Painting the diversity of a world's favorite fruit: A next generation catalog of cultivated bananas

Julie Sardos, Alberto Cenci, Guillaume Martin, Catherine Breton, Valentin Guignon, Ines van den Houwe, Yaleidis Mendez, Gabriel L Sachter-Smith, Rachel Chase, Max Ruas, et al.

Received: 31 May 2024 | Revised: 29 August 2024 | Accepted: 4 September 2024

DOI: 10.1002/ppp3.10581

Plants People Planet PPP Open Access

RESEARCH ARTICLE

Painting the diversity of a world's favorite fruit: A next generation catalog of cultivated bananas

Julie Sardos[1,2] | Alberto Cenci[1,2] | Guillaume Martin[2,3] | Catherine Breton[1,2] | Valentin Guignon[1,2] | Ines Van den Houwe[4] | Yaleidis Mendez[5] | Gabriel L. Sachter-Smith[1,6] | Rachel Chase[1] | Max Ruas[1,2] | Ronan Rivallan[2,3] | Janet Paofa[7] | William Wigmore[8] | David Tilafono Hunter[9] | Angélique D'Hont[2,3] | Nabila Yahiaoui[2,3] | Christophe Jenny[2,3] | Xavier Perrier[2,3] | Nicolas Roux[1,2] | Mathieu Rouard[1,2]

[1]Bioversity International, Montpellier, France

[2]UMR AGAP Institute, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France

[3]CIRAD, UMR AGAP Institute, Montpellier, France

[4]Bioversity International, Leuven, Belgium

[5]USDA–ARS Tropical Agriculture Research Station, Mayaguez, Puerto Rico

[6]Hawaii Banana Source, Waialua, Hawaii, USA

[7]NARI, Southern Regional Centre, Central Province, Port Moresby, Papua New Guinea

[8]Ministry of Agriculture, Rarotonga, Cook Islands

[9]Ministry of Agriculture and Fisheries, Apia, Samoa

Correspondence
Julie Sardos and Mathieu Rouard, Bioversity International, Parc Scientifique Agropolis II, 34397 Montpellier, France.
Email: j.sardos@cgiar.org, m.rouard@cgiar.org

Funding information
CGIAR

Societal Impact Statement

Bananas are nutritious fruits of major importance worldwide. Characterizing their diversity is essential to ensure their conservation and use. A catalog showcasing cultivated bananas genomic diversity was compiled and is to be used as a tool to support the classification of banana cultivars. This research revealed that cultivated banana groups are not all made of identical clones. Materials from recent collecting missions indicated that more banana diversity is expected to be found as the exploration of the banana gene pool continues. These discoveries will drive dynamic conservation strategies for banana genetic resources and should increase their use.

Summary

- Banana is an important food crop cultivated in many tropical and subtropical regions around the world. Because banana cultivars often have low fertility, they are typically propagated clonally, which maintains desirable traits across generations. However, different factors, such as synonymy, incomplete passport data, and environmental effects, complicate the morphological-based assignment of banana cultivars to specific clones or cultivar groups.
- In this study, we applied a previously developed genomic-based tool for fine-scale characterization of banana ancestry, known as in silico chromosome painting, to high-throughput genotyping data from 317 banana accessions. This dataset covers most of the globally conserved, studied, and cultivated cultivar groups and includes both genebanks and new, uncharacterized materials.
- By comparing curated morphological assignation to the genomic patterns resulting from in silico chromosome painting, we compiled a diversity catalog referencing

Julie Sardos and Alberto Cenci contributed equally to the study.

curated passport data, pictures, and chromosome painting patterns of the cultivar groups.

- Examining the genomic patterns obtained, intra-cultivar group variability was discovered. In some cultivar groups, mitotic recombination or deletions accumulated clonally. In addition, at least four cultivar groups encompassed cultivars from distinct sexual events co-existing, notably Pisang Awak with five distinct patterns across two ploidy levels. Finally, additional patterns were discovered in the newest materials of the set, showing that a wider diversity of clones still exists *on farm*.

KEYWORDS
banana, catalogue, clonal diversification, diversity, genomic characterisation

1 | INTRODUCTION

Crop diversity is critical for maintaining the resilience and adaptability of food systems in the face of changing environmental conditions and pests and diseases (McCouch et al., 2020; Smale & Jamora, 2020). Characterizing this diversity is therefore a much-needed effort to reach a comprehensive overview of the genetic diversity existing within a crop species and to ensure that effective conservation strategies are put in place. This knowledge serves as an essential baseline to monitor the evolution of diversity in situ and to identify new material to be conserved ex situ. Initially, crop characterization consisted of morphological assessments, but it later included molecular descriptions using molecular markers such as random amplified polymorphic DNA (RAPD), restriction fragment length polymorphism (RFLP), simple sequence repeat (SSR), and diversity arrays technology (DArT), and single nucleotide polymorphisms (SNPs) (Agarwal et al., 2008; Kilian et al., 2012; Powell et al., 1996). With the recent progress made in the fields involving genomics, an unprecedented level of fine-scale characterization can be reached (McCouch et al., 2020).

Bananas are an important crop for many tropical and subtropical regions around the world. They are a staple food for millions of people and a major source of nutrients, income, and employment for many communities. Currently, 80% of global banana production is limited to a few groups of cultivars, that is, traditional plants that have been cultivated over time through natural and human selection for specific characteristics, such as the dessert Cavendish and the cooking Plantain, but a much wider diversity exists, especially, but not only, in smallholder farms of South Asia and West Oceania, the center of origin of the crop (Simmonds, 1962). Seedlessness, the main trait associated with banana domestication, results from the progressive and joint selection of parthenocarpy and female sterility (Denham et al., 2020). As a result, current-day cultivars are nearly completely sterile, and this trait is maintained through clonal propagation. Domestication started in New Guinea island (Martin, Cottin, et al., 2023) but was a long process that involved the exchange of pre-domesticated forms in a wide South Asia—West Oceania area and the later diffusion of clonal cultivars to Africa, East Oceania, and possibly South America (De Langhe et al., 2015; Perrier et al., 2011). The clonal nature of the crop and the nearly complete female sterility of the cultivars resulted in reduced diversification. Some authors estimated that bananas encompass between 200 and 500 (Stover & Simmonds, 1987) to more than 1000 cultivars (Li & Ge, 2017). These cultivars have been widely circulated across tropical and sub-tropical regions, giving rise to complex distribution patterns combining local-specific diversity and extensive synonymy and homonymy.

This diversity, found in smallholder fields, is conserved through ex-situ national, regional, and international genebanks, which serve as repositories for cultivars, breeding material, and wild relatives and aim at safeguarding as much diversity as possible for present and future generations (Van den Houwe et al., 2020). Continuous efforts are necessary to properly characterize and rationalize the conserved germplasm and to identify gaps in collections. The classification of banana cultivars is complex, as cultivars are diploid or polyploid hybrids originating from crosses between different wild gene pools. To help in the assignation of cultivars, a scoring method was initially developed by the botanists and plant breeders, Norman Simmonds and Kenneth Shepherd (Simmonds & Shepherd, 1955), on the assumption that most banana cultivars were derived from the diploid wild species *Musa acuminata* Colla and *Musa balbisiana* Colla. By considering the different ploidy levels existing in the crop (diploid, triploid, tetraploid) and the scored relative contribution of both wild ancestors (coded A and B, respectively), Simmonds and Shepherd (1955) introduced the concept of genome constitution groups (e.g., AA, AB, AAA, AAB, ABB). Their work also laid the foundations for the definition of an additional taxonomical level, the subgroups, aiming at refining the classification. These subgroups correspond to cultivars considered to be somatic mutants fixed through vegetative propagation of a single clone or possibly siblings of related parents (Simmonds, 1966; Stover & Simmonds, 1987).

After more extensive exploration of banana-growing regions and with the beginning of molecular characterization of wild and cultivated germplasm, contributions of different subspecies of *M. acuminata* were shown (Carreel et al., 2002; Perrier et al., 2011). In addition, other wild ancestors were identified. Notably, it was assessed that some banana cultivars are also hybrids with *Musa schizocarpa* (S genome) or involve undefined *Musa* species of the former *Australimusa* section (T genome) (Carreel et al., 1994; Jarret Robert et al., 1992; Shepherd &

Ferreira, 1984). Using this classification, a catalog of the global banana diversity, the Musalogue (Daniells et al., 2001), documented 14 genome groups of three ploidy levels, subdivided in 37 subgroups across two botanical sections (Daniells et al., 2001; Häkkinen, 2013). All subgroups, preferentially called cultivar groups in our study, such as Cavendish, Plantain, Sucrier, or Pisang Awak, belong to the *Musa* section.

Despite being visionary for its time and even with the wide use of published standard descriptors for bananas (IPGRI, 1996; Taxonomy Advisory Group [TAG], 2016), classifying banana cultivars based on morphology alone is challenging. First, it requires relatively controlled growing conditions because both scores and descriptors were developed in and for ex-situ collections and do not consider variations that can be due to environmental conditions (large sense). Second, the identification requires observations at different stages of the plant's development, including fructification, requiring space and time as well as suitable climatic growing conditions. Moreover, the assignation of cultivars to cultivar groups depends upon trained eyes and the experience of the observers. Consequently, cultivars are regularly inconsistently classified, with a risk of negative impact on their conservation (Vogel Ely et al., 2017). The molecular markers developed later, such as RFLP (Carreel et al., 2002), SSR (Christelová et al., 2017; Hippolyte et al., 2012), or DArT (Risterucci et al., 2009; Sardos et al., 2016), enabled to characterize the genetic bases of the cultivar groups and helped in the assignation process. Relationships between some of the *M. acuminata* subspecies and cultivars were demonstrated, with *M. a. ssp. banksii*, *M. a. ssp. zebrina*, and *M. a. ssp. malaccensis* as main contributors to the A genomes of cultivars (Carreel et al., 1994, 2002; Perrier et al., 2011). However, the use of these technologies for such complex crops does not always allow the unambiguous assignation of accessions due to limited data point density to unravel their extensive hybridization history, further complicated by multiple levels of ploidy.

Recent advances in genomics revealed that cultivar genomes have mosaic patterns of ancestry along their chromosomes resulting from contributions of different *M. acuminata* subspecies, introgressions from *M. schizocarpa* and the presence, in many interspecific groups, of homoeologous chromosome exchanges between the genomes of *M. acuminata* and *M. balbisiana* (Baurens et al., 2019; Cenci et al., 2021; Higgins et al., 2023). Genome ancestry determination and the in silico painting of ancestry along chromosomes not only allowed the allocation of specific genomic patterns to cultivars of different ploidy levels but also revealed unidentified ancestors and the underestimated contribution of *M. schizocarpa* to cultivated bananas (Martin et al., 2020; Martin, Cottin, et al., 2023).

In this study, we applied the genome ancestry mosaic painting method to a large and well-curated panel of accessions, using extensive passport and morphological data from the Bioversity International *Musa* Transit Center (ITC) and recent collecting missions. Then, by characterizing the mosaic patterns observed within cultivar groups and for isolated accessions, we developed a catalog of genomic diversity aiming to be dynamic as banana germplasm characterization expands. This catalog is first to be used for conservation as a baseline of reference for further cultivar group assignation for existing and new genebank materials, as well as for *on-farm* projects. Furthermore, we discussed the implications for farming system, taxonomy, and breeding, where resolving banana cultivar classification could offer valuable new perspectives.

2 | MATERIALS AND METHODS

2.1 | Plant material

A collection of 317 banana accessions was selected for this study, as detailed in Table S1. The sources of these materials were diverse: 264 samples were obtained from the Bioversity International *Musa* Transit Centre (ITC) under the form of lyophilized leaves; 44 samples were collected during collecting missions to Indonesia, Papua New Guinea, Cook Islands, and Samoa (Hermanto et al., 2014; Sardos et al., 2018; Sardos, Paofa, Sachter-Smith, et al., 2019; Sardos, Sachter-Smith, Ghanem, et al., 2019; Sardos, Sachter-Smith, Shandil, et al., 2019); three samples were obtained from an on-farm project in Papua New Guinea (Sardos et al. *in prep.*); and four reference samples were sampled in situ in Papua New Guinea. Young leaf tissues of samples collected in-situ were silica dried on site. A set of two samples from the Centre de Ressources Biologiques: Plantes Tropicales (CRB-PT) were obtained as public sequence data from the study of Martin, Cottin, et al. (2023).

2.2 | Passport data curation

In total, the 317 samples of the study originated from 30 countries and included 302 distinct names, with some samples either lacking names or containing homonyms. Significant curation work was undertaken on the taxonomical classification of accessions used in this study to assign or exclude accessions from cultivar groups before genomic characterization (Table S1). First, passport data available from the *Musa* Germplasm Information System (MGIS—www.crop-diversity.org/mgis) (Ruas et al., 2017) was retrieved. Second, and when available, morphological characteristics and taxonomic expert recommendations (Taxonomic Advisory Group members of MusaNet) obtained within the field verification exercise (Chase et al., 2016; Van den Houwe et al., 2020) were checked to correct or refine the classification of some accessions. When no recent field observations were available, collecting mission reports were checked for collector observations in the field. This was notably the case for the collecting mission reports to Papua New Guinea, Cook Islands, Samoa, and Tanzania (Arnaud & Horry, 1997; Byabachwezi et al., 2005; De Langhe et al., 2001; Irish et al., 2016; Sachter-Smith et al., 2021; Sachter-Smith & Sardos, 2021a, 2021b; Sardos et al., 2017).

2.3 | DNA sequencing and genotyping

This study spanned several years and included the preparation and sequencing of accessions in separate batches. As a result, distinct but

comparable genotyping technologies were utilized according to the most effective approach at each point in time. For all experiments, DNA from each accession was extracted following a cetyltrimethylammonium bromide (CTAB) protocol (modified from Risterucci et al., 2000). The libraries for restriction-site-associated DNA sequencing were built with the PstI, or PstI/MseI restriction enzymes, followed by the addition of barcoded adapters, DNA shearing, amplification, and sequencing. The sequencing data were thus generated using either RADseq, ddRADseq, or GBS techniques (Table S1), following the respective protocols established by Davey et al. (2010) or Elshire et al. (2011). For RADseq, short-insert libraries (300–500 bp) were sequenced to produce 91 bp paired-end reads on an Illumina HiSeq2000 (BGI, Hong Kong, China). For GBS and ddRADseq, libraries were sequenced as 150 bp paired-end reads on an Illumina HiSeq2500 (Genewiz, Azenta Life Sciences, USA) and Illumina NovaSeq 6000 (LGC Genomics GmbH, Germany), respectively.

