

UNIVERSIDADE FEDERAL DE ALAGOAS
CAMPUS ARAPIRACA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRICULTURA E AMBIENTE

MARIA ALICE SILVA OLIVEIRA

**CARACTERIZAÇÃO GENÔMICA DE SEQUÊNCIAS REPETITIVAS EM
ESPÉCIES DIPLOIDES E POLIPLOIDES DO GÊNERO *Stylosanthes* Sw.**

ARAPIRACA

2020

Maria Alice Silva Oliveira

Caracterização genômica de sequências repetitivas em espécies diploides e poliploides do gênero *Stylosanthes* Sw.

Dissertação de Mestrado apresentada ao programa de Pós-Graduação em Agricultura e Ambiente da Universidade Federal de Alagoas – *Campus* Arapiraca, como requisito parcial para obtenção do grau de Mestra em Agricultura e Ambiente.

Orientador: Prof. Dr. André Seco Marques da Silva

Arapiraca
2020

Catálogo na fonte
Universidade Federal de Alagoas – UFAL
Campus Arapiraca

Bibliotecário Responsável: Márcio Thiago dos Santos Albuquerque - CRB - 4 / 2052

O48c Oliveira, Maria Alice Silva

Caracterização genômica de sequências repetitivas em espécies diploides e poliploides do gênero *Stylosanthes Sw.* / Maria Alice Silva Oliveira. – 2020.

72 f. : il. color.

Orientador: André Seco Marques da Silva.

Dissertação (Mestrado) – Programa de Pós-Graduação em Agricultura e Ambiente - PPGAA, Universidade Federal de Alagoas – UFAL, *Campus Arapiraca*, Arapiraca, 2020.

Bibliografia: f.41-48.

Apêndice: f. 49-72

1. Aloploidia. 2. Filogenia. 3. Bioinformática. 4. DNA repetitivo. I. Silva, André Seco Marques da. II. Título.

CDU: 631:575.113

MARIA ALICE SILVA OLIVEIRA

**CARACTERIZAÇÃO GENÔMICA DE SEQUÊNCIAS REPETITIVAS EM ESPÉCIES DIPLOIDES
E POLIPLOIDES DO GÊNERO *Stylosanthes* Sw**

Dissertação submetida ao Corpo Docente do
Programa de Pós-Graduação em Agricultura e
Ambiente da Universidade Federal de Alagoas e
aprovada em 20 de fevereiro de 2020.

Prof. Dr. André Seco Marques da Silva
Universidade Federal de Alagoas – UFAL
Campus Arapiraca
Presidente Titular

Banca Examinadora:

Prof. Dr. Cicero Carlos de Souza Almeida
Universidade Federal de Alagoas – UFAL
Campus Arapiraca
Membro Titular

Prof. Dr. Jakson Leite
Universidade Federal do Maranhão
Campus Bacabal
Membro Titular

AGRADECIMENTOS

Agradeço primeiramente a Deus, pelo dom da vida, por seu amor e cuidado, e por suas infinitas bênçãos;

Aos meus pais e minhas irmãs pelos ensinamentos e pela paciência;

À Universidade Federal de Alagoas por ter me proporcionado essa pós-graduação;

A todos os professores do Programa de Pós-Graduação em Agricultura e Ambiente da Universidade Federal de Alagoas – *Campus* Arapiraca e também a banca avaliadora pelas contribuições realizadas;

Ao Prof. Dr. André Seco Marques da Silva, meu orientador, pela atenção, pelos ensinamentos e por toda disponibilidade;

Ao meu namorado Iago Saturnino de Lima, por todo carinho, paciência e motivação, e por estar sempre ao meu lado diante de qualquer situação;

Aos integrantes do Laboratório de Recursos Genéticos, pelas experiências compartilhadas ao longo desta etapa;

À Fapeal e Capes, pelo apoio financeiro concedido para a realização deste trabalho;

À minha amiga, Alaide Maria, pelo companheirismo durante o curso e pela motivação;

Aos colegas de turma, pelo compartilhamento de conhecimentos e experiências;

A todos que de alguma forma contribuíram direta ou indiretamente para a realização deste trabalho.

Muito obrigada!!!!

“Ninguém comete erro maior do que não fazer nada porque só pode fazer um pouco.” (Edmund Burke)

RESUMO

O genoma nuclear corresponde a maior fração de DNA em eucariotos, e nele estão inseridas as sequências de DNA repetitivo, que podem corresponder em até 90% desse genoma. As sequências repetitivas são grandes aliadas nos estudos evolutivos, ajudando na compreensão da organização e do comportamento evolutivo dos genomas eucarióticos. O advento do sequenciamento de nova geração facilitou a obtenção de uma grande quantidade de sequências e com preços mais acessíveis, possibilitando a descoberta das sequências repetitivas. Dentre as metodologias que surgiram com esse advento está o programa RepeatExplorer, uma ferramenta de alto potencial na descoberta dessas sequências. O gênero *Stylosanthes* Sw. está inserido na família Leguminosae e apesar de possuir alto valor econômico e ambiental ainda possui uma baixa caracterização genética, devido a ocorrência de espécies híbridas e alta complexidade sistemática. É um gênero diverso e com ampla distribuição, sendo o Brasil o seu principal centro de origem e diversidade. O presente estudo buscou compreender a organização do genoma nuclear de nove espécies do gênero *Stylosanthes* Sw., incluindo dois complexos alopoliploides de interesse econômico (Complexo *S. scabra* e Complexo *S. capitata*), através da análise genômica das sequências repetitivas, buscando caracterizar a abundância dessas sequências e entender as relações filogenômicas entre diploides e alopoliploides. Os *reads* genômicos das espécies *S. hamata*, *S. viscosa* e *S. scabra* foram obtidos do estudo de Marques *et al.* (2018), disponíveis no NCBI, e para as demais espécies foi feito o sequenciamento pela GenOne Soluções em Biotecnologia, Rio de Janeiro, Brasil. Os reads das nove espécies foram inseridos no software Geneious, onde foram pareados, convertidos em formato fasta e selecionados. As sequências obtidas foram concatenadas e inseridas no programa RepeatExplorer, sendo caracterizadas individualmente e comparativamente, através da versão RepeatExplorer2 e da ferramenta TAREAN. Foi realizada ainda uma análise filogenética a partir do satelitoma de cada espécie, utilizando o programa AAF. As análises do RepeatExplorer revelaram que as nove espécies de *Stylosanthes* estudadas possuem mais de 50% do seu genoma formado por sequências repetitivas, sendo a classe de elementos transponíveis da ordem LTR a mais abundante. A abundância de DNA satélite variou entre as espécies, e o agrupamento dos reads em gráficos de leitura revelou boa conservação das sequências, representando baixa diversificação no processo evolutivo. A análise filogenética possibilitou a montagem de uma árvore com relações de parentesco entre as espécies esclarecidas e sustentando estudos anteriores. Diferentemente do que nos mostra a literatura, nosso estudo revelou uma estreita relação entre as espécies *S. hamata* e *S. seabrana*, podendo se tratar de sinonímias em processo de especiação. As diferenças encontradas entre os resultados e o que consta na literatura reforça a necessidade de estudos mais aprofundados que revisem a taxonomia e filogenética do grupo, principalmente no âmbito molecular.

Palavras-chave: DNA repetitivo. Alopoliploidia. Filogenômica. Bioinformática.

ABSTRACT

The nuclear genome corresponds to the largest fraction of DNA in eukaryotes, and is inserted as repetitive DNA sequences, which can correspond to up to 90% of this genome. As repetitive sequences are great allies in evolutionary studies, they help to understand the organization and the evolutionary behavior of eukaryotic genomes. The advancement of next generation sequencing facilitates the use of a large number of sequences and more affordable prices, enabling the discovery of repetitive sequences. Among the methodologies that emerged with this advent is the RepeatExplorer program, a tool with high potential in the discovery of these sequences. The genus *Stylosanthes* Sw. is part of the Leguminosae family and despite having high economic and environmental value it still has a low genetic characterization, due to the occurrence of hybrid species and high systematic complexity. It is a diverse genre with wide distribution, with Brazil as its main center of origin and diversity. The study sought to understand the organization of the nuclear genome of nine species of the genus *Stylosanthes* Sw., including two allopolyploid complexes of economic interest (Complex *S. scabra* and Complex *S. capitata*), through the genomic analysis of repetitive sequences, seeking to characterize the abundance of these sequences and understand the phylogenomic relationships between diploids and allopolyploids. The genomic reads of the species *S. hamata*, *S. viscosa* and *S. scabra* were obtained from the study by Marques *et al.* (2018), available at NCBI, and for the other species sequencing was done by GenOne Solutions in Biotechnology, Rio de Janeiro, Brazil. The reads of the nine species were inserted into the Geneious software, where they were paired, converted into a fasta format and selected. The obtained sequences were concatenated and inserted in the RepeatExplorer program, being characterized individually and comparatively, through the RepeatExplorer2 version and the TAREAN tool. A phylogenetic analysis was also carried out using the satellite of each species, using the AAF program. The RepeatExplorer analyzes revealed that the nine studied *Stylosanthes* species have more than 50% of their genome formed by repetitive sequences, with the class of transposable elements of the LTR order being the most abundant. The abundance of satellite DNA varied between species, and the grouping of reads in reading charts revealed good conservation of the sequences, representing low diversification in the evolutionary process. Phylogenetic analysis made it possible to assemble a tree with kinship relationships between the enlightened species and supporting previous studies. Unlike what the literature shows us, our study revealed a close relationship between the species *S. hamata* and *S. seabrana*, which may be synonymous in the process of speciation. The differences found between the results and what appears in the literature reinforces the need for more in-depth studies that review the taxonomy and phylogenetics of the group, especially in the molecular scope.

Keywords: repetitive DNA. Allopolyploid. Phylogenomics. Bioinformatics.

LISTA DE FIGURAS

- Figura 1 – Porcentagem genômica da abundância de cada satélite por espécie estudada. 35
- Figura 2 – Número de reads utilizados de cada espécie para a montagem de SatDNA identificados em análise comparativa entre as espécies e quantidade de pb de cada satélite. 36
- Figura 3 – Árvore filogenética inferida a partir das sequências repetitivas de SatDNA identificadas nos genomas de nove espécies de *Stylosanthes* Sw. e de *A. hypogaea*. As cores dos ramos distinguem as espécies com relações filogenéticas mais próximas. 37
- Figura 4 – Mapas de distribuição geográfica potencial das espécies *S. hamata* (A) e *S. seabrana* (B), em áreas supostamente adequadas ambientalmente. 39

LISTA DE TABELAS

Tabela 1 – Lista das espécies estudadas, nível de ploidia e números de referência e acesso.	26
Tabela 2 – Porcentagens de sequências repetitivas nos genomas de nove espécies de <i>Stylosanthes</i> Sw. e de <i>A. hypogaea</i> .	31
Tabela 3 – Proporção (%) de DNA satélite, pela análise no software Geneious e pela ferramenta TAREAN, e quantidade de satélites identificados por espécie.	33

LISTA DE ABREVIATURAS E SIGLAS

ITS	Espaçadores internos transcritos
LTR	Long Terminal Repeats
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
pb	Pares de base
rDNA	DNA ribossomal
RE	Retrotransposons
satDNA	DNA satélite
SRS	Short Read Sequence
SSR	Simple Sequence Repeats
TE	Elementos transponíveis

SUMÁRIO

1	INTRODUÇÃO	12
2	REVISÃO DE LITERATURA	14
2.1	Família Fabaceae: Gênero <i>Stylosanthes</i> – Caracterização	14
2.1.1	Poliploidia	16
2.1.1.1	<i>Ocorrência aloploide no gênero</i>	16
2.2	Genoma Nuclear	18
2.3	Sequências Repetitivas	19
2.4	Sequenciamento de Nova Geração (NGS)	22
2.4.1	O uso da bioinformática em estudos genômicos	24
2.4.1.1	<i>RepeatExplorer</i>	25
3	MATERIAL E MÉTODOS	26
3.1	Espécies estudadas	26
3.2	Obtenção dos <i>reads</i> genômicos	26
3.3	Caracterização das sequências repetitivas	27
3.4	Análises filogenéticas dos DNAs satélites	27
4	RESULTADOS E DISCUSSÃO	29
4.1	Caracterização das sequências repetitivas	29
4.2	DNA satélite	33
4.3	Análises filogenéticas	37
5	CONSIDERAÇÕES FINAIS	40
	REFERÊNCIAS	41
	APÊNDICE A – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. hamata</i>	49
	APÊNDICE B – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. viscosa</i>	50

APÊNDICE C – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. scabra</i>	53
APÊNDICE D – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. capitata</i> (CG)	56
APÊNDICE E – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. pilosa</i>	59
APÊNDICE F – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. macrocephala</i>	63
APÊNDICE G – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. capitata</i> (Acesso 24)	66
APÊNDICE H – SATÉLITE PRESENTE NO DNA NUCLEAR DE <i>S. seabrana</i>	69
APÊNDICE I – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>S. guianensis</i>	70
APÊNDICE J – SATÉLITES PRESENTES NO DNA NUCLEAR DE <i>A. hypogaea</i>	71

1 INTRODUÇÃO

O tamanho do genoma é o principal traço de biodiversidade notável em eucariotos, sendo as plantas o grupo taxonômico que mais se destaca. Esta característica é importante no entendimento da evolução dos genomas vegetais (KELLY *et al.*, 2012; PELLICER *et al.*, 2018). O genoma nuclear corresponde a maior fração de DNA em eucariotos. É notável uma extensa variação estrutural nos genomas nucleares de plantas, tanto no tamanho, número de cromossomos, disposição e número de genes e no número de cópias de genoma por núcleo, podendo corresponder à origem poliploide frequente e amplificação de retrotransposons (BENNETZEN, 2002; KELLOGG; BENNETZEN, 2004). Essa variação também é resultante da presença de DNAs organelares inseridos de forma independente no DNA nuclear (RIBEIRO, 2016).

No genoma nuclear são encontradas as frações repetitivas, que são sequências nucleotídicas repetidas inúmeras vezes. Essas repetições podem ser em *tandem*, estando dispostas em série, ou intercaladas, apresentando-se de maneira dispersa (RICHARD; KERREST; DUJON, 2008). O DNA repetitivo pode corresponder em até 90% do genoma de plantas superiores, tornando tais plantas um excelente grupo para estudar a dinâmica da evolução dos elementos repetitivos (DODSWORTH *et al.*, 2015).

A função das sequências repetitivas dentro do genoma ainda era desconhecida, devido à escassez de estudos sobre as mesmas. Acreditava-se que apenas serviam para aumentar o tamanho do genoma, sendo compostas de repetições não-codificantes e não-transcritas associadas à heterocromatina, e eram chamadas de DNA “lixo” ou “egoísta”. Sabe-se hoje que o DNA repetitivo possui diversas funções, especialmente na regulação e expressão de genes e genomas, além de funções estruturais e organizacionais nos cromossomos (BISCOTTI *et al.*, 2015; GUARRIDO-RAMOS, 2017).

As sequências de DNA repetitivo são importantes ferramentas em estudos evolutivos, pois a capacidade de acúmulo de variações leva a produção de inúmeras cópias durante a evolução. Conhecer essas sequências e os seus possíveis efeitos nos cromossomos ajuda na compreensão da organização e do comportamento evolutivo dos genomas eucarióticos (MEHROTRA *et al.*, 2014; MEHROTRA; GOYAL, 2015). Além do mais, o mapeamento de sequências repetitivas específicas de um genoma pode auxiliar também na diferenciação de genomas distintos em aloploidos (HAN *et al.*, 2005; HEMLEBEN *et al.*, 2007).

Atualmente a obtenção de uma grande quantidade de sequências se tornou mais fácil e com preços mais acessíveis, tornando as análises por sequenciamento de nova geração

fundamentais (MARDIS, 2008). Desde o advento do NGS (*Next Generation Sequencing*) diversas metodologias surgiram para a análise de genomas de espécies ainda não estudadas, possibilitando a descoberta de sequências repetitivas (MARQUES *et al.*, 2015) e o seu uso em abordagens filogenômicas (DODSWORTH, 2015; DODSWORTH *et al.*, 2015). O programa RepeatExplorer se caracteriza como umas dessas metodologias e fornece uma ferramenta de alto potencial na descoberta de sequências repetitivas em genomas não estudados previamente e na análise comparativa de genomas (NOVÁK *et al.*, 2013).

O gênero *Stylosanthes* Sw. pertence a família Leguminosae e apresenta diversas espécies de interesse econômico para composição de pastagens, adubação verde e recuperação de solos degradados. No entanto, sua sistemática complexa e a ocorrência frequente de espécies híbridas (alopoliploides) dificultam a sua caracterização genética e limita os esforços de melhoramento no gênero (STACE; CAMERON, 1984; VANNI, 2017). É altamente diversificado e polimórfico, compreendendo espécies pantropicais, a maioria delas descritas para a América, com dois centros de diversificação um no México e outro no Brasil, sendo o Brasil o principal centro de origem e diversidade do gênero, com cerca de 30 espécies, 12 das quais endêmicas (COSTA; VALLS, 2010; SANTOS-GARCIA *et al.*, 2012; STACE; CAMERON, 1984; VANNI, 2017).

Dentro da subfamília Papilionoideae, *Stylosanthes* apresenta uma estreita relação filogenética com *Arachis* (gênero do amendoim cultivado), formando linhagens irmãs no clado Pterocarpus (tribo Dalbergieae) (CARDOSO *et al.*, 2013). Essa relação permite um amplo campo promissor para a transferência de marcas já estabelecidas no genoma de *Arachis* (genoma inteiramente sequenciado e anotado).

Partindo da hipótese que grande parte do genoma nuclear das espécies em estudo é formada por DNA repetitivo, e que essa fração é importante na resolução de questões evolutivas dentro do gênero, o presente estudo buscou analisar e compreender a organização do genoma nuclear através de análises genômicas de sequências repetitivas em nove espécies de *Stylosanthes*, incluindo dois complexos aloploiploides: complexo *S. scabra* (*S. hamata* + *S. viscosa*) e complexo *S. capitata* (*S. macrocephala* + *S. pilosa*), caracterizando, comparativamente, a abundância das principais classes de DNA repetitivo em *Stylosanthes*, buscando entender as relações de parentesco (filogenômicas) entre diploides e aloploiploides.

2 REVISÃO DE LITERATURA

2.1 Família Fabaceae: Gênero *Stylosanthes* – Caracterização

O Brasil é um país megadiverso e apresenta uma das mais ricas floras do mundo com cerca de 32.360 espécies de plantas vasculares, das quais 18.082 são consideradas endêmicas. Entre os táxons bem representados na flora brasileira se destaca a família Leguminosae (Fabaceae), a terceira maior família de Angiospermas, citada como uma das mais ricas do Nordeste brasileiro (FORZZA; BAUMGRATZ; BICUDO, 2010).

Botanicamente, essa família se caracteriza como árvores, ervas, arbustos ou trepadeiras, com filotaxia principalmente alterna, com folhas geralmente compostas e frutos do tipo legume (LEWIS *et al.*, 2005). Espécies anuais ou perenes, com folhas trifolioladas e inflorescência do tipo espiga. As flores são protegidas por brácteas e bractéolas, com coloração amarela ou alaranjada, podendo apresentar estrias púrpuras, e algumas podem possuir flores brancas (KARIA, 2008). São registrados para o Brasil 210 gêneros e 2.694 espécies, das quais 1.458 são consideradas endêmicas e 190 consideradas raras (LIMA; QUEIROZ; MORIM, 2010; QUEIROZ; CARDOSO, 2009).

Uma característica ecológica importante da família Leguminosae é a simbiose de suas raízes com bactérias nitrificantes, especialmente nos representantes da subfamília Papilionoideae. Por isso, algumas espécies são utilizadas para a melhoria de solos agrícolas (ex: *Stylosanthes*), método conhecido como “adubação verde”. A família apresenta ainda uma grande importância econômica na alimentação com várias espécies cultivadas como soja (*Glycine max*), ervilha (*Pisum sativum*), feijão (*Phaseolus vulgaris*), amendoim (*Arachis hypogaea*) e grão-de-bico (*Cicer arietinum*), sendo também utilizadas como forrageiras a alfafa (*Medicago sativa*) e outros gêneros como *Desmodium* e *Stylosanthes*. Possui ainda grande potencial ornamental e madeireiro, sendo utilizada também na fabricação de espessantes, corantes, resinas, medicamentos e pesticidas (LEWIS *et al.*, 2005).

O gênero *Stylosanthes* Sw. está inserido no clado Pterocarpus, tribo Dalbergieae, que abrange espécies arbóreas, lianas e herbáceas, e é composto por espécies herbáceas, com folhas trifolioladas, estípulas bidentadas com bainhas amplexicaules, inflorescências espiciformes com flores amarelas, com a presença ou não de eixo rudimentar plumoso e frutos do tipo lomento, com um ou dois artículos com estiletos residuais. O gênero possui cerca de 50 espécies descritas e é dividido em duas seções, com base na presença ou na ausência de eixo plumoso na base de flores e/ou frutos e no número de bractéolas. A seção *Stylosanthes* é

caracterizada por apresentar eixo rudimentar na base e três bractéolas, enquanto que a seção *Stylosanthes* não possui eixo rudimentar na base das flores e/ou frutos e apresenta duas bractéolas (CALLES; SCHULTZE-KRAFT, 2010; COSTA; SARTORI; POTT, 2008; LEWIS *et al.*, 2005).

Devido sua alta diversificação e polimorfismo o gênero *Stylosanthes* Sw. compreende espécies pantropicais, a maioria delas descritas para a América, com dois centros de diversificação um no México e outro no Brasil, sendo o Brasil o principal centro de origem e diversidade do gênero, com cerca de 30 espécies, 12 das quais endêmicas (COSTA; VALLS, 2010; SANTOS-GARCIA *et al.*, 2012; STACE; CAMERON, 1984; VANNI, 2017).

