INGRID REALE ALVES

Estudo da Síntese Translesão em Caulobacter

crescentus

Tese de doutorado apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo para obtenção do título de Doutor em Ciências

São Paulo 2018

INGRID REALE ALVES

Estudo da Síntese Translesão em Caulobacter

crescentus

Tese de doutorado apresentada ao Programa de Pós-Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo para obtenção do título de Doutor em Ciências

Área de concentração: Microbiologia

Orientador: Dr. Rodrigo da Silva Galhardo

Versão Original.

São Paulo 2018

Este trabalho é dedicado à todas grandes mulheres da minha vida: *Hildeth, Joana, Antônia e Maria.* Não existe retribuição para todo o amor, cuidado e leveza que vocês colocam nos meus dias. Este trabalho também é dedicado a todas as *crianças colecionadoras de insetos*.

AGRADECIMENTOS

Agradeço primeiramente ao meu orientador, Prof. Dr. Rodrigo da Silva Galhardo, por me proporcionar trabalhar com um dos mecanismos mais intrigantes presentes na célula. Além da oportunidade, agradeço a paciência, incentivo e conversas construtivas ao longo de todo este trabalho. Definitivamente sem você seria impossível chegar aonde chegamos. Foi uma verdadeira honra ser sua primeira aluna de doutorado e fazer parte do seu grupo.

Agradeço aos professores Dr. Beny Spira e Dra. Marilis Varques pelas dúvidas sanadas e por sempre estarem dispostos a ajudar.

Agradeço ao Prof. Dr. Carlos Menck pelo espaço no seu laboratório e pela sua disponibilidade.

Agradeço ao Prof. Márcio Dias por ter cedido o seu espaço para que tentássemos purificar a proteína ImuA. Em especial, agradeço sua aluna de doutorado Priscila Bury, que me acompanhou nas várias tentativas de conseguir a tão sonhada proteína.

Agradeço ao Prof. Dr. Robson Souza pela disponibilidade em me ajudar com qualquer dúvida de bioinformática.

Agradeço aos meus colegas de laboratório pela ajuda em alguns dos experimentos presentes nesta tese, em especial ao Frank por ser uma verdadeira inspiração para qualquer cientista ao longo de todos os dias.

Agradeço aos alunos Marco Noronha e Aline Freitas que usaram parte da sua iniciação científica para colaborar com os experimentos da minha tese.

Agradeço aos nossos técnicos de laboratório Alexandre e Aline pela ajuda com os materiais.

Agradeço as queridas amigas-cientistas Carina e Ellen que enquanto presentes no nosso laboratório foram fundamentais para a construção deste trabalho.

Agradeço a todos os amigos-cientistas que me ajudaram com experimentos e/ou apenas uma conversa no bar, em especial: Luiz Gustavo, Henrique e Larissa pelas discussões acadêmicas, experimentos compartilhados, torcida nas clonagens e ajuda nas flutuações. Vocês são especiais e foi uma honra aprender com as nossas conversas.

Agradeço aos membros da minha banca de qualificação pela discussão construtiva que me possibilitou pensar em outros caminhos para resultados confusos e ganhar confiança em alguns dos bons resultados apresentados neste trabalho.

Não poderia deixar de agradecer as minhas famílias (a que está longe e a família paulistana que me adotou) e a todos os meus amigos não cientistas que estiveram presentes em todos os momentos da construção deste trabalho. TODOS vocês foram extremamente importantes. Eu tenho certeza que sem a motivação e os puxões de orelha de cada um seria improvável que eu continuasse.

Em especial, eu gostaria de agradecer ao meu avô Osvaldo que sempre me presenteou com insetos em potinhos de conserva e contos amazônicos sobre caranguejos, macacos extravagantes e outros seres extraordinários. Obrigada pela coleção de insetos e por ter sido a minha maior inspiração para me tornar bióloga e uma criança colecionadora de insetos.

Agradeço ao central analítica do IQ-USP pelo apoio experimental.

Agradeço ao CEFAP-USP pela estrutura e apoio durante os experimentos de construção de bibliotecas genômicas e sequenciamento.

Agradeço a CAPES pela bolsa e a FAPESP e CNPQ pelo o financiamento deste trabalho.

"(...) aquele tempo estava chegando ao fim.

Eu me sentia como um cavalo de corrida sem hipódromo ou um campeão universitário repentinamente confrontado com Wall Street e um terno de executivo, seus dias de glória reduzidos a um pequeno troféu dourado na prateleira com uma data gravada, como num túmulo.

Eu via minha vida se ramificando à minha frente como a figueira daquele conto.

Da ponta de cada galho, como um enorme figo púrpura, um futuro maravilhoso acenava e cintilava. Um desses figos era um lar feliz com maridos e filhos, outro era uma poeta famosa, outro, uma professora brilhante, (...) e acima desses figos havia muitos outros que eu não conseguia enxergar.

Me vi sentada embaixo da árvore, morrendo de fome, simplesmente porque eu não conseguia decidir com qual figo eu ficaria. Eu queria todos eles, mas escolher apenas um significava perder todo o resto, e enquanto eu ficava ali sentada, incapaz de tomar decisão, os figos começaram a encolher e ficar pretos e, um por um, desabaram no chão aos meus pés".

Redoma de vidro (p 88 - 89)

Sylvia Plath

RESUMO

ALVES, I. R. ESTUDO DA SÍNTESE TRANSLESÃO EM *Caulobacter crescentus*. Tese de doutorado em Microbiologia – Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, 2018.

Como é de suma importância a integridade da informação contida no DNA, este recebe proteção contra agentes danosos que podem prejudicar sua estrutura. Mesmo em caso de dano, a célula possui um grupo de proteínas que estão envolvidas na correção e mitigação destes danos. O primeiro grupo é um conjunto de proteínas envolvidas no reparo de DNA livre de erro. Caso estas proteínas não consigam minimizar os danos, outro conjunto de proteínas é expresso como uma alternativa ao reparo. Dentre estas, estão as DNA polimerases especializadas em usar uma fita de DNA danificada como molde para replicação. Este mecanismo possibilita à célula sobreviver aos danos potencialmente citotóxicos, às custas de mutagênese. Em bactérias, a reposta ao dano de DNA envolve um conjunto de proteínas que são expressas como parte da resposta SOS. Dentre elas estão enzimas envolvidas na síntese translesão (TLS). Diferentemente de Escherichia coli que possui três polimerases propensas a erro especializadas em TLS, Caulobacter crescentus possui um cassete mutagênico imuABC que está implicado na síntese de DNA usando como molde uma fita danificada. Neste trabalho, estudamos o mecanismo de TLS mediado por ImuABC nesta bactéria, e encontramos uma série de diferencas com o mecanismo de bypass realizado pela principal polimerase implicada em TLS em E. coli (Pol V). As proteínas ImuABC quando expressas em níveis máximos da resposta SOS não são capazes de aumentar as taxas de mutagênese espontânea. O produto do operon imuABC, diferentemente da Pol V, não necessita de RecA para realizar TLS. Apenas a expressão destas proteínas em um background sem o gene recA já é suficiente para que ocorra a mutagênese induzida por UVC. Ao estudar a mutagênese como resposta ao dano de DNA induzido por radiação UVC em níveis genômicos em C. crescentus, notamos que a maioria das mutações encontradas está presente em regiões que possuem pirimidinas adjacentes que sabidamente são extremamente reativas à radiação UVC, levando à formação de fotoprodutos. Nossos dados sugerem que existe uma região no cromossomo circular de C. crescentus que é preferencialmente mutada, e este acúmulo de mutações pode ser consequência do reparo que acontece próximo à origem replicativa, deixando as mutações acumuladas próximas à região de término da replicação.

Palavras chave: Síntese translesão, Dano no DNA, ImuABC, Caulobacter crescentus

ABSTRACT

ALVES, I. R. STUDY OF TRANSLESION DNA SYNTHESIS IN Caulobacter crescentus. PhD thesis in Microbiology – Biomedical Sciences Institute, University of Sao Paulo, Sao Paulo, 2018.

As the integrity of information contained in DNA is of utmost importance, it receives protection against harmful agents that may harm its structure. Even in case of damage, the cell has a group of proteins that are involved in the correction and mitigation of these damages. The first group is a set of proteins involved in error-free DNA repair. If these proteins fail to minimize damage, another set of proteins is expressed as an alternative to repair. Among these are DNA polymerases that specialize in using a damaged DNA strand as a template for replication. This mechanism enables the cell to survive potentially cytotoxic damage at the expense of mutagenesis. In bacteria, the DNA damage response involves a set of proteins that are expressed as part of the SOS response. Among them are enzymes involved in translesion synthesis (TLS). Unlike Escherichia coli that has three TLS error-prone polymerases, Caulobacter crescentus bears the imuABC mutagenic cassette that is involved in DNA synthesis using a damaged template. In this work, we studied the mechanism of TLS mediated by ImuABC in this bacterium, and we found a number of differences relative to the characteristics of the principal polymerase involved in TLS in E. coli (Pol V). ImuABC proteins when expressed at maximum levels of the SOS response are not able to increase the rates of spontaneous mutagenesis. ImuABC, unlike Pol V, does not require RecA to perform TLS. The presence of these proteins in a background without the recA gene is sufficient for UVC-induced mutagenesis to occur. In studying mutagenesis as a response to DNA damage induced by UVC radiation at genomic levels in C. crescentus, we noted that most of the mutations found are present in regions that have adjacent pyrimidines, which are known to be extremely reactive to UVC radiation, leading to the formation of photoproducts. Our data suggest that there is a region on the circular chromosome of C. crescentus that is preferably mutated, and this accumulation of mutations may be a consequence of the repair occurring near the replicative origin, leaving the accumulated mutations close to the replication termination region.

Keywords: Translesion DNA synthesis, DNA damage, ImuABC, Caulobacter crescentus

LISTA DE ABREVIAÇÕES E SIGLAS

 μG – Micrograma

µg/mL – Micrograma por mililitro

 μL – Microlitro

 $\boldsymbol{\mu}\boldsymbol{M}$ - Micromolar

4-NQO – 4-nitroquinolana 1-oxido

6-4 pp – 6,4- Pirimidina pirimidona

AAF - N-acetylaminofluorene

cDNA – DNA complementar

CDS – Sequência de DNA codificante

CEFAP - Centro de Facilidades de Apoio à Pesquisa

CPD – Dímero de Pirimidina Ciclobutano

dG – Desoxiguanosina

DMSO - Dimetilsulfóxido

DNA – Ácido desoxirribonucleico

DO₆₀₀ – Densidade óptica com absorbância em 600 nanômetros

DO550 – Densidade óptica com absorbância em 550 nanômetros

fsDNA – Fita-simples de DNA

g – Força centrífuga relativa

GC – Guanosina e Citosina

GG – Guanosina-Guanosina

 J/m^2 – Joule por metro quadradro

Kbp – Quilo pares de bases

 \mathbf{L} – Litro

 $\mathbf{M} - Molar$

mL - Mililitro

 \mathbf{mM} – Milimolar

MMC – Mitomicina C

MMR – Mismatch repair (reparo de bases mal emparelhadas)

MMS – Metil metanosulfonato

Na-MOPS – MOPS sódio

NAPs - Proteínas de Associação ao Nucleóide

- NFZ Nitrofurazona
- OriC Origem cromossômica (Origem de replicação)
- pb Par de bases
- PCR Reação em Cadeia da Polimerase
- PMSF Fluoreto de Fenilmetilsulfonilo
- Pol II Polimerase II
- Pol III Polimerase III
- Pol IV Polimerase IV
- Pol V Polimerase V
- qPCR Reação em Cadeia da Polimerase Quantitativa
- qRT-PCR Reação em Cadeia da Polimerase em Tempo Real quantitativa
- RecA* Proteína RecA ativa
- Rif^R Rifampicina resistentes
- **RNA** Ácido ribonucleico
- SNP Single Nucleotide Polymorphism
- Ssb Proteína de ligação a fita-simples de DNA
- \mathbf{T} Timina
- TLS Translesion synthesis (síntese translesão)
- UFC Unidade Formadora de Colônias
- UV Ultra-Violeta
- UVC Radiação Ultravioleta C

SUMÁRIO

1 INTRODUÇÃO 1.1 - CROMOSSOMO BACTERIANO: ORGANIZAÇÃO E COMPACTAÇÃO	13 13
1.2 – REPLISSOMO	14
1.2.1 – Replicase (Pol III)	14
1.2.2 – Duplicação do genoma	16
1.3 - MECANISMOS DE RESPOSTA AO DANO NO DNA (DDR)	17
1.3.1 – Bloqueio e colapso da forquilha replicativa	19
1.3.2 – Sistema SOS	19
1.3.3 - Diversidade de polimerases em procariotos e síntese translesão de DNA (TL	LS) 22
1.3.3.1 - Mecanismos de TLS mediado pelas polimerases Pol IV e Pol V	23
1.3.4 - Algumas bactérias utilizam outra Pol III para fazer TLS: o cassete mutagênie <i>imuABC (dnaE2)</i>	co 27
1.4 - <i>Caulobacter crescentus</i> : UM MODELO ALTERNATIVO PARA BIOLOGIA CELULAR E GENÉTICA	x 29
2 OBJETIVOS 2.1 - OBJETIVO GERAL	32 32
2.1.1 – Objetivos específicos	32
3 MATERIAL E MÉTODOS 3.1 - CEPAS BACTERIANAS, PLASMÍDEOS E CONDIÇÕES DE CRESCIMENTO DAS LINHAGENS UTILIZADAS NESTE TRABALHO	33 33
3.2 – TÉCNICAS DE BIOLOGIA MOLECULAR UTLIZADAS EM CLONAGEN	M
	39
3.2.1 – Reação em cadeia da Polimerase (PCR) e Reação de Ligação	39
3.2.2 – Preparação de células quimio-competentes para clonagem	40
3.2.3 – Transformação de células competentes	40
3.4 - TRATAMENTO DO RNA COM DNASE, SÍNTESE DE CDNA E qRT-PCR	342
3.6 - ENSAIOS DE MUTAGÊNESE	44
3.7 -TESTE DE FLUTUAÇÃO	46
3.8 - ENSAIOS DE TRANSDUÇÃO	46
3.10 - EXPRESSÃO DA PROTEÍNA ImuA	47
3.12 - SEQUENCIAMENTO DE GENOMAS	50
3.13 - ANÁLISE DOS GENOMAS DE Caulobacter crescentus	52

3.14 – ESTATÍSTICA	
4 RESULTADOS E DISCUSSÃO	
4.1 - A EXPRESSÃO CONSTITUTIVA DE imuABC EM NÍVEIS MÁ	XIMOS
OBTIDOS NA RESPOSTA SOS NÃO AUMENTA A MUTAGÊNES	E
ESPONTANEA.	
4.2 - A EXPRESSÃO CONSTITUTIVA DE <i>imuABC</i> EM NÍVEIS MÁ RESPOSTA SOS NÃO AUMENTA A MUTAGÊNESE INDUZIDA P NO DNA.	XIMOS DA OR DANOS 57
4 3 - EXPRESSÃO DOS GENES <i>IMUARC</i> EM NÍVEIS MÁXIMOS	
ENCONTRADOS NA RESPOSTA SOS RECUPERA A MUTAGÊNE	SE
INDUZIDA POR UVC NA AUSÊNCIA DE recA	60
4.4 - EXPRESSÃO DA PROTEÍNA ImuA	66
4. 5 - ImuA NÃO INTERFERE NAS FUNÇÕES DESEMPENHADAS EM C. crescentus	POR RecA
4.6 - OUTROS FATORES INDEPENDENTES DE ImuABC SÃO RES POR PARTE DA MUTAGÊNESE INDUZIDA POR UV, E PELA MU INDUZIDA POR MMS EM <i>C. crescentus</i>	SPONSÁVEIS ITAGÊNESE 70
4. 7 - ESTUDO DO ESPECTRO DE MUTAÇÕES INDUZIDAS POR NÍVEL GENÔMICO	UVC EM 75
4.7.1 - Determinação do genótipo da linhagem NA1000 presente no r laboratório	nosso 75
4.7.2 - Taxas de sobrevivência distintas ao UVC alteram a frequência em genomas irradiados	de mutação 77
4.7.3 – Distribuição de mutações induzidas por UV no cromossomo c crescentus	le <i>C</i> . 79
4.7.4 - Espectro de mutações induzidas por UVC no genoma	
4.7.5 - O agrupamento de mutações derivadas do mesmo genoma poo distância percorrida pela TLS	le indicar a
4.7.6 - As mutações induzidas por UV ocorrem majoritariamente em dinucleotídeos de pirimidinas	regiões de 89
4.7.7 – As fitas contínua e descontínua são igualmente mutáveis por l crescentus	uz UV em <i>C</i> .
4.8 - MUTAGÊNESE INDEPENDENTE DE ImuABC A NÍVEL GEN	ÔMICO 93
4.9 - MODELO DE TLS MEDIADO POR ImuABC	
5 CONCLUSÕES & DISCUSSÃO GERAI	02
6 PESPECTIVAS	
REFERÊNCIAS BIBLIOGRÁFICAS	
	113

1 INTRODUÇÃO

Por volta dos anos 50 do século XX o DNA (ácido desoxirribonucleico), molécula responsável pela hereditariedade, foi caracterizada estruturalmente por Watson & Crick (Watson & Crick, 1953). No final do seu artigo emblemático, os dois pesquisadores comentam que a sua organização em dupla hélice possibilita a transmissão de informação de uma célula para outra. A forma desta molécula e suas características únicas permitem a codificação de informações necessárias para a produção de todo o conteúdo celular, assim como a fabricação de novas células.

O conjunto destas características faz do DNA uma molécula extremamente importante. As células possuem componentes especializados em replicar sua informação (proteínas presentes no replissomo – ver tópico a seguir) e um conjunto de mecanismos de reparo de DNA e de tolerância aos danos que podem ser ocasionados durante o processo de replicação ou por agentes nocivos endógenos e/ou exógenos à célula (Friedberg et al., 2006).

1.1 - CROMOSSOMO BACTERIANO: ORGANIZAÇÃO E COMPACTAÇÃO

Em bactérias, o genoma geralmente está organizado em um cromossomo circular que se encontra em diferentes níveis de compactação durante o ciclo celular. Esta organização diferenciada permite uma série de processos como a replicação do DNA, transcrição em moléculas de RNA e recombinação (Badrinarayanan et al., 2015). O mais interessante é que em procariotos todos estes processos acontecem de forma concomitante com a migração destes cromossomos para os polos distintos da célula durante a divisão celular (Badrinarayanan et al., 2015). Hoje se sabe que o nucleóide bacteriano não está organizado de forma aleatória no citoplasma e, dependendo das condições de crescimento, pode mudar a sua distribuição. Isso só acontece porque o DNA está ligado às proteínas de associação ao nucleóide (NAPs: *nucleoid associated proteins*) (Badrinarayanan et al., 2015; Lewis et al., 2000). Ao se associarem ao DNA, estas proteínas modulam a natureza dinâmica do nucleóide, e regulam a expressão gênica (revisto por Dillon & Dornam, 2010).

1.2 - REPLISSOMO

Para que a divisão celular seja completada em todos os organismos, é necessário que aconteça primeiramente a duplicação do genoma. Este processo é responsabilidade de um grupo de proteínas conservadas em todos os domínios da vida, que são fundamentais para a replicação do DNA (Yao & O'Donnell, 2010; O'Donnell et al., 2013). Os principais componentes do replissomo de *Escherichia coli* incluem as proteínas DnaB (helicase), Ssb (*Single-Strand DNA-Binding* ou Proteína de Ligação a Fita-simples de DNA), DnaG (primase) e a holoenzima DNA Polimerase III (associada aos complexos de proteínas *beta-clamp* e *clamp loader*). As funções destas proteínas envolvem respectivamente a separação da dupla-hélice em duas fitas simples de DNA (fsDNA); ligação à fsDNA, estabilização e remoção de estruturas secundárias; confecção de iniciadores para replicação; além da replicação do DNA com alta processividade e precisão (Yao & O'Donnell, 2010).

Na presença de dano no DNA, a Polimerase III (Pol III), especializada em duplicar o genoma, fica impossibilitada de cumprir sua função, e pode ser substituída por outras polimerases que fazem a síntese de DNA na presença de dano. Estas enzimas propensas a erro usam os demais componentes do replissomo para fazer uma síntese através da lesão (Síntese translesão: TLS). Ao se ligar à cinta deslizante beta (*beta-clamp*) e à proteína DnaB, este replissomo especializado em fazer o *bypass* de lesões move-se lentamente e de forma breve até ser substituído pelos componentes do replissomo clássico que prossegue com a síntese de DNA livre de erro (O'Donnell et al., 2013).

1.2.1 – Replicase (Pol III)

A Pol III é uma holoenzima formada por 10 proteínas que podem ser divididas em três conjuntos funcionais diferentes. Em *E. coli* o cerne catalítico está duplicado, e sua constituição inclui as proteínas α (codificada por *dnaE*), ε (*dnaQ*) e θ (*holE*), sendo a proteína α a polimerase responsável pela síntese de DNA, enquanto que ε é encarregado da correção de erros da replicação, e θ é uma proteína estrutural envolvida na estabilização da holoenzima Pol III (O'Donnell, 2006). O outro complexo funcional consiste na cinta deslizante β ou *beta-clamp* (codificada por *dnaN*) (β_2 –*clamp*) que é um fator de processividade que assegura a ligação da Pol III com o DNA e, por fim, o *clamp loader* (DnaX_{cx}) que é um complexo de 6 proteínas: γ (*dnaX* – transcrito completo), τ_2 (também codificado pelo gene dnaX – em um quadro de leitura alternativo), δ (*holA*), δ ' (*holB*), χ (*holC*) e ψ (*holD*) (Robison & Van Oijen, 2013; O'Donnell, 2006). Na ausência de qualquer um dos grupos citados anteriormente, o cerne catalítico das replicases, responsável pela síntese de uma nova fita de DNA, não tem o mesmo nível de afinidade entre seus componentes, e sua ação não é distinguível de outros grupos de Polimerases quanto à processividade. Todavia, quando trabalham em conjunto, podem ser caracterizados como uma replicase livre de erro e altamente processiva (McHenry, 2011a).

O produto do gene *dnaE*, a proteína catalítica α (Pol III), que está implicada na replicação livre de erro em bactérias Gram-negativas, possui todos os domínios comuns às DNA polimerases, sendo eles *palm* (participa da catálise), *thumb* (ajuda na estabilização da ligação com o DNA) e *fingers* (liga-se aos dNTPs que estão sendo recrutados para a extensão da nova fita de DNA) (McHenry, 2011a; Yang & Woodgate, 2007). Esta proteína faz parte da Família C de DNA Polimerases, que agrupa a maioria das enzimas que conseguem sintetizar o DNA com precisão. Porém, o seu domínio *palm* possui um dobramento característico de polimerases da Família X, que é constituída de enzimas que replicam o DNA com u'ma processividade menor quando comparadas com as replicases (McHenry, 2011a).

A duplicação do genoma é catalisada pelo produto do gene *dnaE*. Porém, bactérias com Gram + utilizam além desta proteína o produto do gene *polC* para a síntese de DNA (McHenry, 2011b). Este codifica uma DNA polimerase da mesma família, homóloga à DnaE. Experimentos mostraram que em *Bacillus subtilis* a síntese de DNA é dividida por estas proteínas de acordo com a fita que está sendo sintetizada. Sendo DnaE responsável pela síntese da fita descontínua, enquanto que PolC sintetiza sem interrupção a fita contínua (McHenry, 2011b). Bruck e colaboradores (2003) mostraram que apesar de ambas proteínas estarem envolvidas na duplicação do genoma, o produto do gene DnaE é menos processivo *in vitro* ao replicar a fita descontínua quando comparado com o produto do gene PolC. Entretanto, a taxa de replicação mais lenta não contribui para o aumento da taxa de erro desta polimerase (McHenry, 2011b).

A Pol III duplica o DNA com alta fidelidade, inserindo um nucleotídeo errado a cada 10^5 a 10^6 incorporados (Yao & O'Donnell, 2010). E mesmo ao falhar, existe um mecanismo associado ao cerne catalítico com atividade de exonuclease de revisão 3' - 5' que possui a capacidade de correção, e que diminui a possibilidade de erro de incorporação para 10^{-7} (Yao & O'Donnell, 2010). Quando agregada ao cerne catalítico, o

beta-clamp assegura que a ligação entre o DNA e a Pol III continue íntegra, conferindo processividade à replicação (Yao & O'Donnell, 2010). Por fim, o *clamp loader* usa a energia da hidrólise do ATP para permitir que o *beta-clamp* desempenhe seu papel (O'Donnell, 2006); somado a esta atividade, o *clamp loader* ainda participa da organização do replissomo de uma forma que possibilite a ligação deste com a helicase DnaB durante a incorporação de novos nucleotídeos (Yao & O'Donnell, 2010).

A eficiência da replicação do DNA é, em grande parte, consequência da ação precisa da subunidade α – que possui um sítio catalítico compacto e especializado, e como resultado, este mecanismo é facilmente perturbado pela presença de lesões na fita molde de DNA (Yang & Woodgate, 2007; McHenry, 2011b). Com o objetivo de assegurar o término da duplicação do genoma, as células possuem outro grupo de polimerases, que diferentemente da Pol III especializada na duplicação do genoma, conseguem replicar o DNA usando como molde uma fita lesionada (Sale et al., 2012). A síntese translesão (TLS), que é consequência da ação deste grupo de polimerases especializadas propensas a erro, é um mecanismo de tolerância ao dano de DNA que impede o colapso da forquilha replicativa, quando a replicase não consegue prosseguir com a síntese de DNA (Sale et al., 2012).

1.2.2 – Duplicação do genoma

Na maioria das bactérias, a replicação possui uma origem única no cromossomo que é o ponto de partida para a formação de ambas as forquilhas que irão se mover em sentidos opostos ao replicar o DNA (Revisado por O'Donnell et al., 2013). Em *E. coli,* esta região é chamada de OriC, e neste microrganismo este processo é iniciado quando a proteína DnaA reconhece uma sequência específica, dando início à formação da bolha de replicação ao desenovelar a dupla hélice. Após este passo, a proteína DnaB é recrutada para esta bolha replicativa nascente e se desloca na fita descontínua para dar início a formação do Primossomo (DnaB - helicase + DnaG - primase). Esta estrutura é extremamente importante, uma vez que é responsável pela síntese do iniciador de RNA que é necessário para síntese de DNA da fita contínua e para os fragmentos de Okazaki de aproximadamente 4000 pares de base (pb) que são sintetizados na fita descontínua. Após a síntese do primeiro iniciador, a Pol III é formada (como descrito anteriormente) para começar a síntese de DNA (este mecanismo é revisto Robinson & Van Oijen, 2013).

Quando ligada ao terminal do iniciador, a Pol III replica o DNA com muita eficiência. Estudos de *single molecule* usando fluorescência mostram que sua taxa de replicação é de 500 - 1000 pb por segundo (Tanner et al., 2009; O'Donnell, 2006). Entretanto, a taxa de progressão do replissomo pode ser extremamente variada em diferentes organismos. Mesmo dentro do filo Bacteria, alguns representantes como *E. coli* replicam o DNA em uma velocidade maior quando comparados com outros, como *Caulobacter crescentus*, que replica seu genoma lentamente, com uma velocidade característica de um representante do filo Archaea (O'Donnell et al., 2013).

A replicação do DNA em todos os organismos só pode ocorrer no sentido 5' – 3' como consequência de uma característica das polimerases - que iniciam a replicação a partir de uma hidroxila no carbono 3' do açúcar, logo, apenas a fita *leading* pode ser sintetizada de forma contínua. A outra fita (descontínua) é sintetizada de forma concomitante à primeira, porém em pequenos fragmentos (Okazaki), no sentido oposto ao movimento da DnaB (Robinson & Van Oijen, 2013). Estudos sugerem que esta sincronia ao replicar ambas as fitas só é possível pela presença de uma pequena volta na fita usada como molde, formada dentro da bolha de replicação, que possibilita acesso 5' – 3' para ambos os cernes catalíticos da Pol III presentes no replissomo (Hamdan et al., 2009).

Durante o movimento das forquilhas em direções opostas, estas acabam encontrando sítios de terminação que estão localizados antes da finalização da duplicação do genoma (O'Donnell et al., 2013). Em *E. coli*, as regiões de terminação, chamadas de Ter, também são específicas. Estes sítios estão agrupados em 10 sequências (cinco em cada lado do *replicon*) que são reguladas pela proteína de ligação ao DNA Tus. Após esta interação entre Tus e a fita de DNA, uma barreira polar é formada, impedindo que o replissomo continue se movimentando (Dewar & Walter, 2017).

1.3 - MECANISMOS DE RESPOSTA AO DANO NO DNA (DDR)

Na presença de danos no DNA, as células procarióticas e eucarióticas selecionaram estratégias que restauram a sequência original do DNA com pouco ou nenhum dano a sua estrutura. O conjunto destes mecanismos é comumente chamado de reparo de DNA livre de erro. Geralmente, estes mecanismos estão envolvidos na reparação de lesões, e são recrutados de forma prioritária frente a estes danos para que replicação do DNA prossiga sem nenhum ônus celular. No entanto, nem sempre estes mecanismos são suficientes para

a correção destes erros, levando a célula para vias de tolerância ao dano do DNA, que diferentemente dos mecanismos de reparo livre de erro, evitam o efeito citotóxico das lesões à custa, frequentemente, do aumento da taxa de mutação (Friedberg et al., 2006).

Para organismos aeróbicos, o DNA está embebido em um ambiente rico em água e oxigênio o que facilita o aparecimento de lesões espontâneas (dano endógeno) (Friedberg et al., 2006). Além das lesões ocasionadas por componentes endógenos que podem lesionar esta macromolécula, alguns agentes químicos e físicos são responsáveis por interagir de forma negativa com o DNA, levando a modificações em sua estrutura. Dentre estes componentes genotóxicos, a radiação UV é responsável pela indução de várias lesões citotóxicas, como os dímeros ciclobutano de pirimidina (CPDs: *Cyclobutane Pyrimidine Dimers*) e os fotoprodutos (6,4) pirimidina – pirimidona (Pfeifer et al., 2005). Além destas lesões que são comumente formadas, a radiação UVC (que foi amplamente utilizada neste trabalho) pode levar ao aparecimento de outras lesões minoritárias como os dímeros de purina e monoadutos de pirimidona (Pfeifer et al., 2005).

