

Genômica estrutural, funcional e comparativa da tilápia do Nilo (*Oreochromis niloticus*)

Structural, functional and comparative genomics of Nile tilapia (*Oreochromis niloticus*)

Genómica estructural, funcional y comparada de la tilapia del Nilo (*Oreochromis niloticus*)

Laura Bueno¹
Altino Choupina²

Resumo: A tilápia do Nilo destaca-se na aquacultura devido à sua adaptabilidade a ambientes de água doce e salobra. O avanço significativo na área da genômica culminou na sequenciação e anotação do genoma de *Oreochromis niloticus*, submetido pela Universidade de Maryland no ano de 2019. O presente estudo tem como objetivo apresentar aspectos estruturais, funcionais e comparativos do genoma desta espécie, através da utilização de bancos de dados genômicos, tendo como foco o gene do hormônio do crescimento (*GH*). Estudos associativos genômicos permitiram a identificação de outros cinco genes relacionados ao crescimento. A análise filogenética revelou a evolução deste gene em diversas espécies, enquanto a genômica funcional e comparativa apontam para a necessidade de maior caracterização das relações entre genes.

Palavras-chave: Genômica. *GH*. *Oreochromis niloticus*. Tilápia do Nilo.

Abstract: Nile tilapia stands out in aquaculture due to its adaptability to freshwater and brackish water environments. The significant advance in the area of genomics culminated in the sequencing and annotation of the genome of *Oreochromis niloticus*, submitted by the University of Maryland in 2019. The present study aims to present structural, functional and comparative aspects of the genome of this species, through the use of genomic databases, focusing on the growth hormone (*GH*) gene. Genomic associative studies allowed the identification of another five genes related to growth. Phylogenetic analysis revealed the evolution of this gene in several species, while functional and comparative genomics point to the need for greater characterization of the relationships between genes.

Keywords: Genomics. *GH*. *Oreochromis niloticus*. Nile tilapia.

Resumen: La tilapia del Nilo se destaca en la acuicultura debido a su adaptabilidad a ambientes de agua dulce y salobre. El importante avance en el área de genómica culminó con la secuenciación y anotación del genoma de *Oreochromis niloticus*, presentado por la Universidad de Maryland en 2019. El presente estudio tiene como objetivo presentar aspectos estructurales, funcionales y comparativos del genoma de esta especie, mediante el uso de bases de datos genómicas, centrándose en el gen de la hormona del crecimiento (*GH*). Los estudios genómicos asociativos permitieron identificar otros cinco genes relacionados con el crecimiento. Los análisis filogenéticos revelaron la evolución de este gen en varias especies, mientras que la genómica funcional y comparada apuntan a la necesidad de una mayor caracterización de las relaciones entre genes.

Palabras-clave: Genómica. *GH*. *Oreochromis niloticus*. Tilapia del Nilo.

Submetido 04/01/2024

Aceito 20/02/2024

Publicado 29/02/2024

¹ Graduanda em Biologia e Biotecnologia. Instituto Politécnico de Bragança, Portugal. ORCID: <https://orcid.org/0009-0004-1314-8853>. E-mail: a44361@alunos.ipb.pt

² Professor Doutor em Biotecnologia. Centro de Investigação de Montanha (CIMO), Instituto Politécnico de Bragança, Portugal. ORCID: <https://orcid.org/0000-0002-3956-9398>. E-mail: albracho@ipb.pt

Introdução

A tilápia do Nilo (*Oreochromis niloticus*), nomeada por Linnaeus em 1758, é uma espécie de peixe que se insere na ordem Cichliformes e na família Cichlidae. A versatilidade de adaptação da tilápia a ambientes de água doce e salobra a tornou uma espécie de destaque na aquicultura, resultando em uma produção global estimada em cerca de 4.407,2 mil toneladas métricas em peso vivo no ano de 2020, representando 9% da produção total de aquicultura no cenário mundial naquele ano (FAO, 2022). Dentre os continentes, a Ásia destacou-se como principal polo produtor, contribuindo com aproximadamente 60% da produção global, seguida pela África, que representou aproximadamente 30%, e pelas Américas, com uma parcela de aproximadamente 10% (FAO, 2024).

A área nativa da tilápia do Nilo abrange regiões tropicais e subtropicais da África e do Oriente Médio. Essa espécie possui uma ampla distribuição dentro das bacias dos rios Nilo e Níger e pode ser encontrada nos grandes lagos de Tanganica, Alberto, Eduardo e Jorge, bem como em numerosos pequenos afluentes e lagos em toda a África ocidental e oriental. Além disso, a tilápia estendeu seu habitat para o Oriente Médio, com populações vivendo no Rio Yarkon, em Israel (Nico; Schofield; Neilson, 2013).

Essa espécie foi inicialmente introduzida nos Estados Unidos e outras regiões com o objetivo de ser utilizada na aquicultura para fins de produção de alimentos ou outras finalidades, porém, acabou sendo introduzida em águas abertas. Segundo Nico, Schofield, Neilson (2013), isso pode ter ocorrido devido a fugas acidentais dos tanques ou à liberação deliberada desses peixes em ambientes naturais. Dessa forma, a tilápia do Nilo passou a viver em águas livres, fora do ambiente controlado das fazendas de piscicultura, o que a expôs a ecossistemas naturais, rios, lagos ou outros corpos d'água. Como resultado, este processo acabou por desencadear impactos ecológicos significativos, incluindo a competição com espécies nativas, a alteração dos ecossistemas e um potencial impacto na biodiversidade local.

Além de seu impacto econômico na indústria da aquicultura, as espécies de tilápia africanas, em particular a tilápia do Nilo (*Oreochromis niloticus*), tornaram-se objeto de estudo relevante nas áreas de biologia de pesca, genética e genômica devido às suas notáveis características (Amoussou et al., 2019). Segundo Guyon et al. (2012), as tilápias têm se destacado na ciência por possuírem características como: tempo de geração curto, tamanho grande suficiente para estudos fisiológicos e podem ser facilmente criadas. Além disso, as

tilápias constituem modelos ótimos para a genômica ambiental, uma vez que permitem a análise das interações entre o genoma e meio ambiente e as respostas adaptativas aos estresses ambientais nessa espécie (Cossins; Crawford, 2005). Dessa forma, o estudo da tilápia do Nilo pode contribuir para a compreensão dos mecanismos evolutivos, dado que a tilápia está intimamente relacionada com os peixes ciclídeos dos Grandes Lagos da África Oriental que sofreram radiação. As 2.000 espécies de ciclídeos nesses lagos representam uma coleção de mutantes naturais que podem fornecer informações sobre os mecanismos de especiação e adaptação desta espécie (Kocher, 2004).

O genoma de *Oreochromis niloticus* foi sequenciado pela Universidade de Maryland e submetido na plataforma Ensembl em junho de 2018, apresentando um tamanho de mais de um bilhão de pares de bases. A anotação genética do genoma da tilápia do Nilo, realizada no âmbito do projeto Ensembl, teve início em julho de 2019 e foi concluída em dezembro do mesmo ano. O processo de anotação genética foi realizado com a utilização de uma combinação de alinhamentos proteína-genoma, mapeamento de anotação de uma espécie de referência e alinhamentos de sequenciamento de RNA (Martin et al., 2023).

Portanto, o objetivo deste artigo é de apresentar uma análise abrangente da genômica da tilápia do Nilo, explorando minuciosamente aspectos estruturais, funcionais e comparativos. A investigação desses elementos se revela de importância crucial para a compreensão da complexidade genômica desta espécie e seu profundo impacto na notável capacidade de adaptação, bem como em sua evolução e interações com o ambiente.

