University of São Paulo
"Luiz de Queiroz" College of Agriculture

Genetic mapping in a biparental *Megathyrsus maximus* (Jacq.) population with allele dosage information

Gabriel de Siqueira Gesteira

Thesis presented to obtain the degree of Doctor in Science.
Area: Genetics and Plant Breeding

Piracicaba
2021

Gabriel de Siqueira Gesteira
Bachelor in Agronomy

Genetic mapping in a biparental *Megathyrsus maximus* (Jacq.) population with allele dosage information

Advisor:
Prof. Dr. **ANTONIO AUGUSTO FRANCO GARCIA**

Thesis presented to obtain the degree of Doctor in Science.
Area: Genetics and Plant Breeding

Piracicaba
2021

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Gesteira, Gabriel de Siqueira

Genetic mapping in a biparental *Megathyrsus maximus* (Jacq.) population with allele dosage information / Gabriel de Siqueira Gesteira. – – Piracicaba, 2021 .

131 p.

Tese (Doutorado) – – USP / Escola Superior de Agricultura "Luiz de Queiroz".

1. Mapa de ligação 2. Mapeamento de QTL 3. Poliploide 4. Autotetraploide 5. Forrageira I. Título.

DEDICATORY

To my parents Clóvis and Ana Flora,
my sister Ana Luiza,
and my son Henrique.
I love you!

ACKNOWLEDGEMENTS

To the University of São Paulo (USP), especially the "Luiz de Queiroz" College of Agriculture (ESALQ) for providing the necessary structure for my professional qualification and development of the current work. I am thankful to the the Genetics Department, all professors and staff of both Statistics and Genetics and Plant Breeding Graduate Programs (ESALQ/USP) for contributing to the development of my scientific and professional skills, as well as the helpful contributions during the conduction of this work.

To the Brazilian funding agencies CNPq ("Conselho Nacional de Desenvolvimento Científico e Tecnológico") and CAPES ("Coordenação de Aperfeiçoamento de Pessoal de Nível Superior") for providing the scholarships and financial support to conduct this work, and especially to CAPES for providing me the opportunity of conducting part of my research abroad through the PrInt program. I am also grateful for the additional financial support provided by the UGPN (University Global Partnership Network) to conduct part of this research abroad.

To Embrapa Cattle Beef and UNIPASTO ("Associação para o Fomento à Pesquisa de Melhoramento de Forrageiras") for providing support and conditions to obtain the genotypic and phenotypic data for the mapping population, as well as additional training through courses and laboratory analysis.

To the North Carolina State University (NCSU), the Department of Horticultural Science, and the Bioinformatics Research Center (BRC) for providing the structure to conduct part of this work, as well as the opportunity to develop my scientific skills and share experiences with amazing scientists.

To my advisor, Prof. Dr. Antonio Augusto Franco Garcia, for giving me the opportunity of being part of his excellent scientific group. I am really thankful to all opportunities, training, and support you provided me during the course of my Ph.D., and the inestimable contributions and advices for both my personal, professional, and scientific development. Thank you for inspiring me as an example of professor and scientist.

To Prof. Dr. Zhao-Bang Zeng for providing me the opportunity to join his brilliant group and have an excellent scientific training at NCSU. Thank you for promoting insightful scientific discussions, and for your valuable contributions to my scientific career.

To Prof. Dr. Marcelo Mollinari for providing me excellent training on statistical genetics and linkage mapping in polyploid species, and allowing me to contribute to his awesome work on the MAPpoly package. Thank you for all the scientific discussions and collaborations, as well as valuable opportunities and advices for my professional career.

To Prof. Dr. Guilherme da Silva Pereira for great scientific discussions, collaborations, and valuable advices for my professional development. I am also grateful for your essential help during my move to Raleigh.

To Dr. Mateus Figueiredo Santos for all scientific discussions, collaborations, opportunities, assistance, for sharing his great experience with plant breeding, and for always being accessible. My sincere thanks are extended to all the Embrapa team involved in this work: Dra. Liana Jank, Dra. Mariane de Mendonça Vilela, Dra. Andrea Raposo, and Dra. Lucimara Chiari.

To Prof. Dr. Augusto Garcia, Dr. Kaio Olímpio Dias, and Me. Getulio Caixeta Ferreira for reviewing and providing insightful contributions to the writing and presentation of this thesis.

To all my friends and former colleagues from the Statistical Genetics Lab (ESALQ/USP): João Ricardo Rosa, Luís Felipe Ferrão, Cristiane Taniguti, Elaine Batista, Letícia Lara, Jhonathan Pedroso, Mariana Niederheitmann, Marianella Quezada, Matheus Krause, Rafael Nalin, Rodrigo Amadeu, Mayara Oliveira, Jéssica Zílio, and Pedro Jaloto. I am especially grateful to the current Lab members: Kaio Olímpio, Getúlio Ferreira, Camila Godoy, Vitória Bizão, Wellingson Araújo, Gabriel Mamédio, Rafaela

Graciano, and Laura Polizel. Thank you all for the great scientific discussions, collaborations, support and friendship.

To colleagues and collaborators from other institutions that somehow contributed to the present work. I am especially thankful to Dra. Letícia Castro Lara and Dra. Rebecca Ulbricht Ferreira for sharing experiences about forage research and genomic analysis, and to Dr. Marco Pessoa for sharing the *Urochloa ruziziensis* reference genome on first hand.

To all my friends and colleagues from the Genetics Department, the GVENCK group, ESALQ/USP, and my sincere thanks to all my friends from Piracicaba/SP, Ribeirão Preto/SP, Lavras/MG, and Raleigh/NC. Many thanks to you all for your support and friendship.

To my parents Ana Flora and Clóvis, my sister Ana Luiza, and my son Henrique. I am especially thankful for your support, love, patience, and comprehension. Thank you for inspiring me and making everything possible.

SUMMARY

Resumo 7
Abstract 8
1 Introduction 9
References 10
2 MAPpoly: building multilocus genetic linkage maps in autopolyploids 15
2.1 Introduction 15
References 16
3 Genetic linkage and QTL mapping in *Megathyrsus maximus* (Jacq.) with multiple dosage markers 19
3.1 Introduction 19
3.2 Final considerations 21
References 21
4 Final considerations 27

RESUMO

Mapeamento genético em uma população biparental de *Megathyrsus maximus* (Jacq.) utilizando informação de dosagem alélica

As forrageiras são amplamente difundidas e cultivadas em fazendas em todo o mundo, e utilizadas principalmente para alimentar o gado, constituindo em uma importante fonte de sustentabilidade econômica e ambiental. Uma das gramíneas de maior produtividade utilizadas como forrageira é o capim-colonião (*Megathyrsus maximus* Jacq.), que apresenta alta qualidade nutricional e tolerância a diversos fatores bióticos e abióticos. A espécie combina a vantagem da recombinação genética por meio de cruzamentos sexuais com a capacidade de fixar o vigor híbrido em genótipos superiores e propagá-los por sementes via apomixia. No entanto, pouco se sabe sobre seu comportamento genômico, principalmente devido à alta complexidade de seu genoma autopoliploide. Neste trabalho, implementamos métodos de última geração para construir mapas de ligação em espécies autopoliploides, combinados com uma abordagem multiponto HMM (*Hidden Markov Model*). O software MAPpoly pode construir mapas de ligação considerando níveis de ploidia até 12, importar dados de software de terceiros e exportar mapas e probabilidades condicionais dos genótipos para análises posteriores. O software MAPpoly é fácil de usar e está disponível gratuitamente em versões estáveis e de desenvolvimento. Utilizamos o MAPpoly para construir um mapa de ligação denso e informativo para *M. maximus* considerando marcadores com múltiplas dosagens alélicas, e utilizamos um método de última geração para procurar por QTLs no genoma considerando as características de maior relevância para o melhoramento de *M. maximus*: altura e área da touceira, produção de massa verde, proporção de lâmina foliar, produção de matéria seca foliar e total, densidade volumétrica foliar e total, capacidade de rebrota, e taxa de alongamento foliar. Extraímos o DNA de amostras de folhas de uma população de mapeamento biparental contendo 224 indivíduos e os sequenciamos por meio do protocolo GBS (*Genotyping-by-Sequencing*). Os dados brutos de sequenciamento foram analisados para encontrar variantes e estimar as dosagens alélicas para ambos os pais e todos os indivíduos da população. Cinco genomas de referência de espécies relacionadas foram utilizados durante o processo de mapeamento, devido à ausência de um genoma de referência para *M. maximus*. Em seguida, construímos um mapa de ligação e utilizamos observações fenotípicas das características fenotípicas mais relevantes para isolar o componente genético da população e buscar por QTLs ao longo do genoma, utilizando um modelo de regressão aleatório. Construímos o mapa de ligação mais denso e informativo para a espécie até o momento, com 7095 marcadores abrangendo 1746.18 cM do genoma. Não houve evidência de dupla redução ou pareamento preferencial na população considerada no estudo. Encontramos dez QTLs associados a sete características que são relevantes para o melhoramento de *M. maximus*, com herdabilidades no sentido restrito variando de 0.4127 a 0.1387. A implementação do software e as análises genéticas fornecidas neste trabalho podem ajudar a entender a organização genômica e resolver incertezas a respeito do posicionamento evolutivo e taxonômico de *M. maximus*, bem como auxiliar na construção de um genoma de referência para a espécie. A análise de QTLs fornece uma compreensão mais aprofundada a respeito do controle genético das características relevantes para a espécie, e pode ser utilizada para auxiliar na implementação e aprimoramento de estratégias de seleção baseadas em marcadores moleculares, aumentando a eficiência dos ciclos de seleção em programas de melhoramento de *M. maximus*.

