

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCÊNCIAS
DEPARTAMENTO DE BIOLOGIA MOLECULAR E BIOTECNOLOGIA
CURSO DE GRADUAÇÃO EM BIOTECNOLOGIA

Daniel da Silva Severiano dos Santos

Análise de filogenia e de rede de interações em proteínas Argonautas de bactérias e *arquéias* mesofílicas e termofílicas de interesse biotecnológico.

PortoAlegre
2024

Daniel da Silva Severiano dos Santos

Análise de filogenia e de rede de interações em proteínas Argonautas de bactérias e *arqueias* mesofílicas e termofílicas de interesse biotecnológico.

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do título de bacharel em Biotecnologia da Universidade Federal do Rio Grande do Sul.

Área de habilitação: Biotecnologia Molecular

Orientação: Prof. Dr. Diego Bonatto

PortoAlegre

2024

CIP - Catalogação na Publicação

Santos, Daniel
Análise de filogenia e de rede de interações em
proteínas Argonautas de bactérias e arqueias
mesofílicas e termofílicas de interesse
biotecnológico. / Daniel Santos. -- 2024.
45 f.
Orientador: Diego Bonatto.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Instituto
de Biociências, Curso de Biotecnologia: Biotecnologia
Molecular, Porto Alegre, BR-RS, 2024.

1. Argonautas. 2. Arqueia. 3. Abi. 4. Thermococcus
barophilus. I. Bonatto, Diego, orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os
dados fornecidos pelo(a) autor(a).

Daniel da Silva Severiano dos Santos

**Análise de filogenia e de rede de interações em proteínas Argonautas de bactérias e
arqueias mesofílicas e termofílicas de interesse biotecnológico.**

Trabalho de Conclusão de Curso apresentado como requisito parcial para obtenção do
título de bacharel em Biotecnologia pela Universidade Federal do Rio Grande do Sul.

Aprovada em 14 / 08 /2024

BANCA EXAMINADORA

Prof. Dr. Diego Bonatto

Prof. Charley Chistian Staats

Prof. Dr. Nelson Fagundes

DEDICATÓRIA

Dedico esse trabalho aos muitos que durante minha jornada na UFRGS se conectaram a mim pela paixão do desconhecido, a aqueles que entendem que é um privilégio o conhecimento e aos sagazes que estão sempre em busca do saber.

AGRADECIMENTOS

Como uma vez disse o digníssimo Carlos Drummond de Andrade:

**“No meio do caminho tinha uma pedra
Tinha uma pedra no meio do caminho
Tinha uma pedra
No meio do caminho tinha uma pedra
Nunca me esquecerei desse acontecimento
Na vida de minhas retinas tão fatigadas
Nunca me esquecerei que no meio do caminho
Tinha uma pedra
Tinha uma pedra no meio do caminho
No meio do caminho tinha uma pedra”**

Foram-se as pedras e os caminhos seguem caminhos. Caminhos esses que por ventura espero que ainda existam pedras. E afortunado seria eu que se nesse caminho de pedras possa contar mais uma vez com minha mãe, Jocerlei da Silva dos Santos, que não teve a oportunidade do estudo formal mas que fez da vida uma fonte inesgotável de sabedoria, meu pai, Roberto Severiano dos Santos, que encontrou salvação nos livros, que nunca limitou-se pelas dificuldades de sua infância, que contou com a sorte das palavras gentis e severas da minha avó, Margarida Severiano dos Santos, que criou 12 filhos sozinha, nas dificuldades da miséria social do Ceará, que encontrou tempo e a sabedoria de educar seus filhos aos estudos.

Aos amigos que ficaram e aos que foram, carrego comigo todas palavras de apoio e todos os momentos de escuta. Guardo comigo cada um que tornou possível e viável, intelectualmente e emocionalmente, esse ciclo, que em breve se encerra.

A futura mãe dos meus filhos, Gabrielle Fialho dos Santos, que esteve comigo durante todos meus momentos de incertezas, foi um alicerce de amor e companheirismo durante grande parte desse processo. Ao meu sogro, Batista Fialho, e minha sogra, Simone Fialho, que me acolheram como um filho em diversos momentos.

A todos os professores que tornaram essa experiência de aprendizado uma jornada excitante, emocionante. As professoras que me orientaram durante anos, Professora Fernanda Vianna e Professora Cristianne Matté, dedico os momentos de aprendizado nos laboratórios de Imunogenética e Bioquímica na UFRGS, os momentos mais felizes da minha graduação. Agradeço a banca desse trabalho, Professores Dr. Charley Staats e Dr. Nelson Fagundes, pela correção e olhar atento ao que foi escrito e apresentado, que trouxeram diversos insights, alguns colocados em prática e que certamente contribuíram a uma entrega mais rica. Finalmente, meus agradecimentos ao Professor e Orientador Dr. Diego Bonatto, que durante a pandemia em 2020-2021, me mostrou as maravilhas da bioinformática, se dedicou ao máximo à entrega do ensino com comprometimento, com o detalhe em todas as linhas de códigos e em aulas síncronas e assíncronas incríveis, que abriu meu olhar a ciência de dados e a quem dedico esse trabalho que torna possível meu futuro título de Bacharel.

"Great men are forged in fire. It is the privilege of lesser men to light the flame."

RESUMO

As Argonautas longas, presentes em procariontes (pAgos longos A e B) e eucariotos (eAgos), têm quatro domínios principais: N, PAZ, MID e PIWI, além de dois domínios estruturais de ligação, L1 e L2. Os pAgos são mais diversos que os eAgos, com aproximadamente 30% dos genomas bacterianos e 10% dos genomas sequenciados de arqueias contendo pAgos, dos quais 60% são pAgos curtos. Estes últimos possuem apenas os domínios MID e PIWI, que são inativos.

O domínio PAZ é conectado aos domínios N e MID por dois loops, L1 e L2, permitindo flexibilidade. Nos pAgos curtos, frequentemente há uma fusão com proteínas efetoras contendo o domínio APAZ, que corresponde aos domínios N, L1 e L2 das pAgos longas. Esses sistemas são semelhantes aos de defesa CRISPR-Cas, pois ambos reconhecem ácidos nucleicos invasores, mas os pAgos curtos frequentemente induzem a morte celular para prevenir a propagação do invasor, funcionando como sistemas de infecção abortiva.

Estudos estruturais e bioquímicos mostram que os pAgos de procariontes termofílicos têm atividade de endonuclease guiada por ácidos nucleicos, semelhante à atividade CRISPR-Cas9. Os pAgos longos-A, que possuem uma tétrade catalítica DEDX no domínio PIWI, podem clivar o DNA alvo mediado por guias de DNA. Já os pAgos longos-B carecem dessa tétrade catalítica e o domínio PAZ frequentemente é truncado. Os pAgos curtos são divididos em quatro clados e frequentemente são encontrados em operons junto a um domínio APAZ.

Os pAgos curtos, assim como os pseudo-pAgos curtos, colaboram com proteínas associadas para proteger o hospedeiro contra DNA invasor, como plasmídeos e vírus. Em vez de degradar o DNA invasor diretamente, esses sistemas induzem a morte celular após o reconhecimento mediado por guias, funcionando como mecanismos de defesa abortiva.

Arqueias termofílicas, que são “fósseis vivos” e fornecem insights sobre a evolução primordial, apresentam características bioquímicas e fisiológicas que oferecem oportunidades na biotecnologia. Estas arqueias têm genomas pequenos e são promissoras em aplicações de engenharia metabólica, especialmente em condições onde bactérias e eucariotos falham, como em ambientes de alta temperatura.

Os pAgos termofílicos com atividade de endonuclease têm grande potencial para manipulação genética programável. Portanto, explorar proteínas Argonautas com

propriedades enzimáticas únicas pode levar a avanços na biotecnologia, incluindo aplicações em biologia molecular e medicina. As análises filogenéticas e funcionais indicam que alguns pAgos termofílicos podem funcionar em sistemas de resposta abortiva à infecção, destacando a importância de continuar a caracterizar esses mecanismos e suas aplicações.

Palavras-chave: Proteínas Argonautas; Arquéias Termofílicas; Resposta abortiva a infecção (Abi); TbAgo; *Thermococcus barophilus*.

ABSTRACT

Argonaute (Ago) proteins are present in all domains of life: bacteria, archaea, and eukaryotes. They use small oligonucleotide guides (15-30 nucleotides) to bind to complementary nucleic acid targets, playing crucial roles in gene expression regulation, silencing mobile genome elements, and defense against viruses or plasmids. They are classified into long and short Argonautes based on domain organization and phylogeny.

Long Argonautes, present in prokaryotes (long pAgos A and B) and eukaryotes (eAgos), have four main domains: N, PAZ, MID, and PIWI, along with two structural binding domains, L1 and L2. Prokaryotic Agos (pAgos) are more diverse than eukaryotic Agos (eAgos), with approximately 30% of bacterial genomes and 10% of sequenced archaeal genomes containing pAgos, of which 60% are short pAgos. These short pAgos only possess the inactive MID and PIWI domains.

The PAZ domain is connected to the N and MID domains by two loops, L1 and L2, allowing flexibility. In short pAgos, there is often a fusion with effector proteins containing the APAZ domain, corresponding to the N, L1, and L2 domains of long pAgos. These systems resemble CRISPR-Cas defense systems, as both recognize invading nucleic acids, but short pAgos often induce cell death to prevent the spread of the invader, functioning as abortive infection systems.

Structural and biochemical studies show that pAgos from thermophilic prokaryotes have nucleic acid-guided endonuclease activity similar to CRISPR-Cas9. Long pAgos-A, which possess a catalytic tetrad DEDX in the PIWI domain, can cleave target DNA mediated by DNA guides. Long pAgos-B lack this catalytic tetrad, and the PAZ domain is often truncated. Short pAgos are divided into four clades and are often found in operons alongside an APAZ domain.

Short pAgos, as well as pseudo-short pAgos, collaborate with associated proteins to protect the host against invading DNA, such as plasmids and viruses. Instead of directly degrading the invader DNA, these systems induce cell death after guide-mediated recognition, functioning as abortive defense mechanisms.

Thermophilic archaea, which are "living fossils" and provide insights into primordial evolution, have biochemical and physiological characteristics that offer opportunities in biotechnology. These archaea have small genomes and are promising in metabolic engineering applications, especially in conditions where bacteria and eukaryotes fail, such as high-temperature environments.

Thermophilic pAgos with endonuclease activity have great potential for programmable genetic manipulation. Therefore, exploring Argonaute proteins with unique enzymatic properties can lead to advances in biotechnology, including applications in molecular biology and medicine. Phylogenetic and functional analyses indicate that some thermophilic pAgos may function in abortive infection response systems, highlighting the importance of continuing to characterize these mechanisms and their applications.

Keywords: Argonaute Proteins; Thermophilic Archaea; Abortive Infection (Abi) response; TbAgo; Thermococcus barophilus.

LISTA DE FIGURAS

Figura 1. Parâmetros usados para o alinhamento global das sequências de pAgo utilizando o programa ClustalW.

Figura 2. Árvore filogenética consenso de pAgos gerada por máxima verossimilhança utilizando o software IQ-TREE 2.

Figura 3. Uma composição da árvore filogenética, gerada por máxima verossimilhança à esquerda, e os motivos conservados que correspondem aos Argonautas à direita, deduzidos fazendo uso do programa MEME Suit. A esquerda, a diferenciação dos organismos termofílicos dos mesófilos, foi representado pelos galhos da árvore, vermelhos para os organismos termofílicos e preto os mesofílicos. Proteínas WP anotadas em coloração verde se referem aos organismos bacterianos presentes na árvore filogenética. Entre parênteses ([]) estão os organismos com o domínio paz conservado. A esquerda observa-se diversas caixas coloridas, em que cada caixa faz referências a um motivo conservado diferente. Colorações iguais fazem referência ao mesmo motivo conservado.