As espécies do gênero possuem alta capacidade de fixar nitrogênio e alto teor de proteínas, sendo utilizadas como forrageiras principalmente em consórcio com espécies de gramíneas em pastagens para gado em todo o Brasil. As espécies de *Stylosanthes* são capazes de crescer em regiões muito arenosas, sendo muito resistentes à seca, portanto, servem para recuperação de áreas degradadas (BARCELLOS *et al.*, 2008; SANTOS *et al.*, 2009; STACE; EDYE, 1984). A maioria das espécies é colonizadora, tendo por habitat regiões de baixa precipitação, com solos de baixa fertilidade natural, pobres em cálcio e fósforo e com alto teor de alumínio (MARTUSCELLO *et al.*, 2015).

Na região do Cerrado muitas espécies do gênero estão sendo utilizadas também como cultivar “pioneiro” em solos degradados. Essas características somadas fazem com que *Stylosanthes* seja o gênero com maior número de cultivares dentre as leguminosas tropicais usadas como pastagens em todo o mundo. No Brasil oito cultivares foram liberadas comercialmente no mercado, sendo as espécies *S. guianensis*, *S. capitata* e *S. macrocephala* as mais utilizadas (ANDRADE; ASSIS; SALES, 2010).

Apesar da sua importância econômica, o gênero *Stylosanthes* ainda apresenta pouca caracterização genética e sua taxonomia e sistemática complexa dificulta a delimitação e identificação de espécies. Além do mais, a ocorrência de híbridos poliploides (alopoliploides) no gênero dificulta ainda mais a caracterização das espécies (STACE; CAMERON, 1984; VANNI, 2017). Diferentes centros têm desenvolvido programas de melhoramento genético de *Stylosanthes*, sendo essa uma forrageira com muito potencial, sobretudo para a região Nordeste, devido às suas condições de adaptação a condições áridas (KARIA, 2008; KARIA *et al.*, 2002).

2.1.1 Poliploidia

A poliploidia corresponde à existência de dois ou mais genomas no mesmo núcleo, sendo de comum ocorrência nas plantas e desempenha um importante papel na origem e evolução de plantas silvestres e cultivadas. É considerada a alteração citogenética mais importante na especiação e evolução vegetal. Os indivíduos poliploides podem ser classificados em autopoliploides, originados pela duplicação de um mesmo genoma, e alopoliploides, originados pela duplicação de genomas diferentes, normalmente após um evento de hibridação. Os alopoliploides são mais comuns na natureza que os autopoliploides, desempenhando importante papel na evolução e especiação. Em torno de 40% das espécies cultivadas são poliploides, como alfafa, algodão, batata, batata doce, café, cana de açúcar, fumo, dentre outras, não havendo necessariamente uma relação entre poliploidia e domesticação (SCHIFINO-WITTMANN, 2004).

2.1.1.1 Ocorrência alopoliploide no gênero

As informações cariotípicas para o gênero ainda são incipientes, exceto alguns trabalhos utilizando coloração convencional que revelaram a existência de diploides ($2n = 20$), tetraploides ($2n = 40$) e hexaploides ($2n = 60$) (CAMERON, 1967; STACE; CAMERON, 1984, 1987; VANNI; FERNANDEZ, 2011; VIEIRA; AGUIAR-PERECIN; MARTINS, 1993). Um estudo citotaxonômico de 12 táxons brasileiros revelou que todas eram diploides e com cariótipos muito similares. Os autores observaram pequenas diferenças na morfologia e tamanho dos cromossomos, sugerindo que o gênero deve ter passado por rearranjos cromossômicos, provavelmente inversões pericêntricas, e mudanças no conteúdo de DNA (VIEIRA; AGUIAR-PERECIN; MARTINS, 1993).

Análises filogenéticas baseadas em sequências plastidiais (intron *trnL* e espaçador *trnL-trnF*) revelaram que o gênero *Stylosanthes* é monofilético e dividido em quatro clados (VANDER STAPPEN *et al.*, 1999). Com base na variação intraespecífica nas sequências plastidiais em algumas amostras foi sugerida uma origem alotetraploide para *S. scabra*, *S. ingrata*, *S. sympodialis*, *S. subsericea*, *S. capitata* e *S. fruticosa*, embora ainda não exista clareza de seus parentais. Posteriormente as análises filogenéticas baseadas em sequências nucleares ITS (ITS1–5.8S–ITS2) e marcadores moleculares confirmaram as relações inferidas com dados plastidiais e permitiram identificar o genoma parental de alguns alopoliploides, indicando as espécies diploides *S. viscosa*, *S. hamata*, *S. seabrana*, *S. pilosa* e *S.*

macrocephala como importantes doadoras de genoma no processo de formação desses híbridos (LIU; MUSIAL; THOMAS, 1999; VANDER STAPPEN *et al.*, 2002).

A partir de evidências morfológicas, cariotípicas e moleculares, Maass and Mannetje (2002) descreveram *S. seabrana* como uma nova espécie, enquanto Vanni e Fernandez (2011), com base em características morfológicas e cariotípicas, concluíram que *S. seabrana* deve ser tratada como sinonímia de *S. scabra*. A classificação proposta por Vanni e Fernandez (2011) é atualmente aceita pela GRIN (*Germplasm Resources Information Network*) enquanto que *The Plant List* (TPL) segue o proposto por Maass e Mannetje (2002) (CALLES; SCHULTZE-KRAFT, 2016).

Em estudo realizado por Marques *et al.* (2018), foi evidenciado que o aloploiploide *S. scabra* é um híbrido originado do cruzamento entre os diploides *S. hamata* e *S. viscosa*. Da mesma forma que Liu, Musial e Thomas (1999) evidenciaram que *S. macrocephala* e *S. pilosa* são possíveis doadores de genomas na formação do alotetraploide *S. capitata*. É importante mencionar que tanto *S. macrocephala* como *S. capitata* formam a cultivar Estilosantes Campo Grande, desenvolvida pela Embrapa Gado de Corte no ano 2000, que é uma mistura física de sementes de *Stylosanthes capitata* (correspondendo a 80% da linhagem) e *Stylosanthes macrocephala* (doadora de 20 % da linhagem), originada a partir de espécies sobreviventes numa área de acesso de *Stylosanthes* na fazenda Maracujá em Campo Grande, Mato Grosso do Sul. As sementes se apresentaram vigorosas e bastantes resistentes à antracnose (ANDRADE; ASSIS; SALES, 2010; EMBRAPA, 2000; KARIA, 2008).

A cultivar Estilosantes Campo Grande é muito utilizada para o melhoramento da pastagem e em consórcio com outras espécies. Possui boa adaptação a solos arenosos e com baixa fertilidade. A floração acontece em abril e a maturação dos frutos em maio, logo após ocorre a queda das inflorescências e folhas. Produz sementes em abundância e pode permanecer na pastagem por até dois anos, sendo necessário o manejo e replantio após esse período. Contribui para a fixação de nitrogênio no solo e para a nutrição animal, devido ao alto teor de proteínas, além de apresentar boa resistência à seca (KARIA, 2008; SANTOS, 2014).

Interessantemente, enquanto espécies diploides são encontradas em ambas as seções *Stylosanthes* e *Styposanthes*, espécies poliploides estariam restritas à seção *Styposanthes* e seriam produtos de hibridação interseccional entre diploides, seguido por poliploidização (STACE; CAMERON, 1984; VANNI, 2017). Assim foi sugerido que algumas características morfológicas da seção *Styposanthes* fossem dominantes em relação à seção *Stylosanthes*. No entanto, dados mais recentes de filogenia baseados em ITS parece não sustentar a monofilia

das seções, principalmente devido à posição duvidosa dos aloploidos (VANDER STAPPEN *et al.*, 2002).

A confusão taxonômica e a considerável dificuldade de resolver os problemas do gênero deixa clara a necessidade de estudos mais aprofundados em taxonomia e filogenia, a fim de que se possam esclarecer a origem de novas espécies e a posição que as mesmas ocupam na filogenia do gênero.

2.2 Genoma Nuclear

O tamanho do genoma é o principal traço de biodiversidade notável em eucariotos, variando mais de 64.000 vezes. As plantas terrestres destacam-se dos demais grupos taxonômicos devido à enorme diversidade do tamanho do genoma, onde tal característica é marcante no entendimento da evolução dos genomas de plantas e também pode influenciar as espécies no nível de ecossistema. Os avanços e as melhorias nas novas tecnologias de sequenciamento e análise dos genomas tornam possíveis compreensões mais críticas sobre os mecanismos que sustentam as mudanças no tamanho do genoma, bem como a trajetória evolutiva e a capacidade das plantas responderem às mudanças ambientais (KELLY *et al.*, 2012; PELLICER *et al.*, 2018).

A maior parte do genoma eucariótico está concentrada nos núcleos, onde se encontra dividido em segmentos separados fisicamente, chamados de dupla hélice. O genoma em organismos da mesma espécie tem um número de cromossomos e aspecto característicos. A ploidia (n) do organismo é denominada pelo número de cromossomos no conjunto genômico. Logo, organismos que contêm duas cópias completas do genoma em seus núcleos são denominados diploides, e assim sucessivamente. Em um organismo diploide, os dois membros de um par de cromossomos são denominados cromossomos homólogos, onde as sequências de DNA são praticamente as mesmas, podendo haver uma pequena variação na sequência de nucleotídeos (HESLOP-HARRISON; SCHIMIDT, 2012).

É notável uma extensa variação estrutural nos genomas nucleares de plantas, tanto no tamanho do genoma, número de cromossomo, disposição e número de genes e no número de cópias de genoma por núcleo. Podendo corresponder à origem poliploide frequente e amplificação de retrotransposons (BENNETZEN, 2002; KELLOGG; BENNETZEN, 2004). Essa variação também é resultante da presença de DNAs organelares inseridos de forma independente no DNA nuclear (RIBEIRO, 2016). Em estudo realizado Matsumoto *et al.*

(2005), constatou-se que o genoma nuclear de arroz é composto por 0,20-0,24% inserções de DNA plastidial e 0,18-0,19% de DNA mitocondrial.

A colinearidade em genomas nucleares de espécies estreitamente relacionadas é frequentemente observada, e com o aumento da distância filogenética observam-se mais rearranjos. No entanto, não é possível correlacionar a quantidade de rearranjo e o tempo decorrido, uma vez que a divergência não é perfeita. Os rearranjos do genoma podem alterar os padrões de expressão gênica, ocasionando novas combinações, que dão origem aos híbridos, que podem atuar fortemente na evolução (JUDD *et al.*, 2009; KELLOGG; BENNETZEN, 2004).

Os genomas nucleares nas plantas superiores são relativamente complexos com genes espalhados por múltiplos cromossomos. Além dos genes, os cromossomos de plantas contêm um excesso de DNA repetitivo. As sequências de DNA repetidas centenas ou milhares de vezes no genoma nuclear são abundantes em todos os genomas eucarióticos, podendo representar, na maioria das espécies, mais da metade do conteúdo total do DNA no núcleo da célula. Essas repetições incluem elementos transponíveis e vários tipos de simples repetições em tandem, incluindo DNA satélite e simples repetições de sequência (SSRs). Na maioria das angiospermas, elementos transponíveis, especialmente *long terminal repeats* (LTR) retrotransposons compreendem a grande maioria deste DNA repetitivo (KELLOGG; BENNETZEN, 2004; RIBEIRO, 2016; BISCOTTI; OLMO; HESLOP-HARRISON, 2015).

A variação no conteúdo de DNA nuclear em angiospermas é bastante significativa, sendo ocasionada principalmente pela poliploidia e pelo acúmulo de sequências repetitivas no genoma em combinação com as baixas taxas de remoção de DNA (HIDALGO *et al.*, 2017; PELLICER; FAY; LEITCH, 2010; SCHUBERT; VU, 2016). Diversas são as técnicas moleculares, bioquímicas e de microscopia utilizadas para a análise de genomas, sendo possível verificar seu tamanho, conteúdo de genes, estrutura, organização e composição. Sendo desenvolvidos, cada vez mais, softwares que se destinam a caracterização dos genomas e que possibilitam desvendar a estrutura e funcionamento dos elementos do genoma, tornando possível a compreensão da história evolutiva de grupos e o planejamento de estratégias para manipulação de genomas (HESLOP-HARRISON; SCHIMIDT, 2012).

2.3 Sequências Repetitivas

A existência do DNA repetitivo foi observada pela primeira vez em 1961 por Kit *et al.* e por Seuoka *et al.*, a partir de estudos do DNA genômico de camundongos, através da análise

bandas de DNA que constataram que 10% dessas bandas eram formadas por um componente menor, chamado de “DNA satélite”. O termo “DNA satélite” foi usado pela primeira vez para definir sequências de DNA altamente repetitivas organizadas em arranjos em *tandem* em 1979 por Pech *et al.* e cunhada e popularizada como tal por Singer (GARRIDO-RAMOS, 2017).

Sequências repetitivas são fragmentos de DNA presentes em múltiplas cópias no genoma. Essas sequências estão presentes em todas as plantas superiores e podem representar até 90% do tamanho do genoma em algumas espécies, tornando tais plantas um excelente grupo para estudar a dinâmica da evolução dos elementos repetitivos (DODSWORTH *et al.*, 2015). Em estudo realizado por Gill *et al.* (2014), foi observado que 80% do genoma nuclear do girassol (*Helianthus annuus*) é constituído por sequências repetitivas. Da mesma forma Macas *et al.* (2015) evidenciaram que os retrotransposons do tipo LTR ocupam, sozinhos, 81% do genoma nuclear de *Vicia sylvatica*.

A fração repetitiva é responsável por gerar grandes diferenças entre o tamanho dos genomas, juntamente com os eventos de poliploidização, refletindo na distância evolutiva entre as espécies. Ainda não se tem certeza da origem dessas repetições, mas a característica que as torna úteis é que, em indivíduos diferentes, existem geralmente números diferentes de cópias, servindo como base para estudos que buscam testar os elementos repetitivos como sinal filogenético para evolução de grupos (DODSWORTH *et al.*, 2015; MACAS *et al.*, 2015; MEHROTRA; GOYAL, 2015).

Antes se acreditava que as sequências repetitivas não apresentavam função específica dentro do genoma, apenas servia para aumentar o tamanho do mesmo, sendo compostas de repetições não-codificantes e não-transcritas associadas à heterocromatina sem uso imediato, chamadas de DNA “lixo” ou “egoísta”, e sem papel em qualquer processo biológico. Sabe-se hoje que o DNA repetitivo pode agir na regulação e expressão de genes e genomas, direta ou indiretamente, além da organização de cromossomos, pareamento e segregação, montagem do cinetócoro, no controle do alongamento, tamponamento e replicação de telômeros, na regulação epigenética do estabelecimento e manutenção da heterocromatina, na resposta transcricional durante o estresse (BISCOTTI *et al.*, 2015; GUARRIDO-RAMOS, 2017).

Os elementos repetitivos nos genomas consistem em repetições em série (*tandem repeats*) e elementos móveis intercalados (repetições dispersas). As repetições em *tandem* incluem genes *tandems*, DNA ribossomal (rDNA) e DNA satélite, e no grupo das repetições dispersas estão incluídos todos os transposons, genes de tRNA, e pseudogenes (RICHARD; KERREST; DUJON, 2008).

As sequências de DNA satélite podem ser classificadas em três grupos de acordo com o comprimento da unidade de repetição e tamanho do arranjo: (i) microssatélites (SSR), regiões com repetições de até 10 pb, formando arranjos de 10-100 unidades, (ii) minissatélites com 10-100 pb e (iii) DNA satélite (satDNA) com uma unidade de repetição variável formando arranjos de até 100 Mb e com sequências entre 150-400 pb na maioria das plantas e animais (DODSWORTH *et al.*, 2015; GUARRIDO-RAMOS, 2017).

O DNA satélite (satDNA) é caracterizado pela sua organização genômica em longas matrizes de unidades dispostas em série chamadas monômeros. Os Monômeros são sequências repetidas em *tandem* com mais de 200 nucleotídeos de comprimento que se encontram organizadas em matrizes de centenas ou milhares de cópias, ocupando até vários megabases dentro dos genomas. Essa classe de DNA repetitivo é ainda o principal componente da heterocromatina, podendo formar blocos que aparecem como cromocentros nucleares e bandas cromossômicas (LÓPEZ- FLORES; GARRIDO-RAMOS, 2012; ROBLEDILLO *et al.*, 2018). São encontradas principalmente nas regiões centroméricas e subteloméricas e possuem uma rápida evolução que podem causar barreiras reprodutivas entre organismos e promover a especiação, sendo, portanto muito utilizadas em estudos filogenéticos (GARRIDO-RAMOS, 2017; MEHROTRA; GOYAL, 2015).

Além do DNA satélite, os elementos transponíveis (TEs) também têm sequências reconhecidas pela estrutura característica, que inclui repetições terminais em ambas as extremidades e possuiu capacidade de se mover dentro de um cromossomo. Os elementos transponíveis incluem vários grupos representados especialmente por retrotransposons LTR, não-LTR retrotransposons e DNA transposons. As famílias de TEs podem representar até 10% do DNA em um genoma, e podem ser responsáveis pela repressão ou expressão de genes, de acordo com as condições (WICKER *et al.* 2007).

Os elementos transponíveis (TEs), ou elementos móveis, são fragmentos de DNA que podem se mover de uma parte de um genoma para outra. Eles são categorizados em duas classes com base no recurso estrutural e no padrão de transposição. Os elementos da classe I são conhecidos como retrotransposons e podem se transportar através de um intermediário de RNA e se autorreplicar quando transpostos. Os elementos da classe II, também chamados de transposons de DNA, se movem pelo modo de "cortar e colar" ("*cut-and-paste*"). Os retrotransposons são geralmente os elementos repetitivos mais abundantes nos genomas de plantas, principalmente os de repetição terminal longa (LTR), uma vez que podem aumentar seus números de cópias depois de serem transpostos (GALINDO-GONZÁLEZ *et al.*, 2017; LI *et al.*, 2019).

As sequências de DNA repetitivo são importantes ferramentas em estudos evolutivos, pois a capacidade de acúmulo de variações leva a produção de inúmeras cópias durante a evolução. O DNA repetitivo de um genoma pode se comportar de maneira específica, seja dentro de uma espécie, de um gênero ou de uma família. É possível encontrar ainda sequências não específicas de um determinado grupo (espécie, gênero ou família) presentes em outros grupos. Conhecer essas sequências e os seus possíveis efeitos nos cromossomos ajuda na compreensão da organização e do comportamento evolutivo dos genomas eucarióticos (MEHROTRA *et al.*, 2014; MEHROTRA; GOYAL, 2015).

A fração repetitiva é frequentemente descartada em estudos de bioinformática que analisam os genomas, pois dificulta as análises e montagens, o que leva aos pesquisadores mascararem seus dados, ignorando as mesmas, tanto em estudos com DNA, como com RNA. Estudos cromossômicos têm se mostrado vitais no exame da distribuição e evolução das sequências devido aos desafios da análise de dados de sequência, sendo abertas muitas questões sobre a origem, modo evolutivo e funções que essas sequências podem ter no genoma (BISCOTTI; OLMO; HESLOP-HARRISON, 2015; DODSWORTH *et al.*, 2015).

2.4 Sequenciamento de Nova Geração (NGS)

As pesquisas biológicas tinham uma intensa necessidade de desenvolver ferramentas que permitissem a reconstrução de genomas dos organismos, com a finalidade de conhecer a organização, os padrões de herança e a evolução genômica. A partir de então foram surgindo métodos de sequenciamento de DNA que buscam, cada vez mais, uma produção mais elevada de dados e um custo-benefício melhor. As plataformas de sequenciamento de nova geração são uma alternativa poderosa para estudos de genômica estrutural e funcional (CARVALHO; SILVA, 2010; HEATHER; CHAIN, 2016).

Os métodos de sequenciamento surgiram na década de 70, sendo o método de Sanger o primeiro descrito. Essas tecnologias tornaram-se o princípio da genômica, revolucionando os estudos científicos. O método automatizado de Sanger é considerado uma tecnologia de "primeira geração", que possibilitou a obtenção de importantes informações sobre a caracterização gênica, o tamanho e conteúdo repetitivo do DNA, com uma taxa de erro muito baixa (KOBOLDT, 2013; METZKER, 2010; SANGER; NICKLEN; COULSON, 1997).

Com a necessidade de métodos mais rápidos, baratos e confiáveis surge o sequenciamento de nova geração (NGS), que permite o sequenciamento de DNA em plataformas capazes de gerar informação sobre milhões de pares de bases em uma única

corrida, revolucionando as investigações biológicas (CARVALHO; SILVA, 2010; MARDIS, 2008). As tecnologias de NGS começaram a ser comercializadas em 2005, sendo duas de ampla utilização em todo o mundo: a plataforma 454 FLX da Roche e a Solexa da Illumina, seguidas das plataformas *SOLiD System* (Applied Biosystems), e o *HeliscopeTrue Single Molecule Sequencing* (tSMS), da Helicos (CARVALHO; SILVA, 2010).

O resultado do sequenciamento de nova geração (NGS) é a obtenção de fragmentos de DNA, chamados *reads*, que são utilizados para várias finalidades subsequentes. Tais fragmentos são mais curtos do que aqueles obtidos por pirosequenciamento, na tecnologia *Sanger*, por isso são chamados de “*Short Read Sequence*” (SRS), podendo ocorrer ainda uma variação nesse tamanho de acordo com as especificidades de cada biblioteca, o que constitui um desafio na montagem dos genomas (MARDIS, 2008; METZKER, 2010).

Três bibliotecas principais podem ser construídas com as sequências (*reads*): (i) *reads Single-end*, construídos usando leituras de uma das extremidades da sequência de DNA, e são as leituras mais simples. (ii) *reads paired-end*, obtidos a partir da leitura pareada de cada uma das extremidades das sequências de DNA, logo cada *read* tem seu par complementar. (iii) *reads mate-pair*, constituídos a partir da leitura de fragmentos de DNA de cadeia dupla que são circularizados de modo que as extremidades distantes são fisicamente ligadas e lidas juntas; este tipo de leitura tem potencial para resolver problemas de montagem que envolva sequências de DNA repetitivo (GLENN, 2011).

As pesquisas genéticas e biológicas foram fortemente potencializadas com o uso do NGS, possibilitando o sequenciamento de genomas completos de muitos organismos e estudos comparativos em larga escala, além do desenvolvimento de estudos evolutivos, antes considerados impossíveis. São notáveis as possibilidades que surgiram com esse advento, permitindo ampla investigação biológica, caracterização da diversidade ecológica, e até mesmo a investigação de genomas antigos (MARDIS, 2008; METZKER, 2010).

