

**ORIGINAL ARTICLE**

# Analyzing the expression data (*RNA-Seq*) in *Catharanthus roseus* plant in order to identify key genes of the active substances (vinblastine and vincristine) in different organs (root, leaf and flower) and to investigate their anticancer effects using docking analysis and molecular dynamics.

Masoud Tohidfar<sup>1\*,2</sup>, Yousef Saeedi Honar<sup>1</sup>, Naser Farrokhi<sup>1,2</sup>

<sup>1</sup>Department of Cell & Molecular Biology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran.

<sup>2</sup>Center for International Scientific Studies and Collaboration (CISSC), Ministry of Science, Research and Technology, Tehran, Iran.

**Correspondence**

Masoud Tohidfar  
Email: [m\\_tohidfar@sbu.ac.ir](mailto:m_tohidfar@sbu.ac.ir)

**How to cite**

Tohidfar, M., Saeedi Honar, Y., & Farrokhi, N. (2023). Analyzing the expression data (*RNA-Seq*) in *Catharanthus roseus* plant in order to identify key genes of the active substances (vinblastine and vincristine) in different organs (root, leaf and flower) and to investigate their anticancer effects using docking analysis and molecular dynamics. *Crop Biotechnology*, 13(44), 1-11.

**ABSTRACT**

*Catharanthus roseus* is one of the most important medicinal plants that contains two antitumor substances, vinblastine and vincristine. It is important to identify the involved genes and their expression pattern and anti-tumor effect in different tissues of this plant. By using the expression data of RNA sequencing of different tissues, differential expression genes and their antitumor effects were investigated as in silico. The results showed that the total number of differentially expressed genes in the organs varied between 120 and 1238. The highest number of DEGs compared to the root was related to the leaf and the lowest number was related to the flower. Subsequently, 13 common genes between three different organs and 22 common genes were observed between leaves versus flowers and leaves versus roots. Among them, 6 common genes were observed in all three tissues, and the annotation analysis showed that these genes are involved in the biosynthetic pathway of two important compounds, vinblastine and vincristine. The highest expression of these genes was related to leaves and the lowest was related to roots. Protein network analysis determined that a number of genes that showed the most interaction with other genes were related to the genes of the biocentric pathway of antitumor compounds. Docking and molecular dynamics analysis showed that vinblastine and vincristine, while having good interaction as inhibitors with phosphoglycoprotein (drug resistance protein in tumor cells), also have good stability in interaction with phosphoglycoprotein. Generally DAT, STR, TDC, G10H, D4H, T16H2, Tryptophandecar-boxylase and Strictosidine synthase genes that were in the biosynthesis pathway of vinblastine and vincristine had an effective role in different organs. The obtained results give new insights about the mechanism of treatment with natural products, which can be used to improve the patients.

**KEYWORDS**

Bioinformatics, *Catharanthus roseus*, gene expression, vinblastine and vincristine

© 2023, by the author(s). Published by Payame Noor University, Tehran, Iran.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://cropbiotech.journals.pnu.ac.ir/>

نشر به علمی

## زیست فناوری گیاهان زراعی

«مقاله پژوهشی»

# آنالیز داده‌های بیانی (RNA-Seq) در گیاه *Catharanthus roseus* به منظور شناسایی ژن‌های کلیدی مواد مؤثره (وین بلاستین و وین کریستین) در اندام‌های مختلف (ریشه، برگ و گل) و بررسی اثرات ضد سرطانی آنها با استفاده از آنالیز داکینگ و دینامیک مولکولی

مسعود توحیدفر<sup>۱\*</sup>، یوسف سعیدی هنر<sup>۱</sup>، ناصر فرخی<sup>۲</sup>

<sup>۱</sup>گروه سلولی و مولکولی دانشکده علوم و فناوری زیستی، دانشگاه شهید بهشتی، تهران، ایران.  
<sup>۲</sup>آمرکز مطالعات و همکاری‌های علمی بین‌المللی وزارت علوم تحقیقات و فناوری، تهران، ایران.

### چکیده

گیاه پروانش (*Catharanthus roseus*) یکی از مهمترین گیاهان دارویی است که حاوی دو ماده‌ی ضدتوموری وین‌بلاستین و وین‌کریستین است و شناسایی ژن‌های درگیر و الگوی بیانی و اثرات ضدتوموری آنها در بافت‌های مختلف این گیاه حایز اهمیت است. در همین راستا، با استفاده از داده‌های بیانی (RNA sequencing (RNA-seq) و بافت‌های مختلف به بررسی ژن‌های با بیان افتراقی (Differential Expression Gene) و اثرات ضدتوموری آنها بصورت این سیلکو پرداخته شد. نتایج نشان داد که تعداد کل ژن‌های با بیان افتراقی در اندام‌ها بین ۱۲۰ تا ۱۲۳۸ متغیر بود. بیشترین تعداد DEG نسبت به ریشه مربوط به برگ و کمترین آن متعلق به گل بود. در ادامه، ۱۳ ژن مشترک بین سه اندام مختلف و ۲۲ ژن مشترک بین برگ در برابر گل و برگ در برابر ریشه مشاهده شد که از بین آنها، ۶ ژن مشترک در هر سه بافت بودند. آنالیز مستندسازی نشان داد که این ژن‌ها در مسیر بیوسنتزی دو ترکیب مهم وین‌بلاستین و وین‌کریستین نقش دارند. بیشترین بیان این ژن‌ها مربوط به برگ و کمترین آن مربوط به ریشه بود. آنالیز شبکه‌ی پروتئین مشخص کرد که تعدادی از ژن‌ها که بیشترین برهم‌کنش را با سایر ژن‌ها نشان دادند مربوط به ژن‌های مسیر بیوسنتزی ترکیبات ضد توموری بودند. آنالیز داکینگ و دینامیک مولکولی نشان داد که وین‌بلاستین و وین‌کریستین ضمن اینکه برهم‌کنش خوبی به‌عنوان بازدارنده با فسفولیکوپروتئین (پروتئین مقاومت دارویی در سلول‌های توموری) دارند، از پایداری خوبی هم در برهم‌کنش با فسفولیکوپروتئین برخوردار هستند. بطور کلی، ژن‌های *G10H*، *TDC*، *STR*، *DAT*، *T16H2*، *D4H* و *Strictosidine synthase* و *Tryptophandecar-boxylase* که در مسیر بیوسنتزی وین‌بلاستین و وین‌کریستین بودند، نقش موثری در اندام‌های مختلف داشتند. نتایج بدست آمده بینش‌های جدیدی در مورد مکانیسم درمان با محصولات طبیعی می‌دهد که می‌توان جهت بهبود افراد مبتلا مورد استفاده قرار داد.

