



# Continuous Mapping Identifies Loci Associated With Weevil Resistance [*Cosmopolites sordidus* (Germar)] in a Triploid Banana Population

Brigitte Uwimana<sup>1\*</sup>, Gerald Mwanje<sup>1</sup>, Michael Batte<sup>1</sup>, Violet Akech<sup>1</sup>, Trushar Shah<sup>2</sup>, Marnik Vuylsteke<sup>3</sup> and Rony Swennen<sup>1,4</sup>

<sup>1</sup> International Institute of Tropical Agriculture (IITA), Kampala, Uganda, <sup>2</sup> International Institute of Tropical Agriculture (IITA), International Livestock Research Institute Campus, Nairobi, Kenya, <sup>3</sup> Gnomixx B.V., Melle, Belgium, <sup>4</sup> Department of Crop Biosystems, KU Leuven, Heverlee, Belgium

## OPEN ACCESS

### Edited by:

Hanwei Mei,  
Shanghai Agrobiological Gene Center,  
China

### Reviewed by:

Marcelo Mollinari,  
North Carolina State University,  
United States

Jonathan Fresnedo Ramirez,  
The Ohio State University,  
United States

### \*Correspondence:

Brigitte Uwimana  
B.Uwimana@cgiar.org

### Specialty section:

This article was submitted to  
Plant Breeding,  
a section of the journal  
Frontiers in Plant Science

**Received:** 04 August 2021

**Accepted:** 14 October 2021

**Published:** 29 November 2021

### Citation:

Uwimana B, Mwanje G, Batte M,  
Akech V, Shah T, Vuylsteke M and  
Swennen R (2021) Continuous  
Mapping Identifies Loci Associated  
With Weevil Resistance  
[*Cosmopolites sordidus* (Germar)] in a  
Triploid Banana Population.  
*Front. Plant Sci.* 12:753241.  
doi: 10.3389/fpls.2021.753241

The first step toward marker-assisted selection is linking the phenotypes to molecular markers through quantitative trait loci (QTL) analysis. While the process is straightforward in self-pollinating diploid (2x) species, QTL analysis in polyploids requires unconventional methods. In this study, we have identified markers associated with weevil *Cosmopolites sordidus* (Germar) resistance in bananas using 138 triploid (2n = 3x) hybrids derived from a cross between a tetraploid “Monyet” (2n = 4x) and a 2x “Kokopo” (2n = 2x) banana genotypes. The population was genotyped by Diversity Arrays Technology Sequencing (DArTSeq), resulting in 18,009 polymorphic single nucleotide polymorphisms (SNPs) between the two parents. Marker–trait association was carried out by continuous mapping where the adjusted trait means for the corm peripheral damage (PD) and total cross-section damage (TXD), both on the logit scale, were regressed on the marker allele frequencies. Forty-four SNPs that were associated with corm PD were identified on the chromosomes 5, 6, and 8, with 41 of them located on chromosome 6 and segregated in “Kokopo.” Eleven SNPs associated with corm total TXD were identified on chromosome 6 and segregated in “Monyet.” The additive effect of replacing one reference allele with the alternative allele was determined at each marker position. The PD QTL was confirmed using conventional QTL linkage analysis in the simplex markers segregating in “Kokopo” (AAAA × RA). We also identified 43 putative genes in the vicinity of the markers significantly associated with the two traits. The identified loci associated with resistance to weevil damage will be used in the efforts of developing molecular tools for marker-assisted breeding in bananas.

**Keywords:** banana, banana weevil, continuous mapping, *Cosmopolites sordidus*, *Musa* spp., polyploids, QTL analysis

## INTRODUCTION

Pests and diseases are the main production constraints for bananas and plantains (*Musa* spp.) globally (Jones, 2009). Among them, nematodes and weevils constitute the topmost pests, with *Cosmopolites sordidus* (Germar) or banana weevil root borer (henceforth referred to as banana weevil) causing yield losses of up to 50% (Rukazambuga et al., 1998; Gold et al., 2004). Yield losses

due to banana weevils worsen with every subsequent ratoon of the plantation due to the build-up of the infestation, and 100% losses can be realized because of complete mat disappearance (Gold et al., 2004). Plantains and East African highland cooking bananas (EAHB), which constitute an important source of food and livelihood in sub-Saharan Africa, are highly susceptible to the pest (Gold and Bagabe, 1997; Gold et al., 2001; Kiggundu et al., 2003; Swennen et al., 2013). Banana weevils have been cited as one of the major factors that have caused production decline and disappearance of East African highland bananas (AAAh) in Central Uganda (Rukazambuga et al., 1998; Gold et al., 1999) and in some littoral regions of Kagera in Tanzania (Mbwana and Rukazambuga, 1998). Therefore resistance to banana weevils is one of the breeding goals for the improvement of these two groups of bananas (Brown et al., 2017).

*Cosmopolites sordidus* (Germer) is a beetle that belongs to the subfamily Rhynchophorinae, of the family Curculionidae, of the order Coleoptera (Zimmerman, 1968; Gold et al., 2001). It was first reported as a banana pest in Malaysia as early as 1824 (Zimmerman, 1968). From there, it has spread in almost all the banana-growing regions globally (Gold et al., 2001). Its first report in Africa dates to the early 1900s (Blomme et al., 2013). The reproduction of the banana weevils reduces at temperatures below 12°C, therefore, their damage is more pronounced in banana plantations at an altitude of 1,000 to 1,300 m above sea level or lower, and rarely above 1,600 m above sea level (Gold and Messiaen, 2000). The damages caused by banana weevils arise from the insect larvae feeding on the corm (the true stem of a banana plant) where they bore tunnels. Although rare, feeding on the pseudostem and banana leaves has also been reported, especially under heavy attacks (VILARDEBO, 1973). Consequently, the tunnels in the corm reduce the uptake of nutrients by the plant, hence reducing the plant vigor and delaying flowering. Injuries to the banana corm due to weevils also serve as an entry point for other soil-borne pests and diseases. Ultimately, the yield is affected through the death of the plants by snapping or toppling, production of small bunches, mat disappearance, and eventually, reduced longevity of the plantation (Rukazambuga et al., 1998; Gold and Messiaen, 2000; Gold et al., 2004).

The integrated pest management of the banana weevil includes a combination of cultural, chemical, and biological measures (reviewed by Gold et al., 2001). Among the methods suggested, the use of resistant cultivars is the most sustainable way of controlling banana weevils. Host resistance has been identified in dessert cultivars (AAA) such as “Yangambi Km5,” “Cavendish,” and “Gros Michel,” as well as in some wild and wild-derived diploids (2x) (AA), such as “Calcutta 4,” “TMB2 × 8075-7,” “TMB2 × 7197-2,” and “TMB2 × 6142-1” (Ortiz et al., 1995; Kiggundu et al., 2003). Early and more recent genetic studies, such as those by Ortiz et al. (1995) and Arinaitwe et al. (2015), have endeavored to understand the inheritance of weevil resistance using segregating populations. In a multiple-family plantain population (3x × 2x crosses), the first study concluded that resistance to weevils was governed by partially dominant genes from the resistant parent combined with modifier genes with additive and dosage effect from the plantains (susceptible

parents). The second study reported a single gene controlling the total corm damage in a progeny from a 2x × 2x cross.

Banana improvement relies mainly on classical crossbreeding. This is a three-step process, first focusing on the improvement of the 2x parents from the wild genotypes to get rid of the non-domestication traits. Second, the improved 2x are crossed with triploid (3x) landraces to generate primary tetraploids (4x), and third, the primary 4x are further crossed with the improved 2x to generate secondary 3x hybrids which should combine parthenocarpy with high yield, resistance to pests and diseases, and fruit qualities comparable with those of the original landraces (Lorenzen et al., 2010; Ortiz and Swennen, 2014; Brown et al., 2017). Improved varieties have been developed through this method by a handful number of breeding programs (Tushemereirwe et al., 2015; Tumuhimbise et al., 2016; Dépigny et al., 2018; Tenkouano et al., 2019). However, the process of developing acceptable hybrids is slow, tedious, and expensive, taking up to 2 decades (Batte et al., 2019). One of the main reasons is the long cycle of the crop, which leads to one field evaluation cycle taking 2 to 3 years. Considering the decreasing cost of sequencing-based genotyping, marker-assisted breeding (MAB) has a great potential to improve the efficiency of banana breeding by reducing the breeding cycle by many years (Lorenzen et al., 2011; Nyine et al., 2018).

Identifying marker-trait associations by linkage or association studies is the primary step toward the development of markers for MAB. While mapping by linkage or association studies is well-established in 2x organisms, it is less straightforward when the species, such as banana, is cross-pollinating and polyploid. Although the first banana linkage map was published in 1993 (Fauré et al., 1993), quantitative trait loci (QTL) mapping in bananas has been slow due to the difficulty in developing and maintaining segregating populations in bananas, the extensive efforts required in phenotyping, and the further level of complexity in distinguishing between homolog and chromosomal linkage groups in polyploids. With the release of the banana reference genomes (D’Hont et al., 2012; Martin et al., 2016; Rouard et al., 2018), molecular research toward MAB in bananas is steadily gaining momentum (Sardos et al., 2016; Nyine et al., 2018, 2019; Ahmad et al., 2020).

The generation of accurate genotypic and phenotypic data is the foremost important aspect of developing useful tools for MAB through linkage QTL analysis, genome-wide association studies (GWAS), or genomic selection. Unlike 2x species, polyploids are characterized by multiple allelic combinations (dosages) at the heterozygous loci (Bourke et al., 2018b). For instance, five distinct dosages are possible in a 4x at a bi-allelic marker position, ranging from zero to four copies of the alternative allele. With increasing interest in untargeted sequencing-based genotyping (GBS) due to its relatively low cost, tools to accurately determine the allele dosage using single nucleotide polymorphism (SNP) allele counts are expected to be developed. However, the accurate calling of polyploid genotypes from sequence data requires a sequencing depth of 60x to 80x, estimated in an autotetraploid potato (Uitdewilligen et al., 2013). This would make GBS too expensive for the large-scale genotyping in the study we report, i.e., calling 3x in a banana mapping population of more than 150 segregants.

As an alternative approach, Ashraf et al. (2014) proposed to process allele counts to allele frequency estimates and regress the phenotype on the allele-frequencies. This study has shown that, under additivity assumptions, there is a linear relationship between the phenotype and the allele frequency. Therefore, allele-frequency estimates can be used directly, rather than calling genotypes. Hence, a linear regression of the phenotype on the individual allele frequencies, dubbed “continuous mapping,” will allow (1) the mapping of QTL and (2) the estimation of the allele substitution effect at each SNP.

The purpose of this study was to identify the QTL associated with weevil resistance traits in a 3x F<sub>1</sub> banana progeny derived from a cross between a 4x (“Monyet”) and a 2x (“Kokopo”) accessions using the continuous mapping approach. This is the first study ever on marker-trait association in 4x × 2x banana progeny and on weevil resistance in bananas.

