Check for updates

Received: 29 October 2020 | Revised: 18 January 2021 | Accepted: 22 January 2021

DOI: 10.1002/ppp3.10187

BRIEF REPORT

Plants People Planet PPP Open Access

A digital catalog of high-density markers for banana germplasm collections

Mathieu Rouard[1] | Julie Sardos[1] | Guilhem Sempéré[2,3] | Catherine Breton[1] | Valentin Guignon[1] | Ines Van den Houwe[4] | Sebastien C. Carpentier[4,5] | Nicolas Roux[1]

[1]Bioversity International, Parc Scientifique Agropolis II, Montpellier, France

[2]CIRAD, UMR INTERTRYP, Montpellier, France

[3]INTERTRYP, Univ Montpellier, CIRAD, Montpellier, France

[4]Bioversity International, Leuven, Belgium

[5]Laboratory of Tropical Crop Improvement, Division of Crop Biotechnics, Leuven, Belgium

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

Societal Impact Statement

Global production of bananas, among the top 10 food crops worldwide, is under threat. Increasing the use of germplasm conserved in genebanks is crucial. However, the lack of or difficult access to genetic diversity information limits the efficient utilization of these valuable resources. Here, we present a digital catalog of high-density markers for banana germplasm conserved at the international banana collection. By facilitating access to subsets of genetic diversity information, the catalog has potential to maximize conservation and use of climate-ready varieties and to optimize breeding strategies. The catalog is extendable with data from any banana collection and the software is easily deployable in other crop genebanks.

KEYWORDS

banana, data management, digital genebank, high-density markers, plant conservation

1 | INTRODUCTION

Crop plant diversity managed by genebanks is of great value in the context of the changing needs of agriculture (Smale & Jamora, 2020), but genetic and phenotypic information on this diversity is insufficiently available for most genebanks (McCouch et al., 2013, 2020). The advent of Next Generation Sequencing has enabled—at an ever-decreasing cost—the sequencing of reference genomes of many crops as well as high-density genotyping for large numbers of samples per crop. Genotyping is a powerful tool to help identify gaps or redundancies in germplasm collections, and when combined with phenotyping data, can be used to detect correlations between genome regions and agronomic traits. For some crops, massive sequencing and data processing have been undertaken, as shown in the rice, wheat and barley germplasm collections (Milner et al., 2019; Sansaloni et al., 2020; Wang et al., 2018). These approaches represent increasingly reachable targets for many genebanks worldwide, including the CGIAR international collections (Halewood, Lopez Noriega et al., 2018).

For bananas (*Musa* spp.), the largest ex situ collection is maintained in vitro at one of the CGIAR international genebanks, the International *Musa* Germplasm Transit Centre (ITC), comprised of more than 1,600 accessions (Van den houwe et al., 2020). Then, over 60 national collections worldwide conserve banana diversity and conduct-related research (Figure 1). Bananas (including Plantains) are arguably the world's most important fresh fruit and are a major staple food for hundreds of millions of people in low-income countries. With an estimated world production of 158 million tons annually, the volume of gross banana exports is worth US$12.8 billion to exporting countries (FAOSTAT 2019). Furthermore, most of the global production is by smallholders for their own consumption or for local trade, making it the fourth-most important food crop in the least developed countries (LDCs) as defined by the United Nations, ranked by total production and food consumption.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Author/s, *Plants, People, Planet* © New Phytologist Foundation

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In order to increase understanding of its complex genetics so as to boost crop improvement, the first whole banana genome sequence was released in 2012, for an accession belonging to the *Musa acuminata* species (D'Hont et al., 2012) (Table 1). This original reference has recently been supplemented with of a number of other *Musa* species (Rouard et al., 2018; Wang et al., 2019; Wu et al., 2016). In parallel, high-throughput genotyping methods (i.e., genotyping-by-sequencing (GBS) (Elshire et al., 2011) and restriction-site associated DNA markers (RADSeq) (Davey et al., 2010)) have been used to investigate single nucleotide polymorphisms (SNPs) on various panels of accessions available at the ITC genebank (Cenci et al., 2020; Sardos et al., 2016). In addition, other SNP datasets have been generated from gene expression and proteomics experiments for subsets related to drought tolerance (Cenci et al., 2019; van Wesemael et al., 2019).

