



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE  
CENTRO DE DESENVOLVIMENTO SUSTENTÁVEL DO SEMIÁRIDO  
UNIDADE ACADÊMICA DE ENGENHARIA DE BIOTECNOLOGIA E BIOPROCESSOS  
CURSO DE ENGENHARIA DE BIOTECNOLOGIA E BIOPROCESSOS**

**DÉBORA SOUZA DOS SANTOS**

**MAPEAMENTO E ANÁLISE *IN SILICO* DAS PROTEÍNAS  
REFERENTES AO ESTRESSE SALINO EM ESPÉCIES DO GÊNERO  
DO ALGODOEIRO (*GOSSYPIUM spp.*)**

**SUMÉ - PB  
2021**

**DÉBORA SOUZA DOS SANTOS**

**MAPEAMENTO E ANÁLISE *IN SILICO* DAS PROTEÍNAS  
REFERENTES AO ESTRESSE SALINO EM ESPÉCIES DO GÊNERO  
DO ALGODOEIRO (*GOSSYPIUM spp.*)**

**Monografia apresentada ao Curso de Engenharia de Biotecnologia e Bioprocessos do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharela em Engenharia de Biotecnologia e Bioprocessos.**

**Orientadora: Professora Dra. Ana Verônica Silva do Nascimento.**

**Coorientador: Professor Dr. Horace José Jimenez.**

**SUMÉ - PB  
2021**



S237m Santos, Débora Souza dos.  
Mapeamento e análise in silico das proteínas referentes ao estresse salino em espécies do gênero do algodoeiro (*Gossypium* ssp.). / Débora Souza dos Santos. - 2021.

90 f.

Orientadora: Profa. Dra. Ana Verônica Silva do Nascimento; Coorientador: Prof. Dr. Horace José Jimenez. Trabalho de Conclusão de Curso - (Curso de Bacharelado em Engenharia de Biotecnologia e Bioprocessos) - Universidade Federal de Campina Grande; Centro de Desenvolvimento Sustentável do Semiárido.

1. Cotonicultura. 2. Estresse salino. 3. Salinidade. 4. Algodoeiro - *Gossypium* ssp. 5. Melhoramento genético vegetal. 6. Bioinformática. 7. Primers in silico - desenho. 8. Análise taxonômica de proteínas. 9. Análise físico-química de proteínas. I. Nascimento, Ana Verônica Silva do. II. Jimenez, Horace José. III. Título.

CDU:633.51(043.1)

**Elaboração da Ficha Catalográfica:**

Johnny Rodrigues Barbosa  
Bibliotecário-Documentalista  
CRB-15/626

**DÉBORA SOUZA DOS SANTOS**

**MAPEAMENTO E ANÁLISE *IN SILICO* DAS PROTEÍNAS  
REFERENTES AO ESTRESSE SALINO EM ESPÉCIES DO GÊNERO  
DO ALGODOEIRO (*GOSSYPIUM spp.*)**

**Monografia apresentada ao Curso de Engenharia de Biotecnologia e Bioprocessos do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharela em Engenharia de Biotecnologia e Bioprocessos.**

**BANCA EXAMINADORA:**

---

**Professora Dra. Ana Verônica Silva do Nascimento.  
Orientadora - UAEB/CDSA/UFCG**

---

**Professor Dr. Horace José Jimenez.  
Coorientador**

---

**Professor Dr. Jean César Farias de Queiroz  
Examinador I - UAEB/CDSA/UFCG**

---

**Professor Dr. Aldre Jorge Moraes Barros  
Examinador II - UAEB/CDSA/UFCG**

**Trabalho aprovado em: 27 de maio de 2021**

**SUMÉ - PB**

**Dedico este trabalho de monografia a minha família (Cremilda, Achibaldo, Adna, Jalom, Esther, Vô Oscar e Vó Daura), as pessoas amigas que colaboraram com meu crescimento e aos que me deram oportunidade e me incentivaram.**

## AGRADECIMENTOS

A minha jornada universitária contou com a passagem de muitas pessoas, as quais me deram oportunidades, dificultaram meu caminho me tornando mais forte e me impulsionaram cada vez mais a buscar meus objetivos. A seguir dedicarei para cada uma em específico:

A Universidade Federal de Campina Grande, campus Sumé, por ser o alicerce de minha formação profissional científica, docentes, discentes, especialmente, ao pessoal terceirizado que ganharam meu coração, vocês são maravilhosos. Além das agências de fomento a qual pude fazer parte, como CAPES e CNPq.

A Mérgia que me aceitou num projeto de extensão, o qual não pude fazer parte, mas agradeço mesmo assim. A senhora é luz!

A Alecksandra, por me aceitar fazer parte da família do Laeb, organizar eventos e participar do projeto de extensão sobre sustentabilidade e meio ambiente com Mici.

A Cecir por me receber no projeto de extensão de programação, com uma equipe maravilhosa e divertida, Rave, Pedro, Helô, Bruno e Andressa.

A Glauciane que foi uma mãe para mim, Julio e Eliel no nosso projeto de extensão sobre forno e secador solar.

A Creuza, por liderar o CaririTec tão bem e ter feito parte da equipe e participado do projeto de extensão, projeto voltado para a visibilidade feminina nas exatas.

Ao professor Aldre, por ser o primeiro a me dar uma oportunidade de fazer pesquisa, a qual foi em conjunto com Yasmin, como sua voluntária, ao permitir nosso estágio no laboratório Liep, comandado por Nilton, por sua paciência e dedicação a tudo que se compromete.

A Nahum que viabilizou a aventura de fazer um estágio no UFPE e aos Antônio (Coelho e Demóstenes) da instituição que receberam eu e Yasmin na instituição, nos ensinaram muito sobre a vida e profissão durante nosso estágio na instituição!

A Norma, técnica da UFCG que me ajudou em diversas etapas da minha graduação, principalmente durante o meu projeto com a professora Ilza.

Margareth e Maristela por me receberem e viabilizarem meu estágio no INSA (Instituto Nacional do Semiárido). Obrigada pela paciência e carinho!

A pró Ilza, minha mãe dentro da universidade, pela sua bondade, inteligência e doçura sem igual, sou grata pela oportunidade de trabalhar no laboratório Laqua, ser bolsista num trabalho lindo de iniciação científica sobre manejo de videiras, monitora de química orgânica, e pela orientação de vida!

Ao professor Rafael, meu pai dentro da universidade, pela sua humildade, empatia e singular, grata por ter me permitido trabalhar como voluntária no projeto do Zika vírus ao lado de Mona, me agradecer novamente como bolsista do projeto de iniciação científica sobre o SARS-CoV-2, voluntária na disciplina TEBB (Doenças parasitárias e zoonóticas). Mais que professor, é um amigo querido!

A minha orientadora, Ana Verônica, pela chance de ser monitora de Genética Molecular, além de me ajudar, apoiar e incentivar quando precisei até a entrega do trabalho de conclusão de curso. Além de Horace, meu co-orientador, que me deu todo suporte necessário. Agradeço de coração aos dois!

Agradeço também a todos que fizeram parte comigo do Centro Acadêmico, Empresa Júnior, a tentativa de implementação do IEEE no campus, a tentativa de eleição no Diretório do Centro Estudantil e ao Crea-Pb, por ter participado como representante estadual do meu curso.

Aos amigos que Sumé e a Paraíba me deram: Yasmin, amiga de faculdade e da vida que ganhou meu coração, topa tudo; Matheus, pela sensibilidade e companheirismo; Filipe Marinho por estar sempre pertinho de mim e me escutar; Felipe França, pessoa negativa mais proativa da vida, companheiro da faculdade e dos rolês; Darlan, que trocou de curso mas sempre nos mantivemos próximos; Romildo, pelo carinho e amizade; Elivelton, pela companhia e serenidade que me transmite; Damis, pelas longas conversas com informações desnecessárias. E aos outros que não citei, agradeço também por tudo!

A minha família, que me apoiou a todo instante, aguentou choros, inseguranças e tudo que vem com a maturidade adulta. Achibaldo (painho) te amo, hoje te agradeço por ter me aberto os olhos à um novo horizonte, ao me trazer para Sumé, além de todo apoio que me deu; Cremilda (mainha) te amo e obrigada por todos puxões de orelha, companhia e apoio em tudo que faço; Esther (irmã), parceira da vida e fofocas, minha gêmea mais nova, pessoa que alegra e me irrita todos os dias; Jalom (irmão) que sempre esteve longe, porém presente alegrando todo mundo, bora passar no concurso juntos!; Adna (irmã), apesar das diferenças me ajudou em tudo o que pôde, obrigada por estar sempre perto; e a Oscar (vô) e Daura (vó), dos quais morro de saudades de escutar, fazer almoço, ir tomar açaí e abraçar. Amo todos vocês até o infinito e volta!!! Além deles, agradeço de coração a todos da minha grandiosa e unida família!

## RESUMO

A cotonicultura é uma plantação muito importante no Brasil, detentora do quinto lugar de maior produtora e terceira maior exportadora mundial. Nacionalmente, o plantio de algodão herbáceo concentra-se no centro-oeste e nordeste, onde é majoritariamente cultivado por irrigação, seguido da técnica de sequeiro e sem rega. Os fatores climáticos, alta evapotranspiração e baixa pluviosidade, qualidade da água de irrigação, manejo, características do solo e espécie, são determinantes para existência e agravamento do estresse salino no plantio, sendo um dos mais importantes estresses abióticos, juntamente com o hídrico, a afetar o crescimento, produtividade e qualidade da fibra. Para isto, o melhoramento genético vegetal é a alternativa mais viável para redução das consequências geradas na planta pelo excesso de sais, como uso de produtos químicos na lavoura e perda quantitativa e econômica da produção. Com desenho de primers *in silico* que reconhecerão os domínios conservados no gene relacionados com a tolerância e/ou resistência, viabilizando a manipulação genética para criação de cultivares mais adaptadas a este fator. Assim, a partir da bioinformática, o presente trabalho, analisou taxonomicamente e físico-quimicamente três proteínas de interesse selecionadas, segundo triagem prévia, realizando desenho dos primers, validação destes e seleção dos dez primers mais estáveis, seis das proteínas e quatro degenerados, os quais são seguros, específicos e estão de acordo com padrões qualitativos de referência. De extrema relevância para área experimental, o conhecimento do perfil, ação das proteínas e desenho dos marcadores moleculares realizados, o caminho para se minimizar os problemas gerados pela salinidade ao cultivo é palpável.

**Palavras-chave:** cotonicultura; salinidade; primers.



## ABSTRACT

Cotton culture is a very important plantation in Brazil, holding the fifth place as the largest producer and the third largest exporter in the world. Nationally, the planting of herbaceous cotton is concentrated in the Midwest and Northeast, where it is mostly cultivated by irrigation, followed by the rainfed technique and without irrigation. Climatic factors, high evapotranspiration and low rainfall, irrigation water quality, management, soil and species characteristics, are determinant for the existence and aggravation of saline stress at planting, being one of the most important abiotic stresses, along with water, affect fiber growth, productivity and quality. For this, plant genetic improvement is the most viable alternative to reduce the consequences generated in the plant by excess salts, such as the use of chemical products in the crop and the quantitative and economic loss of production. With design of *in silico* primers that will recognize conserved domains in the gene related to tolerance and/or resistance, enabling genetic manipulation to create cultivars more adapted to this factor. Thus, from bioinformatics, the present work analyzed taxonomically and physico chemically three selected proteins of interest, according to previous screening, designing the primers, validating them and selecting the ten most stable primers, six from the proteins and four degenerates, the which are safe, specific and in accordance with qualitative reference standards. Of extreme relevance for the experimental area, the knowledge of the profile, action of proteins and design of molecular markers carried out, the way to minimize the problems generated by salinity in cultivation is palpable.

**Keywords:** cotton culture; salinity; primers.

## LISTA DE TABELAS

<b>Tabela 1</b>	Informações citogenéticas a respeito dos grupos genômicos do algodoeiro..	<b>20</b>
<b>Tabela 2</b>	Áreas de aplicação do estudo in silico.....	<b>27</b>
<b>Tabela 3</b>	Informação genômicas das espécies do gênero <i>Gossypium</i> depositadas no banco de dados, NCBI.....	<b>33</b>
<b>Tabela 4</b>	Informações sobre as proteínas taxonômicas com função sobre estresse salino em plantas do gênero <i>Gossypium</i> , informações retiradas do NCBI.....	<b>38</b>
<b>Tabela 5</b>	Características físico-químicas das proteínas base usadas na ferramenta BlastP para coleta das outras 15 sequências do servidor NCBI, informações obtidas no servidor ProtParam.....	<b>39</b>
<b>Tabela 6</b>	Primers da proteína pertencente ao domínio Rop_like.....	<b>61</b>
<b>Tabela 7</b>	Primers da proteína pertencente ao domínio P450.....	<b>63</b>
<b>Tabela 8</b>	Primers da proteína pertencente ao domínio bZIP.....	<b>65</b>
<b>Tabela 9</b>	Desenho de primers degenerados das três proteínas dos domínios Rop_like, P450 e bZIP após alinhamento das sequências.....	<b>66</b>

## LISTA DE FIGURAS

<b>Figura 1</b>	Alinhamento de sequência múltipla da proteína de ligação a GTP semelhante, a rac ARAC8, da espécie de algodoeiro <i>Gossypium hirsutum</i> no ClustalW.....	40
<b>Figura 2</b>	Dados do alinhamento do domínio Rop_like no NCBI.....	41
<b>Figura 3</b>	Alinhamento de sequência múltipla da proteína citocromo P450 tipo CYP82D47, da espécie de algodoeiro <i>Gossypium hirsutum</i> no ClustalW.....	41
<b>Figura 4</b>	Dados do alinhamento do domínio da família 82 do citocromo P450 no NCBI.....	43
<b>Figura 5</b>	Alinhamento de sequência múltipla da proteína de fator de choque térmico semelhante a HSF8, da espécie de algodoeiro <i>Gossypium hirsutum</i> no ClustalW.....	43
<b>Figura 6</b>	Dados do alinhamento do domínio da superfamília bZIP no NCBI.....	45
<b>Figura 7</b>	Árvore filogenética bootstrap consensus da proteína do domínio Rop_like, obtida no MEGA-X.....	47
<b>Figura 8</b>	Árvore filogenética bootstrap consensus da proteína do domínio citocromo P450, obtida no MEGA-X.....	48
<b>Figura 9</b>	Árvore filogenética bootstrap consensus da proteína do domínio bZIP, obtida no MEGA-X.....	49
<b>Figura 10</b>	Árvore filogenética bootstrap consensus comparativa dos alinhamentos múltiplos das três proteínas, Rop_like, P450, HSF8, obtida no MEGA-X...	52
<b>Figura 11</b>	Análise tridimensional da proteína dos domínios Rop_like. P450, bZIP, respectivamente, A, B e C.....	53
<b>Figura 12</b>	Perfil de hidrofobicidade em gradiente de cor, regiões em azul mais hidrofílicas e em vermelho mais hidrofóbicas na proteína dos domínios Rop_like. P450, bZIP.....	55
<b>Figura 13</b>	Gráfico de Ramachandran da proteína dos domínios Rop_like. P450, bZIP, respectivamente, A, B e C.....	56
<b>Figura 14</b>	Análise do Z-score da proteína dos domínios Rop_like. P450, bZIP, respectivamente, A, B e C.....	57
<b>Figura 15</b>	Gráfico de energia dos stories da proteína dos domínios Rop_like. P450, bZIP, respectivamente, A, B e C.....	58

<b>Figura 16</b>	Domínio da sequência conservada da proteína referente ao Rop_like, com localização 8 - 180, juntamente com os primers presentes na plataforma neste intervalo, obtido no NCBI.....	<b>61</b>
<b>Figura 17</b>	Domínio da sequência conservada da proteína referente ao P450, com localização 3 - 522, juntamente com os primers presentes na plataforma neste intervalo, obtido no NCBI.....	<b>63</b>
<b>Figura 18</b>	Domínio da sequência conservada da proteína referente ao bZIP, com localização 167 - 215, juntamente com os primers presentes na plataforma neste intervalo, obtido no NCBI.....	<b>65</b>
<b>Figura 19</b>	Alinhamento múltiplo comparativo entre as três proteínas, obtido no servidor Clustal Omega.....	<b>81</b>
<b>Figura 20</b>	Domínio da sequência completa conservada da proteína referente ao Rop_like juntamente com os primers presentes na plataforma, obtido no NCBI.....	<b>88</b>
<b>Figura 21</b>	Domínio da sequência completa conservada da proteína referente ao P450 juntamente com os primers presentes na plataforma, obtido no NCBI.....	<b>89</b>
<b>Figura 22</b>	Domínio da sequência completa conservada da proteína referente ao bZIP juntamente com os primers presentes na plataforma, obtido no NCBI.....	<b>90</b>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>12</b>
<b>2</b>	<b>OBJETIVO.....</b>	<b>14</b>
2.1	OBJETIVO GERAL.....	14
2.2	OBJETIVOS ESPECÍFICOS.....	14
<b>3</b>	<b>REFERENCIAL TEÓRICO.....</b>	<b>15</b>
3.1	ASPECTOS GERAIS.....	15
3.2	ESTRESSE SALINO.....	23
3.3	BIOINFORMÁTICA.....	26
<b>4</b>	<b>METODOLOGIA.....</b>	<b>31</b>
4.1	TRIAGEM E OBTENÇÃO DAS SEQUÊNCIAS.....	31
4.2	ANÁLISE DAS SEQUÊNCIAS.....	31
4.3	ANÁLISE FILOGENÉTICA E AGRUPAMENTO DAS SEQUÊNCIAS.....	31
4.4	PREVISÃO, AVALIAÇÃO E VALIDAÇÃO DA ESTRUTURA TERCIÁRIA DOS MODELOS.....	32
4.5	PRODUÇÃO E VALIDAÇÃO DE PRIMERS.....	32
<b>5</b>	<b>RESULTADOS E DISCUSSÃO.....</b>	<b>33</b>
<b>6</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>68</b>
	<b>REFERÊNCIAS.....</b>	<b>69</b>
	<b>ANEXO.....</b>	<b>81</b>

## 1 INTRODUÇÃO

A família Malvaceae, gênero *Gossypium*, é referente popularmente ao algodoeiro. Cultivo datado a mais de 4 mil anos (SALDANHA; BELO, 2016; DE SOUSA, 2010) com a finalidade de utilização de sua fibra para o setor têxtil e sementes para alimentação animal e extração de óleos vegetais (BLANCO, 2008), entre outras finalidades ligadas a bioenergia e outros setores. Apresenta neste gênero, mais de 50 espécies, sendo quatro as de maior valor econômico e cultivo em grande escala, *G. arboreum L.* e *G. herbaceum L. diplóides*, *G. hirsutum L.* e *G. barbadense L. tetraplóides*, respectivamente em pares, nativas do Velho Mundo e Novo Mundo (DE SOUSA, 2010).

Nacionalmente, o algodoeiro é uma cultura de bastante importância, sendo o quinto maior produtor mundial na safra de 2017/2018, somando com a produção dos quatro países à sua frente 76,66% da produção mundial (DA SILVA, 2020). Nacionalmente, pela última atualização do IBGE, a produção do algodão herbáceo em 2021 tem possibilidade de ultrapassar a do ano anterior, apresentando na primeira safra já quase 6 milhões de toneladas produzidas até abril de 2021 em relação ao total de 7 milhões do anterior, 2020 (IBGE, 2021).

Regionalmente, no Brasil, o centro-oeste concentra 73,3% da produção brasileira, seguido pelo nordeste com 29,3%, sudeste com 3%, norte com 0,3, já o sul não tem dados, de acordo com a produção do ano de 2020 (IBGE (2021). Tais estados, encontram-se no clima semiárido e vegetação caatinga, em sua maior parte, exceto o estado do Maranhão que possui majoritariamente o bioma cerrado, juntamente com os estados do centro-oeste.

Esta área se caracteriza pelas altas temperaturas, eventos de estiagem prolongados, baixa pluviosidade, solos com baixo potencial de retenção de água, alta infiltração e evapotranspiração que faz com que haja uma maior concentração de sais e nutrientes neste solo (HOFFMANN NETO, 1999; DE SENA, 2011; CORREIA, et al., 2011). E conseqüente uma presença maior de poços artesianos e amazonas, nestes locais, pois a acumulação aquífera se dá subterraneamente, principalmente na caatinga, já que o cerrado é o “berço das águas brasileiro”, possuindo muitas nascentes superficiais em seu território (LIMA, 2011).

Apesar da cultura ser adaptada a temperaturas entre 18 e 30°C, com ciclo relativamente curto, podendo ter cultivares com 100-120 dias, alta taxa de produtividade por hectare, no caso 3mil Kg/ha (DE AZEVEDO, SILVA, 2007), este ambiente promove à cultura de algodão um certo nível de estresse abiótico. Normalmente, é usado o regime de sequeiro para plantio, porém apesar de autores indicá-lo (DE AZEVEDO, SILVA, 2007; MATTOS, et al., 2020; MOTA, 2014), pode não compensar o uso desta técnica pela ocorrência de veranicos durante a fase

fenológica, pois o solo pode não manter a umidade necessária para o desenvolvimento da planta. Assim a cultura do algodoeiro por irrigação ganha mais espaço pela maior segurança e produtividade (MOTA, 2014), reforçado por 60% de seu cultivo ser por esta técnica (BLANCO, et al., 2011; PEREIRA, 2012).

Pelo alto teor de sais nas fontes aquíferas usadas na irrigação aliada a um solo com baixa retenção, que gera acúmulo de sais, faz com que a cultura esteja exposta a estresses salinos que irá comprometer a médio e/ou longo prazo fisiologicamente, bioquimicamente, na absorção de nutrientes, e consecutivamente seu desenvolvimento e produtividade (AMORIM, et al, 2010), além do meio ambiente, necessitando de um manejo que minimize as consequências geradas pelo alto teor de salinidade no ambiente e planta, assim como de métodos biotecnológicos que propicie um algodoeiro mais resistente a este fator abiótico.

Devido a esta necessidade, uma alternativa viável é o uso do melhoramento genético vegetal para propiciar cultivares mais adaptadas ambientalmente ao excesso de sais, acarretado devido ao seu cultivo em regiões com este perfil (SOARES FILHO, et al., 2016). Com isto, uma ferramenta que viabiliza este estudo é o desenho de primers *in silico* para uma análise específica e profunda do gênero e genes taxonômicos associados ao estresse salino, para se ter maior precisão na fase experimental de manipulação gênica.

## 2 OBJETIVO

### 2.1 OBJETIVO GERAL

Estudo genético de proteínas associadas ao estresse salino no gênero *Gossypium* para desenho de primers.

### 2.2 OBJETIVOS ESPECÍFICOS

- Busca em banco de dados dos genomas já sequenciados do gênero e de quais espécies são;
- Triagem de proteínas de interesse;
- Análise filogenética, físico-química e estrutural das sequências das proteínas;
- Desenho dos primers referente às proteínas de interesse.



### 3 REFERENCIAL TEÓRICO

#### 3.1 ASPECTOS GERAIS

O algodoeiro branco tem sua origem na Índia, continente asiático, de acordo com documentos históricos (IFRN, 2021), já o naturalmente colorido tem como seu berço o continente americano (SALDANHA; BELO, 2016; DE CARVALHO, 2016), pelos povos Incas no Peru datada a 4500a.C. (NASCIMENTO, 2019), quando foi domesticada por indígenas, os quais selecionaram o fenótipo desejado artificialmente de acordo com marcadores morfológicos por séculos chegando nas variedades e espécies existentes hoje, tendo como descendente de germoplasma pré-colombiano todo algodão colorido cultivado pelo mundo. A domesticação no continente europeu se deu pelos árabes e responsáveis pela comercialização a mais de 4 mil anos (NASCIMENTO, 2019).

A partir da época das grandes navegações, com a descoberta do caminho para Índia no século XVII, esta fibra começa a ter maior visibilidade na Europa e se estabelecendo no mercado a partir da primeira revolução industrial (século XVIII), onde predominantemente se usava lã para produção têxtil (BERGAMIN, 2018).

No Brasil e América como todo, antes da vinda dos europeus para o continente americano, já havia plantio, melhoramento e uso do algodão na têxtil, mas em 1750, os portugueses começaram a produção comercial para exportação da cotonicultura, o qual possuía fibras longas e com produtividade não muito atrativa, no estado do Maranhão no nordeste, com o intuito de reduzir a dependência deste insumo da Inglaterra (BARBOSA, 2005), perdurando do século XVIII até o XX.

Até a década de 1980, antes do declínio, os estados da Bahia e Ceará eram os maiores produtores de algodão arbóreo do nordeste (Embrapa, 2003), que junto com a região sul e sudeste compunham majoritariamente a produção nacional (PETRUCCI, 2011). Porém, durante a mesma década passou por uma decadência na produtividade devido a praga do bicudo e baixa competitividade com o importado, que gerou uma queda do produto no nordeste e maior demanda externa. Somente em meados de 1990, com o deslocamento da área de plantio para áreas do bioma cerrado, centro-oeste, com cultivares mais adaptadas ao clima e mecanização que a cotonicultura se reergueu no mercado interno.

A elevada produtividade em 1997 estimulou seu desenvolvimento, passando de uma década para outra de segundo país maior importador para o quinto maior produtor e terceiro maior exportador desta cultura, perdendo somente dos Estados Unidos, Índia, China, Paquistão, Índia na produção, e para os dois primeiros como exportador (NASCIMENTO, 2019;

ABRAPA, 2019). O fato deste sucesso se deve também pela possibilidade de uso na rotação do plantio de soja afora incentivo por programas governamentais (PETRUCCI, 2011).

Esta planta é cultivada comercialmente em mais de 65 países pelo mundo (DE SOUSA, 2010), tendo como as principais espécies de importância pela fibra para plantio, *G. hirsutum* L., conhecida por algodoeiro herbáceo no Brasil e responsável por 90% da produção mundial, seguida pela *G. barbadense* L., distinta pela alta qualidade da fibra, também tetraplóide e de origem no Novo Mundo, em adição a *G. arboreum* L. e *G. herbaceum* L., duas últimas de menor área de plantio, normalmente em pequena escala, haplóides e do Velho Mundo (FREIRE, et al., 2003; DE SOUSA, 2010).

A mudança do homem nômade para o sedentário, a partir do período neolítico, fez com que estes começassem a investir na agricultura e pecuária, antes sobreviviam com o que ambiente em que estavam oferecia de alimento, tornando animais e vegetais domesticados. Este processo, fez com que com o tempo fossem procurando algumas características mais relevantes, como repressão do sistema de dispersão de sementes, mudanças fisiológicas (alometria e condensação), germinação uniforme e mais acelerada, sincronismo no processo de floração e maturação, modificações bioquímicas, aumento do tamanho de órgãos, ciclo de vida e sistemas de hibridação (EVANS, 1993; apud MANTOVANI, et al., 2008).

Para o algodoeiro, especificamente, a domesticação possibilitou para planta modificações evolutivas de interesse humano em relação a produtividade e ciclo de vida, esta, antes perene, hoje anual e com a possibilidade de duas safras por ano (SILVA, et al., 2009), além de propiciar cultivo e adaptabilidade em zonas áridas e semiáridas, o que anteriormente restringia a regiões temperadas (MANTOVANI, et al., 2008), além de ter adquirido resistência à seca (VIDAL NETO; FREIRE, 2014) e tolerância a salinidade, podendo esta última variar de acordo com as espécies e fase do desenvolvimento da planta (BARROS, 2011).

Caracteristicamente, o gênero *Gossypium* da família Malvaceae (COUTO; RAMOS; DE CASTRO, 2020), é uma planta arbustiva de crescimento indeterminado e com grande variabilidade morfológica entre suas espécies, assim como o ciclo de vida que oscila de cultivar para cultivar, precoce ou tardiamente (de 130 dias até mais de 170 dias), reprodutivamente tem um sistema intermediário em relação as plantas autógamas e alógamas devido seu potencial de fecundação cruzada e autopolinização (DE SOUSA, 2010; VIDAL NETO, FREIRE, 2014; INSTITUTO AGRONÔMICO DE CAMPINAS, 2021), classificada como C3 em relação a usa fotossíntese, precisa de boa umidade no solo no período de germinação (AGEITEC, 2017;

PEREIRA, 2008), oleaginosa, com plantio em solos com pH 5,5 e 6,5, baixa inclinação, altitude até 1500m (RIBEIRO, et al., 1999).

A partir do meristema apical são formados os quatro órgãos principais do algodão, raízes, caule, folhas e flores (NASCIMENTO, 2019). Esta cultura tem quatro fases de crescimento bem definidas: emergência, com início lento (25 a 30 dias), mas após a emergência seu desenvolvimento se torna mais acelerado (CARVALHO, 2007); floração; abertura dos frutos; e, por fim, a colheita (PEREIRA, 2008). Visando durante este período, principalmente da abertura dos frutos ao fim da safra, mantê-la saudável para garantir a produtividade esperada no plantio, a qual sofre grande influência da temperatura e radiação solar, disponibilidade de nutrientes e água, além de fatores externos bióticos e abióticos, causas que se alteram de acordo com o tempo de semeadura e manejo (DA SILVA, 2020).

Em relação ao período de floração, especialmente no nordeste, ocorre a abertura dos botões após 30 dias em cultivares de ciclo de vida precoces, com flores após 15 dias, as quais se mantêm até abertura dos capulhos (70 a 80 dias). Com abertura e fecundação matinal (período de 30 a 40 dias), esta que antes deste processo se caracterizava pela cor de suas pétalas variando do creme ao amarelo, após ficam rosadas seguidas secam, caem e dão início a formação do fruto, imaturo chamado de maçã e após 90 dias torna-se maduro e se chama capulho (DE SOUSA, 2010; VIDAL NETO, FREIRE, 2014).

Outros atributos a estes órgãos reprodutivos, é que possuem pedicelo, isoladas e completas com três brácteas de coloração verde abarcando-a, possui cinco pétalas dispostas juntas em uma, e possui órgãos reprodutores de ambos os sexos, hermafrodita. Destarte, deve-se prestar atenção ao período reprodutivo das espécies, observando que devem ser todas tetraplóides ou diplóides, para fazer com que o período de floração entre elas sejam uniformes e não gere perdas na hibridação (WENDEL; CRONN, 2003), evitando, assim, um manejo inadequado da plantação.

Adicionalmente, o órgão radicular é formado por uma raiz principal vertical e ramificações à esta lateralmente na horizontal, possibilitando a planta maior área de superfície para absorção de nutrientes. A principal é pivotante, encontrando-se nos 20cm a 2,5m de profundidade no solo, porém cerca de 80% da massa total da raiz localiza-se nos 20cm mais superficiais subterraneamente. Já o caule, ereto com 0,8 a 2,5m de altura, com formato cilíndrico à ligeiramente pentangular ou quadrangular, caracteristicamente sublenhoso, e deste ramo de crescimento vertical tem função monopodial, ou seja, possui uma única gema apical, a qual dará origem as ramificações laterais com função simpodial, onde várias gemas participam

de sua formação, nos diversos eixos em que se desenvolve (NASCIMENTO, 2019; DE SOUSA, 2010).

Quanto sua fibra, principal fonte econômica proveniente da planta, tem sua qualidade atribuída a características intrínsecas e extrínsecas, a primeira tem sua expressão controlada geneticamente, como resistência, comprimento e uniformidade deste, precocidade, porte da cultura (baixo ou alto), tolerância a fatores abióticos e bióticos, espessura, já a última, é influenciado pelo ambiente, armazenamento e colheita, manejo, nutrição. Assim, quanto maior, uniforme, grau de cor, fiabilidade, maior interesse há nessas propriedades produtivas e comercial (VIDAL NETO, FREIRE, 2013; FUZATTO, et al., 2014; ANDRADE et al., 2009). Também deve-se observar qual foco de comércio terá, do caroço ou pluma (FUZATTO, et al., 2014).

Em ambiente de estresse, de acordo com ARAÚJO (2018), para estresse hídrico, o algodoeiro produziu fibras qualitativamente superior nas fases iniciais de crescimento e frutificação, maçã. Tem a vantagem de ser hipoalérgica, podendo qualquer pessoa usar sem que haja alguma reação de sensibilidade na pele, podendo ser combinado com outros tipos de fibra (IFRN, 2021). O algodoeiro herbáceo, espécie que tem como porcentagem de semente, em média de 65% e 35% de fibras, é o mais importante na cotonicultura (NEVES, PINTO, 2011; DE ANDRADE JÚNIOR, 2009).

