

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

**AVALIAÇÃO *IN SILICO* DE EVENTOS DE *SPLICING* ALTERNATIVO EM PEIXES-
BOI**

PAMELA KETRYA BARREIROS BAKER

Belém – PA

2022

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

AVALIAÇÃO *IN SILICO* DE EVENTOS DE *SPLICING* ALTERNATIVO EM PEIXES-BOI

PAMELA KETRYA BARREIROS BAKER

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFPA como requisito parcial para obtenção do grau de Doutora em Genética e Biologia Molecular.

Orientadora: Dr^a. Maria Paula Cruz Schneider

Belém
Setembro, 2022

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

AVALIAÇÃO *IN SILICO* DE EVENTOS DE *SPLICING* ALTERNATIVO EM PEIXES-BOI

PAMELA KETRYA BARREIROS BAKER

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFPA como requisito parcial para obtenção do grau de Doutora em Genética e Biologia Molecular.

Banca Examinadora:

Prof^a. Dra. Maria Paula Cruz Schneider (Presidente)

Profa. Dra. Maria Iracilda da Cunha Sampaio (Membro)

Prof. Dr. Leonardo dos Santos Sena (Membro)

Dr. André Luiz Alves de Sá (Membro)

Profa. Dra. Renata Coelho Rodrigues Noronha (Membro)

Profa. Dra. Ândrea Kely Campos Ribeiro dos Santos (Suplente)

BELÉM – PA

Setembro/2022

**AVALIAÇÃO *IN SILICO* DE EVENTOS DE SPLICING ALTERNATIVO EM
PEIXES-BOI**

Tese submetida ao Programa de Pós-Graduação em
Genética e Biologia Molecular da UFPA como
requisito parcial para obtenção do grau de Doutora
em Genética e Biologia Molecular.

Data: 29 de setembro de 2022

Banca Examinadora:

Profa. Dra. Maria Paula Cruz Schneider (Presidente)
Chefe do Laboratório de Genômica e Biotecnologia do Centro de Genômica e Biologia de
Sistemas - UFPA

Profa. Dra. Profa. Dra. Maria Iracilda da Cunha Sampaio (Membro)
Laboratório de Genética e Biologia Molecular – UFPA Campus Bragança

Prof. Dr. Leonardo dos Santos Sena (Membro)
Laboratório de Genética Humana e Médica - UFPA

Dr. André Luiz Alves de Sá (Membro)
Laboratório de Genética Aplicada - UFPA

Profa. Dra. Renata Coelho Rodrigues Noronha (Membro)
Centro de Estudos Avançados da Biodiversidade - UFPA

Dedico esse trabalho a todos aqueles que acreditam no poder do conhecimento científico.

I see trees of green, red roses too.

I see them bloom for me and you.

And I think to myself,

“what a wonderful world!”

Bob Thiele e George David Weiss

AGRADECIMENTOS

Tenho tanto a agradecer que com certeza devo começar me desculpando por não ter espaço hábil para agradecer todos que deveria e o quanto poderia. Todos que devo gratidão sabem que devo tudo a eles e verdadeiramente irão se reconhecer nesse texto. Atenho-me a

compreender que cada um que passou pelo meu caminho até aqui me formou de alguma maneira e eu gosto bastante da forma que venho tomando, mas costumo dizer que gosto mais ainda da minha forma de amanhã, de depois de amanhã. O conhecimento que a ciência proporciona faz com que tenhamos sempre visão de que amanhã vai ser melhor, o mundo, as ideias, nós. Eu sou melhor hoje porque acredito em mim e na ciência que me muda. E tudo muda. Que bom que muda! Se não mudasse o que nos restaria contemplar? E mesmice do nada? Bom é saber o motivo, a razão, e não saber também.

Quem despertou o poder da dúvida em mim foi minha mãe. Obrigada por me apoiar em tudo, mesmo não sabendo as vezes se eu ia dar conta, se você ia dar conta, se nós íamos dar conta. Você é a grande vencedora nisso tudo.

Nessa jornada que não se faz só eu contei com a ajuda de inúmeros líderes. A minha orientadora que foi incansável ao acreditar em mim, sempre se dispondo a uma nova ideia, um novo olhar. Nada faria sem ela. Obrigada professora Dra. Maria Paula Cruz Schneider!

Aos membros fixos e itinerantes do laboratório ao qual faço parte: Lucas Silva, Renan, Ramon, Marjorie, Soraya, Silvanira, Roseane e tantos outros que vão se sentir representados pela sigla do laboratório. Vocês foram pés e mãos no desenvolvimento desse trabalho!

Não poderia deixar de mencionar a Dra. Maíra, que foi incansável na construção desse trabalho juntamente com o Lucas Silva nas análises e em tudo que se possa imaginar. Obrigada mil vezes.

Agradeço o inenarrável apoio familiar que tive ao longo dessa jornada, de todos os lados. Indiscutivelmente eu não poderia ter melhores irmãos, primos, tios e tias nesse mundo.

Agradeço a Universidade Federal do Pará pelo zelo que sempre teve com seus pós-graduandos e pelo empenho do Programa de Pós-Graduação em Genética e Biologia Molecular com seus discentes, em especial a paciência que tiveram com essa aluna.

Aline Oliveira, melhor amiga e irmã de alma, ela sabe de todas as coisas, obrigada por ter me aguentado mais que ninguém nesses longos anos. Sempre nós.

Aos que aqui não citei, mas sabem que viveram comigo nesses anos períodos difíceis e de glórias, saibam que todo dia agradeço a honra de poder dividir o mesmo espaço-tempo nessa terra que vocês. Nunca conseguiria agradecer o suficiente!

Vamos em frente.

SUMÁRIO

INTRODUÇÃO	1
1. REVISÃO DE LITERATURA	3
1.1. O Gênero <i>Trichechus</i>	4
1.2. Estudos genéticos em <i>Trichechus</i> sp.	6
1.3. Mecanismos de <i>splicing</i> e a diversidade biológica	11
1.4. <i>Splicing</i> alternativo e resposta imune	17
2. OBJETIVO	21
2.1. Objetivo geral.....	21
2.2. Estratégias	21
3. MATERIAL E MÉTODOS	22
3.1. Amostras e sequenciamento	22
3.2. Reanálise dos dados do transcrito.....	23
3.3. Comparação com anotação de referência e análise diferencial.....	23
3.4. Análises de enriquecimento funcional	24
3.5. Avaliação da ocorrência de eventos de processamento alternativo	24
3.6. Análise funcional de isoformas alternativas	25
4. RESULTADOS.....	27
4.1. Avaliação da modulação da expressão gênica em peixes-boi em resposta ao evento de maré vermelha	27
4.2. Avaliação da ocorrência de eventos de <i>splicing</i> alternativo em peixes-boi da subespécie <i>Trichechus manatus latirostris</i>	32
4.2.1. Padrões de <i>splicing</i> alternativo obtidos utilizando o pipeline <i>Cufflinks</i>	33
4.2.1.1. Processamento alternativo de genes relacionados ao sistema imune de peixes-boi da subespécie <i>Trichechus manatus latirostris</i>	35
4.2.2. Padrões de <i>splicing</i> alternativo obtidos utilizando o <i>IsoformSwitchAnalyzeR</i>	40
4.2.2.1. Processamento alternativo de genes relacionados ao sistema imune de peixes-boi da subespécie <i>Trichechus manatus latirostris</i>	43
5. DISCUSSÃO.....	45
5.1. O evento da maré vermelha reprime diferencialmente grande quantidade de genes em peixes-boi.....	45

5.2.	Duas estratégias foram necessárias para a verificação da ocorrência de eventos de processamento alternativo em peixes-boi da subespécie <i>Trichechus manatus latirostris</i>	49
5.2.1.	O número de genes diferencialmente modulados é maior do que os genes envolvidos em <i>splicing</i> alternativo	50
5.2.1.1.	O sistema imune dos peixes-boi sofreu intensa modulação transcricional e apresenta grande número de genes processados de forma alternativa	53
5.2.2.	Os padrões de evento de <i>splicing</i> obtidos com a ferramenta <i>IsoformSwitchAnalyzeR</i> diferem do perfil geral observado na literatura.....	56
5.2.2.1.	Proporcionalmente, o número de genes de resposta imune diferencialmente expressos é maior do que o número de genes associados a eventos de <i>splicing</i> alternativo	59
6.	CONCLUSÕES	63
7.	REFERÊNCIAS BIBLIOGRÁFICAS	64

Lista de Figuras

Figura 1: Taxa de variação cronológica de mamíferos marinhos. As barras sólidas mostram os alcances máximos relatados. Ma = milhões de anos atrás (BERTA; SUMICH; KOVACS, 2006)	3
Figura 2. Representação gráfica das espécies de peixes-boi e seus respectivos anos de anotação (Adaptado de JEFFERSON; LEATHERWOOD; WEBBER, 1993).	5
Figura 3: Distribuição geográfica das espécies de <i>Trichechus</i> no mundo com base em seus mitogenomas(DE SOUZA et al., 2021a).....	6
Figura 4. Representação do <i>splicing</i> de mRNA pelo spliceossomo.(WAHL; WILL; LÜHRMANN, 2009)	14
Figura 5. Tipos de <i>splicing</i> alternativo. Éxons são representados por caixas e íntrons, por linhas retas ou segmentadas conectando éxons. Os padrões de <i>splice</i> (à esquerda) observados de duas isoformas comparados conforme indicado pela cor dos padrões de emenda. A classificação correspondente do evento (coluna média) e a abreviatura utilizada (coluna à direita). (Extraído de VITTING-SEERUP; SANDELIN, [s. d.]).	15
Figura 6. Os retângulos em azul escuro representam éxons com <i>splicing</i> constitutivo. Os vermelhos, azul claros e verdes representam <i>splicing</i> alternativo. (a) padrões de <i>splicing</i> alternativo; (b) O <i>splicing</i> alternativo mostra a extensa rede de relações proteína-RNA que regula essa interação envolvendo elementos cis dentro do pré-mRNA e fatores de ação trans que se ligam a esses elementos cis. (Adaptado de Park et al., 2018).	16
Figura 7. Os padrões de <i>splicing</i> observados entre duas isoformas, de forma comparativa, estão indicados por cores, para cada padrão de processamento na coluna à esquerda. A denominação correspondente ao evento está indicada na coluna do meio e a abreviatura utilizada, na coluna à direita (inspirada nos resultados de (VITTING-SEERUP et al., 2014)obtida em (VITTING-SEERUP; SANDELIN, [s. d.]).	26
Figura 8. Categorização funcional dos genes diferencialmente expressos de acordo com o Gene Ontology (GO), excluindo as categorias nominalmente relacionadas ao sistema imune de peixes-boi.	27
Figura 9. Perfil de modulação dos principais grupos enriquecidos na categorização funcional em resposta ao evento de maré vermelha.....	28
Figura 10. Inferência de vias metabólicas moduladas em resposta ao evento de maré vermelha em peixes-boi.	30
Figura 11. Modulação da via de sinalização das proteínas quinase ativada por mitógeno (MAP) quinases em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho, os genes induzidos nas condições experimentais. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados <i>Kegg ontology</i>	31
Figura 12. Modulação da via de sinalização Wnt canônica e não canônica em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho, os genes induzidos nas condições experimentais. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados <i>Kegg ontology</i>	32
Figura 13. Percentual relativo às regiões anotadas a partir dos resultados de BLAST das regiões genômicas.....	33
Figura 14. Categorização das principais funções atribuídas aos genes anotados por meio de alinhamento utilizando a ferramenta BLASTx.....	34

Figura 15. Categorização funcional dos genes possivelmente sofrendo processamento alternativo em peixes-boi.	35
Figura 16. Principais categorias enriquecidas em processos biológicos (BP).	36
Figura 17. Principais categorias enriquecidas em componente celular (CC).	37
Figura 18. Principais categorias enriquecidas em função molecular (MF).	37
Figura 19. Genes processados de forma alternativa avaliados utilizando a ferramenta Kegg ontology.	38
Figura 20. Modulação da via de sinalização TNF em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho estão indicados os genes possivelmente sofrendo splicing alternativo. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados <i>Kegg ontology</i>	38
Figura 21. Modulação da via de sinalização IL-17 em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho estão indicados os genes possivelmente sofrendo splicing alternativo. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados <i>Kegg ontology</i>	39
Figura 22. Modulação da via de sinalização de receptores Toll-like em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho estão indicados os genes possivelmente sofrendo splicing alternativo. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados <i>Kegg ontology</i>	40
Figura 23. Proporção de eventos de splicing alternativo computada pela ferramenta IsoformSwitchAnalyzeR	41
Figura 24. Representação esquemática dos eventos de processamento alternativo (éxon skipping e retenção de íntron) em dois genes, superoxide dismutase 2 (SOD2) e Nedd4 family interacting protein 1 (NDFIP1). As representações foram feitas com base nos resultados de processamento alternativo obtidos <i>in silico</i>	43
Figura 25. Representação esquemática de eventos de processamento alternativo (retenção de íntron) em dois genes, tirosina fosfatase tipo C (PTPRC) e MHC de classe II (HLA-DQA2). As representações foram feitas com base nos resultados de processamento alternativo obtidos <i>in silico</i>	45

Lista de Tabelas

Tabela 1. Padrão de análise de genes diferencialmente expressos na maré vermelha e controle em peixes-boi da Florida (LAZENSKY <i>et al.</i> , 2020).....	9
Tabela 2. Identificação de peixes-boi da Flórida usados no estudo RNA-seq (Lazensky et al., 2020).	22
Tabela 3. Vinte genes mais induzidos em peixes-boi em resposta ao evento de maré vermelha.	28
Tabela 4. Vinte genes mais reprimidos em peixes-boi em resposta ao evento de maré vermelha.....	29

Resumo

A complexidade dos genomas eucariotos não resulta, exclusivamente, de seu conteúdo gênico, mas deve-se, sobretudo, à expressão de diferentes transcritos e proteínas a partir de um mesmo gene. O splicing alternativo é um dos mecanismos mais importantes associados ao processamento de RNAs, expandindo significativamente o conteúdo da informação do genoma para o transcriptoma. A resposta imune de um indivíduo está diretamente ligada ao ambiente e a como esse indivíduo dispõe sua maquinaria celular para reagir a um tipo de infecção. Com a descoberta de que efetivas mudanças no mecanismo de splicing constitutivo ocorriam de acordo com estímulos externos que o organismo sofria, conscientizou-se de que havia uma possibilidade de mudança de função dos linfócitos e de outras células de defesa no processo de resposta imune. Esse mecanismo ocorre através da ativação e desativação das células efetoras do sistema imune por proteínas codificadas através de splicing alternativo. O splicing alternativo é um evento molecular complexo e finamente regulado. Entender as circunstâncias nas quais os eventos de processamento constitutivo são preteridos, os mecanismos associados à regulação do processamento alternativo – e que definem a formação de mRNAs funcionais ou não. Com o objetivo de identificar *in silico* a ocorrência de eventos de processamento alternativo de genes de peixes-boi da subespécie *Trichechus manatus latirostris* reanalizamos um transcriptoma disponível em banco de dados de domínio público e verificamos a ocorrência de processamento alternativo por diferentes estratégias. Nossos resultados evidenciaram maior número de eventos de processamento alternativo utilizando o pipeline Cufflinks em comparação com os resultados obtidos aplicando a ferramenta IsoformSwitchAnalyzeR. A primeira ferramenta, no entanto, não permitiu definir a quais tipos de splicing cada uma das isoformas está relacionada. Observamos ainda que o sistema imune dos peixes-boi não somente sofreu intensa modulação transcricional, como apresenta grande número de genes processados de forma alternativa em resposta ao efeito de toxinas da maré vermelha. Finalmente, observamos que os padrões de evento de splicing obtidos com a ferramenta IsoformSwitchAnalyzeR diferem do perfil geral observado na literatura. Nossos dados precisam ser mais bem explorados, utilizando outras ferramentas de identificação de eventos de splicing, para que possamos assumir que os eventos de processamento alternativo ocorrem de forma distinta em peixes-boi em relação a outros mamíferos. Concluimos que, assim como a modulação da expressão de genes é importante como

resposta dos peixes-boi ao efeito tóxico produzido pelo evento da maré vermelha, aqui evidenciamos que muitos genes são transcritos de forma alternativa também como uma estratégia de resistência aos efeitos nocivos das toxinas.

Palavras-chave: *splicing alternativo*, peixe-boi, expressão diferencial.

Abstract

The complexity of eukaryotic genomes does not result exclusively from their gene content but is mainly due to the expression of different transcripts and proteins from the same gene. Alternative splicing is one of the most important mechanisms associated with RNA processing, significantly expanding the information content from the genome to the transcriptome. An individual's immune response is directly linked to the environment and how that individual disposes his cellular machinery to react to a type of infection. With the discovery that effective changes in the constitutive splicing mechanism occurred according to external stimuli that the organism suffered, he became aware that there was a possibility of change in the function of lymphocytes and other defense cells in the immune response process. This mechanism occurs through the activation and deactivation of the effector cells of the immune system by proteins encoded through alternative splicing. Alternative splicing is a complex and finely regulated molecular event. Understand the circumstances in which constitutive processing events are overlooked, the mechanisms associated with the regulation of alternative processing – and which define the formation of functional or non-functional mRNAs. In order to identify the *in silico* occurrence of alternative splicing events in manatee genes of the subspecies *Trichechus manatus latirostris*, we reanalyzed a transcriptome available in a public domain database and verified the occurrence of alternative splicing by different strategies. Our results showed a greater number of alternative processing events using the Cufflinks pipeline compared to the results obtained using the IsoformSwitchAnalyzeR tool. The first tool, however, did not allow defining which types of splicing each of the isoforms is related to. We also observed that the manatee's immune system not only suffered intense transcriptional modulation, but also presents a large number of genes processed alternatively in response to the effect of red tide toxins. Finally, we observed that the splicing event patterns obtained with the IsoformSwitchAnalyzeR tool differ from the general profile observed in the literature. Our data need to be further explored, using other splicing event identification tools, so that we can assume that alternative processing events occur differently in manatees than in other mammals. We conclude that, just as the modulation of gene expression is important as a response of manatees to the toxic effect produced by the red tide event, here we show that many genes are transcribed in an alternative way also as a strategy of resistance to the harmful effects of toxins.

Keywords: alternative splicing, manatee, differential expression

INTRODUÇÃO

A complexidade dos genomas eucariotos não resulta, exclusivamente, de seu conteúdo gênico, mas deve-se, sobretudo, à expressão de diferentes transcritos e proteínas a partir de um mesmo gene (DE KLERK; 'T HOEN, 2015). O *splicing* alternativo é um dos mecanismos mais importantes associados ao processamento de RNAs, expandindo significativamente o conteúdo da informação do genoma para o transcritoma (CHEN; ZHENG, 2008).

Nos eventos de *splicing* constitutivo, após a sua síntese, os transcritos primários ou pré-mRNAs passam por uma etapa de processamento, gerando o transcrito maduro pela retirada das sequências intrônicas e subsequente junção de éxons (WANG et al., 2015a). No *splicing* alternativo ocorrem variações neste processo, que podem resultar, por exemplo, na utilização alternativa de éxons, uso alternado dos sítios 5' ou 3', e na retenção de íntrons no mRNA maduro.

O *splicing* alternativo é um evento molecular complexo e finamente regulado. Entender as circunstâncias nas quais os eventos de processamento constitutivo são preteridos, os mecanismos associados à regulação do processamento alternativo – e que definem a formação de mRNAs funcionais ou não –, bem como identificar ou mesmo prever a proporção de transcritos finais de cada isoforma constitui ainda um grande desafio. Os sítios de *splicing* dão pistas quanto à possibilidade de ocorrência do processamento alternativo; porém, sabe-se que fatores ambientais ou nutricionais são suficientes para alterar o nível final de transcritos constitutivos e/ou alternativos em um mesmo sítio de *splicing* (CHAUDHARY et al., 2019; HU et al., 2021; LI et al., 2020; PAI; LUCA, 2019).

Em humanos, sabe-se que há um número estimado de 20.000 genes codificadores de proteínas (DUNHAM et al., 2012) e que uma única isoforma de mRNA gerada pode originar aproximadamente dez vezes esse número inicial de genes (DJEALI et al., 2012). Apesar de ser um mecanismo conservado, o *splicing* alternativo é alvo frequente de anomalias genéticas humanas e mutações envolvendo doenças genéticas, ocorrendo geralmente nos silenciadores, nos promotores ou nos locais de quebra de sequência do pré-mRNA, resultando em aberrações proteicas ou isomórficas (PAGANI; BARALLE, 2004).

Estudos relacionados à ligação do mecanismo de *splicing* alternativo com o câncer tem sido largamente explorados (AGRAWAL et al., 2018; BIAMONTI et al., 2014; BONNAL; LÓPEZ-OREJA; VALCÁRCEL, 2020; CLIMENTE-GONZÁLEZ et al., 2017; FAUSTINO; COOPER, 2003) e, com o crescente avanço das tecnologias de sequenciamento, diversos transcritos alternativos têm sido ligados a doença em algum estágio (ZHANG,

Yuanjiao et al., 2021). Tendo em vista de que existe um “*splicing code*”, uma série de características e regras associadas que definem o padrão de *splicing* em qualquer transcrito primário em um dado tipo celular (WANG; BURGE, 2008), pesquisas recentes tentam desenvolver *in silico* uma forma em que a biologia computacional possa prever onde e como ocorrerá um evento de *splicing* alternativo através da técnica computacional chamada *Deep Learning* (JHA; GAZZARA; BARASH, 2017; LEUNG et al., 2014; XIONG et al., 2015) (JHA; GAZZARA; BARASH, 2017; LEUNG et al., 2014; XIONG et al., 2015). Com esse método, Leung e colaboradores conseguiram construir modelos de predição e prever os níveis de *splicing* de éxons individuais em diferentes tecidos estudados de humanos e camundongos (LEUNG et al., 2014).

Tendo em vista as doenças que podem ser desencadeadas pelos erros no processo de *splicing* alternativo, esforços de pesquisas têm sido realizados em busca de estratégias terapêuticas para corrigir esses erros (SCOTTI; SWANSON, 2015). Entre eles, um estudo conduzido por COREY (2017), desenvolveu um medicamento oligonucleotídeo antisense que visa amenizar os efeitos da atrofia muscular espinhal causada por erros de *splicing*. Para se ter uma ideia da força de predição do diagnóstico preditivo do *splicing* alternativo, estudos recentes têm conseguido prever através de padrões pré-estabelecidos a sobrevivência de pacientes com câncer (SHEN et al., 2016; TRINCADO et al., 2016).

Análises em larga escala de isoformas alternativas transcritas vêm contribuindo grandemente com o entendimento da ocorrência e regulação de eventos de *splicing*, rendendo importantes informações a respeito da diversidade de transcritos em eucariotos e das relações de ocorrência de *splicing* alternativo entre espécies e/ou em resposta a diferentes condições ambientais (CHEN et al., 2014; IÑIGUEZ et al., 2017). Partindo de dados de transcritomas (RNA-seq), diversos estudos têm buscado explorar a ocorrência de eventos de *splicing* alternativo. Apesar, no entanto, da importância dos estudos de processamento alternativo e da facilidade em se avaliar sua ocorrência em dados de RNA-seq, tais dados permanecem ainda subexplorados (GUO et al., 2021).

1. REVISÃO DE LITERATURA

Data-se do Eoceno (50 mya) o primeiro fóssil de Sirenia registrado até os dias atuais (Figura 1) (BERTA; SUMICH; KOVACS, 2006). Ao final da última era glacial teve início o período Holoceno (11.784 mya), no qual vivemos até os dias atuais. Esse período foi importante para a biologia evolutiva dos mamíferos aquáticos no geral, pois eles tiveram que se adaptar às condições cada vez mais quentes da última era glacial principal, além de sobreviver ao extenso desmatamento nas latitudes tropicais (BAKER, 2015). Podemos dividir estes animais em três grandes grupos (e um na condição basal) – os cetáceos (Ordem Cetacea; Notoungulata); outra que ocupa o topo da hierarquia – os pinípedes (Ordem Pinnipedia), e ainda o grupo dos Sirênios (peixes-boi) que junto com os cetáceos são exclusivamente aquáticos e completam seu ciclo de vida na água, diferente dos pinípedes (lontras e ursos polares) que são considerados anfíbios.

Estudos de taxonomia e filogenia apontam que a ancestralidade desses três grupos acima citados tem evoluído de três ancestrais comuns terrestres diferentes: artiodátilos, carnívoros e proboscídeos (SPRINGER et al., 2015).

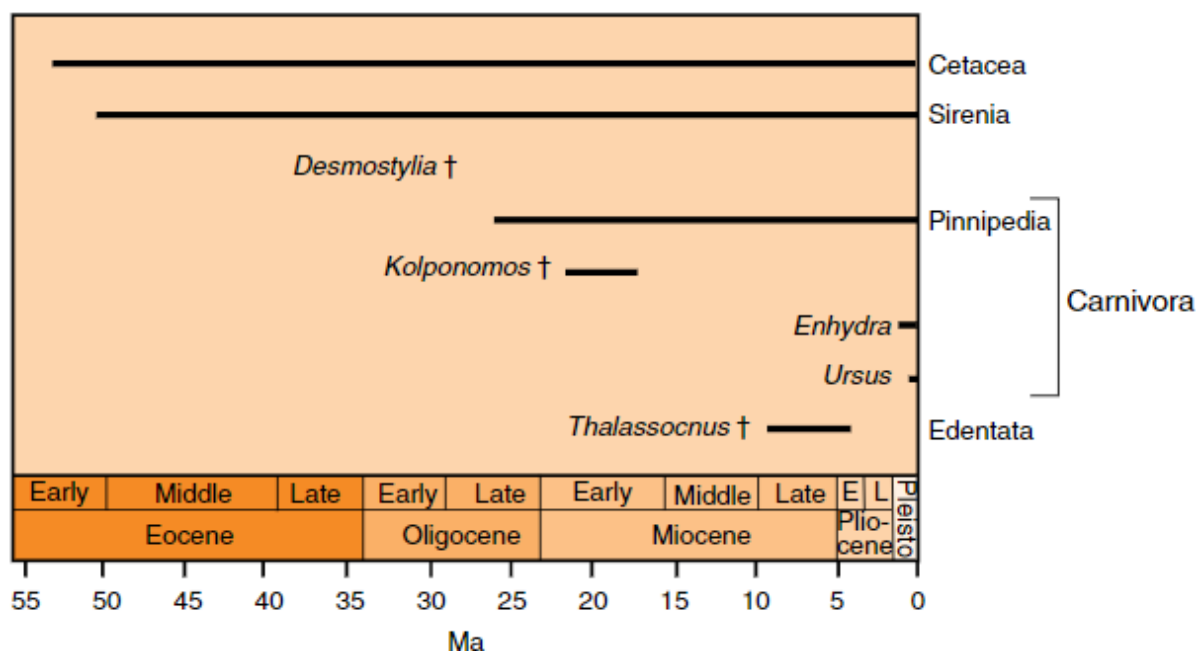


Figura 1: Taxa de variação cronológica de mamíferos marinhos. As barras sólidas mostram os alcances máximos relatados. Ma = milhões de anos atrás (BERTA; SUMICH; KOVACS, 2006)

Acredita-se que o retorno ao ambiente aquático que os mamíferos aquáticos realizaram trouxe não somente adaptações de cunho morfológico e fisiológico, mas também em sua imunologia primária e secundária. (KENNEDY, 1998) relata alguns tipos de infecção em mamíferos aquáticos, incluindo peixes-boi, para espécimes marinhos, especialmente

(KENNEDY, 1998). Ainda assim, estas infecções, tais como a gastroenterite e o vírus da hepatite, ainda se manifestam na maioria dos animais do nosso planeta.

Muitos tipos de infecções têm sido relatados como causas diferentes de morte em mamíferos aquáticos. Na maioria destes casos, infecções alérgicas (incluindo vias digestivas) e por parasitas são as principais motivadoras de morte (KRAUT et al., 1998) dos animais do nosso planeta.

1.1. O Gênero *Trichechus*

A família Trichechidae abrange o gênero *Trichechus* como seu único representante, que por sua vez compreende três espécies: *Trichechus senegalensis*, *Trichechus manatus* e *Trichechus inunguis*. O peixe-boi da África ocidental (*T. senegalensis*) se distribui ao longo atlântico da Costa africana, desde o Sul da Mauritânia até o Centro da Angola e países do interior da Nigéria, Mali e Chade. Devido à preocupação com o risco iminente de extinção da espécie segundo a IUCN (*International Union for Conservation of Nature and Natural Resources*) (KEITH-DIAGNE, 2017), estudos têm sido realizados a fim de preservar os habitats em que os indivíduos ocorrem e avaliar o melhor tipo de manejo dentro das condições desse habitat (BONDE et al., 2012; VIANNA et al., 2006). Como é o caso do estudo de Takouram e colaboradores que avaliou as condições tróficas do lago Ossa, na República dos Camarões, para entender como funciona a dinâmica de disposição de nutrientes no ambiente aquático, que por sua vez irá afetar a turbidez da água e o crescimento de plantas subaquáticas que são o principal alimento de *T. senegalensis*. O estudo demonstra que diante da estação de seca, há baixa disponibilidade de acesso dos indivíduos as áreas de maior crescimento de alimento, prejudicando assim a sobrevivência da população em longo prazo (TAKOUKAM et al., 2021).

O peixe-boi amazônico, *T. inunguis*, se distingue dos demais por habitar em água doce, sendo identificado nos afluentes da bacia amazônica (CANTANHEDE, 2005; SATIZÁBAL et al., 2012). Essa espécie se distribui de acordo com o movimento de enchente-vazante do rio Amazonas (JUNK; SILVA, 1997), procurando forrageio nas várzeas inundáveis e retornando a águas mais profundas quando a vazante do rio se aproxima (Arraut et al., 2010). Essa espécie também está incluída na lista da IUCN (MARMONTEL, M., DE SOUZA, D. & KENDALL, 2016). Recentemente, com a menor frequência de banhistas e movimentação nas praias próximas a capital do Pará, ocorrências dessas espécies têm sido registradas por turistas e moradores, inclusive com comportamento de acasalamento (EMIN-LIMA et al., 2021). Análises de DNA mitocondrial de peixes-boi amazônicos indicaram a

existência de duas populações principais, uma população mais ocidental, que inclui amostras de rios amazônicos peruanos, e uma população oriental, que inclui amostras de rios colombianos e amazônicos no Brasil (SATIZÁBAL et al., 2012). Fatores como isolamento por distância e padrões geográficos históricos podem ser responsáveis por esta subdivisão.