2.4 | Single nucleotide polymorphism (SNP) callings

After demultiplexing with GBSX v1.3 (Herten et al., 2015), FASTQ files (one for each sample) were examined with FastQC v0.11.7. We then used Cutadapt v2.1 to clean them by eliminating Illumina adapter sequences and trimming low-quality ends with a Phred score > 20 (Martin, 2011). Any reads shorter than 30 bp after post-trimming were removed. These reads were subsequently mapped using BWA-MEM v0.7.12 (Li & Durbin, 2010) to the *M. acuminata* DH Pahang genome v4 (Belser et al., 2021; D'Hont et al., 2012), downloaded from the Banana Genome Hub (Droc et al., 2022). Re-alignment was done with the IndelRealigner module from GATK v4.1. We then followed the GATK pipeline recommended for a non-model organism by adding the recalibration step. This consisted of performing an initial round of SNP calling on the original uncalibrated data, selecting the SNPs with the highest confidence, and then executing a round of base recalibration on the original mapped reads files. For duplicate samples, a script using Sambamba software was used to merge the recalibrated bam alignment files. The GATK module HaplotypeCaller v4.1 was then used for SNPs and indels calling. Finally, a script gVCF2vcf_gz.pl was written to combine the individual gVCF files obtained into a single VCF file. The GenomicDB procedure from GATK was used to build the gVCF SNP database, containing all the positions, variants, and non-variants. The snpcluster exclusion procedure was used to process SNP clusters, set for a threshold of three or more SNPs per 10 bp window. The pipeline used to perform SNP analyses is available at https://github.com/CathyBreton/Genomic_Evolution.

2.5 | Genome ancestry mosaic painting

From the resulting VCF files, we used the scripts provided in the VCFHunter v2.1.2 suite (https://github.com/SouthGreenPlatform/VcfHunter). Only bi-allelic sites were conserved, and for each accession, sites with less than 10 reads or more than 1000 reads were converted to missing data as well as heterozygous sites with a minimal frequency in the individual <0.05 (i.e., vcfFilter.1.0.py MinCov:10; MaxCov:1000; minFreq:0.05; MinAl:3; RmAlAlt 1:3:4:5:6:7:8:9:10). In the next step, we conserved the alleles in common with the alleles identified for 11 ancestral gene pools in Martin, Cottin, et al. (2023) (see identification of ancestry informative alleles), using the vcfSelect.py script. Because this dataset was previously obtained from the whole genome scale, it was possible to intersect the genome position of common SNP positions inferred from any genotyping method and map on the same reference genome. Then, the allele ratio in individuals was calculated with the allele_ratio_per_acc.py script, generating one file per accessions containing counted allele ratio according to allocated ancestral gene pools (statistics in Table S1). When necessary, these files were curated to define the ancestry mosaics of unresolved chromosome segments and to infer potential haplotypes. Finally, genome ancestry mosaics SNP ratios and ancestry allocation were curated and refined, and graphical visualizations were drawn using GeMo (Summo et al., 2022). Because the mosaics were inferred using nonphased data, the juxtaposition of segments represents introgressions at the position but may not reflect the real haplotype of a given chromosome.

To characterize each cultivar group at molecular level, one accession—considered by expert co-authors as good representatives of each cultivar group—was selected as a reference (as shown in Table 1). Then, patterns of other accessions were compared against the mosaic pattern of this reference accession. During this comparison, certain accessions showed signs of aneuploidy, which could result from in vitro conservation processes, especially in the cases of deletions or duplications of chromosomes or chromosome arms (Breton et al., 2022). However, these aspects of aneuploidy are not elaborated upon in this study or represented graphically in the catalog (Dataset S1). Other events, usually smaller in size and repeated in several accessions, such as small duplications and deletions, were considered ancestral events that accounted for the creation of different patterns, as described in Martin, Cottin, et al. (2023).

2.6 | Principal component analysis (PCA)

To visualize global diversity and relatedness between groups, clusters, and individuals, we calculated, for each pattern, the respective percentage of each ancestor (Table S2). One representative (as shown in Table 1) by mosaic pattern was used. This data was used to generate a matrix that was subsequently analyzed using PCA with the scikit-learn Python library (Pedregosa et al., 2011) and visualized using the "plotly" graphical rendering Python library.

3 | RESULTS

In our study, we analyzed 317 accessions, precisely curated for their passport data, using genome ancestry mosaic painting to reveal

TABLE 1 Overview of banana cultivar groups and their genomic compositions, alongside the number of accessions and mosaic patterns identified for each cultivar group. It also includes reference accession names with their identifiers and the total number of samples assigned to each cultivar group.

Cultivars group	Genomic composition	Nb of accessions	Nb of mosaics	Reference accessions	Nb of samples
Sucrier (Pisang Mas)	AA	8	1	ITC0653 Pisang Mas	8
Pisang Jari Buaya	AA	8	1	ITC0312 Pisang Jari Buaya	8
M'chare (Mlali)	AA	3	1	ITC1223 Mchare	3
Pisang Lilin	AA	2	1	ITC0395 Lidi	2
Cavendish	AAA	19	1	ITC1471 Zanzebar	19
Gros Michel	AAA	3	1	ITC0724 Cocos	3
Red	AAA	2	1	ITC1833 Shwe Ni	2
Mutika/Lujugira	AAA	24	4	ITC1630 Enjagata	19
				ITC0082 Intokatoke	3
				ITC1770 Siira	1
				ITC0084 Mbwazirume	1
Ilalyi	AAA	6	1	ITC1451 Kitarasa	6
Ambon	AAA	1	1	DYN122 Hom Thong Mokh	1
Orotava	AAA	1	1	DYN121 Hom Sakhon Nakhon	1
Rio	AAA	1	1	ITC0277 Leite	1
Ibota	AAA	2	1	ITC0662 Khai Thong Ruang	2
Kunnan	AB	5	2	ITC1034 Kunnan	3
				ITC1752 Poovilla Chundan	2
Ney Poovan	AB	3	2	ITC0245 Safet Velchi	2
				ITC1751 Adukka Kunnan	1
Plantain	AAB	58	2	ITC0033 Bungaoisan	7
				ITC0007 Asamiensa	42
				Not assigned[a]	9
Maia Maoli/Popoulu	AAB	17	4	ITC0733 Ihi U Maohi	3
				ITC1135 Popoulou (CMR)	12
				COOK009 Torotea	1
				WNB043 Lavugi	1
Iholena	AAB	6	1	ITC0825 Uzakan	6
Laknau	AAB	4	1	ITC0332 Laknao	4
Pome (Prata)	AAB	10	3	ITC0649 Foconah	6
				ITC1723 Ladies Finger	3
				ITC0582 Lady Finger	1
Mysore	AAB	5	1	ITC1613 Karpura Chakkrakeli	5
Silk	AAB	15	2	ITC0348 Silk	11
				ITC0737 Kingala no. 1	4
Pisang Raja	AAB	2	1	ITC0587 Pisang Raja	2
Pisang Awak	ABB/ABBB	17	5	ITC0659 Namwa Khom	7
				ITC1719 Chinia	6
				Ramu Yawa	1
				ITC0213 Pisang Awak	1
				ITC0334 Nzizi	2
Bluggoe[b]	ABB	9	1[b]	ITC0643 Cachaco	9
Monthan[b]	ABB	7	1[b]	ITC1483 Monthan	7
Ney Mannan[c]	ABB	7	1[c]	ITC0361 Blue Java	7

(Continues)

TABLE 1 (Continued)

Cultivars group	Genomic composition	Nb of accessions	Nb of mosaics	Reference accessions	Nb of samples
Peyan[c]	ABB	1	1[c]	ITC0123 Peyan	1
Klue Tiparod	ABB	2	1	ITC0652 Kluai Tiparot	2
Pelipita	ABB	3	1	ITC0472 Pelipita	3
Kalapua	ABB	5	2	ITC2017 Kalapua	3
				Dwarf Kalapua	2
Total		256	46		256

Abbreviation: SNP, single nucleotide polymorphism.
[a]SNP density not high enough for precise determination.
[b]Identical pattern.
[c]Identical pattern.

genomic mosaic patterns at the chromosome level. By comparing patterns and using the curated taxonomical assignment of each accession, we were able to identify reference patterns for cultivar groups defined morphologically in the Musalogue (Daniells et al., 2001) and for the Ilalyi group in De Langhe et al. (2001). Additionally, our analysis identified accessions with mosaic patterns that did not correspond to any known cultivar group, both morphologically and genomically. If these unclassified accessions had unique genomic patterns or matched exclusively with accessions known to be synonyms, that is, the same cultivars with different names, we regarded them as individual accessions (Table S3). However, if the same genomic pattern appeared in several accessions that were not synonyms, we grouped them as clusters of morphological variants. The cultivar group patterns, clusters of accessions, and specific genomic patterns of individual accessions were organized and compiled into a catalog (Dataset S1), as exemplified in Figure 1. Nine remaining patterns for which morphological characterization is still ongoing were presented separately (Dataset S2). Through this process, we identified a total of 83 unique mosaic patterns. Of these, 46 patterns (265 accessions) matched 31 previously defined cultivar groups (Table 1), while the remaining 37 patterns corresponded to 61 additional accessions or clusters of accessions (Tables 1 and S2). Among the cultivar groups for which more than one accession was available, 18 were homogeneous, that is, they were composed of accessions with strictly identical genomic mosaics. Conversely, the nine other cultivar groups displayed heterogeneity, that is, they were composed of accessions with several mosaic patterns.

Overall, the cultivar groups were well differentiated, and their genetic backgrounds were sufficiently discriminating to assign accessions to specific cultivar groups. In the PCA, the different patterns corresponded to the cultivar groups, the cultivar clusters and the individual accessions are well discriminated, and the clustering is consistent with the taxonomic affiliation. The first axis discriminates patterns according to the contribution of the different *M. acuminata* sub-species, notably *M. acuminata* ssp. *banksii* on the left and *M. acuminata* ssp. *malaccensis* on the right while the second axis discriminates the contributions of *M. acuminata* at the bottom and *M. balbisiana* at the top (Figure 2). Besides, the genomic mosaic patterns of the described cultivar groups exhibited a range of ancestral contributions (Table 2). The *M. acuminata* ssp. *banksii*, extended with the accessions "Agutay" *(ssp. errans)* and "borneo" *(ssp. microcarpa)* (referred to as banksii), is the only ancestral genepool for which centromeres were always present, with a minimum of two centromeres being observed in seven patterns. *M. acuminata* ssp. *zebrina* (referred to as zebrina) and *M. schizocarpa* (referred to as schizocarpa) were consistently present across cultivar groups, at least in the form of introgressions, corroborating the findings reported by Martin, Cottin, et al. (2023). However, notable exceptions were observed in the cultivar group Klue Teparod (ABB) from mainland Southeast Asia and the "Auko" clones (ABB) from Papua New Guinea. *M. acuminata* ssp. *malaccensis* (referred to as malaccensis) was also present in 35 patterns (24 cultivar groups). Then, the presence of previously unknown genepools, referred to as m1 and proposed to be *M. acuminata* ssp. *halabanensis* (referred to as halabanensis) in Martin, Cottin, et al. (2023), and m2 (referred to as unknown) were present in 14 (11 cultivar groups) and 18 patterns (15 cultivar groups), respectively. *M. acuminata* ssp. *burmannica*, including ssp. *siamea* (referred to as burmannica) was found to contribute only to the Klue Teparod cultivar group, along with banksii and schizocarpa for the A haplotype. The B genome contributor, *M. balbisiana*, (referred to as balbisiana) was included in two thirds of the patterns within cultivar groups. It contributed a minimum of 10 centromeres in three triploid cultivar groups and up to 34 centromeres in the Pisang Awak-4x-3 pattern. These proportions were in general in line with expectations based on the genomic constitution AB, AAB, or ABB. However, and as noted in Cenci et al. (2021), its proportion varied from strict 1:2, 1:3, 2:3, or 3:4, depending on the presence of homoeologous exchanges between the A and B genomes. Several of the individual accessions exhibited very distinctive patterns. This included accessions with a notable contribution from burmannica, or cultivars with a high contribution of zebrina, as well as tetraploids that exhibited a complete haplotype of the Australimusa (T) genome, now included in the former Callimusa section (Häkkinen, 2013).

3.1 | Genetically homogeneous cultivar groups

Homogeneity was observed across cultivar groups of various ploidy levels and genomic compositions.