Além da radiação, drogas antitumorais como a Mitomicina C (MMC) interagem negativamente com o DNA, formando várias lesões. No ambiente intracelular, a MMC sofre uma redução que leva a sua interação com resíduos de desoxiguanosina (dG) no DNA, que terá como consequência a formação de adutos MMC – dG (Weng et al., 2010). Além disto, esta droga também é capaz de atacar sítios de guanosina – guanosina (GG) o que pode acarretar em ligações cruzadas na mesma fita de DNA ou em fitas de DNA adjacentes (Weng et al., 2010). Estas lesões são potencialmente citotóxicas uma vez que servem de bloqueio tanto para replicação quanto para a transcrição (Weng et al., 2010).

Além destes agentes genotóxicos clássicos, hoje existe uma preocupação com um novo grupo de drogas, os antibióticos. Estas drogas vêm sendo utilizadas no controle de infecções bacterianas que acometem seres humanos, porém uma das consequências do seu uso indiscriminado é o aparecimento de linhagens resistentes, isso porque um dos alvos dos antibióticos é o DNA bacteriano. Os danos causados nesta molécula pelos antibióticos podem ser responsáveis por induzir uma resposta global (reposta SOS), o que ocasionará a expressão de um grupo de proteínas que vai processar este dano, podendo levar ao aumento das taxas de mutação, e desta forma conferir resistência a estas linhagens em ambientes estressantes (Qin et al., 2015).

1.3.1 – Bloqueio e colapso da forquilha replicativa

A persistência de lesões na fita molde de DNA pode bloquear a replicação quando estas distorções entram em choque com o replissomo. Isso acontece porque as polimerases responsáveis pela duplicação do genoma possuem características estruturais que as impossibilitam, quase sempre, de usar como molde uma fita danificada. Esta seletividade assegura que não acontecerá a incorporação de bases erradas durante o processo de elongamento da replicação. Esta especialização é responsável pela duplicação do genoma de forma eficaz e precisa (Yeeles et al., 2013).

A localização de lesões nas fitas contínua ou descontínua pode ter um papel determinante no mecanismo celular que será recrutado para que aconteça o processo de mitigação destes erros. Como a replicação na fita descontínua ocorre com a polimerização de pequenas regiões (fragmentos de Okazaki) pela Pol III, é mais provável que mecanismos de reiniciação da replicação ocorram preferencialmente em decorrência da descontinuidade deste processo, sendo desta forma a TLS um mecanismo mais propício a ser recrutado na fita contínua (Yeeles et al., 2013).

Já para lesões localizadas no molde usado para a replicação da fita contínua, os mecanismos de tolerância são mais complexos. Sabe-se que a presença de danos neste molde impede quase sempre a progressão do replissomo e para que este aparato prossiga, é necessário o *bypass* da lesão por polimerases especializadas que podem replicar o DNA usando como molde uma fita danificada. De forma alternativa, a correção destes danos também pode ser feita via reinício da síntese de DNA à frente desta lesão, seguido pelo reparo pós-replicação (Yeeles et al., 2013). Caso nenhum destes mecanismos seja recrutado, acontecerá o colapso da forquilha replicativa que é extremamente citotóxico.

1.3.2 – Sistema SOS

Ao estudar cepas de *E. coli* irradiadas com UV, Evelyn Witkin propôs, no final da década de 60 do século XX, a existência de uma via comum de genes que provavelmente era regulada por um repressor que seria inativado na presença de dano no DNA (Witkin, 1967). Em meados da década de 70, Radman complementou a descoberta de Witkin ao propor um "mecanismo de replicação celular que tendia ao aumento das taxas de mutação" – e para que isso acontecesse, era necessária a presença dos produtos dos genes *recA* e *lexA* (Radman, 1975). Hoje sabemos que a resposta SOS é o principal mecanismo

de resposta ao dano de DNA em procariotos, tendo como uma das suas consequências a mitigação do efeito citotóxico do dano através do aumento da taxa de mutação via TLS (Schlacher & Goodman, 2007).

Depois de 60 anos de descoberta (Witkin, 1967), hoje se sabe que em *E. coli*, em condições sem estresse, uma proteína repressora (LexA) reprime um grupo de ~ 40 genes (Courcelle et al., 2001) que estão envolvidos em vários processos celulares como inibição da divisão celular, reparo por excisão de nucleotídeos e TLS (Radman, 1975; Courcelle et al., 2001; Erill et al., 2007). Parte destes genes foi identificada pelo grupo de Graham Walker ao realizar uma varredura genética usando o fago Mud. Neste trabalho, eles inseriram este fago carregando o gene *lacZ* desprovido de promotor de forma aleatória no cromossomo de *E. coli*. As linhagens carregando este elemento foram tratadas com MMC e UVC, e tiveram os níveis de expressão dos genes fusionados a este sistema quantificados usando o ensaio de β -galactosidade. Como resultados, eles obtiveram um grupo de genes que eram induzidos por dano (*damage inducible*), sendo este fenótipo abolido em mutantes *recA* e *lexA* (Kenyon & Walker, 1980). Posteriormente, todo o regulon foi determinado em experimentos de microarranjos de DNA (Courcelle et al., 2001).

A proteína RecA, envolvida na resposta SOS, funciona como um sensor de dano ao DNA, ligando-se a regiões em fita simples (fsDNA), que não foram replicadas como consequência do bloqueio da polimerase replicativa frente ao dano. Ao ligar-se à fsDNA, a proteína RecA forma filamentos de nucleoproteína tornando-se ativada (RecA*). Quando o repressor LexA se liga a esse complexo (fsDNA+RecA*) é induzido a um processo de clivagem autocatalítica, que permite a expressão dos genes da resposta SOS, incluindo aqueles que codificam as polimerases capazes de usar uma fita de DNA danificado como molde para a replicação. Após o reparo ou o *bypass* das lesões, RecA torna-se inativa pela ausência de fsDNA (Schlacher & Goodman, 2007; Erill et al., 2007) (Figura 1).



Figura 1. Esquema da resposta SOS em *E. coli* descoberta por Witkin. A) Na ausência de dano no DNA o repressor LexA está ligado ao box SOS, impedindo a transcrição dos genes pela RNA polimerase. B)Na presença de um agente genotóxico, vai haver o acúmulo de fsDNA, que é reconhecido pela proteína RecA. Esta proteína começa a se ligar nestas regiões até se tornar uma nucleoproteína ativa. A ativação de RecA vai mediar a autoclivagem de LexA, e como consequência, seu desligamento do box SOS, permitindo a transcrição dos genes induzidos por esta resposta pela RNA polimerase.

1.3.3 - Diversidade de polimerases em procariotos e síntese translesão de DNA (TLS)

As polimerases podem ser divididas em seis famílias diferentes, com base na homologia entre suas sequências (Jarosz et al., 2007). A Família Y de polimerases reúne proteínas com homólogos em todos os domínios da vida; estas enzimas são especializadas em realizar TLS (Yeeles et al., 2013). Em *E. coli* existem cinco DNA polimerases, sendo três delas induzidas como parte da resposta SOS ao dano de DNA (Pol II, Pol IV e Pol V). Sabe-se que estas proteínas estão envolvidas na via TLS de tolerância ao dano neste organismo (Napolitano et al., 2010).

A Pol II de *E. coli*, codificada pelo gene *polB*, é uma proteína da família B de polimerases que realiza TLS (Wang & Yang, 2009). Nesta enzima estão conservados todos os motivos proteicos presentes nas replicases, assim como a atividade de exonuclease de revisão, o que é incomum para uma TLS-DNA-Polimerase (Wang & Yang, 2009). *In vitro* a Pol II é capaz de fazer o *bypass* de lesões como sítios abásicos, ɛC e adutos AAF (Al Mamun & Humayun, 2006; Becherel & Fuchs, 2001; Paz-Elizur et al., 1996). Foi relatado que *in vivo* a Pol II é responsável por aumentar a frequência de deleções (-2nt) ao sintetizar DNA sobre adutos AAF (Fuchs & Fujii, 2007).

DinB (Pol IV) é a única DNA polimerase da Família Y que é conservada em todos os domínios da vida, incluindo em humanos com a Pol k (Fuchs et al., 2004). Codificada pelo gene dinB, também conhecido na literatura como dinP (ambos aceitos), a Pol IV não possui atividade de exonuclease de revisão 3' – 5' e realiza TLS (Wagner et al., 1999). Jarosz e colaboradores (2006) mostraram que DinB participa do mecanismo de tolerância a danos causados por Nitrofurazona (NFZ) *in vivo* (adutos N^2 -dG), e atua em conjunto com a proteína UmuD'2C na sobrevivência à 4-hidroxiaminoquinolona-n-óxido (adutos de 4-NQO) em E. coli. Assim como para várias polimerases especializadas em TLS, o mutante dinB é viável em E. coli (Goodman & Woodgate, 2013). Em outras bactérias como C. crescentus e Mycobacterium tuberculosis, a expressão de dinB não é induzida por SOS (Galhardo et al., 2005; Rocha et al., 2008; Ordonez et al., 2014), e esta proteína não participa de mecanismos de tolerância ao dano aos agentes descritos para E. coli em C. crescentus (dados não publicados). No entanto em experimentos de mutagênese espontânea realizados recentemente por nosso grupo, os resultados mostraram que o mutante dinB desta bactéria apresenta menos frameshifts -1 espontâneos quando comparado com as linhagens selvagem e mutante imuC (Valencia, 2017). Este mesmo fenótipo já foi associado a proteína DinB em E. coli (Godoy et al., 2007).

A Pol V (UmuD'₂C) de *E. coli* é uma polimerase da Família Y, composta pelo produto gênico do operon *umuDC* regulado pela resposta SOS nesta bactéria (Erill et al., 2007; Ippoliti et al., 2012). Kato e Shinoura (1977), ao realizarem uma varredura genética mostraram que mutações no gene *umuC* resultam na ausência de mutagênese induzida por UV em *E. coli*. Um estudo posterior de Rajagopalan e colaboradores (1992) mostrou que o produto gênico dos genes *umuDC* estão envolvidos no *bypass* de lesões *in vitro*, porém neste trabalho eles acreditavam que estas proteínas, juntamente com a Pol III, faziam o *bypass* das lesões através de um complexo proteico chamado de Mutassomo. Na década de 90, os trabalhos de Reuven e colaboradores (1999), Tang e colaboradores (1999) e Bruck e colaboradores (1996) foram importantes para confirmar o papel dos produtos dos genes *umuDC* na tolerância de dano no DNA como polimerase capaz de realizar TLS em *E. coli* (Reuven et al., 1999; Tang et al., 1999), sendo que a proteína

Hoje sabemos que a Pol V de *E. coli* é composta de duas proteínas diferentes: o dímero UmuD'₂ e a proteína UmuC (Bruck et al., 1996). A proteína UmuD sofre uma autoclivagem estimulada por RecA após o início da resposta SOS, dando origem a UmuD', fundamental para interação com UmuC e a formação da Pol V (UmuD'₂C) (Ippoliti et al., 2012). Dentre as três polimerases envolvidas em mecanismos de tolerância ao dano de DNA em *E. coli* é correto afirmar que a Pol V é a responsável por fazer o *bypass* da maioria das lesões que podem acometer o genoma desta bactéria.

1.3.3.1 - Mecanismos de TLS mediado pelas polimerases Pol IV e Pol V

Existem algumas informações disponíveis na literatura sobre o mecanismo pelo qual a proteína DinB (Pol IV) faz o *bypass* de lesões em *E. coli*. Em 2007, o grupo de Graham Walker mostrou que o dímero UmuD₂ e a proteína RecA modulam a atividade desta polimerase em diversos substratos de DNA. Eles sugerem que as proteínas citadas anteriormente restringem o potencial mutagênico desta polimerase para *frameshifts* -1 ao tornar seu sítio ativo seletivo para o *bypass* de apenas algumas modificações na estrutura do DNA (Godoy et al., 2007). Neste trabalho, eles também sugerem que este controle vai ser mais relaxado em uma fase mais avançada da resposta SOS, e que DinB poderá ser mais mutagênico em fases crônicas desta resposta ao dano em decorrência desta regulação (Godoy et al., 2007). Em um trabalho mais recente, Mallik e colaboradores (Mallik et al., 2015) mostraram que após a indução da resposta SOS, as proteínas DinB e RecA fusionadas com marcadores fluorescentes formam um foco em lesões de quebra-dupla de DNA. Eles sugerem que a interação destas proteínas *in vivo* é importante para restaurar a forquilha replicativa e evitar o seu colapso (Mallik et al., 2015). Scotland e colaboradores (2015) sugeriram que ao superexpressar a proteína DinB, ela ganha acesso a forquilha replicativa (na ausência de danos no DNA), e este mecanismo é consequência de uma troca entre a Pol IV e a Pol III que resulta em queda da sobrevivência em *E. coli* (Scotland et al., 2015). Eles chamam atenção que a troca entre estas enzimas, mediada pela interação com o *beta-clamp*, é um mecanismo interessante que fornece pistas de como a TLS é regulada *in vivo* ao permitir o acesso a forquilha replicativa da proteína DinB mesmo na ausência de danos.

A Pol V é altamente propensa a erro ao replicar o DNA. Por volta da década de 80 do século passado os primeiros modelos de TLS frente à fotoprodutos para esta enzima surgiram na literatura (Bridges & Woodgate, 1985a; Bridges & Woodgate, 1985b). O primeiro deles, chamado de modelo *two-step* (dois passos) (Bridges & Woodgate, 1985b), recebeu este nome porque os autores acreditavam que a TLS de fotoprodutos acontecia em um processo que abrangia dois passos principais, sendo eles: A Pol III seria responsável por inserir um nucleotídeo de forma oposta ao primeiro (3') T (timina) do dímero T-T (consequência da radiação UV), já o segundo passo envolveria a aproximação da proteína RecA da região da lesão que irá possibilitar a interação da Pol III com as proteínas que compõem a Pol V para que em conjunto elas incorporem o próximo nucleotídeo (segundo (5') T que compõe o dímero T-T).

No começo da década de 90, um segundo modelo foi proposto por Echols & Goodman (Echols & Goodman, 1990) para a TLS da Pol V. Eles intitularam este trabalho como *Mutation induced by DNA damage: A many protein affair* (Mutação induzida por dano no DNA: um caso de várias proteínas). O título não poderia ser mais preciso para o que eles propunham neste trabalho. Neste modelo, eles acreditavam que a holoenzima Pol III ao encontrar uma lesão sofria um processo de desaceleração que culminava com o recrutamento e montagem de um complexo proteico chamado por eles de "Mutassomo" – este era composto por uma série de proteínas como RecA, UmuC, UmuD', Ssb, além da Pol III (com *beta-clamp* e *clamp loader* associados). O mutassomo seria responsável por fazer o *bypass* das lesões. Neste modelo, eles já levavam em conta que a proteína RecA era fundamental para que acontecesse a clivagem de UmuD para UmuD' (Bruck et al., 1996). Logo após a esta proposta, ficou claro que além da clivagem desta proteína,

era necessária a presença deste dímero (UmuD'₂) para caracterizar a Pol V como polimerase propensa a erro capaz de fazer o *bypass* de lesões independente da presença da Pol III: o complexo proteico UmuD'₂C purificado quando testado em ensaios de replicação mostra sinais claros de síntese translesão em *E. coli* (Bruck et al., 1999). Além disso, ficou provado que a Pol III não está associada à replicação propensa a erro neste organismo – uma vez que ela é responsável pela duplicação do genoma bacteriano livre de erro e com alta processividade (Yao & O'Donnell, 2010).

Hoje existem dois modelos de TLS para a Pol V aceitos na literatura (Figura 2), e ambos herdaram do modelo de Mutassomo de Echols & Goodman a noção de que a proteína RecA* (quando ativa) tem um papel chave neste mecanismo. Estes autores acreditavam que esta proteína possuía um papel fundamental na mutagênese, já que células deficientes em *lexA* e *recA* não eram mutáveis na presença de radiação UVC, mesmo com o produto dos genes $umuD^+$ e $umuC^+$ sendo expressos em altos níveis de forma constitutiva nestas linhagens (Blanco et al., 1982). O primeiro deles é o modelo chamado de "Trans – ativação da Pol V por RecA" desenvolvido pelo grupo de Myron Goodman em 2006 (Schlacher et al., 2006). A teoria deles é que existe um acúmulo de RecA* ativo em fsDNA em trans (que não está sendo replicada), que quando em contato com a Pol V (à jusante) torna-a ativa (Figura 2). Neste mesmo trabalho, foi testado um modelo alternativo para o acumulo de RecA* em cis que não teve sucesso para TLS. Eles explicam que esta teoria foi pensada com base na argumentação da Navalha de Occam (que explica eventos baseado somente nas premissas que são necessárias para a explicação do mesmo), ou seja, por parcimônia não teria porque RecA* se posicionar a montante da forquilha replicativa (ver modelo alternativo abaixo), já que esta ação poderia bloquear o progresso da Pol V.

O segundo modelo proposto em 2009 por Fujii & Fuchs (Fujii & Fuchs, 2009) assim como o modelo anterior, foca na importância da proteína RecA e do *beta-clamp* para que aconteça TLS mediada pela Pol V. Neste modelo eles fazem uma analogia para explicar a TLS, sendo que de forma figurada a Pol V seria um trem que desliza sobre um trilho que seria composto por proteínas RecA ativa para que o *bypass* de lesões aconteça. Além da nucleoproteína RecA, acredita-se que o *beta-clamp* tenha um papel importante para estabilizar a ligação da fsDNA molde com a Pol V que desliza sobre moléculas de RecA* em *cis*, e as retira do DNA para fazer TLS.



Figura 2. Participação de RecA no processo de TLS mediado pela DNA Pol V. Figura esquemática dos modelos de TLS vigentes para a Pol V de *E. coli*. A) Autoclivagem da proteína UmuD estimulada pela presença de RecA ativa. B) Modelo do Mutassomo proposto pelo grupo de Goodman. Nesta figura apresentamos o modelo vigente em que se acredita que a Pol V necessita da presença de RecA ativa em *trans* para o fazer o *bypass* de lesões que não são replicadas pela Pol III – Eles nomearam este mecanismo como Mutassomo porque além da polimerase propensa a erro, envolve um grande conjunto de proteínas que estão implicadas na replicação livre de erro e que são necessárias para TLS. C) Modelo proposto por Fujii & Fuchs também chama atenção para a necessidade da interação de RecA ativa com a Pol V para que aconteça o *bypass* da lesão. Neste modelo eles fazem a analogia ao trilho (nuclefilamento de RecA ativa) e o trem (Pol V) para explicar como acontece a síntese translesão em *E. coli*.

Como descrito anteriormente, a Pol IV e Pol V foram estudadas exaustivamente nas últimas décadas, uma vez que são as polimerases responsáveis pela síntese translesão em *E. coli* (organismo modelo da biologia molecular). Porém, algumas bactérias não possuem o operon *umuDC* que codifica a Pol V, e utilizam uma outra maquinaria para fazer o *bypass* das lesões (McHenry, 2011a).

1.3.4 - Algumas bactérias utilizam outra Pol III para fazer TLS: o cassete mutagênico *imuABC* (*dnaE2*)

Algumas bactérias possuem uma cópia extra do gene *dnaE*, chamada de *dnaE2*, que codifica uma subunidade α "alternativa" que não está envolvida na duplicação do genoma, e sim no *bypass* de lesões que bloqueariam a síntese de DNA feita pelo produto do gene *dnaE* em microrganismos (McHenry, 2011b). Bactérias que usam a proteína DnaE2 para fazer TLS geralmente não possuem ortólogos de *umuDC* (McHenry, 2011a).

A mutagênese induzida por dano dependente de DnaE2 foi caracterizada primeiro em *Mycobacterium tuberculosis* e *Caulobacter crescentus* (Boshoff et al., 2003; Galhardo et al., 2005). Boshoff e colaboradores observaram que em *Mycobacterium* a segunda cópia do gene *dnaE* (*dnaE2*) é regulada *in vitro* por vários agentes genotóxicos, e a perda desta proteína reduz a sobrevivência à radiação UV e a virulência em camundongos (Boshoff et al., 2003). Este grupo foi o primeiro a associar esta proteína como um importante fator que contribui para o surgimento de mutações de resistência a agentes antibacterianos (Boshoff et al., 2003).

Em 2005 nosso grupo mostrou que o gene *dnaE2* (que também pode ser chamado de *imuC* por ser o terceiro gene do operon *imuABC*) faz parte do cassete mutagênico *imuABdnaE2* induzido por dano em *C. crescentus*, que é responsável pela mutagênese induzida por MMC, e participa de uma das vias de mutagênese induzida por UV neste organismo (Galhardo et al., 2005); o prefixo *"imu"* usado para identificar estes genes é uma referência ao termo *inducible mutagenesis* (mutagênese induzível). Nossos dados indicam que estes genes estão na mesma via de epistasia, uma vez que mutantes simples dos genes *imuA*, *imuB* e *imuC* (*dnaE2*), bem como combinações destas deleções gênicas, inibem a mutagênese induzida pelos agentes descritos anteriormente (Galhardo et al., 2005). Vale a pena ressaltar a importância e o pioneirismo deste trabalho ao associar a ação conjunta destes genes como resposta ao dano no DNA. McHenry em 2011 (McHenry, 2011b) chamou atenção para a grande dispersão destas proteínas em genomas

bacterianos sequenciados até o momento da publicação do seu trabalho e, como é importante que mais estudos se concentrem neste cassete mutagênico como uma forma de entender o mecanismo de ação responsável pela mutagênese induzida por dano mediado por estas proteínas, inclusive para tentar contornar os problemas de resistência bacteriana aos antibióticos. Seguindo a sugestão deste autor, utilizaremos neste trabalho o nome *imuC* para esta segunda cópia de *dnaE*.

O cassete mutagênico regulado pela resposta SOS descoberto por Galhardo e colaboradores é altamente difundido em bactérias (Galhardo et al., 2005). A proteína ImuA é a que possui a sequência menos conservada, com similaridade com as proteínas ImuA' de M. tuberculosis, assim como RecA e SulA (Warner et al., 2010; Galhardo et al., 2005). Em E. coli, RecA está envolvida no reparo por recombinação homóloga, indução da resposta SOS, processamento da proteína UmuD' e associação com a Pol V para fazer o bypass de lesões, enquanto que SulA atua no controle da divisão celular (Altschul et al., 1997; Galhardo et al., 2005). No entanto, o papel de ImuA na tolerância ao dano de DNA em organismos que possuem este operon mutagênico ainda é desconhecido. A proteína ImuB possui similaridade com proteínas da Família Y de polimerases (Galhardo et al., 2005), e possui todos os domínios proteicos típicos desta família que é especializada em TLS. Mudanças em regiões da sequência que codificam resíduos catalíticos conservados em polimerases desta família sugerem que esta proteína não é funcional como polimerase (Warner et al., 2010). Por outro lado, Warner e colaboradores observaram que a proteína ImuB de Mycobacterium interage com o beta*clamp* deste organismo, associação que não é observada entre ImuC e β –*clamp*. Neste mesmo trabalho eles mostraram que ImuB também é capaz de interagir com a proteína ImuC (Warner et al., 2010). Com estas informações em mãos, este grupo criou a hipótese de que ImuB possui um possível papel acessório/regulatório na síntese translesão catalisada pela proteína ImuC (Galhardo et al., 2005, Warner et al., 2010).

Em bactérias do gênero *Pseudomonas*, este cassete mutagênico também está presente. Em *P. putida*, o produto gênico do gene *imuC* não possui efeito significativo no aumento de deleções de 1 par de base (Koorits et al., 2007), enquanto a proteína ImuB parece estar envolvida no acúmulo destas mutações na fase fase estacionária (Koorits et al., 2007), o que demonstra um papel antagônico destas proteínas. Em *Pseudomonas aeruginosa*, Cirz e colaboradores identificaram em experimentos de larga escala genes que são induzidos por dano, e dentre eles estão o que codifica a polimerase DinB (Pol IV), e o operon *imuABC*, e alertaram para a importância destes genes em mediar o

aumento da taxa de mutação e como consequência favorecer o aumento dos casos de resistência bactéria neste organismo (Cirz et al., 2006). Mais recentemente nosso grupo também vem investigando o papel da resposta SOS e destas polimerases na mutagênese induzia por dano que contribui para a resistência a antibióticos (Valencia et al., 2017).

1.4 - Caulobacter crescentus: UM MODELO ALTERNATIVO PARA BIOLOGIA CELULAR E GENÉTICA

O avanço dos estudos em biologia molecular só foi possível graças ao progresso na compreensão de um microrganismo extremamente popular nos meios acadêmicos: *E. coli.* Esta bactéria se tornou modelo para uma série de fenômenos que explicam de forma complexa e precisa mecanismos de regulação gênica, replicação do DNA e mecanismos de reparo de DNA e tolerância ao dano (Friedberg et al., 2006). Porém, mesmo sendo um excelente organismo modelo, esta bactéria não consegue explicar como alguns destes mecanismos funcionam em outros microrganismos.

Buscando suprir estas lacunas, outras bactérias têm sido usadas como modelo para diferentes áreas de estudo da Biologia. Um exemplo disso é *C. crescentus*, pertencente ao grupo Caulobacterales, descrito em 1935 por Henricini e Johnson (Henricini & Johnson, 1935). Neste trabalho, os autores chamam atenção para a morfologia única deste grupo, o que levou a criação de uma nova ordem taxonômica. Dentro desta encontra-se *C. crescentus*. Hoje esta é caracterizada como uma α -proteobacteria, Gram-negativa, de vida livre, que possui vários atrativos fisiológicos e funcionais que a tornam um excelente modelo para a biologia molecular. Este microrganismo possui um ciclo celular com divisão assimétrica, que como consequência, gera dois tipos celulares morfologicamente distintos. Apenas as células fixas, chamadas de talo, conseguem replicar o seu DNA e prosseguir a divisão celular, gerando ciclicamente dois tipos celulares: células flageladas e talo (Collier, 2012). Esta sua característica particular chamou atenção de vários grupos de pesquisa ao redor do mundo que a utilizam para entender o controle do ciclo celular em bactérias e mecanismos de replicação de DNA.

C. crescentus também é um organismo extremamente interessante quando se trata de genética. A holoenzima DNA Polimerase III ainda não foi caracterizada nesta bactéria, porém, análises comparativas entre os genomas sequenciados e depositados em bancos de dados até o momento mostram que a subunidade θ de *E. coli*, codificada pelo gene *holE*, que tem sua função associada com a estabilização da subunidade ε (*dnaQ*) no cerne catalítico e ligação dos componentes da replicase (Taft-Benz & Schaaper, 2004), está confinada somente no grupo das enterobacterias (Dietrich et al., 2014). Logo, estudos usando como modelo outra Pol III são limitados para o entendimento do funcionamento desta holoenzima para a síntese de DNA livre e propensa a erro em procariotos bacterianos (Jensen et al., 2001).

Uma série de ferramentas genéticas já foi desenvolvida para *C. crescentus* para facilitar a manipulação genética deste organismo (e. g. Thanbichler et al., 2007). Estas técnicas podem consolidar esta bactéria como um modelo genético alternativo às enterobactérias.

Em 2010, Marks e colaboradores sequenciaram o genoma de uma linhagem derivada da CB15 (primeira linhagem de *C. crescentus* que teve seu genoma sequenciado em 2001 por Nierman e colaboradores). Esta linhagem foi chamada de NA1000 e o seu genoma apresenta 4.016.942 pb que codificam 3.767 genes que estão espalhados em um cromossomo único circular (sua sequência completa pode ser acessada no link a seguir: http://www.genome.jp/dbget-bin/www_bget?gn:T00841). Este genoma (Marks et al., 2010), será usado como referência ao longo deste trabalho.

Hoje sabemos que o cassete mutagênico *imuABC* é amplamente disperso na natureza, e a presença destes genes em genomas bacterianos é mais frequente que o operon *umuDC*. A compreensão de particularidades de seu mecanismo de ativação pode ser extremamente elucidativa para compreender como a maioria das bactérias faz o *bypass* de lesões e toleram a interação com agentes genotóxicos. Logo, estudar o funcionamento deste mecanismo é fundamental para contribuir com informações que permitam focar em novos alvos moleculares que nos permitam ampliar as ferramentas na luta contra a resistência bacteriana que utilizam polimerases diferentes da Pol V.

Além de genes envolvidos em TLS, nosso grupo tem se esforçado para caracterizar outros genes que fazem parte da resposta SOS em *C. crescentus* (Rocha et al., 2008). Como resultado, em 2015 Lopes-Kulishev e colaboradores (Lopes-Kulishev et al., 2015) mostraram que o produto do gene *mmcB*, induzido por SOS, é necessário para que aconteça a mutagênese induzida por mitomicina C mediada por ImuABC (Galhardo et al., 2005; Lopes-Kulishev et al., 2015). O fenômeno de mutagênese induzida por MMC é inibido na ausência de *mmcB* (que é induzido pela resposta SOS). O mesmo não é observado para danos causados por luz UV. Esta proteína processa o dano induzido por este agente genotóxico. permitindo acesso da maquinaria de TLS (ImuABC). A mutagênese induzia por UV, é independente de processamento por MmcB e parcialmente

dependente de ImuABC. Outra parte do nosso grupo está concentrada em desvendar este componente independente de ImuC e dependente de SOS para a mutagênese induzida por UV. Somado a estes esforços, recentemente publicamos um trabalho mostrando que em *C. crescentus* a TLS mediada por ImuABC não depende da presença de RecA (Alves et al., 2017) – apêndice 8. Estes dados serão descritos ao longo desta tese.

Em 2012, Bos e colaboradores chamaram atenção para um mecanismo de resposta ao dano de DNA presente em *C. crescentus* que é inteiramente independente da ação de TLS-DNA-Polimerases presentes neste organismo. A proteína BapE é induzida em consequência da resposta SOS nesta bactéria e desempenha um mecanismo apoptose-*like* que ao se deparar com danos no DNA não mitigados por mecanismos de reparo livre de erro, direciona a célula para uma morte-programada mediada pela endonuclease BapE (Bos et al., 2012).