FIGURE 1 Diversity of banana bunches at a germplasm collection exhibited at the National Research Centre for Banana (NRCB) in Trichy, India (with genebank curators at the back). Photograph taken by Julie Sardos

While genetic variant information is being produced at a fast pace through various projects and is increasingly processed via standardized bioinformatics workflows, one of the main challenges is the management of an increasing volume of raw and intermediate files that are difficult to handle for many applications. Bioinformatics workflows can produce millions of markers but need to be filtered in multiple ways according to analysis type or user perspective, and working with these data often presents challenges to those without capacity in bioinformatics. Online information systems coping with big data linked to germplasm collections are scarce (König et al., 2020; Mansueto et al., 2017; Raubach et al., 2020; Ruas et al., 2017). Moreover, lack of access to phenotypic information continues to be an additional factor limiting the use of plant genetic resources. Phenotypic data are complex—information on the context under which they were collected is indispensable, and the domain is continuously evolving (Germeier & Unger, 2019). Recognizing these challenges, the availability of easy-to-use, interoperable and flexible solutions to navigate high-density genotyping and phenotyping data online continues to be a key aim for genebanks' delivery of their mission of germplasm documentation and utilization.

In this study, we present an approach used to generate, store and disseminate a catalog of genetic variants of banana and plantain

TABLE 1 An overview of banana

Categories	Description
Geographic origin	South-East Asia and West Oceania
Geographic distribution	Humid tropics and subtropics
Total global production	>158 million tons (FAOSTAT, 2019)
Taxonomy	~75 species and 500–1000 cultivars -
Biology	A giant herb belonging to monocots
	Vegetatively propagated and perennial
	Parthenocarpic and low fertile cultivars
Ploidy	Diploid, triploid and tetraploid
Basic genome information	11 chromosomes[a]
	550–600 million of base pairs
	Approx. 35,000 genes
Common uses	Dessert, cooking, beer, textile, medicine
Nutrition	Rich source of carbs, fiber, potassium, vitamin B6, vitamin C. Some varieties are rich in carotenoids
Main breeding objectives	Drought tolerance, Disease resistance (e.g., Fusarium wilt, Black Leaf streak, banana bacterial wilt (BXW)), Biofortification (e.g., ProVitA), and post-harvest traits (texture, flavor)

[a]Haploid genome of most cultivated bananas and their crop wild relatives in the Eumusa section. Chromosome number can vary $n = 7$ (*M. ingentimusa*), $n = 9$ (*M. becarri*), and $n = 10$ (Callimusa section).

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 List of studies and associated banana accessions available

Purpose of the study/Dataset	Number of accessions (samples)	Collection	Taxonomical coverage	Number of markers	Sequencing technologies	References	Sequence data availability
GWAS panel for parthenocarpy and sterility	106	*ITC*	*M. acuminata* (wild) and AA (cultivated)	7,079,397	Genotyping by Sequencing (GBS)	(Sardos et al., 2016)	PRJNA305234
Genome constitution of ABB, AB	83	*ITC*	ABB, *M. acuminata* and *M. balbisiana*	683,264	RAD sequencing	(Cenci et al., 2020)	PRJNA450532
Genome constitution of AAB	118	*ITC*	AAB, *M. acuminata* and *M. balbisiana*	3,315,168	RAD sequencing	Publication in preparation	PRJNA450532
Panel evaluated for drought tolerance	10 (60)	*ITC*	AAA, AAB, ABB	6,951,307	RNA sequencing	(Cenci et al., 2019)	PRJNA305241
Domestication and acuminata subspecies	254	*ITC*	*M. acuminata* and AA	245,285	RAD sequencing	(Sardos et al., 2021)	PRJNA450532
A/B Structural variations	207	*CRB-PT*	Breeding population	148,329	RAD sequencing	(Baurens et al., 2019)	PRJNA448968 PRJEB28077
Genome ancestry mosaics	25	*CRB-PT, CARBAP*	*M. acuminata*, AA, AAA	191,876	RNA sequencing	(Martin, Cardi, et al., 2020)	SRR956987
Chromosome reciprocal translocations	155	*CRB-PT, ITC*	Breeding populations, *M. acuminata*, AA	120,111	Genotyping by Sequencing (GBS)	(Martin, Baurens, et al., 2020)	PRJNA667853

maintained in the ITC, which is available at https://www.crop-diversity.org/mgis/gigwa and is embedded in the genebank information system through which users can order available germplasm.