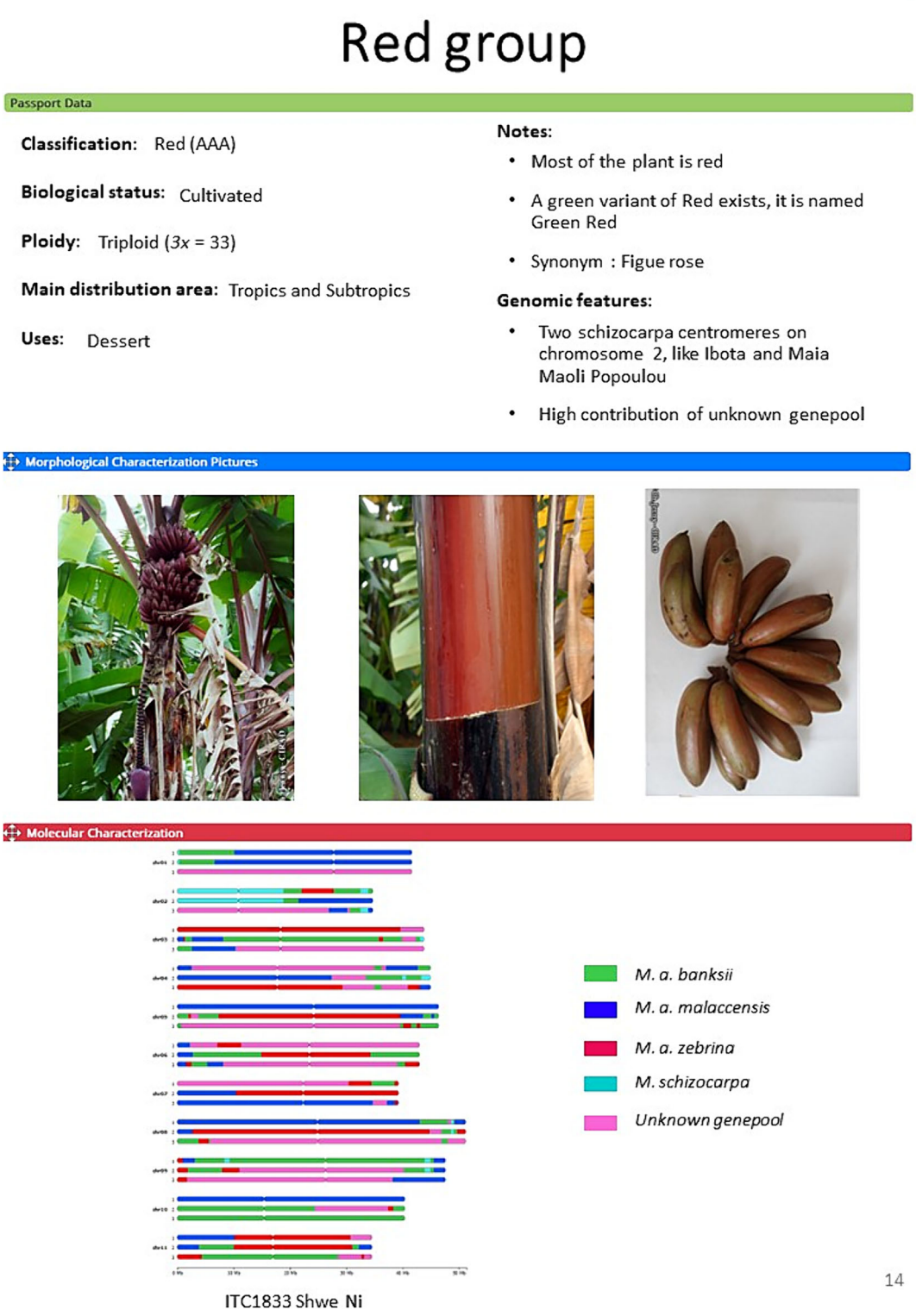

FIGURE 1 Example of the Red cultivar group in the catalog. Each entry is divided into three sections: passport data, morphological pictures, and molecular characterization. The passport data section includes the biological status, ploidy, main distribution area, and uses with notes on the cultivar group and genomic features. Up to three pictures are intended to be representative of the cultivar group. The molecular characterization contains from one to five mosaic patterns with the name of the reference accession(s) used for the mosaics painting. Colored segments show the contribution of each of the ancestral genepools (Martin, Cottin, et al., 2023).

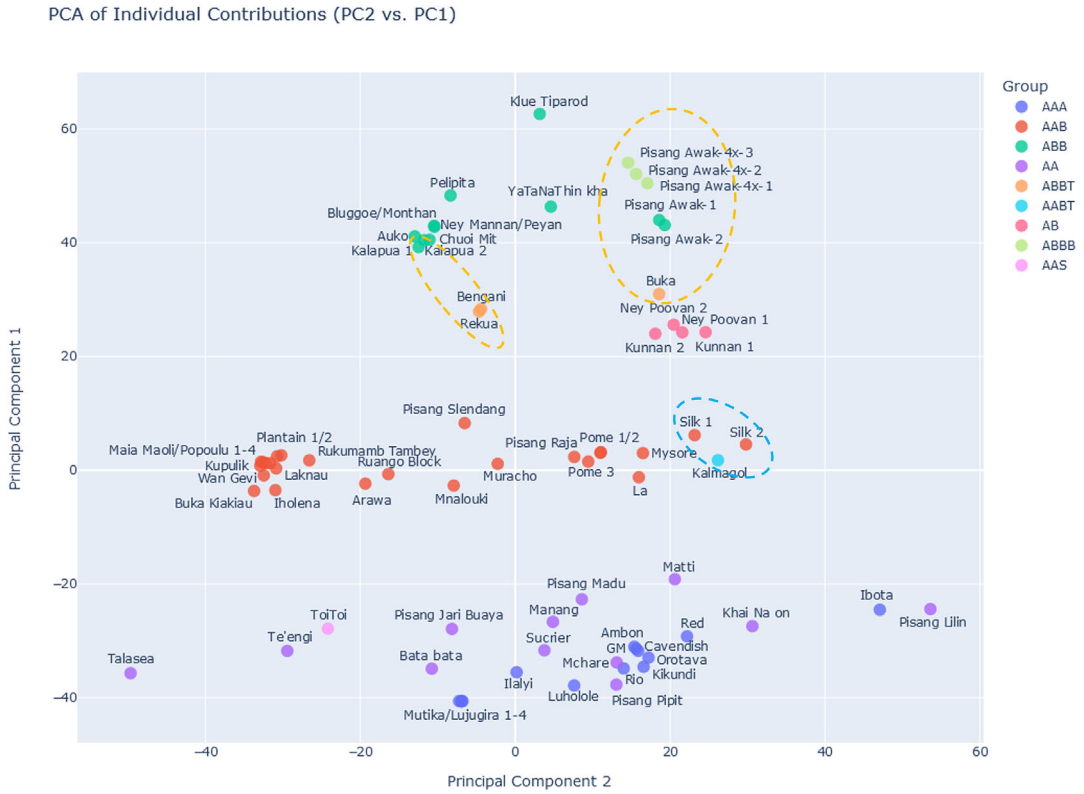

FIGURE 2 Principal component analysis (PCA) of banana cultivar groups, clusters, and individual accessions of the catalog. PCA axis 1 and 2 explain a total of 74.79% (PC1: 49.22%; PC2: 25.58%) of the variance. Dots are colored according to the genome constitution (i.e., AA, AAA, AB, AAB, ABB, ABBB, AAS, AABT, and ABBT). PC1 discriminates the gradient of balbisiana genepool contribution, with no contributions at the bottom and higher contributions at the top. PC2 discriminates between other genepool contributors, positioning *Musa acuminata* malaccensis-rich accessions more on the right and *Musa acuminata* banksii-rich accessions more on the left of the PCA plot.

3.1.1 | Diploid cultivar groups

With an AA genomic constitution, Pisang Jari Buaya (comprising seven accessions) and Sucrier (eight accessions) were homogenous in their mosaic patterns. The patterns of the two accessions classified as Mchare were identical, as well as for the two accessions identified as Pisang Lilin.

3.1.2 | Triploid cultivar groups

In the triploid cultivar groups with an AAA genomic composition, no variation was identified within the 18 accessions of Cavendish analyzed. Similarly, Gros Michel (three accessions), Red (two accessions), and Ibota (two accessions) were homogeneous. After curation, we identified six accessions from Tanzania wrongly assigned to the Mutika/Lujugira group that corresponded to the Ilalyi group as described by De Langhe et al., 2001, and genetically validated in Perrier et al. (2019). We therefore revived the Ilalyi group that was previously removed from the passport data. These accessions shared the same mosaic pattern, with an ancestral basis composed of a combination of zebrina and banksii similar to the one observed in the Mutika/Lujugira group, but with an additional important contribution of malaccensis (five centromeres) (Table 2). Lastly, due to the availability of only one accession for each of the cultivar groups Ambon, Rio, and Orotava, it was not possible to investigate potential variations within their respective mosaic patterns. The study would benefit from more samples to confirm their monoclonal status.

In the triploid cultivar groups with an AAB genomic composition, Laknau (four accessions), Mysore (five accessions), Pisang Raja (two accessions), and Iholena (six accessions) were homogenous.

In the triploid cultivar groups with an ABB genomic composition, the Pelipita (three accessions) and Klue Teparod (two accessions) groups were genetically uniform, but a wider sampling remains necessary to validate this observation. Finally, we identified two patterns corresponding to four cultivar groups with a shared A genome background. The Bluggoe group (10 accessions) and the Monthan group (five accessions) exhibited identical genetic mosaic patterns, despite differences in morphologies. For instance, Bluggoe fruits are mostly straight and horizontal or slightly erect, while Monthan's curve upwards. Similarly, the Ney Mannan (seven accessions) and Peyan (one accession) groups also shared the same mosaic pattern, although we noted slight differences in the levels of heterozygosity within the *M. balbisiana* haplotypes. A larger sample of Peyan representatives and better discrimination of allelic diversity in the B genome would be necessary to provide clearer insights.

TABLE 2 Genepool contributions to various banana cultivar groups and mosaic patterns, detailing the number of centromeres contributed by each genetic source based on the reference chromosome structure.

Cultivars group/pattern	Genome group	Ab	Az	Am	As	Ah	S	U	B	Total	Genepool contribution
Bata Bata cluster	AA	13	4	4	0	1	0	0	0	22	Ab-Az-Am-Ah-S-U
Mchare	AA	6	8	7	0	1	0	0	0	22	Ab-Az-Am-Ah-S-U
Pisang Jari Buaya	AA	7	4	0	0	11	0	0	0	22	Ab-Az-Am-Ah-S
Pisang Lilin	AA	2	1	17	0	0	0	2	0	22	Ab-Az-Am-Ah-S-U
Sucrier	AA	8	4	3	0	1	1	5	0	22	Ab-Az-Am-Ah-S-U
Te'engi cluster	AA	16	1	3	0	0	2	0	0	22	Ab-Az-Am-S
Ambon	AAA	9	7	10	0	0	0	7	0	33	Ab-Az-Am-S-U
Cavendish	AAA	7	9	8	0	2	0	7	0	33	Ab-Az-Am-Ah-S-U
Gros Michel	AAA	9	9	9	0	1	1	4	0	33	Ab-Az-Am-Ah-S-U
Ibota	AAA	4	2	22	0	0	2	3	0	33	Ab-Az-Am-Ah-S-U
Ilalyi	AAA	14	10	5	0	1	2	1	0	33	Ab-Az-Am-Ah-S-U
Mutika/Lujugira 1/2/3/4	AAA	12	18	0	0	1	2	0	0	33	Ab-Az-Ah-S-U
Kikundi cluster	AAA	8	11	12	0	0	0	2	0	33	Ab-Az-Am-S-U
Arawa	AAA	11	15	6	0	0	1	0	0	33	Ab-Az-Am-S-U
Orotava	AAA	10	10	9	0	0	1	3	0	33	Ab-Az-Am-S-U
Red	AAA	5	8	7	0	0	2	11	0	33	Ab-Az-Am-S-U
Rio	AAA	7	14	6	0	1	1	4	0	33	Ab-Az-Am-Ah-S-U
Iholena	AAB	19	3	0	0	0	1	0	10	33	Ab-Az-Am-S-B
Rukumamb Tambey cluster	AAB	17	2	1	0	0	2	0	11	33	Ab-Az-Am-S-B
Arawa cluster	AAB	15	6	1	0	0	1	0	10	33	Ab-Az-Am-S-B
Laknau	AAB	18	2	0	0	0	2	0	11	33	Ab-Az-Am-S-B
MMP-1/2	AAB	18	1	0	0	0	4	0	10	33	Ab-Az-Am-S-B
MMP-3/4	AAB	18	1	0	0	0	4	0	10	33	Ab-Az-S-B
Wan Gevi cluster	AAB	19	1	0	0	0	3	0	10	33	Ab-Az-S-B
Buka Kiakiau	AAB	18	1	0	0	0	4	0	10	33	Ab-Az-S-B
Ruango block cluster	AAB	16	2	4	0	0	1	0	10	33	Ab-Az-Am-S-U-B
Mysore	AAB	5	10	6	0	0	0	1	11	33	Ab-Az-Am-S-U-B
Pisang Raja	AAB	6	9	1	0	0	0	6	11	33	Ab-Az-Am-S-U-B
Plantain-1/2/3	AAB	19	0	1	0	0	1	0	12	33	Ab-Az-Am-S-B
Kupulik cluster	AAB	19	2	0	0	0	1	0	11	33	Ab-Az-S-B
Mnalouki	AAB	13	2	5	0	1	1	0	11	33	Ab-Az-Am-Ah-S-U-B
Pome-1/2/3	AAB	6	8	7	0	1	0	0	11	33	Ab-Az-Am-Ah-S-U-B
Silk-1	AAB	7	2	13	0	0	0	0	11	33	Ab-Az-Am-S-B
Silk-2	AAB	5	2	16	0	0	0	0	10	33	Ab-Az-Am-S-B
Kunnan-1	AB	2	1	8	0	0	0	0	11	22	Ab-Az-Am-S-B
Kunnan-2	AB	4	1	6	0	0	0	0	11	22	Ab-Az-Am-S-B
Ney Poovan-1	AB	3	1	7	0	0	0	0	11	22	Ab-Az-Am-S-B
Ney Poovan-2	AB	3	0	7	0	0	1	0	11	22	Ab-Az-Am-S-B
Bluggoe/Monthan	ABB	8	0	0	0	0	2	0	23	33	Ab-Az-S-B
Kalapua-1/2	ABB	10	1	0	0	0	0	0	22	33	Ab-Az-S-B
Klue Tiparod	ABB	3	0	0	2	0	0	0	28	33	Ab-As-S-B
Ney Mannan/Peyan	ABB	8	0	0	0	0	2	0	23	33	Ab-Az-S-B
Pelipita	ABB	8	0	0	0	0	0	0	25	33	Ab-Az-Am-S-B
Pisang-Awak-1/2	ABB	2	1	7	0	0	1	0	22	33	Ab-Az-Am-S-B
Pisang Awak-4x-3	ABBB	2	1	6	0	0	1	0	34	44	Ab-Az-Am-S-B

(Continues)

TABLE 2 (Continued)

Cultivars group/pattern	Genome group	Ab	Az	Am	As	Ah	S	U	B	Total	Genepool contribution
Pisang-Awak-4x-1/2	ABBB	2	1	7	0	0	1	0	33	44	Ab-Az-Am-S-B
La	AAB	3	13	0	0	0	0	5	11	32	Ab-Az-Am-Ah-S-U-B
Auko	ABB	9	0	0	0	0	2	0	22	33	Ab-S-B
Choi Mit	ABB	9	0	0	0	0	2	0	22	33	Ab-Az-S-B
Choi Xi Mon	ABX	9	1	0	0	0	1	0	11	22	Ab-Az-Am-S-B
Ya Ta Na Thin Kha	ABB	3	0	0	8	0	0	0	22	33	Ab-Az-As-S-B
Khai Na On	AA	3	4	8	0	1	1	5	0	22	Ab-Az-As-Ah-S-U
Manang	AA	8	2	3	4	1	1	3	0	22	Ab-Az-Am-As-Ah-S-U
Matti	AA	4	0	7	11	0	0	0	0	22	Ab-Az-Am-As-S
Pisang Madu	AA	5	0	0	0	11	0	6	0	22	Ab-Az-Am-Ah-S-U
Pisang Pipit	AA	5	7	1	0	1	0	8	0	22	Ab-Az-Am-Ah-S-U
Talasea cluster	AA	18	2	0	0	0	2	0	0	22	Ab-Az-S
ToiToi	AAS	14	3	0	0	1	15	0	0	33	Ab-Az-Ah-S-U
Pisang Slendang	AAB	10	2	1	0	1	3	4	12	33	Ab-Az-Am-Ah-S-U-B
Kalmagol	AABT	4	2	16	0	0	1	0	10	33	Ab-Az-Am-S-B
Bengani	ABBT	10	1	0	0	0	0	0	22	33	Ab-Az-Am-S-B
Rekua	ABBT	10	1	0	0	0	0	0	22	33	Ab-Az-Am-S-B
Buka cluster	ABBT	2	1	7	0	0	1	0	22	33	Ab-Az-Am-S-B
Pisang Buntal	AA	8	2	5	0	1	1	5	0	22	Ab-Az-Am-Ah-S-U
Muku Bugis	AB	7	3	0	0	0	1	0	11	22	Ab-Az-Am-S-B
Mu'u Seribu	AB	10	1	0	0	0	0	0	11	22	Ab-Az-Am-S-U-B
Waga	AAA	18	3	0	0	0	12	0	0	33	Ab-Az-S-U
Pisang Nangka	AAB	11	4	10	0	0	2	2	4	33	Ab-Az-Am-Ah-S-U-B
Bagatow	AAB	18	2	0	0	0	1	0	12	33	Ab-Az-S-B
Muracho	AAB	10	8	1	0	0	0	3	11	33	Ab-Az-Am-S-U-B
Titikaveka Red	AAB	4	9	9	0	1	0	0	10	33	Ab-Az-Am-AH-S-U-B
Pata Tonga	ABB	9	1	0	0	0	0	1	22	33	Ab-Az-S-U-B

Note: The contributors are denoted as follows: Ab for *banksii*, Az for *zebrina*, Am for *malaccensis*, As for *burmannica/siamea*, Ah for *halabanensis*, S for *schizocarpa*, U for unknown, and B for *balbisiana*.