Para se ter um bom manejo do algodoeiro, é necessário se atentar a imprescindibilidade de se observar as suas exigências ambientais, diferenças entre cultivares, época correta de semeadura para que não haja perda na hibridação, a calagem e adubação que deve ser feita, qual técnica de plantio se encaixa melhor no local de plantio, tipos de pragas, doenças e seu controle, rotação recomendada de cultivo, distância e tamanho de poda, época de colheita e os custo versus produtividade na região que irá plantar (INSTITUTO AGRONÔMICO DE CAMPIINAS, 2021).

No nordeste, os tipos de plantio são de produção irrigada, não irrigada e de sequeiro, sendo mais extensiva nesta região do Brasil a primeira e no centro-oeste a segunda, devido a agricultura familiar majoritariamente versus mecanização e produção em grande escala (PORTO, et al., 2019; AGUIAR, MODESTO FILHO, DOS REIS, 2006). É uma planta com exigências ambientais para que se tenha um bom rendimento, solos ácidos, muito úmidos, rasos ou com alto grau de compactação não são bons para este cultivo, clima quente com temperatura variando entre 20 a 26°C, em média, altitude de 200 a 1000m, aumentando seu ciclo em mais de um mês, quanto maior for a elevação (DE OLIVEIRA JÚNIOR, et al., 2009; INSTITUTO AGRONÔMICO DE CAMPINAS, 2021).

As pragas podem assolar esta cultura em várias etapas do ciclo de vida da planta, no início de ser crescimento, como boca-do-ponteiro, percevejo castanho, cigarrinha, boca-da-raiz, pulgão, e na época de floração e frutificação, principalmente, como mosca branca, lagarta-das-maçãs, bicudo, lagarta rosada, ácaro rajado e branco, curuquerê, percevejo manchador, lagarta militar, rosada e *Helicoverpa armigera*. As quais podem ser controladas a partir do manejo com monitoramento, uso de feromônios, rotação de cultura, e principalmente, uso de cultivares resistentes, tolerantes (INSTITUTO AGRONÔMICO DE CAMPINAS, 2021; DA SILVA, 2015).

Já para doenças, pode ser infectado por mancha-angular, ramulose, *Fusarium*, *Verticillium*, nematóides, mosaico das nervuras, e outras não tão frequentes como *ramularia*, *cercospora*, *Trichoderma* e *Bacillus amyloliquefaciens*. O controle destas é a partir de medidas profiláticas ao destruir as plantas infectadas para que interrompa a disseminação, tratamento químico e algodoeiros resistentes ou tolerantes (INSTITUTO AGRONÔMICO DE CAMPINAS, 2021; SUASSUNA, SILVA, BETTIOL, 2019; LANGE, et al., 2018). É importante também a descontaminação de sementes, em alguns casos, com uso de substâncias químicas para o tratamento (HENNING, 2005).

As espécies de maior interesse econômico agregado são: a *G. hirsutum* L., originada do sul do México, com muita variedade de raças, latifolium, punctatum, marie-galante (mocó), a primeira mais cultivada como planta anual; *G. barbadense* L. de origem no Peru, destacando a variedade “rim-de-boi” pela qualidade da fibra e coloridos também, com evidências antigas de cores variando de azul, vermelho, roxo, bronze, verde, marrom, porém consta em banco de germoplasma hoje a marrom, laranja, creme, púrpura, cinza, amarela; *G. arboreum* L. e *G. herbaeum* L., respectivamente, originada e cultivada na África e cultivada na Ásia, a segunda asiática com cultivo na Índia, conhecidas pela alta qualidade da fibra, por em baixa produtividade (DE CARVALHO, 2016; INSTITUTO AGRONÔMICO DE CAMPINAS, 2021).

Além destas, tem a endêmica da caatinga, *G. mustelinum*, que possui grande importância ambiental e social, devido a necessidade de se conservá-la, *in situ e ex situ*, em bancos de germoplasma, trabalho realizado da Embrapa, pois em muitos locais ela não tem resistido ao ambiente, queimadas, desmatamentos para pastos e plantio de coco no litoral. Ocorre naturalmente em todo bioma, com enfoque a sua presença em áreas de preservação, mantendo sua conservação *in situ*, como em Tambaba na Paraíba e em Santa Cruz em Pernambuco. Seu perfil é de uma planta arbustiva de pequeno porte, capulhos pequenos que se desprendem e fazem a dispersão de suas sementes de forma natural, diferente de espécies como *G. hirsutum*

var. *latifolium*, *G. barbadense*, suas fibras são curtas, presença de línter de cor verde e marrom, doravante estudos de seu RNA ribossômico, foi identificado esta espécie como tetraplóide (HOFFMANN, et al., 2020).

As principais diferenças entre cultivares estão no porte da planta, rendimento quantitativo da fibra, resistência e tolerância, tempo de ciclo produtivo (INSTITUTO AGRONÔMICO DE CAMPINAS, 2021). A tendência tem sido a redução de seu tamanho para encurtar o ciclo de vida, torna-la mais produtiva e resistente a fatores ambientais, temperatura e umidade. Também, maior eficiência produtiva com aumento de frutos por planta, adaptação ao manejo, menor densidade foliar, maior detenção dos capulhos, voltada ao cultivo em grande escala e mecanizado, para plantio manual, é preferível de maior porte com mais capulhos e menor retenção da pluma, facilitando a colheita, ciclos tardios fornecendo mais colheitas (VIDAL NETO, FREIRE, 2013).

Dados citogenéticos do gênero *Gossypium* a respeito de grupos genômicos, realizado a partir do pareamento de sequências das espécies do algodoeiro (há mais de 50 espécies), mostrou duas categorias: os organismos diplóides ( $2n=2x=26$  cromossomos) e os alotetraplóides ( $2n=4x=52$  cromossomos), associando dados de duas espécies. Os diplóides formaram três subgrupos e o alotetraplóide somente um, respectivamente, como mostra a Tabela 1 (DE SOUSA, 2010, p. 20).

**Tabela 1** - Informações citogenéticas a respeito dos grupos genômicos do algodoeiro.

	Genoma A, B, E, F	Genoma A, D	Genoma C	Genoma D
Espécies	8	6	11	12
Local onde são encontradas	Afro-Arábia	América	Austrália	América
Informações adicionais	Duas espécies cultivadas: <i>G. herbarium</i> e <i>G. arborium</i> .	Duas espécies cultivadas: <i>G. hirsutum</i> e <i>G. barbadense</i> , as outras quatro são encontradas no Havaí, Brasil, Ilhas Galápagos e México	-	Dez espécies no México e duas no Peru e Ilhas Galápagos

Fonte: SOUSA (2010, p. 20).

Sua aplicabilidade é múltipla, desde a mais com mais atenção, fibra têxtil para produção de tecidos, como brim, llona, voil, algodãozinho, talagarça, tolha, cadarço, atoalhado, riscado, filó, creton, organdi, popeline, encerado, morim, cetim, cambraia, cretone, a uso para

alimentação animal como farinha e óleo para humana, produção de artesanato, cosméticos, área farmacêutica, fabricação de celulose, óleo essencial vegetal, biodiesel (DA SILVA, 2015; SEVERINO, et al., 2019; IFRN, 2021).

Agregado aos seus numerosas fins, hoje, distribuído o cultivo, algodão herbáceo, entre as regiões do centro-oeste com 73% da produção nacional, nordeste com 29,3%, sudeste 3% e sul com quantidade baixa, sem informações, de acordo com o censo das safras de 2020 (IBGE, 2021). A cultivar mais plantada nos estados nordestinos são os herbáceos, *Gossypium hirsutum* L. variedade latifolium Hutch., principalmente para pequenos e médios produtores (Embrapa, 2021), já no centro-oeste é mais voltado para culturas empresarias com alto nível tecnológico (INSTITUTO AGRONÔMICO DE CAMPINAS, 2021).

Importante socioeconomicamente, à nível nacional e internacional, nas safras de em 2017/2018, como quinto maior produtor, reteve 76,66% da produção mundial (DA SILVA, 2020), cultivo que movimentou o valor de 12 bilhões de dólares mundialmente, desde o plantio aos seus subprodutos (DA SILVA, 2020).

Devido a importância da cotonicultura e o tamanho do Brasil, as demandas dentro do cenário nacional relacionada a melhoria de cultivares, pode-se ter três linhas de pesquisa, a primeira com o plantio no cerrado (região centro-oeste predominantemente) em busca do beneficiamento da qualidade da fibra intrínseca e extrínsecamente, maior resistência fitossanitária e aproveitamento no descaroçamento; a segunda com o cultivo no sul e sudeste à procura de maior produtividade, controle fitossanitário, precocidade, maior rendimento e qualidade da fibra; e terceiro o nordeste tem como necessidade maior precocidade e produtividade, fibra mais longa e colorida, sementes sem gossipol, tolerância a fatores abióticos e bióticos inerentes ao semiárido (solo, clima, salinidade, seca) para viabilizar a produção orgânica (VIDAL NETO, FREIRE, 2013). Apesar de demandas específicas, é visível a busca compartilhada pela resistência/tolerância biótica e abiótica (controle fitossanitário), maior rendimento e qualidade da fibra.

Iniciado o melhoramento brasileiro no algodoeiro em 1921, já existia uma rede de pesquisa e melhoramento do gênero, em 1930, situado no nordeste, que foi estimulando o campo da ciência internamente para o estudo desta cultura com decaimento na década de 1980 com a introdução do bicudo nas plantações. Voltou a ser estimular com a reintrodução do plantio no centro-oeste e reestabelecido com a parceria entre Fundação Mato Grosso e Embrapa Algodão, em 1995, que desde então tem ascendido obtivendo sucesso no desenvolvimento de diversas cultivares (DE CARVALHO, 2012), principalmente voltadas as demandas ditas, somando de 1998 a 2010 o registro de cerca de 115 novas cultivares entre as de fibra branca e

colorida (VIDAL NETO, FREIRE, 2013). À exemplo, está o programa Algodão Brasileiro Responsável (ABR), 2012, que com a ONG Better Cotton Initiative (BCI), 2013, se propôs a melhorar a produção de sua fibra, a qual teve sua primeira safra em 2011 (ABRAPA, 2017).

O processo do melhoramento usa de mecanismos adaptativos visando o comportamento de plantas resistentes e tolerantes possuem em contato com o estresse em questão, minimizando os efeitos deste em seu sistema fisiológico, morfológico e bioquímico (PEREIRA, et al., 2012 (a); NOGUEIRA, et al., 2006; AZEVEDO NETO et al., 2010), assim como precisa usar normas de classificação, tal como a High Volume Instrument (HVI) para análise qualitativa da fibra (NASCIMENTO, 2019).

A estrutura genética, alvo do melhoramento genético, consiste no estudo da variação genética e manutenção desta dentro das espécies, obtida evolutivamente por seleção natural e domesticação, por migração geográfica, deriva genética, mutação, assim como também é influenciada pela forma de reprodução do organismo, fluxo gênico e tamanho da população (DE MENEZES, 2009), para usar este conhecimento genômico no entendimento da expressão, função de genes, evolução, dinâmica populacional, taxonomia (classificação de organismos em decorrência das relações evolutivas, identificando espécies e sua proximidade hereditária) (BICUDO, 2004) e manipulação destes.

Como o algodoeiro, reprodutivamente, é parcialmente alógamo e autógamo, suas linhas não são puras, devido seu grau de heterozigose. Assim para os programas de melhoramento, é necessário a adaptação dos métodos usados em plantas alógamas e autógamas em algumas etapas, em outras como hibridação, avaliação e seleção não é preciso. Tais métodos podem ser por seleção genealógica, vigor híbrido, recorrente (usado na Embrapa), hibridação, retrocruzamento, massal. Para o aumento da heterozigose é necessário uso da hibridação como forma de explorar as possibilidades genéticas de tal organismo mais a fundo, podendo os indivíduos transgênicos provenientes do estudo ser oriundos de um só organismo vegetal ou de diversas linhagens do gênero (VIDAL NETO, FREIRE, 2013; CARDOSO, 2018).

Assim, com o advento de técnicas de PCR, conhecimento do código genético a partir do mapeamento, criação de marcadores moleculares mais seguros na década de 1980 (genético com uso de DNA ou RNA e enzimáticos), o que antes era a partir de morfológicos os quais são influenciáveis pelo ambiente, limitado numericamente e às características qualitativas. Com o uso de marcadores, pode-se selecionar com primers locais específicos da sequência para estudo, além de ser possível pelo enorme acervo de dados genéticos em bancos de dados, desenhá-los *in silico*, tornando mais segura a identificação da área correta objetivada (DE MENEZES, 2009; TURCHETTO-ZOLET, 2017; DE ALMEIDA, 2007; SILVEIRA, et al., 2016).



### 3.2 ESTRESSE SALINO

A salinidade é uma realidade da agricultura em todo o mundo (LACERDA, 2016) e é um dos fatores que gera maior preocupação na atualidade (SCHOSSLER, et al., 2012). Fenômeno ocorre em decorrência de fatores, como a geomorfologia do local, onde o solo apresenta caracteristicamente uma grande quantidade de nutrientes e baixo índice de retenção de águas, aliada com uma evapotranspiração maior que a taxa de pluviosidade (RIBEIRO, 2010; ROACHE, BAILEY, BOON, 2006.; FREIRE & FREIRE, 2007), que provoca deficiência da lixiviação do solo e aumenta a concentração destes nutrientes no solo.

Culturas com baixa resistência ao estresse salino (FERREIRA et al., 2001) é outra causa, além da irrigação com aquíferos de baixa qualidade, como uso água salobra para rega do plantio, uma realidade no agronegócio e abastecimento de casas cada vez mais comum (SOARES, 2007) devido a variabilidade temporal e local de fatores hidrogeográficos (D'ALMEIDA et al., 2005), características essas, encontradas principalmente em áreas semiáridas e áridas, como o nordeste brasileiro (LACERDA, 2016; LIMA, 2016). Outros fatores que influenciam seu aparecimento e/ou agravamento, é o manejo incorreto da irrigação, fertilizantes, descanso do solo, rotação de culturas nas entressafras (D'ALMEIDA et al., 2005; EPSTEIN & BLOOM, 2006), assim como, o que acontece é que esta tendência encontrada em acumular sais, presente no mesmo espaço geográfico.

No Brasil, como um todo, estima-se que mais de 17 milhões de hectares estejam enfrentando problemas devido a ocorrência em excesso de sais durante a plantação (RIBEIRO, FREIRE, MONTENEGRO, 2003), e que as maiores parcelas de áreas salinas estejam no nordeste, acumulando 57% do total de área de cultivo no semiárido (HOLANDA, et al., 2007; RIBEIRO, et al., 2009), respectivamente, correspondendo aos estado da Bahia (44%) e Ceará (25%), do total nacional, como os mais assestados, em pesquisa mais antiga (GHEYI; FAGEIRA, 1997).

No nordeste, a água de irrigação, normalmente com salinidade e/ou sodicidade dos solos elevada (SOUSA JÚNIOR, FERNANDES, GHEY, 2008; SUASSUNA, AUDRY, 2009), aliadas as características prévias do solo e resistência da cultura, são fatores determinantes no plantio e manejo a ser adotado em seu cultivo, gerando, possivelmente, como efeito, tanto no solo como nas plantas, em casos mais moderados causa alterações na absorção de nutrientes, e em graves, tornam-se tóxicas em certos níveis, provocando a morte do vegetal e até desertificação do solo. Pois, ao se adicionar mais sais com a rega, fertilizantes e agrotóxicos, afeta a vigorosidade vegetal e solo em níveis antieconômicos para o plantio (MOTA, 2014).

Para as plantas, é reflexo de duas causas: componente osmótico pela redução hídrica interna e alta concentração de solutos, e o iônico pela relação de instável que a presença excessiva dos íons  $\text{Na}^+$  e  $\text{Cl}^-$  geram em outros nutrientes (LIMA, 2016). Este despropósito de sais, principalmente por íons de Na (sódio), Cl (cloro) e cloretos, acima do tolerado pela planta independente de sua resistente ou não a este fator abiótico, provocam redução do seu desenvolvimento, distúrbios fisiológicos (FARIAS et al., 2009), redução da expansão superficial da folha, diretamente relacionada e muitas espécies a fotossíntese e absorção de nutrientes (DA COSTA, 2014).

Esta situação promove uma diminuição do potencial osmótico dentro do vegetal afetando a concentração iônica, trocas gasosas na fotossíntese por fechamento dos estômatos (MUNNS, et al., 2002; FREITAS, 2006; CHE-OTHMAN, MILLAR, TAYLOR, 2017), absorção de outros nutrientes, como o zinco, ferro, cobre, manganês, boro relacionados ao aumento do pH no solo, tornando baixa a biodisponibilidade destes, a absorção radicular de água e manutenção interna desta no vegetal (DIAS, BLANCO, 2010; FREITAS, 2006), assim como baixa concentração de K (potássio), pois K e Na competem por sítios de absorção (FERNANDES, et al., 2002). Outro ponto é a toxicidade na redução da retenção de fitomassa, síntese de moléculas do metabolismo secundário de proteção aos fatores externos, problemas na osmorregulação e integridade das membranas celulares (SCHOSSLER, et al., 2012).

Um mecanismo de interesse é que plantas tolerantes possuem para driblar estas altas concentrações em seu ambiente de cultivo para que não afete a mesma rigorosamente. Um destes processos adaptativos, está na estratégia da planta ao absorver nutrientes conseguir expulsar os cátions de sódio por glândulas salinas (YUAN, et al., 2016), realizando o controle osmótico adequado para seu funcionamento e evitando acúmulo deste em seu limbo foliar (SERTÃO, 2005; MUNNS, et al., 2002).

Outra possibilidade é a adaptação com quantidades elevadas de sais em seu organismo e baixo potencial hídrico, comum as plantas halófitas, que usam vacúolos em seu interior celular da região das folhas para reter todo os sais absorvidos e controlando, assim, a osmose (SILVA, et al., 2000); a partir da produção de moléculas neutras para estabilização de membranas e proteínas preservando o ambiente interno da desidratação em grandes concentrações, e perda de água em baixa (YANCEY, 2005), as quais podem ser a valina, prolina, isoleucina, ácido aspártico, sacarose, frutose, manitol, glicose, pinitol, inositol, betaína entre outras biomoléculas (PARIDA; DAS, 2005). Plantas de interesse para fitorremediação por reduzir a concentração deste nutriente do solo, à exemplo disto tem as plantas presentes em manguezais e *Atriplex numulária* (SANTOS, et al., 2011).

Este sistema de produção de moléculas neutras é extremamente relevante, a exemplo, do aminoácido prolina, não tóxico, relacionado com função de osmoproteção celular, muito estudado pela sensibilidade em resposta ao estresse abiótico e sintetizado a partir do glutamato (SZABADOS; SAVOURÉ, 2009), desencadeando o aumento da resistência da planta ao modular funções mitocondriais, auxiliar na expressão de genes, estabilizar membranas para que não haja diminuição da concentração interna de eletrólitos, manter a concentração de ROS (espécies reativas de oxigênio) dentro dos limites, proporcionando a mínima perda de função possível a planta (KAUR & ASTHIR, 2015).

Já o óxido nítrico (NO), é um radical livre lipofílico, em forma de gás, que atua como mensageiro intracelular em processos desde a germinação da planta até defesa contra estresses bióticos e abióticos, como a hídrico e salinidade (VITAL, et al., 2008), respectivamente, e auxilia no processo reprimir danos oxidativos nas folhas e raízes. A sua produção se adequa aos fatores ambientais que a rodeiam e intensidade dos estresses a que está exposta (MAGALHÃES, et al., 2006), desenvolvendo tolerância pela ativação de proteínas da membrana plasmática (ZHAO, et al., 2004).

Tais conquistas evolutivas são a maneira mais eficiente de se minimizar os efeitos provocados pelo estresse salino (ZHU, et al., 2016), pois a partir de três ferramentas adaptativas conseguem se beneficiar do ambiente, antes impróprio para sua existência, os quais são: a prevenção de perda de água, desidratação; acúmulo do sal em vacúolos no citoplasma para que este não interfira no metabolismo intracelular; e potencial de reparação de lesões (ANOWER, et al., 2017).

Já quando se refere ao solo, está qualitativamente sujeito a dois níveis, tanto em horizontes como nas camadas superficiais (RIBEIRO, et al., 2009). A primeira está vinculada as características e o processo natural do solo em locais com déficit hídrico, alta evapotranspiração, baixa precipitação, intemperismo, arraste e depósito de sais pelo vento (ROACHE, BAILEY, BOON, 2006). Já a segunda é gerada pelo manejo incorreto da cultura durante o plantio, tanto pela irrigação com água de baixa qualidade, adição de compostos químicos na água para controle de pragas deliberadamente, remoção da cobertura vegetal perene, principalmente ateando fogo que deixa o solo exposto e muito danificado (DE OLIVEIRA, et al., 2008).

Apesar de se saber da baixa qualidade da água de irrigação em áreas áridas e semiáridas, suas propriedades estão associadas a sazonalidade climática, em períodos mais chuvosos a concentração de sais é reduzida devido a lixiviação deste, e em épocas mais secas, ocorre o contrário, um acúmulo maior dos níveis de sais (MOTA, 2014). Porém, caso seja feito a

irrigação com manejo adequado, é possível ter duas safras por ano com a cultura do algodão (SILVA, et al., 2009), não necessariamente, assegurando outros parâmetros que assegurem uma boa produtividade nas duas colheitas anuais.

O algodoeiro no nordeste, tem cerca de 23 mil hectares afetados por problemas de salinidade (SILVA, et al., 2008). O comportamento vegetal pode advir de respostas fisiológicas, como alteração de crescimento, capacidade fotossintética e desempenho dos estômatos (PEREIRA, 2012) aos estresses, dentre os quais, a salinidade em que possui uma boa tolerância a esta. Em estudo em 2012 (DE OLIVEIRA, et al., 2012), a cultura do gênero *Gossypium* é capaz de ver uma diminuição em todos fatores, principalmente região foliar (65,8%) e massa seca (65%), que implica num processo de desidratação, ao decorrer do aumento da exposição das plantas à salinidade.

Na fase de germinação, o excesso salino provoca uma diminuição quantitativa e temporal do processo devido a alta umidade do solo e baixa infiltrabilidade (DE VASCONCELOS, 1990). Já durante a fase de crescimento e desenvolvimento, manifestando na planta um tamanho diminuto mais intenso nas secundárias horizontais do que nas raízes primárias verticais (REINHARDT, ROST, 1995), redução da extensão foliar e consecutivamente a capacidade fotossintética devido a diminuição no pigmento da clorofila associada ao fechamento dos estômatos reduzindo a taxa de fixação de carbono (JÁCOME, 1999; JAMIL, et al., 2007; ZANANDRES, et al., 2006).

A fisiologia do algodão branco e colorido são afetadas pelo efeito salino, visível, quantitativamente, nas folhas, altura da planta diâmetro do caule (JÁCOME, et al., 2003; SOUSA JÚNIOR, FERNANDES, GHEYI, 2008). Com isto, apesar de tolerante, o estresse salino influencia demasiadamente de forma negativa a produtividade da cultura e os produtos gerados a partir da colheita (DA SILVA, et al., 2008).

### 3.3 BIOINFORMÁTICA

A bioinformática, é um campo emergente e cada vez mais fundamental para o estudo científico, como ferramenta para desenvolvimento de plataformas e bancos de dados de sequenciamento de aminoácidos e ácidos nucleicos (FERREIRA, 2017), assim como análise de sequências biológicas, genes e proteínas, para predição de função e interações entre as moléculas a partir de sua conformação tridimensional, estudo evolutivo a partir de árvores filogenéticas (SANTOS, 2010), mineração e integração de dados biológicos, design de drogas, primers e moléculas de interesse, distinção entre inibidores e enzimas, classificação de

biomoléculas homólogas pelo alinhamento múltiplo, além de sistematizar dados e relacionar tais informações biológicas (MARTINS, MENEZES, SOUZA, 2015).

Resumidamente, as áreas que englobam a abordagem *in silico* em moléculas biológicas são: alinhamento de sequências, dinâmica molecular, modelagem comparativa, técnicas *ab initio*, “threading” (“ferramenta de predição de enovelamento”), predição de função gênica, filogenia, biologia de sistemas e atracamento molecular. Em termos de aplicações, as análises *in silico* podem ser divididas em três grandes áreas de atuação, compreendidas entre macro e micromoléculas (respectivamente, DNA e proteínas, produto da transcrição e metabólitos), pesquisa funcional e, por último, estrutural (DO REGO, 2012), como mostra a Tabela 2. Já em relação a bioinformática em si, pode-se classifica-la em duas vertentes: a tradicional, relacionada com estudo e análises de sequências de aminoácidos e nucleotídeos, e a estrutural, que trabalha com técnicas conectadas com química computacional que estuda moléculas biológicas tridimensionalmente (VERLI, 2014).

**Tabela 2 - Áreas de aplicação do estudo in silico.**

Análise	Macro e micromoléculas	Funcional	Estrutural
Área de aplicabilidade	Estudos filogenéticos, comparações genômicas, alinhamento, edição, reconhecimento de regiões mutagênicas e epítomos, ilação de coalescência;	Predição de função, sítios ativos, expressão gênica, via metabólica e sua associação com doenças a partir da compreensão gênica;	Estudo das estruturas de enovelamento da proteína, principalmente a primária e terciária, a fim de classificar, comparar e interrelacioná-las;

**Fonte:** REGO (2012, p. 9).

Este termo, bioinformática, surgiu por Pauline Hogeweg, em 1979, em estudos voltados para a área da biologia sistemacional (MARTINS, et al., 2015; SANTOS, 2010). Porém, desde a década de 1960, pesquisadores da área biológica já usavam alguns recursos computacionais para auxílio em seus trabalhos. Esta área das ciências biológicas, ganhou mais visibilidade a partir de 1990 com o projeto do genoma humano e a disseminação da internet, que tornou a informação mais global e acessível, fazendo-a expandir tanto em volume de dados produzidos quanto ferramentas e técnicas para uso *in silico* (DO NASCIMENTO, 2016).

Dentro desta ciência, o conhecimento das proteínas, mais especificamente, de suas fases estruturais de enovelamento, primária e terciária, comparativamente à sequência e função, é um

dos pontos mais investigados, afinal sua conformação está intimamente ligada com sua funcionabilidade química e biológica dentro do metabolismo celular (DO REGO, 2012).

Para estudo da disposição terciária da macromolécula é necessário o uso de softwares e plataformas que façam a predição organizacional de seu arranjo, os quais podem ser por *ab initio* – utiliza informações físico-químicas dos resíduos de aminoácidos para construção do modelo quando normalmente não se tem moléculas homólogas em bancos de dados para uso; modelagem por homologia – usa moléculas análogas sequencialmente e determinadas por testes experimentais; e threading – relaciona padrões de dobras de proteínas, com ou sem ancestralidade comum, para observação de similaridade e construção do arranjo tridimensional (DO REGO, 2012).

O passo seguinte é a validação do modelo, o qual ocorre por meio de ferramentas como o gráfico de Ramachandran e Z-score. O primeiro emprega a avaliação dos ângulos de dobramentos terciários de DNA, RNA e proteína, estes ângulos torcionais correspondem, similarmente ao plano cartesiano de coordenadas, sendo o ângulo *phi* (referente ao carbono alfa e o átomo de nitrogênio do agrupamento amina) – eixo das ordenadas e o *psi* (ângulo existente entre o carbono alfa e o carbono entre o carbono alfa, e grupo carboxila), os quais podem rotacionar 180°.

Dentro deste plano cartesiano, há regiões favoráveis energeticamente, não havendo choques alostéricos dentre os átomos da estrutura tridimensional, valores próximos ficam em regiões secundárias, ainda permissivas, já os valores desfavoráveis são resultado de angulações torcionais inadequados à proteína que foi modelada gerando, assim, os choques alostéricos. Tais áreas dentro do gráfico de Ramachandran permissivas e não permissivas apresentam exceção, somente para resíduos de glicina que tem maior flexibilidade angular e a prolina que é mais rígida, no total, qualitativamente, o gráfico deve possuir 90% ou mais de seus resíduos de aminoácidos nas regiões favoráveis (MORRIS, et al., 1992).

Já o Z-score como ferramenta de refinamento qualitativo da validação da proteína, o qual usa a partir da sequência de aminoácidos fornecidos, os átomos do carbono alfa (C-alfa) que permite até estruturas com baixa resolução. Assim, esta ferramenta estatística fornece um panorama do número de desvios padrão em relação à média populacional em análise, ou seja, este teste verifica o quanto a sequência em estudo varia comparativamente a faixa de pontuação encontrada em proteínas nativas com tamanho similar a de entrada (DATA SCIENCE, 2020; PROSA-WEB, 2021).

Este conceito se aplica também ao gráfico de pontuações de resíduos ou gráfico de energia dos scores, o qual, estatisticamente, plota a energia da sequência de aminoácidos em

função de sua posição, calculando para isso, a média de cada fragmento de resíduos para análise energética (em escuro são janelas de fragmentos maior, 40 resíduos, e em tom mais claro a janela é menor, 10 resíduos). Preferencialmente, deseja-se que os valores sejam negativos, pois desta maneira, há menos áreas problemáticas na sequência (PROSA-WEB, 2021).

Hoje, apesar de existir ferramentas bastante eficientes, mais precisos, e com possibilidades e mais possibilidades de validação computacional, a validação experimental ainda é uma exigência para análise da confiabilidade dos resultados obtidos computacionalmente (VERLI, 2014). Apesar disto, a complementaridade das duas áreas, experimental e computacional, é o que torna o desenvolvimento científico possível e acelerado, pois é necessário de métodos de bancada, como o de espectrometria de massa, cristalografia, para que sequências gênicas possam ser adicionadas nos bancos de dados, e este, por sua vez, traga ao pesquisador uma nova perspectiva do que se está estudando, de forma mais integrada e íntima, criando um ciclo de hipóteses a serem compreendidas por uma diferente ótica, computacional e experimental (E SILVA, NOTARI, DALL'ALBA, 2020).

Na área do melhoramento genético vegetal, torna-se um instrumento primordial para viabilização de qualquer estudo, pois muitas análises elucidativas seriam demoradas e até inviáveis de se fazer à mão, como por exemplo, a integração de dados, como estrutura de peptídeos, vias de sinalização e metabólica, e sequências de DNA/RNA e proteínas, tornando mais compreensível a expressão gênica, localização de proteínas relacionadas com funções de interesse dentro do genoma, funcionamento, desenvolvimento e formação de órgãos vegetais, determinação do ambiente no fenótipo (VASSILEV; et al., 2005).

Particularmente, a filogenia encontra-se compreendida no melhoramento, correlatando e compreendendo a evolução de genes, proteínas e características específicas de grupos de organismos a partir da interpretação das relações filogenéticas, as quais são hereditárias, como os números dos cromossomos, haplóides, diplóides, tetraplóides, alotetraplóides, por exemplo. Assim, as discrepâncias entre os modelos são determinantes para se estabelecer a distância evolutiva entre um e outro vegetal (DE MEDEIROS, 2014).

Esta evolução foi de certa forma acelerada pela domesticação das plantas pelo homem, quando o ser humano passou do nomadismo ao sedentarismo no período neolítico, causou mudanças morfológicas e fisiológicas entre as espécies domesticadas e selvagens nas gerações que se passaram (HANCOCK, 2005), também a prática da seleção e cruzamentos realizados por agricultores das mais diversas civilizações, resultou no surgimento de novas variedades durante as eras humanas. Esta domesticação gerou algumas mudanças que caracteristicamente estão associadas a sete fatores, de acordo com EVANS (1993; apud MANTOVANI, et al.,

2008): supressão do mecanismo de disseminação de sementes, alteração fisionômica (alometria e condensação), processo de germinação mais acelerado e homogêneo, sincronismo nas fases de floração e maturação, modificações bioquímicas, aumento no tamanho dos órgãos, diminuição do ciclo de vida e sistemas de hibridação.