Grandes são os desafios ainda encontrados em plataformas de sequenciamento, principalmente na genômica de plantas, em relação ao custo para a obtenção das bibliotecas e na quantidade de *reads* obtidos, considerados ainda de baixa cobertura, visto que os genomas de plantas são gigantescos. Espera-se que as novas tecnologias que vêm sendo desenvolvidas superem as limitações ainda encontradas e sejam mais acessíveis aos pesquisadores, possibilitando mais estudos e aumentando a quantidade de genomas de plantas descritos (CARVALHO; SILVA, 2010).

2.4.1 O uso da bioinformática em estudos genômicos

A bioinformática é uma ciência nova e em desenvolvimento que apresenta aspectos multidisciplinares, tendo surgido quando se iniciaram o uso de ferramentas computacionais para pesquisas genéticas, bioquímicas e moleculares. Esta ciência envolve várias linhas de conhecimento e tem avançado junto aos avanços das tecnologias computacionais, se tornando imprescindível para o desenvolvimento da biologia molecular, incluindo montagens de genomas completos. Graças ao desenvolvimento de pesquisas com o uso da bioinformática o banco de dados disponíveis no GenBank aumentou exponencialmente (PROSDOCIMI, 2007).

Segundo Verli (2014) a bioinformática teve início na década de 50 quando Watson e Crick publicaram o estudo sobre a molécula de DNA, surgindo daí as bases moleculares para estudos estruturais da replicação e da tradução do material genético. Até a década de 60 foram publicados outros trabalhos que também serviram de base para o desenvolvimento e disseminação da bioinformática, como o de Linus Pauling e Robert Corey e o de Gopalasamudram N. Ramachandran, que forneceram bases sobre a estrutura tridimensional das proteínas. A partir dessas pesquisas a bioinformática se dividiu em duas vertentes: (i) Tradicional ou Clássica e (ii) Estrutural.

As tecnologias de sequenciamento de nova geração evoluem rapidamente, o que leva ao rápido desenvolvimento da bioinformática, através da criação de diversos algoritmos que possam suprir a demanda e melhorar a qualidade da montagem, anotação e análise de genomas. Existem diversas ferramentas computacionais destinadas aos mais diversos tipos de análises, dentre elas podemos citar o NOVOPlasty, CpGAVAS, Geneious, OrganellarGenomeDRAW (OGDRAW), entre outros (DIERCKXSENS; MARDULYN; SMITS, 2016; VERLI, 2014).

O Geneious é umas das plataformas mais utilizadas, pois além da anotação e montagem de genomas, ela integra as mais variadas ferramentas que permitem análise de sequências de DNA, sequências de aminoácidos, alinhamentos em pares e alinhamentos múltiplos, árvores filogenéticas, estruturas 3D, entre outras, através de *plugins* instalados previamente no software. É uma plataforma ideal para os pesquisadores alavancarem os componentes existentes e integrar seus próprios requisitos específicos para a descoberta, análise e visualização de dados biológicos, pois ter acesso as ferramentas corretas é muito importante para o sucesso da pesquisa (KEARSE *et al.*, 2012).

Mesmo com todos os avanços ocorridos nas últimas décadas ainda são encontradas muitas dificuldades para a montagem de genomas eucarióticos, principalmente os de plantas superiores, por possuírem uma maior complexidade devido ao acúmulo de sequências repetitivas e número de pares de bases. Para contornar essas dificuldades faz-se necessário tomar medidas intermediárias que facilitem a construção de bibliotecas genômicas, como a obtenção de *reads mate-pair*, que facilitam encontrar a localização e quantificar as sequências repetitivas dentro dos genomas (PROSDOCINI, 2007; STAATS; MORAIS; MARGIS, 2014).

2.4.1.1 RepeatExplorer

Dentre as ferramentas de bioinformática foi desenvolvido um pipeline bastante eficiente na identificação de famílias repetitivas no DNA nuclear, mesmo em espécies que não possuem um genoma de referência, chamado RepeatExplorer. Esse pipeline se baseia em similaridades através de gráficos, detectando sequências repetitivas através de *reads* de sequências sobrepostas “*all-to-all*”. Através dos agrupamentos feitos pelo pipeline as semelhanças entre os *reads* são transformadas em gráficos virtuais, onde os *reads* são representados por nós e a semelhança entre eles pelas arestas que se encontram conectadas aos nós (GARRIDO-RAMOS, 2017; NOVÁK *et al.*, 2010).

Como consequência aos avanços Novak *et al.* ajustaram o pipeline RepeatExplorer desenvolvendo o Tandem Repeat Analyzer (TAREAN), que possibilita a identificação de SatDNA. O TAREAN usa *reads* de sequências genômicas de baixa cobertura “*low-pass*” e realiza o agrupamento em gráficos. Como resultado do agrupamento dos *reads* são formados os gráficos, os que possuírem estruturas circulares caracterizam sequências em *tandem*. Os *reads* dos clusters são decompostos em k-mers que serão utilizados para reconstruir os monômeros (sequências nucleotídicas) representantes de cada satélite (GARRIDO-RAMOS, 2017; NOVÁK *et al.*, 2017).

Com o desenvolvimento e avanço das ferramentas computacionais novas oportunidades estão sendo abertas, com a finalidade de descobrir os principais detalhes da evolução do DNA satélite, além da obtenção de insights sobre as diferentes famílias de DNA repetitivo presentes num determinado genoma, além da sua abundância, variabilidade, e funções em diferentes processos genéticos e genômicos, favorecendo o desenvolvimento das pesquisas com filogenia e genômica comparativa (GARRIDO-RAMOS, 2017).

3 MATERIAL E MÉTODOS

3.1 Espécies estudadas

Foram estudadas nove espécies do gênero *Stylosanthes* Sw., incluindo dois complexos (Complexo *S. scabra* e Complexo *S. capitata*), listadas a seguir (Tabela 1). Para motivos de comparação dados disponíveis de *Arachis hypogaea* (SRX7041793) foram utilizados.

Tabela 1 – Lista das espécies estudadas, nível de ploidia e números de referência e acesso.

Nome da espécie	Nível de ploidia	Número de referência	Número de acesso do NCBI
<i>S. hamata</i>	Diploide	2701	SRX3517479
<i>S. viscosa</i>	Diploide	2702	SRX3517481
<i>S. scabra</i>	Tetraploide (2n=40)	2703	SRX3517480
<i>S. capitata</i> (cv. Campo Grande)	Tetraploide (2n=40)	2705	
<i>S. pilosa</i>	Diploide	2706	
<i>S. macrocephala</i> (cv. Campo Grande)	Diploide	2707	
<i>S. capitata</i> (acesso 24)	Tetraploide (2n=40)	2708	
<i>S. seabrana</i>	Diploide	2709	
<i>S. guianensis</i>	Diploide	2710	

Fonte: A autora (2019).

3.2 Obtenção dos reads genômicos

Os reads genômicos das espécies *Stylosanthes hamata* (diploide, $2n = 20$), *S. viscosa* (diploide, $2n = 20$) e *S. scabra* (tetraploide, $2n = 40$) que compõem espécies do complexo *S. scabra*, foram obtidos do estudo de Marques *et al.* (2018), disponíveis no NCBI (Tabela 1). Para as demais espécies, *S. pilosa* 'LC 7833' (diploide, $2n = 20$), *S. macrocephala* (cv Campo Grande – diploide, $2n = 20$), *S. capitata* 'cv. Campo Grande' (tetraploide, $2n = 40$), *S. capitata* (acesso 24), *S. seabrana* e *S. guianensis*, foram coletadas folhas jovens e frescas ($\approx 1-5g$) para extração do DNA com o kit Neoplant II. O DNA extraído foi verificado em gel de agarose a 1% (p/v) e medido com um fotômetro NanoDrop 2000 (Thermo Scientific).

O sequenciamento foi realizado pela GenOne Soluções em Biotecnologia, Rio de Janeiro, Brasil, onde foram utilizados 5,0 µg de gDNA de cada espécie para produção da biblioteca de DNA. A biblioteca de sequenciamento foi gerada utilizando o NEBNextUltra II DNA Library Prep Kit para Illumina (New England Biolabs, Inglaterra) seguindo as recomendações dos fabricantes. O DNA genômico foi aleatoriamente fragmentado até um tamanho de 350 pb por Bioruptor e posteriormente selecionados e ligados com o adaptador de comprimento total e filtrados. Por fim, a biblioteca foi analisada para distribuição de tamanho pelo Agilent2100 Bioanalyzer e quantificada via PCR em tempo real. As bibliotecas foram sequenciadas pelo sequenciador NovaSeq 6000 de alta produtividade da Illumina com *paired-end* ($2 \times 150\text{bp}$).

3.3 Caracterização das sequências repetitivas

Os *reads* das nove espécies foram inseridos no software Geneious, onde foram pareados e convertidos em formato fasta, após convertidos foi feita uma seleção dos *reads* de cada amostra para que todos ficassem do mesmo tamanho, permanecendo 1 milhão de *reads* pareados. As sequências obtidas de cada espécie foram concatenadas e inseridas no programa RepeatExplorer.

Todas as espécies no presente estudo foram analisadas individualmente e comparativamente, incluindo *Arachis hypogaea* (SRX7041793), utilizando a versão RepeatExplorer2, seguindo passos padrões do programa e incluindo um *custom repeat database* (banco de dados de repetição personalizado) enriquecido com *repeats* de *Stylosanthes* para facilitar a anotação e comparação dos *repeats* entre as espécies. Adicionalmente, todas as espécies foram analisadas também no TAREAN, que especifica as sequências repetidas em tandem (DNA satélite). Por fim, os *repeats* foram classificados nas principais classes e a abundância e especificidade comparados entre as espécies.

3.4 Análises filogenéticas dos DNAs satélites

A análise filogenética do conjunto de todos os satDNAs de cada espécie (satelitoma) foi realizado utilizando uma abordagem que compara a similaridade de um conjunto de *reads* sem a necessidade de alinhamento e montagem, utilizando o programa AAF (*alignment and assembly free*) (<https://sourceforge.net/projects/aaf-phylogeny>), seguindo os padrões propostos por Fan *et al.* (2015). O satelitoma de *Arachis hypogaea* foi utilizado como grupo

externo na árvore filogenética. Para isso, os *reads* de cada *cluster* de satDNA identificado foram isolados e agrupados por espécie e por família de satDNA.

4 RESULTADOS E DISCUSSÃO

4.1 Caracterização das sequências repetitivas

As análises do RepeatExplorer para as sequências repetitivas revelaram que as nove espécies de *Stylosanthes* estudadas possuem mais de 50% do seu genoma formado por sequências repetitivas (Tabela 2), que está de acordo com a literatura, que diz que grande parte do genoma de eucariotos é composta por sequências repetitivas (BISCOTTI; OLMO; HESLOP-HARRISON, 2015). Foram identificadas 20 famílias de elementos repetitivos, pertencentes às classes de retrotransposons e DNA transposons, além de sequências de rDNA (5S e 45S) e DNA satélite. A espécie que apresentou maior riqueza de famílias foi *S. macrocephala* (2707), com um total de 17, e a espécie com menor abundância de famílias foi *S. guianensis* (2710), apresentando sete famílias. A quantidade de famílias presentes em cada espécie pode estar relacionada ao tamanho dos genomas. A abundância de rDNA variou entre 0,01 – 3,55%, sendo o maior valor correspondente ao 45S_rDNA, na espécie *S. macrocephala* (2707).

A família Athila (LTR – Ty3/Gypsy) representa, em todos os genomas, a maior proporção dentre as famílias de DNA repetitivo observadas (Tabela 2), com valores acima de 24%, fazendo com que o somatório com as demais famílias torne a classe de retrotransposons LTR – Ty3/Gypsy a mais abundante em todos os genomas. A segunda maior porcentagem observada em todos os genomas foi da família SIRE, também pertencente a classe dos retrotransposons da ordem LTR, porém à superfamília Ty1/Copia (Tabela 2).

Algumas famílias foram de ocorrências específicas, como a família Ikeros (LTR – Ty1/Copia), que está presente apenas no genoma de *S. macrocephala* e também no genoma de referência *A. hypogaea*, porém apresentou baixa abundância em *S. macrocephala*. Interessantemente, a família PIF_Harbinger (DNA Transposon) foi detectada apenas nas espécies *S. pilosa*, *S. macrocephala* e *S. capitata* (acesso 24) e em *A. hypogaea*, o que pode corresponder a uma especificidade dentro do grupo com relações mais estreitas de parentesco, que se mantém conservada durante a evolução. A abundância de DNA satélite variou de espécie para espécie, sendo os maiores valores observados em *S. hamata* e *S. scabra*, que apresentaram 3,26% e 4,83%, respectivamente (Tabela 2).

Pode-se observar que a classe de elementos transponíveis (retrotransposons) da ordem LTR é a mais abundante entre as sequências repetitivas dos genomas, evidenciando os estudos de Galindo-González *et al.* (2017) e Li *et al.* (2019) que mostram que os TEs da ordem LTR

são os mais abundantes nos genomas de plantas. As análises também mostraram que a menor fração corresponde a ordem dos Não-LTR. Em todos os genomas foi possível observar ainda que o programa não conseguiu classificar uma fração de elementos repetitivos entre 5 – 18% do total (Tabela 2), provavelmente por não possuírem os domínios necessários para a correta classificação.

Tabela 2 – Porcentagens (%) de sequências repetitivas nos genomas de nove espécies de *Stylosanthes* Sw. e de *A. hypogaea*.

Classes	Famílias	Genoma (%)										
		<i>S. hamata</i> (2701)	<i>S. viscosa</i> (2702)	<i>S. scabra</i> (2703)	<i>S. capitata</i> (2705)	<i>S. pilosa</i> (2706)	<i>S. macrocephala</i> (2707)	<i>S. capitata</i> 24 (2708)	<i>S. seabrana</i> (2709)	<i>S. guianensis</i> (2710)	<i>A. hypogaea</i>	
LTR Ty1/Copia	SIRE	14,950	9,882	13,880	11,680	5,084	1,564	4,871	12,250	0,275	1,097	
	Ikeros						0,022				0,356	
	Bianca	1,208		2,418	1,938	1,133	0,449	1,679	0,750	0,027	0,541	
	Ale		0,138	0,048	0,187	0,065	0,192	0,024			0,159	
	TAR		0,042				0,138				0,022	
	Ivana		0,015	0,157	0,222	0,021	0,115	0,049				
	Tork	0,204	0,756	0,495	0,757	0,575	1,100	0,398	0,276		0,185	
	Total		16,362	10,833	16,998	14,784	6,878	3,58	7,021	13,276	0,302	2,36
LTR Ty3/Gypsy	não-chromovirus	Athila	39,450	31,890	31,730	31,780	24,750	27,220	34,670	37,940	35,220	46,170
		Ogre	0,294	0,106	0,310	0,728	0,397	0,177	0,028	0,088		
		Retand	0,286	1,904	0,401	1,303	1,959	3,021	1,573	0,433	1,424	2,929
	Chromovirus	Tekay	0,151	4,999	0,738	4,002	9,799	6,794	5,919	0,194	0,920	3,565
		Galadriel	0,085		0,128	0,147		0,056	0,018	0,087		0,052
		CRM	0,014			0,118	0,091		0,092	0,037		
	Total		40,28	38,899	33,307	38,078	36,996	37,268	42,3	38,779	37,564	52,716
	LTR não classificado	LTR		2,112								
Ty1_copia								0,016				
Não-LTR	pararetrovirus		0,264		0,112	0,119		0,055			0,299	
	LINE	0,033	0,311	0,422	0,730	0,107	0,367	0,378			0,369	
Total		56,675	52,419	50,727	53,704	44,1	41,215	49,77	52,055	37,866	55,744	
DNA transposons	MuDR_Mutator	0,255	1,400	0,333	0,811	1,148	1,393	0,909	0,155	0,744	0,379	
	EnSpm_CACTA	0,206	0,618	0,119	0,555	0,548	0,375	0,276	0,419	1,182		
	PIF_Harbinger					0,030	0,025	0,016			0,014	
	hAT						0,041					
	Helitron		0,030		0,025						0,039	
DNA transposons não classificado	Classe_1	0,536	0,285	0,617	0,420	0,558	0,959	0,319	0,520	0,033		

Total		0,997	2,333	1,069	1,811	2,284	2,793	1,52	1,094	1,959	0,432
rDNA	45S_rDNA	0,426	0,666	0,312	0,706	0,981	3,551	1,070	1,236	0,658	0,805
	5S_rDNA	0,018	0,051	0,034	0,093	0,152	0,185	0,107	0,064	0,139	0,270
DNA Satélite	Satélite	3,269	1,038	4,835	0,450	0,852	1,079	0,931	0,173	0,693	1,484
Não classificado	NA	5,459	6,002	11,360	10,290	5,464	12,660	11,180	10,010	17,730	8,778
Total de Repeats		66,844	62,509	68,337	67,054	53,833	61,483	64,578	64,632	59,045	67,513

Fonte: A autora (2019).

4.2 DNA satélite

Como já mencionado, a abundância de DNA satélite variou bastante entre as espécies (0,17 – 4,8), porém a quantidade de satélites identificados não possui relação com a porcentagem representada, uma vez que *S. hamata*, por exemplo, apresentou apenas dois satélites e possui a segunda maior porcentagem de SatDNA em seu genoma, como nos mostra a tabela 3.

Tabela 3 – Proporção (%) de DNA satélite, pela análise no software Geneious e pela ferramenta TAREAN, e quantidade de satélites identificados por espécie.

Espécies	Geneious	TAREAN	Satélites
<i>S. hamata</i> (2701)	3,2692	2,2837	2
<i>S. viscosa</i> (2702)	1,0376	0,7720	8
<i>S. scabra</i> (2703)	4,8348	2,8874	7
<i>S. capitata</i> (2705)	0,4495	1,5657	7
<i>S. pilosa</i> (2706)	0,8519	0,7539	10
<i>S. macrocephala</i> (2707)	1,0791	1,1280	8
<i>S. capitata</i> 24 (2708)	0,9305	0,6832	7
<i>S. seabrana</i> (2709)	0,1730	0,5227	1
<i>S. guianensis</i> (2710)	0,6930	0,2749	3
<i>A. hypogaea</i>	1,4838	1,0388	5

Fonte: A autora, (2019).

Em análise mais detalhada feita a partir da ferramenta TAREAN no programa RepeatExplorer foi possível obter a quantidade de *reads* utilizados para a montagem de cada satélite nas análises individuais e na análise comparativa entre as espécies. A figura 1 nos mostra a porcentagem que cada satélite representa em cada espécie, sendo possível observar ainda, que o tamanho dos monômeros não implica necessariamente na proporção que o satélite ocupa no genoma, uma vez que satélites com monômeros menores podem representar uma porcentagem maior no genoma, e vice-versa, como é o caso dos *Clusters* 3 e 157, no genoma de *S. hamata*, onde o *Cluster* 3, com monômero de 128 pb ocupa mais que 2,2% do genoma, enquanto que o *Cluster* 157, que possuiu um monômero de 225 pb, ocupa menos que 0,1%.

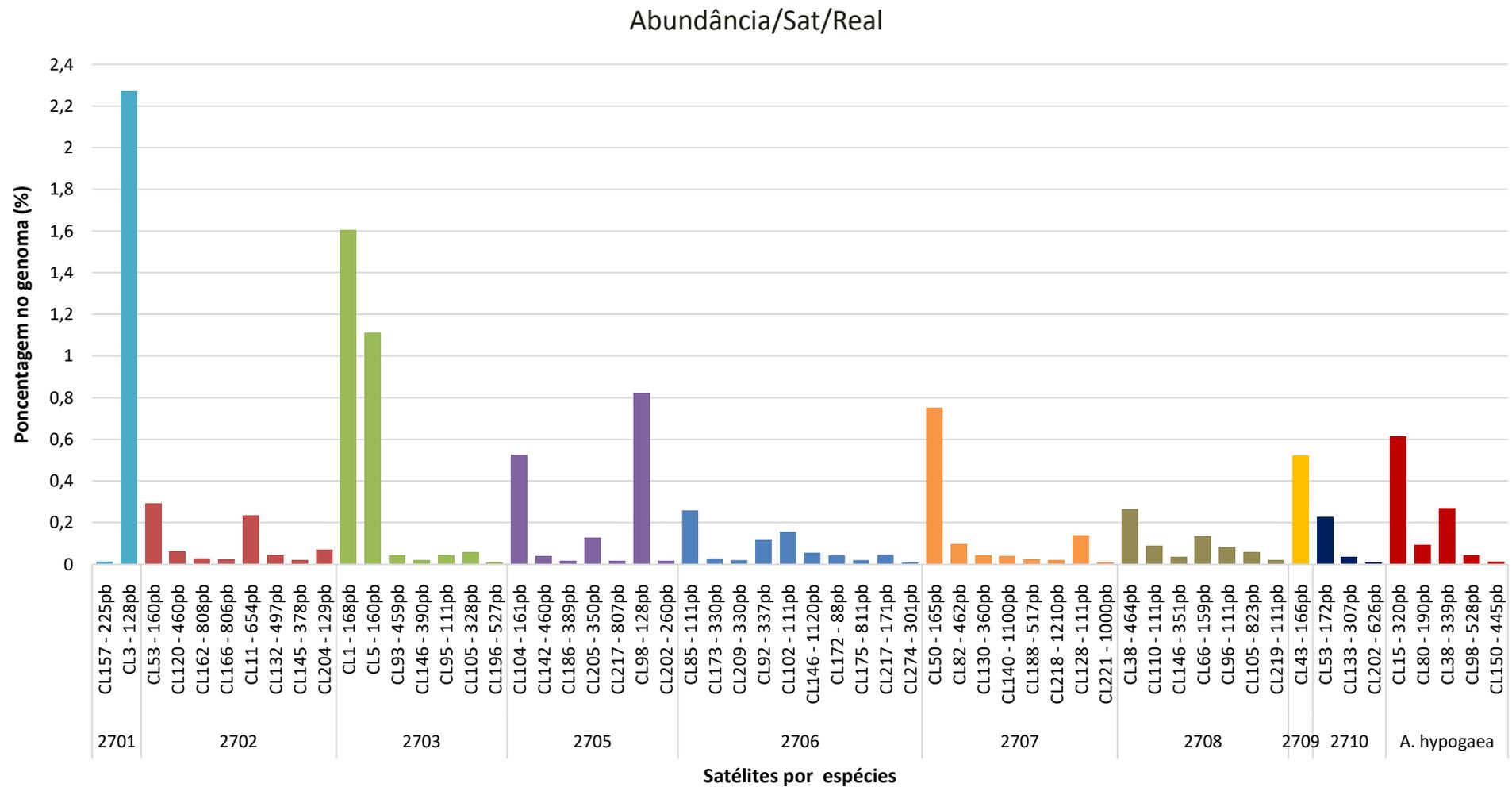
A partir dos gráficos e das reconstruções dos K-mers resultantes do agrupamento dos *reads* (Apêndices) notou-se que no geral as espécies possuem uma boa conservação das

sequências, nos levando a crer que não houve muita diversificação durante o processo evolutivo, até o momento.