2 | MATERIALS AND METHODS

Material used to create the catalog mostly originates from lyophilized leaf tissues of young banana plants distributed by the ITC. Such tissue is the most convenient way to obtain DNA of an acceptable quality and quantity for high-throughput restriction enzyme-associated DNA sequencing methods, as for other omics techniques (Carpentier et al., 2007). Another advantage is that once in stock, the tissues are readily available for distribution, whereas in vitro material takes longer to obtain (i.e., an average of 2 months for proliferating tissues and 4 months for *in vitro* rooted plantlets).

The generated sequence—short reads from Illumina sequencing machines—was processed through bioinformatics workflows composed of open source software that includes quality checks, read mapping on reference genomes, SNP calling and variant effect in genic regions as described in Sardos et al., 2016; Cenci et al., 2020 and Eyland et al., 2020. The outputs of the workflow are enormous text files in the variant call format (VCF). For every accession, another specific file format (i.e., gVCF) containing the full list of variant and non-variant sites is backed up on a server, allowing the system to recall variants with different sampling whenever necessary, thus saving significant time and computing resources.

SNP datasets for the ITC and other collections published in the literature (VCF files) (Table 2) were recorded in a non-relational database browsable via a web application called GIGWA (Sempéré et al., 2019), developed for the purpose of searching large genotyping datasets in an optimized manner. This system, easy to deploy on any platform, is species-agnostic and provides a user-friendly interface to perform advanced data filtering and export for third-party analytical software. It was seamlessly embedded in the *Musa* Germplasm Information System (MGIS https://www.crop-diversity.org/mgis/) (Ruas et al., 2017), the database for global ex situ-held banana genetic resources that provides access to germplasm information for numerous banana collections worldwide, including the ITC.

3 | RESULTS AND DISCUSSION

The diversity of edible bananas has been classified using genome groups according to the relative contribution of their ancestral wild species. Most cultivated bananas derived from hybridization between *Musa acuminata* (A genome) and the *Musa balbisiana* (B genome) species and the most frequent genome combinations are diploids and triploids cultivars denoted: AA, AB, AAA, AAB and ABB. The current catalog of genetic variants spans these species/groups