Metodologia

Para a condução da pesquisa e o estudo dos genes, os dados do genoma da tilápia do Nilo (*Oreochromis niloticus*) foram adquiridos a partir de fontes confiáveis de referência, com ênfase no navegador de genoma “Ensembl” (<https://www.ensembl.org>), uma plataforma especializada em genomas de vertebrados que produz recursos genômicos de alta qualidade, oferecendo suporte para pesquisas em genômica comparativa, evolução, variação de sequência e regulação transcricional. A plataforma Ensembl é um projeto que foi desenvolvido no ano de 1999, alguns anos antes de o esboço do genoma humano ser concluído, pelo EMBL's European Bioinformatics Institute (EMBL-EBI), o Laboratório Europeu de Biologia Molecular do Instituto Europeu de Bioinformática (Martin et al., 2023).

Adicionalmente, empregou-se a plataforma do “National Center for Biotechnology Information (NCBI)” (<https://www.ncbi.nlm.nih.gov/>), ou Centro Nacional de Informação sobre Biotecnologia. Criado em 1988, o NCBI tem como missão desenvolver sistemas de informação para a biologia molecular e hospeda uma extensa variedade de bancos de dados cruciais para os campos de biotecnologia e biomedicina, sendo reconhecido como uma fonte renomada de ferramentas e serviços de bioinformática (Sayers et al., 2022).

Outras informações relevantes para o desenvolvimento deste estudo foram acessadas em plataformas adicionais às referidas anteriormente, sendo que cada uma delas desempenharam um papel substancial no presente artigo. O banco de dados “UniProt Knowledgebase (UniProtKB)” (<https://www.uniprot.org/>) foi acessado para obtenção de sequências e informações funcionais detalhadas de proteínas anotadas (Bateman et al., 2023). Ademais, a plataforma “HUGO Gene Nomenclature Committee at the European Bioinformatics Institute (HGNC)” (<https://www.genenames.org/>) foi utilizada na busca de símbolos e nomenclatura de genes humanos.

Por fim, o genoma da tilápia do Nilo (*Oreochromis niloticus*), empregada no presente estudo, encontra-se devidamente depositada nas plataformas supracitadas, sendo referenciada como “O_niloticus_UMD_NMBU”. No NCBI, o código de acesso para o genoma “O_niloticus_UMD_NMBU” é encontrado como “GCF_001858045.2”. Na plataforma Ensembl o genoma “O_niloticus_UMD_NMBU” é encontrado com o código de acesso “GCA_001858045.3”. Esta abordagem possibilitou assegurar a solidez dos dados utilizados e a conformidade com os padrões reconhecidos de genômica.

Resultados e discussão

Genômica estrutural

As informações que se seguem, referentes à composição gênica do genoma de *Oreochromis niloticus*, foram obtidas através da plataforma Ensembl (Martin et al., 2023).

Atualmente, o genoma da tilápia do Nilo (*Oreochromis niloticus*) é bem caracterizado, com um tamanho de aproximadamente 1.005.681.550 pares de bases. A montagem genômica de *Oreochromis niloticus*, conduzida pelo projeto Ensembl do EMBL-EBI, foi realizada à nível cromossômico. Esse processo envolveu a organização de 3.010 longas sequências contínuas,

ou *contigs*, agrupadas em 2.460 *scaffolds*, resultando na identificação de 22 cromossomos no genoma de *O. niloticus*.

A métrica N50, fundamental para avaliar a qualidade da montagem, representa o tamanho no qual 50% do genoma montado está contido em blocos do tamanho N50 ou maior. Para os *contigs*, o tamanho N50 é de 2.923.640 pares de bases (2,9 MB), enquanto para os *scaffolds* o tamanho N50 é de 38.839.487 pares de bases (38,8 MB). Adicionalmente, a composição percentual de GC, que abrange Guaninas e Citosinas, constitui 40,5% do genoma.

O genoma da tilápia do Nilo contém aproximadamente 33.162 genes, incluindo genes codificantes (genes que contêm fase de leitura aberta – ORF) e não codificantes. Existem 29.189 genes codificantes no genoma de *O. niloticus*, que são responsáveis pela síntese de proteínas. Além disso, o genoma é composto por 4.973 genes não codificantes. Estes genes não codificantes podem desempenhar papéis cruciais na regulação gênica e em outros processos celulares. Dentro dos genes não codificantes, há uma subdivisão em pequenos genes não codificantes (2.346) e genes longos não codificantes (2.581). Os pequenos genes não codificantes são geralmente conhecidos por seu papel na regulação gênica de processos específicos, enquanto os genes longos não codificantes podem ter funções mais complexas e diversas.

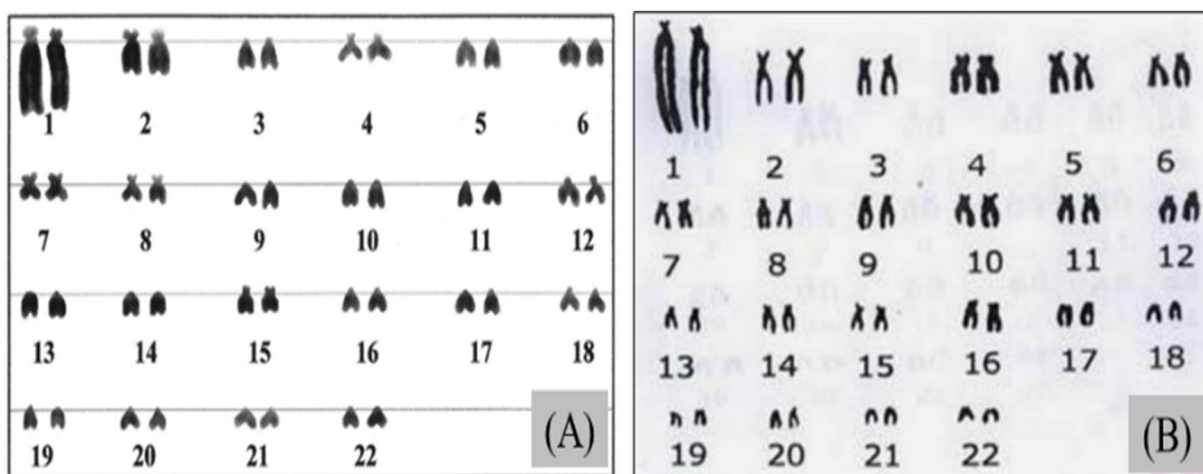
Ademais, o genoma da tilápia do Nilo contém 275 pseudogenes, que são genes que possuem homologia com genes codificantes de proteínas conhecidos, mas contêm um deslocamento de quadro e/ou códon(s) de parada que interrompe a ORF. Eles podem representar genes que sofreram mutações e perderam sua função ao longo do tempo. No total, foram identificadas 82.099 transcrições de genes no genoma da tilápia do Nilo (*Oreochromis niloticus*), o que indica uma complexa rede de regulação e expressão gênica nessa espécie.

Caracterização citogenética

Os autores Supiwong et al. (2013) e Sofy, Layla, Iman (2008) descrevem o cariótipo da tilápia do Nilo. Segundo Supiwong et al. (2013), o perfil cromossômico da tilápia do Nilo é composto por 2 cromossomos grandes acrocêntricos, 2 pequenos submetacêntricos, 10 pequenos acrocêntricos e 30 pequenos telocêntricos (Figura 1-A). Por outro lado, Sofy, Layla, Iman (2008) descrevem o cariótipo da tilápia do Nilo como sendo composto por 2 cromossomos submetacêntricos, 24 subtelocêntricos e 18 telocêntricos (Figura 1-B).