A partir deste conhecimento, o homem moderno, aprimorou este modo de domesticação e melhoramento com técnicas mais concludentes, não baseadas somente no fenótipo, marcadores morfológicos, mas sim no genótipo para que se consiga a expressão fenotípica desejada, surgindo como ferramenta os marcadores moleculares e enzimáticos (FERREIRA, 2017; MANTOVANI, et al., 2008).

Com o início do sequenciamento de genomas até hoje, houve a criação de várias técnicas para maior rapidez, menor custo, desde o método Sanger que fazia a replicação *in vitro* do DNA, fragmentação deste para o sequenciamento por partes, o qual se unifica novamente a partir de ferramentas computacionais; passou pelos algoritmos de grafos, como o de Brujiin e euleriano; e a técnica mais usada hoje em dia, a NGS (Next Generation Sequencing) que utiliza bancos de dados para se ter parte ou todo o DNA sequenciado de algum organismo (FERREIRA, 2017), além da possibilidade de estudo de mutações e funções de partes do gene para estudo de mecanismos regulatórios da fisiologia, crescimento e estresses bióticos e abióticos (JHA, et al., 2015).

Para acesso a estas informações, há diversos bancos de dados como o NCBI (National Center for Biotechnology Information), EMBL (European Molecular Biology Laboratory) e DDBJ (DNA Data Bank of Japan) são as maiores plataformas de dados, responsáveis, até 2017, totalizaram juntas quase 2 bilhões e 800 milhões de sequências depositadas (DDBJ, 2021). Estas utilizam normalmente arquivos com a extensão Fasta para o download de informações que o usuário necessitar, e normalmente, é nesta extensão que são pedidos os arquivos a serem analisados em softwares e outros servidores, como BioEdit, Mega X para alinhamento e produção de árvores filogenéticas; FastPCR e Primer-Blast para desenho de primers; Binfo – RepatMasker Server para mascaramento de regiões da sequência; e Oligo, Bioinfo, Primer3Plus para validação de primers, como exemplos programas e suas funcionalidades dentro da bioinformática clássica para o melhoramento genético vegetal.



## 4 METODOLOGIA

O presente trabalho foi feito inteiramente *in silico* utilizando softwares e servidores para a obtenção e manipulação de dados referentes as proteínas relacionadas ao estresse salino em algodoeiro. Para o estudo das biomoléculas, foi adaptado o estudo de JIMENEZ (2019), já o desenvolvimento do desenho e validação de primers, foi utilizado como base metodológica diversos autores, como QUEIROZ (et al., 2017), YE (2012), VILHENA (2015), KALENDAR, LEE, SCHULMAN (2014) e KIBBE (2007).

### 4.1 TRIAGEM E OBTENÇÃO DAS SEQUÊNCIAS

A triagem das sequências, assim como a obtenção de informações sobre o genoma das espécies de *Gossypium* que tem depositadas e a coleta das sequências de proteínas taxonômicas referente ao estresse salino, foram adquiridas a partir das ferramentas da plataforma NCBI (National Center for Biotechnology Information) (<https://www.ncbi.nlm.nih.gov/pmc/>) e pela ferramenta do mesmo servidor, BlastP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>), coletando 16 sequências juntamente com a pesquisada, adicionando espécies tanto de algodoeiros quanto de outros organismos na triagem, levando em consideração o grau de similaridade entre a proteína escolhida na coleta (JIMENEZ, 2019).

### 4.2 ANÁLISE DAS SEQUÊNCIAS

Para estudo das características das sequências de proteína, foi usado o servidor ProtParam (<https://web.expasy.org/protparam/>) para o estudo físico-químico; o NCBI (<https://www.ncbi.nlm.nih.gov/pmc/>) para obtenção e estudo dos domínios conservados das sequências (JIMENEZ, 2019).

### 4.3 ANÁLISE FILOGENÉTICA E AGRUPAMENTO DAS SEQUÊNCIAS

As sequências proteicas obtidas, foram primeiramente alinhadas no servidor Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) para visualização dos dados, e pelo algoritmo ClustalW do software MEGA X para produção da árvore filogenética das proteínas

separadamente, já para a comparação entre as proteínas foi usado o software BioEdit (FONSECA JÚNIOR, 2014; VIALA, 2014) no alinhamento das sequências totais, tanto para a visualização dos resultados quanto para produção da árvore. Em seguida, foi feita a árvore filogenética no MEGA X, pelo método de máxima parcimônia com bootstrap de mil réplicas, modelo de substituição padrão do programa, realizado tanto o de forma individual com as três proteínas, quanto as três juntas totalizando 48 sequências, à fim de comparação (JIMENEZ, 2019).

#### 4.4 PREVISÃO, AVALIAÇÃO E VALIDAÇÃO DA ESTRUTURA TERCIÁRIA DOS MODELOS

Para estudo da estrutura terciária foi usado o software Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>); análise da estrutura tridimensional de hidropaticidade pelo software Chimera; servidor Molprobitry (<http://molprobitry.biochem.duke.edu/>) para análise de Ramachandran para avaliação da qualidade conformacional dos modelos teóricos protéicos; análise de Z-score e da energia dos scores foi realizada a partir do servidor ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) (JIMENEZ, 2019).

#### 4.5 PRODUÇÃO E VALIDAÇÃO DE PRIMERS

Usado o servidor NCBI (<https://www.ncbi.nlm.nih.gov/pmc/>) para coleta das sequências de RNA dos genes que traduzem as proteínas de interesse. Após esta etapa foi feita o corte da sequência nas regiões específicas do gene relacionada com o domínio conservado, relacionado ao estresse salino (QUEIROZ, et al., 2017). Após esta etapa, foi desenhado os primers nos servidores Primer-Blast no NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>) e Primer3Plus (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>) (YE, 2012; VILHENA, 2015), além de validação do desenho pelo servidor Oligo Calc (<http://biotools.nubic.northwestern.edu/OligoCalc.html>) (KIBBE, 2007). Para o desenho do primer degenerado respectivos as três proteínas, foi feito aliamento das sequências cortadas no MEGA-X pelo algoritmo Clustal W (QUEIROZ, et al., 2017) e em seguida desenhada no software FastPCR (KALENDAR, LEE, SCHULMAN, 2014).

## 5 RESULTADOS E DISCUSSÃO

Para estudo das características e as proteínas que foram depositadas até o ano de 2021 no banco de dados NCBI, além das propriedades de cada uma encontrada na plataforma, Tabela 3 e 4.

**Tabela 3** - Informação genômicas das espécies do gênero *Gossypium* depositadas no banco de dados, NCBI.

Genoma	Espécie	Número de acessos**	Tamanho do genoma	Genes sequenciados
13 cromossomos haplóides	<i>Gossypium tomentosum</i>	GCA_007990495.1	2.193.557.323	trn; ndh; rps; GotoCr; atp; pet; psa; psb; ccs; rpl; rpo; rcb; ycf
13 cromossomos haplóides	<i>Gossypium laxum</i>	GCA_013511315.1	833.895.178	-
13 cromossomos haplóides	<i>Gossypium lobatum</i>	GCA_013467485.1	744.535.494	trn; D9X20; trnl; trnf; rps; ycf; atp; psb; ndh; rpl; psa; pet; cem; rpo; acc; rbc; ccs; mat; clp
13 cromossomos haplóides	<i>Gossypium schwendimanii</i>	GCA_013677275.1	729.428.757	trn; D9X19; ycf; rps; atp; psb; ndh; rpl; psa; pet; cem; rpo; acc; rbc; mat; clp
13 cromossomos haplóides	<i>Gossypium klotzschianum</i>	GCA_013677235.1	670.958.353	trn; rpl; ycf; ndh; rps; rrn; ccs; psa; rpo; pet; psb; clp; cem; acc; rbc; atp; mat; tRNA-Gly

13 cromossomos haplóides	<i>Gossypium davidsonii</i>	GCA_013677245.1	704.223.819	nad; atp; rrn; trn; ccm; mat; rpl; rps; espiga; sdh; cox; mtt; ycf; ndh; psa; ccs; rpo; pet; clp; psb; cem; rbc; atp; mat
13 cromossomos haplóides	<i>Gossypium aridum</i>	GCA_013487665.1	739.119.247	rpl; trn; ycf; ndh; rps; rrn; ccs; psa; clp; rpo; pet; psb; rpl; acc; rbc; atp; mat
13 cromossomos haplóides	<i>Gossypium australe</i>	GCA_005393395.2	1.743.386.794	rpl; trn; ycf; ndh; rrn; rps; ccs; psa; clp; pet; psb; cem; rbc; acc; rpo; atp; mat
13 cromossomos haplóides	<i>Gossypium trilobum</i>	GCA_013467465.1	655.376.848	nad; atp; rrn; trn; ccm; mat; rpl; rps; cox; espiga; mtt; sdh; ycf; ndh; ccs; psa; clp; rpo; pet; psb; cem; acc; rbc; mat
13 cromossomos haplóides	<i>Gossypium armourianum</i>	GCA_013677265.1	780.950.899	rpl; trn; ycf; ndh; rps; rrn; ccs; psa; clp; pet; psb; rbc; atp; rpo
13 cromossomos haplóides	<i>Gossypium harknessii</i>	GCA_013677255.1	732.154.658	rpl; trn; ycf; ndh; rps; rrn; ndh; ccs; psa; clp; pet; psb; cem; rbc; acc; atp; rpo; mat; ccm; ACQ15; sdh; cox; mtt; nad
13 cromossomos haplóides	<i>Gossypium turneri</i>	GCA_008044935.1	755.203.240	clp; trn; ndh; rps; psb; rrn; rpl; ccs; psa; ycf; rpo; pet; psb; cem; acc; rbc; atp; ihb; mat
13 cromossomos haplóides	<i>Gossypium thurberi</i>	GCA_004027125.1	582.006.748	nad; atp; rrn; trn; ccm; mat; rpl; rps; espiga; sdh; cox; mtt; psa; pet; ycf; ccs; ndh; psb; rpo; cem; acc; rbc
13 cromossomos haplóides	<i>Gossypium gossypoides</i>	GCA_013467495.1	664.724.314	trn; pet; ndh; ycf; psb; rps; psa; rpl; mat; GogoCr; rbc; atp; rpo; ccs; clp; cem
13 cromossomos haplóides	<i>Gossypium longicalyx</i>	GCA_010883175.1	1.190.210.753	rpl; rps; trn; ycf; rrn; ccs; ndh; psa; atp; psb; pet; mat; clp; cem; acc; rbc; rpo

26 cromossomos haplóides - 2 conjuntos de genoma	<i>Gossypium mustelinum</i>	GCA_007990455.1	2.315.094.184	rpo; trn; rpl; ndh; rps; ycf; psb; psa; mat; GomuCr; pet; cem; rbc; atp; ccs; acc
		GCA_017165895.1	2.297.221.019	
26 cromossomos haplóides	<i>Gossypium tomentosum</i>	GCA_007990485.1	2.193.557.323	trn; rpl; ndh; rps; ycf; psb; psa; mat; GotoCr; pet; cem; atp; rpo; ccs; clp; acc
26 cromossomos haplóides	<i>Gossypium darwinii</i>	GCA_007990325.1	2.182.957.963	trn; rpl; ndh; rps; ycf; psb; psa; mat; GodaCr; pet; cem; rbc; atp; rpo; ccs; clp; acc
13 cromossomos haplóides - 3 conjuntos de genoma	<i>Gossypium arboreum</i>	GCA_000612285.2	1.694.597.732	LOC10848_ _ _ _; trna;
		GCA_000787975.1	1.862.242.191	
		GCA_013265605.1	94.637.005	
26 cromossomos haplóides - 4 conjuntos de genoma	<i>Gossypium barbadense</i>	GCA_008761655.1	2.195.804.943	rps; ycf; cem; trn; rrn; rpo; rpl; psb; psa; pet; ndh; mat; clp; ccs; atp; acc
		GCA_002928715.1	1.394.237.086	
		GCA_001856525.1	2.566.742.836	
		GCA_002926015.1	775.251.585	
26 cromossomos haplóides - 5 conjuntos de genoma	<i>Gossypium hirsutum</i>	GCA_007990345.1	2.306.070.423	LOC10796_ _ _ _; trna;
		GCA_006980745.1	2.287.866.413	
		GCA_006980775.1	2.308.215.698	
		GCA_000987745.1	2.188.311.125	
		GCA_002504345.1	169.284.829	
13 cromossomos haplóides - 4 conjuntos de genoma	<i>Gossypium raimondii</i>	GCA_000327365.1	761.565.430	LOC10580_ _ _ _;
		GCA_013467475.1	615.032.910	
		GCA_000331045.1	773.767.704	

GCA\_005931075.1

734.884.094

---

**Fonte:** Próprio autor.

---

A partir da Tabela 3, pôde ser visto pelos genes sequenciados mais depositados na plataforma quais genes têm sido mais estudados de cada espécie de algodoeiro, além de se ter um panorama melhor dos dados já sequenciados depositados. Em seguida, foi iniciado a triagem referente a genes com potencial de prover tolerância à resistência ao estresse salino no vegetal, Tabela 4.

Com isto, foi selecionado três proteínas taxonômicas para estudo, Tabela 4, dentre o gênero *Gossypium*, espécie *G. hirsutum*, a qual possui maior importância econômica mundialmente. A primeira, pertencente ao domínio Rop\_like, tem na literatura a confirmação sobre seu potencial que de acordo com FINKELSTEIN (et al., 2002), este domínio está relacionado a resposta ABA (ácido ascórbico), o qual é um fitohormônio que modula processos do crescimento, desenvolvimento e estresses abióticos nas plantas, dentre estes, está o estresse salino.

Já a segunda, do domínio do citocromo P450, tem em sua descrição sobre os mecanismos de ação que os seus domínios conservados estão relacionados, mostra-se diretamente relacionada ao estresse em questão, assim como YAN (et al., 2016) indica em seu estudo com a expressão dessa proteína em plantas transgênicas de *Nicotiana benthamiana* submetidas a condição de redução hídrica e aumento da salina.

E a terceira, do domínio bZIP, respectivamente, está relacionada aos mecanismos de ação à estímulos fisiológicos, dentre os quais, de acordo com WANG (et al., 2021), que indicou associação dos genes deste domínio com estresse abiótico, mais especificamente os JcbZIPs 34, 36, 49 e 50 à resistência ao estresse salino e seca.

Após a validação das funções das três proteínas selecionadas, foi utilizada, a mesma, para estudo físico-químico no servidor ProtParam, Tabela 5, e no BlastP para coletar as sequências análogas para o alinhamento, Figura 1,2,3.

Os resultados do ProtParam são imprescindíveis para o conhecimento da sequência da proteína em estudos experimentais. Sobre o ponto isoelétrico, este é o valor de pH em que a molécula biológica possui bom desempenho em suas funções, imprescindível para estudo experimental, precisamente durante a corrida do gel onde as proteínas são separadas de acordo com seu gradiente de pH na corrida do gel, se movimentando a partir desse diferencial, e

parando no momento em que atingem a carga total igual a zero, ou seja, seu ponto isoelétrico (pI), atingindo a posição estacionária (BRUM, 2007), além do fato de suas propriedades serem influenciadas pelo pH do meio circundante, ou seja, a carga dos resíduos de aminoácidos muda conforme o meio, e por conseguinte, propriedades também, como a eletrostaticidade (VERLI, 2014), possui também, suma importância laboratorial no uso de primers, em que o pH se encontra nos cálculos básicos para definir a temperatura de anelamento ou temperatura de melting destes, primordial na viabilidade prática do desenho, em que tais sequências em estudo, Tabela 5, serão a base para a construção do primer (DE ARRUDA JÚNIOR, 2010).

Já o peso molecular relaciona-se ao estudo experimental da proteína, permitindo identificá-la e isolá-la por cromatografia, por exemplo, pois o peso é característico à quantidade de aminoácidos em sua sequência, sendo uma referência para determinação (GUERRA, et al., 2011). Na cromatografia, a separação de proteínas ocorre conforme o volume e carga molecular, os quais podem ser usados para determinação da massa e consecutivamente o peso molecular, com a eletroforese unidimensional, a qual possui menor capacidade de separação, menos de 50 proteínas geralmente em relação a bidimensional, a massa molecular é o fator determinante na corrida, pois quanto menor a massa, mais rápida é a movimentação da molécula sob o gel. Ao se falar da técnica de eletroforese bidimensional que consiste na separação de proteínas para estudo proteômico, os fatores que influenciam na corrida são o ponto isoelétrico e massa molecular (MORAES, et al., 2013).

O Gravy (*Grand average of hydropathicity*) é o índice de hidropaticidade que expressa a soma de hidropaticidade dos aminoácidos dividida pelo número de resíduos, este índice aponta o quão solúvel a proteína é com valores positivos para hidrofóbicas e negativos para hidrofílicas (DE LIMA, 2015), no caso das proteínas em estudo, Tabela 5, todas são hidrofílicas, o que é um bom fator ao se trabalhar na prática pois a forma de se trabalhar com a biomolécula irá se adequar à tal característica, além de indicar uma estrutura mais compacta e globular, funções biológicas mais dinâmicas, como transporte de gases em que a hemoglobina é responsável, já as hidrofóbicas tendem a ter uma conformação fibrilar em forma de um cordão e função mais estrutural, como o colágeno na formação dos tecidos (MARQUES, 2014).

O índice de estabilidade provém o nível de estabilidade que a biomolécula possui em um tubo de ensaio, a qual é considerada estável abaixo de quarenta e instável acima deste valor (DE LIMA, 2015). Outra característica importante é o índice de meia vida que mostra o tempo que a proteína se mantém ativa *in vitro* e *in vivo*, em diferentes organismos, o que direciona o pesquisador para a manipulação deste material biológico em análise.

**Tabela 4** - Informações sobre as proteínas taxonômicas com função sobre estresse salino em plantas do gênero *Gossypium*, informações retiradas do NCBI.

Proteína	Domínio	Mecanismo de ação	Função	Código de acesso
proteína de ligação a GTP semelhante a rac ARAC8 [ <i>Gossypium hirsutum</i> ]	Rop_like	A subfamília Rop (proteína relacionada com Rho de plantas) desempenha um papel em diversos processos celulares, incluindo organização do citoesqueleto, pólen e crescimento celular vegetativo, respostas hormonais, respostas ao estresse e resistência a patógenos. Rop4 e Rop6; Rop6 também é responsável pela resposta ABA, e a resposta ABA também é regulada por Rop10.	Estresse	NP_001314362.1
citocromo P450 tipo CYP82D47 [ <i>Gossypium hirsutum</i> ]	Citocromo P450 (CYP82 ou Cyp82)	Os genes da família 82 do citocromo P450 (CYP82 ou Cyp82) residem especificamente em dicotiledôneas e geralmente são induzidos por distintos estresses ambientais. Os membros caracterizados incluem: Glycine max CYP82A3 que é induzido por infecção, salinidade e estresse hídrico, e está envolvido na via de sinalização do ácido jasmônico e etileno, aumentando a resistência das plantas.	Estresse salino	NP_001314395.1
proteína de fator de choque térmico semelhante a HSF8 [ <i>Gossypium hirsutum</i> ]	bZIP	Eles atuam em redes de homo e heterodímeros na regulação de um conjunto diverso de processos celulares, incluindo sobrevivência celular, aprendizagem e memória, metabolismo lipídico e progressão do câncer, entre outros. Eles também desempenham papéis importantes nas respostas a estímulos ou sinais de estresse, como citocinas, agentes genotóxicos ou estresses fisiológicos.	Estresse	NP_001314347.1

Fonte: Próprio autor.



**Tabela 5** - Características físico-químicas das proteínas base usadas na ferramenta BlastP para coleta das outras 15 sequências do servidor NCBI, informações obtidas no servidor ProtParam.

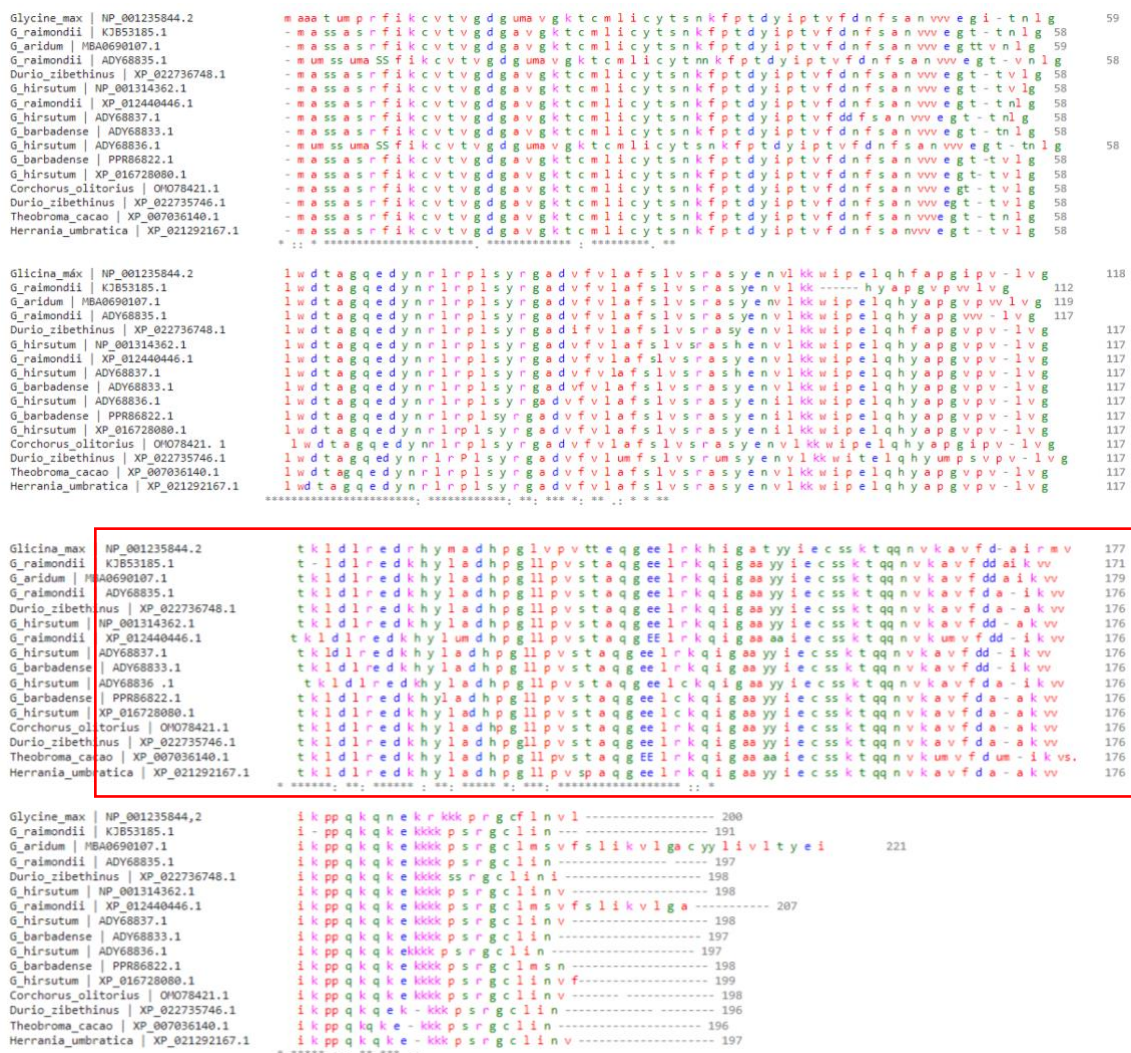
Proteína	Meia vida	Índice de instabilidade	Grande média de hidropatia (GRAVY)	Ponto isoelétrico	Número de aminoácidos	Peso molecular
Rop_like	30 horas	Grau II = 41,63	-0,217	9,02	198	21892,23
P450	(reticulócitos de mamíferos, <i>in vitro</i> ). > 20 horas (fermento, <i>in vivo</i> ).	Grau II = 36,93	-0,211	8,83	522	58489,73
Hsf8	> 10 horas (Escherichia coli, <i>in vivo</i> ).	Grau II = 63,54	-0,727	4,78	515	56661,67

**Fonte:** Próprio autor.

Em seguida, após identificar as três proteínas taxonômicas responsáveis pelo estresse salino no algodoeiro, foi feito o alinhamento de sequência múltipla entre as 16 sequências de cada uma das três proteínas no ClustalW (Figura 1,3,5). É importante obter tanto sequências com alto grau de similaridade, quanto de baixa similaridade para que a árvore tenha uma maior segurança quanto as suas estimativas evolutivas, assim, foi adicionado tais sequências de proteínas, não observando o organismo a qual pertence, mas sim o grau de similaridade com a proteína alvo.

Este passo é importante pois, procurar similaridades entre as sequências de nucleotídeos e/ou aminoácidos, neste caso é de aminoácidos, a fim de determinar se são idênticas o suficiente para inferir em sua homologia, pois nestas ocasiões pode-se estabelecer similaridade funcional (DE LIMA, 2015) e evolutiva.

**Figura 1** - Alinhamento de sequência múltipla da proteína de ligação a GTP semelhante, a rac ARAC8, da espécie de algodoeiro *Gossypium hirsutum* no ClustalW.



Fonte: Próprio autor.

Comparando o alinhamento obtido no ClustalW, Figura 1, com o gráfico de regiões conservadas do domínio Rop\_like da proteína de ligação a GTP semelhante a rac ARAC8 (*Gossypium hirsutum*) referente a estresses abióticos do NCBI, Figura 2, temos que estas sequências estão localizadas na posição 80, 87, 100, 167, 173, 180 da proteína base do alinhamento (NP\_001314362.1), as quais estão selecionadas na Figura 1 em vermelho. A partir da observação dos resíduos conservados, pode-se inferir que dentre as seis posições do NCBI, no alinhamento não é possível encontrar as mesmas posições, somente em posições próximas, 171, 176, 177, 179 mais próximos do 173 e 180 (Figura 2), regiões mais promissoras a fazer parte deste domínio com função relacionada ao estresse salino.

Figura 2 - Dados do alinhamento do domínio Rop\_like no NCBI.

```

Pssm-ID: 206705 Cd Length: 173 Bit Score: 369,18 E-value: 1,85e-132

      10 20 30 40 50 60 70 80
gi 1031987133 8 FIKCVTVGDGAVGKTCMLI C YTSN K FPTYD eu PTVFDNFSANVVV E L T TVNLGLMDTAGQEDYNRLRPLSYRGADV V LAF 87
CDD: cd04133 1 FIKCVTVGDGAVGKTCMLI S YTSN T FPTYD V PTVFDNFSANVVV D L N TVNLGLMDTAGQEDYNRLRPLSYRGADV L LAF 80

      90 100 110 120 130 140 150 160
gi 1031987133 88 SL V S R AS H ENVLKWIPEL Q HYAPGVP V VLVGTKLDLR E DK H. ADHPG LL P VS TAQGEELRKQIGAA Y YIECCSKTQQN 167
CDD: cd04133 81 SL I S K AS Y ENVLKWIPEL R HYAPGVP eu VLVGTKLDLR DK QFF ADHPG AV P IT TAQGEELRKQIGAA A YIECCSKTQQN 160

      170
gi 1031987133 168 VKAVFD A AIKVV I 180
Cdd: cd04133 161 VKAVFD A AIKVV L 173

```

Fonte: Próprio autor.