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

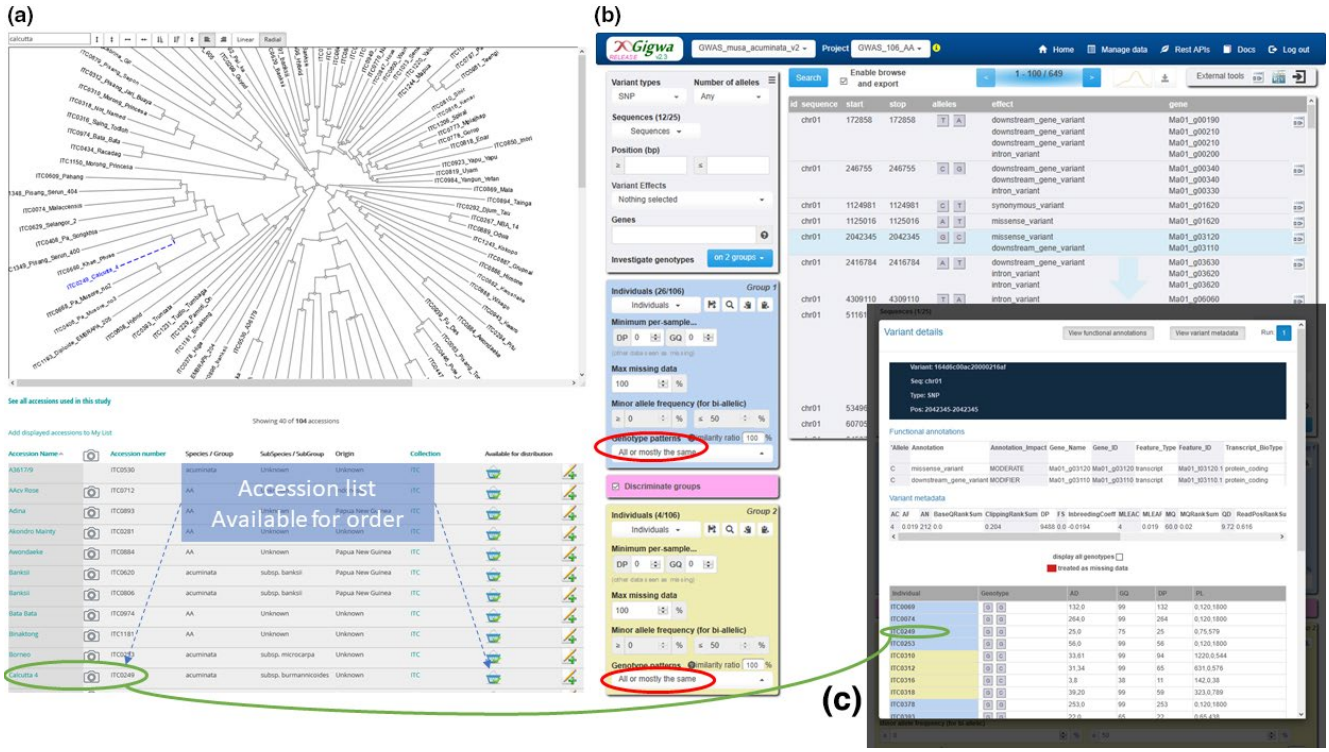

FIGURE 2 Example of a genotyping study in the Musa Germplasm Information System (MGIS). (a) Each study can be opened to access the list of genebank accessions with passport data and a genetic tree. (b) Web interface listing the 649 SNP discriminating between 26 wild banana accessions and cultivated Pisang Jari Buya bananas. (c) Detail view of genetic variants by accession with focus on ITC0249 'Calcutta 4'

for selected subsets of accessions (Table 2). It offers access to datasets with sizes ranging from 245,285 to more than 7 million SNPs depending on the study.

While the system is optimized to explore a large volume of data, it enables efficient filtering options based on a full range of parameters, mostly genetics (e.g., chromosome location, missing data percentage, minor allele frequency, gene mutation effect) but not only. Accession details can be enriched with metadata such as passport data or agronomic traits (e.g., control vs. stress on gene expression analyses), which then become elements which can be filtered. The interface is designed to work with one or two groups of samples, a feature which, when the latter case is used in conjunction with genotype pattern filters, makes it straightforward to identify SNPs discriminating the groups (Figure 2). This is particularly useful to filter by taxonomy or a certain trait between contrasted genotypes to reveal unique alleles held by some accessions. From the user interface, genetic variants of the catalog can be exported in various popular formats (e.g., VCF, BED) for further analyses, or directly imported in other software for genetics analyses. Alternately, content can be programmatically accessed with the Breeding API (BrAPI), a computer–computer programming interface following standard plant specifications (Selby et al., 2019). This solution facilitates essential connections with other information systems (e.g., as implemented in Musabase, the database for banana breeding data https://musabase.org and Genesys, a global information system on Plant Genetic Resources for Food and Agriculture (PGRFA) https://www.genesys-pgr.org).

With regard to types of use of the catalog, it can support various types of research analyses from genetic diversity studies to gene trait association. Of particular interest, a set of SNP markers for a panel of 105 accessions were investigated to provide genebank users with genetic datasets ready for genome-wide association studies (GWAS) once phenotyping data are obtained (Sardos et al., 2016).

The concern that such high levels of genotypic and phenotypic information, associated with germplasm accessions, would enable new breeding techniques (NBTs) that would bypass the access and benefit-sharing (ABS) arrangements linked currently to the distribution of physical material has generated much recent attention (Aubry, 2019; Halewood, Chiurugwi et al., 2018; Smyth et al., 2020). At the moment, any genebank user (e.g., researcher, breeder) can order plants and sequence them without further obligations, and many organizations have already made publicly available such datasets for a wide range of crops. As potential solutions are elaborated (Scholz et al., 2020), an important and challenging crop to breed such as banana should not be ignored, as access to its genetic and phenotypic data may contribute significantly to its progress as a crop (Gaffney et al., 2020).

