



**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA**

THAMIRIS GATTI DEO

Mapeamento da região da apomixia e de QTLs em *Megathyrsus maximus*: uma gramínea forrageira tropical autotetraploide

Mapping of the apomixis region and QTLs in *Megathyrsus maximus*: an autotetraploid tropical forage grass

**CAMPINAS
2019**

THAMIRIS GATTI DEO

Mapeamento da região da apomixia e de QTLs em *Megathyrsus maximus*: uma gramínea forrageira tropical autotetraploide

Mapping of the apomixis region and QTLs in *Megathyrsus maximus*: an autotetraploid tropical forage grass

Tese apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutora em Genética e Biologia Molecular, na Área de Genética Vegetal e Melhoramento

Thesis presented to the Institute of Biology of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Genetics and Molecular Biology, in the area of Plant Genetics and Genetic Improvement

Orientadora: Prof.^a Anete Pereira de Souza

Co-orientadora: Dra. Liana Jank

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELA ALUNA THAMIRIS GATTI DEO E ORIENTADA PELA PROF.^A ANETE PEREIRA DE SOUZA

**CAMPINAS
2019**

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

D44m Deo, Thamiris Gatti, 1994-
Mapeamento da região da apomixia e de QTLs em *Megathyrus maximus* :
uma gramínea forrageira tropical autotetraploide / Thamiris Gatti Deo. –
Campinas, SP : [s.n.], 2019.

Orientador: Anete Pereira de Souza.
Coorientador: Liana Jank.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de
Biologia.

1. Apomixia. 2. Autopoliploide. 3. Genotipagem por sequenciamento. 4.
Capim-guine. 5. Pastagens. 6. Mapeamento de QTL. I. Souza, Anete Pereira
de, 1962-. II. Jank, Liana. III. Universidade Estadual de Campinas. Instituto de
Biologia. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Mapping of the apomixis region and QTLs in *Megathyrus
maximus* : an autotetraploid tropical forage grass

Palavras-chave em inglês:

Apomixis
Autopolyploidy
Genotyping by sequencing
Guinea grass
Pastures
QTL mapping

Área de concentração: Genética Vegetal e Melhoramento

Titulação: Doutora em Genética e Biologia Molecular

Banca examinadora:

Anete Pereira de Souza [Orientador]
Bianca Baccili Zanotto Vigna
João Ricardo Bachega Feijó Rosa
Lucimara Chiari

Maria Imaculada Zucchi

Data de defesa: 13-12-2019

Programa de Pós-Graduação: Genética e Biologia Molecular

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-4468-7070>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3506410828104666>

Campinas, 13 de dezembro de 2019

COMISSÃO EXAMINADORA

Prof.^a Dra. Anete Pereira de Souza (orientadora)

Dra. Bianca Baccili Zanotto Vigna

Dr. João Ricardo Bachega Feijó Rosa

Dra. Lucimara Chiari

Prof.^a Dra. Maria Imaculada Zucchi

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

A Ata da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia.

Dedico aos meus exemplos de vida, amor e humanidade: Alzira, Cláudia, Moacyr e Thiago.

ABSTRACT

Brazil has a prominent role in the international meat market for its production and exportation, once it has the largest commercial cattle herd in the world. Among the main forage grasses of tropical climate that are present in ruminant feed on pasture, *Megathyrsus maximus* (syn. *Panicum maximum*) is one of the most productive, being involved in the feeding of goats, horses, sheep, beef and dairy cattle. Autotetraploidy, apomixis and the recent domestication process indicate a greater complexity involved in the genetic improvement of this species, which makes it essential to clarify the genetic mechanisms that may be involved. And, in this way, the molecular improvement will be able to increase the efficiency of the breeding program and also reduce the time involved in steps on launching cultivars. In this context, the present work aimed to detect genomic regions and to establish genetic relationships between economic importance traits for *M. maximum* with the application of recent genetic-statistical tools. The mapping progeny of this work was obtained from the breeding program of Embrapa Gado de Corte (Campo Grande, Mato Grosso do Sul), being composed of 136 F₁ hybrids originated from the cross between autotetraploid parents ($2n = 4x = 32$), one sexual (S10) and the other apomictic (cv. Mombaça). The individuals were sampled in genotyping-by-sequencing (GBS) libraries constructed using two restriction enzymes (*Pst*I-*Msp*I). The data obtained with these sequencing were analysed using Tassel4-poly. Allele dosage was estimated for SNP markers identified using the SuperMASSA and Updog software, aiming to include in the map all the SNP configurations for a tetraploid. Monomorphic markers and SNPs with over 25% missing data were manually removed using the R software. For the construction of the linkage map, polypmapR and TetraploidSNPMap software were used which enabled the usage of SNPs with allele dosage information. The consensus integrated linkage map contains 8 homology groups (GH) and 858 SNP markers with all doses expected in a tetraploid progeny. The average density was about 1.13 SNPs/cM, with GH 7 presenting the largest size (108.6 cM) and the most markers (159 SNPs); only this group contained markers in all doses. Linkage analysis allowed the detection of a double reduction event, a type of segregation distortion that occurs mainly in autopolyploids. One *locus* in the apomixis region was mapped to HG II at 65 cM, using our linkage map. From similarity analysis to the markers, which are adjacent to the locus linked to apospory, using as base on the *P. virgatum* genome resulted in a gene related to somatic embryogenesis. Using interval mapping, ten QTLs for agronomic traits, such as regrowth capacity and total dry matter were detected, as well as 36 QTLs related to forage quality, such as *in vitro* digestibility of organic matter, lignin, and cellulose. A total of 55 gene families were identified by similarity analysis, highlighting putative genes involved in the regulation of plant growth hormones and lignin synthesis. This thesis provides unprecedented contributions to *M. maximum*, such as the identification of a double reduction event, a high-resolution linkage map constructed from SNPs with allele dosage information, and the mapping of apospory region, agronomic and nutritional traits. In addition to the advance on the knowledge of genetics, these results provide practical information for the breeding program of this species. They are also an incentive for further studies on apomixis and the evolution of polysomic inheritance in this and other tropical grasses.

Keywords: allele dosage; apomixis; autopoliploid; double reduction; *Panicum maximum*; pastures; QTL; yield traits; nutritional traits.

Sumário

PREFÁCIO	11
1 INTRODUÇÃO	12
2 REVISÃO BIBLIOGRÁFICA	14
2.1 Aspectos gerais das gramíneas forrageiras tropicais com ênfase em seu valor nutricional	14
2.2 Desafios da espécie <i>Megathyrus maximus</i>	18
2.3 Mapeamento genético e sua aplicação em <i>M. maximus</i>	24
3 OBJETIVOS	29
3.1 Geral	29
3.2 Específicos	29
4 ARTIGO “High-resolution Linkage Map with Allele Dosage Allows the Identification of Regions Governing Complex Traits and Apospory in Guinea Grass (<i>Megathyrus maximus</i>)”	30
5 RESULTADOS COMPLEMENTARES	70
6 RESUMO DOS RESULTADOS	88
7 CONCLUSÕES	89
8 PERSPECTIVAS	90
9 REFERENCIAS BIBLIOGRAFICAS	91
10 APÊNDICES	105
11 ANEXOS	140
11.1 Anexo I - Termo de aprovação da pesquisa pela Comissão de Bioética pertinente	141
11.2 Anexo II - Cumprimento dos dispositivos da lei nº 9610/98	142
11.3 Anexo III “Relationship between Genetic Variability and Land Use and Land Cover in Populations of <i>Campomanesia adamantium</i> (Myrtaceae)”	143
11.4 Anexo IV “Growth regulators <i>in vitro</i> multiplication of <i>Campomanesia adamantium</i> (Cambess.) O. Berg (Myrtaceae)”	144
11.5 Anexo V “Development and characterization of microsatellite markers in <i>Campomanesia adamantium</i>, a native plant of the Cerrado ecoregions of South America”	145

PREFÁCIO

Esta tese buscou avançar o conhecimento da genética sobre a gramínea forrageira *M. maximus*, desde o processo de divisão meiótica da célula até a identificação de *loci* de caracteres quantitativos, visando uma maior compreensão sobre esse genoma. Desta forma, o presente trabalho inicia com uma breve introdução e posterior aprofundamento com uma revisão bibliográfica, que revela as principais espécies utilizadas como pastagens e os métodos empregados para avaliar o potencial nutricional dessas plantas na alimentação do animal. Após essa abordagem geral, o enfoque é dado a espécie *M. maximus*, também conhecida como *P. maximum*, sua chegada ao Brasil e os desafios genéticos que ela nos apresenta, desde a criação de seu programa de melhoramento genético no país, bem como os trabalhos que foram desenvolvidos até aqui, com o objetivo de obter um maior conhecimento genético-molecular sobre a espécie. E, por fim, o uso de mapas de ligação como alternativa para superar os obstáculos genéticos em conjunto com as inovações metodológicas, no âmbito tecnológico e estatístico. Neste sentido, o Capítulo I apresenta os principais resultados obtidos com as novas tecnologias e metodologias empregadas no mapeamento genético de *M. maximus*, no formato de artigo científico, intitulado “*High-resolution Linkage Map with Allele Dosage Allows the Identification of an Apomixis Region and Complex Traits in Guinea Grass (Megathyrus maximus)*”. Além disso, os resultados preliminares e dificuldades encontradas para obtenção do mapa de ligação presente neste artigo constam em Resultados Complementares. Ao final, consta um resumo dos resultados, conclusões e perspectivas sobre este trabalho.

Em Anexos estão incluídos três artigos disponíveis na literatura, que foram desenvolvidos durante o meu período de graduação e publicados em paralelo ao projeto de Doutorado, sendo eles: “*Relationship between Genetic Variability and Land Use and Land Cover in Populations of Campomanesia adamantium (Myrtaceae)*” no período de 2018; “*Growth regulators in vitro multiplication of Campomanesia adamantium (Cambess.) O. Berg (Myrtaceae)*” e “*Development and characterization of microsatellite markers in Campomanesia adamantium, a native plant of the Cerrado ecoregions of South America*” em 2019.

Neste momento, espero que o leitor esteja instigado a continuar a leitura deste trabalho e que ao final novas ideias e inspirações sejam geradas, para que a pesquisa sobre a genômica de *M. maximus* e de gramíneas poliploides possam continuar em evolução. Obrigada!

1 INTRODUÇÃO

As gramíneas forrageiras são particularmente importantes para a produção global de carne, leite, combustível, biomassa, entre outros. O Brasil, há anos, segue dentre os principais contribuintes nestes setores do mercado (FAOSTAT, 2019). Com um rebanho de 215 milhões de cabeças de gado e 167 milhões de hectares de pastagens, o país possui o maior rebanho bovino comercial do mundo, sendo o maior exportador e segundo maior produtor de carne bovina mundial (ABIEC, 2019). As gramíneas contribuem para a qualidade e produtividade dos produtos de origem animal disponíveis no mercado. Desde seu estabelecimento no Brasil, a forrageira tropical autotetraploide *Megathyrsus maximus* (Jacq.) B. K. Simon & S.W. L. Jacobs (sinonímia *Panicum maximum* Jacq.) é destaque na dieta de ruminantes. O lançamento das primeiras cultivares dessa espécie trouxe a intensificação da produção pecuária no país, com um maior número de animais em uma área menor de pastagem. Isso é atribuído a elevada produtividade e melhor qualidade nutricional dessa forrageira (Jank et al., 2014).

Nesse contexto, há pouco menos de quarenta anos, a Embrapa Gado de Corte localizada em Campo Grande – MS conta com um programa de melhoramento genético dessa espécie, conduzido em parceria com universidades e empresas produtoras de sementes. A partir do banco de germoplasma mantido pela unidade, o programa tem proporcionado o lançamento de cultivares com diferentes atributos que contribuem para a diversificação de pastagens (Jank et al., 2011; Jank et al., 2014; Jank et al., 2017). No entanto, estudos de avaliações agronômicas sobre o potencial forrageiro de acessos do germoplasma e de híbridos existem em grande número, se comparado aos estudos sobre a genética molecular de *M. maximus*. Muitos fatores estão relacionados a isto, como a herança polissômica da espécie e o modo de reprodução por apomixia facultativa, características que exigem ferramentas capazes de superar essas lacunas.

Por isso, a construção de mapas de ligação, a partir de progênies segregantes, têm sido uma importante ferramenta como ponto inicial, para obter conhecimento genômico sobre a distribuição dos *loci* de caracteres quantitativos de espécies (Ott et al., 2015). Estes mapas contemplam estudos genômicos para as espécies em geral, com uma grande relevância em plantas não-modelo e que não possuem seu genoma sequenciado. No entanto, o mapeamento genético em poliploides não está tão avançado quanto em diploides, pois requer métodos especializados que sejam capazes de lidar com marcadores em doses mais altas. Ao contrário dos diploides, os organismos poliploides têm múltiplas conformações de genótipos

heterozigotos. Por exemplo, em um *locus* com dois alelos, um tetraploide como *M. maximus* pode ter três genótipos heterozigotos diferentes, que somados as duas conformações homozigóticas resultam em cinco genótipos diferentes possíveis, sendo eles, AAAA, AAAB, AABB, ABBB e BBBB (van Geest et al., 2017). Nesse contexto, metodologias recentes propuseram a construção de mapas de ligação que incluam os marcadores com doses mais altas e, portanto, mais informativos (Hackett et al., 2017; Bourke et al., 2018a; Mollinari et al., 2019).

Esses marcadores podem ser obtidos por meio da tecnologia de sequenciamento de nova geração (NGS, do inglês *Next Generation Sequencing*), que tem proporcionado avanços em análises genéticas de poliploides, através de métodos que fornecem a redução da complexidade desses genomas e a identificação de milhares de SNPs (*Single Nucleotide Polymorphism*). Estes marcadores são, geralmente, bi-alélicos e apresentam polimorfismo de uma única base, sendo amplamente distribuídos em genomas e presentes em praticamente todos os *loci* gênicos (Perkel, 2008).

Os mapas de ligação, contendo SNPs com informação de dosagem alélica, são fundamentais para análises de QTLs (*loci* de caracteres quantitativos) mais realísticas em relação a arquitetura genética de caracteres agrônômicos importantes. Seus resultados podem contribuir consideravelmente para os programas de melhoramento, visto que o mapeamento de QTLs associa informações da segregação dos marcadores moleculares às médias fenotípicas de cada indivíduo. Dessa forma, a posição e o número de QTLs identificados que influenciam determinado caráter, bem como os efeitos envolvidos são estimados, sendo eles, dominância ou aditividade e interações entre QTL e ambientes (Zeng, 2001; Bourke et al., 2019).

Neste trabalho, o mapeamento genético de *M. maximus* foi realizado a partir de uma progênie F₁ biparental do programa de melhoramento genético da espécie. O contraste entre os genitores e os híbridos gerados permitiu a obtenção de um mapa de ligação robusto e informativo, constituído por marcadores SNP com informação de dosagem alélica, bem como a identificação da região da apomixia e de QTLs relacionados a caracteres agrônômicos e a qualidade da forragem. Esta tese teve o intuito de auxiliar a elucidar a complexidade genética de *M. maximus* e espera-se que os resultados alcançados sejam o primeiro degrau para promover estudos mais aprofundados sobre as diferentes regiões dos QTLs encontrados.

2 REVISÃO BIBLIOGRÁFICA

2.1 Aspectos gerais das gramíneas forrageiras tropicais com ênfase em seu valor nutricional

A família Poaceae (Gramineae) abrange mais de 760 gêneros e 11.500 espécies presentes em todas as zonas climáticas (Soreng et al., 2017). Essa ampla distribuição é favorecida pelo metabolismo C3 ou C4 das gramíneas, onde as plantas C4 dominam as pastagens em climas quentes e secos, mas são substituídas pelas espécies C3 à medida que o clima esfria ou fica mais úmido (Hodkinson, 2018). Gramíneas forrageiras e de gramado tropicais, como as pertencentes aos gêneros *Cynodon*, *Panicum*, *Paspalum*, *Pennisetum* e *Urochloa*, ou plantas de clima temperado, como *Cynosurus*, *Dactylis*, *Lolium*, *Phleum* e *Poa*, são exemplos dessa família (Hodkinson, 2018), cujos centros mais importantes de diversidade genética se localizam na África Oriental, na Eurásia e, em menor grau, na América do Sul (Batello et al., 2008).

Cerca de 70 espécies representam 99% das gramíneas usadas em pastagens cultivadas no mundo e que estão sujeitas ao melhoramento genético (Batello et al., 2008). Os 162,2 milhões de hectares (ha) dedicados a áreas de pastagem no Brasil (**Figura 1**), (ABIEC, 2019) são formados, principalmente, por *Urochloa* spp., *Megathyrsus maximus* (sinonímia *Panicum maximum*), *Pennisetum purpureum*, *Cynodon* spp. e algumas espécies nativas (*Paspalum* spp.) em sistemas extensivos de pastagem, sendo responsáveis, principalmente, pela produtividade e nutrição animal (Jank et al., 2014; Silva et al., 2015). Outras aplicações têm sido exploradas, como o uso em silagem (Costa et al., 2017; Tomaz et al., 2018) e, mais recentemente, a utilização de biomassa como fonte energética para a produção de bioetanol (Mohapatra et al., 2017).

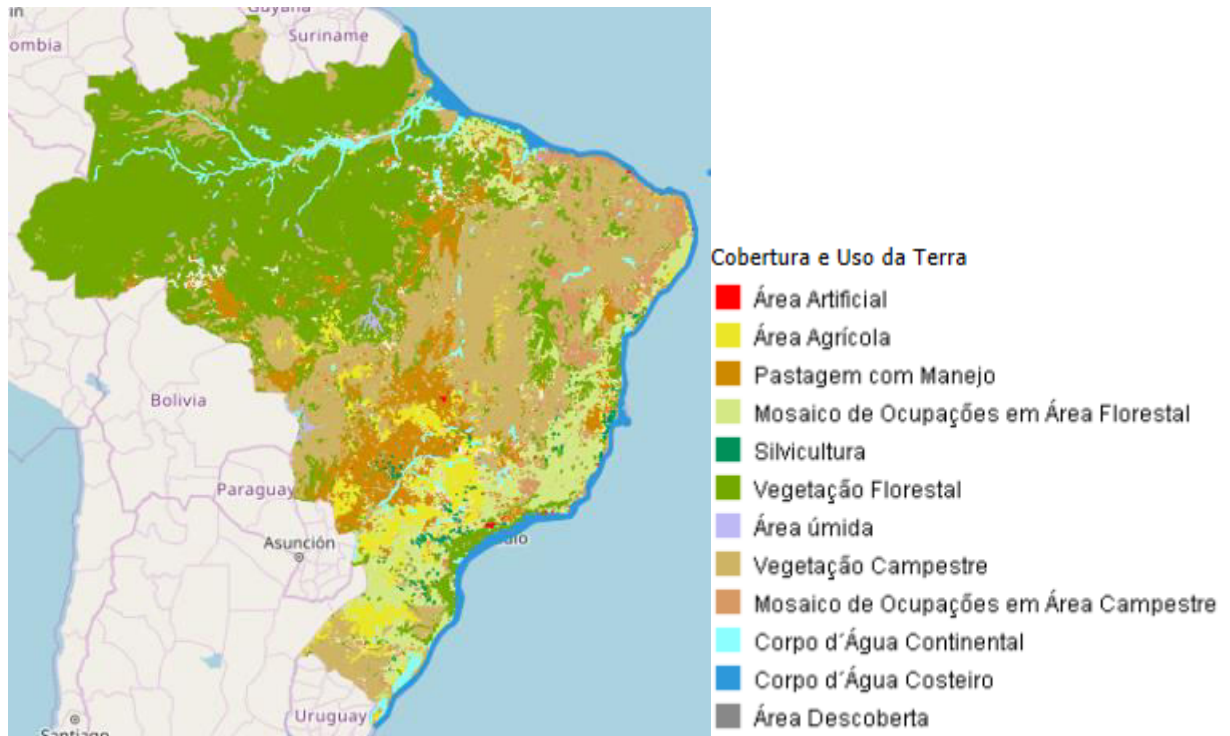


Figura 1. Áreas de uso do solo no Brasil de acordo com diferentes atividades. Fonte: IBGE, 2016.

A manutenção dessas forrageiras em diversas localidades, por todo território nacional, é também atribuída a fertilidade dos solos. A relação entre espécie, clima e solo pode representar diferença quanto aos ganhos na produtividade animal. No caso, as espécies *P. purpureum*, *M. maximus* e *Cynodon* spp. são exigentes quanto à fertilidade do solo, apresentam elevada produtividade e alto potencial de produção de biomassa (Lima et al., 2014; Moraes et al., 2018; Odorico et al., 2018). A gramínea *P. purpureum*, também conhecida como Capim-elefante, tem sido muito empregada na região Sul do país para a alimentação de bovinos leiteiros, na produção de silagem e como feno (Pereira e Cóser, 2001). Algumas cultivares desta gramínea apresentaram bons resultados em solos com baixa fertilidade no Nordeste brasileiro, desde que manejadas adequadamente (Oliveira, 2001). Enquanto, *M. maximus*, popularmente conhecida como Capim Colômbio, é cultivada em todo Brasil e em outros países da América Latina, sendo atrativa para a alimentação de gado de corte, bovinos leiteiros, equinos, caprinos e ovinos (Jank et al., 2011). Entretanto, esta espécie é mais bem adaptada a solos leves e de alta fertilidade (Pereira et al., 2001). Já, *Cynodon* spp. se destaca pela tolerância ao frio e está envolvida na alimentação de equinos e gado de leite, com a expansão de seu uso como silagem e feno. Algumas espécies deste gênero foram identificadas como nativas da América do Sul, possibilitando a exploração de plantas adaptadas às condições edafoclimáticas dos trópicos (Corrêa e Santos, 2003; Vilela, 2005).

Em contrapartida, as forrageiras *Paspalum* spp. e duas espécies do gênero *Urochloa* conhecidas, como *U. decumbens* e *U. humidicola*, são recomendadas para solos com baixa fertilidade. *Paspalum* é um gênero predominantemente das Américas, sendo que o Brasil é um de seus centros de diversidade com mais de 200 espécies nativas e cerca de 75 endêmicas (Reis et al., 2015). Esse gênero se destaca em decorrência do elevado potencial forrageiro e de sua aplicação na cobertura de solo, cujas espécies ocorrem naturalmente nos campos sulinos do país, sendo pastejadas por bovinos (Fávero, 2019). Por fim, o gênero *Urochloa* consiste de espécies conhecidas como braquiárias, que apresentam ampla adaptabilidade ao território nacional, tais como a cv. Basilisk (*U. decumbens*) que está presente em cerca de 70% das pastagens cultivadas, inclusive em solos ácidos e de baixa fertilidade. O uso de *Urochloa* spp. é mais restrito a nutrição de bovinos, principalmente gado de corte (Valle et al., 2009).

Dessa forma, cultivares têm sido desenvolvidas conforme as especificidades descritas acima para cada gênero, a fim de atender as condições muito variadas de clima e solo do país. Isso é possível por meio de programas de melhoramento genético que, de maneira geral, visam genótipos superiores de alta produtividade e de elevada qualidade nutricional, resistentes a pragas e doenças, bem como adaptados a estresses bióticos e abióticos (Valle et al., 2009; Jank et al., 2011). A produção e a qualidade das sementes também são importantes ao melhoramento, visto que a maioria dos capins comercializados são propagados por sementes, cujas cultivares com boa produção de sementes são facilmente estabelecidas no campo e têm sido adotadas pelo mercado mais rapidamente (Jank et al., 2011).

Tais gramíneas forrageiras contribuem em qualidade e na produtividade dos produtos de origem animal, como carne, leite e couro, não apresentando um valor intrínseco diretamente ligado a elas (Valle et al., 2009). O potencial produtivo e nutricional que essas plantas carregam são verificados por meio da produção de biomassa e por análises bromatológicas, nas quais, a partir da matéria seca, são aferidos o teor de carboidratos estruturais, proteína bruta, lipídeos e minerais. A matéria seca do alimento, resultado da retirada das partículas de água, é responsável pelo valor alimentar da forrageira e está intimamente ligada ao desempenho animal (Medeiros e Marino, 2015a). A fibra, constituída por lignina e carboidratos estruturais (celulose, hemicelulose e pectina), representa o componente mais abundante nessas plantas e ocupa espaço no trato digestivo do animal, podendo ser indigestível ou lentamente digestível (Mertens, 2002). A estrutura destes compostos, assim como as proporções de celulose, hemicelulose e lignina, influenciam na disponibilidade de energia do

alimento. Ensaio de fibra em detergente neutro e ácido são utilizados como produto para análises mais aprofundadas que determinam o teor dos compostos da fibra separadamente. Neste caso, pode ser promovido o teor de lignina, polímero orgânico presente na parede celular, que atua como uma barreira mecânica aos microrganismos ruminais e às hidrolases secretadas por estes, impedindo uma maior absorção de nutrientes no rúmen do animal (Medeiros e Marino, 2015a). Já, os carboidratos estruturais fornecem a maior parte da energia na dieta animal, sendo que a celulose é degradada de 25 a 90%, a hemicelulose entre 45 a 90% e a pectina é quase integralmente digerida pelo trato animal (Medeiros e Marino, 2015b). A hemicelulose pode ser verificada facilmente pela subtração do valor de fibra em detergente neutro em relação à fibra em detergente ácido (Medeiros e Marino, 2015a).

Proteínas, lipídeos e minerais estão presentes em menores proporções na forragem. Por meio da avaliação do N₂ (nitrogênio), estima-se o teor de proteína bruta, assim chamado, devido a não diferenciação entre o N presente nos aminoácidos, que corresponde a 60-70% do valor, e o não proteico, equivalente a 20-30%. O animal pode ter um aproveitamento de cerca de 40 a 80% da proteína da dieta, a qual é indispensável para os microrganismos ruminais que a convertem em amônia, sendo que uma parte inferior é metabolizada em aminoácidos e pequenos polipeptídios (Medeiros e Marino, 2015c).

Os lipídios, considerados a principal fonte de reserva de energia, são armazenados como triglicerídeos pelos ruminantes. A determinação do teor de gordura na forragem é realizada pela análise do extrato etéreo, utilizando éter etílico para solubilizar os compostos orgânicos. Destes, 50% pertence a componentes não lipídicos, como clorofila, carotenoides, saponinas e ceras relacionadas a cutícula, cuja energia retirada deles pelos microrganismos ruminais pode ser baixíssima, e a outra metade corresponde aos galactolipídeos e triglicerídeos (Medeiros et al., 2015). O capim possui um baixo valor de lipídeos presente em 3% da matéria seca, suficiente ao consumo animal. Quantidades superiores a 6% podem afetar a degradabilidade da fibra pelas bactérias celulolíticas (Medeiros et al., 2015).

Por fim, os minerais compõem a porção não orgânica da matéria seca e são determinados por meio da análise de cinzas, com a qual ocorre a combustão completa da porção orgânica da forragem. Podem ser divididos em duas categorias: macrominerais (cálcio, fósforo, sódio, potássio, cloro, magnésio e enxofre), cuja necessidade animal é maior, e em microminerais (ferro, selênio, iodo, zinco, cobre, cobalto, manganês e molibdênio), com porções menores necessárias ao organismo (Marino e Medeiros, 2015). Apesar de não

contribuírem com energia à alimentação, os minerais são essenciais em praticamente todos os processos bioquímicos do organismo, desde estrutura óssea à síntese de hormônios e, conseqüentemente, na produtividade animal (Marino e Medeiros, 2015).

Baseado em todas as informações anteriores, é notável o efeito positivo que o valor nutritivo das gramíneas forrageiras proporciona ao desempenho e ao valor do animal. No entanto, estudos genético-moleculares envolvendo caracteres nutricionais e de produção de forragem em gramíneas forrageiras tropicais eram inexistentes até a conclusão desta tese.

2.2 Desafios da espécie *Megathyrsus maximus*

M. maximus, alvo deste trabalho, foi estabelecida no Brasil durante o século XVII e chegou ao nosso território em navios negreiros, onde era utilizada como cama para os escravos. A boa adaptabilidade ao clima e ao solo brasileiros, bem como a capacidade de estabelecimento dessa planta, contribuíram para sua rápida distribuição pelo território, dando origem a cultivar Colonião, nome pelo qual a espécie ficou conhecida popularmente em nosso país. A introdução dessa planta resultou em maior produtividade animal e na disseminação de pastagens por sementes (**Figura 2**) (Combes e Pèrnes, 1970; Savidan et al., 1989; Jank et al., 2013).

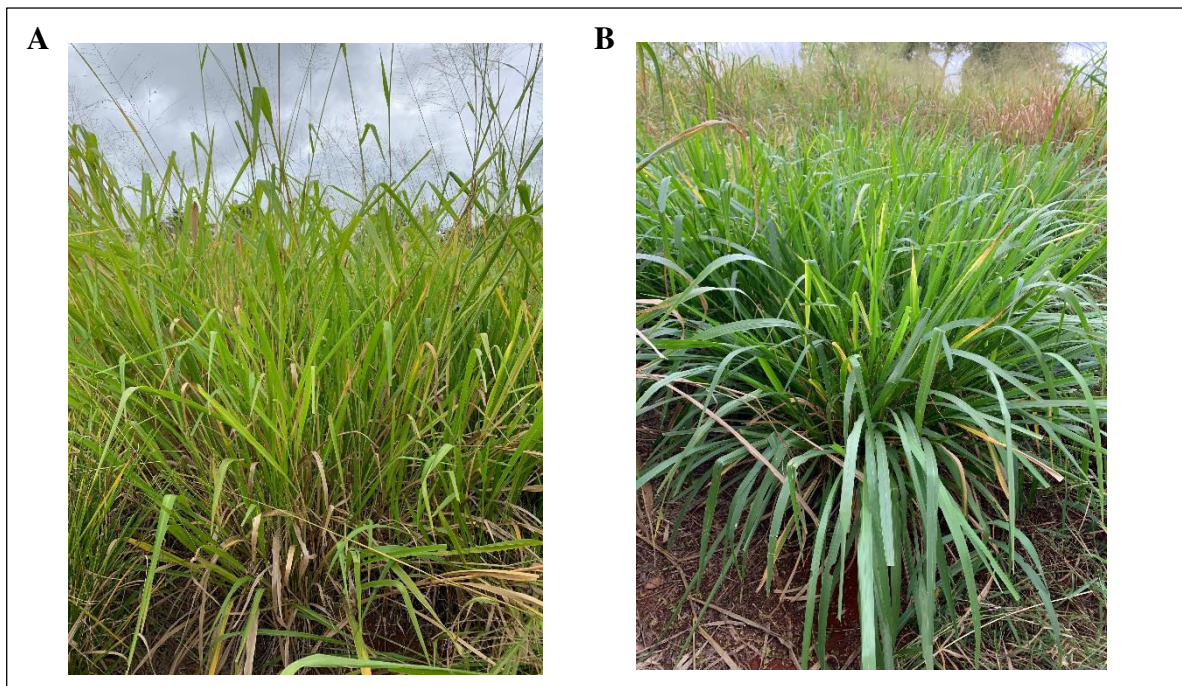


Figura 2. Exemplos de plantas da espécie *M. maximus*. Em A) a cultivar Mombaça, conhecida pelo porte alto e elevada produtividade; B) um dos genitores sexuais (S10), que possui menor porte, sendo usado no programa de melhoramento da espécie, localizado na Embrapa Gado de Corte, MS. Fonte: Liana Jank.

Expedições para coleta de material desta forrageira foram realizadas em seu centro de origem, entre o Quênia e a Tanzânia (África oriental), nos anos de 1967 e 1969, por um grupo de pesquisadores franceses do *Institut de Recherche pour le Développement* (ex-ORSTOM e atual IRD) (Combes e Pernès, 1970; Savidan et al., 1989). Este material foi cedido à Embrapa de Gado de Corte (Campo Grande, Mato Grosso do Sul), bem como outros genótipos de vários países tropicais e deram origem ao banco de germoplasma da espécie, com 426 acessos apomíticos e 417 sexuais (**Figura 3**). Com isso, o programa de melhoramento genético da gramínea adotou, como primeira estratégia, uma ampla avaliação dos caracteres agronômicos e a caracterização morfológica desta coleção (Jank et al., 1994; Jank et al., 1997). Essas análises demonstraram que mais de 40% dos 426 acessos exibiram caracteres agronômicos superiores a cv. Colômbio (Jank et al., 1994), resultado da ampla variabilidade genética existente.



Figura 3. Vista parcial do banco de germoplasma de *Megathyrsus maximus* mantido pela Embrapa Gado de Corte, MS. Fonte: Liana Jank.

Esse germoplasma no Brasil contém plantas com modo reprodutivo apomítico gametofítico, que naturalmente são tetraploides ($2n = 4x = 32$), e genótipos sexuais tetraploides. Essas plantas sexuais foram originadas a partir do cruzamento de um genitor apomítico e uma planta sexual originalmente diploide ($2n = 2x = 16$), que foi artificialmente tetraploidizada com o uso de colchicina pelo IRD. Desta forma, são possíveis cruzamentos entre plantas apomíticas

e sexuais para geração de uma maior variabilidade entre plantas (Warmke, 1954; Savidan e Pernès, 1982; Savidan, 1980; Savidan et al., 1989). Essa estratégia é usada pelo programa de melhoramento de *M. maximus*, de modo que genótipos superiores com modo de reprodução apomítico obtidos na primeira geração de progênies (F_1) já podem ser lançados, após todos os testes, como uma cultivar. A apomixia garante a fixação do vigor híbrido. Por outro lado, os genótipos sexuais superiores são utilizados como genitores em cruzamentos futuros (Jank et al., 2011; Jank et al., 2013).

A apomixia, a autotetraploidia e como lidar com essa herança polissômica têm sido os principais desafios para as pesquisas genéticas da espécie. A apomixia em *M. maximus* é do tipo apospórica seguida de partenogênese, como demonstrado na **Figura 4**. Dessa forma, os gametas reduzidos (n) pela meiose se degeneram e uma célula somática ($2n$) se desenvolve, originando um embrião ($2n$), ou seja, seu conjunto cromossômico é idêntico a planta-mãe. Enquanto que o endosperma da semente apomítica, que fornece nutrientes essenciais para o embrião em desenvolvimento, possui número cromossômico igual a $3n$, pois ocorre a fecundação do núcleo polar com um pólen doado por outra planta (n), processo conhecido como pseudogamia. Diferentemente, o saco embrionário sexual é classificado como *Polygonum* com três ou mais antípodas. Neste caso, o pólen externo (n) fecunda os núcleos polares e o óvulo gerando, portanto, embrião e endosperma híbridos (Savidan, 1982).

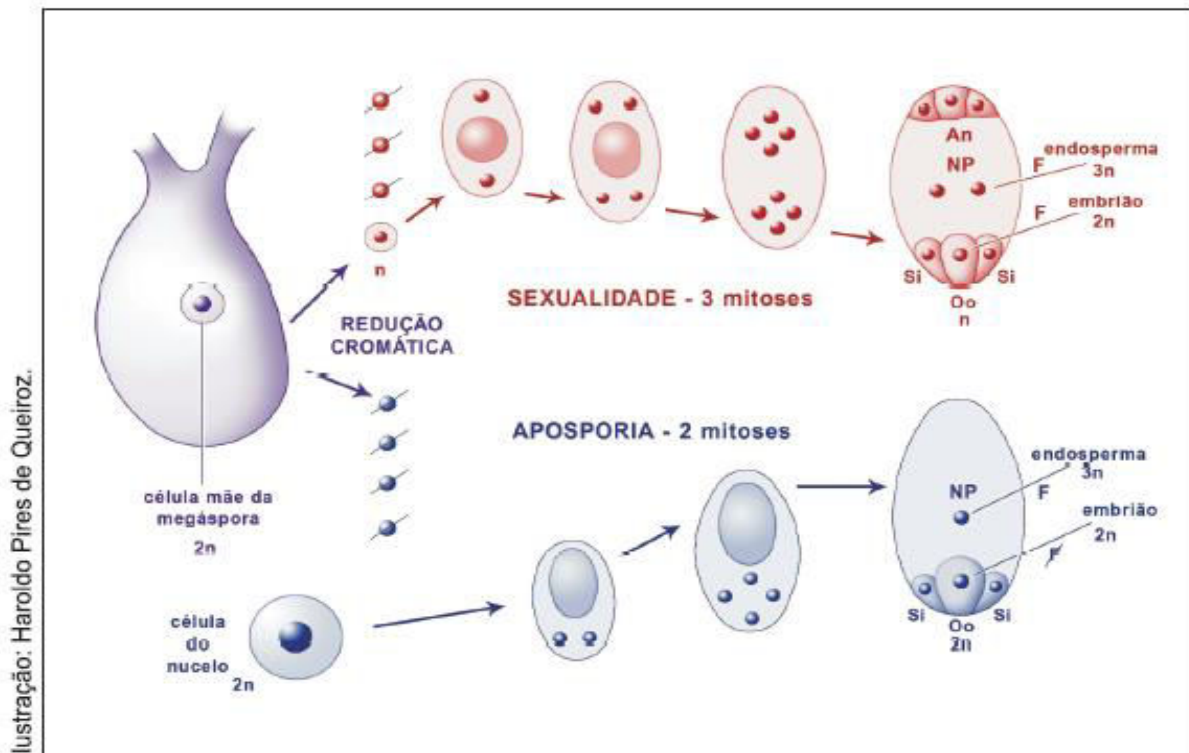


Figura 4. Modo reprodutivo sexual e apomítico de *Megathyrsus maximus*. An: antípoda; NP: núcleo polar; Si: sinérgide; Oo: oosfera; F: fecundação. Fonte: Jank et al., 2008.

O uso bem-sucedido da apomixia no melhoramento exige um conhecimento mais aprofundado sobre a genética envolvida no modo de reprodução de *M. maximus*, bem como dos mecanismos reguladores subjacentes ao desenvolvimento apomítico das sementes, que vêm sendo estudados há anos. A princípio, uma série de combinações de cruzamentos entre plantas apomíticas e sexuais foram realizadas a fim de verificar o tipo de segregação observada nos híbridos para o caráter. Savidan (1981) averiguou uma herança monogênica, em que apenas um gene poderia controlar a apomixia. O resultado de cruzamentos biparentais (genitor apomítico x genitor sexual) para a espécie gera híbridos em uma proporção 1:1, ou seja, cerca de 50% da população corresponde a indivíduos apomíticos e a outra metade a indivíduos sexuais. Evidências a partir de dados moleculares corroboram com o controle genético de um ou poucos genes responsáveis por apomeiose e partenogênese em algumas gramíneas, incluindo *M. maximus* (Ebina et al., 2005), *C. ciliaris* (Yadav et al., 2012; Kumar et al., 2017a), *Pennisetum squamulatum* (Ozias-Akins et al., 1998; Sapkota et al., 2016), *Urochloa* spp. (Pessino et al., 1997; Worthington et al., 2016; Worthington et al., 2019) e *Paspalum notatum* (Martinez et al., 2003; Oliveira et al., 2018). A partir do mapeamento genético e do uso de bibliotecas de BAC (cromossomo artificial bacteriano), foi identificada uma região genômica apospórica (ASGR – *apospory-specific genomic region*) em *Pennisetum* e *Cenchrus*, que é fisicamente extensa, hemizigótica e heterocromática, sendo base para estudos com marcadores moleculares ligados à apomixia em várias espécies (Ozias-Akins et al., 1998; Goel et al., 2003). Esses marcadores podem ser utilizados para a triagem de populações segregantes, bem como para auxiliar o mapeamento de genes responsáveis pelo modo de reprodução e sua caracterização (Bluma-Marques et al., 2014; Worthington et al., 2016; Worthington et al., 2019). Várias são as hipóteses discutidas sobre a base molecular da apomixia. Entre elas, foi sugerido, a partir de plantas apomíticas facultativas, que a apomixia possa ser resultado da desregulação da sexualidade e, por isso, compartilham dos mesmos mecanismos reguladores (Tucker et al., 2003). Além disso, há a possibilidade de controle do caráter por dois ou mais *loci* (Ozias-Akins et al., 1998).

Outros mecanismos genéticos também têm sido discutidos, como a remodelação da cromatina (epigenética) e os elementos transponíveis presentes no cromossomo portador do *locus* da apomixia (Kumar, 2017). Mudanças espaciais e temporais na expressão dos genes envolvidos em vias reprodutivas de plantas sexuais e apomíticas foram relatadas em plantas da

família Brassicaceae (Sharbel et al., 2010). Mecanismos reguladores epigenéticos podem estar envolvidos neste cenário, visto que espécies reativas de oxigênio (ERO) são produzidas em excesso nas plantas, sob estresses causados por condições abióticas e bióticas adversas (Kumar, 2017). Consequentemente, metilação do DNA e/ou modificações nas histonas podem ocorrer e afetar significativamente a expressão gênica (Yan et al., 2010; Kumar et al., 2017b). Na gramínea *P. squamulatum*, que contém estudos mais avançados sobre a apomixia, o gene da espermina/espermidina sintase foi identificado em ASGR (Conner et al., 2008), que auxilia na proteção das moléculas de DNA contra estresses oxidativos decorrentes de ERO. O papel dos retrotransposons no desenvolvimento apomítico também tem sido especulado, mas a sua atuação na expressão diferencial do indivíduo apomítico e sexual ainda não está claro (Akiyama et al., 2011; Huo et al., 2009; Yadav et al., 2012).

Além da apomixia, outro importante desafio relacionado a *M. maximus* está no conhecimento sobre seu genoma. A poliploidização e a subsequente evolução de genomas poliploides é um processo extremamente dinâmico. Os autopoliplóides exibem herança polissômica, visto que o emparelhamento e a recombinação podem ocorrer entre todas as cópias homólogas de cada cromossomo durante a meiose. A compreensão do comportamento de emparelhamento de espécies poliploides também é relevante para o estudo da evolução do genoma (Van de Peer et al., 2017; Bourke et al., 2018b). A autopoliploidia de *M. maximus* é caracterizada pelo número básico de cromossomos igual a 8 com quatro cópias de cada cromossomo homólogo, sendo que as combinações meióticas resultantes podem incluir quadrivalentes, trivalentes, bivalentes e univalentes (Savidan and Pernès, 1982). Geralmente, os genótipos sexuais dessa espécie são diploides com $2n = 2x = 16$ e os genótipos apomíticos são tetraploides com $2n = 4x = 32$ (Combes, 1975; Savidan, 2000; Jank et al., 2008). No entanto, outras ploidias já foram identificadas em menor número de plantas, como triploides ($2n = 3x = 24$), pentaploides ($2n = 5x = 40$), hexaploides ($2n = 6x = 48$), octaploides ($2n = 8x = 64$), nonaploides ($2n = 9x = 72$) e aqueles com irregulares cromossômicas ($2n = 30, 31, 34, 36, 37, 38$) (Combes, 1975; Jank et al., 2008).

Os autopoliplóides ainda exibem mais uma complexidade genômica, que é a distorção de segregação entre *loci* devido a um tipo de meiose mais frequente nestes organismos, chamada de redução dupla, no qual, as cromátides irmãs podem terminar no mesmo gameta. A redução dupla ocorre como resultado da formação de quadrivalentes de cromossomos homólogos (Mather, 1936). Não é possível prever a frequência de qualquer modo

de formação de gametas em termos da fração de recombinação entre dois *loci* e redução dupla (Bailey, 1961; Luo et al., 2004). Por meio de análises citogenéticas, segregações cromossômicas irregulares e a presença de quadivalentes em *M. maximus* foram verificadas em híbridos e genitores do programa de melhoramento da Embrapa Gado de Corte (Pessim et al., 2010; Pessim et al., 2015). No entanto, o evento da redução dupla, resultante dessas especificidades da divisão meiótica, ainda não havia sido reportado para esta espécie.

O conhecimento sobre a genética molecular dessa forrageira está em progresso, porém, ainda é incipiente. Interessantes resultados foram obtidos a partir de dois estudos de diversidade genético-molecular com genótipos do programa de melhoramento brasileiro dessa gramínea. O primeiro estudo contou com 20 marcadores do tipo microssatélite desenvolvidos para a espécie, que foram validados em 25 genótipos de *M. maximus* e em quatro espécies de gramíneas, *U. humidicola*, *U. brizantha*, *P. notatum* e *P. regnelli* (Sousa et al., 2011a). Como resultado, 135 alelos putativos foram obtidos, sendo que o número de alelos variou de 4 a 12, com uma média de 6,7 alelos por *locus*. O coeficiente de Jaccard para verificar a similaridade entre os genótipos demonstrou que todos eles estão intrinsecamente relacionados. Devido a natureza poliploide da espécie, não foi possível estimar o número exato de cópias para cada alelo. Por isso, as análises de desvio do equilíbrio de Hardy-Weinberg e desequilíbrio de ligação não foram aplicadas. O teste de transferibilidade dos marcadores permitiu a amplificação de observou 11 *locus* microssatélites para as quatro espécies. O segundo estudo molecular, avaliou a diversidade e estruturação genética de 396 acessos do banco de germoplasma de *M. maximus* da Embrapa Gado de Corte, também com marcadores microssatélites (Sousa et al., 2011b). A genotipagem com 55 marcadores microssatélites identificou 318 alelos, sendo que o número de alelos por *locus* microssatélite variou de 3 a 16, com uma média de 5,8 alelos por *locus*, evidenciando, portanto, o alto polimorfismo dos marcadores. Os acessos foram estruturados em quatro grupos genéticos, sem concordância com a região geográfica de origem que foram coletados, e uma alta variabilidade genética entre os genótipos foi indicada.

