

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

**MARIA BEATRIZ DOS SANTOS MOTA**

VIAS DE REPARO DO DNA: aspectos evolutivos e o modelo *Aedes aegypti*

Rio de Janeiro  
2015

Maria Beatriz dos Santos Mota

VIAS DE REPARO DO DNA: aspectos evolutivos e o modelo *Aedes aegypti*

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Bioquímica, Instituto de Química, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Bioquímica.

Orientador: Rafael Dias Mesquita

Rio de Janeiro  
2015

M917

Mota, Maria Beatriz dos Santos.

Vias de reparo do DNA: aspectos evolutivos e o modelo *Aedes aegypti* / Maria Beatriz dos Santos Mota. – Rio de Janeiro: UFRJ/ IQ, 2015.

104 f.

Orientador: Rafael Dias Mesquita.

Tese (Mestrado em Ciências) - Universidade Federal do Rio de Janeiro, Instituto de Química, Programa de Pós-Graduação em Bioquímica, 2015.

1. Vias de reparo do DNA. 2. Domínio BRCT. 3. *Aedes aegypti*. I. Mesquita, Rafael Dias. (Orient.). II. Universidade Federal do Rio de Janeiro. Instituto de Química. Programa de Pós-Graduação em Bioquímica. III. Título.

CDD: 572

Maria Beatriz dos Santos Mota

VIAS DE REPARO DO DNA: aspectos evolutivos e o modelo *Aedes aegypti*

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Bioquímica, Instituto de Química, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Bioquímica.

Aprovada em

---

Dr. Rafael Dias Mesquita

---

Dr. Marcelo Alex de Carvalho

---

Dr.<sup>a</sup> Ana Cláudia do Amaral Melo

---

Dr.<sup>a</sup> Glória Regina Cardoso Braz

---

Dr.<sup>a</sup> Renata Schama Lellis (Suplente externo)

---

Dr.<sup>a</sup> Márcia Regina Soares da Silva (Suplente interno)

Dedico esse trabalho aos meus pais Claudio Mota e Barbara Mota e a minha irmã Maria Fernanda, as pessoas mais importantes da minha vida, a quem amo incondicionalmente.

## AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus por estar sempre ao meu lado guiando os meus passos e me ajudando a trilhar o melhor caminho.

Agradeço ao Professor Rafael Mesquita pela orientação, dedicação e atenção. Sua confiança e paciência foram fundamentais durante a realização deste trabalho e para o meu crescimento profissional.

Aos meus amigos do LBVD/BIOINFO, em especial a Thayany por enfrentar essa jornada junto comigo, a Fran e a Larissa por estarem sempre dispostas a me ajudar na parte experimental, e ao André pelas dicas na parte de bioinformática e evolução. Obrigada não só pela ajuda durante o desenvolvimento do trabalho, mas também pelos momentos de descontração e amizade que vocês me proporcionaram, quando eu defender a gente vê.

Agradeço ao Professor Mário Alberto por ceder os ovos de *Aedes aegypti* e ao Carlucio que esteve sempre disposto a me auxiliar em relação eclosão dos mesmos.

Agradeço ao Professor Renato Sampaio pela ajuda fundamental com os PCRs quantitativos, que foram de suma importância para a finalização deste trabalho.

Agradeço também ao Professor Marcelo Alex por disponibilizar seu laboratório no INCA para que eu pudesse realizar parte dos meus experimentos, e o pessoal do INCA, principalmente ao Thales.

Não poderia deixar de agradecer a minha família, ao meu pai Claudio por ser meu exemplo e minha inspiração, a minha mãe Barbara que é minha melhor amiga e me dá colo quando preciso e a minha irmã Maria Fernanda pela amizade, companheirismo e por estar sempre comigo. Agradeço também ao meu namorado Fabrício pela paciência que teve comigo durante esta jornada sempre tentando me passar confiança.

“O sucesso nasce do querer, da determinação e persistência em se chegar a um objetivo. Mesmo não atingindo o alvo, quem busca e vence obstáculos, no mínimo fará coisas admiráveis.” – José de Alencar

## RESUMO

Mota, Maria Beatriz dos Santos. **Vias de Reparo do DNA: Aspectos Evolutivos e o Modelo *Aedes aegypti***, 2015. Dissertação (Mestrado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015

O DNA está constantemente sujeito a danos em virtude da ação de fatores endógenos e exógenos. As vias de reparo do DNA são acionadas para correção dos danos, sendo que o domínio BRCT está presente em diversas proteínas destas vias em organizações estruturais que vão de 1 a 3 repetições. A identificação de ortólogos de TOPBP1 e ECT2 e de seu domínio conservado (tripleto de BRCT) ao longo da escala evolutiva, sugeriram a conservação evolutiva das vias de reparo do DNA. O *Aedes aegypti* é um mosquito vetor capaz de transmitir dengue, febre amarela, chikungunya e zika, seu hábito hematofágico leva a produção de espécies reativas de oxigênio durante a digestão da hemoglobina, podendo assim causar danos ao seu DNA. Neste modelo a busca por ortólogos das proteínas das vias de reparo do DNA e de outras que contivessem o domínio BRCT permitiu a identificação das proteínas XLF, Lig4 (duas isoformas), ECT2, MCPH1, MDC1 e PARP1. Além disso, a procura por indícios de expressão das proteínas das vias de reparo do DNA em *A. aegypti* identificou a expressão de 18 proteínas, sendo a maioria pertencente às vias de reparo que corrigem quebras na dupla fita. A curva de sobrevivência de larvas de *A. aegypti* tratadas com paraquat apresentou resposta dose-tempo dependente, sendo a LD50 do paraquat de 1 mM. A expressão dos genes da via de reparo por união terminal não homóloga (NHEJ), avaliada através de PCR quantitativo, indicou que, com exceção de *DNAPKcs*, todos os genes são expressos, e sugeriu um aumento de expressão quando a via foi estimulada com 0,5 mM de paraquat durante 24 horas.

Palavras chave: vias de reparo do DNA, domínio BRCT, *Aedes aegypti*



## ABSTRACT

Mota, Maria Beatriz dos Santos. **Vias de Reparo do DNA: Aspectos Evolutivos e o Modelo *Aedes aegypti***, 2015. Dissertação (Mestrado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015

The DNA structure is continuously subjected to damage due to the exposure to endogenous and exogenous genotoxic agents. Its integrity depends on the DNA damage response system (DDR). Several proteins involved in DDR enclose the BRCT motifs, which can contain from 1 to 3 BRCT units. The identification of TOPBP1 and ECT2 orthologs and their conserved domain (triplet BRCT) along the evolutionary scale, suggested the conservation of the DDR. The *Aedes aegypti* is the vector of yellow fever, chikungunya, zika and dengue, their blood-feeding behavior leads to the production of reactive oxygen species during hemoglobin digestion that can cause DNA damage. In *A. aegypti* the search for ortholog proteins of DNA repair pathways and those containing the BRCT domain allowed the identification of the proteins XLF, LIG4 (two isoforms), ECT2, MCPH1, MDC1 and PARP1. Furthermore, the search for expression evidences identified 18 proteins, most of them belonging to the repair pathways that fix the double strand breaks. The survival curve of *A. aegypti* (larvae) treated with paraquat showed dose and time-dependent responses with an LD50 of 1 mM. The expression of non-homologous end joining (NHEJ) genes was evaluated by quantitative PCR and indicated that, except for *DNAPKcs*, all genes were expressed. The results suggested higher expression levels after 24 hours of paraquat treatment (0.5 mM).

Keywords: DNA damage response, BRCT domain, *Aedes aegypti*

## LISTA DE FIGURAS

Figura 1: Via de reparo por excisão de base (BER). .....	22
Figura 2: Via de reparo por excisão de nucleotídeo (NER). .....	25
Figura 3: Via de reparo por erro de pareamento (MMR). .....	27
Figura 4: Via de reparo por recombinação homóloga (HR). .....	29
Figura 5: Via de reparo por união terminal não-homóloga (NHEJ). .....	31
Figura 6: Estrutura tridimensional dos domínios BRCT de BRCA1 representada pelo modelo de fitas. ....	33
Figura 7: Estrutura do arranjo em triplete representada pelo modelo de fitas. ....	34
Figura 8: Ciclo de vida do mosquito <i>Aedes aegypti</i> . .....	36
Figura 9: Áreas e países onde há risco de contrair dengue em 2013. ....	37
Figura 10: Países onde foi reportado transmissão de chikungunya e zika. ....	38
Figura 11: Exemplo da metodologia de BLAST recíproco. ....	44
Figura 12: Distribuição evolutiva das vias de reparo. ....	57
Figura 13: Alinhamento entre os três primeiros domínios BRCT presentes na região N-terminal de TOPBP1. ....	61
Figura 14: Alinhamento do triplete de BRCTs de TOPBP1 com a predição de estrutura secundária. ....	62
Figura 15: Dendograma obtido a partir do alinhamento entre os BRCTs individuais do triplete das TOPBP1. ....	63
Figura 16: Alinhamento entre os domínios BRCT presentes na região N-terminal de ECT2. ....	66
Figura 17: Alinhamento do triplete de BRCTs de ECT2 com a predição de estrutura secundária. ....	67
Figura 18: Dendograma obtido a partir do alinhamento entre os BRCTs individuais das ECT2. ....	68
Figura 19: Mapa mostrando a presença das proteínas TOPBP1 e ECT2 ao longo da escala evolutiva. ....	69
Figura 20: Via de reparo por excisão de base (BER) para <i>A. aegypti</i> . ....	73

Figura 21: Via de reparo por excisão de nucleotídeo (NER) para <i>A. aegypti</i> . ....	74
Figura 22: Via de reparo por erro de pareamento (MMR) para <i>A. aegypti</i> .....	75
Figura 23: Via de reparo por recombinação homóloga (HR) para <i>A. aegypti</i> .....	76
Figura 24: Via de reparo por união terminal não homóloga (NHEJ) para <i>A. aegypti</i> . ...	77
Figura 25: Curva de sobrevivência de larvas tratadas com diferentes concentrações de paraquat durante 24h. ....	79
Figura 26: Curva LD50 para o tratamento de larvas com paraquat durante 24h.....	79
Figura 27: Quantificação da expressão gênica de <i>Ku80</i> em larvas de <i>A. aegypti</i> . ....	80
Figura 28: Quantificação da expressão gênica de <i>RAD50</i> em larvas de <i>A. aegypti</i> . ....	81
Figura 29: Quantificação da expressão gênica de <i>Mre11</i> em larvas de <i>A. aegypti</i> . ....	81
Figura 30: Quantificação da expressão gênica de <i>RAD27</i> em larvas de <i>A. aegypti</i> . ....	82
Figura 31: Quantificação da expressão gênica de <i>MLL1</i> em larvas de <i>A. aegypti</i> .....	82
Figura 32: Quantificação da expressão gênica de <i>Lig4</i> em larvas de <i>A. aegypti</i> .....	83

**LISTA DE TABELAS**

Tabela 1: Tabela de iniciadores utilizados no PCR quantitativo.....	50
Tabela 2: Esquema das reações de PCR quantitativo.....	50
Tabela 3: Lista de proteínas das vias de reparo do DNA encontradas em <i>Hemichordata</i> . .....	53
Tabela 4: Listagem das proteínas encontradas nos outros pontos taxonômicos.....	55
Tabela 5: Proteínas das vias de reparo do DNA agrupadas por etapa. ....	56
Tabela 6: Tabela contendo as proteínas TOPBP1 identificadas e utilizadas no alinhamento de seus tripletes de BRCT.....	60
Tabela 7: Proteínas utilizadas no alinhamento dos tripletes de ECT2. ....	64
Tabela 8: Proteínas preditas e ausentes no KEGG para o mosquito <i>A. aegypti</i> . ....	71

**LISTA DE ABREVIACÕES**

5'dRP – 5' desoxirribose fosfato  
6-4PPs - fotoproduto (6-4) pirimidina-pirimidona  
A-NHEJ - Via alternativa reparo por união terminal não-homóloga  
Alka - *DNA-3-methyladenine glycosylase II*  
APE1 - *AP endonuclease 1*  
APE2 - *AP endonuclease 2*  
Artemis - *DNA cross-link repair 1C protein*  
ATM - *Ataxia-telangiectasia mutated*  
ATP – Trifosfato de adenosina  
ATR - *Ataxia telangiectasia and Rad3 related*  
BER - via de reparo por excisão de base  
BLAST - *Basic Local Alignment Search Tool*  
BLM - *bloom syndrome protein*  
BRCA1 - *breast cancer 1 susceptibility protein*  
BRCA2 - *breast cancer 2 susceptibility protein*  
C-NHEJ - Via clássica reparo por união terminal não-homóloga  
CCNH - *cyclin H*  
CDK7 - *cyclin-dependent kinase 7*  
CENT2 - *centrin-2*  
Ct - *threshold cycle*  
CYK-4 - *rac GTPase activating protein 1*  
CPD - dímeros de pirimidina ciclobutano  
CSA - *Cockayne syndrome group A*  
CSB - *Cockayne syndrome group B*  
Cul4 - *cullin 4*  
Dam - *DNA adenine methylase*  
DDB1 - *DNA damage-binding protein 1*  
DDB2 - *DNA damage-binding protein 2*  
DNAPKcs - *DNA-dependent protein kinase catalytic subunit*  
DnaT - *DNA biosynthesis protein (primosomal protein I)*  
Dnl4 - *DNA ligase 4*  
DpoI - DNA polymerase I  
DpoIII - *DNA polymerase III*

DSBs - Quebras na dupla fita  
DSS1 - *26 proteasome complex subunit DSS1*  
ECT2 - *Epithelial cell transforming 2*  
Eme1 - *crossover junction endonuclease EME1*  
ERCC1 - *excision repair cross-complementation group 1*  
EROs – espécies reativas de oxigênio  
Exo1 - *exonuclease 1*  
ExoI - *exodeoxyribonuclease I*  
ExoX - *exodeoxyribonuclease X*  
ExoVII - *exodeoxyribonuclease VII*  
Fen1 - *flap endonuclease-1*  
Fgp - *formamidopyrimidine-DNA glycosylase*  
GGR - reparo do genoma global  
GO - *Gene Ontology*  
HR - reparo por recombinação homóloga  
HR23B - *UV excision repair protein RAD23*  
Ku - *DNA end-binding protein Ku*  
Ku70 - *x-ray repair cross-complementing protein 6*  
Ku80 - *x-ray repair cross-complementing protein 5*  
LD<sub>50</sub> - *lethal dose 50%*  
Lig (NHEJ) - *bifunctional non-homologous end joining protein LigD*  
Lig (NER, MMR, BER) - *DNA ligase (NAD<sup>+</sup>)*  
Lig1 - *DNA ligase 1*  
Lig3 - *DNA ligase 3*  
Lig4 - *DNA ligase 4*  
Lif1 - *ligase-interacting factor 1*  
MBD4 - *methyl-CpG-binding domain protein 4*  
MCPH1 - *microcephalin 1*  
MDC1 - *mediator of DNA damage checkpoint protein 1*  
MFD - *transcription-repair coupling factor*  
MLH1 - *DNA mismatch repair protein MLH1, mutL homolog 1*  
MLH3 - *DNA mismatch repair protein MLH3, mutL homolog 3*  
MMR - reparo por erro de pareamento  
MNAT1 - *CDK-activating kinase assembly factor MAT1*

MPG - *DNA-3-methyladenine glycosylase*  
 Mre11 - *Double-strand break repair protein Mre11*  
 MSH2 - *DNA mismatch repair protein MSH2, mutS homolog 2*  
 MSH3 - *DNA mismatch repair protein MSH3, mutS homolog 3*  
 MSH6 - *DNA mismatch repair protein Msh6, mutS homolog 6*  
 MUS101 - *mutagen-sensitive-101*  
 Mus81 - *crossover junction endonuclease MUS81*  
 MutL - *DNA mismatch repair protein MutL*  
 MutS - *DNA mismatch repair protein MutS*  
 MUTY - *A/G-specific adenine glycosylase*  
 NBN - *nibrin/cell cycle regulatory protein p95*  
 Nbs1 - *nijmegen breakage syndrome protein 1*  
 Nei - *endonuclease VIII*  
 NEIL1 - *endonuclease VIII-like 1*  
 NEIL2 - *endonuclease VIII-like 2*  
 NEIL3 - *endonuclease VIII-like 3*  
 Nej1 - *non-homologous end-joining protein 1*  
 NER – *via de reparo por excisão de nucleotídeo*  
 Nfo - *deoxyribonuclease IV*  
 Mre11 - *meiotic recombination 11*  
 NHEJ - *via de reparo por união terminal não-homóloga*  
 PARP1 - *Poly ADP-ribose polymerase*  
 PAXIP1 - *PAX-interacting protein 1*  
 PCNA - *proliferating cell nuclear antigen*  
 PCRq - *PCR quantitativo*  
 PES1 - *Pescadillo ribosomal biogenesis factor 1*  
 Pfam - *Protein family database*  
 PMS2 - *postmeiotic segregation increased 2*  
 Pol4 - *DNA polymerase IV*  
 Pol $\beta$  - *DNA polymerase beta*  
 Pol $\delta$  - *DNA polymerase delta*  
 Pol $\epsilon$  - *DNA polymerase épsilon*  
 Pol $\lambda$  - *DNA polymerase lambda*  
 Pol $\mu$  - *DNA polymerase um*

PQ - Paraquat

PriA - *primosomal protein N' (replication factor Y)*

PriB - *primosomal replication protein N*

PriC - *primosomal replication protein N''*

pS/T - *resíduos fosforilados de Serina/Treonina*

Rad9 - *cell cycle checkpoint control protein RAD9A*

Rad23B - *UV excision repair protein RAD23 homolog B*

Rad27 - *flap endonuclease-1*

Rad50 - *DNA repair protein RAD50*

Rad51 - *DNA repair protein RAD51*

Rad51B - *RAD51-like protein 1*

Rad51C - *RAD51-like protein 2*

Rad51D - *RAD51-like protein 3*

Rad52 - *DNA repair and recombination protein RAD52*

Rad54 - *DNA repair and recombination protein RAD54 and RAD54-like protein*

Rad55 - *DNA repair protein RAD55*

Rad57 - *DNA repair protein RAD57*

Rad59 - *DNA repair protein RAD59*

RBX1 - *RING-box protein 1*

RecA - *DNA repair protein recA-like 1*

RecB - *exodeoxyribonuclease V beta subunit*

RecC - *exodeoxyribonuclease V gamma subunit*

RecD - *exodeoxyribonuclease V alpha subunit*

RecF - *gap repair protein, DNA replication and repair protein RecF*

RecG - *ATP-dependent DNA helicase RecG*

RecJ - *single-stranded-DNA-specific exonuclease*

RecO - *gap repair protein, DNA repair protein RecO*

RecR - *gap repair protein, recombination protein RecR*

Rev1 - *REVI-like*

RFC - *replication factor C subunit 1*

RuvA - *holliday junction DNA helicase RuvA*

RuvB - *holliday junction DNA helicase RuvB*

RuvC - *crossover junction endodeoxyribonuclease RuvC*

RPA - *replication factor A1*



SSB - *single-strand DNA-binding protein*  
ssDNA - DNA de fita simples  
SMUG - *single-strand selective monofunctional uracil DNA glycosylase*  
Tag - *DNA-3-methyladenine glycosylase I*  
TaxID - *Taxonomy ID*  
TCR - reparo acoplado a transcrição  
TDG/Mug - *TDG/mug DNA glycosylase family protein*  
TFIIH1 - *transcription initiation factor TFIIH subunit 1*  
TFIIH2 - *transcription initiation factor TFIIH subunit 2*  
TFIIH3 - *transcription initiation factor TFIIH subunit 3*  
TFIIH4 - *transcription initiation factor TFIIH subunit 4*  
TONSL - *tonsoku-like protein*  
TOP3 - *DNA topoisomerase III*  
TOPBP1 - *DNA topoisomerase 2-binding protein 1*  
TP53BP1 - *tumor protein p53 binding protein*  
TTDA - *TFIIH basal transcription factor complex TTD-A subunit*  
UNG - *uracil-DNA glycosylase*  
UVRA - *excinuclease ABC subunit A*  
UVRB - *excinuclease ABC subunit B*  
UVRC - *excinuclease ABC subunit C*  
UVRD - *DNA helicase II/ATP-dependent DNA helicase PcrA*  
XLF - *non-homologous end-joining factor 1*  
XPA - *xeroderma pigmentosum group A*  
XPB - *xeroderma pigmentosum group B*  
XPC - *xeroderma pigmentosum group C*  
XPD - *xeroderma pigmentosum group D*  
XPF - *xeroderma pigmentosum group F*  
XPG - *xeroderma pigmentosum group G*  
XRCC1 - *x-ray repair cross-complementing protein 1*  
XRCC2 - *x-ray repair cross-complementing protein 2*  
XRCC3 - *x-ray repair cross-complementing protein 3*  
XRCC4 - *x-ray repair cross-complementing protein 4*  
XRS2 - *DNA repair protein XRS2*  
Xth - *exodeoxyribonuclease III*

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	20
1.1. Mecanismos de Reparo do DNA .....	20
<b>1.1.1. Reparo por Excisão de Base</b> .....	20
<b>1.1.2. Reparo por Excisão de Nucleotídeo</b> .....	23
<b>1.1.3. Reparo por Erro de Pareamento</b> .....	26
<b>1.1.4. Reparo de quebras na dupla fita</b> .....	28
1.1.4.1. Reparo por Recombinação Homóloga .....	28
1.1.4.2. Reparo por União Terminal Não-Homóloga.....	30
1.2. Domínio BRCT.....	32
1.3. Insetos Vetores – <i>Aedes aegypti</i> .....	35
<b>1.3.1. Hematofagia e Estresse Oxidativo</b> .....	39
1.4. Vias de Reparo do DNA e Proteínas contendo domínio BRCT em Insetos....	40
<b>2. OBJETIVO</b> .....	42
2.1. Objetivos específicos .....	42
<b>3. MATERIAIS E MÉTODOS</b> .....	43
3.1. Identificação de proteínas de reparo do DNA e dos tripletes do domínio BRCT ao longo da evolução. ....	43
<b>3.1.1. Distribuição evolutiva das vias de reparo do DNA</b> .....	43
<b>3.1.2. Identificação dos tripletes</b> .....	44
3.2. Estudo das Vias de Reparo no Mosquito <i>A. aegypti</i> .....	46
<b>3.1.3. Identificação de proteínas de reparo e contendo domínio BRCT em <i>A. aegypti</i></b> .....	46
3.1.3.1. Confirmação das proteínas preditas no banco de dados.....	46
3.1.3.2. Identificação de novas proteínas .....	46
3.1.3.3. Busca por proteínas expressas.....	47
<b>3.1.4. Determinação de condições e estímulo e análises de expressão</b> .....	47
3.1.4.1. Eclosão dos ovos .....	47
3.1.4.2. Curva de sobrevivência – Paraquat .....	48
3.1.4.3. LD <sub>50</sub> .....	48
3.1.4.4. Extração e dosagem de RNA .....	48
3.1.4.5. Síntese de cDNA .....	49
3.1.4.6. PCR quantitativo .....	49
<b>4. RESULTADOS</b> .....	52

4.1.	Identificação de proteínas de reparo do DNA e dos tripletes do domínio BRCT ao longo da evolução. ....	52
4.1.1.	<b>Distribuição evolutiva das vias de reparo do DNA.....</b>	<b>52</b>
4.1.2.	<b>Distribuição Evolutiva do Triplete BRCT de TOPBP1 e ECT2 .....</b>	<b>58</b>
4.2.	Estudo das Vias de Reparo no Mosquito <i>A. aegypti</i> .....	70
4.1.3.	<b>Identificação de proteínas de reparo e contendo domínio BRCT em <i>A. aegypti</i>.....</b>	<b>70</b>
4.1.4.	<b>Determinação de condições e estímulo e análises de expressão.....</b>	<b>78</b>
<b>5.</b>	<b>DISCUSSÃO .....</b>	<b>84</b>
5.1.	Identificação de proteínas de reparo do DNA e dos tripletes do domínio BRCT ao longo da evolução. ....	84
5.2.	Identificação de proteínas de vias de reparo e contendo domínio BRCT em <i>A. aegypti</i> .....	90
<b>6.</b>	<b>CONCLUSÃO .....</b>	<b>95</b>
<b>7.</b>	<b>REFERÊNCIAS .....</b>	<b>96</b>

## 1. INTRODUÇÃO

### 1.1. Mecanismos de Reparo do DNA

Alterações à estrutura do DNA ocorrem frequentemente durante a vida celular, sendo consequência de fatores endógenos e da exposição a agentes genotóxicos exógenos. As lesões no DNA podem ocasionar mutações, câncer, morte celular e até mesmo levar à falência do organismo (SANCAR et al., 2004).

As lesões causadas por fatores endógenos podem ser oriundas de erros durante a replicação, como a incorporação errônea de um nucleotídeo; alterações espontâneas das bases acarretando em desaminação, depurinação ou depirimidinação; e a exposição a agentes genotóxicos derivados do metabolismo, como as espécies reativas de oxigênio (EROs) que ocasionam a oxidação e quebra do DNA (CICCIA; ELLEDGE, 2010; FRIEDBERG, 2001).

Dentre as causas ambientais de dano ao DNA destaca-se a radiação ionizante, que causa lesão de forma direta e pela indução de reações que geram EROs (CICCIA; ELLEDGE, 2010). A radiação ultravioleta, que acarreta na formação de dímeros de pirimidina. Além de uma diversidade de substâncias químicas, naturais ou sintéticas, que são capazes de se ligar ao DNA, como fármacos antineoplásicos, agentes alquilantes e toxinas também causam mutações (CICCIA; ELLEDGE, 2010; FRIEDBERG, 2001; MAEDA et al., 2001).

A detecção de alterações na estrutura do DNA ativa os mecanismos de reparo que objetivam corrigir o dano. Ocasionalmente quando o reparo é insuficiente, mecanismos de morte celular são ativados de forma a eliminar mutações potencialmente danosas (BARZILAI; YAMAMOTO, 2004; SANCAR et al., 2004).

#### 1.1.1. Reparo por Excisão de Base

A via de reparo por excisão de base (*Base excision repair* - BER) é responsável pelo reparo de lesões causadas por oxidação, alquilação, desaminação, depurinação das bases nitrogenadas, além de reparar sítios abásicos e quebras na fita simples (WALLACE; MURPHY; SWEASY, 2012; XU et al., 2008) (figura 1).

A BER se inicia com a detecção do erro por uma DNA glicosidase, enzima responsável pelo reconhecimento da base alterada e excisão através da clivagem de

ligações N-glicosídicas, formando um sítio abásico. As DNA glicosidases podem ser monofuncionais ou bifuncionais. Enquanto as monofuncionais apenas atuam como glicosidases, as bifuncionais possuem também atividade de liase, sendo capazes de clivar a fita de DNA na região onde o sítio abásico foi formado (IYAMA; WILSON, 2013; WALLACE; MURPHY; SWEASY, 2012).

Quando a glicosidase é monofuncional ocorre a hidrólise das ligações N-glicosídicas gerando um sítio abásico que é posteriormente reconhecido pela *AP endonuclease 1* (APE1), responsável pela clivagem da fita, na posição 5' imediatamente após o sítio abásico, formando os terminais 5' desoxirribose fosfato (5'dRP) e 3' OH (IYAMA; WILSON, 2013; WALLACE; MURPHY; SWEASY, 2012). Quando a glicosidase é bifuncional, primeiramente, ocorre a hidrólise e formação do sítio abásico e em seguida a fita de DNA é clivada, na posição 3' após o sítio abásico, formando na extremidade 3' um grupamento fosfato ou aldeído  $\beta,\alpha$ -insaturado (PUA), que são, subsequentemente, removidos e substituídos por uma hidroxila (3'-OH) (WALLACE; MURPHY; SWEASY, 2012).

Após a detecção, remoção e processamento das extremidades a BER pode continuar por duas vias: a via curta e via longa. A maioria das lesões é processada pela via curta, independente se o reparo for iniciado por uma glicosidase mono ou bifuncional (ALMEIDA; SOBOL, 2007). No entanto, quando o grupamento da extremidade 5' não for substrato para *DNA polimerase  $\beta$*  (Pol  $\beta$ ) o reparo procede pela via longa. A via longa também é preferencial em situações de baixa concentração de ATP e durante a fase S do ciclo celular (IYAMA; WILSON, 2013).

Na via curta, após o processamento das extremidades, a Pol  $\beta$  remove o açúcar do sítio abásico e preenche o gap inserindo um novo nucleotídeo. Em seguida o complexo formado pelas proteínas *X-ray repair complementing defective repair in Chinese hamster cells 1* e *DNA ligase 3* (XRCC1-Lig3) liga as extremidades finalizando o reparo (SANCAR et al., 2004).

Na via longa, inicialmente, um complexo formado por DNA polimerase  $\beta$ ,  $\delta$ ,  $\epsilon$  (Pol  $\beta$ , Pol  $\delta$ , Pol  $\epsilon$ ), *proliferating cell nuclear antigen* (PCNA) e *flap endonuclease 1* (Fen1) desloca de 2 a 10 nucleotídeos da extremidade 3'. Em seguida, a endonuclease Fen1, presente no complexo, realiza a excisão da cadeia de nucleotídeos enquanto a DNA pol  $\delta/\epsilon$ , com o auxílio de PCNA, sintetiza uma nova cadeia nucleotídica de tamanho semelhante a removida, que é incorporada a fita pela *DNA ligase 1* (Lig1) (SANCAR et al., 2004).

