



**UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA**

Nathalia Maíra Cabral de Medeiros

**Análise filogenética e funcional de dois genes de reparo
homólogos a AP endonuclease em cana-de-açúcar:
ScARP1 e ScARP3.**

Natal

2014

Nathalia Maíra Cabral de Medeiros

**Análise filogenética e funcional de dois genes de reparo
homólogos a AP endonuclease em cana-de-açúcar:
ScARP1 e ScARP3.**

Dissertação apresentada ao
Programa de Pós-Graduação em
Bioquímica da Universidade Federal
do Rio Grande do Norte como
requisito para obtenção do título de
Mestre em Bioquímica.

Orientadora: Kátia Castanho Scorteci

Natal

2014

Catálogo da Publicação na Fonte
Universidade Federal do Rio Grande do Norte - UFRN

Medeiros, Nathalia Maira Cabral de.

Análise filogenética e funcional de dois genes de reparo homólogos a AP endonuclease em cana-de-açúcar: ScARP1 e ScARP3 / Nathalia Maira Cabral de Medeiros. - Natal, 2014.
127f. il.

Orientadora: Profa. Dra. Kátia Castanho Scorteci.

Dissertação (Mestrado) - Universidade Federal do Rio Grande do Norte. Centro de Biociências. Programa de Pós-Graduação em Bioquímica.

1. Cana-de-açúcar - Dissertação. 2. Relação filogenética - Dissertação. 3. Reparo de DNA - Dissertação. I. Scorteci, Kátia Castanho. II. Universidade Federal do Rio Grande do Norte. III. Título.

RN/UF/BSE01

CDU 575.111

Dedico esta obra

À Deus, meus pais, meu irmão, minha família e amigos pelo incentivo e apoio.

Agradecimentos

À minha família pela a paciência e o apoio ao longo deste trabalho, sem eles não conseguiria alcançar os meus objetivos

Á minha orientadora pela auxilio e companheirismo ao longo deste anos.

Aos meus companheiros de laboratório e amigos que me apoiaram e ajudaram.

Aos meus parceiros, co-autores do artigo, sem eles esse trabalho não teria se desenvolvido.

Á Capes pelo auxilio financeiro.

*“Segue o teu destino...
Rega as tuas plantas;
Ama as tuas rosas.
O resto é a sombra
de árvores alheias”*

Fernando Pessoa

RESUMO

O genoma de todos os organismos sofre constantemente a influência de fatores mutagênicos que podem ser de origem endógena e/ou exógena, estes podem resultar em danos ao material genético. Se esses danos não forem corrigidos pode levar ao aparecimento de mutações. As plantas por serem organismos sesséis estão continuamente expostas a estes fatores. Considerando isto, os organismos (animais e vegetais) possuem diferentes vias de reparo de DNA para manter a integridade do material genético. Dentro destas vias, há a via de Reparo por Excisão de Bases (BER) que é composta por diferentes enzimas, e dentro dessa via há a enzima AP endonuclease que é alvo deste estudo. Trabalhos anteriores em cana-de-açúcar identificaram duas sequências de cDNA homólogas a esta proteína que foram denominadas ScARP1 e ScARP3. Com isso, o objetivo deste trabalho foi caracterizar estas duas sequências por meio de análises filogenéticas utilizando sequências presentes dentro do reino *Plantae*, e de análises estruturais dos genes de AP endonuclease por análise *in silico* e por plantas transgênicas contendo cassetes de super-expressão. Além disso, foi realizado transformações e a obtenção plantas transgênicas de *Nicotiana tabacum* contendo cassetes de super-expressão em orientação anti-senso. Foi também analisado a relação filogenética de genes DNA ligase I presentes no organismo vegetal de estudo. Os resultados obtidos permitiram verificar que as sequências ScARP1 e ScARP3 correspondem a uma duplicação, provavelmente devido a um processo de duplicação do genoma como um todo (WGD) que deve ter ocorrido no grupo das gramíneas (*Poaceae*). Reforçando estes dados, foi verificado um possível direcionamento da proteína para organelas diferentes, sendo que a ScARP1 pode ser encontrada no núcleo e a ScARP3 em mitocôndrias e/ou cloroplasto. Com relação as plantas transgênicas contendo o cassete em orientação anti-senso foi observado que estas apresentaram crescimento lento quando comparado com a planta selvagem (não transformada). Além disso, seu fenótipo abrange alterações morfológicas no crescimento foliar, baixa estatura e diminuição na produção de sementes. Entretanto, ainda se faz necessário a obtenção da linhagem homocigota para aprofundar essas observações. Desta forma, estes resultados permitem compreender um pouco melhor do possível papel da enzima AP endonuclease em cana-de-açúcar e em plantas.

Palavras-chave: Relação filogenética. Direcionamento de proteínas. Duplicação. Plantas transgênicas. Expressão gênica.

ABSTRACT

The genome of all organisms constantly suffers the influence of mutagenic factors from endogenous and/or exogenous origin, which may result in damage for the genome. In order to keep the genome integrity there are different DNA repair pathways to detect and correct these lesions. In relation to the plants as being sessile organisms, they are exposed to this damage frequently. The Base Excision DNA Repair (BER) is responsible to detect and repair oxidative lesions. Previous work in sugarcane identified two sequences that were homologous to *Arabidopsis thaliana*: ScARP1 ScARP3. These two sequences were homologous to AP endonuclease from BER pathway. Then, the aim of this work was to characterize these two sequences using different approaches: phylogenetic analysis, *in silico* protein organelle localization and by *Nicotiana tabacum* transgenic plants with overexpression cassette. The *in silico* data obtained showed a duplication of this sequence in sugarcane and Poaceae probably by a WGD event. Furthermore, *in silico* analysis showed a new localization in nuclei for ScARP1 protein. The data obtained with transgenic plants showed a change in development and morphology. Transgenic plants had slow development when compared to plants not transformed. Then, these results allowed us to understand better the potential role of this sequence in sugarcane and in plants in general. More work is important to be done in order to confirm the protein localization and protein characterization for ScARP1 and ScARP3.

Keywords: phylogenetic relationship. Targeting proteins. Duplication. Transgenic plants. Gene expression.

LISTA DE FIGURAS

Figura 1 – **Esquema das diferentes vias de reparo.** Representação esquemática das diferentes vias de Reparo de DNA onde são representados os agentes promotores de lesões ao DNA (parte superior) e na parte mediana alguns exemplos de lesões de DNA induzidas por estes agentes; na parte inferior alguns dos mecanismos de reparo de DNA responsáveis pela remoção dessas lesões. No lado direito da imagem estão representados alguns efeitos das lesões no DNA durante a progressão do ciclo celular. Na parte inferior estão representadas as consequências das mutações surgidas. Abreviações: cis-Pt e MMC - cisplatina e mitomicina C, respectivamente; (6-4) FP e CPD - fotoproducto 6-4 e Ciclobutano Dímero de Pirimidina, respectivamente; BER e NER, Reparo por Excisão de Base e Nucleotídeo, respectivamente; RH, Recombinação Homóloga; EJ, end *joining*. (Adaptado de HOEIJMAKERS, 2009)

.....21

FIGURA 2 – **Esquema da via de excisão de bases.** Representação esquemática da via da excisão de bases (BER) em mamíferos. A via BER se inicia com o reconhecimento e remoção da lesão (representado pela estrela) pela DNA glicosilase. A DNA glicosilase bifuncional é capaz de clivar a ligação fosfodiéster e criar uma ponta 5´ fosfato e a outra ponta 3´ fosfato ou 3´ polinsaturado aldeído (PUA). No caso da remoção da base danificada por uma DNA glicosilase monofuncional, o corte da fita é exercida pela AP endonuclease, criando assim extremidades 3´hidroxila (OH) e 5´desoxiribose-fosfato (dRP). Estas extremidades “não-convencionais” devem ser restauradas para 3´OH e 5´ fosfato através da atividade desoxiribose-fosfato diesterase (dRPase) da Pol β que age sobre 5´ dRP, e da atividade diesterase da AP endonuclease ou da atividade fosfatase da aprataxin (APTX) que agem sobre 3´PUA. O reparo pode se encaminhar para duas vias via curta de BER (*short-patch*) ou via longa de BER (*long-patch*). Durante a via curta de BER a Pol β incorpora um nucleotídeo, em seguida ocorre a ligação com a fita de DNA através do complexo XRCC1/LigIII ou LigI. Na via curta a lesão 5´ é refratária a clivagem pela Pol β ,

sendo então orientada para a via longa de BER, em que Pol β e/ou Pol δ/ϵ realizam o deslocamento da fita, incorporando assim vários nucleotídeos, em seguida ocorre a remoção do “DNA Flap” contendo a lesão 5’ refratária pela Flap endonuclease (FEN) e a ligação a fita de DNA pela LigI. Proteínas de apoio a via de BER estão indicada em cinza (adaptado de BAUTE;DEPICKER,2008).....24

FIGURA 3 - Lesões encontradas nas bases nucleotídicas. Representação estrutura química das diferentes lesões nas bases púricas e pirimídicas, bem como os sítios abásicos. (Adaptado Krokan; Bjørås, 2013).....25

FIGURA 4 – Estrutura tridimensional das enzimas. Nesta figura é apresentado a similaridade estrutural entre as estruturas tridimensionais das enzimas DNase I, HAP1 (humana) e EXOIII (bactéria). As estruturas observadas em α -Helices estão coloridas em azul e as folhas- β na coloração magenta. As setas em amarelo indicam regiões de *loop* que estão presentes nas enzimas HAP1 e EXOIII, mas é ausente na enzima DNase I. (Adaptado de GORMAN et al., 1997).....27

FIGURA 5 – O genoma do cultivar atual de cana-de-açúcar. Organização esquemática do genoma do cultivar atual de cana-de-açúcar. Cada barra representa um cromossomo, as barras em amarelo (cerca de 80% a 90%) representam regiões originadas do *Saccharum officinarum*, as barras em verde (cerca de 10% a 20%) representam regiões originadas do *Saccharum spontaneum*. (Adaptada de GRIVET; ARRUDA,2002).....30

FIGURA 6 - Produção da cana-de-açúcar no Brasil. Representação da divisão da produção de produtos da indústria sucroalcooleira no Brasil, por região, dando em porcentagem (FONTE: CONAB - 2º Levantamento: agosto de 2013).....31

FIGURA 7 - Eventos poliploides ancestrais em plantas com sementes e angiospermas. Duas duplicações ancestrais identificadas pela integração de evidências filogenéticas e do relógio de tempo molecular para a evolução da planta na terra. As figuras ovais em cinza indicam as duplicações do genoma

geralmente aceitas e identificadas em genomas que foram totalmente sequenciados. As figuras em forma de diamante referem-se ao evento de triplicação provavelmente compartilhada por todos os eudicotiledôneas. As barras horizontais indicam regiões que indicam WGD (*whole genome duplication*) ancestral planta semente e WGD ancestral de angiospermas. (Adaptado do artigo JIAO et al., 2011)34

FIGURA 8 – Cassetes de superexpressão. Representação esquemática do cassete de super-expressão em orientação anti-senso (A) e senso (B). A sigla CaMV35S representa o promotor do vírus 35S promotor forte e constitutivo, *nptII* – Gene de resistência ao antibiótico canamicina, ARP1 - sequência de ScARP derivada de cana-de-açúcar, LacZ – gene lacZ de *E. coli* que codifica β-galactosidase, polyAAA – cauda poli-adenilada, os triângulos indicam sítios das enzimas de restrição que foram utilizadas para inserir a sequência de interesse.....44

FIGURA 9 – Transformação utilizando *Agrobacterium tumefaciens*. Esquema do processo de transformação por intermédio da *Agrobacterium tumefaciens*, nas plantas de *Nicotianum tabacum*. O número 1 representa a retirada de folhas *N. tabacum*; 2- Corte dos discos foliares que são colocados em contato com uma solução contendo *A. tumefaciens*, 3- Os disco são colocados em meio Ms solido acrescido de hormônios; 4 – Quando as plântulas começam a se formar são transferidas para um recipiente maior; 5- quando atingem certo tamanho são transferidas para a terra.....46

FIGURA 10 - Domínios conservados das sequências ScARP1 e ScARP3. Representação das superfamílias a que pertence, além dos sítios: catalíticos putativos, sítios de ligação ao DNA, de ligação AP (purínico e apirimídico, ou seja sítios abásicos), de ligação de metais e putativos de ligação com o fosfato. A – sequência de ScARP1. B – sequência de ScARP3.....54

Figura 11 – Alinhamento das sequências de Ap endonucleases. Alinhamento de sequencias de aminoácidos de organismos diferentes:EX3 de Escherichia coli (P09030), Apex1- Humana (P27695) e ScARPS (1 e 3) de cana-de-açúcar. Destacando-se em azul o sítio ativo, em amarelo o sítio de ligação ao cofator

enzimático (magnésio), em cinza um sítio importante para atividade catalítica e em verde o sitio de interação com o substrato da enzima que seria o DNA.....54

FIGURA 12 – Alinhamento das sequências homólogas a AP endonuclease. Representação do alinhamento obtido entre as sequências de aminoácidos de ScARP1, ScARP3 e AtARP. A coloração em vermelho mostra a o local onde as sequências se alinham.....56

FIGURA 13 - Dendograma das sequências scARP1 e scARP3. As sequências foram alinhadas utilizando o programa MAFFT v.6.85 ao nível dos aminoácidos com o L-INS-i e algoritmo CLUSTAL Ω v.1.0.3, com os parâmetros de alinhamento personalizados. Os dendogramas foram calculados com base em uma análise Bayesiana usando MrBayes 3.1.2 e BEAST v.1.7, este é uma árvore conseso gerado através de 1.000.000 árvores geradas no programa.....59

FIGURA 14– Análise estrutural do gene ScARP1. Representação da disposição dos Exons derivados do alinhamento do cDNA de ScARP1 com gene de *Sorghum bicolor* (gi|242060088:65246632-65250387).....60

FIGURA 15 - Análise estrutural do gene ScARP3. Representação da disposição dos Exons derivados do alinhamento do cDNA de ScARP3 com gene de *Sorghum bicolor* (gi|242060088: 65240591-65246229).....61

FIGURA 16 - Domínios conservados das sequências de nucleotídeo ScLig 1.1 e ScLig 1.2. Representação das superfamílias que pertencem as sequências de ScLig 1.1 e ScLig 1.2. Os triângulos em vermelho representam o sitio de ligação ao DNA já os triângulos em azul representam o sítio ativo.....63

FIGURA 17 – Domínios conservados das sequências de aminoácidos ScLig 1.1 e ScLig 1.2. Representação das superfamílias que pertencem as sequências de ScLig 1.1 e ScLig1.2. Os triângulos em vermelho representam o sitio de ligação ao DNA já os triângulos em azul representam o sítio ativo.....64

FIGURA 18 - Dendograma para as sequências com homologia a ScLig 1.1 e ScLig 1.2. Representação da relação filogenética entre as sequências scLig 1.1 e scLig 1.2, com relação a outros organismos vegetais, evidência a formação de dois grupos destacados na figura A árvore foi gerado pelo programa MEGA 5.1 utilizando o modelo de Dayhoff , com o método estatístico neighbor-joining com o teste de filogenia baseado com o método bootstrap com 500 replicatas, sendo o tratamento dos resultados Pairwise deletion.....69

FIGURA 19 – Expressão do gene ScARP em condições de estresse oxidativo. Representação gráfica da expressão de cDNA de scARP1 e scARP3 em tecido foliar quando submetidas a uma condição de estresse oxidativo com H₂O₂ (Adaptado de Maira et al., 2013).70

FIGURA 20 – Hipótese proposta. Representação da hipótese referente ao desfalque na via de excisão de base (BER), devido ao silenciamento da enzima Ap ensonuclease, deste modo a falta desta via de defesa pode gerar quebras de fita simples e dupla o que leva um atraso ou inibição do ciclo celular o que pode resultar em um atraso no desenvolvimento vegetal apresentado nas plantas transgênicas.....81

LISTA DE TABELAS

Tabela 1 – Identificação de proteínas de outros organismos que compartilham homologia com as sequências ScArp1 e ScArp3 de cana-de-açúcar.....42

Tabela 2 - Programa de amplificação por meio da PCR.....50

Tabela 3 - Posição do sitio de inicio da transcrição (TIS) com relação aos scores encontrados das proteínas ScARP1 e ScARP3 com relação as para organela alvo.....57

Tabela 4 - Tamanho das sequências de DNA ligase I de cana-de-açúcar (ScLig 1.1 e ScLig 1.2).....	62
Tabela 5 – Identidade entre as sequências de Dna ligase I de cana-de-açúcar e a sequência de Dna ligase I de Arabidopsis thaliana.....	65
Tabela 6 - Localização subcelular das sequências scLig1.1 e scLig1.2. nucl.- Núcleo, cyto – Citosol, plas- membrana plasmática, chlo – Cloroplasto, nucl. Plas –Membrana nuclear.	66
Tabela 7 – Localização subcelular das sequências homologas a scLig1.1. nucl.- Núcleo, cyto – Citosol, plas- membrana plasmática, chlo – Cloroplasto, nucl. plas –Membrana nuclear, vacu –Vacúolo, mito –Mitocôndria, cysk – Citoesqueleto.....	67
Tabela 8 - Localização subcelular das sequências homologas a ScLig 1.2. nucl.- Núcleo, cyto – Citosol, plas- membrana plasmática, chlo – Cloroplasto, nucl. plas –Membrana nuclear, vacu –Vacúolo, mito –Mitocôndria.....	68

LISTA DE ABREVIATURAS

BER Reparo por Excisão de Bases

WGD Duplicação do genoma como um todo

EROs Espécies reativas de oxigênio

PR Reparo por fotoreativação

DR Reparo direto

NER Reparo por excisão de nucleotídeos

Sítio AP Sítio apurínico/pirimídico

nonF Não-funcionalização

subF Subfuncionalização

neoF Neo- funcionalização

TIS Sítio de início da transcrição

NHEJ Reparo envolvendo recombinação não-homóloga

HR Recombinação homóloga

SSB Quebras de fita simples

DSB Quebra de fita dupla

SUMÁRIO

1 INTRODUÇÃO	18
1.1 REPARO DE DNA E SUA IMPORTÂNCIA PARA AS PLANTAS	18
1.2. Via de reparo por excisão de bases BER em plantas.	22
1.3 AP endonuclease - características e funções.	26
1.4. Cana-de-açúcar e sua importância.	29
1.5. Filogenia – História evolutiva da família Poacea	32
2 JUSTIFICATIVA	37
3 OBJETIVOS	38
4.MATERIAIS E MÉTODOS	39
4.1. Determinação da sequência e análise	39
4.2. Análise filogenética	40
4.3. Análise dos exons e introns das sequências scARP1 e scARP3.	42
4.4. Transformação das plantas utilizando os cassetes de superexpressão.	43
4.5. Análise do desenvolvimento das plantas transformadas e coleta de sementes.	47
4.6. Germinação das sementes transgênicas.	47
4.7. Medição das folhas de Comprimento <i>Nicotiana tabacum</i> .	48
4.8. Extração de DNA das folhas de <i>Nicotiana Tabacum</i> .	49
4.9. PCR.	49
5.0 extração de RNA.	50
5.1. Amplificação do cDNA.	51
5.2. qRT-PCR.	52
6.RESULTADOS	53
6.1.Análise das sequências scARP1 e scARP3.	53
6.2. Análise filogenética das sequências.	57
6.3. Análise de introns e exons das sequências	60
6.4. Obtenção das sequências de DNA ligase I em cana-de-açucar.	61

6.5. Análise filogenética das sequências de scLlg1 e scLig1.2 e localização subcelular.	65
6.6. Expressão dos cDNAs.	70
7. DISCUSSÃO	71
7.1. ScARPs e sua caracterização.	71
7.2. ScARPs e a filogenia.	72
7.3. Plantas transgênicas e sementes.	77
8. CONCLUSÃO	82
9. PERSPECTIVAS	83
REFERÊNCIAS	84

1 INTRODUÇÃO

1.1 Reparo de DNA e sua importância para as plantas

De forma geral, o genoma de todos os organismos vivos é estável e se mantém íntegro ao longo da vida do ser (TUTEJA et al., 2001; MARÉCHAL; BRISSON, 2010; MARÉCHAL et al., 2009). Essa integridade se faz necessária pois o sistema de informação genética é suscetível a sofrer lesões e mutações considerando a existência de um vasto número de agentes mutagênicos presentes no meio ambiente. Além disso, deve-se levar em conta também a existência da taxa de erro intrínseco presente no processo de replicação do DNA, com isso, para evitar tal fato há um sistema de reparo do DNA capaz de verificar esses erros e corrigi-los (ARAVIND et al., 1999). Deste modo, se for observado o genoma de organismos mais simples, a exemplos de alguns vírus, eles codificam múltiplas proteínas cuja função estaria relacionada com o reparo do DNA lesado (FRIEDBERG et al., 1995). Essa análise torna evidente que a necessidade de manutenção da “integridade do genoma” é algo fundamental e característico dos organismos vivos.

O ambiente em que o organismo está inserido além de lhe fornecer condições para sua sobrevivência (água, luz, nutrientes, etc...) também é fonte de agentes estressores, sendo estes de origem abióticos ou bióticos. Estes agentes podem promover uma variedade de lesões no DNA e que, por sua vez, podem vir a comprometer a integridade do genoma se não forem corrigidas (SINGH et al., 2010). Essas lesões podem afetar tanto a replicação do DNA (BRANZEI; FOIANI, 2010) quanto à transcrição (LAGERWERF et al., 2011). Com relação às plantas, estas por serem organismos sésseis, estão constantemente expostas a agentes ambientais e processos endógenos que infligem danos ao DNA e acarretam um estresse genotóxico, na qual podem reduzir a estabilidade do genoma e afetar diretamente o crescimento,

desenvolvimento e conseqüentemente a produtividade do vegetal (SINGH et al., 2010; TUTEJA et al., 2009; ATKINSON; URWIN, 2012). Exemplos destes estresses abióticos seriam as temperaturas extremas (congelamento, calor, frio), disponibilidade de água (seca, inundaçãõ) e toxicidade provocada por íons (salinidade, metais pesados). As plantas também podem ser suscetíveis a diversas infecções ocasionadas por vários tipos de patógenos (estresse biótico) tais como fungos, bactérias e vírus que podem modificar o balanço oxidativo (MAHAJAN; TUTEJA, 2005; HORBACH et al., 2011). Todos esses fatores estressores podem afetar o desenvolvimento vegetativo como também a produção de sementes, conseqüentemente estes fatores têm um papel em limitar a produtividade (ATKINSON; URWIN, 2012). Exemplo disso seria o prejuízo das culturas promovido pelo estresse abiótico que pode ser em torno de 50 % em todo o mundo (BRAY et al., 2000). Assim sendo, para sobreviverem em um ambiente “hostil”, os organismos vegetais desenvolveram mecanismos de defesa únicos em respostas a essas condições ambientais adversas (TUTEJA et al., 2009, KIMURA; SAKAGUCHI, 2006). Esses mecanismos estariam relacionados a vias de sinalização e vias de reparo de DNA que atuam de forma eficiente para promover o reparo de DNA e permitindo assim à sobrevivência da planta sem afetar tanto seu desenvolvimento e produtividade.

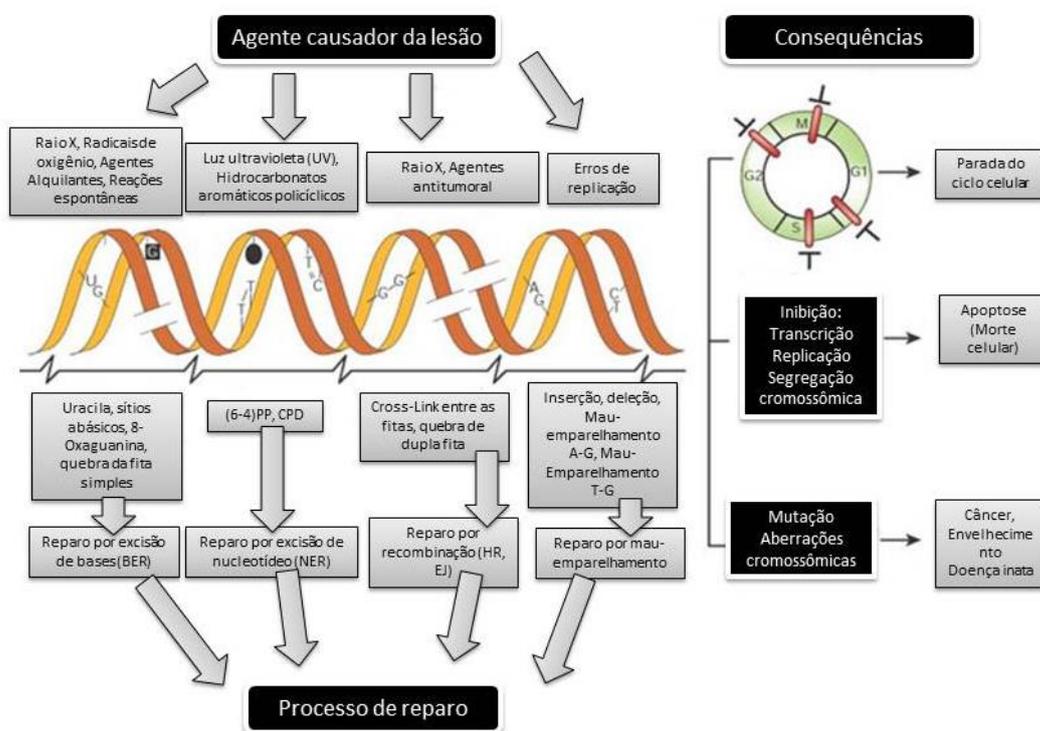
Com relação ao reparo de DNA tem sido observado que estas vias são conservadas em todos os organismos. As vias são classificadas em: via de reparo por fotoreativação (PR) ou também denominada de reparo direto (DR), reparo por excisão de bases (BER), reparo por excisão de nucleotídeos (NER), reparo de quebra de fita dupla (DSB) e reparo por incompatibilidade (MMR) (TUTEJA et al., 2001; SINGH et al., 2011). A figura 1 ilustra que as vias de reparo muitas vezes são específicas para os danos a serem identificados e corrigidos. Não obstante os mecanismos de identificação e correção da lesão no DNA sejam bem descritos em vários organismos como *Escherichia coli*, *Saccharomyces cerevisiae*, ratos e humanos, pouco se sabe com relação a esse processo em plantas.

Com isso, se faz necessário um estudo mais aprofundado sobre essas vias de reparo em plantas, bem como buscar entender a sua regulação (SHI et al., 1997; BOESCH et al., 2011). A importância do estudo nos sistemas vegetais

está diretamente associado ao desenvolvimento e produção das sementes. Rajjou e Debeaujon (2008) verificaram que as vias de reparo de DNA têm um papel importante em diferentes aspectos fisiológicos nas sementes. Kranner et al. (2010) observaram que o tempo prolongado de armazenamento das sementes promove danos ao DNA decorrentes do acúmulo de EROs (espécies reativas de oxigênio). Estes pesquisadores verificaram que as vias de reparo do DNA têm um papel crucial no processo de embebição da semente.

Outra observação importante em plantas foi com relação às proteínas helicases que catalisam a quebra das pontes de hidrogênio, abrindo desta forma a dupla hélice de DNA durante os processos de replicação, recombinação, biossíntese de ribossomos, transcrição e reparo do DNA. Deste modo, as proteínas helicases podem contribuir para o reparo de DNA por favorecer a acessibilidade da região com a lesão. Experimentos com plantas de tabaco contendo cassetes de superexpressão para estas enzimas verificaram que as plantas transgênicas tiveram uma maior tolerância em alta concentração de sal (SANAN-MISHRA et al., 2005; VASHISHT; TUTEJA, 2006).

Figura 1 – **Esquema das diferentes vias de reparo**. Representação esquemática das diferentes vias de Reparo de DNA, onde são representados os agentes promotores de lesões ao DNA (parte superior) e na parte mediana alguns exemplos de lesões de DNA induzidas por estes agentes; na parte inferior alguns dos mecanismos de reparo de DNA responsáveis pela remoção dessas lesões. No lado direito da imagem estão representados alguns efeitos das lesões no DNA durante a progressão do ciclo celular. Na parte inferior estão representadas as consequências das mutações surgidas. Abreviações: cis-Pt e MMC - cisplatina e mitomicina C, respectivamente; (6-4) FP e CPD - fotoproducto 6-4 e Ciclobutano Dímero de Pirimidina, respectivamente; BER e NER, Reparo por Excisão de Base e Nucleotídeo, respectivamente; RH, Recombinação Homóloga; EJ, *end joining*.



Fonte : Adaptado de HOEIJMAKERS, 2009

Desta maneira, o principal objetivo das pesquisas voltadas para a área de reparo de DNA em plantas é adquirir um melhor entendimento deste processo básico, a qual poderá propiciar uma vantagem nas atuais tecnologias de melhoramento vegetal. Este conhecimento terá implicações em aumentar a tolerância a alguns tipos de estresse (abiótico e biótico) e consequentemente promover o aumento da produtividade por permitir que os cultivares consigam tolerar estas condições adversas (BALESTRAZZI; CONFALONIERI, 2011).

1.2. Via de reparo por excisão de bases BER em plantas.

Tendo em vista a importância das vias de reparo para a manutenção da integridade do material genético, se faz necessário entender um pouco melhor a Via de Reparo por Excisão de Bases (BER) que é foco deste trabalho. A via BER tem como função a identificação e correção de lesões oxidativas, alquilação ou oriundas de deaminações nas bases nitrogenadas (GOSSAGE et al., 2012). A correção pode ser de apenas uma base ou em até 13 bases (BARNES; LINDAHL, 2004; FORTINI; DOGLIOTTI, 2007; BAUTE; DEPICKER, 2008; ZHARKOV, 2008).

