cropsci1977.0011183X001700020008x

**Gill, B.S., Friebe, B. and Endo, T.R.** (1991). Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome*. **34**:830-839. doi: 10.1139/g91-128

**Greilhuber, J.** (1977). Why plant chromosomes do not show G-bands. *Theor. Appl. Genet*. **50**:121-124. doi: 10.1007/BF00276805

**Habermann, F.A., Cremer, M., Walter, J., Kreth, G., von Hase, J., Bauer, K., *et al.*** (2001). Arrangements of macro- and microchromosomes in chicken cells. *Chromosome Res*. **9**:569-584. doi: 10.1023/a:1012447318535

**Häkkinen, M.** (2013). Reappraisal of sectional taxonomy in *Musa* (*Musaceae*). *Taxon*. **62**:809-813. doi: 10.12705/624.3

**Häkkinen, M. and Sharrock, S.** (2002). Diversity in the genus *Musa* - Focus on Rhodochlamys. In: INIBAP annual report 2001: Monpellier, France, p. 16-23.

**Han, Y., Zhang, T., Thammapichai, P., Weng, Y. and Jiang, J.** (2015). Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. *Genetics*. **200**:771-779. doi: 10.1534/genetics.115.177642

**Hanson, R.E., Zwick, M.S., Choi, S., Islam-Faridi, M.N., McKnight, T.D., Wing, R.A., *et al*. (1995)** Fluorescent *in situ* hybridization of a bacterial artificial chromosome. *Genome*. **38**:646-651. doi: 10.1139/g95-082

**Hasterok, R., Jenkins, G., Langdon, T., Jones, R.N. and Maluszynska, J.** (2001). Ribosomal DNA is an effective marker of *Brassica* chromosomes. *Theor. Appl. Genet*. 103:486-490. doi: 7/s001220100653

**He, L., Braz, G.T., Torres, G.A. and Jiang, J.M.** (2018). Chromosome painting in meiosis reveals pairing of specific chromosomes in polyploid *Solanum* species. *Chromosoma*. **127**:505-513. doi: 10.1007/s00412-018-0682-9

**He, L., Zhao, H., He, J., Yang, Z., Guan, B., Chen, K., *et al*.** (2020). Extraordinarily conserved chromosomal synteny of *Citrus* species revealed by chromosome - specific painting. *Plant J*. **103**:2225-2235. doi: 10.1111/tpj.14894

**Heslop-Harrison, J.S. and Schwarzacher, T.** (2007). Domestication, genomics and the future for banana. *Ann. Bot*. **100**:1073-1084. doi: 10.1093/aob/mcm191

**Hippolyte, I., Bakry, F., Seguin, M., Gardes, L., Rivallan, R., Risterucci, A.M., *et al*.** (2010). A saturated SSR/DArT linkage map of *Musa acuminata* addressing genome rearrangements among bananas. *BMC Plant Biol.* **10**:65. doi: 10.1186/1471-2229-10-65

**Hippolyte, I., Jenny, C., Gardes, L., Bakry, F., Rivallan, R., Pomies, V., *et al*.** (2012). Foundation characteristics of edible *Musa* triploids revealed from allelic distribution of SSR markers. *Ann. Bot.* **109**:937-951. doi: 10.1093/aob/mcs010

**Horaninow, P.F.** (1862). Prodromus Monographiae Scitaminarum Academiae Caesareae Scientiarum, Petropoli

**Horrocks, M. and Rechtman, R.B.** (2009). Sweet potato (*Ipomoea batatas*) and banana (*Musa* sp.) microfossils in deposits from the Kona Field System, Island of Hawaii. *J. Archaeol. Sci*. **36**:1115-1126. doi: 10.1016/j.jas.2008.12.014

**Horrocks, M., Bedford, S. and Spriggs, M.** (2009). A short note on banana (*Musa*) phytoliths in Lapita, immediately post-Lapita and modern period archaeological deposits from Vanuatu. *J. Archaeol. Sci*. **36:**2048-2054. doi: 10.1016/j.jas.2009.05.024

**Horry, J.P.** (1989). Chimiotaxonomie et organisation génétique dans le genre *Musa* (III). *Fruits.* **44**:573-578.

**Hou, L., Xu, M., Zhang, T., Xu, Z., Wang, W., Zhang, J., *et al*.** (2018). Chromosome painting and its applications in cultivated and wild rice. *BMC Plant Biol*. **18**:110. doi: 10.1186/s12870-018-1325-2

**Houben, A., Kynast, R.G., Heim, U., Hermann, H., Jones, R.N. and Forster, J.W.** (1996). Molecular cytogenetic characterization of the terminal heterochromatic segment of the B-chromosome of rye (*Secale cereale*). *Chromosoma*. **105**:97-103. doi: 10.1007/BF02509519

**Hřibová, E., Čížková, J., Christelová, P., Taudien, S., De Langhe, E. and Doležel J.** (2011). The ITS-5.8S-ITS2 sequence region in the Musaceae: structure, diversity and use in molecular phylogeny. *PLoS One.* **6**:e17863. doi: 10.1371/journal.pone.0017863

**Hřibová, E., Doleželová, M. and Doležel, J.** (2008). Localization of BAC clones on mitotic chromosomes of *Musa acuminata* using fluorescence *in situ* hybridization. *Biol. Plant.* **52**:445-452. doi: 10.1007/s10535-008-0089-1

**Hřibová, E., Doleželová, M., Town, C.D., Macas, J. and Doležel, J.** (2007). Isolation and characterization of the highly repeated fraction of the banana genome. *Cytogenet. Genome Res*. **119**:268-274. doi: 10.1159/000112073

**Hřibová, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J. and Doležel, J.** (2010). Repetitive part of the banana (*Musa acuminata*) genome investigated by low-depth 454 sequencing. *BMC Plant Biol*. **10**:204. doi: 10.1186/1471-2229-10-204

**http://banana-genome.cirad.fr**

**https://github.com/forrestzhang/Chorus**

**Idziak, D., Hazuka, I., Poliwczak, B., Wiszynska, A., Wolny, E. and Hasterok, R.** (2014). Insight into the karyotype evolution of *Brachypodium* species using comparative chromosome barcoding. *PLoS One*. **9**:e93503. doi: 10.1371/journal.pone.0093503

**International Plant Genetic Resources Institute-International Network for the Improvement of Banana and Plantain/Centre de Coopération internationale en recherche agronomique pour le développement [IPGRI-INIBAP/CIRAD].** (1996). Description for Banana (*Musa* spp.). Int. Network for the Improvement of Banana and Plantain, Montpellier, France; Centre de coopération int. en recherche agronomique pour le développement, Montpellier, France; International Plant Genetic Resources Institute Press, Rome

**International trade statistics**. 2019. https://www.wto.org/ Accessed May 2020.

**Jacob, K.C.** (1952). Madras Bananas: A Monograph. Madras: Superintendent Government Press

**Janda J., Šafář J., Kubaláková M., Bartoš J., Kovářová P., Suchánková P., *et al.*** (2006). Advanced resources for plant genomics: a BAC library specific for the short arm of wheat chromosome 1B. *Plant J*. **47**:977-986. doi: 10.1111/j.1365-313X.2006.02840.x

**Janssens, S.B., Vandelook, F., De Langhe, E., Verstraete, B., Smets, E., Van den Houwe, I., *et al*.** (2016). Evolutionary dynamics and biogeography of Musaceae reveal a correlation between the diversification of the banana family and the geological and climatic history of Southeast Asia. *New Phytol*. **210**:1453-1465. doi:10.1111/nph.13856

**Jarret, R.L. Gawel, N., Whittemore, A. and Sharrock, S.** (1992). RFLP-based phylogeny of *Musa* species in Papua New Guinea. *Theoret. Appl. Genet*. **84**:579-584. doi: 10.1007/BF00224155

**Jáuregui, B., de Vicente, M.C., Messeguer, R., Felipe, A., Bonnet, A., Salesses, G., *et al*.** (2001). A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. *Theor Appl Genet*. **102**:1169-1176. doi: 10.1007/s001220000511

**Jeridi, M., Bakry, F., Escoute, J., Fondi, E., Carreel, F., Ferchichi, A., *et al*.** (2011). Homoeologous chromosome pairing between the A and B genomes of *Musa* spp. revealed by genomic *in situ* hybridization. *Ann. Bot.* **108**:975-981. doi: 10.1093/aob/mcr207

**Jiang, J.** (2019). Fluorescence *in situ* hybridization in plants: recent developments and future applications. *Chromosome Res.* **27**:153-165. doi: 10.1007/s10577-019-09607-z

**Jiang, J. and Gill, B.S.** (1994). Different species-specific chromosome translocations in *Triticum timopheevii* and *T. turgidum* support the diphyletic origin of polyploid wheats. *Chromosome Res*. **2**:59-64. doi: 10.1007/BF01539455

**Jiang, J., Gill, B.S., Wang, G.L., Ronald, P.C. and Ward, D.C.** (1995). Metaphase and interphase fluorescence in situ hybridization mapping of the rice genome with bacterial artificial chromosomes. *Proc. Natl. Acad. Sci. USA*. **92**:4487-4491. doi:10.1073/pnas.92.10.4487

**Jiang, J.M. and Gill, B.S.** (2006). Current status and the future of fluorescence *in situ* hybridization (FISH) in plant genome research. *Genome*. **49**:1057-1068. doi: 10.1139/g06-076

**John, H.A., Birnstiel, M.L. and Jones, K.W.** (1969). RNA-DNA hybrids at the cytological level. *Nature*. **223**:582-587. doi: 10.1038/223582a0

**Jones, D.R.** (2000). Diseases of banana, abaca and enset. Wallingford, UK: CABI Publishing.

**Kagy, V., Wong, M., Vandenbroucke, H., Jenny, C., Dubois, C., Ollivier, A., *et al.*** (2016). Traditional banana diversity in Oceania: An endangered heritage. *PLoS One*. **11**:e0151208. doi: 10.1371/journal.pone.0151208

**Karafiátová, M., Bartoš, J. and Doležel, J.** (2016). Localization of Low-Copy DNA Sequences on Mitotic Chromosomes by FISH. In Kianian, S., Kianian, P. (eds). Plant Cytogenetics. *Methods Mol. Biol*., vol 1429. Humana Press, New York, NY. doi: 10.1007/978-1-4939-3622-9_5

**Karamura, D., Karamura, E.B. and Blomme, G**. (2011). General Plant Morphology of *Musa*. In: Pillay, M., Tenkouano, A. (eds.) Banana Breeding Progress and Challenges. CRC Press, New York, 1-20. 1st edition. ISBN: 9780429074363

**Karamura, E., Frison, E., Karamura, D. and Sharrock, S.** (1998). Banana production systems in eastern and southern Africa. In: Picq, C., Foure, E., Frison, E.A., eds. *Bananas and food security*. Montpellier: INIBAP, 401-412.

**Kato, A., Lamb, J.C. and Birchler, J.A.** (2004). Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc. Natl. Acad. Sci. USA*. **101**:13554-13559. doi: 10.1073/pnas.0403659101

**Kervégant, D.** (1935). Le bananier et son exploita-tion. Soc. Edit. Marit. et Colon., Paris.

**Kim, J.S., Childs, K.L, Islam-Faridi, M.N., Menz, M.A., Klein, R.R., Klein, P.E., *et al*.** (2002). Integrated karyotyping of sorghum by *in situ* hybridization of landed BACs. *Genome*. **45**:402-412. doi: 10.1139/g01-141

**Kirch, P.V.** (1997). The Lapita peoples: Ancestors of the Oceanic World. Cambridge, Massachusetts: Blackwell Publishers. ISBN: 978-1-577-18036-4

**Kitavi, M., Downing, T., Lorenzen, J., Karamura, D., Onyango, M., Nyine, M., *et al.*** (2016). The triploid East African Highland Banana (EAHB) genepool is genetically uniform arising from a single ancestral clone that underwent population expansion by vegetative propagation. *Theoret. Appl. Genet.* **129**:547-561. doi: 10.1007/s00122-015-2647-1

**Kopecký, D., Lukaszewski, A.J. and Doležel, J.** (2008). Cytogenetics of Festulolium (*Festuca* × *Lolium* hybrids). *Cytogenet. Genome Res*. **120**:370-383. doi: 10.1159/000121086

**Kopecký, D., Šimoníková, D., Ghesquière, M. and Doležel, J.** (2017). Stability of genome composition and recombination between homoeologous chromosomes in Festulolium (*Festuca* × *Lolium*) cultivars. *Cytogenet. Genome Res*. **151**:106-114. doi: 10.1159/000458746

**Kress, W. J., Prince, L.M., Hahn, W.J., and Zimmer, E.A.** (2001). Unraveling the evolutionary radiation of the families of the Zingiberales using morphological and molecular evidence. *Syst. Biol*. **50**:926-944. doi: 10.1080/106351501753462885

**Kress, W.J. and Specht, C.D.** (2005). Between cancer and capricorn: phylogeny, evolution, and ecology of the tropical Zingiberales. Friis, I., Balslev, H. (Eds.), Plant Diversity and Complexity Patterns—Local, Regional and Global Dimensions, vol. 55, *Biologiske Skrifter*, The Royal Danish Academy of Sciences and Letters, Copenhagen, Denmark, pp. 459-478.

**Křivánková, A., Kopecký, D., Stočes, Š., Doležel, J. and Hřibová, E.** (2017). Repetitive DNA: A Versatile Tool for Karyotyping in *Festuca pratensis* Huds. *Cytogenet. Genome Re*s. **151**:96-105. doi: 10.1159/000462915

**Kubis, S., Schmidt, T., Seymour, J. and Heslop-Harrison, J.S.** (1998). Repetitive DNA Elements as a Major Component of Plant Genomes. *Ann. Bot*. **82**:45-55. doi: 10.1006/anbo.1998.0779

**Lamb, J.C., Danilova, T., Bauer, M.J., Meyer, J.M., Holland, J.J., Jensen, M.D., *et al*.** (2007). Single-Gene Detection and Karyotyping Using Small-Target Fluorescence *in Situ* Hybridization on Maize Somatic Chromosomes. *Genetics*. **175**:1047-1058. doi: 10.1534/genetics.106.065573

**Langer-Safer, P.R., Levine, M. and Ward, D.C.** (1982). Immunological method for mapping genes on *Drosophila* polytene chromosomes. *Proc. Natl. Acad. Sci. USA*. **79**:4381-4385. doi: 10.1073/pnas.79.14.4381

**Lapitan, N.L.V., Brown, S.E., Kennard, W., Stephens, J. and Knudson, D.** (1997). FISH physical mapping with barley BAC clones. *Plant J*. **11**:149-156. doi: 10.1046/j.1365-313X.1997.11010149.x

**Lebot, V., Aradhya, M.K., Manshardt, R.M. and Meilleur, B.A.** (1993). Genetic relationships among cultivated bananas and plantains from Asia and the Pacific. *Euphytica*. **67**:163-175. doi: 10.1007/bf00040618

**Leitch, I.J., Leitch, A.R. and Heslop-Harrison, J.S.** (1991) Physical mapping of plant DNA sequences by simultaneous *in situ* hybridization of two differently labeled fluorescent probes. *Genome*. **34**:329-333. doi: 10.1139/g91-054

**Lescot, T.** (2020). Banana genetic diversity: Estimated world production by type of banana. *FruiTrop*. **269**:98-102.

**Li, H.W.** (1978). The Musaceae of Yunnan. *Acta Phytotax. Sin*. **16**:54-64.

**Li, L.F., Häkkinen, M., Yuan, Y.M., Hao, G. and Ge, X.J.** (2010). Molecular phylogeny and systematics of the banana family (Musaceae) inferred from multiple nuclear and chloroplast DNA fragments, with a special reference to the genus *Musa*. *Mol. Phylogenet. Evol*. **57**:1-10. doi: 10.1016/j.ympev.2010.06.021

**Li, L.F., Wang, H.Y., Zhang, C., Wang, X.F., Shi, F.X., Chen, W.N., *et al*.** (2013). Origins and domestication of cultivated banana inferred from chloroplast and nuclear genes. *PLoS One*. **8**:e80502. doi: 10.1371/journal.pone.0080502

**Lichter, P., Cremer, T., Borden, J., Manuelidis, L. and Ward, D.C.** (1988). Delineation of individual human chromosomes in metaphase and interphase cells by *in situ* suppression hybridization using recombinant DNA libraries. *Hum. Genet.* **80**:224-234. doi: 10.1007/BF01790090

**Lim, K.B., Wennekes, J., de Jong, H.J., Jacobsen, E. and van Tuy, J.M.** (2001). Karyotype analysis of *Lilium longiflorum* and *Lilium rubellum* by chromosome banding and fluorescence *in situ* hybridisation. *Genome.* **44**:911-918. doi: 10.1139/gen-44-5-911

**Linnaeus, C.** (1753). *Musa* L. Species Plantarum, vol. 2, Impensis Laurentii Salvii, Stockholm, p. 1043.

**Lira-Medeiros, C.F., Parisod, C., Fernandes, R.A., Mata, C.S., Cardoso, M.A. and Ferreira, P.C.G.** (2010). Epigenetic Variation in Mangrove Plants Occurring in Contrasting Natural Environment. *PLoS ONE*. **5**:e10326. doi: 10.1371/journal.pone.0010326

**Liu, A.Z., Kress, W.J. and Li, D.Z.** (2002). Insect pollination of *Musella lasiocarpa* (Musaceae): A monotypic genus endemic to Yunnan, China. *Plant Syst. Evol*. **235**:135-146. doi: 10.1007/s00606-002-0200-6

**Liu, A.Z., Kress, W.J. and Li, D.Z.** (2010). Phylogenetic analyses of the banana family (Musaceae) based on nuclear ribosomal (ITS) and chloroplast (*trnL-F*) evidence. *Taxon*. **59**:20-28. doi: 10.1002/tax.591003

**Liu, A.Z., Kress, W.J. and Long, C.L.** (2003). The ethnobotany of *Musella lasiocarpa* (Musaceae), an endemic plant of southwest China. *Econ. Bot*. **57:**279-281. doi: 10.1663/0013-0001

**Liu, W., Rouse, M., Friebe, B., Jin, Y., Gill, B. and Pumphrey, M.O.** (2011). Discovery and molecular mapping of a new gene conferring resistance to stem rust, Sr53, derived from *Aegilops geniculata* and characterization of spontaneous translocation stocks with reduced alien chromatin. *Chromosome Res*. **19**:669-682. doi: 10.1007/s10577-011-9226-3

**Liu, X., Sun, S., Wu, Y., Zhou, Y., Gu, S., Yu, H., *et al*.** (2020a). Dual-color oligo-FISH can reveal chromosomal variations and evolution in *Oryza* species. *Plant J*. **101**:112-121. doi: 10.1111/tpj.14522

**Liu, Y., Wang, X., Wei, Y., Liu, Z., Lu, Q., Liu, F., *et al*.** (2020b). Chromosome Painting Based on Bulked Oligonucleotides in Cotton. *Front. Plant Sci*. **11**:802. doi: 10.3389/fpls.2020.00802

**Long, E.O. and Dawid, I.B.** (1980). Repeated genes in Eukaryotes. *Annu. Rev. Biochem*. **49**:727-764. doi: 10.1146/annurev.bi.49.070180.003455

**Lysák, M. A., Doleželová, M., Horry, J.P., Swennen, R. and Doležel, J.** (1999). Flow cytometric analysis of nuclear DNA content in *Musa*. *Theor. Appl. Genet*. **98**:1344-1350. doi: 10.1007/s001220051201

**Lysák, M.A., Berr, A., Pečinka, A., Schmidt, R., McBreen, K. and Schubert, I.** (2006). Mechanisms of chromosome number reduction in Arabidopsis thaliana and related Brassicaceae species. *Proc. Natl. Acad. Sci. USA.* **103**:5224-5229. doi: 10.1073/pnas.0510791103

**Lysák, M.A., Fransz, P.F., Ali, H.B.M. and Schubert, I.** (2001). Chromosome painting in *A. thaliana*. *Plant J*. **28**:689-697. doi: 10.1046/j.1365-313x.2001.01194.x

**Lysák, M.A., Koch, M.A., Pečinka, A. and Schubert, I.** (2005). Chromosome triplication found across the tribe Brassiceae. *Genome Res*. **15**:516-525. doi: 10.1101/gr.3531105

**Ma, H., Pan, Q., Wang, L., Li, Z., Wan, Y., and Liu, X.** (2011). *Musella lasiocarpa var. rubribracteata* (Musaceae), a New Variety from Sichuan, China. *Novon A Journal for Botanical Nomenclature*. **21**:349-353. doi: 10.2307/23018447

**Macas, J., Navrátilová, A. and Koblížková, A.** (2006). Sequence homogenization and chromosomal localization of VicTR-B satellites differ between closely related *Vicia* species. *Chromosoma*. **115**:437-447. doi: 10.1007/s00412-006-0070-8

**Macas, J., Požárková, D., Navrátilová, A., Nouzová, M. and Neumann, P.** (2000). Two new families of tandem repeats isolated from genus *Vicia* using genomic self-priming PCR. *Mol. Gen. Genet*. **263**:741-751. doi: 10.1007/s004380000245

**MacDaniels, L.H. (**1947). A study of the Fe'i banana and its distribution with reference to Polynesian migrations. *Bull. Bernice P. BishopMus.* 190:1-56.

**Macrae, R., Robinson, R.K. and Sadler, M.J.** (eds) (1993*).* Bananas and Plantains. Encyclopaedia of Food Science, Food Technology and Nutrition, *Academic Press*

**Maidin, S., and Latiff, A.N.** (2015). Nasi lemak packaking: A case study of food freshness and design flexibility. *J. Adv. Manuf. Technol*. **9**:13-19. ISSN 1985-3157.

**Mandáková, T. and Lysák, M. A.** (2008). Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). *Plant Cell*. **20**:2559-2570. doi: 10.1105/tpc.108.062166

**Mandáková, T., Joly, S., Krzywinski, M., Mummenhoff, K. and Lysák, M.A.** (2010). Fast diploidization in close mesopolyploid relatives of *Arabidopsis. Plant Cell*. **22**:2277-2290. doi: 10.1105/tpc.110.074526

**Martin, G., Baurens, F.C., Droc, G., Rouard, M., Cenci, A., Kilian, A., *et al*.** (2016). Improvement of the banana "*Musa acuminata*" reference sequence using NGS data and semi-automated bioinformatics methods. *BMC Genomics*. **17**:1-12. doi: 10.1186/s12864-016-2579-4

**Martin, G., Baurens, F.C., Hervouet, C., Salmon, F., Delos, J.M., Labadie, K., *et al*.** (2020b). Chromosome reciprocal translocations have accompanied subspecies evolution in bananas. *Plant J.* **104**:1698-1711. doi: 10.1111/tpj.15031

**Martin, G., Cardi, C., Sarah, G., Ricci, S., Jenny, C., Fondi, E., *et al*.** (2020a). Genome ancestry mosaics reveal multiple and cryptic contributors to cultivated banana. *Plant J*. **102**:1008-1025. doi: 10.1111/tpj.14683

**Martin, G., Carreel, F., Coriton, O., Hervouet, C., Cardi, C., Derouault, P., *et al.*** (2017). Evolution of the banana genome (*Musa acuminata*) is impacted by large chromosomal translocations. *Mol. Biol. Evol.* **34**:2140-2152. doi: 10.1093/molbev/msx164

**Matheka, J., Tripathi, J.N., Merga, I., Gebre, E. and Tripathi, L.** (2019). A simple and rapid protocol for the genetic transformation of *Ensete ventricosum*. *Plant Methods.* **15**:130. doi: 10.1186/s13007-019-0512-y

**Mbanjo, E.G.N, Tchoumbougnang, F., Mouelle, A.S., Oben, J.E., Nyine, M., Dochez, C., *et al.*** (2012). Molecular marker-based genetic linkage map of a diploid banana population (*Musa acuminata* Colla). *Euphytica.* **188**:369-386. doi: 10.1007/s10681-012-0693-1

**Meltzer, P.S., Guan, X.Y., Burgess, A. and Trent, J.M.** (1992). Rapid generation of region specific probes by chromosome microdissection and their application. *Nat. Genet*. **1**:24-28. doi: 10.1038/ng0492-24

**Meng, Z., Han, J., Lin, Y., Zhao, Y., Lin, Q., Ma, X. *et al*.** (2020). Characterization of a *Saccharum spontaneum* with a basic chromosome number of $x = 10$ provides

new insights on genome evolution in genus *Saccharum*. *Theoret. Appl. Genet.* **133**:187-199. doi: 10.1007/s00122-019-03450-w

**Meng, Z., Zhang, Z.L., Yan, T.Y., Lin, Q.F, Wang, Y., Huang W.Y., *et al.*** (2018). Comprehensively characterizing the cytological features of *Saccharum spontaneum* by the development of a complete set of chromosome-specific oligo probes. *Front. Plant. Sci.* **9**:1624. doi: 10.3389/fpls.2018.01624

**Moore, G., Gale, M.D., Kurata, N. and Flavell, R.B**. (1993). Molecular analysis of small grain cereal genomes: current status and prospects. *Bio/Technology*. **11**:584-589. doi: 10.1038/NBT0593-584

**Mukai, Y., Nakahara, Y. and Yamamoto, M.** (1993). Simultaneous discrimination of the three genomes in hexaploid wheat by multicolor fluorescence *in situ* hybridization using total genomic and highly repeated DNA probes. *Genome*. **36**:489-94. doi: 10.1139/g93-067

**Navrátilová, A., Neumann, P. and Macas, J**. (2003). Karyotype analysis of four *Vicia* species using *in situ* hybridization with repetitive sequences. *Ann. Bot.* **91**:921-926. doi: 10.1093/aob/mcg099

**Němečková, A., Christelová, P., Čížková, J., Nyine, M., Van den houwe, I., Svačina, R., *et al.*** (2018). Molecular and cytogenetic study of East African Highland Banana. *Front. Plant. Sci*. **9**:1371. doi: 10.3389/fpls.2018.01371

**Neumann, P., Navrátilová, A., Koblížková, A., Kejnovský, E., Hřibová, E., Hobza, R., *et al.*** (2011). Plant centromeric retrotransposoms: a structural and cytogenetic perspective. *Mob. DNA.* **2**:4. doi: 10.1186/1759-8753-2-4

**Noumbissié, G.B., Chabannes, M., Bakry, F., Ricci, S., Cardi, C., Njembele, J.C., *et al.*** (2016). Chromosome segregation in an allotetraploid banana hybrid (AAAB) suggests a translocation between the A and B genomes and results in eBSV-free offsprings. *Mol. Breed.* **36**:38-52. doi: 10.1007/s11032-016-0459-x

**Novák, P., Hřibová, E., Neumann, P., Koblížková, A., Doležel, J. and Macas, J.** (2014). Genome-wide analysis of repeat diversity across the family Musaceae. *PLoS One*. **9**:e98918. doi: 10.1371/journal.pone.0098918

**Nyine, M., Uwimana, B., Swennen, R., Batte, M., Brown, A., Christelová, P., *et al.*** (2017). Trait variation and genetic diversity in a banana genomic selection training population. *PLoS One.* **12**:e0178734. doi: 10.1371/journal.pone.0178734

**Olango, T.M., Tesfaye, B., Pagnotta, M.A., Pè, M.E. and Catellani, M.** (2015). Development of SSR markers and genetic diversity analysis in enset (*Ensete ventricosum* (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia. *BMC Genetics*. **16**:98. doi: 10.1186/s12863-015-0250-8

**Ong-Abdullah, M., Ordway, J.M., Jiang, N., Ooi, S.E., Kok, S.Y., Sarpan, N., *et al.*** (2015): Loss of Karma transposon methylation underlies the mantled somaclonal

variant of oil palm. *Nature*. **525**:533-537. doi: 10.1038/nature15365

**Ortiz, R.** (2013). Conventional banana and plantain breeding. *Acta Hortic.* **986**:177-194. doi: 10.17660/ActaHortic.2013.986.19

**Ortiz, R. and Swennen, R.** (2014). From crossbreeding to biotechnology-facilitated improvement of banana and plantain. *Biotechnol. Adv.* **32**:158-169. doi: 10.1016/j.biotechadv.2013.09.010

**Ortiz-Vazquez, E., Kaemmer, D., Zhang, H.B., Muth, J., Rodriguez-Mendiola, M., Arias-Castro, C., *et al.*** (2005). Construction and characterization of a plant transformation-competent BIBAC library of the black Sigatoka-resistant banana *Musa acuminata* cv. Tuu Gia (AA). *Theor. Appl. Genet*. **110**:706-713. doi: 10.1007/s00122-004-1896-1

**Ostberg, C.O., Hauser, L., Pritchard, V.L., Garza, J.C. and Naish, K.A.** (2013). Chromosome rearrangements, recombination suppression, and limited segregation distortion in hybrids between Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and rainbow trout (*O. mykiss*). *BMC Genomics*. **14**:570. doi: 10.1186/1471-2164-14-570

**Osuji, J., Harrisson, G., Crouch, J. and Heslop-Harrison, J.S.** (1997). Identification of the genomic constitution of *Musa* L. lines (Bananas, Plantains and hybrids) using molecular cytogenetics. *Ann. Bot*. **80**:787-793. doi: 10.1006/anbo.1997.0516

**Osuji, J.O., Crouch, J., Harrison, G. and Heslop-Harrison, J.S.** (1998). Molecular cytogenetics of *Musa* species, cultivars and hybrids: Location of 18S-5.8S-25S and 5S rDNA and telomere-like sequences. *Ann. Bot*. **82**:243-248.

**Pan, Q. J., Li, Z.H., Wang, Y., Tian, J., Gu, Y. and Liu, X. X.** (2007). RADP analysis on the genetic diversity of wild and cultivated populations of *Musella lasiocarpa*. *Forest Res.* **20**:668-672.

**Pardue, M.L. and Gall, J.G.** (1969). Molecular hybridization of radioactive DNA to the DNA of cytological preparations. *Proc. Natl. Acad. Sci. USA*. **64**:600-604. doi: 10.1073/pnas.64.2.600

**Pardue, M.L. and Gall, J.G.** (1970). Chromosomal localization of mouse satellite DNA. *Science*. **168**:1356-1358. doi: 10.1126/science.168.3937.1356

**Pečinka, A., Schubert, V., Meister, A., Kreth, G., Klatte, M., Lysák, M. A., *et al.*** (2004). Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. *Chromosoma*. **113**:258-269. doi: 10.1007/s00412-004-0316-2

**Pedersen, C. and Langridge, P.** (1997). Identification of the entire chromosome complement of bread wheat by two-colour FISH. *Genome*. **40**:589-93. doi: 10.1139/g97-077

**Pedersen, C., Rasmussen, S.K. and Linde-Laursen, I.** (1996). Genome and chromosome identification in cultivated barley and related species of the Triticeae (Poaceae) by *in situ* hybridization with the GAA-satellite sequence. *Genome.* **39**:93-104. doi: 10.1139/g96-013

**Perrier, X., Bakry, F., Carreel, F., Jenny, F., Horry, J.P., Lebot, V., *et al.*** (2009). Combining biological approaches to shed light on the evolution of edible bananas. *Ethnobot. Res. Appl.* **7**:199-216. doi: 10.17348/era.7.0.199-216

**Perrier, X., De Langhe, E., Donohue, M., Lentfer, C., Vrydaghs, L., Bakry, F., *et al.*** (2011). Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc. Natl. Acad. Sci. USA.* **108**:11311-11318. doi: 10.1073/pnas.1102001108

**Perrier, X., Jenny, C., Bakry, F., Karamura, D., Kitavi, M., Dubois, C., *et al.*** (2019). East African diploid and triploid bananas: a genetic complex transported from South-East Asia. *Ann. Bot.* **123**:19-36. doi: 10.1093/aob/mcy156

**Pinkel, D., Landegent, J., Collins, C., Fuscoe, J., Segraves, R., Lucas, J., *et al.*** (1988). Fluorescence *in situ* hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocations of chromosome 4. *Proc. Natl. Acad. Sci. USA.* **85**:9138-9142. doi: 10.1073/pnas.85.23.9138

**Ploetz, R.C.** (2015). Fusarium wilt of banana. *Phytopathology*. **105**:1512-1521.

**Price, N.S.** (1995). The origin and development of banana and plantain cultivation. In: Gowen, S., eds. Bananas and Plantains. World Crop Series, Dordrecht: Springer.

**Purseglove, J.W.** (1972). Tropical crops: Monocotyledons. Vol. 2. London: Longmans

**Qu, M., Li, K., Han, Y., Chen, L., Li, Z. and Han, Y.** (2017). Integrated karyotyping of woodland strawberry (*Fragaria vesca*) with oligopaint FISH probes. *Cytogenet. Genome Res*. **153**:158-164. doi: 10.1159/000485283

**Quénéhervé, P., Valette, C., Topart, P., Tezenas du Montcel, H. and Salmon, F.** (2009). Nematode resistance in bananas: screening results on some wild and cultivated accessions of *Musa* spp. *Euphytica.* **165**:123-136. doi: 10.1007/s10681-008-9773-7

**Raboin, L.M., Carreel, F., Noyer, J.L., Baurens, F.C., Horry, J.P., Bakry, F., *et al.*** (2005). Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete donors and n gamete donors. *Mol. Breed*. **16**:333-341. doi: 10.1007/s11032-005-2452-7

**Rayburn, A.L. and Gill, B.S.** (1985). Use of biotin-labeled probes to map specific DNA sequences on wheat chromosomes. *J. Hered*. **76**:78-81. doi: 10.1093/oxfordjournals.jhered.a110049

**Rieseberg, L.H.** (2001). Chromosomal rearrangements and speciation. *Trends Ecol. Evol*. **16**:351-358. doi: 10.1016/S0169-5347(01)02187-5

**Risterucci, A.M., Hippolyte, I., Perrier, X., Xia, L., Caig, V., Evers, M. *et al*.** (2009). Development and assesment of Diversity Array technology for high-throughput DNA analysis in *Musa*. *Theor. Appl. Genet*. **119**:1093–1103. doi: 10.1007/s00122-009-1111-5

**Robinson J.C.** (1996). Bananas and plantains. UK University Press, Cambridge.

**Robinson J.C. and Saúco V.G.** (2010). Bananas and Plantains. Crop Production Science in Horticulture, 19. 2nd edition. ISBN-13: 978-1845936587

**Rosales, F., Arnaud, E. and Coto, J.e.** (1999). A catalogue of wild and cultivated bananas. Atribute to the work of Paul Allen. Montpellier: INIBAP.

**Rouard, M., Droc, G., Martin, G., Sardos, J., Hueber, Y., Guignon, V., *et al.*** (2018). Three new genome assemblies support a rapid radiation in *Musa acuminata* (wild banana). *Genome Biol. Evol.* **10**:3129-3140. doi: 10.1093/gbe/evy227

**Roux, N., Toloza, A., Radecki, Z., Zapata-Arias, F.J. and Doležel, J.** (2003). Rapid detection of aneuploidy in *Musa* using flow cytometry. *Plant Cell Rep.* **21**:483-490. doi: 10.1007/s00299-002-0512-6

**Šafář, J., Bartoš, J., Janda, J., Bellec, A., Kubaláková, M., Valárik, M., *et al*.** (2004). Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J*. **39:**960-968. doi: 10.1111/j.1365-313X.2004.02179.x

**Said, M., Hřibová, E., Danilova, T.V., Karafiátová, M., Čížková, J., Friebe, B., *et al*.** (2018). The *Agropyron cristatum* karyotype, chromosome structure and cross-genome homoeology as revealed by fluorescence *in situ* hybridization with tandem repeats and wheat single-gene probes. *Theor. Appl. Genet*. **131**:2213-2227. doi: 10.1007/s00122-018-3148-9

**Sand, C.** (1989). Petite histoire du peuplement de l'Océanie. In: Ricard, M., ed. Migrations et identités—Colloque CORAIL, Nouméa (PYF), 1988/11/21-22. Papeete: Université Française du Pacifique; 39-40. Available: http://www.documentation.ird.fr/hor/fdi:27803.

**Sandoval, J.A., Côte, F.X. and Escoute, J.** (1996). Chromosome number variations in micropropagated true-to-type and off-type banana plants (*Musa* AAA Grande Naine cv.). *In Vitro Cell. Dev. Biol. Plant.* **32**:14-17. doi: 10.1007/BF02823007

**Sardos, J., Rouard, M., Hueber, Y., Cenci, A., Hyma, K.E., van den Houwe, I., *et al*.** (2016). A Genome-Wide Association Study on the Seedless Phenotype in Banana (*Musa* spp.) Reveals the Potential of a Selected Panel to Detect Candidate Genes in a Vegetatively Propagated Crop. *PLoS ONE*. **11**:e0154448. doi: 10.1371/journal.pone.0154448

**Schmidt, T. and Heslop-Harrison, J.S.** (1998). Genomes, genes and junk: the large-scale organization of plant chromosomes. *Trends Plant Sci*. **3**:195-199. doi: 10.1016/S1360-1385(98)01223-0

**Schubert, I., Fransz, P.F., Fuchs, J. and De Jong, J.H.** (2001). Chromosome painting in plants. *Methods Cell Sci*. **23**:57-69. doi : 10.1023/A:1013137415093

**Schwarzacher, T., Leitch, A.R., Bennett, M.D. and Heslop-Harrison, J.S.** (1989). *In situ* localization of parental genomes in a wide hybrid. *Ann. Bot*. **64**:315-324. doi: 10.1093/oxfordjournals.aob.a087847

**Selvakumar, S. and Parasurama, D.S.** (2020). Maximization of micropropagule production in banana cultivars Grand naine (AAA) and Elakki (AB). *In Vitro Cell. Dev. Biol. Plant*. **56**:515-525. doi: 10.1007/s11627-020-10060-5

**Sharma, S. and Raina, S.N.** (2005). Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. *Cytogenet. Genome Res*. **109**:15-26. doi: 10.1159/000082377

**Sharrock, S.** (1990). Collecting *Musa* in Papua New Guinea. Identification of genetic diversity in the genus *Musa*. In: Jarret, R. L., ed. International Network for the Improvement of Banana and Plantain, Montpellier, France, 140-157.

**Shepherd, K.** (1957). Banana cultivars in East Africa. Trop. Agric. 34:277-286.

**Shepherd, K.** (1999). Cytogenetics of the genus *Musa*. International Network for the Improvement of Banana and Plantain, Montpellier, France, ISBN: 2-910810-25-9.

**Shepherd, K. and Da Silva, K.M.** (1996). Mitotic instability in banana varieties. Aberrations in conventional triploid plants. *Fruit*s. **51**:99-103.

**Shibata, F., Hizume, M. and Kuroki, Y.** (1999). Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequence in the dioecious plant *Rumex acetosa*. *Chromosoma*. **108**:266-270. doi: 10.1007/s004120050377

**Silva, G.S. and Souza, M.M.** (2013). Genomic *in situ* hybridization in plants. *Genet. Mol. Res*. **3**:2953-2965. doi: 10.4238/2013.August.12.11

**Simmonds, N.W.** (1956). Botanical results of the banana collecting expeditions, 1954-5. Kew Bull. **11**:463-489. doi: 10.2307/4109131

**Simmonds, N.W.** (1960). Notes on banana taxonomy. Kew Bull., **14**:198-212.

**Simmonds, N.W.** (1962). The evolution of the bananas. Tropical Science Series. Longmans, London (GBR), 170.

**Simmonds, N.W.** (1966). Bananas. Longman, London, 512 pp.

**Simmonds, N.W. and Shepherd, K.** (1955). The taxonomy and origins of the cultivated bananas. *J. Linn. Soc., Bot*. **55**:302-312.

**Šimoníková, D., Hřibová, E., Čížková, J. and Doležel, J.** Molecular cytogenetics of *Musa*. Unpublished.

**Song, X., Song, R., Zhou, J., Yan, W., Zhang, T., Sun, H., *et al*.** (2020) Development and application of oligonucleotide-based chromosome painting for chromosome 4D of *Triticum aestivum* L. *Chromosome Res.* **28**:171-182. doi: 10.1007/s10577-020-09627-0

**Song, Y.C., Liu, L.H., Ding, Y., Tian, X.B., Yao, Q., Meng, L., *et al*.** (1994). Comparisons of G-banding patterns in six species of the Poaceae. *Hereditas*. **121**:31-38. doi: 10.1111/j.1601-5223.1994.00031.x

**Sotto, R.C. and Rabara, R.C.** (2000). Morphological diversity of *Musa balbisiana* Colla in the Philippines. *InfoMusa*. **9**:28-30.

**Speicher, M.R., Ballard, S.G. and Ward, D.C.** (1996). Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat. Genet*. **12**:368-375. doi: 10.1038/ng0496-368

**Stover, R. H. and Simmonds, N.W.** (1987). Bananas, 3rd edn. London: Longman.

**Stover, R.H.** (1962). Fusarial wilt (Panama disease) of bananas and other *Musa* species. Commonwealth Mycological Inst., Kew, UK

**Subbaraya, U.** (2006). Potential and constraints of using wild *Musa*. Farmers' knowledge of wild *Musa* in India, Food and Agriculture Organization of the United Nations, 33-36.

**Swangpol, S., Volkaert, H., Sotto, R. and Seelanan T.** (2007). Utility of selected non-coding chloroplast DNA sequences for lineage assessment of *Musa* interspecific hybrids. *J. Biochem. Mol. Biol*. **40**:577-587. doi: 10.5483/BMBRep.2007.40.4.577

**Swennen, R.** (1990). Plantain cultivation under West African conditions. A reference manual. Intl. Inst. Trop. Agr., Ibadan, Nigeria

**Tenkouano, A., Vuylsteke, D., Agogbua, J.U., Makumbi, D., Swennen, R. and Ortiz, R.** (2003). Diploid banana hybrids TMB2x5105-1 and TMB2x9128-3 with good combining ability, resistance to black sigatoga and nematodes. *HortScience.* **38**:468-472. doi: 10.21273/HORTSCI.38.3.468

**The *Arabidopsis* Genome Initiative.** (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. **408**:796-815. doi: 10.1038/35048692

**The International *Brachypodium* Initiative.** (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*. **463**:763-768. doi: 10.1038/nature08747

**Therdsak, T., Benchamas, S., Yingyong, P. and Pradit, P.** (2010). Meiotic behavior in microsporocytes of some bananas in Thailand. Kasetsart Journal (Natural Science) **44**:536-543.

**Tobiaw, D.C. and Bekele, E.** (2011). Analysis of genetic diversity among cultivated enset (*Ensete ventricosum*) populations from Essera and Kefficho, southwestern part of Ethiopia using inter simple sequence repeats (ISSRs) marker. *Afr. J Biotechnol*. **10**:15697-15709. doi: 10.5897/AJB11.885

**Tomepke, K., Jenny, C. and Escalant, J.V.** (2004). A review of conventional improvement strategies for *Musa*. *InfoMusa*. **13**:2-6.

**Tsujimoto, H., Mukai, Y., Akagawa, K., Nagaki, K., Fujigaki, J., Yamamoto, M., *et al*.** (1997). Identification of individual barley chromosomes based on repetitive sequences: Conservative distribution of Afa-family repetitive sequences on the chromosomes of barley and wheat. *Gene. Genet. Syst*. **72**:3003-3009. doi: 10.1266/ggs.72.303

**Ude, G., Pillay, M., Nwakanma, D. and Tenkouano, A.** (2002). Genetic Diversity in *Musa acuminata* Colla and *Musa balbisiana* Colla and some of their natural hybrids using AFLP Markers. *Theor. Appl. Genet*. **104**:1246-1252. doi: 10.1007/s00122-002-0914-4

**Umber, M., Pichaut, J.P, Farinas, B., Laboureau, N., Janzac, B., Plaisir-Pineau, K., *et al*.** (2016). Marker-assisted breeding of *Musa balbisiana* genitors devoid of infectious endogenous Banana streak virus sequences. *Mol. Breed*. **36**:74. doi: 10.1007/s11032-016-0493-8

**Valárik, M., Šimková, H., Hřibová, E., Šafář, J., Doleželová, M. and Doležel, J.** (2002). Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (*Musa* spp.). *Chromosome Res*. **10**:89-100. doi: 10.1023/A:1014945730035

**Van Asten, P., Fermont, A.M. and Taulya, G.** (2011). Drought is a major yield loss factor for rainfed East African highland banana. *Agric. Water Manag.* **98**:541-552. doi: 10.1016/j.agwat.2010.10.005

**van Wesemael, J., Kissel, E., Eyland, D., Lawson, T., Swennen, R. and Carpentier, S.** (2019). Using growth and transpiration phenotyping under controlled con-

ditions to select water efficient banana genotypes. *Front. Plant Sci.* **10**: 352. doi: 10.3389/fpls.2019.00352

**Vilarinhos, A., Carrel, F., Rodier, M., Hippolyte, I., Benabdelmouna, A., Triaire, D., *et al*.** (2006). Characterization of translocations in banana by FISH of BAC clones anchored to a genetic map. - In: Abstracts of the International Conference "Plant and Animal Genome XIV". P. 8. Sherago International, San Diego.

**Vilarinhos, A.D., Piffanelli, P., Lagoda, P., Thibivilliers, S., Sabau, X., Carreel, F., *et al*.** (2003). Construction and characterization of a bacterial artificial chromosome library of banana (*Musa acuminata* Colla). *Theor. Appl. Genet*. **106**:1102-1106. doi: 10.1007/s00122-002-1155-2

**Vuylsteke, D., Swennen, R., and De Langhe, E.** (1991). Somaclonal variation in plantains (*Musa* spp., AAB group) derived from shoot-tip culture. *Fruits.* **46**:429-439.

**Vuylsteke, D., Swennen, R., Wilson, G.F. and De Langhe, E.** (1988). Phenotypic variation among *in vitro* propagated plantain (*Musa* spp. cv. AAB). *Scientia Hort*. **36**:79-88. doi: 0.1016/0304-4238(88)90009-X

**Vuylsteke, D.R., Swennen, R.L. and Ortiz, R.** (1993). Development and performance of black sigatoka-resistant tetraploid hybrids of plantain (*Musa* spp., AAB group). *Euphytica.* **65**:33-42. doi: 10.1007/BF00022197

**Wang, Z., Miao, H., Liu, J., Xu, B., Yao, X., Xu, C., *et al*.** (2019). *Musa balbisiana* genome reveals subgenome evolution and functional divergence. *Nat. Plants*. **5**:810-821. doi: 10.1038/s41477-019-0452-6

**WCSP**, World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. 2018. http://wcsp.science.kew.org/. Accessed January 2020.

**Wilson, G.B.** (1945). Cytological studies in the *Musae*; meiosis in some triploid clones. *Genetics*. **31**:241-258.

**Wong, C., Kiew, R., Argent, G.C.G, Set, O., Lee, S. K. and Gan, Y.Y.** (2002). Assessment of the validity of the sections in *Musa* (Musaceae) using AFLP. *Ann. Bot*. **90**:231-238. doi: 10.1093/aob/mcf170

**Wu, T.L. and Kress, W.J.** (2001). Musaceae. In: Wu, C.Y., Raven, P.H. (eds.) Flora of China. Science Press, Beijing, **24**:314-318.

**Xin, H., Zhang, T., Han, Y., Wu, Y., Shi, J., Xi, M., *et al*.** (2018). Chromosome painting and comparative physical mapping of the sex chromosomes in *Populus tomentosa* and *Populus deltoides*. *Chromosoma*. **127**:313-321. doi: 10.1007/s00412-018-0664-y

**Xin, H., Zhang, T., Wu, Y., Zhang, W., Zhang, P., Xi, M., *et al.*** (2020). An extraordinarily stable karyotype of the woody *Populus* species revealed by chromosome painting. *Plant J*. **101**:253-264. doi: 10.1111/tpj.14536

**Yakura, K. and Tanifuji, S.** (1983). Molecular Cloning and Restriction Analysis of *Eco*RI-fragments of *Vicia faba* rDNA- *Plant Cell Physiol.* **24**:1327-1330. doi: 10.1093/oxfordjournals.pcp.a076650

**Yamada, N.A., Rector, L.S., Tsang, P., Carr, E., Scheffer, A., Sederberg, M.C., *et al.*** (2011). Visualization of fine-scale genomic structure by oligonucleotide-based high-resolution FISH. *Cytogenet. Genome Res*. **132**:248-254. doi: 10.1159/000322717

**Yoo, M.J., Szadkowski, E. and Wendel, J.F.** (2013). Homoeolog expression bias and expression level dominance in allopolyploid cotton. *Heredity*. **110**:171-80. doi: 10.1038/hdy.2012.94

**Zatloukalová, P., Hřibová, E., Kubaláková, M., Suchánková, P., Šimková, H., Adoración, C., *et al.*** (2011). Integration of genetic and physical maps of the chickpea (*Cicer arietinum* L.) genome using flow-sorted chromosomes. *Chromosome Res*. **19**:729. doi: 10.1007/s10577-011-9235-2

**Zhang, P., Li, W., Friebe, B. and Gill, B.S.** (2004). Simultaneous painting of three genomes in hexaploid wheat by BAC-FISH. *Genome*. **47**:979-987. doi: 10.1139/g04-042

**Zhao, X.P., Si, Y., Hanson, R.E., Crane, C.F., Price, H.J., Stelly, D.M., *et al.*** (1998). Dispersed repetitive DNA has spread to new genomes since polyploid formation in cotton. *Genome Res*. **8**:479-92. doi: 10.1101/gr.8.5.479. Erratum in: *Genome Res.* (1998) **8**:682.

# 3 AIMS OF THE THESIS

The aim of the Ph.D. thesis was to establish chromosome-specific painting in banana to enable unambiguous identification of chromosomes, enlarge the knowledge about chromosome structural changes in bananas, and perform comparative karyotype analysis within species of the family Musaceae.

The work has started in 2017, when a comprehensive knowledge on chromosome structure and genome rearrangements in *Musa* spp. was lacking. The presence of putative chromosomal translocations has been assumed only after the analysis of chromosome pairing during meiosis in inter-subspecific hybrids of *M. acuminata* by Shepherd (1999), who proposed several translocation groups in *Musa*. Nevertheless, chromosomes included in these translocation events were not known, except of a reciprocal translocation between chromosomes 1 and 4 in *M. acuminata* ssp. *malaccensis* (Martin *et al*., 2017).

However, at the same time, French researchers from Cirad (The French Agricultural Research Centre for International Development) have analysed the genome structure of selected *Musa* species, subspecies and edible cultivars by application of Illumina sequencing approaches (Martin *et al*., 2017; Belser *et al*., 2018; Baurens *et al*., 2019; Dupouy *et al*., 2019; Martin *et al*., 2020a; 2020b), which are mentioned in more detail in the Literature overview.

The aims of the thesis contained four main parts:

**I. Development of oligonucleotide-based chromosome painting FISH in banana (*Musa* spp.)**

The first aim of the thesis was to establish a set of chromosome/chromosome-arm specific oligo painting probes to identify all chromosomes in banana and to anchor pseudomolecules of reference genome sequence of *Musa acuminata* spp. *malaccensis* 'DH Pahang' to individual chromosomes *in situ*, and to create integrated karyotyping in *Musa* spp.

**II. Karyotype reconstruction and evolution in *Musa* spp.**

The second aim of this work was to perform comparative karyotype analysis in a set of twenty edible banana clones and their wild relatives using FISH with chromosome/chromosome-arm specific oligo painting probes and shed a light on chromosome organization and structural chromosome changes accompanying the evolution of the genus *Musa*.

**III. Karyotype evolution within Musaceae**

The third aim of the thesis was to reveal chromosome structure of phylogenetically distinct species covering whole Musaceae family, which differ in genome size and basic chromosome number, and thus elucidate the genome organization at chromosomal level during the evolution and origin of species of this family.

**IV. Establishment of oligo painting FISH in other plant species**

The fourth aim of the thesis was to contribute to development and application of oligo painting FISH in fonio millet (*Digitaria exilis*) to analyse chromosome structure and integrity of newly developed whole genome sequence, and in the genera *Silene* spp. and *Lupinus* spp. to provide the information about sex chromosome evolution and to perform comparative cytogenetic mapping.

# 4 RESULTS

## 4.1 Summary

The first aim of this Ph.D. thesis was to develop a set of chromosome/chromosome-arm specific oligo painting probes in banana (*Musa* spp.). Unique oligomers (oligos) were identified according to Han *et al*. (2015) using a reference genome sequence of *Musa acuminata* ssp. *malaccensis* 'DH Pahang' (Martin *et al*., 2016) and selected with the Chorus program (https://github.com/forrestzhang/Chorus). Eight pseudomolecules corresponded to metacentric chromosomes, thus sets of 20 000 45-mers were designed for individual chromosome arms. However, pseudomolecules 1, 2 and 10 seemed to be acrocentric, in which sets of 20 000 oligomers covering only their long arms could be identified. Due to a lower oligomer density, peri-centromeric regions of all pseudomolecules were excluded from oligomer selection and probe preparation. The chromosome-arm specific oligomer libraries with a density from 0.9 to 2.1 oligos per 1 kb were synthesized as so-called immortal libraries by Arbor Biosciences (Ann Arbor, MI, USA), then fluorescently labeled and used as probes for FISH.

In order to verify their specificity, the newly developed chromosome/chromosome-arm specific oligo probes were hybridized to mitotic metaphase chromosomes of *M. acuminata* ssp. *malaccensis* 'Pahang' (A genome), from which the reference genome sequence was developed. Oligo painting FISH resulted in visible signals covering individual chromosome arms. Subsequently, these painting probes were used for FISH also in closely related banana species, which played a role in the evolution of most banana edible cultivars, *M. balbisiana* 'Tani' (B genome) and *M. schizocarpa* 'Schizocarpa' (S genome), to prove their suitability for comparative karyotype analysis. In *M. balbisiana*, a large translocation of the long arm of chromosome 3 to the long arm of chromosome 1 was detected. Painting probes were also hybridized onto the less condensed meiotic pachytene chromosomes, which provided valuable information about chromosome structure in more detail. Moreover, for better characterization of *Musa* chromosomes, previously developed cytogenetic landmarks, namely probes for 45S rDNA and 5S rDNA, two

satellites CL18 and CL33 and a BAC clone 2G17, were integrated with the painting probes and molecular karyotypes of three diploid *Musa* species were created (Šimoníková *et al*., 2019).

In the second part of the work, karyotypes of twenty representatives of Eumusa section of genus *Musa*, comprising edible banana clones and their probable progenitors, have been studied using developed oligo painting probes. Comparative karyotype analysis in six diploid subspecies of *M. acuminata* and in *M. balbisiana* was performed. In structural homozygous *M. acuminata* ssp. *banksii* 'Banksii' and *M. acuminata* ssp. *microcarpa* 'Borneo', no detectable chromosome translocations were observed compared to the reference genome of *M. acuminata* ssp. *malaccensis* 'DH Pahang'. However, in other subspecies of *M. acuminata* (ssp. *zebrina*, spp. *burmannicoides*, ssp. *siamea* and ssp. *burmannica*), different subspecies-specific chromosome translocations were detected. In *M. acuminata* ssp. *zebrina* 'Maia Oa', a reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8 was observed.

In three closely related subspecies of *M. acuminata* - spp. *burmannicoides*, ssp. *siamea* and ssp. *burmannica*, representing one genetic group, two types of translocations were observed. The first translocation involved a transfer of a part of a long arm of chromosome 8 to a long arm of chromosome 2, the second was a reciprocal translocation involving segments of a long arm of chromosome 1 and short arm of chromosome 9. Detected translocations were presented in both chromosome sets in *M. acuminata* ssp. *burmannicoides* 'Calcutta 4' and *M. acuminata* ssp. *siamea* 'Pa Rayong', whereas only one chromosome set contained these translocations in *M. acuminata* ssp. *burmannica* 'Tavoy'. Moreover, additional subspecies-specific translocations were observed in one chromosome set in *M. acuminata* ssp. *siamea* and *M. acuminata* ssp. *burmannica*, indicating their structural chromosome heterozygosity and hybrid origin. In *M. balbisiana* 'Pisang Klutuk Wulung', the same translocation, which was previously detected by Šimoníková *et al*. (2019) in another *M. balbisiana* accession, was observed.

Chromosome painting probes were hybridized also to edible diploid *Musa* clones 'Huti White', 'Huti (Shumba nyeelu)' and 'Ndyali' from Mchare group (AA

genome composition), confirming their hybrid origin. One chromosome set contained a reciprocal translocation involving segments of long arms of chromosomes 1 and 4. Second chromosome set contained a reciprocal translocation between chromosomes 3 and 8, which was already detected in *M. acuminata* ssp. *zebrina*. In diploid cultivar 'Pisang Lilin' (AA genome constitution), a reciprocal translocation involving long arms of chromosomes 1 and 4 was observed in heterozygous state as well.

In autotriploid dessert bananas 'Cavendish' and 'Gros Michel' (AAA genome constitution), one chromosome set contained a reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8, which was already detected in *M. acuminata* ssp. *zebrina*. Another chromosome set comprised a reciprocal translocation involving segments of long arms of chromosomes 1 and 4, previously detected in Mchare cultivars and in 'Pisan Lilin'. Third chromosome set involved a translocation of whole short arm of chromosome 7 to long arm of chromosome 1, resulting in the presence of telocentric chromosome consisting of long arm of chromosome 7.

Edible bananas from Mutika/Lujugira group, 'Imbogo' and 'Kagera' (AAA genome composition), contained a *zebrina*-type reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8 in two chromosome sets. Moreover, in cultivar 'Imbogo', one chromosome was missing. In this case, chromosome painting revealed a Robertsonian translocation between long arms of chromosomes 1 and 7, whereas short arms of both chromosomes have been completely lost.

In three plantain cultivars '3 Hands Planty', 'Obino l'Ewai' and 'Amou' (AAB genome composition), the B-genome specific translocation, involving a segment of the long arm of chromosome 3 to long arm of chromosome 1, was detected in one chromosome set. Additionally, the plantain clone 'Amou' was found to be an aneuploid, in which one copy of chromosome 2 was missing. Finally, in two triploid cultivars with ABB genome constitution, namely 'Pelipita' and 'Saba sa Hapon', the B-genome specific translocation was found in two chromosome sets, as expected (Šimoníková *et al*., 2020).

In the third part of the work, which is yet to be published, chromosome-specific oligo painting probes were successfully hybridized also to chromosomes of distinct species covering whole Musaceae family (Table 1), which differ in genome size and chromosome number (x=9, 10 or 11).

**Table 1:** List of analysed accessions including their chromosome number and genome size

| Genus | Section | Accession name | Species/ Group | ITC code[a] | Number of chromosomes | Genome size Mb/1C |
|---|---|---|---|---|---|---|
| *Musa* | Rhodochlamys | Musa ornata (red fingers) | *ornata* | 1330 | 2n=2x=22 | 605; unpublished |
| *Musa* | Rhodochlamys | Musa ornata | *ornata* | 0370 | 2n=2x=22 | 600; unpublished |
| *Musa* | Rhodochlamys | Musa laterita | *laterita* | 1575 | 2n=2x=22 | 643; unpublished |
| *Musa* | Australimusa | Musa maclayi type Hung Si | *maclayi* | 0614 | 2n=2x=20 | 722; Bartoš *et al*., 2005 |
| *Musa* | Australimusa | Musa maclayi | *maclayi* | 1207 | 2n=2x=20 | 720; unpublished |
| *Musa* | Australimusa | Musa textilis | *textilis* | 0539 | 2n=2x=20 | 702; Bartoš *et al.*, 2005 |
| *Musa* | Australimusa | Skai | Fe‘i | 0883 | 2n=3x=30 | 733, unpublished |
| *Musa* | Australimusa | Tongkat Langit Papua | Fe‘i | 1716 | 2n=3x=30 | 730, unpublished |
| *Musa* | Callimusa | Musa coccinea | *coccinea* | 0287 | 2n=2x=20 | 705; unpublished |
| *Musa* | Callimusa | Musa beccarii | *beccarii* | 1070 | 2n=2x=18 | 763; Bartoš *et al*., 2005 |
| *Ensete* | - | Ensete ventricosum | *ventricosum* | 1387 | 2n=2x=18 | 592; unpublished |
| *Musella* | - | Musella | *lasiocarpa* | - | 2n=2x=18 | 604, unpublished |

*Code assigned by the International Transit Center (ITC, Leuven, Belgium)

Combination of chromosome-arm specific painting probes enabled the identification of numerous large chromosomal rearrangements in studied accessions. Observed chromosome structures in more distinct wild species corresponded to their evolution and phylogenetic relationships. The oligo painting study of the analysed accessions, representing individual phylogenetic groups of family Musaceae, was enlarged by the application of long-read sequencing technology, Oxford Nanopore, for precise determination of chromosome-translocation breakpoints. Example of chromosome-specific painting FISH on *Musa laterita* from Rhodochlamys section of genus *Musa* is shown on Figure 10. Newly observed translocations have not been detected in any analysed species from section Eumusa.

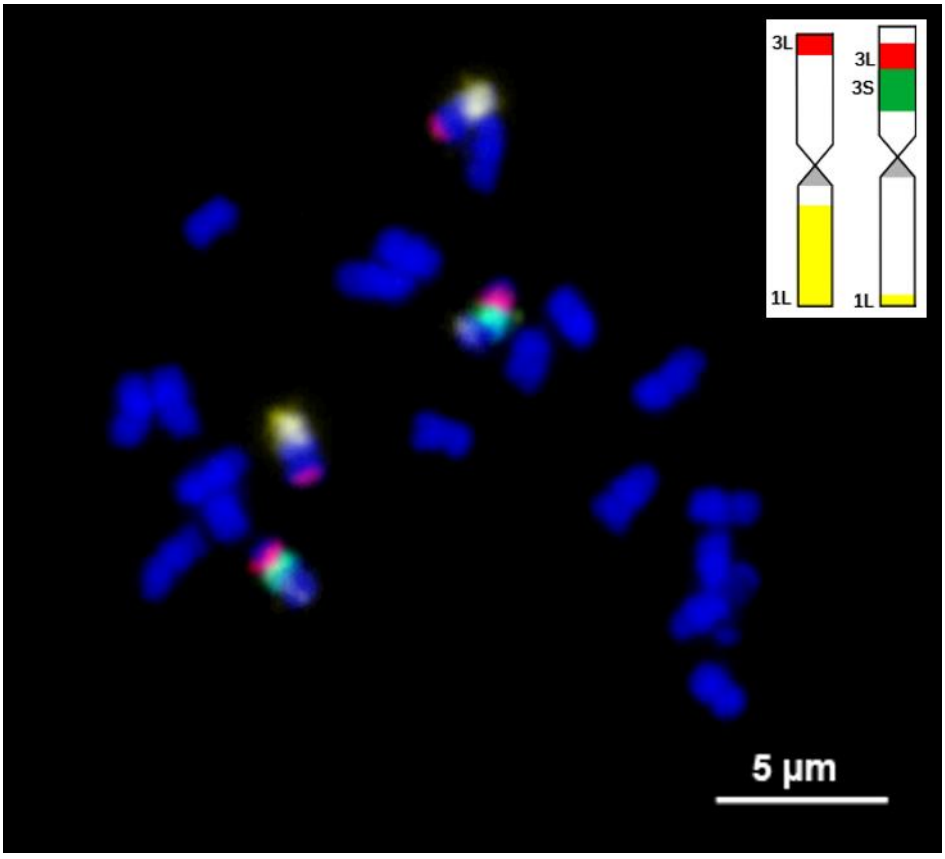



**Figure 10.** Example of oligo painting FISH on mitotic metaphase plate of wild species from Rhodochlamys section of genus *Musa* (*Musa laterita*, 2n=2x=22), probes for long arm of chromosome 3 (3L), short arm of chromosome 3 (3S) and long arm of chromosome 1 (1L) are labeled in red, green and yellow, respectively. Chromosomes were counterstained with DAPI.