3.2 | Genetically heterogeneous cultivar groups

Two types of heterogeneous cultivar groups could be distinguished. The first type, which includes cultivar groups such as Plantain and Mutika/Lujugira, displayed mosaic patterns differentiated by only small chromosomal regions showing variations in allelic ratio. In the second type of heterogeneous cultivar groups, we found two or more mosaic patterns, each exhibiting multiple differences likely resulting from different mechanisms of diversification.

3.2.1 | Mutika/Lujugira

Mutika/Lujugira is a triploid cultivar group with a AAA genomic constitution that is typical to Burundi, Uganda, the Democratic Republic of Congo, Cameroon, Kenya, Rwanda, and Tanzania. For this cultivar group, we conducted an important curation of the passport data, notably by consulting the collecting mission reports when available. In this set of 34 AAA accessions from East Africa, we identified four nearly identical mosaics corresponding to 24 accessions, including well-described Mutika/Lujugira cultivars such as the popular "Mbwazirume" (Shepherd, 1957). These mosaics' main contributors were zebrina, banksii, and schizocarpa. A pattern variation was observed in "Guineo," "Intokatoke," and "Makara," in which a small interstitial region of the second arm of chromosome 10 displayed a mitotic homologous exchange between one of the zebrina haplotypes and the banksii haplotype. In addition, and on the same chromosomal region, the "Siira" accession exhibited a small deletion on the banksii haplotype. Finally, the accession "Mbwazirume" also appeared to be a variant with a switch of the allelic ratio resulting from a mitotic recombination on the first telomere of chromosome 10 (Figure S1). No correlation was found between these variations and the proposed clone sets from Karamura et al. (2010). Finally, three additional mosaics discovered in this set were not assigned to Mutika/Lujugira but corresponded to Tanzanian accessions, including those from the homogeneous Ilalyi group described earlier and four accessions that are discussed later.

3.2.2 | Plantain

The 58 Plantain accessions of the sample were remarkably homogeneous, but a small variation was detected on chromosome 10 between ~8.5 and 13 Mb. We interpreted this change as a diploid region (balbisiana—banksii) resulting from a small deletion of one of the banksii haplotypes present in the original pattern. This predominant variation was detected in 42 accessions out of 49, for which this region could be characterized. Interestingly, the four accessions with origins in Asia, "Bungaoisan" (a medium French) and "Daluyao" (a medium True Horn) from the Philippines, as well as "Mantreken" from Indonesia and "Nendran" (a French) from India, all had the original mosaic pattern without deletion. The three other accessions with a mosaic without the deletion were "Agbagba" from Nigeria (a medium false-horn Plantain widely cultivated in West Africa according to Adheka et al., 2013), "Big Ebanga" (a giant false horn from Cameroon, possibly a synonym of Agbagba), and Maiden Plantain (a French Plantain received from Honduras but of unknown origin).

3.2.3 | Pome

Pome cultivars are AAB triploids that originated in India and are now very popular in Brazil and Hawaii under the name Prata. Ten accessions of the sample displayed a mosaic pattern associated with the Pome group, within which three pattern variations were observed. The first one, identified in six accessions from different countries, shows one deviation from a pure AAB pattern, that is, an A/B recombination (A3:B0) on the first telomere of chromosome 3. The second one, present in three Pome accessions received recently from India, also shows an A/B recombination (A3:B0), but in the interstitial region of the first arm of chromosome 9. This recombination is also present in the third pattern, in addition to another one on the first telomere of chromosome 10 (Dataset S1). This last pattern was identified in one accession from Australia, "Lady Finger (Nelson)," sometimes referred to as belonging to the Nadan group in other collections and which is tetraploid for chromosome 8. Except for the extra chromosome 8 of Pome 3, which has an *M. acuminata—M. schizocarpa* ancestry, the variations observed between the three patterns are linked to A-donor introgressions in B chromosomes. Because all these introgressions correspond to genepools also present in one of the two A genomes, it is difficult to assess whether the three Pome patterns were derived clonally from each other or were obtained through different sexual events. However, the banksii introgression observed on the B chromosome 9 of Pome-2 and Pome-3 occurs frequently in cultivars of AB, AAB, or ABB genomic constitutions, suggesting a common ancestry. This pattern is more likely to have been inherited sexually rather than arising from a new, independent mitotic recombination. The observed variations may have resulted from a combination of clonal diversification and sexual events (at least two from similar parents).

3.2.4 | Maia Maoli/Popoulu (MMP)

After curation of passport data and a morphological trait check, we identified 16 accessions affiliated to the Maia Maoli/Popoulu group corresponding to four genomic patterns. The Maia Maoli/Popoulou patterns, with an AAB genomic composition, are characterized by contributions from banksii, schizocarpa, and zebrina for the A genome, with little to no presence of malaccensis. A striking feature of these four patterns is the absence of B centromeres and the presence of two S centromeres on chromosome 2 (Table 2, Dataset S1). Out of the 16 accessions, MMP-2 was the most frequent pattern, with 12 representatives from both Polynesia (Cook Islands, Samoa, and Hawaii) and Melanesia (Bougainville and New Britain Islands in Papua New Guinea). Three accessions from the Cook Islands, Tahiti, and Samoa exhibited the pattern MMP-1. The differences observed between MMP-1 and MMP-2 are slight. The pattern MMP-1 has a small balbisiana introgression in the interstitial region of the first arm of chromosome 1 and a duplication of the balbisiana first telomere on chromosome 7. The patterns MMP-3 and MMP-4 were identified in one accession each, "Mango Torotea" from Cook Islands and "Lavugi" from New Britain Island, respectively. These patterns are significantly different from MMP-1 and MMP-2, with more than 10 discriminating regions each. They also differ from each other by 15 events. If MMP-1 and MMP-2 may be derived from each other by clonal diversification, this is likely not the case of MMP-3 and MMP-4, which may have resulted from independent sexual events with similar parental contribution.

3.2.5 | Silk

Two closely related patterns were detected in the 15 accessions classified as Silk, confirming previous findings (Sardos et al., 2016). Eleven accessions were displaying the pattern Silk-1, while four displayed the pattern Silk-2. Interestingly, the Silk-2 accessions were all collected in Africa (Burundi, Tanzania, and Congo). The differences observed between the two Silk genomic mosaic patterns were important. Notably, five of their 33 centromeres were of different origins. For example, on chromosome 5, both Silk groups display one B chromosome, but Silk-1 displays two banksii centromeres, while Silk-2 displays one banksii and one zebrina centromere (Table 2). These differences cannot result from clonal diversification. Therefore, the diversity observed within the Silk group results from two different sexual events, probably from parents of similar genetic background.

3.2.6 | Kalapua

The Kalapua group, characterized by an ABB genomic composition, is a popular cooking banana cultivar in Papua New Guinea. Within our sample set of five Kalapua accessions, we identified two distinct genomic mosaics. The first mosaic pattern was present in three of the samples, while the second pattern was found in the remaining two. The

observed differences between these mosaics consist of varying proportions of A and B genomes in four specific genomic regions. The first telomere of chromosome 4 and the interstitial region of the first arm of chromosome 8 are A2:B1 in Kalapua-1 and A1:B2 in Kalapua-2. In addition, on chromosome 9, the first telomere is A1:B2 in Kalapua-1 and A0:B3 in Kalapua-2, while the second telomere is A2:B1 in Kalapua-2 and A1:B2 in Kalapua-1. Kalapua patterns may have resulted from the accumulation of mitotic homoeologous chromosome exchanges. However, mitotic recombination events are rare, and these four cumulated events may alternatively have resulted from two sexual events among similar parents.

3.2.7 | Pisang Awak

The Pisang Awak group comprised two triploid and three tetraploid patterns present in our sample (Figure 2). For the triploid patterns, the pattern Pisang-Awak-1 (PA-1) was found in seven accessions, while Pisang-Awak-2 (PA-2) was found in two accessions from India. Eight variations in the patterns of homologous exchanges between the A and B genomes were observed between these two mosaic patterns. Differences in A/B homoeologous exchanges consisted either in the presence or absence of these events or in variations in the size of common events. This finding is not consistent with two genotypes deriving from each other clonally and rather supports the idea that they were both sexually produced, probably from the same AB parent who produced recombined but unreduced (2x) gametes. Furthermore, A/B homoeologous exchanges enabled us to hypothesize a pedigree relationship with the three tetraploid patterns identified in this cultivar group (four samples). These four accessions, with ABBB genomic composition, were morphologically included in the Pisang Awak group and were produced from unreduced (3x) ABB triploid Pisang Awak gametes crossed with a haploid (1x) B gamete. The A/B introgression patterns observed in the triploid and tetraploid Pisang Awak samples support pedigree relationships between PA-1 and "Ramu Yawa" (Pisang-Awak-4x-1) from Papua New Guinea. Equally, direct ancestry can be inferred between PA-2 and "Pisang Awak" (Pisang-Awak-4x-2) from Sri Lanka. The third tetraploid pattern, discovered in "Foulah 4" and "Nzizi" (Pisang-Awak-4x-3) from Ivory Coast and Nigeria, respectively, did not correspond to any of the triploids Pisang Awak described, suggesting that at least a third triploid form may exist or may have existed (Figure 3).

3.2.8 | Kunnan and Ney Poovan

Two cultivar groups from India with AB genomic composition are defined, the Ney Poovan and the Kunnan groups, but their morphological characteristics are not clear. In our sampling, eight accessions with an AB genomic composition could be affiliated with either cultivar group. These accessions displayed four different mosaic patterns with at least one A introgression into their B genome (consistent with Cenci et al., 2021) and with a malaccensis ancestry dominance as A-donor genome. Two patterns were discovered in both Ney Poovan

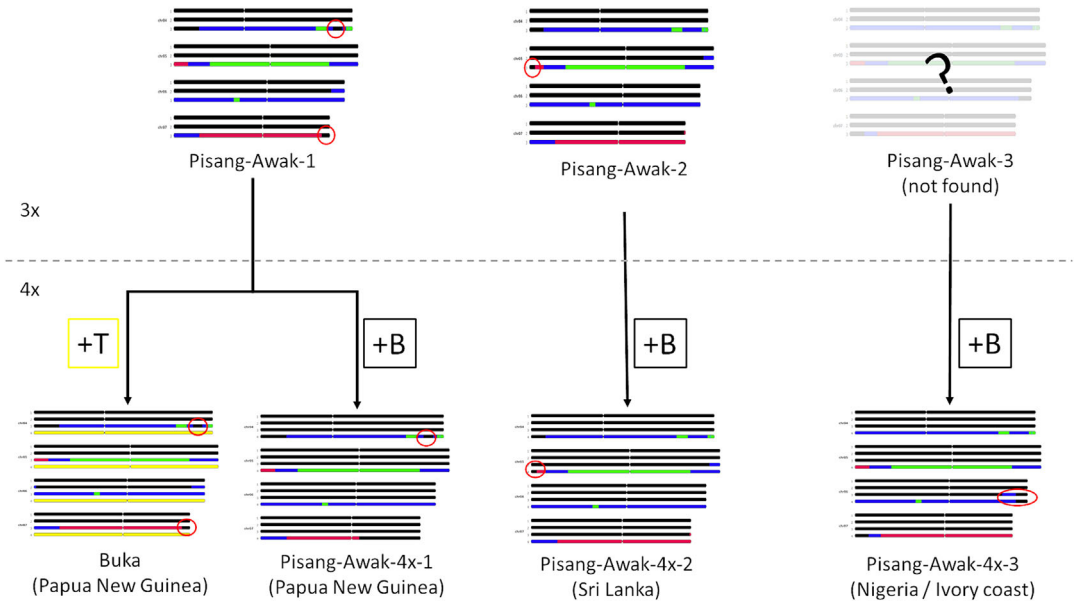

FIGURE 3 Genetic diversification in the Pisang Awak group. This illustration showcases the patterns of five Pisang Awak variants and a Pisang Awak-derived intersectional hybrid for four sets of chromosomes (chromosomes 4, 5, 6, and 7), each marked by distinctive events (red circle) inherited from one of the parents (other differences may result from recombination events that occurred during the production of unreduced (3x gametes). The integration of additional genomes into the triploid (3x) patterns (Pisang Awak-1/3), denoted by specific letters (B for balbisiana, represented in black; T for Australimusa, shown in yellow), has resulted in the formation of closely related tetraploid (4x) varieties. The triploid Pisang Awak-3 is represented with partial transparency as it was not found in the sample but could be inferred from its progeny. Colored segments show the contribution of each of the ancestral genepools.

(three accessions) and Kunnan (four accessions), but the correspondence between the morphological assignation and the patterns was incomplete. Because these cultivar groups were not extensively documented and many synonyms and overlaps exist in India, the true assignation of the accession confusingly named "Kunnan" (ITC1034) but classified as Ney Poovan was difficult to assert.

