



UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE BIOLOGIA

CURSO DE GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS

EVOLUÇÃO DA PROTEÍNA WNT NA CLASSE CALCAREA (PORIFERA)

por

NICOLLE SOUZA LETO

Trabalho de Conclusão de Curso
apresentado ao Instituto de Biologia da
Universidade Federal da Bahia como exigência
para obtenção do grau de Bacharel em Ciências
Biológicas

Orientador: Emilio de Lanna Neto
Coorientadora: Janaina L de Oliveira

Salvador, BA
2023

Data da defesa: 29 de Novembro de 2023

Banca Examinadora

Prof. Dr. Emilio de Lanna Neto

Profa. Dra. Vanessa Rodrigues Paixão Côrtes

Profa. Dra. Paula Carvalhal Lage Von Buettner Ristow

“Everything related to evolution is messy.”
(Wallace Arthur)

AGRADECIMENTOS

À Universidade Federal da Bahia, seu corpo docente, à direção, à administração e demais aos funcionários que possibilitam a realização do curso de Ciências Biológicas.

Aos órgãos fomentadores que possibilitaram o desenvolvimento desta pesquisa: PRPPG-UFBA CNPq e FAPESB.

Aos meus orientadores, professor Dr. Emilio Lanna e professora Dra. Janaina Oliveira, por todo o aprendizado ao longo do meu período no Laboratório de Biologia Evolutiva e do Desenvolvimento - LABED. Ao professor Dr. Lucas Bleicher pelo apoio na realização de parte das análises e no entendimento deste trabalho.

Aos membros egressos e membros atuais do LABED por todas as discussões, reuniões e dicas. Aos amigos por toda cumplicidade ao longo do caminho.

Em especial à minha família: minha mãe, Elisia, e meu pai, Roque, pelo apoio imensurável. Ao Maurício por tudo ao longo desses anos.

RESUMO

Composto por quatro classes: Demospongiae, Hexactinellida, Calcarea e Homoscleromorpha e caracterizado por um corpo simples, sem polarização ântero-posterior (AP) quando adulto, o filo Porifera não é tão simples quanto aparenta. Esponjas possuem importantes genes do desenvolvimento, incluindo o WNT - relacionado à determinação da polaridade celular resultando no estabelecimento do eixo AP. Dentre as classes, Calcarea se destaca pela expansão do número de cópias WNT em seu genoma. Entender a evolução das proteínas WNT nesta classe, portanto, é fundamental para compreender a forma animal. Consequentemente, nos perguntamos: quais inferências evolutivas podem ser feitas a partir da comparação das estruturas primárias e terciárias das proteínas WNT de Calcarea? Nós obtivemos sequências de aminoácidos de WNT a partir de dados públicos, para representantes das duas subclasses, Calcinea e Calcaronea, modelamos as proteínas, recuperamos filogenias de aminoácidos e da estrutura terciária, analisamos os padrões de conservação, coevolução e de substituições. Encontramos 62 sequências para quatro espécies: na subclasse Calcaronea - *Sycon ciliatum* (21), *Sycon coactum* (11) e *Leucosolenia complicata* (17) e na subclasse Calcinea - *Leucetta chagosensis* (13). Nas filogenias, não há relação restrita entre ramos e subclasses, indicando que a multiplicação das cópias ocorreu antes de sua divergência. Observamos dois grupos de estruturas terciárias de WNT em Calcarea: um com estrutura mais conservada próxima ao padrão desta proteína em outros animais e outro com perda de regiões da proteína. Com substituições conjuntas, 111 sítios coevoluem, majoritariamente no domínio C-terminal que inclui uma região de maior flexibilidade a alterações. Seis grupos de aminoácidos coevoluem com manutenção conjunta, possivelmente associados à forma e função da proteína. Concluímos que há maior diversidade de estruturas primárias do que terciárias, sendo dois grupos tridimensionais de proteínas que não apresentam sinal filogenético, porém com uma plausível perda ou alteração de função em um deles.

Palavras chave: Aminoácidos, Estrutura terciária, Conservação, Coevolução, Esponjas.

ABSTRACT

Composed by four classes: Demospongiae, Hexactinellida, Calcarea and Homoscleromorpha, and characterized by a simple body lacking a clear anteroposterior (AP) polarization when adult, the phylum Porifera is not as simple as it seems. Sponges have important developmental genes, including WNT - related to the determination of cell polarity, which, in turn, results in the AP axis establishment. Among the classes, Calcarea stands out as having a remarkable expansion of WNT copies in the genome. Understanding the evolution of WNT proteins in this class is a fundamental step towards understanding the animal form. Consequently, we asked: what evolutionary inferences can be made from comparing the primary and tertiary structures of Calcarea WNT proteins? We obtained WNT amino acid sequences from public data for representatives of the two subclasses, Calcinea and Calcaronea, modeled the proteins, recovered amino acids and tertiary structure phylogenies, analyzed the patterns of conservation, coevolution and substitutions in sites. We found 62 sequences for four species: subclass Calcaronea - *Sycon ciliatum* (21), *Sycon coactum* (11), and *Leucosolenia complicata* (17) and for the subclass Calcinea - *Leucetta chagosensis* (13). There was no restricted relationship between the protein clades and the subclasses, indicating an ancestral expansion of the copies, before the divergence of the subclasses. We observed two groups of WNT based on their tertiary structures in Calcarea: one with a more conserved structure resembling to the standard three-dimensional shape of this protein in other animals, and another group showing different type of losses of regions of the protein. With correlated substitutions, 111 sites coevolved, mostly in the C-terminal domain, which includes a region of greater flexibility to changes. Six groups of amino acids coevolved with joint maintenance, possibly associated with the form and function of the protein. We conclude that there is a greater diversity of primary than tertiary structures of WNT in these sponges, with two groups of proteins that do not present a phylogenetic signal, but with a plausible loss or alteration of function in one of them.

Keywords: Amino acids, Tertiary structure, Conservation, Coevolution, Sponges.

SUMÁRIO

SUMÁRIO	6
INTRODUÇÃO	8
MATERIAL E MÉTODOS	15
Obtenção de dados	15
Análises filogenéticas	16
Análises de conservação e substituição	17
Análises de coevolução de aminoácidos	18
RESULTADOS	20
Agrupamentos de WNT em Calcarea	20
Padrões de conservação e substituição	24
Padrões de coevolução	30
DISCUSSÃO	35
Caracterização, diversidade e evolução das proteínas WNT na classe Calcarea	35
REFERÊNCIAS	45
APÊNDICES	50

Capítulo único

Artigo a ser submetido ao periódico *Proteins: Structure, Function, Bioinformatics*
Homepage: <<https://onlinelibrary.wiley.com/journal/10970134>>

INTRODUÇÃO

As chamadas “vias de sinalização” são fundamentais para controlar o desenvolvimento embrionário e tardio dos corpos dos organismos multicelulares. Cada vez mais sabemos sobre a presença das moléculas relacionadas a estas vias em todos os animais, mas a forma como elas evoluem ainda são pouco compreendidas. Uma destas importantes vias é conhecida como WNT. O gene WNT codifica uma glicoproteína homônima de 350-400 aminoácidos que se liga ao receptor transmembrana Frizzled (Fz) (Croce, McClay, 2009). A proteína WNT é composta por dois domínios: o *carboxy-terminal* (CTD) com 6 pontes dissulfeto e o *amino-terminal* (NTD) com 5 pontes dissulfeto; resultado da interação entre os 22-24 resíduos de Cisteína (Croce, McClay, 2009). Estas proteínas são caracterizadas pelas pontes dissulfeto e também pela presença de um motivo composto pelos aminoácidos RWNC na região C-terminal (Janda, *et al.*, 2012). A estrutura quaternária completa do complexo WNT-Fz inclui também a adição do ácido palmitoleico (PAM) ligado à extremidade do NTD através de um resíduo de Serina, a adição de N-glicano ligado à região de junção de NTD-CTD e a interação proteína-proteína entre WNT-Fz derivada da interação Cisteína-Cisteína no CTD (Fig. 1a-b) (Janda, *et al.*, 2012).

A proteína WNT é comumente relacionada a diferentes vias de sinalização: as vias canônicas e as vias não-canônicas. Apesar disso, essencialmente, as proteínas envolvidas nessas vias são estruturalmente as mesmas. Por isso, alguns autores consideram essa divisão inadequada e atribuem às diferentes atividades do WNT ao contexto celular (Willert, Nusse, 2012). O WNT é central no desenvolvimento animal, pois está envolvido na sinalização atribuída a diversas funções no desenvolvimento inicial, como a determinação de polaridade celular resultando no estabelecimento do eixo ântero-posterior (AP) e funções no desenvolvimento tardio, como a manutenção de células indiferenciadas e tecidos (Komiya, Habas, 2008; Holstein, 2012). Além disso, o WNT atua na padronização dos tecidos durante todo o desenvolvimento, oferecendo informação posicional que leva a diferenciação tecidual (Verkaar, *et al.*, 2012; Li, *et al.*, 2020).

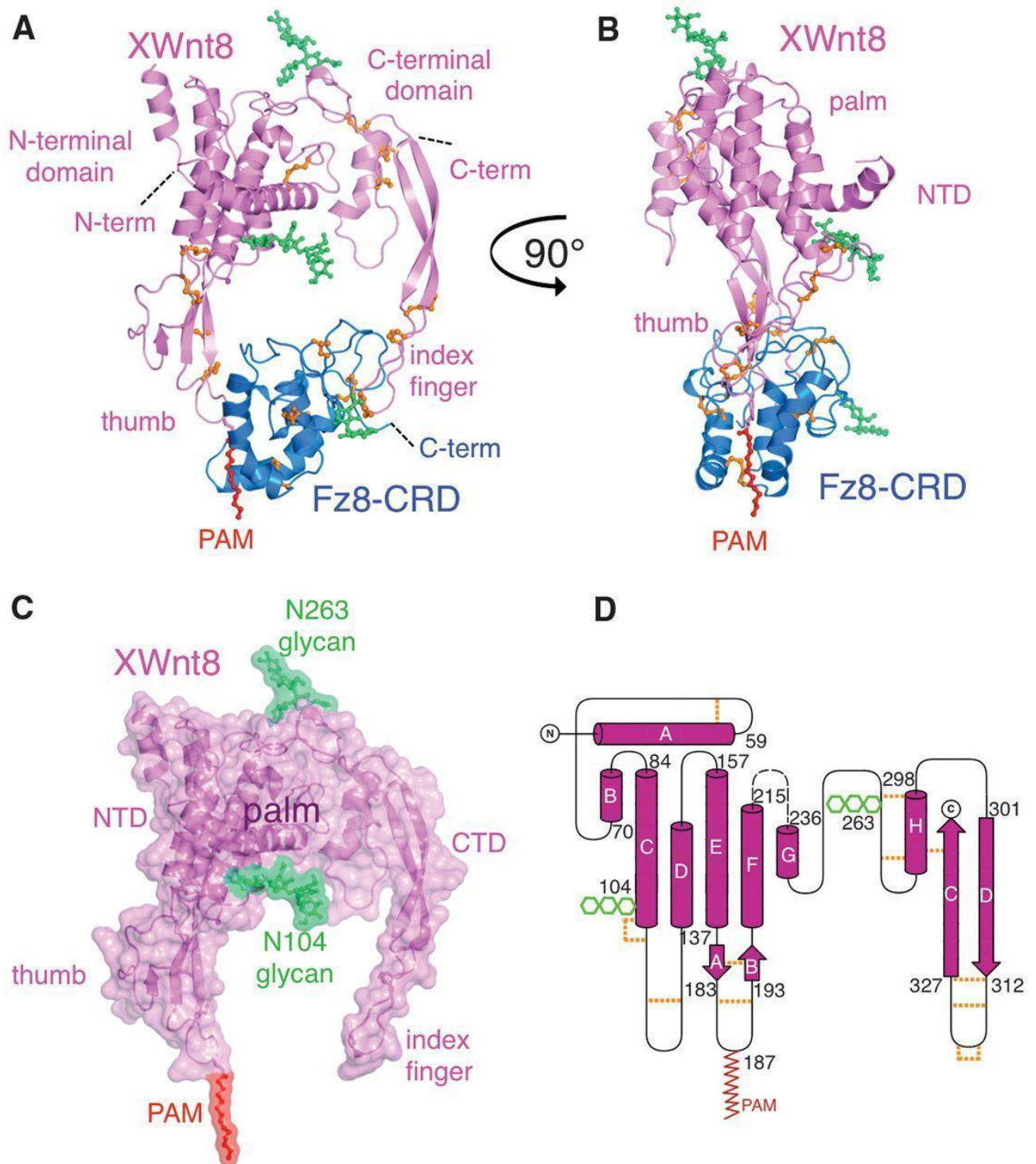


Figura 1 - Estrutura tridimensional do complexo WNT-Frizzled e demais ligantes do WNT8 em complexo com o Frizzled8 em *Xenopus laevis*. (a) Diagramas de fita da interação WNT (rosa) e Frizzled (azul) com ligantes N-Glicano (verde), Ácido Palmitoleico (vermelho) e destaque das pontes dissulfeto (laranja), (b) Diagramas de fita de WNT e Frizzled rotacionado com foco na ligação WNT-Frizzled mediado pelo Ácido Palmitoleico, (c) Representação da superfície da proteína WNT, (d) Estrutura secundária da proteína WNT com destaque de cada interação e seus sítios. (Retirado de Janda *et al.* 2012).

A arquitetura modular dos WNT pode estar relacionada à sua origem evolutiva por fusão de dois domínios independentes: D1, semelhante às saposinas, e D2, semelhante às citocinas (Bazan, *et al.*, 2012). O domínio saposina caracteriza as proteínas

saposinas presentes de protozoários a metazoários mais derivados e são relacionadas a interações lipídicas (Bruhn, 2005). As citoquinas são proteínas que atuam na comunicação entre células (Zhang, An, 2007). O domínio D1 seria então correspondente à região NTD, enquanto o domínio D2 relaciona-se ao domínio CTD (Bazan, *et al.*, 2012). Apesar dessas semelhanças e dessa possível origem evolutiva, a proteína WNT é mais do que a soma de partes. A proteína WNT é restrita aos metazoários e não há estruturas similares em outros grupos, indicando a possibilidade da origem dessa proteína ter sido em um ancestral comum aos animais (Loh, *et al.*, 2016). Conseqüentemente, como a via WNT possui um longo histórico de evolução em conjunto aos metazoários, a compreensão da evolução da proteína nos ajuda a compreender a evolução do grupo e vice-versa (Ryan, Baxevanis, 2007; Bazan, *et al.*, 2012; Rim, *et al.*, 2022).

O grupo não-monofilético dos metazoários não-bilaterais, composto pelos filos Porifera, Ctenophora, Placozoa e Cnidaria, os primeiros clados a divergirem na linhagem dos animais, fornece uma janela ao entendimento da evolução de moléculas envolvidas na formação de características corporais complexas (Renard, *et al.*, 2018). Apesar de algumas discussões recentes questionarem quem seria o grupo-irmão de todos os animais, é atualmente aceito que as esponjas são os animais vivos mais antigos a terem surgido no planeta (ver Li, *et al.*, 2021). Porifera é um filo monofilético e conta, atualmente, com aproximadamente 11 mil espécies (WORMS, 2023), divididas em quatro classes (também monofiléticas): Demospongiae, Hexactinellida, Calcarea e Homoscleromorpha. As esponjas são animais filtradores caracterizados por um plano corporal simples, organizado em função do sistema aquífero (Fig. 2), não possuindo uma polarização AP clara quando adultas, além da ausência de organização tecidual *stricto sensu* (Degnan, *et al.*, 2005; Adamska, 2021). No entanto, sua composição genômica é muito mais complexa do que o esperado. Por exemplo, as esponjas, como um todo, apresentam o conjunto de genes do desenvolvimento clássicos dos metazoários, como PAX, SOX, TGF-Beta e WNT (Srivastava, *et al.*, 2010; Riesgo, *et al.*, 2014; Dunn, *et al.*, 2015). Além disso, as esponjas também possuem padrões complexos do desenvolvimento, como o processo único encontrado na classe Calcarea, a inversão, onde as células inicialmente interiorizadas no embrião são reposicionadas de forma radical (Lanna, Klautau, 2012). A classe Calcarea é caracterizada pela presença de espículas de carbonato de cálcio. É a única classe que possui os cinco tipos de organização dos sistema aquífero (asconoide, siconoide, solenoide, leuconoide e

sileibide, Fig. 2), mas que não apresentam origens evolutivas únicas (i.e., não apresentam sinal filogenético, em que clados evolutivamente próximos apresentam atributos similares delimitados em uma filogenia; Blomberg, *et al.*, 2003) entre as suas duas subclasses, Calcaronea e Calcinea, ou até mesmo entre os *rankings* taxonômicos inferiores (i.e., ordens e famílias; Eerkes-Medrano, Leys, 2006; Ereskovsky, 2010; Cavalcanti, Klautau, 2011; Voigt, *et al.*, 2012).