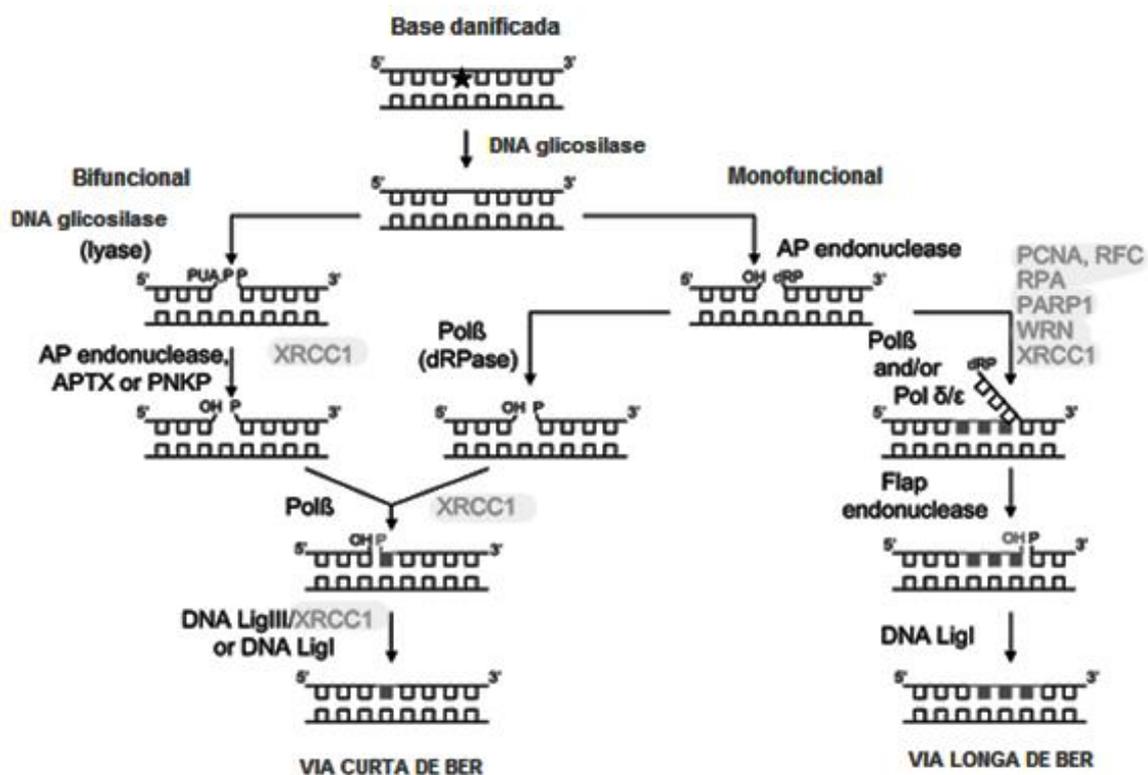
A via BER é formada por várias etapas, como apresentada na figura 2. A identificação das lesões é realizada por enzimas DNA glicosilases onde cada uma reconhece uma lesão específica. Assim, de forma resumida, o mecanismo pode ser descrito da seguinte maneira: primeiramente ocorre o reconhecimento da lesão e a remoção da base danificada criando assim um sítioapurínico/pirimídico (sítio AP). Em seguida ocorre a clivagem da ligação fosfodiéster da fita dupla do DNA e, com isso, a excisão do sítio AP. A lacuna gerada é preenchida com nucleotídeos pela DNA polimerase e depois é religada pela DNA ligase (BAUTE; DEPICKER, 2008) (figura 2). O reconhecimento da base danificada é específico para enzima DNA glicosilase que remove a base lesionada por meio da clivagem da ligação glicosídica, que liga a base nitrogenada ao açúcar (desoxirribose), isso gera a um sítioapurínico/pirimídico que também pode ser nomeado de sítio abásico (sítio AP), as lesões comuns às bases nitrogenadas são apresentadas na figura 3. Cada lesão requisitará uma DNA glicosilase específica para reconhecê-lo e removê-lo. Todavia, esses sítios abásicos também podem surgir por meio de despurinação e despirimidinação espontânea e assim não necessitando da ação da DNA glicosilase para ocorrer (MAYNARD et al., 2009; McCULLOUGH et al., 1999; MOL et al., 1999). O sítio AP pode ser então processado por uma DNA glicosilase bifuncional (com função 3'AP liase) ou por uma AP endonuclease, ambas enzimas clivam a ligação fosfodiéster de modo a promover a remoção do sítio AP. Esta remoção resulta

na produção de uma lacuna e de extremidades do DNA “não-convencionais”. A enzima AP endonuclease gera extremidades 3´OH e a 5´ desoxiribose-fosfato (5´dRP), enquanto as enzimas com função AP liases geram as extremidades 5´ fosfato e a 3´ contendo uma lesão “blocking”, também denominado 3´ fosfato ou 3´ polinsaturado aldeído (PUA), tais extremidades (5´dRP e PUA) atuam impedindo a atividade das enzimas subsequentes na via BER. Portanto, para que a DNA polimerase e a DNA ligase exerçam as suas respectivas funções, se faz necessário o processamento destas extremidades “não-convencionais” para que estas sejam restauradas na conformação ideal (convencional): 3´OH e 5´fosfato. A fração 5´dRP pode ser removido pela atividade 5´ dRPase da DNA polimerase β (Pol β) (MATSUMOTO; KIM, 1995; DETERDING et al., 2000). Enquanto, a enzima AP endonuclease possui a atividade 3´diesterase e é capaz de remover a lesão “blocking” 3´ (IZUMI et al., 2000), a atividade de fosfatase da aprataxin (APTX) também pode atuar da mesma forma sobre a lesão PUA. Após este processamento a lacuna (*gap*) será então preenchida por nucleotídeos através da atuação da enzima DNA polimerase e depois esses fragmentos serão ligados ao esqueleto do DNA por meio da enzima DNA ligase, e por fim a correção da lesão foi realizada e a integridade do genoma foi mantida.

Além disso, a via de excisão de bases pode ser subdividida em duas vias: a via curta de BER (*short-patch*) ou a via longa de BER (*long-patch*) ambas apresentadas na figura 2. A escolha entre qual via será seguida depende de vários fatores como primeiramente a natureza da lesão e o tipo de DNA glicosilase que irá iniciar o processo de correção (FORTINI; DOGLIOTTI, 2007). A via curta é responsável pelo o reparo de um nucleotídeo apenas, enquanto a segunda via é responsável por 2 a 13 nucleotídeos (BAUTE; DEPICKER, 2008). Na via curta de BER, os nucleotídeos são incorporados ao DNA pela Pol β em mamíferos ou pela Pol I em bactéria *Escherichia coli* (SINGHAL et al., 1995). A ligação ao DNA é realizada pelo complexo XRCC1 e pela enzimas LigIII (em mamíferos) ou LigI (em bactérias) (CAPPELLI et al., 1997; NASH et al., 1997). Durante a via longa de BER em mamíferos, a enzima Pol β provavelmente incorpora o primeiro nucleotídeo. No entanto, devido a existência da extremidade 5´ “não-convencional” essa enzima não é capaz de seguir adiante sendo assim, é substituída pela enzima Pol δ ou Pol ϵ . Esta ação resulta na formação de uma

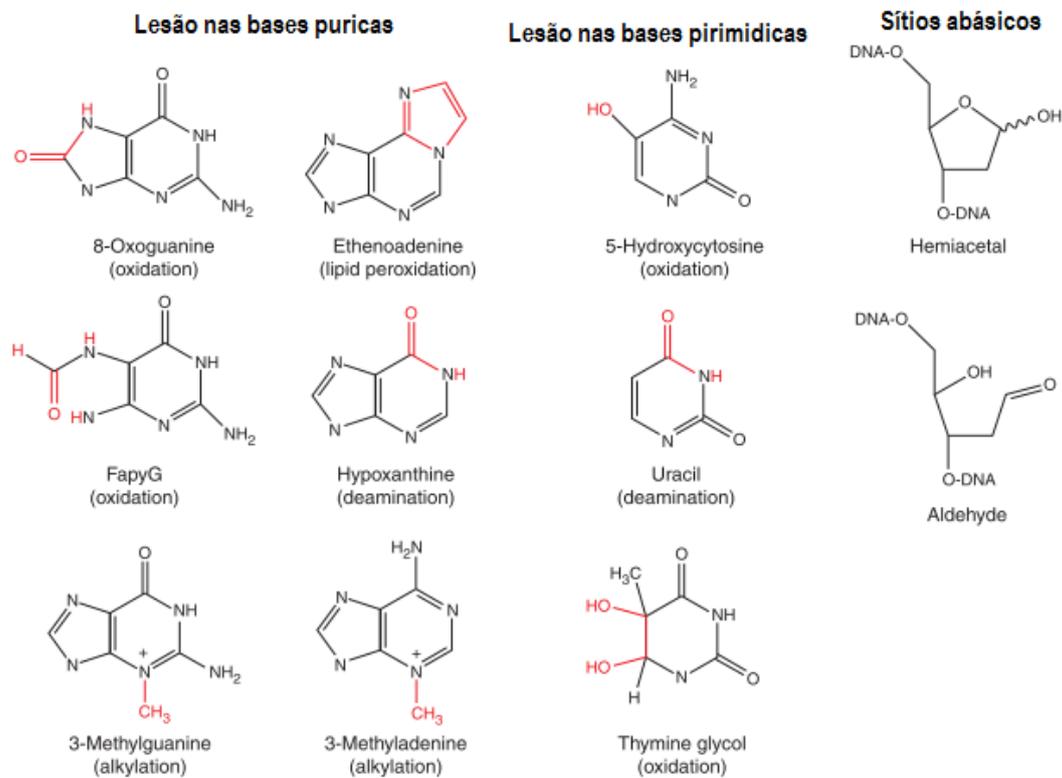
estrutura “Flap” que é então removida pela endonuclease FEN1 e posteriormente ocorre a ligação da fita de DNA por meio da enzima LigI (LEVIN et al., 1997; BAUTE; DEPICKER, 2008). As vias BER longas e curtas foram constituídas a partir de experimentos *in vitro* utilizando proteínas expressas e purificadas (DIANOV; LINDAHL, 1994; KLUNGLAND; LINDAHL, 1997; KUBOTA et al., 1996; NICHOLL et al., 1997; PASCUCCI et al., 1999; SRI-VASTAVA et al., 1998).

Figura 2 - **Esquema da via de excisão de bases.** Representação esquemática da via de excisão de bases (BER) em mamíferos. A via BER se inicia com o reconhecimento e remoção da lesão (representado pela estrela) pela DNA glicosilase. A DNA glicosilase bifuncional é capaz de clivar a ligação fosfodiéster e criar uma ponta 5' fosfato e a outra ponta 3' fosfato ou 3' polinsaturado aldeído (PUA). No caso da remoção da base danificada por uma DNA glicosilase monofuncional, o corte da fita é exercida pela AP endonuclease, criando assim extremidades 3' hidroxila (OH) e 5' desoxiribose-fosfato (dRP). Estas extremidades “não-convencionais” devem ser restauradas para 3'OH e 5' fosfato através da atividade desoxiribose-fosfato diesterase (dRPase) da Pol β que age sobre 5' dRP, e da atividade diesterase da AP endonuclease ou da atividade fosfatase da aprataxin (APTX) que agem sobre 3' PUA. O reparo pode se encaminhar para duas vias via curta de BER (*short-patch*) ou via longa de BER (*long-patch*). Durante a via curta de BER a Pol β incorpora um nucleotídeo, em seguida ocorre a ligação com a fita de DNA através do complexo XRCC1/LigIII ou LigI. Na via longa de BER a Pol β incorpora um nucleotídeo, em seguida ocorre a remoção do “DNA Flap” contendo a lesão 5' refratária pela Flap endonuclease (FEN) e a ligação a fita de DNA pela LigI. Proteínas de apoio a via de BER estão indicada em cinza



Fonte :adaptado de BAUTE ; DEPICKER, 2008.

Figura 3- **Lesões encontradas nas bases nucleotídicas.** Representação estrutura química das diferentes lesões nas bases púricas e pirimídicas, bem como os sítios abásicos.



Fonte: Adaptado Krokan; Bjørås, 2013.

Como apresentado acima, em plantas havia a dúvida se na via BER as duas vias ocorriam (curta e longa), uma vez que não foi identificado no genoma de *Arabidopsis* e de arroz o gene homólogo a enzima Pol β (BRITT, 2002; HAYS, 2002; ROLDAN-ARJONA; ARIZA, 2009). Entretanto, Córdoba-Cañero et al. (2011) demonstraram a partir de experimentos bioquímicos utilizando extrato proteico total de folhas de *Arabidopsis* que as plantas devam possuir as duas vias de reparo BER. Sendo assim, foi proposto que o papel da enzima Pol β deva ser realizado pelo DNA Pol λ (UCHIYAMA et al., 2004). Além disso, também não foi identificado as sequências homólogas para a enzima DNA LigIII, conseqüentemente foi proposto que a função de ligação seria realizada pela DNA ligase I e/ou DNA ligase IV (BAUTE; DEPICKER, 2008). Tendo em vista que poucos estudos têm sido realizados nas plantas para a via de reparo BER, se faz necessário ampliar os estudos nesta área. Desta forma, a caracterização

das enzimas scARP1 e scARP3 (AP endonuclease) podem vir a preencher parcialmente a lacuna de conhecimento existente nessa área.

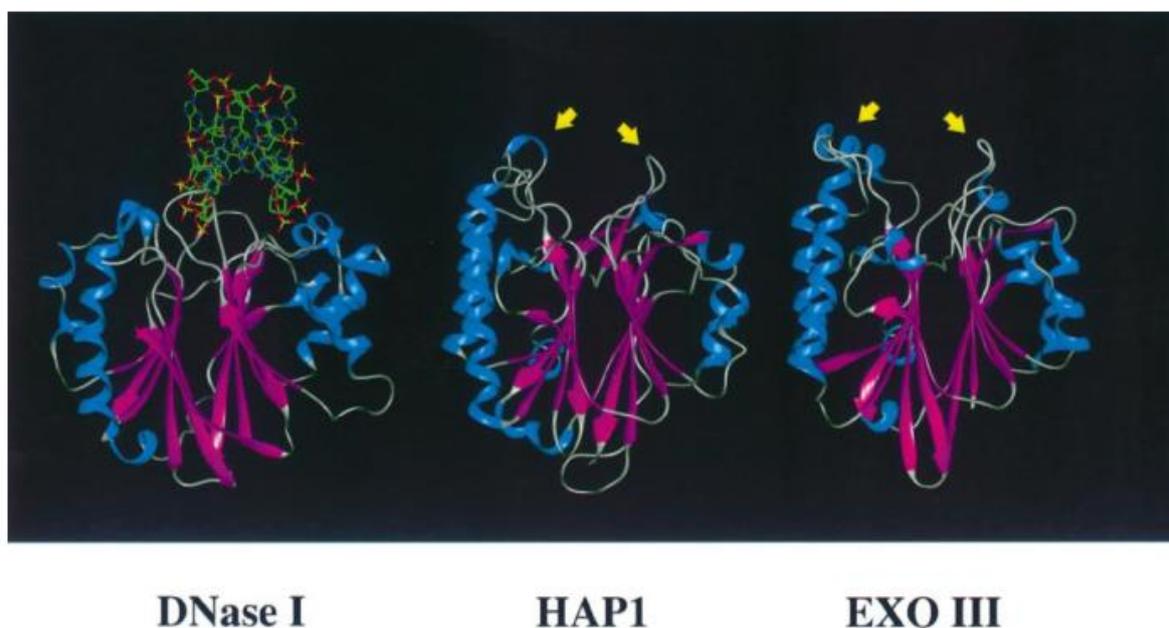
1.3 AP endonuclease - características e funções.

A presença de sítios apurínicos/ apirimídico (AP) é considerada uma lesão comum no DNA (LINDAHL, 1993; LINDAHL; BARNES, 2000; CICCIA; ELLEDGE, 2010). Os sítios APs podem ser formados espontaneamente ou podem ser intermediários centrais da via BER. Tais sítios podem bloquear a replicação e causar mutação. Desta forma, seu reparo é extremamente necessário para manter a integridade genética (LOEB, 1985; SOBOL et al., 2003), sendo tal função desempenhado pela via já mencionada anteriormente - a via de reparo por excisão de bases (BER).

Como já descrito, o sítio AP serve de substrato para a enzima AP endonuclease. Esta classe de enzima faz parte de duas distintas famílias que catalisam o processo de clivagem da ligação fosfodiéster na posição 5' da lesão: a família da endonuclease bacteriana IV (Nfo) e a família de exonuclease III (Xth), que inclui tanto a Xth de bactérias como a AP endonuclease (APE1) encontrada em eucariontes (MOL et al., 2000). O mecanismo de ação destas enzimas, de forma geral, é detectar, reconhecer e clivar a ligação fosfodiéster do DNA na posição 5' do sítio AP, criando assim uma extremidade 3'-OH. Sendo esta extremidade posteriormente processada com a ação de uma DNA polimerase. Contudo, a AP endonuclease também pode atuar como uma 3'->5' exonuclease e assim catalisar o processo de excisão de nucleotídeos para o reparo, particularmente de nucleotídeos com bases oxidadas (CHOU; CHENG, 2002; ISHCHENKO et al., 2003; ISHCHENKO et al., 2004; ISCHENKO et al., 2002; KERINS et al., 2003; DEMPLÉ et al., 1986). Apesar de ambas as famílias aqui mencionadas desempenharem a mesma função, suas estruturas tridimensionais e dependência de cofatores enzimáticos (metais) são distintas. Em estudos iniciais com relação Nfo, este diferenciava da família Xth pela a

resistência a EDTA (CUNNINGHAM et al., 1986; LEVIN et al., 1988), todavia estudos posteriores demonstraram que além desta discrepância inicial, a estrutura física da enzima era igualmente distinta, sendo que a família Xth apresenta duas camadas de folhas beta flanqueado por hélices, enquanto que o cofator enzimático seria Mg^{+2} (MOL et al., 1995), em contraste a família Nfo era formada por enzimas com o núcleo de barril TIM rodeado por hélices, e possui como cofatores enzimáticos de dois a três Zn^{2+} ou um Mn^{+2} (HOSFIELD et al., 1999). Assim sendo, entender esta distinção de ligação de ions metálicos e as atividades destas famílias tornou-se foco de diversos estudos sendo necessário para de fato compreender tal enzima essencial para o desempenho do reparo e manutenção da integridade da molécula de DNA (CVETKOVIC et al., 2010; BARONDEAU et al., 2004; MCMURRAY; TAINER, 2003; HOPFNER et al., 2002; CASTAGNETTO et al., 2002).

Figura 4 – **Estrutura tridimensional das enzimas.** Nesta figura é apresentado a similaridade estrutural entre as estruturas tridimensionais das enzimas DNase I, HAP1 (humana) e EXOIII (bactéria). As estruturas observadas em α -Hélices estão coloridas em azul e as folhas- β na coloração magenta. As setas em amarelo indicam regiões de *loop* que estão presentes nas enzimas HAP1 e EXOIII, mas é ausente na enzima DNase I.



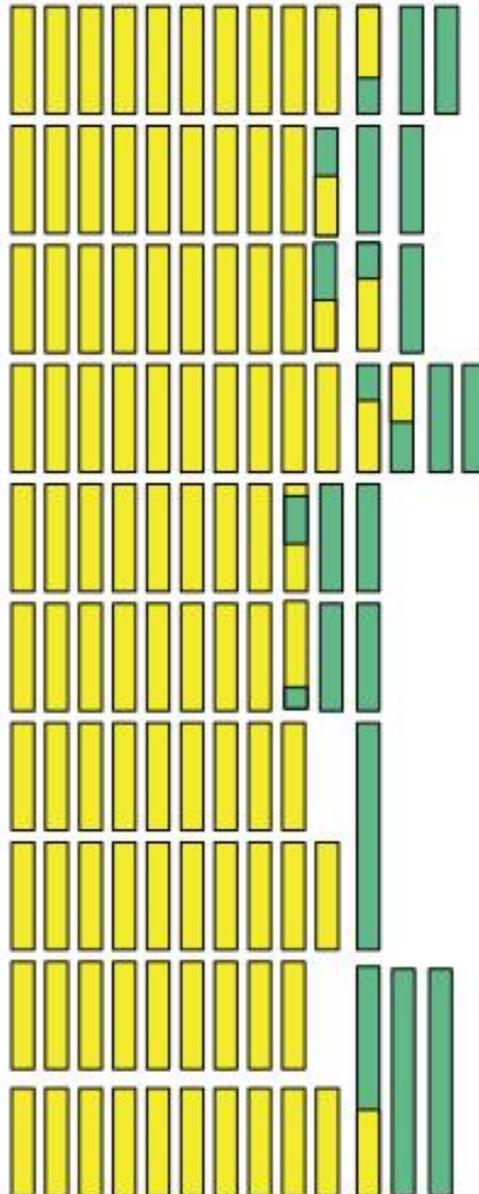
Fonte: Adaptado de GORMAN et al., 1997.

A enzima APE1 humana é uma das AP endonucleases mais conhecidas e que serve de base para estudos comparativos de outras sequências, além disso, esta enzima é também denominada como HAP1 e REF1. As análises comparativas estruturais desta proteína em bactérias e no homem tem mostrado que estas apresentam uma estrutura similar como apresentado na figura 4. Entretanto, existem algumas diferenças que podem ser observadas como: a) a presença de um resíduo de 61 aminoácidos na região N-terminal apenas presente na enzima HAP1 (humana) (DEMPLE et al., 1991; WALKER et al., 1993), b) a presença na região do *loop* compartilhada apenas entre as sequências HAP1 e EXOIII e ausentes na enzima DNase I (GORMAN et al., 1997). Esta região foi proposta como responsável para o reconhecimento de resíduos específicos para a atividade AP endonuclease. Além disso, a estrutura do cristal da APE1 (AP endonuclease humana) mostrou que esta enzima se liga em ambas as curvaturas maior e menor do DNA, além de provocar o processo de “flip out” da desoxirribose abásica que facilitaria a ação da enzima no reconhecimento do sitio AP bem como no seu reparo subsequente (GORMAN et al., 1997; MOL et al., 2000; BEERNINK et al., 2001; GEORGIADIS et al., 2008). No trabalho de Mol et al. (2000) foi proposto uma forma de como a enzima agiria sobre o sítio AP, aonde o aminoácido na posição Asp -210 (Ácido Aspártico posição 210) atuaria no ataque da molécula de água, e o fosfato intermediário seria estabilizado pelo ion Mg^{2+} em contato com os aminoácidos: His-309 (Histidina na posição 309), Asn-174 (Asparagina na posição 174) e Asn-212 (asparaina na posição 212). Os aminoácidos Asp/His atuariam ativando uma molécula de água que iniciaria, por sua vez, um ataque nucleofílico na ligação fosfodiéster, provocando assim a excisão do sitio AP (GORMAN et al., 1997). Existem outros estudos que propõem mecanismos alternativos por meio da observação de mutações no sitio ativo (ERZBERGER; WILSON, 1999; GALBURT et al., 1999; TSUTAKAWA et al., 1999; MUNDLE et al., 2004; MUNDLE et al., 2009), mas ainda não existe, atualmente, um consenso com relação ao modo de atuação do sitio ativo de forma mais detalhada (TSUTAKAWA et al., 2013).

1.4. Cana-de-açúcar e sua importância

A cana-de-açúcar corresponde a uma planta C₄ e pertence ao gênero *Saccharum* da família *Poaceae* e os cultivares utilizados atualmente correspondem a híbridos derivados do cruzamento interespecífico entre *S. officinarum* e *S. spontaneum*. Devido a estes cruzamentos o genoma da cana-de-açúcar tem sido considerado complexo, possuindo de 100 a 130 cromossomos de acordo com o cultivar. Com isso, tem sido estimado que 80% a 90% do genoma são derivados de *S. officinarum*, enquanto cerca de 10% a 20% são oriundos do *S. spontaneum* (D'HONT et al., 1996; PIPERIDIS et al., 2010) (figura 5). Decorrente desse processo de melhoramento genético clássico, o genoma da cana-de-açúcar apresenta-se altamente poliplóide e aneuploide contendo um conjunto completo de genes homólogos que variam de 10 a 12 cópias (alelos) (D'HONT et al., 1996).

Figura 5 – **O genoma do cultivar atual de cana-de-açúcar.** Organização esquemática do genoma do cultivar atual de cana-de-açúcar. Cada barra representa um cromossomo, as barras em amarelo (cerca de 80% a 90%) representam regiões originadas do *Saccharum officinarum*, as barras em verde (cerca de 10% a 20%) representam regiões originadas do *Saccharum spontaneum*.



Fonte: Adaptada de GRIVET; ARRUDA, 2002.

Plantas C₄ são colocadas como ideias como culturas energéticas, porque possuem as seguintes características: alta eficiência de conversão de luz em energia de biomassa, a eficiência elevada do uso da água e alta eficiência de

uso de nitrogênio no nível de folha (TAYLOR et al., 2010). Além disso apresentam a capacidade de crescer em áreas marginais de terra e uma relativamente alta tolerância às limitações do solo, tais como salinidade e encharcamento. As gramíneas com fotossíntese do tipo C₄ possuem um mecanismo de CO₂ concentrando adicional que lhes permita superar espécies C₃, especialmente sob condições de alta luminosidade e temperatura e são, portanto, capazes de gerar maior quantidade biomassa, mesmo em ambientes de recursos limitados. Desta forma, a cana-de-açúcar, devido essas características percorridas anteriormente, apresentam uma vantagem no acúmulo de biomassa podendo ser utilizado para gerações de combustíveis.

Nos trópicos, a cana-de-açúcar está organizada como uma monocultura com grande importância econômica, pois o açúcar acumulado nos seus colmos é utilizado para a produção de açúcar e etanol. Atualmente, o Brasil é o maior produtor mundial de cana-de-açúcar, sendo responsável por 45 % da produção (SCORTECCI et al., 2012). O interesse econômico nesta espécie cresceu significativamente nos últimos anos devido a demanda crescente mundial de uma produção de energia de forma sustentável (CHEAVEGATTI-GIANOTTO et al., 2011).

Figura 6- **Produção da cana-de-açúcar no Brasil**. Representação da divisão da produção de produtos da indústria sucroalcooleira no Brasil, por região, dando em porcentagem



Fonte: CONAB - 2º Levantamento: agosto de 2013..

No Brasil, o estado de São Paulo vem sendo responsável por 87 % da produção brasileira e os estados do nordeste são responsáveis por apenas 8 %. Essa redução na produção está associada à queda de produtividade devido às condições adversas (estresse abiótico e biótico), principalmente decorrentes do solo arenoso que não retém a água, mesmo quando a cultura é irrigada (CONAB, 2014). Assim, para a safra de 2013/14, é estimado que a área plantada para essa cultura em todo Brasil correspondesse a uma área de 8.799.150 mil hectares. Desta maneira a região Sudeste é considerada a maior produtora com 428.298,5 T, seguido pela região Centro-oeste (120.907,5 T), logo depois vem a região Nordeste (54.219,7 T), constituindo as regiões com menor produção seriam as regiões Sul (44.890,1 T) e Norte (3.700,1 T) como apresentado na figura 6.

Deste modo, estes dados mostram que é importante ampliar o conhecimento sobre essa cultura com grande apelo econômico para o Brasil, apesar de ser considerado um organismo com genoma complexo, e que responde de forma diferente a cada região do Brasil.

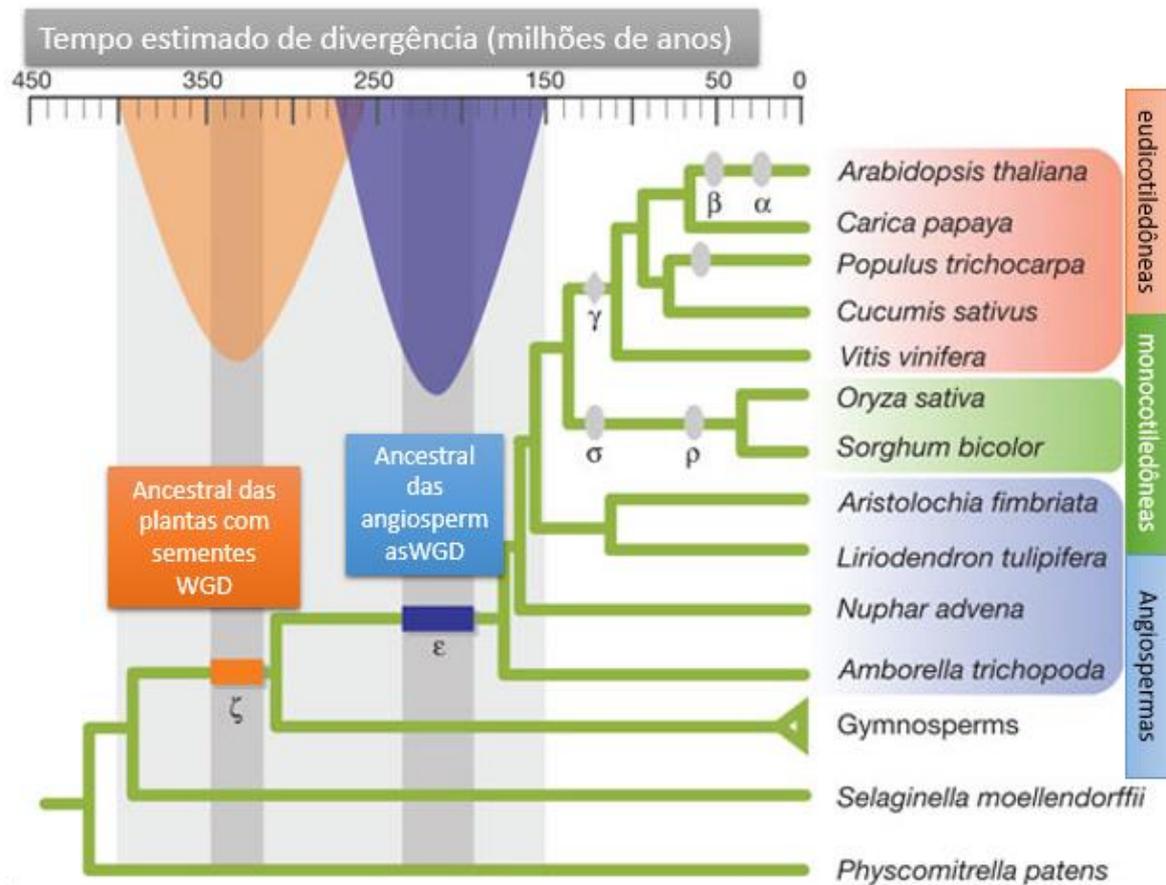
1.5. Filogenia – História evolutiva da família Poacea

A filogenia trata-se do estudo que permite compreender a relação evolutiva entre os diferentes grupos de organismos, com relação ao organismo vegetal seu estudo se torna complexo devido a diversas particularidades existentes no seu genoma, o que torna uma tarefa árdua para interpretar as relações filogenéticas. Assim sendo, é considerado que o grupo das angiospermas corresponde ao maior grupo de plantas terrestres com 800 gêneros com mais de 300.000 espécies vivas (JIAO et al., 2011; BOUCHENAK-KHELLADI et al., 2010). Este grupo é caracterizado pela existência de órgãos

florais que atuam no processo de reprodução. O grupo das angiospermas é dividido em dois subgrupos levando-se em conta o número de cotilédones na plântula: monocotiledôneas (um cotilédone) e eudicotiledôneas (dois cotilédones). Assim, tendo em vista estas diversidades expressas em número de organismos que abrangem as angiospermas, se forcarmos o ponto de vista evolutivo, pode-se concluir que essas plantas apresentam uma grande diversidade de formas e adaptações que lhe conferiram sucesso adaptativo a diversas condições ambientes e que resultou na colonização global.

