



Genome wide identification and characterization of strictosidine synthase-like (*SSL*) genes in wheat (*Triticum aestivum* L.)

Mohammad Hossein Rezadoost^{1*} and Amin Abedi²

1. Assistant Professor, Department of Agricultural Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran (*Corresponding author: rezadoostmh@guilan.ac.ir)
2. Graduate Ph.D., Department of Agricultural Biotechnology, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran

Comprehensive abstract

Introduction

Wheat (*Triticum aestivum* L.) is globally recognized as a crucial food crop. Due to the significant increase in human population in recent years, there is a need for food production to match population growth, especially for primary crops like wheat. Given the decline in wheat cultivation and yield caused by biotic and abiotic stresses, the cultivation of stress-resistant varieties is a cost-effective and fundamental strategy to mitigate their adverse impacts. Identifying resistance genes is essential for developing new resistant varieties using breeding programs. Strictosidine synthase-like (*SSL*) genes with a length of approximately 400 amino acids, play a role in plant immunity regulation and possess an extracellular structural domain resembling animal hemomyosin. Previous studies have shown that all categories of *AtSSL* genes exhibit a response to various biotic and abiotic stressors. At present, our major understanding of the *SSL* gene family in plants is primarily based on research conducted on *Arabidopsis thaliana*. In this study, bioinformatics tools were used to explore the evolutionary relationships and functional roles of the *SSL* gene family in wheat.

Materials and methods

In the first step, the sequence of *SSL* proteins from rice and *Arabidopsis* was used to identify genes encoding wheat *SSL* in the Ensembl Plants database by the PlastP algorithm. Next, phylogenetic relationships were analyzed by MEGA7, the exon-intron structure and intron phase using the Gene Structure view in TBtools-II, and conserved motifs with Multiple Em for Motif Elicitation. Additionally, *cis*-regulatory elements in the promoter region, gene duplication events, and selection pressure were investigated through PlantCare and the Simple Ka/Ks Calculator in TBtools-II. The expression profiles of *TaSSL* genes in response to abiotic stresses were analyzed using the expVIP server. All analyses were conducted using default parameters of the software and servers.

Research findings

This study identified 69 *SSL* genes in the wheat genome, exhibiting a non-uniform distribution across chromosomes. The evolutionary study of this family revealed two main phylogenetic groups in the *SSLs* of different organisms: The first group (I), encompassing exclusively wheat and rice *SSLs* genes, and the second group (II), containing genes from diverse organisms. Further subdivision of group II into three subgroups (A, B, and C) highlighted potential functional divergence among members. The analysis of conserved motifs, gene structure, and intron phase indicated a high degree of conservation for these genes. Furthermore, segmental duplication emerged as the primary driver of wheat *SSL* gene expansion, and these duplicated genes experiencing strong negative selection pressure. The presence of *cis*-regulatory elements responsive to hormones and stresses suggests



intricate regulation of *TaSSL* gene expression. Consistent with this notion, RNA-seq data revealed the inducible expression of *TaSSL* genes in response to abiotic stresses, including cold, heat, drought, and PEG.

Conclusion

The presence of a distinct evolutionary cluster of wheat *SSL* genes, characterized by features typically associated with stress-responsive genes such as a low number of introns, the application of negative selection pressure, the presence of regulatory elements responsive to stresses and hormones, as well as the expression patterns of *TaSSL* genes in response to abiotic stresses indicated their significant role in wheat's stress response mechanisms. Consequently, the findings of this study can provide valuable insights into the functions of *TaSSL* genes, facilitating the identification of potential candidates for producing stress-resistant wheat varieties in future breeding programs.

Keywords: Abiotic stress, Duplication, Evolution, Gene function, *In silico*

Received: October 24, 2023

Accepted: December 08, 2023

Cite this article:

Rezadoost, M. H., & Abedi, A. (2023). Genome wide identification and characterization of strictosidine synthase-like (*SSL*) genes in wheat (*Triticum aestivum* L.). *Cereal Research*, 13(3), 249-268. doi: [10.22124/CR.2024.26392.1805](https://doi.org/10.22124/CR.2024.26392.1805).



شناسایی و مطالعه ویژگی‌های خانواده ژنی شبه استریکتوسیدین سینتاز (SSL) در گستره ژنوم گندم (*Triticum aestivum* L.)

محمدحسین رضادوست^{۱*} و امین عابدی^۲

۱- استادیار، گروه بیوتکنولوژی کشاورزی، دانشکده علوم کشاورزی، دانشگاه گیلان، رشت ایران (* نویسنده مسئول):

rezadoostmh@guilan.ac.ir

۲- دانش‌آموخته دکتری، گروه بیوتکنولوژی کشاورزی، دانشکده علوم کشاورزی، دانشگاه گیلان، رشت، ایران

چکیده جامع

مقدمه: گندم (*Triticum aestivum* L.) در سطح جهانی به‌عنوان یک محصول غذایی مهم شناخته شده است. با توجه به افزایش قابل توجه جمعیت، نیاز به تولید مواد غذایی متناسب با رشد جمعیت، به‌ویژه برای محصولات اصلی مانند گندم وجود دارد. با توجه به کاهش کشت و عملکرد گندم در اثر تنش‌های زیستی و غیرزیستی، کشت رقم‌های مقاوم به تنش‌ها مقرون به صرفه‌ترین و اساسی‌ترین رویکرد برای خنثی کردن اثرات منفی این عوامل است. برای دستیابی به این هدف، شناسایی ژن‌های مقاوم به‌عنوان زیربنای ایجاد رقم‌های مقاوم جدید از طریق برنامه‌های به‌نژادی ضروری است. ژن‌های شبه استریکتوسیدین سینتاز (SSL)، دُمین ساختاری خارج سلولی مشابه با همومیوسین جانوری داشته و با طولی در حدود ۴۰۰ اسید آمینه، در تنظیم ایمنی گیاهان نقش دارند. مطالعات قبلی نشان داده‌اند که همه ژن‌های *AtSSL* در پاسخ به تنش‌های مختلف زیستی و غیرزیستی نقش دارند. با این حال، عمده درک ما از خانواده ژنی SSL در گیاهان فقط بر اساس تحقیقات انجام شده روی آرآبیدوپسیس تالیانا به‌دست آمده است. در این مطالعه، از ابزارهای بیوانفورماتیک به‌منظور رمزگشایی روابط تکاملی و نقش کارکردی خانواده ژنی SSL در گندم استفاده شد.

مواد و روش‌ها: در اولین مرحله، توالی پروتئین‌های SSL برنج و آرآبیدوپسیس از طریق الگوریتم PlastP به‌منظور شناسایی ژن‌های کد کننده SSL گندم در پایگاه داده Ensembl Plants مورد استفاده قرار گرفت. در ادامه روابط فیلوژنتیک با MEGA7، ساختار اگزون-اینترون و فاز اینترونی با TBtools-II in Gene Structure view و موتیف‌های حفاظت شده با Multiple Em for Motif Elicitation تجزیه و تحلیل شد. همچنین، عناصر تنظیمی *Cis* ناحیه پرموتری، مضاعف‌شدگی ژن‌ها و فشار گزینش به‌ترتیب با استفاده از PlantCare و TBtools-II Simple Ka/Ks Calculator مورد مطالعه قرار گرفت. بررسی پروفایل بیان ژن‌های *TaSSL* در پاسخ به تنش‌های غیرزیستی نیز با استفاده از سرور expVIP انجام گرفت. در تمامی تجزیه‌ها از پارامترهای پیش‌فرض استفاده شد.

یافته‌های تحقیق: بر اساس نتایج این مطالعه، ۶۹ ژن SSL در گستره ژنوم گندم یافت شد که به‌صورت غیر یکنواخت روی کروموزوم‌های متفاوت قرار گرفته بودند. مطالعه تکاملی این خانواده نشان داد که SSL‌های موجودات مختلف در دو گروه فیلوژنتیک اصلی قرار می‌گیرند. گروه اول (I)، فقط شامل ژن‌های SSL گندم و برنج بود و در گروه دوم (II)، ژن‌های موجودات مختلف قرار گرفتند. گروه دوم خود به سه زیرگروه A، B و C تقسیم شد و اعضای هر گروه از نظر کارکرد می‌توانند متفاوت

باشند. تجزیه و تحلیل موتیف‌های حفاظت‌شده، ساختار ژنی و فاز اینترونی، حفاظت‌شدگی بالای این ژن‌ها را نشان داد. از سوی دیگر، مضاعف‌شدگی سگمنتال عامل اصلی بسط ژن‌های *SSL* گندم بود و ژن‌های مضاعف‌شده تحت فشار گزینش منفی شدید قرار داشتند. وجود عناصر تنظیمی *Cis* پاسخ به هورمون‌ها و تنش‌ها نشان می‌دهد که تنظیم بیان ژن‌های *TaSSL* بسیار پیچیده است. از سوی دیگر، بررسی پروفایل بیانی ژن‌های *TaSSL* در پاسخ به تنش‌های غیرزیستی بر اساس داده‌های RNA-seq نشان داد که این ژن‌ها در اثر تنش‌های سرما، گرما، خشکی و PEG القاء می‌شوند که در راستای وجود عناصر تنظیمی *Cis* پاسخ به هورمون‌ها و تنش‌ها هستند.

نتیجه‌گیری: وجود گروه تکاملی اختصاصی ژن‌های *SSL* گندم در کنار ویژگی‌های مربوط به ژن‌های پاسخ به تنش مانند تعداد کم اینترون، فشار گزینش منفی اعمال شده بر آن‌ها، وجود عناصر تنظیمی پاسخ به تنش‌ها و هورمون‌ها و نیز الگوی بیان ژن‌های *TaSSL* در پاسخ به تنش‌های غیرزیستی نشان داد که این ژن‌ها نقش مهمی در پاسخ گندم به تنش‌ها دارند. یافته‌های این تحقیق می‌تواند به درک نقش ژن‌های *TaSSL* کمک کند و منجر به شناسایی ژن‌های بالقوه به‌منظور ایجاد رقم‌های گندم مقاوم به تنش در برنامه‌های به‌نژادی آینده شود.

واژه‌های کلیدی: این سیلیکو، تکامل، تنش غیرزیستی، کارکرد ژن، مضاعف‌شدگی

تاریخ دریافت: ۱۴۰۲/۰۸/۰۲

تاریخ پذیرش: ۱۴۰۲/۰۹/۱۷

نحوه استناد به این مقاله:

رضادوست، محمدحسین، و عابدی، امین. (۱۴۰۲). شناسایی و مطالعه ویژگی‌های خانواده ژنی شبه استریکتوسیدین سینتاز (*SSL*) در گستره ژنوم گندم (*Triticum aestivum* L.). *تحقیقات غلات*، ۱۳ (۳)، ۲۴۸-۲۴۹. doi: [10.22124/CR.2024.26392.1805](https://doi.org/10.22124/CR.2024.26392.1805)

مقدمه

متابولیت‌های ثانویه نقش مهمی در بقای گیاه و ایجاد ارتباطات اکولوژیکی بین گونه‌های دیگر دارند. این ترکیبات علاوه بر اینکه برای انسان بسیار با ارزش هستند، به محافظت از گیاهان در برابر پاتوژن‌ها و تنش‌های غیرزیستی نیز کمک می‌کنند (Jan et al., 2021). آلکالوئیدها از مهم‌ترین متابولیت‌های ثانویه گیاهان بوده و نقش بسیار مهمی در مقاومت گیاهان به تنش‌های زیستی و غیرزیستی دارند (Wang et al., 2023a). استریکتوسیدین پیش‌ساز بسیاری از آلکالوئیدها بوده و از چهل سال پیش در بسیاری از گیاهان شناسایی شده است. آنزیم استریکتوسیدین سینتاز (EC 4.3.3.2) (STR) تولید استریکتوسیدین از ترکیب تریپتامین و سکلوگانین را کاتالیز می‌کند. ژن کد کننده این آنزیم برای اولین بار از آسرول (*Rauvolfia serpentine*) جداسازی شد (Kutchan et al., 1988; Yamazaki et al., 2003). آنزیم STR متعلق به گروه آنزیم‌های Nucleophilic attack six-bladed β -propeller (N6P) است. این آنزیم‌ها دارای ویژگی کاتالیتیک مشترک حمله نوکلئوفیلی Hicks et al., (2013).

