

## تجزیه و تحلیل کل رونوشت RNA جهت شناسایی RNA های غیر کد کننده تنظیمی و اهداف آنها در پاسخ به تنش خشکی در گندم (*Triticum aestivum*. L)

محمد علی زارع<sup>۱</sup>، نسیم زرین پنجه<sup>۲\*</sup>، حمیده خلیج<sup>۳</sup>، مصطفی رفیعی پور<sup>۴</sup>، سید رضا میر علیزاده<sup>۵</sup>

<sup>۱</sup> ایران، تهران، دانشگاه پیام نور، گروه بیوتکنولوژی کشاورزی،

<sup>۲</sup> ایران، کرج، مرکز تحقیقات گیاهان دارویی، پژوهشکده گیاهان دارویی

<sup>۳</sup> ایران، تهران، دانشگاه پیام نور، گروه بیوتکنولوژی کشاورزی

<sup>۴</sup> ایران، قزوین، دانشکده علوم، دانشگاه دانش البرز

<sup>۵</sup> ایران، تهران، دانشگاه پیام نور، گروه کشاورزی و محیط زیست

تاریخ دریافت: ۱۴۰۳/۰۵/۱۳ تاریخ پذیرش: ۱۴۰۳/۱۰/۲۹



### چکیده

خشکسالی یکی از عوامل مهم تنش غیر زیستی است، که بر بهره‌وری محصولات زراعی از جمله گندم تأثیر می‌گذارد. با این حال، تحمل به خشکی یک ویژگی پیچیده است که شامل بسیاری از ژن‌ها و مسیرهای سلولی است. شناخت شبکه تنظیم مولکولی گندم در پاسخ به تنش خشکی از اهمیت بالایی در اصلاح مولکولی برخوردار است. RNA های غیر کدکننده با تنظیم بیان ژن بر رشد گیاه و مقاومت در برابر تنش‌های غیرزیستی تأثیر می‌گذارند. برای کمک به درک مکانیسم‌های تنظیمی پیچیده حاکم بر تحمل به خشکی در گندم، ما تجزیه و تحلیل کل توالی رونوشت RNA را در گونه‌ای گندم مقاوم به خشکی انجام دادیم. در مجموع ۲۰۴۶۲ mRNA و ۲۰۹ miRNA با بیان متفاوت شناسایی گردید. از بین ۲۰۴۶۲ ژن با بیان متفاوت، ۱۰۲۵ به عنوان ژن‌های با توانایی مقاومت به تنش خشکی شناسایی گردیدند. ژن‌های هدف شناسایی شده توسط psRNATarget همگی در دسته بندی هستی‌شناسی ژن غنی شدند. نتایج ما مکانیسم‌های مولکولی تحمل به خشکی در گندم را روشن می‌کند و همچنین miRNA های پاسخ‌دهنده به خشکی و ژن‌های هدف آنها را شناسایی می‌کند. این دستاورد ها می‌تواند برای بهبود گندم و گیاهان مرتبط مورد استفاده قرار گیرد، و اطلاعات مفیدی را به عنوان منابع ژنومی ارزشمند در اصلاح مولکولی گندم ارائه کند.

**واژه‌های کلیدی:** تنش خشکی، توالی یابی RNA، گندم، RNA های غیر کدکننده

\* نویسنده مسئول، پست الکترونیکی: [zarinpanjeh@imp.ac.ir](mailto:zarinpanjeh@imp.ac.ir)

### مقدمه

گندم نان یا گندم معمولی (*Triticum aestivum*) از گیاهان گلدار تک‌لپه‌ای یک ساله که به خانواده گندمیان (Gramineae) تعلق دارد. جزء عمده‌ترین گونه‌ای هگزاپلوئیدی است که در سرتاسر دنیا کشت می‌گردد. اهمیت اقتصادی گندم و سهم آن در رژیم غذایی انسان و دام قابل انکار نیست. گندم نان تقریباً ۱۰ هزار سال پیش از تلاقی بین یک گونه تتراپلوئید (AABB) و یک گونه علف بزی دیپلوئید (DD) به وجود آمده است [۳۰]. گندم یکی از گونه‌های بنیادی است که کشت و اهلی کردن آن ارتباط تنگاتنگی با رونق کشاورزی و جوامع آباد داشته است. گندم نان به دلیل عملکرد بالا، کیفیت تغذیه‌ای و فرآوری آن به یکی از محصولات پرکشت تبدیل شده است. تا سال ۲۰۵۰ جمعیت جهان به ۹ میلیارد نفر خواهد رسید. لازمه برآورد نیازهای غذایی این افراد افزایش قابل توجه عملکرد گندم است که از طریق اصلاح ژنتیکی نائل می‌گردد [۲۲]. گیاهان در

محیط‌هایی که دائماً در حال تغییر هستند زندگی می‌کنند که اغلب برای رشد و تکامل آنها نامطلوب یا استرس‌زا هستند. این شرایط محیطی نامطلوب شامل تنش‌های زیستی مانند عفونت پاتوژن‌ها و حمله گیاهخواران، و تنش‌های غیر زیستی مانند خشکی، دمایی، سرمای، کمبود مواد مغذی و وجود مقدار بیش از حد نمک یا فلزات سمی مانند آلومینیوم، آرسنات و کادمیوم موجود در خاک می‌باشد. تنش‌های خشکی، شوری و دمای از عوامل اصلی محیطی موثر بر توزیع جغرافیایی گیاهان در طبیعت هستند که بهره‌وری گیاهان را در کشاورزی محدود و امنیت غذایی را تهدید می‌کنند. با این وجود، توانایی گیاهان برای سازگاری با شرایط محیطی مختلف به طور مستقیم و غیر مستقیم با دو راهبرد شامل، اجتناب از تنش گیاهی و تحمل به تنش گیاهی مرتبط است. اجتناب از تنش گیاهی، یک فاز فیزیولوژیکی غیر فعال مانند دانه‌های بالغ است، در حالی که تحمل به تنش یک تطبیق برگشت پذیر فعال است که به طور کلی به عنوان سازگاری شناخته می‌شود [۱۵]. سازگاری با تنش‌ها به طور اختصاصی از طریق تغییرات عمیق در سطح بیان ژن انجام می‌شود که منجر به تغییر یا اصلاح در ترکیب رونوشت‌های گیاهی، پروتئوم و همچنین متابولوم می‌شود [۱۶]. علاوه بر این، بهبود مقاومت گیاه در برابر تنش‌ها برای بهره‌وری کشاورزی و همچنین برای پایداری محیط‌زیست حیاتی است، زیرا محصولات با مقاومت ضعیف در برابر تنش، آب و کود زیادی مصرف می‌کنند و در نتیجه محیط را به شدت تحت فشار قرار می‌دهند.

خشکسالی یکی از مهم‌ترین عوامل محدود کننده تولید محصولات کشاورزی است که عملکرد محصول را به شدت تحت تاثیر قرار می‌دهد [۱۴]. علاوه بر این، خشکی به عنوان یکی از عوامل اصلی بازدارنده در روند رشد گیاه، می‌تواند مانع تنفس گیاه، فتوسنتز و حرکت روزنه‌ها شود و در نتیجه رشد و متابولیسم فیزیولوژیکی گیاه را تحت تاثیر قرار دهد. در پاسخ به تنش خشکی، گیاهان مکانیسم‌های واکنش به خشکی خود مانند تغییرات مورفولوژیکی و ساختاری، بیان ژن‌های مقاوم به خشکی، سنتز هورمون‌ها و مواد تنظیم‌کننده اسمزی را جهت کاهش تنش خشکی فعال می‌کنند. مطالعات اخیر نشان داده‌اند که RNA های غیرکدکننده (Non-coding RNAs) طولانی و کوچک تعدیل‌کننده‌های مهم تحمل به خشکی در گیاهان هستند. شبکه بین RNA های غیر کدکننده و ژن‌های هدف آنها توسط سایر اجزای آنزیمی موجود در سلول کنترل می‌شود. پیشرفت در تجزیه و تحلیل با توان بالا، مانند توالی‌یابی RNA (High-throughput analysis)، بیوانفورماتیک، تحقیقات علمی را تسریع کرده است [۳۴]. RNAهایی که از رونویسی DNA به وجود می‌آیند، اما به پروتئین ترجمه نمی‌شوند، به عنوان RNA های غیر کد کننده تعریف می‌شوند، نشان داده شده که RNA های غیر کدکننده در مجموعه مختلفی از مکانیسم‌های تنظیمی و اپی ژنتیکی نقش دارند و اثرات تنظیمی خود را عمدتاً از طریق تطبیق جفت باز مکمل با RNA های پیام رسان (mRNA) اعمال می‌کنند. به اثبات رسیده که تجزیه و تحلیل توالی، ابزار مهمی برای بررسی دقیق تفاوت در پاسخ به تنش، بین گیاهان حساس و مقاوم است [۹].

RNA های غیر کدکننده هم از نظر منشأ و هم از نظر عملکرد متنوع هستند. تکثیر، تکامل از عناصر قابل انتقال (Transposable Elements) (TEs)، ساختارهای سنجاق سر تصادفی، توالی‌های کدکننده پروتئین شبه ژن، تکثیر ویروس‌های RNA دار، RNA های دو رشته‌ای مناطق هتروکروماتین، و تکرارهای DNA، نظریه‌هایی برای منشأ ncRNA های مختلف مانند miRNA ها، siRNA ها، piRNA ها و lncRNA ها هستند [۵].

