Chapter

Base Excision Repair in Sugarcane – A New Outlook

Nathalia Maíra Cabral de Medeiros
and Katia Castanho Scortecci

Abstract

The base excision repair (BER) pathway has been associated with genome integrity maintenance. Owing to its central role, BER is present in all three domains of life. The studies in plants, considering BER, have been conducted using *Arabidopsis* and rice models. Therefore, future studies regarding BER are required in other organisms, particularly in crops such as sugarcane, to understand its mechanism, which may reflect the uniqueness of DNA repair in monocots. Our previous results have revealed that sugarcane is an interesting plant for studying this pathway considering the polyploidy genome and genome evolution. This chapter aimed to characterize the BER pathway in sugarcane by using different bioinformatics tools, for example, screening for BER homologs in the sugarcane genome to identify its members. Each sequence obtained was subjected to structural analysis, and certain differences were identified when *Arabidopsis* was compared to other monocots, including sugarcane. Moreover, ROS1, DEM, and DML3 were not identified as a complete sequence in the sugarcane EST database. Furthermore, FEN1 is present as two sequences, namely FEN1A and FEN1B, both featuring different amino acid sequence and motif presence. Furthermore, FEN1 sequence was selected for further characterization considering its evolutionary history, as sequence duplication was observed only in *the Poaceae* family. Considering the importance of this protein for BER pathway, this sequence was evaluated using protein models (3D), and a possible conservation was observed during protein–protein interaction. Thus, these results help us understand the roles of certain BER components in sugarcane, and may reveal the aspects and functions of this pathway beyond those already established in the literature.

Keywords: BER, *Saccharum* spp., DNA repair, *Poaceae*, 3D-model, phylogenetic

1. Introduction

The base excision repair (BER) pathway is linked to the maintenance of genome integrity since BER is an essential genome defense pathway, which acts over a broad range of DNA lesions induced by endogenous or exogenous genotoxic agents [1]. Owing to its central role, BER is present in all three domains of life [2]. As a complex process, BER initiated by the excision of damaged base, proceeds through a sequence of reactions that generate various DNA intermediates and finish with the repair of the initial DNA structure. Nevertheless, BER focuses on repair, deals with DNA demethylation and erases the epigenetic mark 5-methycytosine (5mC) and

converts it to cytosine [3]. Thus, an emerging and crucial role of BER in epigenetic regulation is being investigated and characterized [4–7]. Although various studies have been conducted in animal and microbial systems, BER knowledge regarding plants has been neglected.

Despite these apparent differences in plant research compared to other organisms, knowledge about the BER pathway in plants has gained immense interest in recent years. The results obtained so far reveal that plants possess orthologues of most BER genes previously found in other organisms [8–10]; however, they also retain some plant-specific BER proteins as well as distinct enzyme combinations not observed in other kingdoms (review by [8]). Unfortunately, most of these findings were based on the model *Arabidopsis thaliana*, indicating the importance of amplifying studies on other organisms, particularly important crops [11].

Grasses (*Poaceae*; alternative name *Gramineae*) are undoubtedly an important plant group considering the economic perspective, and provide essential cereals such as *Eragrostis*, *Hordeum*, *Oryza*, *Secale*, *Sorghum*, *Triticum* and *Zea*; stalks such as *Arundo* and *Phragmites*; cane for food and materials for construction such as *Bambusa* and *Phyllostachys* and sugar crops such as *Saccharum* and *Sorghum* [12]. Sugarcane is a crop of noticeable value that can meet the requirements of food, feed fiber, and fuel. Moreover, sugarcane production by weight surpasses that of food crops such as wheat, rice and maize [13]. Despite its importance, this crop has been given less attention in scientific research than other members of *Poaceae* family, such as rice and maize. One reason is the polyploid and heterozygous nature of its genome, leading to lesser research compared to the other grass species studied [14–16].

Furthermore, research has been conducted using the sugarcane expressed sequence tags (ESTs) project (SUCEST), which has identified possible DNA repair genes [17, 18]. BER sequences were predicted, although these investigations were conducted more than 10 years ago [19]. Since then, there have been several improvements in bioinformatics tools as well as in sugarcane genome sequencing [20–24].

More studies are required to unravel the specific features of BER pathway in sugarcane, which may reflect the uniqueness of DNA repair in monocots. A new screening for BER homologs in the sugarcane genome was developed to gain advanced knowledge of BER in this crop. Each sequence was structurally analyzed. Thereafter, some of these sequences were selected for further investigating their evolutionary history. Tri-dimensional models have also been created to verify the conservation of mechanisms and protein–protein interactions in sugarcane BER components. The intriguing results displayed in this chapter raise questions regarding the roles of certain components of BER in sugarcane, just as in monocots, and they might broaden the aspects and functions of this pathway beyond those already established in the scientific literature.

2. Identification of base excision repair's components in sugarcane

The BER components were identified in sugarcane through homology with the bioinformatic tools. In this regard, the SUCEST-FUN database, which assembles distinct sugarcane databases such as the Sugarcane Expressed Sequence Tags genome project (SUCEST-FUN) (http://sucest.lad.ic.unicamp.br/en/) [25]; Sugarcane Gene Index (SGI); SUCAST catalogs and SUCAMET, which include expression data (http://sucest-fun.org); GRASSIUS database [26] and records of the agronomic, physiological and biochemical characteristics of sugarcane cultivars, were used.

BER component	% Identity (Sb/At)	Protein name	Substrate or function	Role in BER	SUCEST-FUN ID
DEM, ROS1 and DMLE	(94.54/54.88, 51.79, 55.05)*	DEMETER, Repressor of silencing 1 and DEMETER-like protein 3	5-methylcytosine (5-meC)	Involved in the initial stage of BER, recognizing the damaged base.	comp89337_c0_seq1
OGG1	(96.64/59.29)	8-oxoguanine-DNA glycosylase 1	8-oxoguanine (8-oxoG)		comp78469_c0_seq3
NTH1	(97.85/ 54.09, 49.31)**	Endonuclease III homolog 1	oxidized pyrimidine		comp79344_c0_seq1
UDG	(96/57)	Uracil-DNA glycosylase	Uracil		comp64547_c0_seq4 and SCEQFL5048B07.g
MBD4L	(88.3/47)	Methyl-CpG-binding domain protein 4-like protein	G:T mismatches within methylated and unmethylated CpG sites. Uracil or 5-fluorouracil in G:U mismatches.		comp78687_c0_seq2
MUTM (1 and 2)****	(94.49/68.64) (86.83/59.11)	Formamidopyrimidine-DNA glycosylase	oxidation products of 8-oxoguanine (8-oxoG)		comp85541_c0_seq1; SCCCLR2C01B12.g
ARP1 (1 and 3)****	(95.09/59.60) (97.42/71.79)	DNA-(apurinic or apyrimidinic site) endonuclease	Ap site	Repair by-products (AP site) of BER or oxidation.	comp79331_c0_seq10; comp86134_c0_seq5
FEN1 (A and B)****	(96/81.69) (82.97/73.07)	Flap endonuclease 1	5′ flap	Involved with the BER's long-patch.	comp79282_c1_seq1 and SCEPRZ1008D03.g; comp85461_c0_seq2 and comp79282_c1_seq1
Pol λ	(95.16/54.70)	DNA POLIMERASE LAMBDA	Resynthesize missing nucleotides	It replaces the Polymerase beta acting on the BER short-patch.	comp80417_c0_seq9
TDP1	(95/45.7)	Tyrosyl-DNA phosphodiesterase 1	Processing of diverse 3′- and 5′-blocking groups at DNA ends	Processing of intermediate BER products	comp89039_c1_seq9 and comp89039_c1_seq3
LIG1	(91.67/70.96)	DNA ligase 1	Seal 5′-PO4 and 3′-OH polynucleotide ends	Involved in the long and BER's short patch.	comp86584_c0_seq5 and comp86584_c0_seq7
LIG4	(97.53/73.25)	DNA ligase 4		Proposed to be involved in BER's short-patch.	comp85403_c0_seq6