### واژه‌های کلیدی

بیوانفورماتیک، گیاه پروانش، بیان ژن، وین بلاستین و وین کریستین.

نویسنده مسئول:

مسعود توحیدفر

رایانامه: [m\\_tohidfar@sbu.ac.ir](mailto:m_tohidfar@sbu.ac.ir)

استناد به این مقاله:

توحیدفر، مسعود، سعیدی هنر، یوسف و فرخی، ناصر (۱۴۰۲). آنالیز داده‌های بیانی (RNA-Seq) در گیاه پروانش (*Catharanthus roseus*) به منظور شناسایی ژن‌های کلیدی مواد مؤثره (وین بلاستین و وین کریستین) در اندام‌های مختلف (ریشه، برگ و گل) و بررسی اثرات ضد سرطانی آنها با استفاده از آنالیز داکینگ و دینامیک مولکولی. فصلنامه علمی زیست فناوری گیاهان زراعی، ۱۳(۴۴)، ۱-۱۱.

## مقدمه

نوین همچون روش‌های زیست‌فناوری می‌تواند مورد توجه قرار گیرد (سهندی، ۱۳۹۵).

فسفولیکوپروتئین در بسیاری از فرایندهای سلولی از قبیل التهاب، تمایز سلول‌های ایمنی، سمیت‌زدایی و ترشح هورمون‌ها نقش دارد. کاهش کارایی درمان و عود بیماری متعاقب آن به علت مقاومت دارویی مهم‌ترین پیامد عملکرد این پروتئین بوده است و به عنوان یکی از پرچالش‌ترین موانع در درمان سرطان مورد توجه است. گلیکوپروتئین‌ها، پروتئین حاوی یک یا تعداد بیشتری واحد اولیگوساکاریدی غیر تکراری (گلیکان) هستند. همچنین، تومورهای تخمدان و سایر تومورها از گلیکوپروتئین‌های خاص و توالی گلیکان مرتبط برای تعدیل پاسخ‌های ایمنی استفاده می‌کنند. گلیکوپروتئین‌ها سلول‌های توموری را که این مولکول‌ها را بیان یا ترشح می‌کنند، قادر می‌سازند تا از حمله سلول‌های ایمنی فرار کنند و سیستم ایمنی را برای تقویت رشد تومور تحریک کنند (فاطمی و همکاران ۱۳۸۶).

یکی از روش‌های افزایش آکالوئیدها (وین بلاستین و وین کریستین)، شناسایی ژن‌های کلیدی و دخیل در سنتز این آکالوئیدها (وین بلاستین و وین کریستین) و افزایش بیان آنها است. هدف از انجام این تحقیق بررسی الگوی بیانی ژن‌های مسیر بیوسنتزی ترکیبات ضدسرطانی (وین بلاستین و وین کریستین) در گیاه پروانش و برهمکنش آنها با استفاده از داکینگ مولکولی با پروتئین‌های مقاومت دارویی نظیر فسفولیکوپروتئین (P-Glycoprotein) در سلول سرطانی است

## روش شناسی پژوهش

در این مطالعه از داده‌های RNA-Seq مربوط به ۳ نمونه از گیاه پروانش استفاده شد. نمونه‌ها شامل اندام‌های مختلف ریشه، گل و برگ بود. داده‌های مربوطه از بانک اطلاعاتی مرکز ملی اطلاعات بیوتکنولوژی (NCBI)، بخش GEO دریافت شد. کیفیت خوانش‌ها با استفاده از نرم افزار FASTQC کنترل شد. خوانش‌های دارای آلودگی‌های آدپتوری و غیره و خوانش‌های با کیفیت کمتر از ۳۰ توسط Trimmomatic ویرایش شدند.

سپس هم‌ردیفی خوانش‌ها با دونه اسمبلی و فایل‌هایی حاشیه‌نویسی شده GTF ([https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org\\_Lusitattissimum](https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Lusitattissimum)) با استفاده از HISAT و

انسان با شناخت گیاهان دارویی از گذشته به شفا بخشی و اثرات درمانی این گیاهان پی برد و به منظور پیشگیری و درمان بیماری‌ها از آنها استفاده کرد. در ایران نیز هزاران گونه گیاهی کشت و کار می‌شود که دارای اثرات دارویی هستند و در درمان بیماری‌ها مورد استفاده قرار می‌گیرند. اما با توجه به محدود بودن زیستگاه طبیعی گیاهان و پایین بودن موثره، استفاده از گیاهان دارویی با محدودیت‌هایی همراه است. بر این اساس محققین جهت بهره‌وری بیشتر از گیاهان دارویی و افزایش تولید متابولیت‌های ثانویه از آنها به استفاده از فناوری زیستی روی آورده‌اند (سهندی ۱ و همکاران ۱۳۹۵).

پروانش (*Catharanthus roseus* L.) متعلق به خانواده خرزهره بوده و ۱۶ کروموزوم دارد ( $2n=16$ ). این گیاه چند ساله است که در مناطق سرد به صورت یکساله کشت می‌شود. این گیاه یکی از مهم‌ترین گیاهان دارویی و طبی است که دارای بیش از ۴۰۰ نوع آکالوئید از نوع ترپنوئید ایندول آکالوئید است (Verpoorte, 2004). دو آکالوئید وین بلاستین و وین کریستین به خاطر اتصال به میکروتوبول‌ها و توقف تقسیم سلولی در طی متافاز میتوز، خاصیت ضد توموری داشته (سلطانی ۱۴۰۰) و مورد توجه زیادی قرار گرفته است (Heijden der Van, 2004).

## پیشینه پژوهش

ارزش اقتصادی دوماه وین بلاستین و وین کریستین بالا است، بطوریکه قیمت هر کیلوگرم وین بلاستین یک میلیون دلار و تولید سالانه آن ۱۲ کیلوگرم و قیمت هر کیلوگرم وین کریستین ۵/۳ میلیون دلار و تولید سالانه آن یک کیلوگرم است (Loyola-Vargas et al, 2007) در حال حاضر این دو ماده از طریق استخراج از گیاه بدست می‌آیند. استخراج آن بعلا غلظت کم این ترکیبات مقرون به صرفه نیست. میزان سنتز این ترکیبات به قدری کم است که برای تولید یک گرم وین بلاستین نزدیک ۵۰۰ کیلوگرم برگ خشک و برای تولید یک گرم وین کریستین حدود ۲ تن برگ خشک نیاز است (همتی ۱۴۰۰). از طرفی روش‌های سنتز هم بخاطر هزینه مناسب نیستند. بنابراین افزایش مقدار آنها از طریق فناوری‌های

Anders *et al.*, 2015; Pertea ) صورت گرفت HTSeq-count  
(*et al.*, 2016).