Foi observado que, assim como na análise pelo RepeatExplorer2, *S. scabra* apresenta a maior abundância de DNA satélite em seu conteúdo genômico, seguido por *S. hamata*, enquanto que *S. guianensis* possui o menor percentual (Tabela 3). Na espécie *S. pilosa* foi possível identificar 10 prováveis satélites (sendo a maior abundância), e as menores incidências de satélites foram observadas nas espécies *S. seabrana*, *S. hamata* e *S. guianensis*, que apresentaram um, dois e três satélites, respectivamente (Tabela 3).

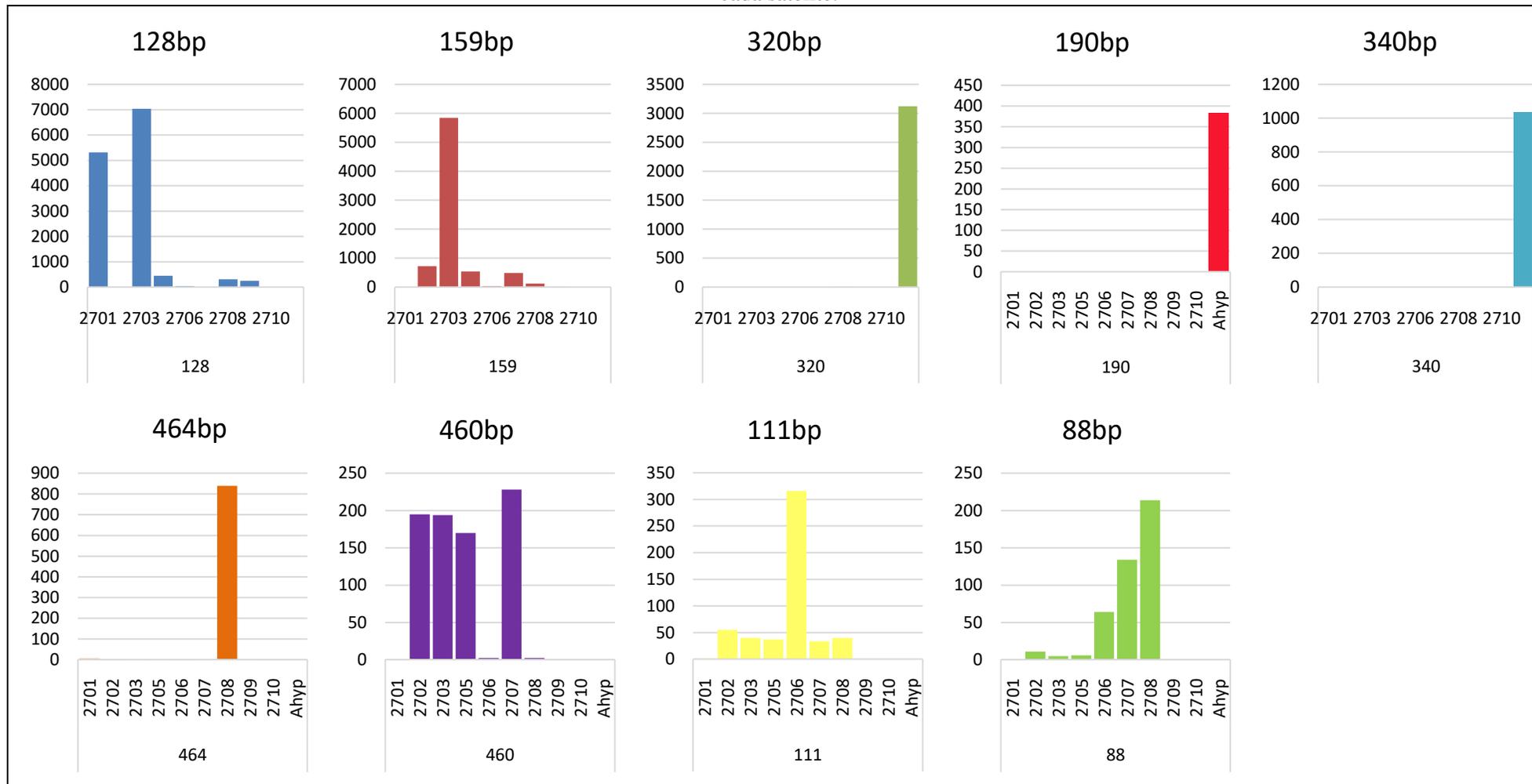
Em análise comparativa entre as espécies foram identificados 9 satélites de DNA, sendo que os satélites com 320, 340 e 190 pb são formados apenas por *reads* do genoma de referência *A. hypogaea*, os demais se mostram diversos no tamanho dos monômeros e na composição dos *reads* de cada espécie, na figura 2 podemos observar a quantidade de *reads* utilizados de cada espécie para a montagem de cada satélite e o tamanho dos monômeros (pb). O satélite com 464 pb foi formado apenas por *reads* de *S. capitata* (acesso 24), e o satélite com 88 pb se mostrou mais abundante nas espécies que compõem o Complexo *S. capitata*. Entre os demais satélites não foi observada especificidade.

Figura 1 – Porcentagem genômica da abundância de cada satélite por espécie estudada.



Fonte: Gráfico resultante da análise de dados de saída do programa RepeatExplorer (2019).

Figura 2 – Número de *reads* utilizados de cada espécie para a montagem de SatDNA identificados em análise comparativa entre as espécies e quantidade de pb de cada satélite.

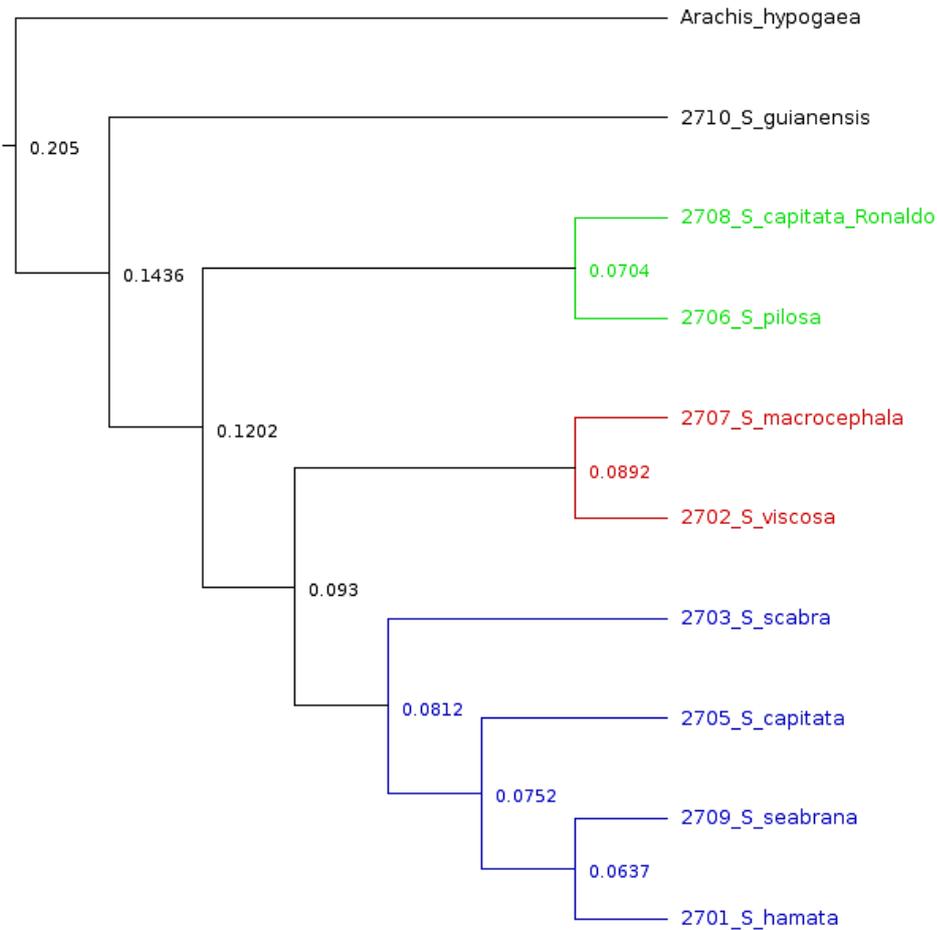


Fonte: Gráficos resultantes das análises de dados de saída do programa RepeatExplorer (2019).

4.3 Análises filogenéticas

A análise filogenética feita a partir das sequências repetitivas de DNA satélite nos revelou uma árvore que sustenta estudos anteriores (Figura 3), como o de Marques *et al.* (2018), onde a posição dos ramos revela uma estreita relação de parentesco entre as espécies *S. hamata* e *S. scabra*, sendo a primeira possível doadora do genoma materno do híbrido *S. scabra*. O estudo realizado por Liu, Musial e Thomas (1999) supunha que as espécies diploides *S. pilosa* e *S. macrocephala* eram possíveis doadoras de genoma na formação do híbrido alotetraploide *S. capitata*, de fato nossas análises resultaram em uma estreita relação entre as espécies *S. pilosa* e *S. capitata* (Acesso 24/Ronaldo), nos levando a acreditar que *S. pilosa* é possível doadora do genoma materno na formação do híbrido *S. capitata*.

Figura 3 – Árvore filogenética inferida a partir das sequências repetitivas de SatDNA identificadas nos genomas de nove espécies de *Stylosanthes* Sw. e de *A. hypogaea*. As cores dos ramos distinguem as espécies com relações filogenéticas mais próximas.



Fonte: A autora (2019).

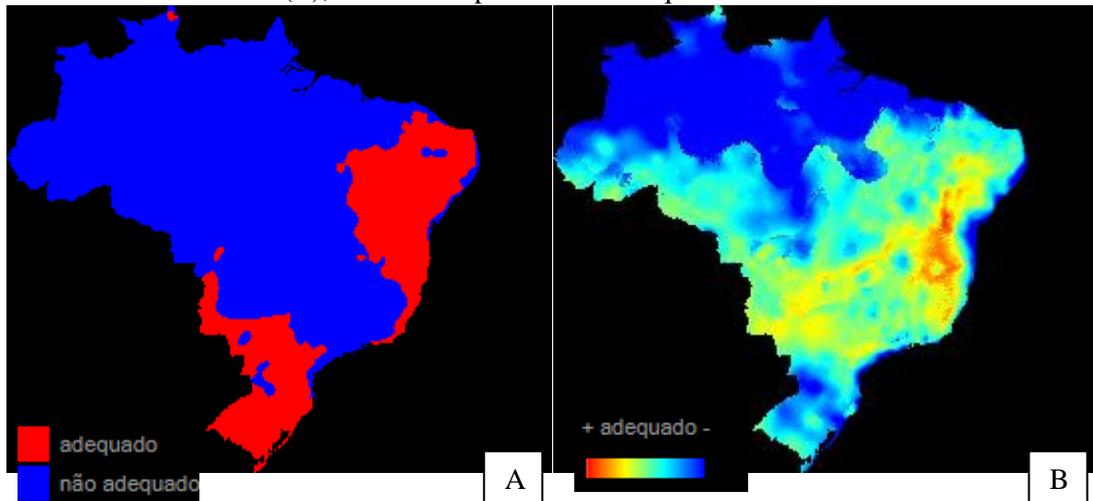
Na árvore, é possível observar também, que apesar de nossos resultados apontarem que *S. pilosa* é doadora do genoma para a formação do híbrido *S. capitata* (Acesso 24/Ronaldo), elas aparecem no mesmo ramo. Esse fato pode ser explicado pela migração das sequências, evento chamado de “Evolução em Concerto”, onde ocorre a transferência horizontal de genes, gerando homogeneidade intraespecífica. Esse fenômeno se caracteriza como uma propriedade na evolução das sequências repetitivas (LIAO, 1999).

Um dado relevante em nossas análises é a diferença da constituição genômica observada entre os dois acessos de *S. capitata*, onde um (*S. capitata* CG) mostrou-se bastante similar ao genoma de *S. scabra*, e o outro (*S. capitata* acesso 24/Ronaldo) revelou características que sustentam estudos anteriores, como o de Liu, Musial e Thomas (1999), onde *S. pilosa* e *S. macrocephala* são possíveis doadoras de genoma na formação desse híbrido. Essa diferença nos leva a acreditar que o primeiro acesso (*S. capitata* CG – 2705) foi erroneamente classificado na separação de sementes para a composição da cultivar, uma vez que a planta de onde foi realizada a extração do DNA dessa espécie foi oriunda de sementes da Cultivar Estilosantes Campo Grande, que consiste em uma mistura física de sementes.

Interessantemente, nossos estudos revelam uma estreita relação de parentesco entre as espécies *S. hamata* e *S. seabrana* (Figura 3), que mesmo apresentando número e abundância de *repeats* diferentes, possuem genomas e características morfológicas bastante similares, nos levando a acreditar que essas espécies são sinonímias botânicas, porém tal resultado difere da literatura, onde temos o estudo de Maass e Manetteje (2002) que diz que *S. seabrana* é uma nova espécie, e o estudo de Vanni e Fernandez (2011) que coloca *S. seabrana* como uma sinonímia botânica de *S. scabra*.

Nossos resultados nos levam a duas hipóteses sobre a espécie *S. seabrana*: a primeira é que a ideia de Vanni e Fernandez (2011) não pode ser sustentada, uma vez que *S. scabra* é um híbrido tetraploide, logo não pode ser sinonímia de *S. seabrana* (diploide), e também porque *S. seabrana* possui maior similaridade com *S. hamata*, como já mencionado. A segunda hipótese é que *S. hamata* e *S. seabrana* se tratariam de sinonímias em processo de especiação, onde já acumulam diferenças moleculares, mas ainda compartilham muitas características morfológicas. Na figura 4 é possível observar a distribuição geográfica potencial das espécies *S. hamata* e *S. seabrana*, onde nota-se que a distribuição de *S. hamata* é mais limitada, e que *S. seabrana* também está distribuída nas mesmas regiões, mas com uma amplitude de áreas adequadas maior, o que pode corresponder ao processo de especiação.

Figura 4 – Mapas de distribuição geográfica potencial das espécies *S. hamata* (A) e *S. seabrana* (B), em áreas supostamente adequadas ambientalmente.



Fonte: Lima, I. B. (2014). Biogeografia da Flora e dos Fungos do Brasil. INCT Herbário Virtual. Disponível em: <http://biogeo.inct.florabrasil.net>. Acesso em: 02 mar. 2020.

5 CONSIDERAÇÕES FINAIS

Diante dos resultados observados é notável que as nove espécies do gênero *Stylosanthes* Sw. caracterizadas nesse estudo possuem alta abundância de DNA repetitivo e mantêm a conservação de muitas sequências, não apresentando grande diversificação durante o processo evolutivo. Se tratando do DNA satélite é visível a baixa especificidade das sequências, sendo maior apenas entre as espécies com alto grau de parentesco, no caso entre os híbridos e seus doadores de genoma. Filogeneticamente, as espécies estudadas, em especial as que compõem os dois complexos (Complexo *S. capitata* e Complexo *S. scabra*), ocupam posições quase esclarecidas, sendo necessários mais estudos filogenéticos que visem esclarecer as posições das demais espécies.

Com relação a espécie *S. capitata* (CG) que compõe a Cultivar Estilosantes Campo Grande, juntamente com a espécie *S. macrocephala*, e é comercializada pela Embrapa Gado de Corte desde o ano de 2000, sugere-se que seja feita uma revisão taxonômica, uma vez que nossos estudos apontam que a mesma é mais similar geneticamente com *S. scabra*. Essa revisão taxonômica é necessária para que os consumidores tenham certeza do produto que estão adquirindo, e também para que a empresa possa adequar as recomendações para a Cultivar, de acordo com as espécies que realmente a compõe.

A divergência entre alguns dos resultados obtidos e aqueles que constam na literatura, como no caso de *S. seabrana* que não se pode afirmar se a mesma é uma espécie distinta ou uma sinonímia de *S. hamata* em processo de especiação, e também sobre caso da espécie *S. capitata* (CG) citado anteriormente, confirma a complexidade e dificuldade de estudos do gênero, e reforça a necessidade de estudos mais aprofundados que revisem a taxonomia e filogenética do grupo, principalmente no âmbito molecular.

REFERÊNCIAS

- ANDRADE, C.M.S.; de ASSIS, G.M.L.; SALES, M.F.L. Estilosantes Campo Grande: leguminosa forrageira recomendada para solos arenosos do Acre. EMBRAPA. **Circular Técnica**, n. 55, 2010. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/879062/estilosantes-campo-grande-leguminosa-forrageira-recomendada-para-solos-arenosos-do-acre>. Acesso em: 23 maio 2018.
- BARCELLOS, A.O. *et al.* Sustentabilidade da produção animal baseada em pastagens consorciadas e no emprego de leguminosas exclusivas, na forma de banco de proteína, nos trópicos brasileiros. **Revista Brasileira de Zootecnia**, Viçosa, MG, n. 37, p. 51-67, 2008. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-35982008001300008. Acesso em: 26 abr. 2018.
- BENNETZEN J. L. Opening the door to comparative plant biology. **Science**, [s. l.], v. 296, p. 60-63, 2002. Disponível em: <https://science.sciencemag.org/content/296/5565/60.summary>. Acesso em: 08 fev. 2019.
- BISCOTTI, M. A.; OLMO, E.; HESLOP-HARRISON, J. S. Repetitive DNA in eukaryotic genomes. **Chromosome Research**, [on-line], v. 23, n. 3, p. 415-420, 2015. Disponível em: <https://link.springer.com/article/10.1007%2Fs10577-015-9499-z>. Acesso em: 12 ago. 2019.
- BISCOTTI, M. A. *et al.* Transcription of tandemly repetitive DNA: functional roles. **Chromosome Research**, [on-line], v. 23, p. 463. 2015. Disponível em: <https://link.springer.com/article/10.1007%2Fs10577-015-9494-4>. Acesso em: 12 ago. 2019.
- CALLES, T.; SCHULTZE-KRAFT, R. New species, nomenclatural changes and recent taxonomic studies in the genus *Stylosanthes* (Leguminosae): An update. **Tropical Grasslands-Forrajões Tropicais**, [on-line], v. 2, p. 122-128, 2016. Disponível em: <http://www.tropicalgrasslands.info/index.php/tgft/article/view/283>. Acesso em: 05 dez. 2018.
- CALLES, T.; SCHULTZE-KRAFT, R. *Stylosanthes* (Leguminosae, *Dalbergieae*) of Venezuela. **Willdenowia**, Berlim, v. 40, p. 305-329, 2010. Disponível em: https://www.researchgate.net/publication/261779414_Stylosanthes_Leguminosae_Dalbergieae_of_Venezuela. Acesso em: 15 set. 2018.
- CAMERON, D.F. Chromosome number and morphology of some introduced *Stylosanthes* species. **Australian Journal of Agricultural Research**, [s. l.], v. 18, p. 375, 1967. Disponível em: <https://www.publish.csiro.au/cp/AR9670375>. Acesso em: 18 set. 2018.
- CARDOSO, D. *et al.* Reconstructing the deep-branching relationships of the papilionoid legumes. **South African Journal of Botany**, [on-line], v. 89, p. 58-75, 2013. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0254629913002585>. Acesso em: 10 maio 2018.
- CARVALHO, M.C. da C.G. de; SILVA, D.C.C.G. da. Sequenciamento de DNA de nova geração e suas aplicações na genômica de plantas. **Ciência Rural**, Santa Maria, v. 40, n. 3, p. 735-744, 2010. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-84782010000300040. Acesso em: 15 ago. 2019.

COSTA, L. C.; SARTORI, A.L.B.; POTT, A. Estudo taxonômico de *Stylosanthes* (Leguminosae- Papilionoideae-Dalbergieae) em Mato Grosso do Sul, Brasil. **Rodriguésia**, Rio de Janeiro, v. 59, p. 547-572, 2008. Disponível em:

[http://www.scielo.br/scielo.php?pid=S2175-](http://www.scielo.br/scielo.php?pid=S2175-78602008000300547&script=sci_abstract&tlng=pt)

[78602008000300547&script=sci_abstract&tlng=pt](http://www.scielo.br/scielo.php?pid=S2175-78602008000300547&script=sci_abstract&tlng=pt). Acesso em: 28 dez. 2018.

COSTA, L.C.; VALLS, J.F.M. *Stylosanthes* Sw. In: FORZZA, R. C. *et al.* (org.). **Catálogo de plantas e fungos do Brasil**. Rio de Janeiro: Andréia Jakobsson Estúdio: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, 2010. p. 831. Disponível em: reflora.jbrj.gov.br/downloads/vol1.pdf. Acesso em: 22 abr. 2018.

DIERCKXSENS, N.; MARDULYN, P.; SMITS, G. NOVOPlasty: *de novo* assembly of organelle genomes from whole genome data. **Nucleic Acids Research**, [on-line], v. 45, n. 4, 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5389512/>. Acesso em: 02 fev. 2019.

DODSWORTH, S. *et al.* Genomic Repeat Abundances Contain Phylogenetic Signal. **Systematic Biology**, [on-line], v. 64, n. 1, p. 112-126, 2015. Disponível em: <https://academic.oup.com/sysbio/article/64/1/112/2848261>. Acesso em: 20 maio 2019.