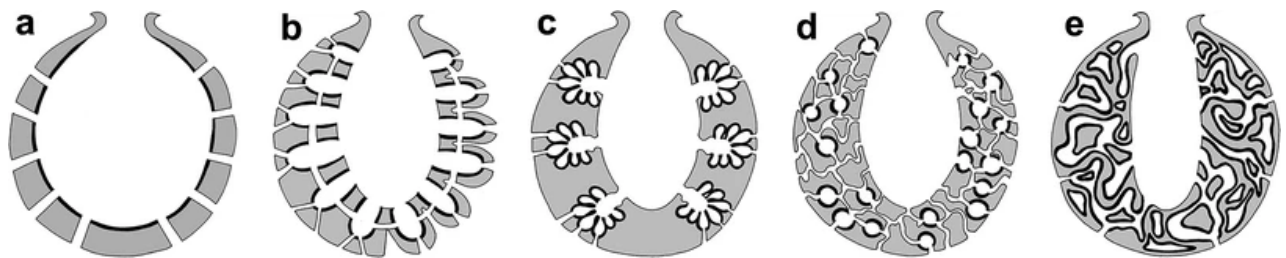


Figura 2 – Diagrama mostrando os diferentes tipos de planos corporais em função do sistema aquífero em Porifera. (a) Asconoide, (b) Siconoide, (c) Sileibide, (d) Leuconoide, (e) Solenoide. Destaques em preto representam a delimitação dos coanócitos. (Retirado de Cavalcanti, Klautau, 2011).

Em Porifera, os genes que compõem a via WNT começaram a ser estudados no início dos anos 2000 com a identificação de presença/ausência em algumas espécies modelo. Inicialmente, foi detectada a presença de receptores *Frizzled* em *Suberites domuncula*, a presença de WNT em *Oscarella* sp. e em *Amphimedon queenslandica* (Adell, *et al.*, 2003; Nichols, *et al.*, 2006; Adamska, *et al.*, 2007). Posteriormente, no fim dos anos 2000 e ao longo dos anos 2010, marcado pela publicação do primeiro genoma de esponjas em 2010, os estudos de presença/ausência se expandiram para análise da estrutura da via e dos genes, análise de padrões de expressão e análises do efeito da inibição dos genes WNT em diversas espécies. Foram realizadas novas análises de estrutura e expressão para *Oscarella* sp. e *Amphimedon queenslandica*, investigações sobre a presença e análise de inibição de WNT em *Ephydatia muelleri*, presença, estrutura e padrão de expressão em *Sycon ciliatum*, presença/ausência e caracterização da via em *Aphrocallistes vastus*, *Spongilla lacustris*, *Petrosia ficiformis*, *Pseudospongosorites suberitoides*, *Ircinia fasciculata*, *Chondrilla nucula*, *Corticium candelabrum*, e, finalmente, estudos sobre a presença, estrutura e expressão de WNT em *Halisarca dujardini* (Adell, *et al.*, 2007; Adamska, *et al.*, 2010; Lapébie, *et al.*, 2009; Windsor, Leys, 2010; Leininger, *et al.*, 2014; Riesgo, *et al.*, 2014, Borisenko, *et al.*, 2016, Windsor-Reid, *et al.*, 2018, Borisenko, *et al.*, 2021).

A partir desses estudos, é possível afirmar atualmente que: (i) três das quatro classes apresentam a proteína WNT e os demais elementos da via, sendo exceção a classe Hexactinellida que não apresenta WNT, possivelmente uma perda secundária no grupo (Riesgo, *et al.*, 2014; Santini, *et al.*, 2023); (ii) enquanto as espécies das classes Demospongiae e Homoscleromorpha apresentam 2-3 cópias de WNT (com algumas exceções pontuais), as espécies da classe Calcarea apresentam 13-21 cópias de WNT, uma possível expansão única para a classe (Lapébie, *et al.*, 2009; Adamska, *et al.*, 2010; Riesgo, *et al.*, 2014; Leininger, *et al.*, 2014; Borisenko, *et al.*, 2016; Windsor-Reid, *et al.*, 2018); (iii) os WNT das esponjas apresentam o padrão clássico de 22-24 resíduos de Cisteína e o motivo RNWC semi-conservado, reforçando a ideia de origem ancestral do WNT a todos os metazoários vivos (Windsor-Reid, *et al.*, 2018); (iv) análises filogenéticas indicam que os WNT das esponjas não se agrupam fortemente aos WNT dos demais metazoários (Lapébie, *et al.*, 2009; Leininger, *et al.*, 2014; Windsor-Reid, *et al.*, 2018); e (v) existem indícios de regionalização de expressão do gene WNT na região posterior durante o estágio larval e na região oscular durante o estágio adulto, sugerindo uma função de polarização nas esponjas para este produto gênico (Fig. 3) (Adamska, *et al.*, 2010; Windsor, Leys, 2010; Leininger, *et al.*, 2014; Borisenko, *et al.*, 2016; Windsor-Reid, *et al.*, 2018). O item (iv) merece destaque, pois ressalta a necessidade e a importância de se entender os WNT nas esponjas. A filogenia de WNT mais robusta (até então) foi realizada com sequências de aminoácidos de dezessete espécies das três classes, sugerindo que há subfamílias específicas das esponjas (Windsor-Reid, *et al.*, 2018). Porém, ainda assim algumas cópias de WNT da demosponja *Halisarca dujardini* e das homoscleromorfos *Corticium candelabrum* e *Oscarella carmela* não se encaixam nessas subfamílias (Windsor-Reid, *et al.*, 2018). Características exclusivas desses genes no primeiro filo a divergir da linhagem animal são esperadas devido ao longo período de tempo que esses genes tiveram para acumular alterações, demonstrando o resultado da complexidade da evolução desses genes nos metazoários ao longo de ao menos 600 milhões de anos (Ryan, Baxanevis, 2007). Consequentemente, é necessário que novos estudos buscando entender a evolução destes genes sejam realizados.

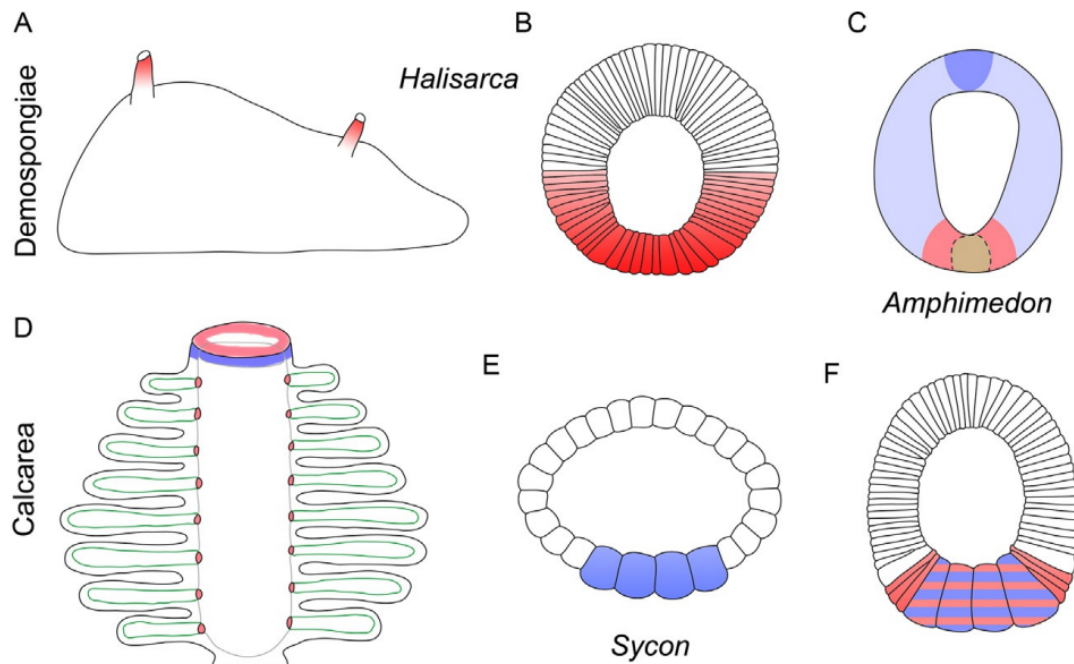


Figura 3 - Resumo dos padrões de expressão de WNT e TGF-Beta em algumas espécies modelo de Porifera. (A-C) Classe Demospongiae, (D-F) Classe Calcarea. (A) *Halisarca dujardini* em estágio adulto, expressão de WNT na região oscular, (B) *Halisarca dujardini* em estágio larval, expressão de WNT na metade posterior, (C) *Amphimedon queenslandica* em estágio larval, expressão de WNT concentrada na região posterior, (D) *Sycon ciliatum* em estágio adulto, expressão de WNT na região oscular, (E) *Sycon ciliatum* em estágio larval inicial pré-inversão, ausência de expressão de WNT, (F) *Sycon ciliatum* em estágio larval tardio, expressão de WNT na região posterior. WNT em vermelho, TGF-Beta em azul. Intensidade da cor representa intensidade de expressão. (Retirado de Borisenko, *et al.*, 2019).

Nos últimos 20 anos, o estudo das ômicas nos primeiros filós a divergirem na linhagem dos metazoários impulsionou o entendimento do que faz os animais serem animais (Renard, *et al.*, 2018). Grandes perguntas começaram a ser elucidadas e geraram novas questões: da discussão de quem seria o grupo-irmão de todos os metazoários até a composição da caixa de ferramentas básica no genoma do último ancestral comum aos metazoários (Harcet, *et al.*, 2010; Riesgo, *et al.*, 2014). Estudos *in silico* que revisitam dados moleculares pré-publicados com novas perguntas e perspectivas são uma etapa essencial para aprofundar o conhecimento que temos nos animais não-bilaterais, em especial em Porifera. O estudo das proteínas da via WNT, com destaque à própria proteína WNT, ainda não foi explorado nas esponjas. Acessar experimentalmente a conformação tridimensional das proteínas é um processo complexo e dispendioso. Porém, uma via para compreendermos essas estruturas terciárias e as suas devidas implicações na célula e nas vias de sinalização é a bioinformática estrutural (Verli, 2014). A bioinformática estrutural é o campo que estuda características biológicas a nível molecular a partir de informações tridimensionais interpretadas computacionalmente (Verli, 2014). Além da importância por si dos estudo das estruturas terciárias, a sua junção

ao estudo de sequências de aminoácidos, i.e., estruturas primárias, fornece ainda mais detalhes da evolução das proteínas, uma vez que as alterações de sequências de aminoácidos de uma proteína tem influência direta na sua conformação tridimensional e, conseqüentemente, na sua função e diversas propriedades (Aravind, *et al.*, 2002; Bhattacharya, *et al.*, 2017).

Nesse sentido, compreender a evolução da proteína WNT das esponjas é uma importante contribuição ao entendimento da evolução animal, especialmente para a classe Calcarea que apresenta peculiaridades em seu desenvolvimento e no número de cópias de WNT. Assim, as perguntas norteadoras para este estudo são: O número expandido de cópias WNT na classe Calcarea é refletido em diversidade estrutural de suas proteínas? A forma tridimensional dos WNT em Calcarea é conservada em relação aos demais metazoários? Quais inferências evolutivas podem ser feitas a partir da comparação da diversidade de estruturas primárias e da diversidade de estruturas terciárias das proteínas WNT? Nosso objetivo foi identificar, comparar e caracterizar os padrões presentes nas estruturas primárias e terciárias da proteína WNT nas duas subclasses de Calcarea através de ferramentas *in silico* visando entender a evolução desta importante proteína animal.

Encontramos 62 sequências para quatro espécies: na subclasse Calcaronea - *Sycon ciliatum* (21), *Sycon coactum* (11) e *Leucosolenia complicata* (18) e na subclasse Calcinea - *Leucetta chagosensis* (13). Não há relação restrita entre os clados de WNT e os táxons, espécies ou subclasses. Isto indica que a multiplicação das cópias ocorreu antes da separação das subclasses. Observamos dois grupos de estruturas terciárias de WNT em Calcarea: um com estrutura mais conservada próxima ao padrão desta proteína em outros animais e outro com perda de regiões da proteína. Com substituições conjuntas, 111 sítios coevoluem, majoritariamente no domínio C-terminal que inclui uma região de maior flexibilidade a alterações. Seis grupos de aminoácidos coevoluem com manutenção conjunta, possivelmente associados à forma e função da proteína. Podemos concluir que existem características, até então, únicas às proteínas WNT da classe Calcarea sem um sinal filogenético forte entre as suas subclasses e que o número expandido de cópias WNT na classe Calcarea não se reflete em uma grande variedade de diversidade tridimensional, havendo dois agrupamentos principais para essas estruturas na classe, possivelmente derivados da variação das restrições da evolução entre as cópias WNT.

MATERIAL E MÉTODOS

Obtenção de dados

A busca de espécies de esponjas da classe Calcarea a serem usadas no presente estudo foi realizada de três maneiras: (1) seleção de artigos no buscador Google Scholar (scholar.google.com.br/) que contivessem identificação prévia de sequências WNT e (2) busca de genomas e transcriptomas publicados para espécies da classe Calcarea listadas no BioProject-NCBI (ncbi.nlm.nih.gov/bioproject/) e no banco de dados do European Molecular Biology Laboratory - EMBL (embl.org/). A seleção das espécies foi feita com base na real disponibilidade dos dados em bancos públicos ou nos materiais suplementares dos artigos e em critérios taxonômicos, mantendo espécies representativas para as subclasses Calcinea e Calcaronea (as duas subclasses da classe Calcarea). Através de (1) encontramos as sequências WNT em aminoácidos, estrutura primária, para três espécies da subclasse Calcaronea: *Sycon ciliatum* (SCI) identificada por Leininger e colaboradores (2014), *Sycon coactum* (SCO) e *Leucosolenia complicata* (LCO) compiladas por Windsor-Reid e colaboradores (2018), identificadas, respectivamente, por Riesgo e colaboradores (2014) e Fortunato e colaboradores (2012). Por meio de (2) encontramos dados de transcriptoma da espécie *Leucetta chagosensis* (LCH), subclasse Calcinea (BioProject-Accession: PRJNA689294 ID: 689294, dados depositados na biblioteca Transcriptome Shotgun Assembly). Para obter as sequências WNT de LCH, realizamos uma análise tBLASTx (blast.ncbi.nlm.nih.gov/Blast.cgi) (Altschul, *et al.*, 1990) comparando todo o seu transcriptoma com as sequências WNT da espécie desmosponja *Amphimedon queenslandica*, cujas sequências WNT foram utilizadas em trabalhos anteriores como guia para a identificação deste gene (Adamska, *et al.*, 2010, Leininger, *et al.*, 2014). A busca retornou 20 sequências significativamente similares a WNT, a partir das quais realizamos a predição de ORFs (*Open Reading Frames*) através do ORFFinder (ncbi.nlm.nih.gov/orffinder/) com tamanho mínimo de 75 nucleotídeos, código genético padrão e qualquer códon de sentido como iniciador (Rombel, *et al.*, 2002). Encontramos 19 ORFs e a partir dessas sequências buscamos semelhanças com famílias, domínios, sítios e motivos proteicos específicos de WNT por meio do InterPro (ebi.ac.uk/interpro/) (Paysan-Lafosse, *et al.*, 2023). Essa análise resultou em 18 sequências WNT. Visando remover sequências redundantes, fizemos a tradução de nucleotídeos para aminoácidos com o BioEdit (Hall, *et al.*, 2011), para realização e

análise de um alinhamento de aminoácidos no *webserver* MAFFT (mafft.cbrc.jp/alignment/server/) (Kato, Toh, 2008). Foi mantida apenas uma das cópias para aquelas sequências que apresentaram sobreposição exatas no alinhamento. Assim, recuperamos 13 proteínas WNT em sequência de aminoácidos, que chamaremos daqui em diante de “estrutura primária”, para *L. chagosensis*. Os nomes para as proteínas de *S. ciliatum*, *S. coactum* e *L. complicata* foram mantidos de acordo com os artigos que identificaram esses genes (Fortunato *et al.*, 2012; Leininger, *et al.*, 2014; Riesgo, *et al.*, 2014; Windsor-Reid, *et al.*, 2018). Para as proteínas WNT de *L. chagosensis*, nomeamos as proteínas como ‘LCHwnt’ de A-M de acordo com a ordem em que a sequência foi isolada em nossa análise.