Figura 4. Mudanças conformacionais no sítio ativo do TtAgo durante o reconhecimento do alvo e catálise . Os resíduos do sítio ativo são mostrados em vermelho; o dedo de glutamato é indicado por um círculo vermelho. Durante as primeiras etapas de ligação do guia e reconhecimento do alvo, o sítio ativo é desconectado; a propagação duplex é acompanhada por mudanças nas conformações dos domínios PIWI e PAZ (indicados com setas vermelhas), plug-in do dedo de glutamato, ligação catalítica de metal e ativação da catálise. Finalmente, a liberação gradual do alvo leva à desconexão do sítio ativo, possibilitando assim o reconhecimento da próxima molécula alvo. Os números de acesso do PDB (do canto superior esquerdo, no sentido horário): 3DLH, 3F73, 4N41, 4NCB, 4NCA, 4N76

Figura 5. Representação gerada via Web Flags dos operons vizinhos as Argonautas (colorido) e o domínio PIWI das Argonautas (preto). Nota-se que diversos espaços em brancos estão representados na figura, estas regiões não foram identificados pelo software como tendo uma função definida. Esses genes foram analisados posteriormente e anotados via HMMER (tabela complementar 2).

Figura 6. Modelo para resposta imune antiviral do sistema SiAgo.

Figura 7. Proteína vizinha do TbAgo (WP_048159988.1) analisada pelo servidor do TMHMM para proteína transmembrana.

Figura 8. SIR2 (SPARSA), TIR (SPARTA) ou Aga2 (sistema SiAgo). Em SPARSA e SPARTA, a ativação de SIR2/TIR leva à conversão de NAD(P)+ em NAM e ADPR(P). No SPARTA, a ativação catalítica requer a tetramerização de quatro complexos SPARTA heterodiméricos ligados ao guia/alvo. Após a detecção do invasor, o complexo SiAgo-SiAga1 é recrutado para a proteína de membrana SiAga2, que induz a despolarização da membrana.

LISTA DE FIGURAS

Tabela 1. Enzimas provenientes de arqueias termofílicas com aplicações em biologia molecular.

Tabela 2. 6EMY_B, 6Y8Q_B, PF08843.16 correspondem as proteínas vizinhas do TbAgo WP_013468137.1, WP_048159988.1, WP_013468141.1 respectivamente.

SUMÁRIO

| | |
|--|----|
| 1. Introdução | |
| 1.1. Argonautas e seus domínios funcionais..... | 16 |
| 1.2. Clados dos Argonautas e seus domínios conservados..... | 17 |
| 1.3. Arquéias termofílicas e suas aplicabilidades biotecnológicas..... | 19 |
| 1.4. Potencial dos Argonautas termotolerantes na biologia sintética..... | 20 |
| 2. Objetivos | |
| 2.1. Objetivo Geral..... | 23 |
| 2.2. Objetivos Específicos..... | 23 |
| 3. Materiais e Métodos | |
| 3.1. Alinhamento de Sequências e construção da Árvore Filogenética..... | 24 |
| 3.2. Análise do contexto genético..... | 25 |
| 3.3. Análise de sequência e estrutura de proteínas..... | 26 |
| 4. Resultado e discussão | |
| 4.1. Anotação das sequências homólogas e construção da árvore filogenética | |
| 4.2. Exploração dos dados e anotação das sequências. | |
| 4.2.1 MEME.....;; | 27 |
| 4.2.2 Web Flag..... | 30 |
| 4.2.3 HMMER e HHPRED.....;; | 31 |
| 4.3. Análises estruturas | |
| 4.3.1 EMBOSS Needle | 32 |
| 4.3.2 TMHMM | 34 |
| 5. Conclusão e Perspectivas..... | 35 |
| 6. Referências..... | 37 |
| 7. Anexos..... | 45 |

1. INTRODUÇÃO

1.1. Argonautas e seus domínios funcionais

As proteínas Argonauta (Ago) são encontradas nos três domínios da vida (bactérias, arqueas e eucariotos). Elas utilizam pequenos guias de oligonucleotídeos (15-30 nucleotídeos) para se ligarem a alvos de ácidos nucleicos complementares, desempenhando papéis cruciais na regulação da expressão gênica, no silenciamento de elementos móveis do genoma e na defesa contra vírus ou plasmídeos. De acordo com a organização dos seus domínios e filogenia, as proteínas Ago são classificadas em Argonautas longas e curtas.(Manakova et al., 2024)

As Argonautas longas, presentes em procariontes (pAgos longos A e B) e eucariotos (eAgos), são compostas por quatro domínios funcionais principais (N, PAZ, MID e PIWI) e dois domínios estruturais de ligação (L1 e L2). A análise do genoma de procariontes revela que os pAgos são mais diversos que os eAgos, uma vez que estão distribuídos em aproximadamente 30% dos genomas bacterianos e 10% dos genomas sequenciados de arqueias, deles, aproximadamente 60% dos pAgos são pAgos curtos, que contêm apenas os domínios MID e PIWI inativos (Manakova et al., 2024; Koopal et al., 2023; Zaremba et al., 2022; Wang et al., 2023; Ni et al., 2023)

O domínio PAZ é conectado ao domínio N e MID por dois loops, L1 e L2, respectivamente. Isto torna o domínio PAZ flexível o suficiente para se mover como um corpo rígido em relação a todos os outros domínios. Nos seus operons ou polipeptídeos, os pAgos curtos são geralmente associados a proteínas efetoras contendo o domínio APAZ (análogo de PAZ). De acordo com modelagem estrutural e dados experimentais, os domínios APAZ correspondem aos domínios N, L1 e L2 das pAgos longas. As proteínas contendo APAZ frequentemente se fundem com domínios efetores Sir2 (regulador silencioso de informação 2), nucleases Mrr ou TIR (receptor Toll-Interleukin-1).(Koopal et al., 2023)

Os APAZ/Agos heterodiméricos ou os APAZ-Agos de cadeia única assemelham-se a pAgos longos sem o domínio PAZ, mas contendo um domínio efetor adicional. Após a detecção de DNA invasor mediada pelo guia dos pAgos curtos, suas enzimas efetoras associadas eliminam a célula hospedeira, prevenindo assim a propagação do invasor. Em vez de degradar diretamente os seus alvos, os pAgos curtos formam complexos com as suas enzimas efetoras associadas para desencadear a morte celular após o reconhecimento mediado por guia do DNA invasor e, assim, funcionam como sistemas de infecção abortivos.

Atualmente, os sistemas de defesa Ago e CRISPR-Cas são os únicos que se baseiam no reconhecimento de ácidos nucleicos guiados por ácidos nucleicos. Embora a consequência desse reconhecimento seja geralmente a degradação do ácido nucleico (Abudayyeh et al., 2016; Gasiunas et al., 2012; Jinek et al., 2012; Kazlauskienė et al., 2017; Niewoehner et al., 2017; Swarts et al., 2014a; Westra et al., 2012), diversas proteínas de membrana e outros efetores tóxicos potenciais são frequentemente encontrados em associação com esses sistemas (Ryazansky et al., 2018; Shah et al., 2019; Shmakov et al., 2018; van Beljouw et al., 2021). Os pAgos com atividade de endonuclease apresentam um grande potencial para a manipulação genética programável. Portanto, é altamente desejável explorar proteínas Argonautas com propriedades enzimáticas únicas, a fim de compreender seus diversos mecanismos catalíticos e promover suas aplicações na biotecnologia.

1.2. Clados dos Argonautas e seus domínios conservados

Os estudos estruturais e bioquímicos de pAgos de procariontes termofílicos mostraram que essas proteínas possuem atividade de endonuclease dirigida por guias de ácidos nucleicos, que é semelhante à atividade CRISPR-Cas9. Entre os pAgos longos-A, 79% possuem uma tétrade catalítica DEDX intacta (Box 1) no domínio PIWI, e a maioria dos pAgos longos-A caracterizados é capaz de clivar o DNA alvo mediado por guias de DNA. No entanto, pelo menos *in vitro*, um pAgo longo-A utiliza guias de RNA para direcionar o DNA, outros utilizam guias de DNA para direcionar o RNA, e alguns não têm uma preferência clara por um tipo específico de guia e alvo; entretanto, todos os pAgos longos-B carecem da tétrade catalítica DEDX e o domínio PAZ frequentemente é parcialmente truncado (PAZ*).

Os pAgos curtos podem ser divididos em quatro clados filogenéticos e são codificados em operons que também codificam um domínio APAZ. Inicialmente, previu-se que o APAZ substituísse funcionalmente o domínio PAZ nos pAgos longos, mas estudos posteriores sugeriram que são homólogos ao domínio N dos pAgos longos. Modelos gerados pelo AlphaFold2 de sistemas de pAgos curtos corroboram que o domínio N e APAZ são homólogos e assumem a mesma posição em relação ao MID-PIWI. O N-terminal do APAZ é geralmente fundido a um (putativo) domínio catalítico. Nos clados S1A e S1B, o APAZ é fundido a um domínio regulador de informação silenciosa 2 (SIR2, também conhecido como Sirtuin). No clado S2A, o APAZ é fundido a um domínio receptor Toll-interleucina (TIR). Nos diferentes sub-clados S2B, o APAZ é fundido a um dos vários domínios (por exemplo, Schlafen/Alba, Mrr-like, DUF4365 ou RecG/DHS-like). No clado

S1A, a proteína contendo o domínio APAZ é fundida ao N-terminal do pAgo curto. Em outras clados, as proteínas contendo o domínio APAZ são codificadas por um gene separado à montante do pAgo curto, o que sugere interdependência funcional. Curiosamente, encontram-se pAgos truncados espalhados por diferentes ramos dos clados longas-A e longas-B. Isso implica que a perda dos domínios N e PAZ ocorreu várias vezes durante a evolução dos pAgos. Conseqüentemente, nem todos os pAgos curtos são relacionados filogeneticamente: o pAgo curto de *Sulfolobus islandicus* (SiAgo) e seus homólogos formam um clado de pAgos que não se agrupa com pAgos longos ou curtos (a clade SiAgo-like), e o pAgo curto AfAgo se agrupa com pAgos do clado longa-B. SiAgo e AfAgo não estão associados a domínios APAZ e dependem de mecanismos funcionais distintos. Portanto, os pAgos curtos filogeneticamente agrupados associados a domínios APAZ são denominados 'pAgos curtos' e os outros pAgos truncados como 'pseudo-pAgos curtos'.(Ofir, Gal et al., 2022; Koopal B et al., 2022)

Similarmente aos pAgos curtos, os pseudo-pAgos curtos se co localizam com outros genes e/ou sequências codificantes. Por exemplo, o Argonauta SiAgo é codificado em um operon com um regulador transcricional previsto, enquanto outras duas proteínas são codificadas na fita oposta do DNA. A presença desses três genes em proximidade ao SiAgo é conservada para homólogos do SiAgo em diferentes cepas. O AfAgo foi caracterizado até agora apenas como uma proteína independente, apesar de ser codificado em um operon contendo sequências estruturais codificantes para duas proteínas hipotéticas. Combinados, esses dados sugerem que tanto os pAgos curtos quanto os pseudo-pAgos curtos funcionam em conjunto com proteínas codificadas em seu contexto genômico.(Zeng et al., 2022; Manakova et.al., 2006; Manakova et.al., 2009)

Também foi observado no Argonautas do organismo *Archaeoglobus fulgidus* DSM4304, a partir de uma análise de genes vizinhos, uma associação de pAgos longos-B com um gene codificador de proteína localizado imediatamente à jusante, correspondendo a uma proteína semelhante ao efetor de membrana SiAga2 do sistema antiviral SiAgo *Sulfolobus islandicus*. Esta observação sugere que os pAgos longos-B cataliticamente inativos, aqui apresentados como pseudo-pAgos curtos, podem funcionar regulando as atividades desses efetores que, quando ativados, podem atuar como toxinas.

Notavelmente, os contextos genéticos dos pAgos também codificam outras proteínas de membrana. As análises de bioinformática mostram que os domínios SIR2 e TIR, recentemente encontrados envolvidos em Abi, também são encontrados em

associação com pAgos, especialmente pAgos curtos e pseudo-pAgos curtos.(Manakova et.al., 2006)

De fato, estudos recentes confirmam que os sistemas de pAgos curtos e pseudo-pAgos formam complexos com suas proteínas associadas para proteger seu hospedeiro contra DNA invasor, como plasmídeos e vírus. Em vez de desencadear a degradação do DNA invasor, semelhante aos pAgos longos, esses sistemas de pAgos curtos e pseudo-pAgos funcionam como sistemas de infecção abortiva, que matam seu hospedeiro para impedir a replicação e a propagação do invasor para outras células.(Zeng et al., 2022)

1.3. Arqueias termófilas e suas aplicabilidades biotecnológicas

Embora o debate ainda rodeie os detalhes da biologia primordial, as arqueias termofílicas são “fósseis vivos” e fornecem um vislumbre deste período crítico da evolução (Whitfield 2004). Muitas vezes ofuscadas pelas suas primas procarióticas, as bactérias, tanto em termos de percepção pública como de estudo científico, às suas características bioquímicas e fisiológicas oferecem oportunidades interessantes para a biotecnologia.