## Via de Reparo por Excisão de Base (BER)

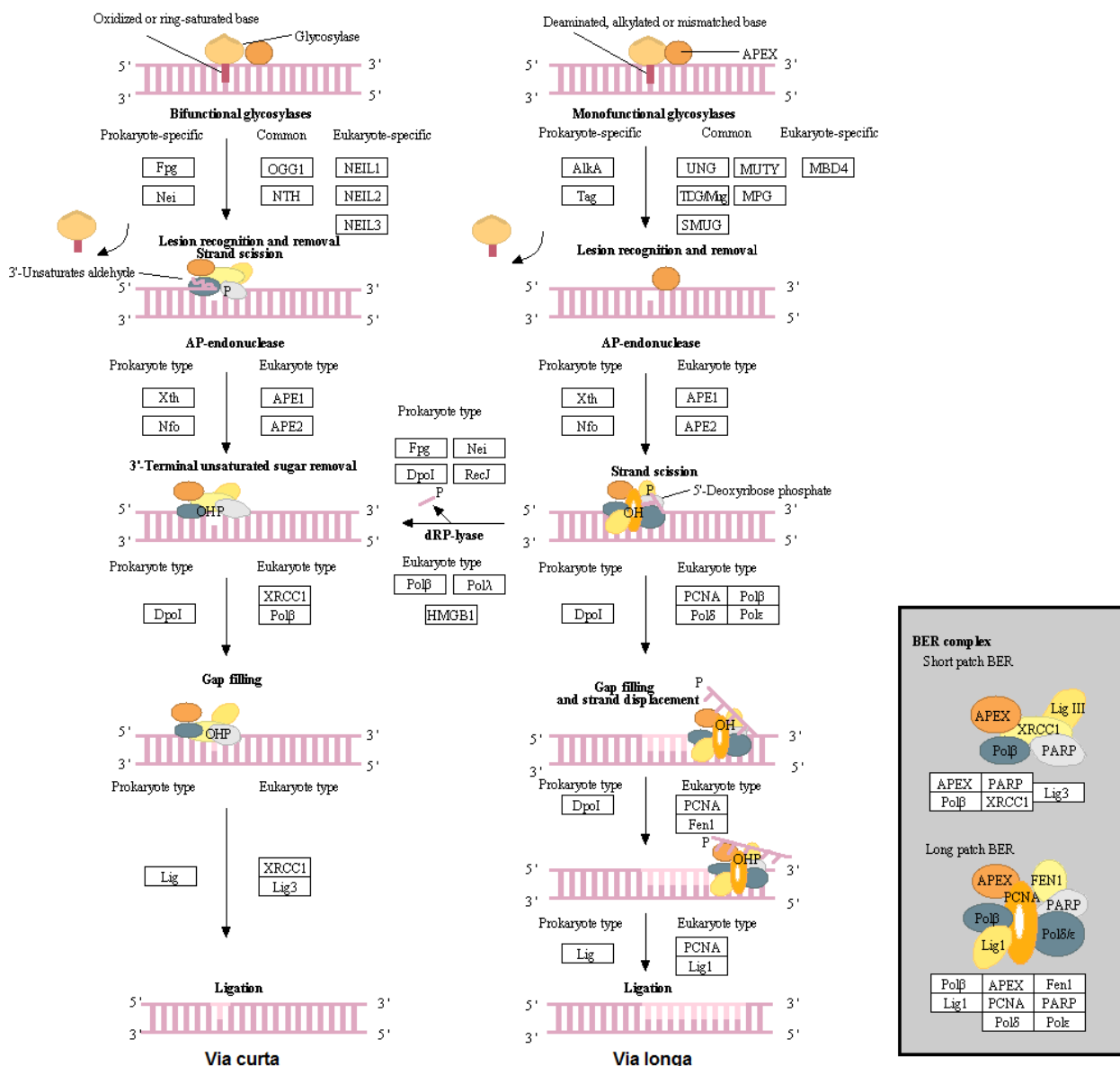


Figura 1: Via de reparo por excisão de base (BER).

A BER pode ser dividida em quatro etapas: Detecção, processamento, síntese e ligação. A etapa de detecção corresponde ao reconhecimento e excisão da base alterada, realizada por glicosidades, que podem ser mono (exemplo: UNG, MUTY) ou bifuncionais (OGG1 e NHT). A segunda etapa é o processamento das extremidades através da incisão na região do sítio abásico. A etapa de síntese repõe o nucleotídeo removido. Por fim, na quarta etapa o reparo é finalizado pela ligação das extremidades. A BER pode ocorrer através da via curta ou da via longa. Na via curta a terceira etapa é orquestrada por XRCC1 e Pol β, e na via longa por PCNA, Pol β, δ, ε e Fen1. As vias também diferem na última etapa, sendo realizada por Lig3 e Lig1, respectivamente.

Adaptado de:

[http://www.kegg.jp/kegg-bin/highlight\\_pathway?scale=1.0&map=map03410&keyword=dna%20repair](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=map03410&keyword=dna%20repair).

### 1.1.2. Reparo por Excisão de Nucleotídeo

A via de reparo por excisão de nucleotídeo (*Nucleotide excision repair* - NER) repara lesões que distorcem a dupla hélice do DNA, grande parte ocasionada por agentes ambientais como raios UV e produtos químicos carcinogênicos (FAGBEMI; ORELLI; SCHÄRER, 2011) (figura 2). Entre as lesões exógenas reparadas pela NER estão os dímeros de pirimidina ciclobutano (CPD) e o fotoproducto (6-4) pirimidina-pirimidona (6-4PPs) gerados pela radiação UV (IYAMA; WILSON, 2013); e os adutos formados pela ligação de compostos químicos ao DNA, como benzopireno, cisplatina e alfatoxina (MAEDA et al., 2001). Entretanto, a NER também é capaz de realizar o reparo de lesões endógenas, como as geradas pelas espécies reativas de oxigênio (EROs) e peroxidação de lipídeos (IYAMA; WILSON, 2013).

O reconhecimento da lesão pode ocorrer por duas sub vias distintas: reparo do genoma global (*global genomic repair* - GGR) e o reparo acoplado à transcrição (*transcription-coupled repair* - TCR). O local onde ocorreu a lesão define qual sub via será utilizada, se o dano estiver em uma região que está sendo transcrita a TCR será ativada, caso esteja localizado em um domínio inativo o reparo irá acontecer via GGR (CLEAVER; LAM; REVET, 2009). Acredita-se que as sub vias apenas diferem na forma de detecção do erro, sendo os passos subsequentes semelhantes para ambas (IYAMA; WILSON, 2013).

Na GGR a detecção de alterações na dupla hélice se inicia por um complexo formado pela proteína sensora *xeroderma pigmentosum group C* (XPC), e as proteínas *UV excision repair protein RAD23 homolog B* (RAD23B) e *centrin-2* (CETN2). (MARTEIJN et al., 2014). O complexo XPC-RAD23-CENT2 se liga a fita simples oposta a lesão e recruta o complexo TFIIH (FAGBEMI; ORELLI; SCHÄRER, 2011). Na TCR a detecção ocorre devido ao bloqueio da elongação da RNA polimerase II mediante a distorção na dupla hélice, servindo como sinal para o recrutamento das proteínas CSA e CSB (IYAMA; WILSON, 2013; MARTEIJN et al., 2014) as quais recrutam o TFIIH para o local da lesão (EARLEY; TURCHI, 2011).

O complexo TFIIH contém dez subunidades proteicas, sendo duas helicases *xeroderma pigmentosum group B* e *xeroderma pigmentosum group D* (XPB e XPD) responsáveis pela abertura da dupla fita de DNA em torno da lesão permitindo o recrutamento dos fatores *replication factor A1* (RPA) e *xeroderma pigmentosum group A* (XPA). O fator RPA se liga a fita oposta a lesão para evitar a junção das fitas e

protegê-la de nucleases (EARLEY; TURCHI, 2011). Especula-se que o fator XPA desempenhe função de verificação do dano devido sua capacidade de detectar nucleotídeos com alterações em suas estruturas, além de servir como suporte para formação do complexo de pré-incisão (EARLEY; TURCHI, 2011; MARTEIJN et al., 2014). A excisão da lesão é catalisada pelas endonucleases *xeroderma pigmentosum group F* (XPF) e *excision repair cross-complementation group 1* (ERCC1) e pela proteína *xeroderma pigmentosum group G* (XPG), que clivam a fita na posição 5' e 3' respectivamente, formando na fita um gap de 22 a 30 nucleotídeos (MARTEIJN et al., 2014). Em seguida o gap é preenchido pela Pol  $\delta$ ,  $\epsilon$  ou  $\kappa$  em cooperação com *proliferating cell nuclear antigen* (PCNA) e *replication factor C subunit 1* (RFC) e a ligação realizada pela Lig1 ou pelo complexo XRCC1-Lig3, sendo as proteínas utilizadas dependentes do estado proliferativo da célula. Em células proliferativas a síntese e a ligação são realizadas pela Pol  $\epsilon$  e Lig1, enquanto em células quiescentes Pol  $\delta$  e  $\kappa$  sintetizam o novo fragmento de DNA e a ligação é função do complexo XRCC1-Lig3 (CLEAVER; LAM; REVET, 2009; MARTEIJN et al., 2014).



## Via de Reparo por Excisão de Nucleotídeo (NER)

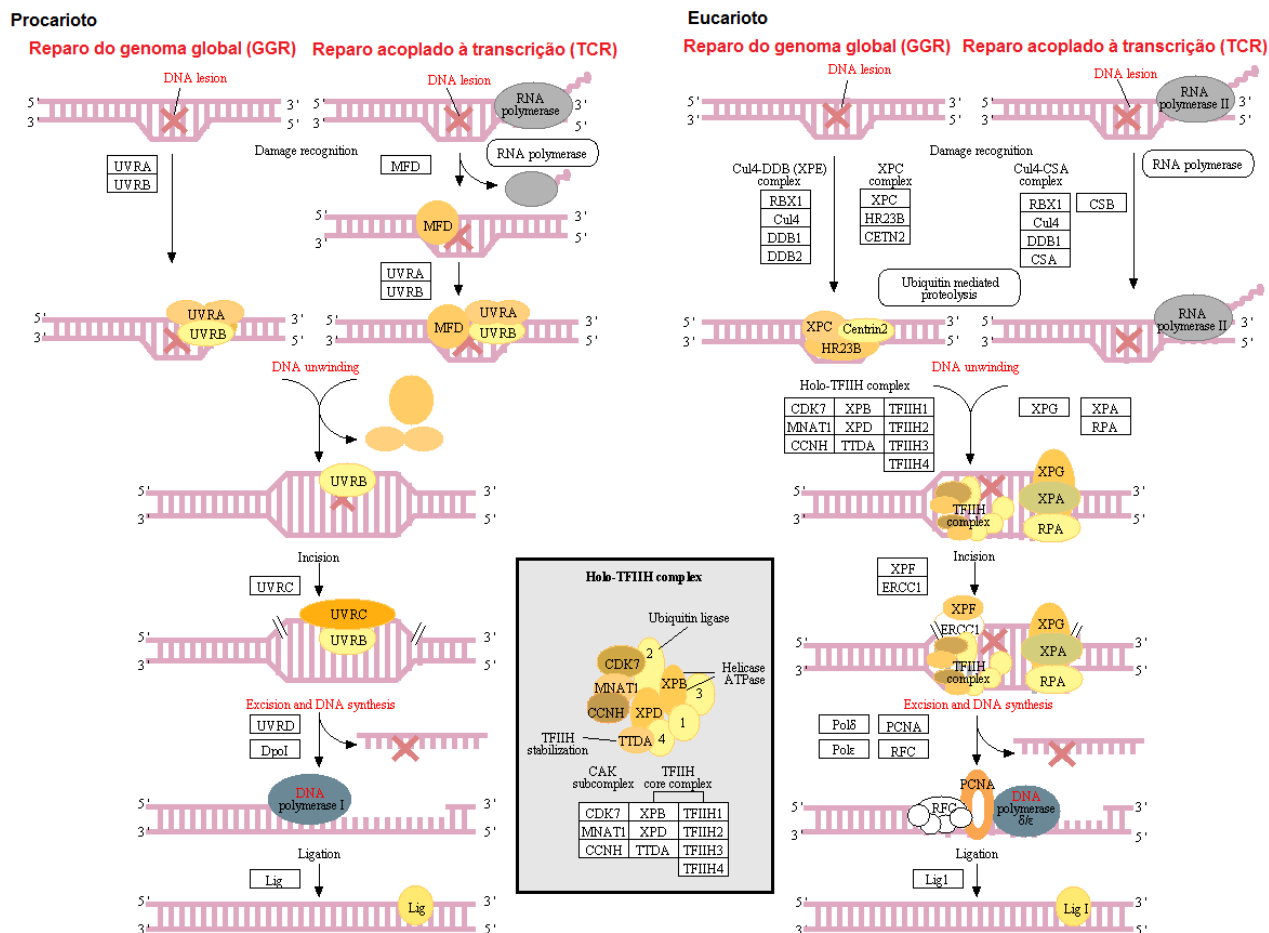


Figura 2: Via de reparo por excisão de nucleotídeo (NER).

A NER pode ser dividida em quatro etapas: detecção, processamento, síntese e ligação. A primeira etapa corresponde ao reconhecimento do dano, que pode ocorrer através do reparo do genoma global (GGR) ou acoplado à transcrição (TCR). O que determina por qual via a detecção irá ocorrer é a co-localização (ou não) do dano e da transcrição. Após a detecção, ambas prosseguem da mesma forma, seguindo para segunda etapa onde ocorre o processamento, que consiste na incisão em torno da lesão para remoção do oligonucleotídeo danificado. Na terceira etapa ocorre a síntese de um novo oligonucleotídeo para repor o danificado. O reparo é finalizado através da ligação das extremidades pela Lig1. Adaptado de [http://www.kegg.jp/kegg-bin/highlight\\_pathway?scale=1.0&map=map03420&keyword=dna%20repair](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=map03420&keyword=dna%20repair).

### 1.1.3. Reparo por Erro de Pareamento

O reparo por erro de pareamento (*Mismatch* - MMR) repara bases mal pareadas e *loops* causados por pequenas inserções ou deleções, ocorridos durante a replicação e que escaparam da detecção pela DNA polimerase (HSIEH; YAMANE, 2008; IYAMA; WILSON, 2013) (figura 3). A MMR também controla a fidelidade da recombinação homóloga (HR) (PAPOULI; CEJKA; JIRICNY, 2004). Mutações nos genes que codificam proteínas envolvidas no reconhecimento do mau pareamento acarretam predisposição ao câncer colorretal não polipóide hereditário ou síndrome de Lynch (IYAMA; WILSON, 2013), estando associadas a mais de 80% dos casos (ERIE; WENINGER, 2014).

A MMR é orquestrada, basicamente, pelas proteínas *DNA mismatch repair protein MutS* (MutS) e *DNA mismatch repair protein MutL* (MutL), em procariotos. A MMR tem início através da detecção de alterações pelos homólogos de MutS, sendo, em eucariotos, função dos heterodímeros MutS $\alpha$  e MutS $\beta$ . Bases mal pareadas, e inserções ou deleções de até duas bases são detectados por MutS $\alpha$  formado pelas proteínas *mutS homolog 2* (MSH2) e *mutS homolog 6* (MSH6), e o reconhecimento de inserções ou deleções de mais de duas bases é realizado por MutS $\beta$ , que compreende as proteínas MSH2 e *mutS homolog 3* (MSH3) (ERIE; WENINGER, 2014; IYAMA; WILSON, 2013). Após a detecção, MutL é recrutado para o local da lesão e forma um complexo com MutS, sinalizando para o início da excisão do dano (MARTIN et al., 2010). Em humanos existem três heterodímeros equivalentes a MutL (MutL $\alpha$ , MutL $\beta$ , MutL $\gamma$ ), sendo MutL $\alpha$ , formado por *DNA mutL homolog 1* (MLH1) e *postmeiotic segregation increased 2* (PMS2), o heterodímero de maior participação na MMR (ERIE; WENINGER, 2014; IYAMA; WILSON, 2013). Antes de iniciar a excisão, para que o reparo seja executado com sucesso, a maquinaria da MMR deve reconhecer a fita recém-sintetizada. Em eucariotos, propõe-se que a fita líder seja reconhecida pela ligação da PCNA, que direciona a MutL $\alpha$  para o local onde deve ocorrer a incisão, enquanto os fragmentos de Okazaki auxiliam na identificação da fita retardatária (ERIE; WENINGER, 2014; IYAMA; WILSON, 2013; STOJIC; BRUN; JIRICNY, 2004). O mau pareamento é retirado através da excisão na fita, ação orquestrada por PCNA em cooperação com a *exonuclease 1* (ExoI) (IYAMA; WILSON, 2013). O passo seguinte da via é o preenchimento do gap formado pela excisão do fragmento contendo o pareamento errôneo, a síntese de uma nova sequência nucleotídica é catalisada pela Pol

$\delta$  na fita líder e na fita retardatária a síntese pela Pol  $\epsilon$  (ERIE; WENINGER, 2014). Finalizando o reparo a Lig1 incorpora a nova sequência a fita de DNA (IYAMA; WILSON, 2013).

### Via de Reparo por Erro de Pareamento (MMR)

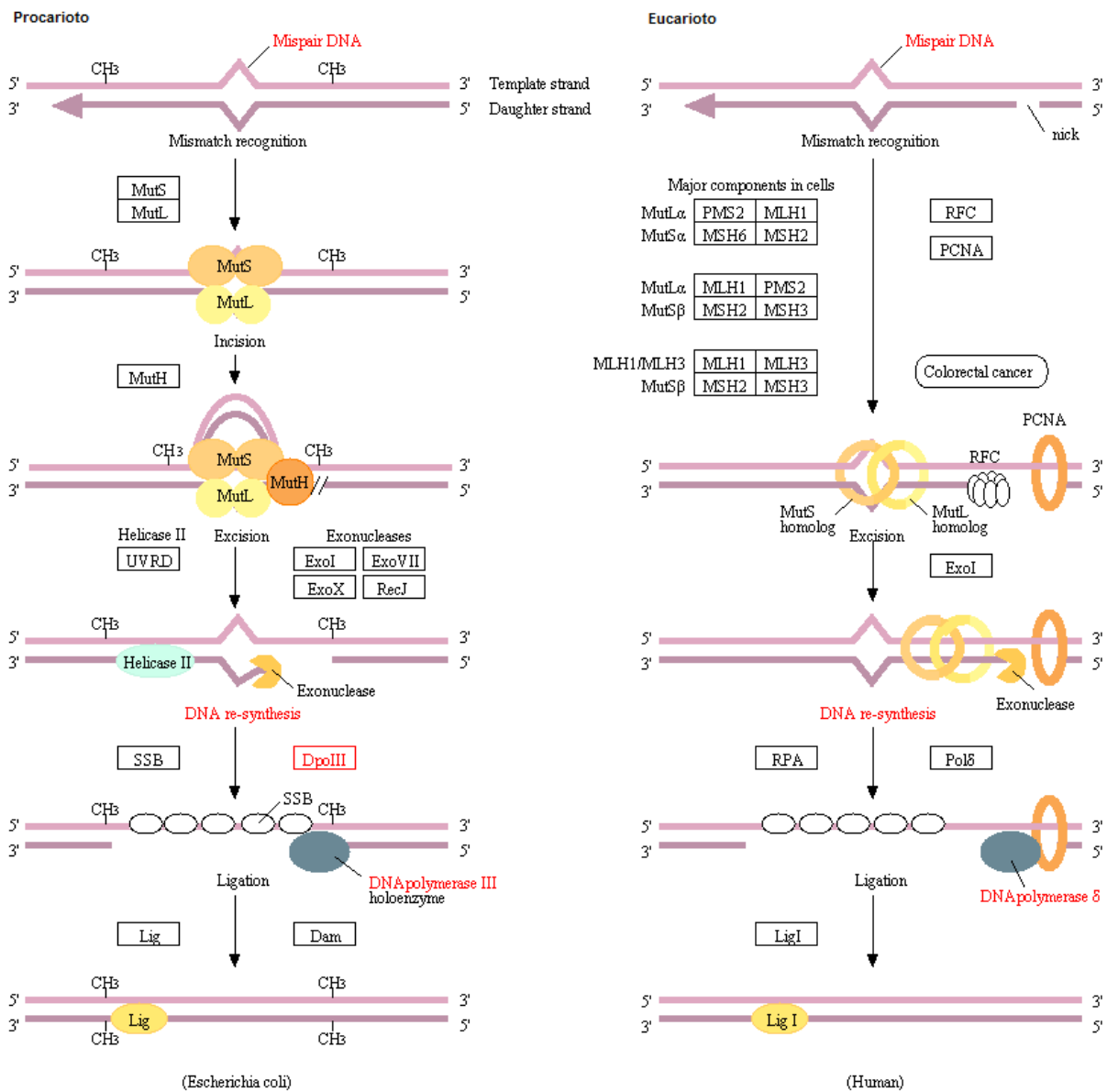


Figura 3: Via de reparo por erro de pareamento (MMR).

A MMR pode ser dividida em quatro etapas: detecção, processamento, síntese e ligação. Esta via se inicia com a detecção do erro através dos heterodímeros MutS e MutL. Em seguida, o processamento consiste na remoção do erro através da ExoI. A etapa de síntese e substituição do fragmento danificado é realizada por RPA e Pol  $\delta$ . Por último, a Lig1 realiza a ligação restaurando o DNA. Adaptado de [http://www.kegg.jp/kegg-bin/highlight\\_pathway?scale=1.0&map=map03430&keyword=dna%20repair](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=map03430&keyword=dna%20repair).

#### 1.1.4. Reparo de quebras na dupla fita

Quebras na dupla fita (DSBs) podem ocasionar a perda de informação genética (HELLEDAY et al., 2007), instabilidade genômica e ativar mecanismos de morte celular quando não reparadas (IYAMA; WILSON, 2013). DSBs podem ocorrer como resposta a ação de agentes genotóxicos endógenos e exógenos, como as espécies reativas de oxigênio e radiação ionizante, como consequência de falhas durante a replicação e também ocorrem naturalmente durante o processo de recombinação V(D)J (processo de recombinação genética, também denominado recombinação somática, que ocorre durante os primeiros passos da maturação de células B e T) e meiose (IYAMA; WILSON, 2013; O'DRISCOLL; JEGGO, 2006; SANCAR et al., 2004).

##### 1.1.4.1. Reparo por Recombinação Homóloga

O reparo por recombinação homóloga (HR) utiliza a região homóloga da cromátide irmã como molde para correção da quebra na dupla fita, o que evita erros e perdas, conferindo maior acurácia ao reparo. Por necessitar de um molde de DNA ocorre apenas nas fases S e G2 do ciclo celular, etapas onde o cromossomo homólogo se encontra disponível para servir como molde. Está diretamente envolvida no reparo de DSBs causadas pelo colapso da forquilha de replicação durante a replicação do DNA em células somáticas, sendo responsável pela manutenção da integridade genômica (HELLEDAY et al., 2007; IYAMA; WILSON, 2013; LOK; POWELL, 2012) (figura 4).

A HR tem início com o reconhecimento da DSB pelo complexo MRN, formado pelas proteínas *meiotic recombination 11* (Mre11), *DNA repair protein RAD50* (Rad50) e *nijmegen breakage syndrome 1* (Nbs1), o qual atua como um sensor para a cinase *ataxia-telangiectasia mutated* (ATM), recrutando-a para o sítio da quebra (IYAMA; WILSON, 2013; OVERCASH et al., 2015). Em seguida, as extremidades das fitas são clivadas por endonucleases para gerar terminais 3' de DNA de fita simples (ssDNA) os quais são estabilizados pela ligação da proteína RPA (IYAMA; WILSON, 2013; LOK; POWELL, 2012). O passo seguinte da via é a formação de filamentos capazes de invadir a cromátide irmã (BENJAMIN H. LOK1, 2012), dessa forma a proteína *breast cancer 2, early onset* (BRCA2) medeia a substituição de RPA por *DNA repair protein RAD51* (RAD51) (IYAMA; WILSON, 2013), proteína que desempenha papel central

na via por ser responsável pela formação dos filamentos de nucleoproteínas necessários para a invasão e a busca por regiões homólogas (HELLEDAY et al., 2007). A invasão acarreta na formação de *D-loop*, nesse ponto o reparo pode seguir via formação de junções de *Holiday* ou através de um modelo que finaliza o reparo sem a formação de junções de *Holiday* (IYAMA; WILSON, 2013). Ao encontrar a região homóloga a Pol  $\delta$  é recrutada para finalizar o reparo (OVERCASH et al., 2015).

### Via de Reparo por Recombinação Homóloga (HR)

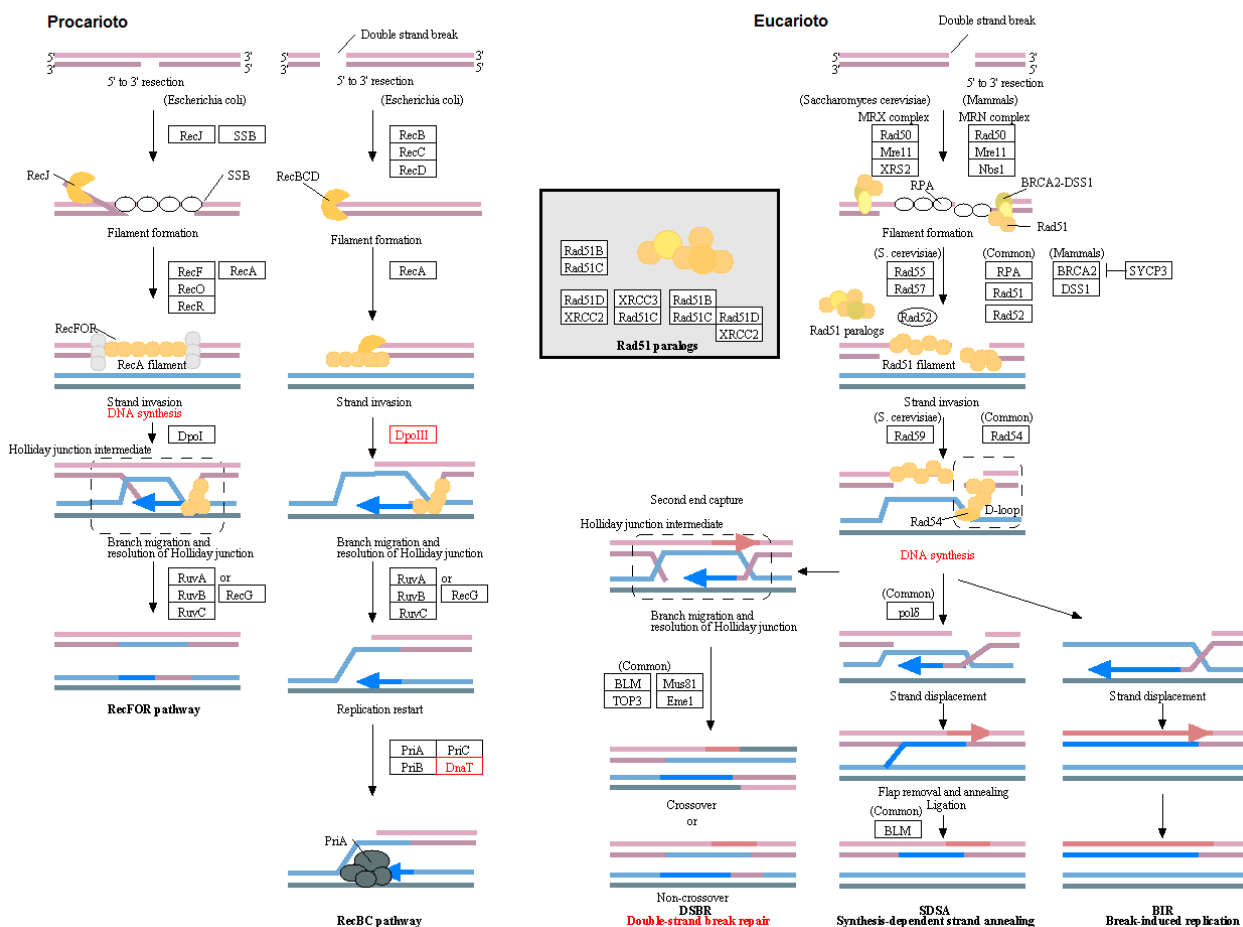


Figura 4: Via de reparo por recombinação homóloga (HR).

A HR pode ser dividida em quatro etapas: Detecção, processamento, síntese e recombinação. Na primeira etapa o complexo MRN detecta a quebra na dupla fita. Na etapa de processamento ocorre a formação de terminais de ssDNA, e invasão do cromossomo homólogo através da Rad51. A recombinação pode ocorrer através da formação de junções de *Holiday* ou não, e a síntese é realizada pela Pol  $\delta$ . Adaptada de [http://www.kegg.jp/kegg-bin/highlight\\_pathway?scale=1.0&map=map03440&keyword=dna%20repair](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=map03440&keyword=dna%20repair).

#### 1.1.4.2. Reparo por União Terminal Não-Homóloga

A via de reparo por união terminal não-homóloga (NHEJ) atua em todas as fases do ciclo celular, sendo predominante na G1. É responsável por reparar as DSBs originadas durante o processo V(D)J e, além disso, algumas de suas proteínas (Ku e DNAPKcs) possuem papel importante na manutenção da integridade dos telômeros (O'DRISCOLL; JEGGO, 2006; IYAMA; WILSON 2013) (figura 5).

A NHEJ pode ser dividida em duas sub vias: a via principal dependente do heterodímero Ku também denominada de via clássica (C-NHEJ) e a via alternativa (A-NHEJ) que usa micro-homologia e proteínas de outras vias, como HR, em seu mecanismo de reparo (FATTAH et al., 2010; IYAMA; WILSON 2013).

A C-NHEJ se inicia com a ligação do heterodímero Ku, formado pelas proteínas *x-ray repair cross-complementing 6* (Ku70) e *x-ray repair cross-complementing 5* (Ku80), as extremidades das fitas danificadas (IYAMA; WILSON, 2013). Ao se ligar, o heterodímero protege a região danificada da ação das exonucleases e recruta a cinase *DNA-dependent protein kinase catalytic subunit* (DNAPKcs) responsável pela fosforilação da endonuclease Artemis (FATTAH et al., 2010), que remove nucleotídeos das extremidades de forma a gerar terminais complementares. Em seguida, Pol  $\mu$  e  $\lambda$  adicionam nucleotídeos necessários para restauração da dupla fita (OVERCASH et al., 2015). A ligação das extremidades é catalisada pelo complexo XRCC4-Lig4 formado pelas proteínas *x-ray repair cross-complementing 4* (XRCC4) e *DNA ligase 4* (Lig4) associado com a proteína *non-homologous end-joining factor 1* (XLF) (FATTAH et al., 2010).