Os modelos tridimensionais das estruturas terciárias de cada proteína foram obtidos por meio da modelagem por homologia no *webserver* SwissModel (swissmodel.expasy.org/) (Kiefer, *et al.*, 2009). A modelagem por homologia é baseada na utilização de uma proteína, cuja estrutura tenha sido acessada de maneira experimental, como molde para um novo modelo, desde que essas proteínas tenham origens evolutivas em comum. A proteína molde foi escolhida utilizando o BLAST com base em dois critérios recomendados e inferidos pela plataforma: a maior identidade (numa escalada de 0 a 1, com mínimo 0.3) e o método de obtenção do molde, priorizando métodos experimentais como a cristalização da estrutura e visualização por difração de raios-X. Três proteínas WNT já foram cristalizadas e acessadas por difração de raios-X: o complexo WNT8-Frizzled de *Xenopus laevis* (Janda, *et al.*, 2012), WNTD de *Drosophila melanogaster* (Chu, *et al.*, 2013) e o complexo WNT8-Frizzled de *Homo sapiens* (Hirai, *et al.*, 2019), depositadas no *Protein Data Bank* com códigos 4F0A, 4KRR, 6AHY, respectivamente. A fim de identificar qual delas apresentaria a maior identidade com as sequências de *Calcarea*, realizamos testes com as três proteínas. A proteína WNT de humanos possui valores mais adequados de identidade (entre 0.26-0.42), portanto, a modelagem dos WNT de *Calcarea* foi realizada com o WNT de humanos.

Análises filogenéticas

A fim de compreender os agrupamentos entre as proteínas WNT de *Calcarea*, recuperamos filogenias de acordo com o nível estrutural analisado, ou seja, para as estruturas primárias e para estruturas terciárias. A primeira abordagem, com a estrutura primária, foi obtida por meio do alinhamento múltiplo de sequências de aminoácidos

(AMS), no *webserver* MAFFT (mafft.cbrc.jp/alignment/server/index.html), seguida pela reconstrução da árvore filogenética por máxima verossimilhança, com uso do modelo *SH like* para nos testes *likelihood ratio test*, e o suporte dos ramos foi obtido por *bootstrap*, com mil réplicas, no PhyML (atgc-montpellier.fr/phyml/) (Guindon, *et al.*, 2010).

Partindo das estruturas terciárias modeladas previamente, recuperamos um cladograma de similaridade estrutural e uma análise de componentes principais que mapeia as semelhanças tridimensionais entre as estruturas. O dendrograma foi obtido através do *software* PROMALS3D com suas configurações padrão (prodata.swmed.edu/promals3d/promals3d.php) (Pei, *et al.*, 2008). O PROMALS3D analisa informação estrutural (terciária) ao comparar estruturas de bancos de dados e as estruturas secundárias geradas a partir da submissão do AMS, obtido anteriormente. Obtivemos através desse *software* um alinhamento de estruturas terciárias e o dendrograma. Numa análise do tipo “All against all”, em que as coordenadas atômicas dos modelos são submetidas e comparadas entre si, geramos uma análise de componentes principais que resulta num agrupamento das estruturas terciárias mais semelhantes entre si através do DaliServer (ekhidna2.biocenter.helsinki.fi/dali/) (Holm, *et al.*, 2023).

Para mapear a diversidade de estruturas tridimensionais na filogenia de aminoácidos, submetemos conjuntos de estruturas ao *webserver* Mol*3DViewer (rcsb.org/3d-view) (Sehnal, *et al.*, 2021) gerando uma imagem com a sobreposição das estruturas tridimensionais para cada um dos clados.

Análises de conservação e substituição

Dividimos as análises de conservação da proteína de acordo ao nível estrutural analisado: as estruturas primárias, que correspondem às sequências de aminoácidos e as estruturas terciárias, que correspondem aos modelos tridimensionais. Para acessar os padrões de conservação das estruturas primárias, realizamos duas análises sítio-específicas que avaliam a manutenção de um determinado aminoácido em uma posição específica de um alinhamento múltiplo de sequências (AMS). Na primeira análise, em que o AMS foi submetido e analisado de acordo à configuração padrão no *software* Skylign (skylign.org/) (Wheeler, *et al.*, 2014), geramos um perfil logos baseado no modelo oculto de Markov (*hidden Markov model logos profile - HMM logos profile*). Nesse modelo, o padrão de conservação de um conjunto de sequências é exibido por meio da frequência

de ocorrência e a probabilidade de ocorrência de um determinado aminoácido em uma posição do alinhamento, além de exibir quão informativo um sítio é no alinhamento. Na segunda análise, um novo AMS foi gerado com as sequências WNT das esponjas em conjunto a mais de 900 sequências WNT depositadas em bancos de dados. Este novo AMS foi submetido e avaliado por meio de uma análise de correlação da manutenção de um determinado aminoácido em uma posição do alinhamento, executada no *software* PFStats (disponível a pedido) (Fonseca-Junior, *et al.*, 2018). Após a identificação e a remoção de *outliers* de todo o alinhamento (três proteínas WNT de esponjas: SClwntI, SClwntO e LCHwntJ), obtivemos a relação de aminoácidos com mais de 80% de conservação.

Submetemos individualmente as estruturas terciárias no software PyMol e ao site Mol*3D Viewer (rcsb.org/3d-view) (Yuan, *et al.*, 2017; Sehnal, *et al.*, 2021), para acessar os padrões de conservação das estruturas terciárias a partir de detalhes do modelo tridimensional, como a posição de cada aminoácido no modelo.

Acessamos os padrões de substituição de aminoácidos observando os resultados secundários das análises de conservação e de coevolução (descritas abaixo). Na análise de conservação, o software PFStats tem como *output* uma tabela em que os sítios conservados ao longo do alinhamento compõem as colunas e o aminoácido contido nesses sítios é exibido nas linhas para cada proteína, entretanto também são exibidos os sítios em que há mudança do aminoácido e o aminoácido contido nele. Por exemplo, o sítio 534 do alinhamento é conservado, $pp=0.85$, com uma Cisteína o ocupando. Porém, para a proteína SCOWntX2 o aminoácido apresentado é uma Leucina. Partindo dessa tabela, nós analisamos os sítios que apresentavam divergências desse tipo. O mesmo foi feito para as tabelas de grupos sob coevolução geradas pelo PFStats.

Análises de coevolução de aminoácidos

Para acessar os padrões de coevolução das estruturas primárias, assim como para os estudos de conservação anteriores, utilizamos análises sítio-específicas. Na primeira análise, utilizamos o AMS e a filogenia de estruturas primárias para obter uma tabela com os pares de sítios coevoluindo com probabilidade a posteriori (pp) >0.95 por meio da análise *Bayesian Graphical Models for co-evolving sites* (BGM) no *software* Datamonkey (datamonkey.org/bgm) (Pond, *et al.*, 2005). A análise BGM detecta, através das interações entre os sítios e seus padrões de substituição correlacionada, a presença de

coevolução entre pares de sítios, esses eventos são mapeados de acordo à filogenia (Pond, *et al.*, 2005). Posteriormente, reconhecemos os sítios mais significativos, ou seja, 1 sítio com mais de um par sob coevolução (por exemplo: o sítio 830 com os pares 830-432, 830-572, 830-606).

Na segunda análise, em que um AMS gerado com as sequências WNT das esponjas em conjunto a mais de 900 sequências WNT depositadas em bancos de dados foi submetido ao software PFStats (disponível a pedido), obtivemos os grupos de aminoácidos coevoluindo, por meio de análise de correlação que considera a manutenção conjunta de grupos de sítios ao longo do alinhamento. Na segunda análise, assim como para os estudos de conservação anteriores, mapeamos os achados do BGM e do PFStats nas estruturas tridimensionais das proteínas utilizando o software PyMol e o site Mol*3D Viewer (rcsb.org/3d-view) (Yuan, *et al.*, 2017; Sehnal, *et al.*, 2021).

RESULTADOS

A partir da busca de espécies e a posterior identificação das cópias WNT em seus genomas e transcriptomas, encontramos 62 proteínas WNT para quatro espécies de esponjas calcárias, sendo: três espécies da subclasse Calcaronea - *Sycon ciliatum* (SCI) com 21 cópias, *Sycon coactum* (SCO) com 11 cópias e *Leucosolenia complicata* (LCO) com 17 cópias; e uma espécie da subclasse Calcinea - *Leucetta chagosensis* (LCH) com 13 cópias. Entre as 62 proteínas, a menor sequência de aminoácidos foi encontrada em LCHwntJ (277 aa) e a maior sequência de aminoácidos foi encontrada em LCOwntQ (502 aa). Apesar dessa discrepância pontual, a maioria das sequências têm mais de 350 aminoácidos (57 proteínas). Obtivemos os modelos tridimensionais para as 62 proteínas utilizando como molde a proteína WNT dos humanos. As identidades molde-modelo se mantiveram entre 24.53% (SCIwntO) e 44.59% (LCHwntC), sendo o maior número de estruturas com identidades entre 30-40% (56 proteínas).

Agrupamentos de WNT em Calcarea

A filogenia recuperada a partir das sequências de aminoácidos não apresentou restrição de espécies em clados de WNT, não havendo sinal filogenético forte; i.e., diferentes proteínas WNT de uma espécie se agrupam com proteínas WNT de outras espécies. Também não há restrição dos clados de proteínas identificados aqui para as subclasses; i.e., proteínas WNT de espécies da subclasse Calcinea se agrupam com proteínas WNT da subclasse Calcaronea sem restrições. Partindo da topologia da árvore, valores de suporte (*bootstrap*) superiores a 0.75 e da sobreposição dos modelos tridimensionais, dividimos seis clados de proteínas WNT na classe Calcarea, denominados C1-C6 e marcado com cores (e.g. C1-Rosa) (Fig. 4a). O cladograma de semelhança estrutural (posição das coordenadas atômicas no modelo tridimensional e alinhamento de estruturas terciárias) das 62 proteínas WNT recuperou padrões semelhantes em relação à filogenia de aminoácidos nos ramos mais internos do diagrama, por exemplo, os grupos-irmão: SCOWntX1 e SCIwntD, LCOwntQ e SCIwntQ, LCOwntS e SCIwntS, SCIwntF e LCOwntF, SCIwntL e SCOWntX5 são sempre encontrados de maneira mais diretamente relacionada (Fig. 5).

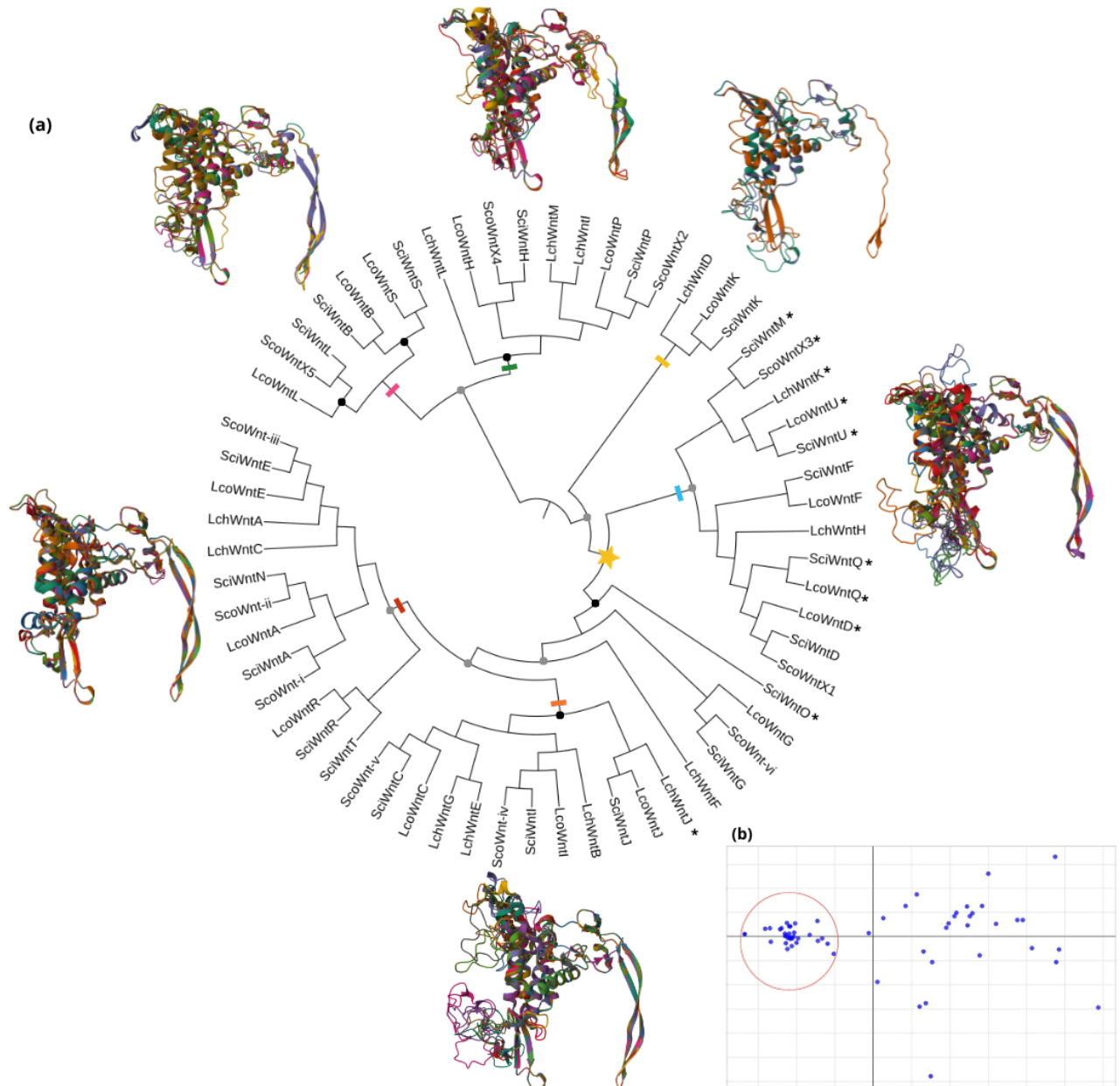


Figura 4 - Agrupamentos das proteínas WNT em Calcarea. (a) Filogenia das 62 proteínas WNT de esponjas calcárias utilizando dados de estruturas primárias com a representação nos seis clados das sobreposição de estruturas tridimensionais para cada clado, topologia da árvore e valores de suporte, C1-Rosa, C2-Verde, C3-Amarelo, C4-Azul, C5-Laranja, C6-Vermelho, (b) Destaque do agrupamento de estruturas terciárias similares obtidas via análise de componentes principais. A estrela na filogenia indica o nó que contém as 32 proteínas agrupadas por proximidade de estrutura terciária ressaltada em (b) [exceto as proteínas deste clado marcadas com asteriscos]. Espécies - Lch: *Leucetta chagosensis*; Lco: *Leucosolenia complicata*; Sci: *Sycon ciliatum*; Sco: *Sycon coactum*. Valores de suporte (*bootstrap*) para os nós mais externos - pontos pretos: maior que 0.85; pontos cinza: maior que 0.75.