BER component	% Identity (Sb/At)	Protein name	Substrate or function	Role in BER	SUCEST-FUN ID
ZDP	(94.8/43.2)	Polynucleotide 3′-phosphatase ZDP	3′-phosphopolynucleotide	Processing of intermediate BER products.	comp78030_c0_seq1
PCNA	(100/85.55 and 86.69) ***	PROLIFERATING CELL NUCLEAR ANTIGEN	A scaffold to recruit the proteins involved in DNA replication, DNA repair, chromatin remodeling, and epigenetics	Involved in the BER's long-patch.	comp82119_c0_seq2 and SCCCCL3140F04.g
PARP1	(96.7/60.6)	Poly [ADP-ribose] polymerase 1	Uses NAD+ as a substrate, synthesizes and transfers ADP-ribose onto aspartic and glutamic acid residues of acceptor proteins	Protects the BER substrate, present in the BER's long-patch.	comp82301_c0_seq8 and SCAGLB1070H02.g
PARP2	(95.5/53.1)	Poly [ADP-ribose] polymerase 2		Not essential for DNA repair in the BER pathway.	comp85410_c0_seq3 and SCJFRT1012D11.g
XRCC1	(97.11/48.2)	X-RAY REPAIR CROSSCOMPLEMENTING PROTEIN 1	interacting with APE1 and stimulating its AP endonuclease activity, prepares the DNA substrate for the DNA polymerase activities.	Involved in the BER's short-patch.	comp81667_c0_seq2, SCVPFL4C09E05.g and SCEZSD1082B05.g
WRN	(93.9/41.5)	WERNER SYNDROME ATP DEPENDENT HELICASE	Helicase enzyme	Interacts with several BER proteins: FEN1, PolB and PARP1	comp74108_c0_seq1, SCUTAM2089F01.g and SCSFLR2031F05.g

In the column % Identity (Sb/At) corresponds to the amino acid sequence identity of sugarcane protein with Sorgum bicolor *and* Arabidopsis thaliana *homologs are shown.*
**Only one sequence was found in* Sorghum bicolor *with high similarity to sugarcane sequence and Arabidopsis's DEM, ROS1 and DMLE.*
***There are two NTH in* Arabidopsis thaliana*: NTH1 (Q9SIC4) and NTH2 (B9DFZ0).*
****There are two PCNA in* Arabidopsis thaliana*: PCNA1 (Q9M7Q7) and PCNA2. (Q9ZW35).*
*****BER components that were identified sequence duplication in sugarcane genome.*

Table 1.
BER components from sugarcane.

Sugarcane BER components identified were compared with sequences belonging to *A. thaliana* and *Sorghum bicolor* (**Table 1**). Subsequently, these sequences were structurally and phylogenetically characterized; hence, their location on the pathway was set (**Table 1**). The following topics will address the particularities that were found relevant to BER and its specificities in sugarcane and monocots.

3. BER components—missing and differences

Sugarcane exhibits almost all components of the BER pathway, even though ROS1, DEM and DML3 were not identified as complete sequences. These DNA glycosylases, which play pivotal roles in epigenetic processes [27], have been well characterized [28–31] and were found in SUCEST-FUN as a single sequence without functional domains. Nevertheless, this result does not indicate a missing enzyme; epigenetic regulation is crucial, particularly in plants, and even more in polyploids organisms [32, 33]. Furthermore, the sugarcane database compiles numerous fragmented sequences that were not assembled and functionally annotated yet, as most data were from the transcriptome [25].

In contrast, differences were observed between the sequences of grasses analyzed (sugarcane and *S. bicolor*) when compared to dicotyledons, *A. thaliana* (**Table 2**). One of these differences, inconsistent with that observed in *A. thaliana,* was that the sequences of the *Poaceae* family present a second flap endonuclease protein 'FEN1B', which differs in size (as they are larger than the canonical Flap endonuclease 1 that receives the suffix A) and lacks the interaction sequence with PCNA. Notably, the sequences FEN1A and FEN1B are found at different *loci* and chromosomes of *S. bicolor*. Duplication in genes related to BER proteins was observed in AP endonucleases (ScARP1 and ScARP3) and MUMT (ScMUTM1 and ScMUTM2), which also reveal structural differences, as observed in FEN1A_CANA and FEN1B_CANA [34–36].

DNA ligase IV revealed certain differences regarding the domain disposition on the sequence (**Table 2**). Additionally, the sequences reveal variable identity (**Table 1**), thereby indicating high similarity within the grass plants. Notably, the BRCT domain is present in the sequences of *A. thaliana* and *S. bicolor*, but not in that of sugarcane. BRCT is a domain related to protein–protein interactions and is present in numerous proteins involved in DNA repair as well as cell cycle control [37–39]. Differences in domain disposition were also perceived in XRCC1, which displayed only one BRCT domain in the *Poaceae* family, whereas two BRCTs were found in the *A. thaliana* sequence. These differences could reflect variations in the protein role in DNA metabolism; these domains are essential because they comprise the activity and binding site of the enzyme.

4. BER's first step - base lesion recognition

BER is initiated by lesion-specific DNA glycosylases. The basic DNA glycosylase enzymatic process involves excision of the modified nucleobase from the DNA by catalyzing the hydrolysis of the N-glycosidic bond [40]. Regarding sugarcane, some of the BER's glycosylases were identified and characterized, suggesting the maintenance of the enzymes in *Saccharum spp.* as well as in conservation of the first step of BER pathway.

The DNA glycosylase OGG1 was identified in sugarcane and is called OGG1_CANA. This glycosylase as well as other sequences belonging to the *Poaceae* and the dicotyledonous, exhibit the conserved domain of the superfamily OGG1 [41].

Query	Accession	Protein domain	lenght (aa)
DML3_ARATH (O49498)	pfam15628	RRM_DME	1044
	cl23768	ENDO3c superfamily	
	cl21423	Perm-CXXC superfamily	
DME_ARATH (Q8LK56)	pfam15628	RRM_DME	1987
	cl23768	ENDO3c superfamily	
	pfam15629	Perm-CXXC	
	cl26620	Glutenin_hmw superfamily	
	cl34047	TonB superfamily	
ROS1_ARATH (Q9SJQ6)	pfam15628	RRM_DME	1393
	cl23768	ENDO3c superfamily	
	pfam15629	Perm-CXXC	
A0A1Z5R5E2_SORBI	pfam15628	RRM_DME	1878
	cl23768	ENDO3c superfamily	
	pfam15629	Perm-CXXC	
comp89337_c0_seq1	_	_	1469
FEN1_ARATH (O65251)	PF00867	N-domain	383
	PF00752	I-domain	
		Interaction with PCNA	
FEN1A_SORBI (C5YUK3)	PF00867	N-domain	380
	PF00752	I-domain	
		Interaction with PCNA	
FEN1B_SORBI (C5WU23)	PF00867	N-domain	428
	PF00752	I-domain	
FEN1A_CANA	PF00867	N-domain	379
	PF00752	I-domain	
		Interaction with PCNA	
FEN1B_CANA	PF00867	N-domain	413
	PF00752	I-domain	
DNLI4_ARATH (Q9LL84)	cl36689	dnl1 superfamily	1219
	cd17722	BRCT_DNA_ligase_IV_rpt1	
	cd17717	BRCT_DNA_ligase_IV_rpt2	
	cl31754	PTZ00121 superfamily	
A0A1Z5REU4_SORBI	cl36689	dnl1 superfamily	1281
	cd17722	BRCT_DNA_ligase_IV_rpt1	
	cl00038	BRCT superfamily	
	cl12940	DNA_ligase_IV superfamily	
DNLI4_CANA	cd07903	Adenylation_DNA_ligase_IV	572
	cl08424	OBF_DNA_ligase_family superfamily	
	pfam04675	DNA_ligase_A_N	

Query	Accession	Protein domain	lenght (aa)
XRCC1_ARATH	PRU00033	BRCT1	352
	PRU00033	BRCT2	
C5Z3V7_SORB	PS50172	BRCT	346
XRCC1_CANA	PS50172	BRCT	346

Table 2.
BER components with distinct features regarding protein domains in sugarcane.

This sequence reveals conservation of glutamine and phenylalanine residues (*Arabidopsis,* residues Q_{324} and F_{328}; sugarcane, Q_{378} and F_{382}) that are responsible for recognition of the damage base [42]. Moreover, site-directed mutagenesis assays in human OGG1 revealed that residues K_{249} and D_{268} (the sugarcane equivalent D_{334} and K_{315}) would also play an essential role in appropriate catalysis of DNA glycosylase [43, 44]. For MUTM, two sequences were identified in sugarcane: ScMUTM1 and ScMUTM2. Similar to OGG1_CANA, these sequences also retain essential residues for their enzymatic activity [36].

In *A. thaliana*, a homolog for Endonuclase III was identified and characterized, and termed as *Arabidopsis thaliana ENDONUCLEASE THREE HOMOLOG 1*(AtNTH1); it presented its enzymatic activity in relation to various substrates, thereby revealing its essential role in plant stress response [45]. A second endonuclease III homolog called AtNTH2, which was found together with AtNTH1 and AtARP in the *A. thaliana* chloroplast nucleus, demonstrating the occurrence of BER pathway in this organelle [46]. Considering grasses, a sequence that would refer to NTH2 remained unidentified. Phylogenetic analyses of this DNA glycosylase revealed duplication of sequences for organisms belonging to the group of dicots, but not for monocots.

