Sari Agriculture Sciences and Natural Resources University

Journal of Crop Breeding Vol. 18, Issue 2, 2026 p: 60-78

OPEN ACCESS

Research Paper

Bioinformatic Analysis of the Heat Shock Transcription Factor (HSFs) Gene Family in *Camelina sativa*

Seyyed HamidReza Hashemipetrudi[1] and Leila Ahangar[2]

1- Department of Genetic Engineering and Biology, Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU), Sari, Iran

2- Associated Professor, Department of Plant Production, Gonbad Faculty of Agricultural Sciences and Natural Resources, Gonbad Kavos University, Gorgan, Iran, (Corresponding author: L.ahangar63@gmail.com)

Received: 18 September, 2025 Revised: 22 December, 2025 Accepted: 24 January, 2026

Extended Abstract

Background: Camelina (*Camelina sativa*) is one of the oldest plants from the Brassicaceae family, commonly known by the names “false flax” or “wild flax”. Camelina is a valuable and high-quality plant-based oil source with great potential for use in various industries to produce beneficial products. As with other plant species, Camelina encounters numerous environmental and non-environmental stresses during its vegetative growth stage. Investigating gene regulatory patterns and gene networks related to signaling and stress responses can help better understand the mechanisms of stress tolerance in this plant. One of the transcription factor families that contributes to appropriate responses to environmental stresses is the Heat Shock Factors (HSFs). In general, the accumulation of HSPs is essential for the survival of cells exposed to various environmental stresses. However, bioinformatics information regarding HSF genes in Camelina has not been reported yet. Therefore, the present study aimed to identify and analyze the gene structures, motifs, and domain conservation, as well as the three-dimensional structures of HSF proteins, in Camelina.

Methods: To identify and analyze the HSF gene family in *C. sativa*, genomic and protein sequence resources were obtained from the NCBI database. Initially, a tBLASTN analysis was performed using HSF protein sequences from the model plant *Arabidopsis* as input against the Camelina genome. After removing redundant sequences, the presence of specific domains was confirmed using the InterProScan tools. Domain searches were also conducted using HMMER v3.0 based on the HSF HMM profile obtained from Pfam database (PF00447) against the *C. sativa* proteome. Duplicate records were merged, and incomplete sequences were excluded. To accurately identify authentic members of the HSF gene family in *C. sativa*, a phylogenetic analysis was conducted using *Arabidopsis thaliana* HSF sequences as references. This analysis effectively distinguished true HSF proteins from those that merely contain the HSF-type winged-helix DNA-binding domain, ensuring the precise classification of HSF family members. The gene structure and physicochemical properties were analyzed using TBtools and ProtParam, while conserved motifs and subcellular localization were identified using MEME and WoLF PSORT, respectively. Phylogenetic analysis was conducted using MEGA 12.0, employing the maximum-likelihood method with bootstrap testing. 3D structures were predicted via Phyre2.0, and protein–protein interactions were analyzed using the STRING database.

Results: Based on the HMM model using the HMMER tool, a total of 137 isoforms of protein sequences containing HSF domains (PF00447) were identified in *C. sativa*. In total, 96 gene loci were determined after removing redundant isoforms. Ultimately, through phylogenetic analysis of these sequences alongside 21 HSF family proteins from *A. thaliana*, 64 protein sequences were confirmed as members of the HSF gene family in Camelina. Structural and physicochemical analysis revealed that the length of these proteins in Camelina ranged from 146 to 496 amino acids, with predicted molecular weights between 17.1 kDa and 55.2 kDa. The isoelectric points varied from 4.7 to 10.2, and the aliphatic index ranged from 58.28 to 76.54. An instability index above 40 was observed in 83% of the sequences, suggesting that these proteins may have lower stability under laboratory or intracellular conditions and are likely to degrade more rapidly. Furthermore, the negative GRAVY values for all CsHSF proteins indicate their hydrophilic nature. According to the subcellular localization analysis of CsHSF proteins, their predominant presence is in the nucleus, followed by the cytoplasm. Based on the results of the phylogenetic tree, ScHSF proteins were classified into three distinct major groups. Gene structure analysis of these groups based on the presence or absence of introns showed that, except for six intronless

Copyright ©2026. Hashemi Petrudi and Ahangar. Published by Sari Agricultural Sciences and Natural Resources University.
This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 Unported License, which allows users to read, copy, distribute, and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

genes, the remaining HSF genes contained introns with more than one phase. The variation in the intron number and phase in some Camelina HSF genes may be due to intron loss or gain during evolution, which likely contributes to the functional diversification of these proteins. Analysis of the relationship between the phylogenetic grouping of HSF proteins and exon number revealed that all groups exhibited a similar distribution in terms of exon count. While protein length and molecular weight varied significantly across groups, the majority of proteins in all groups contained only two or three exons. A total of 15 conserved motifs, ranging from 8 to 50 amino acids in length, were identified across all these proteins. Secondary structure analysis showed that group I proteins predominantly featured regular and continuous alpha-helices. Group II proteins exhibited more diverse secondary structures, including more prominent beta-sheets, extended coils, and a higher proportion of disordered regions. Proteins in group III displayed a combination of structural features from both groups I and II. Furthermore, interaction analysis indicated coordinated behavior among these proteins in response to various environmental conditions.
Conclusion: Bioinformatic identification and analysis of the HSF gene family in *C. sativa* led to the discovery of 64 members with diverse structural and functional characteristics. The relatively high number of family members compared to other species is likely associated with Camelina's allohexaploid nature, consisting of three distinct genomes. Gene structure analysis revealed the presence of conserved DBD and HR-A/B motifs in all members, indicating their central role in responding to environmental stress. Given Camelina's significance as a stress-tolerant oilseed crop, identifying these genes can pave the way for breeding programs aimed at developing cultivars with enhanced tolerance to environmental stresses.

Keywords: Allohexaploid, Gene expression regulation, Gene family Heat Shock Protein (HSP), Oilseed plant

How to Cite This Article: HashemiPetrudi, S. H. R., & Ahangar, L. (2026). Bioinformatic Analysis of the Heat Shock Transcription Factor (HSFs) Gene Family in *Camelina sativa*. *J Crop Breed,* 18(2), 60-78. DOI: 10.61882/jcb.2026.1622

دانشگاه علوم کشاورزی و منابع طبیعی ساری

مقاله پژوهشی

تحلیل بیوانفورماتیکی خانواده عامل رونویسی شوک حرارتی (HSFs) در گیاه کاملینا (*Camelina sativa*)

سیدحمیدرضا هاشمی پطرودی[۱] و لیلا آهنگر[۲]

۱- استادیار، گروه مهندسی ژنتیک و بیولوژی، پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران
۲- دانشیار، گروه تولیدات گیاهی، دانشکده علوم کشاورزی و منابع طبیعی گنبد، دانشگاه گنبد کاووس، گرگان، ایران، (نویسنده مسوول: L.ahangar63@gmail.com)

تاریخ دریافت: ۱۴۰۴/۰۶/۲۷ تاریخ ویرایش: ۱۴۰۴/۱۰/۰۱ تاریخ پذیرش: ۱۴۰۴/۱۱/۰۴
صفحه: ۶۰ تا ۷۸

چکیده مبسوط

مقدمه و هدف: کاملینا (*Camelina sativa*) یکی از قدیمی‌ترین گیاهان زراعی خانواده براسیکا است که با نام‌های کتان کاذب و کتان وحشی شناخته می‌شود. کاملینا یک نمونه از روغن‌های باارزش و با کیفیت بالای گیاهی است که پتانسیل بالایی را برای کاربرد در صنایع گوناگون جهت تولید محصولات مفید دارد. مشابه سایر گونه‌های گیاهی، گیاه کاملینا در طول دوره رویشی خود با تنش‌های محیطی و غیر محیطی متعددی مواجه می‌شود. بررسی الگوهای تنظیمی ژن‌ها و شبکه‌های ژنی مرتبط با سیگنال‌دهی و پاسخ به این تنش‌ها می‌تواند به درک بهتر مکانیسم‌های تحمل به تنش در این گیاهان کمک کند. یکی از خانواده‌های فاکتورهای رونویسی که به ایجاد پاسخ مناسب در برابر تنش‌های زیست‌محیطی کمک می‌کند، فاکتورهای شوک حرارتی (Heat Shock Factor) هستند. اصولاً افزایش تجمع HSPها برای بقای سلول‌هایی که تحت تاثیر تنش‌های مختلف محیطی قرار دارند، ضروری است. با این‌حال، اطلاعات بیوانفورماتیکی مربوط به ژن‌های فاکتور شوک حرارتی (HSF) در کاملینا تاکنون گزارش نشده است. بنابر این، این مطالعه با هدف شناسایی و تجزیه و تحلیل ساختار ژنی، حفاظت‌شدگی موتیف‌ها و دومین‌ها، و همچنین ساختار سه‌بعدی پروتئین‌های HSF در گیاه کاملینا انجام شد.

مواد و روش‌ها : در این مطالعه، برای شناسایی و تحلیل خانواده ژنی HSF در گیاه *C. sativa*، منابع توالی‌های ژنومی و پروتئینی از پایگاه NCBI دریافت شد. ابتدا، آنالیز tBLASTN با استفاده از توالی‌های پروتئینی HSF گیاه مدل آرابیدوپسیس به‌عنوان ورودی، در توالی ژنومی کاملینا انجام شد. پس از حذف توالی‌های تکراری، حضور دمین‌های اختصاصی با نرم‌افزار InterProScan تأیید شد. به‌منظور پوشش جامع‌تر، جستجوی دمین با ابزار HMMER v3.0 بر اساس پروفایل دمین HSF اخذ شده از PFAM با کد دسترسی PF00447 بر روی پروتئوم *C. sativa* نیز اجرا شد. رکوردهای تکراری ادغام و توالی‌های ناقص حذف گردیدند. الگوی اگزون–اینترون این ژن‌ها با استفاده از برنامه TBtools بررسی و ویژگی‌های فیزیکوشیمیایی پروتئین‌ها با ابزار ProtParam محاسبه گردید. برای شناسایی موتیف‌های حفاظت‌شده، برنامه MEME به‌کار رفت. پیش‌بینی مکان سلولی پروتئین‌ها به‌وسیله ابزار WoLF PSORT انجام شد. تجزیه و تحلیل فیلوژنتیکی با استفاده از نرم‌افزار ClustalW انجام و درخت فیلوژنتیکی با MEGA 12.0 بر مبنای روش حداکثر درستنمایی با آزمون بوت استرپ ترسیم شد. در نهایت، ساختار سه‌بعدی پروتئین‌ها نیز با استفاده Phayre 2.0 از پایگاه داده PDB پیش‌بینی گردید و برهمکنش‌های پروتئینی با پایگاه STRING تحلیل شدند.

یافته‌ها: طبق نتایج پژوهش برمبنای مدل HMM در ابزار HAMMER، تعداد ۱۳۷ ایزوفرم از توالی‌های پروتئینی حامل دمین HSFs در گیاه کاملینا شناسایی گردید که پس از حذف نسخه‌های ایزوفرم، تعداد ۹۶ جایگاه ژنی تعیین شدند. در نهایت، در بررسی روابط فیلوژنتیکی این توالی‌ها با تعداد ۲۱ پروتئین خانواده ژنی آرابیدوپسیس تالیانا، تعداد ۶۴ توالی پروتئینی برای خانواده ژنی HSF گیاه کاملینا شناسایی گردید. بررسی ویژگی‌های ساختاری و فیزیکوشیمیایی نشان داد که طول پروتئین‌های این خانواده در کاملینا از ۱۴۶ الی ۴۹۶ اسیدآمینه و وزن مولکولی پیش‌بینی شده نیز بین ۱۷/۱ کیلو دالتون تا ۵۵/۲ کیلو دالتون، نقاط ایزوالکتریک بین ۴/۷ الی ۱۰/۲ و شاخص آلیفاتیک بین ۵۸/۲۸ تا ۷۶/۵۴ متغیر بودند. برآورد شاخص ناپایداری بیش از ۴۰ برای ۸۳ درصد از توالی‌ها نشان‌دهنده‌ی این است که پروتئین در شرایط آزمایشگاهی یا درون‌سلولی پایداری کمتری دارد و احتمالاً سبب تخریب سریع‌تر می‌گردند. همچنین، منفی بودن شاخص آب‌گریزی (GRAVY) برای تمامی پروتئین‌های CsHSF بیانگر آب‌دوست بودن این پروتئین‌ها است. بررسی جایگاه سلولی پروتئین‌های CsHSF در اندامک‌های سلولی نشان داد که بیشترین محل حضور این پروتئین‌ها در هسته و سپس سیتوپلاسم بود. بر اساس نتایج درخت فیلوژنتیکی، پروتئین‌های ScHSF در سه گروه اصلی و متفاوت طبقه‌بندی شدند. بررسی ساختار ژنی این گروه‌ها بر اساس وجود یا عدم وجود اینترون نشان داد که پروتئین‌های خانواده HSF به‌جز شش ژن فاقد اینترون، بقیه ژن‌ها حداقل دارای اینترون‌هایی با بیش از یک فاز بودند. وجود تفاوت در تعداد و فاز اینترون در برخی از ژن‌ها HSF کاملینا می‌تواند به‌دلیل حذف یا به‌دست آوردن اینترون در طی تکامل باشد که احتمالاً سبب ایجاد یک نقش کارکردی برای این پروتئین‌ها شده است. در بررسی ارتباط بین گروه‌بندی فیلوژنتیکی پروتئین‌های HSF و تعداد اگزون‌ها، مشخص شد که تمامی گروه‌ها توزیع مشابهی از نظر تعداد اگزون داشتند. در حالی‌که طول پروتئین‌ها و وزن مولکولی در هر گروه تغییرات قابل توجهی داشتند، اکثریت پروتئین‌ها در تمامی گروه‌ها تنها ۲ یا ۳ اگزون داشتند. در تمام این پروتئین‌ها، ۱۵ موتیف حفاظت‌شده با طول بین ۸ تا ۵۰ آمینواسید شناسایی شدند. بر اساس بررسی ساختار دوم، پروتئین‌های گروه اول عمدتاً دارای مارپیچ‌های آلفا منظم و پیوسته، گروه دوم ساختار دوم متنوع‌تری دارند، شامل صفحات بتا برجسته‌تر، کویل‌های گسترده‌تر، و نواحی نامنظم بیشتر و پروتئین‌های گروه سوم ترکیبی از ویژگی‌های گروه‌های اول و دوم را دارا هستند. علاوه بر این، نتایج برهمکنش بیانگر هماهنگی بین این پروتئین‌ها در پاسخ به شرایط مختلف محیطی هستند.