Em relação aos cromossomos sexuais, na tilápia do Nilo foram identificados loci distintos determinantes do sexo XX/XY tanto no grupo de ligação 1 (LG) quanto no LG23 (Eshel et al., 2012). A determinação do sexo em tilápias é em grande parte genética, embora fatores ambientais também possam influenciar (Bull, 1985).

Figura 1 - (A) Cariótipo da tilápia do Nilo descrito em 2013 e (B) Cariótipo da tilápia do Nilo descrito em 2008.



Fonte: (Supiwong et al., 2013; Sofy; Layla; Iman, 2008).

A caracterização do cariótipo da tilápia do Nilo permanece como um campo em evolução, especialmente considerando que o sequenciamento do genoma da espécie foi realizado apenas em 2019. Dada a relativa recente disponibilidade dessas informações genômicas, é compreensível que a caracterização cromossômica ainda apresente variações e inconsistências entre estudos, como observado nos trabalhos de Supiwong et al. (2013) e Sofy, Layla, Iman (2008).

Caracterização de cromossomos individuais

A análise de cromossomos individuais revela que o grupo de ligação (LG) LG3 é o maior no genoma de *O. niloticus*, com 87.567.345 pares de bases e uma porcentagem de GC de 40,5%. Este grupo de ligação inclui 2.055 genes codificantes de proteínas, 470 genes não codificantes (63 pequenos genes e 407 longos genes não codificantes) e 28 pseudogenes (Martin et al., 2023).

Estudos citogenéticos também indicam que LG7 é o segundo maior grupo de ligação no genoma de *O. niloticus*, com 64.772.279 pares de bases e 40% de conteúdo GC. É composto por 2.082 genes codificantes, 224 genes não codificantes, dentre eles 105 genes pequenos e 119 genes longos, e 14 pseudogenes (Martin et al., 2023).

Por fim, o menor grupo de ligação, LG8, possui 30.527.416 pares de bases e uma porcentagem de GC de 40,5%. Neste cromossomo, foram identificados 993 genes codificantes de proteínas, 80 genes não codificantes (22 pequenos genes e 58 longos genes não codificantes) e 7 pseudogenes (Martin et al., 2023).

Genômica funcional

A genômica funcional desempenha um papel fundamental na compreensão das funções genéticas e das interações entre genes no contexto das redes biológicas que controlam características fenotípicas (Mutch; Wahli; Williamson, 2005). No caso da tilápia do Nilo (*Oreochromis niloticus*), o estudo da genômica funcional ainda está em desenvolvimento devido à recente conclusão da sequenciação de seu genoma em 2019. A escassez de estudos que identifiquem genes de interesse e elucidem suas funções específicas reflete diretamente na limitação da genômica funcional nesta espécie. Portanto, a disponibilidade de dados abrangentes e a compreensão dos mecanismos moleculares subjacentes a características específicas ainda apresentam vasto potencial de descobertas na genômica da tilápia do Nilo.

Caracterização do Gene do Hormônio do Crescimento (*GHI*) na tilápia do Nilo

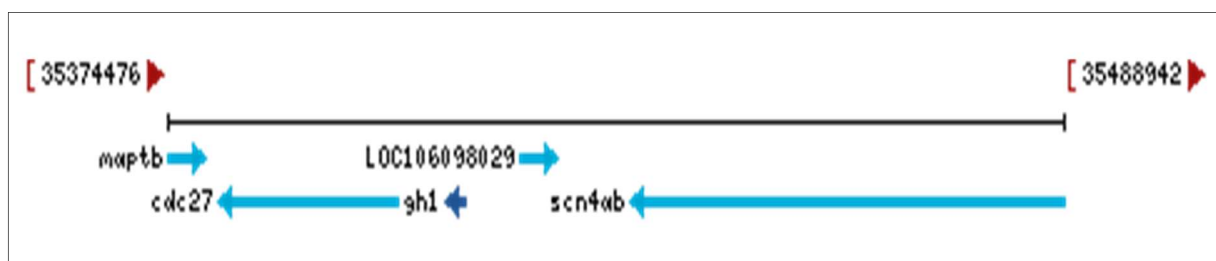
O genoma da tilápia tem sido alvo de análises minuciosas, o que permitiu a identificação de genes de interesse associados a características altamente relevantes, como o crescimento. Um dos genes que tem despertado considerável atenção nos estudos recentes sobre a tilápia do Nilo é o gene do hormônio do crescimento (*GHI*). Este hormônio é um polipeptídeo sintetizado nas células somatotrópicas hipofisárias, que, junto com a prolactina e a somatolactina, constitui a família de hormônios hipofisários com estrutura semelhante e função que parecem ter originado de um gene ancestral comum antes da evolução dos peixes (Kamenskaya; Brykov, 2020).

O aprimoramento das taxas de crescimento natural de peixes tem sido extensivamente investigado na aquicultura, proporcionando avanços significativos na eficiência da produção

animal, aperfeiçoamento da nutrição e a aplicação de estratégias avançadas de seleção genética (Fjalestad; Moen; Gomez-Raya, 2003). Segundo Maclean et al. (2002), os resultados mais impressionantes em peixes transgênicos só foram publicados em 1994, quando Devlin et al. (1994) publicaram os resultados de seus experimentos em salmão do Pacífico. Nos salmonídeos, utilizando genes heterólogos de hormônio de crescimento unidos a fortes promotores de origem de peixe, os autores demonstraram que, mesmo nos peixes do grupo G0 (geração inicial de peixes geneticamente modificados), houve um aumento de mais de 10 vezes na taxa de crescimento em comparação com irmãos não transgênicos, que não receberam os genes modificados.

De acordo com a pesquisa conduzida por Sánchez-Ramos et al. (2006), a compreensão completa dos genes que influenciam as características de crescimento permanece em grande parte desconhecida. No entanto, já foram identificados vários genes candidatos, incluindo o gene do hormônio do crescimento (*GHI*), prolactina (*PRL*), somatolactina (*SL*), fator de crescimento semelhante à insulina (*IGF*) e miostatina (*MSTN*). O símbolo oficial do gene do hormônio do crescimento foi determinado *GHI* (*growth hormone*) pelo HGNC (HUGO Gene Nomenclature Committee at the European Bioinformatics Institute) mas também pode ser referido como *GH*, *GHN* OU *hGH-N*. O gene do hormônio do crescimento (*GHI*) na tilápia do Nilo (*Oreochromis niloticus*) está localizado no cromossomo LG4 com orientação NC_031969.2 (35409936-35411673) no banco de dados do NCBI (Figura 2).

Figura 2 - Localização e orientação do gene *GHI* em *Oreochromis niloticus*, indicada pela seta em azul escuro.