Finally, chromosome-specific oligo painting FISH method was developed in fonio millet (*Digitaria exilis*) and another two genera, *Silene* spp. and *Lupinus* spp. In tetraploid species fonio millet (2n=4x=36), this technique was used to assess the quality of its newly developed genome assembly. A set of eighteen chromosome-specific oligo painting probes, which corresponded to eighteen newly created pseudomolecules, were successfully anchored to individual chromosomes *in situ*. As a result, no cross-hybridization between individual chromosomes was observed, thus a set of eighteen newly assembled pseudomolecules representing individual chromosomes was validated. Moreover, homoeologous chromosomes were unambiguously distinguished in this species (Abrouk *et al*., 2020).

The formation of sex chromosomes in diocieous species from genus *Silene*, which have XX/XY sex determination system, has been also studied using oligo panting FISH. Despite the absence of high quality whole genome sequence, the X chromosome-specific oligo probe, which was designed for *S. latifolia*, successfully generated specific signals mainly in the subtelomeric regions of the X and Y chromosomes in *S. latifolia* and *S. dioica*. However, in gynodioecous *S. vulgaris* and *S. maritima*, the probe hybridized to short arms of the three pairs of autosomes. These results confirmed that sex chromosome evolution in studied dioecious *Silene* spp. was accompanied by multiple chromosomal rearrangements (Bačovský *et al*., 2020).

A set of four oligo painting probes covering both arms of chromosome 6 of the *Lupinus angustifolius* was designed and mapped on mitotic chromosome spreads together with previously identified single copy BAC clones specific to chromosome 6. Using these probes, comparative cytogenetic mapping among the smooth-seeded (*L. angustifolius* L., *L. cryptanthus* Shuttlew., *L. micranthus* Guss.) and rough-seeded (*L. cosentinii* Guss. and *L. pilosus* Murr.) lupin species was performed. Structural chromosome changes, which accompanied the evolution and speciation of lupin species, have been detected between studied species using reciprocal oligo-FISH and BAC-FISH mapping (Bielski *et al*., 2020).

## 4.2 Original papers

4.2.1 Chromosome painting facilitates anchoring reference genome sequence to chromosomes *in situ* and integrated karyotyping in banana (*Musa* spp.)

(Appendix I)

4.2.2 Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure

(Appendix II)

4.2.3 Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate

(Appendix III)

4.2.4 The formation of sex chromosomes in *Silene latifolia* and *S. dioica* was accompanied by multiple chromosomal rearrangements

(Appendix IV)

4.2.5 The puzzling fate of a lupin chromosome revealed by reciprocal oligo-FISH and BAC-FISH mapping

(Appendix V)

### 4.2.1 Chromosome painting facilitates anchoring reference genome sequence to chromosomes *in situ* and integrated karyotyping in banana (*Musa* spp.)

Šimoníková, D., Němečková, A., Karafiátová, M., Uwimana, B., Swennen, R., Doležel, J., Hřibová, E.

*Frontiers in Plant Science* **10**: 1503, 2019

doi: 10.3389/fpls.2019.01503

IF: 4.402

**Abstract:**

Oligo painting FISH was established to identify all chromosomes in banana (*Musa* spp.) and to anchor pseudomolecules of reference genome sequence of *Musa acuminata* spp. *malaccensis* "DH Pahang" to individual chromosomes *in situ*. A total of 19 chromosome/chromosome-arm specific oligo painting probes were developed and were shown to be suitable for molecular cytogenetic studies in genus *Musa*. For the first time, molecular karyotypes of diploid *M. acuminata* spp. *malaccensis* (A genome), *M. balbisiana* (B genome), and *M. schizocarpa* (S genome) from the Eumusa section of *Musa*, which contributed to the evolution of edible banana cultivars, were established. This was achieved after a combined use of oligo painting probes and a set of previously developed banana cytogenetic markers. The density of oligo painting probes was sufficient to study chromosomal rearrangements on mitotic as well as on meiotic pachytene chromosomes. This advance will enable comparative FISH mapping and identification of chromosomal translocations which accompanied genome evolution and speciation in the family *Musaceae*.

### 4.2.2 Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure

Šimoníková, D., Němečková, A., Čížková, J., Brown, A., Swennen, R., Doležel, J., Hřibová, E.

*International Journal of Molecular Sciences* **21**: 7915, 2020

doi: 10.3390/ijms21217915

IF: 4.602

**Abstract:**

Edible banana cultivars are diploid, triploid, or tetraploid hybrids, which originated by natural cross hybridization between subspecies of diploid *Musa acuminata*, or between *M. acuminata* and diploid *Musa balbisiana*. The participation of two other wild diploid species *Musa schizocarpa* and *Musa textilis* was also indicated by molecular studies. The fusion of gametes with structurally different chromosome sets may give rise to progenies with structural chromosome heterozygosity and reduced fertility due to aberrant chromosome pairing and unbalanced chromosome segregation. Only a few translocations have been classified on the genomic level so far, and a comprehensive molecular cytogenetic characterization of cultivars and species of the family Musaceae is still lacking. Fluorescence *in situ* hybridization (FISH) with chromosome-arm-specific oligo painting probes was used for comparative karyotype analysis in a set of wild *Musa* species and edible banana clones. The results revealed large differences in chromosome structure, discriminating individual accessions. These results permitted the identification of putative progenitors of cultivated clones and clarified the

genomic constitution and evolution of aneuploid banana clones, which seem to be common among the polyploid banana accessions. New insights into the chromosome organization and structural chromosome changes will be a valuable asset in breeding programs, particularly in the selection of appropriate parents for cross hybridization.

### 4.2.3 Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate

Abrouk, M., Ahmed, H.I., Cubry, P., Šimoníková, D., Cauet, S., Bettgenhaeuser, J., Gapa, L., Pailles, Y., Scarcelli, N., Couderc, M., Zekraoui, L., Kathiresan, N., Čížková, J., Hřibová, E., Doležel, J., Arribat, S., Bergès, H., Wieringa, J.J., Gueye, M., Kane, N.A., Leclerc, C., Causse, S., Vancoppenolle, S., Billot, C., Wicker, T., Vigouroux, Y., Barnaud, A., Krattinger, S.G.

*Nature communications* **11**: 4488, 2020

doi: 10.1038/s41467-020-18329-4

IF: 12.121

**Abstract:**

Sustainable food production in the context of climate change necessitates diversification of agriculture and a more efficient utilization of plant genetic resources. Fonio millet (*Digitaria exilis*) is an orphan African cereal crop with a great potential for dryland agriculture. Here, we establish high-quality genomic resources to facilitate fonio improvement through molecular breeding. These include a chromosome-scale reference assembly and deep re-sequencing of 183 cultivated and wild *Digitaria* accessions, enabling insights into genetic diversity, population structure, and domestication. Fonio diversity is shaped by climatic, geographic, and ethnolinguistic factors. Two genes associated with seed size and shattering showed signatures of selection. Most known domestication genes from other cereal models however have not experienced strong selection in fonio, providing direct targets to rapidly improve this crop for agriculture in hot and dry environments.

### 4.2.4 The formation of sex chromosomes in *Silene latifolia* and *S. dioica* was accompanied by multiple chromosomal rearrangements

Bačovský, V., Čegan, R., Šimoníková, D., Hřibová, E., Hobza, R.

*Frontiers in Plant Science* **11**: 205, 2020

doi: 10.3389/fpls.2020.00205

IF: 4.402

**Abstract:**

The genus *Silene* includes a plethora of dioecious and gynodioecious species. Two species, *Silene latifolia* (white campion) and *Silene dioica* (red campion), are dioecious plants, having heteromorphic sex chromosomes with an XX/XY sex determination system. The X and Y chromosomes differ mainly in size, DNA content and posttranslational histone modifications. Although it is generally assumed that the sex chromosomes evolved from a single pair of autosomes, it is difficult to distinguish the ancestral pair of chromosomes in related gynodioecious and hermaphroditic plants. We designed an oligo painting probe enriched for X-linked scaffolds from currently available genomic data and used this probe on metaphase chromosomes of *S. latifolia* (2n = 24, XY), *S. dioica* (2n = 24, XY), and two gynodioecious species, *S. vulgaris* (2n = 24) and *S. maritima* (2n = 24). The X chromosome-specific oligo probe produces a signal specifically on the X and Y chromosomes in *S. latifolia* and *S. dioica*, mainly in the subtelomeric regions. Surprisingly, in *S. vulgaris* and *S. maritima*, the probe hybridized to three pairs of autosomes labeling their p-arms. This distribution suggests that sex chromosome evolution was accompanied by extensive chromosomal rearrangements in studied dioecious plants.

### 4.2.5 The puzzling fate of a lupin chromosome revealed by reciprocal oligo-FISH and BAC-FISH mapping

Bielski, W., Książkiwicz, M., Šimoníková, D., Hřibová, E., Susek, K., Naganowska, B.

*Genes* **11**: e1489, 2020

doi: 10.3390/genes11121489

IF: 3.759

**Abstract:**

Old World lupins constitute an interesting model for evolutionary research due to diversity in genome size and chromosome number, indicating evolutionary genome reorganization. It has been hypothesized that the polyploidization event which occurred in the common ancestor of the Fabaceae family was followed by a lineage-specific whole genome triplication (WGT) in the lupin clade, driving chromosome rearrangements. In this study, chromosome-specific markers were used as probes for heterologous fluorescence *in situ* hybridization (FISH) to identify and characterize structural chromosome changes among the smooth-seeded (*Lupinus angustifolius* L., *Lupinus cryptanthus* Shuttlew., *Lupinus micranthus* Guss.) and rough-seeded (*Lupinus cosentinii* Guss. and *Lupinus pilosus* Murr.) lupin species. Comparative cytogenetic mapping was done using FISH with oligonucleotide probes and previously published chromosome-specific bacterial artificial chromosome (BAC) clones. Oligonucleotide probes were designed to cover both arms of chromosome Lang06 of the *L. angustifolius* reference genome separately. The chromosome was chosen for the in-depth study due to observed structural variability among wild lupin species revealed by BAC-FISH and supplemented by *in silico* mapping of recently released lupin genome assemblies. The results highlighted changes in synteny within the Lang06 region between the lupin species, including

putative translocations, inversions, and/or non-allelic homologous recombination, which would have accompanied the evolution and speciation.

## 4.3 Conference presentations

4.3.1 Chromosome oligo painting facilitates the analysis of karyotype evolution in Musaceae
(poster presentation; Appendix VI)

4.3.2 Chromosome-specific oligo painting elucidates large variation in Eumusa genome
(poster presentation; Appendix VII)

4.3.3 Studium struktury chromozómů banánovníku (*Musa* spp.)
(oral presentation)

### 4.3.1 Chromosome oligo painting facilitates the analysis of karyotype evolution in Musaceae

Šimoníková, D., Čížková, J., Karafiátová, M., Němečková, A., Doležel, J., Hřibová, E.

In: Abstracts of the "22nd International Chromosome Conference". Prague, Czech Republic, 2018

**Abstract:**

The family Musaceae comprises three genera: *Musa*, *Ensete* and *Musella*. While *Ensete* (2n=18) and *Musella* (2n=18) are represented by a few and one endemic species, respectively, genus *Musa* contains about 70 different species. Based on plant morphology and basic chromosome number, *Musa* species are classified into four sections. The largest of them, Eumusa with x=11 comprises most of all edible banana cultivars. They originated by intra- and inter-specific crosses between wild diploids *M. acuminata* and *M. balbisiana*. The section Rhodochlamys (x=11) contains ornamental species, which are closely related to those of Eumusa. The section Australimusa (x=10) is represented by species growing in Southeast Asia and contains a peculiar group of edible banana clones known as Fe'i. The section Callimusa is the most diverse and contains species differing in basic chromosome number (x=9, 10) and species which seems to be closely related to *Ensete* and *Musella*. In this work we took the advantage of the availability of whole genome sequence of double haploid *M. acuminata* 'DH Pahang' (2n=22) to design chromosome-specific oligomers suitable for fluorescence *in situ* hybridization (FISH). The oligomers were designed for each of the 22 chromosome arms using the Chorus computer program and synthetized by Arbor Biosciences (Ann Arbor, MI, USA). The oligomers were labeled using reverse transcription either by 5'-biotin or 5'-digoxigenin-specific primers according to Han *et al*. (Genetics 200:771, 2015) and

used for FISH in a set of eleven accessions representing genetic diversity and phylogeny of the *Musaceae* family. This work opens avenues for comparative analysis of structural chromosome changes and sheds light on karyotype evolution in Musaceae.

### 4.3.2 Chromosome-specific oligo painting elucidates large variation in Eumusa genome

Šimoníková, D., Doležel, J., Hřibová, E.

In: Abstracts of the "International Conference on Polyploidy". Ghent, Belgium, 2019

**Abstract:**

The family Musaceae is classified into three genera: *Musa*, *Musella* and *Ensete*. The largest of them, *Musa*, comprises about 70 species, whereas closely related *Musella* (2n = 18) and *Ensete* (2n = 18) are represented by only a few species. Based on basic chromosome number and plant morphology, the genus *Musa* has been divided into four sections: Eumusa (x = 11), Rhodochlamys (x = 11), Australimusa (x = 10) and Callimusa (x = 9, 10). This work is focused on the largest section Eumusa, which showed surprisingly large variation in genome constitution in various subspecies of diploid *M. acuminata* (A genome), *M. balbisiana* (B genome), *M. schizocarpa* (S genome) and their triploid hybrid clones (AAA, AAB and ABB genomes). A set of 18 accessions representing Eumusa section was used for chromosome-specific oligo painting. A reference genome sequence of *M. acuminata* doubled haploid clone 'Pahang' (2n = 22) facilitated the identification of chromosome arm-specific oligomers that were designed using Chorus program and synthetized by Arbor Biosciences (Ann Arbor, MI, USA). The oligos were labeled using reverse transcription by biotin- or digoxigenin-tagged reverse primer and used as probes for fluorescence *in situ* hybridization (Han *et al*., Genetics 200:771, 2015). This powerful approach allowed identification of all chromosomes within a karyotype and discovering chromosomal rearrangements, which accompanied the evolution of the family Musaceae.

## 4.3.3 Studium struktury chromozómů banánovníku (*Musa* spp.)

Šimoníková, D., Hřibová, E.

In: Abstracts of the "53. výroční cytogenetická konference". Olomouc, Czech Republic, 2020

[In Czech]

**Abstract:**

V posledních letech se při studiu rostlinných genomů stala populární cytogenetická metoda zvaná malování chromozomů (chromosome painting)**.** Mikroskopická analýza „malovaných" chromozomů umožňuje kromě identifikace jednotlivých chromozomů také analýzu chromozomálních přestaveb, ke kterým došlo v průběhu evoluce a vzniku druhů.

V cytogenetice člověka malování chromozomů slouží především k diagnostice numerických a strukturních chromozomových aberací, a je využíváno již od 90. let 20. století, kdy se pro přípravu chromozomově specifických sond využívá tzv. průtoková cytometrie či mikrodisekce. Rostliny však na rozdíl od člověka a dalších živočichů obsahují velké množství repetitivních sekvencí, které jsou roztroušeny po celém genomu a znemožňují jednoznačnou identifikaci jednotlivých chromozomů použitím metodických přístupů využívaných v lidské a živočišné cytogenetice.

Řešením se stala metoda BAC FISH založená na klonování fragmentů DNA do bakterií. Vybrané BAC klony, jedinečné pro jednotlivé chromozomy, po naznačení fluorescenční značkou poté umožní identifikaci vybraných chromozomů. Tato metoda je však vhodná pouze pro rostliny s malou velikostí genomu a nízkým obsahem repetic a jejich celogenomová sekvence byla získána sekvenováním fragmentů DNA klonovaných ve vektorech BAC – *Arabidopsis thaliana* a *Brachypodium distachyon*. Chromozomové malování se u ostatních rostlinných druhů rozšířilo až s příchodem sekvenování nové generace, kdy je mnohem jednodušší sestavení celogenomových sekvencí, které jsou využity

pro identifikaci krátkých oligomerů specifických pro jednotlivé chromozomy, které slouží jako cytogenetické markery (tzv. oligo painting). Pro přípravu chromozomově specifických oligo sond se využívají krátké úseky DNA (45 bází) jedinečné pro celý chromozom či jeho vybranou část. V případě banánovníku byly za účelem malování chromozomů vytvořeny cytogenetické sondy specifické pro jednotlivá chromozomální ramena a to s využitím referenční genomové DNA sekvence druhu *M. acuminata*. Krátké úseky DNA jedinečné pro jednotlivá ramena 11 chromozomů tohoto druhu byly fluorescenčně naznačeny a získané sondy byly poté hybridizovány na vybrané druhy reprezentující čeleď banánovníkovitých (*Musaceae*), které se kromě počtu chromozomů liší i velikostí genomu. Byly zjištěny velké rozdíly v genomové struktuře na úrovni chromozomů jak mezi jednotlivými druhy, tak mezi jednotlivými odrůdami jedlých banánovníků. Přítomnost specifických translokací, ke kterým došlo v průběhu evoluce banánovníku, umožnila identifikovat pravděpodobné předky jedlých kultivarů a informace o detailní struktuře chromozomů tak usnadní šlechtitelům výběr vhodných klonů pro šlechtění.

# 5 CONCLUSION

In this work, chromosome-specific oligo painting technique was developed in banana (*Musa* spp.). Using a set of nineteen chromosome/chromosome-arm specific oligo painting probes, DNA pseudomolecules of reference genome sequence of *Musa acuminata* ssp. *malaccensis* 'DH Pahang' were anchored to individual chromosomes *in situ* and all *Musa* chromosomes have been unambiguously identified for the first time. Karyotypes of three diploid *M. acuminata* ssp. *malaccensis* (A genome), *M. balbisiana* (B genome) *and M. schizocarpa* (S genome) species from genus *Musa*, which are progenitors of majority of edible banana clones, were established using a combination of these oligo painting probes and previously developed cytogenetic markers.

Further, oligo painting probes enabled to create karyotypes of twenty representatives of the Eumusa section of genus *Musa*, including wild species commonly used in breeding programs and economically significant edible cultivars. Large differences in chromosome structures, specific for individual species, subspecies and groups of edible accessions were identified. Presence of specific types of large translocations in analysed *Musa* accessions led to the identification of most probable parents of edible cultivated clones. Moreover, structural chromosome heterozygosity confirmed hybrid origin of cultivated clones and some wild subspecies of *M. acuminata*, and pointed to a possible reason of their reduced fertility. Additionally, newly designed oligo painting probes were proved to be suitable for comparative karyotype analysis and identification of structural chromosome changes that accompanied the evolution and speciation within Musaceae family. Twelve accessions representing individual phylogenetic groups of Musaceae species, differing in genome size and chromosome number, were selected for detailed cytogenetic analysis. Extensive chromosomal reaarangements among distantly related species have been observed. To give a precise description of translocation breakpoints, Oxford Nanopore sequencing technology, which provides long sequencing reads, was used and the results will be published later.

Finally, chromosome-specific oligo painting technique was developed and successfully performed also in other plant species, including *Digitaria exilis*, and genera *Lupinus* spp. and *Silene* spp.

The Ph.D. thesis fills the gap in molecular cytogenetic studies of *Musa*. The knowledge about genome structure at chromosomal level and the identification of structural chromosome heterozygosity of cultivated edible bananas and their probable progenitors offers a valuable asset in selecting appropriate parents for cross-hybridization during the breeding of improved cultivars.

# 6 LIST OF ABBREVIATIONS

| | |
|---|---|
| 1C | C (constant)-value, DNA amount of a ‘holoploid genome’ with chromosome number n |
| BAC | bacterial artificial chromosome |
| bp | base pairs |
| CIRAD | The French Agricultural Research Centre for International Development |
| CRBC | chicken red blood cell nuclei |
| DAPI | 4′,6-diamidino-2-phenylindole |
| DArTseq | Diversity Array Technology sequencing |
| DH Pahang | Doubled haploid Pahang |
| DNA | deoxyribonucleic acid |
| EAHB | East African Highland banana |
| eBSV | endogenous Banana streak virus |
| FISH | fluorescence *in situ* hybridization |
| G1 phase | gap/growth 1 phase |
| GISH | genome *in situ* hybridization |
| ITC | International *Musa* Germplasm Transit Center |
| ITS | internal transcribed spacer |
| kb | kilobase pairs |
| LTR | long terminal repeat |
| Mb | mega base pairs |
| n | number of chromosomes in a haploid cell |
| NGS | next generation sequencing |
| NM | Northern Malayan group |

| | |
|---|---|
| NOR | nucleolar organizing region |
| nt | nucleotide |
| PCR | polymerase chain reaction |
| RADseq | Restriction-site associated DNA sequencing |
| rDNA | ribosomal deoxyribonucleic acid |
| RNA | ribonucleic acid |
| rRNA | ribosomal ribonucleic acid |
| SNP | single nucleotide polymorphism |
| spp. | several species |
| ssp. | subspecies |
| SSR | simple sequence repeat |
| ST | Standard group |
| Tm | melting temperature |
| TR4 | Tropical Race 4 |
| var. | variety |
| WGS | whole genome sequencing |
| x | basic chromosome number |

# 7 LIST OF APPENDICES

## Original Papers

Appendix I: Chromosome painting facilitates anchoring reference genome sequence to chromosomes *in situ* and integrated karyotyping in banana (*Musa* spp.)

Appendix II: Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure

Appendix III: Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate

Appendix IV: The formation of sex chromosomes in *Silene latifolia* and *S. dioica* was accompanied by multiple chromosomal rearrangements

Appendix V: The puzzling fate of a lupin chromosome revealed by reciprocal oligo-FISH and BAC-FISH mapping

## Published Posters

Appendix VI: Chromosome oligo painting facilitates the analysis of karyotype evolution in Musaceae

Appendix VII: Chromosome-specific oligo painting elucidates large variation in Eumusa genome

# APPENDIX I

**Chromosome painting facilitates anchoring reference genome sequence to chromosomes *in situ* and integrated karyotyping in banana (*Musa* spp.)**

Šimoníková, D., Němečková, A., Karafiátová, M., Uwimana, B., Swennen, R., Doležel, J., Hřibová, E.

*Frontiers in Plant Science* **10**: 1503, 2019

doi: 10.3389/fpls.2019.01503

IF: 4.402

frontiers in Plant Science

ORIGINAL RESEARCH
published: 20 November 2019
doi: 10.3389/fpls.2019.01503

# Chromosome Painting Facilitates Anchoring Reference Genome Sequence to Chromosomes *In Situ* and Integrated Karyotyping in Banana (*Musa* Spp.)

*Denisa Šimoníková[1], Alžběta Němečková[1], Miroslava Karafiátová[1], Brigitte Uwimana[2], Rony Swennen[3,4,5], Jaroslav Doležel[1] and Eva Hřibová[1]**

*[1] Institute of Experimental Botany, Czech Academy of Sciences, Centre of the Region Hana for Biotechnological and Agricultural Research, Olomouc, Czechia, [2] Banana Breeding, International Institute of Tropical Agriculture, Kampala, Uganda, [3] Bioversity International, Banana Genetic Resources, Heverlee, Belgium, [4] Division of Crop Biotechnics, Laboratory of Tropical Crop Improvement, Katholieke Universiteit Leuven, Leuven, Belgium, [5] Banana Breeding, International Institute of Tropical Agriculture, Arusha, Tanzania*

OPEN ACCESS

**Edited by:**
*Martin A. Lysak, Masaryk University, Czechia*

**Reviewed by:**
*Alexander Betekhtin, University of Silesia of Katowice, Poland*
*Jiming Jiang, University of Wisconsin-Madison, United States*

***Correspondence:**
*Eva Hřibová*
*hribova@ueb.cas.cz*

**Specialty section:**
*This article was submitted to Plant Systematics and Evolution, a section of the journal Frontiers in Plant Science*

**Received:** *29 July 2019*
**Accepted:** *29 October 2019*
**Published:** *20 November 2019*

**Citation:**
*Šimoníková D, Němečková A, Karafiátová M, Uwimana B, Swennen R, Doležel J and Hřibová E (2019) Chromosome Painting Facilitates Anchoring Reference Genome Sequence to Chromosomes In Situ and Integrated Karyotyping in Banana (Musa Spp.). Front. Plant Sci. 10:1503. doi: 10.3389/fpls.2019.01503*

Oligo painting FISH was established to identify all chromosomes in banana (*Musa* spp.) and to anchor pseudomolecules of reference genome sequence of *Musa acuminata* spp. *malaccensis* "DH Pahang" to individual chromosomes *in situ*. A total of 19 chromosome/chromosome-arm specific oligo painting probes were developed and were shown to be suitable for molecular cytogenetic studies in genus *Musa*. For the first time, molecular karyotypes of diploid *M. acuminata* spp. *malaccensis* (A genome), *M. balbisiana* (B genome), and *M. schizocarpa* (S genome) from the Eumusa section of *Musa*, which contributed to the evolution of edible banana cultivars, were established. This was achieved after a combined use of oligo painting probes and a set of previously developed banana cytogenetic markers. The density of oligo painting probes was sufficient to study chromosomal rearrangements on mitotic as well as on meiotic pachytene chromosomes. This advance will enable comparative FISH mapping and identification of chromosomal translocations which accompanied genome evolution and speciation in the family *Musaceae*.

**Keywords: banana, chromosome identification, fluorescence *in situ* hybridization, molecular karyotype, *Musa*, oligo painting FISH**

## INTRODUCTION

Bananas (*Musa* spp.) are grown in tropical and subtropical regions of South East Asia, Africa and South America (Häkkinen, 2013; Janssens et al., 2016). They are one of the world's major fruit crops, a staple and important export commodity for millions of people living mainly in developing countries. Despite the importance and breeding efforts (Ortiz and Swennen, 2014; Brown et al., 2017), little is known about banana genome structure, organization and evolution at chromosomal level across the whole *Musaceae* family.

The genus *Musa* comprises about 75 species and numerous cultivated edible clones. Based on a set of morphological descriptors (IPGRI-INIBAP/CIRAD, 1996) and basic chromosome number (x),

the genus *Musa* has traditionally been divided into four sections: Eumusa (x = 11), Rhodochlamys (x = 11), Australimusa (x = 10), and Callimusa (x = 9, 10) (Cheesman, 1947). Argent (1976) created a separate section Ingentimusa which contains a single species *Musa ingens* with the lowest basic chromosome number (x = 7). However, genotyping using molecular markers revealed close relationship of *M. ingens* with other species of sections Callimusa and Australimusa (Li et al., 2010). Most of the modern edible banana clones originated within section Eumusa after intra- and inter-specific crosses between two wild diploid species *M. acuminata* (donor of A genome) and *M. balbisiana* (donor of B genome). In some cases, diploid *M. schizocarpa* (S genome) contributed to the evolution of edible clones, mainly after cross-breeding with diploid *M. acuminata* (Carreel et al., 1994; Čížková et al., 2013; Němečková et al., 2018). The spontaneous intra- and inter-specific crosses gave arise to seed sterile diploid (AA, AB) and triploid (AAA, AAB, or ABB) edible banana cultivars. Although tetraploid clones (AAAB, AABB) that originated spontaneously are known (Simmonds and Shepherd, 1955; Simmonds, 1956), currently cultivated tetraploid bananas were obtained in the breeding programs.

Species of genus *Musa* have a relatively small genome, ranging from 550 to 750 Mbp/1C (Doležel et al., 1994; Lysák et al., 1999; Asif et al., 2001; Kamaté et al., 2001; Bartoš et al., 2005; Čížková et al., 2015) and until now it was possible to identify only a few chromosomes in their karyotypes. The attempts were hampered by the relatively high number of chromosomes, their small size at mitotic metaphase (1–2 μm) and morphological similarity (Doleželová et al., 1998; Osuji et al., 1998; D'Hont et al., 2000). Chromosome banding, which was found informative in plant species with large and repeat-rich genomes, including wheat and rye (Gill and Kimber, 1977; Gill et al., 1991), did not result in diagnostic chromosome banding patterns in *Musa*, similar to many other plant species (Greilhuber, 1977; Schubert et al., 2001).

The application of fluorescence *in situ* hybridization (FISH), usually done with probes for DNA repeats with chromosome-specific distribution, provided a powerful approach to identify chromosomes in a range of plant species and study chromosome structural changes (e.g., Liu et al., 2011; Danilova et al., 2014; Amosova et al., 2017; Hou et al., 2018). Unfortunately, its use in *Musa* was hampered by the lack of suitable probes (Doleželová et al., 1998; Osuji et al., 1998; Valárik et al., 2002; Hřibová et al., 2007; Čížková et al., 2013). Until now, only NOR-bearing satellite chromosome, two chromosomes with clusters of tandem repeats CL18 and CL33, and two chromosomes bearing 5S rDNA loci can be cytogenetically identified in *M. acuminata* and *M. balbisiana* (Čížková et al., 2013). In *M. schizocarpa*, one chromosome pair bearing NOR and two chromosome pairs bearing tandem repeats CL18 and CL33 and other four chromosome pairs with 5S rDNA loci can be identified cytogenetically (Čížková et al., 2013). Even the mining of the reference genome sequence of *M. acuminata* "DH Pahang" (D'Hont et al., 2012) did not result in identification of sequences suitable as FISH probes useful for unambiguous identification of all *Musa* chromosomes and their anchoring to the genome sequence.

A method for chromosome painting, which allows fluorescent labeling of whole chromosomes, was developed in the late 1980s. This advance revolutionized human cytogenetics and found numerous applications in animal cytogenetics (e.g., Speicher et al., 1996; Cremer and Cremer, 2001; Ferguson-Smith and Trifonov, 2007). The original method was based on FISH with whole chromosome probes obtained from chromosomes isolated by flow cytometric sorting or microdissection. This was the reason why the method failed in plants where a majority of DNA repeats is distributed across the whole genome and only a minority of sequences are unique and chromosome-specific (Schubert et al., 2001). A solution was to use pools of chromosome-specific BAC (Bacterial Artificial Chromosome) clones (Lysák et al., 2001). However, the development of chromosome BAC pools requires whole genome sequence obtained after clone by clone (BAC by BAC) sequencing to identify single or low copy BAC clones useful for painting. Thus, the method is suitable for species with small genomes and containing low amounts of DNA repeats. Till now, painting using chromosome-specific BAC pools was used in dicotyledonous species with small nuclear genomes— *Arabidopsis* and its closely related species (e.g., Lysák et al., 2001; Mandáková and Lysák, 2008; Mandáková et al., 2013) as well as in monocot *Brachypodium distachyon* (Idziak et al., 2014). The attempts to use BAC FISH in banana were not successful due to the lack of a larger number of BAC clones containing single or low copy sequences (Hřibová et al., 2008).

The recent progress in the production of reference genome sequences and in technologies for DNA synthesis provided an alternative opportunity for affordable preparation of whole chromosome probes (chromosome paints) for FISH. The method called oligo painting FISH (Han et al., 2015) is based on *in silico* identification of large numbers of short (usually 45–50 bp) and unique (single copy) sequences in pseudomolecules of individual chromosomes, or their parts, synthesis of oligonucleotides, and their fluorescent labeling. A pool of synthesized and fluorescently labeled oligonucleotides is then used as a probe for FISH. Thus, the oligo painting FISH provides an opportunity to identify individual chromosomes and chromosome regions in *Musa*, perform comparative chromosome analysis and characterize chromosomal rearrangements (Qu et al., 2017; Braz et al., 2018; Xin et al., 2018; Jiang, 2019).

The present study fills the important gap in molecular cytogenetics of *Musa*. The application of oligo painting FISH described here allows anchoring genome sequence to chromosomes *in situ* and unambiguous identification of all *Musa* chromosomes after development of molecular karyotypes by a combined use of oligo painting probes and existing cytogenetic landmarks. Molecular karyotypes are described and compared for the three main genomes of Eumusa section—*M. acuminata* ssp. *malaccensis*, *M. balbisiana*, and *M. schizocarpa*, which contributed to the evolution of many edible banana clones.

## MATERIALS AND METHODS

### Plant Material and Preparations of Chromosome Spreads

Representatives of three species from the section Eumusa were obtained as *in vitro* rooted plants from the International *Musa* Transit Centre (ITC, Bioversity International, Leuven, Belgium). *In vitro* plants were transferred to garden soil and maintained in a heated greenhouse. **Table 1** lists the accessions used in this study.

**TABLE 1 |** List of analyzed accessions, their genomic constitution, genome size, and the number of loci identified on mitotic metaphase chromosomes (data from Čížková et al., 2013).

| Species | Accession name | ITC code[a] | Genomic constitution | Genome size (1C) | Chromosome number (2n) | The number of loci in diploid cells (2n = 22) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | 45S rDNA | 5S rDNA | BAC 2G17 | CL33 | CL18 |
| *M. acuminata* ssp. *malaccensis* | Pahang | 0609 | AA | 594 Mbp[b] | 22 | 2 | 6 | 2 | 4 | 2 |
| *M. balbisiana* | Tani | 1120 | BB | 551 Mbp[b] | 22 | 2 | 6 | 2 | 0 | 4 |
| *M. schizocarpa* | Schizocarpa | 0560 | SS | 671 Mbp[b] | 22 | 2 | 12 | 2 | 4 | 2 |

[a]*Code assigned by the International Transit Centre (ITC, Leuven, Belgium)*

[b]*DNA content was estimated by flow cytometry using Glycine max L. cv. Polanka (2C = 2.5pg DNA) which served as an internal reference standard (Čížková et al., 2013).*

Male buds of *M. acuminata* "Pahang" and *M. balbisiana* "Tani" were obtained from the research station of the International Institute of Tropical Agriculture in Sendusu, Uganda.

Actively growing root tips (~1 cm long) were collected into 50-mM phosphate buffer (pH 7.0) containing 0.2% (v/v) β-mercaptoethanol, pre-treated in 0.05% (w/v) 8-hydroxyquinoline for three hours at room temperature, fixed in 3:1 ethanol:acetic acid fixative overnight, and stored in 70% ethanol. Preparation of protoplast suspensions was performed according to Doležel et al. (1998). Briefly, after digesting root tip segments in a mixture of 2% (w/v) cellulase and 2% (w/v) pectinase in 75-mM KCl and 7.5-mM EDTA (pH 4) for 90 min at 30°C, the suspension of resulting protoplasts was filtered through a 150-µm nylon mesh, pelleted, and washed in 70% ethanol. For further use, the protoplast suspension was stored in 70% ethanol at −20°C. Mitotic metaphase chromosome spreads were prepared by dropping method according to Doležel et al. (1998), the slides were postfixed in 4% (v/v) formaldehyde solution in 2x SSC solution and used for FISH.

Preparation of pachytene chromosome spreads was performed according to Mandáková and Lysák (2008), with minor modifications. Male flowers were fixed in 3:1 ethanol:acetic acid fixative overnight and stored in 70% ethanol at −20°C. Anthers were incubated in 0.3% (w/v) mix of cellulase, cytohelicase, and pectolyase (Sigma Aldrich, Darmstadt, Germany) for 30 min at 37°C. After the incubation in the enzyme mixture, the anthers were dissected in a drop of 60% (v/v) acetic acid on a microscopic slide and spread on the slide placed on a metal hot plate (50°C) after adding 60% (v/v) acetic acid for 25 s. The preparations were fixed in 3:1 ethanol:acetic acid fixative, air-dried, and used for FISH.

## Identification of Specific Oligomers and Labeling of the Oligo Probes

Oligomers specific for individual chromosome arms were identified in the reference genome sequence of *M. acuminata* "DH Pahang" v.2 (Martin et al., 2016) using Chorus pipeline (Han et al., 2015). Sets of 20,000 oligomers (45-mers) per one library were synthesized by Arbor Biosciences (Ann Arbor, Michigan, USA). Labeled oligomer probes were prepared according to Han et al. (2015). Briefly, the oligomer libraries were amplified using emulsion PCR (Murgha et al., 2014), where F primer contained T7 RNA polymerase promoter. The emulsified PCR product was washed with water-saturated diethyl ether and ethyl acetate and purified with QIAquick PCR purification kit (Qiagen, Hilden, Germany). The product (480 ng DNA) was used for T7 *in vitro* transcription with MEGAshortscript T7 Kit (ThermoFisher Scientific/Invitrogen, Waltham, Massachusetts, USA) at 37°C for 4 h. The RNA product was purified using RNeasy Mini Kit (Qiagen) and 42 µg of RNA was reverse-transcribed with either digoxigenin-, biotin-, or CY5-labeled R primer (Eurofins Genomics, Ebersberg, Germany) using Superscript II Reverse Transcriptase and SUPERase-In RNase inhibitor (ThermoFisher Scientific/Invitrogen). The RNA : DNA hybrids were cleaned with Quick-RNA MiniPrep Kit (Zymo Research, Freiburg im Breisgau, Germany) and hydrolyzed with RNase H (New England Biolabs, Ipswich, Massachusetts, USA) and finally with RNase A (ThermoFisher Scientific/Invitrogen). The products were purified with Quick-RNA MiniPrep Kit (Zymo Research) and eluted with nuclease-free water to obtain single-stranded labeled oligomers, which were used as FISH probes.

## Preparation of Other Cytogenetic Markers for FISH

Probes specific for ribosomal DNA sequences were prepared by labeling *Radka1* (part of 26S rRNA gene) and *Radka2* (contains 5S rRNA gene and non-transcribed spacer) DNA clones (Valárik et al., 2002) with biotin-16-dUTP (Roche Applied Science, Penzberg, Germany) or aminoallyl-dUTP-CY5 (Jena Biosciences, Jena, Germany) by PCR using T3 (forward) and T7 (reverse) primers (Invitrogen). Probes for tandem repeats CL18 and CL33 (Hřibová et al., 2010) were amplified using specific primers and labeled with aminoallyl-dUTP-CY5 or fluorescein-12-dUTP (Jena Biosciences, Jena, Germany) by PCR according to Čížková et al. (2013). Single copy BAC clone 2G17 (Hřibová et al., 2008) was labeled by digoxigenin-11-dUTP nick translation following manufacturer's recommendation (Roche Applied Science, Penzberg, Germany).

## Fluorescence *In Situ* Hybridization and Image Analysis

Hybridization mix (30 µl) containing 50% (v/v) formamide, 10% (w/v) dextran sulfate in 2x SSC and 10 ng/µl of labeled probe was added onto slide and denatured for 3 min at 80°C. Hybridization was carried out overnight at 37°C. The sites of

hybridization of digoxigenin- and biotin-labeled probes were detected using anti-digoxigenin-FITC (Roche Applied Science) and streptavidin-Cy3 (ThermoFisher Scientific/Invitrogen), respectively. Chromosomes were counterstained with DAPI and mounted in VECTASHIELD Antifade Mounting Medium (Vector Laboratories, Burlingame, CA, USA). The slides were examined with Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) equipped with Cool Cube 1 camera (Metasystems, Altlussheim, Germany) and appropriate optical filters. The capture of fluorescence signals, merging the layers, and measurement of chromosome length were performed with ISIS software 5.4.7 (Metasystems), the final image adjustment and creation of idiograms were done in Adobe Photoshop CS5.

# RESULTS

## Development of Chromosome Painting Probes and *In Situ* Hybridization

In order to produce chromosome arm-specific painting probes, unique *k*- mers were identified according to Han et al. (2015) in the reference genome sequence of the doubled haploid banana (*M. acuminata* "DH Pahang"; Martin et al., 2016) and analyzed with the Chorus program (https://github.com/forrestzhang/Chorus). While eight pseudomolecules corresponded to metacentric chromosomes, pseudomolecules 1, 2, and 10 appeared to be acrocentric with peri-centromeric region occupying an entire chromosome arm. The density of unique oligomers was lower in peri-centromeric regions in all pseudomolecules (**Supplementary Figure S1**) and these regions were excluded from the selection of oligomers for painting probes. The number of unique oligomers ranged from 79,896 to 127,835 for pseudomolecules 2 and 6, respectively. Sets of 20,000 45-mers specific to individual chromosome arms were then selected in Chorus, synthesized as so called immortal libraries and labeled directly by Cy5 or indirectly by biotin or digoxigenin as described in Materials and Methods. Oligomer libraries were designed to achieve a density of 0.9 to 2.1 oligomers per 1-kb chromosome sequence (**Supplementary Table S1**). To confirm that it is not possible to paint peri-centromeric regions with low oligomer densities and large gaps between low copy oligomers, a painting probe was prepared from peri-centromeric region of pseudomolecule 3. In total, 8,317 oligomers spanning this region (~10.5 Mb long) ensured an average density of ~0.8 oligomers/kb.

First, the painting probes were hybridized to mitotic metaphase chromosomes spreads of *M. acuminata* ssp. *malaccensis* (A genome)—the genotype from which the *Musa* reference genome sequence was developed. FISH with the painting probes resulted in visible signals covering chromosome arms along their lengths (**Figures 1A–F**). This observation confirmed that the probes had the expected parameters. Moreover, because the painting highlighted individual chromosome arms, it was possible to anchor pseudomolecules to individual chromosome arms. This work revealed that in the assembly, pseudomolecules 1, 6, 7 start with long arms and end with short arms, i.e., they are oriented inversely to the way karyotypes are presented, where the short arm of the chromosome is on top and the long arm on the bottom.

Following this, the painting probes were used for FISH in *M. balbisiana* (B genome) and *M. schizocarpa* (S genome) (**Figures 1B**, **G**, **H**). Comparison of chromosome and/or chromosome-arm painting in *M. acuminata* ssp. *malaccensis* and *M. schizocarpa* did not reveal any large chromosome translocations differentiating both species. On the other hand, a large translocation of the long arm of chromosome 3 to long (painted) arm of chromosome 1 was found in *M. balbisiana* (**Figure 1B**).

The small size of condensed mitotic metaphase chromosomes reduces the longitudinal resolution of chromosome painting and hence a chance to discover small structural rearrangements. An alternative is to perform chromosome painting with meiotic pachytene chromosomes (**Figure 2**) which are approximately fifty times longer. When hybridized to pachytene chromosome spreads of *M. acuminata* ssp. *malaccensis*, painting probes provided visible signals and the opportunity to analyze chromosome structure in more detail. This experiment showed that banana chromosomes do not contain large blocks of heterochromatin in distal and subtelomeric regions (**Figure 2**). Taking the advantage of higher spatial resolution, the set of oligo painting probes developed in this work will be suitable to visualize meiotic processes such as crossing over and synapsis. Following this, pachytene chromosome spreads of *M. acuminata* ssp. *malaccensis* were used to evaluate the signal of peri-centromeric painting probe designed for chromosome 3. FISH with the probe did not result in a continuous signal along the whole region. Instead, discontinuous signals, with signal-free gaps along most of the (peri-)centromeric region of chromosome 3 (**Figure 2B**), were observed. Based on this observation, painting probes were not designed for (peri-)centromeric regions of the remaining ten banana chromosomes.

## Integration of Cytogenetic Landmarks and Oligopaints

In order to utilize the existing probes for FISH in *Musa* and develop a highly informative toolbox to characterize *Musa* chromosome structure, the existing cytogenetic landmarks were integrated with the painting probes.

45S rRNA genes mapped to secondary constriction located on non-painted arm of chromosome 10 in all three *Musa* species. The probe for 5S rRNA genes localized to different chromosome regions and on different chromosomes in the three *Musa* species studied. In *M. acuminata* ssp. *malaccensis*, six signals of 5S rDNA were observed on mitotic metaphase plates and were localized in sub-telomeric region of chromosome 1 and long arm of chromosome 8, and in peri-centromeric region on short arm of chromosome 3. Six hybridization signals with 5S rDNA probe were observed also in mitotic metaphase plate of *M. balbisiana*. Two pairs of strong signals were localized in sub-telomeric region of chromosome 2 (non-painted arm) and in peri-centromeric region of the long arm of chromosome 3. Additional weak signal was observed in peri-centromeric region of the long arm of chromosome 6 (**Figures 3** and **4**). In *M. schizocarpa*, three pairs of strong signals and three pairs of weaker signals were observed after FISH with 5S rDNA probe on mitotic metaphase plate. Sub-telomeric region of chromosome 1 (non-painted arm) and peri-centromeric region of short arm of chromosome 3 and long arm of chromosome 4

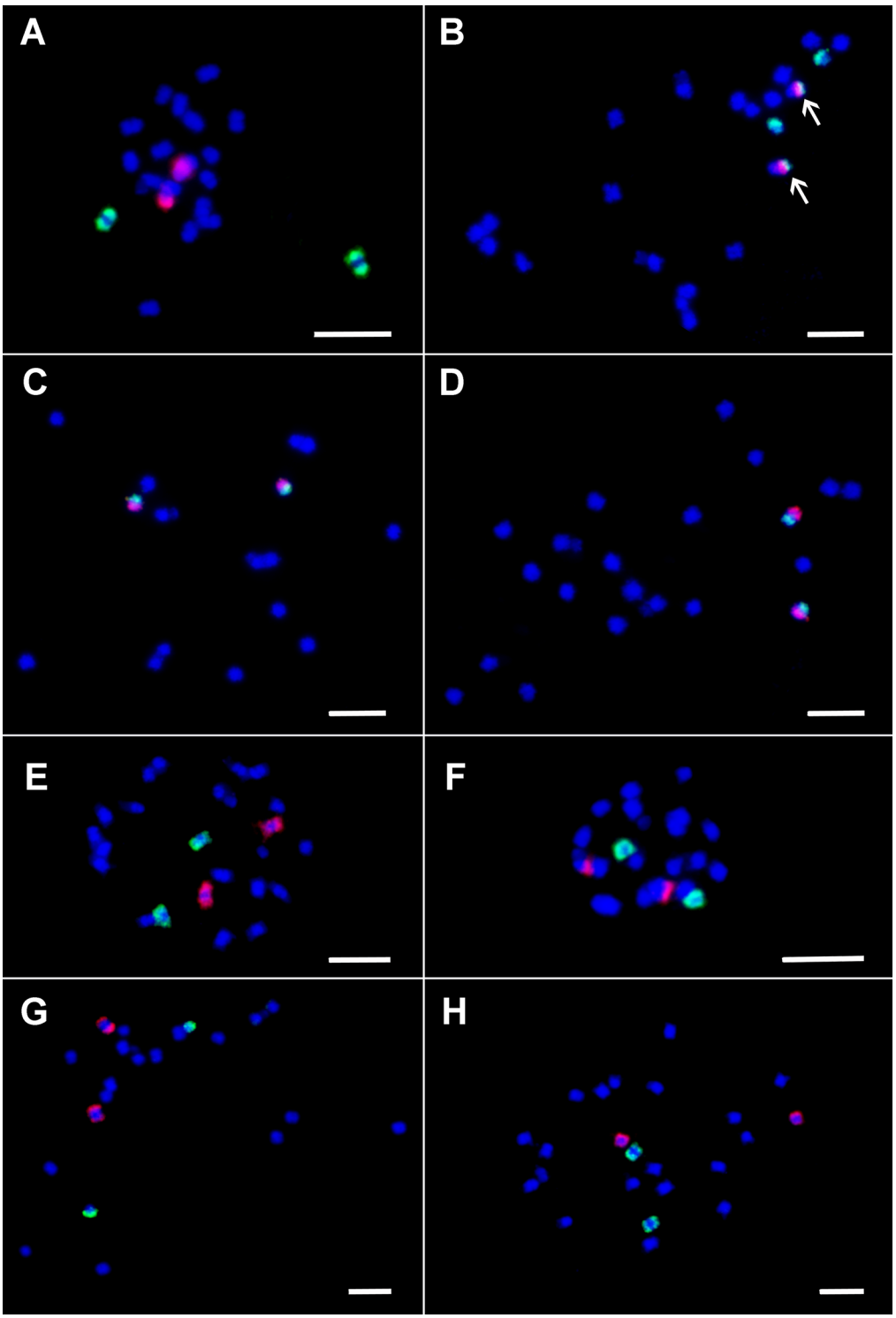



**FIGURE 1 |** Oligo painting FISH on mitotic metaphase plates of three species of *Musa*. **(A)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; chromosome 1 in red, chromosome 3 in green). **(B)** *M. balbisiana* "Tani" (2n = 22, BB; chromosome 1 in red, chromosome 3 in green). **(C)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; short arm of chromosome 4 in green, its long arm in red). **(D)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; short arm of chromosome 5 in red, its long arm in green). **(E)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; chromosome 6 in red, chromosome 7 in green. **(F)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; chromosome 10 in red, chromosome 11 in green. **(G)** *Musa schizocarpa* "Schizocarpa" (2n = 2x = 22, SS; chromosome 8 in red, chromosome 2 in green). **(H)** *Musa schizocarpa* "Schizocarpa" (2n = 2x = 22, SS; chromosome 11 in red, chromosome 9 in green). Chromosomes were counterstained with DAPI (blue). Arrows point to the region of chromosome 3 translocated to chromosome 1 in *M. balbisiana*. Bars = 5 μm.

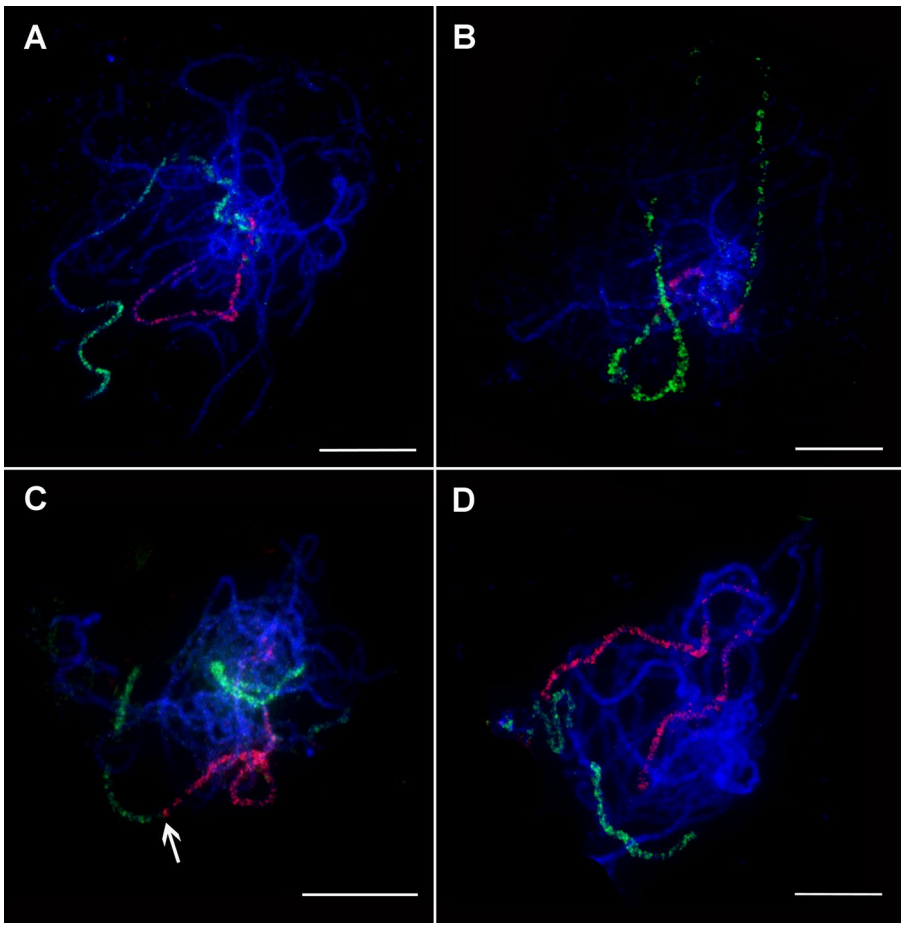



**FIGURE 2 |** Oligo painting FISH on meiotic pachytene chromosome spreads of *Musa*. **(A)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; chromosome 1 in red, chromosome 4 in green). **(B)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; (peri-)centromeric region in red, chromosome 3 in green). **(C)** *M. balbisiana* "Tani" (2n = 22, BB; chromosome 1 in red, chromosome 3 in green). **(D)** *M. balbisiana* "Tani" (2n = 22, BB; chromosome 5 in red, chromosome 11 in green). Chromosomes were counterstained with DAPI (blue). Arrows point to the region translocated from chromosome 3 to chromosome 1 in *M. balbisiana*. Bars = 10 µm.

contained strong signals of 5S rDNA. Additional weaker signals of 5S rDNA probe were observed in peri-centromeric regions of short arm of chromosome 8 and short arm of chromosome 11, as well as on the non-painted arm of chromosome 10 (**Figures 3** and **4**).

Tandem organized repeats CL18, CL33, and BAC clone 2G17 were localized on non-painted arm of chromosome 1 in all three *Musa* species, except satellite CL33, which was not detected on any chromosome in *M. balbisiana*. In contrast, additional signal of tandem repeat CL33 located on non-painted arm of chromosome 2 was observed in *M. acuminata* ssp. *malaccensis* and in *M. schizocarpa* (**Figures 3** and **4**). Finally, additional signal of tandem repeat CL18 was observed in *M. balbisiana* on the non-painted arm of chromosome 2.

## DISCUSSION

Until recently, chromosome painting could be used only in plants whose genomes were sequenced clone by clone (The Arabidopsis Genome Initiative, 2000; The International Brachypodium Initiative, 2010) and chromosome painting was achieved by FISH with pools of single copy BAC clones that covered entire chromosomes. Importantly, chromosome paints developed in one species could be used in related species, providing a powerful approach for comparative karyotype analysis and for tracing karyotype changes during the evolution and speciation (e.g., Lysák et al., 2001; Idziak et al., 2014). Unfortunately, this painting method cannot be used in species with large genomes due to the prevalence of repetitive DNA and in species not closely related to those for which painting using BAC pools was developed.

The progress in DNA sequencing technology and assembly algorithms resulted in a shift from the clone by clone sequencing to shotgun sequencing and a majority of plant genomes has been sequenced in this way (Hamilton and Buell, 2012; Zimin et al., 2017; Belser et al., 2018). The availability of reference genome sequences and the affordable cost of synthesizing short oligonucleotides offered a direct way to develop chromosome paints (Han et al., 2015). Here, thousands of short single copy sequences are identified, bulk synthesized, fluorescently labeled and used as probes for FISH (Han et al., 2015). Pools of labeled

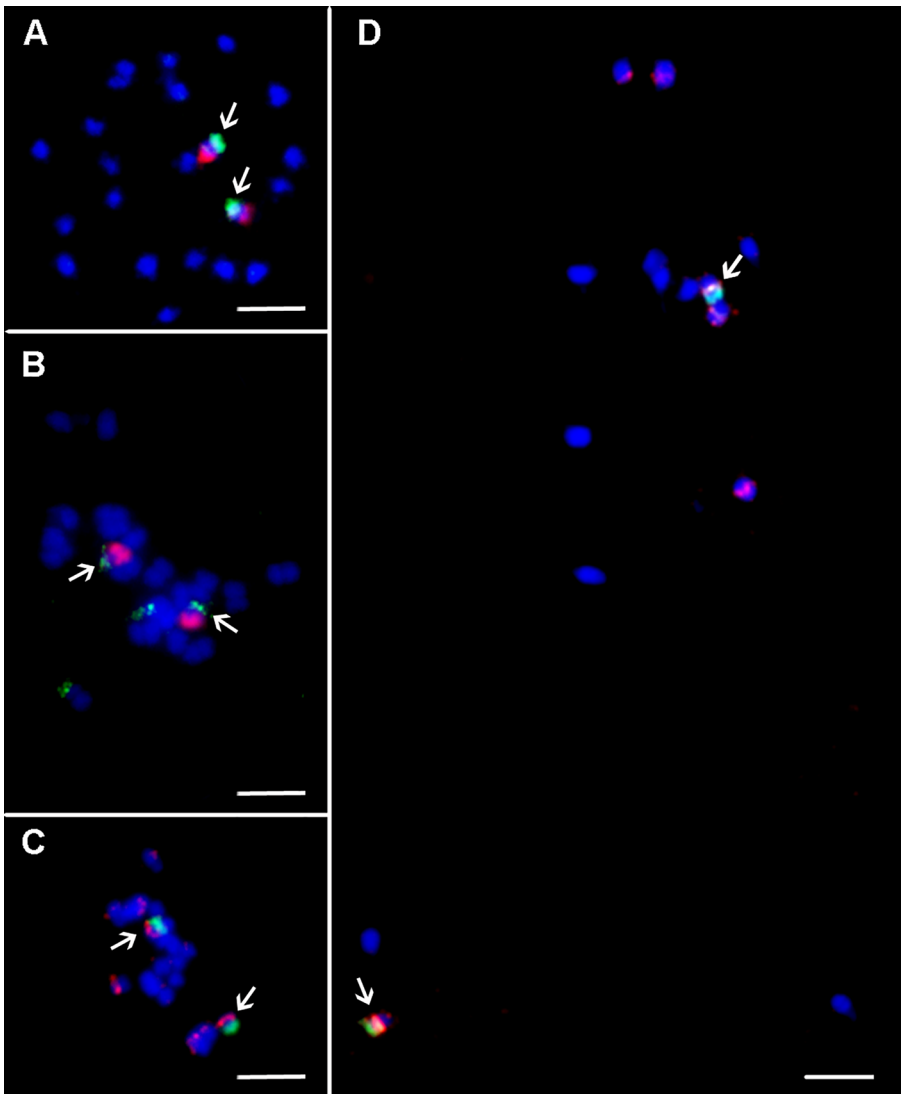



**FIGURE 3 |** Integration of oligo painting FISH and existing cytogenetic markers on mitotic metaphase plates of *Musa*. **(A)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; chromosome 1 in red, BAC clone 2G17 in green). **(B)** *M. schizocarpa* "Schizocarpa" (2n = 22, SS; chromosome 2 in red, tandem repeat CL33 in green). **(C)** *M. schizocarpa* "Schizocarpa" (2n = 22, SS; short arm of chromosome 4 in green, 5S rRNA in red—two loci are localized on long arm of chromosome 4). **(D)** *M. acuminata* ssp. *malaccensis* "Pahang" (2n = 22, AA; 5S rRNA in red, short arm of chromosome 3 in green bears 5S rRNA). Chromosomes were counterstained with DAPI (blue). Bars = 5 µm. Arrows indicate colocalization of oligo painting FISH probes with existing cytogenetic markers.

oligonucleotide probes were found suitable for FISH on somatic metaphase chromosomes, meiotic pachytene chromosomes and interphase nuclei (Han et al., 2015; Filiault et al., 2018; Xin et al., 2018; Albert et al., 2019; Jiang, 2019). Successful applications of oligo painting FISH include construction of molecular cytogenetic karyotypes (Braz et al., 2018; Qu et al., 2017; Meng et al., 2018), identification of large chromosomal rearrangements, analysis of chromosome pairing in meiosis (Han et al., 2015; He et al., 2018; Albert et al., 2019), as well as the visualization of the arrangement of chromosomes in 3D space of interphase nuclei (Albert et al., 2019).

Despite the availability of a reference genome sequence of *M. acuminata ssp. malaccensis* (D'Hont et al., 2012; Martin et al., 2016) and resequencing of more than 120 accessions of *Musa* (Dupouy et al., 2019), DNA pseudomolecules have not been anchored to individual chromosomes and molecular karyotype of *Musa* has not been developed to date. In many plant species, tandem organized repeats serve as useful probes for FISH to identify individual chromosomes and their regions (Hřibová et al., 2007; Badaeva et al., 2015; Koo et al., 2016; Křivánková et al., 2017; Said et al., 2018). The nuclear genome of *Musa* species is relatively small (1C ~ 500–750 Mb; Doležel et al., 1994; Lysák et al., 1999; Asif et al., 2001; Kamaté et al., 2001; Bartoš et al., 2005; Čížková et al., 2015) and until now, only a few tandem organized repeats and rDNA sequences were successfully used as cytogenetic landmarks (Balint-Kurti et al., 2000; Valárik et al., 2002; Hřibová et al., 2007; Hřibová et al., 2010; Čížková et al., 2013; Novák et al., 2014). Moreover, only

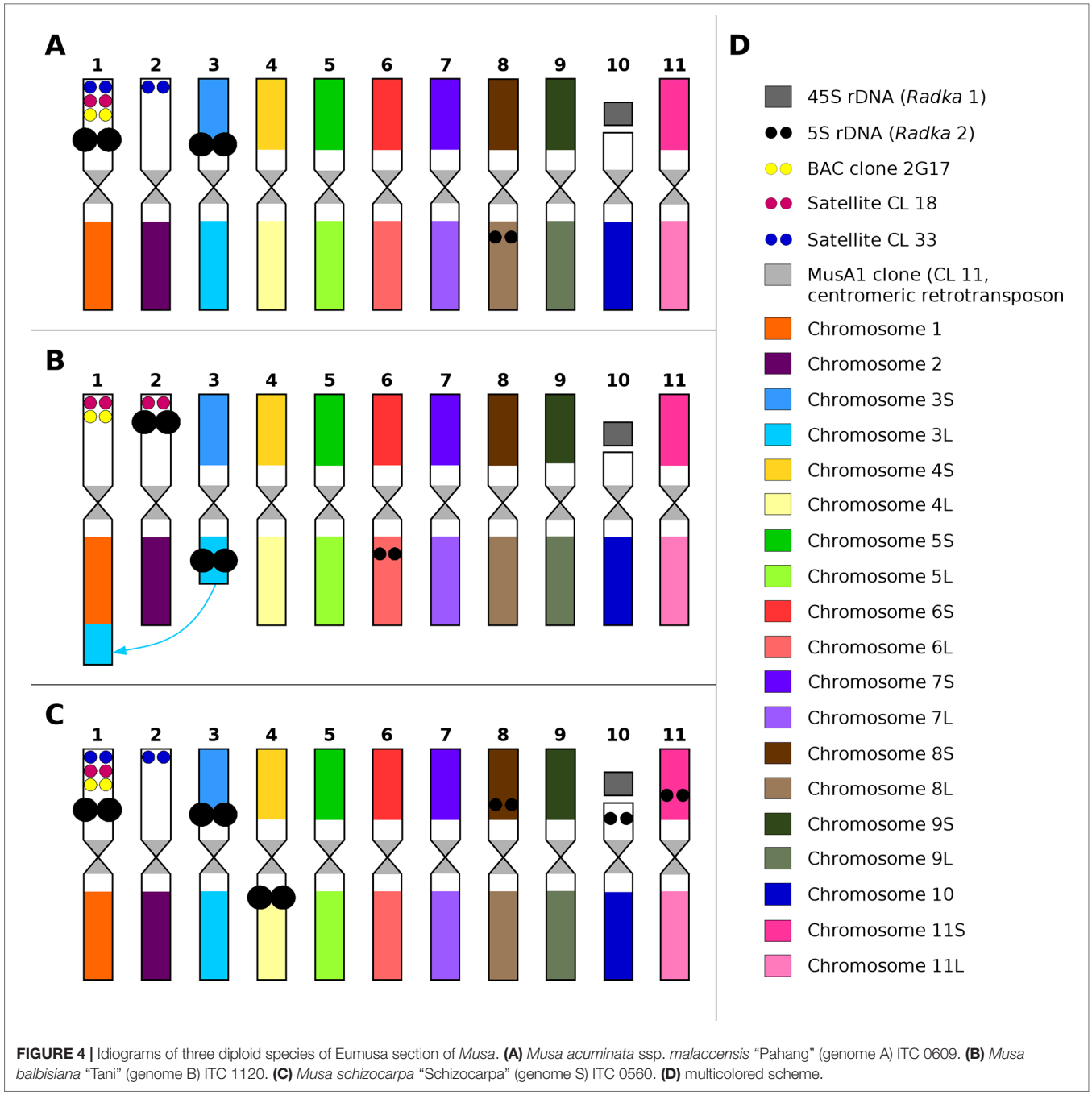



**FIGURE 4 |** Idiograms of three diploid species of Eumusa section of *Musa*. **(A)** *Musa acuminata* ssp. *malaccensis* "Pahang" (genome A) ITC 0609. **(B)** *Musa balbisiana* "Tani" (genome B) ITC 1120. **(C)** *Musa schizocarpa* "Schizocarpa" (genome S) ITC 0560. **(D)** multicolored scheme.

one BAC clone has been used as a cytogenetic marker in *Musa* (Hřibová et al., 2008) and only four BAC clones were localized on pachytene chromosomes (De Capdeville et al., 2009). Thus, the attempts to use of BAC clones for anchoring pseudomolecules to chromosomes in banana as has been done in other species (Jiang et al., 1995; Lapitan et al., 1997; Kim et al., 2002; Idziak et al., 2014), were not successful.