Outros grupos de pesquisa pelo mundo também têm investido esforços sobre a genética da forrageira. Utilizando genótipos de *M. maximus* do germoplasma do Japão, um estudo revelou que seu genoma é estimado em aproximadamente 500 Mpb para um diploide e em 1000 Mpb para um tetraploide (Akiyama et al., 2008). Além disso, a evolução das tecnologias de sequenciamento tem proporcionado maiores avanços na identificação de genes putativos e de marcadores moleculares para a espécie, como exemplo, dois transcriptomas da

espécie foram publicados até o momento. O primeiro transcriptoma contemplou quatro genitores utilizados no programa de melhoramento da Embrapa Gado de Corte, sendo dois deles sexuais (S10 e S12) e os outros dois apomíticos (cv. Mombaça e cv. Tanzania). A partir do tecido foliar foi possível a identificação de aproximadamente 89 mil transcritos que possibilitaram a anotação de genes hipotéticos possivelmente relacionados a via fotossintética C₄, a produção de celulose e lignina (Toledo-Silva et al., 2013). Para o segundo transcriptoma foram utilizados genótipos do melhoramento genético realizado na Índia, e possibilitou a identificação de cerca de 107 mil transcritos obtidos da fase pré-meiótica do desenvolvimento da flor, visando investigar os mecanismos moleculares que regulam a apomixia (Radhakrishna et al., 2018). Porém, maiores informações a respeito desse estudo ainda não foram publicadas. Estes avanços também contribuem para trabalhos voltados diretamente ao melhoramento genético. Recentemente, um estudo avaliou diferentes modelos de predição genômica aplicados a uma população de seleção recorrente de 530 genótipos da espécie que são provenientes do programa de melhoramento da Embrapa Gado de Corte. O efeito da dosagem do alelo na predição também foi considerado e, como resultado, concluiu-se que a seleção genômica pode levar a ganhos adicionais no programa de seleção recorrente de *M. maximus* (Lara et al., 2019). Apesar destes trabalhos, muitas informações ainda são desconhecidas, como a localização de genes responsáveis pelo modo reprodutivo, produtividade e qualidade da forragem, bem como as relações que são estabelecidas entre eles para com o genótipo e o ambiente.

2.3 Mapeamento genético e sua aplicação em *M. maximus*

Embora o primeiro mapa de ligação tenha sido desenvolvido há mais de cem anos (Sturtevant, 1913), o uso desta técnica para estudos genéticos e genômicos ainda persiste na era do *Next Generation Sequencing* (NGS). Isso pode ser atribuído a vários fatores. Um mapa de ligação é, em geral, a primeira representação genômica de uma espécie, construído comumente com base em um cruzamento biparental de interesse, e representações mais avançadas podem ser construídas a partir dele (Ott et al., 2015). A frequência de recombinação entre *locus* nos estudos de herança proporciona a distância genética e o ordenamento dos marcadores nos grupos de homologia do mapa de ligação (Schuster e Cruz, 2005). Esses mapas têm sido aplicados para culturas com importância econômica, sendo notadamente relevantes em espécies que apresentam pouca informação genômica disponível e que não possuem o genoma sequenciado. Entre suas aplicações, os mapas podem ser explorados para conectar e orientar corretamente os *scaffolds* na montagem de genomas (Fierst, 2015), bem como em análises de *loci* de caracteres quantitativos (QTL) que visam maior poder de detecção, devido a associação

entre marcadores moleculares ligados fisicamente a região genômica de interesse (QTL) (Bourke et al., 2018b).

A análise de ligação genética em poliploides é necessária, porém sua abordagem na genética quantitativa tem sido historicamente difícil, tendo início com os trabalhos pioneiros de Mather (1936) e Fisher (1947). Muitas pesquisas voltadas aos estudos de herança polissômica tiveram como objetivo facilitar as análises de mapeamento e, para isso, evitavam vários graus de complexidade (Hackett et al., 1998; Ma et al., 2002). No entanto, algumas características essenciais dessa herança e da análise prática de dados, como o fenômeno da redução dupla, eram desconsiderados (Luo et al., 2004).

As tecnologias para obtenção de marcadores moleculares, os modelos estatísticos aplicados à genética e os estudos de mapeamento evoluem por consequência um do outro. Os estudos de mapeamento eram baseados no uso de marcadores dominantes, como AFLP (*amplified fragment length polymorphism*) e RAPD (*random amplified polymorphic DNA*); e microssatélites (SSR - *simple sequence repeat*), que apesar de codominantes são avaliados como dominantes em estudos com poliploides, todos eles segregando em uma proporção de 1:1 (simplex). Com o advento das tecnologias de NGS, milhares de polimorfismos de base única (SNPs) podem ser identificados pelo sequenciamento de DNA em larga escala com maior qualidade de bases (Poland e Rife, 2012). Entre essas tecnologias, a genotipagem-por-sequenciamento (GBS), descrita por Elshire et al. (2011) e adaptada por Poland et al. (2012), tem sido fortemente empregada em genomas com alta diversidade genética. Essa estratégia de genotipagem baseia-se no desenvolvimento de biblioteca genômica com um *pool* gênico contendo dezenas ou mesmo centenas de indivíduos. A fragmentação do DNA genômico realizado por enzimas de restrição, adequadas ao perfil de digestão das amostras, permite que regiões com maior afinidade à metilação e regiões repetitivas do genoma não sejam aderidos à biblioteca, enquanto regiões com baixo número de cópias podem ser amostradas de duas a três vezes com maior eficiência. Estas enzimas de restrição e o uso de adaptadores *barcodes* únicos para cada indivíduo promovem uma maior facilidade no alinhamento de sequências e, conseqüentemente, auxiliam positivamente as análises de genomas poliploides (Poland & Rife, 2012).

Os marcadores SNP são a forma mais abundante de variação genética entre indivíduos da mesma espécie e compatíveis com plataformas de genotipagem automatizadas (Zargar et al., 2015). Esses marcadores passaram a ser amplamente usados nas análises de

ligação em poliploides e, com isso, estratégias para aumentar a profundidade das *reads* geradas, bem como a acurácia da identificação de SNPs têm sido reportadas. Logo, esses conjuntos de dados reais estimulam o desenvolvimento de novos e mais robustos métodos de análise (Blischak et al., 2017; Pereira et al., 2018; Zych et al., 2019).

Até pouco tempo eram usados somente SNPs em dose única (simplex) no mapeamento de poliploide, na condição 1:1, em que o alelo de referência está em apenas um dos genitores (ABBB x BBBB), ou com a configuração 3:1, sendo uma cópia do referido alelo para cada um dos genitores (ABBB x ABBB). À medida que a ploidia aumenta, torna-se cada vez mais difícil distinguir heterozigotos. Ao considerar um *locus* bialélico autotetraploide, onde A e B são os respectivos alelos de referência e alternativo, espera-se, além dos homozigotos, AAAA (nulliplex) e BBBB (tetraplex), três classes diferentes de heterozigotos: ABBB (simplex), AABB (duplex) e AAAB (triplex). Por isso, a profundidade de sequenciamento é essencial para que a dosagem de cada indivíduo nesses *loci* seja identificada buscando maior acurácia (Pereira et al., 2018). *Softwares* com modelos matemáticos capazes de lidar com profundidade (Pereira et al., 2018; Gerard et al., 2018; Zych et al., 2019) e com as cinco classes de dosagem podem promover avanços significativos para autotetraploides. Este avanço baseia-se na utilização de uma maior quantidade de informação genética, uma vez que é possível que as configurações de doses sejam representadas no mapa, dessa forma, abrangendo uma maior parte do genoma (Hackett et al., 2017; Bourke et al., 2018a; Mollinari et al., 2019). Essa maior quantidade de informação está relacionada a estimação correta das fases de ligação dos marcadores (repulsão ou associação), usadas na construção dos mapas integrados de homólogos, que determinam se os alelos segregantes estão no mesmo homólogo (então, consideramos que os marcadores estão ligados na fase de associação) ou se estão em diferentes homólogos do mesmo cromossomo (assim, estão ligados na fase de repulsão) (Qu e Hancock, 2001). As fases podem ser melhor atribuídas com marcadores de todas as dosagens, pois em caso de somente marcadores simplex, informações limitadas sobre a ligação na fase de repulsão serão fornecidas e a obtenção de um mapa integrado pode não ocorrer (van Geest et al., 2017).

O uso destes mapas de ligação no mapeamento de QTLs, tem como objetivo auxiliar na identificação de regiões do genoma que possuem uma associação estatisticamente significativa a um caráter avaliado a partir de uma progênie biparental, cujos genitores sejam contrastantes para o fenótipo (Li et al., 2011; Hackett et al., 2014; Bourke et al., 2019). Geralmente, esses caracteres são poligênicos e possuem uma distribuição contínua, ou seja,

vários genes (*locus*) são responsáveis por governar uma característica e estão envolvidos em níveis de expressão do fenótipo. Como exemplo, altura da folha, capacidade de rebrota, porcentagem de proteína bruta encontrada no tecido foliar e colmo, entre outros. Esses genes (*locus*) individualmente possuem um pequeno efeito sobre o fenótipo do caráter e em conjunto resultam em um efeito cumulativo na atribuição do caráter (Holland, 2007). Esses efeitos podem ser aditivos, de dominância/ou epistático, em que pode existir forte influência ambiental sobre eles (Allard, 1971).

O primeiro mapa de ligação de *M. maximus* foi construído a partir da cultivar Natsukaze ($2n = 4x = 32$), utilizada em pastagens no Japão (Ebina et al., 2005). Um total de 39 grupos de ligação (GL) foram obtidos, número maior do que esperado para a espécie, a partir de 360 *loci* de marcadores dominantes, sendo 56 combinações de *primers* AFLP e de 41 *primers* RAPD, que cobriram 1703,5 cM. Esse mapa permitiu a identificação do *locus* da apomixia, localizado no GL 8, um grande avanço para a ciência da época, pois marcadores localizados próximos ao *Apo-locus* apresentaram bons resultados para aplicação em programas de melhoramento, na diferenciação entre híbridos sexuais e apomíticos (Ebina et al., 2005).

A partir de uma progênie biparental do programa de melhoramento de *M. maximus* de nosso país, obtida do cruzamento entre S10 e cv. Mombaça (Figura 2), foi proposto um novo mapa genético para a espécie (Toledo-Silva, 2013). Este mapa foi construído com marcadores microssatélites (SSRs) desenvolvidos pelo nosso grupo de pesquisa (Sousa et al., 2011a; Sousa et al., 2011b), além de SSRs altamente polimórficos identificados no transcriptoma de referência da forrageira (Toledo-Silva et al., 2013). Porém, devido ao número limitado de marcadores, foi obtido um mapa genético com baixa resolução e com o número de grupos de ligação não esperado para a espécie.

A fim de aumentar a resolução desse mapa de ligação, Rabello (2017) desenvolveu marcadores SNPs identificados no transcriptoma (Toledo-Silva et al., 2013) e genotipou a progênie com a técnica de espectrometria de massas na plataforma Sequenom iPLEX MassARRAY® (Sequenom Inc., San Diego, California, USA). O mapa foi construído com 122 marcadores, sendo 97 SSRs genotipados por Toledo-Silva (2013) e apenas 25 SNPs oriundos do transcriptoma. Os 32 grupos de ligação esperados para a espécie foram obtidos, porém, o mapa gerado apresentou baixa resolução novamente.

Apesar de todos os esforços, ainda há pouca informação sobre a estrutura e organização do genoma de *M. maximus*. Até o presente trabalho, não havia sido alcançado um mapa de alta resolução e com informação de dosagem alélica para a espécie e também nenhuma informação a respeito de caracteres quantitativos. Em gramíneas forrageiras de pastagens temperadas, QTLs relacionados a produção e qualidade da forragem já foram detectados (Fang et al., 2016; Rajput et al., 2016; Sartie et al., 2018; Khasanova et al., 2019), porém sem o auxílio do mapeamento utilizando marcadores com dosagem alélica. Enquanto isso, apesar de recentes os estudos com QTLs em gramíneas forrageiras tropicais (Ferreira et al., 2019), estamos implementando a dosagem alélica para resultados mais abrangentes e informativos. Além disso, mais informações sobre a aposporia nesta gramínea são acrescentadas, bem como possíveis marcadores associados ao *locus* da aposporia. Os resultados alcançados com esta tese, ajudará a orientar análises de QTLs em autopoliplóides, bem como a aprofundar nossa compreensão da arquitetura genética que controla características importantes de *M. maximus* e de espécies próximas.

3 OBJETIVOS

3.1 Geral

Detectar regiões genômicas e estabelecer as relações genéticas entre caracteres de importância econômica para *Megathyrus maximus* com a aplicação de ferramentas genético-estatísticas recentes.

3.2 Específicos

- Detectar marcadores SNPs com informação de dosagem alélica em uma progênie intraespecífica de *M. maximus*;
- Construir um mapa de ligação consenso e integrado a partir desta progênie;
- Mapear a região do *locus* da aposporia;
- Mapear QTLs associados a caracteres de produção e a qualidade da forragem;
- Investigar possíveis genes candidatos nas regiões dos QTLs e do *locus* da aposporia;
- Estabelecer correlações genéticas entre os caracteres quantitativos importantes para o melhoramento genético de *M. maximus*.

4 ARTIGO

Title

High-resolution Linkage Map with Allele Dosage Allows the Identification of Regions Governing Complex Traits and Apospory in Guinea Grass (*Megathyrus maximus*)

Authors

Thamiris G. Deo¹, Rebecca C. U. Ferreira¹, Letícia A. C. Lara², Aline C. L. Moraes³, Alessandro Alves-Pereira¹, Fernanda A. de Oliveira¹, Antonio A. F. Garcia², Mateus F. Santos⁴, Liana Jank⁴, Anete P. Souza^{1,3*}

¹Center for Molecular Biology and Genetic Engineering, University of Campinas, Campinas, Brazil

²Genetics Department, Escola Superior de Agricultura “Luiz de Queiroz”, University of São Paulo, Piracicaba, Brazil

³Plant Biology Department, Biology Institute, University of Campinas, Campinas, Brazil

⁴Embrapa Beef Cattle, Brazilian Agricultural Research Corporation, Campo Grande, Brazil

* Correspondence: Anete Pereira de Souza anete@unicamp.br

Abstract

Forage grasses are mainly used in animal feed to fatten cattle and dairy herds, and guinea grass (*Megathyrsus maximus*) is considered one of the most productive of the tropical forage crops that reproduce by seeds. Due to the recent process of domestication, this species has several genomic complexities, such as autotetraploidy and aposporous apomixis. Consequently, approaches that relate phenotypic and genotypic data are incipient. In this context, we built a linkage map with allele dosage and generated novel information of the genetic architecture of traits that are important for the breeding of *M. maximus*. From a full-sib progeny, a linkage map containing 858 single nucleotide polymorphism (SNP) markers with allele dosage information expected for an autotetraploid was obtained. The high genetic variability of the progeny allowed us to map 10 quantitative trait *loci* (QTLs) related to agronomic traits, such as regrowth capacity and total dry matter, and 36 QTLs related to nutritional quality, which were distributed among all homology groups (HGs). Various overlapping regions associated with the quantitative traits suggested QTL hotspots. In addition, we were able to map one *locus* that controls apospory (*apo-locus*) in HG II. A total of 55 different gene families involved in cellular metabolism and plant growth were identified from markers adjacent to the QTLs and apospory *locus* using the *Panicum virgatum* genome as a reference in comparisons with the genomes of *Arabidopsis thaliana* and *Oryza sativa*. Our results provide a better understanding of the genetic basis of reproduction by apomixis and traits important for breeding programs that considerably influence animal productivity as well as the quality of meat and milk.

Keywords: apospory, double reduction, forage, polyploid, QTL, SNP, trait correlations.

Introduction

Forage grasses play a fundamental role in the global beef production chain. Brazil is the country with the greatest emphasis on this sector, being the main exporter of beef and having the largest commercial herd of beef cattle in the world, with approximately 215 million heads distributed in 162 million hectares of pasture (ABIEC, 2019). The main factor that led to this scenario was the beginning of tropical forage breeding in the 1980s in Brazil, which although recent, permitted the country to become the world's largest exporter of tropical forage seeds (ITC, 2018).

The African forage grass species *Megathyrsus maximus* (Jacq.) B. K. Simon & S. W. L. Jacobs (syn. *Panicum maximum* Jacq.), also known as guinea grass, is one of the most productive forage grasses reproduced by seeds in the Brazilian market and is also grown in other Latin American countries (Jank et al., 2011). It has been used mainly in intensive systems with high-fertility soils (Valle et al., 2009). Moreover, this forage has a high biomass potential and is promising as a biofuel feedstock (Odorico et al., 2018). The polyploidy and domestication process of this forage grass ensure high genetic variability to be explored (Jank et al., 2011); however, a lack of knowledge of the biology and genetics of the species, including its autotetraploidy and facultative apomictic mode of reproduction (Warmke, 1954), may make breeding more difficult and thus stimulate a need to invest in genetic studies.

The polysomic inheritance in autopolyploids makes genetic research difficult because several types of segregation may be involved (Field et al., 2017) and complexity increases as ploidy increases. Thus, in a segregating population, genetic complexity may influence the segregation and the frequencies of expected genotypes (Field et al., 2017), even causing distortions such as double reduction (DR), which is a type of meiosis in which the sister chromatids are duplicated, forming an unexpected combination of gametes. For example, an autotetraploid genotype with the 'abcd' alleles at a *locus* that can form six expected combinations of gametes, namely, 'ab', 'ac', 'ad', 'bc', 'bd', and 'cd', but a homozygous gamete can be generated, e.g., 'aa', 'bb', 'cc', or 'dd' (Haldane, 1930; Mather, 1935; Haynes and Douches, 1993). How to accommodate DR and its implications in a breeding program as well as the use of these marker *loci* in linkage mapping has been discussed for a long time (Butruille and Boiteux, 2000; Luo et al., 2000; Xu et al., 2013; Layman and Busch, 2018; Bourke et al., 2019) for some economically important species, such as potato (Bradshaw, 2007; Bourke et al., 2015) and alfalfa (Julier et al., 2003), but has not yet been reported in guinea grass.

Linkage maps have been used as the primary source of genetic information for species mainly nonmodel species that do not have their genome sequenced, such as *M. maximus*. The construction of dense linkage maps allows the identification of the structure and evolution of the genome by mapping traits with polygenic and monogenic inheritance and may even contribute to the assembly of the genome of a species (Doerge, 2002; Flint-Garcia et al., 2003; Luo et al., 2004). The majority of linkage maps available for autotetraploids are based on single-dose segregating markers for a parent (Aaaa x aaaa) or a single dose for both parents (Aaaa x Aaaa). Despite the use of complex statistical methods to obtain integrated maps that combine information from both marker patterns, these maps cover only part of the genome because higher-dose markers (AAaa, AAAa, and AAAA) are not included. This limitation results in a considerable loss of genetic information. To overcome this limitation, a new approach allows the assignment of allele dosage information for single nucleotide polymorphism (SNP) markers through exact allele sequencing depth, which generates linkage maps from markers in multiple doses with higher quality, more information and greater applicability, including more efficient detection of *loci* related to traits of economic importance (Serang et al., 2012; Hackett et al., 2014; Pereira et al., 2018; Mollinari and Garcia, 2019).

Mapping of the aposporous apomixis region is extremely important for the genetic breeding of *M. maximus* and other forage grasses, such as *Urochloa* spp., *Paspalum* spp., and *Cenchrus ciliaris*. These species undergo asexual propagation by seeds (Nogler, 1984), which allows the fixation of hybrid vigor in apomictic individuals and their use in the creation of uniform pastures (Jank et al., 2011). Experimental field data showed that aposporous apomixis in tropical forage grasses follows 1:1 Mendelian segregation, indicating monogenic inheritance (Savidan, 1981; Valle et al., 1994; Chen et al., 2000; Savidan, 2000), although a recent study suggests that this reproductive mode should be treated as a quantitative trait (Marcón et al., 2019), and provide evidence of the uncoupling of apomixis in neo-apomictic species, such as guinea grass (Kaushal et al., 2008; Kaushal et al., 2019). Molecularly, the apospory-specific genomic region (ASGR), which is responsible for apospory, is highly conserved among apomictic species (Gualtieri et al., 2006). The influence of some factors, such as epigenetics (Kumar, 2017), the presence of retrotransposons (Akiyama et al., 2011), and gene duplication (Conner et al., 2008), in this region has been reported for other species. Due to the laborious and time-consuming methods required to phenotype apomixis, several studies on forage grasses have searched for markers intrinsically linked to the chromosomal region for this trait (Pessino et al., 1998; Ebina et al., 2005; Bluma-Marques et al., 2014; Thaikua et al., 2016; Vigna et al.,

2016; Worthington et al., 2016, 2019); however, a 100% efficient marker for use in *M. maximus* breeding programs has not yet been identified.

In guinea grass there are no mapping studies of *loci* related to complex traits, such as those involved in forage yield and nutritional quality. Forage yield results from the continuous emission of leaves and tillers, ensuring the restoration of the leaf area after grazing in perennial pastures. Additionally, the nutritional value of a forage is directly related to animal performance and is measured by the crude protein, *in vitro* digestibility, neutral and acid detergent fiber, and lignin percentages (Jank et al., 2011). Thus, the mapping of these and other important quantitative trait *loci* (QTLs) may provide information about the genetic architecture of traits and assist in new strategies for breeding programs of *M. maximus*.

In this context, given the importance of new genomic studies in guinea grass for both biological knowledge and support for breeding programs, our goals were to (i) construct an integrated consensus linkage map from a full-sib progeny of *M. maximus* using allele dosage information, (ii) detect QTLs related to important agronomic and nutritional traits in this progeny, (iii) map the *apo-locus*, and (iv) search for similarity in regions of the markers adjacent to the QTLs and apospory *locus* in *Arabidopsis thaliana*, *Panicum virgatum* and *Oryza sativa*.

Material and methods

Plant material

A full-sib progeny of 136 F₁ hybrids was obtained from a cross between a facultative apomictic genotype of *M. maximus* cv. Mombaça and an obligate sexual genotype, S10. For the crosses, 100 m² of cv. Mombaça was sown in lines with a space of 1 m between the lines. One plant of S10 was planted in each 5 m x 5 m grid. The sexual plants were monitored to ensure the synchronization of flowering with the cv. Mombaça, and the inflorescences that flowered before cv. Mombaça were cut. According to Savidan (1982), 25 m² of the apomictic parent is sufficient to pollinate one sexual plant and to impede contamination from neighboring plants. The S10 plants are wind-pollinated, and the S10 seeds were harvested as they matured.

The sexual parent was derived from sexual x apomictic crosses of an original diploid sexual plant that was duplicated with colchicine (Savidan and Pernès, 1982); thus, both parents were autotetraploid (2n = 4x = 32) (Savidan et al., 1989). In addition to the reproductive mode, the parents have contrasting agronomic and nutritional quality traits (Braz et al., 2017). S10 is

a medium-sized plant (1.4 m tall) with medium width leaves (2.4 cm wide). Both its leaves and stems are glabrous, and its inflorescences consist of a panicle with short primary ramifications and long secondary ramifications throughout. Its spikelets are glabrous and purplish. The cv. Mombaça is a tall plant (1.7-m tall) with wide leaves (3 cm). Its leaves have small amount of hairs, and its stems are glabrous. Its inflorescences comprise a panicle with short primary ramifications and long secondary ramifications only on the inferior ramifications, and its spikelets are glabrous and light purple.

DNA extraction followed the protocol described by Doyle and Doyle (1987), with modifications. DNA samples were visualized on 2% agarose gels to check their quality and integrity, and their concentrations were estimated using a Qubit 3.0 fluorometer (Thermo Scientific, Wilmington, USA).

To retain only true full-sibs, all possible hybrids were previously genotyped with microsatellite markers. This analysis revealed 24 false hybrids, which were excluded at the construction stage of the GBS library.

Experimental design

A field experiment following an augmented block design (ABD) with 160 regular treatments (full-sib progeny) and two checks (the parents ‘Mombaça’ and S10) distributed in eight blocks with two whole replicates was performed at Embrapa Beef Cattle (Brazilian Agricultural Research Corporation), in Campo Grande, Mato Grosso do Sul state, Brazil (20°27’S, 54°37’W, 530 m). Each block consisted of a total of 22 plots (20 individuals and two checks).

Each plant was evaluated for agronomic and nutritional quality traits, totaling 22 traits: i) agronomic traits: green matter (GM - g/plant), total dry matter (TDM - g/plant), leaf dry matter (LDM - g/plant), stem dry matter (SDM - g/plant), regrowth capacity (RC), and percentage of leaf blade (PLB - %) and ii) nutritional quality traits for the leaf and stem: organic matter (OM_L and OM_S, respectively - %), crude protein (CP_L and CP_S - %), *in vitro* digestibility of organic matter (IVD_L and IVD_S - %), neutral detergent fiber (NDF_L and NDF_S - %), acid detergent fiber (ADF_L and ADF_S - %), cellulose (CEL_L and CEL_S - %), silica (SIL_L and SIL_S - %), and permanganate lignin (PL_L and PL_S - %). The agronomic traits were evaluated for six harvests (three harvests in 2013 and three harvests in

2014), but RC was evaluated for only three harvests (one harvest in 2013 and two harvests in 2014). The nutritional quality traits were evaluated for only one harvest in 2014.

Statistical analysis of phenotypic data

Descriptive analyses were performed, and the Box-Cox transformation (Box and Cox, 1964) was applied to correct for nonnormality of the residuals. For traits with multiple harvests (agronomic traits), we fitted the following longitudinal linear mixed model:

$$y_{ijkl} = \mu + h_l + r_{k(l)} + b_{j(l)} + rb_{kj(l)} + t_{i(l)} + \varepsilon_{ijkl}$$

where y_{ijkl} was the phenotypic value of the i^{th} treatment in the j^{th} block and k^{th} replicate at the l^{th} harvest; μ was the fixed overall mean; h_l was the fixed effect of the l^{th} harvest ($l = 1, \dots, L$, with $L = 3$ for RC and $L = 6$ for the other traits); $r_{k(l)}$ was the fixed effect of the k^{th} replicate ($k = 1, \dots, K$, with $K = 2$) at harvest l ; $b_{j(l)}$ was the random effect of the j^{th} block ($j = 1, \dots, J$, with $J = 8$) at harvest l , with $b_{j(l)} \sim N(0, \sigma_b^2)$; $rb_{kj(l)}$ was the random interaction effect of replication k and block j at harvest l , with $rb_{kj(l)} \sim N(0, \sigma_{rb}^2)$; $t_{i(l)}$ was the effect of the i^{th} treatment ($i = 1, \dots, I$, with $I = 162$) at harvest l ; and ε_{ijkl} was the random environmental error. The treatment effects ($t_{i(l)}$) were separated into two groups: $g_{i(l)}$ was the random effect of the i^{th} individual genotype ($i = 1, \dots, I_g$, with $I_g = 160$) at harvest l , and $c_{i(l)}$ was the fixed effect of the i^{th} check ($i = 1, \dots, I_c$, with $I_c = 2$) at harvest l . For genotype effects, the vector $\mathbf{g} = (g_{11}, \dots, g_{I_g L})'$ was assumed to follow a multivariate normal distribution with a mean of zero and genetic variance-covariance (VCOV) matrix $\mathbf{G} = \mathbf{G}_L \otimes \mathbf{I}_{I_g}$, i.e., $\mathbf{g} \sim N(\mathbf{0}, \mathbf{G})$. For residual effects, the vector $\boldsymbol{\varepsilon} = (\varepsilon_{1111}, \dots, \varepsilon_{IJKL})'$ followed a multivariate normal distribution with a mean of zero and residual VCOV matrix $\mathbf{R} = \mathbf{R}_L \otimes \mathbf{I}_{I \cdot J \cdot K}$, i.e., $\boldsymbol{\varepsilon} \sim MVN(\mathbf{0}, \mathbf{R})$.

The VCOV matrices \mathbf{G}_L and \mathbf{R}_L were analyzed considering seven different structures: identity (ID), diagonal (DIAG), compound symmetry (CS), heterogeneous compound symmetry (CS_{Het}), first-order autoregressive (AR1), heterogeneous first-order autoregressive (AR1_{Het}), first-order factor analytic (FA1), and unstructured (US). First, the genetic VCOV matrix (\mathbf{G}_L) was analyzed considering the ID for the residual matrix (\mathbf{R}_L), and posteriorly, the residual matrix (\mathbf{R}_L) was analyzed considering the selected VCOV matrix for genetic effects. Model selection was performed based on the Akaike information criterion (AIC) (Akaike, 1974) and Schwarz information criterion (SIC) (Schwarz, 1978).

For the nutritional quality traits, we fitted the following linear mixed model:

$$y_{ijk} = \mu + r_k + b_j + rb_{jk} + t_i + \varepsilon_{ijk}$$

where y_{ijk} was the phenotypic value of the i^{th} treatment in the j^{th} block and k^{th} replication; μ , r_k , b_j , rb_{kj} , t_i , and ε_{ijk} were as described above but not nested within harvest and with $\varepsilon_{ijk} \sim N(0, \sigma_{\varepsilon}^2)$. The treatment effects (t_i) were separated into two groups: g_i as a random effect, $g \sim N(0, \sigma_g^2)$, and c_i as a fixed effect. All analyses were performed with the R package ASReml-R (Butler et al., 2009).

The heritability of each trait was calculated using the same model as previously mentioned but considering the G_L and R_L matrices as the ID. The equation was

$$\hat{H}^2 = \frac{\sigma_g^2}{\sigma_p^2}$$

where σ_g^2 was the genetic variance and σ_p^2 was the phenotypic variance. The network analysis was carried out using the R package ‘qgraph’ (Epskamp et al. 2012).

Identification of the reproductive mode

The aposporic or sexual reproductive mode was determined for 106 hybrids of the progeny (**Supplementary Table 1**). From the flowers collected during anthesis, we performed an analysis of 30 ovules per hybrid using the clarified ovary method described by Young et al. (1979). Nomarski differential interference contrast microscopy was used to view the ovaries. A chi-square test was performed to verify the Mendelian segregation of this trait according to the expected model of monogenic inheritance in the base package of R (version 3.5.0) (R Core Team, 2018).

GBS library preparation and sequencing

From the extracted DNA, genotyping-by-sequencing (GBS) libraries were built according to Poland et al. (2012), containing 12 replicates for each parent. A total of 200 ng of genomic DNA per sample was digested with a combination of a rare-cutting enzyme (*PstI*) and a frequently cutting enzyme (*MspI*). DNA fragments were ligated to the common and barcode adapters, and the libraries were sequenced as 150 bp single-end reads using the High Output v2 Kit (Illumina, San Diego, CA, USA) for the NextSeq 500 platform (Illumina, San Diego, CA, USA).

SNP calling and allele dosage analysis

First, raw data were checked for quality using NGS QC Toolkit (Patel and Jain, 2012). SNP calling analysis was performed using the TASSEL-GBS v.4 pipeline (Glaubitz et

al., 2014) modified for polyploids (Pereira et al., 2018) that use exact read depths. The default parameters were changed as follows: the minimum number of times a GBS tag must be present was changed to 5, and the minimum count of reads for a GBS tag was changed to 2. This pipeline requires a genome as a reference for SNP calling, but no genome sequence of *M. maximus* is available. To overcome this limitation, the switchgrass genome (*P. virgatum* v1.0, produced by the US Department of Energy Joint Genome Institute) available in the Phytozome database (<http://phytozome.jgi.doe.gov/>) (Goodstein et al., 2012) was chosen because this species is phylogenetically closely related to *M. maximus* (Burke et al., 2016). GBS tags were aligned to the reference genome with Bowtie2 2.3.1 (Langmead et al., 2009) using the following settings: very-sensitive-local, a limit of 20 dynamic programming problems (D) and a maximum of 4 times to align a read (R). Subsequently, only tags that aligned exactly one time were processed. Then, SNP calling was performed under the conditions that the minor allele frequency was greater than 0.05 and the minor allele count was greater than 1,000. Mismatches of duplicated SNPs greater than 0.2 were not merged. Then, in R software (version 3.5.0) (R Core Team, 2018), we selected only the SNPs with a minimum average allele depth equal to or greater than 60 reads. The updog package (Gerard et al., 2018) was used to estimate the allele dosage of these markers, with a fixed ploidy parameter of 4 and the flexdog function considering the F₁ population model. SNPs with less than 0.15 of the posterior proportion of individuals incorrectly genotyped were selected. GBS sequences of each individual were deposited in the NCBI database under number PRJNA563938.

Quality filtering of SNPs

We removed markers with more than 25% missing data and monomorphic markers manually in R software (version 3.5.0) (R Core Team, 2018). Subsequently, we followed Bourke et al. (2018a) to ensure the retention of reliable markers. We first verified the shifted markers for the polysomic inheritance model from which SNP markers that did not correspond to an expected segregation type were removed. A threshold of 5% was used for missing values per marker and per individual. Duplicated markers, which provided no extra information about a *locus*, were also removed in this step. Finally, a principal component analysis (PCA) was performed to identify individuals who deviated from the progeny as well as possible clones.

Linkage map construction

A linkage map was construed using TetraploidSNPMap version 3.0 (Hackett et al., 2017), which allows the use of SNP markers with allele dosage data for autotetraploid species. SNP markers were checked with a chi-square test for goodness of fit, and only markers with a simplex configuration value greater than 0.001 and a segregation value greater than 0.01 for higher dosage were selected for mapping. Some unselected markers were classified as having segregation distortion (SD), being incompatible with the parental dosages (NP) and having DR. To order the selected markers, two-point analysis and multidimensional scaling analysis (MDS) were used to calculate recombination fractions and logarithm of odds (LOD) scores. Outlier markers were removed in this step. Some phases of the linked SNPs were inferred by TetraploidSNPMap software, and other phases were determined manually. The integrated consensus map represented by homology groups (HGs) was plotted using MapChart 2.32 (Voorrips, 2002), in which SNP configurations were identified with different colors.

Monogenic and polygenic trait analysis

QTL mapping of six agronomic traits and sixteen nutritional quality traits was performed with TetraploidSNPMap, applying an interval mapping model (Hackett et al., 2014, 2017). Analyses were conducted for each HG separately using three data files: phenotypic trait data, genotypic data and map data with phase information. The phenotypic data for the reproductive mode, i.e., aposporic or sexual, were considered qualitative due to the evaluation method applied; i.e., apomictic individuals were coded as one, and sexual individuals, as zero. The other phenotypic traits were analyzed as quantitative. QTL positions and significance were evaluated with a 1,000 permutation test. A QTL was declared significant if its LOD score was above the 90% threshold. Simple models were tested for each significant QTL to verify the best QTL model. The lowest SIC (Schwarz, 1978) was the criterion used to define the best model. Using TetraploidSNPMap software, if two or more significant QTLs were identified on the same chromosome, only the one with the greatest effect was considered.

Search for similarity in aposporic and QTL regions

We performed a search for similarity of candidate genes located close to the detected QTL/aposporic *locus* regions. Using the switchgrass genome as a reference and based on chromosomal locations of the markers adjacent to the detected QTLs and aposporic *locus*, we aligned the sequences found in 100 kb regions with Basic Local Alignment Search Tool

(BLAST) (e-value cutoff of $1e-0.5$) against the *A. thaliana* and *O. sativa* genomes through the JBrowse tool in Phytozome (<http://phytozome.jgi.doe.gov/>) (Goodstein et al., 2012).

Results

Phenotypic data

Different VCOV matrices were selected for the agronomic and nutritional traits (**Supplementary Tables 2 and 3**). When the AIC and SIC were not in agreement, we selected the matrices based on the largest difference between the models. For example, considering the GM trait and the G_L matrix, US was selected based on the AIC, and CS was selected based on the SIC (US had 567.32 for the AIC and 702.21 for the SIC, and CS had 581.66 for the AIC and 598.53 for the SIC). The differences between these two selected models were 14.34 for the AIC ($581.66-567.32$) and 103.68 for the SIC ($702.21-598.53$). As the SIC produced the largest difference, this criterion was used, and the CS matrix was selected for the G_L matrix of the GM trait. The heritabilities ranged from 0.19 (PLB) to 0.64 (GM) for the agronomic traits and from 0.06 (SIL_S) to 0.31 (OM_L and CEL_S) for the nutritional quality traits (Table 1). Box-Cox transformation was performed for agronomic traits (GM, TDM, LDM, SDM, RC, PLB) and nutritional traits of the leaves and stems (OM_L, OM_S, CP_S, IVD_L, IVD_S, NDF_S, ADF_L, CEL_L, SIL_L, SIL_S, PL_L, and PL_S).

The correlations between the agronomic and nutritional traits are presented in **Figure 1**. Significant and positive correlations were observed among GM, TDM, SDM, LDM, and RC. PLB presented a negative correlation with other agronomic traits and a positive correlation with NDF, but these correlations were weaker than those found for the other pairs of traits. Complex correlations between nutritional traits were obtained. In particular, stem-related traits were most tightly correlated with leaf-related traits, and NDF_L and NDF_S were closely correlated. The ADF, CEL and PL traits exhibited a significant and positive correlation. The same pattern was observed for the leaf and stem nutritional traits: CP and IVD exhibited a strong negative correlation with PL, ADF, CEL and NDF.

Table 1. Broad-sense heritability obtained for the agronomic and nutritional traits for the F₁ mapping population of guinea grass (*Megathyrsus maximus*) evaluated in this study.

Traits	H²	
Agronomic	Green matter (GM)	0.64
	Total dry matter (TDM)	0.57
	Leaf dry matter (LDM)	0.58
	Stem dry matter (SDM)	0.35
	Percentage of leaf blade (PLB)	0.19
	Regrowth capacity (RC)	0.36
	Nutritional	Leaf organic matter (OM_L)
Stem organic matter (OM_S)		0.15
Leaf crude protein (CP_L)		0.13
Stem crude protein (CP_S)		0.28
Leaf neutral detergent fiber (NDF_L)		0.14
Stem neutral detergent fiber (NDF_S)		0.14
Leaf acid detergent fiber (ADF_L)		0.25
Stem acid detergent fiber (ADF_S)		0.30
Leaf <i>in vitro</i> digestibility of organic matter (IVD_L)		0.18
Stem <i>in vitro</i> digestibility of organic matter (IVD_S)		0.26
Leaf cellulose (CEL_L)		0.26
Stem cellulose (CEL_S)		0.31
Leaf permanganate lignin (PL_L)		0.18
Stem permanganate lignin (PL_S)		0.18
Leaf silica (SIL_L)		0.17
Stem silica (SIL_S)		0.06

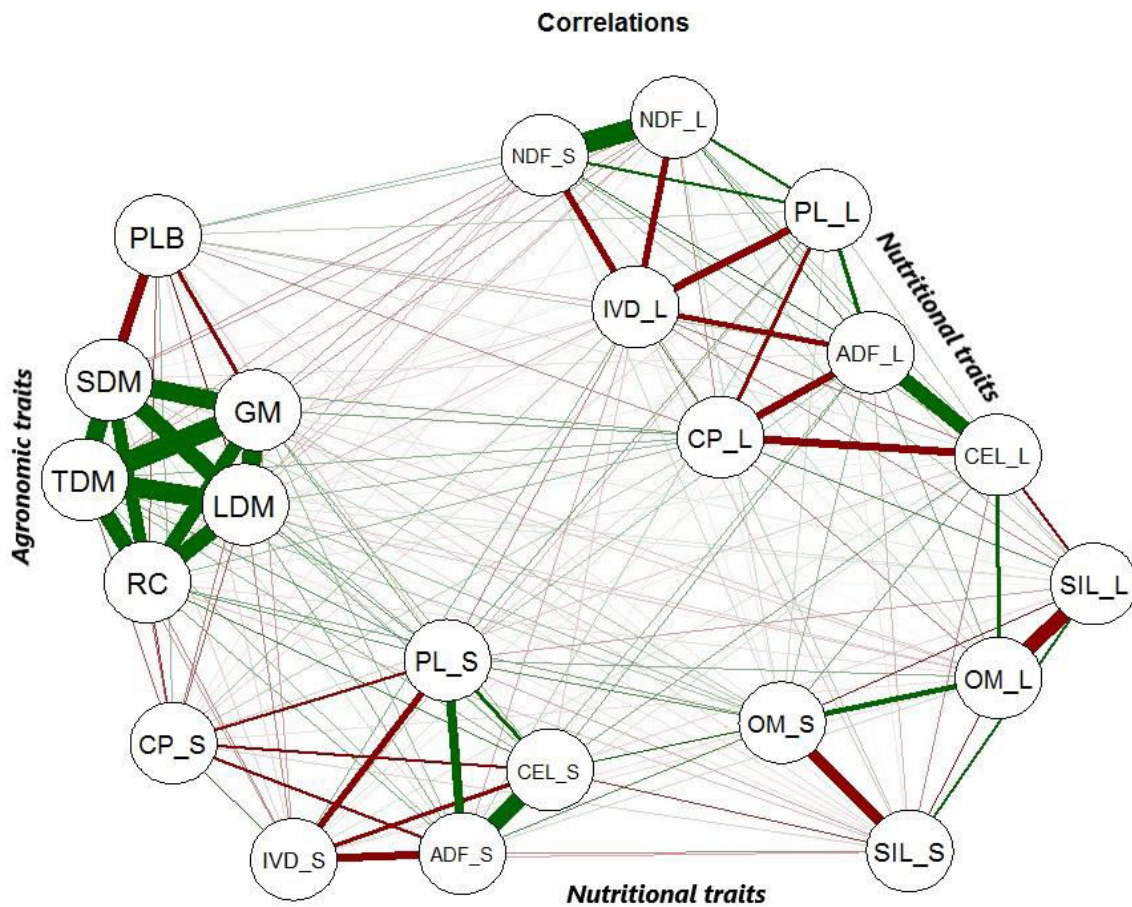


Figure 1. Correlations among all phenotypic traits from the guinea grass mapping population. The green lines correspond to positive correlations between the traits, and the red lines correspond to negative correlations between the traits. The agronomic traits were green matter (GM), total dry matter (TDM), leaf dry matter (LDM), stem dry matter (SDM), percentage of leaf blade (PLB), and regrowth capacity (RC). The leaf and stem nutritional quality traits were the following: organic matter (OM_L and OM_S), crude protein (CP_L and CP_S), neutral detergent fiber (NDF_L and NDF_S), acid detergent fiber (ADF_L and ADF_S), *in vitro* digestibility of organic matter (IVD_L and IVD_S), cellulose (CEL_L and CEL_S), permanganate lignin (PL_L and PL_S), and silica (SIL_L and SIL_S).

Genotyping and linkage analysis

A total of 23,619 SNPs were identified through alignment using the *P. virgatum* genome as the reference, and these presented a mean depth per individual of 247.8 (SD \pm 121.7). After the genotypic analysis, 2,804 SNPs with an average allele depth greater than 60 were selected for further filtering. After allele dosage estimation, 275 SNPs were discarded, in addition to 15 monomorphic markers and four markers with missing data identified manually in R software. Using the updog package, we obtained the allele dosage information for 2,510 SNPs, considering an autotetraploid species (**Table 2**). Simplex (AAAA \times AAAB, ABBB \times

BBBB) was the most commonly found configuration, with 1,654 markers, followed by the duplex (AAAA × AABB, AABB × BBBB) and double-simplex (AAAB × AAAB, ABBB × ABBB) configurations, with 334 and 226 SNPs, respectively. The simplex-duplex (AAAB × AABB) configuration was the most represented among the higher dosages, with 155 markers, and the duplex-simplex (AABB × ABBB) configuration was the least represented, with only six SNPs. Analysis with the polypmapR package revealed that all offspring were 95% compatible with the parents. Two individuals (B107 and C39) were very genetically similar to the parent cv. Mombaça and were removed, one clone (C49) of individual C44 was also removed, and another clone (B127) similar to B126 was also excluded (**Figure 2**). We subsequently discarded another 37 SNPs with missing values and 1,151 duplicated SNPs. This filtering resulted in 132 genotyped offspring with 1,322 markers that were used at the beginning of the linkage analysis in TetraploidSNPMap software. In this step, incompatible markers with the parental allele dosages, SNPs with DR and markers showing SD were not considered. In addition, two-point analysis identified 149 SNPs as duplicated, and MDS analysis identified 27 outliers, which were then excluded. In total, 858 reliable SNPs were included in the linkage map. The apomictic parent, cv. Mombaça, presented 368 exclusive alleles; the sexual parent, S10, presented 275 exclusive alleles; and the two parents shared 215 alleles. TetraploidSNPMap software was used again to rank the 2,510 SNPs based on their expected segregation. The analysis resulted in a total of 114 SNPs with SD, 183 with NP and 243 with DR (**Table 2**).

