



همسانه‌سازی و بررسی بیوانفورماتیکی ژن‌های *CCD4a* و *CCD4b* در زعفران (*Crocus sativus* L.) ایران

محمدجواد حبیب زاده^۱، ابراهیم دورانی علیایی^{۲*}، سیدمهدی زیارت‌نیا^۳ و مصطفی ولی زاده^۴

تاریخ پذیرش: ۲۱ مهر ۱۳۹۸

تاریخ دریافت: ۳۱ فروردین ۱۳۹۸

حبیب زاده، م.، دورانی علیایی، ا.، زیارت‌نیا، س.م.، و ولی زاده، م. ۱۳۹۹. همسانه‌سازی و بررسی بیوانفورماتیکی ژن‌های *CCD4a* و *CCD4b* در زعفران (*Crocus sativus* L.) ایران. زراعت و فناوری زعفران، ۸(۲): ۲۱۱-۲۲۹.

چکیده

زعفران به عنوان یکی از گران‌بهارترین ادویه‌ها و رنگ‌های طبیعی است که در صنایع مختلف از جمله غذایی، دارویی و آرایشی-بهداشتی مورد استفاده قرار می‌گیرد. در سال‌های اخیر، خانواده‌ای از آنزیم‌ها که پیش ماده‌های کاروتنوئیدی را در جایگاه پیوندهای دوگانه برش می‌دهند، در گیاهان شناسایی و معرفی شده‌اند. به این خانواده از آنزیم‌های برش‌دهنده کارتنوئید دی اکسیژناز (CCD) گفته می‌شوند. در این تحقیق به دلیل اهمیت ژن‌های *CCD* در بیوسنتز آپوکاروتنوئیدهای زعفران، دو ایزوفرم از این ژن با استفاده از روش نسخه‌برداری معکوس همسانه‌سازی، تعیین توالی و نتایج با موارد مشابه خارجی مورد مقایسه قرار گرفتند. بررسی‌های بیوانفورماتیکی شامل بررسی ارتباطات خویشاوندی و نیز ساختارهای پروتئینی مورد ارزیابی قرار گرفت. مدل‌سازی سه بعدی این پروتئین‌ها به روش همولوژی مدلیتگ و با استفاده از پایگاه Swiss Model پس از انتخاب الگوی مناسب انجام گرفت. همچنین جهت اعتبارسنجی ساختاری مدل ترسیم شده سه بعدی، پلات رامانچاندرا ترسیم گردید. نتایج حاصله نشان داد دو ایزوفرم *CCD4a* و *CCD4b* هر دو دارای دو اگزون (۶۴۱ و ۱۰۹۹ جفت باز برای *CCD4a* و ۶۱۴ و ۱۰۹۹ جفت باز برای *CCD4b*) و یک اینترون (۶۷۰ جفت باز در *CCD4a* و ۶۶۸ جفت باز در *CCD4b*) هستند. بررسی *In silico* خصوصیات فیزیکوشیمیایی پروتئین‌های *CsCCD4a* و *CsCCD4b* بیانگر این حقیقت بود که پروتئین‌های بدست آمده از این دو ایزوفرم از نظر وزن مولکولی، تعداد اسیدهای آمینه، نقطه ایزوالکتریک، شاخص آلیفاتیک، شاخص ناپایداری و حلالیت مشابه می‌باشند. نتایج بررسی ساختارهای سه بعدی بدست آمده نشان داد این دو ایزوفرم دارای ساختارهای سه بعدی مشابهی می‌باشند. یافته‌های این تحقیق می‌تواند اطلاعات با ارزشی در رابطه با رفتار و واکنش آنزیم *CCD4* در مسیر سنتز آپوکاروتنوئیدهای زعفران فراهم کند و همچنین این نتایج می‌تواند در برنامه‌های آتی اصلاح ژنتیکی زعفران ایران مفید واقع گردد.

کلمات کلیدی: *CCD4*، بررسی فیلوژنتیک، بیوانفورماتیک، مدل‌سازی.

- ۱- دانشجوی دکتری رشته بیوتکنولوژی کشاورزی، گروه به‌نژادی و بیوتکنولوژی گیاهی، دانشکده کشاورزی دانشگاه تبریز، تبریز، ایران.
 - ۲- دانشیار، گروه به‌نژادی و بیوتکنولوژی گیاهی، دانشکده کشاورزی دانشگاه تبریز، تبریز، ایران
 - ۳- استادیار، گروه زیست فناوری مواد غذایی، مؤسسه پژوهشی علوم و صنایع غذایی، مشهد، ایران.
 - ۴- استاد، گروه به‌نژادی و بیوتکنولوژی گیاهی، دانشکده کشاورزی دانشگاه تبریز، تبریز، ایران.
- *- نویسنده مسؤول: (dorani@tabrizu.ac.ir)

مقدمه

زعفران (*Crocus sativus* L.) به عنوان طلای سرخ، گران بهاترین چاشنی غذایی در جهان می‌باشد که ارزش اقتصادی آن به کلاله خشک شده، اندام جنسی ماده گل زعفران، مربوط می‌باشد (Bathaie et al., 2006). عمده ترکیبات دارویی زعفران شامل آپوکاروتنوئیدهای سیس و ترانس کروسین‌ها، پیکروکروسین‌ها و سافرانال می‌باشند (Hosseinpour Azad et al., 2016). کلاله زعفران دارای تعدادی ترکیبات فعال و غیرفعال دیگری از خانواده کاروتنوئیدها شامل زیزانتین، لیکوپین و انواع مختلف آلفا و بتا کاروتن نیز می‌باشد همچنین این گیاه به عنوان منبع غنی از محتویات فنولیکی با فعالیت اثبات شده آنتی‌اکسیدانی می‌باشد (Goli et al., 2012).

کلاله زعفران حاوی مقادیر زیادی (بیش از ۸ درصد از وزن خشک) از آپوکاروتنوئیدهای گلیکوزیده می‌باشد که عامل ایجاد رنگ، طعم و عطر ادویه زعفران می‌باشند (Hosseinpour Azad et al., 2016). آپوکاروتنوئیدها ترکیباتی هستند که فرآیندهای رشد و فیزیولوژیکی حیاتی را در گیاهان تنظیم می‌کنند (Baba et al., 2015). با وجود پیشرفت‌های قابل توجهی که در مورد تولید آپوکاروتنوئیدها در گیاهان صورت گرفته است، مکانیسم تنظیم و بیوسنتز آنها به صورت روشن مشخص نشده است. مطالعات صورت گرفته نشان می‌دهد در کلاله زعفران، کروسیتین از پیش ماده‌های بتا-کاروتن و زیزانتین تولید می‌گردد (Rubio et al., 2009). این مولکول‌های کاروتنوئیدی به واسطه واکنش برشی که توسط تعدادی از آنزیم‌های خانواده CCD^۱ صورت می‌گیرد، تشکیل می‌شوند (Giuliano et al., 2003). تاکنون، انواع مختلفی از آنزیم‌های کاروتنوئید اکسیژناز در گیاهان شناسایی و طبقه‌بندی شده‌اند. اولین ژن شناسایی شده از

خانواده CCDها، ژن *Vp14^۲* ذرت بود که در تشکیل اسیدآبسیک نقش دارد و در حال حاضر تحت عنوان NCED^۳ به عنوان یک کلاس از خانواده بزرگ CCDها طبقه بندی می‌شوند. (Gomez-Gomez et al., 2010) علاوه بر NCED، CCDها در گیاهان به پنج کلاس شامل CCD1، CCD2، CCD4، CCD7 و CCD8 طبقه بندی می‌شوند (Ahrazem et al., 2016).

آنزیم‌های متعلق به کلاس CCD4 در پایانه N خود دارای دمینی برای عبور از پلاستیدها می‌باشند. این آنزیم‌ها پیوند دوگانه را در موقعیت‌های ۱۰-۹ (۱۰'-۹') برش می‌زنند و به نظر می‌رسد که حداقل برای پیش ماده‌های بتاکاروتن از آنزیم‌های CCD1 فعال تر باشند (Rubio et al., 2008)، همچنین آنزیم‌های CCD4 پیش ماده مشخص تری نسبت به CCD1 دارند (Rubio et al., 2009). در مطالعات صورت گرفته برای آنزیم CCD4 دو آنزیم CCD4a و CCD4b شناسایی شده است که این آنزیم‌ها کاروتنوئیدها را در موقعیت‌های مشابهی برش می‌زنند (Giuliano et al., 2003).