O peixe-boi das Índias Ocidentais, *T. manatus*, habita os oceanos tropicais, costas e rios das Américas. Esta espécie é subdividida em indivíduos da Flórida (*T. m. latirostris*) e das Antilhas (*T. m. manatus*) ocorrendo desde a Costa Sul dos Estados Unidos até o estuário amazônico (LUNA et al., 2008, 2012). O gênero *Trichechus* compreende mamíferos aquáticos exclusivamente herbívoros altamente seletivos e considerados k-estrategistas por terem longo cuidado parental com a prole, possuindo baixa taxa reprodutiva comparada a outras espécies (WEISBROD et al., 2021). Por essas e outras características chave como, acidentes com embarcações, encalhe de filhotes e captura acidental em aparelhos de pesca, estão inclusos na lista da IUCN de espécies ameaçadas de extinção (LUNA et al., 2012).

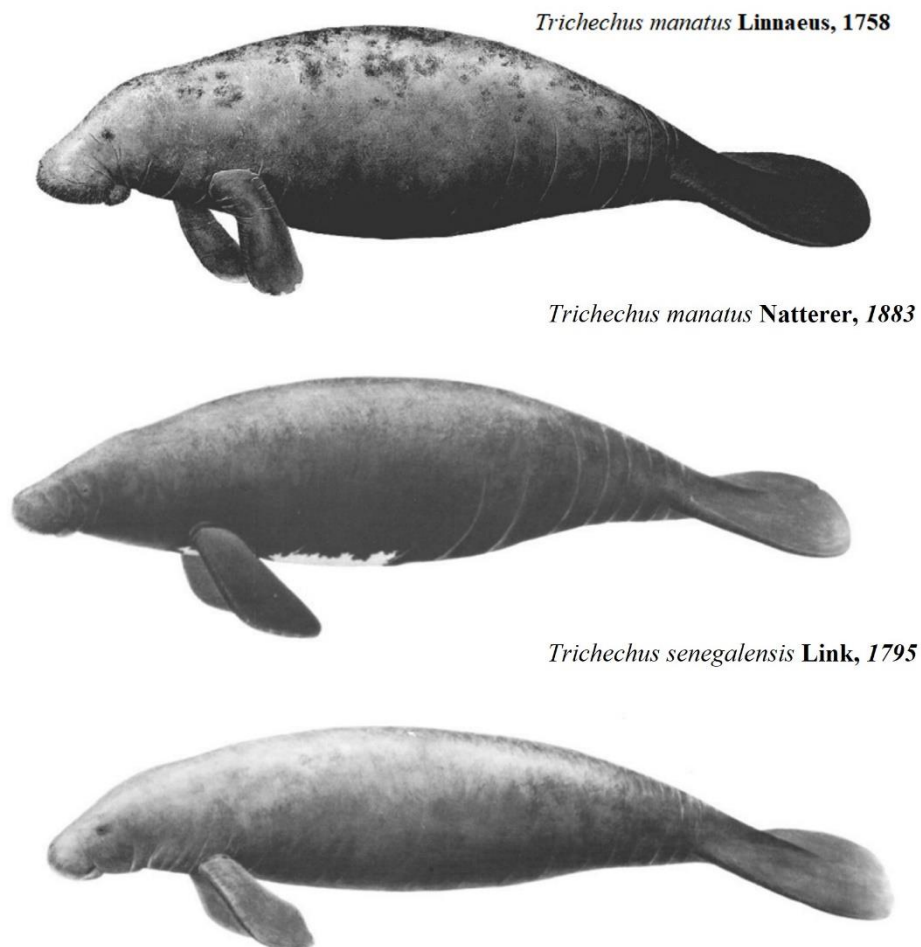


Figura 2. Representação gráfica das espécies de peixes-boi e seus respectivos anos de anotação (Adaptado de JEFFERSON; LEATHERWOOD; WEBBER, 1993).

Por cerca de 40 anos, os dados de peixes-boi da Flórida foram obtidos através de foto registros e identificações de cicatrizes em animais registrados nessas fotografias (BONDE, 2018). Também foram utilizados dados de carcaças via determinação de causa *mortis*, sendo estes naturais ou antropogênicas, sendo essas últimas servindo de alertas às autoridades para ameaças emergentes a esta espécie. Este método serviu para coleta de material biológico, bem como condições de reprodução e determinação de idade (BONDE, 2018). Outros estudos mais recentes vêm monitorando com sensores esses animais em meio selvagem, em via de identificar padrões de alimentação, peso e distribuição nos ecossistemas (BONDE; MCGUIRE; HUNTER, 2012). As espécies de *Trichechus* estão ilustradas quanto à sua distribuição na **Figura 3**.

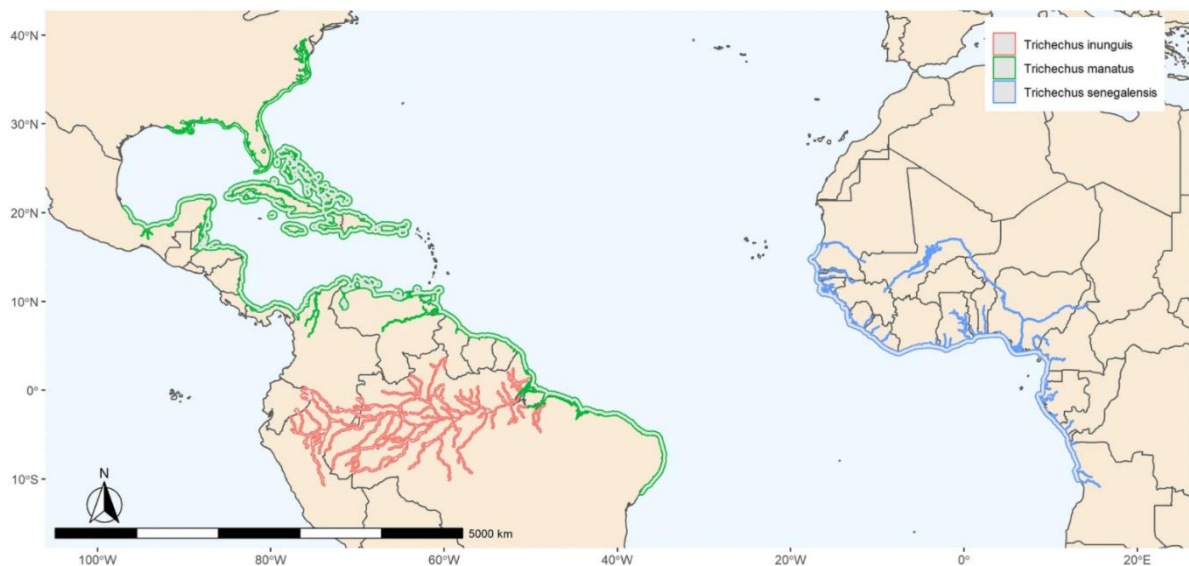


Figura 3: Distribuição geográfica das espécies de *Trichechus* no mundo com base em seus mitogenomas (DE SOUZA et al., 2021a)

1.2. Estudos genéticos em *Trichechus* sp.

Estudos genéticos realizados com mtDNA demonstram que há uma baixa diversidade genética nesse marcador com *Trichechus manatus* (VIANNA et al., 2006). Microsatélites e marcadores neutros mostram também que peixes-boi marinhos possuem populações geograficamente estruturadas (CANTANHEDE, 2005; VIANNA et al., 2006). Luna e colaboradores (LUNA et al., 2012) avaliam também uma baixa diversidade de haplótipos de mtDNA em *T. manatus*, fato que já foi abordado por outros estudos anteriormente (VIANNA et al., 2006). *T. inunguis*, ao contrário das tendências da espécie marinha, mostra uma boa

diversidade genética a baixa estruturação em sua população em sua distribuição (LUNA, Fábila de Oliveira et al., 2021; SATIZÁBAL et al., 2012; VIANNA et al., 2006).

Marcadores genéticos são de suma importância para determinar a ancestralidade entre indivíduos de modo geral e, com a publicação do genoma mitocondrial completo do *T. m. latirostris*, descobriu-se que o genoma mitocondrial anteriormente publicado como sendo de *T. manatus* (ARNASON et al., 2008) na verdade pertencia a uma amostra de *T. inunguis*. Os autores do estudo afirmam ainda que a amostra em erro possa ser provavelmente de origem híbrida entre *T. inunguis* e *T. manatus manatus* (LIMA et al., 2019), tendo em vista a larga semelhança da sequência com *T. inunguis* (VILAÇA; SANTOS, 2020). Uma possível área de hibridização simpátrica é discutida entre especialistas entre as espécies *T. manatus* e *T. inunguis*, onde nas épocas de vazante do rio Amazonas, quando o rio consegue adentrar mais no mar, os indivíduos das espécies conseguem coabitar na mesma área e possivelmente estariam acasalando (DE SOUZA et al., 2021b; GARCIA-RODRIGUEZ et al., 1998; LUNA, Fábila de Oliveira et al., 2021; VIANNA et al., 2006; VILAÇA et al., 2019).

Estratégias que remontam a ancestralidade das espécies citadas são benéficas para entender como elas se distribuem geograficamente e também como pode-se intervir com soluções de manejo sustentável e conservação para todas as espécies de peixes-boi que são sensíveis ao risco de extinção (EXCOFFIER; SMOUSE; QUATTRO, 1992; MORITZ, 1995).

A estrutura genética de populações diferentes foi estudada em peixes-boi das Antilhas mostrando que ambas possuem baixa diversidade genética, mesmo uma das populações sendo tecnicamente mais isolada que a outra. A técnica utilizada para isolar o DNA genômico foi através de amplificações de PCR utilizando 28 microsátélites previamente desenhados para peixes-boi da Flórida (GÓMEZ-CARRASCO et al., 2018).

Outros genes relacionados a evolução genes polimórficos relacionados a imunologia de peixes-boi têm sido realizados recentemente, como o estudo do polimorfismo em TLR4 e TLR8 em *T. m. latirostris* e *T. inunguis* (DE OLIVEIRA et al., 2021).

Em artigo recente publicado por Sá et al. (2022), através da avaliação do gene de DQB de MHC de *T. inunguis* os autores concordam com a manutenção da diversidade genética adaptativa encontrada por outros estudos utilizados marcadores nucleares e mitocondriais nas Américas. Quatro alelos encontrados no estudo, são compartilhados entre *T. manatus* e *T. inunguis* sendo dois desses alelos encontrados nesses animais que ocorrem no Brasil que não ocorrem nos animais que ocorrem na Flórida.

Além das análises genômicas, dados de expressão gênica associados a peixes-boi vêm revelando o perfil transcricional destes organismos. Foi publicado um estudo chave sobre

resposta imune de indivíduos da Flórida envolvendo citocinas, que sabidamente tem importante papel na resposta a infecções bacterianas e virais em mamíferos, bem como em traumas e cura de feridas (FERRANTE; HUNTER; WELLEHAN, 2018). Entre populações saudáveis da costa leste e costa oeste, entre características de sexo e idade, não se observou diferença significativa entre níveis de expressão de citocinas. Porém entre indivíduos da costa leste os níveis de transcritos dos genes GAPDH, IL-2, IL-6 e IFN- γ foram maiores do que para a costa oeste. Esses estudos mostram que tais genes, entre outros citados no artigo, podem ser bons candidatos a investigação imunológica a traumas sofridos ou estresses ambientais (FERRANTE; HUNTER; WELLEHAN, 2018). *Trichechus m. latirostris* torna-se um alvo de comparação evolutiva com humanos, por exemplo (BREAUX et al., 2018). Por serem animais pertencentes aos afroterídeos, grupos dos mais basais clados dos mamíferos eutérios, mantêm um grupo similar de elementos V, subgrupos e similaridades de sequências com humanos, podendo assim realizar-se comparações com os humanos.

Estudos utilizando sequências maduras de mRNA têm sido conduzidos de forma a entender como esses genes transcritos, outrora bem caracterizados por investigações no DNA, podem estar se expressando frente a diferentes condições ambientais às quais esses mamíferos estão sendo expostos (LAZENSKY et al., 2020). Nessas condições é possível avaliar como cassetes de genes são modulados de forma diversa sob condições distintas de estresse e comparar ao genoma que já foi bem descrito na literatura conveniente (HUNTER et al., 2010; LUNA et al., 2012).

Recentemente foi publicado o transcrito do peixe-boi da Flórida avaliando eventos de mortalidade após a proliferação da maré vermelha (LAZENSKY et al., 2020). Nesse estudo, realizou-se o sequenciamento de glóbulos brancos para extração e posterior sequenciamento de RNA de dois grupos, sendo um deles tendo sido exposto a uma floração de maré vermelha no Sudoeste da Flórida e o outro usado como controle. O estudo teve como objetivo central a avaliação transcricional em larga escala. Para tal, os pesquisadores utilizaram como métrica central os maiores e mais reprimidos *fold changes* que encontraram nas análises resultantes. Esses genes foram comparados entre os dois grupos a fim de avaliar que tipo de resposta modulatória o evento “maré vermelha” poderia ter causado nos organismos.

Ainda segundo os autores 158 genes estavam reprimidos e ainda 433 deles induzidos, totalizando 591 genes diferencialmente expressos no grupo afetado pela maré vermelha. Destaque para o gene encontrado como altamente induzido OSCAR (*Osteoclast associated 2C immunoglobulin-like receptor*), e, por conseguinte, o mais induzido PCLO (*Piccolo*

presynaptic cytomatrix protein). OSCAR está classicamente relacionado com a produção e diferenciação de leucócitos, pertencendo à maquinaria do sistema imunológico e a proteína codificada por esse gene pode desempenhar um papel significativo na aterogênese que é mediada pelo estresse oxidativo, sendo a aterogênese a aglutinação de placas de gordura e outros elementos através da inflamação das artérias do corpo do organismo (KIM et al., 2002). Ainda segundo os autores, para OSCAR, ainda foram encontradas múltiplas variantes de transcrição com *splicing* alternativo codificando diferentes isoformas, além de também recrutar monócitos para adesão no processo inflamatório.

Ao lado de dois genes também relacionados a atividades sinápticas e neurais (MTMR2, ANK2) do trabalho anterior (LAZENSKY; HUNTER; AMADOR; et al., 2020), o gene PCLO está ligado a diversas perdas funcionais no cérebro de humanos, como hipoplasia cerebelar tipo 3, que é considerada uma doença rara ligada a má formação do sistema nervoso central ainda na fase embrionária (AHMED et al., 2015). Ainda segundo a pesquisa de Ahmed e colaboradores, PCLO está diretamente ligado a matriz esquelética pré-sináptica e talvez possa estar envolvido na regulação das próprias proteínas pré-sinápticas e suas vesículas, logo, sugere-se que esse gene regule pelo menos uma gama de tipos neuronais no cérebro. Ainda sobre a referida doença, Namavar e colaboradores sugerem que mutações em três genes de subunidade de endonucleases de *splicing* de tRNA esteja ligado a pelo menos três tipos da doença em humanos (NAMAVAR et al., 2011). Outros trabalhos sobre o mesmo gene mostram que o mesmo se associa com desordens de cunho psiquiátrico em humanos como depressão, transtorno bipolar e memória emocional afetada quando há algum distúrbio no funcionamento de PCLO (CHOI et al., 2011; KUEHNER et al., 2011; MBAREK et al., 2017; UNO et al., 2015; WOULDSTRA et al., 2013). Por conseguinte, podemos inferir que sob estresse organismos que tenham alterações nesse gene podem desenvolver perda de funções cognitivas severas, em conjunto com o desencadear de outras alterações que podem surgir com os demais genes em desconformidade que foram relatados por Lazensky e colaboradores (LAZENSKY et al., 2020a).

Tabela 1. Padrão de análise de genes diferencialmente expressos na maré vermelha e controle em peixes-boi da Florida (LAZENSKY et al., 2020)

#	Gene Set Seed	Total # of Genes	Overlapping Genes	Percent Overlap(%)	p-value
1	Neoplasms	4697	157	3	4.62 E-16
2	Inflammation	2351	99	4	2.88

					E-15
3	Cancer	3014	111	3	3.35 E-13
4	Woundsand Injuries	2035	85	4	9.07 E-13
5	Ischemia	993	54	5	2.20 E-12
6	Leukemia	821	48	5	3.48 E-12
7	Arthritis, Rheumatoid	665	42	6	8.04 E-12
8	Neoplasm Metastasis	1947	80	4	1.25 E-11
9	Infection	1914	78	4	3.69 E-11
10	Liver Neoplasms	818	46	5	3.90 E-11
11	Asthma	787	44	5	1.32 E-10
12	BrainIschemia	532	35	6	1.74 E-10
13	Arthritis	657	39	5	2.99 E-10
14	Stomach Neoplasms	783	43	5	3.83 E-10
15	Atherosclerosis	1179	55	4	4.72 E-10
16	Lymphoma	497	33	6	4.75 E-10
17	Lung Neoplasms	1060	51	4	7.86 E-10
18	Intimal Hyperplasia	332	26	7	1.21 E-09
19	Skin Neoplasms	358	27	7	1.31 E-09
20	Colitis	606	36	5	1.51 E-09

¹The total number of genes that belong to a particular pathway based on scanning pubmed.

²The number of genes that are significantly altered in the red tide animals in the particular pathway.

Mais do que quantificar a expressão de genes, os sequenciamentos de RNA (RNA-seq) permitem a descoberta de novos transcritos, detecção de expressão alelo-específica, identificação de RNAs não codificantes, além de possibilitar a busca pela ocorrência de *splicing* alternativo do mRNA (BATTLE et al., 2013; KUKURBA; MONTGOMERY, 2015;

PICKRELL et al., 2010; XIN et al., 2021). Estes últimos dados, que partem igualmente dos dados brutos obtidos no RNA-seq e são obtidos através de análises por pipelines específicos geralmente não são explorados do transcriptoma realizamos no presente trabalho a reanálise dos dados do transcriptoma, explorando eventos de processamento alternativo em resposta às toxinas da maré vermelha em peixes-boi da Flórida. Esta estratégia já rendeu dados significativos de *splicing* alternativo em organismos como o camarão marinho *Litopenaeus vannamei* (ZHANG, Xiaoxi et al., 2019), ostra-do-Pacífico *Crassostrea gigas* (HUANG et al., 2016) e tilápia do Nilo *Oreochromis niloticus* (LI et al., 2020).

Notadamente com aumento da popularidade do sequenciamento por Illumina, muitas ferramentas computacionais tem sido desenvolvidas com o intuito de estimar as isoformas de mRNA e quantificações das variações do *splicing* alternativo através do uso de dados de *short-reads* de RNA-seq (ALAMANCOS; AGIRRE; EYRAS, 2014; CONESA et al., 2016). Essas ferramentas podem levar a dois tipos de análise. A primeira categoria de análise busca estimar a abundância e proporção relativa de isoformas de comprimento total usando dados de *short-reads* de RNA-seq, por meio de um alinhamento com o genoma de referência com as *short-reads*, ou mesmo um transcriptoma, e a partir de então estimar as abundâncias de isoformas de mRNA usando um algoritmo de maximização de expectativas (TRAPNELL et al., 2010; XING et al., 2006). A segunda categoria de análise descobre os eventos de *splicing* alternativo através de dados de RNA-seq, onde as *reads* são alinhadas a éxons específicos ou junções de *splice* são contadas. Dessa forma métodos estatísticos apropriados são aplicados como forma de análise e é possível quantificar o *splicing* alternativo e detectar o *splicing* diferencial entre condições biológicas diversas (KATZ et al., 2010).

1.3. Mecanismos de *splicing* e a diversidade biológica

Didaticamente, a informação contida no genoma é suficiente para a construção de uma proteína viável para executar qualquer tarefa dentro de um organismo vivo. Esse processo ocorre através da leitura e montagem de estruturas que vão apoiar a transcrição de um código que será lido e traduzido para virar então uma proteína funcional.

Esse processo ocorre de acordo com a necessidade do organismo por determinada proteína. A polimerase II é a responsável por fazer a transcrição do DNA que já foi replicado e moldado no DNA do genoma onde está abrigado o gene que representa a receita da proteína a ser produzida. Após a transcrição para RNA, essa fita única passa pelo processo conhecido como *splicing* constitutivo. Esse processo consiste em eliminar as sequências de íntrons e

aproximar as sequências de éxons que são justamente as funcionais. Após esse processo, com um mRNA maduro, ocorrerá a tradução dessa fita para uma proteína (AST, 2005).

Observando a reação imediata do organismo frente às mudanças no ambiente através da transcrição de alguns genes, remete a como a montagem desses genes ocorre, sendo o número de genes inferior ao número de transcritos conhecidos até então (CRAIG VENTER et al., 2001). Para explicar esse fato, existe outro modo de processar as sequências, chamado de *splicing* alternativo, que tem a função de gerar transcritos, até então não anotados no genoma, e torná-los funcionais como uma proteína (ou não) (PAI; LUCA, 2019). O que ocorre nesse processo é nada mais do que o embaralhamento de sequências de íntrons e éxons com o objetivo de gerar uma proteína final distinta da que foi originalmente lida pela polimerase II. Na verdade, sabe-se atualmente que quase todos os transcritos que passam pela leitura da polimerase II sofrem *splicing* alternativo (PIOVESAN et al., 2019). Com isso, podemos entender que o *splicing* alternativo está diretamente ligado às condições ao qual o organismo é submetido e pode responder prontamente a essas alterações.

Os éxons em alguns mamíferos variam de 50 a 300 nucleotídeos (BERGET, 1995) e são cercados por íntrons que chegam a 3.500 nucleotídeos (LONG; DEUTSCH, 1999). Um macrocomplexo de mais de 100 proteínas efetoras e algumas moléculas de RNA (snRNAs) chamado de spliceossomo é o responsável por interagir com sequências específicas localizadas nas fronteiras entre íntrons e éxons, os sítios de *splicing*. Essas extremidades são extremamente conservadas nas extremidades GU (5') e AG na (3'). É nesse ponto que ocorre a separação dos íntrons dos éxons (CHAUDHARY et al., 2019; MATERA; WANG, 2014).

O spliceossomo é um complexo RNA-proteico altamente dinâmico constituído por cinco subunidades de ribonucleoproteínas (RNPs), U1, U2, U4, U5 e U6, além de outras proteínas associadas. Estas subunidades spliceossômicas são denominadas *small nuclear ribonucleoprotein* (snRNPs), e sua atividade é realizada pelo pareamento de bases entre os RNAs que o constituem e as sequências das regiões de ligação no pré-mRNA. O controle dos eventos de *splicing* constitutivo e alternativo, realizado via spliceossomo, depende diretamente dos eventos associados à regulação da sua montagem nos sítios de *splicing* (MATERA; WANG, 2014). A figura 4a mostra elementos conservados de sequências de pré-mRNA de metazoários e leveduras. Os éxons em azul são separados por um íntron em cinza. Na imagem 3b observa-se o ciclo de montagem e desmontagem de íntrons do spliceossomo principal.

A seleção e o reconhecimento dos sítios de *splicing* pela maquinaria dos spliceossomos são influenciados por sequências reguladoras denominadas elementos cis-

acting, que atuam aumentando (*enhancers*) ou diminuindo (*silencers*) a fidelidade no processo de *splicing*. Estes reguladores estão presentes tanto em éxons quanto em íntrons, e suas sequências atuam como regiões de ligação para fatores *trans-acting*. Estes últimos são proteínas reguladoras que atuam coordenando a interação dos snRNPs nos sítios de *splicing* (JUAN; ROCA; ONG, 2014; WANG et al., 2015).

Dentre os muitos fatores *trans-acting*, as proteínas *up-frameshift* (UPF1-3), descobertas inicialmente em *Saccharomyces cerevisiae* (CHANG; IMAM; WILKINSON, 2007), possuem papel relevante no controle de eventos de processamento alternativo. Através de uma via significativamente conservada nos eucariotos, as proteínas UPFs atuam na identificação e degradação de transcritos que apresentem um stop-codon prematuramente inserido. Este importante mecanismo de regulação, conhecido como *nonsense-mediated mRNA-decay* (NMD), funciona como um controle de qualidade, direcionando à degradação mRNAs erroneamente processados (LEWIS; GREEN; BRENNER, 2003). Vale ressaltar que o controle da resposta ao estresse abiótico em plantas, também já foi comprovado por *splicing* alternativo (LALOUM; MARTÍN; DUQUE, 2018).

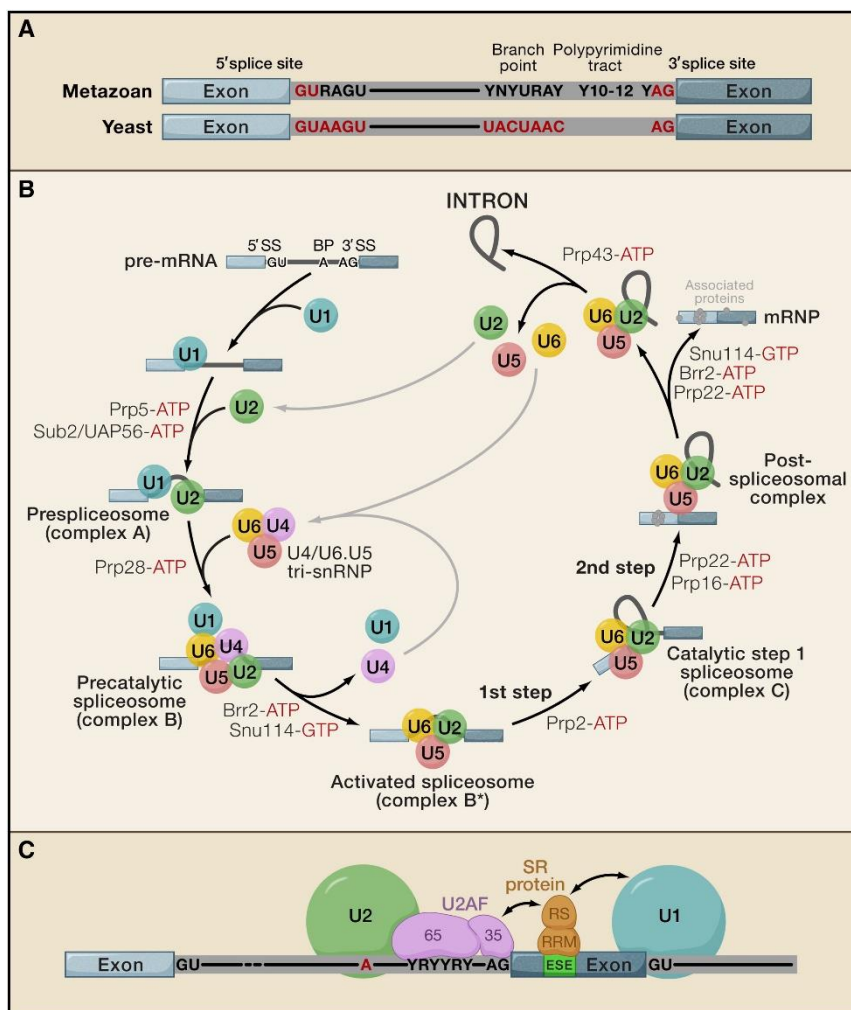


Figura 4. Representação do *splicing* de mRNA pelo spliceossomo.(WAHL; WILL; LÜHRMANN, 2009)

Do ponto de vista da estrutura gênica, o *splicing* alternativo se manifesta nos éxons externos em quatro tipos de eventos Figura 5: (a) exclusão e/ou inclusão do éxon; (b) extremidade 3' alternativa do éxon; (c) extremidade 5' alternativa do éxon; (d) retenção do íntron; (e) e (f) regiões conhecidas como UTRs (*Untranslated Regions*) 5' 3', que são frequentemente encontradas, porém são difíceis de serem caracterizadas devido ao ruído nos dados e a consequente dificuldade de localização dos limites das sequências dos mesmos (Florea, 2006).

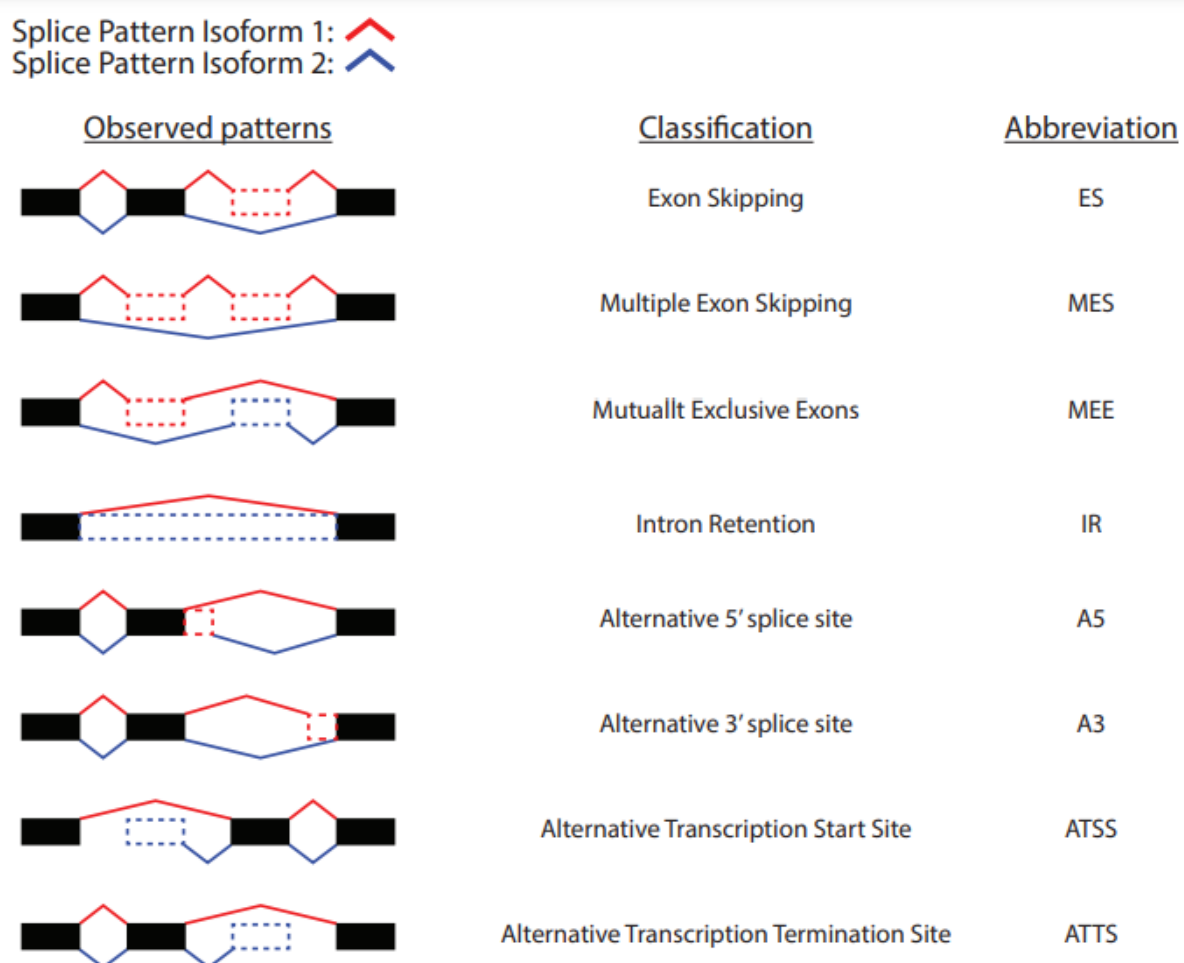


Figura 5. Tipos de splicing alternativo. Éxons são representados por caixas e íntrons, por linhas retas ou segmentadas conectando éxons. Os padrões de *splice* (à esquerda) observados de duas isoformas comparados conforme indicado pela cor dos padrões de emenda. A classificação correspondente do evento (coluna média) e a abreviatura utilizada (coluna à direita). (Extraído de VITTING-SEERUP; SANDELIN, [s. d.]).