## MATERIALS AND METHODS

### Plant Material

This study was carried out at the International Institute of Tropical Agricultural (IITA), Sendusu station in Uganda (0°31′30″N; 32°36′54″E, 1260 m above sea level). A population of 210 F<sub>1</sub> hybrids was generated by crossing “Monyet” (ITC1179, female parent) and “Kokopo” (ITC1233, male parent) through controlled hand-pollination. “Monyet” is a wild banana that belongs to the *zebrina* subspecies (*Musa acuminata*). It is known as a diploid banana (2x = 2n = 22) originating from Indonesia (Martin et al., 2020b), but Arinaitwe et al. (2019) reported it as a 4x (2x = 4n = 44). “Kokopo” is a traditional 2x cultivar of the *M. acuminata* ssp. *banksii* origin (Christelová et al., 2016). From field observations, “Monyet” is resistant to *C. sordidus*, while “Kokopo” is susceptible.

### Ploidy Analysis

The ploidy level of the two parents and their hybrids was determined using a flow cytometry method (Dolezel et al., 1997; Christelová et al., 2016). Nucleic DNA was extracted and stained with 4′,6-diamidino-2-phenylindole (DAPI) as described by Karamura et al. (2016). The DNA was analyzed through a ploidy analyzer (CyFlow® Cube 6, Sysmex Partec, Sysmex Partec GmbH). “Calcutta 4” (2x = 2n = 22) was used as a reference standard 2x and its G<sub>1</sub> peak was set to channel 200. The 2x, 3x, and 4x levels of all F<sub>1</sub> progeny were determined relative to the peak of “Calcutta 4,” with their G<sub>1</sub> peaks set to channels 300 and 400 for the 3x and 4x, respectively.

### Phenotyping

One hundred and forty hybrids together with their parents (“Monyet” and “Kokopo”) and four checks, namely, “Calcutta-4” (AA, ITC0249) and “Yangambi Km5” (AAA, ITC1123) as resistant checks and two East African cooking banana cultivars, “Nakyatengu” and “Kabucurage” (EAHBs, AAAh) as susceptible checks (Kiggundu et al., 2003; Sadik et al., 2010), were screened for their resistance to weevils in a pot experiment. The experiment was set up as described by Sadik et al. (2010).

Briefly, healthy suckers from the field were selected, pared, and dipped in boiling water for 15 s, with the time reduced from the recommended 20–30 s due to the small size of the suckers (Coyne et al., 2010). The suckers were planted in 13 L plastic buckets filled up to 3/4 with sterilized top forest soil, manure, and sawdust (3:1:1). Because not all the hybrids had an equal number of suckers at the same time, the experiment was run in four series (batches), with each series containing the two parents and the four checks. Each series was set up as a randomized complete block design of two blocks with four plants per genotype per block. Each bucket was fitted with a net to avoid the natural infestation of weevils, and the plants were watered every other day (three times a week) and were allowed to establish for 8 weeks.

Trapping using pieces of banana pseudostems were used to collect weevils from an already established plantation of EAHB (Viljoen et al., 2017). The collected weevils were reared in 30 L buckets containing pared corms of a weevil-susceptible cultivar “Nakyatengu” and placed in a cool and dark place. The corms were kept cool by spraying them with water and were changed every 7 days (Viljoen et al., 2017). The male and female weevils were distinguished under a stereomicroscope (K700L, Thomas Scientific) by their morphological features (Gold et al., 2001; Viljoen et al., 2017). Eight weeks after the start of the experiment, 10 weevils (5 males and 5 females) were introduced per plant. Sixty days after the introduction of the weevils, data were collected on corm damage as follows: after uprooting the plants, the corms were pared, and the peripheral damage (PD) was recorded as an estimate of the percentage area damaged by the larvae. Two cross-sections of the corm were generated by cutting at 3 cm (upper cross-section) and 6 cm (lower cross-section) from the collar (Sadik et al., 2010). The corm damage was scored for each cross-section as the percentage of the damaged area of the cortex (upper and lower outer cross-section damage) and the inner cylinder (upper and lower inner cross-section damage). The total cross-section damage (TXD) was estimated by taking the average of the four readings (Sadik et al., 2010). The number of recovered larvae, pupae, and weevils was also recorded per plant.

### Statistical Analysis of the Phenotypic Data

The percentage data (PD and TXD) were logit-transformed before the analysis. The analysis was carried out in Genstat® for Windows™ 20<sup>th</sup> Edition (VSN International, 2019). Just like the QTL analysis, continuous mapping requires trait means rather than individual plot (or unit) values from replicated trials. The Preliminary Single Environment Analysis (PSEA) menu, selected from the Stats | QTLs (linkage/association)| Phenotypic Analysis menu in Genstat® helps to produce adjusted trait means from the experimental data. The design was set to “general.” The term “genotype” was broken into two factors as “genotype\_test” for the F<sub>1</sub> hybrids and “genotype\_extra” for the parents and the checks. The PSEA runs two restricted maximum likelihood (REML) analyses. In the first analysis, the “genotype\_test” factor was fitted as a random term to estimate the variance parameters and heritability, while “genotype\_extra” was fitted as a fixed term. In the second

analysis, the two genotype factors were fitted as a fixed term and unshrunk predicted means were estimated using variance parameters estimated in the first run. The term “batch” was added as an additional fixed term and the term “replication” nested within the batches was added as a random term to the two models, both describing the physical structure of the trial.

## Genotyping

Genotyping was carried out using the genotyping-by-sequencing platform of DArTSeq (DArT®)<sup>1</sup> at the Biosciences eastern and central Africa - International Livestock Research Institute (BecA - ILRI) under the Integrated Genotyping Service and Support (IGSS) project. The sequencing library was generated using the enzyme *Pst*I. The raw short-read sequences for each sample were obtained as fastq files from IGSS. These short reads were then preprocessed to trim off the barcodes and adapters using the FASTX-Toolkit version 0.0.13<sup>2</sup>. The reads were then aligned to the reference genome sequence, version 2 of the “DH Pahang” banana genome (Martin et al., 2016) using Bowtie2 version 2.3.4.1 (Langmead and Salzberg, 2012). The genome analysis tool kit (GATK) version 3.7.0. (Van der Auwera et al., 2013) was used for allele calling per allele dosage using the ploidy information of each genotype. The variant call format (VCF) files generated from the GATK allele calling were filtered for a minor allele frequency of 0.02.

## Continuous Quantitative Trait Loci Mapping

For the continuous QTL mapping, genotypes were not explicitly called, but the allele frequencies obtained from the sequence data were directly used for regression analysis (Ashraf et al., 2014). This method was dubbed “continuous QTL mapping.” For each individual, reference allele counts from the sequencing reads were processed to the reference allele frequency estimates which were directly used as “continuous genotypes,” on which the phenotypes were regressed. The allele frequencies suffer from inaccuracy, which we accounted for by correcting the measurement error as described by Ashraf et al. (2014). For the 3x families, the bias or underestimation in the estimate of the allele effect equals  $\frac{1}{1+2/S_T}$ ,

where  $S_T$  represents the sequencing depth. The above expression of the bias is derived for a constant sequencing depth across all samples. To account for variable depth across samples, the harmonic mean of the depths per sample was used. A regression of the adjusted trait means for PD and TXD, both on the logit scale, on the marker allele frequencies was performed in Genstat® for Windows™ 20<sup>th</sup> Edition (VSN International, 2019). At each marker position, a *p*-value was calculated, assessing the linear regression together with the estimation of the additive effect of replacing the reference allele with the alternative allele. The “DH Pahang” physical map (Martin et al., 2016) was used for the physical ordering of the markers.

## Linkage Mapping and Conventional Quantitative Trait Loci Analysis

To confirm the continuous QTL mapping results, a genetic linkage map was constructed for the simplex markers segregating in “Kokopo” (2x). These markers were of the AAAA (“Monyet”) × RA (“Kokopo”) parental segregation type, with R denoting the reference allele and A the alternative allele. These markers were expected to segregate as AAA (50%) and AAR (50%) in the 3x offspring, thus, allowing the use of the linkage and QTL mapping software available for 2x mapping populations. The F<sub>1</sub> genotypes for these markers were called from the allele counts obtained from the GATK analysis. A maximum likelihood approach was used to infer the genotype for parents and each offspring from the allele counts obtained from the DArTSeq data. At each site *j* for each individual *i*, the likelihood for each of the four possible genotypes is given as:

$$L_{gij} = P(D_{ij}|g_{ij}) \quad (1)$$

where  $D_{ij}$  is the observed sequencing data in the individual *i* at site *j*, and  $g_{ij} \in \{0,1,2,3\}$  is the number of reference alleles contained in the genotype of each individual.

To assign a genotype to a particular individual, the likelihood of each of the four possible genotypes was calculated for the individual. The genotype with the highest likelihood was then assigned to an individual when the most likely genotype was 10 times more likely than the second most likely one. Otherwise, a genotype was not called, and the genotype of the individual was considered missing.

A linkage map was then constructed with outcrossing (CP) specified as a mapping population using the linkage mapping module in Genstat® for Windows™ 20<sup>th</sup> Edition (VSN International, 2019). Linkage grouping was performed at threshold recombination of 0.20. Only linkage groups of at least 10 markers were further analyzed into maps for the QTL mapping in the same software. The option “Single Trait Linkage Analysis (Single Environment)” was chosen to perform the genome-wide scans for QTL effects (Simple and Composite Interval Mapping) for logit\_PD and logit\_TXD. Conditional genotypic probabilities were calculated for every 5 cM based on the Haldane mapping function. The threshold for the QTL detection was set at  $-\log(p) = 2$ . An initial genome-wide scan by simple interval mapping (SIM) was performed to obtain the candidate QTL positions. The candidate QTL were then used as cofactors in the subsequent composite interval mapping (CIM), i.e., a genome-wide scan for QTL effects in the presence of cofactors.