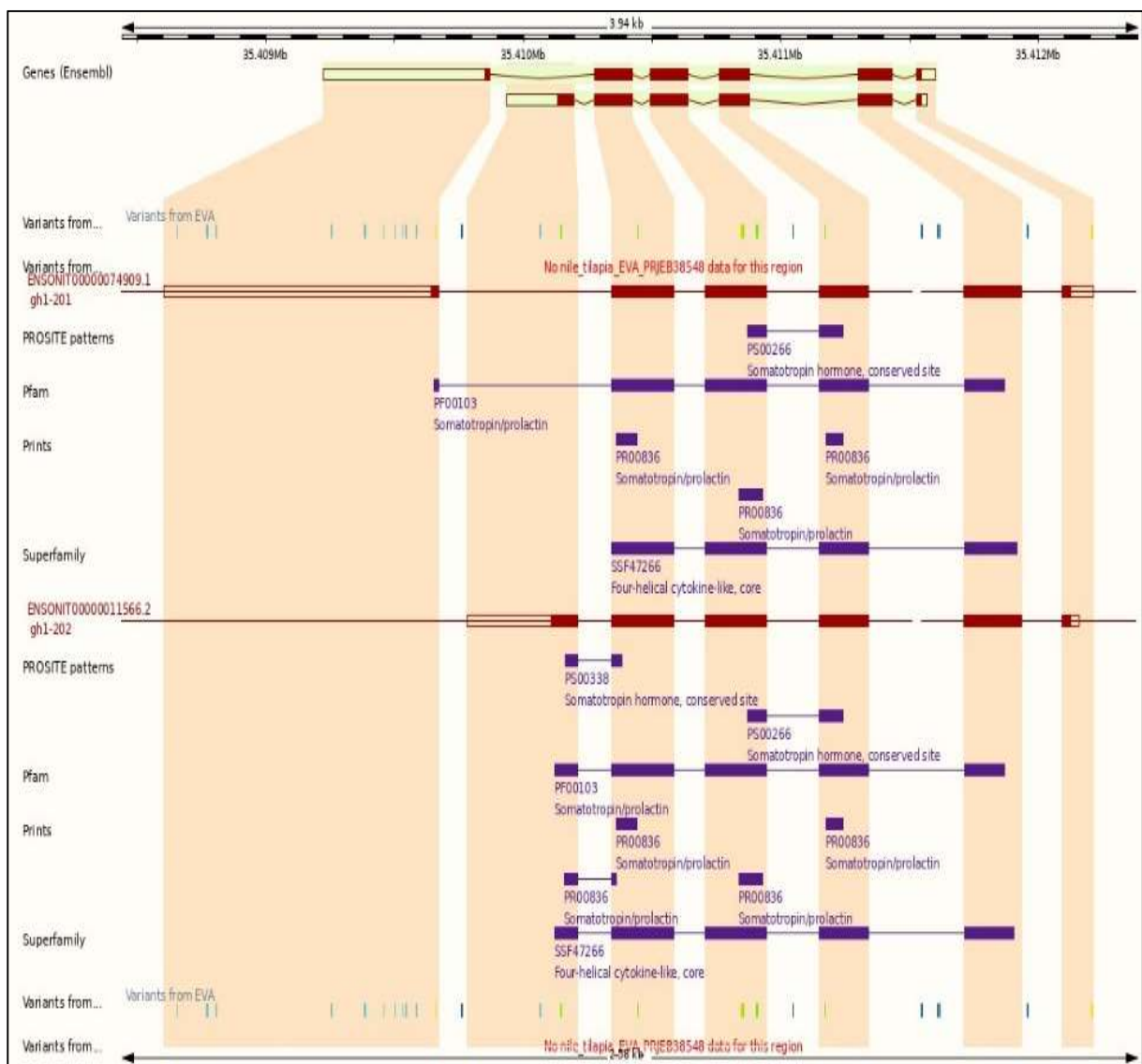


Fonte: NCBI (Sayers et al., 2022).

O *GHI* possui um papel importante em todos os vertebrados, contudo, os efeitos fisiológicos deste hormônio resultam da regulação transcricional deste gene. Na tilápia do Nilo,

Ber e Daniel (1993) relatam a presença de duas transcrições do gene *GHI* no genoma de *O. niloticus*. Os autores ainda afirmam que, com base nos dados obtidos, a presença de duas cópias do gene *GH1* é uma consequência de um evento de duplicação que ocorreu após a separação dos salmonídeos. Na Figura 3 estão representadas as duas cópias do gene *GHI* na tilápia do Nilo, sendo elas *GHI-202* e *GHI-201*. As linhas verticais destacam as posições dos éxons, facilitando a comparação entre as transcrições.

Figura 3 - Transcrições do gene *GHI* em *O. niloticus*, representadas em vermelho por caixas (éxons) e linhas (íntrons).



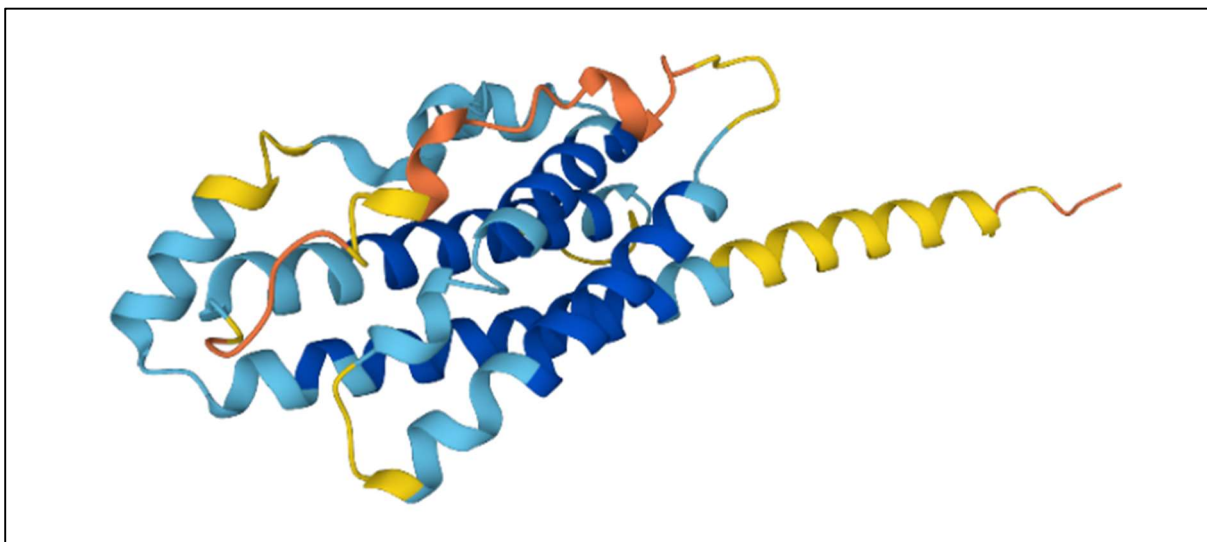
Fonte: Ensembl versão 111 (Martin et al., 2023).

Segundo Kamenskaya e Brykov (2020), o aumento no número de cópias de um mesmo gene está relacionado com a duplicação de todo o genoma, fazendo com que algumas espécies se tornem tetraploides e seu genoma contenha duas cópias de cada gene (Kamenskaya; Brykov, 2020).

A transcrição *GHI-202* possui um tamanho de 841 pares de bases e 6 éxons e codifica a proteína somatotropina (Figura 4) do organismo *Oreochromis niloticus*. A somatotropina na tilápia do Nilo é predominantemente encontrada no espaço extracelular, onde atua como um hormônio de crescimento. Esta proteína exerce sua função estimulando o crescimento do tecido adiposo e está envolvida na regulação positiva do processo de catabolismo lipídico, contribuindo para o controle dos níveis de gordura no corpo (Bateman et al., 2023).

Ademais, a somatotropina possui um peso molecular de 23.409,60 g/mol, com 207 aminoácidos em sua estrutura, tendo um peso médio por resíduo de aminoácido de 113,09 g/mol. Sua carga é de -0,5 e seu ponto isoelétrico é de 6,33.

Figura 4 - Estrutura tridimensional da somatotropina prevista por AlphaFold.



Fonte: Ensembl versão 111 (Martin et al., 2023).

