BioEnergy Research (2022) 15:755–776
https://doi.org/10.1007/s12155-022-10402-8

Linkage Mapping of Biomass Production and Composition Traits in a *Miscanthus sinensis* Population

Raphaël Raverdy[1] · Kristelle Lourgant[1] · Emilie Mignot[2] · Stéphanie Arnoult[2] · Guillaume Bodineau[3] · Yves Griveau[4] · Cristiane H. Taniguti[5] · Maryse Brancourt-Hulmel[1]

Received: 23 August 2021 / Accepted: 24 January 2022 / Published online: 10 March 2022

Abstract

Breeding miscanthus for biomass production and composition is essential for targeting high-yielding genotypes suited to different end-uses. Our objective was to understand the genetic basis of these traits in *M. sinensis*, according to different plant ages and environmental conditions. A diploid population was established in two locations according to a staggered-start design, which distinguished the plant age effect from climatic condition effect. An integrated genetic map of 2602 SNP markers distributed across 19 LGs was aligned with the *M. sinensis* reference genome and spanned 2770 cM. The QTL mapping was based on best linear unbiased predictions estimated across three climatic conditions and at least three ages in both locations. A total of 260 and 283 QTL were related to biomass production and composition traits, respectively. In each location, 40–60% were related to biomass production traits and stable across different climatic conditions and ages and 30% to biomass composition traits. Twelve QTL clusters were established based on either biomass production or composition traits and validated by high genetic correlations between the traits. Sixty-two putative *M. sinensis* genes, related to the cell wall, were evidenced in the QTL clusters of biomass composition traits and orthologous to those of sorghum and maize. Twelve of them were differentially expressed and belonged to gene families related to the cell wall biosynthesis identified in other miscanthus studies. These stable QTL constitute new insights into marker-assisted selection (MAS) breeding while offering a joint improvement of biomass production and composition traits.

Keywords Integrated genetic map · QTL mapping · Orthologous genes · Age effect · Climatic condition effect

Introduction

Miscanthus is a perennial C4 crop that produces valuable lignocellulosic biomass, mainly for bioenergy end-uses, bio-based products, and animal bedding [1–5]. However, only one clone of a *Miscanthus* × *giganteus* interspecific hybrid ($2n = 3x = 57$) is currently available for commercialization in Europe and the USA [6, 7]. It originates from a natural interspecific cross between a tetraploid *M. sacchariflorus* ($2n = 4x = 76$) and a diploid *M. sinensis* ($2n = 2x = 38$) [8] and obtains high biomass yields under various environmental conditions [9, 10]. As *M.* × *giganteus* is sterile, the potential risk of invasiveness from the seeds spreading to a new environment is avoided, but this hampers the breeding of the crop [11, 12]. Moreover, such narrow genetic variability can be risky in case of pest adaptation, and the corresponding phenotypic variability may not be sufficient for the different end-use requirements [8]. The two parents of *M.* × *giganteus,* originating from East Asia, present high genetic variability and are adapted to extended environmental conditions [13–16]. For example, *M. sinensis* reaches the same amount of biomass production as *M.* × *giganteus* under specific conditions [17]: this makes it a relevant candidate for breeding new varieties at an intraspecific level, or for creating new *M.* × *giganteus* clones at an interspecific level,

✉ Maryse Brancourt-Hulmel
maryse.hulmel@inrae.fr

1 BioEcoAgro Joint Research Unit, INRAE, University of Liège, University of Lille, University of Picardie Jules Verne, 80203 Estrées-Mons, France

2 GCIE-Picardie, INRAE, 80203 Estrées-Mons, France

3 GBFOr, INRAE, UE, Ardon, France

4 Institut Jean-Pierre Bourgin, INRAE, University of Paris-Saclay, 78000 Versailles, AgroParisTech, France

5 Department of Genetics, Luiz de Queiroz College of Agriculture, University of São Paulo, Piracicaba, Brazil

despite its self-incompatibility [18]. Breeding miscanthus aims to improve biomass production and composition traits by creating new cultivars, adapted to a range of different environments. However, the optimal time to collect reliable phenotypic data for these traits is a major point of concern in miscanthus breeding. Indeed, the yield-building phase of the plants can last 3 years after crop establishment, and even this can vary between miscanthus species [19]. The plants then reach a yield plateau phase during which biomass production is more stable [20]. Selecting plants only on their phenotype during the yield-building phase may thus be unreliable (Raverdy et al., submitted to BioEnergy Research), especially since the different progenies may reach their full growth potential at different times. Marker-assisted selection programs would thus be a helpful tool for improving the breeding efficiency in miscanthus, as they would make it possible to select young plants on the basis of their genetic information, without waiting for them to reach the yield-plateau phase.

Because of its self-incompatibility, miscanthus is an outcrossing species with a high level of heterozygosity. Therefore, the genetic mapping methods that were initially developed based on inbred line populations have to be adapted to full-sib (F1) populations, as done for other perennial crops such as sugarcane or rubber tree [21–23]. Full-sib diploid populations have a maximum of four different segregating alleles per locus. This makes the haplotype phase estimation more complex than for inbred lines populations, for which a maximum of two alleles segregate.

The initial traditional method for building genetic maps in full-sib populations was the "pseudo-testcross" strategy [24], which was based on separate linkage maps for each parent. Later, Wu et al. [25, 26] developed a method that generates an integrated genetic map and for which the estimation of linkage distances and phases between markers and locations of the quantitative trait loci (QTL) are improved. Moreover, Gazaffi et al. [27] proposed a method based on composite interval mapping and multipoint genetic mapping using molecular markers with different segregation patterns. This last method can reveal QTL that segregates in any pattern and identify dominance effect.

The initial miscanthus genetic map was developed by Atienza et al. [28] based on an *M. sinensis* population, by using random amplified polymorphic DNA (RAPD) markers, and was composed of 28 linkage groups (LGs). This unsaturated map was used to detect QTL for morphological traits, biomass yield, and combustion quality [29–32]: however, the 28 LGs that constitute the genetic map, compared to the base chromosome number in miscanthus ($x = 19$), as well as the small population size of 89 F1 individuals, make it difficult to compare the QTL with recent studies in which saturated maps are presented with the expected number of 19 LGs. Ma et al. [33] created a high-resolution linkage map of *M. sinensis* and identified 19 LGs for the first time. Gifford et al. [34] identified 72 QTL associated with biomass productivity, by using the integrated genetic map of *M. sinensis* constructed by Swaminathan et al. [35]. It was also made up of 19 LGs. Dong et al. [36] developed six high-density parental genetic maps using the pseudo-testcross strategy and two consensus maps which integrated *M. sinensis* and *M. sacchariflorus*. Using these maps derived from three interconnected miscanthus populations, they identified 109 to 288 QTL for 14 biomass production traits which mapped into 86 to 157 meta-QTL. Van der Weijde et al. [37] constructed two parental maps of *M. sinensis* using also the pseudo-testcross strategy and identified 86 QTL related to biomass composition and conversion efficiency traits. Most of these studies identified QTL of *M. sinensis* biparental populations derived from crosses between different heterozygous parents. These populations were evaluated during 2 years in one location. In apple trees, another perennial species, Segura et al. [38] carried out QTL mapping on an F1 progeny and demonstrated that the QTL detected are related to genetic, ontogenetic, and climatic factors: this was possible by a staggered-start design [39–41], which partitioned the year effect into age and growing season effects. The growing season effect itself corresponded to soil and climatic condition effects.

In a previous study, we estimated the heritability values and genetic and phenotypic correlations of an *M. sinensis* population. We highlighted moderate to high heritability values for biomass production and composition traits. Both age and climatic condition effects were considered, based on a staggered-start design carried out in two contrasted locations (Raverdy et al., submitted in BioEnergy Research). We thus expected the phenotypic data of this population to be relevant for undertaking QTL mapping. The objectives of the present study were then (1) to construct an integrated linkage map based on single-nucleotide polymorphism (SNP) markers according to the miscanthus reference genome and (2) to detect QTL for biomass production and composition traits and identify stable QTL, while considering the range of years (i.e., climatic conditions), ages, and locations evaluated. To reach these goals, the same population has therefore been analyzed, based on the staggered-start design that was established in two contrasting locations. The growing season effect corresponded to climatic condition effect as the stands were staggered twice in a single field in both locations. A genotyping-by-sequencing (GBS) approach was initially used to discover SNP markers according to the alignment with the miscanthus reference genome that was released in 2017 [42]. A next step was the development of an integrated linkage map for the population. QTL were then detected for both biomass production and composition traits over 5 consecutive years. To our knowledge, it is the first miscanthus genetic map that considers the alignment with the

miscanthus reference genome. It is also the first miscanthus study for which the QTL of biomass production and composition traits are jointly detected for more than 2 years and in two contrasted locations. This will make it possible to assess marker-trait associations during and after the theoretical establishment phase of miscanthus and according to different ages and climatic conditions. Based on the staggered-start design per location, the "year" effect was partitioned into "plant age" and "climatic condition" effects, which made it possible to detect QTL related to such conditions for the first time in miscanthus.

Materials and Methods

Mapping Population and Experimental Design

Two diploid ornamental *M. sinensis* cultivars, "Malepartus" (Mal) and "Silberspinne" (Sil), were crossed in order to get an F1 full-sib progeny for mapping studies. These two ornamental parents putatively originate from central to southern Japan [15, 16]. Each seed of the population was germinated in vitro, which provided 248 initial genotypes. These plants were propagated in vitro according to a protocol of shoot organogenesis and regeneration [43]. All seedlings were planted in a greenhouse that offered suitable growing conditions before being transplanted to the field. Due to variable genotype ability for in vitro propagation, only 157 genotypes were available for the field trial. However, 248 genotypes were available for genotyping.

The F1 full-sib progeny and the two parental genotypes were cultivated with single plants, in two locations in France. These locations presented contrasting soil and climatic conditions. Experimental design in both locations was based on a staggered-start design [39]. One of the experiments was established at the "GCIE" INRAE (National Institute for Research on Agriculture, Food and Environment) experimental unit of Estrées-Mons (49°53′N, 3°00′E) in a deep loam soil (Orthic Luvisol according to the World Reference Base for Soil Resources, WRB). The other experiment was established at the "GBFOr" INRAE experimental unit of Orléans (47°49′N, 1°54′E) in a sandy soil (Dystric Cambisol, WRB). Each staggered-start design was made up of two stands or groups of genotypes that were organized in two plots established in 2 successive years in the same field: the first group of genotypes (G1) was established in 2014, while the second group (G2) was established in 2015 (Fig. 1). In each location, each plot was adjacent to the other plot in the field and separated by a border row: soil conditions were therefore similar between the two groups. In Estrées-Mons, 157 genotypes and the two parents were cultivated, with 82 genotypes common to each group (G1 and G2) (Fig. 1 and Table S2). The two parents and 104 genotypes of those

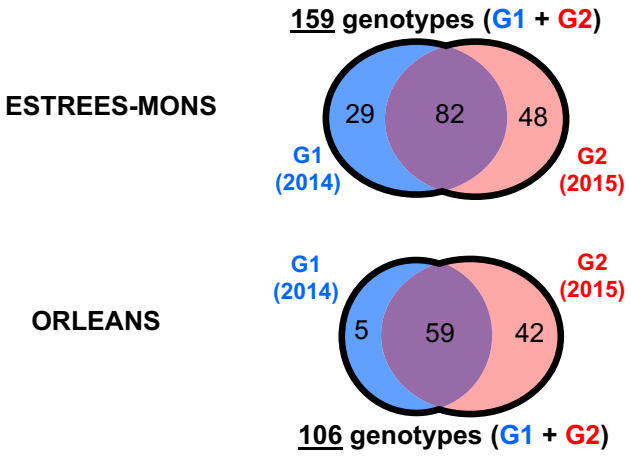

Fig. 1 A staggered-start design was established in Estrées-Mons and Orléans. The total number of *M. sinensis* genotypes (including the two parents) is indicated for each staggered-start design. The number of genotypes for each group (G1 and G2) and the genotypes common to both groups (at the intersection of blue and red circles) are also specified for each location

cultivated in Estrées-Mons were also cultivated in Orléans. These 106 genotypes included 59 common to G1 and G2 (Fig. 1 and Table S2). Finally, 57 genotypes were common both to G1 and G2 and between locations (Table S2). The number of genotypes was unbalanced due to the recalcitrance of some genotypes concerning the propagation and establishment steps, as described previously. In both locations and each plot that corresponded to each group of genotypes sequentially established, single plants were organized in an incomplete randomized block design [44] with five blocks. The genotypes were thus replicated in four of the five blocks on average, except the two parents, which were replicated in all blocks of the two stands. Plant density was 1 plant per square meter, with single plants equally spaced 1 m apart within and between rows. For both locations, more details about the experimental design and the climatic conditions that correspond to the plant cycle are available in Raverdy et al. (submitted in BioEnergy Research).

Development of Single-Nucleotide Polymorphism Markers

The genomic DNA of the 248 individuals of the progeny and of the two parents was extracted from seedlings at the INRAE Gentyane Genomic Service platform (Clermont-Ferrand, Puy-de-Dôme, France), using a sbeadex™ livestock kit (LGC Group, United Kingdom). A GBS approach was used to discover SNP markers which corresponded to the population. It was carried out according to the protocols of Elshire et al. [45] and Cormier et al. [46], at the CIRAD (French Agricultural Research Centre for International Development) Genotyping platform core facility (GPTR,

Montpellier, Hérault, France). The 96-plex libraries were prepared by digestion of the 250 DNA extractions using the *Pst*I restriction enzyme. The single-end sequencing was then carried out in three lanes on a HiSeq™ 3000 sequencer (Illumina® Inc., San Diego, CA, USA), at the "GenoToul GeT" platform (Auzeville, Haute-Garonne, France). The quality check of the reads was conducted using the "FastQC" software v.0.11.2 [47].