Após a remoção dos íntrons, os éxons são realinhados em diversas posições distintas da que originalmente estavam originando vários mRNAs maduros no processo seguinte. Para executar essa função os organismos contam com as proteínas Reguladoras de *Splicing* (SR-*splicing regulators*) que se ligam dentro dos éxons em regiões chamadas ESEs, *Éxonic Splicing Enhancers* para recrutar *snRNA* que por sua vez vai favorecer a inclusão desse éxon no mRNA maduro (Fu, 1995).

Além dos padrões básicos de *splicing* alternativo, são descritos padrões complexos desse mecanismo alternativo envolvendo escolhas éxons ou locais de ligação, conforme exemplificados na Figura 6.

Os sítios de *splicing* mais essenciais dentro do pré-mRNA são os 50SS, 30SS, o sítio de ramificação A e o trato de polipirimidina (Y(n)). Os sítios 50 e 30 conservam os dinucleotídeos GU e AG como o primeiro e os dois últimos nucleotídeos do íntrons, nessa

ordem. Tanto o 30SS quanto o trato de polipirimidina são reconhecidos pelo conjunto de proteínas U2AF. ESEs, ESSs, intensificadores de *splicing* intrônicos (ISEs) e silenciadores de *splicing* intrônicos (ISSs) são motivos reguladores *cis* pré-mRNA que recrutam outras proteínas de ligação ao RNA (SR e hnRNP) para regular a ligação por meio alternativo (PARK et al., 2018) (Figura 6b).

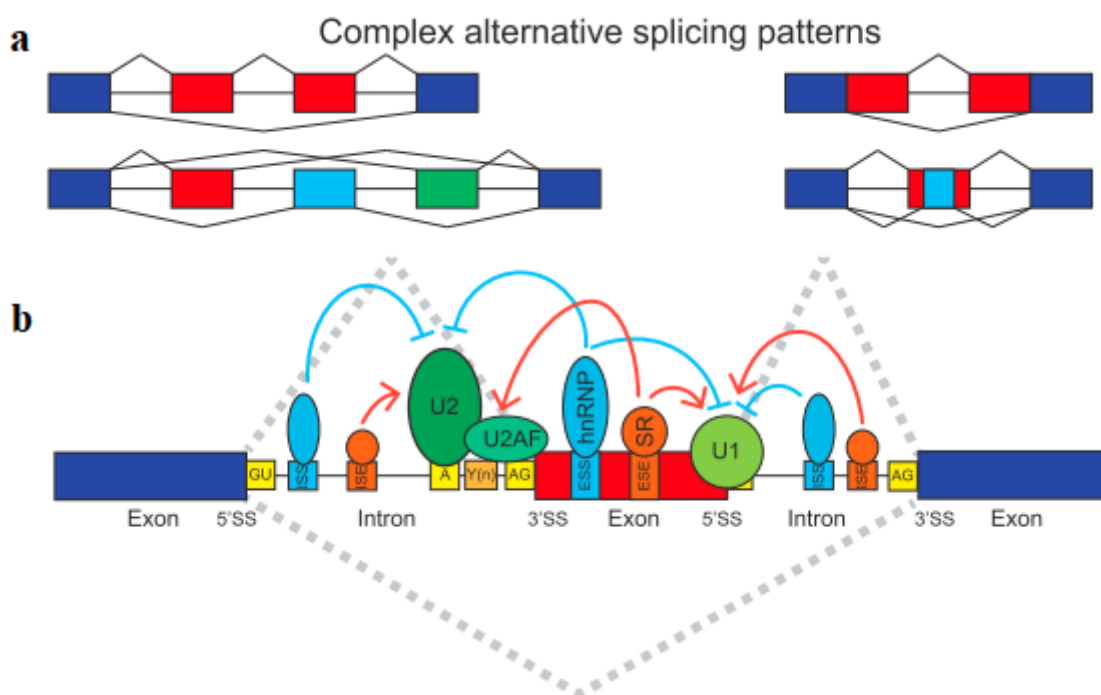


Figura 6. Os retângulos em azul escuro representam éxons com *splicing* constitutivo. Os vermelhos, azul claros e verdes representam *splicing* alternativo. (a) padrões de *splicing* alternativo; (b) O *splicing* alternativo mostra a extensa rede de relações proteína-RNA que regula essa interação envolvendo elementos *cis* dentro do pré-mRNA e fatores de ação *trans* que se ligam a esses elementos *cis*. (Adaptado de Park et al., 2018).

O trabalho de Ehrlich e colaboradores (2019) explora a ocorrência de eventos de *splicing* alternativo no gene Krt10, que codifica para uma queratina epidérmica, em mamíferos terrestres e aquáticos. Sabe-se que a queratina é componente central do citoesqueleto que compõe a epiderme dos mamíferos e é essencial auxiliando a formação de uma barreira contra o meio exterior. O trabalho evidencia duas isoformas do gene Krt10 em mamíferos terrestres. Nos mamíferos aquáticos investigados, este gene não está inteiro, exceto no peixe-boi. Neste último, um ortólogo de Krt10 foi identificado possuindo, no entanto, o quadro de leitura interrompido, sugerindo que Krt10 não codifica para uma queratina completa e funcional. Os dados sugerem que a história evolutiva do gene Krt10 envolveu o surgimento de um padrão de *splicing* alternativo único em mamíferos.

Dessa forma esse mecanismo é de extrema importância dentro da maquinaria biológica adaptativa do organismo. Revelado por estudos envolvendo RNA-seq e adaptações ligadas ao *splicing* alternativo, Ehrlich e colaboradores concluem que há uma diversificação evolutiva do citoesqueleto epidérmico em mamíferos e sugerem uma substituição completa das proteínas epidérmicas quantitativamente predominantes de mamíferos terrestres por queratinas originalmente induzidas por estresse em cetáceos (EHRlich et al., 2019). Tudo isso mostra que através do *splicing* alternativo possa haver um *start* evolutivo para mudanças definitivas no modo com que o organismo expressa suas características (LAZENSKY et al., 2020; PALUSA; ALI; REDDY, 2007). Apesar de alguns estudos sugerirem que à medida que o número de eventos de *splicing* alternativo detectados em estudo de sequenciamentos de larga escala continua a aumentar, argumenta-se que apenas uma pequena fração dessas variantes são reguladas, traduzidas ou têm importância funcional (TRESS; ABASCAL; VALENCIA, 2017). Os autores do estudo ainda, afirmam que a maioria dos éxons envolvidos no processo alternativo estão lá de forma neutra, o que sugeriria que a maioria dos eventos alternativos não resultaria em inovações evolutivas. Obviamente que os autores dos estudos também concordam que os dados podem estar subrepresentados por falhas nas próprias técnicas envolvidas (BLENCOWE, 2017).

1.4. *Splicing* alternativo e resposta imune

A resposta imune de um indivíduo está diretamente ligada ao ambiente e a como esse indivíduo dispõe sua maquinaria celular para reagir a um tipo de infecção. Com a descoberta de que efetivas mudanças no mecanismo de *splicing* constitutivo ocorriam de acordo com estímulos externos que o organismo sofria, conscientizou-se de que havia uma possibilidade de mudança de função dos linfócitos e de outras células de defesa no processo de resposta imune. Esse mecanismo ocorre através da ativação e desativação das células efetoras do sistema imune por proteínas codificadas através de *splicing* alternativo (MARTINEZ; LYNCH, 2013).

Há uma cascata própria para o desencadeamento da resposta imune inata alternativa. As células imunes diferenciadas identificam componentes evolutivamente conservados que são expressos por patógenos por meio de receptores de reconhecimento desses padrões conservados. Com o reconhecimento feito, a proteína MyD88 atua como um adaptador para receber e conectar os sinais recebidos de fora e os sinais que passarão para o interior celular. O controle de complexos de *splicing* SF3A e SF3B entram em ação controlando o *splicing* alternativo, então o gene MyD88 produz uma isoforma de cadeia longa MyD88L e uma curta

MyD88S que vão atuar na ativação e inibição da imunidade inata de células imunes inatas, sendo reguladas pelo receptor *Toll-like* (DE ARRAS; ALPER, 2013; PIZZOLLA et al., 2017). Após a ativação, são desencadeadas outras cascatas de ativação de sinalização intracelular para conduzir a expressão de molecular pró-inflamatórias, interferons (IFN), interleucinas (IL) e quiomocinas, agentes envolvidos na depuração de antígenos (SU; HUANG, 2021).

Não obstante a imunidade inata, a imunidade humoral também é parcialmente perpassada por eventos de *splicing* alternativo. O responsável por produzir os anticorpos que são os principais mediadores da resposta humoral são os linfócitos B, que após a sua maturação, tem como primeiro produto a imunoglobulina IgM, seguida pela IgG, IgE e IgA, produtos resultantes do gene imutável da imunoglobulina após sua ativação. Porém, a transição de IgM para IgD é obtida através de *splicing* alternativo de um transcrito de mRNA primário de um gene de cadeia pesada de Ig, IGH (MAKI et al., 1981; ZHANG; SAXON; MAX, 1992).

A proteína nuclear de ligação ao RNA, chamada de HuR, é conhecida por desempenhar um papel no *splicing* do mRNA bem como na exportação nuclear. A carência de HuR em células B, prejudica a proliferação e diferenciação de linfócitos B, controlando o *splicing* de centenas de genes incluindo também a troca de classes de anticorpos IgM para IgD (DIAZ-MUÑOZ et al., 2015).

Algumas tirosinas kinases já foram identificadas sendo reguladas por *splicing* alternativo em células T, e, pelo menos duas isoformas de cada uma já foram identificadas, com pelo menos uma delas sinalizando essa célula T em humanos. O mesmo acontece com citosinas e interleucinas (LYNCH, 2004). Em análise global utilizando *microarray*, pesquisadores mostraram através de suas análises que a regulação por *splicing* alternativo altera a expressão gênica frente a estimulação das células T e pode ser tão comum quanto a expressão comumente regulada (IP et al., 2007). Todos esses elementos estão relacionados com a sinalização celular para o início ou intermédio da resposta imune do organismo frente a um ataque (LYNCH, 2004) (Martinez and Lynch, 2013). Atualmente, algumas proteínas têm sido estudadas como supostas reguladores do próprio mecanismo de *splicing* alternativo, que é o caso da Tristetrapolina, que desestabiliza os mRNA alvo em sua parte 3'-UTR e mostra ser um potencial gene transcritora e reguladora de *splicing* alternativo na ausência de estímulo externo, segundo demonstrado em ensaios *in vitro* (TU et al., 2019).

O *splicing* alternativo e as isoformas geradas por ele mostra-se atuante no processo de resposta imune contra vírus, que através da translocação de dois fatores reguladores de interferon (IRF3 e IRF7) até o núcleo, levam a transcrição e produção dos tipos I e III de

mRNAs de interferons (IFN). Após a fosforilação desses interferons, dar-se-á início a uma cascata de eventos que culminarão em um estado celular antiviral. Esse estudo anteriormente citado foca exatamente na revisão de como o hospedeiro se comporta com eventos de *splicing* alternativo na regulação da resposta durante uma infecção viral. Os autores discutem ainda que os eventos regulatórios descobertos sobre *splicing* alternativo parecem modular negativamente a resposta antiviral, implicando assim em defeitos na regulação pós-transcricional da resposta imune inata antiviral, que, por sua vez, resultariam em autoimunidade e condições inflamatórias patológicas. (LIAO; GARCIA-BLANCO, 2021).

O meio ambiente ao redor e os mecanismos de *splicing* alternativo tem efeito mútuo um sobre o outro, ambos estão dentro de uma espécie de corrida “armamentista” para se adequar aos ataques sofridos. Na pandemia de COVID-19, causada pelo vírus SARS-CoV-2, pesquisadores revelaram que a proteína NSP16 do vírus se liga aos fatores de *splicing* U1 e U2 em humanos, causando uma supressão de *splicing* de mRNA geral nas células infectadas com o vírus. Mecanismo usado supostamente para prevenir a reação das células com algum tipo de sinalização ou defesa guiadas por ativação dos mecanismos de *splicing* alternativo (BANERJEE et al., 2020).

Com o intuito de dificultar ou mesmo impedir a entrada do vírus na célula hospedeira, pesquisadores trabalharam na modulação preventiva do *splicing* alternativo do pré RNA de ACE2, codificando o receptor da célula para o vírus, excluindo domínios necessários para a entrada do vírus na célula hospedeira mantendo ainda as funções vitais de domínio de ACE2 para o bom funcionamento celular. ACE2 é uma enzima responsável pela conversão de angiotensina 2 e desempenha papel fundamental na manutenção do equilíbrio do sistema renina-angiotensina e também serve de porta entrada para o vírus SARS-CoV-2 (REHMAN; TABISH, 2020).

Extensas revisões sobre abordagens computacionais e experimentais são descritas em uma revisão recente feita por Park e colaboradores (2018), falando das tecnologias recentes que dominam as análises no campo da genética dos transcritos e seus processos, bem como os tipos de análises que estão sendo utilizadas para caracterizar essa sumária de dados obtidos em larga escala.

Tendo em vista a emergente necessidade de aprofundar o entendimento a respeito dos eventos associados ao processamento alternativo de pré-mRNAs, dada a complexidade das informações inseridas neste mecanismo regulatório pós-transcricional, e a disponibilidade do transcrito do peixe-boi da Flórida avaliamos a ocorrência de eventos de *splicing* alternativo em peixes-boi da subespécie *Trichechus manatus latirostris*. Nossas análises tiveram como

foco (i) a identificação de eventos de *splicing* alternativo no peixe-boi e (ii) a quantificação destes eventos exclusivamente em genes associados à resposta imune.

2. OBJETIVO

2.1. Objetivo geral

Identificar *in silico* a ocorrência de eventos de processamento alternativo de genes de peixes-boi da subespécie *Trichechus manatus latirostris* em resposta ao efeito de toxinas da maré vermelha.

2.2. Estratégias

- Reanalisar os dados do transcrito selecionando os genes diferencialmente expressos em resposta ao evento de maré vermelha.
- Identificar variantes de *splicing* alternativo no transcrito do peixe-boi da Flórida por metodologia *in silico*.
- Verificar se as variantes de *splicing* estavam diferencialmente expressas em resposta ao evento de maré vermelha.
- Selecionar genes relacionados ao sistema imune que sofreram processamento alternativo.

3. MATERIAL E MÉTODOS

Os dados brutos utilizados neste trabalho estão disponíveis no NCBI (GEO #GSE86792) e foram publicados por Lazensky e colaboradores (LAZENSKY et al., 2020).

3.1. Amostras e sequenciamento

Amostras de plaquetas da camada leucoplaquetária de peixes-boi foram obtidas de dois grupos: após evento de maré vermelha florescer no sudoeste da Flórida (n = 4) e um grupo controle (n = 7) composto por peixes-boi amostrados em três locais não impactados pelo evento de maré vermelha (LAZENSKY et al., 2020). Os RNAs obtidos foram utilizados em sequenciamento (RNA-seq) e analisados quanto à transcrição gênica. A Tabela 2 apresenta dados gerais sobre data de coleta e condição de saúde dos animais avaliados. A seleção dos peixes-boi incluídos no grupo controle foi feita com base nos valores da proteína de fase aguda amiloide A (SAA) sérica, indicador mais confiável de inflamação ou doença em peixes-boi (HARR et al., 2006). As medições de SAA foram realizadas em soro de peixe-boi (CRAY et al., 2013) sendo considerado ≤ 50 mg/L; exceto para uma amostra.

Tabela 2. Identificação de peixes-boi da Flórida usados no estudo RNA-seq (Lazensky et al., 2020).

Group Name	Same Collection Site	Collection Date	Manatee Identifier/Study ID	Gender	Serum Amyloid A (SAA)* (mg/L)	Health Notes
Red Tide	ZooTampa	3/13/2013	RSW 1307 (16)	M	-	Brevetoxins 4 ng/ml
		2/28/2013	RSW 1310 (17)	M	-	Brevetoxins, not measured
		11/06/2012	103030 (13)	M	-	Brevetoxins 9ng/ml
		02/01/2013	103054 (14)	F	-	Brevetoxins 4ng/nl
Control	Crystal River, FL	12/04/2013	CCR13-14(1)	M	18.96	Excellent condition
		12/04/2013	CCR13-19 (7)	M	25.18	Excellent condition
		12/03/2013	CCR13-10 (10)	M	30.63	Excellent condition
		12/04/2013	CCR13-21 (11)	M	17.72	Excellent condition
	Brevard country, FL	12/10/2013	CBC-13-05 (9)	M	16.77	Excellent condition

	12/10/2013	CBC-13-02 (12)	M	73.98	Excellent/High SAA
ZooTampa	6/21/2014	103150 (8)	F	-	Pneumonia case

*.SAA values were considered high if they were out of the 'normal' range of 0-50 mg/L; SAA values were not collected for manatees at ZooTampa.

3.2. Reanálise dos dados do transcrito

Os dados brutos do RNA-seq depositados foram reanalisados utilizando o genoma do peixe-boi TriManLat1.0 sequenciado pelo Broad Institute (acesso no NCBI por AHIN000000000). O controle de qualidade foi realizado por meio do pacote Trimmomatic v.0.36 (BOLGER; LOHSE; USADEL, 2014). Foi utilizado o pacote HISAT2 (2.0.0) (KIM; LANGMEAD; SALZBERG, 2015) para alinhamento das *reads* com o genoma de referência. Os alinhamentos foram filtrados utilizando o pacote SAMtools (LI et al., 2009), restando apenas leituras devidamente emparelhadas. Os alinhamentos lidos foram processados utilizando o pacote *Cufflinks* v2.2.0 (TRAPNELL et al., 2010) para a montagem das *reads* obtidas em transcritos. Este programa estimou a abundância e procedeu a avaliação de diferencial de expressão nas amostras expostas ao evento de maré vermelha em comparação aos organismos controle.

Cufflinks utiliza o método de montagem de transcritos baseada em anotação de referência (RABT) (ROBERTS et al., 2011) estruturando os dados contra uma anotação de referência conhecida com o intuito de identificar a ocorrência de novos transcritos. A abundância de transcritos foi medida em Fragmentos por Quilobase de éxon por Milhão de Fragmentos mapeados (FPKM - *Fragments Per Kilobase of transcript per Million fragments mapped*). Para cada biblioteca foi gerado um arquivo com a extensão ".bam", contendo todas as *reads* que se alinham ao genoma de referência de todas as amostras. *Cufflinks* foi executado utilizando os parâmetros padrão.

3.3. Comparação com anotação de referência e análise diferencial

Uma vez concluída a montagem das *reads* em transcritos os arquivos gerados foram utilizados como entrada para o *Cuffcompare*. A ferramenta *Cuffcompare* confronta a montagem realizada com transcritos conhecidos de diferentes organismos, explorando diferentes bibliotecas de RNA-seq; também avalia a qualidade da montagem. Foram comparadas as bibliotecas dos animais controle contra as bibliotecas de animais sob efeito do evento da maré vermelha (grupo teste).

O arquivo gerado pelo *Cuffcompare* serviu de entrada para uso do *Cuffdiff*, para identificação de transcritos diferencialmente expressos comparando os dois grupos de amostras. Esta ferramenta assume que os dados apresentam distribuição normal, ou seja, o número de leituras produzidas por cada transcrito é proporcional à sua abundância. *Cuffcompare* e *Cuffdiff* foram executados empregando parâmetros padrão, partindo do genoma de peixe-boi TriManLat1.0 como referência.

3.4. Análises de enriquecimento funcional

Com os resultados obtidos foi gerada uma tabela contendo os genes diferencialmente expressos (GDE), considerando um valor de $p < 0,05$ e valores de p ajustados ($FDR < 0,05$). Um critério de exclusão de 4 vezes ($\log_2$ fold change ≥ 2 ou ≤ -2) foi aplicado como limiar de expressão. Um total de 1.429 genes diferencialmente expressos foi identificado. A anotação funcional dos GDEs foi realizada de acordo com GO (BLAKE; HARRIS, 2008). O genoma de *Homo sapiens* foi utilizado como referência.

A ferramenta ShinyGO v0.76 (GE et al., 2020) foi utilizada para estruturar graficamente os dados de enriquecimento e para verificação das vias metabólicas moduladas em resposta ao evento da maré vermelha. A anotação das vias metabólicas foi realizada com base no *Kegg* (Kyoto Encyclopedia of Genes and Genomes) *Orthology* (KANEHISA; GOTO, 2000). Foram utilizados os parâmetros padrão das ferramentas ShinyGO e *Kegg Orthology* e o genoma de *Homo sapiens* como referência para a anotação.

3.5. Avaliação da ocorrência de eventos de processamento alternativo

Os arquivos de saída do *Cuffdiff* foram utilizados também para identificar eventos de processamento alternativo de genes. Para cada transcrito primário é feita a quantificação de isoformas alternativas detectadas, ou seja, quanto *splicing* diferencial existe entre as isoformas processadas a partir de um único transcrito primário. As isoformas diferencialmente expressas foram identificadas usando um p valor ajustado de 0,05. Foi utilizado critério de exclusão de 4 vezes ($\log_2$ fold change ≥ 2 ou ≤ -2) como limiar para os genes processados de forma alternativa.

No total, 852 regiões apresentando possíveis ocorrências de eventos de processamento alternativo foram identificadas, em 234 genes. Para anotação das regiões genômicas identificadas como processadas alternativamente foram recuperadas as sequências do genoma de *Trichechus manatus latirostris*, disponível no NCBI GenBank (AHIN000000000) por meio do módulo *getfasta* do programa BEDTools v2.30 (QUINLAN; HALL, 2010). A anotação dos

genes presentes nas regiões processadas alternativamente foi realizada por meio de alinhamento BLASTx utilizando o NR database do NCBI.

Realizou-se anotação funcional de acordo com GO (BLAKE; HARRIS, 2008) utilizando *Homo sapiens* como genoma de referência. A ferramenta *ShinyGO* (GE et al., 2020) foi utilizada para estruturar graficamente os dados de enriquecimento e para verificação dos efeitos do evento da maré vermelha sobre o processamento alternativo em vias metabólicas. A anotação das vias metabólicas foi realizada com base no Kegg Orthology (KANEHISA; GOTO, 2000). Foram utilizados os parâmetros padrão das ferramentas *ShinyGO* e Kegg Orthology.

3.6. Análise funcional de isoformas alternativas

Foi aplicada a ferramenta IsoformSwitchAnalyzeR (v1.16.0) para agrupar os eventos de processamento alternativo por tipos específicos de *splicing* (Vitting-Seerup and Sandelin, 2019). Oito categorias de eventos de *splicing* são caracterizadas utilizando este pipeline, incluindo *exon skipping* (ES - uso alternativo de éxon), *alternative 5' splice sites* (A5 - sítios doadores [5'] alternativos), *alternative 3' splice sites* (A3 - sítios aceptores [3'] alternativos), *intron retention* (IR - retenção de íntron), *mutually exclusive exons* (MEE - éxons mutuamente excludentes), *multiple exon skipping* (MES - salto de vários éxons), *alternative transcription start sites* (ATSS - sítio alternativo de início da transcrição) e *alternative transcription termination sites* (A TTS - sítio alternativo de terminação da transcrição), conforme representado na **Error! Reference source not found.**

Foram utilizados os parâmetros padrão da ferramenta (p -value = 0.05). As análises partiram de arquivos GTF (Gene Transfer Format) gerados com a plataforma *Cufflinks*. Por este pipeline, testes de enriquecimento são realizados via *prop.test* em ambiente R e as comparações de enriquecimentos são feitas com *fisher.test*. Os valores de p são corrigidos para testes múltiplos usando o esquema de Benjamin-Hochberg e um FDR < 0,05 é considerado significativo (VITTING-SEERUP; SANDELIN, [s.d.]).



Figura 7. Os padrões de splicing observados entre duas isoformas, de forma comparativa, estão indicados por cores, para cada padrão de processamento na coluna à esquerda. A denominação correspondente ao evento está indicada na coluna do meio e a abreviatura utilizada, na coluna à direita (inspirada nos resultados de (VITTING-SEERUP et al., 2014) obtida em (VITTING-SEERUP; SANDELIN, [s. d.]).

4. RESULTADOS

4.1. Avaliação da modulação da expressão gênica em peixes-boi em resposta ao evento de maré vermelha

Uma reanálise das sequências depositadas no banco de dados de acesso público (GEO #GSE86792) foi feita para se obter os arquivos de entrada necessários para os programas de análise da ocorrência de eventos de *splicing* alternativo. Conforme o corte estabelecido ($\log_2$ fold change ≥ 2 ou ≤ -2) obtivemos 1429 genes diferencialmente expressos, estando 185 induzidos e 1244 reprimidos em resposta ao evento de maré vermelha, comparado aos animais controle. Procedemos à identificação dos grupos enriquecidos por categorização funcional excluindo-se os genes associados à resposta imune – explorados previamente por Lazensky e colaboradores (2020). As principais categorias identificadas por quantidade de genes estão indicadas na **Error! Reference source not found.8**.

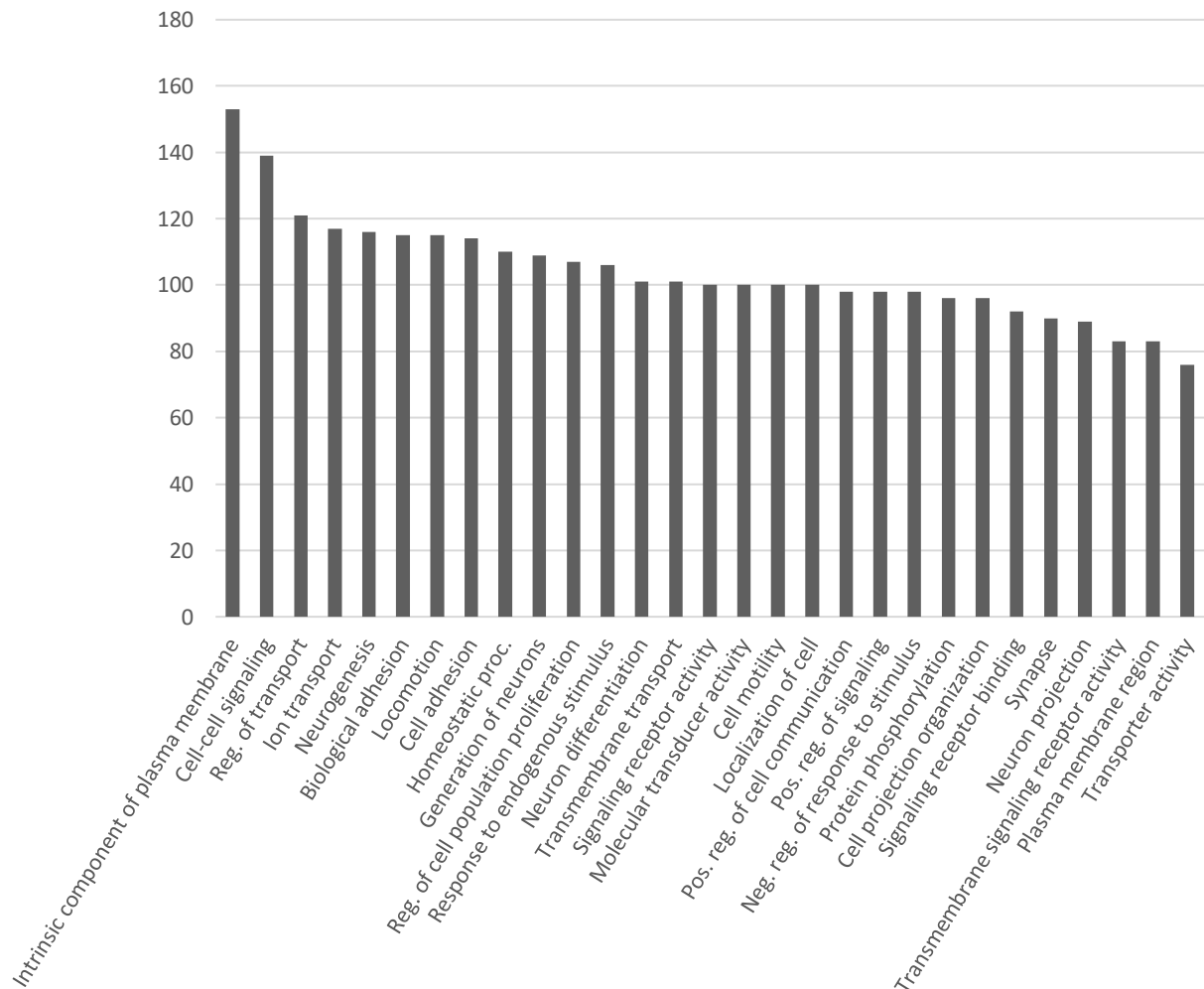


Figura 8. Categorização funcional dos genes diferencialmente expressos de acordo com o Gene Ontology (GO), excluindo as categorias nominalmente relacionadas ao sistema imune de peixes-boi.