نتیجه‌گیری: شناسایی و بررسی بیوانفورماتیکی خانواده ژنی HSFs در گیاه کاملینا منجر به شناسایی ۶۴ عضو شد که دارای ویژگی‌های ساختاری و عملکردی متفاوت هستند. بالا بودن تعداد اعضای این خانواده در گیاه کاملینا نسبت به سایر گیاهان احتمالاً به ماهیت آلوهگزاپلوئیدی آن مرتبط است که از سه ژنوم متفاوت تشکیل شده است. آنالیز ساختار ژنی حاکی از وجود موتیف‌های حفاظت‌شده DBD و HR-A/B در تمام اعضا است که نشان‌دهنده عملکرد محوری آن‌ها در پاسخ به تنش‌های محیطی است. با توجه به اهمیت کاملینا به‌عنوان گیاه دانه روغنی مقاوم به تنش، شناسایی این ژن‌ها می‌تواند زمینه‌ساز برنامه‌های اصلاحی برای تولید ارقام با تحمل بالاتر به تنش‌های محیطی باشد.

واژه‌های کلیدی: آلوهگزاپلوئید، تنظیم بیان ژن، خانواده ژنی، گیاه روغنی، HSP

مقدمه

کاملینا با نام علمی *Camelina sativa* گیاهی کمتر شناخته شده و متعلق به خانواده شب‌بو (Brassicaceae) است که در آسیا و اروپا کشت آن رایج است (Belayneh *et al*., 2015). کاملینا یک نمونه از روغن‌های باارزش و با کیفیت بالای گیاهی است که می‌تواند در صنایع مختلف برای تولیدات مفیدی مورد استفاده قرار بگیرد (Chaturvedi *et al*., 2018, Rezaiean Joybari *et al*., 2024). دوره رشدی این گیاه کوتاه و بین ۸۵ تا ۱۰۰ روز است. از مهم‌ترین مزایای تولید این گیاه

می‌توان به کشت در پاییز و برداشت آن در اواخر فروردین، نیاز آبی کم، سازگاری بهتر با شرایط اقلیمی منطقه، مقاومت به خشکی و سرمای بهاره، حساسیت کمتر به شیوع آفات و عدم حساسیت گیاه به ریزش محصول در زمان برداشت اشاره کرد (Chaturvedi *et al.,* 2018). کاملینا به‌دلیل مقاومت در برابر آفات و بیماری‌ها و همچنین نیاز تغذیه‌ای پایین، به مقدار کود کمتری در مقایسه با سایر گیاهان روغنی مانند کلزا، سویا و آفتابگردان نیاز دارد (Balanusa *et al.,* 2015, Amiri *et al.*, 2023). این گیاه نسبت به خشکسالی مقاوم است و این ویژگی آن را به یک محصول ایده‌آل برای مناطقی با بارندگی ناکافی تبدیل می‌کند (Murphy, 2016). عملکرد دانه گیاه حدود ۳–۱ تن در هکتار با محتوی ۳۸ تا ۴۳ درصد روغن (حاوی ۵۰ تا ۶۰ درصد اسیدهای چرب غیر اشباع و انواع توکوفرول‌های ارزشمند) و ۲۷ تا ۳۲ درصد پروتئین است (Gugel & Falk, 2006). در مقایسه با دیگر دانه‌های روغنی از جمله کلزا، کاملینا دارای بیشترین کارایی مصرف آب است (McVay & Khan, 2011).

گیاهان در طول دوره رشد و نمو خود، با تنش‌های غیر زیستی مختلفی، مانند سرما، گرما، خشکی و شوری مواجه می‌شوند، که رشدونمو، عملکرد و کیفیت آن‌ها را تحت تأثیر قرار می‌دهد. در گیاهان، بسیاری از ژن‌ها می‌توانند تحت تأثیر تنش‌های زیستی و غیرزیستی القاء شوند (Yokotani *et al*., 2013). بیان این ژن‌های القایی با فعال‌سازی طیف وسیعی از ژن‌های پایین‌دست مرتبط با تنش، می‌تواند تحمل گیاه را در برابر تنش غیر زیستی افزایش دهد (Zhang *et al*., 2013). یک مورد از این ژن‌های پاسخگو به تنش، ژن‌های شوک حرارتی هستند که پروتئین‌های شوک حرارتی[1] (HSPs) را کد می‌نمایند. آن‌ها در تاخوردگی، سرهم‌بندی[2]، تثبیت، فعال‌سازی و تخریب پروتئین، در بسیاری از فرآیندهای طبیعی سلولی یا تحت شرایط تنش، دخیل هستند (Schöffl *et al*., 1998). اصولاً افزایش تجمع HSPها برای بقای سلول‌هایی که تحت تاثیر تنش‌های مختلف محیطی قرار دارند، ضروری است (Zhang *et al*., 2013). در سلول‌های یوکاریوتی، بیان ژن‌های شوک حرارتی عمدتاً در سطح رونویسی توسط فاکتورهای شوک حرارتی[3] (HSFs) تنظیم می‌شود (Liu *et al*., 2013, Pandey *et al*., 2015). این فاکتورها از طریق شناسایی عناصر شوک حرارتی[4] (HSEs) موجود در منطقه پروموتر HSPها، بیان آن‌ها را تنظیم می‌کنند (Nover *et al*., 2001). HSPها، نقش‌های قابل‌توجهی در زمینه تحمل به تنش‌های غیرزیستی در گیاهان زراعی ایفا می‌نمایند (Singh *et al*., 2016). از این نظر شناسایی فاکتورهای رونویسی کنترل‌کننده این ژن‌ها (HSF)، اهمیت بسیاری دارد.

HSFها همانند دیگر فاکتورهای رونویسی، یک ویژگی معمول در انتهای آمینی HSFها یعنی یک ساختار مارپیچ–پیچ–مارپیچ[5] در دمین متصل‌شونده به DNA[6] (DBD) دارند؛ این دمین از حفاظت‌شدگی بالایی برخوردار است و می‌تواند برای تشخیص دقیق واحدهای تکراری عناصر سیس HSE (5'-nGAAn-3') مورد استفاده قرار گیرد (Damberger, Pelton *et al*., 1994). در کنار آن، یک دمین حفاظت شده دیگر به‌نام دمین الیگومریزاسیون[7] (OD) قرار دارد. این دمین شامل تکرارهای هپتاد[8] آب‌دوست (HR-A و HR-B) است، که با تشکیل ساختار حلقوی، فرایند اتصال تریمر پروتئین HSF را به نواحی تنظیمی سیس مستقر در پروموتر ژن‌های شوک حرارتی تسهیل می‌نماید. علاوه بر دمین‌های DBD و OD، سیگنال ورود هسته‌ای[9] (NLS) و سیگنال خروج هسته‌ای[10] (NES) با درجه حفاظت‌شدگی کمتر در انتهای کربوکسیلی و در مجاورت دمین فعال‌کننده مستقر هستند (Berberich *et al*., 2012).

لازم به‌ذکر است که HSFها بر اساس ساختار و ترکیب دمین‌هایشان در پلی‌پپتید، به سه دسته A، B و C طبقه‌بندی می‌شوند (Nover *et al*., 2001). کلاس A از HSFها عامل اصلی فعال‌کننده رونویسی در پاسخ به شوک حرارتی هستند و همچنین گزارش شده است که این زیرگروه می‌تواند واکنش رونویسی را در طول تنش‌های غیرزیستی میانجیگری کند (Liu *et al*., 2013). HSFها بیان ژن را برای حفظ هموستازی سلول و کنترل پاسخ استرس گرمایی (Heat Shock) و همچنین سایر عوامل محیطی تنظیم می‌کنند. در طول شرایط استرس‌زا، HSFها رونویسی را فعال نموده، بیش‌بیانی آنها به‌عنوان عاملی برای جمع‌آوری پروتئین‌های آسیب دیده و همچنین ذخیره پروتئین‌های جدید است (Kim *et al.,* 2021).

در بررسی دیگر، مشخص شد که HSFهای گیاهی در آرابیدوپسیس، به غیر از تنش گرمایی، تحت‌تأثیر سایر تنش‌های غیر زیستی مانند سرما، خشکی و شوری نیز القاء می‌شوند و همپوشانی گسترده‌ای بین مسیرهای پاسخ به تنش گرمایی و سایر تنش‌های غیر زیستی وجود دارد (Swindell *et al.*, 2007). علاوه بر این، نشان داده شده است که اعضای HSF به‌طور قابل‌توجهی با رشد گیاه و واکنش‌های مختلف استرس، مانند پاسخ به یخزدگی و نمک، خشکی، و تحمل به دمای بالا مرتبط هستند (Busch *et al*., 2005؛ Garg *et al*., 2002).

در تحقیقات پیشین، نقش HSFها در تنش‌های غیر زیستی مختلف به‌طور گسترده مورد بررسی قرار گرفته است. به‌طور مثال، بیان هترولوگ HSF برنج در گیاه آرابیدوپسیس منجر به افزایش تحمل آرابیدوپسیس تراریخت در برابر تنش غیر زیستی گردید (Yokotani *et al*., 2008). تنوع زیادی در عملکرد HSFهای گیاهی در پاسخ به شوک حرارتی وجود دارد. در بررسی برهم‌کنش‌های HSFها با HSPها و دیگر فاکتورهای رونویسی، مشخص شده است که بیان *HSFA1a* گوجه‌فرنگی برای پاسخ به شوک حرارتی، ضروری است (Mishra *et al*., 2002).

یونسی ملردی و همکاران (Younesi-Melerdi *et al*., 2020) نیز طی مطالعه در گیاه آرابیدوپسیس تحت تنش شوری نشان دادند که بیان ژن *HSFA1a* با مقدار نمک محیط رابطه

1. Heat shock proteins
2. Assembly
3. Heat shock factors
4. Heat shock elements
5. Helix-turn-helix
6. DNA-binding domain
7. Oligomerization domain
8. Heptad repeats
9. Nuclear localization signal
10. Nuclear export signal

مثبت داشت و با افزایش غلظت نمک میزان بیان ژن افزایش چشمگیری نشان داد. عملکرد HSFها به عنوان تنظیم‌کننده اصلی HSR گیاهی عمدتاً در آرابیدوپسیس و گوجه فرنگی شناسایی شده است (Yoshida *et al*., Mishra *et al*., 2002, 2011). در گوجه‌فرنگی، بیان بالای *HSFA1a*، تحمل گرمایی گیاه را بهبود بخشید، در حالی که جهش‌یافته‌های این ژن به تنش حرارتی حساس بودند؛ لذا، بر اساس این فنوتیپ به‌راحتی قابل شناسایی بودند (Mishra *et al*.,2002).

مطالعه انجام شده توسط لامکه و همکاران (Lamke *et al*., 2016) نشان داد که *HSFA2* فعال شدن پایدار چندین ژن درگیر در شوک حرارتی را از طریق متیلاسیون ناحیه ژنومی هدف، تقویت کرد. رونوشت‌های *HSFA2* در شرایط عادی غیرقابل شناسایی هستند. با این حال، پس از وقوع شوک حرارتی، بیان ژن *HSFA2* القا می‌شود، به‌نحوی‌که میزان بیان آن به قوی‌ترین حالت بیان فاکتورهای شوک گرما (HSF) القاشده در گیاهان تبدیل می‌شود (Nishizawa *et al*., 2006).

وانگ و همکاران (Wang *et al*., 2024) طی بررسی خانواده HSF در گیاه *Lycoris radiate* نشان دادند که اکثر *LrHSF*ها در هسته گیاه فعالیت داشتند. علاوه بر این، بیان ژن *LrHSF* را در بافت های مختلف و مراحل رشد گل تحت چهار تنش غیر زیستی مختلف مشاهده نمودند. از سویی، داده‌ها نشان دادند که ژن‌های *LrHSF*، به‌ویژه *LrHSF5*، اساساً در رشد و نمو *L. radiata* و پاسخ آن به تنش‌های غیرزیستی و هورمونی مختلف نقش داشتند. تجزیه و تحلیل شبکه تعامل ژن–ژن وجود اثرات هم‌افزایی بین پاسخ های مختلف ژن‌های *LrHSF* در برابر تنش‌های غیر زنده را نشان داد.