در برخی از پروتئین‌های گیاهان، باکتری‌ها، حشرات و حتی پستانداران، دمین Str_synth (دمین کارکردی آنزیم‌های استریکتوسیدین سینتاز PF03088) شناسایی شده است. بنابراین این پروتئین‌ها دارای رابطه تکاملی با ژن STR هستند. به‌عنوان مثال همومیوسین مگس سرکه، یک گلیکوپروتئین متصل به غشاء بوده و از یک ناحیه کوتاه درون سلولی و دو ناحیه خارج تشکیل شده است که شامل یک دمین مشابه با پروتئین‌های SS گیاهی و یک دمین مشابه به میوسین در انسان بوده و در ایمنی مگس سرکه نقش دارد (Theopold et al., 1996; Fabbri et al., 2000). باید توجه داشت وجود ژن‌های همولوگ STR در ژنوم گیاهان دلیلی بر توانایی آن‌ها در تولید ترکیبات آلکالوئیدی پیچیده نیست. بعنوان مثال آرابیدوپسیس توانایی تولید ترکیبات آلکالوئیدی پیچیده را ندارد با این حال توالی‌یابی ژنوم این گیاه نشان داد که دارای ۱۵ ژن حاوی دمین Str_synth می‌باشد (Facchini et al., 2004). آنزیم STR کد شده توسط این ژن‌ها توانایی کاتالیز واکنش ترکیب تریپتامین و سکلوگانین و تولید استریکتوسیدین را ندارد و فقط

سکلوگانین را به‌صورت دایمر متابولیزه می‌کند که کارکرد این دایمر مشخص نیست. ژن‌های با نقش کارکردی نامشخص که دارای رابطه همولوژی با ژن STR هستند، شبه استریکتوسیدین سینتاز (Strictosidine Synthase-like, SSL) نامیده می‌شوند (Kibble et al., 2009). مطالعه بیش از ۵۰۰ توالی پروتئین SSL موجود در بانک‌های اطلاعاتی نشان داد که این پروتئین‌ها توانایی انجام واکنش اصلی را ندارند، با این حال واکنش هیدرولیتیک را کاتالیز می‌کنند (Hicks et al., 2011).

مطالعات کارکردی ژن‌های SSL نشان دهنده نقش این ژن‌ها در بسیاری از فرآیندهای زیستی گیاهان مانند پاسخ به تنش‌ها می‌باشد. به‌عنوان مثال ژن‌های YLS2 و LAP3 آرابیدوپسیس که به‌دلیل داشتن دمین Str_synth همولوگ آنزیم استریکتوسیدین سینتاز هستند، به‌ترتیب در پاسخ به پیری و نمو گرده نقش دارند (Dobritsa et al., 2009; Yoshida et al., 2001). بیان ژن YLS2 طی پیری و نیز پاسخ به اتیلن، آبسزیک اسید و تاریکی القا می‌شود (Yoshida et al., 2001). از طرفی خاموشی ژن LAP3 باعث نرغیمی در آرابیدوپسیس می‌شود (Dobritsa et al., 2009). ژن OsSTR2L2 برنج کارکرد مشابه با ژن LAP3 آرابیدوپسیس دارد و در نمو بساک و تشکیل دیواره دانه گرده نقش دارد و نقص در این ژن موجب نرغیمی می‌شود (Zou et al., 2017). بیان ژن‌های AT3G51430، AT3G51440 و AT3G51450 آرابیدوپسیس که همولوگ استریکتوسیدین سینتاز هستند در پاسخ به تنش‌های زیستی و غیرزیستی و نیز تیمارهای سالیسیلیک اسید، جاسمونیک اسید و اتیلن به‌صورت معنی‌داری افزایش می‌یابد (Sohani et al., 2009). در گوجه فرنگی miR1916 نقش موثری در تنظیم بیان پس از رونویسی ژن STR دارد و سرکوب بیان این مولکول miRNA موجب القای بیان ژن‌های STR2، UGT و MYB12 و افزایش مقاومت گوجه فرنگی به پاتوژن‌های *Botrytis cinerea* و *Phytophthora infestans* می‌شود (Chen et al., 2019). نتایج مطالعه ژن‌های SSL صنوبر کالیفرنایی (*Populus trichocarpa*) نشان داد که بیان ژن‌های PtrSSL2، PtrSSL6، PtrSSL8، PtrSSL10 و PtrSSL12 در پاسخ به تنش‌های خشکی و سطوح بالای شوری القاء می‌شود. برخی از ژن‌های PtrSSL نیز بیان بالایی در ساقه و برگ دارند (Wang et al., 2023a).

خانواده ژنی SSL در برخی از گیاهان مانند آرابیدوپسیس (Sohani et al., 2009)، برنج (Zou et al., 2017)، ذرت (Gu et al., 2023) و صنوبر کالیفرنایی (Wang et al., 2023a) مطالعه شده است. گندم (*Triticum aestivum* L., $2n = 6x = 42$) (AABBDD) منبع اصلی تغذیه برای حدود ۴۰ درصد از جمعیت جهان است و با بالاترین سطح زیر کشت نسبت به همه گیاهان زراعی (بیش از ۲۱۸ میلیون هکتار)، بیشترین تجارت جهانی را از مجموع تمامی محصولات دیگر دارد. گندم از جایگاه مهمی در تغذیه انسان برخوردار است، به طوری که ۲۰ درصد پروتئین و کالری روزانه را تامین می‌کند. از نظر امنیت غذایی، این محصول پس از برنج، دومین محصول مهم غذایی در کشورهای در حال توسعه است (Giraldo et al., 2019). با این حال تنش‌های زیستی و غیرزیستی بر کیفیت و عملکرد گندم تأثیر جدی دارند. از این رو کشف ژن‌هایی که در مقاومت به تنش‌ها نقش دارند، به عنوان منابع ژنی موثر برای اصلاح مقاومت گندم بسیار مهم هستند (Yuan et al., 2023). با وجود گزارش‌های متعدد بر نقش ژن‌های SSL در پاسخ به تنش‌ها، تا کنون این خانواده ژنی در گندم مطالعه نشده است. بر این اساس در مطالعه حاضر شناسایی و بررسی سیستماتیک خانواده ژنی SSL گندم انجام شد.

مواد و روش‌ها

جستجوی پایگاه داده و شناسایی ژن‌های کد کننده SSL گندم

در مرحله نخست، توالی پروتئینی خانواده ژنی SSL آرابیدوپسیس و برنج به عنوان مهم‌ترین گیاهان مدل دولپه و تک‌لپه از پایگاه داده plant_ensembl (<https://plants.ensembl.org>) دریافت و در ادامه توالی‌های دریافتی از طریق الگوریتم BlastP برای جستجوی ژن‌های SSL گندم در plant_ensembl استفاده شد (Bolser et al., 2016). بعد از دریافت نتایج BlastP، توالی‌های تکراری حذف و بررسی وجود دمین کارکردی Str_synth از طریق سرور SMART (<http://smart.embl-heidelberg.de>) انجام (Letunic & Bork, 2018) و توالی‌های فاقد دمین‌های کارکردی حذف شدند. آنالیز بیوانفورماتیکی مشخصات عمومی پروتئین‌ها شامل طول پروتئین، وزن مولکولی و نقطه ایزوالکتریک با استفاده از سرور ProtParam

هم‌ردیف‌سازی و رسم درخت فیلوژنتیک

طول کامل توالی پروتئین‌های SSL گندم، آرابیدوپسیس (۱۵ ژن)، برنج (۲۴ ژن)، همومیوسین مگس سرکه و انسان، و STR پرپوش (*Catharanthus roseus*) و آسرول با استفاده از سرور Clustal Omega (<https://www.ebi.ac.uk/jdispatcher/msa/clustalo>) هم‌ردیف‌سازی شد (Sievers & Higgins, 2014). سپس نتیجه هم‌ردیف‌سازی برای رسم درخت فیلوژنتیک استفاده و درخت فیلوژنتیک با استفاده از نرم‌افزار MEGA7 و روش نزدیک‌ترین همسایه‌ها (Neighbor joining) رسم شد. به منظور بررسی صحت رسم شده از آزمون بوت‌استرپ با هزار تکرار استفاده شد (Kumar et al., 2016).

بررسی ساختار ژنی، فاز اسپلایسینگ و موتیف‌های

حفاظت شده

ساختار اگزون-اینترون و فاز اسپلایسینگ خانواده ژنی SSL گندم با استفاده از ابزار Gene Structure view نرم‌افزار TBtools-II مطالعه شد (Chen et al., 2023). این آنالیز به دو فایل GFF3 ژنوم گندم و درخت فیلوژنتیک ژن‌های SSL در شکل newich نیاز دارد که فایل اول از سرور biomart دریافت و درخت فیلوژنتیک با استفاده از نرم‌افزار MEGA7 ایجاد شد (Kumar et al., 2016). موتیف‌های حفاظت در توالی پروتئینی SSL گندم با استفاده از سرور MEME (<https://meme-suite.org/meme>) ارزیابی شد (Bailey et al., 2009). پارامترهای این آنالیز شامل طول موتیف بین پنج تا ۵۰ اسید آمینه و حداکثر تعداد موتیف‌های شناسایی شده روی ۲۰ تنظیم شد. نتایج با استفاده از نرم‌افزار TBtools-II ارائه شد (Chen et al., 2023).

آنالیز ناحیه پروموتور و شناسایی عناصر تنظیمی Cis

جهت شناسایی عناصر سیس ناحیه پروموتور ژن‌های SSL گندم، ابتدا ۱۵۰۰ جفت باز ناحیه بالا دست کدون شروع به عنوان ناحیه پروموتوری ژن‌های هدف، از سرور biomart دریافت و سپس با سرور PlantCARE

مکانیسم مضاعف‌شدگی و فشار گزینش

برای شناسایی ژن‌های حاصل از مضاعف‌شدگی سگمنتال (Segmental) دو پیش شرط در نظر گرفته شد که شامل حداقل ۹۰ درصد انطباق بین دو ژن و پوشش ۹۰ درصد ناحیه هم‌ردیف شده به وسیله ژن با توالی کوتاه بود. برای شناسایی ژن‌های حاصل از مضاعف‌شدگی پی در پی (Tandem) جایگاه ژن‌ها روی کروموزوم‌های گندم بررسی و ژن‌هایی که روی یک کروموزوم قرار داشتند (فاصله دو ژن کم‌تر از ۱۰۰ کیلو جفت باز بود و همچنین دو شرط اشاره شده بالا را داشتند) به‌عنوان ژن‌های ایجاد شده از مکانیسم مضاعف‌شدگی تاندم در نظر گرفته شدند. با استفاده از ابزار Simple Ka/Ks Calculator نرم‌افزار TBtools-II نرخ جهش‌های هم‌معنی (Ks) و دگرمعنی (Ka) بین ژن‌های مضاعف شده محاسبه شد (Chen *et al.*, 2023). در ادامه با استفاده از نسبت Ka/Ks فشار گزینش محاسبه شد. بر این اساس، اگر نسبت Ka/Ks بیش‌تر از یک باشد، ژن دارای فشار گزینش مثبت است و اگر این نسبت کم‌تر از یک و برابر با یک باشد، به‌ترتیب بیانگر فشار گزینش منفی و طبیعی است (Wang *et al.*, 2023b).

بررسی بیان ژن‌های SSL با استفاده از داده‌های RNA-seq

از داده‌های RNA-seq پایگاه داده ExpVIP (<http://www.wheat-expression.com>) برای ارزیابی بیان ژن‌های کد کننده SSL گندم در پاسخ به تنش‌های سرما، خشکی، گرما، خشکی/گرما و پلی‌اتیلن گلایکول (PEG) استفاده شد (Borrill *et al.*, 2016). ژن‌هایی که حداقل در یک نمونه دارای بیان $TPM \geq 1$ (Transcripts Per Million) بودند، به‌عنوان ژن‌های دارای بیان انتخاب شدند (Yoshioka *et al.*, 2022). داده‌های دریافت شده به‌صورت $\log_2(TPM+1)$ تبدیل و در ادامه با استفاده از ابزار HeatMap Illustrator نرم‌افزار TBtools-II نقشه حرارتی داده‌های بیانی رسم شد (Chen *et al.*, 2023).

