



CLARA MITRE DO PRADO

FRAÇÃO DO DNA REPETITIVO NA BATATA SILVESTRE
***Solanum calvescens* Bitter E ESPÉCIES RELACIONADAS**

LAVRAS – MG

2021

CLARA MITRE DO PRADO

**FRAÇÃO DO DNA REPETITIVO NA BATATA SILVESTRE *Solanum calvescens*
Bitter E ESPÉCIES RELACIONADAS**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

Prof^ª. Dr^ª. Giovana Augusta Torres
Orientadora

LAVRAS – MG

2021

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Prado, Clara Mitre do.

Fração do Dna Repetitivo na Batata Silvestre *Solanum
calvescens* Bitter e Espécies Relacionadas / Clara Mitre do Prado. -
2021.

46 p.

Orientador(a): Giovana Augusta Torres.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2021.

Bibliografia.

1. Genética Vegetal. 2. Genômica Comparativa. 3. DNA
Repetitivo. I. Torres, Giovana Augusta. II. Título.

CLARA MITRE DO PRADO

**FRAÇÃO DO DNA REPETITIVO NA BATATA SILVESTRE *Solanum calvescens*
Bitter E ESPÉCIES RELACIONADAS**

**REPETITIVE DNA FRACTION IN THE WILD POTATO *Solanum calvescens* Bitter
AND RELATED SPECIES**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Genética e Melhoramento de Plantas, para a obtenção do título de Mestre.

APROVADA em 29 de Abril de 2021

Dr^a Paola Gaiero Guadagna - Universidad de la República – Uruguay

Dr^a Ludmila Cristina Oliveira - Centro de Biologia da Academia Tcheca de Ciências (Biology Centre CAS)

Prof^a. Dr^a. Giovana Augusta Torres

Orientadora

LAVRAS – MG

2021

*À Ciência, que se mantém firme e atuante em benefício à
humanidade, mesmo em tempos de ignorância coletiva.*

Às mulheres que nos antecederam e abriram nosso caminho.

*Aos direitos básicos de saúde, educação e lazer gratuitos à todas
as pessoas.*

Ao direito à vida, que se tornou tão preciosa.

Dedico.

AGRADECIMENTOS

Aos meus pais pelo apoio, paciência e educação investidos em minha formação.

Aos amigos da graduação, que me ensinaram os valores de viver em sociedade e a luta pelos direitos humanos dignos à todas as pessoas.

Aos amigos e colegas do Laboratório de Citogenética Vegetal, que ofereceram apoio, conhecimento e convivência.

Aos amigos do mestrado, Alex, Ana Gabriela, Liliana, Daniela, Ana Luísa, Caio, Laiane, Isabela, Yasmin e Marco Túlio, pelo apoio e convivência virtual que permitiram que a vida fosse mais leve em meio à pandemia.

Aos amigos do RPG, que permitiram finais de semana de escape com muita parceria e risadas.

Às amigas Cynthia e Patrícia, que mesmo a distância não é capaz de dissolver.

Ao meu companheiro Raphael, pelo total apoio, compreensão, parceria e amor.

À Profa. Giovana, por toda confiança, orientação e compreensão que permitiram este degrau de formação acadêmica.

Aos membros da banca pela confiança e participação deste processo.

Aos Técnicos do Laboratório, que permitiram a manutenção do espaço de trabalho.

À todos os técnicos e servidores que permitiram a rápida adaptação às demandas da crise de saúde.

Ao Programa de Pós-graduação em Genética e Melhoramento de Plantas pela confiança.

Ao CNPq pela concessão da bolsa que permitiu a realização e dedicação a este trabalho.

À Universidade Federal de Lavras por toda estrutura e investimento em seus estudantes, professores e servidores.

À sociedade, a qual as Universidades Públicas devem todo seu conhecimento.

*“Nós temos cálcio em nossos ossos,
ferro em nossas veias,
carbono em nossas almas
e nitrogênio em nossos cérebros.
93% pó de estrelas,
com almas feitas de chamas,
nós somos todos estrelas
que tem nomes de pessoas.”*

(Nikita Gill)

RESUMO

A busca de alelos de interesse em espécies nativas de batata é uma estratégia promissora nos programas de melhoramento genético visando cultivares superiores, especialmente com relação a estresses bióticos e abióticos. A espécie triploide nativa *Solanum calvescens* ($2n=3x=36$) é uma das espécies potenciais para uso no melhoramento da batata, porém não tem status taxonômico conclusivo. Ela não é considerada como espécie por alguns autores, sendo definida ora como um citótipo de *Solanum chacoense* ($2n=2x=24$), ora como próxima de *Solanum commersonii* subsp. *malmeanum*, ou até mesmo um híbrido entre essas duas espécies. O cruzamento de *S. calvescens* com a batata cultivada (*Solanum tuberosum*, $2n=4x=48$) produz híbridos viáveis para seleção em programa de melhoramento. Devido a esse potencial, é relevante a definição taxonômica e conhecimento da origem dessa espécie. Portanto, este estudo visou caracterizar a fração repetitiva do DNA desta espécie nativa, de forma comparativa com suas espécies relacionadas, a fim de reunir novas informações genômicas para colaborar com a elucidação da relação entre elas. Para a espécie *S. calvescens* foi realizado sequenciamento de baixa cobertura na plataforma Illumina e determinação do tamanho do genoma por citometria de fluxo. Para as espécies *S. chacoense*, *S. commersonii*, *S. malmeanum* e *S. etuberosum*, foram utilizados dados publicados e disponibilizados pelos autores. As sequências foram analisadas na plataforma *Repeat Explorer*, na qual foram executadas as etapas de pré-processamento, agrupamento e identificação dos elementos repetitivos. O software Geneious foi utilizado para o alinhamento, identificação e quantificação dos satélites. O tamanho do genoma de *S. calvescens* se revelou maior em comparação às espécies relacionadas. A quantidade de DNA repetitivo presente nas espécies variou entre 38,37% do genoma em *S. calvescens*, até 45,37% em *S. chacoense*, com predominância do grupo LTR e superfamília *Ty3/gypsy* em todas as espécies. Todos os satélites mapeados obtiveram identidade inferior a 50%, agrupando-os somente em nível de superfamília. A espécie *S. chacoense* apresentou um satélite espécie-específico, SchacSat10. Com relação ao tamanho do genoma e características da fração repetitiva, *S. calvescens* se diferencia das demais espécies.

Palavras-chave: DNA, fração repetitiva, batata, *Solanum*, genômica, *Repeat Explorer*

ABSTRACT

The search for alleles of interest in wild potato species is a promising strategy in improvement programs targeting superior cultivars, especially with regard to biotic and abiotic stresses. The triploid native species *Solaunum calvescens* ($2n=3x=36$) is one of the potential species for use in potato improvement, but it has no conclusive taxonomic status. Not considered as a species by some authors, being sometimes defined as a cytotype of *Solanum chacoense* ($2n=2x=24$), sometimes as close to *Solanum commersonii* subsp. *malmeanum*, or even a hybrid between these two species. Crossing *S. calvescens* with cultivated potato (*Solanum tuberosum*, $2n=4x=48$) produces viable hybrids for selection in a breeding program. Due to this potential, the taxonomic definition and knowledge of the origin of this species is relevant. Therefore, this study aimed to characterize the repetitive fraction of the DNA of this native species, comparatively with its related species, in order to gather new genomic information to collaborate with the elucidation of the relationship between them. For *S. calvescens*, low pass sequencing was performed on the Illumina platform and genome size was determined by flow cytometry. For the species *S. chacoense*, *S. commersonii*, *S. malmeanum* and *S. etuberosum*, data published and made available by the authors were used. The sequences were analyzed in the Repeat Explorer platform, in which the steps of pre-processing, grouping and identification of repetitive elements were performed. Geneious software was used for satellite alignment, identification and quantification. The genome size of *S. calvescens* proved to be larger compared to related species. The amount of repetitive DNA present in the species ranged from 38.37% of the genome in *S. calvescens* to 45.37% in *S. chacoense*, with a predominance of the LTR group and Ty3/gypsy superfamily in all species. All mapped satellites had less than 50% identity, grouping them only at the superfamily level. The species *S. chacoense* presented a species-specific satellite, SchacSat10. Regarding genome size and repetitive fraction characteristics, *S. calvescens* differs from other species.

Keywords: DNA, repeat, potato, *Solanum*, genomic, *Repeat Explorer*

SUMÁRIO

1	INTRODUÇÃO.....	10
2	REFERENCIAL TEÓRICO.....	12
2.1	Genômica de <i>Solanum</i>	12
2.2	Espécies selvagens da Série <i>Commersoniana</i>	14
2.3	DNA Repetitivo.....	18
2.3.1	Elementos Transponíveis (ETs).....	18
2.3.2	DNA Satélites.....	20
2.3.3	DNA repetitivo em <i>Solanum</i>	21
3	MATERIAL E MÉTODOS.....	23
3.1	Material Genético.....	23
3.2	Citometria de Fluxo.....	25
3.3	Análise da fração repetitiva através da <i>pipeline</i> Repeat Explorer.....	25
3.3.1	Pré-processamento.....	25
3.3.2	Quantificação do DNA repetitivo.....	26
4	RESULTADOS.....	27
4.1	Tamanho do Genoma de <i>Solanum calvescens</i>	27
4.2	Quantificação dos Elementos Transponíveis.....	28
4.3	Sequências em Tandem.....	32
5	DISCUSSÃO.....	35
6	CONCLUSÕES.....	38
	REFERÊNCIAS BIBLIOGRÁFICAS.....	39

1 INTRODUÇÃO

O gênero *Solanum*, família Solanaceae, contém mais de 1500 espécies catalogadas, das quais 160, todas da Seção Petota, produzem tubérculos. Das 20 espécies cultivadas, *S. tuberosum* L. é a mais difundida no mundo, originária da América do Sul, da Cordilheira dos Andes e domesticada entre 2000 e 5000 a.c. através da procura dos tubérculos comestíveis (desprovidos de glicoalcaloides). Foi introduzida na Europa no século XVI pelos colonizadores e reconhecida como cultura de segurança alimentar no século XVIII, em cenários de perda dos cultivos tradicionais de grãos (CACACE e HUARTE, 1996). A partir do século XX, já estava disseminada e consolidada como a terceira cultura alimentar mais importante do mundo (FAOSTAT, 2020), sendo mais valiosa ainda nos países em desenvolvimento, com produções de até 100 milhões de toneladas somente na Índia e China. No Brasil, é a segunda hortaliça de maior importância econômica, com uma produção de cerca de 4,3 milhões de toneladas em área de 140,3 mil hectares em 2017 (IBGE, 2020). Os programas de melhoramento genético têm desenvolvido cultivares visando aumento de produtividade, qualidade do tubérculo e resistência ou tolerância a estresses bióticos e abióticos. Nesse sentido, a busca de alelos de interesse em espécies selvagens constitui uma estratégia promissora para incrementar a variabilidade disponível para os programas (FAO, 2008).

A estratégia de hibridação da batata cultivada (*S. tuberosum*) com espécies selvagens tuberosas tem mostrado importantes resultados para introdução de alelos de interesse, como resistência a pragas, doenças e melhoria das características agronômicas (WEBER et al, 2012), além de ser uma estratégia para o desenvolvimento de cultivares que atendam à crescente demanda por plantio em zonas climáticas com alta temperatura (Brücher, 1999). Todas as espécies nativas que tuberizam, e ocorrem no Brasil, estão inseridas na Série *Commersoniana*. Esta série abrange *S. commersoni* Dunal, encontrada na América do Sul em formas diploides e mais raramente triploides ($2n=2x=24$ e $2n=3x=36$), *S. malmeanum* Bitter, que mais tarde foi renomeada para *S. commersonii* subsp. *malmeanum* ($2n=2x=24$), *S. calvescens* Bitter, encontrada estritamente em Caldas-MG em sua forma triploide ($2n=3x=36$) e *S. chacoense* Bitter, encontrada predominantemente nas regiões ao sul da América do Sul em sua forma diploide $2n=2x=24$ (BRÜCHER, 1999; SIMON et al 2010; GBIF, 2020).

A espécie triploide nativa do Brasil, *S. calvescens*, é uma das espécies potenciais para uso no melhoramento da batata. O cruzamento desta espécie triploide ($2n=3x=36$) com a batata cultivada (*S. tuberosum*; $2n=4x=48$) é viável, produzindo híbridos para retrocruzamento e

seleção de genótipos superiores. No entanto, a origem e status taxonômico de *S. calvescens* são controversos. Com base em aspectos morfológicos, alguns autores não consideram *S. calvescens* uma espécie, mas sim uma variante de *S. chacoense*, (BRÜCHER, 1999; SPOONER et al. 2014), ou de *Solanum malmeanum* (também denominada *Solanum commersonii* subsp. *malmeanum*) (HAWKES, 1978). Há ainda a possibilidade de que seja um híbrido entre os dois taxa, evento já identificado anteriormente no gênero por Pendinen et. al. (2008), ao estudar a natureza alopoliploide da espécie *S. stoloniferum*, tetraploide detentora de genoma tipo AABB. A possibilidade da alotriploidia também é sustentada pelos diversos relatos de híbridos triploides artificiais de espécies selvagens a partir de gametas $2n$, à exemplo de $(2x)$ *S. brevidens* x $(2x)$ *S. chacoense* (EHLENFELDT & HANNEMAN, 1984), $(2x)$ *S. circaefolium* x $(2x)$ *S. tuberosum* haploide (LOUWES et al., 1992), $(2x)$ *S. commersonii* x $(2x)$ *S. gourlayi* (MASUELLI et al., 1992), com a proposta de servirem de ponte para introgressão na batata cultivada.