## Putative Genes for Weevil Resistance

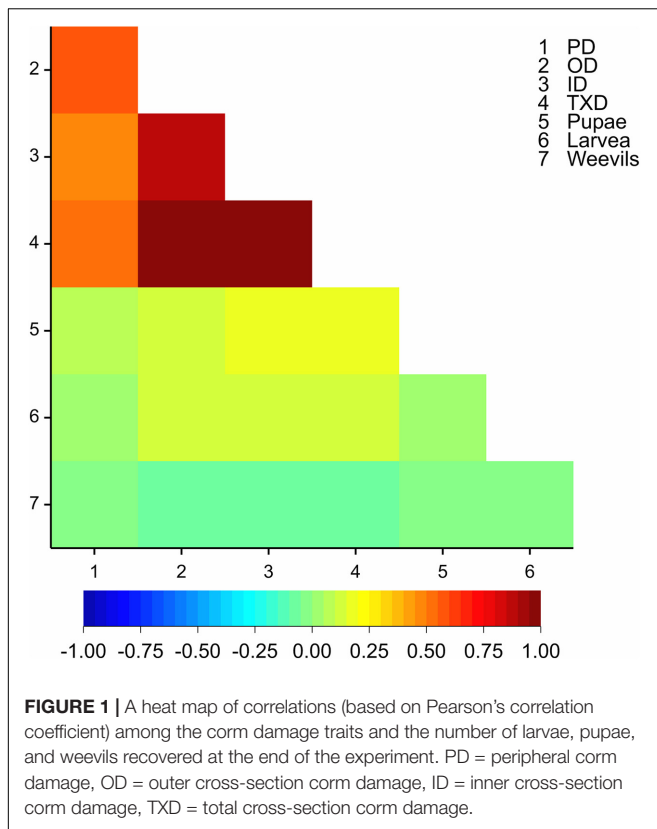
The annotation file for all features on the “DH Pahang” genome (Version 2) was downloaded from the Banana Genome Hub.<sup>3</sup> The annotation file was converted from a generic feature format (gff3) file to a browser extensible data (bed) file format. The list of the identified SNPs with their physical positions on the genome was used in a bedtools (v2.29.0) query with the function closestBed (Quinlan and Hall, 2010). The query identified the

<sup>1</sup><https://www.diversityarrays.com/technology-and-resources/dartreseq/>

<sup>2</sup>[http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)

<sup>3</sup><https://banana-genome-hub.southgreen.fr>





closest gene to the SNP position on either strand. The attributes of identified putative genes were searched in the gene list analysis of the protein analysis through evolutionary relationships gene ontology (PANTHER16.0<sup>4</sup>) using the *M. acuminata* spp. *malaccensis* database (Mi et al., 2020).

## RESULTS

### Ploidy of the Genotypes

Ploidy analysis confirmed “Kokopo” as a 2x and revealed that “Monyet” was a 4x. As expected from a 4x × 2x cross, the majority of the progeny was 3x (94.3%), while 2x and 4x constituted 2.58% and 3.09%, respectively. To avoid the mixture of ploidy, linkage analysis, continuous, and QTL mapping was carried out using the 138 3x F<sub>1</sub> hybrids whose genotypic and phenotypic data were available.

### Phenotypic Data

The outer and inner cross-section corm damage traits were highly correlated ( $r = 0.88$ ,  $p < 0.001$ ), and the two traits were positively correlated with PD ( $r = 0.58$  and  $r = 0.47$ , respectively,  $p < 0.001$ , **Figure 1**). The three corm damage traits (peripheral, inner, and outer cross-section damage), however, were poorly correlated with the number of recovered larvae, pupae, and weevil ( $0.00 < r < 0.22$ , **Figure 1**). Therefore, further

<sup>4</sup><http://pantherdb.org>

**TABLE 1** | The least significant difference (LSD) comparison between the adjusted means for logit-transformed PD (logit\_PD) and TXD (logit\_TXD) for the checks and the parents of the population.

	Logit_PD	Logit_TXD
Calcutta 4	-0.47 <sup>a</sup>	-5.13 <sup>a</sup>
Yangambi Km5	-0.11 <sup>b</sup>	-4.38 <sup>a</sup>
Monyet	0.28 <sup>b</sup>	-2.52 <sup>b</sup>
Kokopo	0.32 <sup>b</sup>	-0.28 <sup>c</sup>
Nakyatengu	0.72 <sup>c</sup>	0.36 <sup>c</sup>
Kabucuragye	0.88 <sup>c</sup>	0.30 <sup>c</sup>
LSD <sub>min</sub>	0.22	0.65
LSD <sub>average</sub>	0.36	1.04
LSD <sub>max</sub>	0.46	1.35

Genotypes having identical letters (a, b, c) do not differ at the 5% significance level.

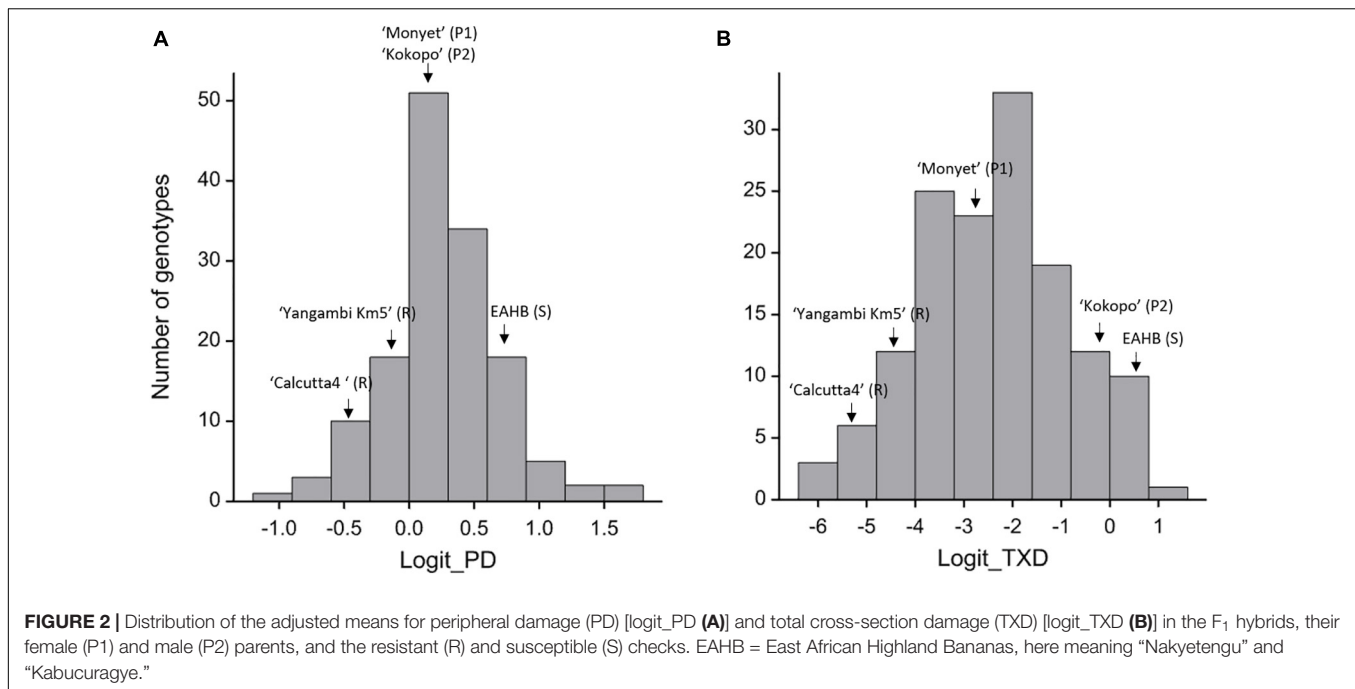
analysis was performed on the corm damage traits only, and the inner and outer cross-section damage were represented by their average as TXD.

The resistant checks, “Yangambi Km5” and “Calcutta 4,” were significantly different from the susceptible checks, “Nakyatengu” and “Kabucuragye” for both PD and TXD ( $p < 0.001$ ). Regarding the two resistant checks, “Calcutta 4” showed a significantly higher level of resistance than “Yangambi Km5” for PD (**Table 1**). The parents were contrasting for TXD, with “Monyet” being significantly less resistant than “Calcutta 4” and “Yangambi Km5,” and “Kokopo” as susceptible as the two EAHBs. However, for PD, the two parents showed equal levels of resistance (**Table 1**).

The two traits, PD and TXD, showed segregation in the population (**Figure 2**). The PD had a population mean of 0.25, ranging from -0.93 to 1.71 on the logit scale (**Figure 2A**), equivalent to 56, 28, and 85% corm damage, respectively. Despite the same level of PD for both parents, the broad-sense heritability for PD was estimated at 0.34, suggesting transgressive segregation in the population. The TXD in the hybrids was skewed toward the resistant phenotype as the majority of the hybrids were resistant with a population means of -2.42 or 8.16% corm damage. The score of the hybrids ranged from -6.06 to 0.90, equivalent to 0.23 and 71% corm damage, respectively (**Figure 2B**). The broad-sense heritability for TXD was estimated at 0.44. Also, for TXD, transgressive segregation in the population was detected: 41 genotypes were significantly more resistant than the resistant parent “Monyet,” and 11 genotypes were significantly more susceptible than the susceptible parent “Kokopo” (**Figure 2B**).

### Genotypic Data

The DArTSeq detected 37,436 unique allelic sequences, of which 28% (10,401) were monomorphic between the two parents and 24% (9,026) were polymorphic between the two parents but not segregating (**Table 2**). Continuous mapping was carried out using 18,009 segregating SNPs, represented by 14,254 SNPs that segregated in “Kokopo” only, 3,067 SNPs that segregated in “Monyet” only, and 688 SNPs that segregated in both parents (**Table 2**). The markers were well distributed over the 11 banana chromosomes of the “DH Pahang” physical map [2nd version, (Martin et al., 2016)]. They ranged from 945 to 2,481 SNPs per



**TABLE 2 |** Number of SNPs per parent genotype combinations in the “Monyet” × “Kokopo” population.

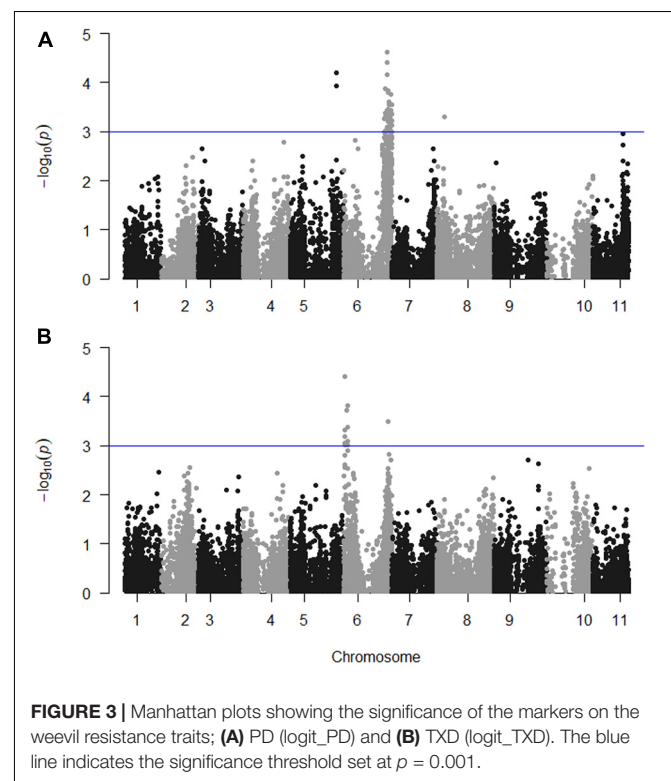
		Genotypes in “Kokopo” (2x) <sup>1</sup>			
		RR	RA	AA	Total
Genotypes in “Monyet”(4x) <sup>1</sup>	RRRR	499	9,635	3,458	13,592
	RRRA	451	189	48	688
	RRAA	2,035	390	206	2,631
	RAAA	243	109	84	436
	AAAA	5,568	4,619	9,902	20,089
	Total	8,796	14,942	13,698	37,436

<sup>1</sup>R: reference allele, A: alternative allele.

chromosome, with the minimum number of markers detected on chromosome 10 and the maximum number on chromosome 8. Only 125 markers came from the unanchored contigs of the “DH Pahang” physical map (**Supplementary Figure 1**).