Figura 3 - Alinhamento de sequência múltipla da proteína citocromo P450 tipo CYP82D47, da espécie de algodoeiro Gossypium hirsutum no ClustalW.

```

Prunus_arnica | CAB4296392.1 --- m d y s l p y v n s i t a f l f t i i v s y y f s --- r r w r p a k l a p t e a k g a w p i f g h l p l l g 53
G_davidsonii | MBA0624583.1 m d l d l s i f g y --- a l v l g i t l - l f l y t k l k r s s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_arboreum | KHG01290.1 m d l d l s i f g y --- a l v l g i t l - l f l y t k l k k s s s g s v g k a p p v a a g a w p i i g h l p l l g 55
G_schwendimani | MBA0866527.1 m d l d f s t f g y --- a v v l g i t l - l f l y t k l k k s s s g s g g k a a p v a a g a w p i i g h l p l l g 55
G_schwendimani | MBA0866525.1 m d l d f s t f g y --- a v v l g i t l - l f l y t k l k k s s s g s g g k a a p v a a g u m w p i i g h l p l l g 55
G_trilobum | MBA0775238.1 m d o n d e d f s t f g y --- u m a v v l g i t l - l f l y t k l k k s n s g s g g k a a p v a a g u m w p i i g h l p l l g 55
G_australe | KAA3488855.1 m d p l d l s t f g y --- u m a v v l c i t l - l f l y t k l k k r s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_australe | KAA3488856.1 m d p l d l s t f g y --- a v l l c i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_arboreum | XP_017643194.1 m d p l d l s t f g y --- a v l l g i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_mustelinum | TYJ35072.1 m d p l d l s t f g y --- a v v l g i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_barbadense | KAB2082570.1 m d p l d l s t f g y --- a v v f g i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_hirsutum | NP_001314395.1 m d p l y l s t f g y --- a v v l c i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_tomentosum | TYH71926.1 m d p l d l s t f g y --- a v v l c i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_mustelinum | TYI82308.1 m d p l d l s t f g y --- a v l c i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_davidsonii | MBA0623572.1 m d p l d l s t f g y --- a v l c i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
G_davidsonii | MBA0623571.1 m d p l d l s t f g y --- a v l c i t l - l f l y t k l k k p s s g s g g k a p p v a a g a w p i i g h l p l l g 55
: * . : : : : * * * * * : * * * * *

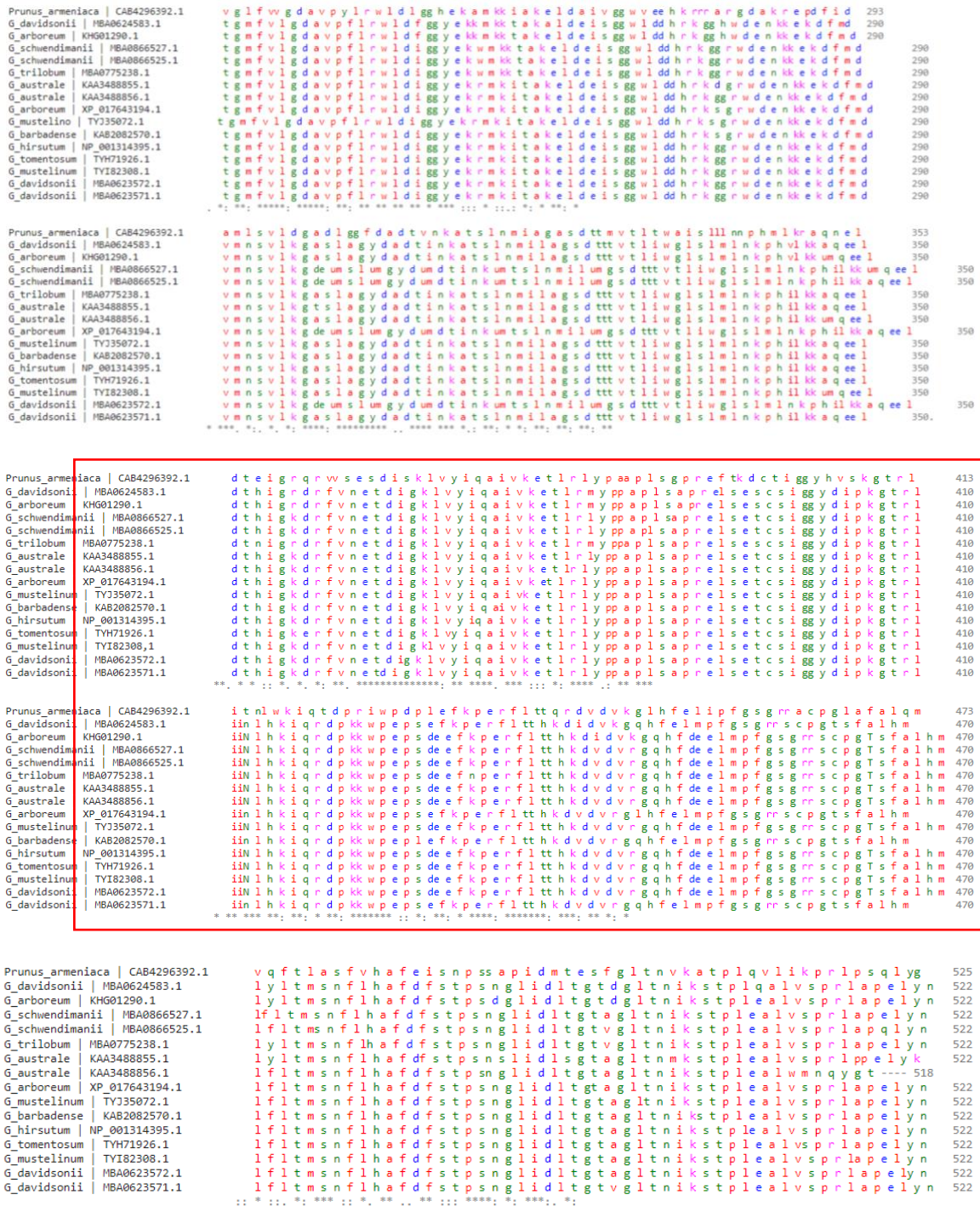
Prunus_arnica | CAB4296392.1 g s t p p h i k l a a n a d k y g p l f t i r i g v y p s l v i s s e i a k e c f t n d l a l s s r p k l a a v d h 113
G_davidsonii | MBA0624583.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s d v a k e i f t v n d m y v s s r s e f a a e h 115
G_arboreum | KHG01290.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s d v a k e i f t v n d m y v s s r s e f a a e h 115
G_schwendimani | MBA0866527.1 g p k t p h e t l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_schwendimani | MBA0866525.1 g p k t p h e t l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_trilobum | MBA0775238.1 g p k t p h e t l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_australe | KAA3488855.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_australe | KAA3488856.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_arboreum | XP_017643194.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_mustelinum | TYJ35072.1 g p k u m p h u e u m l g d l e e k y g p u m y m i r i g v h p u m l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_barbadense | KAB2082570.1 g p k u m p h u e u m l g d l e e k y g p u m y m i r i g v h p u m l v n s s e v u m k e i f t v n d m h v s s r s e f a a e h 115
G_hirsutum | NP_001314395.1 g p k u m p h u e u m l g d l e e k y g p u m y m i r i g v h p u m l v n s s e v u m k e i f t v n d m h v s s r s e f a a e h 115
G_tomentosum | TYH71926.1 g p k u m p h u e u m l g d l e e k y g p u m y m i r i g v h p u m l v n s s e v u m k e i f t v n d m h v s s r s e f a a e h 115
G_mustelinum | TYI82308.1 g p k u m p h u e u m l g d l e e k y g p u m y m i r i g v h p u m l v n s s e v u m k e i f t v n d m y v s s r s e f a a e h 115
G_davidsonii | MBA0623572.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
G_davidsonii | MBA0623571.1 g p k a p h e a l g d l e e k y g p a y m i r i g v h p a l v n s s e v a k e i f t v n d m y v s s r s e f a a e h 115
: * . : : * * * * * : * * * * * : * * * * * : * * * * * : * * * * *

Prunus_arnica | CAB4296392.1 i g y n y a m f g f s p y g q y w r e m r k i t t l e v l s n h r i d q l k k v f v s e i e g s m k l l y n t w a d k k 173
G_davidsonii | MBA0624583.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y n t w a d k k 175
G_arboreum | KHG01290.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f l s e i e g s m k l l y n t w a d k k 175
G_schwendimani | MBA0866527.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
G_schwendimani | MBA0866525.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
G_trilobum | MBA0775238.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
G_australe | KAA3488855.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a e k k 175
G_australe | KAA3488856.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a a k k 175
G_arboreum | XP_017643194.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a e k k 175
G_mustelinum | TYJ35072.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a e k k 175
G_barbadense | KAB2082570.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a e k k 175
G_hirsutum | NP_001314395.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a e k k 175
G_tomentosum | TYH71926.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
G_mustelinum | TYI82308.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
G_davidsonii | MBA0623572.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
G_davidsonii | MBA0623571.1 l g y n y a m f g f s p y g q y w r e m r k i t m l e v l s n h r i d q l k k v f v s e i e g s m k l l y k t w a q k k 175
: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *

Prunus_arnica | CAB4296392.1 k r e s n n r d g v l v e l k q w f d m t l n v i m r t v a g k r y s v a a d e d e k t e a n r v q s a l k e f f y y 233
G_davidsonii | MBA0624583.1 --- n g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_arboreum | KHG01290.1 --- n g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n t a l r d f f h l 230
G_schwendimani | MBA0866527.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_schwendimani | MBA0866525.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_trilobum | MBA0775238.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_australe | KAA3488855.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_australe | KAA3488856.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_arboreum | XP_017643194.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_mustelinum | TYJ35072.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_barbadense | KAB2082570.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_hirsutum | NP_001314395.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_tomentosum | TYH71926.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_mustelinum | TYI82308.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_davidsonii | MBA0623572.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
G_davidsonii | MBA0623571.1 --- d g s g k v l v e m k k h f s d l t l n v i m r t v a g k r y s v a e e d q k - e v l r y n k a l r d f f h l 230
: * * * * * : * * * * * : * * * * * : * * * * * : * * * * *

```



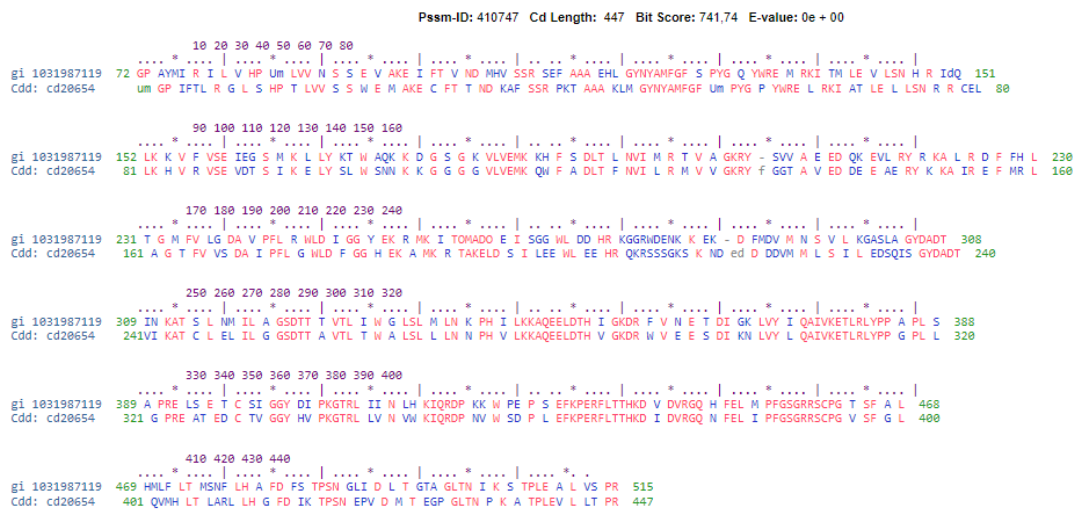


Fonte: Próprio autor.

Já comparando o alinhamento obtido no ClustalW, Figura 3, com o gráfico de regiões conservadas do domínio CYP82 da proteína citocromo P450 tipo CYP82D47 (*Gossypium hirsutum*) referente ao estresse salino do NCBI, Figura 4, temos que estas seqüências estão localizadas na posição 80, 151, 160, 230, 240, 308, 320, 388, 400, 447, 468, 515 da proteína base do alinhamento (NP\_001314395.1), as quais estão selecionadas na Figura 3 em vermelho.

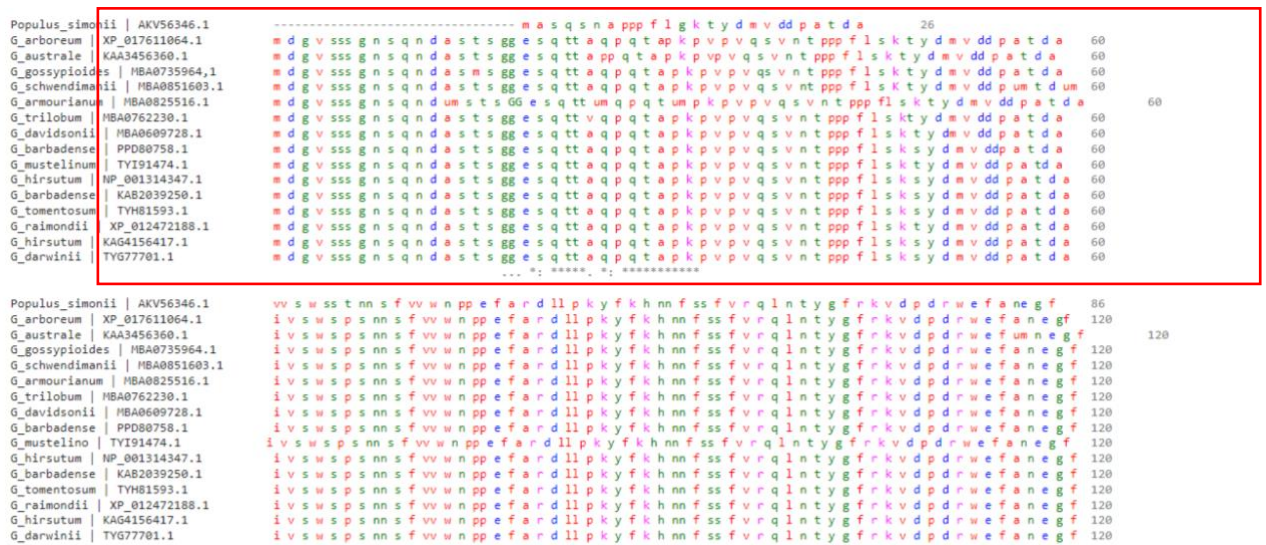
A partir da observação dos resíduos conservados, pode-se inferir que dentre as doze posições do NCBI, no alinhamento só aparece a posição 230, exatamente, menos na *Pronus armeniaca* (CAB4296392.1) que varia para 233, em relação as demais posições é possível encontrar aproximações, 410, 470, 473 para a 400, 468, 473 (Figura 4), regiões mais promissoras a fazer parte deste domínio com função relacionada ao estresse salino.

**Figura 4 -** Dados do alinhamento do domínio da família 82 do citocromo P450 no NCBI.



Fonte: Próprio autor.

**Figura 5 -** Alinhamento de sequência múltipla da proteína de fator de choque térmico semelhante a HSF8, da espécie de algodoeiro *Gossypium hirsutum* no ClustalW.









três alinhamentos individuais num múltiplo (Figura 10,19), para entender a partir das duas ferramentas o grau de similaridade e conservação das proteínas entre as espécies do gênero *Gossypium*, maioritariamente, pois há sequências similares dentro e entre espécies.

Na árvore filogenética, pode-se ver as relações evolutivas entre várias espécies de acordo com as sequências de genes, DNA, RNA, proteínas ou até características fenotípicas. Na árvore, quanto mais próximo um nó do outro (intersecções), mais recente é o antepassado, já sobre os ramos, significa que menor é a mudança e tempo evolutivo. Pela sua estrutura, há a possibilidade de se saber se há um ancestral comum a todos como referencial ou não, caso haja, esta é enraizada e caso contrário ela é não-enraizada (PEREIRA, 2010; SAKAMOTO, 2016).

Esta determinação tem influência direta nas sequências escolhidas para construção da árvore, assim para se determinar a raiz de uma árvore não-enraizada, é só adicionar ao alinhamento um outro grupo de táxon para maior consistência da análise, como foi feito ao incluir sequências com baixa similaridade juntamente com as de alto grau e de organismos diferentes (PEREIRA, 2010; SAKAMOTO, 2016). Este método funciona quando há um galho separado desde o início do ramo principal, o que pode ser visto na Figura 7,8,9,10, já que pela estrutura das árvores obtidas, estas não são enraizadas.

Pelo método de inferência utilizado metodologicamente, o de máxima parcimônia, é baseado numa matriz com os caracteres de interesse nas linhas e táxons nas colunas, e o programa escolhido irá percorrer os locais da árvore procurando critérios que atingem resultados admissíveis considerando cada aminoácido como um caráter particular no cálculo. É um método consistente, porém passível de artefactos, devido a evolução lateral, seleção convergente e outras causas. Aliado a este, foi adicionado o método para medir a confiabilidade de uma árvore inferida, o teste de bootstrap, em que reconstrói a árvores com base na matriz  $n \times m$  ( $m$  sequências, com  $n$  aminoácidos cada uma) e as colunas da matriz são escolhidas aleatoriamente e duplicadas em uma segunda matriz, gerando um novo alinhamento do mesmo tamanho, o qual irá inferir novas sequências usando métodos e parâmetros iguais (PEREIRA, 2010; SAKAMOTO, 2016; SCVORTZOFF, 2012).

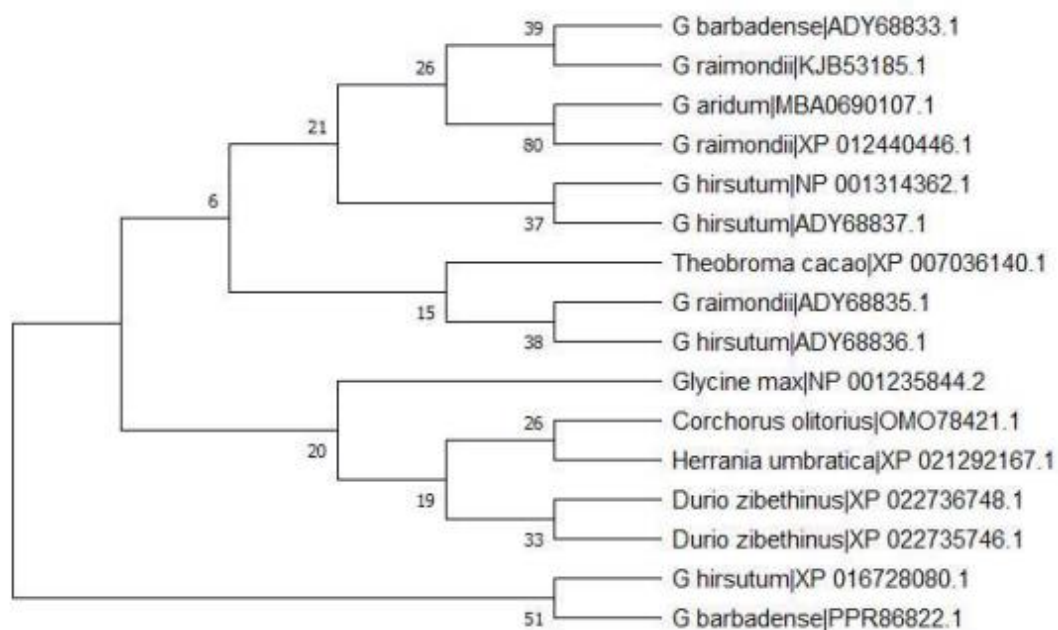
Assim, o resultado do teste será comparado com a árvore original, gerando para cada ramo interior um valor percentual referente a sua qualidade, já que este método observa a distribuição obtida em uma amostra para inferir a distribuição real desta. Deste modo, espera-se que para que seja confiável biologicamente o resultado seja igual ou maior a 95% (SANDERSON, 1995), porém é válido dizer que como este método infere aos próprios dados, não pode são comparáveis entre filogenias distintas (PEREIRA, 2010).



Observando as árvores filogenéticas das três proteínas dos Rop\_like, P450 e bZIP, pode-se inferir taxonomicamente sobre as relações evolutivas entre as mais diversas espécies e o grau de parentesco entre estas, os quais a distância é proporcional a quantidade de substituições de nucleotídeos da sequência gênica (CALDART, et al., 2016).

Para as proteínas do domínio Rop\_like, Figura 7, é possível observar a presença de um outgroup (51) com presença de duas proteínas de espécies distintas do mesmo gênero, *G. barbadense* e *hirsutum*, estas respectivamente, encontram-se presentes em outros clusters dentro da árvore, para a *G. barbadense* encontram-se distantes no 39 e 51, *G. hirsutum* localiza-se no 38,38,51. Também tem o *G. raimondii* nos pontos 39, 80, 38, assim como, contando com a presença do conjunto de aminoácidos mais similares com o gênero *Gossypium* do que entre elas mesmo. A disposição apresentada, revela que este gênero tem boa variabilidade, em parte pela distribuição geográfica as quais estão são originárias e ao fato de sua reprodução, parcialmente alógama e autógama, podendo assim, ser mais similares a proteínas de outras espécies do que entre o mesmo gênero (DE SOUSA, 2010).

**Figura 7** - Árvore filogenética bootstrap consensus da proteína do domínio Rop\_like, obtida no MEGA-X.

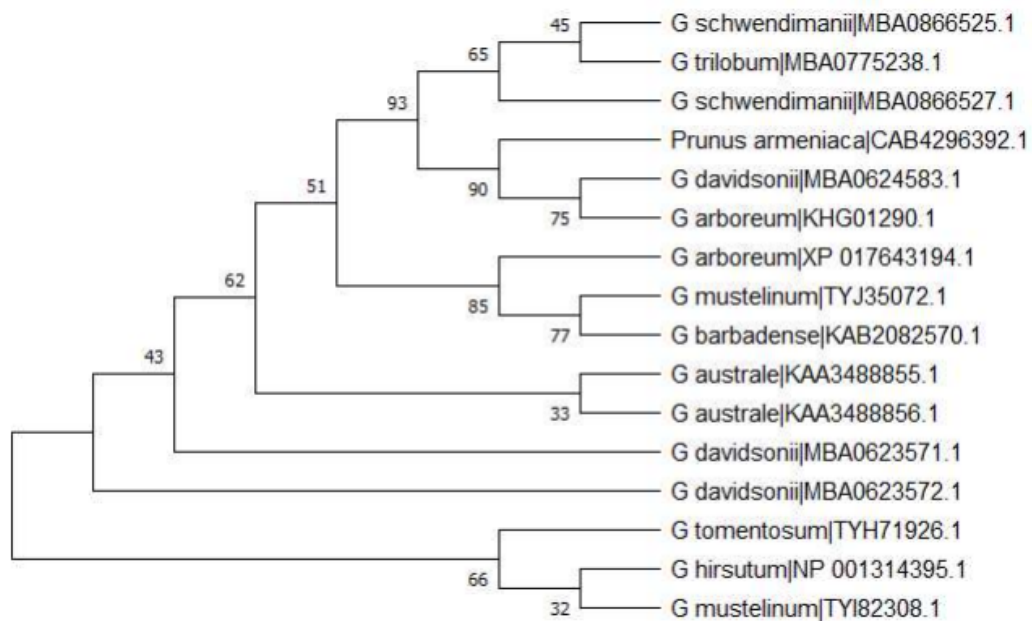


Fonte: Próprio autor.

A proteína do domínio P450, Figura 8, pode-se ver uma homogeneidade maior, já que dentre as 16 sequências, 15 são do gênero *Gossypium*, tal como os clusters (grupos): *G.*

*schwendimanii* e *G. trilobum* (45), *G. davidsonii* e *G. arboreum* (75), *G. mustelinum* e *G. barbadense* (77), *G. Hirsutum* e *G. mustelinum* (32) com ancestral comum o *G. tomentosum* (66). Esta similaridade entre espécies pode estar associada a quantidade de cromossomos, tetraplóides e diplóides, já que a reprodução cruzada só ocorre entre espécie com quantidade de cromossomos comuns (HOFFMANN, et al., 2020), à exemplo disto tem as espécies *G. mustelinum* e *G. barbadense*, tetraplóides que se encontram no mesmo grupo (77) (DE SOUSA, 2010). Além disso, há a presença de outgroup e diferença maior entre as proteínas da mesma espécie do que dentre espécies, como *G. mustelinum* nos nós 32 e 72, *G. arboreum* no 75 e 85, *G. davidsonii* nos 43, 62 e 75.

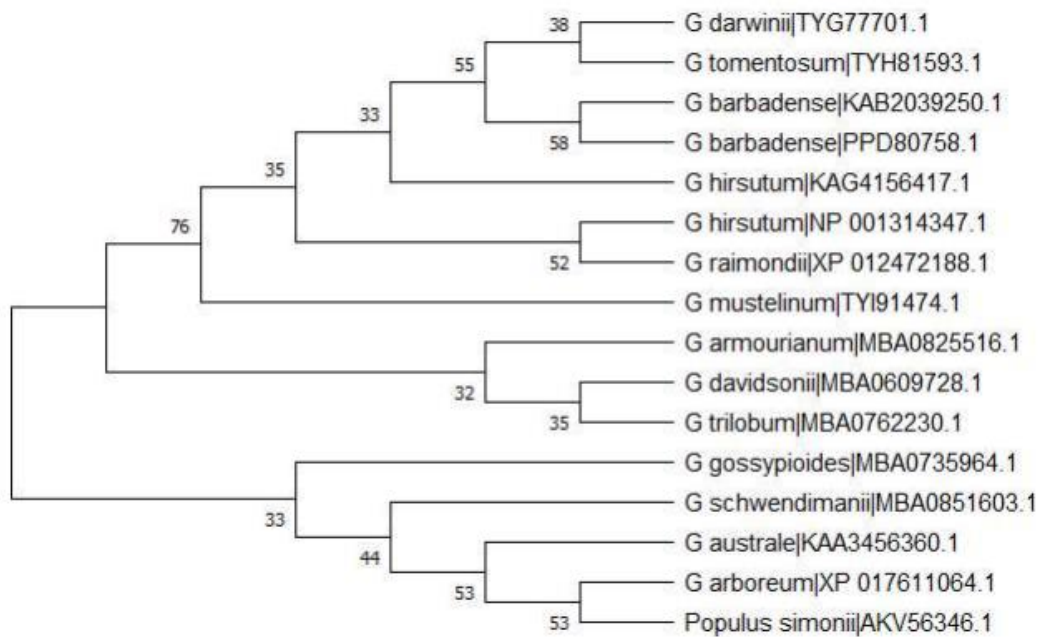
**Figura 8** - Árvore filogenética bootstrap consensus da proteína do domínio citocromo P450, obtida no MEGA-X.



Fonte: Próprio autor.

Já, a árvore das proteínas do domínio bZIP, Figura 9, apresenta disposição similar com a do domínio P450, tendo entre as 16 sequências, 15 pertencente ao gênero *Gossypium*. O outgroup existente é bem espesso em quantidades de sequências que possui, contendo 5 espécies diferentes. Outrora, a similaridade entre estas, as quais formam um grupo, são *G. davidsonii* e *G. tomentosum* (38), *G. hirsutum* e *G. raimondii* (52), *G. davidsonii* e *G. trilobum* (35). Reforçando, o fato do cruzamento como causa.

**Figura 9** - Árvore filogenética bootstrap consensus da proteína do domínio bZIP, obtida no MEGA-X.



**Fonte:** Próprio autor.

No estudo comparativo mais profundo sobre tais proteínas, Figura 10, é possível observar o grau de similaridade genética entre as proteínas em estudo relacionadas a outras espécies, ocasionalmente mais similares entre do que dentre as espécies, reforçando as causas da disposição da árvore, as quais são: boa variabilidade genética pela sua forma de reprodução e reprodução cruzada.

Adicionalmente a isto, na filogenia molecular, a inferência a respeito da classificação taxonômica das espécies é uma concepção válida, afinal a árvore filogenética de gene atribui sua construção em conjunto de homólogos, deduzidos a partir de dados moleculares, os quais podem ter a representação de uma espécie várias vezes dentro da árvore ou a sua exclusão caso não seja símil a biomolécula alvo. Já a árvore de espécie que apesar de ser inferida por sequência de genes em DNA, RNA, possui uma topologia que percebe realmente a história evolutiva das espécies observadas e só pode aparecer cada espécie uma vez, diferentemente da árvore de gene que não pode representar a de espécie, pois é até rasteiro considerar o estudo de genes em espécies como suficiente para estudo da evolução visto que cada gene pode retratar diferentes passos evolutivos para as espécies em estudo, como mutações e fluxo gênico, porém ao se fazer isso com vários genes pode-se conjecturar uma árvore de espécie (SAKAMOTO, 2016).

Dado a este fato de uso de vários genes para conjecturar uma árvore de espécie, pode-se observar na Figura 10, uma árvore formada com três genes, a qual pode nos dar informações mais relevantes sobre as espécies. De acordo com a mesma e o teste de bootstrap, pode-se ver nos galhos internos alguns com altas porcentagens, acima de 95%, valor de referência, em que estão no valor de 99%, dentro deles, nos galhos mais internos ainda, unidades taxonômicas, esta medida vai diminuindo, chegando a galhos terminais com 78% (*G. davidsonii* e *G. arboreum*), 82% (*G. arboreum* com o galho em 76% com *G. mustelinum* e *G. barbadense*) e 80% (*G. aridum* e *G. raimondii*), desta maneira pode-se inferir que tais espécies de algodoeiro são evolutivamente mais próximas biologicamente na realidade. Em relação aos demais ramos e nós terminais, devido a baixa porcentagem de bootstrap, não é válido inferir relação taxonômica-evolutiva entre os genes das espécies presentes, pois apresentam baixa consistência.

É viável atribuir, esta não inferência às árvores filogenéticas de cada proteína isolada, Figura 7,8,9, devido a baixa consistência dos valores atribuídos pelo teste de bootstrap, abaixo de 95%, em que para o domínio Rop\_like, Figura 7, seu galho mais consistente tem 80% (*G. hirsutum* e *G. hirsutum*) é um galho terminal referente a duas proteínas da mesma espécie, para o domínio P450, Figura 8, a média foi a mais alta dentre as três proteínas em análise, chegando a 93%, e por fim, a do domínio bZIP, Figura 9, foi a que possui menor valor, 76%.

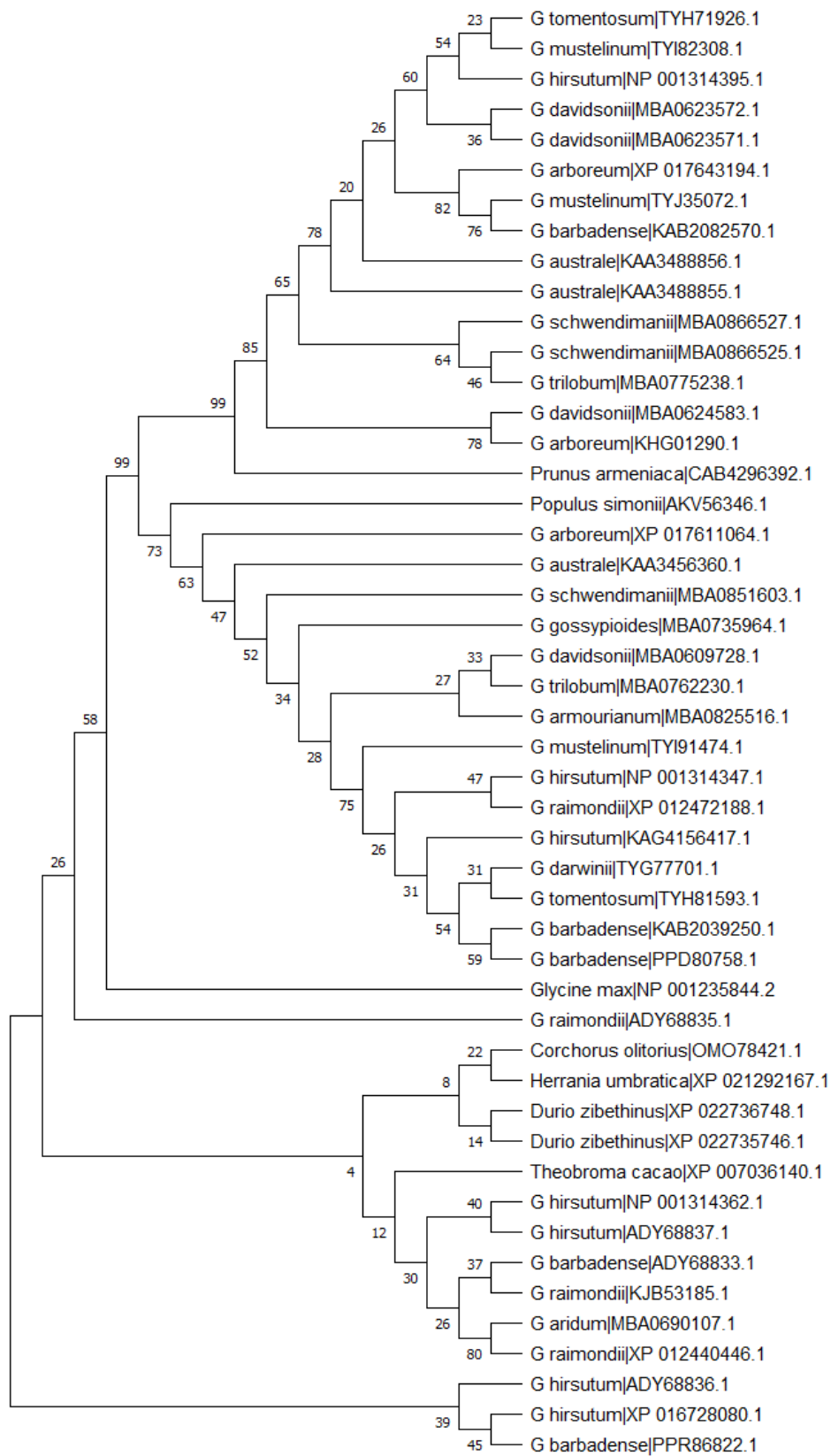
Este valor pode ser atribuído ao baixo grupo de táxons com baixa similaridade com a proteína alvo adicionados às sequências alinhadas para construção da árvore, gerando baixa consistência, o que é minimizado na árvore filogenética comparativa das três proteínas, Figura 10, por ter este maior número de táxons dissemelhantes. Outro fator a se considerar, principalmente na Figura 10, a possibilidade de inferências incorretas em valores altos de bootstrap, como em situações com taxas de mudanças muito divergentes e viés sistemático nos dados (PEREIRA, 2010).

Para o estudo da estrutura tridimensional das proteínas, Figura 11 (a,b,c), é interessante falar que a sua forma terciária revela muitas informações importantes, como a possibilidade de identificar quais resíduos estão mais internos, no núcleo proteico, ou mais na superfície, expostos a interações, além de sua composição molecular com a presença de folhas-beta, alfa-hélice e regiões de looping para caracterizá-la, uso em comparação estrutural onde pode-se identificar sítios funcionais, evoluções e mutações relacionadas a este, e função proteica a partir da sua conformação (DE LIMA, 2015).

Estas biomoléculas são extremamente sensíveis, ao nível em que uma simples mudança em sua sequência pode modificar sua estrutura e conseqüentemente a função desempenhada

pela mesma. Assim, a compreensão entre função e estrutura tridimensional para anotação funcional em maior escala deve ser observadas pelos padrões mais conservados e não padrões sequenciais, pelo fato de haver mais mutações e mudanças neste último, mais dificuldade de predição, além dos conservados serem, normalmente, a região que desempenha alguma funcionalidade (BOARETO, et al., 2012), tal como o fato da estrutura manter-se conservada mais do que a sequência em si, pois é a partir de uma grande porção de aminoácidos que a estrutura secundária e terciária são compostas.