A definição da circunscrição taxonômica e a origem de espécies selvagens constitui subsídio importante na busca de alelos desejáveis dentro do germoplasma selvagem e para o sucesso na introgressão dos mesmos nas espécies cultivadas (GAVRILENKO, 2011). Com o advento e evolução das técnicas de sequenciamento massivo em paralelo de DNA, houve um incremento acentuado dos estudos genômicos, dentre eles os estudos da fração repetitiva. O genoma dos eucariotos apresenta uma fração repetitiva que varia com relação à quantidade e tipo de sequências. Estas sequências repetitivas de DNA são fragmentos que se amplificam de diferentes formas se acumulando no genoma, com alta dinamicidade e plasticidade. São elementos que impactam a diferenciação entre espécies e que participam de processos importantes como função centromérica, telomérica e controle da expressão gênica (MEHROTRA; GOYAL, 2014 e OLIVEIRA, 2015).

Outro aspecto importante do estudo da fração de DNA repetitivo é a possibilidade de estudar essa fração do genoma nuclear para inferências filogenéticas a partir de sequenciamento de baixa cobertura (0,1 a 5%), uma estratégia denominada “genome skimming”. Essa estratégia é mais acessível do ponto de vista técnico e financeiro e permite o estudo de porções do genoma sem necessidade de sua montagem (DODSWORTH, 2015).

Portanto, este estudo visa descrever a fração repetitiva do DNA de *S. calvescens* de forma comparativa com espécies mais relacionadas *Solanum chacoense*, *Solanum commersonii* e *Solanum malmeanum* e com a espécie mais distante *Solanum tuberosum*, para prover novas

informações genéticas para subsidiar a definição do status taxonômico e a origem dessa espécie triploide.

2 REFERENCIAL TEÓRICO

2.1 Genômica de *Solanum*

O gênero *Solanum*, pertencente à família *Solanaceae*, contém mais de 2000 espécies catalogadas. A alta diversidade do gênero levou à sua divisão em seções, subseções, séries e superséries. Todas as 160 espécies que produzem tubérculos, das quais 20 são cultivadas, pertencem à seção Petota (HUNZIKER, 2001).

A seção Petota tem espécies representantes dos seis genomas distintos A, B, C, D, P e E, descritos por Matsubayashi (1991) a partir do estudo de híbridos interespecíficos do gênero. Destes citados, o genoma A é o mais comumente encontrado em várias espécies sul-americanas poliploides que compartilham este genoma. Já os genomas B, C, D, P e E têm essas alcunhas a partir das iniciais da respectiva série aos quais foram identificados pela primeira vez, sendo B da série *Longipedicellata* Busakov, C da série *Conicibaccata* Bitter, D da série *Demissa* Busakov, P da série *Piurana* Hawkes e E da série *Etuberosa*. Com o objetivo de padronizar as interpretações desses genomas, o mesmo autor reconheceu os cinco genomas A, B, C, D e P de espécies que tuberizam e o genoma E de espécies que não tuberizam.

Uma quantidade significativa de trabalho tem sido realizada para o conhecimento aprofundando do genoma da batata. Atualmente, os genomas de referência de batata publicamente disponíveis são do monoploide duplo *Solanum tuberosum* Grupo *phureja* DM1-36, do diploide selvagem *S. commersonii* e do clone diploide de *S. chacoense* - M6.

Em 2011, foi publicado o sequenciamento do genoma do monoploide $2n=2x=24$ (AA) de *S. tuberosum*, o que representou um marco para o aprofundamento dos estudos genômicos desse grupo. As informações detalhadas da constituição genômica da batata, com tamanho genômico estimado em 452,5 Mb, cerca de 39000 genes codificadores de proteínas, 62,2% de DNA repetitivo (sendo 29,4% deste do tipo LTR, longas repetições terminais), permitiram um maior entendimento e investigação da espécie e de grupos relacionados (PGSC, 2011).

Pham et. al (2020) construíram uma versão atualizada deste genoma, unindo as tecnologias mais atuais no sequenciamento genômico de nova geração a um custo menor. A

nova montagem representa 741,6 Mb (87,8%) do genoma estimado de 844 Mb, dos quais 741,5 Mb são contíguos, com excelente qualidade na construção dos 12 cromossomos.

A montagem do genoma de *S. commersonii* ($2n=2x=24$) foi realizada por Aversano et al. (2015). Comparado ao genoma da batata cultivada, o de *S. commersonii* apresentou um nível mais baixo de heterozigosidade, com a maioria dos SNPs identificados localizados em regiões intergênicas. Aproximadamente 44,5% (383Mb) do genoma total (830Mb) foi identificado como sequências repetitivas. Dessa proporção, a maior porção pertence aos retrotransposons de longas repetições terminais (LTRs), ocupando 34% do genoma.

Leisner et al. (2018) publicaram genoma de referência do clone M6 de *S. chacoense*, ancorando 508Mb de aproximadamente 825Mb estimados. Foi feita análise comparativa desse genoma com os genomas de *Capsicum annum* (pimentão) (Kim et al., 2014), *Nicotiana benthamiana* (tabaco) (Bombarely et al., 2012), *S. commersonii* (batata selvagem) (Aversano et al., 2015), *Solanum lycopersicum* (tomate) (The Tomato Genome Consortium, 2012), *Solanum melongena* (berinjela) (Hirakawa et al., 2014) e *S. tuberosum* Group *Phureja* DM1-3 (batata cultivada); Hardigan et al., 2016). Para as três espécies de batatas que tuberizam, o trabalho revelou um conjunto de 1.897 genes específicos a elas candidatos ao processo de tuberação. O genoma de *S. chacoense* M6 também apresentou alta heterozigosidade residual nos cromossomos 4, 8 e 9 em relação aos demais, ou seja, áreas de baixa densidade de genes e alta densidade de DNA repetitivo.

Outro esforço na investigação dos genomas de batata foi realizado por Kyriakidou et. al (2020), no qual os autores realizaram um esboço de montagem do genoma de seis espécies poliploides (variando de 3x a 5x), nativas da América do Norte: *Solanum chaucha* (3x), *S. juzepczukii* (3x), dois genomas de *S. tuberosum* subsp. *andigena* (4x), *S. tuberosum* subsp. *tuberosum* (4x) e *S. curtilobum* (5x). A autora destaca a alta heterozigosidade existente nos genomas. A propagação clonal da batata ao longo de milhares de anos limitou a recombinação genética e levou a estes altos níveis de heterozigosidade, com a provável contribuição da poliploidia e da autoincompatibilidade. Estimou-se que um dos genomas da batata tem cerca de 60,2% de sequências repetitivas, tornando sua montagem ainda mais difícil. Apesar do alto nível de fragmentação dos genomas finais, a autora reafirmou a necessidade e importância do sequenciamento e investigação constante dos genomas da batata como recursos valiosos no melhoramento genético da mesma.

Utilizando-se de um amplo germoplasma do gênero *Solanum* seção *Petota*, com representatividade de espécies cultivadas e selvagens de batata, Hardigan et. al. (2017) estudou

a variação genética existente nas espécies tuberosas e o impacto da domesticação na diversidade genômica do gênero através de SNPs de sequências conservadas. Os autores constataram extensivos eventos de hibridação introgressiva de alelos selvagens nas batatas cultivadas, sugerindo que a hibridação pode ter contribuído para a heterozigosidade e diversificação das mesmas em um período anterior ao gargalo da domesticação. Eles também ressaltaram a importância dos espécimes selvagens como uma rica fonte de alelos na melhoria da batata.

2.2 Espécies selvagens da Série *Commersoniana*

As espécies de batata silvestres que ocorrem no Brasil se encontram organizadas na seção Petota, que contém cerca de 100 espécies nativas e cultivadas da América do Sul. Segundo Brücher (1999), a série *Commersoniana* representa todas as espécies silvestres que tuberizam encontradas no Brasil, podendo também serem encontradas na Argentina, Bolívia, Paraguai e Uruguai, e abrange *S. commersonii*, *S. chacoense*, *S. calvescens* e *S. malmeanum*. Há certa divergência taxonômica para o grupo, pois alguns autores propõem nomenclatura distinta para duas das espécies. *S. malmeanum* (SPOONER, 2014) é encontrada na literatura também como *S. commersonii* subsp. *malmeanum* (KLASEN 2017, 2018; TOMÉ et al. 2009) e *Solanum calvescens* é defendida como citótipo de *S. chacoense* (*S. chacoense* subsp. *calvescens*) (SPOONER, 2014; BRUCHER, 1999; HIJMANS et al. 2007), porém é considerada por Hawkes (1978) como mais próxima morfologicamente de *S. commersonii malmeanum*.

Solanum calvescens, catalogada por Bitter em 1912, é uma espécie selvagem triploide ($2n=3x=36$) originária da região de Caldas-MG e é encontrada em solos ricos em matéria orgânica, próximos ao Rio Verde e a campos de cultivo, encontrada a 1200 m de altitude (HAWKES, 1978; BRÜCHER, 1999). Um levantamento do germoplasma da Embrapa Clima Temperado tem registros de coleta de espécies de batata selvagem e confirma a área de distribuição de *S. calvescens*, que possui apenas um acesso amostrado no Brasil, em Caldas MG, porém com potencial ocorrência em áreas não amostradas na Serra da Mantiqueira do Sul de MG e Sul do estado de SP (MOLIN, 2015a; MOLIN, 2016).

Molin et al. (2015b), estudando a distribuição geográfica das espécies silvestres de batatas brasileiras, verificou a existência de 10 registros de *S. calvescens*, todos em Minas Gerais, porém com áreas ainda sem a devida identificação. Os autores demonstraram

preocupação com a inexistência de estudos que avaliem o seu potencial no melhoramento em relação à presença de genes de resistência a estresses ambientais, pois a mesma espécie é a única que se distribui naturalmente em ambiente tropical brasileiro (Mata Atlântica), um indício de tolerância ao calor.

Atualmente *S. calvescens* se encontra em três conjuntos de dados online (GBIF 2021):

- Royal Botanic Gardens, Kew – Herbarium Specimens, que dispõe de um exemplar de exsicata da mesma.
- Instituto de Botânica Darwinion.
- Herbário da Universidade Federal de Minas Gerais (UFMG).

O status taxonômico e a origem dessa espécie são controversos. *S. calvescens* é similar a *S. commersonii*, especialmente da subespécie *malmeanum*, conhecida na sua forma diploide e triploide, sendo a forma triploide encontrada na Argentina e no Uruguai (HAWKES, 1978). Apesar da alta similaridade, características morfológicas permitiram a diferenciação entre *S. calvescens* e *S. commersonii malmeanum*.

Brücher (1999), ao comparar *S. calvescens* em seu local de origem, Caldas (MG), com *S. commersonii* ($2n=2x=24$) e *S. chacoense* ($2n=2x=24$) da mesma região, afirma que *S. calvescens* não deve ser denominada espécie. Segundo o autor, não existe diferença morfológica da antera entre *S. calvescens* e *S. chacoense* suficiente para distingui-las e que a primeira deve ser considerada um citótipo da segunda. Hijmans et. al. (2007) também consideram *S. calvescens* como uma variante triploide de *S. chacoense* e que o aumento da ploidia seria resultado de um comportamento expansivo no extremo do espectro da espécie.

Fernandes (2016) caracterizou os cromossomos com alta simetria cariotípica da espécie, com sete cromossomos metacêntricos e cinco submetacêntricos e com comprimento total do lote haploide (CTLH) de 18,1 μm . O autor evidencia que as tentativas anteriores de caracterização da taxonomia da espécie se baseiam apenas em características morfológicas, trazendo a necessidade de técnicas complementares citogenéticas e de caracterização genômica modernas para elucidar sua origem.

Genomas triploides e pentaploides podem ser resultado do cruzamento entre espécies diploides e tetraploides ou tetraploides e hexaploides, respectivamente. Pode ocorrer também a fertilização do óvulo por dois núcleos espermáticos do grão de pólen. Já nos alopolioides, cujo complemento cromossômico advém de dois ou mais genomas de espécies distintas, podem emergir naturalmente por duplicação cromossômica após o cruzamento interespecífico HAWKES (1978). Existe ainda a possibilidade de cruzamentos envolvendo gametas não

reduzidos, que possuem o número somático de cromossomos produzidos por meiose anormal (SCHIFINO-WITTMANN e DALL'AGNOL, 2001).

Oliveira et al (1995) constatou que híbridos provenientes do cruzamento entre a batata cultivada *S. tuberosum* e *S. chacoense* apresentam mecanismos na primeira e na segunda divisão meiótica que podem ocasionar formação de gametas não reduzidos.. Fusos paralelos ou fusos fundidos ocorrem na primeira divisão meiótica e promovem a restituição da primeira divisão meiótica, transmitindo cerca de 80% da heterozigosidade parental à progênie (HANNEMAN JUNIOR, 1999), enquanto a citociense prematura ocorre na segunda divisão causando a formação de micrósporos não reduzidos pela omissão da segunda divisão meiótica (MOK & PELOQUIN, 1975).