Além da resposta SOS ao dano de DNA, outros pesquisadores que estudam *C. crescentus* identificaram outro regulon que é também induzido pela presença de lesões de DNA (Modell et al., 2014). Neste trabalho eles caracterizaram um novo inibidor da divisão celular, produto do gene *didA* que é responsável pela filamentação das células na presença de dano no DNA. A proteína DidA se liga à FtsN levando ao bloqueio da citocinese. A regulação deste gene é feita por DriD de forma independente de LexA/RecA (Modell et al., 2014). Este mesmo grupo em 2011 (Modell et al., 2011) caracterizou a proteína SidA que é induzida em consequência da resposta SOS em *C. crescentus*. O produto do gene *sidA* é um inibidor da divisão celular que irá atrasar a progressão do ciclo celular na presença de dano no DNA (Modell et al., 2011).

Mesmo não sendo uma bactéria oportunista e/ou patogênica; *C. crescentus* vem recebendo atenção na comunidade científica ao redor do mundo por ser um excelente modelo para estudar a replicação de DNA, regulação da expressão gênica e controle do ciclo celular. Somada a este repertório de técnicas, a presença de genes que codificam as proteínas ImuABC em seu genoma tornam esta bactéria extremamente interessante para prosseguir os estudos de TLS em bactérias que utilizam modelos diferentes da Pol V.

2 OBJETIVOS

2.1 - OBJETIVO GERAL

• Caracterizar o mecanismo molecular e o padrão de mutagênese da síntese translesão de DNA mediada por ImuABC em *C. crescentus*.

2.1.1 – Objetivos específicos

Objetivando compreender a distribuição de mutações em genomas expostos a radiação UV, propõem-se:

Analisar genomas de isolados das cepas NA1000 (wt) e *imuC* (deficiente em TLS) expostos à luz UV. Desta forma, pretendemos estudar a mutagênese induzida por UV, tanto em seu componente dependente de *imuABC*, quanto aquele independente destes genes. Procuraremos comparar os tipos de mutação, bem como sua localização (posição no cromossomo, preferência por fita transcrita ou não transcrita, preferência por fita *lagging* ou *leading*).

Buscando avaliar se a mutagênese dependente de *imuABC* possui um componente não-induzido, ou seja, se ela acontece de forma independente de lesões, propomos:

- Utilizar cepas que possuem a expressão constitutiva nos níveis da reposta SOS dos genes envolvidos em TLS em *C. crescentus*, para verificar se existe a presença de um fenótipo mutador espontâneo associado a mutagênese *untargeted* (não direcionada a danos no DNA).
- Utilizar estas cepas como *background* para deleção do gene *recA*, para entender se existe dependência do produto deste gene para TLS realizada por ImuC.

3 MATERIAL E MÉTODOS

3.1 - CEPAS BACTERIANAS, PLASMÍDEOS E CONDIÇÕES DE CRESCIMENTO DAS LINHAGENS UTILIZADAS NESTE TRABALHO

As cepas de *C. crescentus* foram crescidas em meio PYE (Peptona 2 g/L; Extrato de levedura 1 g/L; MgSO₄7H₂O 0,2 g/L; CaCl₂ 0,5 mM; 15 g/L de ágar foram adicionados para obtenção de meio sólido) a 30° C em agitação constante. Quando necessário, o meio foi suplementado com gentamicina 0,5 μ g/mL para caldo e 5 μ g/mL para meio-sólido, estreptomicina 10 μ g/mL, espectinomicina 25 μ g/mL e canamicina 5 μ g/mL. Cepas de *E. coli*, usadas como intermediárias de clonagem para a construção dos mutantes e para ensaios de expressão proteica, foram crescidas em meio LB (Peptona 10 g/L; Extrato de levedura 5 g/L; NaCl 10 g/L; sendo 15 g/L de ágar adicionados para obtenção de meio sólido) em agitação constante à 37° C. Quando necessário para adaptações experimentais, linhagens de *E. coli* também foram crescidas a 16° C.

Tabela 1. Linhagens de C. crescentus e E. coli usadas neste trabalho e seus				
respectivos genótipos				

		(Continua)
Linhagens	Genótipo	Referências
C. crescentus		
		Evinger &
NA1000 (RSG 470)	Linhagem parental de C. crescentus	Agabian,
		1977
$AimuC(\mathbf{DSC} 474)$	NA 1000 Aimu CuSpace	Galhardo
$\Delta mu C (KSG 474)$	NA1000 Δ <i>imu</i> C::Spec ⁻¹	et al., 2005
		Construída
PimuA (RSG 536)	NA1000 imuA::pPimuA	por MAL
		Noronha
		Construída
PimuAo ^c (RSG 557)	NA1000 imuA::pPimuAo ^c	por MAL
		Noronha
ΔrecA PimuA (RSG	NA 1000 imuA pPimuA AreaA	Este
106)	$11A1000$ muApr muA $\Delta recA$	trabalho

Δ <i>recA</i> PimuAo ^c	$NA1000 imuA \cdots pDimuAo^{c} AracA$	Este
(RSG 321)	21)	
E. coli		
	F- mcrA Δ (mrr-hsdRMS-mcrBC)	
DH10.8	Φ 80dlacZ Δ M15 Δ lacX74 endA1 recA1 deoR	Hanahah,
Diriop	Δ (ara,leu)7697 araD139 galU galK nupG rpsL	1983
	λ-	
S17	294:RP4-2(Tc:Mu)(Km:Tn7)	Simon et
~		al., 1983
S17pDnaE2del	S17 pDnae2del	Galhardo
	1	et al., 2005
BL21 (DE3) /PGro7	BL21: F ⁻ , $ompT$, $hsdS_B(r_B - m_B -)$, gal, dcm	Nishihara
× ,		et al., 1998
Rosetta™E. coli	F- ompT hsdSB(rB- mB-) gal dcm pRARE	Novagen
	(Cam ^R)	U
	<i>fhuA2</i> [lon] ompT gal (λ DE3) [dcm] Δ <i>hsdS</i> λ	
	$DE3 = \lambda sBamHIo\Delta EcoRI-B$	
	int::(<i>lacI</i> ::PlacUV5::T7 gene1) i21∆ <i>nin5</i> F-	Studier &
	mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80dlacZ Δ M15	Moffatt,
BL21 (DE3)DH10 β	Δ lacX74 endA1 recA1 deoR Δ (ara,leu)7697	1986Hana
	araD139 galU galK nupG rpsL λ-	hah, 1983
S17	294:RP4-2(Tc:Mu)(Km:Tn7)	Simon et
		al., 1983
S17pDnaE2del	S17 pDnae2del	Galhardo
		et al., 2005
BL21 (DE3) /PGro7	BL21: F ⁻ , <i>omp</i> T, <i>hsdS</i> _B ($r_B^- m_B^-$), <i>gal</i> , <i>dcm</i>	Nisninara
	$\mathbf{E} = \operatorname{omn} \mathbf{T} \operatorname{hed} \mathbf{S} \mathbf{D} (\mathbf{r} \mathbf{D} - \mathbf{m} \mathbf{D})$ coldom $\mathbf{n} \mathbf{D} \wedge \mathbf{D} \mathbf{E}$	et al., 1998
Rosetta TM	Γ - omprinsus $\sigma(\Gamma D - \Pi D -)$ gal ucin pKARE	Novagen
	(Calli)	
BL21 (DE3)		
()		

fhuA2 [lon] ompT gal (λ DE3) [dcm] $\Delta hsdS \lambda$ Studier &DE3 = λ sBamHIo Δ EcoRI-BMoffatt,int::(*lacI*::PlacUV5::T7 gene1) i21 $\Delta nin5$ 1986

O estudo da síntese translesão em C. crescentus só é possível pela descoberta dos genes imuABC envolvidos neste mecanismo (Galhardo et al., 2005). Para este trabalho, reconstruímos a linhagem mutante imuC::Spec^R no background de outra linhagem selvagem - NA1000 (fornecida pela professora Marilis Marques). A reconstrução foi necessária após a detecção de uma mutação de aproximadamente 30.000 kpb presente na linhagem NA1000 (parental) do nosso laboratório, usada para a construção deste mutante anteriormente. Este processo foi feito através da conjugação da linhagem de E. coli S17 que já havia sido previamente transformada com o plasmídeo pDnaE2del (Tabela 2) (Galhardo et al., 2005) com a linhagem parental nova. Após a conjugação, selecionamos o primeiro evento de recombinação com canamicina e ácido nalidíxico. Para seleção do segundo evento, que retira o plasmídeo suicida integrado no cromossomo de C. crescentus, usamos seleção em PYE-ágar enriquecido com 3% de sacarose. Para confirmação da mutação, colônias individualizadas foram testadas por PCR com oligos iniciadores dnaE2CasseteSpecFwd e dnaE2CasseteSpecRev que flanqueiam o cassete de espectinomicina que interrompe o gene *imuC* (Tabela 3). Células mutantes amplificam um fragmento de 2,5 Kb (2.000 pb do cassete de espectinomicina + 500 pb do gene *imuC*) enquanto que células selvagens amplificam apenas os 500 pb do gene *imuC* sem o cassete.

Tabela 2. Conjunto de vetores utilizados neste trabalho				
Plasmídeos	Genótipo	Referências		
pBC	Vetor de clonagem clo ^R	Juran & Lazaridis, 2008		
pBCimuAxy l	pBC:: <i>imuAxyl</i>	Este trabalho		
pMCS7	Vetor de integração	Thanbichler et al., 2007		
pRECDEL	Deleção em fase do gene <i>recA</i> clonado no pNPT138	Galhardo et al., 2005		

Tabela 2. Conjunto de vetores utilizados neste trabalho
pPimuA	Fragmento do gene <i>imuA</i> clonado no vetor pMCS7	Este trabalho
pPimuAo ^c	Fragmento do gene <i>imuA</i> com a mutação o ^c clonado no vetor pMCS7	Este trabalho
pDnaE2del	Fragmento do gene <i>dnaE2</i> interrompido pelo cassete Spe ^R no sítio de EcoRI, clonado no vetor pNPTS138	Galhardo et al., 2005
pET28-b	Tag de His na região N- terminal/Trombina/Tag para T7 + sítio adicional de His na região C-terminal	Merck

Para a construção de cepas que expressam de forma constitutiva os genes do operon *imuABC*, introduzimos a mutação o^c (GA \rightarrow CG) por mutagênese sítio dirigida na região promotora do gene *imuA* (Rocha et al., 2008), após este passo clonamos o fragmento com a mutação no plasmídeo pMCS7.

As linhagens controle (P*imuA*) e a que expressa constitutivamente o produto do operon *imuABC* (P*imuA*o^c) possuem uma pequena duplicação que consiste na região promotora do gene *imuA* e parte da CDS deste gene, sendo o diferencial da linhagem P*imuA*o^c a mutação o^c que confere o fenótipo descrito anteriormente. Os insertos de DNA com estas características foram clonados em pMCS7 e posteriormente transferidos por transformação para *E. coli* S17 para serem conjugados com *C. crescentus*. O evento de recombinação que integrou o plasmídeo ao cromossomo bacteriano foi selecionado em meio contendo gentamicina. A mutação o^c na região promotora do gene *imuA* impede a ligação do repressor LexA na região do operador, o que leva a expressão do produto do operon *imuABC* mesmo na ausência de resposta SOS (Rocha et al., 2008). Estas construções foram feitas pelo então aluno de iniciação científica do nosso laboratório Marco Noronha. (Mais detalhes destas construções podem ser observados na Figura 4).

		(Continua)
Oligo	Sequência 5' - 3'	Característica e função
imuAfwd	CATATGATGGAAGCGGGG	Amplificação do gene <i>imuA</i> para
	ACTCGGAC	clonagem no pET28-b
imuArevTH	GGATCCGATTTATCCGAA	Amplificação do gene imuA para
	GCGTCGTC	clonagem no pET28-b
imuApxylfw	AAAGAATTCCTTGCGGCC	Amplificação do gene imuA para
d	CTGAGAGG	clonar no pNPT228XNE
imuApxylre	TTTACTAGTTTATCCGAA	Amplificação do gene imuA para
v	GCGTCGTC	clonar no pNPT228XNE
recAfwd	AGTGAATTCATGACAAGT	Confirma a deleção do gene <i>recA</i>
	CAGGCGGCTTT	
recArevTH	CAGCTCGAGCTAGTCCTC	Confirma a deleção do gene recA
	TTCGCCCTCTT	
pimuAocfw	CATATGTCCGGACGGCGA	Amplificação de imuA com a
	TCTC	mutação o ^c
pimuAocrev	CATATGACCACGCCAAAG	Amplificação de imuA com a
	CCGCCG	mutação o ^c
PimuAmutf	CAGAACAAAAGTG <i>CG</i> AC	Mutagênese sítio-dirigida
OW	ATGGAGTTGGG	
PimuAmutr	CCCAACTCCATGTCGCAC	Mutagênese sítio-dirigida
ev	TTTTGTTCTG	
pimuAocfw2	GTTCATATGCGTAGATGG	Amplificação de <i>imuA</i> com a
	CCACGCTCT	mutação o ^c
pimuAocrev	GTTCATATGCTTCTCGCA	Amplificação de <i>imuA</i> com a
2	GGCCAGTTG	mutação o ^c
Oc check1	GAAGATCAGAACAAAAG	Detecção de recombinantes com a
	TGCG	mutação o ^c
Oc check2	GACAGGGTTTCAGCCTCA	Detecção de recombinantes com a
	TC	mutação o ^c
rpoDfwd	GAAGAACTGGCCGAAAA	RT-PCR (rpoD)
	GCT	

Tabela 3. Conjunto de oligos usados como iniciadores neste trabalho e suas funções

rpoDrev	CGTTCTTGTCCTCGATGA	RT – PCR (rpoD)
	AGTC	
imuBfwd	TCACCACCTCGATGCCGA	RT - PCR (<i>imuB</i>)
imuBrev	TGGCCAGGCTGGTCAAGC	RT - PCR (<i>imuB</i>)
		(continuação)
rpoBfow2		Amplifica região do <i>rpoB</i> que
	ACGCGCCATCAAGGAACG CA	abriga as mutações Rif ^R
rpoBrev	TGGACTGGGCGATGACGT	Amplifica região do <i>rpoB</i> que
	GC	abriga as mutações Rif ^R
dnaE2Casse	CGCTGGGCATGCTGCACG	Confirmação de mutantes
tteSpec	С	dnaE2::spec
Fwd		
dnaE2Casse	AGGCTCATGGCCTGCTCC	Confirmação de mutantes
tteSpec	TGG	dnaE2::spec
Rev		

Usando como background genético as cepas P*imuA* e P*imuA*o^c, usamos o plasmídeo pRECDEL (derivado de pNPTS138) (Galhardo et al., 2005) para deletar o gene *recA* nestas linhagens. Este foi introduzido nestas linhagens por eletroporação. Os eventos de recombinação foram selecionados primeiro em meio PYE com estreptomicina e espectinomicina, para ter certeza da inserção do plasmídeo no cromossomo. Seguido de um segundo evento de recombinação selecionado em PYE-ágar + 3% de sacarose que objetivava selecionar mutantes que perderam o plasmídeo. Para confirmação da deleção deste gene, fizemos PCR das colônias que cresceram a seleção por sacarose e selecionamos aqueles com o tamanho esperado para a deleção do gene *recA* (Tabela 3).

Para a construção do plasmídeo *imu*ApET28b usado para a expressão da proteína ImuA em *E. coli*, amplificamos o gene *imu*A flanqueado por sítios de NdeI e BamHI (iniciadores *imu*Apxylfwd e *imu*Apxylrev que estão descritos na Tabela 3) que permitiu a clonagem deste fragmento em fase com a cauda de histidina C- terminal do vetor pET28b (gentilmente cedido pelo Professor Márcio Dias). Antes de clonar no vetor final, usamos o vetor pBC como intermediário da clonagem: ligamos o inserto contendo o gene *imu*A purificado com o pBC. O resultado da ligação (T4 DNA ligase – Thermo Fisher Scientific, por 16 horas à 4° C) foi transformado em células DH10β. Os clones positivos, previamente selecionados pela coloração azul e branca foram testados por amplificação do gene *imuA*. Para a clonagem no vetor final, três dos clones positivos em pBC foram digeridos com NdeI e BamHI, assim como o plasmídeo pET28-b. O produto da digestão do plasmídeo pBC*imuA* foi purificado e clonado no plasmídeo final. Um novo ciclo de transformação química foi feito e possíveis clones foram testados por amplificação. Como atestamos a presença do gene *imuA*, sequenciamos o inserto para nos certificar da integridade da sua sequência de DNA.

3.2 – TÉCNICAS DE BIOLOGIA MOLECULAR UTLIZADAS EM CLONAGEM

Todas as manipulações genéticas utilizadas para a construção de novas linhagens descritas neste trabalho utilizaram as técnicas descritas a seguir.

3.2.1 – Reação em cadeia da Polimerase (PCR) e Reação de Ligação

Utilizamos a reação em cadeia da polimerase (PCR) para amplificação de genes de interesse e validação de clonagem. Para PCR que visava amplificar insertos gênicos usamos a enzima de alta fidelidade Phusion (BioLabs) de acordo com o protocolo sugerido pela empresa. Para PCR que tinha o objetivo de checar clonagem quanto a presença de insertos positivos, assim como para a amplificação de genes usados para a construção de espectros de mutação utilizado neste trabalho, utilizamos a enzima Taq DNA polimerase (Sigma) de acordo com as instruções do fabricante. A temperatura de anelamento dos primers com o DNA foi utilizada de acordo com as características particulares das sequências de cada par de iniciadores utilizado neste trabalho.

Todas as reações foram acrescidas de DMSO e para facilitar a interação destas polimerases com o DNA rico em GC de *C. crescentus*.

Após a amplificação do inserto de DNA para a clonagem ele foi submetido a um passo adicional de ligação a um vetor de clonagem e posteriormente ao vetor final. Esta ligação foi feita com a enzima T4 Ligase (Sigma) de acordo com as recomendações do fabricante. Para melhorar a eficiência de ligação, mudamos a temperatura de incubação da reação para 4º C (geladeira) por 16 horas.

Os plasmídeos obtidos destas reações de ligação foram extraídos com ajuda do KIT da Promega *Wizard*® *Plus SV Minipreps DNA Purification System*.

3.2.2 – Preparação de células quimio-competentes para clonagem

Neste trabalho as células competentes foram preparadas de acordo com o protocolo proposto por Sambrook & Russell (2006) com algumas modificações.

As células foram recuperadas do freezer -80° C em placas de LB sólido. No dia seguinte, uma colônia isolada foi colocada para crescer em 2 mL de LB caldo por 2 horas em agitação constante à 37° C em um tubo *snap cap* de 13 mL descartável. Após este passo, o pré-cultivo foi transferido para um erlenmeyer de 250 mL e foi acrescido de 500 μ L de MgCl₂ 1 M e 47,5 mL de meio LB líquido pré-aquecido a 37° C. Este cultivo foi deixado agitando à 37° C até atingir a DO₅₅₀ 0,5 (aproximadamente 3 horas).

Enquanto as células chegavam na concentração desejada, 2 tubos tipo Falcon de 50 mL foram deixados imersos em gelo e a centrífuga foi colocada para resfriar a 4º C para o prosseguimento do protocolo.

Com as células na concentração desejada (DO₅₅₀ 0,5), elas foram transferidas para os dois tubos previamente resfriados, e deixadas imersas em gelo por 15 minutos. Após este passo, os tubos foram centrifugados à 3000 g por 12 minutos à 4° C e o sobrenadante foi descartado com cuidado para não perder o sedimento de células e ressuspendido gentilmente com a ajuda de uma micropipeta (1 mL) na solução RF I (50 mL de KCl 1 M; 25 mL MnCl₂.4H₂O 1 M; 30 mL de acetato de potássio 0,5 M pH 6,9; 5 mL de CaCl₂.H₂O 1 M; 75 mL de glicerol. Completar para o volume de 500 mL com água MilliQ).

As células voltaram para o repouso em gelo por mais 15 minutos. Este passo foi seguido de uma nova centrifugação à 3000 g por 12 minutos e o sedimento celular foi ressuspendido em 2 mL da solução RF II (10 mL de Na-MOPS 0,5 M pH 7; 10 mL de KCl 0,5 M; 37,5 mL de CaCl₂.2H₂O 1 M; 75 mL de glicerol; completar o volume com H2O MilliQ para 500 mL) e aliquotadas em tubos de 1,5 mL para serem guardadas no freezer -80° C até o momento de serem utilizadas.

3.2.3 – Transformação de células competentes

As células foram transformadas com o DNA de interesse seguindo os passos a seguir. Antes da transformação a temperatura do banho maria foi ajustada para 42° C e o DNA plasmidial utilizado para a transformação foi retirado do freezer -20° C e deixado

em repouso em gelo para que descongelasse. Após estes passos, as alíquotas com as células competentes foram retiradas do freezer -80° C para dar início a a transformação.

Para cada ensaio de transformação, foi utilizado 50 μ L de célula competente e 5 μ L de DNA plasmidial. Estes foram misturados em um tubo eppendorf e deixados em repouso no gelo por 30 minutos. Após este passso, o tubo da transformação foi submetido a um choque térmico por 2 minutos no banho maria à 42° C e logo após este procedimento, os tubos voltaram para mais 10 minutos em gelo. Para prosseguir com a transformação 500 μ L de LB líquido foi adicionado ao tubo e eles foram levados para serem incubados a 37° C em agutação constante por 90 minutos.

As células transformantes foram plaqueadas em LB sólido acrescido do antibiótico utilizado como marcador de seleção e caso o plasmídeo utilizado possuísse seleção azul e branca, as placas foram acrescidas de Xgal e IPTG para seleção de possíveis clones positivos.

3.3 - EXTRAÇÃO DE RNA COM TRIZOL

Optamos por comparar a quantificação dos transcritos das linhagens selvagem (NA1000), P*imuA* (controle) e P*imuA*o^c para nos certificarmos dos níveis de expressão dos genes do operon *imuABC* após a inserção da mutação o^c nas linhagens P*imuA* e P*imuA*o^c.

Para isso, extraímos o RNA derivado destas linhagens na fase exponencial utilizando 10 mL das culturas, que foram centrifugadas a 5000 g por 10 minutos para retirada do meio de cultura. O sedimento bacteriano foi ressuspendido em 1 mL de Trizol[®] (Thermo Scientific) sem a utilização do vórtex e colocado no freezer -80° C durante a noite para prosseguir a extração no dia seguinte.

Para continuar o protocolo de extração, a mistura homogênea (*pellet* bacteriano + trizol) foi incubada em um banho seco à 65° C por 15 minutos, sendo que a cada 5 minutos os tubos foram vertidos para melhorar a homogeneização. Após este passo, 200 μ L de clorofórmio foram adicionados à solução, e os tubos voltaram a ser vertidos de forma contínua, desta vez por 15 segundos e colocados para descansar em temperatura ambiente por cinco minutos.

A extração de RNA prosseguiu com a centrifugação a 12000 g por 15 minutos em uma centrífuga previamente resfriada a 4º C. Como resultado, foram obtidas três fases (rosa, branca e transparente). Coletamos apenas a fase transparente e repetimos a lavagem com 200 μ L de clorofórmio para nos certificarmos que todo o trizol foi retirado da solução. Para precipitação do RNA, adicionamos a fase translúcida em 500 μ L de isopropanol e levamos para o freezer com temperatura de -80° C durante a noite.

No dia seguinte, retiramos o RNA em fase aquosa do freezer -80° C e esperamos descongelar no gelo. Assim que completado o passo anterior, centrifugamos os tubos por 20 minutos à 4° C. Retiramos o sobrenadante com o auxílio de uma pipeta, para evitar que o sedimento fosse perturbado. Para lavagem, usamos 1 mL de etanol 70% recémpreparado e previamente gelado. Voltamos a centrifugar os tubos por cinco minutos à 4° C. Retiramos o sobrenadante e secamos o pellet por cinco minutos à 55° C em banho seco e ressuspendemos em 50 μ L de água DEPC aquecendo os tubos a 55°C por cinco minutos. Os RNAs foram mantidos a -80° C até serem utilizados.

3.4 - TRATAMENTO DO RNA COM DNASE, SÍNTESE DE CDNA E qRT-PCR

Para avaliar os níveis de expressão do operon *imuABC* nas linhagens selvagem, P*imuA* e P*imuA*o^c usamos o gene *imuB* como alvo de qRT-PCR. Para este propósito, foram usadas duas réplicas biológicas de RNA de cada uma das linhagens mencionadas anteriormente na fase de crescimento exponencial.

Após a extração do RNA, este foi tratado com DNase I (Thermo Scientific) por 40 minutos a 37° C. Para evitar que a DNase I continue agindo durante a síntese de cDNA, 1 μ L de EDTA 1 M foi adicionado à reação, e esta foi incubada por 10 minutos a 65° C em banho seco para inativação desta enzima. Como controle da negativação da presença de DNA após o tratamento descrito, fizemos uma PCR usando os iniciadores do gene *rpoB* com alíquotas das amostras, assim como DNA genômico da linhagem NA1000 (usado como controle positivo). Em caso de ausência de amplificação, seguimos para o passo seguinte da síntese de cDNA.

A síntese de cDNA foi feita com o kit *SuperScript III first-strand synthesis for RT-PCR* (Invitrogen) utilizando 10 μ L do RNA tratado com DNase I em dois passos, como sugerido pelo fabricante. Para confirmar a síntese de cDNA, um novo controle de amplificação foi realizado com alíquotas das amostras após este passo e DNA genômico da linhagem selvagem (NA1000). Em caso de amplificação positiva, as amostras seriam usadas para quantificação de transcritos por qPCR.

Os iniciadores usados na qRT-PCR (*rpoD*fwd e *rpoD*rev; *imuB*fwd e *imuB*rev) tiveram sua eficiência testada previamente pelo então aluno de iniciação científica Marco

Noronha por um ensaio de amplificação por curva padrão e análise da curva de dissociação, o que foi fundamental para o cálculo final de expressão. Todos os iniciadores utilizados neste trabalho assim como suas sequências, podem ser observados na Tabela 3. Para quantificação de transcritos, utilizamos os cDNAs em reações de PCRs independentes para o gene *imuB* assim como para o normalizador, o gene que codifica a subunidade sigma da RNA polimerase (*rpoD*). A linhagem NA1000 em fase logarítmica (DO₆₀₀ 0,5) foi utilizada como referência.

As reações de RT-PCR em tempo real foram realizadas em um volume final de 12,5 μ L contendo 6,25 μ L de *Power SYBR Green PCR master mix*, 1 μ L de cada primer (10 μ M), 50 ng/ μ L de cDNA e 3,25 μ L de água ultrapura (Gibco). As condições de amplificação foram 95° C por 10 minutos, seguidos por 40 ciclos contendo 95° C por 15 segundos e 60° C por um minuto. Os experimentos de RT-PCR foram realizados em triplicata técnica utilizando-se a plataforma *StepOne Plus* (Applied Biosystems). As análises foram realizadas como proposto por Pfaffl em 2001.

3.5 - EXPERIMENTOS DE SOBREVIVÊNCIA A AGENTES GENOTÓXICOS

Ensaios de sobrevivência a diversos agentes genotóxicos foram realizados durante este trabalho. As linhagens NA1000, P*imuA*, P*imuA*o^c, $\Delta recA$ P*imuA* e $\Delta recA$ P*imuA*o^c foram testadas com mitomicina C (MMC) e UV como descrito a seguir. Usamos cultivos de *C. crescentus* na fase exponencial (DO₆₀₀ 0,4). Para os experimentos realizados com UVC, deixamos as células atingirem a concentração desejada e separamos 8 mL do caldo bacteriano em um tubo estéril, que foi centrifugado por cinco minutos a 5000 g para retirada do meio PYE. Os sedimentos bacterianos foram ressuspendidos no mesmo volume de água MilliQ estéril com ajuda do vortex. A solução foi despejada em placas de petri 60 x 15 e irradiada com 0, 50, 100 e 200 J/m² para as linhagens NA1000, P*imuA* e P*imuA*o^c. Os mutantes *recA* foram irradiados com doses menores de UVC (0 e 10 J/m²), uma vez que a ausência da resposta SOS e de recombinação homóloga nestas cepas as deixa mais sensíveis a agentes que danificam o DNA.

Ensaios de sobrevivência ao UVC também foram realizados com a linhagem NA1000 pareada com a *imuC* de forma similar ao ensaio descrito anteriormente com as doses de 0, 150, 200, 250, 300 e 350 J/m^2 .

Para o cálculo de sobreviventes, alíquotas foram retiradas antes e logo após cada irradiação com UVC, submetidas à diluição seriada em PYE, e semeadas em placas de

PYE – ágar, que foram posteriormente mantidas 48 horas à 30° C para contagem de UFC. O cálculo da fração de células que sobreviveram à irradiação foi obtido dividindo a contagem do número de UFC obtidos após a irradiação pelo número obtido em alíquotas não irradiadas (N/N0). Todas as manipulações experimentais e o crescimento das bactérias em PYE – ágar foram realizados na ausência de luz para evitar a fotorreativação, que é um mecanismo de reparo de DNA em bactérias que repara os fotoprodutos (Kimura et al., 2004).

A sobrevivência à MMC também foi feita com células em crescimento exponencial nas linhagens NA1000, P*imuA* e P*imuA*o^c. Após chegar DO₆₀₀ desejada (0,4) o caldo bacteriano foi dividido em alíquotas de 1 mL em tubos Falcon ® de 13 mL de fundo redondo, e receberam doses crescentes de MMC (0; 0,25; 1 e 2 μ g/mL). O tratamento com MMC aconteceu durante uma hora em agitação constante a 30° C. Após este tempo, alíquotas de todos os tubos foram retiradas e diluídas de forma seriada para plaquear em PYE – ágar. A obtenção do número de sobreviventes foi feita de forma similar à descrita para a radiação UVC.

Experimentos de sobrevivência a Metil Metasulfonado (MMS) também foram realizados com as linhagens NA1000 e *imuC*. Estas linhagens foram tratadas com 0 e 15 mM da droga recém diluída. Todo o procedimento metodológico para a sobrevivência a este agente é igual aquele usado para MMC.

3.6 - ENSAIOS DE MUTAGÊNESE

Realizamos ensaios de mutagênese induzida por UV, MMC e MMS. Todos os ensaios foram feitos de forma similar, sendo que suas particularidades serão descritas a seguir.

Para testar a mutagênese induzida por UVC irradiamos as células 50 J/m² (P*imuA* e P*imuA*o^c); 10 J/m² (P*imuA* Δ *recA* e P*imuA*o^c Δ *recA*) e 10, 25, 50 e 75 J/m² (NA1000 e *imuC*). Todas as rodadas de irradiação foram acompanhadas por um controle sem irradiar, que foi usado como indicativo da mutagênese espontânea para cada uma das linhagens descritas anteriormente.