## Via de Reparo por União Terminal Não-Homóloga (NHEJ)

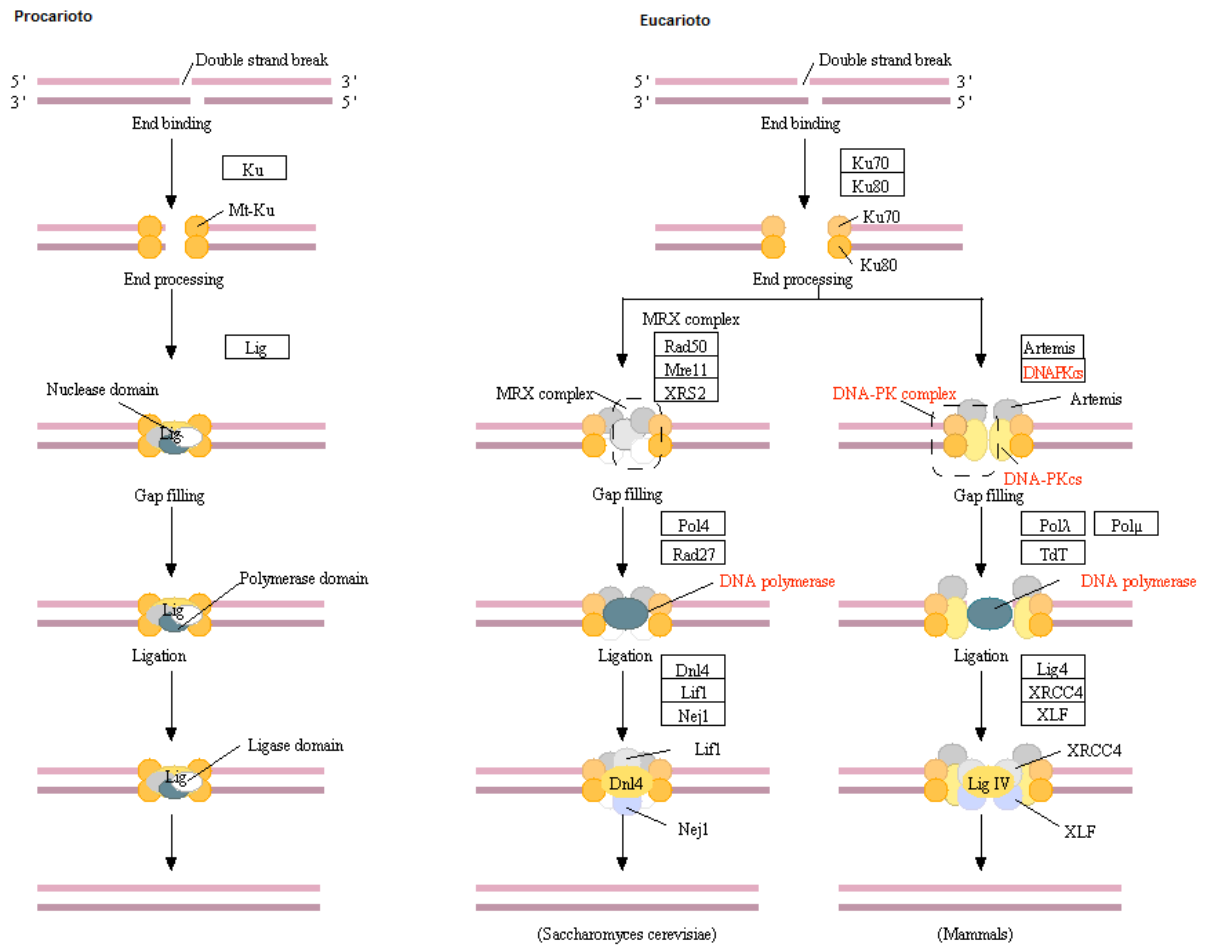


Figura 5: Via de reparo por união terminal não-homóloga (NHEJ).

A NHEJ pode ser dividida em três etapas: detecção, processamento e ligação. A quebra na dupla fita do DNA é detectada pelo complexo Ku (Ku80 e Ku70). Em seguida, a etapa de processamento consiste em gerar terminais complementares através da remoção e adição de nucleotídeos. Na última etapa, as extremidades são reconectadas pelo complexo formado por Lig4, XRCC4 e XLF. Adaptado de [http://www.kegg.jp/kegg-bin/highlight\\_pathway?scale=1.0&map=map03450&keyword=dna%20repair](http://www.kegg.jp/kegg-bin/highlight_pathway?scale=1.0&map=map03450&keyword=dna%20repair).

## 1.2. Domínio BRCT

O domínio BRCT foi inicialmente identificado na porção C-terminal da proteína codificada pelo gene de predisposição ao câncer de mama *BRCA1* (GERLOFF et al., 2012; WOODS et al., 2012), tendo um importante papel na função de supressão tumoral desta proteína (LI et al., 1999; YU et al., 2003). Mutações patogênicas em *BRCA1* aumentam a predisposição a câncer de mama e ovário, sendo encontradas em aproximadamente 50% dos casos de câncer de mama hereditários (EKBLAD et al., 2002; LEUNG; GLOVER, 2011; YU et al., 2003; ZHANG et al., 1998a).

É encontrado em uma diversidade de proteínas procarióticas e eucarióticas (GERLOFF et al., 2012; MESQUITA et al., 2010), das quais a maioria participa das vias de reparo de DNA e dos pontos de checagem do ciclo celular, onde medeiam diversas interações proteína-proteína e proteína-ácido nucleico, além de serem capazes de reconhecer fosfopeptídeos (WOODS et al., 2012). A variedade de funções atribuídas a este domínio está relacionada a diversidade em sua arquitetura, podendo ser encontrado como um domínio individual (single), em múltiplas repetições (tandem) ou em fusão com outros domínios funcionais (FHA e FN3) (LEUNG; GLOVER, 2011).

O domínio BRCT é constituído por, aproximadamente, 90 a 100 aminoácidos (MESQUITA et al., 2010), sua estrutura globular  $\alpha/\beta$  foi determinada, pela primeira vez, por cristalografia, na proteína humana XRCC1, sendo formada por 4 fitas  $\beta$ -pregueadas flanqueadas, de um lado, por um par de  $\alpha$ -hélice ( $\alpha_1$  e  $\alpha_3$ ) e de outro por apenas uma  $\alpha$ -hélice ( $\alpha_2$ ) (ZHANG et al., 1998b). Posteriormente, a determinação da estrutura da região C-terminal de *BRCA1* por cristalografia revelou que ambos os domínios presentes no tandem apresentam estrutura similar ao domínio BRCT de XRCC1 (figura 6) (WILLIAMS; GREEN; GLOVER, 2001).

Quando em tandem, como demonstrado na figura 1, os domínios BRCT são separados por uma região de alça de ligação e se enovelam em um arranjo cabeça-cauda, atuando como uma única unidade funcional capaz de reconhecer proteínas fosforiladas (MESQUITA et al., 2010; SHENG; ZHAO; HUANG, 2011; WILLIAMS; GREEN; GLOVER, 2001).

Em *BRCA1* o reconhecimento de fosfopeptídeos ocorre devido aos resíduos conservados S1655 e K1702, presentes no BRCT N-terminal, que são capazes de realizar ligações de hidrogênio com resíduos fosforilados de Serina/Treonina (pS/T). A região hidrofóbica formada pelos aminoácidos F1704, M1775 e L1839, presente no



domínio BRCT C-terminal, forma a parte central do tandem e promove a especificidade por fenilalanina na posição +3 (CLAPPERTON et al., 2004; MESQUITA et al., 2010; WILLIAMS; GREEN; GLOVER, 2001).

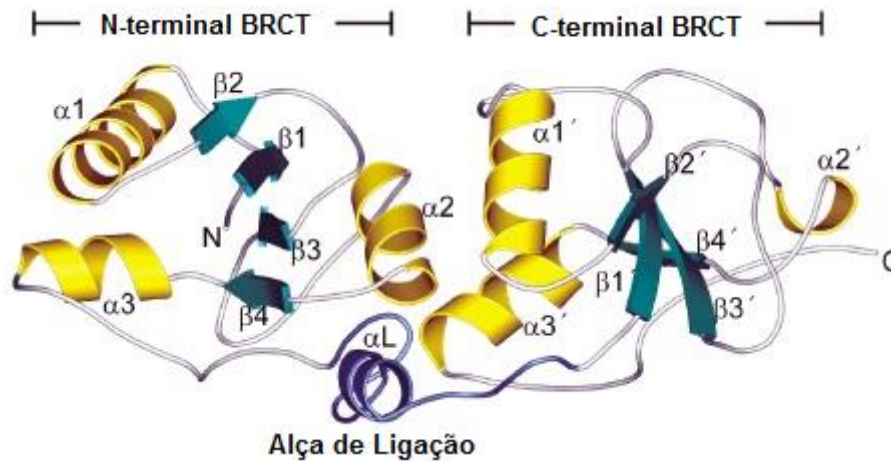


Figura 6: Estrutura tridimensional dos domínios BRCT de BRCA1 representada pelo modelo de fitas.

No estudo foi utilizada a região C-terminal de BRCA1 compreendendo os aminoácidos de 1646-1859. Linker consiste na região de ligação entre os domínios. Adaptado de (WILLIAMS; GREEN; GLOVER, 2001).

Recentemente, estudos cristalográficos, permitiram a caracterização de um arranjo tandem em triplete nas proteínas humanas *DNA topoisomerase 2-binding protein 1* (TOPBP1) e *epithelial cell transforming 2* (ECT2) (figura 7 A e B). Este tipo de arquitetura em triplete apresenta regiões de alça de ligação bastante curtas, entre 17 e 22 aminoácidos, contrastando com outras estruturas de BRCT tandem (dubletes) que, em geral, possuem mais de 30 aminoácidos nesta região (RAPPAS; OLIVER; PEARL, 2011; ZOU et al., 2014). O triplete de BRCT também apresenta diferenças estruturais em relação aos domínios BRCT single e tandem (dublete) (RAPPAS; OLIVER; PEARL, 2011).

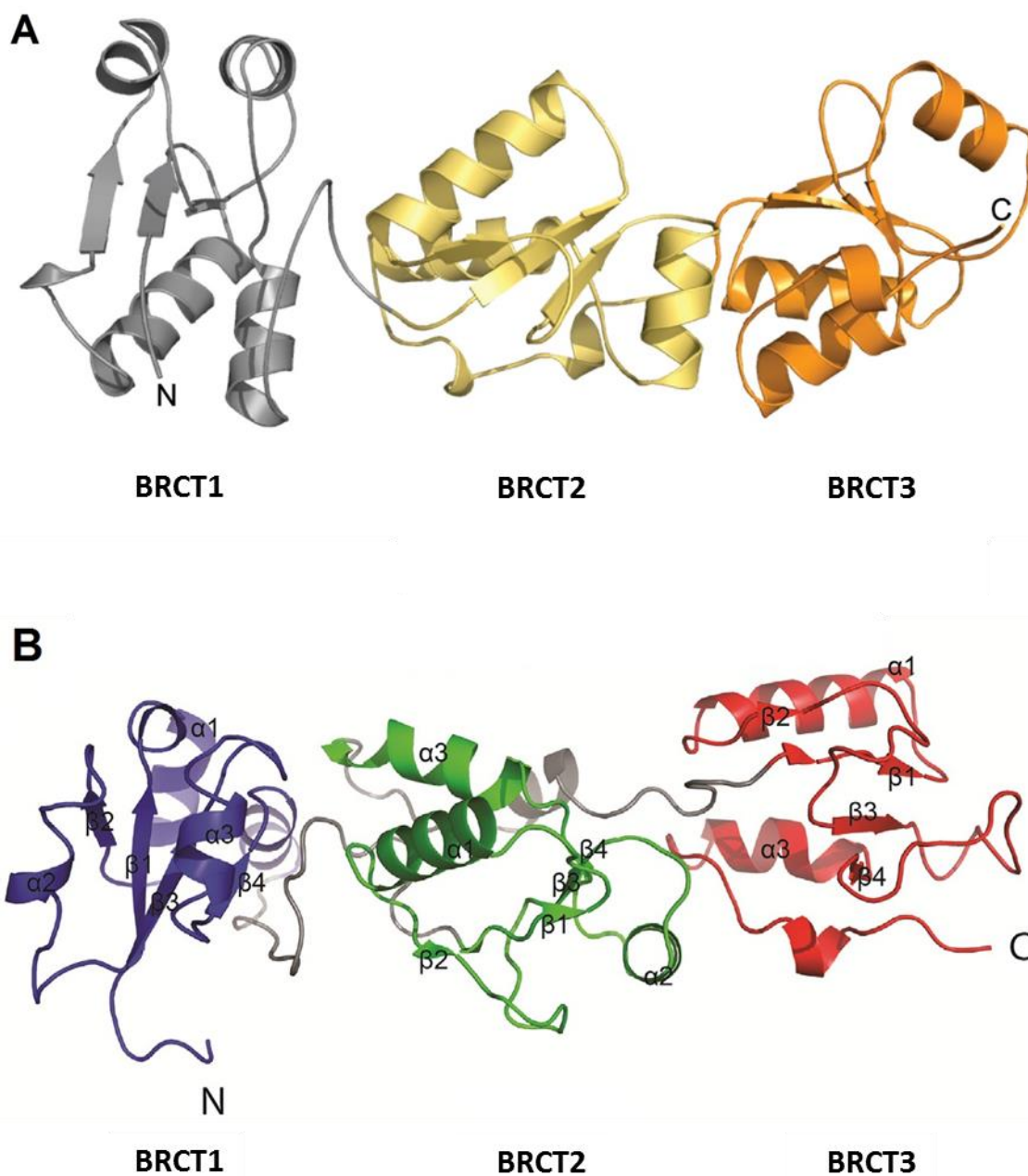


Figura 7: Estrutura do arranjo em triplete representada pelo modelo de fitas.

(A) Estrutura cristalográfica da região N-terminal da TOPBP1 humana mostrando os três primeiros domínios BRCTs que formam a arquitetura em triplete. O estudo utilizou a região os aminoácidos 1-290. (B) Estrutura cristalográfica da região N-terminal da ECT2 humana ilustrando os três domínios BRCT presentes nesta proteína, que formam um arranjo em triplete. A região utilizada no estudo compreendeu os aminoácidos 22-336. Adaptada de (RAPPAS; OLIVER; PEARL, 2011; ZOU et al., 2014).

O arranjo em triplete conserva a função de ligação de proteínas fosforiladas. A proteína TOPBP1 apresenta dois sítios distintos de ligação a fosfopeptídeos, sendo um localizado no segundo BRCT e outro no terceiro (RAPPAS; OLIVER; PEARL, 2011), enquanto ECT2 apresenta apenas um sítio de ligação a fosfopeptídeos presente no segundo BRCT (ZOU et al., 2014). Os tripletes de ECT2 e TOPBP1 são capazes de interagir com proteínas multifosforiladas, sugerindo que a presença de múltiplos BRCTs está relacionada com o reconhecimento de múltiplos sítios de fosforilação (ZOU et al., 2014).

A fosforilação é responsável pela regulação das vias de reparo do DNA, quando um erro é detectado cinases, como ATM e *ataxia telangiectasia and Rad3 related* (ATR), fosforilam proteínas de reparo modificando sua atividade (SIRBU; CORTEZ, 2013). Dessa forma, domínios que reconhecem proteínas fosforiladas, como o BRCT, são importantes nas vias de reparo por participarem da etapa de sinalização do dano (LIM; PAWSON, 2010; WOODS et al., 2012).

### 1.3. Insetos Vetores – *Aedes aegypti*

Os insetos vetores são responsáveis pela transmissão de uma diversidade de doenças infecciosas que causam mais de um milhão de óbitos anualmente (WHO, 2014). Destaca-se nesse grupo de insetos o mosquito da espécie *Aedes aegypti* que pode ser transmissor dos vírus da dengue, febre amarela, chikungunya e zika.

O *A. aegypti* é um membro da família Culicidae, a qual também pertence o *Anopheles gambiae*, mosquito responsável pela transmissão da malária. É um mosquito de hábito diurno e encontra-se bastante adaptado ao ambiente urbano. A fêmea deposita os ovos nas paredes de criadouros contendo água limpa, os quais podem permanecer viáveis por meses. Quando entram em contato com a água os ovos eclodem originando larvas, que se alimentam de micro-organismos e material orgânico particulado. Existem quatro estádios larvais, sendo a disponibilidade de alimento, temperatura e densidade populacional os fatores que determinam a mudança de estágio e o final da fase larvária originando as pupas de onde emergem os mosquitos adultos (CLEMONS et al., 2010; FUNASA, 2001) (figura 8).

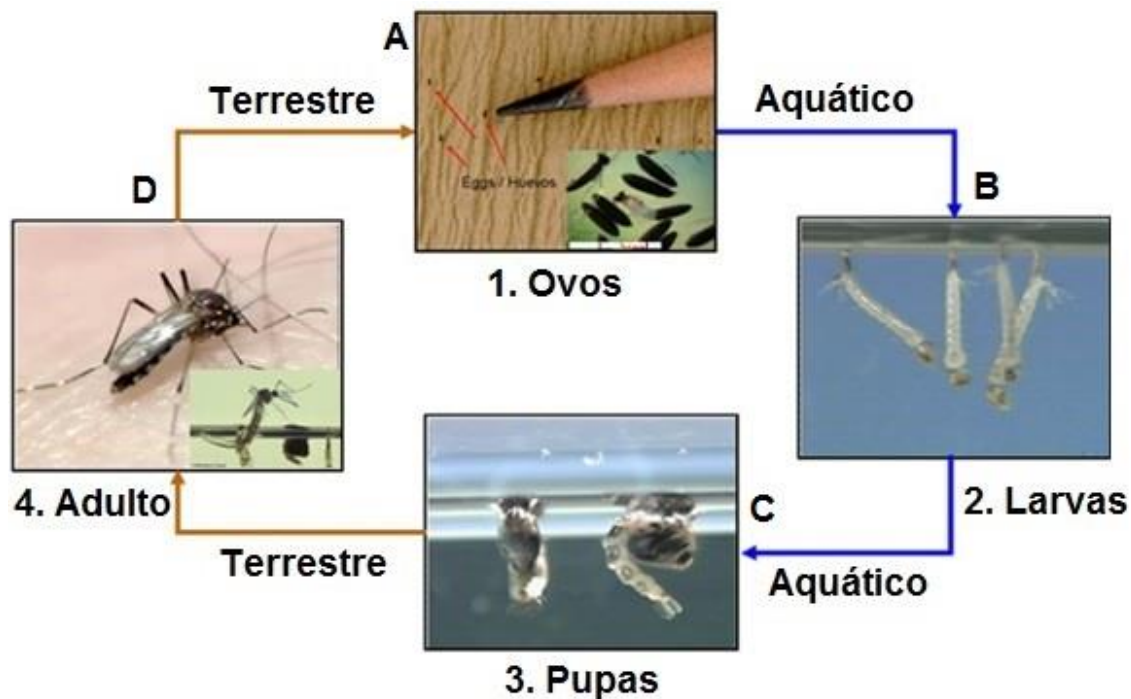


Figura 8: Ciclo de vida do mosquito *Aedes aegypti*.

A figura (A) mostra os ovos de *A. aegypti* e o momento da eclosão dando origem as larvas. As larvas (B) possuem quatro estágios de crescimento, se alimentam de micro-organismos e material orgânico particulado. Ao final do quarto estágio (L4) quando a larva estiver adquirido energia suficiente ocorre a metamorfose que transforma larvas em pupas (C) de onde emergem os mosquitos (D). Adaptado de CDC, 2012.

A dengue atualmente é uma das doenças infecciosas de maior importância no mundo, infectando aproximadamente 390 milhões por ano. É prevalente em regiões tropicais e subtropicais do planeta, se estendendo por 128 países onde 3,9 bilhões de pessoas estão sob o risco de contrair o vírus (WHO, 2015) (figura 9). No Brasil a dengue se estende por todo território, onde estão circulantes os quatro sorotipos (DEN-1, DEN-2, DEN-3 e DEN-4). Em 2013 foram notificados mais de 1 milhão de casos de dengue com mais de 200 óbitos por dengue hemorrágica (PORTAL DA SAÚDE, 2014).



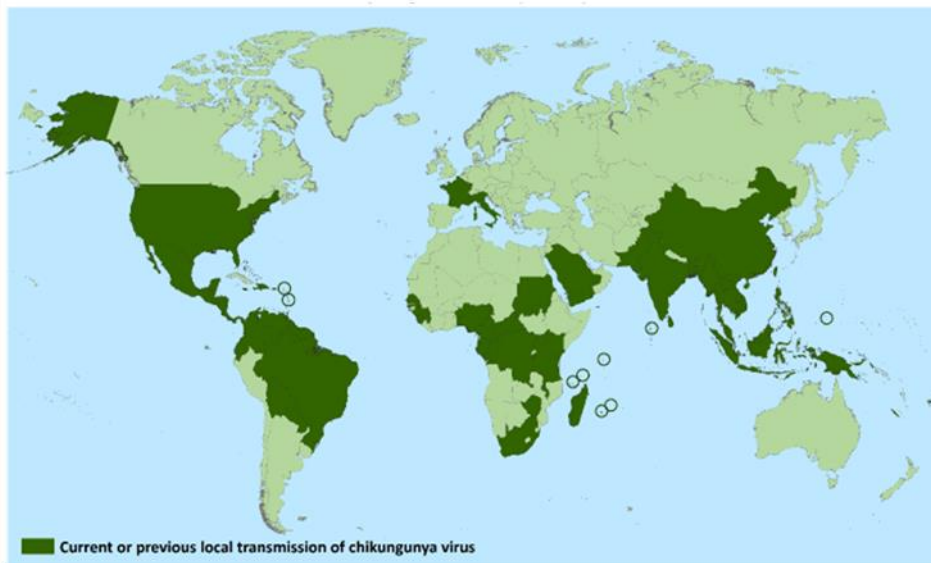
Figura 9: Áreas e países onde há risco de contrair dengue em 2013.  
Adaptado de WHO, 2013.

A infecção por dengue pode levar o indivíduo a um quadro do dengue clássico, forma de baixa letalidade, porém incapacitante ou evoluir para uma forma mais grave e letal da doença, a febre hemorrágica do dengue (FUNASA, 2001). Devido à morbidade causada pela doença a dengue constitui além de um problema de saúde um problema econômico, principalmente em países em desenvolvimento (OVERCASH et al., 2015).

O chikungunya e a zika são doenças infecciosas que têm ganhado destaque no cenário mundial devido a recente disseminação. Ambas foram isoladas e identificadas no continente africano há mais de cinquenta anos. A circulação desses vírus ficou restrita a África e Ásia, onde por um longo período ocorreram apenas surtos esporádicos. Os gráficos (figura 10) apresentam os países que já reportaram transmissão destes vírus em seu território, sendo o Brasil um deles (WHO, 2015; CDC, 2015). Atualmente atenta-se para o risco da disseminação de chikungunya e zika no Brasil, o que poderia gerar uma epidemia grave, uma vez que a população não tem imunidade ao vírus e os mosquitos transmissores (*A. aegypti* e *A. albopictus*) estão presentes em todo território brasileiro.



**Países e territórios que reportaram transmissão de chikungunya  
(Março, 2015)**



**Países que possuem evidência de transmissão de zika vírus  
(Maio, 2015)**

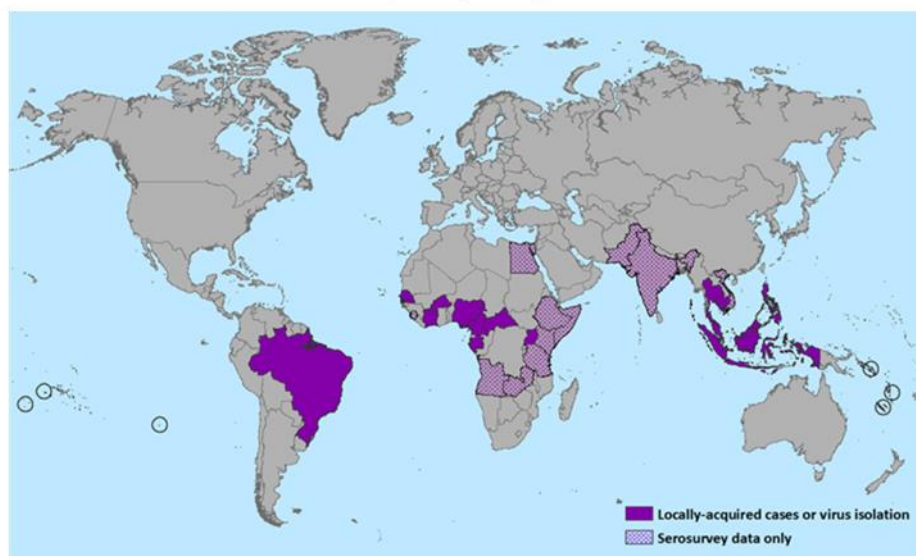


Figura 10: Países onde foi reportado transmissão de chikungunya e zika. O primeiro mapa mostra em verde os países que possuem notificação de transmissão de chikungunya, enquanto o segundo mapa mostra em roxo os países com transmissão de zika reportada em seus territórios e em lilás claro países que possuem apenas dados de inquérito sorológico. Adaptado de CDC, 2015.

A transmissão do vírus acontece na fase de acasalamento do mosquito, onde as fêmeas realizam hematofagia para garantir o desenvolvimento e maturação dos ovos. Ao picar um indivíduo infectado durante o período de viremia, a fêmea é contaminada pelo vírus tornando-se um vetor do mesmo sendo capaz de transmiti-lo através da picada.

### **1.3.1. Hematofagia e Estresse Oxidativo**

A capacidade vetorial dos insetos está intimamente relacionada a hematofagia realizada pelas fêmeas, que são capazes de ingerir, em uma refeição, uma quantidade de sangue que pode ser três vezes maior que seu peso corporal (LIMA et al., 2012; OLIVEIRA et al., 2011).

O sangue é bastante rico em proteínas, sendo a hemoglobina a mais abundante correspondendo a, aproximadamente, 60% da fração proteica (GRAÇA-SOUZA et al., 2006). A hemoglobina é uma proteína globular formada por quatro subunidades de globina ( $\alpha_2\beta_2$ ) ligadas ao grupamento prostético heme. O heme é constituído por um anel de protoporfirina IX e um átomo de  $Fe^{2+}$  que é capaz de se ligar reversivelmente ao oxigênio, conferindo a hemoglobina a função de proteína carreadora de oxigênio (BELCHER et al., 2010).

Nos insetos a hemoglobina é digerida no trato digestivo acarretando na liberação do grupamento heme, que em virtude do seu potencial em promover reações de formação de espécies reativas de oxigênio (EROs), torna-se tóxico quando livre (DANSA-PETRETSKI et al., 1995; OLIVEIRA et al., 2011; PAIVA-SILVA et al., 2006). Em adição, a degradação do heme pela heme oxigenase liberta o ferro que leva a formação de EROs através da reação de Fenton (LIMA et al., 2012; SADRZADEH et al., 1984).

As espécies reativas de oxigênio geradas durante a liberação e hidrólise de heme são capazes de promover a oxidação de lipídeos, proteínas e ácidos nucleicos, além de alterar a permeabilidade e seletividade das membranas plasmáticas devido ao seu caráter anfifílico (GRAÇA-SOUZA et al., 2006).

#### 1.4. Vias de Reparo do DNA e Proteínas contendo domínio BRCT em Insetos

As vias de reparo de DNA têm sido bastante estudadas em organismos modelo, incluindo *Drosophila melanogaster*, porém pouco se sabe sobre estas vias em mosquitos (OVERCASH et al., 2015) apesar de sua dieta diferenciada e seu consequente potencial em gerar danos por EROs, já discutido anteriormente.

Algumas proteínas da via de reparo por recombinação homóloga já foram estudadas em *D. melanogaster*, incluindo as proteínas Mre11 e Rad50 (CIAPPONI et al., 2004). Outras proteínas estudadas foram a Rad51 e Rad54, ambas são essenciais para esta via, sendo a Rad51 de maior importância, assumindo um papel central e crucial (WEI; RONG, 2007). Além disso, o nível de reparo atingido em células somáticas de *D. melanogaster* parece ser proporcional à quantidade de Rad51 (YOO; MCKEE, 2005), porém a super expressão desta proteína acarreta letalidade (YOO; MCKEE, 2004).

Na via por união terminal não homóloga, foi demonstrado que o knockout da proteína Ku70 em *Bombyx mori* gera aumento do reparo via recombinação homóloga (MA et al., 2014). Um estudo em *D. melanogaster* mostrou que indivíduos com Lig4 mutante são viáveis e produzem descendentes férteis, porém a taxa de recombinação homóloga em indivíduos normais apresenta índices abaixo de 15%, enquanto nos mutantes esses índices são superiores a 70% (BEUMER et al., 2008).

Em relação às proteínas contendo domínio BRCT já foram identificadas em *D. melanogaster*: BRCA2, PAXIP1, MDC1, MCPH1, ECT2 e TOPBP1.

BRCA2 participa da recombinação homóloga, via que aparenta ser bastante conservada em *D. melanogaster*. A depleção de BRCA2 em células de *D. melanogaster* aumenta a sensibilidade à radiação ionizante e ao tratamento com hidroxiuréia, além de aumentar as taxas de mortalidade (BROUGH et al., 2008).

A proteína PAXIP1 possui uma função conservada de controle epigenético do desenvolvimento e diferenciação em *D. melanogaster*. É essencial para formação do padrão anterior/posterior durante o desenvolvimento larval, uma vez que controla um mecanismo epigenético de metilação das histonas H3K4 e H3K27 (FANG et al., 2009).

MDC1 também teve seu homólogo (MU2) caracterizado em *D. melanogaster*. A busca por domínios conservados mostrou que MU2 apresenta um domínio FHA na porção N-terminal e na região C-terminal possui dois BRCTs em tandem. A interação desta proteína com Mre11, Rad50 e Nbs ocorre através do domínio FHA, enquanto a



interação com a histona fosforilada H2Av através do domínio BRCT tandem. Essa proteína participa do reconhecimento de danos no DNA em estágios iniciais, sendo fundamental para o reparo do DNA e para o controle de progressão do ciclo celular (DRONAMRAJU; MASON, 2009).