Carputo (2003), estudando o comportamento citogenético de híbridos de *S. commersonii* ($3x=36$) x *S. tuberosum* ($4x=48$), encontrou uma alta variação no número cromossômico, de 52 a 67, com predominância do pentaploide euploide $2n=5x=60$. O estudo também constatou que híbridos pentaploides podem produzir gametas aneuploides viáveis.

Tomé et al. (2009) relata ocorrência de cerca de 26,1% de formação de pólenes não reduzidos para em *Solanum commersonii* subsp. *commersonii* e 21% de pólenes não reduzidos *Solanum commersonii* subsp. *malmeanum*, em decorrência de fusos paralelos e fusos fundidos na primeira divisão meiótica.

A triploidia de *S. calvescens* ($2n=3x=36$) pode ter se originado a partir da união entre um gameta normal e um gameta não reduzido, tanto de cruzamentos intra quanto interespecífico. Pode também ser resultado da união entre gametas normais de uma espécie tetraploide e de uma diploide (FERNANDES, 2016). A possibilidade de alopoliploidia no gênero já foi estudada no gênero por Pendinen (2013), que estudou a série de espécies de batatas mexicanas *Demissa*, pertencente a seção *Petota*, a partir de uma hipótese feita por Matsubayashi (1991) de que todas as espécies desta série seriam alopoliploides. Os autores constataram marcas moleculares que permitiram corroborar com esta afirmativa, enriquecendo a aplicabilidade dos genótipos no melhoramento e propondo melhora na taxonomia da série.

De um modo geral, espécies selvagens de *Solanum* são comumente relacionadas a genes de importância agrônômica, pois frequentemente apresentam resistência a diversos fatores abióticos e doenças que atingem a batata cultivada *Solanum tuberosum* (DEJONG et al., 2001 e SPOONER et al., 2008) e a utilização de híbridos selvagens da batata como ponte para introgressão de alelos originados de germoplasma silvestre é uma prática já conhecida do melhoramento.

Cunha (1992) avaliou 81 clones obtidos do cruzamento entre dihaploides ($2n=2x=24$) de *S. tuberosum* x *S. chacoense* ($2n=2x=24$) e verificou que a produção de flores é maior nos indivíduos diploides em comparação com os indivíduos tetraploides. A autora verificou variação da produção de grãos de pólen viáveis entre as estações, com uma frequência menor em clima de primavera comparada ao clima de outono, porém sem prejudicar o retrocruzamento com *S. tuberosum* ($2x=4x=48$) visando a reconstituição do tetraploide.

Klasen et. al. (2017;2018) analisaram a cruzabilidade entre *S. tuberosum* e espécies da série Commersoniana visando ampliar a base genética do cultivo da batata. Os cruzamentos bem-sucedidos, ou seja, os que produziram frutos foram *S. chacoense* x *S. tuberosum*; *S. commersonii malmeanum* x *S. tuberosum*; *S. commersonii malmeanum* x *S. commersonii commersonii*; *S. chacoense* x *S. commersonii commersonii*; *S. commersonii malmeanum* x *S. chacoense*. O cruzamento de *S. calvescens* com *S. chacoense* foi o único com sucesso reprodutivo envolvendo *S. calvescens* quando essa foi usada como genitor feminino, uma vez que suas anteras eram estéreis.

Um estudo anterior conduzido pelos mesmos autores (KLASEN et. al., 2017), mais aprofundado nas relações entre as espécies silvestres brasileiras (*S. calvescens*, *S. chacoense* e *S. commersonii*), aponta o possível cruzamento entre *S. commersonii malmeanum* e *S. calvescens*, sendo a primeira a doadora dos pólenes, porém com a ressalva de que este cruzamento não gerou frutos maduros e, conseqüentemente, não produziu sementes.

Há diversos trabalhos que geraram híbridos triploides artificiais de espécies selvagens a partir de gametas $2n$, à exemplo de $(2x)$ *S. brevidens* x $(2x)$ *S. chacoense* (EHLLENFELDT & HANNEMAN, 1984), $(2x)$ *S. circaeifolium* x $(2x)$ *S. tuberosum* haploide (LOUWES et al., 1992), $(2x)$ *S. commersonii* x $(2x)$ *S. gourlayi* (MASUELLI et al., 1992). Um exemplo de estudo de resistência foliar à infecção ao patógeno *Phytophthora infestans*, avaliou o grau de esporulação em folhas de vários acessos de batata silvestre, entre elas se encontravam *S. calvescens* e *S. chacoense*. Cerca de 20% da superfície foliar de *S. calvescens* foi infestada, e *S. chacoense* apresentou alta resistência, com apenas alguns níveis de necrose foliar (CASA-COILA et. al., 2012). Simon et al. (2010) avaliaram a capacidade de resistência e reprodução de *S. chacoense*, analisando seu potencial como espécie invasora. Foram avaliados vários acessos de de diferentes locais do mundo, incluindo *S. calvescens* como espécie sinônima, e sua taxa de estabilidade a esses novos ambientes. Entre suas características de resistência principais, se encontram a habilidade de se reproduzir assexuadamente, resistência a insetos e a doenças. Bae et al. (2008) associa que a resistência à murcha de *Verticillium* em híbridos de batata

derivados de *S. chacoense* está ligada a um homólogo do gene de tomate (*Solanum lycopersicum* L.) *Ve.* Há também relatos de resistência ao nematoide do nódulo de raiz, *Meloidogyne incognita* e à sarna comum causada por escabiose de *Streptomyces* (HAWKES 1990).

2.3 DNA Repetitivo

As sequências de DNA repetitivo referem-se aos fragmentos de DNA que se apresentam amplificados, ou seja, multiplicados várias vezes nos genomas. Estes elementos podem ser responsáveis pela grande variação no tamanho dos genomas dos diversos organismos e, devido a sua alta dinamicidade e plasticidade, são elementos atuantes na diferenciação genômica entre espécies. Em vegetais superiores, a fração repetitiva pode constituir até 90% do genoma (MEHROTRA; GOYAL, 2014 e OLIVEIRA, 2015).

Sequências repetitivas foram consideradas “lixo” por muito tempo e descartadas de estudos genômicos. Apenas recentemente foram descobertas evidências de que elas desempenham diversas funções e participam de processos, como movimentação cromossômica, pareamento, recombinação, condensação dos centrômeros, coesão das cromátides irmãs, adesão dos cromossomos ao fuso mitótico, além de interações proteicas com a cromatina e alguns processos de regulação gênica em resposta aos estímulos ambientais (MEHROTRA; GOYAL, 2014).

O DNA repetitivo é dividido em dois grandes grupos, as sequências repetitivas dispersas, que são os elementos transponíveis capazes de criar cópias dispersas ao longo do genoma, e as sequências repetitivas em tandem, que representam blocos de repetições em tandem chamados de satélites (HESLOP-HARRISON; SCHMIDT, 2007).

2.3.1 Elementos Transponíveis (ETs)

As sequências repetitivas dispersas são fragmentos do DNA dos eucariotos que são capazes de se moverem e se amplificarem ao longo do genoma. São os componentes mais abundantes na maioria dos genomas, constituindo cerca de 45% no genoma humano, 52% do

genoma do gambá, podendo também atingir cerca de 85% de alguns grandes genomas de plantas, como o do milho (WICKER et al., 2007; FLORES e RAMOS, 2012).

Sucessivos estudos têm sido feitos sobre a fração repetitiva do DNA, evidenciando informações essenciais, a exemplo de Biémont e Vieira (2006), que demonstraram que elementos transponíveis são expressos em diferentes quantidades, a depender da fase da vida ou do tecido do organismo, uma característica excêntrica para uma parcela do DNA que era considerada “lixo” até bem pouco tempo. Os autores também trouxeram a reflexão de que a dinamicidade do DNA repetitivo nos organismos pode ter influenciado significativamente na evolução dos mesmos, podendo ocorrer até mesmo interações genótipo/ambiente, ou seja, com a expressão de seus elementos transponíveis sendo diferente entre ambientes para o mesmo táxon.

Nos modelos que evidenciam a dinâmica evolutiva dos elementos transponíveis, é prevista uma massiva amplificação dos mesmos a cada evento de transposição ocorrido, com consequente estabilização e equilíbrio definido pela força da seleção natural contra as inserções (BOUTIN et al., 2012). Porém, a taxa de transposição e amplificação de elementos transponíveis, principalmente retrotransposons, não é constante e pode variar entre diferentes organismos e entre as famílias de DNA repetitivo (HAWKINS et al., 2009). Desta forma, vários fatores podem interferir no ganho destas sequências repetitivas dispersas, relacionados com uma resposta genômico/ambiental, como hibridação, poliploidia ou estresse ambiental (GROVER, WENDEL, 2010; KASHKUSH, 2011; BOUTIN et al., 2012).

A multiplicação e inserção de cópias de elementos transponíveis incorporadas dentro de áreas gênicas codificantes podem alterar os padrões de expressão do gene ou até inativá-las (KASHKUSH; FELDMAN; LEVY, 2003), causar fissuras cromossômicas (DERAGON; CAPY, 2000), aumentar o tamanho do genoma hospedeiro, criar alterações na recombinação (FERDOROFF, 2000), desempenhar função regulatória através da transcrição de microRNAs, participar do pareamento e rearranjos cromossômicos, e responder a estímulos ambientais (MEHROTRA; GOYAL, 2014).

Expandindo a abordagem para as relações epigenéticas, as frações repetitivas dispersas se encontram, em sua maioria, em regiões inativadas por metilação (JURKA et al 2007), afirmativa corroborada por outros autores no estudo do milho, onde houve relação direta destes elementos à alta presença de metilações, ou seja, com estas sequências sendo encontradas em áreas hipermetiladas do genoma (SANMIGUEL, 1996; MEYERS, TINGEY, MORGANTE, 2001).

Os mecanismos de transposição que permitem esta dinamicidade classificam estas sequências em: a) Elementos Transponíveis de Classe I, ou retrotransposons, que são sintetizados através da enzima transcriptase reversa, incluindo, obrigatoriamente, uma sequência de RNA como intermediário; e b) Elementos Transponíveis de Classe II, ou transposons, elementos que realizam sua transposição através da excisão de um fragmento do local de origem e consequente inserção em outro, sem deixar cópias no local de origem (WICKER et al., 2007).

Os retrotransposons são os responsáveis pela maior porção dos elementos transponíveis presentes nos genomas, pela sua capacidade de deixar cópias a cada transposição. São divididos em cinco ordens: LTR (Long Terminal Repeat), DIRS (*Dictyostelium Intermediate Repeat Sequence*), Penelope-Like Elements (*PLEs*), LINEs (*Long Interspersed Nuclear Elements*) e por fim os SINEs (*Short Interspersed Nuclear Elements*) (WICKER et al., 2007).

Os retrotransposons LTR contêm dois genes principais, *gag* e *pol*, no qual o primeiro codifica uma poliproteína estrutural e o segundo codifica proteínas catalíticas necessárias à replicação, como transcriptase reversa, RNase e integrase. As famílias Ty1/copia e Ty3/gypsy apresentam ampla frequência e diversificação, sendo as mais importantes dos genomas eucarióticos (WICKER et al., 2007). Esta abundância frequentemente ocorre com a representação de poucas famílias compondo a maior parte da fração repetitiva (WICKER et al., 2007).

Os elementos transponíveis da Classe II, ou transposons de DNA, tem sua excisão e inserção comandada pela transposase. São divididos em Subclasse I, cujos transposons são flanqueados por longas repetições terminais (TIRs – Terminal INverted Repeats) que permitem o corte da dupla fita de DNA e Subclasse II, cujos transposons envolvem o corte de apenas uma das fitas de DNA (JURKA et al., 2007; WICKER et al., 2007)..

2.3.2 DNA Satélites

As sequências de DNA repetitivo que se organizam em tandem são chamadas de DNA satélite. São monômeros arranjados sequencialmente e que podem ser classificados conforme seu comprimento, propriedades e localização no genoma. Podem representar até 20% do DNA nuclear de uma planta, cerca de 50% de alguns genomas de insetos e roedores e menos de 5%

em seres humanos (FLORES e RAMOS, 2012). São divididos em três principais grupos: DNA microssatélite, DNA minissatélite e DNA satélite (GARRIDO-RAMOS, 2017).

Chamados também de *Simple Sequence Repeats* (SSRs), os microssatélites variam de 1 a 5 pb de comprimento, enquanto os minissatélites de 10 a 50 pb e, por fim, as sequências satélites podem possuir monômeros de 150 a 400 pb repetidos em centenas ou milhares de cópias ao longo do genoma (KUBIS; SCHMIDT; HESLOP-HARRISON, 1998; GARRIDO-RAMOS, 2017).

Os satélites de DNA são elementos que possuem comprimento variável e que estão constantemente associados a regiões de heterocromatina, como os centrômeros e subtelômeros (MEHROTRA; GOYAL, 2014). Devido a essa localização, tais elementos são frequentemente ligados a funções de proteção e estabilização da estrutura cromossômica (KUBIS; SCHMIDT; HESLOP-HARRISON, 1998).

Mesmo que muitas famílias de satélite já descritas na literatura sejam sequências específicas para determinados taxons, alguns tipos de sequência satélite são conservadas, desempenhando mesmas funções entre organismos que diferem a nível de reino, a exemplo do DNA ribossômico e algumas sequências teloméricas (HESLOP-HARRISON, 2000). Essas características, associadas à rápida evolução desse tipo de sequência, faz com que a mesma possa ser amplamente utilizada em estudos taxonômicos e filogenéticos, principalmente em espécies vegetais, nas quais esses elementos são abundantes (MEHROTRA; GOYAL, 2014).