گزارشات حاکی از آن هستند که HSFها و ژن‌های کدکننده آن‌ها به‌طور گسترده‌ای در چندین گونه گیاهی نظیر فلفل (Guo *et al*., 2015)، ذرت (Haider *et al*., 2021)، کاهو (Kim *et al*., 2021)، تعیین خصوصیت شده‌اند اما تاکنون گزارشی در مورد ساختار، سازمان‌دهی، روابط فیلوژنتیکی و الگوی بیان این ژن‌ها در گیاه *C. sativa* ارائه نگردیده است. در حالی‌که مطالعه و بررسی الگوی تنظیمی ژن‌ها و شبکه‌های ژنی درگیر در سیگنالیگ و پاسخ به تنش‌های محیطی و غیرمحیطی می‌تواند در درک مکانیسم‌های دخیل در تحمل گیاهان به تنش مؤثر باشد (Zhao *et al*., 2016, Sanyal *et al*., 2016, Rao *et al*., 2016). لذا، با توجه به این‌که تاکنون اطلاعاتی در زمینه بررسی ژن‌های خانواده HSFs در گیاه *C. sativa* صورت نگرفته است، در این تحقیق بررسی این‌سیلیکو خانواده HSFs شامل شناسایی دمین‌های پروتئینی، بررسی ساختار ژنی، روابط تکاملی اورتوگ و پارالوگ مدنظر قرار خواهد گرفت که زمینه شناسایی ژن‌های کاندید پاسخگو به تنش را فراهم می‌نماید.

مواد و روش‌ها

شناسایی خانواده ژنی HSFS در *C. sativa*

برای شناسایی اعضای خانواده HSF در کاملینا (*C.sativa*)، ابتدا توالی‌های ژنومی، پروتئینی و CDS از پایگاه NCBI (https://www.ncbi.nlm.nih.gov) دریافت شدند. شناسایی اولیه با ابزار tBLASTN در نرم‌افزار BioEdit (Hall, 1999) انجام گرفت. برای این‌منظور، از توالی‌های پروتئینی خانواده HSF آرابیدوپسیس به‌عنوان query و با E-value کمتر از 1E-5 استفاده شد. پس از دریافت تمام توالی‌ها و حذف توالی‌های تکراری، در نهایت، توالی‌ها به‌صورت دستی برای تأیید حضور دمین اختصاصی خانواده HSFs با استفاده از پایگاه InterProScan (https://www.ebi.ac.uk/interpro/search/sequence-search) (Jones, Binns *et al*., 2014) مورد آنالیز تکمیلی قرار گرفتند. سپس، به‌منظور پوشش جامع‌تر، جستجوی دمین با ابزار HMMER v3.0 بر اساس پروفایل دمین HSF اخذشده از PFAM با کد دسترسی PF00447 بر روی پروتئوم *C. sativa* اجرا شد و در ادامه در ابزار CDD، حضور دمین‌های HSF در سطح آستانه کمتر از 1E-10 مورد تأیید قرار گرفت. رکوردهای تکراری ادغام، توالی‌های ناقص یا فاقد دامنه‌های کلیدی حذف گردیدند و در صورت وجود ایزوفرم، طولانی‌ترین رونوشت برای تحلیل‌های بعدی انتخاب شد. به‌منظور شناسایی دقیق‌تر اعضای خانواده ژنی HSF و تفکیک آن‌ها از پروتئین‌های دارای دمین HSF (منظور Winged helix DNA-binding domain) با اعضای خانواده ژنیHSF آرابیدوپسیس، آنالیز روابط فیلوژنتیکی پروتئین‌های گیاه کاملینا با پروتئین‌های خانواده ژنی آرابیدوپسیس تالیانا صورت گرفت.

تجزیه و تحلیل فیلوژنتیکی

روابط تکاملی میان ژن‌های *HSFs* در گونه *C. sativa* به‌کمک توالی‌های پروتئینی حاصل از آن‌ها بررسی شد. ابتدا، آنالیز هم‌ردیفی چندگانه با نرم‌افزار ClustalW (Thompson, Gibson *et al*., 2003) انجام گرفت و در نهایت درخت فیلوژنتیکی با استفاده از نرم‌افزار MEGA 12 (Tamura *et al*., 2013) برمبنای روش حداکثر درستنمایی[11] (ML) با آزمون بوت‌استرپ[12] (Felsenstein, 1985) با ۱۰۰ تکرار برای اطمینان از صحت گروه‌بندی ترسیم شد.

تجزیه و تحلیل ساختار ژنی به‌همراه شناسایی موتیف‌های حفاظت‌شده

برای بررسی تنوع و ساختار اعضای خانواده ژنی CSHSFs، با استفاده از برنامه TBtools ساختار اگزون – اینترون مربوط به هر ژن، ترسیم شد (Hu *et al*., 2015). علاوه بر این، توالی‌های آمینواسیدی آنها تحت آنالیز موتیف و پیش‌بینی دمین با برنامه MEME (http://meme-suite.org/tools/meme/) قرار گرفتند (Bailey *et al*., 2009). پارامترهای مورد استفاده شامل شناسایی حداکثر ۲۰ موتیف، و کمترین و بیشترین طول موتیف‌ها به‌ترتیب ۶ و ۵۰ اسیدآمینه بودند. سپس، انوتیشن آن‌ها در پایگاه اطلاعاتی InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/) مورد بررسی قرار گرفت.

تجزیه و تحلیل خصوصیات فیزیکوشیمیایی پروتئین‌ها و جایگاه سلولی آن‌ها

11. Maximum likelihood

12. Bootstrap

نقاط ایزوالکتریک (PI) و وزن مولکولی (MW) پروتئین‌ها نیز با استفاده از ابزار ProtParam از پایگاه اطلاعاتی آنلاین (https://web.expasy.org/protparam/) (Gasteiger *et al.*, 2005) برآورد شدند. برنامه آنلاین WoLF PSORT (https://wolfpsort.hgc.jp/) برای پیش‌بینی جایگاه سلولی پروتئین‌های CSHSFs مورد استفاده قرار گرفت (Horton *et al.*, 2007).

پیش‌بینی ساختار سه‌بعدی و شبکه برهم‌کنش پروتئین‌ها

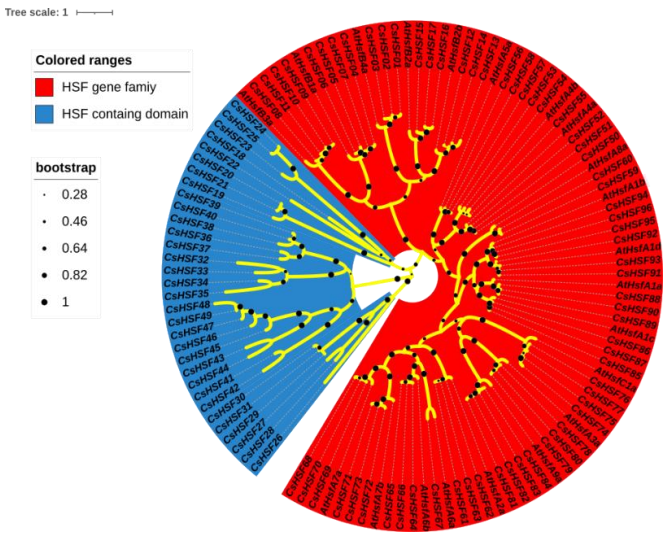
ساختارهای سه‌بعدی پروتئین‌های CsHSF با استفاده از ابزار Phayre 2 بر اساس PDB templateها مدل‌سازی شدند. میان‌کنش‌های عملکردی و فیزیکی بین پروتئین‌های CsHSF با استفاده از پایگاه اطلاعاتی STRING (http://string-db.org/) با شاخص حداقل امتیاز متعامل ۰/۴ ترسیم شدند. در این بخش از گیاه مدل *Arabidopsis thaliana* به‌عنوان گیاه مدل مرجع استفاده شد (Szklarczyk *et al.*, 2019).

نتایج و بحث

شناسایی پروتئین‌های دارای دمین HSF در گیاه *C. sativa*

بر مبنای جستجو بر مبنای مدل HMM در ابزار HAMMER، تعداد ۱۳۷ ایزوفرم از توالی‌های پروتئینی حامل دمین HSF در گیاه کاملینا شناسایی گردیدند که پس از حذف نسخه‌های ایزوفرم، تعداد ۹۶ جایگاه ژنی تعیین گردید. در ادامه به منظور شناسایی دقیق خانواده ژنی HSF در کاملینا، روابط فیلوژنتیکی ۹۶ پروتئین دارای دمین HSF در گیاه *C. sativa* با تعداد ۲۱ پروتئین‌های خانواده ژنی آرابیدوپسیس بررسی شد (شکل ۱). بر اساس نتایج حاصله، توالی‌ها در دو گروه اصلی دسته‌بندی شدند که هر گروه به رنگ خاصی مشخص شده است. به‌طوری‌که گروه دوم بیشترین تعداد از توالی‌ها را شامل شده است. با توجه به نحوه قرارگیری پروتئین‌های کاملینا در گروه دوم به همراه ژن‌های خانواده ژنی AtHSF (رنگ قرمز)، تعداد ۶۴ توالی پروتئینی برای خانواده ژنی HSF گیاه کاملینا شناسایی گردید که مشخصات و شماره دستیابی آنها در جدول ۱ ذکر شده‌اند. بررسی تعداد توالی‌های ژن HSF در گیاهان دیگر نشان می‌دهد که این تعداد شناسایی شده بسیار متفاوت و بیشتر (حدودا سه برابر) از دیگر گیاهان است؛ به عنوان مثال، ۲۱ پروتئین HSF در آرابیدوپسیس (Nover *et al.*, 2001)، ۲۲ مورد در *Allium sativum* (Yang *et al.*, 2024)، ۳۷ مورد در *Secale cereale* (Ren *et al.*, 2023) و ۲۲ مورد در *Lycoris radiate* (Wang *et al.*, 2024) گزارش شده‌اند. بررسی دقیق‌تر نشان می‌دهد که این اختلاف در تعداد ژن احتمالا به سطوح پلوئیدی در این گیاه با توجه به ماهیت آلوهگزاپلوئیدی آن مرتبط باشد که از سه ژنوم متفاوت تشکیل شده است (Blume *et al.*, 2024). در گندم نیز به دلیل ماهیت آلوپلوئیدی تعداد ۵۲–۸۳ توالی HSF در منابع مختلف شناسایی گردیده است (Ye *et al.*, 2020, Zhou *et al.*, 2019, Yun *et al.*, 2023).

توسعه ژنومیکس مقایسه‌ای امکان تجزیه و تحلیل خانواده‌های پروتئینی یکسان را در میان گونه‌های مختلف فراهم آورده است. بر این اساس، نامگذاری ژن‌های HSF کاملینا بر اساس روابط اورتولوژی با آرابیدوپسیس و روابط فیلوژنتیکی ۶۴ توالی پروتئینی HSF کاملینا و ۲۱ پروتئینی AtHSF صورت گرفت. در شکل ۲، جایگاه کروموزومی تمامی این توالی‌ها نشان داده شده است.



شکل ۱– درخت فیلوژنتیکی ۹۶ پروتئین حاوی دمین (HSF) با کد PFAM (PF00447) در گیاه آلوهگزاپلوئید *C. sativa* (CsHSF) و ۲۱ پروتئین خانواده HSF آرابیدپسیس تالیانا (AtHSF). درخت فیلوژنتیکی به کمک نرم‌افزار MEGA 12 با استفاده از توالی کامل پروتئین‌ها توالی آمینواسیدی بر مبنای روش حداکثر درستنمایی با بوت استرپ ۱۰۰ ترسیم شد. خانواده ژنی HSF در گروه قرمز دسته‌بندی گردید.

Figure 1. The phylogenetic tree of 96 domains (HSFs) containing the HSF domain (PFAM code: PF00447) in the allohexaploid plant *C. sative* (CsHSF) and 21 proteins of the *Arabidopsis thaliana* HSF family (AtHSF). The phylogenetic tree was constructed using MEGA 12.0 software, based on the full-length amino acid sequences of HSF proteins, based the Maximum-Likelihood (ML) method with 1,000 bootstrap replicates. The distinct HSF family was grouped into the red cluster.

جدول ۱- مشخصات و شماره دستیابی ۶۴ توالی پروتئینی خانواده ژنی HSF در گیاه *C. sativa*.

Table 1. Characteristics and accession numbers of 64 protein sequences of the HSF gene family in *C. sativa*.