As arqueias termofílicas normalmente têm genomas pequenos (~2 mega pares de bases ou Mpb), simplificando potencialmente a análise de biologia de sistemas e subsequentes esforços de engenharia metabólica(Straub, 2018).

Nesse sentido, nas aplicações onde as plataformas de engenharia metabólica mesófilas mais utilizadas, como *Escherichia coli* e *Saccharomyces cerevisiae* falharam ou tiveram um desempenho insatisfatório, o uso de arqueias termofílicas oferecem uma opção atraente para superar certos problemas de bioprocessamento, como contaminação e infecção por fagos.

Até o momento, as arqueias têm ficado atrás das bactérias e eucariotas no palco da biotecnologia industrial, mas têm sido importantes de várias maneiras, principalmente pelo seu papel na biomineração de cobre (Wheaton et al., 2015) e pelas suas DNA polimerases termoestáveis de alta fidelidade para PCR (Pavlov et al., 2004).

Tabela 1. Enzimas provenientes de arqueias termofílicas com aplicações em biologia molecular. (adaptado de Straub et al., 2018)

| Enzima | Origem | Aplicação | Referências |
|-----------------|-----------------------------|-----------|--------------------|
| DNA polimerases | Thermococcus barophilus Ch5 | PCR | Kwon et al. (2016) |

| | | | |
|---------|-------------------------------------|--|---|
| | Pyrococcus furiosus | PCR, PCR sujeito a erros | Biles and Connolly (2004); Wang et al. (2004) |
| | Sulfolobus solfataricus | PCR sujeito a erros | McDonald et al. (2006) |
| | Thermococcus kodakaraensis | PCR, um passo RT-PCR | Elshawadfy et al. (2014); Okano et al. (2017) |
| | Thermococcus JDF-3 | Sequência de DNA | Arezi, Hansen and Hogrefe (2002) |
| | Thermococcus 9°N-7 | Sequenciamento de DNA, sequência por síntese | Gardner and Jack (2002); Ju et al. (2006) |
| Ligasas | Thermococcus sp. 1519 | Assembleia Gibson | Smagin et al. (2008) |
| | Staphylothermus marinus | Reação em cadeia da ligase | Seo et al. (2007) |
| | Pyrococcus furiosus | Reação em cadeia da ligase | Tanabe et al. (2012) |
| | Methanothermobacter thermotrophicus | 5'-adenilação | Sriskanda et al. (2000) |
| | Thermococcus kodakaraensis | Sequenciamento RNA | Zhang and Tripathi (2017) |
| Inteins | Pyrococcus horikoshii | Purificação de proteína | Lennon, Stanger and Belfort (2016) |
| | Pyrococcus furiosus | Biosensor | Iwai, Lingel and Pluckthun (2001) |

O constante avanço em caracterizar os mecanismos moleculares, a fisiologia e o metabolismo de arqueias termofílicas, possível graças aos avanços de ferramentas "ômicas", torna relevante o interesse em revisitar os organismos previamente caracterizados por suas funções e aplicações biotecnológicas, além de conferir novos olhares as arqueas que possam ter negligenciadas pela falta de entendimento dos seus mecanismos moleculares.

1.4. Potencial dos Argonautas termotolerantes na biologia sintética

Os pAgos termofílicos com atividade endonuclease possuem grande potencial para manipulação genética programável. Portanto, é desejável explorar proteínas de Argonautas com propriedades enzimáticas únicas para compreender seus diversos mecanismos catalíticos e promover suas aplicações em biotecnologia (Fang et al., 2022; Straub et al., 2018). Os estudos estruturais e bioquímicos anteriores de pAgos de procariontes termofílicos mostraram que eles exibem atividade de endonuclease dirigida

por guias de ácidos nucleicos, que é semelhante à atividade CRISPR-Cas9.

Os Agos têm preferências de guia únicas, o domínio PIWI coordena e pré-ordena os nucleotídeos “sementes” do guia (nucleotídeos 2–7 ou 2–8) em uma conformação helicoidal, com bases expostas ao solvente (Bartel DP et al., 2004).

Argonautas euriarqueais termofílicos, que já foram sugeridos como estando mais intimamente relacionados aos Argonautas eucarióticos (Swart et al., 2014). Foi descrito um grupo particular de arqueia termofílicas moderadas do supérfluo **Asgard** que pela sua proximidade evolutiva com eucariontes, acredita-se exibir funções de silenciamento similares aos eAgos e apresentam domínios inativos N-PAZ similar aos pAgos curtos, além de possuir componentes de transposição no operon vizinho ao Argonauta (Bastiaanssen C et al., 2024)

Até o momento, a maioria dos pAgos realiza clivagem específica por meio de um tipo de guia de todos os quatro tipos naturais de pequenos ácidos nucleicos: 5'-P ssDNA, 5'-OH ssDNA, 5'-OH ssRNA e 5'-P ssRNA. Por exemplo, Agos de *Pyrococcus furiosus* (PfAgo) (Rivas et al. 2005; Song et al. 2004; Swarts et al. 2015a), *Aquifex aeolicus* (AaAgo) (Rashid et al. 2007; Yuan et al. 2005, 2006) , *Methanocaldococcus jannaschii* (MjAgo) (Willkomm et al. 2017; Zander et al. 2014, 2017) e *Thermus thermophilus* (TtAgo) (Sheng et al. 2014; Swarts et al. 2014a, 2015b, 2017; Wang et al. 2008a , b, 2009) clivam alvos de ácidos nucleicos complementares usando guias de ssDNA 5'-P. Por sua vez, Agos de *Marinitoga piezophila* (MpAgo) (Doxzen e Doudna 2017; Kaya et al. 2016) e *Thermotoga profunda* (TpAgo) (Kaya et al. 2016) usam o guia 5'-OH ssRNA para clivar alvos de ácido nucleico, enquanto o *Rhodobacter sphaeroides* Ago (RsAgo) (Miyoshi et al. 2016; Olovnikov et al. 2013) utiliza um guia 5'-P ssRNA para reconhecer alvos de ácidos nucleicos cognatos

As análises computacionais prévias identificaram famílias de genes que estão frequentemente codificadas nas vizinhanças genômica dos pAgos e diversos domínios proteicos diretamente fusionados com os pAgos (Makarova et al., 2009; Ryazansky et al., 2018; Swarts et al., 2014b). Embora se preveja que essas proteínas estejam funcionalmente associadas com os pAgos, pouco se sabe sobre seus papéis fisiológicos e mecanismos de ação. Também foi reportado que o pAgo curto do *Sulfolobus islandicus* e suas proteínas associadas colaboram em promover uma resposta antiviral imune robusta mediando a sistema de resposta abortivo contra infecção.

Por meio de análises filogenéticas e funcionais que a Argonauta *Thermococcus barophilus* MP (TbAgo) possui genes próximos que sugerem um mecanismo de função curto e pseudo-curto pAgo em conjunção as proteínas codificadas no seu contexto

genômico, um sistema de ação abortiva a infecção e também um possível mecanismo de transposição dessa maquinaria. As predições das funções dessas proteínas foram feitas *in silico* utilizando modelos previamente descritos para caracterizar um domínios funcionais vizinhos nos Argonautas AfAgo (AfAgo-N/AfAgo-Af1317)(Manakova et al., 2024) e o complexo SiAgo-Aga1-Aga2 (Zeng et al., 2022) que apontam ação de resposta abortiva a infecção (Abi).

2. OBJETIVOS

2.1 Objetivo Geral

Analisar a diversidade evolutiva e seus domínios conservados e prospectar pAgos bacterianos e *arqueias*, mesofílicos e termofílicos, com aplicações biotecnológicas.

2.2 Objetivos Específicos

- Prospectar sequências de Argonautas procariontes(pAGO) usando bancos de dados de genomas de procariotos em bactérias e *arqueias*.
- Estabelecer a relação filogenética das sequências alinhadas dos pAgos.
- Relacionar os domínios conservados e os possíveis papéis funcionais dessas proteínas.

3. MATERIAIS E MÉTODOS

3.1. Alinhamento de Sequências e análise filogenética

Uma busca de similaridade para as sequências primárias de MfAgo, AfAgo, PfAgo, MjAgo, FpAgo e NgAgo foi realizada utilizando a ferramenta BLASTp contra sequências de proteínas no banco de dados NCBI (RefSeq) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). Foram selecionadas sequências da família Ago para alinhamento global utilizando o programa ClustalW (Thompson et al. 1994) de identidade $\geq 30\%$ das sequências primárias. Os dados do alinhamento foram utilizados como entrada para inferir a árvore de máxima verossimilhança, implementado pelo software IQ-TREE 2 (Minh et al. 2020), que selecionou o melhor modelo de análise, através do ModelFinder (S. Kalyaanamoorthy et. al., 2017). Os parâmetros usados para o alinhamento global está indicada nas Figuras 1.

Utilizou-se o MEME, software responsável por descobrir motivos de sequências, no caso do estudo, os argonautas (<https://meme-suite.org/meme/tools/meme>), capaz de apontar motivos novos e sem lacunas (padrões recorrentes e de comprimento fixo). As sequências baixadas do NCBI foram submetidas sem restrição de motivos máximos a serem achados. Se considerou significativo motivos com um E-value ≥ 0.005 .

| | |
|--|------------------------------------|
| Pairwise Alignment | |
| Gap Opening Penalty | <input type="text" value="10.00"/> |
| Gap Extension Penalty | <input type="text" value="0.10"/> |
| Multiple Alignment | |
| Gap Opening Penalty | <input type="text" value="10.00"/> |
| Gap Extension Penalty | <input type="text" value="0.20"/> |
| Weight | |
| Use Negative Matrix | <input type="text" value="OFF"/> |
| Delay Divergent Cutoff (%) | <input type="text" value="30"/> |
| <input type="checkbox"/> Keep Predefined Gap | |

Figura 1. Parâmetros usados para o alinhamento global das sequências de pAgo utilizando o programa ClustalW.

3.2. Análise do contexto genético

Para identificar supostos operons pAgo, as mesmas proteínas utilizadas para análise filogenética foram submetidas à análise de vizinhança genética usando o programa WebFlags (<https://server.atkinson-lab.com/webflags>). A vizinhança analisada consistia em três genes à montante e três genes à jusante do gene pAgo. As proteínas codificadas pelos genes à montante ou à jusante foram ainda caracterizadas com base na sua sequência e/ou estrutura, como indicado nos itens subsequentes.

3.3. Análise de sequência e estrutura de proteínas

A detecção de homologia baseada em sequência foi realizada usando HMMER (<https://www.ebi.ac.uk/Tools/hmmer/>). As potenciais regiões transmembranas foram previstas por meio do programa DeepTMHMM (<https://dtu.biolib.com/DeepTMHMM>). Para análise baseada em estrutura, os modelos gerados pelo programa AlphaFold foram obtidos do banco de dados EBI (<https://www.ebi.ac.uk>) ou construídos usando a implementação ColabFold do AlphaFold (<https://alphafold.ebi.ac.uk/>). As proteínas estruturalmente relacionadas no PDB foram identificadas por pesquisas com FoldSeek (<https://search.foldseek.com/>). Análise de alinhamento das proteínas com proteínas funcionais de interesse foram feitas usando EMBOSS Needle (https://www.ebi.ac.uk/jdispatcher/psa/emboss_needle).

4. RESULTADOS E DISCUSSÃO

4.1 Anotação das sequências homólogas e construção da árvore filogenética

A Pesquisa BLAST contra o banco de dados de proteínas Refseq revelou 30 novas sequências, totalizando 36 Argonautas, dentre elas, 3 bactérias, 32 arqueias e 1 eucarionte, que foi usado como outgroup para enraizamento da árvore. Das bactérias sendo duas termofílicas, uma mesófila e, das arqueias, 12 termofílicas e 20 mesófilas (Tabela Complementar 1). Foram excluídas da análise proteínas hipotéticas e sequências de múltiplas espécies. Dessa forma, foi possível identificar três clusters na análise filogenética dos Argonautas, dois grandes grupos que separam os organismos termófilos dos mesófilos e um grupo menor contando dois organismos termofílicos e um organismo mesofilo (*Archaeoglobus fulgidus*, *Petrotoga miotherma*, *Treponema medium*) (Figura 2).