میکروار ان ای ها (miRNAs) مولکول‌های کوچک درون زاد از کلاس یوکاریوت‌ها، اما اساسی هستند، که طول آن‌ها عمدتاً ۲۰-۲۲ نوکلئوتید است. miRNA ها نقش بیولوژیکی غیر قابل جایگزینی در گیاهان دارند و می‌تواند مراحل مختلف

تکامل، تنظیمات پیچیده رشد و همچنین پاسخ به تنش‌های زیستی یا غیرزیستی در گیاه را تحت تاثیر قرار دهند [۲]. میکرو RNA ها از ژن‌های miRNA (ژن‌های MIR)، توسط RNA پلیمراز دو (II) و RNA پلیمراز سه (III) تولید می‌شوند. این فرایند شامل پردازش و سنتز توسط پروتئین یک شبه دایسر (DCL1) است. سپس دوبلکس miRNA-miRNA توسط انتقال دهنده هستی (HST) به سیتوپلاسم منتقل می‌گردد و در سیتوپلاسم با کمپلکس خاموش‌کننده القا شده توسط RNA (RISC) (RNA-induced silencing complex) که حاوی پروتئین آرگونات یک (Argonaute) (AGO1) می‌باشد، ادغام می‌گردد. این کمپلکس خاموشی ژن را از طریق برش mRNA و سرکوب ترجمه کنترل می‌کند [۶]. RNA های کوچک مداخله گر (siRNA) (short interfering RNA) بسته به نحوه عمل و همچنین بیوژنز به کلاس‌های، ترانس-اکتین siRNA ها (trans-acting siRNAs) (tasiRNA)، siRNA های هتروکروماتیک (heterochromatic siRNAs) (hc-siRNA)، و siRNA های آنتی سنس طبیعی (natural antisense siRNAs) (nat-siRNAs) گروه بندی می‌شوند. siRNA های هتروکروماتیک، مکانیسم تنظیمی مشابه miRNA ها دارند و در جایگاه‌های مستقل از محل سنتز خود عمل می‌کنند. در همان زمان، siRNA های هتروکروماتیک به صورت سیس در مناطق هدف یا همولوگ خود به طور کارآمدی عمل می‌کند [۳۸]. با این حال، siRNA های آنتی سنس طبیعی به نوع سیس و ترانس دسته بندی می‌شوند. تولید نوع سیس شامل رونویسی از رشته‌های مخالف دو RNA در یک مکان مشابه است. به طور همزمان، ترانس شامل رونوشت‌های RNA از جایگاه‌های چندگانه می‌باشد [۱۲]. RNA های غیر کد کننده بلند، رونوشت‌هایی هستند که طول آنها بیش از ۲۰۰ جفت باز است به پروتئین ترجمه نمی‌شوند. آنها به دسته‌های شامل، سیس آنتی سنس طبیعی (cis-natural antisense) (cis-NATs)، رونوشت‌های ترانس آنتی سنس طبیعی (trans-natural antisense transcripts) (trans-NATs) شبه ژن (pseudogenes) سنس یا آنتی سنس (منشأ رشته‌ها)، واگرا، یا همگرا (جهت گیری رونویسی)، و به عنوان اینترونیک یا بین ژنی (جایگاه) طبقه بندی می‌شوند. آنها توسط RNA پلی‌مراز‌های (II, III, IV و V) رونویسی می‌شوند، و تنظیم هدف خود را از طریق روش‌های متعددی از جمله بازسازی کروماتین، سرکوب رونویسی، تقلید، پیرایش RNA و تقویت‌کننده‌های رونویسی کنترل می‌کنند [۱].

اخیراً یافته‌های تحقیقاتی متعدد بیان متفاوت RNA ها غیر کد کننده را در شرایط تنش نشان داده‌اند. این RNA ها ی غیر کد کننده بیان ژن را از طریق شبکه‌های سلولی مرتبط با عملکرد مستقیم در پاسخ به استرس، کنترل می‌کنند [۱۱].

میکرو RNA ها به دلیل عملکرد گسترده اهداف خود، به اهداف جدیدی برای بهبود بهره‌وری گیاهان و تحمل به تنش غیرزیستی تبدیل شده‌اند. به نظر می‌رسد تنظیم میکرو RNA ها به نقش آنها بستگی دارد. میکرو RNA های که در پاسخ به تنش غیر زیستی افزایش می‌یابند ممکن است بیان ژن‌های هدف خود را کاهش دهند، و احتمالاً تنظیم‌کننده‌های منفی در تحمل به تنش هستند (به عنوان مثال، سرکوب‌کننده‌های ژنی پاسخ‌دهنده به استرس). علاوه بر این، میکرو RNA های که تحت تنش کاهش می‌یابند، ممکن است تنظیم‌کننده مثبت در تحمل به تنش باشد [۲۰]. به طور کلی، بیان میکرو RNA ها در پاسخ به تنش غیرزیستی خاص افزایش یا کاهش می‌یابد. علاوه بر این، ممکن است یک میکرو RNA در یک گیاه در پاسخ به تنش‌های غیرزیستی یکسان، هم افزایش و هم کاهش بیان داشته باشد [۱۷]. بنابراین، مسیرهای وابسته به siRNA نقش مهمی در کنترل انتقال ترانسپوزون‌ها در شرایط تنش غیرزیستی ایفا می‌کند. در شرایط کم آبی، ژن القایی آبسزیک اسید (ABA) مرتبط با کم آبی، منجر به سنتز یک RNA مداخله گر می‌شود که نتیجه آن گیاه را به تحمل کم آبی مقاوم می‌کند.

کند. تجزیه و تحلیل نرتن بلات در جوانه گندم نشان داد که بیان چهار RNA مداخله گر به طور قابل توجهی در پاسخ به سرما، گرما، نمک و تنش های کم آبی تغییر کرده است [۳۲].

lncRNA ها عملکرد خود را در پاسخ به تنش های غیرزیستی به روش های مختلف انجام می دهند. اول، برخی از lncRNA های گیاهی از طریق تقلید هدف در پاسخ به تنش غیرزیستی شرکت می کنند. آنها به عنوان RNA های درون زا (ceRNA) با ژن های هدف میکرو RNA ها رقابت می کنند، و از این طریق تعاملات میکرو RNA ها را با اهداف خود مسدود می کنند. lncRNA های آنتی سنس گیاهی با رشته سنس mRNA ها در طی پاسخ به تنش غیرزیستی تعامل دارند. چنین فعل و انفعالاتی منجر به تشکیل دو دوپلکس RNA دو رشته ای می شود، بنابراین بر بیان ژن در رشته مقابل تأثیر می گذارد. برخی از lncRNA ها از طریق مسیر میتلاسیون مستقیم DNA توسط RNA (RNA-directed DNA Methylation) (RdDM) یکسری هم از طریق بازسازی کروماتین و یکسری هم به عنوان پیش ساز برای میکرو RNA ها و RNA های مداخله گر، در پاسخ به تنش های غیر زیستی عمل می کنند [۳۱]. در این تحقیق هدف بررسی و آنالیز پروفایلی از RNA های غیر کدکننده و تعامل آنها با ژن های هدفشان در تنش خشکی است.

## مواد و روشها

### طرح آزمایش

در این طرح جهت مطالعه توالی ترانسکریپتوم نمونه های گندم نان تحت تنش خشکی و نمونه های کنترل از داده های بدست آمده از پلتفرم ایلومینا (Illumina) استفاده گردید. داده های این پلتفرم از پایگاه داده های NCBI تهیه گردید، پس از تعیین و کنترل کیفیت داده ها، سپس آنها با ژنوم مرجع همردیف شدند. در مرحله بعد تجزیه و تحلیل بر روی بیان افتراقی ژن ها صورت گرفت و در نهایت، ژن های با اختلاف بیان و معنی دار از نظر آماری و اثر گذار در تنش خشکی گندم شناسایی و گزارش گردید.

### دریافت داده ها از پایگاه داده های NCBI

داده های این طرح تحقیقی شامل دو گروه کنترل و تحت تنش خشکی که هر کدام شامل چهار نمونه بودند و از پایگاه داده های NCBI تهیه گردیدند. به طور مختصر مراحل کار به این صورت بود، که در دیتابیس های SRA، پایگاه داده های NCBI، و از *T. aestivum* Drought stress جستجو گردید.

### کنترل کیفیت خوانش ها

در این طرح جهت کنترل کیفیت خوانش ها از نرم افزار FastQC در محیط لینوکس، به دلیل در دسترس بودن، کارکردن آسان و سریع بودن در اجرا، استفاده گردید. مزیت این نرم افزار نسبت به بقیه نرم افزار ها شامل این است که می توان داده های با فرمت های SAM، BAM یا FastQ وارد کرد، همچنین ارزیابی سریع با استفاده از نمودارها و جداول خلاصه شده از اینکه

کدام داده‌ها با کیفیت یا بی کیفیت هستند نیز مورد بررسی قرار می‌گیرد و حاصل خروجی این نرم افزار یک فایل HTML به صورت Zip است.

### همردیفی خوانش‌ها با ژنوم مرجع

داده‌های بدست آمده از نمونه‌ها با ژنوم مرجع ([https://plants.ensembl.org/Triticum\\_aestivum/Info/Index](https://plants.ensembl.org/Triticum_aestivum/Info/Index)) (بدست آمده از سایت ensembl plant) و با توالی‌های miRNA در سایت miRbase (<https://mirbase.org>) همردیف گردیدند. جهت همردیفی توالی‌ها از نرم افزار hisat2 استفاده گردید (<http://daehwankimlab.github.io/hisat2/download>).

### تبدیل فایل‌های SAM (Sequence Alignment Map) به BAM (Binary Alignment Map)

فایل‌های SAM/BAM معمولاً خروجی‌های ابزارهای همردیف‌سازها می‌باشد که نیازمند تبدیل فایل SAM به فایل BAM، مرتب‌سازی و نمایه‌سازی یا ادغام هستند. در این طرح از نرم افزار SAMtools (<https://sourceforge.net/projects/samtools>) جهت تبدیل فایل SAM به BAM استفاده گردید، که شامل مرتب‌سازی همردیف‌ها با فرمت BAM است (هدف صرفه‌جویی در فضای ذخیره‌ای است). فایل‌های BAM، حاوی اطلاعاتی شامل، نام و طول توالی مرجع، برنامه‌ای که فایل را ایجاد کرده و مرتب‌سازی می‌باشد.

### مرتب‌سازی (Sorting) و نمایه‌سازی (Indexing) فایل‌های BAM

بعد از اینکه فایل‌های SAM به BAM تبدیل گردیدند در مرحله بعد می‌بایست فایل‌های BAM مرتب‌سازی (جهت دسترسی کارآمدتر به داده‌ها) و نمایه‌سازی (اجازه استخراج سریع همردیف‌های که روی مناطق ویژه ژنوم قرار دارند را می‌دهد) گردند. برای انجام این مراحل از نرم افزار SAMtool در محیط لینوکس استفاده گردید. داده‌های موجود در فایل BAM بر اساس کروموزوم-کانتینگ و داربست‌های ژنوم مرجع مورد استفاده در این طرح مرتب شده‌اند.