گمز-گمز و همکاران (Gomez-Gomez et al., 2010) با جداسازی و مقایسه ساختار ژنتیکی ژن‌های CCD4 از کلاله زعفران و مقایسه آلل‌های آن با دیگر گونه‌های گیاهی به توصیف اینترون‌ها و حفاظت شدگی آن در بین خانواده CCDها پرداختند. مطالعات آن‌ها نشان داد در ژن CCD4 برخلاف ژن‌های CCD1، CCD7 و CCD8 اینترون وجود دارد و به نظر می‌رسد موقعیت اینترون در ساختار ژن‌های CCD4 بسیار متحرک است. در مطالعه‌ای دیگر اهرازم و همکاران (Ahrazem et al., 2010) دو ژن CCD4a و CCD4b را از

۲ - VIVIPAROUS14

۳ - 9-cis-epoxycarotenoid dioxygenas

۱ - Carotenoid Cleavage Dioxygenases

ساختار آگزون و اینترون ژن‌ها و سپس تجزیه و تحلیل فیلوژنتیکی آن‌ها با استفاده از برنامه‌های نرم‌افزاری مرتبط می‌باشد (Wang et al., 2013). روش فیلوژنتیک از اطلاعات مولکولی جهت بیان ارتباطات تکاملی بین ژن‌ها و بین موجودات استفاده می‌شود. به عبارت دیگر هدف از بررسی‌های فیلوژنتیکی کشف ترتیبی از شاخه‌ها به صورت درخت‌هایی است که نشان دهنده بهترین روابط بین منابع مختلف براساس توالی‌های یک ژن می‌باشد (Naghavi et al., 2009). همچنین بررسی‌های فیلوژنتیکی تخمینی از چگونگی اشتقاق اعضای یک خانواده در طی تکامل که به وسیله جهش‌های جایگزینی، حذف، ازدیاد و جهش با آرایش مجدد ایجاد شده و در معرض انتخاب طبیعی تثبیت شده اند (Golovnina et al., 2007).

سرورهای آنلاین و برنامه‌های محاسباتی از ابزارهای رایج در بررسی و شناسایی توالی پروتئین می‌باشند (Sivakumar, 2005). با استفاده از ابزارهای محاسباتی خصوصیات فیزیکوشیمیایی و ساختاری پروتئین‌ها به صورت بهتری قابل فهم می‌باشند. تعدادی از ابزارهای محاسباتی جهت شناخت و پیش‌گویی ساختار پروتئین‌ها به‌وجود آمده‌اند. توالی آمینواسیدی پروتئین مورد نظر می‌تواند به عنوان ورودی در بسیاری از برنامه‌ها و سرورها به کار رفته و اطلاعات بی‌شماری را از پروتئین ارائه دهد (Tramontano et al., 2001).

در این مطالعه با هدف شناخت کامل برخی ژن‌های درگیر در مسیر بیوسنتز ترکیبات آپوکارتونوئیدی زعفران ایران به عنوان مهد تولید این محصول راهبردی و ارزشمند و مقایسه آن با گزارش‌های مشابه در برخی کشورهای پیشرو، در گام نخست دو ژن *CCD4a* و *CCD4b* از زعفران ایران از DNA ژنومی و cDNA جداسازی و همسازسازی گردید. سپس با بررسی توالی‌های اسیدآمینه‌ای و نوکلئوتیدی، بررسی‌های ساختاری شامل پیش‌بینی جایگاه، عملکرد، ساختمان دوم و سوم پروتئین‌ها، تغییرات پس از ترجمه و نیز ارتباطات خویشاوندی این آنزیم‌ها

C. sativus جداسازی نمودند. آن‌ها نشان دادند که *CsCCD4a* با ۵۸۰ اسیدآمینه و *CCD4b* با ۵۶۹ اسیدآمینه دارای ۹۸٪ تشابه (۳۶۹ اسیدآمینه مشابه) می‌باشند (Gomez- et al., 2010).

در پژوهشی که بر روی بیان ژن‌های *CCD* در طول دوره رشد و توسعه کلاله زعفران انجام شد، مشاهده گردید که mRNAهای مسئول ژن‌های دی‌اکسیژناز، الگوهای بیانی متفاوتی در طول مراحل رشد و توسعه کلاله دارند. در این مطالعه نشان داده شد که رونوشت‌های *CCD4a* عملاً در مراحل زرد و نارنجی غیرقابل شناسایی بوده و در مرحله سرخ رنگ کلاله در بالاترین سطح خود قرار می‌گیرند ولی میزان رونوشت‌های *CCD4b* در تمام مراحل رشد و توسعه کلاله به‌طور پیوسته بالا بود (Grilli et al., 2004).

از آنجا که ارزش بالای اقتصادی زعفران به واسطه سطوح بالای آپوکارتونوئیدهای کلاله آن می‌باشد، لذا مطالعه سنتز و تشکیل آپوکارتونوئیدهای حاصل از شکست اکسیداتیوی کاروتنوئیدهای آن مفید خواهد بود (Ahrazem et al., 2010) و هر گونه فعالیت به‌نژادی در زمینه افزایش تولید آپوکارتونوئیدها، مستلزم شناخت و درک تنظیم مسیر سنتز آپوکارتونوئیدها می‌باشد (Bhat et al., 2018). به این منظور، با وجود پیشرفت‌های قابل توجهی که در مورد شناسایی خصوصیات و مشخصات ژن‌های مسیر آپوکارتونوئیدها در گیاهان صورت گرفته است، ولی ساختار ژن‌ها و پروتئین‌های حاصل از بیان خانواده *CCD* در زعفران ایران تاکنون مورد بررسی قرار نگرفته است.

شناخت ساختار به منظور توصیف خانواده‌های ژنی و نیز ایجاد ارتباطات تکاملی مهم می‌باشد (Ahrazem et al., 2010)، انجام بررسی‌های فیلوژنتیکی و بیوانفورماتیکی به منظور بررسی ساختارهای ژنی و پروتئینی امری ضروری است. به‌طور معمول تجزیه و تحلیل فیلوژنتیکی ژن‌ها نیازمند استخراج

مقایسه ژن‌های *CCD4*، نسبت به همسانه‌سازی و توالی‌یابی ژن‌های *CCD4a* و *CCD4b* زعفران ایران اقدام گردید. در این راستا واکنش‌های اتصال محصولات PCR با استفاده از کیت TA cloning (USA, Fermentas) توسط آنزیم T₄ DNA Ligase به داخل ناقل همسانه‌سازی pTZ57 R/T انجام شدند. تراریختگی باکتری *E. coli* سویه DH₅α با استفاده از روش شوک حرارتی (Maniatis et al., 1995) انجام گردید. سلول‌های تراریخته در محیط LB⁻ agar حاوی آنتی‌بیوتیک آمپی‌سیلین (۵۰ μg.ml⁻¹) گزینش شدند. غربالگری جهت تأیید ورود ژن مورد نظر به باکتری به دو روش colony PCR و هضم آنزیمی با آنزیم برشی BamH I (Jena Bioscience, Germany) انجام گرفت. پلاسمیدها از کلون‌های مختلف با استفاده از کیت استخراج پلاسمید TPG (پیشگامان انتقال ژن - ایران) جداسازی شده و پس از خالص‌سازی جهت توالی‌یابی با آغازگرهای یونیورسال M13 و همچنین آغازگرهای اختصاصی به شرکت ماکروژن کره جنوبی (Korea, Macrogen) ارسال شدند.

استخراج RNA و سنتز cDNA

RNA کل از کلاله گل زعفران با استفاده از کیت استخراج RNA (AccuZol™ Total RNA Extraction Solution) (RNA from Bioneer USA) استخراج گردید و پس از تیمار با آنزیم DNase، سنتز cDNA با استفاده از آغازگر تصادفی هگزامر و آنزیم نسخه‌برداری معکوس (RevertAID™ M-MuLV) (USA, Thermo Fisher Scientific) بر اساس دستورالعمل شرکت سازنده صورت گرفت.

بررسی و تحلیل گردید. همچنین نتایج حاصله با در سطح ملکولی با نتایج گزارش شده موجود در بانک ژن مقایسه گردید.

مواد و روش‌ها

استخراج DNA ژنومی

استخراج DNA ژنومی به روش CTAB تغییر یافته (Beiki et al., 2011) و از نمونه‌های پیاز و برگ زعفران که به روش انجمادی خشک شده، انجام گرفت. جهت بررسی کمیت و کیفیت DNA استخراج شده از نانودراپ (BOECO N-1C, Germany) و ژل آگارز یک درصد استفاده شد. به منظور شناسایی توالی‌های ژنی *CCD4a* و *CCD4b* زعفران ایران، اطلاعات مربوط به توالی‌های نوکلئوتیدی DNA و mRNA این ژن‌ها از بانک ژن NCBI استخراج گردید. سپس با استفاده از نرم‌افزار Oligo version 7.60 نسبت به طراحی آغازگر برای نواحی ابتدا و انتهای ژن‌های *CCD4a* و *CCD4b* و همچنین یک آغازگر مشترک (Fp) در ناحیه بالادست ژن‌های فوق‌الذکر اقدام شد (جدول ۱). ژن‌های *CCD4a* و *CCD4b* با واکنش زنجیره‌ای پلیمرز (PCR) از DNA ژنومی و cDNA با آغازگرهای طراحی شده با استفاده از ترکیب آنزیم‌های *Taq polymerase* و *Pfu polymerase* تکثیر شدند. تمامی محصولات حاصل از PCR با استفاده از کیت GeNet Bio از ژل آگارز خالص سازی شدند.