Unlike the previous approaches, chromosome painting using pools of single copy oligomers offers the opportunity to establish a molecular karyotype of *Musa*, making it possible to identify individual chromosomes, follow their behavior during somatic cell cycle and meiosis, perform comparative karyotype analysis, and identify structural chromosome changes. FISH with oligo painting probes developed in this work resulted in visible hybridization signals along chromosomal arms on condensed mitotic metaphase chromosomes (**Figure 1**) as well as on less condensed pachytene chromosomes (**Figure 2**) confirming their usefulness as painting probes in *Musa*.

Only small regions on pachytene chromosomes were free of painting signals. This could be either due to the presence of heterochromatin blocks, or due to gaps in the genome sequence (**Figure 2**). In contrast to chromosome arms, peri-centromeric regions were not labeled. These regions contain large gaps in the genome sequence and large proportion of repetitive DNA sequences in peri-centromeric regions (Hřibová et al., 2010; Neumann et al., 2011; D'Hont et al., 2012; Martin et al., 2016).

We demonstrate that chromosome/chromosome-arm specific oligo painting libraries designed for *M. acuminata* ssp. *malaccensis* can be used for cytogenetic analysis of related species *M. balbisiana* and *M. schizocarpa*, which played an important role in the evolution of many edible banana clones (Carreel et al., 1994; D'Hont et al., 2012; Davey et al., 2013; Čížková et al., 2013). This observation provided an opportunity for comparative karyotype analysis and identification of putative chromosome translocations. In our study, we observed translocation of long arm of chromosome 3 to long arm of chromosome 1 in *M. balbisiana* (B genome) (**Figures 1B** and **2C**). This observation confirms the result of Baurens et al. (2019), which were obtained after anchoring a dense genetic map of *M. balbisiana* "Pisang Klutuk Wulung" to *M. acuminata* ssp. *malaccensis* reference genome sequence (Martin et al., 2016). The authors estimated the size of the translocated region of long arm of chromosome 3 to be ~8 Mb, confirming the sensitivity of oligo chromosome painting.

Co-localization of chromosome painting probes with cytogenetic markers developed earlier for *Musa* (Valárik et al., 2002; Hřibová et al., 2010; Čížková et al., 2013) offered an opportunity to create molecular karyotypes suitable for comparative analysis. The presence of 5S rRNA genes on non-collinear chromosomes in the A, B, and S genomes of *Musa* as described here indicates small chromosomal rearrangements which occurred during *Musa* speciation. On the other hand, the location of tandem organized repeats CL18, CL33, and BAC clone 2G17 on collinear chromosome arms in all three species indicates their structural homology of the chromosome arms. These observations imply that chromosomes containing a particular DNA sequence, e.g., 5S rDNA, cannot be considered as collinear. This shows a potential weakness of comparative karyotype analysis of using only a few cytogenetic markers (Fukui et al., 1994; Murata et al., 1997).

Tandem organized repeats CL18 and CL33 (Hřibová et al., 2010) were located together with 5S rRNA genes on short arms of chromosomes 1 and 2, which lacked oligopainting signals. Genome sequence of *M. acuminata* ssp. *malaccensis* includes three pseudomolecules which are represented by two large regions differing in DNA repeat composition and in density of unique oligomers (**Supplementary Figure 1**, Martin et al., 2016). The constitution of banana pseudomolecules 1, 2, and 10 indicates that they cover only one chromosome arm and a peri-centromeric region. Painting probes created for the three pseudomolecules localized to only one chromosomal arm. One of the pseudomolecules is collinear with acrocentric chromosome 10 and bears 45S rRNA locus on its short arm. The two remaining pseudomolecules represent chromosomes 1 and 2, which seem to be meta or sub-metacentric thus could miss a large sequence region. These observations indicate that these genomic regions were not completely assembled and are missing due to the presence of a large number of various tandem organized sequences.

The improved version of *M. acuminata* "DH Pahang" reference genome sequence represents 450.7 Mbp which corresponds to ~81% of its nuclear genome size estimated by flow cytometry (Čížková et al., 2013). In addition, the reference genome sequence contains a total of 56.6-Mbp sequences, which were not anchored to the 11 pseudomolecules. The most plausible explanation why these sequences were not included in pseudomolecules is that they represent heterochromatin regions, which are difficult to sequence. However, relatively high number of unique oligomers in unanchored scaffolds as observed in this work (**Supplementary Figure 1**) indicates that the unanchored part of the reference genome sequence contains low copy sequences from euchromatic regions. Thus, these regions were probably not anchored due to the absence of DNA markers, or they were too short to be anchored using Bionano optical mapping. The use of long-read sequencing technologies such as Oxford Nanopore in combination with optical mapping (Belser et al., 2018) should further improve the current assembly and shed light on the difficult parts of *M. acuminata* ssp. *malaccensis* genome.

## CONCLUSIONS

In this work, chromosome painting probes were developed for banana (*Musa* spp.) and used to establish molecular karyotypes for three species of *Musa* that were the parents of a majority of cultivated edible banana clones. This advance made it possible to anchor reference genome sequence of banana, *Musa acuminata* spp. *malaccensis* to individual chromosomes. The study also demonstrates the potential of oligo painting FISH for comparative karyotype analysis and identification of structural chromosome changes that accompanied the evolution and speciation in the genus *Musa*.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/ **Supplementary Material**.

## AUTHOR CONTRIBUTIONS

EH and JD conceived the experiments. DŠ, AN, and MK conducted the study and processed the data. BU and RS provided the banana materials. DŠ and EH wrote the manuscript. EH, JD, and DŠ discussed the results and contributed to

manuscript writing. All authors have read and approved the final manuscript.

## ACKNOWLEDGMENTS

We thank Dr. Ines van den Houwe for providing the plant material and Ms. Radka Tušková for excellent technical assistance. This work was supported by the Czech Science Foundation (award No. 19-20303S). The computing was supported by the National Grid Infrastructure MetaCentrum (grant No. LM2010005 under the program Projects of Large Infrastructure for Research, Development, and Innovations). The authors thank all donors who supported this work through their contributions to the CGIAR Fund (http://www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2/), and in particular to the CGIAR Research Program Roots, Tubers and Bananas (RTB-CRP).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01503/full#supplementary-material

**SUPPLEMENTARY FIGURE S1 |** Oligomer coverage of 11 pseudomolecules (labeled **A** – **K**) and concatenated unanchored scaffolds **(L)** in the reference genome of *M. acuminata* 'DH Pahang' (Martin et al., 2016). The oligomers (45 bp) were designed using the Chorus program (Han et al., 2015) and are depicted in black. Position and coverage of tandem repeats CL18 (pink) and CL33 (green) are also shown.

## REFERENCES

Albert, P. S., Zhang, T., Semrau, K., Rouillard, J. M., Kao, Y. H., and Wang, C. R. (2019). Whole-chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal relationships. *Proc. Natl. Acad. Sci. U. S. A.* 116 (5), 1679–1685. doi: 10.1073/pnas.1813957116

Amosova, A. V., Bolsheva, N. L., Zoshchuk, S. A., Twardovska, M. O., Yurkevich, O. Y., Andreev, I. O., et al. (2017). (A) Comparative molecular cytogenetic characterization of seven *Deschampsia (Poaceae)* species. *PloS One* 12 (4), e0175760. doi: 10.1371/journal.pone.0175760

Argent, G. C. G. (1976). The wild bananas of Papua New Guinea. *Notes R. Bot. Gard. Edinburgh.* 35 (1), 77–114.

Asif, M. J., Mak, C., and Othman, R. Y. (2001). Characterization of indigenous *Musa* species based on flow cytometric analysis of ploidy and nuclear DNA content. *Caryologia.* 54 (2), 161–168. doi: 10.1080/00087114.2001.10589223

Badaeva, E. D., Amosova, A. V., Goncharov, N. P., Macas, J., Ruban, A. S., Grechishnikova, I. V., et al. (2015). A set of cytogenetic markers allows the precise identification of all A-genome chromosomes in diploid and polyploid wheat. *Cytogenet. Genome Res.* 146 (1), 71–79. doi: 10.1159/000433458

Balint-Kurti, P., Clendennen, S., Doleželová, M., Valárik, M., Doležel, J., Beetham, P. R., et al. (2000). Identification and chromosomal localization of the *monkey* retrotransposon in *Musa* sp. *Mol. Gen. Genet.* 263 (6), 908–915. doi: 10.1007/s004380000265

Bartoš, J., Alkhimova, O., Doleželová, M., De Langhe, E., and Doležel, J. (2005). Nuclear genome size and genomic distribution of ribosomal DNA in *Musa* and *Ensete* (*Musaceae*): taxonomic implications. *Cytogenet. Genome Res.* 109, 50–57. doi: 10.1159/000082381

Baurens, F. C., Martin, G., Hervouet, C., Salmon, F., Yohomé, D., and Ricci, S. (2019). Recombination and large structural variations shape interspecific edible bananas genomes. *Mol. Biol. Evol.* 36 (1), 97–111. doi: 10.1093/molbev/msy199

Belser, C., Istace, B., Denis, E., Dubarry, M., Baurens, F. C., and Falentin, C. (2018). Chromosome-scale assemblies of plant genomes using nanopore long reads and optical maps. *Nat. Plants.* 4 (11), 879–887. doi: 10.1038/s41477-018-0289-4

Braz, G. T., He, L., Zhao, H., Zhang, T., Semrau, K., and Rouillard, J. M. (2018). Comparative oligo-FISH mapping: An efficient and powerful methodology to reveal karyotypic and chromosomal evolution. *Genetics.* 208 (2), 513–523. doi: 10.1534/genetics.117.300344

Brown, A., Tumuhimbise, R., Amah, D., Uwimana, B., Nyine, M., and Mduma, H. (2017). *The genetic improvement of bananas and plantains (Musa spp.) In Genetic Improvement of Tropical Crops* Vol. pp. Campos, HCaligari, PDS. Cham: Springer, 219–240.

Carreel, F., Fauré, S., González de León, D., Lagoda, P. J. L., Perrier, X., and Bakry, F. (1994). Evaluation of the genetic diversity in diploid bananas (*Musa* sp.). *Genet. Sel. Evol.* 26 (Suppl 1), 125–136. doi: 10.1051/gse:19940709

Cheesman, E. E. (1947). Classification of the bananas. The genus *Ensete* Horan and the genus *Musa* L. Kew. *Bull.* 2 (2), 97–117. doi: 10.2307/4109206

and the genus Musa L. Kew Bull. 2(2):97-117

Čížková, J., Hřibová, E., Humplíková, L., Christelová, P., Suchánková, P., Doležel J. (2013). Molecular analysis and genomic organization of major DNA satellites in banana (Musa spp.). PLoS ONE. 8:e54808. doi: 10.1371/journal.pone.0054808

Čížková, J., Hřibová, E., Christelová, P., Van den Houwe, I., Häkkinen, M., Roux, N., et al. (2015). Molecular and cytogenetic characterization of wild *Musa* species. *PLoS One* 10:e0134096. doi: 10.1371/journal.pone.0134096

Cremer, T., and Cremer, S. (2001). (B) Chromosome territories, nuclear architecture and gene regulation in mammalian cells.. *Nat. Rev. Genet.* 2 (4), 292–301. doi: 10.1038/35066075

D'Hont, A., Paget-Goy, A., Escoute, J., and Carreel, F. (2000). The interspecific genome structure of cultivated banana, *Musa* spp. revealed by genomic DNA *in situ* hybridization. *Theor. Appl. Genet.* 100, 177–183. doi: 10.1007/s001220050024

D'Hont, A., Denoeud, F., Aury, J. M., Baurens, F. C., Carreel, F., Garsmeur, O., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488, 213–217. doi: 10.1038/nature11241

Danilova, T. V., Friebe, B., and Gill, B. S. (2014). Development of wheat single gene FISH map for analyzing homoeologous relationship and chromosomal rearrangements within the triticeae. *Theor. Appl. Genet.* 127 (3), 715–730. doi: 10.1007/s00122-013-2253-z

Davey, M. W., Gudimella, R., Harikrishna, J. A., Sin, L. W., Khalid, N., and Keulemans, J. (2013). A draft *Musa balbisiana* genome sequence for molecular genetics in polyploid, inter- and intra-specific *Musa* hybrids. *BMC Genomics* 14, 683. doi: 10.1186/1471-2164-14-683

De Capdeville, G., Souza Junior, M. T., Szinay, D., Diniz, L. E. C., Wijnker, E., Swennen, R., et al. (2009). The potential of high-resolution BAC-FISH in banana breeding. *Euphytica* 166, 431–443. doi: 10.1007/s10681-008-9830-2

Doležel, J., Doleželová, M., and Novák, F. J. (1994). Flow cytometric estimation of nuclear DNA amount in diploid bananas (*Musa acuminata* and *Musa balbisiana*). *Biol. Plant* 36, 351–357. doi: 10.1007/BF02920930

Doležel, J., Doleželová, M., Roux, N., and Van den houwe, I. (1998). A novel method to prepare slides for high resolution chromosome studies in *Musa* spp. *Infomusa* 7, 3–4.

Doleželová, M., Valárik, M., Swennen, R., Horry, J. P., and Doležel, J. (1998). Physical mapping of the 18S-25S and 5S ribosomal RNA genes in diploid bananas. *Biol. Plant* 41, 497–505. doi: 10.1023/A:1001880030275

Dupouy, M., Baurens, F. C., Derouault, P., Hervouet, C., Cardi, C., Cruaud, C., et al. (2019). Two large reciprocal translocations characterized in the disease resistance-rich *burmannica* genetic group of *Musa acuminata*. *Ann. Bot.* XX, 1–11. doi: 10.1093/aob/mcz078

Ferguson-Smith, M. A., and Trifonov, V. (2007). Mammalian karyotype evolution. *Nat. Rev. Genet.* 8 (12), 950–962. doi: 10.1038/nrg2199

Filiault, D. L., Ballerini, E. S., Mandáková, T., Aköz, G., Derieg, N. J., Schmutz, J., et al. (2018). The *Aquilegia* genome provides insight into adaptive radiation and reveals an extraordinarily polymorphic chromosome with a unique history. *eLife* 7, e36426. doi: 10.7554/eLife.36426

Fukui, K., Kamisugi, Y., and Sakai, F. (1994). Physical mapping of 5S rDNA loci by direct-cloned biotinylated probes in barley chromosomes. *Genome.* 37 (1), 105–111. doi: 10.1139/g94-013

Gill, B. S., and Kimber, G. (1977). Recognition of translocations and alien chromosome transfers in wheat by the Giemsa C-banding technique. *Crop Sci.* 17, 264–266. doi: 10.2135/cropsci1977.0011183X001700020008x

Gill, B. S., Friebe, B., and Endo, T. R. (1991). Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome* 34, 830–839. doi: 10.1139/g91-128

Greilhuber, J. (1977). Why plant chromosomes do not show G-bands. *Theor. Appl. Genet.* 50 (3), 121–124. doi: 10.1007/BF00276805

Häkkinen, M. (2013). Reappraisal of sectional taxonomy in *Musa* (*Musaceae*). *Taxon.* 62 (4), 809–813. doi: 10.12705/624.3

Hamilton, J. P., and Buell, C. R. (2012). Advances in plant genome sequencing. *Plant J.* 70 (1), 177–190. doi: 10.1111/j.1365-313X.2012.04894.x

Han, Y., Zhang, T., Thammapichai, P., Weng, Y., and Jiang, J. (2015). Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. *Genetics.* 200, 771–779. doi: 10.1534/genetics.115.177642

He, L., Braz, G. T., Torres, G. A., and Jiang, J. M. (2018). Chromosome painting in meiosis reveals pairing of specific chromosomes in polyploid *Solanum* species. *Chromosoma* 127, 505–513. doi: 10.1007/s00412-018-0682-9

Hou, L., Xu, M., Zhang, T., Xu, Z., Wang, W., Zhang, J., et al. (2018). *BMC Plant Biol.* 18 (1), 110. doi: 10.1186/s12870-018-1325-2

Hřibová, E., Doleželová, M., Town, C. D., Macas, J., and Doležel, J. (2007). Isolation and characterization of the highly repeated fraction of the banana genome. *Cytogenet. Genome Res.* 119 (3-4), 268–274. doi: 10.1159/000112073

Hřibová, E., Doleželová, M., and Doležel, J. (2008). Localization of BAC clones on mitotic chromosomes of *Musa acuminata* using fluorescence *in situ* hybridization. *Biol. Plant* 52, 445–452. doi: 10.1007/s10535-008-0089-1

Hřibová, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J., and Doležel, J. (2010). Repetitive part of the banana (*Musa acuminata*) genome investigated by low-depth 454 sequencing. *BMC Plant Biol.* 10, 204. doi: 10.1186/1471-2229-10-204

Idziak, D., Hazuka, I., Poliwczak, B., Wiszynska, A., Wolny, E., and Hasterok, R. (2014). Insight into the karyotype evolution of *Brachypodium* species using comparative chromosome barcoding. *PloS One* 9 (3), e93503. doi: 10.1371/journal.pone.0093503

International Plant Genetic Resources Institute-International Network for the Improvement of Banana and Plantain/Centre de Coopération internationale en recherche agronomique pour le développement [IPGRI-INIBAP/CIRAD] International Plant Genetic Resources Institute-International Network for the Improvement of Banana and Plantain/Centre de Coopération internationale en recherche agronomique pour le développement [IPGRI-INIBAP/CIRAD] (1996). Description for Banana (*Musa* spp.). Int. Network for the Improvement of Banana and Plantain, Montpellier, France; Centre de coopération int. en recherche agronomique pour le développement, Montpellier, France; International Plant Genetic Resources Institute Press, Rome.

Janssens, S. B., Vandelook, F., De Langhe, E., Verstraete, B., Smets, E., Van den Houwe, I., et al. (2016). Evolutionary dynamics and biogeography of *Musaceae* reveal a correlation between the diversification of the banana family and the geological and climatic history of Southeast Asia. *New Phytol.* 210 (4), 1453–1465. doi: 10.1111/nph.13856

Jiang, J., Gill, B. S., Wang, G. L., Ronald, P. C., and Ward, D. C. (1995). Metaphase and interphase fluorescence *in situ* hybridization mapping of the rice genome with bacterial artificial chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 92 (10), 4487–4491. doi: 10.1073/pnas.92.10.4487

Jiang, J. (2019). Fluorescence *in situ* hybridization in plants: recent developments and future applications. *Chromosom. Res.* 27 (3), 153–165. doi: 10.1007/s10577-019-09607-z

Křivánková, A., Kopecký, D., Stočes, Š., Doležel, J., and Hřibová, E. (2017). Repetitive DNA: a versatile tool for karyotyping in *Festuca pratensis* Huds. *Cytogenet. Genome Res.* 151 (2), 96–105. doi: 10.1159/000462915

Kamaté, K., Brown, S., Durand, P., Bureau, J. M., De Nay, D., and Trinh, T. H. (2001). Nuclear DNA content and base composition in 28 taxa of *Musa*. *Genome.* 44, 622–627. doi: 10.1139/g01-058

Kim, J. S., Childs, K. L., Islam-Faridi, M. N., Menz, M. A., Klein, R. R., and Klein, P. E. (2002). Integrated karyotyping of sorghum by *in situ* hybridization of landed BACs. *Genome.* 45, 402–412. doi: 10.1139/g01-141

Koo, D. H., Zhao, H., and Jiang, J. (2016). Chromatin-associated transcripts of tandemly repetitive DNA sequences revealed by RNA-FISH. *Chromosome Res.* 24 (4), 467–480. doi: 10.1007/s10577-016-9537-5

Lapitan, N. L. V., Brown, S. E., Kennard, W., Stephens, J. L., and Knudson, D. L. (1997). FISH physical mapping with barley BAC clones. *Plant J.* 11, 149–156. doi: 10.1046/j.1365313X.1997.11010149.x

Li, L. F., Häkkinen, M., Yuan, Y. M., Hao, G., and Ge, X. J. (2010). Molecular phylogeny and systematics of the banana family (*Musaceae*) inferred from multiple nuclear and chloroplast DNA fragments, with a special reference to the genus *Musa*. *Mol. Phylogenet. Evol.* 57 (1), 1–10. doi: 10.1016/j.ympev.2010.06.021

Liu, W., Rouse, M., Friebe, B., Jin, Y., Gill, B., and Pumphrey, M. O. (2011). Discovery and molecular mapping of a new gene conferring resistance to stem rust, Sr53, derived from *Aegilops geniculata* and characterization of spontaneous translocation stocks with reduced alien chromatin. *Chromosome Res.* 19 (5), 669–682. doi: 10.1007/s10577-011-9226-3

Lysák, M. A., Doleželová, M., Horry, J. P., Swennen, R., and Doležel, J. (1999). Flow cytometric analysis of nuclear DNA content in *Musa*. *Theor. Appl. Genet.* 98 (8), 1344–1350. doi: 10.1007/s001220051201

Lysák, M. A., Fransz, P. F., Ali, H. B. M., and Schubert, I. (2001). Chromosome painting in *A. thaliana*. *Plant J.* 28, 689–697. doi: 10.1046/j.1365-313x.2001.01194.x

Mandáková, T., and Lysák, M. A. (2008). Chromosomal phylogeny and karyotype evolution in x = 7 crucifer species (*Brassicaceae*). *Plant Cell* 20, 2559–2570. doi: 10.1105/tpc.108.062166

Mandáková, T., Marhold, K., and Lysák, M. A. (2013). The widespread crucifer species *Cardamine flexulosa* is an allotetraploid with a conserved subgenomic structure. *New Phytol.* 201, 982–992. doi: 10.1111/nph.12567

Martin, G., Baurens, F. C., Droc, G., Rouard, M., Cenci, A., Kilian, A., et al. (2016). Improvement of the banana "*Musa acuminata*" reference sequence using NGS data and semi-automated bioinformatics methods. *BMC Genomics* 17, 1–12. doi: 10.1186/s12864-016-2579-4

Meng, Z., Zhang, Z. L., Yan, T. Y., Lin, Q. F., Wang, Y., Huang, W. Y., et al. (2018). Comprehensively characterizing the cytological features of *Saccharum spontaneum* by the development of a complete set of chromosome-specific oligo probes. *Front. Plant Sci.* 9, 1624. doi: 10.3389/fpls.2018.01624

Murata, M., Heslop-Harrison, J. S., and Motoyoshi, F. (1997). Physical mapping of the 5S ribosomal RNA genes in *Arabidopsis thaliana* by multi-color fluorescence *in situ* hybridization with cosmid clones. *Plant J.* 12 (1), 31–37. doi: 10.1046/j.1365-313X.1997.12010031.x

Murgha, Y. E., Rouillard, J. M., and Gulari, E. (2014). Methods for the preparation of large quantities of complex single-stranded oligonucleotide libraries. *PloS One* 9, e94752. doi: 10.1371/journal.pone.0094752

Němečková, A., Christelová, P., Čížková, J., Nyine, M., Van den houwe, I., Svačina, R., et al. (2018). Molecular and cytogenetic study of East African Highland Banana. *Front. Plant Sci.* 9, 1371. doi: 10.3389/fpls.2018.01371

Neumann, P., Navrátilová, A., Koblížková, A., Kejnovský, E., Hřibová, E., Hobza, R., et al. (2011). Plant cytogenetic perspective. *Mob. DNA* 2 (1), 4. doi: 10.1186/1759-8753-2-4

Novák, P., Hřibová, E., Neumann, P., Koblížková, A., Doležel, J., and Macas, J. (2014). Genome-wide analysis of repeat diversity across the family *Musaceae*. *PloS One* 9 (6), e98918. doi: 10.1371/journal.pone.0098918

Ortiz, R., and Swennen, R. (2014). From crossbreeding to biotechnology-facilitated improvement of banana and plantain. *Biotechnol. Adv.* 32, 158–169. doi: 10.1016/j.biotechadv.2013.09.010

Osuji, J. O., Crouch, J., Harrison, G., and Heslop-Harrison, J. S. (1998). Molecular cytogenetics of *Musa* species, cultivars and hybrids: location of 18S-5.8S-25S and 5S rDNA and telomere-like sequences. *Ann. Bot.* 82, 243–248. doi: 10.1006/anbo.1998.0674

Qu, M., Li, K., Han, Y., Chen, L., Li, Z., and Han, Y. (2017). Integrated karyotyping of woodland strawberry (*Fragaria vesca*) with oligopaint FISH probes. *Cytogenet. Genome Res.* 153, 158–164. doi: 10.1159/000485283

Said, M., Hřibová, E., Danilova, T. V., Karafiátová, M., Čížková, J., Friebe, B., et al. (2018). The *Agropyron cristatum* karyotype, chromosome structure and cross-genome homoeology as revealed by fluorescence *in situ* hybridization with tandem repeats and wheat single-gene probes. *Theor. Appl. Genet.* 131 (10), 2213–2227. doi: 10.1007/s00122-018-3148-9

Schubert, I., Fransz, P. F., Fuchs, J., and De Jong, J. H. (2001). Chromosome painting in plants. *Methods Cell Sci.* 23, 57–69. doi: 10.1023/A:1013137415093

Simmonds, N. W., and Shepherd, K. (1955). The taxonomy and origins of the cultivated bananas. *J. Linn. Soc. Bot.* 55, 302–312. doi: 10.1111/j.1095-8339.1955.tb00015.x

Simmonds, N. W. (1956). Botanical results of the banana collecting expeditions, 1954-5. *Kew Bull.* 11, 463–489. doi: 10.2307/4109131

Speicher, M. R., Ballard, S. G., and Ward, D. C. (1996). Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat. Genet.* 12 (4), 368–375. doi: 10.1038/ng0496-368

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 6814, 796–815. doi: 10.1038/35048692

The International Brachypodium Initiative (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463 (7282), 763–768. doi: 10.1038/nature08747

Valárik, M., Šimková, H., Hřibová, E., Šafář, J., Doleželová, M., and Doležel, J. (2002). Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (*Musa* spp.). *Chromosome Res.* 10 (2), 89–100. doi: 10.1023/A:1014945730035

Xin, H., Zhang, T., Han, Y., Wu, Y., Shi, J., Xi, M., et al. (2018). Chromosome painting and comparative physical mapping of the sex chromosomes in *Populus tomentosa* and *Populus deltoides*. *Chromosoma.* 127, 313–321. doi: 10.1007/s00412-018-0664-y

Zimin, A. V., Puiu, D., Hall, R., Kingan, S., Clavijo, B. J., and Salzberg, S. L. (2017). The first near-complete assembly of the hexaploid bread wheat genome, *Triticum aestivum*. *Gigascience.* 6 (11), 1–7. doi: 10.1093/gigascience/gix097

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# APPENDIX II

## Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure

Šimoníková, D., Němečková, A., Čížková, J., Brown, A., Swennen, R., Doležel, J., Hřibová, E.

*International Journal of Molecular Sciences* **21**: 7915, 2020

doi: 10.3390/ijms21217915

IF: 4.602

International Journal of
*Molecular Sciences*

MDPI

*Article*

# Chromosome Painting in Cultivated Bananas and Their Wild Relatives (*Musa* spp.) Reveals Differences in Chromosome Structure

**Denisa Šimoníková [1], Alžběta Němečková [1], Jana Čížková [1], Allan Brown [2], Rony Swennen [2,3], Jaroslav Doležel [1] and Eva Hřibová [1,*]**

[1] Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Hana for Biotechnological and Agricultural Research, 77900 Olomouc, Czech Republic; simonikova@ueb.cas.cz (D.Š.); nemeckova@ueb.cas.cz (A.N.); cizkova@ueb.cas.cz (J.Č.); dolezel@ueb.cas.cz (J.D.)

[2] International Institute of Tropical Agriculture, Banana Breeding, PO Box 447 Arusha, Tanzania; A.Brown@cgiar.org (A.B.); R.Swennen@cgiar.org (R.S.)

[3] Division of Crop Biotechnics, Laboratory of Tropical Crop Improvement, Katholieke Universiteit Leuven, 3001 Leuven, Belgium

* Correspondence: hribova@ueb.cas.cz; Tel.: +420-585-238-713

Received: 29 September 2020; Accepted: 21 October 2020; Published: 24 October 2020

**Abstract:** Edible banana cultivars are diploid, triploid, or tetraploid hybrids, which originated by natural cross hybridization between subspecies of diploid *Musa acuminata*, or between *M. acuminata* and diploid *Musa balbisiana*. The participation of two other wild diploid species *Musa schizocarpa* and *Musa textilis* was also indicated by molecular studies. The fusion of gametes with structurally different chromosome sets may give rise to progenies with structural chromosome heterozygosity and reduced fertility due to aberrant chromosome pairing and unbalanced chromosome segregation. Only a few translocations have been classified on the genomic level so far, and a comprehensive molecular cytogenetic characterization of cultivars and species of the family Musaceae is still lacking. Fluorescence in situ hybridization (FISH) with chromosome-arm-specific oligo painting probes was used for comparative karyotype analysis in a set of wild *Musa* species and edible banana clones. The results revealed large differences in chromosome structure, discriminating individual accessions. These results permitted the identification of putative progenitors of cultivated clones and clarified the genomic constitution and evolution of aneuploid banana clones, which seem to be common among the polyploid banana accessions. New insights into the chromosome organization and structural chromosome changes will be a valuable asset in breeding programs, particularly in the selection of appropriate parents for cross hybridization.

**Keywords:** chromosome translocation; fluorescence in situ hybridization; karyotype evolution; oligo painting FISH; structural chromosome heterozygosity

## 1. Introduction

Banana represents one of the major staple foods and is one of the most important cash crops with the estimated value of $25 billion for the banana industry. The annual global production of bananas reached 114 million tons in 2017 [1], with about 26 million tons exported in 2019 [2]. Two types of bananas are known—sweet bananas, serving as a food supplement, and cooking bananas, which are characterized by starchier fruits [3]. Edible banana cultivars are vegetatively propagated diploid, triploid, and tetraploid hybrids, which originated after natural cross hybridization between the wild diploids *Musa acuminata* (2n = 2x = 22, AA genome) and *Musa balbisiana* (2n = 2x = 22, BB

genome) and their hybrid progenies. To some extent, other *Musa* species such as *Musa schizocarpa* (2n = 2x = 22, SS genome) and *Musa textilis* (2n = 2x = 20, TT genome) also contributed to the origin of some edible banana clones [4–6].

Based on morphology and geographical distribution, *M. acuminata* has been divided into nine subspecies (*banksii, burmannica, burmannicoides, errans, malaccensis, microcarpa, siamea, truncata,* and *zebrina*) and three varieties (chinensis, sumatrana, and tomentosa) [7–9]. It has been estimated that at least four subspecies of *M. acuminata* contributed to the origin of cultivated bananas [7,10]. Out of them, *M. acuminata* ssp. *banksii*, with the original center of diversity in New Guinea played a major role in this process [6,11,12]. Other subspecies were *M. acuminata* ssp. *burmannica* with the center of diversity in Myanmar [13]; ssp. *malaccensis* with the origin in Malay peninsula [7,14]; and ssp. *zebrina,* which originated in Indonesia [10].

It is believed that human migration together with a different geography of the present archipelago in Southeast Asia during the glacial period, when a drop in the sea level resulted in the interconnection of current islands in Southeast Asia into one land mass [15–18], brought different *M. acuminata* subspecies into close proximity, enabling cross hybridization and giving rise to diploid intraspecific hybrids that were subjected to human selection and propagation [7,8]. The fusion of unreduced gametes produced by diploid edible and partially sterile cultivars with normal haploid gametes from fertile diploids [4,19,20] would give rise to triploids. The number one export dessert banana type Cavendish as well as other important dessert bananas, such as Gros Michel and Pome types, originated according to this scenario by the hybridization of a diploid representative of subgroup "Mchare" (originally named "Mlali") (AA genome; *zebrina*/*microcarpa* and *banksii* ascendance), which served as a donor of an unreduced diploid gamete with the haploid gamete of "Khai" (*malaccensis* ascendance) [12,21,22].

Another group of edible triploid bananas, clones with AAB (so called plantains) or ABB constitution, cover nearly 40% of global banana production, whereas plantains stand for 18% of total banana production [23]. These interspecific triploid cultivars originated after fusion of an unreduced gamete from interspecific AB hybrid with haploid gamete from diploid *M. acuminata* ssp. *banksii* or *M. balbisiana* [7,23]. Their evolution most probably involved several backcrosses [24]. Two important AAB subgroups of starchy bananas evolved in two centers of diversity, the African Plantains and the Pacific (Maia Maoli/Popoulu) Plantains [14].

East African Highland bananas (EAHB) represent an important starchy type of banana for over 80 million people living in the Great Lakes region of East Africa, which is considered a secondary center of banana diversity [25,26]. EAHBs are a subgroup of triploid bananas with AAA constitution, which arose from hybridization between diploid *M. acuminata* ssp. *banksii* and *M. acuminata* ssp. *zebrina*. However, *M. schizocarpa* also contributed to the formation of these hybrids [6]. Interestingly, different EAHB varieties have relatively low genetic diversity, contrary to morphological variation, which occurred most probably due to the accumulation of somatic mutations and epigenetic changes [6,7,12,27,28].

Cultivated clones originating from inter-subspecific and interspecific hybridization and with the contribution of unreduced gametes in the case of triploid clones have reduced or zero production of fertile gametes. This is a consequence of aberrant chromosome pairing during meiosis due to structural chromosome heterozygosity and/or odd ploidy levels. Reduced fertility greatly hampers the efforts to breed improved cultivars [8,23,29,30], which are needed to satisfy the increasing demand for dessert and starchy bananas under the conditions of climate change and increasing pressure of pests and diseases [28].

Traditional breeding strategies for triploid banana cultivars involve the development of tetraploids (4x) from 3x × 2x crosses, followed by the production of secondary triploid hybrids (3x) from 4x × 2x crosses [31–34]. Similarly, diploids also play important role in the breeding strategy of diploid cultivars, 4x × 2x crosses are used to create improved cultivars [33]. It is thus necessary to identify cultivars that produce seeds under specific conditions, followed by breeding for target traits and re-establishing seed-sterile end products. Low seed yield after pollination (e.g., 4 seeds per Matooke (genome AAA) bunch and only 1.6 seeds per Mchare (genome AA) bunch followed by

embryo rescue, with very low germination rate (~2%) illustrates the serious bottleneck for the breeding processes (Brown and Swennen, unpublished).

Since banana breeding programs use diploids as the principal vehicle for introducing genetic variability (e.g., [35–37]), the knowledge of their genome structure at chromosomal level is critical to reveal possible causes of reduced fertility and the presence of non-recombining haplotype blocks, to provide data to identify the parents of cultivated clones that originated spontaneously without direct human intervention, and to select parents for improvement programs. In order to provide insights into the genome structure at chromosome level, we employed fluorescence in situ hybridization (FISH) with chromosome-arm-specific oligo painting probes in a set of wild *Musa* species and edible banana clones potentially useful in banana improvement. Chromosome painting in twenty representatives of the Eumusa section of genus *Musa*, which included subspecies of *M. acuminata*, *M. balbisiana,* and their inter-subspecific and interspecific hybrids, revealed chromosomal rearrangements discriminating subspecies of *M. acuminata* and structural chromosome heterozygosity of cultivated clones. Identification of chromosome translocations pointed to particular *Musa* subspecies as putative parents of cultivated clones and provided an independent support for hypotheses on their origin.

## 2. Results

### *2.1. Karyotype of* M. acuminata *and* M. balbisiana

The first part of the study focused on comparative karyotype analysis in six subspecies of *M. acuminata* and in *M. balbisiana*. Oligo painting FISH in structurally homozygous *M. acuminata* ssp. *banksii* "Banksii" and *M. acuminata* ssp. *microcarpa* "Borneo" did not reveal detectable chromosome translocations (Figure 1A) as compared to the reference banana genome of *M. acuminata* ssp. *malaccensis* "DH Pahang" [38]. Thus, the three subspecies share the same overall organization of their chromosome sets. On the other hand, different chromosome translocations, which were found to be subspecies specific, were identified in the remaining four subspecies of *M. acuminata* (ssp. *zebrina*, ssp. *burmannicoides*, ssp. *siamea,* ssp. *burmannica*).

In *M. acuminata* ssp. *zebrina*, a reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8 was identified (Figure 1B, Supplementary Figure S1A). Three phylogenetically closely related subspecies *M. acuminata* ssp. *burmannicoides*, *M. acuminata* ssp. *siamea,* and *M. acuminata* ssp. *burmannica* shared two translocations (Figure 1C–E, Supplementary Figure S2A,B,E,F). The first of these involved the transfer of a segment of the long arm of chromosome 8 to the long arm of chromosome 2; the second was a reciprocal translocation involving a large segment of the short arm of chromosome 9 and the long arm of chromosome 1. In addition to the translocations shared by representatives of the three subspecies, chromosome painting revealed additional subspecies-specific translocations.

In *M. acuminata* ssp. *siamea* "Pa Rayong", translocation of a small segment of the long arm of chromosome 3 to the short arm of chromosome 4 was detected. Importantly, this translocation was visible only on one of the homologs, indicating structural chromosome heterozygosity and a hybrid origin (Figure 1D, Supplementary Figure S2C). Subspecies-specific translocations involving only one of the homologs were also found in *M. acuminata* ssp. *burmannica* "Tavoy". They included a reciprocal translocation between chromosomes 3 and 8, which was also detected in *M. acuminata* ssp. *zebrina,* and a Robertsonian translocation between chromosomes 7 and 8, which gave rise to a chromosome comprising the long arms of chromosomes 7 and 8 and a chromosome made of the short arms of chromosomes 7 and 8 (Figure 1E, Supplementary Figures S1B and S2D).

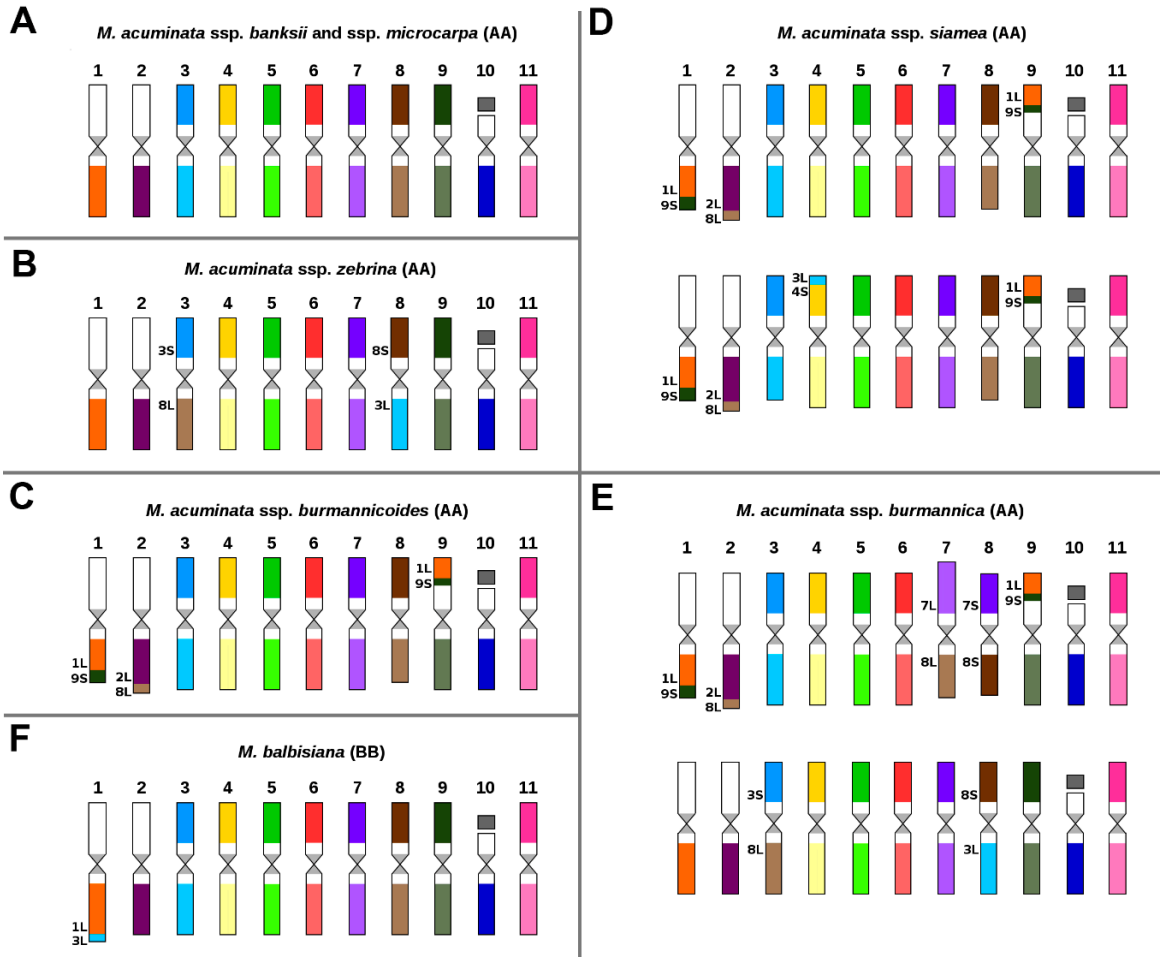



**Figure 1.** Ideograms of six diploid (2n = 2x = 22) subspecies of *Musa acuminata* and diploid (2n = 2x = 22) *Musa balbisiana*: (**A**) *M. acuminata* ssp. *banksii* "Banksii" and ssp. *microcarpa* "Borneo"; (**B**) *M. acuminata* ssp. *zebrina* "Maia Oa"; (**C**) *M. acuminata* ssp. *burmannicoides* "Calcutta 4"; (**D**) *M. acuminata* ssp. *siamea* "Pa Rayong"; (**E**) *M. acuminata* ssp. *burmannica* "Tavoy"; (**F**) *M. balbisiana* "Pisang Klutuk Wulung". Chromosome paints were not used for the short arms of chromosomes 1, 2, and 10. Chromosomes are oriented with their short arms on top and the long arms on the bottom in all ideograms. For better orientation, translocated parts of the chromosomes contain extra labels, which include chromosome number and the chromosome arm that was involved in the rearrangement.

In *M. balbisiana* "Pisang Klutuk Wulung" translocation of a small segment of the long arm of chromosome 3 to the long arm of chromosome 1 was observed (Figure 1F), similar to the translocation identified in our previous work [38] in *M. balbisiana* "Tani". In agreement with a low level of genetic diversity of *M. balbisiana* [18,39], no detectable differences in karyotypes were found between both accessions of the species.

### *2.2. Karyotype Structure of Edible Clones That Originated as Intraspecific Hybrids*

Chromosome painting in diploid cooking banana cultivars belonging to the Mchare group with AA genome, confirmed their hybrid origin. All three accessions analyzed in this work comprised two reciprocal translocations, with each of them observed only in one chromosome set. The first translocation involved the short segments of the long arms of chromosomes 4 and 1 (Figure 2A), while the second one involved a reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8 (Figure 2A, Supplementary Figure S1D). Both translocations were observed in a heterozygous state in all three analyzed Mchare banana representatives. Reciprocal translocation involving short segments of the long arms of chromosomes 4 and 1 was also identified in diploid cultivar "Pisang Lilin" with AA genome (Figure 2B, Supplementary Figure S3C). This clone

is used in breeding programs as a donor of useful agronomical characteristics, such as female fertility or resistance to yellow Sigatoka [40]. Moreover, this accession was heterozygous for the translocation, indicating a hybrid origin (Figure 2B).

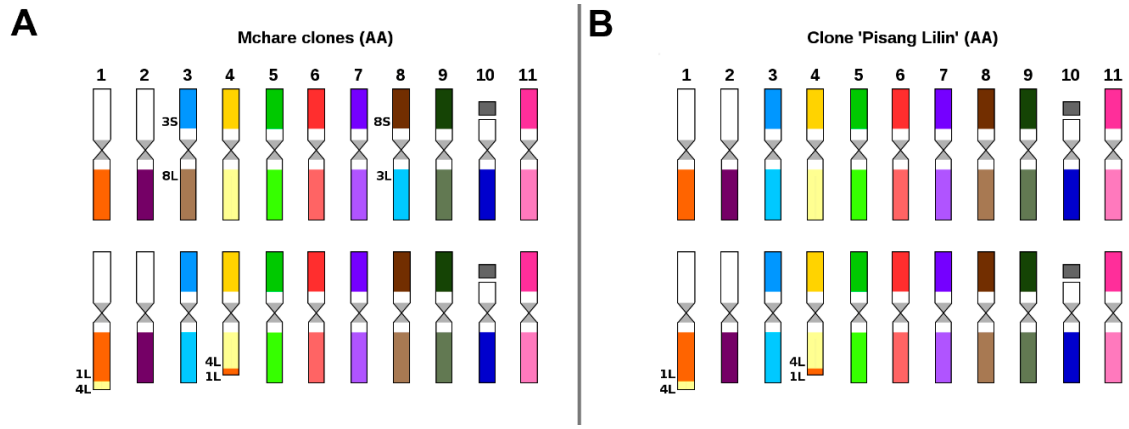



**Figure 2.** Ideograms of diploid (2n = 2x = 22) *Musa* clones: (**A**) *M. acuminata* subgr. Mchare, clones "Huti white", "Huti (Shumba nyeelu)", and "Ndyali"; (**B**) *M. acuminata* clone "Pisang Lilin". Chromosome paints were not used for the short arms of chromosomes 1, 2, and 10. Chromosomes are oriented with their short arms on top and the long arms on the bottom in all ideograms. For better orientation, translocated parts of the chromosomes contain extra labels, which include chromosome number and the chromosome arm that was involved in the rearrangement.

Two representatives of triploid dessert banana cultivars "Cavendish" and "Gros Michel" with AAA genome constitution displayed identical chromosome structure as assessed by chromosome painting. One of the three chromosome sets was characterized by reciprocal translocation between short segments of the long arms of chromosomes 4 and 1 (3A, Supplementary Figure S3D). It is worth mentioning that the same translocation was identified in Mchare cultivars and in "Pisang Lilin" (Figure 2A,B). Another chromosome set of "Cavendish" and "Gros Michel" (Figure 3A, Supplementary Figure S1E,F) contained a reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8, which was also identified in *M. acuminata* ssp. *zebrina* (Figure 1B). In addition to the two translocations, a translocation of the short arm of chromosome 7 to the long arm of chromosome 1 was observed in both accessions of desert banana. The translocation resulted in the formation of a small telocentric chromosome consisting of only the long arm of chromosome 7 (Figure 3A, Supplementary Figure S3E).

East African Highland bananas (EAHB) represent an important group of triploid cultivars with the genome constitution AAA. We have analyzed two accessions of these cooking bananas. Both cultivars ("Imbogo" and "Kagera") contained a reciprocal translocation between the short arm of chromosome 3 and the long arm of chromosome 8, which was also observed in *M. acuminata* ssp. *zebrina.* This translocation was identified only in two out of the three chromosome sets (Figure 3B,C, Supplementary Figure S1C). Karyotype analysis revealed that cultivar "Imbogo" lacked one chromosome and was aneuploid (2n = 32). Chromosome painting facilitated identification of the missing chromosome and suggested the origin of the aneuploid karyotype (Figure 3C, Supplementary Figure S4), which involved a Robertsonian translocation between chromosomes 7 and 1, giving rise to a recombined chromosome containing the long arms of chromosomes 7 and 1. Our observation suggests that the short arms of the two chromosomes were lost (Figure 3C, Supplementary Figure S4). The loss of a putative chromosome comprising two short arms is a common consequence of the Robertsonian translocation [41].

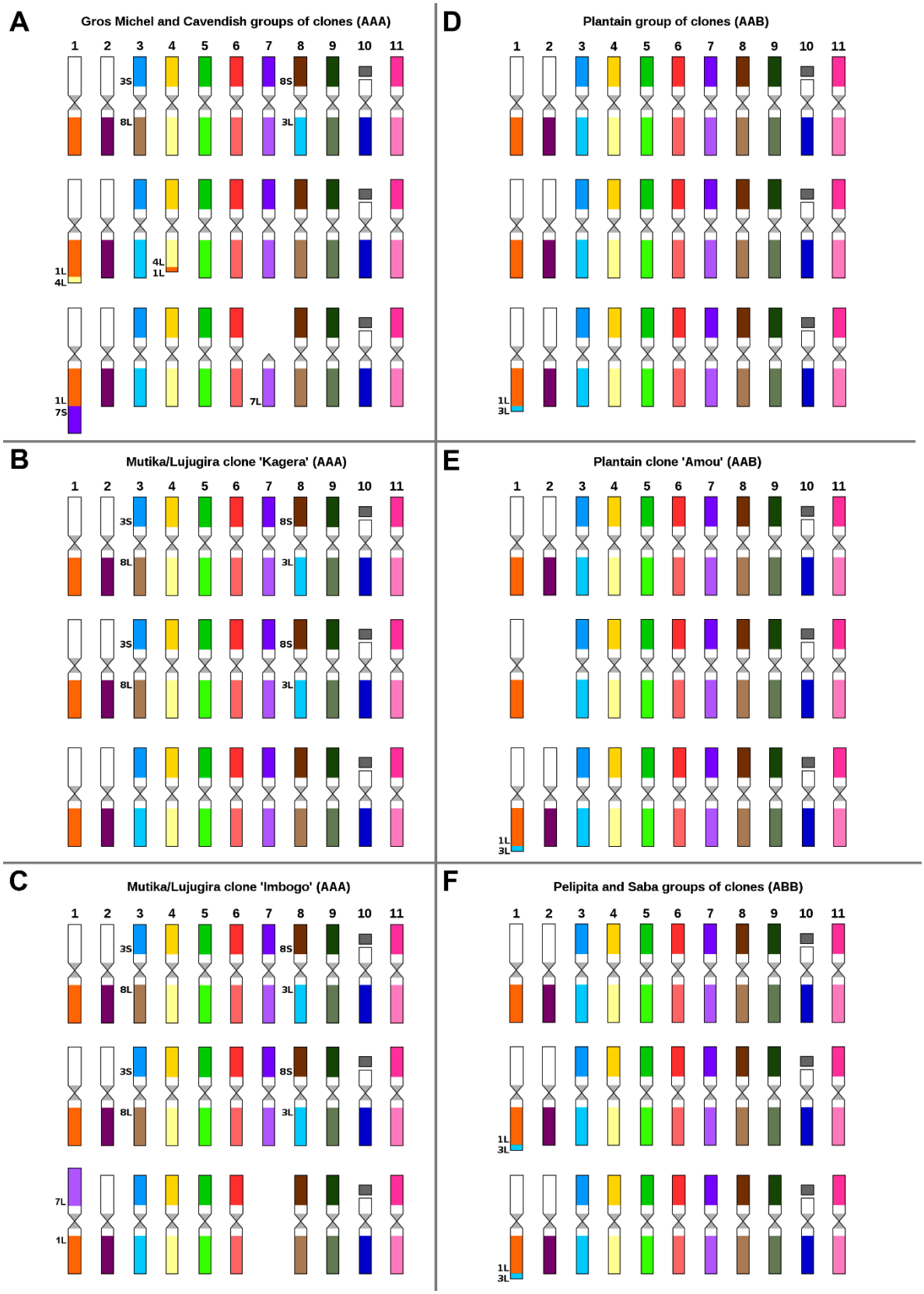



**Figure 3.** Ideograms of triploid (2n = 3x = 33) *Musa* accessions: (**A**) Clones "Gros Michel" and "Poyo"; (**B**) East African Highland banana (EAHB) clone "Kagera"; (**C**) aneuploid EAHB clone "Imbogo" (2n = 3x – 1 = 32); (**D**) plantains "3 Hands Planty" and "Obino l'Ewai"; (**E**) aneuploid plantain clone "Amou" (2n = 3x – 1 = 32); (**F**) clones "Pelipita" and "Saba sa Hapon". Chromosome paints were not used for the short arms of chromosomes 1, 2, and 10. Chromosomes are oriented with their short arms on top and the long arms on the bottom in all ideograms. For better orientation, translocated parts of

the chromosomes contain extra labels, which include chromosome number and the chromosome arm that was involved in the rearrangement.

### *2.3. Karyotype Structure of Interspecific Banana Clones*

Plantains are an important group of starchy type of bananas with AAB genome constitution and originated as hybrids between *M. acuminata* (A genome) and *M. balbisiana* (B genome). Chromosome painting in three cultivars representing two plantain morphotypes-Horn (cultivar "3 Hands Planty") and French type (cultivars "Obino l'Ewai" and "Amou") confirmed the presence of B-genome-specific translocation in one chromosome set in all three accessions, i.e., the translocation of a small segment of the long arm of chromosome 3 to the long arm of chromosome 1 (Figure 3D,E, Supplementary Figure S3B). No other translocation was found in these cultivars. However, the clone "Amou" was found to be aneuploid (2n = 32) as it missed one copy of chromosome 2 (Figure 3E, Supplementary Figure S3F).

In agreement with their predicted ABB genome constitution, B-genome-specific chromosome translocation was also observed in two chromosome sets in triploid cultivars "Pelipita" and "Saba sa Hapon" (Figure 3F, Supplementary Figure S3A). No other translocation was found in these cultivars.

## 3. Discussion

Until recently, cytogenetic analysis in plants was hindered by the lack of available DNA probes suitable for fluorescent painting of individual chromosomes [42,43]. The only option was to use pools of bacterial artificial chromosome (BAC) clones, which were found useful in plant species with small genomes (e.g., [44,45]). The development of oligo painting FISH [46] changed this situation dramatically, and it is now possible to label individual plant chromosomes and chromosomal regions in many species [47–52]. Chromosome-arm-specific oligo painting probes were also recently developed for banana (*Musa* spp.) by Šimoníková et al. [38], who demonstrated the utility of this approach for anchoring DNA pseudomolecules of a reference genome sequence to individual chromosomes in situ and for the identification of chromosome translocations.

In this work, we used oligo painting FISH for comparative karyotype analysis in a set of *Musa* accessions comprising wild species used in banana breeding programs and economically significant edible cultivars (Supplementary Table S1). These experiments revealed chromosomal translocations in subspecies of *M. acuminata* (A genome), their intraspecific hybrids as well as in *M. balbisiana* (B genome) and in interspecific hybrid clones originating from cross hybridization between *M. acuminata* and *M. balbisiana* (Figure 4). A difference in chromosome structure among *M. acuminata* subspecies was suggested earlier by Shepherd [53], who identified seven translocation groups in *M. acuminata* based on chromosome pairing during meiosis. An independent confirmation of this classification was the observation of segregation distortion during genetic mapping in inter-subspecific hybrids of *M. acuminata* [54–57].

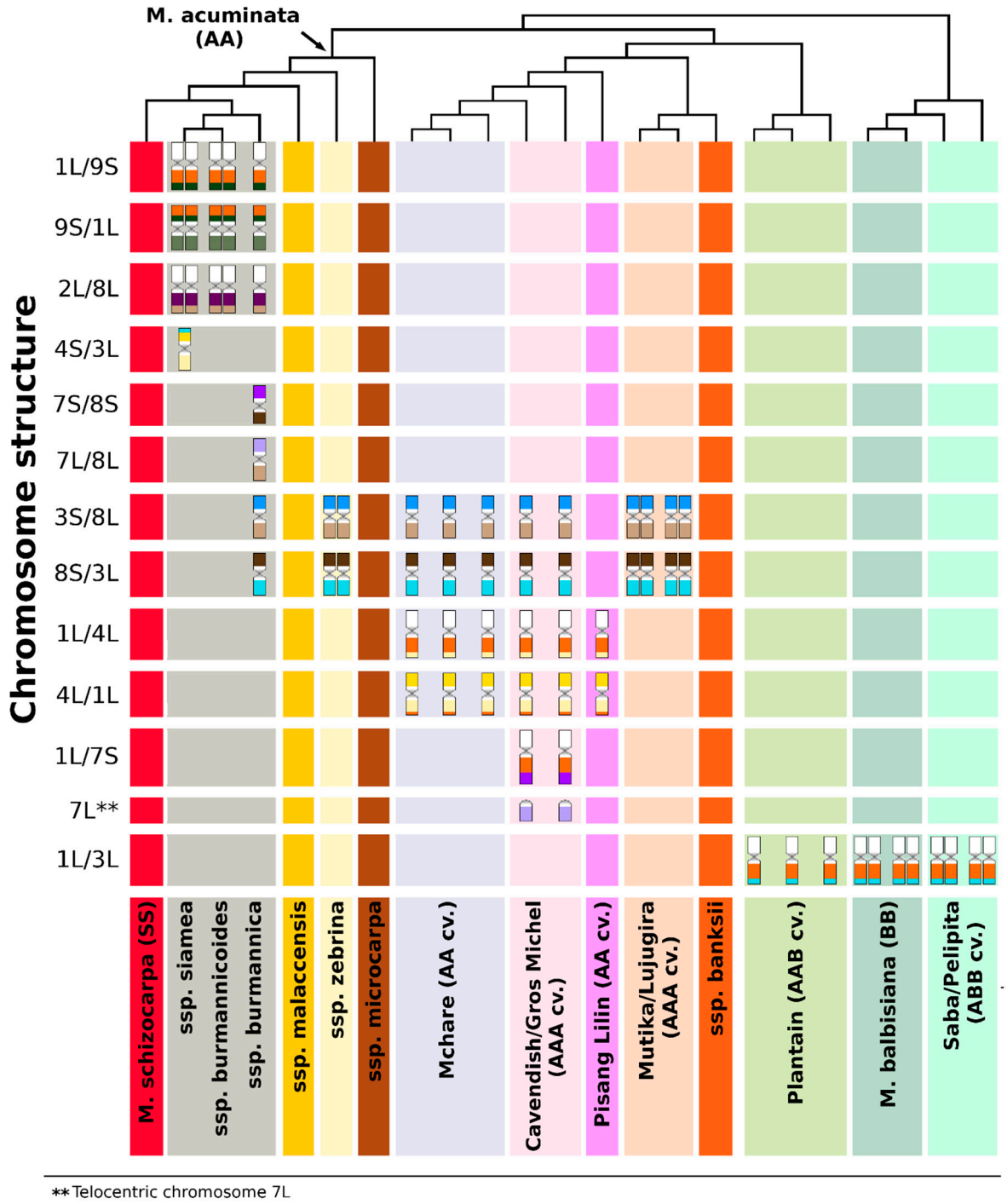



**Figure 4.** Overview of common translocation events revealed by oligo painting fluorescence in situ hybridization (FISH) in *Musa*: diversity tree, constructed using simple sequence repeat (SSR) markers according to Christelová et al. [28], shows the relationships among *Musa* species, subspecies, and hybrid clones. Lineages of closely related accessions and groups of edible banana clones are highlighted in different colors. Individual chromosome structures (displayed as chromosome schemes) are depicted in rows, and their number correspond to the number of chromosomes bearing the rearrangement in the nuclear genome in somatic cell lines (2n).

### *3.1. Structural Genome Variation in Diploid* M. acuminata

In this work, we analyzed one representative of each of six subspecies of *M. acuminata*. We cannot exclude differences in chromosome structure within individual subspecies. However, given the large genetic homogeneity of the six subspecies as clearly demonstrated by molecular markers [6,12,28,58–60], this does not seem probable. We observed a conserved genome structure in *M. acuminata* ssp. *banksii* and *M. acuminata* ssp. *microcarpa,* which did not contain any translocation chromosome when compared to the reference genome of *M. acuminata* ssp. *malaccensis* "DH Pahang" [61]. The genome structure shared by the three subspecies was also observed in *M. schizocarpa* [38] and corresponds to the standard translocation (ST) group as defined by Shepherd [53].

The other chromosome translocation group defined by Shepherd [53], the Northern Malayan group (NM), is characteristic for *M. acuminata* ssp. *malaccensis* and some AA cultivars, including "Pisang Lilin". Our results revealed a reciprocal translocation between chromosomes 1 and 4 in one chromosome set of "Pisang Lilin", thus confirming Shepherd's characterization of this clone as heterozygous, having ST x NM genome structure. Before the *Musa* genome sequence became available, Hippolyte et al. [55] assumed the presence of a duplication of the distal region of chromosome 1 on chromosome 4 in this clone based on comparison of high-density genetic maps. Taking the advantage of the availability of a reference genome sequence and after resequencing genomes of a set of *Musa* species, Martin et al. [8] described heterozygous reciprocal translocation between chromosomes 1 and 4, involving 3 Mb of the long arm of chromosome 1 and a 10 Mb segment of the long arm of chromosome 4 in *M. acuminata* ssp. *malaccensis*. Further experiments indicated preferential transmission of the translocation to the progeny and its frequent presence in triploid banana cultivars [8]. However, it is not clear whether the reciprocal translocation between chromosomes 1 and 4, which we observed in the heterozygous state only, originated in ssp. *malaccensis*, or whether it was transmitted to genomes of some *malaccensis* accessions by ancient hybridization events [8]. Three phylogenetically closely related subspecies *M. acuminata* ssp. *burmannica*, *M. acuminata* ssp. *burmannicoides,* and *M. acuminata* ssp. *siamea*, which have similar phenotype and geographic distribution [12,19] share a translocation of a part of the long arm of chromosome 8 to the long arm of chromosome 2, and a reciprocal translocation between chromosomes 1 and 9. These translocations were identified recently by Dupouy et al. [60] after mapping mate-paired Illumina sequence reads to the reference genome of "DH Pahang". The authors estimated the size of translocated region of chromosome 8 to chromosome 2 to be 7.2 Mb, while the size of the distal region of chromosome 2, which was found translocated to chromosome 8 in wild diploid clone "Calcutta 4" (ssp. *burmannicoides*) was estimated to be 240 kb [60]. The size of the translocated regions of chromosomes 1 and 9 was estimated to be 20.8 and 11.6 Mb, respectively [60]. Using oligo painting, we did not detect the 240 kb distal region of chromosome 2 translocated to chromosome 8, and this may reflect the limitation in the sensitivity of whole chromosome arm oligo painting.