The "TASSEL 5 GBS" pipeline [48] was used in order to analyze the sequencing data. The raw reads of the individuals were initially grouped in tags. By using the "Bowtie 2" v.2.3.2 software [49], tags with a minimum count of 10 reads were retained and aligned with the *M. sinensis* reference genome sequence that was released in December 2017 [42]. The resulting variant call format file, which contains the markers and information regarding the individuals, was then filtered. The filtering was carried out using the "vcf2pop.1.0.py" software [50] and was based on a minor allele frequency threshold of 0.05 and a maximum of 25% missing data per marker. Three marker types were available for the genetic map construction: first, markers that were heterozygous in both parents (ab × ab), called "Bridge" (Bri) markers and that segregated in a 1:2:1 ratio (aa, ab, bb); second, "Mal" markers that were heterozygous in the Malepartus parent and homozygous in the Silberspinne parent (ab × aa); and lastly, "Sil" markers that were homozygous in Malepartus and heterozygous in Silberspinne (aa × ab). The latter two marker types segregated in a 1:1 ratio (aa, ab). According to the classification of Wu et al. [25], "Bri" markers were named B3.7, "Mal" markers were named D1.10, and "Sil" markers were named D2.15.

Genetic Map Construction

An integrated genetic map was constructed based on the mapping population, by using the "OneMap" R package [51, 52]. First, the redundant markers were removed from the analysis. The remaining markers were tested according to expected Mendelian segregation by using a chi-square test with a global $\alpha = 0.05$, corrected for multiple testing with Bonferroni correction. The recombination fraction between all pairs of markers was then determined according to two-point tests [25]. The markers were then grouped based on their position on the reference genome chromosomes and a maximum recombination frequency of 0.35. Based on this grouping, 19 linkage groups (LGs) were obtained, which corresponded to the base chromosome number of miscanthus ($x = 19$). A high homology was highlighted between the markers grouped in each LG and their original position in accordance with the alignment with the reference genome chromosomes. For example, 90% of the markers grouped in LG1 were initially identified in correspondence with chromosome 1 of the miscanthus reference genome sequence. The marker grouping for each of the 19 LGs was thus refined, by only keeping the markers that were in accordance with the corresponding chromosome of the miscanthus reference genome. This ensured that each marker was grouped in the right LG.

The segregation-distorted markers that were kept for the optimization of the marker grouping step were then discarded for marker ordering and phasing. This met the QTL mapping model assumption of Mendelian segregation. Marker ordering was then tested according to three different methods, and the Kosambi mapping function was used [53]. These different marker ordering methods were based either on heuristic algorithms or physical positions within the miscanthus reference genome. They had to be tested independently in order to find the best marker order among them. The best order was defined with the inspection of the expected pattern in the resulting recombination fraction matrix between markers, visualized in heatmaps. The ordering methods have already been investigated in different studies, with the aim of getting the best marker order: it was yielded either by the marker ordering algorithms or the physical positions within a reference genome [54, 55].

The first marker ordering method consisted of the ordering of the most informative markers (1:2:1) using an exhaustive search tool. It consisted of comparing all possible orders, and the remaining markers were positioned according to the "TRY" algorithm [56]. The resulting marker order was unsatisfactory according to the heatmaps (data not shown), even though the "RIPPLE" algorithm [56] was used in order to improve it. A second marker ordering method was thus tested based on the multi-dimensional scaling (MDS) method [57, 58] that was implemented in the OneMap software [51, 52]: although it improved the marker order, this approach did not provide the means to get a satisfactory marker order (data not shown). However, this approach made it possible to refine marker filtering, by removing some markers according to the principal curves method from the "MDSMap" R package [57, 58]. Finally, the third method consisted in ordering the markers according to their positions identified within the miscanthus reference genome. In addition, the marker order was adjusted based on recombination fractions between the markers. This final method was retained, as it yielded the best marker order quality among the three methods.

Once the ordering was defined, the genetic distance was estimated based on multipoint approaches using hidden Markov models [56] that consider outcrossing species [26]. The presence of genotyping errors was managed, as often carried out in mapping studies [54, 59–61]. Thus, a genotyping error probability of 5% was considered in the hidden Markov model emission function. This function was implemented in the OneMap software, which made it possible

to consider uncertainties between observed and estimated genotypes [62].

Phenotypic Data Analysis of Biomass Production and Composition Traits

Five biomass production traits and five biomass composition traits (expressed as a percentage of dry matter, %DM, or cell wall content, %CW) were studied. These phenotypic data were acquired over 5 successive years between 2014 and 2019 in Estrées-Mons and 4 successive years between 2014 and 2018 in Orléans (Fig. 2). In order to name each trait in a relevant manner, a miscanthus ontology was developed at the INRAE BioEcoAgro research unit of Estrées-Mons (https://urgi.versailles.inra.fr/ephesis/ephesis/ontologyportal.do) by using the GnpIS multispecies integrative information system from the INRAE-URGI of Versailles [63]. Four morphological traits were evaluated: canopy height (CH_cm), plant maximum height (HMax_cm), plant stem number (PSNb), and plant circumference (C50_cm). The aboveground biomass yield (ABM_tDMha) was measured after the winter harvest in late February and was expressed as tDM/ha. The biomass composition–related traits were assessed based on near-infrared spectroscopy (NIRS) predictions for all plants of the population and a set of calibration samples for which the composition traits were assessed. These samples were analyzed by the LANO laboratory (Saint-Lô, France) according to a protocol adapted from the Van Soest method [64] and described in Belmokhtar et al. [65]. Three fractions were determined: neutral detergent fiber (NDF), acid detergent fiber (ADF), and acid detergent lignin (ADL). The NDF fraction, which corresponds to the cell wall content (CW), is considered to represent cellulose, hemicelluloses, and lignin. The ADF consists of cellulose and lignin, and the ADL consists of lignin [64]. The cellulose content (CL) was estimated by subtracting ADL from ADF, hemicelluloses content (HEM) was obtained by subtracting ADF from NDF, and finally, lignin content corresponded to ADL. The dry matter content of each calibration sample was determined at 103 °C to express all of the previous values (NDF, ADF, ADL, CL, and HEM) in percentage of dry matter (% DM). The traits were also expressed as percentage of cell-wall, excepting NDF.

For each location considered separately, the staggered-start design made it possible to analyze the phenotyping data by distinguishing the "plant age" effect from the "climatic condition" effect, according to two linear mixed models [66]. An initial model (1), which takes into account the "age" effect, was applied to three data subsets in each location (2016-year, 2017-year or 2018-year data subsets in Fig. 2). A second model (2), which accounts for the "climatic condition" effect, was used with three other data subsets in Orléans (age 1, age 2, or age 3 data subsets in Fig. 2), and a fourth additional subset in Estrées-Mons (age 4 data subset). The two models were assessed using the restricted maximum likelihood (REML) approach, known to be suitable

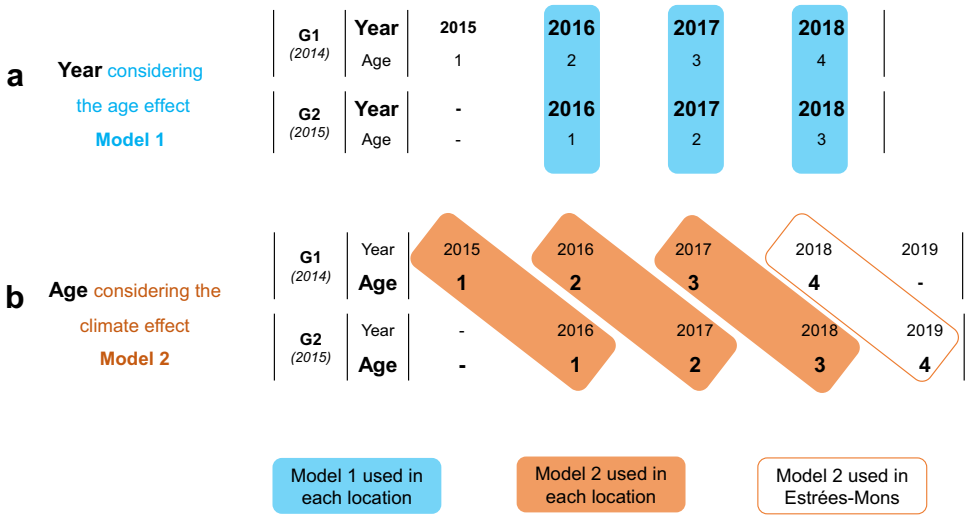

Fig. 2 For each location considered separately, the corresponding staggered-start design was analyzed according to two statistical models: **a**) the age effect model per growing season or year (i.e., climatic condition) for Model 1 and **b**) the climatic condition effect model per age for Model 2. These models were based on two different subsets of the data. For example, case (a) was based on year 2016 considered for G1 and G2, with the age effect modeled according to 2-year-old genotypes in G1 and 1-year-old genotypes in G2. In this case, plants of different ages grew under the same climatic condition during a single year. While case (b) was based, for example, on genotypes with the same age (2-year-old), according to genotypes of G1 which grew in the year 2016 and genotypes of G2 which grew in the year 2017. For this case, plants of the same age grew under two different climatic conditions, related to each year considered for each group. The group year establishment is specified in brackets below each group name

for analyzing unbalanced datasets, and the corresponding function was implemented in the "breedR" R package [67]. Model 1 was specified as follows:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad (1)$$

in which Y_{ijk} represents the phenotypic value measured on plant k of genotype i at age j; μ is the overall mean; α_i is the random effect of genotype i; β_j is the fixed effect of age j; $(\alpha\beta)_{ij}$ is the random interaction between genotype i and age j; and ε_{ijk} is the random residual for plant k of genotype i at age j.

Model 2 was specified as follows:

$$Y_{ikl} = \mu + \alpha'_i + \gamma_l + \left(\alpha'\gamma\right)_{il} + \varepsilon'_{ikl} \quad (2)$$

where each term is similar to those of Model 1, excepting age effect β_j which is replaced by climatic condition effect γ_l of year l, and the interaction between genotype i and age j $(\alpha\beta)_{ij}$ which is replaced by the interaction between genotype i and climatic condition effect of year l $(\alpha'\gamma)_{il}$. The terms α'_i, and ε'_{ikl} were different from the effects given by the previous Model 1 because they were estimated based on different subsets of the data.

In both models, spatial effects were accounted for using an autoregressive correlation structure, based on x and y coordinates in the field, to partition the covariance matrix of each residual ε_{ijk} and ε'_{ikl} into a spatially dependent component and an independent remaining residual variance [68]. In order to minimize the border effect, the trial was surrounded by one row of border plants, which were accounted for in the estimation of the autoregressive model. However, the corresponding genotypes were discarded for subsequent QTL detections, once the best linear unbiased predictions (BLUP) of all genotypes of the trial were calculated.

Block and spatial effects were both initially included in each model. As the models that only considered the spatial effect yielded the best Akaike information criterion (AIC) [69], the model was simplified and the block effect was finally not considered. It can be noted that the low performance of a statistical model that includes a block effect reinforced the similarity in the soil conditions between both stands of the staggered-start design. Based on each model described above, the BLUP of the random genotype (G) effect were estimated in order to carry out the QTL mapping of miscanthus biomass production and composition traits. For each location, the BLUP of the G effect that were estimated using Model 1 were related to the climatic condition of each studied year (i.e., data subset). They were considered as independent from the age effect between G1 and G2 groups, i.e., represented by the fixed effect of the age and the random interaction between genotype and age (Fig. 2). Reciprocally, the BLUP of the G effect estimated using Model 2 were related to each age (i.e., data subset). They were considered as independent from the climatic condition effect between G1 and G2 groups, i.e., represented by the fixed effect of the climatic condition and the random interaction between genotype and climatic conditions (Fig. 2). For biomass production and composition traits, the G effects estimated based on these two models were previously shown to account for the highest part of the variance components analyzed according to the staggered-start design in each location (Raverdy et al., submitted to BioEnergy Research) (Figs. 3 and 4).

For each trait and each condition, the distribution of the BLUP along with the parental values and phenotypic means were observed: as the data were normally distributed, no data transformation was needed (Fig. S5a, S5b, S5c, S5d and S5e).

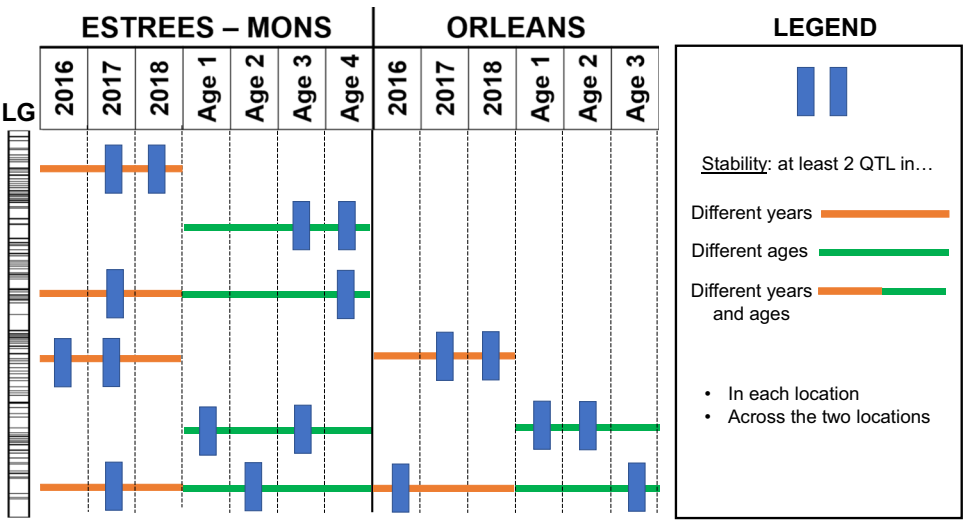

Fig. 3 Illustration of the different types of QTL stability according to the 13 conditions studied

