

**GERAÇÃO DE MODELO CELULAR PARA ESTUDOS  
DA DEFICIÊNCIA DE MUTYH EM CONTEXTOS  
MONO E BIALÉLICO**

**DIMAS PONTES CAFÉ FILHO**

**Dissertação apresentada à Fundação Antônio Prudente para  
obtenção do título de Mestre em Ciências**

**Área de concentração: Oncologia**

**Orientador: Dr. Martín Roffé**

**Co-orientadora: Dra. Dirce Maria Carraro**

**São Paulo**

**2022**

## FICHA CATALOGRÁFICA

**Filho, Dimas.**

**GERAÇÃO DE MODELO CELULAR PARA ESTUDOS DA DEFICIÊNCIA DE MUTYH EM CONTEXTOS MONO E BIALÉLICO. / Dimas Filho. São Paulo, 2022.**

**88f.**

**Dissertação de Mestrado - Fundação Antônio Prudente. Curso de Pós-Graduação em Ciências - Área de concentração: Oncologia.**

**Orientador: Martín Roffé.**

**1. Sistema CRISPR/Cas, 2. Reparo de DNA, 3. Expressão gênica**

**CDU 616**

\*Todos os direitos reservados à FAP. A violação dos direitos autorais constitui crime, previsto no art. 184 do Código Penal, sem prejuízo de indenizações cabíveis, nos termos da Lei nº 9.610/08.

**Nome:** Dimas Pontes Café Filho

**Título:** GERAÇÃO DE MODELO CELULAR PARA ESTUDOS DA DEFICIÊNCIA DE MUTYH EM CONTEXTOS MONO E BIALÉLICO

**Aprovado em:** 20/10/2022

**Banca Examinadora**

Orientador: Dr. Martín Roffé

Instituição: Fundação Antônio Prudente

Co-orientadora: Dra. Dirce Maria Carraro

Instituição: Fundação Antônio Prudente

Membro da banca: Dr. Bryan Eric Strauss

Instituição: Instituto do Câncer do Estado de São Paulo (ICESP)

Membro da banca: Dr. Israel Tojal da Silva

Instituição: Fundação Antônio Prudente

Membro da banca: Dra. Glaucia Noeli Maroso Hajj

Instituição: Fundação Antônio Prudente

**SUPORTE À PESQUISA POR AGÊNCIA DE FOMENTO**

Este trabalho recebeu apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), através de auxílio à Pesquisa - 2019/25453-4

## DEDICATÓRIA

Dedico este trabalho àqueles que me inspiraram durante essa etapa da minha vida, aos meus amigos antigos e novos que me motivaram a seguir em frente. Espero que os esforços com este trabalho ajudem de alguma forma, a trazer mais esperanças aos pacientes oncológicos.

A única prisão real é o medo. E a única  
liberdade real é a liberdade de não ter medo.

Aung San Suu Kyi

## AGRADECIMENTOS

Primeiramente ao meu “Eu” do passado, pela sabedoria e principalmente resiliência que contribuíram e muito para quem eu sou hoje.

Ao meu orientador, Dr. Martín Roffé, pela dedicação e pelas discussões de experimentos.

Agradeço por ter me estimulado a pensar criticamente e por ter contribuído para meu amadurecimento científico e pessoal.

À Dra. Dirce Carraro, por me receber e acolher em seu laboratório e por me auxiliar em diversos momentos. À Dra. Giovanna Torrezan, pela disponibilidade e disposição em dividir seus conhecimentos.

À Dra. Vilma Martins por ver potencial em minha pessoa e ter contribuído para me manter no laboratório de Biologia tumoral e Biomarcadores.

À Denise Yamashita pela paciência em me ensinar do zero, na IC, bom humor e motivação. Papel que foi crucial para minha permanência na pesquisa.

Ao A.C. Camargo Cancer Center, por investir em ciência e ensino. A todos os membros do CIPE, à Pós-Graduação e à Biblioteca. À equipe EAIP, em especial Ester e Gilmara que sempre me socorreram com as burocracias.

Ao Danilo, pela parceria, sempre disposto a me ajudar. A Julia por me auxiliar em diversos momentos. E também ao Felipe que fez parte dessa história. Meu agradecimento maior aos foi pela cumplicidade em todos os momentos, pelos cafés, corridas e apoio emocional. Além de manterem o bom funcionamento do laboratório.

À Barbara pelas conversas e por me ajudar a amadurecer meus pensamentos. À Julia Avian pelo apoio e realismo que me ajudaram a ver com clareza as situações. À Lupinacci pela amizade, apoio e motivação.

À Ferreirinha que através da convivência e amizade que desabrochou com o tempo, resgatou minha parte “louca”. Nossas comemorações mesmo sem motivo aparente foram fundamentais para finalizar os dias sem peso. Foi uma honra ver seu desenvolvimento e continuar compartilhando reflexões e aventuras.

Às “minhas crias” Danielle, Jacqueline e João, que tive o prazer de contribuir com seus desenvolvimentos. E a minha parceira de “crime” do laboratório para a vida, Anny.

Agradeço as minhas parceiras de “*fanfic*” Amanda por contribuir com a ilusão e Ingrid por prever as possibilidades. Além de alegrar os dias no laboratório, apoiaram minha realidade. E ao Pedro pelo companheirismo e por acompanhar cada episódio assiduamente. Tenho orgulho de dizer que temos ótimas histórias para contar.

À Ana Carreta e Emne pelas conversas e momentos de sinceridade e reflexão. Ao meu braço direito, desde a graduação, Dayse e à minha mana Pamela, ambas me fizeram lembrar quem eu sou no momento em que me perdi.

À Patrícia Alves, minha principal referencia de visão holística, quem me educou psicologicamente e emocionalmente, sempre serei grato pela dedicação.

Aos meus amigos de longa data: Leize, Emanuelle, Lucas Alves, Jadson, Thibory Donaire, Vinicius Ernandes, Danny Carvalho, Hayanne, Matheus Martins, Felipe Santana e Álvaro Pieretti pela torcida, confiança e admiração.

Ao Thiago Porfirio (*in memory*) éramos competitivos em tudo, e essa companhia desde a infância sempre me estimulava a fazer o melhor mesmo no caos. Quando faleceu, não havíamos contado um para o outro que entramos no mestrado... fiquei sabendo no velório que também havia passado. Terminei minha dissertação em sua memória.

Agradeço a minha mãe pelo reconhecimento dos meus esforços e principalmente de quem eu sou.

Ao CNPq pelo suporte financeiro (Março/2020 à Maio/2020) e à FAPESP (Junho/2020 a janeiro/2022).

## RESUMO

Café Filho DP. **Geração de modelo celular para estudos da deficiência de *MUTYH* em contextos mono e bialélico.** [Dissertação] São Paulo; Fundação Antonio Prudente; 2022.

Os mecanismos de reparo de DNA são essenciais para proteger e preservar a integridade dos genomas dos organismos vivos. Em humanos, alterações germinativas nos genes de reparo de DNA estão implicadas na predisposição ao câncer. Portadores de variantes patogênicas bialélicas em *MUTYH* têm risco aumentado de desenvolver câncer colorretal, de bexiga e de ovário. No entanto, mutações monoalélicas em *MUTYH* aumentam em mais de nove vezes o risco de desenvolver tumor gástrico e, em menor intensidade, o risco de câncer colorretal e de mama. *MUTYH* é uma glicosilase envolvida no mecanismo de reparo de excisão de bases (BER) e é responsável pelo reconhecimento e excisão de uma adenina pareada erroneamente com uma guanina oxidada (8-oxoG) que não foi reparada corretamente pela OGG1, outra glicosilase envolvida no reparo deste tipo de base modificada. A presença da base modificada 8-oxoG no DNA causa a transversão GC:TA após eventos replicativos do DNA e a ação de *MUTYH* é fundamental para corrigir esse pareamento incorreto e dar tempo à OGG1 reparar lesões 8-oxoG não corrigidas previamente. Os mecanismos envolvidos no risco aumentado de alguns tumores em pacientes portadores de mutações mono e bialélicas ainda são desconhecidos. Através da técnica CRISPR/Cas9 e sua variação CRISPR/Cas9n desenvolvemos construções com guias direcionadas para o éxon 1 ou éxon 2 do gene *MUTYH*, com a finalidade de se estabelecer um modelo em células de câncer gástrico para o estudo da deficiência de *MUTYH* nas condições mono e bialélica. O sequenciamento do DNA dos clones editados validou a eficiência das construções. A análise da cinética de indução da proteína selvagem, através de estresse oxidativo, relevou que *MUTYH* é recrutado rapidamente para o núcleo durante o estresse oxidativo. Desse modo, descrevemos um importante modelo para o entendimento das funções biológicas de *MUTYH* inicialmente em linhagem gástrica e que contribuirá para desvendar os mecanismos biológicos envolvidos na formação de tumores em diferentes tecidos, assim como avaliar a existência de assinaturas mutacionais características.

**Palavras Chave:** 1. Sistemas CRISPR-Cas. 2. Reparo do DNA. 3. Expressão gênica.



## ABSTRACT

Café Filho DP. [Generation of a cellular model for studies of *MUTYH* deficiency in mono and biallelic contexts]. [Dissertação] São Paulo; Fundação Antonio Prudente; 2022.

DNA repair mechanisms are essential to protect and preserve the integrity of the genomes of living organisms. In humans, germline alterations in DNA repair genes are implicated in cancer predisposition. Carriers of a biallelic pathogenic variant in *MUTYH* are at increased risk of developing colorectal, bladder and ovarian cancer. However, monoallelic mutations in *MUTYH* increase by more than nine times the risk of developing gastric tumor and, to a lesser extent, the risk of colorectal and breast cancer. *MUTYH* is a glycosylase involved in the base excision repair (BER) mechanism and is responsible for the recognition and excision of an adenine mispaired with an oxidized guanine (8-oxoG) that has not been correctly repaired by OGG1, another glycosylase involved in the repair of this type of modified base. The presence of the modified 8-oxoG base in the DNA causes the GC:TA transversion after DNA replicative events and the action of *MUTYH* is critical to correct this mismatch, which gives OGG1 time to repair previously uncorrected 8-oxoG lesions. The mechanisms involved in the increased risk of some tumors in patients with mono- and biallelic mutations are still unknown. Through the CRISPR/Cas9 technique and its CRISPR/Cas9n variation, we developed constructs with guides directed to the exon1 or exon2 of the *MUTYH* gene, in order to establish a gastric cancer cellular model for the study of *MUTYH* deficiency in the mono and biallelic contexts. The sequencing of the DNA of edited clones validated the efficiency of the constructions. Analysis of the kinetics of induction of the wild-type protein in response to oxidative stress, revealed that *MUTYH* is rapidly recruited to the nucleus during oxidative stress. Thus, we described an important model for understanding the biological functions of *MUTYH* initially in a gastric cell lineage that will contribute to unravel differences in the biological mechanisms underlying tumors in different tissues and also to evaluate the existence of characteristic mutational signatures.

**Keywords:** 1. CRISPR-Cas systems. 2. DNA repair. 3. Gene expression.

## LISTA DE FIGURAS

<b>Figura 1</b>	Proteína <i>MUTYH</i> e seus domínios funcionais.....	3
<b>Figura 2</b>	Excisão de bases mediada por <i>MUTYH</i> na presença da base modificada 8-oxoG (OG) no DNA.....	8
<b>Figura 3</b>	<i>MUTYH</i> possui múltiplas formas de transcrições que possibilita diversas isoformas .....	10
<b>Figura 4</b>	Sistema CRISPR/Cas9 .....	14
<b>Figura 5</b>	Edição por CRISPR/Cas9-base editors.....	15
<b>Figura 6</b>	Edição por CRISPR/Cas9n .....	16
<b>Figura 7</b>	Mapa circular do plasmídeo RC225939.....	22
<b>Figura 8</b>	Localização das sondas no gene <i>MUTYH</i> .....	27
<b>Figura 9</b>	Corte e purificação dos plasmídeos pSpCas9(BB)-2A-Puro V2.0 (PX459 V2.0) e pSpCas9n(BB)-2A-Puro V2.0 (PX462 V2.0). .....	28
<b>Figura 10</b>	Esquema do plasmídeo PX459 mostrando os sítios de restrição das enzimas BbsI e EcoRI.....	29
<b>Figura 11</b>	Análise de restrição para as colônias com os vetores para a técnica de CRISPR/Cas9.....	30
<b>Figura 12</b>	Análise dos níveis de expressão de <i>MUTYH</i> .....	32

<b>Figura 13</b>	Análise do pool de células AGS modificadas pelo sistema CRISPR/Cas9 .....	33
<b>Figura 14</b>	Análise da expressão de MUTYH em clones de células AGS modificadas por CRISPR/Cas9.....	34
<b>Figura 15</b>	Análise da expressão de MUTYH em clones de células AGS modificadas por CRISPR/Cas9.....	34
<b>Figura 16</b>	Teste de especificidade do anticorpo anti- <i>MUTYH</i> e do plasmídeo para expressão exógena de <i>MUTYH</i> .....	36
<b>Figura 17</b>	Comparação da detecção de <i>MUTYH</i> com anticorpos de diferentes marcas...	37
<b>Figura 18</b>	Detecção de <i>MUTYH</i> nas frações citoplasmáticas e nucleares .....	39
<b>Figura 19</b>	<i>MUTYH</i> nas frações citoplasmáticas e nucleares .....	40
<b>Figura 20</b>	Determinação de <i>MUTYH</i> em extratos fracionados.....	41
<b>Figura 21</b>	PCR para amplificação da região editada pelo sistema CRISPR/Cas9 .....	43
<b>Figura 22</b>	Sequenciamento da região amplificada para o éxon 1 editada pelos sistemas CRISPR/Cas9 e CRISPR/Ca9n.....	44
<b>Figura 23</b>	Sequenciamento da região amplificada para o Exon2 editada pelos sistemas CRISPR/Cas9 e CRISPR/Ca9n.....	45
<b>Figura 24</b>	PCR para amplificação da região editada pelo sistema CRISPR/Cas9 em clones individuais.....	46
<b>Figura 25</b>	Clone knockout heterozigoto da linhagem AGS para o éxon 1 utilizando a ferramenta CRISPR/Cas9n .....	48

**Figura 26** Clone knockout homozigoto da linhagem AGS para o exon1 utilizando a ferramenta CRISPR/ Cas9 ..... 49

## LISTA DE TABELAS

<b>Tabela 1</b>	Alterações monoalélicas e bialélicas de <i>MUTYH</i> e a predisposição ao câncer .....	13
<b>Tabela 2</b>	Linhagens celulares utilizadas e suas aplicações nos experimentos.....	19
<b>Tabela 3</b>	Sequência das guias desenhadas para o gene <i>MUTYH</i> presente no braço curto do cromossomo 1. ....	20
<b>Tabela 4</b>	Sequência dos primers para o gene <i>MUTYH</i> utilizadas para gerar os amplicons ....	24
<b>Tabela 5</b>	Número de clones selecionados e sequenciados de cada grupo.....	47
<b>Tabela 6</b>	Resultado geral dos clones sequenciados e suas respectivas frequências de edição genica .....	47

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>1</b>
1.1	<i>MUTYH</i> e risco de câncer .....	1
1.2	Estresse oxidativo .....	3
1.3	Estresse oxidativo e seu papel em tumores .....	6
1.4	Câncer gástrico .....	11
1.5	O sistema CRISPR/Cas9 para geração de modelos celulares com alterações específicas.....	13
1.6	Justificativa .....	17
<b>2</b>	<b>OBJETIVO</b> .....	<b>18</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b> .....	<b>19</b>
3.1	Cultura de Células .....	19
3.2	Construção dos plasmídeos para a estratégia CRISPR/Cas9 .....	19
3.3	Geração de células nocaute para <i>MUTYH</i> .....	20
3.4	Transfecção.....	21
3.5	Superexpressão de <i>MUTYH</i> .....	21
3.6	Indução de <i>MUTYH</i> por KBRO3.....	22
3.7	Western blot .....	22
3.8	Fracionamento Subcelular .....	23
3.9	Extração e quantificação do DNA genômico .....	24
3.10	Reação em Cadeia da Polimerase.....	24
3.11	Sequenciamento de nova Geração (do inglês next generation sequencing (NGS))	24
<b>4</b>	<b>RESULTADOS</b> .....	<b>26</b>
4.1	Construções dos plasmídeos para gerar células nocaute através da técnica de CRISPR/CAS9 .....	26
4.2	Avaliação da expressão de <i>mutyh</i> em linhagens celulares sob estresse oxidativo	31
4.3	Geração de células nocaute monoalélicas e bialélicas para <i>MUTYH</i> .....	32

4.4	Transfecção com plasmídeo codificando mutyh-MYC-FLAG.....	42
4.5	Extração de DNA genômico para realização de reação de PCR para amplificação da região editada .....	43
4.6	Sequenciamento de amplicon dos pools gerados.....	43
4.7	Sequenciamento dos clones gerados através das ferramentas CRISPR/ CAS9 e CRISPR/CAS9N.....	46
<b>5</b>	<b>DISCUSSÃO.....</b>	<b>51</b>
<b>6</b>	<b>CONCLUSÃO .....</b>	<b>55</b>
<b>7</b>	<b>REFERÊNCIAS.....</b>	<b>56</b>

## **ANEXO**

**Anexo 1** Estadiamento AJCC<sup>a</sup> e definições de TNM

**Anexo 2** Possíveis edições off-targets dos sgrNAs pelo sistema CRISPR

# 1 INTRODUÇÃO

## 1.1 *MUTYH* E RISCO DE CÂNCER

Os mecanismos de reparo de DNA são essenciais para proteger e preservar a integridade dos genomas dos organismos vivos. Em humanos, alterações germinativas nos genes de reparo de DNA estão implicadas na predisposição ao câncer. Portadores de variantes patogênicas bialélicas em *MUTYH* têm risco aumentado de desenvolver câncer colorretal, de bexiga e de ovário (CHATTERJEE e WALKER 2017). *MUTYH* é uma glicosilase conhecida por seu papel no reparo por excisão de base de DNA (BER). É extremamente importante na prevenção de mutações no DNA derivadas da 7,8-diidro-8-oxoguanina (8-oxoG ou OG), que são as principais lesões resultantes da oxidação da guanina. Variantes de *MUTYH* com deficiências na sua atividade levam à acumulação somática de mutações T:A em pares G:C o que eventualmente pode resultar no desenvolvimento de câncer (DE BOER e HOEIJMAKERS 2000; FRIEDBERG 2003). Variantes patogênicas bialélicas do gene *MUTYH* são responsáveis pela síndrome hereditária recessiva chamada de *MUTYH*-associada à polipose (MAP), que é caracterizada pelo altíssimo risco (43-100%) de desenvolvimento de câncer colorretal (CCR). Portadores de alterações monoalélicas do gene *MUTYH* também apresentam um maior risco de desenvolver CCR em relação com a população em geral, embora o risco seja consideravelmente menor em comparação com portadores de alterações bialélicas (WEEDEN e ASSELIN-LABAT 2018). Foi descrito que portadores de variantes patogênicas monoalélicas em *MUTYH* têm um risco aumentado para câncer gástrico, de fígado, de mama e de endométrio, enquanto portadores de variantes patogênicas bialélicas demonstram risco elevado, além de para CCR associado a polipose como comentado acima, também para câncer de ovário e bexiga (RIAZ et al. 2017).

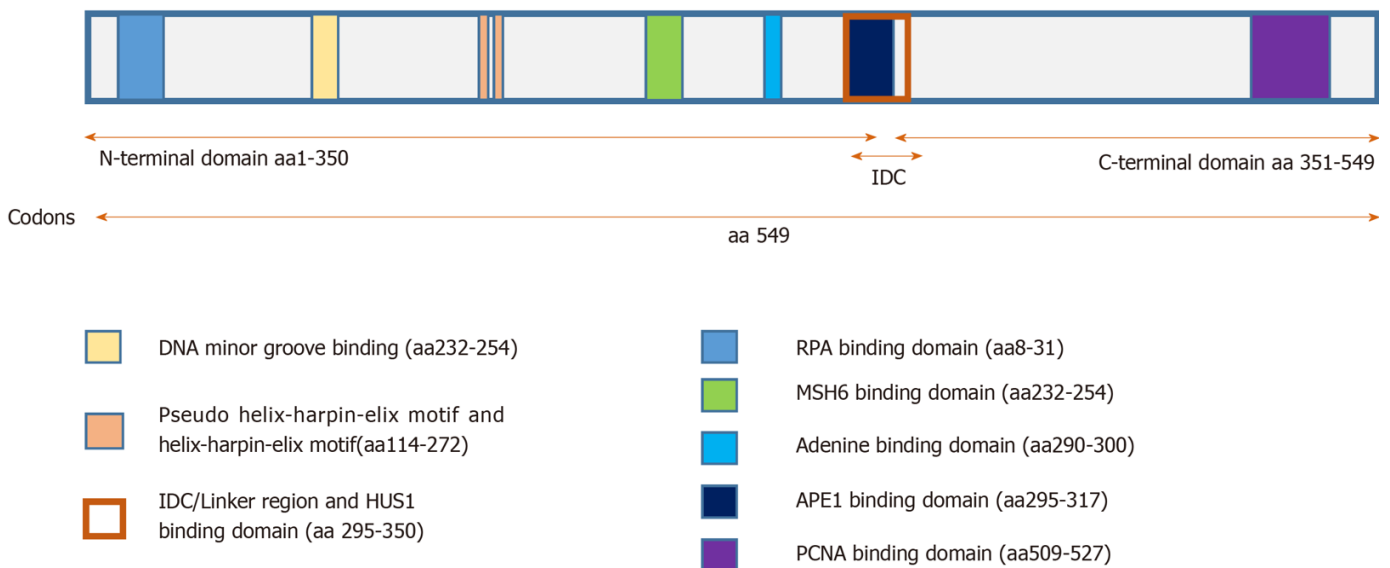
O gene *TP53* é um supressor tumoral responsável por regular a resposta celular diante diversos tipos de estresse, como por exemplo, danos no DNA, ativação de oncogenes e erosão de telômeros. *TP53* é uma proteína que se une ao DNA para regular a transcrição de genes que serão os responsáveis pelas respostas para diversas situações de desequilíbrio homeostático celular. Dessa forma, *TP53* é capaz de regular a resposta das células, mediando



o reparo do DNA ou estimulando a parada de ciclo celular, senescência e eventualmente apoptose (CARVAJAL e MANFREDI 2013). Grande parte dos tumores presentes em humanos possui algum tipo de alteração no gene que codifica TP53 (RIVLIN et al. 2011). Há indícios de que um dos genes alvos de TP53 é o que codifica *MUTYH*. Dessa forma, TP53 poderia regular diretamente a expressão de *MUTYH* através da união a elementos regulatórios no gene *MUTYH*, o que levaria ao aumento da transcrição do mRNA que codifica *MUTYH* (OKA et al. 2008a; OKA e NAKABEPPU 2011).

De fato, em condições de estresse oxidativo, a resposta de TP53 envolveria o aumento de *MUTYH* para primeiramente reparar o acúmulo de 8-oxoG e eventualmente mediar a indução de morte celular, através de um mecanismo que depende de PARP, uma enzima envolvida no reparo do DNA pelos mecanismos NER e BER (do inglês *nucleotide excision repair* e *base excision repair*, respectivamente), e em muitos outros processos celulares, incluindo transcrição e modulação da estrutura da cromatina. É interessante mencionar que a perda da função de TP53 também pode diminuir o reparo mediado pela OGG1, enzima que atua conjuntamente com *MUTYH* (CHATTERJEE et al. 2006a).

Para realizar sua função, *MUTYH* se associa a outras proteínas. Sua estrutura é composta de dois domínios globulares ligados por um conector interdomínios (liga as regiões N- e C-terminal). Na região N-terminal está o sinal de direcionamento mitocondrial (MTS), o domínio de reconhecimento de adenina, o domínio Helix-harpin-helix-GPT (HhH-GPT) e um cluster  $[4\text{Fe-4S}]^{2+}$  que é um cofator de agrupamento de 4 resíduos de cisteína. Em mamíferos, a região conectora interdomínios apresenta 3 resíduos de cisteína com um íon de  $\text{Zn}^{2+}$ , o qual pode desempenhar um papel na coordenação dos domínios N- e C-terminal para o reconhecimento e excisão correto da base adenina em 8-oxoG:A. Modificações pós-traducionais em *MUTYH* podem ser um mecanismo que modula a via de BER. Por exemplo, a fosforilação da Ser524 no sítio de ligação do PCNA (*Proliferating cell nuclear antigen*) modifica a afinidade ao substrato. PARP1 (Poly ADP-ribose polymerase 1) e TP53 podem regular a expressão de *MUTYH* e auxiliar na morte celular mediada pelo gene. O complexo de *mismatch repair* MSH2/MSH6 também apresenta um *crosstalk* com *MUTYH* que interage com SIRT6 (*sirtuin*) que estimula a atividade de PARP-1 por mono-ADP-ribosilação no resíduo da lisina 521 (MAO et al. 2011) (**Figura 1**).



Fonte: CURIA et al. (2020).

**Figura 1** - Proteína *MUTYH* e seus domínios funcionais. RPA: *Replication protein A*; MSH6: *MutS 6 homolog*; APE1: *AP endonuclease 1*

## 1.2 ESTRESSE OXIDATIVO

Em eucariotos, o metabolismo celular para geração de energia é realizado principalmente pelas mitocôndrias durante a respiração aeróbica. Durante este processo são gerados diferentes compostos, normalmente benéficos para a células, necessários para eventos como a transdução de sinal, expressão gênica, ativação de enzimas, formação de ligações dissulfeto durante o dobramento de proteínas e controle da atividade das caspases durante a apoptose. Entretanto, alguns compostos podem se tornar tóxicos quando sua concentração aumenta. As espécies reativas são moléculas formadas inevitavelmente como consequência do metabolismo celular normal e, para evitar que se tornem tóxicas, as células dispõem de mecanismos bioquímicos e genéticos necessários para manter a homeostase. O estresse oxidativo é causado quando ocorre um desbalanço entre a produção das espécies reativas e os sistemas antioxidantes. Tal perturbação pode causar danos em lipídeos, proteína e ácidos nucleicos (DNA e RNA), alterar suas funções e consequentemente levar a diversas doenças. As fontes de estresse oxidativo podem ser endógenas, como as células inflamatórias, peroxissomos e mitocôndrias, ou exógenas como a radiação UV, compostos químicos, poluentes, uso de cigarros e álcool (PIZZINO et al. 2017; SOSA et al. 2013).

As espécies reativas podem ser classificadas em quatro grupos de acordo com seu átomo principal: as espécies reativas de oxigênio (ROS do inglês *Reactive Oxygen Species*), espécies reativas de nitrogênio (RNSs do inglês *Reactive Nitrogen Species*), espécies reativas de enxofre (ERE) e espécies reativas de cloreto (ERC), as quais apresentam diferentes propriedades físico-químicas e meias-vida que variam de nanossegundos até horas, dependendo da estabilidade da molécula (HALLIWELL; GUTTERIDGE 2017). As ROS são os compostos mais abundantemente derivados do metabolismo oxidativo. Entre essas espécies estão o radical hidroxila ( $\bullet\text{OH}$ ), considerado o mais oxidante, o ânion superóxido ( $\text{O}_2^-$ ), e o peróxido de hidrogênio ( $\text{H}_2\text{O}_2$ ), considerado o oxidante mais fraco (BUETTNER 1993; SIMIC; BERGTOLD; KARAM 1989). As principais fontes endógenas de ROS são o vazamento de elétrons durante a cadeia respiratória nas mitocôndrias, a redução de  $\text{O}_2$  pelas NADPH (nicotinamida adenina dinucleotídeo fosfato) oxidases e o processo de formação de ligações dissulfeto durante o dobramento de proteínas no retículo endoplasmático (BROWN e GRIENGLING 2009; HOLTERMAN et al. 2010; MURPHY 2009; YE et al. 2017). Células inflamatórias produzem ROS a partir da redução em oxigênio, onde até 5% do que é metabolizado é convertido em  $\text{O}_2\bullet^-$  (IVANOVA et al. 2013). O óxido nítrico (NO) por exemplo é um RNSs que age como moléculas sinalizadoras. ROS  $\text{O}_2\bullet^-$  em excesso podem danificar ácidos graxos poli-insaturados (PUFAs) e ácidos nucleicos. Quando oxidados, os PUFAs podem atuar como uma sinalização que estimula a inflamação, apoptose e ferroptose (BREITZIG et al. 2016; DODSON; CASTRO-PORTUGUEZ; ZHANG 2019). Outras fontes de ROS ocorrem a partir da atividade do citocromo P450 (CYP), da monoamina oxidase, xantina oxidase, ciclooxigenase, glicolato oxidase, hidroxíácido oxidase, aldeído oxidase e aminoácido oxidase (SIES; JONES 2020). A membrana celular é rica em lipídios poliinsaturados suscetíveis a ação oxidante das espécies reativas que aumentam a permeabilidade da membrana podendo levar a morte celular (HALLIWELL; CHIRICO 1993). A modificação da estrutura de proteínas é bastante afetada pelo acúmulo de aldeídos e cetonas, por exemplo, que podem ser convertidos em ROS causando alteração ou perda da função proteica (LEVINE 2002).