A proteína MCPH1 apresenta três domínios BRCT e participa da sinalização de danos no DNA e controle do ciclo celular, tendo como função a regulação da condensação de cromossomos. Essa proteína também foi identificada em *D. melanogaster*, apresentando uma localização cíclica durante o ciclo celular, estando co-localizada com o DNA na interfase, mas não com os cromossomos mitóticos. A ausência de MCPH1 em *D. melanogaster* resulta em condensação prematura dos cromossomos, gerando instabilidade genômica (BRUNK et al., 2007; RICKMYRE et al., 2007).

A proteína homóloga de ECT2 em *Drosophila* (Pebble) apresenta o mesmo padrão de domínios que a proteína humana, possuindo três domínios BRCT. Tem participação em processos de citocinese e reorganização do esqueleto de actina (SMALLHORN; MURRAY; SAINT, 2004).

A identificação do homólogo de TOPBP1 em *D. melanogaster* (MUS101) ocorreu através de mutações em larvas que geravam hipersensibilidade a agentes causadores de danos no DNA. Outros alelos de MUS101 que causavam diferentes fenótipos foram posteriormente identificados (YAMAMOTO et al., 2000).

O conhecimento das vias de reparo de DNA assim como de sua regulação frente a diversos estímulos é muito importante para entender a fisiologia e as adaptações relacionadas à hematofagia. Além disso, as vias de reparo de quebras em fita dupla têm sido estudadas como ferramenta de manipulação genômica na tentativa de criar insetos transgênicos que sejam incapazes de transmitir doenças. Dessa forma, o completo conhecimento do funcionamento e regulação destas vias é indispensável para o estabelecimento desta técnica (OVERCASH et al., 2015).

## 2. OBJETIVO

O objetivo geral do presente trabalho foi estudar as vias de reparo do DNA, através da determinação da conservação dessas vias ao longo da escala evolutiva, e do estudo no mosquito *Aedes aegypti*.

### 2.1. Objetivos específicos

Estudar a conservação das vias de reparo de DNA ao longo da escala evolutiva.

Identificar em qual ponto da escala evolutiva ocorreu a organização do domínio BRCT em triplete.

Identificar homólogos em *A. aegypti* para as proteínas presentes nas vias de reparo do DNA.

Identificar proteínas contendo domínio BRCT em *A. aegypti*.

Estabelecer condições de estímulo mediante ao estresse oxidativo para estudo das vias de reparo do DNA em *A. aegypti*.

Verificar a expressão das proteínas que participam da via de reparo de união terminal não homóloga em *A. aegypti*, mediante ao estresse oxidativo.

### 3. MATERIAIS E MÉTODOS

#### 3.1. Identificação de proteínas de reparo do DNA e dos tripletes do domínio BRCT ao longo da evolução.

Para estudar a conservação das vias de reparo do DNA foram selecionados pontos taxonômicos que abrangessem toda a escala evolutiva. Para cada um dos táxons escolhidos foi gerado um banco de dados de proteínas a partir da divisão do banco nr.

Os táxons utilizados neste estudo foram: *Mammalia* (TaxID: 40674); *Sauropsida* (TaxID: 8457); *Amphibia* (TaxID: 8292); *Coelacanthimorpha* (TaxID: 118072); *Actinopterygii* (TaxID: 7898); *Chondrichthyes* (TaxID: 7777) ; *Cephalochordata* (TaxID: 7735); *Tunicata* (TaxID: 7712); *Hemichordata* (TaxID: 10219); *Echinodermata* (TaxID: 7586); *Protostomia* (TaxID: 33317); *Platyhelminthes* (TaxID: 6157); *Cnidaria* (TaxID: 6073); *Placozoa* (TaxID: 10226); *Fungi* (TaxID: 4751); *Choanoflagellida* (TaxID: 28009); *Viridiplantae* (TaxID: 33090); *Euglenozoa* (TaxID: 33682); *Amoebozoa* (TaxID: 554915); *Alveolata* (TaxID: 33630); *Archea* (TaxID: 2157); *Bacteria* (TaxID: 2).

##### 3.1.1. Distribuição evolutiva das vias de reparo do DNA

A primeira etapa do estudo da conservação das vias de reparo do DNA foi determinar a distribuição evolutiva das proteínas das vias de reparo do DNA. Dessa forma, foram buscadas, no banco de dados de ortologia KEGG, as proteínas das vias de reparo em cada um dos pontos taxonômicos, identificando as preditas e as ainda não encontradas. Além disso, foram identificadas no KEGG as proteínas das vias de reparo preditas para humanos. Essas proteínas foram utilizadas como *query* para buscar as ausentes nos outros pontos taxonômicos.

O programa MolFuncs (NATARAJAN; JAKOBSSON, 2009) foi utilizado para buscar proteínas das vias de reparo do DNA, ausentes no KEGG, para os pontos taxonômicos selecionados. Esse programa implementou uma estratégia de BLAST recíproco. O primeiro passo dessa estratégia consistiu na realização de um BLAST, entre a proteína *query* (proteínas humanas preditas no KEGG) e o banco de dados de *Mammalia*, com intuito de identificar todas as proteínas de mamíferos ortólogas à *query*. Em seguida foi feito um BLAST entre *query* e os bancos de dados dos táxons,

encontrando proteínas chamadas de “candidatos”. Finalizando, foi efetuado um BLAST entre os candidatos e o banco de dados de *Mammalia*, sendo esperado que o candidato encontrasse alguma das proteínas do grupo de ortólogos, identificados na primeira etapa. Para confirmar se o candidato corresponde à proteína de interesse, os valores de *score* dos BLASTs foram analisados, e a ferramenta *online* Web CD-Search Tool (disponível em <http://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) (MARCHLER-BAUER et al., 2009) foi utilizada para buscar os domínios conservados. A figura 11 mostra um exemplo da metodologia.

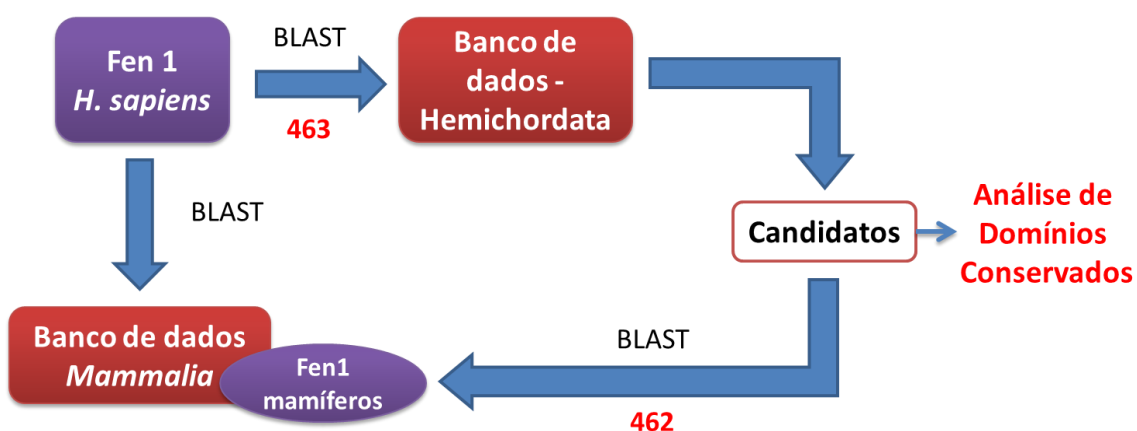


Figura 11: Exemplo da metodologia de BLAST recíproco.

Neste exemplo, a proteína Fen1 estava ausente em *Hemichordata*. O primeiro passo da metodologia consistiu em um BLAST entre a Fen1 humana (*query*) para encontrar a Fen1 em outros mamíferos. Em seguida, foi feito um BLAST da Fen1 humana contra o banco de dados de *Hemichordata*. No terceiro passo foi feito um BLAST entre os candidatos encontrados e o banco de dados de *Mammalia*, sendo esperado que os candidatos encontrassem a Fen1 humana ou a Fen1 de qualquer outro mamífero. Por último, foram analisados os scores dos BLASTs e os domínios conservados para determinar se os candidatos correspondiam a Fen1. Os números abaixo das setas representam o score do melhor resultado de blast.

### 3.1.2. Identificação dos tripletes

Na segunda etapa do estudo da conservação das vias de reparo do DNA foi feita a busca por tripletes de BRCT, nas proteínas TOPBP1 e ECT2, ao longo da evolução. Para identificar os tripletes, as TOPBP1 e ECT2 humanas, que já possuem este arranjo estrutural identificado e cristalizado em humanos, (RAPPAS; OLIVER; PEARL, 2011; ZOU et al., 2014) foram usadas como referência.

Para identificar os tripletes, foram utilizadas como referência as proteínas

TOPBP1 e ECT2 humanas, que já possuem este arranjo estrutural identificado e cristalizado (RAPPAS; OLIVER; PEARL, 2011; ZOU et al., 2014).

Primeiramente, o programa HMMSEARCH foi utilizado para filtrar dos bancos de dados de cada táxon, as proteínas contendo domínio BRCT (PF00533), disponível na base de dados do *Protein family database* (PFAM) (FINN et al., 2014). Posteriormente, foi feito um BLAST entre as proteínas contendo BRCT filtradas e os bancos de dados Uniprot/Swiss-prot, nr\_no\_hypothetical\_95 (banco de dados preparado a partir do banco nr com remoção de sequências sem anotação e mais remoção de sequências redundantes ao nível de 95% de identidade) e Gene Ontology (GO). Além disso, todos os domínios conservados, presentes nas proteínas selecionadas, foram identificados com o HMMSCAN utilizando os bancos de dados PFAM-A e PFAM-B. Todas as etapas citadas acima foram realizadas usando o *software* FAT (SEABRA-JUNIOR et al., 2011), desenvolvido em nosso laboratório, que utiliza localmente os programas HMMER (FINN; CLEMENTS; EDDY, 2011) e BLAST (ALTSCHUL et al., 1990).

Em seguida, foi selecionada uma TOPBP1 de um organismo de cada táxon, sendo escolhida a TOPBP1 da espécie que possuía a maior quantidade de proteínas contendo domínio BRCT no táxon. Quando a espécie não possuía TOPBP1 ou esta se encontrava truncada, foi utilizada a TOPBP1 da segunda espécie com mais proteínas com domínio BRCT. A ferramenta HMMSCAN foi utilizada para localizar a região correspondente aos três primeiros domínios BRCT de TOPBP1, esta parte da proteína foi extraída e depois alinhada utilizando o *software* online PRALINE. A mesma metodologia foi utilizada para ECT2.

O programa JALVIEW (WATERHOUSE et al., 2009) foi utilizado para visualizar o alinhamento múltiplo e carregar a estrutura secundária predita feita com os programas JNET (COLE; BARBER; BARTON, 2008; CUFF; BARTON, 2000), PSIPRED (BUCHAN et al., 2013) e HMMSCAN.

Para comparar os BRCTs presentes nos tripletes foi feito um dendograma desses BRCTs alinhados individualmente. Os BRCTs foram extraídos dos tripletes, sendo a predição de estrutura secundária utilizada para identificar o início e o final de cada BRCT. Os alinhamentos múltiplos foram feitos no PRALINE (SIMOSSIS; HERINGA, 2003, 2005) com as condições padrão, e os dendogramas foram construídos com o programa RAXML (STAMATAKIS, 2006) utilizando a matriz de substituição *Jones Taylor Thornton* (JTT) (JONES; TAYLOR; THORNTON, 1992), sendo realizada uma análise *Bootstrap* com 500 réplicas para avaliar a confiança dos ramos. O *software*

FigTree foi utilizado para visualização das árvores. Esta metodologia foi realizada para o triplete de TOPBP1 e de ECT2.

### 3.2. Estudo das Vias de Reparo no Mosquito *A. aegypti*

Nesta etapa do presente trabalho foi feita a busca por proteínas das vias de reparo de DNA e proteínas contendo o domínio BRCT em *A. aegypti*. Foram realizadas, também, análises de expressão em situações de dano ao DNA.

#### 3.1.3. Identificação de proteínas de reparo e contendo domínio BRCT em *A. aegypti*

##### 3.1.3.1. Confirmação das proteínas preditas no banco de dados

As proteínas relacionadas às vias de reparo do DNA preditas, no banco de dados KEGG, tiveram sua anotação conferida, utilizando o programa FAT, onde foram realizados BLASTs entre estas proteínas e os bancos de dados nr e swiss\_prot, e contra transcriptomas disponíveis para *Aedes aegypti*. Para identificação dos domínios conservados foi utilizada a ferramenta HMMSCAN associada aos bancos PFAM-A e PFAM-B. Estes resultados foram utilizados na curagem manual da anotação.

##### 3.1.3.2. Identificação de novas proteínas

Para buscar as proteínas não identificadas em *A. aegypti*, foram selecionados ortólogos de dois ou mais organismos evolutivamente próximos, os quais foram denominados “iscas”.

Para comparar as iscas com as proteínas preditas no genoma de *A. aegypti*, foram realizados BLASTs entre as iscas e um banco de dados de proteínas preditas, e contra transcriptomas de *A. aegypti* presentes na literatura. Os resultados permitiram a obtenção de candidatos que foram submetidos às mesmas análises e curagem manual descritas no item 3.2.1.1. Os BLASTs foram realizados no programa FAT e os resultados limitados a um valor de e-value  $10^{-10}$ .

Foi feito, também, um BLAST para comparar as iscas com o genoma bruto de *A. aegypti* com o objetivo de identificar regiões genômicas, sem proteínas preditas, que possam codificar um possível homólogo, as quais foram denominadas de regiões candidatas.

### 3.1.3.3. Busca por proteínas expressas

O último passo dessa etapa do presente trabalho consistiu na busca por indícios de expressão das proteínas das vias de reparo do DNA de *A. aegypti*. Para determinar os indícios de expressão, todas as proteínas identificadas nas análises anteriores (preditas no KEGG e encontradas através das metodologias aplicadas no tópico 3.1.3.2) foram submetidas a um BLAST contra transcriptomas disponíveis na literatura para este inseto. O resultado do BLAST indicou quais proteínas estavam sendo expressas nas condições utilizadas para sequenciamento dos transcriptomas da literatura.

## 3.1.4. Determinação de condições de estímulo e análises de expressão

### 3.1.4.1. Eclosão dos ovos

Os ovos de *Aedes aegypti* linhagem *Black Liverpool* foram, gentilmente, cedidos pelo Professor Mário Alberto Cardoso Silva-Neto do Laboratório de Sinalização Celular (LabSiCel), do Instituto de Bioquímica Médica Leopoldo de Meis, da Universidade Federal do Rio de Janeiro, Rio de Janeiro. Os mosquitos foram criados em insetário em ciclo claro/escuro de 12h, temperatura a 28 °C e 80% de umidade relativa do ar. Os adultos foram mantidos em gaiolas com solução de sacarose 10%. As fêmeas foram alimentadas naturalmente em veias da orelha de coelhos, ou artificialmente com sangue de coelho heparinizado. Os cuidados com os animais e os protocolos experimentais seguiram as diretrizes do Comitê de avaliação do uso de animais em pesquisa da UFRJ, CAUAP-UFRJ e do Guia NIH para o uso e cuidado de animais de laboratório (ISBN 0-309-05377-3). Os protocolos foram aprovados pelo CAUAP-UFRJ (registro IBQM067-05/16).

Para realizar a eclosão, um papel de filtro contendo ovos de *A. aegypti* foi colocado no fundo de uma bandeja plástica contendo aproximadamente 1 L de água

filtrada e 3 grãos de ração de gato. A bandeja foi mantida em estufa BOD a 28 °C durante 3 dias, tempo necessário para que as larvas atinjam o estágio L3 de crescimento.

#### 3.1.4.2. Curva de sobrevivência – Paraquat

As larvas em estágio L3 de crescimento foram submetidas ao tratamento com paraquat (SIGMA-ALDRICH), um herbicida com potencial ação pró-oxidante, para construir uma curva de sobrevivência em função da concentração e tempo de exposição.

Para a construção da curva foram adicionadas 5 larvas a tubos Falcon de 15 mL contendo 1 mL de água filtrada com paraquat 0,5 mM, 0,75 mM, 1 mM, 2 mM e 4 mM. Para que as larvas permanecessem alimentadas durante o experimento, ~ 10 mg de ração pulverizada foi acrescida aos tubos Falcon.

As larvas foram mantidas em BOD a 28°C e monitoradas de 2 em 2 horas durante 24 horas para observar a mortalidade. A curva foi obtida a partir da média e erro padrão de três experimentos independentes, cada um deles com pontos em triplicata.

#### 3.1.4.3. LD<sub>50</sub>

A metodologia para o cálculo da LD<sub>50</sub> foi semelhante a descrita para a curva de sobrevivência, alterando apenas as concentrações. As concentrações utilizadas foram: 0,001 mM; 0,01 mM; 0,1 mM; 1 mM; 5 mM; 10 mM e 50 mM de paraquat e a medida foi feita no tempo de 24 h de incubação. O programa GraphPad Prism versão 5.03 foi utilizado para plotar o gráfico, sendo feita uma curva dose-resposta de três parâmetros aplicando o método *Dose-response – Stimulation, log (agonist) vs response (three parameters)*. Este modelo de cálculo assume que a curva possui *slope* padrão igual a 1,0 (*slope de Hill*).

#### 3.1.4.4. Extração e dosagem de RNA

O RNA total foi extraído de três grupos de larvas: controle e tratadas com paraquat 0,5 mM durante 12 e 24 horas. A quantidade de larvas utilizada na extração variou de 20 a 60 larvas.

O RNA total das larvas foi extraído utilizando o kit illustra RNAspin Mini RNA



Isolation da GE healthcare seguindo o protocolo padrão do mesmo. As larvas foram transferidas para um eppendorf e toda água presente foi retirada, em seguida foi adicionado 20  $\mu$ L do tampão de lise para homogeneização e maceração das larvas, realizada com auxílio de um macerador, durante 2 minutos. O protocolo padrão do kit segue descrito brevemente. O homogeneizado foi filtrado para remoção de restos de tecidos, e aplicado em uma membrana de sílica para que ocorra a adsorção do RNA. Em seguida, foi feito o tratamento com DNase I para degradação do DNA e lavagens subsequentes para retirada de contaminantes, proteínas e outros componentes celulares. Ao final do procedimento a concentração de RNA foi determinada por Qubit® Fluorometer (Invitrogen) e a integridade em gel de agarose 1% (TAKAHASHI; OGINO; BABA, 1968), utilizando GelRed como corante fluorescente e padrão peso adequado.

A dosagem de RNA por Qubit foi feita seguindo o protocolo do kit Qubit® RNA Assay, que consiste na adição de um corante fluorescente específico à amostra para posterior leitura fluorimétrica pelo Qubit.

#### 3.1.4.5. Síntese de cDNA

A síntese de cDNA foi realizada em termociclador (Applied Biosystems - Veriti 96-Well Thermal Cycler) utilizando o High-Capacity cDNA Reverse Transcription Kit (Applied biosystems). Para a síntese do cDNA foi utilizado 1  $\mu$ g do RNA extraído e 2  $\mu$ L de Primer randômico, este mix foi incubado no termociclador por 10 minutos a 65°C. Subsequentemente foi adicionado 2  $\mu$ L do Mix 10x (tampão, nucleotídeos e transcriptase reversa) completando o volume a 20  $\mu$ L com água DEPC e incubando novamente no termociclador (10 min à temperatura de 25°C; 2 horas à 37°C e 15 min à 70°C).

#### 3.1.4.6. PCR quantitativo

Para realização do PCR quantitativo (qPCR) foram desenhados iniciadores para os genes preditos na via de reparo por união terminal não homóloga de *Aedes aegypti*. Os desenhos dos iniciadores objetivou a obtenção de amplificadas na faixa de 150-200 bp. Para evitar a amplificação de possíveis contaminantes de DNA genômico todos os iniciadores foram desenhados para anelar no cDNA nas junções de éxons. A tabela 1

mostra a sequência dos iniciadores utilizados.

Tabela 1: Tabela de iniciadores utilizados no PCR quantitativo.

<b>Gene</b>	<b>Iniciador</b>	<b>Tamanho</b>
<i>Ku80</i> AAEL003684	<b>Senso:</b> CAATGTGCTCATGCCAACTT	187
	<b>Anti senso:</b> GTCCAATGCGTCCATTAGGT	
<i>Mre11</i> AAEL000034	<b>Senso:</b> GATGTGATTAGAGGAGAGGACAG	173
	<b>Anti senso:</b> TCACCCAACAGATACGTTTTTC	
<i>Rad27</i> AAEL005870	<b>Senso:</b> TCGGACGAAAAGTAGCAATCG	161
	<b>Anti senso:</b> CTTGATTCCATTTTCCAGCA	
<i>Rad50</i> AAEL005245	<b>Senso:</b> AAGAAATGAAGAAGATTATTGCTGC	172
	<b>Anti senso:</b> ATCAATGCGTCCAGACCTT	
<i>XLf</i> AAEL002939	<b>Senso:</b> CTAAAAGAAATCTAGTAGCGACCAA	186
	<b>Anti senso:</b> CGTAGGACTTCCCGATTCTA	
<i>DNAPKcs</i> AAEL008123	<b>Senso:</b> GCATTCAGAAACGCCTTGAT	199
	<b>Anti senso:</b> GATCTTCACGAGCCAACAGT	
<i>Lig4</i> AAEL017365	<b>Senso:</b> GTGAAATAAGTGCCGTCCTG	162
	<b>Anti senso:</b> GGATAGGAAATATGCTGGTATTCG	
<i>Rp49</i> (endógeno) AAEL003396	<b>Senso:</b> GCTATGACAAGCTTGCCCCCA	190
	<b>Anti senso:</b> AGCTGGAGGTGCTGATGA	

As reações de qPCR foram conduzidas utilizando o reagente HOT FIREPol EvaGreen® qPCR Mix Plus (Solis Biodyne) segundo instruções do fabricante. Resumidamente, reações com volume final de 10  $\mu$ L foram feitas utilizando 2  $\mu$ L do reagente, 4 picomols de cada um dos primers (tabela 1) e 2  $\mu$ L da amostra diluída.

As reações foram realizadas em termociclador LineGene 9600 (Bioer) seguindo o esquema abaixo (tabela 2):

Tabela 2: Esquema das reações de PCR quantitativo

Etapa	Temperatura (°C)	Tempo (min)
1	95	15
2	95	00:15
3	60	00:20
4	72	00:20

As etapas 2 a 4 foram repetidas por 50 ciclos, seguidas por curva de desnaturação.

A validação das reações foi feita utilizando a inclinação e o valor de  $R^2$  de curvas de amplificação com diferentes quantidades de cDNA. Foram utilizadas amostras de cDNA de amostras controle. Os pontos da curva foram preparados através de diluição seriada de fator 5. Foi utilizada a diluição de cDNA 1:100 para as reações.

Os níveis de transcrição de cada gene foram analisados através da diferença de expressão entre o gene alvo e o controle endógeno (*Rp49*) de uma mesma amostra, medido através do *threshold cycle* ( $\Delta Ct$ ). A expressão relativa entre o grupo de larvas tratadas com paraquat por 12 e 24 h e o grupo controle foi analisada através da diferença dos valores de  $\Delta Ct$  de cada amostra tratada contra a média de  $\Delta Ct$  do grupo controle através da fórmula:  $2^{-\Delta\Delta Ct}$  (LIVAK; SCHMITTGEN, 2001).

## 4. RESULTADOS

### 4.1. Identificação de proteínas de reparo do DNA e dos tripletes do domínio BRCT ao longo da evolução.

#### 4.1.1. Distribuição evolutiva das vias de reparo do DNA

Como primeira etapa do estudo da conservação das vias de reparo do DNA foi feita a determinação da distribuição evolutiva das proteínas das vias de reparo do DNA. O *software* MolFuncs (NATARAJAN; JAKOBSSON, 2009) foi utilizado para buscar proteínas das vias de reparo ausentes, no KEGG, para os táxons estudados neste trabalho. Para *Hemichordata*, não foram encontradas proteínas preditas no KEGG, dessa forma foi feita a completa identificação das proteínas das vias de reparo deste táxon, sendo as encontradas listadas na tabela 3. Foram também encontradas proteínas ausentes no KEGG para outros pontos taxonômicos, lista que está apresentada na tabela 4.

Tabela 3: Lista de proteínas das vias de reparo do DNA encontradas em *Hemichordata*.

<b>Proteína</b>	<b>gi</b>	<b>Via de Reparo</b>
OGG	291230163	BER
NTH	585649842	BER
NEIL1	585649842	BER
APE1	2912220201	BER
APE2	291239611	BER
XRCC1	62087142	BER
Pol $\beta$	585712411	BER
Lig3	585652306	BER
UNG	585655422	BER
MUTY	585661637	BER
MPG	291229341	BER
MBD4	442796450	BER
PCNA	585678593	BER, NER, MMR
Pol $\delta$	585722296	BER, NER, MMR, HR
Pol $\epsilon$	585652800	BER, NER
Fen1	585698647	BER
Lig1	585671094	BER, NER, MMR
Pol $\lambda$	585655893	BER, NHEJ
HMBG1	283462216	BER
PARP	PARP1: 585683224 PARP2: 585672552 PARP3: 585704687 PARP4: 585674998	BER
RBX1	291234484	NER
Cul4	585706604	NER
DDB1	585646570	NER
DDB2	291222681	NER
XPC	585697964	NER
HR23B	585700354	NER
CSA	291227378	NER
CSB	291227378	NER
CDK7	291227378	NER
MNAT1	291227378	NER
CCNH	291227378	NER
XPB	291227378	NER
XPD	585720977	NER
TFIIH1	585652251	NER
TFIIH2	585723453	NER
TFIIH3	585702073	NER
TFIIH4	585673521	NER
XPG	585684961	NER
XPA	585654037	NER
RPA	585646261	NER, MMR, HR
XPF	585699715	NER
ERCC1	585700160	NER

Tabela 3: Lista de proteínas das vias de reparo do DNA encontradas em *Hemichordata*.  
Continuação

<b>Proteína</b>	<b>gi</b>	<b>Via de Reparo</b>
RFC	585683961	NER, MMR
PMS2	291233099	MMR
MLH1	291231519	MMR
MLH3	585645287	MMR
ExoI	291241252	MMR
Rad50	585661521	HR, NHEJ
Mre11	585693155	HR, NHEJ
Nbs1	585665417	HR
Rad51	291235734	HR
Rad52	585722477	HR
BRCA2	119395734	HR
BLM	585692617	HR
TOP3	TOP3A 585723571 TOP3B 585652243	HR
Mus81	585647428	HR
XRCC2	291226354	HR
XRCC3	585644175	HR
Rad51B	585687784	HR
Ku70	585646578	NHEJ
Ku80	585701227	NHEJ
Artemis	585660446	NHEJ
DNAPKcs	585664764	NHEJ
TdT	291223873	NHEJ
Lig4	585650410	NHEJ
XRCC4	585706368	NHEJ

Tabela 4: Listagem das proteínas encontradas nos outros pontos taxonômicos.

<b>Ponto Taxonômico</b>	<b>Proteína - gi</b>	<b>Via de Reparo</b>
<i>Amphibia</i>	CSB – 301614708	NER
<i>Tunicata</i>	MNAT1 – 198414152 MLH3 – 699242107 Rad50 – 699237672	NER MMR HR, NHEJ
<i>Echinodermata</i>	Ku70 – 115618058	NHEJ
<i>Platyhelminthes</i>	RBX1 – 93009102 CSA – 674262333 TFIIH3 – 674260349 Rad51B – 675369313 DNAPKcs – 358332995 Lig3 – 674588724 576694139	NER NER NER HR NHEJ BER
<i>Cnidaria</i>	BLM – 749719332	HR
<i>Porifera</i>	Cul4 – 340369300	NER
<i>Fungi</i>	DNAPKcs – 552908684	NHEJ
<i>Choanoflagellida</i>	NEIL1 – 514694550 CSB – 514682141 TFIIH2 – 514700126 Nbs1 – 514695150 Rad51 – 514692457 XRCC3 – 514652869	BER NER NER, HR HR HR
<i>Viridiplantae</i>	Lig3 – 3264984831	BER
<i>Euglenozoa</i>	Cul4 – 72393253 Lig4 – 528221976	NER NHEJ
<i>Bacteria</i>	Fen1 – 516974532 PCNA – 585683961	BER BER, NER, MMR

O resultado final da busca está apresentado na figura 12, que organiza as vias em etapas, agrupando as proteínas para cada uma delas, como apresentado na tabela 5. Alguns táxons, principalmente *Bacteria*, *Archaea* e *Fungi* possuem vias que não são realizadas pelas mesmas proteínas que participam das vias de humano, para muitos destes casos existem proteínas particulares que executam determinada função, sendo consideradas proteínas análogas.

Tabela 5: Proteínas das vias de reparo do DNA agrupadas por etapa.

BER: reparo por excisão de base, NER: por excisão de nucleotídeo, MMR: por erro de pareamento, HR: por recombinação homóloga, NHEJ: por união terminal não homóloga.