همسانه‌سازی و توالی‌یابی

با توجه به این که تاکنون در خصوص توالی ژن‌های *CCD4* زعفران ایران پژوهشی صورت نگرفته و توالی آنها در پایگاه داده‌ها ثبت نشده و تنها اطلاعات مربوط به پژوهش‌های انجام شده در سایر کشورها موجود می‌باشد به منظور بررسی و

۱ - www.ncbi.nih.gov/genbank

۲- Polymerase Chain Reaction

جدول ۱- آغازگرهای مورد استفاده در بررسی ژن‌های CCD4a و CCD4b
Table 1- Primers used in the investigation of CCD4a and CCD4b genes

ژن	آغازگر	توالی
Gene	Primer	Sequence
CCD4a	F4a	5'- ATGGAGTATCGGTTATCCTCC -3'
	R4a	5'- TAC TGC TGT GAC AGC AGC TCA GC -3'
	F4b	5'- ATG GAG TAT CGG TTG TCA TCC TCC T -3'
CCD4b	R4b	5'- TAC TGC TGT GAC AGC AGC TCA GC -3'
	Fp	5'- TCC TTT CGA ATA CTT TTC CTC ATC C -3'

Krogh et al. (http://www.cbs.dtu.dk/services/TMHMM) (2001) و برای بررسی هدف گیری پروتئین از بانک‌های TargetP (Almagro Armenteros et al., 2019) و ChloroP (Emanuelsson et al., 1999) بهره گرفته شد. مدل سازی به روش مدلینگ و با استفاده از پایگاه اطلاعاتی SWISS-MODEL (https://swissmodel.expasy.org/interactive) (Waterhouse et al., 2018; Bienert et al., 2017; Guex et al., 2009) انجام شد. در نهایت به منظور ارزیابی کیفیت مدل سه بعدی با استفاده از پلات رامانچاندرا (Bertoni et al., 2017) مورد ارزیابی قرار گرفت.

بررسی ویژگی های فیلوژنتیکی

ایجاد یک درخت فیلوژنتیکی بر اساس سه مرحله استوار است: الف) هم ردیف کردن توالی‌های DNA یا پروتئین و به دست آوردن داده‌های مقایسه‌ای که جهت ساخت درخت مورد استفاده قرار می‌گیرند. ب) تبدیل داده‌های مقایسه‌ای به درخت بازسازی شده. ج) ارزیابی صحت درخت بازسازی شده.

در این تحقیق، توالی‌ها با نرم افزار Mega 6 (Tamura et al., 2013) هم ردیف شده و با استفاده از روش اتصال مجاور^۱ که یکی از روش‌های رایج ایجاد درخت تکاملی است، رسم گردید. به منظور ارزیابی صحت درخت‌ها از آزمون Bootstrap که یک روش ساده برای آزمودن پایداری نسبی گروه‌ها در درون

بررسی توالی نوکلئوتیدی

توالی‌یابی قطعات مورد نظر توسط شرکت ماکروژن کره جنوبی انجام و نتایج برای حذف آلودگی ناقل با استفاده از نرم‌افزار Chromas version 2.6.2 بررسی و اصلاح گردید. به منظور یافتن توالی‌های مشابه از پایگاه داده NCBI، برنامه NBLAST استفاده شد. همچنین توالی اسید آمینه‌ها با ترجمه توالی‌های نوکلئوتیدی ژن‌های CCD4a و CCD4b به وسیله نرم‌افزارهای ترجمه داده پایگاه مولکولی Expasy (https://web.expasy.org/translate) بدست آمد (Artimo et al., 2012). برای بررسی ویژگی‌های فیزیکیوشیمیایی پروتئین‌ها از نرم‌افزار ProtParam (http://web.expasy.org/protparam) (Gasteiger et al., 2005)، پیش‌بینی ساختار دوم پروتئین‌های بیان شونده از SOPMA (https://npsaprabi.ibcp.fr) (Combet et al., 2000)، پیش‌بینی توالی حفاظت شده با نرم‌افزار Pfam (http://pfam.xfam.org) (El-Gebali et al., 2019)، شناسایی جایگاه درون سلولی از نرم‌افزار ProtComp (http://www.softberry.com) ، الگوی آبگریزی - آبدوستی بر اساس روش کیت و دولیتل (Kyte & Doolittle, 1982) از نرم‌افزار ProtScale (http://web.expasy.org/protscale) (Gasteiger et al., 2005)، وجود سیگنال‌های پپتیدی در پروتئین از نرم‌افزار 4.1 SignalP (Petersen et al., 2011) ، ناحیه تراغشایی و جهت گیری مرتبط با پروتئین CCD4 از نرم‌افزار TMHMM

۱ - Neighbor – joining method

یک درخت است (Nassaj Hoseini & Shamsbakhsh, 2010)، با تکرار پذیری ۱۰۰۰ استفاده شد.

نتایج و بحث

تکثیر، همسانه سازی و توالی‌یابی

نتایج بدست آمده از توالی‌یابی قطعات CCD4 تکثیر شده به روش PCR، و به وسیله آغازگرهای اختصاصی از روی DNA ژنومی و cDNA نشان داد یک اینترون و دو اگزون در توالی هر یک از ژن‌های مورد مطالعه قرار دارد. برای ژن CCD4a مناطق اگزونی ۱ و ۲ به ترتیب ۶۴۱ و ۱۰۹۹ جفت باز و منطقه اینترون شامل ۶۷۰ جفت باز می‌باشد همچنین برای ژن CCD4b طول اگزون ۱، ۶۱۴ جفت باز و طول اگزون ۲، ۱۰۹۹ جفت باز به همراه یک اینترون ۶۶۸ جفت بازی دیده می‌شود. با رونویسی و ترجمه این دو ژن، پروتئین‌هایی به طول ۵۷۹ اسیدآمینه برای CCD4a و ۵۷۰ اسیدآمینه برای CCD4b تشکیل می‌شود.

نتایج بیوانفورماتیکی ساختار پروتئین‌ها

به منظور انجام مطالعات پروتئینی، از تجزیه و تحلیل بیوانفورماتیکی اسیدهای آمینه استفاده گردید. نتایج حاصل از این محاسبات نشان داد که نقطه ایزو الکتریک تئوریتیکال برای CCD4a ۶٫۴۸ و برای CCD4b ۶٫۴۱ می‌باشد. همچنین وزن مولکولی این دو پروتئین به ترتیب ۶۳٫۸ و ۶۳ KDa محاسبه گردید. شاخص ناپایداری پروتئین‌ها به ترتیب برای CCD4a و CCD4b ۴۲٫۶۳ و ۴۱٫۸۰ بود. شاخص آلیفاتیک، به عنوان یک عامل مهم در ارزیابی مقاومت پروتئین‌ها در برابر حرارت، برای CCD4a و CCD4b به ترتیب برابر ۸۲٫۵۹ و ۸۳٫۰۵ محاسبه گردید. از مجموع کل اسیدآمینه‌های CCD4a تعداد ۶۶ اسیدآمینه بار منفی (Asp+ Glu) و ۶۱ اسیدآمینه بار مثبت (Arg+ Lys) داشته و برای CCD4b این تعداد شامل ۶۷

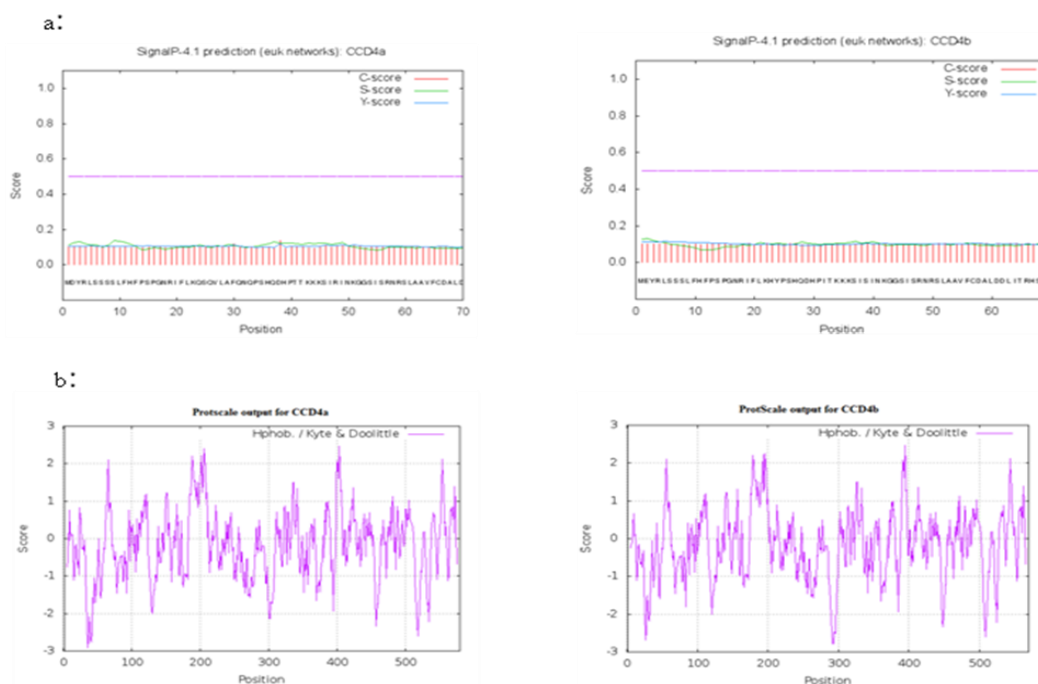
اسیدآمینه بار منفی (Asp+ Glu) و ۶۱ اسیدآمینه بار مثبت (Arg+ Lys) بود (جدول ۲). برنامه SignalP نسخه 4.1 که یکی از ابزارهای مطمئن با صحت ۸۷ درصد (Bendtsen et al., 2004) بین نرم‌افزارها جهت پیش‌بینی سیگنال‌های پپتیدی پروتئین‌ها و مناطق برش آن‌ها می‌باشد، وجود سیگنال پپتیدی در پروتئین‌ها را مشخص نکرد (شکل ۱a). همچنین در خروجی بررسی انجام شده توسط این نرم‌افزار، مقدار نمره D که معیار تمایز بین پروتئین‌های ترش‌چی و غیر ترش‌چی است، ۰٫۴۵۰ برای هر دو پروتئین بود. برخی گزارشات علمی حکایت از آن دارد که توالی پروتئینی با نمره D بالاتر از ۰/۵ احتمالاً دارای سیگنال پپتیدی هستند (Petersen et al., 2011). در الگوی آگریزی-آبوستی برای زنجیره پلی‌پپتیدی CCD4a، کمترین امتیاز (۲٫۹۲۲-) در موقعیت ۳۵ برای اسیدآمینه سرین (S) و بیشترین مقادیر آبوستی برای اسیدآمینه والین در موقعیت ۴۰۳ (با امتیاز ۲٫۴۶۷) و برای زنجیره پلی‌پپتیدی CCD4b، بیشترین امتیاز در موقعیت ۳۹۴ (والین) با امتیاز ۲٫۴۶۷ و کمترین امتیاز آبوستی (۲٫۷۸۸-) در موقعیت‌های ۲۹۲ و ۲۹۳ شامل اسیدآمینه‌های آرژنین و اسپارتیک‌اسید بدست آمد (شکل ۱b).