Assim, dados genômicos obtidos via sequenciamento e localização dos elementos repetitivos em cromossomos vegetais são importantes fontes de informação acerca da composição e história evolutiva dos genomas destas espécies, contribuindo ainda para estudo mais aprofundados de cariotipagem e marcas cromossômicas mais informativas (MACAS; NEUMANN; NAVRÁTILOVÁ, 2007; MACAS et al., 2011; HEITKAM et al., 2020).

2.3.3 DNA repetitivo em *Solanum*

Várias sequências repetitivas têm sido descritas em *Solanum*, especialmente em associação com regiões centroméricas e teloméricas. Torres et al (2011) identificaram duas sequências satélites subteloméricas em *S. tuberosum*, CL14 e CL34. O estudo comparativo com outras espécies revelou a presença dessas sequências em espécies mais relacionadas e a origem mais recente de CL34.

Outro elemento repetitivo, nomeado de pSbTC1, foi descrito por Tek e Jiang (2004) na espécie diploide *Solanum bulbocastanum* Dunal. Identificado como uma repetição telomérica intersticial (ITR), foi localizado no centrômero de várias espécies de *Solanum*. Nas espécies com genomas B e P foi observada uma ampliação massiva de pSbTC1, evidenciando um processo de invasão do centrômero funcional.

A análise molecular de um DNA satélite denominado *Sobo*, detectado apenas em uma espécie diploide de batata (*Solanum bulbocastanum* Dunal.) por Tek et al. (2005) revelou similaridade parcial com sequências LTR, indicando que esse satélite pode ter se originado a partir de rearranjos entre dois LTRs iniciais.

A melhor compreensão da complexidade centromérica existente em *Solanum tuberosum* foi possível a partir de uma série de estudos e investigações citogenéticas e moleculares. Gong et al. (2012) encontraram características peculiares na espécie, onde, dos doze cromossomos básicos, apenas sete apresentaram sequências repetitivas, apresentando arranjos repetitivos específicos para cada cromossomo. Zhang et al. (2014), estudando uma espécie selvagem relacionada, *Solanum verrucosum*, proposta como ancestral da primeira, também relataram centrômeros compostos por DNA repetitivo e centrômeros compostos por poucas cópias. Como ancestral, era esperado que houvesse certa homologia entre estes satélites com *S. tuberosum*, o que não foi verificado. A hipótese dos autores é de que dois centrômeros em *S. verrucosum* sofreram invasão de sequências repetitivas após a diferenciação que originou *S. tuberosum*.

Oliveira (2015) analisou sequências repetitivas centroméricas para a espécie diploide *Solanum chomatophilum*, onde foram identificadas três sequências satélite (Sc83, Sc92/108 Sc111) que apresentaram divergências entre os cromossomos homólogos e variação em distribuição entre os dois acessos da espécie, assim como entre plantas de um mesmo acesso, evidências da alta dinamicidade das sequências repetitivas centroméricas.

Estudando o padrão e organização de sequências centroméricas características dos genomas A e P, Oliveira (2017) utilizou a técnica de FISH a fim de observar o comportamento destes tipos de genomas em espécies aloploiploides (que portam genomas AC e AD). De modo geral, as sequências centroméricas identificadas nestas espécies pertencem ao genoma A, com apenas uma exceção para o genoma D, evidenciando a origem comum destas sequências.

Um estudo de comparação entre as frações repetitivas do genoma de espécies silvestres e cultivadas foi realizado por Gaiero et al. (2019), sendo seis espécies de batata e sete espécies de tomate. As autoras encontraram predominância de elementos repetitivos dispersos, ou elementos transponíveis, do tipo LTR (*Long Terminal Reads*). A família mais abundante foi

Ty3/gypsy para todas as espécies, com diferenças significativas em suas quantidades e arranjos sequenciais, os quais se organizaram em grandes aglomerados nas espécies de batata, em comparação aos arranjos organizados mais uniformemente no genoma das espécies de tomate. Houve também uma similaridade maior nas proporções de DNA repetitivo nas espécies de batata que nas proporções do clado do tomate, um indicativo de equilíbrio alcançado após a diferenciação há cerca de sete milhões de anos atrás. Já entre as espécies de batata, foi relatado diferença contrastante na proporção de sequências satélites, muito mais abundantes nas espécies cultivadas quando comparadas às espécies selvagens.

3 MATERIAL E MÉTODOS

3.1 Material Genético

Para a espécie *Solanum calvescens* ($2n=3x=36$) foi utilizado o acesso coletado na região de Caldas-MG, que se encontra depositado no Banco de Germoplasma de Batata do Departamento de Biologia, setor de Genética e Melhoramento de Plantas da Universidade Federal de Lavras, Lavras-MG. Os tubérculos foram plantados e multiplicados em casa de vegetação. As análises foram feitas no Laboratório de Citogenética Vegetal, Departamento de Biologia da Universidade Federal de Lavras.

O DNA da espécie *Solanum calvescens* foi obtido a partir de folhas jovens de plantas cultivadas em casa de vegetação com uso do Kit de Extração DNeasy® Plant Mini Kit (Qiagen® Inc., Venlo, Netherlands). A qualidade e a concentração do DNA foram avaliadas por meio de Eletroforese em gel de Agarose 1% e em espectrofotômetro NanoDrop™.

O sequenciamento genômico de baixa cobertura (1x) foi feito em plataforma Illumina HiSeq™ 4000, em parceria com o Instituto René Rachou – Fiocruz Minas, em Belo Horizonte – MG. Para tanto, foi feita a montagem da biblioteca com o kit Nextera™ DNA Flex Library Prep com fragmentos Pair-ended de 2x300pb (Illumina™).

As sequências das demais espécies da Série Commersoniana, todas diploides com $2n=2x=24$ (*Solanum chacoense*, *Solanum commersonii*, *Solanum malmeanum*), foram compartilhadas pelo Laboratorio de Evolución y Domesticación de las Plantas do Departamento de Biología Vegetal da Facultad de Agronomía de Montevideo, Uruguai (GAIERO et al, 2019). *S. commersonii* é um genótipo haploide obtido por cultivo de anteras (DEL CASTILLO et al.

2016) a partir do acesso 04.02.3 coletado em Colônia, sudeste do Uruguai em 2004. *S. chacoense* é um genótipo diploide do acesso 07.01.7 coletado em Salto, noroeste de Uruguai em 2007. *S. malmeanum* é um genótipo diploide do acesso R10P1 coletado em Rivera, noroeste de Uruguai em 2017.

As bibliotecas para as três espécies foram preparadas usando o Nextera Library Preparation Kit (Illumina). *S. commersonii* foi sequenciada em Illumina HiSeq2000 na *Applied Bioinformatics, Wageningen University Research*. *S. chacoense* e *S. malmeanum* em Illumina HiSeq4000 no *Beijing Genomics Institute* (BGI) em Hong Kong, sendo essa última com cobertura de aproximadamente 1x (1 Gb de sequência), e sequências de 800 bp.

A espécie *S. etuberosum* ($2n=2x=24$), pertencente à Seção Etuberosum, foi usada como grupo externo para as análises. Os dados de sequenciamento da espécie *S. etuberosum* foram obtidos a partir de sequenciamento de alta cobertura, via Illumina PE HiSeq550bp, compartilhados pelo Biosystematics Group (BERKE, 2016) e subamostrados a 1x para este estudo.

As espécies de *Solanum* selecionadas para o presente estudo e suas principais informações genômicas estão relacionadas na Tabela 1.

Tabela 1 – Espécies de *Solanum* estudadas e suas respectivas classificação, tamanho do genoma, origem dos dados de sequenciamento e código utilizado na plataforma Repeat Explorer (RE).

Classificação	Ploidia	Espécies	Código (RE)	C-values (1C, Mpb)	Origem do Sequenciamento
Seção Petota	3x	<i>Solanum calvescens</i>	<i>Scalv</i>	1470,7*	Dados do estudo
	2x	<i>Solanum chacoense</i>	<i>Schac</i>	617**	Gaiero et al, não publicado
	1x	<i>Solanum commersonii</i>	<i>Scomm</i>	792**	Gaiero et al, não publicado
	2x	<i>Solanum malmeanum</i>	<i>Smalm</i>	811,74	Gaiero et al, não publicado
Seção Etuberosum	2x	<i>Solanum etuberosum</i>	<i>Setub</i>	763**	<i>Biosystematics Group, WUR</i>

*Tamanho do genoma determinado neste estudo.

**Valores conforme Gaiero et al (2019)

Fonte: Do autor (2021)

3.2 Citometria de Fluxo

Tecido foliar jovem de três amostras de acessos de *S. calvescens* e uma amostra da planta padrão de referência interno (*Pisum sativum* L., DNA nuclear = 9,09pg) foram usadas para estimar o conteúdo de DNA nuclear. Os tecidos vegetais foram triturados com o auxílio de um bisturi, em placa de Petri, contendo 1 mL de tampão de extração de núcleos LB01 gelado, para obtenção da suspensão nuclear. Em seguida, a suspensão foi filtrada em malha de 50µm e corada com 25µL de solução 1mg/mL do fluorocromo iodeto de propídeo (Sigma®).

As leituras foram realizadas no citômetro de fluxo FACSCalibur 4 cores (Becton Dickinson – BD) e os histogramas obtidos pelo software CellQuest™ foram analisados no software WinMDI™. A estimativa da quantidade de DNA foi realizada pela posição relativa do pico G1 dos acessos em relação à posição do pico G1 do padrão de referência.

Os cálculos foram realizados de acordo com a equação $Q = E/S \times R$, no qual “Q” é a quantidade de DNA nuclear da amostra, “E” é a posição do pico G1 da amostra, “S” é a posição do pico G1 do padrão interno de referência e “R” é a quantidade de DNA nuclear do padrão. (FERREIRA, 2014).

Os parâmetros Valor 2C, C e Cx foram definidos conforme Greilhuber et al (2005). O valor 2C corresponde ao tamanho do genoma holoploide, ou seja, todo o conteúdo de DNA nuclear, o valor C corresponde ao tamanho do genoma monoploide (2C/2) e Cx corresponde ao tamanho do genoma em relação ao número básico x em poliploides (2C/x).

3.3 Análise da fração repetitiva através da *pipeline* Repeat Explorer

3.3.1 Pré-processamento

Os arquivos de sequenciamento digitais das espécies foram inseridos na plataforma online do Repeat Explorer (RE), *Galaxy RepeatExplorer portal* (NOVÁK et al. 2013) onde se deram todas as etapas de pré-processamento.

Os dados foram submetidos à análise da qualidade das sequências (*reads*) pela ferramenta FASTQC *Quality Reports* (BROWN; PIRRUNG; MCCUE, 2017), seguida de

filtragem pela ferramenta *Filter by quality* nos valores de referência padrão de *Score Phred* 20 em pelo menos 90% das bases.

A partir da inspeção da qualidade dos reads foi determinado o corte das sequências pela ferramenta *Trim sequences* a fim de retirar as regiões de baixa confiabilidade e garantir a qualidade da análise. O comprimento mínimo das sequências após o corte resultou em 80pb para a espécie *S. etuberosum*, 86pb para *S. commersonii*, 110pb para *S. chacoense*, 125pb para *S. malmeanum*, e 175pb para *S. calvescens*.

As sequências *Forward* (F) e *Reverse* (R) foram unidas em um só conjunto de dados através da ferramenta *Interlacer Paired Reads* para cada espécie, finalizando o pré-processamento.

3.3.2 Quantificação do DNA repetitivo

Para determinar a quantidade e descrever os elementos transponíveis nos genomas, foi realizada análise comparativa com cobertura de cerca de 0,1x do genoma para todas as espécies a partir de amostragem realizada pelo RE (TABELA 2).

Tabela 2 – Quantidade de *reads* amostrados para análise de elementos transponíveis (comparativa) e de satélites (TAREAN) em espécies de *Solanum* L..

Ploidia	Espécie	Análise Comparativa	Análise TAREAN
3x	<i>S. calvescens</i>	489.540	1.242.086
2x	<i>S. chacoense</i>	489.119	1.400.772
1x	<i>S. commersonii</i>	1.606.175	1.606.390
2x	<i>S. malmeanum</i>	566.265	1.906.004
2x	<i>S. etuberosum</i>	831.700	9.537.850
Total:		3.982.799	

Fonte: Do autor (2021)

O algoritmo computacional para análise do DNA repetitivo da plataforma RE permite o agrupamento e identificação dos *clusters* de maior similaridade entre si de forma automática e disponibiliza dados preliminares dos elementos repetitivos encontrados. Desta forma, a classificação dos elementos transponíveis presentes na espécie foi feita de forma automática, através do algoritmo *Repeat Explorer 2 Clustering*, com todas as sequências simultaneamente.

Para os agrupamentos não identificados, foi feita a identificação manual das sequências agrupadas pelos clusters anônimos através de comparação contra o banco de dados *Giri Repbase* (BAO; KOJIMA; KOHANY, 2015) e classificados conforme similaridade com as sequências disponíveis.