De forma geral, se analisarmos os genomas dos organismos vegetais pode-se notar que há também uma discrepância com relação a número de cromossomos, tendo uma prevalência de organismo com genoma poliploide, ou seja, genomas com três ou mais conjuntos de cromossomos. Esses conjuntos de cromossomos podem ser provenientes de um cruzamento inter-específico (autopoliploide) ou entre espécies diferentes (intra-específico) que geralmente são estreitamente relacionados (alopoliploides). O estado poliploide pode ser transmitido para a próxima geração (ALBERTIN; MARULLO, 2012). A poliploidia tem sido considerada como um processo importante para moldar a evolução em eucariontes (WENDEL, 2000; OTTO; WHITTON, 2000; VAN DE PEER et al., 2009), já que existem vários organismos poliploides bem descritos tais como o algodão (gênero *Gossypium*) e o trigo (gênero *Triticum*) (OTTO; WHITTON, 2000). Blan e Wolfe (2004) verificaram que nas angiospermas houve um evento que promoveu uma duplicação do genoma por inteiro (WGD, *Whole Genome Duplication*). Este evento provavelmente gerou o efeito de poliploidia (genomas com mais de três conjuntos de cromossomos). Sendo assim, a WGD é uma característica normal no processo de evolução em plantas e vem sendo estimado que pelo menos um WGD tenha ocorrido na ascendência de 70% de todas as angiospermas (BLANC; WOLFE, 2004). Isto pode ser observado na figura 7, onde é possível evidenciar os eventos de duplicação que ocorreram ao longo da evolução das plantas. Tem sido observado que estes eventos têm desempenhado papéis importantes na expansão das angiospermas (BLANC; WOLFE, 2004; TUSKAN et al., 2006; JAILLON et al., 2007; MING et al., 2008; SCHNABLE et al., 2009; SCHMUTZ et al., 2010).

Figura 7- **Eventos poliploides ancestrais em plantas com sementes e angiospermas.** Duas duplicações ancestrais identificadas pela integração de evidências filogenéticas e do relógio de tempo molecular para a evolução da planta na terra. As figuras ovais em cinza indicam as duplicações do genoma geralmente aceitas e identificadas em genomas que foram totalmente sequenciados. A figura em forma de diamante refere-se ao evento de triplicação provavelmente compartilhada por todos os eudicotiledôneas. As barras horizontais indicam regiões que indicam WGD (*whole genome duplication*) ancestral planta semente e WGD ancestral de angiospermas.



.Fonte: Adaptado do artigo JIAO *et al.*, 2011.

Considerando o exposto acima, os eventos de duplicação do genoma como um todo podem resultar em grandes mudanças no padrão de expressão gênica e constituem uma importante fonte de variação adaptativa. Por quase um século foram levantadas algumas hipóteses para explicar o que temos atualmente baseado no processo de WGD (SCHRANZ *et al.*, 2012; JIANG *et al.*, 2013). A partir destes estudos pode-se inferir que o processo de WGD ocorrido nas linhagens de organismos diplóides provavelmente gerou como

consequência um fracionamento do genoma poliplóide, o que provocou rearranjos cromossômicos, inversões de genes, ativação de elementos transponíveis, assim como alterações epigenéticas, estas alterações, provavelmente, tiveram efeito no genoma por interior ou em um conjunto reduzido de pares de genes duplicados (CHEN; NI, 2006; GAETA et al., 2006; TATE et al., 2009; WANG et al., 2010).

Assim, o destino dos pares de genes parálogos (originados pelo o mesmo evento de duplicação de um gene) gerados por um processo de WGD pode ser complexo. Existem alguns casos desses genes terem sido silenciados. Este o processo é denominado como *não-Funcionalização* (nonF), assim como pode ter ocorrido que a função do gene original ter sido dividida entre os genes parálogos originados do processo de duplicação, sendo desta forma denominado este processo como *subFuncionalização* (subF). Além disso, um terceiro exemplo é que esses genes duplicados podem ter adquirido novas funções diferentes daquelas advindas do gene ancestral, neste caso o processo é nomeado como *neoFuncionalização* (neoF) (LYNCH; FORCE, 2000; LYNCH; CONERY, 2000; RODGERS-MELNICK et al., 2012; BLANC; WOLFE, 2004).

A família das gramíneas (*Poaceae*) é formada por cerca de 10.000 espécies e 700 gêneros. Apesar de outras famílias de angiospermas conterem mais espécies e mais gêneros, a família *Poaceae* excede todas as outras famílias em uma característica importante: a dominância ecológica. As gramíneas são encontradas em todas as regiões do mundo e podem se adaptar a habitats tanto de clima temperado quanto de clima tropical (GAUT, 2002). Desta forma, trata-se de um grupo considerado como um dos mais diversificado dentro do grupo das monocotiledôneas e apresenta uma variação no tamanho do genoma entre seus representantes variando assim desde 16,6 pg (*Secale cereale*) até 0,6 pg (*Paspalum notatum*) de DNA por genoma diplóide (BENNETT, 1976). Além disso, essa complexidade pode ser resultante das diferenças no número de WGDs ao longo do processo evolutivo, assim como a presença de altos níveis de sequências repetitivas e elementos transponíveis ativos (TE) (DUBCOVSK; DVORAK, 2007).

O sequenciamento completo e anotação dos genomas de arroz (*Oryza sativa*) (GOFF et al., 2002), milho (SCHNABLE et al., 2009), *Sorghum bicolor* (PATERSON et al., 2009) e *Brachypodium* (*Brachypodium distachyon*) (International Brachypodium Initiative, 2010) permitiu uma análise detalhada da complexidade e diversidade genômica nesta família com grande importância comercial (SCHNABLE et al., 2009). Nos estudos de Luo et al. (2009) foi observado que os mecanismos que regulam a evolução do genoma nas gramíneas podem tanto resultar em uma variação no número de cromossomos como a taxa de mudança evolutiva. Desta forma, em seus estudos foi verificado que o número de mudanças estruturais no genoma do *Sorghum bicolor* foi um pouco maior quando comparada ao genoma de arroz. Entretanto, quando foram analisadas as linhagens da subfamília *Triticeae* verificou que essas mudanças foram bem maiores. Assim, estes autores propuseram que em gramíneas nem todos os genomas se comportaram da mesma maneira ao longo do processo de evolução (Luo et al., 2009). Além disso, as alterações no padrão de transcrição nos genes que sofreram nonF, subF e neoF, decorrentes de um processo de WGD estão susceptíveis a eventos de mutação e/ou mudança nos padrões de epigenética (MEAGHER, 2010). Esta alteração na expressão gênica é considerada uma consequência direta do processo de WGD o que torna mais complexo a análise destes genomas poliploides (DICKINSON et al., 2010).

Com isso, é de extrema importância compreender como o processo de WGD influencia a regulação dos genes que sofreram as duplicações no genoma vegetal tanto para o conhecimento evolutivo como para o melhoramento genético de espécies de interesse.

2. JUSTIFICATIVA.

Considerando que em plantas, o estudo das vias de reparo do DNA tem ocorrido principalmente nos modelos vegetais C_3 - *Arabidopsis thaliana* – planta eudicotiledônea (HAYS, 2002; KIM, 2006; KIMURA; SAKAGUCHI, 2006; LI et al., 2002), e *Oryza sativa* – planta monocotiledônea (KIMURA et al., 2002). É importante ampliar este estudo para outras plantas como a cana-de-açúcar que é uma planta monocotiledônea C_4 .

Em relação à enzima AP endonuclease, alvo deste trabalho, pouco se conhece acerca da sua funcionalidade em plantas, alguns trabalhos se reportam a modelos como a *Arabidopsis* (BABYCHUK et al., 1994). Esses autores clonaram uma sequência homóloga a AP endonuclease denominada AtARP, expressaram e realizaram ensaios funcionais de caracterização desta proteína. Murphy et al. (2009) buscando sequências homólogas no genoma de *Arabidopsis* encontraram mais duas sequências que foram denominadas de Ape1L e Ape2. Entretanto, nesse trabalho, esses autores analisaram apenas o efeito de mutações nesses genes durante o desenvolvimento das sementes.

Estudos *in silico* publicados por Córdoba-Cañero et al. (2009, 2011) utilizando novamente o modelo *Arabidopsis* mostram que em plantas existem as duas vias de reparo BER - via longa e via curta. Apesar de que os dados do genoma não mostrarem nenhuma sequência com homologia a enzima Pol β , os resultados bioquímicos com extrato proteico total mostram que estas duas vias são realizadas em plantas. Deste modo, estes resultados mostram que ainda são escassas informações estruturais e funcionais para a via BER como para sequências homólogas a enzima AP endonuclease e em especial na planta cana-de-açúcar. Assim sendo, se faz necessário a caracterização dos homólogos scARP1 e scARP3 em cana-de-açúcar. Em trabalhos anteriores do nosso grupo de pesquisa foi apresentado dados referentes a uma análise estrutural destas sequências, sendo possível identificar seus domínios conservados que os incluem dentro da família das AP endonucleases e a análise da estrutura tridimensional sugerindo que tais proteínas podem apresentar atividade funcional na célula vegetal (MAIRA, 2011; OLIVEIRA, 2009). Agora

neste trabalho nosso foco é a relação destas enzimas na planta *in vivo* complementando os dados já obtidos anteriormente, bem como tentar entender o processo de duplicação desta sequência no genoma e seu significado evolutivo, além de analisar filogeneticamente outros membros da via de reparo de excisão de bases pertencentes a cana-de-açúcar..

3. OBJETIVO GERAL

Este trabalho tem como objetivo geral estudar estrutural e funcionalmente o papel dos genes Ap endonuclease (ScARP1 e ScARP3), além dos genes de DNA ligase I (ScLig1.1 e ScLig1.2) em plantas de cana-de-açúcar.

3.1 OBJETIVOS ESPECÍFICOS

- Analisar as relações filogenéticas dos homólogos scARP1 e scARP3;
- Realizar a análise estrutural dos genes e proteínas ScARPs;
- Analisar as relações filogenéticas para a DNA ligase I, complementando com os dados obtidos com relação a ScARPs.
- Transformar e obter plantas de *Nicotiana tabacum* com o cassete de super-expressão do gene scARP1 em orientação senso e anti-senso;
- Observar o efeito da expressão do cassete anti-senso no desenvolvimento das plantas transgênicas obtidas, por meio da análise de seu fenótipo;

4. MATERIAIS E MÉTODOS

4.1. Determinação da sequência e análise

Para identificar as sequências homólogas a AP endonuclease de cana-de-açúcar em diferentes organismos foram montadas sequências com alta qualidade (Phred >20) e que foram submetidas a uma análise Blast no NCBI (<http://www.ncbi.nlm.nih.gov>) utilizando BlastX e Tblastn (ALTSCHUL et al., 1997). Em seguida, os domínios funcionais foram anotados utilizando pesquisas contra o domínio conservado do NCBI (CDD), Pfam (<http://www.pfam.edu>), e bases de dados PROSITE (<http://prosite.expasy.ch>). As sequências de ARP de *Homo sapiens* (NP_001231178.1) e Viridiplantae foram obtidas no GeneBank (<http://www.ncbi.nlm.nih.gov>), PDB (www.pdb.org), DFCI (<http://compbio.dfci.harvard.edu> / [TGI](http://tgi.harvard.edu) / plant.html), TAIR (<http://www.arabidopsis.org>) e Phytozome (<http://phyto5.phytozome.net/>). Os dois cDNAs completos de cana-de-açúcar (ScARP1 e ScARP3) foram ainda submetidos ao NETSTART Servidor v.1.0 (PEDERSEN & NIELSEN, 1997) e ORF Finder (NCBI), para realizar previsões da tradução das regiões de sítio de iniciação (TIS) e buscar janela aberta de leitura (*Open Reading frame* -ORF), respectivamente. Para verificar a possível localização sub-celular e identificar as regiões de sinal de endereçamento, essas sequências foram submetidas ao WoLF PSORT Server (HORTON et al., 2007).

Para a caracterização das sequências de cana-de-açúcar (ScARP1 e ScARP3) foram utilizadas as sequências de aminoácidos no formato FASTA da Ap endonuclease humana (APEX1) e a exonuclease III de *Escherichia coli* (EX3) foram retiradas do Uniprot (<http://www.uniprot.org/>), tendo como parâmetro pertencerem a sessão do Swiss-Prot, que contem sequências que foram anotadas manualmente e revisadas, desta forma tais sequências foram alinhadas em as sequências proteicas de cana-de-açúcar, ScARP1 e ScARP3, o alinhamento múltiplo de sequências foi feito por meio do Clustal Ω (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) (SIEVERS et al., 2011), foi

destacado os sítios comuns as quatro sequências por meio das anotações existentes no Uniprot.

As sequências existentes no NCBI referente às enzimas de DNA ligase I (Gene ID: 3978 – *Homo sapiens* - Gene ID: 837333 – *Arabidopsis thaliana*), DNA ligase III (Gene ID: 3980 - *Homo sapiens* – não tem equivalente em *A. thaliana*) e DNA ligase IV (Gene ID: 3981 – *Homo sapiens* - ID: 835822 - *Arabidopsis thaliana*) foram obtidas e depois foram analisadas via a ferramenta Blastn na plataforma de DFCI (<http://compbio.dfci.harvard.edu/tgi/plant.html>) de cana-de-açúcar, com o intuito de buscar sequências homólogas em cana-de-açúcar. Essas sequências foram analisadas via Blastx, ferramenta do NCBI (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) com o intuito de encontrar proteínas similares dentro do banco de dados NR (All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples from WGS projects). Foram então obtidas as sequências ScLig 1 e ScLig 1.2, sendo homólogas a DNA ligase I provenientes de *H. Sapiens* e *A. thaliana*, tendo em vista que não foi encontrado resultado foram analisadas por meio da ferramenta de busca por domínios conservados (CDS) do NCBI (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

4.2. Análise filogenética

As sequências de aminoácidos dos genes ScARP1 e ScARP3, no formato FASTA foram submetidas a análise do programa Blast no NCBI (<http://www.ncbi.nlm.nih.gov>) utilizando BlastX e Tblastn (ALTSCHUL et al., 1997), tendo como alvo a busca de sequências homologas as sequências originadas de cana-de-açúcar, os organismo analisados compreendiam dentro do reino *plantae*. Foram obtidas sequências de 28 Taxa (Tabela 1) que compreendiam organismo desde as algas verdes e plantas vasculares, com esses resultados adquiridos, as sequências foram alinhadas utilizando o programa MAFFT v.6.85 (KATOH; TOH, 2010) ao nível dos aminoácidos com o

L-INS-i e algoritmo CLUSTAL Ω v.1.0.3 (SIEVERS et al., 2011) com os parâmetros de alinhamento personalizados. O modelo de substituição de aminoácido de melhor ajuste foi estimada por meio da utilização da ferramenta ProtTest v.3.2 (DARRIBA et al., 2011) e os dendogramas foram calculados com base em uma análise Bayesiana usando MrBayes 3.1.2 (RONQUIST; HUELSENBECK 2003) e BEAST v.1.7 (DRUMMOND et al., 2012). As inferências foram realizadas com base em 6 funcionamentos independentes, cada uma com três correntes simultâneas com um modelo WAG fixo (WHELAN; GOLDMAN, 2001), utilizando a estimativa do parâmetro de forma gama e proporção dos sítios invariáveis. A cadeia de Markov foi iniciada com uma árvore aleatória e correu para 106 gerações, amostrados a cada 100 gerações. A árvore de consenso foi estimada usando 1.000.000 árvores.

Com relação a sequência de cana-de-açúcar com homologia a DNA ligase I, as sequências de aminoácidos no formato FASTA dos ScLig 1.1 (DFCI - TC128188) e ScLig 1.2 (DFCI - TC150242) foram usadas como consulta dentro do banco de dados de proteínas do Phytozome (<http://www.phytozome.net/search.php>) utilizando a ferramenta Blastp, onde foram obtidas 66 sequências das diferentes espécies. Tais sequências foram então alinhadas utilizando o algoritmo CLUSTAL W. Os dendogramas foram gerados através do programa MEGA 5.1 (TAMURA et al., 2011) utilizando o modelo de Dayhoff com o método estatístico *neighbor-joining*. O teste de filogenia foi baseado com o método *bootstrap* com 500 replicatas, sendo o tratamento dos resultados realizado com *Pairwise deletion*.

Tabela 1 – Identificação de proteínas de outros organismos que compartilham homologia com as sequências ScArp1 e ScArp3 de cana-de-açúcar.

Identificação da Sequência utilizada no Banco de dados	Tamanho da proteína
scARP3_CA065533	273
scARP1_CA291683	325
Oryza_sativa_NP_001044538.1	310
Oryza_sativa_BAD68138.1	307
Oryza_sativa_NP_001044537.1	499
Sorghum_bicolor_XP_002456494.1	325
Zea_mays_NP_001130149.1	323
Zea_mays_ACG32142.1	304
Hordeum_vulgare_BAK03701.1	296
Triticum_aestivum_CJ576969.1	266
Brachypodium_distachyon_XP_003564410.1	298
Brachypodium_distachyon_XP_003564409.1	493
Brachypodium_distachyon_XP_003564408.1	516
Arabidopsis_thaliana_ACB29409.1	538
Arabidopsis_lyrata_XP_002879932.1	550
Arabidopsis_lyrata_XP_002881788.1	430
Vitis_vinifera_XP_002274714.2	479
Glycine_max_XP_003518153.1	507
Ricinus_communis_XP_002530098.1	486
Populus_trichocarpa_XP_002308005.1	331
Populus_trichocarpa_XP_002322649.1	428
Selaginella_moellendorffii_XP_002975564.1	306
Physcomitrella_patens_XP_001770524.1	363
Chlorella_variabilis_EFN56476.1	371
Chlamydomonas_reinhardtii_XP_001703445.1	437
Volvox_carteri_XP_002946858.1	355
Ostreococcus_lucimarinus_XP_001422113.1	273
Ostreococcus_tauri_XP_003078172.1	460
Homo_sapiens_APEX1_SP27695	318

4.3. Análise dos exons e introns das sequências scARP1 e scARP3

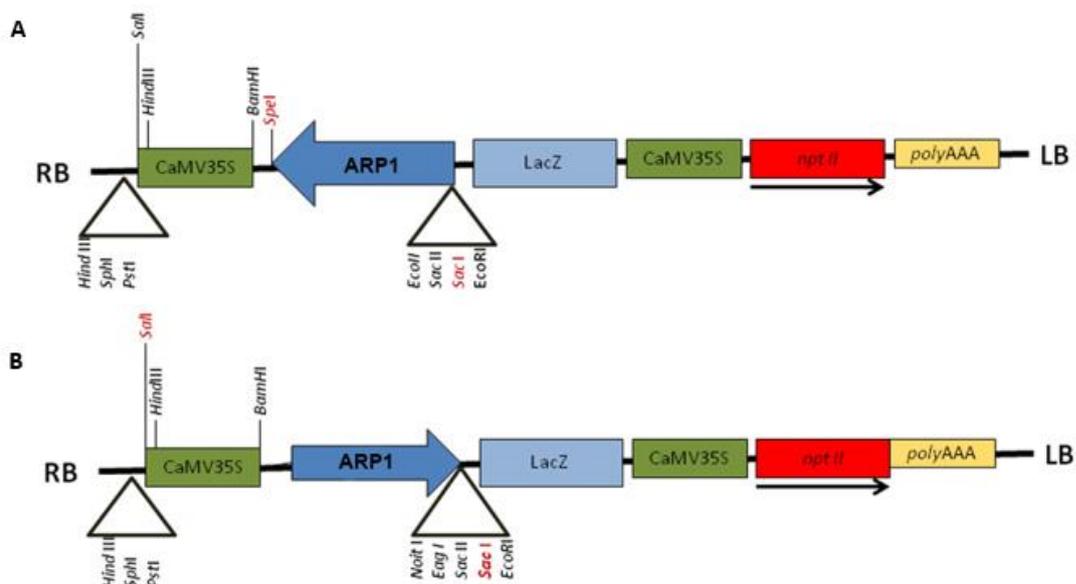
As sequências de cDNA de cana-de-açúcar, ScARP1 e ScARP3, foram analisadas no NCBI (<http://www.ncbi.nlm.nih.gov/>) utilizando a ferramenta Blast (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>), sendo escolhida a opção Blastn, que

procura no banco de dados de nucleotídeos utilizando uma sequência nucleotídica como base de busca, desta forma, procurou-se no banco de dados de sequências de referência de RNAs (refseq-RNA) sequências que apresentariam homologia com as sequências de cana-de-açúcar. Desta forma, teve-se como resultado referente ao ScARP1, o RNAm de uma proteína hipotética de *Sorghum bicolor* (gi:242054696/ref:XM_002456449.1) e, com relação a ScARP3, o RNAm mensageiro também de uma proteína hipotética do mesmo organismo (gi:24205469/ref:XM_002456448.1), com essas sequências de RNAm foi possível descobrir de que parte e em qual cromossomo do *S. bicolor* tais RNAs se originaram, foi encontrado que ambos são oriundos do cromossomo 3 do *S. bicolor* (gi:242060088), na posição 65246632-65250387 pb para o primeiro e 65240591-65246229 pb referente ao segundo RNAm, logo tais seguimentos do cromossomo foram alinhados com as sequências de cDNA das ScARPs e foi verificado as posições que alinhavam e inferido a existência dos exons. O alinhamento foi feito por meio de Clustaw , tendo em vista o alinhamento de códons entre as sequências, a visualização e análise se deu por meio do programa Mega 5.10 (TAMURA et al., 2011).

4.4. Transformação das plantas utilizando os cassetes de superexpressão.

As plantas de *Nicotiana tabacum* foram transformadas por meio da utilização do sistema de *Agrobacterium tumefaciens* ABI carregando o cassete de superexpressão em ambas as orientações - senso e antisenso (figura 8) (AN et al.,1986; BEVAN, 1984).

Figura 8– **Cassetes de superexpressão**. Representação esquemática do cassete de super-expressão em orientação anti-senso (A) e senso (B). A sigla CaMV35S representa o promotor do vírus 35S promotor forte e constitutivo, *nptII* – Gene de resistência ao antibiótico canamicina, ARP1 - sequência de ScARP derivada de cana-de-açúcar, LacZ – gene *lacZ* de *E. coli* que codifica β -galactosidase. polyAAA – cauda poli-adenilada, os triângulos indicam sítios das enzimas de restrição que foram utilizadas para inserir a sequência de interesse.



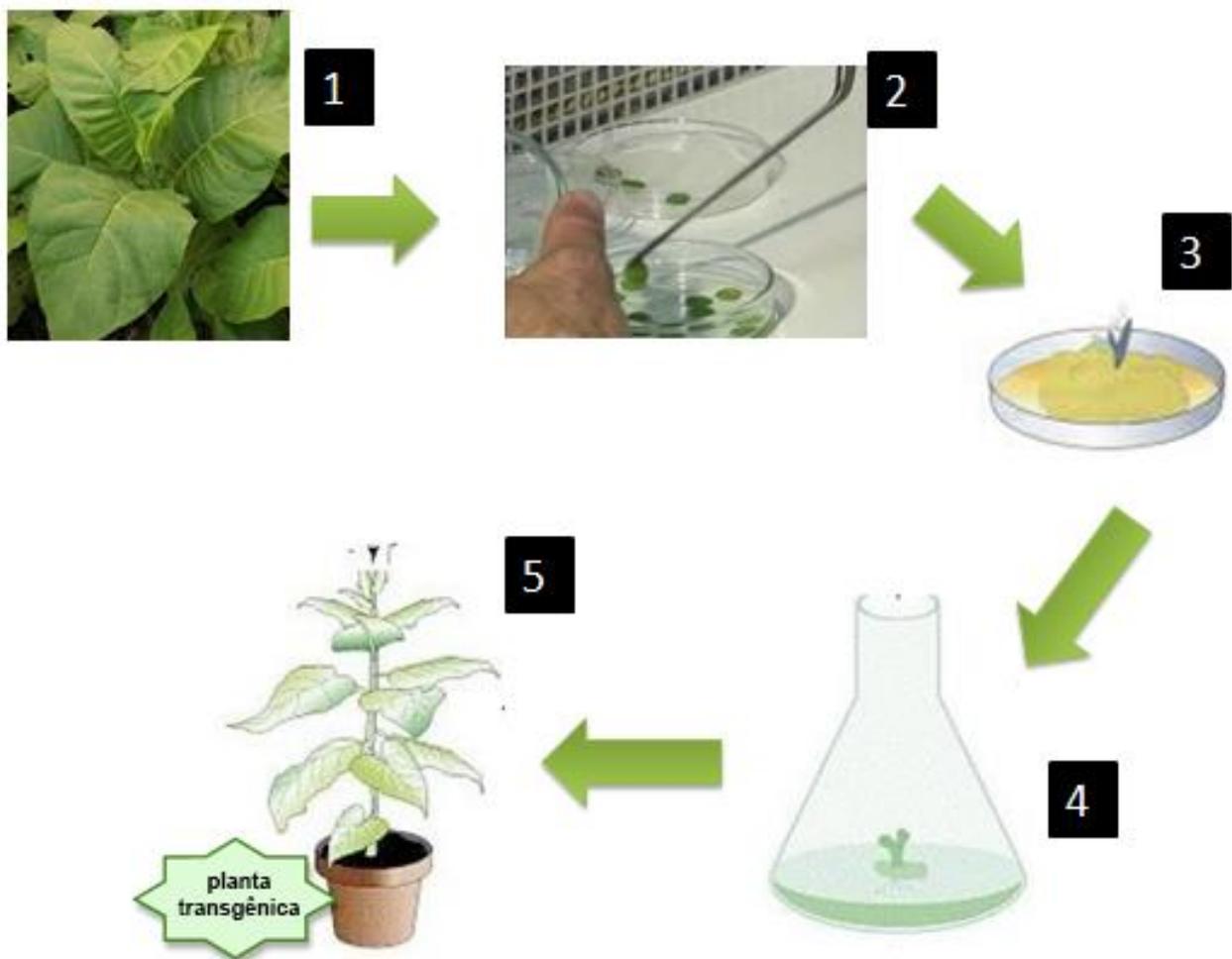
Fonte: Autoria própria.

No protocolo de transformação as bactérias *A. tumefaciens* contendo o cassete de super-expressão (figura 9) foram crescidas em meio líquido Luria-Bertani (LB) acrescido de antibióticos estreptomicina (100 μ g/mL) e rifampicina (100 μ g/mL), em seguida estas foram colocadas para crescer em uma incubadora Tecnal TE420 a 28 °C com agitação a 250 RPM por 48 horas. Decorrido este tempo, a solução contendo as bactérias foi centrifugada por 15 minutos por 5.500 xg e o meio LB foi descartado e as bactérias foram ressuspensas agora em meio Murashige and Skoog Basal Salt mixture (MS) (MURASHIGE; SKOOG, 1962) (4,3 g/L do Sigma 5524), 30 g/L de sacarose em pH 5,8). Todo o meio MS utilizado em todas as etapas foi o MS5524 (Sigma). Esta cultura bacteriana foi colocada agora em contato com os discos foliares das

plantas de tabaco por 5 min (figura 9 item 2). Depois, estes discos foram transferidos para uma placa com meio MS contendo 30 g/L de sacarose, pH 5,8 e permaneceram ali por 48 horas sem a incidência de luz. Decorrido este tempo, estes discos foram agora transferidos para uma nova placa contendo o meio MS com 30 g/L de sacarose em pH 5,8 e acrescidos por canamicina 100 µg/mL, Meropenem 100 µg/mL, benzilaminopurina (BAP) 100 µg/mL e ácido naftaleno acético (NAA ou ANA) 10 µg/mL (figura 9 item 3). Depois de 2 meses, foi observado a formação de calos que foram então transferidos para potes de vidro estéril contendo meio MS acrescidos com os mesmos hormônios anteriores. Quando a plântula foi formada a partir desses calos (gemas) esta foi transferida para outro pote estéril contendo agora somente meio MS com 30 g/L de sacarose, pH 5,8 acrescido com o antibiótico canamicina 100 µg/mL para o desenvolvimento da raiz (figura 9 item 4). Só depois do desenvolvimento do sistema radicular as plântulas foram transferidas para vasos contendo vermiculita e húmus na proporção (1:1) previamente autoclavados (figura 9 item 5).

As plantas foram acompanhadas ao longo do seu desenvolvimento em terra, sendo acrescido fertilizante líquido NPK (Nitrogênio, Fósforo e Potássio), na proporção de nitrogênio 16g/L, fósforo 23 g/L, potássio 31 g/L e enxofre 34,5 g/L, foram utilizados a medida de 15 mL diluídos em 1 litro de água, em intervalos de 15 dias. Os frutos foram recolhidos e as sementes foram retiradas, armazenadas em eppendorfs e mantidas a 4°C.