Palavras-chave: Mapa de ligação, Mapeamento de QTL, Poliploide, Autotetraploide, Forrageira

ABSTRACT

Genetic mapping in a biparental *Megathyrsus maximus* (Jacq.) population with allele dosage information

Forage crops are widespread across farmlands worldwide and primarily used to feed livestock, consisting of an important source of economic and environmental sustainability. One of the highest yielding grasses used as a forage crop is the guineagrass (*Megathyrsus maximus* Jacq.), which presents high nutritional quality and tolerance to many biotic and abiotic factors. The species combines the advantage of genetic recombination through sexual crosses with the ability to fix hybrid vigor in superior genotypes and propagate them by seeds via apomixis. However, little is known about its genomic behavior, mainly due to the high complexity of its autopolyploid genome. In this work, we implemented state-of-the-art methods to construct genetic linkage maps in autopolyploid species, coupled with a multipoint Hidden Markov Model approach. The software MAPpoly can construct genetic linkage maps for ploidy levels up to 12, import data from third-party software, and export maps and genotypic conditional probabilities for further analysis. MAPpoly is easy-to-use and freely available in stable and development versions. We used MAPpoly to construct a dense and informative genetic linkage map for *M. maximus* using multiple dosage markers, then used a state-of-the-art method to search for QTL along the genome considering relevant traits for *M. maximus* breeding: canopy height and area, total yield, proportion of leaf blades, foliar and total dry matter yield, leaf and total volumetric density, regrowth capacity, and leaf elongation rate. We extracted DNA from leaf samples of a biparental mapping population containing 224 individuals and sequenced them through the GBS (Genotyping-by-Sequencing) protocol. Raw sequencing data were analyzed to find variants and call genotype dosages for both parents and all individuals in the population. We used five reference genomes of related species during the mapping process due to the absence of a reference genome for *M. maximus*. Then, we constructed the genetic linkage map and used phenotypic observations of selected traits to isolate the genetic component of the population performance and search for QTL regions along the genome, using a random regression model. We constructed the densest and informative linkage map for the species up to date, with 7095 markers spanning 1746.18 cM of the species genome. There was no evidence of double reduction or preferential pairing in the study population. We found ten QTL associated with seven traits that are relevant to *M. maximus* breeding, with narrow-sense heritabilities ranging from 0.4127 to 0.1387. The software implementation and the genetic analysis provided in this work can help untangle the genomic organization and solve uncertainties regarding *M. maximus* evolutionary and taxonomic placement, as well as help to assemble the species genome. The QTL analysis provides a better understanding of the complex genomic behavior involved in the genetic control of relevant traits and may be used to support marker-based selection strategies, thus increasing the efficiency of selection cycles in *M. maximus* breeding programs.

Keywords: Linkage map, QTL mapping, Polyploid, Autotetraploid, Forage crop

1 INTRODUCTION

The species *Megathyrsus maximus* Jacq. stands out from cultivated forage crops due to its highly favorable agronomic characteristics, including high yield associated with great nutritional quality, which guarantees a good economic return by cultivated area (Jank *et al.*, 2014). The species represents the highest yielding forage crop that propagates by seeds in the Brazilian market. Brazil has approximately 172 mi ha covered by pastures, of which 20% are destinated to *M. maximus* (Martuscello *et al.*, 2007; do Valle *et al.*, 2009). Due to its relevance to livestock and agriculture, and to the highly varying abiotic conditions in Brazil, the breeding and development of elite cultivars adapted to different target regions is necessary.

The propagation of *M. maximus* is most commonly performed by seeds, which can arise from two different reproductive systems: sexual recombination via cross-pollination, and asexual propagation via apomixis (apospory). The former is exclusively associated with natural diploids (2n = 2x = 16), while the latter is associated with natural tetraploid individuals (2n = 4x = 32). The genetic control of apomixis has a simple inheritance with one or few loci in a linkage block, and the cross of sexual and apomictic individuals is enabled through chromosomal duplication of sexual diploids with colchicine. Thus, recombinant progenies of such crosses usually segregate for apomixis in a 1:1 rate (Savidan *et al.*, 1989).

The breeding of *M. maximus* in Brazil started with the introduction of cultivars from Africa, and has been conducted following a recurrent selection scheme between sexual and apomictic accessions, based on phenotypic observations and measurements for the most relevant agronomic traits. Furthermore, genotypes need to pass through an additional step of evaluation of their performance under grazing. Due to the need to evaluate the persistency and repeatability of all relevant traits across many harvests, the entire breeding process can take up to 15 years to the cultivar release (Resende *et al.*, 2014).

The development of next-generation sequencing technologies, such as the GBS (Genotyping-by-Sequencing), enables the obtention of a great quantity of genomic data, thus allowing genomic studies in almost all species (Elshire *et al.*, 2011). In addition to diversity and genetic architecture studies, the availability of high throughput molecular data allows the development of novel techniques for genomic association and prediction in breeding programs. While the development and adoption of such techniques are advanced for model species, such as *Arabidopsis thaliana*, soybean, maize, and wheat, there is a lack of such resources for complex species and polyploids, like *M. maximus*.

The obtention and evaluation of variants along the genome also allows the estimation of genotypes for both diploids and polyploids. In polyploid species, genotypes are usually described through dosages that represent the estimated number of copies of the alternate allele for a given biallelic locus (Garcia *et al.*, 2013; Serang *et al.*, 2012; Gerard *et al.*, 2018; Voorrips *et al.*, 2011; Hackett *et al.*, 2013). For an autotetraploid, there are five possible genotypes with dosages varying between zero and four (Gallais, 2003; Lara *et al.*, 2019). With dosage markers in hands, it is possible to perform further genomic studies, including genetic linkage and QTL (Quantitative Trait Loci) analysis (Hackett *et al.*, 2013; Bourke *et al.*, 2018b; Mollinari and Garcia, 2019).

Linkage analysis has been used to study genomic conformity and the inheritance pattern in targeted populations. Such analysis allows the detection of homology groups, the genetic distances and recombination fractions between markers, as well as the phase configurations in the population founders. When using state-of-the-art methods, such as Mollinari and Garcia (2019), it is also possible to get full haplotypes for all founders and individuals in the populations, study the meiotic process and chromosomal pairing patterns, and use the full genetics of the population in further studies through the genotype probabilities, obtained for all individuals and all genomic positions (Mollinari *et al.*, 2019). The vast majority of linkage maps for autopolyploid species are based on single-dosage markers,

i.e. markers that have only one copy of the reference or the alternate allele for one or both parents, so many multi-dosage markers that may cover important genomic regions are not included in the linkage maps, thus resulting in a loss of genetic information. This source of information is important to connect single-dosage markers and homologs through a multipoint procedure (Mollinari and Garcia, 2019; Mollinari *et al.*, 2019), then allow for a more informative and robust map with greater applicabilities, including further analysis such as QTL mapping for relevant traits Serang *et al.* (2012); Hackett *et al.* (2014); Pereira *et al.* (2018); Mollinari and Garcia (2019); Pereira *et al.* (2020).