بررسی توالی حفاظت شده پروتئین‌های CCD4a و CCD4b، یک توالی را برای هر پروتئین به عنوان الگوی حفاظت شده پروتئین مشخص نمود که این توالی، (۵۷۱-۹۷) برای CCD4a و (۵۶۱-۸۸) برای CCD4b، در خانواده کاروتنوئیداکسیژنازها قرار می‌گیرد. در پیش‌بینی ساختار دوم پروتئین‌های CCD4a و CCD4b، بر اساس نتایج بدست آمده از نرم‌افزار SOPMA برای CCD4a، ۸۹ ساختار مارپیچ آلفا (۱۵٫۳۷٪)، ۱۵۱ Extended strand (۲۶٫۰۸٪)، ۳۱ Beta turn (۵٫۳۵٪) و ۳۰۸ ساختار Random coil و برای CCD4b، ۹۰ ساختار آلفا هلیکس (۱۵٫۷۹٪) و ۱۴۰ Extended strand (۲۴٫۵۶٪)، ۳۸ Beta turn (۶٫۶۷٪) و ۳۰۲ ساختار Random coil مشاهده

جدول ۲ - نتایج حاصل از بررسی خصوصیات فیزیوشیمیایی ژن‌های CCD4b و CCD4a

Table 2- The physicochemical characteristics of the protein sequences resulted from CCD4a and CCD4b

	فرمول شیمیایی Formula	شاخص آلیفاتیک Aliphatic index	تعداد اسیدهای آمینه Number of amino acids	وزن مولکولی Molecular weight (Da)	نقطه ایزوالکتریک pI	شاخص ناپایداری Instability index (II)	میانگین نسبی آبدوستی - آبگریزی Gravy
CCD4a	C ₂₈₆₆ H ₄₄₃₆ N ₇₉₀ O ₈₃₃ S ₁₇	82.59	579	63832.56	6.48	42.63	-0.179
CCD4b	C ₂₈₃₀ H ₄₃₇₀ N ₇₈₀ O ₈₂₁ S ₁₆	83.05	570	62969.52	6.41	41.80	-0.171



شکل ۱- بررسی سیگنال‌های پپتیدی در پروتئین‌های CCD4a و CCD4b. a: نتایج خروجی‌های نرم‌افزار SignalP شامل C-score که محل برش را شناسایی می‌کند، S-score که موقعیت توالی راهنما را تشخیص می‌دهد. Y-score که مشتقی از نتایج نمره‌های C و S با پیش‌بینی دقیق‌تر محل‌های برش نسبت به نمره C خام است، b: الگوی آبگریزی-آبدوستی پروتئین‌های CCD4a و CCD4b

Figure 1- a: Peptide signaling in CCD4a and CCD4b proteins. a: Results of SignalP software for peptide signaling characterization include a C-score that identifies the cleavage site, an S-score that detects the position of the signal peptide. The Y-score, which is a derivative of the C-score combined with the S-score resulting in a better cleavage site prediction than the raw C-score alone, b: Hydrophobic-hydrophilic pattern of proteins.

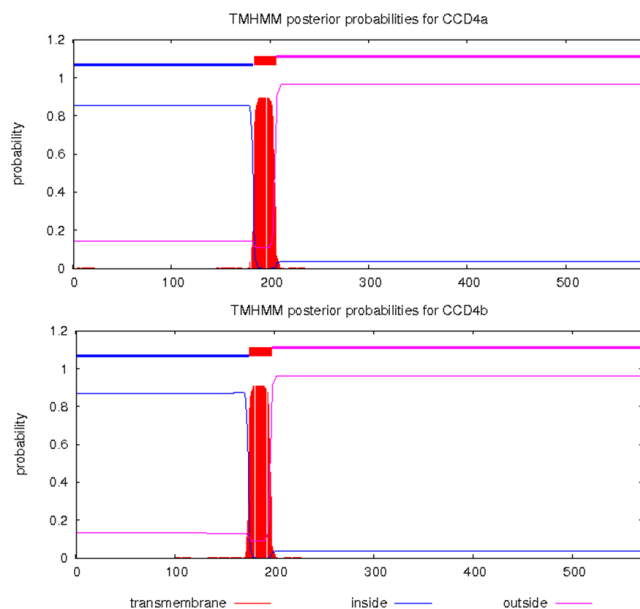
بررسی‌های هدف‌گیری با نرم‌افزارهای ChloroP 1.1 و TargetP 1.1 تنها برای CCD4a وجود سیگنال با هدف‌گیری پلاستییدی را پیش‌بینی نمودند ولی نرم‌افزار ProtComp با امتیاز ۷,۲ برای هر دو پروتئین، جایگاه درون سلولی آن‌ها را کلروپلاست پیش‌بینی نمود. یک مارپیچ تراغشایی در موقعیت اسیدآمینه‌های ۲۰۶ - ۱۸۴ برای CCD4a و ۱۹۷ - ۱۷۵ برای CCD4b توسط نرم‌افزار TMHMM پیش‌بینی گردید. نتایج این نرم‌افزار نشان داد که اسیدآمینه‌های ۱۸۳ - ۱ و CCD4a و ۱۷۴ - ۱ و CCD4b به سمت داخل قرار داشته و در بخش خارج هم اسیدآمینه‌های ۵۷۹ - ۲۰۷ و ۵۷۰ - ۱۹۸ به ترتیب برای CCD4a و CCD4b قرار دارند (شکل-۲).

بررسی‌های هدف‌گیری با نرم‌افزارهای ChloroP 1.1 و TargetP 1.1 تنها برای CCD4a وجود سیگنال با هدف‌گیری پلاستییدی را پیش‌بینی نمودند ولی نرم‌افزار ProtComp با امتیاز ۷,۲ برای هر دو پروتئین، جایگاه درون سلولی آن‌ها را کلروپلاست پیش‌بینی نمود. یک مارپیچ تراغشایی در موقعیت اسیدآمینه‌های ۲۰۶ - ۱۸۴ برای CCD4a و ۱۹۷ - ۱۷۵ برای CCD4b توسط نرم‌افزار TMHMM پیش‌بینی گردید. نتایج این نرم‌افزار نشان داد که اسیدآمینه‌های ۱۸۳ - ۱ و CCD4a و ۱۷۴ - ۱ و CCD4b به سمت داخل قرار داشته و در بخش خارج هم اسیدآمینه‌های ۵۷۹ - ۲۰۷ و ۵۷۰ - ۱۹۸ به ترتیب برای CCD4a و CCD4b قرار دارند (شکل-۲).

