

Caracterização estrutural e funcional das HSP70 no genoma do feijão comum (*Phaseolus vulgaris*): Uma análise *in silico*

Structural and functional characterization of HSP70 in the common bean (*Phaseolus vulgaris*) genome: An *in silico* analysis

Caracterización estructural y funcional de HSP70 en el genoma del frijol (*Phaseolous vulgaris*): Un análisis *in silico*

Recebido: 08/11/2023 | Revisado: 15/11/2023 | Aceitado: 16/11/2023 | Publicado: 19/11/2023

Maria Letícia Pereira Aquino

ORCID: <https://orcid.org/0000-0003-4892-802X>
Universidade Estadual de Alagoas, Brasil
E-mail: leticiapereiraaquino@hotmail.com

Dalton Ferreira Matos

ORCID: <https://orcid.org/0000-0002-6188-7536>
Universidade Estadual de Alagoas, Brasil
E-mail: daltonmatosgn@gmail.com

Genesio José do Amaral Ramos

ORCID: <https://orcid.org/0000-0001-8774-6953>
Universidade Estadual de Alagoas, Brasil
E-mail: genesioramos@alunos.uneal.edu.br

José Alfredo dos Santos Júnior

ORCID: <https://orcid.org/0000-0002-1324-1861>
Centro Universitário Cesmac, Brasil
E-mail: ajrsantus@hotmail.com

José Valdemilson dos Santos Silva

ORCID: <https://orcid.org/0000-0003-4423-8240>
Universidade Estadual de Alagoas, Brasil
E-mail: valdemylson2016@gmail.com

Delma Holanda de Almeida

ORCID: <https://orcid.org/0000-0002-7386-6046>
Universidade Estadual de Alagoas, Brasil
E-mail: delma.holanda@uneal.edu.br

Resumo

O feijão comum (*Phaseolus vulgaris* L.) é a leguminosa mais importante em todo o mundo, com produção global estimada em 26,8 milhões de toneladas métricas em 2016, além de ser cultivado por suas vagens verdes, bem como por suas sementes secas. As plantas, de uma maneira geral, respondem a estresses ambientais com uma série de mecanismos de adaptação fisiológica e molecular. As proteínas de choque térmico (HSPs) são proteínas protetoras contra uma ampla gama de condições de estresse bióticos e abióticos e são proteínas altamente conservadas que ocorrem em todos os organismos na natureza. Foi demonstrado pela primeira vez que as HSP70 eram induzidas por altas temperaturas, mas nos últimos anos elas foram consideradas como proteínas responsivas ao estresse geral, que são induzidas por muitas condições de estresse abiótico. Outros estudos verificaram que as HSP70s são altamente expressas sob condições de calor, frio e estresse por seca. Sendo assim, o aumento da tolerância devido à alta expressão de HSP70s foi demonstrado em algumas espécies de plantas. Diversas ferramentas computacionais têm sido utilizadas para uma ampla análise do genoma *in silico* têm sido utilizadas para identificação de novas proteínas e genes. Embora as tecnologias ômicas tenham se tornado significativamente populares, são poucos os estudos sobre identificação e caracterização da HSP70 em *Phaseolus vulgaris*. Portanto, o objetivo do presente estudo foi realizar uma caracterização estrutural e identificação das HSP70 em *Phaseolus vulgaris* utilizando métodos computacionais de genoma amplo, sendo elas: análise fenética, análise de potencial alergênico e modelagem tridimensional das sequências.

Palavras-chave: Caracterização estrutural; Feijão comum; Proteína de choque térmico.

Abstract

The common bean (*Phaseolus vulgaris* L.) is the most important legume worldwide, with global production estimated at 26.8 million metric tons in 2016, and is cultivated for its green pods as well as its dry seeds. Plants, in general, respond to environmental stresses with a series of physiological and molecular adaptation mechanisms. Heat shock proteins (HSPs) are protective proteins against a wide range of biotic and abiotic stress conditions and are highly

conserved proteins that occur in all organisms in nature. HSP70 was first shown to be induced by high temperatures, but in recent years they have been considered as general stress-responsive proteins, which are induced by many abiotic stress conditions. Other studies have found that HSP70s are highly expressed under conditions of heat, cold and drought stress. Therefore, increased tolerance due to high expression of HSP70s has been demonstrated in some plant species. Several computational tools have been used for a broad analysis of the genome in silico and have been used to identify new proteins and genes. Although omics technologies have become significantly popular, there are few studies on the identification and characterization of HSP70 in *Phaseolus vulgaris*. Therefore, the objective of the present study was to carry out a structural characterization and identification of HSP70 in *Phaseolus vulgaris* using genome-wide computational methods, namely: phenetic analysis, allergenic potential analysis and three-dimensional modeling of the sequences.

Keywords: Structural characterization; Common bean; Heat shock protein.

Resumen

El frijol común (*Phaseolus vulgaris* L.) es la leguminosa más importante a nivel mundial, con una producción global estimada en 26,8 millones de toneladas métricas en 2016, y se cultiva por sus vainas verdes y sus semillas secas. Las plantas, en general, responden al estrés ambiental con una serie de mecanismos de adaptación fisiológicos y moleculares. Las proteínas de choque térmico (HSP) son proteínas protectoras contra una amplia gama de condiciones de estrés biótico y abiótico y son proteínas altamente conservadas que se encuentran en todos los organismos de la naturaleza. Se demostró por primera vez que la HSP70 era inducida por altas temperaturas, pero en los últimos años se ha considerado como proteínas generales que responden al estrés, que son inducidas por muchas condiciones de estrés abiótico. Otros estudios han encontrado que las HSP70 se expresan altamente en condiciones de calor, frío y estrés por sequía. Por lo tanto, se ha demostrado en algunas especies de plantas una mayor tolerancia debido a la alta expresión de HSP70. Se han utilizado varias herramientas computacionales para un análisis amplio del genoma in silico y se han utilizado para identificar nuevas proteínas y genes. Aunque las tecnologías ómicas se han vuelto significativamente populares, existen pocos estudios sobre la identificación y caracterización de HSP70 en *Phaseolus vulgaris*. Por lo tanto, el objetivo del presente estudio fue llevar a cabo una caracterización estructural e identificación de HSP70 en *Phaseolus vulgaris* utilizando métodos computacionales de todo el genoma, a saber: análisis fenético, análisis de potencial alergénico y modelado tridimensional de las secuencias.

Palabras clave: Caracterización estructural; Frijol común; Proteína de choque térmico.

1. Introdução

O feijão comum (*Phaseolus vulgaris* L.) é indiscutivelmente a leguminosa de grãos comestíveis mais importante em todo o mundo, com produção global estimada em 26,8 milhões de toneladas métricas em 2016 (Assefa et al., 2019). O feijão comum é responsável por uma alta proporção da ingestão diária de proteínas em muitos países, particularmente na América Latina, África e partes da Ásia, além de ser uma leguminosa e hortaliça economicamente significativa no Canadá, EUA e Europa, sendo o consumo particularmente alto nos países africanos (Broughton et al. 2003; Buruchara et al. 2011). O feijão comum é muito rico em nutrientes, com proteínas e carboidratos complexos, vitaminas (por exemplo, A, C, folato), fibra alimentar e minerais biologicamente importantes, como Ca, Mg, K, Cu, Fe, Mg e Zn (Broughton et al., 2003; Blair, 2013), além de ajudar a melhorar a saúde do solo e do meio ambiente por meio da fixação simbiótica de nitrogênio (SNF). O feijão comum é cultivado no mundo todo por suas vagens verdes, bem como por suas sementes secas (Adikshita & Kansak 2017).

As plantas, de uma maneira geral, respondem a estresses ambientais com uma série de mecanismos de adaptação fisiológica e molecular (Timperio et al., 2008). As proteínas de choque térmico (HSPs) são proteínas protetoras contra uma ampla gama de condições de estresse bióticos e abióticos e são proteínas altamente conservadas que ocorrem em todos os organismos na natureza (Ajuha et al., 2010). As HSPs foram identificadas pela primeira vez na *Drosophila melanogaster* e até o momento, vários membros têm sido caracterizados em muitas espécies de plantas (Cashikar et al., 2005; Sarkar et al., 2009). Em estudos de Cashikar et al. (2005) e Wang et al. (2009), essas proteínas foram categorizadas em cinco classes de acordo com sua massa molecular, sendo elas sHSP (pequena proteína de choque térmico), HSP60, HSP70, HSP90 e HSP100.

As proteínas de choque térmico de 70 kDa (HSP70) são o grupo mais conservado de chaperonas moleculares de HSP, que são compostas por dois domínios principais (domínios N-terminal ATPase de 45 kDa e 25 kDa de ligação ao peptídeo) e dois subdomínios (15-kDa kDa β -sandwich subdomínios α -helicoidais C-terminal) (Zhu et al., 1996; Masand & Yadav, 2016). Uma fenda de ligação peptídica que é responsável pela interação com o fragmento peptídico hidrofóbico de proteínas

desdobradas de forma controlada por ATP é encontrada no subdomínio β -sandwich das proteínas HSP70 (Richter et al., 2010; Masand & Yadav, 2016).