A segunda cópia do gene *GHI* é o *GHI-201*, também com 6 éxons, porém com um tamanho de 1.258 pares de bases. Esta transcrição também codifica a proteína somatotropina, porém, sua sequência é composta por 192 aminoácidos e cada resíduo de aminoácido tem um peso médio de 113,68 g/mol. Sua carga líquida é positiva, com um valor de 1,0. O ponto

isoelétrico dessa proteína é calculado em 6,75 e possui um peso molecular de aproximadamente 21.826,81 g/mol.

Uma série de outros genes também se relacionam com a taxa de crescimento e desenvolvimento (Wood et al., 2014). Essa associação é reforçada pelos autores Yu et al. (2021), no qual realizaram um estudo utilizando dois grupos compostos por 3.140 peixes que foram criados em lagoas aeradas (normóxia) ou não aeradas (hipóxia noturna). Durante o crescimento, os peixes foram amostrados cinco vezes para determinar os ganhos individuais de peso corporal (PC) com o objetivo de desvendar a arquitetura genômica associada à variação fenotípica na tilápia do Nilo durante adaptação à hipóxia (deficiência de oxigênio) ou normóxia (níveis adequados de oxigênio) e para elucidar o efeito da falta de oxigênio na regulação genética do crescimento.

Os autores ainda realizaram um estudo de associação genômica ampla para identificar SNPs e genes associados aos ambientes hipóxicos e normóxicos na 16ª geração de uma população de tilápias geneticamente melhoradas. Como resultado, Yu et al. (2021) identificaram cinco genes principais que desempenham um papel importante no crescimento e desenvolvimento durante a fase inicial de crescimento de tilápias, sendo eles: *raraa*, *rarab*, *med24*, *brms11a* e *prpf38b*.

Os genes *raraa*, *rarab*, *med24* e *brms11a*, além de serem encontrados em humanos, peixes e camungondos (Hale et al., 2006; Harrison; Wabbersen; Sheperd, 2014) podem também ser encontrados no genoma da tilápia do Nilo (Tabela 1). O gene *Rara* é um gene codificante do receptor alfa do ácido retinóico, um fator de transcrição que desempenha um papel crucial na regulação de genes associados ao crescimento e diferenciação celular (Wang et al., 2017). O mediador da subunidade 24 de transcrição da RNA polimerase II (*med24*) foi identificado como participante do desenvolvimento do sistema nervoso (Harrison; Wabbersen; Sheperd, 2014). Por último, o gene *brms11*, também referido como *brms1*, é um gene codificante do supressor de metástase de câncer de mama 1, sendo um gene de interesse com alta relevância na medicina humana (Zimmermann; Welch, 2020). Além disso, é um gene que apresenta elevada expressão nos testículos e ovários da tilápia do Nilo (Sayers et al., 2022).

Adicionalmente, para o estágio posterior de crescimento, Yu et al. (2021) identificaram 12 regiões sugestivas de marcação de SNPs contendo 22 genes candidatos. Dentro desse conjunto de genes, destaca-se o gene que codifica o uniportador de cálcio mitocondrial (*mcu*),

previamente identificado como um regulador do crescimento e da homeostase do músculo esquelético (Kwong et al., 2018). Outros genes que poderiam estar relacionados com o estágio posterior de desenvolvimento também foram identificados, sendo eles os genes que codificam o transcrito 3 induzido por oncoproteína (*oit3*) e o domínio 1 contendo MAP 6 (*map 3d1*), ambos relacionados com a atividade de ligação ao íon cálcio (Sprague et al., 2006).

Tabela 1 – Principais genes associados ao crescimento e desenvolvimento durante a fase inicial de crescimento na tilápia do Nilo (*Oreochromis niloticus*).

Gene ³	LG ⁴	Posição ⁵	Proteína codificada	Função
<i>raraa</i>	LG8	NC_03197 3.2 (19808860-19949555)	Receptor do ácido retinóico, alfa-a	Codifica o receptor alfa do ácido retinóico, um fator de transcrição que regula genes envolvidos no crescimento e diferenciação celular
<i>rarab</i>	LG4	NC_03196 9.2 (4187766-4216923)	Receptor do ácido retinóico, alfa-b	Codifica o receptor alfa do ácido retinóico, um fator de transcrição que regula genes envolvidos no crescimento e diferenciação celular
<i>med24</i>	LG8	NC_03197 3.2 (19698335-19711822)	Subunidade do complexo mediador 24	Participa do desenvolvimento do sistema nervoso
<i>brms1la</i>	LG19	NC_03198 3.2 (25136031-25141814)	<i>BRMS1</i> como repressor transcrricional a	Atua como um repressor da transcrição, regulando a expressão gênica, está envolvido nos processos biológicos de transcrição e regulação da transcrição

Fonte: NCBI (Sayers et al., 2022).

Genômica comparativa

Após a realização de uma análise da genômica comparativa do gene *GHI*, foi possível obter uma árvore filogenética para este gene. A análise de uma árvore filogenética permite

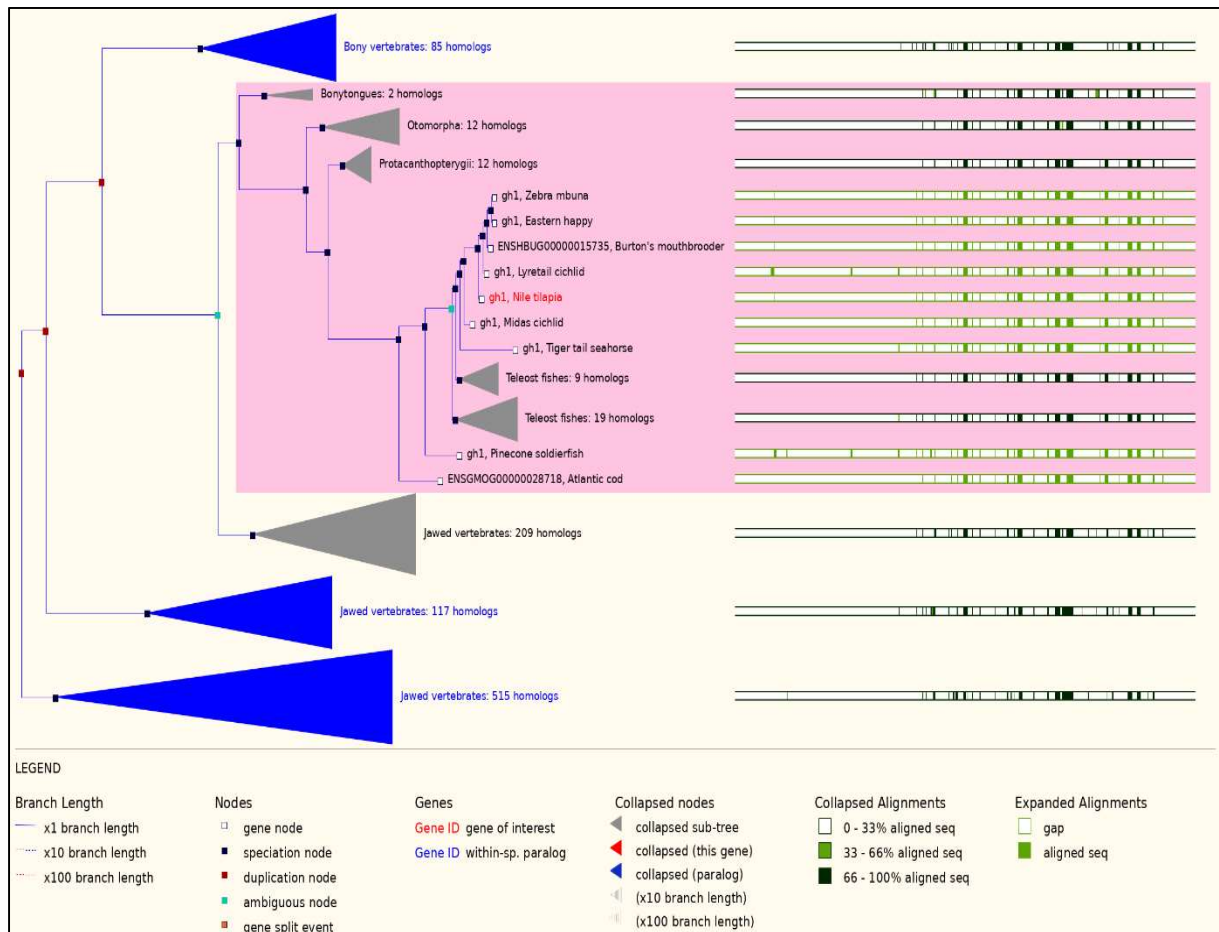
³ Genes associados ao crescimento e desenvolvimento durante a fase inicial de crescimento em *O. niloticus*_UMD_NMBU como genoma de referência para *Oreochromis niloticus*