Para avaliar a fração dos satélites, foram utilizadas coberturas genômicas que variaram entre 0,15 e 1x (Tabela 2). Os satélites foram identificados pelo algoritmo TAREAN (NOVÁK et al., 2017) e confrontadas contra si mesmas através da plataforma Dotmatcher - EMBOSS (LI et al., 2015) para confirmação dos arranjos via análise gráfica. Para verificar a identidade e similaridade entre as sequências satélites identificadas, foi utilizado o alinhamento local múltiplo de sequências através do algoritmo MAFFT (KATOY; STANDLEY, 2013) dentro do software Geneious Prime 2019 (KEARSE et al., 2012). As matrizes de identidade foram avaliadas conforme proposto por Ruiz-Ruano et al. (2016), onde similaridade maior que 95% indica a mesma variante, similaridade entre 80% e 95% indica diferentes variantes de uma mesma família e abaixo de 80% indicam sequências distintas com proximidade ao nível de superfamília.

Para encontrar similaridades dos satélites encontrados com a literatura disponível, foi feito o BLAST (configuração megablast) das sequências consenso identificadas com o banco de dados do NCBI.

A nomenclatura adotada para identificação de cada satélite se deu pelo código de cinco letras utilizado para representar cada espécie, seguido do sufixo “Sat” e posterior número para indicar cada superfamília. Os satélites foram organizados por abundância decrescente de quantidade de *reads*.

A abundância das sequências satélite nos genomas de todas as espécies foi estimada pela ferramenta Map to Reference, no programa Geneious Prime 2019. O gráfico *Bubble Chart* para valores de abundância foi feito via script do R para o pacote GGplot (RSTUDIO, 2021).

4 RESULTADOS

4.1 Tamanho no Genoma de *Solanum calvescens*

Os picos G1 das leituras do padrão de referência e das amostras de *S. calvescens* apresentaram média de 121,5 e 40,2, respectivamente. Pelo cálculo da equação $Q = 40,2/121,5 \times 9,09$, o valor estimado para a quantidade de DNA nuclear total (2C) é 3,0076pg. Portanto, o valor do lote haploide (Valor C) para a espécie é 1,5038pg e 1470,69Mb (1pg = 978Mb). Considerando se tratar de uma espécie triploide, o valor Cx (2C/x) é de 1,003pg e 980,48Mb para o tamanho do genoma monoploide da espécie.

4.2 Quantificação dos Elementos Transponíveis

Do total de 3.982.796 sequências analisadas, 2.753.350 *reads* foram organizados em 296.442 *clusters*, enquanto 1.229.446 *reads* foram classificados como sequências não repetitivas. Destes clusters apresentados pela análise, 391 foram identificados pelo algoritmo do *Repeat Explorer* como mais representativos (mais que 0,01% de abundância). As sequências representadas por genoma cloroplastidial e mitocondrial foram desconsideradas.

A espécie com maior porcentagem de DNA repetitivo foi *S. chacoense* com 45,37%, e a menor foi *S. calvescens*, com 38,37%, a única espécie triploide. As espécies apresentaram padrão semelhante de abundância das diferentes categorias, à exceção das sequências ribossomais e sequências satélites (Tabela 3, Figuras 1 e 2).

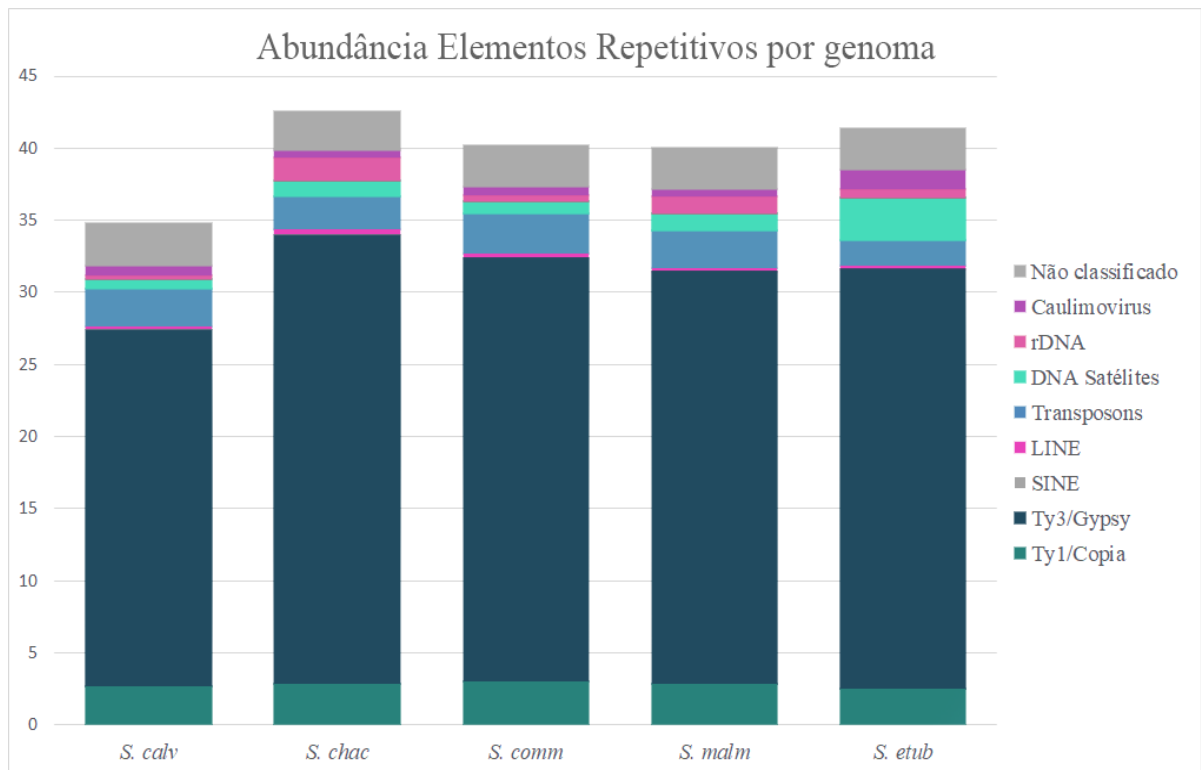
Em todas as espécies, a superfamília de retrotransposons *Ty3/gypsy* foi a mais abundante, variando de 24,72% a 31,15% de abundância, com destaque para a subfamília *Tekay*, variando entre 16,02% a 21,79% de abundância, seguida da subfamília *Athila*, variando entre 4,28% a 6,71% (Tabela 3, Figuras 1 e 2). A segunda superfamília mais abundante foi *Ty1/copia*, variando entre 2,52% a 3,06%, da qual a família mais abundante é *TAR*. Entre os transposons, o mais abundante pertence à Subclasse I e famílias *hAT* e *EnSpm-CACTA* (TABELA 3, FIGURAS 1 e 2).

Tabela 3 - Composição da fração repetitiva do genoma de *Solanum calvescens*, *Solanum chacoense*, *Solanum commersonii*, *Solanum malmeanum* e *Solanum tuberosum*.

		Classificação		<i>S. calv.</i>	<i>S. chac</i>	<i>S. comm</i>	<i>S. malm</i>	<i>S. etub</i>
Classe I	LTR	Ty1/Copia	Ale	0.306	0.344	0.361	0.314	0.267
			Angela	0.227	0.250	0.195	0.212	0.227
			Bianca	0.361	0.389	0.409	0.336	0.377
			Ikeros	0.117	0.122	0.133	0.114	0.117
			Ivana	0.152	0.148	0.173	0.161	0.124
			SIRE	0.153	0.202	0.223	0.209	0.161
			TAR	1.058	1.018	1.140	1.117	0.949
			Tork	0.154	0.200	0.219	0.194	0.173
			Não classificado	0.179	0.165	0.206	0.179	0.128
			Total Ty1/Copia	2.704	2.837	3.060	2.836	2.523
	Ty3/Gypsy	Chromovirus	CRM	0.251	0.319	0.250	0.306	0.074
			Galadriel	0.093	0.053	0.043	0.047	0.016
			Tekay	16.017	19.841	18.900	17.975	21.790
		Non-chromovirus	Athila	4.734	6.714	5.926	6.158	4.276
			Ogre	1.110	1.293	1.250	1.234	0.608
		Não classificado	2.513	2.928	2.932	2.855	2.333	
	Total Ty3/Gypsy	24.718	31.148	29.301	28.576	29.098		
	Caulimovirus	0.639	0.549	0.600	0.567	1.296		
	Não classificado	1.402	1.281	1.522	1.286	0.396		
Total LTR	29.463	35.814	34.483	33.266	33.314			
LINE	0.196	0.326	0.313	0.269	0.202			
SINE	0.022	0.031	0.033	0.035	0.017			
Pararetrovirus	1.939	1.396	1.493	1.346	1.893			
Não classificado	0.228	0.141	0.167	0.139	1.325			
Total Classe I	31.848	37.706	36.488	35.054	36.751			
Classe II	Subclasse I	TIR	CACTA	0.552	0.545	0.632	0.635	0.315
			hAT	0.730	0.655	0.758	0.728	0.390
			Mariner	0.482	0.407	0.505	0.420	0.476
			Mutator	0.306	0.268	0.322	0.316	0.206
			Harbinger	0.321	0.228	0.297	0.267	0.284
	Subclasse II	Helitron	0.061	0.066	0.082	0.073	0.005	
Não classificado	0.161	0.126	0.175	0.158	0.113			
Total Classe II	2.614	2.294	2.770	2.598	1.790			
Tandem Repeats	rDNA	45s	0.226	1.371	0.271	0.899	0.230	
		5s	0.030	0.180	0.120	0.239	0.425	
	Satélites	0.668	1.140	0.833	1.181	2.920		
	Total Tandem	0.923	2.691	1.224	2.319	3.574		
Total classificado	35.385	42.691	40.482	39.970	42.115			
Não classificado	2.981	2.678	2.913	2.816	2.919			
TOTAL	38.366	45.369	43.395	42.786	45.034			

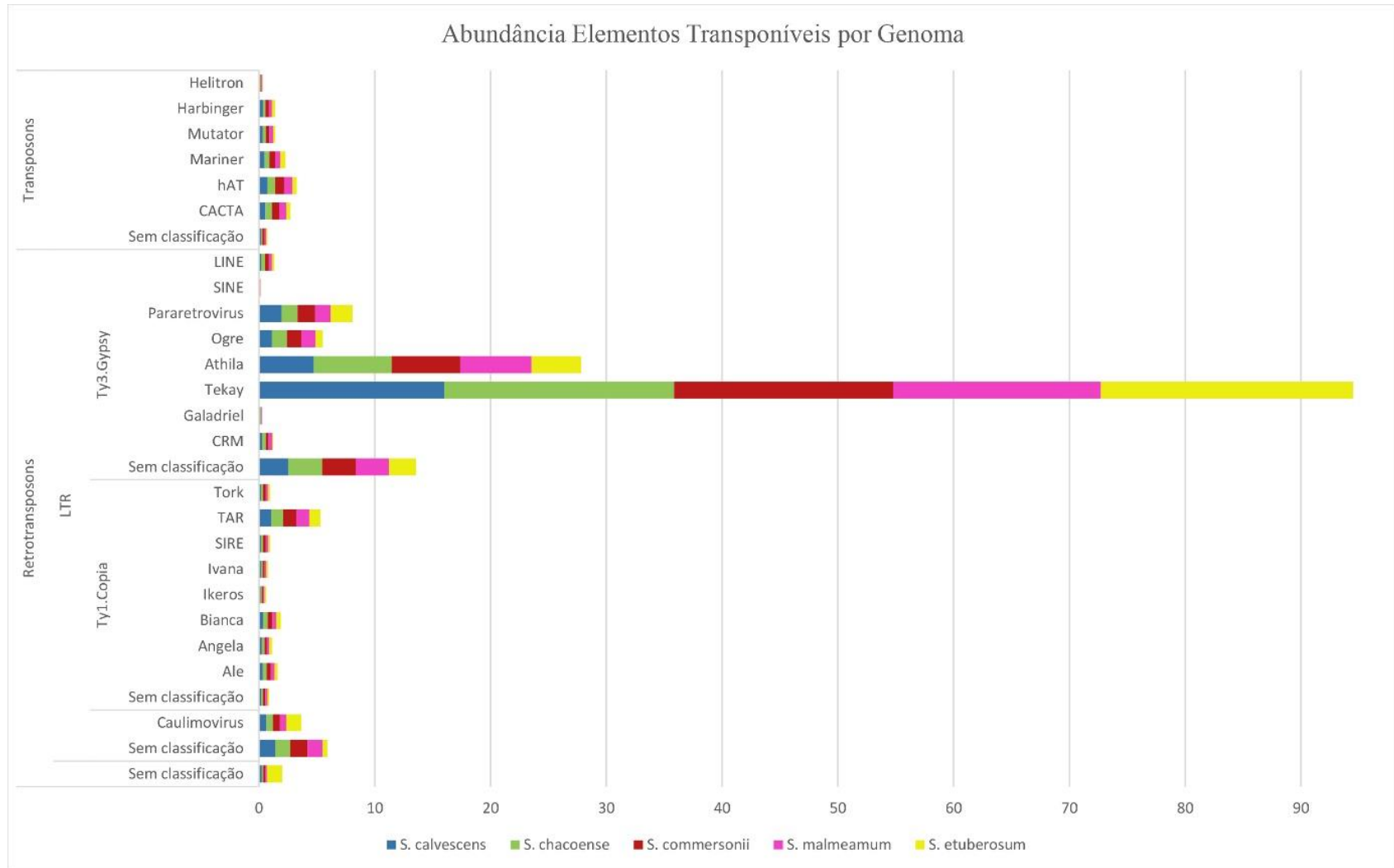
Fonte: Do autor (2021)

Figura 1 - Proporção dos principais elementos repetitivos presentes no genoma de *Solanum calvescens*, *Solanum chacoense*, *Solanum commersonii*, *Solanum malmeanum* e *Solanum etuberosum*



Fonte: Do autor (2021)

Figura 2 - Proporção da abundância dos elementos transponíveis no genoma de *Solanum calvescens*, *Solanum chacoense*, *Solanum commersonii*, *Solanum malmeanum* e *Solanum etuberosum*.