The shared translocations in all three subspecies of *M. acuminata* (*burmannicoides*, *burmannica,* and *siamea*) support their close phylogenetic relationship as proposed by Shepherd [53] and later verified by molecular studies [6,12,28,58,59]. Dupouy et al. [60] coined the idea of a genetically uniform *burmannica* group. However, our data indicate a more complicated evolution of the three genotypes recognized as representatives of different *acuminata* subspecies. First, the characteristic translocations between chromosomes 2 and 8, and 1 and 9 were detected only on one chromosome set in *burmannica* "Tavoy", as compared to those in *M. acuminata* "Calcutta 4" (ssp. *burmannicoides*) and "Pa Rayong" (ssp. *siamea*). Second, we observed two subspecies-specific translocations in ssp. *burmannica* only in one chromosome set, and we detected additional subspecies-specific translocation in *M. acuminata* ssp. *siamea* only in one chromosome set, indicating its hybrid origin. Based on these results, we hypothesize that ssp. *burmannicoides* could be a progenitor of the clones characterized by structural chromosome heterozygosity. Divergence in genome structure between the three subspecies (*burmannica, burmannicoides,* and *siamea*) was demonstrated also by Shepherd [53], who classified some *burmannica* and *siamea* accessions as Northern 2 translocation group of *Musa*, differing

from the Northern 1 group (*burmannicoides* and other *siamea* accessions) by one additional translocation.

We observed subspecies-specific translocations also in *M. acuminata* ssp. *zebrina* "Maia Oa". In this case, chromosome painting revealed a Robertsonian translocation between chromosomes 3 and 8. Interestingly, Dupouy et al. [60] failed to detect this translocation after sequencing mate-pair libraries using Illumina technology. The discrepancy may point to the limitation of the sequencing approach to identify translocations arising by a breakage of (peri)centromeric regions. As these regions comprise mainly DNA repeats, they may not be assembled properly in a reference genome, thus preventing their identification by sequencing. In fact, this problem may also be encountered if subspecies-specific genome regions are absent in the reference genome sequence. We also revealed the *zebrina*-type translocation (a reciprocal translocation between chromosomes 3 and 8) in all three analyzed cultivars of diploid Mchare banana. The presence of a translocation between the long arm of chromosome 1 and the long arm of chromosome 4 on one chromosome set of Mchare indicates a hybrid origin of Mchare, with ssp. *zebrina* being one of the progenitors of this banana group. This agrees with the results obtained by genotyping using molecular markers [28]. Most recently, the complex hybridization scheme of Mchare bananas was also supported by the application of transcriptomic data for the identification of specific single-nucleotide polymorphisms (SNPs) in 23 *Musa* species and edible cultivars [22].

### *3.2. Genome Structure and Origin of Cultivated Triploid* Musa *Clones*

Plantains are an important group of triploid starchy bananas with AAB genome constitution, which originated after hybridization between *M. acuminata* and *M. balbisiana*. As expected, chromosome painting in "3 Hands Planty", "Amou", and "Obino l'Ewai" cultivars revealed B-genome-specific translocation of 8 Mb segment from the long arm of chromosome 3 to the long arm of chromosome 1. Unlike the B-genome chromosome set, the two A-genome chromosome sets of plantains lacked any detectable translocation. Genotyping using molecular markers revealed that the A genomes of the plantain group are related to *M. acuminata* ssp. *banksii* [18,58,62,63]. This is in line with the absence of chromosome translocations we observed in *M. acuminata* ssp. *banksii*.

Triploid cultivar "Pisang Awak", a representative of the ABB group, is believed to contain one A-genome chromosome set closely related to *M. acuminata* ssp. *malaccensis* [12]. However, after simple sequence repeat (SSR) genotyping, Christelová et al. [28] found, that some ABB clones from Saba and Bluggoe-Monthan groups clustered together with the representatives of Pacific banana Maia Maoli/Popoulu (AAB). These results point to *M. acuminata* ssp. *banksii* as their most probable progenitor. Unfortunately, in this case, chromosome painting did not bring useful hints on the nature of the A subgenomes in these interspecific hybrids, as *M. acuminata* ssp. *banksii* and ssp. *malaccensis* representatives do not differ in the presence of specific chromosome translocation.

The presence of B-genome-specific translocation of a small region of long arm of chromosome 3 to the long arm of chromosome 1 [38], observed in all *M. balbisiana* accessions and their interspecific hybrids with *M. acuminata*, seems to be a useful cytogenetic landmark of the presence of B subgenome. In our study, the number of chromosome sets containing B-genome-specific translocation agreed with the predicted genomic constitution (AAB or ABB) of the hybrids. Clearly, one B-genome-specific landmark is not sufficient for the analysis of the complete genome structure of interspecific hybrids at the cytogenetic level. Further work is needed to identify additional cytogenetic landmarks to uncover the complexities of genome evolution after interspecific hybridization in *Musa*. Recently, Baurens et al. [23] analyzed the genome composition of banana interspecific hybrid clones using whole-genome sequencing strategies followed by bioinformatic analysis based on A- and B-genome-specific SNP calling and they also detected the B-genome-specific translocation in interspecific hybrids.

Chromosome painting confirmed a small genetic difference between triploid clones "Gros Michel" and "Cavendish" (AAA genomes) as previously determined by various molecular studies [20,28]. Both clones share the same reciprocal translocation between the long arms of chromosomes 1 and 4 in one chromosome set. Interestingly, diploid Mchare cultivars also contain the same translocation in one

chromosome set. The presence of the translocation was also identified by sequencing genomic DNA both in "Cavendish" and "Gros Michel", as well as in Mchare banana "Akondro Mainty" [8]. These observations confirm a close genetic relationship between both groups of edible bananas as noted previously [12,20]. According to Martin et al. [8,22], a 2n gamete donor, which contributed to the origin of dessert banana clones with AAA genomes, including "Cavendish" and "Gros Michel", belongs to the Mchare (Mlali) subgroup. The genome of this ancient subgroup, which probably originated somewhere around Java, Borneo, and New Guinea, but today is only found in East Africa, is based on *zebrina*/*microcarpa* and *banksii* subspecies [12].

The third chromosome set in triploid "Cavendish" and "Gros Michel" contains a reciprocal translocation between chromosomes 3 and 8, which was detected by oligo painting FISH in the diploid *M. acuminata* ssp. *zebrina*. Our observations indicate that heterozygous representatives of *M. acuminata* ssp. *malaccensis* and *M. acuminata* spp. *zebrina* contributed to the origin of "Cavendish" and "Gros Michel" as their ancestors as suggested earlier [12,28,64]. According to Perrier et al. [12], *M. acuminata* ssp. *banksii* was one of the two progenitors of "Cavendish"/"Gros Michel" group of cultivars. However, using chromosome-arm-specific oligo painting, we observed a translocation of the short arm of chromosome 7 to the long arm of chromosome 1, resulting in a small telocentric chromosome made only of the long arm of chromosome 7, which was, up to now, identified only in these cultivars. The translocation, which gave rise to the small telocentric chromosome, could be a result of processes accompanying the evolution of this group of triploid AAA cultivars. Alternatively, another wild diploid clone, possibly structurally heterozygous, was involved in the origin of "Cavendish"/"Gros Michel" bananas.

We observed reciprocal translocation between chromosomes 3 and 8, which is typical for *M. acuminata* ssp. *zebrina*, in the economically important group of triploid East African Highland bananas (EAHB) with AAA genome constitution. An important role of *M. acuminata* ssp. *zebrina* and *M. acuminata* ssp. *banksii* as the most probable progenitors of EAHB was suggested previously [6,22,27,28,58,65]. Our results, which indicate that EAHB contained two chromosome sets from ssp. *zebrina* and one chromosome set from ssp. *banksii*, point to the most probable origin of EAHB. Hybridization between *M. acuminata* ssp. *zebrina* and ssp. *banksii* could give rise to an intraspecific diploid hybrid with reduced fertility. Triploid EAHB cultivars then could originate by backcross of the intraspecific hybrid (a donor of non-reduced gamete) with *M. acuminata* ssp. *zebrina*, or with another diploid, most probably a hybrid of *M. acuminata* ssp. *zebrina*. Moreover, phylogenetic analysis of Němečková et al. [6] surprisingly also revealed a possible contribution of *M. schizocarpa* to EAHB formation, thus indicating a more complicated origin. Further investigation is needed, and the availability of EAHB genome sequence in particular, to shed more light on the origin and evolution of these triploid clones.

### *3.3. The Origin of Aneuploidy*

To date, aneuploidy in *Musa* has been identified by chromosome counting [5,6,66–69]. Although this approach is laborious and low throughput, it cannot be replaced by flow cytometric estimation of nuclear DNA amounts because of the differences in genome size between *Musa* species, subspecies, and their hybrids (e.g., [5,6,28]). A more laborious approach to achieve high-resolution flow cytometry as used by Roux et al. [70] is too slow and laborious to be practical. Thus, traditional chromosome counting [5,6,66–69] remains the most reliable approach. Obviously, it is not suitable to identify the chromosome(s) involved in aneuploidy and the origin of the aberrations.

Here, we employed chromosome painting to shed light on the nature of aneuploids among the triploid *Musa* accessions. One of the aneuploid clones was identified in the plantain "Amou", in which one copy of chromosome 2 was lost. The origin of aneuploidy in the clone "Imbogo" (AAA genome), a representative of EAHB, involved structural chromosome changes involving the breakage of chromosomes 1 and 7 in centromeric regions, followed by the fusion of the long arms of chromosomes 1 and 7 and the subsequent loss of the short arms of both chromosomes (Supplementary Figure S4). It needs to be noted that these plants were obtained from the International *Musa* Germplasm Transit Center (ITC, Leuven, Belgium), where the clones are stored

in vitro. The loss of the whole chromosome in plantain "Amou" could occur during long-term culture.

To conclude, the application of oligo painting FISH improved the knowledge on genomes of cultivated banana and their wild relatives at chromosomal level. For the first time, a comparative molecular cytogenetic analysis of twenty representatives of the Eumusa section of genus *Musa*, including accessions commonly used in banana breeding, was performed using chromosome painting. However, as only a single accession from each of the six subspecies of *M. acuminata* was used in this study, our results will need to be confirmed by analyzing more representatives from each taxon. The identification of chromosome translocations pointed to particular *Musa* subspecies as putative parents of cultivated clones and provided an independent support for hypotheses on their origin. While we have unambiguously identified a range of translocations, a precise determination of breakpoint positions will have to be done using a long-read DNA sequencing technology [71–73].

The discrepancies in the genome structure of banana diploids observed in our study and published data then point to alternative scenarios on the origin of the important crop. The observation on structural chromosome heterozygosity confirmed the hybrid origin of cultivated banana and some of the wild diploid accessions, which were described as individual subspecies, and serves to inform breeders on possible causes of reduced fertility. Due to aberrant meiosis, gametes produced by structural heterozygotes could be aneuploid and have unbalanced chromosome translocations, duplications, and deletions. Thus, the knowledge of genome structure at the chromosomal level and the identification of structural chromosome heterozygosity will aid breeders in selecting parents for improvement programs.

## 4. Materials and Methods

### *4.1. Plant Material and Diversity Tree Construction*

Representatives of twenty species and clones from the section Eumusa of genus *Musa* were obtained as in vitro rooted plants from the International *Musa* Transit Center (ITC, Bioversity International, Leuven, Belgium). In vitro plants were transferred to soil and kept in a heated greenhouse. Table 1 lists the accessions used in this work. Genetic diversity analysis of banana accessions used in the study was performed using SSR data according to Christelová et al. [28]. A Dendrogram was constructed based on the results of unweighted pair group method with arithmetic mean (UPGMA) analysis implemented in DARwin software v6.0.021 [74] and visualized in FigTree v1.4.0 [75].

**Table 1.** List of *Musa* accessions analyzed in this work and their genomic constitution.

| Species | Subspecies/ Subgroup | Accession Name | ITC Code [a] | Genomic Constitution | Chromosome Number (2n) |
|---|---|---|---|---|---|
| *M. acuminata* | *banksii* | Banksii | 0806 | AA | 22 |
| | *microcarpa* | Borneo | 0253 | AA | 22 |
| | *zebrina* | Maia Oa | 0728 | AA | 22 |
| | *burmannica* | Tavoy | 0072 | AA | 22 |
| | *burmannicoides* | Calcutta 4 | 0249 | AA | 22 |
| | *siamea* | Pa Rayong | 0672 | AA | 22 |
| *M. balbisiana* | - | Pisang Klutuk Wulung | - | BB | 22 |
| Cultivars | unknown | Pisang Lilin | - | AA | 22 |
| | Mchare | Huti White | - | AA | 22 |
| | Mchare | Huti (Shumba nyeelu) | 1452 | AA | 22 |
| | Mchare | Ndyali | 1552 | AA | 22 |
| | Cavendish | Poyo | 1482 | AAA | 33 |
| | Gros Michel | Gros Michel | 0484 | AAA | 33 |
| | Mutika/Lujugira | Imbogo | 0168 | AAA | 32* |
| | Mutika/Lujugira | Kagera | 0141 | AAA | 33 |

| | | | | |
|---|---|---|---|---|
| Plantain (Horn) | 3 Hands Planty | 1132 | AAB | 33 |
| Plantain (French) | Amou | 0963 | AAB | 32* |
| Plantain (French) | Obino l'Ewai | 0109 | AAB | 33 |
| Pelipita | Pelipita | 0472 | ABB | 33 |
| Saba | Saba sa Hapon | 1777 | ABB | 33 |

[a] Code assigned by the International Transit Center (ITC, Leuven, Belgium). * Aneuploidy observed in our study.

### *4.2. Preparation of Oligo Painting Probes and Mitotic Metaphase Chromosome Spreads*

Chromosome-arm-specific painting probes were prepared as described by Šimoníková et al. [38]. Briefly, sets of 20,000 oligomers (45 nt) covering individual chromosome arms were synthesized as immortal libraries by Arbor Biosciences (Ann Arbor, MI, USA) and then labeled directly by CY5 fluorochrome, or by digoxigenin or biotin according to Han et al. [46]. N.B.: In the reference genome assembly of *M. acuminata* "DH Pahang" [76], pseudomolecules 1, 6, 7 are oriented inversely to the traditional way karyotypes are presented, where the short arms are on the top (Supplementary Figure S5, see also Šimoníková et al. [38]).

Actively growing root tips from 3 to 5 individual plants representing each genotype were collected and pre-treated in 0.05% (*w*/*v*) 8-hydroxyquinoline for three hours at room temperature, fixed in 3:1 ethanol:acetic acid fixative overnight at −20 °C and stored in 70% ethanol at −20 °C. After washing in 75 mM KCl and 7.5 mM EDTA (pH 4), root tip segments were digested in a mixture of 2% (*w*/*v*) cellulase and 2% (*w*/*v*) pectinase in 75 mM KCl and 7.5 mM EDTA (pH 4) for 90 min at 30 °C. The suspension of protoplasts thus obtained was filtered through a 150 μm nylon mesh, pelleted and washed in 70% ethanol. For further use, the protoplast suspension was stored in 70% ethanol at −20 °C. Mitotic metaphase chromosome spreads were prepared by a dropping method from protoplast suspension according to Doležel et al. [77], the slides were postfixed in 4% (*v*/*v*) formaldehyde solution in 2 × SSC (saline-sodium citrate) solution, air dried and used for FISH.

### *4.3. Fluorescence In Situ Hybridization and Image Analysis*

Fluorescence in situ hybridization and image analysis were performed according to Šimoníková et al. [38]. Hybridization mixture containing 50% (*v*/*v*) formamide, 10% (*w*/*v*) dextran sulfate in 2 × SSC, and 10 ng/μL of labeled probes was added onto a slide and denatured for 3 min at 80 °C. Hybridization was carried out in a humid chamber overnight at 37 °C. The sites of hybridization of digoxigenin- and biotin-labeled probes were detected using anti-digoxigenin-FITC (Roche Applied Science, Penzberg, Germany) and streptavidin-Cy3 (ThermoFisher Scientific/Invitrogen, Carlsbad, CA, USA), respectively. Chromosomes were counterstained with DAPI and mounted in Vectashield Antifade Mounting Medium (Vector Laboratories, Burlingame, CA, USA). The slides were examined with Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) equipped with Cool Cube 1 camera (Metasystems, Altlussheim, Germany) and appropriate optical filters. The capture of fluorescence signals, merging of the layers, and measurement of chromosome length were performed with ISIS software 5.4.7 (Metasystems). The final image adjustment and creation of ideograms were done in Adobe Photoshop CS5. Different probe combinations hybridizing a minimum of ten preparations with mitotic metaphase chromosome spreads were used for the final karyotype reconstruction of each genotype. A minimum of ten mitotic metaphase plates per slide were captured for each probe combination.

**Supplementary Materials:** Supplementary Materials can be found at www.mdpi.com/1422-0067/21/21/7915/s1.

**Author Contributions:** E.H. and J.D. conceived the project. D.Š., A.N., and J.Č. conducted the study and processed the data. A.B. and R.S. provided the banana materials. D.Š. and E.H. wrote the manuscript. E.H., D.Š., A.B., R.S., and J.D. discussed the results and contributed to manuscript writing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Czech Science Foundation (award no. 19-20303S) and by the Ministry of Education, Youth, and Sports of the Czech Republic (Program INTER-EXCELLENCE, INTER-TRANSFER,

grant award LTT19, and ERDF project "Plants as a tool for sustainable global development" no. CZ.02.1.01/0.0/0.0/16_019/0000827).

**Acknowledgments:** We thank Ines van den Houwe for providing the plant material and all donors who supported this work also through their contributions to the CGIAR Fund (http://www.cgiar.org/who-we-are/cgiar-fund/fund-donors-2/), and in particular to the CGIAR Research Program Roots, Tubers, and Bananas (RTB-CRP). The computing was supported by the National Grid Infrastructure MetaCentrum (grant no. LM2010005 under the program Projects of Large Infrastructure for Research, Development, and Innovations).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Abbreviations

| | |
|---|---|
| BAC | bacterial artificial chromosome |
| DH Pahang | doubled haploid Pahang |
| EAHB | East African Highland banana |
| FISH | fluorescence in situ hybridization |
| FITC | fluorescein isothiocyanate |
| NM group | Northern Malayan group |
| SNP | single-nucleotide polymorphism |
| spp. | species |
| ssp. | subspecies |
| SSR | simple sequence repeat |
| ST group | standard translocation group |
| UPGMA | unweighted pair group method with arithmetic mean |

## References

1. FAOSTAT; Agriculture Organization of the United Nations; FAO, 2017. Available online: http://www.fao.org/home/en/ (accessed on 30 January 2020).
2. International trade statistics, 2019. https://www.wto.org/ (accessed on 20 May 2020).
3. Price, N.S. The origin and development of banana and plantain cultivation. In *Bananas and Plantains*, *World Crop Series*; Gowen, S., Ed.; Springer: Dordrecht, Netherlands, 1995.
4. Carreel, F.; Fauré, S.; Gonzalez de Leon, D.; Lagoda, P.J.L.; Perrier, X.; Bakry, F.; Tezenas du Montcel, H.; Lanaud, C.; Horry, J.P. Evaluation of the genetic diversity in diploid bananas (*Musa* spp.). *Genet. Sel. Evol.* **1994**, *26*, 125–136.
5. Čížková, J.; Hřibová, E.; Humplíková, L.; Christelová, P.; Suchánková, P.; Doležel, J. Molecular analysis and genomic organization of major DNA satellites in banana (*Musa* spp.). *PLoS ONE* **2013**, *8*, e54808.
6. Němečková, A.; Christelová, P.; Čížková, J.; Nyine, M.; Van den Houwe, I.; Svačina, R.; Uwimana, B.; Swennen, R.; Doležel, J.; Hřibová, E. Molecular and cytogenetic study of East African Highland Banana. *Front. Plant Sci.* **2018**, *9*, 1371.
7. Perrier, X.; De Langhe, E.; Donohue, M.; Lentfer, C.; Vrydaghs, L.; Bakry, F.; Carreel, F.; Hippolyte, I.; Horry, J.P.; Jenny, C.; et al. Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc. Natl. Acad. Sci. USA.* **2011**, *108*, 11311–11318.
8. Martin, G.; Carreel, F.; Coriton, O.; Hervouet, C.; Cardi, C.; Derouault, P.; Roques, D.; Salmon, F.; Rouard, M.; Sardos, J.; et al. Evolution of the banana genome (*Musa acuminata*) is impacted by large chromosomal translocations. *Mol. Biol. Evol.* **2017**, *34*, 2140–2152.
9. WCSP, World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew, 2018. Available online: http://wcsp.science.kew.org/ (accessed on 30 January 2020).
10. Rouard, M.; Droc, G.; Martin, G.; Sardos, J.; Hueber, Y.; Guignon, V.; Cenci, A.; Geigle, B.; Hibbins, M.S.; Yahiaoui, N.; et al. Three new genome assemblies support a rapid radiation in *Musa acuminata* (wild banana). *Genome Biol. Evol.* **2018**, *10*, 3129–3140.
11. Sharrock, S. Collecting *Musa* in Papua New Guinea. Identification of genetic diversity in the genus *Musa*. In *International Network for the Improvement of Banana and Plantain*; Jarret, R.L., Ed.; INIBAP: Montpellier, France, 1990; pp. 140–157.
12. Perrier, X.; Bakry, F.; Carreel, F.; Jenny, F.; Horry, J.P.; Lebot, V.; Hippolyte, I. Combining biological approaches to shed light on the evolution of edible bananas. *Ethnobot. Res. Appl.* **2009**, *7*, 199–216.
13. Cheesman, E.E. Classification of the bananas. *Kew Bull.* **1948**, *3*, 145–153.

14. De Langhe, E.; Vrydaghs, L.; De Maret, P.; Perrier, X.; Denham, T.P. Why bananas matter: An introduction to the history of banana domestication. *Ethnobot. Res. Appl.* **2009**, *7*, 165–177.
15. Sand, C. Petite histoire du peuplement de l'Océanie. In *Migrations et Identités—Colloque CORAIL, Nouméa (PYF), 1988/11/21-22. Papeete*; Ricard, M., Ed.; Université Française du Pacifique, 1989; pp. 39–40. Available online: http://www.documentation.ird.fr/hor/fdi:27803 (accessed on 30 January 2020).
16. Denham, T.; Haberle, S.; Lentfer, C. New evidence and revised interpretations of early agriculture in Highland New Guinea. *Antiquity* **2004**, *78*, 839–857.
17. Denham, T. From domestication histories to regional prehistory: Using plants to re-evaluate early and mid-holocene interaction between New Guinea and Southeast Asia. *Food History* **2010**, *8*, 3–22.
18. Kagy, V.; Wong, M.; Vandenbroucke, H.; Jenny, C.; Dubois, C.; Ollivier, A.; Cardi, C.; Mournet, P.; Tuia, V.; Roux, N.; et al. Traditional banana diversity in Oceania: An endangered heritage. *PLoS ONE* **2016**, *11*, e0151208.
19. Simmonds, N.W. *The Evolution of the Bananas. Tropical Science Series*; Longmans: London, UK, 1962; p. 170.
20. Raboin, L.M.; Carreel, F.; Noyer, J.L.; Baurens, F.C.; Horry, J.P.; Bakry, F.; Tezenas Du Montcel, H.; Ganry, J.; Lanaud, C.; Lagoda, P.J.L. Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete donors and n gamete donors. *Mol. Breed.* **2005**, *16*, 333–341.
21. Perrier, X.; Jenny, C.; Bakry, F.; Karamura, D.; Kitavi, M.; Dubois, C.; Hervouet, C.; Philippson, G.; De Langhe, E. East African diploid and triploid bananas: A genetic complex transported from South-East Asia. *Ann. Bot.* **2019**, *123*, 19–36.
22. Martin, G.; Cardi, C.; Sarah, G.; Ricci, S.; Jenny, C.; Fondi, E.; Perrier, X.; Glaszmann, J.-C.; D'Hont, A.; Yahiaoui, N. Genome ancestry mosaics reveal multiple and cryptic contributors to cultivated banana. *Plant J.* **2020**, *102*, 1008–1025.
23. Baurens, F.C.; Martin, G.; Hervouet, C.; Salmon, F.; Yohomé, D.; Ricci, S.; Rouard, M.; Habas, R.; Lemainque, A.; Yahiaoui, N.; et al. Recombination and large structural variations shape interspecific edible bananas genomes. *Mol. Biol. Evol.* **2019**, *36*, 97–111.
24. De Langhe, E.; Hřibová, E.; Carpentier, S.; Doležel, J.; Swennen, R. Did backcrossing contribute to the origin of hybrid edible bananas? *Ann. Bot.* **2010**, *106*, 849–857.
25. Cooper, H.; Spillane, C.; Hodgkin, T. *Broadening the Genetic Bases of Crop Production*. FAO: Rome, Italy, 2001.
26. Tugume, A.K.; Lubega, G.W.; Rubaihayo, P.R. Genetic diversity of East African Highland bananas using AFLP. *Infomusa* **2003**, *11*, 28–32.
27. Kitavi, M.; Downing, T.; Lorenzen, J.; Karamura, D.; Onyango, M.; Nyine, M.; Ferguson, M.; Spillane, C. The triploid East African Highland Banana (EAHB) genepool is genetically uniform arising from a single ancestral clone that underwent population expansion by vegetative propagation. *Theor. Appl. Genet.* **2016**, *129*, 547–561.
28. Christelová, P.; De Langhe, E.; Hřibová, E.; Čížková, J.; Sardos, J.; Hušáková, M.; Van den Houwe, I.; Sutanto, A.; Kepler, A.K.; Swennen, R.; et al. Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. *Biodivers. Conserv.* **2017**, *26*, 801–824.
29. Burke, J.M.; Arnold, M.L. Genetics and the fitness of hybrids. *Annu. Rev. Genet.* **2001**, *35*, 31–52.
30. Batte, M.; Swennen, R.; Uwimana, B.; Akech, V.; Brown, A.; Tumuhimbise, R.; Hovmalm, H.P.; Geleta, M.; Ortiz, R. Crossbreeding East African Highland Bananas: Lessons learnt relevant to the botany of the crop after 21 years of genetic enhancement. *Front. Plant Sci.* **2019**, *10*, 81.
31. Bakry, F.; Horry, J.P. Tetraploid hybrids from interploid 3×/2× crosses in cooking bananas. *Fruit*s **1992**, *47*, 641–655.
32. Tomepke, K.; Jenny, C.; Escalant, J.-V. A review of conventional improvement strategies for *Musa*. *InfoMusa* **2004**, *13*, 2–6.
33. Ortiz, R. Conventional banana and plantain breeding. *Acta Hortic.* **2013**, *986*, 177–194.
34. Nyine, M.; Uwimana, B.; Swennen, R.; Batte, M.; Brown, A.; Christelová, P.; Hřibová, E.; Lorenzen, J.; Doležel, J. Trait variation and genetic diversity in a banana genomic selection training population. *PLoS ONE* **2017**, *12*, e0178734.
35. Amorim, E.P.; Silva, S.O.; Amorim, V.B.O.; Pillay, M. Quality improvement of cultivated *Musa*. In *Banana Breeding: Progress and Challenges*; Pillay, M.; Tenkouano, A., Eds.; CRC Press: New York, NY, USA, 2011; pp. 252–280.
36. Amorim, E.P.; dos Santos-Serejo, J.A.; Amorim, V.B.O.; Ferreira, C.F.; Silva, S.O. Banana breeding at embrapa cassava and fruits. *Acta Hortic.* **2013**, *968*, 171–176.

37. Tenkouano, A.; Vuylsteke, D.; Agogbua, J.U.; Makumbi, D.; Swennen, R.; Ortiz, R. Diploid banana hybrids TMB2x5105-1 and TMB2x9128-3 with good combining ability, resistance to black sigatoga and nematodes. *HortScience* **2003**, *38*, 468–472.
38. Šimoníková, D.; Němečková, A.; Karafiátová, M.; Uwimana, B.; Swennen, R.; Doležel, J.; Hřibová, E. Chromosome painting facilitates anchoring reference genome sequence to chromosomes in situ and integrated karyotyping in banana (*Musa* spp.). *Front. Plant Sci.* **2019**, *10*, 1503.
39. De Langhe, E.; Perrier, X.; Donohue, M.; Denham, T. The original Banana Split: Multi-disciplinary implications of the generation of African and Pacific Plantains in Island Southeast Asia. *Ethnobot. Res. Appl.* **2015**, *14*, 299–312.
40. do Amaral, C.M.; de Almeida dos Santos-Serejo, J.; de Oliveira e Silva, S.; da Silva Ledo, C.A.; Amorim, E.P. Agronomic characterization of autotetraploid banana plants derived from 'Pisang Lilin' (AA) obtained through chromosome doubling. *Euphytica* **2015**, *202*, 435–443.
41. Robertson, W.R.B. Chromosome studies. I. Taxonomic relationships shown in the chromosomes of *Tettigidae* and *Acrididae*: V-shaped chromosomes and their significance in *Acrididae*, *Locustidae*, and *Gryllidae*: Chromosomes and variation. *J. Morphol.* **1916**, *27*, 179–331.
42. Schubert, I.; Fransz, P.F.; Fuchs, J.; De Jong, J.H. Chromosome painting in plants. *Meth. Cell Sci.* **2001**, *23*, 57–69.
43. Jiang, J. Fluorescence in situ hybridization in plants: Recent developments and future applications. *Chromosome Res.* **2019**, *27*, 153–165.
44. Lysák, M.A.; Fransz, P.F.; Ali, H.B.M.; Schubert, I. Chromosome painting in *A. thaliana*. *Plant J.* **2001**, *28*, 689–697.
45. Idziak, D.; Hazuka, I.; Poliwczak, B.; Wiszynska, A.; Wolny, E.; Hasterok, R. Insight into the karyotype evolution of *Brachypodium* species using comparative chromosome barcoding. *PLoS ONE* **2014**, *9*, e93503.
46. Han, Y.; Zhang, T.; Thammapichai, P.; Weng, Y.; Jiang, J. Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. *Genetics* **2015**, *200*, 771–779.
47. Qu, M.; Li, K.; Han, Y.; Chen, L.; Li, Z.; Han, Y. Integrated karyotyping of woodland strawberry (*Fragaria vesca*) with oligopaint FISH probes. *Cytogenet. Genome Res.* **2017**, *153*, 158–164.
48. Braz, G.T.; He, L.; Zhao, H.; Zhang, T.; Semrau, K.; Rouillard, J.M.; Torres, G.A.; Jiang, J. Comparative oligo-FISH mapping: An efficient and powerful methodology to reveal karyotypic and chromosomal evolution. *Genetics* **2018**, *208*, 513–523.
49. He, L.; Braz, G.T.; Torres, G.A.; Jiang, J.M. Chromosome painting in meiosis reveals pairing of specific chromosomes in polyploid *Solanum* species. *Chromosoma* **2018**, *127*, 505–513.
50. Machado, M.A.; Pieczarka, J.C.; Silva, F.H.R.; O'Brien, P.C.M.; Ferguson-Smith, M.A.; Nagamachi, C.Y. Extensive karyotype reorganization in the fish *Gymnotus arapaima* (*Gymnotiformes, Gymnotidae*) highlighted by zoo-FISH analysis. *Front. Genet.* **2018**, *9*, 8.
51. Albert, P.S.; Zhang, T.; Semrau, K.; Rouillard, J.M.; Kao, Y.H.; Wang, C.J.R.; Danilova, T.V.; Jiang, J.; Birchler, J.A. Whole-chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal relationships. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1679–1685.
52. Bi, Y.; Zhao, Q.; Yan, W.; Li, M.; Liu, Y.; Cheng, C.; Zhang, L.; Yu, X.; Li, J.; Qian, C.; et al. Flexible chromosome painting based on multiplex PCR of oligonucleotides and its application for comparative chromosome analyses in *Cucumis*. *Plant J.* **2020**, *102*, 178–186.
53. Shepherd, K. *Cytogenetics of the Genus Musa*; International Network for the Improvement of Banana and Plantain: Montpellier, France, 1999; ISBN 2-910810-25-9.
54. Fauré, S.; Noyer, J.L.; Horry, J.P.; Bakry, F.; Lanaud, C.; de León, D.G. A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). *Theor. Appl. Genet.* **1993**, *87*, 517–526.
55. Hippolyte, I.; Bakry, F.; Seguin, M.; Gardes, L.; Rivallan, R.; Risterucci, A.M.; Jenny, C.; Perrier, X.; Carreel, F.; Argout, X.; et al. A saturated SSR/DArT linkage map of *Musa acuminata* addressing genome rearrangements among bananas. *BMC Plant Biol.* **2010**, *10*, 65.
56. Mbanjo, E.G.N.; Tchoumbougnang, F.; Mouelle, A.S.; Oben, J.E.; Nyine, M.; Dochez, C.; Ferguson, M.E.; Lorenzen, J. Molecular marker-based genetic linkage map of a diploid banana population (*Musa acuminata* Colla). *Euphytica* **2012**, *188*, 369–386.
57. Noumbissié, G.B.; Chabannes, M.; Bakry, F.; Ricci, S.; Cardi, C.; Njembele, J.C.; Yohoume, D.; Tomekpe, K.; Iskra-Caruana, M.L.; D'Hont, A.; et al. Chromosome segregation in an allotetraploid banana hybrid (AAAB) suggests a translocation between the A and B genomes and results in eBSV-free offsprings. *Mol. Breed.* **2016**, *36*, 38–52.

58. Carreel, F.; Gonzalez de Leon, D.; Lagoda, P.; Lanaud, C.; Jenny, C.; Horry, J.; Tezenas du Montcel, H. Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome* **2002**, *45*, 679–692.
59. Hřibová, E.; Čížková, J.; Christelová, P.; Taudien, S.; De Langhe, E.; Doležel, J. The ITS-5.8S-ITS2 sequence region in the Musaceae: Structure, diversity and use in molecular phylogeny. *PLoS ONE* **2011**, *6*, e17863.
60. Dupouy, M.; Baurens, F.C.; Derouault, P.; Hervouet, C.; Cardi, C.; Cruaud, C.; Istace, B.; Labadie, K.; Guiougou, C.; Toubi, L.; et al. Two large reciprocal translocations characterized in the disease resistance-rich *burmannica* genetic group of *Musa acuminata*. *Ann. Bot.* **2019**, *124*, 31–329.
61. Martin, G.; Baurens, F.C.; Droc, G.; Rouard, M.; Cenci, A.; Kilian, A.; Hastie, A.; Doležel, J.; Aury, J.M.; Alberti, A.; et al. Improvement of the banana "*Musa acuminata*" reference sequence using NGS data and semi-automated bioinformatics methods. *BMC Genomics* **2016**, *17*, 1–12.
62. Horry, J.P. Chimiotaxonomie et organisation génétique dans le genre *Musa* (III). *Fruits* **1989**, *44*, 573–578.
63. Lebot, V.; Aradhya, M.K.; Manshardt, R.M.; Meilleur, B.A. Genetic relationships among cultivated bananas and plantains from Asia and the Pacific. *Euphytica* **1993**, *67*, 163–175.
64. Hippolyte, I.; Jenny, C.; Gardes, L.; Bakry, F.; Rivallan, R.; Pomies, V.; Cubry, P.; Tomekpe, K.; Risterucci, A.M.; Roux, N.; et al. Foundation characteristics of edible *Musa* triploids revealed from allelic distribution of SSR markers. *Ann. Bot.* **2012**, *109*, 937–951.
65. Li, L.F.; Wang, H.Y.; Zhang, C.; Wang, X.F.; Shi, F.X.; Chen, W.N.; Ge, X.J. Origins and domestication of cultivated banana inferred from chloroplast and nuclear genes. *PLoS ONE* **2013**, *8*, e80502.
66. Sandoval, J.A.; Côte, F.X.; Escoute, J. Chromosome number variations in micropropagated true-to-type and off-type banana plants (*Musa* AAA Grande Naine cv.). *In Vitro Cell. Dev. Biol. Plant* **1996**, *32*, 14–17.
67. Shepherd, K.; Da Silva, K.M. Mitotic instability in banana varieties. Aberrations in conventional triploid plants. *Fruit*s **1996**, *51*, 99–103.
68. Bartoš, J.; Alkhimova, O.; Doleželová, M.; De Langhe, E.; Doležel, J. Nuclear genome size and genomic distribution of ribosomal DNA in *Musa* and *Ensete* (Musaceae): Taxonomic implications. *Cytogenet. Genome Res.* **2005**, *109*, 50–57.
69. Čížková, J.; Hřibová, E.; Christelová, P.; Van den Houwe, I.; Häkkinen, M.; Roux, N.; Swennen, R.; Doležel, J. Molecular and cytogenetic characterization of wild *Musa* species. *PLoS ONE* **2015**, *10*, e0134096.
70. Roux, N.; Toloza, A.; Radecki, Z.; Zapata-Arias, F.J.; Doležel, J. Rapid detection of aneuploidy in *Musa* using flow cytometry. *Plant Cell Rep.* **2003**, *21*, 483–490.
71. Sedlazeck, F.J.; Rescheneder, P.; Smolka, M.; Fang, H.; Nattestad, M.; von Haeseler, A.; Schatz, M.C. Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* **2018**, *15*, 461–468.
72. Hu, L.; Liang, F.; Cheng, D.; Zhang, Z.; Yu, G.; Zha, J.; Wang, Y.; Xia, Q.; Yuan, D.; Tan, Y.; et al. Location of balanced chromosome-translocation breakpoints by long-read sequencing on the Oxford Nanopore platform. *Front. Genet.* **2020**, *10*, 1330.
73. Soto, D.C.; Shew, C.; Mastoras, M.; Schmidt, J.M.; Sahasrabudhe, R.; Kaya, G.; Andrés, A.M.; Dennis, M.Y. Identification of structural variations in chimpanzees using optical mapping and nanopore sequencing. *Genes* **2020**, *11*, 276.
74. Perrier, X.; Jacquemoud-Collet, J.P. DARwin software. Available online: http://darwin.cirad.fr/ 2006 (accessed on 20 May 2020).
75. FigTree v1.4.0. Available online: http://tree.bio.ed.ac.uk/software/figtree/ (accessed on 20 May 2020).
76. D'Hont, A.; Denoeud, F.; Aury, J.M.; Baurens, F.C.; Carreel, F.; Garsmeur, O.; Noel, B.; Bocs, S.; Droc, G.; Rouard, M.; et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* **2012**, *488*, 213–217.
77. Doležel, J.; Doleželová, M.; Roux, N.; Van den Houwe, I. A novel method to prepare slides for high resolution chromosome studies in *Musa* spp. *Infomusa* **1998**, *7*, 3–4.

# APPENDIX III

**Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate**

Abrouk, M., Ahmed, H.I., Cubry, P., Šimoníková, D., Cauet, S., Bettgenhaeuser, J., Gapa, L., Pailles, Y., Scarcelli, N., Couderc, M., Zekraoui, L., Kathiresan, N., Čížková, J., Hřibová, E., Doležel, J., Arribat, S., Bergès, H., Wieringa, J.J., Gueye, M., Kane, N.A., Leclerc, C., Causse, S., Vancoppenolle, S., Billot, C., Wicker, T., Vigouroux, Y., Barnaud, A., Krattinger, S.G.

*Nature communications* **11**: 4488, 2020

doi: 10.1038/s41467-020-18329-4

IF: 12.121

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-020-18329-4 OPEN

# Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate

Michael Abrouk[1,14], Hanin Ibrahim Ahmed[1,14], Philippe Cubry[2,14], Denisa Šimoníková[3], Stéphane Cauet[4], Yveline Pailles[1], Jan Bettgenhaeuser[1], Liubov Gapa[1], Nora Scarcelli[2], Marie Couderc[2], Leila Zekraoui[2], Nagarajan Kathiresan[5], Jana Čížková[3], Eva Hřibová[3], Jaroslav Doležel[3], Sandrine Arribat[4], Hélène Bergès[4,6], Jan J. Wieringa[7], Mathieu Gueye[8], Ndjido A. Kane[9,10], Christian Leclerc[11,12], Sandrine Causse[11,12], Sylvie Vancoppenolle[11,12], Claire Billot[11,12], Thomas Wicker[13], Yves Vigouroux[2], Adeline Barnaud[2,10]✉ & Simon G. Krattinger[1]✉

Sustainable food production in the context of climate change necessitates diversification of agriculture and a more efficient utilization of plant genetic resources. Fonio millet (*Digitaria exilis*) is an orphan African cereal crop with a great potential for dryland agriculture. Here, we establish high-quality genomic resources to facilitate fonio improvement through molecular breeding. These include a chromosome-scale reference assembly and deep re-sequencing of 183 cultivated and wild *Digitaria* accessions, enabling insights into genetic diversity, population structure, and domestication. Fonio diversity is shaped by climatic, geographic, and ethnolinguistic factors. Two genes associated with seed size and shattering showed signatures of selection. Most known domestication genes from other cereal models however have not experienced strong selection in fonio, providing direct targets to rapidly improve this crop for agriculture in hot and dry environments.

[1] Center for Desert Agriculture, Biological and Environmental Science & Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [2] DIADE, Univ Montpellier, IRD, Montpellier, France. [3] Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Hana for Biotechnological and Agricultural Research, Olomouc, Czech Republic. [4] CNRGV Plant Genomics Center, INRAE, Toulouse, France. [5] Supercomputing Core Lab, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [6] Inari Agriculture, One Kendall Square Building 600/700, Cambridge, MA 02139, USA. [7] Naturalis Biodiversity Center, Leiden, the Netherlands. [8] Laboratoire de Botanique, Département de Botanique et Géologie, IFAN Ch. A. Diop/UCAD, Dakar, Senegal. [9] Senegalese Agricultural Research Institute, Dakar, Senegal. [10] Laboratoire Mixte International LAPSE, Dakar, Senegal. [11] CIRAD, UMR AGAP, Montpellier, France. [12] AGAP, Université de Montpellier, Cirad, INRAE, Institut Agro, Montpellier, France. [13] Department of Plant and Microbial Biology, University of Zurich, Zürich, Switzerland. [14] These authors contributed equally: Michael Abrouk, Hanin Ibrahim Ahmed, Philippe Cubry. ✉email: adeline.barnaud@ird.fr; simon.krattinger@kaust.edu.sa

Humanity faces the unprecedented challenge of having to sustainably produce healthy food for 9–10 billion people by 2050 in a context of climate change. A more efficient use of plant diversity and genetic resources in breeding has been recognized as a key priority to diversify and transform agriculture[1–3]. The Food and Agriculture Organization of the United Nations (FAO) stated that arid and semi-arid regions are the most vulnerable environments to increasing uncertainties in regional and global food production[4]. In most countries of Africa and the Middle East, agricultural productivity will decline in the near future[4], because of climate change, land degradation, and groundwater depletion[5]. Agricultural selection, from the early steps of domestication to modern-day crop breeding, has resulted in a marked decrease in agrobiodiversity[6,7]. Today, three cereal crops alone, bread wheat (*Triticum aestivum*), maize (*Zea mays*), and rice (*Oryza sativa*) account for more than half of the globally consumed calories[8].

Many of today's major cereal crops, including rice and maize, originated in relatively humid tropical and sub-tropical regions[9,10]. Although plant breeding has adapted the major cereal crops to a wide range of climates and cultivation practices, there is limited genetic diversity within these few plant species for cultivation in the most extreme environments. On the other hand, crop wild relatives and orphan crops are often adapted to extreme environments and their utility to unlock marginal lands for agriculture has recently regained interest[2,6,11–14]. Current technological advances in genomics and genome editing provide an opportunity to rapidly domesticate wild relatives and to improve orphan crops[15,16]. De novo domestication of wild species or rapid improvement of semi-domesticated crops can be achieved in less than a decade by targeting a few key genes[6].

White fonio (*Digitaria exilis* (Kippist) Stapf) (Fig. 1) is an indigenous West African millet species with a great potential for agriculture in marginal environments[17,18]. Fonio is a small annual herbaceous C4 plant, which produces very small (~1 mm) grains that are tightly surrounded by a husk[19] (Fig. 1). Fonio is cultivated under a large range of environmental conditions, from a tropical monsoon climate in western Guinea to a hot, arid desert climate in the Sahel zone. Some extra-early maturing fonio varieties produce mature grains in only 70–90 days[19], which makes fonio one of the fastest maturing cereals. Because of its quick maturation, fonio is often grown to avoid food shortage during the lean season (period before main harvest), which is why fonio is also referred to as 'hungry rice'. In addition, fonio is drought tolerant and adapted to nutrient-poor, sandy soils[20]. Despite its local importance for agriculture in West Africa, fonio shows many unfavorable characteristics that resemble undomesticated plants: residual seed shattering, lodging, and lower yields than other cereals[18] (Supplementary Fig. 1). The most likely wild progenitor of cultivated fonio is the tetraploid annual weed *D. longiflora* that is widely distributed in tropical Africa. It has been suggested that fonio was domesticated more than 5000 years ago in the Inner Niger Delta of central Mali. Compared to *D. exilis*, the wild *D. longiflora* shows even heavier seed shattering and hairy spikelets, which suggests that some selection for reduced seed shattering and spikelet hairiness occurred in fonio[21]. In the past few years, fonio has gained in popularity inside and outside of West Africa because of its nutritional qualities.

Here, we present the establishment of a comprehensive set of genomic resources for fonio, which constitutes the first step towards harnessing the potential of this cereal crop for agriculture in harsh environments. These resources include the generation of a high-quality, chromosome-scale reference assembly and the deep re-sequencing of a diversity panel that includes wild and cultivated accessions covering a wide geographic range.

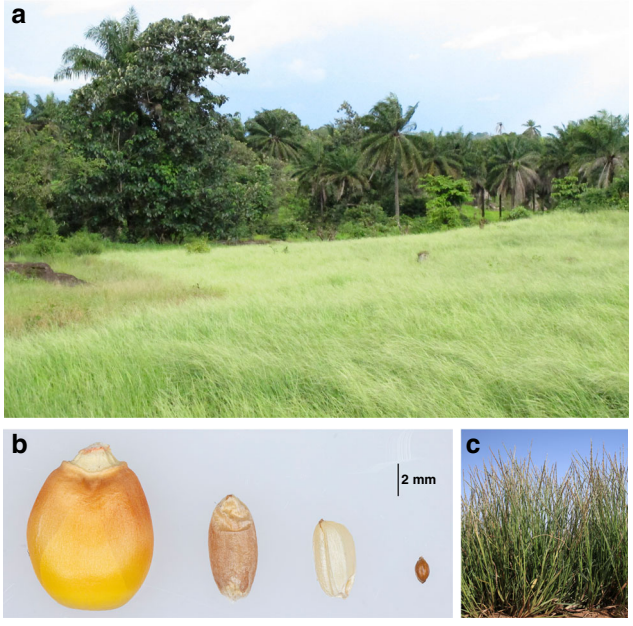



**Fig. 1 Phenotype of fonio (*Digitaria exilis*). a** Field of cultivated fonio in Guinea. **b** Grains of maize, wheat, rice, and fonio (from left to right). **c** Plants of the fonio accession CM05836.

## Results

**Chromosome-scale fonio reference genome assembly**. Fonio is a tetraploid species ($2n = 4\times = 36$)[22] with a highly inbreeding reproductive system[17]. To build a *D. exilis* reference assembly, we chose an accession from one of the driest regions of fonio cultivation, CM05836 from the Mopti region in Mali. The size of the CM05836 genome was estimated to be 893 Mb/1C by flow cytometry (Supplementary Figs. 2 and 3), which is in line with previous reports[22]. The CM05836 genome was sequenced and assembled using deep sequencing of multiple short-read libraries (Supplementary Table 1), including Illumina paired-end (321-fold coverage), mate-pair (241-fold coverage) and linked-read (10× Genomics, 84-fold coverage) sequencing. The raw reads were assembled and scaffolded with the software package DeNovoMAGIC3 (NRGene), which has recently been used to assemble various high-quality plant genomes[23–25]. Integration of Hi-C reads (122-fold coverage, Supplementary Table 2) and a Bionano optical map (Supplementary Table 3) resulted in a chromosome-scale assembly with a total length of 716,471,022 bp, of which ~91.5% (655,723,161 bp) were assembled in 18 pseudomolecules. A total of 60.75 Mb were unanchored (Table 1). Of 1440 Embryophyta single copy core genes (BUSCO v3.0.2), 96.1% were recovered in the CM05836 assembly, 2.9% were missing, and 1% was fragmented. As no genetic *D. exilis* map is available, we used chromosome painting to further assess the quality of the CM05836 assembly. Pools of short oligonucleotides covering each one of the 18 pseudomolecules were designed based on the CM05836 assembly, fluorescently labeled, and hybridized to mitotic metaphase chromosome spreads of CM05836[26]. Each of the 18 libraries specifically hybridized to only one chromosome pair, confirming that our assembly unambiguously distinguished homoeologous chromosomes (Fig. 2a, Supplementary Fig. 4). Centromeric regions contained a tandem repeat with a 314 bp long unit, which was found in all fonio chromosomes (Supplementary Fig. 5). We also re-assembled all the data with the open-source TRITEX pipeline[27] and the two assemblies showed a high degree of collinearity (Supplementary Fig. 6).

**Table 1 Statistics of the fonio genome assembly and annotation.**

| | CM05836 |
|---|---|
| *Length of DeNovoMAGIC3 assembly (Mb)* | *701.662* |
| Number of scaffolds | 8457 |
| N50 (Mb) | 10.741 |
| N90 (Mb) | 1.009 |
| *Length of genome assembly (Mb)* | *716.471*[a] |
| *Total length of pseudomolecules (Mb)* | *655.723* |
| Number of anchored contigs | 18,026 |
| N50 of anchored contigs (kb) | 83.702 |
| Gap size (Mb) | 17.001 (2.6%) |
| Number of genes | 57,021 |
| *Total length of unanchored chromosomes (Mp)* | *60.748* |
| Number of unanchored contigs | 11,129 |
| N50 of unanchored contigs (kb) | 9.191 |
| Gap size (Mb) | 2.962 (4.9%) |
| Number of genes | 2821 |
| *BUSCO* | |
| Complete (%) | 96.1 |
| Duplicated (%) | 84.3 |
| Fragmented (%) | 1 |
| Missing (%) | 2.9 |

[a]DeNovoMAGIC3 + Hi-C + optical map.

We compared the fonio pseudomolecule structure to foxtail millet (*Setaria italica*; $2n = 2\times = 18$), a diploid relative with a fully sequenced genome[28]. The fonio genome shows a syntenic relationship with the genome of foxtail millet, with two homoeologous sets of nine fonio chromosomes (Supplementary Fig. 7). Without a clear diploid ancestor, we could not directly disentangle the two sub-genomes based on genomic information from diploid ancestors[29]. We thus used a genetic structure approach based on full-length long terminal repeat retrotransposons (fl-LTR-RT) as an alternative strategy. A total of 11 fl-LTR-RT families with more than 30 elements were identified and defined as a 'populations', allowing us to apply genetic structure analyses that are often used in population genomics[30]. We searched for fl-LTR-RT clusters that only appeared in one of the two homoeologous sub-genomes. Out of the 11 fl-LTR-RT populations analyzed, two allowed us to discriminate the sub-genomes (Fig. 2b, Supplementary Fig. 8). The two families belong to the Gypsy superfamily and dating of insertion time was estimated to be ~1.56 million years ago (MYA) (58 elements, 0.06–4.67 MYA) and ~1.14 MYA (36 elements, 0.39–1.96 MYA), respectively. The two putative sub-genomes were designated A and B and chromosome numbers were assigned based on the synteny with foxtail millet (Supplementary Fig. 7).

Gene annotation was performed using the MAKER pipeline (v3.01.02) with 34.1% of the fonio genome masked as repetitive. Transcript sequences of CM05836 from flag leaves, grains, panicles, and whole above-ground seedlings (Supplementary Table 4) in combination with protein sequences of publicly available plant genomes were used to annotate the CM05836 assembly. This resulted in the annotation of 59,844 protein-coding genes (57,023 on 18 pseudomolecules and 2821 on unanchored chromosome) with an average length of 2.5 kb and an average exon number of 4.6. The analysis of the four CM05836 RNA-seq samples showed that 44,542 protein coding genes (74.3%) were expressed (>0.5 transcripts per million), which is comparable to the annotation of the bread wheat genome (Supplementary Table 5)[31,32].

**Synteny with other cereals**. Whole genome comparative analyses of the CM05836 genome with other grass species were consistent with the previously established phylogenetic relationships of fonio[20]. A comparison of the CM05836 A and B sub-genomes identified a set of 16,514 homoeologous gene pairs that fulfilled the criteria for evolutionary analyses (Fig. 2c and Supplementary Data 1). The estimation of synonymous substitution rates (Ks) among homoeologous gene pairs revealed a divergence time of the two sub-genomes of roughly 3 MYA (Supplementary Fig. 9 and Supplementary Data 1). These results indicate that *D. exilis* is a recent allotetraploid species. A Ks distribution using orthologous genes revealed that fonio diverged from the other members of the Paniceae tribe (broomcorn millet (*Panicum miliaceum*), Hall's panicgrass (*P. hallii*), and foxtail millet (*S. italica*)) between 14.6 and 16.9 MYA, and from the Andropogoneae tribe (sorghum (*Sorghum bicolor*) and maize (*Z. mays*)) between 21.5 and 26.9 MYA. Bread wheat (*T. aestivum*), goatgrass (*Aegilops tauschii*), barley (*Hordeum vulgare*), rice (*O. sativa*), and purple false brome (*Brachypodium distachyon*) showed a divergence time of 35.3–40 MYA (Fig. 2d). The phylogenetic position of *D. exilis* as a basal taxon of the Paniceae allowed us to reconstruct the hypothetical ancestral genomic state of this tribe (Supplementary Data 2).

The hybridization of two genomes can result in genome instability, extensive gene loss, and reshuffling. As a consequence, one sub-genome may evolve dominancy over the other sub-genome[24,31,33–35]. Using foxtail millet as a reference, the fonio A and B sub-genomes showed similar numbers of orthologous genes: 14,235 and 14,153, respectively. Out of these, 12,541 were retained as duplicates while 1694 and 1612 were specific to the A and B sub-genomes, respectively. The absence of sub-genome dominance was also supported by similar gene expression levels between homoeologous pairs of genes (Mann–Whitney *U* test; $p = 0.66$; Supplementary Fig. 10 and Supplementary Data 3).

**Evolutionary history of fonio and its wild relative**. To get an overview of the diversity and evolution of fonio, we selected 166 *D. exilis* accessions originating from Guinea, Mali, Benin, Togo, Burkina Faso, Ghana, and Niger and 17 accessions of the proposed wild tetraploid fonio progenitor *D. longiflora*[21] from Cameroon, Nigeria, Guinea, Chad, Sudan, Kenya, Gabon, and Congo for whole-genome re-sequencing (Supplementary Table 6). The selection was done from a collection of 641 georeferenced *D. exilis* accessions[36] with the aim of maximizing diversity based on bioclimatic data and geographic origin. We obtained short-read sequences with an average of 45-fold coverage for *D. exilis* (range 36–61-fold) and 20-fold coverage for *D. longiflora* (range 10–28-fold) (Supplementary Data 4). The average mapping rates of the raw reads to the CM05836 reference assembly were 85% and 68% for *D. exilis* and *D. longiflora*, respectively, with most accessions showing a mapping rate of >80% (Supplementary Data 5). After filtering, 11,046,501 high quality bi-allelic single nucleotide polymorphisms (SNPs) were retained (Supplementary Table 7). Nine *D. exilis* and three *D. longiflora* accessions were discarded based on the amount of missing data. The error rate of variant calling (proportion of segregating sites in a re-sequenced CM05836 sample compared to the reference assembly) was 0.04%, which is comparable to other studies[37]. The re-sequenced CM05836 sample showed the lowest genetic divergence from the reference assembly of all re-sequenced accessions (Supplementary Fig. 11a). The SNPs were evenly distributed across the 18 *D. exilis* chromosomes, with a tendency toward a lower SNP density at the chromosome ends (Supplementary Fig. 11b). Approximately 30.2% of the SNPs were gene-proximal (2 kb upstream or downstream of a coding sequence), 9.5% in introns and 6.2% in exons. Of the exonic SNPs, 354,854 (51.6%) resulted in non-synonymous sequence changes, of which 6727 disrupted the coding sequence (premature

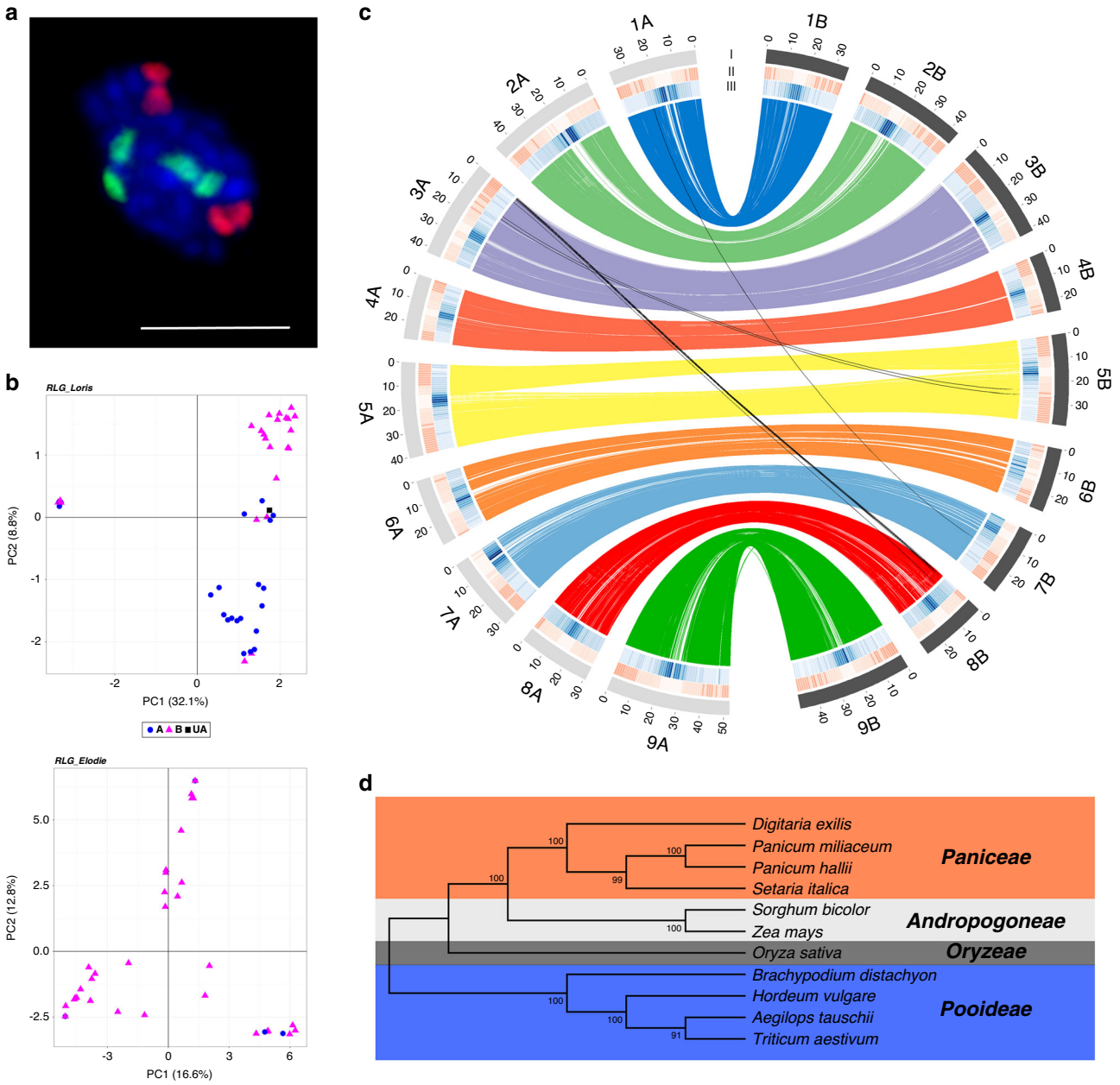



**Fig. 2 Fonio genome features. a** Representative example of oligo painting FISH on mitotic metaphase chromosomes. Shown are probes designed from pseudomolecules 9A (green) and 9B (red) of the CM05836 assembly (scale bar = 5 µm). Oligo painting FISH experiment was repeated independently three times. **b** Principal component analysis (PCA) of the transposable element cluster RLG_Loris (upper panel) that allowed discrimination of the two sub-genomes. Blue dots represent elements found on the A sub-genome; pink triangles represent elements from the B sub-genome; black squares represent elements present on chromosome unanchored. PCA of the transposable element cluster RLG_Elodie (lower panel) that was specific to the B sub-genome. **c** Synteny and distribution of genome features. (I) Number and length of the pseudomolecules. The gray and black colors represent the two sub-genomes. (II, III) Density of genes and repeats along each pseudomolecule, respectively. Lines in the inner circle represent the homoeologous relationships. **d** Maximum likehood tree of 11 *Poaceae* species based on 30 orthologous gene groups. Topologies are supported by 1000 bootstrap replicates. Colors indicate the different clades.

stop codon). The remaining 333,296 (48.4%) exonic SNPs represented synonymous variants. Forty-four percent of the total SNPs (4,901,160 SNPs) were rare variants with a minor allele frequency of <0.01 (Supplementary Fig. 12). The vast majority of the rare variants was found in a few very diverse *D. longiflora* accessions. The mean nucleotide diversity ($\pi$) was $6.19 \times 10^{-4}$ and $3.68 \times 10^{-3}$ for *D. exilis* and *D. longiflora*, respectively. Genome-wide linkage disequilibrium (LD) analyses revealed a faster LD decay in *D. longiflora* ($r^2 \sim 0.16$ at 70 kb) compared to *D. exilis* ($r^2 \sim 0.20$ at 70 kb) (Supplementary Fig. 13).