A árvore filogenética foi obtida através software IQ-TREE 2 (Minh et al., 2020), que inferiu a máxima verossimilhança usando como modelo Q.pfam+F+R4 obtido pelo ModelFinder (S. Kalyaanamoorthy et. al., 2017).

Notou-se aqui o afastamento na árvore filogenética do argonauta WP_048159825.1 do organismo *Thermococcus barophilus* em relação aos demais organismos do mesmo genus *Thermococcus*, *Thermococcus eurythermalis*, *Thermococcus onnurineus*, *Thermococcus thermotolerans*, *Thermococcus thioreducens*, representados pelos argonautas WP_050002102.1, WP_012572468.1, WP_258084286.1, WP_055429304.1, respectivamente. As demais análises no trabalho serão feitas no intuito de buscar compreender o motivo dessa distinção evolutiva entre os argonautas do mesmo genus e buscar compreender os principais mecanismos que estão associados a esse argonauta em questão.

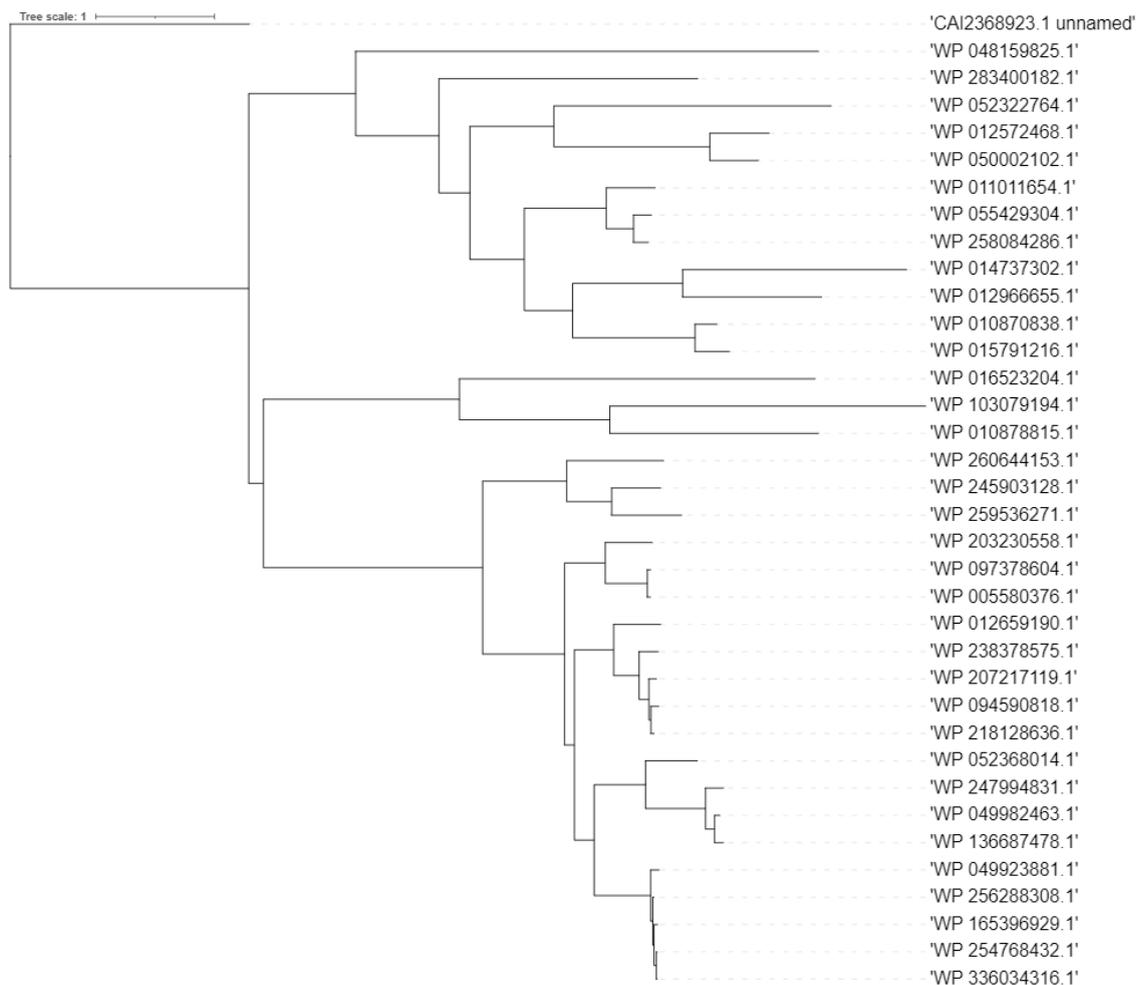


Figura 2. Árvore filogenética consenso de pAgos gerada por máxima verossimilhança utilizando o software IQ-TREE 2.

4.2 Exploração dos dados e anotação das sequências.

4.2.1 MEME

Utilizou-se o programa MEME Suite (<https://meme-suite.org/meme/>) para explorar motivos conservados das sequências dos Argonautas. Verificou-se 56 motivos conservados (E-value < 0.005) dentre os 35 Argonautas analisados. É possível identificar similaridades dos motivos conservados dos organismos com proximidade no filograma (Figura 3), assim como distinções de grandes grupos, como o cluster dos organismos

termofílicos e mesofílicos. Também foi possível identificar uma forte distinção daqueles organismos com seu domínio PAZ conservado daqueles do mesmo cluster que perderam o domínio ou possuem o domínio truncado (Figura 3). A análise dos motivos conservados, representados pelas caixas coloridas (Figura 3), associados à análise estrutural dessas proteínas, estão vinculados a ligações iônicas relevantes para ativação e funcionamento dos domínios dos Argonautas. Sabe-se, por exemplo, que o sítio catalítico de todas as proteínas Ago ativas contém uma tétrade conservada de resíduos de aminoácidos carregados negativamente, DEDX (onde X é D, H ou K) que quelam íons metálicos divalentes catalíticos, Mg²⁺ ou Mn²⁺ (Figura. 4). Mn²⁺ geralmente aumenta a atividade do pAgo, e alguns pAgos (PfAgo) mostraram-se ativos apenas na presença de íons manganês. (Lisitskaya et al., 2018; Makarova et al., 2009)

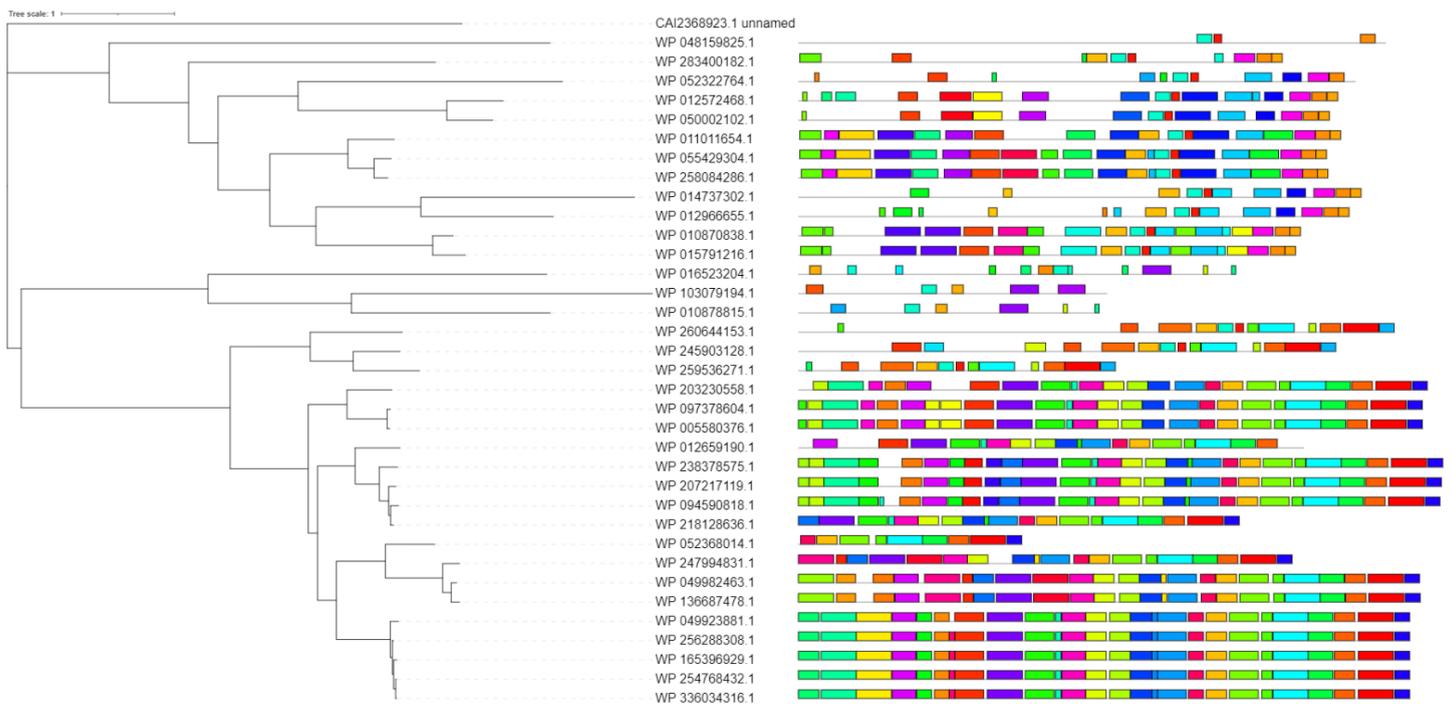


Figura 3. Árvore filogenética inferida por máxima verossimilhança à esquerda, e os motivos conservados que correspondem aos Argonautas à direita, deduzidos fazendo uso do programa MEME Suit. O clado de baixo está representado exclusivamente por organismos mesofílicos, o subclado com 3 organismos na parte central da árvore está representado por 2 bactérias, o argonauta 016523204.1 (*Treponema medium*) mesofílica e 103079194.1 (*Petrotoga miotherma*) termofílica, e uma arquea, o argonauta 010878815.1 (*Archaeoglobus fulgidus*) termofílica. O clado superior está composto exclusivamente por organismos termofílicos, sendo apenas um deles bacteriano, o argonauta WP_283400182.1 (*Desulfurobacterium pacificum*). A esquerda observa-se

diversas caixas coloridas, em que cada caixa faz referências a um motivo conservado diferente. Colorações iguais fazem referência ao mesmo motivo conservado. Nota-se uma forte correlação de motivos conservados em organismos pertencentes ao mesmo clado. O Organismos utilizado para enraizamento da árvore, CAI2368923.1 (*Moneuplotes crassus*), não faz parte dessa análise.