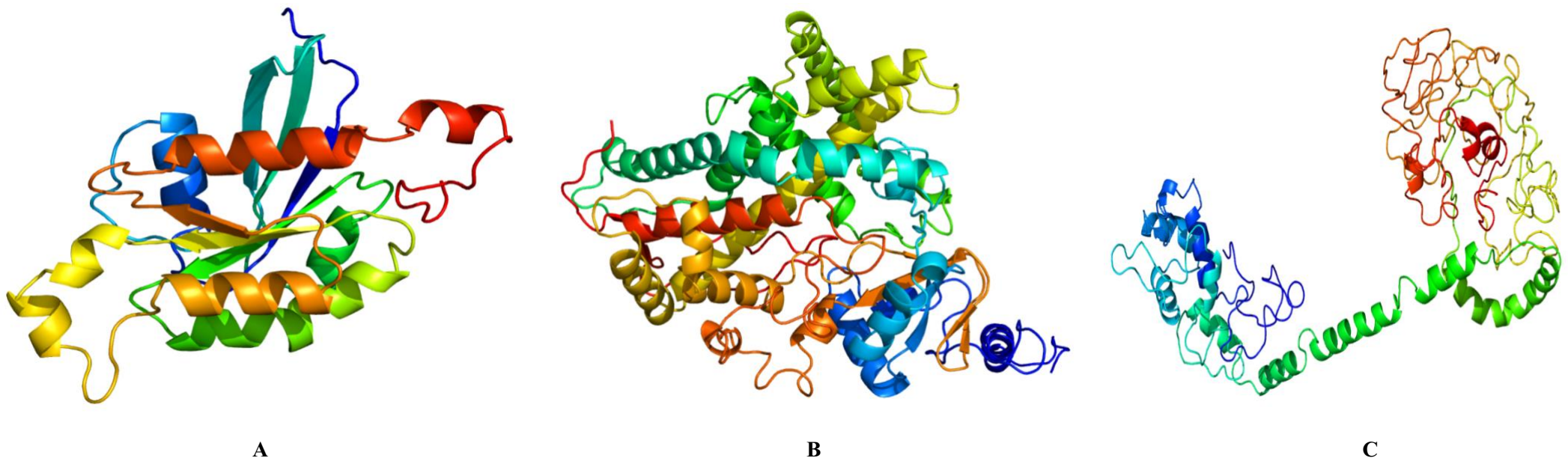
**Figura 10** - Árvore filogenética bootstrap consensus comparativa dos alinhamentos múltiplos das três proteínas, Rop\_like, P450, HSF8, obtida no MEGA-X.



Fonte: Próprio autor.

Dentre os métodos capazes de realizar tal predição, estrutura-função, há o que utiliza a similaridade estrutural, comparando a biomolécula alvo com outra estrutura análoga de função conhecida (LASKOWSKI, WATSON, THORNTON, 2005), alguns motivos provenientes de evolução convergente o divergente, pode ser usado na previsão de função. Uma dificuldade admitida à este método é que muitas vezes há proteínas com sequências análogas e funções divergentes, ou sequências distintas e mesma função e até grupo catalítico e mecanismo iguais, podendo ser, no primeiro caso em famílias distintas, e no segundo, o contrário. Outro fator, é a anotação com base apenas na estrutura local, gerando falsos-positivos, principalmente quando a sequência apresenta baixa resolução e/ou identidade (ROY, YANG, ZHANG, 2012), e a baixa porcentagem das estruturas disponíveis que foram resolvidas experimentalmente, é um fator restritivo à anotação funcional mais generalizada fundamentada no conhecimento da estrutura proteica (DE LIMA, 2015).

**Figura 11** - Análise tridimensional da proteína dos domínios Rop\_like. P450, bZIP, respectivamente, A, B e C.



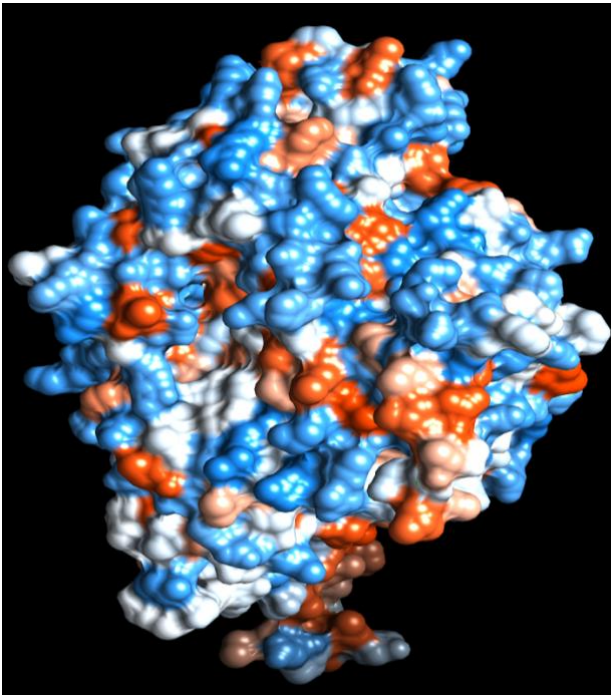
Fonte: Próprio autor.

Nas estruturas das proteínas na Figura 11, pode-se inferir que apesar de serem do mesmo organismo, *G. hirsutum*, elas tem estruturas, tamanho sequência dos números de aminoácidos, Tabela 5, e pertencem a domínios diferentes, porém apresentam mesma função relacionada a tolerância, resistência ao estresse salino, podendo divergir, provavelmente, em seu mecanismo de ação ou até grupo catalítico. Ademais, quanto as características conformacionais, na Figura 11, (a) referente a proteína do domínio Rop\_like tem grande quantidade de folhas-beta em relação a alfa-hélice, e ligações de looping ligando estas entre si; a proteína do domínio P450 tem grande quantidade de folhas-beta em relação a alfa-hélice, e ligações de looping ligando estas entre si (b); e a do domínio bZIP tem grande quantidade de regiões de looping com poucas folhas-beta interligadas por estas e sem presença de alfa-hélices (c).

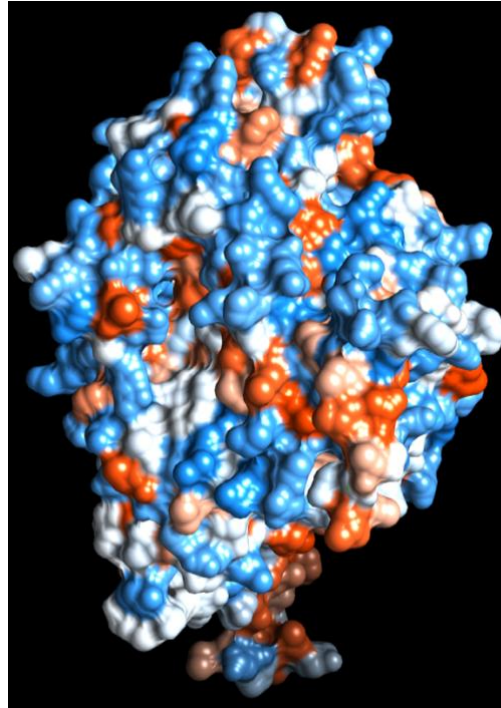
Ainda, sobre a configuração tridimensional das biomoléculas, o perfil de hidrofobicidade, importante para se observar os sítios funcionais, além do comportamento da proteína em meio aquoso vinculando a sua tendência de solubilidade, mais hidrofílica visualmente como região em azul, ou de formar agregados, mais hidrofóbica apresentada como região em vermelho (DO REGO, 2012). Esta característica tem grande relação ao índice de GRAVY, Tabela 5, pois indica o perfil de hidropaticidade, já testado em estrutura secundária, assim como o programa Chimera em análise tridimensional com a estrutura terciária das proteínas. Desta forma, ao analisar a Figura 12 (a,b,c), do domínio Rop\_like, P450 e bZIP, todas retratam a área superficial mais em azul, mostrando que esta tem boa solubilidade em meio aquoso, confirmando previsão obtida quantitativamente pelo teste de GRAVY, e adicionando à esta última análise, o fato de se poder observar quais áreas as proteínas apresentam o perfil mais hidrofílico em sua estrutura.



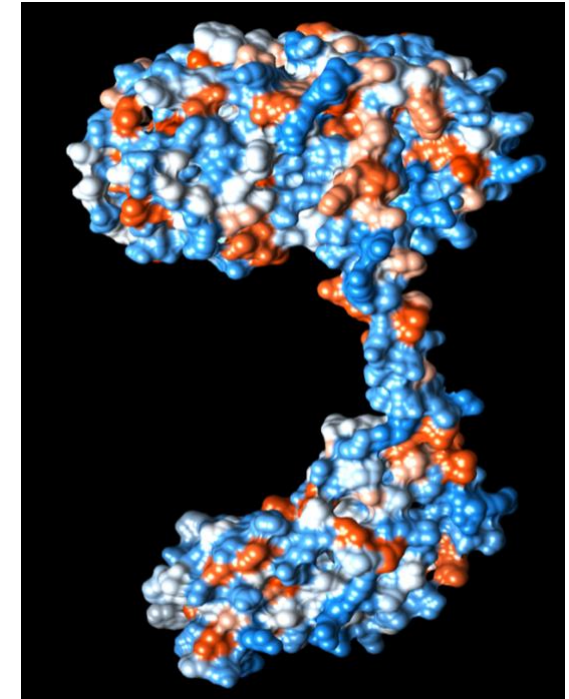
**Figura 12** - Perfil de hidrofobicidade em gradiente de cor, regiões em azul mais hidrofílicas e em vermelho mais hidrofóbicas na proteína dos domínios Rop\_like. P450, bZIP, respectivamente, A, B e C.



A



B



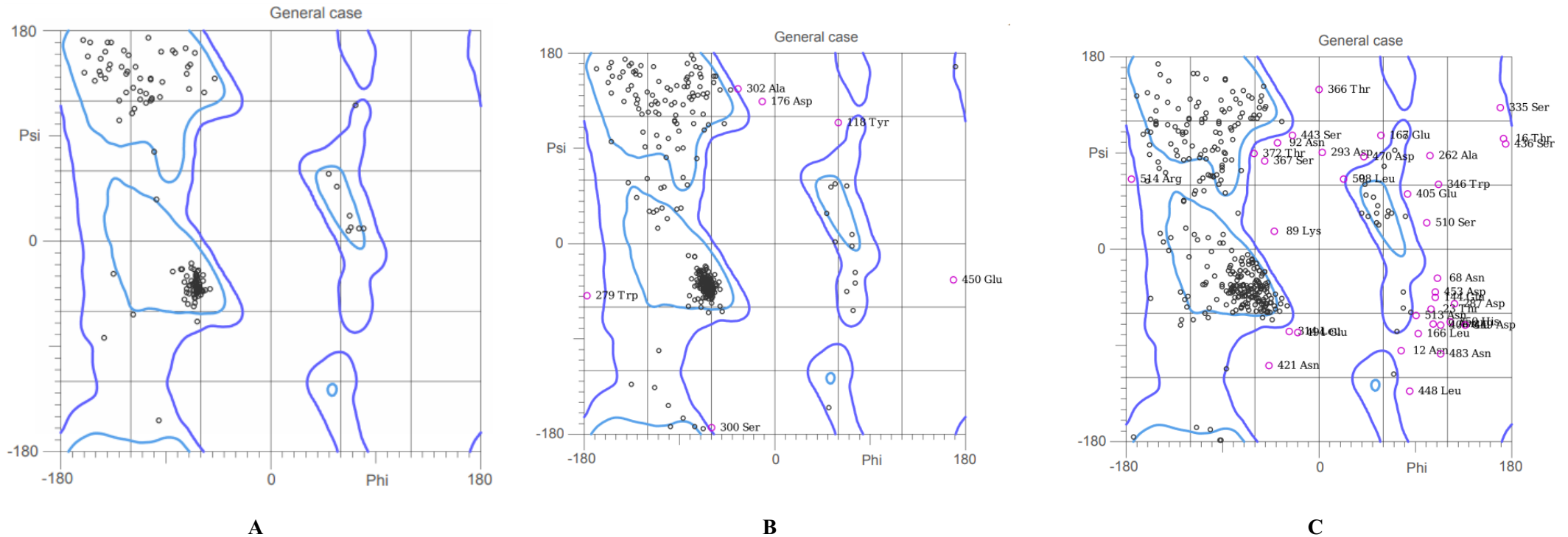
C

Fonte: Próprio autor.

Como etapa de validação do modelo tridimensional das proteínas, o gráfico Ramachandran, mostra as conformações estruturais da alfa-hélice e beta-folha, de forma a mostrar a angulação torcional, phi e psi, da proteína, analisando se esta estabeleceu uma posição favorável à sua estabilidade e atividade biológica, no gráfico estas regiões positivas encontram-se dentro das áreas azuis, as áreas ao redor destas são instáveis e não favoráveis a proteína (DA SILVA, 2010). Válido, enfatizar, que há exceções para glicina e prolina quanto as áreas de permissibilidade dentro

do gráfico, devido a uma maior flexibilidade e rigidez, respectivamente, além disso, para que a resolução da proteína esteja qualitativamente satisfatória, o gráfico deve apresentar valor igual ou maior que 90% de seus resíduos em áreas permissivas (MORRIS, et al., 1992).

**Figura 13** - Gráfico de Ramachandran da proteína dos domínios Rop\_like. P450, bZIP, respectivamente, A, B e C.

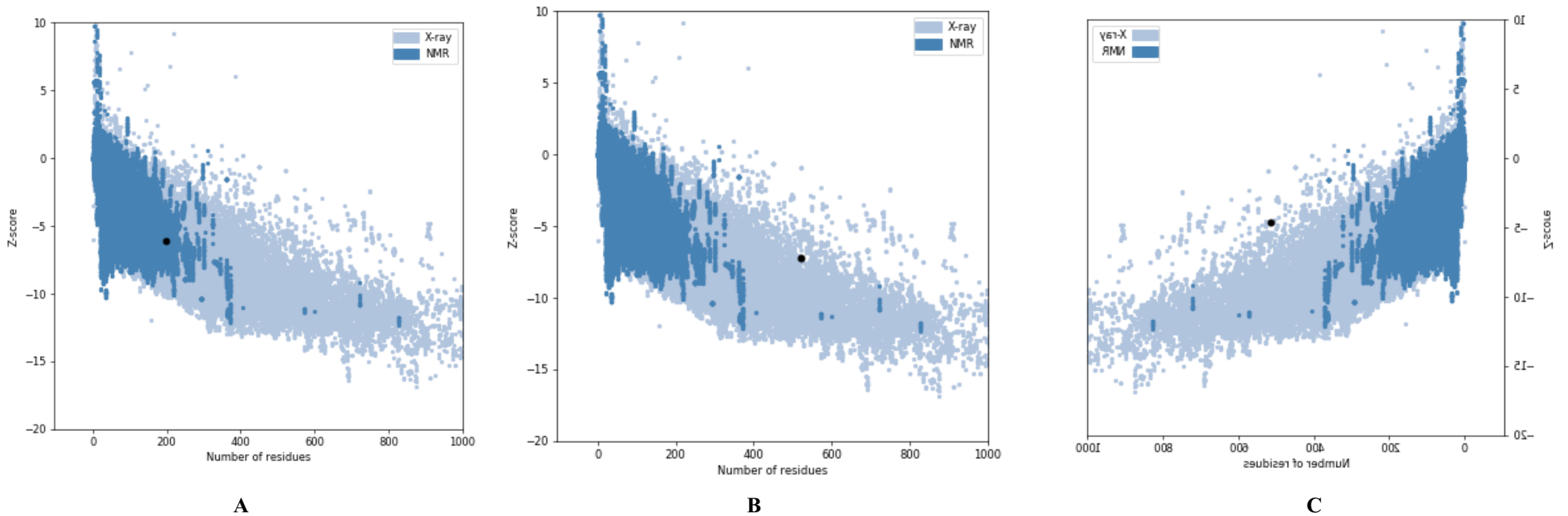


Fonte: Próprio autor.

Atentando-se ao resultado obtido pelo Molprobit, Figura 13, é possível concluir que para a proteína do domínio Rop\_like todos resíduos de aminoácidos estão em conformações favoráveis mostrando que estas estão em equilíbrio, estáveis dentro das áreas azuis (a); para a do domínio P450 vê-se que os resíduos de aminoácidos estão em conformações favoráveis mostrando que estas estão em equilíbrio, estáveis dentro das áreas

azuis, exceto por cinco resíduos: a alanina na posição 302, aspartato – 176, tirosina – 118, serina – 300 e glutamato – 450 (b); e do domínio bZIP, a maioria dos resíduos de aminoácidos estão em conformações favoráveis mostrando que estas estão em relativo equilíbrio, estáveis dentro das áreas azuis, apesar de grande quantidade de resíduo nas áreas brancas, não favoráveis (c), mostrando como modelos com melhor resolução, o (a) seguido pelo (b) e (c), respectivamente, de acordo com a quantidade de resíduos em conformações favoráveis.

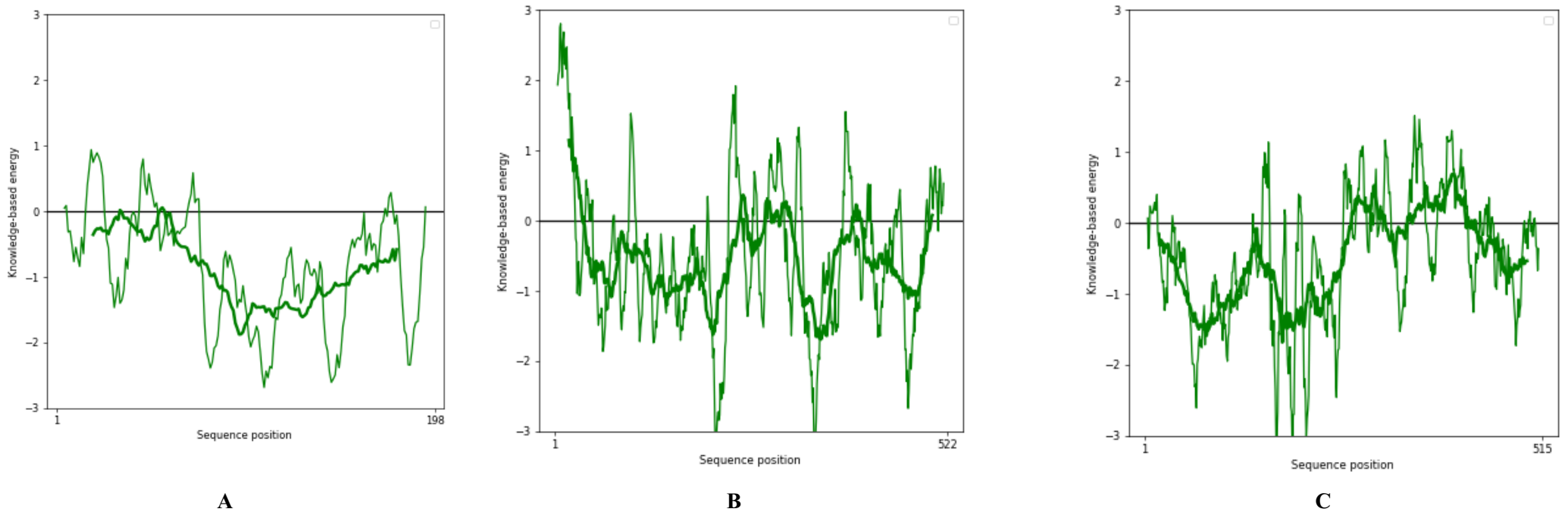
**Figura 14** - Análise do Z-score da proteína dos domínios Rop\_like. P450, bZIP, respectivamente, A, B e C.



Fonte: Próprio autor.

Como ferramentas estatísticas de refinamento qualitativo da validação da proteína, o gráfico de Z-score e o de energia dos scores, mostram a energia conformacional dos resíduos de aminoácidos e sua energia conformacional, a qual espera-se que seja negativo e esteja na faixa azul do gráfico (JIMENEZ, 2019), assim como abaixo da linha preta, a qual representa o Z-score, mostrando a qualidade e confiabilidade do modelo da estrutura 3D (DE OLIVEIRA, SALES, FREIRE 2020), fornecendo uma média de desvio padrão do valor que temos comparativamente a faixa já encontradas em proteínas nativas com o tamanho similar, e o gráfico de energia dos scores, faz a mesma coisa, só que não com a média geral, e sim a partir de fragmentos da sequência total, sendo o mais escuro o com intervalo menor 10 resíduos, e mais claro com hiato maior, 40 resíduos (DATA SCIENCE, 2020; PROSA-WEB, 2021).

**Figura 15** - Gráfico de energia dos stories da proteína dos domínios Rop\_like. P450, bZIP, respectivamente, A, B e C.



**Fonte:** Próprio autor.

Destarte, na Figura 14 (a), domínio Rop\_like, nota-se que o gráfico Z-score encontra-se na região azul mais escura e negativa no gráfico no valor de -6,07, na Figura 15 (a), além da energia mostrar-se favorável, pois está praticamente toda abaixo da linha preta do Z-score, fora um mínimo trecho que está sob a linha, sendo a com intervalo menor, verde claro, apresentando variação maior em relação à linha do Z-score, como esperado. Para a Figura 14 (b), domínio P450, o ponto preto, referente a média total do desvio padrão da proteína, se encontra na região azul e negativa no gráfico no valor de -7,18, na Figura 15 (b), apesar da posição da sequência inicial e final estar bem acima da linha preta do Z-score, o restante da sequência encontra-se próxima a linha ou bem abaixo, mais aconselhado, similar para as duas janelas de intervalo. Já para o domínio bZIP, Figura 14 (c), o valor do desvio padrão encontra-se na região azul, próxima da área branca, negativa no gráfico no valor de -4,66, na Figura 15 (c), vê-se que no início o modelo está favorável, abaixo da linha preta do Z-score, a outra metade da posição da sequência tem intervalos acima, e em seu final volta a se encontra negativa, apresentando uma leve atenuada na análise com intervalos maiores, em verde escuro.

Pode-se constatar a partir da análise estrutural, Figura 13,14,15, com o gráfico de Ramachandran e Z-score que, qualitativamente, a do domínio Rop\_like é mais estavelmente promissora, seguida da P450 e bZIP, respectivamente. Já no estudo tridimensional e nas análises de hidrofobicidade, Figura 11, 12, as três proteínas não apresentaram diferenças significativas, sobre o bom nível de dissolução em meio aquoso, mostrando maior facilidade de manipulação experimental devido a esta característica química, apesar estruturalmente, tamanho sequência e serem de domínios diferentes, ainda apresentam mesma função relacionada a tolerância, resistência ao estresse salino, em comum.

Após, esta etapa, foi desenhado os respectivos primers específicos das sequências estudadas para viabilização do estudo experimental relacionado ao estresse salino nas espécies do algodoeiro (Tabela 6,7,8), dentre as quais é possível ver tanto a sua sequência, o sentido desta, e outras informações importantes para viabilização do trabalho em bancada durante a execução do PCR para que se encontre a região de interesse no gene que traduzirá as proteínas dos domínios Rop\_like, P450 e Hsf8. Além da validação destes desenhos, apresentados pelos fatores físicos RInk, deltaS, deltaG, deltaH, relacionados com as propriedades termodinâmicas de entropia e energia de cada primer apontado. E o domínios de sequências conservadas é ponderada para predição de função (GOSSANI, 2014).

Ao observar os domínios funcionais das sequências de cada proteína relativa a cada domínio em questão (Rop\_like, P450, HSF8) na Figura 16,17,18, partilham da mesma estrutura, acima possui uma régua que mostra o intervalo da sequência gênica em estudo foi delimitada, no caso foi feito cortes nos intervalos funcionais de cada proteína de interesse, respectivamente (8 – 180, 3 – 522, 167 – 215), para um desenho de primer mais específico possível. Esta sequência do gene de tradução de cada proteína, está representada em vermelho, embaixo desta, em preto, estão os éxons que são o segmento que codifica parte da sequência de aminoácidos, e em azul, Figura 16,17, representa os primers existentes na plataforma no intervalo pesquisado, NCBI, não tendo resultado para o domínio, ratificado no gráfico da sequência completa, Figura 20,21,22.

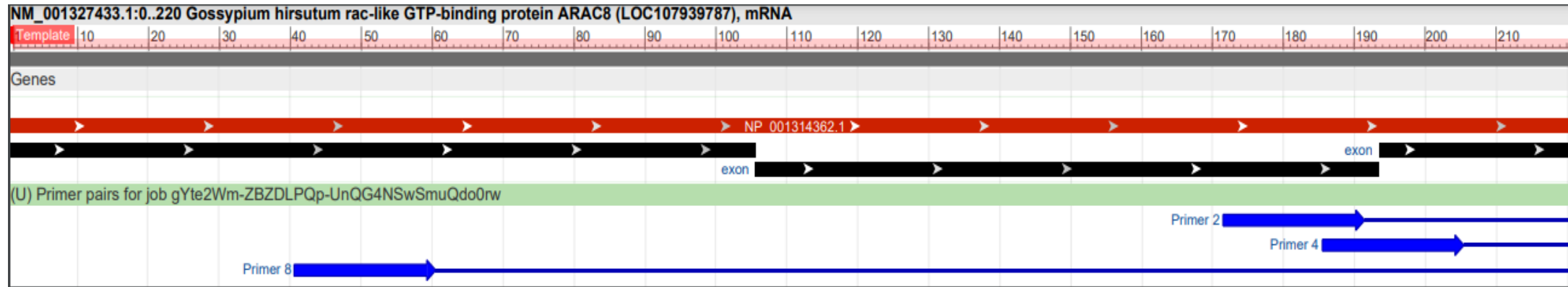
Assim, após o corte da sequência, pode-se analisar nos desenhos dos primers três fatores relacionados a sua qualidade: fator 1 - tamanho (preferível de 18 a 28 bases, mais específicos e estáveis), fator 2 - temperatura de anelamento (entre 50 a 60°C é o ideal), fator 3 - a porcentagem de guaninas e citosinas que possui (é desejado o intervalo de 45 a 55%). Todos estes aspectos tornam o primer específico, seguro e eficiente (MENDONÇA, 2016).

Explorando estas condições de trabalho do primer para avaliação do gene referente ao gene do domínio Rop\_like, Tabela 6. Correlacionando os desenhos obtidos em dois softwares diferentes, é presumível que os resultados atendem as especificações de tamanho, de porcentagem de citosina e guanina (%CG) somente o primer d encontra-se fora do limite esperado, e em relação a temperatura de nivelamento, todas, no sentido positivo e negativo, estão próximos aos 60°C. Assim, os mais promissores são o primer e, 2, b, dois adquiridos na plataforma Primer3Plus e um do NCBI.

Os desenho dos genes do domínio P450, Tabela 7, para o fator 1, todas estão conformes com o intervalo ideal, fator 2, com valores variando com margem de até 0,06°C, e o fator 3 está adequado, exceto pelo primer 7 com 60%, 5% a mais do intervalo de referência. Com base neste parecer, os primers mais favoráveis são os 5, c, cada um proveniente de um servidor.

A Tabela 8, referente ao domínio bZIP, com base nos parâmetros e os dados observados, somente o fator 2 encontra-se ligeiramente acima do intervalo normalizado, aponta como mais promissor, o primer c. Este, coletado no Primer3Plus, já que no Primer-Blast não obteve resultados nas localizações referentes aos domínios de conservação, intervalo referência pequeno (167 – 215).

**Figura 16** - Domínio da sequência conservada da proteína referente ao Rop\_like, com localização 8 - 180, juntamente com os primers presentes na plataforma neste intervalo, obtido no NCBI.



Fonte: Próprio autor.

**Tabela 6** - Primers da proteína pertencente ao domínio Rop\_like.

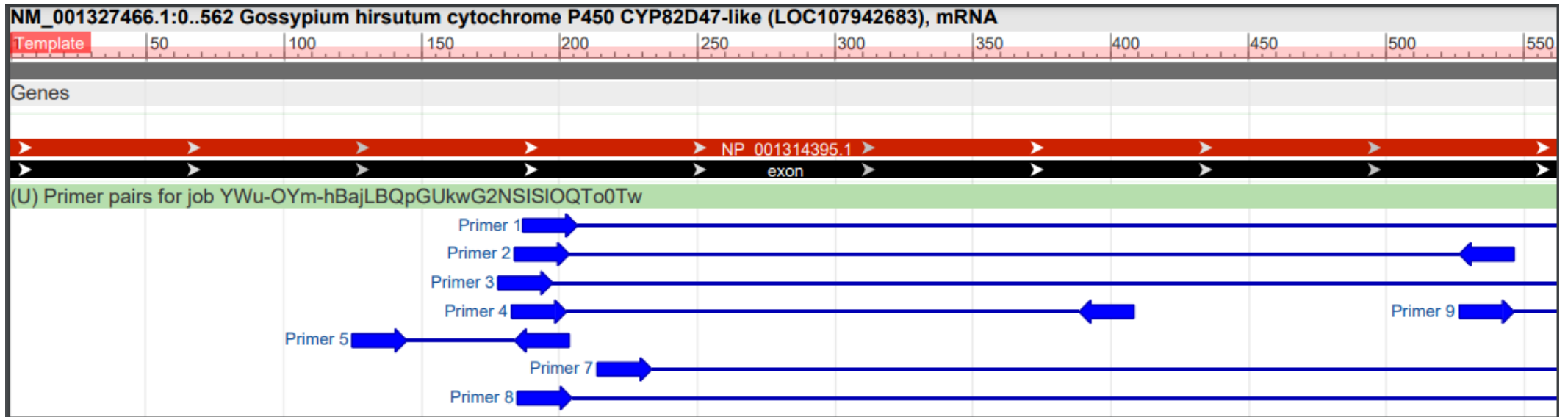
Primers	Tamanho	Sequência 5'→3'	Sentido	Temperatura de anelamento (°C)	%CG (Citosina e Guanina)	RInK Cal/(°K* <i>mol</i> )	DeltaS Kcal. <i>mol</i>	DeltaG Kcal. <i>mol</i>	DeltaH Kcal. <i>mol</i>	Plataforma
Primer 2	20	TTAGGTCTTTGGGACACGGC	Positivo	59,96	55	33,404	533,4	26,9	208,09	Primer-Blast - NCBI
	20	TTCTCATGGCTTGCTCGACT	Negativo	59,39	50					
Primer 4	20	CACGGCTGGACAAGAGGATT	Positivo	60,04	55	33,404	544,9	26,8	211,55	Primer-Blast - NCBI
	20	GTTCTCATGGCTTGCTCGAC	Negativo	59,28	55					
Primer 8	20	TTGGAGATGGAGCTGTTGGC	Positivo	60,32	55	33,404	552,5	26,5	214,83	Primer-Blast - NCBI
	20	ACGTTCTCATGGCTTGCTCG	Negativo	61,02	55					

Primer a	20	GTTGGAGATGGAGCTGTTGG	Positivo	60,7	55	33,404	545,1	25,7	211,35	Primer3Plus
	20	CCGTGTCCCAAAGACCTAAG	Negativo	59,6	55					
Primer b	19	TTGGAGATGGAGCTGTTGG	Positivo	59,8	52	33,404	515,6	24,2	199,95	Primer3Plus
	20	CCGTGTCCCAAAGACCTAAG	Negativo	59,6	55					
Primer c	21	CAGTTGGAGATGGAGCTGTTG	Positivo	60,8	52	33,404	566,4	26,8	218,88	Primer3Plus
	20	CCGTGTCCCAAAGACCTAAG	Negativo	59,60	55					
Primer d	22	TCAAGCGCTTCAAGATTTATCA	Positivo	60	36,4	33,404	561,4	26,7	213,37	Primer3Plus
	20	CCGTGTCCCAAAGACCTAAG	Negativo	59,6	55					
Primer e	20	AGTTGGAGATGGAGCTGTTG	Positivo	57,9	50	33,404	539,5	25,1	208,44	Primer3Plus
	20	CCGTGTCCCAAAGACCTAAG	Negativo	59,6	55					

Fonte: Próprio autor.



**Figura 17** - Domínio da sequência conservada da proteína referente ao P450, com localização 3 - 522, juntamente com os primers presentes na plataforma neste intervalo, obtido no NCBI.



Fonte: Próprio autor.

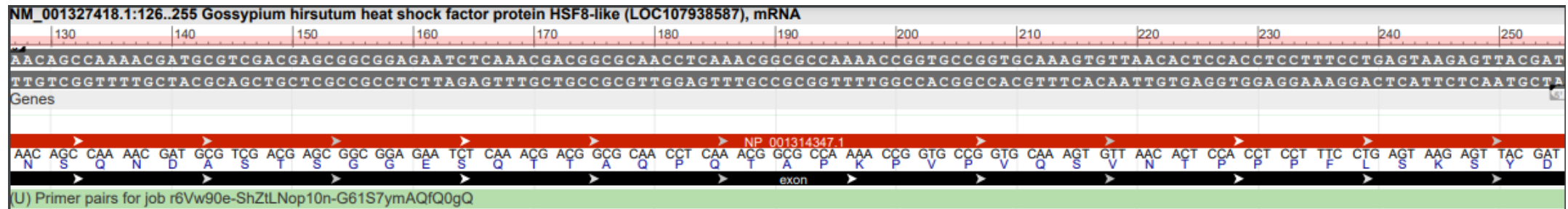
**Tabela 7** - Primers da proteína pertencente ao domínio P450.