This catalog intends to provide open access to genomic resources in an equitable way, ultimately benefiting all, including those in low-income countries (Halewood et al., 2017). It should be noted that it does not include gene functions, but is linked to a genome browser from the banana genome hub which contains gene annotation for references banana genomes (Droc

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2013). Nevertheless, given that banana is not a model plant, most gene functional information are inferred by bioinformatics methods (e.g., homology-based prediction methods). Moreover, given the polyploid genome background of numerous banana cultivars, it may be expected that many agronomic traits are under complex genetic regulation control, necessitating innovative approaches to investigate the role of apparent gene redundancy (Cenci et al., 2019; D'Hont et al., 2012). We have not yet reached the stage where one can easily pick up a gene variant coding for a specific trait and select material of interest to conduct crop improvement. Significant research is still needed to better understand the physiology and genetic architecture of traits in banana. Phenotyping experiments for various traits, including fruit quality, are also still missing, which may inhibit adoption of improved hybrids (Thiele et al., 2020). Furthermore, new plant breeding techniques such as gene editing will have to be fine-tuned for banana, even if some encouraging perspectives have been recently published (Tripathi et al., 2019; Zorrilla-Fontanesi et al., 2020). Finally, regulation frameworks of edited crops are still to be legislated in many countries (Schmidt et al., 2020). While waiting for future policy options, training on the use of such catalogs should be strengthened, particularly for breeders in national programs in those low-income countries with supportive funding schemes.

4 | CONCLUSIONS

A digital catalog of genetic variants is available for banana and is directly linked to the diversity held in the ITC genebank. It is accessible online as a proof of concept for exploration and export of SNP datasets. We adapted the system with the objective of keeping the genetic information connected to the physical material maintained in the genebank. Users can browse genetic information, identify interesting material and order it online for further investigation and use in breeding programs. While many genebanks are wondering if managing high-density markers is in their scope, the GIGWA web application offers a simple and elegant solution. With a reasonable transaction cost, its framework can be extrapolated to any germplasm collection.

Challenges still need to be addressed. First, on a technical side, the datasets are stored by clusters of accessions resulting from individual studies. Merging datasets from various projects and sequencing platforms is a challenging task. From a financial perspective, funder investment is needed to complete the genotyping of the whole collection. Given the relatively small size of the clonal collection of bananas (1,600 accessions compared to the 773,000 accessions managed by the CGIAR collections in total), it would not require a massive investment. Finally, to comply with international rules, further developments will have to take into account the agreements on digital sequence information (DSI) and on access and benefit-sharing (ABS) that are currently being debated in the frameworks of the Convention on Biological Diversity and the International Treaty for Plant Genetic Resources for Food and Agriculture.

ACKNOWLEDGMENTS

We thank the CGIAR Research Program, Roots, Tubers and Bananas (RTB) and the Directorate-general Development Cooperation and Humanitarian from the Belgian Development Cooperation for their financial support.

AUTHORS' CONTRIBUTIONS

M.R. led the writing of the manuscript with critical inputs from J.S. N.R. and S.C.C. I.V.D.H facilitated distribution of plant material. J.S. and M.R. coordinated genotyping sequence production and S.C.C. and M.R. coordinated transcriptomics data sequence production. C.B. and M.R. performed bioinformatics analyses and data management. G.S. and V.G. developed and deployed software for data management. N.R. acquired the project funding and supervised teamwork. All authors contributed to the draft and gave final approval for publication.

ORCID

Mathieu Rouard https://orcid.org/0000-0003-0284-1885
Catherine Breton https://orcid.org/0000-0002-3389-1861

REFERENCES

Aubry, S. (2019). The future of digital sequence information for plant genetic resources for food and agriculture. *Frontiers in Plant Science*, *10*, 1046. https://doi.org/10.3389/fpls.2019.01046

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

Carpentier, S. C., Dens, K., den Houwe, I. V., Swennen, R., & Panis, B. (2007). Lyophilization, a practical way to store and transport tissues prior to protein extraction for 2DE analysis? *Proteomics*, *7*, 64–69. https://doi.org/10.1002/pmic.200700529

Cenci, A., Hueber, Y., Zorrilla-Fontanesi, Y., van Wesemael, J., Kissel, E., Gislard, M., Sardos, J., Swennen, R., Roux, N., Carpentier, S. C., & Rouard, M. (2019). Effect of paleopolyploidy and allopolyploidy on gene expression in banana. *BMC Genomics*, *20*, 244, https://doi.org/10.1186/s12864-019-5618-0