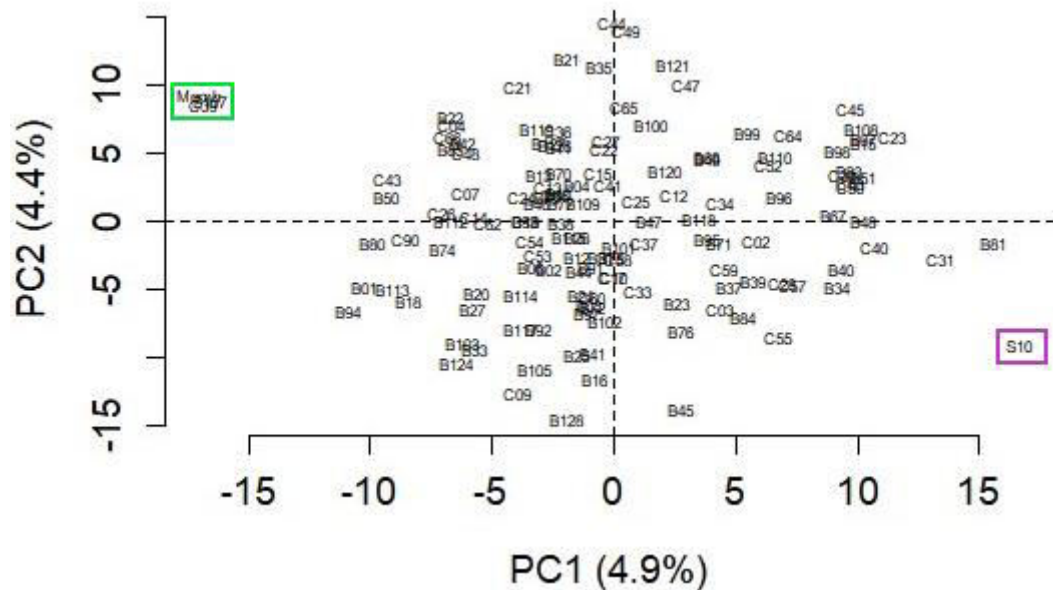


Figure 2. Principal component analysis representing the genetic diversity among the progeny and the parents of the mapping population of guinea grass (*Megathyrsus maximus*). The parents are highlighted in red (S10) and blue (cv. Mombaça).

Table 2. Distribution of SNP markers among genotype classes for a mapping population of guinea grass (*Megathyrsus maximus*).

Marker class	Genotype of the parents	Segregation ratio	Number	DR	NP	Distorted	Mapped
Null	AAAA x BBBB	0	0	0	0	0	0
Simplex	AAAA x AAAB, BBBB x BBBB	1:1	1,654	111	125	18	491
Duplex	AAAA x AABB, AABB x BBBB	1:4:1	334	0	58	40	136
Triplex	AAAA x ABBB, AAAB x BBBB	1:1	35	18	0	0	17
Double-Simplex	AAAB x AAAB, BBBB x ABBB	1:2:1	226	60	0	5	106
X-Double-Simplex	AAAB x ABBB	1:2:1	26	16	0	0	9
Simplex-Duplex	AAAB x AABB	1:5:5:1	155	38	0	9	72
Duplex-Simplex	AABB x ABBB	1:5:5:1	6	0	0	0	1
Double-Duplex	AABB x AABB	1:8:18:8:1	74	0	0	42	26
Total			2510	243	183	114	858

Significant distortion at ($P < 0.001$) and ($P < 0.01$) for simplex and higher-dosage markers. DR corresponds to double reduction and NP are incompatible SNPs with the parental dosages. SNPs with missing data, outliers and duplicates were not considered.

Linkage map

We constructed an integrated consensus linkage map consisting of 858 SNP markers distributed over 756.69 cM in eight HGs, with all possible allele dosage configurations for an autotetraploid species (**Table 2, Figure 3 and Supplementary Table 4**). Considering all integrated consensus HGs, an average density of ~1.13 SNPs/cM was obtained. The largest HG was VII, with 159 SNPs distributed over 108.573 cM, and the smallest HG was VIII, with 49 SNPs present over 70.05 cM. The *interlocus* intervals were relatively small, with a minimum value of 0.003 cM for the majority of the HGs (I, II, IV, VI and VII) and a maximum of 8.65 cM and 7.24 cM on HG V and HG I, respectively (**Table 3**). Among the markers, we identified approximately 30 double-duplex markers (AABB x AABB), which contained all types of doses for autotetraploid progenies (**Table 2, Figure 3 and Supplementary Table 4**) (Hackett et al., 2014).

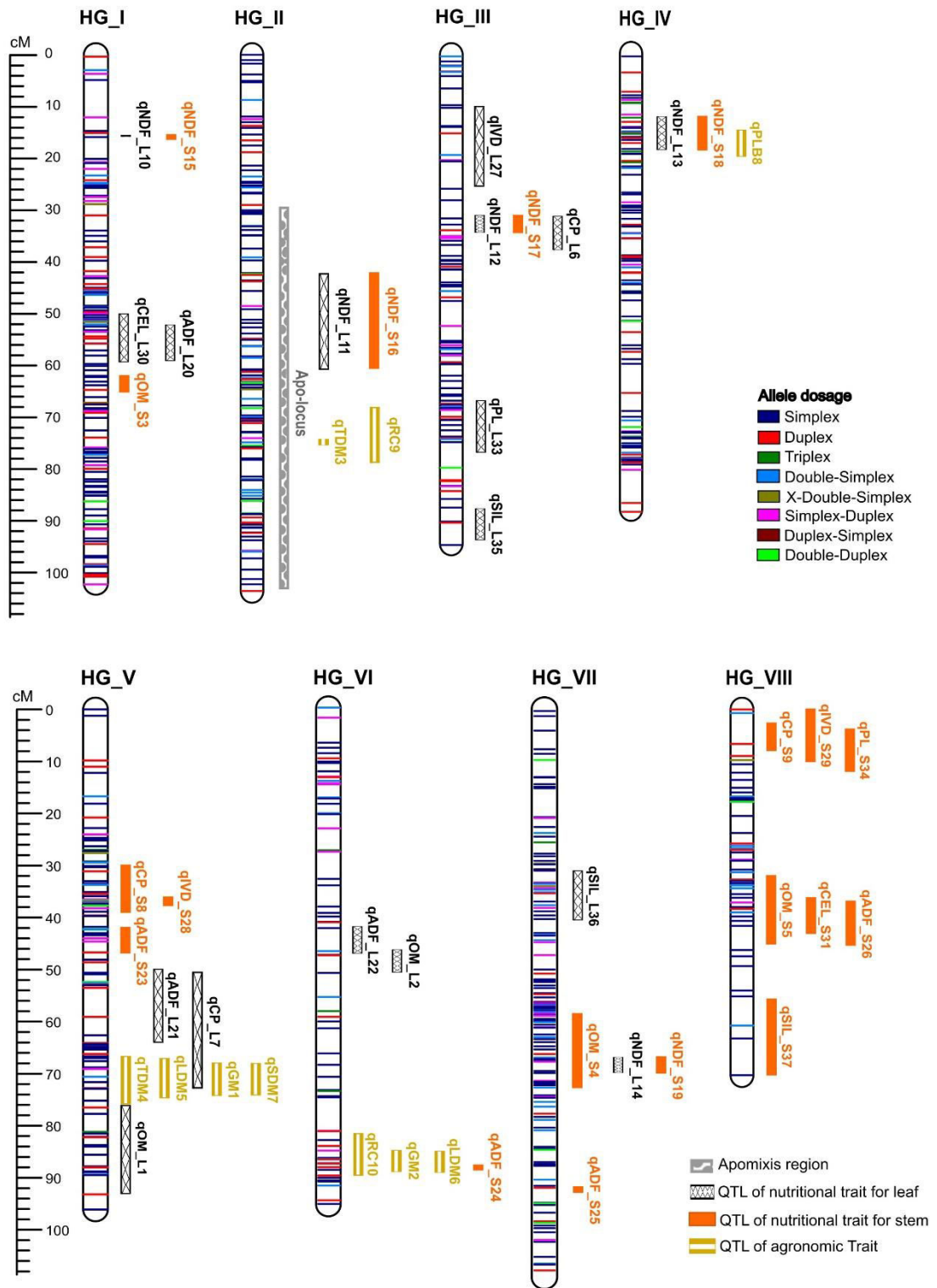


Figure 3. Linkage map constructed for guinea grass (*Megathyrsus maximus*) using SNPs with allele dosage information, including representation of intervals of the highest peaks of QTLs.

Table 3. Summary of the linkage map of guinea grass (*Megathyrsus maximus*) obtained with the S10 × cv. Mombaça population.

HG*	No. mapped SNPs	Length (cM)	Smallest interval (cM)	Longest interval (cM)
I	144	101.74	0.015	7.24
II	128	103.53	0.031	3.43
III	85	94.06	0.027	5.24
IV	115	87.66	0.003	6.34
V	104	96.08	0.008	8.65
VI	74	94.99	0.013	6.42
VII	159	108.57	0.009	5.63
VIII	49	70.05	0.072	6.93
Total	858	756,69		

*Homology group

Apospory mapping

The mode of reproduction of 106 hybrids in the mapping population was determined and indicated that apospory had a segregation ratio of 1:1 based on a chi-square test ($X^2 = 5.43$, $p \geq 0.01$), consistent with the model expected for monogenic inheritance. The *apo-locus* was mapped in the HG II at a peak position of 65 cM, with a high LOD score of 50.06 (**Supplementary Figure 1**). More than 80% of the phenotypic variation in the apomictic reproductive mode was explained, and the simple models best classified the *apo-locus* as a simplex genotype (BBBB × ABBB). As expected, this *locus* was linked to the apomictic parent, cv. Mombaça. Two SNP markers exclusive to this parent, namely, S_14_29023868 and S_10_48091934, were linked to the *apo-locus* at 0.8 cM (**Figure 3**).

Agronomic trait mapping

Ten significant QTLs were mapped for GM, TDM, LDM, SDM, PLB and RC, which were distributed in HGs II, IV, V and VI (**Table 4** and **Supplementary Figure 2**). Here, we will report the QTLs that were identified for the main traits targeted in the *M. maximus* breeding program. TDM was associated with two QTLs in HGs II (qTDM3) and V (qTDM4) with LOD scores of 3.4 and 4.4 and that explained 5.8 and 9.4% of the phenotypic variation,

respectively. For LDM, we detected two QTLs in HGs V (qLDM5) and VI (qLDM6) with LOD scores of 3.9 and 3.0 and that explained 7.9 and 4.5% of the phenotypic variation, respectively. PLB was associated with one QTL in HG IV (qPLB8) with LOD score of 3.8 and that explained 6.8% of the phenotypic variation. We identified two QTLs for RC in HGs II (qRC9) and VI (qRC10) with LOD scores of 4.7 and 3.8 and that explained 10.4% and 6.3% of the phenotypic variation, respectively.

The simple models that best represented the QTLs for each agronomic trait were identified based on the lowest SIC values. The QTL with the greatest effect on TDM, which was present in HG VI, followed a duplex model (AAaa x aaaa) with an additive effect of the A allele from the S10 parent. The same model was verified for LDM with qLDM5, a QTL associated with the S10 parent, and qLDM6, a QTL associated with the cv. Mombaça parent. The QTL with the smallest effect for TDM and the single QTL for PLB followed the double-simplex model (Aaaa x Aaaa) with an additive effect of both parents. The two QTLs detected for RC followed the simplex model, with the S10 parent contributing the allele for qRC9 and the ‘Mombaça’ parent contributing the allele for qRC10.

Table 4. QTLs identified for agronomic traits from the sexual genotype S10 and apomictic cv. Mombaça of guinea grass (*Megathyrsus maximus*).

Agronomic trait	QTL	HG	Position (cM)	LOD	R²	Parents
Green matter (GM)	qGM1	V	71.0	3.56	6.78	both
	qGM2	VI	87.0	2.91	4.29	‘Mombaça’
Total dry matter (TDM)	qTDM3	II	75.0	3.37	5.79	both
	qTDM4	V	70.0	4.41	9.40	S10
Leaf dry matter (LDM)	qLDM5	V	70.0	3.90	7.93	S10
	qLDM6	VI	88.0	2.97	4.53	‘Mombaça’
Stem dry matter (SDM)	qSDM7	V	71.0	3.73	7.24	‘Mombaça’
Percentage of leaf blade (PLB)	qPLB8	IV	17.0	3.81	6.84	both
Regrowth capacity (RC)	qRC9	II	75.0	4.73	10.37	S10
	qRC10	VI	87.0	3.77	6.25	‘Mombaça’

Nutritional trait mapping

Thirty-six significant preliminary QTLs were identified for OM, CP, NDF, ADF, IVD, PL, CEL, and SIL for the leaf and stem. These QTLs were distributed in all HGs (**Table 5** and **Supplementary Figures 3,4**). For CP_L, two QTLs were detected in HGs III (qCP_L6) and V (qCP_L7) with LOD scores of 3.4 and 4.3 and that explained 5.1% and 9.7% of the phenotypic variation, respectively. For CP_S, two QTLs were detected in HGs V (qCP_S8) and VIII (qCP_S9) with LOD scores of 3.8 and 2.8 and that explained 8.4% and 3.2% of the phenotypic variation, respectively. For NDF_L and NDF_S, we detected five QTLs each, all at the same positions in HGs I (qNDF_L10/qNDF_S15), II (qNDF_L11/qNDF_S16), III (qNDF_L12/qNDF_S17), IV (qNDF_L13/qNDF_S18), and VII (qNDF_L14/qNDF_S19). The QTLs with the largest effect on NDF were found in HGs II and IV and were identified with LOD scores of 5.5 and 4.3, respectively; qNDF_L11/qNDF_S16 in HG II explained more phenotypic variation (12%). For the IVD_L trait, only one QTL was found in HG III (qIVD_L27) with LOD score of 3.8 and that explained 5.9% of the phenotypic variation. For IVD_S, two QTLs were obtained in HGs V (qIVD_S28) and VIII (qIVD_S29) with LOD scores of 3.6 and 3.9 and explaining 7.3 and 5.8% of the phenotypic variation, respectively. We detected a single QTL for PL_L located in HG III (qPL_L33) with LOD score of 3.9, and it explained 6.2% of the phenotypic variation. Additionally, PL_S was associated with only one QTL, which was identified in HG VIII (qPL_S34) with LOD score of 2.7 and explained 2.8% of the variation in the trait.

Table 5. QTLs identified for traits related to nutritional quality from the sexual genotype S10 and apomictic cv. Mombaça of guinea grass (*Megathyrsus maximus*).

Nutritional trait	QTL	HG	Position (cM)	LOD	R ²	Parents
Leaf organic matter (OM_L)	qOM_L1	V	87.0	4.15	8.3	S10
	qOM_L2	VI	48.0	3.81	6.7	'Mombaça'
Stem organic matter (OM_S)	qOM_S3	I	63.0	3.85	7.85	S10
	qOM_S4	VII	68.0	4.76	11.44	both
	qOM_S5	VIII	36.0	4.35	7.19	both
Leaf crude protein (CP_L)	qCP_L6	III	34.0	3.43	5.10	both
	qCP_L7	V	65.0	4.32	9.68	both
Stem crude protein (CP_S)	qCP_S8	V	33.0	3.84	8.40	'Mombaça'
	qCP_S9	VIII	6.0	2.84	3.17	S10
Leaf neutral detergent fiber (NDF_L)	qNDF_L10	I	15.0	3.64	6.09	both
	qNDF_L11	II	48.0	5.49	12.08	both
	qNDF_L12	III	33.0	3.51	6.83	both
	qNDF_L13	IV	14.0	4.30	6.84	both
	qNDF_L14	VII	69.0	3.74	7.81	'Mombaça'
Stem neutral detergent fiber (NDF_S)	qNDF_S15	I	15.0	3.61	6.01	both
	qNDF_S16	II	48.0	5.50	12.07	both
	qNDF_S17	III	33.0	3.5	6.81	both
	qNDF_S18	IV	14.0	4.31	6.87	both
	qNDF_S19	VII	69.0	3.75	7.85	'Mombaça'
Leaf acid detergent fiber (ADF_L)	qADF_L20	I	57.0	3.98	8.84	S10
	qADF_L21	V	59.0	4.0	8.07	both
	qADF_L22	VI	45.0	3.08	2.46	both
Stem acid detergent fiber (ADF_S)	qADF_S23	V	45.0	3.48	6.21	'Mombaça'
	qADF_S24	VI	88.0	2.86	4.05	'Mombaça'

Table 5. Continued

Nutritional trait	QTL	HG	Position (cM)	LOD	R²	Parents
	qADF_S25	VII	93.0	3.41	5.90	Both
	qADF_S26	VIII	41.0	3.04	4.05	S10
Leaf <i>in vitro</i> digestibility of organic matter (IVD_L)	qIVD_L27	III	20.0	3.8	5.85	S10
Stem <i>in vitro</i> digestibility of organic matter (IVD_S)	qIVD_S28	V	37.0	3.55	7.31	‘Mombaça’
	qIVD_S29	VIII	6.0	3.91	5.84	
Leaf cellulose (CEL_L)	qCEL_L30	I	55.0	4.69	10.85	both
Stem cellulose (CEL_S)	qCEL_S31	VIII	37.0	3.02	3.37	both
Leaf permanganate lignin (PL_L)	qPL_L33	III	74.0	3.86	6.18	‘Mombaça’
Stem permanganate lignin (PL_S)	qPL_S34	VIII	7.0	2.7	2.78	both
Leaf silica (SIL_L)	qSIL_L35	III	94.0	3.70	5.68	‘Mombaça’
	qSIL_LF36	VII	36.0	4.63	10.63	S10
Stem silica (SIL_S)	qSIL_S37	VIII	61.0	2.96	4.23	both

According to the simple model analysis, the best model for the two QTLs of CP_L was the double-simplex model (Aaaa x Aaaa) with an additive effect of both parents. In contrast, qCP_S8 was best represented by a duplex model (AAaa x aaaa) with a dominant effect of cv. Mombaça. The CP_S9 QTL was represented by a double-simplex model with an additive effect of the S10 parent. The QTL with the greatest effect on NDF was best represented by the double-simplex model with an additive effect from both parents. For IVD_L, a simplex allele was verified in the sexual parent (Aaaa x aaaa). The same model was also observed for IVD_S and PL_L, with the A allele from the apomictic parent associated with the trait. The single QTL for PL_S was explained by a double-simplex model with a dominant effect from both parents.

Search for similarity in aposporic and QTL regions

Some genes that contain conserved protein domains were found in regions flanking the apo-*locus* located in HG II, such as the Spc97/Spc98 family of spindle pole body (SBP)

components, whose proteins assist in the control of the microtubule network (Lin et al., 2015), and the inner centromere protein (ARK-binding region). These two genes play an important role in the segregation of chromosomes during cell division (Kirioukhova et al., 2011; Lin et al., 2015). In addition, a gene possibly involved in the aposporic reproductive mode was similar to somatic embryogenesis receptor-like kinase 1 (SERK1), which is part of a complex associated with the induction of embryo development (Albrecht et al., 2008).

A total of 23 regions were found due to the overlap of QTLs in common regions in the linkage map. In these regions, 55 different gene families from *P. virgatum*, *A. thaliana* and *O. sativa* were identified. Most of these genes may play important roles in cellular metabolism and may be associated with plant growth and development (**Table 6**). Further details about the locations of the candidate genes in their respective QTL regions are provided in **Supplementary Table 5**.

Among the putative genes, many participate in hormonal signaling pathways, such as auxin efflux carrier component in HG I - region 2 (qADF_L20/qCEL_L30), which is related to the intercellular directionality of auxin (Friml et al., 2002); RING/U-box (XERICO) in HG II – region 5 (qNDF_L11/ qNDF_S16), which is involved in the regulation of abscisic acid (ABA) (Ko et al., 2006); gibberellin-regulated family members (GAST1, GASR2, GASR3, and GASR9) in HG III – region 9 (qPL_L33), which are associated with the regulation of gibberellic acid (GA) and ABA (Hedden and Thomas, 2016); ent-copalyl diphosphate synthase in HG IV – region 11 (qNDF_L13/qNDF_S18/qPLB8), which plays a role in GA biosynthesis (Koksal et al., 2011); and ethylene insensitive 3 (EIN3) in HG VIII – region 21 (qCP_S9/qIVD_S29/qPL_S34), which is associated with cell growth and senescence processes (Munné-Bosch et al., 2018) (**Table 6** and **Supplementary Table 5**).

In addition, other genes with a role in plant physiology were also identified, such as the gene encoding the enzyme phosphoenolpyruvate carboxylase (PEPC) in HG I – region 2 (qADF_L20/qCEL_L30), whose function is the catalysis of primary metabolic reactions in plants (Toledo-Silva et al., 2013). Glycosyltransferase was found in HG III – region 7 (qIVD_L27) and is associated with the biosynthesis of polysaccharides and glycoproteins (Hansen et al., 2012). Pectinesterase/pectin methylesterase, which was detected in HG V – region 14 (qGM1/qTDM4/qLDM5/qSDM7/qADF_L21/qCP_L7) and HG VIII – region 23 (qSIL_S37) is related to cellular adhesion and stem elongation (Damm et al., 2016). In addition to these genes, we found a chloroplast envelope transporter in HG VII – region 19

(qOM_S4/qNDF_L14/qNDF_S19). Phosphatidylinositol-4-phosphate 5-kinase, which was present in HG VIII – region 22 (qOM_S5/qADF_S26/qCEL_S31), is involved in coordinating plant growth (Kusano et al., 2008).

Many putative candidate genes present in lignan biosynthesis pathways were verified in some regions of QTLs. Glycosyl hydrolase family 16 was located in HG III - region 8 (qCP_L6/qNDF_L12/qNDF_S17), and the genes associated with secoisolariciresinol dehydrogenase and the NAD(P)-binding Rossmann-fold superfamily were located in region 9 (qPL_L33). UDP-Glycosyltransferase superfamily protein was present in HG V – region 13 (qADF_S23). The 4-coumarate:CoA ligase 3 gene was obtained in HG VI - region 16 (qOM_L2/qADF_L22). The gene encoding NAD(P)-binding Rossmann-fold superfamily was also detected in HG VIII – region 21 (qCP_S9/qIVD_S29/ qPL_S34).

Table 6. Description and function of the genes identified in apomixis and QTL regions from linkage map of guinea grass.

HG	Region	QTL	Gene Description	Function	Reference
I	1	qNDF_L10 qNDF_S15	Rhomboid family protein	Root growth, floral development and fertility	Knopf et al. (2012)
			Phosphoenolpyruvate carboxylase	CO ₂ fixation in the cytoplasm	Toledo-Silva et al. (2013)
I	2	qADF_L20 qCEL_L30	Auxin efflux carrier family protein	Regulator of auxin efflux, differential growth and tropism	Friml et al. (2002)
			Inorganic H pyrophosphatase family protein	Regulation of plant proton-pumping homeostasis	Primo et al. (2019)
			Vacuolar ATP synthase subunit A	Male gametophyte development and Golgi organization	Dettmer et al. (2005)
			Palmitoyltransferase TIP1	Stem cell or root hair and pollen tube growth	Hemsley et al. (2005)
I	3	qOM_S3	PINHEAD	Regulation of cell division and axis determinacy	Newman et al. (2002)
			Plant stearyl-acyl-carrier-protein desaturase family protein	Regulation of oleic acid	Kachroo et al. (2007)
II	4	<i>Apo-locus</i>	Exocyst Complex Protein Exo70	Plant cell morphogenesis	Hála et al. (2008)
			Inner centromere protein (ARK binding region)	Regulation of egg and central cell fate and differentiation	Kirioukhova et al. (2011)
II	4	<i>Apo-locus</i>	Somatic embryogenesis receptor-like kinase 1 (SERK)	Induction of somatic embryogenesis	Albrecht et al. (2008)
			Spc97 / Spc98 family of spindle pole body (SBP) component	Regulation of microtubule network	Lin et al. (2015)

Table 6. Continued

HG	Region	QTL	Gene Description	Function	Reference
II	5	qNDF_L11 qNDF_S16	RING/U-box domain-containing protein (XERICO)	Regulation of abscisic acid (ABA)	Ko et al. (2006)
			Wall-associated receptor kinase galacturonan-binding (GUB_WAK_bind)	Pathogen response and cell expansion	Kohorn et al. (2012)
			Late embryogenesis abundant (LEA) protein-related	Seed maturation and tolerance to abiotic stress in plants	Zhang et al. (2014)
II	6	qTDM3	WRKY family transcription factor family protein	Pathogen defense, senescence and trichome development	Eulgem et al. (2000)
		qRC9	Eukaryotic elongation factor 5A-1	Regulation of cell division, cell growth, and cell death	Feng et al. (2007)
III	7	qIVD_L27	Glycosyltransferase	Biosynthesis of polysaccharides and glycoproteins	Hansen et al. (2012)
III	8	qCP_L6	Xyloglucan:xyloglucosyl transferase	Integral plasma membrane protein and wall-loosening factor	Ndamukong et al. (2009)
		qNDF_L12 qNDF_S17	Xyloglucan endotransglucosylase/hydrolase	Cell wall construction	Yokoyama et al. (2001)
III	9	qPL_L33	Protein thylakoid formation 1, chloroplastic	Vesicle-mediated thylakoid membrane biogenesis	Wang et al. (2004)
			Secoisolariciresinol dehydrogenase	Lignin biosynthesis	Davin & Lewis (2003)
			Gibberellin-regulated family protein	Reproductive development and regulation of growth	Hedden & Thomas (2016)
			NAD(P)-binding Rossmann-fold superfamily protein	Cinnamoyl-CoA reductase activity (lignin biosynthesis)	Pan et al. (2014)

Table 6. Continued

HG	Region	QTL	Gene Description	Function	Reference
III	10	qSIL_L35	Disease resistance protein RPS2 Stripe rust resistance protein	Immune signaling in response to pathogenic fungi Immune signaling in response to root-knot nematodes	Kourelis & van der Hoorn (2018) Dimkpa et al. (2016)
IV	11	qPLB8	Ent-copalyl diphosphate synthase	Gibberellin biosynthesis	Koksal et al. (2011)
		qNDF_L13 qNDF_S18	Phototropic-responsive NPH3 family protein	Phototropic signal response pathway	Roberts et al. (2011)
V	12	qCP_S8 qIVD_S28	Glucuronokinase	UDP-glucuronic acid synthesis (sugar metabolism)	Xiao et al. (2017)
			WRKY DNA-binding domain	Pathogen defense, senescence and trichome development	Eulgem et al. (2000)
			P-glycoprotein	Auxin transport	Geisler & Murphy (2006)
			Protein NRT1/ PTR family 6.2	Transporting different substrates, e.g., nitrate	Corratgé-Faillie & Lacombe (2017)
			Methyl esterase	Hydrolysis of auxin and jasmonic acid	Yang et al. (2008)
	13	qADF_S23	UDP-Glycosyltransferase superfamily protein	Mechanism of normal cell wall lignification	Lin et al. (2016)
V	14	qGM1 qTDM4 qLDM5 qSDM7 qCP_L7 qADF_L21	Early nodulin-like family protein	Cell differentiation and cell wall reorganization during nodulation	Mashiguchi et al. (2009)
			HEXOKINASE	Glucose regulation	Aguilera-Alvarado & Sánchez-Nieto (2017)
			Pectinesterase / Pectin methylesterase	Pectin structure	Pelloux et al. (2007)
			NAD-dependent malic enzyme_mitochondrial precursor	Metabolism in mitochondria	Tronconi et al. (2008)
			NADP-dependent malic enzyme_chloroplast precursor	Metabolism in chloroplasts (C4 plants) or in cytosol	Tronconi et al. (2008)

Table 6. Continued

HG	Region	QTL	Gene Description	Function	Reference
V	15	qOM_L1	Calcium-binding EF-hand family protein	Regulation of cellular and developmental processes	Day et al. (2002)
VI	16	qOM_L2 qADF_L22	4-coumarate:CoA ligase 3	Lignin biosynthesis	Toledo-Silva et al. (2013)
VI	17	qGM2	Galactinol--raffinose galactosyltransferase / Stachyose synthetase	Desiccation protectant in seeds and transporter of sugar in phloem sap	Sengupta et al. (2015)
		qLDM6 qRC10	Nodulin-like / Major Facilitator Superfamily protein	Facilitators of water and ammonia transport	Routray et al. (2015)
		qADF_S24	Peroxidase superfamily protein	Lignin biosynthesis	Toledo-Silva et al. (2013)
VII	18	qSIL_L36	FKBP-type peptidyl-prolyl cis-trans isomerase family protein	Control of cell proliferation and differentiation	Smyczynski et al. (2006)
VII	19	qOM_S4	Chloroplast envelope transporter	Ions transport	Hohner et al. (2016)
		qNDF_L14 qNDF_S19	Endo-1,4-beta-xylanase/ Glycosyl hydrolase family 10 protein	Xylan degradation	Suzuki et al. (2002)
			Alpha-galactosidase	Regulation of cell wall loosening and cell wall expansion	Chrost et al. (2007)
VII	20	qADF_S25	Lipoxygenase	Biosynthesis of polyunsaturated fatty acids	Bannenberget al. (2008)

Table 6. Continued

HG	Region	QTL	Gene Description	Function	Reference
VIII	21	qCP_S9 qIVD_S29 qPL_S34	3-oxoacyl-[acyl-carrier-protein] reductase	Fatty acid biosynthesis	Ding et al. (2015)
			Sugar transporter/spinster transmembrane protein	Transport of lipidic molecules	Niño-González et al. (2019)
			Ethylene insensitive 3 family protein	Modulation of plant growth	Munné-Bosch et al. (2018)
			Glycine-rich cell wall structural transmembrane protein	Component of the cell walls of higher plants	Mousavi & Hotta (2005)
			NAD(P)-binding Rossmann-fold superfamily protein	Cinnamoyl-CoA reductase activity (lignin biosynthesis)	Pan et al. (2014)
VIII	22	qOM_S5	Sterol regulatory element-binding protein	Regulation of sterol biosynthesis	Seo et al. (2008)
		qADF_S26	Alpha carbonic anhydrase	C4 photosynthetic pathway	Toledo-Silva et al. (2013)
		qCEL_S31	Phosphatidylinositol-4-phosphate 5-kinase	Regulator of root hair tip growth	Kusano et al. (2008)
VIII	23	qSIL_S37	Pectinesterase	Pectin structure	Pelloux et al. (2007)

Discussion

Linkage map

We obtained a satisfactory representative linkage map for *M. maximus*, with a large genetic distance observed between the parents, namely, cv. Mombaça (**Figure 2**, quadrant I) and S10 (**Figure 2**, quadrant IV). This genetic distance and the distribution of hybrids can be visualized in the PCA performed with the allele dosage information of all individuals (**Figure 2**). The crossing of contrasting parents makes it possible to check the recombination frequency between *loci*, which is a fundamental principle for observing the segregation of traits in hybrids and promoting the detection of QTLs (Luo et al., 2000).

Because *M. maximus* does not yet have a sequenced genome, we aligned our GBS reads to the allotetraploid genome of *P. virgatum*, which is a species closely phylogenetically related to *M. maximus* (Burke et al., 2016), since *M. maximus* previously belonged to the *Panicum* genus. This genetic proximity was confirmed based on the consistent distribution of the markers in the genetic map (**Figure 3** and **Supplementary Table 4**). In addition, considering our tetraploid mapping population, the use of a tetraploid genome as reference might be more informative than the use of diploid genomes because there is a greater possibility of similar chromosomal rearrangements between these species (Daverdin et al., 2015). However, the possibility of using the genome of *P. virgatum* as reference does not preclude the need for a sequenced genome of *M. maximus*, mainly due to the possibility of identifying exclusive markers of the species, and our linkage map can contribute to the assembly of this reference genome.

After allele depth estimation, only 10% of markers were retained for the next analysis, in which we prioritized a minimum average allele depth of 60 reads to suppress the probable overestimation of allele bias due to our population size. Gerard et al. (2018) recommend that more than 25 reads be used to obtain a strong correlation of true genotypes under high levels of bias and overdispersion, emphasizing that read depth requirements should be based on how many individuals are included in the study. Of the genetic mapping studies involving tropical forage grasses, only one reported the use of SNP markers with allele dosage, in which a minimum overall depth of 25 reads was considered in a biparental progeny of *U. decumbens* containing 217 F₁ hybrids (Ferreira et al., 2019).

Our linkage map is the second for *M. maximus* but the first to be integrated from both parents and containing SNP markers in multiple doses. Updated statistical models and the recent genome sequencing technologies promoted advances in the knowledge of the genetics of polyploid organisms, as demonstrated by the comparison of our genetic map for *M. maximus* with the first map for this species published over ten years ago by Ebina et al. (2005). The first map contained 360 dominant markers obtained with amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) techniques that segregated at a 1:1 ratio and was obtained from an apomictic cultivar used for genetic breeding in Japan. These markers were distributed in 39 linkage groups, which is greater than the number expected for this species ($2n = 4x = 32$); therefore, this map has a limited contribution. In this context, the use of SNPs as codominant markers and their quantitative analysis allowed many gains in our map, being the most real and closest to what we have of the genome of this species; such as coverage of many regions of quantitative traits important to breeders and the genetic effects that influence these traits, as well as alleles of the parents that were determinant of the characters of the hybrid.

In addition, the strategy of greater refinement of SNPs adopted after estimating allele dosage using two programs of mapping prevented overinflation among *loci*. We detected some markers with SD (**Table 2**), but we chose not to use them, aiming to increase the probability of obtaining an exact distance between markers in the HGs. Since genotyping errors are probable, a large amount of missing data and a large number of distorted *loci* may promote the expansion of linkage maps as well as overestimate the recombination fractions and limit the accuracy of the mapping (Cartwright et al., 2007; Gerard et al., 2018). However, SD is a phenomenon commonly found in the genome, and some linkage maps contain distorted markers, including those of grasses such as pearl millet (Sehgal et al., 2012) and napiergrass (Paudel et al., 2018). Therefore, greater knowledge of the occurrence and genetic causes of SD in plants is important for inferring which genes are kept together or separated by SD (Zhu et al., 2006; Anhalt et al., 2008). Thus, future studies might aggregate information on the structure of *loci* with SD in the genome of *M. maximus*.

Most linkage maps in grasses include only single-dose markers with segregation ratios of 1:1 and 3:1; thus, the heterozygous classes were grouped with one of the homozygous classes, resulting in a loss of information (Bourke et al., 2018b). The use of linkage maps with simplex SNPs for most linkage groups and a high density of higher-dosage markers provides

greater confidence in the modeling of allelic effects of QTLs (Hackett et al., 2014). The markers present in our map were mostly simplex, while higher-dosage markers comprised approximately 28.4% of the map, as shown in **Table 2**. Similar values were also verified in the genetic map of *U. decumbens* (Ferreira et al., 2019), probably due to the complexity of scoring and analyzing these types of markers. Our map presented a greater number of alleles exclusive to the apomictic parent than sexual parent, as observed in the previous map of *M. maximus* (Ebina et al., 2005) and in the genetic maps of other forage grasses (Worthington et al., 2016; Ferreira et al., 2019). This difference is likely due to the origin of the genotypes: ‘Mombaça’ and S10 have natural tetraploid genomes, but S10 was obtained from a sexual x apomictic cross of an original diploid sexual plant that was duplicated with colchicine.

Our linkage map was sufficient for the identification of oligogenic trait with a genomic region responsible for 80% of the phenotypic variance and polygenic traits. Therefore, our map will also be useful for the detection of other important characteristics in *M. maximus* and can contribute to the assembly of the genome of this species, as well as to studies about the biology and evolution of other phylogenetically closely related tropical forage grasses, such as those in the *Urochloa* and *Paspalum* genera.

Double reduction in guinea grass

We reported for the first time the occurrence of DR in *M. maximus*. Autotetraploids may undergo this type of segregation when in multivalent pairing, two pairs of chromatids pass to the same pole in anaphase I of meiosis (Haynes and Douches, 1993). In guinea grass, irregular chromosome segregation has already been verified in cytogenetic analysis in hybrids and parents from the breeding program of the Embrapa Beef Cattle (Pessim et al., 2010; Pessim et al., 2015). The distribution of the markers with DR between the dosage types was proportional to the number of SNPs with each configuration. DR has been extensively studied using SNPs with dosage data in autotetraploid linkage maps in potato (Hackett et al., 2017; Bourke et al., 2018a). Approximately 6% of markers in potato have DR (Bourke et al., 2015), corroborating our results (9.68%). Despite some studies suggesting that DR should be included in genetic map construction and in QTL analysis (Li et al., 2010), other studies verified that such markers exert only minor positive effects on the power and accuracy of mapping analysis using single-dose markers (Bourke et al., 2015) and higher-dose markers (Bourke et al., 2016). In addition, statistical models have been created to include DR in linkage mapping (Huang et al., 2019), but no software currently implements them.

The occurrence of DR in *M. maximus* has many implications for breeding programs, being that the effects of DR and how to handle them have been the targets of several studies (Luo et al., 2000; Xu et al., 2013; Layman and Busch, 2018). This type of segregation exposes alleles located in distal regions of the chromosomes to homozygosis and thus is effective in eliminating the lethal alleles in a population (Butruille and Boiteux, 2000). A low rate of DR is sufficient to considerably reduce the equilibrium frequency of a deleterious allele at one *locus* (Luo et al., 2006). As an alternative, DR could be used to accelerate the accumulation of favorable rare alleles through marker-assisted selection (MAS) (Bourke et al., 2015). In addition, it is possible to obtain genotypes with *loci* having a higher homozygosis rate for use in specific crosses (Bourke et al., 2015). In this context, more detailed molecular study could elucidate the influence of DR on the phenotypes of hybrids of our study species.

Apospory mapping and the search for gene similarity

A chi-square test ($X^2 = 5.43$, $p \geq 0.01$) performed for qualitative analysis of the reproductive mode of the 106 hybrids followed the Mendelian inheritance model, corroborating the results obtained with progeny tests in guinea grass performed by Savidan (1980) and Savidan (1981), in which sexual x apomictic progenies exhibited a 1:1 ratio that could be explained by an Aaaa genotype for apomictic because the aposporous apomixis of *M. maximus* is dominant over sexuality. We also proved this finding through SNP markers. Other studies of grasses such as *Pennisetum* (Akiyama et al., 2011), *Paspalum* (Martínez et al., 2003) and *Urochloa* (Valle et al., 1994; Vigna et al., 2016) also verified this segregation for the reproductive mode. Evidence suggests that this *locus* is present in a conserved region of the plant genome; however, further molecular genomic studies on aposporous apomixis in forage grasses are needed because the recent studies have led to other hypotheses, such as a possible influence of epigenetics (Kumar, 2017). Interestingly, more advanced studies on apomixis in *M. maximus* genotypes from a germplasm bank of India reported the decoupling of apomixis into three components, namely, apospory, parthenogenesis and pseudogamous-endosperm. These components are three distinct genetic determinants that determine an individual as apomictic, and recombination might occur among these components, supporting the hypothesis that more major genes are involved in apomixis control (Kaushal et al., 2008; Kaushal et al., 2019). This decoupling in neo-apomictics might be the target of an adaptive mechanism to maintain variability through hybridization and could be beneficial to the breeding program (Kaushal et al., 2019).

The aposporous region in *M. maximus* was previously mapped (Ebina et al., 2005; Bluma-Marques et al., 2014), and similar to our results, no markers were in perfect linkage with the region. Nonetheless, we mapped markers at a shorter distance (0.8 cM) from the apo-*locus* (**Figure 3**). Genetic markers linked to apomixis have been sought in other tropical forage grasses (Vigna et al., 2016; Worthington et al., 2016, 2019), aiming at the efficient and rapid identification of the reproductive mode of progenies. Once identified, such markers may be transferred among forage grasses, based on evidence of conservation of the ASGR. Thus, markers near the apospory region that were identified in our map may be validated and useful for the breeding program of this species.

In addition, we observed that the markers closer to the peak region of the apo-*locus* were in a genomic region of *P. virgatum* similar to a region of the genome of *A. thaliana* that contains the SERK1 gene. This gene is involved in the signaling pathway active during zygotic and somatic embryogenesis in *A. thaliana*, and its overexpression increases the efficiency of somatic embryogenesis initiation (Hecht et al., 2001). In nucellar cells of apomictic genotypes of *Poa pratensis*, the SERK gene is involved in embryo sac development (Albertini et al., 2005). Recently, SERK was reported in *Brachypodium distachyon* grass as having a domain conserved among monocots and plays a prominent role in apomixis (Oliveira et al., 2017). Scarce studies have investigated the genes involved in the regulation of reproductive events in guinea grass, and these previous studies mainly utilized of transcriptome data (Toledo-Silva et al., 2013; Radhakrishna et al., 2018).

QTLs for agronomic and nutritional traits

Traits related to the productivity and quality of forage used for cattle fattening are categorized as complex traits, which are determined by both genetic and environmental factors. The transgressive segregation observed in these traits in *M. maximus* suggests quantitative and polygenic inheritance, consistent with the inheritance of non-Mendelian traits (Miles and Wayne, 2008; Holzman and Hulsey, 2017). A QTL mapping approach is required for the characterization of the genetic architecture of traits. In our study, QTL analysis of autotetraploid progeny was performed using interval mapping (IM) of markers with allele dosage. This same methodology was successfully applied in QTL mapping in signalgrass, another important forage grass (Ferreira et al., 2019). The multiple interval mapping (MIM) method was recently implemented in polyploids and is a new alternative for other mapping studies using data with allele dosage information (Pereira et al., 2019). The mapping method used here considered only

the peak with the largest effect as a QTL, but it is worth mentioning that peaks were present near the peak QTL for all agronomic traits (**Supplementary Figure 2**).

QTLs associated with important agronomic traits were mapped in HGs I, III, VII and VIII, with the phenotypic variation explained ranging from 4.3% to 10.4% (**Table 4**). Because *M. maximus* is undergoing a domestication process, the crop can still be greatly improved by the selection of large-effect QTLs. Therefore, QTL qRC9, located at 75 cM in HG II, which explained 10.3% of the phenotypic variation in RC (**Supplementary Figure 2**, HG_2(B)) and had a predominant additive effect of the female progenitor S10, may be a candidate for the marker-assisted selection program of guinea grass.

We found more than one QTL for TDM, LDM and RC, again supporting the hypothesis of complex genetic control. Conversely, we found only one QTL for PLB in both parents. TDM and LDM showed high broad-sense heritability (< 0.5), followed by RC (0.3) and PLB (0.1), as shown in **Table 1**. Higher heritability values (< 0.85) for LDM and RC and a value for PLB above 0.4 were recently reported in *M. maximus* (Lara et al., 2019), using the generalized heritability formula (Cullis et al., 2006). For this species, a greater amount (g/plant) of TDM and LDM in the progenies has been associated with considerable heritability from the most productive parents (Braz et al., 2013; Braz et al., 2017). Matias et al. (2019) also verified the same pattern in interspecific hybrids from *Urochloa* spp., another genus adapted to tropical conditions. The intermediate heritability of RC and the QTLs associated with this trait that were detected in both parents resulted in hybrids with a good capacity for regrowth. This trait is also considered fundamental in forage grasses because it is directly related to the persistence of the forage after defoliation (Jank et al., 2011).

The negative correlation between PLB and SDM was expected (**Figure 1**), and Braz et al. (2017) verified the higher PLB values in the experiment of this progeny. PLB is related to plant structure, and plants with a high percentage of leaves are desirable because this trait is related to higher forage quality. A higher PLB was observed in the male parent, cv. Mombaça, which is often used as a check in experiments. Since its release in the 1990s, along with cv. Tanzania, cv. Mombaça has promoted pasture intensification in the country due to its very high productivity and forage quality (Jank et al., 2014). Progenies whose female parent is S10 generally also present good yield (Resende et al., 2004). Breeding programs target these traits in search of superior genotypes with greater foliar mass and a higher percentage of leaves due to the higher digestibility of leaves than of stems for animals. Thus, forage breeding is not

restricted to the obtaining of more productive plants; it also contributes to greater efficiency in their transformation into animal production (Valle et al., 2009).

Significant and positive correlations among the traits GM, TDM, SDM, LDM, and RC corroborated the positions of QTLs associated with agronomic traits in the linkage map (**Figures 1** and **3**). The traits GM and LDM exhibited high heritability (**Table 1**) and strong positive correlation (**Figure 1**), which suggests that genetic improvements in these traits could be achieved by selective phenotyping for only one of these traits, without the need to do separation of plant part. We detected qTDM3 and qRC9 in a common region in HG II; qGM1, qTDM4, qLDM5, and qSDM7 in the same region in HG V; and qGM2, qLDM6, and qRC10 in part of a common region in HG VI. Each region containing several QTLs for different traits suggests the occurrence of four QTL hotspots. Interestingly, QTLs for PLB were not detected in any region with other agronomic traits, and a negative correlation has also been observed; however, a positive correlation with NDF was verified, and such a QTL was detected in a similar region with qNDF_L13 and qNDF_S18 in HG IV (**Figure 3**), suggesting a fifth QTL hotspot.