Primers	Tamanho	Sequência 5'→3'	Sentido	Temperatura de anelamento (°C)	%CG (Citosina e Guanina)	RInK Cal/(°K* $\text{mol}^{-1}$ )	DeltaS Kcal.mol	DeltaG Kcal.mol	DeltaH Kcal.mol	Plataforma
Primer 1	20	GCGTTGGGAGACTTGGAAGA	Positivo	59,97	55	33,404	542,8	26,9	210,97	Primer-Blast - NCBI
	20	TACCAGCCACCGTCCTCATA	Negativo	60,03	55					

Primer 2	20	GAAGCGTTGGGAGACTTGGA	Positivo	59,97	55	33,404	542,8	26,9	210,97	Primer-Blast – NCBI
	20	CAACACCTTACCCGAGCCAT	Negativo	60,04	55					
Primer 3	20	CCTCATGAAGCGTTGGGAGA	Positivo	59,75	55	33,404	547,9	26,8	212,59	Primer-Blast - NCBI
	20	GTGCCTTTGGGGATGTCGTA	Negativo	60,04	55					
Primer 4	20	TGAAGCGTTGGGAGACTTGG	Positivo	60,25	55	33,404	537,2	27,1	208,97	Primer-Blast – NCBI
	20	GCGCATTTACGCCAGTATT	Negativo	59,90	50					
Primer 5	20	CCGGTGCATGGCCAATAATC	Positivo	59,69	55	33,404	539,5	26,6	208,99	Primer-Blast – NCBI
	20	TCCAAGTCTCCCAACGCTTC	Negativo	59,97	55					
Primer 7	20	GGACCTGCCTACATGATCCG	Positivo	59,97	60	33,404	559,7	26,8	217,93	Primer-Blast – NCBI
	20	TAAAGCTCAGGAGCAAGGCG	Negativo	60,39	55					
Primer 9	20	ATGGCTCGGGTAAGGTGTTG	Positivo	60,04	55	33,404	535,5	26,8	208,67	Primer-Blast - NCBI
	20	CCAAGGCTTCAAGCGGTGTA	Negativo	60,61	55					
Primer a	20	TTGGTGGTGAATTCGAGTGA	Positivo	60,1	45	33,404	529	25,3	203,4	Primer3Plus
	20	TTGAGCTGATCGATCCTGTG	Negativo	59,9	50					
Primer b	20	CGTTGGGAGACTTGGAAGAG	Positivo	59,8	50	33,404	533	26,1	206,57	Primer3Plus
	20	TCACGCCAGTATTGTCCGTA	Negativo	60,1	50					
Primer c	20	TGGTGGTGAATTCGAGTGAG	Positivo	59,7	50	33,404	537,1	25,6	207,08	Primer3Plus
	20	TTGAGCTGATCGATCCTGTG	Negativo	59,9	50					
Primer d	20	ATGAAGCGTTGGGAGACTTG	Positivo	60,3	50	33,404	531,2	25,9	204,96	Primer3Plus
	20	TCACGCCAGTATTGTCCGTA	Negativo	60,1	50					
Primer e	20	CATGAAGCGTTGGGAGACTT	Positivo	50,3	50	33,404	531,2	25,9	204,96	Primer3Plus
	20	TCACGCCAGTATTGTCCGTA	Negativo	60,1	50					

Fonte: Próprio autor.

**Figura 18** - Domínio da sequência conservada da proteína referente ao bZIP, com localização 167 - 215, juntamente com os primers presentes na plataforma neste intervalo, obtido no NCBI.



Fonte: Próprio autor.

**Tabela 8** - Primers da proteína pertencente ao domínio bZIP.

Primers	Tamanho	Sequência 5'→3'	Sentido	Temperatura de anelamento (°C)	%CG (Citosina e Guanina)	RInK Cal/(°K*mol)	DeltaS Kcal.mol	DeltaG Kcal.mol	DeltaH Kcal.mol	Plataforma
Primer a	18	AGCGGCGGAGAATCTCAA	Positivo	62	55,6	33,404	495,5	24,9	192,44	Primer3Plus
	22	TGGGCTCCAAGAACTATAGCA	Negativo	61,1	45,5					
Primer b	20	GGCGGAGAATCTCAAACGAC	Positivo	62,1	55	33,404	542	26,8	209,12	Primer3Plus
	20	TAGCATCGGTAGCAGGATCA	Negativo	59,4	50					
Primer c	19	GGCGGAGAATCTCAAACGA	Positivo	61,3	52,6	33,404	512,5	25,3	197,72	Primer3Plus
	20	TAGCATCGGTAGCAGGATCA	Negativo	59,4	50					
Primer d	20	GGCGGAGAATCTCAAACGAC	Positivo	62,1	55	33,404	542	26,8	209,12	Primer3Plus
	20	GGTAGCAGGATCATCCACCA	Negativo	60,9	55					
Primer e	19	GGCGGAGAATCTCAAACGA	Positivo	61,3	52,6	33,404	512,5	25,3	197,72	Primer3Plus
	20	GGTAGCAGGATCATCCACCA	Negativo	60,9	55					

Fonte: Próprio autor.

**Tabela 9** - Desenho de primers degenerados das três proteínas dos domínios Rop\_like, P450 e bZIP após alinhamento das sequências.

Primers	Tamanho	Sequência 5'→3'	Sentido	Temperatura de anelamento (°C)	%CG (Citosina e Guanina)	RlnK Cal/(°K* mol)	DeltaS Kcal.mol	DeltaG Kcal.mol	DeltaH Kcal.mol	Qualidade (%)	Plataforma																																																																																																																																																																																																					
Primer 1a	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP																																																																																																																																																																																																					
	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1							Primer 1b	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	58	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1c	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1d	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1e	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	57	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	53,4	-28,1	Primer 1f	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1
Primer 1b	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	58	FastP																																																																																																																																																																																																					
	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1							Primer 1c	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1d	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1e	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	57	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	53,4	-28,1	Primer 1f	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3										
Primer 1c	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	60	FastP																																																																																																																																																																																																					
	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1							Primer 1d	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1e	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	57	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	53,4	-28,1	Primer 1f	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																											
Primer 1d	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP																																																																																																																																																																																																					
	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1							Primer 1e	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	57	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	53,4	-28,1	Primer 1f	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																												
Primer 1e	23	AACTTCAGTGCCAACGTTGTAGT	Positivo	53,6	-28,5	33,404	499,5	29,7	189,6	57	FastP																																																																																																																																																																																																					
	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	53,4	-28,1							Primer 1f	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1	Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																													
Primer 1f	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	60	FastP																																																																																																																																																																																																					
	23	CACCTCTGTAGCTCAATGGCCTT	Negativo	55,3	-29,1							Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																														
Primer 1g	23	AGATGGAGCTGTTGGCAAGACTT	Positivo	54,9	-29,1	33,404	478	30,1	183,4	58	FastP																																																																																																																																																																																																					
	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1							Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8	Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																															
Primer 1h	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP																																																																																																																																																																																																					
	23	TTGTAATCCTCTTGTCCAGCCGT	Negativo	54,5	-28,8							Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5	Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																																																
Primer 1i	23	GGCACAACCTGTGAACTTAGGTCT	Positivo	53,6	-28,2	33,404	491,4	29,3	186,7	58	FastP																																																																																																																																																																																																					
	22	AGACATCTGCACCTCTGTAGCT	Negativo	53,5	-27,5							Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1	Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																																																																	
Primer 1j	22	GAGCTGTTGGCAAGACTTGCAT	Positivo	54,9	-28,3	33,404	457,1	29	175,8	58	FastP																																																																																																																																																																																																					
	23	CGTGTCCCAAAGACCTAAGTTCA	Negativo	53,4	-28,1							Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7	Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																																																																																		
Primer 1k	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP																																																																																																																																																																																																					
	23	CGGTGAAGATTTCTTGGCTACC	Negativo	54,6	-28,7							Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5	Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																																																																																																			
Primer 1l	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP																																																																																																																																																																																																					
	21	TTGCGCATTTACGCCAGTAT	Negativo	54,2	-27,5							Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																																																																																																																				
Primer 1m	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	28,8	181,1	59	FastP																																																																																																																																																																																																					
	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3																																																																																																																																																																																																											

Primer ln	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	—————	28,8	181,1	59	FastP
	23	TTGGCTACCTCACTCGAATTCAC	Negativo	53,6	-28,3							
Primer lo	23	GTAGCCAAGGAAATCTTCACCGT	Positivo	53,8	-28,3	33,404	500	—————	29,9	190	58	FastP
	23	TTTCACGCCAGTATTGTCCGTAA	Negativo	53,3	-28,3							
Primer lp	23	GTAGCCAAGGAAATCTTCACCGT	Positivo	53,8	-28,3	33,404	500	—————	29,9	190	59	FastP
	21	TTGCGCATTTCACGCCAGTAT	Negativo	54,2	-27,5							
Primer lq	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	—————	28,8	181,1	60	FastP
	21	CCAACACCTTACCCGAGCCAT	Negativo	55,2	-27,6							
Primer lr	22	TCCTTACGGACAATACTGGCGT	Positivo	54,1	-28	33,404	489,2	—————	29,7	186,4	59	FastP
	21	CCAACACCTTACCCGAGCCAT	Negativo	55,2	-27,6							
Primer ls	22	CATGGCCAATAATCGGTCACCT	Positivo	53,6	-27,5	33,404	475,2	—————	28,8	181,1	59	FastP
	22	ACGCCAGTATTGTCCGTAAGGA	Negativo	54,1	-28							
Primer lt	20	GGCAAAGCTCCACCTGTAGC	Positivo	54,4	-26,4	33,404	429,7	—————	27,1	165,4	60	FastP
	21	TTGCGCATTTCACGCCAGTAT	Negativo	54,2	-27,5							

Fonte: Próprio autor.

Para a Tabela 9, no desenho do primer degenerado, foi utilizado as sequências cortadas das três proteínas e as usou para gerar um primer mais “geral”, pois tem o objetivo de cobrir todas regiões possíveis referentes a síntese das proteínas de interesse, no gene. Como parâmetros para análise de qualidade deste primer, o intervalo de tamanho é de 22 a 24 pares de bases, temperatura de anelamento aproximadamente a 56°C, e porcentagem de citosina e guanina próxima ou superior a 50% (MALONE, et al., 2006). Assim, de acordo com os resultados que obtidos, pode-se ver uma baixa porcentagem de citosina e guanina, nos desenhos obtidos, a temperatura está ligeiramente abaixo de 56°C, assim como a quantidade de pares de bases. A partir dos dados da Tabela 9, os resultados li, lt, lc, ld, como os primers degenerados mais promissores.

## 6 CONSIDERAÇÕES FINAIS

Os problemas ambientais, lesões que acometem as plantas e decaimento produtivo devido ao estresse salino, principalmente no nordeste, faz-se necessário o uso do melhoramento genético vegetal para minimização destas consequências doravante a criação de uma cultivar tolerante e/ou resistente.

O estudo *in silico*, possibilitou o conhecimento do gênero taxonomicamente, comportamento da planta com contato salino e triagem de três proteínas com função relacionada a este fator abiótico, promovendo estudo físico-químico destas e análise de domínios conservados das sequências para o desenho de primers mais estáveis, seguros, específicos e validação de sua funcionalidade. Além da avaliação estrutural destas para compreensão do comportamento das proteínas e qualidade conformacional de sua estrutura.

Assim, a partir dos resultados computacionais, principalmente os dez primers mais promissores selecionados das proteínas isoladas, as quais são: e, 2, b do domínio Rop\_like; 5, c do domínio P450; c do domínio bZIP, e em conjunto, degenerados, 1i, 1t, 1c, 1d, os quais atestam a fundamental importância destas ferramentas *in silico* para realização experimental da identificação do éxon de interesse no gene com os marcadores moleculares e desenvolvimento da pesquisa de manipulação genética à salinidade em algodoeiro.

## REFERÊNCIAS

AGUIAR, P. H.; MODESTO FILHO, I.; DOS REIS, C. R. Semeadura na época certa garante sucesso da “safrinha”. **Visão Agrícola**, n. 6, p. 78 - 80, 2006.

Algodão no mundo. **ABRAPA**, 2021. Disponível em: <<https://www.abrapa.com.br/Paginas/dados/algodao-no-mundo.aspx>>. Acesso em: 26 de abr. de 2021.

AMORIM, J. R. A.; et al. **Determinação da eficiência de uso da água na parcela de irrigação, no perímetro irrigado Califórnia, em Sergipe**. Aracaju: Embrapa Tabuleiros Costeiros. 2010. 6 p.

ANDRADE, R.; et al. Características tecnológicas de fibras algodoeiro colorido BRS Rubi via fertirrigação In: Congresso Brasileiro De Engenharia Agrícola, 38. 2009, Juazeiro/Petrolina. 2009. **Anais [...]**. Juazeiro: Associação Brasileira de Engenharia Agrícola, 2009.

ANOWER M, R.; et al. Physiological processes associated with salinity tolerance in an alfalfa half-sib family. **Journal of Agronomy and Crop Science**, v. 203, n. 6, p. 506 – 518, 2017.

ARAÚJO, W. P. **Irrigação fenológica em duas cultivares de algodoeiro herbáceo**. 2018. Doutorado (Tese em Engenharia Agrícola) – Universidade Federal de Campina Grande, Campina Grande, 2018, 132p. Disponível em: <<http://dspace.sti.ufcg.edu.br:8080/jspui/bitstream/riufcg/4502/1/WH%C3%89LLYSON%20PEREIRA%20ARA%C3%9AJO%20%E2%80%93%20TESE%20%28PPGEA%29%202018.pdf>>. Acesso em: 28 de abr. de 2021.

AZEVEDO NETO, A. D.; et al. Physiological and biochemical responses of peanut genotypes to water deficit. **Journal of Plant Interactions**, v. 5, n. 1, p. 1 – 10, 2010.

BARBORA, F. B. DA C. Relações de produção na agricultura: algodão no Maranhão (1760 a 1888) e café em São Paulo (1886 a 1929). **Agríc. São Paulo**, São Paulo, v. 52, n. 2, p. 17 – 27, dez. 2005.

BARROS, J. A. S. **Biomassa das plantas de Algodão (*Gossypium hirsutum* L) submetido a diferentes níveis de salinidade e substâncias húmicas, em ambiente protegido**. 2011. Monografia (Bacharel em Engenharia Agrônômica) – Universidade Federal de Alagoas, Rio Largo, 2011, 31p. Disponível em: <<https://ceca.ufal.br//pt-br/graduacao/agronomia/documentos/tcc/tcc-2011/jiose-a-s-barros>>. Acesso em: 18 de abr. de 2021.

BERGAMIN, C. **A revolução industrial e as inovações na produção têxtil: contexto para o desenvolvimento da moda no século XIX**. 2018. Monografia (Bacharel em História) – Universidade Federal de Santa Catarina, 2018, 53p. Disponível em: <<https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/197129/TCC%20Camila%20Bergamin%2009265007.pdf?sequence=1&isAllowed=y>>. Acesso em: 27 de abr. de 2021.

BICUDO, C. E. DE M. Taxonomia. **Biota Neotrop.**, Campinas, v. 4, n. 1, 2004.

BLANCO, B.S, Gossipol e fatores antinutricionais da soja. In: Toxicologia Aplicada á Medicina Veterinária, H. S. Spinosa, S. L. G’orniak, and J. P. Neto, Eds. **Manole**, Barueri, p. 531–545, 2008.

BLANCO, F. F.; et al. Crescimento e produção de cultivares de algodão sob irrigação com déficit. In: Congresso Brasileiro de Algodão & I Cotton Expo, VIII, 2011, São Paulo. **Anais [...]**. Campina Grande: Embrapa Algodão, 2011, p; 1753-1760.

BOARETO, M.; et al. Relationship between global structural parameters and Enzyme Commission hierarchy: Implications for function prediction. **Comput. Biol. Chem.**, v. 40, p. 15 – 19, 2012.

BORÉM, A.; FREIRE, E. C. **Algodão do Plantio a Colheita**. 1 ed. Viçosa: UFV, 2014. 312 p.

BRUM, I. J. B. **Ferramentas de bioinformática para proteômica**. 2007. Dissertação (Mestrado em Biologia Funcional e Molecular) – Universidade Estadual de Campinas, Campinas, 2007, 83p. Disponível em: <[http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/314739/1/Brum\\_ItarajuJuniorBaracuh\\_M.pdf](http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/314739/1/Brum_ItarajuJuniorBaracuh_M.pdf)>. Acesso em: 04 de mai. de 2021.

CALDART, E. T.; et al. Análise filogenética: conceitos básicos e suas utilizações como ferramenta para virologia e epidemiologia molecular. **Acta Scientiae Veterinariae**, v. 44, n. 1392, p. 1 – 20, 2016.

CARDOSO, D. B. O. **Melhoramento genético de algodoeiro colorido: Redes Neurais Artificiais versus métodos convencionais**. 2018. Dissertação (Mestrado em Agronomia) – Universidade Federal de Uberlândia, Uberlândia, 2018, 99p. Disponível em: <<https://repositorio.ufu.br/bitstream/123456789/21630/3/Melhoramentogen%C3%A9ticoalgodoeiro.pdf>>. Acesso em: 29 de abr. de 2021.

CARVALHO, M. C. S. Adubação de cobertura do algodoeiro cultivado em condições de sequeiro na região do Cerrado – Comunicado Técnico, 347. Campina Grande: **Embrapa Algodão**. 5 p, 2007.

CHE-OTHMAN, M. H.; MILLAR, A. H.; TAYLOR, N. L. Connecting salt stress signalling pathways with salinity-induced changes in mitochondrial metabolic processes in C3 plants. **Plant Cell & Environ**, v. 40, n. 12, p. 2875 – 2905, 2017.

CORREIA, R. C.; et al. A região semiárida brasileira. In: VOLTOLLINI, T. V. **Produção de caprinos e ovinos no Semiárido**. Petrolina: Embrapa Semiárido, 2011, p. 21-48.

COUTO, G. H. P.; RAMOS, M. M. DE S.; DE CASTRO, M. T. Eficiência de *Bacillus amyloliquefaciens* e *Trichoderma asperellum* em tratamento de sementes de algodão para ganhos em produtividade. In: Simpósio de TCC – Seminário de Iniciação Científica, XX, 2020, Brasília. **Anais [...]**. Brasília: Centro Universitário ICESP – FACICESP, 2020.

Cultura do algodão herbáceo na agricultura familiar. **EMBRAPA-ALGODÃO**, 2021. Disponível em: <[https://www.spo.cnptia.embrapa.br/conteudo?p\\_p\\_id=conteudoportlet\\_WAR\\_sistemasdeproducaolf6\\_1galceportlet&p\\_p\\_lifecycle=0&p\\_p\\_state=normal&p\\_p\\_mode=view&p\\_p\\_col\\_id=column-2&p\\_p\\_col\\_pos=1&p\\_p\\_col\\_count=2&p\\_r\\_p\\_-76293187\\_sistemaProducaoId=3718&p\\_r\\_p\\_-996514994\\_topicoId=3300](https://www.spo.cnptia.embrapa.br/conteudo?p_p_id=conteudoportlet_WAR_sistemasdeproducaolf6_1galceportlet&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_pos=1&p_p_col_count=2&p_r_p_-76293187_sistemaProducaoId=3718&p_r_p_-996514994_topicoId=3300)>. Acesso em: 01 de mai. de 2021.

D'ALMEIDA, D. M. B. A. D.; et al. Importância relativa dos íons na salinidade de um Cambissolo na Chapada do Apodi, Ceará. **Engenharia Agrícola**, Jaboticabal, v.25, n.3, p.615 - 621, 2005.



- DA AZEVEDO, P. V.; SILVA, F. D. DOS S. Risco climático para o cultivo do algodoeiro na região nordeste do Brasil. **Revista Brasileira de Meteorologia**, v. 22, n. 3, p. 408-416, 2007.
- DA COSTA, A. R. **Nutrição mineral em plantas vasculares: Apontamentos de fisiologia vegetal**. Portugal: Escola de Ciências e Tecnologia da Universidade de Évora. dez. 2014, 139p.
- DA SILVA, A. A. **Adubação nitrogenada em cobertura na cultura do algodão, Vilhena-RO**. 2020. Monografia (Bacharel em Agronomia) – Faculdade da Amazônia, Vilhena, 2020, 31p. Disponível em: <<http://repositorio.famaro.com.br/bitstream/123456789/172/1/TCC%20FINAL%20ANDERSON.pdf>>. Acesso em: 24 de abr. de 2021.
- DA SILVA, E. C.; et al. Physiological responses to salt stress in young umbu plants. **Environmental and Experimental Botany**, v. 63, p. 147 – 157, 2008.
- DA SILVA, E. H. **Variabilidade genética e fisiológica de populações de *Meloidogyne incognita* e identificação de QTLs de uma nova fonte de resistência do algodoeiro (*Gossypium spp.*) a esse nematoide**, 2015. Disponível em: <<https://repositorio.unb.br/handle/10482/17489>>. Acesso: 15 de abr. de 2021.
- DA SILVA, K. N. **Modelagem comparativa da metaloprotease vmbmp de *Bothriopsis moojeni* por homologia estrutural**. 2010. Dissertação (Mestrado em Biotecnologia e Recursos Naturais) – Universidade do Estado do Amazonas, Manaus, 2010, 103p. Disponível em: <<https://pos.uea.edu.br/data/area/titulado/download/18-4.pdf>>. Acesso em: 15 de mai. 2021.
- DDBJ. **Volumes de dados de envio e liberação**, 2021. Disponível em: <<https://www.ddbj.nig.ac.jp/statistics/index-e.html>>. Acesso em: 16 de abr. de 2021.
- DE ALMEIDA, V. C. **Caracterização genética e in situ de *Gossypium barbadense* na região norte do Brasil**. 2007. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Federal do Rio Grande do Norte, Natal, 2007, 77p. Disponível em: <<https://repositorio.ufrn.br/jspui/bitstream/123456789/16772/1/VanessaCA.pdf>>. Acesso em: 26 de abr. de 2021.
- DE ANDRADE JÚNIOR, A. S.; et al. Zoneamento de aptidão climática para o algodoeiro herbáceo no Estado do Piauí. **Revista Ciência Agronômica**, Fortaleza, v.40, n. 2, p. 175 – 184, 2009.
- DE ARRUDA JÚNIOR, R. G. **Temperatura de melting: um estudo comparativo**. 2010. Monografia (Bacharel em Ciência da Computação) – Universidade Federal de Mato Grosso do Sul, Campo Grande, 2010, 55p. Disponível em: <<http://www.facom.ufms.br/wp-content/uploads/2015/11/Temperatura-de-Melting.pdf>>. Acesso em: 05 de mai. de 2021.
- DE CARVALHO, C.; et al. **Anuário brasileiro do algodão 2012**. Santa Cruz do Sul: Editora Gazeta Santa Cruz, 2012. 136 p.
- DE CARVALHO, L. P. Uso de recursos genéticos de algodão gera produtos de vestuário para comércio interno e externo. **Revista RG News**, v. 2, n. 2, p. 29 – 35, 2016.
- DE LIMA, E. B. **Deteção de subfamílias proteicas isofuncionais utilizando integração de dados e agrupamentos espectral**. 2015. Tese (Doutorado em Bioinformática) – Universidade Federal de Minas Gerais, Belo Horizonte, 2015, 194p. Disponível em: <<http://www.pgbioinfo.icb.ufmg.br/defesas/61D.PDF>>. Acesso em: 06 de mai. de 2021.
- DE MEDEIROS, N. M. C. **Análise filogenética e funcional de dois genes de reparo homólogos a AP endonuclease em cana-de-açúcar: ScARP1 e ScARP3**. 2014. Dissertação

(Mestrado em Bioquímica) – Universidade Federal do Rio Grande do Norte, Natal, 2014, 109p. Disponível em: <[https://repositorio.ufrn.br/bitstream/123456789/12627/1/An%c3%a1liseFilogen%c3%a9ticaFuncional\\_Medeiros\\_2014.pdf](https://repositorio.ufrn.br/bitstream/123456789/12627/1/An%c3%a1liseFilogen%c3%a9ticaFuncional_Medeiros_2014.pdf)>. Acesso em: 25 de abr. de 2021.

DE MENDONÇA, I. B. **Acompanhamento das principais técnicas biomoleculares no laboratório de biotecnologia e biologia molecular e no núcleo de genômica e bioinformática.** Relatório de Estágio Obrigatório. Fortaleza, 2016.

DE MENEZES, I. P. P. **Caracterização in situ e diversidade genética de algodoeiros mocós (*Gossypium hirsutum* raça marie galante L. Hutch) da região nordeste do Brasil.** 2009. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade Federal do Rio Grande do Norte, Natal, 2009, 91p. Disponível em: <<https://repositorio.ufrn.br/jspui/bitstream/123456789/16764/1/IvandilsonPPM.pdf>>. Acesso em: 01 de mai. De 2021.

DE OLIVEIRA JÚNIOR, I. S. Algodão herbáceo (*Gossypium hirsutum* r.latifolium Hutch). **Instituto Agronômico de Pernambuco**, Recife, jun. de 2009. Disponível em: <[http://www.ipa.br/pdf/folder\\_algodao.pdf](http://www.ipa.br/pdf/folder_algodao.pdf)>. Acesso em: 28 de abr. de 2021.

DE OLIVEIRA, A. M.; et al. Irrigação com água salina no crescimento inicial de três cultivares de algodão. **Irriga**, Botucatu, v. 13, n. 4, p. 467 – 475, dez. de 2008.

DE OLIVEIRA, C.; SALES, L. S.; FREIRE, J. E. DA C. Modelagem estrutural de uma nova quitinase (GH19) de melão, e análise de docking molecular com N-acetil- $\beta$ (1-4)-D-glicosamina. **Revista Arquivos Científicos (IMMES)**, Macapá, v. 3, n. 1, p. 162 – 171, 2020.

DE OLIVEIRA, F. A.; et al. Sensibilidade do algodoeiro ao cloreto de mepiquat em condições salinas. **Revista Ciência Agronômica**, v. 43, n. 3, p. 484 – 492, 2012.

DE SENA, L. M. M. Conheça e Preserve a Caatinga – O Bioma Caatinga. **Associação Caatinga**, Fortaleza, v. 1, 54p., 2011.

DE SOUSA, L. B. O algodoeiro: alguns aspectos importantes da cultura. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, Mossoró, v. 5, n. 4, p. 19-26, dez. 2010.

DE SOUSA, L. B. O algodoeiro: alguns aspectos importantes da cultura. Mossoró, **Revista Verde**, v. 5, n. 4, p. 19 – 26, dez. de 2010.

DE VASCONCELOS, M. F. **Comportamento de cultivares do algodoeiro herbáceo (*Gossypium hirsutum* L.) sob diferentes percentagens de sódio trocável.** 1990. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal de Campina Grande, Campina Grande, 1990, 76p. Disponível em: <<http://dspace.sti.ufcg.edu.br:8080/xmlui/handle/riufcg/11600>>. Acesso em: 01 de mai. de 2021.

DIAS, N. D.; BLANCO, F. F. Efeitos dos sais no solo e na planta. In: GHEYI, H. R.; DIAS, N. S.; LACERDA, C. F. (Ed.). **Manejo da salinidade na agricultura: Estudos básicos e aplicados.** 1 ed. Fortaleza: Instituto Nacional de Ciência e Tecnologia em Salinidade, 2010 p. 129 – 140.

DO NASCIMENTO, L. **Lapoge database - estrutura relacional de banco de dados biológicos para armazenamento e sistematização de informações genéticas e epidemiológicas.** 2016. Monografia (Bacharel em Ciências Biológicas) – Universidade Federal de Santa Catarina, Florianópolis, 2016, 88p. Disponível em:

<<https://repositorio.ufsc.br/xmlui/bitstream/handle/123456789/174993/TCC%20Leticia%20Nascimento.pdf?sequence=1&isAllowed=y>>. Acesso em: 15 de abr. de 2021.

DO REGO, F. O. R. **Análise *in silico* de proteínas: modelo baseado na apolipoproteína-e relacionada à doença de Alzheimer.** 2012. Monografia (Bacharel em Ciências Biológicas) – Universidade Federal da Bahia, Salvador, 2012, p.46. Disponível em: <<https://repositorio.ufba.br/ri/bitstream/ri/6617/1/Monografia%20-%20Fernanda%20Orpinelli%20R%20Rego.pdf>>. Acesso em: 15 de mai. de 2021.

DO REGO, F. O. R. **Análise *in silico* de proteínas: modelo baseado na apolipoproteína-e relacionada à doença de Alzheimer.** 2012. Monografia (Bacharel em Ciências Biológicas) – Universidade Federal da Bahia, Salvador, 2012, 46p. Disponível em: <<https://repositorio.ufba.br/ri/bitstream/ri/12608/1/Monografia%20-%20Fernanda%20Orpinelli%20R%20Rego.pdf>>. Acesso em: 24 de abr. de 2021.

E SILVA, S. DE A.; NOTARI, D. L.; DALL’ALBA, G. **Bioinformática: Contexto computacional e aplicações.** Caxias do Sul: **Educ**, 2020, 299p.

EPSTEIN, E.; BLOOM, A. J. **Nutrição mineral de plantas: princípios e perspectivas.** 2 ed. Londrina: Editora Planta, 2006, 404p.

FARIAS, S. G. G. et al. Estresse salino no crescimento inicial e nutrição mineral de *Gliricidia* (*Gliricidia sepium* (Jacq.) Kunt ex Steud) em solução nutritiva. **Revista Brasileira de Ciência do Solo**, v. 33, n. 5, p. 1499 – 1505, 2009.

FERNANDES, A. R. et al. Nutrição mineral de mudas de pupunheira sob diferentes níveis de salinidade. **Pesquisa Agropecuária Brasileira**, v. 37, n.11, p. 1613 – 1619, 2002.

FERREIRA, N. A. DE M. **Bioinformática como ferramenta no melhoramento genético de plantas.** 2017. Monografia (Bacharel em Ciências Biológicas) – Universidade Federal do Espírito Santo, Vitória, 2017, 39p. Disponível em: <[https://biologia.ufes.br/sites/cienciasbiologicas.ufes.br/files/field/anexo/mauricio\\_alexander\\_de\\_moura\\_ferreira\\_.pdf](https://biologia.ufes.br/sites/cienciasbiologicas.ufes.br/files/field/anexo/mauricio_alexander_de_moura_ferreira_.pdf)>. Acesso em: 15 de abr. de 2021.

FERREIRA, R. G.; et al. Distribuição da matéria seca e composição química das raízes, caule e folhas de goiabeira submetida a estresse salino. **Pesquisa Agropecuária Brasileira**, v. 36, n. 1, p. 79 – 88, 2001.

FONSECA JÚNIOR, A. A. **Guia Rápido de Bioinformática: PCR, Sequenciamento, Blast e Filogenia para Iniciantes!** 7 ed. São Paulo: Bizantium, 2014, p. 139.

FREIRE, M. B. G. S. & FREIRE, F. J. Fertilidade do solo e seu manejo em solos afetados por sais. In: NOVAIS, R. F.; et al. (Ed.). **Fertilidade do solo.** Viçosa: Sociedade Brasileira de Ciência do Solo, 2007, p. 929 – 954.

FREITAS, J. B. S. **Respostas fisiológicas ao estresse salino de duas cultivares de feijão-caupi.** 2006. Tese (Doutorado em Bioquímica) – Universidade Federal do Ceará, Fortaleza, 2006, 135p. Disponível em: <[http://www.repositorio.ufc.br/bitstream/riufc/10405/1/2006\\_tese\\_jbsfreitas.pdf](http://www.repositorio.ufc.br/bitstream/riufc/10405/1/2006_tese_jbsfreitas.pdf)>. Acesso em: 18 de abr. de 2021.

FUZATTO, M. G.; et al. Algodão (*Gossypium hirsutum* L.). In: AGUIAR, A. T. E.; et al. (Ed.). **Instruções agrícolas para as principais culturas econômicas – Boletim IAC n.º 200.** 7 ed. Campinas: Instituto Agrônomo, 2014, p. 11 – 14.