Via	Deteção	Processamento	Síntese	Ligação/Recombinação (HR)
<b>BER</b>	OGG1 NTH NEIL1 NEIL2 NEIL3 UNG TDG/Mug SMUG MUTY MPG MBD4	APE1 APE2	XRCC1 Pol $\beta$ PCNA Pol $\delta$ Pol $\epsilon$ Fen1	Lig3 Lig1
<b>NER</b>	RBX1 Cul4 DDB1 DDB2 XPC HR23B CENT2 CSA CSB	CDK7 MNAT1 CCNH XPB XPD TFIIH1 TFIIH2 TFIIH3 TFIIH4 TTDA XPA XPG RPA XPF ERCC1	Pol $\delta$ Pol $\epsilon$ PCNA RFC	Lig1
<b>MMR</b>	PMS2 MLH1 MSH2 MSH6 MSH3 MLH1 MLH3 PCNA RFC	ExoI	RPA Pol $\delta$	Lig1
<b>HR</b>	Rad50 Mre11 Nbs1	RPA Rad51 Rad52 BRCA2 DSS1 Rad54 Rad51B Rad51C Rad51D XRCC2 XRCC3	Pol $\delta$	BLM TOP3 Mus81 Eme1
<b>NHEJ</b>	Ku70 Ku80	Artemis DNAPKcs Pol $\lambda$ Pol $\mu$ TdT	X	Lig4 XRCC4 XLF



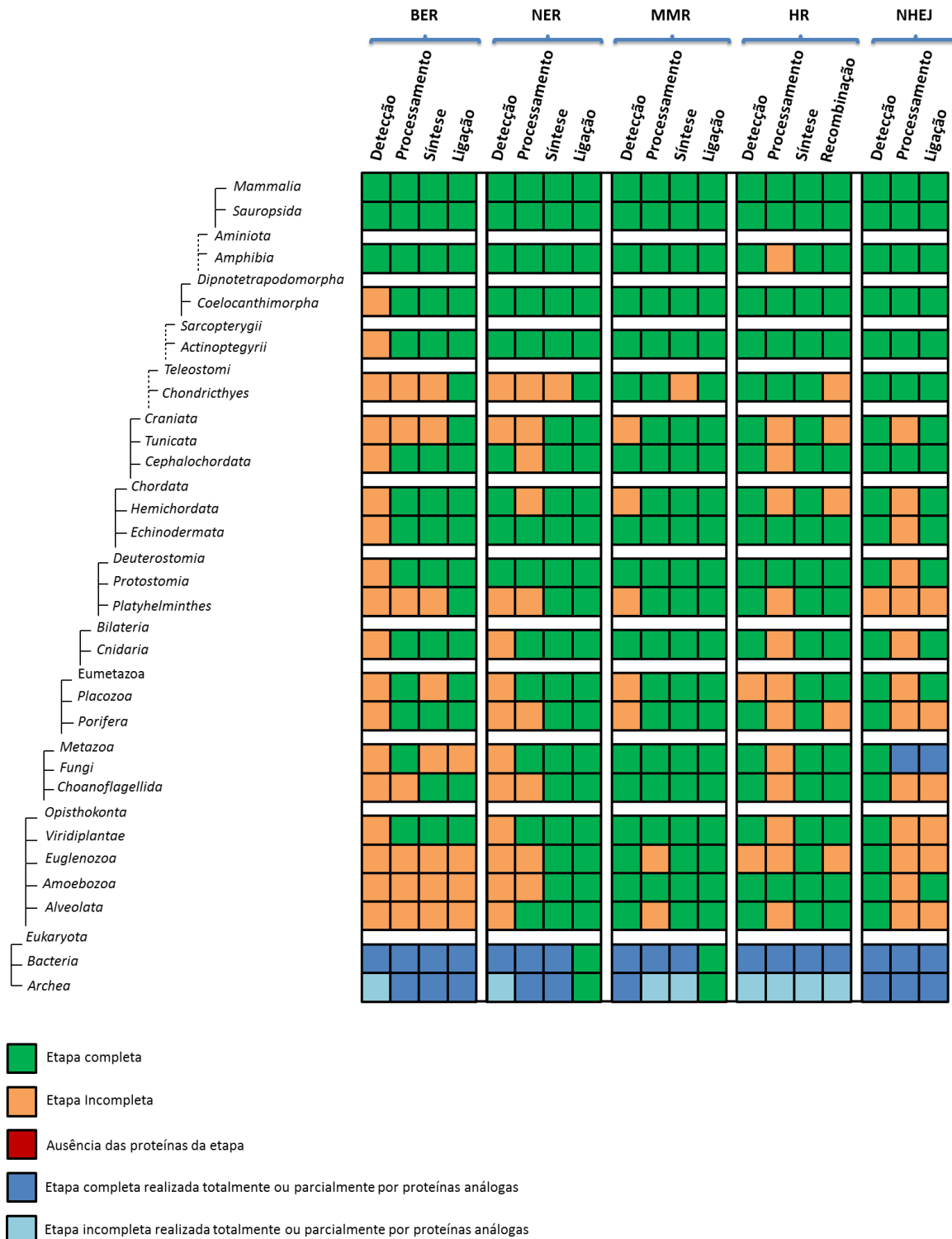


Figura 12: Distribuição evolutiva das vias de reparo.

BER: reparo por excisão de base; NER: reparo por excisão de nucleotídeo, MMR: reparo por erro de pareamento; HR: reparo por recombinação homóloga; NHEJ: reparo por união terminal não homóloga. A figura divide as vias por etapas agrupando as proteínas como descrito na tabela 5. Etapa incompleta significa que não foram identificadas uma ou mais proteínas que participam da etapa em questão. As cores azul escura e clara referem-se a etapas que são funcionais, mas são realizadas por proteínas com baixíssima similaridade (análogas) as proteínas de humanos. Para NER (procaríotos) é sabido não haver ancestral comum com as proteínas de outros prontos taxonômicos (EISEN; HANAWALT, 1999). As linhas pontilhadas representam ligação indireta entre os pontos taxonômicos.

Através da figura 12, é possível observar que as vias de reparo estão bastante conservadas ao longo da evolução. Além disso, a presença de proteínas análogas em procariotos evidencia a conservação da função de reparo. Os resultados também indicam que a via de reparo por erro de pareamento (MMR) parece ser a mais conservada. Em relação à conservação das vias de reparo por excisão de base (BER) e excisão de nucleotídeo (NER), observou-se que a BER se encontra menos conservada do que a NER.

Em *Chondrichthyes*, através das análises descritas na metodologia, não foi possível encontrar a *DNA ligase 1* (Lig1). Entretanto, a Lig1 estava presente em todos os pontos taxonômicos, inclusive em procariotos, sugerindo que as análises realizadas não foram suficientes para identificação de Lig1 neste ponto taxonômico. Dessa forma, na tentativa de encontrar a Lig1 em *Chondrichthyes*, foi feito um BLASTp entre o banco de dados deste táxon e o tsa (banco de dados de transcriptomas), utilizando como *query* a Lig1 de *Actinopterygii*. Através desta metodologia foi possível encontrar a Lig1 em *Chondrichthyes* (*Callorhinchus milii* – gi 398672954), com identidade de 63% e cobertura de 70%.

Em *Fungi*, a via de reparo por união terminal não homóloga (NHEJ) teve sua etapa de processamento marcada com quadrado azul. A cor azul foi atribuída porque, em *S. cerevisiae*, esta etapa é realizada pelo complexo MRX, enquanto em mamíferos a mesma etapa é orquestrada por Artemis e DNAPKcs. Em *Fungi*, o complexo MRX, que compreende as proteínas Rad50, Mre11 e XRS2, também participa da etapa de detecção da via de reparo por recombinação homóloga (HR). Em mamíferos, na HR, o reconhecimento da quebra na dupla fita ocorre através do complexo MRN, semelhante ao MRX, sendo formado pelas proteínas Rad50, Mre11 e Nbs. Apesar da diferença entre a etapa de processamento entre mamíferos e fungos, no decorrer das análises realizadas neste trabalho, foi possível encontrar a DNAPKcs (gi 552908684) em *Fungi*.

#### **4.1.2. Distribuição Evolutiva do Triplete BRCT de TOPBP1 e ECT2**

A segunda etapa do estudo da conservação das vias de reparo do DNA consistiu na busca por tripletes de BRCT ao longo da evolução. Dessa forma, foram buscados, nos pontos taxonômicos escolhidos, os ortólogos das proteínas humanas TOPBP1 e ECT2, que apresentam arranjo em triplete do domínio BRCT. Para identificar estes tripletes, vários resultados foram considerados, incluindo: identificação dos domínios

conservados BRCT (com a ferramenta HMMSCAN), posicionamento sobre a região de triplete da TOPBP1 humana em um alinhamento múltiplo das proteínas completas (utilizando o software online PRALINE), e a predição de estrutura secundária (determinada pelos programas HMMSCAN e PSIPRED). A estrutura secundária de TOPBP1 humana do cristal 2XNK (RAPPAS; OLIVER; PEARL, 2011) foi usada como base para esta análise.

A TOPBP1 foi encontrada em quase todos os pontos taxonômicos estudados, com exceção de *Bacteria*, *Archaea*, *Alveolata* e *Porifera* (tabela 6). Foi feito um alinhamento múltiplo com as TOPBP1, sendo utilizada uma TOPBP1 por táxon (figura 13).

A estrutura em triplete foi encontrada truncada em alguns pontos taxonômicos, apesar dos esforços para buscar um triplete íntegro em outros organismos do mesmo ponto taxonômico (*Fungi*, *Hemichordata* e *Placozoa*).

Foi possível observar uma maior variabilidade nas duas primeiras fitas beta e na primeira alfa-hélice (figura 14A) do BCRT1, enquanto os BRCTs 2 e 3 apresentaram estruturas secundárias muito mais conservadas e semelhantes (figura 14B e 14C).

O BRCT1 não foi identificado em *Fungi*, estando de acordo com diferenças na arquitetura dos domínios BRCT já identificadas na literatura (MESQUITA et al., 2010). Também foi possível observar, em *Hemichordata*, a ausência do BRCT1 e de parte do BRCT2, que só possui a última fita beta e alfa hélice (figura 14A e 14B). O BRCT2 também estava truncado em *Placozoa*, estando ausentes as duas últimas alfa hélices e a última fita beta (figura 14B).

Tabela 6: Tabela contendo as proteínas TOPBP1 identificadas e utilizadas no alinhamento de seus tripletes de BRCT.

<b>Táxons</b>	<b>Espécie - gi</b>
<i>Mammalia</i>	<i>Homo sapiens</i> : 296453012
<i>Sauropsida</i>	<i>Chelonia mydas</i> : 591379654
<i>Amphibia</i>	<i>Xenopus (Silurana) tropicalis</i> : 512848227
<i>Coelacanthimorpha</i>	<i>Latimeria chalumnae</i> : 557019128
<i>Actinopterygii</i>	<i>Danio rerio</i> : 528519341
<i>Chondrichthyes</i>	<i>Callorhinchus milii</i> : 632965729
<i>Cephalochordata</i>	<i>Branchiostoma floridae</i> : 260805410
<i>Tunicata</i>	<i>Oikopleura dioica</i> : 313224695
<i>Hemichordata</i>	<i>Saccoglossus kowalevskii</i> : 585693286
<i>Echinodermata</i>	<i>Strongylocentrotus purpuratus</i> : 390370566
<i>Protostomia</i>	<i>Drosophila melanogaster</i> : 11037277
<i>Platyhelminthes</i>	<i>Schistosoma mansoni</i> : 353230631
<i>Cnidaria</i>	<i>Hydra vulgaris</i> : 449682695
<i>Placozoa</i>	<i>Trichoplax adhaerens</i> : 195999244
<i>Porifera</i>	Não encontrada
<i>Fungi</i>	<i>Mortierella verticillata</i> NRRL 6337: 672824744
<i>Choanoflagellida</i>	<i>Salpingoeca rosetta</i> : 514696680
<i>Viridiplantae</i>	<i>Vitis vinifera</i> : 731427959
<i>Euglenozoa</i>	<i>Trypanosoma cruzi</i> : 407850076
<i>Amoebozoa</i>	<i>Acanthamoeba castellanii</i> str. Neff: 470532208
<i>Alveolata</i>	Não encontrada
<i>Bacteria</i>	Não encontrada
<i>Archaea</i>	Não encontrada





Figura 14: Alinhamento do triplete de BRCTs de TOPBP1 com a predição de estrutura secundária. (A) BRCT1 ampliado mostrando as predições de estrutura secundária. (B) BRCT2 ampliado mostrando as predições de estrutura secundária. (C) BRCT3 ampliado mostrando as predições de estrutura secundária. Seta verde: fita  $\beta$ . Oval vermelho:  $\alpha$  hélice.

A partir do alinhamento e da predição de estrutura secundária, os BRCTs foram extraídos dos triplete e alinhados individualizados para fazer um dendrograma para comparação somente entre os BRCTs deste triplete (figura 15). Em relação ao dendrograma, foi possível separar os BRCTs em três ramos, sendo o agrupamento coerente, apesar da ausência de *bootstrap*, para os ramos dos BRCTs 2 e 3. A diferença significativa do BRCT 1 para os BRCTs 2 e 3, levou a separação deste grupo no único ramo com suporte de *bootstrap* (figura 15). Os três BRCTs de *Euglenozoa* e o primeiro BRCT de *Choanoflagellida* não agruparam como esperado em relação ao alinhamento múltiplo, além dos BRCT3 de *Viridiplantae*, *Amoebozoa* e *Choanoflagellida* não terem ficado exatamente no ramo dos BRCT3, mas nas proximidades.

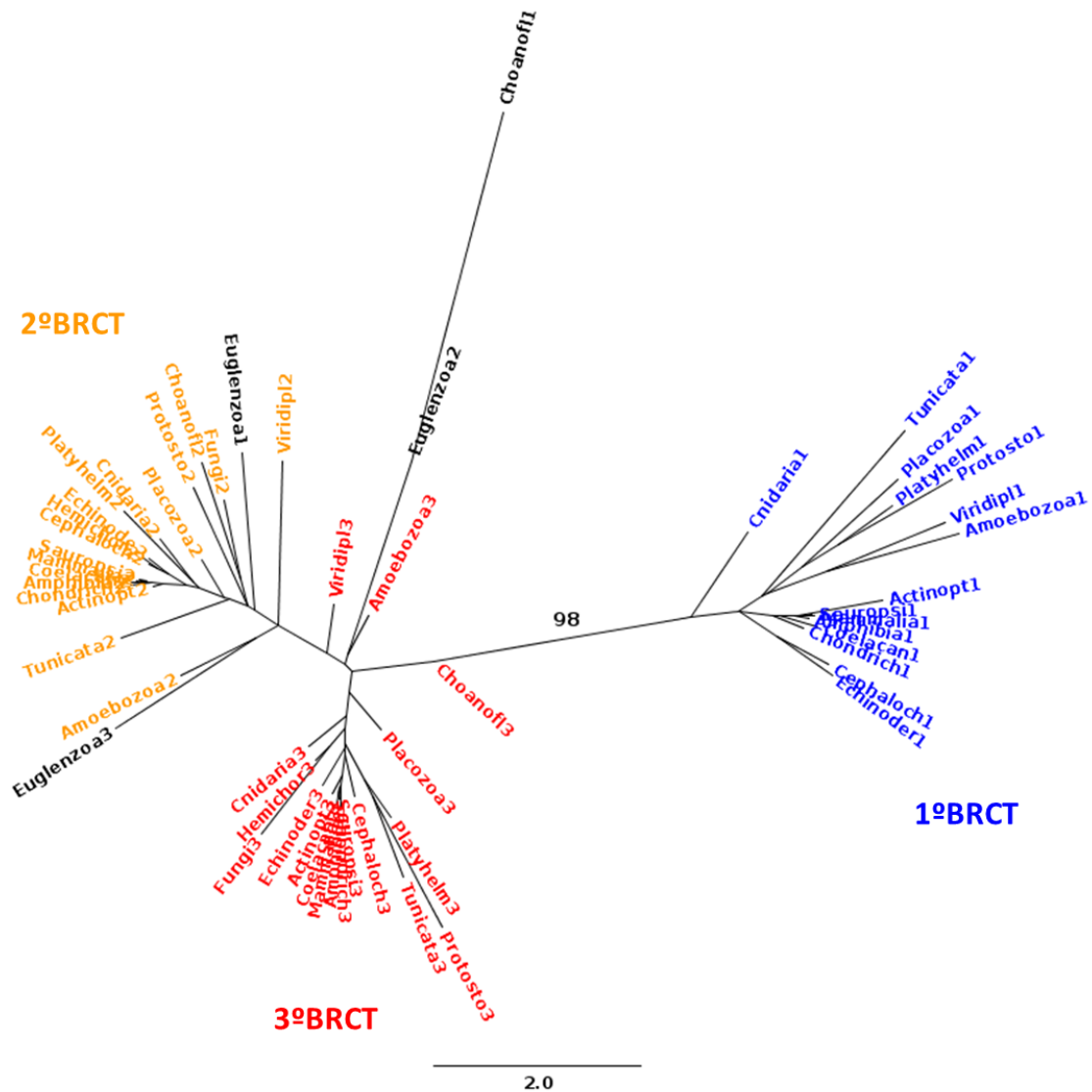


Figura 15: Dendrograma obtido a partir do alinhamento entre os BRCTs individuais do triplete das TOPBP1.

Destacado em azul estão os BRCT1, em laranja BRCT2 e em vermelho BRCT3. Em preto estão marcados os BRCTs que não agruparam no ramo da árvore sugerido pelo alinhamento da região de triplete (Figura 13). 98 indica o valor de *bootstrap* que separa o ramo do BRCT1.

A mesma metodologia foi empregada para ECT2, sendo utilizada a estrutura secundária referente ao cristal 4N40 (ZOU et al., 2014) para ECT2 humana. A proteína ECT2 não foi identificada em oito pontos taxonômicos, sendo eles *Bacteria*, *Archaea*, *Euglenozoa*, *Alveolata*, *Amoebozoa*, *viridiplantae*, *Fungi* e *Porifera* (Tabela 7). O alinhamento múltiplo foi feito com as ECT2 listadas na tabela 7 e o resultado está apresentado na figura 16.

Tabela 7: Proteínas utilizadas no alinhamento dos tripletes de ECT2.

<b>Táxons</b>	<b>Espécie - gi</b>
<i>Mammalia</i>	<i>Homo sapiens</i> : 21735572
<i>Sauropsida</i>	<i>Anolis carolinensis</i> : 637275719
<i>Amphibia</i>	<i>Xenopus (Silurana) tropicalis</i> : 163914811
<i>Coelacanthimorpha</i>	<i>Latimeria chalumnae</i> : 556993915
<i>Actinopterygii</i>	<i>Danio rerio</i> : 51468037
<i>Chondrichthyes</i>	<i>Callorhynchus milii</i> : 632935175
<i>Cephalochordata</i>	<i>Branchiostoma floridae</i> : 260817400
<i>Tunicata</i>	<i>Oikopleura dioica</i> : 313224992
<i>Hemichordata</i>	<i>Saccoglossus kowalevskii</i> : 585711260
<i>Echinodermata</i>	<i>Strongylocentrotus purpuratus</i> : 390358154
<i>Protostomia</i>	<i>Drosophila melanogaster</i> : 24660486
<i>Platyhelminthes</i>	<i>Echinococcus granulosus</i> : 674566404
<i>Cnidaria</i>	<i>Hydra vulgaris</i> : 449686856
<i>Placozoa</i>	<i>Trichoplax adhaerens</i> : 195996387
<i>Porifera</i>	Não encontrada
<i>Fungi</i>	Não encontrada
<i>Choanoflagellida</i>	<i>Salpingoeca rosetta</i> : 514699723
<i>Viridiplantae</i>	Não encontrada
<i>Euglenozoa</i>	Não encontrada
<i>Amoebozoa</i>	Não encontrada
<i>Alveolata</i>	Não encontrada
<i>Bacteria</i>	Não encontrada
<i>Archaea</i>	Não encontrada

O triplete de BRCT não foi observado em *Cnidaria*, *Platyhelminthes* e *Placozoa* (figura 16). O BRCT1 de ECT2 estava ausente em *Placozoa* e *Platyhelminthes*, e em *Cnidaria* encontrava-se truncado, existindo apenas a última fita  $\beta$  e  $\alpha$  hélice. Em *Coelacanthimorpha* este domínio apresentou um padrão diferente dos outros táxons, que pode ser observado pelo *gap* formado nas outras proteínas (figura 17A). Em *Platyhelminthes* o BRCT3 estava truncado, estando ausente a última fita  $\beta$  e  $\alpha$  hélice. Além disso, BRCT3 não foi encontrado em *Placozoa* (figura 17C).

Como observado em TOPBP1, o domínio BRCT1 de ECT2 apresenta maior variação na distribuição da estrutura secundária quando comparado com os BRCTs 2 e 3 (figura 17). Essa diferença foi suficiente para separar o BRCT1 no único ramo do dendograma com suporte de *bootstrap* (figura 18). Os BRCT2 e 3 foram agrupados em ramos distintos, porém sem suporte de *bootstrap*.

No dendograma os BRCTs agruparam como o esperado, com exceção dos BRCT1 de *Cnidaria* (truncado) e *Choanoflagellida*. Além disso, a proteína ECT2 de



*Placozoa* possui somente um domínio BRCT, que no alinhamento foi posicionado junto com outros BRCT2, porém o dendograma sugere que seja o terceiro BRCT (figura 18).



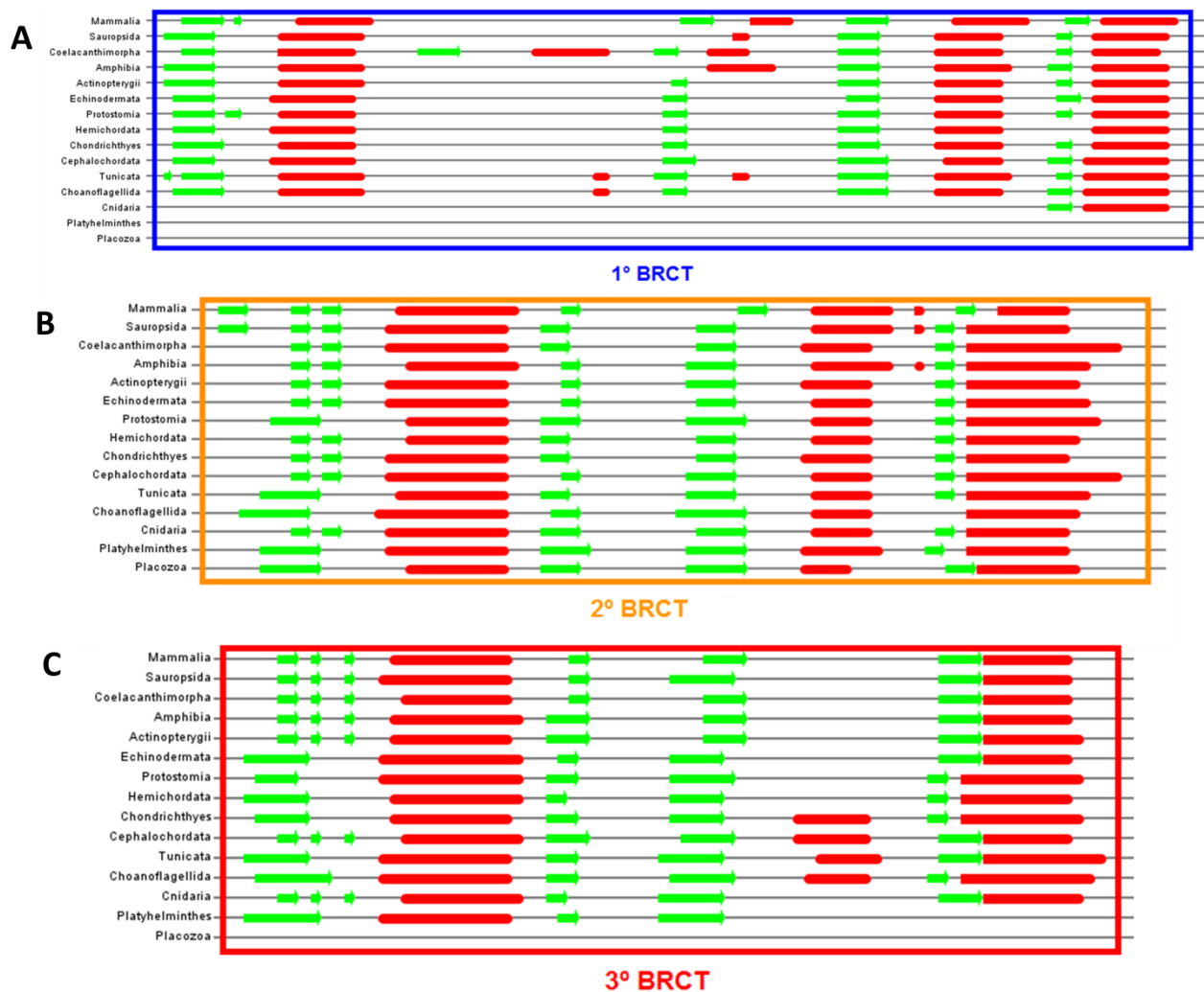


Figura 17: Alinhamento do triplete de BRCTs de ECT2 com a predição de estrutura secundária. (A) BRCT1 ampliado mostrando as predições de estrutura secundária. (B) BRCT2 ampliado mostrando as predições de estrutura secundária. (C) BRCT3 ampliado mostrando as predições de estrutura secundária. Setas verde: fita  $\beta$ . Oval vermelho:  $\alpha$  hélice.

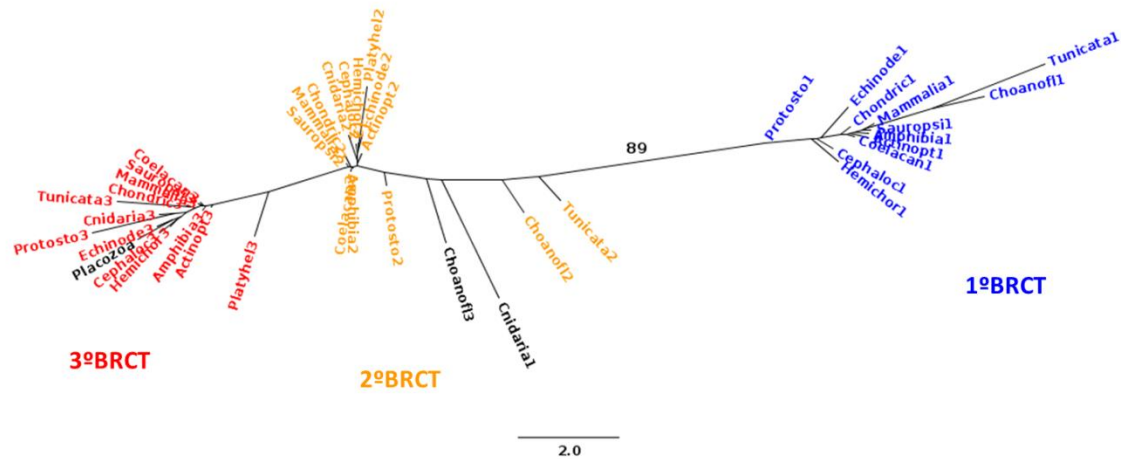


Figura 18: Dendrograma obtido a partir do alinhamento entre os BRCTs individuais das ECT2. Destacado em azul estão os BRCT1, em laranja BRCT2 e em vermelho BRCT3. Em preto estão marcados os BRCTs que não agruparam no ramo da árvore sugerido pelo alinhamento da região de triplete (Figura 16). 89 indica o valor de *bootstrap* que separa o ramo do BRCT1.

Os resultados da identificação das proteínas TOPBP1 e ECT2, e seus triplete de BRCT, foram agrupados em um mapa que resume a presença destas proteínas e do arranjo triplete dos domínios BRCT ao longo da escala evolutiva (figura 19). Através da figura 19 foi possível observar a presença de TOPBP1 e ECT2, e a conservação do arranjo em triplete de BRCT ao longo da escala evolutiva. Como ambas as proteínas não foram encontradas em procariotos, possivelmente se originaram em eucariotos. Em *Porifera* não foi encontrada nem TOPBP1 nem ECT2.

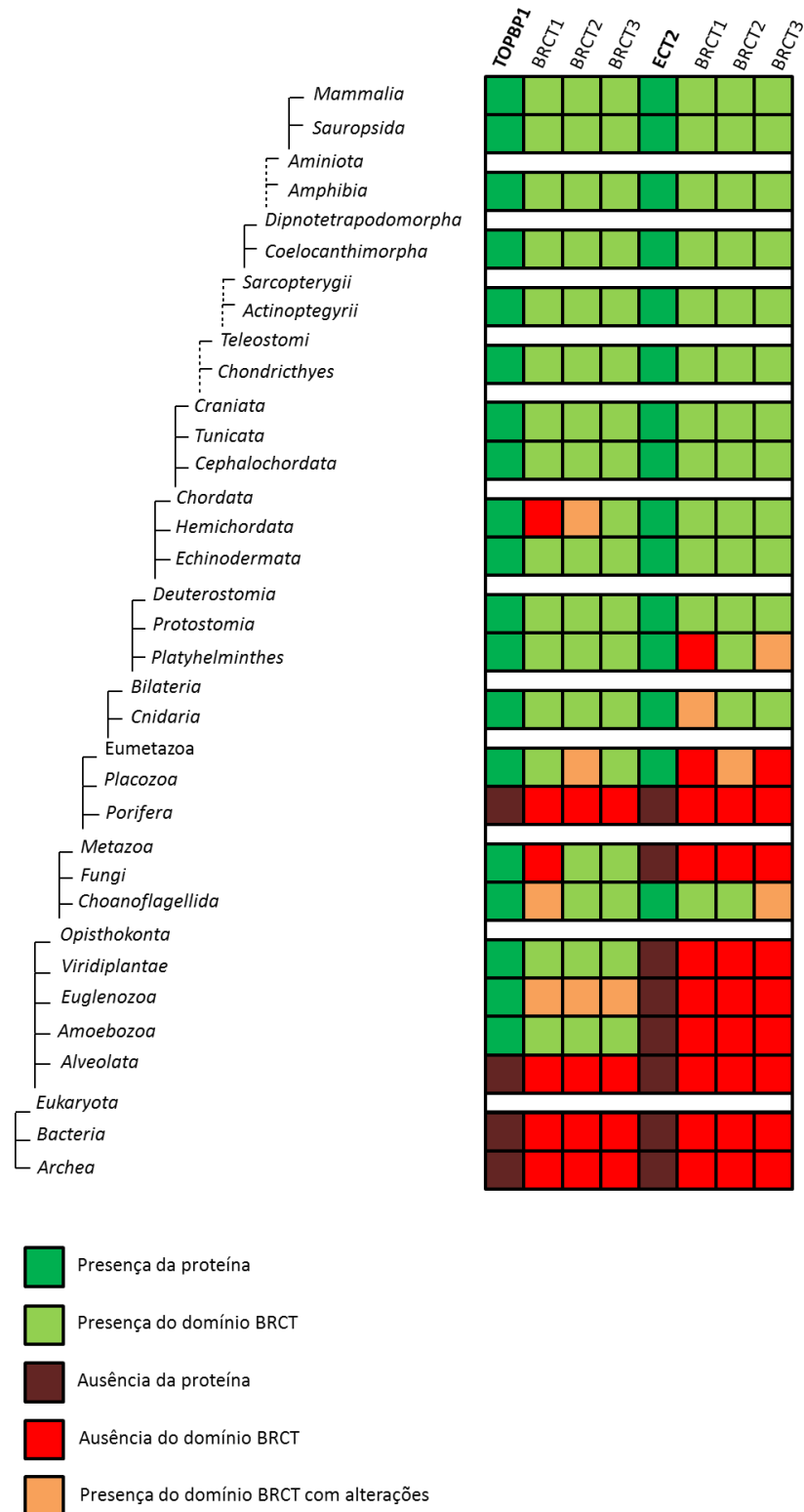


Figura 19: Mapa mostrando a presença das proteínas TOPBP1 e ECT2 ao longo da escala evolutiva.