Foi demonstrado pela primeira vez que as HSP70 eram induzidas por altas temperaturas, mas nos últimos anos elas foram consideradas como proteínas responsivas ao estresse geral, que são induzidas por muitas condições de estresse abiótico (Zhou et al., 2013). Membros da família da HSP70 induzidos por estresse são necessários para dobrar ou degradar proteínas mal dobradas ou danificadas em espécies vegetais sob condições de estresse (Zhou et al., 2013; Frydman, 2001). Nos últimos anos, muitos genes da HSP70 altamente conservados foram identificados e caracterizados em várias espécies de plantas e esses genes foram classificados em subclasses de acordo com sua localização subcelular e homologia de sequência. Essas subclasses são compostas por HSP70 citosólicas, HSP70s plastidiais, HSP70s mitocondriais, HSP70s do retículo endoplasmático e subfamília de HSP91 (Li & Srivastava, 2004; Zhou et al., 2013; Daugaard et al., 2007). Outros estudos verificaram que as HSP70s são altamente expressas sob condições de calor, frio e estresse por seca. Sendo assim, o aumento da tolerância devido à alta expressão de HSP70s foi demonstrado em algumas espécies de plantas (Daugaard et al., 2007; Alvim et al., 2001).

Nos últimos anos, diversas ferramentas computacionais para uma ampla análise do genoma *in silico* têm sido utilizadas para identificação de novas proteínas e genes. Embora as tecnologias ômicas tenham se tornado significativamente populares, são poucos os estudos sobre identificação e caracterização da HSP70 em *Phaseolus vulgaris*. Portanto, o objetivo deste estudo foi realizar uma caracterização estrutural e identificação das HSP70s em *Phaseolus vulgaris*, conduzida a partir da utilização de métodos computacionais de genoma amplo, sendo eles: análise fenética, análise de potencial alergênico e modelagem tridimensional das sequências.

2. Metodologia

2.1 Identificação e Caracterização de HSP70

O genoma do feijão comum foi obtido por meio do GenBank, Banco de Dados situado no *National Center for Biotechnology Information* – NCBI (<https://www.ncbi.nlm.nih.gov>). Tendo como código de acesso: 3885. Para a identificação e caracterização de membros da família das HSP70 foi utilizada a ferramenta tBLASTn (<https://blast.ncbi.nlm.nih.gov>), por intermédio da ferramenta Bioedit (<https://bioedit.software.informer.com>). Para isso utilizou-se sequências “sonda” de HSP70 previamente caracterizadas em outras espécies contendo o domínio “*Heat Shock Proteins*” por alinhamentos junto ao banco de dados do NCBI e UNIPROT, da espécie *Phaseolus vulgaris*.

As sequências candidatas foram traduzidas em sequências de proteínas por intermédio da ferramenta TransDecoder (<https://transdecoder.github.io/>). A busca por domínios conservados foi realizada com o auxílio da ferramenta CD-search (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) e InterProScan (<https://www.ebi.ac.uk/interpro/search/sequence/>) e os novos motivos de PvHSP70s foram obtidos por intermédio da ferramenta MEME 5.3.3 (<http://meme-suite.org/tools/meme>).

Das sequências proteicas com domínio conservado completo (*Heat Shock Proteins*), foram avaliadas: o ponto isoelétrico e a massa molecular, via ferramenta ProtParam (<https://web.expasy.org/protparam/>); a localização subcelular, pela ferramenta ProtComp 9.0 (<http://linux1.softberry.com/berry.phtml>); a presença e o número de segmentos transmembrana, além do peptídeo sinal, via SignalP 5.0 (<http://www.cbs.dtu.dk/services/SignalP/>).

2.2 Análise fenética

As PvHSP70 foram comparadas com HSP70 previamente descritas para outras espécies disponíveis nos bancos de dados do NCBI (<https://www.ncbi.nlm.nih.gov/>) e UniProt (<https://www.uniprot.org/>), gerando uma árvore fenética (Figura 3), a partir do método de Neighbor-Joining (NJ) e bootstrap de 1000 repetições, através da ferramenta MEGA v.7 (Kumar et al., 2016) e visualizados por meio da ferramenta iTol (Letunic; Bork, 2016).

2.3 Potencial alergênico e atividade antimicrobiana e antiviral das HSP70s

A análise de potencial alergênico foi utilizada o programa AllercatPro 2.0 (<https://allercatpro.bii.a-star.edu.sg/>), no qual a abordagem irá se basear de acordo com as orientações de avaliação de alergenicidade para alimententos vegetais, recomendados pela FAO/OMS (2001), sendo que o AllercatPro fornecerá três avaliações: (a) evidência alta; (b) evidência fraca e (c) nenhuma evidência. Ressaltando-se que o programa possui 84% de precisão em comparação com outros métodos atuais que acabam variando entre 51 a 73% (Maurer-Stroh et al., 2019).

A análise de atividade antimicrobiana das sequências de PvHSP70 foram previstas no servidor Collection of Anti-Microbial Peptides - CAMP (<http://www.camp.bicnirrh.res.in/>). A interface de previsão no CAMP permite que os usuários prevejam a atividade antimicrobiana das sequências, sendo selecionados quatro algoritmos (a) *Support Vector Machine* - SVM, (b) *Random Forest* (RF), (c) *Artificial Neural Network* - ANN e (d) *Discriminant Analysis* -DA.

Os quatro algoritmos utilizados, utilizam uma pontuação probabilística variando de 0 a 1 para a previsão dos peptídeos com atividade antimicrobiana (AMP) enquanto o algoritmo DA fornecem apenas as previsões AMP ou de peptídeos que não possuem atividade antimicrobiana (NAMP) (Subramanian et al., 2021).

A análise de potencial antiviral foi realizada através do servidor Meta-Iavp, no qual o servidor pode servir como um recurso computacional para prever novos peptídeos antivirais (Schaduangrat et al., 2019).

2.4 Modelagem tridimensional das sequências

A modelagem comparativa das HSP70 foi realizada utilizando *templates* do Protein Data Bank (PDB) com base na maior identidade e cobertura. No qual essa abordagem envolve 4 etapas principais: (a) atribuição de dobras, que irão reconhecer a semelhança entre o alvo e a estrutura do modelo resolvido; (b) alinhamento da sequência alvo ao modelo; (c) geração de um modelo baseado no alinhamento com o *template* escolhido; e (d) análise de erros considerando o modelo gerado (Webb & Sali, 2016).

Os modelos foram gerados via Modeller 10.1, com 500 modelos teóricos para cada sequência de HSP70. Os modelos teóricos serão visualizados através da ferramenta PyMOL. A ferramenta QmeanDisCo será utilizada para acessar a estimativa da qualidade global dos modelos teóricos (Studer et al. 2021).

Em seguida, foi analisado o potencial eletrostático de superfície, sendo avaliados em relação ao campo de força, como cargas atômicas e raios, por meio do software PDB2PQR Server (http://nbc-222.ucsd.edu/pdb2pqr_2.0.0/). Posteriormente, os resultados foram dispostos no APBS Tools (Jurrus et al., 2018) para obtenção do potencial eletrostático dos modelos. Os potenciais eletrostáticos foram visualizados na superfície molecular da proteína em relação a suas cargas eletronegativas, eletropositivas e neutras, sendo visualizados através do programa gráfico PyMOL (Delano, 2002).

3. Resultados e Discussão

3.1 Identificação e Caracterização das PvHSP70s

Devido as propriedades voltadas para as vias de defesa e desenvolvimento das HSP70, foi necessário a realizar uma análise abrangente desse grupo de proteínas na espécie *P. vulgaris*. Após a remoção das entradas redundantes, 15 genes com domínio *HSP70* foram identificados do banco de dados do genoma do feijão comum. No entanto, alguns genes não possuíam relação com as sequências de interesse, por não possuir o domínio *Heat Shock Proteins*, sendo assim eliminadas. Os genes pertencentes as *HSP70* foram classificadas em *PvHSP70-1* a *PvHSP70-15*, com base na ordem em que foram encontradas no banco de dados do respectivo vegetal.

As sequências polipeptídicas de HSP70 em *Phaseolus vulgaris* (PvHSP70) mostraram-se ser presentes na região citoplasmática, apresentando ponto isoelétrico (PI) entre de 5.08 a 8.19. A massa molecular (MM) variou entre de 47892.78

kDa (PvHSP70-14) a 95630.27 kDa (PvHSP70-3) e o domínio transmembrana apresentou entre 0 e 3 domínios. No entanto apenas as PvHSP70-13 e PvHSP70-15 apresentaram os domínios transmembrana.

As HSP70s, são consideradas proteínas altamente conservadas em várias espécies vegetais, contendo dois domínios funcionais, no qual um desses domínios, possui a função de ligação ao nucleotídeo, na região N-terminal e outros domínios de ligação ao substrato na região C-Terminal. Essas proteínas formam um complexo denominado classificado em multichaperona, em que são conectados a uma proteína chamada de Hop, sendo que a conexão e a interação entres os mecanismos de chaperone são de extrema importância para a viabilidade celular. Além disso, essas proteínas podem ser classificar conforme a sua massa molecular, podendo variar o seu tamanho entre as pequenas de 18 KDa e as consideradas grandes podendo ultrapassar os 70 KDa, assim como foi demonstrado em nosso estudo (Wegele et al., 2004; Tiroli-Cepeda et al., 2014; Liu et al., 2020).