### بررسی ویژگی‌های داده‌ها

در این طرح از نرم افزار FeatureCounts برای بررسی داده‌های که از مراحل قبل تهیه گردیده بودند استفاده گردید (<https://sourceforge.net/projects/subread/files/subread-2.0.3>). از این نرم افزار جهت نسبت دادن خوانش‌ها به نواحی مختلف ژنوم، شناسایی توالی‌های ایترونی، آگزونی و واریانت‌های مختلف استفاده می‌شود. داده‌های ورودی به این نرم افزار شامل یک یا چند فایل همردیفی خوانش‌ها با فرمت SAM و BAM و فایل‌های با فرمت GFF یا SAF که شامل لیستی از ویژگی‌های ژنومی است. خروجی این نرم افزار یک فایل TXT است که شامل اطلاعاتی از قبیل نام ژن، کروموزوم، طول، اندازه و غیره می‌باشد.

## آنالیز بیان افتراقی ژن‌ها (Differentially Expressed Genes (DEGs))

آنالیز بیان افتراقی شامل شناسایی ژن‌ها است که تحت شرایط مختلف آزمایشی قرار گرفته‌اند و تحت این شرایط بیان متفاوتی دارند. جهت انجام آنالیز بیان افتراقی ژن‌ها نرم‌افزارها و محیط‌های برنامه‌نویسی متعددی وجود دارد، که یکی از این برنامه‌ها برنامه R است که یک زبان برنامه‌نویسی رایگان جهت محاسبات آماری و گرافیکی و همچنین بیوانفورماتیک می‌باشد. در این طرح از بسته نرم‌افزاری DESeq2 در محیط R استفاده گردید. این نرم‌افزار نرمال‌سازی داخلی را با محاسبه میانگین ژن‌نومتری برای هر ژن در هر نمونه را انجام می‌دهد، سپس تعداد خوانش‌های هر ژن را تقسیم بر تعداد نمونه‌ها می‌کند. برای شروع کار با این نرم‌افزار نیاز به دو فایل ورودی با فرمت TEX داریم که شامل یک فایل خروجی نرم‌افزار FeatureCounts به صورت ماتریکس و یک فایل لیست نمونه‌ها است. در پایان خروجی این بسته نرم‌افزاری یک فایل خروجی با شناسایی بیان افتراقی ژن‌ها در نمونه‌های کنترل و نمونه‌های تحت تنش خشکی بود. بعد از شناسایی بیان ژن‌ها، در پایگاه داده‌های Ensemble شماره ایدی هر ژن و در پایگاه داده‌های DAVID (Functional Annotation) (Bioinformatics Microarray Analysis (ncifcrf.gov)) نماد ژن‌ها (Gene Symbol) شناسایی گردید. در نهایت در محیط زبان برنامه‌نویسی R ژن‌های با p value کمتر از ۰/۰۵ و لگاریتم بر مبنای ۲ بزرگتر مساوی ۰/۰۵ تا کوچکتر مساوی ۰/۰۵- به عنوان ژن‌های با بیان افتراقی معنادار شناسایی گردیدند.

### جستجو ژن‌های با بیشترین و کمترین سطح معناداری در محیط R

ابزارها و روش‌های مختلفی جهت آنالیز بیان افتراقی ژن‌ها موجود می‌باشد. همانطوری که در مراحل قبل بیان گردید ما از بسته نرم‌افزار DESeq2 استفاده کردیم. در این مرحله جهت مشخص کردن سطوح معنادار بودن با نوشتن دستورات شرطی در این محیط، داده‌های با p value بیشتر از ۰/۰۵ حذف گردید و برنامه برای مابقی داده‌ها اجرا گردید، به این صورت که ابتدا داده‌های با مقدار تغییرفولد (fold chang) لگاریتم در مبنای ۲ بیشتر از ۲ یا کوچکتر از ۲- بدست آمدند. سپس جهت بررسی بیشترین تغییرات، داده‌های با p value بیشتر از ۰/۰۵ به عنوان ژن‌های که افزایش بیان داشتند و داده‌های با p value کمتر از ۰/۰۵- به عنوان ژن‌های با کاهش بیان انتخاب گردیدند.

### شناسایی ژن‌های هدف miRNA

جهت شناسایی ژن‌های هدف miRNA از سایت psRNATarget (<http://plantgrn.noble.org/psRNATarget>) استفاده گردید. جهت انجام این کار توالی بالغ miRNA‌ها به عنوان یک قطعه جستجو در کل دیتا بیس‌های cDNA گندم (cDNA, EnsemblPlants, released 31) جستجو گردید.

### رسم نمودار ستونی

جهت نمایش میزان بیان ژن‌ها با بیان متفاوت، نتایج بدست آمده از محیط R وارد نرم‌افزار گراف پد پریزم (v 9.2) گردید، و نمودار ستونی را برای چندین ژن با بیشترین مقدار لگاریتم بر مبنای ۲ برای هر دو گروه افزایش بیان و کاهش بیان انتخاب گردید.

## رسم نمودار نقطه ای یا نمودار Volcano

جهت رسم این نمودار، از نرم افزار گراف پد پریم (v 9.2) بر اساس مقدار  $p$  value و  $\log_2$  استفاده گردید. سطح معنادار بودن داده ها در این نمودار براساس رنگ تقسیم بندی شدند، که این تقسیم بندی شامل، داده های که  $p$  value آنها بیشتر از ۰/۰۵ و لگاریتم بر مبنای ۲ آنها بین ۲+ تا ۲- بود در نمودار با رنگ آبی نمایش داده شدند. داده های که  $p$  value آنها کمتر از ۰/۰۵ و مقدار لگاریتم بر مبنای ۲ آنها بیشتر از ۲+ بود در نمودار با رنگ سبز نمایش داده شدند. همچنین داده های که  $p$  value آنها کمتر از ۰/۰۵ و مقدار لگاریتم بر مبنای ۲ آنها کمتر از ۲- بود نمایش آنها در نمودار با رنگ قرمز بود، که نشان دهنده ژن های با کاهش بیان بودند. در پایان نقاط متمایز هر ژن تعیین گردید و نماد هر ژن در کنار آن نقاط قرار داده شد.

### هستی شناسی ژن (Gene ontology (GO)) و غنی سازی مسیر ژن ها

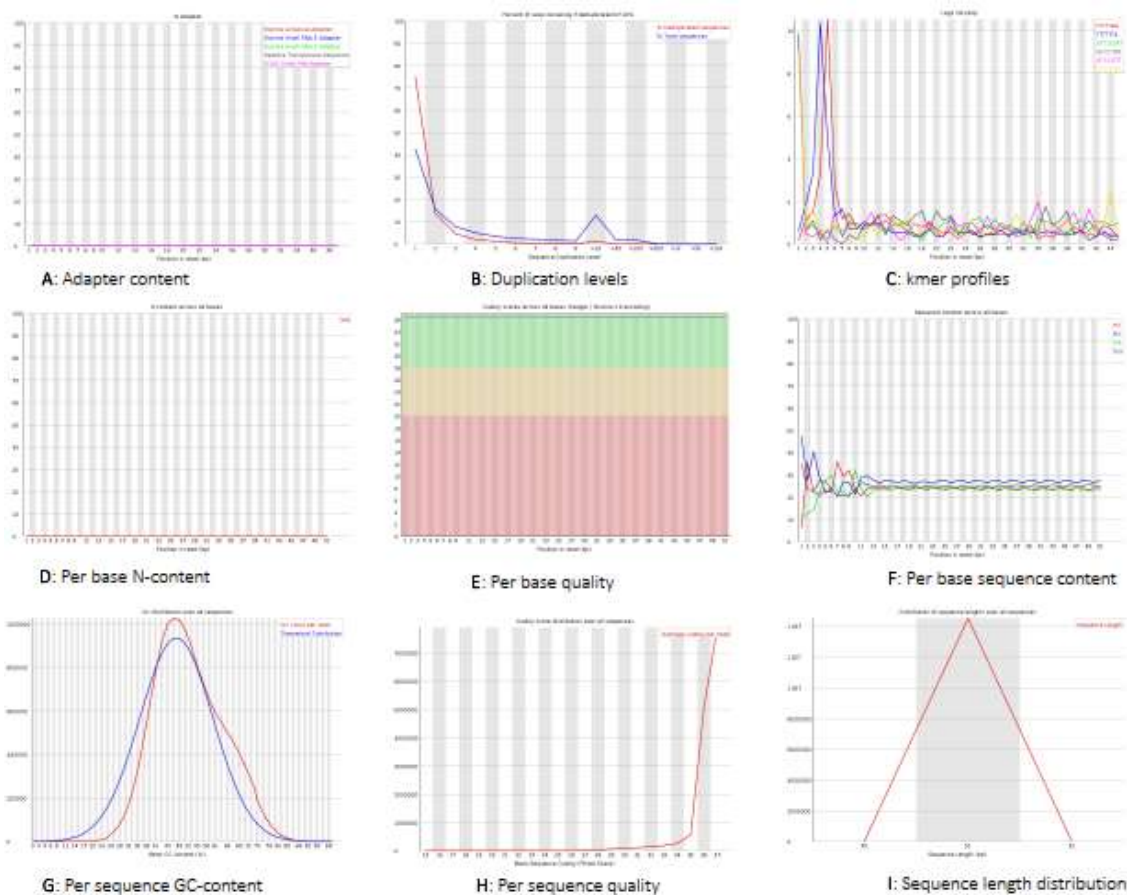
DAVID و pantherdb (<http://www.pantherdb.org>) یک پایگاه داده تحت وب، جهت ارائه عملکرد، اهمیت زیستی یک مجموعه ای از ژن ها می باشد. آنالیزهای غنی سازی عملکردی ژن های با بیان متفاوت، همانند تجزیه و تحلیل GO و تحلیل غنی سازی مسیر های ژنی، با استفاده از DAVID با معیار  $P < 0.05$  cut-off و تعداد غنی سازی ژن بیشتر از ۲ انجام گردید.