Os peroxissomos são responsáveis pela maior parte da produção de  $\text{H}_2\text{O}_2$  e  $\text{O}_2^-$  e também por conter enzimas antioxidantes como a catalase (CAT), a superóxido dismutase (SOD) e glutationala peroxidase (GPX) (BIRBEN et al. 2012; SULLIVAN; CHANDEL 2014). Para assegurar que o processo de produção de espécies reativas não cause danos oxidativos, as células possuem uma variedade de sistemas diretos e indiretos que limitam a formação de

ROS/RNSs ou que removem os metabólitos gerados pelas espécies reativas. Algumas moléculas pequenas não catalíticas eliminam diretamente as espécies reativas, incluindo a bilirrubina endógena, o ácido  $\alpha$ -lipóico, a melatonina, a glutationa (GSH), as vitaminas E, C e A, polifenóis, entre outros (HALLIWELL; GUTTERIDGE 2017). Entre os antioxidantes diretos catalíticos estão a cobre/zinco superóxido dismutase (SOD1), o manganês, o superóxido dismutase mitocondrial (SOD2), e a EC-SOD extracelular (SOD3) que eliminam o  $O_2^{\bullet-}$ , produzindo  $O_2$  e  $H_2O_2$  (SHENG et al. 2014). SOD1 e SOD2 são importantes na ação antioxidante e já foram anteriormente relatados como supressores tumorais (GILL; PISKOUNOVA; MORRISON 2016; HAN et al. 2013; PAPA et al. 2014; SMITH et al. 2011). Assim como as GPXs, as peroxirredoxinas (PRDXs) são responsáveis por reduzir  $H_2O_2$  a  $H_2O$ . As PRDXs são particularmente importantes por serem abundantes, terem uma alta atividade catalítica e reduzir cerca de 90% dos peróxidos celulares (BRIGELIUS-FLOHÉ; MAIORINO 2013; ELKO et al. 2019). Além de diminuir os níveis de ROS, PRDXs e GPXs também contribuem na eliminação de NO e na redução de  $ONOO^-$  (BENHAR 2018). Entre os sistemas antioxidantes indiretos, estão as isoenzimas da heme oxigenase (codificadas por HMOX1/2) que impedem que o heme livre liberado por hemoproteínas durante o processo de estresse oxidativo formem radicais livres (GOZZELINO; JENEY; SOARES 2010); as proteínas que sequestram metais de transição como a ferritina, ferroportina e ceruloplasmina que limitam os níveis intracelulares de ferro livre e suprimem a formação de OH prevenindo danos oxidativos (AROSIO; INGRASSIA; CAVADINI 2009; MANZ et al. 2016; PIETRANGELO 2017); e as enzimas que metabolizam fármacos, como a família da aldeído desidrogenase, a família das aldo-ceto redutases, NADPH quinona oxirredutase, as desidrogenases/redutases de cadeia curta e a família de UDP-glucoronosil transferase, que tem ações antioxidantes indiretas por impedir que produtos de peroxidação lipídica bloqueiem a ação de GSH e também inativam algumas moléculas de sinalização que desencadeiam a apoptose e ferroptose (BREITZIG et al. 2016; DODSON; CASTRO-PORTUGUEZ; ZHANG 2019; HAYES; FLANAGAN; JOWSEY 2005; JIN; PENNING 2007; OPPERMANN 2007; RODRÍGUEZ-ZAVALA et al. 2019; ROWLAND; MINERS; MACKENZIE 2013).

A adaptação molecular das células ao estresse oxidativo ocorre pela ativação de fatores de transcrição que respondem ao status redox celular, como a proteína ativadora 1 (AP-1), o fator induzido por hipóxia  $1\alpha$  (HIF- $1\alpha$ ), o fator 2 relacionado ao fator nuclear eritróide 2 (NRF2), o fator de *heat shock* 1 (HSF1), o fator nuclear  $\kappa B$  (NF- $\kappa B$ ), e o supressor tumoral TP53 (MARINHO et al. 2014). Esses fatores de transcrição provavelmente são ativados

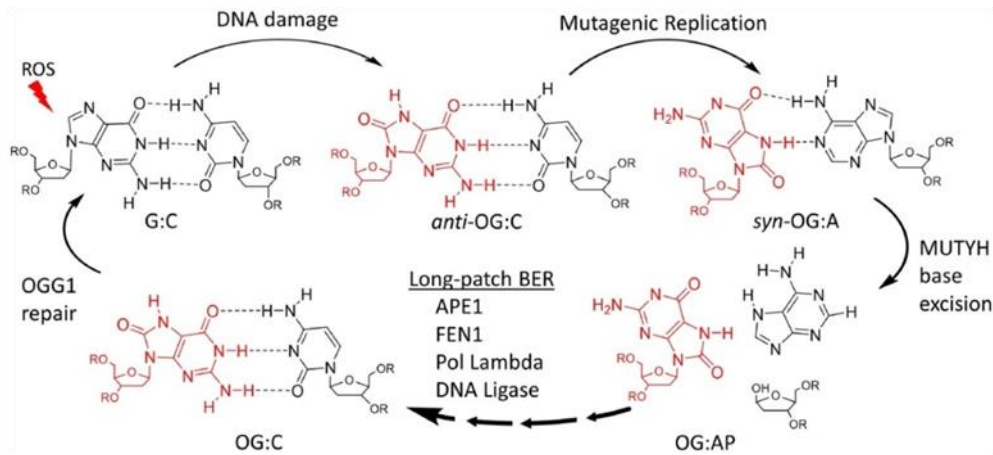
dependendo de quantidades distintas de espécies reativas. Por exemplo, NRF2 é responsável por regular uma ampla gama de genes antioxidantes, sendo a primeira defesa contra o estresse oxidativo seguido por NF- $\kappa$ B e, em seguida, fatores como HIF-1 $\alpha$  (que são ativados primeiramente em hipóxia), as espécies reativas de oxigênio podem atingir níveis bem mais altos, então HIF-1 $\alpha$  aumentaria a resposta antioxidante. Por fim, foi proposto que TP53 age na resposta contra as espécies reativas, quando essas atingem níveis muito mais excessivos, maiores do que os níveis necessários para a ativação dos outros fatores de transcrição, TP53 age parando o ciclo celular e na ativação de vias de morte celular programada (HAYES; DINKOVA-KOSTOVA 2014; SIES; BERNDT; JONES 2017; XIAO et al. 2003).

### **1.3 ESTRESSE OXIDATIVO E SEU PAPEL EM TUMORES**

Do ponto de vista clínico, o estresse oxidativo está relacionado a diversas doenças humanas como doenças neurodegenerativas (p.ex. doença de Alzheimer e doença de Parkinson), doenças inflamatórias, doenças cardiovasculares, alergias, disfunções do sistema imune, diabetes e câncer (SIES 2015). Espécies reativas são produzidas em processos intracelulares e quando o equilíbrio pró- e antioxidante é perturbado, esse estresse pode causar danos no DNA, RNA, lipídios e proteínas. Em concentrações normais, as ROS agem no organismo durante o crescimento, migração e diferenciação celular, porém níveis exacerbados as ROS facilitam a iniciação, promoção e ativação de oncogenes e desestabilizam genes supressores de tumor (KLAUNIG; KAMENDULIS; HOCEVAR 2010). A oxidação do DNA pode gerar a 8-hidroxi-2-desoxiguanosina, um marcador de estresse oxidativo capaz de causar mutações no DNA aumentando o envelhecimento e risco de tumorigênese (MATSUI et al. 2000; VESKOUKIS; TSATSAKIS; KOURETAS 2012). Células inflamatórias como neutrófilos, eosinófilos e macrófagos produzem ROS através da reação de NADPH oxidase. A inflamação crônica causa um aumento considerável da produção de espécies reativas, o que atrai mais células imunes ativadas e contribui com os processos já desregulados, podendo criar uma condição pré-neoplásica favorável (BIRBEN et al. 2012; GUINA et al. 2015). De fato, o processo crônico de inflamação e o estresse oxidativo estão fortemente acoplados e as falhas em controlar esses processos podem causar alterações genéticas e epigenéticas que impulsionam o desenvolvimento de tumores malignos (MURATA et al. 2012). A radiação ionizante e outros

agentes como tabaco, e agentes químicos e infecciosos causam a geração das ROS que podem contribuir para a carcinogênese pela indução de danos ao DNA.

Uma das lesões induzidas por ROS é a formação de bases oxidadas, além de quebra de fita simples e duplas no DNA (IVANOVA et al. 2013). A alteração mais frequente e estável no DNA promovida pela ação de agentes oxidativos é a formação de 8-oxoG e sua alta mutagenicidade se dá pelo fato de parear tanto com a citosina como com a adenina, praticamente com a mesma frequência, resultando em uma transversão, que representa a substituição de uma base purínica (A e G) para uma base pirimídica (C e T) ou vice-versa), ou seja, GC para TA após dois eventos replicativos do DNA (KROKAN; BJORAS 2013; PARKER et al. 2001; TSAI-WU et al. 2000). Essa guanina oxidada é removida pelo mecanismo de reparo por BER. O mecanismo molecular do BER inicia-se, primeiramente, pelo reconhecimento e excisão de bases danificadas pelas DNA glicosilases. Cada glicosilase é específica para um número limitado de bases modificadas. Essas enzimas catalisam a quebra da ligação N-glicosídica entre a base e a desoxirribose, resultando em um sítio AP (local no DNA com ausência de base purina ou pirimidina, denominado sítio apurínico/apirimidínico, de forma espontânea ou devido a danos no DNA), que é reconhecido posteriormente por outro grupo de enzimas, as AP-endonucleases, que fazem a incisão na extremidade 3' ou 5' do sítio AP, gerando uma lacuna. Esta é preenchida através da polimerização e ligação de novos nucleotídeos a sequência de DNA (DENVER 2003) (**Figura 2**). *MUTYH* varre a fita recém sintetizada para localizar e remover a adenina incorporada erroneamente frente a 8-oxoG (par 8-oxoG:A) e *OGG1* que tem a funções de remover 8-oxoG. Ambas foram descritas por trabalhar em conjunto para remover bases mal incorporadas no DNA, mais especificamente por remover 8-oxoguanina (8-oxoG) e evitar consequências mutagênicas que possam ser induzidas pela sua presença (AU et al. 1988; MCLENNAN 1999; MICHAELS; MILLER 1992; SLUPSKA et al. 1996). No caso de mutações em *MUTYH*, essas transversões acontecem em uma alta frequência e resultando em mutações somáticas do tipo nonsense ou sítio de splice em alguns genes (KASTRINOS; SYNGAL 2007).



Fonte: RAETZ; DAVID (2019)

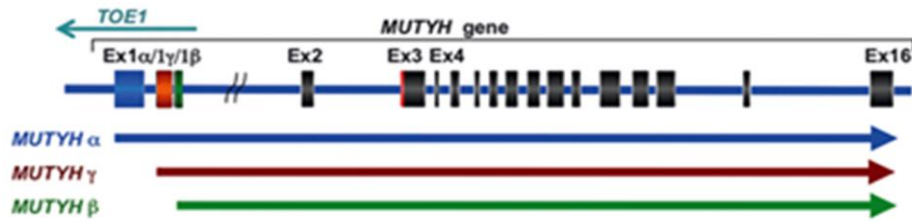
**Figura 2** - Excisão de bases mediada por *MUTYH* na presença da base modificada 8-oxoG (OG) no DNA. 8-oxoG mimetiza bases de timina (T) provocando transversão, as glicosilases OGG1 é responsável por reconhecer e remover 8-oxoG e *MUTYH* reconhece adenina (A) pareada frente a uma base de guanina oxidada dos pares de base 8-oxoG:C e 8-oxoG:A.

Variantes patogênicas germinativas bialélicas estão envolvidas na polipose associada a *MUTYH*, sendo responsáveis pelo aumento no risco de câncer colorretal (AL-TASSAN et al. 2002; JONES 2002; SIEBER et al. 2003). Por fim, foi descoberto que a predisposição a alguns tipos tumorais e outras doenças estavam associados ao acúmulo de 8-oxoG durante o estresse oxidativo em órgãos com *MUTYH* alterado (BANDA et al. 2017).

Em relação a estrutura do gene *MUTYH*, ele é composto por 16 éxons, está localizado no cromossomo 1 (p23.1-p43.3) e codifica uma proteína glicosilase de 65 kDa e 535 aminoácidos (KAIRUPAN; SCOTT 2007). Diferentes transcritos de *MUTYH* já foram identificados, no qual cada um possui uma sequência na região 5' diferente e onde cada transcrito sofre *splicing* alternativo. Os transcritos identificados foram agrupados em três isoformas primárias: *MUTYH*- $\alpha$  (variantes  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$  e  $\alpha_4$ ), *MUTYH*- $\beta$  (variantes  $\beta_1$ ,  $\beta_3$  e  $\beta_5$ ) e *MUTYH*- $\gamma$  (variantes  $\gamma_2$ ,  $\gamma_3$  e  $\gamma_4$ ) (BOLDINOVA et al. 2019; OHTSUBO 2000). A transcrição e *splicing* do pré-mRNA de *MUTYH* produz pelo menos 15 isoformas de mRNA e 9 isoformas diferentes de proteínas (BOLDINOVA et al. 2019). A proteína codificada pelo mRNA das isoformas tipo  $\alpha$  possuem uma sequência de direcionamento mitocondrial (MTS) com 14 resíduos na região N-terminal que são necessários para sua localização na mitocôndria enquanto as proteínas codificadas pelos transcritos tipo  $\beta$  e  $\gamma$  não possuem MTS e tem localização nuclear (TAKAO et al. 1999). Na extremidade 5' de *MUTYH* existe um gene

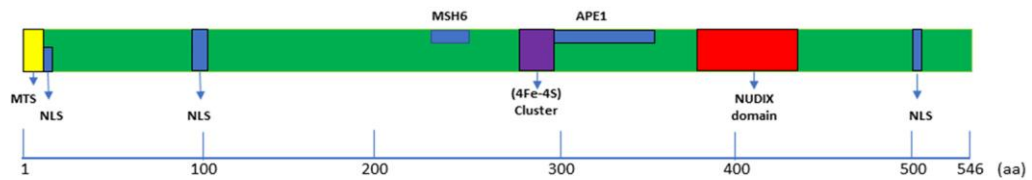
sobreposto, o *TOE1* (target of EGFR member 1) e a transcrição de ambos ocorre em direções opostas (**Figura 3**). O gene *TOE1* inibe a taxa de crescimento celular e induz a expressão de TGF- $\beta$  e do inibidor de CDKN1A (DE BELLE et al. 2003; LARDELLI et al. 2017; OKA; NAKABEPPU 2011).



(A) *MUTYH* gene (1p32.1-p34.3)(B) *MUTYH* transcripts (alternatively spliced)

α5:	CAGCC·ATG·GTCGCTGTGG	GCCATCATG·TGATG	CCTGTGCAGGGATGATTC·GGGGCCCCAGCAG	GCCTG·TTCAGAGAGC
α1:	CAGCC·ATG·GTCGCTGTGG	GCCATCATG·TGATG	GGATGATTC·GGGGCCCCAGCAG	GCCTG·TTCAGAGAGC
α2:	CAGCC·ATG·GTCGCTGTGG	GCCATCATG·TGATG		CAG GCCTG·TTCAGAGAGC
α3:	CAGCC·ATG·GTCGCTGTGG	GCCATCATG·TGATG		GCCTG·TTCAGAGAGC
α4:	CAGCC·ATG·GTCGCTGTGG	GCCATCATG·TGATG		AGAGC
γ1:	TCCCG·GCTCCGGGGAG	GCCATCATG·TGATG	GGATGATTC·GGGGCCCCAGCAG	GCCTG·TTCAGAGAGC
γ2:	TCCCG·GCTCCGGGGAG	GCCATCATG·TGATG	CAG	GCCTG·TTCAGAGAGC
γ3:	TCCCG·GCTCCGGGGAG	GCCATCATG·TGATG		GCCTG·TTCAGAGAGC
γ4:	TCCCG·GCTCCGGGGAG	GCCATCATG·TGATG		AGAGC
βESTs:	TCTCG·CCGCGG	GCCATCATG·TGATG	CCTGTGCAGGGATGATTC·GGGGCCCCAGCAG	GCCTG·TTCAGAGAGC
β1:	TCTCG·CCGCGG	GCCATCATG·TGATG	GGATGATTC·GGGGCCCCAGCAG	GCCTG·TTCAGAGAGC
β2:	TCTCG·CCGCGG	GCCATCATG·TGATG		CAG GCCTG·TTCAGAGAGC
β3:	TCTCG·CCGCGG	GCCATCATG·TGATG		GCCTG·TTCAGAGAGC
β4:	TCTCG·CCGCGG	GCCATCATG·TGATG		AGAGC
β5:	TCCCG·CCCGGGTA·ATTAG	GCCATCATG·TGATG		GCCTG·TTCAGAGAGC

Exon 1 (α, γ, β)                      Exon 2                      Exon 3

(C) Human *MUTYH* proteins

Isoform No (mRNA)	1	15	52	121	MW (aa)	Localization
Isoform 5 (α5)					60,311 (549)	MT
			ACAGMIAECPGAPA			
Isoform 1 (α1)					60,031 (546)	MT
			GMIAECPGAPA			
Isoform 2 (α3)	1	15	52	107	59,034 (535)	MT
			MTS			
Isoform 6 (α2)					59,105 (536)	MT
			A			
Isoform 9 (α4, β4, γ4)					47,000 (429)	MT
Isoform 8 (βESTs)					58,686 (535)	Nu
			ACAGMIAECPGAPA			
Isoform 7 (β1, γ1)					58,407 (532)	Nu
			GMIAECPGAPA			
Isoform 4 (β3, β5, γ3)					57,410 (521)	Nu
Isoform 3 (β2, γ2)					57,481 (522)	Nu
			A			

**Fonte:** Adaptada de OKA; NAKABEPPU (2011).

**Figura 3 - *MUTYH* possui múltiplos transcritos que possibilitam diversas isoformas.**

A) Disposição genômica de *MUTYH*. O gene *MUTYH* produz três transcritos principais, sendo eles:  $\alpha$  (em azul),  $\gamma$  (vermelho escuro),  $\beta$  (verde). (B) As sequências para as regiões 5' (exons 1 a 3) de 15 transcritos diferentes de *MUTYH*. Foram retiradas de:  $\alpha 1$ , NM\_012222.2;  $\alpha 2$ , AB032921.1;  $\alpha 3$ , NM\_001048171.1;  $\alpha 4$ , AB032923.1;  $\alpha 5$ , NM\_001128425.1;  $\gamma 1$ , BI771849.1 (EST);  $\gamma 2$ , NM\_001048172.1;  $\gamma 3$ , NM\_001048174.1;  $\gamma 4$ , AB032929.1;  $\beta 1$ , AB032924.1;  $\beta 2$ , BI768185.1 (EST);  $\beta 3$ , NM\_001048174.1;  $\beta 4$ , BY797353.1 (EST);  $\beta 5$ , AB032926.1;  $\beta$  ESTs, BM911251.1 (EST). Sequências do éxon 3 que codificam aminoácidos extras, são mostradas em vermelho. Os Transcritos  $\alpha 4$ ,  $\gamma 4$  e  $\beta 4$  não possuem 64 nucleotídeos no éxon 3, e  $\beta 5$  possui uma sequência extra (150 nucleotídeos) derivada do íntron 1. (C) Nove isoformas da proteína *MUTYH*. Painel superior: Principais sítios na função da proteína (sinal de direcionamento mitocondrial [MTS], sinal de localização nuclear [NLS], cluster 4Fe-4S e o domínio NUDIX para o reconhecimento de 8-oxoG) (DAVID; O'SHEA; KUNDU 2007; OUT et al. 2010) e os domínios de interação proteína-proteína são mostrados na isoforma 2. Painel inferior: Nove isoformas são mostradas nas regiões amino terminal com resíduos de aminoácidos extras (vermelho). As isoformas 1 á 5 estão nos bancos de dados Entrez Gene e NCBI, e as outras quatro isoformas 6–9 foram nomeadas pela primeira vez por Oka et al.

#### 1.4 ALTERAÇÕES DE *MUTYH* NO CÂNCER GÁSTRICO

O câncer gástrico (CG) é o quinto mais comum tipo de tumor maligno e a quarta causa de morte relacionada a câncer (SUNG et al. 2021), e sua epidemiologia depende da localização anatômica. Os principais fatores de risco associados são a infecção por *H. pylori*, idade avançada, dieta pobre em frutas e vegetais e dietas ricas em sal e comidas processadas, gastrite atrópica crônica, metaplasia intestinal, pólipos gástricos adenomatosos, uso de cigarro, infecção por vírus Epstein-Barr (EBV), histórico na família e algumas síndromes familiares (KURTZ; SHERLOCK 1985; SCHEIMAN; CUTLER 1999). O prognóstico de pacientes com CG está relacionado com o grau tumoral, a extensão tumoral na parede gástrica e o comprometimento de linfonodos (NAKAMUVA et al. 1992; SIEWERT et al. 1998). Pacientes com doença localizada possuem 50% de chances de cura. Porém, uma vez que a doença é diagnosticada em fases mais avançadas, essa chance cai para 10% a 20%. A sobrevida global em 5 anos em pacientes metastáticos, tanto em sítios regionais ou em metástases a distância, pode chegar próximo de zero. Apesar do tratamento para os pacientes com a doença disseminada possa melhorar os sintomas e prolongar a sobrevida, a recidiva da doença é comum (SUNG et al. 2021). Os adenocarcinomas gástricos podem ser classificados em dois

principais tipos: os intestinais e os difusos. O adenocarcinoma intestinal é bem diferenciado e as suas células se arranjam em estruturas tubulares ou glandulares. Termos como tubular, papilar e mucinoso podem ser atribuídos a tipos de adenocarcinomas intestinais. O adenocarcinoma difuso é pouco diferenciado, não apresenta estrutura glandular e pode causar infiltração para a parede gástrica. Alguns tumores podem apresentar características mistas dos tipos intestinais e difusos (HU et al. 2012). A *American Joint Committee on Cancer* (AJCC) classificou o câncer gástrico em quatro graus e os detalhes de cada estadiamento pode ser consultado no **Anexo 1** ao final desta dissertação.

O CG é uma doença heterogênea que apresenta diversos subtipos histológicos e moleculares. Baseado nessas características genômicas, o TCGA (*The Cancer Genoma Atlas*) propôs quatro subtipos de CG: EBV, MSI, GS e CIN (SOHN et al. 2017) que possuem características clinicopatológicas distintas. Entender a biologia tumoral deste tipo de tumor permite acessar um perfil preciso das mutações presentes em cada paciente dentro do microambiente, o que é crucial para auxiliar em decisões terapêuticas e identificar os diferentes prognósticos (DERKS et al. 2020; NAKAMURA et al. 2021).

Em pacientes com CG intestinal avançado positivo para *H. pylori* e metástase linfonodal foram encontradas variantes patogênicas germinativas em *MUTYH* bialélicas e monoalélicas missenses com posterior perda do alelo selvagem no tumor (GUNDERSON; SOSIN 1982). No entanto, pacientes com variantes patogênicas germinativas em *MUTYH* monoalélica podem desenvolver um tumor gástrico esporádico, os quais por exemplo não perdem o alelo selvagem no tumor (KIM et al. 2004). Outros estudos mostraram que pacientes com CG que apresentam baixa expressão de *MUTYH* nos tumores tem uma pior sobrevida quando comparados com os pacientes com tumores com níveis mais altos, podendo atuar como um biomarcador preditivo independente nesses pacientes (SHINMURA et al. 2011).

Apesar da relação de alterações bialélicas (homozigóticas) em *MUTYH* e o desenvolvimento de adenomas colorretais está bem estabelecida na literatura; no caso de variantes patogênicas mutações monoalélicas (heterozigóticas) de *MUTYH* e tumores extracolônicos parece depender muito do histórico familiar, havendo um risco incrementado de cânceres gástricos e hepáticos em relação a população sem alteração no gene, bem como um risco ligeiramente aumentado de cânceres endometrial e de mama (WIN et al. 2016).

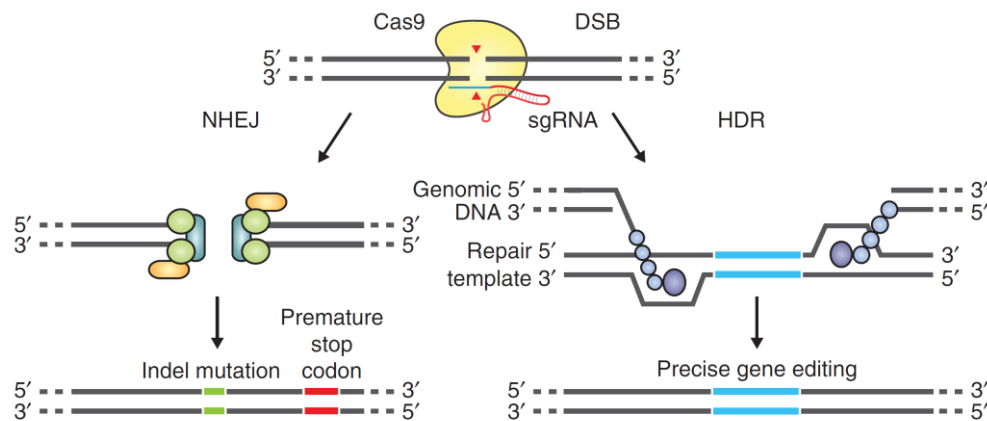
**Tabela 1** - Alterações monoalélicas e bialélicas de *MUTYH* e a predisposição ao câncer. Alterações alélicas no gene *MUTYH* predis põem ao surgimento de diversos tipos de tumores. Dados retirados do artigo científico

Tipo de tumor	Alteração bialélica	Alteração monoalélica
Colorretal	108-129x	1.1- 4.8x
	(43-80%)	(5.6-7.2%)
Mama	Indeterminada	1.0- 1.4x
		(11.0%)
Gástrico	Indeterminada	9.3x
		(2.3-5%)

Fonte: WIN et al. (2016)

### 1.5 O SISTEMA CRISPR/CAS9 PARA GERAÇÃO DE MODELOS CELULARES COM ALTERAÇÕES ESPECÍFICAS

O sistema CRISPR/Cas9 é baseado em um processo imune presente em bactérias chamado CRISPR (*Clustered regularly interspaced short palindromic repeats*)-Cas, que utiliza nucleases guiadas por RNAs para clivar elementos genéticos estranhos (por exemplo, vírus). Esse sistema pode ser reconstituído em células de mamíferos para editar o genoma de maneira direcionada. Cas9 é uma nuclease que é combinada a pequenos RNAs guias (sgRNAs) contendo a sequência da região do DNA de interesse a ser modificado, e quando inserida nas células, causa a clivagem do DNA no local específico de interesse. Essa clivagem sítio-específica induz mecanismos de reparo celular. Um desses mecanismos é o *nonhomologous end joining* (NHEJ), o qual pode resultar em inserções ou deleções (*indel mutations*), alterando a região do DNA clivada (**Figura 4**) (RAN et al. 2013b). Esse sistema é útil para o desenvolvimento de células nocaute, conforme foi utilizado por pesquisadores do laboratório de Biologia Tumoral e Biomarcadores (DE BRITO et al. 2017).



**Fonte:** RAN et al. 2013b

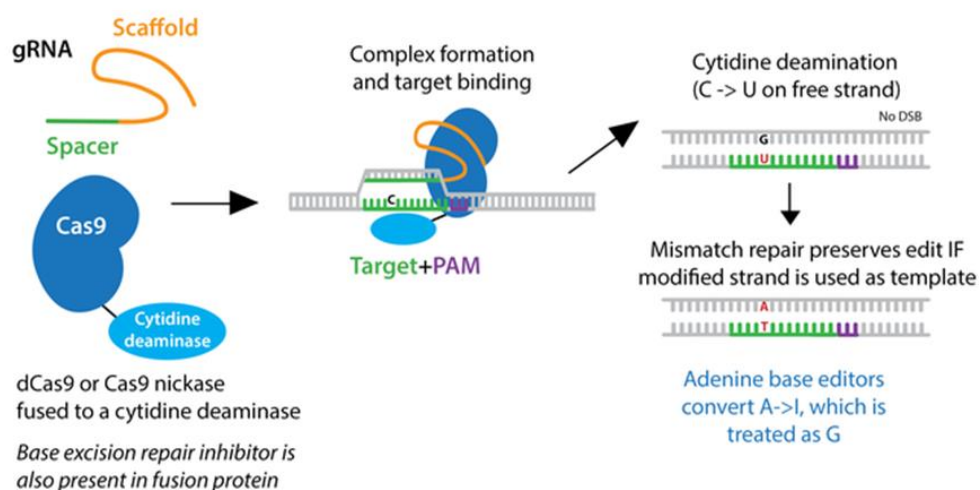
**Figura 4** - Sistema CRISPR/Cas9. Esquema dos mecanismos de reparo induzidos de maneira direcionada pelo sistema CRISPR/Cas9 que levarão à modificação do genoma. É necessário o uso de sgRNAs para direcionar a Cas9 no sítio a ser modificado

Outro mecanismo de reparo do DNA que pode ser explorado para editar o genoma é o *Homology-directed repair* (HDR). Ao contrário do NHEJ, o HDR produz alterações definidas no DNA. No caso do HDR, é necessário adicionar um *repair template* que servirá como molde para o processo de recombinação homóloga. Esse *template* é uma sequência de DNA que contém braços de homologia flanqueando o sítio alvo no DNA e, entre os braços, é incluída a sequência da modificação que será inserida no genoma. Em nosso laboratório foi implementada essa estratégia para gerar células knockin para a mutação de *TP53* p.R337H que é a variante patogênica de TP53 mais frequente no Brasil e está associada com a Síndrome de Li-Fraumeni (GIACOMAZZI et al. 2014). Embora foram geradas células *knockin TP53* p.R337H para um dos alelos, a frequência foi extremamente baixa, impedindo a obtenção dos controles correspondentes. O problema da baixa frequência na indução do mecanismo que precisa de recombinação homóloga foi abordado por vários grupos e diferentes estratégias foram propostas (MARUYAMA et al. 2015), porém nenhuma delas resolveu o problema de maneira eficiente.