Foram selecionados os 10 principais grupos segundo a categorização do GO para avaliação do perfil de modulação (quantificação dos genes induzidos e reprimidos) em resposta ao evento de maré vermelha, considerando os dados de expressão relativa dos genes (**Error! Reference source not found.**). Os principais grupos enriquecidos estão associados a transporte transmembrana, sinalização celular e ao processo de formação e diferenciação de novos neurônios.

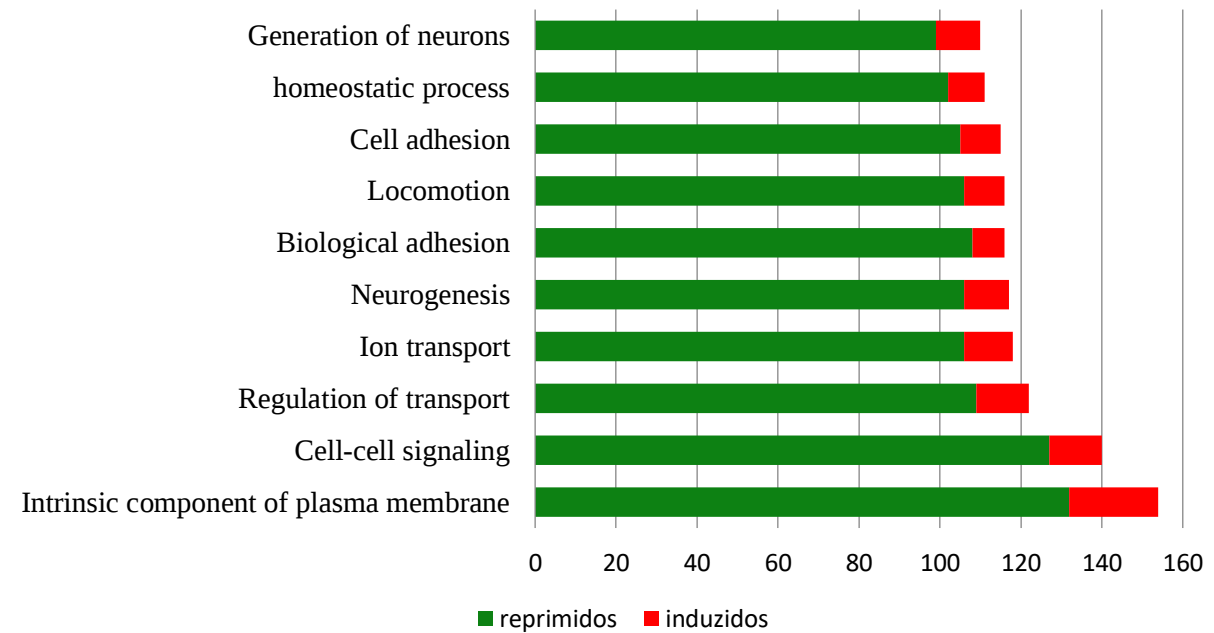


Figura 9. Perfil de modulação dos principais grupos enriquecidos na categorização funcional em resposta ao evento de maré vermelha.

Dentre os genes diferencialmente expressos, os 20 mais induzidos e reprimidos com *p*-valor ajustados estão listados nas Tabelas 3 e 4, respectivamente. O gene mais induzido apresenta modulação cerca de 44 vezes maior ($\log_2 > 5.48$) no grupo de indivíduos expostos à brevetoxina em comparação ao grupo controle (Tabela 3) e corresponde à subunidade β do complexo de proteína adaptadora (AP), que atua selecionando a carga para inclusão em vesículas revestidas de clatrina nas vias secretoras e endocíticas tardias. AP-1 é específico para o tráfego entre a face trans do complexo de Golgi e endossomos (PARK; GUO, 2014).

Tabela 3. Vinte genes mais induzidos em peixes-boi em resposta ao evento de maré vermelha.

GenBank ID	Gene Product	pval	Padj	log2 Fold Change
XM_004387605	Adaptor Related Protein Complex 1 Subunit Beta 1 (AP1B1)	0.051783161	0.308307212	5.485243503
XM_004380199	mitochondrial sodium/hydrogen exchanger 9B2-like	0.086747372	0.40470638	4.684458869
XM_004379435	nuclear transcription factor Y, alpha	0.211316321	0.595622758	4.612754353
XM_004372135	Zinc Finger Protein 618 (ZNF618)	0.277600084	0.670658995	4.344323337

XM_004377354	oxysterol binding protein-like 10 (OSBPL10)	0.303565198	0.692296379	4.279445851
XM_004378916	Neurofibromatosis type 2 (NF2)	0.337424603	0.723684447	4.178720059
XM_004387341	endoplasmic reticulum protein 27	0.151920802	0.523567862	4.178205228
XM_004383315	Albumin	0.434469518	0.794473319	4.131689995
XM_004376693	Ankyrin Repeat And SOCS Box Containing 2 (ASB2)	0.428480376	0.790337013	4.066447213
XM_004380925	iduronate 2-sulfatase	0.063654562	0.344049066	4.065490326
XM_004383019	U-Box Domain Containing 5 (UBOX5)	0.423597209	0.786986025	4.000956806
XM_004374246	Myosin Light Chain 4 (MYL4)	0.439998468	0.797635288	3.989850007
XM_004389843	Peroxisome Proliferator Activated Receptor Delta (PPARD)	0.057321054	0.326634685	3.825369686
XM_004387367	Protein Tyrosine Phosphatase 4A3 (PTP4A3)	0.279085723	0.670693053	3.668788383
XM_004389504	Proteasome 26S Subunit, Non-ATPase 1 (PSMD1)	0.330572042	0.717573914	3.664086036
XM_004390099	septin 5 (SEPT5)	0.141301394	0.506338964	3.622276261
XM_004376911	Lysine Demethylase 6A (KDM6A)	0.571654115	0.888053106	3.569438309
XM_004389447	Down syndrome critical region gene 3 (DSCR3)	0.144253604	0.511216222	3.567228152
XM_004378794	transmembrane protein 56-like	1.42E-07	3.55E-05	3.529461311
XM_004381059	Spermatogenesis Associated 24 (SPATA24)	0.523816695	0.859854357	3.517052552

Dentre os genes mais reprimidos (Tabela 4), os mais significativamente regulados negativamente apresenta repressão com uma alteração de $\log(2) \leq -9.33$ vezes (alteração >635 vezes), demonstrando uma regulação negativa da expressão gênica mais expressiva em relação à regulação positiva.

Tabela 4. Vinte genes mais reprimidos em peixes-boi em resposta ao evento de maré vermelha.

GenBank ID	Gene Product	pval	Padj	log2 Fold Change
XM_004377510	Interleukin 6 (IL6)	7.03E-09	2.59E-06	-9.311017857
XM_004388387	GRP1 (general receptor for phosphoinositides 1)-associated scaffold protein	9.92E-10	4.79E-07	-7.482086875
XM_004388863	Interleukin 1 alpha (IL1A)	1.84E-07	4.39E-05	-7.105210494
XM_004379783	Transient Receptor Potential Cation Channel Subfamily C Member 5 (TRPC5)	0.053886951	0.314787806	-6.568439466
XM_004385437	C-C Motif Chemokine Ligand 4 (CCL4)	1.11E-20	3.00E-17	-6.449817076
XM_004381068	Heparin Binding EGF Like Growth Factor (HBEGF)	3.09E-12	2.52E-09	-6.417384236
XM_004379317	Neurotrimin (NTM)	0.062883767	0.342517983	-6.119373098
XM_004389417	mucin 6, oligomeric mucus/gel-forming (MUC6)	9.99E-08	2.73E-05	-6.065153049
XM_004390150	Hydroxysteroid 17-Beta Dehydrogenase 11 (HSD17B11)	0.053253375	0.312938246	-5.989294709
XM_004380576	MAM and LDL-receptor class A domain-containing protein C10orf112-like	0.059807043	0.334966756	-5.973076767
XM_004381346	hydroxysteroid (17-beta) dehydrogenase 3	0.117888648	0.468447701	-5.948492477
XM_004390077	Tolloid Like 1 (TLL1)	0.05584191	0.321488144	-5.920563445
XM_004383716	FOS Like 1, AP-1 Transcription Factor Subunit (FOSL1)	0.076099047	0.375664682	-5.912286486

XM_004388025	G Protein-Coupled Receptor 61 (GPR61)	0.056776531	0.324289627	-5.894119551
XM_004375166	family with sequence similarity 184, member B (FAM184B)	0.058349261	0.329872459	-5.873953535
XM_004389648	Frizzled Class Receptor 1 (FZD1)	0.073885962	0.371304085	-5.840214134
XM_004379795	Solute Carrier Family 6 Member 14 (SLC6A14)	0.205838327	0.589362136	-7.79379269
XM_004391243	protocadherin alpha 5 (PCDHA5)	0.05333722	0.312953067	-6.74811876
XM_004372319	Energy Homeostasis Associated (ENHO)	0.165539745	0.547147522	-5.96694175
XM_004390914	olfactory receptor 6C75-like	0.063585412	0.344049066	-5.87534205

Foi feita uma análise de enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados *Kegg ontology*, apresentando de forma agrupada as principais funções biológicas realizadas pelos genes diferencialmente expressos (**Error! Reference source not found.**).

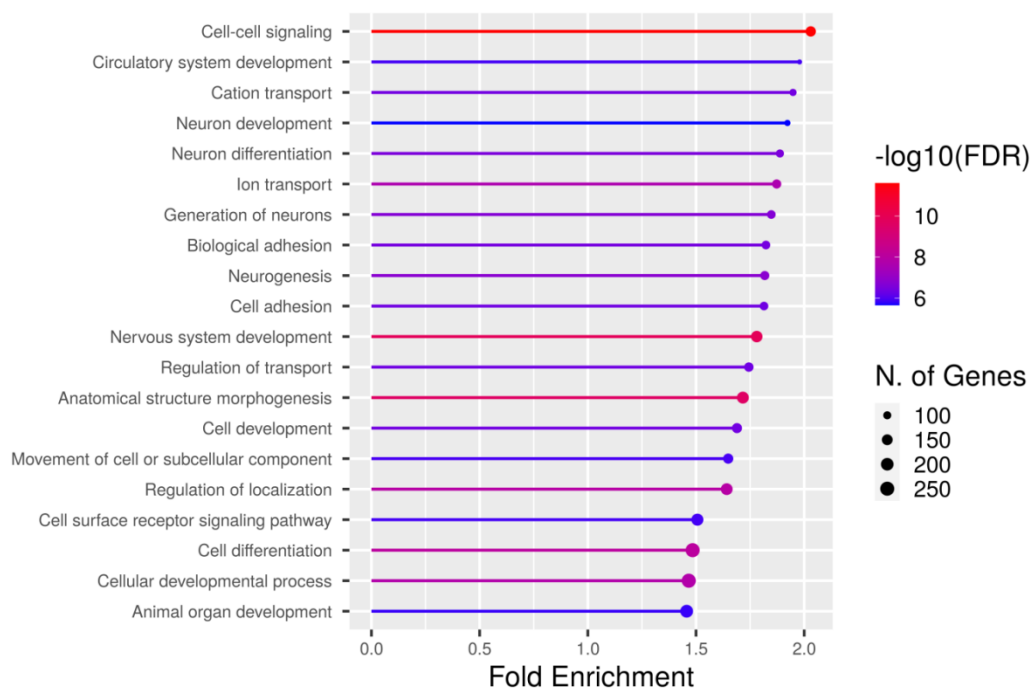


Figura 10. Inferência de vias metabólicas moduladas em resposta ao evento de maré vermelha em peixes-boi.

Dentre as principais vias moduladas, destacamos duas de grande relevância no contexto de sinalização molecular: o sistema de sinalização das proteínas quinase ativada por mitógeno (MAP) quinases (MAPK, *mitogen-activated protein kinase*) (**Error! Reference source not found.**) e a via de sinalização Wnt (NISHINA, 2004; ZHANG; LIU, 2002)(**Error! Reference source not found.**). Dentro da via das MAP quinases tanto a via clássica quanto a via das proteínas quinases ativadas por estresse (SAPK/Jun N-terminais quinases - JNK) estão moduladas, possivelmente afetando processos moleculares como diferenciação e proliferação celular, inflamação e apoptose. As duas vias estão

interconectadas e a via das MAP quinases modula a via de sinalização Wnt: a via Wnt/Ca²⁺ estimula o aumento intracelular de Ca²⁺ através da ativação de proteínas quinase. Na via de sinalização Wnt há modulação diferencial das vias Wnt/β-catenina ou via Wnt “canônica” e das vias Wnt “não-canônicas” da polaridade celular planar (PCP) e Wnt-cálcio (Wnt/Ca²⁺).

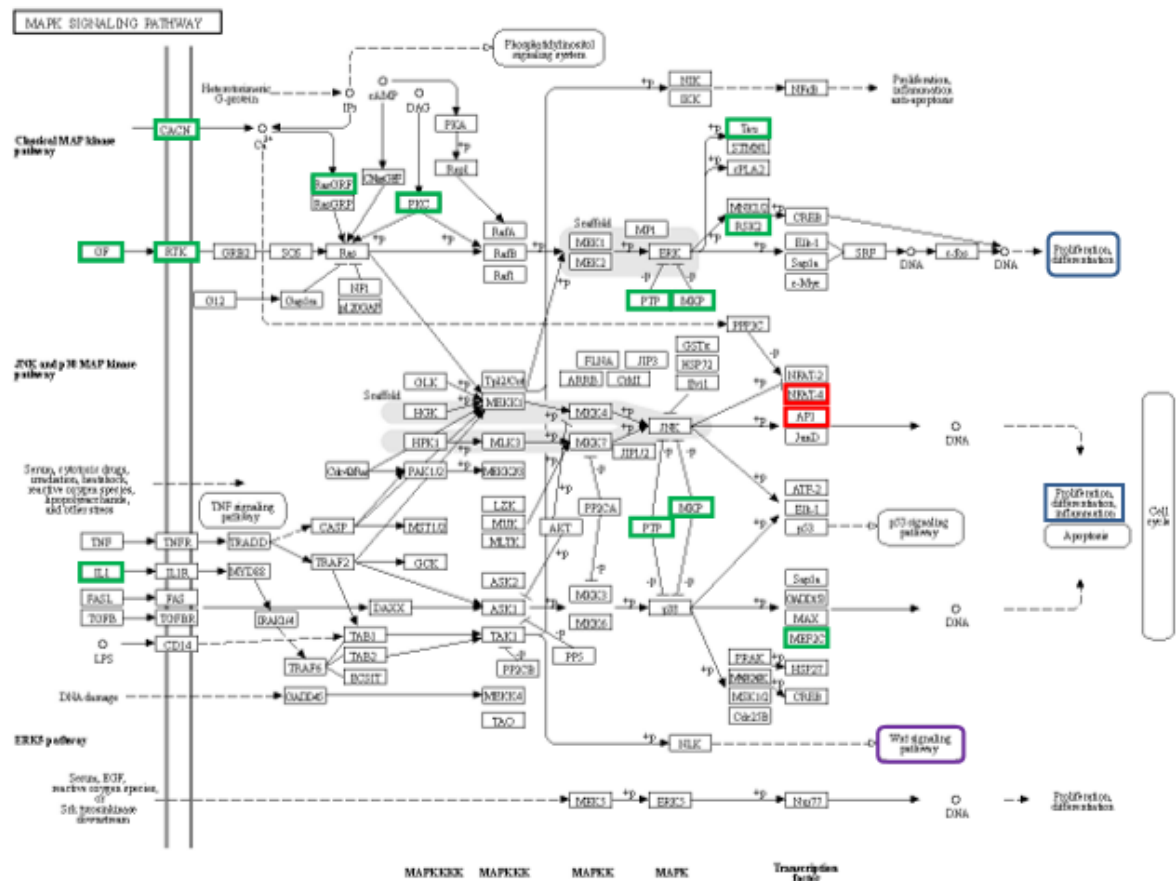


Figura 11. Modulação da via de sinalização das proteínas quinase ativada por mitógeno (MAP) quinases em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho, os genes induzidos nas condições experimentais. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados *Kegg ontology*.

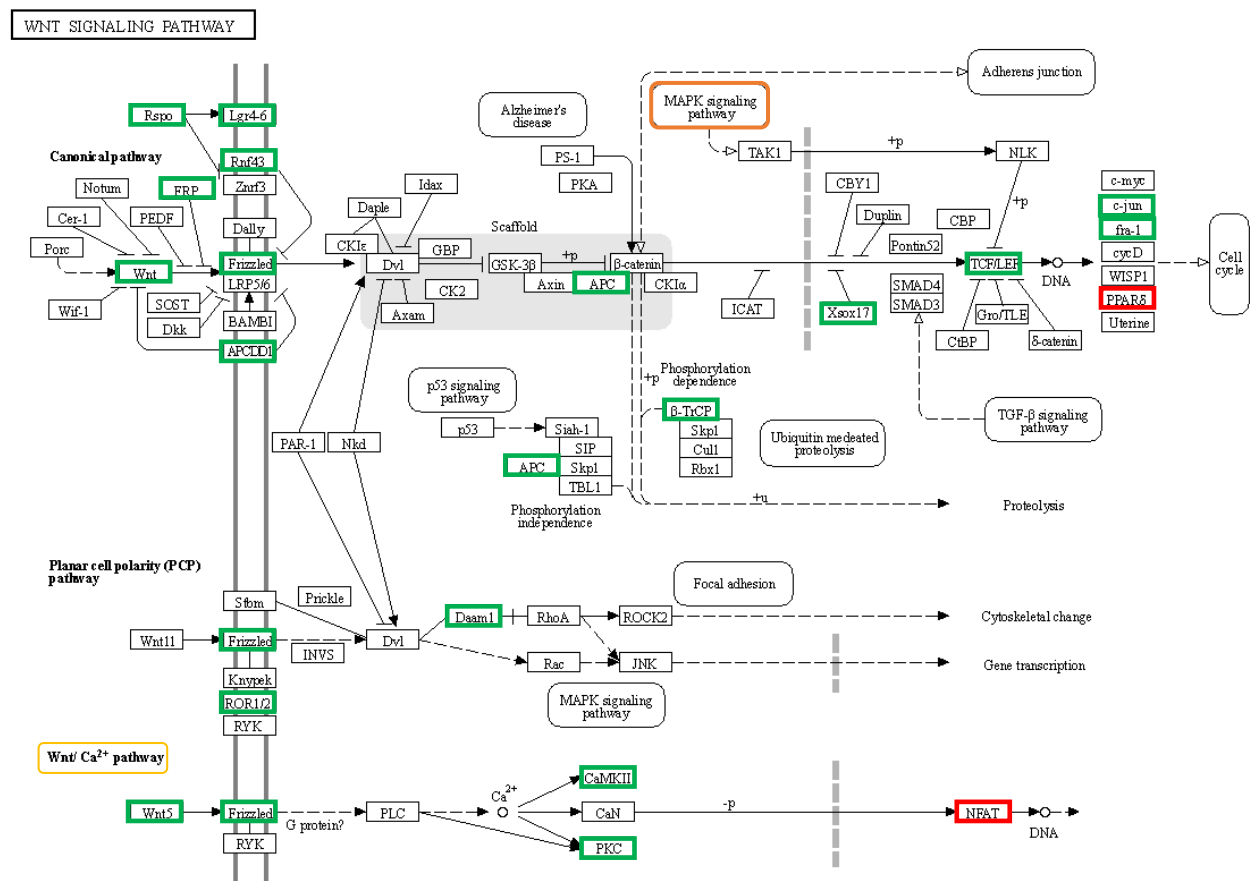


Figura 12. Modulação da via de sinalização Wnt canônica e não canônica em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho, os genes induzidos nas condições experimentais. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados *Kegg ontology*.

4.2. Avaliação da ocorrência de eventos de *splicing* alternativo em peixes-boi da subespécie *Trichechus manatus latirostris*

A primeira avaliação da ocorrência de eventos de processamento alternativo de genes nos dados de peixes-boi foi feita utilizando o pacote *Cufflinks*. Este pacote permite tanto a avaliação de processamento diferencial como a verificação de eventos de *splicing*. O *Cufflinks* precisa dos transcritos alinhados, com a identificação dos sítios de *splice* e cobertura de éxons, para construir as regiões contendo transcritos. A identificação de junções de *splice* e uso diferencial de isoformas foi executada partindo do arquivo de expressão diferencial, tendo sido aplicadas as ferramentas *cuffmerge*, para criar um único arquivo transcritos .gtf e, em seguida, o *cuffdiff*, que efetivamente compara numericamente a ocorrência de transcritos de mRNA em experimentos de RNA-seq e estabelece quais genes estão sendo processados diferencialmente ou estão passando por outros tipos de regulação em nível de isoforma. O

Cuffdiff assume que o número de leituras produzidas individualmente, por transcrito, é proporcional à sua abundância, apresentando distribuição normal.

O pacote *Cufflinks* não identifica, no entanto, o tipo de processamento ocorrido em cada gene identificado. Para tanto, executamos a ferramenta *IsoformSwitchAnalyzeR*, que agrupou os eventos de processamento alternativo por tipos específicos de *splicing*. Por meio do uso desta ferramenta é possível estabelecer se os eventos de processamento pertencem a um dos oito tipos indicados na **Error! Reference source not found.**:

4.2.1. Padrões de *splicing* alternativo obtidos utilizando o pipeline *Cufflinks*

Foram identificadas 852 regiões do genoma sendo processadas de forma alternativa. Destas, 534 estão anotadas segundo as análises realizadas por BLAST; 147 foram caracterizadas como proteínas hipotéticas. 47 regiões foram identificadas como não caracterizadas e outras 42 regiões estão não nomeadas. Estas observações são comparativas a proteínas do banco de dados utilizado como referência. Apenas 82 dentre as regiões com possível processamento alternativo nas condições experimentais não corresponderam a alvos no genoma de referência. A ferramenta BLAST (*Basic Local Alignment Search Tool*) realiza comparações entre sequências biológicas primárias e um banco de dados, retornando as sequências mais similares e a respectiva significância estatística.

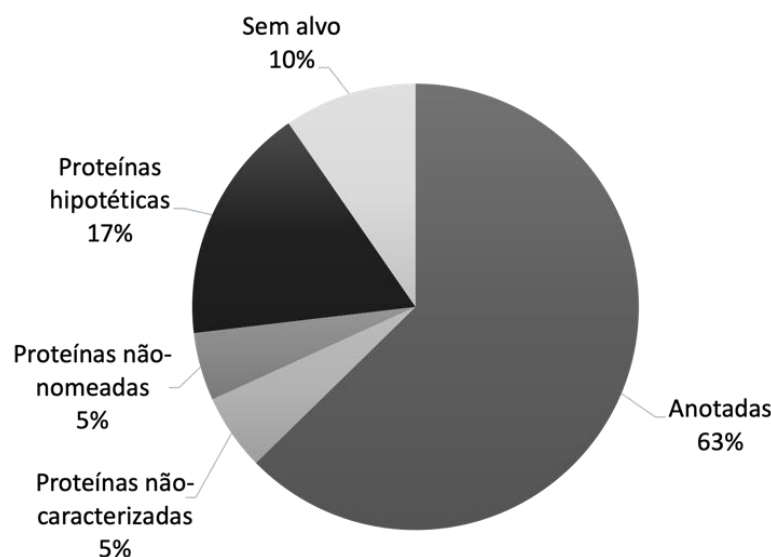


Figura 13. Percentual relativo às regiões anotadas a partir dos resultados de BLAST das regiões genômicas.

Dentre as principais atividades celulares relacionadas estão atividades de transferase, hidrolase e ligação de pequenas moléculas (**Error! Reference source not found.**). Transferase são enzimas da classe EC (*Enzyme Commission*) 2 que catalisam a transferência

de um grupos químicos entre compostos, como por exemplo a transferência de grupos metil, glicosil, acil entre um composto doador e um aceptor. Discutimos duas transferases potencialmente sofrendo splicing alternativo: uma glicosiltransferase (EXT1) e uma acetilgalactosaminiltransferase (GALNTL5).

Hidrolases são enzimas classe EC 3 que utilizam água para quebrar ligações químicas, o que normalmente resulta na divisão de uma molécula maior em moléculas menores. Alguns exemplos comuns de enzimas hidrolases são esterases incluindo lipases, fosfatases, glicosidases, peptidases e nucleosidases. Aqui discutimos o potencial processamento alternativo de três hidrolases: uma endopeptidase (TFPI), uma GTPase (GIT2) e uma metaloprotease (ADAM17). Estes resultados têm por base a anotação do genoma de *Trichechus manatus latirostris*.

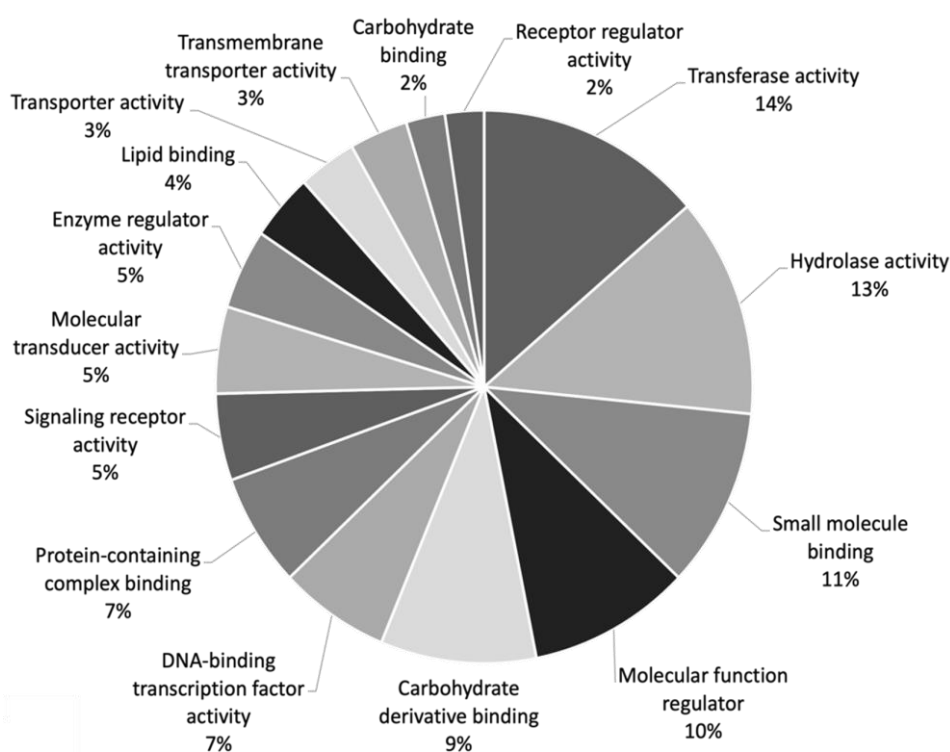


Figura 14. Categorização das principais funções atribuídas aos genes anotados por meio de alinhamento utilizando a ferramenta BLASTx

Utilizando o genoma humano, foi feita a categorização funcional dos genes possivelmente sofrendo processamento alternativo. As análises do GO revelaram que as principais categorias enriquecidas são resposta a estresse (processo biológico, GO:0006950); regulação da resposta a estímulos (processo biológico, GO:0048583); membrana de organelas

(componente celular, GO:0031090) e processos relacionados ao sistema imune (processo biológico, GO:0002376).

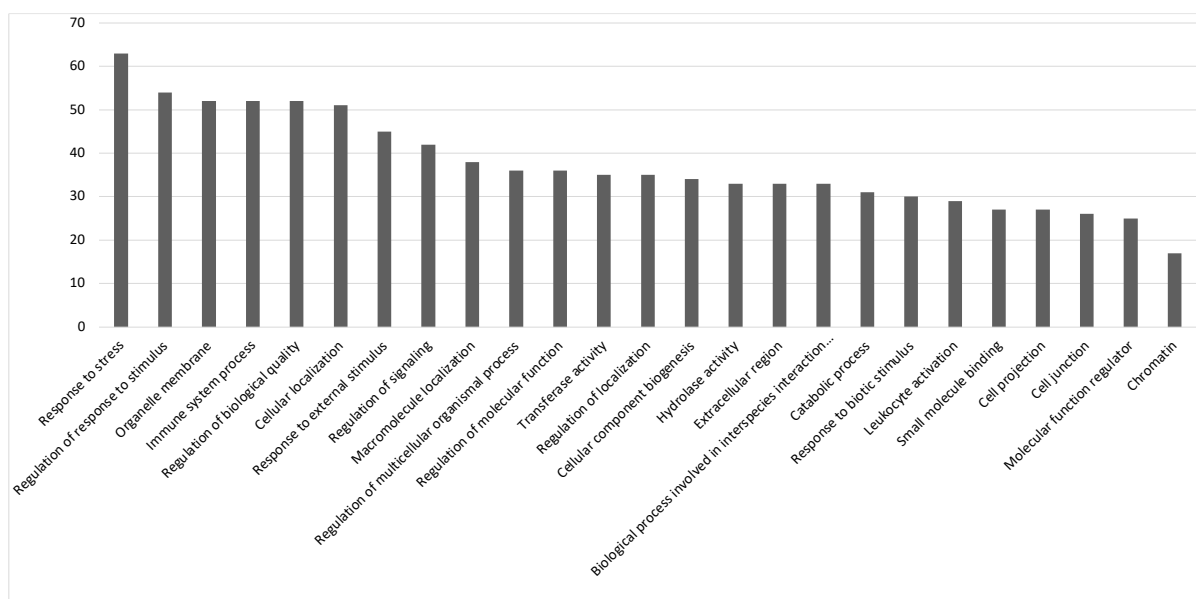


Figura 15. Categorização funcional dos genes possivelmente sofrendo processamento alternativo em peixes-boi.