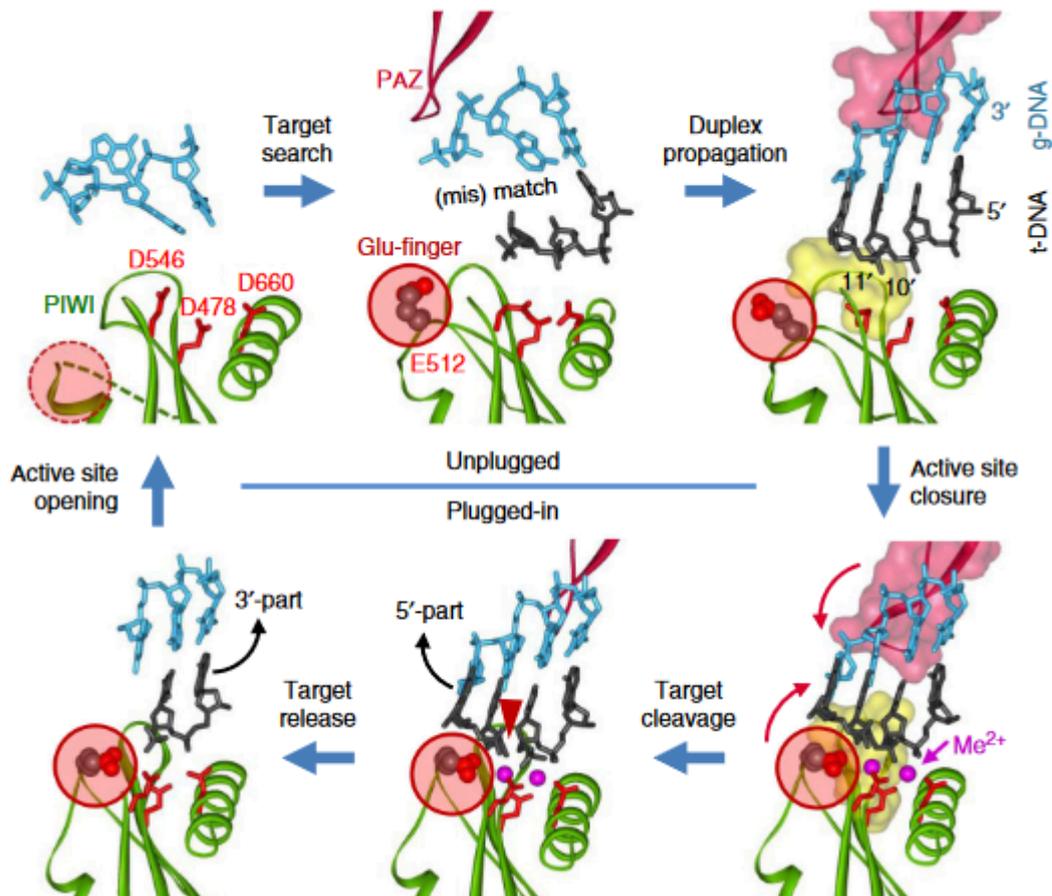


Figura 4. Mudanças conformacionais no sítio ativo do TtAgo durante o reconhecimento do alvo e catálise . Os resíduos do sítio ativo são mostrados em vermelho; o dedo de glutamato é indicado por um círculo vermelho. Durante as primeiras etapas de ligação do guia e reconhecimento do alvo, o sítio ativo é desconectado; a propagação duplex é acompanhada por mudanças nas conformações dos domínios PIWI e PAZ (indicados com setas vermelhas), plug-in do dedo de glutamato, ligação catalítica de metal e ativação da catálise. Finalmente, a liberação gradual do alvo leva à desconexão do sítio ativo, possibilitando assim o reconhecimento da próxima molécula alvo. Os números de acesso do PDB (do canto superior esquerdo, no sentido horário): 3DLH, 3F73, 4N41, 4NCB, 4NCA, 4N76 (adaptado de Lisitskaya et al.,2018)

4.2.2 WebFlag

Para identificar possíveis operons de pAgo, as mesmas proteínas utilizadas para a análise filogenética foram submetidas à análise de vizinhança genética utilizando o programa WebFlags (<https://server.atkinson-lab.com/webflags>). A vizinhança analisada consistiu em três genes à montante e três genes à jusante do gene pAgo. As proteínas codificadas pelos genes à montante ou à jusante foram caracterizadas com base em suas sequências.

Um conjunto de dados totalizando 302 sequências foi usado para inferir possíveis proteínas correspondentes e suas funções celulares (Figura 5 e Tabela Complementar 2.)



Figura 5. Representação gerada via Web Flags dos operons vizinhos as Argonautas (colorido) e o domínio PIWI das Argonautas (preto). Nota-se que diversos espaços em branco estão representados na figura, estas regiões não foram identificadas pelo software como tendo uma função definida. Esses genes foram analisados posteriormente e anotados via HMMER (tabela complementar 2).

Observou-se que proteínas Argonautas com funções truncadas pela perda do domínio PAZ possuem operons vizinhos que transcrevem proteínas efetoras, essas proteínas já foram descritas por formar heterodímeros com pseudo-pAgos, e esses se

assemelham funcionalmente a complexos compostos por APAZ (estruturalmente equivalente aos domínios N-L1-L2) e a proteína MID-PIWI (Makarova et al., 2006; Swarts et al., 2014; Bastiaanssen C. et al., 2024). Utilizando a metodologia de análise dos genes vizinhos dos Argonautas foram identificados dois organismos, o *Thermococcus barophilus* (WP_048159825.1) e o *Treponema medium* (WP_016523204.1) que possuem em estruturas vizinhas o sistema de resposta Abi e o domínio contendo a proteína SIR2-like, respectivamente. Essas observações podem indicar que apesar desses Argonautas terem suas funções truncadas, podem mediar diversas respostas de defesa associadas com as Argonautas completas.

4.2.3 HMMER e HHpred

Foi utilizado a sequência completa dos dados de saída do Web Flag para construir um perfil HMM v3.3.2 via HMMER (<http://hmmer.org/>) (Simon C Potter et al., 2018). O perfil funcional das proteínas vizinhas do *Thermococcus barophilus* sugere que os operons à montante do domínio PIWI WP_048159988.1 e WP_013468141.1, possuem funções associadas ao mecanismo de proteínas antitoxina AbiEi-AbiEii e toxina nucleotidil transferase AbiEii/AbiGii, respectivamente (Tabela complementar 2). Esse sistema de resposta Abi já foi associado ao domínio inativo PIWI SiAgo e suas proteínas associadas SiAga1 e SiAga2. Portanto, existe a possibilidade dos genes vizinhos estarem mediando a resposta de defesa associados ao Argonautas do TbAgo.

Foi conduzido uma aferição das sequências vizinhas do domínio PIWI do TbAgo, usando o banco de dados HHpred (<https://toolkit.tuebingen.mpg.de/tools/hhpred>) para identificar proteínas homologas. Foi constatada uma forte associação das funções anteriormente descritas utilizando o HMMER e utilizou-se a proteína homóloga com menor E-valor e maior Score para análise estrutural posterior (Figura 6).

A análise HHpred (Söding et al., 2005) das proteínas associadas sugere que WP_013468137.1, a sequência imediata à jusante do operon do domínio PIWI do TbAgo, é codificado pela transposase Tn916-like (Tabela 2), que faz parte dos transposons conjugativos (CTNs; também chamados de elementos conjugativos integrativos), que constituem uma classe importante de transposons de DNA que podem se mover de forma autossuficiente entre os genomas bacterianos, estes são enzimas altamente promíscuas, especialmente para integração, podendo ser inseridas em sites aleatórios (Rubio-Cosials et al., 2018).

Tabela 2. 6EMY_B, 6Y8Q_B, PF08843.16 são proteínas homólogas que correspondem as proteínas vizinhas do TbAgo, WP_013468137.1, WP_048159988.1, WP_013468141.1 respectivamente. Análise feita via HHpred .

| Hit | Nome | Probabilidade | E-Value | Score |
|------------|---|---------------|---------|--------|
| 6EMY_B | TIROSINA RECOMBINASE Y-TRANSPOSASE, TN916-LIKE | 100 | 4.1e-30 | 223.29 |
| 6Y8Q_B | Proteína Abi; Antitoxin, DNA-ligante, AbiE. Toxina-antitoxina | 99.74 | 1.7e-15 | 106.24 |
| PF08843.16 | AbiEii; Nucleotidil transferase AbiEii toxina | 99.46 | 3.2e-12 | 111.56 |

4.3 Análises estruturais

4.3.1 EMBOSS Needle

As proteínas WP_048159988.1 e WP_013468141.1 estão à montante e são semelhantes (pontuação 84.5, EMBOSS Needle) ao efector de membrana SiAga2 do sistema antiviral SiAgo da arqueia *Sulfolobus islandicus*. O domínio PIWI (WP_048159825.1) se assemelha ao complexo SiAgo-Aga1 que computa (pontuação 107.5, EMBOSS Needle) (Zeng et al., 2022; Krogh et al., 2001). (Material suplementar 1).

Foi demonstrado que SiAgo e SiAga1 formam um complexo estável que exhibe capacidade de reconhecimento de ácido nucleico dirigida por guia. No complexo, SiAga1 contribui significativamente para a ligação ao ácido nucleico, sugerindo que pode complementar a perda dos domínios N-terminal e PAZ no SiAgo, que funcionam na ligação guia e alvo em pAgos longos.(Kaya et al., 2016; Ma et al., 2004; Song et al., 2004).

É concebível que o domínio APAZ e/ou domínios de ligação ao DNA se fundam com outros pAgos curtos (Makarova et al., 2009; Ryazansky et al., 2018) também pode contribuir para orientar e direcionar a vinculação. Além disso, o domínio MID do SiAgo desempenha um papel na estabilização da guia e, assim, na facilitação da ligação ao alvo, em linha com o facto de o domínio interagir com a extremidade 5' da cadeia guia.(Ma et al., 2005; Miyoshi et al., 2016; Willkomm et al., 2017).

Além disso, os resíduos conservados no domínio MID são essenciais para a mediação de Abi, sugerindo que a capacidade de reconhecimento de ácido nucleico do

complexo SiAgo-Aga1 é crítica para a imunidade. Considerando as interações entre SiAgo-Aga1 e SiAga2, foi proposto um modelo (Figura 6) que explica como o sistema SiAgo media Abi: o complexo SiAgo-Aga1 funciona como o módulo sensor que reconhece os ácidos nucleicos invasores e ativa o SiAga2 por interação direta; além disso, o complexo SiAgo-Aga1-Aga2 na membrana também pode detectar a infecção viral e ser ativado. Então, o SiAga2 ativado desencadeia a despolarização da membrana e causa a morte celular. (Zeng et al., 2022).

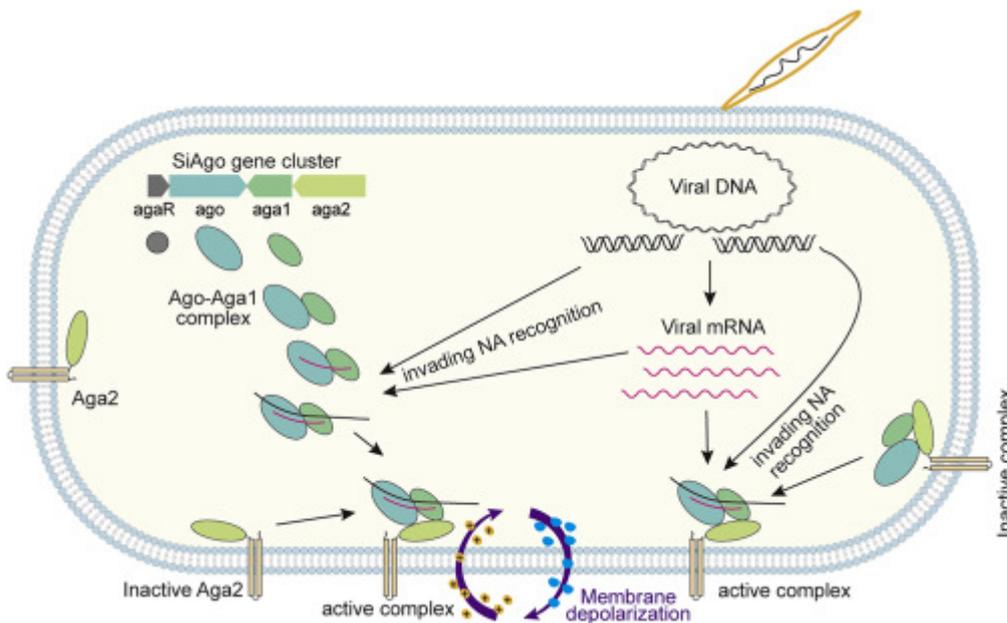


Figura 6. Modelo para resposta imune antiviral do sistema SiAgo. Antes da infecção viral, o complexo SiAgo-Aga1 reside no citoplasma ou liga-se ao SiAga2 na membrana sem desencadeá-lo. Durante a infecção viral, a alta transcrição de genes virais e a rápida replicação do genoma viral fornecem substratos abundantes de RNA e DNA para o complexo SiAgo-Aga1 obter guias e também para o complexo carregado com guia procurar alvos. A ligação ao alvo leva ao recrutamento de mais complexos SiAgo-Aga1 para a proteína transmembrana SiAga2 que é ativada, induzindo a morte de células infectadas através da despolarização da membrana. Além disso, o complexo de membrana SiAgo-Aga1-Aga2 também pode detectar a infecção viral pelo reconhecimento do ácido nucleico e ser ativado.(adaptado de Zeng et al., 2022)

Sendo assim, de acordo com nossa análise preliminar dos dados, deve-se encontrar o Argonautas TbAgo posicionado entre sequências de transposição e proteínas de resistência Abi, indicando possível ação do Argonautas na defesa do organismo e transferência de genes de resistência .