Clustering of QTLs for genetically correlated traits in the same or adjacent regions of HGs in several organisms may be due to physical linkage, pleiotropy or natural selection for coadapted traits (Studer and Doebley, 2011; Wu et al., 2015). We have taken the first step in the identification of *loci* that lead to these trait correlations and the degree to which these patterns affect productivity in *M. maximus*. QTLs co-located in the same region of HGs for agronomic traits have also been identified in some grasses (Fang et al., 2016; Sartie et al., 2018).

Pleiotropy occurs when a gene influences multiple traits; when agronomically important traits are positively associated, more than one trait can be improved simultaneously (Kumar et al., 2017). There are some reports of QTLs with pleiotropic effects in some grasses, such as wheat (Deng et al., 2011), *Setaria* spp. (Mauro-Herrera and Doust, 2016), and a perennial grass (Khasanova et al., 2019).

Linked QTLs for natural selection of coadapted traits can occur in two forms. First, these QTLs can be linked in the coupling phase on the chromosome, i.e., two favorable alleles of both genes are from the same parent, which is useful in breeding programs because can consider two traits as one target. Second, these QTLs are linked in the repulsion phase on the chromosome, where the two alleles of both genes are on opposite chromosomes, possibly due

to favorable alleles of one gene being from one parent and the favorable alleles of another gene being from another parent. In this second case, linkage drag may occur, which needs to be broken for the breeding program (Wu et al., 2015).

All HGs contained QTLs related to nutritional traits, with those related to the leaf and stem being found mainly in HG III and HG VIII, respectively. Again, some traits had more than one peak that could not be considered due to the methodology adopted (**Supplementary Figures 3 and 4**). The phenotypic variance explained by these QTLs ranged between 2.5% (qADF_L22) and 12.1% (qNDF_L11 and qNDF_S16). Interestingly, both parents contributed alleles for most of the identified QTLs, providing evidence that the genotypes have high nutritional quality.

The selection of a superior genotype of perennial tropical forages is based on the analysis of traits through repeated measures over a number of harvests, seasons, and years (Fernandes et al., 2017). The nutritional QTLs found in our study need further investigation because only one harvest was analyzed and interactions could not be measured. In a selection study of *U. decumbens* and *U. humidicola*, Figueiredo et al. (2019) identified a significant Genotype x Harvest Interaction (GHI) effect ($p < 0.05$) for different agronomic and nutritional traits, which reflect differences in the relative performances of genotypes across harvests. GHI also has been reported as significant in *M. maximus*, and the average number of harvests needed for a reliable selection of nutritional traits, such as OM and CP, would be 3 and 5 at accuracy levels of 0.80 and 0.85, respectively (Fernandes et al., 2017). Thus, GHI directly affects the selection of the genotypes and further studies with more harvests are needed to consolidate our results about putative QTLs to nutritional traits.

The heritability of the nutritional traits varied from low (0.06), obtained for SIL_S, to moderate (0.32), obtained for OM_L. This result also reinforces the necessity of more harvests for the efficient selection of superior genotypes for nutritional traits (Fernandes et al., 2017). Traits IVD_L and IVD_S presented the same standard as the crude protein, with the H^2 of the stem being higher than that of the leaf (**Table 1**). Historically, cv. Mombaça has stood out due to its high productivity, but with slightly lower values for forage quality when compared to 'Tanzania', and in biparental crosses, the hybrids obtained from 'Mombaça' also presented these features (Braz et al., 2017). This finding was corroborated by the heritability verified above in our study of agronomic and nutritional traits.

The nutritional quality of forage grasses is important in several aspects and is directly related to the production of meat and milk. *M. maximus* shows a high value of CP compared to other tropical forages. However, the values for lignin and fiber are expected to be low because lignin hampers the enzymatic hydrolysis of cellulose and hemicellulose and, thus, the digestion of the cell wall of the leaf tissue and the stem (Jung and Allen, 1995). The relationship between the biomasses of leaves and stems is important due to its effects on nutritional value and voluntary consumption by animals. The NDF is associated with fibrous fractions and with voluntary consumption. The fractions that are not digested by the animal take up space in the digestive tract, impairing the digestion and consumption of dry matter (Euclides et al., 1999).

NDF_L and NDF_S showed a strong positive correlation, and their respective QTLs were in the same regions in the HGs. In the linkage map, qCEL_L30 and qADF_L20 QTLs were in the same region, and qCEL_S31 and qADF_S26 were both located in HG VIII and were significantly correlated (**Figure 1**). The traits CP and IVD showed a weak positive correlation, but the qCP_S8 and qIVD_S28 QTLs were present in the same region of HG V, and qCP_S9 and qIVD_S29 were identified in the same region of HG VIII. Interestingly, qPL_S34 was verified in HG VIII in the same region as CP and IVD QTLs, and qADF_L21 shared a similar region with qCP_L7 in HG V. In addition, qCP_L7 extended to QTLs related to agronomic traits (qGM1, qTDM4, qLDM5 and qSDM7). A total of 8 probable QTL hotspots have been identified, supporting the need for further studies in search of a specific gene controlling all these traits or several genes acting together.

Notably, in the progenies of *M. maximus* from lower-yielding parents, the nutritional value is generally higher. With the identified QTLs, more in-depth studies of this correlation will be possible. In addition, QTLs related to forage quality have not been found in other important tropical species, such as *Urochloa* spp. Therefore, our results can contribute to the search for important genomic regions in other forage species.

Search for similarity in QTL regions

The search for putative candidate genes was based on all 23 QTL regions. Generally, the same gene families from *A. thaliana*, *O. sativa*, and *P. virgatum* were identified for a common QTL region. These genes are also found in the literature and are described in **Table 6** and **Supplementary Table 5**.

Exploration of genes involved in plant growth and development, especially those related to hormone regulation, is crucial in forage grass breeding programs. Interestingly, the gibberellin family (GAs) was identified in HG III - region 9 (qPL_L33), whose QTL is related to lignin, a complex phenolic polymer deposited in the secondary cell wall of all vascular plants (Zhao, 2016). The interrelations between cell wall components cause cellulose and lignin to be codependent, a normal cellulose deposition pattern may be necessary for lignin assembly, and alterations of lignin content may lead to changes in the cell orientation of cellulose fibrils and, consequently, in digestibility (Anderson et al., 2015; Liu et al., 2016). GAs promote biochemical, physiological and anatomical plant changes (Hedden and Thomas, 2016). The induction of cellulose synthesis by GAs promotes the release of secondary regulators of cell wall proteins and, consequently, can boost lignin deposition and increase lignin content (Zhao, 2016). GAs at increased light levels have been shown to promote cell wall thickness and increase lignin deposition in xylem fibers (Falcioni et al., 2018).

Other important genes identified are associated with pectinesterase/pectin methylesterase and were present in HG V – region 14 (qGM1/qTDM4/qLDM5/qSDM7/qADF_L21/qCP_L7) and HG VIII – region 23 (qSIL_S37), which also contained agronomic QTLs. Pectinesterase is responsible for the hydrolyzation of pectin, the major component of cell walls (Pelloux et al., 2007). In addition, this enzyme is involved in developmental processes such as stem elongation in *A. thaliana* (Damm et al., 2016) and in *B. distachyon* grass (Feng et al., 2015). However, more in-depth research should be performed to ensure an understanding of the signaling pathways of these genes and make this understanding applicable to tropical forage breeding programs.

In conclusion, the present study produced a high-resolution linkage map with allele dosage information obtained from a full-sib progeny of *M. maximus* with high genetic variability. Even without the availability of a sequenced genome for this species, the approach adopted for the construction of our map was sufficient to detect many QTLs associated with agronomic and nutritional traits that are important for forage breeding. Our genetic map also allowed us to map the apo-locus to a single linkage group and provided a more up-to-date study of the mode of reproduction of *M. maximus*. The knowledge about the genetics of these traits that we obtained is the first step in discovering genes involved in relevant biological processes as well as understanding the genetic architecture of relevant traits in this species.

Conflict of interest

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

Author contributions

AG, LJ and AS conceived and designed the experiments. MS and LJ conducted the field experiments. TD, RF, AM, AP, and FO performed the laboratory experiments. TD, RF, LL, AM, AP, and FO analyzed the data. TD, RF and LL wrote the manuscript. All authors read and approved the manuscript.

Funding

This work was supported by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP 2008/52197-4), the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES – Computational Biology Program), Embrapa and UNIPASTO. TD received PhD fellowships from the CAPES Computational Biology Program and an MSc fellowship from FAPESP (2017/17969-5) and the CAPES Computational Biology Program. RF, FO and AP received postdoctoral fellowships from FAPESP (2018/19219-6, 2018/18527-9 and 2018/00036-9, respectively). LL received a postdoctoral fellowship from the CAPES Computational Biology Program. AG and AdS were recipients of a Research Fellowship, and LJ was recipient of a Technological Development scholarship, all received from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Acknowledgments

We acknowledge the members of the Brazilian Agricultural Research Corporation (Embrapa Beef Cattle) for making the progeny of this study available and for carrying out all phenotypic evaluations. This manuscript was previously posted to bioRxiv <https://www.biorxiv.org/content/10.1101/801399v1>.

5 RESULTADOS COMPLEMENTARES

Neste tópico está descrito toda a abordagem empregada até a obtenção do mapa de ligação presente no artigo desta tese “*High-resolution Linkage Map with Allele Dosage Allows the Identification of an Apomixis Region and Complex Traits in Guinea Grass (Megathyrus maximus)*”. Estes foram os primeiros dados de GBS obtidos para *M. maximus* e o uso de *softwares* capazes de lidar com informação de dosagem alélica. Por isso, uma série de testes foram realizados para obter um mapa final com maior acurácia.

Esta seção consiste em três subtópicos para facilitar o entendimento das estratégias adotadas para cada etapa, conforme os programas utilizados (**Tabela 1**). Vale ressaltar que a Parte III foi o método empregado para gerar o mapa de ligação final.

Tabela 1. Programas utilizados em cada parte do projeto.

Parte	Tassel-GBS	Dosagem SuperMASSA	Dosagem updog	Mapa TetraploidSNPMap	Mapa PolymapR
I	- Alinhamento total - *MAF 1%	✓		✓	
II	- Alinhamento total - *MAF 5%		✓	✓	✓
III	- Alinhamento único - *MAF 5%		✓	✓	✓

* frequência mínima do alelo alternativo

O alinhamento das sequências obtidas foi realizado com o auxílio do programa Bowtie v.2.3.1 (Langmead e Salzberg, 2012), utilizando as configurações *very-sensitive-local* com um limite máximo de 20 problemas dinâmicos por busca (D) e máximo de quatro vezes para recomeçar a alinhar uma sequência (R). Os genomas utilizados como referência pertencem a espécies filogeneticamente próximas a *M. maximus*, sendo estes: (i) *Panicum hallii* v3.1; (ii) *Panicum virgatum* v4.1; (iii) *Setaria italica* v2.2; (iv) *Setaria viridis* v2.1 e (v) *Uruchloa ruziziensis* v.1. Os genomas (i) a (iv) estão disponíveis no Phytozome (<http://phytozome.jgi.doe.gov/>) (Goodstein et al., 2012) e o genoma (v) no GenBank (<https://www.ncbi.nlm.nih.gov/genome/>) (Benson et al., 2005) (**Tabela 2**). As sequências obtidas também foram alinhadas com três transcriptomas de *M. maximus*, dois deles montados a partir de genitores do programa de melhoramento da Embrapa – Gado de Corte (Toledo-Silva et al., 2013; Chiari, L., dados não publicados), sendo um deles a partir dos mesmos genitores da progênie deste estudo (Toledo-Silva et al., 2013). O terceiro transcriptoma utilizado foi montado a partir de genitores do programa de melhoramento realizado na Índia (Radhakrishna et al., 2018).

Tabela 2. Dados referentes aos genomas utilizados como referência para o alinhamento das sequências obtidas a partir do GBS.

Genomas	Organização	Ploidia	Montagem	Tamanho (Mb)	Gaps
<i>P. hallii</i>	Autopoliploide	2n	9 cromossomos + <i>scaffolds</i>	535	*
<i>P. virgatum</i>	Alopoliploide	4n	9 cromossomos (K**) + 9 cromossomos (N***) + <i>contigs</i>	1.165,7	*
<i>S. italica</i>	Autopoliploide	2n	336 <i>scaffolds</i>	400,9	1,2%
<i>S. viridis</i>	Autopoliploide	2n	9 cromossomos + <i>scaffolds</i>	395,7	0,2%
<i>U. ruziziensis</i>	Autopoliploide	2n	102.577 <i>scaffolds</i>	732,5	*

*não estimado, K** corresponde ao primeiro genoma e N*** ao segundo genoma.

Para cada alinhamento com os genomas referência citados acima, foi gerado um arquivo com o alinhamento total que corresponde a soma das sequências alinhadas exatamente uma vez e das sequências que alinharam mais de uma vez. A identificação de SNPs foi realizada para cada arquivo através do pipeline TASSEL-GBS v.4, modificado para poliplóides (Glaubitz et al., 2014; Pereira et al., 2018), sendo considerada a contagem alélica mínima de 1.000 e o limite acima de 20% de *mismatches*, a partir do qual SNPs duplicados não foram combinados. O parâmetro de frequência mínima do alelo alternativo (mnMAF) foi de 1%.

De maneira geral, a porcentagem de alinhamento total das sequências de *M. maximus* diferiu em uma margem de 10%, entre as referências (**Figura 1**). O alinhamento com o genoma de *U. ruziziensis* resultou no maior número de sequências alinhadas (32,21%), seguido do genoma de *P. virgatum* (24,23%) (**Figura 1**). O menor número de sequências alinhadas foi obtido com o transcriptoma de Chiari, L. (dados não publicados) (19,04%). A partir dos alinhamentos foi possível observar que mais da metade das sequências não foram alinhadas nenhuma vez. Resultados semelhantes foram relatados em um estudo de uma progênie de *M. maximus* (Lara et al., 2019) e para outras gramíneas que não possuem genoma sequenciado (Ferreira et al., 2019).

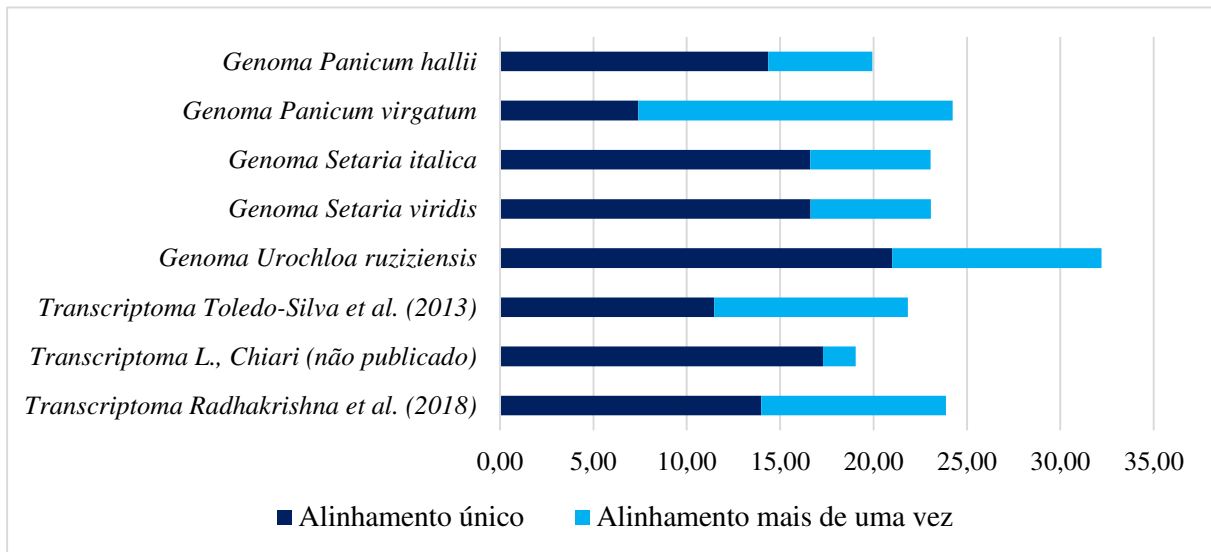


Figura 1. Alinhamento total (%) de sequências obtidas por GBS em *Megathyrsus maximus* em relação aos genomas e transcriptomas de referência utilizados.

A maior porcentagem de alinhamento com o genoma de *U. ruziziensis* seguido de *P. virgatum*, reflete a proximidade filogenética de *M. maximus* a esses gêneros. Visto que *M. maximus* já pertenceu ao clado *Panicum* (Zimmermann et al., 2012) e estudos corroboram que a espécie poderia estar no clado *Urochloa*, dada a proximidade genômica verificada por meio de marcadores cloroplastidiais (Aliscioni et al., 2003). Muitos fatores podem estar relacionados à esta porcentagem de alinhamento geral observado com todos os genomas referência utilizados, como a capacidade de *softwares* em identificar e classificar sequências presentes em regiões duplicadas, a amostragem do sequenciamento e principalmente a carência do próprio genoma da espécie, impossibilitando a amostragem da maior parte da variabilidade genética presente no genoma. No caso dos transcriptomas utilizados, mesmo sendo da própria espécie vale ressaltar que são transcritos obtidos a partir de condições específicas, sendo provenientes apenas de regiões expressas do genoma e, por isso, a representatividade do genoma é menor.

É importante ressaltar, que para fins de comparação, o resultado do alinhamento das sequências de GBS aos genomas e transcriptomas foram apresentados juntos. No entanto, o transcriptoma de Radhakrishna et al. (2018) e o genoma de *U. ruziziensis* não constam nas análises da Parte I, pois os dados destes ainda não estavam disponíveis quando foram realizadas. Posteriormente foram inseridos, por isso constam nas análises da Parte II e III, respectivamente.

Parte I:

Após os alinhamentos, a identificação de SNPs com o programa Tassel modificado para poliplóides (Glaubitz et al., 2014; Pereira et al., 2018) resultou em um maior número de marcadores para os genomas do gênero *Setaria* (**Figura 2**), com mais de 94 mil para *S. viridis* e mais de 93 mil SNPs para *S. italica*. Também foi notório obtermos mais de 90 mil SNPs para o transcriptoma de *M. maximus* de Chiari, L. (dados não publicados).

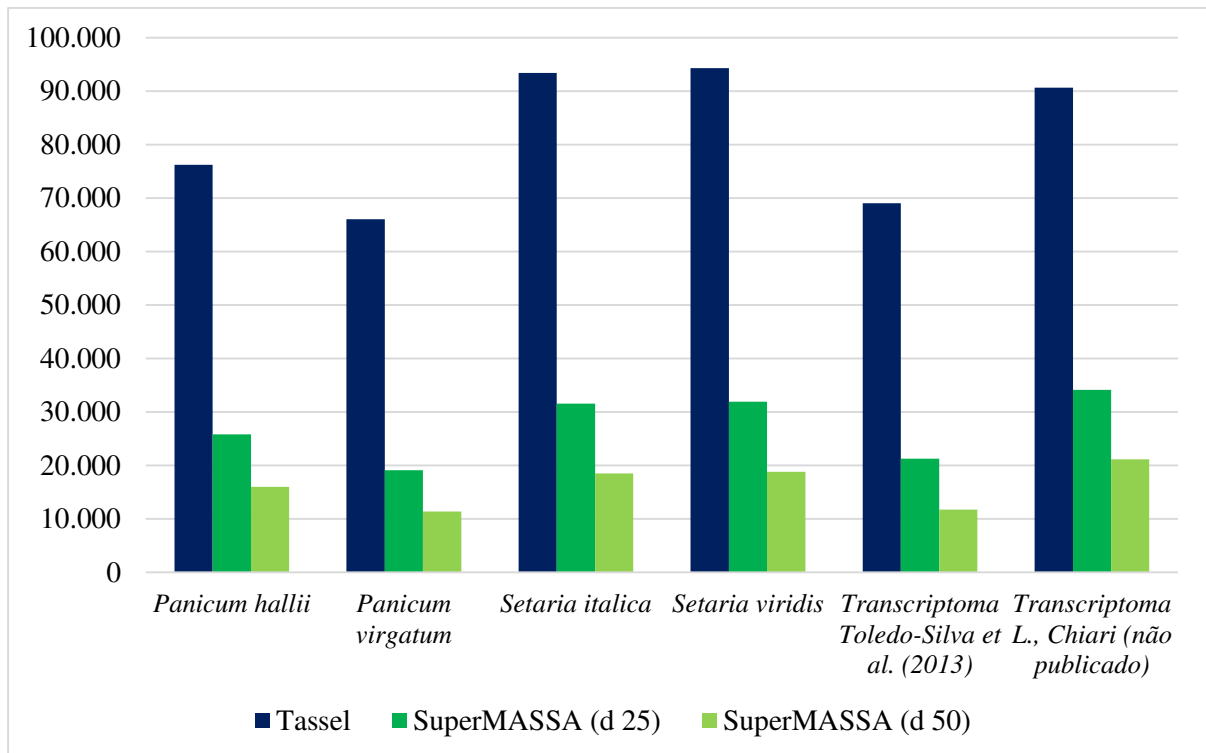


Figura 2. Número de marcadores SNPs obtidos preliminarmente com o *software* Tassel-GBS (Glaubitz et al., 2014; Pereira et al., 2018) e após estimativa da dosagem alélica com o SuperMASSA (Serang et al., 2012). A profundidade alélica mínima foi avaliada com 25 *reads* (d25) e 50 *reads* (d50).

A próxima etapa foi a estimação da dosagem alélica, para isso as profundidades testadas foram 25 e 50 *reads* a partir do *software* SuperMASSA (Serang et al., 2012). Os parâmetros utilizados foram o modelo de população F₁, ploidia fixa igual a 4, probabilidade a *posteriori* mínima de 0,75, *naive* de 0,60 e taxa mínima de chamada de 0,50. Posteriormente, os marcadores foram filtrados manualmente no *software* R v4.6 (R Core Team, 2018), visando remover marcas monomórficas e aquelas com dados perdidos acima de 15 e 25%. O uso de uma porcentagem mais restritiva (15%) para dados perdidos foi empregada para verificar a qualidade dos SNPs. Ao todo, cerca de 49 a 75 mil marcadores foram retirados entre a identificação de SNPs e a estimativa da dosagem alélica para cada conjunto de dados, sendo que apenas os testes de profundidade eliminaram entre 7 e 13 mil de marcadores (**Figura 2**).

Os alinhamentos utilizando *S. viridis*, *S. italica* e o transcriptoma de Chiari, L. (dados não publicados), novamente apresentaram um grande número de SNPs com profundidade de 50 *reads* (**Figura 2**), sendo obtidos mais de 18 mil SNPs a partir dos genomas do gênero *Setaria* e mais de 21 mil a partir do transcriptoma de Chiari L. (dados não publicados). Visando manter os marcadores com elevada profundidade, apenas aqueles com profundidade igual ou maior do que 50 *reads* foram retidos para as próximas análises. Apesar do elevado número de SNPs eliminados, o total de marcadores obtidos para cada conjunto de alinhamento ao final de todas as filtragens (**Figuras 3 e 4**) demonstraram a elevada qualidade do nosso conjunto de dados obtidos por meio de GBS para *M. maximus*.

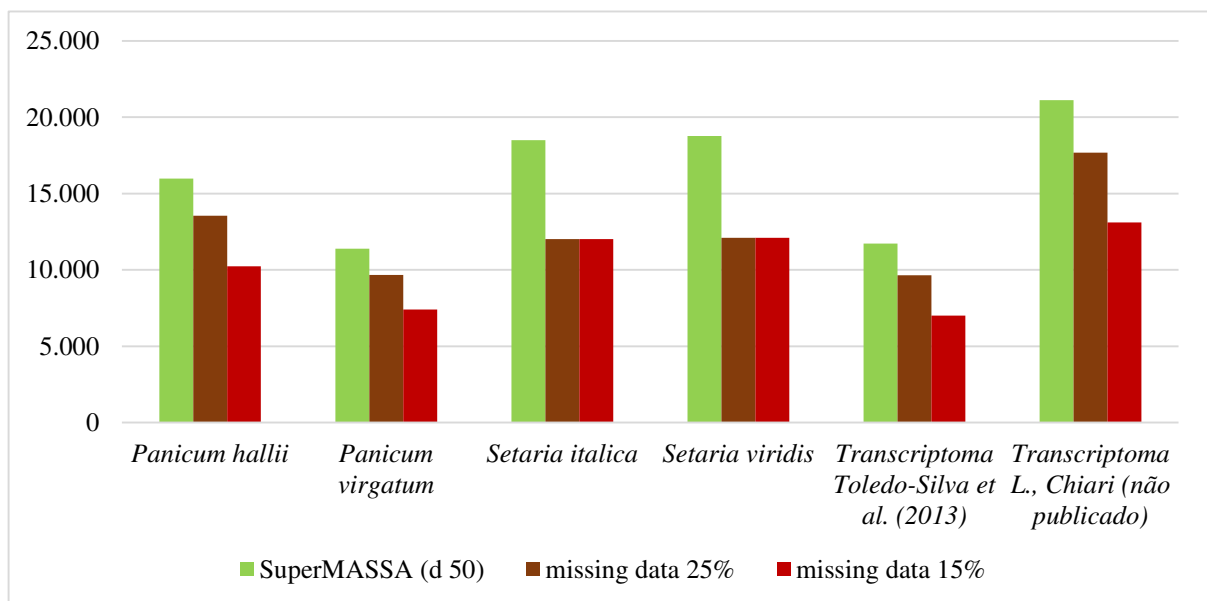


Figura 3. Comparação do número de marcadores SNPs obtido pelo *software* SuperMASSA (Serang et al., 2012) e após os filtros para dados perdidos de 25 e 15%.

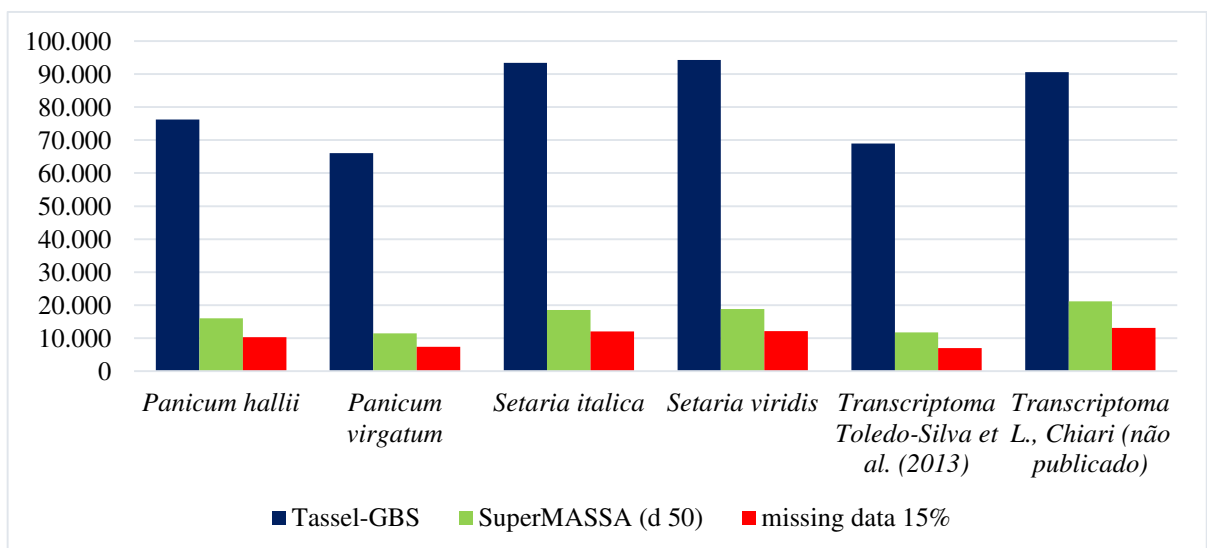


Figura 4. Número de marcadores SNPs obtido em cada etapa de análise.

Ferramentas que são capazes de utilizar marcadores SNPs em doses múltiplas nos estudos de mapeamento genético são recentes e muitos entraves estatísticos persistem dada a complexidade dos genomas poliploides. Mesmo com uma série de filtros de qualidade mais restritivos, o elevado número de marcas ao final foi evidente. No entanto, o programa TetraploidSNPMap permite um arquivo de entrada com um máximo de 8 mil SNPs (Hackett et al., 2017), impossibilitando que alguns conjuntos de dados fossem utilizados. Os conjuntos de SNPs utilizados foram obtidos através dos alinhamentos com o genoma de *P. virgatum* (7.393 SNPs) e com o transcriptoma de Toledo-Silva et al. (2013) (7.009 SNPs). Entretanto, o programa apresentou problemas de processamento com essa quantidade de marcadores.

Cada conjunto de marcadores foi então inserido como arquivo de entrada no programa TetraploidSNPMap versão 3.0 (Hackett et al., 2017), depois foram checados pelo teste de Qui-quadrado, sendo selecionadas as marcas que foram iguais ou maior que 0,001 para SNP Simplex e 0,01 para doses mais altas de SNPs. Após essa seleção, marcadores não-distorcidos foram usados na análise de *cluster* que os dividiram em 8 grupos de homologia (GH). Para a ordenação das marcas foram usadas a análise de dois-pontos e a análise de escala multidimensional (MDS), que calculam as frequências de recombinação e os LOD *scores*. Nesta etapa foram removidos os SNPs *outliers*. Por fim, os grupos de homologia do mapa de ligação foram visualizados usando o pacote LinkageMapView (Ouellette et al., 2017), presente no *software* R v4.6 (R Core Team, 2018).

Após as análises de ligação, o conjunto de marcadores identificados a partir do genoma de *P. virgatum* consistiu de 3.270 SNPs (**Tabela 3**) com um tamanho de 918,34 cM para o mapa de ligação, porém alguns grupos de homologia apresentaram discrepância de marcas em relação aos demais, como os GHs 4, 7 e 8. Em relação ao conjunto de marcadores obtidos a partir do transcriptoma de Toledo-Silva et al. (2013) restaram 3.233 SNPs, que foram bem distribuídos em 8 GHs com um tamanho total de 1.009,1 cM, portanto, este mapa de ligação se mostrou mais representativo ao genoma de *M. maximus*. O GH 1 é o maior com 176,9 cM e o menor é o GH 3, com 102,1 cM. Quanto ao número de marcadores, o GH 7 foi o que apresentou a maior quantidade com 602 SNPs e o GH 8 foi o que apresentou menos com 184 SNPs (**Figura 5; Tabela 3**).

Tabela 3. Análises de ligação obtidas para dois conjuntos de marcadores por meio do programa TetraploidSNPMap (Hackett et al., 2017).

Grupos homologia	Genoma <i>Panicum virgatum</i>		Transcriptoma Toledo-Silva et al. (2013)	
	SNPs	Tamanho (cM)	SNPs	Tamanho (cM)
1	647	111,9	470	176,9
2	467	129,2	402	112,9
3	808	162,7	202	102,1
4	124	83,8	493	115,2
5	659	114,4	503	175,0
6	423	156,3	377	104,4
7	124	101,9	602	110,1
8	18	58,2	184	123,5
	3270	918,34	3233	1009,1

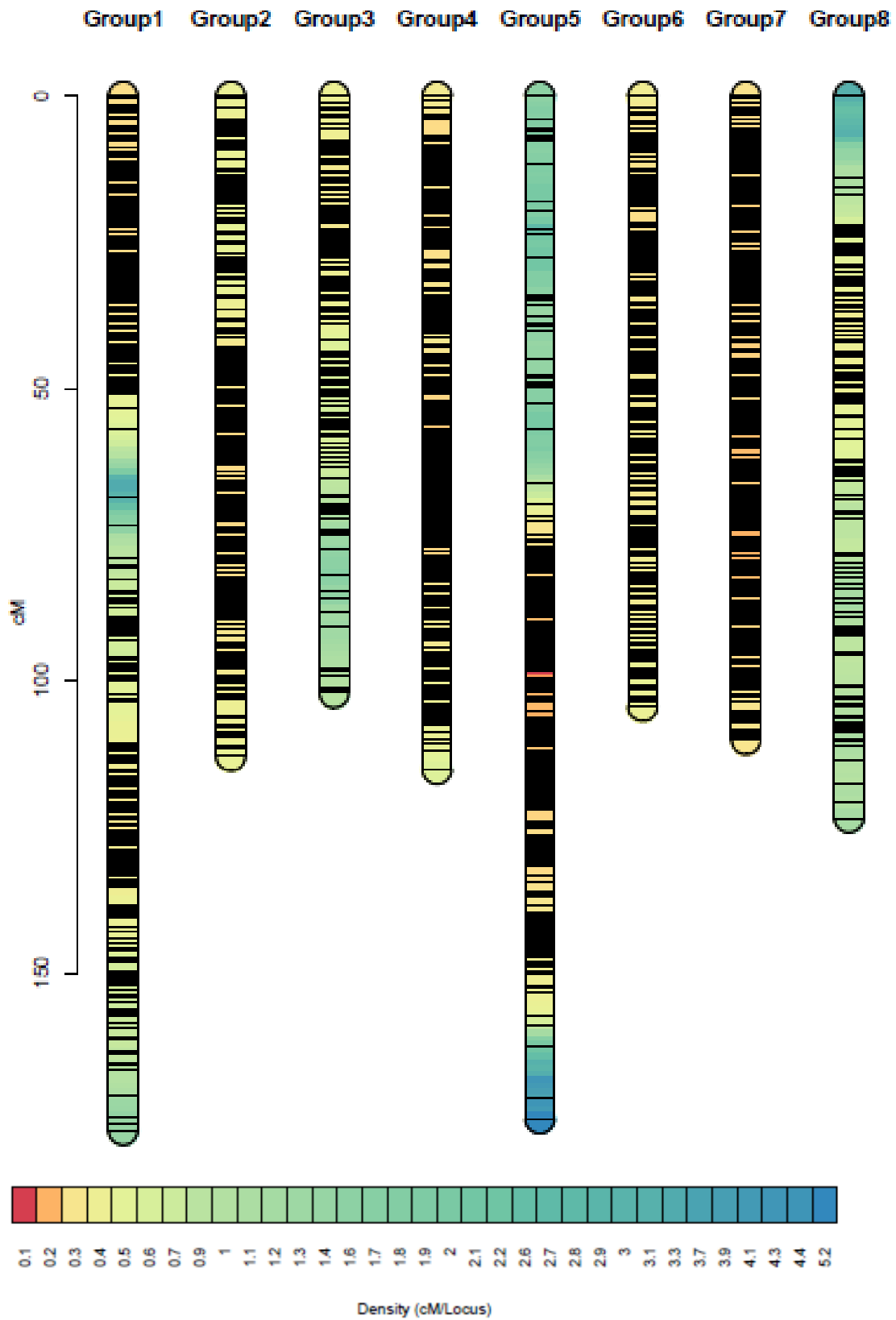


Figura 5. Primeiro mapa de ligação obtido para *M. maximus* com SNPs identificados utilizando o transcriptoma de Toledo-Silva et al. (2013).

Parte II

Dada a primeira estratégia que possibilitou obter um mapa genético a partir de SNPs presentes em regiões potencialmente expressas do genoma (**Figura 5**), fomos instigados a realizar uma segunda abordagem com os conjuntos de dados resultantes dos alinhamentos com os transcriptomas de *M. maximus*. Portanto, utilizamos os conjuntos de dados dos transcriptomas de (1) Chiari, L. (dados não publicados); (2) Toledo-Silva et al. (2013) e (3) Radhakrishna et al. (2018).

A partir de cada alinhamento obtido anteriormente foi realizada a identificação de SNPs com o Tassel-GBS (Glaubitz et al., 2014; Pereira et al., 2018), dessa vez utilizando uma mnMAF de 5%. Apesar da alteração do parâmetro de mnMAF de 1% para 5%, um elevado número de marcadores SNPs foi mantido e com uma maior qualidade atribuída a estes SNPs. Conforme esperado, uma MAF de 5% acarretou em uma diminuição do número de SNPs (**Figura 6**), porém, estudos demonstram que MAFs mais restritivas são imprescindíveis em análises de progênies tetraploides com número amostral menor, visando maior acurácia dos marcadores e diminuir a probabilidade de falsos-positivos no conjunto de dados (Tabangin et al., 2009). O maior número de marcadores foi identificado a partir do alinhamento com o transcriptoma de Radhakrishna et al. (2018), com aproximadamente 89 mil marcas.

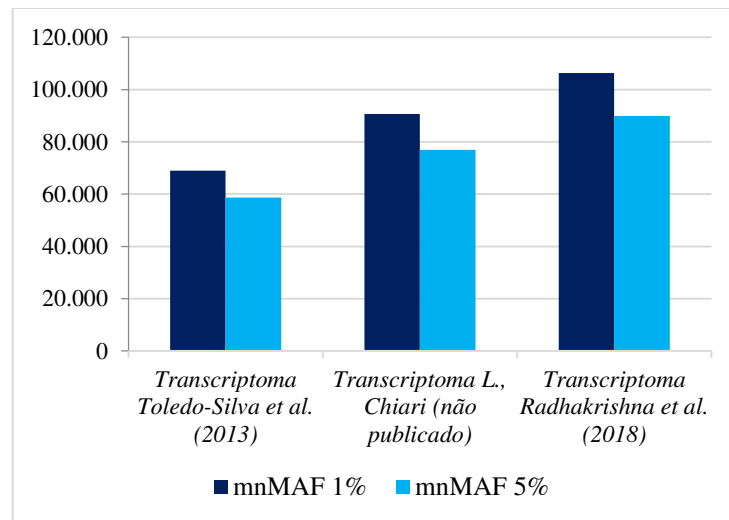


Figura 6. Identificação de SNPs com mnMAF de 5% e 1% através do *software* Tassel-GBS (Glaubitz et al., 2014; Pereira et al., 2018).

Visando maior confiabilidade dos resultados de profundidade alélica em nossos dados, o recente *software* Updog (Gerard et al., 2018) também foi empregado para estimar a

dosagem alélica dos conjuntos de alinhamento resultantes dos transcriptomas (Chiari, L., dados não publicados; Toledo-Silva et al., 2013; Radhakrishna et al., 2018). Todas as análises de profundidade e dosagem alélica associadas a este pacote foram realizadas no *software* R v4.6 (R Core Team, 2018). A profundidade alélica média mínima de cada conjunto de SNPs foi estimada manualmente e foram considerados apenas marcadores com profundidade acima de 60 *reads*. Os argumentos utilizados para o pacote foram ploidia fixa igual a 4, função *flexdog* e modelo F_1 . Para a proporção *a posteriori* de indivíduos genotipados incorretamente foi usado o valor de 0,15. Novamente, marcadores monomórficos e aqueles com dados perdidos acima de 25% foram removidos manualmente com o *software* R v4.6 (R Core Team, 2018).

Após obtida a dosagem alélica dos SNPs foi averiguado que o conjunto de dados (1) manteve a maior quantidade de marcadores (**Figura 7**), um total de 21,1 mil SNPs. Enquanto que para (2) foram retidos 8,3 mil e em (3) 10,2 mil SNPs. Ao considerar os valores obtidos pelo programa SuperMASSA (Serang et al., 2012) foi verificado que o pacote Updog (Gerard et al., 2018) excluiu uma maior quantidade de SNPs, entretanto, deve ser considerado que uma profundidade igual ou maior de 60 *reads* foi utilizada, dando uma maior confiabilidade aos marcadores com dosagem alélica que foram mantidos. A discriminação confiável entre os genótipos heterozigotos em um autotetraploide exige uma profundidade de *reads* de pelo menos 60 (Uitdewilligen et al., 2013).

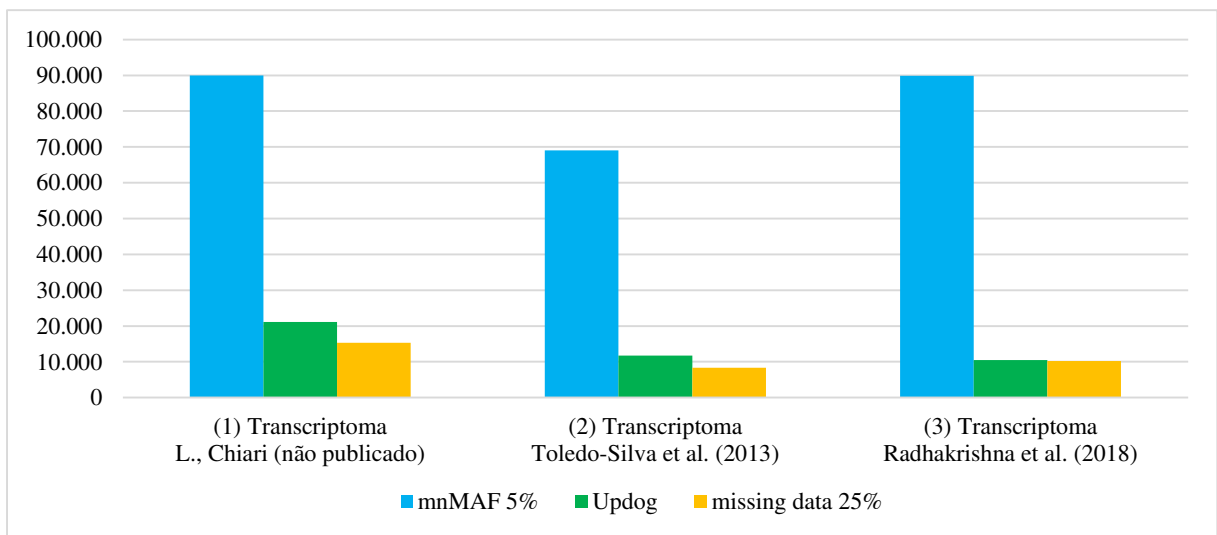


Figura 7. Número de marcadores SNPs obtido em diferentes etapas de análise.

Além disso, após o filtro para dados perdidos foi observado que pouquíssimos marcadores foram eliminados, sendo em (1) 256, (2) 168 e (3) 264 SNPs. Isto evidencia o poder estatístico do updog (Gerard et al., 2018), que lida com viés alélico, *outliers*, superdispersão e erros de sequenciamento para dados GBS (Gerard et al., 2018). Dessa forma, pôde ser inferido

que este pacote poderá contribuir para a construção de mapas genéticos mais robustos e com marcadores mais confiáveis.

Ao final, um total de 33.861 SNPs foram obtidos a partir do alinhamento dos três transcriptomas, sendo que para (1) foram 15.323, (2) 8.360 e (3) 10.178 SNPs. A fim de unir os três conjuntos, primeiro foi necessário diminuir a redundância de SNPs que contém a mesma informação de dosagem. Para isso, foi usado o pacote Biostrings (Pages et al., 2018) que resultou em um total de 12.445 marcadores, um número bastante considerável para as análises de ligação. Com a matriz obtida foi dado início às análises com o polymapR (Bourke et al., 2018), outra recente ferramenta para mapeamento genético, que instigou a sua aplicação pela flexibilidade do *pipeline* no *software* R (R Core Team, 2018) e por não possuir limite na quantidade de SNPs analisados. Primeiramente, uma série de filtragens foram realizadas, os SNPs foram checados se estavam de acordo com o modelo de herança polissômica pela função *checkF1*. Um *threshold* de 5% foi usado para dados perdidos por marcadores e por indivíduo. Marcadores duplicados foram removidos. A análise de componentes principais (PCA) foi realizada para identificar genótipos muito semelhantes, no qual, representa a distribuição dos dados de forma linear. Os genitores foram alocados em quadrantes opostos, inferindo-se, portanto, que os mesmos são contrastantes (**Figura 7C**). Dois indivíduos (B126 e B127) que se apresentaram muito próximos a cv. Mombaça foram removidos, da mesma forma que o indivíduo C44 foi semelhante ao C49, sendo C44 removido.

Uma elevada qualidade dos marcadores SNPs foi visualizada durante as etapas de filtragem (**Figura 8**), nas quais nenhum indivíduo foi eliminado por dados faltantes e todos foram compatíveis com os genitores e apenas 582 SNPs foram retirados com 5% de dados perdidos. Mesmo com o filtro de qualidade realizado manualmente para eliminar SNPs redundantes, os parâmetros do polymapR (Bourke et al., 2018) se apresentaram mais sensíveis e identificaram a presença de mais SNPs redundantes (**Figura 8D**). A eliminação de marcadores redundantes é extremamente necessária, uma vez que pode acarretar grande viés ao mapa genético (Lu et al., 2013).