Assim como a sobrevivência, os experimentos de mutagênese induzida por UV foram feitos no escuro para evitar o reparo dos dímeros por fotoreativação. Neste contexto, bactérias em caldo na fase exponencial (DO₆₀₀: 0,4) crescidas em meio PYE foram separadas em alíquotas de 8 mL que foram centrifugadas para a retirada do meio

de cultura e ressuspendidas no mesmo volume de H₂O MilliQ estéril para serem irradiadas. Após a irradiação com as doses descritas anteriormente, 200 μ L foram retirados de cada ponto, e adicionados a um tubo de 13 mL contendo 800 μ L de PYE líquido concentrado à 1,2 X. O mesmo procedimento foi feito com o controle sem irradiar. Após este passo, os tubos foram incubados por 24 horas longe da luz branca e em agitação constante em uma estufa a 30° C. Para cada experimento de mutagênese utilizamos duplicatas técnicas de cada ponto irradiado, assim como do controle.

No dia seguinte, alíquotas de 100 μ L de cada um dos tubos foram retiradas para diluição seriada e plaqueamento em PYE – ágar, que foi usado para a contagem de sobreviventes após a irradiação. Do volume restante, 700 μ L foi centrifugado a 3000 g por três minutos em temperatura ambiente e ressuspendido em 100 μ L de PYE para plaqueamento em PYE – ágar com 100 μ g/mL de rifampicina. Este antibiótico foi usado para detecção de mutantes Rif^R após o tratamento com todos os agentes testados neste trabalho. Todas as placas foram mantidas na estufa a 30° C por 48 horas para a contagem dos sobreviventes e das colônias rifampicina – resistentes (Rif^R).

A mutagênese induzida por MMC e MMS, que são agentes químicos, possui passos adicionais àqueles descritos para UVC. O tratamento com estas drogas foi feito de forma similar aquele descrito para a sobrevivência a estes agentes. Para a mutagênese após o tratamento, todo o caldo bacteriano (1 mL) que recebeu as doses de 0,25 μ g/mL de MMC (P*imuA* e P*imuA*o^c) e 15 mM de MMS foi lavado duas vezes com PYE (centrifugação 3000 g por cinco minutos e ressuspenção em 1 mL de PYE líquido) para a retirada de toda a droga do meio. Após as lavagens, 200 μ L da cultura tratada foram adicionados em 800 μ L de PYE líquido em tubos de 13 mL. A partir de cada um dos tratamentos foram feitas duas réplicas técnicas para a mutagênese induzida por estas drogas. Estes tubos foram deixados agitando por 24 horas à 30° C para retirada de alíquotas para a sobrevivência e mutagênese Rif^R de forma similar ao descrito para mutagênese UV.

Para calcular a taxa de mutagênese, calculamos a sobrevivência após os tratamentos e normalizamos a contagem de mutantes 700 μ L para 1 mL. Dividimos o número de mutantes normalizado pelo número de viáveis para obter a taxa de mutantes.

3.7 - TESTE DE FLUTUAÇÃO

As taxas de mutação espontânea foram determinadas por testes de flutuação nas linhagens P*imuA* e P*imuA*o^c como proposto por Foster em 2006. Para este trabalho, um pré – inóculo das linhagens testadas foi feito no dia anterior em PYE líquido com 0,5 μ g/mL de gentamicina. Este caldo bacteriano foi diluído para 10² células/mL em 11 culturas independentes em tubos tipo falcon 13 mL com tampa *snap cap* de fechamento duplo (que permite oxigenação dos cultivos). Estes tubos foram deixados agitando por 48 horas à 30° C. Após este passo, três tubos aleatórios foram escolhidos para retirar alíquotas de 100 µL para diluição seriadas usadas para a contagem de sobreviventes. Para mutagênese espontânea 700 µL foram concentrados em 100 µL para serem plaqueados em PYE – ágar com Rifampicina (100 µg/mL). O cálculo de mutantes foi determinado pelo método *Maximum likehood*, utilizando o programa FALCOR (Hall et al., 2009).

3.8 - ENSAIOS DE TRANSDUÇÃO

Para testar a transdução quantitativa nas linhagens P*imuA* e P*imuA*o^c usamos o fago Φ Cr30 *dinB*::spec previamente construído no nosso laboratório, derivado do profago Cr30 que é utilizado para ensaios de transdução em *C. crescentus* (Ely et al., 2015). Antes do ensaio de transdução, a preparação de fagos foi diluída para uma concentração de 10⁻⁴. Cada ensaio era composto de 6 tubos, sendo três controles e três tratados. Os controles estavam relacionados à integridade dos Fagos: 500 µL PYE líquido + 30 µL do fago diluído; 2° e 3° foram usados como controle das linhagens testadas: 500 µL de PYE líquido + 100 µL do pré-inóculo saturado de cada uma das linhagens. Para a transdução em si, adicionamos 30 µL do fago e 100 µl do pré-inóculo em 500 µL de PYE líquido, sendo que cada um dos controles e das transduções foi feita em tubos individuais de 13 mL estéreis. Após este passo, todos os tubos foram deixados agitando por quatro horas à 30° C. Quando chegou ao fim o tempo de agitação, plaqueamos o conteúdo de todos os tubos em placas de PYE-ágar enriquecidas com espectinomicina, estreptomicina e gentamicina que foram deixadas incubando por 72 horas à 30° C para a contagem dos transductantes.

3.9 - ESPECTRO DE MUTAÇÕES

Para determinar o espectro de mutações ocasionadas pela radiação UVC nas linhagens P*imuA* e $\Delta recA$ P*imuA*o^c, sequenciamos o segundo cluster de resistência à rifampicina do gene *rpoB*, onde a vasta maioria das mutações Rif^R se localiza, para determinar as substituições de base que foram induzidas por UVC. Usamos um par de iniciadores (*rpoD*fwd e *rpoD*rev) que flanqueiam uma região de 100 pb que faz parte desta região do gene *rpoB* (Tabela 3). Foram coletadas 31 colônias Rif^R da linhagem P*imuA* e 28 colônias do mutante $\Delta recA$ P*imuA*o^c. Todas as colônias foram provenientes de placas independentes de diversos experimentos em que usamos 10 J/m² de radiação UVC.

O fragmento de interesse foi amplificado por PCR de colônia, purificado usando o kit *Wizard DNA clean UP system* (Promega) e sequenciado usando o método de Sanger com BigDyeTM termination v3.1 cycle sequencing kit (ThermoFisher). As reações de sequenciamento foram amplificadas e preciptadas em nosso laboratório, e a microeletroforese para determinar as sequências foi feita pela Central analítica IQ – USP. As sequências obtidas foram comparadas com a do gene *rpoB* de *C. crescentus* presente nos bancos de dados para determinar as mudanças que conferiram resistência, usando o software Geneious 8.1 (Biomatters)

3.10 - EXPRESSÃO DA PROTEÍNA ImuA

Em parceria com o grupo do professor Dr. Márcio Dias do ICB-USP tentamos as seguintes metodologias para tentar expressar a proteína ImuA de *C. crescentus* em *E. coli*. Como descrito previamente, clonamos o gene *imuA* em fase com o vetor de expressão pET-28b com a *tag* de histidina. O primeiro passo para a expressão foi feito através de um protocolo teste de expressão de proteínas em *E. coli*, padronizado pelo grupo do Prof. Márcio Dias.

Para isso, transformamos previamente as seguintes linhagens de *E. coli* quimiocompetentes: BL21, Rosetta e BL21/pGRO7 (Tabela 1) com o plasmídeo *imu*ApET-28b. De placas frescas de transformação (1 a 3 dias após o crescimento) coletamos uma colônia isolada de cada uma das células transformadas e inoculamos em 2 mL de LB para testar a expressão na temperatura de 37° C. Tubos adicionais foram preparados para testar a expressão a 18° C. Deixamos o caldo inoculado agitando (200 – 220 rpm) nas temperaturas adequadas por aproximadamente 4 horas, quando já era possível observar o crescimento bacteriano por turbidez. Para o teste de expressão realizado na temperatura de 18° C, deixamos a cultura agitando por 30 minutos adicionais a 37° C para partir para o próximo passo.

Para indução da expressão da proteína, adicionamos 1 µL de isopropyl-βthiogalactopyranoside (IPTG) na concentração de 400 nM, e deixamos em agitação constante por quatro horas à 37° C e durante a noite a 18° C. Após a expressão, transferimos 500 µL dos cultivos para tubos de 1,5 mL para centrifugação (12000 g por um minuto). O sobrenadante foi removido com cuidado para não perturbar o sedimento formado. Para a lise, tratamos o sedimento utilizando 1/10 do volume de cultura (50 μ L) do reagente BugBuster (Sigma), 5 µL de Lisonase (uma mistura de DNase 100 mM e Lisozima 100 mM) – este passo tem como objetivo ajudar na lise das células e degradar o DNA genômico. O sedimento foi ressuspendido com ajuda do vórtex e a solução foi deixada na bancada em temperatura ambiente por 15 minutos. Após este passo, o lisado foi centrifugado por três minutos a 12000 g. O sobrenadante, que corresponde à fração solúvel, foi transferido para um tubo novo e ressuspendemos o sedimento em 10 µL de BugBuster (Sigma), e voltamos a centrifugar por três minutos a 12000 g. O sobrenadante foi descartado, e o sedimento foi ressuspendido em 50 µL de BugBuster. Após este passo, corremos a fração solúvel e o pellet ressuspendido em BugBuster em um gel SDS-Page para avaliar o resultado da expressão.

Após o teste de expressão com as três linhagens descritas anteriormente, a linhagem BL21 foi escolhida como a mais promissora para a expressão em larga escala. Para isso, esta linhagem foi transformada com o plasmídeo *imuA*pET28b. Após 24 horas de incubação a 37° C, uma colônia isolada da linhagem transformada foi escolhida para fazer o pré-inoculo em 100 mL de LB + canamicina. No dia seguinte, 10 mL do pré-inóculo foi adicionado em erlenmeyers contendo 1 L de LB (no total de 2 L). Os erlenmeyers foram colocados para agitar a 37° C até os cultivos atingirem DO₆₀₀: 0,6 (fase exponencial), após este passo, foi iniciada a indução com IPTG (200 nM) que aconteceu durante a noite a16° C. No dia seguinte, o caldo bacteriano total foi centrifugado (5000 g por 20 minutos a 4° C). O sedimento que resultou da centrifugação foi transferido para um Becker e tratado com 50 mL com Tris – HCl pH: 7,8; 200 µL Lisozima 100 mM; 200 µL DNase 100 mM e 200 µL de *Phenylmethane sulfonyl fluoride* (PMSF) que é um inibidor de protease. Para homogeneizar o sedimento nos reagentes durante a lise

química, foi utilizado um bastão de vidro e a solução foi deixada descansando por uma hora com o seu conteúdo imerso em gelo. Após este tempo, o conteúdo da lise química imerso em gelo foi sonicado por seis ciclos de 30 segundos, sendo cada um destes ciclos espaçados por tempos de um minuto de descanso com pulsos de três.

O lisado foi centrifugado por uma hora com uma rotação de 15000 g a 4º C, corremos um gel de SDS-Page para avaliar o resultado da expressão e com o auxílio da Aluna de doutorado do professor Marcio Dias, Priscila Bury, purificamos a proteína usando o aparelho AKTA (GE Healthcare).

De forma alternativa, também tentamos purificar a proteína ImuA em *E. coli* usando outro método de lise mecânica. Desta vez utilizamos ciclos de congelamento e descongelamento em nitrogênio líquido (cinco ciclos) no lugar do uso do sonicador.

3.11 - ISOLAMENTO DE COLÔNIAS E EXTRAÇÃO, QUANTIFICAÇÃO E CONTROLE DE QUALIDADE DE DNAS USADOS PARA O SEQUENCIAMENTO DE GENOMAS

As linhagens NA1000 e *imuC* foram expostas a doses crescentes de radiação UVC (0, 150, 200, 250 e 350 J/m²). A sobrevivência a estas doses foi obtida como já descrito anteriormente neste trabalho. A partir dos nossos resultados de sobrevivência escolhemos a dose de 300 J/m² para coleta de colônias isoladas.

De cada experimento de sobrevivência ao UV, recuperamos uma única colônia isolada da dose de 300 J/m^2 , que foi semeada em uma placa nova de PYE – ágar. A partir desta placa, recuperamos uma colônia isolada para a extração de DNA.

Um pré – inoculo de 4 mL foi feito a partir desta colônia, e parte deste caldo (1,35 mL) foi usado para congelar as células no dia seguinte na nossa coleção, e o restante foi centrifugado à 12000 g por 10 minutos para obtenção de sedimento com bactérias para a extração de DNA genômico. Além das colônias isoladas de sobreviventes ao UVC, também congelamos e extraímos o DNA de amostras que não foram irradiadas.

A extração de DNA genômico foi feita usando os kits da Promega *Wizard*TM (*Genomic DNA purification*) ou Qiagen (*Dneasy Blood & tissue kit*). Após a extração, todos os DNAs foram submetidos a controles de qualidade. O primeiro deles foi uma eletroforese em gel de agarose 1% para verificar a integridade da molécula e possível contaminação com RNA. Também quantificamos os DNAs usando o NanoDropTM, que além de estimar a concentração, também nos fornece dados de pureza nas razões 260/280

(para esta razão foram considerados valores próximos 1.8 - que são considerados puros para DNA). Valores próximos 1 podem ser considerados problemáticos para a contaminação com proteínas, fenóis ou demais compostos que possuem absorção a 280 nm. A razão de 260/230 também foi usada (valores que estão fora do intervalo 2 - 2.2 podem ser considerados problemáticos para contaminação com compostos que absorvem a 230 nm).

Para a construção das bibliotecas da Illumina, os DNA genômicos foram quantificados por fluorimetria, usando o Qubit (kit QubitTM dsDNA BR assay – Thermo Fisher Scientific). Apenas os DNAs que estavam com uma concentração acima de 40 ng/µL (obtido com Qubit) e com razões aceitáveis obtidas pelo NanoDrop foram usados para sequenciamento.

3.12 - SEQUENCIAMENTO DE GENOMAS

C. crescentus é uma bactéria que possui o genoma rico em bases GC, por isso, usamos para o preparo das bibliotecas genômicas o kit *TruSeQ PCR free* - Illumina[®]. O protocolo detalhado para a construção pode ser consultado neste site: (https://www.illumina.com/techniques/sequencing/dnasequencing/wholegenomesequen cing.html).

O primeiro passo para a construção de bibliotecas nesta metodologia é a fragmentação de DNA usando o Covaris® para obtenção de fragmentos de 550 pb. Para aperfeiçoar o protocolo de fragmentação sugerido pela Illumina, testamos o ciclo proposto por eles (duty cycle: 10%; intensidade: 2.0; cycles per bust: 200; duração 45 segundos; modo: frequency sweeping; displayed power: S2 9w e temperatura 5,5° - 6° C) e fizemos uma mudança no tempo de 45 para 35 segundos. Após a fragmentação do DNA, corremos um gel de agarose 1,5 % e comparamos os tamanhos dos fragmentos obtidos para os dois tempos.



Figura 3. Teste de fragmentação de DNA genômico no Covaris. A) Marcador de peso molecular; B) Fragmentação com 45 segundos; C) Fragmentação do DNA com 35 segundos e D) Marcador de peso molecular Lambda DNA/Hind III).

O nosso teste mostra que existe um acúmulo de DNA maior na faixa de 250 – 500 pb quando o DNA foi fragmentado com 35 segundos. Por isso, o protocolo de fragmentação foi adequado para este novo tempo no preparo de todas as bibliotecas genômicas de *C. crescentus* construídas neste trabalho. Sem outras adaptações o protocolo para construção das bibliotecas de 550 pb foi o mesmo proposto pelo o manual da Illumina.

De forma adicional, a Illumina sugere dois ensaios para avaliar a qualidade das bibliotecas recém-preparadas antes do sequenciamento. O primeiro deles é uma microeletroforese de capilar feita no Bioanalyzer (Agilent DNA 1000 – Agilent Technologies). Neste caso, alíquotas de 3 μ L das bibliotecas foram mandadas ao CEFAP – USP para a realização do ensaio. Os resultados foram considerados satisfatórios quando indicavam que a maior concentração de DNA estava próxima a 1 Kb.

Para quantificar a concentração de fragmentos no tamanho adequado que estão ligados a adaptadores usados no sequenciamento, a Illumina sugere realizar uma qPCR com o *KAPA Library quantification Kit*, o seu protocolo detalhado pode ser encontrado no link a seguir: http://www.mbl.edu/jbpc/files/2014/05/KAPA Library Quantification Illumina TDS.p df. Para a qPCR, diluímos as bibliotecas de DNA para 0,6 pMol (previamente quantificadas com Qubit) com Tris – HCl 10 nM pH 8 + 0,05% TweenTM. A normalização da concentração de cada amostra foi feita usando o fator de diluição para 0,6 pM e o tamanho de 650 pb (proposto pela Illumina).

O *pool* de amostras sequenciadas foi feito tanto normalizando as amostras para 2 nM, quanto para 4 nM. Esta variação é consequência da heterogeneidade de concentrações de bibliotecas obtidas em diferentes corridas durante este trabalho.

Todas as sequências foram obtidas pelo MiSeq, em parceria com o CEFAP – USP usando cartuchos (v3) de 600 ciclos (2 x 300 pb) e como controle interno 1% de PhiX.

3.13 - ANÁLISE DOS GENOMAS DE Caulobacter crescentus

Analisamos as sequências obtidas na corrida do aparelho MiSeq usando o software Geneious 8.1 (Biomatters limited). Após a transferência dos dados, pareamos as sequências FastQ obtidas dos dois ciclos de sequenciamento. Esta análise nos permitiu obter conjuntos de fragmento de DNA com o mesmo tamanho, assim como a leitura de ambas as pontas do fragmento, gerando desta forma, dados de alta qualidade.

Para a detecção de SNPs, incluímos na nossa biblioteca do Geneious a sequência da linhagem selvagem de *C. crescentus* NA1000 previamente sequenciada (http://www.genome.jp/dbget-bin/www_bget?refseq:NC_011916) por Marks e colaboradores, 2010. Para a chamada de SNPs automatizada pelo Geneious, as sequências foram primeiro mapeadas contra esta referência. Só foram contabilizados como SNPs, indels e mutações *in Tanden* aqueles que estavam presentes em pelo menos 75% dos reads mapeados. Após esta análise, regiões contendo estas variações foram analisadas individualmente quanto a sua cobertura para verificar a solidez dos dados em relação à cobertura total.

Com as variantes confirmadas, fizemos as seguintes análises manuais sobre os SNPs encontrados. Realizamos análises de posicionamento destas no mapa circular do genoma de *C. crescentus* para verificar se existe uma região preferencial de acúmulo de mudanças como consequência a lesões provocadas por luz UV. Olhamos também as bases vizinhas para verificar se as mutações encontradas ocorrem exclusivamente em locais de dinucleotídeos de pirimidina, ou se não necessariamente ocorrem frente aos fotoprodutos.Analisamos também o contexto das mutações para verificar se elas estavam presentes em regiões gênicas e/ou intergênicas.

Por fim, procuramos entender se existe um viés de acúmulo de mutações que é determinado pelo posicionamento dos fotoprodutos nas fitas *leading* ou *lagging*. Esta análise foi feita levando em conta a natureza da replicação do DNA em bactérias em que dois replicons se movimentam em sentido opostos para síntese de DNA. Esta análise só

foi possível para aquelas mutações localizadas em dinucleotídeos de pirimidina, onde assumimos que a fita contendo o dinucleotídeo de pirimidinas era a fita contendo a lesão que foi replicada. Assim, determinamos se as mutações advém de lesões presentes na fita *leading* ou *lagging*.

3.14 – ESTATÍSTICA

Usamos o teste não paramétrico de análise de variância Mann-Whitney para determinar possíveis diferenças estatísticas nos ensaios de mutagênese, sendo considerado valores de p < 0,05 como significativos. As análises estatísticas usadas para validar os dados obtidos pelo sequenciamento dos genomas serão feitas em parceria com o Prof Ricardo Vêncio, da USP, Ribeirão Preto, SP.

4 RESULTADOS E DISCUSSÃO

4.1 - A EXPRESSÃO CONSTITUTIVA DE *imuABC* EM NÍVEIS MÁXIMOS OBTIDOS NA RESPOSTA SOS NÃO AUMENTA A MUTAGÊNESE ESPONTÂNEA.

Para estudar os efeitos da expressão do produto dos genes *imuABC* em níveis fisiológicos obtidos durante a resposta SOS, construímos a linhagem P*imuA*o^c (Tabela 1). Esta linhagem possui uma mutação (o^c) no sítio de ligação do repressor LexA, que impede sua interação com a sequência do operador, permitindo a expressão constitutiva destes genes (Rocha et al., 2008). Como controle, construímos a linhagem P*imuA*, que é geneticamente idêntica à P*imuA*o^c, com excesso da mutação o^c, o que mantém a expressão destes genes reprimida na ausência de sinal indutor da resposta SOS (Figura 4).

Com objetivo de avaliar os níveis de expressão dos genes deste operon em P*imuA* e P*imuA*o^c, e desta forma validar estas construções, fizemos a quantificação relativa de transcritos do gene *imuB*. Nossos resultados mostram que há um aumento de 28 vezes na expressão de *imuB* na linhagem P*imuA*o^c quando comparada com a linhagem controle P*imuA* (Figura 4b). Este dado é corroborado por um resultado anterior do nosso grupo (Galhardo et al., 2005) que mostrou um aumento de aproximadamente 16 vezes na expressão deste gene após irradiação com luz UV.

A quantidade de transcritos do gene *imuB* se mostrou equivalente entre a linhagem selvagem (NA1000) e a linhagem controle, o que sugere que a pequena duplicação do gene *imuA* presente na linhagem P*imuA* não interfere nos níveis de expressão destes genes em condições normais (Tabela 1; Figura 4b). Experimentos adicionais foram realizados com as cepas selvagem, P*imuA* e P*imuA*o^c, e mostraram que a construção destas linhagens também não alterou a dinâmica de crescimento celular (dados não mostrados).

Estudos em *E. coli* mostram que mesmo na ausência da resposta SOS as proteínas envolvidas em TLS levam ao aumento das taxas de mutagênese espontânea quando tem acesso a forquilha replicativa. Nesta bactéria, ensaios de superexpressão do gene *dinB* (Pol IV) através de um plasmídeo heterólogo (F'*lac*) aumenta a taxa de mutação espontânea até 800 vezes (Kim et al., 1997).



Figura 4. Caracterização das cepas P*imuA* e P*imuA*o^c através do ensaio de RT-PCR quantitativa. A) Esquema da construção das linhagens mostrando a região promotora do gene *imuA* com uma pequena duplicação seguida da inserção do plasmídeo pMCS7 no cromossomo de *C. crescentus* e do cassete gênico *imuABC*. A diferença entre as linhagens P*imuA* e P*imuA*o^c pode ser observada nas sequências ao lado da sua identificação que assinalam as duas bases trocadas dentro do operador de LexA (em destaque), tornado o promotor constitutivo. B) Níveis relativos de expressão de *imuB*. Os dados de expressão observados foram gerados usando como controle a linhagem P*imuA*. Sendo estes resultados a média de dois experimentos independentes.

Outros casos envolvendo o operon *umuDC* (Pol V) também já foram relacionados ao aumento da mutagênese espontânea em uma linhagem mutante, chamada de *recA*730. Este mutante é caracterizado por possuir atividade de coprotease constitutiva em relação às proteínas UmuD e LexA, permitindo a formação da proteína Pol V (UmuD'₂C) (Ennis et al., 1995). Como consequência, esta linhagem mostra um aumento da mutagênese espontânea quando comparada com a linhagem selvagem. Este fenótipo mutador é suprimido quando o gene *umuC* é deletado (Caillet-Fauquet & Maenhaut-Michel, 1988). Em 1995, Ennis e colaboradores (1995) caracterizaram bioquimicamente o mutante

*rec*A730 com uma substituição no ácido glutâmico da posição 38 por uma lisina na sequência da proteína.

Entender o acesso destas polimerases propensas a erro à forquilha replicativa na ausência de lesão é extremamente importante. A mutagênese *untargeted*, ou seja, aquela que é consequência da resposta SOS, mas não é direcionada aos sítios com dano no DNA, foi primeiramente caracterizada por Maenhaut-Michael em 1985. Um trabalho posterior deste autor (Caillet-Fauquet & Maenhaut-Michel, 1988) concluiu que o acúmulo destas mutações era consequência do acesso da Pol V à forquilha replicativa, mesmo na ausência de dano no DNA. Hoje se sabe que estas proteínas envolvidas em TLS são responsáveis por uma série de eventos do compartimento celular, incluindo aqueles de interesse médico, como a mutagênese induzida por antibióticos (Gutierrez et al., 2013) e mutagênese induzida por estresse (Galhardo et al., 2007).

Diferentemente de estudos focados nas polimerases responsáveis pelo aumento da mutagênese *untargeted* em *E. coli*, nenhum trabalho foi feito para verificar se os produtos dos genes *imuABC* tem acesso a forquilha replicativa e conseguem aumentar a mutagênese na ausência de dano. Para responder a esta pergunta, comparamos as taxas de mutagênese espontânea das linhagens P*imuA* e P*imuA*o^c em *C. crescentus* (Figura 5). Nossos resultados para as taxas de mutação espontânea obtidos por ensaios de flutuação mostram que não existe diferença estatística para o aparecimento de mutantes Rif^R entre as duas linhagens.

Existem algumas hipóteses que podem explicar porque a superexpressão dos genes *imuABC* é incapaz de aumentar as taxas de mutagênese espontânea. É possível que a funcionalidade das proteínas ImuABC esteja ligada a algum tipo de processamento pós traducional, como acontece com UmuD que precisa sofrer autoclivagem mediada por RecA (Figura 2) antes de formar o complexo funcional da Pol V (Ippoliti et al., 2012). Porém, é improvável que este mecanismo de ativação seja feito pela proteína RecA, já que nenhuma das três proteínas codificadas por *imuABC* possui homologia com alvos de processamento por RecA (UmuD, CI e LexA). Outra possibilidade é que estas proteínas só tenham acesso à forquilha replicativa na presença de lesões, e quando não existe DNA danificado como molde, elas não são recrutadas e/ou são incapazes de ter acesso ao replissomo. Contudo, um estudo recente com *Myxococcus xanthus* superexpressando ~ 100 vezes apenas o produto do gene *imuC* mostrou um aumento discreto da mutagênese espontânea (cerca de quatro vezes) (Peng et al., 2017). Vale a pena lembrar que no nosso modelo estes genes não estavam superexpressos, e sim obedeciam a níveis de expressão

semelhantes àqueles da resposta SOS em *C. crescentus* (Rocha et al., 2008). Estudos adicionais seriam necessários para compreender como funciona a dinâmica de replicação da proteína ImuC e a sua interação com os outros componentes do replissomo durante a síntese de DNA danificado e livre de lesões para compreender o papel do produto de *imuABC* na mutagênese *untargeted*.



Figura 5. Expressão dos genes *imuABC* em níveis máximos obtidos durante a resposta SOS não é capaz de aumentar a taxa de mutagênese espontânea. Este gráfico é representativo da média de cinco experimentos de flutuação independentes, e mostra a ausência de diferença estatística entre as linhagens P*imuA* e P*imuA*o^c (p=0,46017).

4.2 - A EXPRESSÃO CONSTITUTIVA DE *imuABC* EM NÍVEIS MÁXIMOS DA RESPOSTA SOS NÃO AUMENTA A MUTAGÊNESE INDUZIDA POR DANOS NO DNA.

Como já discutido anteriormente, os produtos dos genes *imuABC* são necessários para a mutagênese induzida por dano em *C. crescentus* (Galhardo et al., 2005), fenômeno este que é reflexo da atividade de TLS mediada por estas proteínas. Testamos se a expressão constitutiva destes teria algum efeito na mutagênese induzida por UV. Primeiramente, comparamos a linhagem P*imuA* com a linhagem selvagem (NA1000) em pequenas doses (10 J/m²) de UVC, para verificar se a presença do plasmídeo integrado ao cromossomo da linhagem controle (P*imuA*) afeta a capacidade de induzir mutagênese dependente de TLS (Figura 6). Verificamos que não há diferença entre as linhagens, indicando que a construção não afeta negativamente a capacidade de TLS de *C*.

crescentus. Com o objetivo de verificar se a expressão constitutiva destes genes poderia ocasionar um aumento e/ou disfunção da mutagênese induzida por dano, submetemos as linhagens P*imuA* e P*imuA*o^c a ensaios de mutagênese, e quantificamos os mutantes Rif^R após exposição à MMC (Figura 7) e UV (Figura8).



Figura 6. Mutagênese induzida por UV nas linhagens selvagem (NA1000) e P*imuA*. Este dado é um indicativo que a construção da linhagem P*imuA* não altera a mutagênese induzida por UV. O gráfico mostra a média de três experimentos independentes. O asterisco é um indicativo de diferença estatística entre a mutagênese espontânea e induzida por UV.

Nossos resultados mostram que as linhagens controle (P*imuA*) e P*imuA*o^c não diferem na porcentagem de sobreviventes a MMC (Figura 7a). Para testar a tolerância a este agente, fizemos um ensaio de sobrevivência a diferentes doses de MMC (0; 0,25; 1 e 2 μ g/mL). Os resultados obtidos não mostram diferenças em nenhum dos pontos testados, indicando que a expressão em níveis máximos da resposta SOS dos genes *imuABC* não contribui para um aumento na tolerância aos dados causados por MMC.

Comparamos também as frequências de mutantes espontâneos, e também a mutagênese induzida por MMC entre P*imuA* e P*imuA*o^c, e não observamos diferença significativas entre as duas linhagens (Figura 7b). Além disso, comparamos as taxas de mutagênese espontânea e induzida por MMC de cada linhagem, e nossos resultados mostram que tanto a linhagem P*imuA* quanto a linhagem P*imuA*o^c são proficientes na mutagênese induzida por MMC. Desta forma, não foi observado nenhum incremento de mutagênese relacionado à expressão constitutiva de *imuABC* na linhagem P*imuA*o^c.