A PCA showed a clear genetic differentiation between cultivated *D. exilis* and wild *D. longiflora*. The *D. exilis* accessions clustered closely together, while the *D. longiflora* accessions split into three groups (Fig. 3a). The *D. longiflora* group that showed the greatest genetic distance from *D. exilis* contained three accessions originating from Central (Cameroon) and East Africa (South Sudan and Kenya). The geographical projection of the first principal component (which separated wild accessions from the cultivated accessions) revealed that the *D. exilis* accessions genetically closest to

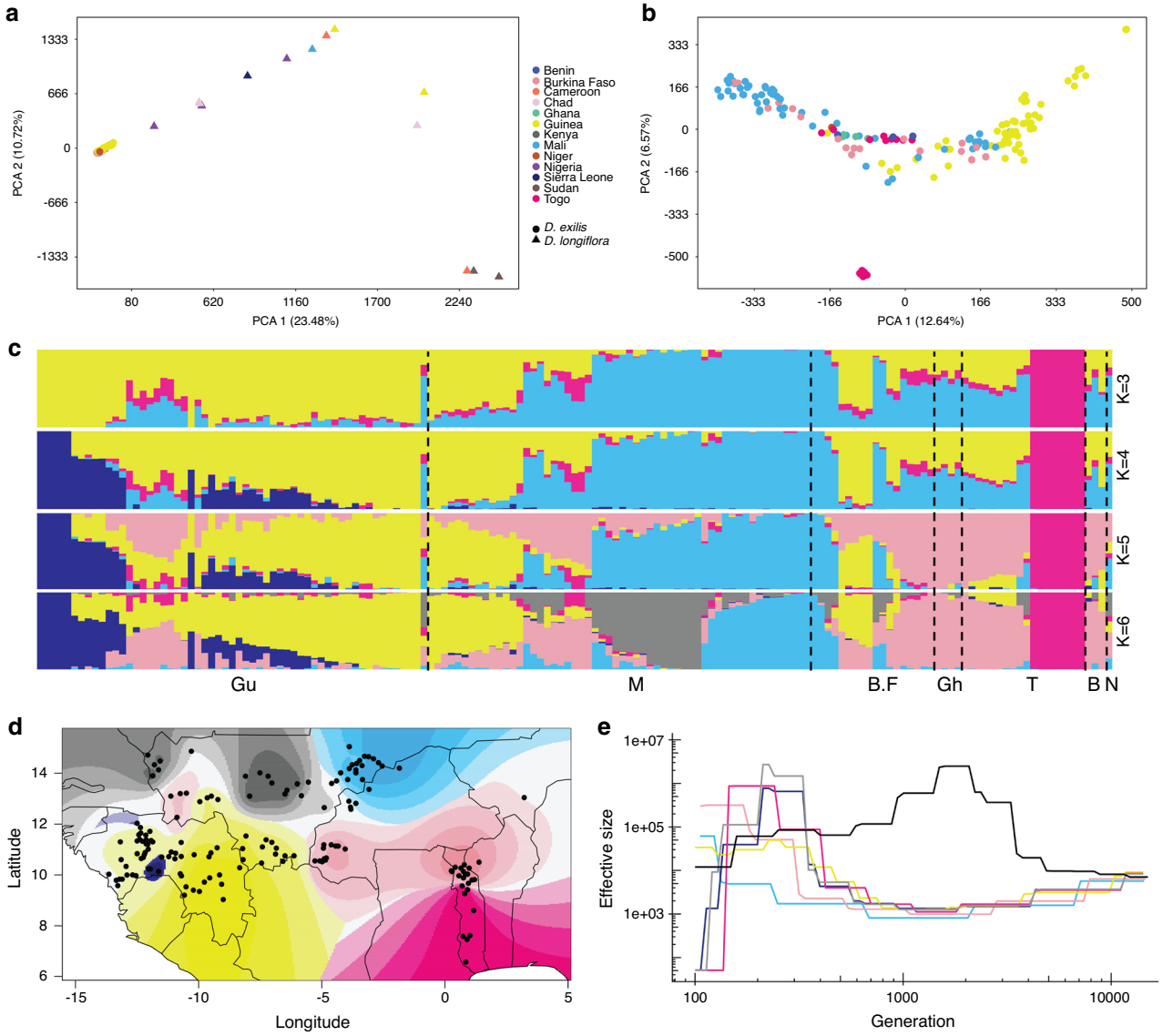



**Fig. 3 Genetic diversity and structure of *D. exilis* and *D. longiflora* diversity panel. a** Principal component analysis (PCA) of 157 *D. exilis* and 14 *D. longiflora* accessions using whole-genome single nucleotide polymorphisms (SNPs). *D. exilis* samples (circles), *D. longiflora* (triangles). **b** PCA of *D. exilis* accessions alone. Colors indicate the country of origin. **c** Population structure (from $K = 3$ to $K = 6$) of *D. exilis* accessions estimated with sNMF. Each bar represents an accession and the bars are filled by colors representing the likelihood of membership to each ancestry. Accessions are ordered from west to east; Guinea (Gu), Mali (M), Burkina Faso (B.F), Ghana (Gh), Togo (T), Benin (B), and Niger (N). **d** Geographic distribution of ancestry proportions of *D. exilis* accessions obtained from the structure analysis at $K = 6$. The colors represent the maximal local contribution of an ancestry. Black dots represent the coordinates of *D. exilis* accessions. **e** Effective population size history of *D. exilis* groups based on $K = 6$ and *D. longiflora* (in black).

*D. longiflora* originated from southern Togo and the west of Guinea (Supplementary Fig. 14a, b).

A PCA with *D. exilis* accessions alone revealed three main clusters. The second principal component separated eight accessions from southern Togo from the remaining accessions (Fig. 3b). In addition, five accessions from Guinea formed a distinct genetic group in the PCA. The remaining accessions were spread along the first axis of the PCA, mainly revealing a grouping by geographic location (Fig. 3b). The genetic clustering was confirmed by genetic structure analyses (Fig. 3c, d). The cross-validation error decreased with increasing $K$ and reached a plateau starting from $K = 6$ (Supplementary Fig. 14c). At $K = 3$, the eight South Togo accessions formed a distinct homogenous population. At $K = 4$, the five accessions from Guinea were separated. With increasing $K$, the admixture plot provided some evidence that natural (climate and geography) and human (ethnicity and language) factors had an effect on shaping the genetic population structure of fonio (Fig. 3c, d, Supplementary Fig. 14d). We observed a significant correlation (Pearson's correlation; $p < 0.05$, $df = 155$) between the genetic population structure (first principal component of PCA) and climate (i.e., mean temperature and precipitation of the wettest quarters) as well as geography (i.e., latitude, longitude, and altitude) (Fig. 3b, Supplementary Fig. 15, Supplementary Data 6). The relationship between genetic differentiation (genetic distance matrix) and climate was still significant when accounting for geographic distance (partial Mantel test; Mantel $r = 0.30$; $p = 0.001$). A significant correlation was also observed between the genetic distance matrix and the dissimilarity matrices of ethnic and linguistic groups (fisher test; $p = 0.0005$, Mantel test; Mantel $r$ ethnic $= -0.19$; Mantel $r$ linguistic $= -0.13$; $p = 0.001$). The effect of ethnic groups remained significant even when we

controlled for geographic and climatic factors (ANCOVA; $p = 0.045$; df = 31) (Supplementary Table 8). Association analyses based on climate variables revealed 38 and 179 loci that might be involved in adaptation to mean temperature and mean precipitation of the wettest quarters, respectively (Supplementary Fig. 16, Supplementary Data 7). Loci associated with temperature contained genes enriched for functions related to hormone metabolic processes, hormone biosynthesis, carbohydrate metabolic processes, homeostatic process, plant development processes, and plant growth (shoot apical meristem maintenance, root growth and development) (Supplementary Data 7). Similarly, genes associated with precipitation were enriched in functions related to plant development and growth (Supplementary Data 7). A genome-wide association study with ethnic groups revealed significant associations for 55 and 227 SNPs with Bambara and Fula ethnic groups, respectively (Supplementary Fig. 17, Supplementary Data 7). In particular, there were three prominent peaks on chromosomes 2A, 2B, and 6A. The peaks on chromosomes 2A and 2B fell into a region that contained a homolog of the Arabidopsis *WSD1* wax ester synthase/diacylglycerol acyltransferase gene[38]. *WSD1* is involved in accumulation of waxes under drought stress, possibly indicating selection for adaptation to drought by certain ethnic groups. Unadmixed populations were mainly found at the geographic extremes of the fonio cultivation area in the north and south, whereas the accessions from the central regions of West Africa tended to show a higher degree of admixture (Supplementary Fig. 18a).

Plotting the spatial distribution of private SNPs (i.e., SNP present only once in a single accession) revealed a hotspot of rare alleles in Togo, Niger, the western part of Guinea, and southern Mali (Supplementary Fig. 18b, c). Rare allele diversity was lower in the eastern part of Guinea, and in southwest Mali. Inference of the *D. exilis* effective population size revealed a decline that started more than 10,000 years ago and reached a minimum between 2000 and 1000 years ago (Fig. 3e). Then, a steep increase of the effective population size occurred to a level that was ~100-fold higher compared to the bottleneck.

**Genomic footprints of selection and domestication**. We used three complementary approaches to detect genomic regions under selection: (i) a composite likelihood-ratio (CLR) test based on site frequency spectrum (SFS)[39], (ii) the nucleotide diversity ($\pi$) ratios between *D. exilis* and *D. longiflora* over sliding genomic windows, and (iii) the genetic differentiation based on $F_{ST}$ calculations, again computed over sliding windows. With the CLR test, 78 regions were identified as candidates for signatures of selection. The genetic diversity ratios and $F_{ST}$ calculations revealed 311 and 208 regions, respectively (Fig. 4, Supplementary Data 8). We then searched for the presence of orthologs of known domestication genes in the regions under selection. The most striking candidate was one of the two orthologs of the rice grain size *GS5* gene[40] (Dexi3A01G0012320 referred to as *DeGS5-3A*) that was detected by genetic diversity ratio and $F_{ST}$-based

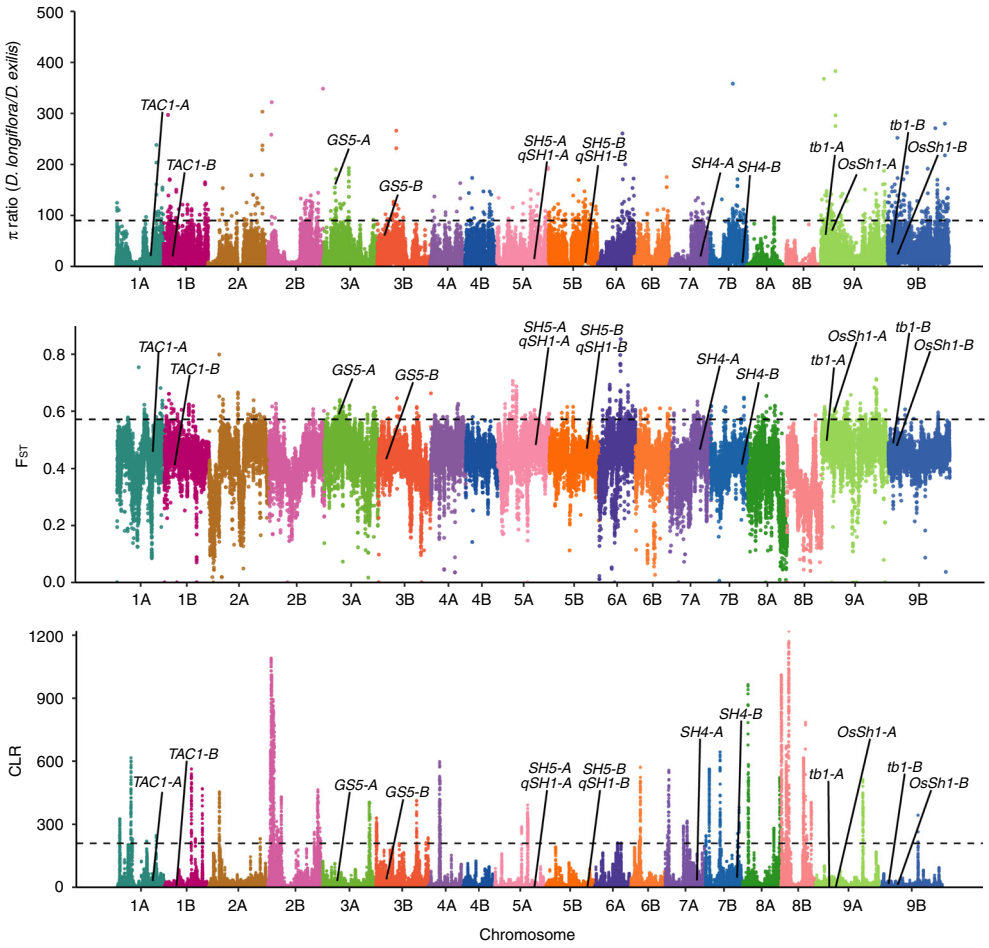



**Fig. 4 Detection of selection in fonio.** Manhattan plots showing detection of selection along the genome based on nucleotide diversity ($\pi$) ratio, $F_{ST}$ and SweeD (from top to bottom). The location of orthologous genes of major seed shattering and plant architecture genes are indicated in the Manhattan plots. The black dashed lines indicate the 1% threshold. Some extreme outliers in the $\pi$ ratio plot are not shown.

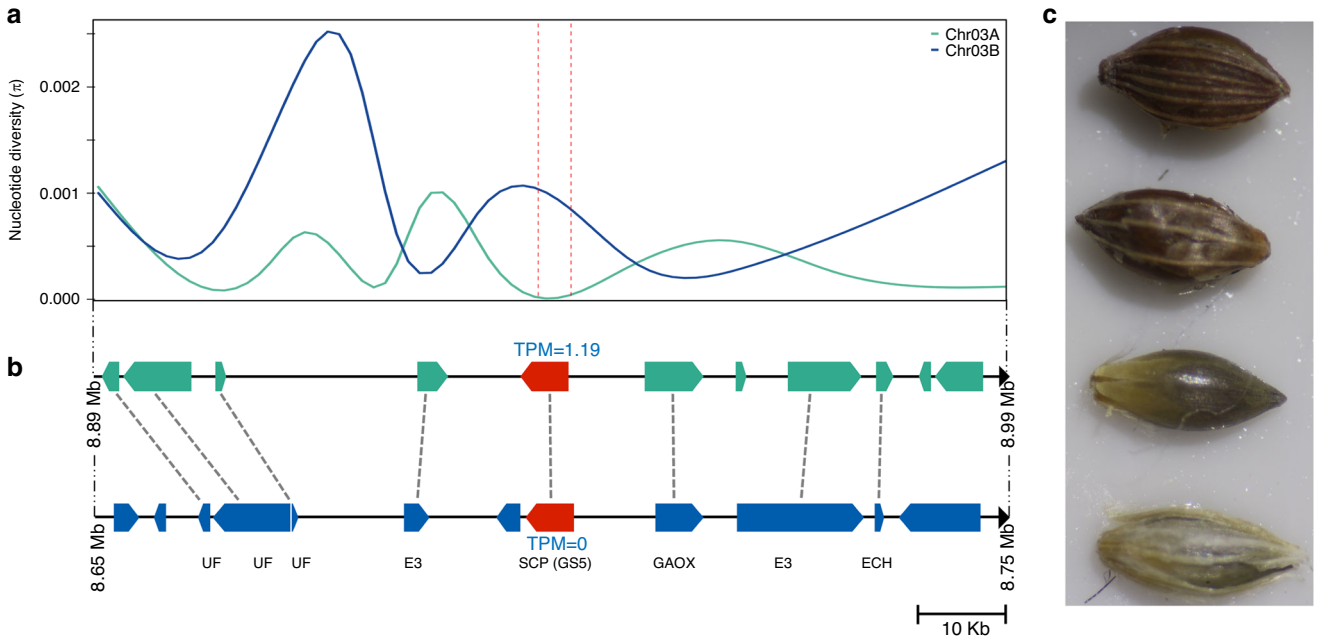



**Fig. 5 Selective sweep at the *GS5* locus in fonio. a** Smoothed representation of nucleotide diversity ($\pi$) in *D. exilis* in a 100 kb window surrounding the ortholog of the rice *GS5* gene. The green curve shows $\pi$ around *DeGS5-3A* on chromosome 3A and the blue curve shows $\pi$ around *DeGS5-3B* on chromosome 3B. The dashed vertical red lines represent the location of the *GS5* genes. The nucleotide diversity was calculated in overlapping 100 bp windows every 25 bp. **b** Schematic representation of the annotated genes in the *GS5* regions and the orthologous gene relationships between chromosomes 3A (green) and 3B (blue). *UF* protein of unknown function, *E3* E3 ubiquitin ligase, *SCP* (*GS5*) serine carboxypeptidase (Dexi3A01G0012320), *GAOX* gibberellin 2-beta-dioxygenase, *ECH* golgi apparatus membrane protein. *DeGS5-3A* and *DeGS5-3B* are indicated in red. The numbers in blue above and below the *GS5* orthologs show their respective expression in transcripts per million (TPM) in grain tissue. **c** Two grains of *D. exilis* (top) and *D. longiflora* (bottom).

calculations (Fig. 4). *GS5* regulates grain width and weight in rice. *D. exilis* showed a dramatic loss of genetic diversity at the *DeGS5-3A* gene (Fig. 5a, b). Domestication of fonio is associated with wider grains of *D. exilis* compared to grains of *D. longiflora* (Fig. 5c). The region of the *GS5* ortholog on chromosome 3B (*DeGS5-3B*) was not identified as being under selection and showed higher levels of nucleotide diversity than the *DeGS5-3A* region (Fig. 5a, b). Only the *DeGS5-3A* but not the *DeGS5-3B* transcript was detected in the *D. exilis* RNA-Seq data from grains. This is in agreement with observations made in rice[40], where a dominant mutation resulting in increased *GS5* transcript levels affects grain size.

Another domestication gene that was detected in the selection scan was an ortholog of the sorghum *Shattering 1* (*Sh1*) gene[41]. *Sh1* encodes a YABBY transcription factor and the non-shattering phenotype in cultivated sorghum is associated with lower expression levels of *Sh1* (mutations in regulatory regions or introns) or truncated transcripts. Domesticated African rice (*O. glaberrima*) carries a 45 kb deletion at the orthologous *OsSh1* locus compared to its wild relative *O. barthii*[42]. Around 37% of the fonio accessions had a 60 kb deletion similar to *O. glaberrima* that eliminated the *Sh1* ortholog on chromosome 9A (*DeSh1-9A*—Dexi9A01G0015055) (Fig. 6a). This deletion was identified by the lack of mapped reads and was confirmed by PCR amplification and Sanger sequencing. The homoeologous region including *DeSh1-9B* (Dexi9B01G0013485) on chromosome 9B was intact (Fig. 6b). Interestingly, accessions with the *DeSh1-9A* deletion showed a minor (7%) but significant reduction in seed shattering (Fig. 6c) compared to accessions with the intact *DeSh1-9A* gene (ANOVA; $p = 0.008$; df = 1). Accessions carrying this deletion were distributed across the whole range of fonio cultivation, which suggests that the deletion is ancient and might have been selected for in certain regions.

## Discussion

Here, we established a set of genomic resources that allowed us to comprehensively assess the genetic variation found in fonio, a cereal crop that holds great promises for agriculture in marginal environments. The analysis of fl-LTR-RT revealed two sub-genome specific transposon clusters that experienced a peak of activity around 1.1–1.5 MYA, indicating that the two fonio sub-genomes hybridized after this period[43]. The Ks analysis estimated that the two sub-genomes diverged prior to these transposon bursts—3 MYA, suggesting that fonio is an allotetraploid species.

The analysis of effective population size revealed a genetic bottleneck that was most likely associated with human cultivation and domestication. The large increase of effective size for fonio after a period of reduction was most probably due to the development and expansion of fonio cultivation. This expansion appears to be a recent event and occurred during the last millennium. The progression of the effective population size resembles the patterns observed for other domesticated crops, with a protracted period of cultivation followed by a marked bottleneck[44,45]. In the case of fonio, this potential bottleneck appears to be milder compared to other crops. It has been observed that the effective population size of the wild rice (*O. barthii*) from West Africa followed a trend similar to the cultivated African rice (*O. glaberrima*), which has been interpreted as a result of environmental degradation[44] rather than human selection. In contrast, no signal of population bottleneck was observed for the proposed wild fonio ancestor *D. longiflora*, indicating that the bottleneck observed in fonio is associated with human cultivation. We also highlighted the strong impact of geographic, climatic and anthropogenic factors on shaping the genetic diversity of fonio. Even if fonio is not a dominant crop across West Africa, it benefits from cultural embedding and plays a key role in ritual systems in many African cultures[46]. For

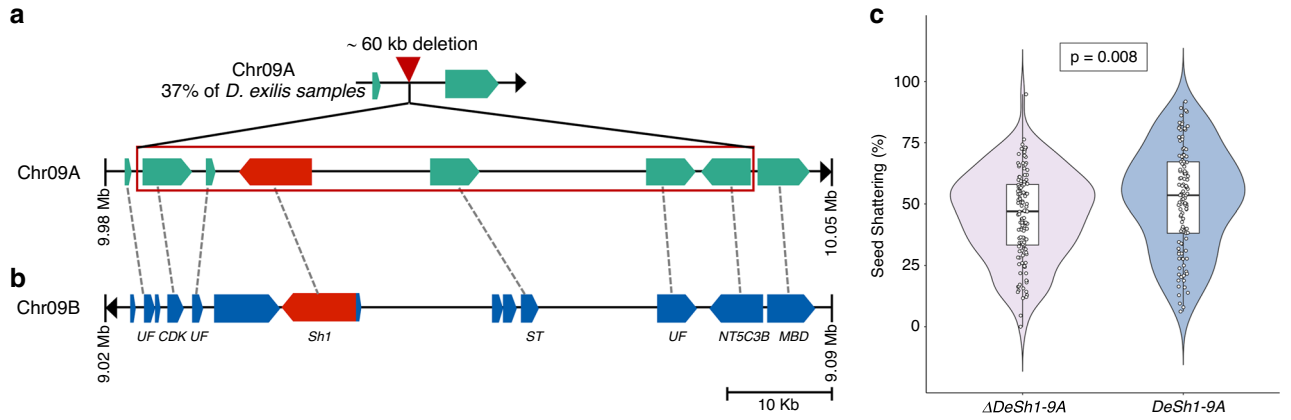



**Fig. 6 Selective sweep at the *Sh1* locus in fonio. a** Schematic representation of genes in the orthologous regions of the sorghum *Sh1* gene. The top most panel shows the region on fonio chromosome 9A with the 60 kb deletion as it is found in 37% of all *D. exilis* accessions. The lower panel shows the genes in accessions without the deletion. The *Sh1* ortholog *DeSh1-9A* is shown in red. **b** The *Sh1* ortholog *DeSh1-9B* on chromosome 9B is present in all *D. exilis* accessions. The dashed lines represent orthologous relationships between the two sub-genomes. *UF* protein of unknown function, *CDK* cyclin-dependent kinase, *Sh1* ortholog of sorghum *Shattering 1* (*Sh1*), *ST* sulfate transporter; *NT5C3B* 7-methylguanosine phosphate-specific 5′-nucleotidase; *MBD* methyl-CpG-binding domain. **c** The violin plots show the probability distribution of the seed shattering percentage in fonio accessions carrying the *DeSh1-9A* deletion (Δ*DeSh1-9A*) in pink compared to accessions with an intact *DeSh1-9A* in blue. The shape of the distribution indicates that the seed shattering percentages for both groups are concentrated around the median. The center of each plot depicts a boxplot, the box represents the interquartile range (IQR), the middle horizontal line represents the median, the vertical lines going down and up from the box are defined as first quartile (−1.5 IQR) and third quartile (+1.5 IQR), respectively. Circles in the plot represent the seed shattering percentage calculated from individual fonio panicles, one circle per panicle, jitter option was added to show all individual observations that overlapped. Mean seed shattering was 45% in Δ*DeSh1-9A* (±0.18 s.d.) and 52% in *DeSh1-9A* (±0.22 s.d.). ($n$ = Three individual panicles of 43 Δ*DeSh1-9A* accessions and 39 *DeSh-9A* accessions, respectively; two-way ANOVA $p = 0.008$; df = 1).

example, we observed a striking genetic differentiation between *D. exilis* accessions collected from northern and southern Togo. This can be attributed to both climatic and cultural differences. While the southern regions of Togo receive high annual rainfalls, the northern parts receive <1000 mm annual rainfall and there is a prolonged dry season[47]. Adoukonou-Sagbadja et al.[47] also noted that there is no seed exchange between farmers of the two regions because of cultural factors.

Despite a genetic bottleneck and the reduction in genetic diversity, fonio still shows many 'wild' characteristics such as residual seed shattering, lack of apical dominance, and lodging. We show that orthologs of most of the well-characterized domestication genes from other cereals were not under strong selection in fonio. Interestingly, an ortholog of the rice *GS5* gene (*DeGS5-3A*) was identified in a selective sweep. Dominant mutations in the *GS5* promoter region were associated with higher *GS5* transcript levels and wider and heavier grains in rice[40]. The *DeGS5-3A* gene showed a complete loss of diversity in the coding sequence in *D. exilis*, suggesting a strong artificial selection for larger grains. In contrast, *DeSh1-9A* showed evidence for a partial selective sweep[48]. The non-shattering phenotype associated with the *Sh1* locus in sorghum is recessive[41] and the deletion of a single *DeSh1* copy only resulted in a quantitative reduction of shattering that might have been selected for in some but not all regions. Whether the 60 kb deletion including *DeSh1-9A* represents standing genetic variation or arose after fonio domestication cannot be determined. The deletion was not identified in any of the 14 re-sequenced *D. longiflora* accessions. Targeting the *DeSh1-9B* locus on chromosome 9B in accessions that carry the *DeSh1-9A* deletion through mutagenesis or genome editing could produce a fonio cultivar with significantly reduced seed shattering, which would form a first important step towards a significant improvement of this crop.

## Methods

**Plant material**. The plant material used in this study comprised a collection of 641 *D. exilis* accessions maintained at the French National Research Institute for Development (IRD), Montpellier, France[36,49]. The collection was established from 1977 to 1988. For genome re-sequencing, a panel of 166 accessions (Supplementary Data 6) was selected based on geographic and bioclimatic data. One plant per accession was chosen for DNA extraction and sequencing. Inflorescences of sequenced *D. exilis* plants were covered with bags to prevent outcrossing. Grains of sequenced plants were collected and kept for further analyses. In addition, 17 accessions of *D. longiflora* were collected from specimens stored at the National Museum of Natural History of Paris (Paris, France), IFAN Ch. A. Diop (Dakar, Senegal), National Herbarium of The Netherlands (Leiden, Netherlands), and Cirad (Montpellier, France) (Supplementary Table 6).

**Flow cytometry for genome size estimation**. The amount of nuclear DNA in fonio was measured by flow-cytometry[50]. In brief, six different individual plants representing accession CM05836 were measured three times on three different days using a CyFlow® Space flow cytometer (Sysmex Partec GmbH, Görlitz, Germany) equipped with a 532 nm green laser. Soybean (*Glycine max* cultivar Polanka; 2C = 2.5 pg) was used as an internal reference standard. The gain of the instrument was adjusted so that the peak representing G1 nuclei of the standard was positioned approximately on channel 100 on a histogram of relative fluorescence intensity when using a 512-channel scale. The low level threshold was set to channel 20 to eliminate particles with the lowest fluorescent intensity from the histogram. All remaining fluorescent events were recorded with no further gating used. 2C DNA contents (in pg) were calculated from the mean G1 peak (interphase nuclei in G1 phase) positions by applying the formula:

$$\text{2C nuclear DNA content} = \frac{(\text{sample G1 peak mean}) \times (\text{standard 2C DNA content})}{(\text{standard G1 peak mean})}. \quad (1)$$

DNA contents in pg were converted to genome sizes in bp using the conversion factor 1 pg DNA = 0.978 Gb[51].

**Reference genome sequencing and assembly**. High molecular weight (HMW) genomic DNA was isolated from a single CM05836 plant with a dark-treatment of 48 h before harvesting tissue from young leaves. Two paired-end and three mate-pair libraries were constructed with different insert sizes ranging from 450 bp to 10 kb. The 450 bp paired-end library was sequenced on an Illumina Hi-Seq 2500 instrument. The other libraries were sequenced on a Illumina NovaSeq 6000 instrument. In addition, one 10× linked read library was constructed and sequenced with an Illumina NovaSeq 6000 instrument, yielding 75-fold coverage. A de novo whole genome assembly (WGA) was performed with the DeNovoMAGIC3 software (NRGene).

For the super-scaffolding, one Hi-C library was generated using the Dovetail™ Hi-C Library Preparation Kit and sequenced on one lane of an Illumina HiSeq

4000 instrument, yielding ~128-fold coverage. The assembly to Hi-C super-scaffolds was performed with the HiRise™ software v2.0. This generated 18 large super-scaffolds with a size between 25.5 and 49.1 Mb (total size = 642.8 Mb with N50 = 39.9 Mb), while the 19th longest super-scaffold was only 1.5 Mb in size, indicating that the 18 longest super-scaffolds correspond to the 18 *D. exilis* chromosomes.

To generate a BioNano optical map, ultra HMW DNA was isolated using the QIAGEN Genomic-tip 500/G kit (Cat no./ID: 10262) from plants grown in a greenhouse with a dark treatment applied during 48 h prior to collecting leaf tissue. Labeling and staining of the HMW DNA were performed according to the Bionano Prep Direct Label and Stain (DLS) protocol (30206—Bionano Genomics) and then loaded on one Saphyr chip. The optical map was generated using the Bionano Genomics Saphyr System according the Saphyr System User Guide (30247—Bionano Genomics). A total of 1114.6 Gb of data were generated, but only 210.6 Gb of data corresponding to molecules with a size larger than 150 kb were retained. Hybrid scaffolding between the WGA and the optical maps was done with the hybridScaffold pipeline (with the Bionano Access v1.4 default parameters). In total, 39 hybrid scaffolds ranging from 550 kb to 40.9 Mb (total length 657.3 Mb with N50 = 22.6 Mb) were obtained (Supplementary Table 3).

To construct the final chromosome-scale assembly, we integrated the Hi-C super-scaffolds and the hybrid scaffold manually. The 39 hybrid scaffolds generated with the optical map were aligned to the 18 Hi-C super-scaffolds produced by HiRise (Supplementary Table 9). The hybrid scaffolds detected two chimeric DeNovoMAGIC3 scafflods that most likely arose from wrong connections in the putative centromeres. The final chromosome-level assembly was constructed as follows: (i) We used the hybrid assembly to break the two chimeric DeNovoMAGIC3 scaffolds, (ii) merged and oriented the hybrid scaffolds using the Hi-C super scaffolds, and (iii) merged the remaining contigs and scaffolds into an unanchored chromosome.

A second genome assembly was produced using the TRITEX pipeline[27]. Briefly, the initial step using Illumina sequencing data (pre-processing of paired-end and mate-pair reads, unitig assembly, scaffolding and gap closing) was done according the instructions of the TRITEX pipeline [https://tritexassembly.bitbucket.io/]. Integration of 10× and Hi-C reads was done differently. For the 10× scaffolding, we used tigmint v1.1.2[52] to detect and cut sequences at positions with few spanning molecules, arks v1.0.3[53] to generate graphs of scaffolds with connection evidence, and LINKS[54] for a second step of scaffolding. For the Hi-C data, we used BWA[55] to map the reads against the previous scaffolds and juicer tools v1.5[56] for the super-scaffolding.

**Preparation of FISH probes and cytogenetic analyses**. Libraries of 45 bp long oligomers specific for each fonio pseudomolecule were designed using the Chorus software v1.1[26] [https://github.com/forrestzhang/Chorus] with the following criteria: -p CGTGGTCGCGTCTCA -l 45 –homology 75 -d 10. The number of oligomers per pseudomolecule was adjusted to ensure uniform fluorescent signals along the entire chromosomes after fluorescence in situ hybridization (FISH). In total, 310,484 oligomers were selected (between 12,647 and 20,000 oligomers per library) and were synthesized by Arbor Biosciences (Ann Arbor, MI, USA). To prepare chromosome painting probes for FISH, the oligomer libraries were labeled directly using 6-FAM or CY3, or indirectly using digoxigenin or biotin. A tandem repeat CL10 with a 314 bp long repetitive unit (155 bp long subunit) was identified after the analysis of fonio DNA repeats with the RepeatExplorer software[57]. A 20 bp oligomer was designed based on the CL10 sequence using the Primer3 software[58] labeled by CY3 and used for FISH.

Chromosome spreads were prepared from actively growing roots (~1 cm long)[59]. Roots were collected into 50 mM phosphate buffer (pH 7.0) containing 0.2% β-mercaptoethanol, pre-treated in 0.05% 8-hydroxyquinoline for three hours at room temperature, fixed in 3:1 ethanol:acetic acid fixative overnight and stored in 70% ethanol at −20 °C. Chromosome spreads prepared from ~20 roots were washed in distilled water, 1× KCl buffer (pH 4) and root tips were incubated in enzyme mixture containing 4% cellulase and 2% macerozyme for 56 min at 37 °C. The reaction was stopped with TE buffer (pH 7.6), root tips were washed in 100% ethanol, then 30 µl of 9:1 ice-cold acetic acid:methanol mixture was added and root tips were broken with tweezers and dropped onto slide placed in a humid box. The preparations were air-dried, postfixed in 4% (v/v) formaldehyde solution in 2× SSC solution and used for fluorescence in situ hybridization (FISH). FISH was performed with the chromosome painting probes and a probe for the CL10 tandem repeat using a hybridization mixture (30 µl) containing 50% formamide, 10% dextran sulfate in 2× SSC and 10 ng/µl of labeled probe. Hybridization was carried out overnight at 37 °C. Digoxigenin-labeled and biotin-labeled probes were detected using anti-digoxigenin-FITC (Roche Applied Science) and streptavidin-Cy3 (ThermoFisher Scientific/Invitrogen), respectively. The preparations were mounted in Vectashield with DAPI (Vector laboratories, Ltd., Peterborough, UK) to counterstain the chromosomes and the slides were examined with an Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) equipped with a Cool Cube 1 camera (Metasystems, Altlussheim, Germany) and appropriate optical filters. The capture of fluorescence signals and measure of chromosome lengths were done using the ISIS software 5.4.7 (Metasystems) and final image adjustment was done in Adobe Photoshop CS5.

**Analysis of fl-LTR-RT**. fl-LTR-RT were identified using both LTRharvest[60] (-minlenltr 100 -maxlenltr 40000 -mintsd 4 -maxtsd 6 -motif TGCA -motifmis 1 -similar 85 -vic 10 -seed 20 -seqids yes) and LTR_finder[61] (-D 40000 -d 100 -L 9000 -l 50 -p 20 -C -M 0.9). Then, the candidate LTRs were filtered using LTR_retriever[62].

fl-LTRs were classified into families using a clustering approach using MeShClust2[63] with default parameters and manually curated with dotter[64] to discard wrongly assigned LTRs. Only LTR-RT families with at least 30 intact copies were considered in genetics analyses, where different families represented 'populations' and each fl-LTR-RT copy of a family was considered as an individual. A multiple sequence alignment was performed within each family using Clustal Omega[65] and the polymorphic sites were identified and converted into variant call format (VCF) using msa2vcf.jar [https://github.com/lindenb/jvarkit]. A principal component analysis (PCA) was performed for each fl-LTR-RT family using the R packages vcfR v.1.8.0[66] and adegenet v2.1.1[67]. Results were visualized with ggplot2 v3.3.2[68].

The approximate insertion dates of the fl-LTR-RTs were calculated using the evolutionary distance between two LTR sequences with the formula $T = K/2\mu$, where $K$ is the divergence rate approximated by percent identity and $\mu$ is the neutral mutation rate. The mutation rate used was $\mu = 1.3 \times 10^{-8}$ mutations per bp per year[69].

**Gene annotation**. A combination of homology-based and de novo approaches was used for repeat annotation using the RepeatModeler software [http://www.repeatmasker.org/RepeatModeler/] and the RepeatExplorer software[57]. The results of RepeatModeler, RepeatExplorer and LTR_retriever were merged into a comprehensive de novo repeat library using USEARCH v.11[70] with default parameters (-id 0.9). The final repeats were classified using the RepeatClassifier module with the NCBI engine and then used to mask the CM05836 assembly.

The protein coding genes were predicted on the masked genome using MAKER (v3.01.02)[71]. Transcriptome reads from flag leaf, grain, panicle and above-ground seedlings were filtered for ribosomal RNA using SortMeRNA v2.1[72] and for adaptor sequences, quality and length using trimmomatic v0.38[73]. Filtered reads were mapped to the reference sequence using STAR v.2.7.0d[74] and the alignments were assembled with StringTie v.1.3.5[75] to further be used as transcript evidence. Protein sequences from *Arabidopsis thaliana*[76], *S. italica*[28], *S. bicolor*[77], *Z. mays*[78] and *O. sativa*[79] were compared with BLASTX[80] to the masked pseudomolecules of CM05836 and alignments were filtered with Exonerate v.2.2.0[81] to search for the accurately spliced alignments. A de novo prediction was performed using ab initio softwares GeneMark-ES v.3.54[82], SNAP v.2006-07-28[83], and Augustus v.2.5.5[84]. Three successive iterations of the pipeline MAKER were performed in order to improve the inference of the gene models and to integrate the final consensus genes. Finally, the putative gene functions were assigned using the UniProt/SwissProt database (Release 2019_08 – [https://www.uniprot.org/]).

**Synteny and comparative genome analysis**. To identify syntenic relationships, BLASTP (*E*-value $1 \times 10^{-5}$) was first used to perform all-versus-all protein comparisons between CM05836 and *S. italica*[28], *Panicum miliaceum*[35], *Panicum hallii*[85], *S. bicolor*[77], *Z. mays*[78], *O. sativa*[79], *B. distachyon*[86], *H. vulgare*[87], *T. aestivum*[31] and *A. tauschii*[88]. Then, MCScanX[89] (-m 25; -s 10) was used for pairwise synteny block detection. For all the comparisons, fonio sub-genomes A and B were analyzed independently and only 1:1 relationships were retained (1:2 relationships for tetraploid broomcorn millet and the ancient tetraploid maize, where individual sub-genomes are not clearly phased).

To estimate the divergence time, we calculated the synonymous substitution rates (*Ks*) for each homoeologous and orthologous pair. Briefly, nucleotide sequences were aligned with clustalW[90] and the *Ks* were calculated using the CODEML program in PAML[91]. Then, time of divergence events was calculated using $T = Ks/2\lambda$, based on the clock-($\lambda$) estimated for grasses of $6.5 \times 10^{-9}$ [92]. We used the maximum likelihood for gene-order analysis (MLGO) tool[93] to reconstruct the hypothetical ancestral genome of the Paniceae tribe from syntenic-block data of *D. exilis*, *S. italica*, *P. miliaceum*, and *P. hallii*. Syntenic blocks were identified using MCScanX[89] using the B sub-genome of *D. exilis* as reference. Then, individual syntenic blocks across all selected species were analyzed according their physical position and orientation in order to combine them into syntenic-block markers[94] (Supplementary Data 2). Syntenic-block markers along with the known species tree were provided as inputs to the MLGO server to infer the ancestral state of the Paniceae.

To analyze genome dominance, the quantification of transcript abundance and expression were calculated using RSEM v1.3.1[95] and genes were considered as expressed when expression was >0.5 TPM in at least one tissue. To study the homoeolog expression bias a two-sided Mann–Whitney *U* test was used to examine if the TPM values of sub-genomes A and B differed significantly.

**Re-sequencing of *D. exilis* and *D. longiflora* accessions**. One or two young leaves from *D. exilis* or *D. longiflora* seedlings were collected in 2 ml Eppendorf tubes, flash-frozen in liquid nitrogen, and ground with glass beads in a SPEX SamplePrep Geno/Grinder 2010. Samples were mixed with 1.2 ml 2× CTAB buffer (2% CTAB, 200 mM Tris/HCl (pH 8), 20 mM EDTA, 1.4 M NaCl, 1.0% PVP, 0.28 M β-

mercaptoethanol) and incubated at 65 °C with periodic mixing for 60 min. Samples were cooled to room temperature and centrifuged at 2000 × *g* for 10 min. Subsequently, 900 µl of the supernatant was transferred to a fresh 2 ml Eppendorf tube and incubated with 800 µl dichloromethane:isoamyl alcohol (24:1) in an overhead-shaker at half-speed at 4 °C for 15 min. After centrifugation at 10,000 × *g* for 15 min, 800 µl of the supernatant was transferred to a fresh 2 ml Eppendorf tube and incubated with 5 µl RNase A (10 mg/ml, EN0531, ThermoFisher Scientific) at 37 °C for 15 min. DNA was precipitated by adding 560 µl isopropanol and mixing the tubes by inversion. Precipitated DNA was pelleted by centrifugation at 4 °C and 10,000 × *g* for 10 min and the supernatant discarded. The DNA pellet was washed in a first ethanol wash (76% ethanol, 200 mM sodium acetate) for 15 min and in a second ethanol wash (76% ethanol, 10 mM ammonium acetate) for 5 min. Subsequently, DNA pellets were air-dried to remove residual ethanol and eluted in 50 µl TE buffer (10 mM Tris/HCl (pH 8), 1 mM EDTA). Extracted DNA was quantified with the Qubit dsDNA HS Assay (Q32851, ThermoFisher Scientific), the purity was confirmed by checking 260/280 and 260/230 ratios on a Nanodrop spectrophotometer, and the integrity was confirmed by analyzing 1 µl per sample on a 1% TAE agarose gel. Library preparation and sequencing were performed by Novogene. Briefly, 1.0 µg DNA per sample was used as input material for the DNA sample preparations. Sequencing libraries were generated using the NEBNext® Ultra II DNA Library Prep Kit following manufacturer's instructions. Libraries were sequenced using an Illumina NovaSeq 6000 system.

**Mapping of re-sequencing data and variant calling**. For quality control of each sample, raw sequence reads were analyzed with the fastqQC tool-v0.11.7 and low-quality reads were filtered with trimmomatic-v0.38[73] using the following criteria: LEADING:20; TRAILING:20; SLIDINGWINDOW:5:20 and MINLEN:50. The filtered paired-end reads were then aligned for each sample individually against the CM05836 reference assembly using BWA-MEM (v0.7.17-r1188)[55] followed by sorting and indexing using samtools (v1.6)[96]. Alignment summary, base quality score and insert size metrics were collected and duplicated reads were marked and read groups were assigned using the Picard tools [http://broadinstitute.github.io/picard/]. Variants were identified with GATK v3.8[97] using the emitRefConfidence function of the HaplotypeCaller algorithm to call SNPs and InDels for each accession followed by a joint genotyping step performed by GenotypeGVCFs. To obtain high confidence variants, we excluded SNPs and InDels with the VariantFiltration function of GATK with the criteria: QD < 2.0; FS > 60.0; MQ < 40.0; MQRankSum < −12.5; ReadPosRankSum < −8.0 and SOR > 4.0. The complete automated pipeline has been compiled and is available on github [https://github.com/IBEXCluster/IBEX-SNPcaller].

A total of 36.5 million variants were called. Raw variants were filtered using GATK v3.8[97,98] and VCFtools v0.1.17[99]. Variants located on chromosome unanchored, InDels, 'SNP clusters' defined as three or more SNPs located within 10 bp, missing data >10%, low and high average SNP depth ($14 \leq DP \geq 42$), and accessions having more than 33% of missing data were discarded. Only biallelic SNPs were retained to perform further analyses representing a final VCF file of 11,046,501 SNPs (Supplementary Table 7). These variants were annotated using snpEff v4.3[100] with the CM05836 gene models.

**Genetic diversity and population structure**. Genetic diversity and population structure analyses were performed using *D. exilis* and *D. longiflora* accessions together, or with *D. exilis* accessions alone. PCA and individual ancestry coefficients estimation were performed using the R package LEA v2.0[101]. The geographical projection of the first PCA axis of the *D. exilis* samples was done and visualized using the Kriging function in the fields v10.3 R package [https://cran.r-project.org/web/packages/fields][102]. For ancestry coefficients analyses, the snmf function was used with 10 independent runs for each $K$ from $K = 2$ to $K = 10$. The optimal $K$, indicating the most likely number of ancestral populations, was determined with the cross-validation error rate. For the two analyses, we considered only SNPs present in at least two accessions (i.e., private SNPs were excluded). We also studied the geographic distribution of private SNPs[103] (i.e., SNP present only once in a single accession). Genome-wide pairwise LD was estimated independently for *D. exilis* (2,617,322 SNPs) and *D. longiflora* (9,839,152 SNPs). LD decay ($r^2$) was calculated using the tool PopLDdecay v3.40[104] in a 500 kb distance and plotted using the ggplot2 v3.3.2 package in R[68].

**Association of climate, geography, and social factors with the population genetics structure**. Natural (i.e., climate, geography) and human (i.e., ethnic and linguistic) factors were tested for an association with the genetic structure of *D. exilis* accessions. Bioclimate data and ethnic and linguistic data were extracted for each of the 166 re-sequenced *D. exilis* accessions from WorldClim v1.4 [https://www.worldclim.org/data/v1.4/worldclim14.html][105], passport data (ethnic groups), and Ethnologue v16 (language, [https://www.ethnologue.com/]). The associations between the genetic structure and different factors were evaluated statistically using Pearson's correlation coefficient test (two-sided) for the bioclimatic and geographic data with the R package stats v3.6.0 (function cor). For the associations of social factors, two-sided Mantel tests (ecodist R package v2.0.5 [https://rdrr.io/cran/ecodist/]) with 999 permutations were performed between the genetic distance matrix of *D. exilis* and the dissimilarity matrix of social factors as inputs. Fisher's exact tests (stats R package v3.6.0) were performed with the structure data at $K = 6$ and accessions having ancestry thresholds of >70% as input. The $p$-values were computed by Monte Carlo simulation. We used analysis of covariance (ANCOVA) to test for the impact of ethnolinguistic factors on the first PCA coordinates while controlling for other covariates: climatic and geographic variables.

Synthetic climate variables were retrieved from the WorldClim v1.4 database[105]. We focused our analysis on two climate variables that were the most pertinent to annual species, i.e., mean temperature of the wettest quarter (bio8) and mean precipitation of the wettest quarter (bio 16), respectively. A GWAS with these climate variables has been performed[106]. Briefly, we first used the impute method from the R package LEA v2.0[101] to fill the gaps in the genotypic matrix in order to achieve greater power. We then filtered out SNPs with a MAF below 5%. We used three different models of association to perform the GWAS, efficient mixed-model association (EMMA)[107] using the R package emma v1.1.2, mixed linear model (MLM) as implemented in the R package gapit v3.0[108] and latent factor mixed models (LFMM) using LFMM2[109] as implemented in lfmm R package v2.0. For gapit (MLM), we considered six principal components, which we chose from the previously described structure analysis. The LFMM v2 method implements a Latent Factor Mixed Model that estimates unknown confounding factors (the latent factors) jointly using the genotypic and phenotypic matrices. For this specific model, the estimation of latent factors was performed using both the phenotypic data (i.e., the climate variable in consideration) and a genotypic matrix of SNPs with a MAF higher than 20%. Once the GWAS step was performed, we used a FDR[110] approach to estimate the significance of the associations for each analysis, retaining a FDR threshold of 5%, resulting in a list of candidate positions for each variable and each GWAS method. We then considered a genomic window of 50 kb upstream and 50 kb downstream for each position as a candidate region, which we denominated quantitative trait locus (QTL). When two QTLs overlapped, they were combined.

To test SNP associations to ethnolinguistic groups, we used the PLINK software v1.90[111]. We filtered out the SNPs with MAF below 5%, more than 10% missing data, and in LD, resulting in 897,796 SNPs. The association of SNPs with each ethnic and linguistic group was tested by a logistic regression model assuming additive genetic effects, using three principal components as covariates to control for population stratification. Significant associations were identified above the threshold of $p$ value $= 1e^{-05}$. Manhattan plots were plotted to display the associations using the qqman R package v0.1.4[112]. GO terms for all fonio protein sequences were assigned using the DeepGOPlus model [https://github.com/bio-ontology-research-group/deepgoplus][113]. We only considered GO labels that have a confident score >0.3. Enrichment analysis for GO biological processes, cellular components, and molecular functions of the genes of interests from climate association analyses was performed using the Fisher exact tests implemented in R package topGO v3.36.0[114] (algorithm = 'weight01'). We reported GO terms with enrichment $p$ value < 0.05 for each GO-term category.

**Demographic history reconstruction**. A Sequentially Markovian Coalescent-based approach[115,116] was used to infer past evolution of the effective population size of *D. exilis* and *D. longiflora*. This analysis was made with the smc++[117] software and additional customized scripts [https://github.com/Africrop/fonio_smcpp]. We excluded SNPs that were not usable (e.g. repeated regions) according to msmc-tools advices [https://github.com/stschiff/msmc-tools]. Analysis was done with all nine individuals from the genetically closest group of *D. longiflora* and also with the six groups of *D. exilis* based on the genetic structure. The smc++ vcf2smc tools was used to generate the input file and smc++ inference was performed for the different sets of genotypes using the smc++ cv command, considering 200 and 25,000 generations lower and upper time points boundaries, respectively. All other options were set to default. We considered a generation time of one year and a mutation rate of $6.5 \times 10^{-9}$. Different sets of distinguished lineages were considered for the different sets of individuals (all the samples were considered as distinguished). Graphical representation was made using the ggplot2 v3.3.2[68] package for R.

**Genome scanning for selection signals**. Detection of selection was performed independently with three different methods: (1) Weir and Cockerham $F_{ST}$; (2) nucleotide diversity ratios ($\pi$) were calculated between *D. exilis* and *D. longiflora* using VCFtools v0.1.17[99] with sliding windows of 50 kb every 10 kb; (3) SweeD v3.3.1[39] was used on *D. exilis* to detect signatures of selective sweeps based on the composite likelihood ratio (CLR) test within non-overlapping intervals of 3000 bp along each chromosome. We grouped CLR peak positions into one common window if two or several peaks were <100 kb apart. Candidate genomic regions were selected for each method based on top 1% values. A list comprising 34 well-characterized domestication genes (Supplementary Data 8) from major crops was selected and compared by BLAST v2.6.0 against gene models and the genome assembly of *D. exilis* accession CM05836. Each putative orthologous gene was validated by local alignment using ClustalW v2.1[90]. We considered the genes as orthologous when ≥ 70% of the coding sequence (CDS) length hit the genomic sequence of *D. exilis* with at least 70% identity. Validated orthologous genes were crossed with genomic regions under selection using bedtools v2.28.0[118]. To assess the effect of Δ*DeSh1* on seeds shattering, fonio panicles with three racemes were collected from mature plants and placed in 15 ml tubes. Individual tubes were

shaken using a Geno Grinder tissue homogenizer (SPEX SamplePrep, Metuchen, NJ) at 1350 rpm speed for 20 s and the percentage of shattering was calculated. ANOVA test was performed with the MVApp [https://mvapp.kaust.edu.sa/][119]. Two specific primer pairs were designed to confirm the 60 kb deletion. The first primer pair amplifies-1 kb fragment within the *DeSh1-9A* gene (forward primer 5′-GTGACCATGCCAAGCAGGCG-3′; reverse primer 5′-GCTAGCTTGAGGTA TTTACGG-3′). The second primer pair was designed in the flanking region of the deletion and was used to amplify ~750 bp fragment in accessions that carry the deletion (forward primer 5′-GCCGTATATAGTTGCCGCATCA-3′; reverse primer 5′-CCTACCGTTAGATCCGTGCGGA-3′).

**Reporting summary**. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Data supporting the findings of this work are available within the paper and its Supplementary Information files. A reporting summary for this Article is available as a Supplementary Information file. The datasets generated and analyzed during the current study are available from the corresponding author upon request. The raw sequencing data used for de novo whole-genome assembly, the raw bionano map, the CM05836 genome assembly, the RNA-seq data for the annotation and the 183 re-sequenced accessions of *D. exilis* and *D. longiflora* for the population genomics analysis are available on EBI-ENA under the study number PRJEB36539 [https://www.ebi.ac.uk/ena/browser/view/PRJEB36539]. The annotation of the CM05836 genome, the Tritex assembly, the probes for the chromosome painting experiment, the gene ontology annotation, the VCF file and *DeSh1* phenotyping data are available on the DRYAD database [https://doi.org/10.5061/dryad.2v6wwpzj0]. The plant coding sequences [https://plants.ensembl.org] and [https://genomevolution.org/coge/], the UniProt/SwissProt database (Release 2019_08–[https://www.uniprot.org/]), the embryophyta_odb9 BUSCO dataset [https://busco-archive.ezlab.org/v2/datasets/embryophyta_odb9.tar.gz], the Ethnologue version 16 language [https://www.ethnologue.com/]), and WorldClim version 1.4 [https://www.worldclim.org/data/v1.4/worldclim14.html] databases were downloaded from source for data analyses. Source data are provided with this paper.

## Code availability

Custom scripts for the SNP calling are available on GitHub: [https://github.com/IBEXCluster/IBEX-SNPcaller]. Custom scripts used to launch the smc++ analysis for fonio genome resequencing data are available on GitHub: [https://github.com/Africrop/fonio_smcpp].

Received: 13 April 2020; Accepted: 16 August 2020;
Published online: 08 September 2020

## References

1. Hickey, L. T. et al. Breeding crops to feed 10 billion. *Nat. Biotechnol.* **37**, 744–754 (2019).
2. Tena, G. Sequencing forgotten crops. *Nat. Plants* **5**, 5 (2019).
3. National Academies of Sciences Engineering and Medicine. *Breakthroughs to Advance Food and Agricultural Research by 2030* (The National Academies Press, Washington, 2019).
4. FAO. *The State of Agricultural Commodity Markets 2018. Agricultural Trade, Climate Change and Food Security* (FAO, Rome, 2018).
5. Dalin, C., Wada, Y., Kastner, T. & Puma, M. J. Groundwater depletion embedded in international food trade. *Nature* **543**, 700–704 (2017).
6. Fernie, A. R. & Yan, J. De novo domestication: an alternative route toward new crops for the future. *Mol. Plant* **12**, 615–631 (2019).
7. Tanksley, S. D. & McCouch, S. R. Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* **277**, 1063–1066 (1997).
8. Gruber, K. Agrobiodiversity: the living library. *Nature* **544**, S8–S10 (2017).
9. Kistler, L. et al. Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. *Science* **362**, 1309–1313 (2018).
10. Wing, R. A., Purugganan, M. D. & Zhang, Q. F. The rice genome revolution: from an ancient grain to Green Super Rice. *Nat. Rev. Genet.* **19**, 505–517 (2018).
11. Eshed, Y. & Lippman, Z. B. Revolutions in agriculture chart a course for targeted breeding of old and new crops. *Science* **366**, eaax0025 (2019).
12. Kantar, M. B. & Runck, B. Take a walk on the wild side. *Nat. Clim. Change* **9**, 731–732 (2019).
13. Dawson, I. K. et al. The role of genetics in mainstreaming the production of new and orphan crops to diversify food systems and support human nutrition. *New Phytol.* **224**, 37–54 (2019).
14. Pironon, S. et al. Potential adaptive strategies for 29 sub-Saharan crops under future climate change. *Nat. Clim. Change* **9**, 758–763 (2019).
15. Wallace, J. G., Rodgers-Melnick, E. & Buckler, E. S. On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. *Annu. Rev. Genet.* **52**, 421–444 (2018).
16. Chen, K., Wang, Y., Zhang, R., Zhang, H. & Gao, C. CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* **70**, 667–697 (2019).
17. Barnaud, A. et al. High selfing rate inferred for white fonio [*Digitaria exilis* (Kippist.) Stapf] reproductive system opens up opportunities for breeding programs. *Genet. Resour. Crop Evol.* **64**, 1485–1490 (2017).
18. Ayenan, M. A. T., Sodedji, K. A. F., Nwankwo, C. I., Olodo, K. F. & Alladassi, M. E. B. Harnessing genetic resources and progress in plant genomics for fonio (*Digitaria* spp.) improvement. *Genet. Resour. Crop Evol.* **65**, 373–386 (2018).
19. Cruz, J. F. & Beavogui, F. *Fonio, an African Cereal* (CIRAD, France, 2016).
20. Adoukonou-Sagbadja, H., Wagner, C., Ordon, F. & Friedt, W. Reproductive system and molecular phylogenetic relationships of fonio millets (*Digitaria* spp., Poaceae) with some polyploid wild relatives. *Trop. Plant Biol.* **3**, 240–251 (2010).
21. Abdul, S. D. & Jideani, A. I. O. Fonio (*Digitaria* spp.) breeding. In *Advances in Plant Breeding Strategies: Cereals* (eds Al-Khayri, J. M., Jain, S. M. & Johnson, D. V.) 47–81 (Springer, 2019).
22. Adoukonou-Sagbadja, H. et al. Flow cytometric analysis reveals different nuclear DNA contents in cultivated Fonio (*Digitaria* spp.) and some wild relatives from West-Africa. *Plant Syst. Evol.* **267**, 163–176 (2007).
23. Avni, R. et al. Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* **357**, 93–97 (2017).
24. Edger, P. P. et al. Origin and evolution of the octoploid strawberry genome. *Nat. Genet.* **51**, 541–547 (2019).
25. Springer, N. M. et al. The maize W22 genome provides a foundation for functional genomics and transposon biology. *Nat. Genet.* **50**, 1282–1288 (2018).
26. Han, Y. H., Zhang, T., Thammapichai, P., Weng, Y. Q. & Jiang, J. M. Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. *Genetics* **200**, 771–779 (2015).
27. Monat, C. et al. TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol.* **20**, 284 (2019).
28. Bennetzen, J. L. et al. Reference genome sequence of the model plant Setaria. *Nat. Biotechnol.* **30**, 555–561 (2012).
29. Tang, H. Disentangling a polyploid genome. *Nat. Plants* **3**, 688–689 (2017).
30. Suguiyama, V. F., Vasconcelos, L. A. B., Rossi, M. M., Biondo, C. & de Setta, N. The population genetic structure approach adds new insights into the evolution of plant LTR retrotransposon lineages. *PLoS ONE* **14**, e0214542 (2019).
31. International Wheat Genome Sequencing Consortium. Shifting the limits in wheat research and breeding through a fully annotated and anchored reference genome sequence. *Science* **361**, eaar7191 (2018).
32. Ramirez-Gonzalez, R. H. et al. The transcriptional landscape of polyploid wheat. *Science* **361**, eaar6089 (2018).
33. Bird, K. A., VanBuren, R., Puzey, J. R. & Edger, P. P. The causes and consequences of subgenome dominance in hybrids and recent polyploids. *New Phytol.* **220**, 87–93 (2018).
34. Schnable, J. C., Springer, N. M. & Freeling, M. Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proc. Natl Acad. Sci. USA* **108**, 4069–4074 (2011).
35. Shi, J. P. et al. Chromosome conformation capture resolved near complete genome assembly of broomcorn millet. *Nat. Commun.* **10**, 464 (2019).
36. Clément, J. & Leblanc, J. M. Collecte IBPGR-ORSTOM de 1977 au Togo (Catalogue ORSTOM, 1984).
37. Ramu, P. et al. Cassava haplotype map highlights fixation of deleterious mutations during clonal propagation. *Nat. Genet.* **49**, 959–963 (2017).
38. Patwari, P. et al. Surface wax esters contribute to drought tolerance in Arabidopsis. *Plant J.* **98**, 727–744 (2019).
39. Pavlidis, P., Zivkovic, D., Stamatakis, A. & Alachiotis, N. SweeD: Likelihood-based detection of selective sweeps in thousands of genomes. *Mol. Biol. Evol.* **30**, 2224–2234 (2013).
40. Li, Y. et al. Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nat. Genet.* **43**, 1266–1269 (2011).
41. Lin, Z. W. et al. Parallel domestication of the Shattering1 genes in cereals. *Nat. Genet.* **44**, 720–724 (2012).
42. Wang, M. H. et al. The genome sequence of African rice (*Oryza glaberrima*) and evidence for independent domestication. *Nat. Genet.* **46**, 982–988 (2014).
43. VanBuren, R. et al. Exceptional subgenome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nat. Commun.* **11**, 884 (2020).
44. Cubry, P. et al. The rise and fall of African rice cultivation revealed by analysis of 246 new genomes. *Curr. Biol.* **28**, 2274–2282 (2018).
45. Liang, Z. et al. Whole-genome resequencing of 472 Vitis accessions for grapevine diversity and demographic history analyses. *Nat. Commun.* **10**, 1190 (2019).