جدول ۳- بررسی ساختار دو بعدی پروتئین CCD4a و CCD4b با نرم افزار SOPMA. پروتئین CCD4a (سمت چپ) و پروتئین CCD4b (سمت راست)

Table 3- Analysis of the secondary structure of the CCD4a protein (left side) and the CCD4b protein (right side) with the SOPMA software

نتیجه SOPMA برای CCD4a				نتیجه SOPMA برای CCD4b			
SOPMA result for CCD4a				SOPMA result for CCD4b			
مارپیچ آلفا	89	is	15.37%	مارپیچ آلفا	90	is	15.79%
Alpha helix				Alpha helix			
مارپیچ ۳۱۰	0	is	0%	مارپیچ ۳۱۰	0	is	0%
310 helix				310 helix			
مارپیچ پی	0	is	0%	مارپیچ پی	0	is	0%
Pi helix				Pi helix			
پل بتا	0	is	0%	پل بتا	0	is	0%
Beta bridge				Beta bridge			
رشته ممتد	151	is	26.08%	رشته ممتد	140	is	24.56%
Extended strand				Extended strand			
چرخش بتا	31	is	5.35%	چرخش بتا	38	is	6.67%
Beta turn				Beta turn			
ناحیه بتا	0	is	0%	ناحیه بتا	0	is	0%
Beta region				Beta region			
مارپیچ تصادفی	308	is	53.20%	مارپیچ تصادفی	302	is	52.98%
Random coil				Random coil			
ساختار مبهم	0	is	0%	ساختار مبهم	0	is	0%
Ambiguous states				Ambiguous states			
ساختارهای دیگر	0	is	0%	ساختارهای دیگر	0	is	0%
Other states				Other states			



شکل ۲- مقایسه ناحیه تراغشایی دو پروتئین با برنامه TMHMM. پروتئین CCD4a (بالا) و پروتئین CCD4b (پایین)
Figure 2- Investigation of transmembrane domains. CCD4a protein (up), CCD4b protein (down).

فیلوژنتیکی نشان داد که آنزیم‌های CCD1 و CCD2 در یک گروه مستقل قرار می‌گیرند و آنزیم‌های CsCCD4 با آنزیم‌های CCD4 سایر گیاهان در یک گروه جداگانه طبقه‌بندی می‌شوند. نکته جالب توجه اینکه در تمامی موارد یک گروه‌بندی واضح از گیاهان تک‌لپه‌ای و دولپه‌ای مشاهده گردید (شکل ۶). نتایج همچنین نشان داد که دو آنزیم CsCCD4a و CsCCD4b توالی‌های مونوفیلیتیک^۱ هستند زیرا دارای جد مشترکی می‌باشند. همچنین توالی‌های مربوط به CsCCD4a، CsCCD4b، CaCCD4a/b، CaCCD4c، TaCCD4c، TaCCD-D4، TaCCD-B4 و TaCCD-A-4 تشکیل یک clade را می‌دهند زیرا همگی آن‌ها از جد مشترکی به وجود آمده‌اند. درخت فیلوژنتیکی بدست آمده، پروتئین‌ها را به دو گروه اصلی تقسیم نمود که پروتئین‌های CsCCD4a و CsCCD4b در یک گروه در کنار یکدیگر قرار گرفته‌اند. بررسی‌های انجام شده نشان می‌دهد پروتئین‌های CCD4 و NCED4 قرار گرفته در گروه اول همگی متعلق به گیاهان رده دولپه‌ای‌ها می‌باشند و گیاهان متعلق به گروه دوم که شامل پروتئین‌های CsCCD4a و CsCCD4b می‌باشند به رده تک‌لپه‌ای‌ها تعلق دارند. همچنین پروتئین‌های CsCCD1، CsCCD2، و CsCCD2L نیز که متعلق به رده گیاهان تک‌لپه‌ای هستند در گروهی جدا از بقیه به عنوان جد مشترک پروتئین CCD4 در سایر گونه‌ها شناخته می‌شود. که نتایج حاصل مشابه گزارشات ارائه شده توسط برخی محققان در اسپانیا می‌باشد (Ahrazem et al., 2015). مقادیر Bootstrapping در محل انشعاب شاخه‌ها درخت به-خصوص در گروه دو که شامل توالی‌های مورد مطالعه در این تحقیق می‌باشد، نشان می‌دهد که درخت فیلوژنی حاضر از درجه اعتبار بالایی برخوردار است.

مدل‌سازی ساختار سه‌بعدی بر اساس انتخاب یک الگو با شباهت بالا با پروتئین هدف با استفاده از پایگاه اطلاعاتی Swiss model صورت گرفت (Brown et al., 2007) (شکل-۳). در انتخاب الگو معیارهایی نظیر وضوح در محدوده ۳ آنگستر، شباهت بالای ۳۵٪ الگو با توالی‌های مورد مطالعه و E-value پایین اعمال گردید تا اعتبار و اطمینان‌پذیری مدل افزایش یابد. الگوی منتخب برای مدل‌سازی پروتئین‌های CCD4a و CCD4b پروتئین VP14 در ذرت (SMTL ID: 3npe.1.A) دارای ۶۰۴ اسیدآمینو می‌باشد و در سال ۲۰۱۰ با قدرت تفکیک ۳٫۲ آنگستر از ذرت کریستالوگرافی شده است (Messing et al., 2010). میزان تطابق الگوی VP14 با پروتئین CCD4a ۴۰٫۸۳٪ و با پروتئین CCD4b ۴۱٫۴۲٪ و میزان شباهت آن برای هر دو پروتئین ۴۰٪ تعیین شد (جدول-۴). لذا با توجه به میزان شباهت پایین الگو با پروتئین‌های مورد نظر، این احتمال وجود دارد که مدل‌سازی به‌صورت ضعیف انجام شده باشد. اما این نکته که آنزیم CCD4 در مسیر سنتز کروسین در زعفران نقش مهم و حیاتی ایفا می‌کند، قطعاً اولین گام در راستای مطالعه جایگاه‌های اتصال و یا لیگاندهای فعال‌کننده یا مهارکننده، مدل‌سازی این پروتئین می‌باشد.

کیفیت پارامترهای ساختاری مدل‌های ساخته شده با پلات رامانچاندرا مورد ارزیابی قرار گرفت که این پلات با محاسبه درصد اسیدآمینوهای قرار گرفته در نواحی مطلوب، مجاز و یا غیرمجاز کیفیت مدل ورودی را تعیین می‌نماید (شکل ۴).

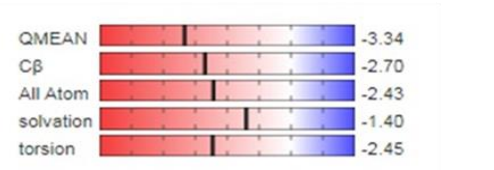
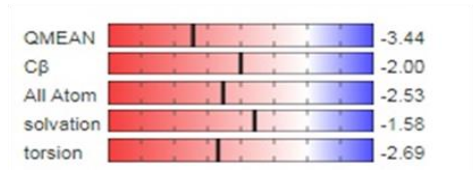
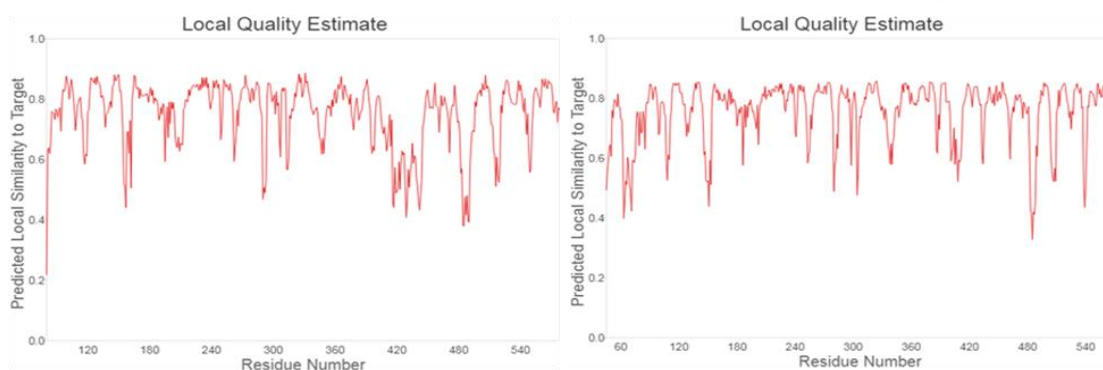
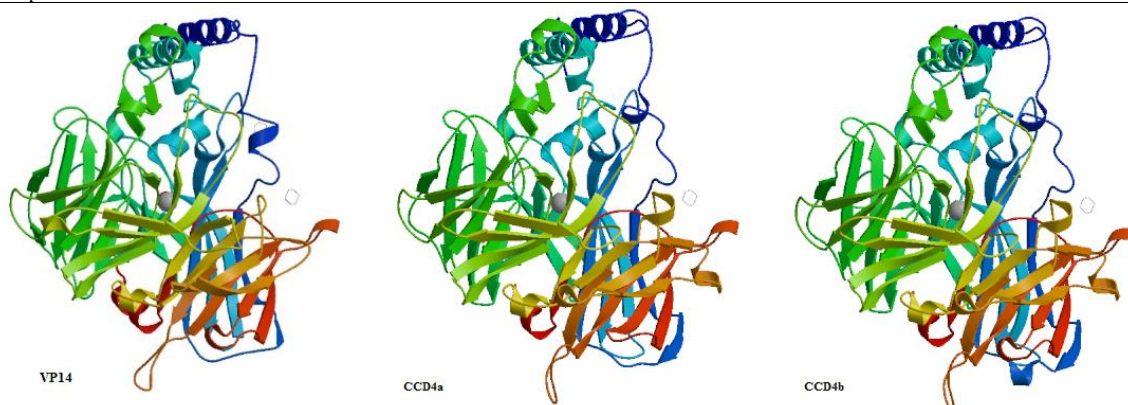
ارتباط فیلوژنتیکی ژن‌ها