Recentemente foi criada uma variação do sistema CRISPR/Cas9 que permite editar bases individuais no DNA genômico (*base editors*) sem o requerimento de recombinação, aumentando a eficiência da técnica (MARX 2018). Esta técnica consiste na fusão de uma Cas9 deficiente para a atividade de nuclease, porém mantendo a função das sgRNAs para alvos específicos no genoma, com enzimas que editarão as bases. Para induzir a conversão C:G>T:A

foi desenvolvido um editor de base composto pela fusão de enzima citidina deaminase de APOBEC1 com Cas9 nickase e a ligação covalente de duas cópias do inibidor de uracila glicosilase (UGI) (**Figura 5**). UGI impede a reversão da modificação introduzida pela citidina deaminase mediada pela uracila N-glicosilase (UNG; um dos mecanismos celulares de BER) (KOMOR et al. 2017). *Adenine base editors* (ADEs) foram gerados para converter A:T>G:C (GAUDELLI et al. 2017). É importante mencionar que existe uma janela de 3 a 9 nucleotídeos, dependendo do sistema, em que os editores podem modificar as bases no sítio do DNA onde foi dirigida a Cas9 pelos sgRNAs. Desse modo, é sempre importante a checagem por sequenciamento das modificações introduzidas.

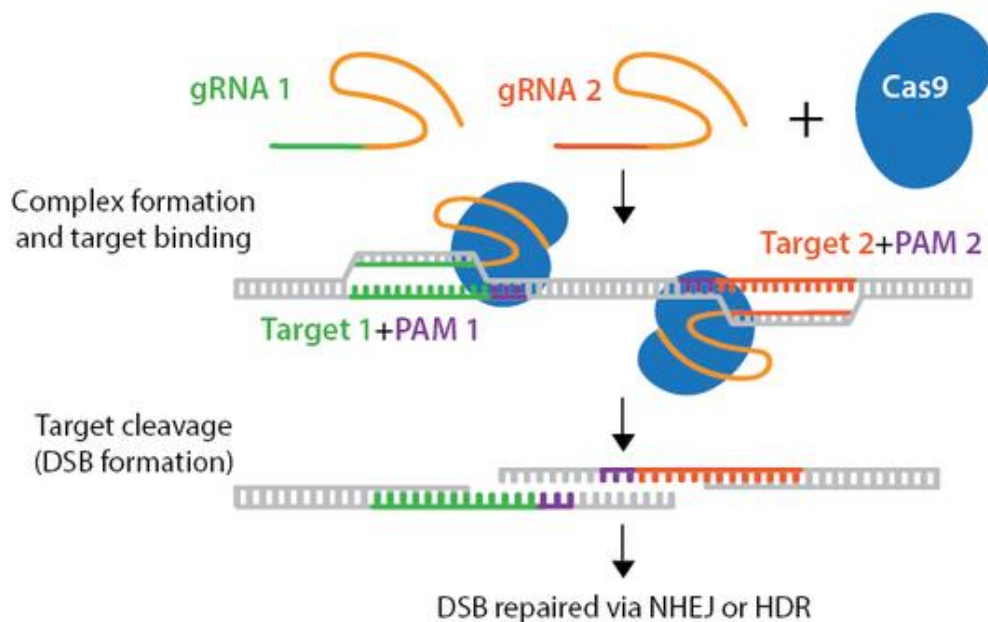
Para funcionar, a Cas9 necessita a presença de uma sequência denominada *Protospacer Adjacent Motif* (PAM) no DNA a ser editado, a que está imediatamente *downstream* da sequência alvo. No caso da Cas9 mais estudada e utilizada, a *S. pyogenes* Cas9 (spCas9), a sequência PAM é NGG. Esse requerimento limita o acesso da spCas9 a algumas regiões do genoma. Para aumentar a acessibilidade da spCas9, foi gerada uma variante que tem uma maior flexibilidade de uso de sequências PAM, denominada XCas9 (HU et al. 2018).



**Fonte:** Adaptada de <https://blog.addgene.org/single-base-editing-with-crispr>.

**Figura 5** - Edição por CRISPR/Cas9-base editors. O spacer é a região do sgRNA complementar para a sequência de DNA genômico que será modificada, além de auxiliar na realização da mutação ao dar suporte à fita de DNA não utilizada enquanto a Cas9 se acopla a uma região PAM do DNA. Com uma fita exposta, a base desejada é modificada de C para U com auxílio da citidina deaminase. Essa mutação é mantida pela UGN (uracil N-glycosylase), gerando uma mutação de C-G para A-T.

Existem diversas variações da técnica CRISPR/Cas9 uma delas utiliza a mutante Cas9 nickase (Cas9n) que possui uma deleção em um dos sítios de clivagem do DNA e dessa maneira introduz quebras de fita simples direcionadas com sgRNA no DNA, em vez das quebras de fita dupla criadas pela enzima Cas do tipo selvagem. Para usar um mutante com nickase, é necessário de dois sgRNAs com fitas complementares/opostas da sequência do DNA alvo, nas proximidades (**Figura 6**). Esses cortes duplos geram um DSB que é reparado usando NHEJ propensa a erros. Essa estratégia de corte duplo reduz os efeitos *off-Target*. Mutantes de nickase também podem ser usados com um modelo de reparo para introduzir edições específicas por meio de reparo por HDR (RAN et al. 2013b).



**Fonte:** Adaptada de <https://www.addgene.org/guides/crispr/#nickase>

**Figura 6** - Edição por CRISPR/Cas9n. Pares de sgRNA adequadamente deslocados guiam Cas9n para cortar simultaneamente ambas as fitas do locus alvo para mediar um DSB, aumentando assim efetivamente a especificidade do reconhecimento do alvo

## 1.6 JUSTIFICATIVA

Na nossa instituição, testes genéticos em genes de predisposição ao câncer são realizados tanto no laboratório da assistência (Laboratório de Diagnóstico Genômico – LDG) como no Grupo de Genômica Clínica e Funcional (GGCF) no nosso centro de pesquisa (CIPE). É comum serem identificados pacientes portadores de variantes genéticas germinativas monoalélicas em *MUTYH*. Em uma análise prévia realizada em centenas de pacientes suspeitos de síndromes de predisposição ao câncer colorretal (não associados a polipose), gástrico e de mama foram detectadas variantes patogênicas monoalélicas em *MUTYH* em 2,2%; 6,4% e 2,1%, respectivamente. Assim, no contexto de variantes patogênicas monoalélicas em *MUTYH*, mesmo havendo mais evidências de risco aumentando no desenvolvimento de câncer gástrico, o entendimento sobre as bases moleculares associadas às diferentes condições alélicas de perda de função de *MUTYH* com o risco aumentado ainda é pouco compreendida.

CRISPR (CRISPR/Cas9) é uma tecnologia recente e versátil que tem a capacidade de fazer alterações o genoma de diferentes organismos com alta especificidade sendo uma ferramenta de biologia molecular e bioquímica bastante poderosa. Sua implementação em um centro de pesquisa de câncer é de enorme valia para contribuir com os estudos relacionados ao câncer. Assim, o desenho desse estudo foca inicialmente na implementação da metodologia CRISPR/Cas9 para entender o papel das variantes patogênicas no gene *MUTYH* na condição monoalélica e bialélica e sua relação com processo tumorigênico do câncer gástrico. Os dados gerados não somente serão importantes para implementação da tecnologia, mas também poderão ser úteis para entender os mecanismos moleculares subjacentes a atuação do *MUTYH* com o alelo inativo em heterozigose ou homozigose em câncer gástrico.



## 2 OBJETIVO

Implementar a técnica de CRISPR/Cas9 como ferramenta para inativação do gene *MUTYH* na condição mono e bialélica usando modelo de câncer gástrico.

### 3 MATERIAL E MÉTODOS

#### 3.1 CULTURA DE CÉLULAS

As linhagens celulares derivadas de adenocarcinoma gástrico AGS (ATCC® CRL-1739™) foram utilizadas na geração do modelo celular com alterações no gene *MUTYH* e experimentos relacionados à indução de *MUTYH* por agente oxidante. Utilizamos a linhagem de carcinoma colorretal HCT116 (ATCC® CCL-246™) como controle de expressão da proteína alvo e a linhagem renal HEK293T (ATCC® CRL-11268™) para controle de eficiência da transfecção. Todas as linhagens foram mantidas nos meios de cultura recomendados pela ATCC.

**Tabela 2** - Linhagens celulares utilizadas e suas aplicações nos experimentos

Linhagem Celular	Tecido	Aplicação
AGS	Gástrico	Geração de modelo celular com deficiência de <i>MUTYH</i> / Indução de <i>MUTYH</i> através de estresse oxidativo
HCT116	Colorretal	Controle de expressão endógena de <i>MUTYH</i>
HEK293T	Renal	Controle da eficiência da transfecção por Lipofectamina 2000

#### 3.2 CONSTRUÇÃO DOS PLASMÍDEOS PARA A ESTRATÉGIA CRISPR/CAS9

Os vetores pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene #62988) e pSpCas9n(BB)-2A-Puro (PX462) V2.0 (Addgene #62987) foram utilizados para incorporação das sequências de DNA que codificam as guias no plasmídeo. Primeiramente, foi realizado um corte com a enzima de restrição BbsI (New England BioLabs #R0539S) e, após o corte, utilizamos fosfatase alcalina intestinal de bezerro (CIP) (New England BioLabs #M0290S) para remover o grupo 5'-fosfato de DNA e prevenir a auto ligação das extremidades. A purificação do DNA foi realizada utilizando o kit *Wizard SV Gel and PCR Clean-UP System* (PROMEGA). Em paralelo, os

oligonucleotídeos codificando as sgRNAs foram fosforiladas nas extremidades 5' pela T4 PNK e posteriormente realizamos sua ligação com o plasmídeo linearizado. Com as sequências dos sgRNAs (**Tabela 3**) inseridas nos plasmídeos. Foram desenhadas guias para o éxon 1 e para o éxon 2 para serem inseridas nas estratégias de edição (CRISPR/Cas9 e CRISPR/Cas9n) pois o éxon 1 é importante para diferenciar as isoformas de *MUTYH* (nucleares), e o éxon 2 está presente em todas as isoformas. Também utilizamos as duas enzimas a Cas9 e Cas9n pois a primeira é mais eficiente, porém com maior probabilidades de efeitos *off-target*, e a Cas9n que reduz os efeitos *off-target*, porém é menos eficiente. Após o processo de ligação do DNA e a transformação das bactérias competentes, cultivadas em meio LB e selecionada com ampicilina (Thermofisher #11593027), escolhemos duas colônias de cada placa e realizamos uma miniprep com o kit *Wizard® Plus SV Minipreps DNA Purification system* (cat A1330, Promega) e submetemos a análise de restrição para checagem. Após a confirmação, foi realizada uma Maxiprep (NucleoBond Xtra Maxi, Macherey – NAGEL; 740414100). As sequências *off-targets* de todas as guias podem ser consultadas no **Anexo 2** no final dessa dissertação.

**Tabela 3** - Sequência das guias desenhadas para o gene *MUTYH* presente no braço curto do cromossomo 1.

Gene	Região Alvo		sgRNA	Sequência	Score
<b>MUTYH</b>	Exon1	Cas9	guide 247	5' CAC CGT CTC CCG CCT GAG TCG TCT G 3' top 5' AAA CCA GAC GAC TCA GGC GGG AGA C 3' bottom	61,69
		Cas9n	guide 221	5' CAC CGC GGG AGA CGA GCG GTG TCA 3' top 5' AAA CTG ACA CCG CTC GTC TCC CGC 3' bottom	54,68
	Exon 2	Ca9	guide 5986	5' CAC CGA GGA AGC CAC GAG CAG CCG T 3' top 5' AAA CAC GCC TGC TGC TGG CTT CCT C 3' bottom	62,14
		Cas9n	guide 6054	5' CAC CGC ATG CTA AGA ACA ACA GTC 3' top 5' AAA CGA CTG TTG TTC TTA GCA TGC 3' bottom	60,39
			guide 5991	5' CAC CGC TTC CTG TGA CCA CTT CCC A 3' top 5' AAA CTG GGA AGT GGT CAC AGG AAG C 3' bottom	60,6

As denominações top e bottom referem-se aos oligonucleotídeos de DNA que são sintetizados para inserir nos plasmídeos correspondentes. O score oferece uma medida da eficiência e da especificidade da guia correspondente (um score maior é melhor).

### 3.3 GERAÇÃO DE CÉLULAS NOCAUTE PARA *MUTYH*

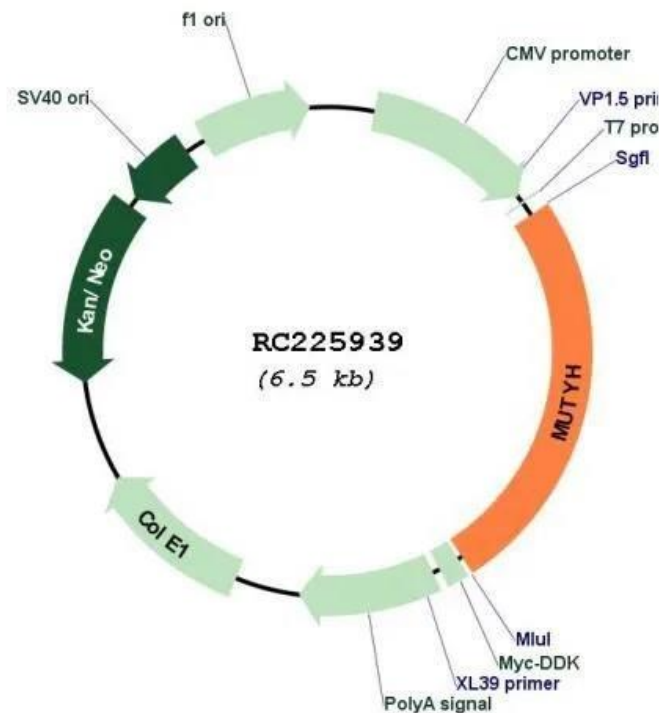
Usamos o sistema de edição genômica CRISPR/Cas9 para gerar células nocaute para o gene *MUTYH*. Para a realização da técnica, utilizamos o plasmídeo pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene ID: 62988) que codifica a nuclease Cas9 selvagem de *Streptococcus pyogenes* (RAN et al. 2013b). Utilizamos também uma estratégia de corte duplo usando o mutante de *nickase* Cas9n, codificada no plasmídeo pSpCas9n(BB)-2A-Puro V2.0 (PX462) V2.0 (Addgene #62987) para minimizarmos a edição fora do alvo. A Cas9 é guiada por sgRNAs com a sequência específica do local a ser modificado no DNA. Os sgRNAs foram desenhados pelo algoritmo <https://www.benchling.com/crispr/> e o DNA codificando as mesmas foi incorporado no plasmídeo para expressão junto com a Cas9. As células AGS foram transfectadas e selecionadas com puromicina, visto que o plasmídeo tem um gene de resistência para esse antibiótico. Após a seleção, isolamos os clones individuais. Esses clones foram checados através de western blot para determinar se foram induzidas células nocaute para *MUTYH*. Posteriormente, as modificações no gene *MUTYH* foram analisadas por sequenciamento.

### 3.4 TRANSFEÇÃO

Utilizamos placas p100 plaqueadas com  $50 \times 10^4$  células. Para cada condição separamos dois tubos falcon com 1 ml de meio Opti-MEM (Invitrogen) em cada. Um tubo recebe 10  $\mu$ l de Lipofectamine 2000 (Invitrogen) e o outro o 8  $\mu$ g do DNA, seguido por incubação por 5 min. Em seguida, o volume contendo o DNA foi inserido no tubo contendo a Lipofectamina 2000 e foi incubado por 20 min para formação de lipossomos. Durante esse tempo, o meio de cultura das células foi trocado para meio Opti-MEM num volume final de 6 ml por placa. Após os 20 min, os 2 ml da solução de lipossomos são aplicados em cada placa. A transfecção foi realizada por 6 h e, após esse tempo, o meio foi trocado pelo meio correspondente da célula com a porcentagem de soro fetal sugerida pela ATCC.

### 3.5 SUPEREXPRESSION DE *MUTYH*

Para controle positivo da expressão de *MUTYH* as células foram transfectadas com 4  $\mu$ g de DNA plasmídial MYH (*MUTYH*) (NM\_001128425) Human Tagged ORF Clone (CAT#: RC225939), variante de transcrição Alfa 5, expressando *MUTYH*-Myc-DDK (Flag) nas células AGS e HEK293T numa concentração de  $20 \times 10^4$ .



**Figura 7** - Mapa circular do plasmídeo RC225939. Mapa do plasmídeo com vetor pCMV6, e suas respectivas enzimas de restrição. Possui gene de resistência à Kanamicina para seleção de colônias após a transformação e o inserto para expressar *MUTYH*-Myc-Flag através da região promotora pCMV6 (Origene).

### 3.6 INDUÇÃO DE *MUTYH* POR $\text{KBrO}_3$

A linhagem AGS expressando *MUTYH* selvagem foram submetidas à uma curva de  $\text{KBrO}_3$  agente altamente oxidante nas concentrações de 100  $\mu\text{M}$ , 200  $\mu\text{M}$ , 300  $\mu\text{M}$ , 400  $\mu\text{M}$  e 500  $\mu\text{M}$  por 24 horas. Também foi utilizado em tratamento de pulso de dano na linhagem celular AGS, na concentração de 20mM por 30 minutos e em seguida recuperadas em meio de cultura recomendado pela ATCC.

### 3.7 WESTERN BLOT

Extratos celulares totais foram obtidos usando o seguinte tampão de lise: 50 mM Tris-HCl pH 8,0; 150 mM NaCl; 1 mM EDTA pH 8,0; 1 mM EGTA pH 8,0; 1% NP-40; 0,5% deoxicolato de sódio; coquetel de inibidores de fosfatase – Thermo Fisher Scientific; coquetel de inibidores

de proteases- Roche. Os extratos totais coletados das linhagens AGS, HCT116 e HEK293T foram centrifugados a 14.000 rpm por 15 min a 4°C e o sobrenadante transferido para novos tubos. Foi realizada a quantificação pela técnica de Bradford. As amostras foram aplicadas em SDS-PAGE, e após a corrida foi feita a transferência para membrana de nitrocelulose. Após a transferência, a membrana foi bloqueada por 1 h com 5% leite em TBS-T. Em seguida, a membrana foi incubada *overnight* com o anticorpo primário a 4°C. Após esse processo, foi colocado o anticorpo secundário acoplado a peroxidase por 1 h e a membrana revelada com kit de quimiluminescência. Foram usados anticorpos para detectar *MUTYH* (Santa Cruz, sc-374571), *MUTYH* (Cell Signaling #6248S) TP53 (Santa Cruz, sc-126), e ERK1/2 (para o *loading control*, já utilizado pelo laboratório de Biologia Tumoral e Biomarcadores).

### 3.8 FRACIONAMENTO SUBCELULAR

As células foram plaqueadas na concentração de  $3 \times 10^6$  em placas de 10mm de circunferência, com volume final de meio em 10ml. No dia seguinte as células foram lavadas com PBS duas vezes e em seguida acrescentar 800 µl de PBS+EDTA 1mM para soltar as células. As células foram raspadas delicadamente sem romper a membrana e coletadas em um eppendorf. Em seguida foi centrifugadas à 3000RPM à 4°C, o sobrenadante foi descartado e acrescentado no eppendorf 300ul do Buffer A (HEPES 10mM PH7,9; NaCl 50mM; Sacarose 0,5M; EDTA 0,1mM; Triton X-100 0,5% e coquetel de inibidor de protease) e ressuspendido levemente e em seguida incubado no gelo por 5 minutos (foi reservado 50 µl e adicionado NP-40 a 0,5% e deoxycholate de sódio a 0,5% em seguida foi centrifugado a máxima velocidade por 15 minutos à 4°C e o sobrenadante foi coletado e congelado como extrato total). Restando 250 µl iniciamos a extração citoplasmática, centrifugando à 400 RPMs (sempre á 4°C) e o sobrenadante foi reservado. O pellet restante foi lavado duas vezes (sem ressuspende) com 400 µl do Buffer B (HEPES 10mM pH7,9; KCL 10mM; EDTA 0,1mM; EGTA 0,1mM e coquetel de inibidor de protease) à 5000 RPM por 10 minutos. O sobrenadante foi descartado e o pellet restante foi ressuspendido com 150 µl do Buffer C (HEPES 10mM pH7,9; NaCl 0,5M; EDTA 0,1mM; EGTA 0,1mM; NP-40 0,1% e coquetel de inibidor de protease) afim de obter o extrato nuclear e iniciado ciclos de vórtex vigorosamente por 1 minuto e mantido à 4°C, esse ciclo de vórtex e gelo foi repetido durante 15 minutos. Em seguida foi centrifugado à máxima

velocidade por 10 minutos à 4°C após esse 30 sobrenadante foi coletado e todas as frações quantificadas por Bradford.

### 3.9 EXTRAÇÃO E QUANTIFICAÇÃO DO DNA GENÔMICO

O DNA genômico de cada clone foi extraído utilizando o kit *DNAEasy* (Qiagen Inc., EUA), seguindo as recomendações do fabricante. A quantificação do DNA e avaliação da pureza após a extração foram realizadas utilizando o espectrofotômetro NanoDrop™ ND-1000 (Thermo Scientific DE, USA) e o fluorômetro Qubit®3.0 (Life Technologies) através do Quant-itds DNA HS assay kit (500) (Life Technologies).

### 3.10 REAÇÃO EM CADEIA DA POLIMERASE

O amplicon do éxon 1 foi desenhado para a região 247-409 (162pb) e para o éxon 2 a região foi 5931-6104 (173 pb). Foram utilizados 20 ng de DNA de cada clone para a reação de PCR e a eficiência da amplificação foi checada em géis de agarose (2%). O produto de PCR foi purificado e processado para a construção de biblioteca de amplicon utilizando o kit *Ion Plus Fragment Library* seguindo as recomendações do fabricante.

**Tabela 4** - Sequência dos primers para o gene *MUTYH* utilizadas para gerar os amplicons.

Primer	Sequência	
Exon1	5' GAA ACT CGC CCA TCG TCA 3'	Forward
	5' ACG GAA GTT CGA CCC ATC 3'	Reverse
Exon2	5' ACG GCT TTG GCT GGG TCT TT 3'	Forward
	5' TTG GGC CAC AAC CTA GTT CC 3'	Reverse

### 3.11 SEQUENCIAMENTO DE NOVA GERAÇÃO (DO INGLÊS NEXT GENERATION SEQUENCING (NGS))

Os amplicons gerados foram tratados com a enzima FuPa para os primers serem parcialmente digeridos e, em seguida, os adaptadores P1 e A juntamente com *barcodes* (que

são sequências únicas capazes de identificar cada amostra após reação de sequenciamento) foram ligados pela enzima Ligase aos amplicons e então, purificados por *beads* magnéticas (Agencourt AMPure XP – Beckman Coulter). As bibliotecas foram quantificadas por PCR em Tempo Real pelo Ion Library TaqMan™ Quantitation Kit (Thermo Fisher Scientific) seguindo a recomendação do fabricante e diluídas a 100 pM. Em seguida, foi realizada a etapa de amplificação da biblioteca com reação de PCR em emulsão utilizando o kit Ion PI™ Hi-Q™ OT2 200 (Thermo Fisher Scientific) na qual ocorre a amplificação clonal dos fragmentos no interior de milhares de micelas geradas nesta solução. O sequenciamento foi então realizado na plataforma Ion S5 ou Proton (Thermo Fisher Scientific) (Thermo Fisher Scientific) seguindo as orientações do fabricante. O sequenciamento é baseado na liberação cíclica dos nucleotídeos em uma ordem conhecida e a incorporação de cada nucleotídeo na cadeia crescente por ligação fosfodiéster libera uma molécula de hidrogênio modificando o pH da solução. A detecção de alterações no pH da solução a cada ciclo do sequenciamento permite a identificação da sequência de bases. O número mínimo de leituras para avaliação da frequência de mutação induzida indicativa de alteração nos alvos do gene *MUTYH* foi de 500 leituras. A chamada de variantes (ou alterações) foi realizada pelo software Torrent Variant Caller (TVC) e as alterações foram analisadas visualmente pelo IGV, que mostra as bases encontradas em cada leitura e suas proporções usando o genoma de referência hg19.



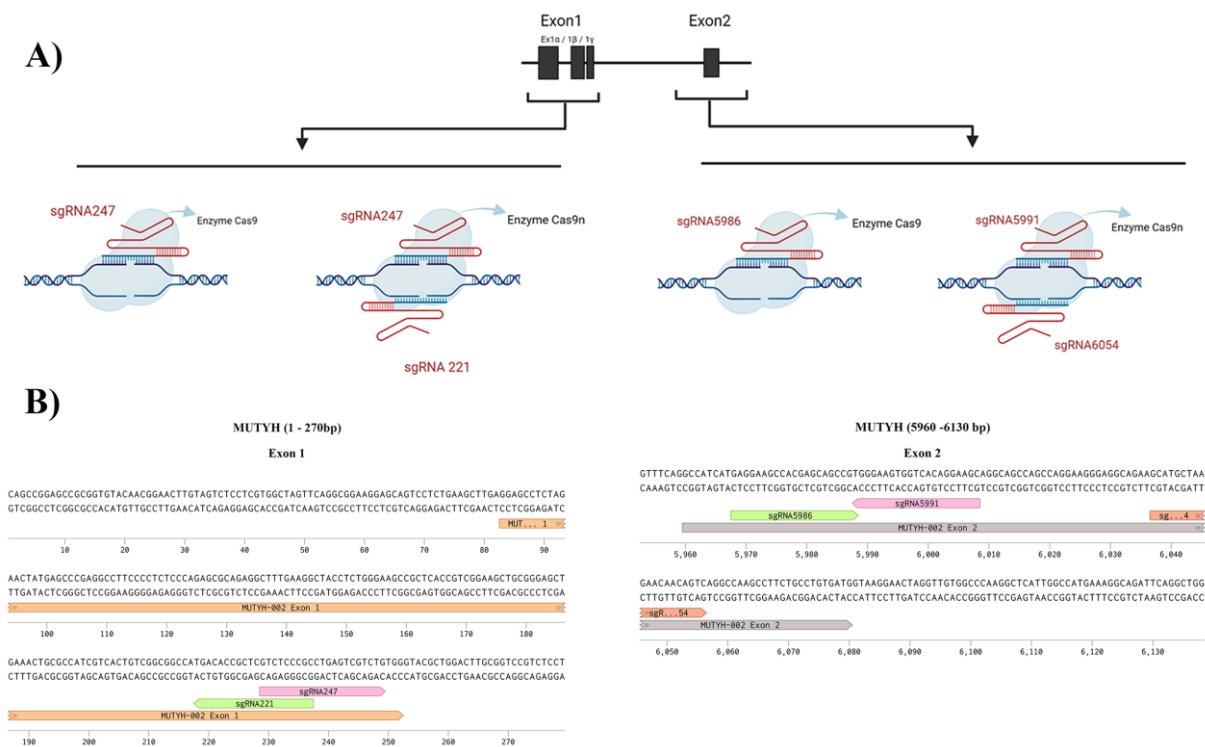
## 4 RESULTADOS

### 4.1 CONSTRUÇÕES DOS PLASMÍDEOS PARA GERAR CÉLULAS NOCAUTE ATRAVÉS DA TÉCNICA DE CRISPR/CAS9

Em células humanas, o gene *MUTYH* transcreve três sequências distintas a partir do éxon 1, produzindo três tipos de transcritos primários,  $\alpha$ ,  $\beta$  e  $\gamma$  (**Figura 3**). Destes três transcritos primários, mais de 15 transcritos são gerados por splicing alternativo no éxon 1 e éxon 3. O mRNA de *MUTYH* do tipo  $\alpha$ 3 é um transcrito principal de *MUTYH* e codifica o *MUTYH* mitocondrial expressa mais abundantemente (TAKAO et al. 1998). Em contraste, *MUTYH* codificado pelo tipo  $\beta$  é a isoforma nuclear mais abundante (OKA; NAKABEPPU 2011; TAKAO et al. 1999).