46. Blench, R. M. Vernacular names for African millets and other minor cereals and their significance for agricultural history. *Archaeol. Anthropol. Sci.* **8**, 1–8 (2016).
47. Adoukonou-Sagbadja, H., Dansi, A., Vodouhe, R. & Akpagana, K. Collecting fonio (*Digitaria exilis* Kipp. Stapf, *D. iburua* Stapf) landraces in Togo. *Plant Genet. Resour. Newsl.* **139**, 63–67 (2004).
48. Meyer, R. S. & Purugganan, M. D. Evolution of crop species: genetics of domestication and diversification. *Nat. Rev. Genet.* **14**, 840–852 (2013).
49. Barnaud, A. et al. Development of nuclear microsatellite markers for the fonio, *Digitaria exilis* (Poaceae), an understudied West African cereal. *Am. J. Bot.* **99**, E105–E107 (2012).
50. Doležel, J., Greilhuber, J. & Suda, J. Estimation of nuclear DNA content in plants using flow cytometry. *Nat. Protoc.* **2**, 2233–2244 (2007).
51. Doležel, J., Bartoš, J., Voglmayr, H. & Greilhuber, J. Nuclear DNA content and genome size of trout and human. *Cytometry* **51A**, 127–128 (2003).
52. Jackman, S. D. et al. Tigmint: correcting assembly errors using linked reads from large molecules. *BMC Bioinforma.* **19**, 393 (2018).
53. Coombe, L. et al. ARKS: chromosome-scale scaffolding of human genome drafts with linked read kmers. *BMC Bioinforma.* **19**, 234 (2018).
54. Warren, R. L. et al. LINKS: scalable, alignment-free scaffolding of draft genomes with long reads. *GigaScience* **4**, 35 (2015).
55. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
56. Dudchenko, O. et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
57. Novak, P., Neumann, P., Pech, J., Steinhaisl, J. & Macas, J. RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics* **29**, 792–793 (2013).
58. Untergasser, A. et al. Primer3–new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
59. Šimoníková, D. et al. Chromosome painting facilitates anchoring reference genome sequence to chromosomes in situ and integrated karyotyping in banana (*Musa* Spp.). *Front. Plant Sci.* **10**, 1503 (2019).
60. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinforma.* **9**, 18 (2008).
61. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).
62. Ou, S. J. & Jiang, N. LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* **176**, 1410–1422 (2018).
63. James, B. T., Luczak, B. B. & Girgis, H. Z. MeShClust: an intelligent tool for clustering DNA sequences. *Nucleic Acids Res.* **46**, e83 (2018).
64. Sonnhammer, E. L. & Durbin, R. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene* **167**, GC1–GC10 (1995).
65. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
66. Knaus, B. J. & Grunwald, N. J. VCFR: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* **17**, 44–53 (2017).
67. Jombart, T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405 (2008).
68. Wickham, H. *ggplot2 - Elegant Graphics for Data Analysis* (Springer International Publishing, 2016).
69. Ma, J. X. & Bennetzen, J. L. Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl Acad. Sci. USA* **101**, 12404–12410 (2004).
70. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).
71. Cantarel, B. L. et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
72. Kopylova, E., Noe, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
73. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
74. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
75. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
76. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796–815 (2000).
77. Paterson, A. H. et al. The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457**, 551–556 (2009).
78. Schnable, P. S. et al. The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**, 1112–1115 (2009).
79. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature* **436**, 793–800 (2005).
80. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
81. Slater, G. S. & Birney, E. Automated generation of heuristics for biological sequence comparison. *BMC Bioinforma.* **6**, 31 (2005).
82. Borodovsky, M. & Lomsadze, A. Eukaryotic gene prediction using GeneMark. hmm-E and GeneMark-ES. *Curr. Protoc. Bioinforma.* **35**, 4.6.1–4.6.10 (2011).
83. Korf, I. Gene finding in novel genomes. *BMC Bioinforma.* **5**, 59 (2004).
84. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215–ii225 (2003).
85. Lovell, J. T. et al. The genomic landscape of molecular responses to natural drought stress in *Panicum hallii*. *Nat. Commun.* **9**, 5213 (2018).
86. The International Brachypodium Initiative. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**, 763–768 (2010).
87. Mascher, M. et al. A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**, 427–433 (2017).
88. Luo, M. C. et al. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature* **551**, 498–502 (2017).
89. Wang, Y. P. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
90. Thompson, J. D., Higgins, D. G. & Gibson, T. J. Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
91. Yang, Z. H. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
92. Gaut, B. S., Morton, B. R., McCaig, B. C. & Clegg, M. T. Substitution rate comparisons between grasses and palms: synonymous rate differences at the nuclear gene Adh parallel rate differences at the plastid gene rbcL. *Proc. Natl Acad. Sci. USA* **93**, 10274–10279 (1996).
93. Hu, F., Lin, Y. & Tang, J. MLGO: phylogeny reconstruction and ancestral inference from gene-order data. *BMC Bioinforma.* **15**, 354 (2014).
94. Ren, L., Huang, W. & Cannon, S. B. Reconstruction of ancestral genome reveals chromosome evolution history for selected legume species. *New Phytol.* **223**, 2090–2103 (2019).
95. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinforma.* **12**, 323 (2011).
96. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
97. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
98. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinforma.* **43**, 11 10 1-11 10 33 (2013).
99. Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
100. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w(1118); iso-2; iso-3. *Fly* **6**, 80–92 (2012).
101. Frichot, E. & Francois, O. LEA: an R package for landscape and ecological association studies. *Methods Ecol. Evol.* **6**, 925–929 (2015).
102. Nychka, D., Furrer, R., Paige, J. & Sain, S. *fields: Tools for Spatial Data*. Retrieved from https://cran.r-project.org/package=fields (2017).
103. Cubry, P., Vigouroux, Y. & Francois, O. The empirical distribution of singletons for geographic samples of DNA sequences. *Front. Genet.* **8**, 139 (2017).
104. Zhang, C., Dong, S. S., Xu, J. Y., He, W. M. & Yang, T. L. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics* **35**, 1786–1788 (2019).
105. Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978 (2005).
106. Cubry, P. et al. Genome wide association study pinpoints key agronomic QTLs in African rice *Oryza glaberrima*. Preprint at https://doi.org/10.1101/2020.01.07.897298 (2020).
107. Kang, H. M. et al. Efficient control of population structure in model organism association mapping. *Genetics* **178**, 1709–1723 (2008).
108. Lipka, A. E. et al. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397–2399 (2012).
109. Caye, K., Jumentier, B., Lepeule, J. & Francois, O. LFMM 2: fast and accurate inference of gene-environment associations in genome-wide studies. *Mol. Biol. Evol.* **36**, 852–860 (2019).
110. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. (Methdol.)* **57**, 289–300 (1995).
111. Purcell, S. et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

112. Turner, S. D. qqman: an R package for visualizing GWAS results usingQ-Q and Manhattan plots. *J. Open Source Softw.* **3**, 731 (2018).
113. Kulmanov, M. & Hoehndorf, R. DeepGOPlus: improved protein function prediction from sequence. *Bioinformatics* **36**, 422–429 (2020).
114. Alexa, A., Rahnenfuhrer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006).
115. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).
116. Schiffels, S. & Durbin, R. Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014).
117. Terhorst, J., Kamm, J. A. & Song, Y. S. Robust and scalable inference of population history froth hundreds of unphased whole genomes. *Nat. Genet.* **49**, 303–309 (2017).
118. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
119. Julkowska, M. M. et al. MVApp-Multivariate analysis application for streamlined data analysis and curation. *Plant Physiol.* **180**, 1261–1276 (2019).

## Acknowledgements

We are grateful to Vinicius M. Lube, Samantha Bazan, Ablaye Ngom, Marie Piquet, Hélène Adam, and Carole Gauron for their technical assistance in Herbarium sampling and photography. Cirad herbarium samples of *Digitaria longiflora* were provided by Samantha Bazan (ALF [http://publish.plantnet-project.org/]). We thank Soukeye Conde for valuable insight, funded by the project Cultivar, reference ANR-16-IDEX-0006, through the Investissements d'avenir program (Labex Agro: ANR-10-LABX-0001-01). This work was supported by the CIRAD-UMR AGAP and IRD UMR DIADE HPC Data Center of the South Green Bioinformatics platform [http://www.southgreen.fr/] and the Center for Desert Agriculture of the King Abdullah University of Science and Technology (KAUST). D.S., J.C., E.H. and J.D. were in part supported by the European Regional Development Fund OPVVV project 'Plants as a tool for sustainable development' number CZ.02.1.01/0.0/0.0/16_019/0000827.

## Author contributions

M.A., A.B., Y.V., and S.G.K. designed research. N.K. and M.A. established the SNP calling pipeline. J.B., M.C., and L.Z. performed molecular analyses and constructed sequencing libraries. M.A., H.I.A., P.C., L.G., N.S., and T.W. performed bioinformatics analyses. H.I.A, P.C. and Y.P. performed association studies. M.A., H.I.A., P.C., Y.V., C.B., A.B., and S.G.K. interpret results. D.S., J.C., E.H., and J.D. performed flow-cytometry and chromosome painting. S.A., S.C., and H.B. constructed and analyzed the CM05836 optical map. Y.P. and H.I.A carried out phenotypic data collection and analyses. J.J.W., M.G., N.A.K., C.L., S.C., S.V., and C.B. contributed biological materials. M.A., H.I.A., and S.G.K. wrote the paper with substantial inputs from P.C., Y.V., and A.B. All authors have read and approved the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at https://doi.org/10.1038/s41467-020-18329-4.

**Correspondence** and requests for materials should be addressed to A.B. or S.G.K.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

# APPENDIX IV

**The formation of sex chromosomes in *Silene latifolia* and *S. dioica* was accompanied by multiple chromosomal rearrangements**

Bačovský, V., Čegan, R., Šimoníková, D., Hřibová, E., Hobza, R.

*Frontiers in Plant Science* **11**: 205, 2020

doi: 10.3389/fpls.2020.00205

IF: 4.402

frontiers in Plant Science

ORIGINAL RESEARCH
published: 28 February 2020
doi: 10.3389/fpls.2020.00205

# The Formation of Sex Chromosomes in *Silene latifolia* and *S. dioica* Was Accompanied by Multiple Chromosomal Rearrangements

*Václav Bačovský[1]*, Radim Čegan[1,2], Denisa Šimoníková[2], Eva Hřibová[2] and Roman Hobza[1,2]**

[1] Department of Plant Developmental Genetics, Institute of Biophysics of the Czech Academy of Sciences, Brno, Czechia, [2] Institute of Experimental Botany, Czech Academy of Sciences, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czechia

OPEN ACCESS

**Edited by:**
*Martin A. Lysak, Masaryk University, Czechia*

**Reviewed by:**
*Ales Kovarik, Academy of Sciences of the Czech Republic (ASCR), Czechia*
*Andreas Houben, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany*

***Correspondence:**
*Václav Bačovský*
*xbacovs@ibp.cz*
*Roman Hobza*
*hobza@ibp.cz*

**Specialty section:**
*This article was submitted to Plant Systematics and Evolution, a section of the journal Frontiers in Plant Science*

**Received:** *07 October 2019*
**Accepted:** *11 February 2020*
**Published:** *28 February 2020*

**Citation:**
*Bačovský V, Čegan R, Šimoníková D, Hřibová E and Hobza R (2020) The Formation of Sex Chromosomes in Silene latifolia and S. dioica Was Accompanied by Multiple Chromosomal Rearrangements. Front. Plant Sci. 11:205. doi: 10.3389/fpls.2020.00205*

The genus *Silene* includes a plethora of dioecious and gynodioecious species. Two species, *Silene latifolia* (white campion) and *Silene dioica* (red campion), are dioecious plants, having heteromorphic sex chromosomes with an XX/XY sex determination system. The X and Y chromosomes differ mainly in size, DNA content and posttranslational histone modifications. Although it is generally assumed that the sex chromosomes evolved from a single pair of autosomes, it is difficult to distinguish the ancestral pair of chromosomes in related gynodioecious and hermaphroditic plants. We designed an oligo painting probe enriched for X-linked scaffolds from currently available genomic data and used this probe on metaphase chromosomes of *S. latifolia* (2n = 24, XY), *S. dioica* (2n = 24, XY), and two gynodioecious species, *S. vulgaris* (2n = 24) and S. *maritima* (2n = 24). The X chromosome-specific oligo probe produces a signal specifically on the X and Y chromosomes in *S. latifolia* and *S. dioica,* mainly in the subtelomeric regions. Surprisingly, in *S. vulgaris* and *S. maritima*, the probe hybridized to three pairs of autosomes labeling their p-arms. This distribution suggests that sex chromosome evolution was accompanied by extensive chromosomal rearrangements in studied dioecious plants.

**Keywords: chromosome painting, double-translocation, pseudo-autosomal region, *Silene*, Y chromosome**

## INTRODUCTION

The genus *Silene* is a model system for sex chromosome evolution, including about 700 species varying greatly in their mating system, ecology and sex determination (Bernasconi et al., 2009). Inside the genus two groups are considered as invaluable for the study of sex chromosome evolution and sex determination; section *Melandrium* and subsection *Otites* (reviewed in Vyskot and Hobza, 2015). Two dioecious plants *S. latifolia* (24, XY) and *S. dioica* (24, XY) from *Melandrium* have heteromorphic sex chromosomes and sex determination similar to mammals (Ming et al., 2007; Charlesworth, 2016). In contrast, related gynodioecious species *S. vulgaris* and *S. maritima* with the same number of autosomes (2n = 24), possess no sex chromosomes having a smaller

genome compared to *S. latifolia* or *S. dioica* (Runyeon and Prentice, 1997; Charlesworth and Laporte, 1998; Široký et al., 2001; Stone et al., 2017).

It is generally accepted that the sex chromosomes are derived from an ordinary pair of autosomes (reviewed in Bachtrog, 2006). As a result of suppressed recombination and accumulation of deleterious mutations, the sex chromosomes differ in their structure, function and gene density. The X chromosome becomes hemizygous and X hemizygosity in males leads to special regulatory mechanisms to equalize the transcription ratio between the X chromosomes and autosomes (Charlesworth and Charlesworth, 2005; Muyle et al., 2017; Darolti et al., 2019). As a result of accumulation of deleterious mutations, the Y chromosome is degenerated and the sex chromosomes may differ even within closely related species as demonstrated in human and chimpanzee (Hughes et al., 2010). Interestingly, newly formed sex chromosomes show the same signs of sex chromosome evolutionary pathways, as described in *Drosophila* (Bachtrog et al., 2009) or stickleback species (Yoshida et al., 2014) in which the ancestral Y chromosome fused with an autosome.

In *S. latifolia* and *S. dioica*, sequence data showed that the sex chromosomes evolved from one pair of autosomes with the divergence of X and Y homologous sequences <10 million years ago (Filatov, 2005), estimating the age of older and younger strata (non-recombining part of the sex chromosomes that differ from each other in level of divergence) around 11 and 6.32 mya (Krasovec et al., 2018). The sex chromosomes in *S. latifolia* and *S. dioica,* vary greatly in size having Y chromosome 1.4 larger than X (heteromorphism) (Vyskot and Hobza, 2015). The Y chromosome has a large non-recombining region and the size of the PAR (pseudo-autosomal region) is less than 10% of its chromosome length (Filatov et al., 2009). Both sex chromosomes accumulated various transposable elements (Bergero et al., 2008b; Kubat et al., 2014) and satellites (Cermak et al., 2008; Kejnovský et al., 2013), and differ in histone modifications and DNA methylation (Rodríguez Lorenzo et al., 2018; Bačovský et al., 2019).

Previous studies suggested that the sex chromosomes in *S. latifolia*, especially the Y chromosome, were derived through multiple rearrangements (Bergero et al., 2008a). Deletion mapping revealed that at least one larger inversion occurred after recombination suppression on the Y chromosome (Zluvova et al., 2005), supported also by findings of four genetically mapped genes between *S. latifolia* and *S. dioica* (Filatov, 2005). Later, Hobza et al. (2007) used physical mapping and confirmed two large inversions on the Y chromosome. These findings were further verified by Y deletion mapping showing that at least one inversion had to have occurred during the formation of the Y chromosome (Kazama et al., 2016), accelerating the recombination suppression (Bergero and Charlesworth, 2009). In contrast, comparative mapping between *S. latifolia* and *S. vulgaris* revealed the existence of one large (SvLG12) and two relatively small (SvLG9, Sv small LG) linkage groups that accompanied the sex chromosomes formation in *S. latifolia* (Bergero et al., 2013; Campos et al., 2017). Yet, it is still not clear what pair(s) of autosomes were included in such translocation and if such linkage groups also exist in other gynodioecious species. Thus, this raises two important questions: if *S. vulgaris* possesses three putative parts of three linkage groups corresponding to the X chromosome in *S. latifolia*, what is the origin of these linkage groups and how many chromosomes were involved in sex chromosome formation?

Recent advances in fluorescence *in situ* hybridization (FISH) experiments have provided a variety of techniques which can be used to study the structure, dynamics and origin of certain loci, chromosomal arms and/or specific chromosomes (reviewed in Cui et al., 2016; Bačovský et al., 2018; Huber et al., 2018). Previous cytogenetic studies in *Silene* species were based mainly on physical mapping of satellite rDNA (Široký et al., 2001), repeats and transposable elements (Cermak et al., 2008; Kejnovsky et al., 2009). Although Lengerova et al. (2004) managed to produce discrete signals using various DNA repeats (satellites, rDNA) and specific BAC clones, this approach proved to be cost ineffective due to the large screening of BACs containing only a low amount of repetitive DNA. As an another option, Hobza et al. (2004) designed a protocol using microdissected X and Y chromosomes from *S. latifolia* for whole chromosome painting. These probes produced relatively discrete signals on both sex chromosomes, but high amount of suppressive unlabeled DNA with very strict hybridization conditions made the use of such method very problematic in other *Silene* species (Hobza and Vyskot, 2007). The recent development of oligo painting probes in plants has proved to be useful in the detection of chromosomal aberrations and in comparative cytogenetics (reviewed in Jiang, 2019). In principal, oligo painting probe can be used to label particular regions containing enough short unique oligo sequences to be computationally isolated, synthesized, pooled and labeled (Han et al., 2015). Such probes, designed from conserved sequences of one species were used e.g. for developing karyotype among genetically related *Solanum* species (Braz et al., 2018), for differentiating of A, B, and D genomes in wheat (Tang et al., 2018), in comparative physical mapping of sex chromosomes in *Populus* (Xin et al., 2018) and in examination of meiotic pairing in polyploid *Solanum* species (He et al., 2018).

In this work, we designed an X chromosome-specific oligo probe enriched by X-linked scaffolds based on the *S. latifolia* female genome. We show that such technique is useful for the detection of discrete signals in sex chromosomes in *S. latifolia* and closely related *S. dioica*. In addition, the probe works well in the related gynodioecious species of *S. vulgaris* and *S. maritima*. Based on our results, we discuss the origin of sex chromosomes from one autosomal pair and the possible application of oligo painting probe in further studies. Our findings support the general hypothesis that multiple chromosomal changes took place during the formation of X and Y chromosomes.

# MATERIALS AND METHODS

## Plant Material

Seedlings of the *Silene* species listed in **Supplementary Table S1** (seeds owned by The Institute of Biophysics of the Czech Academy of Sciences) were used for chromosome preparation following (Bačovský et al., 2019). Young seedlings (average

size = 1 cm) were synchronized for 16 h in 1.125 mM hydroxyurea at RT, washed 2× for 5 min in distilled water and incubated 4 h in distilled water at RT. Cells in metaphase were accumulated by 0.05 mM colchicine at RT 4 h. After 4 h, root tips were stored for 16 h in ice cold water according to Pan et al. (1993). This reduced the number of ball metaphases and increased the mitotic index. As a final step, synchronized seedlings were fixed in freshly prepared Clarke's fixative (ethanol:glacial acetic acid, 3:1, v:v) for 24 h and stored at −20°C in 96% ethanol until use.

## Oligo Painting Probe Selection and Preparation

The oligo painting probe of *S. latifolia*, prepared for X chromosome, was designed using Chorus software as previously described by Han et al. (2015). Briefly, oligo sequences (45 nt; >75% similarity) specific to X chromosome, based on the *S. latifolia* female genome (PRJNA289891; Papadopulos et al., 2015), were selected throughout the X chromosome scaffolds anchored using an X genetic map. Repetitive sequences were discriminated and removed during oligo painting probe design by Chorus pipeline (Han et al., 2015). A total of 12 988 oligo sequences were selected to cover X-linked scaffolds. The oligo sequences were synthesized *de novo* as myTags 20K Immortal library by Arbor Biosciences (Ann Arbor, MI, United States; TATAA Biocenter, Göteborg, Sweden). Labeling and detection of the oligo painting probe followed the published protocol of Han et al. (2015). For labeling of oligo-RNA products, we used universal primers (Eurofins Genomics, Ebersberg, Germany) conjugated with the Cy3 (5′–Cy3–CGTGGTCGCGTCTCA–3′) or primers conjugated with the digoxigenin (5′–DIG CGTGGTCGCGTCTCA–3′), similarly as (Šimoníková et al., 2019). Digoxigenin was detected by FITC conjugated anti-DIG antibody (Roche Life Sciences).

The number of oligo sequences per scaffold, scaffold length, position on genetic map and scaffold ID are included in **Supplementary Table S2**.

## Chromosome and Probe Preparation

Chromosome spreads were obtained from multiple individuals from one population of studied species listed in **Supplementary Table S1**. Chromosomes were prepared as described in Bačovský et al. (2019) with minor modifications. Briefly, fixed root tips were washed 2× in distilled water 5 min, 2× in 0.001M citrate buffer 5 min and digested for 45–50 min in 1% enzyme mix (**Supplementary Table S3**) diluted in 0.001M citrate buffer. Chromosomes were squashed on to slides, freezed in liquid nitrogen and incubated for 5 min in freshly prepared Clarke's fixative. Prepared slides were used directly for fluorescence *in situ* hybridisation (FISH) or stored at −20°C in 96% ethanol until use.

FISH was performed as described by Schubert et al. (2016) using four different stringencies (**Supplementary Table S3**). The centromeric *Silene* tandem arrayed repeat (STAR-C) and subtelomeric tandem repeat (X43.1) were used as reference probes described in Bačovský et al. (2019). STAR-C is primarily located in centromeres on the X chromosome and autosomes, and on the Y in an additional two clusters based on stringency conditions (Hobza et al., 2007). Chromosome pictures were captured with Olympus AX70 microscope equipped with the cold cube camera. After image capture, all channels were processed with the software Adobe Photoshop free version CS2. A color histogram for each X and Y chromosome image was drawn using RGB profiler in ImageJ 1.52i FiJi[1]. These histograms display the distribution of DNA probes along each chromosome arm. RGB profiler was used on the same plot, for each type of green, red or blue selection as described in Mathur et al. (2012). The oligo painting probe was used in at least three individual experiments and each labeling pattern was scored in 30 metaphases/interphases per experiment. We did not observe any variation in signal patterns in studied populations.

# RESULTS

## Test for X Chromosome-Specific Oligo Probe Stringency and Oligo Painting Probe Signal Strength

We developed the X chromosome-specific oligo probe from female genomic sequences described in Papadopulos et al. (2015), using the approach described in Han et al. (2015). A total of 12 988 oligo sequences was selected from the entire currently available genomic data (**Supplementary Figure S1**; Papadopulos et al., 2015), covering on average 2.5–3 oligo sequence/kb (1.8–5.5 oligo sequence/kb) for the selected loci. The total density of selected oligo sequences from the X chromosome is below the recommended level of oligo sequence number per kilobase (0.03 oligo sequence/kb) (Han et al., 2015; Jiang, 2019). Nevertheless, for selected regions an average density of 1.8–5.5 oligo sequence/kb and the average number of oligo sequences is higher than the recommended standard of an oligo painting probe for metaphase chromosomes and single loci, 0.1–0.5 oligo sequence/kb (Han et al., 2015; Jiang, 2019).

To study potential rearrangements accompanying sex chromosome evolution, we used an X chromosome-specific probe in four species in the genus *Silene*, two dioecious (*S. latifolia* and *S. dioica*) and two gynodioecious (*S. vulgaris* and *S. maritima*). In *S. latifolia*, *S. dioica* and *S. vulgaris,* the oligo painting probe yields identical pattern in each species using direct (Cy3-tagged oligo sequences) and indirect labeling (digoxigenin tagged oligo sequences), and different hybridization stringencies (**Supplementary Table S3**). Only minor changes were observed in signal strength if the amount of oligo painting probe in *S. vulgaris* was increased (set to 1 μg per slide due to the weak overall chromosomal coverage). Nevertheless, 87 and 77% stringency produced very faint signal on the chromosomes of *S. maritima* (data not showed), using direct or indirect labeling and using the same amount of DNA (1 μg of oligo painting probe per slide). Therefore, we tested additional two hybridization stringencies, 68 and 62%, respectively, and we detected a similar pattern in *S. maritima* as in *S. vulgaris* (**Figure 1** and **Supplementary Figure S3**) (signal on three pairs of autosomes). Therefore, 68% hybridization stringency was

[1]https://imagej.nih.gov/ij/plugins/

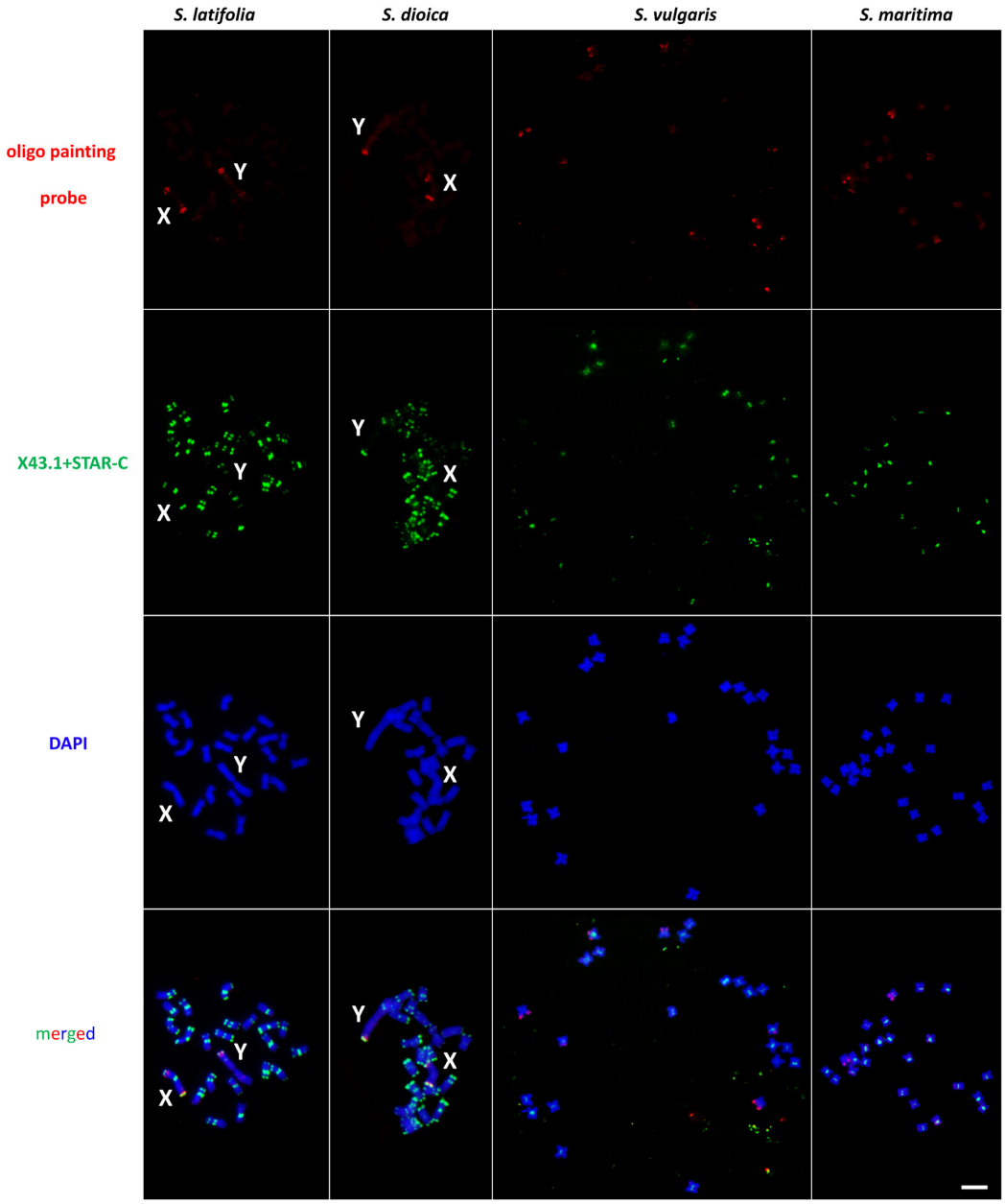



**FIGURE 1 |** Distribution of X chromosome-specific oligo probe on chromosomes in *Silene latifolia*, *S. dioica*, *S. vulgaris* and *S. maritima*. In *S. latifolia* and *S. dioica*, oligo painting probe has discrete signal only on X and Y chromosomes. In *S. vulgaris* and *S. maritima*, three pairs of autosomes can be differentiated from the whole genome. Scale bar = 5 μm.

applied in additional experiments for final analysis in all studied species using 1 μg of X chromosome-specific oligo painting probe per slide (**Figure 1** and **Supplementary Figure S3**).

## X Chromosome-Specific Oligo Probe Pattern in *Silene* Species

The designed oligo painting probe hybridized to both ends of the X chromosome arms (including PAR on the p-arm) and to PAR located on Y q-arm in *S. latifolia* and *S. dioica*. In addition, an extra oligo painting probe signal is clearly visible on the X p-arm, suggesting potential gene-rich locus in this (sub)telomeric region (**Figure 2**). On the Y, the probe colocalizes with X43.1 (sub)telomeric probe band on the Y q-arm (PAR region). The additional oligo painting probe signal was visible on the Y, as an interstitial region, located on the p-arm in both species (**Figure 2**). We did observe extra (weak) signal on the autosomes, using both Cy3- and digoxigenin-conjugated primers and various hybridization stringencies (77, 68, and 62%). Nevertheless, the extra (weak) signal was affected by low hybridization stringency.

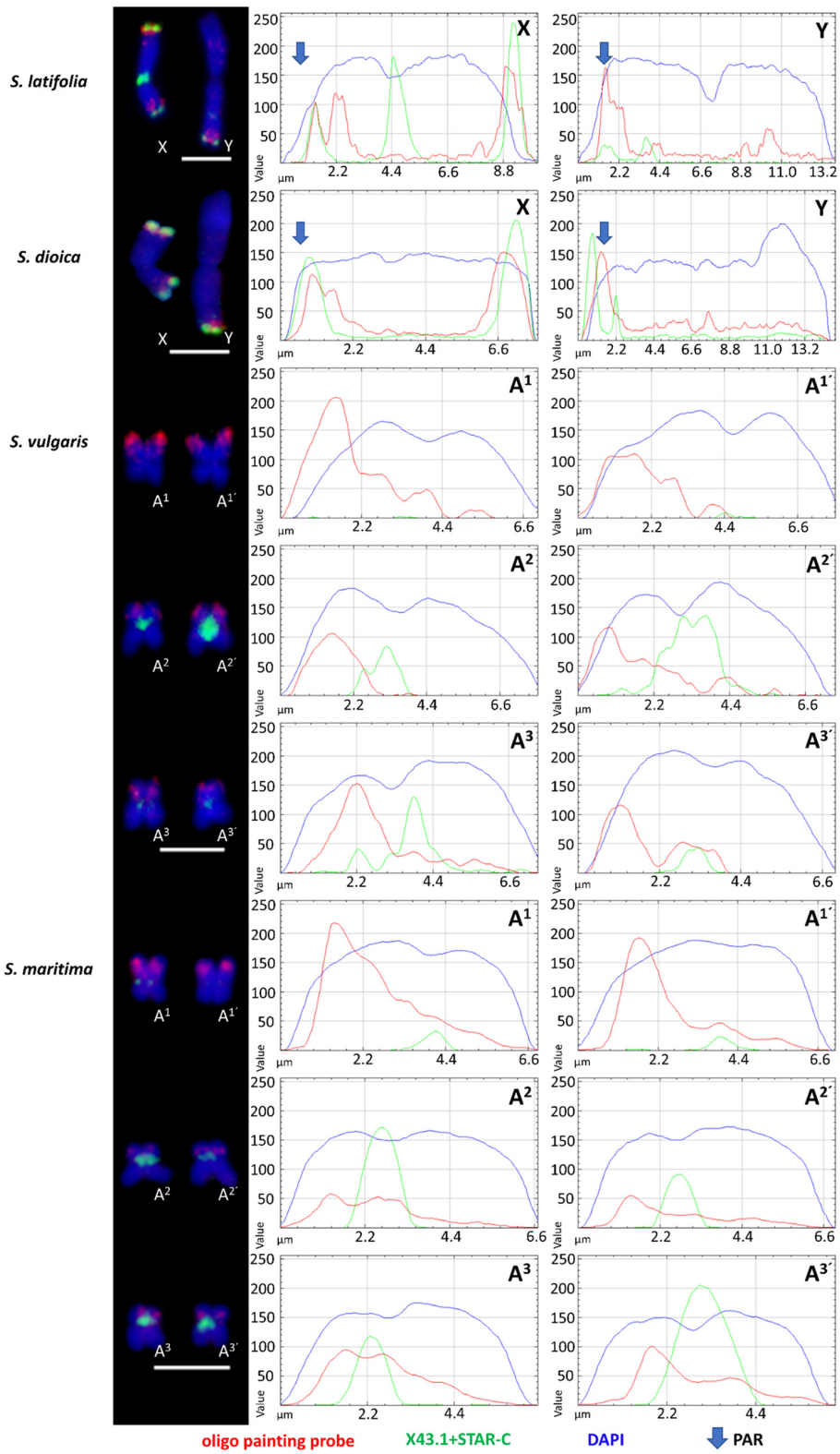



**FIGURE 2 |** Schematic distribution of X chromosome-specific oligo probe on sex chromosomes in *S. latifolia* and *S. dioica* and individual chromosomes in related gynodioecious species, *S. vulgaris* and *S. maritima*. Note the differences between **X** and **Y** chromosomes in *S. latifolia* and *S. dioica*. In S. *vulgaris* and *S. maritima,* the oligo painting probe signal is located on three pair of autosomes, numbered in this study as $\mathbf{A^1}$–$\mathbf{A^3}$ and $\mathbf{A^{1'}}$–$\mathbf{A^{3'}}$. X43.1, a subtelomeric probe, is presented only on sex chromosomes and autosomes in *S. dioica* and *S. latifolia*. Scale bar = 5 µm.

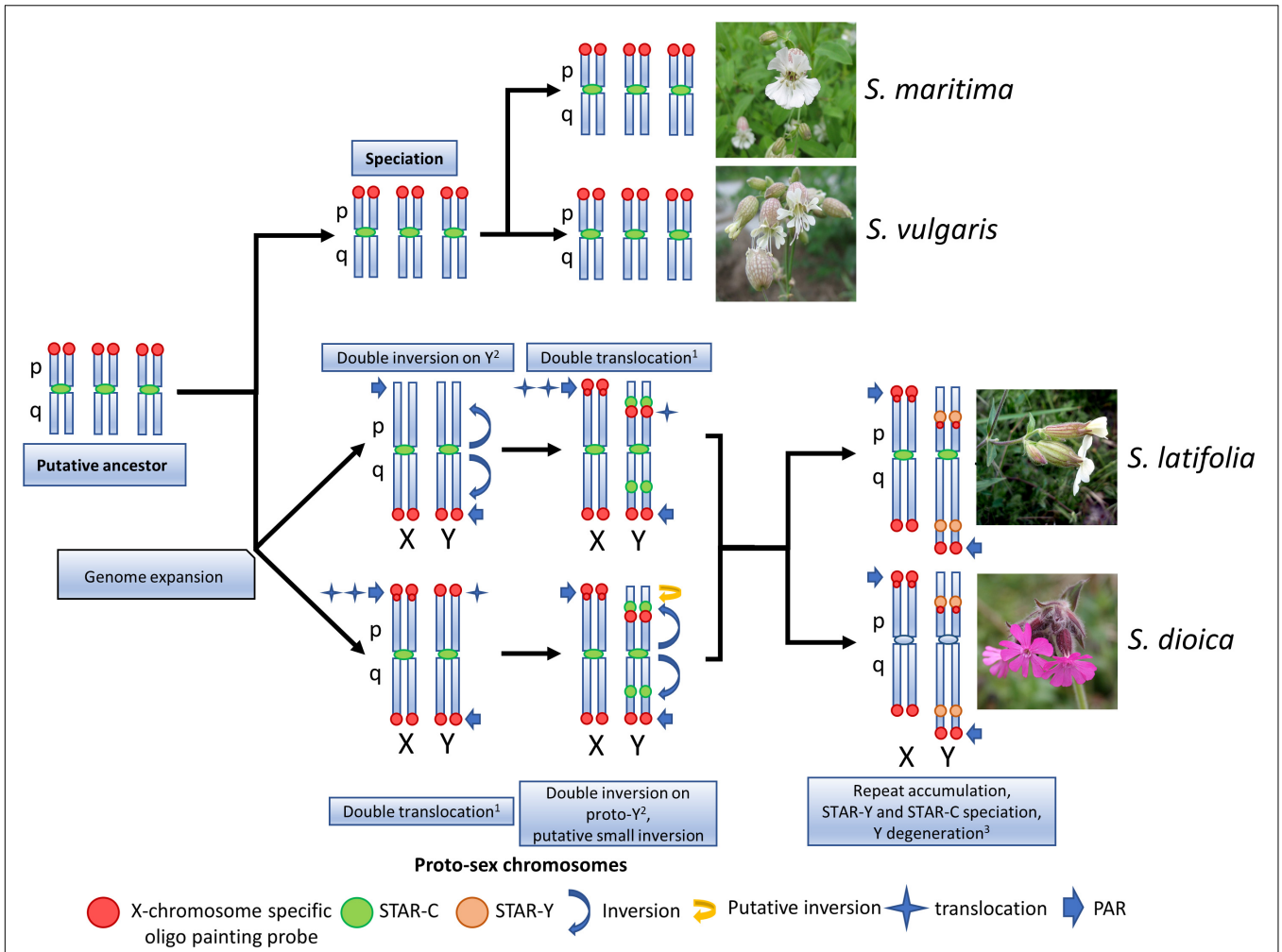



**FIGURE 3 |** Schematic view of the sex chromosomes evolution in *S. latifolia* and *S. dioica*. The sex chromosomes in *S. latifolia* and *S. dioica* underwent at least one double translocation[1] (Bergero et al., 2013; Campos et al., 2017) and double inversion[2] events (Hobza et al., 2007; Zluvova et al., 2007; Bergero et al., 2008a; Kazama et al., 2016). Compared to gynodioecious species (*S. vulgaris* and *S. maritima*), *S. latifolia* and *S. dioica* possess two sex chromosomes having signal from hybridized oligo painting probe. Based on the oligo painting probe signal in this study, two scenarios (supported by the literature) could led to the formation of sex chromosomes in *S. latifolia* and *S. dioica*. If the double translocation moving the segments of STAR-C satellite occurred first, then the autosomal parts have been translocated to proto-sex chromosome X PAR and interstitial Y p-arm region (upper scenario). In the second scenario, the autosomal parts were translocated to the sex chromosomes first, followed by double translocation and at least one putative translocation on the Y p-arm. Flower pictures downloaded on: https://commons.wikimedia.org/wiki/Main_Page.

In *S. latifolia* and *S. dioica* female karyotype, the oligo painting probe produced the same signal on both X chromosomes as on the X chromosome in males (**Supplementary Figure S4**). Thus, the X chromosome-specific oligo painting probe used in this work provides a highly reproducible signal in all studied species.

In *S. vulgaris* and *S. maritima*, application of oligo painting probe differentiated three pairs of autosomes, marked in this study as $A^1$–$A^{3'}$ (**Figure 2**). Compared to *S. maritima*, the decrease in hybridization stringency in *S. vulgaris* did not change the number of loci and signals on the chromosomes. In both gynodioecious species, the oligo painting probe labeled almost the entirety of the p-arms of $A^1$–$A^{1'}$, including (sub)centromere regions. Additionally, the oligo painting probe had a twofold stronger signal on the first pair of autosomes ($A^1$–$A^{1'}$), in *S. vulgaris* and *S. maritima*, than on the second and third ($A^2$–$A^{3'}$) autosomal pairs. The oligo painting probe hybridized to subtelomeric ($A^2$–$A^{2'}$) or more interstitial regions ($A^3$–$A^{3'}$) on these chromosomes (**Figures 1**, **2**).

In interphase, the X chromosome-specific probe differentiated the sex chromosome domains in *S. latifolia* and *S. dioica*, and the $A^1$–$A^{3'}$ autosomal regions in *S. vulgaris* and *S. maritima*. In the first two species, the oligo painting probe differentiated two subdomains located within one nucleus (**Supplementary Figures S2a,b**). In *S. vulgaris* and *S. maritima*, the oligo painting probe labeled three to six subdomains (**Supplementary Figures S2c,d**). Despite the average density being 1.8–5.5/kb in selected regions, the total coverage of the whole chromosome is only 0.03 oligo sequence/kb. The lower coverage is apparent (weaker signal)

in more relaxed chromatin state in early metaphase/prophase in all studied species (**Supplementary Figure S3**), showing that the oligo painting probe labeled (sub)telomere of the sex chromosomes and almost half of their length in the autosomes of *S. vulgaris* and *S. maritima*.

## DISCUSSION

The oligo painting probe specifically hybridized to pseudo-autosomal region (PAR) located on the Y q-arm and interstitial Y p-arm region, and to both sub-telomeric regions on the X chromosome in *S. latifolia* and *S. dioica*, enriched on the X p-arm sub-telomere. This distribution of the oligo painting probe signal correlates with the pattern of specific histone modifications for active chromatin, namely H3K4me1, H3K4me2, H3K4me3, H3K9ac and gene repressive mark H3K27me3 (Bačovský et al., 2019). Since the Y chromosome still contains many active genes (Bergero and Charlesworth, 2011), the additional band on the X p-arm together with the interstitial Y p-arm region probably represent unique/important gene clusters (Hobza et al., 2018). In related *S. vulgaris* and *S. maritima*, the oligo painting probe clearly differentiates three pairs of autosomes. In these gynodioecious species, the oligo painting probe labels three autosomal p-arms as shown on early metaphase/prophase chromosomes (**Supplementary Figure S3**). These chromosomal patterns support previous evidence that parts of three linkage groups in *S. vulgaris* were translocated to the pseudo-autosomal region in *S. latifolia* through a double translocation event (Bergero et al., 2013). Our results support previous studies showing expansion of the PAR region in *S. latifolia* (Bergero et al., 2013; Campos et al., 2017) and newly in *S. dioica* (**Figure 3**). Alternatively, rearrangements could be accompanied by whole chromosome fusion(s) as documented in other species. In Japan Sea stickleback fish, an ancestral Y chromosome fused with the autosome, forming a neo-Y chromosome and the $X_1X_2Y$ sex determination system (compared to closely related Pacific Ocean stickleback with XY system) (Yoshida et al., 2014). In *Rumex hastatulus*, North Carolina male race possesses $XY_1Y_2$, having the older Y (ancestral) chromosome heterochromatinised, and the younger Y with translocated autosomal part (Grabowska-Joachimiak et al., 2015). In *Silene* species "fusion" scenario is unlikely since it usually influences basic karyotype chromosome number (all studied species have $n = 12$). Moreover, existing genetic maps do not suggest large scale fusion events during karyotype evolution in studied species (Bergero et al., 2013). The existence of telomere-like sequences in *S. latifolia* Y chromosome (Uchida et al., 2002) can be explained as remnant of chromosomal inversion as was demonstrated in some *Silene* species (Filatov, 2005; Zluvova et al., 2005). It was reported that such chromosomal rearrangements included at least one paracentric and one pericentric inversion on the Y (Hobza et al., 2007; Bergero and Charlesworth, 2009). On the other hand, additional rearrangements that would suggest alternative scenarios cannot be excluded. Based on deletion mapping one inversion also occurred during the formation of the Y (Kazama et al., 2016). Bergero et al. (2013) and Campos et al. (2017) showed that PAR expanded through two step translocations. Our data support such translocation by the existence of three pair of autosomes (in gynodioecious species) corresponding to sex chromosomes in *S. latifolia* and *S. dioica*. Since sex chromosomes in studied dioecious species originated from one of these three chromosomal pairs, we assume that two additional loci were translocated on proto sex chromosomes in *S. latifolia* and *S. dioica* during sex chromosome evolution (**Figure 3**).

We have also tested the robustness of oligo painting method to study the dynamics of sex chromosomes and PAR in early metaphase/prophase (**Supplementary Figure S3**). In prophase/early metaphase in which the chromosomal (spatial) resolution limit is 10 times higher than in the metaphase and chromosomes are 10 times more extended (reviewed in Figueroa and Bass, 2010), the strength of the signal is relatively low. Therefore, it will be necessary to use more cytogenetic markers (together with oligo sequences) for discrimination of relaxed chromosomes such as e.g. specific antibodies against synaptonemal complexes in meiosis as described in Hurel et al. (2018).

Our X chromosome-specific oligo probe serves as useful tool to study the evolution of sex chromosomes in *S. latifolia*, *S. dioica* and their relatives. Since the genome of *S. latifolia* is still not fully sequenced we would like to leave the possibility that some sequences targeted by the probe might occur more than once in the genome open (e.g. low repetitive or duplicated). Nevertheless, compared to previous attempts and labeling of the sex chromosomes in *Silene* species using e.g. unique BAC clones (Lengerova et al., 2004) or dissected chromosomal probes (Hobza et al., 2004), the oligo painting probe provides an unique signal on both sex chromosomes and is also suitable to study other related species. In addition, the probe simplifies future analysis of chromosome pairing and facilitates the study of the dynamics of the PAR region in interphase or during cell division.

## DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

## AUTHOR CONTRIBUTIONS

VB, RČ, and RH conceived and designed the research. VB, RČ, EH, and DŠ conducted the experiments. DŠ and EH contributed the reagents. VB analyzed the data. VB wrote the manuscript. All the authors read and approved the manuscript.

## FUNDING

This work was supported by the Czech Science Foundation grants No. 19-02476Y and 18-06147S.

## ACKNOWLEDGMENTS

We would like to thank Chris Johnson for the English-language correction. We also like to thank the

reviewers for their thoughtful comments and efforts toward improving our manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00205/full#supplementary-material

**FIGURE S1 |** The distribution of X chromosome-specific oligo probe in X chromosome genetic map in *S. latifolia*. The average coverage was chosen on 2–3 oligo sequences/kb (1.8–5.5). Total size of X chromosome is estimated around 400 Mb.

**FIGURE S2 |** Distribution of X chromosome-specific oligo in interphase in *S. latifolia*, *S. dioica*, *S. vulgaris* and *S. maritima*. Oligo painting probe differentiates two sub-domains in *S. latifolia* **(a)** and *S. dioica* **(b)**, and three to six sub-domains in *S. vulgaris* **(c)** and *S. maritima* **(d)**. Each sub-domain is enlarged and marked by separated colors (white/yellow/orange) in merged channel. X43.1, a sub-telomeric probe, is presented only on sex chromosomes and autosomes in *S. latifolia* and *S. dioica*. Note the distribution of STAR-C in *S. vulgaris* and *S. maritima*. Scale bar = 10 μm.

**FIGURE S3 |** Distribution of X chromosome-specific oligo probe on prophase and early metaphase chromosomes in *S. latifolia*, *S. dioica*, *S. vulgaris* and *S. maritima*. 12 988 has coverage 2.5–3 oligo sequences/kb, reaching 1.8–5.5 oligo sequences/kb on selected loci. Oligo painting probe (in red) hybridizes to very end of X and Y chromosomes in (sub)telomeres in *S. latifolia* and *S. dioica*. In *S. vulgaris* and *S. maritima*, the oligo painting probe labels six pairs of autosomes, hybridizing to their p-arm. Although the signal strength is weaker compared to condensed metaphase chromosomes, the oligo painting probe clearly marks sex chromosomes and autosomes in all studied species. Scale bar = 10 μm.

**FIGURE S4 |** Distribution of the oligo painting probe in *S. latifolia* and *S. dioica* female karyotype. Oligo painting probe was hybridized on metaphase chromosomes and on interphase nuclei in *S. latifolia* **(a,c)** and in *S. dioica* **(b,d)**. The remnant of a cytoplasm (signal not attached to any chromosome) is visible in the bottom of the *S. dioica* **(b)**. Scale bar = 10 μm.

**TABLE S1 |** Plant material.

**TABLE S2 |** Oligo probe sequence scaffolds ID.

**TABLE S3 |** The composition of the enzyme mix and hybridisation stringency.

## REFERENCES


Bachtrog, D. (2006). A dynamic view of sex chromosome evolution. *Curr. Opin. Genet. Dev.* 16, 578–585. doi: 10.1016/j.gde.2006.10.007

Bachtrog, D., Jensen, J. D., and Zhang, Z. (2009). Accelerated adaptive evolution on a newly formed X chromosome. *PLoS Biol.* 7:e1000082. doi: 10.1371/journal.pbio.1000082

Bačovský, V., Hobza, R., and Vyskot, B. (2018). Technical review: cytogenetic tools for studying mitotic. *Methods Mol Biol.* 1675, 509–535. doi: 10.1007/978-1-4939-7318-7-30

Bačovský, V., Houben, A., Kumke, K., and Hobza, R. (2019). The distribution of epigenetic histone marks differs between the X and Y chromosomes in *Silene latifolia*. *Planta* 250, 487–494. doi: 10.1007/s00425-019-03182-7

Bergero, R., and Charlesworth, D. (2009). The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* 24, 94–102. doi: 10.1016/j.tree.2008.09.010

Bergero, R., and Charlesworth, D. (2011). Preservation of the Y Transcriptome in a 10-Million-year-old plant sex chromosome system. *Curr. Biol.* 21, 1470–1474. doi: 10.1016/j.cub.2011.07.032

Bergero, R., Charlesworth, D., Filatov, D. A., and Moore, R. C. (2008a). Defining regions and rearrangements of the *Silene latifolia* Y chromosome. *Genetics* 178, 2045–2053. doi: 10.1534/genetics.107.084566

Bergero, R., Forrest, A., and Charlesworth, D. (2008b). Active miniature transposons from a plant genome and its nonrecombining Y chromosome. *Genetics* 178, 1085–1092. doi: 10.1534/genetics.107.081745

Bergero, R., Qiu, S., Forrest, A., Borthwick, H., and Charlesworth, D. (2013). Expansion of the pseudo-autosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. *Genetics* 194, 673–686. doi: 10.1534/genetics.113.150755

Bernasconi, G., Antonovics, J., Biere, A., Charlesworth, D., Delph, L. F., Filatov, D., et al. (2009). Silene as a model system in ecology and evolution. *Heredity* 103, 5–14. doi: 10.1038/hdy.2009.34

Braz, G. T., He, L., Zhao, H., Zhang, T., Semrau, K., Rouillard, J.-M., et al. (2018). Comparative oligo-FISH mapping: an efficient and powerful methodology to reveal karyotypic and chromosomal evolution. *Genetics* 208, 513–523. doi: 10.1534/genetics.117.300344

Campos, J. L., Qiu, S., Guirao-Rico, S., Bergero, R., and Charlesworth, D. (2017). Recombination changes at the boundaries of fully and partially sex-linked regions between closely related Silene species pairs. *Heredity* 118, 395–403. doi: 10.1038/hdy.2016.113

Cermak, T., Kubat, Z., Hobza, R., Koblizkova, A., Widmer, A., Macas, J., et al. (2008). Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. *Chromosom. Res.* 16, 961–976. doi: 10.1007/s10577-008-1254-2

Charlesworth, D. (2016). Plant sex chromosomes. *Annu. Rev. Plant Biol.* 67, 397–420. doi: 10.1146/annurev-arplant-043015-111911

Charlesworth, D., and Charlesworth, B. (2005). Sex chromosomes: evolution of the weird and wonderful. *Curr. Biol.* 15, R129–R131. doi: 10.1016/j.cub.2005.02.011

Charlesworth, D., and Laporte, V. (1998). The male-sterility polymorphism of Silene vulgaris: analysis of genetic dat from two populations and comparison with Thymus vulgaris. *Genetics* 150, 1267–1282.

Cui, C., Shu, W., and Li, P. (2016). Fluorescence in situ hybridization: cell-based genetic diagnostic and research applications. *Front. Cell Dev. Biol.* 4:89. doi: 10.3389/fcell.2016.00089

Darolti, I., Wright, A. E., Sandkam, B. A., Morris, J., Bloch, N. I., Farré, M., et al. (2019). Extreme heterogeneity in sex chromosome differentiation and dosage compensation in livebearers. *Proc. Natl. Acad. Sci. U.S.A.* 116, 19031–19036. doi: 10.1073/pnas.1905298116

Figueroa, D. M., and Bass, H. W. (2010). A historical andmodern perspective on plant cytogenetics. *Briefings Funct. Genom. Proteom.* 9, 95–102. doi: 10.1093/bfgp/elp058

Filatov, D. A. (2005). Evolutionary history of *Silene latifolia* sex chromosomes revealed by genetic mapping of four genes. *Genetics* 170, 975–979. doi: 10.1534/genetics.104.037069

Filatov, D. A., Howell, E. C., Groutides, C., and Armstrong, S. J. (2009). Recent spread of a retrotransposon in the *Silene latifolia* genome, apart from the Y chromosome. *Genetics* 181, 811–817. doi: 10.1534/genetics.108.099267

Grabowska-Joachimiak, A., Kula, A., Książczyk, T., Chojnicka, J., Sliwinska, E., and Joachimiak, A. J. (2015). Chromosome landmarks and autosome-sex chromosome translocations in *Rumex hastatulus*, a plant with XX/XY1Y2 sex chromosome system. *Chromosom. Res.* 23, 187–197. doi: 10.1007/s10577-014-9446-4

Han, Y., Zhang, T., Thammapichai, P., Weng, Y., and Jiang, J. (2015). Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. *Genetics* 200, 771–779. doi: 10.1534/genetics.115.177642

He, L., Braz, G. T., Torres, G. A., and Jiang, J. (2018). Chromosome painting in meiosis reveals pairing of specific chromosomes in polyploid *Solanum* species. *Chromosoma* 127, 505–513. doi: 10.1007/s00412-018-0682-9

Hobza, R., Hudzieczek, V., Kubat, Z., Cegan, R., Vyskot, B., Kejnovsky, E., et al. (2018). Sex and the flower – developmental aspects of sex chromosome evolution. *Ann. Bot.* 122, 1085–1101. doi: 10.1093/aob/mcy130

Hobza, R., Kejnovsky, E., Vyskot, B., and Widmer, A. (2007). The role of chromosomal rearrangements in the evolution of *Silene latifolia* sex

chromosomes. *Mol. Genet. Genomics* 278, 633–638. doi: 10.1007/s00438-007-0279-0

Hobza, R., Lengerova, M., Cernohorska, H., Rubes, J., and Vyskot, B. (2004). FAST-FISH with laser beam microdissected DOP-PCR probe distinguishes the sex chromosomes of *Silene latifolia*. *Chromosom. Res.* 12, 245–250. doi: 10.1023/B:CHRO.0000021929.97208.1c

Hobza, R., and Vyskot, B. (2007). Laser microdissection-based analysis of plant sex chromosomes. *Methods Cell Biol.* 82, 433–453. doi: 10.1016/S0091-679X(06)82015-7

Huber, D., Voith von Voithenberg, L., and Kaigala, G. V. (2018). Fluorescence in situ hybridization (FISH): history, limitations and what to expect from micro-scale FISH? *Micro Nano Eng.* 1, 15–24. doi: 10.1016/j.mne.2018.10.006

Hughes, J. F., Skaletsky, H., Pyntikova, T., Graves, T. A., van Daalen, S. K. M., Minx, P. J., et al. (2010). Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature* 463, 536–539. doi: 10.1038/nature08700

Hurel, A., Phillips, D., Vrielynck, N., Mézard, C., Grelon, M., and Christophorou, N. (2018). A cytological approach to studying meiotic recombination and chromosome dynamics in Arabidopsis thaliana male meiocytes in three dimensions. *Plant J.* 95, 385–396. doi: 10.1111/tpj.13942

Jiang, J. (2019). Fluorescence in situ hybridization in plants: recent developments and future applications. *Chromosom. Res.* 27, 153–165. doi: 10.1007/s10577-019-09607-z

Kazama, Y., Ishii, K., Aonuma, W., Ikeda, T., Kawamoto, H., Koizumi, A., et al. (2016). A new physical mapping approach refines the sex-determining gene positions on the *Silene latifolia* Y-chromosome. *Sci. Rep.* 6:18917. doi: 10.1038/srep18917

Kejnovsky, E., Hobza, R., Cermak, T., Kubat, Z., and Vyskot, B. (2009). The role of repetitive DNA in structure and evolution of sex chromosomes in plants. *Heredity* 102, 533–541. doi: 10.1038/hdy.2009.17

Kejnovský, E., Michalovova, M., Steflova, P., Kejnovska, I., Manzano, S., Hobza, R., et al. (2013). Expansion of microsatellites on evolutionary young Y chromosome. *PLoS One* 8:45519. doi: 10.1371/journal.pone.0045519

Krasovec, M., Chester, M., Ridout, K., and Filatov, D. A. (2018). The mutation rate and the age of the sex chromosomes in *Silene latifolia*. *Curr. Biol.* 28, 1832.e–1838.e. doi: 10.1016/j.cub.2018.04.069

Kubat, Z., Zluvova, J., Vogel, I., Kovacova, V., Cermak, T., Cegan, R., et al. (2014). Possible mechanisms responsible for absence of a retrotransposon family on a plant Y chromosome. *New Phytol.* 202, 662–678. doi: 10.1111/nph.12669

Lengerova, M., Kejnovsky, E., Hobza, R., Macas, J., Grant, S. R., and Vyskot, B. (2004). Multicolor FISH mapping of the dioecious model plant, *Silene latifolia*. *Theor. Appl. Genet.* 108, 1193–1199. doi: 10.1007/s00122-003-1568-6

Mathur, J., Griffiths, S., Barton, K., and Schattat, M. H. (2012). Chapter eight - green-to-red photoconvertible meosfp-aided live imaging in plants. *Methods Enzymol.* 504, 163–181. doi: 10.1016/B978-0-12-391857-4.00008-2

Ming, R., Wang, J., Moore, P. H., and Paterson, A. H. (2007). Sex chromosomes in flowering plants. *Am. J. Bot.* 94, 141–150. doi: 10.3732/ajb.94.2.141

Muyle, A., Shearn, R., and Marais, G. A. B. (2017). The evolution of sex chromosomes and dosage compensation in plants. *Genome Biol. Evol.* 9, 627–645. doi: 10.1093/gbe/evw282

Pan, W. H., Houben, A., and Schlegel, R. (1993). Highly effective cell synchronization in plant roots by hydroxyurea and amiprophos-methyl or colchicine. *Genome* 36, 387–390. doi: 10.1139/g93-053

Papadopulos, A. S. T., Chester, M., Ridout, K., and Filatov, D. A. (2015). Rapid Y degeneration and dosage compensation in plant sex chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13021–13026. doi: 10.1073/pnas.1508454112

Rodríguez Lorenzo, J. L., Hobza, R., and Vyskot, B. (2018). DNA methylation and genetic degeneration of the Y chromosome in the dioecious plant *Silene latifolia*. *BMC Genomics* 19:540. doi: 10.1186/s12864-018-4936-y

Runyeon, H., and Prentice, H. C. (1997). Genetic differentiation in the Bladder campions, *Silene vulgaris* and *S. uniflora* (Caryophyllaceae), in Sweden. *Biol. J. Linn. Soc.* 61, 559–584. doi: 10.1111/j.1095-8312.1997.tb01807.x

Schubert, V., Ruban, A., and Houben, A. (2016). Chromatin ring formation at plant centromeres. *Front. Plant Sci.* 7:28. doi: 10.3389/fpls.2016.00028

Šimoníková, D., Němečková, A., Karafiátová, M., Uwimana, B., Swennen, R., Doležel, J., et al. (2019). Chromosome painting facilitates anchoring reference genome sequence to chromosomes in situ and integrated karyotyping in banana (Musa Spp.). *Front. Plant Sci.* 10:1503. doi: 10.3389/fpls.2019.01503

Široký, J., Lysák, M. A., Doležel, J., Kejnovský, E., and Vyskot, B. (2001). Heterogeneity of rDNA distribution and genome size in *Silene* spp. *Chromosom. Res.* 9, 387–393. doi: 10.1023/A:1016783501674

Stone, J. D., Koloušková, P., Sloan, D. B., and Štorchová, H. (2017). Non-coding RNA may be associated with cytoplasmic male sterility in *Silene vulgaris*. *J. Exp. Bot.* 68, 1599–1612. doi: 10.1093/jxb/erx057

Tang, S., Tang, Z., Qiu, L., Yang, Z., Li, G., Lang, T., et al. (2018). Developing new oligo probes to distinguish specific chromosomal segments and the A, B, D genomes of wheat (*Triticum aestivum* L.) Using ND-FISH. *Front. Plant Sci.* 9:1104. doi: 10.3389/fpls.2018.01104

Uchida, W., Matsunaga, S., Sugiyama, R., Shibata, F., Kazama, Y., Miyazawa, Y., et al. (2002). Distribution of interstitial telomere-like repeats and their adjacent sequences in a dioecious plant, *Silene latifolia*. *Chromosoma* 111, 313–320. doi: 10.1007/s00412-002-0213-5

Vyskot, B., and Hobza, R. (2015). The genomics of plant sex chromosomes. *Plant Sci.* 236, 126–135. doi: 10.1016/j.plantsci.2015.03.019

Xin, H., Zhang, T., Han, Y., Wu, Y., Shi, J., Xi, M., et al. (2018). Chromosome painting and comparative physical mapping of the sex chromosomes in *Populus tomentosa* and *Populus deltoides*. *Chromosoma* 127, 313–321. doi: 10.1007/s00412-018-0664-y

Yoshida, K., Makino, T., Yamaguchi, K., Shigenobu, S., Hasebe, M., Kawata, M., et al. (2014). Sex chromosome turnover contributes to genomic divergence between incipient stickleback species. *PLoS Genet.* 10:e1004223. doi: 10.1371/journal.pgen.1004223

Zluvova, J., Georgiev, S., Janousek, B., Charlesworth, D., Vyskot, B., and Negrutiu, I. (2007). Early events in the evolution of the *Silene latifolia* Y chromosome: male specialization and recombination arrest. *Genetics* 177, 375–386. doi: 10.1534/genetics.107.071175

Zluvova, J., Janousek, B., Negrutiu, I., and Vyskot, B. (2005). Comparison of the X and Y chromosome organization in *Silene latifolia*. *Genetics* 170, 1431–1434. doi: 10.1534/genetics.105.040444

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# APPENDIX V

**The puzzling fate of a lupin chromosome revealed by reciprocal oligo-FISH and BAC-FISH mapping**

Bielski, W., Książkiwicz, M., Šimoníková, D., Hřibová, E., Susek, K., Naganowska, B.

*Genes* **11**: e1489, 2020

doi: 10.3390/genes11121489

IF: 3.759

genes

MDPI

*Article*

# The Puzzling Fate of a Lupin Chromosome Revealed by Reciprocal Oligo-FISH and BAC-FISH Mapping

**Wojciech Bielski [1,*], Michał Książkiewicz [1], Denisa Šimoníková [2], Eva Hřibová [2], Karolina Susek [1] and Barbara Naganowska [1]**

[1] Department of Genomics, Institute of Plant Genetics, Polish Academy of Sciences, 60-479 Poznan, Poland; mksi@igr.poznan.pl (M.K.); ksus@igr.poznan.pl (K.S.); bnag@igr.poznan.pl (B.N.)

[2] Institute of Experimental Botany of the Czech Academy of Sciences, Centre of the Region Hana for Biotechnological and Agricultural Research, 77900 Olomouc, Czech Republic; simonikova@ueb.cas.cz (D.Š.); hribova@ueb.cas.cz (E.H.)

* Correspondence: wbie@igr.poznan.pl; Tel.: +48-616-550-245

Received: 6 November 2020; Accepted: 8 December 2020; Published: 10 December 2020


**Abstract:** Old World lupins constitute an interesting model for evolutionary research due to diversity in genome size and chromosome number, indicating evolutionary genome reorganization. It has been hypothesized that the polyploidization event which occurred in the common ancestor of the Fabaceae family was followed by a lineage-specific whole genome triplication (WGT) in the lupin clade, driving chromosome rearrangements. In this study, chromosome-specific markers were used as probes for heterologous fluorescence in situ hybridization (FISH) to identify and characterize structural chromosome changes among the smooth-seeded (*Lupinus angustifolius* L., *Lupinus cryptanthus* Shuttlew., *Lupinus micranthus* Guss.) and rough-seeded (*Lupinus cosentinii* Guss. and *Lupinus pilosus* Murr.) lupin species. Comparative cytogenetic mapping was done using FISH with oligonucleotide probes and previously published chromosome-specific bacterial artificial chromosome (BAC) clones. Oligonucleotide probes were designed to cover both arms of chromosome Lang06 of the *L. angustifolius* reference genome separately. The chromosome was chosen for the in-depth study due to observed structural variability among wild lupin species revealed by BAC-FISH and supplemented by in silico mapping of recently released lupin genome assemblies. The results highlighted changes in synteny within the Lang06 region between the lupin species, including putative translocations, inversions, and/or non-allelic homologous recombination, which would have accompanied the evolution and speciation.