نتایج حاصل از درخت فیلوژنتیکی ژن‌های CCD4a و CCD4b بر اساس هم‌ردیفی توالی‌های اسیدآمینوهای آن‌ها از طریق نرم افزار بر خط BLAST رسم شد، نتایج حاصله در درخت فیلوژنتیکی (شکل ۵) مشاهده می‌شود. بررسی

جدول ۴- میزان شباهت اسیدآمینه‌های پروتئین‌های CCD4a و CCD4b با الگو (پروتئین VP14 در ذرت)

Table 4- The similarity of the amino acids of CCD4a and CCD4b with the pattern (VP14 protein in maize)

شرح Description	درصد پوشش Coverage	محدوده Range	میزان شباهت Seq similarity	وضوح Resolution	روش Method	یافت شده توسط Found by	ساختار Oligo-state	میزان تطابق Seq identity	الگو Template	پروتئین Protein
9-cis-epoxycarotenoid dioxygenase1, chloroplast	83%	80- 578	0.42	3.2Å	X-ray	BLAST	monomer	45.23%	3npe.1.A	CCD4a
9-cis-epoxycarotenoid dioxygenase1, chloroplast	85%	71- 569	0.42	3.2Å	X-ray	BLAST	monomer	45.44%	3npe.1.A	CCD4b

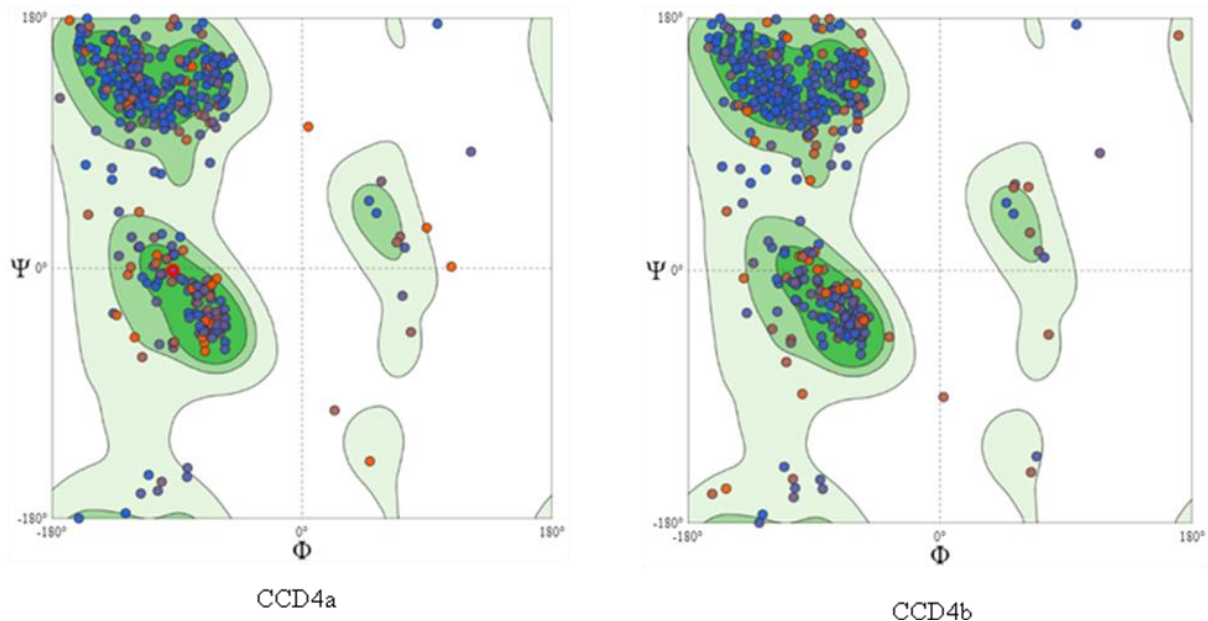


CCD4a

CCD4b

شکل ۳- مدل‌سازی ساختار سه‌بعدی بر اساس انتخاب یک الگو با شباهت بالا با پروتئین هدف با استفاده از پایگاه اطلاعاتی Swiss model ساختار روبان که نشان‌دهنده ساختارهای محافظت شده دوم و سوم (چپ) VP14 (چپ) است که به عنوان الگویی برای مدل‌سازی CsCCD4a (مرکز) و CsCCD4b (راست) مورد استفاده قرار گرفته است. b: نمودارهای ارزیابی کیفی موضعی ساختار و نمودار Z-score برای CCD4a و CCD4b

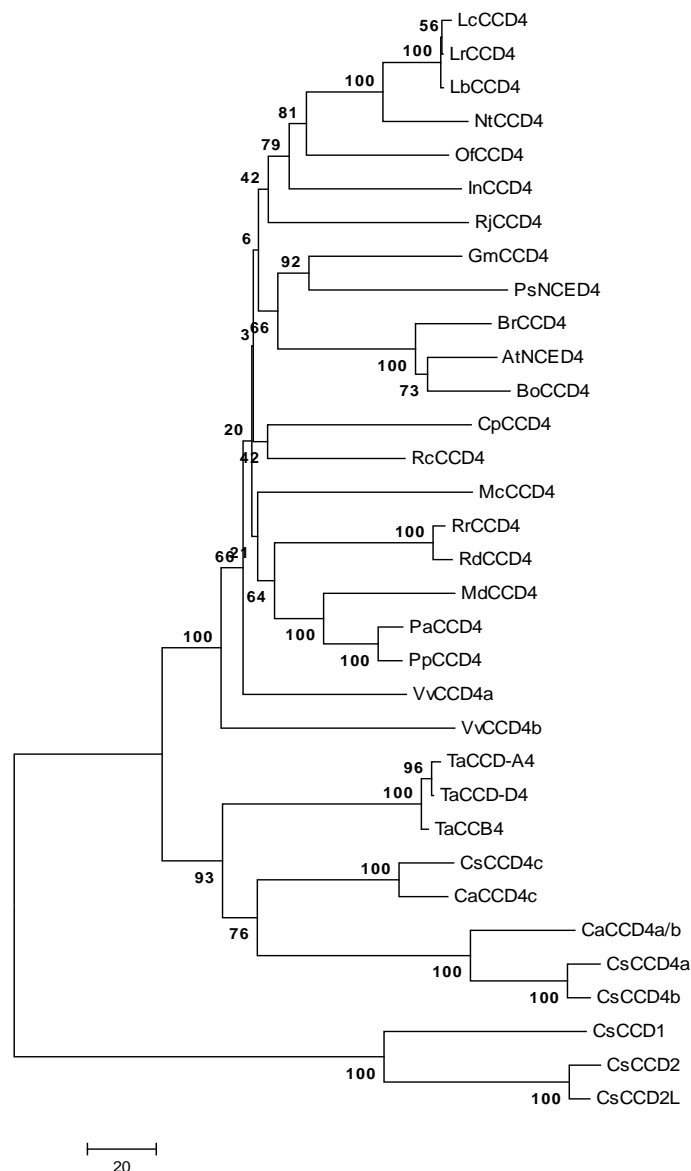
Figure3- Modeling the 3D structure based on the selection of a pattern with high similarity to the target protein using the Swiss model database. a: Ribbon diagrams showing the conserved secondary and tertiary of VP14 (left) used as template for modelling of CsCCD4a (center) and CsCCD4b (right). b: Global Model Quality Estimation and Z-score charts for the CCD4a and CCD4b.



شکل ۴- نمودار رامچاندرا ساختارهای مدل سازی شده CsCCD4a و CsCCD4b
Figure 4- Ramachandran plots for CsCCD4a and CsCCD4b constructed models.

موجود در پایگاه NCBI و دیگر گونه‌های گیاهی مقایسه شده- اند. همان‌طور که قبلاً اشاره شد نتایج وجود یک اینترون و با توالی مشخص را در این خانواده را نشان داد. توالی نوکلئوتیدی بدست آمده از ژن‌های *CCD4a* و *CCD4b* زعفران ایران در پایگاه داده NCBI به شماره دسترسی MK618659 و MK618660 به ثبت رسیدند. این دو ژن به میزان ۹۸٪ با هم تطابق داشته که با توجه به ماهیت تریپلوئیدی *C. sativus* به عنوان گونه‌های آللی مطرح می‌شوند. در تحقیقاتی که تاکنون به انجام رسیده است، در زعفران (*C. sativus*) دو ژن کدکننده *CsCCD4a* و *CsCCD4b* شناسایی و جداسازی شده‌اند. *CsCCD4a* (۵۸۰ اسیدآمین) و *CsCCD4b* (۵۶۹ اسیدآمین) با یکدیگر ۹۸٪ شباهت نشان دادند (Rubio et al., 2008). نتایج تحقیقات انجام شده نشان می‌دهد این دو آنزیم دارای فعالیت برشی در موقعیت‌های ۹-۱۰ (۹'-۱۰') می‌باشند، لذا پیشنهاد گردید که احتمالاً این دو ژن در تولید بتا-یونیون در طول رشد کلاله دخالت دارند (Rubio et al., 2008).