Fonte: Do autor (2021)

4.3 Sequências em Tandem

Inicialmente, o relatório da análise do *TAREAN* apontou 37 possíveis sequências satélite, sendo nove para *S. calvescens*, sete para *S. chacoense*, cinco para *S. commersonii*, seis para *S. malmeanum* e dez para *S. etuberosum*. Destas, 10 sequências satélites foram confirmadas: três para *S. calvescens*, duas para *S. chacoense*, uma para *S. commersonii* e quatro para *S. etuberosum*. O alinhamento via algoritmo MAFFT revelou uma identidade inferior a 80% para todos os satélites. A maior similaridade observada foi de 45,06% entre um satélite de *S. etuberosum* (SetubSat5) e um de *S. commersonii* (ScommSat2), e a menor similaridade de 2,96% entre dois satélites de *S. etuberosum* (SetubSat3 e SetubSat6). Todos os satélites se organizam ao nível de superfamílias (TABELA 4).

Tabela 4 - Classificação em superfamília (SF), comprimento da sequência consenso (Comp. pb) e quantidades de *reads* das sequências de DNA satélite identificadas nos genomas de *Solanum calvescens*, *Solanum chacoense*, *Solanum commersonii*, *Solanum malmeanum* e *Solanum etuberosum*.

Abundância Satélite por Genoma (reads)							
SF	Satélite	Comp. (pb)	<i>S. calv</i>	<i>S. chac</i>	<i>S. comm</i>	<i>S. malm</i>	<i>S. etub</i>
SF1	SetubSat1	181	1831	2683	1128	2567	49922
SF2	ScommSat2	182	2507	3013	1432	1866	28541
SF3	SetubSat3	164	2655	272	313	440	26380
SF4	ScalvSat4	338	19402	4526	9134	9469	55
SF5	SetubSat5	179	1537	794	749	1221	9046
SF6	SetubSat6	2502	1338	621	8696	786	5456
SF7	ScalvSat7	663	1168	533	567	747	1174
SF8	ScalvSat8	525	2117	1	1111	3498	36
SF9	SchaSat9	2508	10	4074	7	46	0
SF10	SchaSat10	177	0	757	0	0	0

Fonte: Do autor (2021)

Com o objetivo de comparar a abundância de cada sequência satélite identificado contra os genomas das espécies, utilizamos o algoritmo de Mapeamento por Referência, o qual permitiu quantificar os monômeros extraídos em *reads* e demonstrar a abundância destas sequências.

Todas as espécies apresentaram ocorrências dos DNA satélites identificados, à exceção dos satélites pertencentes a *S. chacoense*. Um destaque interessante se encontra no satélite identificado em *S. calvescens*, ScalvSat8, que apresentou apenas uma ocorrência na espécie *S. chacoense*. Dentre os satélites originalmente encontrados para a espécie triploide nativa *S. calvescens*, ScalvSat4 se destaca pela presença significativa em todas as espécies mais relacionadas e ausência na *S. etuberosum*. *S. chacoense* apresentou dois satélites praticamente exclusivos, SchacSat9 e SchacSat10. O satélite ScalvSat8 não dispõe de representação significativa no genoma de *S. chacoense*. *S. etuberosum* apresentou a maior abundância relativa em sequências satélites (FIGURA 3, TABELA 4).

As buscas por similaridade aos satélites via BLAST – NCBI revelaram correspondências para todas as sequências satélite encontradas. Para SchacSat10, exclusiva de *S. chacoense*, foram encontradas 15 correspondências com sequências de genomas de *Trypanossoma* (protozoário flagelado), sendo algumas delas com similaridade acima de 95%, ou seja, de mesma variante.

Todas as outras sequências obtiveram altas correspondências para três híbridos da batata cultivada *S. tuberosum*, denominados MSH/14-112, P8 e Solyntus™. Também houveram correspondências para as espécies *Solanum spegazzinii* (batata selvagem diploide), *Solanum lycopersicum* (tomate cultivado), *Solanum pinnatisectum* (batata selvagem diploide), *Solanum demissum* (batata selvagem hexaploide), *S. pennellii* (tomate selvagem) e *S. brevidens* (batata selvagem), em diferentes escalas. O satélite ScalvSat8 apresentou correspondência apenas com híbridos de *S. tuberosum* mencionados acima.

A maioria das correspondências se organiza em nível de família (de 80 a 95% de similaridade) e superfamília (acima de 95%), com maiores representatividades para a batata cultivada, o tomate cultivado e para a espécie selvagem *S. pinnatisectum*.

Figura 3 - Abundância de *reads* das seqüências satélite identificadas em *S. etuberosum* (Setub), *S. malmeanum* (Smalm), *S. commersonii* (Scomm), *S. chacoense* (Schac), *S. calvescens* (Scalv) relativos ao tamanho do genoma de cada espécie



Fonte: Do autor (2021)

5 DISCUSSÃO

O genoma de *Solanum calvescens* foi, em média, duas vezes maior que o genoma das outras espécies, o que pode ser explicado pela sua condição triploide $2n=3x=36$. Mesmo considerando o valor de uma cópia do genoma, Cx , o genoma de *S. calvescens* foi, em média 1,3 vezes maior, sendo mais próximo dos genomas de *Solanum commersonii*, *Solanum malmeanum* e de *Solanum etuberosum* e mais distante do de *Solanum chacoense*. Geralmente, a relação entre tamanho do genoma é ligada à poliploidização (BENNETT; LEITCH, 2005) e à dinâmica de amplificação dos elementos repetitivos, já que os mesmos incidem sobre a quantidade de cópias presentes nos genomas. Em *S. calvescens* a poliploidização parece ser mais relevante, uma vez que a abundância das sequências repetitivas foi menor que a das outras espécies.

A taxa de transposição e amplificação dos elementos transponíveis pode variar entre diferentes organismos e famílias de DNA repetitivo (HAWKINS et al., 2009). A menor abundância de elementos transponíveis em *S. calvescens* pode estar relacionada com a resposta diferenciada dos elementos transponíveis ao ambiente, já que essa espécie está submetida a condições ambientais distintas, e aos eventos envolvidos na poliploidização (GROVER, WENDEL, 2010; KASHKUSH, 2011; BOUTIN et al., 2012).

Genomas vegetais podem ser constituídos de até 90% por sequências repetitivas (MEHROTRA; GOYAL, 2014 e OLIVEIRA, 2015). As variações encontradas em nosso estudo, de aproximadamente 38% a 45% do genoma das espécies, são consideravelmente menores. A porcentagem de 43,4% encontrada para *S. commersonii* corrobora a descrita por Aversano et al. (2015) que relataram 44,5% (383Mb) para o complemento repetitivo, sendo a maior fração repetitiva dominada pelos retrotransposons das sequências terminais longas (LTR) na proporção 34%, valor também encontrado neste estudo (34,5%).

O padrão da abundância relativa dos elementos repetitivos demonstrados na Figura 2 apresenta-se semelhante entre as espécies, à exceção das sequências ribossomais e satélites. Os valores totais da fração repetitiva das espécies *S. chacoense* (45,369%), *S. commersonii* (43,395%) e *S. etuberosum* (45,034%) foram maiores que os encontrados por Gaiero et al (2019), com 29,52%, 27,99% e 28,23% respectivamente, mas com abundâncias relativas para as categorias de elementos repetitivos similares. Consideramos que esta diferença encontrada pelas autoras ocorreu em virtude da maior quantidade de espécies estudadas por elas, interferindo no nível amostral de sequências repetitivas nas etapas de pré-processamento e nivelamento do

RE para eliminar o viés da ploidia. Outra questão abrange o corte das sequências em 75pb padronizado para todas as 13 espécies das autoras. Esses fatores da análise são considerados como impactantes nos resultados gerados na plataforma do RE, como o corte ideal das sequências entre 100-300pb, que podem ocasionar descarte de sequências na análise e consequente menores quantidades finais de DNA repetitivo (NOVÁK, 2020).

Em relação às quantidades de retrotransposons da ordem LTR, a proporção de Ty3/gypsy, a superfamília mais abundante, variou de 24,7% em *S. calvescens* a 31,1% em *S. chacoense*. Para Ty1/copia, as proporções variaram de 2,5% em *S. etuberosum* a 3,1% em *S. commersonii*. Gaiero et al (2019) também descreveu maior abundância de Ty3/gypsy em relação a Ty1/copia. Com relação a Ty3/gypsy foram descritas as proporções de 21,79%, 21,26% e 20,88%, para *S. chacoense*, *S. commersonii* e *S. etuberosum*, respectivamente, enquanto para Ty1/copia as proporções foram 1,75%, 1,83%, 1,41%. De forma geral, a maioria das angiospermas apresentam maior abundância de Ty3/gypsy em relação em Ty1/copia, com raras exceções (MASCAGNI et al., 2020).

Para a classe dos transposons, as famílias que apresentaram maior abundância em todas as espécies foram hAT, variando de 0,39% em *S. etuberosum* a 0,75% em *S. commersonii*, e CACTA, variando de 0,32% em *S. etuberosum* a 0,64% em *S. malmeanum*. A maior abundância destes elementos também foi encontrada por Gaiero et al (2019), para as espécies *S. chacoense*, *S. commersonii* e *S. etuberosum*, com 0,24%, 0,24% e 0,05%, para hAT, e 0,1%, 0,12% e 0,05%, respectivamente.

As espécies apresentaram, baixas abundâncias de sequências satélite, variando de 0,67% em *S. calvescens* a 2,92% em *S. etuberosum*. Estas baixas quantidades de DNA satélite já foram descritas no gênero *Solanum* por Gaiero et al. (2019) para as espécies selvagens de batata *S. chacoense* (1,14%), *S. commersonii* (1,06%) e *S. etuberosum* (1,93%) em comparação a cultivares e espécies de tomate, das quais podem ser encontradas proporções de três a seis vezes maiores de sequências em tandem.

A maioria dos satélites identificados está compartilhada entre as espécies, mesmo que com abundância diferenciada. As exceções são o satélite de *S. calvescens*, ScalvSat8, que apresentou um único read em *S. chacoense*, e os SchacSat9 e SchacSat10, de *S. chacoense*, que foram exclusivos dessa espécie. Porém, como este último foi identificado via BLAST como uma variante presente no gênero *Trypanosoma* sp., esta sequência foi descartada como representante da espécie *S. chacoense*.

A exclusividade de sequências de DNA repetitivo em *S. chacoense*, como visto no satélite SchacSat9, pode indicar um marcador exclusivo da espécie e o distanciamento entre essa espécie e *S. calvescens*. Esta condição pode ser explicada pela “hipótese da biblioteca”, em que espécies ditas filogeneticamente próximas podem compartilhar uma biblioteca ancestral de diferentes sequências, mas que pode ser amplificada diferentemente em cada espécie (Plohl et al., 2012). A diferenciação das duas espécies com relação aos satélites é observada também na sequência ScalvSat8.

A existência de satélites espécie-específicos pode ser usada para identificação rápida de espécies devido ao alto número de cópias e fácil acessibilidade, utilizando de técnicas sensíveis como a amplificação de PCR usando primers que reconhecem especificamente o respectivo satDNA (HEMLEBEN et al. 2007). Em *Solanum*, Schweizer et al. (1988) utilizaram esse recurso para distinguir híbridos somáticos de *Solanum acaule* e *Lycopersicum esculentum*.

Estudando híbridos alotriploides entre *S. commersonii* e *S. tuberosum*, Gaiero et al. (2017) discutem a dificuldade encontrada na pintura cromossômica destes híbridos, a fim de determinar o sucesso da introgressão entre os genomas, algo que pode se estender para outras espécies de batata relacionadas. Dessa forma, a utilização deste satélite exclusivo SchacSat9 como sonda pode ser uma alternativa na elucidação do seu parentesco com *S. calvescens*, caso as técnicas clássicas de GISH apresentem dificuldade semelhante.

A sequência satélite ScalvSat4 apresenta similaridade com uma sequência descrita em *S. demissum* (ZANKE & HEMLEBEN, 1997) como originária e exclusiva de espécies selvagens, podendo ser encontrada em algumas linhagens cultivadas devido a eventos de introgressão em processos de cruzamento com *S. tuberosum*. Esse satélite correspondeu também à família SPG descrita por Gebhardt et al. (1995) em *Solanum spegazzinii*. Segundo os autores, esta família se localiza nas regiões subteloméricas dos cromossomos de batatas e aparece em maior abundância em espécies selvagens, em comparação às espécies cultivadas.

A sequência satélite correspondente a ScalvSat8 não foi correspondida via BLAST com outras espécies, apenas com os três híbridos de *S. tuberosum*, SolyntusTM, MSH/14-112 e P8, sugerindo que pode ter havido introgressão via cruzamentos.