**Keywords:** lupin; FISH; oligo-painting; oligonucleotide probes; comparative-mapping; chromosome evolution; cytogenetics; karyotype evolution; wild species

## 1. Introduction

Legumes (Fabaceae Lindl.) are the third largest family of higher plants with approximately 20,000 species, and second as to the harvested area and total production of 300 million metric tons of grain legumes on 190 million ha [1]. The family is diverse in many aspects, including plant morphology, habitat, and ecology, as well as genome size and evolution [2,3]. It has been assumed that the genome complexity and species diversity were promoted by whole-genome duplications (WGDs), which occurred in ancient legumes before the major diversification events [4]. The WGDs were further followed by polyploidization(s) in particular lineages, advancing their further expansions [5–9]. This is a typical evolutionary scenario, which is believed to have occurred in many angiosperm clades. Moreover, the WGDs were found to be related to global climate changes and periods with high diversification rates [10].

Grain legumes are an important source of nutrients for animal feed and human food production. However, the global market has been dominated by one species, a soybean. One of the proposed alternatives to it is the species from the genus *Lupinus* (Fabaceae), which have so far been used as an important component of animal feed (mainly beef and dairy cattle, sheep, pigs, poultry, finfish, and crustaceans) and soil fertilization (based on the symbiosis with nitrogen-fixing bacteria) [11,12]. In Europe, lupins have attracted wide attention in research and innovation programs, highlighted by the incorporation of *Lupinus* crop representatives into numerous European Union initiatives, such as GLIP, LEGATO, LUPICARP, PROTEIN2FOOD, INCREASE, and others. Lupin breeding is most advanced in Australia and many European countries, especially in those located in the eastern part of the continent. Parallel to the growing use of lupins in the food industry, their genomic studies and the knowledge on molecular and evolutionary mechanisms underlying the high variability observed in the genus have been advancing [13–18].

Genus *Lupinus* evolved about 50–55 mya and consists of approximately 270 species, subdivided into two groups: Old World lupins (OWLs) and New World lupins (NWLs) [8]. OWLs are native to Mediterranean region and North Africa and consist of 12 annual herbaceous species, including three crops: *L. angustifolius* (Narrow-leafed lupin), *L. albus* L. (white lupin), and *L. luteus* L. (yellow lupin) [19]. More importantly, from the evolutionary point of view, OWLs are characterized by high diversity in genome size (2C DNA amounts ranging from 0.97 pg to 2.44 pg) as well as basic (x = 6–9, 13) and somatic ($2n$ = 32–52) chromosome numbers [20,21]. These differences may reflect complex karyotype reorganizations, which occurred during the evolution of this group of plants. Their extant karyotypes were presumably shaped not only by polyploidization, which occurred in the common ancestor of papilionoids, but first of all by whole genome triplication (WGT), which happened at the beginning of the lupine lineage development and were followed by chromosomal rearrangements [13,15,22].

Based on the differences in geographic distribution, morphology (particularly traits of the corolla, pods, and seeds), specific protein polymorphism and alkaloid composition, the OWLs are divided into two groups: smooth-seeded (Malacospermae) group comprising the sections Angustifolius, Albus, Luteus, and Micranthus, and rough-seeded (Scabrispermae) group with the sections Atlanticus and Pilosus [23,24]. Recently identified species *L. mariae-josephi* H. Pascual is similar to Malacospermae in terms of chromosome number, but it is characterized by a unique 'intermediate' seed coat structure having common features with both rough- and smooth-seeded species [25]. Despite the recent progress, the knowledge on the course of evolution of species within OWL clade remains limited, and studies addressing evolutionary karyotype changes, especially those involving non-domesticated species, would facilitate the reconstruction of their phylogeny.

Comparative genome analysis can be done by in silico alignment using genome or transcriptome sequences. The use of chromosome-scale scaffolds can provide *a posteriori* insight into the chromosomal changes that took place during the karyotype evolution of the species of interest. Such approach was used recently in *L. albus*, leading to the conclusion that the current karyotype was shaped by 15 fissions and 21 chromosomal fusions, followed by whole genome triplication resulting in 17 major rearrangements [15]. An alternative strategy to in silico sequence-based analysis is the physical localization of specific DNA sequences on chromosomes by fluorescence in situ hybridization (FISH). FISH has been frequently used as a complementary and validation tool for in silico methods, such as fingerprinting-derived contig construction or de novo genome assembly [26]. Soon after its development in the early 1980s, FISH became the most important technique in plant cytogenetics and has been used frequently until now [27]. Indeed, despite so many high-throughput sequencing technologies developed, the number of FISH-based publications in the Web of Science database has not decreased during the past two decades [28].

To date, the most commonly used FISH probes to anchor genome sequences to chromosomes have been bacterial artificial chromosome (BAC) clones as they carry inserts long enough for proper visualization of hybridization signals [26]. In plants with relatively complex genomes, the utility of BAC clones in cross-species comparative studies is limited, mainly due to the abundance of dispersed

repetitive sequences in the clones. Thus, probes prepared from unique, single-copy sequences are more useful. However, the limiting factor of the use of single-copy sequences as FISH probes was the scarce availability of repeat-free, unique DNA clones, especially in non-model species with non-sequenced genomes. A significant breakthrough has been achieved recently by the innovations in synthesis and labeling of synthetic oligonucleotides, followed by the release of a publicly available software pipeline (Chorus), which together facilitate the development of locus-specific probes for FISH in a cost-effective manner [28,29]. The identification of large numbers of short (usually 45–50 bp), unique sequences across the whole genome assembly is done by the Chorus software, which enables advanced automation of this process. [29]. Designed oligonucleotides are then massively synthesized, labeled, and divided into pools, ready to use as probes for FISH. The oligo-based approach requires the availability of high-quality genome sequence for the oligonucleotide development, carrying also the representatives of the repetitive fraction of the genome. However, despite the rapid development of DNA sequencing techniques and genome sequence assembly algorithms, genomes of many species are not yet available. A solution may be to design probes using genome assembly from a closely related species, which may allow obtaining 'universal' probes that can be used in related species for heterologous hybridization. The 'oligo-painting' FISH provides an opportunity for a very precise cross-species analysis of chromosomal rearrangements [28]. The oligo-based technique in the last few years has proven to be effective in karyotyping and chromosome rearrangements identification in *Cucumis*, *Fragaria*, *Solanum*, or *Musa* species [29–32].

In the present study both traditional (BAC probes) and novel (oligo-painting) FISH approaches were harnessed to provide new insights into karyotype evolution among five OWL species. The species included one domesticated reference *L. angustifolius* ($2n = 40$) and four wild representatives differing in chromosome number, namely *L. cryptanthus* ($2n = 40$), *L. micranthus* ($2n = 52$), *L. cosentinii* ($2n = 32$), and *L. pilosus* ($2n = 42$).

## 2. Materials and Methods

### 2.1. Plant Material

One domesticated and four wild lupin species were used in this study (Table 1). The seeds were provided by the Polish *Lupinus* Gene Bank, Breeding Station Wiatrowo, Poznan Plant Breeders Ltd., Poznań, Poland. These were germinated in Petri dishes at 25 °C to obtain root tips that were suitable for mitotic chromosome isolation. Meiotic pachytene chromosomes were harvested from the young flower buds of the plants cultivated in controlled conditions (16 h of photoperiod, 22 °C; 8 h of night, 18 °C) in the Plant Growing Center of the Institute of Plant Genetics, Polish Academy of Sciences.

**Table 1.** General characteristics of the lupin species used in this study [20,21].

| Group | Section | Species | Accession | Chromosome Number ($2n$) | Genome Size (pg/2C DNA) |
|---|---|---|---|---|---|
| Smooth-seeded | Angustifolius | *L. angustifolius* L. | cv. 'Sonet' | 40 | 1.89 |
| | | *L. cryptanthus* Shuttlew | 96361 | 40 | 1.86 |
| | Micranthus | *L. micranthus* Guss. | 98552 | 52 | 0.98 |
| Rough-seeded | Pilosus | *L. cosentinii* Guss. | 98452 | 32 | 1.42 |
| | | *L. pilosus* Murr. | 98653 | 42 | 1.36 |

### 2.2. BAC Clone DNA Isolation and Labeling

Single copy BAC clones from the *L. angustifolius* nuclear genome BAC library [33] identified as Lang06-specific in the previous study [17] were used. Due to the dispersed mapping pattern of 067H16 BAC clone in wild lupins, one additional probe (059F07) specific to Lang06 [34] was used instead. Moreover, BAC clone 127N17 was also not included in FISH because of its overlapping sequence with the BAC 051D03. The complete list of used BAC clones and their alignment to pseudochromosomes is included in the Supplementary Materials (Supplementary Table S1). DNA isolation from BAC

clones was performed using miniprep kits (QIAprep Spin; Qiagen, Hilden, Germany). BAC DNA thus obtained was labeled by nick-translation (Sigma–Aldrich, St. Louis, MI, USA), using either digoxigenin-11-dUTP (Sigma–Aldrich) or tetramethylrhodamine-5-dUTP (Sigma–Aldrich). BAC clone DNA isolation and labeling were done as described by Susek et al. [35]. To obtain information on localization of BAC clones in the genome assembly, nucleotide sequences of inserts were downloaded from the NCBI database (accession numbers provided in Table S1) and aligned to the *L. angustifolius* pseudomolecules and/or scaffolds [18] using Basic Local Alignment Search Tool (BLAST) implemented in Geneious 9.1.8 program (Biomatters, Ltd., Auckland, New Zealand).

### 2.3. Oligonucleotide Probe Design, Synthesis, and Labeling

The procedure for preparing the oligonucleotide probes consisted of several successive steps shown in Figure 1.

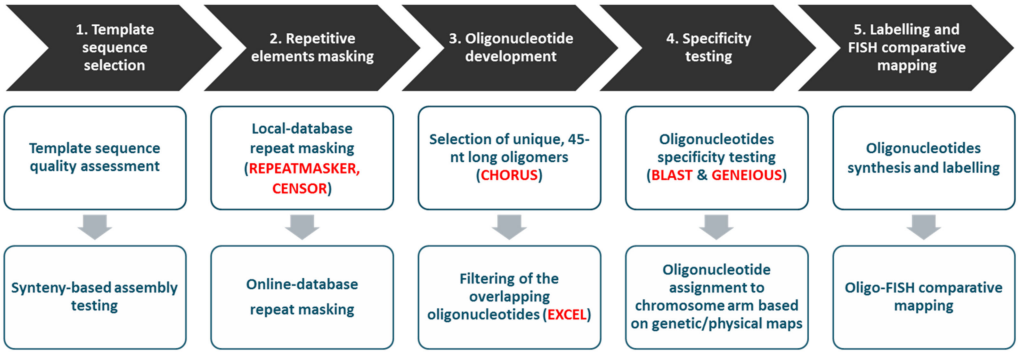



**Figure 1.** General scheme of oligonucleotide-based probe development.

Lang06 pseudochromosome (accession number CP023118) from the reference genome sequence of *L. angustifolius* cv. Tanjil [18] was selected. The total Lang06 pseudochromosome sequence length was 40,902,325 nt. The analysis of the Lang06 pseudochromosome included BLAST alignments to the sequences of related species: *L. albus* transcriptome aligned with *L. albus* genetic map [36], *Arachis duranensis* and *A. ipaensis* genomic sequences [37], and an earlier version of the *L. angustifolius* genome [38]. Based on the Megablast algorithm (word size: 28, e-value: $1 \times 10^{-10}$) performed in Geneious 9.1.8, no candidate miss-assemblies (>100 bp) were detected in the Lang06 pseudochromosome. Repetitive elements were masked subsequently using RepeatMasker [39] and CENSOR programs [40]. Chorus software [29] (Madison, WI, USA; github.com/zhangtaolab/Chorus2/) was used to generate a set of unique, 45 nt oligomers with default parameters (75% homology, dTM 10) based on the template sequence with masked repeats. Obtained oligonucleotides were mapped on the reference *L. angustifolius* cv. Tanjil sequence in Geneious 9.1.8, to determine their location in the genome (pseudochromosome) and the number of expected binding sites. This analysis was supplemented by the specificity test using BLAST with gradually decreasing similarity (to about 75%) and re-mapping of the oligonucleotides. Assignment of developed probes to the particular arms of the Lang06 chromosome was carried out by comparison of the density of markers on the *L. angustifolius* genetic map [41] with the physical distance between these markers. A significant drop in marker density combined with an increase in physical distance was interpreted as (peri)centromere. As lupin centromere regions are composed of many simple sequence repeats, oligonucleotides localized around centromeres were discarded from the probe synthesis.

Four libraries, each comprising 8000–20,000 oligonucleotides (45-mers), were synthesized by Arbor Biosciences (Ann Arbor, MI, USA). For the initial two libraries (O1 and O2), unlabeled 'immortal' oligonucleotides were ordered, and the probe labeling was performed according to Han et al. [29]. Briefly, the libraries were amplified using emulsion PCR [42], with F primer containing T7 RNA polymerase promoter, then washed with water-saturated diethyl ether and ethyl acetate, followed by a

purification step using a QIAquick PCR purification kit (Qiagen). Obtained DNA (~480 ng) was subjected to T7 in vitro transcription using a MEGAshortscript T7 Kit (ThermoFisher Scientific/Invitrogen, Waltham, MA, USA) at 37 °C for 4 h. The next step involved purification of the RNA using RNeasy Mini Kit (Qiagen) and reverse-transcription with either digoxigenin- or biotin-labeled R primer (Eurofins Genomics, Ebersberg, Germany) using Superscript II Reverse Transcriptase and SUPERase-In RNase inhibitor (ThermoFisher Scientific/Invitrogen). Finally, the RNA:DNA hybrids were cleaned with Quick-RNA MiniPrep Kit (Zymo Research, Freiburg im Breisgau, Germany) and hydrolyzed with RNase H (New England Biolabs, Ipswich, MA, USA) and RNase A (ThermoFisher Scientific/Invitrogen). To obtain single-stranded labeled oligonucleotide-probes, the additional purification with a Quick-RNA MiniPrep Kit (Zymo Research) was performed, followed by nuclease-free water elution. The remaining two libraries (O3 and O4) were synthetized by Arbor Biosciences as ready-to-use FISH probes, labeled with either digoxigenin or biotin.

### 2.4. Fluorescence In Situ Hybridization

Preparation of mitotic metaphase spreads and FISH with BAC-based probes (BAC-FISH) was done according to Susek et al. [17] with minor modifications. Young roots (~1 cm long) were treated in a solution of 40% (*v*/*v*) pectinase (Sigma–Aldrich, Darmstadt, Germany), 3% (*w*/*v*) cellulase (Sigma–Aldrich), and 1.5% (*w*/*v*) cellulase 'Onozuka R-10' (Serva, Heidelberg, Germany), in 37 °C for 60 min (lateral roots) or 100 min (primary roots).

Meiotic chromosome preparations were made from anthers collected from young buds. The buds were harvested and fixed in a solution of 96% ethyl alcohol/glacial acetic acid in a ratio of 3:1. The solution was not changed to a new one until the buds were completely discolored and then stored at −20 °C. The fixative was removed by a series of rinses in water and in citrate buffer. Single anthers or small buds devoid of crown petals were isolated using a stereoscopic microscope (SZX7 Olympus, Shinjuku, Tokyo, Japan). The material was digested using an enzyme cocktail, including 10% (*v*/*v*) pectinase (Sigma-Aldrich), 0.1% (*w*/*v*) cellulase (Sigma–Aldrich), and 0.1% (*w*/*v*) cytohelicase (Sigma–Aldrich) and incubated at 37 °C for about 150 min. Then the material was suspended in citrate buffer at 4 °C for 30 min. Finally, individual anthers were suspended in a drop of 60% acetic acid and isolated/transferred to a degreased glass slide. The material was covered with a coverslip and gently squashed. The quality of the material (number of meiotic divisions, degree of chromosome condensation, presence of cytoplasm) was assessed under the phase contrast microscope (BX41/CX41 Olympus). Selected high-quality slides were frozen at −80 °C (or on dry ice), the coverslip was removed and then dehydrated in 99.8% ethyl alcohol cooled to −20 °C for 30 min and dried at room temperature. The final quality was assessed under a phase contrast microscope, and then the slides were stored at −20 °C until use.

To localize the signals in both mitotic and meiotic stage, the chromosomes were counterstained with 4',6-diamidino-2-phenylindole (DAPI) in Vectashield Antifade Mounting Medium (Vector Laboratories, Burlingame, CA, USA). The fluorescent signals were acquired and examined using F-View monochromatic camera attached to an Olympus BX-60 epifluorescence microscope, pseudocolored in Wasabi (Hamamatsu Photonics, Hamamatsu, Shizuoka, Japan), superimposed using Micrografx (Corel Corporation, Ottawa, ON, Canada) Picture Publisher 10 software and GIMP 2.8.20.

Oligo-FISH procedure was as follows: selected slides with meiotic chromosomes were washed in 4% formaldehyde in 2 × Saline Sodium Citrate (SSC) buffer at room temperature for 10 min, then dehydrated in ethanol series for 2 min each (70%, 90%, and 99.8%). Then, the hybridization mix containing 50% (*v*/*v*) formamide, 10% (*w*/*v*) dextran sulfate in 2 × SSC, and 10 ng/µL of the labeled probe was added onto a slide and denatured at 80 °C for 3 min. Hybridization was carried out overnight at 37 °C. The particular probes (labeled with digoxigenin and biotin) were detected using anti-digoxigenin-FITC (Roche Applied Science, Penzberg, Germany) and streptavidin-Cy3 (ThermoFisher Scientific/Invitrogen), respectively. The chromosomes were counterstained with DAPI in Vectashield Antifade Mounting Medium (Vector Laboratories). Fluorescent signals were acquired and examined with Axio Imager Z.2 Zeiss microscope (Zeiss, Oberkochen, Germany) with Cool Cube

1 camera (Metasystems, Altlussheim, Germany). The capture of fluorescence signals and merging the layers were performed with ISIS software 5.4.7 (Metasystems, Heidelberg, Germany). Alternatively, the signals were acquired and examined with the hardware and software as described for BAC-FISH.

## 3. Results

*Oligonucleotide-Based Probe Development and Oligo-FISH*

When selecting a suitable template for the oligonucleotide probes, the following requirements were considered:

- The chromosome region should exhibit at least partial differentiation in related species, evidenced by previous cytogenetic studies or genome/linkage mapping;
- The chromosome region should have a low abundance of repetitive elements to allow the design of unique probes;
- Scaffolding in this region should be strongly supported by linkage mapping to avoid unintentional incorporation of fragments from other chromosomes;
- Chromosome-specific cytogenetic landmarks (i.e., BAC clones) should be available for this region to enable parallel use of two techniques—BAC-FISH and oligo-FISH.

Considering BAC-FISH results obtained in previous studies [17,35] and comparative mapping of *L. angustifolius* and *L. albus* genome assemblies and linkage maps [18,36,37], the pseudochromosome Lang06 sequence of *L. angustifolius* cv. Tanjil was selected as a template for oligonucleotide design. The first set of oligonucleotides was divided into two pools (libraries), specific for both arms (A and B) of the Lang06 chromosome. Based on the results of oligo-FISH, two more pools were later selected from the arm B of Lang06 to allow for fine mapping. The complete list of selected oligonucleotides is available at doi.org/10.5281/zenodo.4226537. A detailed description of each library is shown in Table 2.

**Table 2.** Characteristic of the designed oligonucleotides.

| Oligo Library ID | Lang06 arm (A/B) | Number of Oligonucleotides in Pool | Covered Region in Template Sequence (bp) | Template Coverage | Average Density |
|---|---|---|---|---|---|
| O1 | A | 20,115 | 32,610–11,689,408 | 28.50% | 0.58 oligo/kb |
| O2 | B | 19,926 | 29,400,911–40,902,159 | 28.19% | 0.58 oligo/kb |
| O3 | B | 10,214 | 27,898,181–36,244,755 | 20.41% | 0.82 oligo/kb |
| O4 | B | 8001 | 38,482,724–40,902,115 | 5.91% | 0.30 oligo/kb |

The combined scheme of the chromosomal distribution of the developed oligonucleotide libraries, including localization of the used Lang06-specific BAC clones and the detected repetitive elements, is shown in Figure 2.

The ready-to-use labeled oligonucleotide probes were hybridized to mitotic metaphase chromosome spreads of *L. angustifolius* cv. Tanjil, which served as a reference species and the template for the 45-nt oligomer development. Oligo-FISH resulted in visible signals covering chromosome Lang06 arms (Figure 3A,B). The Oligo1 (O1) probe mapped specifically to the A arm, the Oligo2 (O2) probe to the B arm (Figure 3A), the Oligo3 probe (O3) hybridized to the pericentromeric region of the B arm, while the Oligo4 (O4) probe with the telomere region of the B arm (Figure 3B). These observations confirmed that the probes hybridize specifically to target regions in the *L. angustifolius* genome. Moreover, the presence of locus-specific signals (i.e., the lack of dispersed signals) provided evidence that the repetitive sequence filtering process, based on several rounds of RepeatMasker and CENSOR masking, was effective. The localizations of all probes were also confirmed in meiotic chromosomes by two consecutive oligo-FISH reactions performed on the same slide. Images of oligo-FISH on *L. angustifolius* meiotic sample are presented in Supplementary Figure S1.

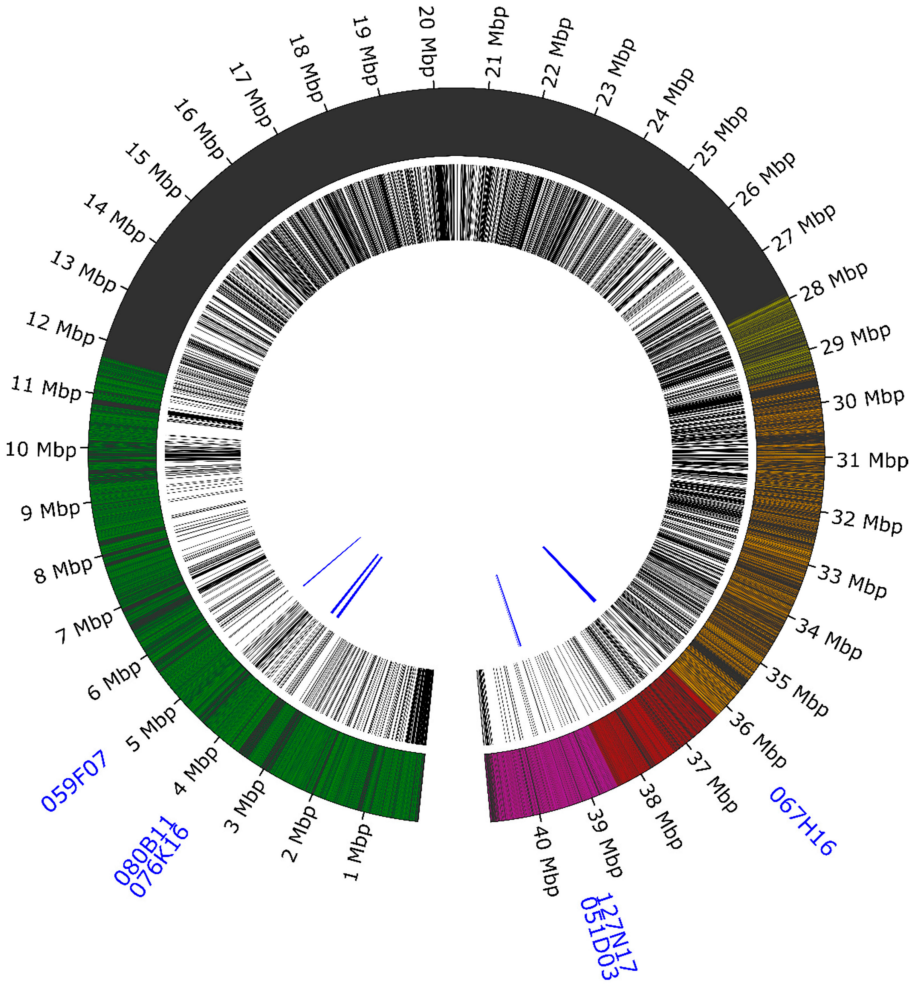



**Figure 2.** Schematic layout of the oligonucleotide libraries and repetitive sequences in pseudochromosome Lang06. The O1 library is marked in green, the O2 library in red, and the O3 library in yellow. Orange highlights the common region for the O2 and O3 libraries, whereas violet covers the region common to O2 and O4 libraries. Repetitive sequences are shown in black in the middle circle. Bacterial artificial chromosome (BAC) clone localization diagram is shown in blue in the inner circle.

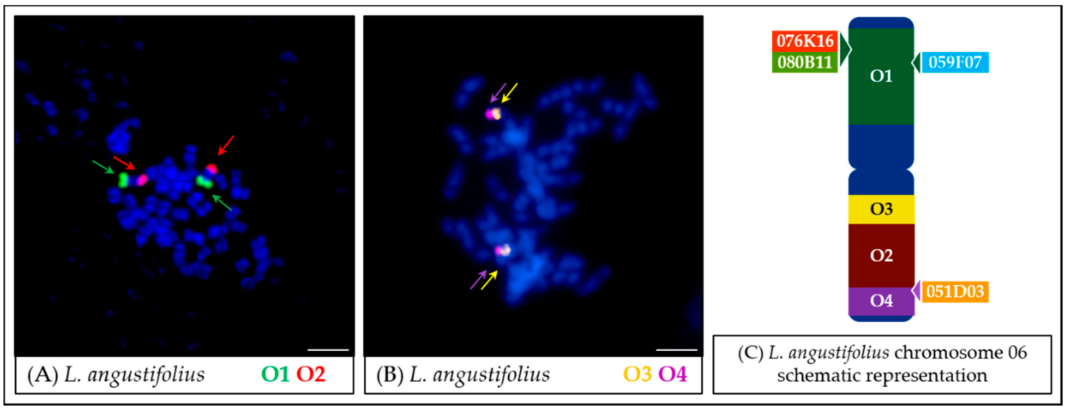



**Figure 3.** Fluorescence in situ hybridization (FISH) mapping of oligonucleotide probes in mitotic chromosomes of *L. angustifolius*. The positions of individual probes are marked by arrows. Probe colors are as follows: green (O1, Lang06 arm (**A**), red (O2, Lang06 arm (**B**)), yellow (O3, pericentromeric region of Lang06 arm B) and purple (O4, telomere region of Lang06 arm B). Scale bar: 5 µm. Chromosome Lang06 schematic representation (**C**), showing the positions of aligned particular oligonucleotide-based or BAC-based probes, was not drawn to scale.

The hybridization pattern of individual oligonucleotide probes in *L. cryptanthus* chromosomes remained identical to the reference species *L. angustifolius*: the O1 probe mapped in the A arm of the Lcry06 chromosome (Figure 4A), while the remaining probes hybridized to the Lcry06 arm B (Figure 4B). The results were verified by comparative mapping of the O1 probe with the Lang06-specific BAC clones: 080B11 (Figure 4C), 076K16 (Figure 4D), 051D03 (Figure 4E), and 059F07. The series of attempts was made to precisely visualize the oligonucleotide probes on meiotic material in studied wild lupin species, but the low signal intensity rendered the attempts unsuccessful.

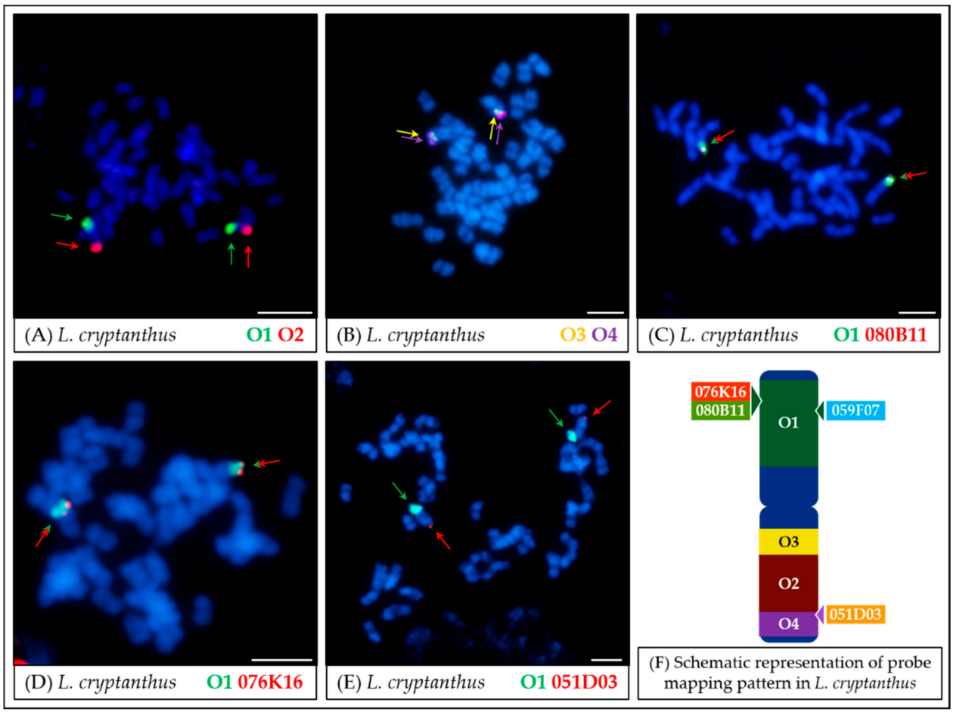



**Figure 4.** FISH mapping of oligonucleotide probes (**A**,**B**) and oligonucleotide combined with BAC clones (**C**–**E**) on mitotic metaphase chromosomes of *L. cryptanthus*. The positions of individual probes are marked by arrows. Probe colors are as follows: green (O1, Lang06 arm A), red (O2, Lang06 arm B), yellow (O3, pericentromeric region of Lang06 arm B) and purple (O4, telomere region of Lang06 arm B). Scale bar: 5 µm. Schematic representation of probe mapping pattern in *L. cryptanthus* chromosomes (**F**), showing observed positions of particular oligonucleotide-based or BAC-based probes, was not drawn to scale.

Comparative cytogenetic mapping in *L. micranthus* chromosomes showed that the structure of the arm A of the chromosome Lmic06 remained unchanged compared to *L. angustifolius* (Figure 5A). It was also revealed that the O1 probe co-localized with BAC clones 076K16, 059F07 (Figure 5D,E), and 080B11 (not shown). Structural differences in *L. micranthus* were revealed using probes from the arm B of the Lang06 chromosome; namely, the O2 probe was mapped to both arms of the chromosome. Because of the difference from Lmic06, it is named here as Lmic06′ (Figure 5A). BAC clone 051D03 detected on the B arm of Lang06 was co-localized with O2 on the A arm of Lmic06′ (Figure 5F). The O3 and O4 probes split into two separate arms of this Lmic06′ chromosome (Figure 5B,C). In case of the O3 probe, beside the two major loci, minor weaker signals were also observed in the chromosomes other than Lmic06 or Lmic06′ (Figure 5B,C).

The intensity of oligo-FISH signals was noticeably weaker in *L. cosentinii*. The specificity of the O1 probe to the arm A of the Lcos06 chromosome was preserved, whereas both O2 and O4 probes hybridized to two loci, the first localized in the arm B of the same chromosome as the O1 probe (Lcos06) and the second in a different chromosome (Lcos06′) (Figure 6A,B). The O3 probe revealed signals dispersed over multiple loci on *L. cosentinii* chromosomes. To confirm the results of oligo-FISH,

combinations of selected BAC clones with oligonucleotide probes were analyzed. These experiments showed that clone 059F07 shared the locus with the O1 probe (Figure 6C). Both 059F07 and 080B11 shared the chromosome (Lcos06) with the O2 probe (Figure 6D,E). Other BAC clones (051D03, 076K16) hybridized to multiple loci.

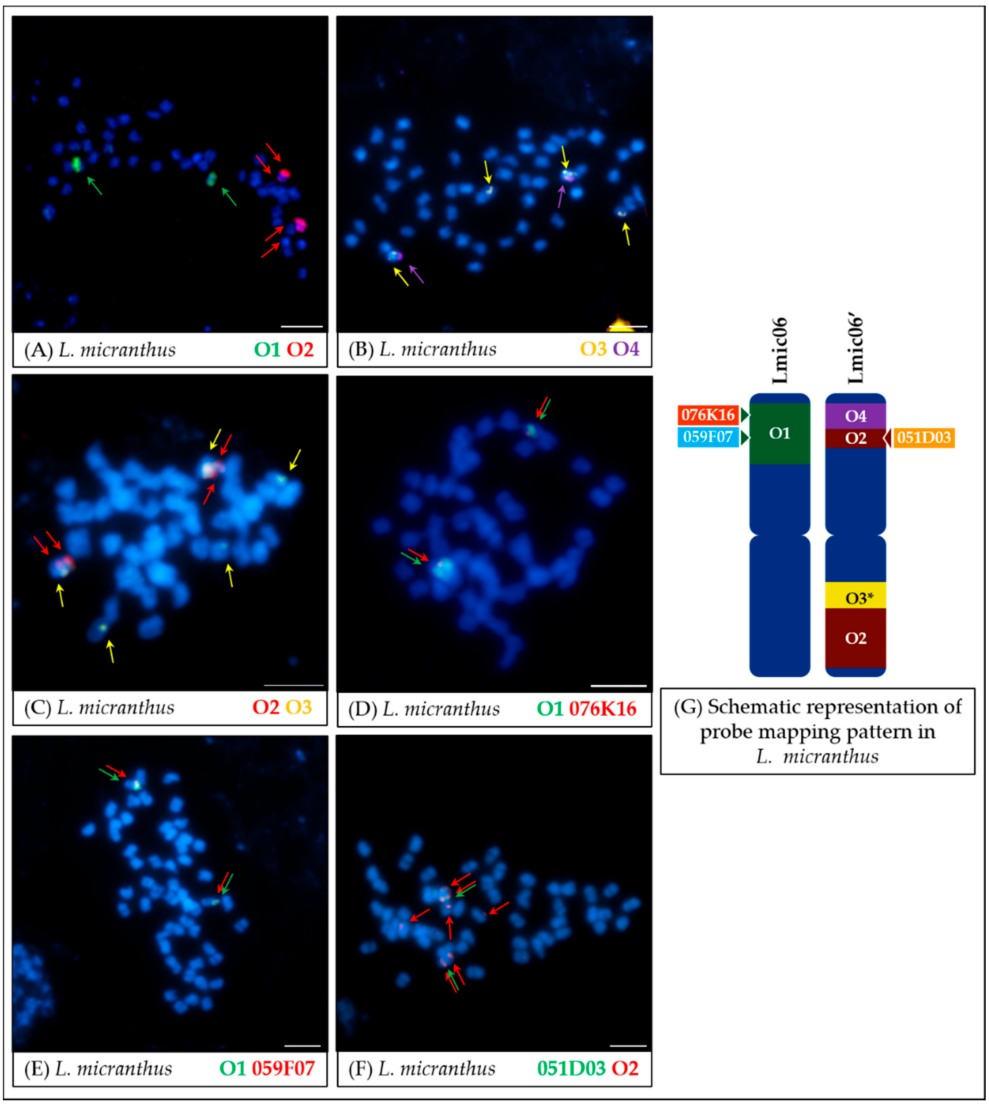



**Figure 5.** FISH mapping of oligonucleotide probes (**A**–**C**) and oligonucleotide probes combined with BAC clones (**D**–**F**) on mitotic chromosomes of *L. micranthus*. The positions of individual probes are marked by arrows. Probe colors are as follows: green (O1, Lang06 arm A), red (O2, Lang06 arm B), yellow (O3, pericentromeric region of Lang06 arm B), and purple (O4, telomere region of Lang06 arm B). Scale bar: 5 µm. Schematic representation of probe mapping pattern in *L. micranthus* chromosomes (**G**), showing observed positions of particular oligonucleotide-based or BAC-based probes, was not drawn to scale. O3*—beside two major loci, minor signals were also noticed.

In the last of the analyzed species, *L. pilosus*, a unique karyotyping pattern was observed. The O1 probe hybridized to the arm A of the chromosome Lpil06 and co-localized with BAC clone 080B11 (Figure 7A,C) and 059F07 as in the reference species, while the BAC clone 076K16 mapped to a different chromosome (Lpil06', Figure 7D). Although the remaining probes (O2, O3, and O4) hybridized to the arm B of the Lpil06 chromosome (as in reference species), their chromosome arrangement was reversed compared to the *L. angustifolius*: O3 hybridized to the near-telomere region, whereas O4 in the pericentromeric region (see enlarged fragment of Figure 7B). Moreover, beside the two major loci,

minor weaker signals in other chromosomes were noticed during the O3 probe mapping (Figure 7B). Interestingly, the O2 probe hybridized to two separate chromosomes (Lpil06 and Lpil06'' in which the O2 probe was co-localized with BAC clone 051D03, Figure 7F), and both of these were different from the chromosome Lpil06' on which BAC 076K16 was mapped (Figure 7D). Noteworthy, BAC clone 051D03 hybridizing to a different chromosome than Lpil06 was confirmed in our previous research using FISH comparative mapping with BAC clone 080B11 [17].

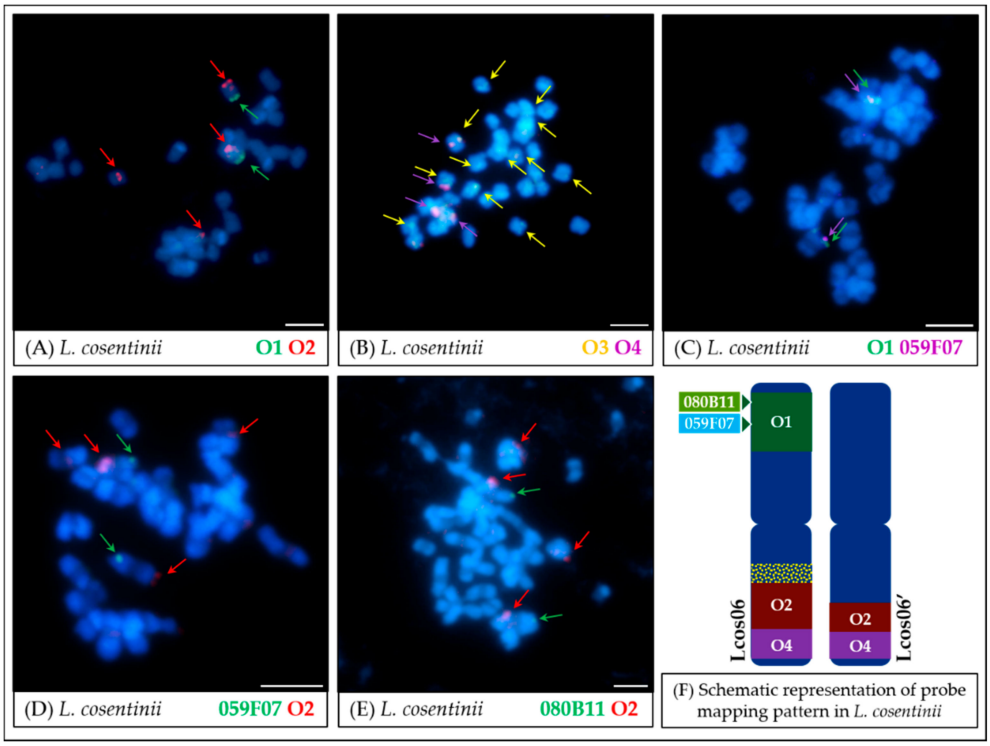



**Figure 6.** FISH mapping results of oligonucleotide probes (**A**,**B**) and oligonucleotide combined with BAC clones (**C**–**E**) in mitotic chromosomes of *L. cosentinii*. The positions of individual probes are marked by arrows. Probe colors are as follows: green (O1, Lang06 arm A), red (O2, Lang06 arm B), yellow (O3, pericentromeric region of Lang06 arm B), and purple (O4, telomere region of Lang06 arm B). Scale bar: 5 µm. Schematic representation of probe mapping pattern in *L. cosentinii* chromosomes (**F**), showing observed positions of particular oligonucleotide-based or BAC-based probes, was not drawn to scale. The O3 probe hybridized to multiple loci.

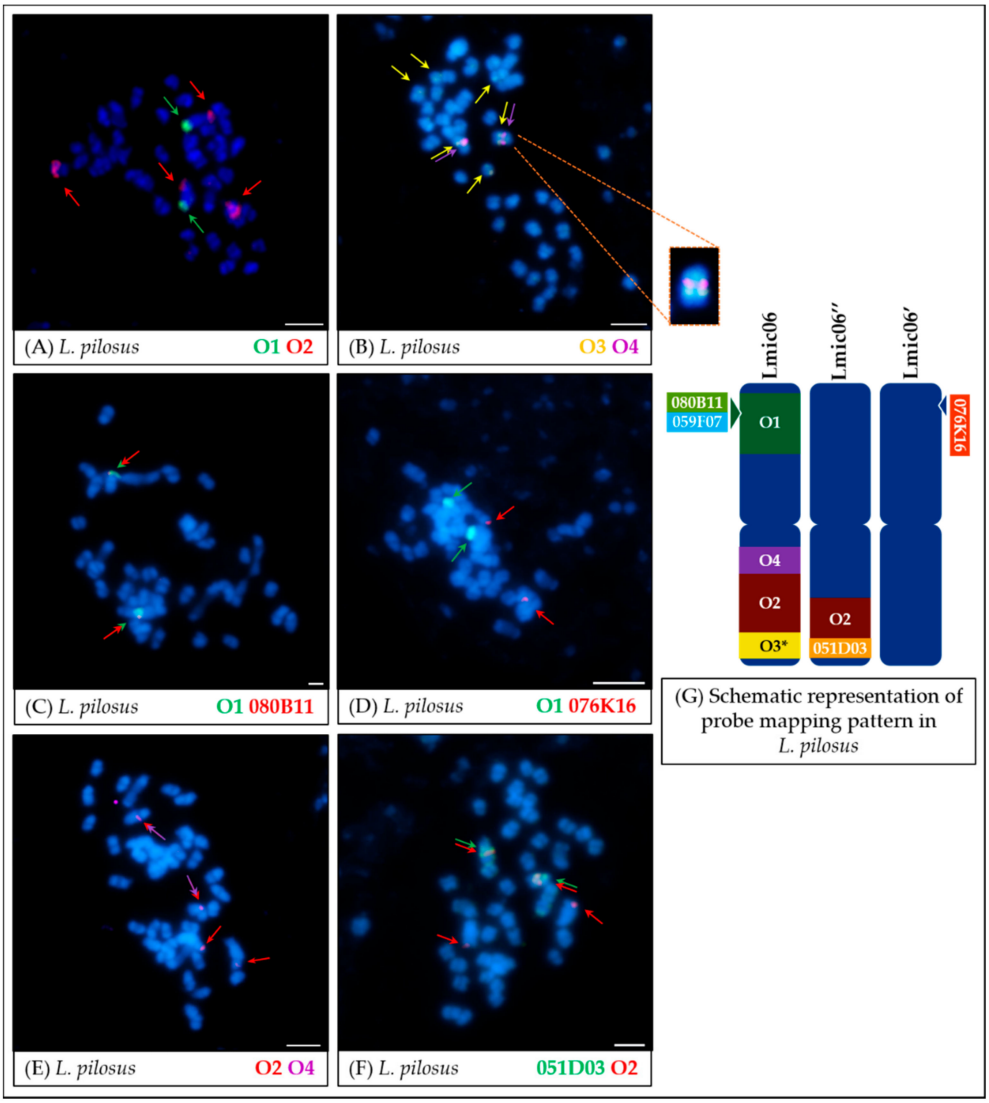



**Figure 7.** FISH mapping of oligonucleotide probes (**A**,**B**) and oligonucleotide probes combined with BAC clones (**C**–**F**) on mitotic chromosomes of *L. pilosus*. The positions of individual probes are marked by arrows. Probe colors are as follows: green (O1, Lang06 arm A), red (O2, Lang06 arm B), yellow (O3, pericentromeric region of Lang06 arm B), and purple (O4, telomere region of Lang06 arm B). Scale bar: 5 µm. Schematic representation of probe mapping pattern in *L. pilosus* chromosomes (**G**), showing observed positions of particular oligonucleotide-based or BAC-based probes, was not drawn to scale. O3*—besides two major loci, minor signals were also noticed. The fragment of Figure 7B showing the intra-chromosomal inversion was magnified (marked with orange, dashed line).

## 4. Discussion

### *4.1. Development of Oligonucleotide Probe Sets*

The BAC-FISH method provided preliminary insight into the diversification of karyotype structure among the studied lupin species [17,35]. Due to the fact that BAC clones covered only short fragments of chromosomes, there was a need to develop a new type of probe applicable for comparative mapping and covering substantially larger genomic regions, including the entire chromosome arms. Such an approach would allow for a more detailed examination of the differences between species. One of the possible solutions is a massive synthesis of oligonucleotide probes covering the unique regions of the chromosome [43]. As recently demonstrated in *Cucumis* [29], such probes can be used in heterologous

FISH to visualize structural chromosomal differences between species that differentiated up to 12 Mya. However, a lower intensity of fluorescent signals should be taken into account.

Whole-genome triplication, which is believed to predate *Lupinus* lineage evolution, occurred about 24.6 Mya [38], whereas differentiation of particular OWL species has been dated from about 1 to 10 mya [8,13,44]. Given these estimations, the evolutionary age of closely related nodes in OWL clade fits within the suggested limitations for the use of heterologous oligo-FISH. The optimal approach for OWL comparative mapping would comprise the hybridization of oligonucleotide probes designed for all *L. angustifolius* chromosomes. Such a strategy was implemented in *Zea mays* L. [45], when sets of oligonucleotide probes were developed, each of them designed to bind specifically to a single chromosome, making it possible to identify all 10 chromosomes of this species in consecutive FISH reactions. The undoubted constraint of such an approach when used in species with higher number of chromosomes would be linear multiplication of costs associated with the synthesis of chromosome-specific oligonucleotide sets. Moreover, such studies in OWLs are currently hampered by the uncertainties in super scaffold assembly constituting one of the reference *L. angustifolius* genome versions, highlighted by differences between the two recently published versions of the sequence [18,38]. Therefore, according to the assumption that the high quality of the template directly translates into the quality of the designed probes [28,29], Lang06 sequence was preselected for oligonucleotide probe design in this work. Our in silico comparative analysis indicated that this pseudochromosome does not contain significant errors (missing or incorrect fragments from other chromosomes), which could hinder the specificity of the developed probes. Moreover, our previous BAC-based studies highlighted this chromosome as a good candidate to track large-scale rearrangements [17]. Indeed, when *L. angustifolius* and *L. albus* genome assemblies were aligned to each other, Lang06 was found to be split between Chr16 and Chr23 in the latter species [13,15].

The oligonucleotide probes designed in this study were characterized by parameters similar to those used in the studies of *Cucumis* and *Solanum* species [29,31], namely the length (45 nt), homology (>75%), and the difference between the probes and hairpins Tm (dTM 10). Oligo–FISH performed on the mitotic and meiotic *L. angustifolius* chromosomes showed that all four probes mapped specifically in the Lang06 chromosome (Figure 3), consistent with the assumptions made during the probe design (Figure 2). The high specificity of developed probes highlighted the correctness of the Lang06 pseudochromosome assembly, at least in the scaffolds covered by the probes. However, it should be noted that some potential discrepancies (such as the incorrect orientation of sequence fragments covered by individual probes as well as the absence or multiplication of specific regions of the sequence) could go unnoticed due to the established parameters of the FISH reaction (stringency) and the characteristic of the designed probes. Thus, oligonucleotides labeled with a common fluorescent label are the source of a uniform signal, regardless of their arrangement within a single pool. The absence of fragments up to 1 Mbp in the template sequence may also go unnoticed due to the resolution limitations of the FISH performed on mitotic metaphase chromosomes [46]. This issue can be partially resolved by analyzing the fluorescence signal in a less condensed chromatin stage. To exemplify, such an oligo-FISH approach performed on *Musa acuminata* chromosomes revealed minor discrepancies in genomic sequence orientation in the form of the inversion of arms of the 1st, 6th, and 7th chromosomes in relation to the karyotype [32]. It is also worth emphasizing that, in our study, the mapping pattern of the developed probes in Lang06 reflected their organization in the pseudochromosome sequence in the libraries specific for individual arms (sets of oligonucleotide probes O1 and O2), as well as in the opposite regions of the B Lang06 arm (O3 and O4).

### 4.2. Comparative Mapping of Wild Lupin Species Using Oligonucleotide Probes

The Oligo–FISH results in *L. cryptanthus* (Figure 4) were identical to those obtained for the reference species, including the mapping pattern of oligonucleotide probes together with BAC clones. Each of the four oligonucleotide probes specifically mapped to a single region of the Lcry06 chromosome, and the intensity of hybridization signals was higher as compared to other species (*L. micranthus*, *L. cosentinii*,

and *L. pilosus*). This was consistent with the previous studies involving BAC-based probes [17,35] and in line with the hypothesis that *L. cryptanthus* is a wild form of *L. angustifolius* [21]. Genomic sequences of Lang06 and Lcry06 chromosomes seem very similar because there were no significant (and observable with the methods used) changes in the regions marked by oligonucleotide probes.

On the other hand, in *L. micranthus*, the oligo-FISH analysis showed the existence of significant structural genomic differences between this species and *L. angustifolius* (Figure 5). Similar to the BAC–FISH results [17], probes from different arms of the chromosome Lang06 landed onto two separate *L. micranthus* chromosomes (Lmic06 and Lmic06'). The novel information provided by the oligonucleotide-based approach was that the O2 probe mapped to both arms of the Lmic06', on the A arm co-localizing with both the O4 and BAC clone 051D03, and on the B arm with the O3 probe. Moreover, weaker signals, which were noticed during O3 probe mapping, might be related to the propagation of repetitive elements or duplication/insertion of short sequence fragments in the *L. micranthus* genome, collinear to the pericentromeric regions of Lang06. It should be noted that to visualize both 051D03 and O2 probe in FISH, the stringency was lowered (to about 65%); hence additional signals were visible for O2 (Figure 5F).

The analysis of the hybridization pattern of the oligonucleotide probes in *L. cosentinii* in reference to *L. angustifolius* (Figures 3 and 6) revealed that only one probe from the Lang06 arm A mapped to a single locus, retaining the reference pattern. Two probes from the B arm (O2 and O4) were mapped together in the B arm of Lcos06 chromosome, but also hybridized to another chromosome (Lcos06'). This might be the result of duplication and/or translocation of the arm B of Lang06 chromosome (containing O1 and O2 probes) to Lcos06' chromosome. The O3 probe in *L. cosentinii* was the only oligonucleotide probe that hybridized to multiple loci. It is probably a reflection of a specific type(s) of repetitive sequences in the *L. cosentinii* genome, contributing to the 'dispersed' mapping pattern of the O3 probe (and BAC clones 051D03 and 076K16). It is possible that some of the oligonucleotides in the pericentromeric region of *L. angustifolius* (which the O3 probe was designed to target) belong to the group of repetitive sequences represented more abundantly (or grouped in appropriate clusters) in the genome of *L. cosentinii* than in the reference species. The presence of specific repetitive elements may also explain weak, additional signals observed during O3 probe mapping in *L. micranthus* and *L. pilosus* chromosomes.

The most recently diversified [8] among the analyzed set of species is *L. pilosus*, which generally revealed a similar Oligo–FISH mapping pattern to the reference *L. angustifolius* (Figure 7). However, the segment order in the arm B of Lpil06 chromosome was reversed as compared to the Lang06, most likely due to paracentric inversion. It was also noted that the O2 probe hybridized to loci on two chromosomes (Lpil06 and Lpil06'), highlighting the remnants of a hypothetical translocation/duplication. O1 probe in *L. pilosus* retained its reference hybridization pattern contrary to the BAC clone 076K16 from the same arm, which was mapped on a different chromosome. Sequence alignment of 076K16 clone to the *L. angustifolius* genome with anchored oligonucleotides showed a significant (>30%) decrease in the frequency of oligonucleotides comprising the O1 probe in the region matching BAC clone sequence, as compared to the average oligonucleotide density of the O1 probe. Most likely, this decrease was due to the specific sequence properties in this genome region (e.g., the presence of palindromes or low complexity sequences) that disrupted the design of short oligonucleotide probes but not necessarily interfering with the hybridization of longer sequences, such as BAC clones. Noteworthy, alignment of 076K16 clone sequence to the *L. albus* genome assembly [15] resulted in three distinct, high-score matches (hits on chromosomes Lalb05, Lalb09, and Lalb16) with numerous rearrangements, which implies complex evolutionary reshuffling (Supplementary Table S2). The decrease in the frequency of oligonucleotides, along with potential evolutionary sequence changes, may explain the fact that no signal was detected at the same locus when the O1 probe and BAC clone 076K16 were mapped simultaneously in *L. pilosus* (Figure 7). On the other hand, the difference in loci between the BAC clone and the O1 probe itself may indicate the translocation of a short sequence fragment (including this clone) to another chromosome. Such a

phenomenon has been observed, among others, in *L. angustifolius* for the *LanFTc2* gene, which is located in the Lang17 chromosome region with a very high degree of collinearity within legumes. However, none of these collinear regions contains a corresponding homolog of this gene, despite the presence of such homologs for other genes adjacent to *LanFTc2* [47,48].

## 5. Conclusions

I. *L. cryptanthus* (Figures 4 and 8) is the only species tested with no significant structural differences detected compared to *L. angustifolius*, which supports the hypotheses of a close relationship between these two lupins.

II. In the case of *L. micranthus* (Figures 5 and 8), evolutionally the oldest among the studied species, the probes specific to the Lang06 chromosome landed on two different chromosomes, which may represent the pattern in the common ancestor of Old World lupins. During the course of evolution and speciation, the two genome fragments were translocated to one chromosome.

III. In *L. cosentinii* (Figures 6 and 8), hybridization of both O2 and O4 to two different chromosomes, as well as the highest number of probes (BACs and oligonucleotides) dispersed on multiple loci, might be the result of duplication and/or translocation of the arm B fragment of Lang06 chromosome (containing O1 and O2 probes) to Lcos06' chromosome.

IV. Significant synteny changes detected in *L. pilosus* (Figures 7 and 8) were probably the result of a series of rearrangements, including translocation, paracentric inversion, and/or non-allelic homologous recombination, leading to the separation of probes derived from Lang06 into three individual *L. pilosus* chromosomes.

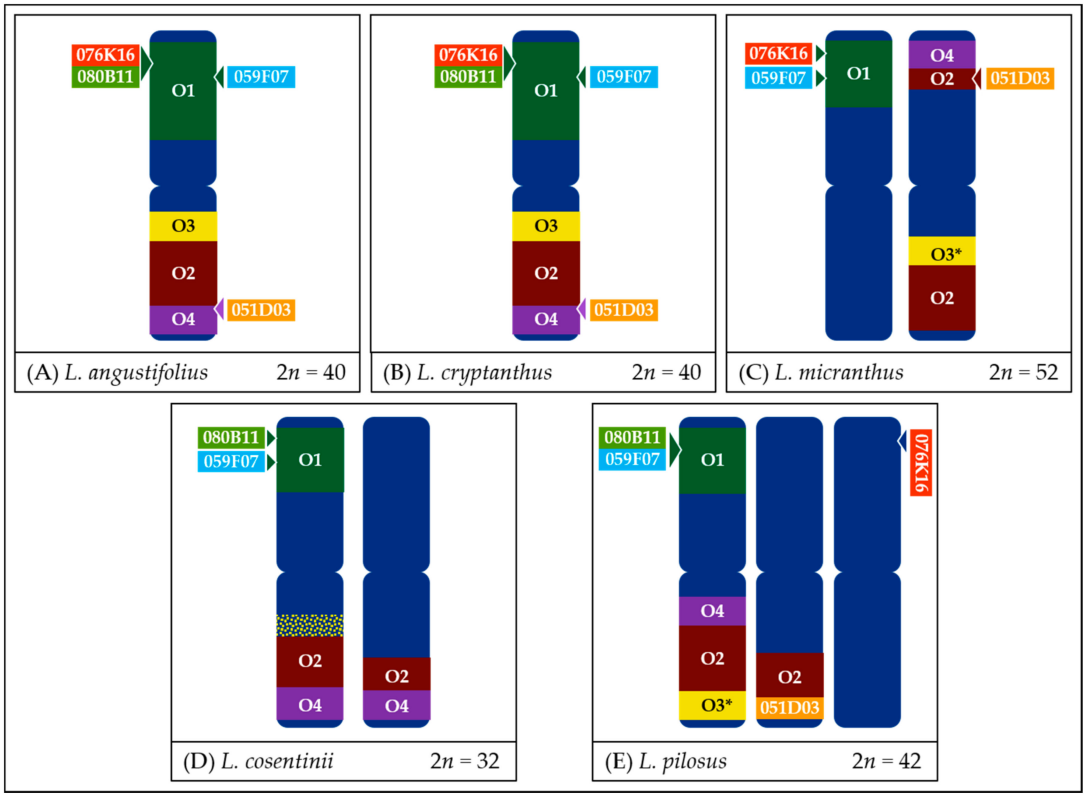



**Figure 8.** Schematic representation of probe mapping pattern, showing observed positions of particular oligonucleotide-based or BAC-based probes in *L. angustifolius* (**A**), *L. cryptanthus* (**B**), *L. micranthus* (**C**), *L. cosentinii* (**D**), and *L. pilosus* (**E**) chromosomes. In the case of probe O3 in *L. micranthus* and *L. pilosus*, beside two major loci, minor signals were also noticed. In *L. cosentinii*, the O3 probe hybridized to multiple loci. Chromosome schemes and probes length are not drawn to scale.

**Supplementary Materials:** The following are available online at http://www.mdpi.com/2073-4425/11/12/1489/s1, Figure S1: FISH mapping of oligonucleotide probes in meiotic chromosomes of *L. angustifolius*, Table S1: Position of BAC clones in the pseudochromosome Lang06 of *L. angustifolius* cv. Tanjil mapped using BLAST in Geneious R9.1.8., Table S2: Alignment of BAC clone 076K16 to *L. albus* genome mapped using BLAST in White Lupin Genome Sequence Server 1.0.11. The detailed list of used oligonucleotide probes is available online at doi.org/10.5281/zenodo.4226537.

**Author Contributions:** Conceptualization, W.B., M.K., K.S. and B.N.; Data curation, W.B.; Funding acquisition, W.B.; Investigation, W.B. and D.Š.; Methodology, W.B., D.Š., E.H. and K.S.; Project administration, W.B.; Resources, W.B. and E.H.; Supervision, B.N.; Validation, W.B.; Visualization, W.B.; Writing–original draft, W.B. and M.K.; Writing–review & editing, W.B., M.K., D.Š., E.H., K.S. and B.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by the National Science Centre, Poland (grant PRELUDIUM12 no. 2016/23/N/NZ2/01509) and ERDF project "Plants as a tool for sustainable global development" no. CZ.02.1.01/0.0/0.0/16_019/0000827.