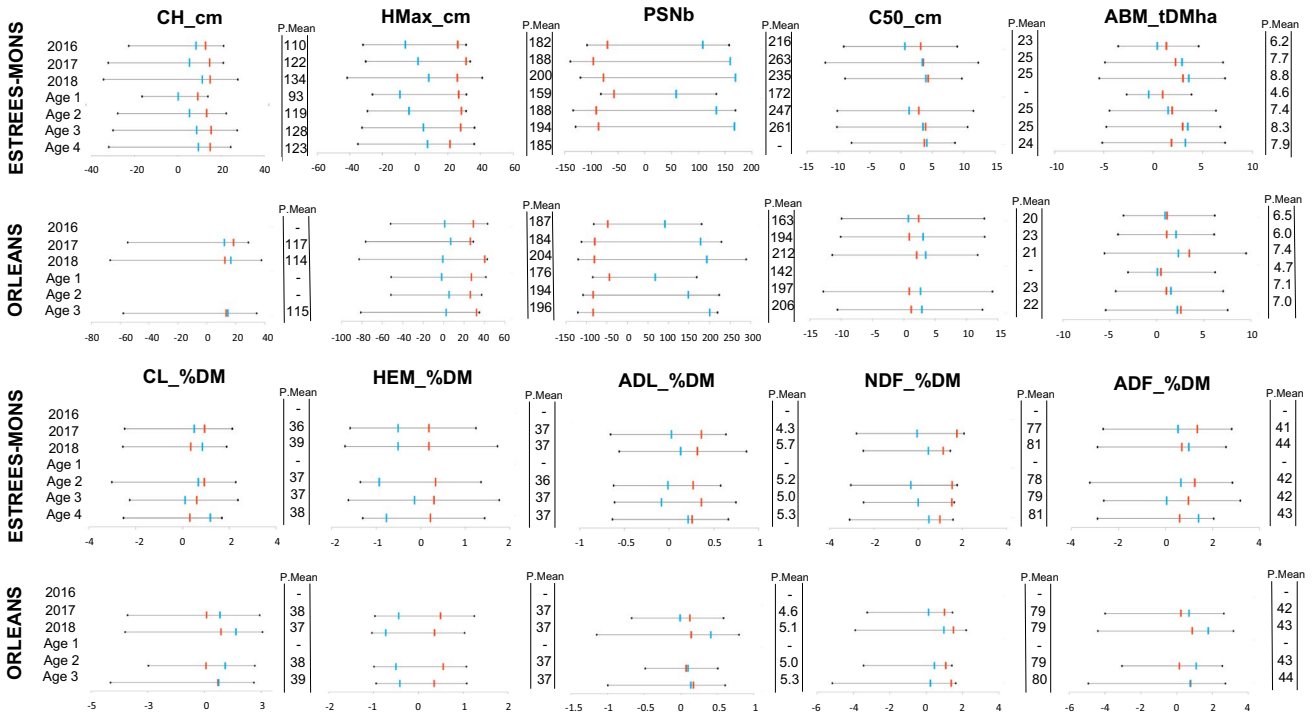

Fig. 4 Summarized distributions of best linear unbiased predictors (BLUP) and the associated phenotypic means (P.Mean), for biomass production and composition traits for each condition. The red and blue lines represent the parental BLUP, which correspond to Malepartus and Silberspinne, respectively

Quantitative Trait Loci Mapping

The BLUP of the five biomass production traits and the five biomass composition traits (expressed as %DM or %CW) previously cited were used for QTL mapping. Each data subset of the staggered-start design per location was considered separately (Fig. 2). As miscanthus is an outcrossing species, a specific CIM model adapted to full-sib progeny was carried out by using the "fullsibQTL" R package [27, 70]. Outcrossing segregation patterns were considered in the model based on a multipoint approach which estimated three genetic effects according to three contrasts: two contrasts concerned the additive effects of the QTL alleles for each parent, and the third one was related to the intra-locus interaction (dominance) between additive effects on each parent. The conditional multipoint probabilities of QTL genotypes were obtained for every 1-cM interval. For each linkage group, the cofactors were selected using multiple linear regression with a stepwise procedure. The associated model selection was based on the AIC [69]. The QTL were defined for a threshold based on a significance level of 5% across distributed LOD scores that were obtained by selecting the second LOD profile peak from 1000 permutations [71]. However, although some QTL were defined according to the same method, the corresponding threshold was based on a significance level of 10% in order to detect a supplemental set of stable QTL between conditions (Fig. S1). The QTL impacted by this methodology were displayed with a "*" in the results section (Fig. 5). The three genetic effects previously detailed, the linkage phases, segregation patterns, and the proportion of genotypic variation explained by each QTL (R^2) were estimated based on the CIM model. The QTL confidence intervals were calculated using the 2-LOD drop-off method [72].

As QTL mapping was carried out for each climatic condition and each age in each location separately, it was possible to highlight QTL that were stable according to these conditions. The QTL were defined as stable for a given trait when they co-localized under at least two conditions (Fig. 3). QTL could be stable across different climatic conditions in each location, for example, a QTL detected in 2017 and 2018 in Estrées-Mons. QTL could also be stable across different ages in each location, for example, if a QTL was detected at age 3 and age 4 in Estrées-Mons. Moreover, QTL could also be stable across climatic conditions and ages in a given location: for example, when a QTL was detected in 2017 and at age 4 in Estrées-Mons. Finally, stable QTL were also identified when they co-localized under at least four conditions across the two locations: for example, (1) two QTL detected under two climatic conditions in each location, (2) two QTL detected for two ages in each location, and, lastly,

(3) two QTL detected for one climatic condition and one age in each location (Fig. 3).

For each trait, the proportion of stable QTL across climatic conditions, ages, and/or locations was then determined: in Estrées-Mons, for example, the number of stable QTL that was detected under at least two climatic conditions was divided by the total number of QTL detected across the different climatic conditions in Estrées-Mons, in order to calculate the corresponding percentage.

QTL clusters were identified for a given climatic condition or a given age in each location. These clusters corresponded to the co-localization of at least three QTL for different traits: the biomass composition traits that were either expressed as %DM or %CW were not considered as different traits. QTL clusters were made up of biomass production traits, biomass composition traits, and both biomass production and composition traits. The reliability of these QTL clusters was verified according to the Pearson correlation coefficients based on BLUP. They were computed by using the "stats" R package and visualized using the "corrplot" R package [73].

Identification of Putative Cell Wall–Related Genes in *M. sinensis*

Putative cell wall–related genes were identified in *M. sinensis* based on the orthologous relationships between *M. sinensis* and sorghum and between *M. sinensis* and maize. These two species are indeed relatives of miscanthus and can be considered as genetic models for miscanthus. This offers the opportunity to take into account the genetic knowledge that is currently available for maize and sorghum.

Orthologous *M. sinensis* genes were initially identified based on two cell wall–related gene lists, hereafter referred to as search lists, that were composed of 2148 candidate genes for sorghum and 2470 for maize (Virlouvet, personal communication). Secondly, these *M. sinensis* genes that are located between the two markers flanking each QTL cluster, i.e., flanking the region where the most extreme QTL confidence intervals within a cluster overlapped, were selected. This selection was based on the annotation file Msinensis_497_v7.1.gene.gff3 from Phytozome, the Plant Comparative Genomics portal of the Department of Energy's Joint Genome Institute (https://phytozome.jgi.doe.gov), which contains 67,789 genes. Among these positional *M. sinensis* genes, the genes that were orthologous to cell wall–related genes in sorghum and maize could finally be identified by comparison with the initial search lists.

Results

Single-Nucleotide Polymorphism Calling and Filtering

The sequencing of GBS libraries produced 1,161,537,843 raw reads based on a read length of 150 bp. Nineteen individuals out of the 248 were removed from the analysis, due to the low quality of the corresponding sequencing data: a total of 229 individuals was thus considered for the genetic map construction. The SNP calling and the alignment with the *M. sinensis* reference genome made it possible to identify a final set of 149,741 biallelic SNP markers. The filtering of these markers then provided a selection of 9330 high-quality SNP markers available for genetic mapping. The details of the markers according to their segregation type and coding [25] are available in Table 1.

Genetic Map Construction

The grouping of the SNP markers resulted in 3774 SNP markers distributed across 19 LGs. Based on the initial marker dataset, redundant markers were removed, and markers unlinked to any of the 19 LGs were thus not considered for the analysis. It can be noted that LG12 was partitioned into two LGs, because it was not possible to group all three marker types together due to the presence of only four B3.7 available markers (i.e., heterozygous in both parents): for that reason, LG12a was made up of "Bri" and "Mal" markers on the one hand, while LG12b was made up of "Bri" and "Sil" markers on the other. Finally, all LGs corresponded to the 19 chromosomes of the miscanthus reference genome, when checking each marker provenance according to their physical position in relation to the original alignment step. Segregation-distorted markers were initially kept in order to optimize the grouping phase, but they were then removed before the ordering step:

Table 1 Characteristics of SNP markers after the filtering steps and which are available for the construction of the integrated genetic map

	Bridge (Bri)	Malepartus (Mal)	Silberspinne (Sil)	Total
SNPs number	2,488	3,833	3,009	9,330
Marker type	B3.7	D1.10	D2.15	-
Parental cross type	ab × ab	ab × aa	aa × ab	-
Observed bands in the progeny	a, 2ab, b (ab × ab)	a, ab (ab × a)	a, ab (a × ab)	-
SNPs segregation type	1:2:1	1:1	1:1	-

3262 non-distorted SNP markers thus constituted the 19 LGs. Moreover, a supplemental analysis using the principal curves method led to the removal of an extra set of 290 markers. After the ordering of the SNP markers according to the physical positions within the miscanthus reference genome, a few badly ordered markers were moved or removed according to the recombination fractions between them. The final integrated map was made up of 2602 markers and had a total length of 2770 cM (Table 2 and Fig. S2). The different markers on the map were as follows: 613 "Bri" markers of B3.7 type, 1256 "Mal" markers of D1.10 type, and 733 "Sil" markers of D2.15 type. In the 19 LGs, LG1 was the largest with a length of 217.3 cM and LG12a and LG12b were the shortest with a length of 49.3 and 32.6 cM, respectively. The average inter-marker distance was 1.06 cM when considering the 19 LGs. LG4 showed the highest density with a mean interval of 0.72 cM between markers, which was in accordance with the highest number of 271 markers mapped on this LG (Table 2). For each LG, the heatmap showed the good quality of the marker order (Fig. S3).

Transgressive Segregation Was Observed for Biomass Production and Composition Traits

The distribution of the BLUP and phenotypic means of the progeny were observed for biomass production and composition traits for each condition (Fig. 4; Fig. S4; Fig. S5a, S5b, S5c, S5d and S5e). These conditions were related to the climatic condition that occurred across the year, the age of the plants, and the location in which they were established. This resulted in a total of 13 conditions. Parental BLUP were reported for each trait and each condition. For both biomass production traits and composition traits, transgressive segregation was observed for each condition, except for the plant stem number (PSNb). In fact, the two parents of the population were chosen and crossed based on their highly contrasted stem number, as shown in Fig. 4. High genetic variability was observed for each of the biomass production traits: this was illustrated at age 3 by the BLUP range expressed as a percentage of its corresponding phenotypic mean. These percentages ranged between 35% for the plant maximum height (HMax_cm) in Estrées-Mons and 186% for the aboveground biomass yield (ABM_tDMha) in Orléans. This variability tended to increase over the years (i.e., climatic conditions) and ages for most of the biomass production traits.

Regarding biomass composition traits, their genetic variability was lower than the genetic variability observed for biomass production traits: the BLUP range, expressed as a percentage of its corresponding phenotypic mean, varied between 5% for NDF_%DM in Estrées-Mons and 30% for ADL_%DM in Orléans. In contrast to the biomass

Table 2 Descriptive statistics related to the integrated genetic map. The marker type is coded according to the notation of Wu et al. [25]

Linkage group (LG)	Number of markers	Length (cM)	Average inter- marker interval (cM)	Marker segregation type	Marker type
1	105	217.3	2.07	1:2:1; 1:1	B3.7; D1.10; D2.15
2	256	191.6	0.75	1:2:1; 1:1	B3.7; D1.10; D2.15
3	161	152.6	0.95	1:2:1; 1:1	B3.7; D1.10; D2.15
4	271	194.4	0.72	1:2:1; 1:1	B3.7; D1.10; D2.15
5	205	202.7	0.99	1:2:1; 1:1	B3.7; D1.10; D2.15
6	201	193.1	0.96	1:2:1; 1:1	B3.7; D1.10; D2.15
7	240	212.3	0.88	1:2:1; 1:1	B3.7; D1.10; D2.15
8	152	148.0	0.97	1:2:1; 1:1	B3.7; D1.10; D2.15
9	69	109.4	1.59	1:2:1; 1:1	B3.7; D1.10; D2.15
10	94	130.2	1.39	1:2:1; 1:1	B3.7; D1.10; D2.15
11	103	128.1	1.24	1:2:1; 1:1	B3.7; D1.10; D2.15
12a	60	49.3	0.82	1:2:1; 1:1	B3.7; D1.10
12b	5	32.6	6.52	1:2:1; 1:1	B3.7; D2.15
13	110	102.3	0.93	1:2:1; 1:1	B3.7; D1.10; D2.15
14	88	108.3	1.23	1:2:1; 1:1	B3.7; D1.10; D2.15
15	38	120.1	3.16	1:2:1; 1:1	B3.7; D1.10; D2.15
16	112	105.0	0.94	1:2:1; 1:1	B3.7; D1.10; D2.15
17	108	124.2	1.15	1:2:1; 1:1	B3.7; D1.10; D2.15
18	67	79.9	1.19	1:2:1; 1:1	B3.7; D1.10; D2.15
19	157	168.5	1.07	1:2:1; 1:1	B3.7; D1.10; D2.15
Total	**2602**	**2770**	**1.06**		

production traits, this genetic variability did not increase across the years or ages.

Stable QTL Were Mainly Detected Across Climatic Conditions and Ages for Biomass Production Traits, While Few Stable QTL Were Detected for Biomass Composition Traits

A total of 260 QTL was detected for biomass production traits (Tables 3 and S1) and 283 QTL for biomass composition traits (Tables 4 and S1). The number of QTL was reported for each trait under each condition (Tables 3 and 4). For each location, stable QTL (i.e., QTL that co-localized under at least two conditions) were highlighted across the years, ages, and for both years and ages. An example is given with the solid red triangle in LG8 (37 cM) in Fig. 5: in Estrées-Mons, three stable QTL for aboveground biomass yield (ABM_tDMha) co-localized in 2016, 2017, and 2018, and three QTL also co-localized for age 2, age 3, and age 4. Six QTL thus co-localized for this trait and were either related to years and/or ages.

For biomass production traits considered in each location, the average proportion of stable QTL detected according to different climatic conditions or different ages was similar and around 30% (Table 3). These proportions were mainly consistent between traits and conditions. The stable QTL were not necessarily the QTL that displayed the highest percentage of genotypic variance accounted for (R^2), or the most significant QTL. In each location, the proportion of stable QTL detected by considering years and ages together was higher than stable QTL that only considered years or ages separately. This proportion was 59% in Estrées-Mons and 39% in Orléans. Thus, QTL detected in a given year (e.g., for the climatic condition in 2017 when G1 plants were 3 years old and G2 plants were 2 years old), were also detected for a given age (i.e., age 3 for plants grown in 2017 for G1 and 2018 for G2). Regarding the stability of QTL across the two locations, no stable QTL were identified for biomass production traits (Table 3).