Figura 9 - **Transformação utilizando *Agrobacterium tumefaciens***. Esquema do processo de transformação por intermédio da *Agrobacterium tumefaciens*, nas plantas de *Nicotianum tabacum*. O número 1 representa a retirada de folhas *N. tabacum*; 2- Corte dos discos foliares que são colocados em contato com uma solução contendo *A. tumefaciens*, 3- Os disco são colocados em meio Ms solido acrescido de hormônios; 4 – Quando as plântulas começam a se formar são transferidas para um recipiente maior; 5- quando atingem certo tamanho são transferidas para a terra.



Fonte: Autoria própria.

4.5. Análise do desenvolvimento das plantas transformadas e coleta de sementes.

As sementes obtidas foram coletadas das plantas T0 (plantas que sofreram o processo de transformação). Estas sementes foram colocadas para germinar em meio MS (descrito no item 4.4) acrescido com o antibiótico canamicina na concentração final de 100 µg/mL para selecionar as sementes que continham o cassete de superexpressão. Quando estas plântulas apresentavam mais ou menos 10 cm de altura, estas foram transferidas agora para a terra (planta T1). Estas plantas cresceram em terra e suas sementes também foram recolhidas (sementes T2). Estas sementes também foram germinadas na presença de meio MS acrescido com o antibiótico canamicina na concentração de 100 µg/mL.

Assim, as plantas potencialmente transgênicas contendo o cassete de super-expressão no sentido anti-senso foram analisadas segundo o seu fenótipo, tendo como parâmetros de observação a comparação com os controles (tabaco selvagem). Neste período foi analisado o tempo de germinação, morfologia da folha e de outros órgãos vegetais, tempo de desenvolvimento, quantidade de sementes produzidas. Tais parâmetros foram analisados em todas as gerações (T0, T1 e T2).

4.6. Germinação das sementes transgênicas

As sementes de *N. tabacum* selvagens e de plantas potencialmente transgênicas foram transferidas para tubos cônicos estéreis de 1,5 mL em condições estéreis em fluxo laminar (cabine de segurança biológica, PA 410, Pachane) e foram desinfestadas com 1,0 mL de uma solução Triton 0,1 %/ Alcool 70% (1:1000). Estas sementes foram agitadas manualmente durante cinco minutos e depois esta solução foi retirada e foi adicionado 1 mL de álcool a 95%. As sementes foram agitadas e esta solução foi novamente descartada. Este

processo foi realizado mais duas vezes para posterior transferência destas sementes para um papel filtro previamente autoclavado. Neste papel, as sementes secaram a temperatura ambiente no fluxo laminar por uns 5-10 min. Em seguida, estas sementes foram transferidas para placas de petri contendo meio Murashige and Skoog (MS) Sigma 5524 sólido (4,3 g/L acrescido com sacarose 30 g/L e agar 8 g/L pH 5,8) e acrescido com o antibiótico de canamicina 100 µg/mL. As placas contendo as sementes desinfestadas foram mantidas a uma temperatura de 4 °C por um período de 48 h. Decorrido este período, estas placas foram transferidas para BOD a 23 °C em fotoperíodo de 12 h (MA 403 Marconi). A taxa de germinação foi calculada como o número de sementes que germinou (Sg) com relação a número total de sementes presente na placa (ST):

$$\text{Sg/ ST} \times 100 = \text{Taxa de germinação em \%}$$

A observação das sementes e o cálculo da taxa de germinação se iniciou quando foi observada a germinação das primeiras sementes (emergência da radícula), a partir deste dia foi observado em intervalos de dois dias a germinação das sementes. Também foi realizada uma análise do fenótipo das sementes por meio da observação do tempo de germinação das sementes no período de 30 dias. Os dados coletados foram registrados em tabelas e devem ser submetidos a análises estatísticas.

4.7. Medição do comprimento das folhas de *Nicotiana tabacum*.

Para a análise da área foliar foi utilizada a sexta folha de plantas de *Nicotiana tabacum* com 5 meses de idade, tanto das plantas transgênicas da geração T1 (linhagens 14, 1 e 2) quanto da planta selvagem (controle).. Tais análises foram realizadas utilizando-se o programa de processamento de imagem e análise no java -IMAGEJ (<http://imagej.nih.gov/ij/index.html>). Para esta análise foram retiradas 5 folhas de cada planta pertencente a cada linhagem (14, 1 e 2) como também da planta selvagem (controle). Depois que o resultado foi obtido foi realizado o cálculo da média aritmética dos valores e desvio padrão r

foi analisado no programa Excel 2013, e os gráficos box-plots foram desenvolvidos pelo programa ioGAS 4.4

4.8. Extração de DNA das folhas de *Nicotiana tabacum*

Para a confirmação da presença do cassete de super-expressão foram inicialmente retirado amostras das folhas das plantas *N. tabacum* potencialmente transgênicas e selvagem. Estas amostras foram armazenadas em tubos cônicos de 1,5 mL no freezer -18 °C. Para o processo de extração de DNA foi adicionado a cada tubo 400 µL do tampão de extração para PCR (Tris pH 7,5 200 mM/ EDTA 25 mM/ NaCl 250 mM/ SDS 0,5%). A amostra foi macerada nesse tubo cônico contendo tampão e com o auxílio de um pistilo. Após este processo, foi adicionado 400 µL de clorofórmio. Este tubo foi homogeneizado manualmente, e em seguida centrifugado (Mikro 200R, Hettich Zentrifugen) por 10 minutos 11.519 xg a 20 °C. Decorrido este tempo, o sobrenadante foi transferido para um novo tubo e a este foi adicionado 400 µL de isopropanol. O tubo foi então misturado e centrifugado por 15 minutos a 11.519 xg a 20 °C. Após este tempo, o precipitado formado foi então lavado com 150 µL de etanol 70% e novamente centrifugado por 10 minutos a 11.519 xg a 20 °C. O sobrenadante foi retirado e o tubo foi colocado na estufa a 50 °C para secar por 2-5 min. Após esta etapa, o precipitado foi ressuspendido em 100 µL de água milliQ previamente autoclavada e estocado a -18 °C até o momento da metodologia de PCR para verificação do cassete de superexpressão.

4.9. PCR

As amostras de DNA obtida no item acima foram submetidas a metodologia de PCR. Primeiramente, foi feito um Mix com os seguintes componentes advindos do kit Gotaq Flexi DNA polymerase da Promega - 1X de

5x Green GoTaq Flexi Buffer, MgCl₂ 2,5 mM, dNTP 0,7 mM, primer *upstream* de canamicina (0,6 pmol), primer *downstream* de canamicina (0,6 pmol) e GoTaq DNA polymerase (2 U). Deste mix foram transferidos 22,5 µL para cada tubo cônico de 0,2 mL e a este foi adicionado 2,5 µL de DNA proveniente da extração de DNA de cada amostra (plantas potencialmente transgênicas e selvagem). Depois, o tubo contendo o Mix + DNA foi transferido para a máquina de PCR (eppendorf Mastercycler) com o programa descrito a seguir:

Tabela 2 – Programa de amplificação por meio da PCR.

Etapa	Temperatura	Tempo	Número de ciclos
Desnaturação inicial	95°C	2 minutos	1 ciclo
Desnaturação	95°C	45 segundos	35 ciclos
Anelamento	42-65 °C	45 segundos	35 ciclos
Extensão	72 °C	1 minuto	35 ciclos
Extensão final	72 °C	5 minutos	1 ciclo
Soak	4 °C	indefinido	1 ciclo

Após a reação de amplificação, as amostras foram separadas em um gel de agarose 1 % (em tampão TAE 1X) em corrente de 50 mA e 60 V por 1 hora e meia. As amostras foram acrescidas de Sybr green I strain Lonza (1:100). Este gel foi visualizado por meio do programa *I-pix loccus biotecnologia molecular imaging*.

5.0 Extração de RNA

Foram retirados diferentes tecidos das plantas de cana-de-açúcar (folhas, raízes, flores). Estes foram triturados em nitrogênio (N₂) líquido até obtenção de um pó. Ao término da maceração, as amostras foram colocadas em tubos previamente gelados e foram acrescentados 2 mL de tampão Trizol (Invitrogen). Esse material foi congelado rapidamente em N₂ líquido e descongelado a temperatura ambiente por 5 minutos. Ao término deste intervalo de tempo, foi adicionado 400 µL de clorofórmio, vortexado por 15 segundos e mantido a

temperatura ambiente por 15 minutos. Após esse tempo, estes tubos foram centrifugados durante 15 minutos a 5.000 xg a 4 °C (centrifuga 5804 R eppendorf). O sobrenadante foi transferido para novos tubos, nestes foram acrescentados de 500 µL de isopropanol. O material foi misturado e mantido a temperatura ambiente por 15 minutos, em seguida os tubos foram centrifugado durante 15 minutos a 5.000 xg e a 4 °C. O sobrenadante foi então descartado e o precipitado lavado com 100 µL de etanol a 70% e seco a temperatura ambiente. As amostras foram ressuspensas em 60 µL de H₂O ultra pura com DEPC 0,1 % e o material foi guardado no freezer a – 80 °C para posterior utilização.

As amostras foram quantificadas em espectrofotômetro, obtendo a absorbância a 260 nm e calculando a concentração a partir da seguinte fórmula: [RNA] = D.O. x valor da diluição x correção p/ RNA (40) = µg/mL. Para a verificação da qualidade do RNA, este foi analisado tanto por espectrofotometria como por meio de uma eletroforese em gel de agarose 2 % contendo 0,5 µg/mL de brometo de etídio (SAMBROOK; RUSSEL, 2001).

5.1. Amplificação do cDNA

De modo a produzir o cDNA, foram utilizados 10 µg de RNA total . Primeiramente as amostras (10µg de RNA total) foram tratadas com DNase (Ambion) de acordo com o protocolo do fabricante. A partir do material tratado foi realizado a produção do cDNA utilizando 3 µg de RNA total previamente tratado com DNase. Para esta reação foi utilizado o *High Capacity cDNA Reverse Transcription Kits* (life Technologies) de acordo com as recomendações do fabricante. O produto final foi diluído com a adição de 180 µL H₂O ultra pura autoclavada contendo 0,1% de DEPC.

A qualidade do cDNA foi avaliada através da realização de uma PCR, utilizando *primers* para a seqüência de um gene com expressão constitutiva (EF fator de alongação). Nesta reação foram usados 15 ng do cDNA diluído, 10 µL de Sybergreen (Applied Biosystem) e 0,05 pM de *primer foward*, 0,05 pM do

primer reverse num volume final de 10 μ L. Esta reação foi testada no aparelho Step one (Applied Biosystem).

5.2. qRT-PCR

As reações de PCR em tempo real foram realizadas utilizando um volume final de 20 μ L, para isso, foram utilizando cDNA (15 ng), *primer forward* (0,05 pMol), do *primer reverse* (0,05 pMol) e 10 μ L de Syber-green (Applied Biosystem). A amplificação foi conduzida utilizando o programa padrão em incubações iniciais a 50°C por 2 min, 95°C por 2 min e seguidas de 40 ciclos de 95°C por 15 s e 60°C por 30 s. Os experimentos foram realizados utilizando 3 replicatas biológicas e para cada uma foram realizadas 3 repetições técnicas. Nesses experimentos foram utilizados como controle endógeno primers para os genes Fator de Elongação e o gene Ubiquitina.

Primers utilizados:

scARP1RTF 5' GGTAGTGGATCGCCTATTCACGTT 3';
scARP1RTR 5' AAGATGTGTTATTCGGTGGGATGC 3';
scARP3RTF 5' AGCGATGCTTATTCCTGAGGGTTT 3';
scARP3RTR 5' TGGATATTCCAATCTGCTCAGTGC 3';
EF1 α RTF 5' AAGGCCCGTTATGATGAGATTGTG 3';
EF1 α RTR 5' CAAAACCAGAGATTGGGACGAAAG 3';
UBQRTF 5'-GTCAAGACCCTCACTGGCAAGACT 3';
UBQRTR 5'-GATTCCTCCTTGTCCTGGATCTT 3'.

6. RESULTADOS

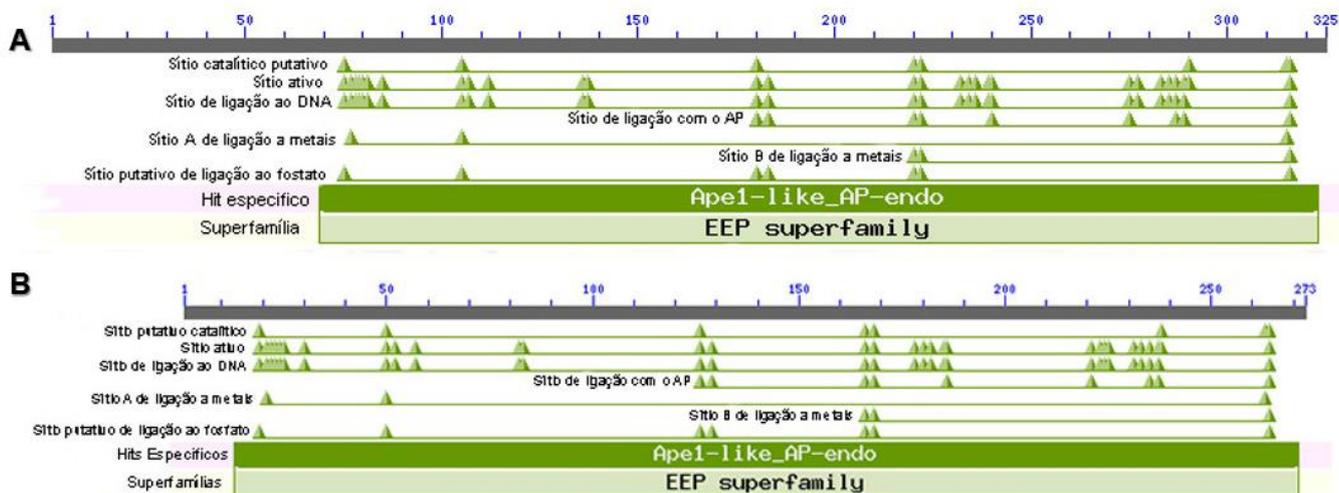
6.1. Análise das sequências ScARP1 e ScARP3.

As sequências alvo de estudo deste trabalho foram resultado de um estudo prévio onde Agnez-Lima et al. (2001) as identificaram no banco de dados SUCEST (<http://www.fapesp.br>). Estas duas sequências são homólogas a proteína AP endonuclease de *Arabidopsis thaliana* - AtARP (At2g4160.2) - e assim foram denominados em ScARP1 e ScARP3, sendo Sc proveniente do gênero *Saccharum* e ARP indicando a homologia a enzima AP endonuclease de *A. thaliana*.

Primeiramente, dispendo das sequências cana-de-açúcar, foram identificados os domínios característicos da superfamília Exonuclease-Endonuclease-Phosphataseb (EEP), e dentro desta superfamília está incluso as subfamília *Escherichia coli* exonuclease III (ExoIII-like_AP-endo) e a subfamília *Human Ape1-like subfamília da família ExoIII purinic/apirimidinic (AP) endonucleases (Ape1-like_AP-endo)*. O domínio EEP inclui todas as enzimas que compartilham um mecanismo comum: a clivagem catalítica de ligações fosfodiéster. Com relação às sequências de ScARPs foi observado que as proteínas estão inclusas nesta superfamília anteriormente descrita (figura 10). Nesta figura se evidencia os domínios conservados anteriormente descritos em ambas as sequências ScARP1 e ScARP3.

Posteriormente, foi realizada a comparação das ScARPS com as sequências da Apex 1 humana e a Exonuclease III (Exo3) de *E. coli*, por meio de um alinhamento demonstrado na figura 11, pode-se notar a conservação de sítios essenciais para a manutenção da atividade da enzimas, tal como o sítio ativo, sítios de ligação com metais, que seriam cofatores enzimáticos, e de sítios de interação com o substrato. Desta forma, esta relação entre todas as sequências, tanto nas sequências de cana-de-açúcar quanto as provenientes de outros organismos (humano e bactéria), fortalece a inclusão das ScARPs dentro do grupo das enzimas AP endonucleases.

Figura 10 - **Domínios conservados das sequências ScARP1 e ScARP3**. Representação das superfamílias a que pertence, além dos sítios: catalíticos putativos, sítios de ligação ao DNA, de ligação AP (purínico e apirimídico, ou seja sítios abásicos), de ligação de metais e putativos de ligação com o fosfato. A – sequência de ScARP1. B – sequência de ScARP3.



Fonte: Autoria própria.

Figura 11 – **Alinhamento das sequências de Ap endonucleases**. Alinhamento de sequências de aminoácidos de organismos diferentes: EX3 de Escherichia coli (P09030), Apex1- Humana (P27695) e ScARPS (1 e 3) de cana-de-açúcar. Destacando-se em azul o sítio ativo, em amarelo o sítio de ligação ao cofator enzimático (magnésio), em cinza um sítio importante para atividade catalítica e em verde o sítio de interação com o substrato da enzima que seria o DNA.

```

sp|P09030|EX3_ECOLI      -----MPK-RGKKGVAEDGDELRTEPEAKKSKTAAK-----NDKEAAGEGPALYE
sp|P27695|APEX1_HUMAN   MSSTSSQRGYPARRGSAANMSS-TSSQRGYPAR-KSEPWTRLTHQERRPQWVAYNPRITMR
scARP1                   -----MR
scARP3

sp|P09030|EX3_ECOLI      -----MKFVSNINGLRARPH----QLEAIVEKHQPDVIGLQETKVVHDDM
sp|P27695|APEX1_HUMAN   DPPDQKTSFSGKPATLKI CSWNVVDGLRAWIKKKGLD---WVKEEAPDILCLQETKCSENK
scARP1                   PP-----PLSTDTSSMKILSWNVNGLQTMV-QSGFSADELVGRNFDVLCLETHLEERN
scARP3                   PP-----PLSKDTRALKIMSWNVNGLKALLKRGFVSQQLAQREDFVLCLETKMQEKD
                          *: *:*:*:*: . . . *:*:*:*:

sp|P09030|EX3_ECOLI      FPLEEVA-KLGYNVF----YHGQKGHYGVALLTKETPIAVRRGFPGDDEEAQRRIIMAEI
sp|P27695|APEX1_HUMAN   LPAEL-QELPGLSHQYNSAPSDKEGYSVGLLSRQCPLKVSYGI GDEEHDQEGRVIVAEF
scARP1                   VDLFRN-LVPEYDITYNSCSVARLGYSGTAVISRVPISVQYGLGIPEDHQRLLITLFE
scARP3                   VEVIKDTLLDGYTNSFFTCVSKLGYSGTAVISRVRKPLSIKYLGLGIPDHDEGRVVIVFEF
                          . *: *:*:*:*: *: *: *: *: *: *:

sp|P09030|EX3_ECOLI      PSLNGVIVINGYFPQGESRDHPKPKAKAQFYQNLQNYLETELKRDNFVLMGDMNISP
sp|P27695|APEX1_HUMAN   DSF---VLVTAIVPNAGRGLVRLEYRQ--RWDEAFRKF-LKGLASRKLPLVCGDLNVAH
scARP1                   DDF---YLVNAYIPNSGRGLRRLNRYR-NEWDPFSDF-IKKLECSKPFIVAGDLNCAH
scARP3                   DDF---YLLTAIVPNSGDGLKRLTYRV-TEWDPSLGNV-MKELEKSKAVILTGLNCAH
                          .: :*:*:*:*: .: .: .: .: .: * :*: *:*:*:

sp|P09030|EX3_ECOLI      TLDIGIGEENRKRRLRTGKCSFLPEEREW-MDRLMSWGLVDFRHRANPQTADRFSWFDY
sp|P27695|APEX1_HUMAN   EEIDLRLNPKGNKK-----NAGFTPQERQGFGEELLQAVPLADSFRHLYPNTPYAYTFWY
scARP1                   ESIDIHNPQAKTE-----AAGFTVEERESFEENLSKGLDITFRKQHPN-AVAYTFWGE
scARP3                   QEIDIHDPAGNRK-----SAGFTNEERESFETHFLSQGFVDITFRKQHRN-VVAYSYWGY
                          .:*: .: .: .: *:*:*:*: .: .: *:*:*:*:

sp|P09030|EX3_ECOLI      RSKGFDDNRGLRIDLLASQPLAECVETGIDYEIRSMKPSDHPVWATFRR
sp|P27695|APEX1_HUMAN   MMNARSKNVGWRLDYFLLSHSLPALCDS----KIRSKALGSDHCPITLYLAL
scARP1                   --NQRITNKGWRLDYFLASESIADKVHDS----YTLPDVSLSDHSPIGLVLKL
scARP3                   RHNARKTNKGWRLDYFLVSESIAEKVHDS----YILPDISASDHSPLGLVVLKL
                          . * * * * * .: .: .: *:*:*:*:

```

Fonte: Autoria própria.

Tendo em vista, a presença dos domínios característicos de uma AP endonucleases, foi analisado então o alinhamento das duas proteínas ScARPs com a proteína AtARP do modelo *Arabidopsis*, da qual elas seriam homólogas. A figura 12 representa o alinhamento obtido com o programa ClustalΩ. Nesse alinhamento podemos observar que não ocorreu um alinhamento destas sequências na porção 5' com a sequência proteica de *Arabidopsis*. Este resultado levanta uma hipótese que a parte que não alinha poderia indicar um possível sinal de endereçamento para organelas.

Para verificarmos esta hipótese, primeiramente foi realizada uma análise dos prováveis sítios de início da transcrição (TIS) com o auxílio do programa NETSTART Servidor v.1.0 e depois analisado com o programa WoLF PSORT Server. Esse programa calcula um *score* (valor de pontuação) para esse provável endereçamento. Os valores que o programa considera como provável são valores acima de 5,0. Assim quanto maiores que 5, maior será a probabilidade do endereçamento para o organela ser “real”. Os resultados obtidos para as duas sequências de cana-de-açúcar estão apresentados na tabela 3. Pode-se verificar que para a sequência ScARP1 o primeiro TIS (sítio de início da transcrição) foi encontrado na posição 92 com um valor de pontuação de 6,0, sugerindo que a proteína putativa pode ter um direcionamento para mitocôndria. O segundo TIS encontrado foi na posição 149 e com uma pontuação de 7,0 que sugere um provável direcionamento da proteína alvo para o núcleo da célula. O terceiro TIS apresentou valores inferiores de 5,0, assim não foram considerados. Sendo o sítio escolhido seria aquele que apresentava o maior *score*, desta forma sendo o segundo TIS (posição 149).

Para a sequência ScARP3, o TIS foi identificado na posição 170 com os valores de 6.5 ou 6.0 de pontuação. Estes valores foram associados com o direcionamento para mitocôndrias ou para cloroplasto/mitocôndrias, respectivamente (Tabela 3).

Figura 12 – Alinhamento das seqüências homólogas a AP endonuclease. Representação do alinhamento obtido entre as seqüências de aminoácidos de ScARP1, ScARP3 e AtARP . A coloração em vermelho mostra a o local onde as seqüências se alinham.

```

CLUSTAL O(1.2.0) multiple sequence alignment

scARP1 -----
ARP      QSSAIYVAKFLVVPLRSLRVGSSFVGVGVGTRSFNKRILMSNATAFSINNSKRKELKIPGA
scARP3   -----

scARP1 -----
ARP      AIDQNC HQMGSDTDRDEMGLQDDRKEIEAMTVQELRSTLRKLGVPVKGRKQELISTLRL
scARP3   -----

scARP1 -----
ARP      HMDSNLPDQKETSSSTRSDSVTIKRKISNREEPTEDECTNSEAYDIEHGEKRVKQSTEKN
scARP3   -----

scARP1 -----
ARP      -----MSSTSSQRGYPARRGSAANMSS---TSSQRGYPARKSEPWTRL
scARP3   LKAKVSAKAIAKEQKSLMRTGKQ--IQSKEETSSTISSELLKTEI ISSPSQSEPWTVL
-----

scARP1 -----
ARP      THQERRPQWVAYNPRIMRPPPLSTDTSSMKILSNVNVNGLQTMV-QSGFSADELVGRNFD
scARP3   AHKKPQKDWKAYNDKIMRPPPLDEGTCVKVMTWNVNGLRGLLKFEFSALQLAQRENFD
-----
*****  *  .:*.::*****:  :  .**  .*.  **:*

scARP1 -----
ARP      VLCLQETHLEERNVDLFLK-NLVPEYDITYWSCSVARLGYSGTAVISRVQPI SVQYGLGIP
scARP3   ILCLQETKLVKDVVEIKKTLIDGYDHSFNCSVSKLGYSGTAIISRKPLSVRYGTGLS
-----
VLCLQETKMQEKDVEVIKDTLLDGYTNSFFTCSVSKLGYSGTAISRVKPLSIKYGLGIP
:*****:  :*:  *  .:  *  :*****:*****:*****:  *

scARP1 -----
ARP      EHDQEGRLITLLEFDDFYLVNAYIPNSGRGLRRLNRYRVNEWDPFSDFIKKLECSKPVIVA
scARP3   GHDTEGRIVTAEFDSFYLINTYVPNSGDGLKRLSYRIEENDRTLNSHIKELEKSKPVVLT
-----
**  ***:  *  **  ***:  :*****  ***:  **  **  :*****  **  ***:

scARP1 -----
ARP      GDLNCARESIDIHNPQAKTEAAGFTVEERESFEENLSSKGLIDTFRKQHPNAVAYTFNGE
scARP3   GDLNCAHEEIDI FNPAGNKR SAGFTIEERQSFGANLLDKGFVDTFRKQHPGVVGYTYWGY
-----
GELNCAHQEIDIHDPAGNRKSAGFTNEERESFETHFLSQGFVDTFRKQHRV VVAYSYWGY
*:*****:  ***:  *  .:  :*****  ***:  .:  :*****  .:  ***:

scARP1 -----
ARP      --NQRI TNKGWRLDYFLASESIADKVHDSYTLDPVSLSDHSPIGLVLKL
scARP3   RHGGRKTNKGWRLDYFLVSQSI AANVHDSYILPDINGS DHCPIGLILKL
-----
RHNARKTNKGWRLDYFLVSESI AEKVHDSYILPDISADHSP LGLVLKL
*  *****  .:  **  :*****  ***:  **  .:  ***:

```

Fonte: Autoria própria.

Tabela 3 - Posição do sitio de inicio da transcrição (TIS) com relação aos scores encontrados das proteínas ScARP1 e ScARP3 com relação as para organela alvo.

Posição pb	Sequência	Núcleo	Cloroplasto e Mitochondria	somente Cloroplasto	somente Mitochondria	Citoplasma ou Mitochondria
92	ScARP1_1ªATG	3.0	5.3	3.5	6.0	3.8
149	ScARP1_2ªATG	7.0	3.0	1.0	0	0
235	ScARP1_3ªATG	2.0	1.0	2.0	0	0
260	ScARP1_4ªATG	3.0	5.0	0	0	0
170	ScARP3_1ªATG	3.0	6.0	6.5	4.5	0

Considerando estes resultados, em cana-de-açúcar há duas sequências homólogas com provável direcionamento dessas proteínas para organelas diferentes quando comparado com o modelo vegetal *Arabidopsis*.

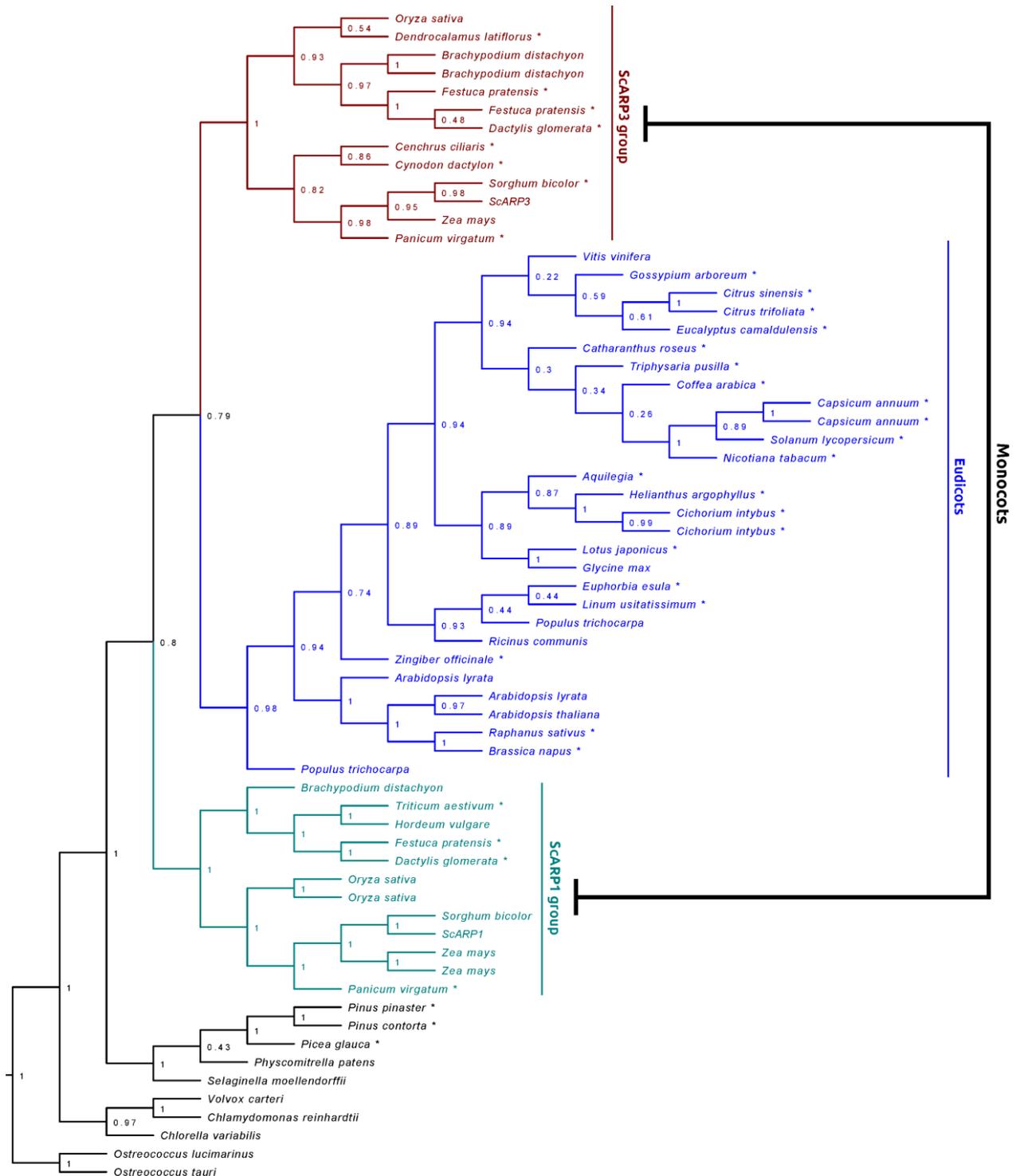
6.2. Análise filogenética das sequências.