Para compreender completamente a dinâmica molecular desses organismos e suas Argonautas, bem como suas possíveis aplicações biotecnológicas, é necessário

compreender como esses domínios efetores interagem em resposta a estímulos ou superexpressão. (van Houte et al., 2016).(Fusco et al., 2015; Quax et al., 2013).

4.3.2 TMHMM

A análise da sequência da proteína usando o servidor TMHMM (Krogh et al., 2001) foi inconclusivo em prever que a proteína WP_048159988.1, vizinha do TbAgo, contém regiões transmembranas (Figura 7). Novas análises *in vitro* são necessárias para validar a hipótese de que a proteína está presente na membrana celular, uma vez que o mecanismo descrito por SiAgo2 foi confirmado como sendo mediado por uma proteína transmembrana.(Figura 7.)

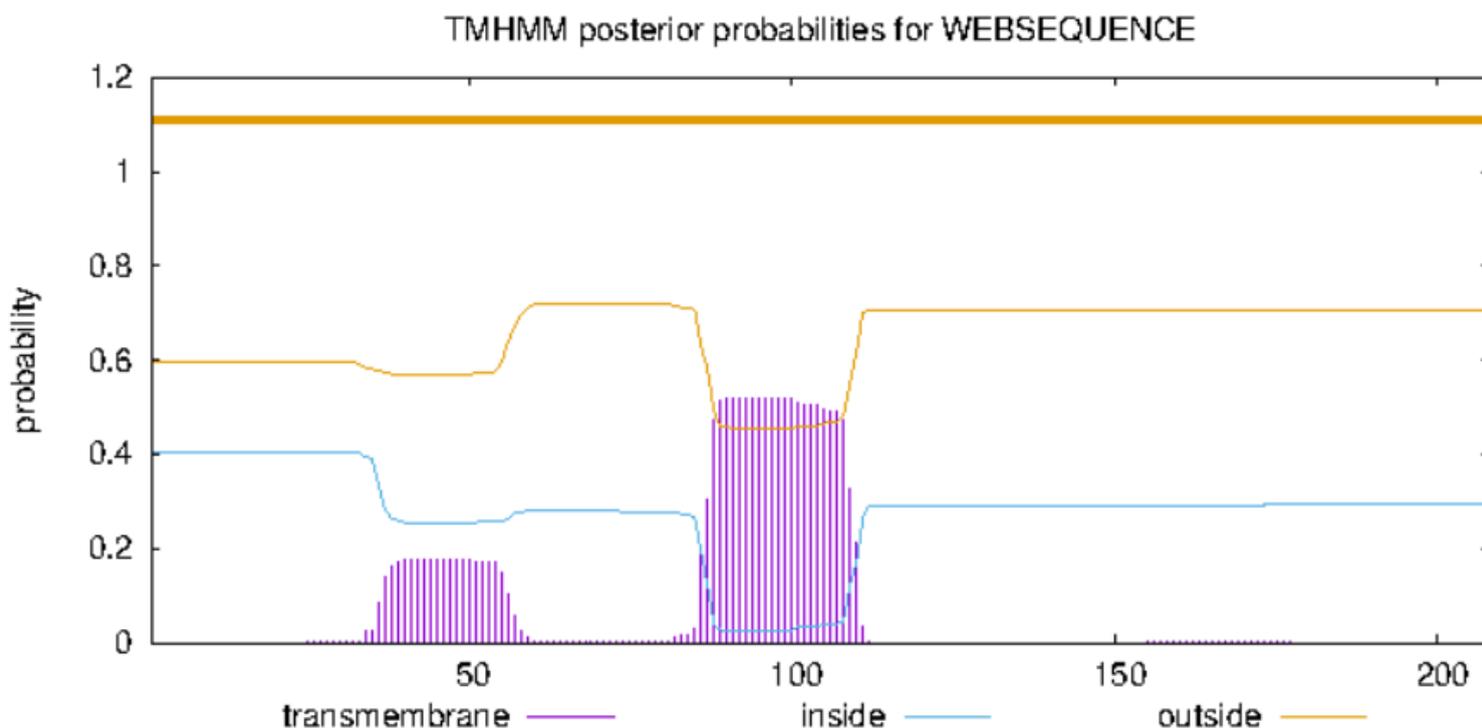


Figura 7. Proteína vizinha do TbAgo (WP_048159988.1) analisada pelo servidor do TMHMM para proteína transmembrana.

Compreende-se, até os dias atuais, que os mecanismos de resposta a infecção mediadas por proteínas efetoras podem ocorrer através de três sistemas distintos (figura 8). Este estudo visa compreender o mecanismos de atuação do TbAgo, usando ferramentas que analisam as funções moleculares dessa proteína, usando proteínas semelhantes e com suas ações moleculares mais bem definidas na literatura. Portanto é necessário que se explore ainda mais essas relações a fim de encontrar dados que

sustentem a hipótese de um possível mecanismo molecular similar ao sistema SiAgo ou a outro mecanismo de resposta Abi.

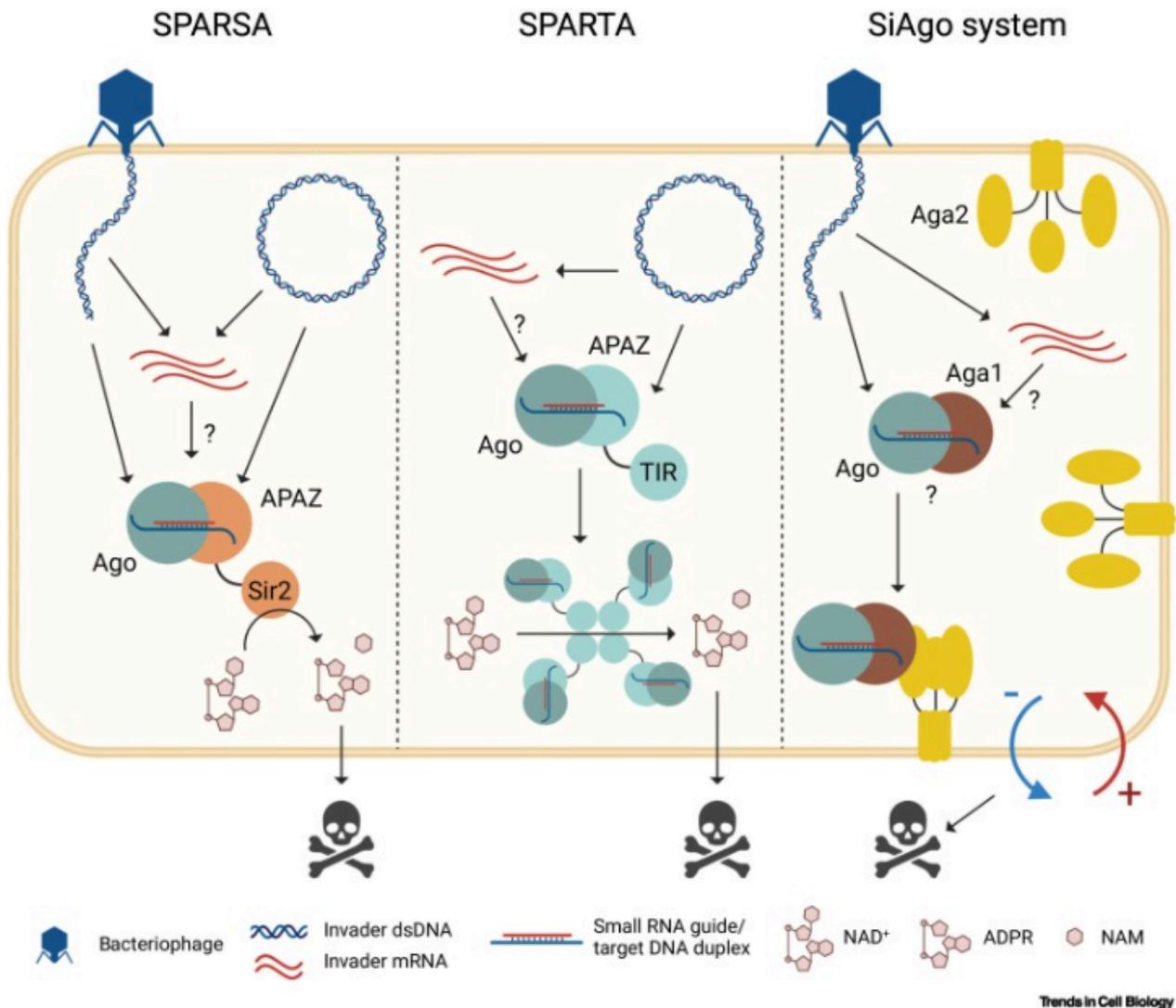


Figura 8. SIR2 (SPARSA), TIR (SPARTA) ou Aga2 (sistema SiAgo). Em SPARSA e SPARTA, a ativação de SIR2/TIR leva à conversão de NAD(P)+ em NAM e ADPR(P). No SPARTA, a ativação catalítica requer a tetramerização de quatro complexos SPARTA heterodiméricos ligados ao guia/alvo. Após a detecção do invasor, o complexo SiAgo-SiAga1 é recrutado para a proteína de membrana SiAga2, que induz a despolarização da membrana.(adaptado de Koopal et al., 2023)

5.CONCLUSÃO E PERSPECTIVAS

O estudo das proteínas Argonauta (Ago) têm revelado uma complexa diversidade funcional e estrutural nos três domínios da vida: bactérias, arqueas e eucariotos. Este

trabalho buscou explorar a diversidade evolutiva e os domínios conservados das Argonautas, especialmente em organismos procariontes mesofílicos e termofílicos, com vistas a suas possíveis aplicações biotecnológicas.

A análise filogenética e de sequência revelou uma clara distinção entre os Argonautas termofílicos e mesofílicos, além da identificação de domínios conservados, como PAZ, MID e PIWI, que são essenciais para a função dessas proteínas. A identificação de novos pAgos, particularmente em arqueias, sugere um vasto campo ainda inexplorado, com potencial significativo para a biotecnologia.

Os Argonautas termofílicos demonstraram propriedades promissoras para manipulação genética programável, similar à atividade CRISPR-Cas9. Estes pAgos não apenas clivam ácidos nucleicos de forma precisa, mas também exibem potencial para serem utilizados em condições extremas, o que pode ser vantajoso para aplicações industriais onde as plataformas mesofílicas convencionais falham.

O avanço no entendimento das funções e mecanismos moleculares dos pAgos abre várias frentes de pesquisa e aplicações práticas:

- **Desenvolvimento de Ferramentas de Edição Genética:** Explorar a atividade endonuclease dos pAgos termofílicos para desenvolver novas ferramentas de edição genética que possam operar em ambientes extremos ou em condições onde as enzimas atuais não são eficazes.
- **Estudos Funcionais de pAgos e suas Proteínas Associadas:** Realizar estudos funcionais detalhados para entender a interação entre pAgos e suas proteínas associadas, como os domínios APAZ, SIR2 e TIR, e como estas interações podem ser aproveitadas para desenvolver novos mecanismos de defesa contra invasores genômicos.
- **Biotecnologia Industrial:** Implementar pAgos em processos biotecnológicos industriais, especialmente aqueles que envolvem biomineração, PCR de alta fidelidade, e biossensores, aproveitando as propriedades termoestáveis e de alta especificidade dessas proteínas.
- **Caracterização de Novos Organismos:** Continuar a caracterização de Argonautas em novos organismos, especialmente em arqueias que têm sido negligenciadas, para descobrir novas enzimas e mecanismos que possam ser utilizados em diversas aplicações biotecnológicas.

- **Síntese de Engenharia Metabólica:** Utilizar pAgos em engenharia metabólica para otimizar a produção de biocombustíveis, produtos químicos e outros bioprodutos em organismos termofílicos, que podem oferecer vantagens sobre os sistemas mesofílicos tradicionais.