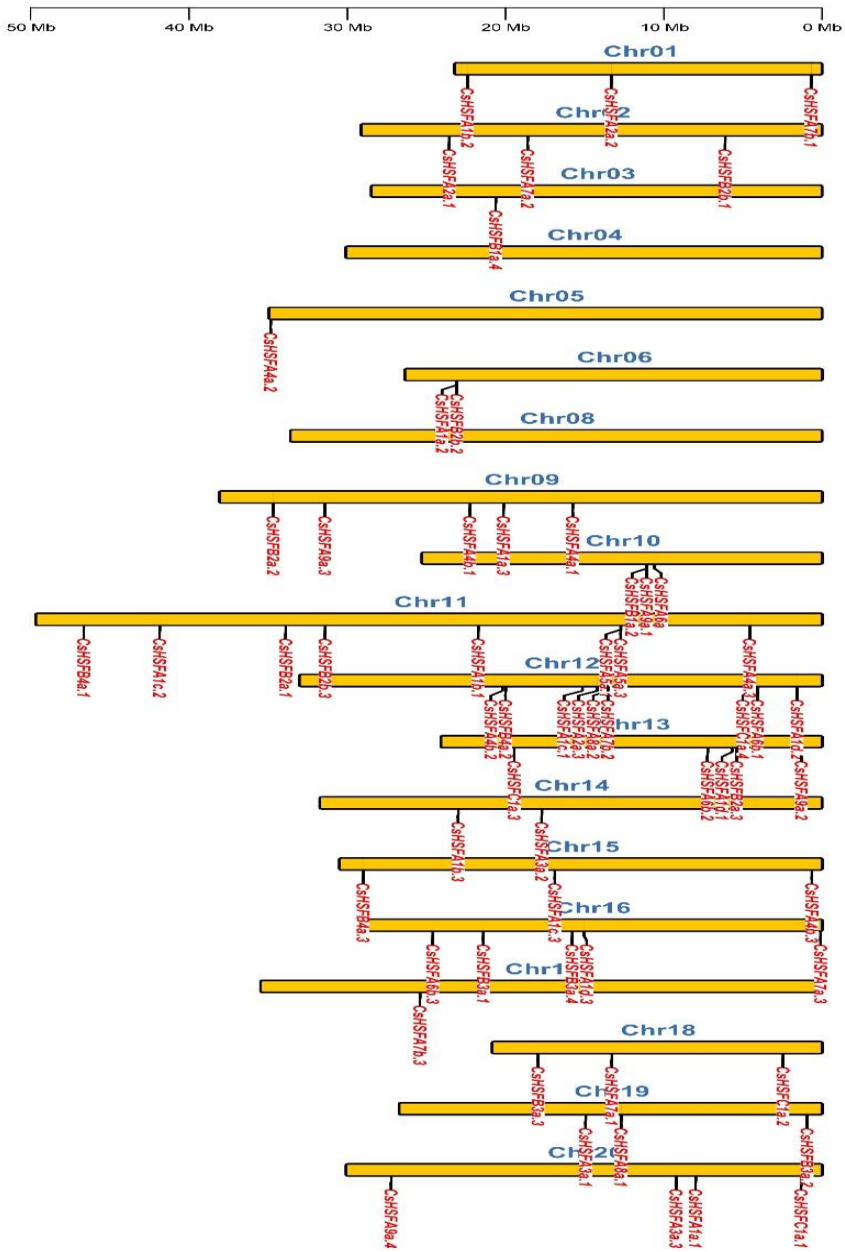
Sequence ID	Locus ID	mRNA ID	Protein ID	Length	Chromosome	Exon count
CsHSFA1a.1	LOC104769970	XM_019234213.1	XP_019089758.1	1455	20	4
CsHSFA1a.2	LOC104793084	XM_010436321.1	XP_010434623.1	441	6	2
CsHSFA1a.3	LOC104715472	XM_010441646.2	XP_010439948.1	693	9	2
CsHSFA1b.1	LOC104728202	XM_010455554.2	XP_010453856.1	384	11	2
CsHSFA1b.2	LOC104701663	XM_010422084.1	XP_010420386.1	402	1	3
CsHSFA1b.3	LOC104742226	XM_010494310.2	XP_010492612.1	1086	14	2
CsHSFA1c.1	LOC104733535	XM_010480354.2	XP_010478656.1	1341	12	2
CsHSFA1c.2	LOC104725927	XM_010462749.2	XP_010461051.1	1056	11	2
CsHSFA1c.3	LOC104747002	XM_010501475.2	XP_010499777.1	999	15	2
CsHSFA1d.1	LOC104735365	XM_010487329.2	XP_010485631.1	627	13	5
CsHSFA1d.2	LOC104729459	XM_010465435.2	XP_010463737.1	870	12	2
CsHSFA1d.3	LOC104750627	XM_010506362.2	XP_010504664.1	1125	16	2
CsHSFA2a.1	LOC104756193	XM_010431332.2	XP_010429634.1	453	2	1
CsHSFA2a.2	LOC104789585	XM_010419094.2	XP_010417396.1	1020	1	3
CsHSFA2a.3	LOC104731690	XM_010474339.2	XP_010472641.1	1176	12	4
CsHSFA3a.1	LOC104766383	XM_019228997.1	XP_019084542.1	1014	19	3
CsHSFA3a.2	LOC104741830	XM_010492562.2	XP_010490864.1	1491	14	3
CsHSFA3a.3	LOC104772350	XM_019234439.1	XP_019089984.1	405	20	1
CsHSFA4a.1	LOC104715324	XM_010441495.2	XP_010439797.1	405	9	1
CsHSFA4a.2	LOC104788675	XM_010436184.2	XP_010434486.1	882	5	3
CsHSFA4a.3	LOC104721613	XM_010451152.2	XP_010449454.1	675	11	2
CsHSFA4b.1	LOC104711738	XM_010443374.2	XP_010441676.1	831	9	2
CsHSFA4b.2	LOC104732306	XM_010483236.1	XP_010481538.1	1413	12	3
CsHSFA4b.3	LOC104744385	XM_010496392.1	XP_010494694.1	1416	15	3
CsHSFA5a.1	LOC104723082	XM_010451840.2	XP_010450142.1	411	11	1
CsHSFA5a.2	LOC104699542	XM_010436905.2	XP_010435207.1	885	7	3
CsHSFA5a.3	LOC104723078	XM_010451819.1	XP_010450121.1	504	11	3
CsHSFA6a	LOC104718351	XM_010447577.2	XP_010445879.1	381	10	2
CsHSFA6b.1	LOC104733175	XM_010468274.2	XP_010466576.1	588	12	4
CsHSFA6b.2	LOC104735719	XM_010490017.2	XP_010488319.1	1455	13	3
CsHSFA6b.3	LOC104752248	XM_010513451.2	XP_010511753.1	1050	16	3
CsHSFA7a.1	LOC104761432	XM_010517328.2	XP_010515630.1	1056	18	2
CsHSFA7a.2	LOC104753860	XM_010428470.2	XP_010426772.1	993	2	2
CsHSFA7a.3	LOC104748983	XM_010505598.2	XP_010503900.1	882	16	3
CsHSFA7b.1	LOC104781611	XM_010414855.2	XP_010413157.1	1416	1	4
CsHSFA7b.2	LOC104731596	XM_010470554.2	XP_010468856.1	381	12	2
CsHSFA7b.3	LOC104758004	XM_010514450.2	XP_010512752.1	1074	17	2
CsHSFA8a.1	LOC104766182	XM_019227694.1	XP_019083239.1	1227	19	2
CsHSFA8a.2	LOC104731597	XM_010472451.1	XP_010470753.1	510	12	1
CsHSFA9a.1	LOC104718434	XM_010448066.2	XP_010446368.1	1176	10	5
CsHSFA9a.2	LOC104734394	XM_010484511.2	XP_010482813.1	1335	13	3
CsHSFA9a.3	LOC104713426	XM_010444678.2	XP_010442980.1	405	9	1
CsHSFA9a.4	LOC104771806	XM_019235968.1	XP_019091513.1	1032	20	5
CsHSFB1a.1	LOC104704728	XM_010438958.1	XP_010437260.1	1122	7	3
CsHSFB1a.2	LOC104718438	XM_010448396.1	XP_010446698.1	567	10	2
CsHSFB1a.3	LOC104701524	XM_010438959.1	XP_010437261.1	696	7	2
CsHSFB1a.4	LOC104777629	XM_010433788.1	XP_010432090.1	1086	3	2
CsHSFB2a.1	LOC104728615	XM_010459760.2	XP_010458062.1	828	11	2
CsHSFB2a.2	LOC104714080	XM_010445736.2	XP_010444038.1	1050	9	3
CsHSFB2a.3	LOC104735299	XM_010485593.2	XP_010483895.1	375	13	2
CsHSFB2b.1	LOC104718514	XM_010423820.2	XP_010422122.1	1128	2	2
CsHSFB2b.2	LOC104793083	XM_010436278.2	XP_010434580.1	756	6	2
CsHSFB2b.3	LOC104724827	XM_010457280.2	XP_010455582.1	1110	11	3
CsHSFB3a.1	LOC104753895	XM_010510387.2	XP_010508689.1	405	16	1
CsHSFB3a.2	LOC104763923	XM_010519361.2	XP_010517663.2	1401	19	2
CsHSFB3a.3	LOC104762324	XM_010519360.1	XP_010517662.1	963	18	2
CsHSFB3a.4	LOC104750710	XM_010507653.2	XP_010505955.1	846	16	3
CsHSFB4a.1	LOC104726792	XM_010463203.2	XP_010461505.1	951	11	2
CsHSFB4a.2	LOC104732282	XM_010480803.2	XP_010479105.1	1416	12	3
CsHSFB4a.3	LOC104747816	XM_010501909.2	XP_010500211.1	948	15	4
CsHSFC1a.1	LOC104768556	XM_019229880.1	XP_019085425.1	1257	20	3
CsHSFC1a.2	LOC104760336	XM_010515257.2	XP_010513559.1	1050	18	3
CsHSFC1a.3	LOC104737167	XM_010490261.2	XP_010488563.1	1128	13	2
CsHSFC1a.4	LOC104730056	XM_010468566.2	XP_010466868.1	693	12	3

بررسی روابط تکاملی، موتیف‌ها، دمین‌ها و ساختار ژنی خانواده CsHSF

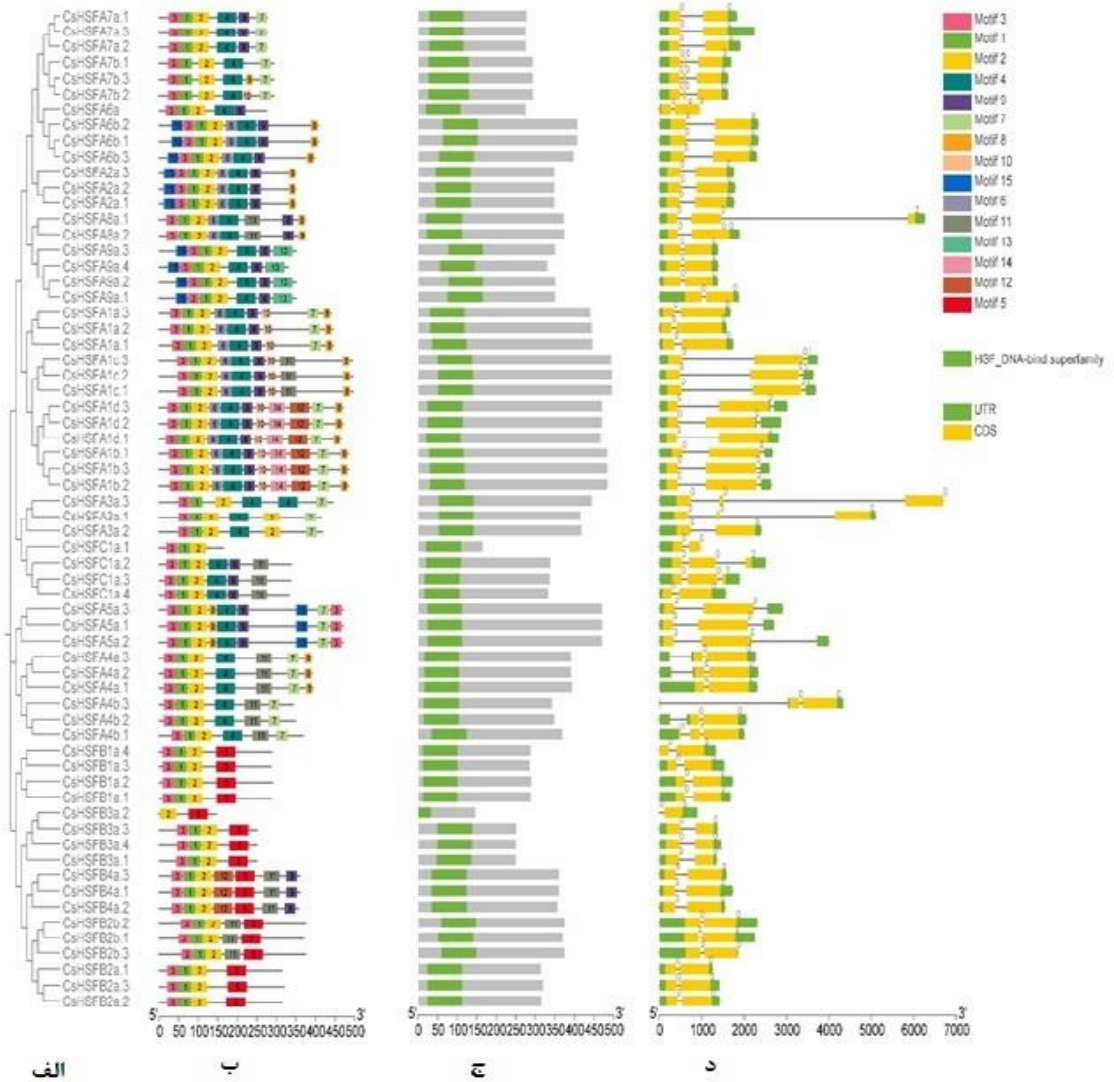
بررسی ساختار ژنی بر اساس وجود یا عدم وجود اینترون می‌تواند اطلاعات زیادی را در جهت تأیید روابط فیلوژنتیکی از نظر تکامل ژنومی ارائه دهد (Moore & Purugganan 2003). برای ترسیم ساختار ژنی از نرم‌افزار TBtools استفاده شد. همانگونه که در شکل ۳ نشان داده شده است، پروتئین‌های خانواده HSF بر اساس اندازه و طول اینترون با هم تفاوت دارند. اینترون‌ها بر اساس اختلال در آخرین کدون یک اگزون، به سه مرحله یا فاز دسته‌بندی می‌گردند. مرحله اول که در آن اینترون بعد از نوکلئوتید سوم کدون جای میگیرد فاز صفر نام دارد. مرحله دوم در آن اینترون بین باز اول و دوم کدون درج می‌شود

و فاز اول نام دارد. سپس، جایی‌که اینترون بین باز دوم و سوم یک کدون درج می‌شود فاز دوم نام دارد (Kim *et al*., 2021). به‌جز شش ژن فاقد اینترون، بقیه ژن‌ها حداقل دارای اینترون‌هایی با بیش از یک فاز هستند. وجود تفاوت در تعداد و فاز اینترون در برخی از ژن‌های HSF کاملینا می‌تواند به‌دلیل حذف یا به‌دست آوردن اینترون در طی تکامل باشد که احتمالاً سبب ایجاد یک نقش کارکردی برای این پروتئین‌ها شده است (Azizi & Zare 2022). در کل، افزایش و کاهش اینترون‌ها در تکامل ژن‌های HSF مشاهده می‌شود، اما در مقایسه با تکثیر و کاهش ژن، نقش کمتری در تکامل دارد. ژن‌های HSF تغییرات تکاملی مختلفی را در گونه‌های گیاهی متحمل شده‌اند که این تغییرات ساختاری به تنوع عملکردی ژن‌های HSF کمک و سازگاری آن‌ها با شرایط رشدی و محیطی را تقویت می‌کنند (Zafar *et al*., 2016).