For the biomass composition traits evaluated in each location, the proportion of stable QTL was globally higher for the QTL related to the different climatic conditions, than for

Table 3 Number of QTL detected for each of the biomass production traits for each condition. The number of stable QTL across conditions was written in brackets and the related proportions in italic. Here, the "year" refers to the climatic condition that occurred in each year

			CH_cm	HMax_cm	PSNb	C50_cm	ABM_tDMha	Sum
ESTREES-MONS	**Year**	2016	2	6	6	7	6	27
		2017	6	6	6	3	4	25
		2018	5	4	5	5	7	26
		Total (Stable)	**13 (6)**	**16 (0)**	**17 (4)**	**15 (5)**	**17 (7)**	**78 (22)**
		Stable (%)	***46%***	***0%***	***24%***	***33%***	***41%***	***28%***
	Age	Age 1	6	7	7	NA	5	25
		Age 2	7	4	4	4	6	25
		Age 3	4	5	3	6	4	22
		Age 4	4	4	NA	5	5	18
		Total (Stable)	**21 (4)**	**20 (5)**	**14 (5)**	**15 (6)**	**20 (7)**	**90 (27)**
		Stable (%)	***19%***	***25%***	***36%***	***40%***	***35%***	***30%***
	Year and Age	**Total (Stable)**	**34 (20)**	**36 (21)**	**31 (17)**	**30 (21)**	**37 (20)**	**168 (99)**
		Stable (%)	***59%***	***58%***	***55%***	***70%***	***54%***	***59%***
ORLEANS	**Year**	2016	NA	3	5	4	4	16
		2017	3	3	4	3	3	16
		2018	5	4	5	3	2	19
		Total (Stable)	**8 (0)**	**10 (4)**	**14 (5)**	**10 (4)**	**9 (3)**	**51 (16)**
		Stable (%)	***0%***	***40%***	***36%***	***40%***	***33%***	***31%***
	Age	Age 1	NA	5	6	NA	3	14
		Age 2	NA	3	3	3	2	11
		Age 3	3	3	3	5	2	16
		Total (Stable)	**3 (0)**	**11 (4)**	**12 (6)**	**8 (0)**	**7 (2)**	**41 (12)**
		Stable (%)	***0%***	***36%***	***50%***	***0%***	***29%***	***29%***
	Year and Age	**Total (Stable)**	**11 (2)**	**21 (8)**	**26 (8)**	**18 (9)**	**16 (9)**	**92 (36)**
		Stable (%)	***18%***	***38%***	***31%***	***50%***	***56%***	***39%***
ESTREES-MONS and ORLEANS	**Year**	**Total (Stable)**	**21 (0)**	**26 (0)**	**31 (0)**	**25 (0)**	**26 (0)**	**129 (0)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***
	Age	**Total (Stable)**	**24 (0)**	**31 (0)**	**26 (0)**	**23 (0)**	**27 (0)**	**131 (0)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***
	Year and Age	**Total (Stable)**	**45 (0)**	**57 (0)**	**57 (0)**	**48 (0)**	**53 (0)**	**260 (0)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***

Table 4 Number of QTL detected for each of the biomass composition traits under each condition. The number of stable QTL across conditions was written in brackets and the related proportions in italic. Here, the "year" refers to the climatic condition that occurred in each year

			NDF_%DM	ADF_%DM	CL_%DM	HEM_%DM	ADL_%DM	CL_%CW	HEM_%CW	ADL_%CW	Sum
ESTREES-MONS	Year	2016	-	-	-	-	-	-	-	-	
		2017	3	4	3	3	4	1	4	5	27
		2018	4	5	5	3	2	2	4	3	28
		Total (Stable)	**7 (2)**	**9 (4)**	**8 (0)**	**6 (0)**	**6 (2)**	**3 (0)**	**8 (0)**	**8 (0)**	**55 (8)**
		Stable (%)	***29%***	***44%***	***0%***	***0%***	***33%***	***0%***	***0%***	***0%***	***15%***
	Age	Age 1	-	-	-	-	-	-	-	-	-
		Age 2	3	5	4	3	5	3	4	3	30
		Age 3	5	4	5	4	2	5	5	4	34
		Age 4	5	3	1	7	3	7	3	8	37
		Total (Stable)	**13 (0)**	**12 (0)**	**10 (4)**	**14 (2)**	**10 (0)**	**15 (0)**	**12 (2)**	**15 (0)**	**101 (8)**
		Stable (%)	***0%***	***0%***	***40%***	***14%***	***0%***	***0%***	***17%***	***0%***	***8%***
	Year and Age	**Total (Stable)**	**20 (3)**	**21 (5)**	**18 (6)**	**20 (7)**	**16 (0)**	**18 (2)**	**20 (13)**	**23 (4)**	**156 (40)**
		Stable (%)	***15%***	***24%***	***33%***	***35%***	***0%***	***11%***	***65%***	***17%***	***26%***
ORLEANS	Year	2016	-	-	-	-	-	-	-	-	-
		2017	5	5	5	1	4	5	2	3	30
		2018	2	5	2	3	4	4	4	6	30
		Total (Stable)	**7 (0)**	**10 (0)**	**7 (2)**	**4 (0)**	**8 (2)**	**9 (2)**	**6 (0)**	**9 (0)**	**60 (6)**
		Stable (%)	***0%***	***0%***	***29%***	***0%***	***25%***	***22%***	***0%***	***0%***	***10%***
	Age	Age 1	-	-	-	-	-	-	-	-	-
		Age 2	5	3	3	5	4	3	4	5	32
		Age 3	2	2	3	6	9	6	5	2	35
		Total (Stable)	**7 (0)**	**5 (0)**	**6 (0)**	**11 (0)**	**13 (0)**	**9 (0)**	**9 (2)**	**7 (0)**	**67 (2)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***22%***	***0%***	***3%***
	Year and Age	**Total (Stable)**	**14 (4)**	**15 (4)**	**13 (7)**	**15 (2)**	**21 (6)**	**18 (5)**	**15 (7)**	**16 (2)**	**127 (37)**
		Stable (%)	***29%***	***27%***	***54%***	***13%***	***29%***	***28%***	***47%***	***13%***	***29%***
ESTREES-MONS and ORLEANS	Year	**Total (Stable)**	**14 (0)**	**19 (0)**	**15 (0)**	**10 (0)**	**14 (0)**	**12 (0)**	**14 (0)**	**17 (0)**	**115 (0)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***
	Age	**Total (Stable)**	**20 (0)**	**17 (0)**	**16 (0)**	**25 (0)**	**23 (0)**	**24 (0)**	**21 (0)**	**22 (0)**	**168 (0)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***
	Year and Age	**Total (Stable)**	**34 (0)**	**36 (0)**	**31 (0)**	**35 (0)**	**37 (0)**	**36 (0)**	**35 (0)**	**39 (0)**	**283 (0)**
		Stable (%)	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***	***0%***

the QTL related to the different ages (Table 4). However, these proportions were relatively low and never exceeded 15% on average, with some traits showing no stable QTL at all. As for biomass production traits, the consideration of successive climatic conditions and ages together made it possible to increase the proportion of stable QTL for biomass composition traits to up to around 30% in each location. No stable QTL were identified across the two locations, similarly to biomass production traits (Tables 4 and 5).

QTL Clusters Were Identified for Biomass Production Traits and Biomass Composition Traits

Many QTL for one trait co-localized with QTL for another trait under each of the 13 conditions. Therefore, QTL clusters were defined when at least three QTL for different traits co-localized. This led to the identification of 12 QTL clusters, which were either specific to biomass production traits, biomass composition traits, or both (Fig. 5 and Table 6).

Five QTL clusters were identified for biomass production traits and were located in LG4, LG7, and LG18 (Fig. 5 and Table 6). Four of these clusters were made up of QTL for canopy height (CH_cm), plant circumference (C50_cm), and aboveground biomass yield (ABM_tDMha). One of the clusters, located in LG7, was made up of QTL for canopy height, plant maximum height (HMax_cm), and aboveground biomass yield. All these clusters for biomass production traits were identified in Estrées-Mons, in 2018, and at ages 1, 2, and 4. In LG18, a particularly stable cluster, based on the same component traits, was detected in 2018, at age 2 and age 4. The reliability of the clusters was confirmed based on the large positive correlations between the BLUP of the traits that belonged to the clusters identified for each condition. These correlations were all statistically significant at the 0.05 probability level and ranged from 0.65 to 0.96 (Fig. 6).

Regarding biomass composition traits, six QTL clusters were identified (Fig. 5 and Table 6): three of them were located in LG4, LG5, and LG15 and were made up of QTL for cellulose, hemicellulose, and lignin contents (expressed as %DM or %CW). Two other clusters were made up of QTL for ADF_%DM, cellulose (as %DM or %CW), and hemicelluloses (as %CW) and were located in LG13 and LG15. The last cluster was located in LG16 and was made up of QTL related to ADF_%DM, CL_%DM, and ADL_%DM. As for biomass production traits, the QTL clusters were confirmed by the high, positive, or negative correlations between the biomass composition traits, statistically significant at the 0.05 probability level (Fig. 6): they ranged from − 0.55 to − 0.96 for negative values and from 0.48 to 0.98 for positive values (except for the correlations between cellulose and lignin, which ranged from 0.37 to 0.47). Stable QTL clusters were detected in LG15, for QTL detected in 2017 and at age 2 in Orléans. The other QTL clusters were detected in 2017, at age 2 and age 3 in Estrées-Mons, and in 2018 in Orléans.

Interestingly, a QTL cluster made up of both biomass production and composition traits was detected in LG15,

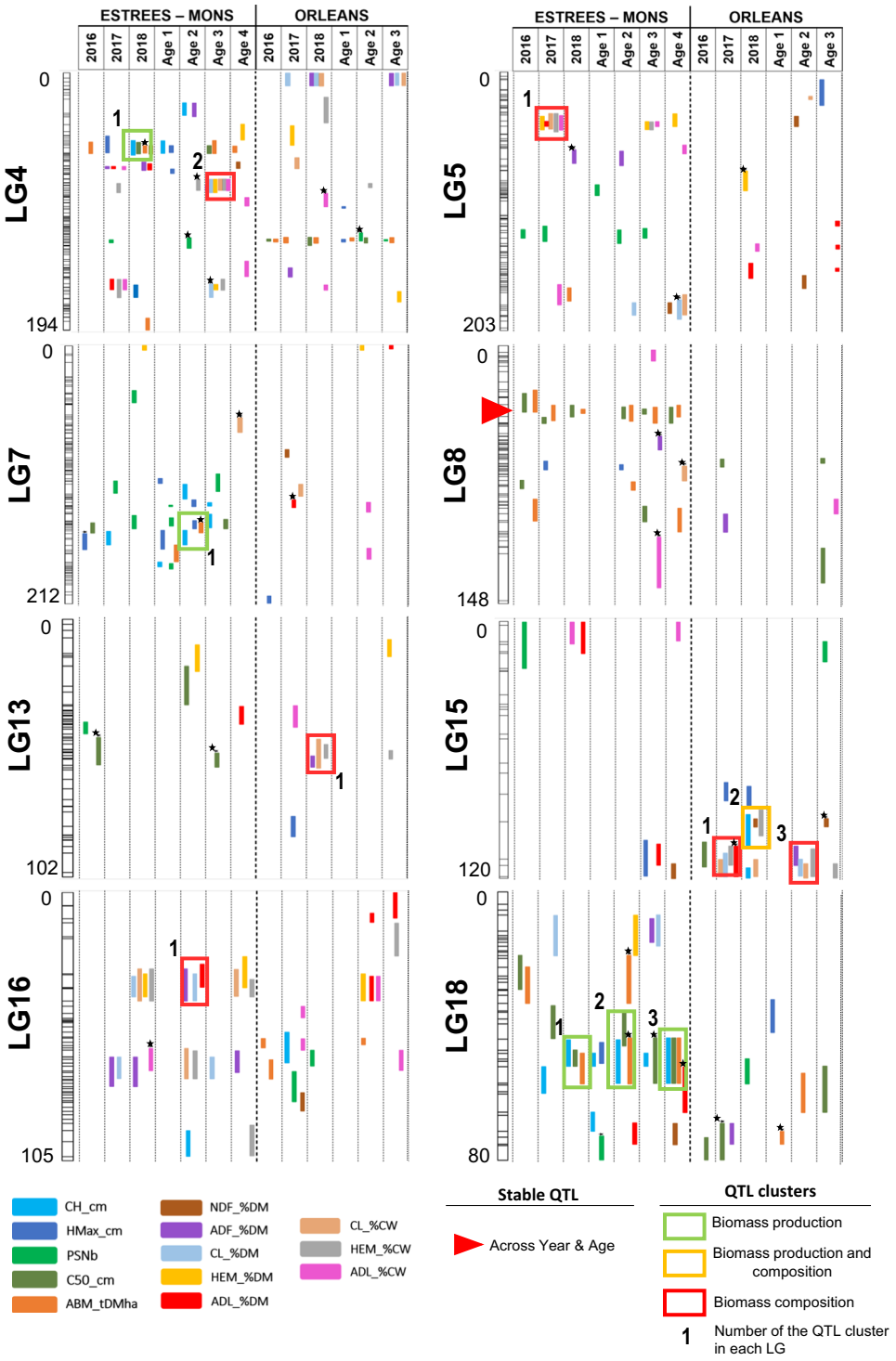

Fig. 5 Representation of eight LGs out of 19. The QTL were detected according to 13 conditions. A part of the stable QTL is illustrated, as well as QTL clusters. The QTL detected for a threshold based on a 10% significance level are marked with a "*" (see the "Materials and Methods" section for more details). The length of each LG is specified to the left of the LG in cM

for the year 2018 (i.e., climatic condition) in Orléans (Fig. 5). These QTL were related to canopy height (CH_cm), NDF_%DM and hemicelluloses (HEM_%CW). This was consistent with the statistically significant correlations between the corresponding BLUP: 0.44 between CH_cm and NDF_%DM and − 0.55 between CH_cm and HEM_%CW (Fig. 6).