### آنالیز DEGs

برای این منظور ماتریس بیان ژن هر نمونه در مقایسه با نمونه بعدی مورد بررسی قرار گرفت. شناسایی ژن های با بیان متفاوت با استفاده از بسته DESeq2 نرم افزار R بر اساس مدلی مبتنی بر توزیع دوجمله ای منفی انجام شد. در این آنالیز ژن های با مقادیر  $P < 0.05$  که با روش FDR تصحیح شده بودند جداسازی شد.

### آنالیز هستی شناسی<sup>۱</sup>

آنالیز هستی شناسی ژن های با بیان افتراقی با به کارگیری ابزار آنالیز agriGO<sup>۲</sup>، با استفاده از روش<sup>۳</sup> SEA انجام شد و دسته بندی DEGs بر اساس سه پارامتر اجزاء سلولی، عملکرد مولکولی و فرایندهای زیستی صورت گرفت (Tian *et al.*, 2017).

### آنالیز غنی سازی مسیر KEGG

تجزیه غنی سازی مسیر برای شناسایی مسیرهایی که ژن های با بیان افتراقی در آنها درگیرند با استفاده از پایگاه داده KEGG<sup>۴</sup> انجام شد (Kanehisa & Goto, 2000).

### شبکه پروتئینی و هاب ژن ها

برای بررسی برهم کنش های پروتئین - پروتئین، از پایگاه String استفاده شد. برای شناسایی ژن های هاب، نرم افزار Cytoscape گزینه Cytohubba براساس الگوریتم محاسباتی MNC بکاررفت.

### آماده سازی مولکول کوچک زیستی (لیگاند)

برای استفاده از ساختار وین بلاستین لازم است ساختار سه بعدی آن طراحی شود زیرا تنها در پایگاه داده Pubchem ساختار دو بعدی آن موجود است. نرم افزار تخصصی این حوزه مجموعه نرم افزار «کم بایو آفیس» است که تمام چالش های موجود در طراحی ساختار لیگاند را برطرف می کند. در نهایت ساختار لازم بدست آمد.

### آماده سازی گیرنده (رسپتور)

برای بدست آوردن ساختار رسپتورها (فسفوگلیکوپروتئین و لیوین) از پایگاه داده پروتئین PDB استفاده شد.

### بررسی ساختار پروتئین با نرم افزار پایمول

این نرم افزار بر اساس زبان برنامه نویسی پایتون ارائه شده و می توان با استفاده از دستور نوشتن همانند نرم افزار پایتون فرمان انجام فعالیت را صادر کرد. و برای ساخت پپتیدهای بهینه برای آنالیزهای ضد میکروبی استفاده شده است.

### نرم افزار کایمرا

نرم افزار کایمرا یک بسته نرم افزاری شیمی و بیوانفورماتیک برای تجسم و شبیه سازی تعاملی، نمایش، تجزیه و تحلیل و آنالیز ساختار مولکولی و داده های مرتبط با آن ها استفاده شده است.

### نرم افزار مادلر

از نرم افزار مادلر به منظور همولوژی مدلینگ اتوماتیک پروتئین ها استفاده شد. نرم افزار مادلر یکی از نرم افزارهای جامع به منظور همولوژی و یا مقایسه ی مدلینگ های ساختار سه بعدی پروتئین است.

### داکینگ مولکولی

یکی از سرورهای آنلاین کاربردی در زمینه داکینگ مولکولی سرور سوئیس داک است (<http://www.swissdock.ch>). این سرور با استفاده از گیرنده و مولکول کوچک زیستی که در اختیار آن قرار داده شده به بررسی برهم کنش آن ها می پردازد.

### شبیه سازی دینامیک مولکولی

برای دینامیک مولکولی از گرومکس، با نام مختصر (Groningen Machine for Chemical Simulations) استفاده شد. گرومکس یک بسته نرم افزاری تطبیق پذیر و متنوع برای انجام دینامیک مولکولی است که از طریق حل معادلات حرکت نیوتنی امکان شبیه سازی سامانه هایی متشکل از صدها تا میلیون ها ذره را فراهم می کند.

3. Singular Enrichment Analysis  
4. <https://www.genome.jp/kegg/>

1. GO Enrichment  
2. <http://bioinfo.cau.edu.cn/agriGO/>

## آنالیز مسیرهای سیگنالینگ

پس از مشخص شدن ارتباط بین ژن‌ها با استفاده از پایگاه داده Enrichr مسیرهای سیگنالینگ مشخص شدند (Subramanian, 2007).

## یافته های پژوهش

برای آنالیزیابی، از داده‌های RNA-Seq مربوط به پایگاه داده Gene Expression Omnibus (GEO) DataSets با شماره شناسه GSE57326 و پلت فرم GPL22656 استفاده شد. این داده‌ها مربوط به بیان ژن در سه نمونه ریشه، برگ و گل بود. در مجموع ۲۱۰۱۸۱۳۱۴ خوانش حاصل شد که بین ۷۰ تا ۱۵ میلیون در هر نمونه متغیر بود. از این میان ۷۱۰۶۰۴۳۸، ۸۰۸۷۹۶۷۴ و ۶۱۸۹۶۴۳۲ میلیون خوانش بترتیب مربوط به نمونه ریشه، برگ و گل بود (جدول ۱).