A figura mostra, também, a presença dos três domínios BRCT que compõem o triplete. As alterações no domínio BRCT são referentes aos BRCTs truncados ou aos que apresentaram diferenças entre a posição no alinhamento e o agrupamento no dendograma. As linhas pontilhadas representam ligação indireta entre os pontos taxonômicos.

## 4.2. Estudo das Vias de Reparo no Mosquito *A. aegypti*

### 4.1.3. Identificação de proteínas de reparo e contendo domínio BRCT em *A. aegypti*

As vias de reparo de DNA consideradas neste trabalho somam em eucariotos e procariotos, aproximadamente, 196 proteínas no banco de dados KEGG. Estas vias são realizadas por 81 e 75 proteínas em humanos e *D. melanogaster*, respectivamente. O mosquito *A. aegypti* possui 56 destas proteínas identificadas no KEGG, as quais tiveram sua anotação confirmada através das análises realizadas (tabela 8). Nas figuras seguintes as proteínas presentes no KEGG estão marcadas em verde, enquanto as não identificadas estão em branco.

Em relação às proteínas contendo o domínio BRCT estão preditas no KEGG para humanos 24 proteínas, para *D. melanogaster* 10 e para *A. aegypti* 8 proteínas (tabela 8).

Tabela 8: Proteínas preditas e ausentes no KEGG para o mosquito *A. aegypti*.

\*Proteínas que participam das vias de reparo e possuem domínio BRCT.

Proteínas das vias de reparo do DNA		Proteínas com domínio BRCT	
Preditas	Ausentes	Preditas	Ausentes
OGG1	NEIL1	CTDP1	ANKRD32
NHT	NEIL2	MCPH1	BARD1
SMUG	NEIL3	Nbs1*	BRCA1
APE1	UNG	PES1	DBF4
XRCC1*	TDG/Mug	REV1	ECT2
HMGB1	MUTY	RFC1	Lig3*
PCNA	MPG	TOPBP1	Lig4*
Pol $\delta$	MBD4	XRCC1*	MDC1
Pol $\epsilon$	APE2		PARP1*
Fen1	Pol $\beta$		PARP4*
APEX	Pol $\lambda$ *		PAXIP1
PARP*	Lig3*		Pol $\lambda$
RBX1	Lig1*		Pol $\mu$
Cul4	DDB2		RAP1
DDB1	CENT2		TDT
XPC	CSA		TP53BP1
HR23B	CSB		
CDK7	XPG		
MNAT1	MSH3		
CCNH	MLH3		
XPB	Ku70		
XPD	Artemis		
TTDA	Pol $\mu$		
TFIIH1	TdT		
TFIIH2	Lig4		
TFIIH3	XRCC4		
TFIIH4	XLF		
XPA	Rad51B		
RPA	Rad52		
XPF			
ERCC1			
PCNA			
RFC*			
PMS2			
MLH1			
MSH6			
MSH2			
ExoI			
SSB			
Ku80			
DNAPKcs			
Rad50			
Mre11			
Nbs1*			
Rad51			
Rad51C			
Rad51D			
XRCC2			
XRCC3			
BRCA2			
DSS1			
Rad54			
BLM			
TOP3			
Mus81			
Eme1			

Na busca por proteínas ausentes nos KEGG, foi possível identificar 7, sendo as proteínas contendo o domínio BRCT: *Epithelial cell transforming 2* (ECT2), *Microcephalin 1* (MCPH1), *Mediator of DNA damage checkpoint protein 1* (MDC1) e *Poly ADP-ribose polymerase* (PARP1), e as proteínas pertencentes a via de reparo por união terminal não homóloga: *Non-homologous end-joining factor 1* (NHEJ1/XLF) e *DNA ligase 4* (Lig4), sendo encontradas duas isoformas de Lig4, as quais também possuem domínio BRCT.

Foram encontrados indícios de expressão para 18 proteínas em transcriptomas disponíveis na literatura, sendo elas: Lig4 (identificada previamente nas análises), FEN1, CDK7, XPD, BLM, BRCA2, DSS1, Eme1, Mre11, Mus81, RAD50, Rad51, Rad54, RPA, TOP3, DNAPKcs, Ku80, Rad27. As figuras que seguem (figuras 20-24) representam as vias de reparo ilustradas pelo KEGG. Em verde estão as proteínas preditas e em branco as ausentes, as setas vermelhas indicam as proteínas que possuem indícios de expressão e as setas azuis apontam para as identificadas nas análises realizadas no presente trabalho.



### Via de Reparo por Excisão de Base (BER)

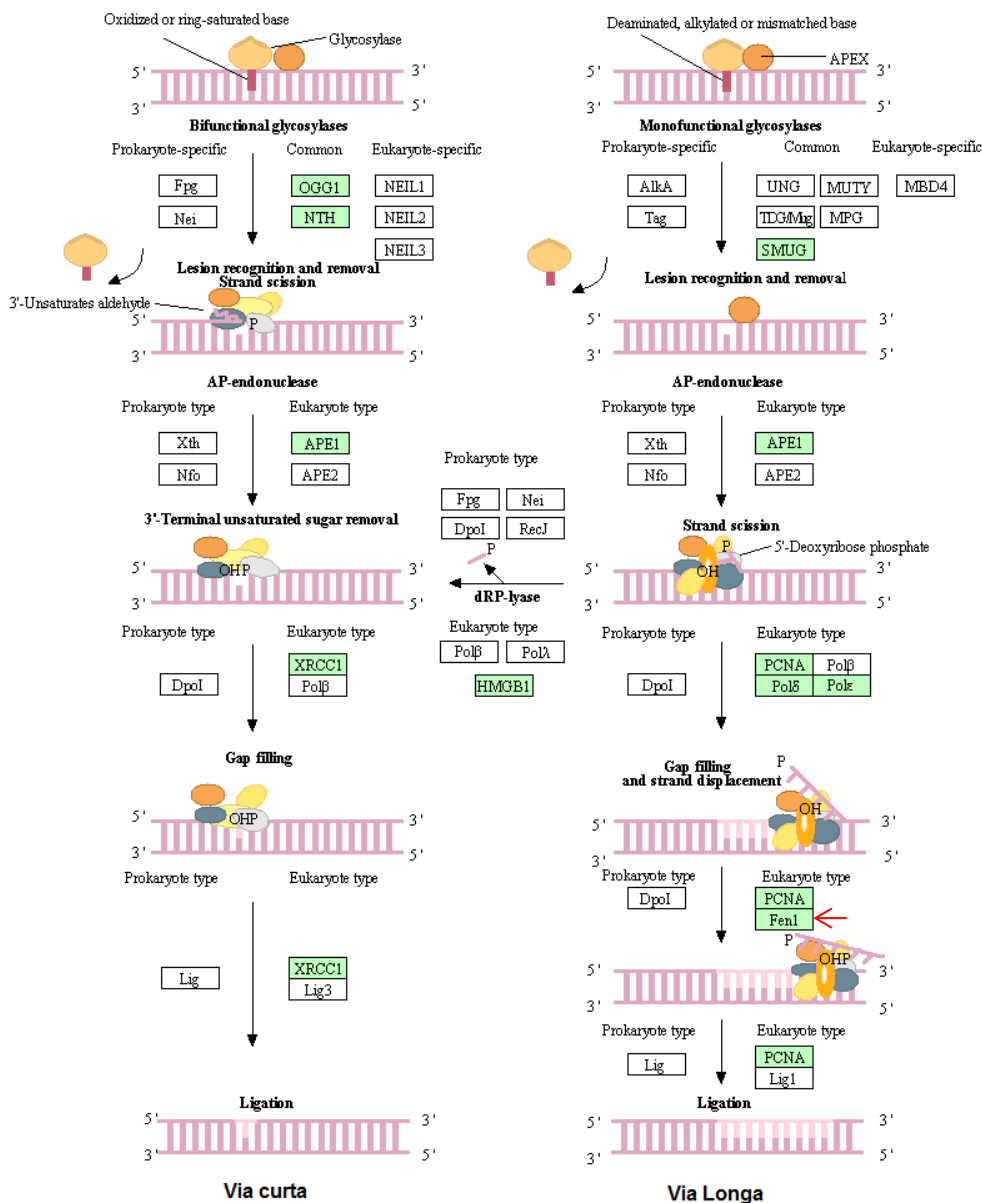


Figura 20: Via de reparo por excisão de base (BER) para *A. aegypti*.

As proteínas preditas aparecem destacadas nos quadrados verdes enquanto as não identificadas estão apresentadas nos quadrados brancos. A seta vermelha indica a proteína cuja expressão foi confirmada pelas análises apresentadas neste trabalho. Imagem gerada pelo banco de dados KEGG (<http://www.kegg.jp/kegg/pathway/ko/ko03410.html>) e adaptada.

## Via de Reparo por Excisão de Nucleotídeo (NER)

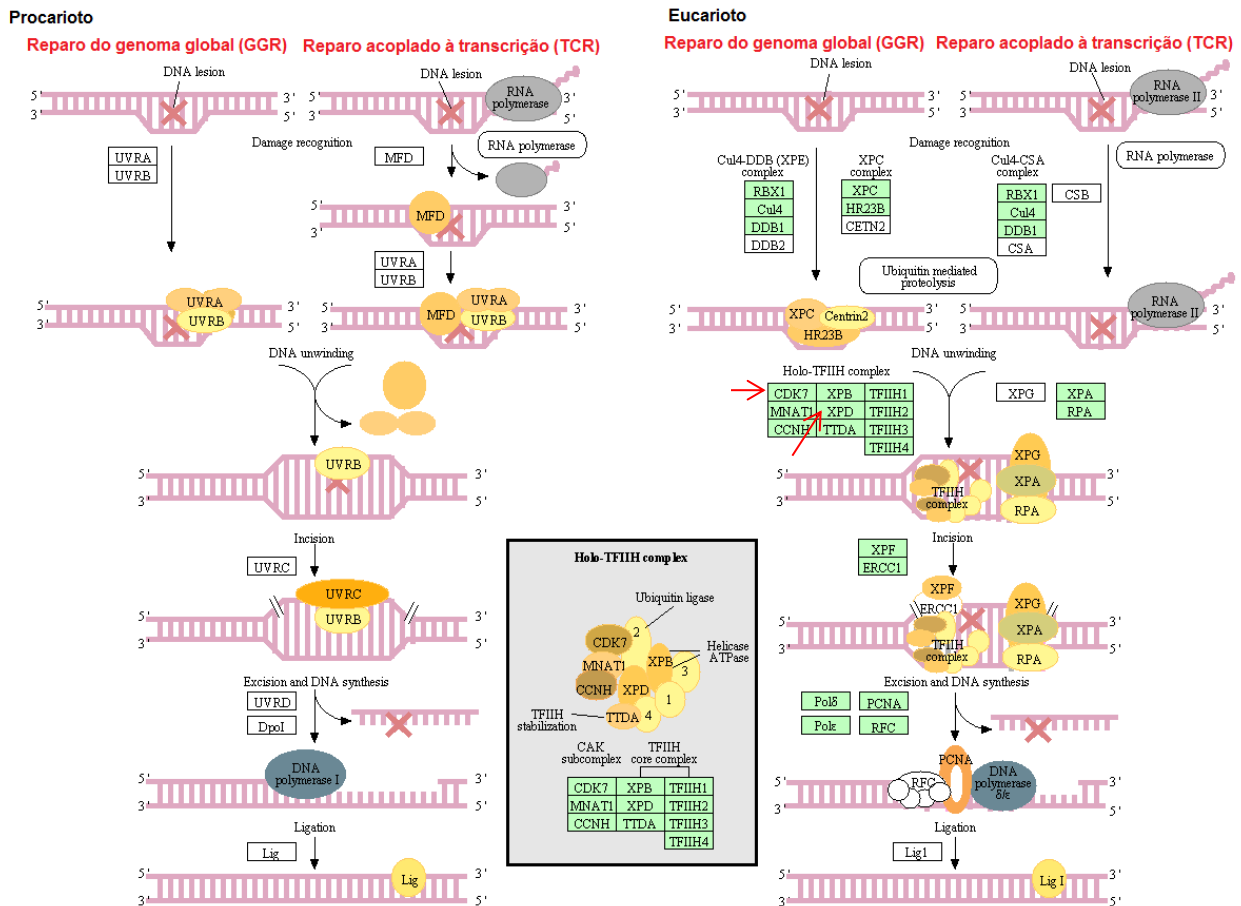


Figura 21: Via de reparo por excisão de nucleotídeo (NER) para *A. aegypti*.

As proteínas preditas aparecem destacadas nos quadrados verdes enquanto as não identificadas estão apresentadas nos quadrados brancos. As setas vermelhas indicam as proteínas cuja expressão foi confirmada pelas análises apresentadas neste trabalho. Imagem gerada pelo banco de dados KEGG (<http://www.kegg.jp/kegg/pathway/ko/ko03420.html>) e adaptada.

## Via de Reparo por Erro de Pareamento (MMR)

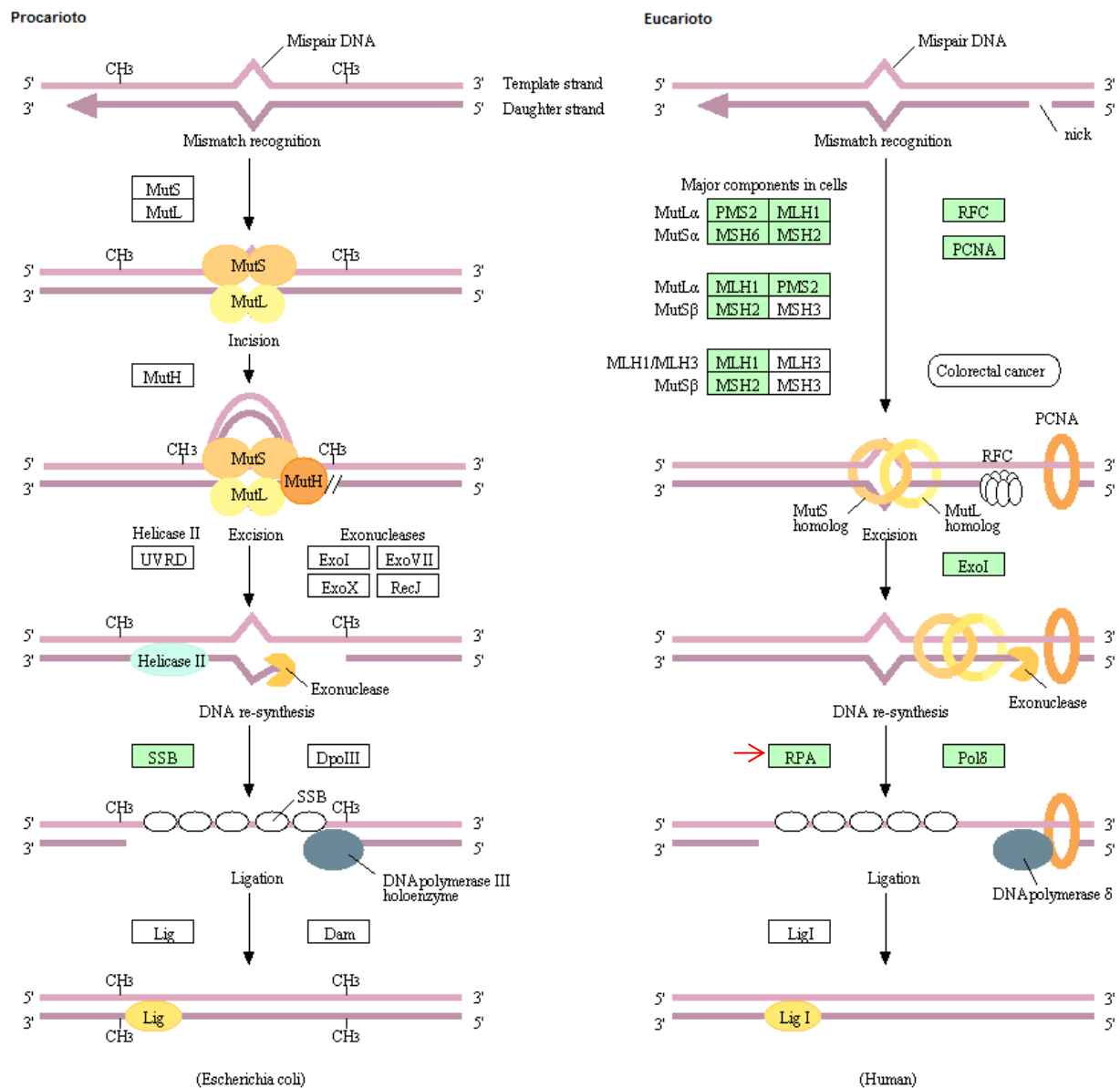


Figura 22: Via de reparo por erro de pareamento (MMR) para *A. aegypti*.

As proteínas preditas aparecem destacadas nos quadrados verdes enquanto as não identificadas estão apresentadas nos quadrados brancos. A seta vermelha indica a proteína cuja expressão foi confirmada pelas análises apresentadas neste trabalho. Imagem gerada pelo banco de dados KEGG (<http://www.kegg.jp/kegg/pathway/ko/ko03430.html>) e adaptada.



## Via de Reparo por União Terminal Não-Homóloga (NHEJ)

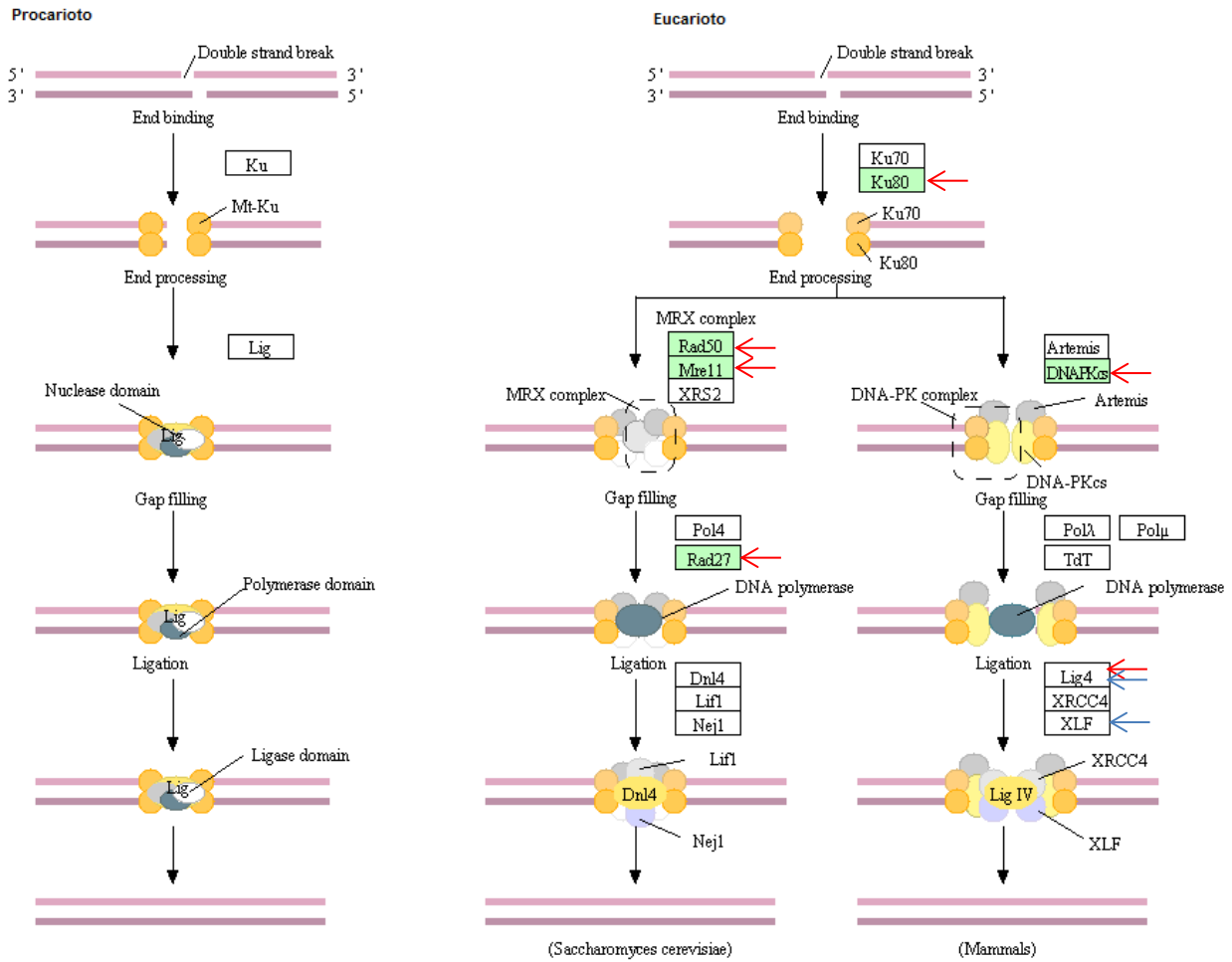


Figura 24: Via de reparo por união terminal não homóloga (NHEJ) para *A. aegypti*.

As proteínas previstas aparecem destacadas nos quadrados verdes enquanto as não identificadas estão apresentadas nos quadrados brancos. As setas vermelhas indicam as proteínas cuja expressão foi confirmada pelas análises apresentadas neste trabalho e as setas azuis as que foram identificadas através das metodologias de bioinformática descritas. Imagem gerada pelo banco de dados KEGG (<http://www.kegg.jp/kegg/pathway/ko/ko03450.html>) e adaptada.

#### 4.1.4. Determinação de condições e estímulo e análises de expressão

Com o objetivo de determinar condições de estímulo para o estudo das vias de reparo do DNA, larvas de *A. aegypti* foram submetidas ao tratamento com paraquat, um herbicida capaz de gerar estresse oxidativo. As larvas foram tratadas com diferentes concentrações de paraquat para determinação de uma curva de sobrevivência em função da concentração e tempo de tratamento.

A curva de sobrevivência (figura 25) demonstrou que o tratamento com paraquat possui resposta dose-tempo dependente, ou seja, a mortalidade aumenta à medida que se aumenta a concentração e o tempo de exposição das larvas.

Os dados obtidos foram utilizados na escolha das condições de estímulo para o estudo das vias de reparo, sendo escolhida a concentração de 0,5 mM nos tempos 12 e 24 horas. Para estudar as vias de reparo do DNA, é necessário induzir danos que ativem as vias sem ativar mecanismos de morte celular. Esta concentração pareceu ser ideal para estudo das vias de reparo, pois não apresentou altas taxas de mortalidade durante o tempo de tratamento, levando a acreditar que o dano é suficiente para ativação das vias, porém insuficiente para gerar apoptose.

Além da curva de sobrevivência foi calculada a LD<sub>50</sub> do paraquat em 24 horas de exposição a esse herbicida. A LD<sub>50</sub> corresponde a concentração necessária para matar 50% das larvas, dessa forma quanto menor a concentração maior a toxicidade. A LD<sub>50</sub>, encontrada para o paraquat durante 24h de tratamento, foi determinada através do gráfico apresentado na figura 26, sendo 1 mM.

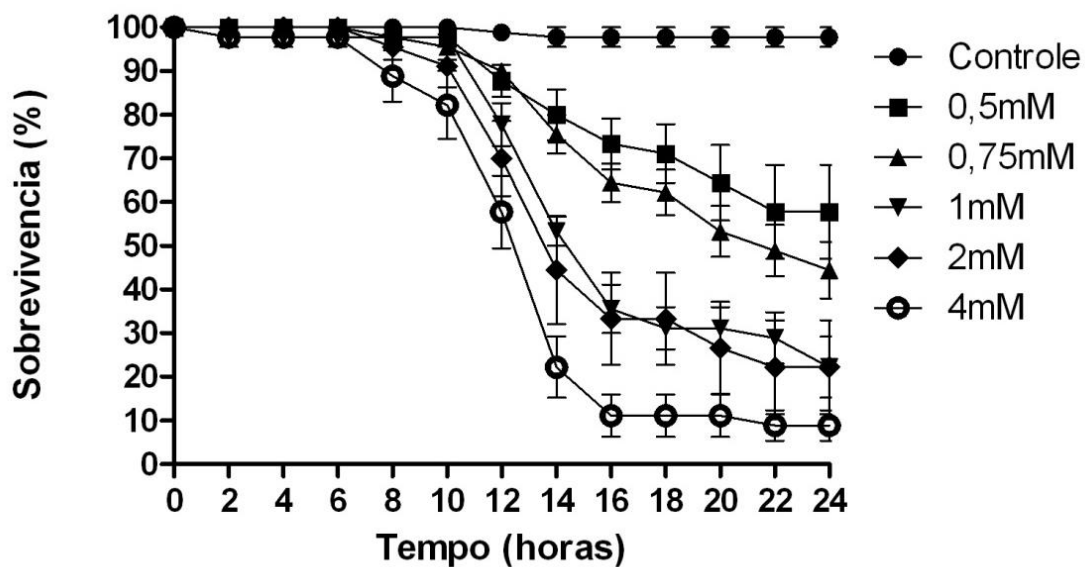


Figura 25: Curva de sobrevivência de larvas tratadas com diferentes concentrações de paraquat durante 24h.

O gráfico é resultado de média e erro padrão de três experimentos independentes em triplicata. Os pontos onde as barras de erro não são visíveis possuem erro igual a zero.

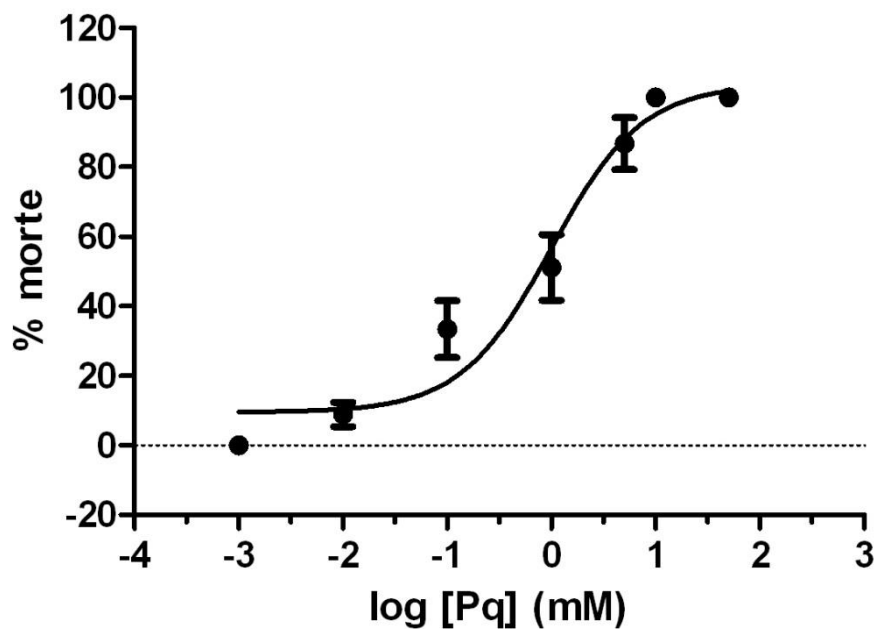


Figura 26: Curva LD50 para o tratamento de larvas com paraquat durante 24h.

A curva é resultado de média e erro padrão de três experimentos independentes em triplicata.  $LD_{50} = 1 \text{ mM}$ . Os pontos onde as barras de erro não são visíveis possuem erro igual a zero.

Utilizando as condições de estímulo iniciais, estabelecidas através da curva de sobrevivência, foram feitos PCRs quantitativos (qPCR) para todos os genes conhecidos, em *A. aegypti*, da via de reparo por união terminal não homóloga (NHEJ), sendo eles: *Ku80*, *Rad50*, *Mre11*, *DNAPKcs*, *Rad27*, *Lig4* e *XLF*. Dessa forma, as larvas foram divididas em três grupos, sendo um grupo controle e dois grupos tratados com 0,5 mM de paraquat durante 12h e 24h. Os resultados obtidos estão apresentados nas figuras 27 a 32. Todos os qPCR foram realizados em triplicata técnica para duas amostras independentes, cada uma delas representando uma triplicata biológica agrupada. Devido ao número de amostras (duas) o cálculo estatístico ainda não foi significativo para nenhum resultado.

O único gene que não foi validado, e não teve a sua expressão confirmada através do PCRq foi *DNAPKcs*.

Os resultados sugerem, como tendência, um aumento na expressão, de todos os genes, nas amostras tratadas com 0,5 mM de paraquat durante 24h. Nas amostras tratadas com 0,5 mM de paraquat por 12h, a diferença na expressão em relação ao controle não foi visível para a maioria dos genes, sugerindo um aumento para *Lig4*, *XLF*, *Mre11* e *Ku80* e uma leve redução para *Rad27*. Os dados também sugerem uma redução na expressão de *Rad50*, em relação ao controle, na amostra tratada durante 12 horas.

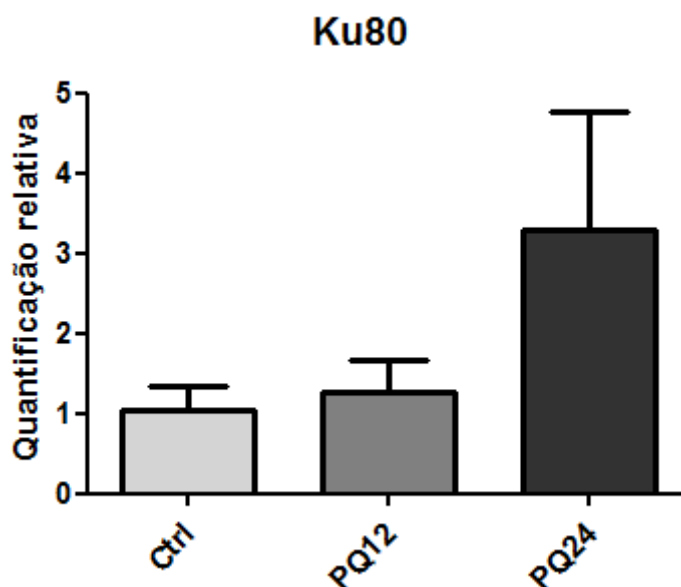


Figura 27: Quantificação da expressão gênica de *Ku80* em larvas de *A. aegypti*. Ctrl: Controle, PQ12: Tratamento com paraquat por 12 horas e PQ24: Tratamento com paraquat por 24 horas.