Encontramos um grupo com 32 proteínas fortemente correlacionadas devido à proximidade entre as suas estruturas terciárias (Fig. 4b). Observando este grupo tridimensionalmente próximo dentre os seis clados de WNT em Calcarea, notamos que essas 32 proteínas estão majoritariamente dispostas em dois clados: o clado C5-Laranja

e o clado C6-Vermelho (Fig. 4a). A sobreposição das estruturas terciárias entre as 13 proteínas do clado C6-Vermelho gera uma conformação quase exata entre as estruturas, indicando uma forte manutenção dessa forma (Fig. 4a). Em contrapartida, o clado C1-Rosa com apenas sete proteínas sobrepostas tem uma conformação bem mais irregular (Fig. 4a).

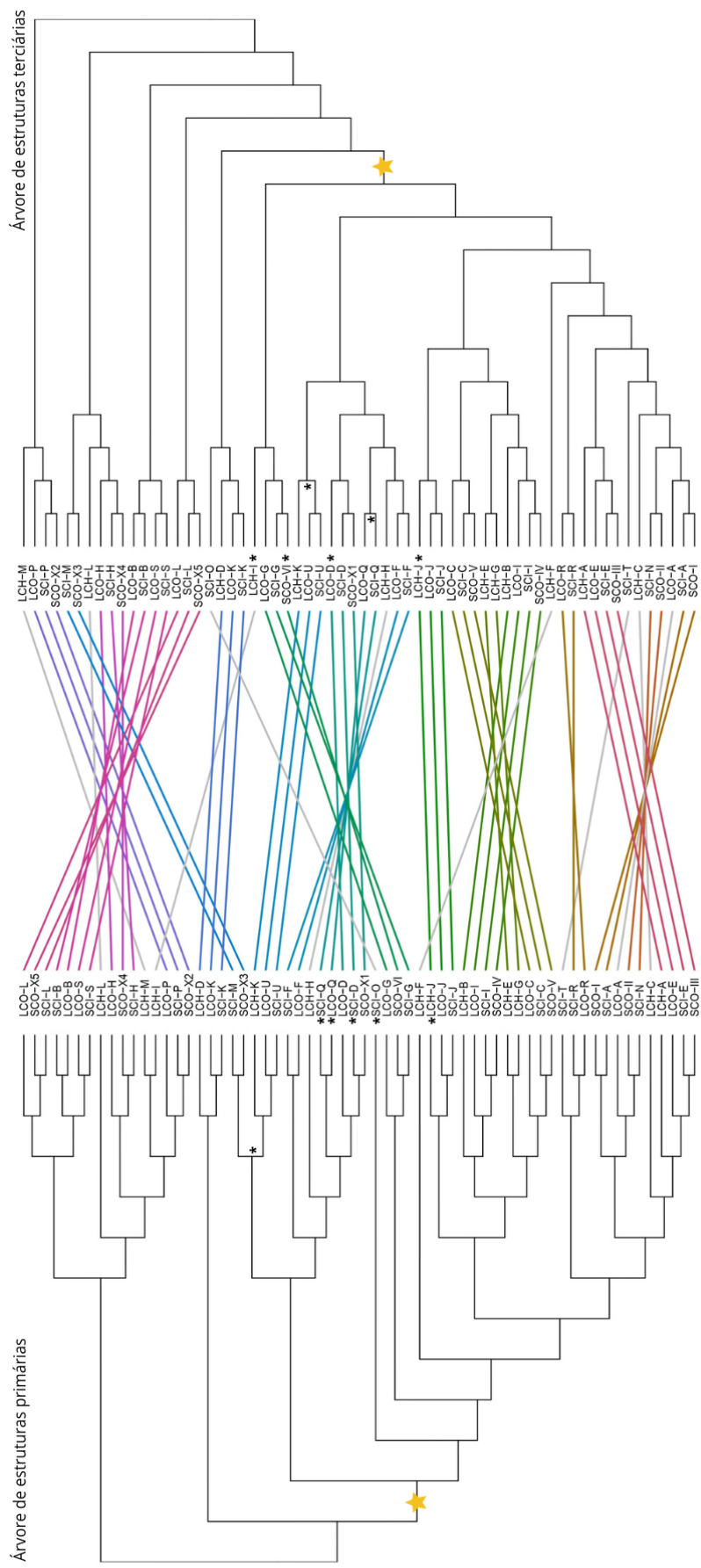


Figura 5 - Comparação dos cladogramas das proteínas WNT de Calcarea gerados a partir dos dados de estrutura primária (gerada no PhyML) em baixo e cladograma do alinhamento estrutural (gerado no

PROMALS3D) em cima. As estrelas representam o ramo mais externo a conter proteínas agrupadas por proximidade tridimensional (indicada na Fig. 4b) e os asteriscos representam proteínas do ramo que não estão neste grupo e quando apresentados no nó referem-se a todo o ramo. As conexões entre as filogenias refletem as posições dos nós terminais em cada filogenia, o esquema de cores é gerado aleatoriamente para facilitar a visualização e não está relacionado aos cladogramas anteriores. Espécies - LCH: *Leucetta chagosensis*; LCO: *Leucosolenia complicata*; SCI: *Sycon ciliatum*; SCO: *Sycon coactum*.

Padrões de conservação e substituição

Encontramos a conservação característica dos resíduos de Cisteínas nas estruturas primárias e a sua consequente formação de pontes dissulfeto nas estruturas terciárias para as 62 proteínas WNT da classe Calcarea. A partir do alinhamento das sequências de aminoácidos, identificamos 24 resíduos de Cisteína, 12 resíduos na região NTD (nas posições do alinhamento 277, 321, 347, 359, 413, 426, 428, 499, 590, 593, 599 e 604) e 12 resíduos na região CTD (nas posições do alinhamento 724, 740, 776, 780, 781, 796, 798, 810, 816, 822, 826 e 836) (Fig. 6a). Nove sítios foram identificados no perfil logos contendo resíduos de Cisteínas (C) com frequência 100% sendo eles: em NTD - 44, 108, 413, e em CTD - 724, 776, 780, 781 e 836. Outros quatorze sítios foram identificados no perfil logos contendo resíduos de Cisteína que apresentam frequência acima de 90%: em NTD - 277, 347, 359, 426, 428, 499, 590, 592, 599, 604, e em CTD - 796, 798, 822 e 826. Identificamos através do perfil logos e do alinhamento de aminoácidos que o motivo RWNC é parcialmente conservado (Tab. 1). Não há um padrão claro de perdas deste motivo entre as espécies ou subclasses, pois em todas as quatro espécies nas duas subclasses há proteínas com o motivo completo e proteínas com substituições e/ou adições entre os aminoácidos do motivo. Não há padrão de substituição para um aminoácido específico entre as proteínas que sofreram alterações no motivo. Notamos que, nessas 28 proteínas com alterações no motivo, o aminoácido W tende a ser mantido, sendo encontrado em 25 proteínas. Nos cladogramas delimitados na filogenia (Fig. 4a), a conservação do motivo RWNC não segue um padrão claro, o motivo é conservado integralmente em 5 de 7 proteínas em C1-Rosa, em 7 de 9 proteínas em C2-Verde, em 3 de 3 proteínas em C3-Amarelo, em 8 de 13 proteínas em C4-Azul, em 11 de 12 proteínas em C5-Laranja e em 7 de 13 proteínas no clado C6-Vermelho (Tab. 1). Identificamos, por meio do PFStats, 26 sítios conservados, com $pp \geq 0.8$ ao longo do alinhamento, desses, 19 sítios contêm resíduos de Cisteína e os demais contêm resíduos de Triptofano (um sítio), Ácido glutâmico (um sítio), Glicina (três sítios), Asparagina (um sítio) e Arginina (um sítio).

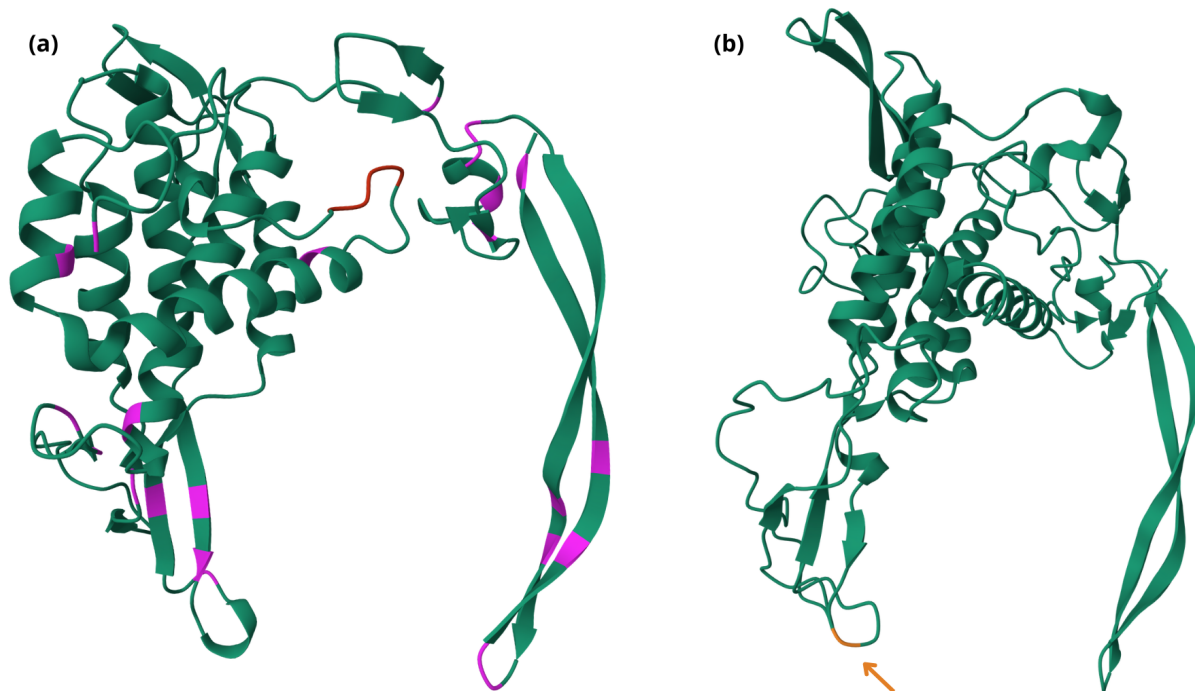


Figura 6 - Padrões de conservação de estrutura primária mapeados nas estruturas terciárias dos WNT de *Calcearia*. (a) Representação da estrutura terciária da proteína SCLwntA, resíduos de Cisteína em rosa, motivo RWNC em vermelho; (b) representação da estrutura terciária da proteína SCLwntD com destaque em laranja para a Fenilalanina no sítio 249 sem a formação da alça.

Para o funcionamento adequado do complexo WNT-Fz, uma Serina (Ser) localizada no domínio NTD deve estar presente para mediar a junção WNT-PAM. Na região terminal do NTD forma-se uma sub-estrutura côncava em que esse aminoácido ocupa a posição central. No WNT cristalizado de *Xenopus laevis*, a Serina está na posição 257 e é visível na superfície da molécula (Janda, *et al.*, 2012). Observamos a Serina localizada em NTD e acessível na superfície molecular para 34 das 62 proteínas (Tab. 1). Observamos também que em 20 proteínas há, na posição ideal da Serina, uma substituição por uma Treonina (Thr), ainda assim mantendo o formato de alça e a visibilidade na superfície molecular desse novo aminoácido (Tab. 1). A conformação mais comum encontrada nas alças que possuem a Treonina no centro é Glicina-Treonina-Leucina. Porém, quatro proteínas não tinham nem Serina nem Treonina na posição central da sub-estrutura côncava: SCOWntV com uma Cisteína no sítio 230, SCLwntD com uma Fenilalanina no sítio 249, SCLwntO com uma Prolina no sítio 140 e LCOwntC com uma Cisteína no sítio 244. Entretanto, apesar de não estar na posição central da sub-estrutura côncava que se liga diretamente ao PAM, SCLwntD possui uma Treonina no sítio 250 e SCLwntO possui uma Serina no sítio 139. Dessas quatro estruturas, três mantêm o formato que possibilita a ligação ao PAM com exceção de

SClwntD (Fig. 6b). Serina ou Treonina não aparecem conservadas nos resultados do PFStats (Material suplementar).

Na extremidade de uma das alfa-hélices do domínio NTD, próximo à região de interação WNT-PAM-Fz, há uma posição na estrutura tridimensional relacionada a interações de van der Waals na ligação WNT-Fz que contém resíduos de Arginina ou Lisina nas proteínas WNT dos animais (Janda, *et al.*, 2012). Observamos o resíduo Arginina nesta posição em 10 das 62 proteínas, o resíduo Lisina em 36 das 62 proteínas e nas demais 16 proteínas havia outro resíduo de aminoácido (Tab. 1). Essas 16 proteínas estão distribuídas em todas as quatro espécies, entretanto na espécie LCH apenas uma proteína apresentou essa substituição. Das 16 proteínas com ausência de Arginina ou Lisina nessa posição, 13 estão distribuídas fora do grupo tridimensionalmente correlacionado (i.e., estrela; Tab. 1, Fig. 4).