### جدول ۱. اطلاعات کلی مربوط به داده‌های RNA-Seq در گیاه

پروانش مربوط به شناسه GSE57326

نمونه	تعداد خوانش‌ها	نرخ نقشه‌برداری (%)
ریشه	۷۱۰۶۰۴۳۸	۸۹/۲۵
برگ	۸۰۸۷۹۶۷۴	۹۰/۷۶
گل	۶۱۸۹۶۴۳۲	۹۴/۸۹

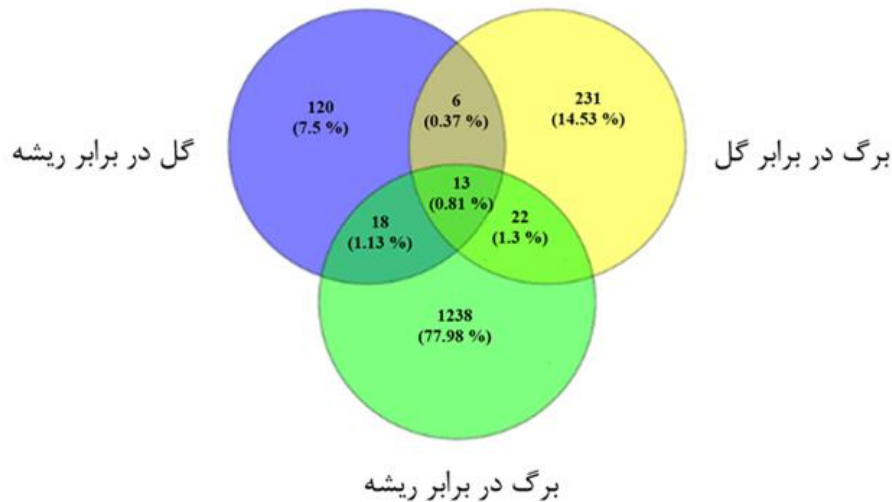
در مجموع به طور متوسط ۸ درصد داده‌ها در مرحله پیرایش حذف شدند که نشان‌دهنده کیفیت بسیار مناسب داده‌ها بود. بیشترین تعداد کل خوانش‌ها مربوط به نمونه برگ و کمترین آن برای نمونه گل گزارش شده است. در ادامه، خوانش‌های با میانگین ۹۱/۵۵ درصد با ژنوم هم‌ردیف شدند (جدول ۱).

## شناسایی ژن‌های با بیان افتراقی (DEGs)

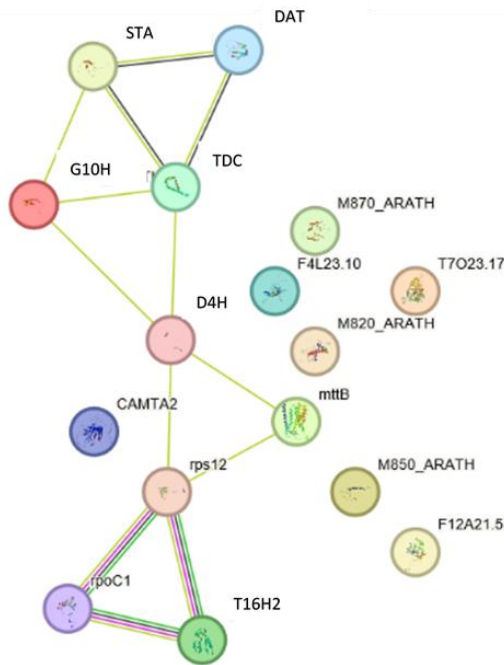
جهت شناسایی ژن‌هایی با بیان افتراقی معنی‌دار، آنالیز تفاوت بیان ژن بین بافت‌ها صورت گرفت و پارامترهای بیان افتراقی مانند مقادیر  $p$ -values، FDR و مقادیر  $\log_2$  fold change (FC) از ماتریس داده‌های شمارش شده ترانسکریپتوم با استفاده از بسته نرم‌افزاری DESeq2 و edgeR محاسبه شدند. نتایج نشان داد که تعداد کل ژن‌ها با بیان افتراقی در اندام‌ها نسبت به هم بین ۱۲۰ تا ۱۲۳۸ متغیر بود. بیشترین تعداد DEG ها برای اندام برگ نسبت به ریشه و کمترین آن متعلق به نمونه گل در برابر ریشه بود (شکل ۱).

برای بررسی ژن‌های مشترک، نمودار Venn بین اندام‌های مختلف رسم شد (شکل ۱). نتایج نشان داد که ۱۳ ژن مشترک بین سه اندام مختلف و ۲۲ ژن مشترک بین برگ در برابر گل و برگ در برابر ریشه مشاهده شد. همچنین ۶ ژن مشترک بین گل در برابر ریشه و برگ در برابر گل مشاهده شد، و سایر ژن‌ها به طور اختصاصی نسبت به بافت پاسخ دادند.

از جمله ژن‌های مشترک در هر سه بافت شامل DAT، Tryptophandecarboxylase، T16H2، D4H، G10H، TDC، STR و Strictosidine synthase بودند که آنالیز مستند سازی نشان داد که در مسیر بیوسنتزی دو ترکیب مهم وین بلاستین و وین کریستین نقش دارند. ولی تغییرات بیان آنها فرق داشت بطوریکه که بیشترین بیان آنها مربوط به برگ و کمترین آن مربوط به ریشه بود. درحالی‌که ژن‌های مشترک مربوط به ریشه به طور اختصاصی در فرآیندهای تردوکس، سلول، ترانسپورت، متابولیسم هورمونی و کینازهای گیرنده سیگنالینگ تغییر بیان نشان دادند.



شکل ۱. ژن های با بیان افتراقی و نمودار ون دیاگرام در اندام های مختلف در گیاه پروانش



شکل ۲. شبکه پروتئینی بین ژن های مشترک بافت های برگ، ریشه و گل در گیاه پروانش

بررسی مسیرهای KEGG برای ژنهای مشترک پروانش نشان داد که در بین این مسیرها ژن های مرتبط با بیوسنتز متابولیت های ثانویه و بیوسنتز وین بلاستین و وین کریستین که مختص این نوع گونه گیاه است در پاسخ به نوع اندام تغییرات بیانی متفاوتی داشتند.

### هستی شناسی (GO) ژن های مشترک

آنالیز هستی شناسی بر روی ژن های مشترک انجام شد. این ژن ها در سه دسته اصلی GO شامل فرآیندهای بیولوژیکی، اجزای سلولی و عملکردهای مولکولی تقسیم شدند. در اینجا دسته بندی اصطلاحات GO از فرآیندهای بیولوژیکی صورت گرفت. در این حالت ، ۹ ژن کدکننده ترانسکریپت های متعلق به دو ترکیب مهم وین بلاستین و وین کریستین وجود داشت که به طور معنی داری در برگ افزایش بیان داشتند و بالعکس در ریشه کاهش بیان نشان دادند. علاوه بر این در برگ نسبت به ریشه TF های با بالاترین تغییرات مشاهده شد که در مسیر عملکردی وین بلاستین و وین کریستین نقش داشتند. از ۵ گروه عملکردی مرتبط با ۶ ژن اختصاصی که به طور مشترک در ریشه بیان شدند، به طور معنی دار در فرآیندهای پاسخ به تنش ها، پاسخگو بودند.