Avaliamos o número de genes que potencialmente sofrem processamento alternativo e que estão diferencialmente expressos e identificamos apenas 4 genes: o gene CXCR2, que codifica para um receptor 2 de quimiocina (motivo C-X-C) e o gene HRH4, que codifica para um receptor de histamina (H4) apresentam sua isoforma constitutiva induzida. O gene NFKB1, que codifica para o fator nuclear kappa β e o gene TRIB1 (*Tribbles 1*), que codifica para um membro da família de pseudoquinases apresentam a isoforma constitutiva reprimida em resposta ao evento de maré vermelha.

4.2.1.1. Processamento alternativo de genes relacionados ao sistema imune de peixes-boi da subespécie *Trichechus manatus latirostris*

A categorização funcional identificou 53 genes associados ao sistema imune dos peixes-boi apresentando processamento alternativo. Os genes são JARID2 (*jumonji and AT-rich interaction domain containing 2*), POU2F2 (*POU class 2 homeobox 2*), APBB1IP (*amyloid beta precursor protein binding family B member 1 interacting protein*), B4GALT1 (*beta-1,4-galactosyltransferase 1*), TESC (*tescalcin*), FLT3LG (*fms related receptor tyrosine kinase 3 ligand*), MKNK2 (*MAPK interacting serine/threonine kinase 2*), SIRPB1 (*signal regulatory protein beta 1*), TLR8 (*toll like receptor 8*), CD33 (*CD33 molecule*), VSIR (*V-set immunoregulatory receptor*), PPP3CB (*protein phosphatase 3 catalytic subunit beta*), NFKB1 (*nuclear factor kappa B subunit 1*), RIPOR2 (*RHO family interacting cell polarization*

regulator 2), MAPK14 (mitogen-activated protein kinase 14), HAVCR1 (hepatitis A virus cellular receptor 1), ID2 (inhibitor of DNA binding 2), CD2 (CD2 molecule), EXOSC9 (exosome component 9), CXCL6 (C-X-C motif chemokine ligand 6), CD68 (CD68 molecule), NFATC1 (nuclear factor of activated T cells 1), GCH1 (GTP cyclohydrolase 1), RAP1GAP2 (RAP1 GTPase activating protein 2), CASP9 (caspase 9), LMO2 (LIM domain only 2), TLR2 (toll like receptor 2), HERC5 (HECT and RLD domain containing E3 ubiquitin protein ligase 5), RAB20 (RAB20, member RAS oncogene family), S100A8 (S100 calcium binding protein A8), PTPRJ (protein tyrosine phosphatase receptor type J), DSN1 (DSN1 component of MIS12 kinetochore complex), ADAM17 (ADAM metalloproteinase domain 17), BCL2L11 (BCL2 like 11), PIK3AP1 (phosphoinositide-3-kinase adaptor protein 1), AZGP1 (alpha-2-glycoprotein 1, zinc-binding), S100A9 (S100 calcium binding protein A9), PTX3 (pentraxin 3), PF4 (platelet factor 4), SLC9B2 (solute carrier family 9 member B2), RAB31 (RAB31, member RAS oncogene family), STX8 (syntaxin 8), TRIB1 (tribbles pseudokinase 1), PDE12 (phosphodiesterase 12), CXCR2 (C-X-C motif chemokine receptor 2), SETD2 (SET domain containing 2, histone lysine methyltransferase), EXT1 (exostosin glycosyltransferase 1), LRRK2 (leucine rich repeat kinase 2), ATG7 (autophagy related 7), CLEC6A (C-type lectin domain containing 6A), NFAM1 (NFAT activating protein with ITAM motif 1), PTPRC (protein tyrosine phosphatase receptor type C) e CCL5 (C-C motif chemokine ligand 5).

As análises do ontologia gênica utilizando exclusivamente os genes associados à resposta imune revelaram que a maior parte dos genes associados a processos biológicos estão associados à sinalização autócrina (**Error! Reference source not found.**). As categorias de componente celular (**Error! Reference source not found.**) e função molecular (**Error! Reference source not found.**) foram pouco enriquecidas.

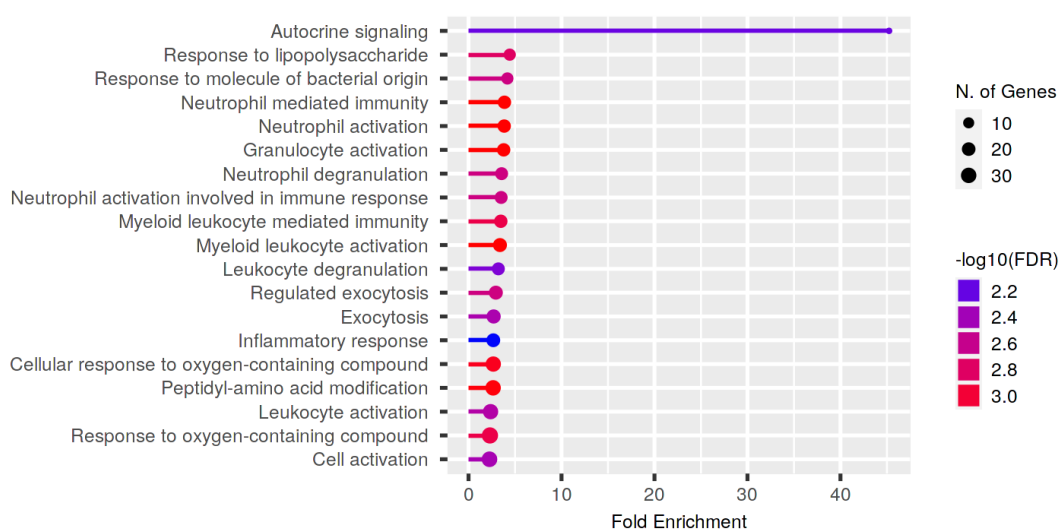


Figura 16. Principais categorias enriquecidas em processos biológicos (BP).

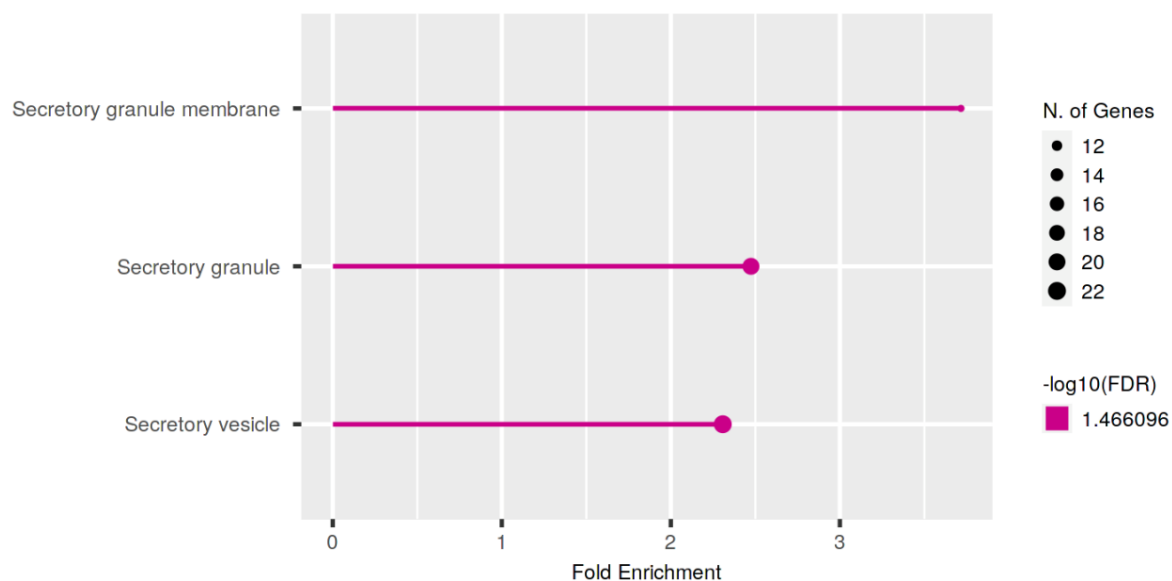


Figura 17. Principais categorias enriquecidas em componente celular (CC).

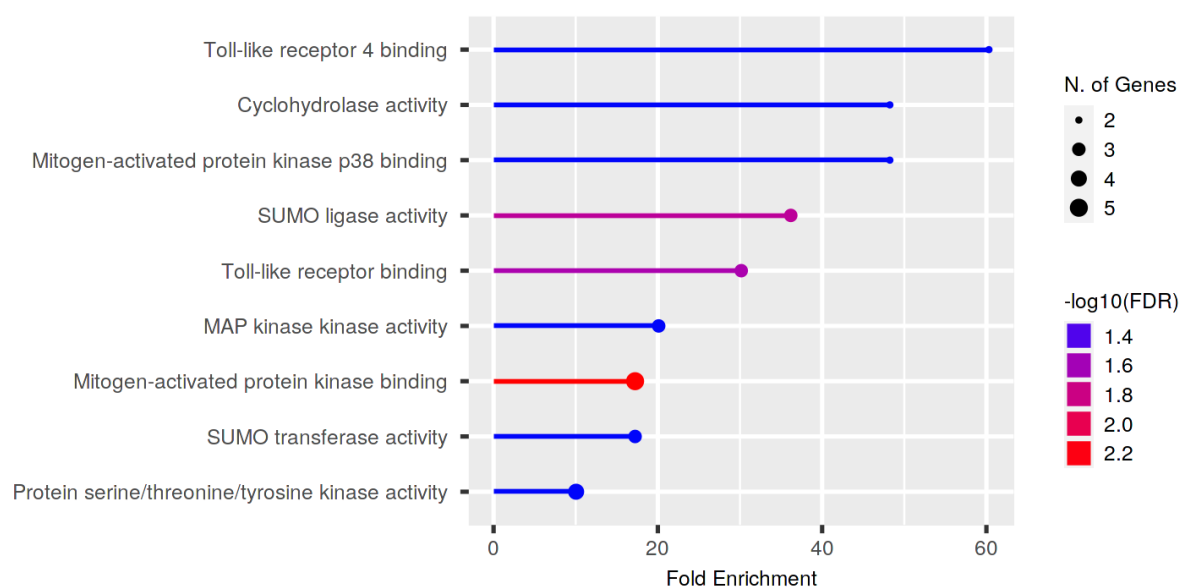


Figura 18. Principais categorias enriquecidas em função molecular (MF).

Foi feita uma análise de enriquecimento de vias metabólicas com os genes possivelmente processados de forma alternativa utilizando o banco de dados *Kegg ontology*. De forma geral se destacam a atividade de transporte por vesículas, biossíntese de glicanos, além do acumulo de transcritos rocessados de forma alternativa de geens relacionados a vias clássicas do sistema imune como receptores toll-like, interleucinas e fatores de necrose tumoral (**Error! Reference source not found.**).

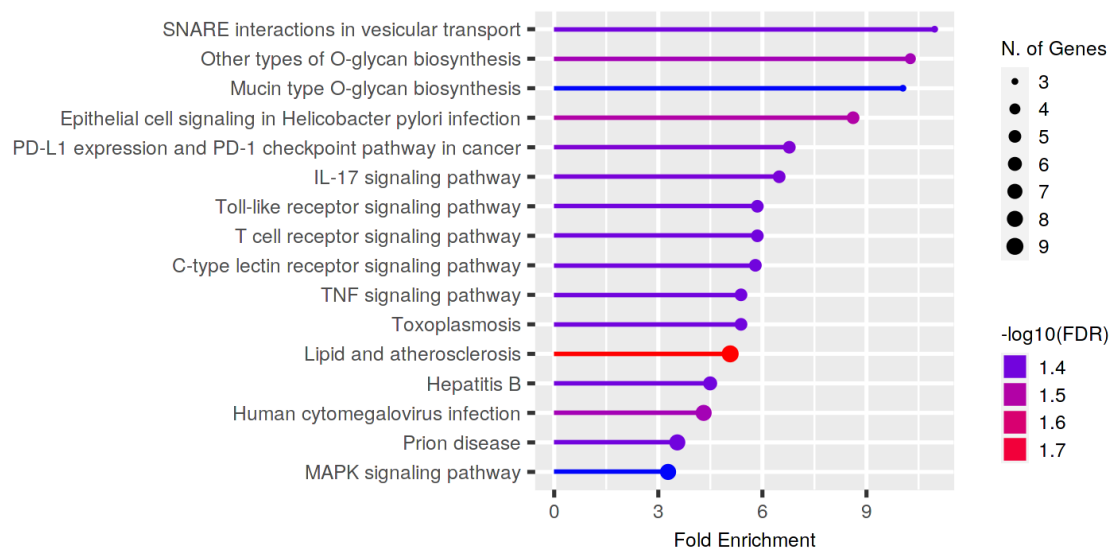


Figura 19. Genes processados de forma alternativa avaliados utilizando a ferramenta Kegg ontology.

Avaliando as principais vias moduladas, destacamos que houve acúmulo de transcritos das isoformas alternativamente processadas para três importantes vias relacionadas ao sistema imune: a via de sinalização TNF (**Error! Reference source not found.**), IL-17 (**Error! Reference source not found.**) e a via de receptores Toll-like (**Error! Reference source not found.**).

TNF SIGNALING PATHWAY

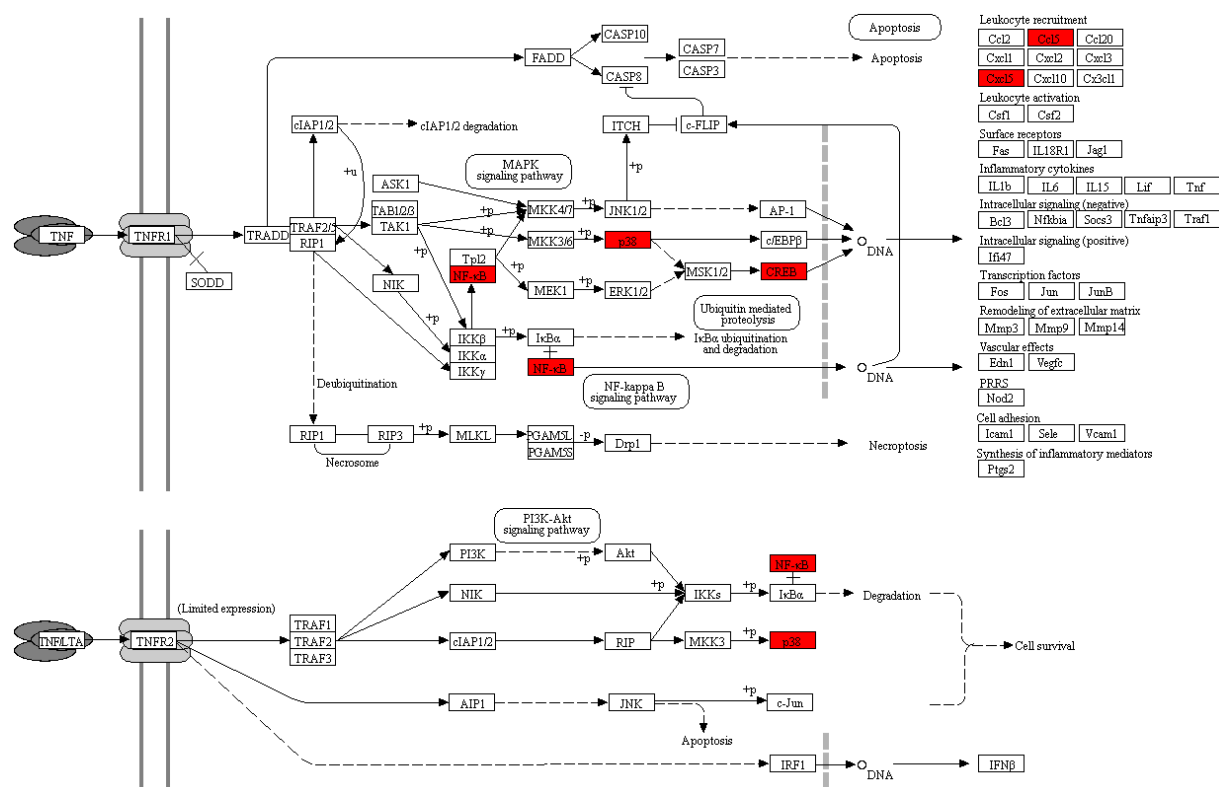


Figura 20. Modulação da via de sinalização TNF em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho estão indicados os genes possivelmente sofrendo splicing alternativo. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados *Kegg ontology*.

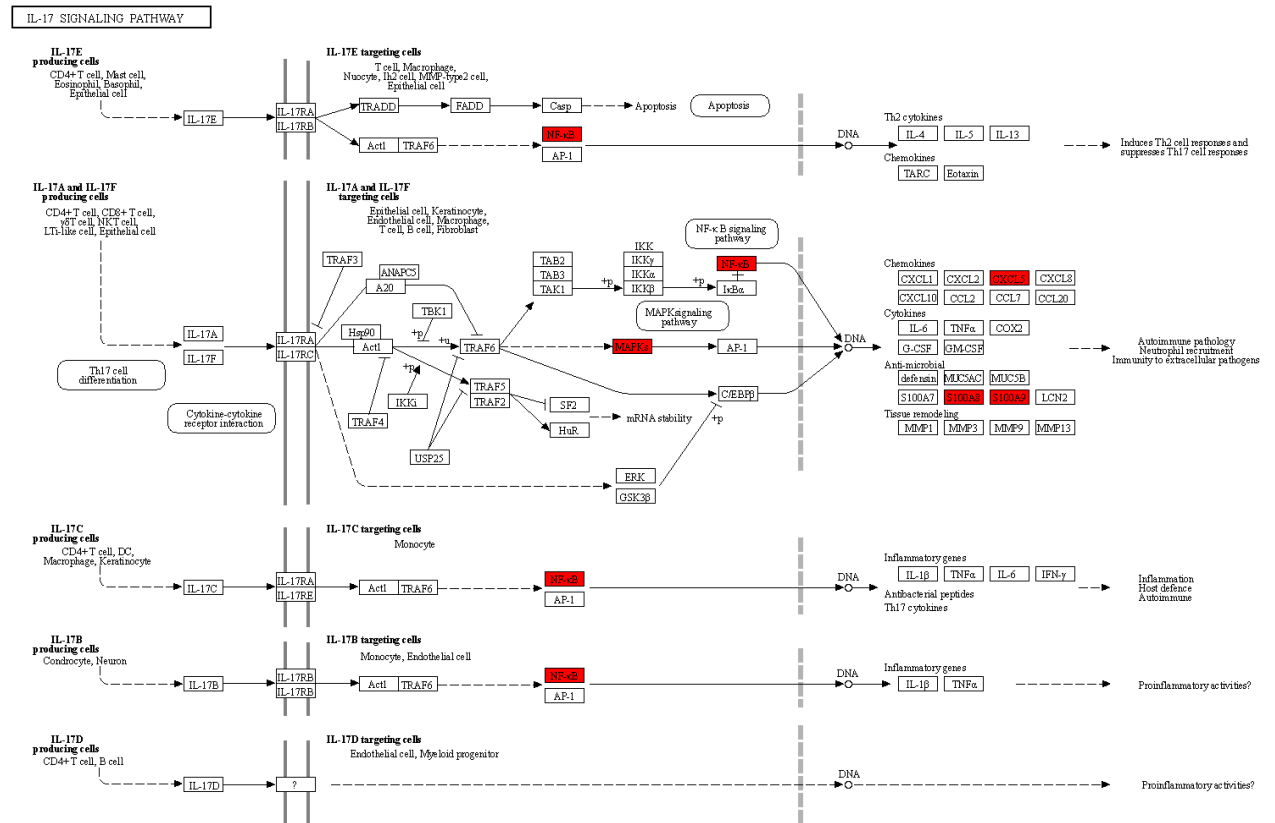


Figura 21. Modulação da via de sinalização IL-17 em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho estão indicados os genes possivelmente sofrendo splicing alternativo. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados *Kegg ontology*

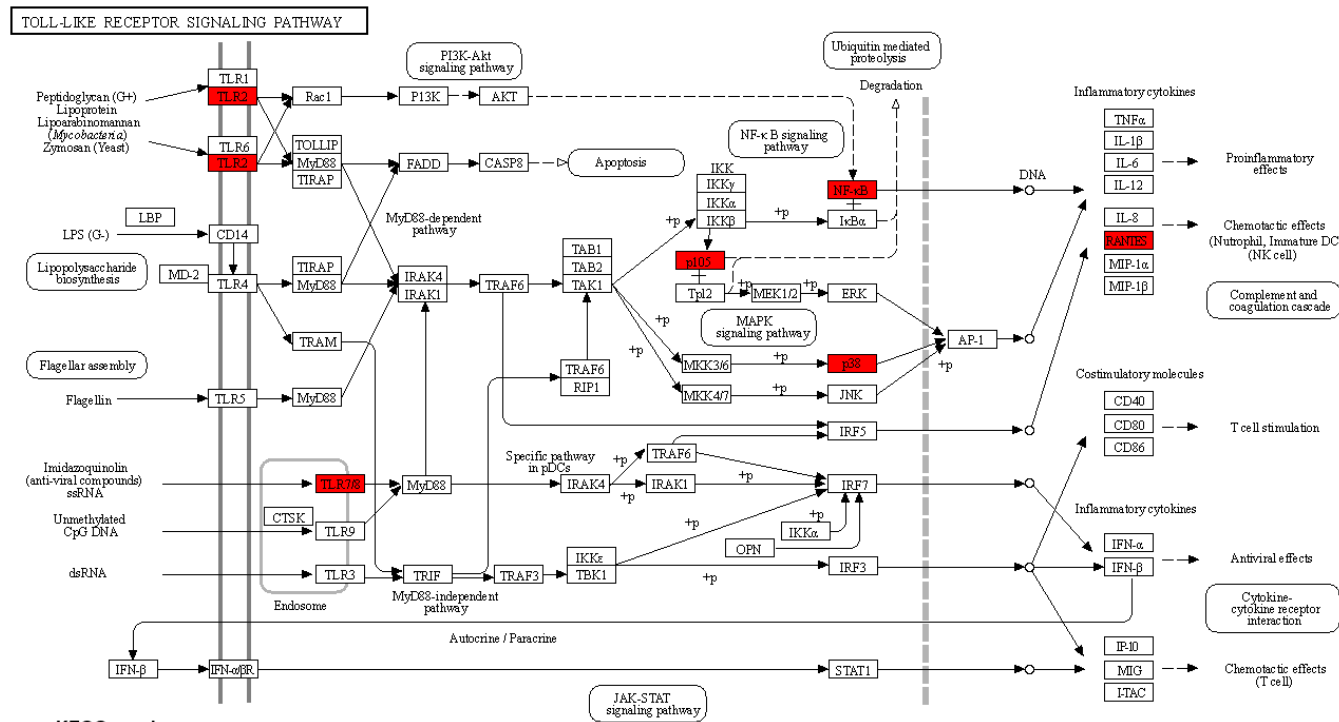


Figura 22. Modulação da via de sinalização de receptores Toll-like em peixes-boi em resposta ao evento de maré vermelha. Em verde estão indicados os genes reprimidos nas condições experimentais. Em vermelho estão indicados os genes possivelmente sofrendo splicing alternativo. A análise resulta do enriquecimento de vias metabólicas com os genes diferencialmente expressos utilizando o banco de dados *Kegg ontology*.

4.2.2. Padrões de splicing alternativo obtidos utilizando o IsoformSwitchAnalyzeR

Os arquivos de saída do *Cuffdiff* foram avaliados utilizando a ferramenta *IsoformSwitchAnalyzeR*, que permite a categorização dos eventos de splicing por tipo de processamento (**Error! Reference source not found.**). O programa avaliou oito tipos de processamento alternativo não tendo identificado variantes do tipo MEE (éxon mutuamente exclusivo) ou MES (salto múltiplos de éxons). O evento de maior ocorrência foi ATSS (sítios alternativos de início de transcrição) e o tipo com menos isoformas identificadas os eventos de éxon *skipping* (ES).

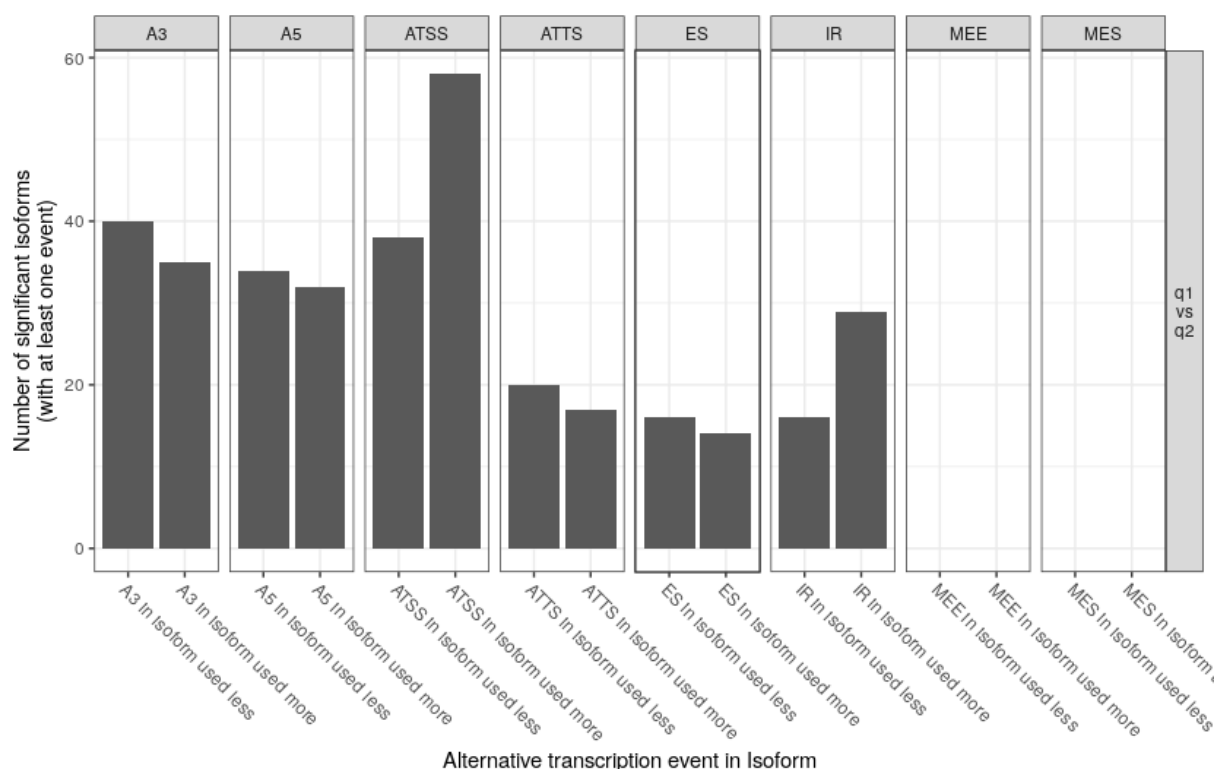


Figura 23. Proporção de eventos de *splicing* alternativo computada pela ferramenta IsoformSwitchAnalyzeR. A ferramenta identificou 97 genes que potencialmente sofrem processamento alternativo. Foi feita a anotação destes genes identificados por meio da ferramenta BLAST. Do total de genes identificados, 33% não identificamos correspondência de anotação. Dentre os genes anotados identificamos atividades relacionadas a oxidação-redução, transporte relacionado a vesículas, transporte de íons, apoptose, regulação de processos metabólicos. Listamos alguns dos genes que potencialmente sofrem processamento alternativo indicando a quais eventos de *splicing* estes estão associados (**Tabela 5**).

Tabela 5. Lista de genes identificados pela ferramenta IsoformSwitchAnalyzeR e que potencialmente sofrem *splicing* alternativo em peixes-boi.

Locus	Gene description	Alternative splicing event					
		A3	A5	ATSS	ATTS	ES	IR
AHIN01002447.1:0-11523	superoxide dismutase 2		X			X	X
AHIN01003784.1:10690-23854	serine/threonine-protein kinase Sgk1			X	X	X	X
AHIN01080675.1:38-49743	Nedd4 family interacting protein 1	X		X	X	X	X
AHIN01045573.1:3-29359	protein tyrosine phosphatase receptor type C	X	X	X		X	X
AHIN01029676.1:24461-	peroxiredoxin 6	X	X	X			X

35997							
AHIN01026963.1:1233-64665	ubiquitin protein ligase E3A	X	X	X			
AHIN01024640.1:4-4467	crystallin beta-gamma domain containing 1	X	X	X			
AHIN01024024.1:19649-32552	DnaJ heat shock protein family (Hsp40) member A1	X	X	X	X		X
AHIN01022638.1:24-58229	microtubule actin crosslinking factor 1	X	X	X	X	X	X
AHIN01021425.1:913-9910	CYLD lysine 63 deubiquitinase	X		X	X	X	
AHIN01017379.1:28699-57442	pantothenate kinase 3	X	X	X	X		X
AHIN01002446.1:5652-16893	superoxide dismutase [Mn], mitochondrial-like (predicted)			X			X
AHIN01003274.1:1025-80084	syntaxin binding protein 5	X	X	X	X		
AHIN01004971.1:2142-25264	zinc finger RANBP2-type containing 2	X	X	X	X		
AHIN01006201.1:6080-16016	C1D nuclear receptor corepressor	X		X			
AHIN01006591.1:2-12640	sulfotransferase family 1C member 4	X	X	X	X	X	X
AHIN01007237.1:13144-28838	thioredoxin domain containing 9 (PREDICTED)		X	X	X		
AHIN01008105.1:42-19101	calmodulin 2 (predicted)	X		X			X
AHIN01011313.1:23321-76669	Fas cell surface death receptor	X	X	X	X	X	X
AHIN01011597.1:24559-56891	transmembrane 9 superfamily member 3 (predicted)	X	X	X			X

Nossos resultados evidenciam que para um mesmo gene, como regra, mais de um evento de processamento alternativo está associado. Dois dentre os genes processados de forma alternativa estão representados de maneira esquemática quanto aos eventos de *exon skipping* e retenção de íntron (**Error! Reference source not found.**).