Figura 7. Efeitos da expressão constitutiva de *imuABC* na sobrevivência e mutagênese induzida por mitomicina C. A) Porcentagem de sobreviventes a diferentes doses de MMC. B) Mutagênese induzida por MMC ($0,25 \mu g/mL$) nas linhagens P*imuA* e P*imuA*o^c. Não há diferença estatisticamente significativa nas frequências de mutantes espontâneos e induzidos por MMC quando as duas linhagens são comparadas. Ambas as linhagens se mostraram proficientes na mutagênese mediada por dano (MMC), mostrando taxas que diferiram da mutagênese espontânea (P*imuA* p=0,01 e P*imuA*o^c p=0,03).

Resultados similares aos descritos anteriormente para MMC de tolerância e indução da mutagênese também foram encontrados quando usamos radiação UV como agente estressor (Figura 8). As linhagens P*imuA* e P*imuA*o^c não apresentaram diferenças significativas na resistência à luz UV, assim como não foram estatisticamente distintas em relação às taxas de mutagênese espontânea e induzida por UV. A radiação UV (50 J/m²), assim como a MMC, também foi eficiente em induzir a mutagênese mediada por SOS nas linhagens P*imuA* e P*imuA*o^c quando comparado com os dados de mutagênese espontânea das duas linhagens (Figura 8B).

Nossos resultados para tolerância e mutagênese induzida por MMC e UV mostram que a expressão dos genes *imuABC* em níveis máximos da resposta SOS não afeta o fenótipo descrito por Galhardo e colaboradores quando caracterizaram a mutagênese mediada por estes genes (Galhardo et al., 2005). Além disso, nossos resultados sugerem que não existe nenhum acréscimo de tolerância a estes agentes genotóxicos que esteja relacionada a esta expressão diferenciada destes genes. Em conjunto, estes dados indicam que a expressão de *imuABC* obtida após o tratamento com os agentes genotóxicos é suficiente para a TLS, e consequente mutagênese.



Figura 8. Efeitos da expressão constitutiva de *imuABC* na sobrevivência e mutagênese induzida por radiação UVC. A) Porcentagem de sobreviventes a diferentes doses de UV. B) Mutagênese induzida por UV (50 J/m²) nas linhagens P*imuA* e P*imuA*o^c. Não há diferença estatisticamente significativa nas frequências de mutantes espontâneos e induzidos por luz UV quando as duas linhagens são comparadas. Ambas as linhagens se mostraram proficientes na mutagênese mediada por dano UV, mostrando taxas que diferiram da mutagênese espontânea (P*imuA* p=0,006 e P*imuA*o^c p=0,006).

4.3 - EXPRESSÃO DOS GENES *IMUABC* EM NÍVEIS MÁXIMOS ENCONTRADOS NA RESPOSTA SOS RECUPERA A MUTAGÊNESE INDUZIDA POR UVC NA AUSÊNCIA DE *recA*

Em *E. coli* existem alguns modelos que tentam explicar como funciona o mecanismo em que uma enzima propensa a erro assume a forquilha replicativa e faz o *bypass* de lesões. Todos estes modelos foram introduzidos anteriormente neste trabalho e possuem um ponto de sobreposição. Existe necessidade da proteína RecA no complexo proteico responsável pela TLS (Schlacher et al., 2006; Fujii & Fuchs, 2009). Outro ponto em comum nestes modelos é que eles estão restritos a *E. coli*, que possui como principal enzima responsável pelo *bypass* de lesões a Pol V (Goodman, 2014). Diferentemente do operon *imuABC* presente em *C. crescentus, umuDC* estão restritos a um pequeno grupo de organismos dentro do filo Bacteria. Esta restrição filogenética chama atenção para necessidade de entender diferentes modelos de síntese translesão como aqueles que possuem outras TLS-DNA-Polimerases como ImuABC, por exemplo.

Para a Pol V, RecA é fundamental para várias funções, começando pela autoclivagem do repressor LexA, o que permite a expressão do operon *umuDC* durante a resposta SOS (Radman, 1975; Erill et al., 2012). Posteriormente, é necessária para um novo mecanismo de processamento, que vai mediar a autoclivagem da subunidade UmuD para UmuD' que é necessária para a formação do complexo proteico Pol V (UmuD'₂C) (Burckhardt et al., 1988), e por fim, existe a necessidade desta proteína associada ao DNA para que o mecanismo de TLS em si ocorra (Schlacher et al., 2006; Fujii & Fuchs, 2009). Além da Pol V, o produto do gene *dinB* (PolIV) também necessita de RecA na forquilha para que aconteça a replicação de DNA mediada por esta polimerase em *E. coli* (Godoy et al., 2007; Mallik et al., 2015).

Para testar se o cassete mutagênico *imuABC* presente em *C. crescentus* também necessita da proteína RecA para o mecanismo de tolerância ao dano, construímos no *background* das linhagens P*imuA* e P*imuA*o^c mutantes $\Delta recA$, chamados de $\Delta recA$ P*imuA* e $\Delta recA$ P*imuA*o^c (Tabela 1). A mutação em *recA* previne a indução da resposta SOS por agentes genotóxicos, porém a linhagem contendo a mutação P*imuA*o^c expressará *imuABC* em níveis máximos, tendo em vista que o sítio de ligação de LexA está alterado. Assim sendo, podemos aferir se há a necessidade de RecA para outra atividade que não seja a indução do SOS em si, no que diz respeito à TLS mediada por *imuABC*. Escolhemos para estes experimentos somente a luz UVC, uma vez que nosso grupo publicou um trabalho mostrando que para que aconteça a mutagênese induzida por MMC, é necessária a presença da endonuclease MmcB, que é regulada pela resposta SOS. Sem a ação desta proteína a proteína ImuC não teria substrato para fazer o *bypass* deste tipo lesão específica causado por este agente (Lopes-Kulishev et al., 2015).

Nossos resultados podem ser observados a seguir (Figura 9). Primeiramente, pode-se observar que a ausência de *recA* abole a mutagênese induzida por luz UV. Entretanto, na linhagem *recA* (que não possui resposta SOS) com a expressão constitutiva dos genes *imuABC* (*ArecA* PimuAo^c), o fenômeno de mutagênese induzido por radiação UVC (10 J/m²) é estatisticamente igual ao observado na linhagem P*imuA*. Estes resultados permitem duas conclusões. Primeiramente, pode-se inferir que não há nenhum outro gene do regulon SOS cujo aumento da expressão seja necessário para a TLS de fotoprodutos em *C. crescentus*. Além disso, fica evidente que RecA é dispensável para o mecanismo de TLS mediado por ImuABC, sendo somente necessária para a indução da resposta SOS.

Existem algumas hipóteses que podem explicar a ausência da necessidade de RecA para a TLS mediada por ImuABC. A mais provável está na função ainda obscura da proteína ImuA. Sabemos que esta proteína possui regiões de homologia com RecA e SulA (Figura 9), que são proteínas induzidas por SOS em *E. coli* e estão envolvidas em múltiplas funções dentro da célula, incluindo processamento de polimerases e ativação da TLS e controle da divisão celular, respectivamente (Galhardo et al., 2005). Talvez o acúmulo de ImuA na forquilha replicativa na presença de uma lesão seja fundamental para o recrutamento das proteínas ImuB (que é uma polimerase da família Y) e ImuC (polimerase da família C) para que ocorra a TLS. Para testar essa hipótese, o ideal seria purificar estas proteínas e fazer ensaios de atividade para testar se a proteína ImuA possui afinidade com a fsDNA (substrato para TLS-DNA-Polimerases) e verificar se ImuB e/ou ImuC possuem sítios de interação com o *beta–clamp* e com outros componentes do replissomo, como as helicases, que são características fundamentais para que as polimerases sejam funcionais.

RecA – E. coli	AFU91764.1	RAWKIETIWYSIMRLGEDRSMDVETISTGSLSLD
RecA – C. crescentus	ACL94606.1	MTSQAALKLVAKEEGDKQRALEAALAQIDRAFGKGSVMKLGEKGKVEIESVSTGSLGLD
ImuA – C. crescentus	YP_002518691.1	MELGMAGSREARLAALRGRIAAMEAGTRTPTPVLSFGEPSID
		: : :* *:*
RecA – E. coli	AFU91764.1	ALGAGGLPMGRIVEIYGPESSGKTTLTLQVIAAAQREGKTCAFIDAEHALDPIYARKLG
RecA – C. crescentus	ACL94606.1	ALGIGGLPKGRIVEVYGPESSGKTTLALHVVAEVQKAGGTAAFVDAEHALDPSYAYKLG
ImuA – C. crescentus	YP_002518691.1	CFPGGGLPLGGWHEVTGAGLEDETGAAPAAFVTQLIRGLTDRKGGAVVW
		.: **** * *: *:* : * * * . :
RecA – E. coli	AFU91764.1	DIDNLLCSQPDTGEQALEICDALARSGAVDVIVVDSVAALTPKAEIEGEIGDSHMGLAA
RecA – C. crescentus	ACL94606.1	NLDNLLVSQPDNGEQALEITDTLVRSGAVDIVVVDSVAALTPKAEIEGEMGDSLPGLQA
ImuA – C. crescentus	YP_002518691.1	ARRADLFAPGLLGLGFPAARLIQVRARDEAETLSLLEDALSTQGVAAAVAEAEAPDLTA
		*: * .*: : **:.:**
RecA – E. coli	AFU91764.1	MMSQAMRKLAGNLKQSNTLLIFINQIRMKIGVMFGNPETTTGGNALKFYASVRLDIRRI
RecA – C. crescentus	ACL94606.1	LMSQALRKLTASINKANTIVIFINQIRHKIGVMYGSPETTTGGNALKFYASVRLDIRRT
ImuA – C. crescentus	YP_002518691.1	RRLQLACEKRGGFGVVLHRRPYGGRAGGKPRLVSGSASFSRWRIAPA * :: * * * *.*:* *
RecA – E. coli	AFU91764.1	AVKEGENVVGSETRVKVVKNKIAAPFKOAEFOILYGEGINFYGELVDLGVKEKLIEKAG
RecA – C. crescentus	ACL94606.1	SVKARDEIVGNNVRVKVVKNKVAPPFREVEFDIMYGEGISKLGEVIDLGVKAGIIDKAG
ImuA – C. crescentus	YP 002518691.1	SGPPPDDIGRP
	_	: ::: :**:: *
RecA – E. coli	AFU91764.1	WYSYKGEKIGQGKANATAWLKDNPETAKEIEKKVRELLLSNPNSTPDLCRRWNG
RecA – C. crescentus	ACL94606.1	WFSYGSQRIGQGRDNVREFLKNNPDVAADIEKAVRKSSQKIEEELLVGGPEEGEED
ImuA – C. crescentus	YP_002518691.1	GWILQAQEAGHGPHPFRLVSQLADHDVAAAEAGRRFG

Figura 9. Alinhamento múltiplo das sequências das proteínas RecA de *E. coli* e *C. crescentus* e da proteína ImuA da bactéria estudada neste trabalho. O alinhamento foi feito com ClustalW2 online (http://www.genome.jp/tools-bin/clustalw).

Usando como modelo bacteriano *M. tuberculosis* que utiliza os genes *imuA*' e *imuBC* (neste organismo estes genes não estão organizados em operon) para fazer TLS,

Warner e colaboradores sugeriram um mecanismo parcial para a TLS feita pelos produtos destes genes. Eles acreditam que as proteínas ImuB e ImuC funcionam em conjunto para replicar o DNA lesionado, sendo ImuB uma proteína acessória fundamental para a TLS realizada por ImuC. Porém neste modelo eles não sugerem nenhum papel funcional para a proteína ImuA (Warner et al., 2010).



Figura 10. A expressão constitutiva de *imuABC* restaura a mutagênese induzida por luz UV no mutante *recA*. A mutagênese induzida por luz UVC na linhagem *recA* PimuA não difere da contagem de mutantes espontâneos (p= 0,29). Porém, a mutagênese dependente de fotoprodutos é restaurada na linhagem que expressa as proteínas ImuABC em níveis fisiológicos ($\Delta recA$ PimuAo^c) (p= 0,00205). As taxas de mutagênese induzida (p= 0,17) não diferem entre as linhagens $\Delta recA$ PimuAo^c e PimuA (usada como controle positivo do fenômeno de mutagênese induzida por UV).

Em 2005 quando Galhardo e colaboradores (Galhardo et al., 2005) descreveram os três genes responsáveis pela mutagênese induzida por dano em *C. crescentus*, foi caracterizado um espectro de mutações característico da ação desta via de TLS. Outras polimerases que são propensas a erro tendem a deixar uma "assinatura de mutações" que é característica da forma que ela faz o *bypass* das lesões. Um exemplo deste mecanismo é a Pol η , que quando ausente do genoma humano existe uma diminuição bruscas das substituições AT (Martomo et al., 2005). Com objetivo de verificar se os tipos de mutação encontrados na linhagem *PimuA* na *recA PimuA*o^c são equivalentes, o que sugeriria que o mecanismo de TLS operando na linhagem selvagem e na linhagem *recA* com expressão

constitutiva de *imuABC* é o mesmo, amplificamos e sequenciamos o gene rpoB de colônias Rif^R de ambas as linhagens.

Obtivemos sequências da região do gene *rpoB* de 31 mutantes Rif^R da linhagem P*imuA* e 28 da cepa *recA* P*imuA*o^c. Analisamos cada uma delas comparando com a sequência íntegra do banco de dados, e obtivemos o espectro de mutações causadas pela exposição à luz UVC (Figura 11 e Tabela 4). Como já introduzido neste trabalho, a radiação UVC é responsável por gerar prioritariamente lesões chamadas de fotoprodutos (Pfeifer et al., 2005). Estas modificações no DNA acontecem prioritariamente em citosinas adjacentes, sendo as sequências TT e TC mais reativas que CT e CC (Rastogi et al., 2010).

Nossos dados mostram que a maioria das mutações encontradas em ambas as linhagens são transições G:C \rightarrow A:T, aparecendo em uma frequência de 71% das mutações da P*imuA* e 64,3 da linhagem $\Delta recA$ P*imuA*o^c. Estes dados são consistentes com aqueles observados por Galhardo e colaboradores em 2005 (Galhardo et al., 2005). No seu trabalho esta mudança também foi a mais frequente (42,9%) para colônias Rif^R recuperadas após a irradiação com 90 J/m² de UVC. Neste mesmo trabalho, ficou determinado que a assinatura de mutações da proteína ImuC envolve também as mutações G:C \rightarrow C:G. Este tipo de mudança era totalmente abolido no espectro de mutações no mutante $\Delta imuC$ (Galhardo et al., 2005). No nosso sequenciamento encontramos este tipo de mutação tanto na linhagem P*imuA* (6,5%) quanto na cepa $\Delta recA$ P*imuA*o^c (7,1%).

Outra possível assinatura de mutações dependentes de ImuC descrita anteriormente por Galhardo e colaboradores em 2005 é o aparecimento de mutações *in tandem*, que neste caso são mutações agrupadas em bases adjacentes ou duas mutações separadas por uma base. Quando sequenciamos esta região do gene *rpoB*, encontramos um total de 6,5% destas mutações para a linhagem P*imuA* e 3,6% Δ *recA* P*imuA*o^c. Este tipo particular de mutação também foi muito frequente quando sequenciamos o genoma completo de *C. crescentus* (linhagem selvagem) após a irradiar com doses maiores de UV (300 J/m²) (ver adiante).

Apesar das variações encontradas nas frequências das mutações, todas as mudanças indicadas como fazendo parte do espectro de mutações de ImuC foram amostradas nas duas linhagens, o que é um indicativo de que elas foram consequência da mesma via de TLS, ou seja, houve restauração do fenômeno de TLS mediado por *imuABC* na linhagem *recA*.

1 /////10				
Tipo de mutação	PimuA	∆recA PimuAo ^c		
G:C→A:T	71,0	64,3		
G:C→C:G	6,5	7,1		
A:T→G:C	9,7	25		
A:T→C:G	0	0		
A:T→T:A	6,5	0		
Mutações in tandem	6,5	3,6		
Total	100%	100%		

Tabela 4. Tipos de mutações induzidas por UVC nas linhagens PimuA e ΔrecA PimuA o^c

Quando comparamos o espectro de mutações induzidas por UVC obtidos para as linhagens PimuA e $\Delta recA$ PimuAo^c (Tabela 4) é possível observar que elas estão representadas em proporções semelhantes entre as duas cepas. Pequenas diferenças podem se dever ao baixo número amostral e aleatoriedade dos ensaios de mutagenicidade. As mutações A:T \rightarrow T:A foram representadas apenas na linhagem PimuA. Este tipo de mutação não foi apontado por Galhardo e colaboradores (Galhardo et al., 2005) como fazendo parte da assinatura TLS mediada por ImuABC, e a maioria destas mutações não estão localizadas em regiões de dímeros de timina (Figura 11), e provavelmente não são derivadas do *bypass* de fotoprodutos pelo produto dos genes *imuABC*. Este mesmo tipo de substituição também foi pouco amostrado no espectro de mutações presentes em genomas derivados da linhagem NA1000 após ser irradiado com 300 J/m² (ver adiante).



Figura 11. Região do gene *rpoB* mostrando de forma esquemática mutações presentes em mutantes Rif^R após a exposição à radiação UVC (10 J/m²). As cepas P*imuA* e $\Delta recA$ *PimuA*o^c estão mostradas acima, com a sequência original do gene *rpoB* em preto, e as mutações encontradas posicionadas acima desta sequência. Nesta figura é possível observar as mudanças listadas na (Tabela 4).

4.4 - EXPRESSÃO DA PROTEÍNA ImuA

Para tentar entender o papel da proteína ImuA na síntese translesão dependente de *imuABC*, construímos o plasmídeo *imuA*pET28-b (Tabela 2) que possui *imuA* clonado em fase com uma cauda de histidina, para a purificação desta proteína ao usar uma coluna de níquel que possui afinidade por este aminoácido. A expressão foi feita através da indução por IPTG e os nossos resultados podem ser vistos a seguir.

Com o objetivo de escolher o melhor hospedeiro para expressão da proteína ImuA, fizemos um ensaio piloto com as seguintes linhagens de *E. coli* BL21, BL21 (pGRO7) e Rosetta. Nossos resultados indicaram a linhagem BL21 como a mais adequada para a expressão da proteína ImuA (dados não mostrados).

A proteína ImuA tem massa molecular predita de 25,45 kDA (sem a cauda de histidina) e coeficiente de extinção de 1,14. Para obtenção da proteína ImuA, a expressamos em larga escala em BL21 (4 L), e purificamos usando uma coluna de níquel que possui afinidade por histidina. Os nossos resultados da purificação identificaram um pico pequeno de 130 mAU presente no intervalo das frações 6 à 12. Para tentar determinar se este resultado era correspondente com o tamanho da proteína ImuA corremos um gel de acrilamida 10% que pode ser observado na Figura 12.



Figura 12. Gel de acrilamida 10% com as amostras das frações 12 à 6 (decrescente da esquerda para direita) da purificação da proteína ImuA. Neste gel podemos ver a presença de uma banda de 25 kDA, presumivelmente da proteína ImuA, nas frações indicadas pela seta vermelha (frações 11, 10, 9 e 8) quando comparadas com a banda de mesma massa do marcador. No gel também temos o *Flow through* (FL), Lisado (L) e a fração recuperada com 5% do tampão b (5%b)

Nossos resultados mostram pouca especificidade e baixo rendimento para a expressão da proteína ImuA usando a BL21 em um protocolo em larga escala. A discrepância entre os resultados do ensaio piloto e dos dados mostrados anteriormente pode ser consequência de mudanças protocolares em relação à lise celular. Quando testamos a melhor linhagem para expressar ImuA, o protocolo de lise era inteiramente químico (usando o reagente BugBuster). Porém, o uso deste protocolo é inviável para

grandes volumes que são necessários para expressão de proteínas e purificação.

Tentando eliminar este viés experimental, resolvemos usar um novo protocolo para expressão e purificação das proteínas. Desta vez utilizamos a lise mecânica com nitrogênio líquido, e para aumentar nossas chances de sucesso, voltamos a incluir as três linhagens de *E. coli* como receptores para o vetor *imu*ApET28-b. Estes resultados podem ser observados na Figura 13.



Figura 13. Gel de acrilamida 10%. Expressão da proteína ImuA em BL21, BL21 (pGRO7) e Rosetta. Para testar a eficiência do método de expressão usando lise mecânica o nitrogênio líquido. Comparamos o NI (não induzido) com a indução por 1 e 2 horas com IPTG para as três linhagens. Não é possível observar diferenças na indução da expressão quando comparado com os controles nas três linhagens.

Mesmo mudando o protocolo de extração, o experimento mostrou-se ineficiente. Quando comparamos os níveis de expressão do não induzido, com os tratamentos que sofreram indução por IPTG (1 e 2 horas) não observamos diferença na densidade das bandas presentes no gel (Figura 13) mostrando que este protocolo de expressão usando lise mecânica mediada por hidrogênio líquido não foi eficiente para obtenção das proteínas. Uma possível explicação é a presença de códons raros na sequência *imuA* para a expressão em *E. coli* (Figura 14). Esta análise mostrou que existe a presença de cinco trincas incomuns neste organismo, o que pode ter contribuído para o baixo rendimento da expressão (para fazer a análise da sequência de *imuA* para trincas raras usamos este link: http://nihserver.mbi.ucla.edu/RACC/). Outros problemas podem estar relacionados à natureza da proteína que está sendo expressa. Ainda não sabemos sua função, porém por homologia (Figura 9) e parcimônia é possível que ela tenha um mecanismo de ação análogo a da proteína RecA (Schlacher et al., 2006). Baseado nestas informações e possíveis características da proteína ImuA, talvez o método mais eficaz para purificar esta proteína seja usando um substrato de DNA ligado a ela para aumentar a sua estabilidade. Para fugir dos códons raros presentes em *E. coli* como organismo hospedeiro, uma outra possibilidade seria clonar esta proteína em um vetor de expressão diretamente em *C. crescentus*, ou substituir esses códons raros através de mutagênese sítio-dirigida.

A expressão da proteína que codifica a subunidade catalítica da Pol V (UmuC) foi árdua. Em um trabalho reflexivo sobre a importância do mecanismo de TLS mediado por Pol V Goodman (um dos responsáveis por um dos modelos de TLS para esta proteína já descrito) descreve como foi difícil expressar e purificar UmuC (Goodman, 2014). Neste mesmo trabalho ele ressalta como a proteína purificada foi fundamental para entender o mecanismo de TLS em *E. coli*. Expressar e purificar as proteínas ImuABC é o próximo grande passo para compreender como a maioria das bactérias (que não possuem *umuDC*) são capazes de realizar TLS. Esforços futuros do nosso grupo se concentrarão em otimizar a obtenção de ImuA recombinante para estudos bioquímicos e estruturais.

Sequência do Gene imuA(735 nucleotídeos)

Figura 14. A sequência do gene *imuA* (*Caulobacter crescentus*) com os cinco códons raros em *E. coli*, com seu posicionamento destacado em vermelho na sequência.

4. 5 - ImuA NÃO INTERFERE NAS FUNÇÕES DESEMPENHADAS POR RecA EM *C. crescentus*

Para tentar verificar se existe alguma diferença nas taxas de recombinação homóloga quando o operon *imuABC* está sendo expresso em níveis da resposta SOS, fizemos ensaios de transdução quantitativa usando as linhagens *PimuA* (controle) e *PimuA*o^c. Contagens diferenciadas de transductantes entre as duas linhagens poderia ser um indicativo de que a proteína ImuA quando expressa em níveis máximos, poderia competir com RecA por substratos, afetando assim suas funções, como a recombinação homóloga.

Nossos resultados podem ser observados na Figura 15, e mostram que nesta condição (com expressão em níveis máximos da resposta SOS), a expressão de ImuABC não afeta a frequência de recombinação homóloga. Isso pode se dever ao fato de que os níveis de ImuA atingidos durante a resposta SOS não promovem competição com RecA por substratos, ou simplesmente indicar que ImuA não se liga às regiões de fsDNA. Outros dados obtidos por nosso grupo também indicam que ImuA não interfere na capacidade de indução da resposta SOS mediada por RecA. Novas abordagens serão necessárias para testar o papel dessa proteína.



Figura 15. Transdução quantitativa. Teste da eficiência da transdução das linhagens *PimuA* e *PimuA*o^c com o fago Cr30 *dinB*::spec.

4.6 - OUTROS FATORES INDEPENDENTES DE ImuABC SÃO RESPONSÁVEIS POR PARTE DA MUTAGÊNESE INDUZIDA POR UV, E PELA MUTAGÊNESE INDUZIDA POR MMS EM *C. crescentus*

A mutagênese induzida por UV é apenas parcialmente dependente de *imuABC*, conforme demonstrado anteriormente para células irradiadas com 90 J/m² de UV (Galhardo et al., 2005). Nestes experimentos foi mostrado que 1/3 das mutações induzidas

por UVC ocorrem de maneira independente de *imuABC*. Para tentar entender se este fenômeno é dose-dependente, fizemos ensaios de mutagênese com doses progressivas de radiação (10, 25, 50 e 75 J/m²) nas linhagens NA1000 (selvagem) e no mutante *imuC*. Nossos resultados confirmam que existe um componente adicional à *imuABC* que é responsável por parte da mutagênese induzida por UV, tendo em vista que em todas as doses testadas pode-se observar um aumento da mutagênese por luz UV na linhagem *imuC*.

Quando comparamos as taxas de mutação nas duas linhagens para entender os dois fenômenos, observamos que existe diferença entre as cepas selvagem e mutante nas doses de 25, 50 e 75 J/m² mostrando que quanto maior a dose, maior prevalência de TLS. Por outro lado, o componente independente de *imuC* não exibe a dependência de dose, sendo constante em todas as doses testadas.

Estes dados são corroborados por dois trabalhos que discutem a importância do acumulo de dano para autoclivagem de LexA mediada por RecA em *E. coli* (Krishna et al., 2007; Sassanfar & Roberts, 1990). Neste organismo sabe-se que as enzimas que são responsáveis por TLS não são expressas de forma prioritária no decorrer desta resposta, sendo induzidas somente após a expressão de genes envolvidos em mecanismo de reparo de DNA menos propensos a erro (Kreuzer, 2013). Esta regulação está relacionada com a força de ligação diferenciada do repressor LexA pela região operadora de diferentes genes expressos em consequência da resposta SOS em *E. coli* (Krishna et al., 2007; Sassanfar & Roberts, 1990).


Figura 16. Mutagênese induzida por UV nas linhagens NA1000 e *imuC*. Nesta figura podemos observar as diferentes frequências de mutantes Rif^R para as duas linhagens, sendo o asterisco (*) indicativo de diferenças significativas entre as cepas NA1000 e *imuC* nas doses de 25 (p=0,01876), 50 (p=0,001) e 75 J/m² (p=0,00453). Comparamos também diferenças entre as taxas de mutagênese espontânea e induzida por UV nas duas linhagens, com **a** indicando diferenças estatística para a cepa selvagem (NA1000) em 10 (p=0,00005), 25 (p=0,00001), 50 (p=0,00001) e 75 J/m² (p=0,00004) e **b** indicando variações na linhagem mutante para *imuC (imuC)* nas doses: 10 (p=0,00001), 25 (p=0,00001) e 75 J/m² (p=0,00256). Média de pelo menos cinco experimentos independentes.

Como já introduzido anteriormente, *C. crescentus* possui outra polimerase propensa a erro (homólogo do gene *dinB*). Porém esforços passados do nosso grupo mostram que esta proteína não está envolvida no mecanismo de tolerância ao dano nesta bactéria (dados não publicados). E diferentemente de *imuABC*, *dinB* não é induzido como consequência da resposta SOS em *C. crescentus* (Galhardo et al., 2005; Rocha et al., 2008).

O mais interessante é que mesmo dentro da resposta SOS, existem outras proteínas que podem participar da resposta ao dano de DNA neste organismo. Uma delas é a proteína BapE, que é uma endonuclease de *C. crescentus*, que está envolvida em um mecanismo de apoptose-*like*, que seria uma via alternativa de resposta ao dano (Bos et al., 2012). Ainda dentro de genes da resposta SOS envolvidos neste mecanismo de mutagênese indução da mutagênese UV, nosso grupo tem estudado de forma paralela a este trabalho um possível papel de fotoproduto liases, que reparariam estes danos residuais (dados em desenvolvimento).

Hoje sabemos que a resposta SOS é um mecanismo universal de resposta ao dano de DNA em bactérias. Outros grupos que também estudam *C. crescentus* mostraram que existem genes não regulados por LexA que participam de mecanismos de tolerância ao dano nesta bactéria (Modell et al., 2014). A ação destes genes pode se dar em paralelo ou em conjunto com genes regulados por LexA, em mecanismos ainda não elucidados, para tolerar danos induzidos por UVC.

Além de estudar a mutagênese induzida por UV, resolvemos testar MMS para entender se os genes *imuABC* estão envolvidos no mecanismo de tolerância ao dano deste agente. Sabe-se que em *E. coli* o produto do operon *umuDC* é responsável pela mutagênese induzida por MMS (Grzesiuk & Janion, 1996); que quando em contato com o DNA é responsável por modificar a guanina para 7-metilguanina e adenina para 3-metiladenina. Drogas com efeito alquilante, como o MMS, podem levar a erro de pareamento e bloqueio da replicação (Lundin et al., 2005). *C. crescentus* não possui homólogos dos genes que codificam a Pol V, sendo o produto dos genes do cassete mutagênico *imuABC* a única via de TLS mediada por polimerases propensas a erro com fenótipo de tolerância responsáveis pela mutagênese induzida por dano nesta bactéria (Galhardo et al., 2005).

Nossos resultados para mutagênese induzida por MMS mostram que esta independe da TLS mediada por ImuABC em *C. crescentus* (Figura 17). Quando comparamos as linhagens selvagem e deficiente em TLS, não encontramos diferenças estatísticas entre as frequências de mutantes Rif^R. No entanto, se olharmos com cuidado os resultados, é possível verificar que existe uma tendência biológica de aumento da taxa de mutagênese induzida por MMS na linhagem selvagem quando comparada com o mutante *imuC*, porém este aumento não é estatisticamente distinto. Este resultado pode indicar que mesmo presente, a TLS não é o mecanismo prevalente para a mutagênese induzida por este agente nestas condições experimentais.