در بررسی ارتباط بین گروه‌بندی فیلوژنتیکی پروتئین‌های HSF و تعداد اگزون‌ها، مشخص شد که تمامی گروه‌ها (گروه اول: CsHSFB1a.1، CsHSFB1a.2، CsHSFB1a.3، CsHSFB1a.4، CsHSFB2a.1، CsHSFB2a.2، CsHSFB2a.3، CsHSFB3a.1، CsHSFB3a.2، CsHSFB3a.3، CsHSFB3a.4، CsHSFB4a.1، CsHSFB4a.2، CsHSFB4a.3، CsHSFB2b.1، CsHSFB2b.2، CsHSFB2b.3، گروه دوم: CsHSFB4a.1، CsHSFB4a.2، CsHSFB4a.3، CsHSFB4b.1، CsHSFB4b.2، CsHSFB4b.3، CsHSFB5a.1، CsHSFB5a.2، CsHSFB5a.3 و گروه سوم توزیع مشابهی از نظر تعداد اگزون داشتند و اکثریت ژن‌های این خانواده در کاملینا تنها دارای ۲ یا ۳ اگزون بودند. بررسی‌های مشابه در *A. thaliana* نیز نشان داده‌اند که بیشتر ژن‌های این خانواده دارای ۲–۳ اگزون هستند (Sultan *et al*., 2016). این موضوع بیانگر الگوی محافظت‌شده‌ی ساختار ژنیHSFها در گونه‌های مختلف گیاهی است. علاوه بر این، نتایج نشان دادند که اگرچه طول پروتئین‌ها و وزن مولکولی در هر گروه تغییرات قابل‌توجهی داشتند، اکثریت پروتئین‌ها در تمامی گروه‌ها تنها ۲ یا ۳ اگزون داشتند. به‌طور خاص، گروه اول یک پروتئین با ۱ اگزون، و در گروه سوم پروتئین‌هایی با حداکثر ۳ اگزون نیز مشاهده شدند، اما هیچ‌یک از این ویژگی‌ها همبستگی آشکاری با تقسیم‌بندی فیلوژنتیکی نداشتند. از این رو، تعداد اگزون‌ها در این گروه‌ها به‌طور کلی مشابه باقی ماند. این یافته‌ها نشان می‌دهند که در حالی‌که ویژگی‌های فیزیکوشیمیایی مانند طول و وزن مولکولی پروتئین‌ها در بین گروه‌ها متفاوت هستند، تعداد اگزون‌ها در این پروتئین‌ها تقریبا ثابت است و بنابر این تعداد اگزون‌ها احتمالا عامل تعیین‌کننده اصلی در عملکرد HSFها در کاملینا نیستند. به‌عبارت دیگر، این نتایج می‌توانند نشان دهند که در انتخاب ژن‌های مرتبط به پاسخ به تنش‌های محیطی، تعداد اگزون‌ها نقش کلیدی ندارد و سایر ویژگی‌های ساختاری و تنظیمی اهمیت بیشتری دارند. جهت بررسی بیشتر و مشخص‌کردن میزان تشابه و تفاوت توالی پروتئین‌های HSF در *C. sativa* از نرم افزار MEME استفاده شد. بر این اساس، ۱۵ موتیف حفاظت‌شده با طول بین ۸ تا ۵۰ آمینواسید شناسایی شدند. موتیف ۲ در تمامی توالی‌ها ، موتیف ۱ و ۳ در ۶۳ توالی، موتیف ۴ در ۴۶ توالی و موتیف ۹ در ۳۷ توالی پروتئینی یافت شدند (شکل ۲). بر اساس بررسی جایگاه این موتیف‌ها، موتیف‌های اصلی (به‌ویژه موتیف‌های ۱، ۲ و ۳) در ناحیه‌ی DNA- (DBD) bindingdomain قرار دارند که مسئول اتصال به عناصر پاسخ شوک حرارتی (HSEs) است. موتیف ۴ و برخی موتیف‌های دیگر در بخش HR-A/Bdomain شناسایی شدند که برای الیگومریزاسیون و تشکیل تریمر HSF ضروری است. علاوه بر این، تعدادی از موتیف‌ها در نزدیکی نواحی Nuclear Localization Signal (NLS) و Nuclear Export Signal (NES) قرار دارند که در تنظیم انتقال هسته‌ای پروتئین نقش دارند. این نتایج نشان می‌دهند که موتیف‌های حفاظت‌شده‌ی شناسایی‌شده با دامنه‌های عملکردی شناخته‌شده‌ی HSFها هم‌پوشانی دارند و احتمالاً در حفظ عملکرد اساسی این فاکتورهای رونویسی (از جمله اتصال به DNA، تشکیل کمپلکس فعال و تنظیم جابه‌جایی هسته‌ای) نقش کلیدی ایفا می‌کنند (Lin *et al.*, 2011, Scharf *et al.*, 2012).



شکل ۲- جایگاه کروموزومی ۶۴ ژن خانواده HSF در گیاه *Camelina sativa*
Figure 2. Chromosome locations of 64 HSF gene families in *Camelina sativa*



شکل ۳- بررسی روابط تکاملی، شناسایی موتیف و دمین پروتئین‌های خانواده HSF بهمراه ساختار ژنی در *Camelina sativa*. الف) درخت فیلوژنی ۶۴ پروتئین CsHSF، ب) موتیف‌هایی که در توالی‌های پروتئینی شناسایی شدند. ج) شناسایی دمین‌های HSF در توالی پروتیئنی به‌همراه جایگاه آنها و در نهایت د) الگوی ساختار ژنی که جعبه‌های زرد نمایانگر اگزون و خطهای مشکی نمایانگر اینترون‌ها هستند. مناطق ترجمه نشده (UTR) بالادست و پایین‌دست با جعبه‌های سبز نشان داده شده‌اند.

Figure 3. Investigating evolutionary relationships, motif, and domain patterns of HSF family proteins along with their gene structure in *Camelina sativa.* a) The phylogenetic tree of the 64 HSF family, b) Motif identification in HSF protein sequences, c) Identification of HSF domains in the protein sequence, along with their locations, D) The gene structure of HSFs in *C. sativa.* Yellow boxes represent exons, and black lines indicate introns. The upstream and downstream untranslated regions (UTR) are shown as green boxes.

بررسی خصوصیات فیزیکوشیمیایی

بررسی ۶۴ توالی HSF نشان داد که طول پروتئین‌های این خانواده از ۱۴۶ (CsHSFB3a) تا ۴۹۶ (CsHSFA1b.3 و CsHSFA1c.2) اسیدآمینه متفاوت بود (جدول ۲). وزن مولکولی پیش‌بینی شده نیز بین ۱۷/۱ کیلو دالتون در توالی CsHSFB3a تا ۵۵/۲ کیلو دالتون در توالی CsHSFA1c.2 متغیر بود. شاخص آلیفاتیک که در یک پروتئین به‌عنوان حجم نسبی اشغال‌شده توسط زنجیره‌های جانبی آلیفاتیک، مانند آلانین، والین، ایزولوسین و لوسین، تعریف می‌شود که ممکن است به‌عنوان یک عامل مثبت برای افزایش پایداری حرارتی پروتئین‌های کروی در نظر گرفته شود که در این تحقیق بین ۵۸/۲۸ در توالی CsHSFA6b.1 تا ۷۶/۵۴ در توالی CsHSFA8a.1 متغیر بود. نقاط ایزوالکتریک پیش‌بینی شده (pI) بین ۴/۷ (CsHSFA1c) و ۱۰/۲ (CsHSFB4a) متغیر بود. شاخص ناپایداری ۵۹ پروتئین از ۶۴ پروتئین CsHSF (۸۳/۶ درصد) بیشتر از ۴۰ پیش‌بینی شده است که نشان‌دهنده‌ی این است که پروتئین در شرایط آزمایشگاهی یا درون‌سلولی پایداری کمتری دارد و احتمالاً سریع‌تر تخریب می‌گردد. این ویژگی می‌تواند به معنای تنظیم سریع و عملکرد موقتی نیز باشد، یعنی CsHSFها در پاسخ به تنش‌ها به‌سرعت فعال شده و سپس حذف می‌شوند تا تعادل سلولی حفظ گردد. پروتئین‌های ناپایدار در اغلب موارد برای حفظ ساختار و عملکرد خود نیاز به تعامل با پروتئین‌های دیگر دارند. بنابر این، احتمالاً عملکرد واقعی CsHSFها وابسته به تشکیل کمپلکس‌های چندپروتئینی است. همچنین، ناپایداری بالا می‌تواند نشانه‌ای از نقش CsHSFها در پاسخ‌های سریع و گذرا به تنش‌های محیطی مثل گرما، خشکی یا اکسیداتیو باشد که در گیاهی مقاوم به شرایط تنش مانند کاملینا این نوع ناپایداری می‌تواند مزیت تکاملی محسوب گردد (Gamage *et al*., 2019).

شاخص آب‌گریزی (GRAVY) پروتئین که به‌صورت مثبت و منفی مشخص می‌گردد در تمامی پروتئین‌های CsHSF منفی بوده، که نشان می‌دهد پروتئین‌های CsHSF آب‌دوست هستند در حالی‌که مثبت بودن این شاخص، دلالت بر آبگریز بودن پروتئین‌ها دارد (جدول ۲). مطالعات قبلی نشان می‌دهند که بیشتر پروتئین‌های HSF در گیاهان مختلف از نوع آب‌دوست، اسیدی، ناپایدار با شاخص آلیفاتیک بالا هستند که ویژگی‌های بارز پروتئین‌های مرتبط با تنش هستند (Rao *et al*., 2008). پروتئین آب‌دوست معمولاً در محیط‌های سلولی محلول‌ترند. مطالعات وانگ و همکاران (Wang *et al*., 2024) بر روی خانواده HSF در گیاه *Lycoris radiate* نشان دادند که طول پروتئین‌های این خانواده از ۴۶۸–۲۱۵ اسیدآمینه و وزن مولکولی پیش‌بینی شده نیز بین ۵۱/۳ — ۲۴/۹ کیلودالتون متغیر بود. همچنین، شاخص آلیفاتیک بین ۸/۹ — ۴/۷۷ متغیر بود. تفاوت مشاهده شده در مشخصات فیزیکوشیمیایی، مانند طول پروتئین خانواده HSF در کاملینا و گیاهان دیگر، احتمالاً ناشی از ساختار خاص کاملینا به‌عنوان یک گونه آلوهگزاپلوئید و به تبع وجود نسخه‌های متعدد ژنی است (Bird *et al*., 2025, Blume *et al*., 2024). همچنین، تنوع در نواحی غیر محافظت‌شده پروتئین‌ها مانند linker regions و activation domains و فشارهای انتخابی ناشی از شرایط محیطی (مانند خشکی و سرما) می‌توانند در ایجاد این تنوع نقش داشته باشند (Chaudhary *et al*., 2023).

بررسی جایگاه سلولی پروتئین‌های CsHSF در اندامک‌های سلولی با استفاده از سایت آنلاین WoLF PSORT نشان می‌دهد که بیشترین محل حضور این پروتئین‌ها در هسته و سپس سیتوپلاسم است. وانگ و همکاران (Wang *et al*., 2024) و هایر و همکاران (Haider *et al.,* 2021) نیز به‌ترتیب هسته را محل حضور خانواده HSF در گیاه *Lycoris radiate* و *Zea mays* عنوان نمودند. همچنین، نتایج بررسی *LrHSF*ها در گیاه *Lycoris radiata* نشان دادند که اکثر این پروتئین‌ها در بخش هسته قرار داشتند، در حالی‌که LrHSF9 در سیتوپلاسم قرار داشت (Wang *et al*., 2024).

جدول ۲- خصوصیات فیزیکوشیمیایی پروتئین‌های HSF در *Camelina sativa* تعداد اسیدآمینه، وزن مولکولی، نقطه ایزوالکتریک، شاخص ناپایداری، شاخص آلیفاتیک و متوسط هیدروپاتی کل پروتئین‌ها با استفاده از ابزار Expasy-ProtParam برآورد شد.

Table 2. Physicochemical Properties of HSF Proteins in *Camelina sativa*. The amino acid number, molecular weight, isoelectric point, instability index, aliphatic index, and average hydropathicity of total proteins were estimated using the Expasy-ProtParam tool.