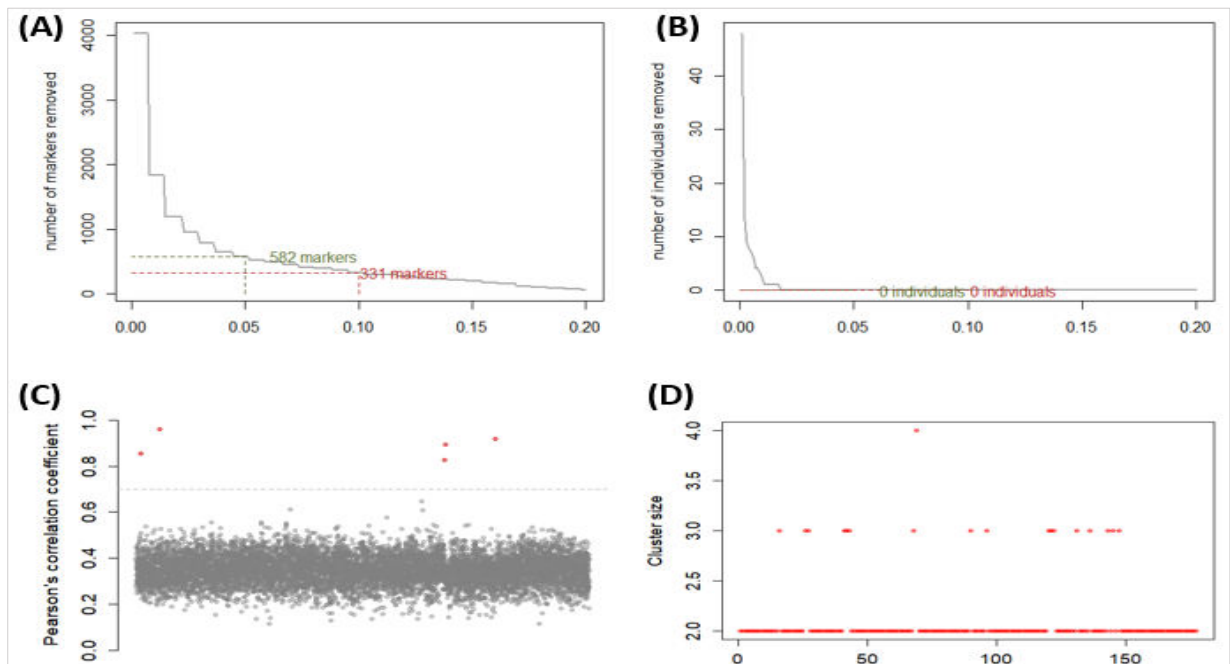


Figura 8. Etapas de filtragem do polypmapR (Bourke et al., 2018). (A) *Threshold* usado para dados perdidos por marcadores e (B) por indivíduo. (C) *Threshold* para combinar indivíduos muito semelhantes. (D) Marcadores duplicados.

A partir de 133 híbridos e 11.837 marcadores foi iniciada as análises de ligação. A princípio é construído um mapa de ligação para cada genitor, onde marcadores simplex (Aaaa) x nulliplex (aaaa) foram usados como um “esqueleto” para a montagem prévia dos grupos de homologia de cada mapa. Os SNPs com essa dosagem foram representados na **Figura 9A**, em um *threshold* de LOD *scores*. Conforme o aumento do LOD *score*, maior o número de grupos de SNPs, os quais podem ser considerados como grupos de homologia (GH) e de ligação (GL). Obtivemos a formação de quatro grupos no LOD igual a 3, sendo que um deles apresentou um número muito maior de SNPs. Com o aumento do LOD para 4, foi observado a formação de 18 grupos, sendo que um deles se manteve com a maior quantidade de SNPs. No entanto, não foi observada a presença dos oito grupos de homologia esperados para *M. maximus* e tão pouco dos 32 grupos de ligação. O polypmapR (Bourke et al., 2018) também permite a aferição de grupos manualmente, dessa forma, conseguimos um máximo de cinco GHs e 20 GLs fixando o LOD igual a 7 (**Figura 9C**). Averiguamos que com a diminuição do LOD ocorria o aumento de GLs, porém diminuía o número de GHs (**Figura 9B e 9C**). Apesar disso, tentamos separar os GLs variando os valores de LOD e observando como se comportavam os grupos e as ligações entre eles. Mesmo com a aferição manual, o pacote não aceitou dar continuidade as análises de ligação. Inúmeras tentativas foram realizadas manualmente por meio do acréscimo de marcadores com outras doses, entretanto, nenhum resultado satisfatório foi obtido.

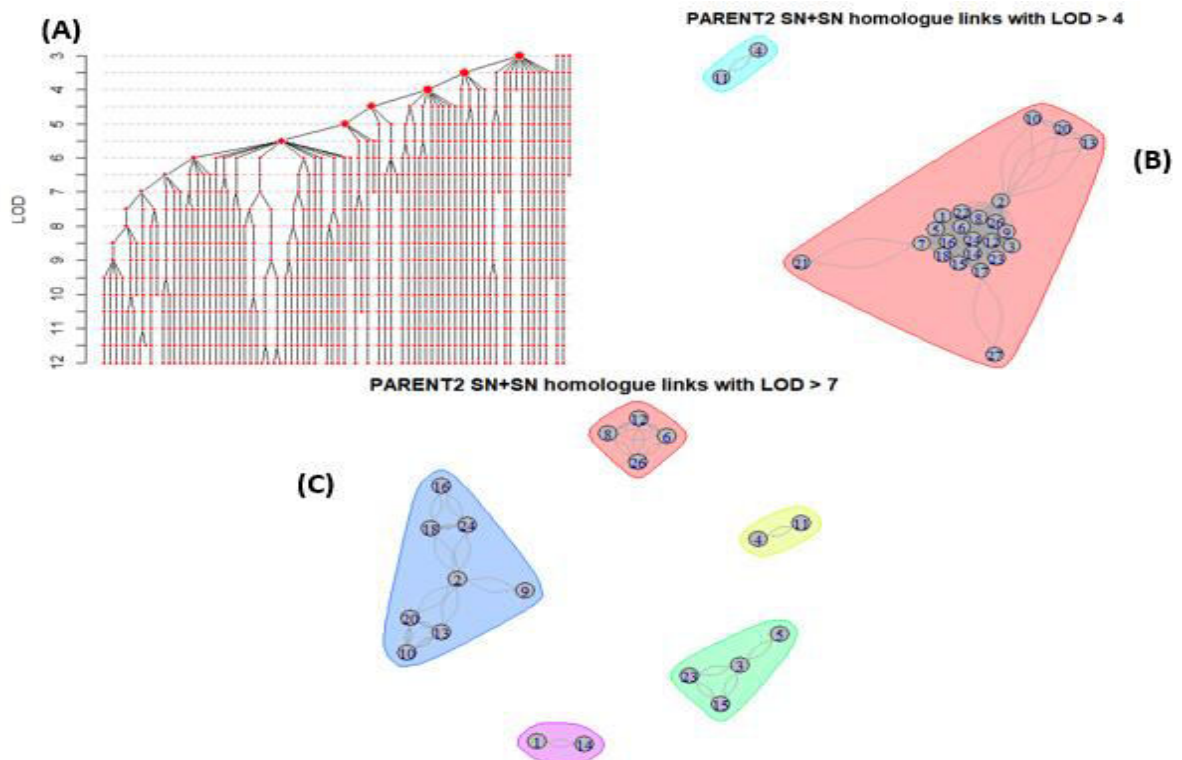


Figura 9. Análises de ligação com o pacote PolymapR. (A) *Threshold* de LOD scores. (B) Representação gráfica do comportamento dos GH e GL para o genitor sexual com LOD 4 e (C) LOD 7.

Com a possibilidade de um grande viés estar inserido aos dados, empregamos somente o conjunto de dados obtido a partir do alinhamento com o transcriptoma de Toledo-Silva et al. (2013), montado a partir dos genitores da nossa população de estudo. As etapas do pacote citadas anteriormente também foram realizadas para este conjunto. A partir de 8.360 SNPs permaneceram 4.309 SNPs e estes apresentaram uma melhor distribuição entre os grupos (**Figura 10**). Devido ao programa fazer uso da estratégia de apenas ao final das análises de ligação unir os mapas dos genitores, nossos dados retornavam com GHs repetidos e também com a ausência de outro GH, não sendo possível gerar um mapa de ligação integrado devido à problemas do próprio programa, os quais foram confirmados pelos autores do pacote.

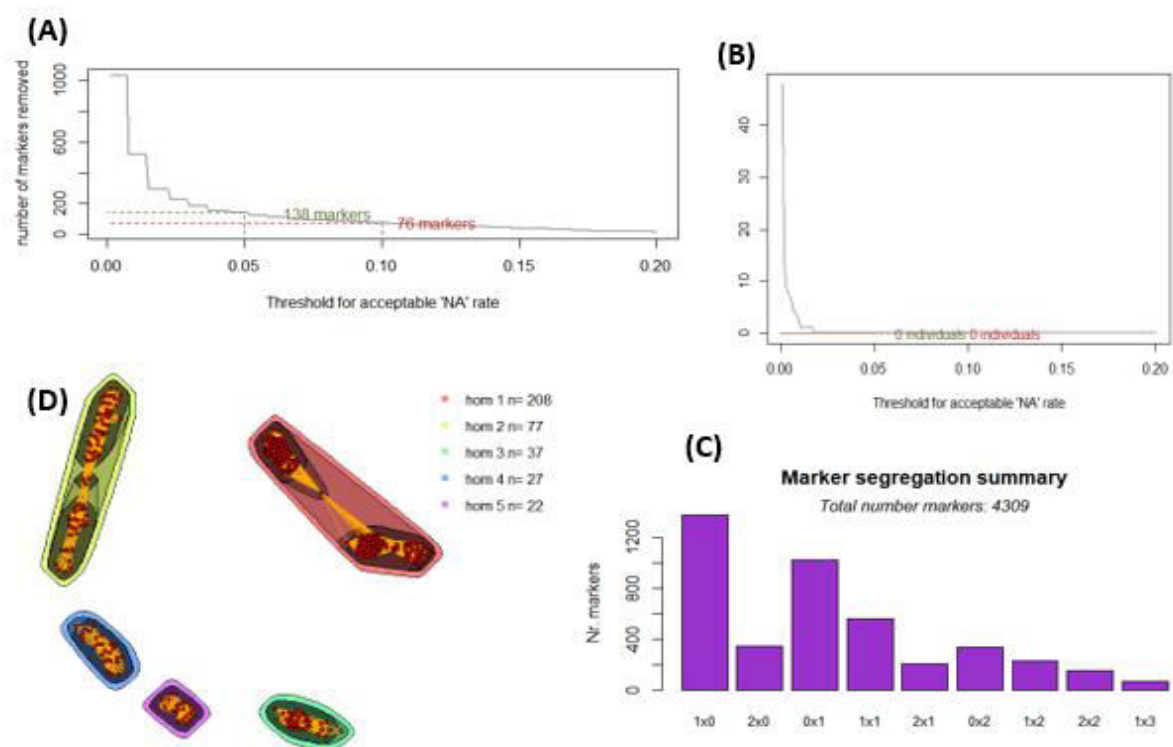


Figura 10. Etapas de filtragem e análise de ligação para o conjunto de dados alinhado com o transcriptoma de Toledo-Silva et al. (2013) e dosagem alélica obtida com o updog (Gerard et al., 2018). **(A)** *Threshold* usado para dados perdidos por marcadores e **(B)** por indivíduo. **(C)** Número de SNPs após filtragem e segregação esperada conforme dosagem. **(D)** Representação gráfica dos GHs.

Em uma nova alternativa, utilizamos o conjunto de dados obtido do alinhamento com este mesmo transcriptoma, porém a partir da dosagem alélica obtida com o SuperMASSA (Serang et al., 2012). Por meio da filtragem realizada no polymapR (Bourke et al., 2018) foi verificado no conjunto gerado pelo SuperMASSA (Serang et al., 2012) um número muito maior de dados faltantes (**Figura 11**). Dessa forma, é possível inferir que os modelos estatísticos empregados ao updog (Gerard et al., 2018) garantiram uma maior acurácia aos nossos dados. As análises de ligação com esse conjunto se dispuseram de forma idêntica ao conjunto anterior, isto é, os GLs dos genitores não conseguiram se agrupar como o esperado. Acreditamos que o fato de ser um pacote criado a partir de dados de simulação (Bourke et al., 2018), em que temos um conjunto ideal de marcadores, pode ter influenciado a não aplicabilidade dos nossos dados ao polymapR (Bourke et al., 2018). Apesar disso, este pacote demonstrou enorme potencial para a construção de mapas genéticos de poliploides e soluções poderão vir conforme a atualização de novas versões.

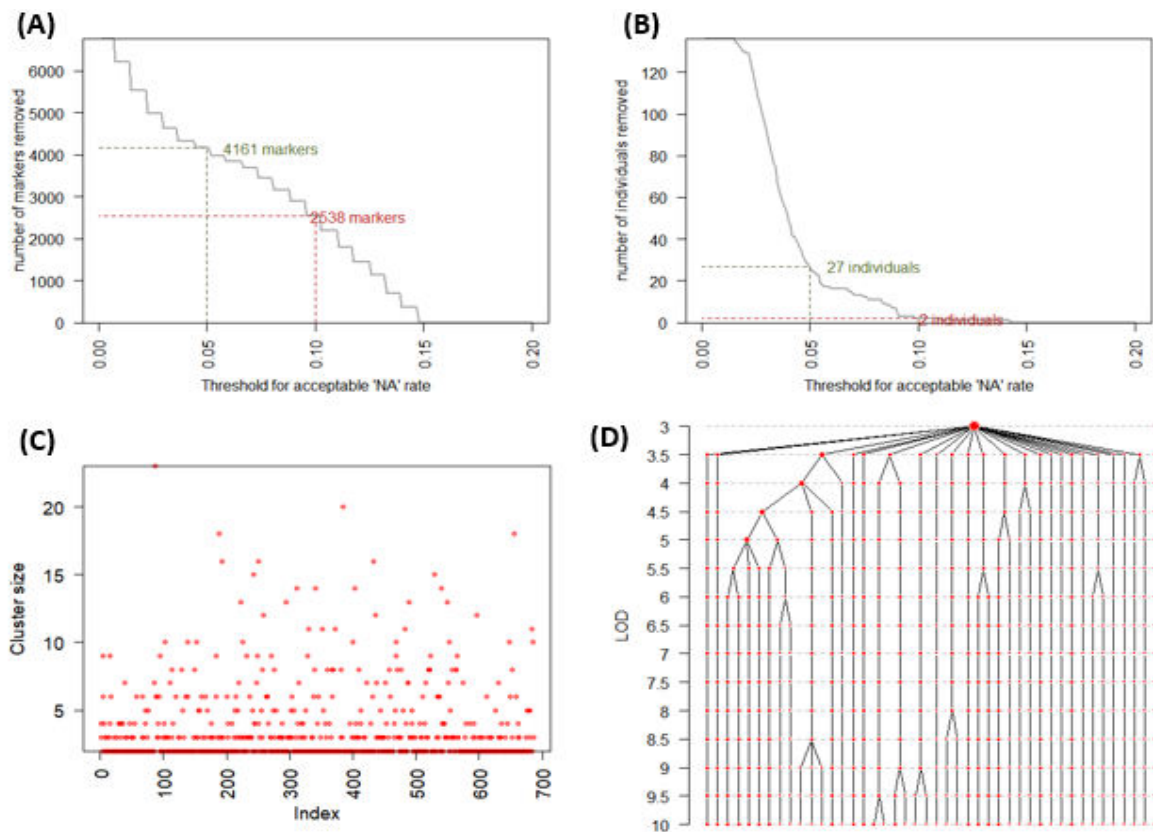


Figura 11. Etapas de filtragem e análise de ligação para o conjunto de dados alinhado com o transcriptoma de Toledo-Silva et al. (2013) e dosagem alélica obtida com o SuperMASSA (Serang et al., 2012). **(A)** *Threshold* usado para dados perdidos por marcadores e **(B)** por indivíduo. **(C)** Marcadores duplicados. **(D)** *Threshold* de LOD scores.

Em outros trabalhos que fizeram uso de transcriptomas como genoma referência, foram empregadas diferentes abordagens para lidar com um possível viés associado aos dados. Por exemplo, a busca de sintenia dos transcritos com espécies muito próximas filogeneticamente (van Geest et al., 2017), ou uma montagem do transcriptoma de *reads* únicas (Singh et al., 2017). No primeiro caso, em um genoma com ploidia maior que $2n$, estaríamos acrescentando muitos marcadores redundantes ao mapa genético, já que não é possível identificar múltiplas cópias de genes. O segundo caso refere-se à capacidade de *softwares* em lidarem com a duplicação genômica dos poliploides. Qualquer estratégia adotada, neste momento, carregará a possibilidade de um viés alélico significativo.

Os transcriptomas utilizados neste ensaio, usaram metodologias distintas quanto ao número de réplicas, ao tipo de sequenciamento e aos *softwares* e parâmetros usados para a montagem dos transcritos (Chiari, L., dados não publicados; Toledo-Silva et al., 2013; Radhakrishna et al., 2018). Todas essas implicações, sem dúvida, atuaram na dificuldade de formação dos GHs, pois continham um grande viés atrelado a esses SNPs. Por isso, a parte

seguinte do nosso trabalho não abordou esses transcriptomas, enfatizando que os genomas podem agregar uma maior robustez aos mapas de ligação, visto que quaisquer SNPs de regiões codantes e não codantes podem estar presentes.

Parte III:

Importante salientar que a metodologia abordada nesta Parte foi a que obtivemos um resultado satisfatório para o trabalho e que promoveu o Capítulo I.

A partir dos alinhamentos obtidos anteriormente com os genomas utilizados como referência, isto é, *P. hallii*, *P. virgatum*, *S. italica*, *S. viridis* e *U. ruziziensis*, que contém o alinhamento total das *tags* GBS, um outro arquivo foi gerado manualmente somente com as *tags* alinhadas uma única vez, visando eliminar possíveis SNPs presentes em regiões duplicadas do genoma. Cada conjunto de dados gerado foi, então, utilizado para a identificação de SNPs com o *software* Tassel-GBS modificado para poliploides (Glaubitz et al., 2014; Pereira et al., 2018), com a mnMAF mantida em 5%. O maior número de SNPs foi verificado a partir do genoma de *U. ruziziensis*, cerca de 87 mil, seguido dos alinhamentos com os genomas do gênero *Setaria*, 63 mil marcadores. Uma menor quantidade foi observada ao alinhamento de *P. virgatum*, aproximadamente 24 mil SNPs (**Figura 12**).

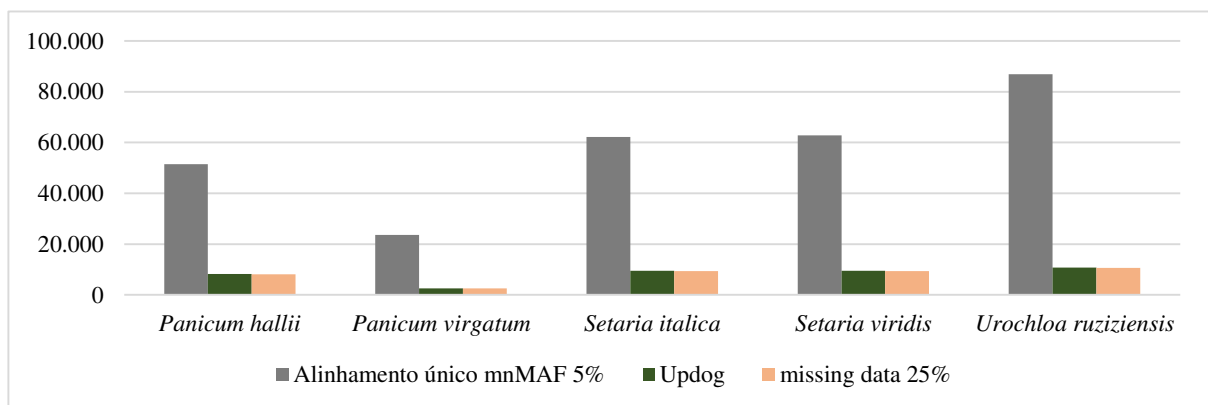


Figura 12. Número de marcadores SNPs obtido em cada etapa de análise.

Todas as análises de profundidade e dosagem alélica seguiram a metodologia empregada na Parte II com o pacote updog (Gerard et al., 2018) e feito uso dos mesmos parâmetros. Em relação a quantidade de SNPs obtida com o SuperMASSA (Serang et al., 2012) e com o updog (Gerard et al., 2018), verificamos que o primeiro programa gerou em média 3 mil marcadores a mais, entretanto, a qualidade das marcas se apresentaram superior com o updog (Gerard et al., 2018) atribuído aos poucos SNPs que foram descartados na etapa de

filtragem para dados perdidos (**Figura 12**). Por isso, indicamos o seu uso para determinar a dosagem alélica de autopoliploides como *M. maximus* a partir de dados NGS.

Cada conjunto de marcadores obtido pelo updog (Gerard et al., 2018) foi utilizado como arquivo de entrada no polymapR (Bourke et al., 2018). Todas as etapas de filtrações com o uso deste pacote seguiram os parâmetros mencionados na Parte II. Após os filtros, o *threshold* de LOD scores contribuiu para averiguar a distribuição dos SNPs dentro dos grupos. Conforme a **Figura 13**, foi observado que no LOD igual a 3 ocorreu a formação de um único grupo para os conjuntos a partir dos genomas de *S. italica* e de *S. viridis*, e três grupos para os conjuntos a partir dos genomas de *P. hallii* e de *U. ruziziensis*, sendo um deles com uma quantidade muito maior de marcadores. No LOD igual a 4 foi verificado a presença de mais de 20 grupos para estes conjuntos de marcadores.

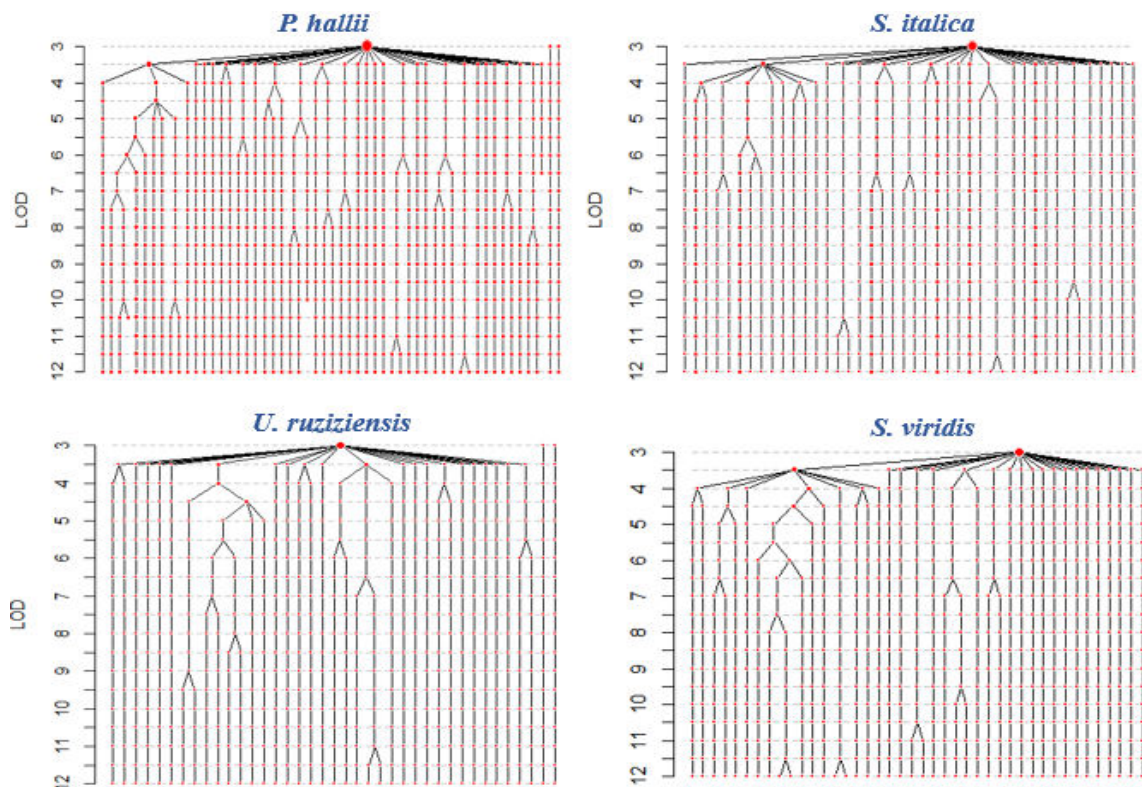


Figura 13. *Threshold* de LOD scores a partir de análises de ligação pelo pacote polymapR (Bourke et al., 2018).

Somente os SNPs identificados pelo alinhamento a partir do genoma de *P. virgatum* demonstraram uma distribuição próxima a esperada (**Figura 14**), corroborando a proximidade filogenética desta espécie a *M. maximus*. Em LOD igual a 3 foi possível observar oito grupos de homologia, sendo que em cinco deles, foi verificado quatro ou oito grupos menores, representativos dos grupos de ligação. Apesar da continuidade das análises de ligação,

novamente não houve a formação do mapa integrado com a mesma situação decorrida na Parte II.

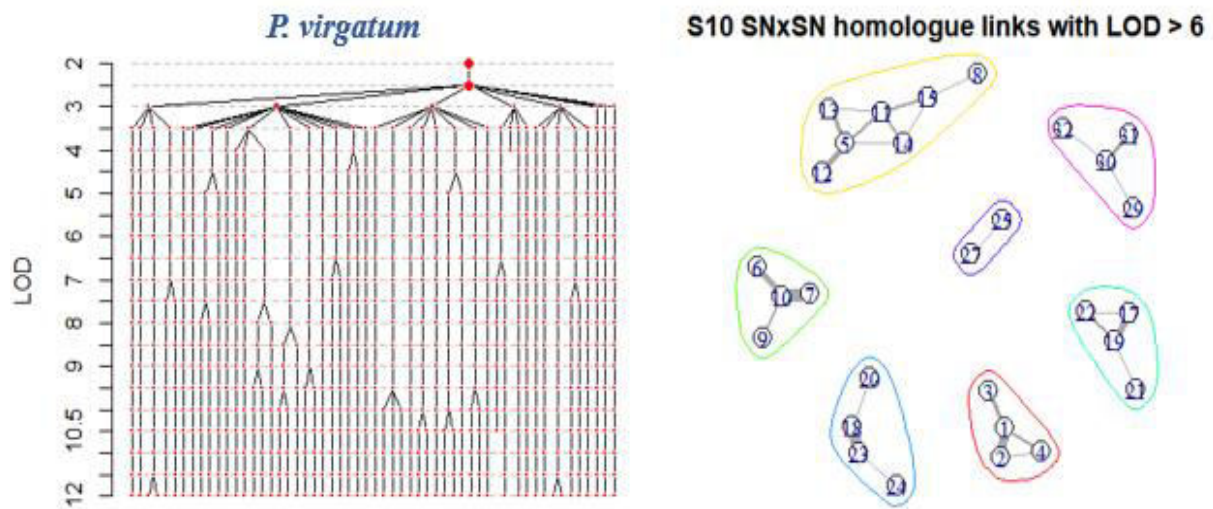


Figura 14. Análises de ligação para o conjunto de dados alinhado a partir do genoma de *P. virgatum*.

A partir disso, duas estratégias foram traçadas. A primeira delas foi inserir o conjunto de SNPs, obtidos após o updog, no programa TetraploidSNPMap (Hackett et al., 2017). No entanto, os marcadores não apresentaram uma boa distribuição entre os GHs, sendo que a maioria deles se mantiveram no GH 1. Na segunda estratégia utilizamos como arquivo de entrada para o TetraploidSNPMap (Hackett et al., 2017) o conjunto de marcadores obtidos após a filtragem realizada pelo polymapR (Bourke et al., 2018), que resultou em um arquivo com 132 indivíduos e 1.322 marcadores. Dessa vez, os SNPs se distribuíram uniformemente entre os 8 GHs com um total de 858 marcadores, portanto, este mapa de ligação obtido deu origem ao artigo. Dessa forma, foi possível concluir que a utilização das sequências alinhadas uma única vez ao genoma, a chamada de SNPs utilizando o *software* Tassel-GBS modificado para poliploides (Glaubitz et al., 2014; Pereira et al., 2018), a estimação da dosagem alélica com o pacote updog (Gerard et al., 2018), seguida de filtragens de qualidade mais sensíveis com o pacote polymapR (Bourke et al., 2018), e posterior análises de ligação com o programa TetraploidSNPMap (Hackett et al., 2017), resultaram em um mapa robusto e de maior acurácia para *M. maximus*, visto que, os valores utilizados nestes parâmetros da Parte III são aqueles exigidos para mapas de ligação de alta qualidade.

6 RESUMO DOS RESULTADOS

Os resultados obtidos estão resumidos a seguir e indicam que os objetivos propostos foram alcançados:

Artigo

- O uso de *softwares* desenvolvidos especificamente para poliploides permitiu estimar a dosagem alélica dos marcadores SNPs.
- O fenômeno da redução dupla, comum em autopoliploides, foi identificado pela primeira vez em *M. maximus*.
- Um mapa de ligação com 858 SNPs em múltiplas doses foi obtido a partir de uma progênie de cruzamento biparental do programa de melhoramento genético da Embrapa Gado de Corte – MS.
- A região da apomixia foi identificada com a presença de um *locus* segregante.
- Um total de 46 QTLs inéditos para *M. maximus* foram detectados, sendo 10 QTLs para caracteres agronômicos e 36 QTLs relacionados a qualidade da forragem.
- As fortes correlações entre os caracteres analisados permitiram encontrar zonas *hotspots* de QTLs.

Resultados Complementares

- A proximidade filogenética do genoma de referência utilizado aliado aos filtros de qualidade no conjunto de SNPs foram determinantes para a construção do mapa de ligação, em relação a fatores como maior quantidade de SNPs e/ou maior porcentagem de alinhamento total.
- A aplicação de vários filtros de qualidade no conjunto de SNPs promoveu a diminuição de SNPs redundantes e com distorção de segregação, sendo uma alternativa para a construção de mapas em poliploides.
- Recentes programas, como updog e PolymapR, se mostraram promissores para lidar com dados de NGS para *M. maximus*.

7 CONCLUSÕES

No cenário em que forrageiras poliploides exercem importante atividade econômica, principalmente em países como o Brasil, que é movido pelo setor da agropecuária, o conhecimento genético molecular sobre estas culturas deixa de ser uma opção. A poliploidia e o modo reprodutivo por apomixia, presentes em gramíneas forrageiras, são algumas das várias lacunas a serem desvendadas pela ciência. Esta tese buscou avançar o conhecimento sobre a genética da gramínea forrageira tropical *M. maximus* em vários aspectos, cujos resultados poderão ser inseridos no programa de melhoramento da espécie. As consequências práticas da distorção de segregação proporcionada pela redução dupla ainda não estão bem elucidadas em progênies de autotetraploides, bem como considerá-la em programas de melhoramento. Porém, termos a comprovação da sua existência nas progênies de *M. maximus* poderá contribuir no entendimento da distribuição e expressão dos caracteres, para os quais antes não havia resposta sobre a singularidade de alguns híbridos em relação à expressão de determinado caráter. O mapa de ligação obtido com marcadores de diferentes dosagens, proporcionou uma base a estudos de mapeamento mais acurados. O mapeamento do *locus* da apomixia por análise qualitativa reiterou que poucos genes são responsáveis pelo modo reprodutivo. A análise de similaridade dessa região também contribuiu com nossos resultados, pois encontramos poucos genes presentes nela, sendo dois genes relacionados à rede de microtúbulos e centrômero, responsáveis pela separação das cromátides durante a divisão celular e um gene presente na embriogênese somática. Possivelmente, essa região possui uma baixa taxa de recombinação devido a função desses genes. A identificação da arquitetura genômica de vários caracteres envolvidos na produção e qualidade nutricional da forragem e a análise de correlações sugerem uma complexa rede gênica. Além disso, QTLs de um caráter presente em ambos os genitores sugerem um arraste de ligação (em inglês, *drag linkage*), que poderá ser quebrado pelo programa de melhoramento (Wu et al., 2015), como proteína bruta nas folhas (CP_L), celulose nas folhas (CEL_L) e no colmo (CEL_S), lignina em permanganato no colmo (PL_S), sílica no colmo (SIL_S) e porcentagem de folhas em relação aos colmos (PLB). Já no caso do QTL ser derivado apenas de um genitor, cujo caráter é destaque e benéfico ao programa de melhoramento, como matéria seca do colmo (SDM), digestibilidade *in vitro* do colmo (IVD_S), lignina em permanganato nas folhas (PL_L) na cv. Mombaça e digestibilidade *in vitro* nas folhas (IVD_L) em S10, são importantes estudos mais aprofundados sobre essas regiões para verificar a expressão gênica. Por fim, este trabalho trouxe um conhecimento genético robusto que poderá ser útil também para estudos evolutivos de poliploides, bem como em programas

de melhoramento de outras plantas autotetraploides e para várias espécies de gramíneas forrageiras.

8 PERSPECTIVAS

A partir das análises de ligação dos marcadores com dados de dosagem alélica, o indício do evento de redução dupla como uma forma de segregação na gramínea em estudo, é uma abertura à frente de novas pesquisas voltadas ao conhecimento mitótico, meiótico e de herança polissômica. Como também, o mapa de ligação construído pode fornecer um embasamento na montagem do genoma da espécie. A partir das regiões de QTLs e da aposporia detectadas, inúmeras direções podem ser tomadas, a começar pela verificação de marcadores moleculares ligados aos caracteres quanto à sua eficiência, para posterior implementação no programa de melhoramento de *M. maximus* e até de outras gramíneas. Inclusive, a validação dos marcadores que foram detectados próximos ao *locus* da aposporia, em conjunto com outros marcadores intrinsecamente ligados a essa região presentes na literatura, pode ser implementado a seleção assistida por marcadores ao programa. Já em relação aos genes identificados, estudos de filogenômica podem proporcionar um maior conhecimento acerca da genética evolutiva da espécie, bem como, validá-los na espécie e averiguar as relações existentes entre eles para suportar a hipótese de QTLs *hotspots* ou novas suposições pertinentes. O uso de bibliotecas BACs (*bacterial artificial chromosome*) tem ganhado espaço no reconhecimento de genes a partir de estudos em QTLs para gramíneas forrageiras temperadas e em outras plantas de genoma complexo, podendo ser uma alternativa viável para o nosso caso. Além disso, a constante atualização das tecnologias de sequenciamento e de métodos estatísticos, com maior ênfase à realidade poliploide, poderão avançar mais os estudos de mapeamento em *M. maximus*, como por exemplo, o uso de *software* que suporte mapeamento múltiplo por intervalo, ou seja, a detecção de mais de um QTL sobre um caráter no mesmo cromossomo. Outras análises ainda podem ser exploradas, como realizar o mapeamento de QTLs de acordo com as avaliações fenotípicas para cada corte separadamente, com o intuito de verificar se os QTLs aqui identificados se mantêm estáveis a cada corte/avaliação. E, até mesmo, é possível realizar uma análise verificando a estabilidade dos QTLs para o período das águas e da seca. Por meio dos recentes avanços metodológicos, esses QTLs mapeados também podem ser empregados em trabalhos da espécie sobre seleção genômica e associação genômica ampla (GWAS), visando maior acurácia dos modelos.

9 REFERENCIAS BIBLIOGRAFICAS

- ABIEC-Associação Brasileira das Indústrias Exportadoras de Carne. (2019). Beef Report Perfil da pecuária no Brasil [online]. Disponível em: <http://www.abiec.com.br/controle/uploads/arquivos/sumario2019portugues.pdf>.
- Aguilera-Alvarado, G. P., and Sánchez-Nieto, S. (2017). Plant hexokinases are multifaceted proteins. *Plant Cell Physiol.* 58(7), 1151-1160.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716-723.
- Akiyama, Y., Goel, S., Conner, J. A., Hanna, W. W., Yamada-Akiyama, H., & Ozias-Akins, P. (2011). Evolution of the apomixis transmitting chromosome in *Pennisetum*. *BMC Evol. Biol.* 11(1), 289.
- Akiyama, Y., Yamada-Akiyama, H., Yamanouchi, H., Takahara, M., Ebina, M., Takamizo, et al. (2008). Estimation of genome size and physical mapping of ribosomal DNA in diploid and tetraploid guineagrass (*Panicum maximum* Jacq.). *Grassland science*, 54(2), 89-97.
- Albertini, E., Marconi, G., Reale, L., Barcaccia, G., Porceddu, A., Ferranti, F., and Falcinelli, M. (2005). SERK and APOSTART. Candidate genes for apomixis in *Poa pratensis*. *Plant Physiol.* 138(4), 2185-2199.
- Albrecht, C., Russinova, E., Kemmerling, B., Kwaaitaal, M., and de Vries, S. C. (2008). Arabidopsis SOMATIC EMBRYOGENESIS RECEPTOR KINASE proteins serve brassinosteroid-dependent and-independent signaling pathways. *Plant physiol.* 148(1), 611-619.
- Allard, R. W. (1971). Princípios do melhoramento genético das plantas. São Paulo: Edgard Blüchner, 381p.
- Anderson, N. A., Tobimatsu, Y., Ciesielski, P. N., Ximenes, E., Ralph, J., et al. (2015). Manipulation of guaiacyl and syringyl monomer biosynthesis in an Arabidopsis cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and modified cell wall structure. *Plant Cell.* 27(8), 2195-2209.
- Anhalt, U. C. M., Heslop-Harrison, P., Byrne, S., Guillard, A., and Barth, S. (2008). Segregation distortion in *Lolium*: evidence for genetic effects. *Theor. Appl. Genet.* 117, 297-306.
- Bailey, N. T. (1961). Introduction to the mathematical theory of genetic linkage. *The Clarendon Pms, Oxford*.
- Bannenberg, G., Martínez, M., Hamberg, M., and Castresana, C. (2009). Diversity of the enzymatic activity in the lipoxygenase gene family of *Arabidopsis thaliana*. *Lipids*, 44(2), 85.
- Batello, C., Mannelje, L., Martinez, A., & Suttie, J. (2008). Plant genetic resources of forage crops, pasture and rangelands. Thematic background study. *FAO report*, 5-7.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., Wheeler, D. L. (2005). GenBank. *Nucleic. Acids Res.*, 33: D34–D38.
- Blischak, P. D., Kubatko, L. S., & Wolfe, A. D. (2017). SNP genotyping and parameter estimation in polyploids using low-coverage sequencing data. *Bioinformatics*, 34(3), 407-415.
- Bluma-Marques, A. C., Chiari, L., Agnes, D. C., Jank, L., & Pagliarini, M. S. (2014). Molecular markers linked to apomixis in *Panicum maximum* Jacq. *Afr. J. Biotechnol.* 13, 2198-2202.
- Bourke, P. M., Hackett, C. A., Voorrips, R. E., Visser, R. G., & Maliepaard, C. (2019). Quantifying the power and precision of QTL analysis in autopolyploids under bivalent and multivalent genetic models. *G3*, g3-400269.
- Bourke, P. M., van Geest, G., Voorrips, R. E., Jansen, J., Kranenburg, T., Shahin, A., et al. (2018a). polymapR-linkage analysis and genetic map construction from F1 populations of outcrossing polyploids. *Bioinformatics*. 34, 3496-3502.
- Bourke, P. M., Voorrips, R. E., Visser, R. G., & Maliepaard, C. (2018b). Tools for genetic studies in experimental populations of polyploids. *Front. Plant Sci.* 9, 513.

- Bourke, P. M., Voorrips, R. E., Visser, R. G. F., and Maliepaard, C. (2015). The double-reduction landscape in tetraploid potato as revealed by a high-density linkage map. *Genetics* 201, 853-863.
- Box, G. E. P., and Cox, D. R. (1964). An analysis of transformations. *J. R. Stat. Soc. Ser. B (Methodol.)* 26, 211-252.
- Bradshaw, J. E. (2007). The canon of potato science: 4. Tetrasomic inheritance. *Potato Res.* 50, 219-222.
- Braz, T. G. D. S., Fonseca, D. M. D., Jank, L., Resende, M. D. V. D., Martuscello, J. A., and Simeão, R. M. (2013). Genetic parameters of agronomic characters in *Panicum maximum* hybrids. *Rev. Bras. Zootecnia*, 42(4), 231-237.
- Braz, T. G. D. S., Martuscello, J. A., Jank, L., Fonseca, D. M. D., Resende, M. D. V., and Evaristo, A. B. (2017). Genotypic value in hybrid progenies of *Panicum maximum* Jacq. *Ciênc. Rural* 47, e20160599.
- Burke, S. V., Wysocki, W. P., Zuloaga, F. O., Craine, J. M., Pires, J. C., Edger, P. P., et al. (2016). Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). *BMC Plant Biology*, 16(1), 140.
- Butler, D. G., Cullis, B. R., Gilmour, A. R., and Gogel, B. J. (2009). ASREML-R Reference Manual. Release 3.0. Australia: Queensland Department of Primary Industries.
- Butruille, D. V., and Boiteux, L. S. (2000). Selection-mutation balance in polysomic tetraploids: impact of double reduction and gametophytic selection on the frequency and subchromosomal localization of deleterious mutations. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6608-6613.
- Cartwright, D. A., Troggo, M., Velasco, R., and Gutin, A. (2007). Genetic mapping in the presence of genotyping errors. *Genetics* 176, 2521-2527.
- Chen, L., Guan, L., Kojima, A., and Adachi, T. (2000). The mechanisms of appearance of aposporous initial cell and apomictic embryo sac formation in *Paspalum notatum*. *Cytologia* 65, 333-341.
- Chrost, B., Kolukisaoglu, U., Schulz, B., and Krupinska, K. (2007). An α -galactosidase with an essential function during leaf development. *Planta*, 225(2), 311-320.
- Combes, D. (1975) Polymorphisme et modes de reproduction dans la section des Maximae du genre *Panicum* (Gramineae) en Afrique. *Mém ORSTOM* 77:1-99.
- Combes, D. & J. Pernès, 1970. Variations dans les nombres chromosomiques du *Panicum maximum* Jacq. en relation avec le mode de reproduction. *C. R. Seances Acad. Sci. D.* 270: 782-785.
- Conner, J. A., Goel, S., Gunawan, G., Cordonnier-Pratt, M. M., Johnson, V. E., et al. (2008) Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of *Pennisetum* and *Cenchrus*. *Plant Physiol.* 147: 1396-1411.
- Corratgé-Faillie, C., and Lacombe, B. (2017). Substrate (un) specificity of Arabidopsis NRT1/PTR FAMILY (NPF) proteins. *J. Exp. Bot.* 68(12), 3107-3113.
- Corrêa, L. A. & Santos, P. M. Manejo e utilização de plantas forrageiras dos gêneros *Panicum*, *Brachiaria* e *Cynodon*. São Carlos: Embrapa Pecuária Sudeste, 2003. 36 p. (Embrapa Pecuária Sudeste. *Documentos*, 34).
- Coser, A. C. & Pereira, A. V. Forrageiras para corte e pastejo. Juiz de Fora: Embrapa Gado de Leite, 2001. 37 p. (Embrapa Gado de Leite. *Circular técnica*, 66).
- Costa, N. R., Andreotti, M., Crusciol, C. A. C., Lima, C. G. D. R., Castilhos, A. M. D., et al. (2017). Yield and nutritive value of the silage of corn intercropped with tropical perennial grasses. *Pesqui. Agropecu. Bras.* 52(1), 63-73.
- Cullis, B. R., Smith, A. B., and Coombes, N. E. (2006). On the design of early generation variety trials with correlated data. *Journal of agricultural, biological, and environmental statistics.* 11(4), 381.