With the genotype probabilities in hand, it is possible to study the behavior of complex traits through QTL mapping. It consists of methods that combine both phenotypic and genomic data with statistical genetics models to allow the study of the genetic architecture of complex traits (Pereira *et al.*, 2020). The study of the genetic architecture through QTL mapping is the first step towards more advanced studies, such as fine mapping and differential expression studies, but can also be used as a basis for some decision-making processes in breeding programs, including QTL cloning and marker-assisted selection (Jaganathan *et al.*, 2020). Thus, the study has the potential to reduce the time and costs related to selection cycles in breeding programs.

QTL mapping has been widely used to study the behavior of complex traits for diploid species (Lander and Botstein, 1989; Zeng, 1994; Jiang and Zeng, 1995, 1997; Kao and Zeng, 1997; Kao *et al.*, 1999; Zeng *et al.*, 1999; Piepho, 2000; Wu *et al.*, 2002c,a; Malosetti *et al.*, 2004; Silva *et al.*, 2012), and more recently, for polyploid crops (Hackett *et al.*, 2001; Ming, 2001; Wu *et al.*, 2004; Pastina *et al.*, 2011; Hackett *et al.*, 2013; Margarido *et al.*, 2015; Balsalobre *et al.*, 2017; Ferreira *et al.*, 2019; Pereira *et al.*, 2020; Deo *et al.*, 2020; Cappai *et al.*, 2020). Although other mapping studies were reported for *M. maximus* (Ebina *et al.*, 2005; Deo *et al.*, 2020), there is still no consensus regarding the taxonomic placement of the species, and little is known about its evolutionary and genomic behaviors, as well as the molecular pathways that drive phenotypic expression for relevant traits.

The present work aims to: (a) implement the state-of-the-art methods for linkage mapping in autopolyploid populations proposed by Mollinari and Garcia (2019) in a friendly and easy-to-use statistical software; (b) obtain polyploid genotypes in the form of dosage markers for a biparental mapping population of *M. maximus*; (c) construct a highly dense and informative genetic linkage map for *M. maximus* using multiple dosage markers; (d) study the genetic architecture of relevant traits for *M. maximus* through a MIM (Multiple Interval Mapping, Kao *et al.* (1999); Pereira *et al.* (2020)) QTL mapping model.

References

Balsalobre, T. W. A., G. da Silva Pereira, G. R. A. Margarido, R. Gazaffi, F. Z. Barreto, C. O. Anoni, C. B. Cardoso-Silva, E. A. Costa, M. C. Mancini, H. P. Hoffmann, A. P. de Souza, A. A. F. Garcia, and M. S. Carneiro, 2017 Gbs-based single dosage markers for linkage and qtl mapping allow gene mining for yield-related traits in sugarcane. BMC Genomics **18**: 72.

Bourke, P. M., G. van Geest, R. E. Voorrips, J. Jansen, T. Kranenburg, A. Shahin, R. G. F. Visser, P. Arens, M. J. M. Smulders, and C. Maliepaard, 2018b Polymapr-linkage analysis and genetic map construction from f1 populations of outcrossing polyploids. Bioinformatics **34**: 3496–3502.

Cappai, F., R. R. Amadeu, J. Benevenuto, R. Cullen, A. Garcia, A. Grossman, L. F. V. Ferrão, and P. Munoz, 2020 High-resolution linkage map and qtl analyses of fruit firmness in autotetraploid blueberry. Frontiers in Plant Science **11**: 1767.

Deo, T. G., R. C. U. Ferreira, L. A. C. Lara, A. C. L. Moraes, A. Alves-Pereira, F. A. de Oliveira, A. A. F. Garcia, M. F. Santos, L. Jank, and A. P. de Souza, 2020 High-resolution linkage map with allele dosage allows the identification of regions governing complex traits and apospory in guinea grass (megathyrsus maximus). Frontiers in Plant Science **11**: 15.

do Valle, C. B., L. Jank, and R. M. S. Resende, 2009 O melhoramento de forrageiras tropicais no brasil. Revista Ceres **56**: 460–472.

Ebina, M., H. Nakagawa, T. Yamamoto, H. Araya, S. ichi Tsuruta, M. Takahara, and K. Nakajima, 2005 Co-segregation of aflp and rapd markers to apospory in guineagrass (panicum maximum jacq.). Grassland Science **51**: 71–78.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell, 2011 A robust, simple genotyping-by-sequencing (gbs) approach for high diversity species. PLoS ONE **6**: e19379.

Ferreira, R. C. U., L. A. de Castro Lara, L. Chiari, S. C. L. Barrios, C. B. do Valle, J. R. Valério, F. Z. V. Torres, A. A. F. Garcia, and A. P. de Souza, 2019 Genetic mapping with allele dosage information in tetraploid urochloa decumbens (stapf) r. d. webster reveals insights into spittlebug (notozulia entreriana berg) resistance. Frontiers in Plant Science **10**: 92.

Gallais, A., 2003 *Quantitative genetics and breeding methods in autopolyploid plants*. Institut national de la recherche agronomique, Paris.

Garcia, A. A. F., M. Mollinari, T. G. Marconi, O. R. Serang, R. R. Silva, M. L. C. Vieira, R. Vicentini, E. A. Costa, M. C. Mancini, M. O. S. Garcia, M. M. Pastina, R. Gazaffi, E. R. F. Martins, N. Dahmer, D. A. Sforça, C. B. C. Silva, P. Bundock, R. J. Henry, G. M. Souza, M.-A. van Sluys, M. G. A. Landell, M. S. Carneiro, M. A. G. Vincentz, L. R. Pinto, R. Vencovsky, and A. P. Souza, 2013 Snp genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. Scientific Reports **3**: 3399.

Gerard, D., L. F. V. Ferrão, A. A. F. Garcia, and M. Stephens, 2018 Genotyping polyploids from messy sequencing data. Genetics **210**: 789–807.

Hackett, C. A., J. E. Bradshaw, and G. J. Bryan, 2014 Qtl mapping in autotetraploids using snp dosage information. Theoretical and Applied Genetics **127**: 1885–1904.

Hackett, C. A., J. E. Bradshaw, and J. W. McNicol, 2001 Interval mapping of quantitative trait loci in autotetraploid species. Genetics **159**: 1819–1832.

Hackett, C. A., K. McLean, and G. J. Bryan, 2013 Linkage analysis and qtl mapping using snp dosage data in a tetraploid potato mapping population. PLoS ONE **8**: e63939.

Jaganathan, D., A. Bohra, M. Thudi, and R. K. Varshney, 2020 Fine mapping and gene cloning in the post-ngs era: Advances and prospects. Theoretical and Applied Genetics **133**: 1791–1810.

Jank, L., S. C. Barrios, C. B. do Valle, R. M. Simeão, and G. F. Alves, 2014 The value of improved pastures to brazilian beef production. Crop and Pasture Science **65**: 1132.

Jiang, C. and Z. B. Zeng, 1995 Multiple trait analysis of genetic mapping for quantitative trait loci. Genetics **140**: 1111–1127.

Jiang, C. and Z.-B. Zeng, 1997 Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. Genetica **101**: 47–58.

Kao, C.-H. and Z.-B. Zeng, 1997 General formulas for obtaining the mles and the asymptotic variance-covariance matrix in mapping quantitative trait loci when using the em algorithm. Biometrics **53**: 653.

Kao, C.-H., Z.-B. Zeng, and R. D. Teasdale, 1999 Multiple interval mapping for quantitative trait loci. Genetics **152**: 1203–1216.

Lander, E. S. and D. Botstein, 1989 Mapping mendelian factors underlying quantitative traits using rflp linkage maps. Genetics **121**: 185–199.