QTL Effects of Biomass Production and Composition Traits Were Found to Be Stronger in Orléans Than in Estrées-Mons

According to the QTL identified for each of the biomass production and composition traits, the minimum and maximum for additive effects of each parent and dominance effects

Table 5 Summary statistics of the QTL identified for each trait in both locations

Trait	Location	Additive effect for Mal		Additive effect for Sil		Dominance effect		Proportion of significant effects from all QTL of the trait			R^2 range (%)	LOD range
		Min	Max	Min	Max	Min	Max	Mal	Sil	Dominance		
CH_cm	Estrées-Mons	−8.23	6.28	−5.04	6.06	−4.06	6.72	0.35	0.29	0.36	1.4–15.9	4.8–11.2
	Orléans	−6.31	4.05	−4.98	6.53	−8.10	9.16	0.32	0.32	0.36	4.5–24.5	5.4–11.0
HMax_cm	Estrées-Mons	−9.21	3.94	−6.59	8.73	−5.40	5.28	0.26	0.42	0.32	3.5–16.9	4.7–11.8
	Orléans	−11.78	5.95	−13.39	9.73	−14.11	8.23	0.28	0.38	0.34	0.1–19.6	5.2–11.4
PSNb	Estrées-Mons	−18.53	33.95	−29.46	12.09	−24.58	21.79	0.34	0.29	0.37	3.6–17.4	4.6–11.1
	Orléans	−37.33	21.53	−20.86	34.68	−35.97	39.52	0.27	0.33	0.40	0.5–19.9	5.6–13.8
C50_cm	Estrées-Mons	−1.54	1.15	−1.84	1.69	−1.94	1.24	0.23	0.43	0.34	2.2–16.8	4.6–11.4
	Orléans	−0.94	1.83	−1.69	2.32	−3.07	2.63	0.14	0.40	0.46	2.8–20.1	5.4–9.8
ABM_tDMha	Estrées-Mons	−0.97	0.87	−0.98	0.97	−1.14	0.94	0.25	0.39	0.36	3.9–19.5	4.7–11.2
	Orléans	−0.79	1.19	−0.64	1.25	−0.82	1.20	0.26	0.39	0.35	6.4–20.1	5.9–13.5
NDF_%DM	Estrées-Mons	−0.32	0.20	−0.28	0.29	−0.32	0.33	0.36	0.31	0.33	3.2–16.1	4.8–12.6
	Orléans	−0.19	0.72	−0.94	0.54	−0.47	0.52	0.32	0.29	0.38	2.3–21.0	5.7–12.4
ADF_%DM	Estrées-Mons	−0.58	0.87	−0.60	0.61	−0.30	0.43	0.38	0.31	0.31	2.8–19.3	4.6–11.9
	Orléans	−0.95	0.57	−0.46	0.41	−0.56	0.50	0.35	0.32	0.32	8.9–26.1	5.7–13.6
CL_%DM	Estrées-Mons	−0.45	0.27	−0.31	0.23	−0.31	0.53	0.34	0.29	0.37	3.2–15.9	4.8–9.4
	Orléans	−0.73	0.19	−0.32	0.35	−0.56	0.40	0.31	0.28	0.41	3.7–24.0	5.6–14.1
HEM_%DM	Estrées-Mons	−0.30	0.33	−0.36	0.22	−0.17	0.34	0.36	0.36	0.27	3.4–15.8	4.6–10.4
	Orléans	−0.16	0.19	−0.16	0.23	−0.24	0.18	0.28	0.31	0.41	0.9–17.2	5.3–10.9
ADL_%DM	Estrées-Mons	−0.11	0.12	−0.10	0.13	−0.09	0.07	0.34	0.37	0.29	4.2–14.1	4.8–9.6
	Orléans	−0.44	0.29	−0.25	0.24	−0.15	0.22	0.33	0.31	0.35	0.7–21.8	5.1–15.6
CL_%CW	Estrées-Mons	−0.31	0.24	−0.29	0.32	−0.26	0.17	0.32	0.37	0.32	1.0–12.0	4.7–7.8
	Orléans	−0.41	0.50	−0.46	0.47	−0.36	0.41	0.32	0.37	0.32	1.7–29.4	5.2–14.0
HEM_%CW	Estrées-Mons	−0.29	0.45	−0.43	0.28	−0.40	0.36	0.35	0.33	0.33	5.4–12.4	4.6–9.8
	Orléans	−0.21	0.44	−0.43	0.39	−0.52	0.34	0.32	0.38	0.30	5.9–25.2	5.8–12.4
ADL_%CW	Estrées-Mons	−0.11	0.09	−0.07	0.16	−0.11	0.11	0.31	0.31	0.38	4.0–14.8	4.7–8.6
	Orléans	−0.19	0.14	−0.13	0.23	−0.17	0.17	0.38	0.32	0.29	7.2–21.9	5.4–13.0

Table 6 List of the QTL clusters that were identified in different LGs and conditions. See the "Materials and Methods" section for trait names. EM = Estrées-Mons and ORL = Orléans

Cluster	Flanking marker interval positions (cM)	Trait type of the cluster	Traits in the cluster	R^2 range (%)	Condition
LG4_cluster1	51.0–62.2	Production	CH, C50, ABM	7.1– 16.4	EM_2018
LG4_cluster2	80.3–90.9	Composition	CL, HEM, ADL	11.3–15.9	EM_Age3
LG5_cluster1	32.7–48.7	Composition	CL, HEM, ADL	9.3–14.8	EM_2017
LG7_cluster1	144.7–164.3	Production	CH, HMax, ABM	6.1–15.9	EM_Age2
LG13_cluster1	47.2–59.0	Composition	ADF, CL, HEM	5.9–11.9	ORL_2018
LG15_cluster1	93.7–119.3	Composition	CL, HEM, ADL	6.7–29.4	ORL_2017
LG15_cluster2	80.7–111.0	Production/Composition	CH, NDF, HEM	6.2–10.0	ORL_2018
LG15_cluster3	93.7–119.3	Composition	CL, HEM, ADF	16.0–24.4	ORL_Age2
LG16_cluster1	18.0–42.6	Composition	ADF, CL, ADL	6.8–11.4	EM_Age2
LG18_cluster1	43.7–57.0	Production	CH, C50, ABM	9.1–13.3	EM_2018
LG18_cluster2	35.8–57.0	Production	CH, C50, ABM	7.0–12.4	EM_Age2
LG18_cluster3	43.7–57.0	Production	CH_ C50, ABM	7.5–10.1	EM_Age4

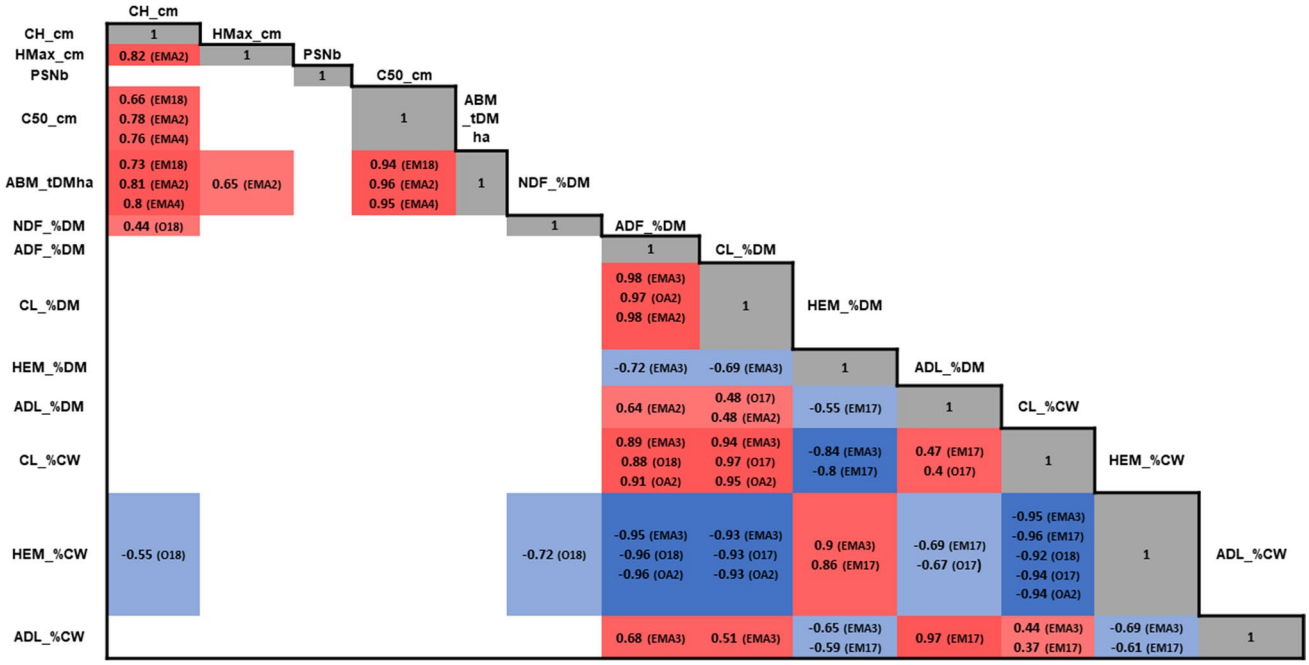

Fig. 6 Significant correlations between BLUP that corresponded to traits within the QTL clusters, among all significant correlations. The significance level was assessed according to a p value < 0.05. The blanks represent correlations that were not shown because the corresponding traits were not identified in the clusters. The conditions under which the QTL clusters were identified are described as EM17, Estrées-Mons in 2017; EM18, Estrées-Mons in 2018; EMA2, Estrées-Mons at age 2; EMA3, Estrées-Mons at age 3; EMA4, Estrées-Mons at age 4; O17, Orléans in 2017; O18, Orléans in 2018; and OA2, Orléans at age 2. See the "Materials and Methods" section for details of trait names

were reported for each location in Table 5, as well as the ranges of R^2 and LOD threshold values. The proportion of significant additive and dominance effects found based on all QTL corresponding to each trait was also reported.

The QTL were detected according to a LOD threshold that ranged from 5.0 to 11.5 on average. For biomass production traits, the highest percentage of genotypic variance explained by the QTL (R^2) of each trait ranged from 15.9 to 19.5% in Estrées-Mons and from 19.6 to 24.5% in Orléans. The QTL detected in Orléans thus tended to explain more genotypic variance than the QTL detected in Estrées-Mons: this was consistent with the values of additive and dominance effects, which were often higher in Orléans than in Estrées-Mons. For example, a maximal dominance effect of 8.2 cm was highlighted for total plant height in Orléans, compared to a BLUP of 5.3 cm in Estrées-Mons. Interestingly, most of the significant QTL effects identified for each production trait in both locations were either due to dominance effects or additive effects of the Silberspinne parent. This was notably observed for plant circumference (C50_cm), for which 43% of the significant effects originated from the Silberspinne allelic effect in Estrées-Mons and 46% originated from the dominance effect in Orléans.

Concerning biomass composition traits, the highest percentage of genotypic variation explained by the QTL of each trait ranged from 12 to 19.3% in Estrées-Mons and from 17.2 to 29.4% in Orléans. As for biomass production traits, the QTL identified in Orléans explained a higher percentage of the genotypic variation than the QTL identified in Estrées-Mons: this was also consistent with the higher allelic and dominance effects often observed in Orléans than in Estrées-Mons. As an example, the maximum dominance effect for cellulose (%CW) was 0.41% in Orléans compared to 0.17% in Estrées-Mons. In contrast to the proportion of significant effects observed for biomass production traits, the maximum proportion of significant effects observed for biomass composition traits were either due to the Malepartus allelic effect, Silberspinne allelic effect, or dominance effect, and that was the case in both locations. This can be shown for hemicellulose content (%DM) in Orléans, with 41% of the significant effects originating from the dominance effect. Concerning cellulose content (%CW) in Estrées-Mons, 37% of the significant effects originated from the Silberspinne allelic effect. Regarding lignin content (%CW), 38% of the significant effects originated from the Malepartus allelic effect.

Some *M. sinensis* Genes Within the QTL Clusters Were Orthologous to Sorghum and Maize Cell Wall–Related Genes

A total of 809 and 494 M. *sinensis* genes within the QTL clusters were determined to be orthologous to genes in sorghum and maize, respectively. The QTL clusters contained 62 of these *M. sinensis* genes that were orthologous to cell wall–related genes in sorghum and maize (Table 7), knowing that the QTL clusters were either related to miscanthus biomass production or composition traits (Tables 6 and 7). It must be noted that some genes are repeated in Table 7 as they belong to different QTL clusters. The QTL cluster located in LG5 did not contain any *M. sinensis* genes that were orthologous to cell wall–related genes in sorghum and maize. Regarding the other clusters, some underlying genes were identified for different climatic conditions and different ages (Tables 6 and 7). The 62 M. *sinensis* genes belonged to three main categories: 15 genes (24%) coded for enzymes involved in polysaccharide biosynthesis, 11 genes (18%) coded for enzymes involved in the phenylpropanoid pathway that provides precursors for the biosynthesis of lignin, and five genes (8%) coded for cell-wall proteins. Finally, 17 transcription factors (27%) were also identified based on the miscanthus literature review (Table 7). Finally, twelve genes among these 62 genes belong to families that were previously found to contain genes involved in the secondary cell wall (SCW) biosynthesis of miscanthus and are highlighted in bold in Table 7

Discussion

The stability of the QTL detected for biomass production and composition traits was investigated based on 13 different conditions related to the staggered-start design established in each of the two locations. The QTL of both types of traits were found to be more stable for successive climatic conditions and ages considered together, compared to successive climatic conditions or successive ages considered separately. The evaluation of each climatic condition and each age was made possible based on the staggered-start design: usually, other types of designs such as "single-start" designs lead to the evaluation of plants for a given year, in which the related age and climatic condition cannot be distinguished. According to our design, the biomass production traits appeared to be more stable than the biomass composition traits across the conditions evaluated. However, there was no stable QTL highlighted across both contrasted locations. The QTL clusters representing co-localizations of QTL for biomass production and/or composition traits were identified across 13 different conditions. The corresponding intervals were screened for the underlying genes that correspond to orthologous cell wall–related genes known in sorghum and maize.