DODSWORTH, S. Genome skimming for next-generation biodiversity analysis. **Trends Plant Science**, [on-line], v. 20, p. 525-7, 2015. Disponível em: [https://www.cell.com/trends/plant-science/fulltext/S1360-1385\(15\)00176-4](https://www.cell.com/trends/plant-science/fulltext/S1360-1385(15)00176-4). Acesso em: 11 maio 2019.

EMBRAPA. Estilosantes Campo Grande: estabelecimento, manejo e produção animal. **Comunicado técnico**, n. 61. Campo Grande: Embrapa Gado de Corte, 2000. Disponível em: <https://www.embrapa.br/busca-de-publicacoes/-/publicacao/323701/estilosantes-campo-grande-estabelecimento-manejo-e-producao-animal>. Acesso em: 04 jun. 2018.

FAN, H. *et al.* An assembly and alignment-free method of phylogeny reconstruction from next-generation sequencing data. **BMC Genomics**, [on-line], v. 16, p. 522. Disponível em: <https://bmcbgenomics.biomedcentral.com/articles/10.1186/s12864-015-1647-5>. Acesso em: 22 fev. 2020.

FORZZA, R. C. *et al.* Síntese da diversidade brasileira. In: FORZZA, R. C. *et al.* (org.). **Catálogo de plantas e fungos do Brasil**. Rio de Janeiro: Andréia Jakobsson Estúdio: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, 2010. p. 871. Disponível em: reflora.jbrj.gov.br/downloads/vol1.pdf. Acesso em: 22 abr. 2018.

GALINDO-GONZÁLEZ, L. *et al.* LTR-retrotransposons in plants: engines of evolution. **Gene**, [s. l.], v. 626, p. 14-25, 2017. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0378111917303220?via%3Dihub>. Acesso em: 19 mar. 2019.

GARRIDO-RAMOS, M. Satellite DNA: An evolving topic. **Genes**, [on-line], v. 8, p. 230, 2017. Disponível em: <https://www.mdpi.com/2073-4425/8/9/230>. Acesso em: 10 jun. 2019.

GILL, N. *et al.* Sequence-based analysis of structural organization and composition of the cultivated sunflower (*Helianthus annuus* L.) genome. **Biology**, [on-line], v. 3, n. 2, p. 295-319, 2014. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4085609/>. Acesso em: 16 ago. 2019.

GLENN, T.C. Field guide to next-generation DNA sequencers. **Molecular Ecology Resources**, [on-line], v. 11, n. 5, p. 759-769, 2011. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-0998.2011.03024.x>. Acesso em: 14 ago. 2019.

HAN, F.P. *et al.* Rapid and repeatable elimination of a parental genome-specific DNA repeat (pGcIR-1a) in newly synthesized wheat allopolyploids. **Genetics**, [on-line], v. 170, p. 1239-1245, 2005. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1451191/>. Acesso em: 03 jul. 2018.

HEATHER, J. M.; CHAIN, B. The sequence of sequencers: the history of sequencing DNA, **Genomics**, [on-line], v. 107, n. 1, p. 1-8, 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4727787/>. Acesso em: 14 ago. 2019.

HEMLEBEN, V. *et al.* Plant highly repeated satellite DNA: molecular evolution, distribution and use for identification of hybrids. **Systematics and Biodiversity**. [on-line], v. 5, p. 277-289, 2007. Disponível em: <https://www.tandfonline.com/doi/abs/10.1017/S147720000700240X>. Acesso em: 11 maio 2018.

HESLOP-HARRISON, J. S.; SCHMIDT, T. Plant nuclear genome composition. **eLS**, [on-line], p. 1-9, 2012. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470015902.a0002014.pub2>. Acesso em: 12 ago. 2019.

HIDALGO, O. *et al.* Genomic gigantism in the whisk-fern family (Psilotaceae): *Tmesipteris obliqua* challenges record holder *Paris japonica*. **Botanical Journal of the Linnean Society**, [s. l.], v. 183, p. 509-514, 2017. Disponível em: <https://academic.oup.com/botlinnean/article/183/4/509/3739733>. Acesso em: 12 ago. 2019.

JUDD, W. S. *et al.* **Sistemática vegetal: um enfoque filogenético**. Tradução de André Olmos Simões *et al.* 3. ed. São Paulo: Artmed, 2009. 632p.

KARIA, C.T. *et al.* Caracterização morfológica de acessos do gênero *Stylosanthes* do Banco Ativo de Germoplasma da Embrapa Cerrados – Coleção 1994/1995. **Embrapa Cerrados**. Planaltina, DF, 2002. Disponível em: <https://www.infoteca.cnptia.embrapa.br/bitstream/doc/567951/1/bolpd72.pdf>. Acesso em: 13 abr. 2018.

KARIA, C.T. **Genetic and morpho-agronomic characterization of *Stylosanthes guianensis* (Aubl.) Sw. germplasm**. 2008. 140f. Tese (Doutorado em Agronomia) - Universidade Federal de Goiás, Departamento de Ciências Agrárias, Goiânia, 2008. Disponível em: <https://repositorio.bc.ufg.br/tede/handle/tde/436>. Acesso em: 29 mar. 2018.

KEARSE, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. **Bioinformatics**, [on-line], v. 28, n. 12, p. 1647-1649, 2012. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3371832/pdf/bts199.pdf>. Acesso em: 04 abr. 2019.

KELLOGG, E. E. A.; BENNETZEN, J. J. L. The evolution of nuclear genome structure in seed plants. **American Journal of Botany**, [on-line], v. 91, n. 10, p. 1709-1725, 2004. Disponível em: <https://bsapubs.onlinelibrary.wiley.com/doi/pdf/10.3732/ajb.91.10.1709>. Acesso em: 02 mar. 2019.

KELLY, L. J. *et al.* Why size really matters when sequencing plant genomes. **Plant Ecology & Diversity**, [on-line], v. 5, p. 415-425, 2012. Disponível em: https://www.researchgate.net/publication/239793382_Why_size_really_matters_when_sequencing_plant_genomes. Acesso em: 02 fev. 2019.

KOBOLDT, D.C. The next-generation sequencing revolution and its impact on genomics. **Cell**, [on-line], v. 155, n. 1, p. 27–38, 2013. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0092867413011410>. Acesso em: 08 dez. 2018.

LEWIS, G.P. *et al.* **The legumes of the world**. Kew: Royal Botanic Gardens, 2005.

LI, S.F. *et al.* The landscape of transposable elements and satellite DNAs in the genome of a dioecious plant spinach (*Spinacia oleracea* L.). **Mobile DNA**. [on-line], v. 10, n. 3, 2019. Disponível em: <https://mobilednajournal.biomedcentral.com/articles/10.1186/s13100-019-0147-6>. Acesso em: 19 mar. 2019.

LIAO, D. Concerted Evolution: Molecular Mechanism and Biological Implications. **The American Journal of Human Genetics**. [on-line], v. 64, p. 24–30, 1999. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1377698/>. Acesso em: 28 fev. 2020.

LIMA, H.C.; QUEIROZ, L.P.; MORIM, M.P. Fabaceae. In: FORZZA, R. C. *et al.* (org.). **Catálogo de plantas e fungos do Brasil**. Andreia Jakobsson Estúdio: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, 2010. v.2. Disponível em: <reflora.jbrj.gov.br/downloads/vol1.pdf>. Acesso em: 22 abr. 2018.

LIU, C.J.; MUSIAL, J.M.; THOMAS, B.D. Genetic relationships among *Stylosanthes* species revealed by RFLP and STS analyses. **Theoretical and Applied Genetics**, [on-line], n. 99, p. 1179-1186, 1999. Disponível em: <https://link.springer.com/article/10.1007%2Fs001220051322>. Acesso em: 17 mar. 2018.

LÓPEZ- FLORES, I.; GARRIDO- RAMOS, M.A. The Repetitive DNA Content of Eukaryotic Genomes. **Genome Dynamic Basel**, Karger, v. 7, p. 1-28, 2012. Disponível em: <https://www.karger.com/Article/Abstract/337118>. Acesso em: 20 jan. 2019.

MAASS B.L.; MANNETJE L. ‘t. *Stylosanthes seabrana* (Leguminosae: Papilionoideae), a new species from Bahia, Brazil. **Novon**, [on-line], v. 12, p. 497–500, 2002. Disponível em: <https://www.biodiversitylibrary.org/page/643076#page/498/mode/1up>. Acesso em: 11 nov. 2018.

- MACAS, J. *et al.* Depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe fabaeae. **Plos One**, San Francisco, v. 10, n. 11, p. 1-23, 2015. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143424>. Acesso em: 19 ago. 2018.
- MARDIS, E. R. Next-Generation DNA Sequencing Methods. **Annual Review of Genomics and Human Genetics**, [s. l.], v. 9, p. 387-402, 2008. Disponível em: <https://www.annualreviews.org/doi/abs/10.1146/annurev.genom.9.081307.164359>. Acesso em: 17 jul. 2018.
- MARQUES, A. *et al.* Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. **Proceedings of the National Academy of Sciences of the USA**, [on-line], v. 112, p. 13633-13638, 2015. Disponível em: <https://www.pnas.org/content/112/44/13633.short>. Acesso em: 16 set. 2018.
- MARQUES, A. *et al.* Origin and parental genome characterization of the allotetraploid *Stylosanthes scabra* Vogel (Papilionoideae, Leguminosae), an important legume pasture crop. **Annals of Botany**, [on-line], n. 22, p. 1- 17, 2018. Disponível em: <https://academic.oup.com/aob/article/122/7/1143/5048901>. Acesso em: 03 mar. 2019.
- MARTUSCELLO, J.A. *et al.* Diversidade genética em acessos de *Stylosanthes capitata*. **Boletim de Indústria Animal**, Nova Odessa, v. 72, n. 4, p. 284-289, 2015. Disponível em: <http://www.iz.sp.gov.br/pdfsbia/1450723374.pdf> . Acesso em: 28 mar. 2018.
- MATSUMOTO, T. *et al.* The map-based sequence of the rice genome. **Nature**, [s. l.], v. 436, n. 7052, p. 793–800, 2005. Disponível em: https://www.nature.com/articles/nature03895?utm_medium=affiliate&utm_source=commission_on_junction&utm_campaign=3_nsn6445_deeplink_PID2190813&utm_content=deeplink. Acesso em: 15 dez. 2018.
- MEHROTRA, S. *et al.* Significance of Satellite DNA Revealed by Conservation of a Widespread Repeat DNA Sequence Among Angiosperms. **Applied Biochemistry and Biotechnology**, [on-line], v. 173, n. 7, p. 1790-17801, 2014. Disponível em: <https://link.springer.com/article/10.1007%2Fs12010-014-0966-3>. Acesso em: 10 fev. 2019.
- MEHROTRA, S.; GOYAL, V. Repetitive Sequences in Plant Nuclear DNA: Types, Distribution, Evolution and Function. **Genomics Proteomics Bioinformatics**, [on-line], v. 12, p. 164-171, 2015. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4411372/>. Acesso em: 22 mar. 2019.
- METZKER, M.L. Sequencing technologies - the next generation. **Nature Reviews Genetics**, [s. l.], v. 11, p. 31-46, 2010. Disponível em: <https://www.nature.com/articles/nrg2626>. Acesso em: 10 nov. 2018.
- NASCIMENTO, M.P.S.C.B. *et al.* Avaliação de acessos de *Stylosanthes scabra* Vog no Piauí. In: REUNIAO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 38, Piracicaba. **Anais [...]**. Piracicaba: SBZ, 2001. p. 184-185. Disponível em: <https://ainfo.cnptia.embrapa.br/digital/bitstream/item/68983/1/2-0518.PDF>. Acesso em: 25 mar. 2018.

NOVÁK, P. *et al.* Graph-based clustering and characterization repetitive sequences in next generation sequencing data. **BMC Bioinformatics**, [on-line], v. 11, n. 378, p. 1-12, 2010. Disponível em: <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-378>. Acesso em: 03 ago. 2019.

NOVÁK, P. *et al.* TAREAN: a computacional toll for identification and characterization of satellite DNA from unassembled short reads. **Nucleic Acids Research**, [on-line], v. 45, n. 12, p. 1-10, 2017. Disponível em: <https://academic.oup.com/nar/article/45/12/e111/3574061>. Acesso em: 15 ago. 2019.

NOVÁK, P. *et al.* Repeat explorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. **Bioinformatics**, [on-line], v. 29, p. 792-793, 2013. Disponível em: <https://academic.oup.com/bioinformatics/article/29/6/792/184407>. Acesso em: 17 jul. 2018.

PELLICER, J. *et al.* Genome size diversity and its impact on the evolution of land plants. **Genes**, [on-line], v. 9, n. 88, p. 1-14, 2018. Disponível em: <https://www.mdpi.com/2073-4425/9/2/88>. Acesso em: 12 ago. 2019.

PELLICER, J.; FAY, M. F.; LEITCH, I. J. The largest eukaryotic genome of them all? **Botanical Journal of the Linnean Society**, [s. l.], v. 164, p. 10-15, 2010. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1095-8339.2010.01072.x>. Acesso em: 12 ago. 2019.

PROSDOCIMI, F. **Introdução à Bioinformática**. Brasília: Biotecnologia Ciencia & Desenvolvimento, 2007. 77p. Disponível em: <https://www.passeidireto.com/arquivo/62955027/introducao-a-bio-informatica-francisco-prosdocimi-2007>. Acesso em: 16 ago. 2019.

QUEIROZ, L.P.; CARDOSO, D.B.O.S. Leguminosae. *In*: GIULIETTI, A. M. *et al.* **Plantas raras do Brasil**. Belo Horizonte, MG: Conservação Internacional, 2009. Disponível em: https://biowit.files.wordpress.com/2010/11/plantas_raras_do_brasil.pdf. Acesso em: 04 nov. 2018.

RIBEIRO, S.B. **Sequenciamento e caracterização parcial do genoma de cagaiteira (*Eugenia dysenterica* DC.)**. 2016. 80f.: Dissertação (Mestrado em genética e melhoramento de plantas) - Universidade Federal de Goiás, Escola de Agronomia (EA), Goiânia, 2016. Disponível em: <http://repositorio.bc.ufg.br/tede/handle/tede/6232>. Acesso em: 10 ago. 2019.

RICHARD, G-F; KERREST, A.; DUJON, B. Comparative Genomics and Molecular Dynamics of DNA Repeats in Eukaryotes. **Microbiology and Molecular Biology Reviews**, [on-line], v. 72, n. 4, p. 686-727, 2008. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2593564/>. Acesso em: 12 ago. 2019.

ROBLEDILLO, L.A. Satellite DNA in *Vicia faba* is characterized by remarkable diversity in its sequence composition, association with centromeres, and replication timing. **Scientific Reports**, [on-line], v. 8, n. 5838, 2018. Disponível em: <https://www.nature.com/articles/s41598-018-24196-3>. Acesso em: 05 ago. 2019.

SANGER F.; NICKLEN S.; COULSON A.R. DNA sequencing with chain-terminating inhibitors. **Proceedings of the National Academy of Sciences of the United States of America**, [on-line], v. 74, p. 5463-5467, 1997. Disponível em: <https://www.pnas.org/content/pnas/74/12/5463.full.pdf>. Acesso em: 14 ago. 2019.

SANTOS, L.G. dos. **Análise comparativa da pureza genética das leguminosas forrageiras *Stylosanthes capitata* vog. e *Stylosanthes macrocephala* M.B Ferr. Et Sousa Costa utilizando marcadores moleculares**. São Paulo, 2014. 49f. Dissertação (Mestrado em Biologia Comparada) - Universidade de São Paulo, São Paulo, 2014. Disponível em: <https://teses.usp.br/teses/disponiveis/59/59139/tde-29122014-214827/en.php>. Acesso em: 22 nov. 2018.

SANTOS, M.O. *et al.* Isolation and characterization of microsatellite loci in the tropical forage legume *Stylosanthes guianensis* (Aubl.) Sw. **Conservation Genetics Research**, [on-line], n. 1, p. 43-46, 2009. Disponível em: <https://link.springer.com/article/10.1007%2Fs12686-009-9010-2>. Acesso em: 11 set. 2018.

SANTOS-GARCIA, M.O. *et al.* Identification of *Stylosanthes guianensis* varieties using molecular genetic analysis. **Aob Plants**, [on-line], 2012. Disponível em: <https://academic.oup.com/aobpla/article/doi/10.1093/aobpla/pls001/174418>. Acesso em: 23 mar. 2018.

SCHIFINO-WITTMANN, M.T. Poliploidia e seu impacto na origem e evolução das plantas silvestres e cultivadas. **Revista brasileira de Agrociência**, Pelotas, v. 10, n. 2, p. 151-157, 2004. Disponível em: <http://www2.ufpel.edu.br/faem/agrociencia/v10n2/artigo02.pdf>. Acesso em: 15 ago. 2019.

SCHUBERT, I.; VU, G.T.H. Genome Stability and evolution: attempting a holistic view. **Trends in Plant Science**, [on-line], v. 21, n. 9, p. 749-757, 2016. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S1360138516300681>. Acesso em: 06 ago. 2019.

STAATS, C.C.; MORAIS, G.L.de; MARGIS, R. Projetos Genoma. *In*: VERLI, H. (org.). **Bioinformática da biologia à flexibilidade molecular**. São Paulo: SBBq, cap. 4, p. 63-79, 2014. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/166105/001012172.pdf?sequence=1>. Acesso em: 15 ago. 2019.

STACE, H.M.; CAMERON, D.F. Cytogenetic review of taxa in *Stylosanthes hamata* sensu lato. **Trop Grasslands**, [s. l.], v. 2, n. 4, p. 182-188, 1987. Disponível em: http://tropicalgrasslands.info/public/journals/4/Historic/Tropical%20Grasslands%20Journal%20archive/PDFs/Vol_21_1987/Vol_21_04_87_pp182_188.pdf. Acesso em: 15 mar. 2018.

STACE, H.M.; CAMERON, D.F. Cytogenetics and the evolution of *Stylosanthes*. *In*: STACE, H. M.; EDYE, L. A. (ed.). **The biology and agronomy of *Stylosanthes***. Sidney, AU: Academic Press, 1984. p. 49-72. Disponível em: <http://www.sidalc.net/cgi-bin/wxis.exe/?IsisScript=catalco.xis&method=post&formato=2&cantidad=1&expresion=mfn=033760>. Acesso em: 08 mar. 2018.

STACE, H.M.; EDYE, L.A. **The Biology and agronomy of *Stylosanthes***. Londres: Academic Press, 1984. Disponível em: https://books.google.com.br/books?hl=pt-BR&lr=&id=DZk7YiopHWEC&oi=fnd&pg=PP1&dq=The+Biology+and+agronomy+of+Stylosanthes.+Academic+Press.&ots=OIxHVSRDcy&sig=WWfuqrgdduOGMywjykJdu_534L0#v=onepage&q=The%20Biology%20and%20agronomy%20of%20Stylosanthes.%20Academic%20Press.&f=false. Acesso em: 10 out. 2018.

VANDER STAPPEN, J. *et al.* Interspecific and progeny relationships in the genus *Stylosanthes* inferred from chloroplast DNA sequence variation. **Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie**, [s. l.], n. 322, p. 481-490, 1999. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0764446999800985>. Acesso em: 24 abr. 2018.

VANDER STAPPEN, J. *et al.* Phylogenetic analysis of *Stylosanthes* (Fabaceae) based on the internal transcribed spacer region (ITS) of nuclear ribosomal DNA. **Plant Systematics and Evolution**, [on-line], n. 234, p. 27-51, 2002. Disponível em: <https://link.springer.com/article/10.1007/s00606-002-0193-1>. Acesso em: 04 maio 2018.

VANNI, R. O. The genus *Stylosanthes* (Fabaceae, Papilionoideae, Dalbergieae) in South America. **Boletín de la Sociedad Argentina de Botánica**, Córdoba, v. 52, p. 549-586, 2017. Disponível em: <http://dx.doi.org/10.31055/1851.2372.v52.n3.18033>. Acesso em: 22 dez. 2019.

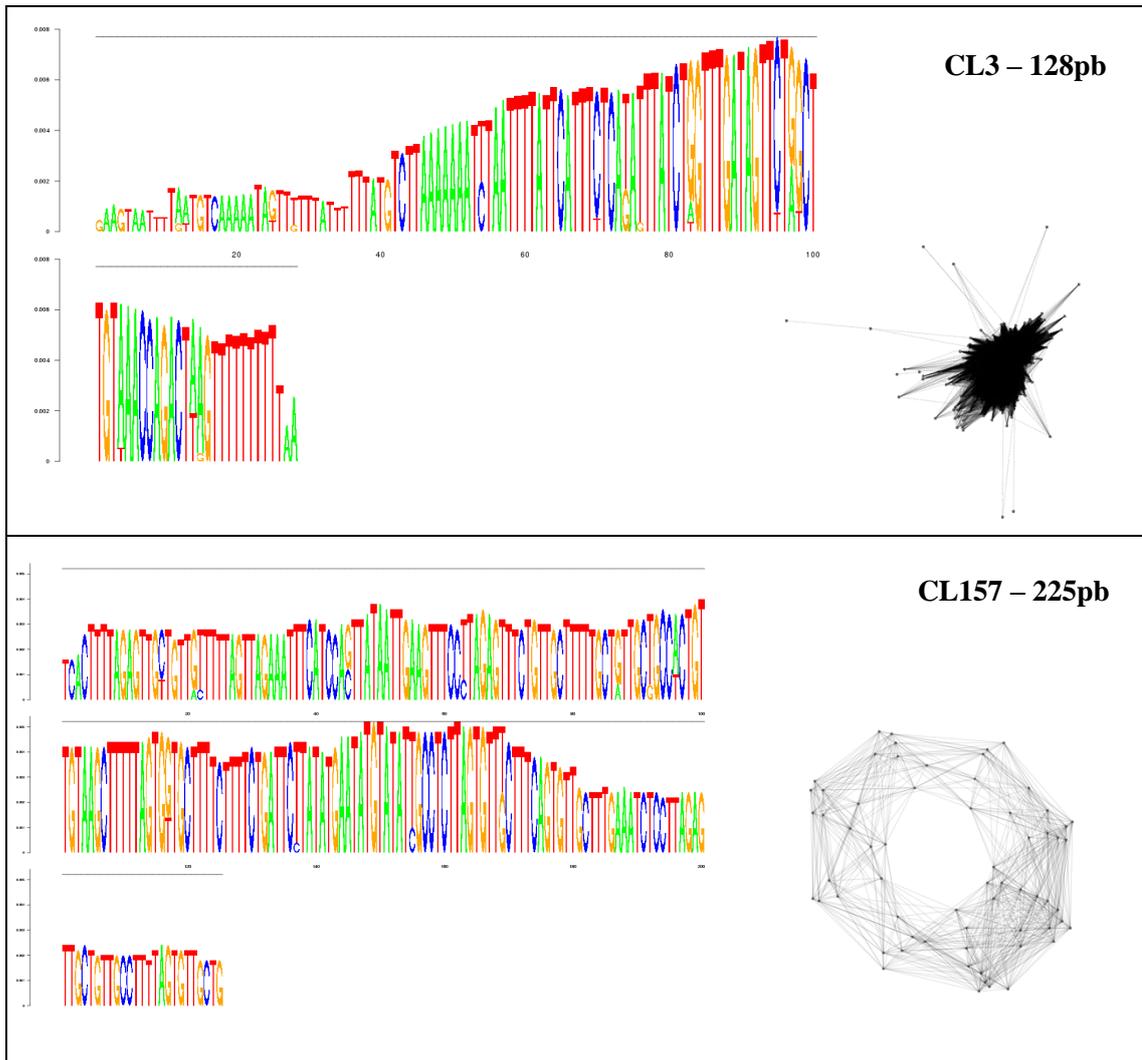
VANNI, R.O.; FERNANDEZ, A. The true identity of *Stylosanthes seabrana* BL Maass & L. 't Mannelje (Leguminosae Papilionoideae). **Caryologia**, [on-line], n. 64, p. 247-250, 2011. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/00087114.2011.10589789>. Acesso em: 20 fev. 2019.