Os resultados de Wang et al. (2022), possuem similaridades com os dados desse estudo (Tabela 1), além do mais em relação ao seu número de aminoácidos (aa), esse grupo de proteínas são consideradas polípticas variando entre 45 a 657, para a espécie *Dendrobium officinale*, no entanto desse ter em mente que esse número pode variar entre espécies como *Ziziphus jujuba* com 481 a 917 (Panzade et al., 2020), 143 a 886 na espécie *Gossypium hirsutum* (Zhou et al., 2023) e entre outras espécies vegetais.

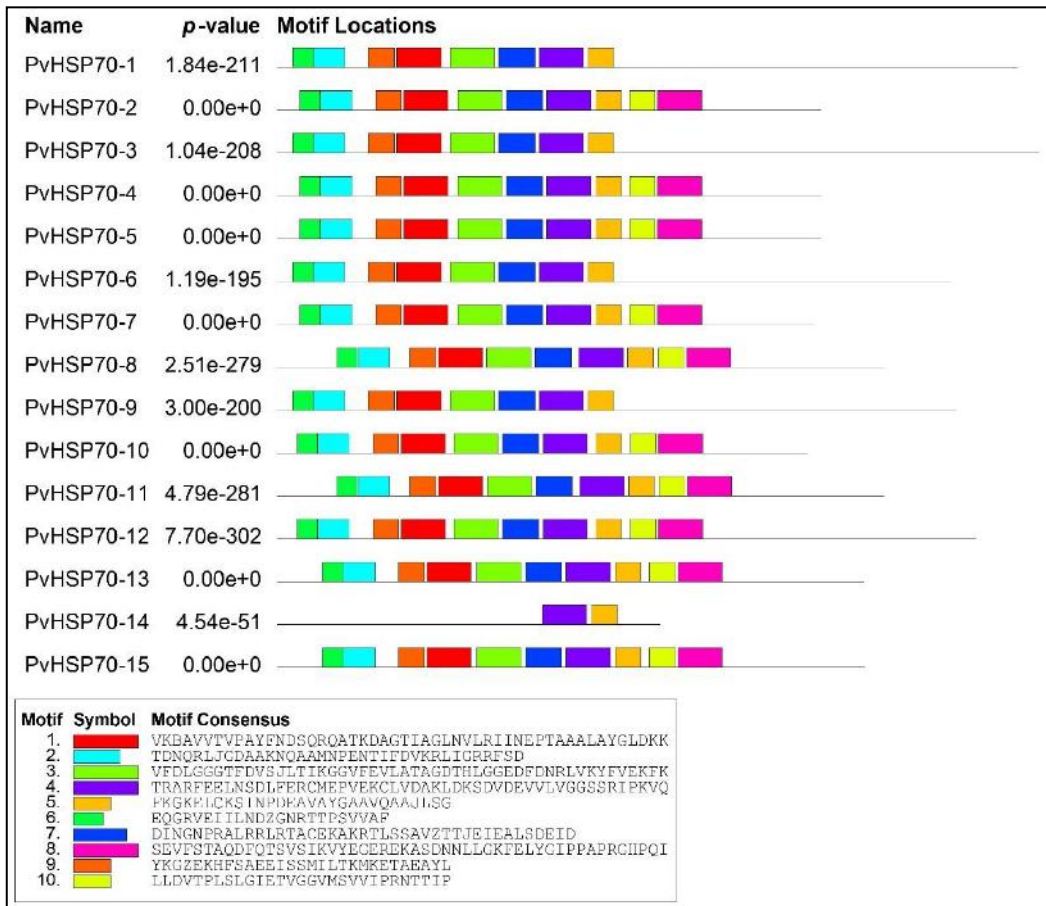
Tabela 1 - Informações detalhadas das HSP70 de *Phaseolus vulgaris* identificadas utilizando a ferramenta tBLASTn por intermédio da ferramenta Bioedit.

Nome dos Genes	ID do Gene	Características da proteína						Localização subcelular
		Localização nos Cromossomos	Nº de aa	MM (kDa)	PI	GRAVY	TM	
PvHSP70-1	Phvul.010G150800	Cromossomo 10	841	94	5.46	-0.514	0	Citoplasma
PvHSP70-2	Phvul.009G076700	Cromossomo 9	617	68,71	5.59	-0.281	0	Citoplasma
PvHSP70-3	Phvul.005G171400	Cromossomo 5	865	95,63	5.09	-0.488	0	Citoplasma
PvHSP70-4	Phvul.008G011900	Cromossomo 8	617	68,99	6.28	-0.307	0	Citoplasma
PvHSP70-5	Phvul.008G011400	Cromossomo 8	617	69,26	7.05	-0.325	0	Citoplasma
PvHSP70-6	Phvul.008G267400	Cromossomo 8	765	84,93	5.36	-0.304	0	Citoplasma
PvHSP70-7	Phvul.008G011600	Cromossomo 8	609	67,74	5.86	-0.252	0	Citoplasma
PvHSP70-8	Phvul.006G219500	Cromossomo 6	689	73,72	5.20	-0.284	0	Citoplasma
PvHSP70-9	Phvul.006G042400	Cromossomo 6	771	86,37	5.37	-0.350	0	Citoplasma
PvHSP70-10	Phvul.006G170200	Cromossomo 6	602	67,09	5.96	-0.317	0	Citoplasma
PvHSP70-11	Phvul.006G219600	Cromossomo 6	689	73,66	5.24	-0.299	0	Citoplasma
PvHSP70-12	Phvul.006G170300	Cromossomo 6	793	88,07	8.19	-0.306	0	Citoplasma
PvHSP70-13	Phvul.002G328904	Cromossomo 2	666	73,53	5.11	-0.440	3	Citoplasma
PvHSP70-14	Phvul.002G046600	Cromossomo 2	435	47,89	5.83	0.087	0	Citoplasma
PvHSP70-15	Phvul.002G301304	Cromossomo 2	666	73,41	5.08	-0.423	3	Citoplasma

Fonte: Autores (2023).

Para analisar a distribuição dos motivos conservados, 10 motivos conservados foram revelados nos genes *PvHSP70s* (Figura 1).

Figura 1 - Análise dos motivos conservados das HSP70s em *Phaseolus vulgaris*, geradas a partir da ferramenta MEME.



Fonte: Autores (2023).

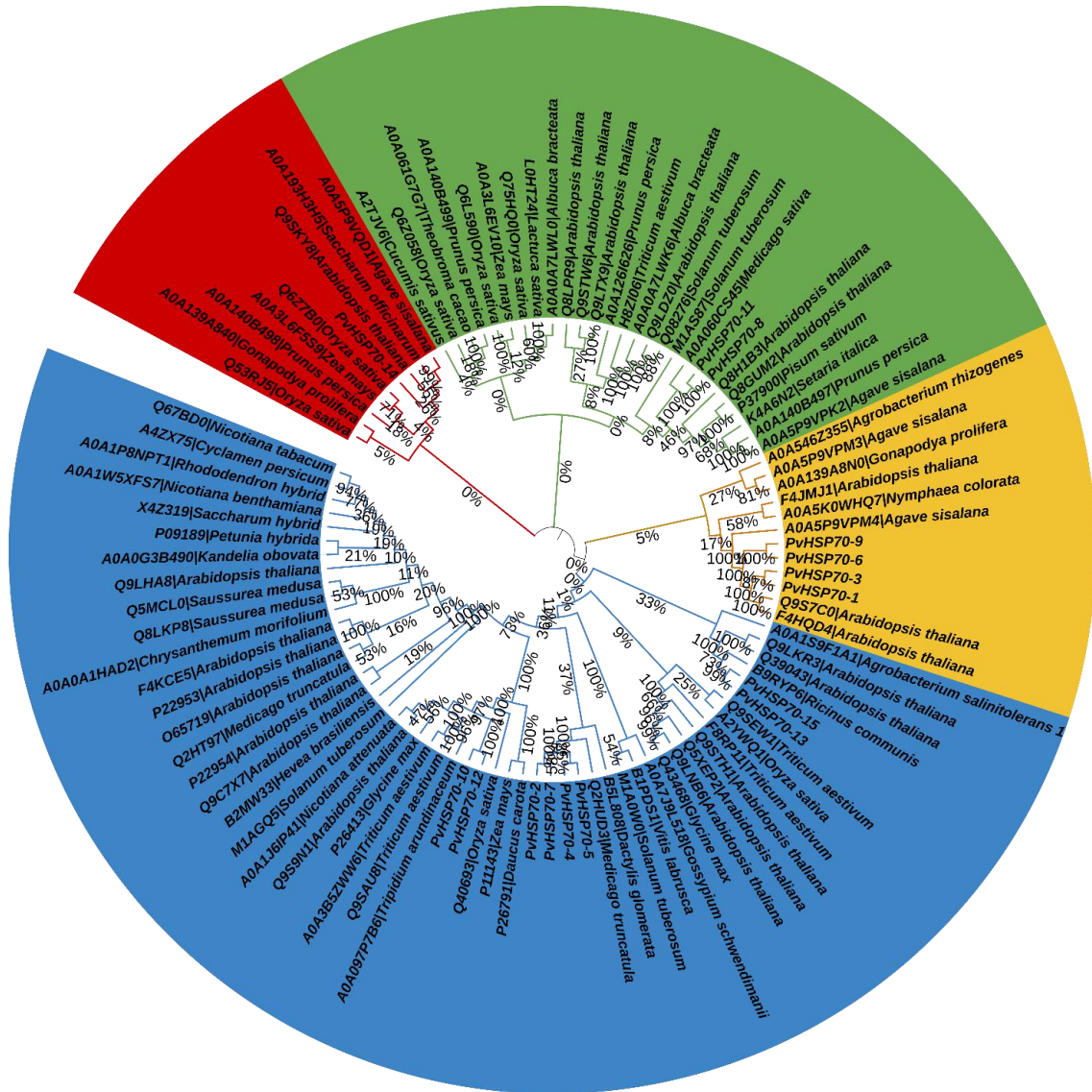
Como pode-se observar na figura, os números dos motivos variaram de 2 a 10, e o comprimento do motivo variou de 23 (motivo 6) a 50 (motivo 1, 3, 4 e 8) aminoácidos, representados na Tabela presente na Figura 1. 10 genes *PvHsp70* tinham 10 motivos, 4 genes *PvHsp70* tinham 4 motivos, e apenas 1 gene, continha apenas 1 motivo. Tendo em vista que o resultado dos motivos conservados possui similaridade com alguns estudos, como é o caso do trabalho de Liu et al. (2022), onde o número de motivos conservados e quantidade de aminoácidos em cada motivo consenso, foi similar ao nosso estudo, no entanto os pesquisadores verificaram os clados que continham um maior número de motivos, possuindo essa particularidade que não foi verificado em nosso resultado.