بدون شک، در گسترش کاربرد بیوتکنولوژی در گیاهان، توسعه مهندسی ژنتیک یکی از روش‌های مورد توجه می‌باشد که ابزاری مفید و سریع به منظور رسیدن به مطالبات مورد انتظار در اصلاح گونه‌ها گیاهی فراهم می‌کند (Bilas et al., 2016). تاکنون پیشرفت‌های مهمی در رابطه با ژن‌های کدکننده کاروتنوئیدها در گیاهان مختلف مانند سیب (*Malus domestica*), *MdCCD4* chrysanthemum, *CmCCD4a* (*Chrysanthemum morifolium*), رز (*Rose*) *RdCCD4* (*Osmanthus damascene*), *Osmanthus* (*Osmanthus fragrans*), *OfCCD4* و آرابیدوپسیس (*AtCCD4*) صورت گرفته است (Huang et al., 2009; Gomez-Gomez et al., 2010). با این حال در رابطه با ویژگی‌های ساختاری پروتئین-های کدشونده به وسیله‌ی این ژن‌ها در زعفران اطلاعات کمی در دسترس می‌باشد. در این مطالعه، به منظور شناخت بیشتر از ساختار ژنومی زعفران ایران، ژن‌های *CsCCD4a* و *CsCCD4b* جداسازی، توالی یابی و مورد بررسی قرار گرفتند. نتایج حاصله با داده‌های



شکل ۵- درخت فیلوژنتیکی توالی اسیدآمینهای CsCC4b و CsCCD4a با سایر CCD در *Crocus sativus* و CCD4 در سایر گیاهان

Figure 5. Phylogenetic tree of the sequence of CsCCD4a and CsCC4b with other CCD in *Crocus sativus* and CCD4 in other plants.

Accession number used are: *Arabidopsis thaliana*, AtNCED4 (NP_193652.1). *Brassica rapa*, BrCCD4 (XP_009132578.1). *Brassica oleracea* var. *alboglabra*, BoCCD4 (AJP16417.1). *Crocus ancyrensis*, CaCCD4a/b (AKN09910.1). CaCCD4c (AKN09911.1). *Crocus sativus*, CsCCD1 (.....). CsCCD2 (ACD62475.1). CsCCD2L (ALM23547.1). CsCCD4c (AEO50759.1). *Carica papaya*, CpCCD4 (.....). *Glycine max*, GmCCD4 (XP_003516508.1). *Ipomoea nil*, InCCD4 (FAA01245.1). *Lycium barbarum*, LbCCD4 (AIX87510.1). *Lycium chinense*, LcCCD4 (AIY62809.1). *Lycium ruthenicum*, LrCCD4 (AIX87534.1). *Momordica charantia*, McCCD4 (AFU91490.1). *Malus domestica*, MdCCD4 (ABY47995.1). *Manihot esculenta*, MeCCD4 (XP_021595193.1). *Nicotiana tabacum*, NtCCD4 (NP_001313051.1). *Osmanthus fragrans*, OfCCD4 (ABY60887.1). *Prunus avium*, PaCCD4 (XP_021809777.1). *Prunus persica*, PpCCD4 (AGL08676.1). *Pisum sativum*, PsNCED4 (BAC10552.1). *Ricinus communis*, RcCCD4 (XP_002519944.1). *Rosa x damascene*, RdCCD4 (ABY60886.1). *Rhododendron japonicum* f. *flavum*, RjCCD4 (BAT32880.1). *Rosa rugosa*, RrCCD4 (AKT74335.1). *Triticum aestivum*, TaCCD-A4 (ANT73645.1). TaCCD-B4 (ANT73646.1). TaCCD-D4 (ANT73647.1). *Vitis vinifera*, VvCCD4a (AFJ94675.1). VvCCD4b (AFJ94676.1).

از نتایج پژوهش‌های صورت گرفته بر روی ساختار ژن‌های *CCD4* می‌توان چنین نتیجه‌گیری نمود که خانواده *CCD4* بسیار پویا بوده و با وجود تغییرات نسبی مشاهده شده در توالی *C. sativus* از سری ژن‌های خانواده *CCD* در جنس *Crocus* می‌باشند (Ahrazem et al., 2010). نتایج بررسی‌های مولکولی در این مطالعه نشان داد که مقدار دامنه وزن مولکولی، نقطه ایزوالکتریک، شاخص آلیفاتیک و خاصیت آبگریزی و حالیت در هر دو پروتئین مورد مطالعه نزدیک به هم می‌باشد و نتایج حاصل نشان دهنده شباهت ساختار پروتئینی آنزیم‌های مورد مطالعه می‌باشد.

اطلاع از نقطه ایزوالکتریک بدین منظور است که یک پروتئین می‌تواند برای انتخاب و بهینه‌سازی روش‌های مورد استفاده برای خالص‌سازی پروتئین اعم از کروماتوگرافی تبادل یونی (IEC^۳) و الکتروفورز ایزوالکتریک فوکوسینگ (IEF^۴) بسیار مفید باشد. به علاوه پروتئین‌هایی که در محدوده pH ایزوالکتریک قرار داشته و مجموع بار مثبت و منفی رشته پپتیدی آن‌ها با یکدیگر برابر می‌باشد، توانایی نگهداری آب را نداشته و ناگزیر رسوب می‌کنند. شاخص ناپایداری پروتئین‌ها، تخمینی از ثبات پروتئین در شرایط لوله آزمایش می‌باشد (Mirhosseini et al., 2016) و این شاخص به نوع اسیدآمین‌های موجود بستگی دارد. شاخص ناپایداری محاسبه شده برای پروتئین‌های مورد بررسی نشان می‌دهند که پروتئین‌هایی که شاخص ناپایداری آن‌ها کمتر از ۴۰ باشند با ثبات و در غیر این صورت بی‌ثبات خواهند بود. پروتئین‌های پایدار در شرایط تنش مثل تنش حرارتی می‌توانند ساختار سوم خود و در نتیجه عملکردشان را در شرایط تنش حفظ کنند (Pettersen et al., 2004). شاخص ناپایداری برای پروتئین‌های *CCD4a* و *CCD4b* کمی

محققان اسپانیایی در مطالعه ژن‌های *CCD4* در *C. sativus* برای ژن *CsCCD4a* استخراج شده از DNA ژنومی سه توالی متفاوت را گزارش کردند یک کپی حاوی یک اینترون، یک کپی فاقد اینترون و در نهایت یک کپی شامل توالی فاقد اینترون کوتاه شده است (Ahrazem et al., 2010). در حالی که در این تحقیق، تنها توالی *CsCCD4a* حاوی یک اینترون در DNA ژنومی بدست آمد. که دلیل این تغییرات بر اساس پژوهش‌های صورت گرفته تحرک بسیار بالای ساختار اینترون-اگزون در ژن‌های *CCD4* بیان شده است (Ahrazem et al., 2010).

از طرفی در پژوهشی که محققان اسپانیایی بر روی تنوع موجود در توده‌های جمع‌آوری شده زعفران به دو روش AFLP^۱ و MS-AFLP^۲ انجام دادند نتایج نشان داد که علی‌رغم تنوع ژنتیکی کم ولی تنوع اپی‌ژنتیکی در بین نمونه‌های زعفران بالا بود به طوری که نمونه‌ها به دو گروه اصلی شامل نمونه‌های اسپانیایی و نمونه‌های سایر کشورها شامل ایران، هند، ترکیه و افغانستان طبقه بندی گردید. در این تحقیق نمونه‌های ایران شامل ۹ نمونه در ۵ ژنوتیپ مختلف قرار گرفتند که با هیچ گروه دیگری هم‌پوشانی نداشته و بیشترین تعداد اختلاف ژنتیکی در بین آن‌ها مشاهده گردید. در این پژوهش امکان تأثیر دو عامل منشأ جغرافیایی و از بین رفتن سطوح زیرکشت زعفران در شرایط محیطی مشابه، به عنوان عوامل مؤثر بر وضعیت اپی‌ژنتیکی کنونی نمونه‌های زعفران معرفی گردیده است (Busconi et al., 2015). در این مطالعه نیز وجود اختلاف در توالی ژن‌های *CCD4* زعفران ایران نسبت به زعفران اسپانیا نیز تأییدی بر اختلاف ژنتیکی بین آن‌ها می‌تواند باشد.

۱- Amplified Fragment Length Polymorphism

۲-Methyl Sensitive Amplified Fragment Length Polymorphism

۳ - Ion Exchange Chromatography

۴ - Isoelectric Focusing Electrophoresis

(Rubio et al., 2008). همچنین آبریز بودن پیش ماده آزریم‌های CCD بیانگر این است که این آزریم‌ها باید به غشاهای سلولی متصل شوند تا بدین طریق به سوبستراهای خود دسترسی پیدا کنند (Wang et al., 2013). دیگر مطالعات صورت گرفته در این خصوص نشان‌دهنده این است که CsCCD4a و CaCCD4b حاوی بخش آبریز هستند که این نقاط آبریز امکان عبور این آزریم‌ها را از غشا فراهم می‌کنند (Ma et al., 2013). انجام مطالعات با TMHMM نشان داد که CsCCD4a و CsCCD4b هر کدام دارای یک مارپیچ می‌باشند که در پایانه N قرار دارند و با توجه به نتایج بدست آمده پروتئینی غشایی بوده که حاوی یک سیگنال پپتیدی در پایانه N خود می‌باشد که با نتایج حاصل از نرم‌افزارهای تعیین سیگنال پپتیدی مطابقت دارد.