Para gerar células nocautes para *MUTYH* desenhamos small guide RNAs (sgRNAs) para o exon 1 pois este transcreve principalmente a isoforma tipo  $\beta$  (nuclear) de interesse e para o exon 2, a modo de alterar todas as isoformas. Dessa maneira poderemos conseguir determinar os efeitos das isoformas não nucleares na interpretação dos ensaios funcionais. O DNA codificando os sgRNAs foram inseridas nas construções como parte da estratégia de edição (CRISPR/Cas9 e CRISPR/Cas9n), conforme exemplificado na **figura 8**.

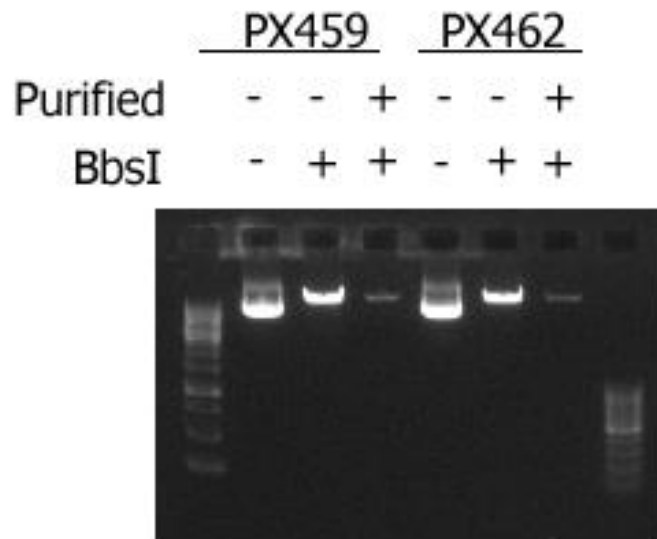
## sgRNA *MUTYH* Location



**Figura 8** - Regiões no gene *MUTYH* reconhecidas pelos sgRNAs A) Organização genômica do gene *MUTYH*; O éxon 1 do gene *MUTYH* produz três tipos de transcritos principais,  $\alpha$ ,  $\beta$  e  $\gamma$  sendo a isoforma  $\beta$  a mais abundante no núcleo. As distintas isoformas tem diferenças no éxon 1 e as sgRNAs desenhadas originalmente foram dirigidas contra a região codificante das isoformas nucleares. Isso não quer dizer que uma alteração no éxon 1 não tenha um efeito nas demais isoformas, porém, para garantir o sucesso da estratégia, adicionalmente desenhemos sgRNAs para uma região codificante no éxon 2 que está presente em todas as isoformas. B) Locais de inserção das guias nas sequências do éxon 1 e do éxon 2 de *MUTYH*; Os sgRNAs para a sequência de DNA genômico modificada são expressas a partir do mesma construção que codifica a /Cas9 ou Cas9n e geram um DSB que é reparado usando NHEJ.

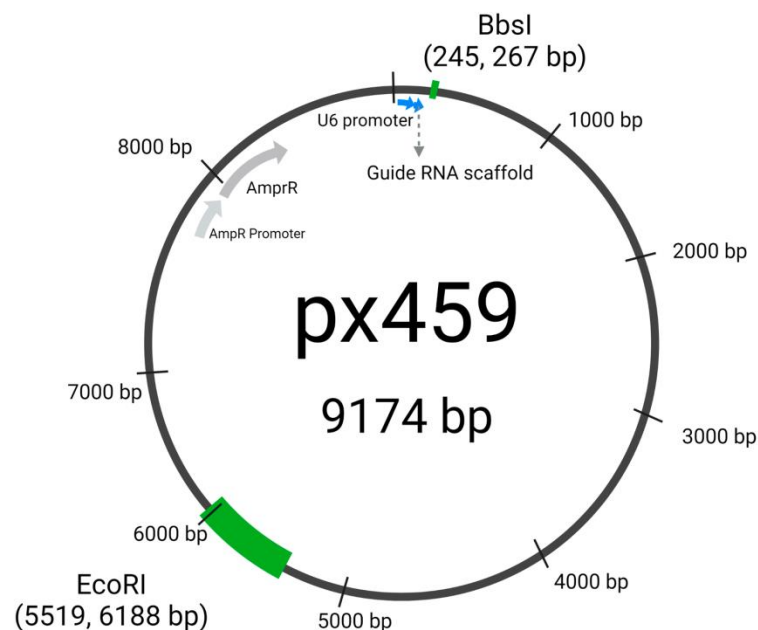
Para incorporação das sequências de DNA que codificam as guias no plasmídeo, primeiramente foi realizado um corte com a enzima de restrição BbsI e, após o corte, utilizamos fosfatase alcalina intestinal de vitela (CIP) para remover o grupo 5'-fosfato de DNA e prevenir a auto ligação das extremidades, seguido pela posterior purificação do DNA. 5  $\mu$ l da amostra foi corrida em gel de agarose 1%. Após a purificação com o kit “Wizard® SV Gel and PCR Clean- Up System”, procedemos para a etapa de ligação do plasmídeo linearizado

com o DNA das guias. Os DNAs codificando as guias foram fosforilados na extremidade 5' pela T4 PNK. Na etapa de transformação, aplicamos a metade da reação de ligação para transformação - conforme padronizado no laboratório de Biologia Tumoral e Biomarcadores. O material foi ressuspendido em meio LB sem antibiótico e aplicados em placas LB-agar com antibiótico correspondente e deixadas à 37°C *overnight*. Como não houve formação de colônias, repetimos o processo de corte com a enzima de restrição e remoção do grupo 5'-fosfato do DNA. Após reservar 5 µl da reação, seguimos com a purificação, 5 µl da eluição foi retirada e corrida em gel de agarose 1% junto com alíquota retirada antes da purificação, a fim de detectar uma possível falha no processo. Notamos que após a purificação houve uma perda significativa do material após a etapa de eluição (**Figura 9**). Isso explicaria a dificuldade na transformação das bactérias devido à baixa concentração de DNA plasmidial, fazendo com que não houvesse crescimento de colônias bacterianas. Como solução, utilizamos todo o volume da etapa de ligação para a transformação, houve crescimento discreto de colônias.



**Figura 9** - Corte e purificação dos plasmídeos pSpCas9(BB)-2A-Puro V2.0 (PX459 V2.0) e pSpCas9n(BB)-2A-Puro V2.0 (PX462 V2.0). Os plasmídeos foram cortados com a enzima de restrição BbsI e purificados para realização da ligação com as sgRNAs desenhadas para o gene *MUTYH*.

Após o processo de transformação de bactérias competentes, foi realizado o cultivo em meio LB com ampicilina como antibiótico de seleção. Realizamos uma miniprep com o kit Wizard® Plus SV Minipreps DNA Purification system (cat A1330, Promega) e selecionamos duas colônias de cada condição para extrair o DNA plasmidial. A verificação dessa reação foi feita por corte com as enzimas de restrição BbsI e EcoRI, que geraria um fragmento de 22 pares de base e um de 669, respectivamente, no caso dos plasmídeos px459 e px462 sem as guias (**Figura 10**). Quando a sequência guia é incorporada no Guide RNA scaffold, o plasmídeo perde os sítios para enzima BbsI, gerando apenas um DNA linearizado.

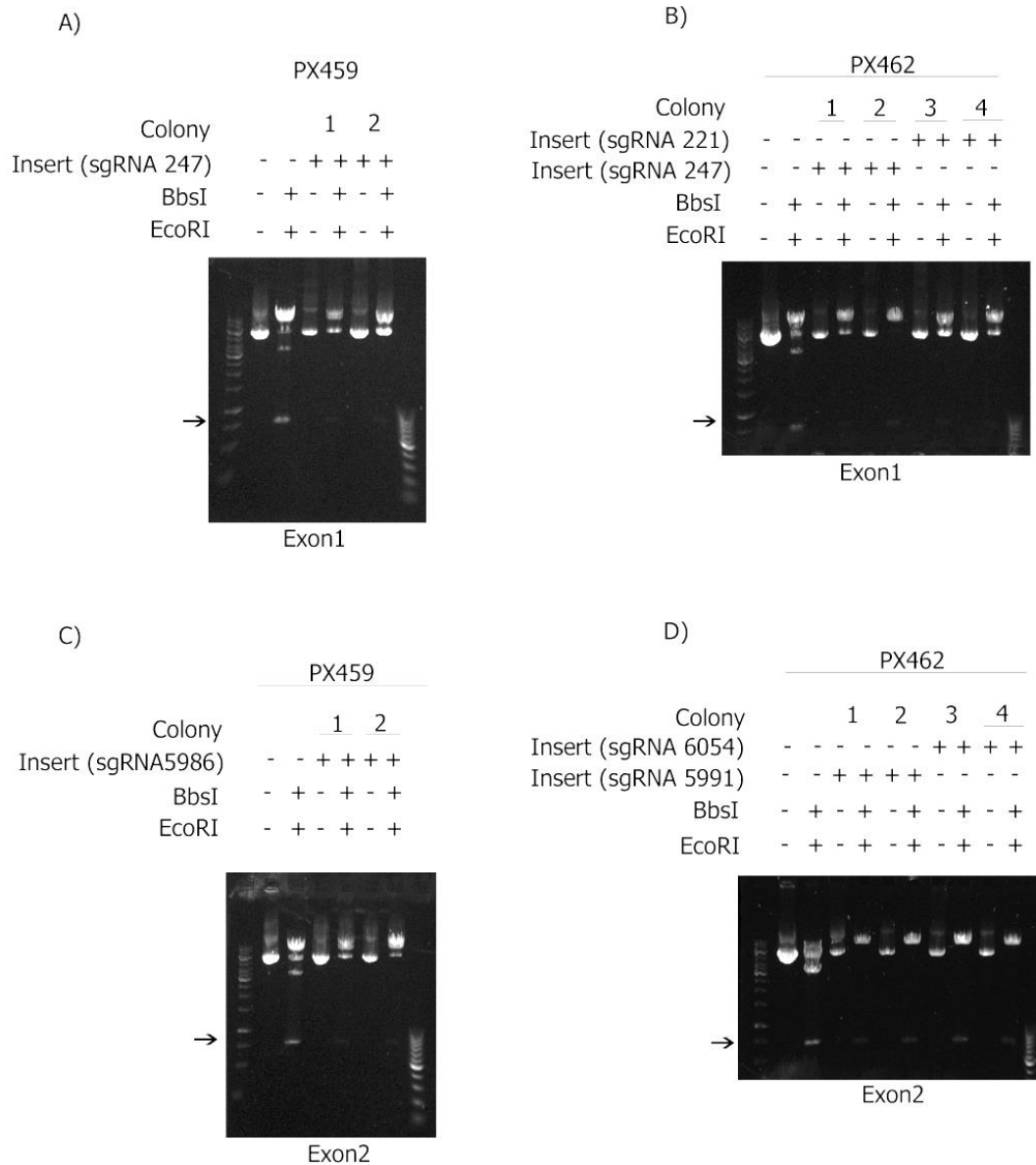


Fonte: Adaptado de <https://www.addgene.org/190542/>

**Figura 10** - Esquema do plasmídeo PX459 mostrando os sítios de restrição das enzimas BbsI e EcoRI. Após o corte no sítio da enzima BbsI e a desfosforilação o inserto (sgRNA) fosforilado é inserido e o mesmo não repõe este sítio da enzima BbsI. O plasmídeo PX462 é idêntico, porém, codifica a Cas9n.

Como resultado, podemos observar que os plasmídeos digeridos perderam o sítio do corte da enzima BbsI mostrando que a inserção das guias foi eficiente (**Figura 11 A-D**). Caso as guias não tivessem sido inseridas, o sítio de BbsI estaria intacto e teria gerado três fragmentos do plasmídeo. Sendo assim na ausência do inserto, teria bandas de aproximadamente 3000, 5500, 669 e 22 pares de bases (o fragmento de 22bp não é observado no gel); E na presença

do inserto, bandas de 8900 e 669. Após a confirmação, uma das colônias foi expandida e submetida a uma Maxi-Prep.



**Figura 11** - Análise de restrição para as colônias com os vetores para a técnica de CRISPR/Cas9.

O produto da reação foi usado para transformar bactérias. Os plasmídeos de duas colônias foram utilizados para checar se a reação de ligação funcionou. Os plasmídeos foram cortados com as enzimas BbsI e EcoRI. Na ausência do inserto, teria bandas de aproximadamente 3000, 5500, 669 e 22 pares de bases (o fragmento de 22bp não é observado no gel). Quando o inserto codificando as sgRNAs foi corretamente ligado, o sítio BbsI é perdido, resultando em dois fragmentos 8900pb e 669pb clivado somente pela enzima EcoRI. (A) Inseto codificando guias para o éxon 1 no plasmídeo PX459 que codifica a Cas9; (B) Inseto codificando guias para o éxon 1 no plasmídeo PX462 que codifica a Cas9n; (C) Inseto codificando guias para o éxon 2 no plasmídeo PX459 que

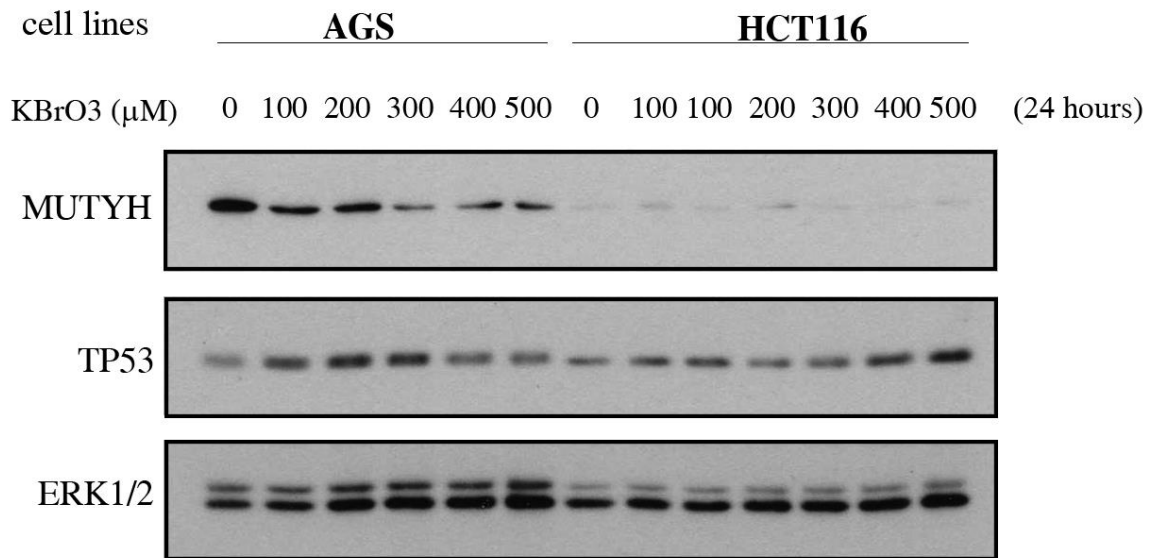
codifica a Cas9; (D) Inseto codificando guias para o éxon 2 no plasmídeo PX462 que codifica a Cas9n. As amostras foram corridas em gel de agarose 0,7%.

## 4.2 AVALIAÇÃO DA EXPRESSÃO DE *MUTYH* EM LINHAGENS CELULARES SOB ESTRESSE OXIDATIVO

Uma parte essencial do presente projeto é geração de modelos celulares para estudar os efeitos das alterações de *MUTYH*. Outro ponto importantes é entender como alterações de *MUTYH*, especialmente as monoalélicas, estão relacionadas com o processo tumoral em tumores diferentes do colorretal. Para isso, antes da edição verificamos a expressão endógena de *MUTYH* em diferentes linhagens celulares derivadas de tipos tumorais em que alterações de *MUTYH* foram verificadas. Assim, inicialmente avaliamos duas linhagens celulares AGS (adenocarcinoma gástrico) e como controle de expressão de *MUTYH* endógeno utilizamos a linhagem HCT116 (carcinoma colorretal) utilizada na literatura para experimentos e expressão de *MUTYH* (OKA et al. 2014). Como a função de *MUTYH* está relacionada com a correção de 8-oxoG, incluímos também o tratamento com o agente oxidante  $\text{KBrO}_3$  a modo de induzir a geração de 8-oxoG. Através de western blot observamos que a expressão de *MUTYH* foi consideravelmente maior em células AGS, derivadas de tumor gástrico, do que em células HCT116, derivadas de carcinoma colorretal. Além disso, nas células AGS verificamos uma diminuição da expressão de *MUTYH* após 24 horas de tratamento com concentrações crescentes de  $\text{KBrO}_3$ . A relação de *MUTYH* com o supressor tumoral TP53 também foi avaliada, já que existem evidências de que *MUTYH* seria um alvo da resposta de TP53 (CHATTERJEE et al. 2006b; OKA et al. 2008b; OKA; NAKABEPPU 2011). É importante mencionar que as duas células escolhidas possuem TP53 selvagem.

Observamos que o tratamento com  $\text{KBrO}_3$  induziu um aumento da expressão de TP53 nas células AGS, no entanto, os níveis induzidos de TP53 foram diferentes dependendo da concentração de  $\text{KBrO}_3$ . Nas células HCT166, o aumento de TP53 foi menor e não foi observado para todas as concentrações avaliadas (**Figura 12**).

Devido à proteína *MUTYH* na linhagem AGS ser facilmente detectável em comparação com as células HCT116. Seguimos somente com a linhagem de interesse, AGS (gástrica). De maneira inesperada, a expressão de *MUTYH* em células HCT166 é baixa, o que é diferente com mostrado na literatura (OKA et al. 2014).

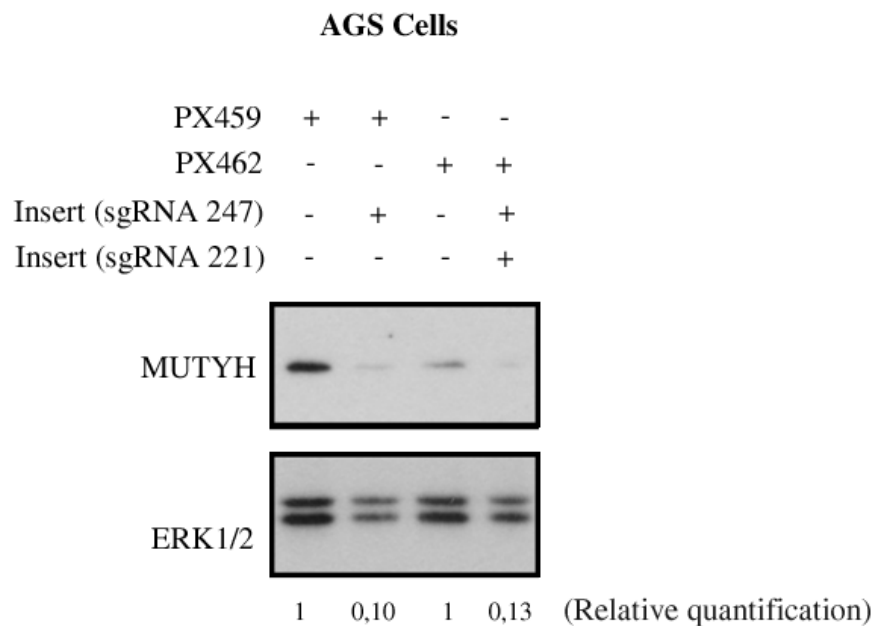


**Figura 12** - Análise dos níveis de expressão de *MUTYH*. Células AGS e HCT116 foram tratadas por 24 h com as concentrações indicadas de KBrO<sub>3</sub>. Extratos celulares foram usados para detectar por western blot as proteínas *MUTYH*, TP53 e ERK em gel de 10%. O anticorpo da Santa Cruz (sc-374571) foi utilizado para *MUTYH*. ERK1/2 foi usado como loading control.

### 4.3 GERAÇÃO DE CÉLULAS NOCAUTE MONOALÉLICAS E BIALÉLICAS PARA *MUTYH*

Após a transfecção e seleção das células AGS que incorporaram o plasmídeo, o pool de células foi analisado por western blot (**Figura 13**). Observamos que houve uma diminuição da expressão de *MUTYH* nas células expressando os sgrNAs, o qual é consistente com a perda de expressão esperada.





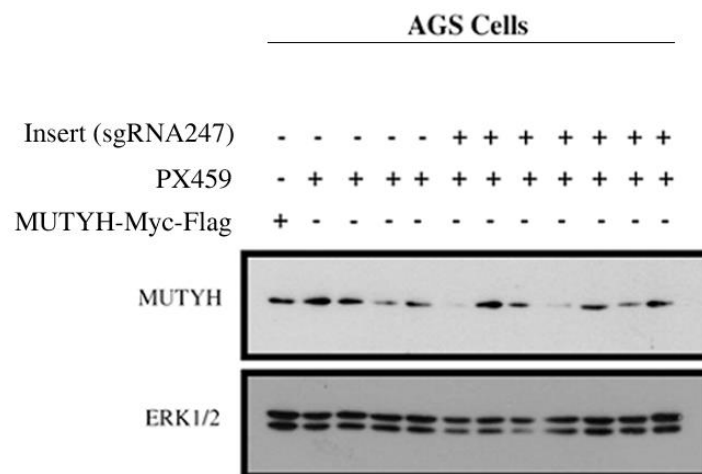
**Figura 13** - Análise do pool de células AGS modificadas pelo sistema CRISPR/Cas9. Células AGS foram transfectadas para expressar Cas9 selvagem (PX459) ou Cas9n (PX462) junto com sgRNAs direcionados para a região do DNA que codifica o primeiro éxon da isoforma alpha5 de *MUTYH*. Após a seleção das células contendo o plasmídeo que tem um cassette de resistência a puromicina, extratos celulares foram usados para western blot para detectar *MUTYH*. ERK1/2 foi usado como loading control. Embaixo do WB está a quantificação relativa respeito às células que não expressam as guias de RNA.

Sabendo que existem grandes chances de que no pool de células AGS houvesse células nocaute para *MUTYH*, procedemos para isolar clones individuais e analisar a expressão de *MUTYH* por western blot (**Figura 14**). No caso das células transfectadas para expressar a Cas9 selvagem, observamos uma expressão de *MUTYH* muito variável em quatro clones obtidos de células que não expressam o sgRNA, porém alguns dos clones expressando o sgRNA apresentaram níveis muito baixos de *MUTYH* condizentes com o esperado, confirmando que realmente o gene *MUTYH* foi editado. No caso dos clones obtidos com a Cas9n e dois sgRNAs, observamos um efeito similar, além de que vários clones de células que expressam os sgRNAs tiveram níveis bem reduzidos (**Figura 15**).

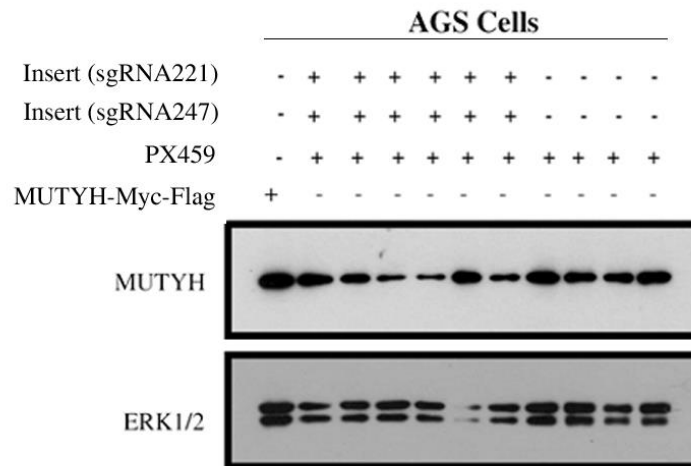
A partir desses resultados surgiram algumas questões. Em primeiro lugar, embora observamos clones com níveis muito baixos de *MUTYH*, em nenhum deles verificamos ausência absoluta. Uma das possibilidades é de que não conseguimos células nocaute dos dois alelos. Por outro lado, a banda fraca remanescente poderia corresponder às isoformas mitocondriais de *MUTYH* que não foram nocauteadas e que o anticorpo também poderia

detectar se expressas. É importante notar que a região modificada com os sgRNAs direcionados contra o primeiro éxon da isoforma  $\alpha 5$  corresponde a uma região *upstream* da região que codifica o primeiro éxon das isoformas mitocondriais. Dessa maneira, sequenciamos o DNA desses clones que tiveram uma grande redução da expressão de *MUTYH* para entender quais foram as alterações no genoma.

Em segundo lugar, nesse set de experimentos incluímos um controle para verificar a especificidade do anticorpo. Basicamente, transfectamos células AGS com um plasmídeo para expressar exogenamente *MUTYH* com tags C-terminais de Myc e Flag (*MUTYH*-Myc-Flag). No entanto, no western blot para *MUTYH* não observamos uma superexpressão de *MUTYH*, nem fomos capazes de detectar uma banda com uma menor migração, embora sutil, causada pelos tags (**Figura 14 e Figura 15**). Dessa maneira ficamos preocupados pela especificidade do anticorpo e/ou plasmídeo para expressão de *MUTYH*-Myc-Flag e realizamos alguns experimentos para avaliá-la.



**Figura 14** - Análise da expressão de *MUTYH* em clones de células AGS modificadas por CRISPR/Cas9. Clones individuais de células AGS expressando Cas9. A expressão de *MUTYH* foi analisada por western blot. ERK1/2 foi usado como loading control. A primeira canaleta corresponde a células AGS transfectadas para expressar *MUTYH*-Myc-Flag como controle positivo para *MUTYH*.

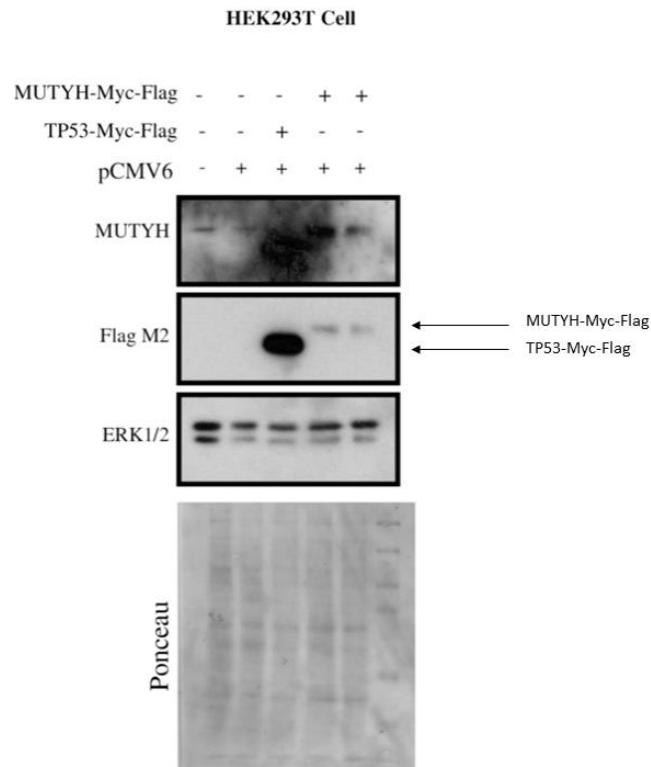


**Figura 15** - Análise da expressão de *MUTYH* em clones de células AGS modificadas por CRISPR/Cas9n. Clones individuais de células AGS expressando Cas9n. A expressão de *MUTYH* foi analisada por western blot. ERK1/2 foi usado como *loading control*. A primeira canaleta corresponde a células AGS transfectadas para expressar *MUTYH-Myc-Flag* como controle positivo para *MUTYH*.