Todavia, é por meio da análise de motivos e por intermédio da ferramenta MEME, que se verifica a função biológica de um determinado gene (Boubakri et al., 2022). Como é o caso dos motivos em nosso estudo que possui uma gama de funções, como: Acompanhante de dobramento de proteína dependente de ATP (GO:0140662); ligação de ATP (GO:0005524); possui as funções de acompanhante de dobramento de proteína dependente de ATP (GO:0140662); ligação de proteína de choque térmico (GO:0031072); atividade de hidrólise de ATP (GO:0016887) e entre outras funções biológicas.

3.2 Análise fenética

Para explorar a relação evolutiva da família HSP70s, as sequências de proteínas de HSP70s de 100 espécies de plantas diferentes, incluindo a do feijão comum, *Medicago sativa* e *Prunus persica*, foram utilizadas para construir a árvore fenética (Figura 2).

Figura 2 - Fenograma das HSPs70 de *P. vulgaris*, comparativamente com sequências do banco de dados Uniprot, representadas por seus códigos de acesso, seguido da espécie vegetal, construídas pelo método NJ.



Fonte: Autores (2023).

Como apresentado na Figura 2, as HSP70s foram divididas em 4 grupos parafiléticos conforme a sua relação fenética, classificados do Grupo 1 (em azul) ao Grupo 4 (em vermelho). As PvHSP70s mostraram-se presentes em todos os grupos.

Destes 4 grupos apresentados, o grupo 1 apresentou o maior número de PvHSP70, contendo 8 membros dessas proteínas, apresentando valores de bootstrap variando entre 73% a 100%, no entanto o menor grupamento foi o grupo 4 contendo apenas um membro para as PvHSP70, com valor de bootstrap de 26%, para a PvHSP70-14. As outras subfamílias continham 4 PvHSP70 para o grupo 2, no qual os valores de bootstrap variaram entre 17% a 100% e o grupo 3, continha

apenas 2 membros dessa família, possuindo valores de bootstrap altos, constando valor de 100% para as PvHSP70-8 e PvHSP70-11.

Ainda em relação a Figura 2, para cada grupo apresentado há correspondentes em relação as famílias botânicas, sendo: o primeiro grupo corresponde às famílias *Solanaceae*, *Primulaceae*, *Ericaceae*, *Poaceae*, *Rhizophoraceae*, *Brassicaceae*, *Asteraceae*, *Euphorbiaceae*, *Poaceae* e *Leguminosae*. O segundo grupo corresponde às famílias *Brassicaceae*, *Leguminosae*, *Asparagaceae*, *Crassulaceae*. O terceiro grupo corresponde às famílias *Crassulaceae*, *Rosaceae*, *Poaceae*, *Brassicaceae*, *Leguminosae*, *Asparagaceae*, *Malvaceae*, *Cucurbitaceae*. O último e quarto grupamento correspondem às famílias *Asparagaceae*,

Deve-se levar em consideração que os resultados em nosso estudo correspondem a diversos estudos da literatura científica, no qual a família o grupo de proteínas HSP70 corresponde a 4 grupos parafiléticos, como é o caso do estudo de Panzade et al. (2020) e Zhou et al. (2023), no qual o maior grupamento é o grupo 1, assim como foi demonstrado neste estudo, além do mais as proteínas homologas de cada espécie pode possuir um gene duplicado. Além do mais a espécie do milho (*Zea mays*) apresenta alta homologia com algumas espécies, como é o caso do arroz, isso se dá por ser um grão e pertencer a mesma família botânica (Jiang et al., 2021), no entanto nesse estudo as PvHSP70, possui uma alta afinidade com algumas espécies, como é o caso da espécie *Glycine max*, isso se dá pelo fato de pertencerem ao mesmo grupo botânico, no entanto algumas outras espécies, possuem similaridade, possivelmente por serem grãos, como arroz, milho, feijão e trigo.

3.3 Potencial alergênico e atividade antimicrobiana e antiviral das HSP70s

Diversas proteínas vegetais são incríveis, como é o caso da soja, feijão, trigo, no qual contem determinados aminoácidos essenciais a alimentação humana (Marcus, 2013), tendo em vista que uma proteína de qualidade com sabor doce, textura e disponibilidade de proteínas vegetais são consideradas desafios para a indústria de alimentos conseguir produzir determinados produtos de qualidade utilizando essas proteínas (Kumar et al., 2022).

No entanto alguns alimentos conseguem induzir reação alérgicas a seres humanos, sendo assim considerados um grande problema na saúde pública, chegando a proporções quase epidêmicas em diversas regiões do mundo, comprometendo a qualidade de vidas, podendo ser influenciada pela genética do indivíduo, o ambiente, como também as interações genoma-ambiente (Sicherer & Sampson, 2018).

Além do mais existem outros fatores que irá determinar uma resposta alérgica a uma determinada proteína alimentar envolvem diferentes fatores, como a quantidade (dose) ingerida, a via de exposição, propriedades do alérgeno, matriz alimentar, modificação de alérgenos pelo processamento de alimentos e entre outros fatores. Tendo em vista que diversas famílias possuem entre um ou dois alérgenos considerados potentes, enquanto a maioria das proteínas não são alergênicas (Costa et al., 2022).

Diante disso, através da análise *in silico* para o potencial alergênico das PvHSP70s foi utilizado o servidor AllerCatPro 2.0. Demonstrou-se que essas proteínas do feijão comum, possuem homologia a cinco proteínas alergênicas pertencentes às espécies *Malassezia sympodialis* (Mala s 10), *Tyrophagus putrescentiae* (Tyr p 28), *Aedes aegypti* (Aed a 8), *Corylus avellana* (Cor a 10) e *Dermatophagoides pteronyssinus* (Der p 28). Das 15 PvHSP70 analisadas, 10 apresentaram evidência alta para atividade alergênica (PvHSP70-2, PvHSP70-4, PvHSP70-5, PvHSP70-7, PvHSP70-8, PvHSP70-10, PvHSP70-11, PvHSP70-12, PvHSP70-13 e PvHSP70-15), com porcentagem (%) de identidade, do epítipo na estrutura 3D, variando entre 94,1% a 100%.

As *Heat Shock Protein 70*, possui uma relação com doenças alérgicas (Guo et al., 2021), na literatura científica foram identificados alérgenos em diversas leguminosas, como é o caso da *Oryza sativa*, possuindo o alérgeno Bet v I, no

entanto, ainda existem algumas lacunas dos alérgenos presentes em outras espécies como é o caso do feijão comum, como é o caso das HSP70 (Tabassum et al., 2020).

As proteínas classificadas em *Heat Shock Protein 70* (HSP70), são consideradas proteínas com forte atividade alergênica, podendo induzir asma alérgica em modelos de animais (Guo et al., 2021), no qual vários membros das HSP70, são alérgenos conhecidos em ácaros domésticos (Aki et al., 1994), mosquitos picadores, mosca negra, e barata (Chuang et al., 2010), possuindo alérgenos em potenciais, destacando o Touro p 28, cupin_1 (Hao et al., 2022), Cor a 10 (Hirano et al., 2016), Pru av 1 (Kesici et al., 2020) e entre outros alérgenos como demonstrado nesse estudo.

Isso acontece devido que todos os alérgenos consegue compartilha um grau de identidade em suas sequencias de aminoácidos, podendo ser expressados em outras espécies (Barre et al., 2023), no qual a avaliação de suas alergenicidade em alimentos, pode ser realizado por meio de métodos preditivos, como é o caso de modelos

Ademais, todos os alérgenos compartilham um determinado grau de identidade em suas sequências de aminoácidos (Barre et al., 2023). A sua avaliação do potencial alergênico em um determinado alimento pode ser realizada através de alguns métodos preditivos, como é o caso dos modelos *in vivo*, *in vitro* e *in silico* (Verhoeckx et al., 2016).