Three main points will be discussed in this section: (1) the stability of biomass production and composition traits highlighted across the ages and the climatic conditions of the successive years studied, based on the staggered-start design established in each location; (2) the clusters of QTL for biomass production and composition traits that are consistent with the moderate to high genetic correlations highlighted between these traits; and (3) the different orthologous cell wall–related genes that are known in sorghum and maize, two relatives of miscanthus, and that were found in the regions of the clusters highlighted.

Stable QTL of Biomass Production and Composition Traits Were Highlighted Across Climatic Conditions and Ages Based on the Staggered-Start Design of Each Location, While No Stable QTL Were Detected Across the Two Locations

In each location, stable QTL that corresponded to the QTL detected for a given trait that co-localized under at least two conditions across different climatic conditions and/or across different ages, were identified for biomass production and composition traits. The assessment of QTL stability was an important objective of the study. This is why the staggered-start design was analyzed according to two different linear mixed models, related to each climatic condition and each age. It also made it possible to consider all the genotypes of the population in each location. Regarding biomass production traits, the different climatic conditions and ages considered together in each location led to highlight 59% and 39% of stable QTL in Estrées-Mons and Orléans, respectively. These results can be compared to those reported by Gifford et al. [34] and Dong et al. [36], in which each of the years studied in their experimental designs was not partitioned into age and climatic condition effects. Gifford et al. [34] studied 13 biomass production traits in a *M. sinensis* population over 2 successive years, among which they identified 61% of stable QTL: 22 QTL re-discovered in 2012 out of 36 QTL detected in 2011. Dong et al. [36] established three interconnected miscanthus populations and carried out four different QTL analysis methods, either related to CIM or stepwise analyses. This led to the detection of 288, 264, 133, and 109 QTL for 14 biomass production traits across 2 years. In 2013, they re-identified from 48 to 56% of the QTL that had already been detected in 2012. When climatic conditions and ages were considered separately in each location of our study, around 30% of stable QTL were identified either over the years or across the ages, regardless of the location. Accordingly, these lower proportions result from

Table 7 List of the miscanthus genes that were related to orthologous cell wall–related genes in sorghum and maize (Virlouvet et al., personal communication). Each miscanthus gene was detected in at least one cluster that was detected for a specific condition. When a gene is detected in two clusters, it corresponds to two different conditions. Accordingly, the corresponding stability type is specified. Two types of orthologous relationships were assessed: miscanthus with sorghum and miscanthus with maize. A miscanthus gene can thus correspond to an orthologous sorghum gene that can also have a maize ortholog (in blue). In addition, a miscanthus gene can correspond to an orthologous maize gene that can also have a sorghum ortholog (in green). The black color corresponds to a miscanthus gene that was directly identified based on orthologous relationships with both sorghum and maize. A miscanthus gene ID starts with the root "Misin-", a sorghum gene ID starts with the root "Sobic.0-" and contains a "G" among the gene numbers and a maize gene ID starts with the root "Zm00001d-". A miscanthus gene written with a * corresponds to an identical miscanthus gene, which is displayed in two different rows as the gene was identified in two clusters with different trait types. The miscanthus genes highlighted in bold belong to a gene family for which cell-wall candidate genes were identified in miscanthus by Hu et al. [79] and Zeng et al. [80]

Cluster & (Condition)	Stability type	Miscanthus Gene ID: Misin-	Gene ID of the search list (maize or sorghum list): Sobic.0- or Zm00001d-	Corresponding maize or sorghum orthologous gene ID: Zm00001d- or Sobic.0-	Category	Function	Gene family
LG4_Cluster1 (EM_2018)	-	04G064100	02G068000	019089	Signaling	Unknown	HIP
		04G064800	02G068600	007765	Phenylpropanoid pathway	Hydrolase	P4H
		04G065100	02G069000	019094; 007762; 033748	Polysaccharides	Transferase	GALT
		04G064400	007768	01G051300	Regulation of transcription	Transcription factor	TFbHLH
		04G067000	007781	02G063800	Polysaccharides	Transferase	CHAT
		04G067300	007781	02G063800	Polysaccharides	Transferase	CHAT
		04G067500	007781	02G063800	Polysaccharides	Transferase	CHAT
		04G068400	007781	02G063800	Polysaccharides	Transferase	CHAT
LG4_Cluster2 (EM_Age3)	-	04G120500	02G126600	NA	Oxidation-reduction process	Oxidoreductase	Unknown
		04G129500	**02G139900**	**005305; 043716; 043714; 043719**	**Polysaccharides**	**Synthase**	**CslAL**
		04G144200	**02G158900**	**019957**	**Phenylpropanoid pathway**	**4-coumarate-CoA ligase**	**4CL**
		04G147100	**02G154600**	**019994**	**Cell wall protein**	**Arabinogalactanprotein**	**FLA**
		04G148800	02G168700	NA	Oxidation-reduction process	Oxidoreductase	Unknown
		04G152300	**02G175600**	**NA**	**Phenylpropanoid pathway**	**Transferase**	**HCT**
		04G161400	01G531300	027368	Polysaccharides	Hydrolase	XTH
LG7_Cluster1 (EM_Age2)	-	07G442700	04G282700	017601	Polysaccharides	Reductase	UER
		07G451700	04G274800	NA	Unknown	Unknown	Unknown
		07G453000	04G273800	051520	Regulation of transcription	Transcription factor	TFMYB
		07G455800	04G271300	017720; 051542	DNA regulation	DNA topoispmerase	DNATopo
		07G458100	04G269400	017753	Pectin	Pectin acetylesterase	PAEL
		07G458600	04G268700	017762	Signaling	Hydrolase	CYP707
		07G460900	04G267000	017782	Regulation of transcription	Transcription factor	TFMYB_related
		07G472200	04G228400	018119; 052192	Regulation of transcription	Transcription factor	TFbHLH
		07G475800	04G232300	018090	Cell wall protein	Transporter	Annexin
		07G479100	**04G235900**	**NA**	**Lignification**	**Laccases**	**LAC**
		07G479400	04G236300	052247; 018061	Amino acids	Kinase	SK
		07G480900	04G237800	052260	Polysaccharides	Transferase	GAT
		07G453800	**017712**	**04G271800**	**Regulation of transcription**	**Transcription factor**	**TFWRKY**
		07G477700	018074	04G234500	Cell wall protein	Protease	Aspartyl
LG13_Cluster1 (ORL_2018)	-	**13G113600**	**07G137400**	**050406; 032188**	**Polysaccharides**	**Synthase**	**CslAL**
		13G113800	07G137600	032187; 050409	Cell wall protein	Dehydrogenase	MDH
		13G114900	07G138900	032178	Regulation of transcription	Transcription factor	TFMYB
		13G118000	**07G141200**	**050417**	**Phenylpropanoid pathway**	**Cinnamoyl-CoA reductase**	**CCR**
		13G118600	**07G142200**	**050455**	**Phenylpropanoid pathway**	**Hydroxycinnamoyltransferase**	**HCT**
		13G125700	07G148300	032079	Signaling	Peroxidase	Unknown
		13G130700	07G153700	049991	Regulation of transcription	Transcription factor	TFGeBP
		13G123300	050082	07G146200	Pectin	Pectin methylesterase	PME
		13G134100	050020	07G156200	Polysaccharides	Transferase	FUT
LG15_Cluster1; LG15_Cluster3	Year & Age	15G168200*	08G157000	041418	Regulation of transcription	Transcription factor	IAA
		15G174600*	**08G166100**	**041299**	**Polysaccharides**	**Transferase**	**GT**
(ORL_2017 ; ORL_Age2)		**15G175500***	**08G164800**	**041472**	**Regulation of transcription**	**Transcription factor**	**TFNAC**
		15G183300*	08G190200	031213	Cell wall protein	Unknown	LigA
		15G184700*	09G109300	009888	DNA regulation	DNA primase	DNAP
		15G185800*	08G169650; 08G169800	041504	Polysaccharides	Transferase	GALT
		15G189500*	08G168300	041491	Regulation of transcription	Transcription factor	TFNF-YA
		15G193200	08G185400	031179; 036973; 004870	Signaling	GTP Binding	Ras-related
		15G197500	08G191500	031223	Polysaccharides	Transferase	GlcAT

Table 7 (continued)

LG15_Cluster2 (ORL_2018)	-	15G160900	08G149650	041376; 030911	Regulation of transcription	Transcription factor	TFG2-like
		15G168200*	08G157000	041418	Regulation of transcription	Transcription factor	IAA
		15G174600*	**08G166100**	**041299**	**Polysaccharides**	**Transferase**	**GT**
		15G175500*	**08G164800**	**041472**	**Regulation of transcription**	**Transcription factor**	**TFNAC**
		15G183300*	08G190200	031213	Cell wall protein	Unknown	LigA
		15G184700*	09G109300	009888	DNA regulation	DNA primase	DNAP
		15G185800*	08G169650; 08G169800	041504	Polysaccharides	Transferase	GALT
		15G189500*	08G168300	041491	Regulation of transcription	Transcription factor	TFNF-YA
LG16_Cluster1 (EM_Age2)	-	16G042000	09G029800	035442	Polysaccharides	Glucanase	Glucanase
		16G042400	09G030100	035440	Regulation of transcription	Transcription factor	TFbZIP
		16G044400	09G039100	024809	Regulation of transcription	Transcription factor	TFMYB_related
		16G048700	09G036500	NA	Regulation of transcription	Transcription factor	TFMYB
		16G051300	09G023401	NA	Unknown	Unknown	Unknown
LG18_Cluster1; LG18_Cluster2; LG18_Cluster3 (EM_2018; EM_Age2; EM_Age4)	Year & Age	18G256000	10G254300	014701	Regulation of transcription	Transcription factor	TFG2like
		18G256300	10G254900	014705	Regulation of transcription	Transcription factor	TFC3H
		18G260800	**10G256400**	**014721**	**Polysaccharides**	**Transferase**	**GT**
LG18_Cluster2 (EM_Age2)	-	18G246100	10G245300	014601; 014602	Phenylpropanoid pathway	Peroxidase	PRD
		18G246400	10G245400	014603	Phenylpropanoid pathway	Peroxidase	PRD
		18G246600	10G245800	014608	Phenylpropanoid pathway	Peroxidase	PRD
		18G247400	10G246400	014613	Polysaccharides	Hydrolase	XTH
		18G250300	10G249200	014656	Regulation of transcription	Transcription factor	TFGATA
		18G249600	036499	10G248200	Unknown	Unknown	Unknown

the partition of each year studied into the corresponding climatic condition on the one hand, and the age on the other.

Regarding biomass composition traits, a higher proportion of stable QTL was also identified across the climatic conditions and ages when they were considered together rather than separately. However, these different proportions ranged from 3 to 29%, which was relatively low compared to biomass production traits. Van der Weijde et al. [37] studied an *M. sinensis* population for traits related to biomass composition and conversion efficiency. They reported 23% of stable QTL in 2 successive years: 20 out of 86 QTL were detected in 2013 and 2014. These proportions confirm that, in miscanthus, the QTL of biomass composition traits seem to be less stable across different years (and ages) than those of biomass production traits. This may be explained by strong genotype x climatic condition or genotype x age interactions, as significant genotype x year interactions have already been highlighted for miscanthus biomass composition traits [74, 75].

In each location studied, the proportion of stable QTL for both types of traits across the climatic conditions and/or ages was rather low: this could be expected, as biomass production and composition traits can be affected by the variability related to plant age and environmental factors, such as the related climatic conditions that occur each year [74, 75]. However, these stable QTL mapped across the climatic conditions and/or ages would lead to the identification of relevant targets for MAS programs. Among these, a relevant example was highlighted in LG8 (37 cM) for plant circumference (C50_cm) and aboveground biomass yield (ABM_tDMha), symbolized with a solid red triangle in Fig. 5: these stable QTL are relevant in terms of their stability over different ages and climatic conditions of successive years in Estrées-Mons, especially as they are stable for age 2, 3, and 4. It means that future genetic material could be screened at a young age, in order to select individuals that show beneficial alleles according to this QTL. It could thus speed up miscanthus breeding when based on an early selection of such individuals.

The proportion of stable QTL depends on environmental conditions, plant age and the genetic material considered, which is specific to each miscanthus study [34, 36, 37]. However, these prior studies did not detect QTL for more than 3 years after establishment and did not distinguish the age effect from climatic condition effect. Segura et al. [38] used a staggered-start design and carried out QTL mapping in order to dissect the apple tree architecture into genetic, ontogenetic, and environmental effects. This made it possible to determine the genetic determinism of related traits, with regard to tree ontogeny and climatic conditions. To our knowledge, the present study uses a staggered-start design in miscanthus for the first time, in order to detect stable QTL in different climatic conditions and/or at different ages. Moreover, a staggered-start design was established in each of the two contrasting locations, which led to the examination of QTL stability across locations as well, as had never been done before in miscanthus.

Accordingly, when considering the different years (i.e., climatic conditions) and ages together across locations, no stable QTL were identified for biomass production and composition traits. Thus, it shows that the QTL detected are specific to each location studied. These QTL are relevant for miscanthus breeding programs, as they express themselves in specific conditions related to a given location. It indicates

that other environmental effects interact with the genetic basis of biomass production and composition traits across locations. The different climatic and soil conditions could explain that, as the staggered-start design was established in a deep loam soil in Estrées-Mons, it was established in a sandy soil in Orléans. In addition, the climate in Estrées-Mons is more influenced by the ocean than in Orléans: the differences in climatic conditions between both locations are presented according to different periods of the plant cycle in Raverdy et al. (submitted to BioEnergy Research). Thus, significant genotype × location interactions may explain the lack of stable QTL across locations. In a study comparing different miscanthus species across five different locations in Europe, Clifton-Brown et al. [76] and Lewandowski et al. [77] indeed highlighted genotype × location interactions for biomass production and biomass composition traits. Moreover, 53 additional progenies were grown in Estrées-Mons compared to Orléans (Fig. 1): the genetic variability was therefore not identical between the two locations. This can also be a reason why the genotypic variances (R^2) explained by the QTL detected in Orléans were mostly higher than those explained by the QTL detected in Estrées-Mons. The establishment effect can impact QTL detection power, as miscanthus is mature from around 2 to 3 years [19] or 5 years after establishment [78]: a substantial number of QTL were detected from young to old plants in our study, which suggests that the effect due to the establishment may be limited within the location. However, the different establishment conditions between locations could also explain the lack of stable QTL across locations. Tejera et al. [41] used a staggered-start design and showed that the *M.* × *giganteus* yield response to fertilization was influenced by establishment conditions in each location but not by the plant age.