## Continuous Mapping

Continuous mapping identified 44 SNP markers showing a significant ( $p < 0.001$ ) linear regression with PD (logit\_PD) and 11 SNPs showing a significant ( $p < 0.001$ ) linear regression with TXD (logit\_TXD, **Figure 3**). Forty-one (93%) of the SNPs significantly associated with PD and all 11 SNPs significantly associated with TXD were located on chromosome 6, with the SNPs associated with TXD located at two distinct regions on the two arms of the chromosome. The remaining three SNPs associated with PD were on chromosomes 5 (2 SNPs) and 8 (1 SNP, **Figure 3** and **Table 3**). The SNPs associated with the resistance to weevils at the peripheral level segregated in the parental genotype “Kokopo,” except for two SNPs on chromosome 5 segregating in “Monyet.” Conversely, the SNP



markers associated with the resistance to weevils at the corm cross-section level (logit\_TXD) segregated in “Monyet” only (**Table 3**). This opposite parental segregation suggests that the two traits have different mechanisms of resistance in this population. The proportion of the phenotypic variance explained

**TABLE 3** | Single nucleotide polymorphism markers significantly associated with PD (Logit\_PD) and TXD (logit-TXD) due to weevils at a significance level of  $\alpha = 0.001$  based on the continuous QTL mapping in a population of 138 triploid F<sub>1</sub> progeny derived from the “Monyet” × “Kokopo” cross having 18,009 SNP markers segregating.

Trait	SNP ID	Chr. <sup>1</sup>	Position (bp)	-Log(p)	Effect <sup>2</sup>	Expl. var. (%) <sup>3</sup>	Alleles <sup>4</sup>		Parental genotypes		
							R	A	Monyet	Kokopo	
Peripheral damage (logit_PD)	chr05_36340987	5	36,340,987	4.20	0.91	12.59	C	A	RRRA	RR	
	chr05_36341014	5	36,341,014	3.93	0.88	11.66	G	A	RRRA	RR	
	chr06_31421785	6	31,421,785	3.26	0.88	9.27	C	T	AAAA	RA	
	chr06_32076809	6	32,076,809	3.87	0.66	11.47	T	C	AAAA	RA	
	chr06_32727886	6	32,727,886	3.37	0.91	9.83	T	C	AAAA	RA	
	chr06_32727887	6	32,727,887	3.30	0.91	9.55	C	T	AAAA	RA	
	chr06_32728262	6	32,728,262	3.02	-0.56	8.64	C	T	RRRR	RA	
	chr06_33159407	6	33,159,407	3.00	0.71	8.45	A	T	RRRR	RA	
	chr06_33545364	6	33,545,364	4.41	0.63	13.30	A	C	RRRR	RA	
	chr06_33654798	6	33,654,798	3.09	0.74	8.69	G	T	AAAA	RA	
	chr06_33938938	6	33,938,938	4.62	0.69	14.62	A	C	RRRR	RA	
	chr06_34092319	6	34,092,319	3.44	0.82	9.89	G	A	AAAA	RA	
	chr06_34170984	6	34,170,984	4.16	-0.72	13.24	C	G	RRRR	RA	
	chr06_34306968	6	34,306,968	3.23	0.73	9.17	C	T	AAAA	RA	
	chr06_34306976	6	34,306,976	3.79	0.79	11.10	G	C	AAAA	RA	
	chr06_34307019	6	34,307,019	3.32	-0.75	9.49	G	T	RRRR	RA	
	chr06_34317794	6	34,317,794	3.06	0.76	8.56	G	C	AAAA	RA	
	chr06_34437000	6	34,437,000	3.46	-0.66	10.04	T	A	RRRR	RA	
	chr06_34516898	6	34,516,898	3.35	0.75	9.74	A	G	AAAA	RA	
	chr06_34517231	6	34,517,231	3.21	0.72	9.10	G	T	AAAA	RA	
	chr06_34598599	6	34,598,599	3.20	-0.75	9.05	T	A	RRRR	RA	
	chr06_34625789	6	34,625,789	3.84	0.77	11.26	A	C	AAAA	RA	
	chr06_34726334	6	34,726,334	3.45	0.71	10.11	C	A	AAAA	RA	
	chr06_35043294	6	35,043,294	3.06	0.50	8.58	T	C	AAAA	RA	
	chr06_35061540	6	35,061,540	3.27	0.72	9.30	C	T	AAAA	RA	
	chr06_35061548	6	35,061,548	3.26	0.72	9.26	A	G	AAAA	RA	
	chr06_35175016	6	35,175,016	3.55	0.63	10.34	G	A	AAAA	RA	
	chr06_35288600	6	35,288,600	3.33	0.71	9.50	G	T	AAAA	RA	
	chr06_35632650	6	35,632,650	3.60	0.85	10.43	G	A	AAAA	RA	
	chr06_35734979	6	35,734,979	3.12	0.61	8.87	G	A	AAAA	RA	
	chr06_35748038	6	35,748,038	3.31	0.70	9.43	A	T	AAAA	RA	
	chr06_35831106	6	35,831,106	3.06	0.66	8.66	A	G	AAAA	RA	
	chr06_36279592	6	36,279,592	3.23	0.79	9.16	T	G	AAAA	RA	
	chr06_36411988	6	36,411,988	3.00	-0.76	8.38	A	T	RRRR	RA	
	chr06_36421307	6	36,421,307	3.22	-0.79	9.19	G	A	RRRR	RA	
	chr06_36510718	6	36,510,718	3.76	0.73	11.07	T	G	AAAA	RA	
	chr06_37071546	6	37,071,546	3.43	0.76	10.12	C	A	RRRR	RA	
	chr06_37168283	6	37,168,283	3.00	0.52	8.45	A	G	AAAA	RA	
	chr06_37301458	6	37,301,458	3.37	1.07	9.72	A	C	AAAA	RA	
	chr06_37301468	6	37,301,468	3.22	1.06	9.20	G	T	AAAA	RA	
	chr06_37301474	6	37,301,474	3.55	1.11	10.34	T	C	AAAA	RA	
	chr06_37510753	6	37,510,753	3.13	0.74	8.82	T	C	AAAA	RA	
	chr06_37510754	6	37,510,754	3.13	0.74	8.82	G	C	AAAA	RA	
	chr08_5892162	8	5892162	3.29	-1.44	9.46	A	T	RRRR	RA	
	Total cross-section damage (Logit_TXD)	chr06_634567	6	63,4567	3.01	-1.64	8.54	C	G	RRAA	RR
		chr06_634591	6	63,4591	3.06	-1.65	8.73	G	T	RRAA	RR
		chr06_634597	6	63,4,597	3.19	1.75	9.18	A	T	RRAA	RR
chr06_699603		6	699,603	3.31	1.88	9.53	G	C	RRAA	RR	
chr06_880914		6	880,914	4.41	1.81	13.31	A	T	RRAA	RR	

(Continued)

TABLE 3 | (Continued)

Trait	SNP ID	Chr. <sup>1</sup>	Position (bp)	-Log(p)	Effect <sup>2</sup>	Expl. var. (%) <sup>3</sup>	Alleles <sup>4</sup>		Parental genotypes	
							R	A	Monyet	Kokopo
	chr06_2412450	6	2,412,450	3.71	-2.15	10.83	A	C	RRAA	RR
	chr06_2795989	6	2,795,989	3.82	-2.31	11.20	G	A	RRAA	RR
	chr06_2824699	6	2,824,699	3.09	-1.80	8.68	A	G	RRAA	RR
	chr06_3048794	6	3,048,794	3.01	-1.44	8.40	G	T	RRAA	AA
	chr06_3332180	6	3,332,180	3.38	1.33	9.94	C	A	RRAA	RR
	chr06_34681454	6	34,681,454	3.49	-1.50	10.15	G	C	RRAA	RR

<sup>1</sup>Chromosome; <sup>2</sup>additive substitution effect for the reference allele; <sup>3</sup>percentage of explained variance by the SNP marker; <sup>4</sup>R: reference allele, A: alternative allele.

by significant markers was estimated at each SNP site. For PD, it ranged from 8 to 15%, with a peak at chr06\_33938938 on chromosome 6. For TXD, the explained variance ranged from 8 to 13% with a peak at chr06\_880914 on chromosome 6 (Table 3). The analysis estimated the additive effect of replacing one reference allele by the alternative allele at each marker position, with the alternative allele associated with a negative effect on the traits, which translates into resistance at eight SNPs for PD and seven SNPs for TXD (Table 3).

### “Kokopo” Linkage Map and Linkage Quantitative Trait Loci Analysis

To confirm the QTL for logit\_PD identified using continuous mapping on chromosome 6, we used simplex markers segregating AAAA (“Monyet”) × RA (“Kokopo”) to build a linkage map and map the QTL. Out of 4,619 SNPs showing an AAAA × RA parental genotype combination, 816 had less than 10% of missing data and did not show segregation distortion from the expected 1:1 segregation ratio at a significance level of  $\alpha = 0.001$  ( $\chi^2$  test for goodness of fit). The linkage grouping (LG) performed at threshold recombination of 0.20 resulted in 13 linkage groups (LGs) of at least 10 markers comprising 786 SNPs and spanning a total length of 2,343.7 cM. The number of markers per LG ranged from 10 up to 192 markers, the shortest LG being 25 cM and the longest being 521 cM in length. The LGs corresponded well with the banana chromosomes based on the chromosomal origin of the markers on “DH Pahang” (Martin et al., 2016), except for LG 8 containing markers from chromosomes 3 and 8 (Figure 4 and Table 4). Such a pattern suggested a translocation in “Kokopo” from chromosome 3 to 8, relative to the physical map of “DH Pahang.” However, the linkage map does not give any evidence of a reciprocal translocation from chromosome 8 to 3, although some genetic information might have been lost due to segregation distortion, given that chromosome 3 was the smallest, made of 13 markers (Table 4). Sixty percent of the AAAA × RA markers had a segregation distortion at a significance level of 0.05 with a distortion of 0.45 to 0.76 per chromosome (Figure 4B). Nevertheless, chromosomes 3 or 8 did not show particularly higher numbers of distorted markers which could be associated with a chromosomal dislocation, as only 45% of the markers were distorted on each of the two chromosomes at a significance level of 5% (Figure 4B). The translocation from chromosome 3 to 8 was further tested using

linkage disequilibrium (LD) analysis [Genstat® for Windows™ 20<sup>th</sup> Edition (VSN International, 2019)]. LD was determined for the markers on LG 38 (Table 4) in the coupling phase. The pairwise LD and LD decay plots from this analysis indicated a seamless continuation of the markers from chromosome 8 to chromosome 3 on the same LG (Supplementary Figure 2).