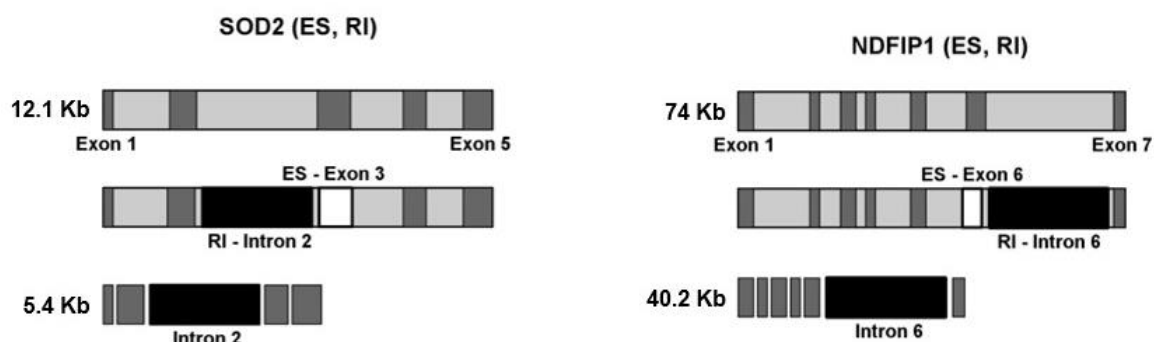


Figura 24. Representação esquemática dos eventos de processamento alternativo (éxon skipping e retenção de íntron) em dois genes, superoxide dismutase 2 (SOD2) e Nedd4 family interacting protein 1 (NDFIP1). As representações foram feitas com base nos resultados de processamento alternativo obtidos *in silico*.

4.2.2.1. Processamento alternativo de genes relacionados ao sistema imune de peixes-boi da subespécie *Trichechus manatus latirostris*

Procedemos à busca por genes associados à resposta imune dentre aqueles que potencialmente apresentam processamento alternativo em nossas condições experimentais. A **Tabela 6** lista os genes e o tipo de processamento identificado.

Tabela 6. Lista de genes identificados pela ferramenta IsoformSwitchAnalyzeR associados ao sistema imune de peixes-boi e que potencialmente sofrem *splicing* alternativo.

Locus	Gene description	Alternative splicing event					
		A3	A5	ATSS	ATTS	ES	IR
AHIN01045573.1:3-29359	protein tyrosine phosphatase receptor type C	X	X	X		X	X
AHIN01154856.1:639-1797	MHC class II antigen (HLA-DQA2) gene (predito)			X	X		X
AHIN01052580.1:9184-41292	E2F associated phosphoprotein	X	X	X	X		
AHIN01050480.1:27-25810	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein epsilon		X		X		
AHIN01047395.1:9120-40188	pyruvate dehydrogenase kinase 1	X		X			
AHIN01046490.1:234-89458	lymphocyte antigen 75	X	X	X	X		X
AHIN01080675.1:38-49743	Nedd4 family interacting protein 1	X		X	X	X	X
AHIN01041547.1:3088-27778	protein tyrosine phosphatase receptor type E	X	X	X	X		

AHIN01029228.1:8-31737	lysosomal trafficking regulator	X	X	X	X		
AHIN01009197.1:72-29883	switching B cell complex subunit SWAP70		X	X			X
AHIN01153712.1:20902-44604	Y-box binding protein 3	X	X	X	X		
AHIN01151138.1:616-46354	zinc finger MYM-type containing 5	X	X	X	X	X	X
AHIN01087228.1:88-25646	kelch like family member 24	X	X	X	X	X	
AHIN01083197.1:3381-38355	STEAP4 metalloredutase	X	X	X	X		
AHIN01081908.1:422-16820	protein tyrosine phosphatase 4A1			X		X	X
AHIN01078162.1:98829-116633	thrombospondin 1 (predito)	X	X	X	X		X
AHIN01075891.1:69108-96696	family with sequence similarity 46 member C (predicted)	X	X	X		X	
AHIN01072367.1:40-29660	dedicator of cytokinesis 11	X		X	X	X	
AHIN01062283.1:26258-45535	NFKB inhibitor zeta	X		X			X
AHIN01062014.1:25187-30617	mbt domain containing 1 (predito)	X		X		X	X
AHIN01058886.1:3172-20721	nuclear factor, erythroid 2 like 3			X			X
AHIN01056051.1:16-17050	TNF alpha induced protein 8	X					X
AHIN01053422.1:21003-37687	integral membrane protein 2B	X	X	X		X	X

Vinte e três dentre os 97 genes identificados pela ferramenta IsoformSwitchAnalyzeR estão relacionados ao sistema imune dos peixes-boi. Somente para o fator de transcrição do tipo zinc finger contendo domínio MYM tipo 5 (RAKHRA; RAKHRA, 2021), que regula atividades como degradação de proteínas mediada por ubiquitina, transdução de sinal e migração celular, apresenta acúmulo de transcritos referentes à ocorrência de todos os seis eventos de *splicing* alternativo contabilizados.

Muitos genes que sofrem processamento alternativo atuam na regulação do sistema imune incluindo fatores de transcrição e reguladores do tráfego lisossomal. Dentre os genes

associados ao sistema imune evidenciamos de forma esquemática os possíveis eventos de *splicing* preditos para um receptor da proteína tirosina fosfatase tipo C (PTPRC) e um MHC de classe II (HLA-DQA2) (**Error! Reference source not found.**).



Figura 25. Representação esquemática de eventos de processamento alternativo (retenção de íntron) em dois genes, tirosina fosfatase tipo C (PTPRC) e MHC de classe II (HLA-DQA2). As representações foram feitas com base nos resultados de processamento alternativo obtidos *in silico*.

5. DISCUSSÃO

5.1. O evento da maré vermelha reprime diferencialmente grande quantidade de genes em peixes-boi

A reanálises de genes diferencialmente expressos revelou, como padrão, maior abundância de genes reprimidos em resposta ao evento de maré vermelha comparado ao número de genes induzidos. Como previamente publicado, a maioria deles tem função associada à resposta imune dos peixes-boi (LAZENSKY et al., 2020a). Expandindo as análises sobre os genes diferencialmente expresso observamos aqui que o estresse sofrido pelos peixes-boi alterou a modulação de genes relacionados a transporte transmembrana, incluindo genes de transporte de íons e transporte vesicular, bem como genes associados a sinalização molecular. Este padrão de expressão revela uma resposta molecular às toxinas do evento da maré vermelha: enquanto o sistema de transporte transmembrana tenta, como supomos, evitar o transporte de toxinas para o interior celular pela repressão de transportadores, a repressão de cascatas de ativação modula negativamente eventos cruciais como diferenciação e proliferação celular numa tentativa de, possivelmente, evitar o gasto energético. Há também evidências da repressão da neurogênese, evento comumente observado em células de mamíferos sob estresse. Os genes modulados positivamente estão relacionados, principalmente, ao sistema imune e recepção de estímulos.

Os eventos de maré vermelha são florações do dinoflagelado *Karenia brevis*, organismo produtor uma potente classe de toxinas, denominadas brevetoxinas, que são liberadas após a lise celular em águas oceânicas ou estuarinas, ou por aerossolização, na atmosfera (HOAGLAND et al., 2020). Ainda que vários peixes-boi sobrevivam à exposição subletal destas toxinas, muitos não resistem: alguns episódios de mortalidade de mamíferos marinhos em grande escala em decorrência dos efeitos de brevetoxinas foram registrados, como a morte de pelo menos 149 peixes-boi em 1996 na costa da Flórida (BOSSART et al., 1998), e um episódio envolvendo 34 peixes-boi em 2002 (FLEWELLING et al., 2005). No primeiro evento citado, a presença de lesões pulmonares indicou intoxicação por via inalatória e, no segundo evento, registrado em 2002, as altas concentrações de brevetoxinas no estômago indicaram intoxicação por fonte alimentar. Bossart e colaboradores (1998) constataram que o sistema imunológico dos peixes-boi desempenhou um papel significativo no evento de intoxicação por brevetoxicose. A toxina foi encontrada em macrófagos e linfócitos no pulmão, fígado e outros tecidos resultando em apoptose celular e comprometimento imunológico crônico.

Em 2020, a análise do perfil transcricional de peixes-boi após contato com brevetoxinas mostrou que a resposta imune, proliferação celular e apoptose foram as vias celulares predominantemente desreguladas. Vias neurodegenerativas também foram particularmente afetadas tanto em relação aos genes induzidos quanto aos genes reprimidos. Nossos resultados, após a exclusão dos genes relacionados à resposta imune – e que representam o maior grupo de genes diferencialmente expressos tanto em relação a genes induzidos quanto a genes reprimidos (LAZENSKY et al., 2020a) – indicaram elevado número de genes que codificam para genes relacionados a componentes da membrana plasmática como transportadores do tipo ABC (ABCC8, ABCC10), carreadores de soluto (SLC6A16, SLC12A3, SLC26A3, SLC22A17, SLC4A3) e canais de íons (ASIC2, ASIC1). Também elevado número de genes associados à transdução de sinais como proteínas de ligação de cálcio (CALCOCO1, CPNE6, CALB2, SYT3), proteínas de superfície celular neuronal (NLGN2, HTR3E), a família de glicoproteínas secretadas WNT (WNR5A, WNT2B) e proteínas quinase (RPS6KA1, MAGI2). O terceiro grupo de destaque compreende genes que codificam para proteínas regulatórias de transporte como SHANK1, que apresenta múltiplos domínios de repetição de anquirina; o receptor de prostaglandina PTGER3, que codifica para uma proteína membro da família de receptores acoplados à proteína G; e o gene CASP8, que codifica para regulador do ciclo celular.

Há grande predominância de genes reprimidos em resposta ao evento da maré vermelha. Dentre os genes induzidos foram observados os genes CASP8, CALCOCO1, além de genes como FBXW11, que codifica para um membro da família de proteínas F-box com atividade E3 ubiquitina ligase e participa da patogênese de múltiplos tumores (LIU et al., 2022), e o gene ATXN3, que contém um domínio de ubiquitina-protease em sua porção N-terminal (JOHNSON et al., 2020) revelando a ativação de genes que codificam para proteínas relacionadas à sinalização celular e regulação de processos biológicos.

A exclusão dos genes relacionados ao sistema imune foi feita de acordo com os grupos segundo a categorização do GO. Por tal motivo ainda muitos genes relacionados ao sistema imune permanecem em nossos resultados: diversos genes relacionados ao sistema imunológico estão agrupados em mais de uma categoria funcional. A categorização funcional dos genes super-representados, avaliados por termos GO (*Gene Ontology*), agrupa os genes diferencialmente expressos segundo a funcionalidade, inferindo funções biológicas.

Obtivemos em nossas análises 1429 genes diferencialmente expressos. Definimos como padrão de corte $\log_2$ fold change ≥ 2 ou ≤ -2 . O trabalho de Lazensky e colaboradores (2020) não apresenta o padrão de corte estabelecido tendo o referido trabalho confirmado como diferencialmente expressos 591 genes. Foi utilizada, portanto, maior estringência no corte.

A análise de genes modulados positiva e negativamente dentro de cada categoria evidencia, para os grupos analisados, não haver um perfil homogêneo de repressão ou indução por categoria, com predominância para genes reprimidos em relação aos induzidos. Fazendo uma análise mais específica em relação aos genes com maiores valores de acúmulos de transcritos observamos que dentre os genes mais induzidos estão alguns associados à atividade de transporte. O gene que codifica para uma subunidade do complexo de proteína adaptadora (AP1B1) atua no tráfego intracelular mediado por vesículas, por vias de transporte endocíticas e secretoras (PARK; GUO, 2014); o gene mitocondrial transportador de sódio/hidrogênio 9B2-like pertence a uma família de carreadores de solutos (HOLMES; SPRADLING REEVES, 2016). Alguns fatores de transcrição também foram identificados entre os genes mais reprimidos, como o fator de transcrição nuclear Y, essencial na regulação de numerosos processos biológicos como proliferação celular e apoptose (LI et al., 2018), e o fator de transcrição do tipo Zinc Finger 618 (ZNF618), que atua de forma cooperativa com a proteína ligase u3-ubiquitina UHRF2 regulando a metilação de DNA em células diferenciadas (LIU et al., 2016).

Entre os genes reprimidos mais significativamente regulados estão, além dos genes que codificam para interleucinas IL6 e IL1A e da quimiocina CCL4, moduladas em resposta a processos inflamatórios (CHANG; CHEN, 2016; GOSHEN; YIRMIYA, 2009; WILLENBERG et al., 2002), o gene GRP1, um ativador de GTPases, membro da família das citohesinas, associado à disseminação e adesão celular, quimiotaxia e tráfego de proteínas (POIRIER et al., 2005); e o gene TRPC5, um receptor de canais iônicos transdutor de estímulos térmicos. Ambos estão, portanto, associados à atividade celular de transporte. O gene HBEGF codifica para um membro da família de fatores de crescimento do fator de crescimento epidérmico que estimulam o crescimento e a diferenciação. É uma glicoproteína de membrana (RAAB; KLAGSBRUN, 1997).

Das análises de enriquecimento de vias metabólicas realizadas com o *Kegg ontology* destacamos duas vias importantes para a sinalização celular. A via clássica de sinalização das MAP quinases (também conhecida como via de quinases reguladas por sinal extracelular - ERK) está reprimida sob as condições experimentais avaliadas. Esta via regula a proliferação celular em um processo complexo: uma vez ativadas, as quinases translocam para o núcleo e transativam fatores de transcrição, alterando a expressão gênica modulando de diversos eventos celulares como crescimento, diferenciação ou mitose (ZHANG; LIU, 2002). De forma interessante, alguns genes da via das proteínas quinases ativadas por estresse (SAPK/Jun N-terminais quinases - JNK), responsiva a uma variedade de estresses ambientais, citocinas inflamatórias e fatores de crescimento estão ativados: o fator de transcrição NFAT-4, envolvido na geração e sobrevivência de células T e o fator de transcrição AP-1, que desempenha um papel conservado na resposta imediata ao estresse controlando a proliferação e sobrevivência celular (AGRON et al., 2017; NISHINA, 2004; OUKKA et al., 1998). A indução destes dois fatores de transcrição sugere regulação sobre eventos de diferenciação, proliferação, mas, sobretudo, inflamação.

A via das MAP quinases está diretamente interligada à via de sinalização Wnt. As proteínas Wnt mediam a transdução de pelo menos três principais vias de sinalização que desempenham papéis centrais em muitas etapas iniciais e tardias do desenvolvimento como destino celular, proliferação e migração; estão envolvidas também em eventos embriológicos importantes, incluindo desenvolvimento de membros. Na primeira dentre as vias principais, denominada via Wnt/ β -catenina ou via Wnt “canônica”, observamos a repressão de genes fundamentais como Wnt e Frizzled (os ativadores das vias), TCF (fator nuclear), um regulador/antagonista APC (CROCE; MCCLAY, 2008), e genes efetores como fator de transcrição c-jun e fra-1, superexpressos em células de câncer colorretal (BEJJANI et al.,

2021). O gene que codifica para o fator de transcrição PPAR gama (receptor ativado por proliferador de peroxissoma gama do) que está induzido em nossos resultados (cerca de 14-fold; $\log_2 = 3.82$), é regulado negativamente enquanto a via canônica Wnt/beta-catenina é regulada positivamente em células cancerígenas (LECARPENTIER et al., 2017). Observamos, portanto, um padrão de modulação invertido em relação a células de câncer de cólon.

As vias Wnt “não-canônicas” da polaridade celular planar (PCP) e a via Wnt-cálcio (Wnt/Ca²⁺) também estão reprimidas. Assim como a via canônica, essas duas vias não canônicas são ativadas pelos mesmos eventos iniciais: a interação de um ligante Wnt com seu receptor cognato Frizzled seguido pela ativação do efector citoplasmático. No entanto, diferentes proteínas adicionais estão envolvidas, seja para regular a interação Wnt/Frizzled ou para transduzir o sinal para o núcleo (CROCE; MCCLAY, 2008). Em nossos resultados no final da via Wnt-cálcio observou-se o acúmulo diferencial de transcritos do gene NFAT em nossos resultados (cerca de 4-fold; $\log_2 = 2.01$). Este gene codifica para um fator nuclear de células T ativadas que regula negativamente a sinalização canônica de Wnt/ β -catenina ajudando a reprimir a via canônica (HUANG et al., 2011). Nossos resultados evidenciam que, ainda que as funções celulares estejam fortemente reprimidas em resposta aos efeitos causados pelo evento da maré vermelha, a atividade transcricional busca alternativas para sustentar a atividade celular.

5.2. Duas estratégias foram necessárias para a verificação da ocorrência de eventos de processamento alternativo em peixes-boi da subespécie *Trichechus manatus latirostris*

Para gerar a lista de genes processados de forma alternativa foram utilizadas ferramentas livres e de código aberto, que permitem a identificação de novos genes, além de isoformas de genes conhecidos através da quantificação de diferenças na quantidade de RNA transcrito. As *reads* geradas no transcrito alinhadas ao genoma de referência pelo pacote *Cufflinks* (Trapnell et al., 2010) foram mapeadas agrupando as leituras em possíveis transcritos e, em seguida, gerando a montagem final do transcrito. Os dados de diferencial de expressão apenas quantificam a presença ou ausência de transcritos para determinados genes e, como resultado de análises estatísticas, estas são definidas ou não como diferencialmente expressas. O pacote *Cufflinks* prevê estruturas de transcrição usando um gráfico de sobreposição entre as leituras das *reads* seqüenciadas, buscando alinhamentos compatíveis (padrões de emenda semelhantes). Este gráfico de sobreposição representa

transcrições putativas (TRAPNELL et al., 2010). Nas análises de *splicing* alternativo, as leituras são montadas a partir de duas ou mais condições biológicas. A estimativa da abundância de transcritos é realizada atribuindo a cada leitura uma probabilidade de pertencer a qualquer uma das isoformas (um valor que depende da estimativa de abundância atual de isoformas) e atribuindo *reads* a isoformas (GHOSH; CHAN, 2016). A ferramenta *Cuffdiff* realiza testes para mudanças relativas de abundância de isoformas, chamadas de sobrecarga pós-transcricional (TRAPNELL et al., 2010).

Como resultado das análises realizadas com o pacote *Cufflinks* obtivemos uma lista de genes e regiões gênicas que possivelmente sofrem processamento alternativo. Aqui consideramos a possibilidade de ocorrência destes processamentos alternativos. Somente após validações, utilizando material biológico, poderemos afirmar ou refutar a ocorrência dos eventos de *splicing* alternativo. Nesta lista não temos, no entanto, a identificação do tipo de processamento sofrido. Com este propósito utilizamos os arquivos de saída desta ferramenta e utilizamos a ferramenta *IsoformSwitchAnalyzeR*, que identificou, anotou e visualizou os eventos de *splicing* alternativo de acordo com o tipo de processamento sofrido.

5.2.1. O número de genes diferencialmente modulados é maior do que os genes envolvidos em *splicing* alternativo

Identificamos menor quantidade de genes processados de forma alternativa do que genes diferencialmente expressos. Do total de 852 regiões do genoma potencialmente sofrendo *splicing* alternativo, 534 estão anotadas no genoma de *H. sapiens*. Dentre estas 534 regiões identificamos 234 genes, o que significa que para alguns genes mais de uma isoforma alternativa foi identificada.

Dentre os genes possivelmente apresentando mais de uma isoforma e categorizados como transferases, identificamos possível ocorrência de *splicing* alternativo em genes como a exostosina glicosiltransferase do tipo 1 (EXT1; *exostosin glycosyltransferase 1*) e a proteína 5 tipo polipeptídeo N-acetilgalactosaminiltransferase (GALNTL5; *polypeptide N-acetylgalactosaminyltransferase-like protein 5*). Para estes dois genes já foi verificada a ocorrência de processamento alternativo: uma mutação (c.1284+2del) no gene EXT1 resulta em um evento de *éxon skipping* (éxon 4) durante o *splicing* de mRNA, resultando na expressão reduzida de EXT1/EXT2. Este evento de processamento alternativo fornece evidências genéticas moleculares para um novo mecanismo patogênico de exostoses múltiplas hereditárias (HME) (GUO et al., 2019). Quanto ao gene GALNTL5, foi verificada que uma variante de *splicing* em amostras de testículo de touro leiteiro da raça holandesa. Um evento

de *exon skipping* ocorrida no éxon 2 resultou em uma proteína predita muito menor do que a proteína resultante do processamento constitutivo. Ainda que os níveis transcricionais da isoforma alternativa tenha sido similar ao da isoforma constitutiva, a proteína resultante do processamento alternativo pode não ser funcional (LI et al., 2016).

Entre as hidrolases, o gene que codifica para um inibidor de endopeptidases denominado inibidor da via do fator tecidual (TFPI; *tissue factor pathway inhibitor*) está associado ao bloqueio da iniciação da coagulação em humanos. O controle se dá pelo *splicing* alternativo na extremidade 3' do éxon 2 deste gene, que origina 2 isoformas: TFPI α e TFPI β (ELLERY et al., 2014). Também uma pequena GTPase identificada como GIT2, que possui atividade de conversão de GTP a GDP, apresenta ao menos 10 isoformas alternativas expressas com alto grau de especificidade de tecido a depender da isoforma alternativa, de modo que diferentes tecidos expressam complementos muito distintos de variantes de GIT2. A família de proteínas GIT provavelmente regula muitos receptores acoplados à proteína G. Dois eventos de *splicing*, especificamente, foram avaliados, um gerando uma isoforma funcional, que altera a expressão do receptor β 2-adrenérgico em células de superfície, em humanos, e uma segunda isoforma, que apresenta uma proteína terminada prematuramente (PREMONT et al., 2000).

Uma terceira hidrolase, que codifica para o gene ADAM17 (*a disintegrin and metalloprotease domain -containing protein 17*), uma metaloprotease que atua sobre células imunes e não imunes clivando vários substratos, incluindo ligantes do receptor do fator de crescimento epidérmico (EGFR), moléculas de adesão, citocinas pró-inflamatórias e quimiocinas e seus receptores. É um regulador chave dos níveis de superfície de TNF α solúvel em macrófagos pró-inflamatórios e células dendríticas (ADU-AMANKWAAH et al., 2021). Foi observado em camundongos que uma substituição C794T no éxon 7 do gene ADAM17 resulta em uma isoforma aberrante que altera drasticamente o fenótipo, incluindo defeitos do segmento anterior do olho, ausência das glândulas meibomianas e defeitos nas válvulas cardíacas semilunares.

A análise de enriquecimento funcional de genes revelou grande número de genes relacionados a resposta a estresse sendo processados de forma alternativa nos peixes-boi. O Gene Ontology requer a disponibilidade de dados suficientes de anotação funcional para o transcriptoma em estudo. Esta ferramenta contém dados de anotação para a maioria das espécies modelo, mas não possui em seu banco de dados a anotação de peixes-boi e, por isso, utilizamos o genoma humano como referência.

Apenas 4 genes apresentam modulação constitutiva e alternativa dentre os genes contidos nas quatro principais categorias do GO: CXCR2 e HRH4 induzidos, e NFKB1 e TRIB1 reprimidos em resposta ao evento de maré vermelha. O receptor 2 de quimiocina (motivo C-X-C) está entre os 10 genes mais induzidos em função dos efeitos da maré vermelha e atua na regulação do sistema imune (LAZENSKY et al., 2020a). As quimiocinas são uma superfamília de proteínas de moléculas pequenas e regulam seletivamente o recrutamento e a ativação de subconjuntos de leucócitos para locais preferenciais por meio de quimiotaxia. O gene CXCR2 é um membro da superfamília de receptores acoplados à proteína G e atua como potente promotor da angiogênese, proliferação e invasão tumoral (WANG et al., 2015b). Já foi avaliado que este gene possui 11 éxons e que o quadro de leitura aberta é de 1065 bp, quadro inteiramente localizado no último éxon. Os éxons restantes formam 11 regiões 5' não traduzidas diferentes por meio de *splicing* alternativo (KORBECKI et al., 2022). Ainda não se sabe o significado da formação de um número tão grande de diferentes mRNAs, mas estes eventos podem estar relacionados à regulação da expressão de CXCR2.

HRH4 codifica para um receptor de histamina (H4) que atua em processos imunológicos, inflamatórios e na proliferação celular e é predominantemente expresso em células hematopoiéticas e imunes. Está possivelmente associado à proliferação de células tumorais, sobretudo de carcinomas do cólon (ZHANG et al., 2012). Em humanos o gene forma uma proteína funcional de 390 aminoácidos e estão descritas duas isoformas alternativas no UniProt. A transcrição de HRH4 está associada à ativação da via Erk1/2 das MAP quinases (KARLSTEDT; JIN; PANULA, 2013), não modulada em nossos resultados. Os eventos de processamento alternativo podem estar atuando na regulação desta via nos peixes-boi em nossas condições experimentais.

NFKB1 (Nuclear Factor Kappa B Subunit 1) codifica para um fator de transcrição envolvido em processos celulares como desenvolvimento e neuro degeneração. O gene humano possui 24 éxons e codifica uma proteína de 968 aminoácidos. Foram detectados transcritos alternativos por *éxon skipping* entre os éxons 4 ou 5, que levam ao mRNA não funcional truncado relacionado à adaptação ao estresse provocado pelas altitudes andinas (SONG et al., 2018). Em nossos resultados a repressão deste gene, somado ao acúmulo de transcritos da isoforma alternativa pode estar associada a uma tentativa de resposta aos efeitos nocivos das toxinas da maré vermelha. Por fim, o acúmulo de transcritos alternativos do gene TRIB1 também foi contrabalanceado com a repressão da isoforma constitutiva. Este gene pertence à família de pseudoquinases, um sub-ramo da subfamília de proteína quinase ativada

por Ca^{2+} /calmodulina. Os avanços da pesquisa em genes homólogos de *Tribbles* (TRIB) estabeleceram o consenso de que essa família de proteínas desempenha papéis em diversas condições biológicas e regula redes de sinalização intracelular, atuando diretamente nas principais vias de sinalização celular, como as vias MAPK e NF- κ B. O gene TRIB1 é moderadamente expresso em vários órgãos e tecidos, incluindo a glândula tireóide, medula óssea, fígado e pulmão. Possui três éxons e, além de seu processamento constitutivo possui uma isoforma alternativa validada em humanos, encontrada em baixa abundância. A isoforma alternativa consiste na associação dos éxons 2 e 3, apresenta um encurtamento na região 5'UTR e uma região de codificação 5' adicional não encontrada na isoforma constitutiva (DANGER; FESEHA; BROUARD, 2022).

Em nossos resultados observamos que muitas proteínas quinase possivelmente sofrem processamento alternativo. Estes genes, que apresentam acúmulo de transcritos de suas isoformas geradas por *splicing* alternativo, não são os mesmos genes observados nos resultados de diferencial de expressão. Nossos resultados sugerem que, somada à modulação imposta pela transcrição dos genes em sua isoforma constitutiva, há um rol de genes que contribuem na modulação da via das proteínas quinase, mas que agem exclusivamente pela ativação de sua isoforma não-constitutiva. Estes resultados ressaltam haver, portanto, um número ainda maior de genes modulados na **Error! Reference source not found.** – e que estão ativos nesta via pela transcrição de isoformas alternativas. Observamos, portanto, que o evento da maré vermelha não somente ativa a transcrição de genes, mas que este evento tóxico impõe ainda uma seleção modulatória: ou os genes estão diferencialmente expressos em sua isoforma constitutiva, ou estão diferencialmente expressos somente na sua isoforma alternativa, promovendo assim a resposta ao estresse gerado pelas toxinas do ambiente. Como citado anteriormente, apenas 4 genes apresentam modulação tanto de sua isoforma constitutiva quanto da isoforma alternativa nos principais grupos gênicos.

5.2.1.1. O sistema imune dos peixes-boi sofreu intensa modulação transcricional e apresenta grande número de genes processados de forma alternativa

Numericamente, entre os genes relacionados ao sistema imune, 3 (TRIB1, CXCR2, NFKB1) apresentam transcritos diferencialmente expressos tanto em sua isoforma constitutiva quanto na isoforma alternativa.

Os genes listados que possivelmente sofreram processamento alternativo e que estão relacionados ao sistema imune dos peixes-boi incluem somente aquelas regiões que foram anotadas no genoma de referência. Por limitações da ferramenta utilizada para a ontologia gênica, que não dispõe do genoma de peixe-boi em seu banco de dados, a anotação foi feita utilizando o genoma humano.

Nos três domínios abrangidos pela ontologia gênica (GO) observamos que, dentre os genes com função associada à resposta imune em peixes-boi para *processo biológico*, o termo mais enriquecido foi sinalização autócrina, função onde uma célula produz uma molécula sinalizadora para ela mesma responder. Quanto a *componente celular*, somente a atividade secretória foi enriquecida, enquanto em relação à *função molecular*, houve enriquecimento de receptores do tipo toll e quinases.

A análise geral de enriquecimento de vias metabólicas realizadas com o *Kegg ontology* apresentou um número elevado de genes relacionados ao transporte por vesículas. O gene STX8, por exemplo, é membro da família SNARE (receptor da proteína de ligação ao fator sensível a N-etilmaleimida solúvel) e está envolvido no tráfego (geração, transporte e fusão) de vesículas na sinapse imunológica. Stx8 está localizada principalmente nos compartimentos endossomais e lisossômicos de reciclagem tardia, com pouca expressão nos compartimentos endossomais iniciais (BHAT et al., 2016). Também entre os genes diferencialmente expressos identificamos alguns relacionados ao tráfego por vesículas evidenciando o papel deste tipo de transporte em resposta ao estresse sofrido pelos peixes-boi.

Avaliando as principais vias relacionadas a resposta imune identificando os genes que possivelmente apresentam processamento alternativo em nossas condições experimentais. Destacamos a via de sinalização TNF, IL-17 e a via de receptores Toll-like. Membros da superfamília TNF são citocinas pleiotrópicas que atuam na mediação da inflamação, vigilância imunológica, desenvolvimento e homeostasia do tecido linfóide, destruição da matriz extracelular, apoptose, além de outras funções. Os principais sinais transduzidos pelos membros da superfamília TNF incluem a ativação envolvendo o fator nuclear- κ B (NF- κ B), e modulando as vias quinase N-terminal JUN, proteína quinase ativada por mitógeno p38 e ERK1/ERK2 (AGGARWAL, 2003). A atuação do TNF se dá via receptores (TNFR). Em nossos resultados identificamos o possível processamento alternativo no gene que codifica para a quinase p38, que controla a síntese de TNF e também do fator de transcrição CREB (CRE-binding protein), ativado via p38 e responsável pela regulação de eventos celulares como relacionado ao metabolismo/proliferação neuronal, plasticidade sináptica e função mitocondrial (JENSEN et al., 2017). Inclui também o processamento alternativo do fator

nuclear- κ B, previamente discutido. Três variantes de proteína de p38 foram descritas até o momento: as isoformas 1 e 3 são proteínas de 359 aminoácidos com a inclusão variável do éxon 10 (isoforma 1) ou éxon 11 (isoforma 3) (WELLS et al., 2006).