Em especial, através do método *in silico*, é possível predizer um determinado alérgeno de uma proteína, como é o caso das HSP70s em diversos alimentos como Morangueiro 1 (Kesici et al., 2020), Milho, Soja (Chen et al., 2016), Avelã (Costa et al., 2016) e entre outras espécies, como é o caso da espécie *Phaseolus vulgaris*, que demonstrou possuir homologia para cinco alérgenos em potenciais, no qual acabam implicando na mesma atividade funcional em outras espécies (Anisimova et al., 2022). Diante disso, no nosso estudo foi possível retratar a presença de alérgenos de diferentes espécies expressos no feijão-fava, com informações detalhadas na Tabela 2.

Tabela 2 - Análise do potencial alergênico, gerada pela ferramenta AllercatPro 2.0.

Nome	Espécie	Proteína alergênica	Atividade alergênica	% da identidade do epítipo na estrutura 3D	Evidência da atividade alergênica
PvHSP70-1	<i>Malassezia sympodialis</i>	Mala s 10	<u>AG(2)</u>	76.9	Evidência fraca
PvHSP70-2	<i>Tyrophagus putrescentiae</i>	Tyr p 28	-	100.0	Evidência forte
PvHSP70-3	<i>Malassezia sympodialis</i>	Mala s 10	<u>AG(2)</u>	78.9	Evidência fraca
PvHSP70-4	<i>Tyrophagus putrescentiae</i>	Tyr p 28	-	100.0	Evidência forte
PvHSP70-5	<i>Tyrophagus putrescentiae</i>	Tyr p 28	-	100.0	Evidência forte
PvHSP70-6	<i>Malassezia sympodialis</i>	Mala s 10	<u>AG(2)</u>	64.7	Evidência fraca
PvHSP70-7	<i>Tyrophagus putrescentiae</i>	Tyr p 28	-	100.0	Evidência forte
PvHSP70-8	<i>Aedes aegypti</i>	Aed a 8	-	94.1	Evidência forte
PvHSP70-9	<i>Malassezia sympodialis</i>	Mala s 10	<u>AG(2)</u>	69.2	Evidência fraca
PvHSP70-10	<i>Tyrophagus putrescentiae</i>	Tyr p 28	-	100.0	Evidência forte
PvHSP70-11	<i>Aedes aegypti</i>	Aed a 8	-	94.1	Evidência forte
PvHSP70-12	<i>Tyrophagus putrescentiae</i>	Tyr p 28	-	100.0	Evidência forte
PvHSP70-13	<i>Corylus avellana</i>	Cor a 10	-	100.0	Evidência forte
PvHSP70-14	<i>Dermatophagoides pteronyssinus</i>	Der p 28	-	46.2	Evidência fraca
PvHSP70-15	<i>Corylus avellana</i>	Cor a 10	-	100.0	Evidência forte

Fonte: Autores (2023).

Em relação à atividade antimicrobiana e antiviral das PvHSP70, foram selecionadas as 15 sequências contendo o domínio *heat shock protein* e analisadas posteriormente usando a ferramenta CAMPR₃ e Meta-iAVP, com o intuito de prever seu potencial antimicrobiano e antiviral. Todas as HSP70s em *P. vulgaris* foram analisadas considerando três preditores antimicrobianos, e apenas dois para antiviral, estando detalhado na Tabela 3.

Tabela 3 - Atividade antimicrobiana e antiviral das HSP70s em *Phaseolus vulgaris*, geradas por intermédio das ferramentas CAMPR₃ e Meta-iAVP.

Nome	Probabilidade da atividade antimicrobiana (AMP)			Probabilidade da atividade antiviral (AVP)		
	SVM	RF	ANN	Prediction	AVP	Non-AVP
PVHSP70-1	1.00	0.82	NAMP	AVP	0.954	0.046
PVHSP70-2	1.00	0.82	NAMP	Non-AVP	0.044	0.956
PVHSP70-3	1.00	0.81	NAMP	Non-AVP	0.000	1.000
PVHSP70-4	1.00	0.84	NAMP	AVP	1.00	0.000
PVHSP70-5	1.00	0.85	NAMP	Non-AVP	0.324	0.676
PVHSP70-6	1.00	0.82	AMP	AVP	0.774	0.226
PVHSP70-7	1.00	0.83	NAMP	AVP	0.990	0.010
PVHSP70-8	1.00	0.81	AMP	Non-AVP	0.000	1.000
PVHSP70-9	1.00	0.82	NAMP	Non-AVP	0.000	1.000
PVHSP70-10	1.00	0.85	NAMP	AVP	0.998	0.002
PVHSP70-11	1.00	0.81	AMP	Non-AVP	0.000	1.000
PVHSP70-12	1.00	0.85	AMP	AVP	0.828	0.172
PVHSP70-13	1.00	0.83	AMP	AVP	0.998	0.002
PVHSP70-14	1.00	0.80	NAMP	AVP	0.818	0.182
PVHSP70-15	1.00	0.83	AMP	Non-AVP	0.316	0.684

Fonte: Autores (2023).

As sequencias de PvHSP70s analisadas foi utilizada três parâmetros durante a análise, para o qual foi obtido as seguintes predições: *Support Vector Machine* (SVM), com valor para todas as PvHSP70 de 1,00. O preditor classificado em *Random Forest Classifier* (RF), com valores entre 0,80 a 0,85, *Artificial Neural Network* (ANN), apresentaram classificação AMP, para apenas 5 das PvHSP70, onde 10 das PvHSP70 apresentaram classificação NAMP que são sequências que não apresentam atividade antimicrobiana.

Em relação à análise da atividade antiviral em HSP70s, neste estudo, o software Meta-iAVP previu a atividade dessas proteínas, no qual o resultado da previsão antiviral sugeriu que 8 PvHSP70s possuem atividade antiviral (AVP), enquanto 7 não possuem atividade antiviral (Non-AVP). Das HSP70s que possuem atividade antiviral, 3 apresentam atividade antiviral promissoras, por possuir taxas mais altas, em comparação das outras PvHSP70s. Das que apresentam atividade antiviral promissoras, estão: PVHSP70-7, PVHSP70-10 e PVHSP70-13 apresentam as maiores pontuações para essa atividade, com informações detalhadas na Tabela 3.

As plantas conseguir sintetizar uma grande diversidade de proteínas que podem estar envolvidas nos mecanismos de defesa contra vários patógenos, como bactérias, vírus, fungos e entre outros corpos estranhos sob diferentes condições ambientais. Tendo em vista que a maioria dessas proteínas são enzimas, lectinas, ciclotídeos, proteínas inativadoras de ribossomos e entre outras proteínas como é o caso das proteínas relacionado a patogenicidade (Jain et al., 2022).

Uma vez que as plantas estão ligadas diretamente ao solo, elas devem desenvolver meios de defesa contra diversos agentes bióticos, por conta disso elas desenvolveram uma rede complexa de sinalização para obterem respostas celulares

apropriadas. Assim como os mamíferos, sua atividade biológica depende de interações específicas entre moléculas orgânicas, íons e receptores para se comunicarem e responderem a esses sinais de estresse, contra microrganismos ou vírus (Salas et al., 2015; Musidlak et al., 2017; Guo et al., 2019).

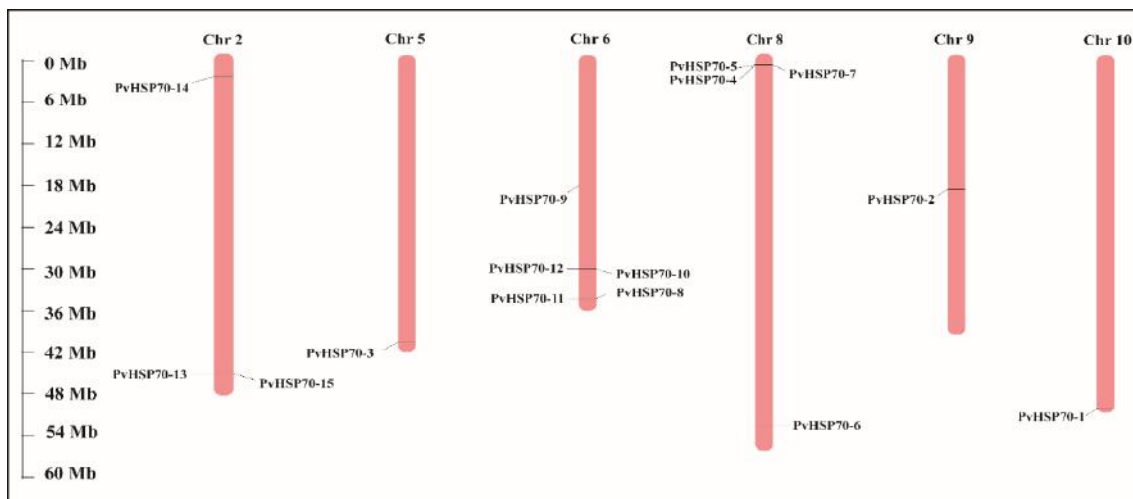
Isso ocorre devido as suas propriedades de defesa contra esses corpos estranhos, além do mais essas proteínas de defesa, possuem o potencial de serem aplicadas em diversos campos como medicina, na construção de drogas terapêuticas, produtos de consumo humano, além de fornecer novos insights sobre técnicas em pesquisa avançada para o desenvolvimento de ferramentas voltadas para proteínas antivirais e antimicrobianas (Naikoo et al., 2021).