VERLI, H. **O que é bioinformática?**. Bioinformática da Biologia à flexibilidade molecular. São Paulo: SBBq, cap. 1, p. 2-12, 2014. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/166105/001012172.pdf?sequence=1>. Acesso em: 15 ago. 2019.

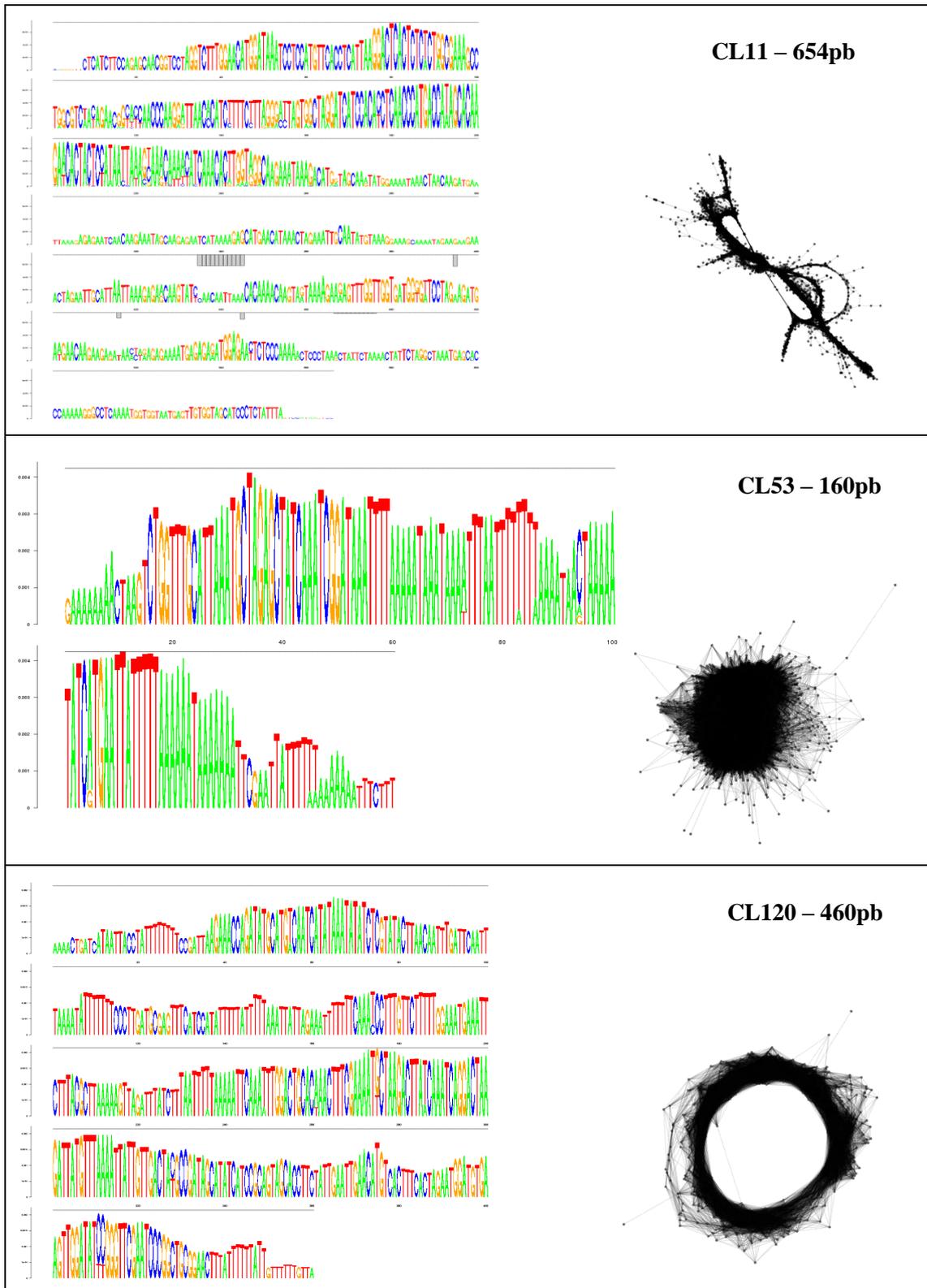
VIEIRA, M.L.C.; AGUIAR-PERECIN, M.L.R. de; MARTINS, O.S. A cytotaxonomic study in twelve Brazilian taxa of *Stylosanthes* Sw., Leguminosae. **Cytologia**, [s. l.], n. 58, p. 305-311, 1993. Disponível em: https://www.jstage.jst.go.jp/article/cytologia1929/58/3/58_3_305/_article/-char/ja/. Acesso em: 11 jan. 2019.

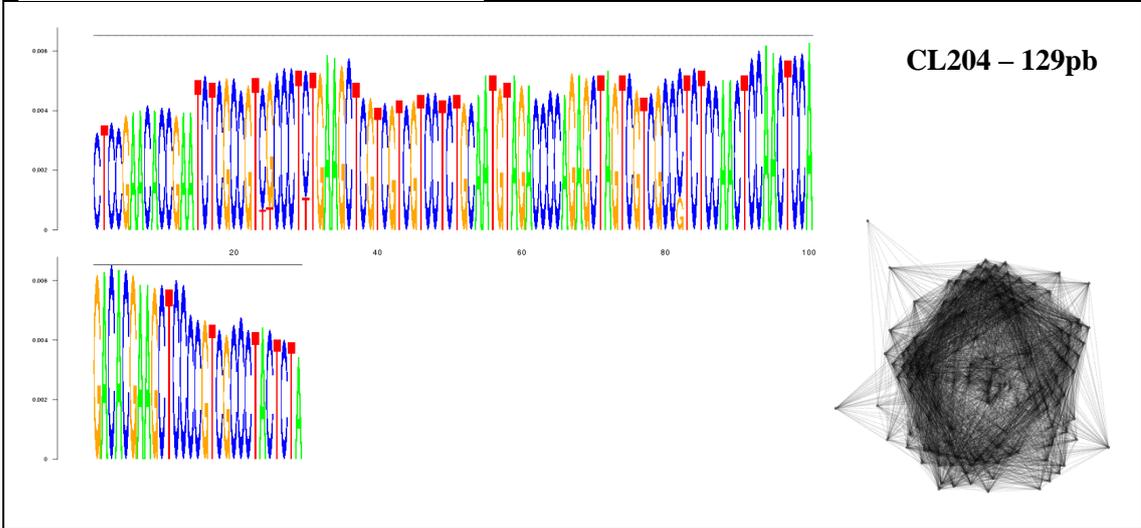
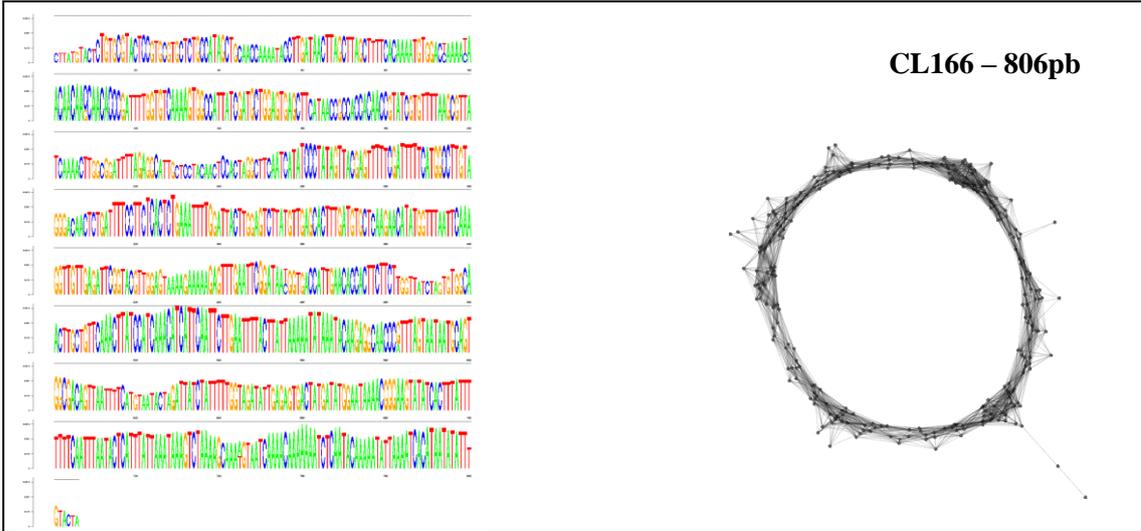
WICKER, T. *et al.* A unified classification system for eukaryotic transposable elements. **Nature Reviews Genetics**, [on-line], v. 8, n. 12, p. 973-982, 2007. Disponível em: https://www.nature.com/articles/nrg2165?utm_medium=affiliate&utm_source=commission_junction&utm_campaign=3_nsn6445_deeplink_PID2190813&utm_content=deeplink. Acesso em: 14 jun. 2019.

APÊNDICE A – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. hamata*

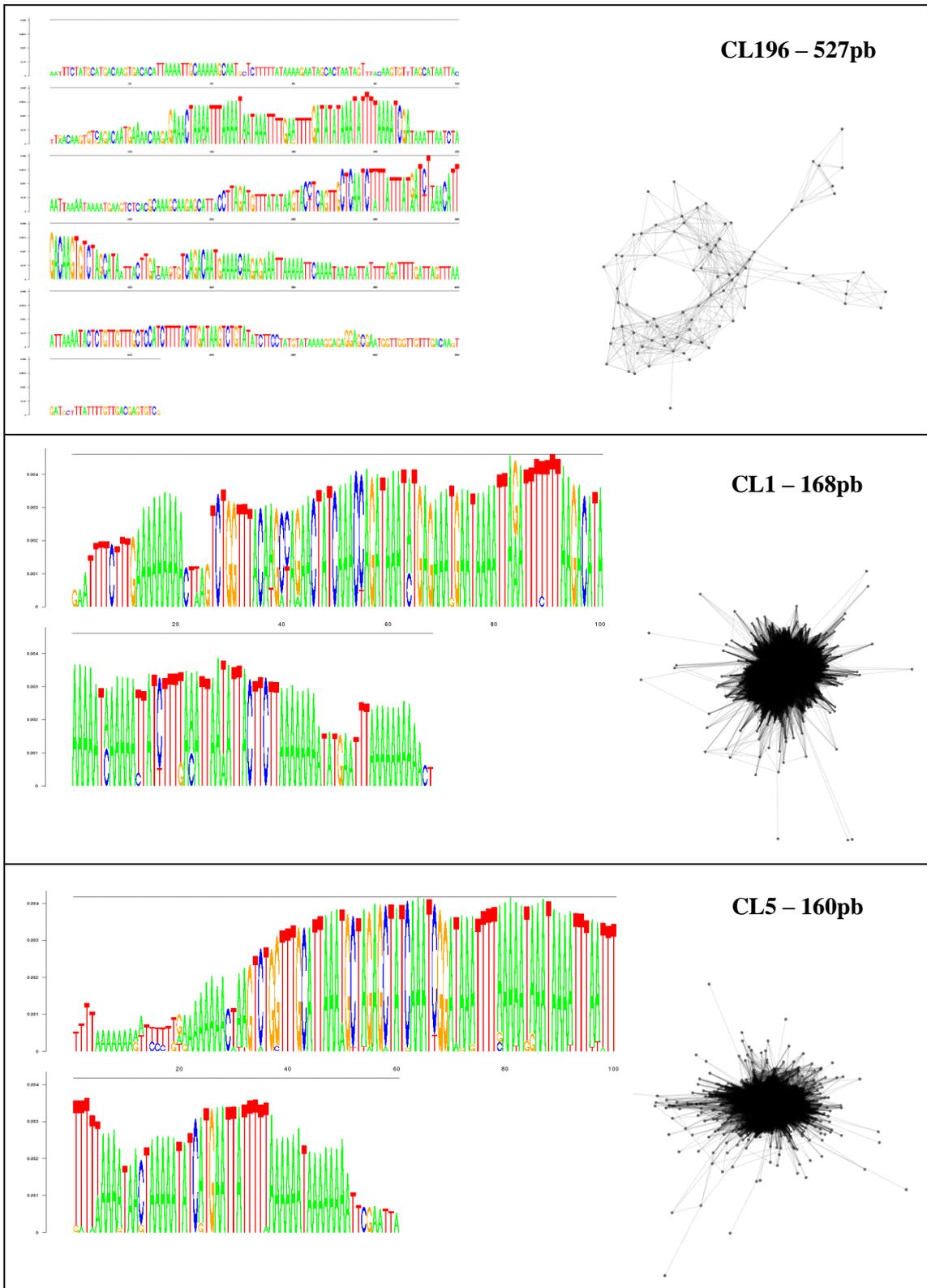


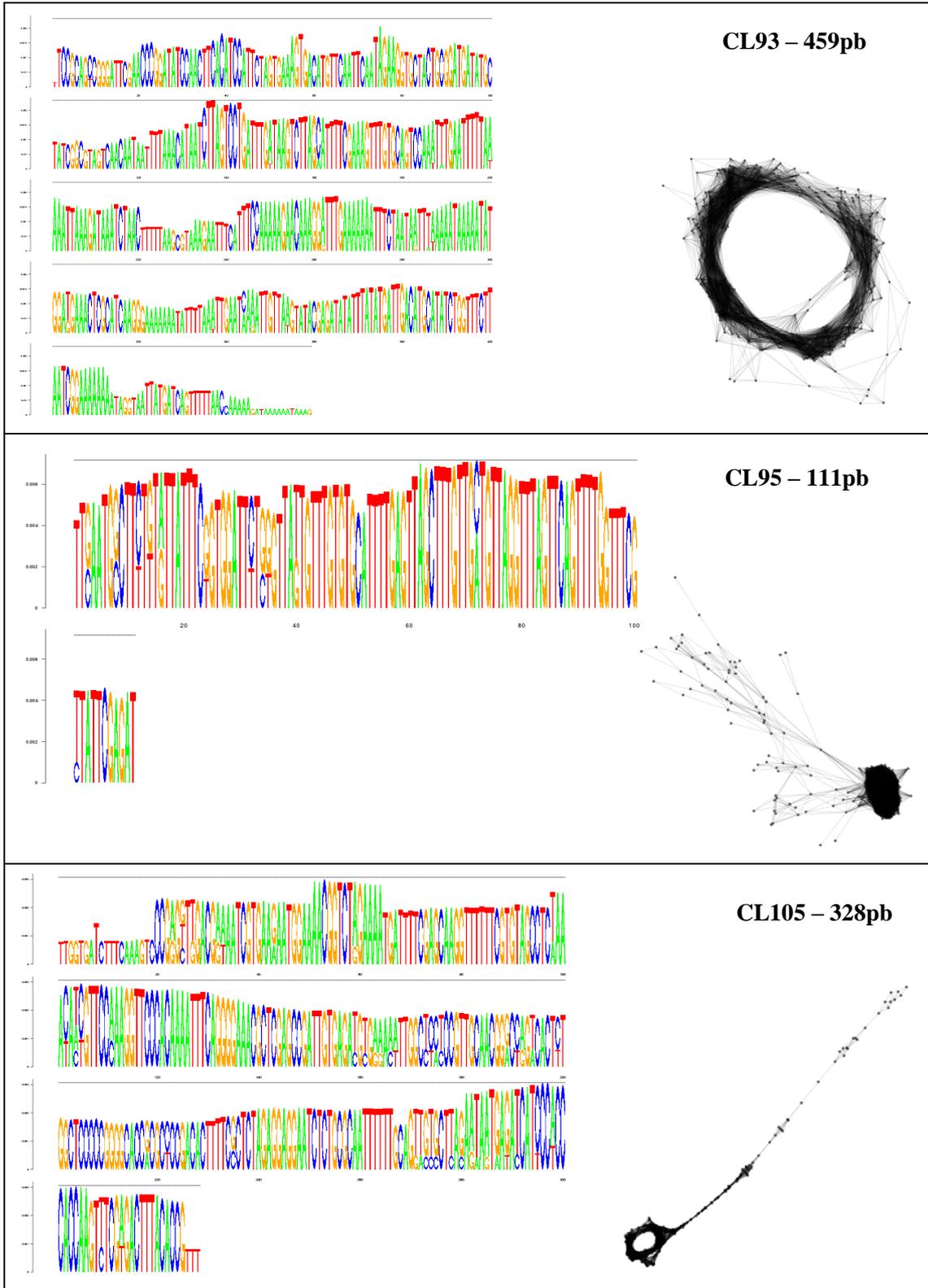
APÊNDICE B – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. viscosa*