به منظور فهم بهتر ارتباط بین ژن های مشترک شبکه برهمکنش پروتئینی بر اساس الگوریتم MNC رسم شد (شکل ۲). در بین بافت های مختلف ژن های DAT، STR، TDC، G10H، D4H، T16H2 بیشترین برهمکنش را با سایر ژن ها نشان دادند. ژن STR علاوه بر این که با سایر ژن های مشترک پاسخ به بافتها با درجه بالایی برهمکنش می دهد، بیشترین کاهش بیان را در ریشه داشت و به شدت تحت تاثیر بافت بود (شکل ۲).

**آماده سازی لیگاند**

وین‌بلاستین پروانش طبق جدول ۲ با احتمال ۰.۹۶۳ توانایی مهار فسفوگلیکوپروتئین را دارد. که برای ادامه کار، فسفوگلیکوپروتئین به عنوان پروتئین هدف انتخاب شد.

**آماده سازی رسپتور فسفوگلیکوپروتئین**

ابتدا ساختار فسفوگلیکوپروتئین با شماره دسترسی ( pdb id 3G5U) از پایگاه داده پروتئین دریافت شد. پس از آنالیز ساختار، مشخص شد که این پروتئین حاوی آمینواسیدهای هرز در جایگاه فعال خود است برای رفع این مشکل ابتدا ساختار توسط نرم افزار مدلر و سوئیس مدل به رفع این آمینواسیدهای هرز پرداخته شد. سپس با استفاده از سرور ucla آنالیز نمودار رامچاندرا پلات صورت گرفت و نرم افزار و سرور با یک دیگر مقایسه شدند (شکل ۳). نتایج نشان داد که سرور سوئیس مدل بر اساس هوش مصنوعی و الگوریتم آلفا فولد تولنایی بهتری برای رفع مشکل آمینواسیدهای هرز را دارند.

**داکینگ مولکولی**

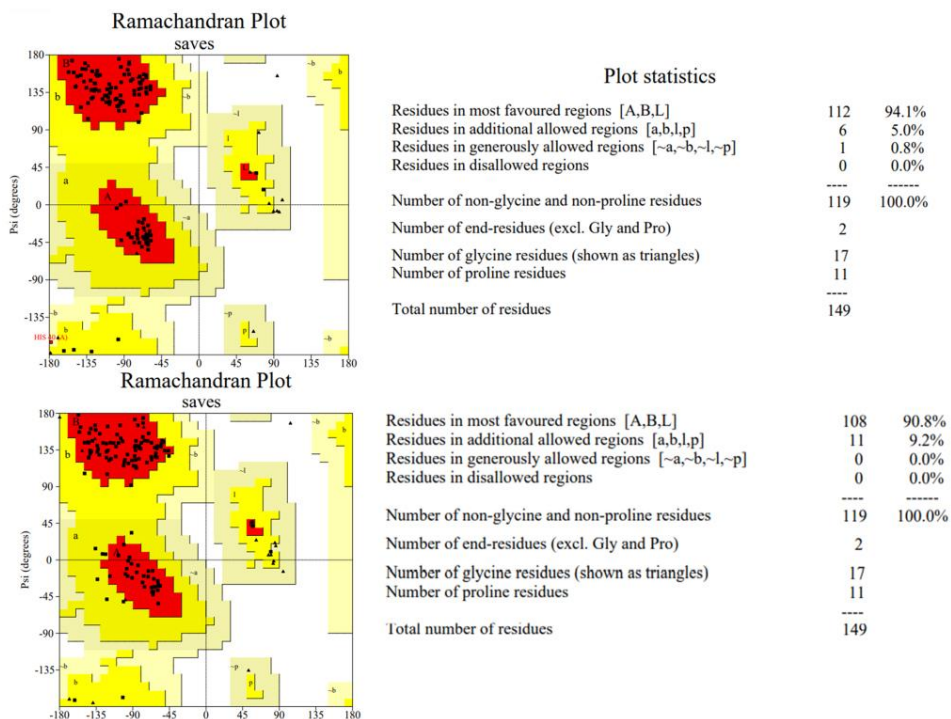
پس از آماده سازی ساختار رسپتورها و لیگاندها، داکینگ مولکولی بر روی ساختار فسفوگلیکوپروتئین صورت گرفت. با استفاده از نرم افزار اتوداک وینا ابتدا فراخوانی رسپتورها انجام شد. کاهش سطح انرژی ساختارها و فرآیند داکینگ انعطاف پذیر با نرم افزار کایمرا صورت گرفت. نتایج نشان داد که از بین متابولیت‌های غربال شده، متابولیت‌های ثانویه گیاه پروانش به خوبی تولنایی برهم‌کنش را با رسپتور فسفوگلیکوپروتئین دارند. از بین متابولیت‌های پروانش، وین‌بلاستین و وین‌کریستین بهترین برهم‌کنش را نشان دادند (جدول ۳ و ۴).

با استفاده از سرور way2drug خواص متابولیت‌های ثانویه از جمله وین‌بلاستین، وین‌کریستین در گیاه پروانش طبق جدول ۲ مورد تجزیه و تحلیل قرار گرفت، نتایج نشان داد که خواص ضد سرطانی، افزایش بیان ژن TP53، خواص ضد میکروبی و مهار فسفوگلیکوپروتئین در متابولیت‌های وین‌بلاستین، وین‌کریستین، ویندولین، اولاندین و نیرین به میزان بالایی وجود دارد.

**جدول ۲.** خواص ضد سرطانی و میکروبی متابولیت‌های ثانویه در گیاه پروانش

ردیف	احتمال مهارکنندگی	رسپتور
۱	۰.۹۸۹	TP53 expression enhancer
۲	۰.۹۷۵	Tubulin antagonist
۳	۰.۹۶۶	CYP3A4 substrate
۴	۰.۹۶۳	P-glycoprotein substrate
۵	۰.۹۶۰	Lactose synthase inhibitor
۶	۰.۹۵۲	CYP3A substrate
۷	۰.۹۴۷	Xenobiotic-transporting ATPase inhibitor
۸	۰.۹۳۸	Cytostatic
۹	۰.۹۳۶	P-glycoprotein inhibitor
۱۰	۰.۹۳۱	Antineoplastic alkaloid
۱۱	۰.۸۷۲	Antineoplastic
۱۲	۰.۸۵۰	Beta tubulin antagonist
۱۳	۰.۸۰۹	Antineoplastic (Cervical cancer)
۱۴	۰.۷۶۸	Anticarcinogenic
۱۵	۰.۷۰۷	Antimitotic

بر اساس اطلاعات سرور way2drug که با استفاده از الگوریتم‌های هوش مصنوعی فعالیت می‌کند متابولیت