Figura 17. Mutagênese induzida por MMS. O gráfico mostra que existe um componente independente de *imuC* que é responsável pela mutagênese induzida por MMS em *C. crescentus*. A linhagem NA1000 (selvagem) possui taxas de mutação estatisticamente distintas quando comparamos a mutagênese espontânea e induzida por MMS (**a**: p=0, 00122). Também encontramos diferença entre as taxas de mutagênese espontânea e induzida por MMS na linhagem *imuC*::spec (**b**: p=0,00047). Quando comparamos as duas linhagens em relação a diferenças na mutagênese espontânea (p=0,06944) e induzida por MMS (ns: p=0,15866) não encontramos diferenças estatísticas. Média de três experimentos independentes.

O MMS é uma droga alquilante, e quando interage com o DNA leva à formação de lesões que são potencialmente citotóxicas, já que podem bloquear a progressão da forquilha replicativa (Lundin et al., 2005). Na presença deste tipo de dano, a TLS pode ser recrutada para continuar com a replicação e impedir o colapso da forquilha. Nossos resultados (Figura 17) mostram que a mutagênese induzida por MMS é independente de *imuC*. Além desta TLS-DNA-Polimerase, nosso grupo também investigou se o produto do gene *dinB* participa da sobrevivência e da mutagênese mediada por esta droga (dados não publicados). Os resultados obtidos para o mutante *dinB* são semelhantes a estes encontrados para *imuC*. Sabe-se que a proteína DinB (Pol IV) é recrutada para fazer o *bypass* de lesões alquilantes em *E. coli* e participa da tolerância a agentes alquilantes (Bjedov et al., 2007). A proteína Pol V também participa da tolerância a danos ocasionados por este agente neste organismo (Bagg et al., 1981). Estes dados mostram a importância da TLS para a tolerância de danos alquilantes em *E. coli*. No entanto, em *C. crescentus* nossos dados sugerem que a TLS mediada por ImuC e mesmo por DinB tem um papel secundário no mecanismo de tolerância a este agente.

Para descartar a participação dos genes imuC e dinB na tolerância a agentes alquilantes em *C. crescentus*, seria interessante testar no futuro se existe ambos podem atuar sobre estas lesões, analisando o duplo mutante. É possível que neste caso uma proteína compense a ausência da outra.

4. 7 - ESTUDO DO ESPECTRO DE MUTAÇÕES INDUZIDAS POR UVC EM NÍVEL GENÔMICO

Estudos de mutagênese são essenciais para entender como as células se comportam na presença de agentes nocivos. Comumente, usamos genes como marcadores da mutagenicidade. Um exemplo é o gene rpoB que é a base molecular para o aparecimento de mutantes que são resistentes a Rifampicina (Rif^R). Sendo que 95% das mutações que determinam este fenótipo estão restritas a uma sequência de apenas 81 pb (Telenti et al., 1993). Estes marcadores são extremamente funcionais, porém são uma simplificação do que acontece globalmente no genoma na presença de estresse. Wyrick & Robert discutem em seu artigo de 2015 (Wyrick & Robert, 2015) as vantagens das abordagens genômicas para estudar mecanismos de reparo de DNA e mutagênese. Os autores chamam atenção para as características do cromossomo: seja em níveis de compactação (exercido por proteínas de associação ao nucleóide) ou através do dinamismo que ocorre em regiões regulatórias para a expressão de genes. Todos estes eventos estão intrinsicamente ligados aos mecanismos de reparo de DNA (Wyrick & Robert, 2015). Logo, a abordagem genômica pode nos fornecer pistas adicionais e somar características ao espectro de mutação que é característico de cada agente. Desta forma, utilizamos o sequenciamento do genoma de bactérias irradiadas para análise do processo de TLS em C. crescentus.

4.7.1 - Determinação do genótipo da linhagem NA1000 presente no nosso laboratório

Para analisar o espectro de mutações induzidas por UVC em nível genômico em *C. crescentus*, primeiramente sequenciamos o genoma de 4 isolados da linhagem NA1000 (selvagem) e dois isolados da linhagem *imuC* que não foram irradiados com UVC e foram usados como controle.

Esta abordagem é necessária para que exista a caracterização de possíveis mutações presentes na linhagem NA1000 usada em nosso laboratório quando comparada com a sequência depositada no banco de dados (NC_011916.1). Além disso, nos permite

inferir o número de mutações espontâneas que podem estar presentes nos genomas de bactérias que passaram por um ciclo de descongelamenteo, crescimento em meio líquido, e isolamento de uma colônia. Nossos resultados mostram a presença de seis mutações que foram encontradas em todos os genomas usados como controle da linhagem selvagem (listadas na Tabela 5), que são características da linhagem selvagem presente no nosso laboratório. Este mesmo grupo de mutações também foi encontrado na linhagem *imuC*. Além destas mudanças, uma deleção de um pouco mais de 2 Kb também foi observada na nossa linhagem parental (Tabelas 5). Fora estas mutações comuns, nenhum dos seis isolados apresentou nenhuma mutação exclusiva, indicando que a frequência de mutações espontâneas que ocorrem durante o crescimento descrito acima é baixa, e não deve afetar a interpretação dos resultados obtidos na análise de células irradiadas que será mostrada adiante.

Estas mutações serão subtraídas do conjunto de mutações induzidas por UVC de cada genoma estudado a seguir, como uma forma de determinar apenas as mutações que são consequência da radiação.

iunçao proteica						
Locus	Gene	Função				
CCNA_00944	S/I	Determinação do comprimento do gancho				
		flagelar				
CCNA_00987	csbD	Proteína de resposta ao estresse				
CCNA_01467	ccoN	Subunidade I: citocromo C oxidase tipo ccb3				
CCNA_01803	S/I	Complexo de piruvato desidrogenase				
CCNA_03557	S/I	Proteína hipotética				
CCNA_03876	rho	Fator de terminação da transcrição rho				
CCNA_03239,	S/I;	Proteína de ligação à putrecine; Glutamina				
CCNA_03240,	glnA3;	sintetase; Glutamato-1-semialdeído 2,1-				
CCNA_03241	S/I	aminomutase				

Tabela 5. Lista de mutações encontradas nos genomas usados como controle sem irradiar nas linhagens NA1000 e *imuC* mostrando os loci afetados, nome do gene a

*S/I: Gene sem identificação

4.7.2 - Taxas de sobrevivência distintas ao UVC alteram a frequência de mutação em genomas irradiados

Para entender a importância do mecanismo de tolerância ao dano de DNA mediado por ImuABC em *C. crescentus*, fizemos ensaios de sobrevivência a radiação usando doses progressivas de UVC (0, 150, 200, 250 e 300 J/m²) tanto para a coleta de colônias da linhagem selvagem como para cepa mutante *imuC*. Como mostrado anteriormente (Figura 16), a TLS mediada por ImuC torna-se prevalente em doses mais altas de radiação, por isso optamos por utilizar a dose de 300 J/m² como uma forma de otimizar a coleta de colônias que sobreviveram à radiação através de TLS.

Coletamos 10 colônias isoladas da linhagem selvagem que foram irradiadas com 300 J/m² de UVC em meio PYE (não transparente), e apresentaram taxa de sobrevivência de 10%. Como resultado dos sequenciamentos, encontramos apenas 2 SNPs (*single nucleotide polymorphism*) nos dez genomas sequenciados. Os mapas destes dois genomas com estes SNPs únicos podem ser observados na Figura 18.

Desta forma, ficou evidente que necessitávamos de doses mais efetivas de UV para que a mutagênese pudesse ser estudada satisfatoriamente. Para aprimorar o experimento fizemos mudanças no protocolo de irradiação, que incluem um passo adicional de centrifugação do caldo bacteriano, com sua posterior substituição por água MilliQ estéril. Com a irradiação das bactérias em meio translúcido, foi possível potencializar o efeito da luz UVC, como pode ser observado na Figura 19. Este efeito protetor em relação a radiação pode ser explicado pela absorção dos raios UVC por aminoácidos presentes no meio de cultura.

Após a adequação do protocolo de irradiação, sequenciamos 10 genomas da linhagem selvagem (NA1000) com sobrevivência de aproximadamente 0,001% (300 J/m² – Figura 19). O aumento da suscetibilidade nessas condições reflete na diferença da taxa de SNP/genoma entre os diferentes experimentos. Quando os isolados foram obtidos de ensaios com 10% de sobrevivência ao UV, observamos uma média de 0,2 SNP/genoma. Já quando os isolados sequenciadas são derivadas de tratamentos que geram 0,001% de sobrevivência ao UV, temos um aumento para 5,1 SNP/genoma, conforme detalhado a seguir



Figura 18. Mapas de SNPs encontrados nas linhagens RSG483 e RSG574 (derivadas da NA1000) após serem irradiadas com 300 J/m². O DNA foi extraído a partir de colônias isoladas que são derivadas de experimentos com taxa de sobrevivência próxima a 10%. Os mapas foram gerados com o Geneious 8.1, usando como modelo o cromossomo circular de *C. crescentus*.



Figura 19. Curva de sobrevivência ao UVC. Porcentagem de sobreviventes a doses crescentes de radiação (150, 200, 300 e 350 J/m²). Estes dados são a média de pelo menos quatro experimentos independentes que foram realizados após a padronização do protocolo de irradiação quando retiramos o meio de cultura e substituímos por água MilliQ estéril.

4.7.3 – Distribuição de mutações induzidas por UV no cromossomo de C. crescentus

C. crescentus é um excelente modelo para o estudo do ciclo celular em bactérias. Isso porque esta α-proteobacteria possui uma divisão celular assimétrica que a permite a partir de uma célula talo gerar dois tipos morfo-fisiológicos de células distintos (talo e flagelada) a cada ciclo de divisão (Collier, 2012). Esta característica o torna um excelente organismo para estudos de replicação do DNA e de progressão e divisão celular (Schrader & Shapiro, 2015).

Ao compilar todas as mutações induzidas por UVC em todos os genomas derivados da NA1000 irradiada com luz UV, foi construído um "Pan-Genoma", mostrando todas as mutações, que pode ser observado na Figura 20 (Os genomas individuais não mostrados ao longo do texto estão presentes em apêndices no final deste trabalho). A organização espacial das mutações é extremamente interessante, já que mostra um acúmulo espelhado de mudanças na sequência de DNA em regiões intermediárias entre a origem e o término da replicação.

Para estudar a organização espacial das mutações induzidas por UVC no genoma de *C. crescentus*, dividimos o cromossomo circular de 4042929 pb em 16 regiões de 256683 pares de base cada. As 51 mudanças encontradas em relação a linhagem parental podem ser observadas na Figura 21 (A Tabela mostrando os limites das 16 regiões e a quantidade de mutações que está presente em cada uma delas pode ser observada no Apêndice 7). Nossos resultados mostram que duas regiões genômicas possuem a maioria das mutações. A primeira delas é o intervalo 1010733 a 1263415 pb com 7 mutações e a segunda região é 2779514 a 3032196 pb com 15 mudanças.

Quando dividimos o genoma em dois, separando as mutações que estão mais próximas a origem de replicação (proximal) ou ao término da duplicação do genoma (distal), mostramos na Figura 25b que 78% das mutações totais encontram-se mais próximas à região de término da replicação, e apenas 22% destas mudanças estão localizadas próximo a origem de replicação.



Figura 20. Pan-Genoma com a somatória de todas as mutações derivadas de isolados de NA1000 irradiados com UVC. A e B, assinalados pelas chaves são regiões que aparentemente são preferencialmente mutadas no cromossomo de *C. crescentus* quando irradiado. As cores diferentes representam mutações derivadas de de genomas diferentes.

Os mecanismos de replicação e divisão celular ocorrem ao mesmo tempo em bactérias e como consequência existe uma intensa maquinaria regulatória que trabalha para que eles aconteçam em sincronia (Taylor et al., 2017). Em *C. crescentus*, para que a replicação seja iniciada, é preciso que aconteça primeiramente um processo de diferenciação celular. Este fenômeno é consequência da distinção das células flageladas (móveis) que são incapazes de replicar o DNA em células talo (que possuem replicação ativa) (Mohl et al., 2001).

Após a síntese de DNA que acontece nas células talo, a divisão celular irá produzir novamente duas células morfologicamente distintas (flagelada ou talo). Como implícito pela dinâmica do ciclo celular incomum desta bactéria, existe uma regulação adicional que acontece no início da replicação de DNA (Taylor et al., 2017). Este passo extra é consequência da ação de uma proteína chamada GapR que é uma NAP de *C. crescentus*, que se liga preferencialmente em sequências ricas em AT que estão localizadas na região promotora de genes que controlam o ciclo celular (Ricci et al., 2016). De forma independente, Arias-Cartin e colaboradores (2016) mostraram que este tipo de NAP (GapR) possui um viés de ligação ao longo do cromossomo bacteriano. Sendo que regiões próximas ao início da replicação são alvos preferenciais de ligação desta proteína em *C. crescentus*.

Assim como o cromossomo eucariótico, o nucleóide bacteriano está associado a um conjunto de proteínas (histonas-*like*) chamadas de NAPs, que permitem a compactação do DNA durante a divisão celular. Estas proteínas são caracterizadas pela sua ligação quase sempre aleatória com DNA, e a sua ausência afeta o fitness celular (Badrinarayanan et al., 2015).

Estudos *in vitro* mostraram que a proteína HU (NAP) de *E. coli* tem afinidade por DNA lesionado e pode estar implicada em mecanismos de reparo de DNA (Oberto et al., 2009). A ausência de HU neste organismo foi associada a sensibilidade a radiação UV, pelo comprometimento do reparo por recombinação homóloga (Li & Waters, 1998). Em 2009, Oberto e colaboradores caracterizaram em *E. coli* um regulon induzido por HU que controla um grupo de genes diversos, incluindo a expressão de proteínas que estão relacionadas a anaerobiose, estresse à ácido, alta osmolaridade e, mais interessante ainda, esta proteína estaria envolvida na modulação do repressor LexA, que é o regulador negativo da resposta SOS em bactérias e controla a indução da expressão de várias proteínas envolvidas na resposta ao dano de DNA. Estudos recentes em *C. crescentus* identificaram uma NAP, chamada de GapR. Na ausência desta proteína, a célula cresce com dificuldades e apresentada fenótipos relacionados ao estresse (Ricci et al., 2016; Taylor et al., 2017). Diferentemente das NAPs de *E. coli* e *B. subtilis* (Badrinarayanan et al., 2015), GapR possui regiões de ligação preferencial no cromossomo de *C. crescentus* com viés de interação maior com regiões próximas a origem de replicação do que ao término (Arias-Cartin et al., 2016). A interação desta proteína com estas regiões pode ser uma das explicações para a distribuição dasos mutações observadas neste trabalho.

Em *C. crescentus* não existem homólogos para o operon *hupAB* que codifica HU (Arias-Cartin et al., 2016). Porém sabemos hoje que GapR caracterizada como uma NAP nesta bactéria aparentemente está relacionada de alguma forma com a resposta ao estresse. No seu trabalho publicado em 2016, Arias-Cartin e colaboradores fizeram uma análise por RNA-seq do mutante $\Delta gapR$ e da linhagem selvagem (NA1000). Quando comparados, a ausência desta proteína é responsável pela regulação positiva de 25 genes que estão envolvidos com a resposta ao estresse, sendo que 16 deles participam de mecanismos de reparo ao dano de DNA e oito são induzidos como consequência da resposta SOS - incluindo proteínas envolvidas em TLS (*imuABC*) (Galhardo et al., 2005) e outros mecanismos de tolerância ao dano nesta bactéria (*bapE* e *didA*) (Bos et al., 2012; Modell et al., 2014).

A ligação de forma preferencial de GapR próximo a região do início da replicação, e o acúmulo de mutações induzidas por dano próximas a regiões de término, podem sugerir algumas hipóteses. Como discutido por Arias-Cartin e colaboradores (2016) GapR aparentemente possui um papel regulatório para o início da replicação de DNA em *Caulobacter*. O mutante $\Delta gapR$ é viável, porém apresenta fenótipos pleiotrópicos de perda de *fitness*. Assim como HU, talvez a presença desta proteína module positivamente algumas repostas ao dano de DNA, que por sua vez, corrijam lesões de forma eficiente não deixando substrato para TLS.



Figura 21. Localização dos SNPs totais nas 16 subregiões do genoma de *C. crescentus* (A). Em B) mostramos que ao dividir o genoma em duas regiões que são caracterizadas pela sua proximidade com a origem (proximal) e terminação da replicação (distal) existe um viés para o acumulo de mutações próximas ao término da replicação.

Existem alguns trabalhos que associam as NAPs de outras bactérias com a modulação positiva do reparo por recombinação homóloga – que é um mecanismo de

tolerância ao dano de DNA livre de erro (Li & Waters, 1998) além da modulação de outros mecanismos de reposta ao dano de DNA que não são propensos a erro (Oberto et al., 2009). Baseado nestas informações, é possível que o UVC seja eficaz em lesionar o DNA de forma uniforme ao longo do cromossomo de *C. crescentus*. Contudo, o diferencial para a menor quantidade de mutações próximas à origem de replicação pode ser a presença da proteína de associação ao nucleóide GapR, que por sua vez poderia modular positivamente o reparo de DNA livre de erro nestas regiões, levando a uma distribuição não uniforme de mutações ao longo do cromossomo (Figuras 20 e 21). Por outro lado, também é hipótese plausível que NAPs como a GapR atuem como barreiras físicas para a luz UV, diminuindo o número de lesões nas regiões em que estas proteínas se liguem preferencialmente.

Foster e colaboradores em 2013 estudando mutações espontâneas em linhagens de *E. coli* deficientes em reparo de erros de emparelhamento (MMR) em nível genômico mostraram um padrão de acúmulo de mudanças próximas a regiões de término da replicação, similar ao encontrado neste trabalho (Figura 25). Eles explicaram a mutabilidade preferencial desta região ao assumir que como a replicação do DNA possui uma origem única e a síntese é bidirecional, regiões que estão próximas ao início da replicação (DNA recém-replicado) possuem mais substratos para recombinação homóloga, e por mais tempo, que regiões próximas ao término (Foster et al., 2013). A recombinação homóloga é uma outra alternativa para célula tolerar lesões no DNA que bloqueiam a progressão da replicação, sendo este mecanismo livre de erro (Fuchs, 2016). Este fenômeno também poderia promover o viés de mutações para regiões mais distais à origem de replicação.

4.7.4 - Espectro de mutações induzidas por UVC no genoma.

Para estudar o espectro de mutações induzidas por UVC em nível genômico na linhagem NA1000, compilamos todas as substituições encontradas e comparamos com dados já publicados pelo nosso grupo (Galhardo et al., 2005; Alves et al., 2017 – este trabalho) para tentar correlacionar as mutações encontradas nos genomas com aquelas com assinatura de TLS mediada por ImuC. Nossos resultados podem ser observados na Tabela 6.

Tipo de mutação	Genoma (NA1000)	NA1000*	imuC*	PimuA [#]
G:C→A:T	29,4	42,9	88,6	71
G:C→C:G	15,7	28,5	0	6,5
A:T→G:C	25,5	14,3	11,4	9,7
A:T→C:G	11,8	2,9	0	0
A:T→T:A	2,0	0	0	6,5
Indel	5,9	0	0	0
Mutações in tandem	9,8	10,4	0	6,5
Total	100% (51)	100% (35)	100% (35)	100% (31)
* Galhardo et al.,				

Tabela 6. Espectro de mutações induzidas por UVC em C. crescentus

2005

[#]Este trabalho

Em 2005, Galhardo e colaboradores publicaram o espectro de mutações induzidas por UVC das linhagens NA1000 (selvagem), *imuC* e *imuB*. Parte deste resultado pode ser observada na Tabela 6. Para gerar estes dados, eles compilaram as substituições de base encontradas no segundo *hotspot* do gene *rpoB* que confere resistência a Rifampicina quando irradiadas com 90 J/m² (Galhardo et al., 2005). A partir destes dados, foi possível determinar uma assinatura de mutações que são ocasionadas pela TLS mediada por ImuABC sendo elas substituições G:C→C:G e mutações *in tandem*. Estas mudanças desaparecem do espectro de substituições do mutante *imuC* que é deficiente em TLS mediada por ImuABC.

Estas mutações que são consideradas características da assinatura de ImuC voltam a aparecer no espectro de mudanças da linhagem $\Delta recA$ PimuA (Alves et al., 2017 – este trabalho: Tabela 4) mesmo sendo derivadas de experimentos de irradiação com uma dose diferente (10 J/m²). Assim como Galhardo e colaboradores (2005), neste trabalho também usamos mutantes Rif^R para determinar o conjunto de substituições encontrados na linhagem deficiente em SOS e com TLS ativa.

Quando comparamos as mutações encontradas em genomas derivados da linhagem NA1000 com as substituições achadas em mutantes Rif^R na mesma cepa, é possível observar algumas diferenças. Quando analisamos as mutações que estão presentes em nível genômico, encontramos algumas deleções de um único nucleotídeo, que estão ausentes de todos os demais espectros mostrados na Tabela 6. Trabalhos anteriores já associaram *frameshifts* de -1 par de base como consequência da TLS mediada pela Pol IV em *E. coli* (Strauss et al., 2000). *C. crescentus* possui um ortólogo do gene *dinB* porém não foi observado nenhum efeito deste gene sobre a mutagênese induzida por UVC. Não é impossível que as deleções mostradas na Tabela 6 possam ser consequência da ação desta proteína quando acionada nestas condições específicas ou pode ser que seja uma assinatura da mutagênese mediada por ImuC que só foi possível observar em escala genômica. Deve-se levar em conta que inserções e deleções não são toleradas no gene *rpoB*, de forma que os espectros de mutação anteriormente determinados são incapazes de fornecer informação quanto a este tipo de mutação. Para teste destas hipóteses, o ideal seria sequenciar isolados das linhagens *dinB* e *imuC* irradiados com UV, para verificar se este tipo de deleção desaparece dos seus respectivos espectros de mutação.

Com sequenciamento genômico, também encontramos uma substituição A:T \rightarrow T:A que também está presente no espectro de mutações da linhagem P*imuA* (Tabela 6). Como é pouco frequente, é possível que este tipo de mudança esteja associado à mutagênese espontânea independente de danos causados pela radiação, já que não está presente em regiões de dinucleotídeos de pirimidinas (Figura 11). O aparecimento desta mutação também pode ser consequência de outro mecanismo de resposta ao dano de DNA, mais precisamente ao UVC, diferente da TLS mediada por ImuC.

Como já mencionado, uma das características da mutagênese UVC mediada por ImuC é a presença de mutações *in tandem*. Na compilação das mutações encontradas em genomas derivados de isolados obtidos em condições de irradiação com 0,001% de sobrevivência, estes tipos peculiares de substituições agrupadas voltaram a aparecer com 9,8% das mutações totais. Este valor é muito similar ao encontrado por Galhardo e colaboradores em 2005 – quando eles computaram 10,4% de substituições *in tandem* da soma total das mutações totais presentes na linhagem NA1000. O mais interessante neste dado é a correlação dos tipos específicos de mutações *in tandem* encontradas nos genomas irradiados e nos dois espectros publicados anteriormente por Galhardo e colaboradores em 2005, e Alves e colaboradores em 2017. Além da frequência, os mesmos tipos de substituições se repetem em relação a mutações agrupadas; o que sugere que são resultados do mecanismo de TLS mediado pela ImuC.

Além disso, as trocas G:C \rightarrow C:G voltam a ser observadas, com uma frequência de 15,7%. Lembrando que este tipo específico de mutação desaparece do espectro de

mutações da linhagem *imuC* publicado em 2005 (Galhardo et al., 2005), e está associado à assinatura de TLS de ImuC em *C. crescentus*.

4.7.5 - O agrupamento de mutações derivadas do mesmo genoma pode indicar a distância percorrida pela TLS

Pouco se sabe como estas polimerases tem acesso a forquilha replicativa, porém é claro que esta regulação fina para sua ativação é um mecanismo celular para prevenir o acúmulo de mutações mediada pela sua ação na ausência de danos no DNA. Já foi descrito que Pol V de *E. coli* na presença do fator de processividade (*beta-clamp*) e da proteína RecA faz o *bypass* de lesões (Schlacher et al., 2006; Fujii & Fuchs, 2009). Em 2004, Fujii & Fuchs mostraram como ocorre a troca da Pol III pela Pol V em um sítio de lesão. Neste trabalho, os autores chamam atenção para a baixa processividade da Pol V – mesmo ligada ao *beta-clamp* - e como consequência desta característica, foi sugerido a TLS é um evento breve, que não se prolonga por muitos pares de bases de extensão, e logo após o *bypass* da lesão a Pol III retoma a forquilha replicativa, e prossegue com a replicação livre de erro (Fujii & Fuchs, 2004). Como resultado, as mutações decorrentes da síntese de DNA propensa a erro tendem a se acumular em regiões pequenas próximas a lesão, pelo menos no que diz respeito à Pol V de *E. coli*. Como a TLS em *C. crescentus* usa uma polimerase que não pertence à família Y, esse quadro pode ser diferente.

Alguns dos nossos resultados mostram que existe um acúmulo de mutações induzidas pela radiação UVC em regiões pequenas. Um exemplo desta natureza está mostrado na Figura 22, onde possível observar o acúmulo de mutações em pequenas regiões do genoma. Estes eventos podem estar associados à síntese translesão. Como já discutido, estas proteínas são pouco processivas e percorrem pequenas distâncias no genoma – esta característica é uma consequência da sua baixa afinidade com os demais componentes do replissomo (Goodman & Woodgate, 2013). O primeiro dos eventos que chamamos atenção, assinalado por A na Figura 22, tem uma distância de 4276 pares de bases entre as mutações. Se acreditarmos que estas mudanças na sequência do DNA são consequência da TLS, é mais provável que a polimerase responsável por esta síntese tenha tido acesso a forquilha replicativa uma única vez e como consequência da TLS foi responsável por estas duas mutações.



Figura 22. Mapa genômico de uma linhagem derivada da cepa selvagem NA1000 após ser irradiada com UVC. Nesta figura podemos observar todas as mutações, com destaque para dois pontos assinalados por A e B, onde a distância entre estas mudanças é pequena. Este fenômeno pode ser associado a eventos de TLS.

4.7.6 - As mutações induzidas por UV ocorrem majoritariamente em regiões de dinucleotídeos de pirimidinas

A radiação UVC é responsável por gerar fotoprodutos na molécula de DNA ao reagir com bases pirimídicas adjacentes (Rastogi et al., 2010). Estas lesões são altamente citotóxicas, já que impedem a progressão da replicação feita pela Pol III. Em *E. coli*, a Pol V é a polimerase responsável por fazer a TLS destes danos, isso acontece quando esta TLS-DNA-Polimerase é recrutada e assume a forquilha bloqueada evitando o colapso da replicação, tendo como consequência o aumento da taxa de mutação (Schlacher & Goodman, 2007).

Para verificar se existe uma correlação entre o contexto da sequência de DNA e o aparecimento de mutações induzidas por UVC, analisamos as mutações provenientes dos experimentos com aproximadamente 0,001% de sobrevivência. Nossos resultados podem ser observados na Figura 23, e mostram que 86% das mutações catalogadas estão em regiões de pelo menos duas pirimidinas adjacentes. Este resultado é bastante expressivo, sugerindo que as mutações são majoritariamente decorrência de TLS diretamente sobre lesões no DNA. Entretanto, 14% das mutações observadas se dão por outros mecanismos que não a síntese de DNA frente aos fotoprodutos em dinucleotídeos de pirimidina.

Outros mecanismos como o reparo por excisão de nucleotídeos (NER) também são recrutados na presença de dano causado por UV. O NER é orquestrado por um grupo de quatro genes (*uvrABCD*) que estão envolvidos no reconhecimento do dano, excisão e resíntese do DNA usando como molde uma fita não danificada (Kisker et al., 2013). Como este mecanismo é pouco passível de erro, é improvável que estas mutações seja uma consequência de falhas do NER.

Na presença de fotoprodutos, as replicases são incapazes de prosseguir com a replicação do DNA e as mutações derivadas destas lesões são consequência da ação de polimerases especializadas que conseguem utilizar uma fita de DNA danificado como molde para prosseguir a replicação. Estas enzimas são propensas a cometer erros durante a síntese de DNA devido a características morfológicas que as permitem fazer TLS (Sale et al., 2012). Possivelmente estas proteínas estão sendo recrutadas nestes sítios de lesões para impedir o colapso da forquilha, e como são propensas a erros deixam pelo caminho um rastro de mutações que está associado à sua ação. É possível que as demais mutações que não se encontram em dinucleotídeos de pirimidinas sejam consequência desta atividade em curtos trechos de DNA contendo fotoprodutos. Em outras palavras, seriam

fruto da TLS de lesões vizinhas, e consequência da natureza propensa a erro das polimerases envolvidas nesse processo.



Figura 23. Porcentagem de mutações localizadas em sequências de pirimidinas adjacentes, que são alvo da radiação UVC para a formação de fotoprodutos. Aproximadamente 86% das mutações encontradas em genomas irradiados derivados da linhagem selvagem estão localizados em dinucleotídeos deste tipo. Dados contabilizados apenas dos experimentos com 0,1% de sobrevivência a UVC – linhagem selvagem (NA1000).

Dentre as sequências de pirimidinas adjacentes próximas às regiões de mutação, fizemos um levantamento para saber se existe sequências de dinucleotídeos mais propensas à mutagênese por luz UV. Nossos resultados podem ser observados na Figura 24, e mostram que a sequência CC é a mais frequentemente mutagenizada por luz UVC (44%), seguida pelas sequências TC com 28%, CT com 16% e TT com 12%. Ao analisarmos as frequências destas sequências de dinucleotídeos no genoma de *C. crescentus* encontramos os seguintes números: 39,3% de CC, 27,3% de TC, 21,4% de CT e 11,9% de TT. Quando comparamos estes valores com aqueles encontrados para a frequência de dinucleotídeos que estão no contexto de mutações induzidas por irradiação com UVC, existe claramente uma manutenção das mesmas proporções (Figura 24). Estes dados sugerem que a frequência de mutações nos diferentes dinucleotídeos reflete a abundância destes dinucleotídeos no genoma de *C. crescentus*, e não a reatividade individual destes duplos com a radiação UVC, ou mesmo uma característica da maquinaria de TLS mediada por ImuABC. Esta hipótese difere dos dados revisados por Rastogi e colaboradores (2010), que mostram que as sequências TT e TC são mais

reativas a radiação UVC. Os nossos dados indicam que a natureza rica em GC do genoma de *C. crescentus* (Marks et al., 2010) é a característica preponderante para determinação do espectro de dinucleotídeos que estão no contexto de mutações induzidas por UVC.