- Damm, T., Commandeur, U., Fischer, R., Usadel, B., and Klose, H. (2016). Improving the utilization of lignocellulosic biomass by polysaccharide modification. *Process Biochemistry*, 51(2), 288-296.
- Daverdin, G., Bahri, B. A., Wu, X., Serba, D. D., Tobias, C., Saha, M. C., et al. (2015). Comparative relationships and chromosome evolution in switchgrass (*Panicum virgatum*) and its genomic model, foxtail millet (*Setaria italica*). *BioEnerg. Res.* 8, 137-151.
- Davin, L. B., and Lewis, N. G. (2003). An historical perspective on lignan biosynthesis: monolignol, allylphenol and hydroxycinnamic acid coupling and downstream metabolism. *Phytochemistry Reviews*, 2(3), 257.
- Day, I. S., Reddy, V. S., Ali, G. S., and Reddy, A. S. N. (2002). Analysis of EF-hand-containing proteins in *Arabidopsis*. *Genome Biol.* 3(10), research0056-1.
- Deng, S., Wu, X., Wu, Y., Zhou, R., Wang, H., Jia, J., and Liu, S. (2011). Characterization and precise mapping of a QTL increasing spike number with pleiotropic effects in wheat. *Theor. Appl. Genet.* 122(2), 281-289.
- Dettmer, J., Schubert, D., Calvo-Weimar, O., Stierhof, Y. D., Schmidt, R., and Schumacher, K. (2005). Essential role of the V-ATPase in male gametophyte development. *Plant J.* 41(1), 117-124.
- Dimkpa, S. O., Lahari, Z., Shrestha, R., Douglas, A., Gheysen, G., and Price, A. H. (2015). A genome-wide association study of a global rice panel reveals resistance in *Oryza sativa* to root-knot nematodes. *J. Exp. Bot.* 67(4), 1191-1200.
- Ding, W., Lin, L., Zhang, B., Xiang, X., Wu, J., Pan, Z., and Zhu, S. (2015). OsKASI, a β -ketoacyl-[acyl carrier protein] synthase I, is involved in root development in rice (*Oryza sativa* L.). *Planta*, 242(1), 203-213.
- Doerge, R. W. (2002). Mapping and analysis of quantitative trait *loci* in experimental populations. *Nat. Rev. Genet.* 3, 43-52.
- Ebina, M., Nakagawa, H., Yamamoto, T., Araya, H., Tsuruta, S. I., Takahara, M., et al. (2005). Co-segregation of AFLP and RAPD markers to apospory in guinea grass (*Panicum maximum* Jacq.). *Grassl. Sci.* 51, 71-78.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One.* 6(5), e19379.
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V.D. and Borsboom, D. (2012) qgraph: network visualizations of relationships in psychometric data. *J Stat Softw* 48:1–18.
- Euclides, V. P. B., Thiago, L. R. L. D. S., Macedo, M. C. M., and Oliveira, M. P. D. (1999). Consumo voluntário de forragem de três cultivares de *Panicum maximum* sob pastejo. *Rev. Bras. Zootec.* 28, 1177-1185.
- Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends in plant science*, 5(5), 199-206.
- Falcioni, R., Moriwaki, T., Oliveira, D. M. D., Andreotti, G. C., Souza, L. A., et al. (2018). Increased gibberellins and light levels promotes cell wall thickness and enhance lignin deposition in xylem fibers. *Front. Plant Sci.* 9, 1391.
- Fang, X., Dong, K., Wang, X., Liu, T., He, J., Ren, R., et al. (2016). A high density genetic map and QTL for agronomic and yield traits in Foxtail millet [*Setaria italica* (L.) P. Beauv.]. *BMC Genomics.* 17(1), 336.
- FAOSTAT – Food and agriculture organization of the united nations statistics. Disponível em: <http://www.fao.org/faostat/en/#data/QC> (último acesso em 14 de outubro de 2019).
- Favero, A. P. Avaliação da citologia, genética e biologia reprodutiva de espécies de *Paspalum* para o estabelecimento de um programa de hibridações. São Carlos, SP: Embrapa Pecuária Sudeste, 2019. 22 p. (Embrapa Pecuária Sudeste - *Circular Técnica*).

- Feng, H., Chen, Q., Feng, J., Zhang, J., Yang, X., and Zuo, J. (2007). Functional characterization of the Arabidopsis eukaryotic translation initiation factor 5A-2 that plays a crucial role in plant growth and development by regulating cell division, cell growth, and cell death. *Plant Physiol.*, 144(3), 1531-1545.
- Feng, Y., Yin, Y., and Fei, S. (2015). Down-regulation of BdBRI1, a putative brassinosteroid receptor gene produces a dwarf phenotype with enhanced drought tolerance in *Brachypodium distachyon*. *Plant Sci.* 234, 163-173.
- Ferreira, R. C. U., Lara, L. A. D. C., Chiari, L., Barrios, S. C. L., do Valle, C. B., Valério, et al. (2019). Genetic mapping with allele dosage information in tetraploid *Urochloa decumbens* (Stapf) R. D. Webster reveals insights into spittlebug (*Notozulia entriana* Berg) resistance. *Front. Plant Sci.* 10, 1-16.
- Field, D. L., Broadhurst, L. M., Elliott, C. P., and Young, A. G. (2017). Population assignment in autopolyploids. *Heredity* 119, 389-401.
- Fierst, J. L. (2015). Using linkage maps to correct and scaffold de novo genome assemblies: methods, challenges, and computational tools. *Front. Genet.* 6, 220.
- Figueiredo, U. J., Berchembrock, Y. V., Valle, C. B., Barrios, S. C. L., Quesenberry, K. H., Muñoz, P. R., et al. (2019). Evaluating early selection in perennial tropical forages. *Crop Breed. Appl. Biotechnol.* 19 (3), 291–299.
- Fisher, R. A. (1947). The theory of linkage in polysomic inheritance. *Philos Trans R Soc Lond B Biol Sci.* 233, 55–87.
- Flint-Garcia, S. A., Thornsberry, J. M., and Buckler, E. S. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357-374.
- Friml, J., Wiśniewska, J., Benková, E., Mendgen, K., and Palme, K. (2002). Lateral relocation of auxin efflux regulator PIN3 mediates tropism in Arabidopsis. *Nature*, 415(6873), 806.
- Geisler, M., and Murphy, A. S. (2006). The ABC of auxin transport: the role of p-glycoproteins in plant development. *FEBS letters*, 580(4), 1094-1102.
- Gerard, D., Ferrão, L. F. V., Garcia, A. A. F., & Stephens, M. (2018). Genotyping polyploids from messy sequencing data. *Genetics.* 210, 789-807.
- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PLoS One.* 9, e90346.
- Goel, S., Chen, Z., Conner, J. A., Akiyama, Y., Hanna, W. W. & Ozias-Akins, P. (2003) Physical evidence that a single hemizygous chromosomal region is sufficient to confer aposporous embryosac formation in *Pennisetum squamulatum* and *Cenchrus ciliaris*. *Genetics.* 163:1069-1082.
- Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178-D1186.
- Gualtieri, G., Conner, J. A., Morishige, D. T., Moore, L. D., Mullet, J. E., and Ozias-Akins, P. (2006). A segment of the apospory-specific genomic region is highly microsyntenic not only between the apomicts *Pennisetum squamulatum* and buffelgrass, but also with a rice chromosome 11 centromeric-proximal genomic region. *Plant Physiol.* 140, 963-971.
- Hackett, C. A., Boskamp, B., Vogogias, A., Preedy, K. F., & Milne, I. (2017). TetraploidSNPMap: software for linkage analysis and QTL mapping in autotetraploid populations using SNP dosage data. *J. Hered.* 108, 438-442.
- Hackett, C. A., Bradshaw, J. E., & Bryan, G. J. (2014). QTL mapping in autotetraploids using SNP dosage information. *Theor. Appl. Genet.* 127, 1885-1904.
- Hackett, C. A., Bradshaw, J. E., Mayer, R. C., McNicol, J. W. & Milbourne, D. (1998). Linkage analysis in tetraploid species: a simulation study. *Genet. Res.* 71, 143–154.

- Hála, M., Cole, R., Synek, L., Drdová, E., Pečenková, T., Nordheim, A., et al. (2008). An exocyst complex functions in plant cell growth in *Arabidopsis* and tobacco. *Plant Cell*. 20(5), 1330-1345.
- Haldane, J. B. S. (1930). Theoretical genetics of autopolyploids. *J. Genet.* 22, 359-372.
- Hansen, S. F., Harholt, J., Oikawa, A., and Scheller, H. V. (2012). Plant glycosyltransferases beyond CAZy: a perspective on DUF families. *Front. Plant Sci.* 3, 59.
- Haynes, K. G., and Douches, D. S. (1993). Estimation of the coefficient of double reduction in the cultivated tetraploid potato. *Theor. Appl. Genet.* 85, 857-862.
- Hecht, V., Vielle-Calzada, J. P., Hartog, M. V., Schmidt, E. D., Boutilier, K., Grossniklaus, U., and de Vries, S. C. (2001). The *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 gene is expressed in developing ovules and embryos and enhances embryogenic competence in culture. *Plant Physiol.* 127(3), 803-816.
- Hedden P., and Thomas S. G. (2016). Annual Plant Review: The Gibberellins, 1st Edn, Vol. 49 New York, NY: Wiley Blackwell.
- Hemsley, P. A., Kemp, A. C., and Grierson, C. S. (2005). The TIP GROWTH DEFECTIVE1 S-acyl transferase regulates plant cell growth in *Arabidopsis*. *Plant Cell*. 17(9), 2554-2563.
- Hodkinson, T. R. (2018). Evolution and Taxonomy of the Grasses (Poaceae): A Model Family for the Study of Species-Rich Groups. *Annu. Plant Rev. Online.* 1-39.
- Höhner, R., Aboukila, A., Kunz, H. H., and Venema, K. (2016). Proton gradients and proton-dependent transport processes in the chloroplast. *Front. Plant Sci.* 7, 218.
- Holland, J. B. (2007). Genetic architecture of complex traits in plants. *Curr. Opin. Plant Biol.* 10 (2), 156-161.
- Holzman, R., and Hulsey, C. D. (2017). Mechanical transgressive segregation and the rapid origin of trophic novelty. *Scientific Rep.* 7, 40306.
- Huang, K., Wang, T., Dunn, D. W., Zhang, P., Cao, X., Liu, R., et al. (2019). Genotypic frequencies at equilibrium for polysomic inheritance under double-reduction. *G3 (Bethesda)* 9, 1693-1706.
- Huo, H., Conner, J. A., & Ozias-Akins, P. (2009). Genetic mapping of the apospory-specific genomic region in *Pennisetum squamulatum* using retrotransposon-based molecular markers. *Theor. Appl. Genet.* 119(2), 199-212.
- ITC - International Trade Centre. (2018). Trade map - List of exporters for the Seeds of forage plants for sowing. [Online]. Available at: https://www.trademap.org/Country_SelProduct_TS.aspx. (Accessed June 27, 2019).
- Jank, L., R.A. Barbosa, C.M.S. de Andrade, J.R. Valério, M.C.M. Macedo, et al. 2017. O capim-BRS Quênia (*Panicum maximum* Jacq.) na diversificação e intensificação das pastagens. Brasília, DF.
- Jank, L., Barrios, S. C., do Valle, C. B., Simeão, R. M., & Alves, G. F. (2014). The value of improved pastures to Brazilian beef production. *Crop Pasture Sci.* 65(11), 1132-1137.
- Jank, L., Valle, C. B., Simeão, R. M., Chiari, L., Jungmann, L., & Xavier, K. G. Forage Breeding at Embrapa Beef Cattle (2013). In: Jank, L., Chiari, L., Valle, C. B., & Simeão, R. M. *Forage breeding and biotechnology*. 1.ed. Campo Grande: Embrapa, p.1-23.
- Jank, L., Valle, C. D., & Resende, R. M. S. (2011). Breeding tropical forages. *Crop Breed. Appl. Biotechnol.*, 11(SPE), 27-34.
- Jank, L., Resende, R. M. S., Valle, C. B., et al. Melhoramento genético de *Panicum maximum* (2008). In: Resende, R. M. S., Valle, C. B., & Jank, L. *Melhoramento de Forrageiras Tropicais*. 1.ed. Campo Grande: Embrapa, p.55-87.

- Jank, L., Calixto, S., Costa, J. C. G., Savidan, Y. H., & Curvo, J. B. E. (1997) Catalog of the characterization and evaluation of the *Panicum maximum* germplasm: morphological description and agronomical performance. Campo Grande, Embrapa Gado de Corte. 53p. (*Documentos*, 68).
- Jank, L.; Savidan, Y.; Souza, M.T.; & Costa, J. G. C. (1994). Avaliação do germoplasma de *Panicum maximum* introduzido da África. 1. Produção forrageira. *R. Bras. Zootec.* 23, 433-440.
- Jia, G., Wang, H., Tang, S., Zhi, H., Liu, S., Wen, Q., et al. (2017). Detection of genomic *loci* associated with chromosomal recombination using high-density linkage mapping in *Setaria*. *Scientific Rep.* 7, 1-8.
- Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., et al. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, 473(7345), 97.
- Julier, B., Flajoulot, S., Barre, P., Cardinet, G., Santoni, S., Huguet, T., et al. (2003). Construction of two genetic linkage maps in cultivated tetraploid alfalfa (*Medicago sativa*) using microsatellite and AFLP markers. *BMC Plant Biol.* 3, 9.
- Jung, H. G., and Allen, M. S. (1995). Characteristics of plant cell walls affecting intake and digestibility of forages by ruminants. *J. Anim. Sci.* 73, 2774-2790.
- Kachroo, A., Shanklin, J., Whittle, E., Lapchyk, L., Hildebrand, D., and Kachroo, P. (2007). The Arabidopsis stearyl-acyl carrier protein-desaturase family and the contribution of leaf isoforms to oleic acid synthesis. *Plant Mol. Biol.* 63(2), 257-271.
- Kaushal, P., Malaviya, D. R., Roy, A. K., Pathak, S., Agrawal, A., Khare, A., et al. (2008). Reproductive pathways of seed development in apomictic guinea grass (*Panicum maximum* Jacq.) reveal uncoupling of apomixis components. *Euphytica* 164, 81–92.
- Kaushal, P., Dwivedi, K. K., Radhakrishna, A., Srivastava, M. K., Kumar, V., Roy, A. K., et al. (2019). Partitioning apomixis components to understand and utilize gametophytic apomixis. *Front. Plant Sci.* 10, 256.
- Khasanova, A., Lovell, J. T., Bonnette, J., Weng, X., Jenkins, J., Yoshinaga, et al. (2019). The genetic architecture of shoot and root trait divergence between mesic and xeric ecotypes of a perennial grass. *Front. Plant Sci.* 10, 366.
- Kirioukhova, O., Johnston, A. J., Kleen, D., Kägi, C., Baskar, R., Moore, J. M., et al. (2011). Female gametophytic cell specification and seed development require the function of the putative Arabidopsis INCENP ortholog WYRD. *Development*, 138(16), 3409-3420.
- Knopf, R. R., Feder, A., Mayer, K., Lin, A., Rozenberg, M., Schaller, A., and Adam, Z. (2012). Rhomboid proteins in the chloroplast envelope affect the level of allene oxide synthase in *Arabidopsis thaliana*. *Plant J.* 72(4), 559-571.
- Ko, J. H., Yang, S. H., and Han, K. H. (2006). Upregulation of an Arabidopsis RING-H2 gene, XERICO, confers drought tolerance through increased abscisic acid biosynthesis. *Plant J.* 47(3), 343-355.
- Kohorn, B. D., and Kohorn, S. L. (2012). The cell wall-associated kinases, WAKs, as pectin receptors. *Front. Plant Sci.* 3, 88.
- Köksal, M., Hu, H., Coates, R. M., Peters, R. J., and Christianson, D. W. (2011). Structure and mechanism of the diterpene cyclase ent-copalyl diphosphate synthase. *Nature Chem. Biol.* 7(7), 431.
- Kourelis, J., and van der Hoorn, R. A. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell.* 30(2), 285-299.
- Kumar, S. (2017). Epigenetic control of apomixis: a new perspective of an old enigma. *Adv. Plants Agric. Res.* 7, 00243.
- Kumar, S., Saxena, S., & Gupta, M. G. (2017a) Marker-assisted screening of F2 mapping population of *Cenchrus ciliaris* L. segregating for apomictic and sexual mode of reproduction. *Crop Breed. Appl. Biotechnol.* 17: 10-17.

- Kumar, S., Beena, A. S., Awana, M., & Singh, A. (2017b). Salt-induced tissue-specific cytosine methylation downregulates expression of HKT genes in contrasting wheat (*Triticum aestivum* L.) genotypes. *DNA Cell Biol.* 36(4), 283-294.
- Kumar, J., Gupta, D. S., Gupta, S., Dubey, S., Gupta, P., and Kumar, S. (2017). Quantitative trait *loci* from identification to exploitation for crop improvement. *Plant cell reports*, 36(8), 1187-1213.
- Kusano, H., Testerink, C., Vermeer, J. E., Tsuge, T., Shimada, H., Oka, A., et al. (2008). The Arabidopsis phosphatidylinositol phosphate 5-kinase PIP5K3 is a key regulator of root hair tip growth. *Plant Cell.* 20(2), 367-380.
- Langmead, B., Trapnell, C., Pop, M., et al. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology*, 10(3), R25.
- Lara, L. A., Santos, M. F., Jank, L., Chiari, L., Vilela, M. M., Amadeu, R. R., et al. (2019). Genomic selection with allele dosage in *Panicum maximum* Jacq. *G3*. g3.200986.202018.
- Layman, N. C., and Busch, J. W. (2018). Bottlenecks and inbreeding depression in autotetraploids. *Evolution* 72, 2025-2037.
- Leitch, A. R. & Leitch, I. J. Genomic plasticity and the diversity of polyploid plants. *Science*. 320, 481–483 (2008).
- Li, J., Das, K., Fu, G., Tong, C., Li, Y., Tobias, C., et al. (2010). EM algorithm for mapping quantitative trait *loci* in multivalent tetraploids. *Int. J. Plant Genom.* 2010, 1-10.
- Lima, M. A., Gomez, L. D., Steele-King, C. G., Simister, R., Bernardinelli, O. D., et al. (2014). Evaluating the composition and processing potential of novel sources of Brazilian biomass for sustainable biorenewables production. *Biotechnol. Biofuels*. 7(1), 10.
- Lin, T. C., Neuner A., and Schiebel, E. (2015). Targeting of g-tubulin complexes to microtubule organizing centers: Conservation and divergence. *Trends in Cell Biology*. 25:296–307.
- Lin, J. S., Huang, X. X., Li, Q., Cao, Y., Bao, Y., Meng, X. F., et al. (2016). UDP-glycosyltransferase 72B1 catalyzes the glucose conjugation of monolignols and is essential for the normal cell wall lignification in *Arabidopsis thaliana*. *Plant J.* 88(1), 26-42.
- Liu, J., Im Kim, J., Cusumano, J. C., Chapple, C., Venugopalan, N., Fischetti, R. F., and Makowski, L. (2016). The impact of alterations in lignin deposition on cellulose organization of the plant cell wall. *Biotechnology for biofuels*, 9(1), 126.
- Luo, Z. W., Hackett, C. A., Bradshaw, J. E., McNicol, J. W., & Milbourne, D. (2000). Predicting parental genotypes and gene segregation for tetrasomic inheritance. *Theor. Appl. Genet.* 100(7), 1067-1073.
- Luo, Z. W., Zhang, R. M., & Kearsley, M. J. (2004). Theoretical basis for genetic linkage analysis in autotetraploid species. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7040-7045.
- Luo, Z. W., Zhang, Z., Leach, L., Zhang, R. M., Bradshaw, J. E., and Kearsley, M. J. (2006). Constructing genetic linkage maps under a tetrasomic model. *Genetics* 172, 2635-2645.
- Ma, C. X., Casella, G., Shen, Z. J., Osborn, T. C., & Wu, R. (2002). A unified framework for mapping quantitative trait *loci* in bivalent tetraploids using single-dose restriction fragments: A case study from alfalfa. *Genome Res.* 12, 1974–1981.
- Marcón, F., Martínez, E. J., Rodríguez, G. R., Zilli, A. L., Brugnoli, E. A., and Acuña, C. A. (2019). Genetic distance and the relationship with heterosis and reproductive behavior in tetraploid bahiagrass hybrids. *Mol. Breed.* 39, 89.
- Marino, C. T., & Medeiros, S. R. (2015). Minerais e vitaminas na nutrição de bovinos de corte. In: Medeiros, S. R. de; Gomes, R. C.; & Bungenstab, D. J. *Nutrição de bovinos de corte: fundamentos e aplicações*. Brasília, DF: Embrapa.

- Martinez, E. J., Hopp, H. E., Stein, J., Ortiz, J. P. A., Quarin, C. L. (2003). Genetic characterization of apospory in tetraploid *Paspalum notatum* based on the identification of linked molecular markers. *Mol Breed.* 12:319–327.
- Mashiguchi, K., Asami, T., and Suzuki, Y. (2009). Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes in Arabidopsis. *Bioscience, biotechnology, and biochemistry.* 73(11), 2452-2459.
- Massa, A. N., Manrique-Carpintero, N. C., Coombs, J., Haynes, K. G., Bethke, P. C., et al. (2018). Linkage analysis and QTL mapping in a tetraploid russet mapping population of potato. *BMC Genet.* 19, 1-13.
- Mather, K. (1936). Segregation and linkage in autotetraploids. *J. Genet.* 32, 287–314.
- Matias, F. I., Alves, F. C., Meireles, K. G. X., Barrios, S. C. L., do Valle, C. B., Endelman, J. B., et al. (2019). On the accuracy of genomic prediction models considering multi-trait and allele dosage in Urochloa spp. interespecific tetraploid hybrids. *Mol. Breeding.* 39: 100.
- Medeiros, S. R., & Marino, C. T. (2015). Lipídios na nutrição de ruminantes. In: Medeiros, S. R., Gomes, R. C., & Bungenstab, D. J. *Nutrição de bovinos de corte: fundamentos e aplicações.* Brasília, DF: Embrapa.
- Medeiros, S. R., & Marino, C. T. (2015a) Valor nutricional dos alimentos na nutrição de ruminantes e sua determinação. In: Medeiros, S. R., Gomes, R. C., & Bungenstab, D. J. *Nutrição de bovinos de corte: fundamentos e aplicações.* Brasília, DF: Embrapa.
- Medeiros, S. R., & Marino, C. T. (2015b). Carboidratos na nutrição de gado de corte. In: Medeiros, S. R., Gomes, R. C., & Bungenstab, D. J. *Nutrição de bovinos de corte: fundamentos e aplicações.* Brasília, DF: Embrapa.
- Medeiros, S. R., & Marino, C. T. (2015c). Proteínas na nutrição de bovinos de corte. In: Medeiros, S. R., Gomes, R. C., & Bungenstab, D. J. *Nutrição de bovinos de corte: fundamentos e aplicações.* Brasília, DF: Embrapa.
- Mengist, M. F., Alves, S., Griffin, D., Creedon, J., McLaughlin, M. J., Jones, P. W., et al. (2018). Genetic mapping of quantitative trait *loci* for tuber-cadmium and zinc concentration in potato reveals associations with maturity and both overlapping and independent components of genetic control. *Theor. Appl. Genet.* 131, 929-945.
- Mertens, D. R. (2002). Gravimetric determination of amylase-treated neutral detergent fiber in feeds with refluxing in beakers or crucibles: collaborative study. *J. AOAC Int.*, 85(6), 1217-1240.
- Milano, E. R., Payne, C. E., Wolfrum, E., Lovell, J., Jenkins, J., Schmutz, J., et al. (2018). Quantitative trait *loci* for cell wall composition traits measured using near-infrared spectroscopy in the model C4 perennial grass *Panicum hallii*. *Biotechnol. Biofuels* 11, 25.
- Miles, C., and Wayne, M. (2008) Quantitative trait *locus* (QTL) analysis. *Nature Educ.* 1(1):208.
- Mohapatra, S., Mishra, C., Behera, S. S., & Thatoi, H. (2017). Application of pretreatment, fermentation and molecular techniques for enhancing bioethanol production from grass biomass – A review. *Renew. Sust. Energ. Rev.* 78, 1007-1032.
- Mollinari, M., & Garcia, A. A. F. (2019). Linkage analysis and haplotype phasing in experimental autopolyploid populations with high ploidy level using hidden Markov models. *G3.*
- Mollinari, M., Olukolu, B. A., Pereira, G. D. S., Khan, A., Gemenet, D., Yengo, C., & Zeng, Z. B. (2019). Unraveling the hexaploid sweetpotato inheritance using ultra-dense multilocus mapping. *bioRxiv*, 689638.
- Morais, R. F., Morais, C. S. B., de Morais, L. F., & Almeida, J. C. C. (2018). Energy balance of elephant grass biomass for power generation by direct biomass combustion. *Afr. J. Biotechnol.* 17(13), 405-410.
- Mousavi, A., and Hotta, Y. (2005). Glycine-rich proteins. *Applied biochemistry and biotechnology*, 120(3), 169-174.

- Munné-Bosch, S., Simancas, B., and Müller, M. (2018). Ethylene signaling cross-talk with other hormones in *Arabidopsis thaliana* exposed to contrasting phosphate availability: Differential effects in roots, leaves and fruits. *J. Plant Physiol.* 226, 114-122.
- Ndamukong, I., Chetram, A., Saleh, A., and Avramova, Z. (2009). Wall-modifying genes regulated by the *Arabidopsis* homolog of trithorax, ATX1: repression of the XTH33 gene as a test case. *Plant J.* 58(4), 541-553.
- Newman, K. L., Fernandez, A. G., and Barton, M. K. (2002). Regulation of axis determinacy by the *Arabidopsis* PINHEAD gene. *Plant Cell.* 14(12), 3029-3042.
- Niño-González, M., Novo-Uzal, E., Richardson, D. N., Barros, P. M., and Duque, P. (2019). More Transporters, More Substrates: The *Arabidopsis* Major Facilitator Superfamily Revisited. *Molecular plant.*
- Nogler, G. A. (1984). Gametophytic apomixis. In: *Embryology of Angiosperms*, ed. B. M. Johri (Berlin, Heidelberg: Springer), 475-518.
- Odorico, F. H., de Araújo Morandim-Giannetti, A., Lucarini, A. C., & Torres, R. B. (2018). Pretreatment of Guinea grass (*Panicum maximum*) with the ionic liquid 1-ethyl-3-methyl imidazolium acetate for efficient hydrolysis and bioethanol production. *Cellulose*, 25(5), 2997-3009.
- Oliveira, F. A., Vigna, B. B. Z., da Silva, C. C., Favero, A. P., de Pina Matta, F., Azevedo, A. L. S., & Souza, A. P. (2018). Apomixis-related genes identified from a coexpression network in *Paspalum notatum*, a Neotropical grass. *bioRxiv*, 369280.
- Oliveira, E. J., Koehler, A. D., Rocha, D. I., Vieira, L. M., Pinheiro, M. V. M., et al. (2017). Morpho-histological, histochemical, and molecular evidences related to cellular reprogramming during somatic embryogenesis of the model grass *Brachypodium distachyon*. *Protoplasma*, 254(5).
- Oliveira, M. C. Gramafante: uma nova cultivar de capim elefante para o Vale do São Francisco. (2001). Embrapa Semiárido. *Boletim de Pesquisa e Desenvolvimento.* 14.
- Ott, J., Wang, J., & Leal, S. M. (2015). Genetic linkage analysis in the age of whole-genome sequencing. *Nat. Rev. Genet.* 16(5), 275.
- Ozias-Akins, P., Roche, D., & Hanna, W. W. (1998) Tight clustering and hemizyosity of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus that may have no allelic form in sexual genotypes. *Proc. Natl. Acad. Sci. U.S.A.* 95: 5127–5132.
- Pan, H., Zhou, R., Louie, G. V., Mühlemann, J. K., Bomati, E. K., et al. (2014). Structural studies of cinnamoyl-CoA reductase and cinnamyl-alcohol dehydrogenase, key enzymes of monolignol biosynthesis. *Plant Cell.* 26(9), 3709-3727.
- Patel, R. K., and Jain, M. (2012). NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7, e30619.
- Paudel, D., Kannan, B., Yang, X., Harris-Shultz, K., Thudi, M., Varshney, R. K., et al. (2018). Surveying the genome and constructing a high-density genetic map of napiergrass (*Cenchrus purpureus* Schumach). *Scientific Rep.* 8, 14419.
- Pelloux, J., Rusterucci, C., and Mellerowicz, E. J. (2007). New insights into pectin methylesterase structure and function. *Trends in plant science.* 12(6), 267-277.
- Pereira, G. D. S., Gemenet, D. C., Mollinari, M., Olukolu, B. A., Wood, J. C., Diaz, F., et al. (2019). Multiple QTL mapping in autopolyploids: a random-effect model approach with application in a hexaploid sweetpotato full-sib population. *bioRxiv*, 622951.
- Pereira, G. S., Garcia, A. A. F., & Margarido, G. R. A. (2018). A fully automated pipeline for quantitative genotype calling from next generation sequencing data in autopolyploids. *BMC Bioinform.* 19, 398.

- Pereira, A. V., Valle, C. B., Ferreira, R. P., & Miles, J. W. Melhoramento de forrageiras tropicais. (2001). In: Nass, L. L., Valois, A. C. C., Melo, I. S., & Valadares-Inglis, M. C. *Recursos genéticos e melhoramento de plantas*. Rondonópolis: Fundação MT. 549-601.
- Perkel, J. SNP genotyping: six technologies that keyed a revolution. *Nature*. (2008): 447.
- Pernès, J. (1975) Organization évolutive d'un groupe agamique: la section des maximae du genre *Panicum* (Graminées). Paris: ORSTOM. 106 p. (*Mémoires ORSTOM*, 75).
- Pessim, C., Pagliarini, M. S., Jank, L., Kaneshima, A. M. D. S., and Bonato, A. B. M. (2010). Meiotic behavior in *Panicum maximum* Jacq. (Poaceae: Panicoideae: Paniceae): hybrids and their genitors. *Acta Sci. Agron*. 32(3), 417-422.
- Pessim, C., Pagliarini, M. S., Silva, N., and Jank, L. (2015). Chromosome stickiness impairs meiosis and influences reproductive success in *Panicum maximum* (Poaceae) hybrid plants. *Genet. Mol. Res.* 14(2), 4195-4202.
- Pessino, S. C., Evans, C., Ortiz, J. P. A., Armstead, I., Valle, C. B. D., and Hayward, M. D. (1998). A genetic map of the apospory-region in *Brachiaria* hybrids: identification of two markers closely associated with the trait. *Hereditas* 128, 153-158.
- Pessino, S. C., Ortiz, J., Leblanc, O., Valle, C. B., & Hayward, M. D. (1997) Identification of a maize linkage group related to apomixis in *Brachiaria*. *Theor. Appl. Genet.* 94: 439-444.
- Poland, J. A., & Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome*. 5(3), 92-102.
- Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One*. 7, e32253.
- Primo, C., Pizzio, G. A., Yang, J., Gaxiola, R. A., Scholz-Starke, J., and Hirschi, K. D. (2019). Plant proton pumping pyrophosphatase: the potential for its pyrophosphate synthesis activity to modulate plant growth. *Plant Biology*.
- Punnuri, S. M., Wallace, J. G., Knoll, J. E., Hyma, K. E., Mitchell, S. E., Buckler, E. S., et al. (2016). Development of a high-density linkage map and tagging leaf spot resistance in pearl millet using genotyping-by-sequencing markers. *Plant Genome* 9, 1-13.
- Qu, L., & Hancock, J. F. (2001). Detecting and mapping repulsion-phase linkage in polyploids with polysomic inheritance. *Theor. Appl. Genet.* 103(1), 136-143.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rabello, C. F. (2017). *Caracterização genético-molecular em Panicum maximum: identificação de marcadores moleculares SNPs e mapeamento genético*. Dissertação (Mestrado em Genética e Biologia Molecular) – Instituto de Biologia, Universidade Estadual de Campinas.
- Radhakrishna, A., Dwivedi, K. K., Srivastava, M. K., Roy, A. K., Malaviya, D. R., & Kaushal, P. (2018). Transcriptomic data of pre-meiotic stage of floret development in apomictic and sexual types of guinea grass (*Panicum maximum* Jacq.). *Data brief*. 18, 590-593.
- Rajput, S. G., Santra, D. K., & Schnable, J. (2016). Mapping QTLs for morpho-agronomic traits in proso millet (*Panicum miliaceum* L.). *Mol. Breed.* 36, 1-18.
- Reis, P. A., da Silva, A. L. H., da Silva, A. S., Valls, J. F. M., Gomes, S. M., & Oliveira, R. C. (2015). *Paspalum minutispiculatum* (Poaceae, Panicoideae): A new annual species from Central Brazil. *Syst. Bot.* 40(3), 746-754.

- Resende, R. M. S., Jank, L., Valle, C. B. D., and Bonato, A. L. V. (2004). Biometrical analysis and selection of tetraploid progenies of *Panicum maximum* using mixed model methods. *Pesq Agropec Bras.* 39, 335-341.
- Roberts, D., Pedmale, U. V., Morrow, J., Sachdev, S., Lechner, E., Tang, X., et al. (2011). Modulation of phototropic responsiveness in Arabidopsis through ubiquitination of phototropin 1 by the CUL3-Ring E3 ubiquitin ligase CRL3NPH3. *Plant Cell.* 23(10), 3627-3640.
- Routray, P., Masalkar, P. D., Roberts, D. M. (2015). "Nodulin intrinsic proteins: facilitators of water and ammonia transport across the symbiosome membrane" in Biological Nitrogen Fixation, ed. F. J. de Bruijn (John Wiley & Sons, Inc.), 695–704.
- Sapkota, S., Conner, J. A., Hanna, W. W., Simon, B., Fengler, K., Deschamps, S., et al. (2016). *In silico* and fluorescence in situ hybridization mapping reveals collinearity between the *Pennisetum squamulatum* apomixis carrier-chromosome and chromosome 2 of sorghum and foxtail millet. *PLoS one*, 11(3), e0152411.
- Sartie, A. M., Easton, H. S., Matthew, C., Rolston, M. P., & Faville, M. J. (2018). Seed yield in perennial ryegrass (*Lolium perenne* L.): comparative importance of component traits and detection of seed-yield-related QTL. *Euphytica.* 214(12), 226.
- Savidan, Y. (1980). Chromosomal and embryological analyses in sexual x apomictic hybrids of *Panicum maximum* Jacq. *Theor. Appl. Genet.* 58(2), 153-156.
- Savidan, Y. (1981). Genetics and utilization of apomixis for the improvement of guinea grass (*Panicum maximum*) In: International Grassland Congress, 14., 1981, Lexington. Proceedings ... Boulder: Westview Press, 1983. 182-184.
- Savidan, Y.H. (1982). Nature et hérédité de l'apomixie chez *Panicum maximum* Jacq. Paris: ORSTOM.
- Savidan, Y., & Pernès, J. (1982). Diploid-tetraploid-dihaploid cycles and the evolution of *Panicum maximum* jacq. *Evolution.* 36, 596-600.
- Savidan, Y. H., Jank, L., Costa, J. C. G., & Valle, C. B. (1989). Breeding *Panicum maximum* in Brazil. 1. Genetic resources, modes of reproduction and breeding procedures. *Euphytica*, 41(1-2), 107-112.
- Savidan, Y. (2000). Apomixis: genetics and breeding. *Plant Breed. Rev.* 18, 13-86.
- Schuster, I., & Cruz, C. D. (2005). Estatística genômica aplicada a populações derivadas de cruzamentos controlados. *Crop Breeding and Applied Biotechnology*, 5(2).
- Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461-464.
- Sehgal, D., Rajaram, V., Armstead, I. P., Vadez, V., Yadav, Y. P., Hash, C. T., and Yadav, R. S. (2012). Integration of gene-based markers in a pearl millet genetic map for identification of candidate genes underlying drought tolerance quantitative trait *loci*. *BMC Plant Biol.* 12(1), 9.
- Sengupta, S., Mukherjee, S., Basak, P., and Majumder, A. L. (2015). Significance of galactinol and raffinose family oligosaccharide synthesis in plants. *Front. Plant Sci.* 6, 656.
- Seo, P. J., Kim, S. G., and Park, C. M. (2008). Membrane-bound transcription factors in plants. *Trends in plant science*, 13(10), 550-556.
- Serang, O., Mollinari, M., & Garcia, A. A. F. (2012). Efficient exact maximum a posteriori computation for Bayesian SNP genotyping in polyploids. *PLoS One.* 7, e30906.
- Sharbel TF, Voigt ML, Corral JM, Galla G., Kumlehn J, et al. (2010). Apomictic and sexual ovules of *Boechera* display heterochronic global gene expression patterns. *Plant Cell.* 22, 655–671.
- Silva, S. C., Sbrissia, A. F., & Pereira, L. E. T. (2015). Ecophysiology of C4 forage grasses - understanding plant growth for optimising their use and management. *Agriculture*, 5(3), 598-625.

- Singh, J., Kalberer, S. R., Belamkar, V., Assefa, T., Nelson, M. N., et al. (2018). A transcriptome-SNP-derived linkage map of *Apios americana* (potato bean) provides insights about genome re-organization and synteny conservation in the phaseoloid legumes. *Theor. Appl. Genet.* 131(2), 333-351.
- Smyczynski, C., Roudier, F., Gissot, L., Vaillant, E., Grandjean, O., Morin, H., et al. (2006). The C terminus of the immunophilin PASTICCINO1 is required for plant development and for interaction with a NAC-like transcription factor. *J. Biol. Chem.* 281(35), 25475-25484.
- Soreng, R. J., Peterson, P. M., Romaschenko, K., Davidse, G., Teisher, J. K., et al. (2017). A worldwide phylogenetic classification of the Poaceae (Gramineae) II: An update and a comparison of two 2015 classifications. *J. Syst. Evol.* 55(4), 259-290.
- Sousa, A. C. B.; Jungmann, L.; Campos, T.; Sforça, D. A.; Boaventura, L. R.; Silva, G. M. B.; Zucchi, M. I.; Jank, L.; & Souza, A. P. (2011a). Development of microsatellite markers in Guineagrass (*Panicum maximum* Jacq.) and their transferability to other tropical forage grass species. *Plant Breed.* 130.
- Sousa, A. C. B., Jank, L., de Campos, T., Sforça, D. A., Zucchi, M. I., & Souza, A. P. (2011b). Molecular diversity and genetic structure of guineagrass (*Panicum maximum* Jacq.), a tropical pasture grass. *Trop. Plant Biol.* 4(3-4), 185-202.
- Studer, A. J., and Doebley, J. F. (2011). Do large effect QTL fractionate? A case study at the maize domestication QTL teosinte branched1. *Genetics.* 188(3), 673-681.
- Sturtevant, A.H. (1913). The linear arrangement of six sex-linked factors in *Drosophila*, as shown by their mode of association. *J. Exp. Zool. A Ecol. Genet. Physiol.* 14, 43-59.
- Suzuki, M., Kato, A., Nagata, N., and Komeda, Y. (2002). A xylanase, AtXyn1, is predominantly expressed in vascular bundles, and four putative xylanase genes were identified in the *Arabidopsis thaliana* genome. *Plant Cell Physiol.* 43(7), 759-767.
- Tabangin, M. E., Woo, J. G., & Martin L. J. (2009). The effect of minor allele frequency on the likelihood of obtaining false positives. *BMC Proc.* 3. (7).
- Thaikua, S., Ebina, M., Yamanaka, N., Shimoda, K., Suenaga, K., and Kawamoto, Y. (2016). Tightly clustered markers linked to an apospory-related gene region and quantitative trait loci mapping for agronomic traits in *Brachiaria* hybrids. *Grassl. Sci.* 62, 69-80.
- Toledo-Silva, G. (2013b). *Estudos genético-moleculares em Panicum maximum: mapeamento genético-molecular e análise de transcriptoma via RNA-seq.* Tese de doutorado em Genética e Biologia Molecular – Instituto de Biologia, Universidade Estadual de Campinas.
- Toledo-Silva, G., Cardoso-Silva, C. B., Jank, L., & Souza, A. P. (2013). *De novo* transcriptome assembly for the tropical grass *Panicum maximum* Jacq. *PLoS one*, 8(7).
- Tomaz, P. K., Araujo, L. C., Sanches, L. A., Santos-Araujo, S. N., de Lima, T. O., Lino, A. D. A., & Ferreira, E. M. (2018). Effect of sward height on the fermentability coefficient and chemical composition of Guinea grass silage. *Grass Forage Sci.* 73(3), 588-598.
- Tronconi, M. A., Fahnenstich, H., Weehler, M. C. G., Andreo, C. S., Flügge, U. I., Drincovich, M. F., and Maurino, V. G. (2008). Arabidopsis NAD-malic enzyme functions as a homodimer and heterodimer and has a major impact on nocturnal metabolism. *Plant Physiol.* 146(4), 1540-1552.
- Tucker, M. R., Araujo, A. C. G., Paech, N. A., Hecht, V., Schmidt, E. D., Rossell, J. B., et al. (2003). Sexual and apomictic reproduction in *Hieracium* subgenus *Pilosella* are closely interrelated developmental pathways. *Plant Cell.* 15(7), 1524-1537.
- Valle, C. B., Jank, L. & Resende, R. M. S. (2009). O melhoramento de forrageiras tropicais no Brasil. *Rev. Ceres.* 56, 460-472.
- Valle, C. D., Glienke, C., and Leguizamon, G. O. C. (1994). Inheritance of apomixis in *Brachiaria*, a tropical forage grass. *Apomixis Newsletter*, 7, 42-43.

- Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nat. Rev. Genet.* 18(7), 411.
- Van Geest, G., Bourke, P. M., Voorrips, R. E., Marasek-Ciolakowska, A., Liao, Y., Post, A., et al. (2017). An ultra-dense integrated linkage map for hexaploid chrysanthemum enables multi-allelic QTL analysis. *Theor. Appl. Genet.* 130 (12), 2527-2541.
- Vigna, B. B. Z., Santos, J. C. S., Jungmann, L., do Valle, C. B., Mollinari, M., et al. (2016). Evidence of allopolyploidy in *Urochloa humidicola* based on cytological analysis and genetic linkage mapping. *PLoS One* 11, e0153764.
- Vilela, D. *Potencial das pastagens de Cynodon na pecuária de leite*. In: Vilela, D.; Resende, J. C. & Lima, J. (Ed.). *Cynodon: forrageiras que estão revolucionando a pecuária brasileira*. Juiz de Fora: Embrapa Gado de Leite, 2005. p. 191-223.
- Voorrips, R. E. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77-78.
- Wang, Q., Sullivan, R. W., Kight, A., Henry, R. L., Huang, J., Jones, A. M., and Korth, K. L. (2004). Deletion of the chloroplast-localized Thylakoid formation1 gene product in *Arabidopsis* leads to deficient thylakoid formation and variegated leaves. *Plant Physiol.* 136(3), 3594-3604.
- Warmke, H. E. (1954). Apomixis in *Panicum maximum*. *Am. J. Bot.* 41, 5–11.
- Worthington, M., Heffelfinger, C., Bernal, D., Quintero, C., Zapata, Y. P., Perez, J. G., et al. (2016). A parthenogenesis gene candidate and evidence for segmental allopolyploidy in apomictic *Brachiaria decumbens*. *Genetics.* 203(3), 1117-1132.
- Worthington, M., Ebina, M., Yamanaka, N., Heffelfinger, C., Quintero, C., Zapata, Y. P., ... & de la Hoz, J. F. (2019). Translocation of a parthenogenesis gene candidate to an alternate carrier chromosome in apomictic *Brachiaria humidicola*. *BMC genomics*, 20(1), 41.
- Wu, B., Mao, D., Liu, T., Li, Z., & Xing, Y. (2015). Two quantitative trait loci for grain yield and plant height on chromosome 3 are tightly linked in coupling phase in rice. *Mol. Breed.* 35(8), 156.
- Xiao, W., Hu, S., Zhou, X., Yao, R., Luo, J., Yuan, C., et al. (2017). A glucuronokinase gene in *Arabidopsis*, AtGlcAK, is involved in drought tolerance by modulating sugar metabolism. *Plant Mol. Biol. Rep.*, 35(2), 298-311.
- Xu, F., Lyu, Y., Tong, C., Wu, W., Zhu, X., Yin, D., et al. (2013). A statistical model for QTL mapping in polysomic autotetraploids underlying double reduction. *Brief. Bioinform.* 15, 1044-1056.
- Yadav, C. B., Kumar, S., Gupta, M. G., & Bhat, V. (2012). Genetic linkage maps of the chromosomal regions associated with apomictic and sexual modes of reproduction in *Cenchrus ciliaris*. *Mol. Breed.* 30(1), 239-250.
- Yan, H., Kikuchi, S., Neumann, P., Zhang, W., Wu, Y., Chen, F., & Jiang, J. (2010). Genome-wide mapping of cytosine methylation revealed dynamic DNA methylation patterns associated with genes and centromeres in rice. *Plant J.* 63(3), 353-365.
- Yang, Y., Xu, R., Ma, C. J., Vlot, A. C., Klessig, D. F., and Pichersky, E. (2008). Inactive methyl indole-3-acetic acid ester can be hydrolyzed and activated by several esterases belonging to the AtMES esterase family of *Arabidopsis*. *Plant Physiol.* 147(3), 1034-1045.
- Yokoyama, R., and Nishitani, K. (2001). A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of *Arabidopsis*. *Plant Cell Physiol.* 42(10), 1025-1033.
- Young, B. A., Sherwood, R. T., and Bashaw, E. C. (1979). Cleared-pistil and thick-sectioning techniques for detecting aposporous apomixis in grasses. *Can. J. Bot.* 57, 1668-1672.

- Zargar, S. M., Raatz, B., Sonah, H., Bhat, J. A., Dar, Z. A., Agrawal, G. K., & Rakwal, R. (2015). Recent advances in molecular marker techniques: insight into QTL mapping, GWAS and genomic selection in plants. *J. Crop Sci. Biotechnol.* 18(5), 293-308.
- Zhang, X., Lu, S., Jiang, C., Wang, Y., Lv, B., Shen, J., and Ming, F. (2014). RcLEA, a late embryogenesis abundant protein gene isolated from *Rosa chinensis*, confers tolerance to *Escherichia coli* and *Arabidopsis thaliana* and stabilizes enzyme activity under diverse stresses. *Plant Mol. Biol.* 85(4-5), 333-347.
- Zhao, Q. (2016). Lignification: flexibility, biosynthesis and regulation. *Trends Plant Sci.* 21(8), 713-721.
- Zhu, C., Wang, C., and Zhang, Y. M. (2006). Modeling segregation distortion for viability selection I. Reconstruction of linkage maps with distorted markers. *Theor. Appl. Genet.* 114, 295-305.
- Zych, K., Gort, G., Maliepaard, C. A., Jansen, R. C., & Voorrips, R. E. (2019). FitTetra 2.0 – improved genotype calling for tetraploids with multiple population and parental data support. *BMC bioinformatics*, 20(1), 148.