Lara, L. A., M. F. Santos, L. Jank, L. Chiari, M. de M Vilela, R. R. Amadeu, J. P. R. dos Santos, G. da S Pereira, Z.-B. Zeng, and A. A. F. Garcia, 2019 Genomic selection with allele dosage in panicum maximum jacq. G3 Genes|Genomes|Genetics **9**: 2463–2475.

Malosetti, M., J. Voltas, I. Romagosa, S. Ullrich, and F. van Eeuwijk, 2004 Mixed models including environmental covariables for studying qtl by environment interaction. Euphytica **137**: 139–145.

Margarido, G. R. A., M. M. Pastina, A. P. Souza, and A. A. F. Garcia, 2015 Multi-trait multi-environment quantitative trait loci mapping for a sugarcane commercial cross provides insights on the inheritance of important traits. Molecular Breeding **35**: 175.

Martuscello, J. A., L. Jank, D. M. da Fonseca, C. D. Cruz, and D. de Noronha Figueiredo Vieira da Cunha, 2007 Repetibilidade de caracteres agronômicos em panicum maximum jacq. Revista Brasileira de Zootecnia **36**: 1975–1981.

Ming, R., 2001 Qtl analysis in a complex autopolyploid: Genetic control of sugar content in sugarcane. Genome Research **11**: 2075–2084.

Mollinari, M. and A. A. F. Garcia, 2019 Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden markov models. G3 Genes|Genomes|Genetics **9**: 3297–3314.

Mollinari, M., B. A. Olukolu, G. da S. Pereira, A. Khan, D. Gemenet, G. C. Yencho, and Z.-B. Zeng, 2019 Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. G3 Genes|Genomes|Genetics **10**: 281–292.

Pastina, M. M., M. Malosetti, R. Gazaffi, M. Mollinari, G. R. A. Margarido, K. M. Oliveira, L. R. Pinto, A. P. Souza, F. A. van Eeuwijk, and A. A. F. Garcia, 2011 A mixed model qtl analysis for sugarcane multiple-harvest-location trial data. Theoretical and Applied Genetics **124**: 835–849.

Pereira, G. S., A. A. F. Garcia, and G. R. A. Margarido, 2018 A fully automated pipeline for quantitative genotype calling from next generation sequencing data in autopolyploids. BMC Bioinformatics **19**: 398.

Pereira, G. S., D. C. Gemenet, M. Mollinari, B. A. Olukolu, J. C. Wood, F. Diaz, V. Mosquera, W. J. Gruneberg, A. Khan, C. R. Buell, G. C. Yencho, and Z.-B. Zeng, 2020 Multiple qtl mapping in autopolyploids: a random-effect model approach with application in a hexaploid sweetpotato full-sib population. Genetics **215**: 579–595.

Piepho, H.-P., 2000 A mixed-model approach to mapping quantitative trait loci in barley on the basis of multiple environment data. Genetics **156**: 2043–2050.

Resende, R. M. S., M. D. Casler, and M. D. V. de Resende, 2014 Genomic selection in forage breeding: Accuracy and methods. Crop Science **54**: 143–156.

Savidan, Y. H., L. Jank, J. C. G. Costa, and C. B. do Valle, 1989 Breeding panicum maximum in brazil. 1. genetic resources, modes of reproduction and breeding procedures. Euphytica **41**: 107–112.

Serang, O., M. Mollinari, and A. A. F. Garcia, 2012 Efficient exact maximum a posteriori computation for bayesian snp genotyping in polyploids. PLoS ONE **7**: e30906.

Silva, L. C. E., S. Wang, and Z.-B. Zeng, 2012 Multiple trait multiple interval mapping of quantitative trait loci from inbred line crosses. BMC Genetics **13**: 67.

Voorrips, R. E., G. Gort, and B. Vosman, 2011 Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics **12**: 172.

Wu, R., C.-X. Ma, and G. Casella, 2002a Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. Genetics **160**: 779–792.

Wu, R., C.-X. Ma, and G. Casella, 2004 A bivalent polyploid model for mapping quantitative trait loci in outcrossing tetraploids. Genetics **166**: 581–595.

Wu, R., C.-X. Ma, S. S. Wu, and Z.-B. Zeng, 2002c Linkage mapping of sex-specific differences. Genetical Research **79**: 85–96.

Zeng, Z. B., 1994 Precision mapping of quantitative trait loci. Genetics **136**: 1457–1468.

Zeng, Z.-B., C.-H. Kao, and C. J. Basten, 1999 Estimating the genetic architecture of quantitative traits. Genetical Research **74**: 279–289.

2 MAPPOLY: BUILDING MULTILOCUS GENETIC LINKAGE MAPS IN AUTOPOLYPLOIDS

2.1 Introduction

Polyploidy – a condition where any species presents more than two sets of chromosomes – is widespread across many organism kingdoms, including Bacteria, Fungi, Plantae, and Animalia. This condition is even more frequent in plants, whose incidence in cultivated species is around 30% (Salman-Minkov *et al.*, 2016) including economically important crops such as wheat, cotton, sugarcane, vegetables (e.g., potato and sweet-potato), forages, and trees. Approximately 50-70% of of angiosperms have undergone polyploidy during their evolutionary process (Chen *et al.*, 2006). Besides playing an important role in the evolutionary behavior of many species, polyploidy may also bring other consequences, such as auto-incompatibility, irregular meiosis, high levels of inbreeding depression, and highly heterozygotic genomes. As a consequence of its high genomic complexity, there is a delay in the development of methods and tools to study polyploid organisms, especially the autopolyploids (Garcia *et al.*, 2013; Mollinari *et al.*, 2019).

Genetic maps are important tools for a variety of genetic endeavors, including evolutionary studies, genome assembly, QTL mapping, the study of inheritance, and haplotype transmission. The construction of genetic maps in diploids is well established, with several software available (e.g., Lander *et al.* (1987); Stam (1993); Margarido *et al.* (2007)), whereas it is a challenging task for autopolyploids. For instance, considering an autotetraploid species under random bivalent meiosis, an individual can produce six different gametes containing two homologs each (Hackett, 2001). When combining two gametes, the number of possible genotypes is 36 (Figure 2.1). For hexaploids and octaploids, the numbers of possible genotypes are 400 and 4900, respectively. This massive number of genotypes poses a severe hindrance in the detection of recombination events across generations in these species.

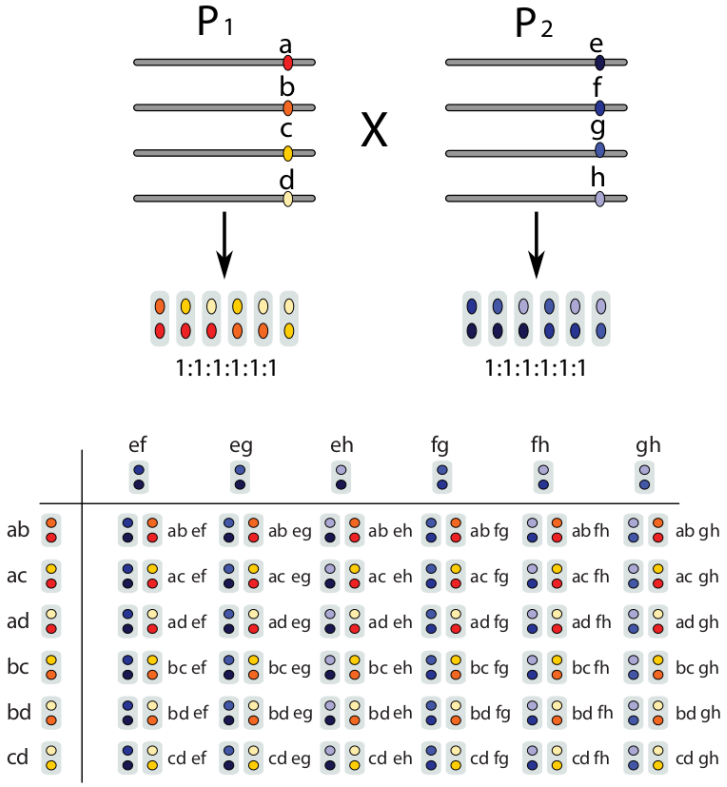

Figure 2.1. Representation of all possible genotypes in an autotetraploid biparental cross.