Cenci, A., Sardos, J., Hueber, Y., Martin, G., Breton, C., Roux, N., Swennen, R., Carpentier, S. C., & Rouard, M. (2020). Unravelling the complex story of intergenomic recombination in ABB allotriploid bananas. *Annals of Botany*, *127*, 7–20. https://doi.org/10.1093/aob/mcaa032

D'Hont, A., Denoeud, F., Aury, J.-M., Baurens, F.-C., Carreel, F., Garsmeur, O., Noel, B., Bocs, S., Droc, G., Rouard, M., Da Silva, C., Jabbari, K., Cardi, C., Poulain, J., Souquet, M., Labadie, K., Jourda, C., Lengellé, J., Rodier-Goud, M., … Wincker, P. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature*, *488*, 213. https://doi.org/10.1038/nature11241

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

Droc, G., Lariviere, D., Guignon, V., Yahiaoui, N., This, D., Garsmeur, O., Dereeper, A., Hamelin, C., Argout, X., Dufayard, J.-F., Lengelle, J., Baurens, F.-C., Cenci, A., Pitollat, B., D'Hont, A., Ruiz, M., Rouard, M.,

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

& Bocs, S. (2013). The banana genome hub. *Database*, *2013*. https://doi.org/10.1093/database/bat035

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

Eyland, D., Breton, C., Sardos, J., Kallow, S., Panis, B., Swennen, R., Paofa, J., Tardieu, F., Welcker, C., Janssens, S. B., & Carpentier, S. C. (2020). Filling the gaps in gene banks: Collecting, characterizing and phenotyping wild banana relatives of Papua New Guinea. *Crop Science*, *61*(1), 137–149. https://doi.org/10.1002/csc2.20320

Food and Agriculture Organization of the United Nations (2019). *FAOSTAT statistical database*. Rome: FAO, c1997.

Gaffney, J., Tibebu, R., Bart, R., Beyene, G., Girma, D., Kane, N. A., Mace, E. S., Mockler, T., Nickson, T. E., Taylor, N., Zastrow-Hayes, G. (2020). Open access to genetic sequence data maximizes value to scientists, farmers, and society. *Global Food Security*, *26*, 100411. https://doi.org/10.1016/j.gfs.2020.100411

Germeier, C. U., & Unger, S. (2019). Modeling crop genetic resources phenotyping information systems. *Frontiers in Plant Science*, *10*, 728. https://doi.org/10.3389/fpls.2019.00728

Halewood, M., Chiurugwi, T., Sackville Hamilton, R., Kurtz, B., Marden, E., Welch, E., Michiels, F., Mozafari, J., Sabran, M., Patron, N., Kersey, P., Bastow, R., Dorius, S., Dias, S., McCouch, S., & Powell, W. (2018). Plant genetic resources for food and agriculture: opportunities and challenges emerging from the science and information technology revolution. *New Phytologist*, *217*, 1407–1419. https://doi.org/10.1111/nph.14993

Halewood, M., Lopez Noriega, I., Ellis, D., Roa, C., Rouard, M., & Sackville, H. N. (2017). *Potential implications of the use of digital sequence information on genetic resources for the three objectives of the Convention on Biological Diversity*. A submission from CGIAR to the Secretary of the Convention on Biological Diversity (CBD).

Halewood, M., Lopez Noriega, I., Ellis, D., Roa, C., Rouard, M., & Sackville, H. R. (2018). Using genomic sequence information to increase conservation and sustainable use of crop diversity and benefit-sharing. *Biopreservation and Biobanking*, *16*, 368–376. https://doi.org/10.1089/bio.2018.0043

König, P., Beier, S., Basterrechea, M., Schüler, D., Arend, D., Mascher, M., Stein, N., Scholz, U., & Lange, M. (2020). BRIDGE—A visual analytics web tool for barley genebank genomics. *Frontiers in Plant Science*, *11*, 701. https://doi.org/10.3389/fpls.2020.00701