The interval QTL mapping followed by a composite interval QTL mapping confirmed the QTL for the PD on chromosome 6 at the significance level of 0.01 [ $-\log(p) = 2$ , Figure 5]. The QTL had a peak at marker chr06\_34317794, which was one of the SNPs identified by continuous mapping to be significantly ( $p < 0.001$ ) associated with PD (Table 3). In contrast, the linkage QTL analysis did not identify any QTL for the total TXD, which can be explained by the fact that “Monyet,” the parent that did not segregate in the linkage map, is the parent contributing to the differences in the TXD phenotype among the segregants (Table 3).

### Putative Genes Associated With Weevil Resistance

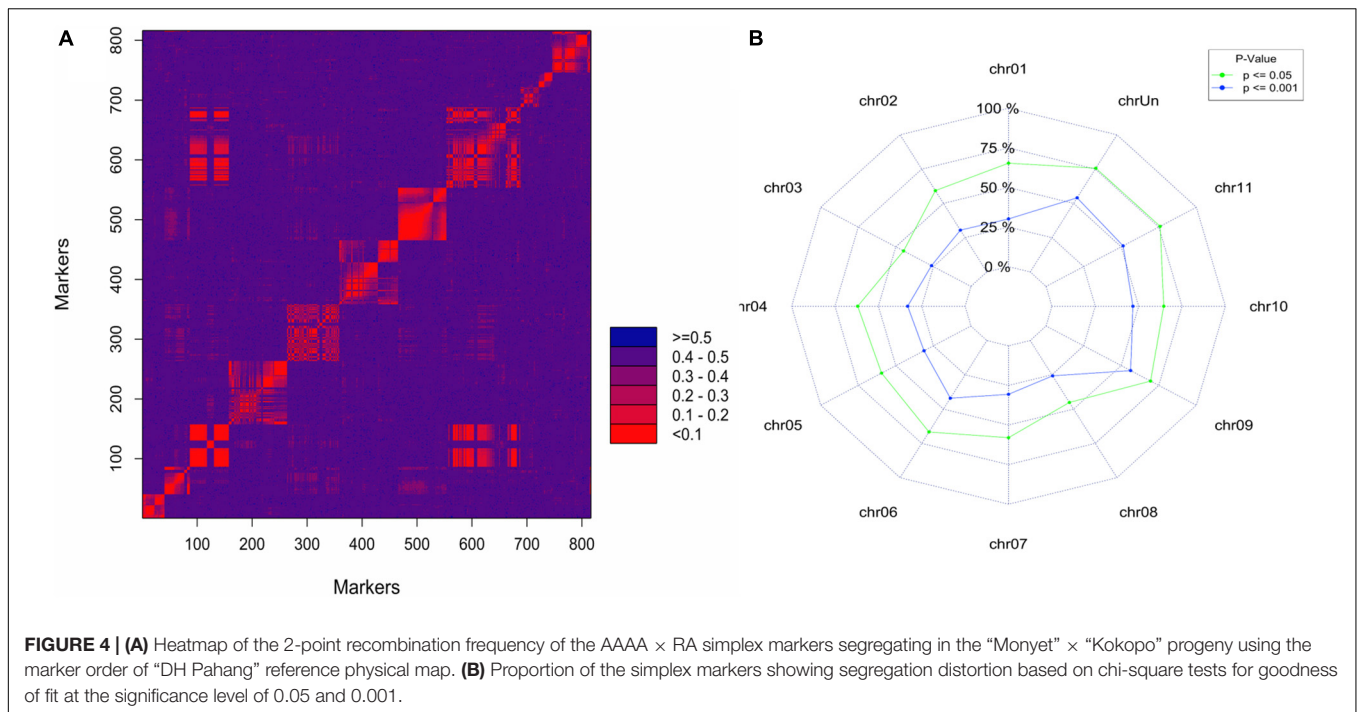
Using the physical position of the SNP significantly associated with the two traits to identify the nearest gene in the annotation file yielded 43 putative genes, with 33 of them associated with PD, 8 associated with TXD, and 1 putative gene shared between the two traits (Supplementary Table 1). Thirty-two of the SNPs were located within the gene region, seven were within less than 1 kbp, and the remaining seven were between 1 and 4.5 kbp from the gene. In terms of biological processes, the identified putative genes associated with the two traits were mostly involved in cellular and metabolic processes (Figure 6 and Supplementary Table 1). Three putative genes were identified as responsive to stimuli. These were *Ma05\_g24110* (LOW QUALITY PROTEIN: transformation/transcription domain-associated protein-like), *Ma06\_g33600* (Calcium-dependent protein kinase 15), and *Ma06\_g38560* (Rho GDP-dissociation inhibitor 1), all in the vicinity of the SNPs significantly associated with PD damage.

## DISCUSSION

### Ploidy Analysis

Ploidy verification is a common practice in banana research because possible discrepancies and unexpected ploidy levels have





**TABLE 4 |** Information on the “Kokopo” genetic linkage map based on simplex AAAA × RA<sup>1</sup> markers segregating in the “Monyet” × “Kokopo” population.

LG <sup>1</sup> number	LG <sup>2</sup> name	Corresponding chromosome <sup>3</sup>	Map length (cM)	Number of markers
1	1	1	140.7	40
2	2	2	130.4	41
3	38	3 and 8	521.3	192
4	3	3	35.2	13
5	4a	4	182.8	71
6	4b	4	124.9	32
7	5	5	302.4	90
8	6	6	367.0	107
9	7	7	226.0	88
10	9	9	64.0	26
11	10a	10	61.9	15
12	10b	10	24.7	10
13	11	11	162.4	60
Total			2343.7	785

<sup>1</sup>R: reference allele, A: alternative allele; <sup>2</sup>LG: linkage group; <sup>3</sup>corresponding chromosome based on “DH Pahang” physical map (Martin et al., 2016).

been reported (Nsabimana and van Staden, 2006; Karamura et al., 2016). The flow cytometry analysis of “Monyet” has shown that it is a 2x banana,<sup>5</sup> though it was first reported as a 4x by Arinaitwe et al. (2019). While ploidy discrepancies arise mostly from an unreliable estimation of ploidy based on morphological features, we hypothesize that the 4x “Monyet” used in this study originated from a genome duplication event of a 2x “Monyet.” This suggestion is supported by the segregation

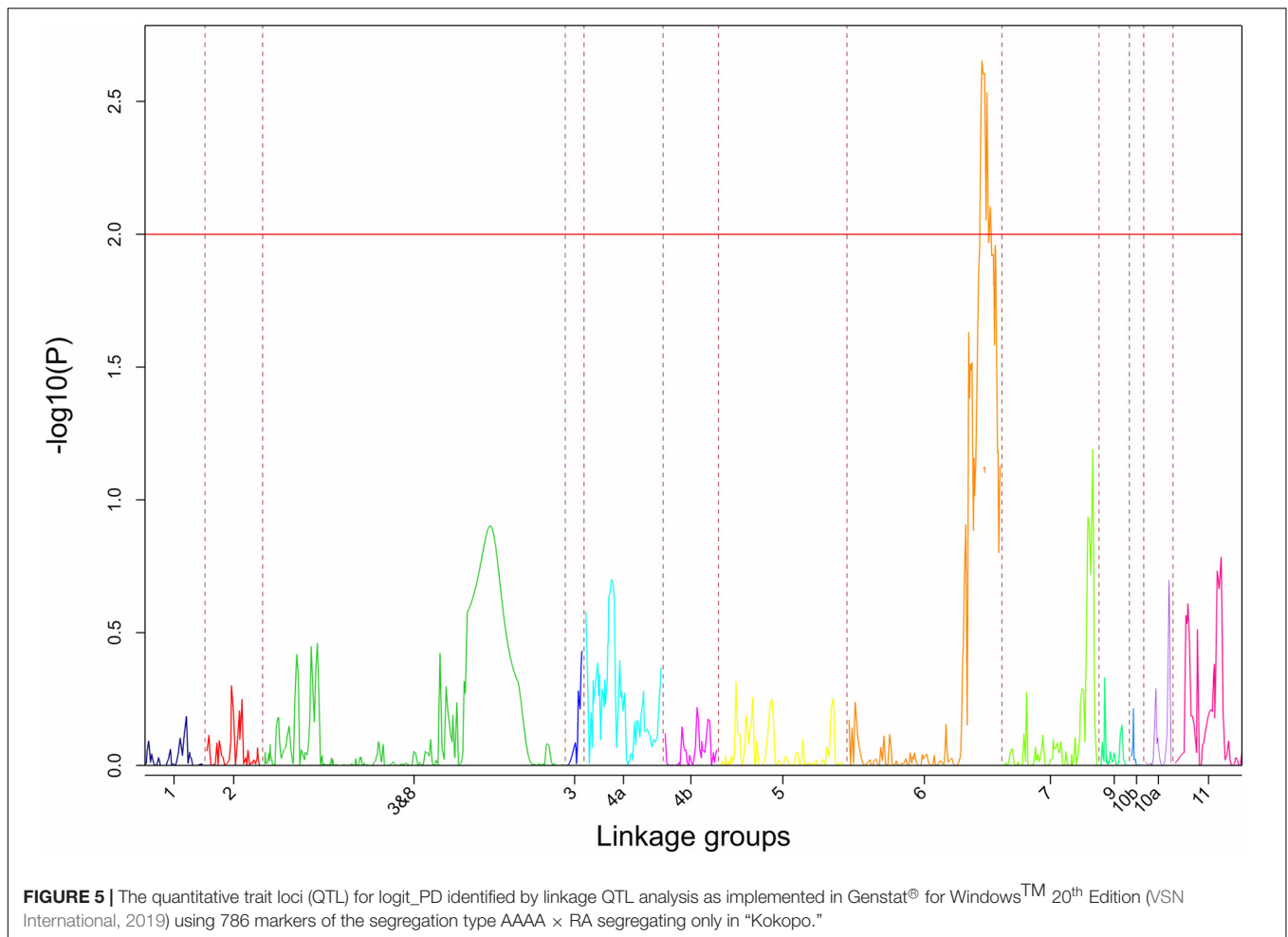
<sup>5</sup><https://www.crop-diversity.org/mgis/accession/01BEL0841179>

of the markers in “Monyet” where nulliplex (RRRR), duplex (RRAA), and quadruplex (AAAA) loci make up 97% of the SNPs (Table 2). However, the 4x “Monyet” plants in the collection at IITA Sendusu reveal no morphological features of the enlarged pseudostem and drooping leaves typical for 4x bananas. As a result of the 4x of “Monyet,” the progeny of “Monyet” × “Kokopo” was 96% 3x, as expected from a cross between 4x and 2x bananas (Brown et al., 2017).