Considering any biparental cross, a fully informative (multiallelic) marker would allow the distinction between all homologs, facilitating the estimation of phase configurations and recombination frequencies. However, modern molecular genotyping techniques are usually based on sequencing, such as the GBS (genotyping-by-sequencing), where the most common form of variation is the SNP (single nucleotide polymorphism, Brumfield *et al.* (2003)). The information provided by such marker is remarkably low, since it allows the detection of only two different alleles. In a diploid biparental cross, there are up to four different alleles, two in each parent; for an autotetraploid species, there are up to eight different alleles, four in each parent, and so on. While these markers are almost fully informative for diploid species, polyploid species are restricted to the few genotypic classes provided by the dosage combinations of two alleles, without phase information and frequently associated with significant error rates (Voorrips *et al.*, 2011; Gerard *et al.*, 2018). Also, the loss of information increases proportionally to the ploidy level under this scenario (Mollinari and Garcia, 2019). Considered this, methods for building genetic linkage maps in polyploids are expected to be able to assess the multiallelic information using biallelic information. In other words, linkage analysis is capable of estimating the correct phase configurations and recombination frequencies between all markers in a given genome fragment, and traceback all haplotypes to the population founders, independent of markers' informativeness.

Approximation methods were used as attempts to overcome the complexity of polyploid organisms. One of the most common methods is the application of diploid-based mapping procedures using simplex markers only, which present the same segregation of diploid species (Wu *et al.*, 1992). More recently, expanded multi-dosage methods were proposed and implemented to build genetic linkage maps for some polyploids, such as autotetraploids (Hackett *et al.*, 2013), triploids and hexaploids (Bourke *et al.*, 2018b). While the methods for autotetraploids work well, for higher ploidy levels they are extremely dependent on simplex markers and still restricted to hexaploid species. Considering the information loss when using such procedures, a considerable proportion of wrong estimated phase configurations is expected, which may impact the estimation of recombination frequencies and genotypic probabilities. In addition, the lack of a multipoint approach is crucial for polyploid species, thus jointly estimating map parameters is the best way to increase informativeness and assess their multiallelic nature.

Trying to overcome these challenges, Mollinari and Garcia (2019) proposed a general method to build genetic linkage maps in autopolyploids with any even ploidy level. The procedure considers all information available on molecular markers, including genotypic errors, and relies on a multipoint approach to reestimate distances and account for genotypic probabilities. Here we present the first release of *MAPpoly*, a software developed to construct genetic linkage maps in autopolyploid species based on the Mollinari and Garcia (2019) procedure. *MAPpoly* can handle different datasets, including the most common VCF format, and is capable of building genetic linkage maps considering ploidy levels from 2 to 12. The software can also use parallelization to achieve high computational efficiency, reorder submaps to fix wrong ordering and phasing, estimate parental and individual haplotypes, and detect preferential pairing during meiosis.

References

Bourke, P. M., G. van Geest, R. E. Voorrips, J. Jansen, T. Kranenburg, A. Shahin, R. G. F. Visser, P. Arens, M. J. M. Smulders, and C. Maliepaard, 2018b Polymapr-linkage analysis and genetic map construction from f1 populations of outcrossing polyploids. Bioinformatics **34**: 3496–3502.

Brumfield, R. T., P. Beerli, D. A. Nickerson, and S. V. Edwards, 2003 The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology & Evolution **18**: 249–256.

Chen, L., Q. Lou, Y. Zhuang, J. Chen, X. Zhang, and J. N. Wolukau, 2006 Cytological diploidization and rapid genome changes of the newly synthesized allotetraploids cucumis × hytivus. Planta **225**: 603–614.

Garcia, A. A. F., M. Mollinari, T. G. Marconi, O. R. Serang, R. R. Silva, M. L. C. Vieira, R. Vicentini, E. A. Costa, M. C. Mancini, M. O. S. Garcia, M. M. Pastina, R. Gazaffi, E. R. F. Martins, N. Dahmer, D. A. Sforça, C. B. C. Silva, P. Bundock, R. J. Henry, G. M. Souza, M.-A. van Sluys, M. G. A. Landell, M. S. Carneiro, M. A. G. Vincentz, L. R. Pinto, R. Vencovsky, and A. P. Souza, 2013 Snp genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. Scientific Reports **3**: 3399.

Gerard, D., L. F. V. Ferrão, A. A. F. Garcia, and M. Stephens, 2018 Genotyping polyploids from messy sequencing data. Genetics **210**: 789–807.

Hackett, C. A., 2001 A comment on xie and xu: 'mapping quantitative trait loci in tetraploid species'. Genetical Research **78**: 187–189.

Hackett, C. A., K. McLean, and G. J. Bryan, 2013 Linkage analysis and qtl mapping using snp dosage data in a tetraploid potato mapping population. PLoS ONE **8**: e63939.

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly, S. E. Lincoln, and L. Newburg, 1987 Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1**: 174–181.

Margarido, G. R. A., A. P. Souza, and A. A. F. Garcia, 2007 Onemap: Software for genetic mapping in outcrossing species. Hereditas **144**: 78–79.

Mollinari, M. and A. A. F. Garcia, 2019 Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden markov models. G3 Genes|Genomes|Genetics **9**: 3297–3314.

Mollinari, M., B. A. Olukolu, G. da S. Pereira, A. Khan, D. Gemenet, G. C. Yencho, and Z.-B. Zeng, 2019 Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. G3 Genes|Genomes|Genetics **10**: 281–292.

Salman-Minkov, A., N. Sabath, and I. Mayrose, 2016 Whole-genome duplication as a key factor in crop domestication. Nature Plants **2**: 16115.

Stam, P., 1993 Construction of integrated genetic linkage maps by means of a new computer package: Joinmap. The Plant Journal **3**: 739–744.

Voorrips, R. E., G. Gort, and B. Vosman, 2011 Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics **12**: 172.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore, and S. D. Tanksley, 1992 The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theoretical and Applied Genetics **83**: 294–300.

3 GENETIC LINKAGE AND QTL MAPPING IN *MEGATHYRSUS MAXIMUS* (JACQ.) WITH MULTIPLE DOSAGE MARKERS

3.1 Introduction

Megathyrsus maximus (Jacq.) B.K. Simon & S.W.L. Jacobs (Syn. *Panicum maximum* Jacq., *Urochloa maxima* (Jacq.) R. Webster), commonly known as guinea grass, is a forage crop widely used in cattle beef production due to its high yield and outstanding nutritional quality. The species originated in east Africa, but had great adaptation to different tropical and sub-tropical land areas and became widely cultivated in many countries in South America. In the Brazilian market, it is situated among the most productive grasses that propagate by seeds (Jank *et al.*, 2011). The species occurs in two natural forms: sexual diploid (2n=2x=16) and apomictic tetraploid (2n=4x=32) genotypes. Other chromosomal numbers, as well as hexaploids and aneuploids, were also reported in the literature with lower frequencies (Warmke, 1951; Jauhar, 1969; Savidan, 1980; Hamoud *et al.*, 1994; Giussani *et al.*, 2001; Jain *et al.*, 2003; Akiyama *et al.*, 2008).

Since natural tetraploid genotypes of *M. maximus* undergo apomixis (i.e., asexual propagation by seeds, Nogler (1984)), it is possible to fix superior genotypes and their hybrid vigor while maintaining uniform pastures by using genetically identical seeds (Jank *et al.*, 2011). Most of the *M. maximus* breeding programs have been taking advantage of sexual and apomictic genotypes by combining them in crossing schemes. Thus, sexual genotypes are used to allow recombinations throughout the crosses, while apomixis is used to fix the best genotypes and produce seeds on a large scale through asexual propagation. Sexual and tetraploid genotypes were initially diploids which had their chromosomes duplicated with colchicine to allow viable crossings (Nakagawa and Hanna, 1992; NAkagawa *et al.*, 1993; Savidan, 1980). Previous studies have shown that progenies derived from tetraploid sexual vs. apomictic crosses segregate in a 1:1 rate for apomixis (Savidan, 1978, 1981; Ebina *et al.*, 2005; Bluma-Marques *et al.*, 2014; Deo *et al.*, 2020), though other studies suggest a quantitative genetic control for this trait (Kaushal *et al.*, 2008, 2019; Marcón *et al.*, 2019).