Sugarcane NTH1, called NTH1_CANA, belongs to the Helix-hairpin-Helix (HHH) superfamily [47]. Furthermore, regarding *Escherichia coli*'s endonuclease III protein, the Helix-Hairpin-Helix domain has iron–sulfur binding sites [4Fe-4S] [48]. These sites comprised four conserved cysteines that would act on redox chemistry and DNA binding [49], and both motif and sites are conserved in the NTH1_CANA. Moreover, conservation of aspartic acid (D) at the active site, which is a residue preserved in other DNA glycosylases besides NTH1, such as UNG and MBD4L [50], was also evidenced in sugarcane.

Another glycosylase identified was UDG_CANA, which was conserved in the domain belonging to the UDG superfamily, more precisely concerning family-1 [51, 52]. Additionally, it conserved aspartic acid (D) as an active site [51]. It is known that the human UNG gene encodes two forms of the protein, one directed towards the mitochondria (UNG1) and another towards the nucleus (UNG2) [53]. The *A. thaliana* UNG (AtUNG) seems to be homologous to these two types of UNGs, being proven to act on mitochondrial DNA [54]. Most grass sequence annotations of computational prediction that directed the UNGs to both the nucleus and the mitochondria, raised the question whether there is only one UNG for both organelles in plants.

Ramiro-Merina et al. [55] demonstrated that *A. thaliana* encodes a monofunctional DNA glycosylase homologous to mammalian MBD4, known as MBD4-like or AtMBD4L. Nota et al. [56] indicated that the activation of AtMBD4L induces the expression of a late gene from the BER AtLIG1 pathway and reveals the mechanism by which it increases the plant's tolerance to oxidative stress. In relation to sugarcane, one fragment features the same domain and active site as AtMBD4L, implying a probable functional protein in *Saccharum spp.*

5. AP site removal—AP endonuclease role in sugarcane

AP endonuclease is an essential enzyme for BER pathway as this enzyme identifies and process AP (apurinic/apyrimidinic) site [57]. These AP sites may be a result of the action of DNA glycosylases or it may be spontaneously generated. Unrepaired AP sites can lead to mutations during semiconservative replication, which indicates the importance of the role of AP endonuclease in maintenance of the genetic code [58].

In *A. thaliana*, three AP endonucleases are homologous to APE1 (HUMAN AP ENDONUCLEASE 1), namely AtAPE1L, AtAPE2, and AtARP [59]. Each of these presents their specifics based on enzymatic activity, regulation, and sub-cellular localization. For sugarcane, two sequences were identified, and their three-dimensional structures were inferred: ScARP1 and ScARP3 [35]. By examining the sequences of ScARPs (1 and 3), we can observe the conservation of essential sites for the catalysis and binding of metals (enzymatic cofactors) [34].

ScARP1 has greater similarity with AtARP (60%), whereas AtAPE1L and AtAPE2 reveal a correspondence below 50%. These values may indicate diversity in structure, amino acid composition, and perhaps function. ScARP3 reveals a divergence compared with ScARP1. ScARP3 is closer to AtARP, presenting an even higher percentage of identity (75%). Maíra et al. [35] demonstrated that the sequence ScARP3 would be closer to the group of dicotyledonous plants, whereas ScARP1 would be included within the monocots, more precisely together with representatives of the *Poaceae* family.

Medeiros et al. [34] purified ScARP1 and verified the enzymatic activity of this sugarcane enzyme against several substrates. This study found the capability of ScARP1 to process AP sites; however, other enzymatic activities (exonuclease, phosphatase, and 3′-phosphodiesterase) were not confirmed. The AP endonuclease activity complementation assay in extracts of *A. thaliana* demonstrated that ScARP1 was capable of complementing around 40% of the activity of AtARP from $arp^{-/-}$ mutant plant extracts [34].

6. Flap endonuclease (FEN1)—BER's long-patch in sugarcane

FEN1 is a structure-specific nuclease that can remove flap structures and is involved in different DNA metabolic pathways, including DNA replication, DNA repair, apoptotic DNA degradation, and maintenance of telomere stability [60]. In case of BER, FEN1 in complex with proliferating cell nuclear antigen (PCNA) plays a pivotal role in the long patch as it removes a short flap structure generated by Pol β activity [61].

Regarding plants, it is known that two FEN1 counterparts were identified in rice (*Oryza sativa*, OsFEN1a, OsFEN1b). Functional complementation assays revealed that only OsFEN1a would be able to complement the *fen1/rad27* mutant in yeast, suggesting that these two genes may be functionally distinct [62]. In addition, OsFEN1a, expressed in *Escherichia coli*, presents flap-endonuclease and 5′-exonuclease activity [63]. In *A. thaliana*, only one FEN1 homolog, namely SAV6, was identified [63]. Biochemical characterization of SAV6 protein (also called FEN1) revealed that, unlike animal FEN1, the SAV6 protein has flap-endonuclease and gap endonuclease activity but does not reveal 5′ exonuclease activity; however, similar to human FEN1 (hFEN1), SAV6 is also necessary for maintaining the genome integrity and responding to plant DNA damage [64].

As observed in *O. sativa*, sugarcane has two FEN1 sequences, namely FEN1A_CANA and FEN1B_CANA. Considering the structure of the Flap endonuclease, it is known that

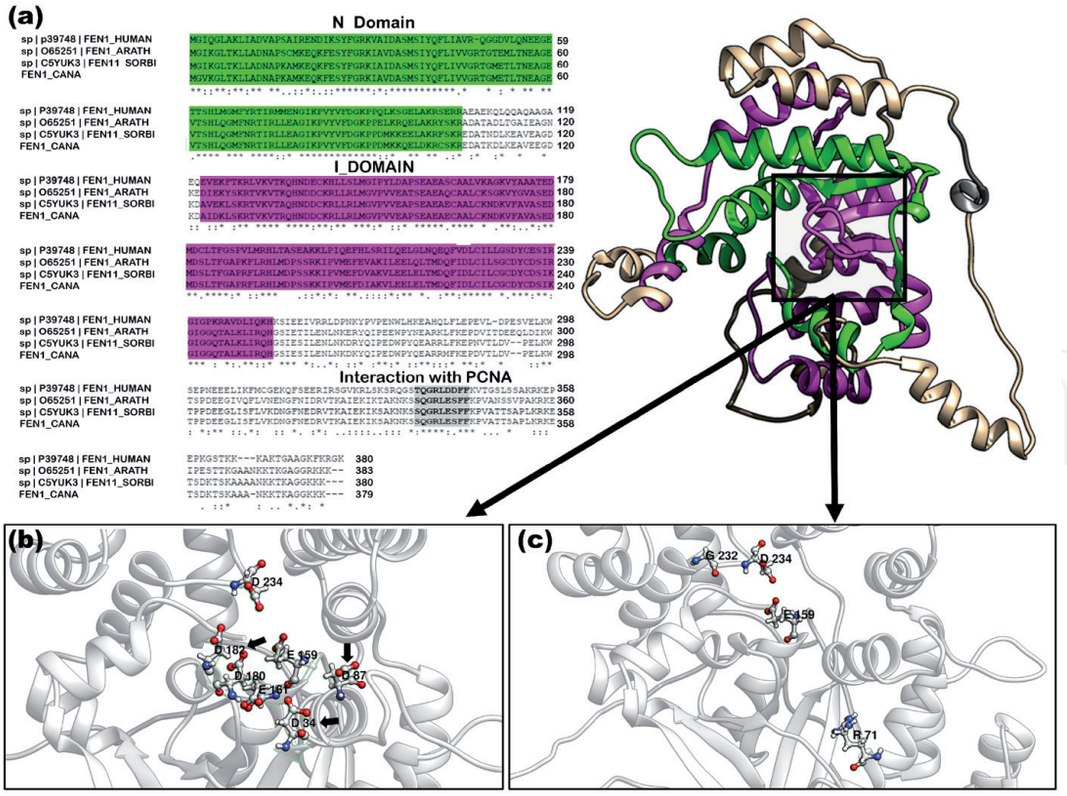

Figure 1.
The proposed model for FEN1 of sugarcane. (a) It was represented the alignment obtained using Clustal omega for FEN1 sequences of Homo sapiens*,* Arabidopsis thaliana, Sorghum bicolor *and sugarcane. The colors in the alignment and in the model, correspond to the N-terminal region (green), internal region (I) (purple) and the segment that interacted with PCNA (gray). Metal-binding sites (b) and DNA binding sites (c) are highlighted. The black arrows in (b) indicate the probable active site of the enzyme.*

human FEN1 comprises the N-terminal domain and the intermediate (Domain I) in addition to a C-terminal region, which is important for the interaction of FEN1 with other proteins, such as PCNA and WRN (**Figure 1**) [65, 66]. The FEN1A_CANA sequence preserves the domains described previously; however, the FEN1B_CANA and other *Poaceae* similar sequences analyzed do not possess the binding domain for PCNA in its C-terminal region, which may affect its mechanism of action in the plant cell.