3.3 | Other cultivars (clusters and individual accessions)

Some accessions of the set were ambiguous in classification, with morphological similarity with well-known cultivar groups, but not complying enough to discriminating criteria to be considered as part of these defined cultivar groups.

3.3.1 | Similar to Mutika/Lujugira

Four accessions were collected in Tanzania with unclear classifications (De Langhe et al., 2001) and comprise two distinct mosaic patterns that share similarities with the Mutika/Lujugira group. Notably, they include a significant malaccensis component (12 centromeres and 6 centromeres, respectively), similar to what was observed in the Ilalyi group (Table 2). The first pattern, Kikundi, is identified in three accessions. "Ntebwa" and "Ntindi I," both from Tanzania's Usambara region, differ in their uses; "Ntebwa" is used for cooking, and "Ntindii I" serves both as a cooking (flour) and dessert banana. The third accession, "Kikundi," differs by the pinkish color of the pseudostem contrasting with the green observed in the two others. The second pattern, Luholole, is represented by a single accession from the Morogo district of Tanzania.

3.3.2 | Similar to Plantain

Two accessions can be linked morphologically to the Plantain group. The accession "Kupulik" was collected in the late 1980's in Papua New Guinea (Island of New Ireland in the Bismark Archipelago) as a Horn type Plantain, but it is not a Plantain. Its mosaic shares similarities with both Plantain and Iholena, but the malaccensis component is absent in "Kupulik," like two patterns of the Maia Maoli/Popoulu (Figure 2). Two other accessions originating from Papua New Guinea, "Bubun" and "Navente 2," exhibited the same pattern. Then, the cultivar "Mnalouki" from the Comoros (Perrier et al., 2019) shares two haplotypes with Plantain cultivars and was proposed to be a progeny of a Plantain (2x gamete) × Mchare (1x gamete) (Martin, Baurens, et al., 2023).

3.3.3 | Similar to Iholena

Some level of morphological confusion exists around the Iholena group (Arnaud & Horry, 1997; Kagy et al., 2016; Sachter-Smith et al., 2021; Sachter-Smith & Sardos, 2021a). In our set, five accessions sharing some, but not all, morphological features of the Iholena exhibited two different and distinct mosaic patterns. The accessions "Rukumamb Tambey," "Tigua," and "Balabolo 1" form the Rukumamb Tambey cluster. They share the bunch shape and the color of the flesh with Iholena, but their fruits do not turn yellow when ripe, and the lower surface of their new leaf is green. The accessions "Arawa" and "Mamae Upolu" displayed a second mosaic pattern and were also different in their morphology (notably "Arawa," which had an overall more diploid look at collect). "Mamae Upolu," collected in Samoa, differs from Iholena by its slightly more upward fruits, the green lower surface of the new leaf, and the more yellow color of the flesh. Despite their morphological proximity, these two mosaic patterns differ from Iholena by the ancestry of five and six centromeres, respectively, and the notable presence of one malaccensis centromere that is absent in Iholena (Table 2).

3.3.4 | Similar to Maia Maoli/Popoulu

We observed six banana accessions that morphologically resemble the Maia Maoli/Popoulu (MMP) cultivars but with three different patterns. They were all collected in Papua New Guinea and surrounding islands and are composed of three different mosaic patterns that share a common background with the four MMP patterns previously identified and with other AAB cooking bananas. The first pattern, named here Wan Gevi, is composed of two accessions. The second pattern is represented by a unique accession, "Buka Kiakiau." Unlike the two other patterns, the third pattern, named Ruango Block and made of three accessions, contains malaccensis as a contributor (four centromeres).

3.3.5 | Clusters of accessions with distinctive morphotypes

We noted in our set some clusters of accessions that may correspond to morphological variants of the same genomic pattern. It was notably the case of three sets of diploid accessions. The first set, composed of three accessions from the Philippines and Malaysia, was called here the Bata-Bata cluster. The second set was composed of four accessions from Papua New Guinea and was named here the Te'engi cluster. Then, the Talasea cluster was composed of two accessions collected in the Papua New Guinea outer islands, one being likely the reddish variant of the other. Two morphological variants were also observed in the ABBT Buka cluster, "Buka" being a shorter cultivar than "Bukayawa."

3.4 | Individual profiles

Finally, we listed in the catalog individual accessions that cannot be morphologically assigned to described cultivar groups and that have specific genomic patterns (Table S3). Several diploids of AA genomic

compositions presented interesting characteristics. The cultivar "Khai Na On" is the male (1x gamete) parent of the "Gros Michel" group (Hippolyte et al., 2012; Martin, Cottin, et al., 2023; Raboin et al., 2005). The cultivar "Pisang Madu" is the keystone to the genome ancestry mosaic painting approach as it enabled the discovery and the identification of diagnostic SNPs for two uncharacterized ancestors of cultivated bananas, and it is related to the Cavendish group. Its genome is indeed composed of a full haplotype of the unknown ancestor m1, proposed to be halabanensis from Indonesia (Martin, Cottin, et al., 2023). Additionally, it also contains six centromeres belonging to the unknown ancestor (Martin, Cottin, et al., 2023; Sardos et al., 2022). "Manang," an accession from the Philippines, has a significant contribution from burmannica (four centromeres), the same as "Matti" from India, also included in the catalog with nearly a full haplotype of burmannica.

For the triploids, the "Chuoi Mit" cultivar from Vietnam exhibits an ABB genomic composition, featuring an A genome that closely resembles the A genome found in the Bluggoe/Monthan pattern. The proximity of these cultivar groups has already been reported based on homoeologous exchanges between A and B subgenomes (Cenci et al., 2021) and is now confirmed analyzing the A mosaic pattern. "Pisang Slendang," an Indonesian AAB cultivar, shares one of its A haplotypes with both the Bluggoe and Monthan groups, as well as with "Chuoi Mit." "Chuoi Xi Mon" bears an unidentified genome differing from the characterized unknown genepool.

The cultivar "Auko" (synonym "Vunapope") from Papua New Guinea, with an ABB genomic composition, has a unique A mosaic pattern among all cultivated bananas (no zebrina ancestry). It is only made of banksii and schizocarpa. This finding supports the hypothesis of the early domestication of banana in New Guinea involving both *M. a*. ssp *banksii* and *M. schizocarpa* originated (Carreel et al., 2002), which are native from New Guinea (Martin, Cottin, et al., 2023). On the other side of the spectrum, the cultivar "La" from Vietnam with an AAB genomic composition has 14 zebrina centromeres and only one small schizocarpa introgression. The "Ya Ta Na Thin kha" accession collected in Myanmar was provided with a poor classification (*Musa*) and was identified as a triploid ABB in our analysis. Its genomic composition is rich in burmannica, like the Klue Teparod group. However, the pattern of "Ya Ta Na Thin kha" is different, notably with a small introgression of zebrina in the first arm of chromosome 9 that is absent in Klue Teparod, which exhibits a balbisiana introgression in the same region. These two patterns shared recombination breakpoints, indicating common evolutionary history, even suggesting that "Ya Ta Na Thin kha" could be one of the genotypes at the origin of this cultivar group (Figure 4). However, in the absence of morphological description available, we were not able to assess whether "Ya Ta Na Thin kha" belongs to the Klue Teparod group.

Until recently, only one triploid cultivar with a full S haplotype was known. The cultivar "Toitoi" was collected on the island of Bougainville in Papua New Guinea. Its unrecombined schizocarpa haplotype suggested it resulted from an unreduced (2x) gamete with an AA genomic composition like many AA diploids cultivated in the country and a regular (1x) S gamete, probably from a wild specimen of *M. schizocarpa* endemic to New Guinea. A second cultivar with an AAS genomic composition named "Waga" was found since then, still in Papua New Guinea, and has a different mosaic pattern that interestingly shows two *M. acuminata* ssp. *banksii* introgressions on the chromosome 4 of the *M. schizocarpa* haplotype that probably results from a different type of cross (Dataset S2). Several tetraploid cultivars with a T genome were discovered in Papua New Guinea. However, the SNPs assigned to T are representative of the whole former Australimusa section of the *Musa* species, from which arose the Fei (Fe'i or Fehi) bananas, without providing any specific insight into the exact species or genepool involved. Among these tetraploid cultivars, "Kalmagol," which morphologically resembles the Silk group, had a genomic composition of AABT with the AAB genome that could correspond to the pattern Silk-2. Among the three patterns with an ABBT genomic composition, the cultivars "Buka" and "Bukayawa" displayed an ABB genome that could have derived from the PA-1 pattern (Figure 3), while the cultivar "Bengani" likely derived from the pattern Kalapua-2. In Cook Island, "Rekua," a second ABBT cultivar with an ABB genome like the Kalapua patterns, was discovered. It is different from "Bengani," and its ABB genome may have arisen from Kalapua-1.

4 | DISCUSSION

4.1 | Clonal diversification at the genomic scale

Domestication of bananas was a gradual process in which the selection for edible pulp led to today's parthenocarpic and highly sterile cultivars (Simmonds, 1966). In this scenario, cultivar groups were expected to correspond to cultivars clonally derived from each other with the clonal accumulation of both point mutations and epigenetic variations as main mechanisms of diversification (Simmonds, 1966). For example, in West and Central Africa and in East Africa, respectively, despite the extremely high levels of intra-cultivar group phenotypic diversity observed in the Plantain and Mutika/Lujugira groups (Karamura et al., 1998; Tézenas Du Montcel et al., 1983), they were both found genetically homogenous (Kitavi et al., 2016; Noyer et al., 2005), but with significant levels of epigenetic variations (Kitavi et al., 2020; Noyer et al., 2005). Here, we identified larger genomic variations in these cultivar groups, fixed through vegetative propagation. Specifically, deletions were observed in the Mutika-3 and Plantain-2 patterns, and another event interpreted as a homologous exchange due to mitotic recombination was inferred in the Mutika-2 and Mutika-4 patterns. The patterns Plantain-2 and Mutika-2 are in accordance with the findings of Martin, Cottin, et al. (2023) for these two cultivar groups, but the high number of accessions analyzed in our set enabled the discovery of other variants. In the Plantain group, the deletion, frequently observed in African accessions, is absent in the Asian accessions, suggesting that this event occurred after the introduction of the first Plantain cultivar(s) in Africa. This supports previous hypotheses (De Langhe et al., 2015; Perrier et al., 2011) on the origin of Plantain in Southeast Asia. For the Mutika/Lujugira cultivars studied here, the three genomic variants characterized on

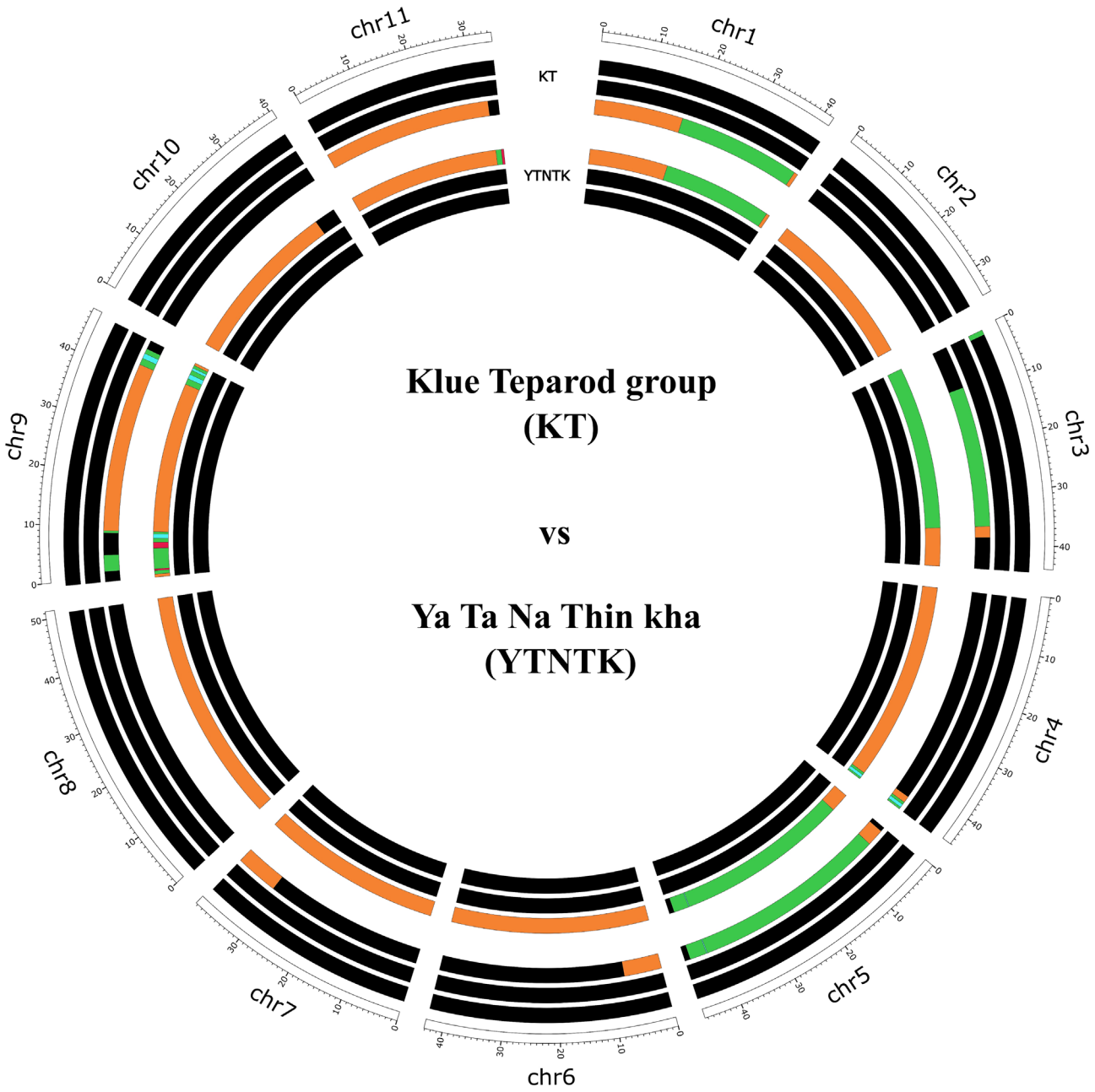

FIGURE 4 Genome ancestry mosaic painting applied to the Klue Tiparod (ABB, 3x) and "Ya Ta Na Thin kha" (collected as *Musa* spp), revealing related pattern and a possible pedigree relationship. The colors of segments correspond to ancestral contributions (black: *Musa balbisiana*, green: *Musa acuminata* banksii genetic group, and orange: *M. a. burmannica* including ssp. siamea).

chromosome 10 were found in only a small portion of the samples. However, most of the Mutika/Lujugira analyzed were introduced to the ITC from Rwanda and Burundi and likely do not represent the entire diversity of this cultivar group. Further genomic characterization of the Mutika/Lujugira germplasm across a wider geographical range is recommended for future studies.