Tabela 1 - Sumário de resultados dos sítios significativos para manutenção da forma, resíduos de Cisteína, e sítios significativos para adição de ligantes na estrutura quaternária, WNT-PAM e WNT-Fz, por região da proteína WNT em *Calcarea*

Região da proteína	Domínio NTD				Região ligadora (Linker)	Domínio CTD		Clado na filogenia
	Resíduo na região terminal do NTD, ligação WNT-PAM	Resíduo na base da a-hélice, interação WNT-Fz	Motivo RWNC completo	Presença de aminoácidos sob coevolução de sítios da região 440 do alinhamento	Presença de aminoácidos sob coevolução de sítios da região 660 do alinhamento	Perdas no domínio CTD, estrutura β -hairpin	Grupo 4 de coevolução por conservação C-C: interação com Fz	C1-C6 na filogenia de estruturas primárias
<i>Leucetta chagosensis</i> (Calcinea)								
LCHwntA	Thr247	Lys242	Sim				C366, -	C6-Vermelho
LCHwntB	Thr246	Lys241	Sim	Sim			C366, C367	C5-Laranja
LCHwntC	Ser247	Lys242	Sim				C366, C367	C6-Vermelho
LCHwntD	Ser217	Arg212	Sim	Sim		Sim	C338, C339	C3-Amarelo
LCHwntE	Thr222	Lys217	Sim				C345, C346	C5-Laranja
LCHwntF	Ser294	Lys289	Sim	Sim			C433, C434	-
LCHwntG	Ser220	Lys215	Sim				C343, C344	C5-Laranja
LCHwntH	Ser273	Lys268		Sim	Sim		C405, C406	C4-Azul
LCHwntI	Ser218 Ser219	Arg213		Sim		Sim	C329, C330	C2-Verde
LCHwntJ	Thr140	Lys135		Sim				C5-Laranja

LCHwntK	Ser305	Lys300				Sim	P437, A442	C4-Azul
LCHwntL	Ser232	Asn227				Sim	H348, -	C2-Verde
LCHwntM	Ser256	Lys251				Sim	L368, T369	C2-Verde

Leucosolenia complicata
(Calcaronea)

LCOwntA	Thr217	Lys212					C335, C336	C6-Vermelho
LCOwntB	Ser209	Lys204	Sim			Sim	C329, C330	C1-Rosa
LCOwntC	C244	Lys239	Sim				C372, C373	C5-Laranja
LCOwntD	Thr251, Ser253	Lys246	Sim	Sim	Sim		C374, C375	C4-Azul
LCOwntE	Thr233	Lys228	Sim				C356, C357	C6-Vermelho
LCOwntF	Thr283, Ser284	Arg278					C440, C441	C4-Azul
LCOwntG	Ser277	Arg272	Sim				C402, C403	-
LCOwntH	Ser226	Lys221				Sim	C340, C341	C2-Verde
LCOwntI	Thr231	Lys226	Sim		Sim		C364, C365	C5-Laranja
LCOwntJ	Thr243	Lys238	Sim				C367, C368	C5-Laranja
LCOwntK	Ser220	Glu215	Sim	Sim		Sim	C344, C345	C3-Amarelo
LCOwntL	Ser310	Ser305				Sim	V446, A451	C1-Rosa
LCOwntP	Ser263	Val258		Sim	Sim	Sim	K394, K397	C2-Verde
LCOwntQ	Ser364	Arg359	Sim	Sim			C484, C485	C4-Azul
LCOwntR	Thr256	Lys251	Sim				C375, C376	C1-Rosa
LCOwntS	Ser215	Lys210	Sim				C346, C347	C6-Vermelho
LCOwntU	Ser158, Ser159	Arg153				Sim	G286, L289	C4-Azul

Sycon ciliatum (Calcaronea)

SCIwntA	Thr215	Lys210	Sim				C333, C334	C6-Vermelho
SCIwntB	Ser219	Thr214		Sim		Sim	D341, T342	C1-Rosa
SCIwntC	C230	Lys215	Sim				C358, C359	C5-Laranja
SCIwntD	Phe249, Thr250	Lys245	Sim	Sim	Sim		C392, C393	C4-Azul
SCIwntE	Thr256	Lys251	Sim				C378, C379	C6-Vermelho
SCIwntF	Ser289, Thr290	Gln283		Sim	Sim		C446, C447	C4-Azul
SCIwntG	Ser274	Arg269	Sim				C402, C402	-
SCIwntH	Ser207	Asp202	Sim			Sim	C318, C319	C2-Verde

SClwntI	Thr226	Lys221	Sim		Sim			C5-Laranja
SClwntJ	Thr256	Lys251	Sim	Sim			C380, C381	C5-Laranja
SClwntK	Ser251	Lys246	Sim	Sim		Sim	C377, C378	C3-Amarelo
SClwntL	Ser229	His224				Sim	M356, Y362	C1-Rosa
SClwntM	Ser238	Ser233				Sim	P360, W361	C4-Azul
SClwntN	Ser224, Phe225	Lys219					C349, C350	C6-Vermelho
SClwntO	Ser139	Gly135						-
SClwntP	Ser234	Phe229		Sim		Sim	- , W367	C2-Verde
SClwntQ	Ser224, Phe225	Arg219	Sim	Sim			C344, C345	C4-Azul
SClwntR	Thr264	Lys259	Sim	Sim			C382, C383	C6-Vermelho
SClwntS	Ser284	Asn279				Sim	C416, C417	C1-Rosa
SClwntT	Thr251	Lys246				Sim	C370, C371	C6-Vermelho
SClwntU	Ser206, Ser 206	Arg201			Sim	Sim	S341, V344	C4-Azul

<i>Sycon coactum</i> (Calcaronea)								
SCOwntI	Thr222	Lys217				Sim	C340, C341	C6-Vermelho
SCOwntII	Thr229	Lys224					C350, C352	C6-Vermelho
SCOwntIII	Thr225	Lys220					C347, C348	C6-Vermelho
SCOwntIV	Thr225	Lys220	Sim				C354, C355	C5-Laranja
SCOwntV	C230	Lys225	Sim				C356, C357	C5-Laranja
SCOwntVI	Ser272	Arg267	Sim	Sim	Sim		C399, C400	-
SCOwntX1	Thr249, Ser251	Thr224	Sim		Sim		C383, C384	C4-Azul
SCOwntX2	Ser239	Tyr234		Sim		Sim	K370, W373	C2-Verde
SCOwntX3	Ser214, Ser216	Ser209				Sim	P334, -	C4-Azul
SCOwntX4	Ser226, Ser228	Lys221	Sim			Sim	C336, C337	C2-Verde
SCOwntX5	Ser293, Ser295	Gln228				Sim	D419, Y423	C1-Rosa

Vinte e cinco modelos tridimensionais apresentaram perdas significativas no domínio CTD, sofrendo a descaracterização da dupla de folhas- β que formam a estrutura β -*hairpin* (Tab. 1, Fig. 7). Todas essas proteínas estão contidas nos clados C1-Rosa

(exceção de LCOwntS que apresenta as duas folhas- β), C2-Verde, C3-Amarelo e parte do C4-Azul (exceção de SClwntM, SCOwntX3, LCHwntK, LCOwntU e SClwntU) (Fig. 4). Novamente, não há sinal filogenético para essas perdas que ocorrem em todas as espécies: em SCI 9 das 21 cópias apresentaram perdas, em LCO 6 de 17 cópias apresentaram perdas, em SCO 5 das 11 cópias apresentaram perdas, e em LCH 5 das 13 cópias apresentaram perdas.

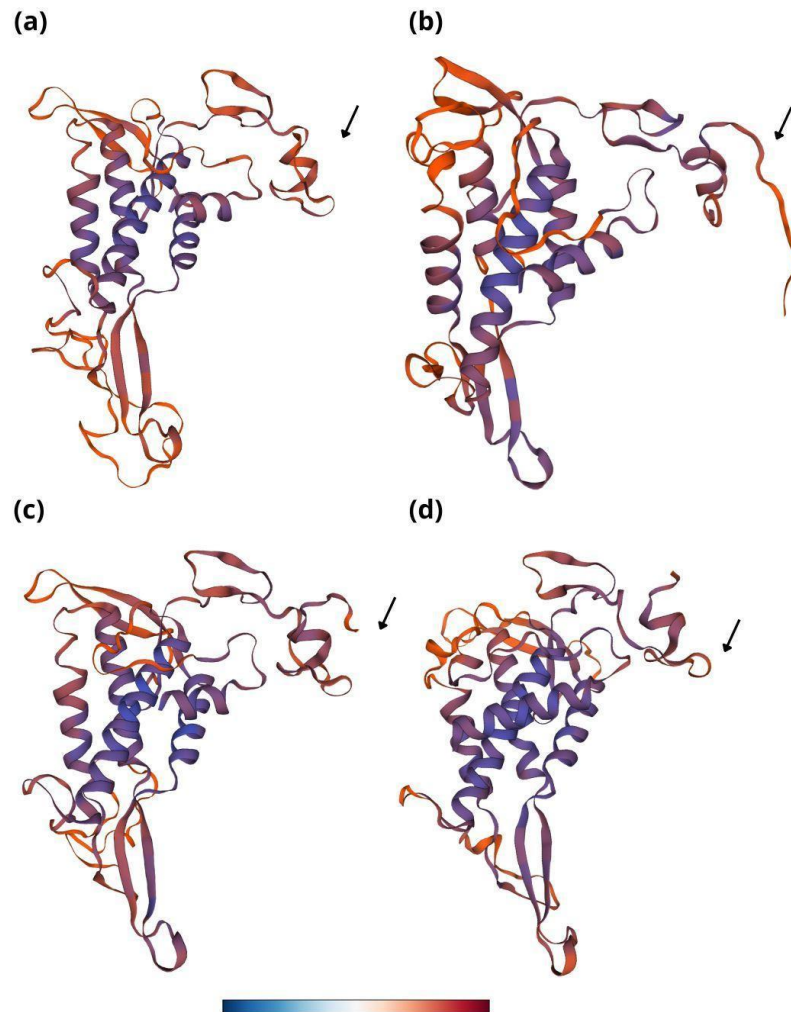


Figura 07- Casos extremos de perdas no domínio CTD (index finger) nos modelos tridimensionais das proteínas WNT de *Calcareia*. (a) LCHwntD; (b) LCHwntI; (c) LCOwntK; e (d) LCOwntU. A escala de cores é referente ao intervalo de confiabilidade global da modelagem, em que azul é maior confiabilidade e vermelho menor confiabilidade. As setas indicam a região onde se esperaria o domínio CTD.

Apesar da grande conservação encontrada nos sítios que contém resíduos de Cisteína, algumas proteínas exibem substituições pontuais nesses sítios onde esperaríamos a presença desse resíduo. Dois sítios do alinhamento se destacam, C530 e C534, com substituições em seis proteínas. Em C530: LCOwntS sofre deleção, SClwntH e SCOwntX4 sofreram substituição para Treonina, LCOwntH sofre substituição para

Valina, LCHwntI sofre substituição para Lisina, LCHwntM sofre substituição para Isoleucina. Em C534: SCOWntX2 sofre uma substituição para Leucina, SClwntP sofre substituição para Metionina, LCOwntP sofre substituição para Alanina, SClwntS sofre substituição para Isoleucina, LCOwntS sofre deleção, LCOwntH sofre substituição para Isoleucina. Essas proteínas com alterações em C530 e C534 estão nos clado C1-Rosa e C2-Verde que possuem estruturas tridimensionais menos ordenadas na sobreposição. Desse clado, as proteínas SCOWntX2, SCOWntX4, SClwntH, SClwntP, SClwntS, LCHwntM, LCOwntH, LCOwntP sofreram também perdas significativas no domínio CTD.

Padrões de coevolução

Encontramos 111 pares de aminoácidos coevoluindo com $pp \geq 0.95$ por meio da análise BGM. Nessa análise, o padrão de coevolução observado se dá pelo viés de substituição conjunta, ou seja, quais são as chances de um par de sítios sofrer mudanças em seus aminoácidos ao mesmo tempo e de maneira relacionada. Dessa forma, esses sítios apresentaram grande variedade de aminoácidos ao longo do alinhamento. Partindo desses pares, notamos que alguns sítios se repetiam mais que outros, a exemplo, o sítio 830 aparece nove vezes pareado aos sítios 326, 385, 432, 563, 572, 606, 633, 741 e 786 (Fig. 7a-b). Com esse mesmo padrão, outros dois sítios se destacaram: 338 e 385. Além disso, duas regiões apresentam pares intercalados: a região 440 tem os pares 442-443, 443-446 e 444-447 (Fig. 7c); a região 660 tem os pares 661-662, 662-663, 661-664 e 662-665 (Fig. 7b). Os domínios CTD e NTD apresentaram pares de sítios coevoluindo intra e inter-regiões, a exemplo do par 830-432 (CTD-NTD, respectivamente), mas os pares de sítios na região ligadora (Linker) coevoluem apenas entre si, a exemplo 662-665 (Linker-Linker) (Fig. 7b). Ao longo do alinhamento, a maior parte das 62 proteínas não apresenta aminoácidos nas regiões 440 e 660, ou seja, ao longo dos sítios 440 e 660 no alinhamento, a maior parte das proteínas estudadas apresentam *gaps*, ausência de aminoácidos. Vinte proteínas das quatro espécies apresentam aminoácidos nos sítios da região 440 do alinhamento, doze proteínas das quatro espécies apresentam aminoácidos nos sítios da região 660 do alinhamento e apenas seis apresentam ambas as regiões (Tab. 1). Não há um padrão no agrupamento dessas proteínas nas filogenias de estruturas primárias ou terciárias, elas estão distribuídas ao longo dos seis clados (Fig. 4a). Notamos que essas proteínas que apresentaram aminoácidos coevoluindo na região 440 do alinhamento também demonstraram um prolongamento na região NTD, como em SClwntF (Fig. 7c).

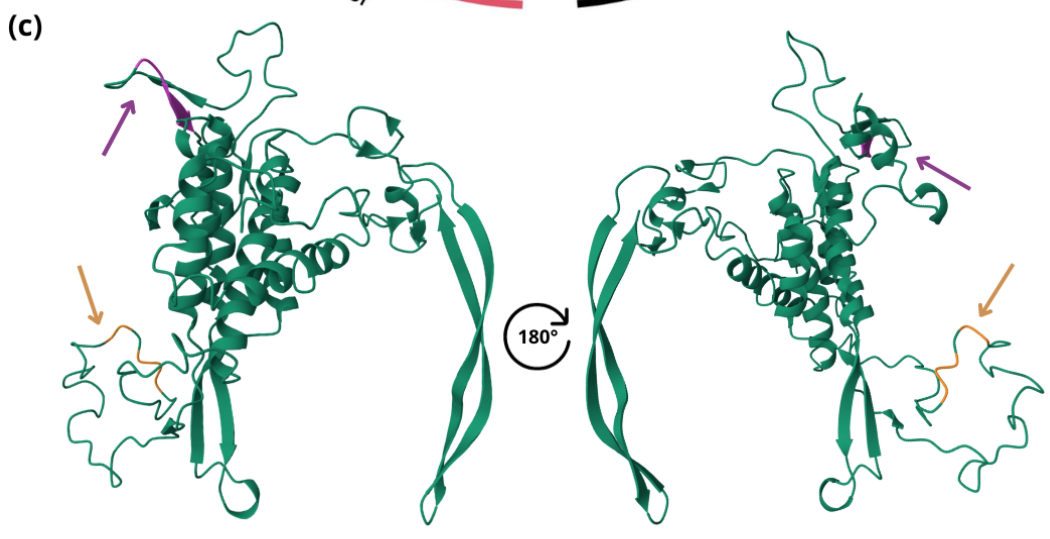
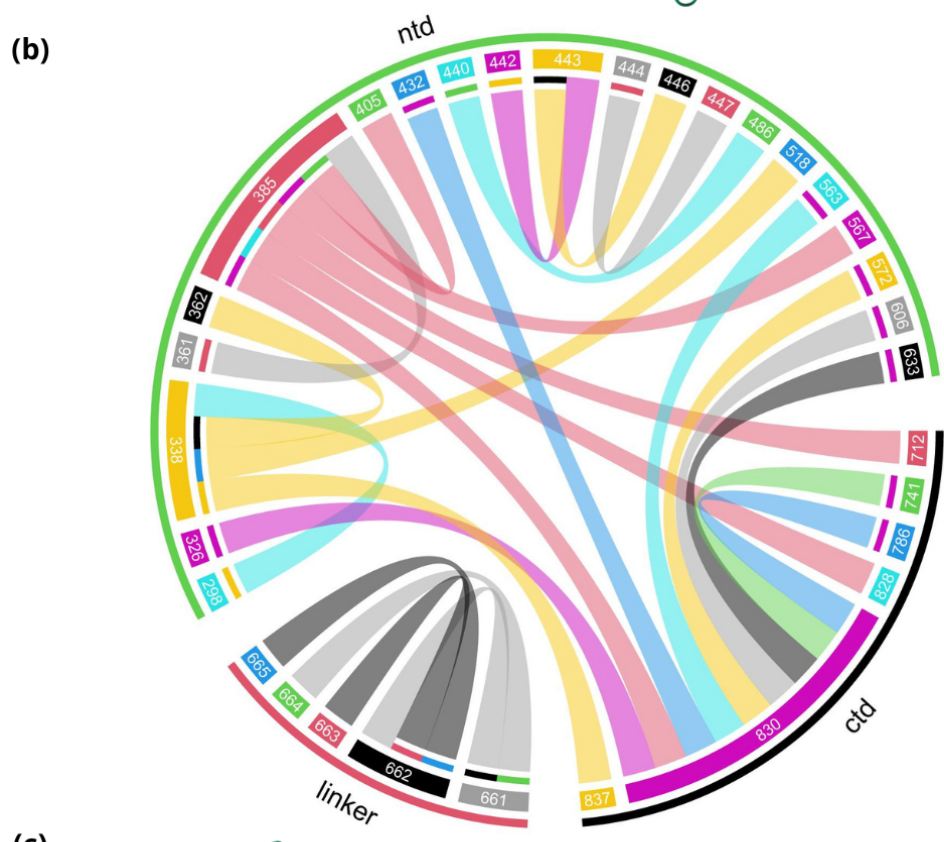
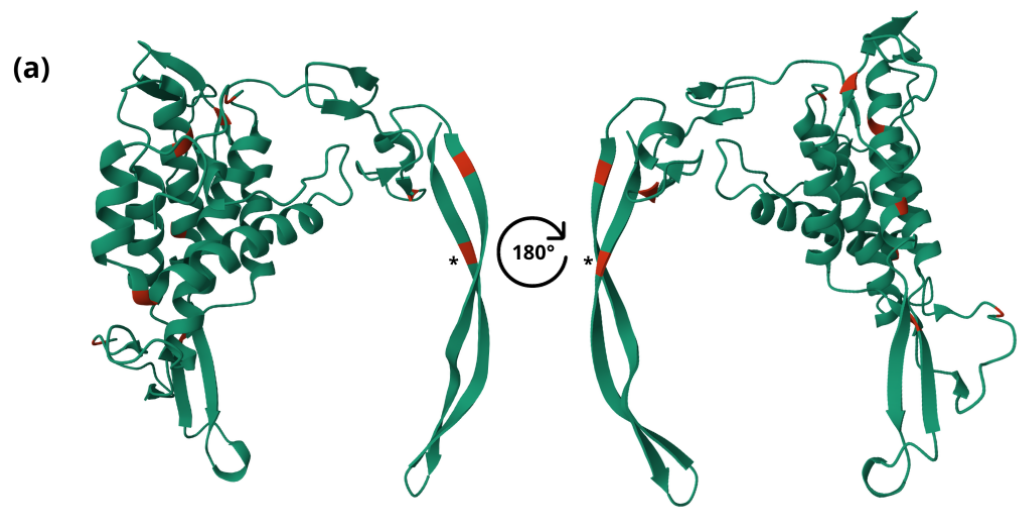