Considering the protein structure, FEN1 is a nuclease that features two regions: the N-terminal region and I-region [67]. The alignment of FEN1 from *H. sapiens*, *A. thaliana* and *S. bicolor* as well as FEN1A's sugarcane ascertained the conservation of these regions (**Figure 1**). Notably, the region of interaction with the PCNA (**Figure 1a**) that is in a loop, in that way, more exposed and facilitating its possible interaction with PCNA. The sugarcane FEN1A model presents the conservation of metal-binding sites (Mg^{+2} ion; **Figures 1b** and **c**). The residues D_{34}, D_{87}, and D_{182}, considering the equivalent residues in human FEN1 [68], may be responsible for the catalytic activity of the enzyme (**Figure 1b** and **c**).

7. PCNA role in plants

Studies on PCNA have revealed that it plays a crucial role in DNA replication as well as in DNA repair, cell cycle regulation and apoptosis [69–71]. In *A. thaliana*, two PCNAs, AtPCNA1 and AtPCNA2, are present, which differ from each other in eight amino acids, in addition to the fact that AtPCNA2 has an extra residue in the protein length [72]. Of these eight different amino acids, four are identical to the residues found in *Brassica napus* and human PCNAs [73].

Considering the difference between AtPCNAs, Anderson et al. [72] demonstrated that co-expression of POLH (DNA polymerase eta - Pol η) and AtPCNA2 (and not AtPCNA1) was necessary to restore normal resistance to UV radiation in the yeast RAD30 mutant. The difference was in lysine (K) 201 present in AtPCNA1, which would inhibit the ubiquination of lysine 164, thus affecting its connection with Pol η and not being able to act on trans-lesion synthesis (TLS) and restore the progression of the replication fork. The lysine at position 201 of AtPCNA1 belonged to the group comprising amino acids with electrically charged side chains. In the case of K, this could be endowed with a positive charge, whereas the corresponding one at AtPCNA2 would be an asparagine (N) that belonged to the group of amino acids with polar side chains without being loaded. In PCNA_CANA, the corresponding residue in question would be a glutamine that concerns the same group as N, which leads to the conclusion that sugarcane PCNA would be closer to AtPCNA2 than AtPCNA1 and could, as such, act in the TLS.

The three-dimensional model of sugarcane's PCNA is revealed as homotrimeric architecture in the form of a ring, comprising three identical chains of PCNA, as indicated by different colors in the sugarcane PCNA model (**Figure 2**).

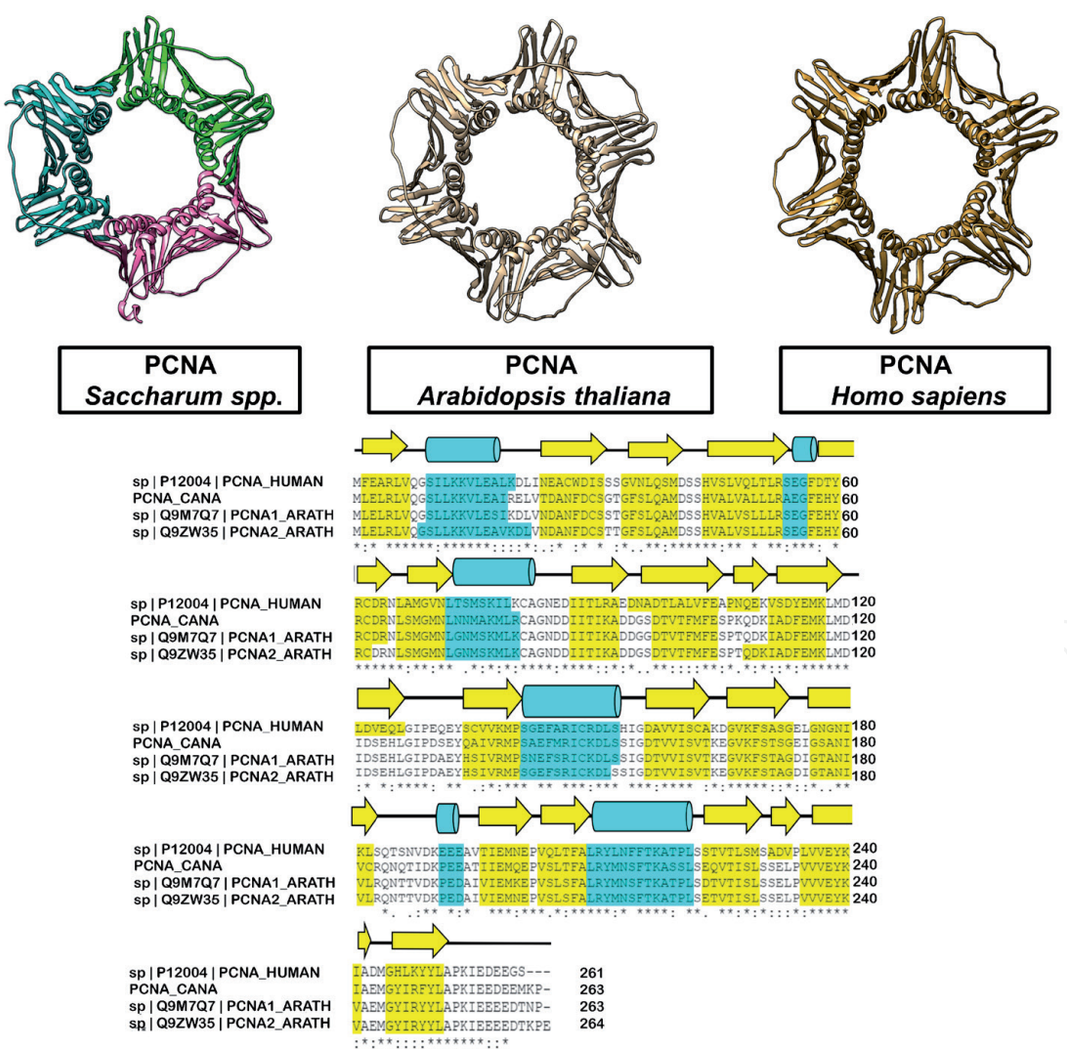

Figure 2.
Three-dimensional model and protein sequences of plant and human PCNAs. The 3D models are depicted above the alignment presenting conservation of the structure in ring-shaped homotrimeric architectures. The model of the putative sugarcane PCNA; the structures highlighted in blue, green and pink are individual chains of PCNA that together compose the homotrimeric ring. Below the models, the corresponding alignment of the PCNA sequences of Arabidopsis thaliana, Homo sapiens *and* Saccharum spp.*, highlighting the secondary structures (yellow arrow, beta sheet; blue cylinder, alfa helix) is presented.*

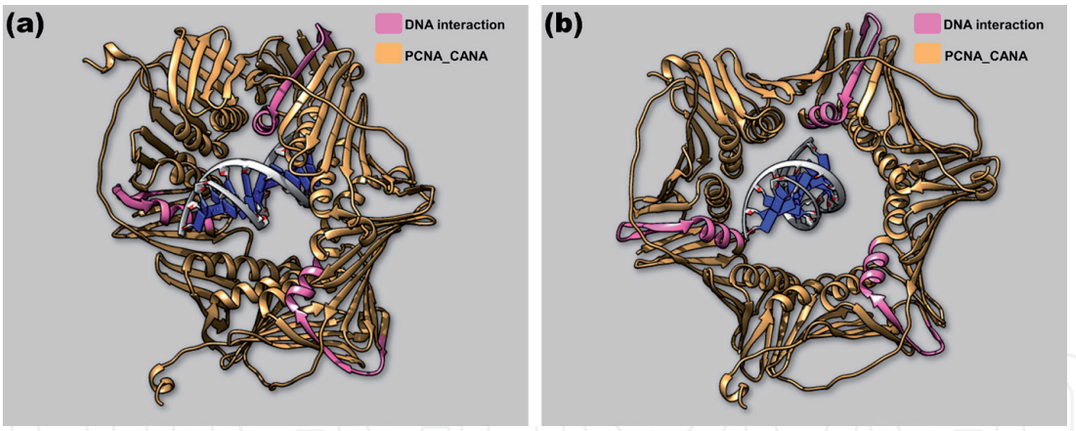

Figure 3.
Models proposed for sugarcane PCNA associated with DNA. The region of the PCNA that interacts with the DNA, facing the inside of the ring of the homotrimeric structure, is depicted in pink. The double strand that constitutes a helix of predominantly blue color represents the three-dimensional structure of DNA. (a) View of the sugarcane PCNA model (in orange) interacting with the DNA (structure in blue and white) seen from the side. (b) Frontal view of the model.