The advancements in molecular technology enabled the detection of variants in DNA and RNA sequences, which allowed the identification of sources of variations on both molecular and phenotypic bases. This knowledge has been supporting studies on DNA recombination, molecular paths, and its interactions, helping to understand the mechanisms that drive phenotypic expression, organism differentiation, and speciation (Metzker, 2009; Elshire *et al.*, 2011; Poland and Rife, 2012). Many of these investigations are based upon analyzes on the genetic diversity of a population, linkage and QTL mapping, genome-wide association studies, or whole-genome prediction; thus, their outcomes have the potential to significantly change the way breeding programs are planned and conducted (Poland and Rife, 2012). Therefore, methods to obtain, evaluate, and analyze molecular datasets are widespread and well developed, especially for diploid species. However, there has been a delay in the development and extension of such technology and methods for polyploid species, mainly due to its genomic complexity and lack of resources (Garcia *et al.*, 2013).

Single nucleotide polymorphism (SNP) are the most abundant form of variation in the genome, usually consisting of biallelic markers (Brumfield *et al.*, 2003). The evaluation of such variants along the genome also allows to access the allele abundance and the estimation of the genotypes in a polyploid species (Garcia *et al.*, 2013; Serang *et al.*, 2012; Gerard *et al.*, 2018; Voorrips *et al.*, 2011; Hackett *et al.*, 2013).

Thus, individual genotypes can be represented with different dosages ranging from zero up to the ploidy level of the species. The dosage value usually represents the estimated count of the reference allele that an individual carries for a given biallelic locus. As an example, an autotetraploid species may present

individual dosages ranging from 0 to 4, which would represent the genotypes aaaa, aaaA, aaAA, aAAA, and AAAA for a biallelic marker (Lara *et al.*, 2019). Despite providing more information than single dosage markers (i.e., only nulliplex, simplex, and double-simplex genotype combinations), dosage-based genotypes still lack the complete genetic information for an individual, especially for polyploids. This complete information would include multiple allele information and their distribution across individual haplotypes and along the genome, their phase configurations with adjacent locus, and the origin and recombination events that generated each haplotype in an individual genetic set. Fortunately, state-of-the-art methods are capable of recovering the complete genetic information from the same datasets that generate dosage genotypes, usually based on genomic sequences or linkage analysis. The former takes sequence-based information to detect unique haplotype sequences and recover phase configurations or multiallelic information (Motazedi *et al.*, 2018, 2019; van Geest *et al.*, 2020; Moeinzadeh *et al.*, 2020), while the latter uses additional information from the population structure to model other genetic phenomena, such as the expected Mendelian segregation and recombination frequencies, to get individual haplotypes (Hackett *et al.*, 2013; Bourke *et al.*, 2018b; Mollinari and Garcia, 2019).

Linkage analysis has been widely used to understand genetic conformity and the inheritance pattern in targeted mapping populations. In addition to the identification of linkage groups, the recombination frequencies, physical distances, and the phase configuration between a set of genetic variants, linkage analysis allows to recover the full genetic information and study the meiotic process involved in the haplotypic inheritance for a given population (Mollinari and Garcia, 2019; Mollinari *et al.*, 2019). With the full genetic information in hand, it is also possible to search for QTL along the genome by making use of the joint genotype probabilities of all individuals (Pereira *et al.*, 2020). Similar to the development of molecular technologies, methods to build genetic linkage maps and search for QTL are well developed for diploid species, while its extension for polyploid species appeared sometime later. Only recently, autopolyploid species were benefit from the extension of methods to construct integrated genetic linkage maps based on multi-dosage information, primarily for tetraploids (Hackett *et al.*, 2013) and hexaploids (Bourke *et al.*, 2018b), then extended to take advantage of the Hidden Markov Model to get multilocus estimates for higher ploidy levels (Mollinari and Garcia, 2019). The same was observed for QTL mapping models Hackett *et al.* (2013, 2014); Chen *et al.* (2018); Pereira *et al.* (2020). Such methods are well developed and established for diploids (Wu *et al.*, 2002b,c) and as such, high-quality and dense linkage maps have been vastly reported for these species. On the other hand, a few polyploid species benefit from high-quality, dense, and integrated linkage maps, while most were constructed based on single-dosage markers (Wu *et al.*, 1992) or using diploid-based methods (Shirasawa *et al.*, 2017; Ferreira *et al.*, 2019; Balsalobre *et al.*, 2017). Thus, they lack the informativeness provided by novel sequencing technologies coupled with large populations, good reference genomes, and state-of-the-art statistical methods.

Several investigations have been conducted to study the *M. maximus* molecular behavior, including genetic diversity studies (Giussani *et al.*, 2001; Aliscioni *et al.*, 2003; González and Morton, 2005; Akiyama *et al.*, 2008; Salariato *et al.*, 2010; II, 2011; Morrone *et al.*, 2012; Hunt *et al.*, 2014; Kellogg, 2015; Burke *et al.*, 2016; Tomaszewska *et al.*, 2021), linkage and QTL mapping (Ebina *et al.*, 2005; Deo *et al.*, 2020), transcriptome and RNAseq analysis (Yamada-Akiyama *et al.*, 2009; Toledo-Silva *et al.*, 2013; Wedow *et al.*, 2019; Radhakrishna *et al.*, 2018), genomic selection (Lara *et al.*, 2019), and cytogenomics (Tomaszewska *et al.*, 2021). However, there is still no consensus regarding the taxonomic placement of the species, and little is known about its evolutionary and genomic behaviors, as well as the molecular pathways that drive phenotypic expression. Although other linkage and QTL studies were reported for *M. maximus*, they lack the density and informativeness provided by more recent statistical methods, yet few autopolyploid species have already benefited from them (Ferreira *et al.*, 2019; Mollinari *et al.*, 2019; Cappai *et al.*, 2020;

Oloka *et al.*, 2021).

Given *M. maximus* relevance to agriculture and the lack of studies involving recent statistical methods, more studies are necessary to unravel the species genomic complexity, its inheritance patterns, molecular pathways, and their relation to phenotypic expression. Thus, the objectives of this study were to: (a) detect DNA polymorphisms in a *M. maximus* mapping population; (b) construct a state-of-the-art genetic linkage map with phased parental and progeny haplotypes; (c) use a state-of-the-art method to search for QTL regions along the genome for relevant traits.

3.2 Final considerations

We were able to detect DNA variants and map them to the densest and informative linkage map of *M. maximus* up to date. We also provided phased haplotypes for all individuals in the mapping population and detected genomic regions that are significantly associated with the genetic control of important traits. The present investigation provides new insights into the genomic behavior and evolutionary pathway of *M. maximus*, producing more evidence for its evolutionary placement among other relative grasses, and providing more support to the latest taxonomic classification of the species. Thus, new speculations can be drawn over the genomic origin and relatedness between *M. maximus* and other related species, helping to solve the puzzle on the relationship between the basic chromosomal number and the evolutionary pathways between Panicoidae species. The maps can help to assemble the reference genome for the species, thus encouraging and providing better information for future studies. The detected QTL may help breeding programs to increase efficiency across selection cycles throughout marker-based selection strategies for the most relevant traits, and also provide the basis for further investigations that can help unravel the genetic behavior of quantitaive traits for *M. maximus*.

References

Akiyama, Y., H. Yamada-Akiyama, H. Yamanouchi, M. Takahara, M. Ebina, T. Takamizo, S. ichi Sugita, and H. Nakagawa, 2008 Estimation of genome size and physical mapping of ribosomal dna in diploid and tetraploid guineagrass (panicum maximum jacq.). Grassland Science **54**: 89–97.

Aliscioni, S. S., L. M. Giussani, F. O. Zuloaga, and E. A. Kellogg, 2003 A molecular phylogeny of panicum (poaceae: Paniceae): Tests of monophyly and phylogenetic placement within the panicoideae. American Journal of Botany **90**: 796–821.