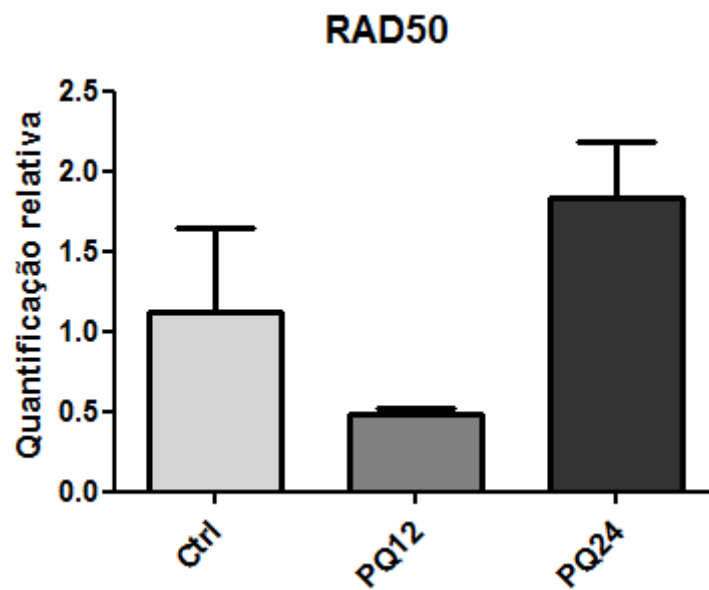


Figura 28: Quantificação da expressão gênica de *RAD50* em larvas de *A. aegypti*. Ctrl: Controle, PQ12: Tratamento com paraquat por 12 horas e PQ24: Tratamento com paraquat por 24 horas.

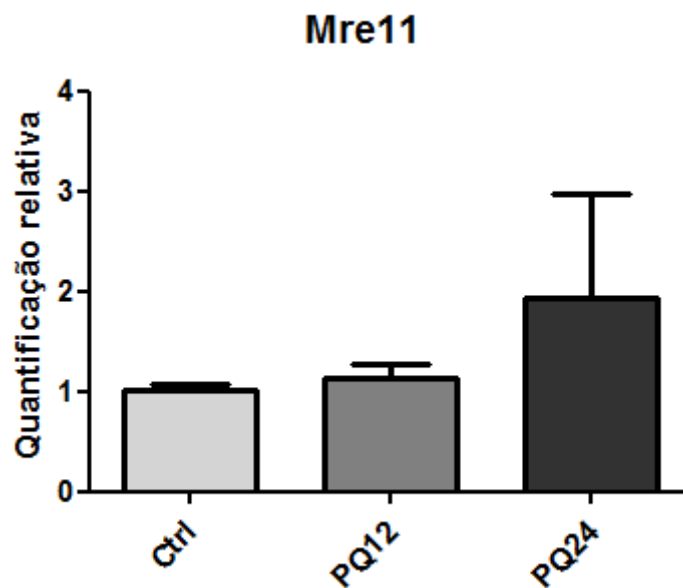


Figura 29: Quantificação da expressão gênica de *Mre11* em larvas de *A. aegypti*. Ctrl: Controle, PQ12: Tratamento com paraquat por 12 horas e PQ24: Tratamento com paraquat por 24 horas.

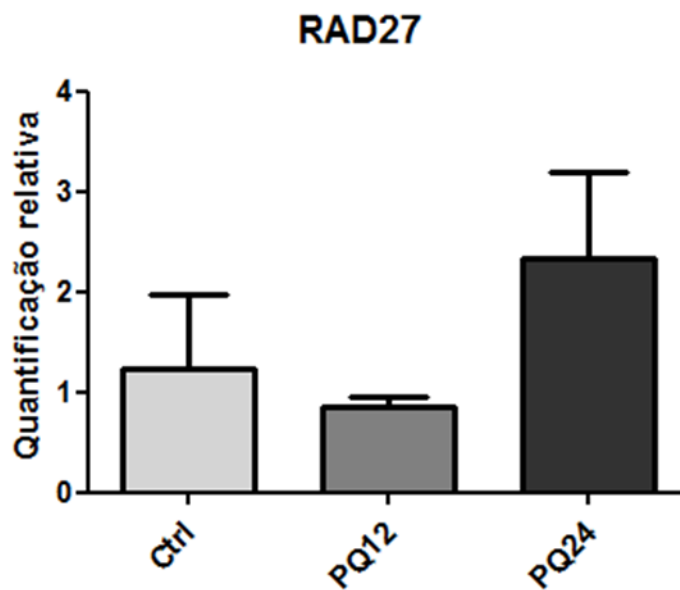


Figura 30: Quantificação da expressão gênica de *RAD27* em larvas de *A. aegypti*.  
Ctrl: Controle, PQ12: Tratamento com paraquat por 12 horas e PQ24: Tratamento com paraquat por 24 horas.

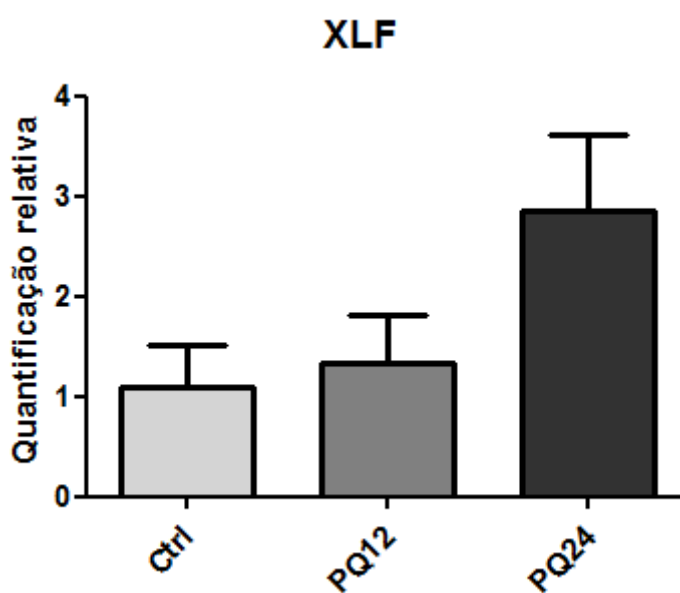


Figura 31: Quantificação da expressão gênica de *XLF* em larvas de *A. aegypti*.  
Ctrl: Controle, PQ12: Tratamento com paraquat por 12 horas e PQ24: Tratamento com paraquat por 24 horas.

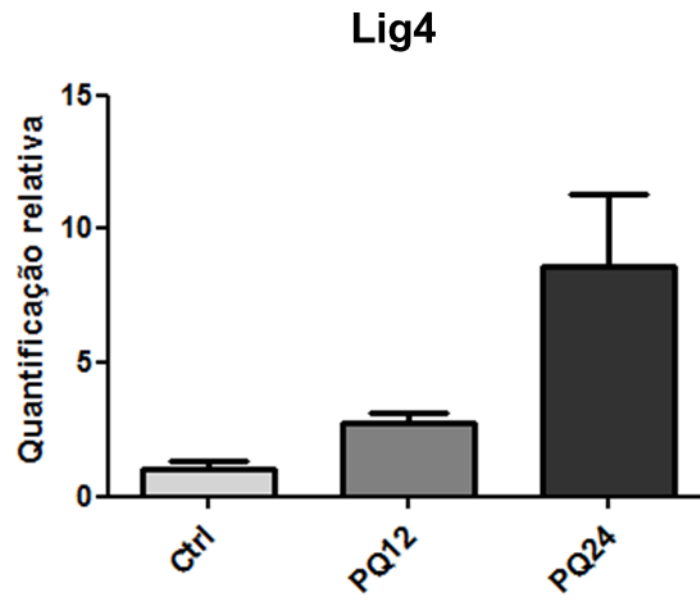


Figura 32: Quantificação da expressão gênica de *Lig4* em larvas de *A. aegypti*.  
Ctrl: Controle, PQ12: Tratamento com paraquat por 12 horas e PQ24: Tratamento com paraquat por 24 horas

## 5. DISCUSSÃO

### 5.1. Identificação de proteínas de reparo do DNA e dos tripletes do domínio BRCT ao longo da evolução.

A estrutura do DNA está constantemente exposta a alterações causadas por fatores genotóxicos endógenos e exógenos. Quando um dano é detectado as vias de reparo do DNA são ativadas objetivando a correção do dano (SANCAR et al., 2004). Proteínas das vias de reparo estão presentes nos três reinos (*Archaea*, *Bacteria* e *Eukaryota*), entretanto uma particularidade encontrada em eucariotos é a integração dos mecanismos de reparo com os pontos de checagem do ciclo celular e apoptose (ARAVIND; WALKER; KOONIN, 1999).

A distribuição evolutiva das vias de reparo do DNA apresentada, na figura 12, sugere significativa conservação destas vias, principalmente em eucariotos, fato observado, também, para um padrão de domínios conservados como os tripletes das proteínas TOPBP1 e ECT2 (figura 19).

Em *Chondrichthyes* foi observada ausência da *DNA Ligase I* (Lig1), proteína que participa da etapa de ligação das vias de reparo por excisão de base (BER), excisão de nucleotídeo (NER) e por erro de pareamento (MMR). Como esta proteína se encontra presente em todos os pontos taxonômicos estudados neste trabalho (figura 12), além de sua função ser de extrema importância na manutenção da integridade do DNA, seria bastante improvável que tenha sido perdida neste ponto da escala evolutiva. Dessa forma, foi feita outra busca utilizando BLASTp contra o banco de dados tsa, e como *query* a Lig1 de *Actinopterygii*. Através desta metodologia foi possível encontrar a Lig1 neste ponto taxonômico. Este táxon possui apenas uma espécie com genoma sequenciado (*Callorhinchus milii*) (VENKATESH et al., 2014), o que leva a suposição de que a qualidade do genoma e a predição gênica não foram suficientes para identificação desta proteína nas primeiras análises.

Em relação a quebras na dupla fita do DNA em *Bacteria*, já foi notificado que, em *E. coli*, o reparo deste tipo de dano ocorre, majoritariamente, através da recombinação homóloga (HR) (EISEN; HANAWALT, 1999). No entanto, existem estudos que mostraram que existe reparo por união terminal não homóloga (NHEJ) em bactérias (PITCHER; BRISSETT; DOHERTY, 2007; SHUMAN; GLICKMAN, 2007; WELLER et al., 2002), apesar de não ter sido encontrada nenhuma proteína referente a

esta via nas nossas análises. Ortólogos funcionais de proteínas desta via foram identificadas em bactérias, como por exemplo, os ortólogos das proteínas do complexo Ku (DELLA et al., 2004; DOHERTY; JACKSON; WELLER, 2001). Análises na sequência destas proteínas, em bactérias, demonstraram que estas são bem menores do que as eucarióticas e que, geralmente, não apresentam os domínios vWA e SAP. Além disso, acredita-se que estas proteínas formem homodímeros e não heterodímeros como em eucariotos. A identificação destes ortólogos em bactérias sugere que Ku é uma proteína antiga cujo surgimento antecede a ramificação de eucariotos e procariotos, porém duplicações de genes e fusão de domínios acarretaram na evolução ao heterodímero Ku70/Ku80 (ARAVIND; KOONIN, 2001; DOHERTY; JACKSON; WELLER, 2001; PITCHER; BRISSETT; DOHERTY, 2007). Além de ortólogos funcionais de Ku, também foi reportada a existência de uma *DNA ligase ATP dependente* (LigD). Entretanto, ainda não foram encontradas em *Archaea* e *Bacteria* ortólogos para as proteínas DNAPKcs e XRCC4 (EISEN; HANAWALT, 1999).

A NHEJ de *Fungi* também é orquestrada por proteínas diferentes em relação à via de humanos. A única etapa realizada por proteínas ortólogas é a detecção, a qual participam Ku70 e Ku80. A etapa de processamento em fungos é realizada pelo complexo MRX que compreende as proteínas Rad50, Mre11 e XRS2, mesmo complexo que participa da detecção no reparo por recombinação homóloga. Finalizando a via, têm-se as proteínas Dnl4, Lif1 e Nej1, ortólogos funcionais de Lig4, XRCC4 e XLF, respectivamente, realizando a etapa de ligação (CRITCHLOW; JACKSON, 1998; DESHPANDE; WILSON, 2007; WU; TOPPER; WILSON, 2008). Aparentemente ainda não havia sido encontrado em fungos ortólogos da proteína DNAPKcs (SHRIVASTAV; DE HARO; NICKOLOFF, 2008), uma cinase de papel crucial nesta via. Entretanto, nas análises realizadas neste trabalho foi encontrada uma proteína ortóloga a DNAPKcs no fungo da espécie *Rhizophagus irregularis* (gi 552908684).

Os resultados mostraram que a via de reparo por erro de pareamento encontra-se bastante conservada, fato que já havia sido reportado anteriormente (EISEN; HANAWALT, 1999; HARFE; JINKS-ROBERTSON, 2000). A primeira etapa desta via em bactérias é realizada pelas proteínas MutS e MutL enquanto em eucariotos esta etapa é orquestrada por complexos proteicos com atividade MutS e MutL “like” ((EISEN; HANAWALT, 1999; HARFE; JINKS-ROBERTSON, 2000; LIN; NEI; MA, 2007). Análises evolutivas anteriores sugeriram que, ambos *MutS* e *MutL*, foram adquiridos por eucariotos e/ou *Archaea* a partir de transferência horizontal, sendo a

origem desta via nestes pontos taxonômicos atribuídas a um ancestral comum de bactérias (LIN; NEI; MA, 2007). Estes fatos explicam a cor azul escura atribuída a esta etapa em *Bacteria* (Figura 12), pois este taxón possui apenas as proteínas MutL e MutS, que são homólogas dos complexos proteicos que estão presente em eucariotos.

Na via de reparo por excisão de base, a etapa inicial de detecção aparentemente está incompleta na maioria dos táxons. Esta etapa é realizada por glicosidases que reconhecem a base alterada e clivam ligações N-glicosídicas retirando a base da fita formando um sítio abásico (WALLACE; MURPHY; SWEASY, 2012). Existem diversas classes de glicosidases estando à maioria delas presentes nos três reinos (EISEN; HANAWALT, 1999). No entanto, existem glicosidases que são específicas de procariotos e eucariotos, como por exemplo, as glicosidases que realizam excisão de bases oxidadas (Fpg/Nei) são específicas de procariotos, entretanto vertebrados possuem as glicosidases NEIL1, NEIL2 e NEIL3 que são homólogas a Fpg/Nei (DAS et al., 2006; MORLAND et al., 2002).

A via de reparo por excisão de nucleotídeo apresenta mecanismos bioquímicos semelhantes entre procariotos e eucariotos, porém as proteínas que atuam nesta via em eucariotos apresentam grandes diferenças na sequência de aminoácidos em relação às proteínas de procariotos (EISEN; HANAWALT, 1999; MORITA et al., 2010). Em procariotos está via não é orquestrada pela mesma maquinaria utilizada por eucariotos. A ligação é a única etapa em que a mesma proteína (Lig1) participa em ambas as vias. Apesar das diferenças em relação à maquinaria utilizada por cada via foram encontradas, em procariotos, as proteínas PCNA e RFC que participam da etapa de síntese em conjunto com as polimerases Pol  $\delta$  e Pol  $\epsilon$ , sendo em procariotos esta realizada pelas proteínas UVRD e DpoI. Em relação à etapa de processamento das extremidades, já foram encontradas em procariotos as proteínas XPB e XPD, helicases que compõem o complexo TFIIH (essencial para o processo de transcrição e para o reparo por excisão de nucleotídeos) (LIN; CHOI; GRALLA, 2005; LIU et al., 2008; SPIES, 2014). Não foram encontrados outros componentes do complexo TFIIH em *Bacteria* e *Archaea*, trabalhos recentes sugerem que em *Bacteria* e *Archaea* estas helicases estejam formando complexos com outras proteínas (BALASINGHAM et al., 2012; BISWAS et al., 2009).

Em procariotos a via de reparo por recombinação homóloga também é orquestrada por análogos das proteínas utilizadas pela mesma via em eucariotos, sendo marcada em azul escuro (*Archaea* está marcada em azul claro por não apresentar todas

as proteínas procarióticas preditas no KEGG) (figura 12). Um exemplo é a proteína RecA presente em bactérias que realiza a função da proteína Rad51. Foi proposto que os genes da família *recA/Rad51* se originaram a partir de um único ancestral comum, através de duplicação gênica, perda de genes e endossimbiose. A duplicação gênica de um gene *recA* ancestral, antes da divisão em *Archaea* e eucariotos, originou duas linhagens de genes *Rad51* “like” *RAD $\alpha$*  e *RAD $\beta$* , sendo que em *Bacteria* foi mantida uma cópia de *recA* “like”. Em *Archaea* é possível observar a presença de dois genes *RADA* e *RADB*. Em eucariotos duplicações em *RAD $\alpha$*  originaram os genes *Rad51* e *DMC1* enquanto duplicações em *RAD $\beta$*  levaram a origem dos parálogos de *Rad 51* (*Rad51B*, *Rad51C*, *Rad51D*, *XRCC2*, *XRCC3*) (CHINTAPALLI et al., 2013; LIN et al., 2006).

Recentemente foi mostrado que a organização dos domínios BRCT em tandem sofreu uma expansão evolutiva de procariotos para metazoários (MESQUITA et al., 2010), demonstrando que à medida que se aumenta a complexidade de um organismo maior é a necessidade de possuir mecanismos para manter o DNA intacto (MESQUITA et al., 2010; SHENG; ZHAO; HUANG, 2011).

O número de proteínas contendo domínio BRCT por genoma está correlacionado com a complexidade do organismo, sendo encontrada apenas uma proteína em *E. coli*, 12 em *S. cerevisiae*, 24 em *H. sapiens* e 28 em *O. sativa* (SHENG; ZHAO; HUANG, 2011; WOODS et al., 2012). Em procariotos não foram encontradas proteínas contendo BRCT em tandem, sendo identificado individualizado em uma *DNA ligase* homóloga a *DNA ligase 4* (Lig4) (MESQUITA et al., 2010).

Neste trabalho foi feita a busca, ao longo da escala evolutiva, das proteínas que apresentam BRCT em triplete, já caracterizado em humanos nas proteínas TOPBP1 e ECT2, em seguida alinhamentos foram feitos para identificar a conservação destes tripletes.

A proteína TOPBP1, em humanos, coordena a ativação dos pontos de checagem em resposta ao dano através do acoplamento ao complexo 9-1-1 (formado pelas proteínas Rad9, Hus1 e Rad1), na região de quebra na fita simples, ativando o complexo ATR-ATRIP (RAPPAS; OLIVER; PEARL, 2011).

O primeiro domínio BRCT de TOPBP1 foi identificado, pela primeira vez, em um estudo cristalográfico da região N-terminal desta proteína em humanos (RAPPAS; OLIVER; PEARL, 2011). Estruturalmente, é levemente diferente dos outros BRCTs, apresentando um maior intervalo entre suas estruturas secundárias, os BRCTs 2 e 3

apresentam sítios de ligação a fosfopeptídeos, enquanto o BRCT1 não possui. Através de experimentos de mutação e interação foi determinado que o BRCT2 constitui o principal sítio para ligação do resíduo pSer387 de Rad9, enquanto o BRCT 3 apresenta baixa afinidade a esse resíduo (RAPPAS; OLIVER; PEARL, 2011).

A proteína TOPBP1 foi identificada na maioria dos pontos taxonômicos estudados neste trabalho, não sendo encontrada apenas em *Bacteria*, *Archaea*, *Alveolata* e *Porifera*, o que condiz com os resultados publicados em 2010 por Mesquita e colaboradores, que reportou a ausência desta proteína em procariotos (MESQUITA et al., 2010).

Em poríferos foi encontrada apenas uma espécie com genoma sequenciado (*Amphimedon queenslandica*) (SRIVASTAVA et al., 2010), assim a ausência de TOPBP1 neste ponto taxonômico pode ser consequência de problemas de cobertura no genoma sequenciado ou da predição gênica. Apesar da identificação das proteínas das outras vias de reparo (Figura 12) sugerir boa qualidade dos dados, as análises realizadas neste trabalho encontraram apenas seis proteínas contendo domínio BRCT (XRCC1, MCPH1, REV1, PESC, PARP1, Lig4) em *Porifera*, quando o esperado era aproximadamente o mesmo número encontrado para *Placozoa*, que contém 17 proteínas com domínio BRCT (TOPBP1, PAXIP1, XRCC1, BRCA1, RFC1, MDC1, NBN, PARP1, PES1, ECT2, Lig4, TONSL, TP53B, REV1, MCPH1, PARP4, Pol  $\mu$ ).

O dendograma (figura 15) separou os BRCTs em três ramos, de forma coerente com o alinhamento múltiplo (figura 13). Apenas os três domínios BRCTs de *Euglenozoa* e o primeiro de *Choanoflagellida* não agruparam da forma que era esperado, sugerindo que os domínios BRCTs em questão apresentam diferenças mais significativas em relação aos BRCTs dos outros táxons, apesar da ausência de *bootstrap* na sua localização deixar margem para dúvida. Como somente o ramo dos BRCTs1 possuem suporte de *bootstrap* poderíamos sugerir que nenhum destes domínios é um BRCT1, mas não temos como afirmar se são os BRCTs 2 ou 3. Em relação à *Euglenozoa*, como todos os BRCTs não agruparam corretamente no dendograma pode-se pressupor que, apesar de terem sido usados os três primeiros BRCTs encontrados, estes não correspondem aos três primeiros BRCTs de TOPBP1, sugerindo que TOPBP1 esteja truncada.

O alinhamento dos tripletes de TOPBP1 (figura 13) mostrou conservação das estruturas secundárias, sugerindo conservação destes domínios e consequentemente sugerindo a preservação da função de reconhecimento de peptídeos multi-fosforilados



presente no triplete humano (RAPPAS; OLIVER; PEARL, 2011). Entretanto, para as proteínas que possuem o triplete truncado não é possível afirmar que esta função se encontra conservada, principalmente para a TOPBP1 de *Placozoa* que, apesar de possuir o triplete, tem o BRCT2 incompleto, sendo este o domínio que apresenta maior afinidade ao resíduo pSer387 de Rad9. Quanto ao primeiro BRCT, apesar de não ter sido determinado seu papel no reconhecimento de fosfopeptídeos multifosforilados, pode-se especular que sua ausência, como ocorre em *Fungi* e *Hemichordata*, acarretaria em perda da função, uma vez que não há descrito a função de reconhecimento de peptídeos multi-fosforilados por tandens de BRCT.

Através dos resultados encontrados é possível especular sobre a origem da organização em triplete dos domínios BRCT de TOPBP1, que possivelmente, se originou em eucariotos unicelulares, uma vez que é possível encontrar esta arquitetura em *Amoebozoa* e, mesmo com diferenças, em *Euglenozoa* e *Choanoflagellida*.

Em humanos, a proteína ECT2 desempenha um importante papel na citocinese devido a sua capacidade de auto inibição e regulação positiva por CYK-4. A interação com o resíduo pS164 de CYK-4 ocorre através de um sítio conservado de ligação a fosfo-serina presente no segundo BRCT (ZOU et al., 2014).

Esta proteína não foi identificada em oito pontos taxonômicos, sendo eles: *Bacteria*, *Archaea*, *Euglenozoa*, *Alveolata*, *Amoebozoa*, *viridiplantae*, *Fungi* e *Porifera*. Os resultados condizem com o publicado por Mesquita e colaboradores (MESQUITA et al., 2010). A possível justificativa para ECT2 não ter sido encontrada em poríferos é a mesma dada para a ausência de TOPBP1 neste ponto taxonômico, a presença de somente um genoma sequenciado.

O agrupamento dos BRCTs no dendograma (figura 18) se mostrou coerente com o alinhamento múltiplo (figura 16), sendo os BRCTs que não agruparam no ramo esperado: o BRCT1 de *Cnidaria*, o BRCT3 de *Choanoflagellida* e o único BRCT de *Placozoa*. O BRCT1 de *Cnidaria* encontra-se truncado, existindo apenas a última fita beta e alfa hélice, justificando o fato de não ter agrupado corretamente no dendograma (figura 17A). Em relação ao BRCT3 de *Choanoflagellida*, pode-se supor que este apresente poucas diferenças em relação ao seu BRCT2 gerando o posicionamento próximo destes, entretanto devido à ausência de *bootstrap* não se pode afirmar tal suposição. A proteína ECT2 de *Placozoa* possui apenas um domínio BRCT que alinhou na posição referente ao BRCT2 (figura 17B), enquanto no dendograma agrupou junto

com o BRCT3. Dessa forma não é possível afirmar se o único domínio BRCT encontrado no homólogo de ECT2 em *Placozoa* corresponde ao BRCT2 ou ao BRCT3.

Em *Coelacanthimorpha*, o primeiro domínio BRCT apresenta um padrão diferente dos outros táxons, que inclui uma inserção de 31 aminoácidos após a primeira alfa-hélice. No alinhamento esta inserção gera um espaçamento nas outras proteínas (figura 17A). Esta mesma região foi encontrada em outras isoformas de ECT2 de *Actinopterygii* e *Sauropsida* (resultados não mostrados), o que sugere que pode ter ocorrido erro de predição nestas proteínas, já que seria pouco provável que mutações alterando sítios de *splicing* e gerando aumento da região codificante acontecessem de forma independente, mas produzindo modificações semelhantes em poucas espécies tão distantes evolutivamente.

A conservação das estruturas secundárias e dos aminoácidos observadas no alinhamento de ECT2 (figura 16) sugere que a função de reconhecimento de peptídeos multi-fosforilados (RAPPAS; OLIVER; PEARL, 2011) esteja preservada. Porém, apesar de todas as proteínas ECT2 estudadas neste trabalho apresentarem o segundo domínio conservado (domínio onde ocorre a ligação a fosfopeptídeos), não é possível afirmar que todas possuem esta função inalterada, uma vez que outras regiões do triplete auxiliam a interação, como por exemplo, um núcleo hidrofóbico formado pelos segundo e terceiro BRCTs (ZOU et al., 2014). Dessa forma, a ausência de alguma região do triplete pode comprometer a afinidade a fosfopeptídeos.

Através dos resultados obtidos é possível observar que, ambas as proteínas se originaram em eucariotos (figura 19), sendo que TOPBP1 parece ter se originado em protistas e ECT2 em *Choanoflagellida*. Dessa forma, é possível sugerir que o triplete de BRCT se originou em eucariotos.

## 5.2. Identificação de proteínas de vias de reparo e contendo domínio BRCT em *A. aegypti*

A expectativa para a etapa bioinformática era encontrar ortólogos ainda não identificados, em *A. aegypti*, das proteínas das vias de reparo e proteínas com o domínio BRCT.

Através das análises bioinformáticas foi possível conferir a anotação de todas as proteínas preditas para *A. aegypti* no banco de dados KEGG, realizar a identificação de

proteínas ausentes, e identificar proteínas que possuem indícios de expressão através de transcriptomas presentes na literatura.

Os homólogos que foram identificados neste trabalho são: ECT2, MCPH1, MDC1, PARP1, XLF e duas possíveis isoformas de Lig4.

As proteínas XLF e Lig4 pertencem à via de reparo por união terminal não homóloga (NHEJ), responsável por reparar quebras na dupla fita do DNA. Ambas atuam na etapa final desta via, que corresponde à ligação das extremidades.

A possível duplicação de Lig4 foi reportada por Overcash e colaboradores (OVERCASH et al., 2015), mas carece de confirmação, uma vez que pode se tratar de erro de predição gênica.

De uma maneira geral, o pequeno número de novas proteínas encontradas pode ser consequência de limitações da predição gênica e/ou do próprio genoma.

Com relação às proteínas as quais foram encontrados indícios de expressão (condições fisiológicas: bloqueio da via IMD ou Toll, infecção bacteriana, e estresse oxidativo) (Fen1, CDK7, XPD, BLM, BRCA2, DSS1, Eme1, Mre11, Mus81, Rad50, Rad51, Rad54, RPA, TOP3, DNAFKas, Ku80, Rad27 e Lig4), a maioria está relacionada as vias de reparo de quebras na dupla fita, sugerindo que estas provavelmente são as mais utilizadas por este inseto. A alimentação hematofágica, realizada pelas fêmeas, representa um grande desafio oxidativo devido à digestão da hemoglobina, que leva liberação de heme e  $Fe^{2+}$ , que são capazes de induzir reações de formação de espécies reativas de oxigênio (EROs) (DANSA-PETRETSKI et al., 1995; GRAÇA-SOUZA et al., 2006). Quando em excesso EROs pode causar danos oxidativos ao DNA além de gerar quebras na dupla fita (BARZILAI; YAMAMOTO, 2004; IYAMA; WILSON, 2013). A hematofagia e sua consequente geração de EROs pode explicar o fato da maioria das proteínas, que foram encontrados indícios de expressão, serem pertencentes as vias de reparo que reparam quebras na dupla fita.

O passo seguinte deste estudo foi investigar a ação do estresse oxidativo sob as vias de reparo, sendo analisados, por PCR quantitativo, os genes referentes às proteínas preditas na via de reparo por união terminal não homóloga (NHEJ), incluindo os encontrados nas análises bioinformáticas (Lig4 e XLF).

O agente oxidativo utilizado para o estudo foi o herbicida paraquat (1,1'-dimethyl-4,4'-bipyridinium), o qual atua induzindo o estresse oxidativo através do aumento na produção de radicais livres, associado à depleção dos mecanismos antioxidantes. O paraquat é reduzido durante o ciclo redox formando o radical  $PQ^+$  que

é rapidamente re-oxidado formando  $PQ^{2+}$ , gerando ânion superóxido ( $O_2^-$ ). A redução subsequente do ânion superóxido leva a geração de peróxido de hidrogênio ( $H_2O_2$ ) e radicais hidroxila ( $\cdot OH$ ) (BLACK et al., 2008; FUKUSHIMA et al., 2002; GAWARAMMANA; BUCKLEY, 2011).

A curva de sobrevivência de larvas tratadas com paraquat apresentou resposta dose-tempo dependente, onde a mortalidade aumenta à medida que se aumenta a concentração e o tempo de tratamento (figura 25).