Additionally, this model exhibits sequence and structural similarity with other PCNAs, as observed in **Figure 2**. Compared to *A. thaliana* and *H. sapiens* PCNAs, sugarcane's PCNA overlapped its secondary structure. Structure conservation is also observed in the PCNA models (**Figure 2**), which display the same homotrimeric ring predicted for the sugarcane model.

PCNA is generally called a sliding clamp, since it was predicted that the double strand of DNA would pass through the opening of the PCNA ring and would serve as an anchoring platform for several proteins involved in DNA metabolism [74]. The attainable preservation of this function was verified in the *Saccharum spp*. model. The 5L7C crystal [75], which is a model of human PCNA, was used for comparison with PCNA's sugarcane. By aligning the crystal with the model (root mean square deviation, RMSD = 0.653 Å), it was possible to ascertain the probable conservation of its interaction with the DNA (depicted in pink). Therefore, we identified a DNA binding site in the model, which faces the interior of the ring orifice, where the double strand of DNA should pass (depicted in blue) (**Figure 3**).

8. Sugarcane protein models—conservation throughout plants

PCNA and FEN1A were proteins identified in sugarcane, which were presumed to interact with each other [76]. This is due to PCNA interaction sequence detected in the N-terminal segment of FEN1A. To verify the truthfulness of this interaction, three-dimensional models were created for PCNA and FEN1A sugarcane proteins. These models were assessed for the conservation of secondary structure, active sites, and residue interactions with the substrate. Based on this analysis, the role of these sugarcane proteins can be established.

PCNA, as previously mentioned, would serve as a scaffold, and moreover, various functions can be performed ranging from DNA methylation to base excision repair. Thus, using the IUL1 crystal that comprises the human PCNA associated with FEN1 [77], the possibility of the sugarcane's predicted models of these proteins that may interact with each other was verified. This result demonstrates that FEN1 of sugarcane is associated with the homotrimeric ring of PCNA (**Figure 4**). The sequence of interaction with PCNA differs, revealing that this sequence is in the interface of PCNA and FEN1 interaction.

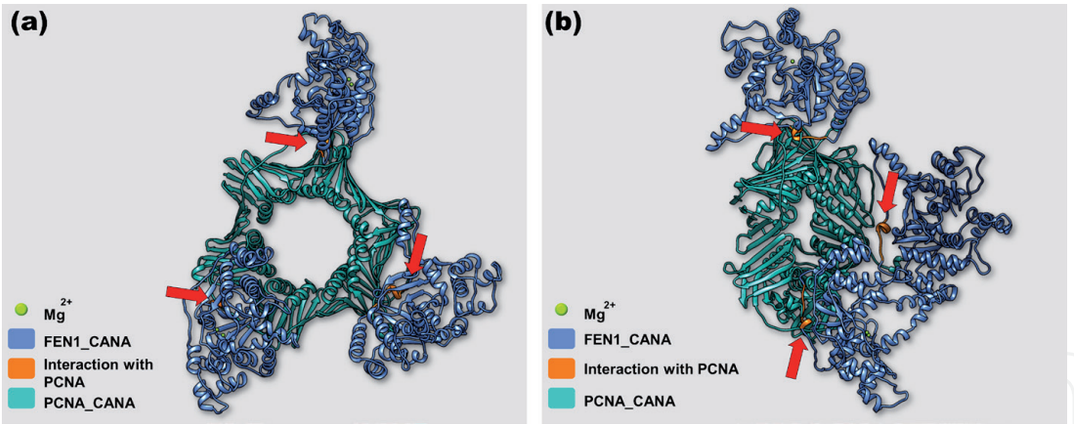

Figure 4.
Proposed complex of sugarcane PCNA and FEN1. (a) Lateral view of the complex. (b) Frontal view of the complex. FEN1 models are in blue, PCNA in green and Mg^{+2} ions are highlighted in yellow. The sequence of FEN1 that interacts with PCNA is highlighted in orange and is indicated with red arrows.

9. BER pathway—evolutionary analysis in the grass outlook

Overall, the phylogenetic analyses revealed differences in the presence or absence of duplication of BER pathway components. In few cases, duplication was observed in dicotyledons and not in monocotyledons, for example, NTH, PCNA and DNA ligase 1. Herein, structural difference was noted (size, presence or absence of certain conserved domains), indicating diverse DNA repair mechanisms between plants.

Singh et al. [78] compared the plant genomes available at that time, thus aiming to compare the genes involved in DNA repair and recombination. They found that FEN1, in the genome of monocotyledons (corn, rice, *S. bicolor*, and *Brachypodium distachyon*) presented two copies and that such copies would not be products of intra-genomic duplication. In particular, these copies were subtypes of FEN1, FEN1A and FEN1B. Singh et al. [78] also identified one copy of FEN1 in dicots, namely *A. thaliana*, *Medicago truncatula*, *Vitis venifera*, and *Papaver somniferum*; however, *Glycine max* presented two copies of FEN1; in such case, these copies were products of intra-genomic duplication.

A new analysis regarding FEN1 in plants, particularly sugarcane, was conducted. It was discovered that FEN1B was only found within *Poales*, specifically *Panicoideae* (**Figure 5**). Important crops such as *Oryza sativa*, *Zea mays* and *S. bicolor* display FEN1B as well as FEN1A. Evolutionary analyses revealed that FEN1A and FEN1B had distinct assembly. Moreover, the flap endonucleases (FEN1A and FEN1B) of the same species were not located at the same branch in the phylogenetic tree. Nonetheless, FEN1 was duplicated in some eudicot groups, as in *Noccaea caerulescens* and *Nicotiana tabacum*; however, these sequences have all the regions required for a functional FEN1.

Although the absence of region may compromise the enzymatic activity of FEN1B, the other residues, domains and active sites were conserved. These findings raise questions regarding the maintenance of FEN1B in the genome of these organisms, its functions and its role in BER.

Maíra et al. [35] proposed that a whole genome duplication event (WGD) would be related to the duplication observed in the AP endonuclease sequence in the grasses group; however, further studies indicate that duplications are present in other plant groups in addition to *Poaceae*. The BER's duplication genes do not cover all the components of this pathway; on the contrary, a few sequences—ARP, MUTM and FEN1—could be set as duplications. Issues regarding the maintenance

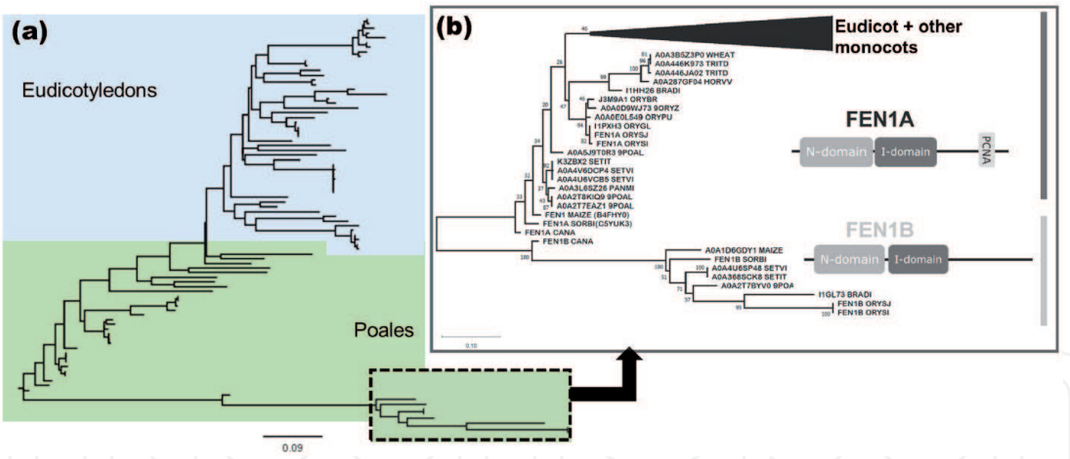

Figure 5.
FEN1 evolutionary analysis by maximum likelihood method. The evolutionary history was inferred using the maximum likelihood method and JTT matrix-based model conducted in MEGA X. the percentage of trees in which the associated taxa clustered together is presented next to the branches. (a) Phylogenetic tree comprising plant FEN1 sequences; green color represents the Poales group and blue color represents the eudicotyledons group. (b) Phylogenetic tree focus on Poales group, in which FEN1A and FEN1B clusters are displayed on distinct branches. The FEN1A and FEN1B domains are displayed next to their respective clusters.

of these sequences in plant genomes, particularly sugarcane, need to be responded to essentially comprehend the evolutionary aspect of the BER pathway in monocots.