As citocinas da família da interleucina 17 (IL-17) desempenham papel central no controle de infecções, especialmente contra infecções de origem bacteriana e fúngica. Se não controladas, essas citocinas inflamatórias contribuem para a patologia de inúmeras condições inflamatórias autoimunes e crônicas. A maioria dos membros da família IL-17 caracterizados até hoje mediam a sinalização através de receptores heterodiméricos compostos por IL-17RA e uma subunidade que confere especificidade de ligante ou sinalização (MONIN; GAFFEN, 2018). As marés vermelhas da Flórida ocorrem quando as espécies de dinoflagelados naturais, *K. brevis*, “florescem” e emitem neurotoxinas de algas, que foram responsáveis por numerosos episódios de mortalidade de peixes-boi e outros animais (LAZENSKY et al., 2020a). Nossos resultados evidenciam a influência exercida pelas toxinas destes dinoflagelados no processamento alternativo de genes da família da interleucina 17, incluindo quinases, quimiocinas e dois genes de resposta imune a patógenos, S100A8/A9. Também conhecidos como MRP8 e MRP14, respectivamente, são proteínas de ligação ao Ca^{2+} pertencentes à família S100. São constitutivamente expressos em neutrófilos e monócitos como sensores de Ca^{2+} , exercendo papel crítico na modulação da resposta inflamatória ao estimular o recrutamento de leucócitos e induzir a secreção de citocinas (WANG et al., 2018).

Na via de sinalização de receptores Toll-like o gene TLR2 sofre, possivelmente, processamento alternativo, assim como TLR7/8. Um requisito fundamental de qualquer sistema imunológico é a capacidade de reconhecer a invasão de patógenos nos tecidos e, em seguida, iniciar uma resposta imune eficaz contra o patógeno. Essa detecção é detectada por uma série de receptores de reconhecimento de patógenos especializados, que inclui a família de receptores semelhantes a Toll (TLR) (PHILBIN et al., 2005). Os receptores Toll-like (TLRs) 2 são transdutores de sinal e detectam lipoproteínas de membrana de um amplo espectro de bactérias sendo, portanto, ativados por patógenos. Supõe-se que seja um receptor central de reconhecimento de padrões nas respostas imunes do hospedeiro à invasão microbiana (LIEN et al., 1999). Cinco variantes de *splicing* foram identificados em humanos. Em uma delas foi identificada como funcional ainda que o éxon 2 tenha sido removido (*éxon skipping*). Todas as variantes de *splicing* continham a sequência completa do éxon 3 sugerindo ser esta porção essencial para a proteína. A variação da expressão das diferentes isoformas foi vista como dependente de tempo (HAEHNEL et al., 2002). Sugere-se que o

splicing alternativo possivelmente ocorrendo neste receptor seja uma resposta a proteínas presentes nas toxinas dos dinoflagelados.

Os membros da família TLR TLR7, TLR8 e TLR9 formam um cluster filogeneticamente relacionado ativados por componentes ribonucleicos/ácido nucleico ou análogos dessas moléculas provenientes de patógenos. Em mamíferos esses TLRs são dependentes da via endocítica e estão localizados em vesículas lisossômicas (PHILBIN et al., 2005). Os receptores TLR7/8 estão listados entre os genes sofrendo processamento alternativo, juntamente com outros genes associados a transporte por meio de vesículas. Todas estas informações sugerem haver necessidade da transcrição de isoformas não-constitutivas para ativar a defesa imune dos peixes-boi nas condições experimentais.

5.2.2. Os padrões de evento de *splicing* obtidos com a ferramenta IsoformSwitchAnalyzeR diferem do perfil geral observado na literatura

Os eventos de éxon mutuamente exclusivo (MEE), onde duas isoformas do mesmo gene contém dois éxons mutuamente exclusivos e que não são encontrados em nenhuma das outras isoformas desse gene, e salto múltiplos de éxons (MES), no qual comparado com o pré-RNA hipotético, vários éxons consecutivos são ignorados na isoforma analisada, não foram identificados em nossos resultados. O evento com maior número de ocorrências, ATSS, corresponde a isoformas com sítios alternativos de início de transcrição: em comparação ao pré-RNA constitutivo, o sítio ATSS usado é, por definição, mais à jusante do que o do pré-RNA. O segundo evento mais contabilizado, A3, corresponde a sítio acceptor alternativo 3', que corresponde a um sítio doador usado, por definição, mais à jusante do que o pré-RNA (o éxon a jusante é mais curto). Depois, o evento com maior quantificação foi A5, que corresponde a sítio doador alternativo 5' (o sítio doador usado é, por definição, mais a montante do que o pré-RNA - o éxon a montante é mais curto). O evento seguinte, com maior número de eventos, foi retenção de íntrons, seguido do evento de sítios alternativos de terminação de transcrição (ATTS), evento no qual o sítio ATTS usado é, por definição, mais a montante do que o pré-RNA. Por último, o evento com menos isoformas contabilizadas foi éxon *skipping*. Neste caso ocorre o salto de um éxon por vez em cada isoforma alternativa formada.

De forma intrigante, os eventos de *splicing* alternativo mais comuns em mamíferos são os éxon *skipping* e os sítios alternativos de processamento 5' e 3', e em contraste, retenção de íntron é o evento de *splicing* alternativo mais raro encontrado em mamíferos (NAKKA et al., 2018; SAMMETH; FOISSAC; GUIGÓ, 2008). Ainda que, geralmente, as buscas por eventos

de processamento alternativo em animais se restrinja aos cinco principais eventos reconhecidos (éxons skipping, éxons mutuamente exclusivos, sítio de junção alternativo 5' e 3' e retenção de íntrons), nossos dados ainda estariam em divergência com o padrão regular de processamento alternativo em mamíferos. Antes que assumamos que estes eventos ocorram de forma distinta em peixes-boi em relação a outros mamíferos iremos realizar novas análises (utilizando outra ferramenta), bem como faremos uma reanálise dos nossos dados (utilizando o IsoformSwitchAnalyzeR).

Dentre os genes identificados como possivelmente sofrendo processamento alternativo destacamos o gene que codifica para uma superóxido dismutase 2 (SOD2). Este desempenha papel vital no controle de espécies celulares reativas de oxigênio (ROS) que são geradas tanto em condições ótimas quanto em condições de estresse. Níveis excessivos de ROS podem causar danos irreversíveis a lipídios, proteínas, DNA e alterar o estado redox, podendo causar morte celular. Para manter a homeostase tanto sob condições regulares quanto em condições de estresse, sistemas de defesa enzimáticos (superóxido dismutases, catalases e peroxidases) e não enzimáticos (glutathione e outros compostos contendo tiol, ácido ascórbico, antocianina, etc.) que podem diminuir os níveis de ROS por eliminação ou oxidação são modulados. Em plantas foi observado o processamento alternativo de SOD2 em resposta a stress abiótico, sugerindo um mecanismo de resistência a condições limitantes (SAINI et al., 2021). Os eventos observados estão relacionados à região 3' e 5'UTR, com eventos de éxon *skipping* nos íntrons 3, 7 e 8. Em nossos resultados identificamos a possível ocorrência de éxon *skipping* (éxon 3) e também a possível ocorrência de retenção de íntron (íntron 2) e sítio doador alternativo 5'. Na **Error! Reference source not found.** está representado o padrão de processamento do gene SOD2 em resposta ao evento de maré vermelha em peixes-boi: a região em preto indica o íntron possivelmente retido em virtude do estresse causado pela toxinas, segundo nossas condições experimentais; em branco está indicado o éxon possivelmente perdido e, por fim, apresentamos uma representação do possível processado alternativo no caso da ocorrência de ambos processamentos ao mesmo tempo. Não incluímos na representação o processamento do sítio doador alternativo 5'.

SGK1 codifica para uma serina/treonina quinase pertencente à família AGK Kinase. É ubiquamente expresso e desempenha um papel fundamental na regulação hormonal de vários canais iônicos, transportadores, bombas e fatores de transcrição ou reguladores. A expressão de SGK1 é modulada por estresse celular e hormônios, incluindo glicocorticóides e mineralocorticóides. Evidências sugerem que o aumento da expressão ou função de SGK1 está relacionado à hipótese de estresse patogênico da depressão maior (DATTILO et al.,

2020). Análises por RT-PCR em várias linhas celulares e tecidos humanos confirmaram a expressão das 3 variantes alternativas de SGK1, que diferiram exclusivamente em seus primeiros éxons. A estrutura da região regulatória dos três variantes indicou responsividade comum aos glicocorticóides, mas atribuiu diferenças em resposta à hipóxia e diferenciação celular (SIMON et al., 2007). Em nossos resultados este gene somente não apresenta sítio doador alternativo 3' e 5'. Os demais eventos de processamento alternativo estão associados podendo formar mais de uma isoforma distinta da constitutiva.

As proteínas do tipo Nedd-4 são ubiquitina-ligase E3 com domínios transmembrana, que regulam as principais decisões de tráfego, incluindo o direcionamento de proteínas para proteossomos ou lisossomos. No processo de ubiquitinação, geralmente são as ubiquitina ligases da classe E3 atuam determinando o alvo proteico. As proteínas da família Nedd4 demonstraram ter diversas funções em diferentes organismos, incluindo regulação negativa de proteínas de membrana por endocitose, exocitose de certas proteínas de membrana, regulação de fatores de transcrição e clivagem endoproteolítica. Dada a infinidade de papéis para as proteínas da família Nedd4, é provável que existam vários mecanismos reguladores para iniciar, terminar e direcionar sua atividade (CONWAY; KINSMAN; KRAMER, 2022; FLASZA et al., 2002). A possível regulação por *splicing* alternativo foi verificada no gene humano da família Nedd4, que gera pelo menos seis isoformas de *splicing* alternativo. Uma das isoformas identificadas corresponde a evento no sítio doador alternativo 3', outra isoforma corresponde a evento no sítio 5'; as demais, corresponderam a eventos de éxon *skipping* e afetam a estrutura do domínio da proteína. A análise da abundância relativa das diferentes isoformas em uma variedade de tecidos encontrou evidências de que os eventos de *splicing* alternativo podem estar sendo regulados de maneira específica em cada tecido (FLASZA et al., 2002). Em nossos resultados observamos a possível formação de isoformas do tipo ATSS, com sítios alternativos de início de transcrição, A3, com sítio doador usado mais à jusante do que o definido no pré-RNA, ATTS, gerando isoforma processada mais a montante do que o pré-RNA constitutivo, além de eventos de retenção de íntron e éxon *skipping*. Na **Error! Reference source not found.** representamos a isoforma formada caso os eventos de éxon *skipping* e retenção de íntron acontecessem simultaneamente. A confirmação da ocorrência de ambos os eventos, simultaneamente ou independentemente, demandam validações por experimentos *in vivo*.

A identificação dos eventos associados ao gene da família Nedd4 evidencia um potencial papel regulatório frente ao estresse causado pelas toxinas da maré vermelha. Foi observado em células murinas que a presença de neurotoxinas que provocam estresse

oxidativo em neurônios, como camptotecina, peróxido de hidrogênio e zinco, induzem a expressão da proteína NEDD4 (KWAK et al., 2012). Nossos resultados evidenciam tanto a modulação diferencial de isoformas que atuam contra reativas de oxigênio (ROS), bem como observamos a expressão diferencial de genes associados ao sistema nervoso como neurogênese e diferenciação neuronal (**Error! Reference source not found.**) sugerindo a ocorrência da modulação por processamento alternativo de forma a regular os efeitos nocivos das toxinas sobre o sistema nervoso dos peixes-boi.

A real ocorrência destes eventos de *splicing* precisa passar por validação para que se estabeleça a relação entre a contagem de transcritos e a efetiva ocorrência do evento de processamento. Um possível efeito biológico resultante deste processamento não constitutivo também requer validação.

5.2.2.1. Proporcionalmente, o número de genes de resposta imune diferencialmente expressos é menor do que o número de genes associados a eventos de *splicing* alternativo

Dentre os genes diferencialmente expressos em resposta às toxinas da maré vermelha, conforme evidenciado por Lazensky e colaboradores (2020), os relacionados à resposta imune estavam significativamente modulados estando, na maior parte das vezes, reprimidos. Dedicamos especial atenção aos genes que sofreram processamento alternativo e que estavam relacionados à resposta imune para expandir o conhecimento quanto à relação da resposta dos peixes-boi às toxinas das algas.

Na **Error! Reference source not found.** destacamos que 23 dentre os 97 genes associados a eventos de *splicing* alternativo pela ferramenta IsoformSwitchAnalyzeR estão associados à resposta imune dos peixes-boi. Alguns apresentam função regulatória, como o gene que codifica para o regulador do tráfego lisossomal (do inglês *lysosomal trafficking regulator*) e atua na regulação de múltiplos aspectos da atividade lítica das células NK (*natural killer*) (GIL-KRZEWSKA et al., 2016), e o fator de transcrição do tipo zinc finger contendo domínio MYM tipo 5 (RAKHRA; RAKHRA, 2021), ainda pouco estudado em relação à sua atividade regulatória sobre o sistema imune. Eventos de processamento alternativo foram descritos em camundongos para membros da família dos reguladores do tráfego lisossomal por salto múltiplos de éxons (BOONE et al., 2010) e retenção de íntron (BARBOSA et al., 1997). Nenhuma das isoformas alternativas foi avaliada quanto à sua funcionalidade. Enquanto o gene que codifica para o regulador do tráfego lisossomal não apresentou diferencial de expressão para sua isoforma constitutiva, o gene que codifica para o

fator de transcrição do tipo zinc finger MYM tipo 5 acumulou transcritos de forma diferencial sugerindo um papel difuso na regulação do sistema imune, tanto pela ativação da isoforma constitutiva quanto de isoformas alternativas.

Dois receptores tipo tirosina fosfatase também foram associados a eventos de processamento alternativo e possuem função regulatória pela catalise da desfosforilação de proteínas tirosina (Tyr)-fosforiladas. Enquanto a atividade do gene tirosina fosfatase tipo E (PTPRE) não está bem estabelecida na literatura, o gene do receptor da proteína tirosina fosfatase tipo C (PTPRC) codifica para o receptor do antígeno leucocitário comum (CD45). CD45 afeta a adesão celular, migração, sinalização de citocinas, desenvolvimento celular e estado de ativação. Eventos de processamento alternativo foram observados em ambos: para PTPRE, uma isoforma não transmembranar foi identificada, diferindo da isoforma regular, que possui domínio transmembrana e é expresso na superfície celular e pode estar associada à regulação do sistema hematopoiético (ELSON; LEDER, 1995). Para o gene PTPRC foi descrita a ocorrência de processamento alternativo dos éxons 4, 5 e 6 gerando pelo menos 8 isoformas, sendo que cinco dentre elas são funcionais em humanos. Os eventos de *splicing* alternativo parecem estar associados a alterações na sensibilidade a estímulos externos podendo contribuir para autoimunidade e imunodeficiência (HERMISTON; XU; WEISS, 2003). Além disso, avanços recentes sugerem que a modulação da função CD45 pode ter benefícios terapêuticos em muitos estados de doença.

A **Error! Reference source not found.** apresenta uma possível isoforma decorrente da retenção do íntron 21 no gene que codifica para PTPRC em peixes-boi. Para o mesmo gene foi observada ainda a possibilidade de ocorrência de processamento alternativo por *exon skipping*, A3 (sítio aceptor alternativo 3'), A5 (sítio doador alternativo 5') e ATSS (sítios alternativos de início de transcrição). O gene que codifica para PTPRC está diferencialmente expresso, enquanto PTPRE não apresentou diferencial de expressão de sua isoforma constitutiva. Nossos resultados sugerem que tanto o perfil transcricional quanto a ocorrência de processamento alternativo para PTPRC podem também ser uma modulação direcionada pela exposição aos estímulos externos nocivos decorrentes da exposição às toxinas.

Um gene que de forma putativa codifica para um MHC de classe II (HLA-DQA2), proteína de membrana presente em células apresentadoras de antígeno, foi associado a eventos de *splicing* em nossos resultados. O complexo principal de histocompatibilidade humano (MHC) ou HLA codifica antígenos de superfície celular classe I e classe II que estão também relacionados à atividade regulatória de respostas imunes funcionais (AUFRAY et al., 1987). Os genes MHC de classe II codificam as duas cadeias que formam os

heterodímeros funcionais HLA-DR, HLA-DQ, HLA-DP, HLA-DM e HLA-DO (HOUTMAN et al., 2021). HLA-DQA2 pertence à família de cadeias alfa HLA classe II. A proteína codificada forma um heterodímero com uma cadeia beta de classe II. Ele está localizado em vesículas intracelulares e desempenha um papel central no carregamento de peptídeos de moléculas do MHC classe II, ajudando a liberar o peptídeo de cadeia invariante associado à classe II (CLIP) do sítio de ligação do peptídeo. As moléculas de classe II são expressas em células apresentadoras de antígenos (linfócitos B, células dendríticas, macrófagos) e são usadas para apresentar peptídeos antigênicos na superfície celular para serem reconhecidos pelas células T CD4 (ÁLVARO-BENITO; FREUND, 2020).

Não identificamos o acúmulo diferencial de transcritos do gene MHC Classe II HLA-DQA2, mas nossos resultados evidenciaram a ocorrência de processamento alternativo deste gene por ATSS, ATTS e IR. Na **Error! Reference source not found.** está representada de forma esquemática o possível evento de IR, que resulta da retenção do íntron 3 na isoforma alternativa final. Por meio de experimentos de Northern blot e dados de sequência de mRNA foi demonstrada a existência de mRNA de DQA2 processado de forma alternativa em linhagens de células linfoblastóides B humanas, que se destacam pela resposta imunológica realizada pela produção de anticorpos. Em contraste, não foi encontrada esta isoforma diferencial em linhagem de monócitos ou em células T ativadas (YU; SHEEHY, 1991).

O gene que codifica para uma proteína 8 induzida pelo fator de necrose tumoral alfa (TNFAIP8), um tipo de proteína antiapoptótica em tumorigênese ou doenças inflamatórias. A expressão de TNFAIP8 é encontrada de forma ubíqua na maioria dos tecidos humanos, incluindo medula óssea, células imunes, pulmão, pâncreas, rim, fígado. Sua expressão é induzida em resposta a sinais inflamatórios mediados pelo TNF α (CHENG et al., 2022). Foi observado em células dendríticas que TNFAIP8 regula a via JNK, das proteínas quinases ativadas por estresse (LIU et al., 2021). Esta via está sendo regulada negativamente em nossos resultados como visto nos resultados de diferencial de expressão. Nos dados referentes a processamento alternativo este mesmo gene está possivelmente associado a eventos de IR e A3. Análise em larga escala evidenciou a ocorrência de eventos de *splicing* alternativo na região promotora de TNFAIP8 em humanos. Não foram relatados os resultados biológicos decorrentes destes eventos de *splicing* (XIONG et al., 2018). Até o momento oito isoformas estão descritas (NITURE et al., 2019).

Por fim, o gene que codifica para uma proteína do tipo Nedd-4, abordado anteriormente (**Error! Reference source not found.**), também tem atividade relacionada à

resposta imune dos peixes boi atuando diretamente tanto na resposta ao stress quanto na neurogênese e diferenciação neuronal.

Proporcionalmente, um número maior de genes associados ao sistema imune de peixes-boi acumularam transcritos de forma alternativa comparado ao número de genes do sistema imune que apresentaram significativo diferencial de expressão. Enquanto nos resultados de diferencial de expressão 98 genes dentre 591 foram correlacionados ao sistema imune, pelas análises obtidas com a ferramenta IsoformSwitchAnalyzeR temos 23 genes dentre os 97 anotados acumulando transcritos de isoformas não-constitutivas. Se anteriormente foi relatada a importância da expressão de genes de resposta imune como resposta dos peixes-boi ao efeito tóxico produzido pelas algas aqui evidenciamos que muitos genes são transcritos de forma alternativa também como uma estratégia de resistência aos efeitos nocivos das toxinas.

6. CONCLUSÕES

As análises de expressão gênica e processamento alternativo evidenciaram que o sistema imune é intensamente modulado nos peixes-boi em resposta às toxinas de algas. Um número maior de genes relacionados à atividade e regulação do sistema imune sofre *splicing* alternativo comparativamente ao número de genes com acúmulo diferencial de transcritos. A transcrição é, de forma geral, reprimida em resposta às toxinas; um número menor de genes é induzido nestas condições experimentais. Muitas vias associadas à regulação de genes foram reprimidas sugerindo um efeito robusto das toxinas sobre os genes com atividade regulatória nos peixes-boi.

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ADU-AMANKWAAH, J. et al. ADAM17, A Key Player of Cardiac Inflammation and Fibrosis in Heart Failure Development During Chronic Catecholamine Stress. **Frontiers in Cell and Developmental Biology**, v. 9, 13 dez. 2021.
- AGGARWAL, B. B. Signalling pathways of the TNF superfamily: a double-edged sword. **Nature Reviews Immunology** 2003 3:9, v. 3, n. 9, p. 745–756, 2003.
- AGRAWAL, A. A. et al. Targeting splicing abnormalities in cancer. **Curr. Opin. Genet. Dev.**, v. 48, p. 67–74, 1 fev. 2018.
- AGRON, M. et al. Regulation of AP-1 by MAPK Signaling in Metal-Stressed Sea Anemone. **Cellular Physiology and Biochemistry**, v. 42, n. 3, p. 952–964, 2017.
- AHMED, M. Y. et al. Loss of PCLO function underlies pontocerebellar hypoplasia type III. **Neurology**, v. 84, n. 17, p. 1745–1750, 28 abr. 2015.
- ÁLVARO-BENITO, M.; FREUND, C. Revisiting nonclassical HLA II functions in antigen presentation: Peptide editing and its modulation. **HLA**, v. 96, n. 4, p. 415–429, 1 out. 2020.
- ARNASON, U. et al. Mitogenomic relationships of placental mammals and molecular estimates of their divergences. **Gene**, v. 421, n. 1–2, p. 37–51, 15 set. 2008.
- AUFFRAY, C. et al. Structure and expression of HLA-DQ α and -DX α genes: interallelic alternate splicing of the HLA-DQ α gene and functional splicing of the HLA-DX α gene using a retroviral vector. **Immunogenetics** 1987 26:1, v. 26, n. 1, p. 63–73, jan. 1987.
- BAKER, P. **VARIABILIDADE DO GENE DQB DO COMPLEXO PRINCIPAL DE HISTOCOMPATIBILIDADE (MHC) DE CLASSE II EM PEIXES-BOI DA AMAZÔNIA (*Trichechus inunguis*, SIRENIA, MAMMALIA)**. Dissertação de Mestrado—Belém: UFPA, fev. 2015.
- BANERJEE, A. K. et al. SARS-CoV-2 Disrupts Splicing, Translation, and Protein Trafficking to Suppress Host Defenses. **Cell**, v. 183, n. 5, p. 1325–1339.e21, 25 nov. 2020.
- BARBOSA, M. D. F. S. et al. Identification of Mutations in Two Major mRNA Isoforms of the Chediak-Higashi Syndrome Gene in Human and Mouse. **Human Molecular Genetics**, v. 6, n. 7, p. 1091–1098, 1 jul. 1997.
- BATTLE, A. et al. Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. **Genome Research**, v. 24, n. 1, p. gr.155192.113, 3 out. 2013.
- BEJJANI, F. et al. Fra-1 regulates its target genes via binding to remote enhancers without exerting major control on chromatin architecture in triple negative breast cancers. **Nucleic Acids Research**, v. 49, n. 5, p. 2488–2508, 18 mar. 2021.

- BERGET, S. M. Exon Recognition in Vertebrate Splicing. **Journal of Biological Chemistry**, v. 270, n. 6, p. 2411–2414, fev. 1995.
- BERTA, A.; SUMICH, J. L.; KOVACS, K. M. **Marine Mammals: Evolutionary Biology**. 2nd. ed. [s.l.] Elsevier Ltd, 2006. v. 45
- BHAT, S. S. et al. Syntaxin 8 is required for efficient lytic granule trafficking in cytotoxic T lymphocytes. **Biochimica et Biophysica Acta (BBA) - Molecular Cell Research**, v. 1863, n. 7, p. 1653–1664, 1 jul. 2016.
- BIAMONTI, G. et al. The alternative splicing side of cancer. **Semin. Cell Dev. Biol.**, v. 32, p. 30–36, 2014.
- BLAKE, J. A.; HARRIS, M. A. The Gene Ontology (GO) project: structured vocabularies for molecular biology and their application to genome and expression analysis. **Current protocols in bioinformatics**, v. Chapter 7, n. SUPPL. 23, 2008.
- BLENCOWE, B. J. The Relationship between Alternative Splicing and Proteomic Complexity. **Trends in Biochemical Sciences**, v. 42, n. 6, p. 407–408, 1 jun. 2017.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114, 8 ago. 2014.
- BONDE, R. K. Sirenian Life History. **Encyclopedia of Marine Mammals**, p. 859–861, 1 jan. 2018.
- BONDE, R. K.; MCGUIRE, P. M.; HUNTER, M. E. A review of the key genetic tools to assist imperiled species conservation: analyzing West Indian manatee populations. **Journal of Marine Animals and Their Ecology**, v. 5, n. 1, p. 8–19, 2012.
- BONNAL, S. C.; LÓPEZ-OREJA, I.; VALCÁRCEL, J. Roles and mechanisms of alternative splicing in cancer-implications for care. **Nat. Rev. Clin. Oncol.**, v. 17, n. 8, p. 457–474, 1 ago. 2020.
- BOONE, M. et al. The lysosomal trafficking regulator interacting protein-5 localizes mainly in epithelial cells. **Journal of Molecular Histology** 2010 41:1, v. 41, n. 1, p. 61–74, 1 abr. 2010.
- BOSSART, G. D. et al. Brevetoxicosis in manatees (*Trichechus manatus latirostris*) from the 1996 epizootic: gross, histologic, and immunohistochemical features. **Toxicologic pathology**, v. 26, n. 2, p. 276–282, 1998.
- BREAUX, B. et al. The Florida manatee (*Trichechus manatus latirostris*) T cell receptor loci exhibit V subgroup synteny and chain-specific evolution. **Developmental & Comparative Immunology**, v. 85, p. 71–85, 1 ago. 2018.

CANTANHEDE, A. M. Phylogeography and population genetics of the endangered Amazonian manatee , *Trichechus inunguis* Natterer , 1883 (Mammalia , Sirenia). p. 401–413, 2005.

CHANG, T. T.; CHEN, J. W. Emerging role of chemokine CC motif ligand 4 related mechanisms in diabetes mellitus and cardiovascular disease: friends or foes? **Cardiovascular Diabetology**, v. 15, n. 1, p. 117, 24 ago. 2016.

CHANG, Y. F.; IMAM, J. S.; WILKINSON, M. F. The nonsense-mediated decay RNA surveillance pathway. **Annual review of biochemistry**, v. 76, p. 51–74, 2007.

CHAUDHARY, S. et al. Alternative splicing and protein diversity: Plants versus animals. **Frontiers in Plant Science**, v. 10, p. 708, 31 maio 2019.

CHEN, L. et al. Correcting for Differential Transcript Coverage Reveals a Strong Relationship between Alternative Splicing and Organism Complexity. **Molecular Biology and Evolution**, v. 31, n. 6, p. 1402, 2014.

CHEN, L.; ZHENG, S. Identify Alternative Splicing Events Based on Position-Specific Evolutionary Conservation. **PLoS ONE**, v. 3, n. 7, p. 2806, 30 jul. 2008.

CHENG, X. et al. TNFAIP8 modulates the survival and immune activity of Th17 cells via p53/ p21/ MDM2 pathway after acute insult. **Cytokine: X**, v. 4, n. 1, p. 100062, 1 mar. 2022.

CHOI, K. H. et al. Gene expression and genetic variation data implicate PCLO in bipolar disorder. **Biological psychiatry**, v. 69, n. 4, p. 353–359, fev. 2011.

CLIMENTE-GONZÁLEZ, H. et al. The functional impact of alternative splicing in cancer. **Cell Rep.**, v. 20, n. 9, p. 2215–2226, 29 ago. 2017.

CONWAY, J. A.; KINSMAN, G.; KRAMER, E. R. The Role of NEDD4 E3 Ubiquitin–Protein Ligases in Parkinson’s Disease. **Genes 2022, Vol. 13, Page 513**, v. 13, n. 3, p. 513, 14 mar. 2022.

COREY, D. R. Nusinersen, an antisense oligonucleotide drug for spinal muscular atrophy. **Nature Neuroscience 2017 20:4**, v. 20, n. 4, p. 497–499, 13 fev. 2017.

CRAIG VENTER, J. et al. The sequence of the human genome. **Science**, v. 291, n. 5507, p. 1304–1351, 16 fev. 2001.

CRAY, C. et al. Assessement of serum amyloid A levels in the rehabilitation setting in the Florida manatee (*Trichechus manatus latirostris*). **Journal of zoo and wildlife medicine : official publication of the American Association of Zoo Veterinarians**, v. 44, n. 4, p. 911–917, dez. 2013.

CROCE, J. C.; MCCLAY, D. R. Evolution of the Wnt Pathways. **Methods in molecular biology (Clifton, N.J.)**, v. 469, p. 3, 2008.

DANGER, R.; FESEHA, Y.; BROUARD, S. The Pseudokinase TRIB1 in Immune Cells and Associated Disorders. **Cancers** **2022**, **Vol. 14**, **Page 1011**, v. 14, n. 4, p. 1011, 17 fev. 2022.

DATTILO, V. et al. The Emerging Role of SGK1 (Serum- and Glucocorticoid-Regulated Kinase 1) in Major Depressive Disorder: Hypothesis and Mechanisms. **Frontiers in Genetics**, v. 11, p. 826, 5 ago. 2020.

DE ARRAS, L.; ALPER, S. Limiting of the innate immune response by SF3A-dependent control of MyD88 alternative mRNA splicing. **PLoS genetics**, v. 9, n. 10, out. 2013.