نام ژن Gene name	طول پروتئین Protein length (aa)	وزن مولکولی Molecular weight (kDa)	نقطه ایزوالکتریک Isoelectric point	شاخص ناپایداری Instability index	شاخص آلیفاتیک Aliphatic index	متوسط هیدروپاتی کل GRAVY
CsHSFA1a.1	446	49.8	5.2	60.66	70.61	-0.636
CsHSFA1a.2	444	49.9	5.2	62.58	73.56	-0.617
CsHSFA1a.3	440	49.5	5.14	62.18	73.98	-0.618
CsHSFA1b.1	484	53.5	4.81	52.77	68.6	-0.675
CsHSFA1b.2	485	53.6	4.81	51.28	69.28	-0.665
CsHSFA1b.3	484	53.6	4.89	50.91	68.41	-0.693
CsHSFA1c.1	496	55.2	4.7	59.84	62.28	-0.778
CsHSFA1c.2	496	55.1	4.73	61.86	62.3	-0.771
CsHSFA1c.3	494	55	4.72	60.95	61.76	-0.78
CsHSFA1d.1	466	51.4	4.99	51.73	76.07	-0.499
CsHSFA1d.2	471	52	4.94	51.73	75.88	-0.507
CsHSFA1d.3	471	52	5	52.46	75.48	-0.515
CsHSFA2a.1	349	39.2	5.11	46.04	68.65	-0.573
CsHSFA2a.2	349	39.2	5.12	48.64	66.7	-0.572
CsHSFA2a.3	349	39.2	5.12	46.67	66.42	-0.596
CsHSFA3a.1	416	46.7	4.96	51.79	62.57	-0.687
CsHSFA3a.2	418	46.8	4.83	54.02	63.21	-0.691
CsHSFA3a.3	444	50.2	5.22	48.85	62.59	-0.745
CsHSFA4a.1	394	45.1	5.27	65.2	61.07	-0.886
CsHSFA4a.2	391	44.8	5.32	64.16	61.76	-0.876
CsHSFA4a.3	391	44.9	5.38	65	61.51	-0.898
CsHSFA4b.1	369	42.3	5.56	59.59	64.88	-0.919
CsHSFA4b.2	349	40	5.45	57.11	64.13	-0.941
CsHSFA4b.3	343	39.5	5.3	59.61	64.11	-0.964
CsHSFA5a.1	470	52.8	8.52	54.75	64.17	-0.884
CsHSFA5a.2	470	52.9	8.52	53.85	64.38	-0.889
CsHSFA5a.3	471	52.9	7.77	55.41	65.69	-0.859
CsHSFA6a	275	32.3	6.99	51.18	67.31	-0.818
CsHSFA6b.1	408	47	5.06	62.42	58.28	-0.928
CsHSFA6b.2	408	47	4.97	63.47	59.95	-0.9
CsHSFA6b.3	398	45.8	5.01	63.09	59.25	-0.909
CsHSFA7a.1	277	32	5.82	60.97	66.17	-0.701
CsHSFA7a.2	276	32	5.95	58.04	66.41	-0.667
CsHSFA7a.3	275	31.9	5.96	60.37	66.65	-0.699
CsHSFA7b.1	294	33.9	5.67	70.55	65.34	-0.694
CsHSFA7b.2	293	33.7	6.03	69.73	66.59	-0.668
CsHSFA7b.3	293	33.7	5.78	64.88	67.88	-0.657
CsHSFA8a.1	373	42.1	4.89	41.13	76.54	-0.693
CsHSFA8a.2	374	42.2	4.85	41.55	76.34	-0.692
CsHSFA9a.1	350	39.6	6.44	43.08	74.86	-0.699
CsHSFA9a.2	351	39.5	6.49	40.85	71.6	-0.711
CsHSFA9a.3	351	39.6	5.99	44.38	72.96	-0.718
CsHSFA9a.4	330	37.9	6.24	42.71	73.45	-0.746
CsHSFB1a.1	287	31.3	5.78	38.41	61.43	-0.718
CsHSFB1a.2	289	31.7	5.65	39.32	63.67	-0.746
CsHSFB1a.3	286	31.2	5.78	39.24	60.98	-0.726
CsHSFB1a.4	288	31.4	5.94	37.99	60.21	-0.758
CsHSFB2a.1	314	35.3	6.98	60.69	58.89	-0.839
CsHSFB2a.2	316	35.6	6.2	61.18	60.98	-0.808
CsHSFB2a.3	320	36	6.62	62.43	60.84	-0.819
CsHSFB2b.1	370	39.3	4.75	54.46	69.43	-0.453
CsHSFB2b.2	375	39.5	4.78	52.37	70.08	-0.425
CsHSFB2b.3	375	39.5	4.79	55.13	67.73	-0.47
CsHSFB3a.1	250	28.9	5	50.28	61.92	-0.965
CsHSFB3a.2	146	17.1	5.82	54.38	60.68	-1.184
CsHSFB3a.3	251	29.1	5.13	55.07	62.47	-0.946
CsHSFB3a.4	251	29	5.12	54.54	62.47	-0.938
CsHSFB4a.1	361	41.1	7.88	57.71	62.35	-0.733
CsHSFB4a.2	357	40.5	8.41	53.41	63.05	-0.702
CsHSFB4a.3	361	40.9	8.71	57.33	63.43	-0.716
CsHSFC1a.1	165	19.5	6.21	36.08	75.64	-0.687
CsHSFC1a.2	339	38.2	5.69	54.47	63.51	-0.723
CsHSFC1a.3	337	37.9	5.68	56.2	64.75	-0.698
CsHSFC1a.4	332	37.7	5.82	54.25	66.02	-0.699

پیش‌بینی ساختار دوم پروتئین CsHSF

ساختار ثانویه هشت توالی پروتئین HSF در *C. sativa* با استفاده از نرم‌افزار Psipred در شکل ۴ نشان داده شده است. دو توالی CsHSFA4b.1 و CsHSFA5a.1 از گروه ۱، دو توالی CsHSFB1a.3 و CsHSFB2b.2 از گروه دوم، و دو توالی CsHSFA7b.3 و CsHSFC1a.4 از گروه سوم به صورت تصادفی انتخاب شدند. تحلیل تطبیقی ساختار پروتئین‌های مورد بررسی از سه گروه نشان می‌دهد که هر گروه دارای الگوی ساختاری و عملکردی متمایز و در عین حال مکمل است. پروتئین‌های گروه اول عمدتاً دارای مارپیچ‌های آلفا منظم و پیوسته هستند که با نقش کلاسیک فاکتورهای رونویسی HSF هم‌راستاست؛ این ساختارها در نواحی DBD و HR-A/B برای اتصال به DNA و تشکیل دایمرهای فعال ضروری‌اند. در مقابل، پروتئین‌های گروه دوم ساختار دوم متنوع‌تری دارند، شامل صفحات بتا برجسته‌تر، کویل‌های گسترده‌تر، و نواحی نامنظم بیشتر که نشان‌دهنده‌ی انعطاف‌پذیری فضایی و قابلیت تعامل با اجزای مختلف سلولی است. همچنین، این گروه دارای ویژگی‌های عملکردی برجسته‌تری مانند پپتیدهای سیگنال، هلیکس‌های غشایی، و جایگاه‌های اتصال فلز هستند که بیانگر نقش‌های تنظیمی، ترشحی، یا پاسخ سریع به تنش‌های محیطی است. پروتئین‌های گروه سوم، به‌ویژه CsHSFA7b.3 و CsHSFC1a.4، ترکیبی از ویژگی‌های گروه‌های اول و دوم را دارا هستند. ساختار دوم آن‌ها شامل مارپیچ‌های آلفا، صفحات

بتا، و نواحی کویل است که به‌صورت متعادل در توالی توزیع شده‌اند. حضور دامنه‌های اتصال به DNA در CsHSFA7b.3، همراه با ساختارهای غشایی و سیگنال‌دهی در هر دو پروتئین، نشان‌دهنده‌ی عملکرد چندگانه این گروه است؛ از یک سو، در تنظیم رونویسی و از سوی دیگر در تعاملات غشایی، تنظیم‌های پس‌ازرونویسی، و پاسخ‌های محیطی. چنین تنوع عملکردی در خانواده HSF پیش‌تر نیز گزارش شده است (Scharf *et al*., 2012). بیانگر پیچیدگی تنظیمی این فاکتورها در گیاهان، به‌ویژه در شرایط تنش گرمایی و محیطی است. در مجموع، می‌توان گفت که گروه اول نمایانگر عملکرد کلاسیک رونویسی، گروه دوم نشان‌دهنده‌ی فعالیت در تنظیم‌های غشایی و سیگنال‌دهی، و گروه سوم حامل ویژگی‌های تلفیقی و تطبیقی است که احتمالاً در تنظیم‌های چندلایه و پاسخ‌های پیچیده نقش دارند.

مطالعات ساختاری در گونه‌های مدل مانند *A. thaliana* و *Oryza sativa*نشان داده‌اند که مارپیچ‌های آلفا ساختار غالب در فاکتورهای HSF هستند، به‌ویژه در نواحی DBD و HR-A/Bکه برای اتصال به DNA و تشکیل دایمرهای فعال ضروری‌اند (Scharf *et al*., 2012, Guo *et al*., 2016). همچنین، مشابهت الگوی ساختاری پروتئین‌های گروه اول با گونه‌های دیگر نشان‌دهنده‌ی حفظ ساختار عملکردی در طول تکامل است. در مقابل، تنوع در ساختار پروتئین‌های گروه دوم و سوم که در مطالعات برنج نیز گزارش شده است، بیانگر آن است‌که برخی اعضای خانواده HSF دارای ویژگی‌های تنظیمی و غشایی برجسته‌تری هستند (Shamshad *et al*., 2023). همچنین بررسی مطالعات بین گونه‌ها نشان می‌دهد که ساختار آلفا–هلیکس در عملکرد رونویسی نقش مرکزی دارد؛ بنابر این، تنوع ساختاری در سایر اعضا می‌تواند با عملکردهای تنظیمی، ترشحی، یا پاسخ به تنش‌های محیطی مرتبط باشد.

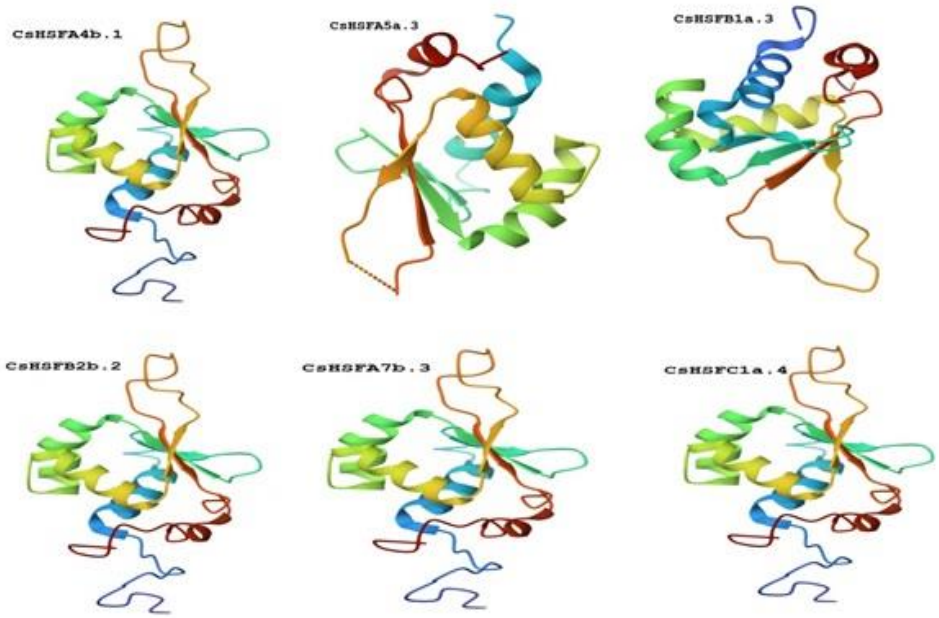
پیش‌بینی ساختار سه‌بعدی پروتئین CsHSF

جهت پیش‌بینی ساختار سه‌بعدی پروتئین‌های CsHSF، مدل‌سازی همولوژی پروتئین با استفاده از Phayre 2 بر اساس PDB انجام شد. ابتدا، یک الگوی مناسب برای هر پروتئین یافت شد و سپس یک مدل با استفاده از الگو تولید شد (شکل ۵). با توجه به این‌که ساختار سه‌بعدی، عملکرد اختصاصی پروتئین را تضمین می‌کند، این امر می‌تواند بیانگر شباهت‌های عملکردی در میان پروتئین‌های CsHSF باشد. تحلیل ساختار سه‌بعدی پروتئین‌های CsHSFA4b.1 و CsHSFA5a.1 از گروه اول، پروتئین‌های CsHSFB1a.3 و CsHSFB2b.2 از گروه دوم و پروتئین‌های CsHSFA7b.3 و CsHSFC1a.4 از گروه سوم نشان می‌دهد که الگوی فضایی این پروتئین‌ها با طبقه‌بندی عملکردی آن‌ها هم‌راستا است. در پروتئین‌های گروه اول، مارپیچ‌های آلفا در نواحی ابتدایی و مرکزی غالب‌اند و دامنه اتصال به DNA (DBD) به‌وضوح قابل تشخیص است؛ این ناحیه شامل سه مارپیچ آلفا متوالی با بسته‌بندی هیدروفوبیک است که با ساختار گزارش‌شده در AtHSFA1 و AtHSFA2 در *A. thaliana* مشابه است (Scharf *et al.,* 2012). ناحیه HR-A/B نیز با مارپیچ‌های موازی در بخش میانی ساختار دیده می‌شود که برای دایمرسازی ضروری است. در مقابل، پروتئین‌های گروه دوم دارای ساختار دوم متنوع‌تری هستند؛ مارپیچ‌های آلفا با صفحات بتا و کویل‌های سطحی ترکیب شده‌اند و دامنه DBD کوتاه‌تر و فشرده‌تر است. این ویژگی‌ها با ساختار AtHSFB1 در آرابیدوپسیس هم‌خوانی دارد که عملکرد تنظیمی و غشایی دارد. در پروتئین‌های گروه سوم، ساختار سوم تلفیقی از عناصر دو گروه دیگر است؛ در CsHSFA7b.3 دامنه DBD به‌صورت کلاسیک حفظ شده است، در حالی‌که در CsHSFC1a.4 عناصر غشایی، پپتیدهای سیگنال، و صفحات بتا برجسته‌تر هستند. این ساختار تطبیقی با AtHSFC1 در آرابیدوپسیس مشابه است که در پاسخ‌های چندلایه به تنش‌های محیطی نقش دارد. بنابراین، ساختار سه‌بعدیCsHSFها نه‌تنها طبقه‌بندی عملکردی آن‌ها را تأیید می‌کند، بلکه حفظ الگوهای تکاملی در مقایسه با گونه‌های مدل را نیز نشان می‌دهد.