Within those two cultivar groups, we did not observe obvious correlations between the small genomic variations detected and the striking phenotypic features of these cultivar groups, a pattern also observed with epigenetic variations (Kitavi et al., 2020; Noyer et al., 2005). However, these events constitute valuable markers for tracing the evolutionary history of Plantain and Mutika/Lujugira. Additionally, these genomic variations notably generate gene copy number variations, as identified in multiple crops (Gabur et al., 2019; Stein et al., 2017; Yakushiji et al., 2006), and may be linked to interesting traits, such as disease resistance, as identified in a few somaclonal variants of Cavendish exhibiting deletions on chromosome 5 (Hou et al., 2022). Therefore, we speculate that the specific regions of chromosome 10, where structural variations were identified in the Mutika/Lujugira and Plantain groups, merit further investigation.

4.2 | Sexuality still matters in cultivated bananas

Most edible bananas do not produce viable seeds, either in cultivation or in the wild, unless female flowers are heavily and artificially pollinated (De Langhe, 2009). However, somatic mutation in bananas may not have been the only source of intra-cultivar variations. The occasional occurrence of seeds in the fruits of some cultivars has been documented, especially in Papua New Guinea (Arnaud & Horry, 1997) (see Figure S2). The diversity of genome patterns observed within several cultivar groups, including well-known cultivar groups such as Silk, Pome, Maia Maoli/Popoulu, and Pisang Awak, seemed difficult to explain only by clonal diversification. The variations observed, with centromeres of different origins and/or the accumulation of high numbers of recombination, rather support a meiotic origin to these differences. Despite prolonged asexual reproduction and selection for seedlessness, bananas have maintained some residual sexual reproductive capacity, such as *Ensete ventricosum* in its sister genus (Denham et al., 2020; Tamrat et al., 2022). Furthermore, similarly to what was inferred between cultivars in grape (Myles et al., 2011), some of the different patterns observed within these cultivar groups might correspond to siblings from the same parents. This important finding supports Kagy et al. (2016), who proposed an enlarged vision of cultivar groups and considered that sets of closely related cultivars could ensue from different sexual events within similar or closely related parents. In addition, the Pisang Awak example, with the occurrence of tetraploid siblings of triploid cultivars, illustrates that original clones can also be sources of sexual diversification within cultivar groups (Figure 3). Pisang Awak are more prone to set seeds (Simmonds, 1962) (Figure S2), and the tetraploid cultivars observed here show that farmers continue to select new banana cultivars that arise accidentally from seeds, further contributing to expanding diversity as observed in grapes (Bouby et al., 2013), date palms (Gros-Balthazard et al., 2016), and other vegetatively propagated crops (Zohary & Spiegel-Roy, 1975).

4.2.1 | Progressive genepool introgression and secondary diversification

Interestingly, several of the accessions not affiliated with any cultivar group provided valuable insights by revealing the progressive incorporation of exotic genepools. This is particularly notable in East Africa, where the Ilayi group, along with Kikundi and Luhohole patterns, share a common genomic background with the Mutika/Lujugira group. These groups are characterized by an important contribution of banksii and zebrina but supplemented by malaccensis. Such pattern was also observed in "Mnalouki," when compared with Plantain, as well as in the Ruango Block cluster when compared with Maia Maoli/Popoulu and, to some extent, in the Rukumamb Tambey and Arawa clusters when compared with Iholena. Considering that banana's domestication center arose in New Guinea island (Martin, Cottin, et al., 2023) where only *M. acuminata* ssp. *banksii* and *M. schizocarpa* can be found, the genomic makeup of "Auko," lacking the zebrina genepool, supports that the addition of zebrina to the genomic backgrounds of cultivars likely resulted from secondary diversification events. However, the common contribution of zebrina to all other cultivars suggests that the addition of zebrina precluded the insertion of malaccensis in banana cultivars. This scenario is consistent with the geographic distribution of *M. acuminata* subspecies and is in line with the correlation observed between the wild subspecies geographical ranges and the wild ancestors' contribution to local cultivars (Martin, Cottin, et al., 2023; Sardos et al., 2022). Therefore, these accessions related to known cultivar groups but with the additional contribution of malaccensis may result from more recent sexual diversification, such as "Mnalouki," possibly a sibling of Plantain (Martin, Baurens, et al., 2023).

In addition, the incorporation of a supplementary genepool (with a T haplotype) into cultivars from the Silk, Pisang Awak, and Kalapua groups was also observed in tetraploid accessions that were collected in the Pacific. Interestingly, the Silk and Pisang Awak groups originated in India and Southeast Asia, respectively, while wild and cultivated specimens of the former Australimusa section can be found only in an area going from Sulawesi (eastern Indonesia) to the Pacific Islands. This suggests that these hybridizations occurred more recently, after the introduction of these cultivar groups in the distribution range of the ex-Australimusa specimens. This illustrates the importance of conserving local genetic resources, as they can still play an active role in crop diversification.

4.3 | Implications for farming system, taxonomy, breeding, and conservation

4.3.1 | Farming systems

In clonal crops, vegetative propagation is an efficient way to preserve and multiply favorable genotypes that would not be maintained through sexual reproduction (McKey et al., 2010b). In bananas, selection for seedlessness, hybridization, and polyploidy have put populations of this crop onto the road of purely clonal evolution (McKey et al., 2010a), and farmers have historically selected and preserved clonal varieties upon noticing changes in traits in the field (Karamura et al., 2010). In other clonal crops, such as cassava (Elias et al., 2000), yams (Scarcelli et al., 2006), or Enset (Shigeta, 1996), farming practices enabling the incorporation of seedlings into the cultivated stocks have been documented. In banana, it was previously assessed that diversification also occurred through residual sexual events among cultivars (De Langhe et al., 2010; Martin, Baurens, et al., 2023, this study), and the occasional occurrence of seeds in some cultivars under traditional farming systems has been recorded (McKey et al., 2010b). However, to our knowledge, no description of farming practices enabling the capture of seedlings exists. The results presented here therefore highlight the need to investigate farming practices in selected areas where more traditional forms of agriculture still exist to understand how sexual reproduction contributes to diversification.

The co-existence of genomic variations with two origins, clonal and sexual, in the overall diversity of cultivated bananas has implications in a context where monoclonal agriculture puts banana

cultivation at risk in the face of biotic and abiotic stresses. These variations are valuable sources of diversity that can be overlooked by farmers. As stated before, deletions and duplications are sources of gene copy number variations that can result in differential phenotypes, just as homologous exchanges do. The use of this intra-cultivar group diversity can be an innovative way to diversify agrosystems by ensuring the co-existence of clonal and sexual variants in farmer cultivar portfolios. Because cultivar adoption is affected by a combination of sensory characteristics, agronomic properties, and environmental and socio-cultural factors affecting production (Madalla, 2021), planting different genomic patterns associated with the same cultivar group could allow to overcome part of these constraints while introducing diversity in farmers' fields. Equally, the cultivars sharing morphological characteristics with known cultivar groups and which were found hybridized or introgressed with exotic genepools could allow the introduction of additional genetic diversity, hence with putative beneficial new traits, while enhancing the likelihood of acceptance by farmers.

4.3.2 | Classification of cultivated bananas

This study shows that intra-cultivar group diversification is made of a combination of sexual and clonal diversification and pleads for a relaxed definition of the cultivar group concept as the set of closely genetically related individuals sharing peculiar morphological characteristics. Classification criteria should be revised combining morphological assessment and in silico chromosome painting results, for example, using genomic determination keys (Figure S3). Besides, classification would benefit from considering indigenous traditional knowledge from farmers, for example, in incorporating local uses in the discriminant characteristics such as the beer cultivar types among Mutika/Lujugira (Karamura, 1999).

Our findings also raise new questions about the current taxonomy. For example, it may be debatable to consider Bluggoe and Monthan as two separate cultivar groups while they share seemingly identical genomic backgrounds. From a classification point of view, while Bluggoe and Monthan have some morphological differences that initially justified their separation into separate groups, this situation may be considered as in other groups, such as Plantain or Mutika/Lugugira, which are single groups despite their significant intra-group phenotypic diversity. Equally, the scientific community interested in banana classification and taxonomy should consider and agree on the creation of new cultivar groups as (clusters of) new genomic patterns are discovered. In some cases, cultivar groups may also be enlarged in a way they would incorporate the "grey zones" of cultivars resembling the "core" accessions of the cultivar groups but differing in their genomic patterns.

In such revision effort, some accessions would remain alone, constituting de facto cultivar groups with only one cultivar, this being relative to the current sample and possibly revisited with the addition of new cultivars. The catalog presented here (Dataset S1), conceived as a dynamic and evolutive document, constitutes a valuable supporting tool for this task.

4.3.3 | Breeding requires to maximize diversity

The chromosome painting approach was shown to be useful to understand the formation of current varieties (Martin, Baurens, et al., 2023), an essential point to ease successful breeding. It could also be a useful tool for breeders, both to support the selection of parents and the selection of hybrids (Cenci et al., 2023). The catalog presented here and the association of the patterns discovered in active genebank accessions open the door to an optimized use of banana diversity for breeding crosses. The occurrence of numerous individual accessions that cannot be affiliated to existing cultivar groups and that display unique and peculiar genomic backgrounds presents a fresh perspective for the use of original accessions as parents. Additionally, the intra-cultivar group diversity could also be a valuable resource for breeding. The selection of parents among the different clones or variants within a given cultivar group could allow the incorporation of new genetic variation in existing breeding schemes and may result in the incorporation of possible useful traits in the obtained progenies.

4.3.4 | Let us keep characterizing and collecting

The characterization of ancestral origin along chromosomes (mosaic genomes) was proven efficient to support germplasm characterization (Ahmed et al., 2019; Martin et al., 2020; Santos et al., 2019; Wu et al., 2021). The present study is aimed at providing an efficient tool to support the community of banana researchers and workers in the task of classifying cultivars. This tool was also found helpful for the resolution of taxonomical issues that may arise. In this study, we found that about 25% of the taxonomic information displayed in the existing passport data required corrections, or clarifications, when details were missing (Table S1). This catalog also constitutes a tool to support the routine management of banana genebanks through molecular characterization. Combining high-throughput genotyping with this tool constitutes a much faster way to classify germplasm when compared with morphological characterization. Additionally, creating a baseline by genotyping germplasm that enters collections can help track in vitro-induced aneuploids (found here but data not shown), synonyms, and potential duplicates.

In this study, we could not characterize the entire banana genepool. For example, the chromosome painting results obtained for the many unique diploid cultivars from Papua New Guinea are not presented here but may be subject to a separate study. Also, the characterization of the popular Saba cultivars from the Philippines was not conducted because of a lack of material available in the genebank. Nevertheless, this catalog constitutes a baseline that can be further enriched as prospections and genomic characterization continue. An online version (https://www.crop-diversity.org/mgis5/cultivar-group-catalogue) aims to be dynamic, continually expanded, and updated as characterizations of the ancestral gene pools and sequencing technologies improve.

Importantly, we provided evidence for the richness of patterns identified in a small number of accessions and that much more

diversity exists in regions that have not been explored or have been underexplored. Screening only a few new accessions from recent collecting missions was sufficient to discover new genetic profiles that did not match defined cultivar groups, maybe justifying the creation of new cultivar groups or blurring the lines of the defined ones. Now that germplasm characterization has entered the area of genomics, it is more than likely that further prospection of banana diversity in farmers' fields will enable the discovery of new variants and new genotypes. Obviously, gaps remain in our perception of banana diversity, and additional collecting missions are necessary to enrich our understanding of banana diversity and to fill the gaps in the collections.

AUTHOR CONTRIBUTIONS

Conceptualization: Julie Sardos, Alberto Cenci, and Mathieu Rouard. *Provided material*: Ines Van den Houwe, Janet Paofa, William Wigmore, and David Tilafono Hunter. *Library preparation and GBS genotyping*: Ronan Rivallan. *Passport data curation*: Julie Sardos, Christophe Jenny, Rachel Chase, Max Ruas, Yaleidis Mendez, Gabriel L. Sachter-Smith, and Mathieu Rouard. *Methodology and bioinformatic analysis*: Guillaume Martin and Catherine Breton. *Mosaic analyses and curation*: Alberto Cenci, Julie Sardos, and Mathieu Rouard. *Catalog design*: Mathieu Rouard and Julie Sardos. *Online catalog development*: Valentin Guignon. *Data release in MGIS/Gigwa*: Catherine Breton and Mathieu Rouard. *Writing—original draft preparation*: Julie Sardos and Mathieu Rouard. *Writing—review and editing*: Xavier Perrier, Guillaume Martin, Nabila Yahiaoui, Christophe Jenny, Rachel Chase, Gabriel L. Sachter-Smith, Alberto Cenci, and Nicolas Roux. *Funding acquisition*: Angélique D'Hont, Nicolas Roux, and Mathieu Rouard. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS

We thank the International Transit Center for providing banana samples, BGI and LGC for their services for the RAD sequencing, and the AGAP Institute's genotyping platform for the GBS. We acknowledge the CRP Roots, Tubers and Bananas and the Genebank Initiative of the CGIAR for the funding of the genotyping experiments, as well as the Crop Trust for their support of the collecting missions. This work was technically supported by the CIRAD—UMR AGAP HPC Data Centre of the South Green Bioinformatics platform. We are grateful to the taxonomic information provided by the Taxonomy Advisory Group of MusaNet (https://musanet.org/who-we-are/taxonomic-advisory-group) for providing on a regular basis critical insight to verify the true-to-typeness of the accessions.

We dedicate this work to two eminent banana scientists who passed away during the preparation of this manuscript. Professor Edmond De Langhe devoted his life and passion to bananas until the end and leaves an unvaluable contribution to our understanding of this crop. Dr. Hughes Tezenas du Montcel, a renowned expert in banana research, significantly enriched the field through his extensive collection of banana plants and laid the groundwork for the Musa Germplasm Information System. This work could not have been possible without their dedication to banana germplasm collection and conservation.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The sequencing reads were deposited in the NCBI SRA associated with the BioProject PRJNA450532. SNP datasets were recorded in a database (dbname: Catalogue_Musa_Mosaic) browsable via a web application https://gigwa.cgiar.org/gigwa (Rouard et al., 2022; Sempéré et al., 2019).