O contínuo desenvolvimento de técnicas "ômicas" e a aplicação de ferramentas avançadas de bioinformática e biologia estrutural são essenciais para avançar neste campo. A exploração dos pAgos, especialmente aqueles de ambientes extremos, promete revolucionar a biotecnologia e abrir novas possibilidades para a ciência e a indústria.

6. REFERÊNCIAS

- Altenhoff, Adrian M., e Christophe Dessimoz. "Phylogenetic and Functional Assessment of Orthologs Inference Projects and Methods". *PLoS Computational Biology* 5, nº 1 (16 de janeiro de 2009): e1000262. <https://doi.org/10.1371/journal.pcbi.1000262>.
- Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell*. 2004;116(2):281-297. doi:10.1016/S0092-8674(04)00045-5
- Bastiaanssen C, Bobadilla Ugarte P, Kim K, et al. RNA-guided RNA silencing by an Asgard archaeal Argonaute. *Nat Commun*. 2024;15(1):5499. doi:10.1038/s41467-024-49452-1
- B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R. Lanfear (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.*, 37:1530-1534. <https://doi.org/10.1093/molbev/msaa015>
- Chong, Yuesheng, Qian Liu, Fei Huang, Dong Song, e Yan Feng. "Characterization of a Recombinant Thermotolerant Argonaute Protein as an Endonuclease by Broad Guide Utilization". *Bioresources and Bioprocessing* 6, nº 1 (5 de junho de 2019): 21. <https://doi.org/10.1186/s40643-019-0254-8>.
- Dong, Huarong, Fei Huang, Xiang Guo, Xiaoyi Xu, Qian Liu, Xiao Li, e Yan Feng. "Characterization of Argonaute Nucleases from Mesophilic Bacteria *Paenibacillus Borealis* and *Brevibacillus Laterosporus*". *Bioresources and Bioprocessing* 8, nº 1 (19 de dezembro de 2021): 133. <https://doi.org/10.1186/s40643-021-00478-z>.

- Doron, Shany, Sarah Melamed, Gal Ofir, Azita Leavitt, Anna Lopatina, Mai Keren, Gil Amitai, e Rotem Sorek. “Systematic Discovery of Antiphage Defense Systems in the Microbial Pangenome”. *Science (New York, N.Y.)* 359, nº 6379 (2 de março de 2018): eaar4120. <https://doi.org/10.1126/science.aar4120>.
- Doxzen, Kevin W., e Jennifer A. Doudna. “DNA Recognition by an RNA-Guided Bacterial Argonaute”. *PLOS ONE* 12, nº 5 (17 de maio de 2017): e0177097. <https://doi.org/10.1371/journal.pone.0177097>.
- Esyunina, Daria, Anastasiia Okhtienko, Anna Olina, Vladimir Panteleev, Maria Prostova, Alexei A. Aravin, e Andrey Kulbachinskiy. “Specific Targeting of Plasmids with Argonaute Enables Genome Editing”. *Nucleic Acids Research* 51, nº 8 (8 de maio de 2023): 4086–99. <https://doi.org/10.1093/nar/gkad191>.
- Arezi B, Hansen CJ, Hogrefe HH. Efficient and high fidelity incorporation of dye-terminators by a novel archaeal DNA polymerase mutant. *J Mol Biol* 2002;322:719–29.
- Fang, Gang, Nitin Bhardwaj, Rebecca Robilotto, e Mark B. Gerstein. “Getting Started in Gene Orthology and Functional Analysis”. *PLoS Computational Biology* 6, nº 3 (26 de março de 2010): e1000703. <https://doi.org/10.1371/journal.pcbi.1000703>.
- Fang, Mengjun, Zhipeng Xu, Di Huang, Muhammad Naeem, Xiangcheng Zhu, e Zhinan Xu. “Characterization and Application of a Thermophilic Argonaute from Archaeon *Thermococcus Thioreducens*”. *Biotechnology and Bioengineering* 119, nº 9 (2022): 2388–98. <https://doi.org/10.1002/bit.28153>.
- Felsenstein, Joseph. “Confidence Limits on Phylogenies: An Approach Using the Bootstrap”. *Evolution* 39, nº 4 (1985): 783–91. <https://doi.org/10.2307/2408678>.
- García-Quintans, Nieves, Laurie Bowden, José Berenguer, e Mario Mencía. “DNA interference by a mesophilic Argonaute protein, CbcAgo”. *F1000Research* 8 (9 de janeiro de 2020): 321. <https://doi.org/10.12688/f1000research.18445.2>.
- Huang, Shan, Kaihang Wang, e Stephen L. Mayo. “Genome Manipulation by Guide-Directed Argonaute Cleavage”. *Nucleic Acids Research* 51, nº 8 (8 de maio de 2023): 4078–85. <https://doi.org/10.1093/nar/gkad188>.
- Iwai H, Lingel A, Pluckthun A. Cyclic green fluorescent protein produced in vivo using an artificially split PI-PfuI intein from *Pyrococcus furiosus*. *J Biol Chem* 2001;276:16548–54.
- Jiang, Xiaoman, Yang Liu, Qi Liu, e Lixin Ma. “Characterization of a Programmable Argonaute Nuclease from the Mesophilic Bacterium *Rummeliibacillus*

- Suwonensis”. *Biomolecules* 12, nº 3 (março de 2022): 355. <https://doi.org/10.3390/biom12030355>.
- Ka, Donghyun, Hyejin Oh, Eunyoung Park, Jeong-Han Kim, e Euiyoung Bae. “Structural and Functional Evidence of Bacterial Antiphage Protection by Thoeris Defense System via NAD⁺ Degradation”. *Nature Communications* 11, nº 1 (4 de junho de 2020): 2816. <https://doi.org/10.1038/s41467-020-16703-w>.
- Kaya E, Doxzen KW, Knoll KR, et al. A bacterial Argonaute with noncanonical guide RNA specificity. *Proceedings of the National Academy of Sciences*. 2016;113(15):4057-4062. doi:10.1073/pnas.1524385113
- Koopal, Balwina, Sumanth K. Mutte, e Daan C. Swarts. “A long look at short prokaryotic Argonautes”. *Trends in Cell Biology* 33, nº 7 (1º de julho de 2023): 605–18. <https://doi.org/10.1016/j.tcb.2022.10.005>.
- Kropocheva, E. V., L. A. Lisitskaya, A. A. Agapov, A. A. Musabirov, A. V. Kulbachinskiy, e D. M. Esyunina. “Prokaryotic Argonaute Proteins as a Tool for Biotechnology”. *Molecular Biology* 56, nº 6 (2022): 854–73. <https://doi.org/10.1134/S0026893322060103>.
- Koopal B, Potocnik A, Mutte SK, et al. Short prokaryotic Argonaute systems trigger cell death upon detection of invading DNA. *Cell*. 2022;185(9):1471-1486.e19. doi:10.1016/j.cell.2022.03.012
- Kwon KM, Kang SG, Sokolova TG et al.. Characterization of a family B DNA polymerase from *Thermococcus barophilus* Ch5 and its application for long and accurate PCR. *Enzyme Microb Technol* 2016;86:117–26.
- Kuzmenko, Anton, Anastasiya Oguienko, Daria Esyunina, Denis Yudin, Mayya Petrova, Alina Kudinova, Olga Maslova, et al. “DNA Targeting and Interference by a Bacterial Argonaute Nuclease”. *Nature* 587, nº 7835 (novembro de 2020): 632–37. <https://doi.org/10.1038/s41586-020-2605-1>.
- Lennon CW, Stanger M, Belfort M. Protein splicing of a recombinase intein induced by ssDNA and DNA damage. *Genes & Dev* 2016;30:2663–68.
- Li, Wenqiang, Yang Liu, Ruyi He, Longyu Wang, Yaping Wang, Wanting Zeng, Zhiwei Zhang, Fei Wang, e Lixin Ma. “A Programmable pAgo Nuclease with RNA Target Preference from the Psychrotolerant Bacterium *Mucilaginibacter Paludis*”. *Nucleic Acids Research* 50, nº 9 (20 de maio de 2022): 5226–38. <https://doi.org/10.1093/nar/gkac315>.
- Li, Xiao, Huarong Dong, Xiang Guo, Fei Huang, Xiaoyi Xu, Nuolan Li, Yu Yang, Tianbao Yao, Yan Feng, e Qian Liu. “Mesophilic Argonaute-based isothermal detection of

- SARS-CoV-2". *Frontiers in Microbiology* 13 (2022).
<https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2022.957977>.
- Lisitskaya, Lidiya, Alexei A. Aravin, e Andrey Kulbachinskiy. "DNA interference and beyond: structure and functions of prokaryotic Argonaute proteins". *Nature Communications* 9 (4 de dezembro de 2018): 5165.
<https://doi.org/10.1038/s41467-018-07449-7>.
- Lisitskaya, Lidiya, Ekaterina Kropocheva, Aleksei Agapov, Maria Prostova, Vladimir Panteleev, Denis Yudin, Sergei Ryazansky, et al. "Bacterial Argonaute Nucleases Reveal Different Modes of DNA Targeting in Vitro and in Vivo". *Nucleic Acids Research* 51, n° 10 (9 de junho de 2023): 5106–24.
<https://doi.org/10.1093/nar/gkad290>.
- Lisitskaya, Lidiya, Yeonoh Shin, Aleksei Agapov, Anna Olina, Ekaterina Kropocheva, Sergei Ryazansky, Alexei A. Aravin, Daria Esyunina, Katsuhiko S. Murakami, e Andrey Kulbachinskiy. "Programmable RNA Targeting by Bacterial Argonaute Nucleases with Unconventional Guide Binding and Cleavage Specificity". *Nature Communications* 13, n° 1 (8 de agosto de 2022): 4624.
<https://doi.org/10.1038/s41467-022-32079-5>.
- Lu Z, Xia R, Zhang S, et al. Evolution of optimal growth temperature in Asgard archaea inferred from the temperature dependence of GDP binding to EF-1A. *Nat Commun.* 2024;15(1):515. doi:10.1038/s41467-024-44806-1
- M N, Sa G, Rj C, W Y. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell.* 2005;121(7). doi:10.1016/j.cell.2005.04.024
- Ma JB, Ye K, Patel DJ. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. *Nature.* 2004;429(6989):318-322. doi:10.1038/nature02519
- Ma JB, Yuan YR, Meister G, Pei Y, Tuschl T, Patel DJ. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein. *Nature.* 2005;434(7033):666-670. doi:10.1038/nature03514
- McDonald JP, Hall A, Gasparutto D et al.. Novel thermostable Y-family polymerases: applications for the PCR amplification of damaged or ancient DNAs. *Nucleic Acids Res* 2006;34:1102–11.
- Makarova KS, Wolf YI, van der Oost J, Koonin EV. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of

- defense against mobile genetic elements. *Biology Direct*. 2009;4(1):29. doi:10.1186/1745-6150-4-29
- Makarova, Kira S., Nick V. Grishin, Svetlana A. Shabalina, Yuri I. Wolf, e Eugene V. Koonin. “A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action”. *Biology Direct* 1, nº 1 (16 de março de 2006): 7. <https://doi.org/10.1186/1745-6150-1-7>.
- Malone, Colin D., e Gregory J. Hannon. “Small RNAs as Guardians of the Genome”. *Cell* 136, nº 4 (20 de fevereiro de 2009): 656–68. <https://doi.org/10.1016/j.cell.2009.01.045>.
- Manakova, Elena, Edvardas Golovinas, Reda Pocevičiūtė, Giedrius Sasnauskas, Arunas Silanskas, Danielis Rutkauskas, Marija Jankunec, et al. “The missing part: the *Archaeoglobus fulgidus* Argonaute forms a functional heterodimer with an N-L1-L2 domain protein”. *Nucleic Acids Research* 52, nº 5 (9 de janeiro de 2024): 2530–45. <https://doi.org/10.1093/nar/gkad1241>.
- Miesen P, Joosten J, Rij RP van. PIWIs Go Viral: Arbovirus-Derived piRNAs in Vector Mosquitoes. *PLOS Pathogens*. 2016;12(12):e1006017. doi:10.1371/journal.ppat.1006017
- Miyoshi T, Ito K, Murakami R, Uchiumi T. Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat Commun*. 2016;7:11846. doi:10.1038/ncomms11846
- Ming, Dengming, Michael E Wall, e Kevin Y Sanbonmatsu. “Domain motions of Argonaute, the catalytic engine of RNA interference”. *BMC Bioinformatics* 8 (30 de novembro de 2007): 470. <https://doi.org/10.1186/1471-2105-8-470>.
- Ni D, Lu X, Stahlberg H, Ekundayo B. Activation mechanism of a short argonaute-TIR prokaryotic immune system. *Sci Adv*. 2023;9(29):eadh9002. doi:10.1126/sciadv.adh9002
- Ofir, Gal, Ehud Herbst, Maya Baroz, Daniel Cohen, Adi Millman, Shany Doron, Nitzan Tal, et al. “Antiviral Activity of Bacterial TIR Domains via Immune Signalling Molecules”. *Nature* 600, nº 7887 (dezembro de 2021): 116–20. <https://doi.org/10.1038/s41586-021-04098-7>.
- Ofir G, Herbst E, Baroz M, et al. Antiviral activity of bacterial TIR domains via signaling molecules that trigger cell death. Published online January 6, 2021:2021.01.06.425286. doi:10.1101/2021.01.06.425286