#### 4.4 TRANSFECCÃO COM PLASMÍDEO CODIFICANDO *MUTYH-MYC-FLAG*

Como não foi possível observar de forma evidente a superexpressão da proteína *MUTYH*, além dos resultados conflitantes das expressões proteicas, surgiram algumas dúvidas em relação ao que de fato estávamos observando nas análises proteicas. Se o anticorpo anti-*MUTYH* que estávamos usando era específico ou se a taxa de incorporação do plasmídeo codificando *MUTYH-Myc-Flag* nas células era baixa sendo necessário adaptar o protocolo de transfecção. Para responder essa pergunta, usamos a linhagem HEK293T por sua alta transfectabilidade. Além disso, utilizamos outro plasmídeo correspondendo ao mesmo vetor que o utilizado para expressar *MUTYH-Myc-Flag*, porém codificando TP53-Myc-Flag. Esse último plasmídeo foi usado com muito êxito em diversos experimentos no laboratório de Biologia Tumoral e Biomarcadores. De forma intrigante, o western blot realizado para detectar o *tag de flag* demonstrou que *MUTYH-Myc-Flag* é expresso, mas em quantidades muito inferiores em comparação com TP53-Myc-Flag (**Figura 16**). Quanto ao anticorpo anti-*MUTYH*, foi detectada uma pequena superexpressão de *MUTYH* na linhagem HEK293T, quando comparado com a expressão endógena nas outras células. Em conclusão, a expressão de

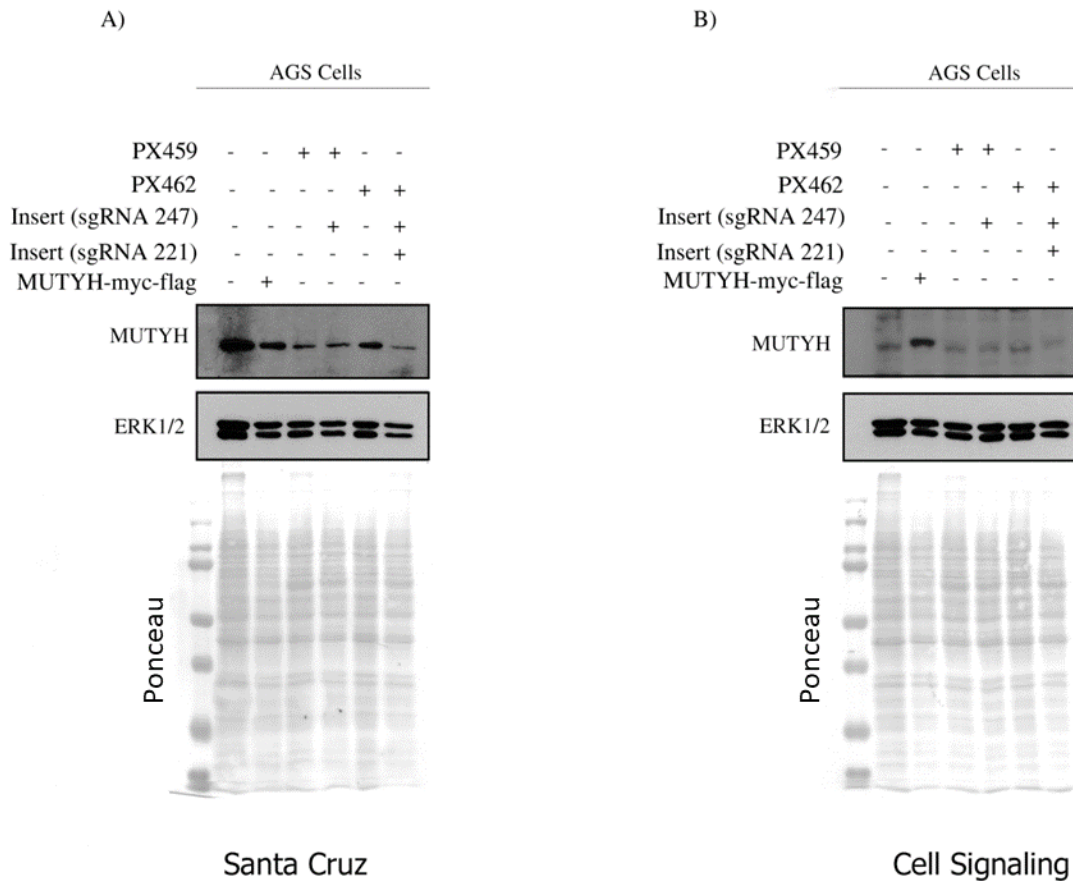
*MUTYH*-Myc-Flag é muito baixa e provavelmente pode ter sido mascarada nas células AGS que expressam grandes quantidades de *MUTYH* basal.



**Figura 16** - Teste de especificidade do anticorpo anti-*MUTYH* e do plasmídeo para expressão exógena de *MUTYH*. Transfecção transiente para expressar *MUTYH*-Myc-Flag ou TP53-Myc-Flag em linhagens HEK293T. Anticorpo Anti-Flag M2 (Sigma #F1804) foi usado para detectar a expressão das proteínas que foram incorporadas. A expressão de *MUTYH*-Myc-Flag foi analisada por western blot com anti-*MUTYH* (Santa Cruz). ERK1/2 foi usado como loading control.

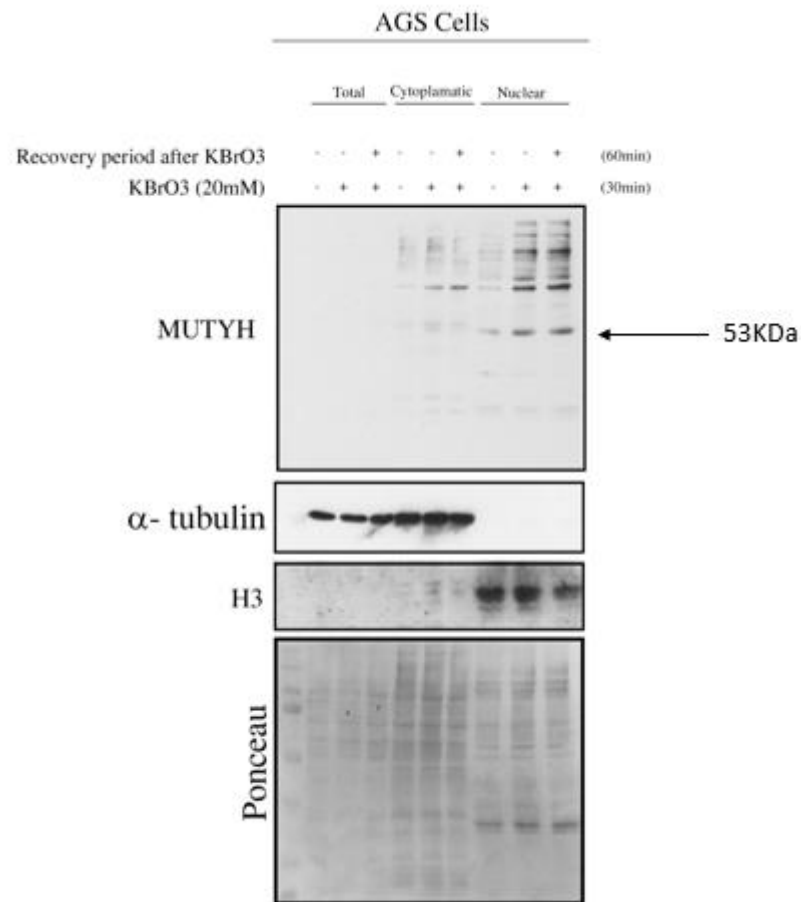
Para validar o anticorpo, adquirimos um novo anti-*MUTYH* (Cell signaling; #6248S). Realizamos um western blot utilizando as mesmas células que sofreram edição para o éxon 1 (**Figura 13**) e aplicamos a mesma sequência em duplicata no mesmo gel (concentração de 10%) para testar os dois anticorpos anti-*MUTYH* em paralelo (**Figura 17**). Após a transferência para a membrana de nitrocelulose, a mesma foi cortada ao meio e incubamos uma com o anticorpo da Santa Cruz (SC- 374571) num título de 1:1000 conforme descrito anteriormente e na outra membrana incubamos com anticorpo da Cell Signaling (#6248S) com a mesma titulação. O resultado foi divergente entre as membranas, porém o anticorpo da Cell Signaling foi capaz de detectar a superexpressão (*MUTYH*-Myc-Flag) de maneira mais clara, enquanto

na membrana que foi incubada com o anticorpo da Santa Cruz observamos padrões diferentes.



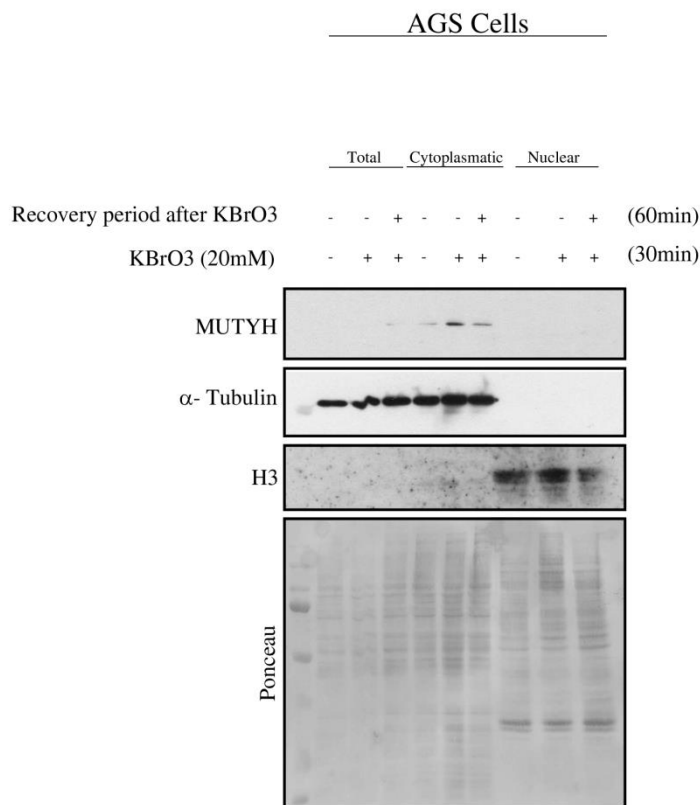
**Figura 17** - Comparação da detecção de *MUTYH* com anticorpos de diferentes marcas. Os extratos das células AGS transfectadas para expressar px459 (Cas9) ou px462 (Cas9n) junto com sgRNAs direcionadas ao Exon1 do gene *MUTYH*, foram submetidos à western blot em duplicata para avaliar a especificidade dos anticorpos da (A) Santa Cruz (SC- 374571) e da (B) Cell Signaling (#62485). A expressão de *MUTYH*-MYC-Flag foi utilizada como controle positivo. ERK1/2 foi usado como *loading control*.

A fim de validar o anticorpo cinética da expressão de *MUTYH* em resposta ao estresse oxidativo, realizamos um pulso com 20 mM de  $\text{KBrO}_3$  por trinta minutos à 37°C. Após o tratamento, as células foram lavadas com PBS e foi colocado o respectivo meio de cultura para recuperação durante 1 h à 37°C antes de prosseguir para o fracionamento subcelular. Logo após os tempos de recuperação, foi iniciado o fracionamento das proteínas citoplasmáticas e nucleares e, por fim, quantificadas e submetidas a uma análise de western blot. Conseguimos observar a expressão da proteína *MUTYH* predominantemente na fração nuclear. Além disso, observamos que *MUTYH* se manteve no núcleo após uma hora de recuperação. Observamos que a expressão basal é muito baixa no controle sem o tratamento e durante o estresse oxidativo houve um aumento considerável na expressão que se manteve mesmo após 1 h de recuperação (**Figura 18**). Utilizamos como controle de carregamento a  $\alpha$ -tubulina para o extrato total e citoplasmático, enquanto a histona H3 foi utilizada como controle de carregamento para a fração nuclear. O não aparecimento de  $\alpha$ -tubulina nas canaletas que contém o extrato nuclear indica que não houve contaminação citoplasmática, ou seja, foi bem purificada. Podemos concluir também que o anticorpo da Cell Signaling (#6248S) detecta primariamente a fração nuclear de *MUTYH* (**Figura 18**).



**Figura 18** - Detecção de *MUTYH* nas frações citoplasmáticas e nucleares usando anticorpo da Cell Signaling. As células AGS foram submetidas a um pulso de 20 mM de KBrO<sub>3</sub> durante 30 minutos e em seguida foram lavadas com PBS e recuperadas por 1 hora. Os extratos proteicos foram fracionados para avaliar a expressão da proteína nas frações citoplasmática e nuclear e posteriormente submetidas a uma análise por western blot utilizando o anticorpo da Cell Signaling (#62485). Alfa-tubulina foi utilizada como *loading control* do extrato total e citoplasmático. A histona H3 foi utilizada como *loading control* para a fração nuclear.

Como última checagem entre os anticorpos, utilizamos os mesmos extratos fracionados em um novo gel porém utilizamos com o anticorpo da Santa Cruz para marcar *MUTYH*. Observamos que o anticorpo detectou uma expressão discreta somente no extrato citoplasmático que se torna um pouco mais evidente somente na canaleta que possui o tratamento de KBrO<sub>3</sub> (**Figura 19**). Com base nesses experimentos concluímos que o anticorpo da Cell Signaling é o mais apropriado para nossa abordagem por ser possível detectar a superexpressão da proteína (*MUTYH*-myc-Flag) e no extrato ser possível a visualização da proteína na fração nuclear além da cinética sob estresse oxidativo.

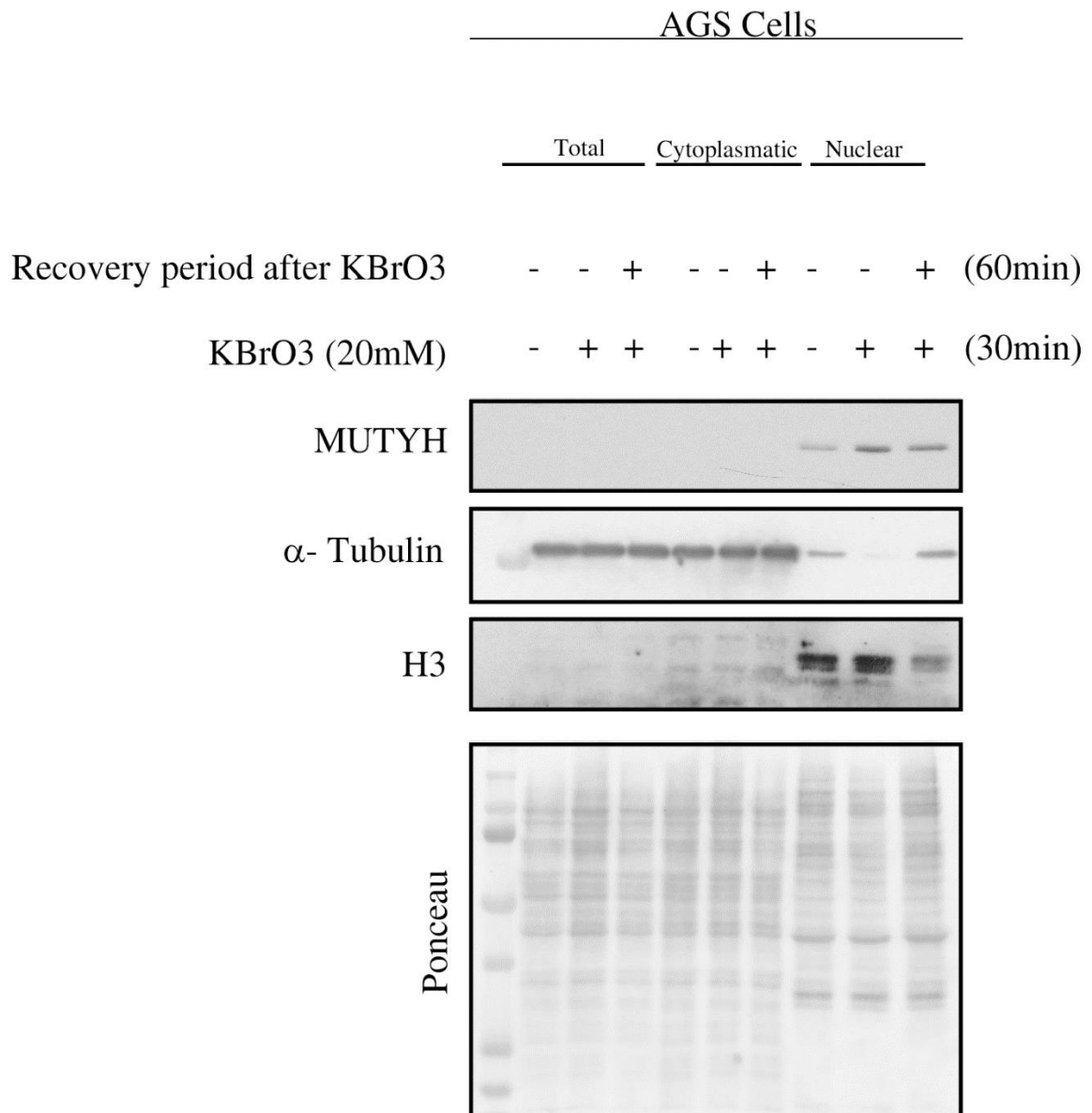


**Figura 19** - *MUTYH* nas frações citoplasmáticas e nucleares usando anticorpo da Santa Cruz.

Os extratos obtidos anteriormente das células AGS foram submetidos novamente a uma análise por western blot, porém utilizamos o anticorpo anti-*MUTYH* da Santa Cruz (SC-374571) para avaliar a especificidade do anticorpo e comparar com a detecção do outro anticorpo.  $\alpha$ -tubulina foi utilizada como *loading control* do extrato total e citoplasmático. A histona H3 foi utilizado como *loading control* para a fração nuclear.



Para confirmar o resultado obtido, repetimos esse experimento com o anticorpo da Cell Signaling e o resultado foi reprodutível e evidente que durante o pulso de 20 mM a expressão de *MUTYH* aumenta na fração nuclear não sendo possível a visualização nas demais frações (**Figura 20**).



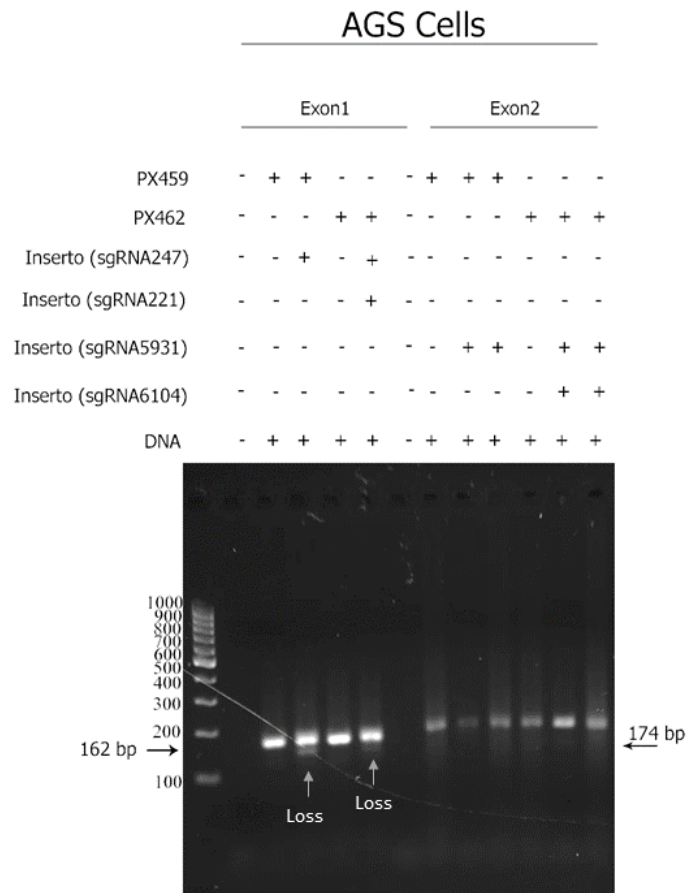
**Figura 20** – Repetição da determinação de *MUTYH* em extratos fracionados com anticorpo da Cell Signaling. Extratos proteicos fracionados das células AGS. Utilizando o anticorpo da Cell Signaling. Alfa-tubulina foi utilizada como *loading control* do extrato total e citoplasmático. A histona H3 foi utilizada como *loading control* para a fração nuclear.

#### 4.5 EXTRAÇÃO DE DNA GENÔMICO PARA REALIZAÇÃO DE REAÇÃO DE PCR PARA AMPLIFICAÇÃO DA REGIÃO EDITADA

Um dos resultados que esperávamos observar com a seleção dos clones era a diminuição da expressão de *MUTYH* ou uma menor expressão que poderia ser um indicativo de que ao menos um dos alelos foi comprometido pela edição. No entanto, para saber realmente se houve modificação, tornou-se necessário sequenciar o pool na região genômica a ser modificada.

A partir do pool das células geradas com as guias para o éxon 1 e éxon 2 foi extraído o DNA genômico de ambos para realização de uma PCR para amplificação da região editada (**Tabela 2**). Devido às dificuldades de confirmação por western blot, as células geradas para o nocaute do éxon 2 de *MUTYH* foram submetidas diretamente para extração de DNA genômico para amplificação e posteriormente sequenciamento.

O produto das reações de PCR gerou amplicons nos tamanhos de 162 bp correspondentes as guias desenhadas para o éxon 1 e 174 bp para as guias referentes ao éxon 2. Os diferentes amplicons foram enviados para sequenciamento de amplicon (**Figura 21**). Podemos observar que as células AGS editadas com CRISPR/Cas9 e Cas9n e as respectivas guias direcionadas para o éxon 1 e éxon 2 apresentaram duas bandas, ou seja, podem existir células que foram editadas. Amostra que possui a Cas9 com a guia sgRNA247 direcionada para o éxon 1 (terceira e quinta canaleta) apresentou duas bandas na altura aproximadamente de 164bp e uma menor indicando que parte da fita de DNA está menor, ou seja, perda de bases, evidenciando que nesse pool a taxa de eficiência pode ter sido maior.



**Figura 21** - PCR para amplificação da região editada pelo sistema CRISPR/Cas9. As canaletas sem DNA indicam o controle negativo. As setas indicam as bandas de peso molecular 162 bp e 174 bp para o éxon 1 e éxon 2, respectivamente. As amostras foram corridas em gel de agarose 2%.

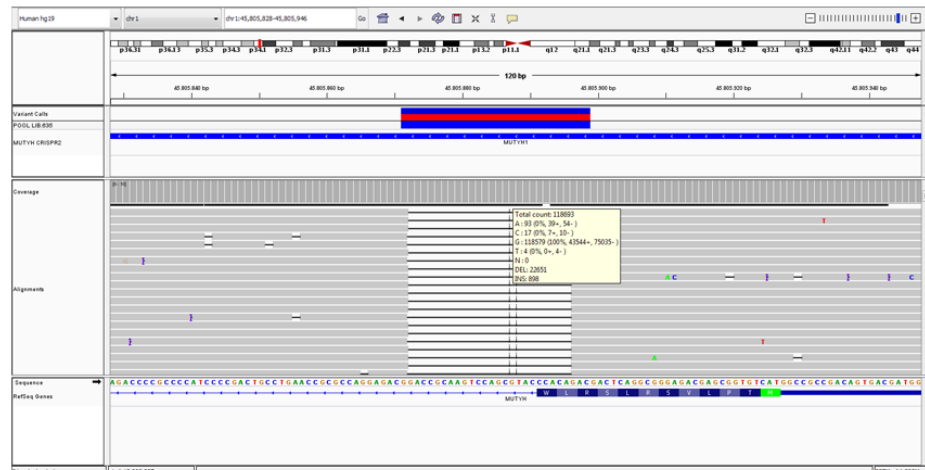
#### 4.6 SEQUENCIAMENTO DE AMPLICON DOS POOLS GERADOS

Os amplicons dos pools correspondentes foram sequenciados na plataforma Ion Torrent (ThermoScientific). Detectamos alelos editados tanto no éxon 1 e no éxon 2, o que indica que o sistema CRISPR/Cas9 funcionou adequadamente nas células AGS. É importante mencionar que os dois sistemas, com Cas9 ou com Cas9n, funcionaram, porém, gerando alterações com diferentes frequências (**Figura 22 e Figura 23**).

# AGS

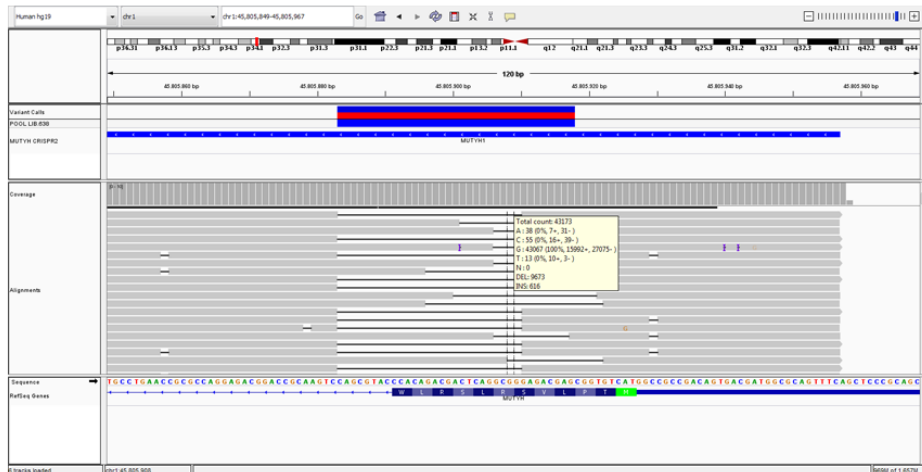
## Exon1

PX459



Deletion 19%

PX462

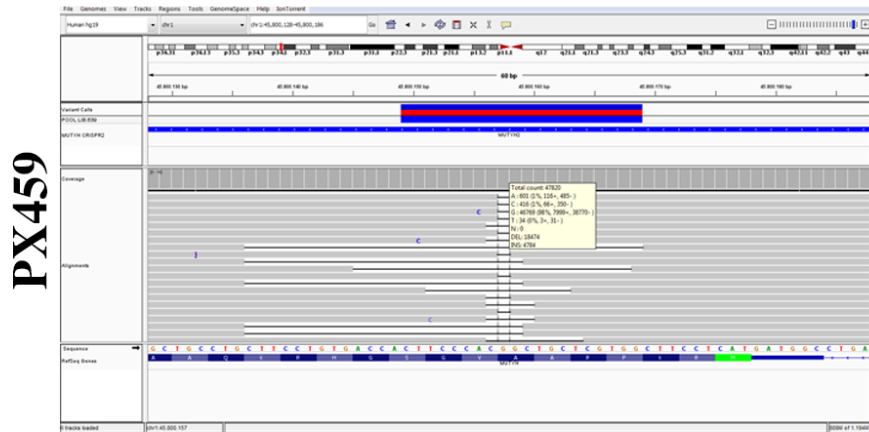


Deletion 22%

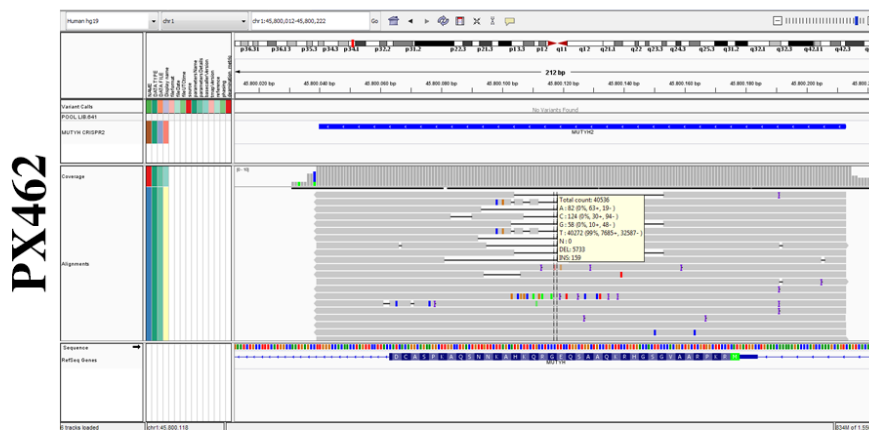
**Figura 22** - Sequenciamento da região amplificada para o éxon 1 editada pelos sistemas CRISPR/Cas9 e CRISPR/Ca9n. Sequenciamento do pool gerado pelo sistema CRISPR/Cas9 (PX459) para o éxon 1, observamos uma taxa de deleção de 19% nesse pool; e na estratégia CRISPR/Ca9n (PX462) cerca de 22% de deleção.

# AGS

## Exon2



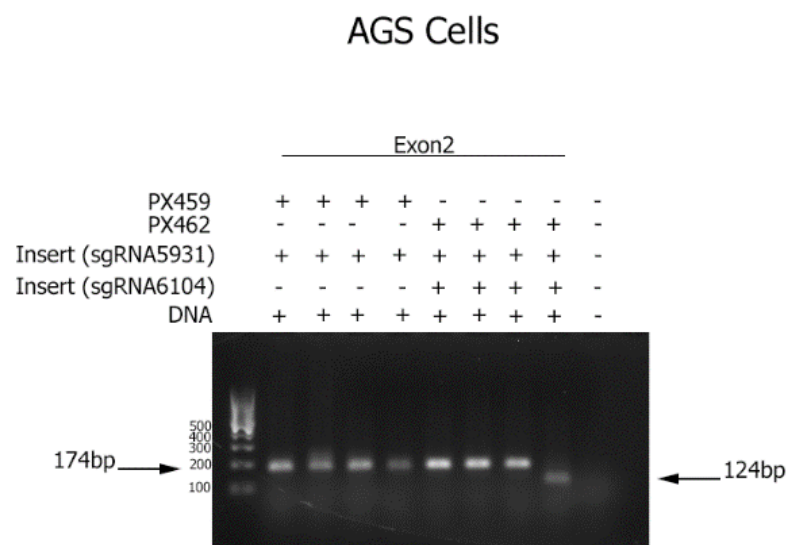
Deletion 8%



Deletion 14%

**Figura 23** - Sequenciamento da região amplificada para o Exon2 editada pelos sistemas CRISPR/Cas9 e CRISPR/Ca9n. Sequenciamento do pool gerado pelo sistema CRISPR/Cas9 (PX459) com frequência de deleções para o exon2 de 8% e pela CRISPR/Ca9n (PX462) com 14% de deleção.