Notably, the fate of the vast majority of duplicate genes resulting from segmental duplication includes the nonfunctionalization of a member of the pair [79, 80], which should occur within a few million years in the absence of any intrinsic advantage of duplicate copying [80, 81]. Specifically, plant genomes, on average, reveal 65% of their annotated genes that are duplicated [82]. Most of these copies are derived from ancient WGD events in the terrestrial plant lineage [82]. Li et al. [83] investigated the fate of duplicate genes from 40 different species of flowering plants; of these, all species experienced at least one or more WGD events throughout their evolutionary history. The loss of genes was observed immediately after genome duplication, so that the genes quickly returned to the state of a single copy [83]; however, some of these genes have preserved their state of multiple copies. Such genes belong to families of genes involved in the response to biotic and abiotic stress, and are therefore important for the adaptation of the plant to the environment. Thus, it is possible to correlate the duplication and retention of these copies with an adaptive advantage such that genes can confer to the plant, allowing it to act more efficiently in response to environmental variations. DNA repair genes are linked to this hypothesis, since they are necessary to maintain the stability of the genome and preserve genetic information. In addition to the fact that several of these genes, already described in this chapter, act in other processes of adaptive importance such as response to oxidative stress.

10. Conclusions

In sugarcane as well as in other plants, except for the plant models, few studies have focused on the characterization and structural analysis of individual components of metabolic pathways. Moreover, it should be considered that the traditional breeding strategy lags behind the demand for commercial needs due to insufficient knowledge on characteristics related to stress tolerance, inefficient selection techniques and low genetic variation and fertility. The evident deficiency of biotechnology will be

supplemented with studies aimed at the biochemical and functional characterization of important pathways and their components, such as the DNA repair pathway, for instance, the BER pathway. Therefore, it is necessary to emphasize the importance of this chapter in other plant species apart from sugarcane, provide supplementary information, and raise questions on the components of the BER pathway and its evolutionary issue regarding monocots and dicots.

Acknowledgements

The authors wish to thank to Coordenação de Aperfeiçoamento Pessoal de Nível Superior-CAPES, Conselho Nacional de Desenvolvimento Científico e Tecnológico-CNPq, and Ministério de Ciência, Tecnologia, Inovação (MCTI) for financial support.

Conflict of interest

The authors declare no conflict of interest.

Author details

Nathalia Maíra Cabral de Medeiros and Katia Castanho Scortecci*
Departamento de Biologia Celular e Genética – Centro de Biociencias, Universidade Federal do Rio Grande do Norte, Brazil

*Address all correspondence to: kacscort@yahoo.com

References

[1] Wallace SS. Base Excision Repair: A Critical Player in Many Games. DNA Repair. 2014;19:14-26. DOI: 10.1016/j.dnarep.2014.03.030.

[2] Lindahl T, Wood RD. Quality Control by DNA Repair. Science. 1999;286:1897-1905. DOI: 10.1126/science.286.5446.1897.

[3] Krokan HE, Bjørås M. Base excision repair. Cold Spring Harb Perspect Biol. 2013;5:a012583. DOI: 10.1101/cshperspect.a012583.

[4] Drohat AC, Coey CT. Role of Base Excision "repair" Enzymes in Erasing Epigenetic Marks from DNA. Chemical Reviews. 2016;166:12711-12729. DOI: 10.1021 acs.chemrev.6b00191.

[5] Kohli RM, Zhang Y. TET Enzymes, TDG and the Dynamics of DNA Demethylation. Nature. 2013;502:472-9 DOI: 10.1038/nature12750.

[6] Fleming AM, Ding Y, Burrows CJ. Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. Proc Natl Acad Sci U S A. 2017;114:2604-2609. DOI:10.1073/pnas.1619809114.

[7] Bellacosa A, Drohat AC. Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites. DNA Repair (Amst). 2015; 32:33-42. DOI: 10.1016/j.dnarep.2015.04.011.

[8] Roldan-Arjona T, Ariza RR, Córdoba-Cañero D. DNA base excision repair in plants: an unfolding story with familiar and novel characters. Frontiers in plant science. 2019;10:1055. DOI: 10.3389/fpls.2019.01055.

[9] Kimura S, Sakaguchi K. DNA repair in plants. Chemical Reviews. 2006;106:753-766. DOI: 10.1021/cr040482n.

[10] Britt AB. DNA damage and repair in plants. Annu Rev Plant Physiol Plant Mol Biol. 1996;47:75-100. DOI: 10.1146/annurev.arplant.47.1.75.

[11] Manova V, Gruszka D. DNA damage and repair in plants – From models to crops. Frontiers in Plant Science. 2015;6:885. DOI: 10.3389/fpls.2015.00885.

[12] Sarwar H, Sarwar MF, Sarwar M, Qadri NA, Moghal S. The importance of cereals (Poaceae: Gramineae) nutrition in human health: A review. Journal of cereals and oilseeds. 2013;4:32-35 .DOI: 10.5897/jco12.023.

[13] FAO. Food and Agriculture Data [Internet].2018. Available from: http://www.fao.org/faostat/en/#home [Accessed: 2020-06-19].

[14] Thirugnanasambandam PP, Hoang N V., Henry RJ. The challenge of analyzing the sugarcane genome. Frontiers in Plant Science. 2018;9:616. DOI: 10.3389/fpls.2018.00616.

[15] Souza GM, Berges H, Bocs S, Casu R, D'Hont A, Ferreira JE, Henry R, Ming R, Potier B, Van Sluys MA, Vincentz M, Paterson AH. The Sugarcane Genome Challenge: Strategies for Sequencing a Highly Complex Genome. Trop Plant Biol. 2011;4:145-156. DOI: 10.1007/s12042-011-9079-0.

[16] Dal-Bianco M, Carneiro MS, Hotta CT, Chapola RG, Hoffmann HP, Garcia AAF, Souza GM. Sugarcane improvement: How far can we go? Current Opinion in Biotechnology. 2012;23:265-270. DOI: 10.1016/j.copbio.2011.09.002.

[17] Lima WC, Medina-Silva R, Galhardo RS, Menck CFM. Distribution of DNA repair-related ESTs in sugarcane. Genet Mol Biol. 2001;24:141-146. DOI: 10.1590/S1415-47572001000100019.

[18] Costa RMA, Lima WC, Vogel CIG, Berra CM, Luche DD, Medina-Silva R, Galhardo RS, Menck CFM, Oliveira VR. DNA repair-related genes in sugarcane expressed sequence tags (ESTs). Genet Mol Biol. 2001;24:131-140. DOI: 10.1590/S1415-47572001000100018.

[19] Agnez-Lima LF, de Medeiros SRB, Maggi BS, Quaresma GAS. Base excision repair in sugarcane. Genet Mol Biol. 2001;24:123-129. DOI: 10.1590/S1415-47572001000100017.

[20] Okura VK, de Souza RSC, de Siqueira Tada SF, Arruda P. BAC-pool sequencing and assembly of 19 Mb of the complex sugarcane genome. Front Plant Sci. 2016;7:342. DOI: 10.3389/fpls.2016.00342.

[21] De Mendonça VM, Del Bem LE, Van Sluys MA, De Setta N, Kitajima JP, Cruz GMQ, Sforça DA, De Souza AP, Ferreira PCG, Grativol C, Cardoso-SilvaCB, VicentiniR, VincentzM. Analysis of Three Sugarcane Homo/Homeologous Regions Suggests Independent Polyploidization Events of *Saccharum officinarum* and *Saccharum spontaneum*. Genome Biol Evol. 2017;9:266-278, DOI: 10.1093/gbe/evw293.

[22] Grativol C, Regulski M, Bertalan M, McCombie WR, Da Silva FR, Zerlotini Neto A, Vicentini R, Farinelli L, Hemerly AS, Martienssen RA, Ferreira PCG. Sugarcane genome sequencing by methylation filtration provides tools for genomic research in the genus Saccharum. Plant J. 2014;79:162-172. DOI: 10.1111/tpj.12539.