⁴ Grupo de Ligação

⁵ Em pares de bases (pb)

estudar os processos evolutivos das espécies ao distinguir os genes ortólogos e parálogos que estão diretamente relacionados com os processos de especiação e duplicação durante a evolução dos genomas (Chauve; Doyon; El-mabrouk, 2008). A árvore filogenética para o gene *GHI* é composta por 1007 genes, apresentando 812 nós de especiação, 170 nós de duplicação e 24 genes com informações ambíguas. Os quadrados vermelhos representam os nós de duplicação e os quadrados azuis representam os nós de especiação, dando origem aos genes parálogos e ortólogos. Outra classe de nó, ambígua, é mostrada como um quadrado azul mais claro (Figura 5).

Figura 5 - Árvore filogenética do gene *GHI*, destacado em vermelho.



Fonte: Ensembl versão 111 (MARTIN et al., 2023).

A análise da árvore filogenética construída para o gene *GHI* revela uma história evolutiva notável deste gene em diversas espécies. Dentro dessa árvore, destacam-se espécies

notáveis, como: Midas Cichlid (*Amphilophus citrinellus*), Lyretail cichlid (*Neolamprologus brichardi*), Burton's mouthbrooder (*Astatotilapia burtoni*), Eastern happy (*Astatotilapia calliptera*) e Zebra mbuna (*Maylandia zebra*), que compartilham homologies e características semelhantes ao gene *GHI* da tilápia do Nilo.

Todas essas espécies fazem parte da família Cichlidae e surgiram por meio de eventos de especiação. Os nós de especiação representam momentos cruciais na árvore filogenética, marcando o momento em que ancestrais comuns se dividiram em diferentes populações, originando espécies independentes (Martin et al., 2023). O processo de especiação é evidenciado pela divergência genética ao longo do tempo e pela adaptação a ambientes específicos, resultando em características distintas e na formação de espécies geneticamente isoladas (Campbell; Poelstra; Yoder, 2018).

Os ramos desta árvore filogenética ilustram a evolução do gene *GHI* ao longo das eras, refletindo as relações entre os alelos do gene nas diferentes espécies. Isso nos permite compreender a diversificação deste gene em várias linhagens, bem como sua conservação ou divergência ao longo da evolução. Além disso, a árvore filogenética destaca a conservação genética em algumas espécies, sugerindo a importância do gene *GHI* em suas biologies. Paralelamente, evidencia a divergência genética em outros ramos, indicando adaptações genéticas específicas em diferentes ambientes ou funções desempenhadas pela somatotropina em cada espécie.

O gene *GHI* possui 272 ortólogos, que evoluíram por eventos de especiação em diferentes espécies, mantendo funções semelhantes (Chauve; Doyon; El-mabrouk, 2008). Além disso, este gene possui 3 parálogos (Tabela 2), ou seja, genes homólogos que evoluíram por duplicação genética dentro da espécie *O. niloticus* e adquiriram funções diferentes ao longo do tempo (Chauve; Doyon; El-mabrouk, 2008). As duplicações podem ser consideradas os eventos primários da evolução do genoma e foram bem reconhecidas na era pré-genômica (Koonin, 2005). Tais duplicações podem ser observadas pelos nós vermelhos na árvore filogenética do gene *GHI* (Figura 5).

Tabela 2 – Parálogos do gene *GHI* na tilápia do Nilo (*Oreochromis niloticus*).

Gene parálogo	Descrição	Localização	Taxonomia do ancestral
<i>smtla</i>	Somatolactina	LG14:15,838,823-15,842,586:1	Vertebrados com mandíbula (<i>Gnathostomata</i>)
<i>prl</i>	Prolactina <i>O.niloticus</i>	LG4:5,664,900-5,685,911:-1	Vertebrados com mandíbula (<i>Gnathostomata</i>)
<i>prl2</i>	Semelhante à prolactina <i>O.niloticus</i>	LG17:490,449-494,489:1	Vertebrados com mandíbula (<i>Gnathostomata</i>)

Fonte: Ensembl versão 111 (Martin et al., 2023).

Os genes parálogos do *GHI* em *O. niloticus* surgiram por meio de eventos de duplicação genética em um ancestral dos vertebrados com mandíbula (*Gnathostomata*). Após os eventos de duplicação, passaram a codificar proteínas como a somatolactina e prolactina, que estão envolvidas nos processos de crescimento, desenvolvimento e de transdução de sinal, participam da atividade hormonal e estão localizadas em região extracelular (Sayers et al., 2022).

Os autores Kawauchi et al. (2002) identificaram pela primeira vez o gene *GH* em lampreia marinha sem mandíbula (*Petromyzon marinus*), uma espécie representante de um dos dois grupos de vertebrados mais antigos existentes. Esta descoberta foi vista como evidência de que as duplicações genéticas que deram origem ao *GH* (hormônio do crescimento), *SL* (somatolactina) e *PRL* (prolactina) ocorreram em um ancestral dos gnatostomados. Além disso, os autores sugerem que o gene *GH* é o precursor da família *GH* e que seu gene foi duplicado durante a evolução inicial dos gnatostomados para formar a somatolactina e a prolactina, podendo ter contribuído para a expansão dos vertebrados para novos ambientes.

Diversos estudos têm apontado para outros aspectos relevantes da genômica comparativa relacionados aos genes estudados neste trabalho. De acordo com Hale et al. (2006), os genes *raraa* e *rarab* também desempenham papéis fundamentais no processo de desenvolvimento do peixe-zebra (*Danio rerio*). Além disso, Harrison, Wabbersen, Sheperd

(2014) corroboraram a importância do gene *med24*, destacando sua ortologia também em seres humanos e camundongos, além do peixe-zebra.

Além dessas investigações, Ber e Daniel (1992), ao comparar as sequências no ponto de início da transcrição do gene *GH* em tilápia do Nilo, carpa, truta arco-íris e salmão do atlântico, identificaram uma região de notável homologia antes do TATA Box, uma sequência de DNA encontrada na região promotora central dos genes. Os autores ainda relatam que, apesar da similaridade na estrutura geral, o gene *GH* da tilápia do Nilo apresenta um tamanho do transcrito primário significativamente mais curto (1666 nt) em comparação com outros genes *GH* de peixes já caracterizados.