Sabendo que existem células com alelos modificados no pool, procedemos a gerar clones individuais a partir dos pools. Após o isolamento dos clones, o DNA genômico foi amplificado para sequenciar as modificações nos clones individuais. De maneira interessante, um dos clones obtidos com Cas9 direcionada para o exon2, mostrou um amplicon de menor tamanho em relação aos demais amplicons, indicativo de que ocorreu uma deleção considerável nos alelos, e significando a obtenção de células AGS modificadas para *MUTYH* (Figura 24).



**Figura 24** - PCR para amplificação da região editada pelo sistema CRISPR/Cas9 em clones individuais. Um dos clones obtidos pela Cas9n direcionada para o éxon 2 indica perda significativa em seu tamanho. As amostras foram corridas em gel de agarose 2%. Amplificação da região editada pelos sistemas CRISPR/Cas9 (PX459) e CRISPR/Cas9n (PX462) para o éxon 2.

#### 4.7 SEQUENCIAMENTO DOS CLONES GERADOS ATRAVÉS DE CRISPR/ CAS9 E CRISPR/CAS9N

Dividimos a seleção dos clones em quatro grupos conforme mostra a **Tabela 5** e selecionamos um total de 160 clones dos quais 16 foram sequenciados.

**Tabela 5** - Número de clones selecionados e sequenciados de cada grupo. Quantidade de clones por grupos correspondentes às estratégias CRISPR utilizadas e seus respectivos éxons alvos.

Grupo	Estratégia de edição	Éxon	Numero de clones selecionados	Numero de clones sequenciados	
1	CRISPR/Cas9	1	40	4	
2	CRISPR/Cas9n	1	40	4	
3	CRISPR/Cas9	2	40	4	
4	CRISPR/Cas9n	2	40	4	
			160	16	<b>Total</b>

As análises do sequenciamento dos clones revelaram um padrão de três leituras distintas em um mesmo clone conforme exemplificado na **Tabela 6**, esses resultados sugerem uma hiperploídia no cromossomo 1 na linhagem AGS. Porém dois clones se mostraram promissores com 98% de deleção e outro com 51% de taxa de edição mimetizando a condição variante patogênica bialélica e monoalélica.

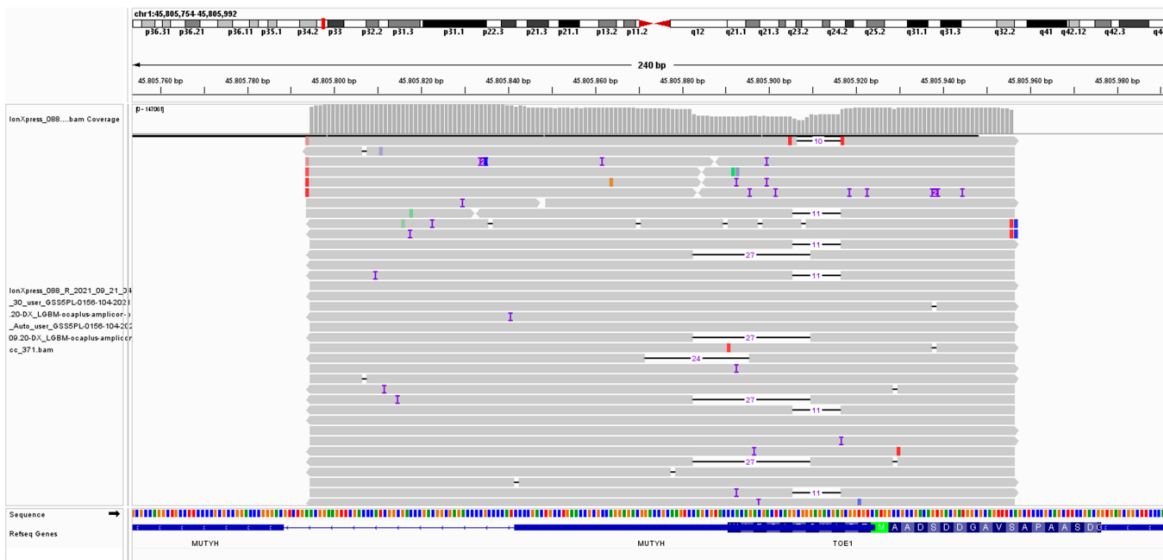
**Tabela 6** - Resultado geral dos clones sequenciados e suas respectivas frequências de edição genica. Clones da linhagem AGS gerados pelo sistema CRISPR/Cas9 sequenciados e suas respectivas frequências globais de edição genica nos exons correspondentes as guias.

Resultado do sequenciamento dos clones da linhagem AGS							
CRISPR	Enzima	Inserto/ Guia	Éxon	Alterações			
				Alelo 1	Alelo 2	Alelo 3	Frequência de edição (%)
CRISPR	Cas9	sgRNA247	1	Selvagem	Selvagem	del 1 base	27%
		sgRNA247	1	inserção 1 base	deleção 1 base	deleção 24 bases	98%
		sgRNA247	1	Selvagem	deleção 1 base	deleção 11 bases	65%
		sgRNA247	1	Selvagem	Selvagem	Selvagem	0%
	Ca9n	sgRNA247 + sgRNA221	1	Selvagem	deleção 10 bases	deleção 15 bases	48%
		sgRNA247 + sgRNA221	1	deleção 26 bases	deleção 11 bases	Selvagem	51%
		sgRNA247 + sgRNA221	1	deleção 1 base	Selvagem	Selvagem	15%
		sgRNA247 + sgRNA221	1	Selvagem	Selvagem	Selvagem	0%
	Cas9	sgRNA5931	2	Vazio	Vazio	Vazio	Vazio
		sgRNA5931	2	deleção 10 bases	Selvagem	Selvagem	6%
		sgRNA5931	2	Selvagem	Selvagem	Selvagem	0%
		sgRNA5931	2	Selvagem	deleção 4 bases	deleção 12 bases	35%
	Ca9n	sgRNA5991 + sgRNA6104	2	Selvagem	Selvagem	Selvagem	0%
		sgRNA5991 + sgRNA6104	2	Selvagem	Selvagem	Selvagem	0%
		sgRNA5991 + sgRNA6104	2	Selvagem	deleção 24 bases	Selvagem	15%
		sgRNA5991 + sgRNA6104	2	deleção 10 bases	Selvagem	Selvagem	27%

Na **Figura 25** está apresentado um dos clones sequenciado utilizando a ferramenta CRISPR/ Cas9n com duas guias direcionadas para o éxon 1. Foi possível observar uma deleção pequena de 11 bases que vai resultar uma troca de uma leucina por uma prolina na posição 4 gerando um códon de parada 34 códons depois (NM\_001128425.1:c.11\_21del; p.(Leu4Profs\*34)). Outra deleção maior de 26 bases também foi observada e a alteração na proteína não pode ser aferida (NM\_001128425.1:c.18\_36+8del (p.sp!)). Essas edições representam 50% das leituras e o restante são selvagens.

## AGS- Cell

### Clone - PX462 Insert (sgRNA247 + sgRNA221)



**Figura 25** - Clone knockout heterozigoto da linhagem AGS para o éxon 1 utilizando a ferramenta CRISPR/Cas9n. Sequenciamento do amplicon obtido de clone que foi editado utilizando a técnica CRISPR/Cas9n com guias direcionadas para o éxon 1 do gene *MUTYH* na linhagem AGS.

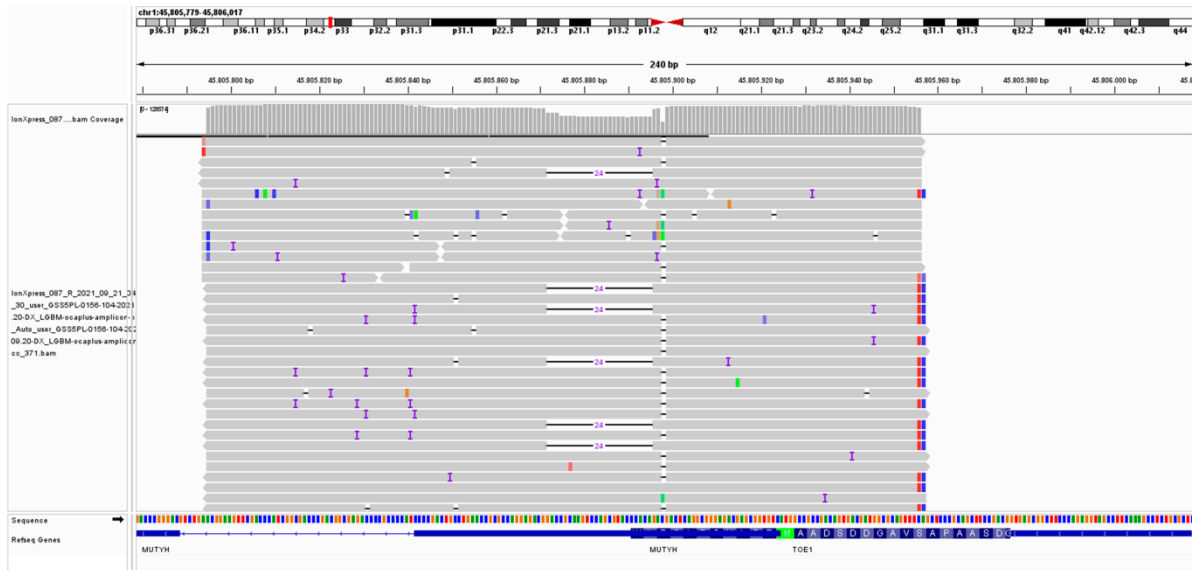
Utilizando a ferramenta CRISPR/Cas9 com uma guia direcionada para o éxon 1 observamos a presença de reads com deleção de 24 pb que estaria resultando em uma mutação *in frame*. Outro grupo de reads evidenciaram a deleção de uma base (**Figura 26**). A deleção de uma base resulta na troca de uma arginina na posição 10 por uma leucina e ocasiona um frameshift resultando em um códon de parada seis aminoácidos depois (NM\_001128425.1:c.29delG; p.(Arg10Leufs\*6)). Em geral pelo frameshift ser muito cedo provavelmente essa proteína não será produzida, uma vez que o RNA mensageiro deve ser



degradado pela via NMD (Nonsense-Mediated Decay). Também foi detectada uma deleção maior de 26 bases que sua alteração na proteína não pode ser aferida (NM\_001128425.1:c.18\_36+8del (p.sp1?)). E essas deleções representam 98% das leituras.

## AGS - Cells

### Clone PX459 Insert (sgRNA247)



**Figura 26** - Clone knockout homozigoto da linhagem AGS para o exon1 utilizando a ferramenta CRISPR/ Cas9. Sequenciamento do amplicon obtido de clone que foi editado utilizando a técnica CRISPR/Cas9 com guias direcionadas para o éxon 1 do gene *MUTYH* na linhagem AGS.

## 5 DISCUSSÃO

As espécies reativas de oxigênio podem ser formadas como subprodutos do metabolismo bioenergético ou da exposição ambiental. 8-oxoG é uma das lesões de DNA mais prevalentes induzidas por ROS e o reparo mediado por *MUTYH* de pares incorretos de A:8-oxoG desempenha um papel crucial na prevenção e na manutenção da integridade genética. As consequências funcionais da deficiência de *MUTYH* não são totalmente compreendidas até o momento, mas sabe-se que a deficiência de *MUTYH* previne a formação de quebras de fita simples (SSBs) e a morte celular sob o estresse oxidativo (FOTI et al. 2012).

Sabe-se que variantes patogênicas bialélicas (homozigótica) no gene *MUTYH*, ocorrem em aproximadamente 0,04% da população caucasiana e estão associadas com um aumento de 18 a 100 vezes no risco de desenvolver adenomas e CRC em comparação com a população em geral (CLEARY et al. 2009). Um estudo recente relatou risco aumentado dos portadores de variante patogênica bialélica em *MUTYH* de desenvolver câncer de bexiga e câncer de ovário, mas nenhuma associação no desenvolvimento de tumor cerebral, pancreático, rim ou próstata (WIN et al. 2016).

Variantes patogênicas monoalélicas (heterozigóticas) de *MUTYH* estão associadas a um aumento moderado do risco de desenvolver CRC além de um aumento de 11% para o desenvolvimento de câncer de mama e 5% aumentado para câncer gástrico (WIN et al. 2014). Diversos estudos científicos vêm demonstrando que a redução da atividade de *MUTYH* em ambos os alelos ou apenas em um dos alelos, pode aumentar o risco de diferentes tipos de cânceres.

Câncer gástrico é um problema de saúde pública, como bem descrito na literatura. As células gástricas estão expostas a todo o momento ao estresse oxidativo, sendo agravado pela dieta, tabagismo e alcoolismo (SUNG et al. 2021). Em uma análise multivariada feita por Shinmura et al. revelou que a expressão reduzida de *MUTYH* é um preditor de um mau prognóstico de sobrevivência entre os pacientes com câncer gástrico (SHINMURA et al. 2011).

Para realizar estudos funcionais em relação à deficiência de uma proteína podemos utilizar ferramentas de edição gênica, sendo a melhor opção para desvendar esses mecanismos. A programação da nuclease Cas9 associada a CRISPR para modificar locais

específicos do genoma, permite uma nova maneira de interrogar a função do gene ou variante patogênica. Através da geração de modelos celulares com variantes genéticas como realizado por Lentsch et al. que utilizou linhagens celulares de câncer pancreático para nocautear a mutação de *KRAS* c.35G>A (p.G12D) e observar que, em condições regulares, o crescimento dos clones nocauteados se assemelhava à de células que possuíam *KRAS* selvagem, indicando que é possível eliminar a mutação *KRAS* G12D por CRISPR/Cas9 (LENTSCH et al. 2019). A ferramenta de edição genica CRISPR-Cas9 possui um enorme potencial para ampliar a compreensão dos processos biológicos que levam à tumorigênese.

Neste estudo, realizamos a construção dos plasmídeos com as enzimas Cas9 do tipo selvagem e a Cas9 do tipo nickase, com guias direcionadas para o exon 1 ou exon 2 do gene *MUTYH*, a fim de gerar células nocautes. As construções foram checadas pelo teste de restrição e em seguida após a confirmação iniciamos a geração do modelo celular. Obtivemos uma media de 15% de eficiência de edição no pool de células. Para uma edição genica bem-sucedida, a escolha do método para introduzir o complexo Cas9/sgRNA nas células é crítico. Uma alternativa de otimizar a transfecção e aumentar a eficiência, seria trocar o método de lipofectamina 2000 por eletroporação aprimorada chamada *Nucleofection*<sup>®</sup> da Lonza, que utiliza uma combinação de parâmetros específicos elétricos que permite e transfere a molécula diretamente no núcleo das células, a eficiência de transfecção é significativamente maior quando utilizada pela metodologia lipofectamina 2000 (YANG et al. 2010).

Após isolar os clones, os genomas dos clones individuais foram submetidos ao sequenciamento de amplicons. Observamos que grande parte dos clones apresentaram *reads* de três tipos de leitura em cada clone, sugerindo que as células AGS podem ser hiperdiploide no cromossomo 1. Não foi possível localizar na literatura se de fato no cromossomo 1 há um terceiro alelo nessa linhagem, de qualquer forma, segundo a ATCC o número cromossômico modal é de 49, significando que possui copias a mais de cromossomos, ocorrendo em 60% das células. Segundo análise de cariótipo mencionada na ATCC relatam que geralmente há três cópias para cromossomo 20 e uma única cópia para X,8 e 18 e ocasionalmente há três cópias para o cromossomo 14 e não constam maiores informações que expliquem esse achado referente ao cromossomo 1. Todavia há estudos utilizando a mesma linhagem AGS que demonstraram que uma menor expressão de *MUTYH* resulta em uma menor atividade supressora para proliferação dos clones AGS, e não mencionaram nenhuma anormalidade referente ao número de copias do gene *MUTYH*, ou na função da proteína (SHINMURA et al.

2011). Uma possível solução para verificar a ploidia dessa linhagem e validar o resultado dos sequenciamentos, seria realizar um *OncoScan* Microarray Assay (OMA); esse ensaio permite detectar desequilíbrios cromossômicos e variantes de número de cópias.

No clone gerado utilizando a enzima Cas9 do tipo selvagem obtivemos duas leituras com deleções próximas de tamanhos diferentes e em posições diferentes no exon 1 e uma leitura selvagem, enquanto o outro clone gerado com a enzima Cas9n possui duas deleções de tamanhos diferentes no exon 1. Do ponto de vista de efeito basicamente podemos esperar um alelo funcional, ou seja, podemos esperar que essa célula se comporte como um mutante monoalélico. No clone gerado com a enzima Cas9 o sequenciamento mostra uma taxa de deleção de 98% no exon1, nesse caso podemos esperar ausência da atividade funcional de *MUTYH* mimetizando a condição da variante patogênica em homozigose. Os dois modelos podem ser comparados com a célula selvagem futuramente em ensaios funcionais para verificar se realmente os efeitos das deleções são como esperados. Poderá ser utilizado ensaios de dano e reparo de DNA, como por exemplo pelo *Comet Assay*, um ensaio de eletroforese em gel de célula única, para medir quebras de fita de DNA como uma forma de avaliar indiretamente o reparo de DNA sob estresse oxidativo (FANG et al. 2015).

Realizamos ensaios de indução da proteína *MUTYH* para entender a cinética frente ao estresse oxidativo utilizando linhagens celulares derivadas de tumor gástrico e colorretal (AGS e HCT116, respectivamente). A linhagem HCT116 é muito utilizada na literatura para detecção da proteína *MUTYH* em ensaios funcionais, devido à *MUTYH* estar associada à Polipose Adenomatosa Familiar (NIELSEN et al. 1993). Deste modo, utilizamos essas células a fim de comparar os efeitos na expressão e indução da proteína entre as linhagens. Os resultados gerados pela detecção da proteína foram inconclusivos, devido as divergências nas detecções da proteína entre as linhagens celulares, mesmo repetindo um mesmo extrato celular total a marcação do anticorpo variava. Portanto uma nova abordagem será necessária para compreender o mecanismo de indução e detecção da proteína *MUTYH*.

Sabe-se que OGG1 interage com *MUTYH* para excisão da 8-oxoG enquanto *MUTYH* remove adenina pareada erroneamente (MAYNARD et al. 2008). Amouroux et al. 2010 observaram que após um pulso de dano oxidativo de 20 mM de  $\text{KBrO}_3$  por 30 minutos, houve um pico no recrutamento de OGG1 para o núcleo após 3 horas em recuperação e uma redução nas horas seguintes voltando ao estado basal após 8 horas. Essa cinética foi possível observar por western blot utilizando a fração nuclear. Repetimos esse ensaio para detectar *MUTYH* e

aproveitamos para comparar as especificidades dos anticorpos. Observamos que durante o estímulo com o agente oxidante, *MUTYH* é recrutado para o núcleo e se mantém após 1 hora em recuperação do dano. Podemos inferir que *MUTYH* é recrutado primeiramente em relação a OGG1 na via BER, durante o estresse oxidativo.

Nossos experimentos de cinética usando a linhagem AGS com *MUTYH* selvagem indicam que os níveis basais de *MUTYH* são baixos e é recrutado para o núcleo durante o estresse oxidativo e o recrutamento se mantém mesmo após o estímulo. Com esses dados poderá ser realizado ensaios comparando os níveis de recrutamento de *MUTYH* nas células que possuem edição gênica. Visto que ambas as guias são direcionadas para o exon1, que transcreve as três principais isoformas nucleares de *MUTYH*.

Poderá ser realizada microscopia confocal com marcações para observar a formação de 8-oxG utilizando anticorpo para detecção de bases de guanina oxidadas (após a fixação deverá tratar com RNase-A para permitir a detecção de DNA modificado). Com a padronização do anticorpo anti *MUTYH*, poderá ser avaliado o direcionamento de *MUTYH* combinado com o tratamento de pulso de estresse oxidativo. Também poderemos realizar uma superexpressão utilizando o plasmídeo *MUTYH*-Myc-Flag, acompanhado de imunofluorescência com o anticorpo anti-Flag como controle indireto na localização da proteína.

A geração desse modelo celular poderá contribuir com diversas investigações da deficiência de *MUTYH*, como por exemplo avaliar assinaturas mutacionais típicas de lesões 8-oxoG que resultam em transversão G:C para T:A após os eventos replicativos do DNA. Através de ensaio de sobrevivência celular à estresse oxidativo, será possível verificar se as alterações genéticas induzidas no gene *MUTYH* das células AGS levam de fato a perda da função de *MUTYH*. Além disso, as construções do sistema CRISPR/Cas9 e CRISPR/Cas9n ambos com guias direcionadas para o exon1 ou exon2, poderão ser utilizados também para investigações da deficiência de *MUTYH* em outras linhagens, como por exemplo de colorretal, uma vez que *MUTYH* também está associado à síndrome FAP.

## 6 CONCLUSÃO

Com a confirmação do sequenciamento, os resultados apontam que as metodologias da técnica de CRISPR/Cas9 funcionaram. Os clones foram isolados e sequenciados, e obtivemos dois clones em potencial para mimetizar a inativação de *MUTYH* nas condições monoalélica e bialélica.

## 7 REFERÊNCIAS

- Al-Tassan N, Chmiel NH, Maynard J, et al. Inherited variants of MYH associated with somatic G:C→T:A mutations in colorectal tumors. **Nat. Genet.** 2002; 30:227–232.
- Arosio P, Ingrassia R, Cavadini P. Ferritins: A family of molecules for iron storage, antioxidation and more. **Biochim. Biophys. Acta - Gen. Subj.** 2009; 1790:589–599.
- Au KG, Cabrera M, Miller JH, Modrich P. Escherichia coli mutY gene product is required for specific A-G----C.G mismatch correction. **Proc. Natl. Acad. Sci.** 1988; 85:9163–9166.
- Banda DM, Nuñez NN, Burnside MA, Bradshaw KM, David SS. Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine. **Free Radic. Biol. Med.** 2017; 107:202–215.
- Benhar M. Roles of mammalian glutathione peroxidase and thioredoxin reductase enzymes in the cellular response to nitrosative stress. **Free Radic. Biol. Med.** 2018; 127:160–164.
- Birben E, Sahiner UM, Sackesen C, Erzurum S, Kalayci O. Oxidative Stress and Antioxidant Defense. **World Allergy Organ. J.** 2012; 5:9–19.
- Boldinova, Khairullin, Makarova, Zharkov. Isoforms of Base Excision Repair Enzymes Produced by Alternative Splicing. **Int. J. Mol. Sci.** 2019; 20:3279.
- Breitzig M, Bhimineni C, Lockey R, Kolliputi N. 4-Hydroxy-2-nonenal: a critical target in oxidative stress? **Am. J. Physiol. Physiol.** 2016; 311:C537–C543.
- Brigelius-Flohé R, Maiorino M. Glutathione peroxidases. **Biochim. Biophys. Acta - Gen. Subj.** 2013; 1830:3289–3303.
- Brown DI, Griendling KK. Nox proteins in signal transduction. **Free Radic. Biol. Med.** 2009; 47:1239–1253.
- Buettner GR. The Spin Trapping of Superoxide and Hydroxyl Free Radicals with DMPO (5,5-Dimethylpyrrolidine-N-oxide): More About Iron. **Free Radic. Res. Commun.** 1993; 19:s79–s87.
- Carvajal LA, Manfredi JJ. Another fork in the road—life or death decisions by the tumour suppressor p53. **EMBO Rep.** 2013; 14:414–421.
- Chatterjee A, Mambo E, Osada M, Upadhyay S, Sidransky D. The effect of p53 -RNAi and p53 knockout on human 8-oxoguanine DNA glycosylase (hOgg1) activity. **FASEB J.** 2006a; 20:112–114.
- Chatterjee A, Mambo E, Osada M, Upadhyay S, Sidransky D. The effect of p53-RNAi and p53

- knockout on human 8-oxoguanine DNA glycosylase (hOgg1) activity. **FASEB J.** 2006b; 20:112–114.
- Chatterjee N, Walker GC. Mechanisms of DNA damage, repair, and mutagenesis. **Environ. Mol. Mutagen.** 2017; 58:235–263.
- Cleary SP, Cotterchio M, Jenkins MA, et al. Germline MutY human homologue mutations and colorectal cancer: a multisite case-control study. **Gastroenterology** 2009; 136:1251–60.
- Curia MC, Catalano T, Aceto GM. MUTYH: Not just polyposis. **World J. Clin. Oncol.** 2020; 11:428–449.
- David SS, O’Shea VL, Kundu S. Base-excision repair of oxidative DNA damage. **Nature** 2007; 447:941–50.
- De Belle I, Wu J-X, Sperandio S, Mercola D, Adamson ED. In Vivo Cloning and Characterization of a New Growth Suppressor Protein TOE1 as a Direct Target Gene of Egr1. **J. Biol. Chem.** 2003; 278:14306–14312.
- De Boer J, Hoeijmakers JHJ. Nucleotide excision repair and human syndromes. **Carcinogenesis** 2000; 21:453–460.
- De Brito G, Lupinacci FC, Beraldo FH, et al. Loss of prion protein is associated with the development of insulin resistance and obesity. **Biochem. J.** 2017; 474:2981–2991.
- Denver DR. An Evolutionary Analysis of the Helix-Hairpin-Helix Superfamily of DNA Repair Glycosylases. **Mol. Biol. Evol.** 2003; 20:1603–1611.
- Derks S, De Klerk LK, Xu X, et al. Characterizing diversity in the tumor-immune microenvironment of distinct subclasses of gastroesophageal adenocarcinomas. **Ann. Oncol.** 2020; 31:1011–1020.
- Dodson M, Castro-Portuguez R, Zhang DD. NRF2 plays a critical role in mitigating lipid peroxidation and ferroptosis. **Redox Biol.** 2019; 23:101107.
- Elko EA, Cunniff B, Seward DJ, et al. Peroxiredoxins and Beyond; Redox Systems Regulating Lung Physiology and Disease. **Antioxid. Redox Signal.** 2019; 31:1070–1091.
- Fang L, Neutzner A, Turtschi S, Flammer J, Mozaffarieh M. Comet assay as an indirect measure of systemic oxidative stress. **J. Vis. Exp.** 2015; e52763.
- Foti JJ, Devadoss B, Winkler JA, Collins JJ, Walker GC. Oxidation of the guanine nucleotide pool underlies cell death by bactericidal antibiotics. **Science** 2012; 336:315–9.
- Friedberg EC. DNA damage and repair. **Nature** 2003; 421:436–440.
- Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in



- genomic DNA without DNA cleavage. **Nature** 2017; 551:464–471.
- Giacomazzi J, Graudenz MS, Osorio CABT, et al. Prevalence of the TP53 p.R337H Mutation in Breast Cancer Patients in Brazil. **PLoS One** 2014; 9:e99893.
- Gill JG, Piskounova E, Morrison SJ. Cancer, Oxidative Stress, and Metastasis. **Cold Spring Harb. Symp. Quant. Biol.** 2016; 81:163–175.
- Gozzelino R, Jeney V, Soares MP. Mechanisms of Cell Protection by Heme Oxygenase-1. **Annu. Rev. Pharmacol. Toxicol.** 2010; 50:323–354.
- Guina T, Biasi F, Calfapietra S, Nano M, Poli G. Inflammatory and redox reactions in colorectal carcinogenesis. **Ann. N. Y. Acad. Sci.** 2015; 1340:95–103.
- Gunderson LL, Sosin H. Adenocarcinoma of the stomach: Areas of failure in a re-operation series (second or symptomatic look) clinicopathologic correlation and implications for adjuvant therapy. **Int. J. Radiat. Oncol.** 1982; 8:1–11.
- Halliwell B, Chirico S. Lipid peroxidation: its mechanism, measurement, and significance. **Am. J. Clin. Nutr.** 1993; 57:715S-725S.
- Halliwell B, Gutteridge JMC. Free Radicals in Biology and Medicine. In: **Acta Crystallogr. Sect. D Struct. Biol.** 5. ed. [s.l.] Oxford University Press, 2017. 73p.384–385.
- Han L, Lee SW, Yoon JH, et al. Association of SOD1 and SOD2 single nucleotide polymorphisms with susceptibility to gastric cancer in a Korean population. **APMIS** 2013; 121:246–256.
- Hayes JD, Dinkova-Kostova AT. The Nrf2 regulatory network provides an interface between redox and intermediary metabolism. **Trends Biochem. Sci.** 2014; 39:199–218.
- Hayes JD, Flanagan JU, Jowsey IR. Glutathione Transferases. **Annu. Rev. Pharmacol. Toxicol.** 2005; 45:51–88.
- Holterman CE, Franovic A, Payette J, Lee S. ETS-1 oncogenic activity mediated by transforming growth factor alpha. **Cancer Res.** 2010; 70:730–40.
- Hu B, El Hajj N, Sittler S, Lammert N, Barnes R, Meloni-Ehrig A. Gastric cancer: Classification, histology and application of molecular pathology. **J. Gastrointest. Oncol.** 2012; 3:251–61.
- Hu JH, Miller SM, Geurts MH, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. **Nature** 2018; 556:57–63.
- Ivanova D, Bakalova R, Lazarova D, Gadjeva V, Zhelev Z. The impact of reactive oxygen species on anticancer therapeutic strategies. **Adv. Clin. Exp. Med.** 2013; 22:899–908.
- Jin Y, Penning TM. Aldo-Keto Reductases and Bioactivation/Detoxication. **Annu. Rev.**

**Pharmacol. Toxicol.** 2007; 47:263–292.

Jones S. Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C->T:A mutations. **Hum. Mol. Genet.** 2002; 11:2961–2967.