- Olovnikov, Ivan, Ken Chan, Ravi Sachidanandam, Dianne K. Newman, e Alexei A. Aravin. "Bacterial Argonaute Samples the Transcriptome to Identify Foreign DNA". *Molecular Cell* 51, n° 5 (12 de setembro de 2013): 594–605. <https://doi.org/10.1016/j.molcel.2013.08.014>.
- Parker, James S., Eneida A. Parizotto, Muhan Wang, S. Mark Roe, e David Barford. "Enhancement of the Seed-Target Recognition Step in RNA Silencing by a PIWI/MID Domain Protein". *Molecular Cell* 33, n° 2 (30 de janeiro de 2009): 204–14. <https://doi.org/10.1016/j.molcel.2008.12.012>.
- Parker, James S., S. Mark Roe, e David Barford. "Crystal Structure of a PIWI Protein Suggests Mechanisms for siRNA Recognition and Slicer Activity". *The EMBO Journal* 23, n° 24 (8 de dezembro de 2004): 4727–37. <https://doi.org/10.1038/sj.emboj.7600488>.
- Parker, J., Roe, S. & Barford, D. Structural insights into mRNA recognition from a PIWI domain–siRNA guide complex. *Nature* 434, 663–666 (2005). <https://doi.org/10.1038/nature03462>
- Potter, Simon C, Aurélien Luciani, Sean R Eddy, Youngmi Park, Rodrigo Lopez, e Robert D Finn. "HMMER web server: 2018 update". *Nucleic Acids Research* 46, n° W1 (2 de julho de 2018): W200–204. <https://doi.org/10.1093/nar/gky448>.
- Qi, Jialing, Zhangji Dong, Yunwei Shi, Xin Wang, Yinyin Qin, Yongming Wang, e Dong Liu. "NgAgo-Based Fabbp11a Gene Knockdown Causes Eye Developmental Defects in Zebrafish". *Cell Research* 26, n° 12 (dezembro de 2016): 1349–52. <https://doi.org/10.1038/cr.2016.134>.
- Rashid, Umar Jan, Dirk Paterok, Alexander Koglin, Holger Gohlke, Jacob Piehler, e Julian C.-H. Chen. "Structure of Aquifex Aeolicus Argonaute Highlights Conformational Flexibility of the PAZ Domain as a Potential Regulator of RNA-Induced Silencing Complex Function". *The Journal of Biological Chemistry* 282, n° 18 (4 de maio de 2007): 13824–32. <https://doi.org/10.1074/jbc.M608619200>.
- Rubio-Cosials, Anna, Eike C. Schulz, Lotte Lambertsen, Georgy Smyshlyaev, Carlos Rojas-Cordova, Kristoffer Forslund, Ezgi Karaca, Aleksandra Bebel, Peer Bork, e Orsolya Barabas. "Transposase-DNA Complex Structures Reveal Mechanisms for Conjugative Transposition of Antibiotic Resistance". *Cell* 173, n° 1 (22 de março de 2018): 208-220.e20. <https://doi.org/10.1016/j.cell.2018.02.032>.

- Ryazansky, Sergei, Andrey Kulbachinskiy, e Alexei A. Aravin. "The Expanded Universe of Prokaryotic Argonaute Proteins". *mBio* 9, nº 6 (18 de dezembro de 2018): 10.1128/mbio.01935-18. <https://doi.org/10.1128/mbio.01935-18>.
- S. Kalyaanamoorthy, B.Q. Minh, T.K.F. Wong, A. von Haeseler, L.S. Jermini (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods*, 14:587-589. <https://doi.org/10.1038/nmeth.4285>
- Miesen P, Joosten J, van Rij RP (2016) PIWIs Go Viral: Arbovirus-Derived piRNAs in Vector Mosquitoes. *PLOS Pathogens* 12(12): e1006017. <https://doi.org/10.1371/journal.ppat.1006017>
- Smagin VA, Mardanov AV, Bonch-Osmolovskaia EA et al.. Isolation and characteristics of new thermostable DNA ligase from archaea of the genus *Thermococcus*. *Prikl Biokhim Mikrobiol* 2008;44:523–8.
- Song JJ, Smith SK, Hannon GJ, Joshua-Tor L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*. 2004;305(5689):1434-1437. doi:10.1126/science.1102514
- Sriskanda V, Kelman Z, Hurwitz J et al.. Characterization of an ATP-dependent DNA ligase from the thermophilic archaeon *Methanobacterium thermoautotrophicum*. *Nucleic Acids Res* 2000;28:2221–8.
- Straub, Christopher T, James A Counts, Diep M N Nguyen, Chang-Hao Wu, Benjamin M Zeldes, James R Crosby, Jonathan M Conway, et al. "Biotechnology of extremely thermophilic archaea". *FEMS Microbiology Reviews* 42, nº 5 (25 de junho de 2018): 543–78. <https://doi.org/10.1093/femsre/fuy012>.
- Swarts, Daan C., Kira Makarova, Yanli Wang, Kotaro Nakanishi, René F. Ketting, Eugene V. Koonin, Dinshaw J. Patel, e John van der Oost. "The Evolutionary Journey of Argonaute Proteins". *Nature Structural & Molecular Biology* 21, nº 9 (setembro de 2014): 743–53. <https://doi.org/10.1038/nsmb.2879>.
- Tanabe M, Ishino S, Yohda M et al.. Structure-based mutational study of an archaeal DNA ligase towards improvement of ligation activity. *ChemBioChem* 2012;13:2575–82.
- Tamura, Koichiro, Glen Stecher, e Sudhir Kumar. "MEGA11: Molecular Evolutionary Genetics Analysis Version 11". *Molecular Biology and Evolution* 38, nº 7 (1º de julho de 2021): 3022–27. <https://doi.org/10.1093/molbev/msab120>.
- Thompson, J. D., D. G. Higgins, e T. J. Gibson. "CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting,

- Position-Specific Gap Penalties and Weight Matrix Choice”. *Nucleic Acids Research* 22, n° 22 (11 de novembro de 1994): 4673–80. <https://doi.org/10.1093/nar/22.22.4673>.
- Willkomm, Sarah, Adrian Zander, Alexander Gust, e Dina Grohmann. “A Prokaryotic Twist on Argonaute Function”. *Life (Basel, Switzerland)* 5, n° 1 (16 de fevereiro de 2015): 538–53. <https://doi.org/10.3390/life5010538>.
- Xu, Xiaoyi, Hao Yang, Huarong Dong, Xiao Li, Qian Liu, e Yan Feng. *Characterization of Argonaute Nucleases from Mesophilic Bacteria Pseudobutyrvibrio ruminis*, 2023. <https://doi.org/10.21203/rs.3.rs-3532969/v1>.
- Yuan, Yu-Ren, Yi Pei, Hong-Ying Chen, Thomas Tuschl, e Dinshaw J. Patel. “A Potential Protein-RNA Recognition Event along the RISC-Loading Pathway from the Structure of *A. aeolicus* Argonaute with Externally Bound siRNA”. *Structure (London, England: 1993)* 14, n° 10 (outubro de 2006): 1557–65. <https://doi.org/10.1016/j.str.2006.08.009>.
- Yuan, Yu-Ren, Yi Pei, Jin-Biao Ma, Vitaly Kuryavyi, Maria Zhadina, Gunter Meister, Hong-Ying Chen, Zbigniew Dauter, Thomas Tuschl, e Dinshaw J. Patel. “Crystal Structure of *A. Aeolicus* Argonaute, a Site-Specific DNA-Guided Endoribonuclease, Provides Insights into RISC-Mediated mRNA Cleavage”. *Molecular Cell* 19, n° 3 (5 de agosto de 2005): 405–19. <https://doi.org/10.1016/j.molcel.2005.07.011>.
- Zander, Adrian, Sarah Willkomm, Sapir Ofer, Marleen Van Wolferen, Luisa Egert, Sabine Buchmeier, Sarah Stöckl, et al. “Guide-Independent DNA Cleavage by Archaeal Argonaute from *Methanocaldococcus Jannaschii*”. *Nature Microbiology* 2, n° 6 (20 de março de 2017): 17034. <https://doi.org/10.1038/nmicrobiol.2017.34>.
- Zeng, Zhifeng, Yu Chen, Rafael Pinilla-Redondo, Shiraz A. Shah, Fen Zhao, Chen Wang, Zeyu Hu, et al. “A short prokaryotic Argonaute activates membrane effector to confer antiviral defense”. *Cell Host & Microbe* 30, n° 7 (13 de julho de 2022): 930-943.e6. <https://doi.org/10.1016/j.chom.2022.04.015>.
- Zhang L, Tripathi A. Archaeal RNA ligase from *Thermococcus kodakarensis* for template dependent ligation. *RNA Biol* 2017;14:36–44.
- Zheng, Yaqi, Yang Zhao, Kunming Dong, Li Miao, Xiaojian Zhou, Yong Gong, e Likui Zhang. “A novel Mre11 protein from the hyperthermophilic euryarchaeon *Thermococcus barophilus* Ch5 possesses 5'–3' exonuclease and endonuclease activities”. *International Journal of Biological Macromolecules* 272 (1° de junho de 2024): 132654. <https://doi.org/10.1016/j.ijbiomac.2024.132654>.

- Zaremba M, Dakineviciene D, Golovinas E, et al. Short prokaryotic Argonautes provide defence against incoming mobile genetic elements through NAD⁺ depletion. *Nat Microbiol.* 2022;7(11):1857-1869. doi:10.1038/s41564-022-01239-0
- Willkomm S, Oellig CA, Zander A, et al. Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein. *Nat Microbiol.* 2017;2:17035. doi:10.1038/nmicrobiol.2017.35
- Wang X, Li X, Yu G, et al. Structural insights into mechanisms of Argonaute protein-associated NADase activation in bacterial immunity. *Cell Res.* 2023;33(9):699-711. doi:10.1038/s41422-023-00839-7

7. ANEXOS

Material suplementar 1. Análises EMBOSS Needle.

- https://drive.google.com/file/d/1muWO8zEcFW576UiEep3MhWyW0i8mJ-3z/view?usp=drive_link (TbAgo(PIWI) - SiAgo-Aga1)
- https://drive.google.com/file/d/1iXBCLMpLDqdO7HcUoB5jdWpw5ZEC-zwj/view?usp=drive_link (TbAgo_upstream(Abi) - SiAgo2)

Tabelas complementares 1 & 2. Blastp e HMMER.

- https://docs.google.com/spreadsheets/d/1BtXRf2MrVge0tQo28DyPYJJ5dmi21RJS5AicczYmUTo/edit?usp=drive_link

Sequências Argonautas Alinhadas.

- https://drive.google.com/file/d/1XXa1tJIC0ekVrI8_-mgiLWH7tyP9fpy7/view?usp=drive_link

Árvore Filogenética.

- https://drive.google.com/file/d/1mYY-z03gxVSPwrvBI12uYW-8rm1dEXKe/view?usp=drive_link (newick_export)
- https://drive.google.com/file/d/18weArJsEiCoB_tuwHySsurNWtylITQBO/view?usp=drive_link (mtsx)

Sequências genes vizinhos (WebFlags).

- https://drive.google.com/file/d/1eaEhLYQey4YQTLQXhiSjAsP-zjQHSgSv/view?usp=drive_link