DE KLERK, E.; 'T HOEN, P. A. C. Alternative mRNA transcription, processing, and translation: insights from RNA sequencing. **Trends in genetics : TIG**, v. 31, n. 3, p. 128–139, 1 mar. 2015.

DE OLIVEIRA, T. M. et al. TLR4 and TLR8 variability in Amazonian and West Indian manatee species from Brazil. **Genetics and Molecular Biology**, v. 44, n. 2, 2021.

DE SOUZA, É. M. S. et al. The evolutionary history of manatees told by their mitogenomes. **Scientific Reports** **2021 11:1**, v. 11, n. 1, p. 1–10, 11 fev. 2021a.

DE SOUZA, É. M. S. et al. The evolutionary history of manatees told by their mitogenomes. **Scientific Reports**, v. 11, n. 1, p. 3564, 11 dez. 2021b.

DIAZ-MUÑOZ, M. D. et al. The RNA-binding protein HuR is essential for the B cell antibody response. **Nature Immunology** **2015 16:4**, v. 16, n. 4, p. 415–425, 23 fev. 2015.

DJEBALI, S. et al. Landscape of transcription in human cells. **Nature** **2012 489:7414**, v. 489, n. 7414, p. 101–108, 5 set. 2012.

DUNHAM, I. et al. An integrated encyclopedia of DNA elements in the human genome. **Nature** **2012 489:7414**, v. 489, n. 7414, p. 57–74, 5 set. 2012.

EHRLICH, F. et al. Differential Evolution of the Epidermal Keratin Cytoskeleton in Terrestrial and Aquatic Mammals. **Molecular Biology and Evolution**, v. 36, n. 2, p. 328–340, 1 fev. 2019.

ELLERY, P. E. R. et al. Translation of human tissue factor pathway inhibitor- β mRNA is controlled by alternative splicing within the 5' untranslated region. **Arteriosclerosis, Thrombosis, and Vascular Biology**, v. 34, n. 1, p. 187–195, jan. 2014.

ELSON, A.; LEDER, P. Identification of a cytoplasmic, phorbol ester-inducible isoform of protein tyrosine phosphatase E (delayed early response/induction/differentiation). **Cell Biology**, v. 92, p. 12235–12239, 1995.

EMIN-LIMA, R. et al. AMAZONIAN MANATEES (*Trichechus inunguis*) INHABITING AN EQUATORIAL METROPOLIS: HISTORICAL RECORDS AND MATING ACTIVITY NEAR BELÉM, NORTHERN BRAZIL = SOBRE A PRESENÇA DO PEIXE-BOI

AMAZÔNICO (*Trichechus inunguis*) PRÓXIMO À BELÉM: REGISTROS HISTÓRICOS E ATIVIDADE REPRODUTIVA. **Boletim do Laboratório de Hidrobiologia**, v. 31, n. 2, p. 1–7, 22 dez. 2021.

EXCOFFIER, L.; SMOUSE, P. E.; QUATTRO, J. M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. **Genetics**, v. 131, n. 2, p. 479–491, 1992.

FAUSTINO, N. A.; COOPER, T. A. Pre-mRNA splicing and human disease. **Genes Dev.**, v. 17, n. 4, p. 419–437, 15 fev. 2003.

FERRANTE, J. A.; HUNTER, M. E.; WELLEHAN, J. F. X. DEVELOPMENT AND VALIDATION OF QUANTITATIVE PCR ASSAYS TO MEASURE CYTOKINE TRANSCRIPT LEVELS IN THE FLORIDA MANATEE (*TRICHECHUS MANATUS LATIROSTRIS*). **Journal of Wildlife Diseases**, v. 54, n. 2, p. 283–294, 1 abr. 2018.

FLASZA, M. et al. Alternative Splicing Determines the Domain Structure of WWP1, a Nedd4 Family Protein. **Biochemical and Biophysical Research Communications**, v. 290, n. 1, p. 431–437, jan. 2002.

FLEWELLING, L. J. et al. Red tides and marine mammal mortalities. **Nature** **2005 435:7043**, v. 435, n. 7043, p. 755–756, 8 jun. 2005.

GARCIA-RODRIGUEZ, A. I. et al. Phylogeography of the West Indian manatee (*Trichechus manatus*): How many populations and how many taxa? **Molecular Ecology**, v. 7, p. 1137–1149, 1998.

GE, S. X. et al. ShinyGO: a graphical gene-set enrichment tool for animals and plants. **Bioinformatics (Oxford, England)**, v. 36, n. 8, p. 2628–2629, 15 abr. 2020.

GHOSH, S.; CHAN, C. K. K. Analysis of RNA-seq data using TopHat and cufflinks. **Methods in Molecular Biology**, v. 1374, p. 339–361, 2016.

GIL-KRZEWSKA, A. et al. Chediak-Higashi syndrome: LYST domains regulate exocytosis of lytic granules, but not cytokine secretion by NK cells. **The Journal of allergy and clinical immunology**, v. 137, n. 4, p. 1165, 1 abr. 2016.

GÓMEZ-CARRASCO, G. et al. Genetic Diversity and Structure From Antillean Manatee (*Trichechus manatus manatus*) in the Southern Gulf of México: Comparison Between Connected and Isolated Populations: <https://doi.org/10.1177/1940082918795560>, v. 11, 31 ago. 2018.

GOSHEN, I.; YIRMIYA, R. Interleukin-1 (IL-1): A central regulator of stress responses. **Frontiers in Neuroendocrinology**, v. 30, n. 1, p. 30–45, jan. 2009.

GUO, W. et al. 3D RNA-seq: a powerful and flexible tool for rapid and accurate differential expression and alternative splicing analysis of RNA-seq data for biologists. **RNA biology**, v. 18, n. 11, p. 1574–1587, 2021.

GUO, X. et al. A novel splice mutation induces exon skipping of the EXT1 gene in patients with hereditary multiple exostoses. **International Journal of Oncology**, v. 54, n. 3, p. 859, 1 mar. 2019.

HAEHNEL, V. et al. Transcriptional Regulation of the Human Toll-Like Receptor 2 Gene in Monocytes and Macrophages. **The Journal of Immunology**, v. 168, n. 11, p. 5629–5637, 1 jun. 2002.

HARR, K. et al. Comparison of methods used to diagnose generalized inflammatory disease in manatees (*Trichechus manatus latirostris*). **Journal of zoo and wildlife medicine : official publication of the American Association of Zoo Veterinarians**, v. 37, n. 2, p. 151–159, 2006.

HERMISTON, M. L.; XU, Z.; WEISS, A. CD45: A Critical Regulator of Signaling Thresholds in Immune Cells. <http://dx.doi.org/10.1146/annurev.immunol.21.120601.140946>, v. 21, p. 107–137, 28 nov. 2003.

HOAGLAND, P. et al. Lessening the Hazards of Florida Red Tides: A Common Sense Approach. **Frontiers in Marine Science**, v. 7, p. 538, 9 jul. 2020.

HOLMES, R. S.; SPRADLING REEVES, K. D. Evolution of Vertebrate Solute Carrier Family 9B Genes and Proteins (SLC9B): Evidence for a Marsupial Origin for Testis Specific SLC9B1 from an Ancestral Vertebrate SLC9B2 Gene. **Journal of Phylogenetics & Evolutionary Biology**, v. 4, n. 3, 2016.

HOUTMAN, M. et al. Haplotype-Specific Expression Analysis of MHC Class II Genes in Healthy Individuals and Rheumatoid Arthritis Patients. **Frontiers in Immunology**, v. 12, p. 3315, 17 ago. 2021.

HU, Z. et al. Characterization and Comparative Transcriptomic Analysis of Skeletal Muscle in Pekin Duck at Different Growth Stages Using RNA-Seq. **Animals : an open access journal from MDPI**, v. 11, n. 3, p. 1–21, 1 mar. 2021.

HUANG, B. et al. Genome-Wide Analysis of Alternative Splicing Provides Insights into Stress Adaptation of the Pacific Oyster. **Marine biotechnology (New York, N.Y.)**, v. 18, n. 5, p. 598–609, 1 out. 2016.

HUANG, T. et al. Nuclear Factor of Activated T Cells (NFAT) Proteins Repress Canonical Wnt Signaling via Its Interaction with Dishevelled (Dvl) Protein and Participate in Regulating

Neural Progenitor Cell Proliferation and Differentiation. **The Journal of Biological Chemistry**, v. 286, n. 43, p. 37399, 10 out. 2011.

IÑIGUEZ, L. P. et al. Identification and analysis of alternative splicing events in *Phaseolus vulgaris* and *Glycine max*. **BMC genomics**, v. 18, n. 1, 22 ago. 2017.

IP, J. Y. et al. Global analysis of alternative splicing during T-cell activation. **RNA**, v. 13, n. 4, p. 563–572, 1 abr. 2007.

JEFFERSON, T. A.; LEATHERWOOD, S.; WEBBER, M. A. **MARINE MAMMALS OF THE WORLD FAO SPECIES IDENTIFICATION GUIDE** . Roma: [s.n.].

JENSEN, P. et al. TNF α affects CREB-mediated neuroprotective signaling pathways of synaptic plasticity in neurons as revealed by proteomics and phospho-proteomics. **Oncotarget**, v. 8, n. 36, p. 60223, 9 set. 2017.

JHA, A.; GAZZARA, M. R.; BARASH, Y. Integrative deep models for alternative splicing. **Bioinformatics**, v. 33, n. 14, p. i274–i282, 15 jul. 2017.

JOHNSON, S. L. et al. Ubiquitin-interacting motifs of ataxin-3 regulate its polyglutamine toxicity through hsc70-4-dependent aggregation. **eLife**, v. 9, p. 1–22, 1 set. 2020.

JUNK, W. J.; SILVA, V. M. F. DA. Mammals, Reptiles and Amphibians. p. 409–417, 1997.

KANEHISA, M.; GOTO, S. KEGG: kyoto encyclopedia of genes and genomes. **Nucleic acids research**, v. 28, n. 1, p. 27–30, 1 jan. 2000.

KARLSTEDT, K.; JIN, C.; PANULA, P. Expression of histamine receptor genes *Hrh3* and *Hrh4* in rat brain endothelial cells. **British Journal of Pharmacology**, v. 170, n. 1, p. 58, set. 2013.

KEITH-DIAGNE, L. **THE IUCN RED LIST OF THREATENED SPECIES™- African Manatee Assessment, Errata version** . , abr. 2017.

KENNEDY, S. Morbillivirus infections in aquatic mammals. **Journal of Comparative Pathology**, v. 119, n. 3, p. 201–225, 1 out. 1998.

KIM, D.; LANGMEAD, B.; SALZBERG, S. L. HISAT: a fast spliced aligner with low memory requirements. **Nature Methods** 2015 12:4, v. 12, n. 4, p. 357–360, 9 mar. 2015.

KIM, N. et al. A novel member of the leukocyte receptor complex regulates osteoclast differentiation. **Journal of Experimental Medicine**, v. 195, n. 2, p. 201–209, 21 jan. 2002.

KORBECKI, J. et al. CXCR2 Receptor: Regulation of Expression, Signal Transduction, and Involvement in Cancer. **International Journal of Molecular Sciences** 2022, Vol. 23, Page 2168, v. 23, n. 4, p. 2168, 16 fev. 2022.

KRAUT, N. et al. Requirement of the mouse *I-mfa* gene for placental development and skeletal patterning. **The EMBO Journal**, v. 17, n. 21, p. 6276–6288, 2 nov. 1998.

KUEHNER, C. et al. PCLO rs2522833 impacts HPA system activity in healthy young adults. **Translational psychiatry**, v. 1, n. 5, p. e10, maio 2011.

KUKURBA, K. R.; MONTGOMERY, S. B. RNA Sequencing and Analysis. **Cold Spring Harbor protocols**, v. 2015, n. 11, p. 951, 1 nov. 2015.

KWAK, Y. D. et al. Upregulation of the E3 ligase NEDD4-1 by Oxidative Stress Degrades IGF-1 Receptor Protein in Neurodegeneration. **The Journal of Neuroscience**, v. 32, n. 32, p. 10971, 8 ago. 2012.

LALOUM, T.; MARTÍN, G.; DUQUE, P. Alternative Splicing Control of Abiotic Stress Responses. **Trends in Plant Science**, v. 23, n. 2, p. 140–150, 1 fev. 2018.

LAZENSKY, R. et al. Investigating the gene expression profiles of rehabilitated Florida manatees (*Trichechus manatus latirostris*) following red tide exposure. **PLoS ONE**, v. 15, n. 7, 1 jul. 2020a.

LAZENSKY, R. et al. Investigating the gene expression profiles of rehabilitated Florida manatees (*Trichechus manatus latirostris*) following red tide exposure. **PLOS ONE**, v. 15, n. 7, p. e0234150, 2 jul. 2020b.

LECARPENTIER, Y. et al. Interactions between PPAR Gamma and the Canonical Wnt/Beta-Catenin Pathway in Type 2 Diabetes and Colon Cancer. **PPAR Research**, v. 2017, 2017.

LEUNG, M. K. K. et al. Deep learning of the tissue-regulated splicing code. **Bioinformatics**, v. 30, n. 12, 15 jun. 2014.

LEWIS, B. P.; GREEN, R. E.; BRENNER, S. E. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 1, p. 189–192, 7 jan. 2003.

LI, B. J. et al. Genome-Wide Characterization of Alternative Splicing Events and Their Responses to Cold Stress in Tilapia. **Frontiers in Genetics**, v. 11, p. 244, 18 mar. 2020.

LI, G. et al. The animal nuclear factor Y: an enigmatic and important heterotrimeric transcription factor. **American Journal of Cancer Research**, v. 8, n. 7, p. 1106, 2018.

LI, H. et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics (Oxford, England)**, v. 25, n. 16, p. 2078–2079, ago. 2009.

LI, H. et al. A novel splice variant of the bovine GALNTL5 gene identified in Chinese Holstein bull testis tissue and its mRNA expression. **Genetics and Molecular Research**, v. 15, n. 2, 2016.

LIAO, K. C.; GARCIA-BLANCO, M. A. Role of Alternative Splicing in Regulating Host Response to Viral Infection. **Cells 2021, Vol. 10, Page 1720**, v. 10, n. 7, p. 1720, 8 jul. 2021.

LIEN, E. et al. Toll-like Receptor 2 Functions as a Pattern Recognition Receptor for Diverse Bacterial Products. **Journal of Biological Chemistry**, v. 274, n. 47, p. 33419–33425, nov. 1999.

LIMA, C. S. et al. A hybrid swarm of manatees along the Guianas coastline, a peculiar environment under the influence of the Amazon River plume. **Anais da Academia Brasileira de Ciências**, v. 91, 26 ago. 2019.

LIU, C. et al. F-Box Protein 11 Suppresses Cell Proliferation and Aerobic Glycolysis in Glioblastomas by Mediating the Ubiquitin Degradation of Cdc25A. **Journal of Neuropathology & Experimental Neurology**, v. 81, n. 7, p. 511–521, 20 jun. 2022.

LIU, S. Q. et al. TNF- α -induced protein 8-like 2 negatively regulates the immune function of dendritic cells by suppressing autophagy via the TAK1/JNK pathway in septic mice. **Cell Death & Disease** **2021 12:11**, v. 12, n. 11, p. 1–13, 30 out. 2021.

LIU, Y. et al. Zinc Finger Protein 618 Regulates the Function of UHRF2 (Ubiquitin-like with PHD and Ring Finger Domains 2) as a Specific 5-Hydroxymethylcytosine Reader. **Journal of Biological Chemistry**, v. 291, n. 26, p. 13679–13688, jun. 2016.

LONG, M.; DEUTSCH, M. Intron—exon structures of eukaryotic model organisms. **Nucleic Acids Research**, v. 27, n. 15, p. 3219–3228, 1 ago. 1999.

LUNA, F. DE O. et al. Genetic Connectivity of the West Indian Manatee in the Southern Range and Limited Evidence of Hybridization With Amazonian Manatees. **Frontiers in Marine Science**, v. 7, 6 jan. 2021.

LUNA, F. O. et al. Vista do Captura e utilização do peixe-boi marinho (*Trichechus manatus manatus*) no litoral Norte do Brasil. **21**, p. 115–123, 2008.

LUNA, F. O. et al. Phylogeographic implications for release of critically endangered manatee calves rescued in Northeast Brazil. **Aquatic Conservation: Marine and Freshwater Ecosystems**, v. 22, n. 5, p. 665–672, jul. 2012.

LYNCH, K. W. Consequences of regulated pre-mRNA splicing in the immune system. **Nature Reviews Immunology** **2004 4:12**, v. 4, n. 12, p. 931–940, dez. 2004.

MAKI, R. et al. The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. **Cell**, v. 24, n. 2, p. 353–365, 1981.

MARMONTEL, M., DE SOUZA, D. & KENDALL, S. **THE IUCN RED LIST OF THREATENED SPECIES™- African Manatee Assessment, Errata version**. [s.l.: s.n.].

MARTINEZ, N. M.; LYNCH, K. W. Control of Alternative Splicing in Immune Responses: Many Regulators, Many Predictions, Much Still to Learn. **Immunological reviews**, v. 253, n. 1, p. 216, maio 2013.

MATERA, A. G.; WANG, Z. A day in the life of the spliceosome. **Nature Reviews Molecular Cell Biology** 2014 15:2, v. 15, n. 2, p. 108–121, 23 jan. 2014.

MBAREK, H. et al. Genome-Wide Significance for PCLO as a Gene for Major Depressive Disorder. **Twin research and human genetics : the official journal of the International Society for Twin Studies**, v. 20, n. 4, p. 267–270, ago. 2017.

MONIN, L.; GAFFEN, S. L. Interleukin 17 Family Cytokines: Signaling Mechanisms, Biological Activities, and Therapeutic Implications. **Cold Spring Harbor Perspectives in Biology**, v. 10, n. 4, 1 abr. 2018.

MORITZ, C. Uses of Molecular Phylogenies for Conservation on JSTOR. **Philosophical Transactions: Biological Sciences**, v. 349, n. 1327, p. 113–118, 1995.

NAKKA, K. et al. Diversification of the muscle proteome through alternative splicing. **Skeletal Muscle** 2018 8:1, v. 8, n. 1, p. 1–18, 6 mar. 2018.

NAMAVAR, Y. et al. Classification, diagnosis and potential mechanisms in Pontocerebellar Hypoplasia. **Orphanet Journal of Rare Diseases**, v. 6, n. 1, p. 1–14, 12 jul. 2011.

NISHINA, H. Physiological Roles of SAPK/JNK Signaling Pathway. **Journal of Biochemistry**, v. 136, n. 2, p. 123–126, 1 ago. 2004.

NITURE, S. et al. Oncogenic Role of Tumor Necrosis Factor α -Induced Protein 8 (TNFAIP8). **Cells**, v. 8, n. 1, 1 jan. 2019.

OUKKA, M. et al. The Transcription Factor NFAT4 Is Involved in the Generation and Survival of T Cells. **Immunity**, v. 9, n. 3, p. 295–304, set. 1998.

PAGANI, F.; BARALLE, F. E. Genomic variants in exons and introns: identifying the splicing spoilers. **Nature Reviews Genetics** 2004 5:5, v. 5, n. 5, p. 389–396, maio 2004.

PAI, A. A.; LUCA, F. Environmental influences on RNA processing: Biochemical, molecular and genetic regulators of cellular response. **Wiley interdisciplinary reviews. RNA**, v. 10, n. 1, 1 jan. 2019.

PALUSA, S. G.; ALI, G. S.; REDDY, A. S. N. Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. **The Plant Journal**, v. 49, n. 6, p. 1091–1107, 22 fev. 2007.

PARK, E. et al. The Expanding Landscape of Alternative Splicing Variation in Human Populations. **The American Journal of Human Genetics**, v. 102, n. 1, p. 11–26, 4 jan. 2018.

PARK, S. Y.; GUO, X. Adaptor protein complexes and intracellular transport. **Bioscience Reports**, v. 34, n. 4, 1 ago. 2014.

PHILBIN, V. J. et al. Identification and characterization of a functional, alternatively spliced Toll-like receptor 7 (TLR7) and genomic disruption of TLR8 in chickens. **Immunology**, v. 114, n. 4, p. 507–521, abr. 2005.

PICKRELL, J. K. et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. **Nature** **2010 464:7289**, v. 464, n. 7289, p. 768–772, 10 mar. 2010.

PIOVESAN, A. et al. Human protein-coding genes and gene feature statistics in 2019. **BMC Research Notes**, v. 12, n. 1, p. 1–5, 4 jun. 2019.

PIZZOLLA, A. et al. Pattern recognition receptor immunomodulation of innate immunity as a strategy to limit the impact of influenza virus. **Journal of leukocyte biology**, v. 101, n. 4, p. 851–861, abr. 2017.

POIRIER, M.-B. et al. General Receptor for Phosphoinositides 1, a Novel Repressor of Thyroid Hormone Receptor Action that Prevents Deoxyribonucleic Acid Binding. **Molecular Endocrinology**, v. 19, n. 8, p. 1991–2005, 1 ago. 2005.

PREMONT, R. T. et al. The GIT Family of ADP-ribosylation Factor GTPase-activating Proteins: FUNCTIONAL DIVERSITY OF GIT2 THROUGH ALTERNATIVE SPLICING. **Journal of Biological Chemistry**, v. 275, n. 29, p. 22373–22380, 21 jul. 2000.

QUINLAN, A. R.; HALL, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. **Bioinformatics**, v. 26, n. 6, p. 841–842, 15 mar. 2010.

RAAB, G.; KLAGSBRUN, M. Heparin-binding EGF-like growth factor. **Biochimica et Biophysica Acta (BBA) - Reviews on Cancer**, v. 1333, n. 3, p. F179–F199, dez. 1997.

RAKHRA, G.; RAKHRA, G. Zinc finger proteins: insights into the transcriptional and post transcriptional regulation of immune response. **Molecular Biology Reports**, v. 48, n. 7, p. 5735, 1 jul. 2021.

REHMAN, S. U.; TABISH, M. Alternative splicing of ACE2 possibly generates variants that may limit the entry of SARS-CoV-2: A potential therapeutic approach using SSOs. **Clinical Science**, v. 134, n. 10, p. 1143–1150, 1 maio 2020.

ROBERTS, A. et al. Identification of novel transcripts in annotated genomes using RNA-Seq. **Bioinformatics (Oxford, England)**, v. 27, n. 17, p. 2325–2329, set. 2011.

SÁ, A. L. A. DE et al. Novel insights on aquatic mammal MHC evolution: Evidence from manatee DQB diversity. **Developmental and Comparative Immunology**, v. 132, p. 104398, 1 jul. 2022.

SAINI, A. et al. Splice variants of superoxide dismutases in rice and their expression profiles under abiotic stresses. **International Journal of Molecular Sciences**, v. 22, n. 8, p. 3997, 2 abr. 2021.

SAMMETH, M.; FOISSAC, S.; GUIGÓ, R. A general definition and nomenclature for alternative splicing events. **PLoS computational biology**, v. 4, n. 8, ago. 2008.

SATIZÁBAL, P. et al. Phylogeography and sex-biased dispersal across riverine manatee populations (*Trichechus inunguis* and *Trichechus manatus*) in South America. **PloS one**, v. 7, n. 12, p. e52468, 20 jan. 2012.

SCOTTI, M. M.; SWANSON, M. S. RNA mis-splicing in disease. **Nature Reviews Genetics** **2015 17:1**, v. 17, n. 1, p. 19–32, 23 nov. 2015.

SHEN, S. et al. SURVIV for survival analysis of mRNA isoform variation. **Nature Communications**, v. 7, 9 jun. 2016.

SIMON, P. et al. Differential Regulation of Serum- and Glucocorticoid-Inducible Kinase 1 (SGK1) Splice Variants Based on Alternative Initiation of Transcription. **Cellular Physiology and Biochemistry**, v. 20, n. 6, p. 715–728, 2007.

SONG, J. et al. Novel Form of Alternative Splicing of NFKB1. Its Role in Polycythemia and Adaptation to High Altitude in Andean Aymara. **Blood**, v. 132, n. Supplement 1, p. 2316, 29 nov. 2018.

SPRINGER, M. S. et al. Interordinal gene capture, the phylogenetic position of Steller's sea cow based on molecular and morphological data, and the macroevolutionary history of Sirenia. **Molecular Phylogenetics and Evolution**, v. 91, 2015.

SU, Z.; HUANG, D. Alternative Splicing of Pre-mRNA in the Control of Immune Activity. **Genes** **2021, Vol. 12, Page 574**, v. 12, n. 4, p. 574, 15 abr. 2021.

TAKOUKAM, A. K. et al. African manatee (*Trichechus senegalensis*) habitat suitability at Lake Ossa, Cameroon, using trophic state models and predictions of submerged aquatic vegetation. **Ecology and Evolution**, v. 11, n. 21, p. 15212, 1 nov. 2021.

TRAPNELL, C. et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. **Nature Biotechnology** **2010 28:5**, v. 28, n. 5, p. 511–515, 2 maio 2010.

TRESS, M. L.; ABASCAL, F.; VALENCIA, A. Alternative Splicing May Not Be the Key to Proteome Complexity. **Trends in Biochemical Sciences**, v. 42, n. 2, p. 98–110, 1 fev. 2017.

TRINCADO, J. L. et al. The prognostic potential of alternative transcript isoforms across human tumors. **Genome Medicine**, v. 8, n. 1, p. 1–14, 17 ago. 2016.

TU, Y. et al. Tristetraprolin specifically regulates the expression and alternative splicing of immune response genes in HeLa cells. **BMC Immunology**, v. 20, n. 1, p. 1–13, 2 maio 2019.

UNO, K. et al. The Piccolo Intronic Single Nucleotide Polymorphism rs13438494 Regulates Dopamine and Serotonin Uptake and Shows Associations with Dependence-Like Behavior in Genomic Association Study. **Current molecular medicine**, v. 15, n. 3, p. 265–274, 2015.

VIANNA, J. A. et al. Phylogeography, phylogeny and hybridization in trichechid sirenians: Implications for manatee conservation. **Molecular Ecology**, v. 15, p. 433–447, 2006.

VILAÇA, S. T. et al. Manatee genomics supports a special conservation area along the Guianas coastline under the influence of the Amazon River plume. **Estuarine, Coastal and Shelf Science**, v. 226, p. 106286, out. 2019.

VILAÇA, S. T.; SANTOS, F. R. Complete mitochondrial genome of the Florida manatee (*Trichechus manatus latirostris*, Sirenia). **Genetics and Molecular Biology**, v. 42, n. 4, 3 fev. 2020.

VITTING-SEERUP, K. et al. spliceR: an R package for classification of alternative splicing and prediction of coding potential from RNA-seq data. **BMC Bioinformatics**, v. 15, n. 1, p. 81, 23 dez. 2014.

VITTING-SEERUP, K.; SANDELIN, A. IsoformSwitchAnalyzeR: Analysis of changes in genome-wide patterns of alternative splicing and its functional consequences. [s.d.].

WAHL, M. C.; WILL, C. L.; LÜHRMANN, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. **Cell**, v. 136, n. 4, p. 701–718, 20 fev. 2009.

WANG, S. et al. S100A8/A9 in inflammation. **Frontiers in Immunology**, v. 9, n. JUN, p. 1298, 11 jun. 2018.

WANG, Y. et al. Mechanism of alternative splicing and its regulation. **Biomedical Reports**, v. 3, n. 2, p. 152, mar. 2015a.

WANG, Z. et al. The prognostic value of CXC-chemokine receptor 2 (CXCR2) in gastric cancer patients. **BMC Cancer**, v. 15, n. 1, p. 1–8, 23 out. 2015b.

WANG, Z.; BURGE, C. B. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. **RNA**, v. 14, n. 5, p. 802–813, 1 maio 2008.

WEISBROD, T. C. et al. Manatee *Trichechus manatus latirostris* calf mortality in Florida: a retrospective review of pathology data from 2009-2017. **Diseases of Aquatic Organisms**, v. 147, p. 111–126, 16 dez. 2021.

WELLS, C. A. et al. Alternate transcription of the Toll-like receptor signaling cascade. **Genome Biology**, v. 7, n. 2, p. R10, 17 fev. 2006.

WILLENBERG, H. S. et al. Role of Interleukin-6 in Stress Response in Normal and Tumorous Adrenal Cells and during Chronic Inflammation. **Annals of the New York Academy of Sciences**, v. 966, n. 1, p. 304–314, jun. 2002.

WOUDSTRA, S. et al. Modulatory effects of the piccolo genotype on emotional memory in health and depression. **PloS one**, v. 8, n. 4, p. e61494, 2013.

XIN, R. et al. isoCirc catalogs full-length circular RNA isoforms in human transcriptomes. **Nature Communications** 2021 12:1, v. 12, n. 1, p. 1–11, 12 jan. 2021.

XING, Y. et al. An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. **Nucleic Acids Research**, v. 34, n. 10, p. 3150–3160, 1 maio 2006.

XIONG, H. Y. et al. The human splicing code reveals new insights into the genetic determinants of disease. **Science**, v. 347, n. 6218, 9 jan. 2015.

XIONG, Y. et al. Profiles of alternative splicing in colorectal cancer and their clinical significance: A study based on large-scale sequencing data. **EBioMedicine**, v. 36, p. 183, 1 out. 2018.

YU, L. P.; SHEEHY, M. J. The cryptic HLA-DQA2 (“DX alpha”) gene is expressed in human B cell lines. **The Journal of Immunology**, v. 147, n. 12, 1991.

ZHANG, C. et al. Deletion and Down-Regulation of HRH4 Gene in Gastric Carcinomas: A Potential Correlation with Tumor Progression. **PLoS ONE**, v. 7, n. 2, p. e31207, 20 fev. 2012.

ZHANG, K.; SAXON, A.; MAX, E. E. Two unusual forms of human immunoglobulin E encoded by alternative RNA splicing of epsilon heavy chain membrane exons. **The Journal of experimental medicine**, v. 176, n. 1, p. 233–243, 1 jul. 1992.

ZHANG, W.; LIU, H. T. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. **Cell Research**, v. 12, n. 1, p. 9–18, mar. 2002.

ZHANG, Y. et al. Alternative splicing and cancer: a systematic review. **Signal Transduction and Targeted Therapy** 2021 6:1, v. 6, n. 1, p. 1–14, 24 fev. 2021.