نتایج و بحث

شناسایی خانواده ژنتیکی SSL گندم

گیاهان بر اساس وجود طیف وسیعی از ژن‌های ساختاری و پاسخ دهنده به تنش‌ها در ژنوم خود، قادر به

حیات و پاسخ به شرایط متغیر محیطی هستند. از جمله این ژن‌ها می‌توان به خانواده SSL اشاره کرد که نقش مهمی در پاسخ به تنش‌های زیستی و غیرزیستی دارند. در این مطالعه با استفاده از الگوریتم BlastP، ۶۹ ژن کد کننده SSL در گستره ژنوم گندم شناسایی شد (جدول ۱). این ژن‌ها به‌صورت غیریکنواخت روی کروموزوم‌های زیرژنوم A، B و D قرار گرفتند، به‌طوری که زیرژنوم‌های A، B و D به‌ترتیب داری ۱۸، ۲۱ و ۲۵ ژن SSL بودند. پنج ژن نیز به‌علت مشخص نبودن کروموزوم در اسکافولد قرار گرفتند. به‌جز کروموزوم‌های 6A، 6B، 1D و 6D که فاقد ژن SSL بودند، سایر کروموزوم‌ها حداقل دارای یکی از اعضای این خانواده ژنی بودند (جدول ۱).

با مطالعه ویژگی‌های فیزیوشیمیایی پروتئین‌های SSL گندم مشخص شد که طول این پروتئین‌ها با میانگین ۳۲۸،۶، از ۱۸۶ اسید آمینه در SSL34 تا ۴۱۴ اسید آمینه در SSL12، SSL30 و SSL50 متغیر است. همچنین میانگین وزن این پروتئین‌ها ۳۵،۶۷ کیلودالتون محاسبه شد. پروتئین‌های SSL28 با وزن ۲۰،۶ کیلودالتون و SSL50 با وزن ۴۶،۶۹ کیلودالتون به‌ترتیب کم‌ترین و بیش‌ترین وزن مولکولی را داشتند. نقطه ایزوالکتریک نظری پروتئین‌های SSL با میانگین ۷،۴۹ در دامنه ۵،۰۴ در SSL4 تا ۹،۵۲ در SSL3 قرار داشت (جدول ۱). مقایسه ویژگی‌های فیزیوشیمیایی پروتئین‌های SSL گندم، صنوبر کالیفرنایی، ذرت و آرابیدوپسیس نشان داد که وزن مولکولی SSL‌های گندم و ذرت کم‌تر از آرابیدوپسیس و صنوبر کالیفرنایی است. میانگین وزن مولکولی SSL‌های ذرت، آرابیدوپسیس و صنوبر کالیفرنایی به‌ترتیب ۳۶،۵۲، ۴۰،۶۲ و ۴۰،۹۵ کیلودالتون بود. میانگین طول پروتئین‌های SSL در ذرت و گندم به‌ترتیب دارای ۳۲۸،۲ و ۳۴۰ اسید آمینه بود که در این مورد نیز نسبت به آرابیدوپسیس و صنوبر آمریکایی طول کم‌تری داشتند، زیرا میانگین طول پروتئین‌های SSL این دو گیاه ۳۶۷ و ۳۷۰ محاسبه شد. میانگین نقطه ایزوالکتریک گندم نزدیک به صنوبر و بیش‌تر از میانگین آن برای ذرت و آرابیدوپسیس بود که مقدار آن برای ذرت، صنوبر و آرابیدوپسیس به‌ترتیب ۶،۵۷، ۷،۲۱ و ۶،۶۷ گزارش شده است. از آنجایی که ذرت و گندم هر دو از خانواده غلات هستند، بنابراین می‌توان گفت ویژگی‌های فیزیوشیمیایی مشابه بیش‌تری نسبت به صنوبر کالیفرنایی و آرابیدوپسیس دارند.

جدول ۱- مشخصات فیزیکیوشیمیایی خانواده ژنی *TaSSL*
 Table 1. Physico-chemical characteristics of *TaSSL* gene family

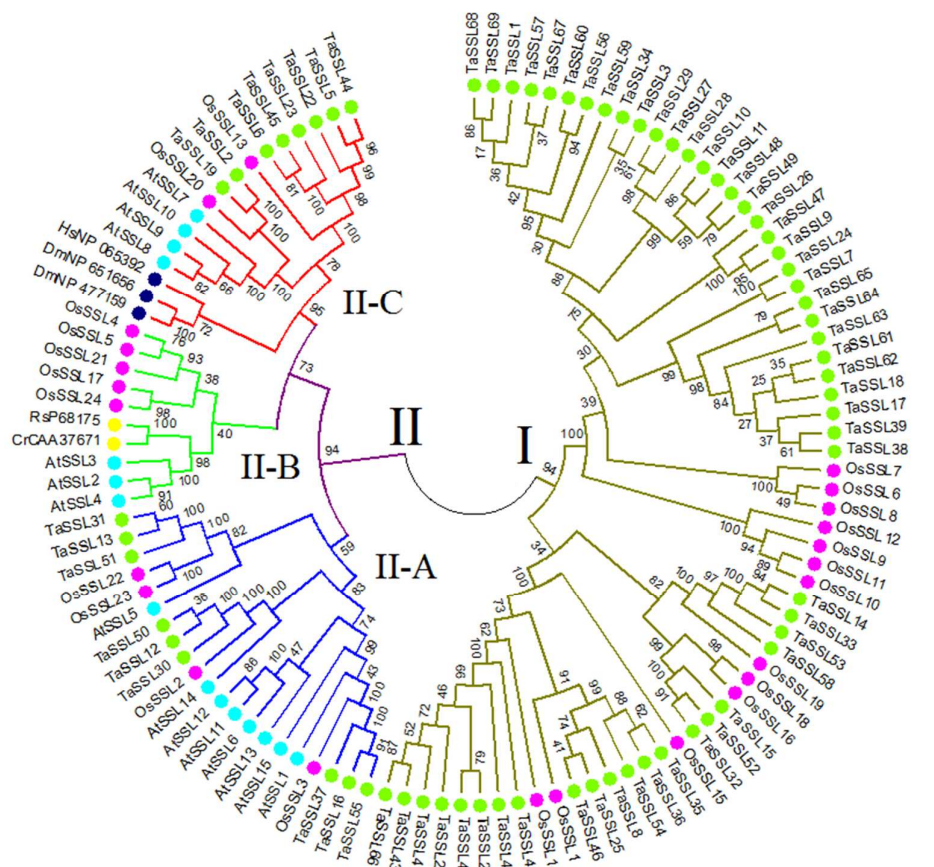
Gene stable ID	Gene name	Chromosome	Gene start (bp)	Gene end (bp)	Strand	Length	Weight (kDA)	pI
TraesCS1A02G026500	TaSSL1	1A	12639162	12640199	+	346	37.72	8.28
TraesCS1A02G203800	TaSSL2	1A	365536062	365536994	+	245	27.01	6.48
TraesCS1A02G437900	TaSSL3	1A	587769449	587770204	+	187	20.62	9.52
TraesCS2A02G078000	TaSSL4	2A	35582909	35584365	-	344	36.65	6.27
TraesCS2A02G170200	TaSSL5	2A	125251161	125253229	-	397	42.4	5.33
TraesCS2A02G170300	TaSSL6	2A	125457444	125459448	-	321	34.58	5.23
TraesCS2A02G260200	TaSSL7	2A	409642762	409643864	+	347	37.68	9.47
TraesCS3A02G256600	TaSSL8	3A	478662356	478666993	+	337	36.34	6.17
TraesCS3A02G298200	TaSSL9	3A	532640354	532641654	+	318	34.81	7.24
TraesCS3A02G496700	TaSSL10	3A	722394773	722396059	+	346	37.73	9.4
TraesCS3A02G496900	TaSSL11	3A	722438264	722439476	-	351	38.18	9.28
TraesCS4A02G089000	TaSSL12	4A	94979877	94981689	-	414	46.58	6.92
TraesCS5A02G131200	TaSSL13	5A	295586795	295589905	+	254	27.17	5.18
TraesCS5A02G188200	TaSSL14	5A	389724313	389726710	-	364	39.21	7.8
TraesCS5A02G188300	TaSSL15	5A	389763146	389764957	-	378	40.58	8.39
TraesCS5A02G390200	TaSSL16	5A	585489225	585491855	-	399	43.68	5.97
TraesCS7A02G138300	TaSSL17	7A	90586924	90587553	+	210	23.02	8.63
TraesCS7A02G138700	TaSSL18	7A	90663981	90664610	-	210	23.06	8.95
TraesCS1B02G217600	TaSSL19	1B	394852348	394854665	+	368	39.31	5.13
TraesCS2B02G092900	TaSSL20	2B	53519486	53520805	-	343	36.6	7.23
TraesCS2B02G093000	TaSSL21	2B	53577415	53578743	+	344	36.65	6.43
TraesCS2B02G196500	TaSSL22	2B	174591805	174593791	-	368	39.21	5.22
TraesCS2B02G196600	TaSSL23	2B	174662494	174664338	-	366	39	5.42
TraesCS2B02G281700	TaSSL24	2B	390113753	390115093	-	345	37.28	9.2
TraesCS3B02G289800	TaSSL25	3B	465068549	465070005	+	337	36.4	6.27
TraesCS3B02G338500	TaSSL26	3B	544923468	544924750	-	344	37.34	7.23
TraesCS3B02G558800	TaSSL27	3B	793018848	793019713	+	210	22.98	8.95
TraesCS3B02G559000	TaSSL28	3B	793045532	793046161	+	188	20.6	8.79
TraesCS3B02G559100	TaSSL29	3B	793061658	793062881	+	346	37.63	9.11
TraesCS4B02G215300	TaSSL30	4B	454923053	454924792	+	414	46.59	6.65
TraesCS5B02G133600	TaSSL31	5B	249711888	249714643	-	375	40.87	6.18
TraesCS5B02G195300	TaSSL32	5B	352577250	352579265	+	347	37.1	7.87
TraesCS5B02G195400	TaSSL33	5B	353013738	353016479	+	365	39.22	6.39
TraesCS5B02G330900	TaSSL34	5B	515187558	515188179	+	186	20.88	6.48
TraesCS5B02G382800	TaSSL35	5B	561692255	561693729	-	342	36.93	7.79
TraesCS5B02G383000	TaSSL36	5B	561712281	561713669	-	340	36.73	7.08
TraesCS5B02G395100	TaSSL37	5B	572546647	572549222	-	400	43.83	5.91
TraesCS7B02G040400	TaSSL38	7B	40346759	40352481	+	349	37.79	9.02
TraesCS7B02G040900	TaSSL39	7B	40442512	40443808	-	349	37.8	9.13
TraesCS2D02G076000	TaSSL40	2D	32520264	32521551	-	344	36.76	6.09
TraesCS2D02G076200	TaSSL41	2D	32557965	32559316	+	344	36.73	6.54
TraesCS2D02G076300	TaSSL42	2D	32594073	32595754	-	347	37.58	5.41
TraesCS2D02G076400	TaSSL43	2D	32598127	32599584	+	344	36.84	7.23
TraesCS2D02G177700	TaSSL44	2D	121059689	121061706	-	368	38.97	5.04
TraesCS2D02G177900	TaSSL45	2D	121354288	121356255	-	323	34.77	5.6
TraesCS3D02G256900	TaSSL46	3D	359092168	359093581	+	358	38.61	6.61
TraesCS3D02G303900	TaSSL47	3D	418417313	418418550	-	321	34.87	6.83
TraesCS3D02G503900	TaSSL48	3D	591624629	591625564	+	312	34.21	9.34
TraesCS3D02G504000	TaSSL49	3D	591653654	591654892	-	346	37.71	9.33
TraesCS4D02G215800	TaSSL50	4D	369668365	369670201	+	414	46.69	7.31
TraesCS5D02G139700	TaSSL51	5D	223229427	223232332	+	374	40.8	6.51
TraesCS5D02G202700	TaSSL52	5D	307967730	307969560	+	347	37.18	7.25