Através da curva de sobrevivência foram escolhidas condições para estudo das vias de reparo do DNA, sendo escolhida a concentração de 0,5 mM nos tempos 12h e 24h. Para estudar estas vias é necessário gerar danos suficientes para induzir o reparo, porém sem ativar os mecanismos de apoptose. Dessa forma, a concentração de 0,5 mM parece ser a mais interessante para estudo devido às taxas de mortalidade observada na curva, (aproximadamente 40% no tempo de tratamento de 24 h).

Além da curva de sobrevivência a  $LD_{50}$  do paraquat foi mensurada para 24h de tratamento sendo encontrado um valor de 1 mM (figura 26).

O efeito do paraquat já foi determinado em células de mamíferos e também de inseto, sendo a sensibilidade variável com a concentração e o tempo de exposição. Como por exemplo, o tratamento de cultura primária de fibroblastos de pele de ratos com paraquat apresentou  $LD_{50}$  de 3 a 4 mM (SALMON et al., 2005). Culturas de células de *Aedes albopictus* tratadas com paraquat apresentaram  $LD_{50}$  de 1 a 2  $\mu M$  (FALLON; KURTZ; CARROLL, 2013). Estudos utilizando *D. melanogaster* demonstraram que uma concentração de 20 mM de paraquat acarreta na morte de 50% destes insetos quando tratados com essa concentração por 55h (MINOIS et al., 2012). As diferenças na  $LD_{50}$  ocorrem, possivelmente, devido a diferenças não só no organismo testado como também no estágio de desenvolvimento, ou mesmo no tecido ou cultura de célula, não sendo possível uma comparação dos dados de  $LD_{50}$  obtidos com os presentes na literatura para esta molécula.

A partir dos resultados da etapa bioinformática e da curva de sobrevivência, foram reunidas informações para desenho de iniciadores e condições de estímulo oxidante, que permitiram realizar PCRs quantitativos (qPCR) para avaliar a expressão dos genes de NHEJ em larvas de *A. aegypti* tratadas com paraquat.

Como dito anteriormente o paraquat causa estresse oxidativo, o que pode gerar danos oxidativos no DNA, principalmente quebras na fita dupla. Dessa forma, desejava-se pesquisar a variação da expressão dos genes participantes das vias de reparo em

quebras na fita dupla. A NHEJ foi à escolha inicial por conter as duas proteínas identificadas por bioinformática (LIG4 e XLF). Não foi possível realizar qPCR para as duas isoformas de Lig4 encontradas neste trabalho em virtude da semelhança entre a sequência de aminoácidos, o que dificultou o desenho de iniciadores específicos. Dessa forma, o par de iniciadores desenhado amplifica ambas as isoformas.

A expressão de todos os genes, com exceção de *DNAPKcs*, foi confirmada através do PCR quantitativo. Como os PCRs foram feitos em duplicatas, em relação aos índices de expressão determinados pode-se apenas sugerir variação de expressão, sendo necessárias mais replicatas para confirmação e cálculo estatístico.

O maior índice de expressão foi observado nas amostras tratadas com 0,5 mM de paraquat durante 24h, resultado que condiz com o que era esperado, uma vez que o paraquat induz estresse oxidativo gerando danos no DNA.

Nas amostras tratadas com 0,5 mM de paraquat por 12h, a diferença na expressão em relação ao controle não foi muito visível para a maioria dos genes, sendo maior para *Lig4*, *XLF*, *Mre11* e *Ku80* e levemente menor para *Rad27*.

Apenas o gene *Rad50* apresentou uma maior diferença em relação ao controle, sendo observada menor expressão na amostra tratada durante 12h (figura 28). Esse gene, junto com *Mre11*, participa do complexo MRN que atua na etapa de detecção da quebra na dupla fita da via de reparo por recombinação homóloga (HR). O balanço entre NHEJ/HR é determinado pelo ciclo celular, enquanto HR ocorre nas fases S e G2 do ciclo, quando o cromossomo homólogo encontra-se disponível para servir como molde, NHEJ pode ocorrer em qualquer fase, sendo predominante em G1. No entanto, em geral, existe competição entre os complexos de HR e NHEJ pelo sítio de quebra (IYAMA; WILSON, 2013). Estudos em células CHO mostraram que NHEJ é dominante sob HR, devido a abundância e alta afinidade de Ku, que ocupa rapidamente as extremidades quebradas iniciando o reparo (MANSOUR et al., 2008). A preferência pelo reparo por NHEJ reportada por Mansour e colaboradores, pode justificar a redução da expressão de *Rad50* nas amostras tratadas com 0,5 mM de paraquat durante 12 horas. Pode-se pressupor que, como NHEJ é dominante, em uma fase inicial do dano esta seja suficiente para o reparo, entretanto, a medida que se aumenta o tempo de tratamento os níveis de expressão aumentam, indicando que ambas as vias estão sendo necessária para reparar o dano. Porém, para confirmação destas hipóteses seria necessário determinar os índices de expressão de outras proteínas da via de HR, como por exemplo, de *Rad51*

que possui papel central nesta via e obter mais replicatas para ter suporte estatístico para as diferenças observadas.

## 6. CONCLUSÃO

Os resultados de distribuição das proteínas das vias de reparo e os alinhamentos dos triplete de TOPBP1 e ECT2 permitem concluir que, as vias de reparo do DNA encontram-se conservadas em eucariotos. Além disso, a presença de proteínas análogas em procariotos evidencia a conservação da função de reparo do DNA.

A distribuição evolutiva das proteínas TOPBP1 e ECT2 demonstrou que ambas, possivelmente, se originaram em eucariotos unicelulares, uma vez que TOPBP1 foi encontrada em protistas e ECT2 em *Choanoflagellida*. Além disso, os alinhamentos múltiplos de TOPBP1 e ECT2 sugerem que a arquitetura em triplete de BRCT, também, teve origem em eucariotos unicelulares.

Em *A. aegypti* foi possível identificar as proteínas Lig4 (duas isoformas) e XLF, ambas pertencentes à via de reparo por união terminal não homóloga. Foram, também, identificadas proteínas contendo domínio BRCT, sendo elas: ECT2, MCPH1, MDC1, PARP1, além das duas isoformas de Lig4 que também apresentam domínio BRCT. Foi possível identificar indícios de expressão de 18 proteínas, sendo a maioria pertencente às vias de reparo de quebras na dupla fita.

A curva de sobrevivência de larvas de *A. aegypti* quando estimuladas por paraquat, apresentou resposta dose-tempo dependente, sendo a LD50 de 0,95 mM.

Os PCRs quantitativos, realizados para os genes da via de reparo por união terminal não homóloga demonstraram que todos, com exceção da *DNAPKcs*, estão sendo expressos, sendo os índices de expressão maiores nas amostras tratadas com 0,5 mM de paraquat durante 24h.

## 7. REFERÊNCIAS

- ALMEIDA, K. H.; SOBOL, R. W. A unified view of base excision repair: Lesion-dependent protein complexes regulated by post-translational modification. **DNA Repair**, v. 6, n. 6, p. 695–711, 2007.
- ALTSCHUL, S. F. et al. Basic local alignment search tool. **Journal of Molecular Biology**, v. 215, n. 3, p. 403-410, 1990.
- ARAVIND, L.; KOONIN, E. V. Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. **Genome Research**, v. 11, n. 8, p. 1365–1374, 2001.
- ARAVIND, L.; WALKER, D. R.; KOONIN, E. V. Conserved domains in DNA repair proteins and evolution of repair systems. **Nucleic Acids Research**, v. 27, n. 5, p. 1223–1242, 1999.
- BALASINGHAM, S. V. et al. Enzymatic activities and DNA substrate specificity of mycobacterium tuberculosis DNA helicase XPB. **PLoS ONE**, v. 7, n. 5, 2012.
- BARZILAI, A.; YAMAMOTO, K. I. DNA damage responses to oxidative stress. **DNA Repair**, v. 3, p. 1109–1115, 2004.
- BELCHER, J. D. et al. Heme degradation and vascular injury. **Antioxidants & redox signaling**, v. 12, n. 2, p. 233–248, 2010.
- BENJAMIN H. LOK1, 2 AND SIMON N. POWELL1. Molecular Pathways: Understanding the Role of Rad52 in Homologous Recombination for Therapeutic Advancement. **Changes**, v. 29, n. 23, p. 997–1003, 2012.
- BEUMER, K. J. et al. Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. **Proceedings of the National Academy of Sciences of the United States of America**, v. 105, n. 50, p. 19821–19826, 2008.
- BISWAS, T. et al. DNA-dependent ATPase activity of bacterial XPB helicases. **Biochemistry**, v. 48, n. 12, p. 2839–2848, 2009.
- BLACK, A. T. et al. Increased oxidative stress and antioxidant expression in mouse keratinocytes following exposure to paraquat. **Toxicology and Applied Pharmacology**, v. 231, n. 3, p. 384–392, 2008.
- BROUGH, R. et al. Functional analysis of Drosophila melanogaster BRCA2 in DNA repair. **DNA Repair**, v. 7, n. 1, p. 10–19, 2008.
- BRUNK, K. et al. Microcephalin coordinates mitosis in the syncytial Drosophila embryo. **Journal of cell science**, v. 120, n. Pt 20, p. 3578–3588, 2007.
- BUCHAN, D. W. A. et al. Scalable web services for the PSIPRED Protein Analysis Workbench. **Nucleic acids research**, v. 41, n. Web Server issue, 2013.



CENTERS OF DISEASE CONTROL AND PREVENTION, Chikungunya virus, 2015. Disponível em: <http://www.cdc.gov/chikungunya/index.html>.

CENTERS OF DISEASE CONTROL AND PREVENTION, Mosquito Life-Cycle, 2012. Disponível em: [http://www.cdc.gov/Dengue/entomologyEcology/m\\_lifecycle.html](http://www.cdc.gov/Dengue/entomologyEcology/m_lifecycle.html).

CENTERS OF DISEASE CONTROL AND PREVENTION, Zika Virus, 2015. Disponível em: <http://www.cdc.gov/zika/index.html>.

CHINTAPALLI, S. V et al. Reevaluation of the evolutionary events within recA/RAD51 phylogeny. **BMC genomics**, v. 14, n. 1, p. 240, 2013.

CIAPPONI, L. et al. The Drosophila Mre11/Rad50 complex is required to prevent both telomeric fusion and chromosome breakage. **Current Biology**, v. 14, n. 15, p. 1360–1366, 2004.

CICCIA, A.; ELLEDGE, S. J. The DNA Damage Response: Making It Safe to Play with Knives. **Molecular Cell**, v. 40, n. 2, p. 179–204, 2010.

CLAPPERTON, J. A et al. Structure and mechanism of BRCA1 BRCT domain recognition of phosphorylated BACH1 with implications for cancer. **Nature structural & molecular biology**, v. 11, n. 6, p. 512–518, 2004.

CLEAVER, J. E.; LAM, E. T.; REVET, I. Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity. **Nature reviews. Genetics**, v. 10, n. 11, p. 756–768, 2009.

CLEMONS, A. et al. Aedes aegypti: An emerging model for vector mosquito development. **Cold Spring Harbor Protocols**, v. 5, n. 10, 2010.

COLE, C.; BARBER, J. D.; BARTON, G. J. The Jpred 3 secondary structure prediction server. **Nucleic acids research**, v. 36, n. Web Server issue, 2008.

CRITCHLOW, S. E.; JACKSON, S. P. DNA end-joining: From yeast to man. **Trends in Biochemical Sciences**, 1998.

CUFF, J. A.; BARTON, G. J. Application of multiple sequence alignment profiles to improve protein secondary structure prediction. **Proteins**, v. 40, n. 3, p. 502–511, 2000.

DANSA-PETRETSKI, M. et al. Antioxidant role of Rhodnius prolixus heme-binding protein: Protection against heme-induced lipid peroxidation. **Journal of Biological Chemistry**, v. 270, n. 18, p. 10893–10896, 1995

DAS, A. et al. NEIL2-initiated, APE-independent repair of oxidized bases in DNA: Evidence for a repair complex in human cells. **DNA Repair**, v. 5, n. 12, p. 1439–1448, 2006.

DELLA, M. et al. Mycobacterial Ku and ligase proteins constitute a two-component NHEJ repair machine. **Science (New York, N.Y.)**, v. 306, n. 5696, p. 683–685, 2004.

DESHPANDE, R. A.; WILSON, T. E. Modes of interaction among yeast Nej1, Lif1 and Dnl4 proteins and comparison to human XLF, XRCC4 and Lig4. **DNA Repair**, v. 6, n. 10, p. 1507–1516, 2007.

DOHERTY, A. J.; JACKSON, S. P.; WELLER, G. R. Identification of bacterial homologues of the Ku DNA repair proteins. **FEBS Letters**, v. 500, n. 3, p. 186-188, 2001.

DRONAMRAJU, R.; MASON, J. M. Recognition of double strand breaks by a mutator protein (MU2) in *Drosophila melanogaster*. **PLoS Genetics**, v. 5, n. 5, 2009.

EARLEY, J. N.; TURCHI, J. J. Interrogation of nucleotide excision repair capacity: impact on platinum-based cancer therapy. **Antioxidants & redox signaling**, v. 14, n. 12, p. 2465–2477, 2011.

EISEN, J. A; HANAWALT, P. C. A phylogenomic study of DNA repair genes, proteins, and processes. **DNA Repair**, v. 435, n. 3, p. 171–213, 1999.

EKBLAD, C. M. S. et al. Characterisation of the BRCT domains of the breast cancer susceptibility gene product BRCA1. **Journal of Molecular Biology**, v. 320, n. 02, p. 431–442, 2002.

ERIE, D. A.; WENINGER, K. R. Single molecule studies of DNA mismatch repair. **DNA Repair**, v. 20, p. 71–81, 2014.

FAGBEMI, A. F.; ORELLI, B.; SCHÄRER, O. D. Regulation of endonuclease activity in human nucleotide excision repair. **DNA Repair**, v. 10, n. 7, p. 722–729, 2011.

FALLON, A. M.; KURTZ, C. M.; CARROLL, E. M. The oxidizing agent, paraquat, is more toxic to *Wolbachia* than to mosquito host cells. **In Vitro Cellular and Developmental Biology - Animal**, v. 49, n. 7, p. 501–507, 2013.

FANG, M. et al. *Drosophila* ptp is essential for anterior/posterior patterning in development and interacts with the PcG and trxG pathways. **Development (Cambridge, England)**, v. 136, n. 11, p. 1929–1938, 2009.

FATTAH, F. et al. Ku regulates the non-homologous end joining pathway choice of DNA double-strand break repair in human somatic cells. **PLoS genetics**, v. 6, n. 2, p. e1000855, 2010.

FINN, R. D. et al. The protein families database. **Nucleic Acids Research**, v. 42, p. 222-230, 2014.

FINN, R. D.; CLEMENTS, J.; EDDY, S. R. HMMER web server: Interactive sequence similarity searching. **Nucleic Acids Research**, v. 39, n. SUPPL. 2, 2011.

FRIEDBERG, E. C. How nucleotide excision repair protects against cancer. **Nature reviews. Cancer**, v. 1, n. October, p. 22–33, 2001.

FUNDAÇÃO NACIONAL DE SAÚDE, Dengue instruções para pessoal de combate ao vetor: manual de normas técnicas. - 3. ed., rev. - Brasília: Ministério da Saúde, 2001.

FUKUSHIMA, T. et al. Mechanism of cytotoxicity of paraquat. **Environmental health and preventive medicine**, v. 7, n. 3, p. 89–94, 2002.

GAWARAMMANA, I. B.; BUCKLEY, N. A. Medical management of paraquat ingestion. **British Journal of Clinical Pharmacology**, v. 72, n. 5, p. 745–757, 2011.

GERLOFF, D. L. et al. BRCT domains: A little more than kin, and less than kind. **FEBS Letters**, v. 586, n.17, p. 2711–2716, 2012.

GRAÇA-SOUZA, A. V. et al. Adaptations against heme toxicity in blood-feeding arthropods. **Insect Biochemistry and Molecular Biology**, v. 36, p. 322–335, 2006.

HARFE, B. D.; JINKS-ROBERTSON, S. DNA mismatch repair and genetic instability. **Annual Review of Genetics**, v. 34, p. 359–399, 2000.

HELLEDAY, T. et al. DNA double-strand break repair: From mechanistic understanding to cancer treatment. **DNA Repair**, v. 6, p. 923–935, 2007.

HSIEH, P.; YAMANE, K. DNA mismatch repair: Molecular mechanism, cancer, and ageing. **Mechanisms of Ageing and Development**, v. 129, p. 391–407, 2008.

IYAMA, T.; WILSON, D. M. DNA repair mechanisms in dividing and non-dividing cells. **DNA Repair**, v. 12, p. 620–636, 2013.

JONES, D. T.; TAYLOR, W. R.; THORNTON, J. M. The rapid generation of mutation data matrices from protein sequences. **Computer applications in the biosciences : CABIOS**, v. 8, n. 3, p. 275–282, 1992.

LEUNG, C. C. Y.; GLOVER, J. N. M. BRCT domains: Easy as one, two, three. **Cell Cycle**, v. 10, p. 2461–2470, 2011.

LI, S. et al. Binding of CtIP to the BRCT repeats of BRCA1 involved in the transcription regulation of p21 is disrupted upon DNA damage. **Journal of Biological Chemistry**, v. 274, p. 11334–11338, 1999.

LIM, W. A.; PAWSON, T. Phosphotyrosine Signaling: Evolving a New Cellular Communication System. **Cell**, v. 142, p. 661–667, 2010.

LIMA, V. L. A et al. The antioxidant role of xanthurenic acid in the *Aedes aegypti* midgut during digestion of a blood meal. **PLoS ONE**, v. 7, n. 6, p. 1–8, 2012.

LIN, Y. C.; CHOI, W. S.; GRALLA, J. D. TFIIF XPB mutants suggest a unified bacterial-like mechanism for promoter opening but not escape. **Nature structural & molecular biology**, v. 12, n. 7, p. 603–607, 2005.

LIN, Z. et al. Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. **Proceedings of the National**

**Academy of Sciences of the United States of America**, v. 103, n. 27, p. 10328–10333, 2006.

LIN, Z.; NEI, M.; MA, H. The origins and early evolution of DNA mismatch repair genes--multiple horizontal gene transfers and co-evolution. **Nucleic acids research**, v. 35, n. 22, p. 7591–7603, 2007.

LIU, H. et al. Structure of the DNA Repair Helicase XPD. **Cell**, v. 133, n. 5, p. 801–812, 2008.

LIVAK, K. J.; SCHMITTGEN, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. **Methods (San Diego, Calif.)**, v. 25, n. 4, p. 402–408, 2001.

LOK, B. H.; POWELL, S. N. Molecular pathways: Understanding the role of Rad52 in homologous recombination for therapeutic advancement. **Clinical Cancer Research**, v. 18, n. 23, p. 6400–6406, 2012.

MA, S. et al. CRISPR/Cas9 mediated multiplex genome editing and heritable mutagenesis of BmKu70 in Bombyx mori. **Scientific reports**, v. 4, p. 4489, 2014.

MAEDA, T. et al. Nucleotide Excision Repair Genes are Upregulated by Low- Dose Artificial Ultraviolet B: Evidence of a Photoprotective SOS Response? **The Society for Investigative Dermatology**, v. 53, p. 1490–1497, 2001.

MANSOUR, W. Y. et al. Hierarchy of nonhomologous end-joining, single-strand annealing and gene conversion at site-directed DNA double-strand breaks. **Nucleic Acids Research**, v. 36, n. 12, p. 4088–4098, 2008.

MARCHLER-BAUER, A. et al. CDD: Specific functional annotation with the Conserved Domain Database. **Nucleic Acids Research**, v. 37, n. SUPPL. 1, 2009.

MARTEIJN, J. A et al. Understanding nucleotide excision repair and its roles in cancer and ageing. **Nat Rev Mol Cell Biol**, v. 15, n. 7, p. 465–481, 2014.

MARTIN, L. M. et al. DNA mismatch repair and the DNA damage response to ionizing radiation: Making sense of apparently conflicting data. **Cancer Treatment Reviews**, v. 36, n. 7, p. 518–527, 2010.

MESQUITA, R. D. et al. Tandem BRCT Domains: DNA's Praetorian Guard. **Genes & Cancer**, v. 1, n. 11, p. 1140–1146, 2010.

MINOIS, N. et al. Spermidine promotes stress resistance in Drosophila melanogaster through autophagy-dependent and -independent pathways. **Cell Death and Disease**, v. 3, 2012.

MORITA, R. et al. Molecular mechanisms of the whole DNA repair system: a comparison of bacterial and eukaryotic systems. **Journal of nucleic acids**, v. 2010, 2010.

MORLAND, I. et al. Human DNA glycosylases of the bacterial Fpg/MutM superfamily: an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA. **Nucleic acids research**, v. 30, n. 22, p. 4926–4936, 2002.

NATARAJAN, S.; JAKOBSSON, E. Functional equivalency inferred from “authoritative sources” in networks of homologous proteins. **PLoS ONE**, v. 4, n. 6, 2009.

O’DRISCOLL, M.; JEGGO, P. A. The role of double-strand break repair - insights from human genetics. **Nature reviews. Genetics**, v. 7, n. January, p. 45–54, 2006.

OLIVEIRA, J. H. M. et al. Blood meal-derived heme decreases ROS levels in the midgut of *Aedes aegypti* and allows proliferation of intestinal microbiota. **PLoS Pathogens**, v. 7, n. 3, 2011.

OVERCASH, J. M. et al. Understanding the DNA damage response in order to achieve desired gene editing outcomes in mosquitoes. **Chromosome Research**, 2015.

PAIVA-SILVA, G. O. et al. A heme-degradation pathway in a blood-sucking insect. **Proceedings of the National Academy of Sciences**, v. 103, n. 21, p. 8030-8035, 2006.

PAPOULI, E.; CEJKA, P.; JIRICNY, J. Advances in Brief Dependence of the Cytotoxicity of DNA-Damaging Agents on the Mismatch Repair Status of Human Cells. **Trends in Cell Biology**, p. 3391–3394, 2004.

PITCHER, R. S.; BRISSETT, N. C.; DOHERTY, A. J. Nonhomologous end-joining in bacteria: a microbial perspective. **Annual review of microbiology**, v. 61, p. 259–282, 2007.

PORTAL DA DAÚDE, Casos de Dengue. Brasil, Grandes Regiões e Unidades Federadas, 1990 a 2013. Disponível em:  
<http://portalsaude.saude.gov.br/images/pdf/2014/julho/31/Dengue-classica-at---2013.pdf>

PORTAL DA DAÚDE, Óbitos por Febre Hemorrágica da Dengue. Brasil, Grandes Regiões e Unidades Federadas, 1990 a 2013. Disponível em:  
<http://portalsaude.saude.gov.br/images/pdf/2014/julho/31/--bitos-por-casos-graves-at---2013.pdf>

RAPPAS, M.; OLIVER, A. W.; PEARL, L. H. Structure and function of the Rad9-binding region of the DNA-damage checkpoint adaptor TopBP1. **Nucleic Acids Research**, v. 39, n. 1, p. 313–324, 2011.

RICKMYRE, J. L. et al. The *Drosophila* homolog of MCPH1, a human microcephaly gene, is required for genomic stability in the early embryo. **Journal of cell science**, v. 120, n. 20, p. 3565–3577, 2007.

SADRZADEH, S. M. H. et al. Hemoglobin. A biologic Fenton reagent. **Journal of Biological Chemistry**, v. 259, n. 23, p. 14354–14356, 1984.

SALMON, A. B. et al. Fibroblast cell lines from young adult mice of long-lived mutant strains are resistant to multiple forms of stress. **American journal of physiology. Endocrinology and metabolism**, v. 289, n. 1, p. 23–29, 2005.

SANCAR, A. et al. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. **Annual review of biochemistry**, v. 73, p. 39–85, 2004.

SEABRA-JUNIOR, E. S. et al. FAT - Functional Analysis Tool. Brasil. **Software n. 11083-6, INPI**, 2011.

SHENG, Z. Z.; ZHAO, Y. Q.; HUANG, J. F. Functional evolution of BRCT domains from binding DNA to protein. **Evolutionary Bioinformatics**, v. 2011, p. 87–97, 2011.

SHRIVASTAV, M.; DE HARO, L. P.; NICKOLOFF, J. A. Regulation of DNA double-strand break repair pathway choice. **Cell research**, v. 18, n. 1, p. 134–147, 2008.

SHUMAN, S.; GLICKMAN, M. S. Bacterial DNA repair by non-homologous end joining. **Nature reviews. Microbiology**, v. 5, n. 11, p. 852–861, 2007.

SIMOSSIS, V. A.; HERINGA, J. The PRALINE online server: Optimising progressive multiple alignment on the web. **Computational Biology and Chemistry**, v. 27, n. 4-5, p. 511–519, 2003.

SIMOSSIS, V. A.; HERINGA, J. PRALINE: A multiple sequence alignment toolbox that integrates homology-extended and secondary structure information. **Nucleic Acids Research**, v. 33, n. SUPPL. 2, 2005.

SIRBU, B. M.; CORTEZ, D. DNA damage response: three levels of DNA repair regulation. **Cold Spring Harbor perspectives in biology**, v. 5, n. 8, 2013.

SMALLHORN, M.; MURRAY, M. J.; SAINT, R. The epithelial-mesenchymal transition of the *Drosophila* mesoderm requires the Rho GTP exchange factor Pebble. **Development (Cambridge, England)**, v. 131, n. 11, p. 2641–2651, 2004.

SPIES, M. Two steps forward, one step back: Determining XPD helicase mechanism by single-molecule fluorescence and high-resolution optical tweezers. **DNA Repair**, v. 20, p. 58–70, 2014.

SRIVASTAVA, M. et al. The *Amphimedon queenslandica* genome and the evolution of animal complexity. **Nature**, v. 466, n. 7307, p. 720–726, 2010.

STAMATAKIS, A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. **Bioinformatics**, v. 22, n. 21, p. 2688–2690, 2006.

STOJIC, L.; BRUN, R.; JIRICNY, J. Mismatch repair and DNA damage signalling. **DNA Repair**, v. 3, p. 1091–1101, 2004.

TERUAKI IYAMA; DAVID M. WILSON III. DNA repair mechanisms in dividing and non-dividing cells. **Changes**, v. 29, n. 8, p. 997–1003, 2012.

- VENKATESH, B. et al. Elephant shark genome provides unique insights into gnathostome evolution. **Nature**, v. 505, n. 7482, p. 174–9, 2014.
- WALLACE, S.; MURPHY, D.; SWEASY, J. Base excision repair and cancer. **Cancer letters**, v. 327, p. 73–89, 2012.
- WATERHOUSE, A. M. et al. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. **Bioinformatics**, v. 25, n. 9, p. 1189–1191, 2009.
- WEI, D. S.; RONG, Y. S. A Genetic Screen For DNA Double-Strand Break Repair Mutations in Drosophila. **Genetics**, v. 177, p. 63–77, 2007.
- WELLER, G. R. et al. Identification of a DNA nonhomologous end-joining complex in bacteria. **Science (New York, N.Y.)**, v. 297, n. 5587, p. 1686–1689, 2002.
- WORLD HEALTH ORGANIZATION, Chikungunya, 2015. Disponível em: <http://www.who.int/mediacentre/factsheets/fs327/en/>.
- WORLD HEALTH ORGANIZATION, Dengue, countries or areas at risk, 2013. Disponível em: [http://gamapserver.who.int/mapLibrary/Files/Maps/Global\\_DengueTransmission\\_ITHR\\_iskMap.png?ua=1](http://gamapserver.who.int/mapLibrary/Files/Maps/Global_DengueTransmission_ITHR_iskMap.png?ua=1)
- WORLD HEALTH ORGANIZATION, Dengue and severe dengue, 2015. Disponível em: <http://www.who.int/mediacentre/factsheets/fs117/en/>.
- WORLD HEALTH ORGANIZATION, Vector-borne diseases, 2014. Disponível em: <http://www.who.int/mediacentre/factsheets/fs387/en/>.
- WILLIAMS, R. S.; GREEN, R.; GLOVER, J. N. Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. **Nature structural biology**, v. 8, n. 10, p. 838–842, 2001.
- WOODS, N. T. et al. Charting the Landscape of Tandem BRCT Domain-Mediated Protein Interactions. **Science Signaling**, v. 5, n. 242, p. rs6, 2012.
- WU, D.; TOPPER, L. M.; WILSON, T. E. Recruitment and Dissociation of Nonhomologous End Joining Proteins at a DNA Double-Strand Break in *Saccharomyces cerevisiae*. **Genetics**, v. 178, p. 1237–1249, 2008.
- XU, G. et al. Base excision repair, aging and health span. **Mechanisms of Ageing and Development**, v. 129, p. 366–382, 2008.
- YAMAMOTO, R. R. et al. The *Drosophila* mus101 gene, which links DNA repair, replication and condensation of heterochromatin in mitosis, encodes a protein with seven BRCA1 C-terminus domains. **Genetics**, v. 156, n. 2, p. 711–721, 2000.
- YOO, S.; MCKEE, B. D. Overexpression of *Drosophila* Rad51 protein (DmRad51) disrupts cell cycle progression and leads to apoptosis. **Chromosoma**, v. 113, n. 2, p. 92–101, 2004.

YOO, S.; MCKEE, B. D. Functional analysis of the *Drosophila* Rad51 gene (*spn-A*) in repair of DNA damage and meiotic chromosome segregation. **DNA Repair**, v. 4, n. 2, p. 231–242, 2005.

YU, X. et al. The BRCT domain is a phospho-protein binding domain. **Science (New York, N.Y.)**, v. 302, n. 2003, p. 639–642, 2003.

ZHANG, H. et al. BRCA1 physically associates with p53 and stimulates its transcriptional activity. **Oncogene**, v. 16, p. 1713–1721, 1998a.

ZHANG, X. et al. Structure of an XRCC1 BRCT domain: A new protein-protein interaction module. **EMBO Journal**, v. 17, n. 21, p. 6404–6411, 1998b.

ZOU, Y. et al. Crystal structure of triple-BRCT-domain of ECT2 and insights into the binding characteristics to CYK-4. **FEBS Letters**, v. 588, n. 17, p. 2911–2920, 2014.