**Acknowledgments:** The authors would like to express their gratitude to Jaroslav Doležel and other members of his laboratory for their kind support. We would also like to thank Arbor Biosciences for the synthesis of the oligonucleotide probes.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Gepts, P.; Beavis, W.D.; Brummer, E.C.; Shoemaker, R.C.; Stalker, H.T.; Weeden, N.F.; Young, N.D. Legumes as a model plant family. genomics for food and feed report of the cross-legume advances through genomics conference. *Plant Physiol.* **2005**, *137*, 1228–1235. [CrossRef] [PubMed]
2. Doyle, J.J.; Luckow, M.A. The rest of the iceberg. legume diversity and evolution in a phylogenetic context. *Plant Physiol.* **2003**, *131*, 900–910. [CrossRef] [PubMed]
3. Schmid, R.; Lewis, G.; Schrire, B.; Mackinder, B.; Lock, M. Legumes of the world. *TAXON* **2006**, *55*, 251. [CrossRef]
4. Bertioli, D.; Moretzsohn, M.C.; Madsen, L.H.; Sandal, N.; Leal-Bertioli, S.C.M.; Guimarães, P.M.; Hougaard, B.K.; Fredslund, J.; Schauser, L.; Nielsen, A.M.; et al. An analysis of synteny of Arachis with Lotus and Medicago sheds new light on the structure, stability and evolution of legume genomes. *BMC Genom.* **2009**, *10*, 45. [CrossRef] [PubMed]
5. Schmutz, J.; Cannon, S.B.; Schlueter, J.A.; Ma, J.; Mitros, T.; Nelson, W.; Hyten, D.L.; Song, Q.; Thelen, J.J.; Cheng, J.; et al. Genome sequence of the palaeopolyploid soybean. *Nature* **2010**, *463*, 178–183. [CrossRef] [PubMed]
6. Soltis, D.E.; Visger, C.J.; Marchant, D.B.; Soltis, P.S. Polyploidy: Pitfalls and paths to a paradigm. *Am. J. Bot.* **2016**, *103*, 1146–1166. [CrossRef] [PubMed]
7. Cannon, S.B.; McKain, M.R.; Harkess, A.; Nelson, M.N.; Dash, S.; Deyholos, M.K.; Peng, Y.; Joyce, B.; Stewart, C.N.; Rolf, M.; et al. Multiple polyploidy events in the early radiation of nodulating and nonnodulating legumes. *Mol. Biol. Evol.* **2015**, *32*, 193–210. [CrossRef]
8. Drummond, C.S.; Eastwood, R.J.; Miotto, S.T.S.; Hughes, C.E. Multiple continental radiations and correlates of diversification in Lupinus (Leguminosae): Testing for key innovation with incomplete Taxon sampling. *Syst. Biol.* **2012**, *61*, 443–460. [CrossRef]
9. LPWG. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny: The Legume Phylogeny Working Group (LPWG). *TAXON* **2017**, *66*, 44–77. [CrossRef]
10. Ren, R.; Wang, H.; Guo, C.; Zhang, N.; Zeng, L.; Chen, Y.; Zhang, X.; Qi, J. Widespread whole genome duplications contribute to genome complexity and species diversity in angiosperms. *Mol. Plant* **2018**, *11*, 414–428. [CrossRef]
11. Petterson, D.S. The use of Lupins in feeding systems—Review. *Asian-Australas. J. Anim. Sci.* **2000**, *13*, 861–882. [CrossRef]
12. Lucas, M.M.; Stoddard, F.L.; Annicchiarico, P.; Frias, J.; Emartinez-Villaluenga, C.; Esussmann, D.; Duranti, M.M.; Eseger, A.; Zander, P.; Pueyo, J. The future of lupin as a protein crop in Europe. *Front. Plant Sci.* **2015**, *6*, 705. [CrossRef] [PubMed]

13. Xu, W.; Zhang, Q.; Yuan, W.; Xu, F.; Aslam, M.M.; Miao, R.; Li, Y.; Wang, Q.; Li, X.; Zhang, X.; et al. The genome evolution and low-phosphorus adaptation in white lupin. *Nat. Commun.* **2020**, *11*, 1–13. [CrossRef] [PubMed]
14. Iqbal, M.M.; Erskine, W.; Berger, J.D.; Udall, J.A.; Nelson, M.N. *Genomics of Yellow Lupin (Lupinus luteus L.)*; Springer Science and Business Media LLC: Berlin, Germany, 2020; pp. 151–159.
15. Hufnagel, B.; Marques, A.; Soriano, A.; Marquès, L.; Divol, F.; Doumas, P.; Sallet, E.; Mancinotti, D.; Carrère, S.; Marande, W.; et al. High-quality genome sequence of white lupin provides insight into soil exploration and seed quality. *Nat. Commun.* **2020**, *11*, 1–12. [CrossRef] [PubMed]
16. Wyrwa, K.; Książkiewicz, M.; Koczyk, G.; Szczepaniak, A.; Podkowiński, J.; Naganowska, B. A tale of two families: Whole genome and segmental duplications underlie glutamine synthetase and phosphoenolpyruvate carboxylase diversity in narrow-leafed Lupin (*Lupinus angustifolius* L.). *Int. J. Mol. Sci.* **2020**, *21*, 2580. [CrossRef]
17. Susek, K.; Bielski, W.; Wyrwa, K.; Hasterok, R.; Jackson, S.A.; Wolko, B.; Naganowska, B. Impact of chromosomal rearrangements on the interpretation of Lupin Karyotype evolution. *Genes* **2019**, *10*, 259. [CrossRef]
18. Zhou, G.; Jian, J.; Wang, P.; Li, C.; Tao, Y.; Li, X.; Renshaw, D.; Clements, J.; Sweetingham, M.; Yang, H. Construction of an ultra-high density consensus genetic map, and enhancement of the physical map from genome sequencing in *Lupinus angustifolius*. *Theor. Appl. Genet.* **2018**, *131*, 209–223. [CrossRef]
19. Gladstones, J. Distribution, origin, taxonomy, history and importance. In *Lupins as Crop Plants: Biology, Production, and Utilization*; Gladstones, J.S., Atkins, C.A., Hamblin, J., Eds.; CAB International: Wallingford, CT, USA, 1998; pp. 1–36.
20. Pazy, B.; Heyn, C.; Herrnstadt, I.; Plitmann, U. Studies in populations of the old world Lupinus species. I. chromosomes of the East Mediterranean lupines. *Isr. J. Bot.* **1977**.
21. Naganowska, B.; Wolko, B.; Śliwińska, E.; Kaczmarek, Z. Nuclear DNA content variation and species relationships in the genus Lupinus (Fabaceae). *Ann. Bot.* **2003**, *92*, 349–355. [CrossRef]
22. Masterson, J. Stomatal size in fossil plants: Evidence for polyploidy in majority of angiosperms. *Science* **1994**, *264*, 421–424. [CrossRef]
23. Plitmann, U.; Pazy, B. Cytogeographical distribution of the Old World Lupinus. *Webbia* **1984**, *38*, 531–540. [CrossRef]
24. Gladstones, J. *Lupins of the Mediterranean Region and Africa*; Western Australian Department of Agriculture: Kensington, Australia, 1974.
25. Mahé, F.; Pascual, H.; Coriton, O.; Huteau, V.; Perris, A.N.; Misset, M.-T.; Ainouche, A. New data and phylogenetic placement of the enigmatic Old World lupin: Lupinus mariae-josephi H. Pascual. *Genet. Resour. Crop. Evol.* **2010**, *58*, 101–114. [CrossRef]
26. Jiang, J.; Gill, B.S. Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. *Genome* **2006**, *49*, 1057–1068. [CrossRef] [PubMed]
27. Langer-Safer, P.R.; Levine, M.; Ward, D.C. Immunological method for mapping genes on Drosophila polytene chromosomes. *Proc. Natl. Acad. Sci. USA* **1982**, *79*, 4381–4385. [CrossRef]
28. Jiang, J. Fluorescence in situ hybridization in plants: Recent developments and future applications. *Chromosom. Res.* **2019**, *27*, 153–165. [CrossRef]
29. Han, Y.; Zhang, T.; Thammapichai, P.; Weng, Y.; Jiang, J. Chromosome-specific painting in Cucumis species using bulked oligonucleotides. *Genetics* **2015**, *200*, 771–779. [CrossRef]
30. Qu, M.; Li, K.; Han, Y.; Chen, L.; Li, Z.; Han, Y. Integrated karyotyping of Woodland Strawberry (*Fragaria vesca*) with Oligopaint FISH probes. *Cytogenet. Genome Res.* **2017**, *153*, 158–164. [CrossRef]
31. Braz, G.T.; He, L.; Zhao, H.; Marand, A.P.; Semrau, K.; Rouillard, J.-M.; Torres, G.A.; Jiang, J. Comparative Oligo-FISH mapping: An efficient and powerful methodology to reveal karyotypic and chromosomal evolution. *Genetics* **2017**, *208*, 513–523. [CrossRef]
32. Šimoníková, D.; Němečková, A.; Karafiátová, M.; Uwimana, B.; Swennen, R.; Doležel, J.; Hřibová, E. Chromosome painting facilitates anchoring reference genome sequence to chromosomes in situ and integrated karyotyping in banana (*Musa* Spp.). *Front. Plant. Sci.* **2019**, *10*, 10. [CrossRef]
33. Kasprzak, A.; Šafář, J.; Janda, J.; Doležel, J.; Wolko, B.; Naganowska, B. The bacterial artificial chromosome (BAC) library of the narrow-leafed lupin (*Lupinus angustifolius* L.). *Cell. Mol. Biol. Lett.* **2006**, *11*, 396–407. [CrossRef]

34. Wyrwa, K.; Książkiewicz, M.; Szczepaniak, A.; Susek, K.; Podkowiński, J.; Naganowska, B. Integration of *Lupinus angustifolius* L. (narrow-leafed lupin) genome maps and comparative mapping within legumes. *Chromosom. Res.* **2016**, *24*, 355–378. [CrossRef] [PubMed]
35. Susek, K.; Bielski, W.; Hasterok, R.; Naganowska, B.; Wolko, B. A first glimpse of Wild Lupin karyotype variation as revealed by comparative cytogenetic mapping. *Front. Plant. Sci.* **2016**, *7*, 1152. [CrossRef] [PubMed]
36. Książkiewicz, M.; Nazzicari, N.; Yang, H.; Nelson, M.N.; Renshaw, D.; Rychel, S.; Ferrari, B.; Carelli, M.; Tomaszewska, M.; Stawiński, S.; et al. A high-density consensus linkage map of white lupin highlights synteny with narrow-leafed lupin and provides markers tagging key agronomic traits. *Sci. Rep.* **2017**, *7*, 1–15. [CrossRef] [PubMed]
37. Bertioli, D.J.; Cannon, S.B.; Froenicke, L.; Huang, G.; Farmer, A.D.; Cannon, E.K.S.; Liu, X.; Gao, D.; Clevenger, J.; Dash, S.; et al. The genome sequences of Arachis duranensis and Arachis ipaensis, the diploid ancestors of cultivated peanut. *Nat. Genet.* **2016**, *48*, 438–446. [CrossRef] [PubMed]
38. Hane, J.K.; Ming, Y.; Kamphuis, L.G.; Nelson, M.N.; Garg, G.; Atkins, C.A.; Bayer, P.E.; Bravo, A.; Bringans, S.; Cannon, S.; et al. A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*), an emerging health food: Insights into plant-microbe interactions and legume evolution. *Plant. Biotechnol. J.* **2016**, *15*, 318–330. [CrossRef] [PubMed]
39. Smit, A.; Hubley, R.; Green, P. RepeatMasker Open-4.0. 2013–2015. Available online: http://www.repeatmasker.org (accessed on 8 December 2020).
40. Kohany, O.; Gentles, A.J.; Hankus, L.; Jurka, J. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinform.* **2006**, *7*, 474. [CrossRef]
41. Kamphuis, L.G.; Hane, J.K.; Nelson, M.N.; Gao, L.; Atkins, C.A.; Singh, K.B. Transcriptome sequencing of different narrow-leafed lupin tissue types provides a comprehensive uni-gene assembly and extensive gene-based molecular markers. *Plant. Biotechnol. J.* **2015**, *13*, 14–25. [CrossRef]
42. Murgha, Y.E.; Rouillard, J.-M.; Gulari, E. Methods for the preparation of large quantities of complex single-stranded oligonucleotide libraries. *PLoS ONE* **2014**, *9*, e94752. [CrossRef]
43. Beliveau, B.; Joyce, E.; Apostolopoulos, N.; Yilmaz, F.; Fonseka, C.; McCole, R.; Chang, Y.; Li, J.; Senaratne, T.; Williams, B. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21301–21306. [CrossRef]
44. Eastwood, R.; Drummond, C.; Schifino-Wittmann, M.; Hughes, C. *Diversity and Evolutionary History of Lupins-Insights from New Phylogenies. Lupins for Health and Wealth*; International Lupin Association: Canterbury, New Zealand, 2008; pp. 346–354.
45. Albert, P.S.; Marand, A.P.; Semrau, K.; Rouillard, J.-M.; Kao, Y.-H.; Wang, C.-J.R.; Danilova, T.V.; Jiang, J.; Birchler, J.A. Whole-chromosome paints in maize reveal rearrangements, nuclear domains, and chromosomal relationships. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 1679–1685. [CrossRef]
46. Jiang, J.; Gill, B.S. Nonisotopic in situ hybridization and plant genome mapping: The first 10 years. *Genome* **1994**, *37*, 717–725. [CrossRef] [PubMed]
47. Książkiewicz, M.; Rychel, S.; Nelson, M.N.; Wyrwa, K.; Naganowska, B.; Wolko, B. Expansion of the phosphatidylethanolamine binding protein family in legumes: A case study of *Lupinus angustifolius* L. Flowering Locus T homologs, LanFTc1 and LanFTc2. *BMC Genom.* **2016**, *17*, 820. [CrossRef] [PubMed]
48. Nelson, M.N.; Książkiewicz, M.; Rychel, S.; Besharat, N.; Taylor, C.M.; Wyrwa, K.; Jost, R.; Erskine, W.; Cowling, W.A.; Berger, J.; et al. The loss of vernalization requirement in narrow-leafed lupin is associated with a deletion in the promoter and de-repressed expression of a Flowering Locus T(FT) homologue. *New Phytol.* **2016**, *213*, 220–232. [CrossRef] [PubMed]

# APPENDIX VI

## Chromosome oligo painting facilitates the analysis of karyotype evolution in Musaceae

Šimoníková, D., Čížková, J., Karafiátová, M., Němečková, A., Doležel, J., Hřibová, E.

In: Abstracts of the "22nd International Chromosome Conference". Prague, Czech Republic, 2018

# Chromosome oligo painting facilitates the analysis of karyotype evolution in *Musaceae*

Šimoníková D., Čížková J., Karafiátová M., Němečková A., Doležel J., Hřibová E.

Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, Olomouc, Czech Republic

## Introduction

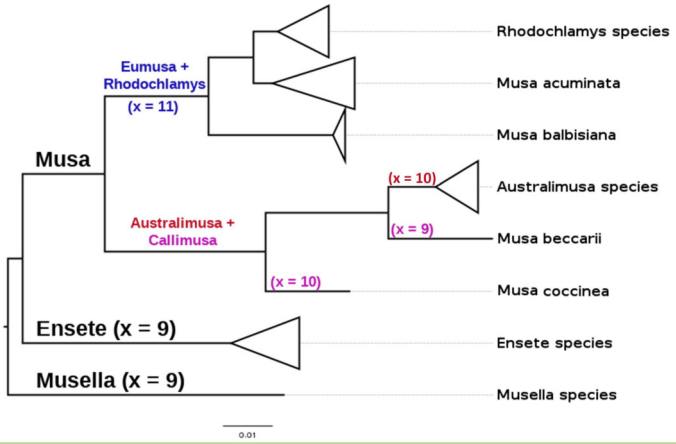
- Banana (*Musa* spp.) is one of the world's major fruit crops, a staple food and a significant export commodity for millions of people living in subtropical and tropical regions
- The banana genome (1Cx = 550 - 750 Mbp) is divided into morphologically similar chromosomes (1 - 2 µm long)
- Cultivated banana clones differ in ploidy (2x, 3x and 4x) and genomic constitution
- The majority of diploid and triploid clones originate from inter- or intraspecific crosses between *M. acuminata* (AA genome) and *M. balbisiana* (BB genome)
- Hybridization between different genomes leads to chromosomal rearrangements, which lead to abnormalities in meiosis and decreased fertility

## Aims of the study

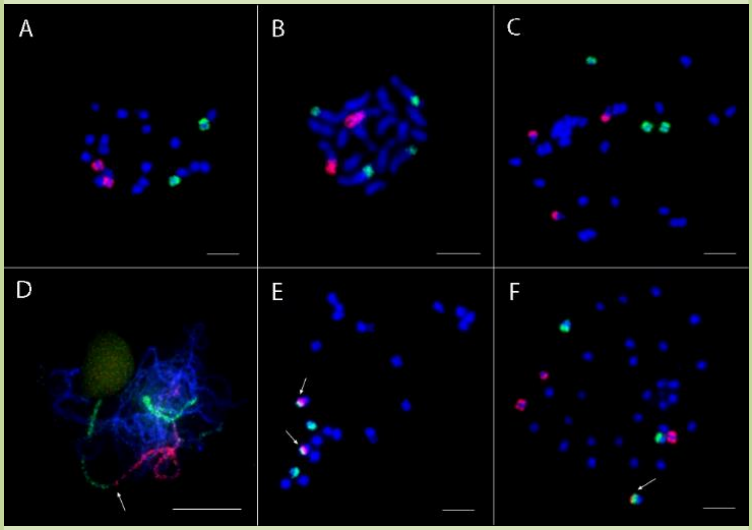
- Anchoring pseudomolecules to individual chromosomes and karyotype reconstruction using chromosome oligo painting
- Studying karyotype evolution in selected *Musaceae* accessions
- Identification of chromosomal rearrangements in edible hybrid clones (triploid plantains with AAB genome)

## Plant material

15 accessions representing the phylogeny of *Musaceae* family differing in basic chromosome number and genome size




## Results



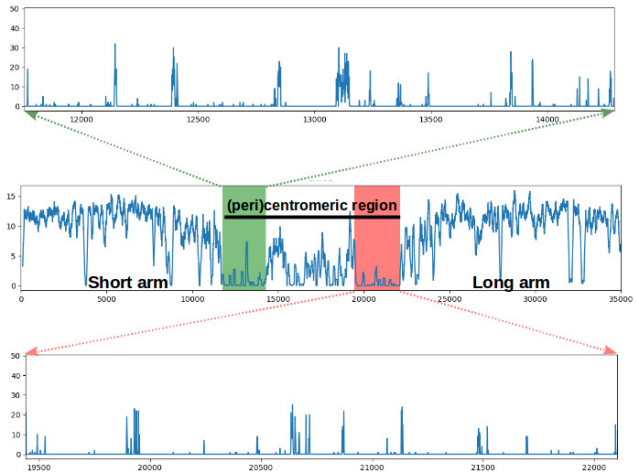

Examples of oligo painting on mitotic metaphase or meiotic pachytene chromosomes. (A) *M. acuminata* ssp. *banksii* (2n = 22, AA; chromosome 7 in red, chromosome 11 in green). (B) *M. beccarii* (2n = 18; chromosome 1 in red, chromosome 3 in green). (C) Plantain 'Obino l'Ewai' (2n = 3x = 33, AAB; chromosome 10 in red carrying a secondary constriction, chromosome 11 in green). (D) *M. balbisiana* 'Tani' (2n = 22, BB; chromosome 1 in red, chromosome 3 in green). (E) *M. balbisiana* 'Pisang Klutuk Wulung' (2n = 22, BB; chromosome 1 in red, chromosome 3 in green). (F) Plantain '3 Hands Planty' (2n = 3x = 33, AAB; chromosome 3 in red, chromosome 1 in green). Chromosomes were counterstained with DAPI (blue). Arrows point to a translocated region of chromosome 3 to chromosome 1 in *M. balbisiana* as well as to a chromosome 1 from B genome in plantain. Bars = 5 µm.

## Experimental design

**Oligo painting:**

1) *In silico* analysis

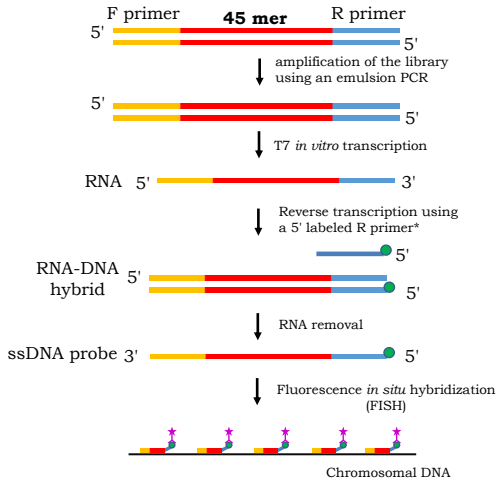
- Use of reference genome sequence *Musa acuminata* 'DH Pahang'
- Identification of 45nt oligos specific to a single chromosome/arm
- Selection of specific oligos (max 20 000 oligos per probe) using Chorus program (Han *et al.*, 2015)
- Massively parallel *de novo* synthesis by Arbor Biosciences (Ann Arbor, USA)




Example of oligo density specific to a pseudomolecule of chromosome 3 of *M. acuminata* 'DH Pahang' using Chorus program. Oligos coverage is marked in blue. Areas with low coverage (pericentromeric regions) were eliminated from probe preparation.

2) Probe preparation

Amplification and labeling of synthesized oligos:




* R primer is labeled indirectly (digoxigenin, biotin) or directly (FITC, CY3, CY5, DY-415)

## Conclusions

- Oligos specific to *M. acuminata* 'DH Pahang' can be utilized for analysis of other species of *Musaceae* family
- Density of oligo painting probes is sufficient to study chromosomal rearrangements on mitotic as well as on meiotic pachytene chromosomes
- Large chromosomal rearrangements can be visualized on highly condensed mitotic chromosomes
- Smaller chromosomal rearrangements (e.g. 1 - 4 translocation, published by Martin *et al.*, 2017) need to be analysed on more decondensed and larger meiotic pachytene chromosomes
- Oligo painting will enable studying the evolution of karyotypes in *Musaceae* and/or verification of *in silico* identified genome rearrangements

This research has been supported by National Program of Sustainability I (LO1204) and computing was supported by the National Grid Infrastructure MetaCentrum (LM2010005).

# APPENDIX VII

## Chromosome-specific oligo painting elucidates large variation in Eumusa genome

Šimoníková, D., Doležel, J., Hřibová, E.

In: Abstracts of the “International Conference on Polyploidy”. Ghent, Belgium, 2019

# Chromosome oligo painting elucidates large variation in *Eumusa* genome

Šimoníková Denisa, Doležel Jaroslav, Hřibová Eva

Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Šlechtitelů 31, Olomouc, Czech Republic

(contact: simonikova@ueb.cas.cz)

## Introduction

- Banana (*Musa* spp.) is one of the world's major fruit crops, a staple food and a significant export commodity for millions of people living in subtropical and tropical regions
- The banana genome (1Cx = 550 - 750 Mbp) is divided into morphologically similar chromosomes (1 - 2 µm long)
- Cultivated banana clones differ in ploidy (2x, 3x and 4x) and genomic constitution
- The majority of diploid and triploid clones originate from inter- or intra-specific crosses between diploid *M. acuminata* (genome A) and *M. balbisiana* (genome B)
- Hybridization between different genomes leads to chromosomal rearrangements, which lead to abnormalities in meiosis and decreased fertility

## Aims of the study

- Anchoring pseudomolecules to individual chromosomes
- Molecular karyotype reconstruction using chromosome arm-specific oligo painting probes in combination with previously identified cytogenetic markers (Čížková *et al.*, 2013)
- Comparative analysis of molecular karyotypes in selected accessions representing *Eumusa* section

## *Musaceae* family

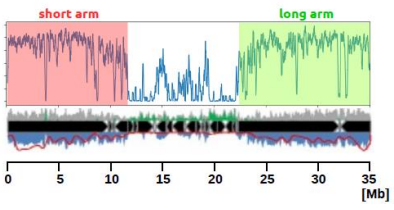
- *Genus: Musa*
  - *Section:* ***Eumusa* (x = 11): *M. acuminata, M. balbisiana, M. schizocarpa***
  - *Section: Rhodochlamys* (x = 11)
  - *Section: Australimusa* (x = 10)
  - *Section: Callimusa* (x = 9, 10)
- *Genus: Ensete* (x = 9)
- *Genus: Musella* (x = 9)

## Experimental design

### Oligo painting

1) *In silico* analysis

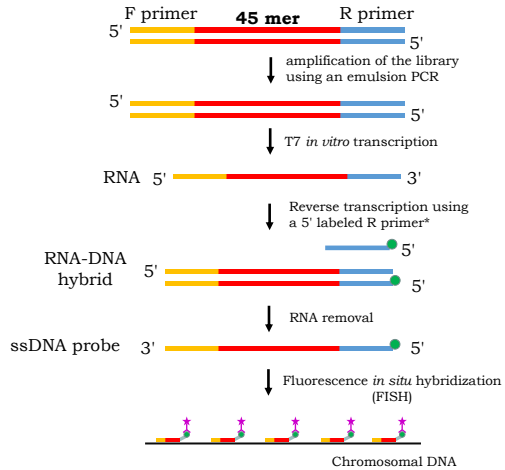
- Use of reference genome sequence *Musa acuminata* 'DH Pahang'
- Identification of unique oligos (45mers) covering a chromosome or specific region
- Selection of max 20 000 oligos per probe using Chorus pipeline (Han *et al.*, 2015)
- Massively parallel *de novo* synthesis by Arbor Biosciences (Ann Arbor, USA)



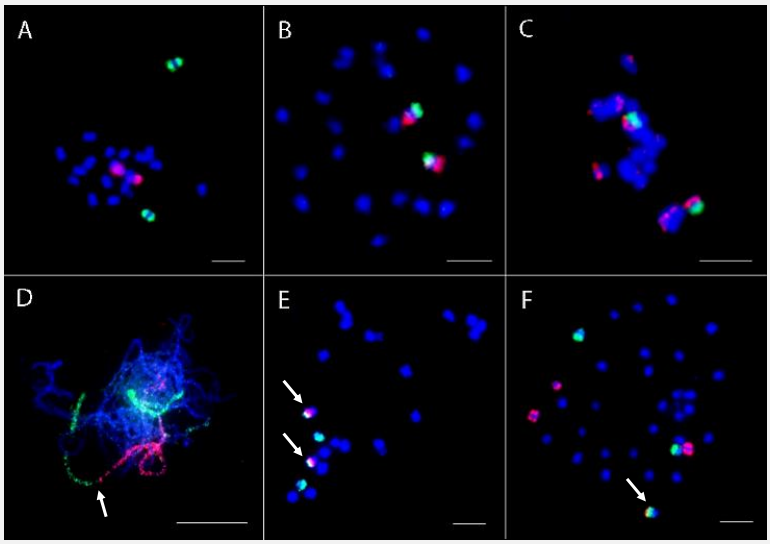

Example of oligo density specific to pseudomolecule 3 of *M. acuminata* 'DH Pahang' using Chorus pipeline. Oligos coverage is marked in blue. Areas with low oligo's coverage (pericentromeric regions) were eliminated from probe design.

2) Probe preparation

- Amplification and labeling of synthesized oligos:

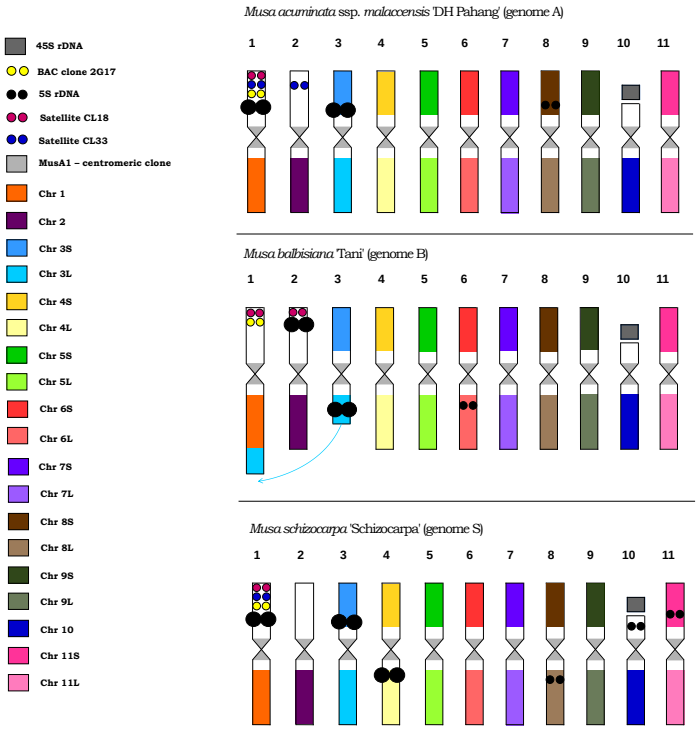



## Results




Examples of oligo painting on mitotic metaphase or meiotic pachytene chromosomes. (A) *M. acuminata* ssp. *malaccensis* 'DH Pahang' (2n = 22, AA; chromosome 1 in red, chromosome 3 in green). (B) *M. acuminata* ssp. *malaccensis* 'DH Pahang' (2n = 22, AA; chromosome 1 in red, BAC clone 2G17 in green). (C) *Musa schizocarpa* 'Schizocarpa' (2n = 2x = 22, SS; short arm of chromosome 4 in green, 5S rDNA loci in red – two of them were localized on long arm of chromosome 4). (D) *M. balbisiana* 'Tani' (2n = 22, BB; chromosome 1 in red, chromosome 3 in green). (E) *M. balbisiana* 'Pisang Klutuk Wulung' (2n = 22, BB; chromosome 1 in red, chromosome 3 in green). (F) Plantain '3 Hands Planty' (2n = 3x = 33, AAB; chromosome 3 in red, chromosome 1 in green). Chromosomes were counterstained with DAPI (blue). Arrows point to a translocated region of chromosome 3 to chromosome 1 in *M. balbisiana* as well as to a chromosome 1 from B genome in plantain '3 Hands Planty'. Bars = 5 µm (resp. 10 µm in picture C).

## Idiograms of three diploid *Eumusa* accessions




## Conclusions

- For the first time, molecular karyotype of diploid *M. acuminata*, *M. balbisiana*, *M. schizocarpa* and their triploid hybrid clones from *Eumusa* section was revealed, specific DNA sequences were successfully mapped on individual chromosomes using chromosome-specific oligo painting
- Oligos specific to *M. acuminata* 'DH Pahang' can be utilized for studying the evolution of karyotypes in whole *Musaceae* family and/or identification of chromosome translocations
- Density of oligo painting probes is sufficient to study chromosomal rearrangements on mitotic as well as on meiotic pachytene chromosomes

This work has been supported by the Grant Agency of the Czech Republic (award No. 19-20303S). The computing was supported by the National Grid Infrastructure MetaCentrum (grant No. LM2010005 under the program Projects of Large Infrastructure for Research, Development, and Innovations).

**Palacký University Olomouc**

**Faculty of Science**

**Department of Botany**

**and**

**Centre of Plant Structural and Functional Genomics**

**Institute of Experimental Botany AS CR**

**Centre of the Region Haná for Biotechnological and Agricultural Research**

**Olomouc**

**Denisa Šimoníková**

# Karyotype evolution in bananas (*Musa* spp.)

**P1527 Biology - Botany**

**Summary of Ph.D. Thesis**

Olomouc

2021

Ph.D. thesis was carried out at the Department of Botany, Faculty of Science, Palacký University Olomouc, in years 2017–2021.

Candidate: **Mgr. Denisa Šimoníková**

Supervisor: **Mgr. Eva Hřibová, Ph.D.**

Reviewers: **Prof. Dr. Andreas Houben**
Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany

**Prof. Mgr. Martin Lysák, Ph.D., DSc.**
Masaryk University, Brno, Czech Republic

**RNDr. Jiří Macas, Ph.D.**
Institute of Plant Molecular Biology, Biology Centre CAS, České Budějovice, Czech Republic

The evaluation of the Ph.D. thesis was written by…….....
……...

The summary of the Ph.D. thesis was sent for distribution on…….....

The oral defence will take place on……in front of the commission for the Ph.D. study of the program Botany in……..
……...

The Ph.D. thesis is available in the Library of the biological departments of the Faculty of Science of Palacký University Olomouc, Šlechtitelů 11, Olomouc-Holice.

**Prof. Ing. Aleš Lebeda, DrSc.**
Chairman of the Commission for the Ph.D.
Thesis of the study program Botany
Department of Botany, Faculty of Science
Palacký University Olomouc

# Content

# 1 Introduction

Bananas and plantains are one of the major staple food crops for millions of people living in tropical and subtropical developing countries and represent important world trade commodity. In 2019, about 159 million tons of bananas and plantains were produced worldwide (FAOSTAT, 2019). However, majority of the bananas is consumed in a place of their origin, only about 10-20 % is exported. In 2019, around 26 million tons of bananas were distributed around the world (International trade statistics, 2019). Two types of bananas are available at the market, sweet and cooking bananas. While sweet bananas serve as a food supplement, cooking bananas are starchier and have to be processed before consuming.

Edible banana cultivars, which belong to Eumusa section of genus *Musa*, originated after natural inter- or intra-specific cross hybridization between two wild species *Musa acuminata* Colla (2n=2x=22, AA genome) and *Musa balbisiana* Colla (2n=2x=22, BB genome). Cultivated bananas differ in ploidy level and genomic constitution. Crosses resulted mostly in seed sterile diploid (AA, AB), triploid (AAA, AAB, ABB) or even tetraploid hybrids (AAAB, AABB), obtained in the breeding programs (Simmonds and Shepherd, 1955; Simmonds, 1956). Moreover, molecular studies revealed the contribution of *Musa schizocarpa* (Eumusa section, 2n=2x=22, SS genome), which is very close to *M. acuminata*, and diploid *Musa textilis* (Australimusa section, 2n=2x=20, TT genome) to the origin of edible banana clones after crosses with *M. acuminata* (Carrell *et al*., 1994; Čížková *et al*., 2013; Němečková *et al.*, 2018). The most common edible cultivars are triploid bananas, that emerged after the fusion of unreduced gamete from edible diploid, almost sterile cultivar with haploid gamete from a fertile diploid (Simmonds, 1962; Carreel *et al.*, 1994; Raboin *et al*., 2005).

Based on geographical distribution, nine subspecies (namely *banksii, burmannica, burmannicoides, errans, malaccensis, microcarpa, siamea, truncata*, and *zebrina*) and three varieties (namely chinensis, sumatrana, and tomentosa) have been recognized within *M. acuminata* (Simmonds, 1962; Perrier *et al*., 2009; 2011; Martin *et al*., 2017; WCSP, 2018). Due to climatic changes during glaciation periods, different *M. acuminata* subspecies got in close proximity, resulting in cross hybridization and the origin of intra-specific hybrids, which were further selected and propagated by humans (Perrier *et al*., 2011; Martin *et al*., 2017). Clarifying the evolution of individual subspecies of *M. acuminata*, which played important role as genetic contributors to modern cultivars, is crucial for understanding the domestication and diversification of banana.

Several cytogenetic studies on chromosome pairing during meiosis in *Musa* (Wilson, 1945; Simmonds, 1962; Dessauw, 1987; Fauré *et al*., 1993; Shepherd, 1999; Therdsak *et al*., 2010) indicated that evolution of edible cultivated banana clones and some wild species was accompanied by chromosomal rearrangements, particularly translocations, which lead to differences in chromosome structure (Dodds, 1943; Fauré *et al*., 1993; Shepherd, 1999). Structural chromosome heterozygosity causes irregularities in meiosis (as irregular chromosomal pairing) and subsequent development of low fertility or complete sterility in hybrids and aberrant chromosome numbers in the progeny (Jáuregui *et al*., 2001; Rieseberg, 2001; Ostberg *et al*., 2013).

Based on chromosome pairing during meiosis in inter-subspecific hybrids of *M. acuminata*, seven (to eight) translocation groups of structurally homogenous accessions have been distinguished among *M. acuminata* subspecies by Shepherd (1999). One Standard group (ST) and six groups differing from Standard group by 1 to 4 translocations only partially corresponded to the classification of individual subspecies. These findings were supported by several authors, who observed segregation distortion during genetic mapping in inter-subspecific hybrids (Fauré *et al*., 1993; Hippolyte *et al*., 2010; Mbanjo *et al*., 2012; Noumbissié *et al*., 2016).

Despite huge socio-economic significance of bananas, a little is known about their genome structure and organization at chromosomal level, which is important for further genetic improvement of edible clones in breeding programs. Moreover, the evolution of whole Musaceae family is still unclear. However, the availability of banana reference genome sequence enabled the implementation of relatively new cytogenetic method called oligo painting FISH in banana studies and the identification of all chromosomes within a karyotype. The aim of this work is to shed light on genome organization and karyotype evolution in Musaceae family, especially in edible banana clones, using fluorescence *in situ* hybridization with whole chromosome oligo painting probes.

# 2 Aims of the thesis

## I. Development of oligonucleotide-based chromosome painting FISH in banana (*Musa* spp.)

The first aim of the thesis was to establish a set of chromosome/chromosome-arm specific oligo painting probes to identify all chromosomes in banana and to anchor pseudomolecules of reference genome sequence of *Musa acuminata* spp. *malaccensis* 'DH Pahang' to individual chromosomes *in situ*, and to create integrated karyotyping in *Musa* spp.

## II. Karyotype reconstruction and evolution in *Musa* spp.

The second aim of this work was to perform comparative karyotype analysis in a set of twenty edible banana clones and their wild relatives using FISH with chromosome/chromosome-arm specific oligo painting probes and shed a light on chromosome organization and structural chromosome changes accompanying the evolution of the genus *Musa*.

## III. Karyotype evolution within Musaceae

The third aim of the thesis was to reveal chromosome structure of phylogenetically distinct species covering whole Musaceae family, which differ in genome size and basic chromosome number, and thus elucidate the genome organization at chromosomal level during the evolution and origin of species of this family.

## IV. Establishment of oligo painting FISH in other plant species

The fourth aim of the thesis was to contribute to development and application of oligo painting FISH in fonio millet (*Digitaria exilis*) to analyse chromosome structure and integrity of newly developed whole genome sequence, and in the genera *Silene* spp. and *Lupinus* spp. to provide the information about sex chromosome evolution and to perform comparative cytogenetic mapping.

# 3 Materials and methods

### Plant material and preparation of chromosome spreads (*Musa* spp.)

*In vitro* rooted plants of *Musa* spp. accessions used in this work were obtained from the International *Musa* Transit Centre (ITC, Biodiversity International, Leuven, Belgium). *In vitro* plants were transferred to garden soil and maintained in a heated greenhouse. Male buds of *M. acuminata* 'Pahang' and *M. balbisiana* 'Tani' were obtained from the research station of the International Institute of Tropical Agriculture in Sendusu, Uganda.

Protoplast suspensions were obtained from actively growing *Musa* root tips, which were pre-treated in 0,05% (w/v) 8-hydroxyquinoline and fixed in 3:1 ethanol:acetic acid fixative. In the next step, dropping technique was used for the preparation of mitotic metaphase chromosome spreads according to Doležel *et al*. (1998). Preparation of pachytene chromosome spreads was done according to Mandáková and Lysák (2008), with minor modifications.

### Identification of specific oligomers and labeling of the oligo probes (*Musa* spp.)

Oligomers specific for individual chromosome arms were identified in the reference genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang' v.2 (Martin *et al*., 2016) using Chorus pipeline (Han *et al*., 2015). Sets of 20 000 oligomers (45-mers) per one library were synthesized by Arbor Biosciences (Ann Arbor, MI, USA) and labeled directly by CY5 fluorochrome, or by digoxigenin or biotin according to Han *et al*. (2015) and used for FISH.

### Preparation of other cytogenetic markers (*Musa* spp.)

Probes specific for ribosomal DNA sequences were prepared by labeling Radka1 (part of 26S rRNA gene) and Radka2 (contains 5S rRNA gene and non-transcribed spacer) DNA clones (Valárik *et. al*., 2002) with biotin-16-dUTP or aminoallyl-dUTP-CY5 by PCR using T3 and T7 primers. Probes for tandem repeats CL18 and CL33 (Hřibová *et al*., 2010) were amplified using specific primers and labeled with aminoallyl-dUTP-CY5 or fluorescein-12-dUTP by PCR according to Čížková *et al*. (2013). Single copy BAC clone 2G17 (Hřibová *et al*., 2008) was labeled by digoxigenin-11-dUTP nick translation.

**Fluorescence *in situ* hybridization and image analysis (*Musa* spp.)**

Hybridization mixture containing 50% (v/v) formamide, 10% (w/v) dextran sulfate in 2x SSC and 10 ng/µl of labeled probe was added onto a slide and denatured for 3 min at 80 °C. Hybridization was carried out overnight at 37 °C. The sites of hybridization of digoxigenin- and biotin-labeled probes were detected using anti-digoxigenin-FITC and streptavidin-Cy3, respectively. Chromosomes were counterstained with DAPI and mounted in Vectashield Antifade Mounting Medium. The slides were examined with Axio Imager Z.2 Zeiss microscope equipped with Cool Cube 1 camera and appropriate optical filters. The capture of fluorescence signals, merging the layers and measurement of chromosome length were performed with ISIS software 5.4.7, the final image adjustment and creation of idiograms were done in Adobe Photoshop CS5.

# 4 Summary of results

The first aim of this Ph.D. thesis was to develop a set of chromosome/chromosome-arm specific oligo painting probes in banana (*Musa* spp.). Unique oligomers (oligos) were identified according to Han *et al*. (2015) using a reference genome sequence of *Musa acuminata* ssp. *malaccensis* 'DH Pahang' (Martin *et al*., 2016) and selected with the Chorus program (https://github.com/forrestzhang/Chorus). Eight pseudomolecules corresponded to metacentric chromosomes, thus sets of 20 000 45-mers were designed for individual chromosome arms. However, pseudomolecules 1, 2 and 10 seemed to be acrocentric, in which sets of 20 000 oligomers covering only their long arms could be identified. Due to a lower oligomer density, peri-centromeric regions of all pseudomolecules were excluded from oligomer selection and probe preparation. The chromosome-arm specific oligomer libraries with a density from 0.9 to 2.1 oligos per 1 kb were synthesized as so-called immortal libraries by Arbor Biosciences (Ann Arbor, MI, USA), then fluorescently labeled and used as probes for FISH.

In order to verify their specificity, the newly developed chromosome/chromosome-arm specific oligo probes were hybridized to mitotic metaphase chromosomes of diploid *M. acuminata* ssp. *malaccensis* 'Pahang' (A genome), from which the reference genome sequence was developed. Oligo painting FISH resulted in visible signals covering individual chromosome arms. Subsequently, these painting probes were used for FISH also in closely related banana species, which played a role in the evolution of most banana edible cultivars, diploid *M. balbisiana* 'Tani' (B genome) and *M. schizocarpa* 'Schizocarpa' (S genome), to prove their suitability for comparative karyotype analysis. In *M. balbisiana*, a large translocation of the long arm of chromosome 3 to the long arm of chromosome 1 was detected. Painting probes were also hybridized onto the less condensed meiotic pachytene chromosomes, which provided valuable information about chromosome structure in more detail. Moreover, for better characterization of *Musa* chromosomes, previously developed cytogenetic landmarks, namely probes for 45S rDNA and 5S rDNA, two satellites CL18 and CL33 and a BAC clone 2G17, were integrated with the painting probes and molecular karyotypes of three diploid *Musa* species were created (Šimoníková *et al*., 2019).

In the second part of the work, karyotypes of twenty representatives of Eumusa section of genus *Musa*, comprising edible banana clones and their probable progenitors, have been studied using developed oligo painting probes. Large differences in chromosome structures, specific for individual species, subspecies and groups of edible accessions were identified

(Figure 1). Presence of specific types of large translocations in analysed *Musa* accessions led to the identification of most probable parents of edible cultivated clones. Moreover, structural chromosome heterozygosity confirmed hybrid origin of cultivated clones and some wild subspecies of *M. acuminata*, and pointed to a possible reason of their reduced fertility (Šimoníková *et al*., 2020). In the third part of the work, which is yet to be published, chromosome-specific oligo painting probes were successfully hybridized also to chromosomes of twelve distinct species covering whole Musaceae family, which differ in genome size and chromosome number (x=9, 10 or 11).

Combination of chromosome-arm specific painting probes enabled the identification of numerous large chromosomal rearrangements in studied accessions. Observed chromosome structures in more distinct wild species corresponded to their evolution and phylogenetic relationships. The oligo painting study of the analysed accessions, representing individual phylogenetic groups of family Musaceae, was enlarged by the application of long-read sequencing technology, Oxford Nanopore, for precise determination of chromosome-translocation breakpoints.

Finally, chromosome-specific oligo painting FISH method was developed in fonio millet (*Digitaria exilis*) and another two genera, *Silene* spp. and *Lupinus* spp. In tetraploid species fonio millet (2n=4x=36), this technique was used to assess the quality of its newly developed genome assembly. As a result, no cross-hybridization between individual chromosomes was observed, thus a set of eighteen newly assembled pseudomolecules representing individual chromosomes was validated. Moreover, homoeologous chromosomes were unambiguously distinguished in this species (Abrouk *et al*., 2020).

The formation of sex chromosomes in diocieous species from genus *Silene*, which have XX/XY sex determination system, has been also studied using oligo panting FISH. The results confirmed that sex chromosome evolution in studied dioecious *Silene* spp. was accompanied by multiple chromosomal rearrangements (Bačovský *et al*., 2020).

A set of four oligo painting probes covering both arms of chromosome 6 of the *Lupinus angustifolius* was designed and mapped on mitotic chromosome spreads together with previously identified single copy BAC clones specific to chromosome 6. Using these probes, comparative cytogenetic mapping among the smooth-seeded and rough-seeded lupin species was performed. Structural chromosome changes, which accompanied the evolution and speciation of lupin species, have been detected (Bielski *et al*., 2020).

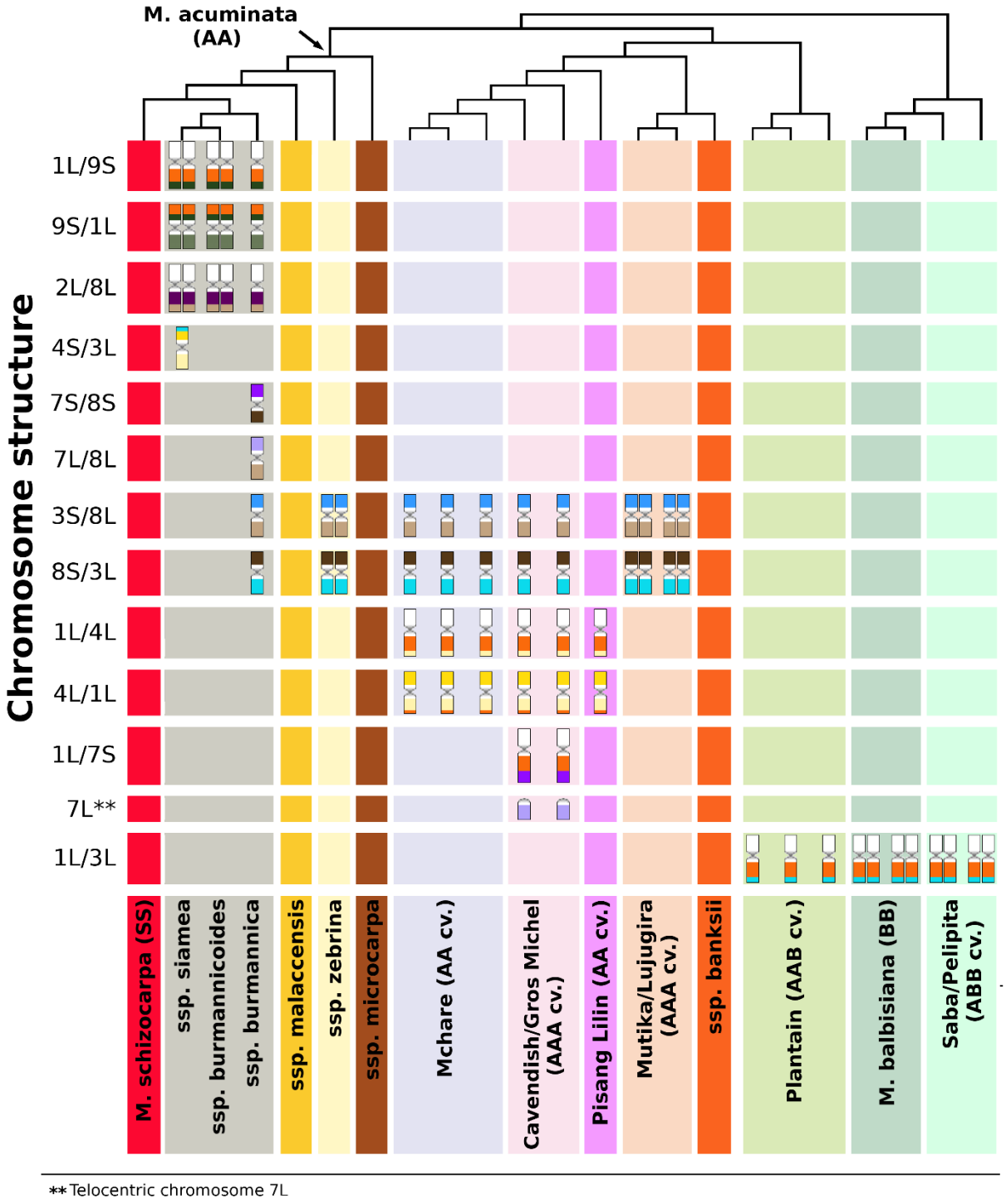



**Figure 1.** Overview of common translocation events revealed by oligo painting FISH in *Musa*: Diversity tree, constructed using SSR markers according to Christelová *et al*. (2017) shows the relationships among *Musa* species, subspecies and hybrid clones. Lineages of closely related accessions and groups of edible banana clones are highlighted in different colors. Individual chromosome structures (displayed as chromosome schemes) are depicted in rows, and their number correspond to the number of chromosomes bearing the rearrangement in the nuclear genome in somatic cell lines (2n; Šimoníková *et al*., 2020).

# 5 Summary

The Ph.D. thesis provides a comparative analysis of chromosome structure in selected banana species and edible banana cultivars from genus *Musa*, and their closely related genera *Musella* and *Ensete* using recently developed oligo painting FISH technique. Oligo painting FISH is based on the identification of short (45-nt long), chromosome-specific oligos from a reference genome sequence, which are synthetized in parallel as a pool, fluorescently labeled and used as probes for fluorescence *in situ* hybridization (FISH). Chromosome/chromosome-arm specific oligo painting probes developed using the reference genome sequence of *M. acuminata* ssp. *malaccensis* 'DH Pahang' (A genome) were used for unambiguous identification of individual chromosomes and anchoring of the pseudomolecules to chromosomes *in situ*. Further, the cross-hybridization of the oligo painting probes specific to *M. acuminata* ssp. *malaccensis* to closely related genomes of *M. balbisiana* (B genome) and *M. schizocarpa* (S genome) proved the use of the probes for comparative karyotyping in banana (*Musa* spp.).

In the second part of the work, available chromosome/chromosome-arm specific oligo painting probes were used for the identification of large structural chromosome changes (translocations) in a set of twenty economically important diploid and triploid edible banana cultivars from the genus *Musa* and in species and subspecies, which were probably involved in their origin. Large differences in chromosome structure between individual species and subspecies of *Musa* were detected. Specific translocations, which occurred during the evolution of *Musa*, enabled the identification of putative progenitors of edible banana clones. Observed structural chromosome heterozygosity supported the hybrid origin of selected accessions and pointed to possible reason for their low fertility. These findings will be a valuable asset for banana breeders in selecting parents for further crosses.

FISH technique with previously developed chromosome-arm specific oligo painting probes was also applied to study chromosome structure of phylogenetically distinct species to *M. acuminata* within whole Musaceae family, differing in genome size and basic chromosome number. This part of the work elucidated the genome organization at chromosomal level during the evolution and origin of species of this family. Finally, oligo painting probes were designed for the analysis of chromosome structures in other plant species, e. g. in fonio millet (*Digitaria exilis*), and genera *Lupinus* spp. and *Silene* spp.

# 6 References


**Abrouk, M., Ahmed, H.I., Cubry, P., Šimoníková, D., Cauet, S., Bettgenhaeuser, J., *et al.*** (2020). Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate. *Nat. commun*. **11**:4488. doi: 10.1038/s41467-020-18329-4

**Bačovský, V., Čegan, R., Šimoníková, D., Hřibová, E. and Hobza, R.** (2020). The formation of sex chromosomes in *Silene latifolia* and *S. dioica* was accompanied by multiple chromosomal rearrangements. *Front. Plant Sci*. **11**:205. doi: 10.3389/fpls.2020.00205

**Bielski, W., Książkiwicz, M., Šimoníková, D., Hřibová, E., Susek, K. and Naganowska, B.** (2020). The puzzling fate of a lupin chromosome revealed by reciprocal oligo-FISH and BAC-FISH mapping. *Genes*. **11**:e1489. doi: 10.3390/genes11121489

**Carreel, F., Fauré, S., González de León, D., Lagoda, P.J.L., Perrier, X., Bakry, F., *et al.*** (1994). Evaluation of the genetic diversity in diploid bananas (*Musa* sp.*). Genet. Sel. Evol*. **26**:125-136. doi: 10.1051/gse:19940709

**Čížková, J., Hřibová, E., Humplíková, L., Christelová, P., Suchánková, P. and Doležel J.** (2013). Molecular analysis and genomic organization of major DNA satellites in banana (*Musa* spp.). *PLoS One*. **8**:e54808. doi: 10.1371/journal.pone.0054808

**Dessauw, D.** (1987). Etude des facteurs de la ste´rilite´ du bananier (*Musa* spp) et des relations cytotaxinomiques entre *M. acuminata* Colla et *M. balbisiana* Colla. PhD thesis, University of Paris-Sud Centre d'Orsay.

**Dodds, K.S.** (1943): Genetical and cytological studies of *Musa*. V. Certain edible diploids. *J. Genet*. **45**:113-138. doi: 10.1007/BF02982931

**Doležel, J., Doleželová, M., Roux, N., and Van den houwe, I.** (1998). A novel method to prepare slides for high resolution chromosome studies in *Musa* spp. *Infomusa*. **7**:3-4.

**FAOSTAT,** Agriculture Organization of the United Nations, FAO. 2019. http://www.fao.org/home/en/. Accessed June 2021.

**Fauré, S., Noyer, J.L., Horry, J.P., Bakry, F., Lanaud, C. and de León, D.G.** (1993). A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). *Theor. Appl. Genet.* **87**:517-526. doi: 10.1007/BF00215098

**Han, Y., Zhang, T., Thammapichai, P., Weng, Y. and Jiang, J.** (2015). Chromosome-specific painting in *Cucumis* species using bulked oligonucleotides. *Genetics*. **200**:771-779. doi: 10.1534/genetics.115.177642

**Hippolyte, I., Bakry, F., Seguin, M., Gardes, L., Rivallan, R., Risterucci, A.M., *et al.*** (2010). A saturated SSR/DArT linkage map of *Musa acuminata* addressing genome rearrangements among bananas. *BMC Plant Biol.* **10**:65. doi: 10.1186/1471-2229-10-65

**Hřibová, E., Doleželová, M. and Doležel, J.** (2008). Localization of BAC clones on mitotic chromosomes of *Musa acuminata* using fluorescence *in situ* hybridization. *Biol. Plant*. **52**:445-452. doi: 10.1007/s10535-008-0089-1

**Hřibová, E., Neumann, P., Matsumoto, T., Roux, N., Macas, J. and Doležel, J**. (2010). Repetitive part of the banana (*Musa acuminata*) genome investigated by low-depth 454 sequencing. *BMC Plant Biol*. **10**:204. doi: 10.1186/1471-2229-10-204

**https://github.com/forrestzhang/Chorus**

**Christelová, P., De Langhe, E., Hřibová, E., Čížková, J., Sardos, J., Hušáková, M., *et al.*** (2017). Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. *Biodivers. Conserv.* **26**:801-824. doi: 10.1007/s10531-016-1273-9

**International trade statistics**. 2019. https://www.wto.org/ Accessed May 2020.

**Jáuregui, B., de Vicente, M.C., Messeguer, R., Felipe, A., Bonnet, A., Salesses, G., *et al.*** (2001). A reciprocal translocation between 'Garfi' almond and 'Nemared' peach. *Theor Appl Genet*. **102**:1169-1176. doi: 10.1007/s001220000511

**Mandáková, T. and Lysák, M. A.** (2008). Chromosomal phylogeny and karyotype evolution in x=7 crucifer species (Brassicaceae). *Plant Cell*. **20**:2559-2570. doi: 10.1105/tpc.108.062166

**Martin, G., Baurens, F.C., Droc, G., Rouard, M., Cenci, A., Kilian, A., *et al.*** (2016). Improvement of the banana "*Musa acuminata*" reference sequence using NGS data and semi-automated bioinformatics methods. *BMC Genomics*. **17**:1-12. doi: 10.1186/s12864-016-2579-4

**Martin, G., Carreel, F., Coriton, O., Hervouet, C., Cardi, C., Derouault, P., *et al.*** (2017). Evolution of the banana genome (*Musa acuminata*) is impacted by large chromosomal translocations. *Mol. Biol. Evol.* **34**:2140-2152. doi: 10.1093/molbev/msx164

**Mbanjo, E.G.N, Tchoumbougnang, F., Mouelle, A.S., Oben, J.E., Nyine, M., Dochez, C., *et al.*** (2012). Molecular marker-based genetic linkage map of a diploid banana population (*Musa acuminata* Colla). *Euphytica.* **188**:369-386. doi: 10.1007/s10681-012-0693-1

**Němečková, A., Christelová, P., Čížková, J., Nyine, M., Van den houwe, I., Svačina, R., *et al.*** (2018). Molecular and cytogenetic study of East African Highland Banana. *Front. Plant. Sci*. **9**:1371. doi: 10.3389/fpls.2018.01371

**Noumbissié, G.B., Chabannes, M., Bakry, F., Ricci, S., Cardi, C., Njembele, J.C., *et al.*** (2016). Chromosome segregation in an allotetraploid banana hybrid (AAAB) suggests a translocation between the A and B genomes and results in eBSV-free offsprings. *Mol. Breed.* **36**:38-52. doi: 10.1007/s11032-016-0459-x

**Ostberg, C.O., Hauser, L., Pritchard, V.L., Garza, J.C. and Naish, K.A.** (2013). Chromosome rearrangements, recombination suppression, and limited segregation distortion in hybrids between Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) and rainbow trout (*O. mykiss*). *BMC Genomics*. **14**:570. doi: 10.1186/1471-2164-14-570

**Perrier, X., Bakry, F., Carreel, F., Jenny, F., Horry, J.P., Lebot, V., *et al.*** (2009). Combining biological approaches to shed light on the evolution of edible bananas. *Ethnobot. Res. Appl.* **7**:199-216. doi: 10.17348/era.7.0.199-216

**Perrier, X., De Langhe, E., Donohue, M., Lentfer, C., Vrydaghs, L., Bakry, F., *et al.*** (2011). Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc. Natl. Acad. Sci. USA.* **108**:11311-11318. doi: 10.1073/pnas.1102001108

**Raboin, L.M., Carreel, F., Noyer, J.L., Baurens, F.C., Horry, J.P., Bakry, F., *et al.*** (2005). Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete donors and n gamete donors. *Mol. Breed*. **16**:333-341. doi: 10.1007/s11032-005-2452-7

**Rieseberg, L.H.** (2001). Chromosomal rearrangements and speciation. *Trends Ecol. Evol*. **16**:351-358. doi: 10.1016/S0169-5347(01)02187-5

**Shepherd, K.** (1999). Cytogenetics of the genus *Musa*. International Network for the Improvement of Banana and Plantain, Montpellier, France, ISBN: 2-910810-25-9.

**Simmonds, N.W.** (1956). Botanical results of the banana collecting expeditions, 1954-5. Kew Bull. **11**:463-489. doi: 10.2307/4109131

**Simmonds, N.W.** (1962). The evolution of the bananas. Tropical Science Series. Longmans, London (GBR), 170.

**Simmonds, N.W. and Shepherd, K.** (1955). The taxonomy and origins of the cultivated bananas. *J. Linn. Soc., Bot*. **55**:302-312.

**Šimoníková, D., Němečková, A., Čížková, J., Brown, A., Swennen, R., Doležel, J., *et al.*** (2020). Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure. *Int. J. Mol. Sci*. **21**:7915. doi: 10.3390/ijms21217915

**Šimoníková, D., Němečková, A., Karafiátová, M., Uwimana, B., Swennen, R., Doležel, J., *et al.*** (2019). Chromosome painting facilitates anchoring reference genome sequence to chromosomes *in situ* and integrated karyotyping in banana (*Musa* spp.) *Front. Plant Sci*. **10**:1503. doi: 10.3389/fpls.2019.01503

**Therdsak, T., Benchamas, S., Yingyong, P. and Pradit, P.** (2010). Meiotic behavior in microsporocytes of some bananas in Thailand. Kasetsart Journal (Natural Science) **44**:536-543.

**Valárik, M., Šimková, H., Hřibová, E., Šafář, J., Doleželová, M. and Doležel, J.** (2002). Isolation, characterization and chromosome localization of repetitive DNA sequences in bananas (*Musa* spp.). *Chromosome Res*. **10**:89-100. doi: 10.1023/A:1014945730035

**WCSP**, World Checklist of Selected Plant Families. Facilitated by the Royal Botanic Gardens, Kew. 2018. http://wcsp.science.kew.org/. Accessed January 2020.

**Wilson, G.B.** (1945). Cytological studies in the *Musae*; meiosis in some triploid clones. *Genetics*. **31**:241-258.

# 7 List of author's publications

## 7.1 Original papers

Šimoníková, D., Němečková, A., Karafiátová, M., Uwimana, B., Swennen, R., Doležel, J., Hřibová, E. (2019). Chromosome painting facilitates anchoring reference genome sequence to chromosomes *in situ* and integrated karyotyping in banana (*Musa* spp.) *Front. Plant Sci*. **10**:1503. doi: 10.3389/fpls.2019.01503

Šimoníková, D., Němečková, A., Čížková, J., Brown, A., Swennen, R., Doležel, J., Hřibová, E. (2020). Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure. *Int. J. Mol. Sci*. **21**:7915. doi: 10.3390/ijms21217915

Abrouk, M., Ahmed, H.I., Cubry, P., Šimoníková, D., Cauet, S., Bettgenhaeuser, J., Gapa, L., Pailles, Y., Scarcelli, N., Couderc, M., Zekraoui, L., Kathiresan, N., Čížková, J., Hřibová, E., Doležel, J., Arribat, S., Bergès, H., Wieringa, J.J., Gueye, M., Kane, N.A., Leclerc, C., Causse, S., Vancoppenolle, S., Billot, C., Wicker, T., Vigouroux, Y., Barnaud, A., Krattinger, S.G. (2020). Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate. *Nat. commun*. **11**:4488. doi: 10.1038/s41467-020-18329-4

Bačovský, V., Čegan, R., Šimoníková, D., Hřibová, E., Hobza, R. (2020). The formation of sex chromosomes in *Silene latifolia* and *S. dioica* was accompanied by multiple chromosomal rearrangements. *Front. Plant Sci*. **11**:205. doi: 10.3389/fpls.2020.00205

Bielski, W., Książkiwicz, M., Šimoníková, D., Hřibová, E., Susek, K., Naganowska, B. (2020). The puzzling fate of a lupin chromosome revealed by reciprocal oligo-FISH and BAC-FISH mapping. *Genes*. **11**:e1489. doi: 10.3390/genes11121489

## 7.2. Published abstracts

Šimoníková, D., Čížková, J., Karafiátová, M., Němečková, A., Doležel, J., Hřibová, E.: Chromosome oligo painting facilitates the analysis of karyotype evolution in Musaceae. In: Abstracts of the "22nd International Chromosome Conference". Prague, Czech Republic, 2018. (poster presentation)

Šimoníková, D., Doležel, J., Hřibová, E.: Chromosome-specific oligo painting elucidates large variation in Eumusa genome. In: Abstracts of the "International Conference on Polyploidy". Ghent, Belgium, 2019. (poster presentation)

Šimoníková, D., Hřibová, E.: Studium struktury chromozómů banánovníku (*Musa* spp.). In: Abstracts of the "53. výroční cytogenetická konference". Olomouc, Czech Republic, 2020. (oral presentation)

# 8 Souhrn (Summary in Czech)

**Název práce:** Evoluce karyotypu banánovníku (*Musa* spp.)

Cílem této disertační práce bylo objasnění struktury chromozomů u vybraných druhů a jedlých kultivarů banánovníku z rodu *Musa*, a také jejich blízce příbuzných rodů *Musella* a *Ensete* za pomoci nedávno vyvinuté metodiky pro malování chromozomů ("oligo painting FISH"). "Oligo painting FISH" využívá celogenomovou sekvenci pro identifikaci krátkých (45 bází), chromozomově specifických oligomerů, které jsou poté nasyntetizovány, fluorescenčně naznačeny a použity jako sondy pro fluorescenční *in situ* hybridizaci (FISH). Poprvé tak bylo pomocí sond specifických pro jednotlivé chromozomy nebo jejich ramena identifikováno všech 11 chromozomů, ke kterým byly ukotveny DNA pseudomolekuly referenční genomové sekvence druhu *M. acuminata* ssp. *malaccensis* 'DH Pahang' (A genom). Oligo malovací sondy navržené pro *M. acuminata* ssp. *malaccensis* byly úspěšně hybridizovány i na blízce příbuzné druhy *M. balbisiana* (B genom) a *M. schizocarpa* (S genom), což poukázalo na jejich možné využití pro analýzu struktury chromozomů i u dalších zástupců banánovníku (*Musa* spp.).

Druhá část práce byla zaměřena na využití těchto oligo malovacích sond pro identifikaci velkých chromozomálních přestaveb (translokací) u dvaceti zástupců ekonomicky významných diploidních a triploidních jedlých typů banánovníku z rodu *Musa* a druhů, resp. poddruhů, které se pravděpodobně podílely na jejich vzniku. Mezi jednotlivými druhy, resp. poddruhy i odrůdami byly zjištěny velké rozdíly v genomové struktuře na úrovni chromozomů. Specifické chromozomové translokace, ke kterým došlo v průběhu evoluce, umožnily identifikaci pravděpodobných předků jedlých typů banánovníku. U vybraných testovaných položek byl potvrzen jejich hybridní charakter, což poukazuje na možnou příčinu jejich snížené fertility. Tyto poznatky o detailní struktuře jednotlivých chromozomů jsou cennou informací pro šlechtitele při výběru klonů vhodných pro další křížení.

Sondy specifické pro jednotlivé chromozomy nebo chromozomální ramena byly následně využity pro studium struktury chromozomů i u vybraných zástupců reprezentujících čeleď banánovníkovité (Musaceae), lišících se kromě velikosti genomu i základním chromozomovým číslem. Cílem bylo objasnění organizace genomu na úrovni chromozomů v průběhu evoluce a vzniku druhů v čeledi Musaceae. V neposlední řadě byly oligo malovací sondy navrženy pro studium struktury chromozomů i u dalších rostlinných druhů, např. rosičky útlé (*Digitaria exilis*), a rodů lupina (*Lupinus* spp.) a silenka (*Silene* spp.).