Table 1. Continued		جدول ۱- ادامه						
TraesCS5D02G202800	TaSSL53	5D	308217655	308221035	+	365	39.48	8.23
TraesCS5D02G389100	TaSSL54	5D	458458282	458459667	-	356	38.37	6.53
TraesCS5D02G400000	TaSSL55	5D	465200400	465203233	-	399	43.65	5.91
TraesCS7D02G075900	TaSSL56	7D	45028549	45029902	-	346	37.66	8.27
TraesCS7D02G076000	TaSSL57	7D	45056264	45057289	+	342	37.53	8.9
TraesCS7D02G076300	TaSSL58	7D	45105633	45107975	-	362	39.6	8.31
TraesCS7D02G076500	TaSSL59	7D	45192155	45192727	+	191	21.52	8.6
TraesCS7D02G076600	TaSSL60	7D	45221301	45221930	+	210	23.25	8.02
TraesCS7D02G139900	TaSSL61	7D	89486370	89487624	+	349	37.82	9.03
TraesCS7D02G140000	TaSSL62	7D	89522303	89523349	+	349	37.75	9.13
TraesCS7D02G140500	TaSSL63	7D	89730530	89731768	-	346	37.51	9.2
TraesCS7D02G140600	TaSSL64	7D	89790365	89791410	-	308	33.19	8.53
TraesCSU02G104000	TaSSL65	Un	91490467	91491513	+	349	37.76	9.08
TraesCSU02G189700	TaSSL66	Un	285394503	285395933	-	344	36.84	7.23
TraesCSU02G223800	TaSSL67	Un	330012358	330013691	+	346	37.71	8.28
TraesCSU02G240600	TaSSL68	Un	358915197	358915768	+	190	21.23	9.32
TraesCSU02G255900	TaSSL69	Un	396475176	396475747	+	190	21.12	9.19

مطالعه فیلوژنتیک خانواده ژنی *SSL* گندم

به‌منظور درک بهتر روابط تکاملی ژن‌های *SSL* گندم، درخت فیلوژنتیک پروتئین‌های *SSL* برنج، گندم و آرابیدوپسیس به‌همراه توالی آنزیم STR پرپوش و آسرول و نیز پروتئین همومیوسین انسان و مگس سرکه بر اساس روش نزدیک‌ترین همسایه‌ها رسم شد. نتایج نشان داد که درخت فیلوژنتیک به دو گروه اصلی تقسیم می‌شود (شکل ۱). گروه اول با ۶۵ عضو بزرگ‌تر بوده و تنها شامل *SSL*‌های گندم و برنج است. گروه دوم از ۴۸ توالی ایجاد شده و تمامی موجودات اشاره شده دارای نماینده هستند (شکل ۱). گروه دوم را می‌توان به سه زیرگروه تقسیم کرد. زیرگروه A با ۲۱ عضو شامل *SSL*‌های گندم، برنج و آرابیدوپسیس، زیرگروه B شامل پنج *SSL* برنج، سه *SSL* آرابیدوپسیس و پروتئین‌های STR پرپوش و آسرول، و زیرگروه C شامل دو *SSL* برنج، سه *SSL* آرابیدوپسیس، هشت *SSL* گندم و نیز سه همومیوسین مربوط به انسان و مگس سرکه است (شکل ۲). مطالعه تکاملی خانواده ژنی *SSL* در ذرت، آژیلوپس، سورگوم، برنج و آرابیدوپسیس نیز نشان داد که از پنج زیرگروه، سه زیرگروه اختصاصی گیاهان تک‌لپه بوده و ژن‌های *SSL* آرابیدوپسیس محدود به دو گروه می‌باشد. حضور ژن‌های *SSL* تک‌لپه و دو لپه در گروه دوم نشان می‌دهد که این ژن‌ها از یک جد مشترک تکامل پیدا کرده‌اند (Wang et al., 2018). اما ایجاد گروه فیلوژنتیک اول که مختص گندم و برنج است، می‌تواند حاکی از تکامل مستقل خانواده ژنی *SSL* در

گیاهان تک‌لپه پس از انشقاق از گیاهان دولپه باشد که در نتیجه فرآیندهایی مانند حذف ژن یا تکثیر اختصاصی ژن در یک گونه یا لینه ایجاد می‌شود (Liu et al., 2020). از آنجایی که ژن‌های اورتولوگ گونه‌های مختلف موجود در یک گروه فیلوژنتیک می‌توانند کارکرد مشابه داشته باشند، از این‌رو می‌توان فرض کرد که ژن‌های *SSL* گندم، برنج و آرابیدوپسیس موجود در گروه‌های II-B و II-C به‌ترتیب در بیوسنتز آلکالوئیدها و پاسخ ایمنی گیاه نقش داشته باشند (Guo et al., 2008). در گروه II-B ژن *STR* پرپوش و آسرول به‌عنوان ژن کلیدی در بیوسنتز استریکتوسیدین ساخته می‌شوند. با این‌حال نتایج مطالعه ژن‌های *SSL* این زیرگروه نشان داده است که توانایی تولید استریکتوسیدین از تریپتامین و سکلوگانین را ندارد و تنها می‌تواند سکلوگانین را به‌صورت دایمر تولید کند. گروه II-C از ژن‌های *SSL* گیاهی و ژن‌های همومیوسین تشکیل شده است که در ایمنی نقش دارند (Theopold et al., 1996). بر اساس مطالعه انجام گرفته روی ژن‌های *SSL* آرابیدوپسیس این زیرگروه، نقش این ژن‌ها در ایجاد مقاومت آرابیدوپسیس به تنش‌های زیستی و غیرزیستی است. بر این اساس، ژن‌های *AtSSL7*، *AtSSL8*، *AtSSL9* و *AtSSL10* را شبیه همومیوسین نیز می‌نامند (Sohani et al., 2009). می‌توان پیش‌بینی کرد که ژن‌های *SSL* گندم و برنج این زیرگروه نیز نقش دفاعی در گیاه داشته باشند. با این‌حال این امر نیاز به مطالعه بیشتر دارد.



شکل ۱- درخت فیلوژنتیک خانواده ژنی *TaSSL* گندم (*Ta*)، برنج (*Os*) و *Arabidopsis thaliana* (*At*)، STR آسرول (*Rs*) و پریوش (*Cr*)، همومیوسین انسان (*Hs*) و مگس سرکه (*Dm*). ژن‌های مربوط به هر موجود با رنگ متفاوت نشان داده شده است. رسم درخت به روش نزدیک‌ترین همسایه‌ها (NJ) و نرم‌افزار MEGA7 با ۱۰۰۰ تکرار بوت‌استرپ انجام شد.

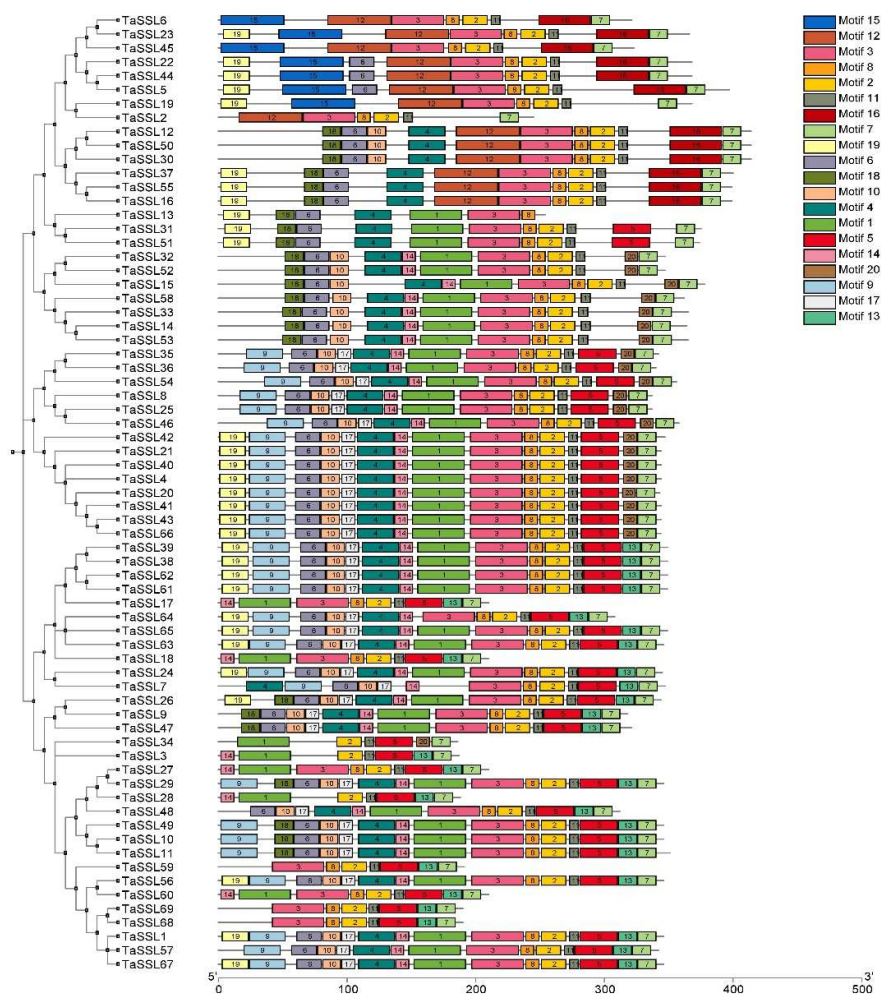
Figure 1. Phylogenetic tree of the *TaSSL* gene family of *Triticum aestivum* (*Ta*), *Oryza sativa* (*Os*) and *Arabidopsis thaliana* (*At*), STR of *Rauwolfia serpentine* (*Rs*) and *Catharanthus roseus* (*Cr*), hemomucin of human and fruit fly. Genes related to each organism are shown with a different color. The tree was constructed using MEGA7 software using neighbor-joining (NJ) method with 1000 bootstraps.

سه و هشت شناسایی شد. موتیف پنج با فراوانی هفت در *TaSSL6*، *TaSSL23*، *TaSSL45*، *TaSSL22*، *TaSSL5*، *TaSSL44* و *TaSSL19* تشخیص داده شد که از نظر فیلوژنتیک در گروه II-C قرار داشته و رابطه نزدیکی با ژن‌های همومیوسین دارند (شکل ۲). بررسی کارکردی این موتیف‌ها از طریق سرور SMART نشان داد که موتیف‌های دو، سه، هشت و یازده مربوط به دمین کارکردی Str_synth (PF03088) هستند. همچنین موتیف پنج مربوط به دمین انتهایی آمینو شبیه استریکتوسیدین سینتاز (PF20067) است.

شناسایی موتیف‌های حفاظت شده و ساختار اگزون-

اینترون خانواده ژنی *SSL* گندم

برای درک بهتر میزان حفاظت شدگی موجود در خانواده ژنی *SSL* گندم، ساختار موتیف‌های حفاظت شده این پروتئین‌ها ارزیابی شد. در نتیجه این بررسی ۲۰ موتیف حفاظت شده شناسایی شد. بررسی فراوانی و توزیع این موتیف‌ها حاکی از حفاظت شدگی بالای *SSL*‌های گندم داشت (شکل ۲). موتیف‌های دو، هفت و یازده در ۶۸ پروتئین *SSL* شناسایی شد و تنها *TaSSL12* فاقد این موتیف‌ها بود. همچنین به جز *TaSSL3*، *TaSSL28* و *TaSSL34* در سایر پروتئین‌های این خانواده، موتیف‌های