Balsalobre, T. W. A., G. da Silva Pereira, G. R. A. Margarido, R. Gazaffi, F. Z. Barreto, C. O. Anoni, C. B. Cardoso-Silva, E. A. Costa, M. C. Mancini, H. P. Hoffmann, A. P. de Souza, A. A. F. Garcia, and M. S. Carneiro, 2017 Gbs-based single dosage markers for linkage and qtl mapping allow gene mining for yield-related traits in sugarcane. BMC Genomics **18**: 72.

Bluma-Marques, A. C., L. Chiari, D. C. Agnes, L. Jank, and M. S. Pagliarini, 2014 Molecular markers linked to apomixis in panicum maximum jacq. African Journal of Biotechnology **13**: 2198–2202.

Bourke, P. M., G. van Geest, R. E. Voorrips, J. Jansen, T. Kranenburg, A. Shahin, R. G. F. Visser, P. Arens, M. J. M. Smulders, and C. Maliepaard, 2018b Polymapr-linkage analysis and genetic map construction from f1 populations of outcrossing polyploids. Bioinformatics **34**: 3496–3502.

Brumfield, R. T., P. Beerli, D. A. Nickerson, and S. V. Edwards, 2003 The utility of single nucleotide polymorphisms in inferences of population history. Trends in Ecology & Evolution **18**: 249–256.

Burke, S. V., W. P. Wysocki, F. O. Zuloaga, J. M. Craine, J. C. Pires, P. P. Edger, D. Mayfield-Jones, L. G. Clark, S. A. Kelchner, and M. R. Duvall, 2016 Evolutionary relationships in panicoid grasses based on plastome phylogenomics (panicoideae; poaceae). BMC Plant Biology **16**: 140.

Cappai, F., R. R. Amadeu, J. Benevenuto, R. Cullen, A. Garcia, A. Grossman, L. F. V. Ferrão, and P. Munoz, 2020 High-resolution linkage map and qtl analyses of fruit firmness in autotetraploid blueberry. Frontiers in Plant Science **11**: 1767.

Chen, J., F. Zhang, L. Wang, L. Leach, and Z. Luo, 2018 Orthogonal contrast based models for quantitative genetic analysis in autotetraploid species. New Phytologist **220**: 332–346.

Deo, T. G., R. C. U. Ferreira, L. A. C. Lara, A. C. L. Moraes, A. Alves-Pereira, F. A. de Oliveira, A. A. F. Garcia, M. F. Santos, L. Jank, and A. P. de Souza, 2020 High-resolution linkage map with allele dosage allows the identification of regions governing complex traits and apospory in guinea grass (megathyrsus maximus). Frontiers in Plant Science **11**: 15.

Ebina, M., H. Nakagawa, T. Yamamoto, H. Araya, S. ichi Tsuruta, M. Takahara, and K. Nakajima, 2005 Co-segregation of aflp and rapd markers to apospory in guineagrass (panicum maximum jacq.). Grassland Science **51**: 71–78.

Elshire, R. J., J. C. Glaubitz, Q. Sun, J. A. Poland, K. Kawamoto, E. S. Buckler, and S. E. Mitchell, 2011 A robust, simple genotyping-by-sequencing (gbs) approach for high diversity species. PLoS ONE **6**: e19379.

Ferreira, R. C. U., L. A. de Castro Lara, L. Chiari, S. C. L. Barrios, C. B. do Valle, J. R. Valério, F. Z. V. Torres, A. A. F. Garcia, and A. P. de Souza, 2019 Genetic mapping with allele dosage information in tetraploid urochloa decumbens (stapf) r. d. webster reveals insights into spittlebug (notozulia entreriana berg) resistance. Frontiers in Plant Science **10**: 92.

Garcia, A. A. F., M. Mollinari, T. G. Marconi, O. R. Serang, R. R. Silva, M. L. C. Vieira, R. Vicentini, E. A. Costa, M. C. Mancini, M. O. S. Garcia, M. M. Pastina, R. Gazaffi, E. R. F. Martins, N. Dahmer, D. A. Sforça, C. B. C. Silva, P. Bundock, R. J. Henry, G. M. Souza, M.-A. van Sluys, M. G. A. Landell, M. S. Carneiro, M. A. G. Vincentz, L. R. Pinto, R. Vencovsky, and A. P. Souza, 2013 Snp genotyping allows an in-depth characterisation of the genome of sugarcane and other complex autopolyploids. Scientific Reports **3**: 3399.

Gerard, D., L. F. V. Ferrão, A. A. F. Garcia, and M. Stephens, 2018 Genotyping polyploids from messy sequencing data. Genetics **210**: 789–807.

Giussani, L. M., J. H. Cota-Sánchez, F. O. Zuloaga, and E. A. Kellogg, 2001 A molecular phylogeny of the grass subfamily panicoideae (poaceae) shows multiple origins of c4 photosynthesis. American Journal of Botany **88**: 1993–2012.

González, A. T. and C. Morton, 2005 Molecular and morphological phylogenetic analysis of brachiaria and urochloa (poaceae). Molecular Phylogenetics and Evolution **37**: 36–44.

Hackett, C. A., J. E. Bradshaw, and G. J. Bryan, 2014 Qtl mapping in autotetraploids using snp dosage information. Theoretical and Applied Genetics **127**: 1885–1904.

Hackett, C. A., K. McLean, and G. J. Bryan, 2013 Linkage analysis and qtl mapping using snp dosage data in a tetraploid potato mapping population. PLoS ONE **8**: e63939.

Hamoud, M. A., S. A. Haroun, R. D. MacLeod, and A. J. Richards, 1994 Cytological relationships of selected species ofpanicum l. Biologia plantarum **36**: 37.

Hunt, H. V., F. Badakshi, O. Romanova, C. J. Howe, M. K. Jones, and J. S. P. Heslop-Harrison, 2014 Reticulate evolution in panicum (poaceae): the origin of tetraploid broomcorn millet, p. miliaceum. Journal of Experimental Botany **65**: 3165–3175.

II, G. P. W. G., 2011 New grass phylogeny resolves deep evolutionary relationships and discovers c4 origins. New Phytologist **193**: 304–312.

Jain, A., S. N. Zadoo, A. K. Roy, P. Kaushal, and D. R. Malaviya, 2003 Meiotic system and probable basic chromosome number of panicum maximum jacq. accessions. Cytologia **68**: 7–13.

Jank, L., C. Valle, and R. Resende, 2011 Breeding tropical forages. Crop Breeding and Applied Biotechnology **11**: 27–34.

Jauhar, P. P., 1969 Aneuploid alteration of chromosome number in the panicum maximum complex. Indian Journal of Genetics and Plant Breeding **29**: 342–347.

Kaushal, P., K. K. Dwivedi, A. Radhakrishna, M. K. Srivastava, V. Kumar, A. K. Roy, and D. R. Malaviya, 2019 Partitioning apomixis components to understand and utilize gametophytic apomixis. Frontiers in Plant Science **10**: 256.

Kaushal, P., D. R. Malaviya, A. K. Roy, S. Pathak, A. Agrawal, A. Khare, and S. A. Siddiqui, 2008 Reproductive pathways of seed development in apomictic guinea grass (panicum maximum jacq.) reveal uncoupling of apomixis components. Euphytica **164**: 81–92.

Kellogg, E., 2015 *Flowering Plants. Monocots Poaceae*. Springer International Publishing, Cham.

Lara, L. A., M. F. Santos, L. Jank, L. Chiari, M. de M Vilela, R. R. Amadeu, J. P. R. dos Santos, G. da S Pereira, Z.-B. Zeng, and A. A. F. Garcia, 2019 Genomic selection with allele dosage in panicum maximum jacq. G3 Genes|Genomes|Genetics **9**: 2463–2475.

Marcón, F., E. J. Martínez, G. R. Rodríguez, A. L. Zilli, E. A. Brugnoli, and C. A. Acuña, 2019 Genetic distance and the relationship with heterosis and reproductive behavior in tetraploid bahiagrass hybrids. Molecular Breeding **39**: 89.

Metzker, M. L., 2009 Sequencing technologies - the next generation. Nature Reviews Genetics **11**: 31–46.

Moeinzadeh, M.-H., J. Yang, E. Muzychenko, G. Gallone, D. Heller, K. Reinert, S. Haas, and M. Vingron, 2020 Ranbow: a fast and accurate method for polyploid haplotype reconstruction. PLOS Computational Biology **16**: e1007843.