Tendo em vista os diferentes endereçamentos celulares obtidos para as duas sequências ScARPs em cana-de-açúcar foi realizada uma análise filogenética, tentando observar se tal fenômeno se reflete em outros organismos.

Assim sendo, foi realizada primeiramente uma pesquisa de similaridade no GenBank obtendo-se 28 sequências proteicas homólogas a ARP pertencentes ao taxon *Viridiplantae* (tabela 1). Nesta análise foram incluído os genes paralogas e ortólogos. Foram obtidas sequências das algas verdes unicelulares *Ostreococcus lucimarinus* e *Ostreococcus Tauri* como grupos externos. O modelo mais adequado para o alinhamento obtido utilizando o modelo de probabilidade WAG foi baseado num método de máxima probabilidade aproximada para estimar uma matriz de substituição de aminoácidos de proteínas globulares (WHELAN; GOLDMAN, 2001).

O dendograma consenso de todas as seis árvores independentes realizadas está representado na figura 13. Nos ramos basais, a árvore da proteína ARP foi semelhante para a filogenia *Viridiplantae*, na qual pode ser observado que os organismos unicelulares, tais como algas verdes e Volvox, Chlamydomonas, além do grupo externo, e Chlorella como um indivíduo de ramo parafiléticos a um clado embriófitos (plantas terrestres). Dentro deste último grupo, dois subgrupos estão claramente identificados, os de plantas terrestres primitivas, representados pela taxa *Physcomitrella* (*Bryophyta*) e *Selaginella* (*Lycopodiophyta*) (figura 13), este clado também incluiu as espécies coníferas representadas por *Picea* e *Pinus* (*Pinophyta*). O outro subgrupo que foi observado inclui as sequências de Angiospermas, separados em dois cladados distintos. Curiosamente, as duas proteínas ScARPs foram agrupados em dois diferentes ramos (figura 13). Pode-se observar que enquanto ScARP1 foi incluída em um ramo em conjunto com outros membros da família *Poaceae* (Monocotiledoneas), a sequência proteica ScARP3 foi agrupada com outros membros da família *Poaceae*, que por sua vez, estariam próximos de um grande clado que incluem tanto os membros de dicotiledôneas e como da família *Poaceae*.

Figura 13- **Dendograma das sequências scARP1 e scARP3.** As sequências foram alinhadas utilizando o programa MAFFT v.6.85 ao nível dos aminoácidos com o L-INS-i e algoritmo CLUSTAL Ω v.1.0.3, com os parâmetros de alinhamento personalizados. Os dendogramas foram calculados com base em uma análise Bayesiana usando MrBayes 3.1.2 e BEAST v.1.7. Esta árvore foi obtida com um consenso com 1.000.000.



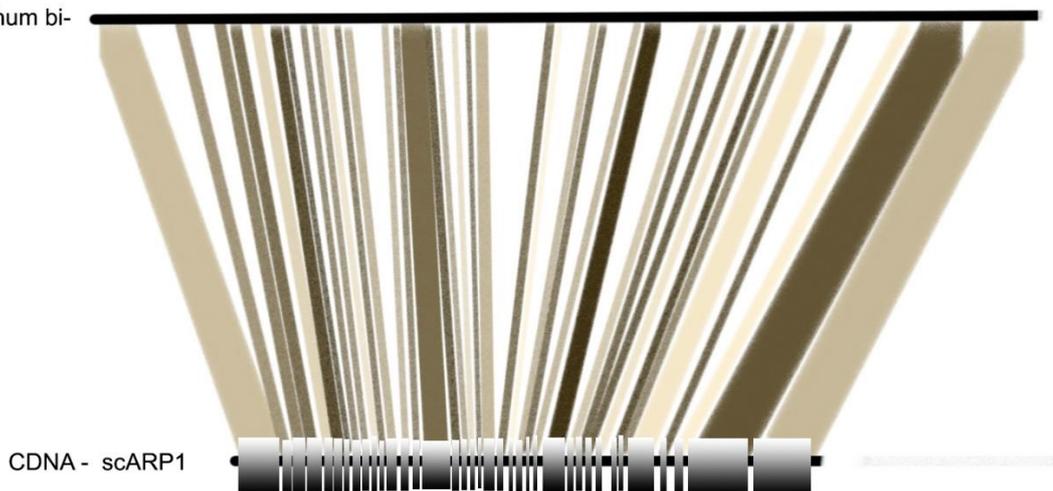
Fonte: Autoria própria.

6.3. Análise de introns e exons das sequências

Considerando que na figura 13 foi observado que as duas sequências de cana-de-açúcar estão próximas filogeneticamente das sequências encontradas em *Sorghum*, então utilizamos essas sequências de modo a identificarmos comparativamente uma possível organização estrutural dos genes em cana-de-açúcar e do grupo em questão. Assim, esta organização estrutural pode ser inferida quando se alinhou primeiramente o cDNA ScARP1 o gene de *S. bicolor* (gi:242060088 - posição 65246632-65250387) o cDNA XM_002456449.1 correspondente. Na figura 14, observamos que estas sequências são formadas provavelmente por 36 exons com tamanhos variando de 47 pb até 415 pb .

Figura 14 – **Análise estrutural do gene ScARP1**. Representação da disposição dos Exons derivados do alinhamento do cDNA de ScARP1 com gene de *Sorghum bicolor* ,(gi|242060088:65246632-65250387).

gi|242060088:65246632-65250387 Sorghum bicolor chromosome 3



Fonte: Autoria própria.

No caso do cDNA ScARP3, foi observado, por meio do alinhamento com o cDNA do RNAm de Sorgum (gi:24205469/ref:XM_002456448.1) e seu gene correspondente (gi:242060088 na posição 65240591-65246229), que poderia se inferir a existência de 32 exons. Estes exons teriam um tamanho variando de 39 pb até 238 pb

Figura 15 – **Análise estrutural do gene ScARP3**. Representação da disposição dos Exons derivados do alinhamento do cDNA de ScARP3 com gene de *Sorghum bicolor* ,(gi|242060088:65240591-65246229).



Fonte: Autoria própria.

6.4. Obtenção das sequências de DNA ligase I em cana-de-açúcar.

Como encontramos para as sequências ScARPs uma duplicação e um novo direcionamento de organelas. Pretendemos também analisar filogeneticamente a via de reparo BER como um todo na planta de cana-de-açúcar. Desta forma, nessa primeira etapa foram escolhidas as DNA ligases que

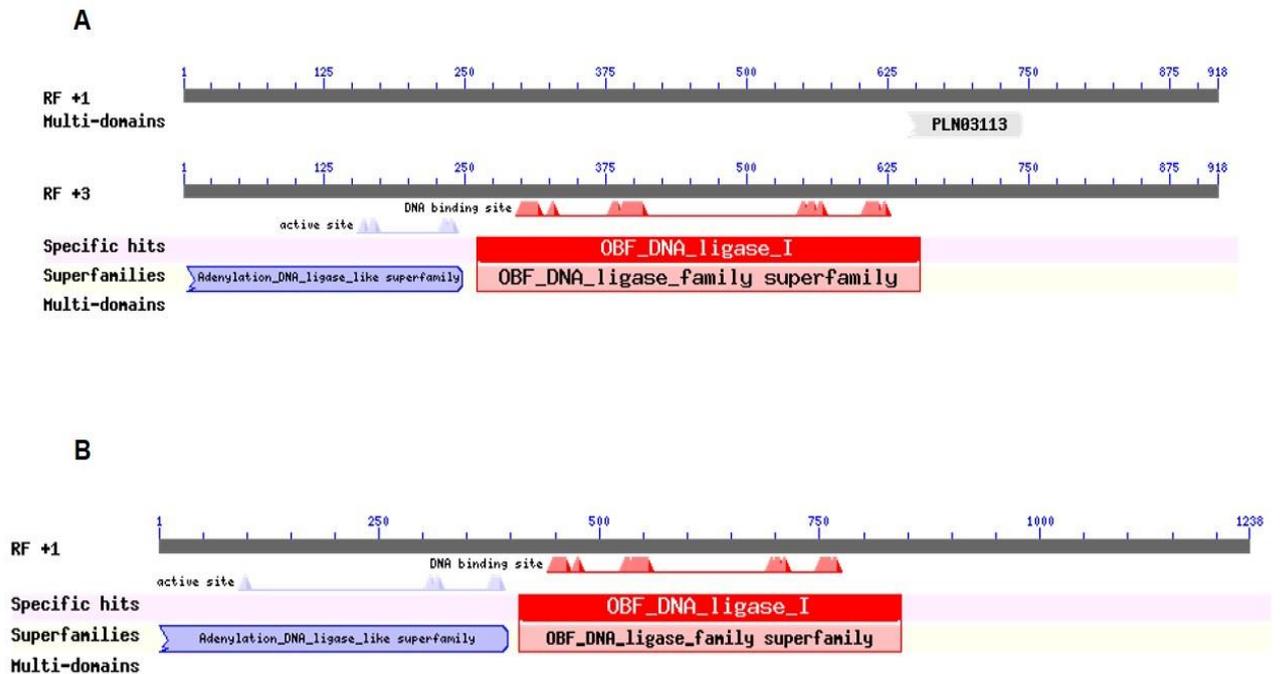
compõem a via BER. Com relação a DNA ligase III e DNA ligase IV não foi encontrado nenhum resultado da presença de sequências homólogas destas enzimas dentro do banco de dados das sequências de cana-de-açúcar no DFCI (<http://compbio.dfci.harvard.edu/tgi/plant.html>). Para a sequência com homologia a DNA ligase I foram identificadas duas sequências homólogas, uma similar a sequência de DNA ligase I humana que foi denominada de ScLig 1.1, e outra similar a sequência de DNA ligase I de *A. thaliana* que foi nomeada de ScLig 1.2.

As sequências de DNA ligase I provenientes de cana-de-açúcar apresentam 916 nucleotídeos, isso referente a ScLig 1.1, enquanto a outra sequência scLig 1.2 apresenta 1238 (figura 16), tais dados podem ser observados na tabela 4. Com relação às proteínas putativas ScLig 1.1 e ScLig 1.2 estas apresentam 214 e 321 aminoácidos respectivamente (figura 17). Analisando estas sequências observa-se nos domínios conservados a presença dos sítios ativos e sítios de ligação com o DNA, além das sequências estarem enquadradas no grupo ATP- dependente DNA ligase I com o domínio de ligação oligonucleotideo/oligossarideo (*OB*)-*fold* (OBF_DNA_ligase_I), sendo um grupo que abrange polinucleotídeos dependentes de ATP e enzimas As enzimas DNA Ligases que catalisam a formação de ligações fosfodiéster, usando substratos de ácidos nucleicos, tais sequências pertencem também a superfamília domínio de adenilação de proteínas similares a DNA ligases dependentes de ATP (Adenylation_DNA_ligase_like) e DNA ligase I com o domínio de ligação oligonucleotideo/oligossarideo (*OB*)-*fold* (OBF_DNA_ligase_family) (figura 16 e 17).

Tabela 4 – Tamanho das sequências de DNA ligase I de cana-de-açúcar (ScLig 1.1 e ScLig 1.2).

Nome	Tamanho (nucleotídeos, proteínas)
ScLig 1.1	916 nucléotídeos 214 aminoácidos
ScLig 1.2	1238 nucleotídeos 321 aminoácidos

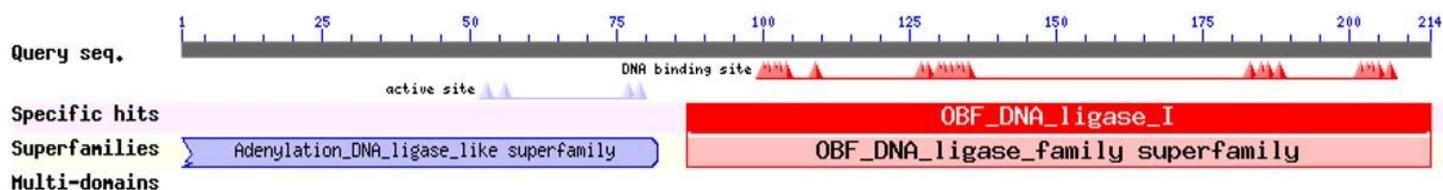
Figura 16. **Domínios conservados das sequências de nucleotídeo ScLig 1.1 e ScLig 1.2.** Representação esquemática dos domínios obtidos para as sequências de nucleotídeo ScLig 1.1 e ScLig 1.2 mostrando as superfamílias a que pertencem. Os triângulos em vermelho representam o sítio de ligação ao DNA já os triângulos em azul representam o sítio ativo. A – sequência nucleotídica de ScLig 1.1. B- sequência nucleotídica de ScLig 1.2.



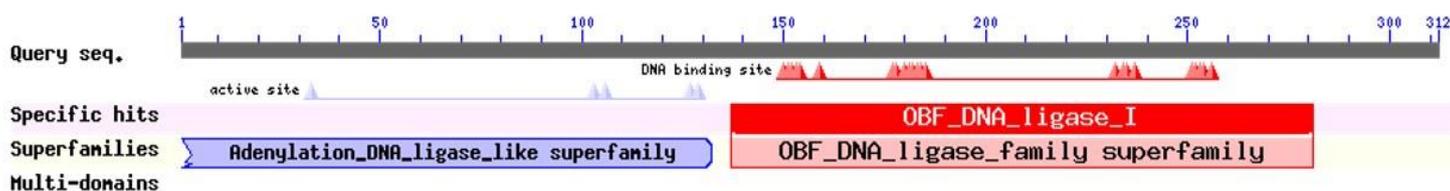
Fonte: Autoria própria.

Figura 17 – **Domínios conservados das sequências de aminoácidos ScLig 1.1 e ScLig 1.2.** Representação esquemática dos domínios conservados obtidos para as sequências de aminoácidos SscLig 1.1 e ScLig 1.2 mostrando as superfamílias a que pertencem. Os triângulos em vermelho representam o sitio de ligação ao DNA já os triângulos em azul representam o sítio ativo. A – sequência proteica de ScLig 1.1. B- sequência proteica de ScLig 1.2.

A



B



Fonte: Autoria própria.

Com o intuito de verificar a conservação entre estas sequências proteicas de cana-de-açúcar foi realizado um alinhamento utilizando o programa ClustalΩ entre as sequências de cana-de-açúcar e a DNA ligase I de *A. thaliana*. Neste alinhamento observa-se que a identidade foi de 78 % com relação a ScLig 1.1 enquanto a ScLig 1.2 apresenta identidade de 38%, quando colocadas para alinhar entre elas (ScLigs) obteve-se a identidade de 41%, tais dados são evidenciados na tabela 7. Assim sendo, com os dados do alinhamento e as análises anteriores, evidencia-se a caracterização destas sequências de cana-de-açúcar como enzimas incluídas dentro do grupo das DNA ligases I.

Tabela 5 – Identidade entre as sequências de DNA ligase I de cana-de-açúcar e a sequência de DNA ligase I de *Arabidopsis thaliana* quando alinhadas entre si.

Sequências	Identidade DNA ligase I (<i>Arabidopsis thaliana</i>)	Identidade ScLig 1.1	Identidade ScLig 1.2
ScLig1.1	78%	-	41%
ScLig1.2	38%	41%	-

6.5. Análise filogenética das sequências de ScLig 1.1 e ScLig 1.2 e localização subcelular.

Com base nos resultados obtidos acima para as sequências ScARP, realizamos a mesma abordagem para as sequências ScLig. Na tabela 6 observamos que a sequência proteica para ScLig 1.1 pode ser direcionada tanto para o cloroplasto como para o citosol, enquanto a sequência ScLig 1.2 se direciona apenas para o núcleo.

Essa análise foi ampliada para as sequências homólogas identificadas nos bancos de dados para outras plantas e que foram utilizadas para as análises filogenéticas. Nessas tabelas observa-se que para as sequências com homologia a ScLig 1.1 são direcionadas ou para núcleo ou para cloroplasto ou mitocôndria (Tabela 7), enquanto que as sequências com homologia a ScLig 1.2 são direcionadas para o núcleo ou núcleo/cloroplasto (Tabela 8).

O dendograma permite observar que a sequência ScLig 1.2 ficou localizada em um grupo específico, na qual ficou denominado de grupo 1 (figura 18), tal grupo é representado por um quadrado cinza, e as sequências que o compõem compartilhavam, em sua maioria, o mesmo endereçamento subcelular da sequência de DNA ligase I de cana-de-açúcar, especificamente a ScLig1.2,

como foi expresso na tabela 8. A sequência ScLig 1.1, por outro lado, ficou localizada em outro grupo diferente que foi nomeado de grupo 2 (figura 18), grupo em destaque em quadrado de coloração verde, tal como observado no grupo 1, as sequências pertencentes ao grupo 2 compartilham endereçamentos subcelulares em comum (tabela 7). Nota-se que ambas as sequências de ScLigs ficaram agrupadas no clado pertencente ao grupo das gramíneas constituído pelos representantes *Setaria italica*, *Oryza sativa*, *Brachypodium distachyon*, *Sorghum bicolor* e *Zea mays*. Contudo, este grupo das gramíneas (incluindo as ScLig) se subdividiu em dois grandes grupos (grupo 1 e 2), o que sugere a possibilidade de duplicações de genes e possíveis alterações no direcionamento da atuação da enzima, como apresentado nas tabelas 7 e 8. No entanto, pode-se verificar que nem todas as sequências enquadradas dentro das gramíneas sofreram duplicação, assim sendo, existem sequências presentes no grupo 2 que não tem representantes no grupo 1, indicando que só existe uma cópia desta sequência, mas essa observação só se deteve a poucas sequências, sendo que a grande maioria apresentou a duplicação, ou seja foi observado sequências correspondentes nos dois grupos. De modo geral esse resultado obtido para as ScLigs foi semelhante ao encontrado para as sequências ScARPs, pois nestas também foi observado a existência de direcionamento subcelular distintos. Contudo tais dados ainda são superficiais para de fato afirmar a existência de uma duplicação do genoma como um todo dentro do grupo das gramíneas, o que se faz necessário aprofundar análises na filogenia destas sequências e de outras a serem adicionadas em trabalhos futuros.

Tabela 6 - Localização subcelular das sequências ScLig 1.1 e ScLig 1.2. nucl.- Núcleo, cyto – Citosol, plas- membrana plasmática, chlo – Cloroplasto, nucl. Plas –Membrana nuclear.

Nome	Localização subcelular
ScLig 1.2	nucl: 9.0, cyto: 2.0
ScLig 1.1	chlo: 5.0, cyto: 5.0,

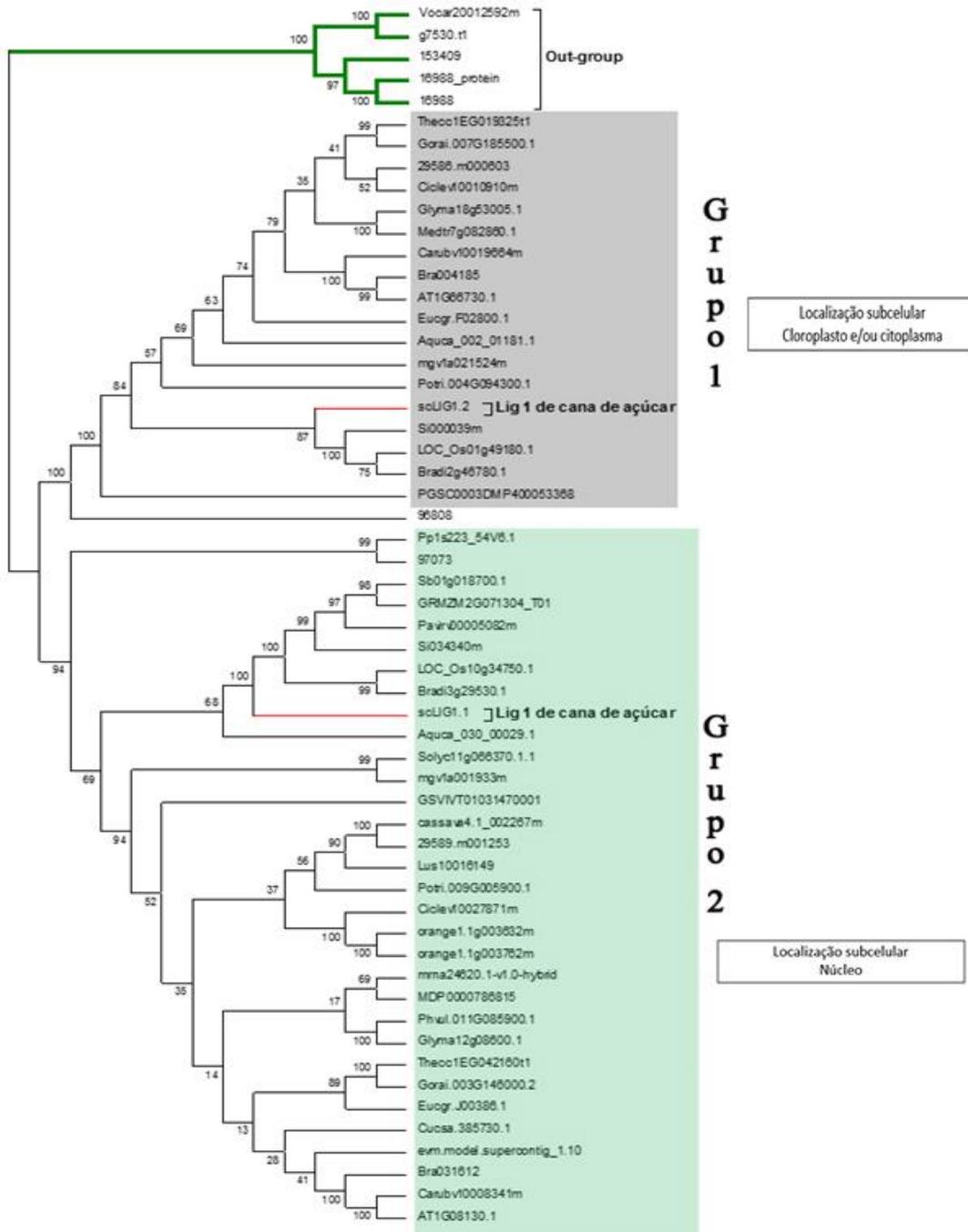
Tabela 7 - Localização subcelular das sequências homologas a scLig1.1. nucl.- Núcleo, cyto – Citosol, plas- membrana plasmática, chlo – Cloroplasto, nucl. plas –Membrana nuclear, vacu – Vacúolo, mito –Mitocôndria, cysk –Citoesqueleto.

Espécie	Número de acesso	Localização subcelular
<i>Physcomitrella patens</i>	Pp1s223_54V6.1	chlo: 6.0, cyto: 5.0,
<i>Selaginella moellendorffii</i>	97073	E.R.: 5.0
<i>Sorghum bicolor</i>	Sb01g018700.1	chlo: 14.0
<i>Zea mays</i>	GRMZM2G071304_T01	chlo: 9.5, chlo_mito: 7.3
<i>Panicum virgatum</i>	Pavirv00005082m	nucl: 7.0, cyto: 5.0
<i>Setaria itálica</i>	Si034340m	cyto: 5.0,
<i>Oryza sativa</i>	LOC_Os10g34750.1	cyto: 7.0, nucl: 5.0
<i>Brachypodium distachyon</i>	Bradi3g29530.1	cyto: 7.0
<i>Aquilegia coerulea</i>	Aquca_030_00029.1	nucl: 6.0, cyto: 5.0,
<i>Solanum lycopersicum</i>	Solyc11g066370.1.1	chlo: 13.5, chlo_mito: 7.5
<i>Mimulus guttatus</i>	mgv1a001933m	nucl: 5.0
<i>Vitis vinífera</i>	GSVIVT01031470001	cyto: 7.0, chlo: 6.5,
<i>Manihot esculenta</i>	cassava4.1_002267m	chlo: 9.5, chlo_mito: 5.5,
<i>Ricinus communis</i>	29589.m001253	chlo: 4.5
<i>Linum usitatissimum</i>	Lus10016149	chlo: 11.5, chlo_mito: 6.5,
<i>Populus trichocarpa</i>	Potri.009G005900.1	nucl: 6.0
<i>Citrus clementina</i>	Ciclev10027871m	chlo: 9.5, chlo_mito: 6.0
<i>Citrus sinensis</i>	orange1.1g003762m orange1.1g003632m	chlo: 13.5, chlo_mito: 7.5 chlo: 12.5, chlo_mito: 7.0
<i>Fragaria vesca</i>	mrna24620.1-v1.0-hybrid	mito: 4.0
<i>Malus domestica</i>	MDP0000786815	cyto: 7.0, cysk: 3.0
<i>Theobroma cacao</i>	Thecc1EG042160t1	chlo: 11.5, chlo_mito: 7.0
<i>Gossypium raimondii</i>	Gorai.003G146000.2	chlo: 12.5, chlo_mito: 7.3
<i>Eucalyptus grandis</i>	Eucgr.J00386.1	chlo: 12.0
<i>Cucumis sativus</i>	Cucsa.385730.1	chlo: 11.5, chlo_mito: 7.0
<i>Carica papaya</i>	evm.model.supercontig_1.10	cyto: 5.0
<i>Brassica rapa</i>	Bra031612	chlo: 10.5, chlo_mito: 6.0
<i>Capsella rubella</i>	Carubv10008341m	chlo: 8.5, chlo_mito: 6.0
<i>Arabidopsis thaliana</i>	AT1G08130.1	chlo: 8.5, chlo_mito: 6.0
<i>Phaseolus vulgaris</i>	Phvul.011G085900.1	chlo: 12.0
<i>Glycine max</i>	Glyma12g08600.1	chlo: 12.5, chlo_mito: 7.3

Tabela 8 - Localização subcelular das sequências homologas a ScLig 1.2. nucl.- Núcleo, cyto – Citosol, plas- membrana plasmática, chlo – Cloroplasto, nucl. plas –Membrana nuclear, vacu – Vacúolo, mito –Mitocôndria.

Espécie	Número de acesso	Localização subcelular
<i>Setaria itálica</i>	Si000039m	nucl: 10.0,
<i>Oryza sativa</i>	LOC_Os01g49180.1	chlo: 6.0
<i>Brachypodium distachyon</i>	Bradi2g46780.1	nucl: 8.0, chlo: 6.0
<i>Theobroma cacao</i>	Thecc1EG019325t1	nucl: 8.0
<i>Gossypium raimondii</i>	Gorai.007G185500.1	nucl: 11.0, chlo: 3.0
<i>Ricinus communis</i>	29586.m000603	chlo: 10.5, chlo_mito: 7.0
<i>Citrus clementina</i>	Ciclev10010910m	nucl: 12.0
<i>Glycine max</i>	Glyma18g53005.1	nucl: 7.0
<i>Medicago truncatula</i>	Medtr7g082860.1	chlo: 11.0,
<i>Capsella rubella</i>	Carubv10019664m	nucl: 11.5, cyto_nucl: 7.5,
<i>Brassica rapa</i>	Bra004185	nucl: 13.0
<i>Arabidopsis thaliana</i>	AT1G66730.1	nucl: 12.0
<i>Eucalyptus grandis</i>	Eucgr.F02800.1	nucl: 12.0
<i>Aquilegia coerulea</i>	Aquca_002_01181.1	nucl: 9.0
<i>Mimulus guttatus</i>	mgv1a021524m	nucl: 4.5
<i>Solanum tuberosum</i>	PGSC0003DMP400053368	chlo: 4.0, cyto: 4.0
<i>Selaginella moellendorffii</i>	96808	cyto: 9.0

Figura 18- **Dendograma para as sequências com homologia a ScLig 1.1 e ScLig 1.2.** Representação da relação filogenética entre as sequências ScLig 1.1 e ScLig 1.2, com relação a outros organismos vegetais, evidência a formação de dois grupos destacados na figura, além disso as caixas de texto informa a localização subcelular que as sequências de cada grupo compartilham em comum. A árvore foi gerado pelo programa MEGA 5.1 utilizando o modelo de Dayhoff , com o método estático *neighbor-joining* com o teste de filogenia baseado com o método bootstrap com 500 replicatas, sendo o tratamento dos resultados *Pairwise* deletivo

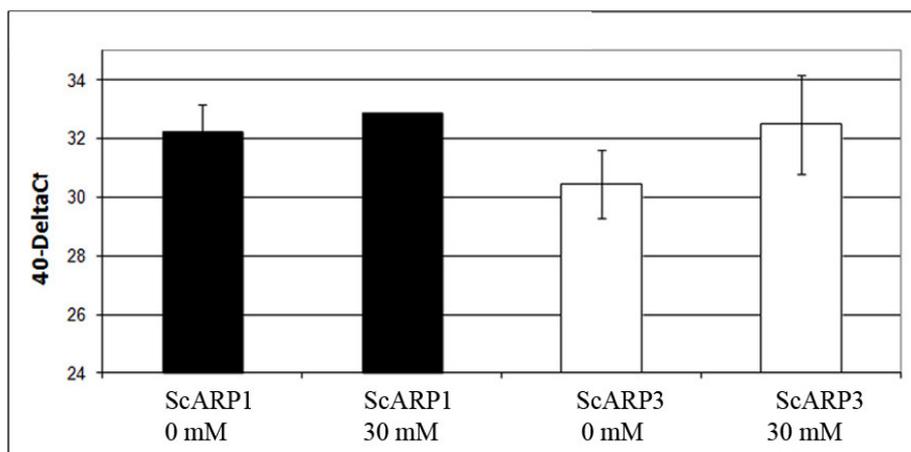


Fonte: Autoria própria.