A cultivar SolyntusTM teve seu genoma sequenciado por Van Lieshout et al. (2020). O desenvolvimento da cultivar envolveu cruzamentos para introgressão de alelos de outras espécies relacionadas de batata, entre elas, uma linhagem de *S. chacoense* que foi cruzada logo no início do processo, o que explicaria a identificação dos majoritariamente encontrados em espécies selvagens. Os outros dois híbridos de *S. tuberosum*, MSH/14-112 e P8, que também

apresentaram correspondências com as sequências satélite encontradas, foram desenvolvidos por Tiwari et al. (2013; 2018) na Índia e também envolveram processos de cruzamento com espécies selvagens visando à resistência à requeima.

Por fim, a espécie *Solanum malmeanum* não apresentou um satélite identificado e confirmado após a análise do TAREAN e confirmação via Dotmatcher, porém não se pode descartar a possibilidade de sequências satélite com baixo número de cópias estarem presentes.

A caracterização da fração repetitiva aponta para distanciamento evolutivo entre *S. calvescens* e as demais espécies, tanto o tamanho do genoma encontrado quanto a abundância das sequências repetitivas sugerem uma resposta ao estresse ambiental e aos eventos de poliploidização. A presença de satélite específico de *S. chacoense* reforça a hipótese de *S. calvescens* como espécie e não citótipo da mesma, e estudos complementares citogenéticos e genômicos podem contribuir ainda mais para a taxonomia do grupo.

6 CONCLUSÕES

A fração repetitiva total do genoma de *Solanum calvescens*, *Solanum chacoense*, *Solanum commersonii*, *Solanum malmeanum* e *Solanum etubersoum*, foram quantificados em, respectivamente, 38,4%, 45,4%, 43,4%, 42,8% e 45% do genoma.

Os elementos repetitivos mais abundantes em todas as espécies foram os retrotransposons de longas repetições terminais (LTR), sendo a super família Ty3/gypsy mais abundante que Ty1/gypsy.

As espécies compartilham oito superfamílias de DNA satélite, e uma é específica de *S. chacoense*.

A comparação da fração repetitiva aponta para a diferenciação de *S. calvescens* com relação às outras espécies, corroborando a hipótese de que se trata de uma espécie e não um citótipo.

REFERÊNCIAS BIBLIOGRÁFICAS

- AVERSANO R., CONTALDI F., ERCOLANO M. R. The *Solanum commersonii* genome sequence provides insights into adaptation to stress conditions and genome evolution of wild potato relatives. **The Plant Cell** 27: 954–968. 2015.
- BAE, J., HALTERMAN, D., and JANSKY, S. H.. Development of a molecular marker associated with *Verticillium* wilt resistance in diploid interspecific potato hybrids. **Mol. Breed.** 22:61–69. 2008.
- BAO, W.; KOJIMA, K. K.; KOHANY, O. Repbase Update, a database of repetitive elements in eukaryotic genomes. **Mobile DNA**, v. 6, n. 1, p. 11, 2 jun. 2015.
- BENNETT M.D., LEITCH I.J. Nuclear DNA amounts in Angiosperms: progress, problems and prospects. **Annals of Botany** 95:45-90. 2005.
- BERKE, L., GRANDONT, L., AFLITOS, S. A., BACHEM, C. W. B., BECKER, F. F. M., BOUWMEESTER, K., VAN DE GEEST, H. C., GOVERS, F., DE JONG, H., PETERS, S. A., SANCHEZ PEREZ, G. F., SCHIJLEN, E. G. W. M., VAN DEN BERG, R., & SCHRANZ, M. E. (2016). You say tomato, I say potato: high-quality genome assembly of the sister-group species *Solanum etuberosum* provides insights into genome and trait evolution. 28-29. **Solanaceae Conference**, Davis, California, United States. 2016.
- BIÉMONT, C., VIEIRA, C. Genetics: junk DNA as na evolutionary force. **Nature**, London, v. 443, n. 521-524, 2006.
- BOMBARELY, A., ROSLI, H.G., VREBALOV, J., MOFFETT, P., MUELLER, L.A. and MARTIN, G.B. A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. **Mol. Plant Microbe Interact.** 25, 1523–1530. 2012.
- BOUTIN T.S., LE ROUZIC A., CAPY P. How does selfing affect the dynamics of selfish transposable elements? **Mobile DNA Journal** 3:5. 2012.
- BROWN, J.; PIRRUNG, M.; MCCUE, L. A. FQC Dashboard: integrates FastQC results into a web-based, interactive, and extensible FASTQ quality control tool. **Bioinformatics**, v. 33, n. 19, p. 3137–3139, 1 out. 2017.
- BRÛCHER, H. Batatas Silvestres Brasileiras - Seção Tuberarium do gênero na Região Sul do Brasil. **EMBRAPA Hortaliças**, Brasília, DF, n. 21, 14 p, 1999.
- CACACE, J. E.; HUARTE, M. Descubriendo la pala. **Balcare: PROPAPA** – Fundacion Argeninta. 1996.
- CARPUTO, D. Cytological and breeding behavior of pentaploids derived from 3x × 4x crosses in potato. **Theoretical and Applied Genetics** 106(5): 883–888. 2003.
- CASA-COILA, V. H., GOMES, C., LIMA-MEDINA, I., SOMAVILLA, L., & CASTRO, C. (2012). Resistência foliar in vitro de espécies silvestres de batata a *Phytophthora infestans* (Mont.) de Bary. In **Embrapa Clima Temperado-Artigo em anais de congresso (ALICE)**.

In: Congresso de La Asociación Latinoamericana de La Papa, 25.; Encontro Nacional de Produção e Abastecimento de Batata, 14., 2012, Uberlândia: Asociación Latinoamericana de la Papa, 2012.

CUNHA, A. L. Caracterização agrônômica e produção de pólen 2n de haploides de *Solanum tuberosum* x *Solanum chacoense*. **Dissertação (Mestrado em Agronomia) UFLA**, Lavras-MG, 1992.

DEL CASTILLO, D. L. et al. Cranial development and directional asymmetry in Commerson's dolphin, *Cephalorhynchus commersonii commersonii*: 3D geometric morphometric approach. **Journal of Mammalogy**, v. 97, n. 5, p. 1345-1354, 2016.

DERAGON, J. M., CAPY, P. Impacto of transposable elements on the human genoma. **Annals of Medicine, Helsinki**, v. 32, p. 264-273, 2000.

DODSWORTH, Steven. Genome skimming for next-generation biodiversity analysis. **Trends in plant science**. 20. 2015.

EHLENFELDT M. K, HANNEMAN R. E Jr. The use of Endosperm Balance Number and 2n gametes to transfer exotic germplasm in potato. **Theoretical and Applied Genetics** 68:155–161. 1984.

FAO – Food and Agriculture Organization. Forthcoming (c). **The State of Agricultural Commodity Markets**. Rome. 2008.

FAOSTAT – Food and Agriculture Organization of the United Nations - **Statistics Division**. Disponível em: <<http://www.fao.org/faostat/en/#data/QC>> Acesso em: 16 Fev. 2020.

FERDOROFF, N. Transposons and genome evolution in plants. **Proceedings of the National Academy of Sciences of the United States of America**. Washington, v. 20, n. 13, p. 700-702, 2000.

FERNANDES, S. B. Caracterização Citogenética da Espécie Selvagem *Solanum calvescens* Bitt. **Monografia**. UFLA. 2016.

GAIERO, P., MAZZELLA, C., VILARÓ, F. et al. Pairing analysis and in situ Hybridisation reveal autopolyploid-like behaviour in *Solanum commersonii* × *S. tuberosum* (potato) interspecific hybrids. **Euphytica** **213**, 137 (2017).

GAIERO, P., VAIO, M., PETERS, S. A., SCHRANZ, M. E., DE JONG, H., & SPERANZA, P. R. Comparative analysis of repetitive sequences among species from the potato and the tomato clades. **Annals of Botany**. 2019.

GARRIDO-RAMOS, M. Satellite DNA: An Evolving Topic. **Genes**, v. 8, n. 9, p. 230, 18 set. 2017.

GAVRILENKO, T. Application of molecular cytogenetics in fundamental and applied research of potato. **Genetics, Genomics and Breeding of Potato**/Ed. J. Bradeen, C. Kole. Published by 'Science Publishers'. 1st edition. USA, 184-206. 2011.

GBIF. *Solanum calvescens* Bitter in GBIF Secretariat. **GBIF Backbone Taxonomy**. Checklist dataset <https://doi.org/10.15468/39omei> accessed via GBIF.org. 2021.

GEBHARDT, C., EBERLE, B., LEONARDS-SCHIPPERS, C., WALKEMEIER, B., & SALAMINI, F. Isolation, characterization and RFLP linkage mapping of a DNA repeat family of *Solanum spigazzinii* by which chromosome ends can be localized on the genetic map of potato. **Genetical Research**, 65(01), 1. 1995.

GONG, Z. et al. Repeatless and repeat-based centromeres in potato: implications for centromere evolution. **The Plant Cell**, v. 24, n. 9, p. 3559-3574, 2012.

GREILHUBER J. et al. The Origin, Evolution and Proposed Stabilization of the Terms ‘Genome Size’ and ‘C-Value’ to Describe Nuclear DNA Contents, **Annals of Botany**, v. 95, n.1, p. 255–260, 2005.

GROVER C.E., WENDEL J.F. Recent insights into the mechanisms of genome size change in plants. **Journal of Botany**. 2010.

HANNEMAN JUNIOR, R. E. The reproductive biology of the potato and its implication for breeding. **Potato Research, Wageningen**, v. 42, n. 1, p. 283-312, 1999.

HARDIGAN, Michael A. et al. Genome diversity of tuber-bearing *Solanum* uncovers complex evolutionary history and targets of domestication in the cultivated potato. **Proceedings of the National Academy of Sciences**, v. 114, n. 46, p. E9999-E10008, 2017.

HAWKES, J. G. History of the potato. In: HARRIS, P. M. (Ed.). *The Potato Crop*. Londres: **Chapman and Hall Ltd**. p. 1-14. 1978.

HAWKES, J. G. *The potato: evolution, biodiversity, and genetic resources*. Washington, DC: **Smithsonian Institution Press**. 1990.

HAWKINS J.S., PROULX S.R., RAPP R.A., WENDEL J.F. Rapid DNA loss as a counterbalance to genome expansion through retrotransposon proliferation in plants. **Proceedings of the National Academy of Sciences USA** 106:17811-17816. 2009.

HEITKAM, T.; WEBER, B.; WALTER, I.; LIEDTKE, S.; OST, C.; SCHMIDT, T. Satellite DNA landscapes after allotetraploidisation of quinoa (*Chenopodium quinoa*) reveal unique A and B subgenomes. **The Plant Journal**, n. March, p. tpj.14705, 24 jan. 2020.

HEMLEBEN, V., KOVARIK, A., TORRES-RUIZ, R. A., VOLKOV, R. A., & BERIDZE, T.. Plant highly repeated satellite DNA: Molecular evolution, distribution and use for identification of hybrids. **Systematics and Biodiversity**, 5(3), 277–289. 2007.

HESLOP-HARRISON, J. (Pat); SCHMIDT, T. Plant Nuclear Genome Composition. **Encyclopedia of Life and Science**, p. 1–8, 2007.

HIJMANS, R., T. GAVRILENKO, S. STEPHENSON, J. BAMBERG, A. SALAS, and D. M. SPOONER. Geographic and environmental range expansion through polyploidy in wild potatoes (*Solanum* section *Petota*). **Glob. Ecol. Biogeogr.** 16:485–495. 2007.

HIRAKAWA, H., SHIRASAWA, K., MIYATAKE, K. et al.. Draft genome sequence of eggplant (*Solanum melongena* L.): the representative solanum species indigenous to the old world. **DNA Res.** 21, 649–660. 2014.

IBGE - Instituto Brasileiro de Geografia e Estatística. **Levantamento sistemático da produção agrícola.** Disponível em: <https://www.ibge.gov.br/estatisticasnovportal/economicas/agricultura-e> Acesso em: 16 Fev. 2020.

JIANG, J; BIRCHLER J. A.; PARROTT, W.A.; DAWE, R.K. A molecular view of plant centromeres. **Trends in Plant Science.**,8 (12): 570-575, 2003.

JURKA, J.; KAPITONOV, V.V.; KOHANY, O.; JURCA, M.V. Repetitive sequences in complex genomes: structure and evolution. **Annu. Rev. Genomics Hum. Genet.** 8:241-259, 2007.

KASHKUSH K., FELDMAN M., LEVY A. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. **Nature Genetics** 33:102–106. 2003.

KATOH, K.; STANDLEY, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. **Molecular Biology and Evolution**, v. 30, n. 4, p. 772–780, 1 abr. 2013.

KEARSE, M.; MOIR, R.; WILSON, A.; STONES-HAVAS, S.; CHEUNG, M.; STURROCK, S.; BUXTON, S.; COOPER, A.; MARKOWITZ, S.; DURAN, C.; THIERER, T.; ASHTON, B.; MEINTJES, P.; DRUMMOND, A. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, v. 28, n. 12, p. 1647–1649, 2012.

KIM, S., PARK, M., YEOM, S.I. et al. Genome sequence of the hot pepper provides insights into the evolution of pungency in *Capsicum* species. **Nat. Genet.** 46, 270–278. 2014.

KLASEN, G. L.; CASTRO, C. M.; HEIDEN, G. Cruzabilidade entre parentes silvestres da batata, visando ampliar a base genética da cultura. In: **Embrapa Clima Temperado-Artigo em anais de congresso (ALICE)**. In: Congresso De Iniciação Científica, 26.; Encontro De Pós-Graduação Ufpel, 19.; Semana Integrada De Ensino, Pesquisa E Extensão, 3, Pelotas. Anais. Pelotas: UFPel, 2017.