Mansueto, L., Fuentes, R. R., Borja, F. N., Detras, J., Abriol-Santos, J. M., Chebotarov, D., Sanciangco, M., Palis, K., Copetti, D., Poliakov, A., Dubchak, I. (2017). Rice SNP-seek database update: New SNPs, indels, and queries. *Nucleic Acids Research*, *45*, D1075–D1081. https://doi.org/10.1093/nar/gkw1135

Martin, G., Baurens, F.-C., Hervouet, C., Salmon, F., Delos, J.-M., Labadie, K., Perdereau, A., Mournet, P., Blois, L., Dupouy, M., Carreel, F., Ricci, S., Lemainque, A., Yahiaoui, N., … D'Hont, A. (2020). Chromosome reciprocal translocations have accompanied subspecies evolution in bananas. *The Plant Journal*, *104*, 1698–1711. https://doi.org/10.1111/tpj.15031

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

McCouch, S., Baute, G. J., Bradeen, J., Bramel, P., Bretting, P. K., Buckler, E., Burke, J. M., Charest, D., Cloutier, S., Cole, G., Dempewolf, H. (2013). Agriculture: Feeding the future. *Nature*, *499*, 23–24. https://doi.org/10.1038/499023a

McCouch, S., Navabi, K., Abberton, M., Anglin, N. L., Barbieri, R. L., Baum, M., Bett, K., Booker, H., Brown, G. L., Bryan, G. J., Cattivelli, L. (2020). Mobilizing crop biodiversity. *Molecular Plant*, *13*, 1341–1344. https://doi.org/10.1016/j.molp.2020.08.011

Milner, S. G., Jost, M., Taketa, S., Mazón, E. R., Himmelbach, A., Oppermann, M., Weise, S., Knüpffer, H., Basterrechea, M., König, P., Basterrechea, M., König, P., Schüler, D. (2019). Genebank genomics s the diversity of a global barley collection. *Nature Genetics*, *51*, 319–326. https://doi.org/10.1038/s41588-018-0266-x

Raubach, S., Kilian, B., Dreher, K., Amri, A., Bassi, F. M., Boukar, O., Cook, D., Cruickshank, A., Fatokun, C., Haddad, N. E., Humphries, A., Jordan, D., Kehel, Z., Kumar, S., Labarosa, S. J., Nguyen, L. H., Mace, E., McCouch, S., McNally, K., … Shaw, P. D. (2020). From bits to bites: Advancement of the germinate platform to support genetic resources collections and pre-breeding informatics for crop wild relatives. *Crop Science*, 1–29. https://doi.org/10.1002/csc2.20248

Rouard, M., Droc, G., Martin, G., Sardos, J., Hueber, Y., Guignon, V., Cenci, A., Geigle, B., Hibbins, M. S., Yahiaoui, N., Baurens, F. C., Berry, V., Hahn, M. W., D'Hont, A., & Roux, N. (2018). Three new genome assemblies support a rapid radiation in *Musa acuminata* (wild banana). *Genome Biology and Evolution*, *10*, 3129–3140. https://doi.org/10.1093/gbe/evy227

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*, https://doi.org/10.1093/database/bax046

Sansaloni, C., Franco, J., Santos, B., Percival-Alwyn, L., Singh, S., Petroli, C., Campos, J., Dreher, K., Payne, T., Marshall, D., Kilian, B. (2020). Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nature Communications*, *11*, 4572 https://doi.org/10.1038/s41467-020-18404-w

Sardos, S., Breton, C., Perrier, X., Van den Houwe, I., Paofa, J., Rouard, M., & Roux, N. (2021). Wild to domesticates: genomes of edible diploid bananas hold traces of several undefined genepools. *bioRxiv*, *428762*. https://doi.org/10.1101/2021.01.29.428762

Sardos, J., Rouard, M., Hueber, Y., Cenci, A., Hyma, K. E., van den Houwe, I., Hribova, E., Courtois, B., & Roux, N. (2016). A Genome-Wide Association Study on the seedless phenotype in banana (Musa spp.) reveals the potential of a selected panel to detect candidate genes in a vegetatively propagated crop. *PLoS One*, *11*, e0154448. https://doi.org/10.1371/journal.pone.0154448

Schmidt, S. M., Belisle, M., & Frommer, W. B. (2020). The evolving landscape around genome editing in agriculture. *EMBO Reports*, *21*, e50680. https://doi.org/10.15252/embr.202050680