6.6. Expressão dos cDNAs

No trabalho de Lima et al. (2009) foi verificado que o padrão de expressão para ambos os cDNAs, ScARP1 e ScARP3, foram os mesmos para os diferentes tecidos analisados – folha, raiz, ápice meristemático e flor. Entretanto, o nível de expressão entre os dois genes foi diferente, sendo maior para ScARP3. Considerando esta diferença no nível de expressão e que as sequências ScARPs são uma provável duplicação com um provável direcionamento para organelas diferentes (tabela 3). Neste trabalho foi analisado o padrão de expressão destas duas sequências frente a uma condição de estresse oxidativo. Considerando que a via BER é responsável pela detecção e correção de lesões oxidativas, para essa análise foi utilizado então plantas de cana-de-açúcar submetidas anteriormente a 0 mM e 30 mM de H₂O₂ por oito horas. O tratamento das plantas de cana-de-açúcar foi realizado por Barreto (2013). A figura 19 representa o resultado de expressão para esses dois cDNAs nas condições de 0 mM (controle) e 30 mM de H₂O₂. Nesse material pode-se observar que existe uma tendência para a expressão diferencial do cDNA ScARP3 na condição de estresse oxidativo (figura 19). Entretanto, a expressão para o cDNA ScARP1 manteve-se a mesma na condição controle e e na presença do agente estressor.

Figura 19 - **Expressão do gene ScARP em condições de estresse oxidativo.** Representação gráfica da expressão de cDNA de scARP1 e scARP3 em tecido foliar quando submetidas a uma condição de estresse oxidativo com H₂O₂



Fonte: Adaptado de Maira et al., 2013.

7. DISCUSSÃO

7.1. ScARPs e sua caracterização.

AGNEZ-LIMA et al. (2001) identificaram duas sequências homólogas a AP endonuclease de *Arabidopsis thaliana* no genoma de cana-de-açúcar, estas foram denominadas de: ScARP1 e ScARP3. Tendo como base estes resultados prévios, este trabalho explorou tais sequências homólogas buscando obter informações com relação a sua funcionalidade, caracterizando-as, e também com relação à filogenia.

Considerando a possibilidade de estas sequências serem fruto de uma duplicação e posterior divergência, evidenciado nos resultados de endereçamento subcelular e dados de níveis de expressão gênica, com base nos informações obtidas neste trabalho, seria possível amplificar o escasso conhecimento em relação ao reparo de DNA em plantas e caracterização de um enzima chave para a via de reparo de excisão de bases.

Primeiramente, foi realizado a caracterização *in silico* das sequências de ScARPs. Neste verificou-se a manutenção dos domínios conservados que são caracterizados como importantes para a atividade da enzima AP endonuclease (Mol et al., 2000). Quando colocado as sequências de cana-de-açúcar junto com outras sequências da enzima AP endonucleases em outras espécies vegetais em um alinhamento, pode-se observar que houve uma conservação significativa com o domínio APE1. Tal domínio é referente a enzima AP endonuclease, além deste foi verificado a conservação de outros motivos também caracterizados como importantes para a atividade enzimática, sítios de ligação ao DNA e os sítios catalíticos (Mol et al., 2000). Esses dados reforçam que as sequências ScARP1 e ScARP3 possam ser funcionais. Entretanto, no modelo *Arabidopsis* até o presente momento apenas a sequência AtARP foi caracterizada como tendo atividade AP endonuclease e com os domínios de atividade redox Jun e

Fos (BABYCHUK et al., 1994). Murphy et al. (2009) identificaram no genoma do modelo *Arabidopsis* mais dois tipos de AP endonucleases homólogos a enzima presente em bactéria, humano e levedura: *AtAPE1L*, *AtAPE2*, tais genes apresentam estruturas diferentes se comparada a *AtARP* previamente descrita. Os trabalhos realizados com mutantes para os genes *AtAPE1L* e *AtAPE2* foram associados somente com o desenvolvimento embrionário (MURPHY et al., 2009), e apenas *AtARP* apresenta a função de reparo de DNA (BABYCHUK et al., 1994). Desta forma, torna-se evidente que os dados em cana-de-açúcar com relação a enzima AP endonuclease atuariam de modo a suplementar as informações presente na literatura atual que se aplica, praticamente, ao organismo modelo *Arabidopsis*.

7.2. ScARPs e a filogenia.

Os resultados obtidos com a análise filogenética mostraram que as sequências de cana-de-açúcar pertencem a clados parafiléticos diferentes, contudo essa observação não se detém tão somente as ScARPs, pois outras sequências com homologia a AP endonuclease pertencentes ao táxon das *Poaceae* também se organizaram da mesma maneira. Estes dados permitem observar que houve uma duplicação do gene ARP precocemente na linhagem das gramíneas, provavelmente sendo fruto de uma duplicação do genoma por inteiro (WGD). Sabe-se que a duplicação de genes e os processos de WGDs são fatos bastantes presentes na filogenia das plantas (TANG et al., 2010; JIAO et al., 2011; WANG et al., 2011), sendo associado provavelmente a inovações adaptativas e diversificação dos grupos vegetais (SCHRANZ et al., 2012). Tendo-se em vista a família das gramíneas que apresenta mais de 10.000 espécies distribuídas globalmente e são à base de grande parte da dieta humana, essas inovações adaptativas que foram mencionadas anteriormente estariam diretamente relacionadas com essa distribuição abrangente e sucesso adaptativos de seus representantes. Além disso, infere-se que um evento de WGD, denominado de *rho*, na base das *Poaceae* (SALSE et al., 2008, PATERSON et al., 2004) tal evidência é suportada pelo dados de duplicação

encontrados para alguns genes dessa família (GAUT et al., 1999), entre estes genes, podemos citar os pares de genes duplicados que contribuíram para traços-chave como a espiguetta floral (PRESTON et al., 2009) e as sementes de endosperma com amido (WU et al., 2008).

Em uma análise mais detalhada da árvore filogenética pode-se também notar que a proteína ScARP3 está mais próxima das ARPs de eudicotiledonas por apresentar similaridade a elas, por outro lado ScARP1 não apresenta tal semelhança isso seria devido, provavelmente, a modificações na sequência, e provavelmente a um novo direcionamento de organelas (KRENEK et al., 2013; YANG et al., 2000). Os processos de WGDs estão frequentemente, seguidos por diversificação de genes, uma vez que a duplicação deles podem desempenhar um papel importante na evolução, como podem, por exemplo, conferir adaptações novas, tais como sub- ou neofuncionalização subjacentes a evolução de várias funções ocasionando, com isso, a diversificando dos genes. Por outro lado, o processo de duplicação também pode ser fortemente prejudicial (CONRAD; ANTONARAKIS, 2007; MAKINO; MCLYSAGHT, 2010) e em humanos têm sido associados com doenças como a de Parkinson (SINGLETON et al., 2003) e câncer (SEEGER et al., 1985; DAVOLI; DE LANGE, 2011). Com relação as plantas, alguns pesquisadores acreditam que a poliploidia poderia estar relacionado com a redução do potencial evolutivo (*apud* WAGNER, 1970; *apud* GRANT, 1981; *apud* STEBBINS, 1950; BARKER; ARRIGO, 2012; COMAI, 2000) tendo assim uma visão discrepante com relação a esse evento de duplicação do genoma, eles propuseram que a maioria dos poliploide desapareceriam rápido devido ao fato de eles compartilharem atributos em comum com o ancestral diploide, sendo assim, eles sofreriam exclusão competitiva por terem mais de um organismo com as mesmas características compartilhando o mesmo nicho (ARRIGO; BARKER, 2012). Assim sendo, do ponto de vista genético, os poliploides sofreriam maiores dificuldades evolutivas em consequência da ineficiência da seleção quando alelos se apresentam em múltiplas copias (OTTO; WHITTON, 2000). E poderiam apresentar perdas por autoincompatibilidade e potencialmente baixos tamanhos populacionais, pois estariam relacionados com a dificuldade de gerar descendentes férteis (ARRIGO; BARKER, 2012; BROWNFIELD; KÖHLER, 2011; RAMSEY; SCHEMSKE, 2002).

Contudo, sabe que o evento de duplicações de todo o genoma (WGDs) ocorreram nas linhagens de plantas (BOWERS et al., 2003; PATERSON; BOWERS; CHAPMAN; 2004; JAILLON et al., 2007), animais (JAILLON et al., 2004;) e fungos (KELLIS; BIRREN; LANDER; 2004), com consequências tiveram a origem de inovações evolutiva (ZHANG ; COHN, 2008) e que renunciando a diversificação das espécies (SCANNELL et al., 2006).

Tendo em vista que as ScARPs são, de fato, genes duplicados, a retenção de um gene duplicado pode ser associada com uma espécie de proteção contra mutação deletéria ou com a aquisição de novas funções, e estas podem ser do tipo neofuncionalização ou subfuncionalização (PRÍNCIPE; PICKETT, 2002). No entanto, a inovação funcional pode ser conseguida não só pela evolução das sequências de codificação, mas também por modificações nas regiões reguladoras do gene, conduzindo a uma alteração da rede reguladora (DE SMET; VAN DE PERR, 2012). Desta forma, há um consenso geral de que mudanças evolutivas importantes podem acontecer principalmente no nível de regulação gênica (PRÍNCIPE; PICKETT, 2002; DE SMET; VAN DE PERR, 2012). No caso das sequências ScARP, esta mudança no direcionamento das proteínas de scARPs pode estar associado com um processo de inovação que envolve também a regulação, pois o seu direcionamento para o núcleo pode vir a gerar uma alteração no padrão de expressão gênica. Com os dados obtidos foi possível notar que ambas as scARPs tem como alvos potenciais o cloroplasto e a mitocôndria, no entanto, somente a scARP1 pode se direciona para o núcleo da célula. O fato da ScARP3 se direcionar para o cloroplasto e/ou mitocôndria poderia estar relacionado com a tendência de aumento da expressão deste mesmo gene quando exposto ao estresse oxidativo, desta forma tal Ap endonuclease poderia ter uma função de auxiliar ou mesmo suplementar no reparo de DNA que ocorre em tais organelas sendo estas mais propensas a sofrerem danos pois ambas possuem a cadeia transportadora de elétrons em suas membranas, e desta forma produzem comportos como EROS (espécies reativas de oxigênio) que se não forem neutralizados podem vim a causar danos em diversas biomoléculas, inclusive o DNA (FOYER; NOCTOR, 2003) além disso, o estresse oxidativo afeta diretamente tais organelas pois estaria relacionada a esse desequilíbrio da produção e neutralização dos EROS

(LUSHCHAK, 2011), desta forma, pode-se inferir que a existência de duas Apendonucleases seria um mecanismo adaptativo de garantir a integridade do genoma. Tendo em vista que cana-de-açúcar, por ser uma planta C₄, e seu valor comercial está associado com a sua capacidade de acúmulo de biomassa advindo da fotossíntese e que tal processo de metabólico gera também EROS (FOYER; SHIGEOKA, 2011), desta forma o entendimento de como a planta mantém a integridade do seu genoma, beneficiando a produtividade, se faz necessário. Entretanto, estes resultados e inferências precisam ser melhor explorados para confirmar tais sobreposição aqui discutidas.

Com base na conservação de proteína com duplo alvo para alguns genes parálogos em *Arabidopsis* e *O. sativa*, Morgante et al. (2009) propuseram a hipótese de que o duplo direcionamento poderia ter implicações importantes na evolução, pois trata-se de um processo comum e pode expandir as funções de uma determinada proteína na célula, bem como pode favorecer a manutenção do gene duplicado. Tal fato pode ser evidenciado com relação as ScARPs, pois ambas expressam de forma diferenciada e com direcionamento diferente o que pode estar relacionado com expansão de função dentro da célula vegetal da cana-de-açúcar. Xu et al. (2013) observaram que a duplicação surgiu precocemente na evolução das plantas e os seus dados mostraram que é um mecanismo conservado para alguns genes de *Physcomitrella*, *Arabidopsis* e arroz. Enfatizando essa observação, neste estudo 11 dos 16 genes analisados que apresentaram duplicação demonstraram esta conservação, apesar de uma duplicação do gene seguido por neofuncionalização que foi observada em arroz. Estes resultados sugerem que essa nova segmentação pode acompanhar os eventos de duplicação de genes. Um caso semelhante também foi observado no gene denominado de *FERTILIZAÇÃO ENDOSPERMA INDEPENDENTES (FIE)* (LUO et al., 2009). Sabe-se que *Arabidopsis* tem um gene único *FIE*, todavia na cultura do milho, sorgo e arroz, duas sequências do gene *FIE* foram encontrados, o que indicaria uma duplicação. Os padrões de expressão dos dois transcritos de milho são bastante diferentes, e um deles, *ZmFIE1*, é específico do endosperma, sendo inferido que esta subfuncionalização provavelmente surgiu após a divergência linhagem grama, já que os padrões de expressão dos genes de *FIE* atuais são similares aos de arroz (DICKINSON et al., 2012).

Somado a estes dados relacionados com as AP endonucleases em cana-de-açúcar, outro alvo deste trabalho para as relações filogenéticas foi a enzima DNA ligases I. Esta enzima está presente na via de excisão de bases, bem como também se apresenta em outros processos celulares, tal como em outras vias de reparo e na via de replicação do DNA. As DNA ligases desempenham um papel essencial na manutenção da integridade genômica juntando as quebras na ligação fosfodiéster do DNA que ocorrem durante a replicação e a recombinação, sendo esta ação uma consequência do dano ao DNA e seu posterior reparo. Três genes humanos, *LIG1*, *LIG3* e *LIG4*, codificam DNA-ligases dependentes de ATP. Estas enzimas apresentam regiões catalíticas que atuam de modo a gerar a mesma reação em três etapas, entretanto apesar de ter um processo catalítico em comum tais enzimas apresentam diferentes domínios que medeiam esta atividade (ELLENBERGER; TOMKINSON, 2008). A DNA-ligase I une os fragmentos de Okazaki, como parte do complexo de replicação do DNA, e os genes *LIG1* são essenciais em seres humanos (*Homo sapiens*), *Saccharomyces cerevisiae* e *Arabidopsis thaliana* (BARNES et al., 1992; TOMKINSON et al., 1992; BABIYCHUK et al., 1998; TAYLOR et al., 1998). A DNA ligase I é uma enzima essencial tanto para o reparo de quebra de fita simples (SSB) em todos os eucariontes como no reparo de quebra de fita dupla (DSB) em plantas e animais (LIANG et al., 2008; WATERWORTH et al., 2009). A proteína *LIG1* também apresenta atividade na mitocôndria em *Saccharomyces cerevisiae* e em *Arabidopsis* (WILLER et al., 1999; SUNDERLAND et al., 2006). No trabalho de Córdoba-Cañero et al. (2011) foi mostrado que a proteína *LIG1* é especificamente necessária para a etapa final da via de excisão de bases (BER). Os dados apresentados neste estudo também indicam que *LIG1* é direcionado para locais de quebra de DNA independentemente de como eles são formados, e que outras ligases de DNA não pode substituir as suas funções para a finalização destas quebras. Portanto, é provável que a maioria dos eventos BER em células de *Arabidopsis* são finalizados pela proteína *LIG 1* (Córdoba-Cañero et al., 2011).

No caso do modelo cana-de-açúcar foi buscado os genes homólogos das três *LIGs* anteriormente mencionados, porém somente foi encontradas sequências referentes a *LIG1* denominadas como *ScLig 1.1* e *ScLig 1.2*. Estes resultados

indicam uma possível duplicação para este gene, tal como foi observado anteriormente para as sequências ScARPs. Quando analisado o direcionamento subcelular também foi observado duplo endereçamento. As análises filogenéticas mostram que as sequências se organizam em dois grupos distintos e isso não se detém unicamente as DNAs ligases de cana-de-açúcar, pois se observa a duplicação e a separação de genes também ocorrem nas sequências outras plantas analisadas. Tais evidências corroboram com o que foi observado para com as sequências ScARPs suportando assim a hipótese do processo de WGD na linhagem das gramíneas. Tais observações geram novas perguntas: Será que esses eventos de duplicação também ocorrem em outros genes referentes a via de reparo por excisão de bases? E nas outras vias de reparo? Análises posteriores sobre tais genes são necessárias para compreender melhor esse processo em cana-de-açúcar.

Assim, nossos resultados, sugerem que os dois homólogos ScARP, bem como os ScLIGs, em cana-de-açúcar provavelmente surgiram a partir de um processo de WGD na linhagem das gramíneas e, como previsto pela modelagem estrutural visto em trabalhos anteriores, são, teoricamente funcionais. Além disso, as diferenças nas regiões reguladoras e os níveis de expressão do transcrito pode indicar um processo subfuncionalização. No entanto, mais estudos envolvendo testes de complementação e atividade da proteína *in vitro* serão de grande importância para a compreensão do papel da ScARPs parálogos em células de cana-de-açúcar.

7.3. Plantas transgênicas e sementes

As plantas submetidas a transformação exibiram um fenótipo alterado se comparado as plantas controle não transformadas, deve-se levar em conta que o cassete no sentido anti-senso em tais plantas deveriam resultar em um silenciamento parcial ou total do gene ARP1 (dados não mostrados). Tal fenótipo também foi observado em um trabalho anterior do grupo (Dissertação de Andrea Lima de Oliveira, 2009). No trabalho de Oliveira (2009) foi obtido apenas um

exemplar da planta transgênica hemizigota, e observou-se que esta planta também tinha alterações na folha, tal como a presença de folhas lanceoladas, crescimento lento e abortamento das flores. Nossos resultados mostram que estes fenótipos estão se mantendo também na descendência T1 e T2. Fortalecendo que provavelmente este padrão modificado possa ser devido a presença da construção em orientação anti-senso. As plantas obtidas, até o momento, são da geração T0 e T1, com o genótipo heterozigoto e homozigoto para o cassete de super-expressão no sentido anti-senso, isto é, estas plantas estão com o cassete de super-expressão se segregando. Os dados com a germinação sugerem que as linhagens obtidas tenham o padrão de segregação de acordo com as leis de Mendel para uma sequência, isto é 1:2:1. Este padrão de segregação fortalece a hipótese que estas linhagens apresentem apenas uma cópia do cassete de super-expressão em seu genoma. As plantas da geração T0 apresentaram crescimento mais horizontal do que vertical, de modo que as plantas são baixas, folhas com colorações diferenciadas bem como com enrugamento. Li et al. (2010) trabalhando com plantas transgênicas resistentes a estresse salino observaram também folhas enrugadas. Os resultados obtidos por esse grupo pode ser associado com as alterações que foram observadas por nós, considerando que em condições de estresse salino ocorre também um aumento das EROs, alterando assim o equilíbrio redox, e nessas condições são necessárias diferentes enzimas para reduzir esses efeitos e provavelmente para corrigir as lesões no DNA que possa ter surgido nesse desbalanço redox. Outro exemplo de modificações foi o encontrado por Handayani et al. (2007). Este grupo produziu plantas transgênicas com um oncogene sendo este derivado do próprio T-DNA e também verificaram alterações morfológicas nas folhas, alterações no tamanho da planta. Nesse estudo, estas modificações poderiam ser relacionadas também com uma condição de estresse que a planta teria sido submetida e que alterou o seu desenvolvimento.

As modificações observadas tanto no desenvolvimento como na morfologia podem ser associadas a uma possível instabilidade provocada no genoma causada pelo fato de silenciar um gene pertencente a via de reparo de BER, tendo em vista que a principal via para proteger o DNA de danos endógeno ou exógeno é o reparo por excisão de bases (BER) (CÓRDOBA-CAÑERO et al.,

2009). Desta forma, a ausência do gene poderia afetar tanto o desenvolvimento vegetal como também proveria as morfologias anormais das folhas observados até o momento. Sabendo-se que as propriedades químicas das enzimas que compõem a via de BER, na qual, em sua maioria, preferem substratos compostos por fitas duplas de DNA, acredita-se que tal via possa atuar na fase G1 do ciclo celular, durante esta fase a atividade de BER manteria a transcrição sem erros e prepararia o DNA para a replicação através da remoção de danos ao genoma. No entanto, se o dano a uma base do DNA não for removido antes da iniciação da replicação do DNA, a integridade do genoma poderia ser assegurada por um sistema de reparo chamado Síntese de translesão do DNA- (TLS) que envolve a ação de DNA Polimerase, cuja função seria de executar a síntese de DNA ao longo de um ampla gama de lesões da base (DIANOV; HURBCHER, 2013). A via de reparo BER é fortemente apoiada pelo mecanismo de TLS (PARSON; ELDER, 2003). No entanto, quebras de fita simples ou SSBs (*single strand breaks* – quebras de fita simples) se não reparados pela via de reparo BER tem o potencial de atingir a forquilha de replicação do DNA e com isso, gerar quebras de fita dupla (DSB- *double stream break*) (SOBOL et al., 1996). Ocorrendo este tipo de lesão, então será necessário um reparo envolvem quebras de fita dupla a ser reparado pela a via de reparo por recombinação não homologa (NHEJ) ou recombinação homóloga (HR). Contudo, isso não é o suficiente para substituir a via de reparo BER, tal como foi observado em ratos deficientes ou em Pol β , DNA-ligase III ou XRCC1, enzimas estas envolvidas na reparação dos SSBs e que resultou em letalidade precoce dos embriões (SOBOL et al., 1996; GAO et al., 2011; TEBBS et al., 2003). Nossos resultados com as plantas transgênicas mostraram que as plantas não morreram, isto é, não apresentaram a letalidade como fenótipo, apenas um desenvolvimento lento e em algumas linhagens foi observado um aborto nas flores. Este fenótipo pode ser associado com uma redundância gênica, ou seja, a presença de outros genes que poderiam realizar funções similares e que ainda não tenham sido identificados e caracterizados. Essa redundância gênica pode assim garantir a manutenção de processos essenciais para a viabilidade do organismo. Além disso, deve-se enfatizar que podem existir genes específicos para plantas ainda não identificados e caracterizados por não terem homólogos em outros organismos, tal com humano ou bactérias que normalmente servem como base

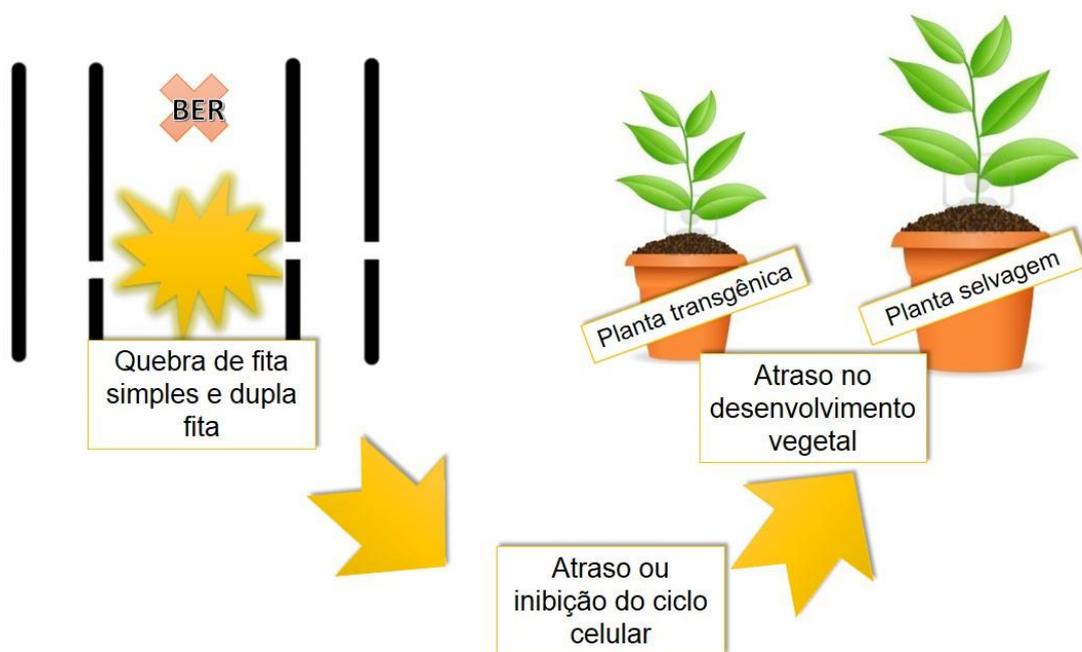
para busca de homólogos em genomas vegetais. Um exemplo é a presença de uma terceira DNA ligase em *Arabidopsis*, chamada LIG6. Esta proteína trata-se de uma enzima específica de plantas para a família das ligases identificada e sequências de DNA putativos (BONATTO et al., 2005). De forma geral, danos ao DNA normalmente estão relacionados com o bloqueio da transcrição, o que conferia o aumento da senescência ou apoptose, o que resultaria, por sua vez, em um aceleração do envelhecimento do organismo induzidos pelo o aumento dos danos no DNA (MITCHELL et al., 2003; AKBARI; KROKAN, 2008; SINCLAIR; OBERDOERFFER, 2009). No entanto, o que foi observado nas plantas originadas da transformação com o cassete em orientação anti-senso até agora é o processo inverso, isto é, um atraso no seu desenvolvimento, o que pode vir a sustentar a hipótese anteriormente citada, mas o processo pelo qual se explica isso ainda não pode ser inferido unicamente dispondo de plantas homozigotas.

Assim sendo, somado a tudo que foi dito anteriormente, muito pouco se sabe sobre o mecanismo bioquímico de reparo de DNA em plantas. Com o que foi analisado até o momento pode-se concluir a partir dos dados que a ausência parcial ou total da enzima AP endonuclease, a via de reparo de DNA por BER não pode ser desempenhada com sucesso, pois o sitio abásico não pode ser retirado de modo que pode gerar mutações ou mesmo quebra de fitas simples ou duplas do DNA durante os processos de replicação e transcrição. Tais danos poderiam levar a um bloqueio ou atraso no ciclo celular o que poderia explicar o atraso no desenvolvimento vegetal das plantas transgênicas tanto da geração T0 e T1 (figura 32). Todavia, análises futuras devem ser feitas com plantas homozigotas para essas construções de modo que essa hipótese possa ser testada e partir disso aprofundar as análises bioquímicas.

Contudo, com as observações até agora feitas com relação às plantas transgênicas pode-se notar que tal proteína afeta o desenvolvimento da planta como um todo de modo a causar fenótipos bem evidentes, porém é importante obter resultados com as plantas homozigotas e com as plantas contendo somente o T-DNA. Além disso, será importante obtermos resultados da atividade enzimática, substrato e condições de atividade, como a real localização e expressão gênica desta sequência como de outras sequências associadas a via

de reparo BER para compreendermos a função deste gene e da proteína em plantas.

Figura 20 – **Hipótese proposta.** Representação hipótese referente ao desfalque na via de excisão de base (BER), devido ao silenciamento da enzima AP endonuclease, deste modo a falta desta via de defesa pode gerar quebras de fita simples e dupla o que leva um atraso ou inibição do ciclo celular o que pode resultar em um atraso no desenvolvimento vegetal apresentado nas plantas transgênicas.



Fonte: Autoria própria.

8.CONCLUSÃO

Até o momento, pouco se sabia sobre a filogenia em cana-de-açúcar, assim como para o processo de reparo por excisão de bases deste organismo. Desta forma, as conclusões deste trabalho são:

- Existe a possibilidade que a duplicação do genoma como um todo (WGD) poderia não ter só ocorrido na cana-de-açúcar, como foi a proposto inicial, mas envolveria também a família das *Poaceae*, sendo observados nos dados envolvendo as enzimas ScARPs e ScLigs, contudo se faz necessário análise de outros genes para suportar tal hipótese;
- Fenótipo de organismos *in vivo* transformados com cassete de super-expressão na orientação anti-senso sugerem que ocorreu uma redução ou inativação da enzima ScARP permitindo observar alterações no desenvolvimento;
- O alvo principal deste trabalho foi a enzima AP endonuclease que é importante para a Via BER. A caracterização se deu por meio da análise da presença dos domínios conservados que permitiram observar a conservação e uma provável função da enzima. Contudo também foi observado variação na expressão do RNAm, o que pode sugerir uma alteração na regulação dos genes alvo. Além disso, esse trabalho corrobora com a amplificação do conhecimento nesta área de reparo em plantas e em especial no modelo de cana-de-açúcar.

9. PERSPECTIVAS

Para trabalhos futuros, se faz necessário ampliar os resultados até agora obtidos com esse trabalho atual. Desta forma planeja-se continuar a análise do desenvolvimento das plantas homizigotas a serem obtidas, análise do direcionamento destas proteínas por meio da construção transducional com um gene repórter, expressão de proteínas e ensaios bioquímicos. Complementando com análise de expressão diferencial dos genes (DNA glicosilases e AP endonuclease) em condições de estresse abiótico. Deste modo será possível concluir as observações que foram obtidas neste trabalho e dados anteriores do nosso grupo de pesquisa.

REFERENCIAS

AGNEZ-LIMA, L. F et al. Base excision repair in sugarcane. **Genetic and Molecular Biology**, [S.l.], v. 24, p.123-129, 2001.

AKBARI, M., KROKAN, H. E. Cytotoxicity and mutagenicity of endogenous DNA base lesions as potential cause of human aging. **Mech. Age. Dev.**, [S.l.], v. 129, p.353-365, 2008.

ALBERTIN, W.; MARULLO, P. Polyploidy in fungi: evolution after whole-genome duplication. **Proc Biol Sci.**, [S.l.], v. 279, n.1738, p. 2497-2509, 2012.

ALTSCHUL, S. F. et al. a new generation of protein database search programs. **Nucleic Acids Research**, [S.l.], v. 25, p. 3389-3402, 1997.