In this study, each staggered-start design makes it possible to highlight a higher proportion of stable QTL for a range of climatic conditions and ages considered together rather than separately. However, the stability of QTL under these conditions is higher for biomass production traits than for biomass composition traits, studied together for the first time in a miscanthus mapping population. Across locations, no stable QTL were identified, which may be due to different environmental conditions such as climate, soil, and establishment effects. This brings new insights into miscanthus breeding, as stable QTL are needed from different genetic material evaluated across different ages and climatic conditions: the comparison of stable QTL between studies would lead to the identification of the most significant genomic regions associated with biomass production and composition traits. Such QTL that are specific to a given location will benefit to the breeding of miscanthus, notably to target the conditions encountered in a particular region.

Clusters of QTL for Biomass Production and Composition Traits Were Consistent with the Moderate to High Genetic Correlations Highlighted Between These Traits

The QTL clusters identified for biomass production and composition traits were in agreement with the moderate to high genetic correlations between the traits. The QTL clusters related to biomass production traits were identified in LG4, LG7, and LG18. They were made up of QTL that overlapped at similar positions, for traits such as canopy height, total plant height, plant circumference, and aboveground biomass yield. The corresponding significant and moderate to high genetic correlations suggest that QTL overlapping is not random. Moreover, the stability of the QTL cluster is shown in LG18, as QTL were detected in 2018 and at ages 2 and 4 in Estrées-Mons. This is possible based on each staggered-start design evaluated over 5 years in two locations. Gifford et al. [34] identified QTL clusters in LG3 and LG6, which were re-identified in 2 subsequent years and that were consistent with high genetic and phenotypic correlations as well. These clusters were made up of QTL related to the plant circumference, stem diameter, plant stem number, aboveground biomass yield, or characteristics of the leaves such as leaf width, length, and area. These QTL identified for leaf-related traits are relevant: as canopy height refers to the height of the different leaves of the plant that contributes to yield, the QTL identified for canopy height in our study can in fact be related to different phenotypic characteristics of the leaves. However, none of their different QTL clusters were common to our QTL clusters. Dong et al. [36] identified different QTL clusters in their three interconnected miscanthus populations: these clusters were related to plant height, plant circumference, stem volume and density, and the aboveground biomass yield. They were identified in various LG depending on the population and were in agreement with the moderate to high phenotypic and genetic correlations between these traits. For one of their populations originating from a cross between an *M. sinensis* and an *M. sacchariflorus* cultivar, they identified QTL clusters in LG4 and LG7: these LGs were common to the LGs in which we identified QTL clusters for the same type of traits related to plant height, plant circumference, and aboveground biomass yield. However, an investigation of the QTL cluster positions in their study would be based on the alignment with the *M. sinensis* reference genome in order to determine if the same genomic regions are involved.

Regarding biomass composition traits, we identified QTL clusters in LG4, LG5, LG13, LG15, and LG16, which were also in agreement with the significant moderate to high correlations between these traits. The stability of the clusters is also notable, because two clusters were identified in 2017 and for age 2 in Orléans. This is made possible based on the staggered-start design as well. They were located in LG15 and made up

of traits related to cellulose, hemicelluloses, lignin, and ADF contents. The co-localization of QTL related to ADF content with those related to cellulose and lignin content is not surprising, as ADF content represents the sum of cellulose and lignin contents [64]. Van der Weijde et al. [37] identified a major QTL cluster for traits related to conversion efficiency and composition traits: this cluster was located on chromosome 6 according to the *Sorghum bicolor* reference genome that was used in the construction of their two parental miscanthus genetic maps. In miscanthus, the corresponding chromosomes are chromosomes 11 and 12, as the miscanthus genome has been shown to be the result of chromosomal duplication and fusion based on the sorghum genome [33, 35, 42]. However, none of our QTL clusters is common with their QTL clusters, because we did not identify QTL clusters in LG11 and LG12.

Our study was conducted by considering biomass production and composition traits together: this led to the identification of a QTL cluster in LG15, which was made up of both biomass production and composition traits. The corresponding traits were canopy height, NDF (%DM), and hemicellulose content: the moderate and significant correlations of canopy height with these composition traits (respectively, 0.44 and − 0.55) tend to validate the existence of this cluster. However, further analysis according to the genes underlying this cluster is necessary to confirm this assumption.

The QTL clusters identified for biomass production and composition traits could be explained by different genetic factors, such as the pleiotropic effects of the genes underlying these QTL or linked genes. Sometimes, these clusters can originate from genomic regions with segregation distortion, but this may not be possible in our study as we carefully filtered the distorted markers for the construction of our integrated genetic map. The staggered-start design led to the identification of QTL clusters located in LG4, LG5, LG7, LG13, LG15, LG16, and LG18 for a range of climatic conditions and ages that consider biomass production and composition traits together, for the first time in miscanthus.

Orthologous Cell Wall–Related Genes Previously Identified in Sorghum and Maize Enabled the Identification of Putative Cell Wall–Related Genes in *M. sinensis*

Some of the 62 M. *sinensis* genes that were identified in the QTL clusters based on the orthologous cell wall–related genes known in sorghum and maize belong to specific gene families. Twelve genes among the 62 genes belong to families that were previously found to contain genes involved in the secondary cell-wall (SCW) biosynthesis of miscanthus [79, 80]. Hu et al. [79] carried out a transcriptome analysis of genes involved in secondary cell-wall biosynthesis in developing internodes of *M. lutarioriparius*: they highlighted different gene members in specific gene families. These families included genes encoding 4-coumarate-CoA ligase (4CL) and cinnamoyl-CoA reductase (CCR), both involved in the biosynthesis of several classes of phenylpropanoids, as well as laccase (LAC), involved in the polymerization of lignin [81]. They also identified the cellulose synthase–like (CSL) and glycosyltransferase (GT) gene families that are involved in the biosynthesis of cellulose and hemicellulose components in plants. Finally, three other gene families in common with Hu et al. [79] were identified: the fasciclin-like arabinogalactan (FLA) gene family, for which genes are involved in cell wall modification and assembly; the NAC transcription factor (TFNAC) and WRKY transcription factor (TFWRKY) families that contain transcriptional factors for the regulation of secondary cell wall development. Based on genetic and transcriptional analyses in *M.* × *giganteus*, Zeng et al. [80] identified several genes that are common to the genes we highlighted: the 4CL and CCR families that were also reported by Hu et al. [79], as well as the shikimate hydroxycinnamoyl transferase (HCT) family.

Based on these different comparisons, we hypothesize that twelve *M. sinensis* genes out of the 62 genes previously identified are involved in secondary cell wall development. This hypothesis is supported by the fact that these genes were mainly located in the QTL clusters composed of *M. sinensis* biomass composition traits, especially for the clusters located in LG4, LG13, and LG15. Cluster 2 in LG4 and clusters 1 and 3 in LG15 were particularly notable, as the R^2 of the related QTL mainly ranged from 11.3 to 29.4% (Table 6).

Conclusion

In this study, an integrated genetic map of 2770 cM was constructed based on 2602 SNP markers distributed across 19 LGs and was aligned with the released *M. sinensis* reference genome. This integrated genetic map, which was highly saturated, led to the identification of 260 and 283 QTL related to biomass production and composition traits, respectively. The staggered-start design established in each of the two contrasting French locations led to the detection of QTL that were stable across different climatic conditions and different ages. For both types of traits, a higher stability of the QTL was found when the climatic conditions were considered together with the different ages, than when they were considered separately. These differences were highlighted based on the distinction of the plant age effect from climatic condition effect. For a given location, the most stable QTL identified across different climatic conditions and different ages would be interesting for miscanthus breeders, as they are stable regardless of the condition assessed in our experiment. They are thus important resources to carry out future MAS programs. This would be true especially for the QTL which were found to be stable at age 3 and age 4, as they could be relevant for screening young plants without the need to wait for their

mature age. This would be more suited to biomass production traits, as the biomass composition traits were found to be less stable across the conditions. However, no stable QTL were identified across locations: it highlights that the QTL detected in this study were specific to the conditions encountered in Estrées-Mons or in Orléans, and it shows the relevance of carrying out the study in two locations. They may be explained by the existence of QTL that correspond to the genotype × age and genotype × climatic condition interaction effects. These effects were specifically assessed in the two models carried out for the analysis of the staggered-start designs, but their corresponding mapping has not been carried out yet, and their future detection would be desirable.

Clusters of QTL were then identified for biomass production and composition traits under different conditions: this means that linked genes or pleiotropic effects from the genes underlying these QTL would make it possible to jointly improve these different traits. Moreover, these QTL clusters contained 62 M. *sinensis* genes that were orthologous to cell wall–related genes in sorghum and maize. Twelve of these genes were identified as putatively involved in secondary cell wall biosynthesis. In summary, all these QTL clusters which correspond to different traits or stability types and their underlying candidate genes constitute targets of interest for miscanthus breeders, in order to evaluate and create new miscanthus cultivars that would be adapted to different environments, with a high biomass yield and a composition suited to bioenergy, biomaterials or animal bedding.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12155-022-10402-8.

Acknowledgements The authors wish to acknowledge each member of the staff who worked for the establishment and phenotyping of each trial in INRAE experimental stations: the GCIE–Picardie team in Estrées-Mons, the GBFOr team in Orléans, and the IJPB team in Versailles. The authors are grateful to the teams of the INRAE Gentyane Genomic Services platform (Clermont-Ferrand, Puy-de-Dôme, France), the CIRAD Genotyping platform core facility (GPTR, Montpellier, Hérault, France) and the GenoToul GeT platform (Auzeville, Haute-Garonne, France) for DNA extraction, GBS data and sequencing, respectively. The authors also acknowledge David Pot and Mathieu Reymond for providing lists of cell wall–related genes identified in sorghum and maize. They thank Catherine Lapierre for her expertise regarding the lignin biosynthesis. Finally, they thank Rebecca James and Andrea Rau who edited the English of the manuscript.

Author Contribution Maryse Brancourt-Hulmel carried out the conception and the design of the study. Kristelle Lourgant, Emilie Mignot, Stéphanie Arnoult, Guillaume Bodineau, and Raphaël Raverdy prepared the material and carried out phenotyping data management and collection. Yves Griveau realized NIRS data collection and predictions. Raphaël Raverdy, Maryse Brancourt-Hulmel, Kristelle Lourgant, and Cristiane H. Taniguti contributed to the analysis of the data. Raphaël Raverdy wrote the manuscript and all authors commented on the manuscript.

Funding The funding for this research was provided by the French National Research Agency (Agence Nationale de la Recherche, ANR), grant ANR-11-BTBR-0006-BFF in the frame of the program Investments for the Future.

Data Availability This work was performed with the Plant Bioinformatics Facility of INRAE-URGI: https://urgi.versailles.inra.fr/ephesis/ephesis/ontologyportal.do

https://doi.org/10.15454/1NVRNJ for datasets related to phenotyping resources.

https://doi.org/10.15454/ET8MJE for dataset related to genotyping resources.

Declarations

Ethics Approval and Consent to Participate Not applicable.

Consent for Publication All authors agree to publish this article in BioEnergy Research.

Competing Interests The authors declare no competing interests.

References

1. Jones M, Walsh M (2001) Miscanthus for energy and fibre. James & James (Science Publishers) Ltd
2. Heaton EA, Clifton-Brown J, Voigt TB et al (2004) Miscanthus for renewable energy generation: European Union experience and projections for Illinois. Mitig Adapt Strateg Glob Chang 9:433–451. https://doi.org/10.1023/B:MITI.0000038848.94134.be
3. Johnson M, Tucker N, Barnes S, Kirwan K (2005) Improvement of the impact performance of a starch based biopolymer via the incorporation of *Miscanthus giganteus* fibres. Ind Crops Prod 22:175–186. https://doi.org/10.1016/j.indcrop.2004.08.004
4. Girones J, Vo L, Arnoult S et al (2016) Miscanthus stem fragment – reinforced polypropylene composites: Development of an optimized preparation procedure at small scale and its validation for differentiating genotypes. Polym Test 55:166–172. https://doi.org/10.1016/j.polymertesting.2016.08.023
5. Lewandowski I, Clifton-Brown J, Kiesel A et al (2018) *Miscanthus*. Perenn Grasses Bioenergy Bioprod 23:35–59. https://doi.org/10.1016/b978-0-12-812900-5.00002-3
6. Acikel H (2011) The use of *Miscanthus (Giganteus)* as a plant fiber in concrete production. Sci Res Essays 6:2660–2667. https://doi.org/10.5897/SRE10.1139
7. Anderson E, Arundale R, Maughan M et al (2011) Growth and agronomy of *Miscanthus* × *giganteus* for biomass production. Biofuels 2:167–183. https://doi.org/10.4155/bfs.10.80
8. Hodkinson TR, Klaas M, Jones MB et al (2015) Miscanthus: a case study for the utilization of natural genetic variation. Plant Genet Resour Characterisation Util 13:219–237. https://doi.org/10.1017/S147926211400094X
9. Clifton-Brown JC, Neilson B, Lewandowski I, Jones MB (2000) The modelled productivity of *Miscanthus* × *giganteus* (GREEF et DEU) in Ireland. Ind Crop Prod 12:97–109
10. Heaton EA, Dohleman FG, Long SP (2008) Meeting US biofuel goals with less land: the potential of miscanthus. Glob Chang Biol 14:2000–2014. https://doi.org/10.1111/j.1365-2486.2008.01662.x
11. Greef J, Deuter M (1993) Syntaxonomy of *Miscanthus x giganteus* Greed et Deu. Angew Bot 67:87–90
12. Linde-Laursen I (1993) Cytogenetic analysis of *Miscanthus'Giganteus'*, an interspecific hybrid. Hereditas 119:297–300. https://doi.org/10.1111/j.1601-5223.1993.00297.x