Figura 8 - Visualização de alguns sítios de aminoácidos coevoluindo em modelos tridimensionais das proteínas de WNT de *Calcarea*. (a) Pares, em vermelho, do sítio 830 do alinhamento (asterisco) na proteína SClwntA; (b) Circos-plot demonstrando os sítios coevoluindo por viés de substituição conjunta; (c) Regiões do alinhamento: 440, destaque em laranja no sítio (seta laranja), e 600, destaque em rosa no sítio (seta rosa) na proteína SClwntF.

Encontramos seis grupos de aminoácidos coevoluindo por conservação, ou seja, sendo mantidos de forma conjunta em determinados sítios do alinhamento com $pp \geq 0.8$. Esses seis grupos não são relacionados aos clados C1-C6 mencionados anteriormente. Esse padrão de coevolução sugere que esses sítios são necessários à manutenção da forma e da função da proteína. Denominados de G1 a G6, os grupos de aminoácidos coevoluindo nos sítios do alinhamento são: (Grupo 1 - G1) Ácido glutâmico 110, Glicina 1425, Glutamina 115, Fenilalanina 1416; (Grupo 2 - G2) Fenilalanina 2773, Glicina 102, Glicina 476; (Grupo 3 - G3) Ácido aspártico 943, Asparagina 1296, Asparagina 1292; (Grupo 4 - G4) Cisteína 2781, Cisteína 2787; (Grupo 5 - G5) Leucina 53, Glutamina 60, e, (Grupo 6 - G6) Ácido aspártico 921, Asparagina 923 (Fig. 9). Ressaltamos que nenhuma das 62 proteínas WNT de *Calcarea* possuíam simultaneamente de forma integral todos os aminoácidos de todos os seis grupos indicados anteriormente nos sítios sob coevolução, ou seja, em nenhuma das 62 proteínas há simultaneamente todos os aminoácidos dos grupos G1-G6. Por exemplo, algumas proteínas não apresentaram G1-G6 integralmente devido a substituições: SClwntH e SCOWntX4 não apresentaram os aminoácidos Ácido aspártico 943-Asparagina 1296-Asparagina 1292 do grupo G3, pois apresentaram outros aminoácidos nesses sítios. Além disso, notamos que no grupo G4, as mesmas proteínas que sofrem substituição da dupla de Cisteínas também têm perda da dupla de folhas- β no domínio CTD. Outro exemplo é o de perdas de aminoácidos nos sítios coevoluindo demarcados em G1-G6, como em LCOwntU em que há perda dos dois aminoácidos do grupo G5.

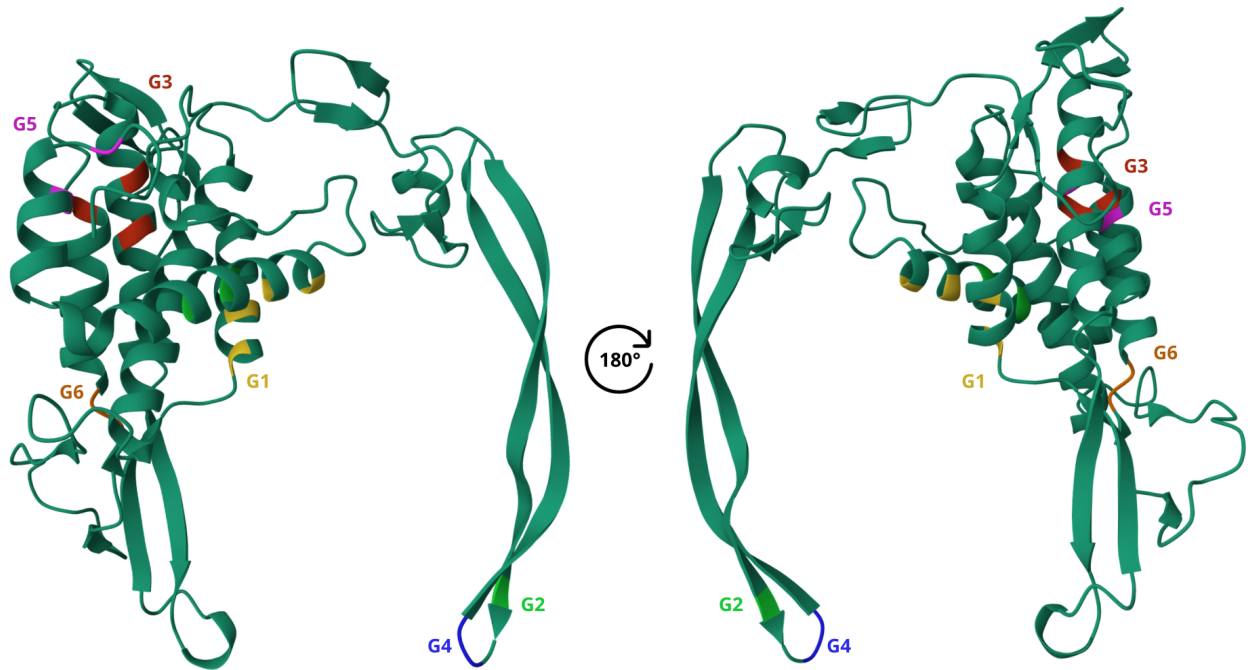


Figura 9 - Visualização dos seis grupos de aminoácidos coevoluindo por conservação nas proteínas WNT de Calcarea, exemplificada no modelo tridimensional da proteína SCLwntA. Destacamos que existem substituições nos sítios 115 e 1416 do alinhamento em G1. (Q115R, F1416V) e substituições nos sítios 1296 e 1292 do alinhamento em G3 (N1296H, N1292A). G1 - amarelo, G2 - Verde claro, G3 - Vermelho, G4 - Azul, G5 - Rosa, G6 - Laranja.

DISCUSSÃO

A via de sinalização do WNT é essencial para o desenvolvimento animal e está relacionada à padronização corporal (Willert, Nusse, 2012). A proteína WNT é restrita aos metazoários e aparece no primeiro filo a divergir na linhagem animal, os poríferos, no qual vem sendo associada a funções do desenvolvimento, como a polarização de seu sistema aquífero (Adell, *et al.*, 2007; Adamska, *et al.*, 2010; Lapébie, *et al.*, 2009; Windsor, Leys, 2010; Leininger, *et al.*, 2014; Riesgo, *et al.*, 2014, Borisenko, *et al.*, 2016, Windsor-Reid, *et al.*, 2018, Borisenko, *et al.*, 2021). Uma das classes do grupo apresenta um número de cópias WNT elevado em relação às demais espécies estudadas, mesma classe que possui características do desenvolvimento únicas em relação às demais esponjas, fatos que instigam o estudo dessas proteínas nas duas subclasses desse grupo (Leininger, *et al.*, 2014; Riesgo, *et al.*, 2014). Neste trabalho, nós respondemos à pergunta ‘*Quais inferências evolutivas podem ser feitas a partir da comparação da diversidade de estruturas primárias e da diversidade de estruturas terciárias das proteínas WNT?*’. Descrevemos, pela primeira vez para a subclasse Calcinea, a presença de 13 cópias WNT na espécie *Leucetta chagosensis*. Delimitamos seis clados (C1-C6) de proteínas WNT através da topologia da filogenia de estruturas primárias e sobreposição de estruturas terciárias, também identificamos dois grandes agrupamentos de proteínas WNT em Calcarea quanto à manutenção da forma tridimensional - um que mantém a forma “tradicional” e outro que apresenta alterações. Identificamos padrões de conservação de aminoácidos, como a manutenção dos sítios relacionados à manutenção da forma na proteína WNT em Calcarea, os resíduos de Cisteína; e padrões de coevolução de aminoácidos, como as alterações na região ligadora. Nossos resultados apoiam a ideia de que em Calcarea há mais diversidade de estruturas primárias do que de estruturas terciárias.

Caracterização, diversidade e evolução das proteínas WNT na classe Calcarea

As proteínas WNT da classe Calcarea apresentam o número esperado de aminoácidos, 350-400, e as características definidoras da proteína: 22-24 resíduos de Cisteína e o motivo RWNC parcialmente conservado (Croce, McClay, 2009; Janda, *et al.*, 2012; Windsor-Reid, *et al.*, 2018). Das 62 proteínas estudadas, 57 têm mais de 350 aminoácidos, o que sugere adequação e consistência na obtenção de dados, pois a qualidade dos transcritos e a obtenção de Open Reading Frames (ORFs) pode impactar

negativamente na completude da sequência de aminoácidos das proteínas causando limitações nas análises. O padrão de Cisteínas e o motivo RWNC parcialmente conservado estão presentes nas espécies deste estudo. As espécies da subclasse Calcaronea - *Sycon ciliatum* (SCI), *Sycon coactum* (SCO) e *Leucosolenia complicata* (LCO), cuja identificação se deu previamente em estudos já publicados, e uma espécie da subclasse Calcinea - *Leucetta chagosensis* (LCH), cuja identificação se deu por meio de similaridade de sequências entre os transcritos da espécie e os WNT identificados em outra esponja, *Amphimedon queenslandica* (Adamska, *et al.*, 2010; Leininger, *et al.*, 2014, Windsor-Reid, *et al.*, 2018). Apresentamos a primeira identificação de WNT na subclasse Calcinea para espécie *L. chagosensis* que contém 13 cópias de WNT com padrão de resíduos de Cisteína conservados e motivo RWNC parcialmente conservado.

Um primeiro resultado importante do presente estudo diz respeito a esse número de cópias elevado de WNT na classe Calcarea. Partindo da presente nova identificação de proteínas WNT em uma espécie da subclasse Calcinea, que possui número similar de cópias ao identificado previamente para as espécies da subclasse Calcaronea, podemos indicar que a expansão do número de cópias se deu antes da separação das duas subclasses. Este número é muito diferente do observado nas outras classes de esponjas. Enquanto em Hexactinellida nenhum WNT é observado no genoma (Santini, *et al.*, 2023), nas classes Homoscleromorpha e Demospongiae não foram identificadas, até então, espécies com número de cópias WNT tão expressivos quanto a média apresentada na classe Calcarea, apesar de existirem duas exceções - a homoscleromorfa *Oscarella carmela* com oito cópias e a demosponja *Halisarca dujardini* com dez cópias (Leininger, *et al.*, 2014; Borisenko, *et al.*, 2018). No entanto, há chances de que os números de cópias de WNT possam ser diferentes em todas as classes. Enquanto em Calcarea apenas a espécie *Sycon ciliatum* teve a identificação de WNT a partir de seu genoma, para as três demais espécies a identificação foi feita a partir de transcriptomas. O mesmo se aplica às outras classes, pois há poucos genomas publicados para as demosponjas e homoscleromorfas, mais uma vez podendo haver mais cópias ainda não identificadas nas espécies das duas classes. Entretanto, a demosponja *Amphimedon queenslandica*, uma espécie modelo no grupo Porifera para “ômicas”, que possui seu genoma publicado e relativamente bem estudado, apresenta apenas três cópias WNT identificadas (Adamska, *et al.*, 2010).

Com o presente trabalho, verificamos mais uma similaridade que une os WNT das esponjas aos demais metazoários: a estrutura tridimensional típica do WNT nos demais metazoários é encontrada nos WNT da classe Calcarea por meio de modelagem por homologia. Das 62 proteínas estudadas, obtivemos modelos com identidade molde-modelo superior a 30% para 56 delas, como recomendado pela plataforma SwissModel, o que nos dá segurança para comparar os modelos obtidos neste estudo (Kiefer, *et al.*, 2009). Com o objetivo de evitar que artefatos estruturais surgissem em resposta ao uso de moldes diferentes, optamos por utilizar um único molde: o WNT de *Homo sapiens*, PDB:6AHY (Hirai, *et al.*, 2018). Apesar da grande distância evolutiva entre esponjas e humanos, o maior valor de identidade para a modelagem das proteínas WNT de Calcarea pode ser atribuído à melhor resolução de cristalização do WNT de humanos. Isso significa que a estrutura tridimensional da proteína dos humanos é mais bem resolvida (a proteína WNT humana cristalizada tem uma resolução de 2.8 Angstrom (Å)), se tornando um molde mais adequado para uma sequência de aminoácido do que uma outra proteína com menor resolução. Além disso, existe um número limitado de proteínas WNT acessadas experimentalmente para os animais como um todo, além dos humanos, há cristalizações para *Drosophila melanogaster* e *Xenopus laevis*, dois organismos que também são evolutivamente distantes das esponjas (Janda, *et al.*, 2012; Chu, *et al.*, 2013; Hirai, *et al.*, 2018). Os modelos tridimensionais das proteínas WNT em Calcarea apresentam a forma característica dos WNT nos demais animais: os dois domínios CTD e NTD e uma região ligadora. A forma típica da proteína WNT, que se dá pela formação de pontes dissulfeto em decorrência da disposição de resíduos de Cisteína ao longo de toda sua extensão, está presente em todas as proteínas WNT de Calcarea modeladas (Janda, *et al.*, 2012). Esses resíduos são conservados em sítios específicos das proteínas WNT desde as esponjas calcárias até metazoários mais derivados, reforçando a ideia de origem ancestral do gene WNT a todos os metazoários. Forma e função estão estritamente relacionadas nas proteínas. Devido às diversas funções base no desenvolvimento desempenhadas pela proteína WNT é esperado que a sua forma seja conservada ao longo da evolução (Bhattacharya, *et al.*, 2017). Portanto, identificar a manutenção da forma da proteína em uma classe de esponjas é um indicativo de que essas proteínas WNT podem atuar de forma semelhante nas esponjas como agem nas vias WNT de outros animais.