OLIVEIRA, A. L. de. **Caracterização de um cDNA homólogo a uma AP endonuclease em cana-de-açúcar**. 2005-2007.Dissertação (Mestrado em Genética e Biologia Molecular) - Universidade Federal do Rio Grande do Norte, Natal, 2007.

APE, K. L.; HIRT, H. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. **Annu. Rev. Plant Biol.** [S.l.], v. 55, p. 373-399, 2004.

ARRIGO, N.; BARKER, M. S. Rarely successful polyploids and their legacy in plant genomes **Current Opinion in Plant Biology**, [S.l.], v. 15, n. 2, p. 140-146, 2012.

ATKINSON, N. J.; URWIN, P. E.The interaction of plant biotic and abiotic stresses: from genes to the field. **J. Exp. Bot.**, [S.l.], v. 63, p. 3523-3543, 2012.

EWING, B., GREEN P. Base-calling of automated sequencer traces using phred, II. Error probabilities. **Genome Research**, [S.l.], v. 8, p. 186-194, 1998.

BABIYCHUK, E. et al. The *Arabidopsis thaliana* Apurinic endonuclease Arp reduces human transcription factors Fos and Jun. **Proceedings National Academy Science Letters**, [S.l.], v. 91, p. 3299-3303, 1994.

BALESTRAZZI A. et al. Genotoxic stress and DNA repair in plants: emerging functions and tools for improving crop productivity. **Plant Cell Rep.**, [S.I.], v. 30, n. 3, p. 287-95, 2011.

BANITT, I.; WOLFSON, H. J. ParaDock: a flexible non-specific DNA–rigid protein docking algorithm. **Nucleic Acids Research**, [S.I.], v. 39, p.135, 2011.

BARNES, D. E.; LINDAHL, T. Repair and genetic consequences of endogenous DNA base damage in mammalian cells. **Annu. Rev. Genet.** [S.I.], v. 38, p. 445-476, 2004.

BARONDEAU, D. P. et al. (Nickel superoxide dismutase structure and mechanism. **Biochemistry**, [S.I.], v. 43, p. 8038-8047, 2004.

BAUTE, J.; DEPICKER, A. Base excision repair and its role in maintaining genome stability. **Crit Rev Biochem Mol Biol**, [S.I.], v. 43, n. 4, p. 239-76, 2008.

BEERNINK, P. T. et al. Twodivalent metal ions in the active site of a new crystal form of human apurinic/apyrimidinic endonuclease, Ape1. Implications for the catalytic mechanism. **J. Mol. Biol.** [S.I.], v. 307, p. 1023-1034, 2001.

BENNETT, M. D. DNA amount, latitude, and crop plant distribution. **Env. Exp. Bot.**, [S.I.], v. 16, p. 93-108, 1976.

BEVAN, M. Binary Agrobacterium vectors for plant transformation. **Nucleic Acids Res.**, [S.I.], v. 12, p. 8711-8721, 1984.

BLANC, K. G.; WOLFE, K. H. Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. **Plant Cell**, [S.I.], v. 16, p. 1679-1691, 2004.

_____. Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. **Plant Cell**, [S.I.], v. 16, p. 1667-1678, 2004.

BONATTO, D.; BRENDEL, M.; HENRIQUES, J. A. P. A new group of plant-specific ATP-dependent DNA ligases identified by protein phylogeny, hydrophobic cluster analysis and 3-dimensional modelling. **Funct. Plant Biol.**, [S.I.], v. 32, p. 161-174, 2005.

BOWERS, J. E. et al. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. **Nature**, [S.I.], v. 422, p. 433-438, 2003.

BRANZEI, D.; FOIANI, M. Maintaining genome stability at the replication fork. **Nat. Rev. Mol. Cell. Biol.**, [S.I.], v. 1, n. 3, p. 208-219, 2010.

BRAY, E. A.; BAILEY-SERRES, J.; WERETILNYK, E. Responses to abiotic stresses. **Biochemistry and Molecular Biology of Plants**. Edited by GRUISSEM, W., BUCHANNAN, B. AND JONES, R. pp. 1158– 1249. American Society of Plant Physiologists, Rockville, MD. Celebi-Toprak, 2000.

BRITT, A. B. Molecular genetics of DNA repair in higher plants Damage. **Trends in plant Science**, [S.I.], v. 4, n. 1, p. 1360-1385, 1999.

BRITT, A. Repair of damaged bases. **The Arabidopsis Book**. Disponível em: <http://www.arabidopsis.org/publications/arabidopsis>. 2002. Acesso em: 20 de setembro 2013. DOI:10.119/tab.005.

BROWNFIELD, L.; KÖHLER, C. Unreduced gamete formation in plants: mechanisms and prospects. **J Exp Bot**, [S.I.], v. 62, p. 1659-1668, 2011.

CAPPELLI, E. et al. Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. **J Biol Chem**, [S.I.], v. 272, p. 23970-23975, 1997.

CARRIE, C.; GIRAUD, E.; WHELAN, J. Protein transport in organelles: dual targeting of proteins to mitochondria and chloroplasts. **FEBS Journal**, [S.I.], v. 276, p. 1187-1195, 2009.

CASTAGNETTO, J. M. et al. The Metalloprotein Database and Browser at The Scripps Research Institute. **Nucleic Acids Res.**, [S.I.], v. 30, p. 379-382, 2002.

CHEAVEGATTI-GIANOTTO, A. et al. Sugarcane (*Saccharum X officinarum*): a Reference Study for the Regulation of Genetically Modified Cultivars in Brazil. **Tropical Plant Biol.**, [S.I.], v. 4, p. 62-89, 2011.

CHEN, W.; ZHU, T. Networks of transcription factors with roles in environmental stress response. **Trend in Plant Science**, [S.I.], v. 9, p. 591-596, 2004.

CHEN, Z. et al. Structure and control of the actin regulatory WAVE complex. **Nature**, [S.I.], v. 468, p. 533-538, 2010.

CHEN, Z. J.; NI, Z. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. **Bioessays**, [S.I.], v. 28, p. 240-252, 2006.

CHENG, M. et al. Genetic transformation of wheat mediated by *Agrobacterium tumefaciens*. **Plant Physiol.**, [S.I.], v. 115, p. 971-980, 1997.

CHOU, K. M., CHENG, Y. C. An exonucleolytic activity of human apurinic/apyrimidinic endonuclease on 3' mispaired DNA. **Nature**, [S.I.], v. 415, p. 655-659, 2002.

CICCIA, A.; ELLEDGE, S. J. The DNA damage response. Making it safe to play with knives. **Mol. Cell**, [S.I.], v. 40, p. 179-204, 2010.

COMAI, L. The advantages and disadvantages of being polyploid. **Nature Reviews**, [S.I.], v. 6, p. 836-846, 2005.

CONRAD B.; ANTONARAKIS, S. E Gene duplication: A drive for phenotypic diversity and cause of human disease. **Annu Rev Genomics Hum Genet**, [S.I.], v. 8, p.17-35, 2007.

CÓRDOBA-CAÑERO, D. et al. Single-nucleotide and long-patch base excision repair of DNA damage in plants. **Plant Journal**, [S.I.], v. 60, p. 716-728, 2009.

CÓRDOBA-CAÑERO, D.; ROLDÁN-ARJONA, T.; ARIZA, R. R. Arabidopsis ARP endonuclease functions in a branched base excision DNA repair pathway completed by LIG1. **Plant J.** [S.I.], v. 68, n. 4, p. 693-702, 2011.

CULLIGAN, K. M. et al. ATR and ATM play both distinct and additive roles in response to ionizing radiation. **Plant J.** [S.I.], v. 48, p. 947-961, 2006.

CUNNINGHAM, R. P. et al. Endonuclease IV (nfo) mutant of *Escherichia coli*. **J. Bacteriol.**, [S.I.], v. 168, p. 1120-1127, 1986.

CVETKOVIC, A. et al. Microbial metalloproteomes are largely uncharacterized. **Nature**, [S.I.], v. 466, p. 779-782, 2010.

CZECHOWSKI, T. et al. Real-time RT-PCR profiling of over 1400 Arabidopsis transcription factors: Unprecedented sensitivity reveals novel root and shoot-specific genes. **Plant Journal**, [S.I.], v. 38, p. 366-379, 2004.

D'HONT, A. et al. Characterisation of the double genome structure of modern sugarcane cultivars (*Saccharum* spp) by molecular cytogenetics. **Mol Gen Genet**, [S.I.], v. 250, p. 405-413, 1996.

DARRIBA, D. et al. ProtTest 3: fast selection of best-fit models of protein evolution. **Bioinformatics**, [S.I.], v. 27, p. 1164-1165, 2011.

DAVOLI, T.; DE LANGE, T. The Causes and Consequences of Polyploidy in Normal Development and Cancer. **Annual Review of Cell and Developmental Biology**, [S.I.], v. 27, p. 585-610, 2001.

DE BUCK, S. et al. Agrobacterium tumefaciens transformation and cotransformation frequencies of Arabidopsis thaliana root explants and tobacco protoplasts. **Mol. Plant-Microbe Interact.**, [S.I.], v. 11, p. 449-457, 1998.

DE BUCK, S. et al. The T-DNA integration pattern in Arabidopsis transformants is highly determined by the transformed target cell. **Plant J.**, [S.I.], v. 60, p. 134-145, 2009.

DE NEVE, M. et al. T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. **Plant J.**, [S.I.], v. 11, p. 15-29, 1997.

DE SMET, R.; VAN DE PEER, Y. Redundancy and rewiring of genetic networks following genome-wide duplication events. **Current Opinion in Plant Biology**, [S.I.], v. 15, p. 168-176, 2012.

DEMPLE, B.; HERMAN, T.; CHEN, D. S. Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease. Definition of a family of DNA repair enzymes. **Proc. Natl. Acad. Sci.**, [S.I.], v. 88, p. 11450-11454, 1991.

DEMPLE, B.; JOHNSON, A.; FUNG, D. Exonuclease III and endonuclease IV remove 3' blocks from DNA synthesis primers in H₂O₂-damaged Escherichia coli. **Proc. Natl. Acad. Sci.**, [S.I.], v. 83, p. 7731-7735, 1986.

DETERDING, L. J. et al. Mapping of the 5'-2-deoxyribose-5-phosphate lyase active site in DNA polymerase β by mass spectrometry. **J Biol Chem**, [S.I.], v. 275, p. 10463-10471, 2000.

DEVIC, M. et al. Efficient PCR walking on plant genomic DNA. **Plant Physiology and Biochemistry**, [S.I.], v. 35, p. 331-339, 1997.

D'HONT, A.; GLASZMANN, J. C. Sugarcane genome analysis with molecular markers: a first decade of research. International Society of Sugar Cane. **Technologists**, [S.I.], v. 24, p. 556-559, 2001.

D'HONT A. et al. The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. **Nature**, [S.I.], v. 488, n. 7410, p. 213-217, 2012.

DIANOV, G. L.; HÜBSCHER, U. Mammalian base excision repair: the forgotten archangel. **Nucleic Acids Res.**, [S.I.], v. 41, n. 6, p. 3483-90, 2013.

DICKINSON, H.; COSTA, L.; GUTIERREZ-MARCOS, J. Epigenetic neofunctionalisation and regulatory gene evolution in grasses. **Trends in Plant Science**, [S.I.], v. 17, n. 7, p. 389-394, 2012.

ISHCHENKO, A. A.; YANG, X.; RAMOTAR, D.; SAPARBAEV, M. The 3'→5' exonuclease of Apn1 provides an alternative pathway to repair 7, 8-dihydro-8-oxodeoxyguanosine in *Saccharomyces cerevisiae*. **Mol. Cell Biol.**, [S.I.], v. 25, p. 6380-6390, 2005.

DRUMMOND, A. J.; RAMBAUT, A. Beast: Bayesian evolutionary analysis by sampling trees. **BMC Evolutionary Biology**, [S.I.], v. 7, p. 214, 2007.

DRUMMOND, A. J. et al. Bayesian phylogenetics with BEAUti and the BEAST 1.7. **Molecular Biology and Evolution**, [S.I.], v. 29, p. 1969-1973, 2012.

DUBCOVSKY, J.; DVORAK, J. genome plasticity a key factor in the success of polyploid wheat under domestication. **Science**, [S.I.], v. 316, p. 862-1866, 2007.

ELLENBERGER, T.; TOMKINSON, A. E. Eukaryotic DNA ligases: structural and functional insights. **Annu. Rev. Biochem.**, [S.I.], v. 7, p. 313-338, 2008.

ERZBERGER, J. P.; WILSON, D. M. The role of Mg²⁺ and specific amino acid residues in the catalytic reaction of the major human abasic endonuclease. New insights from EDTA-resistant incision of acyclic abasic site analogs and site-directed mutagenesis. **J. Mol. Biol.**, [S.l.], v. 290, n. 3, p. 447-457, 1999.

ESWAR, N. et al. Comparative protein structure modeling using MODELLER. **Current Protocols in Protein Science**, [S.l.], v. 50, p. 2.9.1–2.9.31, 2007.

EWING, B. et al. Base-calling of automated sequencer traces using phred. I. **Accuracy assessment, Genome Research**, [S.l.], v. 8, p. 175-185, 1998.

FORTINI, P.; DOGLIOTTI, E. Base damage and single-strand break repair: mechanisms and functional significance of short- and long-patch repair subpathways. **DNA Repair**, [S.l.], v. 6, p. 398-409, 2007.

FOSTER, S. S. et al. Cell cycle- and DNA repair pathway-specific effects of apoptosis on tumor suppression. **Proc Natl Acad Sci U S A.**, [S.l.], v. 109, n. 25, p. 9953-8, 2012.

FOYER, C. H.; NOCTOR, G. Redox sensing and signalling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria, **Physiol. Plant.**, [S.l.], p.119355-364, 2003.

FOYER, C. H.; SHIGEOKA, S. Understanding Oxidative Stress and Antioxidant Functions to Enhance Photosynthesis. **Plant Physiology**, [S.l.], v. 155, p. 93-100, 2011.

SILVA, F. L. de. **Caracterização funcional de cDNAs obtidos em biblioteca subtrativa**. 2006-2009. Dissertação (Mestrado em Ciências Biológicas) - Universidade Federal do Rio Grande do Norte, Natal, 2009.

GAETA R. et al. Genomic changes in resynthesized *Brassica napus* and their effect on gene expression and phenotype. **Plant Cell**, [S.l.], v. 28, p. 240-252, 2006.

GALBURT, E. A. et al. A novel endonuclease mechanism directly visualized for I-Pol. **Nat. Struct. Biol.**, [S.l.], v. 6, p.1096-1099, 1999.

GAO, Y. et al. DNA ligase III is critical for mtDNA integrity but not Xrcc1-mediated nuclear DNA repair. **Nature**, [S.l.], v. 471, p. 240-244, 2011.

GAUT, B. S. et al. Patterns of genetic diversification within the Adh gene family in the grasses (Poaceae). **Mol Biol Evol**, [S.I.], v. 16, p. 1086-1097, 1999.

GAUT, B. S. Evolutionary dynamics of grass genomes. **New Phytologist**, [S.I.], v. 154, p. 15-28, 2002.

GEORGIADIS, M. M et al. Evolution of the redox function in mammalian apurinic/aprimidinic endonuclease. **Mutat. Res.**, [S.I.], v. 643, p. 54-63, 2008.

GIGLIA-MARI, G.; ZOTTER, A.; VERMEULEN, W. DNA Damage Response **COLD SPRING HARBOR PERSPECTIVES IN BIOLOGY** , [S.I.], v. 3, n. 1, p.??, 2011.

GILL, S. S.; TUTEJA, N. Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. **Plant Physiol Biochem**, [S.I.], v. 48, n. 12, p. 909-30, 2010.

GOFF S.A. et al A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). **Science**, [S.I.], v. 296, p. 92-100, 2002.

GORDON, D.; ABAJIAN, C.; GREEN, P. Consed: A graphical tool for sequence finishing. **Genome Research**, [S.I.], v. 8, p. 195-202, 1998.

GORMAN, M. A. et al. The crystal structure of the human DNA repair endonuclease HAP1 suggests the recognition of extra-helical deoxyribose at DNA abasic sites. **EMBO J.**, [S.I.], v. 16, p. 6548-6558, 1997.

GOSSAGE, L. et al. Base excision repair factors are promising prognostic and predictive markers in cancer. **Current molecular pharmacology**, [S.I.], v. 5, n. 1, p. 115–24, 2012.

GRANT, V. **Plant Speciation**. [S.I.]: Columbia University Press, 1981.

GRIVET, L.; ARRUDA, P. Sugarcane genomics: depicting the complex genome of an important tropical crop. **Current Opinion in Plant Biology**, [S.I.], v. 5, p. 122-127, 2001.

GRIVET, L.; GLASZMANN, J.; D'HONT, A. Molecular evidence of sugarcane evolution and domestication. In: MOTLEY, T. J. (Ed). **Darwin's harvest: new approaches to the origins, evolution and conservation of crops.** Columbia: University Press, 2006. p.49-66.

GUTIERREZ, R. A. et al. Phylogenetic profiling of the *Arabidopsis thaliana* proteome: What proteins distinguish plants from mother organisms? **Genome Biol.**, [S.I.], v. 5, p. 53, 2004.

HANDAYANI, N. S. N.; TANAKA, N.; YOSHIDA, K. Phenotype of Transgenic Tobacco Plants (*Nicotiana tabacum* cv . *Petit Havana* SR-1) Expressing 1724orf13 Gene of *Agrobacterium rhizogenes* strain. **Indonesian Journal of Biotechnology**, [S.I.], v. 12, n. 2, p. 980-987, 2007.

HAYS, J. B. *Arabidopsis thaliana*, a versatile model system for study of eukaryotic genome-maintenance functions. **DNA Repair**, [S.I.], v. 1, p. 579-600, 2002.

HESS, B. P-LINCS: a parallel linear constraint solver for molecular simulation. **Journal of Chemical Theory and Computation**, [S.I.], v. 4, p. 116, 2008.

HESS B. et al. GROMACS 4: algorithms for highly efficient load-balanced, and scalable molecular simulation. **Journal of Chemical Theory and Computation**, [S.I.], v. 4, p. 435-447, 2008.

HOBBS, S. L. A.; WARKENTIN, T. D.; DELONG, C. M. O. Transgene copy number can be positively or negatively associated with transgene expression. **Plant Mol. Biol.**, [S.I.], v. 21, p. 17-26, 1993.

HOEIJMAKERS, J. H. DNA damage, aging, and cancer. **New England Journal Medicine**, [S.I.], v. 361, p. 1475-1485, 2009.

HOPFNER, K. P. et al. The Rad50 zinc-hook is a structure joining Mre11 complexes in DNA recombination and repair. **Nature**, [S.I.], v. 418, p. 562-566, 2002.

HORBACH, R. et al. When and how to kill a plant cell: infection strategies of plant pathogenic fungi. **J Plant Physiol**, [S.I.], v. 168, p. 51-62, 2011.

HORTON P. et al. WoLF PSORT: protein localization predictor. **Nucleic Acids Research**, [S.I.], v. 35, p. 585–587, 2007.

HOSFIELD, D. J. et al. Structure of the DNA repair enzyme endonuclease IV and its DNA complex. Double-nucleotide flipping at abasic sites and three-metal-ion catalysis. **Cell**, [S.I.], v. 98, p. 397-408, 1999.

HU, W. et al. TaASR1, a transcription factor gene in wheat, confers drought stress tolerance in transgenic tobacco Plant, **Cell and Environment** ., [S.I.] v: 36, p: 1449–1464, 2013.

HUMPHREY, W. F., DALKE, A.; SCHULTEN, K. V. M. D. Visual molecular Dynamics. **Journal of Molecular Graphics**, [S.I.], v. 14, p. 33-38, 1996.

INTERNATIONAL BRACHYPODIUM INITIATIVE .Genome sequencing and analysis of the model grass Brachypodium distachyon. **Nature**, [S.I.], v. 463, p. 763-768, 2010.

ISCHENKO, A. A.; SAPARBAEV, M. K. Alternative nucleotide incision repair pathway for oxidative DNA damage. **Nature**, [S.I.], v. 415, p. 183-187, 2002.

ISHCHENKO, A. A. et al. Characterisation of new substrate specificities of Escherichia coli and Saccharomyces cerevisiae AP endonucleases. **Nucleic Acids Res.**, [S.I.], v. 31, p. 6344-6353, 2003.

GROS, L.; ISHCHENKO, A. A., IDE, H.; ELDER, R. H.; SAPARBAEV, M. K. The major human AP endonuclease (Ape1) is involved in the nucleotide incision repair pathway. *Nucleic Acids Research*, [S.I.],v. 32, N. 1, p. 73±81, 2004.

IZUMI, T. et al. Requirement for human AP endonuclease 1 for repair of 3'-blocking damage at DNA single-strand breaks induced by reactive oxygen species. **Carcinogenesis**, [S.I.], v. 21, p. 1329-1334, 2000.

JAILLON, O. et al. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. **Nature**, [S.I.], v. 431, p. 946-957, 2004.

JAILLON, O. J. et al. For the french-italian public consortium for grapevine genome characterization. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. **Nature**, [S.I.], v. 449, p. 463, 2007.

JIANG, W. et al. Prevalent Role of Gene Features in Determining Evolutionary Fates of Whole-Genome Duplication Duplicated Genes in Flowering Plants. **Plant Physiology**, [S.I.], v. 161, n. 4, p. 1844-1861, 2013.

JIAO Y. et al. Ancestral polyploidy in seed plants and angiosperms. **Nature**, [S.I.], v. 473, p. 97-100, 2011.

JONES, J. D. G. et al. T-DNA structure and gene expression in petunia plants transformed by *Agrobacterium tumefaciens* C58 derivatives. **Mol. Gen. Genet.**, [S.I.], v. 207, p. 478-485, 1987.

KATOH K., STANDLEY, D.M. Parallelization of the MAFFT multiple sequence alignment program. **Bioinformatics**, [S.I.], v. 26, p. 1899-1900, 2010.

KELLIS, M.; BIRREN, B. W.; LANDER, E. S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. **Nature**, [S.I.], v. 428, p. 617-624, 2004.

BARRETO, K. F. M. **Prospecção de proteínas em resposta ao estresse oxidativo em cana-de-açúcar, (*saccharum sp*)**. 2008-2011. Dissertação (Mestrado em Ciências Biológicas)- Universidade Federal do Rio Grande do Norte, Natal, 2011.

KERINS, S. M., COLLINS, R., MCCARTHY, T. V. () Characterization of an endonuclease IV 3'-5' exonuclease activity. **J. Biol. Chem.** 278, 3048-3054.2003.

KIKUCHI S, SATOH K, NAGATA T, KAWAGASHIRA N, DOI K, KISHIMOTO N, YAZAKI J, ISHIKAWA M, YAMADA H, OOKA H, HOTTA I, KOJIMA K, NAMIKI T, OHNEDA E, YAHAGI W, SUZUKI K, LI CJ, OHTSUKI K, SHISHIKI T, OTOMO Y, MURAKAMI K, IIDA Y, SUGANO S, FUJIMURA T, SUZUKI Y, TSUNODA Y, KUROSAKI T, KODAMAT, MASUDAH, KOBAYASHI M, XIE Q, LUM, NARIKAWA R, SUGIYAMA A, MIZUNO K, YOKOMIZO S, NIIKURA J, IKEDA R, ISHIBIKI J, KAWAMATAM, YOSHIMURAA, MIURA J, KUSUMEGI T, OKA M, RYU R, UEDA M, MATSUBARA K, KAWAI J, CARNINCI P, ADACHI J, AIZAWA K, ARAKAWA T, FUKUDA S, HARA A, HASHIZUME W, HAYATSU N, IMOTANI K, ISHII Y, ITOH M, KAGAWA I, KONDO S, KONNO H, MIYAZAKI A, OSATO N, OTA Y, SAITO R, SASAKI D, SATO K, SHIBATA K, SHINAGAWA A, SHIRAKI T, YOSHINO M, HAYASHIZAKI Y, YASUNISHI A. Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. **Science** 301:376-379.2003.

KIM S.I., VEENA , GELVIN S.B. Genome-wide analysis of Agrobacterium T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. **Plant J**, 51: 779–791. 2007.

KIMURA S., SAKAGUCHI K. DNA repair in plants. **Chemical reviews**, 106, 753-766. 2006.

KIMURA, S., UCHIYAMA, Y., KASAI, N., NAMEKAWA, S., SAOTOME, A., UEDA, T., ANDO, T., ISHIBASHI, T., OSHIGE, M., FURUKAWA, T., YAMAMOTO, T., HASHIMOTO, J., SAKAGUCHI, K.. A novel DNA polymerase homologous to *Escherichia coli* DNA polymerase I from a higher plant, rice (*Oryza sativa* L.). **Nucleic Acids Res.** 30,:1585–1592. 2002.

KLUNGLAND, A. AND LINDAHL, T. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). **EMBO J.** 16, 3341–3348. 1997.

KRANNER I., MINIBAYEVA F.V., BECKETT R.P., SEAL C.E. What is stress? Concepts, definitions and applications in seed science. **New Phytol.** 188(3):655-73. 2010.

KROKAN, H. E.; BJØRÅS, M. Base excision repair. **Cold Spring Harbor perspectives in biology**, 5, (4), 012583. Apr/ 2013.

KUBOTA, Y., NASH, R.A., KLUNGLAND, A., SCHAR, P., BARNES, D.E. AND LINDAHL, T. Reconstitution of DNA base excision-repair with purified human proteins: interaction between DNA polymerase b and the XRCC1 protein. **EMBO J.**, 15, 6662–6670. 1996.

LAGERWERF, S., VROUWE, M. G., OVERMEER, R. M., FOUSTERI, M. I., MULLENDERS, L.H. DNA damage response and transcription. **DNA Repair**, 10 (7): 1568-7864. 2011.

LEVIN, D.S., BAI, W., YAO, N., O'DONNELL, M., AND TOMKINSON, A.E. An interaction between DNA ligase I and proliferating cell nuclear antigen: implications for Okazaki fragment synthesis and joining. **Proc Natl Acad Sci USA** 94:12863–12868. 1997.

LEVIN, J. D., JOHNSON, A. W., DEMPLE, B. Homogeneous *Escherichia coli* endonuclease IV. Characterization of an enzyme that recognizes oxidative damage in DNA. **J. Biol. Chem.** 263, 8066–8071. 1988.

LI B., LI N., DUAN X., WEI A., YANG A., ZHANG J. Generation of marker-free transgenic maize with improved salt tolerance using the FLP/FRT recombination system. **Journal of biotechnology**, 145 (2), 206–13.15/jan/ 2010.

LIANG C, JAISWAL P, HEBBARD C, AVRAHAM S, BUCKLER ES, CASSTEVENS T, HURWITZ B, MCCOUCH S, NI J, PUJAR A, RAVENSCROFT D, REN L, SPOONER W, TECLE I, THOMASON J, TUNG CW, WEI X, YAP I, YOUENS-CLARK K, WARE D, STEIN L. Gramene: a growing plant comparative genomics resource. **Nucleic Acids Res.** 36:D947-53.2008.

LINDAHL, T. Instability and decay of the primary structure of DNA. ,**Nature** 362, 709–715.1993.

LINDAHL, T., BARNES, D. E. Repair of endogenousDNA damage. Cold Spring Harb. **Symp. Quant. Biol.** 65, 127–133.2000.

LOEB, L. A. Apurinic sites as mutagenic intermediates. *Cell* 40, 483–484.1985.
LUO M., PLATTEN D., CHAUDHURY A., PEACOCK W.J., DENNIS E.S. Expression, imprinting, and evolution of rice homologs of the polycomb group genes. **Molecular Plant**, 2, 711–723. 2009.

LUO MC, DEAL KR, AKHUNOV ED, AKHUNOVA AR, ANDERSON OD, ANDERSON JA, BLAKE N, CLEGG MT, COLEMAN-DERR D, CONLEY EJ, CROSSMAN CC, DUBCOVSKY J, GILL BS, GU YQ, HADAM J, HEO HY, HUO N, LAZO G, MA Y, MATTHEWS DE, MCGUIRE PE, MORRELL PL, QUALSET CO, RENFRO J, TABANAO D, TALBERT LE, TIAN C, TOLENO DM, WARBURTON ML, YOU FM, ZHANG W, DVORAK J. Genome comparisons reveal a dominant mechanism of chromosome number reduction in grasses and accelerated genome evolution in Triticeae. **Proc. Natl. Acad. Sci. U.S.A.** 106, 15780–15785.2009.

LUSHCHAK, V.I. Adaptive response to oxidative stress: Bacteria, fungi, plants and animals. **Comparative Biochemistry and Physiology**, Part C 153 175–190. 2011.

LYNCH M, CONERY J.S. The evolutionary fate and consequences of duplicate genes. **Science** 290: 1151–1155. 2000.

LYNCH M., FORCE A.G. The origin of interspecific genomic incompatibility via gene duplication. **Am Nat**, 156:590-605. 2000.

LYNCH, M., FORCE, A. The probability of duplicate gene preservation by subfunctionalization. **Genetics**, 154, 459–473. 2000.

LYNCH, M. AND CONERY, J.S. The evolutionary fate and consequences of duplicate genes. **Science** 290, 1151–1155.2000.

LYNCH, M. AND FORCE, A. The probability of duplicate gene preservation by subfunctionalization. **Genetics** 154, 459–473.2000.

MACINDOE G., MAVRIDIS L., VENKATRAMAN V., DEVIGNES M., RITCHIE D.W. HexServer: An FFT-based protein docking server powered by graphics processors. **Nucleic Acids Research**, 38, 445-449. 2010.

MAHAJAN S., TUTEJA N. Cold, salinity and drought stresses: an overview. **Arch. Biochem. Biophys.** 444, 139-158. 2005.

MAHER, R. L., AND BLOOM, L. B. (2007) Pre-steady-state kinetic characterization of the AP endonuclease activity of human AP endonuclease 1. **J. Biol. Chem.** 282, 30577–30585.2007.

MAIRA, N., TORRES, T. M., DE OLIVEIRA, A. L., DE MEDEIROS, S. R. B., AGNEZ-LIMA, L. F., LIMA, J. P. M. S., SCORTECCI, K. C. (, Identification, characterisation and molecular modelling of two AP endonucleases from base excision repair pathway in sugarcane provide insights on the early evolution of green plants. **Plant Biology**. doi: 10.1111/plb.12083.2013.