شکل ۳. نمودار رامچاندرا ن پلات پروتئین فسفولیکوپروتئین

جدول ۳. نتایج داگینگ متابولیت وین بلاستین با رسپتور فسفولیکوپروتئین

ViewDock - C:\Users\AL\Desktop\پروانش\docking result\vincristin\7.pdbqt

S	Score	RMSD l.b.	RMSD u.b.	HBonds (all)	HBond Ligand Atoms	HBond Receptor Atoms
V	-8.8	0.0	0.0	3	3	3
V	-8.6	2.548	9.109	1	1	1
V	-8.1	2.489	5.151	2	2	2
V	-7.8	1.898	4.382	3	3	2
V	-7.4	2.283	4.596	2	2	2
V	-7.3	2.414	5.292	0	0	0
V	-7.3	2.669	4.953	2	2	2
V	-6.7	2.209	8.571	3	3	3
V	-6.5	2.487	3.855	0	0	0
V	-6.5	2.572	5.083	1	1	1

جدول ۴. نتایج داگینگ متابولیت وین کریستین با رسپتور فسفولیکوپروتئین

ViewDock - C:\Users\AL\Desktop\پروانش\docking result\vinblastin\10.pdbqt

S	Score	RMSD l.b.	RMSD u.b.	HBonds (all)	HBond Ligand Atoms	HBond Receptor Atoms
V	-9.2	0.0	0.0	1	1	1
V	-8.9	1.624	2.193	1	1	1
V	-8.3	2.061	9.157	1	1	1
V	-7.7	2.154	4.541	1	1	1
V	-7.5	2.28	4.817	0	0	0
V	-6.9	2.015	5.036	1	1	1
V	-6.8	2.184	4.541	1	1	1
V	-6.7	2.454	9.057	2	2	2
V	-6.4	2.112	8.707	0	0	0



## شبیه‌سازی دینامیک مولکولی

میزان ثابت بودن لیگاند در جایگاه فعال فسفوگلیکوپروتئین با استفاده از نمودار RMSD مشخص می‌شود (شکل ۴). نتایج RMSD ثابت کرد که وین‌بلاستین از پایداری خوبی در برهم‌همکنش با فسفوگلیکوپروتئین برخوردار است.

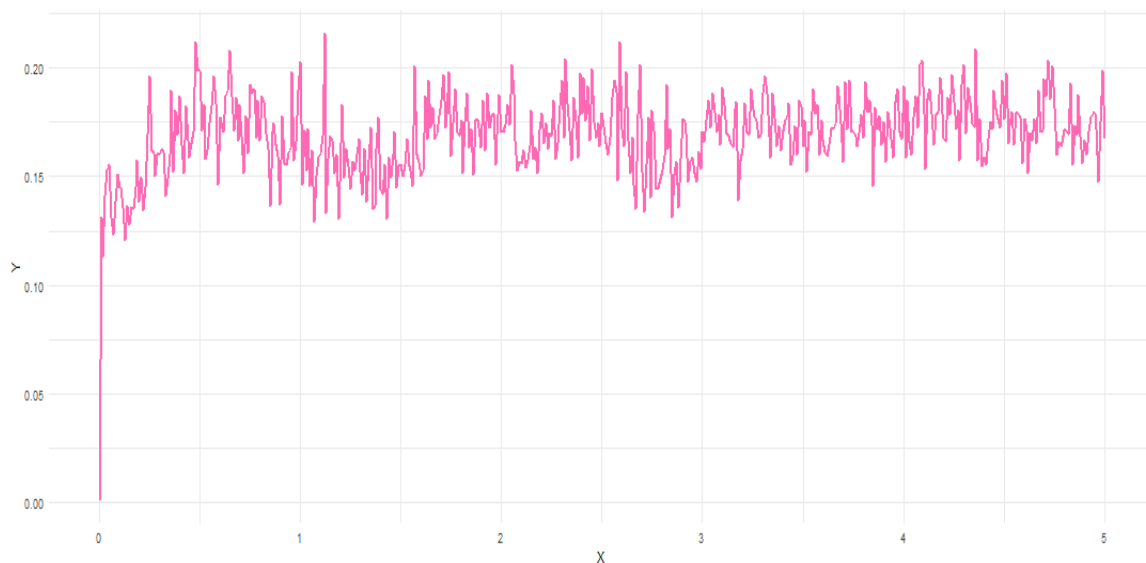
ایجاد می‌شود و فرد بیمار داروی وین‌بلاستین را مصرف می‌کند علاوه بر مهار فسفوگلیکوپروتئین، افزایش بیان TP53 رخ می‌دهد بنابراین این متابولیت ارزشمند به صورت هدفمند سرکوب ژن‌های مقاومت دارویی را انجام می‌دهد و منجر به افزایش بیان ژن‌های می‌شود که توسط فسفوگلیکوپروتئین محدود شده‌اند.

## آنالیز شبکه ژنی

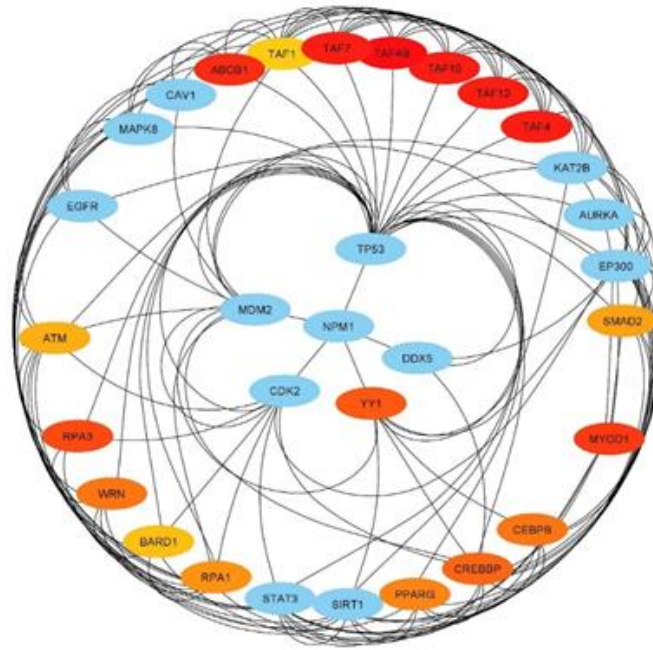
پس از آنکه مشخص شد که وین‌بلاستین توانایی مهار فسفوگلیکوپروتئین را دارد، آنالیز شبکه ژنی نشان داد (شکل ۵) که ارتباط هم‌بیانی بین افزایش بیان ژن MDM2 و کاهش بیان TP53 وجود دارد. همانگونه که در جدول (۲) مشخص شد که وین‌بلاستین توانایی افزایش بیان ژن TP53 را دارد از نظر تئوری بیان ژن و آنالیز هم‌بیانی ژن‌ها، فسفوگلیکوپروتئین منجر به سرکوب ژن TP53 می‌شود. نکته قابل توجه این است که در زمانی که مقاومت دارویی