Tendo em vista que as plantas possuem uma complexa interação de fatores bioquímicos, genéticos e celulares, no qual o genoma vegetal conseguem condicionar múltiplas proteínas visando o combate desses corpos estranhos, no qual as proteínas foram projetadas para serem molecular versáteis, desempenhando múltiplas funções, atuando como defesa, fazendo parte do sistema imunológico do organismo, como parte de uma rede enzimática necessária durante o metabolismo celular (SALAS et al., 2015; Musidlak et al., 2017). Devido aos seus efeitos inibitórios, as proteínas vegetais de defesa atuam como um promissor agente antimicrobiano e antiviral para humanos e outros animais (Jain et al., 2022).

3.4 Distribuição cromossômica

Em seguida, foi realizado uma análise de mapeamento cromossômico das *PvHSP70* usando a sequência do genoma do feijão comum como referência. As 15 sequências de *PvHSP70* foram mapeadas em cromossomos utilizando a ferramenta MG2C.

Figura 3 - Distribuição cromossômica de genes *HSP70* em *Phaseolus vulgaris*, gerada na ferramenta MG2C.



Fonte: Autores (2023).

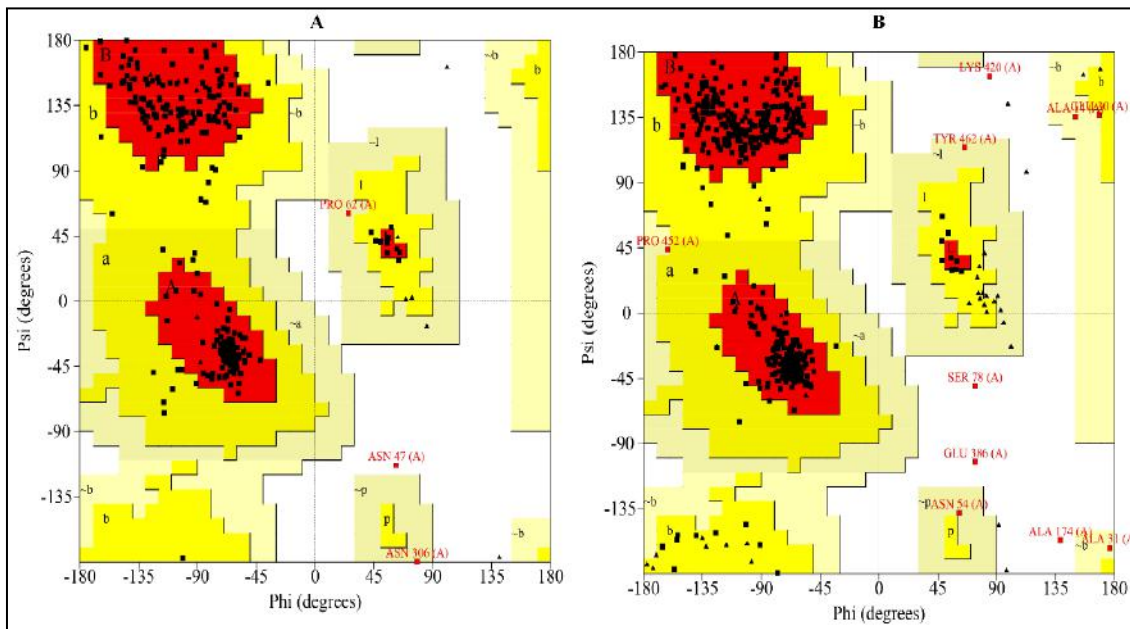
Como pode-se observar na Figura 3, as 15 *PvHSP70s* foram encontradas distribuídos nos 10 cromossomos, exceto nos cromossomos 1, 4, 7 e 11 que não possuíam nenhum gene para o referido gene analisado. Além do exposto, os 15 genes foram distribuídos em 6 cromossomos com um máximo de 3 genes no cromossomo 2; 1 gene nos cromossomos 5, 9 e 10; 4 genes no cromossomo 8; e 5 genes no cromossomo 6. Curiosamente, muitos *PvHSP70* foram agrupados em cromossomos da feijão, com distâncias físicas muito curtas, como é o caso das *PvHSP70-4*, *PvHSP70-5* e *PvHSP70-7*, estando agrupadas no cromossomo 8. De acordo com a literatura científica a distribuição de genes HSP70s nos cromossomos se comportam de forma heterógena, como é o caso da espécie *Gossypium hirsutum* foi encontrado 32 genes HSP70s para a referida espécie, no qual foi verificado

que as HSP70s foram distribuídas de forma conservadora nos cromossomos Chr5, Chr6, Chr9, Chr10, Chr11 e Chr13 (Zhou et al., 2023), Já na espécie *Pyropia yezoensis* foi identificado 15 genes distribuídos de forma heterogênea em três cromossomos, no qual o cromossomo 1 possui a maior quantidade desses genes na referida espécie (Yu et al., 2021).

3.5 Modelagem tridimensional das sequências

Através do gráfico do gráfico de Ramachandran, foram selecionados os possíveis modelos que apresentaram maior porcentagem referentes aos resíduos de aminoácidos presentes na região mais favorável, como também na região permitida, indicando assim as regiões dos resíduos localizados em regiões energeticamente favoráveis (LIU et al., 2019), sendo calculados os ângulos ϕ (Phi) e ψ (Psi), para as proteínas selecionadas. As duas proteínas escolhidas foram as PvHSP70-12 e PvHSP70-13, possuindo características estereoquímicas e geometria 3D de boa qualidade.

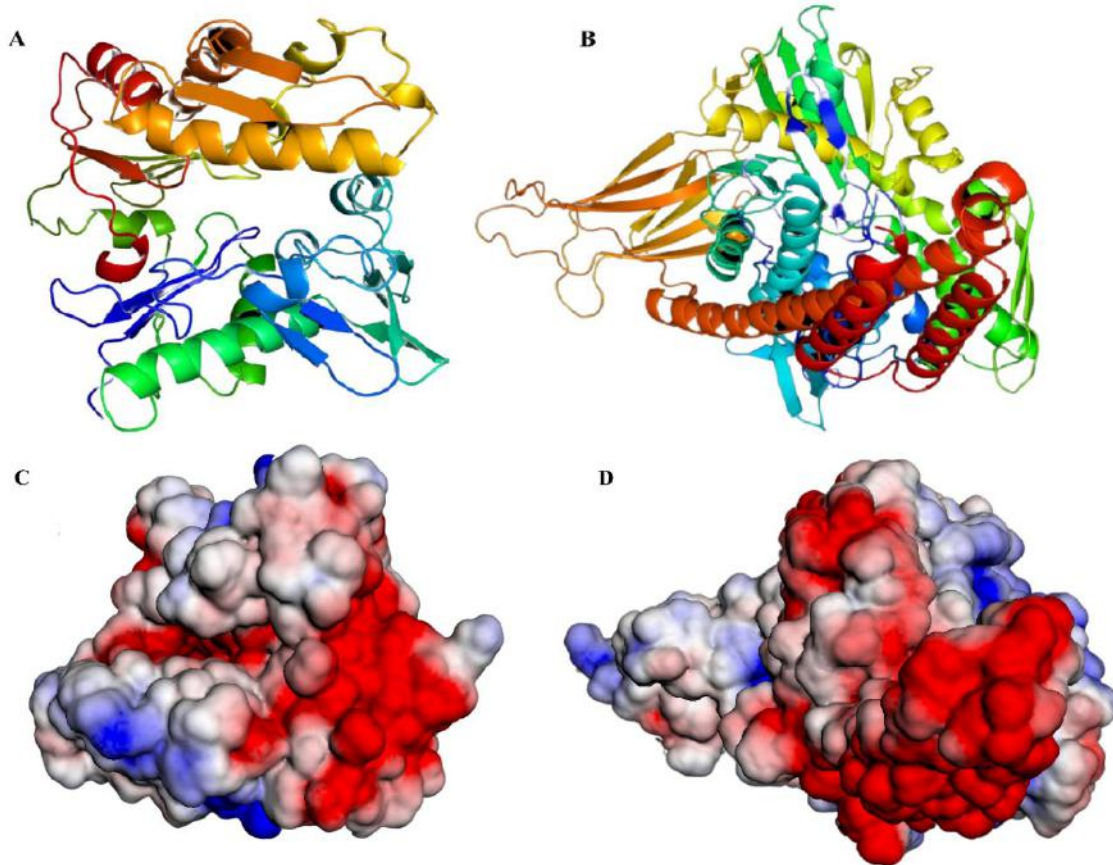
Figura 4 - Análise estereoquímica dos modelos tridimensionais das PvHSP70-12 (A) e PvHSP70-13 (B), representando os dois modelos com >90 dos aminoácidos nas regiões mais favoráveis.