Mollinari, M. and A. A. F. Garcia, 2019 Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden markov models. G3 Genes|Genomes|Genetics **9**: 3297–3314.

Mollinari, M., B. A. Olukolu, G. da S. Pereira, A. Khan, D. Gemenet, G. C. Yencho, and Z.-B. Zeng, 2019 Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. G3 Genes|Genomes|Genetics **10**: 281–292.

Morrone, O., L. Aagesen, M. A. Scataglini, D. L. Salariato, S. S. Denham, M. A. Chemisquy, S. M. Sede, L. M. Giussani, E. A. Kellogg, and F. O. Zuloaga, 2012 Phylogeny of the paniceae (poaceae: Panicoideae): Integrating plastid dna sequences and morphology into a new classification. Cladistics **28**: 333–356.

Motazedi, E., D. de Ridder, R. Finkers, S. Baldwin, S. Thomson, K. Monaghan, and C. Maliepaard, 2018 Tripoly: Haplotype estimation for polyploids using sequencing data of related individuals. Bioinformatics **34**: 3864–3872.

Motazedi, E., C. Maliepaard, R. Finkers, R. Visser, and D. de Ridder, 2019 Family-based haplotype estimation and allele dosage correction for polyploids using short sequence reads. Frontiers in Genetics **10**: 335.

Nakagawa, H. and W. W. Hanna, 1992 Induced sexual tetraploids for breeding guineagrass (panicum maximum jacq.). Japanese Journal of Grassland Science **38**: 152–159.

NAkagawa, H., N. Shimizu, and W. W. Hanna, 1993 Cytology of" natsukaze" guineagrass, a natural apomictic hybrid between a sexual and an apomictic plant. Japanese Journal of Grassland Science **39**: 374–380.

Nogler, G. A., 1984 pp. 475–518 in *Gametophytic Apomixis*, Springer Berlin Heidelberg.

Oloka, B. M., G. da Silva Pereira, V. A. Amankwaah, M. Mollinari, K. V. Pecota, B. Yada, B. A. Olukolu, Z.-B. Zeng, and G. C. Yencho, 2021 Discovery of a major qtl for root-knot nematode (meloidogyne incognita) resistance in cultivated sweetpotato (ipomoea batatas). Theoretical and Applied Genetics **134**: 1945–1955.

Pereira, G. S., D. C. Gemenet, M. Mollinari, B. A. Olukolu, J. C. Wood, F. Diaz, V. Mosquera, W. J. Gruneberg, A. Khan, C. R. Buell, G. C. Yencho, and Z.-B. Zeng, 2020 Multiple qtl mapping in autopolyploids: a random-effect model approach with application in a hexaploid sweetpotato full-sib population. Genetics **215**: 579–595.

Poland, J. A. and T. W. Rife, 2012 Genotyping-by-sequencing for plant breeding and genetics. The Plant Genome **5**: 92–102.

Radhakrishna, A., K. K. Dwivedi, M. K. Srivastava, A. Roy, D. Malaviya, and P. Kaushal, 2018 Transcriptomic data of pre-meiotic stage of floret development in apomictic and sexual types of guinea grass (panicum maximum jacq.). Data in Brief **18**: 590–593.

Salariato, D. L., F. O. Zuloaga, L. M. Giussani, and O. Morrone, 2010 Molecular phylogeny of the subtribe melinidinae (poaceae: Panicoideae: Paniceae) and evolutionary trends in the homogenization of inflorescences. Molecular Phylogenetics and Evolution **56**: 355–369.

Savidan, Y., 1978 Genetic control of facultative apomixis and application in breeding panicum maximum. Communication to the XIVth International Congress of Genetics .

Savidan, Y., 1980 Chromosomal and embryological analyses in sexual x apomictic hybrids ofpanicum maximum jacq. Theoretical and Applied Genetics **58**: 153–156.

Savidan, Y., 1981 Genetics and utilization of apomixis for the improvement of guineagrass (panicum maximum jacq.). Proc. XIVth Int. Grassl. Congr **1981**: 182–184.

Serang, O., M. Mollinari, and A. A. F. Garcia, 2012 Efficient exact maximum a posteriori computation for bayesian snp genotyping in polyploids. PLoS ONE **7**: e30906.

Shirasawa, K., M. Tanaka, Y. Takahata, D. Ma, Q. Cao, Q. Liu, H. Zhai, S.-S. Kwak, J. C. Jeong, U.-H. Yoon, H.-U. Lee, H. Hirakawa, and S. Isobe, 2017 A high-density snp genetic map consisting of a complete set of homologous groups in autohexaploid sweetpotato (ipomoea batatas). Scientific Reports **7**: 44207.

Toledo-Silva, G., C. B. Cardoso-Silva, L. Jank, and A. P. Souza, 2013 De novo transcriptome assembly for the tropical grass panicum maximum jacq. PLoS ONE **8**: e70781.

Tomaszewska, P., M. S. Vorontsova, S. A. Renvoize, S. Z. Ficinski, J. Tohme, T. Schwarzacher, V. Castiblanco, J. J. de Vega, R. A. C. Mitchell, and J. S. P. Heslop-Harrison, 2021 Complex polyploid and hybrid species in an apomictic and sexual tropical forage grass group: genomic composition and evolution in urochloa (brachiaria) species. bioRxiv .

van Geest, G., J. Willemsen, K. R. Gujjula, P. Siddavatam, C. R. Willis, R. Ramadhar, R. E. Voorrips, and P. Arens, 2020 Micro-haplotyping in polyploids using massively parallel amplicon sequencing. Research Square pp. 2693–5015.

Voorrips, R. E., G. Gort, and B. Vosman, 2011 Genotype calling in tetraploid species from bi-allelic marker data using mixture models. BMC Bioinformatics **12**: 172.

Warmke, H. E., 1951 Cytotaxonomic investigations of some varieties of panicum maximum and of p. purpurascens in puerto rico 1. Agronomy Journal **43**: 143–149.

Wedow, J. M., C. R. Yendrek, T. R. Mello, S. Creste, C. A. Martinez, and E. A. Ainsworth, 2019 Metabolite and transcript profiling of guinea grass (panicum maximum jacq) response to elevated [co2] and temperature. Metabolomics **15**: 51.

Wu, K. K., W. Burnquist, M. E. Sorrells, T. L. Tew, P. H. Moore, and S. D. Tanksley, 1992 The detection and estimation of linkage in polyploids using single-dose restriction fragments. Theoretical and Applied Genetics **83**: 294–300.

Wu, R., C.-X. Ma, I. Painter, and Z.-B. Zeng, 2002b Simultaneous maximum likelihood estimation of linkage and linkage phases in outcrossing species. Theoretical Population Biology **61**: 349–363.

Wu, R., C.-X. Ma, S. S. Wu, and Z.-B. Zeng, 2002c Linkage mapping of sex-specific differences. Genetical Research **79**: 85–96.

Yamada-Akiyama, H., Y. Akiyama, M. Ebina, Q. Xu, S. ichi Tsuruta, J. Yazaki, N. Kishimoto, S. Kikuchi, M. Takahara, T. Takamizo, S. ichi Sugita, and H. Nakagawa, 2009 Analysis of expressed sequence tags in apomictic guineagrass (panicum maximum). Journal of Plant Physiology **166**: 750–761.

4 FINAL CONSIDERATIONS

The present work presents the software implementation of MAPpoly, the best freely available tool for constructing integrated genetic linkage maps using multilocus information in autopolyploid species up to date. We also use it to construct the densest and informative linkage map for the autotetraploid *M. maximus* forage crop. The new genetic linkage map provides insights on the species evolutionary placement and taxonomic classification, as well as knowledge about its genomic behavior and genetic conformity. We associate the linkage map with phenotypic observations and state-of-the-art methods to identify QTL regions associated with important traits for the species. The QTL identified in this study can be used to implement and enhance marker-based selection strategies in *M. maximus* breeding programs, thus supporting the increase on the efficiency of selection cycles for the most important traits. The results of this work can be used in further investigations, such as the assembly of the species genome, QTL refinement through fine-mapping, QTL-by-environment and gene expression studies.