10 APÊNDICES**Informações suplementares do Capítulo I****Title**

High-resolution Linkage Map with Allele Dosage Allows the Identification of an Apomixis Region and Complex Traits in Guinea Grass (*Megathyrsus maximus*)

Authors

Thamiris G. Deo, Rebecca C. U. Ferreira, Letícia A. C. Lara, Aline C. L. Moraes, Alessandro Alves-Pereira, Fernanda A. de Oliveira, Antonio A. F. Garcia, Mateus F. Santos, Liana Jank, Anete P. Souza

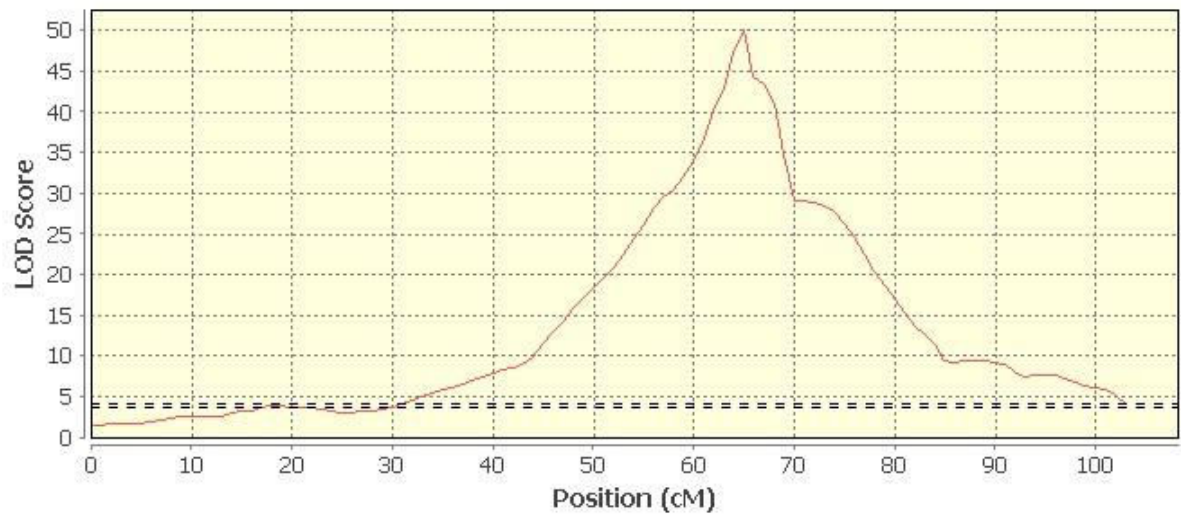
Supplementary material

Figure S1. Identification of the apomixis region located in HG II from guinea grass mapping population. Dotted line indicate the LOD thresholds of 95% obtained after the permutation tests.

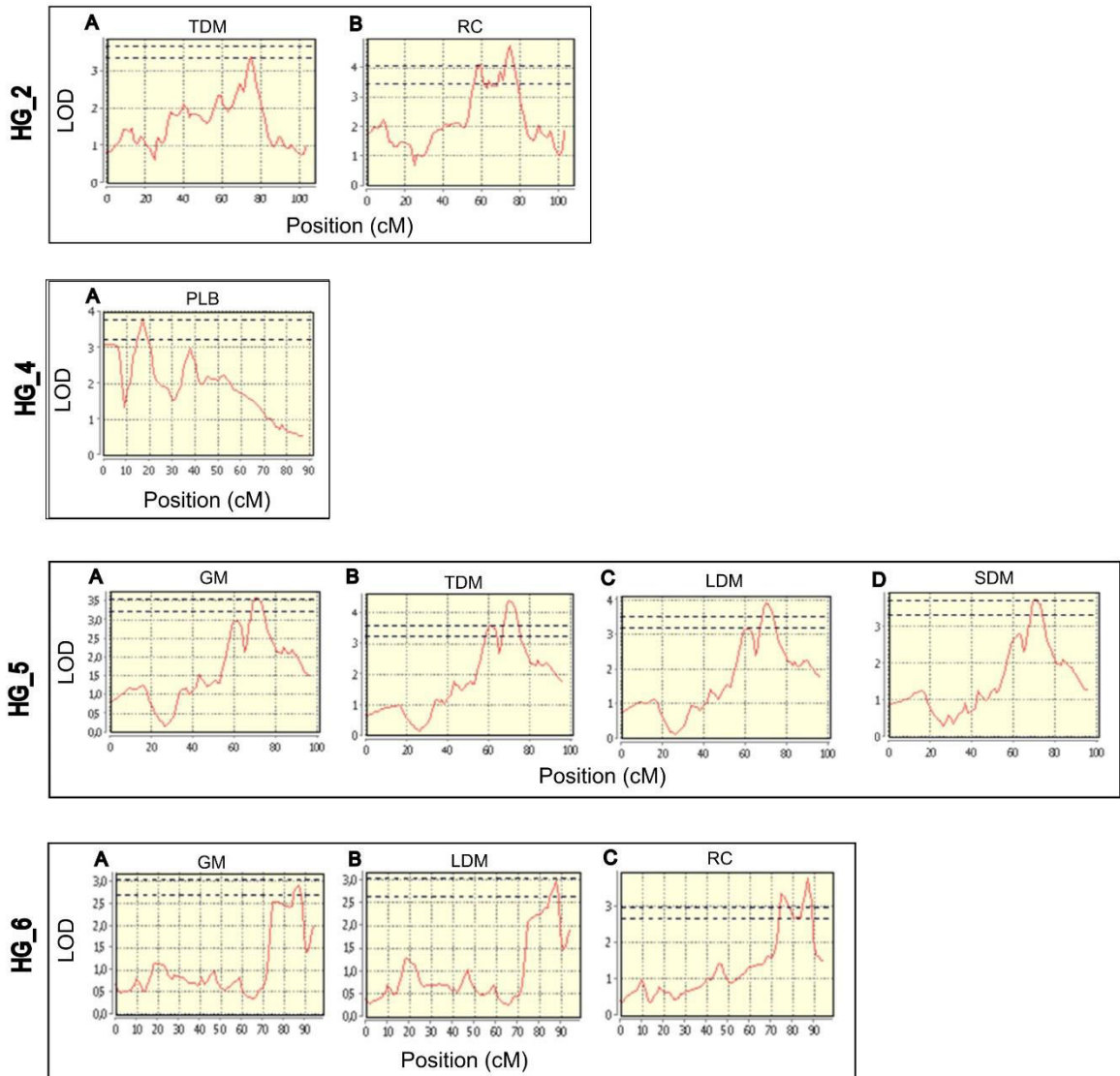


Figure S2. Interval mapping for agronomic traits from the guinea grass (*Megathyrsus maximus*) mapping population in HGs II, IV, V and VI

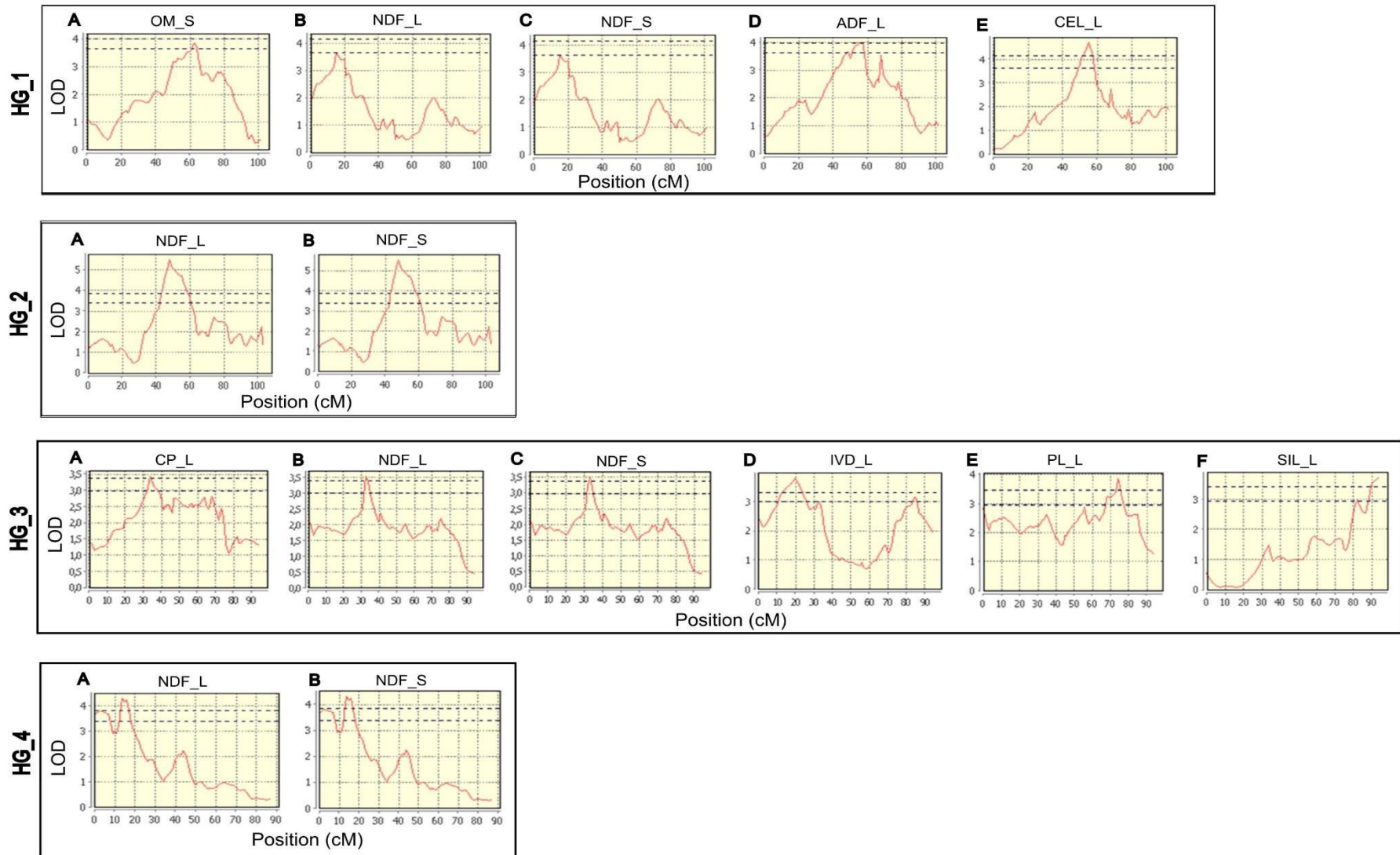


Figure S3. Interval mapping for forage quality from the guinea grass (*Megathyrus maximus*) mapping population in HGs I to IV.

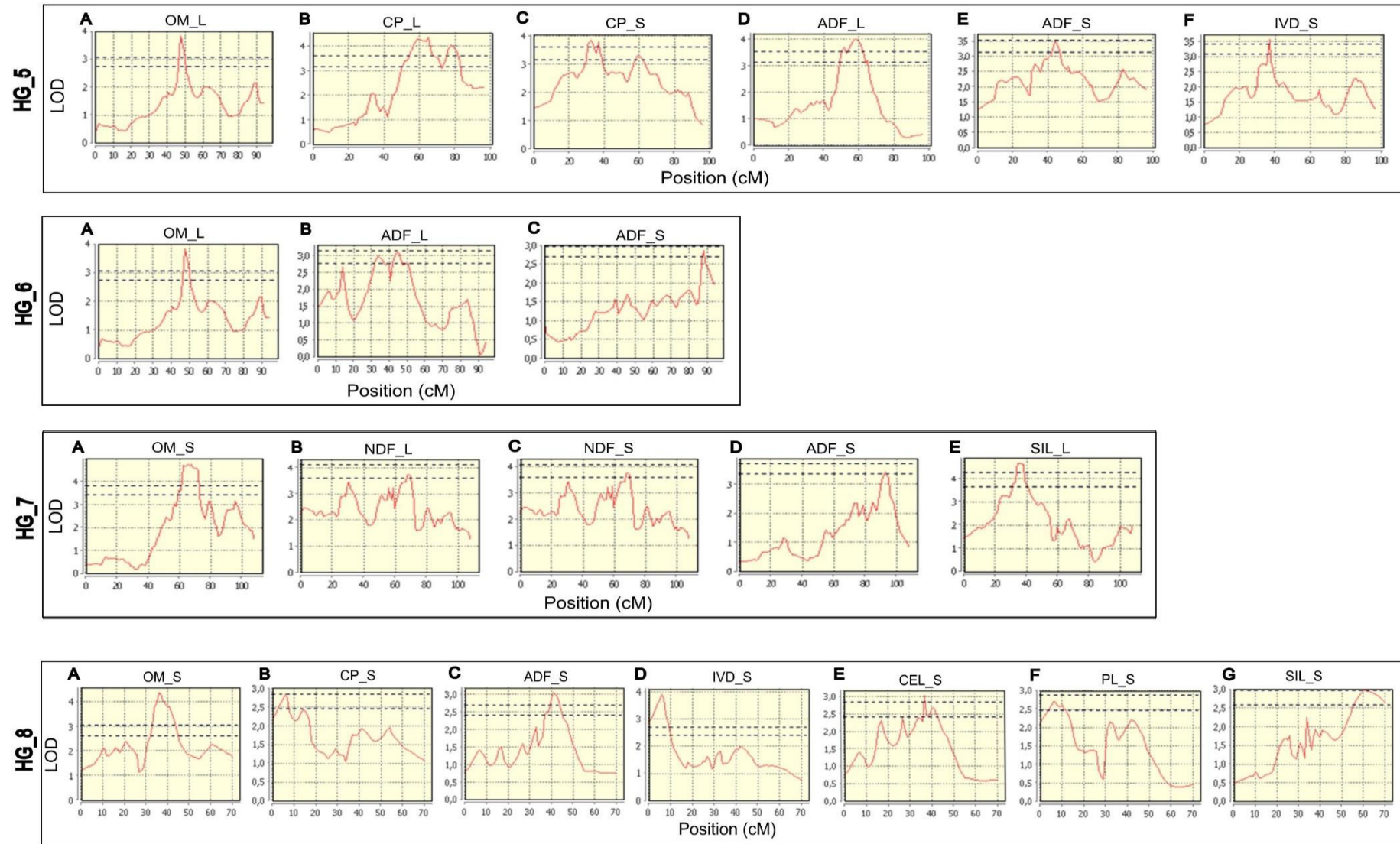


Figure S4. Interval mapping for forage quality from the guinea grass (*Megathyrsus maximus*) mapping population in HGs V to VIII.

Table S1. Reproductive mode of the hybrids from mapping population of guinea grass (*Megathyrsus maximus*).

Genotype	RM^a	Genotype	RM	Genotype	RM	Genotype	RM
B01	Apo ^b	B46	Apo	B102	Sex	C24	Apo
B02	Sex ^c	B47	Apo	B103	Sex	C25	Apo
B03	ND ^d	B48	Apo	B105	ND	C26	Apo
B04	ND	B49	Sex	B108	Apo	C27	Sex
B06	Apo	B50	Sex	B109	Apo	C28	Sex
B07	ND	B51	Apo	B110	Sex	C30	Sex
B10	ND	B66	Apo	B112	ND	C31	ND
B11	Apo	B67	Sex	B113	Sex	C33	Apo
B12	Sex	B68	Apo	B114	Apo	C34	Apo
B13	Sex	B70	Apo	B115	Apo	C36	Apo
B15	ND	B71	ND	B117	Sex	C37	ND
B16	ND	B72	Sex	B118	Sex	C40	Apo
B18	Apo	B73	Apo	B119	ND	C41	Apo
B20	Sex	B74	Sex	B120	Apo	C43	ND
B21	Sex	B75	Sex	B121	ND	C44	Apo
B22	Sex	B76	Apo	B124	Apo	C45	Sex
B23	Sex	B77	ND	B126	Apo	C47	Sex
B24	Apo	B80	ND	B128	Apo	C48	Sex
B25	Sex	B81	Sex	C02	Apo	C50	ND
B27	ND	B83	ND	C03	Apo	C52	Apo
B28	Apo	B84	Sex	C04	Sex	C53	Apo
B31	ND	B85	Sex	C07	Apo	C54	Sex
B33	Sex	B91	Apo	C09	Apo	C55	Apo
B34	Apo	B92	Apo	C10	Apo	C57	Apo
B35	Sex	B93	Apo	C12	Apo	C58	Apo
B37	Apo	B94	Apo	C13	Sex	C59	ND
B38	Apo	B95	ND	C14	Apo	C60	Apo
B39	ND	B96	Sex	C15	Apo	C62	Sex
B40	Apo	B97	Apo	C16	Sex	C63	Apo
B41	Apo	B98	ND	C17	Sex	C64	Apo
B42	ND	B99	Apo	C21	Apo	C65	Sex
B44	Apo	B100	Apo	C22	Sex	C66	Sex
B45	Apo	B101	Apo	C23	Apo	C90	ND

^aReproductive mode; ^bApomictic; ^cSexual; ^dNot determined.

Table S2. AIC and SIC values for G_L and R_L matrices for agronomic traits*. The lowest values and selected variance-covariance (VCOV) structures are indicated in bold. NC means not converged.

GM						TDM						LDM					
G_L matrix			R_L matrix			G_L matrix			R_L matrix			G_L matrix			R_L matrix		
VCOV	AIC	SIC	VCOV	AIC	SIC	VCOV	AIC	SIC	VCOV	AIC	SIC	VCOV	AIC	SIC	VCOV	AIC	SIC
ID	1610.4	1632.8	ID	581.7	598.5	ID	1730.0	1752.5	ID	864.3	886.7	ID	1738.9	1761.3	ID	823.5	840.3
DIAG	1619.6	1670.2	DIAG	569.4	614.4	DIAG	1738.2	1788.7	DIAG	836.7	887.2	DIAG	1747.5	1798.0	DIAG	798.6	843.5
CS	581.7	598.5	CS	-272.8	-239.1	CS	871.8	888.7	CS	216.8	250.5	CS	823.5	840.3	CS	87.4	115.5
CSHet	587.4	632.4	CSHet	-295.7	-233.9	CSHet	871.0	915.8	CSHet	170.4	232.1	CSHet	828.6	879.1	CSHet	46.4	108.1
AR1	NC	NC	AR1	-546.0	-512.3	AR1	864.3	886.7	AR1	241.7	275.3	AR1	818.8	846.9	AR1	NC	NC
AR1Het	NC	NC	AR1Het	-568.0	-506.2	AR1Het	NC	NC	AR1Het	168.9	230.6	AR1Het	NC	NC	AR1Het	-92.1	-30.4
FA1	585.3	635.8	FA1	NC	NC	FA1	870.0	926.1	FA1	NC	NC	FA1	NC	NC	FA1	NC	NC
US	567.3	702.2	US	-713.8	-573.3	US	873.9	1003.0	US	-9.5	130.8	US	NC	NC	US	NC	NC
SDM						PLB						RC					
G_L matrix			R_L matrix			G_L matrix			R_L matrix			G_L matrix					
VCOV	AIC	SIC	VCOV	AIC	SIC		AIC	SIC		AIC	SIC		AIC	SIC			
ID	965.4	980.1	ID	835.8	850.5	ID	1954.4	1971.3	ID	1831.0	1861.4	ID	377.2	421.6			
DIAG	966.4	990.9	DIAG	824.8	854.3	DIAG	1950.1	1995.0	DIAG	1821.8	1872.3	DIAG	378.1	427.5			
CS	835.8	850.5	CS	806.6	831.2	CS	1848.6	1871.0	CS	1839.5	1867.6	CS	343.0	387.3			
CSHet	834.3	858.8	CSHet	798.0	837.4	CSHet	1836.1	1886.6	CSHet	1822.4	1878.5	CSHet	345.7	400.0			
AR1	836.0	850.8	AR1	826.0	845.7	AR1	1839.0	1861.4	AR1	1840.8	1868.8	AR1	344.9	394.2			
AR1Het	834.4	859.0	AR1Het	815.4	849.8	AR1Het	1827.5	1878.0	AR1Het	1823.7	1879.9	AR1Het	345.7	400.0			
FA1	836.1	865.7	FA1	NC	NC	FA1	1838.1	1911.1	FA1	NC	NC	FA1	346.8	406.0			
US	840.3	879.6	US	793.3	837.6	US	NC	NC	US	1833.3	1968.0	US	352.1	421.1			

*Agronomic traits: green matter (GM), total dry matter (TDM), leaf dry matter (LDM), stem dry matter (SDM), percentage of leaf blade (PLB) and regrowth capacity (RC).

Table S3. AIC and SIC values for G_L and R_L matrices for nutritional traits*.

Leaf								
	OM_L	CP_L	NDF_L	ADF_L	IVD_L	CEL_L	PL_L	SIL_L
AIC	342.34	500.52	354.96	345.76	329.68	528.47	348.05	343.99
SIC	353.84	515.86	362.62	357.26	341.19	536.14	359.55	355.49
Stem								
	OM_S	CP_S	NDF_S	ADF_S	IVD_S	CEL_S	PL_S	SIL_S
AIC	328.04	310.15	1019.01	666.55	332.18	467.56	319.97	337.44
SIC	339.44	321.55	1026.67	677.95	343.58	475.16	331.37	348.83

*Nutritional quality traits for the leaf and stem: organic matter (OM_L and OM_S), crude protein (CP_L and CP_S), neutral detergent fiber (NDF_L and NDF_S), acid detergent fiber (ADF_L and ADF_S), *in vitro* digestibility of organic matter (IVD_L and IVD_S), cellulose (CEL_L and CEL_S - %), permanganate lignin (PL_L and PL_S) and silica (SIL_L and SIL_S).

Table S4. Name, position and dosage type of the SNP markers in the eight HGs from the guinea grass (*Megathyrus maximus*) mapping population.

HG	Name	Position	Dosage type
1	D_1_1699583	0.0	Duplex
1	S_2_3709976	0.0	Simplex
1	SS_1_3590376	2.637	Double-Simplex
1	SD_1_3590319	3.270	Simplex-Duplex
1	S_1228_24403	4.458	Simplex
1	SD_1_34354292	11.694	Simplex-Duplex
1	DD_16_50296225	11.729	Double-Duplex
1	S_1_20942263	14.279	Simplex
1	D_7_33646464	14.727	Duplex
1	S_1_20942225	15.455	Simplex
1	S_1_46594647	19.720	Simplex
1	SD_3_46295633	20.324	Simplex-Duplex
1	S_3_46295663	20.502	Simplex
1	SS_3_46295676	20.533	Double-Simplex
1	SD_1_37202168	21.574	Simplex-Duplex
1	SS_7_33646470	22.881	Double-Simplex
1	D_7_3806733	23.752	Duplex
1	SS_1_44355249	24.281	Double-Simplex
1	SS_14_42178687	24.492	Double-Simplex
1	S_1_44355241	24.763	Simplex
1	S_3_46295651	24.964	Simplex
1	S_1_34368018	25.056	Simplex
1	S_7_8484743	25.345	Simplex
1	S_14_42178733	26.828	Simplex
1	S_25255_590	26.920	Simplex
1	S_25255_596	26.937	Simplex
1	SD_1_44355270	27.000	Simplex-Duplex
1	SD_7_7288432	27.791	Simplex-Duplex
1	XSS_7_7112702	28.367	X-double-simplex
1	D_7_8484782	30.588	Duplex
1	S_7_7288434	33.534	Simplex
1	S_7_13783366	34.537	Simplex
1	S_7_17402908	35.649	Simplex
1	D_2_30230956	36.712	Duplex
1	D_16_6264738	38.621	Duplex
1	S_7_13569595	39.304	Simplex
1	D_7_13569568	41.276	Duplex
1	SD_7_13783398	42.330	Simplex-Duplex
1	S_8_13535572	42.711	Simplex
1	S_8_26205934	42.717	Simplex
1	D_7_21223239	43.752	Duplex
1	S_7_42982235	44.530	Simplex

1	D_18_32894865	44.653	Duplex
1	SS_8_21900921	45.098	Double-Simplex
1	S_8_26350175	45.148	Simplex
1	S_8_21765090	45.415	Simplex
1	S_8_26205926	45.763	Simplex
1	SS_8_13535567	45.869	Double-Simplex
1	S_7_50011977	48.014	Simplex
1	SS_2327_14615	48.165	Double-Simplex
1	S_2684_9874	48.279	Simplex
1	SS_7_44107582	48.833	Double-Simplex
1	SD_7_42982203	48.841	Simplex-Duplex
1	S_8_50636764	48.864	Simplex
1	D_7_43180011	49.151	Duplex
1	SS_2684_9877	49.219	Double-Simplex
1	SD_7_43180023	49.531	Simplex-Duplex
1	S_18_24146718	49.831	Simplex
1	S_7_40289321	50.340	Simplex
1	SS_7_43176646	50.649	Double-Simplex
1	S_7_40289322	50.656	Simplex
1	SD_7_44783718	50.830	Simplex-Duplex
1	S_8_20762396	50.864	Simplex
1	SS_18_24146717	50.997	Double-Simplex
1	XSS_12_43118788	51.091	X-double-simplex
1	SS_12_43118785	51.638	Double-Simplex
1	S_7_49627673	52.016	Simplex
1	S_7_44783739	52.698	Simplex
1	S_7_44783736	52.895	Simplex
1	S_7_49627665	52.920	Simplex
1	SD_8_45824313	53.038	Simplex-Duplex
1	D_17_76964365	53.899	Duplex
1	D_8_50074072	54.338	Duplex
1	SS_7_44783696	55.227	Double-Simplex
1	D_8_47736933	55.257	Duplex
1	S_6_49366931	56.609	Simplex
1	S_8_45824276	57.606	Simplex
1	S_8_50074068	59.168	Simplex
1	S_11_49575150	59.621	Simplex
1	S_1_56966880	60.376	Simplex
1	S_18_32894856	61.452	Simplex
1	S_729_21504	61.656	Simplex
1	S_8_50508635	62.641	Simplex
1	S_2_56901087	63.310	Simplex
1	D_11_49575112	64.192	Duplex
1	S_8_2051486	65.583	Simplex
1	S_7_3423898	65.643	Simplex
1	XSS_8_2051472	66.730	X-double-simplex

1	D_2_70040496	66.858	Duplex
1	S_8_2253997	67.040	Simplex
1	S_17_3201062	67.730	Simplex
1	S_17_3201057	67.915	Simplex
1	S_1_60892985	67.929	Simplex
1	SD_1_50988808	68.334	Simplex-Duplex
1	D_2_39424176	68.649	Duplex
1	S_7_33646453	69.563	Simplex
1	S_2_39424144	72.024	Simplex
1	D_2_85236088	73.445	Duplex
1	SD_1_64485840	75.280	Simplex-Duplex
1	S_2_3156786	75.550	Simplex
1	S_1_62300248	75.597	Simplex
1	S_2_68490644	76.131	Simplex
1	S_1_60071226	76.237	Simplex
1	S_2_77029350	76.395	Simplex
1	SS_1_64558324	76.717	Double-Simplex
1	S_1_60071184	77.271	Simplex
1	S_1_62143423	78.047	Simplex
1	SD_1_65328685	78.721	Simplex-Duplex
1	D_1_64558302	79.418	Duplex
1	S_1_62300212	79.434	Simplex
1	S_1_64485793	79.954	Simplex
1	S_4_77344309	81.405	Simplex
1	S_1_65495271	81.926	Simplex
1	S_1_65095317	82.669	Simplex
1	S_2_547804	82.930	Simplex
1	S_2_77029356	83.764	Simplex
1	S_1_79601248	83.993	Simplex
1	S_1_78235672	84.451	Simplex
1	S_2_95905275	84.572	Simplex
1	D_1_77920034	84.572	Duplex
1	DD_1_77920004	85.664	Double-Duplex
1	S_2_80368458	87.154	Simplex
1	S_1_77919996	87.248	Simplex
1	S_1_75031903	88.426	Simplex
1	DD_11_20928668	89.473	Double-Duplex
1	S_9_1930273	90.119	Simplex
1	D_2_93749617	90.901	Duplex
1	SD_1_75031893	90.931	Simplex-Duplex
1	SD_1_77919999	91.074	Simplex-Duplex
1	S_2_83608442	92.791	Simplex
1	S_2_85513969	92.849	Simplex
1	S_1_71892997	93.361	Simplex
1	D_1_75031912	93.852	Duplex
1	S_1_73439738	96.183	Simplex

1	S_1_70306265	96.502	Simplex
1	SS_2_83608460	96.507	Double-Simplex
1	D_2_81571646	97.792	Duplex
1	SS_2_82390933	98.027	Double-Simplex
1	S_18_34565671	98.269	Simplex
1	S_1_70523932	98.275	Simplex
1	S_18_34565672	99.464	Simplex
1	D_2_82390912	99.848	Duplex
1	D_3_467216	100.170	Duplex
1	SD_1_73439779	101.743	Simplex-Duplex
2	S_25255_594	0.000	Simplex
2	S_6_792739	0.982	Simplex
2	S_11_56205134	1.674	Simplex
2	S_1_12465709	3.774	Simplex
2	S_6_3949967	4.959	Simplex
2	S_5_1904772	5.281	Simplex
2	SS_16_4829835	8.710	Double-Simplex
2	S_11_56205163	11.924	Simplex
2	SD_16_4829809	12.389	Simplex-Duplex
2	S_14_57446669	13.043	Simplex
2	D_6_5985167	13.669	Duplex
2	S_15_3102249	14.064	Simplex
2	S_4_80448279	15.374	Simplex
2	D_11259_1360	16.499	Duplex
2	S_7_63937	17.521	Simplex
2	D_11259_1400	18.807	Duplex
2	S_5_4702409	21.434	Simplex
2	S_5339_10060	22.275	Simplex
2	SS_16_3751233	23.525	Double-Simplex
2	S_6_1289718	24.452	Simplex
2	S_14_57446676	24.526	Simplex
2	S_7_63926	24.665	Simplex
2	S_5_8789063	25.153	Simplex
2	S_15_3102241	25.272	Simplex
2	SS_15_3102236	25.587	Double-Simplex
2	S_11477_2711	25.627	Simplex
2	S_5_5371347	26.594	Simplex
2	S_5_13410437	26.844	Simplex
2	D_5_13410454	28.985	Duplex
2	S_10_73051980	30.041	Simplex
2	S_5_13410451	30.385	Simplex
2	S_5_13710428	30.732	Simplex
2	SS_5_17512332	33.000	Double-Simplex
2	S_1_24472654	33.087	Simplex
2	S_5_16799423	33.829	Simplex
2	S_5_16564616	34.756	Simplex

2	S_5_16799422	34.903	Simplex
2	S_1_29018318	37.363	Simplex
2	SS_5_18860119	39.094	Double-Simplex
2	S_3_26025802	39.733	Simplex
2	S_1_29018279	42.054	Simplex
2	DD_5_19426677	42.329	Double-Duplex
2	D_5_24044946	42.497	Duplex
2	S_4_25415348	43.529	Simplex
2	D_15_59445300	43.681	Duplex
2	S_5_24044921	45.610	Simplex
2	SD_17_1782342	48.535	Simplex-Duplex
2	S_3_26025770	49.111	Simplex
2	S_11_64437852	51.094	Simplex
2	S_17_1782327	51.794	Simplex
2	S_17_1782294	52.648	Simplex
2	S_1_38339276	53.824	Simplex
2	D_1_37607072	54.788	Duplex
2	S_16_47309024	54.978	Simplex
2	S_1_37607071	56.189	Simplex
2	SS_11_41266975	56.283	Double-Simplex
2	S_11_41266993	58.176	Simplex
2	SS_64_29736	58.549	Double-Simplex
2	S_10_28552900	60.699	Simplex
2	S_5_52166873	60.955	Simplex
2	S_11_41266996	61.142	Simplex
2	D_1_37607021	61.206	Duplex
2	S_6_5099445	61.861	Simplex
2	S_6_22187422	61.929	Simplex
2	SS_5_16668963	62.365	Double-Simplex
2	D_9_23315063	62.565	Duplex
2	D_5_16668949	63.076	Duplex
2	SS_16_46921584	63.230	Double-Simplex
2	S_6_21186594	63.230	Simplex
2	DD_10_66985802	63.264	Double-Duplex
2	S_6_22187426	63.724	Simplex
2	S_14_29023868	64.157	Simplex
2	S_10_48091934	64.188	Simplex
2	S_5_55915746	64.341	Simplex
2	XSS_6_21186575	64.611	X-double-simplex
2	SS_10_66985814	66.362	Double-Simplex
2	S_6_22187398	67.677	Simplex
2	DD_5_51532611	68.183	Double-Duplex
2	S_3_58324136	69.638	Simplex
2	SS_5_63144891	69.865	Double-Simplex
2	S_10_66985808	70.172	Simplex
2	S_10_66985781	70.257	Simplex

2	S_5_63144924	70.288	Simplex
2	S_5_63144872	70.337	Simplex
2	D_5_46088876	70.478	Duplex
2	S_5_36319246	70.706	Simplex
2	D_17_47688006	71.135	Duplex
2	D_5_52166861	71.135	Duplex
2	S_6_5083067	72.775	Simplex
2	S_9_93635876	72.910	Simplex
2	SD_5_63144886	74.021	Simplex-Duplex
2	SS_6_5099415	74.758	Double-Simplex
2	DD_5_63144870	75.446	Double-Duplex
2	S_5_63351703	75.700	Simplex
2	D_5_63144879	75.952	Duplex
2	S_6_42493237	77.897	Simplex
2	S_9_14916835	78.237	Simplex
2	S_13_52657368	81.371	Simplex
2	SS_5_31869446	81.941	Double-Simplex
2	S_6_42493267	82.113	Simplex
2	SS_5_31869450	83.952	Double-Simplex
2	SS_5_64169045	84.488	Double-Simplex
2	SS_5_68376990	85.146	Double-Simplex
2	S_6_52964607	85.692	Simplex
2	DD_5_9342905	86.050	Double-Duplex
2	DD_5_66642334	88.472	Double-Duplex
2	S_6_44588399	88.642	Simplex
2	S_5_64169031	88.700	Simplex
2	D_5_66642368	89.312	Duplex
2	D_5_66642331	90.319	Duplex
2	D_5_66642371	90.448	Duplex
2	S_6_50803323	90.658	Simplex
2	S_5_66642307	91.188	Simplex
2	S_5_66642299	91.299	Simplex
2	S_5_68198441	92.219	Simplex
2	D_6_50935272	92.309	Duplex
2	S_5_9342886	93.010	Simplex
2	S_5_68377002	93.741	Simplex
2	SD_5_9342629	95.709	Simplex-Duplex
2	SS_6_50935276	95.924	Double-Simplex
2	S_5_9342932	97.182	Simplex
2	S_5_66644417	99.366	Simplex
2	S_18_106668337	99.446	Simplex
2	S_14_54212097	101.225	Simplex
2	S_5_63144876	102.236	Simplex
2	D_3_57916915	103.531	Duplex
2	D_3_57916903	103.531	Duplex
2	SD_11_72796448	103.531	Simplex-Duplex

3	SS_17_32339857	0.000	Double-Simplex
3	S_17_32339864	0.964	Simplex
3	S_3_898136	1.755	Simplex
3	SS_14_65502910	1.924	Double-Simplex
3	D_8_3841166	2.938	Duplex
3	SS_3_898102	2.965	Double-Simplex
3	S_103_32460	3.808	Simplex
3	S_103_32476	6.186	Simplex
3	S_562_19851	6.236	Simplex
3	S_12_17473598	9.394	Simplex
3	S_12_17473571	9.864	Simplex
3	S_13_38639521	13.410	Simplex
3	S_5_9243093	13.571	Simplex
3	D_3_7946740	14.796	Duplex
3	SS_2_3629174	18.977	Double-Simplex
3	SD_729_21503	20.019	Simplex-Duplex
3	S_2_3629175	20.244	Simplex
3	S_2_3629181	25.479	Simplex
3	S_3_45154783	27.693	Simplex
3	S_3_78499323	31.194	Simplex
3	S_3_7073447	32.446	Simplex
3	D_2_3156828	33.460	Duplex
3	SD_3_15814711	34.633	Simplex-Duplex
3	SD_4_42742062	35.029	Simplex-Duplex
3	S_5_10896697	35.540	Simplex
3	S_3_15814701	36.306	Simplex
3	S_12_63925623	38.390	Simplex
3	S_5_10896682	39.249	Simplex
3	S_2_74300311	39.389	Simplex
3	S_3_57916904	39.994	Simplex
3	D_5_10896679	40.536	Duplex
3	S_2_80824255	41.007	Simplex
3	S_4_42742065	41.084	Simplex
3	S_2_80824258	43.513	Simplex
3	S_3_50286568	44.043	Simplex
3	S_5_10896688	44.155	Simplex
3	S_3_59581065	44.291	Simplex
3	SS_3_59807865	45.226	Double-Simplex
3	D_3_57232456	46.376	Duplex
3	S_2_80824304	47.079	Simplex
3	SD_3_64308814	51.855	Simplex-Duplex
3	S_4_69971270	54.796	Simplex
3	S_1901_2105	55.124	Simplex
3	SD_3_64490202	55.635	Simplex-Duplex
3	S_18_26106378	56.050	Simplex
3	S_18_26106423	56.155	Simplex

3	SS_3_64490218	56.376	Double-Simplex
3	S_3_57232439	57.239	Simplex
3	SD_3_68084635	57.567	Simplex-Duplex
3	D_4_75935654	57.621	Duplex
3	D_1_24971519	58.924	Duplex
3	D_3_68026563	58.963	Duplex
3	S_2_86215170	59.231	Simplex
3	S_3_59807868	59.302	Simplex
3	S_3_72126203	61.473	Simplex
3	S_4_77078208	62.463	Simplex
3	S_11_65382979	63.929	Simplex
3	S_18348_1172	65.072	Simplex
3	S_4_81497247	65.445	Simplex
3	S_2_11403862	66.275	Simplex
3	D_7_55760558	66.936	Duplex
3	S_16_8476951	67.127	Simplex
3	SS_13215_770	67.458	Double-Simplex
3	S_7_55760567	67.722	Simplex
3	SD_3_68913834	68.065	Simplex-Duplex
3	D_15_7485846	69.366	Duplex
3	D_13215_742	69.870	Duplex
3	S_14_35885386	70.016	Simplex
3	S_4_88316891	70.997	Simplex
3	S_3_81117146	72.002	Simplex
3	S_3_87883967	73.212	Simplex
3	D_13215_743	73.521	Duplex
3	SS_3_81103476	73.699	Double-Simplex
3	S_3_82892072	74.335	Simplex
3	DD_4_94916073	79.217	Double-Duplex
3	D_3_84786589	81.616	Duplex
3	D_4_93287306	81.766	Duplex
3	S_3_81102375	82.667	Simplex
3	SD_3_87784991	82.761	Simplex-Duplex
3	D_4_101910741	83.655	Duplex
3	S_2716_12810	85.249	Simplex
3	S_4_93149506	86.920	Simplex
3	S_3_89096243	89.633	Simplex
3	D_4_101472580	89.794	Duplex
3	S_3_91817947	94.058	Simplex
4	S_14_20389967	0.000	Simplex
4	S_14_20390652	0.000	Simplex
4	D_15_26475097	3.147	Duplex
4	D_14_20390654	6.759	Duplex
4	S_14_20389944	7.549	Simplex
4	S_14_19604308	8.015	Simplex
4	SD_18_53338540	8.415	Simplex-Duplex

4	T_14_5066444	8.941	Triplex
4	T_13_8006566	9.002	Triplex
4	SD_14_46855625	11.226	Simplex-Duplex
4	T_14_5066402	11.845	Triplex
4	D_14_20390818	12.591	Duplex
4	S_13_3521904	13.552	Simplex
4	SD_2355_2824	13.707	Simplex-Duplex
4	S_13_38639517	13.846	Simplex
4	S_11_3868770	14.478	Simplex
4	SS_15_3109466	14.556	Double-Simplex
4	SS_15_3109465	14.604	Double-Simplex
4	D_15_3109484	14.798	Duplex
4	T_13_8006568	14.866	Triplex
4	S_13_33760544	15.474	Simplex
4	DD_15_3109474	15.547	Double-Duplex
4	D_13_3267077	15.676	Duplex
4	S_13_25407230	16.145	Simplex
4	D_14_20389959	16.703	Duplex
4	S_14_20389857	17.751	Simplex
4	T_13_8006575	18.258	Triplex
4	S_6_52294050	18.692	Simplex
4	S_9_27474255	18.819	Simplex
4	D_13_38639518	20.089	Duplex
4	T_15_3109450	20.394	Triplex
4	S_13_38639493	21.201	Simplex
4	SS_13_8006527	21.507	Double-Simplex
4	S_8_28393021	22.783	Simplex
4	S_13_8006523	26.180	Simplex
4	S_280_36722	26.629	Simplex
4	SD_14_45090610	28.109	Simplex-Duplex
4	S_13_50908458	28.663	Simplex
4	S_13_46816949	28.773	Simplex
4	S_16_21220638	28.864	Simplex
4	S_14_45090563	29.077	Simplex
4	S_16_21220648	29.595	Simplex
4	S_13_48127882	29.652	Simplex
4	S_16_21220660	29.652	Simplex
4	S_16938_1641	30.120	Simplex
4	S_16_21220654	30.123	Simplex
4	S_14_48623677	31.251	Simplex
4	S_14_37858277	31.320	Simplex
4	D_13_48127868	32.395	Duplex
4	S_14_53737303	32.730	Simplex
4	S_13_46816952	32.765	Simplex
4	SD_14_45090569	34.041	Simplex-Duplex
4	SS_5_13158078	34.130	Double-Simplex

4	S_13_50908425	35.002	Simplex
4	D_13_47692259	35.096	Duplex
4	S_1_57004838	38.304	Simplex
4	D_1012_16963	38.462	Duplex
4	D_13_48127837	38.462	Duplex
4	D_14_44253015	38.828	Duplex
4	S_13_48127889	39.038	Simplex
4	S_18_6808489	39.443	Simplex
4	SD_14_53569797	40.131	Simplex-Duplex
4	SS_14_40547946	40.618	Double-Simplex
4	S_14_45090642	40.647	Simplex
4	D_13_47711178	41.575	Duplex
4	D_13_38639533	41.670	Duplex
4	S_14_46359041	43.077	Simplex
4	SS_13_48127863	43.611	Double-Simplex
4	S_13_48127858	43.858	Simplex
4	T_14_28099456	45.229	Triplex
4	S_14_48423390	45.264	Simplex
4	S_14_40547948	45.599	Simplex
4	S_14_48623674	46.992	Simplex
4	S_9_33701780	50.054	Simplex
4	S_14_53255647	50.879	Simplex
4	DD_13_51892778	50.986	Double-Duplex
4	S_14_53255591	51.046	Simplex
4	D_14_53255624	53.067	Duplex
4	S_14_55258228	55.619	Simplex
4	S_3327_13962	56.318	Simplex
4	D_3_13212960	56.856	Duplex
4	S_181_17574	58.292	Simplex
4	S_4_6784752	59.169	Simplex
4	D_14_55258272	64.752	Duplex
4	S_13_58203708	68.269	Simplex
4	S_13_59819187	69.401	Simplex
4	SS_28113_260	70.141	Double-Simplex
4	DD_13_66115038	71.401	Double-Duplex
4	SD_14_65489226	72.206	Simplex-Duplex
4	S_14_60572186	72.279	Simplex
4	S_1_74435942	72.369	Simplex
4	S_13_12223623	72.534	Simplex
4	S_14_65489200	73.151	Simplex
4	DD_14_65489204	73.269	Double-Duplex
4	S_1_74435927	73.281	Simplex
4	S_13_70647344	73.422	Simplex
4	S_181_17599	73.567	Simplex
4	S_14_65962750	74.475	Simplex
4	SS_13_60436791	74.935	Double-Simplex