## آنالیز مسیرهای سیگنالینگ

با استفاده از ۲۰ ژن‌ها که از آنالیز شبکه ژنی فسفوگلیکوپروتئین بدست آمد، آنالیز مسیرهای سیگنالینگ این ژن‌ها ثابت می‌کند زمانی که فسفوگلیکوپروتئین در ایجاد شبکه با ژن‌های سرکوب‌گر است DNA دچار آسیب‌های جدی می‌شود (شکل ۶). مسیرهای متفاوتی ممکن است در زمان بیان ژن MDR1 رخ دهد که با استفاده از پایگاه داده Enrichr شناسایی آن‌ها امکان‌پذیر است.



شکل ۴. آنالیز نمودار RMSD



شکل ۵. آنالیز شبکه ژنی با استفاده از سایتواسکب به منظور شناسایی اثرات متقابل بین ژن های هاب

#### The Overlap Between Signal Transduction Pathways that Contribute to a Range of LMNA Laminopathies WP4879

DNA IR-damage and cellular response via ATR WP4016

Pregnane X receptor pathway WP2876

DNA Repair Pathways Full Network WP4946

Integrated Cancer Pathway WP1971

G1 to S cell cycle control WP45

Irinotecan pathway WP229

Transcriptional cascade regulating adipogenesis WP4211

Homologous recombination WP186

Retinoblastoma gene in cancer WP2446

شکل ۶. آنالیز مسیرهای سیگنالینگ برای ۲۰ ژن هاب حاصل از آنالیز شبکه ژنی

### نتیجه گیری و پیشنهادها

در تحقیق حاضر، الگوی بیانی در اندام های مختلف در گیاه پروانش و شناسایی ژن های درگیر در مسیر بیوسنتزی وین کریستین و وین بلاستین با استفاده از داده های RNA-Seq بررسی شد. در این مطالعه تعداد ژن های با بیان افتراقی در اندام های مختلف برگ، گل و ریشه متفاوت بود، اما در مجموع تعداد ژن های با بیان افتراقی در اندام برگ در برابر ریشه بیشترین بود که نشان می دهد فعل و انفعالات سلولی در برگ نسبت به ریشه فعالتر است. زیرا بیشترین تعداد متابولیت ها در گیاه پروانش در برگ تولید می شود (سجادی ۱۳۷۹). در مطالعه حاضر ۶ و ۱۳

ژن مشترک به ترتیب برای برگ در برابر گل با گل در برابر ریشه و برای برگ در برابر گل با گل در برابر ریشه و برگ در برابر ریشه بدست آمد. نتایج آنالیز GO نشان داد که فرآیندهای زیستی اختصاصی پروانش، فرآیندهای متابولیکی و فرآیندهای سلولی مرتبط با متابولیت های ثانویه بیشترین تعداد ژن ها را داشتند که حاکی از وضعیت فعال متابولیکی سلول است.

بیشترین تغییر بیان ژن در هر سه هسته مشترک ژن های پاسخگو به اندام، مربوط به ژن های دخیل در مسیر بیوسنتزی وین کریستین و وین بلاستین مشاهده شد که بیشترین آن مربوط به ژن STR بود. بطوریکه در برگ و ریشه به ترتیب بیشترین و کمترین

توانایی آنها در مهار مقاومت دارویی فسفولیپوپروتئین (P-Glycoprotein) آنالیز داکینگ و دینامیک مولکولی انجام شد. مطالعه داکینگ ثابت کرد که این دو ماده می‌توانند فسفولیپوپروتئین را مهار کنند. با توجه به اینکه دو متابولیت ثانویه وین‌بلاستین، وین‌کریستین به خوبی با رسپتور فسفولیپوپروتئین متصل شدند هنوز مشخص نیست که اگر فرمولاسیون دارویی با این سه ترکیب طراحی شود آیا ممکن است مقاومت تومورهای بدخیم بخصوص در بخش شیمی درمانی بشکند؟ و یا پروتئین فسفولیپوپروتئین که با وین‌بلاستین اینترکشن قوی داده است آیا از مهاجرت تومورهای تخمدان به سایر بخش‌ها جلوگیری می‌شود؟ در مجموع فرآیند داکینگ مولکولی گام اولیه‌ای است برای پاسخ به این سوالات که در آینده قطعا این موارد، پژوهشگران را به سمت خود جذب خواهد کرد. نتایج دینامیک مولکولی با استفاده از نرم‌افزار GROMACS2022، پایداری و عملکرد کمپلکس وین‌بلاستین و وین‌کریستین را با پذیرنده نشان داده است. این کمپلکس، RMSD کمتر و تغییرات متعادل تری داشت، که تایید کننده ثبات کمپلکس وین‌بلاستین با فسفولیپوپروتئین بود.

در مجموع، تکنیک داکینگ برای یافتن ترکیب صحیح لیگاند (پپتید) و گیرنده طراحی شده است و اکنون برای موارد دیگری هم مورد استفاده قرار گرفته است. فرآیند اتصال یک مولکول کوچک (پپتید) به پروتئینی هدف (پاتونژن) آن ساده نیست، چندین عامل بر تعامل بین آنها تاثیر می‌گذارد. پروتکل‌های سریع و ارزان داکینگ را می‌توان با تکنیک‌های دقیق همانند دینامیک مولکولی برای پیش بینی قابل اعتمادتر کمپلکس‌های پروتئین-لیگاند ترکیب کرد. بنابراین، ترکیبی از دو تکنیک در یک پروتکل که در آن داکینگ برای غربالگری سریع کتابخانه‌های بزرگ و دینامیک مولکولی برای بررسی پایداری استفاده می‌شود می‌تواند راهگشای خوبی برای شناسایی پپتیدهای هدف باشد. با توجه به اطلاعات بدست آمده نتایج به خوبی برهم‌کنش پروتئین‌های هم خانواده مقاومت دارویی و همچنین پذیرنده را نشان داد. این فرآیند تاثیر متابولیت‌های ثانویه بر روی مقاومت دارویی را که مدتی است که مورد توجه محققان حوزه طراحی دارو قرار گرفته است را نشان داده است. با توجه به پیشرفت‌های اخیر بیوانفورماتیک و هوش مصنوعی توجه به فرمولاسیون دارویی و غربالگری مجازی پیشرفت‌های زیادی در این حوزه حاصل شده است و آینده نیز با توجه به فراگیر شدن علم بیوانفورماتیک، پیشرفت الگوریتم‌ها نیز اتفاق می‌افتد و در نهایت حوزه بیوانفورماتیک روز به روز گسترده‌تر خواهد شد.