Kairupan C, Scott RJ. Base excision repair and the role of MUTYH. **Hered. Cancer Clin. Pract.** 2007; 5:199.

Kastrinos F, Syngal S. Recently Identified Colon Cancer Predispositions: MYH and MSH6 Mutations. **Semin. Oncol.** 2007; 34:418–424.

Kim CJ, Cho YG, Park CH, et al. Inactivating mutations of the Siah-1 gene in gastric cancer. **Oncogene** 2004; 23:8591–8596.

Klaunig JE, Kamendulis LM, Hocevar BA. Oxidative Stress and Oxidative Damage in Carcinogenesis. **Toxicol. Pathol.** 2010; 38:96–109.

Komor AC, Zhao KT, Packer MS, et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. **Sci. Adv.** 2017; 3:

Krokan HE, Bjoras M. Base Excision Repair. **Cold Spring Harb. Perspect. Biol.** 2013; 5:a012583–a012583.

Kurtz RC, Sherlock P. The diagnosis of gastric cancer. **Semin. Oncol.** 1985; 12:11–8.

Lardelli RM, Schaffer AE, Eggens VRC, et al. Biallelic mutations in the 3' exonuclease TOE1 cause pontocerebellar hypoplasia and uncover a role in snRNA processing. **Nat. Genet.** 2017; 49:457–464.

Lentsch E, Li L, Pfeffer S, et al. CRISPR/Cas9-Mediated Knock-Out of KrasG12D Mutated Pancreatic Cancer Cell Lines. **Int. J. Mol. Sci.** 2019; 20:5706.

Levine RL. Carbonyl modified proteins in cellular regulation, aging, and disease<sup>2,3</sup> 2Guest Editor: Earl Stadtman 3This article is part of a series of reviews on “Oxidatively Modified Proteins in Aging and Disease.” The full list of papers may be found on the homepag. **Free Radic. Biol. Med.** 2002; 32:790–796.

Manz DH, Blanchette NL, Paul BT, Torti FM, Torti S V. Iron and cancer: recent insights. **Ann. N. Y. Acad. Sci.** 2016; 1368:149–161.

Mao Z, Hine C, Tian X, et al. SIRT6 promotes DNA repair under stress by activating PARP1. **Science** 2011; 332:1443–6.

Marinho HS, Real C, Cyrne L, Soares H, Antunes F. Hydrogen peroxide sensing, signaling and regulation of transcription factors. **Redox Biol.** 2014; 2:535–562.

- Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. **Nat. Biotechnol.** 2015; 33:538–542.
- Marx V. Base editing a CRISPR way. **Nat. Methods** 2018; 15:767–770.
- Matsui A, Ikeda T, Enomoto K, et al. Increased formation of oxidative DNA damage, 8-hydroxy-2'-deoxyguanosine, in human breast cancer tissue and its relationship to GSTP1 and COMT genotypes. **Cancer Lett.** 2000; 151:87–95.
- Maynard S, Schurman SH, Harboe C, De Souza-Pinto NC, Bohr VA. Base excision repair of oxidative DNA damage and association with cancer and aging. **Carcinogenesis** 2008; 30:2–10.
- McLennan AG. The MutT motif family of nucleotide phosphohydrolases in man and human pathogens (review). **Int. J. Mol. Med.** 1999;
- Michaels ML, Miller JH. The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). **J. Bacteriol.** 1992; 174:6321–6325.
- Murata M, Thanan R, Ma N, Kawanishi S. Role of Nitrate and Oxidative DNA Damage in Inflammation-Related Carcinogenesis. **J. Biomed. Biotechnol.** 2012; 2012:1–11.
- Murphy MP. How mitochondria produce reactive oxygen species. **Biochem. J.** 2009; 417:1–13.
- Nakamura Y, Kawazoe A, Lordick F, Janjigian YY, Shitara K. Biomarker-targeted therapies for advanced-stage gastric and gastro-oesophageal junction cancers: an emerging paradigm. **Nat. Rev. Clin. Oncol.** 2021; 18:473–487.
- Nakamuva K, Ueyarna T, Yao T, et al. Pathology and prognosis of gastric carcinoma. Findings in 10,000 patients who underwent primary gastrectomy. **Cancer** 1992; 70:1030–1037.
- Nielsen M, Lynch H, Infante E, Brand R. **MUTYH-Associated Polyposis.** [s.l: s.n.].
- Ohtsubo T. Identification of human MutY homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. **Nucleic Acids Res.** 2000; 28:1355–1364.
- Oka S, Ohno M, Tsuchimoto D, Sakumi K, Furuichi M, Nakabeppu Y. Two distinct pathways of cell death triggered by oxidative damage to nuclear and mitochondrial DNAs. **EMBO J.** 2008a; 27:421–432.
- Oka S, Ohno M, Tsuchimoto D, Sakumi K, Furuichi M, Nakabeppu Y. Two distinct pathways of

cell death triggered by oxidative damage to nuclear and mitochondrial DNAs. **EMBO J.** 2008b; 27:421–432.

Oka S, Leon J, Tsuchimoto D, Sakumi K, Nakabeppu Y. MUTYH, an adenine DNA glycosylase, mediates p53 tumor suppression via PARP-dependent cell death. **Oncogenesis** 2014; 3:e121–e121.

Oka S, Nakabeppu Y. DNA glycosylase encoded by MUTYH functions as a molecular switch for programmed cell death under oxidative stress to suppress tumorigenesis. **Cancer Sci.** 2011; 102:677–682.

Oppermann U. Carbonyl Reductases: The Complex Relationships of Mammalian Carbonyl- and Quinone-Reducing Enzymes and Their Role in Physiology. **Annu. Rev. Pharmacol. Toxicol.** 2007; 47:293–322.

Out AA, Tops CMJ, Nielsen M, et al. Leiden open variation database of the MUTYH gene. **Hum. Mutat.** 2010; 31:1205–1215.

Papa L, Hahn M, Marsh EL, Evans BS, Germain D. SOD2 to SOD1 Switch in Breast Cancer. **J. Biol. Chem.** 2014; 289:5412–5416.

Parker A, Gu Y, Mahoney W, Lee S-H, Singh KK, Lu A-L. Human Homolog of the MutY Repair Protein (hMYH) Physically Interacts with Proteins Involved in Long Patch DNA Base Excision Repair. **J. Biol. Chem.** 2001; 276:5547–5555.

Pietrangelo A. Ferroportin disease: pathogenesis, diagnosis and treatment. **Haematologica** 2017; 102:1972–1984.

Pizzino G, Irrera N, Cucinotta M, et al. Oxidative Stress: Harms and Benefits for Human Health. **Oxid. Med. Cell. Longev.** 2017; 2017:1–13.

Raetz AG, David SS. When you're strange: Unusual features of the MUTYH glycosylase and implications in cancer. **DNA Repair (Amst).** 2019; 80:16–25.

Ran FA, Hsu PD, Lin C-Y, et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. **Cell** 2013a; 154:1380–1389.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. **Nat. Protoc.** 2013b; 8:2281–2308.

Riaz N, Blecua P, Lim RS, et al. Pan-cancer analysis of bi-allelic alterations in homologous recombination DNA repair genes. **Nat. Commun.** 2017; 8:857.

Rivlin N, Brosh R, Oren M, Rotter V. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. **Genes and Cancer** 2011; 2:466–474.

- Rodríguez-Zavala JS, Calleja LF, Moreno-Sánchez R, Yoval-Sánchez B. Role of Aldehyde Dehydrogenases in Physiopathological Processes. **Chem. Res. Toxicol.** 2019; 32:405–420.
- Rowland A, Miners JO, Mackenzie PI. The UDP-glucuronosyltransferases: Their role in drug metabolism and detoxification. **Int. J. Biochem. Cell Biol.** 2013; 45:1121–1132.
- Scheiman JM, Cutler AF. Helicobacter pylori and gastric cancer. **Am. J. Med.** 1999; 106:222–226.
- Sheng Y, Abreu IA, Cabelli DE, et al. Superoxide Dismutases and Superoxide Reductases. **Chem. Rev.** 2014; 114:3854–3918.
- Shinmura K, Goto M, Suzuki M, et al. Reduced expression of MUTYH with suppressive activity against mutations caused by 8-hydroxyguanine is a novel predictor of a poor prognosis in human gastric cancer. **J. Pathol.** 2011; 225:414–423.
- Sieber OM, Lipton L, Crabtree M, et al. Multiple Colorectal Adenomas, Classic Adenomatous Polyposis, and Germ-Line Mutations in MYH. **N. Engl. J. Med.** 2003; 348:791–799.
- Sies H. Oxidative stress: a concept in redox biology and medicine. **Redox Biol.** 2015; 4:180–183.
- Sies H, Berndt C, Jones DP. Oxidative Stress. **Annu. Rev. Biochem.** 2017; 86:715–748.
- Sies H, Jones DP. Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. **Nat. Rev. Mol. Cell Biol.** 2020; 21:363–383.
- Siewert JR, Böttcher K, Stein HJ, Roder JD. Relevant Prognostic Factors in Gastric Cancer. **Ann. Surg.** 1998; 228:449–461.
- Simic MG, Bergtold DS, Karam LR. Generation of oxy radicals in biosystems. **Mutat. Res. Mol. Mech. Mutagen.** 1989; 214:3–12.
- Slupska MM, Baikalov C, Luther WM, Chiang JH, Wei YF, Miller JH. Cloning and sequencing a human homolog (hMYH) of the Escherichia coli mutY gene whose function is required for the repair of oxidative DNA damage. **J. Bacteriol.** 1996; 178:3885–3892.
- Smith CM, Chen Y, Sullivan ML, Kochanek PM, Clark RSB. Autophagy in acute brain injury: feast, famine, or folly? **Neurobiol. Dis.** 2011; 43:52–9.
- Sohn BH, Hwang J-E, Jang H-J, et al. Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. **Clin. Cancer Res.** 2017; 23:4441–4449.
- Sosa V, Moliné T, Somoza R, Paciucci R, Kondoh H, LLeonart ME. Oxidative stress and cancer: An overview. **Ageing Res. Rev.** 2013; 12:376–390.

- Sullivan LB, Chandel NS. Mitochondrial reactive oxygen species and cancer. **Cancer Metab.** 2014; 2:17.
- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. **CA. Cancer J. Clin.** 2021; 71:209–249.
- Takao M, Aburatani H, Kobayashi K, Yasui A. Mitochondrial targeting of human DNA glycosylases for repair of oxidative DNA damage. **Nucleic Acids Res.** 1998; 26:2917–2922.
- Takao M, Zhang Q-M, Yonei S, Yasui A. Differential subcellular localization of human MutY homolog (hMYH) and the functional activity of adenine:8-oxoguanine DNA glycosylase. **Nucleic Acids Res.** 1999; 27:3638–3644.
- Tsai-Wu J-J, Su H-T, Wu Y-L, Hsu S-M, Wu CHH. Nuclear localization of the humanmutY homologuehMYH. **J. Cell. Biochem.** 2000; 77:666–677.
- Veskoukis AS, Tsatsakis AM, Kouretas D. Dietary oxidative stress and antioxidant defense with an emphasis on plant extract administration. **Cell Stress Chaperones** 2012; 17:11–21.
- Weeden CE, Asselin-Labat M-L. Mechanisms of DNA damage repair in adult stem cells and implications for cancer formation. **Biochim. Biophys. Acta - Mol. Basis Dis.** 2018; 1864:89–101.
- Win AK, Dowty JG, Cleary SP, et al. Risk of colorectal cancer for carriers of mutations in MUTYH, with and without a family history of cancer. **Gastroenterology** 2014; 146:1208-11.e1–5.
- Win AK, Reece JC, Dowty JG, et al. Risk of extracolonic cancers for people with biallelic and monoallelic mutations in MUTYH. **Int. J. Cancer** 2016; 139:1557–1563.
- Xiao GG, Wang M, Li N, Loo JA, Nel AE. Use of Proteomics to Demonstrate a Hierarchical Oxidative Stress Response to Diesel Exhaust Particle Chemicals in a Macrophage Cell Line. **J. Biol. Chem.** 2003; 278:50781–50790.
- Yang A-P, Ma J-X, Qin W-M, Xie Y-H, Jin Y-X. [Amaxa Nucleofector(TM) nuclear transfection apparatus transfers L1210 cell line]. **Zhongguo shi yan xue ye xue za zhi** 2010; 18:1604–8.
- Ye Z-W, Zhang J, Ancrum T, Manevich Y, Townsend DM, Tew KD. Glutathione S -Transferase P-Mediated Protein S-Glutathionylation of Resident Endoplasmic Reticulum Proteins Influences Sensitivity to Drug-Induced Unfolded Protein Response. **Antioxid. Redox Signal.** 2017; 26:247–261.

**Anexo 1 - Estadiamento AJCC<sup>a</sup> e definições de TNM.**

<b>Estágio</b>	<b>TNM</b>	<b>Descrição</b>
IA	T1, N0, M0	T1 = Tumor invade a lâmina própria, mucosa muscular, ou submucosa.
		–T1a = Tumor invade a lâmina própria ou mucosa muscular
		–T1b = Tumor invade a submucosa.
		N0 = Sem comprometimento de linfonodos regionais.
		M0 = Sem metástase a distância.
IB	T1, N1, M0	T1 = Tumor invade a lâmina própria, mucosa muscular, ou submucosa.
		–T1a = Tumor invade a lâmina própria ou mucosa muscular
		–T1b = Tumor invade a submucosa.
		N1 = Metástase em 1 ou 2 linfonodos regionais.
		M0 = Sem metástase a distância.
	T2, N0, M0	T2 = Tumor invade a muscularis própria. <sup>b</sup>
	N0 = Sem comprometimento de linfonodos regionais.	
	M0 = Sem metástase a distância.	
IIA	T1, N2, M0	T1 = Tumor invade a lâmina própria, mucosa muscular, ou submucosa.
		–T1a = Tumor invade a lâmina própria ou mucosa muscular
		–T1b = Tumor invade a submucosa.
		N2 = Metástase em 3 a 6 linfonodos regionais.
		M0 = Sem metástase a distância.
	T2, N1, M0	T2 = Tumor invade a muscularis própria. <sup>b</sup>
		N1 = Metástase em 1 ou 2 linfonodos regionais.
		M0 = Sem metástase a distância.
	T3, N0, M0	T3 = Tumor penetra o tecido conectivo subseroso sem invasão do peritônio
N0 = Sem comprometimento de linfonodos regionais.		
M0 = Sem metástase a distância.		
IIB	T1, N3a, M0	T1 = Tumor invade a lâmina própria, mucosa muscular, ou submucosa.
		–T1a = Tumor invade a lâmina própria ou mucosa muscular
		–T1b = Tumor invade a submucosa.
		N3a = Metástase em 7 a 15 linfonodos regionais.
		M0 = Sem metástase a distância.
		T2 = Tumor invade a muscularis própria. <sup>b</sup>

	T2, N2, M0	N2 = Metástase em 3 a 6 linfonodos regionais. M0 = Sem metástase a distância.
	T3, N1, M0	T3 = Tumor penetra o tecido conectivo subROSo sem invasão do peritônio N1 = Metástase em 1 ou 2 linfonodos regionais. M0 = Sem metástase a distância.
	T4a, N0, M0	T4a = Tumor invade a sROSa (peritônio visceral). N0 = Sem comprometimento de linfonodos regionais. M0 = Sem metástase a distância.
IIIA	T2, N3a, M0	T2 = Tumor invade a muscularis própria. <sup>b</sup> N3a = Metástase em 7 a 15 linfonodos regionais. M0 = Sem metástase a distância.
	T3, N2, M0	T3 = Tumor penetra o tecido conectivo subROSo sem invasão do peritônio N2 = Metástase em 3 a 6 linfonodos regionais. M0 = Sem metástase a distância.
	T4a, N1, M0	T4a = Tumor invade a sROSa (peritônio visceral). N1 = Metástase em 1 ou 2 linfonodos regionais. M0 = Sem metástase a distância.
	T4a, N2, M0	T4a = Tumor invade a sROSa (peritônio visceral). N2 = Metástase em 3 a 6 linfonodos regionais. M0 = Sem metástase a distância.
	T4b, N0, M0	T4b = Tumor invade estruturas/órgãos adjacentes. <sup>c,d</sup> N0 = Sem comprometimento de linfonodos regionais. M0 = Sem metástase a distância.
	IIIB	T1, N3b, M0
T2, N3b, M0		T2 = Tumor invade a muscularis própria. <sup>b</sup> N3b = Metástase em 16 ou mais linfonodos regionais. M0 = Sem metástase a distância.
T3, N3a, M0		T3 = Tumor penetra o tecido conectivo subROSo sem invasão do peritônio N3a = Metástase em 7 a 15 linfonodos regionais. M0 = Sem metástase a distância.
T4a, N3a, M0		T4a = Tumor invade a sROSa (peritônio visceral). N3a = Metástase em 7 a 15 linfonodos regionais. M0 = Sem metástase a distância.
		T4b = Tumor invade estruturas/órgãos adjacentes. <sup>c,d</sup>



	T4b, N1, M0	N1 = Metástase em 1 ou 2 linfonodos regionais. M0 = Sem metástase a distância.
	T4b, N2, M0	T4b = Tumor invade estruturas/órgãos adjacentes. <sup>c,d</sup> N2 = Metástase em 3 a 6 linfonodos regionais. M0 = Sem metástase a distância.
IIIC	T3, N3b, M0	T3 = Tumor penetra o tecido conectivo subROSo sem invasão do peritônio N3b = Metástase em 16 ou mais linfonodos regionais. M0 = Sem metástase a distância.
	T4a, N3b, M0	T4a = Tumor invade a sROSa (peritônio visceral). N3b = Metástase em 16 ou mais linfonodos regionais. M0 = Sem metástase a distância.
	T4b, N3a, M0	T4b = Tumor invade estruturas/órgãos adjacentes. <sup>c,d</sup> N3a = Metástase em 7 a 15 linfonodos regionais. M0 = Sem metástase a distância.
	T4b, N3b, M0	T4b = Tumor invade estruturas/órgãos adjacentes. <sup>c,d</sup> N3b = Metástase em 16 ou mais linfonodos regionais. M0 = Sem metástase a distância.
IV	Qualquer T, Qualquer N, M1	TX = Tumor primário não pode ser avaliado. T0 = Sem evidência do tumor primário. Tis = Carcinoma in situ: tumor intraepitelial sem invasão da lâmina própria, T1 = Tumor invade a lâmina própria, mucosa muscular, ou submucosa. -T1a = Tumor invade a lâmina própria ou mucosa muscular -T1b = Tumor invade a submucosa. T2 = Tumor invade a muscularis própria. <sup>b</sup> T3 = Tumor penetra o tecido conectivo subROSo sem invasão do peritônio T4 = Tumor invade a sROSa (peritônio visceral) ou estruturas adjacentes. <sup>c,d</sup> -T4a = Tumor invade a sROSa (peritônio visceral). -T4b = Tumor invade estruturas/órgãos adjacentes. NX = Linfonodos regionais não podem ser avaliados. N0 = Sem comprometimento de linfonodos regionais. N1 = Metástase em 1 ou 2 linfonodos regionais. N2 = Metástase em 3 a 6 linfonodos regionais. N3 = Metástase em mais de 7 linfonodos regionais. -N3a = Metástase em 7 a 15 linfonodos regionais. -N3b = Metástase em 16 ou mais linfonodos regionais. M1 = Metástase a distância.

T = tumor primário; N = linfonodo regional; M = metástase a distância.

<sup>a</sup> AJCC: Stomach. In: Amin MB, Edge SB, Greene FL, et al. eds. *AJCC Cancer Staging Manual*. 8<sup>th</sup> ed. New York, NY: Springer 2017, p.203–20.

<sup>b</sup> Um tumor pode penetrar a muscularis própria com extensão para os ligamentos gastrocólico ou gastrohepático, ou para o omentum maior ou menor, sem perfuração do peritônio visceral cobrindo essas estruturas. Nesses casos, o tumor é classificado como T3. Se houver perfuração do peritônio visceral cobrindo os ligamentos gástricos ou o omentum, o tumor deve ser classificado como T4.

<sup>c</sup> As estruturas adjacentes do estomago incluem o baço, colón transverso, fígado, diafragma, pâncreas, parede abdominal, glândula adrenal, rins, intestino menor, e retroperitônio.

<sup>d</sup> Extensões intratumorais para o duodeno ou esôfago não são consideradas invasão de uma estrutura adjacente, mas é classificada usando a profundidade da maior invasão em qualquer um desses sítios.

Anexo 2 - Possíveis edições off-targets dos sgRNAs pelo sistema CRISPR

MUTYH	exon1	NGG SpCas9			
Position	Strand	Sequence	PAM	Specificity Score	Efficiency Score
247		TCTCCCGCTGAGTCGTCTG	TGG	87,3500061	61,69628985

MUTYH spCas9 exon1\_247 REV aaacCAGACGACTCAGGCGGGAGAC  
MUTYH spCas9 exon1\_247 FWD CACCGTCTCCCGCTGAGTCGTCTG

Sequence	PAM	Score	Gene	Chromosome	Strand	Position	Mismatches	On-target
TCTCCCGCTGAGTCGTCTG	TGG	100	ENSG00000132773	chr1	-1	45340225	0	true
ACTGCCCGGAGTCGTCTG	AGG	0,99598475		chr18	1	4749172	4	false
CCTACTGCCGAGTCGTCTG	AGG	0,874468199	ENSG00000271855	chr2	1	9556450	4	false
CTTTCCGCTGTGTCGTCTG	TGG	0,815232558	ENSG00000180287	chr1	1	242089889	4	false
TTTCGCGCCGGGTCGTCTG	AAG	0,750829186		chr12	-1	39145224	4	false
GCCACCGCTGAGTCGCCTG	CAG	0,7231875	ENSG00000066455	chr14	1	92794237	4	false
TCTCCGCTGCCTCGTCTG	CAG	0,566146479	ENSG00000114491	chr3	-1	124730372	3	false
TCTCCAGCTGAGTGTCTG	AGG	0,524560483	ENSG00000224092	chr9	1	1978434	3	false
CCTCTGCCAGAGTCGGCTG	CAG	0,422232863		chr10	-1	1513402	4	false
TCTCAACCTGAGTCGTCTA	GGG	0,318060471		chr5	-1	38062204	4	false
GCTCTGGCCTGAGTTGCTG	CAG	0,290263191		chr10	-1	78946476	4	false
TCTCTGGGTGAGTGTCTG	AGG	0,261592937		chr1	1	29323506	4	false
GCTCCACCTGGGTCGGCTG	AGG	0,257482892	ENSG00000196364	chr16	-1	1261267	4	false
TCTCCAGCAAGAGTCGTCTC	CAG	0,254152313		chr1	-1	2278417	4	false
TCTCCGGCTGCGTCTCTG	CAG	0,242512217		chr3	-1	42998174	3	false
CCTGCCCTGAGTCATCTG	AGG	0,221432937		chr17	1	2906176	4	false
GCCCTCTGAGTCATCTG	GGG	0,220667986		chr4	-1	153181837	4	false
TCCGACGCTGAGTCGTATG	CAG	0,218077001		chr17	-1	29584541	4	false
TCTCTGGCTCAGTCGCCTG	GAG	0,216214272		chr3	-1	4029352	4	false
TCTCCCATGAGTCCTCTG	GAG	0,213423926		chr1	1	76689377	3	false
TTTCACTCTGAGTCCTCTG	TGG	0,212440736		chr1	-1	34479210	4	false
CCTCCGCTGAGCCGCTCG	CAG	0,210819149	ENSG00000053747	chr18	-1	23689652	3	false
TCTGCAGCTGCGTCGTCTG	TGG	0,20936374		chr4	1	186552933	4	false
TCTCCCTTCTGAGTCGCCTA	AAG	0,20668576		chr16	-1	48957979	4	false
TCACTCACCTGAGTCCTCTG	TGG	0,203275436		chr8	-1	141333658	4	false
TCTACCACCTGATTCGGCTG	CAG	0,185914242	ENSG00000240230	chr7	-1	968094	4	false
GCTCACGCTCAGTCCTCTG	AGG	0,185834016		chrX	1	45191025	4	false
ACTCCAGCCAGAGTCTTCTG	TGG	0,182581478		chr5	-1	26861513	4	false
TCTCTACCCGAGTCTTCTG	AAG	0,18096061		chr11	1	128598426	4	false
TGTTCTCTGAGGCTCTG	CAG	0,175142482		chr5	1	142795	4	false
TCTCCCTCAGTAGTCTG	AAG	0,171522421		chr8	1	120804781	4	false
ACTTCGGCTGCGTCCTCTG	CGG	0,16842067	ENSG00000099812	chr19	1	758471	4	false
TCTTCTCTGTGTGGTCTG	TGG	0,153507355		chr2	1	128714542	4	false
TCTACCGCCGAAGTCATCTG	CGG	0,148440866		chr10	1	7332096	4	false
CCTCCGCTGGGTCGTCTCT	GAG	0,146644889	ENSG00000156113	chr10	-1	77638502	4	false
TCCCAAGCTGAGTGGTCTT	GGG	0,143937537		chr6	1	52337101	4	false
CCTCCGCTGTCTCGCCTG	GAG	0,139652934	ENSG00000138080	chr2	1	27082799	4	false
TCTCACTCTGAGTTCCTCTG	CAG	0,127439739		chr18	-1	53283102	4	false
TCTCTGCTGAATCTTCTG	GGG	0,124415151		chr16	-1	8350283	3	false
GCTCCGCTCAGTGGGCTG	GGG	0,112710611		chr16	-1	55695431	4	false
TCCCAAGCTAAGTCCTCTG	GAG	0,102977505		chr9	1	98228037	4	false
TCTCTGGATGAGTCATCTG	CAG	0,101572243		chr2	1	179239377	4	false
TCTACTCTGGGTCATCTG	TGG	0,101431556	ENSG00000073910	chr13	-1	32228572	4	false
TCTCCGGCTGTGTAGTCTT	GAG	0,094628178		chr1	-1	40884915	4	false
CCTCCCTCTGAGTCCCTG	GAG	0,094225542		chr3	-1	10379511	4	false
TCGCCAGCTGTGTCTCTG	TGG	0,092224456		chrY	1	8509388	4	false
TCACTGCTGCGTCTTCTG	CAG	0,092224456		chr2	1	65227494	4	false
TTTCCCGCTGAATCTTCTG	AGG	0,085694546		chr6	1	50722533	4	false
TCTCTGGCTGTCTCTCTG	TGG	0,084051228		chr19	1	1969430	4	false
TCTCCGCTGGCTCTTCTG	GGG	0,081981591		chr5	-1	179165813	3	false

MUTYH	exon2	NGG SpCas9			
Position	Strand	Sequence	PAM	Specificity Score	Efficiency Score
5986		1 AGGAAGCCACGAGCAGCCGT	GGG	72,9058572	62,14738492

MUTYH Cas9 exon2\_5986 REV      aaacACGGCTGCTGCTGGCTTCCTC  
MUTYH Cas9 exon2\_5986 FWD      CACCGAGGAAGCCACGAGCAGCCGT