4	S_13_69001336	74.959	Simplex
4	S_181_17577	75.270	Simplex
4	SS_14_63490016	76.263	Double-Simplex
4	D_9_24886482	76.709	Duplex
4	SD_14_65211200	76.728	Simplex-Duplex
4	S_14_65296892	77.318	Simplex
4	D_14_63162192	77.397	Duplex
4	SS_14_65962751	77.467	Double-Simplex
4	S_11_4190684	77.630	Simplex
4	S_14_11396267	77.740	Simplex
4	S_14_65997409	78.049	Simplex
4	D_14_65296854	78.175	Duplex
4	S_14_65296893	78.646	Simplex
4	SD_14_65594982	79.618	Simplex-Duplex
4	D_4_54130123	85.971	Duplex
4	D_4_54130131	87.658	Duplex
5	S_10_4350590	0.000	Simplex
5	S_10_4669435	1.197	Simplex
5	D_10_11085637	9.843	Duplex
5	D_10_11085663	11.016	Duplex
5	S_10_11085611	12.212	Simplex
5	SS_10_20303178	16.737	Double-Simplex
5	S_10_19857331	18.117	Simplex
5	D_5_2484801	20.762	Duplex
5	S_2_92807999	22.773	Simplex
5	SD_2364_6949	23.962	Simplex-Duplex
5	S_2364_6943	24.675	Simplex
5	S_10_19857281	24.874	Simplex
5	S_5_70521598	25.182	Simplex
5	S_9_23228629	26.280	Simplex
5	T_16_8476957	26.963	Triplex
5	S_9_25341246	27.160	Simplex
5	XSS_16_8476923	27.627	X-double-simplex
5	T_9_67715959	29.198	Triplex
5	S_10_26787329	29.334	Simplex
5	SS_18_112080423	29.407	Double-Simplex
5	D_10_20303168	29.417	Duplex
5	S_10_21980151	30.058	Simplex
5	S_9_28761393	30.178	Simplex
5	S_16_8476927	30.329	Simplex
5	D_4_51232833	31.119	Duplex
5	S_9_25341227	33.011	Simplex
5	S_10_26787325	33.457	Simplex
5	S_9_73798057	33.583	Simplex
5	SS_14_37064124	33.659	Double-Simplex
5	S_1_78192598	35.051	Simplex

5	T_16_8476921	35.148	Triplex
5	D_10_32705488	35.390	Duplex
5	S_4_51232779	35.907	Simplex
5	SD_10_32705444	36.486	Simplex-Duplex
5	DD_10_34384853	36.704	Double-Duplex
5	SS_16_24183522	36.712	Double-Simplex
5	D_10_54203087	36.928	Duplex
5	SS_10_35012573	37.022	Double-Simplex
5	D_9_74576648	37.320	Duplex
5	S_9_67461096	37.434	Simplex
5	DD_6559_6725	37.726	Double-Duplex
5	SD_16_24183528	38.152	Simplex-Duplex
5	S_11_4190708	38.905	Simplex
5	D_10_65093670	39.562	Duplex
5	S_15_3109475	39.636	Simplex
5	S_10_58309651	39.706	Simplex
5	S_9_78242747	41.785	Simplex
5	DD_9_15185238	42.239	Double-Duplex
5	SS_9_75026837	42.314	Double-Simplex
5	S_13_74046758	43.024	Simplex
5	S_9_72144897	43.356	Simplex
5	SD_13_74046809	43.954	Simplex-Duplex
5	SD_9_78242732	44.587	Simplex-Duplex
5	D_6_26559105	46.717	Duplex
5	S_6_26559092	48.145	Simplex
5	D_10_71964458	48.619	Duplex
5	S_5_7408718	50.609	Simplex
5	S_9_81392552	50.895	Simplex
5	DD_18_5394882	52.368	Double-Duplex
5	SS_9_81576794	52.638	Double-Simplex
5	D_9_81576812	52.905	Duplex
5	S_10_76866399	52.917	Simplex
5	S_9_81392558	52.965	Simplex
5	D_10_76866474	53.538	Duplex
5	D_14_54212099	59.142	Duplex
5	S_10_84979582	62.589	Simplex
5	D_4_15884221	64.078	Duplex
5	S_10_85946619	64.318	Simplex
5	S_10_84979593	64.502	Simplex
5	S_9_76198147	64.942	Simplex
5	S_10_89868575	65.119	Simplex
5	S_9_91721410	65.431	Simplex
5	D_10_89868586	66.216	Duplex
5	S_9_91721406	66.686	Simplex
5	S_4_15884218	66.981	Simplex
5	S_4_101334904	66.995	Simplex

5	S_3782_11675	67.567	Simplex
5	DD_10_89968763	67.592	Double-Duplex
5	S_10_89968752	68.835	Simplex
5	SD_5_29538219	69.100	Simplex-Duplex
5	SD_4_46385718	69.141	Simplex-Duplex
5	SS_10_85946581	70.607	Double-Simplex
5	S_11_63464540	71.561	Simplex
5	D_10_93089848	72.795	Duplex
5	S_9_95401063	72.908	Simplex
5	S_10_93341467	75.175	Simplex
5	D_9_109089413	76.541	Duplex
5	S_3782_11677	77.708	Simplex
5	T_2_30788698	81.205	Triplex
5	S_10_20303154	81.507	Simplex
5	S_9_105743969	82.099	Simplex
5	D_10_97483989	82.227	Duplex
5	S_9_109089432	83.741	Simplex
5	S_9_111746576	83.991	Simplex
5	S_15_3102242	85.565	Simplex
5	S_9_114815055	87.807	Simplex
5	D_9_109089438	87.962	Duplex
5	S_10_98908365	88.757	Simplex
5	S_10_99501367	88.875	Simplex
5	D_10_99501373	88.889	Duplex
5	SS_10_99501380	89.417	Double-Simplex
5	S_9_115230680	89.484	Simplex
5	D_9_111746581	93.151	Duplex
5	S_9_115230695	96.078	Simplex
6	SS_15_23320222	0.000	Double-Simplex
6	S_15_26475103	0.000	Simplex
6	SD_15_42148824	1.907	Simplex-Duplex
6	S_11259_1370	6.656	Simplex
6	S_5_2340112	7.664	Simplex
6	S_5_19426662	8.692	Simplex
6	D_2497_17706	9.746	Duplex
6	S_5_2340208	10.284	Simplex
6	S_6_41795655	10.525	Simplex
6	S_16_4829806	10.643	Simplex
6	S_16_2741333	12.238	Simplex
6	SD_16_2741372	13.159	Simplex-Duplex
6	D_16_2741351	13.342	Duplex
6	SS_6_41795642	14.017	Double-Simplex
6	SD_2497_17745	14.498	Simplex-Duplex
6	SD_15_1475675	14.666	Simplex-Duplex
6	SS_15_6600399	17.155	Double-Simplex
6	S_6_41795641	17.371	Simplex

6	S_16_4829837	18.385	Simplex
6	SS_5_2340214	20.170	Double-Simplex
6	S_14_4902170	20.447	Simplex
6	SD_2497_17748	23.146	Simplex-Duplex
6	T_1012_16966	27.269	Triplex
6	SD_15_6600353	27.615	Simplex-Duplex
6	S_2497_17739	32.783	Simplex
6	S_9_67574800	34.010	Simplex
6	S_6_792755	38.130	Simplex
6	S_3_45154796	39.341	Simplex
6	S_6_792760	39.985	Simplex
6	S_15_46815780	41.021	Simplex
6	D_3327_13989	41.071	Duplex
6	S_6_41795684	42.198	Simplex
6	SS_10_84824590	46.640	Double-Simplex
6	S_15_26475137	47.305	Simplex
6	D_15_26475143	47.362	Duplex
6	SS_15_26475147	47.362	Double-Simplex
6	S_10_84824585	50.822	Simplex
6	SS_3327_13950	55.367	Double-Simplex
6	T_3327_13990	58.069	Triplex
6	D_16_61789379	59.211	Duplex
6	T_3327_13977	59.211	Triplex
6	S_5_14128123	60.129	Simplex
6	S_3_7073431	61.350	Simplex
6	S_5_62621793	66.190	Simplex
6	S_16_61789378	68.373	Simplex
6	S_3327_13969	70.712	Simplex
6	S_12_57629442	73.169	Simplex
6	T_12_57629411	73.432	Triplex
6	S_16_61789369	74.399	Simplex
6	S_12_57629434	74.553	Simplex
6	SD_12_57629407	80.972	Simplex-Duplex
6	D_4_93162493	81.064	Duplex
6	S_7_46721684	82.843	Simplex
6	D_1012_16969	83.922	Duplex
6	SD_12_57629445	84.780	Simplex-Duplex
6	S_4_93162483	86.210	Simplex
6	SS_16_72289026	86.355	Double-Simplex
6	D_5_62621828	86.491	Duplex
6	SS_16_72289038	87.236	Double-Simplex
6	D_4_93162474	87.262	Duplex
6	D_1012_16924	87.989	Duplex
6	S_15_9009526	88.478	Simplex
6	D_1012_16942	89.552	Duplex
6	S_729_21463	90.040	Simplex

6	S_1012_16920	90.481	Simplex
6	SS_4_6784758	90.509	Double-Simplex
6	S_5_62621797	90.696	Simplex
6	SS_8932_1127	91.457	Double-Simplex
6	D_16_61789374	94.322	Duplex
6	S_1012_16955	94.979	Simplex
6	D_1012_16961	94.992	Duplex
6	S_1012_16948	94.992	Simplex
6	S_4_93162495	94.992	Simplex
6	S_7_46721739	94.992	Simplex
7	S_18_120027251	0.000	Simplex
7	S_17_87865446	1.001	Simplex
7	S_17_85450080	3.758	Simplex
7	S_17_83220333	7.375	Simplex
7	S_17_87865449	8.325	Simplex
7	DD_18_121213690	9.513	Double-Duplex
7	S_17_85450110	12.778	Simplex
7	S_18_121213684	12.888	Simplex
7	S_18_112080439	14.192	Simplex
7	S_18_118763675	14.696	Simplex
7	S_3_22411506	14.998	Simplex
7	S_18_120027266	20.625	Simplex
7	SD_3_22411481	20.815	Simplex-Duplex
7	S_17_80826651	22.456	Simplex
7	SS_18_111104468	23.678	Double-Simplex
7	S_18_112080411	24.448	Simplex
7	T_17_80826654	25.494	Triplex
7	S_18_101338018	25.526	Simplex
7	S_17_76129467	27.662	Simplex
7	S_17_76694291	28.193	Simplex
7	S_18_100802406	29.129	Simplex
7	D_17_76129446	29.138	Duplex
7	XSS_17_74821220	29.561	X-double-simplex
7	SS_18_100802431	29.607	Double-Simplex
7	S_18_99756507	29.763	Simplex
7	D_17_76129443	29.782	Duplex
7	S_17_76129436	30.711	Simplex
7	DS_6_8578250	30.896	Duplex-Simplex
7	S_18_99698133	30.978	Simplex
7	SD_18_95426449	33.284	Simplex-Duplex
7	S_18_95426455	33.490	Simplex
7	SS_18_95426454	33.647	Double-Simplex
7	XSS_18_95426406	34.030	X-double-simplex
7	SD_17_70977112	34.396	Simplex-Duplex
7	SS_17_70985284	34.712	Double-Simplex
7	SS_17_70720677	35.173	Double-Simplex

7	D_18_82948002	35.433	Duplex
7	S_18_91402915	36.986	Simplex
7	SS_17_67201201	37.739	Double-Simplex
7	S_18_95426463	37.739	Simplex
7	SD_17_67201234	38.154	Simplex-Duplex
7	S_14_60572187	38.904	Simplex
7	S_17_68810770	39.654	Simplex
7	S_17_67201205	40.408	Simplex
7	S_18_81378340	43.141	Simplex
7	SS_17_58621805	43.403	Double-Simplex
7	S_18_82947993	43.574	Simplex
7	SS_17_52267612	44.492	Double-Simplex
7	SD_17_52267600	44.906	Simplex-Duplex
7	SD_6_17981071	47.355	Simplex-Duplex
7	S_6_17981050	50.198	Simplex
7	D_18_80987976	50.970	Duplex
7	S_18_77621401	52.063	Simplex
7	S_15_47637463	52.153	Simplex
7	S_18_77867012	52.361	Simplex
7	S_18_80987999	52.603	Simplex
7	S_18_77621371	53.321	Simplex
7	S_18_77621368	53.781	Simplex
7	S_18_77621378	54.762	Simplex
7	S_17_33650152	54.835	Simplex
7	D_17_33650153	55.030	Duplex
7	S_17_13634622	55.643	Simplex
7	SD_17_26437733	56.386	Simplex-Duplex
7	D_17_33650164	56.493	Duplex
7	S_17_13634643	56.554	Simplex
7	SS_17_36643243	56.633	Double-Simplex
7	SD_17_26437730	56.826	Simplex-Duplex
7	S_17_32339855	56.982	Simplex
7	SS_18_63638321	57.077	Double-Simplex
7	S_11_1645159	57.127	Simplex
7	SS_17_35348690	57.246	Double-Simplex
7	SD_17_57050330	57.324	Simplex-Duplex
7	S_17_20035813	57.377	Simplex
7	XSS_18_63638309	57.426	X-double-simplex
7	S_17_20079074	57.579	Simplex
7	S_18_64979507	57.689	Simplex
7	S_17_33650161	57.831	Simplex
7	SS_17_26437760	57.976	Double-Simplex
7	S_17_35348688	58.015	Simplex
7	S_17_44470361	58.065	Simplex
7	SS_18_64979505	58.123	Double-Simplex
7	SS_18_77021672	58.394	Double-Simplex

7	T_17_57050318	58.409	Triplex
7	SD_18_77621374	58.749	Simplex-Duplex
7	S_15_26156810	58.963	Simplex
7	SD_17_20035861	59.222	Simplex-Duplex
7	S_3_84786608	59.767	Simplex
7	S_18_34888648	59.767	Simplex
7	S_18_34888674	59.767	Simplex
7	S_9_33701744	59.785	Simplex
7	S_11_3868749	60.029	Simplex
7	S_15_26156799	60.183	Simplex
7	S_18_34888650	60.274	Simplex
7	SS_17_20079028	60.401	Double-Simplex
7	S_18_62834739	60.869	Simplex
7	SD_17_33650171	61.383	Simplex-Duplex
7	S_15_26156795	61.548	Simplex
7	S_17_18361948	62.816	Simplex
7	SS_18_32919135	63.252	Double-Simplex
7	S_18_24252645	63.748	Simplex
7	S_17_21483242	64.256	Simplex
7	S_17_13097198	65.098	Simplex
7	S_18_34888651	65.654	Simplex
7	D_17_18361947	66.582	Duplex
7	SS_18_26036161	67.255	Double-Simplex
7	S_18_33217951	67.563	Simplex
7	S_18_33217952	67.921	Simplex
7	SD_17_13612927	68.078	Simplex-Duplex
7	S_12_35392613	69.943	Simplex
7	S_17_13097222	70.334	Simplex
7	SD_18_77621392	71.783	Simplex-Duplex
7	D_17_13612873	71.849	Duplex
7	S_17_13612871	72.002	Simplex
7	S_14_59955367	72.067	Simplex
7	S_17_15078753	72.441	Simplex
7	S_10_48855583	72.723	Simplex
7	SS_6985_893	73.057	Double-Simplex
7	S_17_10017618	74.566	Simplex
7	DD_12_35392604	74.607	Double-Duplex
7	SD_13_60913111	74.941	Simplex-Duplex
7	S_17_18449096	75.122	Simplex
7	SS_18_14147558	75.900	Double-Simplex
7	SS_18_77621399	76.768	Double-Simplex
7	D_13_60913087	78.230	Duplex
7	S_13_60913084	78.764	Simplex
7	SS_18_13282119	79.433	Double-Simplex
7	S_18_15238741	80.935	Simplex
7	SS_14_59955373	81.397	Double-Simplex

7	S_3_84786627	84.574	Simplex
7	S_10_72717337	84.969	Simplex
7	DD_10_63937305	85.231	Double-Duplex
7	S_4_56990197	87.556	Simplex
7	S_5_14128106	87.556	Simplex
7	S_8932_1173	87.556	Simplex
7	S_10_84824582	87.556	Simplex
7	S_4_93162488	87.994	Simplex
7	S_2982_12078	88.228	Simplex
7	S_13_60913101	88.287	Simplex
7	SS_18_14729584	91.031	Double-Simplex
7	S_18_14729586	92.179	Simplex
7	D_18_14147593	92.586	Duplex
7	S_18_10821450	95.454	Simplex
7	DD_18_9872330	95.550	Double-Duplex
7	DD_1218_7538	95.733	Double-Duplex
7	S_18_11521178	95.938	Simplex
7	S_2982_12082	97.411	Simplex
7	D_4_1173803	99.122	Duplex
7	DD_18_999595	99.367	Double-Duplex
7	S_18_7493129	99.875	Simplex
7	S_18_11521181	100.356	Simplex
7	S_18_5394843	101.194	Simplex
7	SD_18_6697980	102.749	Simplex-Duplex
7	S_18_4864802	103.136	Simplex
7	S_18_999568	105.952	Simplex
7	SS_18_999547	107.284	Double-Simplex
7	S_4_102693171	107.474	Simplex
7	D_4_1173774	108.573	Duplex
7	SS_17_3201106	108.573	Double-Simplex
7	S_5_72020398	108.573	Simplex
8	D_6_45098574	0.000	Duplex
8	SS_11_68578576	0.659	Double-Simplex
8	D_11_4020718	6.579	Duplex
8	D_11_68249454	8.890	Duplex
8	XSS_11_64513242	9.732	X-double-simplex
8	S_11_64513224	10.475	Simplex
8	S_11_64513227	12.084	Simplex
8	S_5_30149706	13.483	Simplex
8	S_11_68240954	15.013	Simplex
8	S_4_74112838	15.865	Simplex
8	SS_5_53808986	16.694	Double-Simplex
8	S_4_74112851	17.021	Simplex
8	S_17_17529539	17.216	Simplex
8	S_11_69371944	17.506	Simplex
8	DD_11_4020716	17.745	Double-Duplex

8	S_5_53808985	20.449	Simplex
8	S_5_30149755	23.688	Simplex
8	D_11_64513223	25.658	Duplex
8	SS_5_30149703	26.142	Double-Simplex
8	SS_11_68240975	26.608	Double-Simplex
8	D_5_53809007	26.853	Duplex
8	S_11_4020727	27.377	Simplex
8	SD_11_69371976	28.844	Simplex-Duplex
8	S_17_17529536	29.008	Simplex
8	S_11_68578581	30.691	Simplex
8	SS_4_83368321	31.150	Double-Simplex
8	D_840_28909	32.555	Duplex
8	S_5_30149716	32.941	Simplex
8	S_4_83368339	33.289	Simplex
8	SS_11_4020710	33.746	Double-Simplex
8	SS_11_68249658	34.264	Double-Simplex
8	S_11_64513215	35.370	Simplex
8	S_11_60525476	36.148	Simplex
8	SD_11_68240957	36.958	Simplex-Duplex
8	S_4640_5336	37.030	Simplex
8	S_3636_3330	37.920	Simplex
8	D_12_63723745	38.167	Duplex
8	SS_12_39928936	38.872	Double-Simplex
8	S_14369_618	39.717	Simplex
8	S_9_62017560	40.520	Simplex
8	S_6_8578247	41.513	Simplex
8	S_12_64417025	46.112	Simplex
8	S_12_39928937	47.308	Simplex
8	S_4640_5330	49.180	Simplex
8	S_840_28893	53.889	Simplex
8	S_12_16675989	55.047	Simplex
8	SS_12_16676006	60.624	Double-Simplex
8	S_729_21469	63.119	Simplex
8	S_729_21481	70.052	Simplex

Table S5. Genes identified in the apomixis and QTLs regions based on similarity to the *Panicum virgatum* genome.

HG	Regions	Name	Gene	Description
I	1	qNDF_L10 qNDF_S15	Pavir.1KG151000	RHOMBOID FAMILY PROTEIN
			AT5G38510	Rhomboid-related intramembrane serine protease family protein
			OsRhmbd6	Putative Rhomboid homologue
I	2	qADF_L20 qCEL_L30	AT1G70940/AT1G77110/AT5G16530/ AT5G15100/AT5G57090	Auxin efflux carrier family protein
			Os02g50960/Os12g04000/Os08g41720	Auxin efflux carrier component
			AT1G78900	vacuolar ATP synthase subunit A
			Os06g43680	palmitoyltransferase TIP1
			Pavir.4KG247000	PHOSPHOENOLPYRUVATE CARBOXYLASE
			Os01g02050	PHOSPHOENOLPYRUVATE CARBOXYLASE
			Os06g39640	PINHEAD
			AT1G15690	Inorganic H pyrophosphatase family protein
Os02g09150	Inorganic H ⁺ pyrophosphatase			
I	3	qOM_S3	AT3G02620/AT3G02610/AT3G02630	Plant stearoyl-acyl-carrier-protein desaturase family protein
			Os08g10010/Os04g31070	acyl-desaturase_chloroplast precursor
			Pavir.1NG248700	EXOCYST COMPLEX PROTEIN EXO70
			AT5G50380	exocyst subunit exo70 family protein F1
			AT1G72470	exocyst subunit exo70 family protein D1
			AT1G54090	exocyst subunit exo70 family protein D2
			AT3G14090	exocyst subunit exo70 family protein D3
			Os02g30230/Os01g69230/Os08g41820	exo70 exocyst complex subunit domain containing protein

HG	Regions	Name	Gene	Description
II	4	<i>Apo-locus</i>	AT3G43610	Spc97 / Spc98 family of spindle pole body (SBP) component
			Pavir.6KG334200.1	Inner centromere protein, ARK binding region (INCENP_ARK-bind)
			AT1G71830	Somatic embryogenesis receptor-like kinase 1 (SERK)
II	5	qNDF_L11 qNDF_S16	Pavir.5NG291500.1	Wall-associated receptor kinase galacturonan-binding (GUB_WAK_bind)
			Pavir.6KG334900	RING/U-box domain-containing protein (XERICO)
			AT3G19430/AT5G54370/AT5G60520	Late embryogenesis abundant (LEA) protein-related
II	6	qTDM3 qRC9	Pavir.3KG465900	WRKY TRANSCRIPTION FACTOR 3-RELATED
			Pavir.3KG466000	cell cycle arrest protein BUB3 (BUB3) // translation initiation factor 5A (EIF5A)
			AT1G13950	eukaryotic elongation factor 5A-1
			AT4G26640	WRKY family transcription factor family protein
			AT5G56270	WRKY DNA-binding protein 2
			AT1G13960	WRKY DNA-binding protein 4
			AT5G07100	WRKY DNA-binding protein 26
			AT2G38470	WRKY DNA-binding protein 33
			AT3G01080	WRKY DNA-binding protein 58
			Os01g61080	WRKY24
			Os04g39570	WRKY35
			Os05g27730	WRKY53
Os03g33012	WRKY81			
III	7	qIVD_L27	Pavir.1NG033800	Glycosyltransferase
			AT2G41640/AT3G18170/AT3G10320/	Glycosyltransferase family 61 protein

HG	Regions	Name	Gene	Description
			AT3G57380	
			Os10g35020/Os02g22650/Os06g49320	Glycosyltransferase
III	8	qCP_L6	Pavir.1NG040300	xyloglucan:xyloglucosyl transferase
		qNDF_L12	AT1G11545	xyloglucan endotransglucosylase/hydrolase 8
		qNDF_S17	AT1G14720	xyloglucan endotransglucosylase/hydrolase 28
			AT4G18990	xyloglucan endotransglucosylase/hydrolase 29
			AT1G32170	xyloglucan endotransglucosylase/hydrolase 30
			AT1G10550	xyloglucan:xyloglucosyl transferase 33
III	9	qPL_L33	Pavir.2KG486900	PROTEIN THYLAKOID FORMATION 1, CHLOROPLASTIC
			Pavir.2NG571900	Secoisolariciresinol dehydrogenase
			AT2G14900/AT2G39540/AT5G59845	Gibberellin-regulated family protein
			AT3G02885	GAST1 protein homolog 5
			Os10g02625	Gibberellin regulated protein
			Os03g41060	GASR2 - Gibberellin-regulated
				GASA/GAST/Snakin family protein precursor
			Os03g55290	GASR3 - Gibberellin-regulated
				GASA/GAST/Snakin family protein precursor
			Os07g40240	GASR9 - Gibberellin-regulated
				GASA/GAST/Snakin family protein precursor
			AT4G03140/AT3G26760/AT3G26770/AT3G51680	NAD(P)-binding Rossmann-fold superfamily protein
III	10	qSIL_L35	Pavir.2NG549800	LEUCINE-RICH REPEAT-CONTAINING PROTEIN
			Os01g72680/Os09g14100	Disease resistance protein RPS2
			AT1G12220	Disease resistance protein (CC-NBS-LRR class) family
			Os03g14900/Os12g10410	NB-ARC/LRR disease resistance protein
			Os11g37880/Os11g37850/Os11g37774	Stripe rust resistance protein

HG	Regions	Name	Gene	Description
IV	11	qPLB8	Pavir.7KG092300	Ent-copalyl diphosphate synthase
		qNDF_L13	AT5G67440/AT5G64330	Phototropic-responsive NPH3 family protein
		qNDF_S18		
V	12	qCP_S8	Pavir.5KG199700	Glucuronokinase
		qIVD_S28	AT3G01640	Glucuronokinase G
			Pavir.5NG207200	WRKY DNA-binding domain
			Os01g18584	WRKY9
			Os02g53100	WRKY32
			AT1G68150	WRKY DNA-binding protein 9
			AT4G04450	WRKY family transcription factor
			Os05g04640	WRKY5
			Os05g49210	WRKY43
			Os01g14440	WRKY1
			AT1G62300	WRKY family transcription factor
			AT4G22070	WRKY DNA-binding protein 31
			AT4G01720	WRKY family transcription factor
			AT1G18860	WRKY DNA-binding protein 61
			AT3G62150	P-glycoprotein 21
			AT1G02530	P-glycoprotein 12
			AT1G02520	P-glycoprotein 11
			AT4G01830	P-glycoprotein 5
			AT5G46540	P-glycoprotein 7
			AT4G01820	P-glycoprotein 3
			AT3G28380	P-glycoprotein 17
			AT1G10680	P-glycoprotein 10
			AT4G25960	P-glycoprotein 2
			AT1G27940	P-glycoprotein 13

HG	Regions	Name	Gene	Description
			Pavir.5NG245000	PROTEIN NRT1/ PTR FAMILY 6.2
			Os02g48570/Os01g04950/Os04g36040	Peptide transporter PTR2
			AT2G23610	methyl esterase 3
			AT2G23580	methyl esterase 4
			AT2G23560	methyl esterase 7
			AT2G23550	methyl esterase 6
			AT4G37150	methyl esterase 9
			AT2G23590	methyl esterase 8
			AT2G23620	methyl esterase 1
V	13	qADF_S23	AT4G01070/AT2G18570/AT5G66690/ AT2G18560/AT2G18570/AT5G66690 Os02g14680/Os02g14590/Os02g14540	UDP-Glycosyltransferase superfamily protein UDP-glucuronosyl and UDP-glucosyl transferase domain containing protein
V	14	qGM1 qTDM4 qLDM5 qSDM7 qCP_L7 qADF_L21	Pavir.5NG490300 Pavir.5NG490600 Pavir.5NG491000 Pavir.5NG491300 AT1G62750 AT4G27520 Os06g46740 AT5G53870 AT2G25060 AT4G32490 Os01g53930/Os05g44760/Os05g45590/ Os05g09500/Os05g31110/Os01g52450/ Os01g09460/Os07g09890 AT4G29130	EARLY NODULIN-LIKE PROTEIN 3- RELATED HEXOKINASE Pectinesterase / Pectin methylesterase MALIC ENZYME Translation elongation factor EFG/EF2 protein Early nodulin-like protein 2 Early nodulin 20 precursor Early nodulin-like protein 1 Early nodulin-like protein 14 Early nodulin-like protein 4 Hexokinase Hexokinase 1

HG	Regions	Name	Gene	Description
			AT1G47840	Hexokinase 3
			AT3G20040	Hexokinase
			AT4G37840	Hexokinase-like 3
			AT2G19860	Hexokinase 2
			Os01g53990/Os05g44600/Os01g44340/ Os04g46740/Os08g10604/Os01g19440	Pectinesterase
			AT2G21610	Pectinesterase 11
			AT3G29090	Pectin methylesterase 31
			Os07g31380/Os05g09440/Os10g35960	NAD-dependent malic enzyme_mitochondrial precursor
			Os01g54030/Os02g44550/Os01g52500	NADP-dependent malic enzyme
			AT2G19900	NADP-malic enzyme 1
			AT1G79750	NADP-malic enzyme 4
			AT4G00570	NAD-dependent malic enzyme 2
			AT5G25880	NADP-malic enzyme 3
			AT5G11670	NADP-malic enzyme 2
			AT2G13560	NAD-dependent malic enzyme 1
			Os01g09320	NADP-dependent malic enzyme_chloroplast precursor
V	15	qOM_L1	Pavir.5KG761500	EF-HAND CALCIUM-BINDING DOMAIN CONTAINING PROTEIN
			AT4G12860	Calcium-binding EF-hand family protein
VI	16	qOM_L2	AT1G65060	4-coumarate:CoA ligase 3
		qADF_L22	AT3G21240	4-coumarate:CoA ligase 2
VI	17	qGM2	Pavir.8NG293200	Galactinol--raffinose galactosyltransferase /
		qLDM6		Stachyose synthetase
		qRC10	Os07g09010/Os10g08850/Os03g58580/ Os12g44060/Os12g44060/Os12g44070/ Os08g42010/Os03g58580/Os07g09010/	Nodulin
		qADF_S24		

HG	Regions	Name	Gene	Description
			Os12g29950	
			Os03g47810	Nodulin family protein
			AT1G74780	Nodulin-like / Major Facilitator Superfamily protein
			Pavir.8KG076200	ALPHA-DIOXYGENASE 2
			AT3G01420	Peroxidase superfamily protein
			AT1G73680	Alpha dioxygenase
VII	18	qSIL_L36	Pavir.9NG559100	PEPTIDYL-PROLYL CIS-TRANS ISOMERASE PASTICCINO1
			AT3G54010	FKBP-type peptidyl-prolyl cis-trans isomerase family protein
			Os03g25140	Peptidyl-prolyl isomerase PASTICCINO1
VII	19	qOM_S4	Pavir.9KG151500	Chloroplast envelope transporter
		qNDF_L14	Pavir.9KG151700	Endo-1,4-beta-xylanase
		qNDF_S19	Pavir.9KG151800	ALPHA-GALACTOSIDASE 1
			AT1G58370/AT4G08160/AT1G58370/AT1G10050/AT4G38650	Glycosyl hydrolase family 10 protein / carbohydrate-binding domain-containing protein
			Os05g23350/Os01g04300/Os05g25560	Glycosyl hydrolase family 10 protein
			Os10g35110/Os07g48160/Os10g35070/Os07g26900/Os10g35070/Os07g26900/Os07g48160/Os10g35110	Alpha-galactosidase precursor
			AT5G08370/AT5G08370	Alpha-galactosidase 2
			AT5G08380/AT5G08380	Alpha-galactosidase 1
VII	20	qADF_S25	Pavir.9NG150900	Linoleate 9S-lipoxygenase / Linoleate 9-lipoxygenase
			Os03g49380/Os12g37350/Os03g52860/Os03g49260/Os05g23880/Os08g39850/Os11g36719/Os03g49380	lipoxygenase
			Os06g04420/Os02g19790	lipoxygenase 4

HG	Regions	Name	Gene	Description
			AT3G45140	lipoxygenase 2
			AT1G55020/AT1G55020	lipoxygenase 1
			AT1G17420/AT1G17420	lipoxygenase 3
VIII	21	qCP_S9	Pavir.6KG366500	3-oxoacyl-[acyl-carrier-protein] reductase
		qIVD_S29	Pavir.6KG367000	Sugar transporter/spinster transmembrane protein
		qPL_S34	Pavir.6KG367600	ethylene-insensitive protein 3
			Os08g39830/Os03g20780/Os07g48630	Ethylene-insensitive 3
			AT1G73730	ETHYLENE-INSENSITIVE3-like 3
			AT5G10120	Ethylene insensitive 3 family protein
			AT3G20770	Ethylene insensitive 3
			Os10g02800	Glycine-rich cell wall structural transmembrane protein
			AT2G17845/AT3G55310/AT3G55290/ AT1G62610	NAD(P)-binding Rossmann-fold superfamily protein
			LOC_Os08g39960/Os04g30760/Os12g13930	3-oxoacyl-reductase_chloroplast precursor
VIII	22	qOM_S5	Pavir.6KG306400	STEROL REGULATORY ELEMENT-BINDING PROTEIN
		qADF_S26		
		qCEL_S31	Pavir.6KG306300	ALPHA CARBONIC ANHYDRASE 5
			AT1G08065	Alpha carbonic anhydrase 5
			AT2G28210	Alpha carbonic anhydrase 2
			AT4G20990	Alpha carbonic anhydrase 4
			AT1G08080	Alpha carbonic anhydrase 7
			AT2G26420	Phosphatidylinositol-4-phosphate 5-kinase 3
			Os03g24160/Os07g46490/AT2G26420	Phosphatidylinositol-4-phosphate 5-kinase
VIII	23	qSIL_S37	Pavir.6NG098000	PECTINESTERASE 8
			Os08g10604/Os05g44600/Os01g44340/ Os07g41650/Os09g39760/Os01g53990	pectinesterase

11 ANEXOS

11.1 Anexo I - Termo de aprovação da pesquisa pela Comissão de Bioética pertinente



CIDADE UNIVERSITÁRIA "ZEFERINO VAZ",
11 DE NOVEMBRO DE 2019.

CIBio: 06/2019

IDENTIFICAÇÃO

Doutorado: THAMIRIS GATTI DEO
Comissão de Pós-Graduação – Genética e Biologia Molecular
Instituto de Biologia – UNICAMP.

PROJETO

"Mapeamento da região da apomixia e de QTLs em *Megathyrus maximus*: uma gramínea forrageira tropical autotetraplóide."

PARECER

Projeto aprovado pela CIBio / CBMEG sob número 14/2003 – Mapeamento e análise da diversidade genética em plantas e outros organismos.

Coordenador: Profa. Dra. Anete Pereira de Souza.

PROFA. DRA. MÔNICA BARBOSA DE MELO
Presidente da CIBio / CBMEG – UNICAMP

11.2 Anexo II - Cumprimento dos dispositivos da lei nº 9610/98

Cumprimento dos dispositivos da lei n.o. 9.610/98

Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Mapeamento da região da apomixia e de QTLs em *Megathyrus maximus*: uma gramínea forrageira tropical autotetraploide**, não infringem os dispositivos da Lei n.o 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 13 de dezembro de 2019

Assinatura: Thamiris Gatti Deo

Nome do(a) autor(a): **Thamiris Gatti Deo**

RG n.º 40887637-2

Assinatura: Anete Pereira de Souza

Nome do(a) orientador(a): **Anete Pereira de Souza**

RG n.º 8680325-6

11.3 Anexo III



Article

Relationship between Genetic Variability and Land Use and Land Cover in Populations of *Campomanesia adamantium* (Myrtaceae)

Bruno do Amaral Crispim ¹, Miklos Maximiliano Bajay ², Adrielle Ayumi de Vasconcelos ³, Thamis Gatti Deo ³, Ramilla dos Santos Braga ⁴, Mariana Pires de Campos Telles ^{4,5}, Maria do Carmo Vieira ⁶, Thiago de Oliveira Carnevali ⁶, Julio César Jut Solórzano ⁷ and Alexeia Baruffatti Grisolia ^{1,*}

- ¹ Faculty of Biological and Environmental Sciences, Federal University of Grande Dourados, Dourados, Mato Grosso do Sul 79804-970, Brazil; brunocrispim.bio@gmail.com
- ² Departamento de Genética, ESALQ/Universidade de São Paulo-USP, Piracicaba, São Paulo 13418-900, Brazil; mmbajay@gmail.com
- ³ Department of Genetics, Evolution, Microbiology and Immunology, Institute of Biology, University of Campinas (UNICAMP), Campinas, São Paulo 13083-862, Brazil; adriayumi@gmail.com (A.A.d.V.); thamisdeo@hotmail.com (T.G.D.)
- ⁴ Laboratório de Genética & Biodiversidade, Departamento de Genética, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, Goiás 74045-155, Brazil; ramillabraga@gmail.com (R.d.S.B.); tellesmpc@gmail.com (M.P.d.C.T.)
- ⁵ Escola de Ciências Agrárias e Biológicas, Pontifícia Universidade Católica de Goiás, Goiânia, Goiás 74175-120, Brazil
- ⁶ Faculty of Agrarian Sciences, Federal University of Grande Dourados, Dourados, Dourados, Mato Grosso do Sul 79804-970, Brazil; mariavieira@ufgd.edu.br (M.d.C.V.); thiagocarnevali@live.com (T.d.O.C.)
- ⁷ Postgraduate Program in Natural Resources, State University of Mato Grosso do Sul, Dourados, Mato Grosso do Sul 79804-970, Brazil; julio.jut@gmail.com
- * Correspondence: alexeiagrisolia@ufgd.edu.br; Tel.: +55-67-3410-2223

Received: 30 May 2018; Accepted: 15 July 2018; Published: 27 September 2018



Abstract: *Campomanesia adamantium* is an endemic plant of Cerrado biome that has potential for cultivation because its fruits have culinary and medicinal uses. However, genetic diversity studies using molecular markers with Cerrado species are scarce, and the inadequate extractive exploitation of fruits and the expansion of agricultural frontiers may also affect genetic variability. Therefore, studies in this field are of interest as they can provide sources for conservation and breeding programs. In this context, we investigated the genetic diversity of native populations of *C. adamantium* from different sites and the relationship between genetic variability and the land use and land cover of each site. A total of 207 plants were sampled in seven sites and characterized with seven polymorphic microsatellite markers. The use and coverage of land were mapped based on aerial images, and the land was classified into different categories. The genetic diversity was high in all populations, with low levels of differentiation due to allele sharing, mainly in Mato Grosso do Sul and Paraguay populations. The geographically closest populations were more genetically similar. The use and coverage of land indicated that intense agriculture promotes a significant decrease in genetic variability.

Keywords: genetic variability; guavira; heterologous amplification; population genetics; SSR

11.4 Anexo IV

1280

Brazilian Applied Science Review**Reguladores de crescimento na multiplicação *in vitro* de *Campomanesia adamantium* (Cambess.) O. Berg (Myrtaceae)****Growth regulators *in vitro* multiplication of *Campomanesia adamantium* (Cambess.) O. Berg (Myrtaceae)**

Recebimento dos originais: 19/01/2019
 Aceitação para publicação: 21/02/2019

Ademir Goelzer

Bacharel em Biotecnologia
 Instituição: Universidade Federal de Lavras
 Endereço: Campus Universitário, 4, Lavras-MG, Brasil
 E-mail: ademigoelzer@gmail.com

Thamiris Gatti Déo

Bacharel em Biotecnologia
 Instituição: Universidade Estadual de Campinas
 Endereço: Avenida Cândido Rondon, 400 - Cidade Universitária, Campinas - SP, Brasil
 E-mail: thamisrdeo@hotmail.com

Graciela Beatris Lopes

Bacharel em Biotecnologia
 Instituição: Universidade Federal de Lavras
 Endereço: Campus Universitário, 4, Lavras-MG, Brasil
 E-mail: gracielabealopes@gmail.com

Cláudia Roberta Damiani

Doutora em Biotecnologia Molecular pela Università di Pisa (Pisa – Itália)
 Instituição: Universidade Federal da Grande Dourados
 Endereço: Rodovia Dourados/Itahum, Km 12 - Unidade II, Dourados-MS, Brasil
 E-mail: claudiadamiani@ufgd.edu.br

RESUMO

Campomanesia adamantium (guavira), é frutífera nativa do Cerrado, encontrada em grande maioria no estado silvestre e apresenta limitações na propagação relacionada à recalcitrância das sementes. Considerando a ausência de manejo e tratamentos culturais, o estabelecimento da cultura *in vitro* a partir de material vegetativo está sujeito a altos índices de contaminação fúngica e bacteriana. Visando a propagação, objetivou-se avaliar a capacidade de multiplicação *in vitro* de guavira, utilizando segmentos caulinares retirados de plântulas da germinação *in vitro*. Foram desenvolvidos três experimentos, sendo avaliados o efeito de 5,0 µM de TDZ (*thiadizuram*), 1,0 µM de ANA (ácido naftaleno acético) e a combinação dos mesmos, bem como, o efeito isolado de BAP (6-benzilaminopurina) e 2iP (2-isopentenil adenina), nas concentrações de 0; 2,5; 5,0; 7,5 e 10 mg L⁻¹. Os resultados obtidos demonstraram que o uso de 5,0 µM de TDZ aumentou o comprimento das brotações, número de gemas e taxa de multiplicação, enquanto a combinação de TDZ e ANA exerceu um efeito negativo sobre o crescimento das brotações, número de folhas e taxa de multiplicação. O cultivo dos explantes em meio contendo BAP aumentou o número de brotações, ao aumento da concentração do regulador, entretanto, reduziu o comprimento das brotações.

11.5 Anexo V



Development and characterization of microsatellite markers in *Campomanesia adamantium*, a native plant of the Cerrado ecoregions of South America

Bruno do Amaral Crispim¹ , Thamiris Gatti Déo², Juliana dos Santos Fernandes², Adrielle Ayumi de Vasconcelos², Maria do Carmo Vieira³ , Thiago de Oliveira Carnevali², Miklos Maximiliano Bajay⁴ , Maria Imaculada Zucchi⁵ , and Alexeia Barufatti^{1,2,6}

Manuscript received 13 May 2019; revision accepted 28 June 2019.

¹ Faculdade de Ciências Exatas e Tecnologia, Universidade Federal da Grande Dourados, Dourados, Mato Grosso do Sul 79804-970, Brazil

² Faculdade de Ciências Biológicas e Ambientais, Universidade Federal da Grande Dourados, Dourados, Mato Grosso do Sul 79804-970, Brazil

³ Faculdade de Ciências Agrárias, Universidade Federal da Grande Dourados, Dourados, Mato Grosso do Sul 79804-970, Brazil

⁴ Universidade Estadual de Santa Catarina, Rua Cel. Fernandes Martins, 270 - Progresso - Laguna, Santa Catarina 88.790-000, Brazil

⁵ Agência Paulista de Tecnologia dos Agronegócios, Pólo Centro-Sul (APTA), Piracicaba, São Paulo, Brazil

⁶ Author for correspondence: barufattialexia@gmail.com

Citation: Crispim, B. A., T. G. Déo, J. dos Santos Fernandes, A. A. Vasconcelos, M. C. Vieira, T. O. Carnevali, M. M. Bajay, M. I. Zucchi, and A. Barufatti. 2019. Development and characterization of microsatellite markers in *Campomanesia adamantium*, a native plant of the Cerrado ecoregions of South America. *Applications in Plant Sciences* 7(9): e11287.

doi:10.1002/aps3.11287

PREMISE: A novel set of nuclear microsatellite markers was developed and characterized for *Campomanesia adamantium* (Myrtaceae) and tested for cross-amplification in the related species *C. sessiliflora*.

METHODS AND RESULTS: Forty-one primer pairs were designed for simple sequence repeat loci, of which 36 successfully amplified and were polymorphic. The number of alleles ranged from two to 14, with an average of 8.14 alleles per locus. Additionally, cross-amplification was tested in *C. sessiliflora*; more than 55.5% of the microsatellite loci amplified, confirming the use of these microsatellite markers in a related species.

CONCLUSIONS: We developed a set of microsatellite markers that will be useful for future studies of genetic diversity and population structure of *C. adamantium* and a closely related species, which will aid in future conservation efforts.

KEY WORDS *Campomanesia sessiliflora*; cross-amplification; genetic structure; guavira; microsatellite loci; Myrtaceae; simple sequence repeat (SSR) markers.

The Cerrado is the second largest of Brazil's major biomes, after Amazonia, and this biome is characterized by central Brazil's plateau of woodlands, savannas, grasslands, and gallery and dry forests. The climate is seasonal—wet from October to March and dry from April to September—and mild throughout the year, with temperatures ranging from 22°C to 27°C. Average annual rainfall is 1500 mm (Klink and Machado, 2005). The Cerrado biome features a high diversity of native plants with high economic potential for the pharmaceutical and food markets. This biome is a hotspot of biodiversity because endemic species have progressively been threatened by deforestation and agricultural expansion. Research on the genetic and floral biology of these species is extremely important to promote their conservation. Of several native plants of the Cerrado, *Campomanesia adamantium* (Cambess.) O. Berg (Myrtaceae) should be highlighted because of its environmental significance. This species can resist flooding and has been used for spoiled-area reforestation projects (Hardt et al., 2006). Also known

as guavira or gabirola, *C. adamantium* is commonly found in Brazil and Paraguay and is notable for the potential commercial use of its fruits to produce food and beverages. Many researchers have reported important medicinal properties of this plant, including antioxidant (Coutinho et al., 2008), antimicrobial (Pavan et al., 2009), and antitumor properties (Pascoal et al., 2014).

To start a genetic breeding program for any plant species, it is necessary to assess its genetic variability. Genetic diversity in *C. adamantium* was detected previously using random amplified polymorphic DNA (RAPD) markers (Assis et al., 2013) and by measuring phenotypic characteristics (Resende and Teixeira, 2015). Microsatellite markers have been shown to be efficient tools to assess genetic variation at the individual and population levels because of their hypervariability, wide genomic distribution, co-dominant inheritance, reproducibility, and multi-allelic nature (Haq et al., 2014). Recent research has validated the transferability of microsatellite markers from closely related species of *C. adamantium*, including

Applications in Plant Sciences 2019 7(9): e11287; <http://www.wileyonlinelibrary.com/journal/AppsPlantSci> © 2019 Crispim et al.

Applications in Plant Sciences is published by Wiley Periodicals, Inc. on behalf of the Botanical Society of America. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.