## Resistance to Weevil Damage in the Population

The two EAHB used as susceptible checks (“Nakyatengu” and “Kabucuragye”) in this study were equally susceptible to both PD and TXD, as expected (Kiggundu et al., 2003; Sadik et al., 2010). “Yangambi Km5” and “Calcutta 4,” the two resistant checks recommended to be used in weevil screening (Viljoen et al., 2017), have been reported as equally resistant to the pest using percentage coefficient of infestation, peripheral corm damage, and internal corm damage, scored qualitatively and quantitatively under natural and artificial infestation (Ortiz et al., 1995; Sadik et al., 2010). In this study, we found the two genotypes equally resistant for TXD, but “Calcutta 4” was significantly more resistant than “Yangambi Km5” for PD. Therefore, we strongly suggest “Calcutta 4” to be used as a reference when screening banana genotypes for their response to weevil infestation, especially for new hybrids developed by the banana improvement programs.

While “Monyet” and “Kokopo” were equally affected for PD, they contrasted in terms of cross-section damage, with “Monyet” being more resistant than “Kokopo.” The transgressive segregation displayed in the progeny is in line with the presence



of more than one loci in the repulsion phase in the two parents segregating for the traits (deVicente and Tanksley, 1993). Thus, the population segregated even for PD, although the parents did not contrast for this trait. The resistant genotypes in this study should be further investigated for their male and female fertility, agronomic characteristics, and their response to other pests and diseases common in bananas, such as nematodes, black Sigatoka, fusarium wilt, and banana bacterial wilt. In separate studies, “Monyet” and “Kokopo” have been found contrasting in terms of their response to black Sigatoka (caused by *Pseudocercospora fijiensis*) (Kimunye et al., 2021), banana bacterial wilt (caused by *Xanthomonas vesicola* pv. *musacearum*) (Nakato et al., 2019), and to banana fusarium wilt (caused by *Fusarium oxysporum* f. sp. *cubense* race 1) (Arinaitwe et al., 2019). The segregants from this cross should be tested for possible multiple resistance to these diseases, in addition to resistance to weevils.

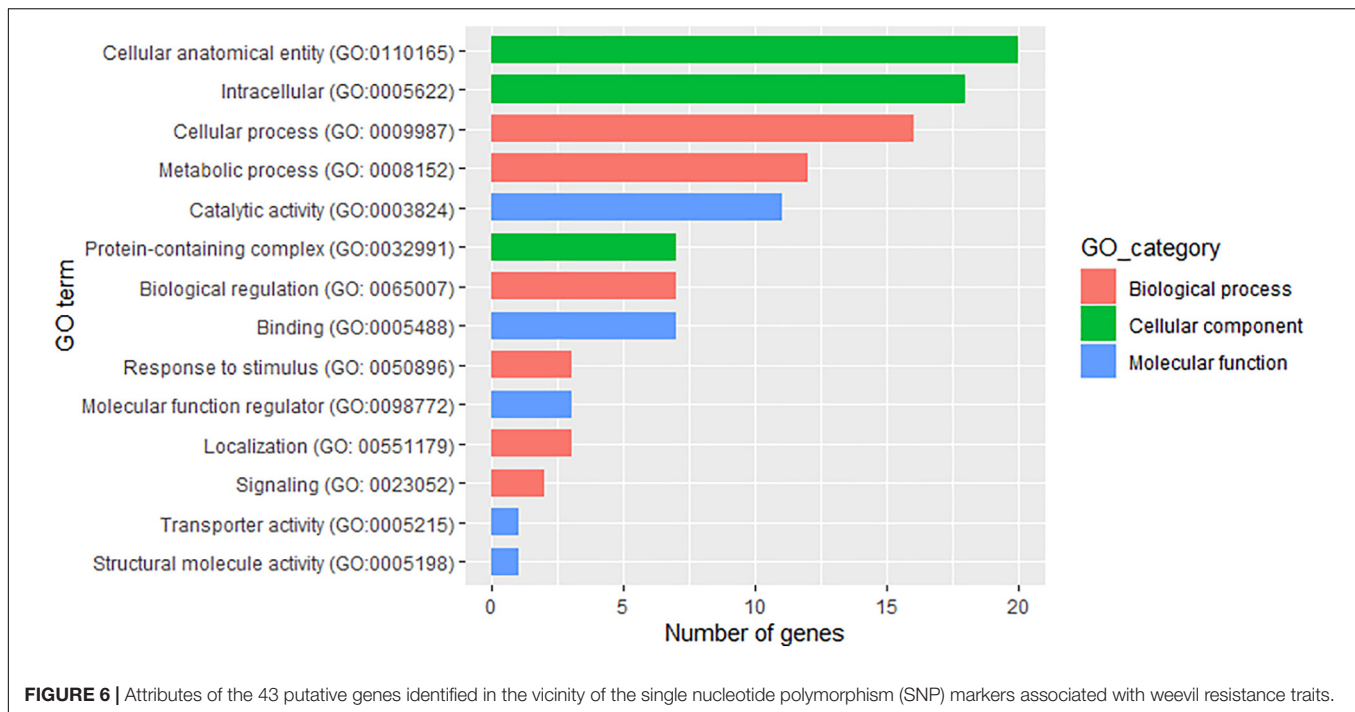
### Quantitative Trait Loci Analysis

We have mapped the loci associated with weevil resistance in a 3x banana population derived from a cross between 4x and 2x genotypes, using DArTSeq SNP markers and an alternative QTL mapping strategy called continuous mapping. To our knowledge, our study is the first to report loci associated with

weevil resistance in bananas, as the previous study that attempted to analyze the genetics for weevil resistance in bananas had solely relied on segregation analysis without the use of markers (Arinaitwe et al., 2015). The latter study suggested that corm damage was under the control of a single dominant gene in a 2x population, while our study identified at least two QTL involved in weevil resistance.

Self-fertilizing 2x species have been the subject of most developments for genetic and genomic tools for decades, at the expense of non-conventional species that are polyploids or out-crossing, or both (Bourke et al., 2018b). With the advances in sequencing and genotyping, tools for polyploids have raised interest (Bourke et al., 2018b). Still, most of the tools developed for polyploids apply for even ploidy levels, especially 4x which make up the majority of the polyploid species (Comai, 2005; Ferrão et al., 2021; Garreta et al., 2021); in contrast, resources to construct linkage maps and perform QTL analyses in odd ploidy levels are limited.

Bourke et al. (2018a) developed the polymapR software, which is a powerful tool for the construction of linkage maps in populations of all ploidy levels with allele dosage genotype calls, including populations from 4x × 2x crosses. However, the robustness of this tool relies on the accurate calling of the allele



dosage genotypes. The population in this study was genotyped at a low sequencing depth as carried out by DArTSeq, which is often associated with genotype miscalling in polyploids. These hitches were circumvented by using the physical positions of the segregating markers instead and regressing the phenotype on the allele frequencies, called continuous mapping (Ashraf et al., 2014). So far, only a few studies have applied continuous mapping successfully in polyploid genetic and genomic analysis with sequence data of low sequencing depth to counter the effect of low-depth sequencing on genotype calling in polyploids (Sverrisdóttir et al., 2017; Zheng et al., 2020). In this study, we identified the QTL for PD and TXD by using continuous mapping. One of the QTL, segregating in the 2x parent “Kokopo” could be confirmed using the conventional QTL mapping method based on linkage analysis. However, the QTL linkage analysis used only a subset of the data that conformed to the conventional segregation of a BC<sub>1</sub> population. Continuous mapping allows the use of all the markers irrespective of the segregation type of the parents, hence encompassing the optimum genetic information of the population.

One of the major drawbacks of continuous mapping is the reliance on an existing grouping and order of the markers in the form of a linkage map or a physical map. In our study, a physical map of bananas based on the “DH Pahang” reference genome (Martin et al., 2016) made the study possible. Also, continuous mapping results do not give information about the genetic inheritance pattern of the trait, such as the effect of dominance, epistasis, and linkage disequilibrium as a result of linkage. While continuous mapping does not take into consideration the linkage disequilibrium among the markers intrinsically, the significance of contiguous markers on the same chromosome gives confidence about the presence of a QTL at a given position. This was the case for the markers on chromosome 6. Therefore, the markers

on chromosomes 5 and 8 associated with PD could be discrepant since the flanking markers are not significantly associated with the traits. We expect continuous mapping to be of great advantage in QTL mapping studies in bananas, especially for consumer-preferred quality traits in 3x banana hybrids derived from 4x by 2x crosses as developed in EAHBs and plantains.

The identified loci associated with the resistance to weevil damage on chromosomes 5, 6, and 8 do not coincide with any known QTL for resistance to pests and diseases in *Musa*. This is mainly because marker-trait association in bananas is still in its infancy stage (Ahmad et al., 2020), although the advances in deciphering *Musa* genomes are promising to mend the situation.<sup>6</sup> Ten percent or more of corm damage translates into a bunch yield reduction of 20–45% (Rukazambuga et al., 1998). Therefore, despite the low explained variance by the identified QTL, the additive allele substitution effect at each locus as determined in this study has potential in banana breeding, especially at a polyploid level. Moreover, this knowledge on allelic effect instead of the genotype effect will facilitate the validation of the QTL in other genetic backgrounds.