13. Sun Q, Lin Q, Yi Z-L et al (2010) A taxonomic revision of *Miscanthus s.l.* (Poaceae) from China. Bot J Linn Soc 164:178–220. https://doi.org/10.1111/j.1095-8339.2010.01082.x
14. Sacks EJ, Juvik JA, Lin Q, et al (2013) The gene pool of *Miscanthus* species and its improvement. In: Paterson A. (eds) Genomics of the Saccharinae. Plant Genetics and Genomics: Crops and Models, vol 11. Springer, New York, pp 73–101
15. Clark LV, Brummer JE, Głowacka K et al (2014) A footprint of past climate change on the diversity and population structure of *Miscanthus sinensis*. Ann Bot 114:97–107. https://doi.org/10.1093/aob/mcu084
16. Clark LV, Ryan Stewart J, Nishiwaki A et al (2015) Genetic structure of *Miscanthus sinensis* and *Miscanthus sacchariflorus* in Japan indicates a gradient of bidirectional but asymmetric introgression. J Exp Bot 66:4213–4225. https://doi.org/10.1093/jxb/eru511
17. Anzoua KG, Suzuki K, Fujita S et al (2015) Evaluation of morphological traits, winter survival and biomass potential in wild Japanese *Miscanthus sinensis* Anderss. populations in northern Japan. Grassl Sci 61:83–91. https://doi.org/10.1111/grs.12085
18. Głowacka K (2011) A review of the genetic study of the energy crop miscanthus. Biomass Bioenerg 35:2445–2454. https://doi.org/10.1016/j.biombioe.2011.01.041
19. Clifton-brown JC, Chiang Y, Hodkinson TR (2008) Miscanthus genetic resources and breeding potential. In: Vermerris W (ed) Genetic Improvement of Bioenergy Crops. Springer Science, Berlin/Heidelberg, pp 273–290
20. Zub HW, Brancourt-Hulmel M (2010) Agronomic and physiological performances of different species of *Miscanthus*, a major energy crop. A review Agron Sustain Dev 30:201–214. https://doi.org/10.1051/agro/2009034
21. Conson ARO, Taniguti CH, Amadeu RR et al (2018) High-resolution genetic map and QTL analysis of growth-related traits of *Hevea brasiliensis* cultivated under suboptimal temperature and humidity conditions. Front Plant Sci 9:1–16. https://doi.org/10.3389/fpls.2018.01255
22. Balsalobre TWA, da Silva PG, Margarido GRA et al (2017) GBS-based single dosage markers for linkage and QTL mapping allow gene mining for yield-related traits in sugarcane. BMC Genomics 18:1–19. https://doi.org/10.1186/s12864-016-3383-x
23. Souza LM, Gazaffi R, Mantello CC et al (2013) QTL mapping of growth-related traits in a full-sib family of rubber tree (*Hevea brasiliensis*) evaluated in a sub-tropical climate. PLoS ONE 8. https://doi.org/10.1371/journal.pone.0061238
24. Grattapaglia D, Sederoff R (1994) Genetic linkage maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers. Genetics 137:1121–1137
25. Wu R, Ma CX, Painter I, Zeng ZB (2002) Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. Theor Popul Biol 61:349–363. https://doi.org/10.1006/tpbi.2002.1577
26. Wu R, Ma CX, Wu SS, Zeng ZB (2002) Linkage mapping of sex-specific differences. Genet Res 79:85–96. https://doi.org/10.1017/S0016672301005389
27. Gazaffi R, Margarido GRA, Pastina MM et al (2014) A model for quantitative trait loci mapping, linkage phase, and segregation pattern estimation for a full-sib progeny. Tree Genet Genomes 10:791–801. https://doi.org/10.1007/s11295-013-0664-2
28. Atienza SG, Satovic Z, Petersen KK et al (2002) Preliminary genetic linkage map of *Miscanthus sinensis* with RAPD markers. Theor Appl Genet 105:946–952. https://doi.org/10.1007/s00122-002-0956-7
29. Atienza SG, Satovic Z, Petersen KK et al (2003) Identification of QTLs influencing agronomic traits in *Miscanthus sinensis* Anderss. I. Total height, flag-leaf height and stem diameter. Theor Appl Genet 107:123–129. https://doi.org/10.1007/s00122-003-1220-5
30. Atienza SG, Satovic Z, Petersen KK et al (2003) Identification of QTLs associated with yield and its components in *Miscanthus sinensis* Anderss. Euphytica 132:353–361. https://doi.org/10.1023/A:1025041926259
31. Atienza SG, Satovic Z, Petersen KK et al (2003) Identification of QTLs influencing combustion quality in *Miscanthus sinensis* Anderss. II. Chlorine and potassium content. Theor Appl Genet 107:857–863. https://doi.org/10.1007/s00122-003-1218-z
32. Atienza SG, Satovic Z, Petersen KK et al (2003) Influencing combustion quality in *Miscanthus sinensis* Anderss.: identification of QTLs for calcium, phosphorus and sulphur content. Plant Breed 122:141–145. https://doi.org/10.1046/j.1439-0523.2003.00826.x
33. Ma XF, Jensen E, Alexandrov N et al (2012) High resolution genetic mapping by genome sequencing reveals genome duplication and tetraploid genetic structure of the diploid *Miscanthus sinensis*. PLoS ONE 7. https://doi.org/10.1371/journal.pone.0033821,pp1-11
34. Gifford JM, Chae WB, Swaminathan K et al (2015) Mapping the genome of *Miscanthus sinensis* for QTL associated with biomass productivity. GCB Bioenergy 7:797–810. https://doi.org/10.1111/gcbb.12201
35. Swaminathan K, Chae WB, Mitros T et al (2012) A framework genetic map for *Miscanthus sinensis* from RNAseq-based markers shows recent tetraploidy. BMC Genomics 13:142. https://doi.org/10.1186/1471-2164-13-142
36. Dong H, Liu S, Clark LV et al (2018) Genetic mapping of biomass yield in three interconnected miscanthus populations. GCB Bioenergy 10:165–185. https://doi.org/10.1111/gcbb.12472
37. Van der Weijde T, Kamei CLA, Severing EI et al (2017) Genetic complexity of miscanthus cell wall composition and biomass quality for biofuels. BMC Genomics 18:1–15. https://doi.org/10.1186/s12864-017-3802-7
38. Segura V, Durel C, Costes E (2009) Dissecting apple tree architecture into genetic, ontogenetic and environmental effects : QTL mapping. Tree Genet Genomes 165–179. https://doi.org/10.1007/s11295-008-0181-x,pp
39. Loughin TM (2006) Improved experimental design and analysis for long-term experiments. Crop Sci 46:2492–2502. https://doi.org/10.2135/cropsci2006.04.0271
40. Segura V, Cilas C, Segura V et al (2008) Dissecting apple tree architecture into genetic, ontogenetic and environmental effects : mixed linear modelling of repeated spatial and temporal measures. New Phytol. https://doi.org/10.1111/j.1469-8137.2007.02374.x,pp165-179
41. Tejera M, Boersma N, Vanloocke A et al (2019) Multi-year and multi-site establishment of the perennial biomass crop *Miscanthus* × *giganteus* using a staggered start design to elucidate N response. Bioenergy Res 12:471–483. https://doi.org/10.1007/s12155-019-09985-6
42. Mitros T, Session AM, James BT et al (2020) Genome biology of the paleotetraploid perennial biomass crop miscanthus. Nat Commun 11:1–11. https://doi.org/10.1038/s41467-020-18923-6,pp1-11
43. Rambaud C, Arnoult S, Bluteau A et al (2013) Shoot organogenesis in three *Miscanthus* species and evaluation for genetic uniformity using AFLP analysis. Plant Cell Tissue Organ Cult 113:437–448. https://doi.org/10.1007/s11240-012-0284-9
44. Dagnelie P (2012) Principes d'expérimentation : planification des expériences et analyse de leurs résultats, Les Presse
45. Elshire RJ, Glaubitz JC, Sun Q et al (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE 6:1–10. https://doi.org/10.1371/journal.pone.0019379
46. Cormier F, Lawac F, Maledon E et al (2019) A reference high-density genetic map of greater yam (*Dioscorea alata L.*). Theor Appl Genet 132:1733–1744. https://doi.org/10.1007/s00122-019-03311-6

47. Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc
48. Glaubitz JC, Casstevens TM, Lu F et al (2014) TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. PLoS ONE 9. https://doi.org/10.1371/journal.pone.0090346
49. Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. https://doi.org/10.1038/nmeth.1923
50. Garsmeur O, Droc G, Antonise R et al (2018) A mosaic monoploid reference sequence for the highly complex genome of sugarcane. Nat Commun 9:1–10. https://doi.org/10.1038/s41467-018-05051-5
51. Margarido GRA, Souza AP, Garcia AAF (2007) OneMap: Software for genetic mapping in outcrossing species. Hereditas 144:78–79. https://doi.org/10.1111/j.2007.0018-0661.02000.x
52. Margarido GRA, Mollinari M, Broman K, et al (2020) OneMap: Software for constructing genetic maps in experimental crosses: full-sib, RILs, F2 and backcrosses. R package version 2.2.0
53. Kosambi DD (1943) The estimation of map distances from recombination values. Ann Eugen 12:172–175. https://doi.org/10.1111/j.1469-1809.1943.tb02321.x
54. Mollinari M, Garcia AAF (2019) Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden Markov models. G3 Genes. Genomes, Genet 9:3297–3314. https://doi.org/10.1534/g3.119.400378
55. Mollinari M, Olukolu BA, da Pereira GS et al (2020) Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. G3 Genes. Genomes, Genet 10:281–292. https://doi.org/10.1101/689638
56. Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. Proc Natl Acad Sci U S A 84:2363–2367. https://doi.org/10.1073/pnas.84.8.2363
57. Preedy KF, Hackett CA (2016) A rapid marker ordering approach for high-density genetic linkage maps in experimental autotetraploid populations using multidimensional scaling. Theor Appl Genet 129:2117–2132. https://doi.org/10.1007/s00122-016-2761-8
58. Preedy KF, Hackett CA, Boskamp B (2018) MDSMap: High density linkage maps using multi-dimensional scaling. 1–10
59. Hackett CA, Broadfoot LB (2003) Effects of genotyping errors, missing values and segregation distortion in molecular marker data on the construction of linkage maps. Heredity (Edinb) 90:33–38. https://doi.org/10.1038/sj.hdy.6800173
60. Cartwright DA, Troggio M, Velasco R, Gutin A (2007) Genetic mapping in the presence of genotyping errors. Genetics 176:2521–2527. https://doi.org/10.1534/genetics.106.063982
61. Bilton TP, Schofield MR, Black MA, et al (2018) Accounting for errors in low coverage high-throughput sequencing data when constructing genetic maps using biparental outcrossed populations. bioRxiv 209:65–76. https://doi.org/10.1101/249722
62. Taniguti CH (2021) Building highly saturated genetic maps with OneMap 3.0: new approaches using workflows
63. Steinbach D, Alaux M, Amselem J et al (2013) GnpIS: An information system to integrate genetic and genomic data from plants and fungi. Database 2013:1–9. https://doi.org/10.1093/database/bat058
64. Van Soest PJ, Wine RH (1967) Use of detergents in the analysis of fibrous feeds. IV. Determination of plant cell-wall constituents. J AOAC Int 50:50–55. https://doi.org/10.1093/jaoac/50.1.50
65. Belmokhtar N, Arnoult S, Chabbert B et al (2017) Saccharification performances of miscanthus at the pilot and miniaturized assay scales: Genotype and year variabilities according to the biomass composition. Front Plant Sci 8:1–13. https://doi.org/10.3389/fpls.2017.00740
66. Henderson CR (1984) Applications of linear models in animal breeding. Univ Guelph, Guelph
67. Muñoz F, Sanchez L (2019) BreedR: Statistical methods for forest genetic resources analysts. R package version 0.12–4
68. Costa e Silva J, Dutkowski GW, Gilmour AR, (2001) Analysis of early tree height in forest genetic trials is enhanced by including a spatially correlated residual. Can J For Res 31:1887–1893. https://doi.org/10.1139/cjfr-31-11-1887
69. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19:716–723. https://doi.org/10.1109/TAC.1974.1100705
70. Gazaffi R, Amadeu RR, Mollinari M, et al (2020) FullsibQTL: an R package for QTL mapping in biparental populations of outcrossing species
71. Chen L, Storey JD (2006) Relaxed significance criteria for linkage analysis. Genetics 173:2371–2381. https://doi.org/10.1534/genetics.105.052506
72. Lander ES, Botstein S (1989) Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185. https://doi.org/10.1093/genetics/121.1.185
73. Wei T, Simko V (2017) R package "corrplot": Visualization of a correlation matrix (version 0.84)
74. Allison GG, Morris C, Clifton-Brown J et al (2011) Genotypic variation in cell wall composition in a diverse set of 244 accessions of *Miscanthus*. Biomass Bioenerg 35:4740–4747. https://doi.org/10.1016/j.biombioe.2011.10.008
75. Arnoult S, Brancourt-Hulmel M (2015) A review on miscanthus biomass production and composition for bioenergy use: Genotypic and environmental variability and implications for breeding. Bioenergy Res 8:502–526. https://doi.org/10.1007/s12155-014-9524-7
76. Clifton-Brown JC, Lewandowski I, Andersson B et al (2001) Performance of 15 miscanthus genotypes at five sites in Europe. Agron J 93:1013–1019. https://doi.org/10.2134/agronj2001.9351013x
77. Lewandowski I, Andersson B, Basch G et al (2003) Environment and harvest time affects the combustion qualities of miscanthus genotypes. Agron J 1274–1280. https://doi.org/10.2134/agronj2003.1274
78. Lesur C, Jeuffroy MH, Makowski D et al (2013) Modeling long-term yield trends of *Miscanthus*×*giganteus* using experimental data from across Europe. F Crop Res 149:252–260. https://doi.org/10.1016/j.fcr.2013.05.004
79. Hu R, Xu Y, Yu C et al (2017) Transcriptome analysis of genes involved in secondary cell wall biosynthesis in developing internodes of *Miscanthus lutarioriparius*. Sci Rep 7:1–16. https://doi.org/10.1038/s41598-017-08690-8
80. Zeng X, Sheng J, Zhu F et al (2020) Genetic, transcriptional, and regulatory landscape of monolignol biosynthesis pathway in *Miscanthus* × *giganteus*. Biotechnol Biofuels 13:1–14. https://doi.org/10.1186/s13068-020-01819-4
81. Vanholme R, Demedts B, Morreel K et al (2010) Lignin biosynthesis and structure. Plant Physiol 153:895–905. https://doi.org/10.1104/pp.110.155119