Dividimos seis clados de proteínas WNT na classe Calcarea, a partir da filogenia de estruturas primárias, por meio da topologia da árvore e dos valores de bootstrap. A composição dos clados não é restrita entre as espécies ou subclasses uma vez que um mesmo clado apresenta proteínas das diferentes espécies da subclasse Calcaronea e Calcinea, e.g., C4-Azul. Apenas C1-Rosa é composto exclusivamente por proteínas de espécies de uma única subclasse, Calcaronea. Entretanto, não há uma implicação direta de origem exclusiva desse clado nessa subclasse ou de uma perda secundária para a subclasse Calcinea. Para tal afirmação, há a necessidade uma amostragem maior para Calcinea, pois o não agrupamento dessa subclasse no clado C1-Rosa, pode ser particular às proteínas WNT de *L. chagosensis*. Como ocorre, por exemplo, no clado C3-Amarelo, composto por Calcaronea e Calcinea, em que há 3 das 4 espécies, mas não há proteínas WNT de *S. coactum*. A partir da divisão desses clados na filogenia de estruturas primárias, nós sobreposamos as estruturas terciárias a fim de verificar o alinhamento dos modelos no espaço tridimensional. Notamos que os clados C1-Rosa, com sete proteínas, e C2-Verde, com nove proteínas, apresentam um elevado grau de desordem entre as sobreposições na região central da molécula e especialmente na região NTD, onde forma-se um emaranhado entre as estruturas das proteínas. O mesmo ocorre de forma ainda mais destacada em C4-Azul, com 13 proteínas, também na região central e na região CTD. Uma das possíveis explicações para essa observação seria o relaxamento da seleção purificadora sobre essas cópias. Isso permitiria a manutenção de mudanças no gene, como inserções, resultando nesses aminoácidos que não se ajustam adequadamente na estrutura tridimensional através da modelagem nessas regiões e a posterior desordem no alinhamento estrutural. Após um evento de expansão do número de cópias de um gene, quando há redundância funcional entre as cópias, as limitações impostas à mutações, como as inserções e deleções, é reduzida para parte das cópias, aumentando as chances de diferenciação (Pointing, Russell, 2003). O clado C5-Laranja, com 12 proteínas, apresenta a desordem de sobreposição na região NTD e na região ligadora, porém as regiões CTD e central da molécula estão bem alinhadas. Em contraponto, o clado C6-Vermelho, com 13 proteínas, apresenta um alinhamento no espaço tridimensional bastante ordenado. Propomos que essa ordem está relacionada à conservação restrita da estrutura tridimensional. A seleção purificadora, também definida como restrição funcional, atua restringindo as mutações aceitáveis em um sítio de forma que não haja prejuízo à função e estrutura de um gene ou proteína (Graur, Li, 2000). O resultado da ação da seleção purificadora intensa sob um gene é a diminuição da taxa de

ocorrência de substituições (Graur, Li, 2000). Podemos inferir indiretamente que as proteínas WNT que são estruturalmente mais conservadas estão sob ação desse tipo de seleção, resultando na manutenção de sua forma e, conseqüentemente, função.

Esses desdobramentos evolutivos em decorrência da expansão de cópias podem ser verificados com o aparecimento de dois grandes agrupamentos das proteínas quanto à forma tridimensional, outra evidência, encontrada neste trabalho, de que em *Calcarea* há mais diversidade de estrutura primária do que estrutura terciária. Podemos dividir as proteínas WNT da classe *Calcarea* em um grupo estruturalmente muito próximo e um grupo de estruturas terciárias que ficam dispersas, a partir de uma PCA. Ao mapear o grupo estruturalmente muito próximo na filogenia de estruturas primárias, notamos que essas proteínas estão majoritariamente contidas nos grupos C5-Laranja e C6-Vermelho, clados que consideramos anteriormente como estruturalmente conservados em relação às proteínas WNT cristalizadas de outros metazoários. O agrupamento de proteínas tridimensionalmente distantes estão contidas majoritariamente nos clados C1-Rosa, C2-Verde, C3-Amarelo e C4-Azul. Essas duas observações reforçam a ideia de estruturas terciárias mais conservadas nessas 32 proteínas, enquanto as demais 30 proteínas sofrem acúmulos e perdas estruturais que as tornam menos correlacionadas tridimensionalmente e menos conservadas estruturalmente. Uma hipótese para a formação desses dois agrupamentos de proteínas WNT na classe *Calcarea* é que a expansão do número de cópias permitiu o relaxamento da seleção purificadora para um grupo dessas proteínas, levando à manutenção de mutações na região NTD com a conseqüente possibilidade de perda da forma e função tradicional associada à essas proteínas na via WNT (Grauer, Li, 2000). Outra possibilidade é contrária a anterior: a expansão do número de cópias permitiu que um grupo mantivesse a função original da proteína enquanto outro grupo pudesse ganhar nova função a partir dessas mudanças (Grauer, Li, 2000). Entretanto, essas inferências são indiretas e para confirmar perda ou ganho de função, estudos que investiguem a ação desses genes, através da dinâmica de expressão da via WNT devem ser realizados em espécies calcárias de forma comparativa a outras esponjas que não têm essa expansão. Por exemplo, investigar se existem diferenças significativas no local e na duração de expressão entre cópias WNT contidas no grupo conservado estruturalmente e cópias que sofrem alterações estruturais para *Calcarea* e comparar esses padrões às espécies demosponjas que apresentam menos cópias.

Outra observação quanto à diversidade de WNT na classe Calcarea está relacionada à perda de parte do domínio CTD em algumas proteínas. Das 62 proteínas, 25 apresentaram perdas significativas no domínio CTD, dessas cinco proteínas tiveram valor de identidade molde-modelo inferior a 30%. Como essa quantidade de proteínas foi relativamente baixa, podemos, portanto, associar essas perdas à evolução no grupo e que isso não foi um artefato da modelagem. As proteínas que apresentam esse padrão estão nos clados C1-Rosa, C2-Verde, C3-Amarelo e C4-Azul, os mesmos clados que contém proteínas estruturalmente distantes e que apresentam alterações na região NTD. É plausível associar a perda de partes do domínio CTD, que participa diretamente da interação WNT-Fz, a possibilidade de perda de função dessas proteínas WNT nas vias em que ela atuaria normalmente. Notamos que as proteínas que possuem essa perda encontram-se neste segundo agrupamento das proteínas mais distantes tridimensionalmente. Podemos considerar esse grupo como menos conservado, em função das perdas no domínio CTD e acúmulos e perdas estruturais no domínio NTD.

Os WNT de Porifera não se agrupam fortemente às 13 subfamílias WNT dos demais metazoários (ver Lapébie, *et al.*, 2009; Adamska, *et al.*, 2010; Leininger, *et al.*, 2014; Borisenko, *et al.*, 2016; Windsor-Reid, *et al.*, 2018). Em uma tentativa anterior de agrupar os WNT das esponjas em subfamílias específicas, Windsor-Reid e colaboradores (2018) dividiram as subfamílias: PorWntA, PorWntB e PorWntC que agrupam os WNT das espécies de demosponjas, WntI e WntII que agrupam os WNT das homoscleromorfos e CalcareaWnt (CwntA/N-CwntU) que agrupa os WNT das espécies calcárias. A delimitação desses 20 agrupamentos (CwntA-CwntU) para os WNT das espécies calcárias foi feita apenas com espécies da subclasse Calcaronea (SCI, SCO, LCO) (Windsor-Reid, *et al.*, 2018). Em nossa filogenia de estruturas primárias, nós recuperamos todas essas subfamílias. Os grupos CwntA/N, CwntB, CwntD, CwntF, CwntG, CwntH, CwntI, CwntL, CwntM, CwntO, CwntR, CwntS e CwntT foram exatamente iguais em composição de proteínas e o agrupamento entre os ramos terminais. Nos nós mais internos, esses 12 agrupamentos apresentavam apenas proteínas WNT da subclasse Calcaronea. Nós recuperamos um ramo mais bem resolvido, (LCOwntG(SCIwntG,SCOWntVI)), em comparação à politomia do grupo CwntG, delimitado por Windsor-Reid e colaboradores (2018). Os grupos CwntC, CwntE, CwntF, CwntJ, CwntK, CwntP, CwntQ e CwntU mantiveram a composição de proteínas WNT das espécies de Calcaronea, mas apresentaram também proteínas WNT de *L. chagosensis*

(Calcinea). Em todos esses oito ramos compatíveis com esses oito agrupamentos das subfamílias, as proteínas WNT de LCH agruparam mais externamente em relação aos ramos das proteínas de espécies de Calcareonea. A inserção de proteínas de LCH (Calcinea) entre alguns grupos da subfamília CalcareaWnt reforça a ideia de que a expansão de WNT em Calcarea ocorreu antes da divergência entre as subclasses. A filogenia apresentada por Windsor-Reid e colaboradores (2018) é pouco resolvida nos nós mais externos, mas há dois ramos mais bem resolvidos: o agrupamento de CwntE, CwntA/N, CwntT, CwntR, CwntJ, CwntI, CwntC e CwntG, aqui denominado “(a)”; e o agrupamento de CwntD, CwntQ e CwntF, aqui denominado “(b)”. Mesmo com a baixa resolução dos nós mais externos na filogenia de estruturas primárias aqui apresentadas, recuperamos integralmente os dois ramos (a) e (b) com valor de suporte superior a 0.75. O agrupamento (a), em nossa filogenia, corresponde ao nó que une C5-Laranja e C6-Vermelho, dois clados que consideramos bem conservados quanto à estrutura terciária. O clado C4-Azul de nossa filogenia de estrutura primária, com valor de suporte superior a 0.75, contém o agrupamento (b) junto aos correspondentes de CwntM e CwntU. Apesar dessa compatibilidade, os ramos que representam os agrupamentos da subfamília não necessariamente representam os clados C1-C6 que delimitamos aqui, cada clado C1-C6 pode conter mais de um desses 20 agrupamentos da subfamília CalcareaWnt. Na segunda filogenia gerada neste estudo, relacionando as estruturas terciárias das 62 proteínas, não recuperamos exatamente os nós mais externos que correspondem aos seis clados (C1-C6) mencionados anteriormente, entretanto as relações entre as proteínas contidas dentro desses clados se mantêm, por exemplo, as proteínas dos clados C5-Laranja e C6-Vermelho se mantêm associadas em ambas as filogenias. Os diferentes agrupamentos dos nós mais externos podem ser associados ao uso de dados e métodos distintos para recuperar as duas filogenias. Nessa filogenia de estruturas terciárias, recuperamos também os nós mais internos que correspondem em partes às subfamílias CalcareaWnt delimitadas por Windsor-Reid e colaboradores (2018). A maior divergência encontrada entre as duas filogenias geradas aqui em comparação às subfamílias CalcareaWnt, foi a posição da subfamília CwntO - composta apenas por SClwntO - que é isolada em relação às demais na filogenia de estruturas primárias e também no estudo de Windsor-Reid e colaboradores (2018). Porém, na filogenia de estruturas terciárias gerada neste estudo, CwntO aparece agrupada à CwntD. Ao tratar de proteínas com tempo de divergência muito antigo, ou seja, relações evolutivas de origem muito distante, alinhamentos de aminoácidos podem não ser estatisticamente acurados, já que

a divergência evolutiva da estrutura primária ocorre de forma muito mais rápida em relação a estrutura terciária (Johson, *et al.*, 1990a, Thorne, *et al.*, 1996). Além disso, dados de estruturas terciárias abrigam mais informações sobre uma proteína do que dados de estruturas primárias (Johson, *et al.*, 1990b). Nesse sentido, comparar as estruturas terciárias pode ser uma solução para recuperar relações evolutivas mais fidedignas (Johson, *et al.*, 1990a). Isso poderia indicar que filogenias de estrutura terciária para as proteínas WNT têm maior capacidade de recuperar as similaridades entre as subfamílias de Porifera e podem ser utilizadas para recuperar as relações entre os WNT de Porifera e dos demais metazoários. Ademais, entre as subfamílias deve haver maior diversidade de composição de aminoácidos do que de estruturas terciárias, reforçando a hipótese de que em *Calcarea* há mais diversidade de estrutura primária do que estrutura terciária.

Ao não se agruparem fortemente aos demais WNT dos metazoários, os WNT de Porifera devem, portanto, apresentar peculiaridades que surgiram no grupo ao longo de milhões de anos de evolução após a divergência na linhagem dos metazoários. Identificamos uma dessas peculiaridades a nível de estrutura terciária, uma substituição de aminoácidos numa área de interação WNT-Frizzled. Encontramos nesse trabalho, em 20 proteínas nos seis clados (C1-C6) para todas as quatro espécies nas duas subclasses, a substituição do aminoácido Serina pelo aminoácido Treonina na região NTD onde há ligação com o ácido palmitoléico. Apesar de Serina e Treonina serem aminoácidos polares e possuírem um grupo hidroxila (-OH), esses são aminoácidos com chance de permutabilidade reduzida, ou seja, a chance de ocorrerem substituições que levam à mudança Serina->Treonina é reduzida, segundo o método de Argyle (Grauer, Li, 2000). Não devemos, entretanto, descartar a possibilidade de que o aminoácido Treonina permita a ligação WNT-Fz. A confirmação dessa ligação só pode ser feita a partir de um estudo específico da dinâmica dessa ligação. Não há marcação filogenética para origem dessa substituição entre as proteínas de *Calcarea*, o que implica em uma origem ancestral - ao menos - à divisão das duas subclasses do grupo. Para compreender a origem dessa substituição, deve-se então avaliar qual resíduo de aminoácido está presente nessa posição para as demais proteínas WNT reconhecidas no filo Porifera.

A estrutura tridimensional de uma proteína se dá pela interação entre seus átomos, levando a um complexo processo de dobradura, que resulta na proximidade entre sítios de maneira não linear (Fares, McNally, 2006). A evolução de uma proteína está associada

também a processos de coevolução entre seus sítios, levando a conformações estruturais específicas que resultam em restrições funcionais (Fares, Travers, 2006). Nós identificamos sítios coevoluindo sob dois processos distintos: coevolução por conservação conjunta e coevolução por viés de substituição conjunta. Identificamos seis grupos de sítios coevoluindo por conservação (G1-G6) com sítios distribuídos ao longo de toda a proteína. A coevolução por conservação desses sítios, indica a sua importância para a manutenção forma e função da proteína. Sítios ligados funcionalmente e estruturalmente tendem a ser mantidos sob restrições seletivas mais fortes, ou seja, os aminoácidos contidos nesses sítios devem ser mantidos a fim de manter a função e estrutura da proteína e alterações nessas posições podem ter resultados drásticos na conformação tridimensional (Fares, McNally, 2006; Codoner, Fares, 2008). Destacamos, especialmente, a coevolução dos grupos G3 e G6 com resíduos de aminoácidos relacionados à ligação WNT-N-Glicano e do grupo G4 com resíduos de aminoácidos relacionados à forma da proteína WNT. Os grupos G3 e G6 contém, respectivamente, um resíduo do aminoácido Ácido aspártico e dois resíduos do aminoácido Asparagina e um resíduo do aminoácido Ácido Aspártico e um resíduo do aminoácido Asparagina. Resíduos de Asparagina estão relacionados à adição de glicanos que compõem a estrutura quaternária da proteína WNT em *X. laevis* (Janda, *et al.*, 2012). A posição do sítio na estrutura primária para esses resíduos não se sobrepõem perfeitamente aos resíduos Asn104 e Asn263 de *X. laevis*, porém esses resíduos de Asparagina estão em sítios próximos ao que seria esperado para a ligação do N-Glicano na conformação tridimensional, como em LCOwntA com Asn166 e Asn285 e em LCHwntB Asn193 e Asn315. O grupo G4 contém uma dupla de resíduos de Cisteína localizadas no domínio CTD, os quais formam as pontes dissulfeto que mantém a estrutura característica da proteína WNT.