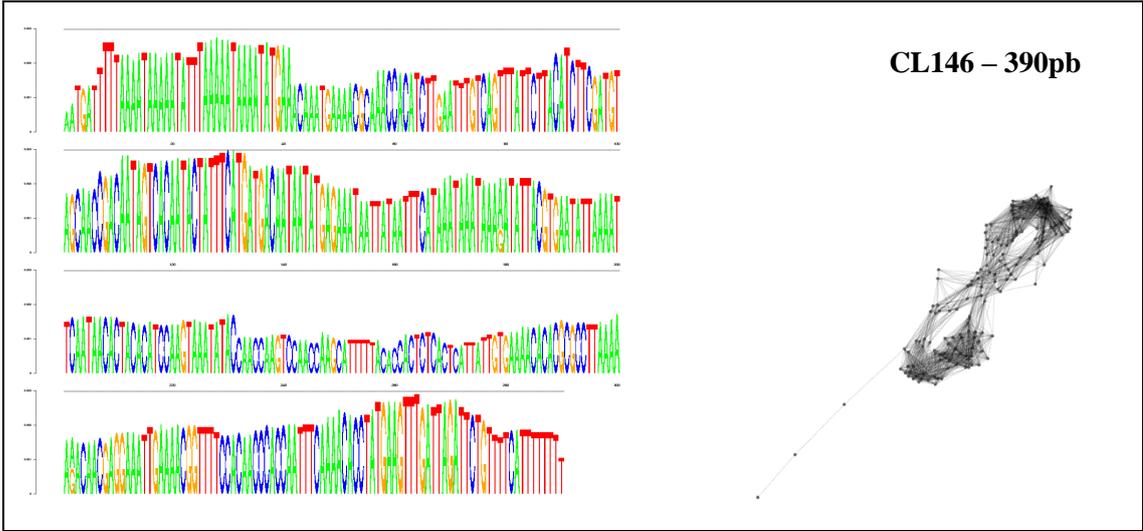


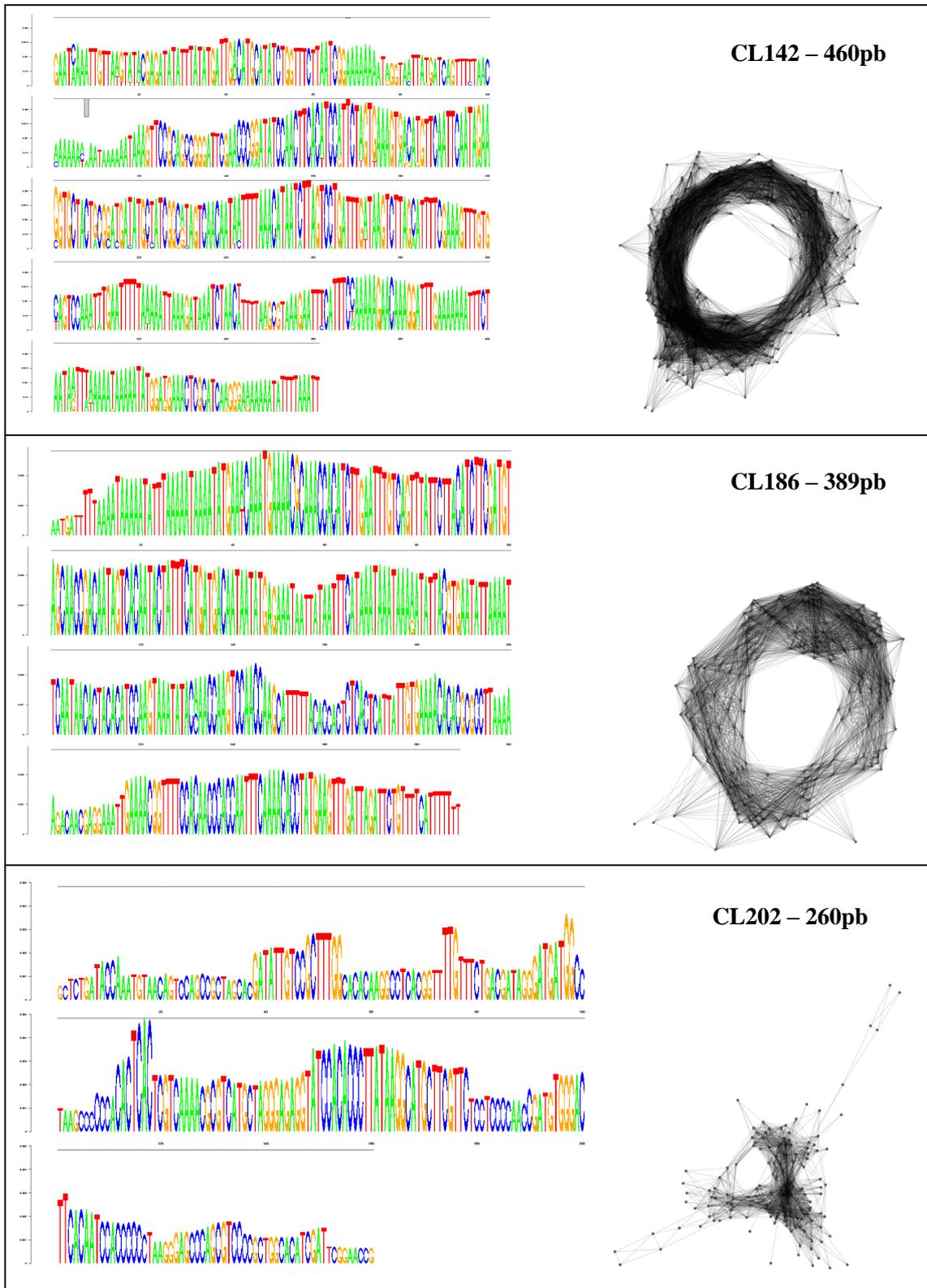


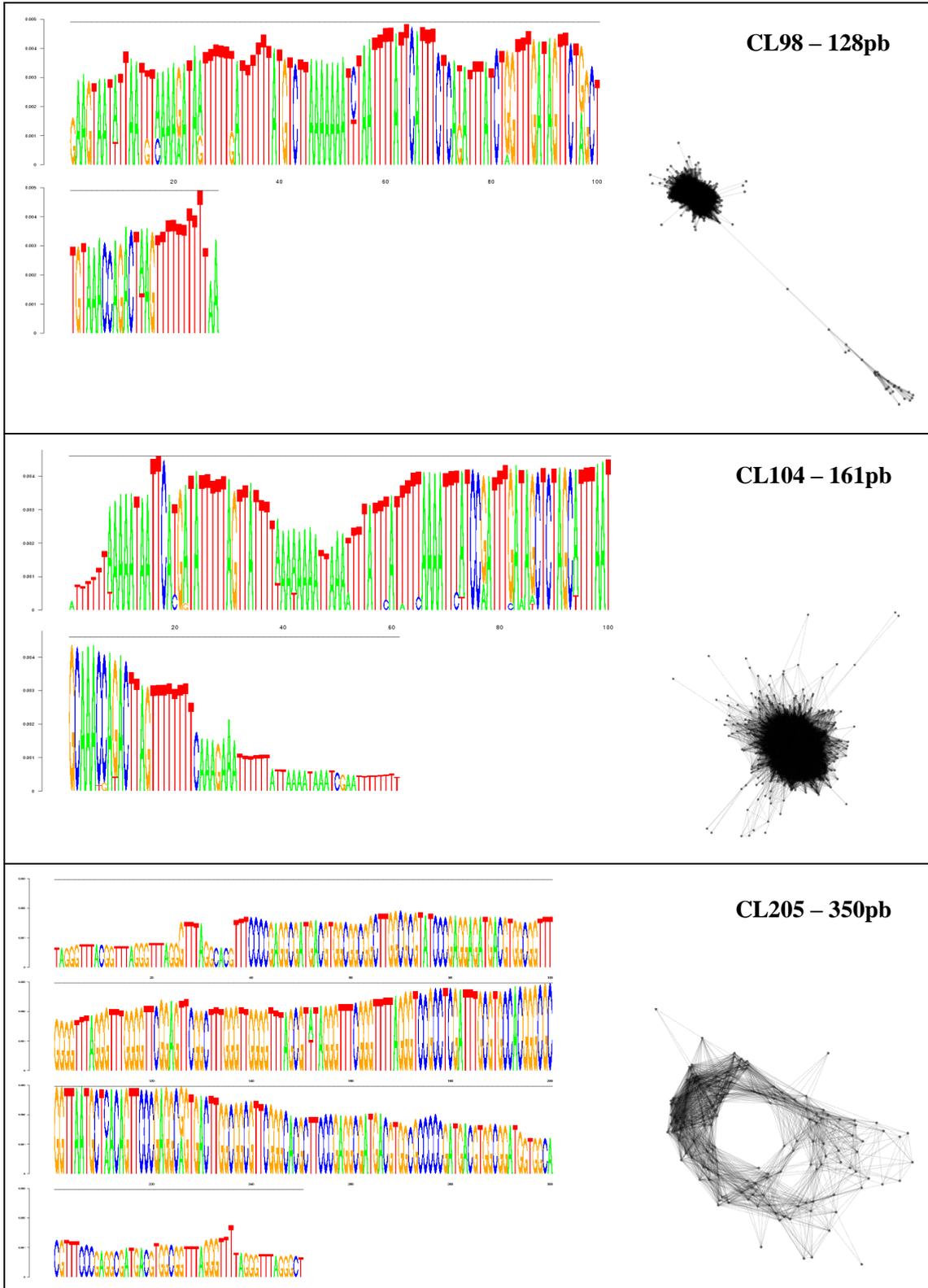
APÊNDICE C – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. scabra*

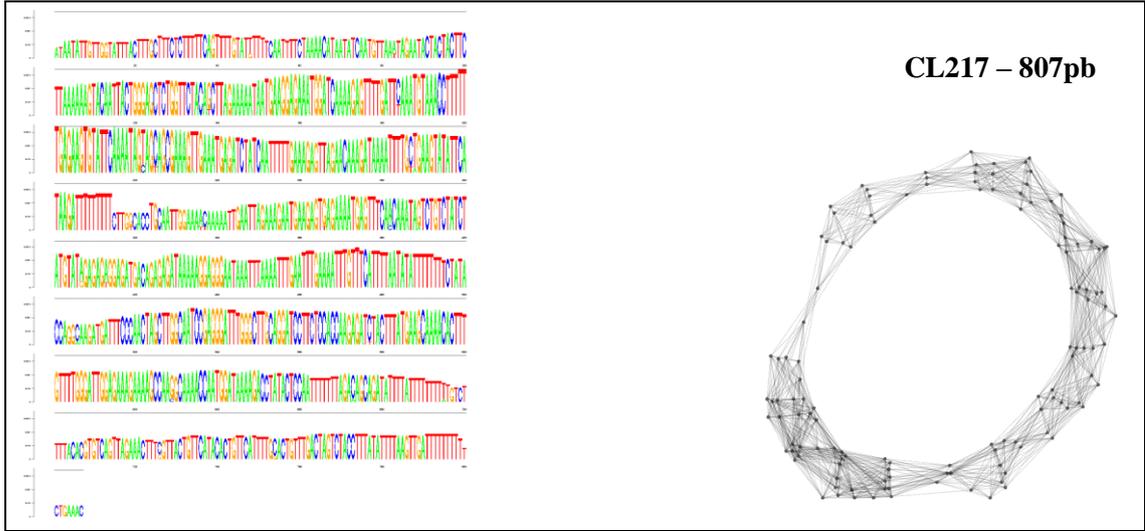


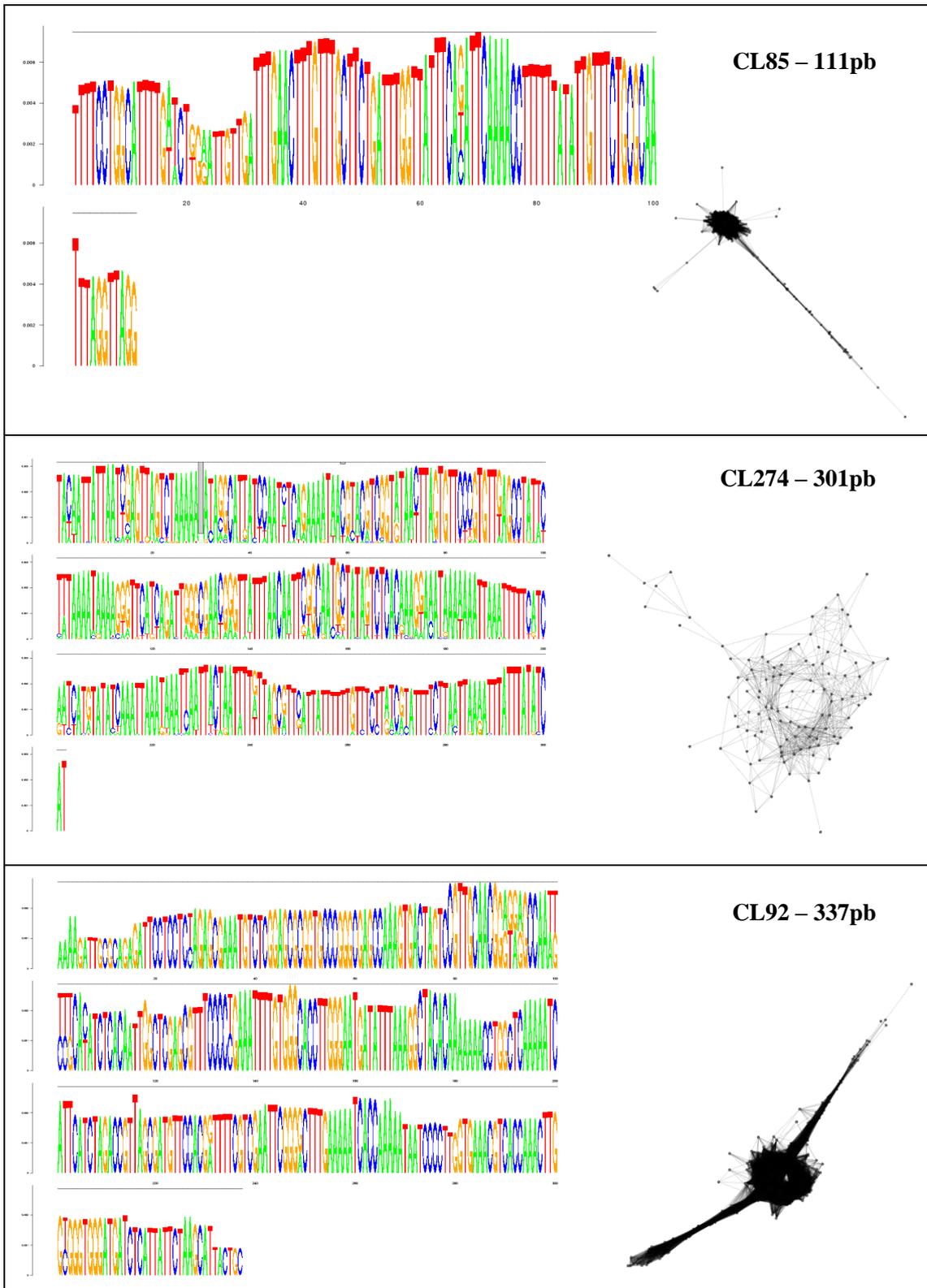


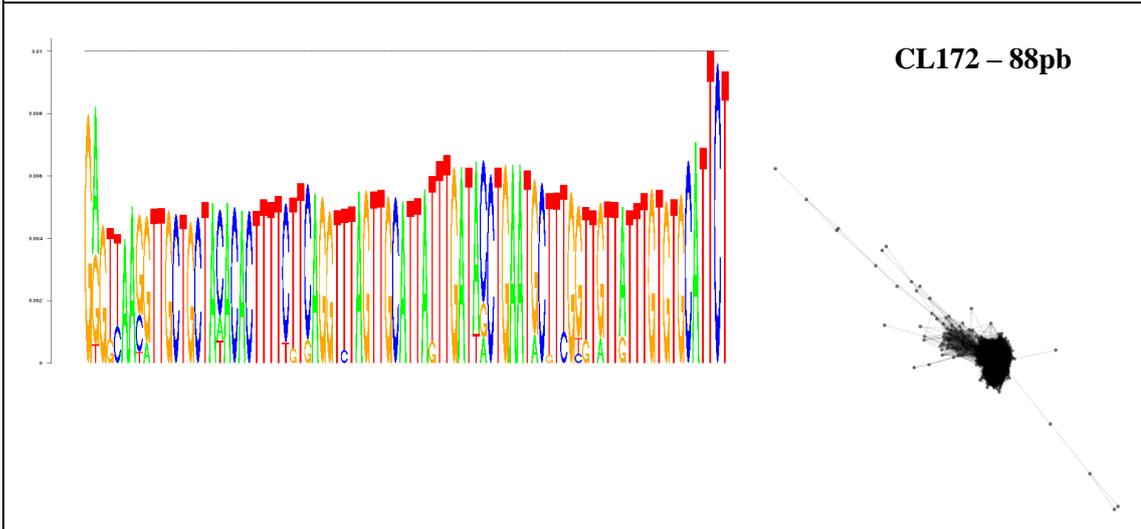
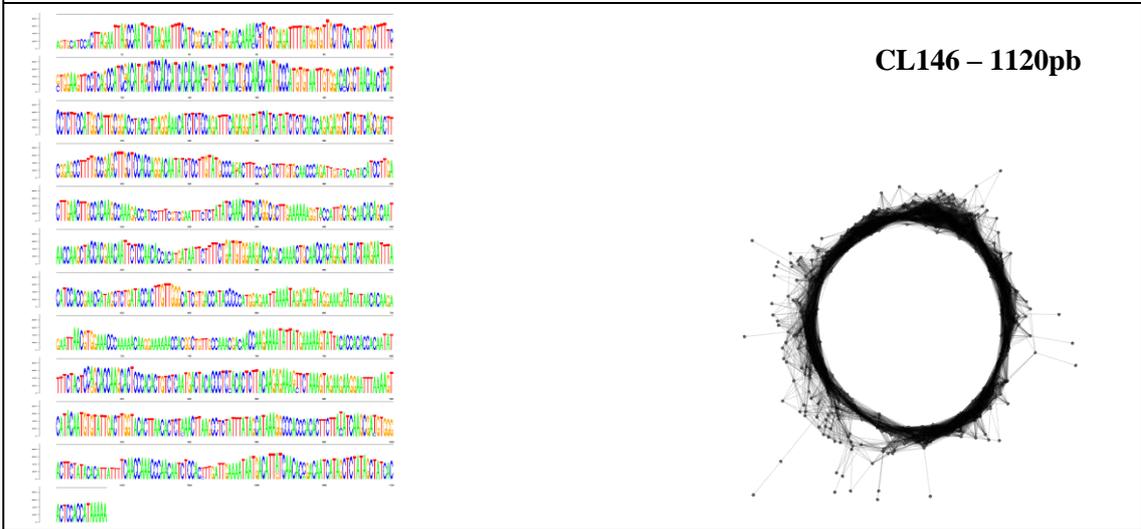
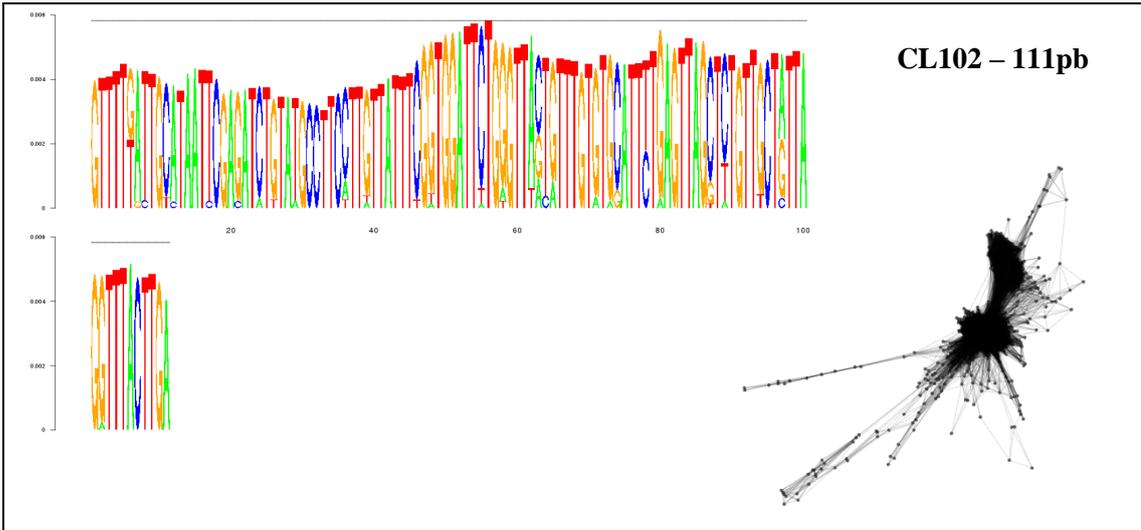


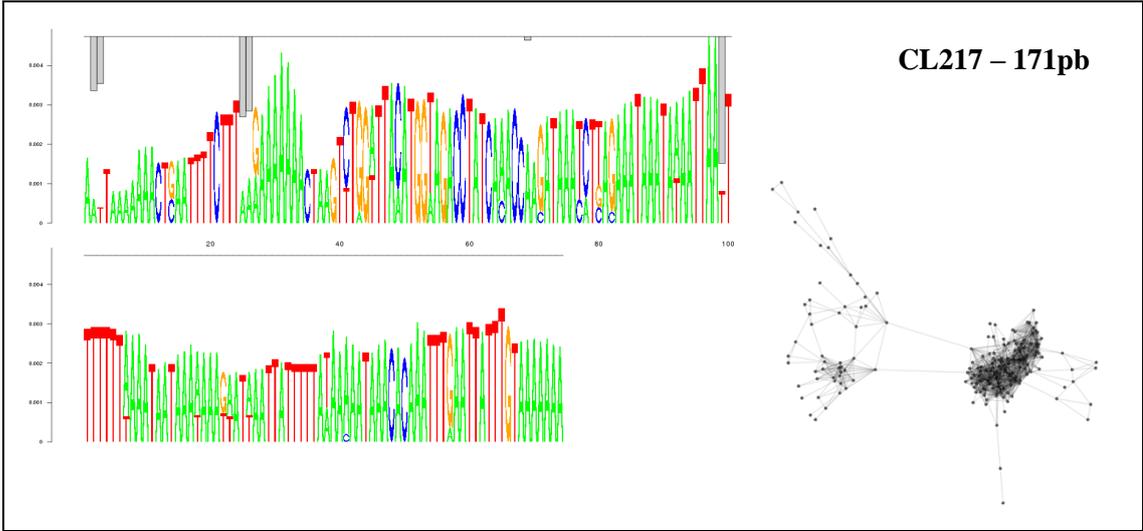
APÊNDICE D – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. capitata* (CG)