Parte das mutações (14% - Figura 23) que não estão localizadas em dímeros podem ser consequência de fenômenos de mutação espontânea. No entanto, parte delas deve ser consequência da ação destas polimerases que são propensas a erro que assumem a forquilha na presença de uma lesão e persistem com a replicação pouco precisa até a replicase assumir a forquilha replicativa.



Figura 24. Frequência de sequências de pirimidinas adjacentes. A) Sequências de dinucleotídeos de pirimidina encontrados no contexto das mutações induzidas por UVC em genomas de *C. crescentus* (Dados contabilizados apenas de genomas de isolados obtidos em condições de 0,001% de sobrevivência a UVC – linhagem selvagem: NA1000). B) Frequência de dinucleotídeos de pirimidina no genoma de *C. crescentus*.

4.7.7 – As fitas contínua e descontínua são igualmente mutáveis por luz UV em C. crescentus

Para a duplicação do material genético, organismos vivos utilizam as fitas senso e antisenso como moldes. Em 2012, Furusawa criou um modelo implicando as diferenças de fidelidade entre a replicação destas duas fitas. Neste trabalho ele considera a natureza estocástica das mutações, porém julga que diferenças no mecanismo de replicação das fitas *leading* (contínuo) e *lagging* (descontínuo) podem contribuir para um viés no aparecimento de mutações como consequência do mecanismo de replicação do DNA. Isto acontece porque durante a síntese contínua de DNA, a replicase se mantem na forquilha replicativa e a presença de lesões que bloqueiam é mais perceptível, podendo ser facilmente corrigida, já ao replicar de forma descontínua o DNA as várias entradas da replicase podem mascarar estas lesões e dificultar o processamento destas.

Quando analisamos nossos dados de genomas provenientes de células irradiadas com UVC, fizemos uma triagem para determinar se existia alguma tendência de acúmulo de mutações que foram consequência da replicação contínua ou descontínua de DNA. Esta análise só foi possível para mutações que estão posicionadas em dinucleotídeos de pirimidina, nas quais assumimos que a fita onde a lesão estava originalmente posicionada é conhecida. Após determinar se o dinucletídeo de pirimidina estava presente na fita + ou - (Apêndice 8) calculamos a porcentagem de lesões presentes na fita contínua e descontínua (Figura 25).

Nossos resultados mostram que não existe um viés neste aspecto para o aparecimento de mutações induzidas por UVC em *C. crescentus*, com 45% das lesões que originaram mutações presentes na fita descontínua, e 55% presentes na fita contínua.

Nossos resultados sugerem que o aparecimento de mutações induzidas por UVC no genoma de *C. crescentus* não tem influência da maquinaria replicativa. Outros trabalhos mostram que o acúmulo de mutações na fita *lagging* é bem documentado na literatura. Em 1993, Veaute & Fuchs ao testar o que eles chamaram de efeito assimétrico da mutagênese, usaram plasmídeos contendo lesões na fita contínua e descontínua. Os resultados deles mostraram que para a mutagênese causada por adutos AFF existe um aumento de pelo menos 20% na frequência de mutações quando as lesões estavam localizadas na fita descontínua (este aumento da frequência foi observado na presença e ausência da resposta SOS).

No começo da década passada, o grupo de Fijalkowska publicou dois trabalhos (Maliszewska-Tkaczyk et al., 2000 & Gawel et al., 2002) para tentar relacionar o posicionamento de lesões induzidas por UVC e seu potencial mutagênico quando presentes na fita contínua e/ou descontínua.

No trabalho de 2000, ao estudar linhagens de *E. coli* que expressam construtivamente os genes *umuDC* usando como marcador genético o gene *lacZ*, estes autores observaram que existe uma prevalência de mutações que são acumuladas pela replicação de DNA usando a fita decontínualesionada por UVC (Maliszewska-Tkaczyk et al., 2000). Porém em 2002 quando eles testaram o mesmo sistema para determinação de viés de replicação contínua e descontínua de DNA após a radiação UV em linhagens deficientes em NER ($\Delta uvrA$) não foi possível observar diferenças na mutagênese usando como marcador o gene *lacZ*.

De forma geral, a maquinaria responsável por fazer o *bypass* de lesões causadas por UVC em *C. crescentus* (*imuABC*) parece não diferir entre a replicação contínua e descontínua quando se trata da fidelidade da replicação ou acesso ao substrato, ou seja, o aumento das taxas de mutação independente da fita que está sendo replicada.



Figura 25. Frequência de mutações que foram derivadas da replicação de DNA contínua (leading) e descontínua (lagging).

4.8 - MUTAGÊNESE INDEPENDENTE DE ImuABC A NÍVEL GENÔMICO

Sequenciamos seis linhagens derivadas da cepa *imuC* também irradiadas com 300 J/m² para entender aspectos da mutagênese dependente e independente de *imuABC*. Devido ao baixo número de mutações, nossos resultados ainda são preliminares, mas já nos fornecem algumas pistas. Assim como nos genomas derivados da linhagem selvagem, as mutações no Pan-Genoma da linhagem *imuC* também aparentam estar agrupadas mais próximo às regiões de término da replicação (Figura 26), o que pode implicar que a disposição das mutações no cromossomo bacteriano não é uma consequência da TLS mediada por estas proteínas, uma vez que as mutações independentes de *imuABC* apresentam o mesmo padrão.

Por outro lado, quando comparamos os números de SNPs, observamos uma diminuição no número de mutações na linhagem *imuC* em comparação com NA1000, o que sugere que grande parte destas mutações são consequência da TLS mediada por *imuABC*. Entre os seis genomas sequenciados dos mutantes deficientes em TLS, encontramos uma soma de sete SNPs em relação aos isolados *imuC* não-irradiados usados

como controle, o que equivale uma média de 1,16 SNP/genoma sequenciado. Este valor é equivalente a aproximadamente 1/5 da média das mutações dos genomas derivados da linhagem selvagem que foram irradiados com 300 J/m² (5,1 SNP/genoma). No trabalho de 2005, Galhardo e colaboradores caracterizaram o produto dos genes *imuABC* como responsáveis pelos mecanismos de tolerância ao dano de DNA em *C. crescentus*, estes autores mostraram que parte da mutagênese induzida por UVC era independente de *imuABC*. Ao longo deste trabalho também mostramos que a TLS dependente destas proteínas se torna um mecanismo prevalente de acordo com o aumento da dose de radiação (Figura 16). Com isso, podemos concluir que ao usarmos uma dose de 300 J/m² para obter isolados para o sequenciamento, seria esperado que a síntese translesão fosse responsável por grande parte da mutagênese, com a concomitante diminuição do fator independente de TLS na mutagênese induzida por UVC a nível genômico.



Figura 26. Mapa das mutações encontradas nos seis genomas sequenciados derivados da linhagem *imuC*. Assim como no Pan-Genoma das mutações derivadas da linhagem selvagem (NA1000), existe um acúmulo preferencial das mutações localizadas próximo a região de término da replicação.

4.9 - MODELO DE TLS MEDIADO POR ImuABC

Diferentemente da TLS mediada por umuDC (Pol V), não sabemos como o produto dos genes imuABC fazem o bypass de lesões no DNA. A construção de um modelo de TLS para estas proteínas só será possível quando conhecermos a fundo a bioquímica de cada uma das proteínas componentes deste sistema. Hoje, temos em mãos algumas informações que podem ser úteis para pensarmos em um modelo especulativo de síntese translesão dependente destas proteínas. Sabemos que a proteína ImuA possui similaridade com as proteínas RecA e SulA de E. coli. A expressão destas proteínas é dependente da resposta SOS neste organismo, e elas estão envolvidas em diversas funções relacionadas a mecanismos de reparo e tolerância ao dano no DNA (Galhardo et al., 2005; Rocha et al., 2008; Schlacher et al., 2006; Fulconis et al., 2006; Fujii & Fuchs, 2009). Sendo que a proteína RecA está implicada em ambos os mecanismos de TLS mediado pela Pol V (Schlacher et al., 2006; Fujii & Fuchs, 2009). Nossos resultados mostram que a ausência de RecA não altera a mutagênese induzida por UVC em C. crescentus (Figura 9), ou seja, não existe necessidade desta proteína para que a TLS seja funcional quando mediada por ImuABC. Acreditamos que a similaridade de RecA com ImuA, somada aos nossos dados (Galhardo et al., 2005; este trabalho), sugere que esta proteína tenha um possível papel análogo a RecA na TLS mediada por ImuABC: ImuA se ligaria a regiões de lesões funcionando como um sensor de dano no DNA (Figura 27).

As demais proteínas presentes neste cassete mutagênico são polimerases. ImuB é uma polimerase da Família Y (Galhardo et al., 2005). Esta família de proteínas agrupa polimerases especializadas em replicar DNA danificado (Sale et al., 2012). Já a proteína ImuC é produto do gene *dnaE2 (imuC)* que é uma segunda cópia do gene *dnaE* especializado em fazer o *bypass* de lesões (Galhardo et al., 2005). A proteína DnaE é a replicase reponsável pela síntese de DNA livre de erro (O'Donnell, 2006). Sabemos que as proteínas ImuB e ImuC também são necessárias para a mutagênese induzida por dano em *C. crescentus*. Para uma polimerase ser funcional, ela precisa, além de possuir domínios funcionais, ter sítios de interação com algumas proteínas do replissomo, como o *beta-clamp* (Yao & O'Donnell, 2010). Em 2010, Warner e colaboradores mostraram através de um ensaio de duplo híbrido que a proteína ImuB de *M. tuberculosis* interage com o Beta *beta-clamp* e com a proteína ImuC, porém estes autores chamam atenção para que alguns dos resíduos catalíticos característicos de polimerase estão ausentes na estrutura de ImuB – sugerindo que ela não seria funcional para a replicação de DNA

propensa a erro. Estes autores sugerem que esta proteína tem provavelmente um papel acessório na TLS realizada por ImuC.

De forma resumida nosso modelo é apenas uma complementação para o que foi proposto por Warner e colaboradores (2010) para *M. tuberculosis*. Com isso, sugerimos que a TLS realizada por ImuABC começa com a indução da expressão destas proteínas por SOS - regulado por LexA e RecA (Galhardo et al., 2005; Rocha et al., 2008). A proteína ImuA funciona como um sensor de dano de DNA, e recruta as proteínas ImuB e ImuC para a o forquilha bloqueada pela lesão. ImuB funciona como um intermediário para a TLS, uma vez que interage com a proteína ImuA e ImuC, e com o *beta-clamp* possibilitando a ligação deste complexo proteico com o replissomo. (Figura 27).

Nosso modelo especulativo pode ser observado na Figura 27. Para ter certeza da função de cada uma destas proteínas na síntese translesão seria necessário extrapolar as barreiras dos ensaios genéticos e tentar entender como elas interagem com o DNA danificado e entre si. É possível que mesmo dentro das bactérias que utilizam este sistema existam algumas diferenças no funcionamento deste complexo. A proteína ImuA' de *M. tuberculosis* é extremamente diferente da ImuA de *C. crescentus*. Há indicativos também que a proteína ImuC de *C. crescentus* possua um sítio de interação com o *beta-clamp* então não é possível descartar que o produto do gene *imuC* seja o componente do cassete mutagênico que esteja relacionado com a interação indispensável de polimerases com o replissomo.



Figura 27. Síntese translesão mediada por ImuABC. Para a configuração deste modelo usamos como base os trabalhos de Warner et al., 2010 e nossos esforços durante os últimos anos para tentar entender o funcionamento destas proteínas. A) Os genes *imuABC* são expressos na presença de dano de DNA em resposta a autoclivagem do repressor LexA mediada pela presença da nucleoproteína RecA ativa que reconhece regiões de fsDNA. B) A Pol III, responsável pela duplicação do DNA genômico, não consegue continuar a replicação na presença de uma lesão. A proteína ImuA se acumula nestas regiões de fsDNA e C) recruta as proteínas ImuB e ImuC para fazer a TLS. Dados em *M. tuberculosis* mostram que ImuB interage com o *beta-clamp* e com a proteína ImuC porém é incapaz de fazer a polimerização, uma vez que parte dos seus resíduos catalíticos estão ausentes, deixando o papel do *bypass* para a proteína ImuC (Warner et al., 2010; este trabalho).

5 CONCLUSÕES & DISCUSSÃO GERAL

 Diferentemente da Pol V, a indução da expressão dos genes *imuABC* em níveis máximos obtidos na resposta SOS é insuficiente para aumentar as taxas de mutagênese espontânea.

Algumas bactérias como *C. crescentus* não possuem ortólogos do operon *umuDC*. Elas utilizam o produto do cassete mutagênico ImuABC para fazer o *bypass* na presença de lesões no DNA, impedindo o colapso da forquilha – que é um evento potencialmente citotóxico. Alguns trabalhos mostraram que a Pol V pode assumir a forquilha replicativa mesmo na ausência de dano, e como consequência, levar ao aumento as taxas de mutação espontânea. Nossos dados para *C. crescentus* sugerem que ImuABC não tem acesso a forquilha replicativa e/ou não é mutagênica quando está sintetizando o DNA usando como molde uma fita de DNA integra, sem lesões.

• A expressão de *imuABC* em níveis máximos obtidos na resposta SOS não promove mudanças na tolerância e mutagênese a MMC e UVC em *C. crescentus*.

Apenas o aumento da expressão dos genes que fazem parte do cassete mutagênico não foi suficiente para alterar as taxas de sobrevivência nem para aumentar a mutagênese induzida por UVC e MMC em *C. crescentus*.

 ImuABC não precisa de RecA na forquilha para fazer o *bypass* de lesões induzidas por UVC.

Diferentemente dos modelos de TLS mediados pela Pol IV e Pol V que já foram propostos para *E. coli*, nossos resultados mostraram que mesmo na ausência de RecA, a expressão de *imuABC* é capaz de restaurar a mutagênese induzida por UVC em um *background* sem a indução da resposta SOS. Estes dados sugerem um mecanismo de TLS novo para bactérias, que pode ser importante para elucidar como bactérias que não possuem ortólogos dos genes *umuDC* acumulam mutações.

 Existe uma complementação dos espectros de mutação induzidas por UVC nas linhagens PimuA e ΔrecA PimuAo^c. Ao sequenciarmos regiões do gene *rpoB* (que quando mutado confere resistência à rifampicina) das linhagens *PimuA* (controle) e $\Delta recA$ *PimuA*o^c (TLS em níveis da resposta SOS), observamos uma equivalência das mutações, o que sugere que a compensação do fenótipo observado na linhagem $\Delta recA$ *PimuA*o^c quando comparada com a cepa controle é de fato uma consequência da mutagênese mediada por ImuABC.

 Ensaios de mutagênese induzida por UVC com doses crescentes mostram que existem dois mecanismos que são responsáveis pela mutagênese causada por este agente. Com o aumento da dose de UVC, a TLS fica mais prevalente.

Estudos em *E. coli* sugerem que a força de ligação do repressor LexA e a região operadora dos genes que são induzidos em consequência da resposta SOS é responsável por modular a expressão diferenciada de grupos de genes diferentes na presença de dano no DNA. Ainda não foram publicados dados da força dos operadores dos genes que fazem parte da resposta SOS em *C. crescentus*, porém é possível inferir que o acúmulo de dano no DNA causado pelo o aumento da dose de UVC pode estar relacionado ao aumento da ativação da TLS em comparação com o outro mecanismo de resposta a radiação UVC, que é mais ativo em doses menores.

• Existe um componente independente de *imuABC* que está envolvido na mutagênese induzida por MMS em *C. crescentus*.

Quando comparamos as taxas de mutagênese induzida por MMS entre as linhagens selvagem (NA1000) e o mutante *imuC*, não observamos diferenças entre as frequências de mutantes induzidos por dano entre as duas cepas. Este dado sugere que outro mecanismo, diferente da TLS mediada por estas proteínas, é responsável pela mutagênese induzida por este agente.

• Sobrevivências diferentes ao UVC levam a taxas de mutação diferentes.

Quando sequenciamos linhagens derivadas da cepa selvagem (NA1000) irradiadas com 300 J/m² obtemos taxas de sobrevivência diferentes, usando diferentes metodologias experimentais. Condições que levam à 10% de sobrevivência promovem aumentos pequenos na mutagênese, enquanto que condições onde a toxicidade da luz UV é maior geram mais mutações

• No cromossomo circular de *C. crescentus* existem duas regiões espelhadas mais próximas ao término que são preferencialmente mutadas por UVC.

Observamos um padrão claro em que as mutações estão agrupadas em espelho próximo a regiões próximas ao término da replicação. Este fenômeno pode refletir diferenças na eficiência de mecanismos de reparo de DNA entre regiões próximas ao início da replicação e o término da replicação.

• Mutações induzidas por UVC derivadas da linhagem selvagem de *C. crescentus* não estão preferencialmente associadas com a replicação contínua ou descontínua.

A replicação do DNA em *C. crescentus* acontece a partir de uma origem única e segue de forma bidirecional. A replicação da fita descontínua é consistente com várias entradas da replicase que poderia afetar a percepção de lesões pela maquinaria replicativa. Porém nossos dados mostram que não existe um viés de fita replicada nas mutações induzidas por UVC em genomas irradiados.

• Mutações induzidas por UV na linhagem *imuC* mostram uma organização espacial similar à encontrada na linhagem NA1000.

Nossos dados de sequenciamento genômico da linhagem deficiente em TLS ainda são preliminares, pois o número de mutações encontradas é pequeno, necessitando de sequenciamentos adicionais. Porém, a distribuição das substituições encontradas neste genoma parece seguir o mesmo padrão encontrado no cromossomo da cepa selvagem. Isso pode ser um indicativo de que a organização espacial das mutações é consequência da diferença de eficiência de mecanismos de reparo ao longo do cromossomo, e não apenas um viés relacionado a TLS mediada por ImuC.

6 PESPECTIVAS

A mutagênese induzida por *imuABC* é um fenômeno recente na literatura científica, e difere de outros modelos de TLS elaborados para *E. coli*. Os primeiros passos para entender como a proteína ImuC faz o *bypass* de lesões foi dado em conjunto pelo nosso grupo e de estudos com *M. tuberculosis*. A compreensão genética é importante, porém a bioquímica pode abrir portas sobre o papel individual de cada uma destas três proteínas que fazem parte do cassete mutagênico durante a TLS.

Com a caracterização genética feita, o próximo passo é focar na ação destas proteínas a nível bioquímico para tentar entender como funciona o *bypass* de lesões mediado por estas proteínas.

Al Mamun, A. A., Humayun, M. Z. 2006. *Escherichia coli* DNA polymerase II can efficiently bypass 3, N(4) - ethenocytosine lesions in vitro and in vivo. Mutation Research. 593: 164 – 176.

Alves I. R., Lima-Noronha, M. A., Silva, L. G., Fernández-Silva, F. S., Freitas, A. L. D., Marques, M. V., Galhardo, R. S. 2017. Effect of SOS-induced levels of imuABC on spontaneous and damage-induced mutagenesis in *Caulobacter crescentus*. DNA Repair (Amst).59:20-26. doi: 10.1016/j.dnarep.2017.09.003.

Arias-Cartin, R., Dobihal, G. S., Campos, M., Surovtsev, I. V., Parry, B., Jacobs-Wagner, C. 2016. Replication fork passage drives asymmetric dynamics of a critical nucleoid-associated protein in Caulobacter. The EMBO Journal 36, 301-318. doi: 10.15252/embj.201695513.

Badrinarayanan, A., Le, T. B., and Laub, M. T. 2015. Bacterial chromosome organization and segregation. Annual Review of Cell and Developmental Biology. 31: 171–199. <u>http://doi.org/10.1146/annurev-cellbio-100814-125211.</u>

Bagg, A., Kenyon, C. J., & Walker, G. C. (1981). Inducibility of a gene product required for UV and chemical mutagenesis in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 78(9), 5749–5753.

Becherel, O. J., and Fuchs, R. P. 2001. Mechanism of DNA polymerase II-mediated frameshift mutagenesis. PNAS. 98: 8566 – 8571.

Bjedov, I., Dasgupta, C. N., Slade, D., Le Blastier, S., Selva, M., & Matic, I. 2007. Involvement of *Escherichia coli* DNA Polymerase IV in Tolerance of Cytotoxic Alkylating DNA Lesions *in Vivo*. *Genetics*, *176*(3), 1431–1440. http://doi.org/10.1534/genetics.107.072405.

Bos, J., Yakhnina, A. A., & Gitai, Z. (2012). BapE DNA endonuclease induces an apoptotic-like response to DNA damage in *Caulobacter*. Proceedings of the National Academy of Sciences of the United States of America, 109(44): 18096–18101. http://doi.org/10.1073/pnas.1213332109.

Boshoff, H. I., Reed, M. B., Barry, C. E. 3rd, Mizrahi, V. 2003. DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*.Cell. 18;113(2):183-93.

Bridges, B. A., and Woodgate, R. 1985. Mutagenic repair in *Escherichia coli*: products of the recA gene and of the *umuD* and *umuC* genes act at different steps in UV-induced mutagenesis. Proceedings of the National Academy of Sciences of the United States of America, 82(12): 4193–4197.

Bridges, B. A., and Woodgate, R. 1985. The two-step model of bacterial UV mutagenesis. Mutat Res. 150(1-2): 133 – 9.

Bruck, I., Goodman, M. F., O'donnell, M. 2003. The Essential C Family DnaE Polymerase Is Error-prone and Efficient at Lesion Bypass. J Biol Chem. 7; 278(45): 44361 – 8. doi: 10.1074/jbc.M308307200.

Bruck, I., Woodgate, R., McEntee, K., Goodman, M. F. 1996. Purification of a soluble UmuD'C complex from *Escherichia coli*. Cooperative binding of UmuD'C to single-stranded DNA. The Journal of Biological Chemistry. 271: 10767 – 10774.

Burckhardt, S. E., Woodgate, R., Scheuermann, R. H., & Echols, H. 1988. UmuD mutagenesis protein of Escherichia coli: overproduction, purification, and cleavage by RecA. *Proceedings of the National Academy of Sciences of the United States of America*, 85(6), 1811–1815.

Christen, B., Abeliuk, E., Collier, J. M., Kalogeraki, V. S., Passarelli, B., Coller, J. A., Shapiro, L. 2011. The essential genome of a bacterium. Molecular Systems Biology. 7,528. <u>http://doi.org/10.1038/msb.2011.58</u>.

Cirz, R. T., O'Neill, B. M., Hammond, J. A., Head, S. R., & Romesberg, F. E. 2006. Defining the *Pseudomonas aeruginosa* SOS Response and Its Role in the Global Response to the Antibiotic Ciprofloxacin. Journal of Bacteriology, 188(20): 7101–7110. http://doi.org/10.1128/JB.00807-06.

Collier, J. 2012. Regulation of chromosomal replication in *Caulobacter crescentus*. Plasmid. <u>V:67(2)</u>. 76 – 87. <u>https://doi.org/10.1016/j.plasmid.2011.12.007</u>.

Costes, A., and Lambert, S. A. E. 2013. Homologous Recombination as a Replication Fork Escort: Fork-Protection and Recovery. Biomolecules. 3(1): 39 - 71. http://doi.org/10.3390/biom3010039.

Courcelle, J., Khodursky, A., Peter, B., Brown, P. O., and Hanawalt, P. C. 2001. Comparative gene expression profiles following UV exposure in wild-type and SOS-deficient *Escherichia coli*. Genetics. 158(1): 41–64.

Dewar, J. M., and Walter, J. C. 2017. Mechanisms of DNA replication termination. Nature Reviews Molecular Cell Biology. 18: 507 – 516. doi:10.1038/nrm.2017.42.

Dietrich, M., Pedró, L., García, J., Pons, M., Hüttener, M., Paytubi, S., Madrid, C., Juárez, A. 2014. Evidence for Moonlighting Functions of the θ Subunit of Escherichia coli DNA Polymerase III. J. Bacteriol.. 196(5): 1102 – 1112. doi: 10.1128/JB.01448-13.

Dillon, S. C., Dorman, C. J. 2010. Bacterial nucleoid-associated proteins, nucleoid structure and gene expression. Nat Rev Microbiol. 8 (3):185-95. doi: 10.1038/nrmicro2261.

Drake, J.W., Charlesworth, B., Charlesworth, D., Crow, J.F. 1998. Rates of spontaneous mutation. Genetics. 148: 1667 – 1686.

Echols, H, and Goodman, M. F. 1990. Mutation induced by DNA damage: a many protein affair. Mutation Research/DNA Repair. <u>36:</u> 301-311.

Ely, B., Gibbs, W., Diez, S., and Ash, K. 2015. The *Caulobacter crescentus* transducing phage Cr30 is a unique member of the T4-like family of myophages. Current Microbiology. 70(6): 854–858. <u>http://doi.org/10.1007/s00284-015-0799-5</u>.

Ennis, D. G., Levine, A. S., Koch, W. H., Woodgate, R. 1995. Analysis of recA mutants with altered SOS functions, Mutat. Res. Repair. 336: 39 – 48, http://dx.doi.org/10.1016/0921-8777(94)00045-8.

Erill, I., Campoy, S., and Barbé, J. 2007. Aeons of distress: an evolutionary perspective on the bacterial SOS response. FEMS Microbiology Reviews. 31: 637–656. doi:10.1111/j.1574-6976.2007.00082.x.

Evinger, M., and Agabian, N. 1977. Envelope-associated nucleoid from *Caulobacter crescentus* stalked and swarmer cells. Journal of Bacteriology. 132(1): 294 – 301.

Fleming, A. 1929. On the Antibacterial Action of Cultures of a Penicillium, with Special Reference to their Use in the Isolation of *B. influenzæ*. British Journal of Experimental Pathology. 10(3): 226 - 236.

Foster, P. L. 2006. Methods for Determining Spontaneous Mutation Rates. Methods in Enzymology. 409: 195 – 213. <u>http://doi.org/10.1016/S0076-6879(05)09012-9</u>.

Foster, P. L., Hanson, A. J., Lee, H., Popodi, E. M., and Tang, H. 2013. On the Mutational Topology of the Bacterial Genome. G3: Genes|Genomes|Genetics. 3(3): 399–407. <u>http://doi.org/10.1534/g3.112.005355</u>.

Friedberg, E. C., Walker, G. C., Siede, W., Wood, R. D., Schultz, R. A., Ellenberger, T. 2006. DNA Repair and Mutagenesis. Ed. AMS Press. p 9 – 47. Washington, DC.

Fuchs, R. P. 2016. Tolerance of lesions in *E. coli*: Chronological competition between Translesion Synthesis and Damage Avoidance. **DNA Repair.** (Amst). 44:51-58. doi: 10.1016/j.dnarep.2016.05.006.

Fuchs, R. P., and Fujii, S. 2007. Translesion synthesis in *Escherichia coli*: lessons from the NarI mutation hot spot. DNA Repair. 6: 1032 – 1041.

Fuchs, R. P., Fujii, S., and Wagner, J. 2004. Properties and functions of *Escherichia coli*: pol IV and pol V. Advances in Protein Chemistry. 69: 229 – 264.a

Fujii, S., and Fuchs, R. P. 2004. Defining the position of the switches between replicative and bypass DNA polymerases. The EMBO Journal. 23(21): 4342 – 4352. http://doi.org/10.1038/sj.emboj.7600438. b.

Fujii, S., and Fuchs, R. P. 2009. Biochemical basis for the essential genetic requirements of RecA and the β -clamp in Pol V activation. Proceedings of the National Academy of Sciences of the United States of America. 106(35): 14825 – 14830. http://doi.org/10.1073/pnas.0905855106.

Fulconis, R., Mine, J., Bancaud, A., Dutreix, M., and Viovy, J-L. 2006. Mechanism of RecA-mediated homologous recombination revisited by single molecule nanomanipulation. 25(18): 4293 4304. The EMBO Journal. http://doi.org/10.1038/sj.emboj.7601260.

Furusawa, M. 2012. Implications of fidelity difference between the leading and the lagging strand of DNA for the acceleration of evolution. Frontiers in Oncology. 2. 144. http://doi.org/10.3389/fonc.2012.00144.

Galhardo, R. S., Rocha, R. P., Marques, M. V., Menck, C. F. M. 2005. An SOS-regulated operon involved in damage-inducible mutagenesis in *Caulobacter crescentus*. Nucleic acid research. 33(8): 2603 – 2614.

Gawel, D., Maliszewska – Tkaczyk, M., Jonczyk, P., Schaaper, R. M., Fijalkowska, I. 2002. Lack of strand bias in UV-induced mutagenesis in *Escherichia coli*. Journal of bacteriology. P 4449 – 4454.

Godoy, V. G., Jarosz, D. F., Simon, S. M., Abyzov, A., Ilyin, V., & Walker, G. C. 2007. UmuD and RecA Directly Modulate the Mutagenic Potential of the Y-family DNA Polymerase DinB. Molecular Cell. 28(6): 1058 – 1070. http://doi.org/10.1016/j.molcel.2007.10.025.

Goodman, M. F. 2014. The Discovery of Error-prone DNA Polymerase V and Its Unique Regulation by RecA and ATP. The Journal of Biological Chemistry., 289(39): 26772–26782. http://doi.org/10.1074/jbc.X114.607374.

Goodman, M. F., and Woodgate, R. 2013. Translesion DNA Polymerases. Cold Spring Harbor Perspectives in Biology. 5(10): <u>http://doi.org/10.1101/cshperspect.a010363</u>.

Goosen, N., Moolenaar, G. F. 2007. Repair of UV damage in bacteria. DNA Repair (Amst); 7(3): 353-79. doi: <u>10.1016/j.dnarep.2007.09.002</u>.

Grzesiuk, E., Janion, C. 1996. MMS-induced mutagenesis and DNA repair in *Escherichia coli* dnaQ49: contribution of UmuD' to DNA repair. Mutat Res. 15: 362 (2): 147-54.

Gutierrez, A., Laureti, L., Crussard, S., Abida, H., Rodríguez-Rojas, A., Blázquez, J., Baharoglu, Z., Mazel, D., Darfeuille, F., Vogel, J., Matic, I. 2013. β -Lactam antibiotics promote bacterial mutagenesis via an RpoS-mediated reduction in replication fidelity, Nat. Commun. 4 - 1610, <u>http://dx.doi.org/10.1038/ncomms2607</u>.