## نتایج

### کنترل کیفیت داده ها

در اولین مرحله داده های هر نمونه از نظر کیفیت کنترل گردیدند و نتایج حاصل از این کنترل کیفیت که شامل ۹ مرحله آنالیز بوده است که برای تمام نمونه ها انجام گرفته است و تنها آنالیزهای مربوط به نمونه ERR6110489 نشان داده شده است (شکل ۱). این ۹ مرحله آنالیز کیفیت داده ها به شرح زیر می باشد. شکل A) کنترل از نظر وجود یا عدم وجود آداپتور، به دلیل اینکه پلتفرم ایلامینا از آداپتور ها برای تعیین توالی استفاده می کند، پس بایستی قبل از شروع آنالیز داده ها پیرایش و حذف گردند. پس این نرم افزار حضور و عدم حضور آداپتور ها را به صورت خطوط قرمز رنگ در یک گراف به ما نشان می دهد، به این صورت که اگر در داده ها آداپتور حضور داشته باشد بر اساس میزان آنها به موازات محور  $y$  نمایان می شوند، و در صورت عدم حضور آداپتور این خطوط قرمز رنگ چسبیده به محور  $x$  امتداد می یابد که تایید کننده عدم حضور آداپتور در خوانش ها است. شکل B) کنترل مضاعف شدگی خوانش ها، منظور از مضاعف شدگی یعنی به تعداد دفعات مشخصی که درصدی از خوانش ها در فایل خروجی تکرار گردیده اند. این مضاعف شدگی ها در توالی یابی RNA طبیعی است و نشان دهنده توالی رونوشت های با بیان بالا است. شکل C) کنترل توالی های با تکرارهای چندتایی، این توالی ها می توانند توالی های هموپلیمر (همانند AAAAAAAAAA)، دو نوکلئوتیدی (همانند CACACACACA) و یا سه نوکلئوتیدی (همانند CATCATCATCAT) باشند. متأسفانه این توالی های با تکرار چندتایی حاوی اطلاعات محدودی بوده و مکان یابی صحیح آنها به مرجع مشکل است. شکل D) کنترل بازهای مبهم، در طی فرآیند توالی یابی اگر توالی یک باز شناسایی نگردد، در طی خوانش توالی ها آن باز شناسایی نشده با حرف N نشان داده می شود و در خوانش های که تعداد زیادی N

شناسایی گردیده بایستی حذف گردند. شکل E) کنترل کیفیت خوانش‌ها، کیفیت بازها میزان اطمینان از خوانش‌ها را نشان می‌دهد. نمودار کیفیت توالی‌یابی به ازای هر باز است، که نمایشی از کیفیت بازها را در هر موقعیت بازی در طول خوانش ارائه می‌دهد. در این نمودار محور X و Y به ترتیب نشان‌دهنده امتیاز کیفی و موقعیت خوانش است. زمانی که میانگین خوانش‌ها در قسمت سبز رنگ، نارنجی و قرمز رنگ نمودار قرار بگیرند، به ترتیب نشان‌دهنده کیفیت خوب، متوسط و ضعیف خوانش‌ها است. شکل F) کنترل درصد بازهای فراخوانی شده، این نمودار نشان‌دهنده درصد بازهای فراخوانی شده در هر موقعیت کل خوانش‌ها برای هر چهار نوکلئوتید است. زمانی که مقدار هر باز با مقدار آن در موجود مورد نظر مشابه باشد، نمودار به صورت یک خط راست نمایش داده می‌شود، اما اگر اختلاف بین مقدار بازها بیشتر از ۱٪ در هر موقعیت خوانش باشد، یک هشدار توسط نرم‌افزار صادر می‌گردد. معمولاً بیشترین اختلاف در ابتدا و انتهای نمودار مشاهده می‌گردد.



شکل ۱- گراف‌های مربوط به کنترل کیفیت نمونه ERR6110489

شکل G) کنترل تعداد خوانش‌ها در برابر درصد محتوای GC و توزیع میانگین آنها برای هر باز، معمولاً محتوای GC خوانش‌ها از یک توزیع نرمال تبعیت می‌کنند و نمودار آن به صورت خط راست می‌باشد، اما نمودار توزیع غیر نرمال محتوای GC به صورت نمودار نامعلوم می‌باشد، که نشان دهنده آلودگی کتابخانه است. شکل H) کنترل کل خوانش‌ها در برابر میانگین نمره کیفیت، این نمودار نشان دهنده خلاصه‌ای از میانگین نمره کیفیت نسبت به طول خوانش‌ها در کل خوانش‌ها می‌باشد. شکل I) کنترل توزیع طول خوانش‌ها در تمام توالی‌ها، این کنترل از این جهت مهم است که ایلامینا خوانش‌های با طول‌های غیر یکنواختی ایجاد می‌کند، و پیرایش بر اساس کیفیت یا آداپتور منجر به تولید قطعات بسیار کوتاه می‌گردد. نتایج بررسی کیفیت داده‌های خام توالی RNA نمونه‌های ما، با استفاده از معیارهای آنالیزی این نرم‌افزار به عنوان یک داده با کیفیت مطابقت داشت، که همین امر منجر به اجازه انجام آنالیزهای بعدی برای ادامه کار ما گردید.

### همردیفی خوانش‌ها با ژنوم مرجع

از نرم‌افزار *hisat2* جهت همردیفی خوانش‌ها استفاده گردید که فایل خروجی آن یک فایل *SAM* می‌باشد (شکل ۲). در فایل *SAM* اطلاعاتی در رابطه با کیفیت و موقعیت خوانش‌های همردیف شده با ژنوم مرجع وجود دارد. جهت ادامه کار لازم بود فایل *SAM* به فایل *BAM* توسط نرم‌افزار *SAMtools* تبدیل گردد، فایل *BAM* فایل فشرده شده فایل *SAM* می‌باشد که اطلاعات آن به صورت فشرده‌تر و منظم‌تر هستند و جهت مراحل بعدی آنالیز این فایل مناسب‌تر است (شکل ۳). در ادامه کار با استفاده از همین نرم‌افزار درصد همردیفی و تعداد خوانش‌ها برای هر کدام از نمونه‌ها محاسبه گردید (جدول ۱).

```
@PG ID:hisat2 PN:hisat2 VN:2.1.0 CL:"/usr/bin/hisat2-align-1 --wrapper basic-0 -x
Triticum_aestivum_genome -S ERR6110415.hisat.sam -1 /tmp/3510.inpipe1 -2 /tmp/3510.inpipe2"
ERR6110415.1 99 2A 304979062 60 1S50M = 304979151 141 NGGCGATG
TAAATATAGGTTGAAACATAAACCAATCACCTTACTCTAATACT #FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF A
S:i:-1 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:50 YS:i:0 YT:Z:CP NH:i:1
ERR6110415.1 147 2A 304979151 60 51M = 304979062 -141 ACACGTGT
GCTAGCTTAGGCCACTGACAAAATACTAAGAACACATAAAAAC FFFFFFFF,F:F:F:FF:FF:FFFF:FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF A
S:i:0 ZS:i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-1 YT:Z:CP NH:i:1
ERR6110415.2 83 7B 411055729 60 50M1S = 411055655 -125 CTTACAGT
CCCACTAGGACCTCTCTACCTTGCACTCCCTGTGGAACGTAN FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF# A
S:i:-1 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:50 YS:i:0 YT:Z:CP NH:i:1
ERR6110415.2 163 7B 411055655 60 51M = 411055729 125 CTGAAGTC
GAATGATGAGTTCAGAAATCTGTTGCTTAAGAAGTGATGCTG FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF:FFFF A
S:i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-1 YT:Z:CP NH:i:1
ERR6110415.3 83 6B 271824269 60 50M1S = 271823882 -438 CTTTGAG
AGTATTGCTTCCITTTATGCTATTGAGCGGTGCTTTGCAGAN FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF# A
S:i:-1 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:50 YS:i:0 YT:Z:CP NH:i:1
ERR6110415.3 163 6B 271823882 60 51M = 271824269 438 CACAACAC
TGCGGCAGCAITTTCTTGGAGCTCACAAGCAATTTCTTGCTCC FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF:FFFF A
S:i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-1 YT:Z:CP NH:i:1
ERR6110415.4 99 4A 91789615 60 1S50M = 91789755 192 NAGCAAAG
AATTACAGAACTGAAACAAATAACAAGTTGCCAAAAACGGAAC #FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF A
S:i:-6 XN:i:0 XM:i:1 XO:i:0 XG:i:0 NM:i:1 MD:Z:14A35 YS:i:0 YT:Z:CP NH:i:1
ERR6110415.4 147 4A 91789755 60 51M = 91789615 -192 TAATAACT
CGAAAAAGAGGTTTGTGATCAAAACAATCCAGCTAACCCAG FFFFF:FFFFFFFF:FFFF:FFFFFFFFFFFFFFFF:FFFF,FFF:FFFF A
S:i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-6 YT:Z:CP NH:i:1
```

شکل ۲- یک نمونه فایل *SAM* حاصل خروجی نرم‌افزار *hisat2* جهت همردیفی خوانش‌ها با ژنوم مرجع