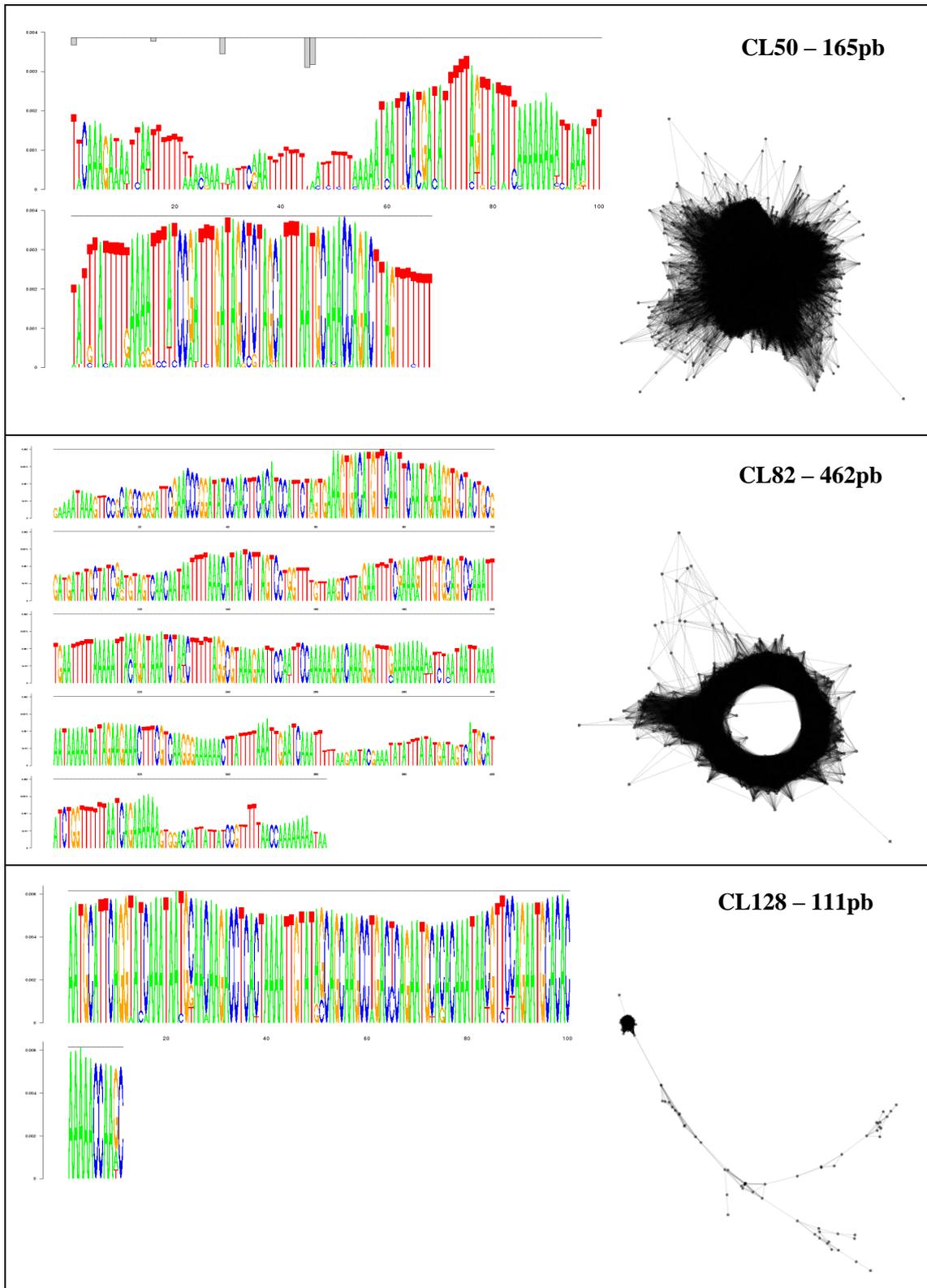


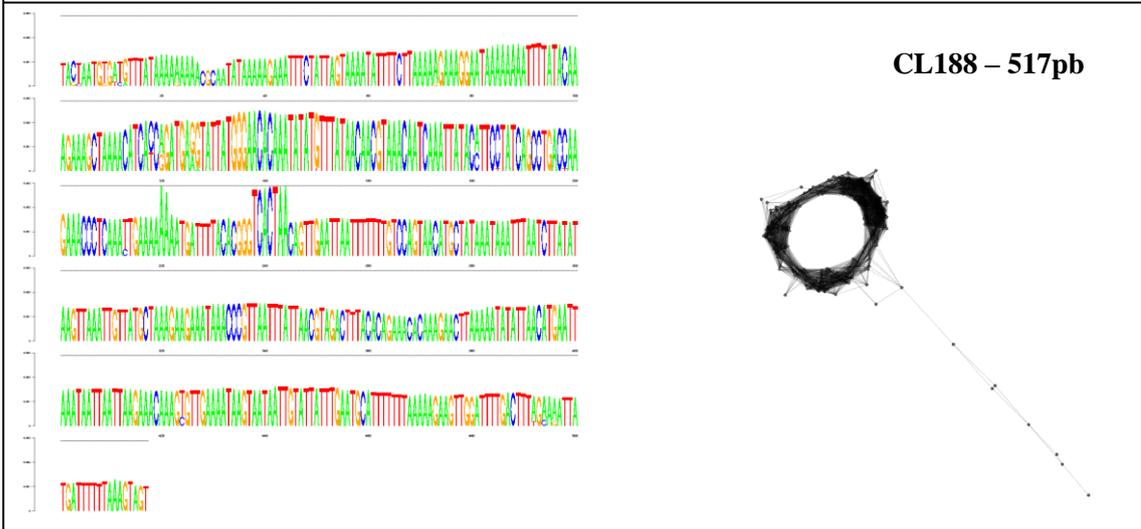
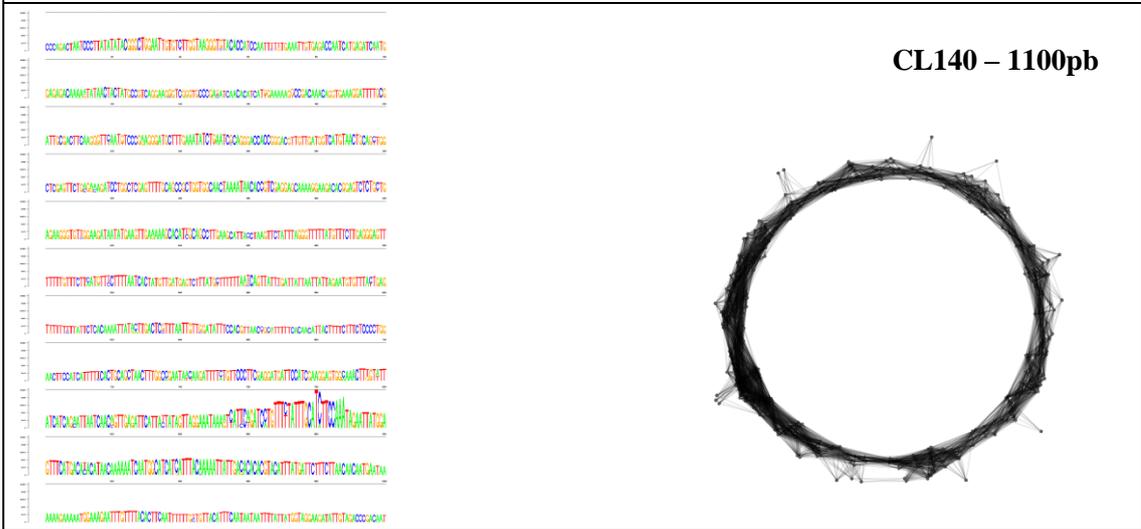
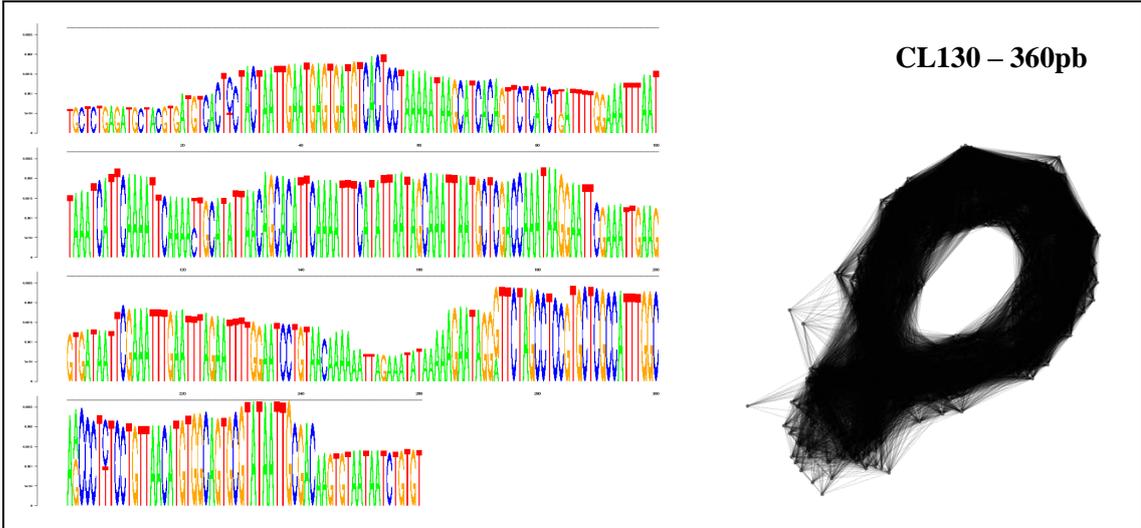


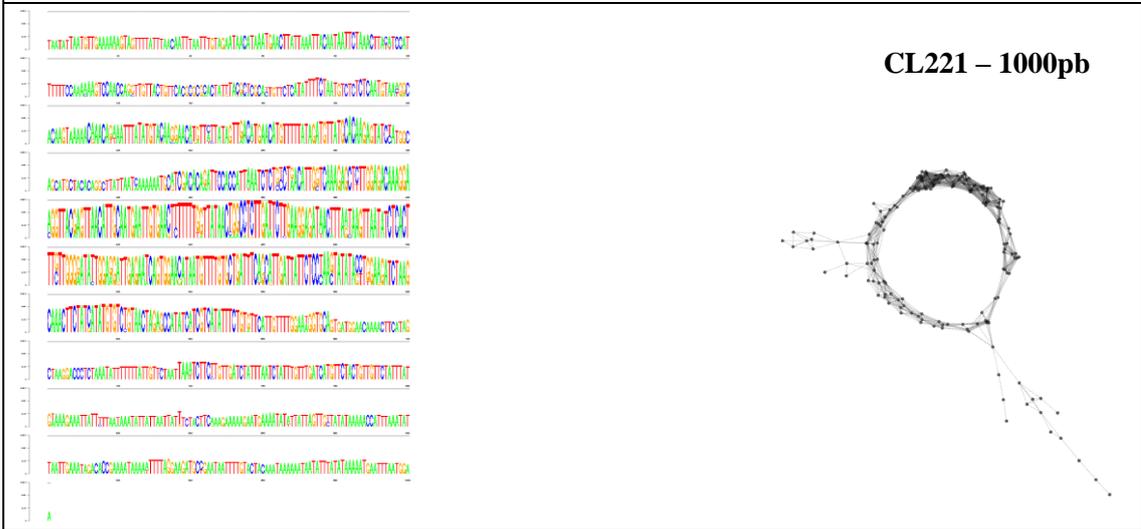
APÊNDICE E – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. pilosa*

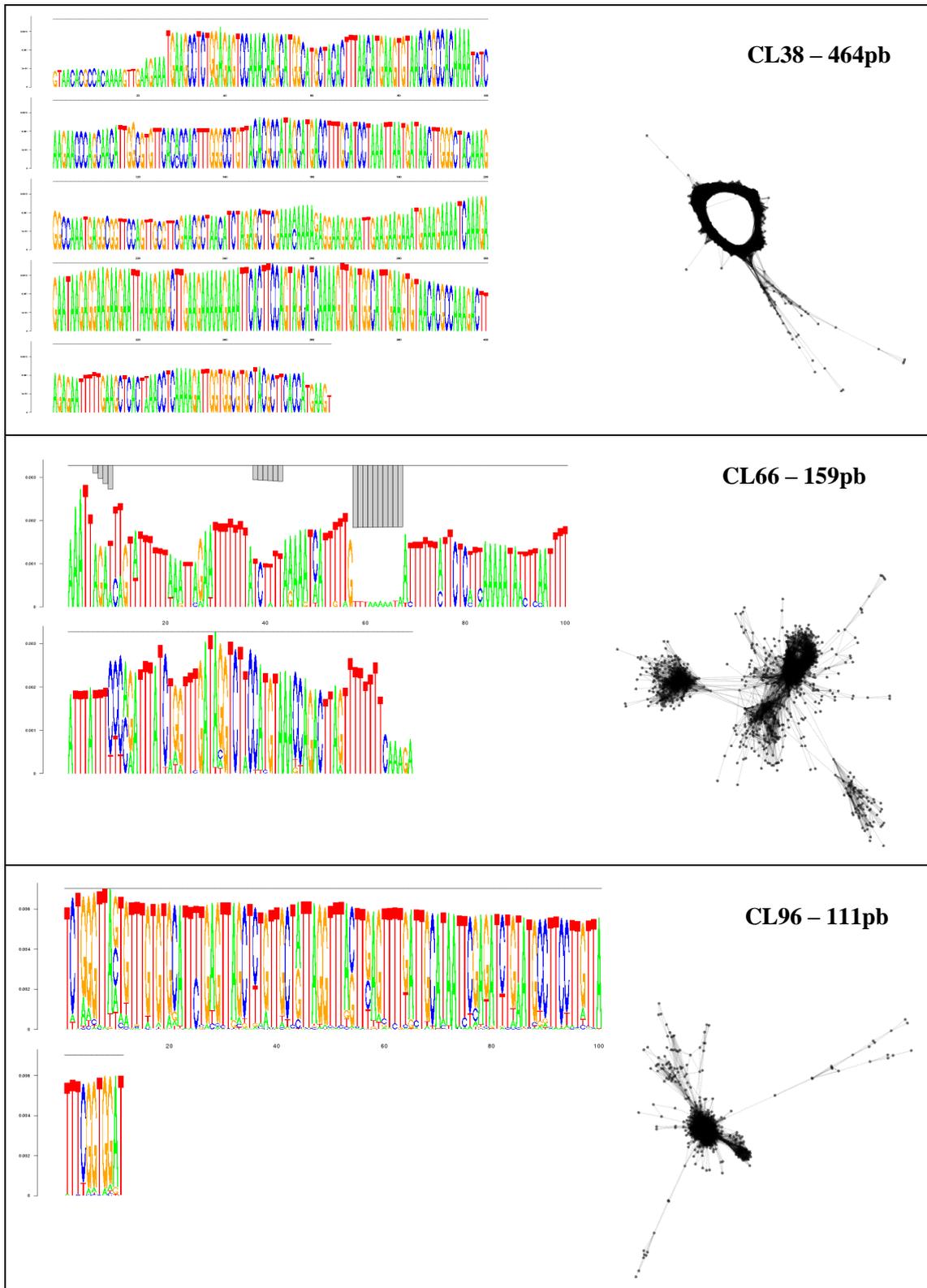


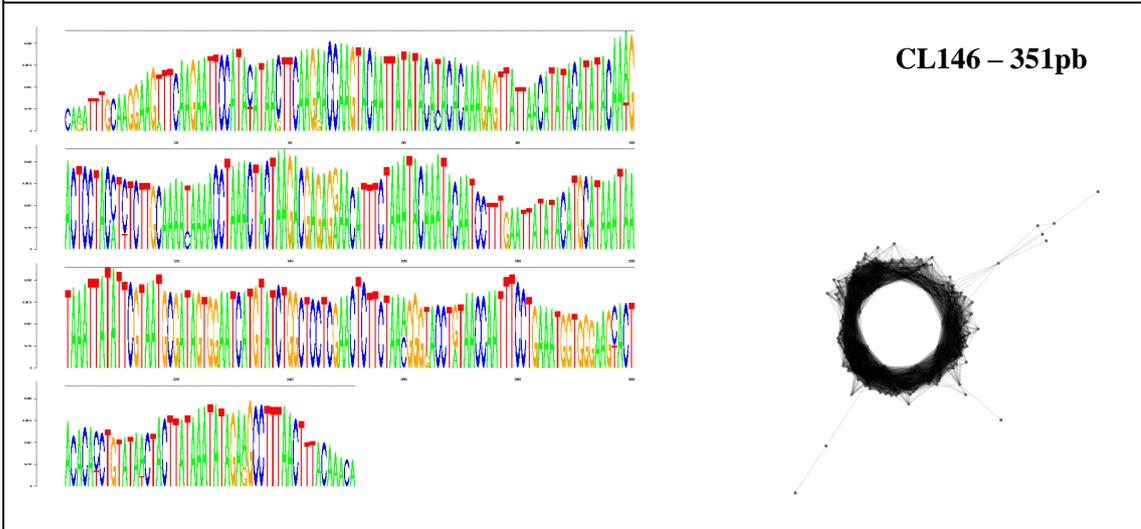
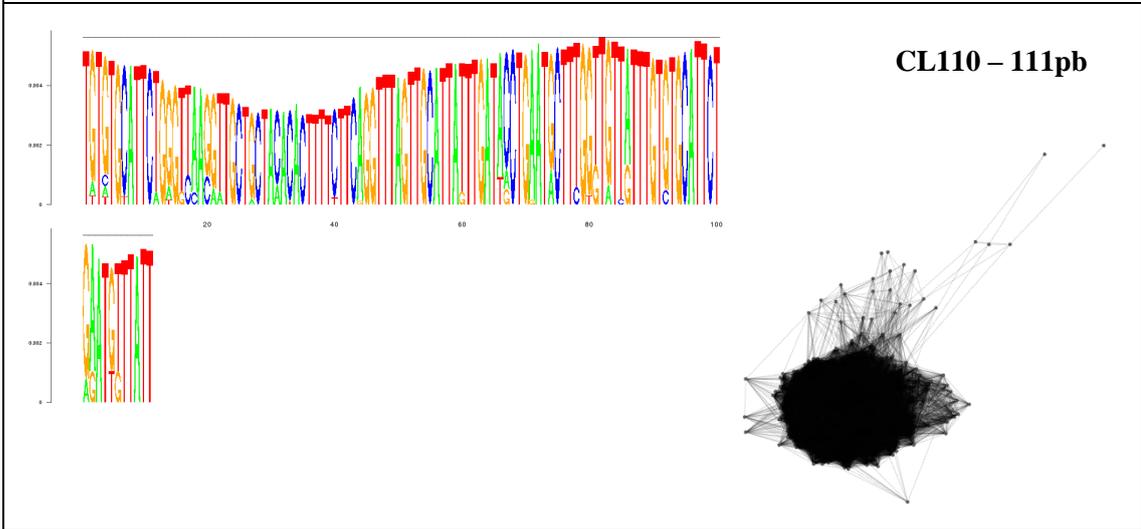
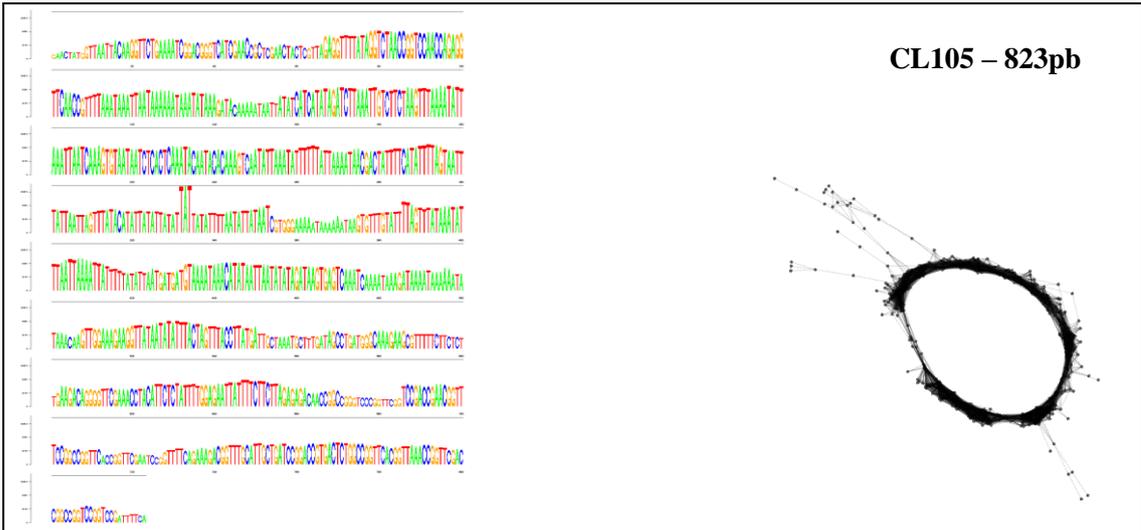


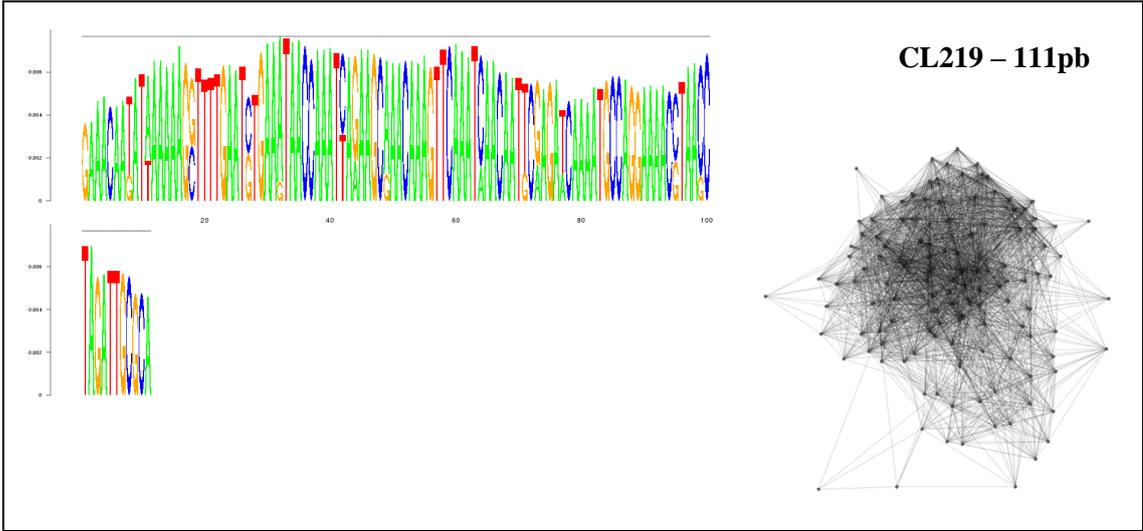
APÊNDICE F – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. macrocephala*

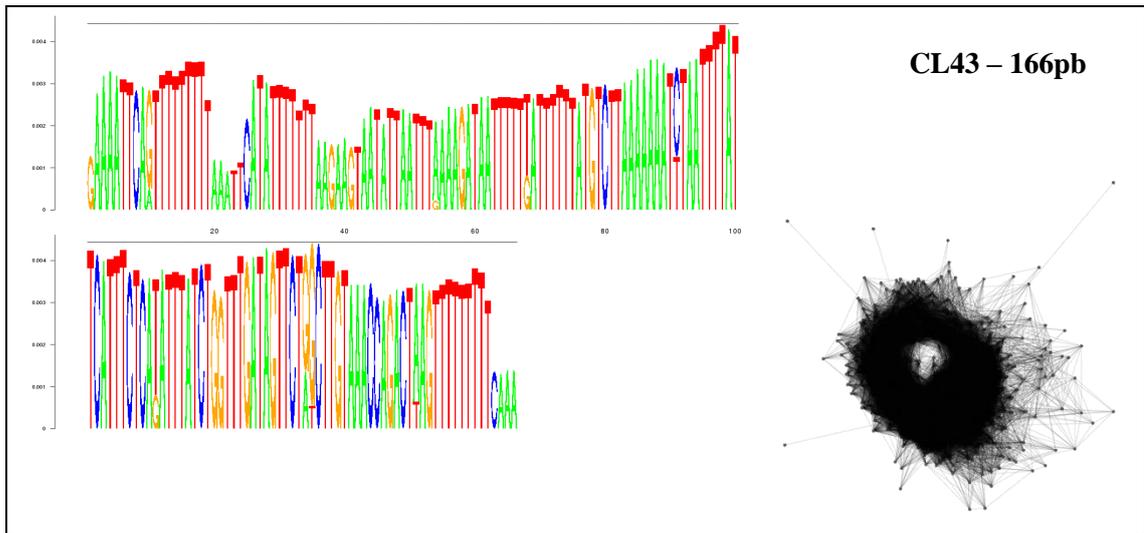




APÊNDICE G – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. capitata* (Acesso 24)





APÊNDICE H – SATÉLITE PRESENTE NO DNA NUCLEAR DE *S. seabrana*

APÊNDICE I – SATÉLITES PRESENTES NO DNA NUCLEAR DE *S. guianensis*