Fonte: Autores (2023).

A estereoquímica dos dois modelos analisados foi estimada plotando os ângulos Phi e Psi no gráfico de Ramachandran, sendo possível observar que todos os modelos gerados possuem >90% dos aminoácidos nas posições mais favoráveis para os modelos variaram a porcentagem nas regiões permitidas, estando estas representadas na Figura 4. Em seguida, foram obtidos os dois modelos propostos, empregando abordagens de modelagem para construir a estrutura da proteína 3D,

Figura 5 - Estruturas tridimensionais das *heat shock protein*. A – PvHSP70-12 e B - PvHSP70-13, sendo demonstrado as alfa-hélices e as folhas betas. C e D representam o potencial eletrostático de superfície calculado com ABPS (PyMol), sendo as regiões carregadas negativamente representadas em vermelho.



Fonte: Autores (2023).

A Figura 5 demonstra a presença de estruturas α -hélice e folhas- β (A-B). A estrutura 3D pode elucidar na compreensão da estrutura tridimensional da proteína, onde as estruturas possuem monômeros uniformes com alfas hélices(α) e folhas beta (β). Isso acontece porque as estruturas são normalmente semelhantes, possuindo funções parecidas (TABUSAM et al., 2022).

De acordo com Rosenzweig et al. (2019), as HSP70s analisadas possuem similaridade com a PvHSP70-13, no entanto o autor retrata em seu estudo que as proteínas, sendo estruturalmente distintas, com poucas ou nenhuma homologia entre elas, pode diferir no seu modo de ligação com alguns ligantes, no entanto todos funcionam como um estabilizador de conformação abertas entre todas as HSP70s analisadas.

Com isso, de acordo com o design da estrutura de umas determinadas proteínas, pode-se determinar o ligante de ligação, que é de suma importância, para ligar aos resíduos externos ou internos. Além do mais, a estrutura 3D das proteínas pode prever determinadas conexões físicas, papéis funcionais distintos, sequências divergentes independentes, como também propriedades biológicas conservadas (Tabusam et al., 2022).

Em seguida, foi analisado o potencial eletrostático para as duas PvHSP70s (Figura 5 C-D), os nossos resultados demonstraram que o potencial eletrostático possui predominância de regiões negativas (vermelho) em regiões que possuem partes neutras (branco) e positivas (azuis). De acordo com Tomiczek et al. (2020), retratam que esse grupo de proteínas é

regulado eletrostaticamente, pela presença de um patch de potencial eletrostático uniformemente positivo em torno da hélice II do domínio J e um patch complementar de potencial negativo em torno do subdomínio IIa, ou seja, o potencial eletrostático tem a capacidades de influenciar a estrutura tridimensional das HSP70, como foi demonstrado pelos autores.

Além do mais, as interações das HSP70 fornecem um excelente sistema de modelos, voltados para estudos relacionado ao seu potencial eletrostático, juntamente de suas interações proteína-peptídeos (Kajander et al., 2009). Diante do que foi exposto, através dos resultados é possível observar que os modelos tridimensionais, são modelos em potencial a serem utilizados em estudos voltados a aplicação biotecnológica, no entanto necessita-se de estudos mais aprofundados sobre as PvHSP70-12 e PvHSP70-13.

4. Considerações Finais

As Heat shock proteins (proteínas de choque térmico) são conhecidas por sua atividade de defesa em vários estresses bióticos e abióticos, que acometem diversos organismos, como é o caso da espécie *Phaseolus vulgaris*. Neste estudo foi identificado um total de 15 genes *PvHSP70s* distribuídos em 6 cromossomos no genoma do feijão comum, possuindo atividade de defesa antimicrobiana e antiviral, podendo ser consideradas proteínas promissoras e aplicadas em estudos voltados a aplicações biotecnológicas. No entanto necessita-se de estudos mais aprofundados sobre a atividade de defesa das HSP70s no feijão comum, por meio de análise in vitro e in vivo, corroborando assim a sua eficácia.

Espera-se que essas novas informações a respeito dos genes das HSPs no feijão comum sejam úteis para mais investigações aprofundadas de possíveis respostas funcionais ao desenvolvimento, fisiologia relacionados à defesa na determinada espécie. Além disso, a presente pesquisa facilitará a caracterização funcional detalhada de genes HSPs de culturas de leguminosas em estudos futuros.

Agradecimentos

Os autores agradecem o apoio de agências brasileiras de fomento à pesquisa como CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) sob o Código Financeiro 001, e CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

Referências

- Adikshita, S., & Kansak, S. (2017). Evaluation of fungicides against angular leaf spot of French bean caused by *Phaeoisariopsis griseola*. *International Journal of Farm Sciences*, 6, 1–5. <https://www.indianjournals.com/ijor.aspx?target=ijor:ijfs&volume=6&issue=4&article=001>.
- Ahuja I., de Vos R. C., Bones A. M., & Hall R. D. (2010) Plant molecular stress responses face climate change. *Trends Plant Sci* 15(12):664–674. 10.1016/j.tplants.2010.08.002.
- Aki, T., Fujikawa, A., Wada, T., Jyo, T., Shigeta, S., Murooka, Y., & Ono, K. (1994). Cloning and expression of cDNA coding for a new allergen from the house dust mite, *Dermatophagoides farinae*: homology with human heat shock cognate proteins in the heat shock protein 70 family. *The Journal of Biochemistry*, 115(3), 435-440.
- Alvim F. C., Carolino S. M., Cascardo J. C., Nunes C. C., Martinez C. A., Otoni W. C., & Fontes E. P. (2001) Enhanced accumulation of BiP in transgenic plants confers tolerance to water stress. *Plant Physiol* 126(3):1042–1054. 10.1104/pp.126.3.1042.
- Anisimova, O. K., Kochieva, E. Z., Shchennikova, A. V., & Filyushin, M. A. (2022). Thaumatin-like protein (TLP) genes in garlic (*Allium sativum* L.): Genome-wide identification, characterization, and expression in response to *Fusarium proliferatum* infection. *Plants*, 11(6), 748.
- Assefa, T., Assibi Mahama, A., Brown, A. V., Cannon, E. K., Rubyogo, J. C., Rao, I. M., & Cannon, S. B. (2019). A review of breeding objectives, genomic resources, and marker-assisted methods in common bean (*Phaseolus vulgaris* L.). *Molecular Breeding*, 39, 1-23.
- Barre, A., Sénéchal, H., Nguyen, C., Granier, C., Poncet, P., & Rougé, P. (2023). Structural Basis for the IgE-Binding Cross-Reacting Epitopic Peptides of Cup s 3, a PR-5 Thaumatin-like Protein Allergen from Common Cypress (*Cupressus sempervirens*) Pollen. *Allergies*, 3(1), 11-24.
- Blair, M. W. (2013). Mineral biofortification strategies for food staples: the example of common bean. *Journal of agricultural and food chemistry*, 61(35), 8287-8294.