|                      |              |  |
|----------------------|--------------|--|
| ERR6110415.bam       | 1,419,887 KB | mrserver:~\$ samtools view ERR6110415.bam   head -n 10                       |
| ERR6110415sorted.bam | 930,385 KB   | ERR6110415.1 99 2A 304979062 60 1550M = 304979151 141 WGCGATGTAA             |
| ERR6110471.bam       | 1,508,779 KB | ATATAGGTTGAAACATAAACCAATCACTTACTTAATACT ##### AS:                            |
| ERR6110471sorted.bam | 939,736 KB   | i:-1 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NH:i:0 MD:Z:50 YS:i:0 YT:Z:CP NH:i:1        |
| ERR6110485.bam       | 1,781,315 KB | ERR6110415.1 147 2A 304979151 60 51M = 304979062 -141 ACACGTGTCT             |
| ERR6110485sorted.bam | 1,109,312 KB | AGCTTAGGCCACTGCACAAATACTAAGAACACATAAAAG ##### F:F:F:FF:FF:FFFFF,FF:##### AS: |
| ERR6110489.bam       | 1,502,965 KB | i:0 ZS:i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-1 YT:Z:CP NH:i:1 |
| ERR6110489sorted.bam | 982,779 KB   | ERR6110415.2 83 7B 411055729 60 50M1S = 411055655 -125 CTTACGTCOC            |
| ERR6110513.bam       | 1,650,776 KB | ACTAGGACTCTCTACCTTGCACTCCCTGCGAAGGTAM ##### AS:                              |
| ERR6110513sorted.bam | 1,103,043 KB | i:-1 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NH:i:0 MD:Z:50 YS:i:0 YT:Z:CP NH:i:1        |
| ERR6110525.bam       | 1,907,098 KB | ERR6110415.2 163 7B 411055655 60 51M = 411055729 125 CTGAAGTCGAA             |
| ERR6110525sorted.bam | 1,279,592 KB | TGATAGGTTGAGAAATCTGTGCTTAAGAGTGTGCTG ##### AS:                               |
| ERR6110529.bam       | 1,451,422 KB | i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-1 YT:Z:CP NH:i:1        |
| ERR6110529sorted.bam | 934,539 KB   | ERR6110415.3 83 6B 271824269 60 50M1S = 271823882 -438 CTTTGGAGAT            |
| ERR6110531.bam       | 1,743,931 KB | ATTGHTCCITTTATGCTATTGAGCGGTGCTTTCGAGAN ##### AS:                             |
| ERR6110531sorted.bam | 1,122,767 KB | i:-1 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NH:i:0 MD:Z:50 YS:i:0 YT:Z:CP NH:i:1        |
|                      |              | ERR6110415.3 163 6B 271823882 60 51M = 271824269 438 CACACACTGC              |
|                      |              | GGCAGCAITTTCTTGGAGCTCACAGCAATTTCTTCTCOC ##### AS:                            |
|                      |              | i:0 XN:i:0 XM:i:0 XO:i:0 XG:i:0 NM:i:0 MD:Z:51 YS:i:-1 YT:Z:CP NH:i:1        |
|                      |              | ERR6110415.4 99 4A 91789615 60 1550M = 91789755 192 NAGCAAGANT               |
|                      |              | TACAGAACTGAAACAAATAACAAGTTGCCAAAAAGGAAAC ##### AS:                           |
|                      |              | i:-6 XN:i:0 XM:i:1 XO:i:0 XG:i:0 NH:i:1 MD:Z:14A35 YS:i:0 YT:Z:CP NH:i:1     |
|                      |              | ERR6110415.4 147 4A 91789755 60 51M = 91789615 -192 TAATACTCGA               |
|                      |              | AAAGAGGTTTTGCTGATCAACAAATATCCAGCTAACAG ##### AS:                             |

شکل ۳- یک نمونه از فایل BAM که توسط نرم افزار SAMtools ایجاد شده

جدول ۱- نتایج حاصل از همردیفی خوانش‌ها با ژنوم مرجع

| نام نمونه    | درصد همردیفی | خوانش    |
|--------------|--------------|----------|
| SAMEA8938573 | 84.66%       | 17874321 |
| SAMEA8938577 | 87.30%       | 18756247 |
| SAMEA8938576 | 91.19%       | 21325871 |
| SAMEA8938572 | 88.56%       | 19856274 |
| SAMEA8938601 | 92.40%       | 24598365 |
| SAMEA8938582 | 89.52%       | 24889500 |
| SAMEA8938589 | 83.20%       | 25023541 |
| SAMEA8938607 | 82.40%       | 23542840 |

### نتایج حاصل از نرم افزار FeatureCount

برای اینکه بتوانیم از فایل BAM بدست آمده از مرحله قبل در نرم افزار FeatureCount استفاده کنیم ابتدا این فایل توسط نرم افزار SAMtools نمایه سازی و مرتب گردیدند. فایل مرتب گردیده توسط نرم افزار FeatureCount مورد ارزیابی قرار گرفت. در این مرحله از آنالیز تعداد خوانش‌ها متصل گردیده به ژنوم، توالی‌های اینترونی، اگزونی، واریانت‌های مختلف و سایر ویژگی‌ها مورد ارزیابی قرار گرفتند (شکل ۴).



خروجی بسته DEseq2 یک فایل اکسل می باشد که اسامی ژن ها براساس شماره Ensembl آنها نمایش داده است برای اینکه در ادامه آنالیز بتوانیم عملکرد و نقش هر ژن را در پایگاه داده ها بررسی نماییم نیاز به تبدیل آنها به اسم ژن است، که جهت تبدیل آنها به اسم ژن (gene symbol) از پایگاه داده های pantherdb به آدرس <http://www.pantherdb.org> و DAVID به آدرس <https://david.ncifcrf.gov> استفاده گردید و بخشی از نتایج آن در جدول ۳ نمایش داده شده است.

جدول ۳- بخشی از نتایج حاصل از جستجوی نام ژن ها در پایگاه DAVID

| ژن های مرتبط | نام ژن  | ENSEMBL_GENE_ID    |
|--------------|---|--------------------|
| RG           | wall-associated receptor kinase 2-like(LOC123067981)            | TraesCS1A02G017700 |
| RG           | uncharacterized LOC123070268(LOC123070268)                      | TraesCS1A02G022800 |
| RG           | RRM domain-containing protein (A0A3B5XUY2_WHEAT)                | TraesCS1A02G064500 |
| RG           | uncharacterized LOC123188981(LOC123188981)                      | TraesCS1A02G064500 |
| RG           | Cupin type-1 domain-containing protein (A0A3B5XV32_WHEAT)       | TraesCS1A02G066100 |
| RG           | uncharacterized LOC123190597(LOC123190597)                      | TraesCS1A02G083300 |
| RG           | Fungal lipase-like domain-containing protein (A0A3B5XWZ9_WHEAT) | TraesCS1A02G094700 |
| RG           | triacylglycerol lipase OBL1-like(LOC123191521)                  | TraesCS1A02G094700 |
| RG           | bidirectional sugar transporter SWEET3a-like(LOC123110707)      | TraesCS1A02G129600 |
| RG           | kiwellin-1-like(LOC123121375)                                   | TraesCS1A02G167700 |

در مرحله آنالیز بیان افتراقی از مجموع حدود ۱۱۲۳۳ ژن که در گروه تنش خشکی نسبت به گروه کنترل شناسایی گردید، بر اساس معیار  $pvalue < 0.05$  و  $\log_2(\text{fold change})$  بزرگتر از ۲- و کوچکتر از ۲+، حدود ۵۴۸۸ ژن افزایش بیان و حدود ۵۷۴۵ ژن کاهش بیان نشان دادند (

جدول) که در پایگاه داده های pantherdb و DAVID اسم برخی از آنها نمایش داده شد و یکسری دیگر از ژن ها به صورت ناشناخته بدون اسم مشخصی نمایش داده شدند، که ما همان اسم Ensembl آنها را گزارش کردیم.

جدول ۴- بخشی از ژن های که براساس pvalue و Log2fold change بیان متفاوتی نشان دادند.

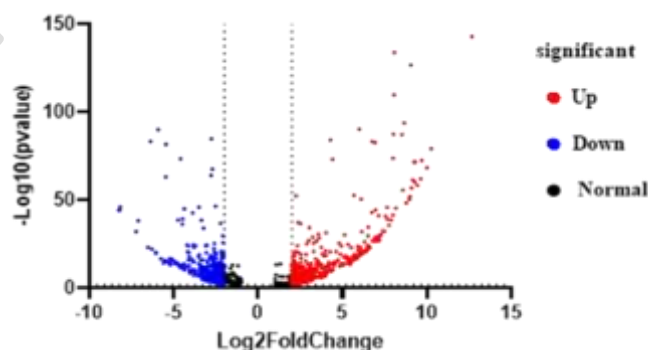
| dif  | Log2FC       | pvalue   | Gene Symbol | Ensembl gene       |
|------|--------------|----------|-------------|--------------------|
| down | -8.208732915 | 3.45E-45 | WRKY        | TraesCS6D02G136200 |
| down | -8.216786734 | 3.03E-45 | -           | TraesCS1B02G350800 |
| up   | 9.683383695  | 1.95E-73 | MYC2        | TraesCS1B02G208000 |
| down | -5.755749622 | 1.52E-17 | -           | TraesCS7A02G064000 |
| up   | 7.861761829  | 1.75E-39 | PYL         | TraesCS1D02G126900 |
| up   | 3.285964436  | 2.37E-26 | -           | TraesCS7D02G073700 |
| down | -3.172375052 | 1.45E-06 | -           | TraesCS7B02G108500 |
| up   | 3.940430114  | 4.89E-14 | -           | tae-MIR1119        |
| up   | 3.771658881  | 1.65E-09 | -           | tae-MIR9654a       |

همچنین در این مطالعه حدود ۱۲۰ miRNA با بیان متفاوت شناسایی گردید که از بین آنها ۸۰ ژن افزایش بیان و ۴۰ ژن کاهش بیان نشان دادند که در جدول ۵ بخشی از آنها نمایش داده شده است. از نمودار Volcano برای نشان دادن بیان افتراقی ژن‌ها استفاده گردید.

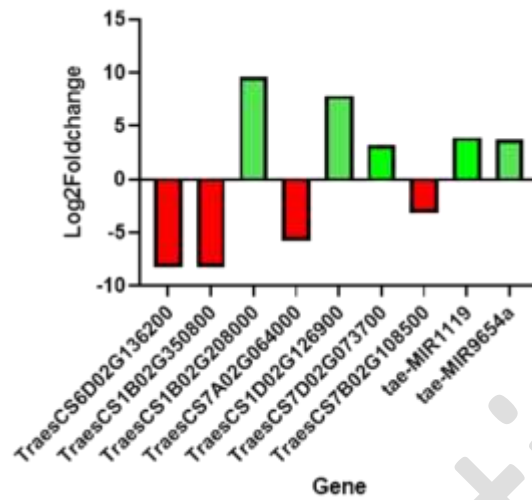
جدول ۵- نمایش بخشی از miRNA ها که براساس pvalue و Log2fold change بیان متفاوتی نشان دادند

| miRNA ID        | Pvalue      | log2FC       | Regulated |
|-----------------|-------------|--------------|-----------|
| tae-miR165a-5p  | 0.012160178 | -1.978340911 | down      |
| tae-miR419      | 1.32E-13    | -4.084396532 | down      |
| tae-miR482a-5p  | 2.95E-10    | -3.666568232 | down      |
| tae-miR530b     | 0.000528321 | -2.335490692 | down      |
| tae-miR858a     | 2.95E-11    | 2.009874     | up        |
| tae-miR9772     | 1.29E-10    | 2.098474     | up        |
| tae-miR1120c-5p | 1.55E-12    | 3.026864912  | up        |
| tae-miR1119     | 4.89E-14    | 3.940430114  | up        |
| tae-miR4391     | 3.91E-08    | -2.934152997 | down      |
| tae-miR5037c    | 0.014521237 | -2.908919153 | down      |

برای رسم این نمودار از پارامترهای  $-\log_{10}(p \text{ value})$  و  $\log_2 \text{ fold change}$  استفاده گردید. این نمودار شامل سه طیف رنگی است، که رنگ قرمز افزایش بیان (یعنی  $p \text{ value} < 0.05$  و  $\text{Log}_2 > 2$ )، رنگ آبی کاهش بیان (یعنی  $p \text{ value} < 0.05$  و  $\text{Log}_2 < -2$ ) و رنگ مشکی عدم تغییر بیان ژن (یعنی  $p \text{ value} > 0.05$ ) را نشان می‌دهند (شکل ۵). منظور از ژن‌های که در منطقه عدم تغییر قرار می‌گیرند، یعنی به طور معناداری دچار افزایش یا کاهش بیان نشده‌اند. همچنین برای برخی از ژن‌ها که بیان متفاوت تری داشتند نمودار ستونی رسم گردید (شکل ۶).