[23] Riaño-Pachón DM, Mattiello L. Draft genome sequencing of the sugarcane hybrid SP80-3280. F1000Research. 2017;6:861. DOI: 10.12688/f1000research.11859.2

[24] Piperidis N, D'Hont A. Sugarcane genome architecture decrypted with chromosome-specific oligo probes. Plant J. 2020;103:2039-2051. DOI: 10.1111/tpj.14881.

[25] Vettore AL, da Silva FR, Kemper EL, Souza GM, da Silva AM, Ferro MIT, Henrique-Silva F, Giglioti ÉA, Lemos MVF, Coutinho LL, Nobrega MP, Carrer H, França SC, Bacci M, Goldman MHS, Gomes SL, NunesLR, CamargoLEA, SiqueiraWJ, Van Sluys MA, Thiemann OH, Kuramae EE, Santelli R V., Marino CL, Targon MLPN, Ferro JA, Silveira HCS, Marini DC, Lemos EGM, Monteiro-Vitorello CB, Tambor JHM, Carraro DM, Roberto PG, MartinsVG, GoldmanGH, deOliveiraRC, Truffi D, Colombo CA, Rossi M, de Araujo PG, Sculaccio SA, Angella A, Lima MMA, de Rosa VE, Siviero F, Coscrato VE, Machado MA, Grivet L, Di Mauro SMZ, Nobrega FG, Menck CFM, Braga MDV, Telles GP, Cara FAA, Pedrosa G, Meidanis J, Arruda P. Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. Genome Res. 2003;13:2725-2735. DOI:10.1101/gr.1532103.

[26] Yilmaz A, Nishiyama MY, Fuentes BG, Souza GM, Janies D, Gray J, Grotewold E. GRASSIUS: A platform for comparative regulatory genomics across the grasses. Plant Physiol. 2009;149:171-180. DOI:10.1104/pp.108.128579.

[27] Liu R, Lang Z. The mechanism and function of active DNA demethylation in plants. Journal of Integrative Plant Biology. 2020;62:148-159. DOI:10.1111/jipb.12879.

[28] Agius F, Kapoor A, Zhu JK. Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. Proc Natl Acad Sci U S A. 2006;103:11796-11801. DOI:10.1073/pnas.0603563103.

[29] Tang K, Lang Z, Zhang H, Zhu JK. The DNA demethylase ROS1 targets genomic regions with distinct chromatin

modifications. Nat Plants. 2016;2:1-10. DOI:10.1038/nplants.2016.169.

[30] Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marín MI, Martínez-Macías MI, Ariza RR, Roldán-Arjona T. DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. Proc Natl Acad Sci U S A. 2006;103:6853-6858. DOI:10.1073/pnas.0601109103.

[31] Liu R, How-Kit A, Stammitti L, Teyssier E, Rolin D, Mortain-Bertrand A, Halle S, Liu M, Kong J, Wu C, Degraeve-Guibault C, Chapman NH, Maucourt M, Hodgman TC, Tost J, Bouzayen M, Hong Y, Seymour GB, Giovannoni JJ, Gallusci P. A DEMETER-like DNA demethylase governs tomato fruit ripening. Proc Natl Acad Sci U S A. 2015;112:10804-10809. DOI:10.1073/pnas.1503362112.

[32] Ding M, Chen ZJ. Epigenetic perspectives on the evolution and domestication of polyploid plant and crops. Current Opinion in Plant Biology. 2018;42: 37-48. DOI:10.1016/j.pbi.2018.02.003.

[33] Mason AS, Wendel JF. Homoeologous Exchanges, Segmental Allopolyploidy, and Polyploid Genome Evolution. Frontiers in Genetics. 2020;11:1014. DOI:10.3389/fgene.2020.01014.

[34] Medeiros NMC, Córdoba-Cañero D, García-Gil CB, Ariza RR, Roldán-Arjona T, Scortecci KC. Characterization of an AP endonuclease from sugarcane – ScARP1. Biochem Biophys Res Commun. 2019;514:926-932. DOI: 10.1016/j.bbrc.2019.04.156

[35] Maira N, Torres TM, de Oliveira AL, de Medeiros SRB, Agnez-Lima LF, Lima JPMS, Scortecci KC. 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 Biol. 2014;16:622-631. DOI:10.1111/plb.12083.

[36] Scortecci KC, Lima AFO, Carvalho FM, Silva UB, Agnez-Lima LF, de Medeiros SRB. A characterization of a MutM/Fpg ortholog in sugarcane-A monocot plant. Biochem Biophys Res Commun. 2007;361:1054-1060. DOI: 10.1016/j.bbrc.2007.07.134.

[37] Callebaut I, Mornon JP. From BRCA1 to RAP1: A widespread BRCT module closely associated with DNA repair. FEBS Lett. 1997;400:25-30. DOI: 10.1016./S0014-5793(96)01312-9.

[38] Reubens MC, Rozenzhak S, Russell P. Multi-BRCT Domain Protein Brc1 Links Rhp18/Rad18 and γH2A To Maintain Genome Stability during S Phase. Mol Cell Biol. 2017;37:e00260-17. DOI: 10.1128/MCB.00260-17.

[39] Polo LM, Xu Y, Hornyak P, Garces F, Zeng Z, Hailstone R, Matthews SJ, Caldecott KW, Oliver AW, Pearl LH. Efficient Single-Strand Break Repair Requires Binding to Both Poly(ADP-Ribose) and DNA by the Central BRCT Domain of XRCC1. Cell Rep. 2019;26:573-581. DOI:10.1016/j.celrep.2018.12.082.

[40] Krokan HE, Standal R, Slupphaug G. DNA glycosylases in the base excision repair of DNA. Biochemical Journal. 1997;325:1-16. DOI:10.1042/bj3250001.

[41] Dany AL, Tissier A. A functional OGG1 homologue from *Arabidopsis thaliana*. Mol Gen Genet. 2001; 265:293-301. DOI:10.1007/s004380000414.

[42] García-Ortiz MV, Ariza RR, Roldán-Arjona T. An OGG1 orthologue encoding a functional 8-oxoguanine DNA glycosylase/lyase in *Arabidopsis thaliana*. Plant Mol Biol. 2001;47:795-804. DOI: 10.1023/A:1013644026132.

[43] Nash HM, Lu R, Lane WS, Verdine GL. The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: Direct identification, ablation and chemical reconstitution. Chem Biol. 1997;4:693-702. DOI: 10.1016/s1074-5521(97)90225-8.

[44] Bjørås M, Seeberg E, Luna L, Pearl LH, Barrett TE. Reciprocal "flipping" underlies substrate recognition and catalytic activation by the human 8-oxo-guanine DNA glycosylase. J Mol Biol. 2002;317:171-177. DOI:10.1006/jmbi.2002.5400.

[45] Roldan-Arjona T, Garcia-Ortiz M V., Ruiz-Rubio M, Ariza RR. cDNA cloning, expression and functional characterization of an *Arabidopsis thaliana* homologue of the *Escherichia coli* DNA repair enzyme endonuclease IIIcDNA cloning, expression and functional characterization of an *Arabidopsis thaliana* homologue of the *Escherichia coli* DNA repair enzyme endonuclease III. Plant Mol Biol. 2000;44:43-52. DOI: 10.1023/A:1006429114451.

[46] Gutman BL, Niyogi KK. Evidence for base excision repair of oxidative DNA damage in chloroplasts of *Arabidopsis thaliana*evidence for base excision repair of oxidative DNA damage in chloroplasts of *Arabidopsis thaliana*. J Biol Chem. 2009;284:17006-17012. DOI: 10.1074/jbc.M109.008342.

[47] Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. Cell. 2002;110:33-42. DOI:10.1016/S0092-8674(02)00807-3.

[48] Kuo CF, McRee DE, Cunningham RP, Tainer JA. Crystallization and crystallographic characterization of the iron-sulfur-containing DNA-repair enzyme endonuclease III from *Escherichia coli*. J Mol Biol. 1992;227:347-351. DOI:10.1016/0022-2836(92)90703-M.

[49] Thayer MM, Ahern H, Xing D, Cunningham RP, Tainer JA. Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. EMBO J. 1995;14:4108-4120. DOI: 10.1002/j.1460-2075.1995.tb00083.x.

[50] Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. Cell. 2002;110:33-42. DOI: 10.1016/S0092-8674(02)00807-3.