ORCID

Julie Sardos https://orcid.org/0000-0001-5505-9198
Alberto Cenci https://orcid.org/0000-0002-4390-5948
Guillaume Martin https://orcid.org/0000-0002-1801-7500
Catherine Breton https://orcid.org/0000-0002-3389-1861
Valentin Guignon https://orcid.org/0000-0003-0903-6811
Gabriel L. Sachter-Smith https://orcid.org/0000-0003-1007-0873
Rachel Chase https://orcid.org/0000-0002-5859-6455
Nabila Yahiaoui https://orcid.org/0000-0002-4454-3926
Christophe Jenny https://orcid.org/0000-0001-6165-6888
Xavier Perrier https://orcid.org/0000-0002-0092-003X
Mathieu Rouard https://orcid.org/0000-0003-0284-1885

REFERENCES

Adheka, J. G., Dhed'a, D. B., Sivirihauma, C., Karamura, D., Langhe, E., Swennen, R., & Blomme, G. (2013). Plantain collection and morphological characterization in Democratic Republic of Congo: Past and present activities and prospects. In *Banana systems in the humid highlands of sub-Saharan Africa* (Vol. 1). CABI. https://doi.org/10.1079/9781780642314.0001

Agarwal, M., Shrivastava, N., & Padh, H. (2008). Advances in molecular marker techniques and their applications in plant sciences. *Plant Cell Reports*, *27*, 617–631. https://doi.org/10.1007/s00299-008-0507-z

Ahmed, D., Comte, A., Curk, F., Costantino, G., Luro, F., Dereeper, A., Mournet, P., Froelicher, Y., & Ollitrault, P. (2019). Genotyping by sequencing can reveal the complex mosaic genomes in gene pools resulting from reticulate evolution: A case study in diploid and polyploid citrus. *Annals of Botany*, *123*, 1231–1251. https://doi.org/10.1093/aob/mcz029

Arnaud, E., & Horry, J.-P. (1997). *Musalogue: A catalogue of Musa germplasm-Papua New Guinea collecting missions 1988–1989*. IPGRI.

Baurens, F.-C., Martin, G., Hervouet, C., Salmon, F., Yohomé, D., Ricci, S., Rouard, M., Habas, R., Lemainque, A., Yahiaoui, N., & D'Hont, A. (2019). Recombination and large structural variations shape interspecific edible bananas genomes. *Molecular Biology and Evolution*, *36*, 97–111. https://doi.org/10.1093/molbev/msy199

Belser, C., Baurens, F.-C., Noel, B., Martin, G., Cruaud, C., Istace, B., Yahiaoui, N., Labadie, K., Hřibová, E., Doležel, J., Lemainque, A., Wincker, P., D'Hont, A., & Aury, J. M. (2021). Telomere-to-telomere gapless chromosomes of banana using nanopore sequencing. *Communications Biology*, *4*, 1047. https://doi.org/10.1038/s42003-021-02559-3

Bouby, L., Figueiral, I., Bouchette, A., Rovira, N., Ivorra, S., Lacombe, T., Pastor, T., Picq, S., Marinval, P., & Terral, J. F. (2013). Bioarchaeological insights into the process of domestication of grapevine (*Vitis vinifera* L.) during Roman times in Southern France. *PLoS ONE*, *8*, e63195. https://doi.org/10.1371/journal.pone.0063195

Breton, C., Cenci, A., Sardos, J., Chase, R., Ruas, M., Rouard, M., & Roux, N. (2022). A protocol for detection of large chromosome variations in

banana using next generation sequencing. In J. Jankowicz-Cieslak & I. L. Ingelbrecht (Eds.), *Efficient screening techniques to identify mutants with TR4 resistance in banana: Protocols* (pp. 129–148). Springer. https://doi.org/10.1007/978-3-662-64915-2_9

Byabachwezi, M. S. R., Mkulila, S., & Karamura, D. A. (2005). *Tanzania Musa expedition 2005: Morogoro-Mbeva region, December 2005*. INIBAP.

Carreel, F., de Leon, D. G., Lagoda, P., Lanaud, C., Jenny, C., Horry, J. P., & du Montcel, H. T. (2002). Ascertaining maternal and paternal lineage within *Musa* by chloroplast and mitochondrial DNA RFLP analyses. *Genome*, *45*, 679–692. https://doi.org/10.1139/g02-033

Carreel, F., Fauré, S., De León, D. G., Lagoda, P. J. L., Perrier, X., Bakry, F., du Montcel, H. T., Lanaud, C., & Horry, J. P. (1994). Evaluation de la diversité génétique chez les bananiers diploïdes (Musa sp). *Genetics Selection Evolution*, *26*, S125. https://doi.org/10.1186/1297-9686-26-S1-S125

Cenci, A., Martin, G., Breton, C., D'Hont, A., Yahiaoui, N., Sardos, J., & Rouard, M. (2023). Visualizing and inferring chromosome segregation in the pedigree of an improved banana cultivar (gold finger) with genome ancestry mosaic painting. *Horticulturae*, *9*, 1330. https://doi.org/10.3390/horticulturae9121330

Cenci, A., Sardos, J., Hueber, Y., Martin, G., Breton, C., Roux, N., Swennen, R., Carpentier, S. C., & Rouard, M. (2021). Unravelling the complex story of intergenomic recombination in ABB allotriploid bananas. *Annals of Botany*, *127*, 7–20. https://doi.org/10.1093/aob/mcaa032

Chase, R., Sardos, J., Ruas, M., Van den Houwe, I., Roux, N., Hribova, E., & Dolezel, J. (2016). The field verification activity: A cooperative approach to the management of the global Musa in-vitro collection at the International Transit Centre. *Acta Horticulturae*, *1114*, 61–65.

Christelová, P., Langhe, E. D., Hřibová, E., Čížková, J., Sardos, J., Hušáková, M., Van den Houwe, I., Sutanto, A., Kepler, A. K., Swennen, R., Roux, N., & Doležel, J. (2017). Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. *Biodiversity and Conservation*, *26*, 801–824. https://doi.org/10.1007/s10531-016-1273-9

Daniells, J., Jenny, C., Karamura, D., & Tomekpe K. (2001). *Musalogue (a catalogue of Musa germplasm): Diversity in the genus Musa*. International Plant Genetic Resources Institute. ISBN: 2-910810-42-9

Davey, J. W., Davey, J. L., Blaxter, M. L., & Blaxter, M. W. (2010). RADSeq: Next-generation population genetics. *Briefings in Functional Genomics*, *9*, 416–423. https://doi.org/10.1093/bfgp/elq031

De Langhe, E. (2009). Relevance of banana seeds in archaeology. *Ethnobotany Research and Applications*, *7*, 271–281. https://doi.org/10.17348/era.7.0.271-281

De Langhe, E., Hřibová, E., Carpentier, S., Doležel, J., & Swennen, R. (2010). Did backcrossing contribute to the origin of hybrid edible bananas? *Annals of Botany*, *106*, 849–857. https://doi.org/10.1093/aob/mcq187

De Langhe, E., Karamura, D., & Mbwana, A. (2001). *Tanzania Musa Expedition 2001*. INIBAP/IPGRI, Future Harvest.

De Langhe, E., Perrier, X., Donohue, M., & Denham, T. P. (2015). The original banana split: Multi-disciplinary implications of the generation of African and Pacific Plantains in Island Southeast Asia. *Ethnobotany Research and Applications*, *14*, 299–312. https://doi.org/10.17348/era.14.0.299-312

Denham, T., Barton, H., Castillo, C., Crowther, A., Dotte-Sarout, E., Florin, S. A., Pritchard, J., Barron, A., Zhang, Y., & Fuller, D. Q. (2020). The domestication syndrome in vegetatively propagated field crops. *Annals of Botany*, *125*, 581–597. https://doi.org/10.1093/aob/mcz212

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

Droc, G., Martin, G., Guignon, V., Summo, M., Sempéré, G., Durant, E., Soriano, A., Baurens, F. C., Cenci, A., Breton, C., Shah, T., Aury, J. M., Ge, X. J., Harrison, P. H., Yahiaoui, N., D'Hont, A., & Rouard, M. (2022). The banana genome hub: A community database for genomics in the Musaceae. *Horticulture Research*, *9*, uhac221. https://doi.org/10.1093/hr/uhac221

Elias, M., Rival, L., & McKey, D. (2000). Perception and management of cassava (*Manihot esculenta* Crantz) diversity among Makushi Amerindians of Guyana (South America). *Journal of Ethnobiology*, *20*, 239–265.

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, *6*, e19379. https://doi.org/10.1371/journal.pone.0019379

Gabur, I., Chawla, H. S., Snowdon, R. J., & Parkin, I. A. P. (2019). Connecting genome structural variation with complex traits in crop plants. *Theoretical and Applied Genetics*, *132*, 733–750. https://doi.org/10.1007/s00122-018-3233-0

Gros-Balthazard, M., Newton, C., Ivorra, S., Pierre, M.-H., Pintaud, J.-C., & Terral, J.-F. (2016). The domestication syndrome in *Phoenix dactylifera* seeds: Toward the identification of wild date palm populations. *PLoS ONE*, *11*, e0152394. https://doi.org/10.1371/journal.pone.0152394

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

Hermanto, C., Sutanto, A., Daniells, J., Riska, H. S. & Hilman, Y. (2014). *Triangle banana exploration report, Central Maluku and Lesser Sunda Islands, Indonesia. 16 February–6 March 2013*. Bioversity International.

Herten, K., Hestand, M. S., Vermeesch, J. R., & Van Houdt, J. K. J. (2015). GBSX: A toolkit for experimental design and demultiplexing genotyping by sequencing experiments. *BMC Bioinformatics*, *16*, 73. https://doi.org/10.1186/s12859-015-0514-3

Higgins, J., Osorio-Guarín, J. A., Olave-Achury, C., Toloza-Moreno, D. L., Enriquez, A., Di Palma, F., Yockteng, R., & De Vega, J. J. (2023). Characterizing subgenome recombination and chromosomal imbalances in banana varietal lineages. *Annals of Botany*, *133*, 349–364.

Hippolyte, I., Jenny, C., Gardes, L., Bakry, F., Rivallan, R., Pomies, V., Cubry, P., Tomekpe, K., Risterucci, A. M., Roux, N., Rouard, M., Arnaud, E., Kolesnikova-Allen, M., & Perrier, X. (2012). Foundation characteristics of edible Musa triploids revealed from allelic distribution of SSR markers. *Annals of Botany*, *109*, 937–951. https://doi.org/10.1093/aob/mcs010

Hou, B.-H., Tsai, Y.-H., Chiang, M.-H., Tsao, S. M., Huang, S. H., Chao, C. P., & Chen, H. M. (2022). Cultivar-specific markers, mutations, and chimerisim of Cavendish banana somaclonal variants resistant to *Fusarium oxysporum* f. sp. cubense tropical race 4. *BMC Genomics*, *23*, 470.

IPGRI. (1996). *Descriptors for banana (Musa spp.)*. International Plant Genetic Resources Institute.

Irish, B., Rios, C., Roux, N., Sardos, J., & Goenaga, R. (2016). *Characterization of the Musa spp*. Taxonomic Reference Collection at the USDA-ARS Tropical Agriculture Research Station. USDA-ARS.

Jarret Robert, L., Gawel, N., Whittemore, A., & Sharrock, S. (1992). RFLP-based phylogeny of Musa species in Papua New Guinea. *Theoretical and Applied Genetics*, *84*, 579–584. https://doi.org/10.1007/BF00224155

Kagy, V., Wong, M., Vandenbroucke, H., Jenny, C., Dubois, C., Ollivier, A., Cardi, C., Mournet, P., Tuia, V., Roux, N., Doležel, J., & Perrier, X. (2016). Traditional banana diversity in Oceania: An endangered heritage. *PLoS ONE*, *11*, e0151208. https://doi.org/10.1371/journal.pone.0151208

Karamura, D., Karamura, E., Tushemereirwe, W., Rubaihayo, P. R., & Markham, R. (2010). Somatic mutations and their implications to the conservation strategies of the East African highland bananas (*Musa* spp.). *Acta Horticulturae*, *879*, 615–622. https://doi.org/10.17660/ActaHortic.2010.879.68

Karamura, D. A. (1999). Numerical taxonomic studies of the East African highland bananas (Musa AAA - East Africa) in Uganda. PhD Thesis. University of Reading, United Kingdom.

Karamura, E., Frison, E., Karamura, D., & Sharrock, S. (1998). Banana production systems in eastern and southern Africa. In *Bananas and food security* (pp. 401–412). INIBAP.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., Caig, V., Heller-Uszynska, K., Jaccoud, D., Hopper, C., Aschenbrenner-Kilian, M., Evers, M., Peng, K., Cayla, C., Hok, P., & Uszynski, G. (2012). Diversity arrays technology: A generic genome profiling technology on open platforms. *Methods in Molecular Biology (Clifton, N.J.)*, *888*, 67–89.

Kitavi, M., Cashell, R., Ferguson, M., Lorenzen, J., Nyine, M., McKeown, P. C., & Spillane, C. (2020). Heritable epigenetic diversity for conservation and utilization of epigenetic germplasm resources of clonal East African Highland banana (EAHB) accessions. *Theoretical and Applied Genetics*, *133*, 2605–2625. https://doi.org/10.1007/s00122-020-03620-1

Kitavi, M., Downing, T., Lorenzen, J., Karamura, D., Onyango, M., Nyine, M., Ferguson, M., & Spillane, C. (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. *TAG. Theoretical and Applied Genetics. Theoretische und angewandte Genetik*, *129*, 547–561. https://doi.org/10.1007/s00122-015-2647-1

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with burrows-wheeler transform. *Bioinformatics (Oxford, England)*, *26*, 589–595. https://doi.org/10.1093/bioinformatics/btp698

Li, L.-F., & Ge, X.-J. (2017). Origin and domestication of cultivated banana. *Ecological Genetics and Genomics*, *2*, 1–2. https://doi.org/10.1016/j.egg.2016.10.001

Madalla, N. (2021). *Farmers traits preferences for improved banana cultivars in Tanzania and Uganda. Introductory paper at the Faculty of Landscape Architecture, Horticulture and Crop Production Science*. Swedish University of Agriculture Science.