شکل ۴- پیش‌بینی ساختارهای دوم پروتئینی خانواده HSF در *C. sativa*. برنامه PSIPRED برای رسم ساختار دوم استفاده شد.

Figure 4. Secondary structure prediction of the HSF protein family in *C. sativa*. The PSIPRED program was used to predict and visualize the secondary structure.



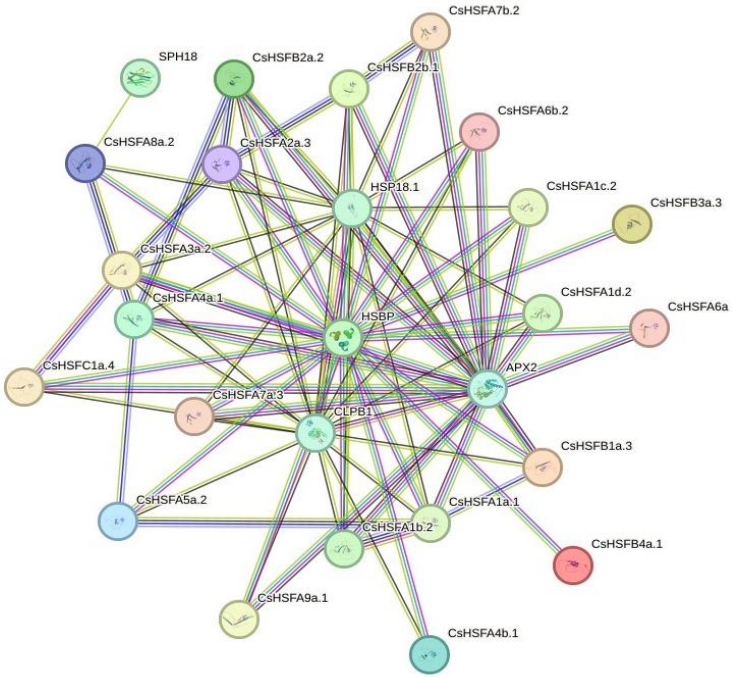
شکل ۵- پیش‌بینی ساختارهای سوم پروتئینی خانواده HSF در *Camelina sativa*. برنامه Phayre 2 بر اساس PDB برای پیش‌بینی ساختار سه‌بعدی پروتئینی استفاده شد.

Figure 5. Tertiary structure prediction of the HSF ҂rotein family in *Camelina sativa.* The Phayre 2 base on PDB program was used to predict and model the tertiary structures of the proteins.

شبکه برهم‌کنش پروتئین–پروتئین بین پروتئین‌های CsHSF

شبکه برهم‌کنش پروتئین–پروتئین بین پروتئین‌های CsHSF با استفاده از پایگاه اطلاعاتی STRING بر اساس همولوژی با گیاه آرابیدوپسیس ترسیم شد (شکل ۶). برهمکنش پروتئین‌ها با خطوط رنگی مختلف نشان داده شده است. تفاوت رنگ بیانگر تفاوت پروفایل بیان و موقعیت سلولی پروتئین‌ها است.

تحلیل شبکه تعاملات عملکردی پروتئین‌های خانواده CsHSF نشان می‌دهد که هر گروه ساختاری دارای الگوی ارتباطی متمایز در پاسخ به تنش حرارتی است. پروتئین‌های CsHSFB1a.3 و CsHSFB2b.2 در گروه دوم، بیشترین تعامل مستقیم را با تنظیم‌کننده‌های پاسخ به گرما مانند HSBP و سایر فاکتورهای مرتبط با بیان ژن‌های HSP دارند. این موقعیت شبکه‌ای نشان‌دهنده‌ی نقش تنظیمی سریع و مستقیم این کلاس در مسیرهای پاسخ به تنش است. در مقابل، پروتئین‌های CsHSFA7b.3 و CsHSFC1a در گروه سوم، در مرکز شبکه قرار گرفته‌اند و با تعداد بیشتری از گره‌ها در ارتباط‌اند؛ این الگوی تعامل پیچیده بیانگر نقش تطبیقی و چندلایه آن‌ها در تنظیم‌های محیطی، تعاملات بین پروتئینی، و مسیرهای هورمونی است. پروتئین‌های CsHSFA4b.1 و CsHSFA5a.1 در گروه اول، تعاملات محدودتری دارند و بیشتر در حاشیه شبکه قرار گرفته‌اند که با نقش تخصصی آن‌ها در تنظیم مستقیم بیان ژن‌های پاسخ به تنش هم‌راستا است. این الگوهای تعامل با مطالعات تطبیقی در *A. thaliana* هم‌خوانی دارند، جایی‌که AtHSFBها در تنظیم سریع پاسخ‌های حرارتی فعال‌اند، AtHSFCها در تنظیم‌های ترکیبی نقش دارند، و AtHSFAها عملکرد کلاسیک رونویسی را حفظ کرده‌اند (Scharf *et al*., 2012). همچنین، بر اساس برهمکنش‌های مشاهده شده، ژن *HSP18* بیشترین تعاملات را با سایر ژن‌ها داشت که بررسی‌ها نشان داد که این پروتئین ممکن است نقش مهمی در فرآیند تمایز جوانه‌های گل و کاهش تنش حرارتی داشته باشد (Tsukaya *et al*., 1993). مطالعه‌ای بر روی نخود فرنگی نشان داده است که یکی از اعضای HSPهای کوچک، HSP18.1، با اتصال به پروتئین‌های دیگر و حفظ آنها در حالتی‌که می‌توانند دوباره تا شوند، از تجمع حرارتی پروتئین‌ها جلوگیری می‌کند (Lee *et al*., 1997). بنابر این، تحلیل شبکه‌ایCsHSFها نه‌تنها تفاوت عملکردی بین کلاس‌ها را آشکار می‌سازد، بلکه حفظ الگوهای تکاملی در تعاملات پروتئینی را نیز نشان می‌دهد.



شکل ۶- شبکه برهم‌کنش پروتئین-پروتئین بین پروتئین‌های CsHSF با استفاده از پایگاه اطلاعاتی STRING ترسیم شد. گره‌های رنگی توپر، پروتئین‌های با ساختار سه‌بعدی شناخته‌شده را نشان می‌دهد.

Figure 6. The protein–protein interaction network of CsHSF in *Camelina sativa* was constructed using the STRING database. Colored solid nodes represent proteins with experimentally determined or modeled three-dimensional structures.

نتیجه‌گیری کلی

میان مسیرهای سیگنالی مرتبط با تنش‌های غیرزیستی و فرایندهای رشد و نمو گیاه، برهم‌کنش‌های پیچیده‌ای وجود دارند که تنظیم هماهنگ بیان ژن را ایجاب می‌کنند. در این میان، فاکتورهای رونویسی شوک حرارتی به عنوان یکی از خانواده‌های ژنی کلیدی، نقش تنظیمی برجسته‌ای در تعدیل پاسخ‌های مولکولی به تنش‌های محیطی همچون گرما، خشکی و شوری ایفا می‌کنند. خانواده HSF نه‌تنها در فعال‌سازی مجموعه‌ای از ژن‌های محافظتی مرتبط با پایداری پروتئین‌ها و حفظ هموستازی سلولی نقش دارد، بلکه در تنظیم فرآیندهای رشد، تکوین اندام‌ها و سازگاری فیزیولوژیک گیاه نیز مشارکت دارد. در گیاه کاملینا که یک ژنوم آلوهگزاپلوئید دارد، تنوع و چندگانگی ژنی حاصل از پلی‌پلوئیدی، فرصت‌های ویژه‌ای را برای تقسیم عملکرد (subfunctionalization) و نوکارکردی شدن (neofunctionalization) اعضای خانواده HSF فراهم کرده است؛ امری که می‌تواند به بهبود پاسخ به تنش‌ها و افزایش تحمل‌پذیری این گونه روغنی نسبت به شرایط نامساعد محیطی منجر شود. در این مطالعه، ۶۴ ژن از خانواده HSF در ژنوم *C. sativa* شناسایی شدند که بر اساس همولوژی در سه گروه اصلی طبقه‌بندی شدند. پروتئین‌های واقع، در زیرگروه‌های مختلف از نظر خصوصیات فیزیکوشیمیایی مانند GRAVY، نقطه ایزوالکتریک از حفاظت‌شدگی بالایی برخوردار است که این امر می‌تواند بیانگر شباهت‌های عملکردی این پروتئین‌ها در زیرگروه‌های مختلف‌شان باشد. ساختار دوم تمامی پروتئین‌ها بیانگر درصد بالای مارپیچ الفا و پیچیدگی کویل‌مانند بود. در تمام این پروتئین‌ها، ۱۵ موتیف حفاظت‌شده با طول بین ۸ تا ۵۰ آمینواسید شناسایی شدند. علاوه بر این، نتایج برهمکنش بیانگر هماهنگی بین این پروتئین‌ها در پاسخ به شرایط مختلف محیطی هستند. شناسایی و تحلیل عملکردی ژن‌های HSF در کامیلنا، چشم‌اندازهای ارزشمندی را برای اصلاح ژنتیکی و به‌کارگیری ابزارهای زیست‌فناورانه فراهم می‌سازند تا بتوان از این ژن‌ها در توسعه ارقام مقاوم به تنش‌های محیطی و بهبود کارایی زراعی استفاده نمود. این نتایج نه‌تنها در درک سازوکارهای مولکولی سازگاری گیاهان نقش دارند، بلکه می‌توانند مبنای طراحی راهبردهای مؤثر در به‌نژادی و بهبود صفات وابسته به تنش در کشاورزی پایدار باشند.

References

Amiri, R., Rostami-Ahmadvandi, H., & Sayyahfar, M. (2023). Evaluation of the Genetic Variation of Oil-Rich Camelina Advanced Lines under Rainfed Conditions. *Journal of Crop Breeding,* 15(47), 152-164. https://doi.org/10.61186/jcb.15.47.152 [In Persian]

Azizi, S., & Zare, N. (2022). Genome-wide identification and functional analysis of lipoxygenase (LOX) gene family in some Fabaceae species using bioinformatics methods. Crop Biotechnology, 11(3), 77-97. https://doi.org/10.30473/CB.2022.64490.1882

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., ... & Noble, W. S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research, 37(suppl_2), W202-W208. https://doi.org/ 10.1093/nar/gkp335

Balanuca, B., Stan, R., Hanganu, A., Lungu, A., & Iovu, H. (2015). Design of new camelina oil-based hydrophilic monomers for novel polymeric materials. *Journal of the American Oil Chemists' Society,* 92(6), 881-891. https://doi.org/ 10.1007/s11746-015-2654-z

Belayneh, H. D., Wehling, R. L., Cahoon, E., & Ciftci, O. N. (2015). Extraction of omega-3-rich oil from Camelina sativa seed using supercritical carbon dioxide. *The Journal of Supercritical Fluids,* 104, 153-159. https://doi.org/ 10.1002/fsn3.572

Bird, K. A., Brock, J. R., Grabowski, P. P., Harder, A. M., Healy, A. L., Shu, S., ... & Kliebenstein, D. J. (2025). Allopolyploidy expanded gene content but not pangenomic variation in the hexaploid oilseed Camelina sativa. *Genetics,* 229(1), 1-44. https://doi.org/10.1093/genetics/iyae183

Blume, R. Y., Hotsuliak, V. Y., Nazarenus, T. J., Cahoon, E. B., & Blume, Y. B. (2024). Genome-wide identification and diversity of FAD2, FAD3 and FAE1 genes in terms of biotechnological importance in Camelina species. *BMC Biotechnology,* 24(1), 107. https://doi.org/10.1186/s12896-024-00936-4

Busch, W., Wunderlich, M., & Schöffl, F. (2005). Identification of novel heat shock factor-dependent genes and biochemical pathways in Arabidopsis thaliana. *The Plant Journal,* 41(1), 1-14. https://doi.org/ 10.1111/j.1365-313X.2004.02272.x

Chaturvedi, S., Bhattacharya, A., Khare, S.K., & Kaushik, G. (2018) Camelina sativa: An Emerging Biofuel Crop. pp, 1-38. In: Hussain C.(eds). Handbook of Environmental Materials Management. *Springer*, *Cham* https://doi.org/10.1007/978-3-319-58538-3_110-1

Chaudhary, R., Higgins, E.E., Eynck, C., Sharpe, A.G. & Parkin, I.A.P. (2023) Mapping QTL for vernalization requirement identified adaptive divergence of the candidate gene *Flowering Locus C* in polyploid *Camelina sativa*. *Plant Genome*, 16(4), e20397. https://doi.org/10.1002/tpg2.20397

Gamage, D. G., Gunaratne, A., Periyannan, G. R., & Russell, T. G. (2019). Applicability of instability index for in vitro protein stability prediction. *Protein and peptide letters,* 26(5), 339-347. https://doi.org/10.2174/0929866526666190228144219.