Conclusão

A tilápia do Nilo (*Oreochromis niloticus*) revela-se uma espécie de peixe com uma notável capacidade de adaptação a diferentes ambientes aquáticos, o que a tornou uma espécie de destaque na aquicultura global. Além disso, essa espécie ocupa um papel relevante em estudos genômicos, destacando-se por possuir tempo de geração curto, tamanho grande suficiente para estudos fisiológicos e por sua facilidade de criação.

O genoma da tilápia do Nilo exibe um tamanho superior a um bilhão de pares de bases, demonstrando a imensa complexidade de sua composição genética. Através dos bancos de dados genômicos, foi possível identificar 22 cromossomos que compõem a estrutura genômica de *O. niloticus*. A análise individual dos cromossomos revelou que o grupo de ligação LG3 é o mais extenso no genoma de *O. niloticus*, enquanto o LG8 é o menor.

Duas cópias do gene *GH1* da tilápia do Nilo foram identificadas, ambas codificando a proteína somatotropina, sendo elas: *GHI-202* e *GHI-201*. Essas transcrições surgiram por meio de eventos de duplicação de todo o genoma após a separação dos salmonídeos. Ao examinar a literatura existente, foi possível identificar outros genes associados à taxa de crescimento e desenvolvimento inicial da tilápia do Nilo, como *raraa*, *rarab*, *med24* e *brms11a*.

Foi possível obter uma árvore filogenética do gene *GHI*, elucidando espécies que exibem homologias e semelhanças próximas ao gene *GHI* de *O. niloticus*, como Midas Cichlid (*Amphilophus citrinellus*), Lyretail cichlid (*Neolamprologus brichardi*), Burton's mouthbrooder (*Astatotilapia burtoni*), Eastern happy (*Astatotilapia calliptera*) e Zebra mbuna

(*Maylandia zebra*) que surgiram por meio de eventos de especiação e se dividiram em diferentes populações, preservando funções semelhantes.

Também foi possível discernir os genes parálogos do gene *GHI* na tilápia do Nilo por meio da utilização da árvore filogenética, que evoluíram por eventos de duplicação genética na era pré-genômica e adquiriram funções diferentes ao longo do tempo. Por meio da literatura existente, foi possível identificar que esses genes surgiram como resultado de eventos de duplicação que ocorreram em um ancestral dos vertebrados com mandíbula (Gnathostomata). Após a ocorrência dos eventos de duplicação genética, os genes parálogos passaram a codificar proteínas como a somatolactina e a prolactina, ambas desempenhando papéis fundamentais no crescimento, desenvolvimento e em processos de transdução de sinal.

Além disso, por meio de uma análise comparativa utilizando a literatura existente, foi possível identificar que os genes associados ao estágio inicial do desenvolvimento da tilápia do Nilo, examinados no contexto do presente estudo, também apresentam características homólogas às de outras espécies. Os genes *raraa* e *rarab*, além de desempenharem papéis fundamentais na tilápia do Nilo, também ocupam uma posição significativa no processo de desenvolvimento do peixe-zebra (*Danio rerio*). Além disso, a ortologia do gene *med24* também surge como notável em humanos e camundongos, além do peixe-zebra.

Em conclusão, os estudos realizados no presente artigo demonstram a importância da pesquisa bibliográfica e da exploração de bancos de dados genômicos como ferramentas essenciais no estudo dos mecanismos genéticos e evolutivos da tilápia do Nilo. É importante ressaltar que, embora a anotação genômica da tilápia do Nilo tenha sido concluída em 2019, a genômica funcional ainda oferece muitas possibilidades de descobertas. Dessa forma, são necessários estudos adicionais para compreender aspectos aprofundados referentes às características genéticas do organismo *Oreochromis niloticus*.

Referências

- AMOUSSOU, T. Olivier *et al.* An insight into advances in fisheries biology, genetics and genomics of African tilapia species of interest in aquaculture. **Aquaculture Reports**, v. 14, p. 100188, jul. 2019. Disponível em: <https://doi.org/10.1016/j.aqrep.2019.100188>. Acesso em: 28 out. 2023.
- BATEMAN, Alex *et al.* UniProt: the universal protein knowledgebase in 2023. **Nucleic Acids Research**, 21 nov. 2022. Disponível em: <https://doi.org/10.1093/nar/gkac1052>. Acesso em: 23 jan. 2024.

BER, Raphael; DANIEL, Violet. Sequence analysis suggests a recent duplication of the growth hormone encoding gene in *Tilapia nilotica*. **Gene**, v. 125, n. 2, p. 143-150, mar. 1993. Disponível em: [https://doi.org/10.1016/0378-1119\(93\)90321-s](https://doi.org/10.1016/0378-1119(93)90321-s). Acesso em: 12 fev. 2024.

BER, Raphael; DANIEL, Violet. Structure and sequence of the growth hormone-encoding gene from *Tilapia nilotica*. **Gene**, v. 113, n. 2, p. 245-250, abr. 1992. Disponível em: [https://doi.org/10.1016/0378-1119\(92\)90402-b](https://doi.org/10.1016/0378-1119(92)90402-b). Acesso em: 13 nov. 2023.

BULL, J. J. Sex determining mechanisms: an evolutionary perspective. **Experientia**, v. 41, n. 10, p. 1285-1296, out. 1985. Disponível em: <https://doi.org/10.1007/bf01952071>. Acesso em: 23 jan. 2024.

CAMPBELL, C. Ryan; POELSTRA, J. W.; YODER, Anne D. What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. **Biological Journal of the Linnean Society**, v. 124, n. 4, p. 561-583, 11 jun. 2018. Disponível em: <https://doi.org/10.1093/biolinnean/bly063>. Acesso em: 23 jan. 2024.

CHAUVE, Cedric; DOYON, Jean-Philippe; EL-MABROUK, Nadia. Gene family evolution by duplication, speciation, and loss. **Journal of Computational Biology**, v. 15, n. 8, p. 1043-1062, out. 2008. Disponível em: <https://doi.org/10.1089/cmb.2008.0054>. Acesso em: 12 fev. 2024.

COSSINS, Andrew R.; CRAWFORD, Douglas L. Fish as models for environmental genomics. **Nature Reviews Genetics**, v. 6, n. 4, p. 324-333, abr. 2005. Disponível em: <https://doi.org/10.1038/nrg1590>. Acesso em: 28 out. 2023.

DEVLIN, Robert H. *et al.* Extraordinary salmon growth. **Nature**, v. 371, n. 6494, p. 209-210, set. 1994. Disponível em: <https://doi.org/10.1038/371209a0>. Acesso em: 1 nov. 2023.

ESHEL, O. *et al.* Linkage and physical mapping of sex region on LG23 of Nile tilapia (*Oreochromis niloticus*). **G3 Genes|Genomes|Genetics**, v. 2, n. 1, p. 35-42, jan. 2012. Disponível em: <https://doi.org/10.1534/g3.111.001545>. Acesso em: 23 jan. 2024.

FAO fisheries & aquaculture. Disponível em: https://www.fao.org/fishery/statistics-query/en/global_production. Acesso em: 9 fev. 2024.

FAO. **The state of world fisheries and aquaculture 2022: Towards Blue Transformation**. Roma: FAO, 2022. *E-book*. ISBN 9789251363645. Disponível em: <https://doi.org/10.4060/cc0461en>. Acesso em: 24 jan. 2024.

FJALESTAD, Kjersti Turid; MOEN, Thomas; GOMEZ-RAYA, Luis. Prospects for genetic technology in salmon breeding programmes. **Aquaculture Research**, v. 34, n. 5, p. 397-406, 28 mar. 2003. Disponível em: <https://doi.org/10.1046/j.1365-2109.2003.00823.x>. Acesso em: 1 nov. 2023.