[51] Pearl LH. Structure and function in the uracil-DNA glycosylase superfamily. Mutat Res - DNA Repair. 2000;460:165-181. DOI:10.1016/S0921-8777(00)00025-2.

[52] Lee HW, Dominy BN, Cao W. New family of deamination repair enzymes in Uracil-DNA glycosylase superfamily. J Biol Chem. 2011;286: 31282-31287. DOI:10.1074/jbc.M111.249524.

[53] Krokan HE, Kavli B, Sarno A, Slupphaug G. Enzymology of Genomic Uracil Repair. Genomic Uracil. 2018; 89-126. DOI:10.1142/9789813233508_0004.

[54] Boesch P, Ibrahim N, Paulus F, Cosset A, Tarasenko V, Dietrich A. Plant mitochondria possess a short-patch base excision DNA repair pathway. Nucleic Acids Res. 2009; 37:5690-5700. DOI:10.1093/nar/gkp606.

[55] Ramiro-Merina Á, Ariza RR, Roldán-Arjona T. Molecular characterization of a putative plant homolog of MBD4 DNA glycosylase. DNA Repair (Amst). 2013;12:890-898. DOI:10.1016/j.dnarep.2013.08.002.

[56] Nota F, Cambiagno DA, Ribone P, Alvarez ME. Expression and function of AtMBD4L, the single gene encoding the nuclear DNA glycosylase MBD4L in Arabidopsis. Plant Sci. 2015;235:122-129. DOI:10.1016/j.plantsci.2015.03.011.

[57] Wiederhold L, Leppard JB, Kedar P, Karimi-Busheri F, Rasouli-Nia A, Weinfeld M, Tomkinson AE, Izumi T, Prasad R, Wilson SH, Mitra S, Hazra TK. AP endonuclease-independent DNA base excision repair in human cells. Mol Cell. 2004;15:209-220. DOI:10.1016/j.molcel.2004.06.003.

[58] Talpaert-Borlè M. Formation, detection and repair of AP sites. Mutat Res - Fundam Mol Mech Mutagen. 1987;181:45-56. DOI:10.1074/jbc.M111.249524.

[59] Murphy TM, Belmonte M, Shu S, Britt AB, Hatteroth J. Requirement for abasic endonuclease gene homologues in Arabidopsis seed development. PLoS One. 2009;4:e4297. DOI: 10.1371/journal.pone.0004297.

[60] Tsutakawa SE, Classen S, Chapados BR, Arvai AS, Finger LD, Guenther G, Tomlinson CG, Thompson P, Sarker AH, Shen B, Cooper PK, Grasby JA, Tainer JA. Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. Cell. 2011;145:198-211. DOI:10.1016/j.cell.2011.03.004.

[61] Lin Y, Beard WA, Shock DD, Prasad R, Hou EW, Wilson SH. DNA polymerase β and flap endonuclease 1 enzymatic specificities sustain DNA synthesis for long patch base excision repair. J Biol Chem. 2005;280:3665-3674. DOI:10.1074/jbc.M412922200.

[62] Kimura S, Furukawa T, Kasai N, Mori Y, Kitamoto HK, Sugawara F, Hashimoto J, Sakaguchi K. Functional characterization of two flap endonuclease-1 homologues in rice. Gene. 2003;314:63-71. DOI:10.1016/S0378-1119(03)00694-2.

[63] Kimura S, Ueda T, Hatanaka M, Takenouchi M, Hashimoto J, Sakaguchi K. Plant homologue of flap endonuclease-1: Molecular cloning, characterization, and evidence of expression in meristematic tissues. Plant Mol Biol. 2000;42:415-427. DOI:10.1023/A:1006349511964.

[64] Zhang Y, Wen C, Liu S, Zheng L, Shen B, Tao Y. Shade avoidance 6 encodes an arabidopsis flap endonuclease required for maintenance of genome integrity and development. Nucleic Acids Res. 2016;44:1271-1284. DOI:10.1093/nar/gkv1474.

[65] Karanja KK, Livingston DM. C-terminal flap endonuclease (rad27) mutations: Lethal interactions with a DNA ligase I mutation (cdc9-p) and suppression by proliferating cell nuclear antigen (POL30) in *Saccharomyces cerevisiae*. Genetics. 2009;183:63-78. DOI:10.1534/genetics.109.103937.

[66] Zheng L, Zhou M, Chai Q, Parrish J, Xue D, Patrick SM, Turchi JJ, Yannone SM, Chen D, Shen B. Novel function of the flap endonuclease 1 complex in processing stalled DNA replication forks. EMBO Rep. 2005;6:83-89. DOI:10.1038/sj.embor.7400313.

[67] Lieber MR. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. BioEssays. 1997;19:233-240. DOI:10.1002/bies.950190309.

[68] Shen B, Nolan JP, Sklar LA, Park MS. Essential amino acids for substrate binding and catalysis of human flap endonuclease. J Biol Chem. 1996; 271: 9173-9176. DOI: 10.1074/jbc.271.16.9173.

[69] Moldovan GL, Pfander B, Jentsch S. PCNA, the Maestro of the Replication Fork. Cell. 2007;129:665-679. DOI: 10.1016/j.cell.2007.05.003.

[70] Strzalka W, Ziemienowicz A. Proliferating cell nuclear antigen (PCNA): A key factor in DNA replication and cell cycle regulation. Annals of Botany. 2011;107:1127-1140. DOI:10.1093/aob/mcq243.

[71] Boehm EM, Gildenberg MS, Washington MT. The Many Roles of PCNA in Eukaryotic DNA Replication. Enzymes. 2016;39:231-254. DOI:10.1016/bs.enz.2016.03.003.

[72] Anderson HJ, Vonarx EJ, Pastushok L, Nakagawa M, Katafuchi A, Gruz P, Di Rubbo A, Grice DM, Osmond MJ, Sakamoto AN, Nohmi T, Xiao W, Kunz BA. Arabidopsis thaliana Y-family DNA polymerase η catalyses translesion synthesis and interacts functionally with PCNA2. Plant J. 2008;55:895-908. DOI:10.1111/j.1365-313X.2008.03562.x.

[73] Markley NA, Young D, Laquel P, Castroviejo M, Moloney MM. Molecular genetic and biochemical analysis of *Brassica napus* proliferating cell nuclear antigen functionMolecular genetic and biochemical analysis of *Brassica napus* proliferating cell nuclear antigen function. Plant Mol Biol. 1997;34:693-700. DOI:10.1023/A:1005817220344.

[74] Kelman Z. PCNA: Structure, functions and interactions. Oncogene. 1997;14:629-40. DOI: 10.1038/sj.onc.1200886.

[75] De March M, Merino N, Barrera-Vilarmau S, Crehuet R, Onesti S, Blanco FJ, De Biasio A. Structural basis of human PCNA sliding on DNA. Nat Commun. 2017;8:13935. DOI: 10.1038/ncomms13935.

[76] Gomes X V., Burgers PMJ. Two modes of FEN1 binding to PCNA regulated by DNA. EMBO J. 2000;19:3811-3821. DOI:10.1093/emboj/19.14.3811.

[77] Sakurai S, Kitano K, Yamaguchi H, Hamada K, Okada K, Fukuda K, Uchida M, Ohtsuka E, Morioka H, Hakoshima T. Structural basis for recruitment of human flap endonuclease 1 to PCNA. EMBO J. 2005;24:683-693. DOI: 10.1038/sj.emboj.7600519.

[78] Singh SK, Roy S, Choudhury SR, Sengupta DN. DNA repair and recombination in higher plants: Insights from comparative genomics of arabidopsis and rice. BMC Genomics. 2010;11:443. DOI:10.1186/1471-2164-11-443.

[79] Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science. 2000;290: 1151-1155. DOI: 10.1126/science.290.5494.1151.

[80] Lynch M, O'Hely M, Walsh B, Force A. The probability of preservation of a newly arisen gene duplicate. Genetics. 2001;159:1789-1804.

[81] Cheng F, Wu J, Cai X, Liang J, Freeling M, Wang X. Gene retention, fractionation and subgenome differences in polyploid plants. Nature Plants. 2018;4:258-268. DOI:10.1038/s41477-018-0136-7.

[82] Panchy N, Lehti-Shiu M, Shiu SH. Evolution of gene duplication in plants. Plant Physiol. 2016;171:2294-2316. DOI:10.1104/pp.16.00523.

[83] Li Z, Defoort J, Tasdighian S, Maere S, Van De Peer Y, De Smet R. Gene duplicability of core genes is highly consistent across all angiosperms. Plant Cell. 2016;28:326-344. DOI: 10.1105/tpc.15.00877.