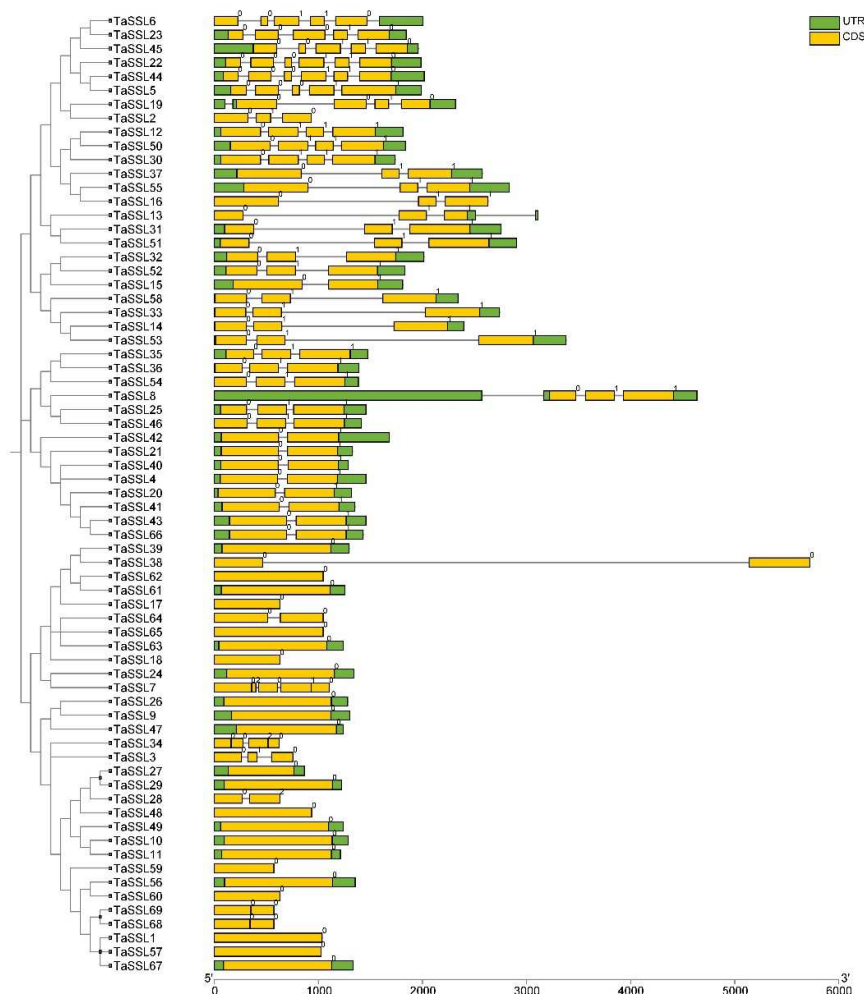
شکل ۲- موتیف‌های حفاظت شده پروتئین‌های TaSSL. هر موتیف با یک رنگ مجزا مشخص شده است.

Figure 2. The conserved motifs of the TaSSL proteins. Different motifs are presented in different colors.

اینترن به‌عنوان بخشی از استراتژی سازگاری برای پاسخ سریع به تنش‌های زیستی و غیرزیستی از ژن‌های غنی از اینترن منشاء گرفته‌اند (Liu *et al.*, 2021). بنابراین، احتمالاً ژن‌های SSL گندم در پاسخ به تنش‌ها می‌توانند نقش داشته باشند. مقایسه ساختار ژنی خانواده SSL، گندم با ذرت و صنوبر کالیفرنایی نشان داد که برخلاف ژن‌های SSL گندم و ذرت که صفر تا پنج اینترن (عموماً فاقد اینترن) دارند، خانواده ژنی SSL صنوبر کالیفرنایی دارای دو تا شش اینترن بوده و ژن‌های SSL فاقد اینترن در این گیاه وجود ندارد (Gu *et al.*, 2023; Wang *et al.*, 2023a). این نتایج نشان دهنده مسیر تکاملی متفاوت ژن‌های SSL گندم و ذرت با صنوبر کالیفرنایی به‌عنوان یک گیاه دولپه متفاوت بوده و نتایج آنالیز فیلوژنتیک را تأیید می‌کند.

تفاوت ساختاری اگزون-اینترن منبع مهمی برای تنوع خانواده ژنی و تنوع زیستی گیاهی است. این تفاوت و تنوع ساختاری بیان ژن را تحت تاثیر قرار می‌دهد (Xu *et al.*, 2012). مطالعه ساختار اگزون-اینترن خانواده ژنی SSL، گندم نشان داد که بیش‌ترین تعداد اینترن مربوط به زیر گروه II-C فیلوژنتیک است (شکل ۳). در این زیر گروه به استثناء ژن *TaSSL2* که دو اینترن دارد، سایر ژن‌ها دارای چهار یا پنج اینترن هستند. ژن‌های زیر گروه II-A و II-B دو یا سه اینترن دارند. همچنین کم‌ترین تعداد اینترن در گروه I فیلوژنتیکی مشاهده شد و به جز ژن *TaSSL7* که چهار اینترن داشت، سایر ژن‌های این گروه از صفر تا سه اینترن دارند (شکل ۳).

ژن‌های یوکاریوتی از نظر تعداد اینترن به سه گروه فاقد اینترن، دارای اینترن کم و غنی از اینترن‌ها تقسیم می‌شوند. از نظر تکاملی، ژن‌های دارای اینترن کم یا فاقد



شکل ۳- ساختار اگزون- اینترون ژن‌های *TaSSL* گندم. نواحی غیر کد کننده، اگزون‌ها و اینترون‌ها به ترتیب با جعبه‌های سبز و زرد و خطوط سیاه مشخص شده‌اند.

Figure 3. The exon – intron structure of *TaSSL* in *T. aestivum*. UTR, exons and introns are presented by green and yellow boxes and black lines, respectively.

بالا بودن فراوانی فاز اینترونی صفر و یک در ژن‌های *SSL* حفاظت‌شدگی بالای ساختار ژنی آن‌ها را نشان می‌دهد (Yan *et al.*, 2019).

آنالیز ناحیه پروموتوری خانواده ژنی *SSL* گندم

عناصر تنظیمی سیس توالی‌هایی در بالادست ۵' ژن‌ها هستند که بیان ژن را تحت تاثیر قرار می‌دهند. اگرچه این توالی‌ها قادر به کد کردن پروتئین نیستند، با این حال جایگاه کارکردی را ایجاد کرده و از طریق تعامل با عوامل رونویسی بیان ژن در مراحل مختلف حیات گیاه را تنظیم می‌کنند (Hernandez-Garcia & Finer, 2014). از این‌رو برای پیش‌بینی نقش بالقوه ژن‌های *SSL* گندم، در این مطالعه ۱۵۰۰ جفت باز بالادست ژن‌های *SSL* گندم بررسی و مشخص شد که در مجموع ۱۲۹۶ عنصر سیس

اسپلایسینگ اینترون به سه فاز صفر، یک و دو تقسیم می‌شود. در فاز صفر، اینترون بین دو کدون، در فاز یک، اینترون بین نوکلئوتید اول و دوم کدون و در فاز دو، اینترون بین نوکلئوتید دوم و سوم کدون قرار دارد (Wu *et al.*, 2023). بررسی فاز اینترونی نشان داد که در ژن‌های *SSL* گندم، ۵۸ درصد مربوط به فاز صفر، ۴۰ درصد مربوط به فاز یک و دو درصد مربوط به فاز دو هستند (شکل ۳). فاز اینترونی با میزان حفاظت‌شدگی توالی در جایگاه اسپلایسینگ با تکامل اینترون‌های اسپلایزومی مرتبط است. بالاترین میزان حفاظت‌شدگی مربوط به فاز صفر می‌باشد. همچنین اینترون فاز یک حفاظت‌شدگی بالایی دارد، در حالی‌که اینترون‌های فاز دو کم‌ترین میزان حفاظت‌شدگی را نشان داده‌اند. بنابراین،

پاسخ به تنش‌ها را به‌خود اختصاص دادند. MYB و MYC در پاسخ به خشکی و ARE در پاسخ به شرایط بی‌هوای نقش دارند. بیش‌ترین فراوانی و نوع عنصر تنظیمی در پروموتور ژن‌های *TaSSL1*، *TaSSL9* و *TaSSL60* مشاهده شد. پروموتور ژن *TaSSL1* دارای ۲۲ نوع عنصر تنظیمی با فراوانی ۵۳ بود. در ناحیه تنظیمی ژن‌های *TaSSL9* و *TaSSL60* ۲۱ نوع عنصر تنظیمی شناسایی شد که به‌ترتیب دارای فراوانی ۴۱ و ۳۹ بودند. عناصر تنظیمی سیس پاسخ به تنش‌ها و هورمون‌ها در پروموتور خانواده ژنی *SSL* ذرت و صنوبر کالیفرنایی نیز شناسایی شده است. این نتایج نشان دهنده نقش ژن‌های *SSL* در سازگاری گیاهان به شرایط محیطی است (Gu *et al.*, 2023; Wang *et al.*, 2023a).

در پروموتور این ژن‌ها وجود دارد (جدول ۲). این عناصر سیس مربوط به ۱۴ نوع عنصر تنظیمی پاسخ به هورمون‌ها و ۱۶ نوع عنصر تنظیمی پاسخ به تنش‌ها هستند و در پاسخ به هورمون‌های آبسزیک اسید، سالیسیلیک اسید، اتیلن، اکسین، جیبرلین و تنش‌های سرما، خشکی، کادمیون، دهیدراسیون، زخم، پاتوژن، الیستور و شرایط بی‌هوای نقش دارند (جدول ۲). از میان عناصر سیس پاسخ به هورمون‌ها، بیش‌ترین فراوانی مربوط به *ABRE*، *ABRE4* و *ABRE3A* بود که به‌ترتیب فراوانی ۱۰۱، ۷۵ و ۷۵ داشتند و هر سه در پاسخ به هورمون آبسزیک اسید نقش دارند. از سوی دیگر عناصر تنظیمی *MYB* با فراوانی ۲۵۶، *MYC* با فراوانی ۱۶۸ و *ARE* با فراوانی ۱۰۰ بیش‌ترین تعداد عناصر سیس

جدول ۲- عناصر تنظیمی سیس شناسایی شده در پروموتور ژن‌های *TaSSL* و کارکرد آن‌ها
Table 2. Cis-regulatory elements in *TaSSL* genes promoters and their function

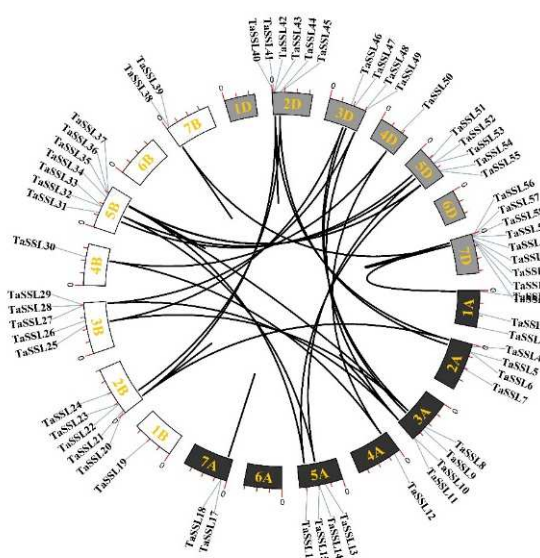
Type	Cis Elements	Function	Frequency
Hormone	ABRE2	cis-acting element involved in the abscisic acid responsiveness	1
	AT~ABRE	ABA response	3
	SARE	cis-acting element involved in salicylic acid responsiveness	3
	TATC-box	cis-acting element involved in gibberellin-responsiveness	5
	AuxRR-core	cis-acting regulatory element involved in auxin responsiveness	6
	CARE	ABA response, GA response	10
	GARE-motif	gibberellin-responsive	18
	ERE	ethylene-responsive element	23
	P-box	gibberellin-responsive element	23
	TCA-element	cis-acting element involved in salicylic acid responsiveness	32
	TGA-element	auxin-responsive element	34
	ABRE3a	cis-acting element involved in the abscisic acid responsiveness	75
	ABRE4	cis-acting element involved in the abscisic acid responsiveness	75
	ABRE	cis-acting element involved in the abscisic acid responsiveness	101
	DRE	dehydration-responsive	1
	AP-1	Cd response	2
	box S	wounding and pathogen response	12
	DRE1	drought response	14
	Stress	TC-rich repeats	cis-acting element involved in defense and stress responsiveness
WUN-motif		wound-responsive element	19
MYB recognition site		drought response	26
GC-motif		enhancer-like element involved in anoxic specific inducibility	26
MBS		MYB binding site involved in drought-inducibility	41
W box		elicitor responsive cis-element	44
LTR		cis-acting element involved in low-temperature responsiveness	48
DRE core		drought response	57
WRE3		wounding	57
ARE		cis-acting regulatory element essential for the anaerobic induction	100
MYC	drought response	168	
MYB	water response, drought response	256	

مضاعف‌شدگی و فشار گزینش

پدیده مضاعف‌شدگی از مهم‌ترین ویژگی‌های ساختار ژنومی گیاهان است و به تکامل ژنوم کمک می‌کند. از آنجا که مضاعف‌شدگی موجب ایجاد ژن‌های جدید می‌شود، بنابراین نقش مهمی در تکامل و بسط خانواده‌های ژنی ایفا می‌کند (Cannon *et al.*, 2004). گسترش خانواده‌های ژنی از طریق سه فرآیند مضاعف‌شدگی سگمنتال، تاندم و مضاعف‌شدگی کامل ژنوم (Whole genome duplication) اتفاق می‌افتد (Magwanga *et al.*, 2018). در مضاعف‌شدگی سگمنتال، ژن‌های مضاعف شده در سراسر ژنوم پراکنده می‌شوند، در حالی که در مضاعف‌شدگی تاندم ژن‌های مضاعف‌شده تمایل دارند با یکدیگر ایجاد خوشه کنند (Qanmber *et al.*, 2019). نتایج نشان داد که ۷۳ جفت ژن مضاعف‌شده در خانواده ژنی SSL گندم وجود دارد که ۶۴ جفت ژن در اثر مضاعف‌شدگی سگمنتال و نه جفت ژن در اثر مضاعف‌شدگی تاندم ایجاد شده‌اند (شکل ۴، جدول ۳). از این‌رو مضاعف‌شدگی سگمنتال عامل اصلی بسط این خانواده ژنی در گندم است. مضاعف‌شدگی سگمنتال در بسط و تکامل خانواده‌های ژنی نقش مهمی دارد و برای سازگاری گیاه به محیط و گونه‌زایی ضروری است (Yu *et al.*, 2023). بر این اساس، می‌توان گفت که ژن‌های

TaSSL حاصل از مضاعف‌شدگی سگمنتال می‌تواند در پاسخ به تنش‌ها نقش داشته باشند.