Martin, G., Baurens, F.-C., Labadie, K., Hervouet, C., Salmon, F., Marius, F., Paulo-de-la-Reberdiere, N., van den Houwe, I., Aury, J. M., D'Hont, A., & Yahiaoui, N. (2023). Shared pedigree relationships and transmission of unreduced gametes in cultivated banana. *Annals of Botany*, *131*, 1149–1161. https://doi.org/10.1093/aob/mcad065

Martin, G., Cardi, C., Sarah, G., Ricci, S., Jenny, C., Fondi, E., Perrier, X., Glaszmann, J. C., D'Hont, A., & Yahiaoui, N. (2020). Genome ancestry mosaics reveal multiple and cryptic contributors to cultivated banana. *The Plant Journal*, *102*, 1008–1025. https://doi.org/10.1111/tpj.14683

Martin, G., Cottin, A., Baurens, F.-C., Labadie, K., Hervouet, C., Salmon, F., Paulo-de-la-Reberdiere, N., van den Houwe, I., Sardos, J., Aury, J. M., D'Hont, A., & Yahiaoui, N. (2023). Interspecific introgression patterns reveal the origins of worldwide cultivated bananas in New Guinea. *The Plant Journal*, *113*, 802–818. https://doi.org/10.1111/tpj.16086

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*, 10. https://doi.org/10.14806/ej.17.1.200

McCouch, S., Navabi, Z. K., Abberton, M., Anglin, N. L., Barbieri, R. L., Baum, M., Bett, K., Booker, H., Brown, G. L., Bryan, G. J., Cattivelli, L., Charest, D., Eversole, K., Freitas, M., Ghamkhar, K., Grattapaglia, D., Henry, R., Valadares Inglis, M. C., Islam, T., … Rieseberg, L. H. (2020). Mobilizing crop biodiversity. *Molecular Plant*, *13*, 1341–1344. https://doi.org/10.1016/j.molp.2020.08.011

McKey, D., Elias, M., Pujol, B., & Duputié, A. (2010a). Ecological approaches to crop domestication. In *Biodiversity in agriculture: Domestication, evolution, and sustainability* (Vol. 377). Cambridge University Press.

McKey, D., Elias, M., Pujol, B., & Duputié, A. (2010b). The evolutionary ecology of clonally propagated domesticated plants. *New Phytologist*, *186*, 318–332. https://doi.org/10.1111/j.1469-8137.2010.03210.x

Myles, S., Boyko, A. R., Owens, C. L., Brown, P. J., Grassi, F., Aradhya, M. K., Prins, B., Reynolds, A., Chia, J. M., Ware, D., Bustamante, C. D., & Buckler, E. S. (2011). Genetic structure and domestication history of the grape. *Proceedings of the National Academy of Sciences*, *108*, 3530–3535. https://doi.org/10.1073/pnas.1009363108

Noyer, J. L., Causse, S., Tomekpe, K., Bouet, A., & Baurens, F. C. (2005). A new image of plantain diversity assessed by SSR, AFLP and MSAP markers. *Genetica*, *124*, 61–69. https://doi.org/10.1007/s10709-004-7319-z

Pedregosa, F., Varoquaux, G., Gramfort, A., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dobourg, V., Vanderpalst, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Perrier, X., Jenny, C., Bakry, F., Karamura, D., Kitavi, M., Dubois, C., Hervouet, C., Philippson, G., & de Langhe, E. (2019). East African diploid and triploid bananas: A genetic complex transported from South-East Asia. *Annals of Botany*, *123*, 19–36. https://doi.org/10.1093/aob/mcy156

Perrier, X., Langhe, E. D., Donohue, M., Lentfer, C., Vrydaghs, L., Bakry, F., Carreel, F., Hippolyte, I., Horry, J. P., Jenny, C., Lebot, V., Risterutcci, A. M., Tomekpe, K., Doutrelepont, H., Ball, T., Manwaring, J., de Maret, P., & Denham, T. (2011). Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences*, *108*, 11311–11318. https://doi.org/10.1073/pnas.1102001108

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., & Rafalski, A. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding*, *2*, 225–238. https://doi.org/10.1007/BF00564200

Raboin, L.-M., Carreel, F., Noyer, J.-L., Baurens, F. C., Horry, J. P., Bakry, F., Montcel, H. T. D., Ganry, J., Lanaud, C., & Lagoda, P. J. L. (2005). Diploid ancestors of triploid export banana cultivars: Molecular identification of 2n restitution gamete donors and n gamete donors. *Molecular Breeding*, *16*, 333–341. https://doi.org/10.1007/s11032-005-2452-7

Risterucci, A. M., Grivet, L., N'Goran, J. A. K., Pieretti, I., Flament, M. H., & Lanaud, C. (2000). A high-density linkage map of *Theobroma cacao* L. *Theoretical and Applied Genetics*, *101*, 948–955. https://doi.org/10.1007/s001220051566

Risterucci, A.-M., Hippolyte, I., Perrier, X., Xia, L., Caig, V., Evers, M., Huttner, E., Kilian, A., & Glaszmann, J. C. (2009). Development and assessment of diversity arrays technology for high-throughput DNA analyses in Musa. *TAG. Theoretical and Applied Genetics. Theoretische Und Angewandte Genetik*, *119*, 1093–1103. https://doi.org/10.1007/s00122-009-1111-5

Rouard, M., Sardos, J., Sempéré, G., Breton, C., Guignon, V., van den Houwe, I., Carpentier, S. C., & Roux, N. (2022). A digital catalog of high-density markers for banana germplasm collections. *Plants, People, Planet*, *4*, 61–67. https://doi.org/10.1002/ppp3.10187

Ruas, M., Guignon, V., Sempere, G., Sardos, J., Hueber, Y., Duvergey, H., Andrieu, A., Chase, R., Jenny, C., Hazekamp, T., Irish, B., Jelali, K., Adeka, J., Ayala-Silva, T., Chao, C. P., Daniells, J., Dowiya, B., Effa Effa, B., Gueco, L., … Rouard, M. (2017). MGIS: Managing banana (*Musa* spp.) genetic resources information and high-throughput genotyping data. *Database: The Journal of Biological Databases and Curation*, *2017*, bax046.

Sachter-Smith, G. L., Paofa, J., & Sardos, J. (2021). *Bananas of West New Britain - Papua New Guinea*.

Sachter-Smith, G. L., & Sardos, J. (2021a). *Bananas of Cook Islands: A catalog of banana diversity seen on the islands of Rarotonga and Aitutaki*. Bioversity International.

Sachter-Smith, G. L., & Sardos, J. (2021b). *Bananas of Samoa: A catalogue of banana diversity seen on the island of Upolu*. Bioversity International.

Santos, J. D., Chebotarov, D., McNally, K. L., Bartholomé, J., Droc, G., Billot, C., & Glaszmann, J. C. (2019). Fine scale genomic signals of admixture and alien introgression among Asian rice landraces. *Genome Biology and Evolution*, *11*, 1358–1373. https://doi.org/10.1093/gbe/evz084

Sardos, J., Breton, C., Perrier, X., van den Houwe, I., Carpentier, S., Paofa, J., Rouard, M., & Roux, N. (2022). Hybridization, missing wild ancestors and the domestication of cultivated diploid bananas. *Frontiers in Plant Science*, *13*, 969220. https://doi.org/10.3389/fpls.2022.969220

Sardos, J., Christelová, P., Čížková, J., Paofa, J., Sachter-Smith, G. L., Janssens, S. B., Rauka, G., Ruas, M., Daniells, J. W., Doležel, J., & Roux, N. (2018). Collection of new diversity of wild and cultivated bananas (*Musa* spp.) in the Autonomous Region of Bougainville, Papua New Guinea. *Genetic Resources and Crop Evolution*, *65*, 2267–2286. https://doi.org/10.1007/s10722-018-0690-x

Sardos, J., Paofa, J., Janssens, S., Sachter-Smith, G., Rauka, G., & Roux, N. (2017). *Banana collecting mission in the Autonomous Region of Bougainville (AROB), Papua New Guinea*. Bioversity International.

Sardos, J., Paofa, J., Sachter-Smith, G. L., Togabalaguguwa, E., Hribova, E., Van Den Houwe, I., & Roux, N. (2019). *Report from the banana collecting mission to the Province of West New Britain, Papua New Guinea from September 30th to October 12th, 2019*. Bioversity International.

Sardos, J., Perrier, X., Doležel, J., Hřibová, E., Christelová, P., van den Houwe, I., Kilian, A., & Roux, N. (2016). DArT whole genome profiling provides insights on the evolution and taxonomy of edible banana (*Musa* spp.). *Annals of Botany*, *mcw170*, 1269–1278. https://doi.org/10.1093/aob/mcw170

Sardos, J., Sachter-Smith, G., Ghanem, M., Hribova, E., Van Den Houwe, I., Roux, N., & Wigmore, W. (2019). *Report from the banana collecting mission to Rarotonga and Aitutaki, Cook Islands*. Bioversity International.

Sardos, J., Sachter-Smith, G., Shandil, A., Van Den Houwe, I., Hribova, E., Matalavea, P., Vaai, A., Iosefa, T., Hunter, D.T., & Roux, N. (2019). *Report from the banana collecting mission to Samoa* from July 22nd to August 2nd, 2019. Bioversity International.

Scarcelli, N., Tostain, S., Mariac, C., Agbangla, C., da, O., Berthaud, J., & Pham, J. L. (2006). Genetic nature of yams (*Dioscorea* sp.) domesticated by farmers in Benin (West Africa). *Genetic Resources and Crop Evolution*, *53*, 121–130. https://doi.org/10.1007/s10722-004-1950-5

Sempéré, G., Pétel, A., Rouard, M., Frouin, J., Hueber, Y., de Bellis, F., & Larmande, P. (2019). Gigwa v2—Extended and improved genotype investigator. *GigaScience*, *8*. https://doi.org/10.1093/gigascience/giz051

Shepherd, K. (1957). Banana cultivars in East Africa. *Tropical Agriculture (Trinidad)*, *34*, 277–286.

Shepherd, K., & Ferreira, F. (1984). *The PNG biological foundation's banana collection at Laloki, Port Moresby, Papua New Guinea*. Newsletter-Regional Committee for Southeast Asia (IBPGR).

Shigeta, M. (1996). Creating landrace diversity: The case of the Ari People and Ensete (*Ensete ventricosum*) in Ethiopia. In *Redefining nature*. Routledge.

Simmonds, N. W. (1962). *The evolution of the bananas*. Longmans.

Simmonds, N. W. (1966). *Bananas*. Longmans.

Simmonds, N. W., & Shepherd, K. (1955). The taxonomy and origins of the cultivated bananas. *Journal of the Linnean Society of London, Botany*, *55*, 302–312. https://doi.org/10.1111/j.1095-8339.1955.tb00015.x

Smale, M., & Jamora, N. (2020). Valuing genebanks. *Food Security*, *12*, 905–918. https://doi.org/10.1007/s12571-020-01034-x

Stein, A., Coriton, O., Rousseau-Gueutin, M., Samans, B., Schiessl, S. V., Obermeier, C., Parkin, I. A. P., Chèvre, A. M., & Snowdon, R. J. (2017). Mapping of homoeologous chromosome exchanges influencing quantitative trait variation in *Brassica napus*. *Plant Biotechnology Journal*, *15*, 1478–1489. https://doi.org/10.1111/pbi.12732

Stover R. H., & Simmonds, N. W. (1987). *Bananas*. Longman.

Summo, M., Comte, A., Martin, G., Perelle, P., Weitz, E. M., Droc, G., & Rouard, M. (2022). GeMo: A web-based platform for the visualization and curation of genome ancestry mosaics. *Database*, *2022*, baac057.

Tamrat, S., Borrell, J. S., Shiferaw, E., Wondimu, T., Kallow, S., Davies, R. M., Dickie, J. B., Nuraga, G. W., White, O., Woldeyes, F., Demissew, S., & Wilkin, P. (2022). Reproductive biology of wild and domesticated *Ensete ventricosum*: Further evidence for maintenance of sexual reproductive capacity in a vegetatively propagated perennial crop. *Plant Biology (Stuttgart, Germany)*, *24*, 482–491. https://doi.org/10.1111/plb.13390

Taxonomy Advisory Group (TAG). (2016). *Minimum list of descriptors for Musa*. Musanet.

Tézenas Du Montcel, H., De Langhe, E., & Swennen, R. (1983). Essai de classification des bananiers plantains (AAB). *Fruits*, *38*, 461–474.

Van den Houwe, I., Chase, R., Sardos, J., Ruas, M., Kempenaers, E., Guignon, V., Massart, S., Carpentier, S., Panis, B., Rouard, M., & Roux, N. (2020). Safeguarding and using global banana diversity: A holistic approach. *CABI Agriculture and Bioscience*, *1*, 15. https://doi.org/10.1186/s43170-020-00015-6

Vogel Ely, C., de Loreto Bordignon, S. A., Trevisan, R., & Boldrini, I. I. (2017). Implications of poor taxonomy in conservation. *Journal for Nature Conservation*, *36*, 10–13. https://doi.org/10.1016/j.jnc.2017.01.003

Wu, Y., Lin, F., Zhou, Y., Wang, J., Sun, S., Wang, B., Zhang, Z., Li, G., Lin, X., Wang, X., Sun, Y., Dong, Q., Xu, C., Gong, L., Wendel, J. F., Zhang, Z., & Liu, B. (2021). Genomic mosaicism due to homoeologous exchange generates extensive phenotypic diversity in nascent allopolyploids. *National Science Review*, *8*, nwaa277. https://doi.org/10.1093/nsr/nwaa277

Yakushiji, H., Kobayashi, S., Goto-Yamamoto, N., Tae-Jeong, S., Sueta, T., Mitani, N., & Azuma, A. (2006). A skin color mutation of grapevine, from black-skinned pinot noir to white-skinned pinot blanc, is caused by deletion of the functional VvmybA1 allele. *Bioscience, Biotechnology, and Biochemistry*, *70*, 1506–1508. https://doi.org/10.1271/bbb.50647

Zohary, D., & Spiegel-Roy, P. (1975). Beginnings of fruit growing in the old world. *Science*, *187*, 319–327. https://doi.org/10.1126/science.187.4174.319

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sardos, J., Cenci, A., Martin, G., Breton, C., Guignon, V., Van den Houwe, I., Mendez, Y., Sachter-Smith, G. L., Chase, R., Ruas, M., Rivallan, R., Paofa, J., Wigmore, W., Hunter, D. T., D'Hont, A., Yahiaoui, N., Jenny, C., Perrier, X., Roux, N., & Rouard, M. (2025). Painting the diversity of a world's favorite fruit: A next generation catalog of cultivated bananas. *Plants, People, Planet*, *7*(1), 263–283. https://doi.org/10.1002/ppp3.10581