GHEYI, H. R.; FAGERIA, N. K. Efeitos dos sais sobre as plantas. In: GHEYI, H. R.; DIAS, N. DA S.; LACERDA, C. F. **Manejo e controle da salinidade na agricultura irrigada**. Campina Grande: UFCG, p. 125-131, 1997.

GOSSANI, C. M. D. **Evolução de genes que conferem resistência a múltiplas drogas em fungos**. 2014. Dissertação (Mestrado em Biociências e Biotecnologia) – Universidade Estadual do Norte Fluminense Darcy Ribeiro, Compos dos Goytacazes, 2014, 59p. Disponível em: <[https://uenf.br/posgraduacao/biociencias-biotecnologia/wp-content/uploads/sites/12/2020/12/dissertacao\\_cristiani\\_gossani.pdf](https://uenf.br/posgraduacao/biociencias-biotecnologia/wp-content/uploads/sites/12/2020/12/dissertacao_cristiani_gossani.pdf)>. Acesso em: 06 de mai. de 2021.

GUERRA, R. A. T.; et al. **Ciências biológicas: Cadernos CB Virtual 1**. João Pessoa: Ed. Universitária, 2011. 516p.

HANCOCK, J. F. Contributions of domesticated plant studies to our understanding of plant evolution. **Annals of Botany**, v. 96, n. 6, p. 953 – 963, 2005.

HENNING, A. A. **Patologia e tratamento de sementes: Noções Gerais – Documentos 264**. 2 ed. Londrina: Embrapa. 2005. 52p.

HOFFMANN NETO, E. G. **Tecnologias para a produção de forragem em solos de cerrado do Brasil central**. 1999. Monografia (Bacharel em Agronomia) – Universidade Federal de Santa Catarina, Florianópolis, 1999, 67p. Disponível em: <<https://repositorio.ufsc.br/bitstream/handle/123456789/117714/138632.pdf?sequence=1>>. Acesso em: 06 de abr. de 2021.

HOFFMANN, L. V.; et al. **Algodão nativo brasileiro: mitigação do risco de extinção – Documentos 279**. 1ed. Campina Grande: Embrapa Algodão, 2020, 21p.

HOLANDA, A. C.; et al. Desenvolvimento inicial de espécies arbóreas em ambientes degradados por sais. **Revista de Biologia e Ciências da Terra**, Campina Grande, v. 7, n. 1, p. 39-50, 2007. Disponível em: <<http://joaootavio.com.br/bioterra/workspace/uploads/artigos/ambientesdegradadosporsais-51817e5ee3d71.pdf>>. Acesso em: 21 fev. 2011.

IFRN. **Fibras**, 2021. Disponível em: <<http://www2.ifrn.edu.br/teciteca/pdf/fibras.pdf>>. Acesso em: 21 de abr. de 2021.

INSTITUTO AGRONÔMICO DE CAMPINAS. **Algodão: Gossypium hirsutum L.**, 2021. Disponível em: <[http://www.iac.sp.gov.br/imagem\\_informacoestecnologicas/31.pdf](http://www.iac.sp.gov.br/imagem_informacoestecnologicas/31.pdf)>. Acesso em: 20 de abr. de 2021.

JÁCOME, A. G. **Crescimento e produção de genótipos de algodoeiro, em um solo salinosódico**. 1999. Dissertação (Mestrado em Engenharia Agrícola) – Universidade Federal de Campina Grande, Campina Grande, 1999, 112p. Disponível em: <<http://dspace.sti.ufcg.edu.br:8080/xmlui/bitstream/handle/riufcg/9902/ALECKSANDRA%20GOMES%20J%20%81COME%20-%20DISSERTA%20%87%20%83O%20PPGEA%201999.pdf?sequence=1&isAllowed=y>>. Acesso em: 28 de abr. de 2021.

JAMIL, M.; et al. Salinity reduced growth PSII photochemistry and chlorophyll content in radish. **Scientia Agricola** v. 64, n. 2, p. 111 – 118, 2007.

JHA, U. C. et al. Functional Genomics: Applications in Plant Science. In: BARH, D.; KHAN, M. S.; DAVIES, E. **Omics in plant science**. New Delhi: Springer India, p. 65 – 111, 2015.

- JIMENEZ, H. J. **Análise Molecular in silico e Palinológica de espécies de Amaryllidaceae J. ST. – HIL.** 2019. Tese (Doutorado em Agronomia) – Universidade Federal Rural de Pernambuco. Recife, 2019, 111p. Disponível em: <<http://www.tede2.ufrpe.br:8080/tede/handle/tede2/8381>>. Acesso: 04 de abr. de 2021.
- KALENDAR, R.; LEE, D.; SCHULMAN, A. H. FastPCR Software for PCR, In Silico PCR, and Oligonucleotide Assembly and Analysis. In: Valla S., Lale R. (eds) DNA Cloning and Assembly Methods. Methods in Molecular Biology (Methods and Protocols). **Humana Press**, Totowa, v. 1116, p. 271 – 302, 2014.
- KAUR, G.; ASTHIR, B. Proline: a key player in plant abiotic stress tolerance. **Biologia Plantarum**, v. 59, n. 4, p. 609 – 619, 2015.
- KIBBE, W. A. OligoCalc: an online oligonucleotide properties calculator. **Nucleic Acids Research**. v. 35, p. 43 – 46, 2007.
- LACERDA, J. DE J. **Estresse salino e seus efeitos no crescimento inicial de clones de *Eucalyptus spp.*** 2016. Dissertação (Mestrado em Agronomia) – Universidade Estadual do Sudoeste da Bahia, Vitória da Conquista, 2016, 121p. Disponível em: <<http://www2.uesb.br/ppg/ppgagronomia/wp-content/uploads/2019/02/DISSERTACAO-JANDERSON-versao-definitiva.pdf>>. Acesso em: 18 de abr. de 2021.
- LANGE, A. **Eficiência do *Bacillus amyloliquefaciens* na promoção de crescimento e produtividade no milho.** 2018. Monografia (Bacharel em Agronomia) – Universidade Federal de Mato Grosso, Mato Grosso, 2018, 35p. Disponível em: <<https://bdm.ufmt.br/handle/1/919>>. Acesso em: 02 de mai. de 2021.
- LASKOWSKI, R. A.; WATSON, J. D.; THORNTON, J. M. ProFunc: a server for predicting protein function from 3D structure. **Nucleic Acids Res.**, v. 33, n. 2, p; 89 – 93, 2005.
- Levantamento Sistemático da Produção Agrícola – Tabela 1618: Área plantada, área colhida e produção, por ano da safra e produto das lavouras. **IBGE**, 2021. Disponível em: <<https://sidra.ibge.gov.br/tabela/1618#resultado>>. Acesso em: 06 de abr. de 2021.
- LIMA, J. E. F. W. Situação e perspectivas sobre as águas do cerrado. **Ciência Cult.**, São Paulo, v. 63, n. 3, p. 27-29, 2011.
- LIMA, N. DA S. **Cultivo de pimentão em sistema hidropônico sob estresse salino.** 2016. Tese (Doutorado em Engenharia Agrícola) – Universidade Federal Rural de Pernambuco, Recife, 2016, 90p. Disponível em: <[http://www.pgea.ufrpe.br/sites/ww3.pgea.ufrpe.br/files/documentos/nadielanimas\\_tese.pdf](http://www.pgea.ufrpe.br/sites/ww3.pgea.ufrpe.br/files/documentos/nadielanimas_tese.pdf)>. Acesso em 18 de abr. de 2021.
- MAGALHÃES, J. R.; et al. Origem do óxido nítrico em plantas e seu papel como sinalizador de estresses. In: FERNANDES, M. S. (Ed.). **Nutrição Mineral de Plantas**. 22 ed. Viçosa: Sociedade Brasileira de Ciência do Solo, 2006, p. 201 – 214.
- MALONE, G.; et al. Prospecção de genes em bibliotecas de cDna. **Revista Brasileira de Agrociência**, Pelotas, v. 12, n. 1, p. 7 – 13, mar. 2006.
- MANTOVANI, A.; et al. **Origem e evolução de plantas cultivadas**. 1 ed. Brasília: Embrapa, 2008, 914p.
- MARQUES, M. R. F. **Bioquímica**. 1 ed. Florianópolis: Universidade Federal de Santa Catarina, 2014, 182p.

MARTINS, A. C. P.; et al. A evolução da bioinformática e sua contribuição para a saúde. In: Fórum de Ensino, Pesquisa, Extensão e Gestão, IX, 2015, Montes Claros. **Anais [...]**. Montes Claros: Universidade Estadual de Montes Claros – Unimontes, 2015.

MATTOS, L. C.; et al. A saga do algodão no semiárido nordestino: histórico, declínio e as perspectivas de base agroecológica. **Revista Desenvolvimento e Meio Ambiente**, UFPR, v. 55, p. 556-580, dez. 2020.

MORAES, C. S.; et al. **Série em biologia celular e molecular: métodos experimentais no estudo de proteínas**. Rio de Janeiro: Fiocruz, 2013. 84p.

MORRIS, A. L.; et al. Stereochemical quality of protein structure coordinates. **Proteins: Structure, Function, and Bioinformatics**, v. 12, p. 345-364, 1992.

MOTA, A. F. **Cultivo do algodoeiro irrigado com águas salinas de acordo com a fase de desenvolvimento da cultura**. 2014. Dissertação (Mestrado em Manejo de Solo e Água) – Universidade Rural do Semi-Árido, Mossoró, 2014, 64p. Disponível em: <<https://ppgmsa.ufersa.edu.br/wp-content/uploads/sites/59/2014/10/Disserta%C3%A7%C3%A3o-ANDYGLEY-FERNANDES-MOTA.pdf>>. Acesso em 07 de abr. de 2021.

MOTA, A. F. **Cultivo do algodoeiro irrigado com águas salinas de acordo com a fase de desenvolvimento da cultura**. 2014. Dissertação (Mestrado em Manejo de Solo e Água) – Universidade Federal Rural do Semi-Árido, Mossoró, 2014, 64p. Disponível em: <<https://ppgmsa.ufersa.edu.br/wp-content/uploads/sites/59/2014/10/Disserta%C3%A7%C3%A3o-ANDYGLEY-FERNANDES-MOTA.pdf>>. Acesso em: 15 de abr. de 2021.

MUNNS, R.; et al. Avenues for increasing salt tolerance of crops, and the role of physiologically based selection traits. **Plant and Soil**, v. 247, p. 93 – 105, 2002.

NASCIMENTO, A. G. G. **Genótipos e terços de desenvolvimento de capulhos: afetam a qualidade da fibra do algodoeiro (*Gossypium hirsutum* L.)**. 2019. Monografia (Bacharel em Engenharia Agrônômica) – Universidade Federal de Uberlândia, Uberlândia, 2019, 32p. Disponível em: <<https://repositorio.ufu.br/bitstream/123456789/24029/3/Gen%c3%b3tiposTer%c3%a7osDesenvolvimento.pdf>>. Acesso em: 23 de abr. de 2021.

NEVES, M. F.; PINTO, M. J. A. **The brazilian cotton chain: challenges and strategies**. São Paulo: ABRAPA, 2011, 55p.

NEVES, M. F.; PINTO, M. J. A.; MOURA, A. DE A. Tendências e agenda estratégica da cadeia do algodão. **ABRAPA**, 2017. Disponível em: <<https://www.abrapa.com.br/Paginas/NoticiaAbrapa.aspx?noticia=308>>. Acesso em: 29 de abr. de 2021.

NOGUEIRA, R. J. M. C.; et al. Comportamento estomático e potencial da água da folha em amendoim cv. BRS 151-L7 submetido a estresse hídrico. **Revista Brasileira de Oleaginosas e Fibrosas**, v. 10, p. 985 – 991, 2006.

O que é um Z-score?. **DATA SCIENCE**, 2020. Disponível em: <<https://datascience.eu/pt/matematica-e-estatistica/o-que-e-um-z-score/>>. Acesso em: 24 de abr. de 2021.

OLIVEIRA, E. M. M.; DE OLIVEIRA, T. C.; DE LIMA, I. S.; DOS SANTOS, T. F. **Documentos 114: Utilização de ferramentas de bioinformática na construção de primers para detecção de sequências específicas de DNA**. Rio de Janeiro: Embrapa, 2011, 20p.

OLIVEIRA, E. M. M.; et al., Desenho de primers degenerados através da bioinformática – Comunicado Técnico 208. **Embrapa**, Rio de Janeiro, mai. de 2021.

Página de ajuda do ProSA-web. **PROSA-WEB**, 2021. Disponível em: <[https://prosa.services.came.sbg.ac.at/prosa\\_help.html](https://prosa.services.came.sbg.ac.at/prosa_help.html)>. Acesso em: 24 de abr. de 2021.

PARIDA, A. K.; DAS, A. B. Salt tolerance and salinity effects on plants: a review. **Ecotoxicology and Environmental Safety**, v. 60, n. 3, p. 324 – 349, 2005.

PEREIRA, A. T. A. Qualidade via água. Cultivar - Grandes Culturas. **Revista Globo Rural**, Pelotas, n. 103, p. 14 – 16, 2008.

PEREIRA, E. R. DE L. **Tolerância de genótipos do algodão colorido ao estresse salino**. 2012. Dissertação (Mestrado em Ciências Agrárias) – Universidade Estadual da Paraíba, Campina Grande, 2012, 105p. Disponível em: <[http://pos-graduacao.uepb.edu.br/ppgca/download/outros\\_documentos/DISSERTA%C3%87%C3%83O%20-%20ELIZANDRA%20RIBEIRO%20DE%20LIMA%20PEREIRA%20-%202012.pdf](http://pos-graduacao.uepb.edu.br/ppgca/download/outros_documentos/DISSERTA%C3%87%C3%83O%20-%20ELIZANDRA%20RIBEIRO%20DE%20LIMA%20PEREIRA%20-%202012.pdf)>. Acesso em 05 de abr. de 2021.

PEREIRA, G. G. **Busca de um tamanho ótimo de gene e proteína para maximização da qualidade da filogenia resultante**. 2010. Dissertação (Mestrado em Genética e Biologia Molecular) – Universidade estadual de Campinas, Campinas, 2010, 87p. Disponível em: <[http://repositorio.unicamp.br/bitstream/REPOSIP/317448/1/Sampaio\\_ClaudioLuisMarques\\_M.pdf](http://repositorio.unicamp.br/bitstream/REPOSIP/317448/1/Sampaio_ClaudioLuisMarques_M.pdf)>. Acesso em: 06 de mai. de 2021.

PEREIRA, J. W. DE L.; et al. Mudanças bioquímicas em genótipos de amendoim submetidos a déficit hídrico moderado. **Revista Ciência Agrônômica**, Ceará, v.43, n. 4, p. 766 – 773, 2012 (a).

PETRUCCI, F. **Algodão transgênico na agricultura: benefícios e controvérsias**. 2011. Monografia (Bacharel em Ciências Econômicas) – Universidade Estadual de Campinas, Campinas, 2011, 46p. Disponível em: <<http://www.bibliotecadigital.unicamp.br/document/?down=000852580>>. Acesso em: 25 de abr. de 2021.

PORTO, E. R.; et al. Agricultura bioassalina: desafios e alternativas para o uso de águas salobras e salinas no semiárido brasileiro – Documentos 121. 1ed. Jaguariúna: **Embrapa Meio Ambiente**, out. de 2019, 39p.

QUEIROZ, J. A. DA S.; et al. Desenho e Validação de Primers In Silico para Detecção do Vírus Sincicial Respiratório Humano. **Revista Fimca**, v. 4, n.1, p. 17 – 30, dez. de 2017.

REINHARDT, D. H.; ROST, T. L. Developmental changes of cotton root primary tissues induced by salinity. **International Journal of Plant Sciences**, v. 156, n. 4, p. 505 – 513, 1995.

RIBEIRO, A. C.; GUIMARÃES, P. T. G.; ALVAREZ, V. V. H. (Ed.) Recomendações para o uso de corretivos e fertilizantes em Minas Gerais. 5ª aproximação. Viçosa: **Comissão de Fertilidade do Solo do Estado de Minas Gerais**, 359 p. 1999.

RIBEIRO, M. R. Origem e classificação dos solos afetados por sais. In: GHEYI, H. R.; DIAS, N. DA S.; LACERDA, C. F. de (ed). **Manejo da salinidade na agricultura: Estudos básicos e aplicados**. Fortaleza, INCTSal, p. 11-19, 2010.

RIBEIRO, M. R.; et al. Química dos solos salinos e sódicos. In: MELO, V. F.; ALLEONI, L. R. F. (ed.). **Química e mineralogia do solo. Parte II – Aplicações**. 1 ed. Viçosa: Sociedade Brasileira de Ciência do Solo, 2009, p. 449 – 484.

RIBEIRO, M. R.; FREIRE, F. J.; MONTENEGRO, A. A. Solos halomórficos no Brasil: Ocorrência, gênese, classificação, uso e manejo sustentável. In: CURI, N.; MARQUES, J. J.; GUILHERME, L. R. G. G.; LIMA, J. M.; LOPES, A. S.; ALVAREZ V. V. H. (ed.). **Tópicos em ciência do solo**. v. 3. Viçosa: Sociedade Brasileira de Ciência do Solo, p.165 - 208, 2003.

ROACHE, M. C.; BAILEY, P. C.; BOON, P. I. Effects of salinity on the decay of the freshwater macrophyte, *Triglochin procerum*. **Aquatic Botany**, v. 84, n. 1, p. 45 – 52, 2006.

ROY, A.; YANG, J. & ZHANG, Y. COFACTOR: an accurate comparative algorithm for structurebased protein function annotation. **Nucl. Acids Res.**, v. 40, n. W1, p. 471 – 477, 2012.

SAKAMOTO, T. **Ferramentas para análise filogenética e de distribuição taxonômica de genes ortólogos**. 2016. Tese (Doutorado em Bioinformática) Universidade Federal de Minas Gerais, Belo Horizonte, 2016, 114p. Disponível em: <<http://www.pgbioinfo.icb.ufmg.br/defesas/106D.PDF>>. Acesso em: 8 de mai. de 2021.

SALDANHA, C. I.; BELO, M. A. DE A. Produção Animal Universidade Brasil – Fossipol: Toxicidade dos produtos do algodão. **Universidade Brasil - Boletim técnico**, Descalvado - SP, 25p., dez. 2016.

SANDERSON, M. J. Objections to bootstrapping phylogenies: a critique. **Systematic Biology**, v. 44, n. 3, p. 299 – 320, 1995.

SANTOS, K. C. F.; et al. Atividade biológica em solo salino sódico saturado por água sob cultivo de *Atriplex nummularia*. **Revista Ciência Agronômica**, v.42, n.3, p.619 – 627, 2011.

SANTOS, L. A. **Uso de ferramentas de bioinformática para estudos de epidemiologia molecular, filogeografia e filodinâmica viral**. 2010. Dissertação (Mestrado em Biotecnologia em Saúde e Medicina Investigativa) – Fundação Oswaldo Cruz - Fiocruz, Salvador, 2010, 140p. Disponível em: <<https://www.arca.fiocruz.br/bitstream/icict/4237/1/Luciane%20Amorim%20Santos%20Uso%20de%20ferramentas%20de%20bioinform%C3%A1tica....pdf>>. Acesso em: 16 de abr. de 2021.

SCHOSSLER, T. R.; et al. Salinidade: Efeitos na fisiologia e na nutrição mineral de plantas. **Enciclopédia Biosfera**, Goiânia, v. 8, n. 15, p. 1563 – 1578, 2012.

SCVORTZOFF, M. V. **Relações filogenéticas e evolução cromossômica em espécies do gênero *Oxalis* (oxalidaceae)**. 2012. Tese (Doutorado em Biologia Vegetal) – Universidade Federal de Pernambuco, 2012, 164p. Disponível em: <<https://repositorio.ufpe.br/bitstream/123456789/11803/1/Tese%20de%20doutorado%20M%20Vaio.pdf>>. Acesso: 07 de mai. de 2021.

SERTÃO, M. A. J. **Uso de corretivos e cultivo do capim *Urocloa* (*Urocloa mosambicensis* (Hack.) Dandy) em solos degradados do Semi-árido**. 2005. Dissertação (Mestrado em Zootecnia) – Universidade Federal de Campina Grande, Patos, 2005, 75p. Disponível em: <[http://www.cstroid.sti.ufcg.edu.br/ppgz/dissertacoes/dissertacao\\_2005/n\\_04\\_maria\\_auxiliadora\\_justino\\_sertao/n\\_04\\_maria\\_auxiliadora\\_justino\\_sertao.pdf](http://www.cstroid.sti.ufcg.edu.br/ppgz/dissertacoes/dissertacao_2005/n_04_maria_auxiliadora_justino_sertao/n_04_maria_auxiliadora_justino_sertao.pdf)>. Acesso em: 25 de abr. de 2021.



SEVERINO, L. S.; et al. Produto: Algodão – Parte 01: Caracterização e desafios tecnológicos. **Embrapa Algodão**, Campina Grande, p. 1 – 29, 2019.

SILVA, F. A. M.; et al. Efeito do estresse salino sobre a nutrição mineral e o crescimento de mudas de Aroeira (*Myracrodruon urundeuva*) cultivadas em solução nutritiva. **Cerne**, v. 6, n. 1, p. 52 – 59, 2000.

SILVEIRA, J. A. G.; et al. Mecanismos biomoleculares envolvidos com a resistências ao estresse salino em plantas. In: GHEYI, H. R.; et al. Manejo da salinidade na agricultura: estudos básicos e aplicados. 2ed. **INCTSal**, Fortaleza, p. 181 – 197, 2016.

SOARES FILHO, W. DOS S.; et al. Melhoramento genético e seleção de cultivares tolerantes à salinidade. In: Manejo da salinidade na agricultura: estudos básicos e aplicados, GHEYI, H. R.; DIAS, N DA S., LACERDA, C. F. DE; GOMES FILHO, E. **INCTSal**, Fortaleza, p.259-274, 2016.

SOARES, T. M. **Utilização de águas salobras no cultivo da alface em sistema hidropônico NFT como alternativa agrícola condizente ao semi-árido brasileiro**. 2007. Tese (Doutorado em Agronomia) – Universidade de São Paulo, Piracicaba, 2007, 267p. Disponível em: <<https://www.teses.usp.br/teses/disponiveis/11/11143/tde-29112007-093534/publico/TeseTalesSoares.pdf>>. Acesso em: 20 de abr. de 2021.

SOUSA JUNIOR, S. P.; FERNANDES, P. D.; GHEYI H. R. Uso de água salina no crescimento do algodoeiro colorido BRS Verde sob adubação nitrogenada. **Engenharia Ambiental**, Espírito Santo do Pinhal, v. 5, n. 1, p. 28 – 46, 2008.

SUASSUNA, J.; AUDRY, P. Qualidade da água na irrigação do trópico semi-árido: um estudo de caso. **Fundação Joaquim Nabuco**, 2019. Disponível em: <<https://www.fundaj.gov.br/index.php/artigos-joao-suassuna/9242-a-qualidade-da-agua-na-irrigacao-do-tropico-semi-arido-um-estudo-de-caso>>. Acesso em: 02 de mai. de 2021.

SUASSUNA, N. D.; SILVA, J. C.; BETTIOL, W. Uso do Trichoderma na cultura do algodão. In: MEYER, M. C.; MAZARO, S. M.; DA SILVA, J. C. (Ed.). **Trichoderma: uso na agricultura**. Brasília: Embrapa, 2019. p. 361 – 379.

SZABADOS, L.; SAVOURÉ, A. Proline: a multifunctional amino acid. **Trends in Plant Science**, v. 15, n. 2, p. 89 – 97, 2009.

TURCHETTO-ZOLET, A. C.; et al. **Marcadores moleculares na Era Genômica: Metodologias e aplicações**. 1ed. Ribeirão Preto: Sociedade Brasileira de Genética – SGB, 2017, 181p.

VASSILEV, D. et al. Application of bioinformatics in plant breeding. **Biotechnology and Biotechnological Equipment**, v. 19, p. 139 – 152, 2005.

VERLI, H. **Bioinformática da biologia à flexibilidade molecular**. 1 ed. São Paulo: Sociedade Brasileira de Bioquímica e Biologia Molecular - SBBq, 2014, 282p.

VIALA, V. L. **Análise combinada do transcriptoma de glândula de veneno e do proteoma do veneno da espécie Pseudonaja textilis (Elapidae: Serpentes)**. 2014. Tese (Doutorado em Ciências na Área de Tecnologia Nuclear – Aplicações) – Instituto de Pesquisas Energéticas e Nucleares, Universidade de São Paulo, São Paulo, 2014, 173p. Disponível em: <[http://pelicano.ipen.br/PosG30/TextoCompleto/Vincent%20Louis%20Viala\\_D.pdf](http://pelicano.ipen.br/PosG30/TextoCompleto/Vincent%20Louis%20Viala_D.pdf)>. Acesso em: 28 de abr. de 2021.

VIDAL NETO, F. DAS C.; FREIRE, E. C. Melhoramento genético do algodoeiro. In: VIDAL NETO, F. das C.; CAVALCANTI, J. J. V. (Ed.). **Melhoramento genético de plantas no Nordeste**. Brasília: Embrapa, p. 49 - 83, 2013.

VILHENA, A. P. Brimer: **Um sistema web para gerenciamento de primers para o laboratório de tecnologia biomolecular da UFPA**. 2015. Monografia (Bacharel em Sistemas de Informação) – Universidade Federal do Sul e Sudeste do Pará, Marabá, 2015, 70p. Disponível em: <<https://faceel.unifesspa.edu.br/images/works/TCC/2015/BRIMER-UM-SISTEMA-WEB-PARA-GERENCIAMENTO-DE-PRIMERS-PARA-O-LABORATORIO-DE-TECNOLOGIA-BIOMOLOCLAR-DA-UFPA.pdf>>. Acesso em: 28 de abr. de 2021.

VITAL, S. A.; et al. Opposing roles for superoxide and nitric oxide in the NaCl stress-induced upregulation of antioxidant enzyme activity in cotton callus tissue. **Environmental and Experimental Botany**, v. 62, n. 1, p. 60 – 68, 2008.

WANG, Z.; et al. Caracterização de todo o genoma de fatores de transcrição bZIP e seus padrões de expressão em resposta à seca e estresse salino em *Jatropha curcas*. **International Journal of Biological Macromolecules**, mai. de 2021.

WENDEL, J. F.; CRONN, R. C. Polyploidy and the evolutionary history of cotton. **Advances in Agronomy**, v. 78, p. 139 – 186, 2003.

YAN, Q.; et al. *GmCYP82A3*, um gene da família do citocromo P450 da soja envolvido na via de sinalização de ácido jasmônico e etileno, aumenta a resistência das plantas a estresses bióticos e abióticos. **Journals Plos One**, set. de 2016.

YANCEY, P. H. Organic osmolytes compatible metabolic and counteracting cytoprotectants in high osmolarity and other stress. **The Journal of Experimental Biology**, v. 208, n. 15, p. 2819 – 2830, 2005.

YE, J.; et al. Primer-BLAST: Uma ferramenta para projetar primers específicos para a reação em cadeia da polimerase. **BMC Bioinformática**, v. 13, n. 134, p. 1 – 11, 2012.

YUAN, Z.; et al. Specialized microbiome of a halophyte and its role in helping non-host plants to withstand salinity. **Scientific Reports**, v. 6, n. 32467, p. 1 – 13, 2016.

ZANCAN, W. L. A.; CHITARRA, L. G.; CHITARRA, G. S. Fungos associados à podridão de maçãs do algodoeiro na região de Primavera do Leste – MT. **Agricultural Sciences**, Uberlândia, v. 27, n. 4, p. 518 – 525, 2011.

ZHAO, L., et al. Nitric oxide functions as a signal in salt resistance in the calluses from two ecotypes of reed. **Plant Physiol**, v. 134, n. 2, p. 849 – 857, 2004.

ZHENG, Z.; et al. A pequena GTPase ROP10 associada à membrana plasmática é um regulador negativo específico das respostas ao ácido abscísico em *Arabidopsis*. **The Plant Cell**, v. 14, 11 ed., p. 2787 – 2797, nov. de 2002.

ZHU, M.; et al. Evaluating Predictive Values of Various Physiological Indices for Salinity Stress Tolerance in Wheat. **Journal of Agronomy and Crop Science**, v. 202, n. 2, p. 115 – 124, 2016.

## ANEXO

O alinhamento múltiplo comparativo entre as 16 sequências usadas no alinhamento de cada uma das três proteínas, do domínio bZIP, CP450 e Rop\_like, totalizando 48 em análise pelo servidor Clustal Omega, e os gráficos de conservação de domínios da sequência completa dos genes.