Sequence	PAM	Score	Gene	Chromosome	Strand	Position	Mismatches	On-target
AGGAAGCCACGAGCAGCCGT	GGG	100	ENSG00000132781	chr1		-1 45334486	0	true
AGGAGGCCAGCAGCAGCCGG	GAG	1,600606061		chr8		-1 22635836	3	false
AGCAAGCCAGAAGCAGCCGT	AGG	1,444340538		chr20		-1 17223922	3	false
AGAAAGCCACGAGCAGCCGT	CAG	1,071101615	ENSG00000186951	chr22		-1 46238920	3	false
AGGAGGCCAGGAGCAGCCCT	GAG	1,062078035		chr14		1 1,06E+08	3	false
AGGAGGCCAGGAGCAGCCCT	GAG	1,062078035		chr14		1 1,06E+08	3	false
GGGACGCCAGGAGCAGCCGT	CAG	0,907117279		chr6		1 1,59E+08	4	false
GGGAAGCCAGGAGCAGCCGT	CAG	0,903609249		chr4		1 83771281	3	false
CCGAGCCACGGGAGCAGCCGT	TAG	0,815232558	ENSG00000004777	chr19		1 35775524	4	false
ACGGAGCCAGGGAGCAGCCGT	CGG	0,757879225		chr11		-1 46391313	4	false
AGGAGGCCAAGTGCAGCCGT	TAG	0,692825215		chr10		1 21495974	4	false
GGGTGCCACGAGCAGCCCT	GGG	0,64125		chr20		1 1300735	4	false
ATGCAGCCAGGAGCAGCCAT	GGG	0,6040625		chr16		-1 88058748	4	false
AAGAAGCTAGGAGCAGCCCT	AAG	0,57739368		chr17		-1 52680919	4	false
AGCAGGCCATGAGCAGCCAT	GAG	0,569310169		chr15		-1 93067810	4	false
TGCAAGCCCCGAGCAGCCGC	GGG	0,56287496		chr9		1 1,32E+08	4	false
AAGAAGCCAGCAGCAGCCGT	GGG	0,552771536	ENSG00000250496	chr4		-1 99022857	4	false
GGGACGCCACGAGCAGCCGA	GGG	0,525151075		chr21		-1 42786076	4	false
CAGAAGCCAGGAGCAGCCAT	GAG	0,420648198	ENSG00000148399	chr9		1 1,38E+08	4	false
CGGAAGTCAAGAGCAGCCCT	TGG	0,412809469		chr22		1 40652319	4	false
AGGTAGCCACGAGCAGCTGT	GAG	0,398231214		chr2		1 626287	3	false
TGGAATGCCACGAGCAGCCGG	CGG	0,363228822	ENSG00000244268	chr3		-1 1,53E+08	4	false
AGTAGGCCACGAGCAGCCCT	GGG	0,346946139	ENSG00000005206	chr19		-1 2345254	4	false
GGGACGCCACGAGCAGCCCT	TAG	0,342929348		chr22		1 17798530	4	false
AGGAAGAGAGAGCAGCCAT	TGG	0,337753202		chr6		-1 1,7E+08	4	false
AGGGAGTACGAGCAGCCGA	TAG	0,33446875		chr19		1 5635769	4	false
AGAAAGCCAGGAGCAGCCCT	AAG	0,332967293		chr9		-1 1,37E+08	4	false
AGGAAGCCCTGAGCAGCCGT	TAG	0,327951525		chr8		-1 56198085	3	false
GGGAAGCCAGGATCAGCCGG	CGG	0,320905809		chrX		1 8731719	4	false
AGGGAGGCCAAGCAGCCGC	AGG	0,311118646		chr16		1 89747167	4	false
ATGAAGCCACGAGCAGAAGT	CGG	0,304403397		chr10		1 3047354	3	false
AGGAAGCCACTGAGCAGCCTT	AGG	0,291009076		chr17		-1 47313754	4	false
AGTAGGCCAGGAGCAGCCGT	AAG	0,285354847		chr12		-1 1,25E+08	4	false
AGGCAGCCCGAGGAGCCGT	GGG	0,281258699		chr18		-1 10055448	3	false
AGGTATCCACGAGCAGCCGC	AGG	0,278667457	ENSG00000182685	chr16		1 2209972	4	false
AGGAGGCCAGGTGCAGCCAT	CAG	0,258121353		chr4		-1 32932712	4	false
AGTAAGCCAGGAGCAGCCAC	TAG	0,257545569		chrX		1 1,03E+08	4	false
AGTAAGCCAGGAGCAGCCCA	GGG	0,257545569		chr20		-1 48272242	4	false
AGAAACCCAGGAGCTGCCGT	CAG	0,255874923	ENSG00000258847	chr14		1 65957702	4	false
AGGGAGCCAGGAGCAGCCCTG	TGG	0,252037406		chr2		1 9110772	4	false
AGGCAGCCAGGAGCAGCCAG	CAG	0,252037406		chr10		-1 3199785	4	false
AGGCAGCCAGGAGCAGCCAG	GGG	0,252037406		chr5		1 1,69E+08	4	false
AGGAAGCCACACAGCCGT	GGG	0,248799279		chr15		-1 51093782	3	false
AGGAGGCCAGGAGCAGCCCA	GAG	0,243493765		chr14		-1 1,06E+08	4	false
AGGAGGCCAGGAGCAGCCCA	GAG	0,243493765		chr14		-1 1,06E+08	4	false
AGGAAGCCACGAGCAGCTGT	CAG	0,238855848		chr9		1 1,02E+08	3	false
AGGCAGGCCAGGAGCAGCTGT	GGG	0,225245782		chr7		1 31051342	4	false
GGGAAACCCAGGAGCAGCTGT	CAG	0,222324795		chr6		1 4373657	4	false
TGGAAACCATGAGCAGCTGT	GGG	0,222324795		chr1		1 1,72E+08	4	false
AATAAGCCACTAGCAGCAGT	GGG	0,220868987		chr10		-1 43792677	4	false

MUTYH

exon1

NGG SpCas9n

Position	Strand	Sequence	PAM	Specificity Score	Efficiency Score
247	1	TCTCCCGCTGAGTCGTCTG	TGG	87,3500661	61,69628985
221	-1	GCGGGAGACGAGCGGTGCA	TGG	84,9130721	54,6894905

MUTYH Cas9n exon1_247(221) REV	aaacCAGACGACTCAGCGGGAGAC
MUTYH Cas9n exon1_247(221) FWD	CACCGTCTCCCGCTGAGTCGTCTG
MUTYH Cas9n exon1_221(247) REV	aaacTGACACCGCTCGTCTCCCGC
MUTYH Cas9n exon1_221(247) FWD	CACCGGGAGACGAGCGGTGCA

Sequence 247	PAM	Score	Gene	Chromosome	Strand	Position	Mismatch	On-target
TCTCCCGCTGAGTCGTCTG	TGG	100	ENSG00000132773	chr1	-1	45340225	0	true
ACTGCCCGGAGTCGTCTG	AGG	0,99598475		chr18	1	4749172	4	false
CCTACTGCCGGAGTCGTCTG	AGG	0,874468199	ENSG00000271855	chr2	1	9556450	4	false
CTTTCCGCCTGTGTCTCTG	TGG	0,815232558	ENSG00000180287	chr1	1	2,42E+08	4	false
TTTTCCGGCCGGGTCTCTG	AAG	0,750829186		chr12	-1	39145224	4	false
GCCACCGCTGAGTCGCCTG	CAG	0,7231875	ENSG00000066455	chr14	1	92794237	4	false
TCTCCCGCTGCTCTCTG	CAG	0,566146479	ENSG00000114491	chr3	-1	1,25E+08	3	false
TCTCCAGCTGAGTCGTCTG	AGG	0,524560483	ENSG00000224092	chr9	1	1978434	3	false
CCTCTGCCAGAGTCGGCTG	CAG	0,42232863		chr10	-1	1513402	4	false
TCTAACCTGAGTCGTCTG	GGG	0,318060471		chr5	-1	38062204	4	false
GCTCTGGCTGAGTTGTCTG	CAG	0,290263191		chr10	-1	78946476	4	false
TCTCTGGGTGAGTGGTCTG	AGG	0,261592937		chr1	1	29323506	4	false
GCTCCACCTGGGTCTGGCTG	AGG	0,257482892	ENSG00000196364	chr16	-1	1261267	4	false
TCTCCAGCAAGAGTCGTCTG	CAG	0,254152313		chr1	-1	2278417	4	false
TCTCCCGCTGCTCTCTG	CAG	0,242512217		chr3	-1	42998174	3	false
CCTGCCTCTGAGTCATCTG	AGG	0,221432937		chr17	1	2906176	4	false
GCCCCCTCTGAGTCATCTG	GGG	0,220667986		chr4	-1	1,53E+08	4	false
TCCGCAGCCTGAGTCGTATG	CAG	0,218077001		chr17	-1	29584541	4	false
TCTCTGGCTCAGTCGCCTG	GAG	0,216214272		chr3	-1	4029352	4	false
TCTCCCCATGAGTCCTCTG	GAG	0,213423926		chr1	1	76689377	3	false
TTTCACTCTGAGTCCTCTG	TGG	0,212440736		chr1	-1	34479210	4	false
CCTCCGCCTGAGCCGCTCG	CAG	0,210819149	ENSG00000053747	chr18	-1	23689652	3	false
TCTGCAGCTGCGTGGCTG	TGG	0,20936374		chr4	1	1,87E+08	4	false
TCTCCCTCTGAGTCGCCTA	AAG	0,20668576		chr16	-1	48957979	4	false
TCACTCACCTGAGTCCTCTG	TGG	0,203275436		chr8	-1	1,41E+08	4	false
TCTACCACCTGATTCGGCTG	CAG	0,185914242	ENSG00000240230	chr7	-1	968094	4	false
GCTCAGCCTCAGTCCTCTG	AGG	0,185834016		chrX	1	45191025	4	false
ACTCAGCAGAGTCCTCTG	TGG	0,182581478		chr5	-1	26861513	4	false
TCTCTACCCGGAGTCCTCTG	AAG	0,18096061		chr11	1	1,29E+08	4	false
TGTTCTCTGAGGCGTCTG	CAG	0,175142482		chr5	1	142795	4	false
TCTTCCCTCAGTAGTCTG	AAG	0,171522421		chr8	1	1,21E+08	4	false
ACTTCGCCTGCGTCTCTG	CGG	0,16842067	ENSG00000099812	chr19	1	758471	4	false
TCTTCTCTGTGTGGTCTG	TGG	0,153507355		chr2	1	1,29E+08	4	false
TCTACCGCGAAGTCATCTG	CGG	0,148440866		chr10	1	7332096	4	false
CCTCCCGCTGGGTCTCTCT	GAG	0,146644889	ENSG00000156113	chr10	-1	77638502	4	false
TCCCAAGCTGAGTGGTCTT	GGG	0,143937537		chr6	1	52337101	4	false
CCTCCCGCTGCTCGCTG	GAG	0,139652934	ENSG00000138080	chr2	1	27082799	4	false
TCTCACTCTGAGTTGCCTG	CAG	0,127439739		chr18	-1	53283102	4	false
TCTCTGCCTGAATCTCTG	GGG	0,124415151		chr16	-1	8350283	3	false
GCTCCCGCTCAGTGGGCTG	GGG	0,112710611		chr16	-1	55695431	4	false
TCCCAAGCTAAGTCCTCTG	GAG	0,102977505		chr9	1	98228037	4	false
TCTCTGGATGAGTCATCTG	CAG	0,101572243		chr2	1	1,79E+08	4	false
TCTACCTCTGGGTCTCTG	TGG	0,101431556	ENSG00000073910	chr13	-1	32228572	4	false
TCTCCCGCTGTGTAGTCTT	GAG	0,094628178		chr1	-1	40884915	4	false
CCTCCCTCTGAGTCCCTCTG	GAG	0,094225542		chr3	-1	10379511	4	false
TGCGCAGCTGTGTCTCTG	TGG	0,092224456		chrY	1	8509388	4	false
TCACTGCTGCTCTCTCTG	CAG	0,092224456		chr2	1	65227494	4	false
TTTCCCTGAATCTCTCTG	AGG	0,085694546		chr6	1	50722533	4	false
TCTCTGGCTGTGTCTCTG	TGG	0,084051228		chr19	1	1969430	4	false
TCTCCCGCTGCTCTCTG	GGG	0,081981591		chr5	-1	1,79E+08	3	false

Sequence 221	PAM	Score	Gene	Chromosome	Strand	Position	Mismatch	On-target
GCGGGAGACGAGCGGTGTCA	TGG	100	ENSG00000132773	chr1		45340251	0	true
GCGGCAGACTGGCGGTGTCA	AGG	1,365955696		chr10		1 77662689	3	false
GCAGGAGACAGGGGTGTCA	TGG	1,085740393		chr12		-1 3319802	3	false
CCGGGAGCCGAGCGGTGTCA	GGG	0,981117534		chr16		-1 84408508	3	false
GCGGGAGCGCGCGGTGTCC	CGG	0,775531746		chr14		1 1,04E+08	3	false
TGGGCAGACGAGTGTGTCA	AGG	0,666032609		chr10		-1 1,11E+08	4	false
CAGGGAGACAGCGGTGTGA	GGG	0,641947011		chr2		1 96820141	4	false
GCCGGCCTCGAGCGGTGTCA	CGG	0,573703296	ENSG00000189401	chrX		-1 70063342	4	false
GGGGGCGGGAGCGGTGTCA	GAG	0,551002484		chr5		1 1009622	4	false
GAGAGAGATGAGTGTGTCA	GAG	0,403051705		chr4		-1 1,88E+08	4	false
GCAGGAAACAGGGGTGTCA	TGG	0,390462178		chr10		1 71579243	4	false
GAGGGAGAGAGAGGTGTCA	GGG	0,388192656		chr19		1 45285476	4	false
GAGAGAGACGGCGGTGTGA	CAG	0,368529031	ENSG00000102103	chrX		1 48898050	4	false
GGGGGAGACAGCGGTGTCA	GAG	0,36325442		chr1		1 25024293	4	false
GGGGGAGCGGGCGGTGTCA	CAG	0,359235475	ENSG00000161653	chr17		1 44006834	4	false
CCTGGAACGAGCGTGTCA	TAG	0,333142451		chr6		-1 37102611	4	false
GCAGGAGACTTGGGTGTGTCA	AAG	0,311595017		chr17		-1 77946443	4	false
CCGGGAGAGGGCGGTGTCC	GGG	0,301653696		chr6		-1 1,37E+08	4	false
GTGGGAGACACTGTGTGTCA	AGG	0,290570895		chr8		1 80961211	4	false
GCTGGACAGGAGCGGTATCA	CAG	0,286475974	ENSG00000108774	chr17		-1 42128719	4	false
GCTGGACAGGAGCGGTATCA	TAG	0,286475974	ENSG00000213731	chr10		-1 74423796	4	false
GCTGGAGACATGGGTGTCA	GAG	0,278724543		chr8		-1 1,28E+08	4	false
GGGGGAGACAGCGGTGTGC	TAG	0,271059097		chr20		-1 39146995	4	false
GCAGGAGACGGGAGTGTCA	GGG	0,262133355		chr1		1 1,93E+08	3	false
GAGGGAGAGAAGCGGTGTCA	GGG	0,262081669		chr20		1 63092719	4	false
GCAGGGGACAAGCGATGTCA	GGG	0,255874923		chr8		-1 5478220	4	false
GGGGGTGAGGAGAGGTGTCA	GAG	0,239267008		chr12		1 94529441	4	false
GCTGGAGCCAGAGTGTCA	CAG	0,224201868		chr15		1 77825647	4	false
TCGGGAGCCGTGCGGTGTCA	GGG	0,218942797		chr16		-1 86965496	4	false
GAGGGAGATGAACGGTGTCA	GGG	0,218152458		chr5		1 10460593	4	false
GAGGGAGATGGGTGTGTCA	GGG	0,217451081	ENSG00000105486	chr19		-1 48169619	4	false
GCGGGAGCGAGCTGTGTCC	AGG	0,20820582	ENSG00000165084	chr8		-1 68330879	3	false
CAGGGAGAGGAGCGGGTCA	TGG	0,206845442		chr8		1 55372067	4	false
GGGGGAGGGAAGGTGTCA	GAG	0,194577468		chr15		-1 93153206	4	false
GCGAGAGGAGTGGTGTCC	CGG	0,194072736		chr13		-1 1,07E+08	4	false
GAGGGAGGGGAGCGGGTCA	GAG	0,188135804		chr14		1 1,06E+08	4	false
GCGGAGAAGATGGGTGTCA	GGG	0,184200003		chr7		-1 322531	4	false
ACGGGAGAAGATGGTCTCA	CAG	0,179179864		chr3		1 1,7E+08	4	false
GTGGGAGGAGAGGTTTCA	AGG	0,173430778		chr20		1 22108827	4	false
AGGGGAGGAGGCTGTGTCA	GGG	0,168044786		chr12		1 1,12E+08	4	false
GTGTGAGACACCGGAGTCA	CAG	0,162958378	ENSG00000171241	chr16		-1 46595611	4	false
CCAGGAGACGTGTGTGTCA	CAG	0,15208195		chr11		-1 1,14E+08	4	false
GAGGGAGAAGAGAGGTGTGA	GAG	0,151628051		chr15		1 89672039	4	false
GTGTGAGACGGCTGTGTCA	GAG	0,148040672	ENSG00000173218	chr1		1 1,16E+08	4	false
ACGAGAGACGAGAGGGTCA	TGG	0,133974153		chr4		1 43372204	4	false
GCGGGAGACCGAGGTGTCA	GGG	0,13083914	ENSG00000150093	chr10		1 32957794	4	false
GCGGGTGGGGAGCGGTGGCA	CAG	0,122327774		chr16		-1 12118134	4	false
GAGGGAGAAGAGGGTGTCA	GGG	0,11344589		chr6		1 39923782	4	false
GTGGGAGATGAGGGTGTCA	GGG	0,11344589		chr3		-1 38894083	4	false
GTGGGAGCAAGGGGGTCA	AAG	0,11200033		chr12		-1 87728395	4	false

MUTYH		exon2	NGG SpCas9n			
Position	Strand	Sequence	PAM	Specificity Score	Efficiency Score	
6054	1	GCATGCTAAGAACAACAGTC	AGG	73,4556496	60,39010601	
5991	-1	CTTCCTGTGACCACTTCCCA	CGG	47,1142723	60,6056946	

MUTYH Cas9n exon2\_6054(5991) FWD CACCCGATGCTAAGAACAACAGTC  
MUTYH Cas9n exon2\_6054(5991) REV aaacGACTGTTGTTCTTAGCATGC  
MUTYH Cas9n exon2\_5991(6054) REV aaacTGGGAAGTGGTCACAGGAAGC  
MUTYH Cas9n exon2\_5991(6054) FWD CACCGCTTCTGTGACCACTTCCCA

Sequence 6054	PAM	Score	Gene	Chromosome	Strand	Position	Mismatches	On-target
GCATGCTAAGAACAACAGTC	AGG	100	ENSG00000132781	chr1	-1	45334418	0	true
ACATTCAAAGAACACAGTC	CAG	1,825175809		chr11	1	1,04E+08	3	false
TAAAGCTTAGAACACAGTC	AGG	1,490585774		chr15	-1	45037381	4	false
CAATACTAAAAACAACAGTC	AGG	1,45825		chr9	1	74869199	4	false
GTCAGCAAAGAACACAGTC	AGG	0,948269911		chr2	1	80140172	4	false
CAAGGCTAAGGACAACAGTC	CAG	0,894654977		chrX	1	1,05E+08	4	false
GAATCCTCAGAACACAGTT	TAG	0,86875		chr6	-1	18938404	4	false
GAATGCTGGAAACAACAGTC	AGG	0,85307625		chr2	-1	1,64E+08	4	false
CCATTCTTAGAGCAACAGTC	AGG	0,822887324		chr20	-1	1611277	4	false
GAATGCTGACTACAACAGTC	TAG	0,8021981		chr12	1	1,07E+08	4	false
GAGTGTGAGAACCACTGTC	CAG	0,731005743		chr12	-1	82800429	4	false
GCTTCCTCAGAACACTGTC	TAG	0,693518269		chr6	1	1,3E+08	4	false
GCATGAGAAGAACACAGTC	CAG	0,63080972		chr6	-1	21074020	3	false
CCATGCTAAGAACACAGCT	AAG	0,6255		chr7	-1	92550435	3	false
GGAAACTAAGAACACAGGC	TGG	0,619993094		chr9	-1	1,3E+08	4	false
ACATGCCGAGAACACAGTG	TGG	0,600378217		chr22	1	42334625	4	false
TCATGGAAAAACAACAGTC	TGG	0,59204552		chr4	1	1,15E+08	4	false
TCATGCCAAAAACAACAGTT	AGG	0,566353146		chr5	1	39237196	4	false
GCATGCTTAGGACAACAGAC	CAG	0,562043147		chr14	1	25031860	3	false
GCATGCTTAGTAGAACAGTC	AAG	0,555226531		chr20	-1	40756146	3	false
GGATGCAAAAGACAACAGTC	CAG	0,552771536	ENSG00000083857	chr4	-1	1,87E+08	4	false
GCAGTCCAAGAACACAGTT	CAG	0,548453615		chr12	-1	479402	4	false
TCATGCTACCAACAACAGTT	TGG	0,49465803		chr4	-1	1,58E+08	4	false
GGATGCTGGGAGCAACAGTC	AGG	0,48023778		chr7	-1	44560231	4	false
GTATCCTGAGAACCACAGTC	CAG	0,475		chr2	1	30721795	4	false
ACTTGCTAGGAACAACAGTC	AGG	0,466831832		chr2	1	2,08E+08	4	false
GCTTGATAAAAAACAACAGTA	AAG	0,466385242		chr3	1	1,84E+08	4	false
GTATACAAAGAAAAACAAGTC	AAG	0,43797375		chr8	1	4983110	4	false
GCATGCTCAGCACAGCAGTC	GAG	0,411362445		chr8	1	27467117	3	false
GCATGCTTAGGACAGCAGTC	AGG	0,411362445		chr7	-1	1,34E+08	3	false
GGGTGCTAGGAACAACAGGC	TGG	0,395355187		chr15	1	60920561	4	false
GCAAGCAAAAAACAACAGAC	AAG	0,373494281		chr7	-1	1,06E+08	4	false
GAATGATAATAAAAAACAAGTC	CAG	0,364081312		chrX	1	1,43E+08	4	false
GTATGCTAGAAAAACAAGTC	AAG	0,357525436		chr2	1	1,9E+08	4	false
GCATGCACAAAAACAACAGAC	TAG	0,337753202		chr10	1	1,06E+08	4	false
GCAAGCTCAGTACAACAGGC	AGG	0,32953125		chr9	1	89754738	4	false
GCATGCTGGTACAACAGTC	CAG	0,318573763		chr7	1	1,54E+08	4	false
TCATGCTCAGAACAACAGCT	AAG	0,314061871		chr1	1	84301401	4	false
GCATACTAAGAACCAACATTA	TAG	0,313718788		chr13	-1	89828947	3	false
GACTGCTAAAAACAAGTC	AGG	0,297562541		chr21	1	42415305	4	false
GCATGAGGAGATCAACAGTC	TGG	0,293223721	ENSG00000162923	chr1	1	2,24E+08	4	false
GCCTGCTCAGAACCAACAGAA	AGG	0,286584383	ENSG00000179088	chr12	-1	1,03E+08	4	false
GCCTGCCAAGATCAACAGTG	CGG	0,284516558		chr5	-1	1889193	4	false
AAATGCTAAGAAAAACAAGAC	AAG	0,280184879		chr6	1	24021443	4	false
ACAGGCTAAGAAAAACAAGAC	AGG	0,273136203		chr7	1	41472365	4	false
CCATTCTAAGAAAAACAAGGC	AAG	0,269743207		chr12	1	14711785	4	false
CCACGCTGAGAACCACAGTC	AGG	0,26673995		chrX	1	1,3E+08	4	false
GAAAGCAAAGAACACAATC	AGG	0,260603689		chr22	1	36912505	4	false
GCCTCCTAAGCACATCAGTC	AAG	0,25993425		chr9	-1	87514132	4	false
GAATGCTGAGAACCACAGTA	TAG	0,244251994		chr12	-1	85383784	4	false

Sequence 5991	PAM	Score	Gene	Chromosome	Strand	Position	Mismatches	On-target
CTTCCTGTGACCACTTCCCA	CGG	100	ENSG00000132781	chr1		1 45334481	0	true
CTTTCTGTGTCCACTTCCCA	GAG	6,161619718		chr10		1 1,21E+08	2	false
CTTCCTGAGGCCACTTCCCA	GAG	5,028448276	ENSG00000136153	chr13		-1 75788886	2	false
CTTTCTGTCACTTCCCA	GGG	3,869666667		chr5		-1 1,13E+08	2	false
CTTCCTGTGGCCACTTCCCT	CAG	3,316855909	ENSG00000204539	chr6		-1 31115389	2	false
CATCTTGTGTCCACTTCCCA	AGG	2,639366516		chr11		1 3130258	3	false
CTTCCTGTGTCTTCCCA	GGG	2,039793072		chr4		1 1,89E+08	2	false
CGTCATGTGACCACTTCCCC	GGG	1,873049645	ENSG00000261606	chr15		1 74830362	3	false
CTGCCTGTGCCACTTCCCC	CAG	1,609602648		chr16		1 67201971	3	false
CTCCCTGTTCCCACTTCCCA	TAG	1,5345273		chr22		1 32775456	3	false
CTGCCTGTGCTCACTTCCCA	TGG	1,444340538		chr15		-1 92518011	3	false
CCACCTGCGTCCACTTCCCA	CAG	1,425166355		chr3		-1 13307982	4	false
CACTCTGGGACCACTTCCCA	CAG	1,422115385		chr8		1 1,24E+08	4	false
CTGACTGTGACCACTTGCCA	CAG	1,389709056		chr10		1 1,17E+08	3	false
CAGGCTGTGGCCACTTCCCA	AGG	1,376650053		chr9		-1 1,36E+08	4	false
ACACATGTGACCACTTCCCA	GAG	1,356225869		chr6		-1 35181806	4	false
GAGGCTGTGACCACTTCCCA	GAG	1,325518868		chr15		-1 95194267	4	false
CTTCCTGCCTCCACTTCCCA	AAG	1,324893309		chr10		1 49600514	3	false
CTGCCTGAGACTTCCCA	GAG	1,17887122		chr9		-1 3846851	3	false
CTTCCTGTGACCTTCCCC	AGG	1,144104851		chr19		1 42350143	2	false
CTTGCTGTGACCACTTCCCA	CAG	1,048360262		chr4		-1 7657188	3	false
GTTTCTTTGTCCACTTCCCA	AAG	0,99598475		chr16		1 72869530	4	false
TTACTGTGACCACTTCCCA	AAG	0,967381888		chr15		1 95739386	3	false
ATCCCTGTGACCACTTCCCA	AGG	0,967381888		chr20		-1 35364804	3	false
CCACATCTGACCACTTCCCA	GAG	0,955825847		chr1		-1 15073893	4	false
CTTCCTCTCCCACTTCCCA	AAG	0,932638593		chr3		1 1,08E+08	3	false
CTACATGTGCCACTTCCCA	TAG	0,924515551		chr9		-1 37644988	4	false
TAACCTGTAACCACTTCCCA	CAG	0,9132825		chr15		1 59338675	4	false
CGTGCTGGGACCACTTCCCC	AGG	0,879031065	ENSG00000111181	chr12		-1 198772	4	false
ATAGCTGTGATCACTTCCCA	CAG	0,87421833		chr6		1 1,63E+08	4	false
GGTCTCGTGACCACTTCCCA	GAG	0,865587349	ENSG00000279038	chr22		-1 38265129	4	false
CTCCCTGTGACAACCTTCCCT	TGG	0,859852881		chr17		1 40489159	3	false
CACCCTGTGTCCACTTCCCT	AGG	0,827636945		chr17		1 75512045	4	false
CTTCCTGATAACTTCCCA	CAG	0,822860153		chr5		1 1,7E+08	3	false
CTGGCAGTGCCCACTTCCCA	GAG	0,81893398		chr1		1 36385480	4	false
CCTCCTGGGACACTTCCCA	GGG	0,8021981		chr1		-1 3466873	4	false
TTAGCTGTGACTACTTCCCA	GGG	0,796410829		chr1		-1 33985262	4	false
CCTCCTGAGGCCACTTCCCT	GAG	0,790868822		chr18		-1 39552596	4	false
CTGGCTGTGGCCACTTCCCC	GGG	0,788917088	ENSG00000183150	chr12		1 12662260	4	false
CTTCCTTTGGCCACTTCCCA	AGG	0,748203178		chr5		-1 79872558	3	false
CCACCTGTGCCACTTACCA	AGG	0,688134882		chr10		1 1,01E+08	4	false
CCCCCTGAGACCACTTCCAA	CAG	0,639582009		chr8		1 1,43E+08	4	false
CTTACTGTGATCTTCCCA	AGG	0,638640845		chr4		1 1,89E+08	3	false
ATTGCTGTGGCTTCCCA	CAG	0,631726959		chr5		1 1,7E+08	4	false
CTTCCAGTGGCCACTTCCCTA	AGG	0,614156892	ENSG00000113739	chr5		-1 1,73E+08	3	false
ATGCCTGTGACCACTTCCCT	CAG	0,613763337		chr8		-1 22714305	4	false
TTTCTTCTGACCACTTCCCA	CAG	0,61104935		chr20		-1 60986978	4	false
CCACCTGTGCCCTTCCCA	GAG	0,610728971		chr5		1 1137236	4	false
GTTCTGGGGCACTTCCAA	TAG	0,597418895		chr2		-1 1,64E+08	4	false
CTTCCTGTAACACTTCCCA	AAG	0,592837031		chrX		-1 8855388	3	false