## Putative Genes for Weevil Resistance

Weevil resistance in bananas has been hypothesized to be a result of antibiosis (detering the weevils), antixenosis (non-preference), and corm hardness. Antibiosis was not supported on excised plants (Night et al., 2010) and corm hardness was found uncorrelated with field resistance (Ortiz et al., 1995). Although we cannot conclude on the mechanism of resistance in this study, our results suggest that the resistance for PD differs from the one for cross-section. Despite the strong positive

<sup>6</sup><https://banana-genome-hub.southgreen.fr/>

correlation between the two traits, PD segregated in “Kokopo” while TXD segregated solely in “Monyet.” The two traits were associated with the markers on chromosome 6, but none of the significant markers was shared among them. Although one marker associated with TXD (chr06\_34681454) was in the same vicinity as the markers associated with PD (chr06\_34625789 and chr06\_34726334, **Table 3**), the three SNPs were not located within or close to the same putative genes (**Supplementary Table 1**). Therefore, breeding for weevil resistance in bananas should consider the two traits independently. We identified 43 putative genes co-localized with weevil resistance SNPs. The genes of interest for weevil resistance would be those associated with plant defense mechanisms. The three putative genes responsive to stimuli, *Ma05\_g24110* (LOW QUALITY PROTEIN: transformation/transcription domain-associated protein-like), *Ma06\_g33600* (Calcium-dependent protein kinase 15), and *Ma06\_g38560* (Rho GDP-dissociation inhibitor 1) fall under this category. Among them, *Ma06\_g33600* belongs to the Calcium-dependent protein kinase gene family (CDPKs) already known to play a vital role in the plant defense against biotic and abiotic stress (Dubiella et al., 2013; Schulz et al., 2013). Likewise, *Ma06\_g34300*, a putative wound-induced protein co-localizing with the SNP significantly associated with PD, is of interest, as it is putatively involved in wound repair, which might be useful to corm damage due to weevil larvae. This putative gene has a sequence identity closely similar to *LOC4332560* identified in *Oryza sativa* L. ssp. *japonica* (Kikuchi et al., 2003). The regulation of these genes should be further investigated in known resistant vs. susceptible banana genotypes to determine their potential functionality concerning the response of bananas to *C. sordidus*.

## Chromosomal Translocation

Chromosomal translocations are common in bananas and have played an important role in the evolution and speciation of the *Musa* genome (Martin et al., 2017, 2020a; Šimoníková et al., 2020). Chromosome 8 has been reported in many translocation events in various *M. acuminata* spp. (Dupouy et al., 2019). The linkage analysis of the simplex AAAA × RA markers segregating in “Kokopo” suggested a translocation from chromosome 3 to chromosome 8, relative to the “DH Pahang” physical map. The translocation was identified through the construction of a “Kokopo” linkage map and confirmed by LD analysis. Such a translocation has been reported in *M. acuminata* spp. *zebrina* and *M. acuminata* ssp. *burmannica* (Šimoníková et al., 2020), but not in *M. acuminata* spp. *banksii*. Unlike the above-mentioned studies, our data did not give evidence of a reciprocal translocation from chromosome 8 to chromosome 3. This finding suggests that more translocations are yet to be discovered in *Musa*. The segregation distortion of the AAAA × RA markers

## REFERENCES

- Ahmad, F., Martawi, N. M., Poerba, Y. S., de Jong, H., Schouten, H., and Kema, G. H. J. (2020). Genetic mapping of *Fusarium* wilt resistance in a wild banana *Musa acuminata* ssp. *malaccensis* accession. *Theor. Appl. Genet.* 133, 3409–3418. doi: 10.1007/s00122-020-03677-y
- Arinaitwe, I. K., Barekye, A., Kubiriba, J., Sadik, K., Karamura, E., and Edema, R. (2015). Genetic analysis of weevil (*Cosmopolites sordidus*) resistance in an F2 diploid banana population. *J. Plant Breed. Genet.* 3, 77–91.
- Arinaitwe, I. K., Teo, C. H., Kayat, F., Tumuhimbise, R., Uwimana, B., Kubiriba, J., et al. (2019). Evaluation of banana germplasm and genetic analysis of an F1 population for resistance to *Fusarium oxysporum* f. sp. *cubense* race 1. *Euphytica* 215, 175–186. doi: 10.1007/s10681-019-2493-3

was pronounced with 60% of the genome-wide markers distorted at a significance level of 0.05. However, there was no indication of a particular higher segregation distortion for the markers on chromosomes 3 or 8 as compared with the rest of the chromosomes, which would be associated with the translocation.

## DATA AVAILABILITY STATEMENT

The phenotypic and genotypic data in this study are publicly available on Musabase: Phenotypic data: [https://musabase.org/breeders/search?dataset\\_id=21](https://musabase.org/breeders/search?dataset_id=21) (the user should click on “Related Trial Phenotypes” and then click on the “Download Phenotypes”). Genotypic data: [https://musabase.org/breeders\\_toolbox/protocol/12](https://musabase.org/breeders_toolbox/protocol/12) (the user should expand the “Genotype Data” section and then click on “Download All Genotype Data VCF”).

## AUTHOR CONTRIBUTIONS

BU, MB, and RS conceived and designed the study. BU, MB, GM, and VA investigated while BU, MV, and TS curated and analyzed the data. MB, VA, BU, and RS provided the resources. BU and RS acquired the funds. MB, BU, MV, and TS provided the methodology. RS supervised the study. BU, MV, and TS wrote the original draft. BU, MB, GM, VA, TS, MV, and RS reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

## FUNDING

This study was financially supported by the Bill and Melinda Gates Foundation under grant number OPP1213871.

## ACKNOWLEDGMENTS

We would like to thank Moses Nyine for his contribution to the bioinformatic analysis of the sequence data.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2021.753241/full#supplementary-material>



- Ashraf, B. H., Jensen, J., Asp, T., and Janss, L. L. (2014). Association studies using family pools of outcrossing crops based on allele-frequency estimates from DNA sequencing. *Theor. Appl. Genet.* 127, 1331–1341. doi: 10.1007/s00122-014-2300-4
- Batte, M., Swennen, R., Uwimana, B., Akech, V., Brown, A., Tumuhimbise, R., et al. (2019). Crossbreeding East African Highland Bananas: lessons learnt relevant to the botany of the crop after 21 years of genetic enhancement. *Front. Plant Sci.* 10:81. doi: 10.3389/fpls.2019.00081
- Blomme, G., Ploetz, R., Jones, D., De Langhe, E., Price, N., Gold, C., et al. (2013). A historical overview of the appearance and spread of *Musa* pests and pathogens on the African continent: highlighting the importance of clean *Musa* planting materials and quarantine measures. *Ann. Appl. Biol.* 162, 4–26. doi: 10.1111/aab.12002
- Bourke, P. M., Voorrips, R. E., Visser, R. G. F., and Maliepaard, C. (2018b). Tools for genetic studies in experimental populations of polyploids. *Front. Plant Sci.* 9:513. doi: 10.3389/fpls.2018.00513
- Bourke, P. M., van Geest, G., Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., et al. (2018a). polypmapR—linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. *Bioinformatics* 34, 3496–3502. doi: 10.1093/bioinformatics/bty371
- Brown, A., Tumuhimbise, R., Amah, D., Uwimana, B., Nyine, M., Mduma, H., et al. (2017). “Bananas and plantains (*Musa* spp.),” in *Genetic Improvement of Tropical Crops*, eds H. Campos and P. D. S. Caligari (Cham: Springer), 219–240.
- Christelová, P., De Langhe, E., Hřibová, E., Čížková, J., Sardos, J., Hušáková, M., et al. (2016). Molecular and cytological characterization of the global *Musa* germplasm collection provides insights into the treasure of banana diversity. *Biodivers. Conserv.* 26, 801–824. doi: 10.1007/s10531-016-1273-9
- Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* 6, 836–846. doi: 10.1038/nrg1711
- Coyne, D., Wasukira, A., Dusabe, J., Rotifa, I., and Dubois, T. (2010). Boiling water treatment: a simple, rapid and effective technique for nematode and banana weevil management in banana and plantain (*Musa* spp.) planting material. *Crop Protect.* 29, 1478–1482. doi: 10.1016/j.cropro.2010.08.008
- deVicente, M. C., and Tanksley, S. D. (1993). QTL analysis of transgressive segregation in an interspecific tomato cross. *Genetics* 134, 585–596. doi: 10.1093/genetics/134.2.585
- Dépigny, S., Noupadja, P., Tomekpe, K., Bonneviot, P., Lescot, T., and Bakry, F. (2018). “CARBAP K74”: a triploid plantain-like hybrid designed to promote sustainable plantain-based cropping systems. *Acta Hort.* 1196, 63–70. doi: 10.17660/ActaHortic.2018.1196.7
- D’Hont, A., Denoeud, F., Aury, J.-M., Baurens, F.-C., Carreel, F., Garsmeur, O., et al. (2012). The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. *Nature* 488, 213–217. doi: 10.1038/nature11241
- Dolezel, J., Lysák, M. A., Houwe, I. V. D., Dolezelová, M., and Roux, N. (1997). Use of flow cytometry for rapid ploidy determination in *Musa* species. *InfoMusa* 6, 6–9.
- Dubiella, U., Seybold, H., Durian, G., Komander, E., Lassig, R., Witte, C.-P., et al. (2013). Calcium-dependent protein kinase/NADPH oxidase activation circuit is required for rapid defense signal propagation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8744–8749. doi: 10.1073/pnas.1221294110
- Dupouy, M., Baurens, F.-C., Derouault, P., Hervouet, C., Cardi, C., Cruaud, C., et al. (2019). Two large reciprocal translocations characterized in the disease resistance-rich burmannica genetic group of *Musa acuminata*. *Ann. Bot.* 20, 1–11. doi: 10.1093/aob/mcz078
- Fauré, S., Noyer, J. L., Horry, J. P., Bakry, F., Lanaud, C., and González de León, D. (1993). A molecular marker-based linkage map of diploid bananas (*Musa acuminata*). *Theor. Appl. Genet.* 87, 517–526. doi: 10.1007/bf00215098
- Ferrão, L. F. V., Amadeu, R. R., Benevenuto, J., de Bem Oliveira, I., and Munoz, P. R. (2021). Genomic selection in an outcrossing autotetraploid fruit crop: lessons from blueberry breeding. *Front. Plant Sci.* 12:676326. doi: 10.3389/fpls.2021.676326
- Garreta, L., Cerón-Souza, I., Palacio, M. R., and Reyes-Herrera, P. H. (2021). MultiGWAS: an integrative tool for Genome Wide Association Studies (GWAS) in tetraploid organisms. *Ecol. Evol.* 11, 7411–7426. doi: 10.1101/2020.08.16.252791
- Gold, C. S., and Bagabe, M. I. (1997). Banana weevil, *Cosmopolites sordidus* Germar (Coleoptera: Curculionidae), infestations of cooking- and beer-bananas in adjacent plantations in Uganda. *Afr. Entomol.* 5, 103–108.
- Gold, C. S., Karamura, E. B., Kiggundu, A., Bagamba, F., and Abera, A. M. K. (1999). Geographic shifts in the highland cooking banana (*Musa* spp., group AAA-EA) production in Uganda. *Int. J. Sustain. Dev. World Ecol.* 6, 45–59. doi: 10.1080/13504509.1999.9728471
- Gold, C. S., and Messiaen, S. (2000). “Le charançon du bananier *Cosmopolites sordidus*,” in *Parasites et RAVAGEURS DES MUSA*, eds J. Sarah, J. Pinochet, and J. Stanton (Montpellier: INIBAP).
- Gold, C. S., Pena, J. E., and Karamura, E. B. (2001). Biology and integrated pest management for the banana weevil *Cosmopolites sordidus* (Germar) (Coleoptera: Curculionidae). *Integr. Pest Manag. Rev.* 6, 79–155. doi: 10.1023/A:1023330900707
- Gold, S., Kagezi, G. H., Night, G., and Ragama, P. E. (2004). The effects of banana weevil, *Cosmopolites sordidus*, damage on highland banana growth, yield and stand duration in Uganda. *Ann. appl. Biol.* 145, 263–269.
- Jones, D. R. (2009). Disease and pest constraints to banana production. *Acta Hort.* 828, 21–36. doi: 10.17660/ActaHortic.2009.828.1
- Karamura, D., Tumuhimbise, R., Muhangi, S., Nyine, M., Pillay, M., Reuben, S. T., et al. (2016). Ploidy level of the banana (*Musa* spp.) accessions at the germplasm collection centre for the East and Central Africa. *Afr. J. Biotechnol.* 15, 1692–1698. doi: 10.5897/ajb2016.15442
- Kiggundu, A., Gold, C. S., Labuschagne, M. T., Vuylsteke, D., and Louw, S. (2003). Levels of host plant resistance to banana weevil, *Cosmopolites sordidus* (Germar) (Coleoptera: Curculionidae), in African *Musa* germplasm. *Euphytica* 133, 267–277. doi: 10.1023/a:1025725124849
- Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., et al. (2003). Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* 301, 376–379. doi: 10.1126/science.1081288
- Kimunye, J., Were, E., Swennen, R., Viljoen, A., and Mahuku, G. (2021). Sources of resistance to *Pseudocercospora fijiensis*, the cause of black Sigatoka in banana. *Plant Pathol.* 70, 1651–1664. doi: 10.1111/ppa.13408
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lorenzen, J., Hearne, S., Mbanjo, G., Nyine, M., and Close, T. (2011). Use of molecular markers in banana and plantain improvement. *Acta Hort.* 897, 231–236. doi: 10.17660/ActaHortic.2011.897.25
- Lorenzen, J., Tenkouano, A., Bandyopadhyay, R., Vroh, B., Coyne, D., and Tripathi, L. (2010). Overview of banana and plantain (*Musa* spp.) improvement in africa: past and future. *Acta Hort.* 879, 595–603.
- Martin, G., Baurens, F.-C., Droc, G., Rouard, M., Cenci, A., Kilian, A., et al. (2016). Improvement of the banana “*Musa acuminata*” reference sequence using NGS data and semi-automated bioinformatics methods. *BMC Genomics* 17:243. doi: 10.1186/s12864-016-2579-4
- Martin, G., Cardi, C., Sarah, G., Ricci, S., Jenny, C., Fondi, E., et al. (2020b). Genome ancestry mosaics reveal multiple and cryptic contributors to cultivated banana. *Plant J.* 102, 1008–1025. doi: 10.1111/tpj.14683
- Martin, G., Baurens, F.-C., Hervouet, C., Salmon, F., Delos, J.-M., Labadie, K., et al. (2020a). Chromosome reciprocal translocations have accompanied subspecies evolution in bananas. *Plant J.* 104, 1698–1711. doi: 10.1111/tpj.15031
- Martin, G., Carreel, F., Coriton, O., Hervouet, C., Cardi, C., Derouault, P., et al. (2017). Evolution of the banana genome (*Musa acuminata*) is impacted by large chromosomal translocations. *Mol. Biol. Evol.* 34, 2140–2152. doi: 10.1093/molbev/msx164
- Mbwana, A. S. S., and Rukazambuga, N. D. T. M. (1998). “Banana IPM in Tanzania,” in *Mobilizing IPM for Sustainable Banana Production in Africa*, eds E. Frison, C. S. Gold, E. B. Karamura, and R. A. Sikora (Montpellier: INIBAP), 237–245.
- Mi, H., Ebert, D., Muruganujan, A., Mills, C., Albuo, L.-P., Mushayamaha, T., et al. (2020). PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res.* 49, D394–D403. doi: 10.1093/nar/gkaa1106
- Nakato, G. V., Christelová, P., Were, E., Nyine, M., Coutinho, T. A., Doležel, J., et al. (2019). Sources of resistance in *Musa* to *Xanthomonas campestris* pv. musacearum, the causal agent of banana xanthomonas wilt. *Plant Pathol.* 68, 49–59. doi: 10.1111/ppa.12945
- Night, G., Gold, C. S., and Power, A. G. (2010). Survivorship and development rates of banana weevils reared on excised plant material of different banana cultivars. *Int. J. Trop. Insect Sci.* 30, 77–83. doi: 10.1017/S1742758410000111