شکل ۵: نمودار Volcano جهت نمایش بیان افتراقی ژن‌ها. از دو پارامتر pvalue و log2FC جهت بیان افتراقی استفاده گردید.



شکل ۶: نمودار میله ای برای نمایش ژن‌های انتخاب‌گردیده که بیشترین افزایش و کاهش بیان در بین ژن‌های دیگر نشان دادند. (بر اساس معیار Log2FoldChange).

### تجزیه و تحلیل هستی‌شناسی ژن

در مجموع از بین ۱۱۲۳۳ ژن بابت تفاوت ۱۰۲۵ ژن براساس دو پارامتر p value و log2FC انتخاب گردیدند. برای این ۱۰۲۵ ژن آنالیز هستی‌شناسی (GO) انجام گرفت، که نتیجه هستی‌شناسی در سه دسته شامل فرایند‌های بیولوژیکی (Biological processes) (BP)، اجزاء سلولی (Cellular components) (CC) و عملکرد سلول (Molecular function) (MF) تقسیم شدند. اصطلاح هستی‌شناسی برای BP شامل فرایندهای سلولی، متابولیک و فرایند‌های آلی هست، فعالیت‌های کاتالیتیکی، اتصال و فاکتور رونویسی جزء MF هستند، برای CC شامل غشاء و اجزای غشای هست. از بین ژن‌های که آنالیز هستی‌شناسی انجام گرفت ۳۹۴ تا از آنها در مسیرهای مختلف KEGG نقشه‌یابی گردیدند، این مسیرها شامل مسیرهای متابولیسم نیتروژن، آلانین، اسپاراتات گلوامات، سنتز دی‌ترپنوئیدها، فنیل پروپنوید و گلیسروفسفولیپیدها هستند. علاوه بر مسیرهای بالا در مسیرهای تعامل پاتوژن-گیاه (Plant-pathogen interaction)، مسیرهای سیگنالینگ گیاهی، متابولیسم نشاسته، ساکاروز و سنتز فلاونوئیدها نقش دارند.

### ژن‌های با بیان متفاوت مرتبط با تنش خشکی

زمانی گیاه تحت عوامل محیطی خارج سلول قرار می‌گیرد، سیگنال‌های به داخل سلول ارسال می‌کند، در ابتدا منجر به فعال شدن مسیرهای تعامل پاتوژن-گیاه و سپس فعال شدن یکسری از پروتئین‌های دیگر (مثل پروتئین کینازها) می‌شود. در این مطالعه یکسری ژن‌های با بیان متفاوت مرتبط با تنش خشکی در مسیر تعامل پاتوژن-گیاه، که شامل پروتئین‌های مقاوم به بیماری، ترئونین کیناز، زایلاناز، کالمودلین و رسپتور کیناز مرتبط با براسینواستروئیدها شناسایی گردید. مسیر سیگنالینگ میتوژن فعال‌کننده پروتئین کیناز (Mitogen-activated protein kinases) (MAPK)، یکی از مهمترین مسیرهای

سیگنالینگ در یوکاریوت‌ها است، که نقش آن هماهنگی پاسخ سلولی جهت رشد و تکامل طبیعی گیاه، پاسخ ایمنی و پاسخ به تنش‌های غیر زیستی است. در مطالعه ما طی غنی‌سازی ژن‌های با بیان متفاوت، ژن‌های که در مسیر MAPK نقش ایفا کردند شامل سرین/ترئونین کیناز (FLS2)، فاکتور رونویسی (WRKY) و رسپتور آبسزیک اسید (PLY) بودند. تنش‌های محیطی خارج از سلول، با تحریک فیتوهورمون‌ها منجر به بالا بردن سازگاری گیاه با تنش‌های محیطی می‌شود. ژن‌های که در این مطالعه در ارتباط با فیتوهورمون‌ها شناسایی گردیدند شامل، جاسمونات حاوی دومین پروتئینی ZIM (Jasmonate domain-containing proteins) (JAZ)، فاکتور‌های اتصال به عناصر پاسخ دهنده به آبسزیک اسید (ABA) (ZIM domain-containing proteins) (ABF)، پروتئین‌های خانواده SAUR، پروتئین (Ethylene-insensitive protein) (EIN2) و فاکتور رونویسی TGA بودند. ژن‌های EIN2، MYC2 و EIN2 هم در مسیر سیگنالینگ MAPK و هم در مسیر انتقال سیگنالینگ هورمونی گیاهی نقش دارند. این ژن‌ها احتمالاً تولید فیتوهورمون‌ها را از طریق شرکت در اتصالات سیگنالینگ خودبخودی جهت بهبود مقاومت به خشکی گیاه گندم تحریک می‌کنند. علاوه بر ژن‌های بالا ژن‌های مرتبط با متابولیسم گلوتامات و مسیر متابولیسم نیتروژن شناسایی گردید، که ژن‌های متابولیسم گلوتامات شامل گلوتامات دهیدروژن (GDH)، گلوتامات سنتتاز و ژن‌های مسیر متابولیسم نیتروژن شامل نیتروژن ردوکتاز، نیتروژن ترانسپورتر (NRT) و کربونیک انهدراز (CAH) بودند.

### پیش‌بینی ژن‌های هدف miRNA

نتایج با آنالیز برخط توالی miRNA‌های انتخابی شامل tae-miR9654a-3p و tae-miR1119 جهت پیش‌بینی ژن‌های هدف آنها، از طریق psRNATarget، صدها ژن هدف شناسایی گردید، که از بین ژن‌های هدف آنها بیشتر فاکتور‌های رونویسی و آنتی‌اکسیدان‌تی در ارتباط با تنش‌های غیر زیستی بودند، نتایج این آنالیز نشان داد که این miRNAها ژن‌های هدف خود را از طریق مکانیسم برش تنظیم می‌کنند.

### بحث

تنش خشکی شدیدترین محدودیت رشد و تکامل گیاه گندم در سطح مزرعه هست، همچنین از مهمترین فاکتورها غیرزیستی کاهش دهنده محصولات می‌باشد [۱۰]. در مطالعه حاضر، یک تجزیه و تحلیل بر روی داده‌های ترانسکریپتوم (RNAseq) نمونه‌های گندم نان تحت تنش خشکی و نرمال به منظور شناسایی ژن‌های درگیر در تنش خشکی صورت گرفت. نتایج حاصل از مطالعه این نمونه‌ها با توجه به دو معیار Log fold chang بین -۲ تا +۲ و pvalue کوچکتر از ۰/۰۵ منجر به شناسایی یکسری از ژن‌ها با بیان افتراقی گردید که یکسری افزایش بیان و یکسری دیگر کاهش بیان نشان دادند. بررسی بیشتر یکسری از این ژن‌ها حاکی از اهمیت کلیدی آنها در شرایط تنش خشکی بوده است. در این مطالعه، تعداد زیادی از ژن‌ها با بیان افتراقی مرتبط با تنش خشکی، طی هستی‌شناسی ژن، در سه مسیر غنی‌سازی گردید. این مسیرها شامل مسیر تعامل پاتوزن-گیاه، مسیر انتقال سیگنالینگ گیاهی و مسیر سیگنالینگ MAPK بودند. بعضی از این ژن‌ها غنی‌سازی شده در مسیر تعامل پاتوزن-گیاه، شامل ترئونین پروتئین کیناز FLS2، پروتئین کیناز ۱ مرتبط با رسپتور ۱ حساس به براسینوستروئید (BAK1)، WRKY33، سرین/ترئونین پروتئین کیناز (FRK1) بودند که علاوه بر این مسیر در مسیر سیگنالینگ MAPK هم نقش دارند. یکسری دیگر از ژن‌ها با بیان افتراقی متمایز در مسیر انتقال سیگنال هورمون گیاهی نقش داشتند، این ژن‌ها

شامل پروتئین رسپتور PR1 آبسزیک اسید، فاکتور رونویسی MYC2 و EIN2 مسیر سیگنالینگ MAPK بودند هم در بین ژن‌های با بیان متمایز غنی سازی گردید. فاکتورهای رونویسی TGA و JAZ همراه با ژن‌های که در بالا بیان شد، در تنظیم باز و بسته شدن استروما و رشد ریشه جهت سازگاری با تنش محیطی نقش دارند. این مسیر سازگاری اینطور عمل می‌کند که فاکتورهای FLS2 و BAK1 بیان FRK1 و PR1 را با تاثیر مستقیم بر روی WRKY23 مهار می‌کنند، بدین صورت یک مکانیسم پاسخ به تنش در گیاه ایجاد می‌کنند. PLY نه تنها سازگاری گیاه به تنش را از طریق مسیر پروتئین کیناز فعال شده توسط میتوزن (MPK1) فعال می‌کند، بلکه می‌تواند مستقیماً از طریق فاکتور ABF بر روی بسته شدن استروما اثر بگذارد. علاوه بر این تعامل بین JAZ و MYC2 می‌تواند بطور غیرمستقیم بر روی رشد ریشه گیاه و بهبود سازگاری آنها با تنش اثر بگذارد. EIN2 می‌تواند از طریق فعال کردن بیان اندوکیتیناز B پاسخ دفاعی در گیاه ایجاد کند [۳۶].