Hall, B. M., Ma, C.-X., Liang, P., & Singh, K. K. 2009. Fluctuation AnaLysis CalculatOR: a web tool for the determination of mutation rate using Luria–Delbrück fluctuation analysis. Bioinformatics 25(12): 1564–1565. http://doi.org/10.1093/bioinformatics/btp253.

Hamdan, S. M., Loparo, J. J., Takahashi, M., Richardson, C. C., & van Oijen, A. M. 2009. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature*, *457*(7227), 336–339. <u>http://doi.org/10.1038/nature07512</u>.

Hanahan, D. 1989. Biologically pure *Escherichia coli* cell line which is a *deoR* mutant and which is more transformation efficient with foreign plasmids than *deoR Escherichia coli* cell lines, process for obtaining these cell lines. Methods of use. U.S. patent 4, 851, 348.

Henrici, A. T., and Johnson, D. E. 1935. Studies of Freshwater Bacteria: II. Stalked Bacteria, a New Order of Schizomycetes . Journal of Bacteriology. 30(1): 61–93.

Inamura, E., Katayama, T., Taguchi, S. 2017. Absorption of Low-Dose Ultraviolet Radiation by Mycosporine-like Amino Acids Induced by the Dinoflagellate *Prorocentrum micans*. Plankton and Benthos Research. Vol. 12 No. 1 p. 15-24. http://doi.org/10.3800/pbr.12.15.

Ippoliti, P. J., DeLateur, N. A., Jones, K. M., Beuning, P. J. 2012. Multiple strategies for translesion synthesis in bactéria. Cells. 1: 799 – 831. doi: 10.3390/cells1040795.

Jarosz, D. F, Beuning, P. J., Cohen, S. E., Walker, G. C. 2007. Y-family DNA polymerases in *Escherichia coli*. Trends Microbiol.15(2): 70 - 7.

Jarosz, D. F., Godoy, V. G., Delaney, J. C., Essigmann, J. M., Walker, G. C. 2006. A single amino acid governs enhanced activity of DinB DNA polymerases on damaged templates. Nature letters. 439: 225 – 228.

Jenal, U., and Fuchs, T. 1998. An essential protease involved in bacterial cell-cycle control. The EMBO Journal. 17(19): 5658 – 5669. http://doi.org/10.1093/emboj/17.19.5658.

Jensen, R. B., Wang, S. C., & Shapiro, L. 2001. A moving DNA replication factory in *Caulobacter crescentus*. *The EMBO Journal*, 20(17), 4952–4963. http://doi.org/10.1093/emboj/20.17.4952

Juran, B. D., & Lazaridis, K. N. 2008. Genetics and Genomics of PBC. Clinics in Liver Disease, 12(2): 349–ix. <u>http://doi.org/10.1016/j.cld.2008.02.007</u>.

Kato, T., and Shinoura, Y. 1977. Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. Mol Gen Genet. 156(2): 121 - 31.

Kenyon, C. J., and Walker, G. C. 1980. DNA-damaging agents stimulate gene expression at specific loci in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America. 77(5): 2819–2823.

Kim, S.-R., Maenhaut-Michel, G., Yamada, M., Yamamoto, Y., Matsui, K., Sofuni, T., Ohmori, H. 1997. Multiple pathways for SOS-induced mutagenesis in *Escherichia coli*: An overexpression of *dinB/dinP* results in strongly enhancing mutagenesis in the absence of any exogenous treatment to damage DNA. Proceedings of the National Academy of Sciences of the United States of America. 94(25): 13792–13797.

Kimura, S., Tahira, Y., Ishibashi, T., Mori, Y., Mori, T., Hashimoto, J., and Sakaguchi, K. 2004. DNA repair in higher plants; photoreactivation is the major DNA repair pathway in non-proliferating cells while excision repair (nucleotide excision repair and base excision repair) is active in proliferating cells. Nucleic Acids Research, 32(9), 2760–2767. http://doi.org/10.1093/nar/gkh591.

Kisker, C., Kuper, J., and Van Houten, B. 2013. Prokaryotic Nucleotide Excision Repair. Cold Spring Harbor Perspectives in Biology. *5*(3): a012591. http://doi.org/10.1101/cshperspect.a012591.

Koorits, L., Tegova, R., Tark, M., Tarassova, K., Tover, A., Kivisaar, M. 2007. Study of involvement of ImuB and DnaE2 in stationary-phase mutagenesis in *Pseudomonas putida*. DNA repair. 6: 863 – 868.

Kreuzer, K. N. 2013. DNA Damage Responses in Prokaryotes: Regulating Gene Expression, Modulating Growth Patterns, and Manipulating Replication Forks. Cold Spring Harbor Perspectives in Biology., *5*(11): a012674. http://doi.org/10.1101/cshperspect.a012674.

Krishna, S., Maslov, S., and Sneppen, K. 2007. UV-Induced Mutagenesis in *Escherichia coli* SOS Response: A Quantitative Model. PLoS Computational Biology. 3(3): e41. http://doi.org/10.1371/journal.pcbi.0030041.

Lang, K. S., Hall, A. N., Merrikh, C. N., Ragheb, M., Tabakh, H., Pollock, A. J., Merrikh, H. (2017). Replication-Transcription Conflicts Generate R-Loops that Orchestrate Bacterial Stress Survival and Pathogenesis. *Cell*, *170*(4), 787–799.e18. http://doi.org/10.1016/j.cell.2017.07.044

.Lewis, P. J., Thaker, S. D., and Errington, J. 2000. Compartmentalization of transcription and translation in *Bacillus subtilis*. The EMBO Journal, 19(4): 710–718. http://doi.org/10.1093/emboj/19.4.710.

Li, S., and Waters, R. 1998. *Escherichia coli* Strains Lacking Protein HU Are UV Sensitive due to a Role for HU in Homologous Recombination. Journal of Bacteriology. 180(15): 3750–3756.

Liang-dong Lu(吕亮东), Qing Sun(孙青), Xiao-yong Fan(范小勇), Yi Zhong(钟怡), Yu-feng Yao(姚玉峰), Guo-Ping Zhao(赵国屏) (2010). Mycobacterial MazG Is a Novel NTP Pyrophosphohydrolase Involved in Oxidative Stress Response. The Journal of Biological Chemistry. 285(36): 28076 – 28085. http://doi.org/10.1074/jbc.M109.088872.

Lopes-Kulishev, C. O., Alves, I. R., Valencia, E. Y., Pidhirnyj, M. I., Fernández-Silva, F. S., Rodrigues, T. R., Guzzo, C. R., Galhardo, R. S. 2015. Functional characterization of two SOS-regulated genes involved in mitomycin C resistance in *Caulobacter crescentus*. DNA Repair (Amst). 33:78-89. doi: 10.1016/j.dnarep.2015.06.009.

Lundin, C., North, M., Erixon, K., Walters, K., Jenssen, D., Goldman, A. S. H., and Helleday, T. 2005. Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable *in vivo* DNA double-strand breaks. *Nucleic Acids Research*, *33*(12), 3799–3811. <u>http://doi.org/10.1093/nar/gki681</u>.

Maliszewska-Tkaczyk, M., Jonczyk, P., Bialoskorska, M., Schaaper, R. M., Fijalkowska, I. J. 2000. SOS mutator activity: Unequal mutagenesis on leading and lagging strands. PNAS. vol. 97 no. 23. doi: 10.1073/pnas.220424697.

Mallik, S., Popodi, E. M., Hanson, A. J., and Foster, P. L. 2015. Interactions and Localization of *Escherichia coli* Error-Prone DNA Polymerase IV after DNA Damage. Journal of Bacteriology. *197*(17), 2792–2809. <u>http://doi.org/10.1128/JB.00101-15</u>.

Marchesi, F., Lowrie, D., Cole, S., Colston, M.J., Matter, L., Schopfer, K., Bodmer, T. Detection of rifampicin-resistance mutations in *Mycobacterium tuberculosis*. 1993. The Lancet. 341(8846): 647 – 651.

Marks, M. E., Castro-Rojas, C. M., Teiling, C., Du, L., Kapatral, V., Walunas, T. L., and Crosson, S. 2010. The Genetic Basis of Laboratory Adaptation in *Caulobacter*
crescentus . Journal of Bacteriology. *192*(14): 3678 – 3688. http://doi.org/10.1128/JB.00255-10

Marlen Adler, Mehreen Anjum, Otto G. Berg, Dan I. Andersson, Linus Sandegren; High Fitness Costs and Instability of Gene Duplications Reduce Rates of Evolution of New Genes by Duplication-Divergence Mechanisms, *Molecular Biology and Evolution*, Volume 31, Issue 6, 1 June 2014, Pages 1526–1535, <u>https://doi.org/10.1093/molbev/msu111</u>

Martins-Pinheiro, M., Marques, R. C., and Menck, C. F. 2007. Genome analysis of DNA repair genes in the alpha proteobacterium *Caulobacter crescentus*. BMC Microbiology. 7, 17. <u>http://doi.org/10.1186/1471-2180-7-17</u>.

Martomo, S. A., Yang, W. W., Wersto, R. P., Ohkumo, T., Kondo, Y., Yokoi, M.,Gearhart, P. J. 2005. Different mutation signatures in DNA polymerase η - and MSH6-deficient mice suggest separate roles in antibody diversification. Proceedings of the National Academy of Sciences of the United States of America. 102(24): 8656–8661. http://doi.org/10.1073/pnas.0501852102.

McHenry, C. S. 2011. Bacterial replicases and related polymerases. Current Opinion in Chemical Biology. 15(5): 587–594. <u>http://doi.org/10.1016/j.cbpa.2011.07.018</u>. b.

McHenry, C. S. 2011. Breaking the rules: bacteria that use several DNA polymerase IIIs. EMBO Reports. 12(5): 408 – 414. <u>http://doi.org/10.1038/embor.2011.51</u>. a.

Merrikh, H., Machon, C., Grainger, W. H., Grossman, A. D., Soultanas, P. 2011. Codirectional replication-transcription conflicts lead to replication restart. Nature. 470, 554–557. doi:10.1038/nature09758.

Modell, J. W., Hopkins, A. C., and Laub, M. T. 2011. A DNA damage checkpoint in *Caulobacter crescentus* inhibits cell division through a direct interaction with FtsW. Genes & Development..25(12), 1328–1343. <u>http://doi.org/10.1101/gad.2038911</u>.

Modell, J. W., Kambara, T. K., Perchuk, B. S., and Laub, M. T. 2014. A DNA Damage-Induced, SOS-Independent Checkpoint Regulates Cell Division in *Caulobacter crescentus*. PLoS Biology. 12(10), e1001977. http://doi.org/10.1371/journal.pbio.1001977.

Mohl, D. A., Easter, J., and Gober, J. W. 2001. The chromosome partitioning protein, ParB, is required for cytokinesis in *Caulobacter crescentus*. Molecular Microbiology. 42: 741–755. doi:10.1046/j.1365-2958.2001.02643.x.

Napolitano, R., Janel-Bintz, R., Wagner, J., and Fuchs, R. P. P. 2000. All three SOSinducible DNA polymerases (Pol II, Pol IV and Pol V) are involved in induced mutagenesis. The EMBO Journal. 19(22): 6259–6265. http://doi.org/10.1093/emboj/19.22.6259

Napolitano, R., Janel-Bintz, R., Wagner, J., Fuchs, R. P. P. 2000. All three SOS – inducible DNA polymerases (Pol II, Pol IV e Pol V) are involved in induced mutagenesis. The EMBO jornal. 19(22): 6259 – 6265.

Nierman, W. C., Feldblyum, T. V., Laub, M. T., Paulsen, I. T., Nelson, K. E., Eisen, J., Fraser, C. M. 2001. Complete genome sequence of *Caulobacter crescentus*. Proceedings

of the National Academy of Sciences of the United States of America, 98(7): 4136 – 4141. http://doi.org/10.1073/pnas.061029298.

Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., & Yura, T. 1998. Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese Cedar Pollen, Cryj2, in *Escherichia coli*. Applied and Environmental Microbiology., 64(5), 1694–1699.

O'Donnell, M. 2006. Replisome Architecture and Dynamics in *Escherichia coli*. The Journal Of Biological Chemistry. Vol. 281, NO. 16, pp. 10653–10656. doi: 10.1074/jbc.R500028200.

O'Donnell, M., Langston, L., Stillman, B. 2013. Principles and concepts of DNA replication in bacteria and eukarya. Cold Spring Harbor Perspectives in Biology. 5(7): 10.1101/cshperspect.a010108 a010108. <u>http://doi.org/10.1101/cshperspect.a010108</u>.

Oberto, J., Nabti, S., Jooste, V., Mignot, H., and Rouviere-Yaniv, J. 2009. The HU Regulon Is Composed of Genes Responding to Anaerobiosis, Acid Stress, High Osmolarity and SOS Induction. PLoS ONE., 4(2), e4367. http://doi.org/10.1371/journal.pone.0004367.

Ordonez, H., Uson, M. L., & Shuman, S. 2014. Characterization of three mycobacterial DinB (DNA polymerase IV) paralogs highlights DinB2 as naturally adept at ribonucleotide incorporation. *Nucleic Acids Research*, 42(17), 11056–11070. http://doi.org/10.1093/nar/gku752.

Bjedov, I., Dasgupta, C. N., Slade, D., Le Blastier, S., Selva, M., & Matic, I. (2007). Involvement of *Escherichia coli* DNA Polymerase IV in Tolerance of Cytotoxic Alkylating DNA Lesions *in Vivo*. *Genetics*, *176*(3), 1431–1440. <u>http://doi.org/10.1534/genetics.107.072405</u>.

Blanco M., Herrera G., Collado P., Rebollo J. E., Botella L. M. 1982. Influence of RecA protein on induced mutagenesis. Biochimie. 64(8-9):633-6.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402.

P. Caillet-Fauquet, G. Maenhaut-Michel. 1988. Nature of the SOS mutator activity: genetic characterization of untargeted mutagenesis in Escherichia coli, Mol. Gen. Genet. 213: 491–498, <u>http://dx.doi.org/10.1007/BF00339621</u>.

Paz-Elizur, T., Takeshita, M., Goodman, M., O'Donnell, M., and Livneh, Z. 1996. Mechanism of translesion DNA synthesis by DNA polymerase II. Comparison to DNA polymerases I and III core. Journal of Biological Chemistry. 271: 24662 – 24669.

Pfeifer, G. P., You, Y. H., Besaratinia A. 2005. Mutations induced by ultraviolet light. <u>Mutat Res.</u> 1;571(1-2):19-31.

Pfeifer, G. P., You, Y. H., Besaratinia, A. (2015). Mutations induced by ultraviolet light. Mutat Res. 571(1-2): 19 - 31. Qin, T.-T., Kang, H.-Q., Ma, P., Li, P.-P., Huang, L.-Y., & Gu, B. 2015. SOS response and its regulation on the fluoroquinolone resistance. Annals of Translational Medicine., 3(22), 358. <u>http://doi.org/10.3978/j.issn.2305-5839.2015.12.09</u>.

R. Peng, J. Chen, W. Feng, Z. Zhang, J. Yin, Z. Li, Y. Li. 2017. Error-prone DnaE2 balances the genome mutation rates in *Myxococcus xanthus* DK1622. Front. Microbiol. 8. 1–10, <u>http://dx.doi.org/10.3389/fmicb.2017.00122</u>.

Radman, R. 1975. SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. Basic Life Sci. 5A: 355 - 67.

Rajagopolan, M., Lu, C., Woodgate, R., O'Donnel, M., Goodman, M. F. 1992. Activity of the purified mutagenesis proteins UmuC, UmuD' and RecA in replicative by-pass of an abasic DNA lesions by DNA polymerase III. PNAS. 89: 10777 – 107781.

Rastogi, R. P., Richa, Kumar, A., Tyagi, M. B., & Sinha, R. P. 2010. Molecular Mechanisms of Ultraviolet Radiation-Induced DNA Damage and Repair. Journal of Nucleic Acids. 592980. <u>http://doi.org/10.4061/2010/592980</u>.

Reuven, N. B., Arad, G., Maor –Shashani, A., Livneh, Z. 1999. The mutagenesis protein UmuC is a DNA polymerase activated by UmuD', RecA and SSB and is specialized for translesion replication. The jornal of Biological Chemistry. 274(45): 31763 – 31766.

Ricci, D. P., Melfi, M. D., Lasker, K., Dill, D. L., McAdams, H. H., and Shapiro, L. 2016. Cell cycle progression in *Caulobacter* requires a nucleoid-associated protein with high AT sequence recognition. Proceedings of the National Academy of Sciences of the United States of America., 113(40): E5952 – E5961. <u>http://doi.org/10.1073/pnas.1612579113</u>.

Robinson, A., van Oijen, A. M. 2013. Bacterial replication, transcription and translation: mechanistic insights from single-molecule biochemical studies. Nat Rev Microbiol. 11(5): 303 - 15. doi: 10.1038/nrmicro2994.

Rocha, R. P.; Paquola, A. C. M.; Marques, M. V.; Menck, C. F. M.; Galhardo, R. S. 2008. Caracterization of the SOS regulon of *Caulobacter crescentus*. Journal of Bacteriology. 190(4): 1209 – 1218.

Sale, J. E., Lehmann, A. R., Woodgate, R. 2012. Y-family DNA polymerases and their role in tolerance of cellular DNA damage. Nature Reviews. Molecular Cell Biology, 13(3): 141–152. http://doi.org/10.1038/nrm3289

Sanger, F., Nicklen, S., and Coulson, A. R. 1977. DNA sequencing with chainterminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America. 74(12): 5463 – 5467.

Sassanfar, M, Roberts, J. W. 1990.. Nature of the SOS-inducing signal in *Escherichia coli*. The involvement of DNA replication. J Mol Biol. 5: 212 (1): 79 - 96. doi:10.1016/0022-2836(90)90306-7.

Schlacher, K., Cox, M. M., Woodgate, R., Goodman, M. F. 2006. RecA acts in *trans* to allow replication of damaged DNA by DNA polymerase V. *Nature* 442: 883 - 887. doi:10.1038/nature05042.

Schlacher, K., Goodman, M. F. 2007. Lessons from 50 years of SOS DNA-damageinduced mutagenesis. Nature Reviews Molecular Cell Biology. 8: 587 - 594. doi: 10.1038/nrm2198.

Schrader, J. M., and Shapiro, L. 2015. Synchronization of *Caulobacter Crescentus* for Investigation of the Bacterial Cell Cycle. Journal of Visualized Experiments : JoVE, (98): 52633. Advance online publication. http://doi.org/10.3791/52633.

Schrader, J. M., Li, G.-W., Childers, W. S., Perez, A. M., Weissman, J. S., Shapiro, L., & McAdams, H. H. 2016. Dynamic translation regulation in *Caulobacter* cell cycle control. Proceedings of the National Academy of Sciences of the United States of America., 113(44): E6859 – E6867. <u>http://doi.org/10.1073/pnas.1614795113</u>.

Scotland, M. K., Heltzel, J. M. H., Kath, J. E., Choi, J.-S., Berdis, A. J., Loparo, J. J., & Sutton, M. D. 2015. A Genetic Selection for *dinB* Mutants Reveals an Interaction between DNA Polymerase IV and the Replicative Polymerase That Is Required for Translesion Synthesis. PLoS Genetics., 11(9): e1005507. http://doi.org/10.1371/journal.pgen.1005507.

Simon, R., Priefer, U., Pühler, A. 1983. A Broad Host Range Mobilization System for *In Vivo* Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Nature Biotechnology 1, 784 – 791. doi:10.1038/nbt1183-784.

Sommer, S., Bailone, A., and Devoret, R. 1993. The appearance of the UmuD'C protein complex in *Escherichia coli* switches repair from homologous recombination to SOS mutagenesis. Molecular Microbiology. 10: 963–971. doi:10.1111/j.1365-2958.1993.tb00968.x.

Soultanas, P. 2011. The replication-transcription conflict. Transcription. 2(3): 140–144. http://doi.org/10.4161/trns.2.3.15908.

Strauss, B. S., Roberts, R., Francis, L., and Pouryazdanparast, P. 2000. Role of the *dinB* Gene Product in Spontaneous Mutation in *Escherichia coli* with an Impaired Replicative Polymerase. Journal of Bacteriology. 182(23): 6742 – 6750.

Studier, F. W., Moffatt, B. A. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 189(1): 113 - 30.

Taft-Benz, S. A., Schaaper, R. M. 2004. J. Bacteriol. 186 (9) 2774-2780. doi: 10.1128/JB.186.9.2774-2780.

Tang, M., Shen, X., Frank, E. G., O'Donnell, M., Woodgate, R., Goodman, M. F. 1999. UmuD'₂C is an error-prone DNA polymerase, *Escherichia coli* pol V. PNAS. 96: 8919 – 8924.

Taylor, J. A., Panis, G., Viollier, P. H., and Marczynski, G. T. 2017. A novel nucleoidassociated protein coordinates chromosome replication and chromosome partition. Nucleic Acids Research. 45(15), 8916 – 8929. http://doi.org/10.1093/nar/gkx596.

Thanbichler, M., Iniesta, A. A., & Shapiro, L. (2007). A comprehensive set of plasmids for vanillate- and xylose-inducible gene expression in *Caulobacter crescentus*. *Nucleic Acids Research*, *35*(20), e137. http://doi.org/10.1093/nar/gkm818.

Thanbichler, M., Iniesta, A. A., and Shapiro, L. 2007. A comprehensive set of plasmids for vanillate and xylose - inducible gene expression in *Caulobacter crescentus*. Nucleic Acids Research. *35*(20): e137. http://doi.org/10.1093/nar/gkm818.

Valencia, E. Y., Esposito, F., Spira, B., Blázquez, J., & Galhardo, R. S. (2017). Ciprofloxacin-Mediated Mutagenesis Is Suppressed by Subinhibitory Concentrations of Amikacin in Pseudomonas aeruginosa. *Antimicrobial Agents and Chemotherapy*, *61*(3), e02107–16. http://doi.org/10.1128/AAC.02107-16.

Veaute, X, and Fuchs, R. P. 1993. Greater susceptibility to mutations in lagging strand of DNA replication in Escherichia coli than in leading strand. Science. 261(5121): 598 -600.

Wagner, J., Gruz, P., Kim, S., Yamada, M., Matsui, K., Fuchs, R. P. P., Nohmi, T. 1999. The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA Pol IV, involved in mutagenesis. Molecular Cell. 4: 281 – 286.

Wang, F., & Yang, W. 2009. Structural Insight Into Translesion Synthesis By DNA Pol II. Cell. 139(7), 1279–1289. <u>http://doi.org/10.1016/j.cell.2009.11.043</u>.

Warner, D. F., Ndwandwe, E. D., Abrahams, G. L., Kana, B. D., Machowski, E. E. 2010. Essential roles for *imuA*' and *imuB*- encoded accessory factors in DnaE2-dependent mutagenesis in *Mycobacterium tuberculosis*. PNAS. 1 - 6.

Watson, J. D.; Crick, F. H. C. 1953. A Structure for Deoxyribose Nucleic Acid. Nature. 171: 737-738.

Weng, M., Zheng, Y., Jasti, V. P., Champeil, E., Tomasz, M., Wang, Y., Tang, M. 2010. Repair of mitomycin C mono- and interstrand cross-linked DNA adducts by UvrABC: a new model. Nucleic Acids Research. 38(20): 6976–6984. http://doi.org/10.1093/nar/gkq576.

Witkin, E. M. 1967. The radiation sensitivity of *Escherichia coli* B: A hypothesis relating filament formation and prophage induction. PNAS. 57: 1275 – 1279.

Wyrick, J. J., and Roberts, S. A. 2015. Genomic Approaches to DNA repair and Mutagenesis. DNA Repair. 36. 146–155. <u>http://doi.org/10.1016/j.dnarep.2015.09.018</u>.

Yang, W. & Woodgate, R. 2007. What a difference a decade makes: Insights into translesion DNA synthesis. Proceedings of the National Academy of Sciences of the United States of America. 104(40): 15591–15598. http://doi.org/10.1073/pnas.0704219104.

Yao, N. Y. & O'Donnell, M. 2010. SnapShot: The Replisome. Cell, 141(6): 1088–1088.e1. <u>http://doi.org/10.1016/j.cell.2010.05.042</u>.

Yeeles, J. T. P., Poli, J., Marians, K. J., and Pasero, P. 2013. Rescuing Stalled or Damaged Replication Forks. Cold Spring Harbor Perspectives in Biology. 5(5): a012815. http://doi.org/10.1101/cshperspect.a012815

APÊNDICES

sem irradiar, mostrando o tipo da mutação, posição e o fenótipo				
Locus	Mutação	Posição	Natureza da mutação	
CCNA_00944	Inserção (+G)	1028157	Mudança da matriz de leitura	
CCNA_00987	Deleção (-A)	1068007	Mudança da matriz de leitura	
CCNA_01467	T - C	1582587	S: Fenilalanina por Ácido Glutâmico	
CCNA_01803	T - C	1936766	Mutação silenciosa	
CCNA_03557	C - G	3731598	S: Arginina por Prolina	
CCNA_03876	T - A	4059568	S: Ácido glutâmico por Ácido aspártico	
CCNA_03239, CCNA_03240, CCNA_03241	Deleção (- 2,627 pb)			
S. Substituição de aminoácido				

APÊNDICE 1. Lista de mutações encontradas nos genomas usados como controle sem irradiar. mostrando o tipo da mutação, posição e o fenótipo

S: Substituição de aminoácido



APÊNDICE 2. Mapas dos genomas derivados da linhagem selvagem NA1000 após irradiação com UVC. Mapa NA1000 1 e NA1000 4 foram gerados usando o Geneious 8.1.



APÊNDICE 3. Mapas dos genomas derivados da linhagem selvagem NA1000 após irradiação com UVC. Mapa NA1000 5 e RSG 479 foram gerados usando o Geneious 8.1.



APÊNDICE 4. Mapas dos genomas derivados da linhagem selvagem NA1000 após irradiação com UVC. Mapa RSG 480 e RSG 481 foram gerados usando o Geneious 8.1.



APÊNDICE 5. Mapas dos genomas derivados da linhagem selvagem NA1000 após irradiação com UVC. Mapa RSG 488 e RSG 489 foram gerados usando o Geneious 8.1.



APÊNDICE 6. Mapa de genoma derivado da linhagem selvagem NA1000 após irradiação com UVC. Mapa RSG 490 foi gerado usando o Geneious 8.1.

Mapa Genômico					
Região	Intervalo (pb)	Número de mutações			
9	0 - 252683	3			
10	252684 -505366	1			
11	503367 - 758049	1			
12	758049 - 1010732	1			
13	1010733 - 1263415	7			
14	1263416 - 1516098	4			
15	1516099 - 1768781	1			
16	1768782 - 2021464	6			
1	2021465 - 2274147	3			
2	2274148 - 2526830	1			
3	2526831 - 2779513	3			
4	2779514 - 3032196	15			
5	3032197 - 3284879	1			
6	3284880 - 3537562	4			
7	3537563 -3790245				
8	3790246 - 4042929				
Total		51			

APÊNDICE 7. Localização das mutações representadas no Pan Genoma da NA1000 nos 16 intervalos do mapa cromossômico de *C. crescentus*.

Linhagem	Posição	Fita +	Fita -	Posição da lesão
RSG490	1051154	CTC	GAG	Lagging
RSG490	1118066	CCT	GGA	Lagging
RSG490	1385074	TTG	AAC	Lagging
RSG490	2548615	GGA	CCT	Lagging
RSG490	2727311	CAT	GTC	Lagging
RSG489	513995	CTG	GAC	Lagging
RSG489	1370096	ATC	TAG	Lagging
RSG489	1890603	CATA	GTAT	-
RSG489	1959428	GAAG	CTTC	Leading
RSG489	2009904	TCC	AGG	Leading
RSG488	2733261	TGG	ACC	Lagging
RSG488	2795628	CCG	GGC	Leading
RSG488	2947165	TCA	AGT	Leading
RSG488	3026719	GCC	CGG	Leading
RSG488	3332821			-
RSG479	205800	CAG	GTC	Leading
RSG479	434104	CAG	GTC	Leading
RSG479	1128647	TCA	AGT	Lagging
RSG479	1254082	TAC	ATG	-
RSG479	1515372	GAC	CTG	Leading
RSG479	2368980	CTTC	GAAG	Leading
RSG480	1191154	CCC	GGG	Lagging
RSG480	3510768	GAT	CTA	Lagging
RSG481	170745	GCC	CGG	Lagging
RSG481	1254082	TAC	ATG	-
RSG481	2201742	GAA	CTT	Lagging
NA1000 1	1260469	TAC	ATG	
NA1000 1	1345766	GAT	CTA	Leading
NA1000 1	2053632	TCG	AGC	Leading
NA1000 1	3044829	GCC	CGG	Leading
NA1000 1	3370908	CCG	GGC	Leading
NA1000 2	1196930	GTCC	CATC	Lagging
NA1000 2	1192654	TCC	AGG	Lagging
NA1000 2	1254082	TGC	ACG	
NA1000 2	3026 719	GCC	CGG	Leading
NA1000 2	3026769	ATA	TAT	
NA1000 2	3026791	CTA	GAT	Leading
NA1000 2	3026793	A-AC	T-TC	Lagging

APÊNDICE 8. Lista das vizinhanças dos dímeros e seu posicionamento em relação a Fita + ou Fita – e a determinação da fita contendo a lesão usada como molde para a replicação da região com a mutação

NA1000 2	3026804	ATT	TAA	Leading
NA1000 2	3026808	GCC	CGG	Leading
NA1000 2	3026812	CCG	GGC	Leading
NA1000 2	3026835	ATC	TCG	Leading
NA1000 2	3026852	GCC	CGG	Leading
NA1000 2	3026859	GGC	CCG	Lagging
NA1000 2	3030 603	GCC	CGG	Leading
NA1000 4	1644971	CCT	GGA	Lagging
NA1000 4	2248735	TCG	AGC	Leading
NA1000 4	3347457	CTC	GAG	Leading
NA1000 5	65217	GGATC	CCTAG	Leading
NA1000 5	899256	C-CA	G-GT	Lagging
NA1000 5	2868064	TGA	ACT	Lagging

Rosa: Posição do dímero (Fita + ou Fita -); **Azul:** Lesão posicionada na fita Lagging;

APÊNDICE 9