- Nsabimana, A., and van Staden, J. (2006). Ploidy investigation of bananas (*Musa* spp.) from the National Banana Germplasm Collection at Rubona–Rwanda by flow cytometry. *S. Afr. J. Bot.* 72, 302–305. doi: 10.1016/j.sajb.2005.10.004
- Nyine, M., Uwimana, B., Akech, V., Brown, A., Ortiz, R., Doležel, J., et al. (2019). Association genetics of bunch weight and its component traits in East African highland banana (*Musa* spp. AAA group). *Theor. Appl. Genet.* 132, 3295–3308. doi: 10.1007/s00122-019-03425-x
- Nyine, M., Uwimana, B., Blavet, N., Høibová, E., Vanrespaille, H., Batte, M., et al. (2018). Genomic prediction in a multiploid crop: genotype by environment interaction and allele dosage effects on predictive ability in banana. *Plant Genome* 11:170090. doi: 10.3835/plantgenome2017.10.0090
- Ortiz, R., and Swennen, R. (2014). From crossbreeding to biotechnology-facilitated improvement of banana and plantain. *Biotechnol. Adv.* 32, 158–169. doi: 10.1016/j.biotechadv.2013.09.010
- Ortiz, R., Vuylsteke, D., Dumpe, B., and Ferris, R. S. B. (1995). Banana weevil resistance and corm hardness in *Musa germplasm*. *Euphytica* 86, 95–102.
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Rouard, M., Droc, G., Martin, G., Sardos, J., Hueber, Y., Guignon, V., et al. (2018). Three new genome assemblies support a rapid radiation in *Musa acuminata* (wild banana). *Genome Biol. Evol.* 10, 3129–3140. doi: 10.1093/gbe/evy227
- Rukazambuga, N. D. T. M., Gold, C. S., and Gowen, S. R. (1998). Yield loss in East African highland banana (*Musa* spp., AAA-EA group) caused by the banana weevil, *Cosmopolites sordidus* Germar. *Crop Protect.* 17, 581–589. doi: 10.1016/S0261-2194(98)00056-8
- Sadik, K., Nyine, M., and Pillay, M. (2010). A screening method for banana weevil (*Cosmopolites sordidus* Germar) resistance using reference genotypes. *Afr. J. Biotechnol.* 9, 4725–4730.
- Sardos, J., Rouard, M., Hueber, Y., Cenci, A., Hyma, K. E., van den Houwe, I., et al. (2016). A Genome-wide association study on the seedless phenotype in banana (*Musa* spp.) reveals the potential of a selected panel to detect candidate genes in a vegetatively propagated crop. *PLoS One* 11:e0154448. doi: 10.1371/journal.pone.0154448
- Schulz, P., Herde, M., and Romeis, T. (2013). Calcium-dependent protein kinases: hubs in plant stress signaling and development. *Plant Physiol.* 163, 523–530. doi: 10.1104/pp.113.222539
- Šimoníková, D., Němčková, A., Čížková, J., Brown, A., Swennen, R., Doležel, J., et al. (2020). Chromosome painting in cultivated bananas and their wild relatives (*Musa* spp.) reveals differences in chromosome structure. *Int. J. Mol. Sci.* 21:7915.
- Sverrisdóttir, E., Byrne, S., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., et al. (2017). Genomic prediction of starch content and chipping quality in tetraploid potato using genotyping-by-sequencing. *Theor. Appl. Genet.* 130, 2091–2108. doi: 10.1007/s00122-017-2944-y
- Swennen, R., Blomme, G., Van Asten, P., Lepoint, P., Karamura, E., Njukwe, E., et al. (2013). “Mitigating the impact of biotic constraints to build resilient banana systems in Central and Eastern Africa,” in *Agro-Ecological Intensification of Agricultural Systems in the African Highlands*, eds B. Vanlauwe, P. Van Asten, and G. Blomme (New York, NY: Routledge), 85–104.
- Tenkouano, A., Lamien, N., Agogbua, J., Amah, D., Swennen, R., Traore, S., et al. (2019). Promising high-yielding tetraploid plantain-bred hybrids in West Africa. *Int. J. Agron.* 2019:3873198.
- Tumuhimbise, R., Buregyeya, H., Barekye, A., Ssali, R. T., Talengera, D., Kubiriba, J., et al. (2016). Selection of cooking banana genotypes for yield and black Sigatoka resistance in different locations in Uganda. *J. Plant Breed. Crop Sci.* 8, 60–71. doi: 10.5897/JPBBS2016.0559
- Tushemereirwe, W., Batte, M., Nyine, M., Tumuhimbise, R., Barekye, A., Tendo, S., et al. (2015). *Performance of NARITA Banana Hybrids in The Preliminary Yield Trial for Three Cycles in Uganda*. Kampala: ITA, NARO.
- Uitdewilligen, J. G., Wolters, A.-M., D’hoop, B. B., Borm, T. J., Visser, R. G., and van Eck, H. J. (2013). A Next-generation sequencing method for genotyping-by-sequencing of highly heterozygous autotetraploid potato. *PLoS One* 8:e62355. doi: 10.1371/journal.pone.0062355
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., et al. (2013). From fastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Prot. Bioinform.* 43, 11.10.11–11.10.33. doi: 10.1002/0471250953.bi1110s43
- Vilardbo, A. (1973). Le coefficient d’infestation, critere d’évaluation du degre d’attaques des bananeraies par *Cosmopolites sordidus* Germ. le charancon noir du bananier. *Fruits* 28, 417–431.
- Viljoen, A., Mahuku, G., Massawe, C., Ssali, R. T., Kimunye, J., Mostert, G., et al. (2017). *Banana Pests and Diseases Field Guide for Disease Diagnostics and Data Collection Improvement*. Ibadan: IITA.
- VSN International (2019). *Genstat for Windows*, 18th Edn. Hemel Hempstead: VSN International.
- Zheng, W., Shen, F., Wang, W., Wu, B., Wang, X., Xiao, C., et al. (2020). Quantitative trait loci-based genomics-assisted prediction for the degree of apple fruit cover color. *Plant Genome* 13:e20047. doi: 10.1002/tpg2.20047
- Zimmerman, E. C. (1968). The *Cosmopolites banana* weevils (Coleoptera: Curculionidae; Rhynchophorinae). *Pac. Insects* 10, 295–299.

**Conflict of Interest:** MV was employed by the company Gnomixx B.V.

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

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Uwimana, Mwanje, Batte, Akech, Shah, Vuylsteke and Swennen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.