علاوه بر مسیرهای مهمی که در بالا به آنها اشاره گردید، به طور چشمگیری تعدادی از ژن‌ها با بیان متمایز در مسیر متابولیسم نیتروژن و متابولیسم اسیدهای آمینه غنی سازی گردید. همانطور که مشخص هست نیتروژن یک عنصر ضروری برای رشد و تکامل گیاه محسوب می‌شود، علاوه بر این در تنظیم فرآیندهای فیزیولوژیکی و بیوشیمیایی، همانند سازگاری، ترمیم و خشی کردن اثرات تنش خشکی نقش دارند. نیترات موجود در خاک از طریق ترانسپورتر NRT به ریشه انتقال می‌یابد و در سیتوپلاسم به نیتريت احیا می‌گردد [۲۶]. در کلروپلاست یا پلاست ها نیتريت از طریق آنزیم نیتريت ردوکتاز به نیتروژن آمونیوم تبدیل می‌گردد [۲۹]. نیتروژن آمونیوم وارد چرخه GS/GOGAT می‌شود و به نیتروژن آلی تبدیل می‌گردد، که در سنتز سایر اسیدهای آمینه و پروتئین‌ها نقش دارند [۲۷]. مطالعات انجام گرفته نشان می‌دهد که تنش‌های خارجی می‌تواند منجر به تجمع نیتريت در ریشه گیاهان گردد، که نتیجه آن بالا بردن مقاومت گیاه به تنش‌های خشکی می‌باشد [۲۵]. چن و همکاران در سال ۲۰۱۲ نشان دادند که تنش خشکی ایجاد شده با پلی اتیلن گلیکول از طریق تغییر در سطح بیان ترانسپورتر نیترات باعث تجمع نیتريت در ریشه گیاه و نهایتاً بهبود مقاومت به تنش خشکی در گیاه می‌گردد [۳]. نیتريك اكسيد القا شده توسط نیترات ایجاد مقاومت به انواع تنش‌ها را در گیاه ایجاد می‌کند [۱۳].

در مطالعه ما، آنزیم‌های در ارتباط با برداشت نیتروژن مانند NRT و NR در نمونه‌های تحت تنش خشکی افزایش یافتند. که نشان دهنده این هست که برداشت نیتروژن در نمونه گندم مقاوم به خشکی افزایش می‌یابد. علاوه بر این یکسری از مطالعات نشان دادند که در گیاهانی که تحت تنش اسموتیک قرار می‌گیرند، به طور عمده تجمع پرولین از طریق مسیر بیوسنتز پرولین، همراه با گلوتامات به عنوان پیش نیاز جهت حفظ رشد طبیعی افزایش می‌یابد [۲۸]. در گیاهان، گلوتامات از طریق چرخه GS/GOGAT به نیتروژن آلی تبدیل می‌گردد [۷]. آسپاراتات سنتاز (AS) از طریق تنظیم متابولیسم نیتروژن، مقاومت گیاه را به بیماری و تنش‌ها افزایش می‌دهد [۸]. در این مطالعه، GDH، GS و GOGAT به طور قابل ملاحظه‌ای در نمونه مقاوم به خشکی کاهش بیان نشان دادند. علاوه بر این، آسپاراتات سنتاز هم در نمونه مقاوم به خشکی کاهش بیان نشان داد.

شناخت اساسی شبکه‌های تنظیمی به واسطه miRNAs تحت شرایط تحمل به تنش خشکی می‌تواند به تشریح بیشتر مسیرهای مولکولی که گیاه را به تنش‌های خشکی سازگار می‌کند کمک کند. در سال‌های اخیر miRNAs گیاهی که در پاسخ به سازگاری گیاه به تنش خشکی نقش دارند به طور کاربردی شناسایی گردیده‌اند. رونوشت miR-319 طی تنش خشکی افزایش بیان نشان می‌دهد. افزایش بیان این ژن همراه با تغییرات مرفولوژیکی هست که تحمل گیاه به تنش خشکی و شوری را بالا می‌برد [۲۳]. در این مطالعه رونوشت Tra-miR-319 در نمونه تنش خشکی افزایش بیان نشان داد. ژن‌های

هدف این miRNA شامل ژن های AsPCF5، AsPCF6، AsPCF8 و فاکتور های رونویسی AsTCP14 بودند که تحت بیان miR-319 کاهش بیان نشان دادند [۳۵]. افزایش بیان Tra-miR-1119 در گیاه گندم ترانسژنیک TaMIR1119 تحت تنش خشکی منجر به بهبود وضعیت فنوتیپ و بیومس (وزن تر) گیاه می‌گردد، که همراه با افزایش عملکرد فتوسنتز، تجمع اسمولیت و بهبود هموستاز گونه های فعال اکسیژن (ROS= Reactive oxygen species) در گونه های ترانسژنیک گردید. در مطالعات صورت گرفته توسط محققین دیگر، نشان داده شد که در گیاه گندم تحت تنش خشکی، افزایش بیان این ژن منجر به افزایش بیان آنزیم های آنتی اکسیدانت نسبت به نمونه های کنترل گردید [۳۳]. در مطالعات ما هم در نمونه تحت تنش سطح بیان این miRNA افزایش یافت. ژن های هدف آن هم بیشتر فاکتور های رونویسی شامل (TaPCF، TaZP، TabHLH49) و ژن های آنتی اکسیدانتی (NtSOD2، NtCAT1، NtFeSOD) بودند. پژوهش های صورت گرفته بر اساس تجزیه و تحلیل عملکرد ژن نشان داد که، ژن های هدف miR9654a-3p مرتبط با انتقال وزیکولی پروتئین در سلول ها هست. در همین حال، تجزیه و تحلیل بیان ژن تحت تنش خشکی نشان داد که روند افزایش بیان miR9654a-3p برخلاف ژن هدف آن است که کاهش بیان ژن را نشان می‌دهد [۴]. در مطالعات ما هم نشان داده شد که بیان این miR در شرایط تنش خشکی در نمونه مقاوم به خشکی افزایش می‌یابد. ژن های هدف این miR شامل پروتئین مرتبط با وزیکول، همولکوس ریبونکلوپروتئین هسته ای و شیکیمات دهیدروژناز است.

آکوپورین ها (AQP) پروتئین های غشایی ضروری هستند که در فرایندهای همانند رسیدگی دانه، جوانه زنی، حرکت روزنه ها، فتوسنتز و فرایند گلدهی نقش دارند، همچنین نه تنها در انتقال طولانی آب از ریشه به برگ ها نقش دارند، بلکه در انتقال بین غشایی و درون غشایی آب به عنوان یک مکانیسم پاسخ به تنش هم نقش دارند [۲۱]. تجمع زیاد گونه های فعال اکسیژن (ROS) به علت تنش های محیطی می‌تواند به صورت جدی بر روی تکامل گیاه تداخل ایجاد کند، همچنین اثر منفی بر روی عملکرد محصول می‌گذارد [۱۹]. ژن پراکسیداز (PD) می‌تواند تعادل ROS در موجودات را با کاتالیز پراکسید هیدروژن (H<sub>2</sub>O<sub>2</sub>) تنظیم کند، و از این طریق نقش مهمی در رشد، تکامل و پاسخ به تنش غیر زیستی اعمال بکند [۳۳]. در مطالعات صورت گرفته توسط ما ta-miR-417 ژن پراکسیداز را هدف قرار می‌دهد و سطح بیان آن را کاهش می‌دهد. با مطالعه داده های بیانی ژن ها ما متوجه شدیم که این miR ژن TraesCS2D02G260300 که همان ژن پراکسیداز هست را هدف قرار می‌دهد. با مطالعاتی که توسط محققین صورت گرفته TraesCS2D02G260300 می‌تواند توسط تنش های مختلف غیرزیستی القا گردد. به عنوان مثال، سطح بیان این ژن در شرایط کمبود فسفر [۲۴]، گرما [۱۸] و تنش شوری [۳۷] کاهش یابد.

## نتیجه گیری

علی‌رغم مطالعات گسترده ای که بر روی تنش خشکی در گیاه گندم صورت گرفته است، هنوز مکانیسم های مولکولی درگیر در این مسیر به طور کامل شناسایی نگردیده است. با توجه به نقش کلیدی ژنتیک، اپی ژنتیک و ژن های مسئول در تنش خشکی، مطالعه در سطح ژنوم و ترانسکریپتوم می‌تواند مکانیسم های مولکولی اثر گذار در ایجاد مقاومت به تنش خشکی در گیاه گندم را شناسایی کند. ما در این مطالعه با استفاده از داده های ترانسکریپتوم بدست آمده از روش توالی یابی RNA توانستیم ژن ها و miRNA های با بیان متفاوت را در گیاه گندم مقاوم به تنش خشکی را شناسایی کنیم. این نتایج بیشتر مکانیسم های مولکولی اثرگذار بر روی تنش خشکی را شرح می‌دهد. ژن های هدف این miRNA های پاسخ دهنده به تنش

خشکی شامل فاکتورهای متعدد با نقش‌های مشخص شده به تنش خشکی بودند. این miRNAها و ژن‌های هدف آنها ممکن است از طریق برنامه‌های به‌نژادی یا مهندسی ژنتیک جهت بهبود تحمل به تنش خشکی در ارقام مختلف گندم و سایر گیاهان مرتبط مورد استفاده قرار گیرند.

**تضاد منافع:** نویسندگان اعلام می‌کنند که هیچ تضاد منافی ندارند.

## منابع

1. Akay A, Jordan D, Navarro IC, Wrzesinski T, Ponting CP, Miska EA, Haerty W (2019). Identification of functional long non-coding RNAs in *C. elegans*. *BMC biology*, 17:1-14.
2. Anjali N, Nadiya F, Thomas J, Sabu KK (2019). Identification and characterization of drought responsive microRNAs and their target genes in cardamom (*Elettaria cardamomum* Maton). *Plant Growth Regulation*, 87:201-216.
3. Chen CZ, Lv XF, Li JY, Yi HY, Gong JM (2012). Arabidopsis NRT1. 5 is another essential component in the regulation of nitrate reallocation and stress tolerance. *Plant Physiology*, 159:1582-1590.
4. Chen X, Yang Y, Ran L ping, Dong Z di, Zhang E jin, Yu X run, Xiong F (2017). Novel insights into miRNA regulation of storage protein biosynthesis during wheat caryopsis development under drought stress. *Frontiers in Plant Science*, 8:1707.
5. Contreras-Cubas C, Palomar M, Arteaga-Vázquez M, Reyes JL, Covarrubias AA (2012). Non-coding RNAs in the plant response to abiotic stress. *Planta*, 236:943-58.
6. Gangadhar BH, Venkidasamy B, Samynathan R, Saranya B, Chung IM, Thiruvengadam M (2021). Overview of miRNA biogenesis and applications in plants. *Biologia*, 76:2309-2327.
7. Guan M, Møller IS, Schjoerring JK (2015). Two cytosolic glutamine synthetase isoforms play specific roles for seed germination and seed yield structure in Arabidopsis. *Journal of experimental botany*, 66:203-212.
8. Horst RJ, Doehlemann G, Wahl R, Hofmann J, Schmiedl A, Kahmann R, Kamper J, Sonnewald U, Voll LM (2010). *Ustilago maydis* infection strongly alters organic nitrogen allocation in maize and stimulates productivity of systemic source leaves. *Plant physiology*, 152:293-308.
9. Hu J, Zeng T, Xia, Q. Qian Q, Yang C, Ding Y, Chen L, Wang W (2018). Unravelling miRNA regulation in yield of rice (*Oryza sativa*) based on differential network models. *Sci Rep*, 8:8498 (2018). <https://doi.org/10.1038/s41598-018-26438-w>
10. Islam W, Idrees A, Waheed A, Zeng F (2022). Plant responses to drought stress: microRNAs in action. *Environmental Research*, 215:114282.