Scholz, A., Hillebrand, U., Freitag, J., Cancio, I., Ribeiro, C., Haringhuizen, G., Oldham, P., Saxena, D., Seitz, C., Thiele, T., & van Zimmeren, E. (2020). *Finding compromise on ABS & DSI in the CBD: Requirements & policy ideas from a scientific perspective*. https://doi.org/10.13140/RG.2.2.35180.80001

Selby, P., Abbeloos, R., Backlund, J. E., Basterrechea Salido, M., Bauchet, G., Benites-Alfaro, O. E., Birkett, C., Calaminos, V. C., Carceller, P., Cornut, G., Vasques Costa, B. (2019). BrAPI—an application programming interface for plant breeding applications. *Bioinformatics*, *35*, 4147–4155. https://doi.org/10.1093/bioinformatics/btz190

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

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

Smyth, S. J., Macall, D. M., Phillips, P. W. B., & de Beer, J. (2020). Implications of biological information digitization: Access and benefit sharing of plant genetic resources. *The Journal of World Intellectual Property*, *23*, 267–287. https://doi.org/10.1111/jwip.12151

Thiele, G., Dufour, D., Vernier, P., Mwanga, R. O. M., Parker, M. L., Geldermann, E. S., Teeken, B., Wossen, T., Gotor, E., Kikulwe, E.,

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Tufan, H. (2020). A review of varietal change in roots, tubers and bananas: Consumer preferences and other drivers of adoption and implications for breeding. *International Journal of Food Science & Technology*. https://doi.org/10.1111/ijfs.14684

Tripathi, J. N., Ntui, V. O., Ron, M., Muiruri, S. K., Britt, A., & Tripathi, L. (2019). CRISPR/Cas9 editing of endogenous banana streak virus in the B genome of Musa spp. overcomes a major challenge in banana breeding. *Communications Biology*, *2*, 46. https://doi.org/10.1038/s42003-019-0288-7

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

van Wesemael, J., Kissel, E., Eyland, D., Lawson, T., Swennen, R., & Carpentier, S. C. (2019). Using growth and transpiration phenotyping under controlled conditions to select water efficient banana genotypes. *Frontiers in Plant Science*, *10*, 352. https://doi.org/10.3389/fpls.2019.00352

Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., Li, M., Zheng, T., Fuentes, R. R., Zhang, F., Mansueto, L., Copetti, D., Sanciangco, M., Palis, K. C., Xu, J., Sun, C., Fu, B., Zhang, H., Gao, Y., … Leung, H. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature*, *557*, 43–49. https://doi.org/10.1038/s41586-018-0063-9

Wang, Z., Miao, H., Liu, J., Xu, B., Yao, X., Xu, C., Zhao, S., Fang, X., Jia, C., Wang, J., Zhang, J., Li, J., Xu, Y., Wang, J., Ma, W., Wu, Z., Yu, L., Yang, Y., Liu, C., … Jin, Z. (2019). Musa balbisiana genome reveals subgenome evolution and functional divergence. *Nature Plants*, *5*, 810–821. https://doi.org/10.1038/s41477-019-0452-6

Wu, W., Yang, Y.-L., He, W.-M., Rouard, M., Li, W.-M., Xu, M., Roux, N., & Ge, X.-J. (2016). Whole genome sequencing of a banana wild relative Musa itinerans provides insights into lineage-specific diversification of the Musa genus. *Scientific Reports*, *6*, 31586. https://doi.org/10.1038/srep31586

Zorrilla-Fontanesi, Y., Pauwels, L., Panis, B., Signorelli, S., Vanderschuren, H., & Swennen, R. (2020). Strategies to revise agrosystems and breeding to control Fusarium wilt of banana. *Nature Food*, *1*, 599–604. https://doi.org/10.1038/s43016-020-00155-y

How to cite this article: Rouard M, Sardos J, Sempéré G, et al. A digital catalog of high-density markers for banana germplasm collections. *Plants*, *People*, *Planet*. 2022;4:61–67. https://doi.org/10.1002/ppp3.10187

25722611, 2022, 1, Downloaded from https://nph.onlinelibrary.wiley.com/doi/10.1002/ppp3.10187 by CIRAD - DGDRS - DIST, Wiley Online Library on [15/02/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License