- Boubakri, H., Chihaoui, S. A., Najjar, E., Barhoumi, F., & Jebara, M. (2022). Comprehensive identification, evolutionary patterns and the divergent response of PRX genes in *Phaseolus vulgaris* under biotic and abiotic interactions. *3 Biotech*, *12*(8), 175.
- Broughton, W. J., Hernández, G., Blair, M., Beebe, S., Gepts, P., & Vanderleyden, J. (2003). Beans (*Phaseolus* spp.)—model food legumes. *Plant and soil*, *252*, 55-128.
- Buruchara, R., Chirwa, R., Sperling, L., Mukankusi, C., Rubyogo, J. C., Mutohi, R., & Abang, M. M. (2011). Development and delivery of bean varieties in Africa: the Pan-Africa Bean Research Alliance (PABRA) model. *African crop science journal*, *19*(4), 227-245.
- Cashikar A.G., Duenwald M., & Lindquist S.L. (2005) A chaperone pathway in protein disaggregation. Hsp26 alters the nature of protein aggregates to facilitate reactivation by Hsp104. *J Biol Chem* *280*(25):23869–23875. 10.1074/jbc.M502854200.
- Chen, M., Xu, J., Devis, D., Shi, J., Ren, K., Searle, I., & Zhang, D. (2016). Origin and functional prediction of pollen allergens in plants. *Plant physiology*, *172*(1), 341-357.
- Chuang, J. G., Su, S. N., Chiang, B. L., Lee, H. J., & Chow, L. P. (2010). Proteome mining for novel IgE-binding proteins from the German cockroach (*Blattella germanica*) and allergen profiling of patients. *Proteomics*, *10*(21), 3854-3867.
- Costa, J., Mafra, I., Carrapatoso, I., & Oliveira, M. B. P. (2016). Hazelnut allergens: Molecular characterization, detection, and clinical relevance. *Critical reviews in food science and nutrition*, *56*(15), 2579-2605.
- Costa, J., Villa, C., Verhoeckx, K., Cirkovic-Velickovic, T., Schrama, D., Roncada, P., & Holzhauser, T. (2022). Are physicochemical properties shaping the allergenic potency of animal allergens? *Clinical Reviews in Allergy & Immunology*, *62*(1), 1-36.
- Daugaard M., Rohde M., & Jaattela M. (2007) The heat shock protein 70 family: highly homologous proteins with overlapping and distinct functions. *FEBS Lett* *581*(19):3702–3710. 10.1016/j.febslet.2007.05.039
- Frydman J. (2001) Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu Rev Biochem* *70*:603–647. 10.1146/annurev.biochem.70.1.603
- Guo, W., Zhan, X., Jiang, F., & Xi, Y. (2021). Analysis of allergen components and identification of bioactivity of HSP70 in pollen of *Populus deltoides*. *Proteome science*, *19*, 1-14.
- Guo, Z., Li, Y., & Ding, S. W. (2019). Small RNA-based antimicrobial immunity. *Nature Reviews Immunology*, *19*(1), 31-44.
- Hao, M., Xijiri, Zhao, Z., & Che, H. (2022). Identification of allergens in white-and red-fleshed pitaya (*Selenicereus undatus* and *Selenicereus costaricensis*) seeds using bottom-up proteomics coupled with immunoinformatics. *Nutrients*, *14*(9), 1962.
- Hirano, K., Hino, S., Oshima, K., Nadano, D., Urisu, A., Takaiwa, F., & Matsuda, T. (2016). Evaluation of allergenic potential for rice seed protein components utilizing a rice proteome database and an allergen database in combination with IgE-binding of recombinant proteins. *Bioscience, Biotechnology, and Biochemistry*, *80*(3), 564-573.
- Jain, M., Amera, G. M., Muthukumar, J., & Singh, A. K. (2022). Insights into biological role of plant defense proteins: A review. *Biocatalysis and Agricultural Biotechnology*, *40*, 102293.
- Jiang, L., Hu, W., Qian, Y., Ren, Q., & Zhang, J. (2021). Genome-wide identification, classification and expression analysis of the Hsf and Hsp70 gene families in maize. *Gene*, *770*, 145348.
- Kajander, T., Sachs, J. N., Goldman, A., & Regan, L. (2009). Electrostatic interactions of Hsp-organizing protein tetratricopeptide domains with Hsp70 and Hsp90: computational analysis and protein engineering. *Journal of Biological Chemistry*, *284*(37), 25364-25374.
- Kesici, M., Ipek, A., Ersoy, F., Ergin, S., & Gülen, H. (2020). Genotype-dependent gene expression in strawberry (*Fragaria x ananassa*) plants under high temperature stress. *Biochemical genetics*, *58*, 848-866.
- Kumar, M., Tomar, M., Potkule, J., Punia, S., Dhakane-Lad, J., Singh, S., & Kennedy, J. F. (2022). Functional characterization of plant-based protein to determine its quality for food applications. *Food Hydrocolloids*, *123*, 106986.
- Li, Z., & Srivastava, P. (2004). Heat-shock proteins. *Current protocols in immunology, Appendix 1*. <https://doi.org/10.1002/0471142735.ima01ts58>.
- Liu, J., Han, D., & Shi, Y. (2019). Gene cloning, expression, and antifungal activities of permatin from naked oat (*Avena nuda*). *Probiotics and antimicrobial proteins*, *11*, 299-309.
- Liu, Q., Liang, C., & Zhou, L. (2020). Structural and functional analysis of the Hsp70/Hsp40 chaperone system. *Protein Science*, *29*(2), 378-390.
- Liu, X., Chen, H., Li, S., & Wang, L. (2022). Genome-wide identification of the Hsp70 gene family in grape and their expression profile during abiotic stress. *Horticulturae*, *8*(8), 743.
- Marcus, J. (2013). Protein Basics: Animal and vegetable proteins in food and health. *Culinary nutrition*, 189-230.
- Panzade, K. P. et al. (2021) Genome-wide analysis of Hsp70 and Hsp100 gene families in *Ziziphos jujuba*. *Cell Stress and Chaperones*, *26*(2), 341-353
- Masand S., & Yadav S.K. (2016) Overexpression of MuHSP70 gene from *Macrotyloma uniflorum* confers multiple abiotic stress tolerance in transgenic *Arabidopsis thaliana*. *Mol Biol Rep* *43*(2):53–64. 10.1007/s11033-015-3938-y.

- Musidlak, O., Nawrot, R., & Goździcka-Józefiak, A. (2017). Which plant proteins are involved in antiviral defense? Review on in vivo and in vitro activities of selected plant proteins against viruses. *International Journal of Molecular Sciences*, 18(11), 2300.
- Naikoo, G. A., Mustaqeem, M., Hassan, I. U., Awan, T., Arshad, F., Salim, H., & Qurashi, A. (2021). Bioinspired and green synthesis of nanoparticles from plant extracts with antiviral and antimicrobial properties: A critical review. *Journal of Saudi Chemical Society*, 25(9), 101304.
- Richter, K., Haslbeck, M., & Buchner, J. (2010) The heat shock response: life on the verge of death. *Mol Cell* 40(2):253–266. 10.1016/j.molcel.2010.10.006.
- Rosenzweig, R., Nilleghoda, N. B., Mayer, M. P., & Bukau, B. (2019). The Hsp70 chaperone network. *Nature reviews molecular cell biology*, 20(11), 665-680.
- Salas, C. E., Badillo-Corona, J. A., Ramírez-Sotelo, G., & Oliver-Salvador, C. (2015). Biologically active and antimicrobial peptides from plants. *BioMed research international*, 2015.
- Sarkar, N.K., Kim, Y.K., & Grover, A. (2009) Rice sHsp genes: genomic organization and expression profiling under stress and development. *BMC Genom* 10:393. DOI: 10.1186/1471-2164-10-393.
- Shevchenko, M., Servuli, E., Albakova, Z., Kanevskiy, L., & Sapozhnikov, A. (2021). The role of heat shock protein 70 kDa in asthma. *Journal of Asthma and Allergy*, 757-772.
- Tabassum, R., Dosaka, T., Ichida, H., Morita, R., Ding, Y., Abe, T., & Katsube-Tanaka, T. (2020). FLOURY ENDOSPERM11-2 encodes plastid HSP70-2 involved with the temperature-dependent chalkiness of rice (*Oryza sativa* L.) grains. *The Plant Journal*, 103(2), 604-616.
- Tabusam, J., Shi, Q., Feng, D., Zulfiquar, S., Shen, S., Ma, W., & Zhao, J. (2022). HSP70 gene family in Brassica rapa: Genome-wide identification, characterization, and expression patterns in response to heat and cold stress. *Cells*, 11(15), 2316.
- Timperio, A.M., Egidi, M.G., & Zolla, L. (2008) Proteomics applied on plant abiotic stresses: role of heat shock proteins (HSP). *J proteom* 71(4):391–411. DOI: 10.1016/j.jprot.2008.07.005.
- Tirolí-Cepeda, A. O., Lima, T. B., Balbuena, T. S., Gozzo, F. C., & Ramos, C. H. (2014). Structural and functional characterization of the chaperone Hsp70 from sugarcane. Insights into conformational changes during cycling from cross-linking/mass spectrometry assays. *Journal of proteomics*, 104, 48-56.
- Tomiczek, B., Delewski, W., Nierzwicki, L., Stolarska, M., Grochowina, I., Schilke, B., & Marszałek, J. (2020). Two-step mechanism of J-domain action in driving Hsp70 function. *PLoS Computational Biology*, 16(6), e1007913.
- Verhoeckx, K., Broekman, H., Knulst, A., & Houben, G. (2016). Allergenicity assessment strategy for novel food proteins and protein sources. *Regulatory Toxicology and Pharmacology*, 79, 118-124.
- Wang, H., Dong, Z., Chen, J., Wang, M., Ding, Y., Xue, Q., & Ding, X. (2022). Genome-wide identification and expression analysis of the Hsp20, Hsp70 and Hsp90 gene family in *Dendrobium officinale*. *Frontiers in Plant Science*, 13, 979801.
- Wang, W., Vinocur, B., Shoseyov, O., & Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in plant science*, 9(5), 244–252. DOI: 10.1016/j.tplants.2004.03.006.
- Wegele, H., Müller, L., & Buchner, J. (2004). Hsp70 and Hsp90—a relay team for protein folding. *Reviews of physiology, biochemistry and pharmacology*, 1-44.
- Yu, X., Mo, Z., Tang, X., Gao, T., & Mao, Y. (2021). Genome-wide analysis of HSP70 gene superfamily in *Pyropia yezoensis* (Bangiales, Rhodophyta): Identification, characterization and expression profiles in response to dehydration stress. *BMC Plant Biology*, 21, 1-14.
- Zhou, S.J., Jing, Z., & Shi, J.L. (2013) Genome-wide identification, characterization, and expression analysis of the MLO gene family in *Cucumis sativus*. *Genet mol res* 12(4):6565–6578. 10.4238/2013.December.11.8.
- Zhou, X., Su, L., Tang, R., Dong, Y., Wang, F., Li, R., & Li, H. (2023). Genome-wide analysis of Hsp40 and Hsp70 gene family in four cotton species provides insights into their involvement in response to *Verticillium dahliae* and abiotic stress. *Frontiers in Genetics*, 14, 1120861.
- Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., & Hendrickson, W. A. (1996). Structural analysis of substrate binding by the molecular chaperone DnaK. *Science (New York, N.Y.)*, 272(5268), 1606–1614. 10.1126/science.272.5268.1606.