MAKINO T, MCLYSAGHT A. Ohnologs in the human genome are dosage balanced and frequently associated with disease. **Proc Natl Acad Sci USA** 107(20):9270–9274. 2010.

MANNERS J.M. , CASU R.E. Transcriptome analysis and functional genomics of sugarcane. **Tropical Plant Biology**, 4, 9-21. 2011.

MARÉCHAL, A., PARENT, J., VÉRONNEAU-LAFORTUNE, F., JOYEUX, A., LANG, B. F., BRISSON N. Whirly proteins maintain plastid genome stability in Arabidopsis. **PNAS** 106 (34) 14693-14698. 2009.

MARÉCHAL, A., BRISSON, N. Recombination and the maintenance of plant organelle genome stability. **New Phytologist**, 186: 299–317.2010.

MATSUMOTO Y, KIM K. Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. **Science**,269(5224):699-702.1995.

MATSUOKA, S., GARCIA, A.A.F., CALHEIROS, G.C. Hibridação em cana-de-açúcar. In: **Borém A (ed) Hibridação Artificial de Plantas. Viçosa: Editora UFV**, 221–254. 1999.

MAYNARD S., SCHURMAN S.H., HARBOE C., DE SOUZA-PINTO N.C., BOHR V.A. I. Base excision repair of oxidative DNA damage and association with cancer and aging. **Carcinogenesis**, 30, (1), 2–10. Jan/2009.

MCCULLOUGH A.K., DODSON M.L., LLOYD R.S. Initiation of base excision repair: glyco- sylase mechanisms and structures. *Annu. Rev. Biochem.*, 68, 255–285. 1999.

MCMURRAY, C. T., TAINER, J. A. Cancer, cadmium and genome integrity. **Nat. Genet.** 34, 239–241.2003.

MEAGHER, R. The Evolution of epitype. **Plant Cell** 22, 1658–1666.2010.
MING, R., HOU S., FENG, Y., YU, Q., DIONNE-LAPORTE A, SAW, J.H., SENIN, P., WANG, W., LY, B.V., LEWIS, K.L., et al.. The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya*Linnaeus).**Nature**,452:991–996. 2008.

MITCHELL J.R., HOEIJMAKERS J.H., NIEDERNHOFER L.J. Divide and conquer: nucleotide excision repair battles câncer and ageing.**Curr. Opin. Cell. Biol.** 15:232– 240. 2003.

MOL C.D., IZUMI T., MITRA S., TAINER J.A. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination. **Nature**, 403, 451-456. 2000.

MOL C.D., PARIKH S.S., PUTNAM C.D., LO T.P., TAINER J.A .DNA repair mechanisms for the recognition and removal of damaged DNA bases. **Annu. Rev. Biophys. Biomol. Struct.**, 28, 101–128. 1999.

MOL, C. D., HOSFIELD, D. J., TAINER, J. A. () Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair. The 3' ends justify the means. **Mutat. Res.** 460, 211–229.2000.

MOL, C. D., IZUMI, T., MITRA, S., AND TAINER, J. A. DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination. **Nature**, 403, 451–456.2000.

MOL, C. D., KUO, C. F., THAYER, M. M., CUNNINGHAM, R. P., TAINER, J. A. Structure and function of the multifunctional DNA-repair enzyme exonuclease III. **Nature**, 374, 381–386.1995.

MORGANTE C.V, RODRIGUES R.A.O., MARBACH P.A.S., BORGONOVIC C.M., MOURA D.S., SILVA-FILHO M.C. Conservation of dual-targeted proteins in Arabidopsis and rice points to a similar pattern of gene-family evolution. **Molecular Genetics & Genomics**, 281, 525–538. 2009.

MUNDLE, S. T., DELANEY, J. C., ESSIGMANN, J. M., AND STRAUSS, P. R. Enzymatic mechanism of human apurinic/aprimidinic endonuclease against a THF AP site model substrate. *Biochemistry* 48, 19–26.2009.

MUNDLE, S. T., FATTAL, M. H., MELO, L. F., CORIOLAN, J. D., O'REGAN, N. E., AND STRAUSS, P. R. (2004) Novel role of tyrosine in catalysis by human AP endonuclease 1. **DNA Repair**, 3, 1447–1455.2004.

MURPHY T.M., BELMONTE M., SHU S., BRITT A.B., HATTEROTH J. Requirement for abasic endonuclease gene homologues in Arabidopsis seed development. **PLoS ONE**, 4, 4297. 2009.

NARUSAKA Y., NARUSAKA M., SEKI M., FUJITA M., ISHIDA J., NAKASHIMA M., ENJU A., SAKURAI T., SATOU M., KAMIYA A., PARK P., KOBAYASHI M., SHINOZAKI K. Expression profiles of Arabidopsis phospholipase A IIA gene in response to biotic and abiotic stresses. **Plant and Cell Physiology** , 44, 1246–1252. 2003.

NASH, R.A., CALDECOTT, K.W., BARNES, D.E., LINDAHL, T.. XRCC1 protein interacts with one of two distinct forms of DNA ligase III. **Biochemistry** ,36:5207–5211. 1997.

NATHALIA MAÍRA CABRAL DE MEDEIROS. Modelagem de proteínas de SCARP1 e SCARP3 homologas a ARP de Arabidopsis thaliana em cana-de-açúcar. **Trabalho de Conclusão de Curso** (Graduação em Ciências Biológicas) - Universidade Federal do Rio Grande do Norte. 2011.

NATURE PUBLISHING GROUP HOEIJMAKERS, J. H. J. Genome maintenance mechanisms for preventing cancer. **Nature**, 411, 366–374 .2001.

NICHOLL, I.D., NEALON, K. AND KENNY, M.K. Reconstitution of human base excision repair with purified proteins. **Biochemistry**, 36, 7557–7566. 1997.

NICOTRA A.B., ATKIN O.K., BONSER S.P., DAVIDSON A.M., FINNEGAN E.J., MATHESIUS U., POOT P., PURUGGANAN M.D., RICHARDS C.L., VALLADARES F., VAN KLEUNEN M. Plant phenotypic plasticity in a changing climate. **Trends in plant science**, v. 15, n. 12, p. 684–92, dez. 2010.

O'BRIEN J.A., DAUDI A., BUTT V.S., BOLWELL G.P. Reactive oxygen species and their role in plant defence and cell wall metabolism. **Planta**, 236(3):765-79.2012.

OHNO S. Evolution by Gene Duplication. **Springer-Verlag**. 1970.

OLSSON M. H. M., SONDERGAARD C. R., ROSTKOWSKI M., JENSEN J.H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. **Journal of Chemical Theory and Computation**, 2, 525–537. 2011.

OTTO SP, WHITTON J.. Polyploid incidence and evolution. **Annu. Rev. Genet.** 34:401–37.2000.

OTTO, S.P, WHITTON, J. Polyploid incidence and evolution. **Annu Rev Genet**, 34:401-437. 2000.

PANT B.D., BUHTZ A., KEHR J., SCHEIBLE W.R. MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. **Plant Journal**, 53, 731-738. 2008.

PARSONS, J.L. & ELDER, R.H. DNA N-glycosylase deficient mice: a tale of redundancy. **Mutation Research.**, 531, 165–175.15/May/ 2003.

PASCUCCI, B., STUCKI, M., JONSSON, Z.O., DOGLIOTTI, E. AND HUBSCHER, U. Long patch base excision repair with purified human proteins. DNALigase I as patch size mediator for DNA polymerases d and e. **J. Biol. Chem.** 274, 33696–33702. 1999.

PATERSON A.H, BOWERS J.E., CHAPMAN B.A. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. **Proc Natl Acad Sci U S A**, 101:9903-9908. 2004.

PATERSON AH, BOWERS JE, BRUGGMANN R, DUBCHAK I, GRIMWOOD J, GUNDLACH H, HABERER G, HELLSTEN U, MITROS T, POLIAKOV A, SCHMUTZ J, SPANNAGL M, TANG H, WANG X, WICKER T, BHARTI AK, CHAPMAN J, FELTUS FA, GOWIK U, GRIGORIEV IV, LYONS E, MAHER CA, MARTIS M, NARECHANIA A, OTILLAR RP, PENNING BW, SALAMOV AA, WANG Y, ZHANG L, CARPITA NC, FREELING M, GINGLE AR, HASH CT, KELLER B, KLEIN P, KRESOVICH S, MCCANN MC, MING R, PETERSON DG, MEHBOOB-UR-RAHMAN, WARE D, WESTHOFF P, MAYER KF, MESSING J, ROKHSAR DS. The Sorghum bicolor sequence and the diversification of grasses. **Nature** 457, 551–556.2009.

PATERSON, A.H., BOWERS, J.E., CHAPMAN, B.A. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. **Proc Natl Acad Sci USA** 101:9903–9908. 2004

PEDERSEN A.C., NIELSEN H. Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. **In Proceeding of the 5th International Conference on Intelligent Systems for Molecular Biology**, 5, 226–233.1997.

PETTERSEN E. F., GODDARD T. D., HUANG C. C., COUCH G.S., GREENBLATT D.M., MENG E.C., FERRIN T.E. UCSF Chimera—A visualization system for exploratory research and analysis. **Journal of Computational Chemistry**, 25, 1605–1612. 2004.

PIPERIDIS G, PIPERIDIS N, D'HONT. A Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. **Mol Genet Genomics** 284(1):65–73.2010.

PIPERIDIS, G., PIPERIDIS, N., D'HONT, A. Molecular cytogenetic investigation of chromosome composition and transmission in sugarcane. **Mol Genet Genomics**, 284(1):65–73.2010.

POEHLMANN A, ROESSNER A. Importance of DNA damage checkpoints in the pathogenesis of human cancers. **Pathol Res Pract**, 206(9):591-601.2010.

POREBSKI S., BAILEY L.G., BAUM B.R. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. **Plant Molecular Biology Reporter**, 15, 8-15. 1997.

POSADA D., BUCKLEY T.R. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. **Systems Biology**, 53, 793-808. 2004.

PRESTON J.C., CHRISTENSEN A., MALCOMBER S.T., KELLOGG E.A. Mads-box gene expression and implications for developmental origins of the grass spikelet. **Am J Bot**, 96:1419-1429. 2009.

PRINCE V.E., PICKETT F.B. Splitting pairs: the diverging fates of duplicated genes. **Nature Reviews Genetics**, 3, 827–837. 2002.

PROOST, S., BELA, M. V., STERCKA, L., BILLIAUA, K., PARYSA T. V., VAN DE PEERA, Y., VANDEPOELEA, K. PLAZA, a comparative genomics resource to study gene and genome evolution in plants. **Plant Cell** 21:3718 –3731. 2009.

RAJJOU L., DEBEAUJON I. Seed longevity: survival and maintenance of high germination ability of dry seeds. **C R Biol**, 331:796–805. 2008.

RAMAKERS C., RUIJTER J.M., DEPREZ R.H., MOORMAN A.F. Assumption free analysis of quantitative real-time polymerase chain reaction (PCR) data. **Neuroscience Letters**, 13, 62–66. 2003.

RAMSEY J, SCHEMSKE D. Neopolyploidy in flowering plants. **Annu Rev Ecol Syst**, 33:589-639. 2002

RODGERS-MELNICK, E., MANE, S. P., DHARMAWARDHANA, P., SLAVOV, G. T., CRASTA, O. R., STRAUSS, S. H., BRUNNER, A. M., DIFAZIO, S. P. Contrasting patterns of evolution following whole genome versus tandem duplication events in *Populus*. **Genome Res**, 22:95-105. 2012.

ROLDÁN-ARJONA T., ARIZA R.R. Repair and tolerance of oxidative DNA damage in plants. **Mutation Research**, 681, 169-179. 2009.

RONQUIST F., HUELSENBECK J. MrBayes 3: bayesian phylogenetic inference under mixed models. **Bioinformatics**, 19, 1572–1574. 2003.

ROY S., CHOUDHURY S.R., SINGH S.K., DAS K.P. AtPol λ , a homolog of mammalian DNA polymerase λ *Arabidopsis thaliana*, is involved in the repair of UV-B induced DNA damage through the dark repair pathway. **Plant Cell Physiology**, 52, 448–467. 2011.

RUTTER M.T., CROSS K.V., VAN WOERT P.A. Birth, death and sub-functionalization in the Arabidopsis genome. **Trends in Plant Science**, 17, 204–212. 2012.

SALSE J, BOLOT S, THROUDE M, JOUFFE V, PIEGU B, QURAIISHI UM, CALCAGNO T., COOKE R., DELSENY M., FEUILLET C. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. **Plant Cell**, 20:11-24. 2008.

SAMBROOK J.J., RUSSELL D.D.W. Molecular Cloning: a Laboratory Manual, 3rd, **Cold Spring Harbor**, NY. 2001.

SANAN-MISHRA N., PHAM X.H, SOPORY S.K., TUTEJA N. Pea DNA helicase 45 overexpression in tobacco confers high salinity tolerance without affecting yield. **Proc Natl Acad Sci USA**, 102:509–514.2005.

SCANNELL, D.R., BYRNE, K.P., GORDON, J.L., WONG, S., WOLFE. K.H. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. **Nature** 440: 341–345. 2006.

SCHMUTZ, J., CANNON, S.B., SCHLUETER, J., MA, J., MITROS, T., NELSON, W., HYTEN, D.L., SONG, Q., THELEN, J.J., CHENG, J., et al. Genome sequence of the palaeopolyploid soybean. **Nature**, 463:178–183. 2010.

SCHNABLE PS, WARE D, FULTON RS, STEIN JC, WEI F, PASTERNAK S, LIANG C, ZHANG J, FULTON L, GRAVES TA, MINX P, REILY AD, COURTNEY L, KRUCHOWSKI SS, TOMLINSON C, STRONG C, DELEHAUNTY K, FRONICK C, COURTNEY B, ROCK SM, BELTER E, DU F, KIM K, ABBOTT RM, COTTON M, LEVY A, MARCHETTO P, OCHOA K, JACKSON SM, GILLAM B, CHEN W, YAN L, HIGGINBOTHAM J, CARDENAS M, WALIGORSKI J, APPLEBAUM E, PHELPS L, FALCONE J, KANCHI K, THANE T, SCIMONE A, THANE N, HENKE J, WANG T, RUPPERT J, SHAH N, ROTTER K, HODGES J, INGENTHRON E, CORDES M, KOHLBERG S, SGRO J, DELGADO B, MEAD K, CHINWALLA A, LEONARD S, CROUSE K, COLLURA K, KUDRNA D, CURRIE J, HE R, ANGELOVA A, RAJASEKAR S, MUELLER T, LOMELI R, SCARA G, KO A, DELANEY K, WISSOTSKI M, LOPEZ G, CAMPOS D, BRAIDOTTI M, ASHLEY E, GOLSER W, KIM H, LEE S, LIN J, DUJMIC Z, KIM W, TALAG J, ZUCCOLO A, FAN C, SEBASTIAN A, KRAMER M, SPIEGEL L, NASCIMENTO L, ZUTAVERN T, MILLER B, AMBROISE C, MULLER S, SPOONER W, NARECHANIA A, REN L, WEI S, KUMARI S, FAGA B, LEVY MJ, MCMAHAN L, VAN BUREN P, VAUGHN MW, YING K, YEH CT, EMRICH SJ, JIA Y, KALYANARAMAN A, HSIA AP, BARBAZUK WB, BAUCOM RS, BRUTNELL TP, CARPITA NC, CHAPARRO C, CHIA JM, DERAGON JM, ESTILL JC, FU Y, JEDDELOH JA, HAN Y, LEE H, LI P, LISCH DR, LIU S, LIU

Z, NAGEL DH, MCCANN MC, SANMIGUEL P, MYERS AM, NETTLETON D, NGUYEN J, PENNING BW, PONNALA L, SCHNEIDER KL, SCHWARTZ DC, SHARMA A, SODERLUND C, SPRINGER NM, SUN Q, WANG H, WATERMAN M, WESTERMAN R, WOLFGRUBER TK, YANG L, YU Y, ZHANG L, ZHOU S, ZHU Q, BENNETZEN JL, DAWE RK, JIANG J, JIANG N, PRESTING GG, WESSLER SR, ALURU S, MARTIENSSEN RA, CLIFTON SW, MCCOMBIE WR, WING RA, WILSON RK. The B73 maize genome: complexity, diversity and dynamics. **Science** 326, 1112–1115.2009.

SCHNABLE, P.S., WARE, D., FULTON, R.S., STEIN, J.C., WEI, F., PASTERNAK, S., LIANG, C., ZHANG, J., FULTON, L., GRAVES, T.A., et al. The B73 maize genome: Complexity, diversity, and dynamics. **Science**,326:1112–1115. 2009.

SCHRANZ, M. E.; MOHAMMADIN, S.; EDGER, P. P. Ancient whole genome duplications , novelty and diversification : the WGD Radiation Lag-Time Model. **Current Opinion in Plant Biology**, 15, 2, 147–153.2012.

SCORTECCI K.C., CRESTE S., CALSA T. JR., XAVIER, M.A., LANDELL , M.G.A.L., FIGUEIRA, A., BENEDITA, V.A. Challenges, Opportunities and Recent Advances in Sugarcane Breeding, in: I.Y. Abdurakhmonov (Ed.). **Plant Breeding, Intech**, ISBN978-953-307-932-5, DOI: 10.5772/28606. 2012.

SEEGER R.C., BRODEUR G.M., SATHER H., DALTON A., SIEGEL S.E., WONG K.Y., HAMMOND D. Association of multiple copies of the N-myc oncogene with rapid progression of neuroblastomas. **N Engl J Med** ,313(18):1111–1116. 1985.

SHARMA Y.K, DAVIS K.R. The eVects of ozone on antioxidant responses in plants. **Free Radic Biol Med** 23:480–488.1997.

SHEN M. Y., SALI A. Statistical potential for assessment and prediction of protein structures. **Protein Science**, 15, 2507–2524. 2006.

SHI L, KENT R, BENICE N, BRITT AB. Developmental expression of a DNA repair gene in Arabidopsis. **Mutation Research** 384: 145–156. 1997.

SIBERIL Y., DOIREAU P., GANTET P. Plant bZIP G-box binding factors. Modular structure and activation mechanisms. **European Journal of Biochemistry**, 268, 5655–5666. 2001.

SIEVERS F., WILM A., DINEEN D., GIBSON T.J., KARPLUS K., LI W., LOPEZ R., MCWILLIAM H., REMMERT M., SODING J., THOMPSON J.D., HIGGINS D.D. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, 7, 539. 2011.

SIEVERS, F., WILM, A., DINEEN, D.G., GIBSON,T.J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SÖDING, J., THOMPSON, J.D., HIGGINS, D.G. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, 7:539-75.2011.

SINCLAIR DA, OBERDOERFFER P.The ageing epigenome: damaged beyond repair?.**Age. Res. Rev.** 8:189–198. 2009.

SINGH S.K., ROY S., CHOUDHURY S.R., SENGUPTA D.N.DNA repair and recombination in higher plants: insights from comparative genomics of Arabidopsis and rice. **BMC Genomics**, 11:443, 1471-2164.2010.

SINGH, S. K. CHOUDHURY, S. R., ROY, S. , SENGUPTA D. N.Understanding DNA repair and recombination in higher plant genome: Information from genome-wide screens in Arabidopsis and rice. **Plant Signaling & Behavior**, 6 (1), 120–122, 1 jan/ 2011.

SINGHAL, R.K., PRASAD, R., ANDWILSON, S.H. DNA polymerase β conducts the gap-filling step in uracil-initiated base excision repair in a bovine testis nuclear extract. **J Biol Chem**, 270:949–957. 1995.

SINGLETON A.B, FARRER M., JOHNSON J., SINGLETON A., HAGUE S., KACHERGUS J., HULIHAN M., PEURALINNA T., DUTRA A., NUSSBAUM R., LINCOLN S., CRAWLEY A., HANSON M., MARAGANORE D., ADLER C., COOKSON M.R., MUENTER M., BAPTISTA M., MILLER D., BLANCATO J., HARDY J., GWINN-HARDY K. Alpha-Synuclein locus triplication causes Parkinson's disease. **Science** 302(5646):841.2003.

SOBOL, R. W., KARTALOU, M., ALMEIDA, K. H., JOYCE, D. F., ENGELWARD, B. P., HORTON, J. K., PRASAD, R., SAMSON, L. D., WILSON, S. H. Base excision repair intermediates induce p53-independent cytotoxic and genotoxic responses. **J. Biol. Chem.** 278, 39951–39959.2003.

SOBOL,R.W., HORTON,J.K., KUHN,R., GU,H., SINGHAL,R.K., PRASAD,R., RAJEWSKY,K.,WILSON,S.H. Requirement of mammalian DNA polymerase-bin base-excision repair. **Nature**, 379, 183–186. 11/ January /1996.

Souza, G. M., Berges, H., Bocs, S., Casu, R., D'Hont, A., Ferreira, J. E., Henry, R. Ming, R., Potier, B., Sluys, M.V., Vincentz, M., Paterson, A.H. The Sugarcane Genome Challenge: Strategies for Sequencing a Highly Complex Genome. **Tropical Plant Biol.** 4:145–156. 2011.

SRIVASTAVA, D.K., BERG, B.J., PRASAD, R., MOLINA, J.T., BEARD, W.A., TOMKINSON, A.E., WILSON, S.H. Mammalian abasic site base excision repair. Identification of the reaction sequence and rate-determining steps. **J. Biol. Chem.**, 273, 21203–21209. 1998.

STEBBINS, G.L. Variation and Evolution in Plants. **Columbia University Press.** 1950.

SUNDERLAND, P.A., WEST, C.E., WATERWORTH, W.M. AND BRAY, C.M. An evolutionarily conserved translation initiation mechanism regulates nuclear or mitochondrial targeting of DNA ligase 1 in *Arabidopsis thaliana*. **Plant J.** 47, 356–367. 2006.

TANG H., BOWERS J. E., WANG X., PATERSON A.H. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. **Proceedings of the National Academy of Sciences**, 107, 472–477. 2010.

TATE, J.A., JOSHI, P., SOLTIS, K.A., SOLTIS, P.S., SOLTIS, D.E. On the road to diploidization? Homoeolog loss in independently formed populations of the allopolyploid *Tragopogon miscellus* (Asteraceae). **BMC Plant Biol.** 9: 80. 2009.

TEBBS, R.S., THOMPSON, L.H., CLEAVER, J.E. Rescue of Xrcc1 knockout mouse embryo lethality by transgene-complementation. **DNA Repair (Amst)**, 2, 1405–1417. 9/August/2003.

THE ARABIDOPSIS GENOME INITIATIVE. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. **Nature** 408: 796–815. 2000.

TSUTAKAWA, S. E., JINGAMI, H., AND MORIKAWA, K. () Recognition of a TG mismatch. The crystal structure of very short patch repair endonuclease in complex with a DNA duplex. **Cell**, 99, 615–623. 1999.

TSUTAKAWA, S. E., SHIN, D. S., MOL, C. D., IZUMI, T., ARVAI, A. S., MANTHA, A. K., SZCZESNY, B., IVANOV, I. N., HOSFIELD, D. J., MAITI, B., PIQUE, M. E., FRANKEL, K. A., HITOMI, K., CUNNINGHAM, R. P., MITRA, S., TAINER, J. A. Conserved Structural Chemistry for Incision Activity in

Structurally Non-homologous Apurinic/Apyrimidinic Endonuclease APE1 and Endonuclease IVDNA Repair Enzymes. **J. Biol. Chem**, 288:8445-8455. 2013.

TUSKAN, G. A., DIFAZIO, S., JANSSON, S., BOHLMANN, J., GRIGORIEV, I., HELLSTEN, U., PUTNAM, N., RALPH, S., ROMBAUTS, S., SALAMOV, A., SCHEIN, J., L. STERCK, A. AERTS, R. R. BHALERAO, R. P. BHALERAO, D. BLAUDEZ, W. BOERJAN, A. BRUN, A. BRUNNER, V. BUSOV, M. CAMPBELL, J. CARLSON, M. CHALOT, J. CHAPMAN, G.-L. CHEN, D. COOPER, P. M. COUTINHO, J. COUTURIER, S. COVERT, Q. CRONK, R. CUNNINGHAM, J. DAVIS, S. DEGROEVE, A. DÉJARDIN, C. DEPAMPHILIS, J. DETTER, B. DIRKS, I. DUBCHAK, S. DUPLESSIS, J. EHLTING, B. ELLIS, K. GENDLER, D. GOODSTEIN, M. GRIBSKOV, J. GRIMWOOD, A. GROOVER, L. GUNTER, B. HAMBERGER, B. HEINZE, Y. HELARIUTTA, B. HENRISSAT, D. HOLLIGAN, R. HOLT, W. HUANG, N. ISLAM-FARIDI, S. JONES, M. JONES-RHOADES, R. JORGENSEN, C. JOSHI, J. KANGASJÄRVI, J. KARLSSON, C. KELLEHER, R. KIRKPATRICK, M. KIRST, A. KOHLER, U. KALLURI, F. LARIMER, J. LEEBENS-MACK, J.-C. LEPLÉ, P. LOCASCIO, Y. LOU, S. LUCAS, F. MARTIN, B. MONTANINI, C. NAPOLI, D. R. NELSON, C. NELSON, K. NIEMINEN, O. NILSSON, V. PEREDA, G. PETER, R. PHILIPPE, G. PILATE, A. POLIAKOV, J. RAZUMOVSKAYA, P. RICHARDSON, C. RINALDI, K. RITLAND, P. ROUZÉ, D. RYABOY, J. SCHMUTZ, J. SCHRADER, B. SEGERMAN, H. SHIN, A. SIDDIQUI, F. STERKY, A. TERRY, C.-J. TSAI, E. UBERBACHE, P. UNNEBERG, J. VAHALA, K. WALL, S. WESSLER, G. YANG, T. YIN, C. DOUGLAS, M. MARRA, G. SANDBERG, Y. VAN DE PEER, D. ROKHSAR. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). **Science**, 313:1596–1604. 2006.

TUTEJA N, AHMAD P, PANDA BB, TUTEJA R. Genotoxic stress in plants: shedding light on DNA damage, repair and DNA repair helicases. **Mutat Res**, 681(2-3):134-49. 2009

TUTEJA N., SINGH M.B., MISRA M.K., BHALLA P.L., TUTEJA N. Molecular mechanisms of DNA damage and repair: Progress in plants. **Critical Reviews in Biochemistry and Molecular Biology**, 36, 337–97. 2001.

UCHIYAMA Y, KIMURA S, YAMAMOTO T, ISHIBASHI T, SAKAGUCHI K.. Plant DNA polymerase lambda, a DNA repair enzyme that functions in plant meristematic and meiotic tissues. **European Journal of Biochemistry**, 271: 2799–2807. 2004.

VAN DE PEER Y, MAERE S, MEYER A. The evolutionary significance of ancient genome duplications. **Nat. Rev. Genet.** 10, 725 –732. 2009.

VANDEPOELE, K., AND VAN DE PEER, Y. Exploring the plant transcriptome through phylogenetic profiling. **Plant Physiol.** 137: 31–42. 2005.

VASHISHT A.A, TUTEJA N. Stress responsive DEAD-box helicases: a new pathway to engineer plant stress tolerance. **J Photochem Photobiol B**,84(2):150-60.2006

VESPA L., COUVILLION M., SPANGLER E., SHIPPEN D.E. ATM and ATR make distinct contributions to chromosome end protection and the maintenance of telomeric DNA in *Arabidopsis*, **Genes Dev.** 19, 2111–2115. 2005.

WAGNER, W.H.: Biosystematics and evolutionary noise. **Taxon**, 19:146-151.26. 1970.

WALKER, L. J., ROBSON, C. N., BLACK, E., GILLESPIE, D., HICKSON, I. D. Identification of residues in the human DNA repair enzyme HAP1 (Ref-1) that are essential for redox regulation of Jun DNA binding. **Mol. Cell Biol.** 13, 5370–5376.1993.

WANG X., TANG H., PATERSON A. H. Seventy Million Years of Concerted Evolution of a Homoeologous Chromosome Pair, in Parallel, in Major Poaceae Lineages. **Plant Cell**, 23, 27–37. 2011.

WANG, Y., JHA, A.K., CHEN, R., DOONAN, J.H., YANG, M. Polyploidy-associated genomic instability in *Arabidopsis thaliana*. **Genesis**,48:254–263. 2010.

WATERHOUSE, A.M., PROCTER, J.B., MARTIN, D.M.A, CLAMP, M. BARTON, G. J. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. **Bioinformatics**, 25 (9): 1189-1191. 2009.

WATERWORTH, W.M., KOZAK, J., PROVOST, C.M., BRAY, C.M., ANGELIS, K.J. AND WEST, C.E. DNA ligase 1 deficient plants display severe growth defects and delayed repair of both DNA single and double strand breaks. **BMC Plant Biol.** 9, 79. 2009.

WENDEL, J. F. Genome evolution in polyploids. **Plant Mol. Biol.** 42, 225–249. . 2000.

WHELAN S., GOLDMAN N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. **Molecular Biology and Evolution**, 18, 691–699. 2001.

WILLER, M., RAINEY, M., PULLEN, T. AND STIRLING, C.J. The yeast CDC9 gene encodes both a nuclear and a mitochondrial form of DNA ligase I. **Curr. Biol.** 9, 1085–1094. 1999.

WU Y., ZHU Z., MA L., CHEN M. The preferential retention of starch synthesis genes reveals the impact of whole-genome duplication on grass evolution. **Mol Biol Evol**, 25: 1003-1006. 2008.

XU L., CARRIE C., LAW S., MURCHA M.W., WHELAN J. Acquisition, conservation and loss of dual-targeted proteins in land plants. **Plant Physiology**, 161, 644-662. 2013.

ZHANG, G., COHN, M.J. Genome duplication and the origin of the vertebrate skeleton. **Curr Opin Genet Dev** 18:387–393. 2008.

ZHARKOV D.O. Base excision DNA repair. **Cellular and Molecular Life Sciences**, 65, 1544-1565. 2008.