11. Jha UC, Nayyar H, Jha R, Khurshid M, Zhou M, Mantri N, Siddique KHM, (2020). Long non-coding RNAs: Emerging players regulating plant abiotic stress response and adaptation. *BMC Plant Biology*, 20:1-20.
12. Jin H, Vacic V, Girke T, Lonardi S, Zhu JK (2008). Small RNAs and the regulation of cis-natural antisense transcripts in Arabidopsis. *BMC molecular biology*, 9:1-13.
13. Kaya C, Ashraf M, Alyemeni MN, Ahmad P (2020). The role of endogenous nitric oxide in salicylic acid-induced up-regulation of ascorbate-glutathione cycle involved in salinity tolerance of pepper (*Capsicum annuum* L.) plants. *Plant Physiology and Biochemistry*, 147:10-20.
14. Khan MA, Iqbal M, Akram M, Ahmad M, Hassan MW, Jamil M (2013). Recent advances in molecular tool development for drought tolerance breeding in cereal crops: a review. *Zemdirbyste-Agriculture*, 100:325-334.
15. Kosová K, Vítámvás P, Prášil IT, Renaut J (2011). Plant proteome changes under abiotic stress-contribution of proteomics studies to understanding plant stress response. *Journal of proteomics*, 74:1301-22.
16. Koyro HW, Ahmad P, Geissler N (2012). Abiotic Stress Responses in Plants: An Overview. In: Ahmad, P., Prasad, M. (eds) *Environmental Adaptations and Stress Tolerance of Plants in the Era of Climate Change*. Springer, New York, NY. <https://doi.org/10.1007/978-1-4614-0815-4-1>.
17. Kruszka K, Pieczynski M, Windels D, Bielewicz D, Jarmolowski A, Szweykowska-Kulinska Z, Vazquez F (2012). Role of microRNAs and other sRNAs of plants in their changing environments. *Journal of plant physiology*, 169:1664-1672.
18. Kumar RR, Goswami S, Sharma SK, Kala YK, Rai GK, Mishra DC, Grover M, Singh GP, Pathak H, Rai A, Chinusamy V, Rai RD (2015). Harnessing next generation sequencing in climate change: RNA-Seq analysis of heat stress-responsive genes in wheat (*Triticum aestivum* L.). *Omics: a journal of integrative biology*, 19:632-47.
19. Li Q, Zheng Q, Shen W, Cram D, Fowler DB, Wei Y, Zou J (2015). Understanding the biochemical basis of temperature-induced lipid pathway adjustments in plants. *The plant cell*, 27:86-103.
20. Liu X, Zhang X, Sun B, Hao L, Liu C, Zhang D, Tang H, Li C, Li Y, Shi Y, Xie X, Song Y, Wang T, Li Y (2019). Genome-wide identification and comparative analysis of drought-related microRNAs in two maize inbred lines with contrasting drought tolerance by deep sequencing. *PLoS One*, 14: e0219176.
21. Madrid-Espinoza J, Brunel-Saldias N, Guerra FP, Gutiérrez A, Del Pozo A (2018). Genome-wide identification and transcriptional regulation of aquaporin genes in bread wheat (*Triticum aestivum* L.) under water stress. *Genes*, 9:497.
22. Mayer KF, Rogers J, Doležel J, Pozniak C, Eversole K, Feuillet C, Gill B, Friebe B, Lukaszewski AJ, Sourdille P (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, 345: p.1251788.

23. Nag A, King S, Jack T (2009). miR319a targeting of TCP4 is critical for petal growth and development in Arabidopsis. *Proceedings of the National Academy of Sciences*, 106:22534-22539.
24. Oono Y, Kobayashi F, Kawahara Y, Yazawa T, Handa H, Itoh T, Matsumoto T (2013). Characterisation of the wheat (*Triticum aestivum* L.) transcriptome by de novo assembly for the discovery of phosphate starvation-responsive genes: gene expression in Pi-stressed wheat. *BMC genomics*, 14:1-14.
25. Pan QN, Geng CC, Li DD, Xu SW, Mao DD, Umbreen S, Loake GJ, Cui BM (2019). Nitrate reductase-mediated nitric oxide regulates the leaf shape in Arabidopsis by mediating the homeostasis of reactive oxygen species. *International Journal of Molecular Sciences*, 20:2235.
26. Parker JL, Newstead S (2014). Molecular basis of nitrate uptake by the plant nitrate transporter NRT1. 1. *Nature*, 507:68-72.
27. Redinbaugh MG, Campbell WH (1998). Nitrate regulation of the oxidative pentose phosphate pathway in maize (*Zea mays* L.) root plastids: induction of 6-phosphogluconate dehydrogenase activity, protein and transcript levels. *Plant Science*, 134:129-140.
28. Rhodes D, Handa S, Bressan RA (1986). Metabolic changes associated with adaptation of plant cells to water stress. *Plant physiology*, 82:890-903.
29. Sétif P, Hirasawa M, Cassan N, Lagoutte B, Tripathy JN, Knaff DB (2009). New insights into the catalytic cycle of plant nitrite reductase. *Electron transfer kinetics and charge storage. Biochemistry*, 48:2828-2838.
30. Tanno K ichi, Willcox G (2006). How fast was wild wheat domesticated? *Science*, 311:1886-1886.
31. Wang J, Meng X, Dobrovolskaya OB, Orlov YL, Chen M (2017). Non-coding RNAs and their roles in stress response in plants. *Genomics, Proteomics & Bioinformatics*, 15:301-312.
32. Yao Y, Ni Z, Peng H, Sun F, Xin M, Sunkar R, Zhu JK, Sun Q (2010). Non-coding small RNAs responsive to abiotic stress in wheat (*Triticum aestivum* L.). *Functional & integrative genomics*, 10:187-90.
33. You J, Chan Z (2015). ROS regulation during abiotic stress responses in crop plants. *Frontiers in Plant Science*, 6:1092.
34. Yu Y, Zhang Y, Chen X, Chen Y (2019). Plant noncoding RNAs: hidden players in development and stress responses. *Annual review of cell and developmental biology*, 35:407.
35. Zeeshan M, Qiu CW, Naz S, Cao F, Wu F (2021). Genome-wide discovery of miRNAs with differential expression patterns in responses to salinity in the two contrasting wheat cultivars. *International Journal of Molecular Sciences*, 22:12556.

36. Zhang J, Huang D, Zhao X, Zhang M, Wang Q, Hou X, Di D, Su B, Wang S, Sun P (2022). Drought-responsive WRKY transcription factor genes IgWRKY50 and IgWRKY32 from *Iris germanica* enhance drought resistance in transgenic *Arabidopsis*. *Frontiers in Plant Science*, 13:983600.
37. Zhang Y, Liu Z, Khan AA, Lin Q, Han Y, Mu P, Liu Y, Zhang H, Li L, Meng X, Ni Z, Xin M (2016). Expression partitioning of homeologs and tandem duplications contribute to salt tolerance in wheat (*Triticum aestivum* L.). *Scientific reports*, 6:1-10.
38. Zheng B, Wang Z, Li S, Yu B, Liu JY, Chen X (2009). Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. *Genes & development*, 23:2850-2860.

## Total RNA transcript analysis to identify regulatory non-coding RNAs and their targets in response to drought stress in wheat (*Triticum aestivum* L.)

Zare M.A.<sup>1</sup>, Zarinpanjeh N.<sup>\*2</sup>, Khalaj H.<sup>1</sup>, Rafiepour M.<sup>3</sup>, Mir Alizadeh S.R.<sup>4</sup>

<sup>1</sup>Department of Agriculture, Payame Noor University, Tehran, Iran, <sup>2</sup>Medicinal Plants Research Center, Institute of Medicinal Plants, ACECR, Karaj, Iran, <sup>1</sup>Department of Agriculture, Payame Noor University, Tehran, Iran, <sup>3</sup>Danesh Alborz University, Faculty of Science, Qazvin, Iran, <sup>4</sup>Department of Agriculture and Environment, Payame Noor University, Tehran, Iran

\*zarinpanjeh@imp.ac.ir

### Abstract

Drought is a significant abiotic stress factor that impacts the productivity of crops, including wheat. However, drought tolerance is a complex trait involving numerous genes and cellular pathways. Understanding the molecular regulatory network of wheat in response to drought stress is crucial for molecular breeding. Non-coding RNAs affect plant growth and resistance to abiotic stresses by regulating gene expression. To help understand the complex regulatory mechanisms governing drought tolerance in wheat, we performed a total RNA transcriptome sequence analysis in a drought-tolerant wheat cultivar. A total of 20462 mRNAs and 209 miRNAs with different expressions were identified. Among 20462 genes with different expressions, 1025 were identified as genes that withstand drought stress. The target genes identified by psRNATarget were all enriched in gene ontology categories. Significant differences exist in the number of target genes regulated by differently expressed miRNAs, especially in signal transduction, stress response, and antioxidant pathways. Our results elucidate wheat's molecular mechanisms of drought tolerance and identify drought-responsive miRNAs and their target genes. These insights can be utilized to enhance wheat and related crops, providing valuable genomic resources for molecular breeding.

**Keywords:** Drought stress, RNA sequencing, Wheat, Non-coding RNAs