Identificamos 111 pares de sítios coevoluindo por viés de substituição conjunta. Os sítios que estão no domínio NTD coevoluem intra e inter-regiões, porém preferencialmente entre si, os sítios da região ligadora coevoluem exclusivamente intra-região e os sítios no domínio CTD coevoluem intra e inter-regiões, mas preferencialmente entre si também. Esse padrão de evolução intra-região pode estar associado à história evolutiva da proteína WNT, onde a composição em dois domínios da proteína se dá pela possível origem como o resultado da junção de uma saposina e uma citoquina (Bazan, *et al.*, 2012). Considerando o argumento anterior de que mudanças em sítios chave para forma/função podem ser drásticos para a proteína, podemos inferir que

os sítios que estão coevoluindo sob viés de substituição conjunta e que não estão diretamente ligados à forma/função das proteínas WNT são mantidos sob restrições seletivas menos fortes, resultando o que leva à possibilidade de ocorrência desse tipo de coevolução (Codoner, Fares, 2008). Um exemplo para esse argumento, é a região ligadora, que apresentou grande presença de coevolução entre seus sítios. A região ligadora possui maior flexibilidade molecular, em relação à sua forma tridimensional, por não estar envolvida diretamente na interação WNT-Fz ou na manutenção da forma da proteína (Janda, *et al.*, 2012; Hirai, *et al.*, 2018).

Por fim, nossos dados indicam que a expansão do número de cópias WNT ocorre em toda a classe Calcarea, antes da divisão entre as duas subclasses. Há maior diversidade de estruturas primárias do que terciárias na classe, um reflexo da evolução das proteínas em que as estruturas primárias sofrem mudanças com maior facilidade do que as estruturas terciárias. Notamos dois agrupamentos de estruturas terciárias, em um há manutenção da forma da proteína WNT e em outro há alterações da estrutura, um possível reflexo da alteração das restrições da evolução sob essas cópias. A proteína WNT está relacionada à formação e evolução do sistema aquífero, que organiza o plano corporal do grupo Porifera. Indutivamente, a expansão de proteínas WNT em Calcarea poderia estar relacionada a sua maior diversidade de planos corporais. Com a expansão de cópias e a alteração das restrições evolutivas sobre parte delas, alguns genes poderiam sofrer grandes mudanças (como as proteínas dos clados C1-C4) e serem cooptados para exercer novas funções, podendo até mesmo participar de outras vias com novos receptores. Esse é o primeiro estudo que avaliou a evolução da proteína WNT através de estruturas primárias e terciárias para o filo Porifera, o primeiro grupo a divergir na linhagem dos metazoários. Um passo no entendimento da complexidade dessas proteínas em uma classe com características únicas no filo, a classe Calcarea. Compreender a evolução dessa proteína a nível de estrutura primária e terciária em animais tão antigos é uma janela para compreendermos o funcionamento da via de sinalização e suas consequências no desenvolvimento do plano corporal dos animais.

REFERÊNCIAS

- Adell, T., Nefkens, I., & Müller, W. E. (2003). Polarity factor 'Frizzled' in the demosponge *Suberites domuncula*: identification, expression and localization of the receptor in the epithelium/pinacoderm. *FEBS Letters*, 554(3), 363-368.
- Adell, T., Thakur, A. N., & Müller, W. E. (2007). Isolation and characterization of Wnt pathway-related genes from Porifera. *Cell Biology International*, 31(9), 939-949.
- Adamska, M., Degnan, S. M., Green, K. M., Adamski, M., Craigie, A., Larroux, C., & Degnan, B. M. (2007). Wnt and TGF- β expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PloS One*, 2(10), e1031.
- Adamska, M., Larroux, C., Adamski, M., Green, K., Lovas, E., Koop, D., ... & Degnan, B. M. (2010). Structure and expression of conserved Wnt pathway components in the demosponge *Amphimedon queenslandica*. *Evolution & Development*, 12(5), 494-518.
- Adamska, M. (2021) Porifera. In: *Handbook of Marine Model Organisms in Experimental Biology*. CRC Press, p. 67-77.
- Aravind, L., Mazumder, R., Vasudevan, S., & Koonin, E. V. (2002). Trends in protein evolution inferred from sequence and structure analysis. *Current Opinion in Structural Biology*, 12(3), 392-399.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410.
- Bhattacharya, R., Rose, P. W., Burley, S. K., & Prlić, A. (2017). Impact of genetic variation on three dimensional structure and function of proteins. *PloS One*, 12(3), e0171355.
- Bazan, J. F., Janda, C. Y., & Garcia, K. C. (2012). Structural architecture and functional evolution of Wnts. *Developmental Cell*, 23(2), 227-232.
- Blomberg, S. P., Garland Jr, T., & Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: behavioral traits are more labile. *Evolution*, 57(4), 717-745.
- Borisenko, I., Adamski, M., Ereskovsky, A., & Adamska, M. (2016). Surprisingly rich repertoire of Wnt genes in the demosponge *Halisarca dujardini*. *BMC Evolutionary Biology*, 16, 1-7.
- Borisenko, I., Podgornaya, O. I., & Ereskovsky, A. V. (2019). From traveler to homebody: Which signaling mechanisms sponge larvae use to become adult sponges?. *Advances in Protein Chemistry and Structural Biology*, 116, 421-449.
- Borisenko, I., Bolshakov, F. V., Ereskovsky, A., & Lavrov, A. I. (2021). Expression of Wnt and TGF-Beta Pathway Components during Whole-Body Regeneration from Cell Aggregates in Demosponge *Halisarca dujardini*. *Genes*, 12(6), 944.
- Bruhn, H. (2005). A short guided tour through functional and structural features of saposin-like proteins. *Biochemical Journal*, 389(2), 249-257.
- Cavalcanti, F., Klautau, M. 2011. Solenoid: a new aquiferous system to Porifera. *Zoomorphology* 130(4):255-260.

- Chu, M. L. H., Ahn, V. E., Choi, H. J., Daniels, D. L., Nusse, R., & Weis, W. I. (2013). Structural studies of Wnts and identification of an LRP6 binding site. *Structure*, 21(7), 1235-1242.
- Codoñer, F. M., & Fares, M. A. (2008). Why should we care about molecular coevolution?. *Evolutionary Bioinformatics*, 4, 117693430800400003.
- Croce, J. C., & McClay, D. R. (2009). Evolution of the Wnt pathways. *Wnt Signaling*, 3-18.
- Degnan, B. M., Leys, S. P., & Larroux, C. (2005). Sponge development and antiquity of animal pattern formation. *Integrative and Comparative Biology*, 45(2), 335-341.
- Dunn, C. W., Leys, S. P., & Haddock, S. H. (2015). The hidden biology of sponges and ctenophores. *Trends in Ecology & Evolution*, 30(5), 282-291.
- Eerkes-Medrano, D. I., & Leys, S. P. (2006). Ultrastructure and embryonic development of a syconoid calcareous sponge. *Invertebrate Biology*, 125(3), 177-194.
- Ereskovsky, A. V. (2010). *The comparative embryology of sponges*. Springer Science & Business Media.
- Fares, M. A., & Travers, S. A. (2006). A novel method for detecting intramolecular coevolution: adding a further dimension to selective constraints analyses. *Genetics*, 173(1), 9-23.
- Fares, M. A., & McNally, D. (2006). CAPS: coevolution analysis using protein sequences. *Bioinformatics*, 22(22), 2821-2822.
- Fonseca-Júnior, N. J., Afonso, M. Q., Oliveira, L. C., & Bleicher, L. (2018). PFstats: a network-based open tool for protein family analysis. *Journal of Computational Biology*, 25(5), 480-486.
- Graur, D., & Li, W. H. (2000). *Molecular evolution*. Second edition. Sinauer Associates, Sunderland, MA.
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307-321.
- Hall, T., Biosciences, I., & Carlsbad, C. J. G. B. B. (2011). BioEdit: an important software for molecular biology. *GERF Bull Biosci*, 2(1), 60-61.
- Harcet, M., Roller, M., Četković, H., Perina, D., Wiens, M., Müller, W. E., & Vlahoviček, K. (2010). Demosponge EST sequencing reveals a complex genetic toolkit of the simplest metazoans. *Molecular Biology and Evolution*, 27(12), 2747-2756.
- Hirai, H., Matoba, K., Mihara, E., Arimori, T., & Takagi, J. (2019). Crystal structure of a mammalian Wnt-frizzled complex. *Nature Structural & Molecular Biology*, 26(5), 372-379.
- Holm, L., Laiho, A., Törönen, P., & Salgado, M. (2023). DALI shines a light on remote homologs: One hundred discoveries. *Protein Science*, 32(1), e4519.
- Holstein, T. W. (2012). The evolution of the Wnt pathway. *Cold Spring Harbor Perspectives in Biology*, 4(7), a007922.

- Janda, C. Y., Waghray, D., Levin, A. M., Thomas, C., & Garcia, K. C. (2012). Structural basis of Wnt recognition by Frizzled. *Science*, 337(6090), 59-64.
- Johnson, M. S., Sutcliffe, M. J., & Blundell, T. L. (1990a). Molecular anatomy: phyletic relationships derived from three-dimensional structures of proteins. *Journal of Molecular Evolution*, 30, 43-59.
- Johnson, M. S., Sali, A., & Blundell, T. L. (1990b). Phylogenetic relationships from three-dimensional protein structures. *Methods in Enzymology*, 183, 670-690.
- Katoh, K., & Toh, H. (2008). Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics*, 9(4), 286-298.
- Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., & Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*, 37(suppl_1), D387-D392.
- Komiya, Y., & Habas, R. (2008). Wnt signal transduction pathways. *Organogenesis*, 4(2), 68-75.
- Lanna, E., & Klautau, M. (2012). Embryogenesis and larval ultrastructure in *Paraleucilla magna* (Cnidaria, Scyphozoa), with remarks on the epilarval trophocyte epithelium ("placental membrane"). *Zoomorphology*, 131, 277-292.
- Lapébie, P., Gazave, E., Ereskovsky, A., Derelle, R., Bézac, C., Renard, E., ... & Borchiellini, C. (2009). WNT/ β -catenin signalling and epithelial patterning in the homoscleromorph sponge *Oscarella*. *PloS One*, 4(6), e5823.
- Leininger, S., Adamski, M., Bergum, B., Guder, C., Liu, J., Laplante, M., ... & Adamska, M. (2014). Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans. *Nature Communications*, 5(1), 3905.
- Li, Y., Shen, X. X., Evans, B., Dunn, C. W., & Rokas, A. (2021). Rooting the animal tree of life. *Molecular Biology and Evolution*, 38(10), 4322-4333.
- Li, X., Ortiz, M. A., & Kotula, L. (2020). The physiological role of Wnt pathway in normal development and cancer. *Experimental Biology and Medicine*, 245(5), 411-426.
- Loh, K. M., van Amerongen, R., & Nusse, R. (2016). Generating cellular diversity and spatial form: Wnt signaling and the evolution of multicellular animals. *Developmental Cell*, 38(6), 643-655.
- Nichols, S. A., Dirks, W., Pearse, J. S., & King, N. (2006). Early evolution of animal cell signaling and adhesion genes. *Proceedings of the National Academy of Sciences*, 103(33), 12451-12456.
- Renard, E., Leys, S. P., Wörheide, G., & Borchiellini, C. (2018). Understanding animal evolution: the added value of sponge transcriptomics and genomics: the disconnect between gene content and body plan evolution. *BioEssays*, 40(9), 1700237.
- Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., & Leys, S. P. (2014). The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Molecular Biology and Evolution*, 31(5), 1102-1120.

- Rim, E. Y., Clevers, H., & Nusse, R. (2022). The Wnt pathway: from signaling mechanisms to synthetic modulators. *Annual Review of Biochemistry*, 91, 571-598.
- Rombel, I. T., Sykes, K. F., Rayner, S., & Johnston, S. A. (2002). ORF-FINDER: a vector for high-throughput gene identification. *Gene*, 282(1-2), 33-41.
- Ryan, J. F., & Baxevanis, A. D. (2007). Hox, Wnt, and the evolution of the primary body axis: insights from the early-divergent phyla. *Biology Direct*, 2, 1-11.
- Santini, S., Schenkelaars, Q., Jourda, C., Duchesne, M., Belahbib, H., Rocher, C., ... & Renard, E. (2023). The compact genome of the sponge *Oopsacas minuta* (Hexactinellida) is lacking key metazoan core genes. *BMC Biology*, 21(1), 1-21.
- Sehnal, D., Bittrich, S., Deshpande, M., Svobodová, R., Berka, K., Bazgier, V., ... & Rose, A. S. (2021). Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research*, 49(W1), W431-W437
- Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E., Mitros, T., ... & Rokhsar, D. S. (2010). The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature*, 466(7307), 720-726.
- Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., ... & Bateman, A. (2023). InterPro in 2022. *Nucleic Acids Research*, 51(D1), D418-D427.
- Pei, J., Tang, M., & Grishin, N. V. (2008). PROMALS3D web server for accurate multiple protein sequence and structure alignments. *Nucleic Acids Research*, 36(suppl_2), W30-W34.
- Ponting, C. P., & Russell, R. R. (2002). The natural history of protein domains. *Annual Review of Biophysics and Biomolecular Structure*, 31(1), 45-71.
- Pond, S. L. K., Frost, S. D., & Muse, S. V. (2005). HyPhy: hypothesis testing using phylogenies. *Bioinformatics*, 21(5), 676-679.
- Thorne, J. L., Goldman, N., & Jones, D. T. (1996). Combining protein evolution and secondary structure. *Molecular Biology and Evolution*, 13(5), 666-673.
- Verli, H. (2014). O que é Bioinformática?. *Bioinformática: da biologia à flexibilidade molecular*. Cap. 1, p. 2-12.
- Verkaar, F., Cadigan, K. M., & van Amerongen, R. (2012). Celebrating 30 years of Wnt signaling.
- Voogd, N.J.; Alvarez de, B.; Boury-Esnault, N.; Carballo, J.L.; Cárdenas, P.; Díaz, M.-C.; Dohrmann, M.; Downey, R.; Goodwin, C.; Hajdu, E.; Hooper, J.N.A.; Kelly, M.; Klautau, M.; Lim, S.C.; Manconi, R.; Morrow, C.; Pinheiro, U.; Pisera, A.B.; Ríos, P.; Rützler, K.; Schönberg, C.; Vacelet, J.; van Soest, R.W.M.; Xavier, J. (2023). World Porifera Database. Accessed at <https://www.marinespecies.org/porifera> on 2023-08-26. doi:10.14284/359
- Voigt O, Wülfing E, Wörheide G. 2012. Molecular Phylogenetic Evaluation of Classification and Scenarios of Character Evolution in Calcareous Sponges (Porifera, Class Calcarea). *PLoS One* 7(3):e33417.

Yuan, S., Chan, H. S., & Hu, Z. (2017). Using PyMOL as a platform for computational drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 7(2), e1298.

Wheeler, T. J., Clements, J., & Finn, R. D. (2014). Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. *BMC Bioinformatics*, 15, 1-9.

Willert, K., & Nusse, R. (2012). Wnt proteins. *Cold Spring Harbor Perspectives in Biology*, 4(9), a007864.

Windsor, P. J., & Leys, S. P. (2010). Wnt signaling and induction in the sponge aquiferous system: evidence for an ancient origin of the organizer. *Evolution & Development*, 12(5), 484-493..

Windsor-Reid, P. J., Matveev, E., McClymont, A., Posfai, D., Hill, A. L., & Leys, S. P. (2018). Wnt signaling and polarity in freshwater sponges. *BMC Evolutionary Biology*, 18, 1-14.

Zhang, J. M., & An, J. (2007). Cytokines, inflammation and pain. *International Anesthesiology Clinics*, 45(2), 27.

APÊNDICES

O material suplementar listados abaixo podem ser acessados no link:

<<https://drive.google.com/drive/folders/1mnRrLuH7mJzIb0SxmTA9FBBecHrtFOCf?usp=sharing>>.

Apêndice 1 - Alinhamento Múltiplo de Sequências - Estrutura primária

Apêndice 2 - Alinhamento Múltiplo de Sequências - Estrutura terciária

Apêndice 3 - Perfil Logos

Apêndice 4 - Modelos tridimensionais

Apêndice 5 - Tabela de conservação de aminoácidos

Apêndice 6 - Tabela de coevolução de aminoácidos