نسبت *Ka/Ks* یا فشار گزینش، شاخص مهمی برای بررسی تکامل ژن‌ها است. نتایج این مطالعه نشان داد که دو جفت ژن مضاعف‌شده فشار گزینش مثبت و سایر ژن‌های مضاعف‌شده فشار گزینش منفی را متحمل شده‌اند. میانگین فشار گزینش منفی و مثبت نیز به ترتیب ۰,۱۷، ۰,۲۹ بود (جدول ۳). گزینش منفی موجب می‌شود که سرعت تکامل خانواده‌های ژنی در سطح پروتئین کاهش یافته و نقش کارکردی آن‌ها حفظ شود (Ahmad *et al.*, 2022). از سوی دیگر فشار گزینش مثبت موجب ایجاد ژن‌های چند کارکردی در نتیجه جهش‌های ایجاد شده طی تکامل می‌شود (Zhang *et al.*, 2023). از این‌رو می‌توان گفت کارکرد ژن‌های مضاعف‌شده *SSL*، که فشار گزینش منفی را تجربه کرده‌اند، طی تکامل خود حفظ شده و ژن‌های تحت کنترل فشار گزینش مثبت می‌توانند کارکردهای متفاوتی داشته باشند. در نتیجه از یک سو ژن‌های *SSL* کارکردهای پیشین خود را انجام می‌دهند و از سوی دیگر، نقش‌های زیستی جدید ایجاد شده در این ژن‌ها می‌تواند توانایی‌های گندم در جنبه‌های مختلف زیستی مانند عملکرد و یا مقاومت در برابر تنش‌ها را افزایش دهد.



شکل ۴- جایگاه کروموزومی ژن‌های *TaSSL* گندم. کروموزوم‌های هر زیرژنوم با رنگ متفاوت نشان داده شده‌اند. منحنی‌های سیاه رنگ ژن‌های مضاعف شده را به هم متصل کرده‌اند.

Figure 4. Chromosomal locations of *T. aestivum* *SSL* genes. Each sub-genome chromosomes are represented by different colored boxes. Black curves connecting the genes indicate duplications.

جدول ۳- فشار گزینش و نوع مضاعف‌شدگی جفت ژن‌های *TaSSL* مضاعف‌شده
 Table 3. Selection pressure and the type of duplication of *TaSSL* gene pairs

Gene 1	Gene 2	Ka	Ks	Ka\ Ks	Selection	Duplication
TaSSL56	TaSSL1	0.01	0.17	0.07	Negative	Segmental
TaSSL57	TaSSL1	0.02	0.09	0.25	Negative	Segmental
TaSSL67	TaSSL1	0.01	0.00	3.01	Positive	Segmental
TaSSL10	TaSSL11	0.02	0.19	0.08	Negative	Tandem
TaSSL49	TaSSL11	0.01	0.19	0.07	Negative	Segmental
TaSSL37	TaSSL16	0.01	0.13	0.09	Negative	Segmental
TaSSL55	TaSSL16	0.00	0.12	0.04	Negative	Segmental
TaSSL17	TaSSL18	0.01	0.16	0.07	Negative	Tandem
TaSSL20	TaSSL21	0.03	0.13	0.20	Negative	Tandem
TaSSL4	TaSSL21	0.03	0.12	0.22	Negative	Segmental
TaSSL40	TaSSL21	0.03	0.13	0.23	Negative	Segmental
TaSSL41	TaSSL21	0.03	0.13	0.21	Negative	Segmental
TaSSL43	TaSSL21	0.03	0.10	0.25	Negative	Segmental
TaSSL8	TaSSL25	0.02	0.11	0.15	Negative	Segmental
TaSSL10	TaSSL29	0.02	0.12	0.21	Negative	Segmental
TaSSL11	TaSSL29	0.03	0.19	0.14	Negative	Segmental
TaSSL49	TaSSL29	0.02	0.13	0.16	Negative	Segmental
TaSSL12	TaSSL30	0.01	0.06	0.10	Negative	Segmental
TaSSL50	TaSSL30	0.01	0.07	0.08	Negative	Segmental
TaSSL14	TaSSL33	0.02	0.05	0.35	Negative	Segmental
TaSSL35	TaSSL36	0.01	0.15	0.09	Negative	Tandem
TaSSL8	TaSSL36	0.04	0.27	0.14	Negative	Segmental
TaSSL39	TaSSL38	0.01	0.05	0.17	Negative	Tandem
TaSSL20	TaSSL4	0.02	0.10	0.23	Negative	Segmental
TaSSL20	TaSSL40	0.02	0.09	0.18	Negative	Segmental
TaSSL4	TaSSL40	0.02	0.09	0.25	Negative	Segmental
TaSSL20	TaSSL41	0.01	0.12	0.09	Negative	Segmental
TaSSL4	TaSSL41	0.02	0.11	0.17	Negative	Segmental
TaSSL40	TaSSL41	0.02	0.12	0.16	Negative	Segmental
TaSSL43	TaSSL41	0.01	0.09	0.13	Negative	Segmental
TaSSL20	TaSSL43	0.02	0.12	0.14	Negative	Segmental
TaSSL4	TaSSL43	0.03	0.12	0.23	Negative	Segmental
TaSSL40	TaSSL43	0.02	0.11	0.23	Negative	Segmental
TaSSL22	TaSSL44	0.03	0.09	0.32	Negative	Segmental
TaSSL25	TaSSL46	0.01	0.07	0.21	Negative	Segmental
TaSSL8	TaSSL46	0.02	0.09	0.20	Negative	Segmental
TaSSL9	TaSSL47	0.02	0.19	0.12	Negative	Segmental

Table 3. Continued

جدول ۳- ادامه

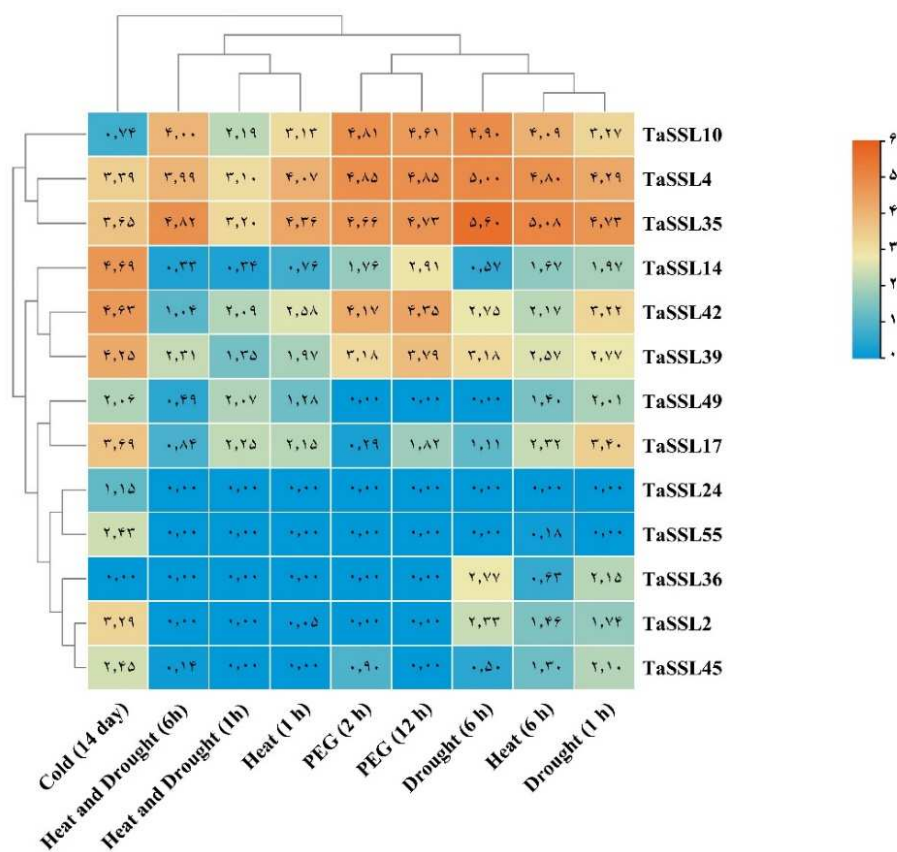
Gene 1	Gene 2	Ka	Ks	Ka\ Ks	Selection	Duplication
TaSSL10	TaSSL49	0.02	0.13	0.14	Negative	Segmental
TaSSL44	TaSSL5	0.01	0.06	0.21	Negative	Segmental
TaSSL12	TaSSL50	0.00	0.06	0.03	Negative	Segmental
TaSSL31	TaSSL51	0.01	0.05	0.23	Negative	Segmental
TaSSL32	TaSSL52	0.01	0.07	0.17	Negative	Segmental
TaSSL14	TaSSL53	0.01	0.05	0.25	Negative	Segmental
TaSSL33	TaSSL53	0.02	0.05	0.37	Negative	Segmental
TaSSL35	TaSSL54	0.02	0.14	0.13	Negative	Segmental
TaSSL36	TaSSL54	0.02	0.15	0.11	Negative	Segmental
TaSSL37	TaSSL55	0.01	0.13	0.09	Negative	Segmental
TaSSL56	TaSSL57	0.03	0.16	0.16	Negative	Tandem
TaSSL68	TaSSL59	0.03	0.14	0.21	Negative	Segmental
TaSSL69	TaSSL59	0.03	0.14	0.20	Negative	Segmental
TaSSL45	TaSSL6	0.02	0.11	0.16	Negative	Segmental
TaSSL38	TaSSL61	0.02	0.11	0.17	Negative	Segmental
TaSSL39	TaSSL61	0.02	0.13	0.17	Negative	Segmental
TaSSL62	TaSSL61	0.01	0.10	0.12	Negative	Tandem

TaSSL63	TaSSL61	0.03	0.16	0.20	Negative	Tandem
TaSSL65	TaSSL61	0.05	0.34	0.16	Negative	Segmental
TaSSL38	TaSSL62	0.02	0.14	0.10	Negative	Segmental
TaSSL39	TaSSL62	0.02	0.17	0.10	Negative	Segmental
TaSSL63	TaSSL62	0.03	0.15	0.17	Negative	Tandem
TaSSL65	TaSSL62	0.05	0.33	0.16	Negative	Segmental
TaSSL38	TaSSL63	0.04	0.17	0.22	Negative	Segmental
TaSSL39	TaSSL63	0.04	0.17	0.24	Negative	Segmental
TaSSL65	TaSSL63	0.03	0.25	0.11	Negative	Segmental
TaSSL38	TaSSL65	0.06	0.35	0.16	Negative	Segmental
TaSSL20	TaSSL66	0.02	0.12	0.14	Negative	Segmental
TaSSL21	TaSSL66	0.03	0.10	0.25	Negative	Segmental
TaSSL4	TaSSL66	0.03	0.12	0.23	Negative	Segmental
TaSSL40	TaSSL66	0.02	0.11	0.23	Negative	Segmental
TaSSL41	TaSSL66	0.01	0.09	0.13	Negative	Segmental
TaSSL43	TaSSL66	0.00	0.00	NaN	nnnnn	Segmental
TaSSL56	TaSSL67	0.01	0.16	0.06	Negative	Segmental
TaSSL57	TaSSL67	0.02	0.09	0.18	Negative	Segmental
TaSSL69	TaSSL68	0.01	0.00	3.57	Positive	Segmental