بیان را نشان داد. نتایج مربوط به شبکه برهمکنش پروتئین-پروتئین نشان داد که ژن‌های DAT، STR، TDC، G10H، D4H، T16H2 بیشترین برهمکنش را با سایر ژن‌ها دارد که نقش مهمی را در مسیر بیوستتزی وین‌بلاستین و وینکریستین ایفا می‌کنند (Salami *et al.*, 2020). نتایج مربوط به آنالیز KEGG برای ژن‌های مشترک سه مقایسه نشان داد که در بین این مسیرها ژن‌های مرتبط با بیوستتزی متابولیت‌های ثانویه نقش مهمی دارند. در بافت برگ مسیرهای دیگر هم تغییر بیان معنی‌داری نشان دادند. این حالت نشان می‌دهد که بیشترین متابولیت‌ها در برگ فعال هستند. در حالیکه در ریشه به‌طور اختصاصی‌تر مسیرهای متابولیت‌های ثانویه کمتری تغییر کردند (Liu *et al.*, 2022).

دو ماده وین‌بلاستین و وینکریستین در مسیر ترپنوئید ایندول آلکالوئیدی به میزان کم در گیاه ساخته می‌شوند. فراهم نبودن سنتز آن بروش شیمیایی باعث شد تا استفاده از روش‌های زیست‌فناوری در راستای افزایش متابولیت‌های ثانویه از اهمیت بسزایی برخوردار شود. برخی از محرک‌های رشد به‌نظر می‌رسد توانایی افزایش این ترکیبات را داشته باشند. یکی از این محرک‌های رشد اسید سالیسیلیک است که سیستم دفاعی گیاه را تحریک می‌کند که در نهایت باعث بهبود عملکرد و افزایش محتوای کلی آلکالوئیدهای وین‌بلاستین و وین‌کریستین می‌شود. (Guo *et al.*, 2014; سلطانی، ۱۴۰۰).

از بین ۹ ژنی که در هر سه بافت تغییرات بیانی داشتند، به ترتیب بیشترین و کمترین بیان آنها مربوط به برگ و ریشه بود. آنالیزها نشان داد که این ۹ ژن در مسیر بیوستتزی وین‌بلاستین و وین‌کریستین نقش دارند. با توجه به اینکه بیشترین میزان این ترکیبات در برگ است به همین دلیل بیشترین تغییرات بیانی را در برگ نشان دادند. در این میان ژن STR، علاوه بر اینکه، بالاترین برهمکنش را به بقیه ژن‌های پاسخ به بافتها نشان داد، به شدت تحت تاثیر بافت بود. بنابراین انتخاب دقیق استراتژی تنظیم بیان ژن در بافت‌های مختلف می‌تواند نقش مهمی در افزایش متابولیت‌های ثانویه همچون وینکریستین و وین‌بلاستین داشته باشد (Soltani *et al.*, 2021).

در این تحقیق، تعدادی ژن به عنوان کلنید برای بهبود عملکرد وین‌کریستین و وین‌بلاستین معرفی شد که می‌تواند برای برنامه‌های اصلاحی و مهندسی ژنتیک مورد استفاده قرار گیرد. در ادامه به منظور بررسی اثرپذیری خواص ضد سرطانی متابولیت‌های ثانویه گیاه پروانش (وین‌بلاستین و وین‌کریستین) و

## سپاسگزاری

این پروژه با حمایت مالی مرکز مطالعات و همکاریهای علمی بین‌المللی وزارت علوم تحقیقات و فناوری انجام شده است.

## References

- Anders, S., Pyl, P.T., Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31, 166-169.
- Bomzan, D., Shilpashree, H., Nagegowda, D. (2022). Agrobacterium-Mediated in Planta Transformation in Periwinkle. doi: 10.1007/978-1-0716-2349-7\_22.
- Fatemi, F., Dadkhah, A., Honardoost, M., Ebrahimi, M., Hedayati, M., Shadnough, M. (2007). Mechanism of drug resistant in cancer. *Research in Medicine*, 31, 91-97. (in Persian)
- Guo, X.R., Chang, B.W., Zu, Y.G., Tang, Z.H. (2014). The impacts of increased nitrate supply on *Catharanthus roseus* growth and alkaloid accumulations under ultraviolet-B stress. *J Plant Interact* 9(1), 640-646.
- Heijden, R., Jacobs, D.L., Snoeijer, W., Hallard, D., Verpoorte, R. (2004). The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem* 11(5), 607-628.
- Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1), 27-30.
- Loyola-Vargas, V.M., Broeckling, C.D., Badri, D.V., Vivanco, J.M. (2007). Effect of transporters on the secretion of phytochemicals by the roots of *Arabidopsis thaliana*. *Planta* 225, 301-310.
- Noble, R.L. (1990). The discovery of the vinca alkaloids-chemotherapeutic agents against cancer. *Biochem Cell Biol*, 68(12):1344-1351.
- Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650-1667.
- Sajjadi, E., and Hallard, D. (1959). Comparison of different organs of *Parvanesh* plant in terms of indole alkaloids production. *Pajohesh Sazandegi*, 13, 30-33. (in Persian)
- Soltani, N., Nazarian-Firouzabadi, F., Shafeinia, A., Sadr, A.S., Shirali, M. (2021). Expression analysis of some genes involved in Terpenoid Indole Alkaloids pathway in periwinkle (*Catharanthus roseus*) following foliar application of salicylic acid. *Novin Genetic*, 16, 25-32. (in Persian)
- Sahandi Khalifeh-Kandy, A., Pazhouhandeh, M and Mohajjel-Shoja H. (2016). The effect of rolC gene the medicinal plant *Catharanthus roseus*. *Genetic engineering and biosafety*, 5, 41-50. (in Persian)
- Subramanian, A., Kuehn, H., Gould, J., Tamayo, P. & Mesirov, J. P. GSEA-P: A desktop application for gene set enrichment analysis. *Bioinformatics* 23(23), 3251-3253.
- Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., & Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Research*, 45(W1), W122–W129.
- Verpoorte, R. (2004). The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem* 11(5), 607-628.