GUYON, Richard *et al.* A high-resolution map of the Nile tilapia genome: a resource for studying cichlids and other percomorphs. **BMC Genomics**, v. 13, n. 1, p. 222, 2012. Disponível em: <https://doi.org/10.1186/1471-2164-13-222>. Acesso em: 28 out. 2023.

HALE, Laura A. *et al.* Characterization of the retinoic acid receptor genes *raraa*, *rarab* and *rarg* during zebrafish development. **Gene Expression Patterns**, v. 6, n. 5, p. 546-555, jun. 2006. Disponível em: <https://doi.org/10.1016/j.modgep.2005.10.007>. Acesso em: 6 nov. 2023.

HARRISON, Colin; WABBERSEN, Tara; SHEPHERD, Iain T. In vivo visualization of the development of the enteric nervous system using a *Tg(-8.3bpHox2b:kaede)* transgenic zebrafish. **genesis**, v. 52, n. 12, p. 985-990, 4 nov. 2014. Disponível em: <https://doi.org/10.1002/dvg.22826>. Acesso em: 6 nov. 2023.

KAMENSKAYA, D. N.; BRYKOV, V. A. Fish growth hormone genes: structure and divergence. **Russian Journal of Marine Biology**, v. 46, n. 4, p. 233-242, jul. 2020. Disponível em: <https://doi.org/10.1134/s1063074020040033>. Acesso em: 12 fev. 2024.

KAWAUCHI, Hiroshi *et al.* Identification of growth hormone in the sea lamprey, an extant representative of a group of the most ancient vertebrates. **Endocrinology**, v. 143, n. 12, p. 4916-4921, 1 dez. 2002. Disponível em: <https://doi.org/10.1210/en.2002-220810>. Acesso em: 12 fev. 2024.

KOCHER, Thomas D. Adaptive evolution and explosive speciation: the cichlid fish model. **Nature Reviews Genetics**, v. 5, n. 4, p. 288-298, abr. 2004. Disponível em: <https://doi.org/10.1038/nrg1316>. Acesso em: 28 out. 2023.

KOONIN, Eugene V. Orthologs, paralogs, and evolutionary genomics. **Annual Review of Genetics**, v. 39, n. 1, p. 309-338, dez. 2005. Disponível em: <https://doi.org/10.1146/annurev.genet.39.073003.114725>. Acesso em: 20 jan. 2024.

KWONG, Jennifer Q. *et al.* The mitochondrial calcium uniporter underlies metabolic fuel preference in skeletal muscle. **JCI Insight**, v. 3, n. 22, 15 nov. 2018. Disponível em: <https://doi.org/10.1172/jci.insight.121689>. Acesso em: 20 jan. 2024.

MACLEAN, N. *et al.* Transgenic tilapia and the tilapia genome. **Gene**, v. 295, n. 2, p. 265-277, ago. 2002. Disponível em: [https://doi.org/10.1016/s0378-1119\(02\)00735-7](https://doi.org/10.1016/s0378-1119(02)00735-7). Acesso em: 1 nov. 2023.

MARTIN, Fergal J. *et al.* Ensembl 2023. **Nucleic Acids Research**, v. 51, n. D1, p. 933-941, 6 jan. 2023. Disponível em: <https://doi.org/10.1093/nar/gkac958>. Acesso em: 19 jan. 2024.

MUTCH, David M.; WAHLI, Walter; WILLIAMSON, Gary. Nutrigenomics and nutrigenetics: the emerging faces of nutrition. **The FASEB Journal**, v. 19, n. 12, p. 1602-1616, out. 2005. Disponível em: <https://doi.org/10.1096/fj.05-3911rev>. Acesso em: 30 out. 2023.

NICO, L.; SCHOFIELD, P.; NEILSON, M. **Oreochromis niloticus (linnaeus, 1758)**. 18 dez. 2013. Disponível em: <https://nas.er.usgs.gov/queries/factsheet.aspx?SpeciesID=468>. Acesso em: 27 out. 2023.

PONZONI, Raul W. *et al.* Genetic improvement of Nile tilapia (*Oreochromis niloticus*) with special reference to the work conducted by the WorldFish Center with the GIFT strain. **Reviews in Aquaculture**, v. 3, n. 1, p. 27-41, 23 fev. 2011. Disponível em: <https://doi.org/10.1111/j.1753-5131.2010.01041.x>. Acesso em: 4 nov. 2023.

SÁNCHEZ-RAMOS, I. *et al.* Identificación de RFLP en genes relacionados con el crecimiento en dorada *sparus aurata* L., 1758. **Boletín. Instituto Español de Oceanografía.**, v. 21, n. 1-4, p. 253-259, 1 dez. 2005. Disponível em: <http://hdl.handle.net/10261/313270>. Acesso em: 2 nov. 2023.

SAYERS, Eric W. *et al.* Database resources of the national center for biotechnology information. **Nucleic Acids Research**, v. 50, n. D1, p. 20-26, 1 dez. 2021. Disponível em: <https://doi.org/10.1093/nar/gkab1112>. Acesso em: 19 jan. 2024.

SOFY, H.; LAYLA, A.; IMAN, M.K.A. Karyotypic diversity of some tilapia species. **Nature and Science**, v. 6, n. 1, p. 19-27, 2008. Disponível em: https://www.sciencepub.net/nature/0601/03_0346_Sofy_Karyotypic.pdf. Acesso em: 29 out. 2023.

SPRAGUE, J. The Zebrafish Information Network: the zebrafish model organism database. **Nucleic Acids Research**, v. 34, n. 90001, 1 jan. 2006. Disponível em: <https://doi.org/10.1093/nar/gkj086>. Acesso em: 20 jan. 2024.

SUPIWONG, Weerayuth *et al.* Chromosomal characteristic of Nile tilapia (*Oreochromis niloticus*) from mitotic and meiotic cell division by t-lymphocyte cell culture. **Cytologia**, v. 78, n. 1, p. 9-14, 2013. Disponível em: <https://doi.org/10.1508/cytologia.78.9>. Acesso em: 29 out. 2023.

WANG, Xueju *et al.* Retinoic acid receptor alpha drives cell cycle progression and is associated with increased sensitivity to retinoids in T-cell lymphoma. **Oncotarget**, v. 8, n. 16, p. 26245-26255, 17 fev. 2017. Disponível em: <https://doi.org/10.18632/oncotarget.15441>. Acesso em: 12 fev. 2024.

WOOD, Andrew R. *et al.* Defining the role of common variation in the genomic and biological architecture of adult human height. **Nature Genetics**, v. 46, n. 11, p. 1173-1186, 5 out. 2014. Disponível em: <https://doi.org/10.1038/ng.3097>. Acesso em: 4 nov. 2023.

YU, Xiaofei *et al.* Genome-wide association analysis of adaptation to oxygen stress in Nile tilapia (*Oreochromis niloticus*). **BMC Genomics**, v. 22, n. 1, 9 jun. 2021. Disponível em: <https://doi.org/10.1186/s12864-021-07486-5>. Acesso em: 4 nov. 2023.

ZIMMERMANN, Rosalyn C.; WELCH, Danny R. BRMS1: a multifunctional signaling molecule in metastasis. **Cancer and Metastasis Reviews**, v. 39, n. 3, p. 755-768, 30 mar. 2020. Disponível em: <https://doi.org/10.1007/s10555-020-09871-0>. Acesso em: 20 jan. 2024.