توجهی دارد. بر این اساس می‌توان چهار الگوی بیان مشاهده کرد. الگوی اول مربوط به بیان ژن‌ها در پاسخ به تنش سرما است که کاملاً از سایر تنش‌ها متفاوت بوده و به‌جز ژن‌های *TaSSL24* و *TaSSL36* سایر ژن‌ها بیان متوسط تا بالایی را در پاسخ به تنش سرمایی داشتند (شکل ۵). گروه دوم مربوط به پاسخ به تنش گرمایی (یک ساعت پس از تنش) و تنش توأم خشکی و گرما (یک و شش ساعت پس از تنش) است. بر اساس میزان بیان سه ژن *TaSSL10*، *TaSSL4* و *TaSSL35* نقش اصلی در پاسخ به این تنش‌ها را بر عهده داشتند (شکل ۵). گروه سوم مربوط به تنش اسمزی القاء شده به‌وسیله PEG است که ژن‌های موثر در پاسخ به این تنش شامل *TaSSL10*، *TaSSL4*، *TaSSL35*، *TaSSL14*، *TaSSL42* و *TaSSL39* بودند (شکل ۵). گروه چهارم مربوط به تنش خشکی (یک و شش ساعت پیش از تنش) و تنش گرمایی (شش ساعت پس از تنش) است. در این گروه می‌توان گفت که تنها دو ژن *TaSSL24* و *TaSSL55* فاقد بیان بودند و سایر ژن‌های مورد مطالعه طیفی از بیان متوسط تا بالا را نشان دادند (شکل ۵). مطالعه بیان ژن‌های *SSL* در سایر گیاهان مانند ذرت، صنوبر کالیفرنایی و آرابیدوپسیس نیز نشان داده است که این ژن‌ها می‌توانند به‌عنوان ژن‌های موثر در پاسخ به تنش‌ها شناخته شوند (Sohani et al., 2009; Gu et al., 2023a; Wang et al., 2023).

پاسخ ژن‌های *SSL* گندم به تنش‌های غیرزیستی
 به‌منظور ارزیابی پاسخ ژن‌های *SSL* به تنش‌های غیرزیستی، از داده‌های RNA-seq سرور expVIP استفاده شد. بر اساس میزان بیان، ژن‌ها به سه گروه کم ($\log_2 TPM < 2$)، متوسط ($2 \leq \log_2 TPM \leq 4$) و بالا ($4 \leq \log_2 TPM \leq 6$) تقسیم شدند (Sarcheshmeh et al., 2023). بر این اساس، ژن‌های مورد مطالعه را می‌توان به دو گروه کلی تقسیم کرد. در گروه اول ژن‌هایی قرار دارند که در پاسخ به همه تنش‌ها بیان متوسط تا بالایی از خود نشان داده‌اند (شکل ۵). این گروه شامل ژن‌های *TaSSL4*، *TaSSL10* و *TaSSL35* می‌باشند که به‌جز بیان *TaSSL10* در شرایط تنش سرمایی در پاسخ به سایر تنش‌ها القاء شده‌اند (شکل ۵). گروه دوم شامل ژن‌های *TaSSL14*، *TaSSL42*، *TaSSL39*، *TaSSL49*، *TaSSL17*، *TaSSL24*، *TaSSL55*، *TaSSL36* و *TaSSL2* هستند. بیان هر یک از ژن‌های این گروه به تنش‌ها متفاوت بوده و می‌توان گفت برخی از آن‌ها اختصاصی یک یا چند تنش هستند (شکل ۵). به‌عنوان مثال، ژن *TaSSL36* بیان متوسطی در پاسخ به تنش خشکی در ساعت‌های مختلف داشته و در سایر تنش‌ها فاقد بیان است. چنین حالتی در مورد ژن *TaSSL55* نیز مشاهده می‌شود که بر اساس میزان بیان می‌توان گفت این ژن در شرایط تنش سرمایی القاء می‌شود (شکل ۵).

گروه‌بندی بیان بر اساس نوع تنش نیز نتایج قابل



شکل ۵- الگوی بیان ژن‌های *TaSSL* گندم تحت شرایط تنش‌های متفاوت. جعبه‌های رنگی میزان بیان را نشان می‌دهند. نقشه حرارتی بر اساس مقادیر $\text{Log}_2(\text{TPM} + 1)$ رسم شده است.

Figure 5. The expression pattern of *TaSSL* genes under diverse abiotic stresses. The color boxes indicate expression values. The heatmap was generated using $\text{Log}_2(\text{TPM} + 1)$ values.

نتیجه‌گیری کلی

در این مطالعه برای اولین بار خانواده ژنی *SSL* گندم از نظر تکاملی و کارکردی مورد مطالعه قرار گرفت. بر اساس نتایج درخت فیلوژنی مشخص شد که ژن‌های *SSL* گندم، اگرچه دارای جد مشترک با گیاهان تک‌لپه و دولپه است، اما طی تکامل برخی از ژن‌ها مسیر متفاوتی را طی کرده‌اند که نتیجه آن ایجاد گروه مربوط به گیاهان تک‌لپه می‌باشد. از سوی دیگر، مطالعه موتیف‌های حفاظت‌شده و ساختار ژنی نشان داد که خانواده ژنی *SSL* گندم حفاظت شده است و اینترون‌های کمی دارد. عامل اصلی بسط این خانواده ژنی مضاعف‌شدگی سگمنتال است و طی تکامل فشار گزینش منفی بر ژن‌های مضاعف‌شده اعمال شده است. وجود عناصر تنظیمی پاسخ به هورمون‌ها و تنش‌ها در پروموتور ژن‌های *TaSSL* و نیز الگوی بیانی پیچیده این ژن‌ها در پاسخ به تنش‌های سرما، گرما، اسمز و خشکی

نشان دهنده نقش ژن‌های *TaSSL* در پاسخ به تنش‌های غیرزیستی است. با توجه به این که تنش‌های زیستی و غیرزیستی، عملکرد گندم را به‌عنوان یکی از مهم‌ترین گیاهان زراعی کاهش داده و امنیت غذایی را تهدید می‌کنند، بنابراین شناسایی ژن‌های پاسخ به تنش نقش مهمی در برنامه‌های اصلاحی گندم با هدف افزایش مقاومت به تنش‌ها دارد. ویژگی‌هایی نظیر فشار گزینش منفی اعمال شده بر ژن‌های *TaSSL*، عناصر تنظیمی پاسخ به تنش‌ها و هورمون‌ها، تعداد کم اینترون و پاسخ ژن‌های *TaSSL* به تنش‌های غیرزیستی نقش این ژن‌ها را در پاسخ گندم به شرایط نامساعد محیطی نشان می‌دهند. این مطالعه اولین گام در درک مسیر تکاملی و کارکردی ژن‌های *TaSSL* است و بنابراین از این اطلاعات می‌توان جهت انتخاب ژن‌های موثر در تولید رقم‌های گندم مقاوم به تنش‌ها در مطالعات آینده استفاده کرد.

تضاد منافع

نویسنده (گان) تایید می‌کنند که این تحقیق در غیاب هر گونه روابط تجاری یا مالی که می‌تواند به‌عنوان تضاد منافع بالقوه تعبیر شود، انجام شده است.

تا کنون به‌طور کامل به هیچ‌زبانی و در هیچ نشریه یا همایشی چاپ و منتشر نشده و هیچ اقدامی نیز برای انتشار آن در هیچ نشریه یا همایشی صورت نگرفته و نخواهد گرفت.

رعایت اخلاق در نشر

نویسنده (گان) اعلام می‌کنند که در نگارش این مقاله به‌طور کامل از اخلاق نشر از جمله سرقت ادبی، سوء رفتار، جعل داده‌ها و یا ارسال و انتشار دوگانه، پیروی کرده‌اند. همچنین این مقاله حاصل یک کار تحقیقاتی اصیل بوده و

اجازه انتشار مقاله

نویسنده (گان) با چاپ این مقاله به‌صورت دسترسی باز موافقت کرده و کلیه حقوق استفاده از محتوا، جدول‌ها، شکل‌ها، تصویرها و غیره را به ناشر واگذار می‌کنند.

References

- Ahmad, H. M., Alafari, H. A., Fiaz, S., Alshaya, D. S., Toor, S., Ijaz, M., Rasool, N., Attia, K. A., Zaynab, M., & Azmat, S. (2022). Genome-wide comparison and identification of myosin gene family in *Arabidopsis thaliana* and *Helianthus annuus*. *Heliyon*, 8(12), e12070. doi: [10.1016/j.heliyon.2022.e12070](https://doi.org/10.1016/j.heliyon.2022.e12070).
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., Duvaud, S., Flegel, V., Fortier, A., & Gasteiger, E. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Research*, 40(W1), W597-W603. doi: [10.1093/nar/gks400](https://doi.org/10.1093/nar/gks400).
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W., & Noble, W. S. (2009). MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Research*, 37, W202-W208. doi: [10.1093/nar/gkp335](https://doi.org/10.1093/nar/gkp335).
- Bolser, D., Staines, D. M., Pritchard, E., & Kersey, P. (2016). Ensembl plants: integrating tools for visualizing, mining, and analyzing plant genomics data. *Plant Bioinformatics: Methods & Protocols*, 1374, 115-140. doi: [10.1007/978-1-4939-3167-5_6](https://doi.org/10.1007/978-1-4939-3167-5_6).
- Borrill, P., Ramirez-Gonzalez, R., & Uauy, C. (2016). ExpVIP: A customizable RNA-seq data analysis and visualization platform. *Plant Physiology*, 170(4), 2172-2186. doi: [10.1104/pp.15.01667](https://doi.org/10.1104/pp.15.01667).
- Cannon, S. B., Mitra, A., Baumgarten, A., Young, N. D., & May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biology*, 4, 10. doi: [10.1186/1471-2229-4-10](https://doi.org/10.1186/1471-2229-4-10).
- Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., Liu, Y., Feng, J., Chen, H., & He, Y. (2023). TBtools-II: A “one for all, all for one” bioinformatics platform for biological big-data mining. *Molecular Plant*, 16(11), 1733-1742. doi: [10.1016/j.molp.2023.09.010](https://doi.org/10.1016/j.molp.2023.09.010).
- Chen, L., Meng, J., He, X. L., Zhang, M., & Luan, Y. S. (2019). Solanum lycopersicum microRNA1916 targets multiple target genes and negatively regulates the immune response in tomato. *Plant, Cell & Environment*, 42(4), 1393-1407. doi: [10.1111/pce.13468](https://doi.org/10.1111/pce.13468).
- Dobritsa, A. A., Nishikawa, S.-I., Preuss, D., Urbanczyk-Wochniak, E., Sumner, L. W., Hammond, A., Carlson, A. L., & Swanson, R. J. (2009). LAP3, a novel plant protein required for pollen development, is essential for proper exine formation. *Sexual Plant Reproduction*, 22(3), 167-177. doi: [10.1007/s00497-009-0101-8](https://doi.org/10.1007/s00497-009-0101-8).
- Fabbri, M., Delp, G., Schmidt, O., & Theopold, U. (2000). Animal and plant members of a gene family with similarity to alkaloid-synthesizing enzymes. *Biochemical & Biophysical Research Communications*, 271(1), 191-196. doi: [10.1006/bbrc.2000.2598](https://doi.org/10.1006/bbrc.2000.2598).
- Facchini, P. J., Bird, D. A., & St-Pierre, B. (2004). Can *Arabidopsis* make complex alkaloids? *Trends in Plant Science*, 9, 116-122. doi: [10.1016/j.tplants.2004.01.004](https://doi.org/10.1016/j.tplants.2004.01.004).
- Giraldo, P., Benavente, E., Manzano-Agugliaro, F., & Gimenez, E. (2019). Worldwide research trends on wheat and barley: A bibliometric comparative analysis. *Agronomy*, 9, 352. doi: [10.3390/agronomy9070352](https://doi.org/10.3390/agronomy9070352).
- Gu, L., Cao, Y., Chen, X., Wang, H., Zhu, B., Du, X., & Sun, Y. (2023). The genome-wide identification, characterization, and expression analysis of the strictosidine synthase-like family in maize (*Zea mays* L.). *International Journal of Molecular Sciences*, 24, 14733. doi: [10.3390/ijms241914733](https://doi.org/10.3390/ijms241914733).