**Figura 19** - Alinhamento múltiplo comparativo entre as três proteínas, obtido no servidor Clustal Omega.

```

Populus_simonii|AKV56346.1          -----ma          2
G_arboreum|XP_017611064.1          md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_australe|KAA3456360.1            md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_gossypioides|MBA0735964.1        md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_schwendimani|MBA0825516.1        md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_armourianum|MBA0825516.1        md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_trilobum|MBA0762230.1            md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_davidsonii|MBA0609728.1          md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_barbadense|PPD080758.1           md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_mustelinum|TYI91474.1            md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_hirsutum|NP_001314347.1          md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_barbadense|KAB2039250.1          md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_tomentosum|TYH81593.1           md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_raimondii|XP_012472188.1         md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_hirsutum|KAG4156417.1           md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
G_darwinii|TYG77701.1              md-----gv-----ss----sgnsqndastsggesqtt--aaqqtapkpvpv 36
Glycine_max|NP_001235844.2         -----           0
G_raimondii|KJB53185.1             -----           0
G_aridum|MBA0690107.1              -----           0
G_raimondii|ADY68835.1             -----           0
Durio_zibethinus|XP_022736748.1    -----           0
G_hirsutum|NP_001314362.1          -----           0
G_raimondii|XP_012440446.1         -----           0
G_hirsutum|ADY68837.1              -----           0
G_barbadense|ADY68833.1            -----           0
G_hirsutum|ADY68836.1              -----           0
G_barbadense|PPR86822.1            -----           0
G_hirsutum|XP_016728000.1          -----           0
Corchorus_olitorius|OM078421.1    -----           0
Durio_zibethinus|XP_022735746.1    -----           0
Theobroma_cacao|XP_007036140.1     -----           0
Herrania_umbriatica|XP_021292167.1 -----           0
Prunus_armeniaca|CAB4296392.1      -----mdyslpyvnsitaflftiivsyysf---rwrwpaklapteakgawpifghlp--1 51
G_davidsonii|MBA0624583.1          mdlldsifgy---avlgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_arboreum|KHG01290.1              mdlldsifgy---avlgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_schwendimani|MBA0866527.1        mdlldfstfgy---avlgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_schwendimani|MBA0866525.1        mdlldfstfgy---avlgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_trilobum|MBA0775238.1            mdlldfstfgy---avlgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_australe|KAA3488855.1            mdpldstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_australe|KAA3488856.1            mdpldstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_arboreum|XP_017643194.1          mdpldstfgy---avlgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_mustelinum|TYJ35072.1            mdpldstfgy---avfgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_barbadense|KAB2082570.1          mdpldstfgy---avfgitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_hirsutum|NP_001314395.1          mdplydstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_tomentosum|TYH71926.1           mdpldstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_mustelinum|TYI82308.1            mdpldstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_davidsonii|MBA0623572.1          mdpldstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53
G_davidsonii|MBA0623571.1          mdpldstfgy---avlcitl-lflytklkrsssgsggkappvaagawpiighlp--1 53

```

```

Populus_simonii|AKV56346.1      sqsnappfflgktydmvddpatdaviswsstnnsfvwnpppefardllpkyfkhnfss- 61
G_arboreum|XP_017611064.1      qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_australe|KAA3456360.1       qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_gossypoides|MBA0735964.1     qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_schwendimani|MBA0851603.1   qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_armourianum|MBA0825516.1     qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_trilobum|MBA0762230.1       qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_davidsonii|MBA0609728.1     qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_barbadense|PPD80758.1       qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_mustelinum|TYI91474.1       qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_hirsutum|NP_001314347.1     qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_barbadense|KAB2039250.1     qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_tomentosum|TYH81593.1       qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_raimondii|XP_012472188.1    qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_hirsutum|KAG4156417.1       qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
G_darwinii|TYG77701.1        qsvntpppflsktydmvddpatdaviswspsnnsfvwnpppefardllpkyfkhnfss- 95
Glycine_max|NP_001235844.2    -----maaataprfik-----vtvgdgvgtcmlicytsnkfptd 37
G_raimondii|KJ353185.1       -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_aridum|MBA0690107.1        -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_raimondii|ADY68835.1       -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
Durio_zibethinus|XP_022736748.1 -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_hirsutum|NP_001314362.1     -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_raimondii|XP_012440446.1    -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_hirsutum|ADY68837.1        -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_barbadense|ADY68833.1       -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_hirsutum|ADY68836.1        -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_barbadense|PPR86822.1      -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_hirsutum|XP_016728080.1     -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
Corchorus_olitorius|OM078421.1 -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
Durio_zibethinus|XP_022735746.1 -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
Theobroma_cacao|XP_007036140.1 -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
Hernania_umberratica|XP_021292167.1 -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
Prunus_armeniaca|CAB4296392.1 -----massasrfik-----vtvgdgvgtcmlicytsnkfptd 36
G_davidsonii|MBA0624583.1     lggstpph--iklaamadkypolftirl--gvypslvissselake-----cftndlals 103
G_arboreum|KHG01290.1        lggpkaph--eslgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_schwendimani|MBA0866527.1   lggpkaph--eslgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_trilobum|MBA0775238.1       lggpktph--etlgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_australe|KAA3488855.1       lggpktph--etlgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_australe|KAA3488856.1       lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_arboreum|XP_017643194.1     lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_mustelinum|TYJ35072.1      lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_barbadense|KAB2082570.1     lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_hirsutum|NP_001314395.1     lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_tomentosum|TYH71926.1      lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_mustelinum|TYI82308.1      lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_davidsonii|MBA0623572.1     lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
G_davidsonii|MBA0623571.1     lggpkaph--ealgldekygpaymiri--gvhpalvvnssdvake-----iftvndmvs 105
.                               . . . . . : . : * :
Populus_simonii|AKV56346.1      -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--han-- 108
G_arboreum|XP_017611064.1     -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_australe|KAA3456360.1       -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_gossypoides|MBA0735964.1     -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_schwendimani|MBA0851603.1   -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_armourianum|MBA0825516.1     -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_trilobum|MBA0762230.1       -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_davidsonii|MBA0609728.1     -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_barbadense|PPD80758.1       -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_mustelinum|TYI91474.1       -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_hirsutum|NP_001314347.1     -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_barbadense|KAB2039250.1     -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_tomentosum|TYH81593.1       -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_raimondii|XP_012472188.1    -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_hirsutum|KAG4156417.1       -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
G_darwinii|TYG77701.1        -----fvrlntygfvkvdppdrwefanegflrgqkhlrrnissrpkpah--hgh-- 142
Glycine_max|NP_001235844.2    yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 89
G_raimondii|KJ353185.1       yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_aridum|MBA0690107.1        yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 89
G_raimondii|ADY68835.1       yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
Durio_zibethinus|XP_022736748.1 yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_hirsutum|NP_001314362.1     yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_raimondii|XP_012440446.1    yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_hirsutum|ADY68837.1        yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_barbadense|ADY68833.1       yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_hirsutum|ADY68836.1        yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_barbadense|PPR86822.1      yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
G_hirsutum|XP_016728080.1     yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
Corchorus_olitorius|OM078421.1 yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
Durio_zibethinus|XP_022735746.1 yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
Theobroma_cacao|XP_007036140.1 yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
Hernania_umberratica|XP_021292167.1 yipt--vfdnfsanvvvgei--tnlg--lwdtagq----edynrlrplsyrgadvflafsl 88
Prunus_armeniaca|CAB4296392.1 srpklaavdhlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_davidsonii|MBA0624583.1     srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_arboreum|KHG01290.1        srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_schwendimani|MBA0866527.1   srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_trilobum|MBA0775238.1       srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_australe|KAA3488855.1       srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_australe|KAA3488856.1       srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_arboreum|XP_017643194.1     srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_mustelinum|TYJ35072.1      srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_barbadense|KAB2082570.1     srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_hirsutum|NP_001314395.1     srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_tomentosum|TYH71926.1      srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_mustelinum|TYI82308.1      srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_davidsonii|MBA0623572.1     srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
G_davidsonii|MBA0623571.1     srsefaaaehlgnyamgfspsyga--fwremnk-----itmlevlsnrhdqkkvfv 158
.                               . : . : * :

```

Populus\_simonii|AKV56346.1 -----qqqqphganssvaacvevgkfgleeeverlkrdknvlmqelvr1rqq 156  
 G\_arboreum|XP\_017611064.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_australe|KAA3456360.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_gossypoides|MBA0735964.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_schwendimani|MBA0851603.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_armourianum|MBA0825516.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_trilobum|MBA0762230.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_davidsonii|MBA0609728.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_barbadense|PPD80758.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_mustelinum|TYI91474.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_hirsutum|NP\_001314347.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_barbadense|KAB2039250.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_tomentosum|TYH81593.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_raimondii|XP\_012472188.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_hirsutum|KAG4156417.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 G\_darwinii|TYG77701.1 -----qatqashgasssvdacevevgkfgleeeverlkrdknvlmqelvr1rqq 190  
 Glycine\_max|NP\_001235844.2 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 123  
 G\_raimondii|KJ853185.1 vsrasyenvlkk-----hyagpvvvlvgt-lid-----1----- 116  
 G\_aridum|MBA0690107.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 124  
 G\_raimondii|ADY68835.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 Durio\_zibethinus|XP\_022736748.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_hirsutum|NP\_001314362.1 vsrashenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_raimondii|XP\_012440446.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_hirsutum|ADY68837.1 vsrashenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_barbadense|ADY68833.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_hirsutum|ADY68836.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_barbadense|PPR86822.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 G\_hirsutum|XP\_016728080.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 Corchorus\_olitorius|OM078421.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 Durio\_zibethinus|XP\_022735746.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 Theobroma\_cacao|XP\_007036140.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 Herrania\_umbatica|XP\_021292167.1 vsrasyenvlkkwipdqhfyagpvv-lvgtklid-----1----- 122  
 Prunus\_armeniaca|CAB4296392.1 evttflqelyktwstaeikresnndgvlvelkqw-----fgdrtlnvilmvagkry 209  
 G\_davidsonii|MBA0624583.1 eiegsmklllyntwadkk-----ngsgkvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_arboreum|KHG01290.1 eiegsmklllyntwadkk-----ngsgkvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_schwendimani|MBA0866527.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_schwendimani|MBA0866525.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_trilobum|MBA0775238.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_australe|KAA3488855.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_australe|KAA3488856.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_arboreum|XP\_017643194.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_mustelinum|TYJ35072.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_barbadense|KAB2082570.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_hirsutum|NP\_001314395.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_tomentosum|TYH71926.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_mustelinum|TYI82308.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_davidsonii|MBA0623572.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207  
 G\_davidsonii|MBA0623571.1 eiegsmklllyntwadkk-----dsgskvlvemkhh-----fsdtlnvimrtvagkry 207

Populus\_simonii|AKV56346.1 qatdsqatqmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----d----- 204  
 G\_arboreum|XP\_017611064.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_australe|KAA3456360.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_gossypoides|MBA0735964.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_schwendimani|MBA0851603.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_armourianum|MBA0825516.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_trilobum|MBA0762230.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_davidsonii|MBA0609728.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_barbadense|PPD80758.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_mustelinum|TYI91474.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_hirsutum|NP\_001314347.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_barbadense|KAB2039250.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_tomentosum|TYH81593.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_raimondii|XP\_012472188.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_hirsutum|KAG4156417.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 G\_darwinii|TYG77701.1 qatdnqlqvmvqrlqgmeqrqqqmsflakamspgflaqfvqqn-----e----- 238  
 Glycine\_max|NP\_001235844.2 ----- 123  
 G\_raimondii|KJ853185.1 ----- 116  
 G\_aridum|MBA0690107.1 ----- 124  
 G\_raimondii|ADY68835.1 ----- 122  
 Durio\_zibethinus|XP\_022736748.1 ----- 122  
 G\_hirsutum|NP\_001314362.1 ----- 122  
 G\_raimondii|XP\_012440446.1 ----- 122  
 G\_hirsutum|ADY68837.1 ----- 122  
 G\_barbadense|ADY68833.1 ----- 122  
 G\_hirsutum|ADY68836.1 ----- 122  
 G\_barbadense|PPR86822.1 ----- 122  
 G\_hirsutum|XP\_016728080.1 ----- 122  
 Corchorus\_olitorius|OM078421.1 ----- 122  
 Durio\_zibethinus|XP\_022735746.1 ----- 122  
 Theobroma\_cacao|XP\_007036140.1 ----- 122  
 Herrania\_umbatica|XP\_021292167.1 ----- 122  
 Prunus\_armeniaca|CAB4296392.1 faadedek-tearrvqsalk-----effyyvlgfvvdavpflrwdlgyekmkkitak 263  
 G\_davidsonii|MBA0624583.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_arboreum|KHG01290.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_schwendimani|MBA0866527.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_schwendimani|MBA0866525.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_trilobum|MBA0775238.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_australe|KAA3488855.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_australe|KAA3488856.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_arboreum|XP\_017643194.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_mustelinum|TYJ35072.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_barbadense|KAB2082570.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_hirsutum|NP\_001314395.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_tomentosum|TYH71926.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_mustelinum|TYI82308.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_davidsonii|MBA0623572.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260  
 G\_davidsonii|MBA0623571.1 vgaeedqk-----evlryrkalr-----dffhltgmfvlgdavpflrwdlgyekmkkitak 260



```

Populus_simonii|AKV56346.1      ----nrrriteankkrrlkqedvpenedsgshdgaivkyqplmneaqaqlrqiakmdas 260
G_arboreum|XP_017611064.1      ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_australe|KAA3456360.1       ----snrrriseankkrrlkqdg---sdtasaasdgaikiyqplne--akemlrqimkgdss 289
G_gossypioides|MBA0735964.1    ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqpsne--akemlrqimkgdss 289
G_schwendimani||MBA0851603.1  ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_armourianum|MBA0825516.1    ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_trilobum|MBA0762230.1       ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_davidsonii|MBA0609728.1     ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_barbadense|PPD80758.1       ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_mustelinum|TYI91474.1       ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_hirsutum|NP_001314347.1     ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_barbadense|KAB2039250.1     ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_tomentosum|TYH81593.1      ----snrrriseankkrrlkqdv---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_raimondii|XP_012472188.1    ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_hirsutum|KAG4156417.1      ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
G_darwinii|TYG77701.1        ----snrrriseankkrrlkqdg---ndtsaasdgaikiyqplne--akemlrqimkgdss 289
Glycine_max|NP_001235844.2    ----- 123
G_raimondii|KJ853185.1       ----- 116
G_aridum|MBA0690107.1        ----- 124
G_raimondii|ADY68835.1       ----- 122
Durio_zibethinus|XP_022736748.1 ----- 122
G_hirsutum|NP_001314362.1    ----- 122
G_raimondii|XP_012440446.1    ----- 122
G_hirsutum|ADY68837.1       ----- 122
G_barbadense|ADY68833.1      ----- 122
G_hirsutum|ADY68836.1       ----- 122
G_barbadense|PPR86822.1     ----- 122
G_hirsutum|XP_016728080.1    ----- 122
Corchorus_olitorius|OM078421.1 ----- 122
Durio_zibethinus|XP_022735746.1 ----- 122
Theobroma_cacao|XP_007036140.1 ----- 122
Herrania_umbatica|XP_021292167.1 ----- 122
Prunus_armeniaca|CAB4296392.1 ----- 303
G_davidsonii|MBA0624583.1    ----- 300
G_arboreum|KHG01290.1       ----- 300
G_schwendimani||MBA0866527.1 ----- 300
G_schwendimani||MBA0866525.1 ----- 300
G_trilobum|MBA0775238.1     ----- 300
G_australe|KAA3488855.1     ----- 300
G_australe|KAA3488856.1     ----- 300
G_arboreum|XP_017643194.1    ----- 300
G_mustelinum|TYJ35072.1     ----- 300
G_barbadense|KAB2082570.1    ----- 300
G_hirsutum|NP_001314395.1    ----- 300
G_tomentosum|TYH71926.1     ----- 300
G_mustelinum|TYI82308.1     ----- 300
G_davidsonii|MBA0623572.1    ----- 300
G_davidsonii|MBA0623571.1    ----- 300

```

```

Populus_simonii|AKV56346.1      sklesydnnlg-----akdsqssshstsvtlqevlptsgqsahipaafgisr 308
G_arboreum|XP_017611064.1      prldsfnknhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_australe|KAA3456360.1       prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_gossypioides|MBA0735964.1    prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_schwendimani||MBA0851603.1  prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_armourianum|MBA0825516.1    prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_trilobum|MBA0762230.1       prldsfnnnhenfmihgdvsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_davidsonii|MBA0609728.1     prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_barbadense|PPD80758.1       pkldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_mustelinum|TYI91474.1       prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_hirsutum|NP_001314347.1     prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_barbadense|KAB2039250.1     pkldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_tomentosum|TYH81593.1      pkldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_raimondii|XP_012472188.1    prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_hirsutum|KAG4156417.1      prldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
G_darwinii|TYG77701.1        pkldsfnnnhenflihgdsassglggnsasrvsgvtlkevppstgscstyvpatswlp 349
Glycine_max|NP_001235844.2    -----redr-----h-----ymad 132
G_raimondii|KJ853185.1       -----redk-----h-----y1ad 125
G_aridum|MBA0690107.1        -----redk-----h-----y1ad 133
G_raimondii|ADY68835.1       -----redk-----h-----y1ad 131
Durio_zibethinus|XP_022736748.1 -----redk-----h-----y1ad 131
G_hirsutum|NP_001314362.1    -----redk-----h-----y1ad 131
G_raimondii|XP_012440446.1    -----redk-----h-----y1ad 131
G_hirsutum|ADY68837.1       -----redk-----h-----y1ad 131
G_barbadense|ADY68833.1      -----redk-----h-----y1ad 131
G_hirsutum|ADY68836.1       -----redk-----h-----y1ad 131
G_barbadense|PPR86822.1     -----redk-----h-----y1ad 131
G_hirsutum|XP_016728080.1    -----redk-----h-----y1ad 131
Corchorus_olitorius|OM078421.1 -----redk-----h-----y1ad 131
Durio_zibethinus|XP_022735746.1 -----redk-----h-----y1ad 131
Theobroma_cacao|XP_007036140.1 -----redk-----h-----y1ad 131
Herrania_umbatica|XP_021292167.1 -----redk-----h-----y1ad 131
Prunus_armeniaca|CAB4296392.1 -----lgefdadtvn-----katslnmiagsdttvtliw1-----s-----l1ln 341
G_davidsonii|MBA0624583.1    -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_arboreum|KHG01290.1       -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_schwendimani||MBA0866527.1 -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_schwendimani||MBA0866525.1 -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_trilobum|MBA0775238.1     -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_australe|KAA3488855.1     -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_australe|KAA3488856.1     -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_arboreum|XP_017643194.1    -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_mustelinum|TYJ35072.1     -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_barbadense|KAB2082570.1    -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_hirsutum|NP_001314395.1    -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_tomentosum|TYH71926.1     -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_mustelinum|TYI82308.1     -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_davidsonii|MBA0623572.1    -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338
G_davidsonii|MBA0623571.1    -----lagydadtin-----katslnmilagsdttvtliw1-----s-----lmln 338

```

Populus_simonii AKV56346.1	qggsapiseiqsslqiassekvtsaqfp---dismlvgaqgatsipqadvimphvsqm	365
G_arboreum XP_017611064.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_australe KAA3456360.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_gossypiooides MBA0735964.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_schwendimani MBA0851603.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_armourianum MBA0825516.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_triobum MBA0762230.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_davidsonii MBA0609728.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_barbadense PPD80758.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_mustelinum TYI91474.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_hirsutum NP_001314347.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_barbadense KAB2039250.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_tomentosum TYH81593.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_raimondii XP_012472188.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_hirsutum KAG4156417.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
G_darwinii TYG77701.1	hpsaaaiseiqpspcttsseeittrfpp---dmsalvgaqkpaamsipqtdismpeasem	406
Glycine_max NP_001235844.2	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	173
G_raimondii KJBS3185.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	167
G_aridum MBA0690107.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	175
G_raimondii ADY68835.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
Durio_zibethinus XP_022736748.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_hirsutum NP_001314362.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_raimondii XP_012440446.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_hirsutum ADY68837.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_barbadense ADY68833.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_hirsutum ADY68836.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_barbadense PPR86822.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
G_hirsutum XP_016728080.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
Corchorus_olitorius OM078421.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
Durio_zibethinus XP_022735746.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
Theobroma_cacao XP_007036140.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
Herrania_umbatica XP_021292167.1	hpgllpvst---aqeelrkqigaayyiecsskt-qqnkvavf-----dda	172
Prunus_armeniaca CAB4296392.1	npml---k---raqlndteigrarvsesdsklyyiqav-----ket	381
G_davidsonii MBA0624583.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_arboreum KHG01290.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_schwendimani MBA0866527.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_schwendimani MBA0866525.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_triobum MBA0775238.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_australe KAA3488855.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_australe KAA3488856.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_arboreum XP_017643194.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_mustelinum TYJ35072.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_barbadense KAB2082570.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_hirsutum NP_001314395.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_tomentosum TYH71926.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_mustelinum TYI82308.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_davidsonii MBA0623572.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378
G_davidsonii MBA0623571.1	kphvl---k---kaeelrdthigrdrfvnetdigrlyyiqav-----ket	378

Populus_simonii AKV56346.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	425
G_arboreum XP_017611064.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_australe KAA3456360.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_gossypiooides MBA0735964.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_schwendimani MBA0851603.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_armourianum MBA0825516.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_triobum MBA0762230.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_davidsonii MBA0609728.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_barbadense PPD80758.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_mustelinum TYI91474.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_hirsutum NP_001314347.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_barbadense KAB2039250.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_tomentosum TYH81593.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_raimondii XP_012472188.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_hirsutum KAG4156417.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
G_darwinii TYG77701.1	pevmqevsvdipaenyvesetgnngfidpstlvngsipqlddl-----pdidaf	457
Glycine_max NP_001235844.2	irmvixppqkqekkkkpsrgclnfv-----s1-----	200
G_raimondii KJBS3185.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	191
G_aridum MBA0690107.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	204
G_raimondii ADY68835.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	197
Durio_zibethinus XP_022736748.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	198
G_hirsutum NP_001314362.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	198
G_raimondii XP_012440446.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	201
G_hirsutum ADY68837.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	198
G_barbadense ADY68833.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	197
G_hirsutum ADY68836.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	197
G_barbadense PPR86822.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	198
G_hirsutum XP_016728080.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	199
Corchorus_olitorius OM078421.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	198
Durio_zibethinus XP_022735746.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	196
Theobroma_cacao XP_007036140.1	ikvvi-ppqkqekkkkpsrgclnfv-----s1-----	196
Herrania_umbatica XP_021292167.1	akvvi-ppqkqekkkkpsrgclnfv-----s1-----	197
Prunus_armeniaca CAB4296392.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	426
G_davidsonii MBA0624583.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_arboreum KHG01290.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_schwendimani MBA0866527.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_schwendimani MBA0866525.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_triobum MBA0775238.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_australe KAA3488855.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_australe KAA3488856.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_arboreum XP_017643194.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_mustelinum TYJ35072.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_barbadense KAB2082570.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_hirsutum NP_001314395.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_tomentosum TYH71926.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_mustelinum TYI82308.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_davidsonii MBA0623572.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423
G_davidsonii MBA0623571.1	lrlyppapls---aprelsetcsiggy---dipkgrliinlhkiqdp-----kk	423

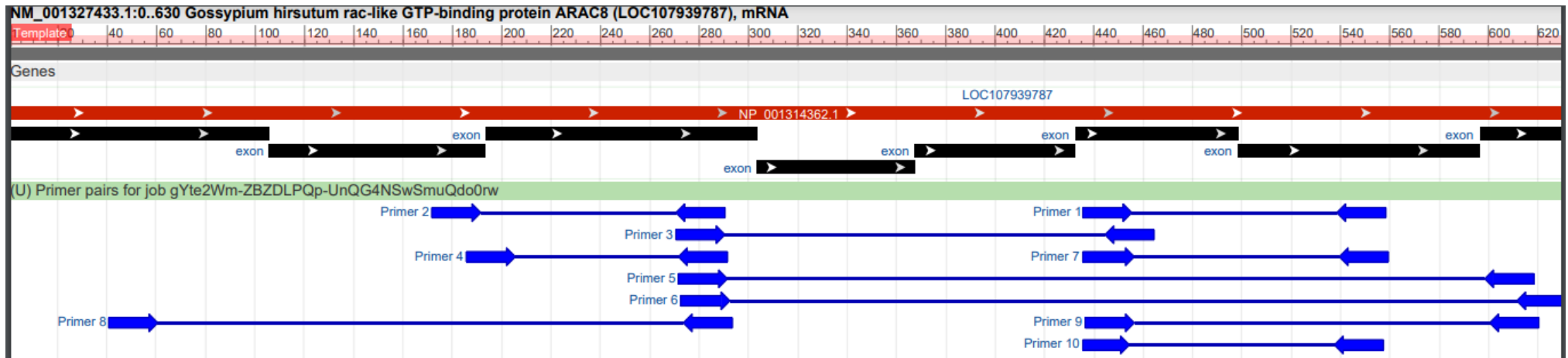
Populus_simonii AKV56346.1	wddllvqspdpediesssveakngndvqpnangwdka-----qhvdlqiteq----me	474
G_arboreum XP_017611064.1	wedllvpspepediestpmdskgegneqapmdkgwdes-----qhdndlteq----mg	506
G_australe KAA3456360.1	wedllvpspgpediestpmdskgegneqapmdkkrwdes-----qhdndlteq----mg	506
G_gossypioides MBA0735964.1	wedllapspgpediestpmdskgegneqapmdkgwdes-----qhdndlteq----mg	506
G_schwendimani MBA0851603.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qhdndlteq----mg	506
G_armourianum MBA0825516.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qhdndlteq----mg	506
G_trilobum MBA0762230.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qhdndlteq----mg	506
G_davidsonii MBA0609720.1	wedllvpspgpediestpmdskgernewapmdkgwdes-----qhdndlteq----mg	506
G_barbadense PPD80758.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_mustelinum TVI91474.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_hirsutum NP_001314347.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_barbadense KAB2039250.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_tomentosum TYH81593.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_raimondii XP_012472188.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_hirsutum KAG4156417.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
G_darwinii TVG77701.1	wedllvpspgpediestpmdskgegneqapmdkgwdes-----qymndlteq----mg	506
Glycine_max NP_001235844.2	-----	200
G_raimondii KJ853185.1	-----	191
G_aridum MBA0690107.1	-----ikvlgacyyliiv-----lty-----e	220
G_raimondii ADY68835.1	-----	197
Durio_zibethinus XP_022736748.1	-----	198
G_hirsutum NP_001314362.1	-----	198
G_raimondii XP_012440446.1	-----ikvlga-----	207
G_hirsutum ADY68837.1	-----	198
G_barbadense ADY68833.1	-----	197
G_hirsutum ADY68836.1	-----	197
G_barbadense PPR86822.1	-----	198
G_hirsutum XP_016728080.1	-----	199
Corchorus_olitorius OMO78421.1	-----	198
Durio_zibethinus XP_022735746.1	-----	196
Theobroma_cacao XP_007036140.1	-----	196
Herrania_umbtratica XP_021292167.1	-----	197
Prunus_armeniaca CAB4296392.1	wpdplefkperflthkdvdvkgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	486
G_davidsonii MBA0624583.1	wpepsefkperflthkdidvkgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_arboreum KHG01290.1	wpepsefkperflthkdidvkgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_schwendimani MBA0866527.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_schwendimani MBA0866525.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_trilobum MBA0775238.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_australe KAA3488855.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_australe KAA3488856.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_arboreum XP_017643194.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_mustelinum TVJ35072.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_barbadense KAB2082570.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_hirsutum NP_001314395.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_tomentosum TYH71926.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_mustelinum TVI82308.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_davidsonii MBA0623572.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483
G_davidsonii MBA0623571.1	wpepsefkperflthkdvdvrgqhfelmpfsgnrscpgtsfalhmlftmsnflhaf	483



Populus_simonii AKV56346.1	11-----tsdkk1-----	482
G_arboreum XP_017611064.1	11-----gsdnngs-----	515
G_australe KAA3456360.1	11-----gsdnrs-----	515
G_gossypoides MBA0735964.1	11-----gsdnngs-----	515
G_schwendimani MBA0851603.1	11-----gsennrs-----	515
G_armourianum MBA0825516.1	11-----gsddnrs-----	515
G_trilobum MBA0762230.1	11-----gsdnrs-----	515
G_davidsonii MBA0609728.1	11-----gsdnrs-----	515
G_barbadense PPD080758.1	11-----gsdnngs-----	515
G_mustelinum TYI91474.1	11-----gsdnrs-----	515
G_hirsutum NP_001314347.1	11-----gsdnrs-----	515
G_barbadense KAB2039250.1	11-----gsdnngs-----	515
G_tomentosum TYH81593.1	11-----gsdnrs-----	515
G_raimondii XP_012472188.1	11-----gsdnrs-----	515
G_hirsutum KAG4156417.1	11-----gsdnrs-----	515
G_darwinii TYG77701.1	11-----gsdnrs-----	515
Glycine_max NP_001235844.2	-----	200
G_raimondii KJB53185.1	-----	191
G_aridum MBA0690107.1	i-----	221
G_raimondii ADY68835.1	-----	197
Durio_zibethinus XP_022736748.1	-----	198
G_hirsutum NP_001314362.1	-----	198
G_raimondii XP_012440446.1	-----	207
G_hirsutum ADY68837.1	-----	198
G_barbadense ADY68833.1	-----	197
G_hirsutum ADY68836.1	-----	197
G_barbadense PPR86822.1	-----	198
G_hirsutum XP_016728080.1	-----	199
Corchorus_olitorius QMO78421.1	-----	198
Durio_zibethinus XP_022735746.1	-----	196
Theobroma_cacao XP_007036140.1	-----	196
Herrania_umbratica XP_021292167.1	-----	197
Prunus_armeniaca CAB4296392.1	isnpssapidmtesfgltnvkatplqvlikprlpsqlyg	525
G_davidsonii MBA0624583.1	fstpsnglidltgtdgltnikstplqalvvsprlapelyn	522
G_arboreum KHG01290.1	fstpsdglidltgtdgltnikstplealvvsprlapelyn	522
G_schwendimani MBA0866527.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_schwendimani MBA0866525.1	fstpsnglidltgtvgltnikstplealvvsprlapelyn	522
G_trilobum MBA0775238.1	fstpsnglidltgtvgltnikstplealvvsprlapelyn	522
G_australe KAA3488855.1	fstpsnslidlsagtagltnmkstplealvvsprlppelyk	522
G_australe KAA3488856.1	fstpsnglidltgtagltnikstplealwmnqvgt----	518
G_arboreum XP_017643194.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_mustelinum TYJ35072.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_barbadense KAB2082570.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_hirsutum NP_001314395.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_tomentosum TYH71926.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_mustelinum TYI82308.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_davidsonii MBA0623572.1	fstpsnglidltgtagltnikstplealvvsprlapelyn	522
G_davidsonii MBA0623571.1	fstpsnglidltgtvgltnikstplealvvsprlapelyn	522

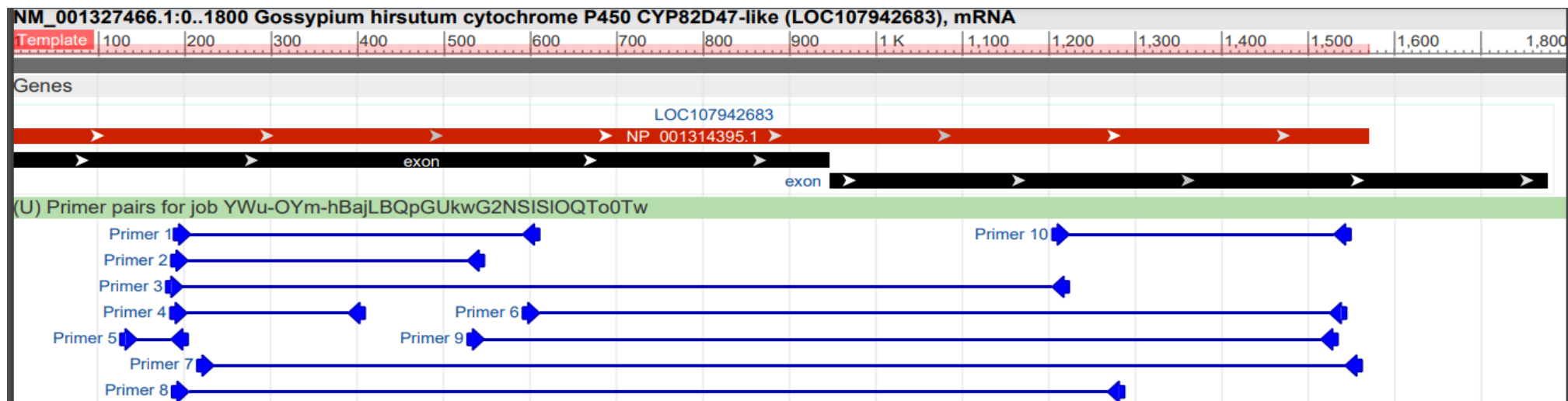
Fonte: Próprio autor.

Figura 20 - Domínio da sequência completa conservada da proteína referente ao Rop\_like juntamente com os primers presentes na plataforma, obtido no NCBI.



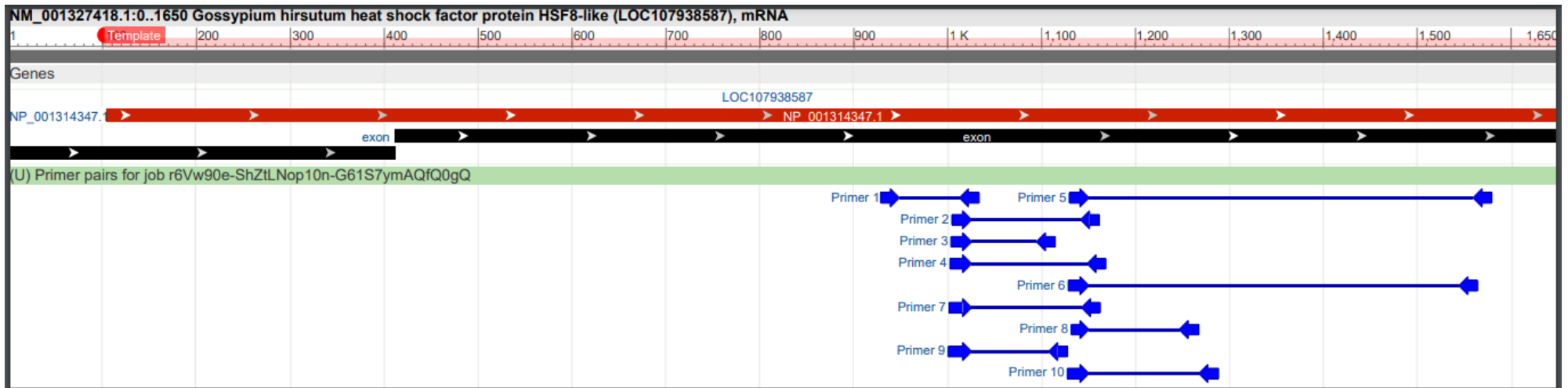
Fonte: Próprio autor.

**Figura 21** - Domínio da sequência completa conservada da proteína referente ao P450 juntamente com os primers presentes na plataforma, obtido no NCBI.



Fonte: Próprio autor.

**Figura 22** - Domínio da sequência completa conservada da proteína referente ao bZIP juntamente com os primers presentes na plataforma, obtido no NCBI.



Fonte: Próprio autor.