Garg, A. K., Kim, J. K., Owens, T. G., Ranwala, A. P., Choi, Y. D., Kochian, L. V., & Wu, R. J. (2002). Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proceedings of the National Academy of Sciences,* 99(25), 15898-15903. https://doi.org/ 10.1073/pnas.252637799

Gasteiger, E., Hoogland, Ch., Gattiker, A., Duvaud, S., Wilkins, M.E., Appel, R.D., & Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. The proteomics protocols handbook. *J. M. Walker. New York City, New York, United States, Humana Press,* 571-607. https://doi.org/ 10.1385/1-59259-890-0:571

Gugel, R. K., & Falk, K. C. (2006). Agronomic and seed quality evaluation of Camelina sativa in western Canada. *Canadian journal of plant science,* 86(4), 1047-1058. https://doi.org/ 10.4141/P04-081

Guo, M., Liu, J.H., Ma, X., Luo, D.X., Gong, Z.H., & Lu, M.H. (2016) The plant heat stress transcription factors (HSFs): structure, regulation, and function in response to abiotic stresses. *Frontiers in Plant Science,* 7, 114. https://doi.org/10.3389/fpls.2016.00114

Guo, M., Lu, J. P., Zhai, Y. F., Chai, W. G., Gong, Z. H., & Lu, M. H. (2015). Genome-wide analysis, expression profile of heat shock factor gene family (CaHsfs) and characterisation of CaHsfA2 in pepper (Capsicum annuum L.). BMC plant biology, 15(1), 1-20. https://doi.org/ 10.1186/s12870-015-0512-7

Hahn, A., Bublak, D., Schleiff, E., & Scharf, K.D. (2011) Crosstalk between Hsp90 and Hsp70 chaperones and heat stress transcription factors in tomato. *Plant Cell,* 23, 741–755. https://doi.org/ 10.1105/tpc.110.076018

Haider, S., Rehman, S., Ahmad, Y., Raza, A., Tabassum, J., Javed, T., ... & Mahmood, T. (2021). In silico characterization and expression profiles of heat shock transcription factors (HSFs) in maize (Zea mays L.). *Agronomy,* 11(11), 2335. https://doi.org/ 10.3390/agronomy11112335

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., & Nakai, K. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Research,* 35(suppl_2), W585-W587. https://doi.org/ 10.1093/nar/gkm259

Hu, Y., Han, Y. T., Wei, W., Li, Y. J., Zhang, K., Gao, Y. R., ... & Feng, J. Y. (2015). Identification, isolation, and expression analysis of heat shock transcription factors in the diploid woodland strawberry Fragaria vesca. *Frontiers in Plant Science,* 6, 736. https://doi.org/ 10.3389/fpls.2015.00736

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C., ... & Hunter, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics,* 30(9), 1236-1240. https://doi.org/ 10.1093/bioinformatics/btu031

Kim, T., Samraj, S., Jiménez, J., Gómez, C., Liu, T., & Begcy, K. (2021). Genome-wide identification of heat shock factors and heat shock proteins in response to UV and high intensity light stress in lettuce. *BMC Plant Biology,* 21(1), 185. https://doi.org/ 10.1186/s12870-021-02959-x

Lämke, J., Brzezinka, K., Altmann, S., & Bäurle, I. (2016). A hit-and-run heat shock factor governs sustained histone methylation and transcriptional stress memory. *The EMBO Journal*, *35*(2), 162-175. https://doi.org/ 10.15252/embj.201592593

Lee, G. J., Roseman, A. M., Saibil, H. R., & Vierling, E. (1997). A small heat shock protein stably binds heat-denatured model substrates and can maintain a substrate in a folding-competent state. *The EMBO Journal*, *16*(3), 659-671.

Lin, Y.X., Jiang, H.Y., Chu, Z.X., Tang, X.L., Zhu, S.W., & Cheng, B.J. (2011) Genome-wide identification, classification and analysis of heat shock transcription factor family in maize. *BMC Genomics*, 12, 76. https://doi.org/10.1186/1471-2164-12-76

Liu, Y., Zhang, C., Chen, J., Guo, L., Li, X., Li, W., ... & Zhang, L. (2013). Arabidopsis heat shock factor HsfA1a directly senses heat stress, pH changes, and hydrogen peroxide via the engagement of redox state. *Plant Physiology and Biochemistry,* 64, 92-98. https://doi.org/ 10.1016/j.plaphy.2012.12.013.

McVay, K. A., & Khan, Q. A. (2011). Camelina yield response to different plant populations under dryland conditions. *Agronomy Journal*, *103*(4), 1265-1269. https://doi.org/ 10.2134/agronj2011.0057

Mishra, S. K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., & Scharf, K. D. (2002). In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. *Genes & development,* 16(12), 1555-1567. https://doi.org/10.1101/gad.228802

Mishra, S. K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., & Scharf, K. D. (2002). In the complex family of heat stress transcription factors, HsfA1 has a unique role as master regulator of thermotolerance in tomato. *Genes & Development,* 16(12), 1555-1567. https://doi.org/ 10.1101/gad.228802

Moore, R. C., & Purugganan, M. D. (2003). The early stages of duplicate gene evolution. *Proceedings of the National Academy of Sciences*, *100*(26), 15682-15687. https://doi.org/ 10.1073/pnas.2535513100.

Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K., & Shigeoka, S. (2006). Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *The Plant Journal*, *48*(4), 535-547. https://doi.org/ 10.1111/j.1365-313X.2006.02889.x

Nover, L., Bharti, K., Doring, P., Mishra, S.K., Ganguli, A., & Scharf, K.D. (2001) Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? *Cell Stress Chaperones,* 6, 177–189. https://doi.org/ 10.1379/1466-1268

Pandey, G. K., Kanwar, P., Singh, A., Steinhorst, L., Pandey, A., Yadav, A. K., ... & Luan, S. (2015). Calcineurin B-like protein-interacting protein kinase CIPK21 regulates osmotic and salt stress responses in Arabidopsis. *Plant Physiology,* 169(1), 780-792. https://doi.org/ 10.1104/pp.15.00623.

Rao, P. K., Roxas, B. A., & Li, Q. (2008). Determination of global protein turnover in stressed mycobacterium cells using hybrid-linear ion trap-fourier transform mass spectrometry. *Analytical Chemistry,* 80(2), 396-406. https://doi.org/ 10.1021/ac701690d.

Ren, Y., Ma, R., Xie, M., Fan, Y., Feng, L., Chen, L., ... & Wang, B. (2023). Genome-wide identification, phylogenetic and expression pattern analysis of HSF family genes in the Rye (*Secale cereale* L.). *BMC Plant Biology,* 23(1), 441. https://doi.org/ 10.1186/s12870-023-04418-1

Rezaian joybari, N., Majidian, P., Majidian, P., & Ghorbani, H.R. (2024). Investigating the Effect of Iron Nanoparticle and Putrescine Stimulants on some Functional and Physiological Traits of Camelina (*Camelina Sativa*). *Journal of Crop Breeding*. *15*(48), 178-188. https://doi.org/10.61186/jcb.15.48.178 [In Persian]

Scharf, K.D., Berberich, T., Ebersberger, I., & Nover, L. (2012) The plant heat stress transcription factor (Hsf) family: structure, function and evolution. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1819(2), 104–119. https://doi.org/10.1016/j.bbagrm.2011.10.002

Scherf, U., Ross, D. T., Waltham, M., Smith, L. H., Lee, J. K., Tanabe, L., ... & Weinstein, J. N. (2000). A gene expression database for the molecular pharmacology of cancer. *Nature Genetics,* 24(3), 236-244. https://doi.org/10.1038/73439.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., & Lohmann, J.U. (2005) A gene expression map of Arabidopsis thaliana development. *Nature Genetics,* 37(5), 501-506. https://doi.org/ 10.1038/ng1543

Shamshad, A., Rashid, M., & Zaman, Q. (2023) In-silico analysis of heat shock transcription factor *(OsHSF)* gene family in rice *(Oryza sativa* L.*). BMC Plant Biology,* 23, 395. https://doi.org/10.1186/s12870-023-04399-1

Singh, R. K., Jaishankar, J., Muthamilarasan, M., Shweta, S., Dangi, A., & Prasad, M. (2016). Genome-wide analysis of heat shock proteins in C4 model, foxtail millet identifies potential candidates for crop improvement under abiotic stress. *Scientific Reports,* 6(1), 32641. https://doi.org/ 10.1038/srep32641.

Sultan, S., Ali, M., Nawaz, S., Ali, M.A., & Shahzad, A. (2016) Genome wide analysis of Heat Shock Factors (*HSF*) gene family of *Arabidopsis thaliana*. *Journal of Biology, Agriculture and Healthcare,* 6(1), 33–40.

Swindell, W. R., Huebner, M., & Weber, A. P. (2007). Transcriptional profiling of Arabidopsis heat shock proteins and transcription factors reveals extensive overlap between heat and non-heat stress response pathways. *BMC Genomics,* 8(1), 125. https://doi.org/10.1186/1471-2164-8-125.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution,* 30(12), 2725-2729.

Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2003). Multiple sequence alignment using ClustalW and ClustalX. *Current protocols in bioinformatics,* (1), 2-3. https://doi.org/ 10.1093/molbev/mst197

Tsukaya, H., Takahashi, T., Naito, S., & Komeda, Y. (1993). Floral organ-specific and constitutive expression of an Arabidopsis thaliana heat-shock HSP18. 2:: GUS fusion gene is retained even after homeotic conversion of flowers by mutation. *Molecular and General Genetics MGG,* 237(1), 26-32. https://doi.org/10.1007/BF00282780

Wang, N., Shu, X., Zhang, F., Song, G., & Wang, Z. (2024) Characterization of the Heat Shock Transcription Factor Family in *Lycoris radiata* and Its Potential Roles in Response to Abiotic Stresses. *Plants (Basel).*13(2), 271. https://doi.org/ 10.3390/plants13020271.

Xu, P., Gua, Q., Pang, X., Zhang, P., Kong, D., & Liu, J. (2020) New Insights into Evolution of Plant Heat Shock Factors (Hsfs) and Expression Analysis of Tea Genes in Response to Abiotic Stresses. *Plant*, 9(3), 311. https://doi.org/10.3390/plants9030311

Yang, Q.Q., Yang, F., Liu, C.Y., Zhao, Y.Q., Lu, X.J., Ge, J., Zhang, B.W., Li, M.Q., Yang, Y., & Fan, J.D. (2024) Genome-wide analysis of the HSF family in *Allium sativum* L. and AsHSFB1 overexpression in Arabidopsis under heat stress. *BMC Genomics,* 25(1), 1072. https://doi.org/10.1186/s12864-024-11002-w.

Ye, J., Yang, X., Hu, G., Liu, Q., Li, W., Zhang, L., & Song, X. (2020). Genome-wide investigation of heat shock transcription factor family in wheat (Triticum aestivum L.) and possible roles in anther development. *International Journal of Molecular Sciences,* 21(2), 608. https://doi.org/10.3390/ijms21020608

Yokotani, N., Ichikawa, T., Kondou, Y., Matsui, M., Hirochika, H., Iwabuchi, M., & Oda, K. (2008) Expression of rice heat stress transcription factor OsHsfA2e enhances tolerance to environmental stresses in transgenic Arabidopsis. *Planta*. 227(5), 957-967. https://doi.org/ 10.1007/s00425-007-0670-4

Yoshida, T., Ohama, N., Nakajima, J., Kidokoro, S., Mizoi, J., Nakashima, K., ... & Yamaguchi-Shinozaki, K. (2011). Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Molecular Genetics and Genomics,* 286(5), 321-332. https://doi.org/ 10.1007/s00438-011-0647-7

Younesi-Melerdi, E., Nematzadeh, G., & Pakdin-Parizi, A. (2020). Expression analysis of some genes involved in signaling networks of Aeluropus littoralis (Gouan) Parl. under salinity stress. *Environmental Stresses in Crop Sciences,* 13(4), 1259-1270. https://doi.org/ 10.22077/ESCS.2020.2443.1646

Yun, L., Zhang, Y., Li, S., Yang, J., Wang, C., Zheng, L., ... & Gao, J. (2023). Phylogenetic and expression analyses of HSF gene families in wheat (Triticum aestivum L.) and characterization of TaHSFB4-2B under abiotic stress. *Frontiers in Plant Science,* 13, 1047400. https://doi.org/10.3389/fpls.2022.1047400

Zafar, S. A., Hussain, M., Raza, M., Muhu-Din Ahmed, H. G., Rana, I. A., Sadia, B., & Atif, R. M. (2016). Genome wide analysis of heat shock transcription factor (HSF) family in chickpea and its comparison with Arabidopsis. *Plant Omics,* 9(2), 136-141.

Zhang, J., Li, J., Liu, B., Zhang, L., Chen, J., & Lu, M. (2013) Genome-wide analysis of the Populus Hsp90 gene family reveals differential expression patterns, localization, and heat stress responses. BMC Genomics, 14(1), 532. https://doi.org/ 10.1186/1471-2164-14-532

Zhou, M., Zheng, S., Liu, R., Lu, J., Lu, L., Zhang, Ch., Liu, Z., Luo, C., Zhang, L., Yant, L., & Wu, Y. (2019). Genome-wide identification, phylogenetic and expression analysis of the heat shock transcription factor family in bread wheat (*Triticum aestivum* L.). *BMC Genom,* 20, 505. https://doi.org/10.1186/s12864-019-5876-x