- Guo, J., Wu, J., Ji, Q., Wang, C., Luo, L., Yuan, Y., Wang, Y., & Wang, J. (2008). Genome-wide analysis of heat shock transcription factor families in rice and Arabidopsis. *Journal of Genetics & Genomics*, 35, 105-118. doi: [10.1016/S1673-8527\(08\)60016-8](https://doi.org/10.1016/S1673-8527(08)60016-8).
- Hernandez-Garcia, C. M., & Finer, J. J. (2014). Identification and validation of promoters and cis-acting regulatory elements. *Plant Science*, 217, 109-119. doi: [10.1016/j.plantsci.2013.12.007](https://doi.org/10.1016/j.plantsci.2013.12.007).
- Hicks, M. A., Barber, A. E., & Babbitt, P. C. (2013). The Nucleophilic Attack Six-Bladed β -Propeller (N6P) Superfamily. In: Orengo, C., & Bateman, A. (Eds.). *Protein Families: Relating Protein Sequence, Structure, and Function*. Wiley Online Library. pp. 125-158. doi: [10.1002/9781118743089.ch6](https://doi.org/10.1002/9781118743089.ch6).
- Hicks, M. A., Barber, A. E., Giddings, L. A., Caldwell, J., O'connor, S. E., & Babbitt, P. C. (2011). The evolution of function in strictosidine synthase-like proteins. *Proteins: Structure, Function, & Bioinformatics*, 79(11), 3082-3098. doi: [10.1002/prot.23135](https://doi.org/10.1002/prot.23135).
- Jan, R., Asaf, S., Numan, M., Lubna, & Kim, K.-M. (2021). Plant secondary metabolite biosynthesis and transcriptional regulation in response to biotic and abiotic stress conditions. *Agronomy*, 11(5), 968. doi: [10.3390/agronomy11050968](https://doi.org/10.3390/agronomy11050968).
- Kibble, N. A., Sohani, M. M., Shirley, N., Byrt, C., Roessner, U., Bacic, A., Schmidt, O., & Schultz, C. J. (2009). Phylogenetic analysis and functional characterisation of strictosidine synthase-like genes in *Arabidopsis thaliana*. *Functional Plant Biology*, 36, 1098-1109. doi: [10.1071/FP09104](https://doi.org/10.1071/FP09104).
- Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology & Evolution*, 33(7), 1870-1874. doi: [10.1093/molbev/msw054](https://doi.org/10.1093/molbev/msw054).
- Kutchan, T., Hampp, N., Lottspeich, F., Beyreuther, K., & Zenk, M. (1988). The cDNA clone for strictosidine synthase from *Rauvolfia serpentina* DNA sequence determination and expression in *Escherichia coli*. *FEBS Letters*, 237, 40-44. doi: [10.1016/0014-5793\(88\)80167-4](https://doi.org/10.1016/0014-5793(88)80167-4).
- Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., Rouzé, P., & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, 30(1), 325-327. doi: [10.1093/nar/30.1.325](https://doi.org/10.1093/nar/30.1.325).
- Letunic, I., & Bork, P. (2018). 20 years of the SMART protein domain annotation resource. *Nucleic Acids Research*, 46(D1), D493-D496. doi: [10.1093/nar/gkx922](https://doi.org/10.1093/nar/gkx922).
- Liu, H., Lyu, H. M., Zhu, K., Van De Peer, Y., & Cheng, Z. M. (2021). The emergence and evolution of intron-poor and intronless genes in intron-rich plant gene families. *The Plant Journal*, 105, 1072-1082. doi: [10.1111/tpj.15088](https://doi.org/10.1111/tpj.15088).
- Liu, M., Dong, H., Wang, M., & Liu, Q. (2020). Evolutionary divergence of function and expression of laccase genes in plants. *Journal of Genetics*, 99, 1-16. doi: [10.1007/s12041-020-1184-0](https://doi.org/10.1007/s12041-020-1184-0).
- Magwanga, R. O., Lu, P., Kirungu, J. N., Lu, H., Wang, X., Cai, X., Zhou, Z., Zhang, Z., Salih, H., & Wang, K. (2018). Characterization of the late embryogenesis abundant (LEA) proteins family and their role in drought stress tolerance in upland cotton. *BMC Genetics*, 19, 6. doi: [10.1186/s12863-017-0596-1](https://doi.org/10.1186/s12863-017-0596-1).
- Qanmber, G., Liu, J., Yu, D., Liu, Z., Lu, L., Mo, H., Ma, S., Wang, Z., & Yang, Z. (2019). Genome-wide identification and characterization of the PERK gene family in *Gossypium hirsutum* reveals gene duplication and functional divergence. *International Journal of Molecular Sciences*, 20, 1750. doi: [10.3390/ijms20071750](https://doi.org/10.3390/ijms20071750).
- Sarcheshmeh, M. K., Abedi, A., & Aalami, A. (2023). Genome-wide survey of catalase genes in *Brassica rapa*, *Brassica oleracea*, and *Brassica napus*: Identification, characterization, molecular evolution, and expression profiling of BnCATs in response to salt and cadmium stress. *Protoplasma*, 260, 899-917. doi: [10.1007/s00709-022-01822-6](https://doi.org/10.1007/s00709-022-01822-6).
- Sievers, F., & Higgins, D. G. (2014). Clustal omega. *Current Protocols in Bioinformatics*, 48, 313. doi: [10.1002/0471250953.bi0313s48](https://doi.org/10.1002/0471250953.bi0313s48).
- Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., & Kasprzyk, A. (2009). BioMart—biological queries made easy. *BMC Genomics*, 10, 1-12. doi: [10.1186/1471-2164-10-22](https://doi.org/10.1186/1471-2164-10-22).
- Sohani, M., Schenk, P., Schultz, C., & Schmidt, O. (2009). Phylogenetic and transcriptional analysis of a strictosidine synthase-like gene family in *Arabidopsis thaliana* reveals involvement in plant defence responses. *Plant Biology*, 11, 105-117. doi: [10.1111/j.1438-8677.2008.00139.x](https://doi.org/10.1111/j.1438-8677.2008.00139.x).
- Theopold, U., Samakovlis, C., Erdjument-Bromage, H., Dillon, N., Axelsson, B., Schmidt, O., Tempst, P., & Hultmark, D. (1996). Helix pomatia lectin, an inducer of *Drosophila* immune

- response, binds to hemomucin, a novel surface mucin. *Journal of Biological Chemistry*, 271, 12708-12715. doi: [10.1074/jbc.271.22.12708](https://doi.org/10.1074/jbc.271.22.12708).
- Wang, R., Zhao, W., Yao, W., Wang, Y., Jiang, T., & Liu, H. (2023a). Genome-wide analysis of strictosidine synthase-like gene family revealed their response to biotic/abiotic stress in poplar. *International Journal of Molecular Sciences*, 24(12), 10117. doi: [10.3390/ijms241210117](https://doi.org/10.3390/ijms241210117).
- Wang, T., Song, H., Zhang, B., Lu, Q., Liu, Z., Zhang, S., Guo, R., Wang, C., Zhao, Z., & Liu, J. (2018). Genome-wide identification, characterization, and expression analysis of superoxide dismutase (SOD) genes in foxtail millet (*Setaria italica* L.). *3 Biotech*, 8(12), 486. doi: [10.1007/s13205-018-1502-x](https://doi.org/10.1007/s13205-018-1502-x).
- Wang, Y., Ruan, Q., Zhu, X., Wang, B., Wei, B., & Wei, X. (2023b). Identification of alfalfa *SPL* gene family and expression analysis under biotic and abiotic stresses. *Scientific Reports*, 13, 84. doi: [10.1038/s41598-022-26911-7](https://doi.org/10.1038/s41598-022-26911-7).
- Wu, X., Zhou, C., Li, X., Lin, J., Aguila, L. C. R., Wen, F., & Wang, L. (2023). Genome-wide identification and immune response analysis of mitogen-activated protein kinase cascades in tea geometrid, *Ectropis griseascens* Warren (Geometridae, Lepidoptera). *BMC Genomics*, 24, 344. doi: [10.1186/s12864-023-09446-7](https://doi.org/10.1186/s12864-023-09446-7).
- Xu, G., Guo, C., Shan, H., & Kong, H. (2012). Divergence of duplicate genes in exon–intron structure. *Proceedings of the National Academy of Sciences*, 109, 1187-1192. doi: [10.1073/pnas.1109047109](https://doi.org/10.1073/pnas.1109047109).
- Yamazaki, Y., Urano, A., Sudo, H., Kitajima, M., Takayama, H., Yamazaki, M., Aimi, N., & Saito, K. (2003). Metabolite profiling of alkaloids and strictosidine synthase activity in camptothecin producing plants. *Phytochemistry*, 62(3), 461-470. doi: [10.1016/S0031-9422\(02\)00543-5](https://doi.org/10.1016/S0031-9422(02)00543-5).
- Yan, F., Zhou, H., Yue, M., Yang, G., Li, H., Zhang, S., & Zhao, P. (2019). Genome-wide identification and transcriptional expression profiles of the F-box gene family in common walnut (*Juglans regia* L.). *Forests*, 10(3), 275. doi: [10.3390/f10030275](https://doi.org/10.3390/f10030275).
- Yoshida, S., Ito, M., Nishida, I., & Watanabe, A. (2001). Isolation and RNA gel blot analysis of genes that could serve as potential molecular markers for leaf senescence in *Arabidopsis thaliana*. *Plant & Cell Physiology*, 42(2), 170-178. doi: [10.1093/pcp/pce021](https://doi.org/10.1093/pcp/pce021).
- Yoshioka, Y., Suzuki, G., Zayasu, Y., Yamashita, H., & Shinzato, C. (2022). Comparative genomics highlight the importance of lineage-specific gene families in evolutionary divergence of the coral genus, *Montipora*. *BMC Ecology & Evolution*, 22, 71. doi: [10.1186/s12862-022-02023-8](https://doi.org/10.1186/s12862-022-02023-8).
- Yu, J., Yuan, Y., Dong, L., & Cui, G. (2023). Genome-wide investigation of NLP gene family members in alfalfa (*Medicago sativa* L.): Evolution and expression profiles during development and stress. *BMC genomics*, 24, 320. doi: [10.1186/s12864-023-09418-x](https://doi.org/10.1186/s12864-023-09418-x).
- Yuan, Y., Yin, X., Han, X., Han, S., Li, Y., Ma, D., Fang, Z., Yin, J., & Gong, S. (2023). Genome-wide identification, characterization and expression analysis of the *TaDUF724* gene family in wheat (*Triticum aestivum*). *International Journal of Molecular Sciences*, 24(18), 14248. doi: [10.3390/ijms241814248](https://doi.org/10.3390/ijms241814248).
- Zhang, L., Li, S., Fang, X., An, H., & Zhang, X. (2023). Genome-wide analysis of LysM gene family members and their expression in response to *Colletotrichum fructicola* infection in Octoploid strawberry (*Fragaria x ananassa*). *Frontiers in Plant Science*, 13, 1105591. doi: [10.3389/fpls.2022.1105591](https://doi.org/10.3389/fpls.2022.1105591).
- Zou, T., Li, S., Liu, M., Wang, T., Xiao, Q., Chen, D., Li, Q., Liang, Y., Zhu, J., & Liang, Y. (2017). An atypical strictosidine synthase, *OsSTRL2*, plays key roles in anther development and pollen wall formation in rice. *Scientific Reports*, 7, 6863. doi: [10.1038/s41598-017-07064-4](https://doi.org/10.1038/s41598-017-07064-4).