KLASEN, G. L.; CASTRO, C. M.; HEIDEN, G. Relações de cruzabilidade entre a batata e parentes silvestres do Brasil. In: **Embrapa Clima Temperado-Artigo em anais de congresso (ALICE)**. In: Congresso De Iniciação Científica, 27.; Encontro De Pós-Graduação Ufpel, 20.; Semana Integrada De Ensino, Pesquisa E Extensão, 4, Pelotas. Anais. Pelotas: UFPel, 2018.

KUBIS, S. E., SCHMIDT, T., HESLOP-HARRISON, J. S. Repetitive DNA elements as a major component of plant genomes. **Annals of Botany**, Oxford, v. 82, p. 45-55, 1998.

- KYRIAKIDOU, M., ANGLIN, N.L., ELLIS, D. *et al.* Genome assembly of six polyploid potato genomes. **Sci Data** 7, 88. 2020.
- LEISNER, C. P., HAMILTON, J. P., CRISOVAN, E. Genome sequence of M6, a diploid inbred clone of the high glycoalkaloid-producing tuber-bearing potato species *Solanum chacoense*, reveals residual heterozygosity. **The Plant Journal** 94: 562–570. 2018.
- LI, W.; COWLEY, A.; ULUDAG, M.; GUR, T.; MCWILLIAM, H.; SQUIZZATO, S.; PARK, Y. M.; BUSO, N.; LOPEZ, R. The EMBL-EBI bioinformatics web and programmatic tools framework. **Nucleic acids research**, v. 43, n. W1, p. W580-4, 1 jul. 2015.
- LOUWES K M, HOEKSTRA R, MATTHEIJ W M.. Interspecific hybridization between the cultivated potato *Solanum tuberosum* subspecies *tuberosum* L. and the wild species *S.circaeifolium* subsp. *circaeifolium* Bitter exhibiting resistanceto Phytophthorainfestans (Mont.) de Bary and Globoderapallida (Stone) Behrens. 2. Sexual hybrids. **Theoretical and Applied Genetics** 84:363–370. 1992.
- MACAS, J.; KEJNOVSKÝ, E.; NEUMANN, P.; NOVÁK, P.; KOBLÍŽKOVÁ, A.; VYSKOT, B. Next Generation Sequencing-Based Analysis of Repetitive DNA in the Model Dioceous Plant *Silene latifolia*. **PLoS ONE**, v. 6, n. 11, p. e27335, 9 nov. 2011.
- MACAS J., NEUMANN, P., NAVRÁTILOVÁ, A. Repetitive DNA in the pea (*Pisum sativum* L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and *Medicago truncatula*. **BCM Genomics** 8: 427. 2007.
- MASCAGNI, F.; VANGELISTI, A.; GIORDANI, T.; CAVALLINI, A.; NATALI, L. A computational comparative study of the repetitive DNA in the genus *Quercus* L. **Tree Genetics & Genomes**, v. 16, n. 1, p. 11, 23 fev. 2020.
- MASUELLI R W, CAMADRO E L, MENDIBURU A O. 2n gametes in *Solanum commersonii* and cytological mechanisms of triplandroid formation in triploid hybrids of *Solanum commersonii* × *S. gourlayi*. **Genome**, 35:864–869. 1992.
- MATSUBAYASHI, M. Phylogenetic relationships in the potato and its related species. In: T Tsuchiya, PK Gupta (eds) *Chromosome Engineering in Plants: Genetics, Breeding, Evolution, Part B*. Elsevier, Amsterdam, The Netherlands, pp 93–118. 1991.
- MEHROTRA, Shweta; GOYAL, Vinod. Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. **Genomics, proteomics & bioinformatics**, v. 12, n. 4, p. 164-171, 2014.
- MEYERS, B. C., TINGEY, S. V., MORGANTE, M. Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. **Genome Research**, New York, v.11, p. 1660-1676, 2001.
- MOK, D. W. S. A.; PELOQUIN, S. J. The inheritance of three mechanisms of diplandroid (2n pollen) formation in diploid potatoes. **Heredity**, Baltimore, v. 35, n. 2, p. 295- 302, 1975.

MOLIN, L. H. D.; CASTRO, C. M.; HEIDEN, G. . Acesso de espécies silvestres no Banco Ativo de Germoplasma de batata da Embrapa Clima Temperado. In: **8º Congresso Brasileiro de Melhoramento de Plantas, Goiânia. Anais**, 2015a.

MOLIN, LHD; CASTRO, CAROLINE MARQUES; HEIDEN, GUSTAVO. Distribuição geográfica de batatas silvestres (*Solanum*, Solanaceae) nativas do Brasil. In: **Embrapa Clima Temperado-Artigo em anais de congresso (ALICE)**. In: Congresso de Iniciação Científica; Encontro de Pós-Graduação. Pelotas. [Anais.]. Pelotas: UFPel, 2015b.

MOLIN, L. H. D.; COSTA, F. A.; CASTRO, C. M.; HEIDEN, G. Parentes silvestres de batata (*Solanum* sect. *Petota*, Solanaceae) no Brasil: preenchendo lacunas de coletas em bancos de germoplasma. In: **5º Congresso Brasileiro de Recursos Genéticos**, 2016.

NOVÁK, P. et al. RepeatExplorer: A Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. **Bioinformatics** v. 29, n.6, p. 792–793. 1367-4811 (Electronic) n1367-4803 (Linking). 2013.

NOVÁK, P., ÁVILA ROBLEDDILLO, L., KOBLÍŽKOVÁ, A., VRBOVÁ, I., Neumann, P., & Macas, J. TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. **Nucleic Acids Research**, 45(12), e111–e111. 2017.

NOVÁK, P.; HŘIBOVÁ, E.; NEUMANN, P.; KOBLÍŽKOVÁ, A.; DOLEŽEL, J.; MACAS, J. Genome-Wide Analysis of Repeat Diversity across the Family Musaceae. **PLoS ONE**, v. 9, n. 6, p. e98918, 16 jun. 2014.

NOVÁK, P.; NEUMANN, P.; MACAS, J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. **Nature protocols**, v. 15, n. 11, p. 3745-3776, 2020.

OLIVEIRA, G. K. de. Distribuição de sequências centroméricas de *Solanum* nos genomas A, C e D de espécies aloploidoides. 2017. 50 p. **Dissertação - Universidade Federal de Lavras, Lavras, Repositório da UFLA**. 2017.

OLIVEIRA, M. N.; DAVIDE, L. C.; PINTO, C. A. P. B. Mechanisms of 2n potato pollen formation in dihaploid *Solanum tuberosum* L. x *S. chacoense* Bitt. hybrid clones. **Brazilian Journal of Genetics**, v. 18, n. 3, p. 445–450, 1995.

OLIVEIRA, L. C. DNA repetitivo centromérico e disperso em espécies de *Solanum*. 2015. 113 f. **Tese (Doutorado em Genética e Melhoramento de Plantas)**. Universidade Federal de Lavras-UFLA. Lavras- MG. 2015.

PENDINEN, G., GAVRILENKO, T., JIANG, J., SPOONER, D. M.. Allopolyploid speciation of the Mexican tetraploid potato species *Solanum stoloniferum* and *S. hjertingii* revealed by genomic in situ hybridization. **Genome**. 2008.

PEREIRA, J. A. Isolamento e análise de sequências de DNA repetitivo em espécies da Tribo *Meliponini*. **Locus Repositório Institucional da UFV**. 2018.

- PHAM, G. M., HAMILTON, J. P., WOOD, J. C., BURKE, J. T., ZHAO, H., VAILLANCOURT, B., ... & BUELL, C. R. Construction of a chromosome-scale long-read reference genome assembly for potato. **GigaScience**, 9(9). 2020.
- PLOHL, M., MEŠTROVIĆ, N. & MRAVINAC, B. Satellite DNA evolution. **Genome Dynamics**. 7:126–152. 2012.
- PGSC - POTATO GENOME SEQUENCING CONSORTIUM et al. Genome sequence and analysis of the tuber crop potato. **Nature**, v. 475, n. 7355, p. 189-195, 2011.
- RSTUDIO Team (2021). RStudio: Integrated Development for R. **RStudio, PBC**, Boston, MA URL <http://www.rstudio.com/>.
- RUIZ-RUANO, F. J.; LÓPEZ-LEÓN, M. D.; CABRERO, J.; CAMACHO, J. P. M. High throughput analysis of the satellitome illuminates satellite DNA evolution. **Scientific Reports**, v. 6, n. 1, p. 28333, 7 set. 2016.
- SANMIGUEL, P. GAUT, B. S., TIKHONOV, A., JIN, K. et. al. Nested retrotransposons in the intergenic regions of the maize genome. **Science**. 274 :765-768. 1996.
- SCHIFINO-WITTMANN, M. T.; DALL'AGNOL, M. Gametas Não Reduzidos no Melhoramento De Plantas. **Ciência Rural**, v. 31, n. 1, p. 169–175, 2001.
- SCHWEIZER, G., GANAL, M., NINNEMANN, H. *et al.* Species-specific DNA sequences for identification of somatic hybrids between *Lycopersicon esculentum* and *Solanum acaule*. *Theoret. Appl. Genetics* 75, 679–684. 1988.
- SIMON, R., XIE, C., CLAUSEN, A., JANSKY, S., HALTERMAN, D., CONNER, T., SPOONER, D. (2010). Wild and Cultivated Potato (*Solanum* sect. *Petota*) Escaped and Persistent Outside of its Natural Range. **Invasive Plant Science and Management**, 3(3), 286-293. doi:10.1614/IPSM-D-09-00043.1 2010.
- SPOONER, D. M. et al. Systematics, diversity, genetics, and evolution of wild and cultivated potatoes. **The botanical review**, v. 80, n. 4, p. 283-383, 2014.
- SPOONER, David M. et al. Genomic origins of potato polyploids: GBSSI gene sequencing data. **Crop Science**, v. 48, n. Supplement_1, p. S-27-S-36, 2008.
- TEK, A. L. et al. Sobo, a recently amplified satellite repeat of potato, and its implications for the origin of tandemly repeated sequences. **Genetics**, v. 170, n. 3, p. 1231-1238, 2005.
- TEK, A. L.; JIANG, J. The centromeric regions of potato chromosomes contain megabase-sized tandem arrays of telomere-similar sequence. **Chromosoma**, v. 113, n. 2, p. 77-83, 2004.
- THE TOMATO GENOME CONSORTIUM. The tomato genome sequence provides insights into fleshy fruit evolution. **Nature**, 485, 635–641. 2012.

TIWARI J. K., POONAM, KUMAR V., SINGH B. P., SHARMA S., LUTHRASK and BHARDWAJ V (2013) Evaluation of potato somatic hybrids of dihaploid *S. tuberosum*(+) *S. pinnatisectum* for late blight resistance. **Potato J40**: 176-179. 2013.

TIWARI, Jagesh & LUTHRA, S. & DEVI, Sapna & KUMAR, Vinod & ALI, Nilofer & ZINTA, Rasna & CHAKRABARTI, Swarup. Development of Advanced Back-Cross Progenies of Potato Somatic Hybrids and Linked ISSR Markers for Late Blight Resistance with Diverse Genetic Base- First Ever Produced in Indian Potato Breeding. **Potato Journal**. 45. 17-27. 2018.

TOMÉ, L. G. O., ALVES, A. A., SALGADO, C. C., DAVIDE, L. C., PINTO, C. A. B. P. Pólen 2n e mecanismos meióticos de formação em *Solanum commersonii* ssp. **Ciência e Agrotecnologia**, 33(2), 473-477. 2009.

TORRES, GIOVANA. A. et al. Organization and evolution of subtelomeric satellite repeats in the potato genome. **G3: Genes, Genomes, Genetics**, v. 1, n. 2, p. 85-92, 2011.

VAN LIESHOUT, N., VAN DER BURGT, A., de VRIES, M. E., TER MAAT, M., EICKHOLT, D., ESSELINK, D., FINKERS, R. Solyntus, the New Highly Contiguous Reference Genome for Potato (*Solanum tuberosum*); **Genes|Genomes|Genetics**, g3.401550. 2020.

WATANABE, K. Potato genetics, genomics, and applications. **Breeding Sci.** 65, 53–68. 2015.

WEBER, B. N.; HAMERNIK, A. J.; JANSKY, S. H. Hybridization barriers between diploid *Solanum tuberosum* and wild *Solanum raphanifolium*. **Genetic Resources and Crop Evolution**, v. 59, n. 7, p. 1287–1293, 2012.

WICKER, T., SABOT, F., HUA-VAN, A. BENNETZEN, J. L., CAPY, P., CHALOUB, B., FLAVELL, A., LEROY, P. MORGANTE, M., PANAUD, O., PAUX, E., SANMIGUEL, P., SCHULMAN, A. A unified classification system for eukaryotic transposable elements. **Nature Rev. Genet** 8: 973-982. 2007.

ZANKE, C., & HEMLEBEN, V. A new *Solanum* satellite DNA containing species-specific sequences which can be used for identification of genome parts in somatic hybrids of potato. **Plant Science**, 126(2), 185–191.1997.

ZHANG, H. et al. Boom-Bust Turnovers of Megabase-Sized centromeric DNA in *Solanum* Species: Rapid Evolution of DNA Sequences Associated with Centromeres. **Plant Cell**. 26(4):1436-1447. 2014.