وجود سیگنال پپتیدی در پایانه N با هدف‌گیری کلروپلاستی بوده که توسط بررسی‌های مولکولی تأیید گردید. مطالعات جایگاه درون سلولی نشان داد که محل این آزریم‌ها کلروپلاست می‌باشد. با توجه به این که کلاله فاقد کلروپلاست است، به نظر می‌رسد به احتمال زیاد کروموپلاست و به‌طور اختصاصی تر پلاستوگلوبول‌ها جایگاه این آزریم‌ها هستند (Rubio et al., 2008). آزریم‌های CCD4 در پلاستیدها واقع شده‌اند و به‌طور اختصاصی در پلاستوگلوبول‌ها (Ma et al., 2008; Rubio et al., 2013) جایی که بیشتر آزریم‌های کاروتنوژنیک قرار دارند (Nogueira et al., 2013). با توجه به نتایج حاصل در این مطالعه در خصوص پروتئین‌های CsCCD4a و CsCCD4b پیشنهاد می‌گردد که پایانه N این پروتئین‌ها می‌تواند به عنوان یک سیگنال پپتیدی عمل نموده که پس از ساخته شدن پلی‌پپتید آن را به سمت جایگاه پیش‌بینی شده (پلاستوگلوبول) هدایت می‌نماید. پس از عبور از غشای پلاست و گلوبول به‌واسطه مارپیچ تراغشایی و با توجه به الگوی آبدوستی-آبریزی، پروتئین‌ها از طریق این بخش به غشای کلروپلاستی (ب‌طور

بیشتر از ۴۰ می‌باشد که جزء پروتئین‌های ناپایدار تحت شرایط آزمایشگاهی طبقه‌بندی می‌شوند.

شاخص آلیفاتیک عامل مهم دیگری به منظور برآورد مقدار آبریزی پروتئین و یک عامل مثبت برای پایداری حیاتی پروتئین و ساختار سه‌بعدی آن محسوب می‌شود که در واقع عبارت است از حجم نسبی پروتئین که با زنجیره‌های آلیفاتیک اشغال شده‌اند (V, A, I, L,) (Fallah Ziarani et al., 2017). پروتئین‌هایی با شاخص آلیفاتیک بسیار بالا (بیشتر از ۱۰۰) ممکن است در دامنه دمایی بسیار بالا از خود ثبات نشان دهند. محاسبه شاخص آلیفاتیک به عنوان یک فاکتور مثبت در مقاومت به حرارت نشان داد که هر دو پروتئین CCD4a و CCD4b به طور مشابهی از ثبات بالایی در برابر دماهای بالا برخوردار می‌باشند.

غشاهای لیپیدی محیطی مناسب برای آزریم‌هایی هستند که سوبستراهای آبریزی همانند کاروتنوئیدها را متابولیزه می‌کنند (Ahrazem et al., 2015). در *in vivo*، کاروتنوئیدها می‌توانند در بین غشا دولایه لیپیدی یا پلاستییدی ذخیره و انباشته شوند (Cunningham et al., 1998). از آنجا که آزریم‌های CCD از کاروتنوئیدها به عنوان پیش ماده استفاده می‌کنند، تعیین مشخصات آبریزی-آبدوستی این آزریم‌ها و خاصیت تراغشایی آن‌ها مهم می‌باشد. نتایج حاصل از نرم‌افزار ProtScale نشان داد که CsCCD4a و CCD4b حاوی بخش‌های آبریز (شکل ۱b، دمین‌های آبریز در بالا و دمین‌های آبدوست در زیر خط صفر قرار دارند) هستند. متوسط خاصیت آبریزی (GRAVY^۱) کل محاسبه شده برای CCD4a و CCD4b منفی می‌باشد و این بدین معنی است که پروتئین‌های مورد مطالعه غیرقطبی هستند. وجود یک دمین آبریز در CsCCD4a و CsCCD4b امکان برقراری ارتباط با لیپیدهای غیرقطبی را فراهم می‌سازد

۱ - Grand average of hydropathicity

ارزیابی مدل پیشنهاد شده در Swiss Model بر اساس پارامترهای^۱ GMQE و QMEAN4 می‌باشد. تابع QMEAN4 یک تابع حسابی مرکب برای ارزیابی کیفیت global (کل ساختار) و local (هر اسیدآمین) ساختار مدل‌سازی شده می‌باشد. نمودار کیفیت local برای هر اسیدآمین در مدل (محور X) شباهت مورد انتظار به ساختار بومی^۲ (محور Y) را نشان می‌دهد. به طور معمول، اسیدآمین‌هایی که امتیاز زیر صفر را نشان بدهند، از کیفیت پایینی برخوردارند. تابع QMEAN4 شامل چهار توصیف ساختاری پتانسیل زاویه چرخش، پتانسیل فاصله اتم‌ها، اثر متقابل کربن بتا و پتانسیل حلال پوشی می‌باشد (شکل ۴b). GMQE یک ارزیابی کیفی است که خصوصیات هم‌ترازی مدل - الگو را ترکیب می‌کند. نمودارهای ارزیابی کیفی موضعی ساختار و میزان Z-score در مورد تابع حسابی مرکب QMEAN4 تمامی اتم‌ها، کربن بتا، حلالیت و زوایای پرخشی در شکل ۴b نشان داده شده است. ارزیابی کیفیت ساختار مدل شده پس از بهینه سازی انرژی به وسیله پلات رامچاندرا در صورت گرفت پلات رامچاندرا یک روش بسیار پرکاربرد جهت پلات زوایای پرخشی پروتئین و ارزیابی کیفیت استرئوشیمیایی ساختار مدل‌سازی شده می‌باشد. این پلات با محاسبه درصد اسیدآمین‌های قرار گرفته در نواحی مطلوب، مجاز و یا غیرمجاز کیفیت مدل ورودی را تعیین می‌نماید (Fallah Ziarani et al., 2017). همان‌طور که در شکل ۵ مشاهده می‌شود، ۹۱،۱۵٪ اسیدآمین‌ها در CsCCD4a و ۸۹،۶۷٪ اسیدآمین‌ها در CsCCD4b برای ساختارهای مدل شده در مناطق مطلوب و ۱،۳۴٪ و ۲،۴۱٪ اسیدآمین‌ها به ترتیب برای CsCCD4a و CsCCD4b در مناطق مجاز دارند. با جمع نواحی مطلوب و مجاز ۹۲،۴۹٪ برای CCD4a و ۹۲،۰۸٪ برای CCD4b بدست می‌آید و این بدان معنی است که هر دو مدل از کیفیت

اختصاصی تر پلاستوگلوبول) متصل می‌شوند.

ساختار سه بعدی یک پروتئین که بر اساس مدل‌سازی همولوژیکی آن رسم می‌گردد، به میزان زیادی خواص عملکردی آن را مشخص می‌نماید. شکل فضایی بخش‌های متفاوت هر پروتئین، در واقع تعیین کننده کاربرد و عملکرد ویژه همان بخش از پروتئین بوده و هر نوع تغییری در این ساختار می‌تواند باعث اخلال در عملکرد آن شود. اساس پیش‌بینی توسط مدل‌سازی همولوژیکی این است که توالی پروتئین با یک یا چند پروتئین با ساختار شناخته شده شباهت داشته باشد. این روش، مدل‌سازی مقایسه‌ای (CA) نیز نامیده می‌شود (Yousefi, 2017). در روش مدل‌سازی مقایسه‌ای، ساختمان پروتئین بر اساس مشابهت توالی ساختمان با توالی یا ساختمان‌های شناخته شده با روش‌های تجربی پیش‌بینی می‌شود. در واقع این روش بر این اصل استوار است که اگر دو پروتئین از نظر توالی تشابه بالایی داشته باشند، احتمالاً ساختمان سه بعدی بسیار مشابهی خواهند داشت. ساختار سه بعدی پروتئین منبع مهم اطلاعاتی برای درک بهتر عملکرد پروتئین و واکنش آن با اجزای دیگر (لیگاند، پروتئین و غیره) می‌باشد. به علت سختی و پرهزینه بودن فرآیند کریستالوگرافی به خصوص در مورد پروتئین‌های مهم، پیشگویی ساختار آن‌ها از طریق ابزار *In silico* به منزله میانبری برای مطالعات بعدی و بررسی جزئیات ساختاری آنها خواهد بود. به عبارت دیگر طراحی یک لیگاند مؤثر برای فعال سازی یا مهار پروتئین در یک مسیر خاص، با شبیه‌سازی ساختار سوم آن پروتئین مقرون به صرفه بوده و روند مطالعات را تسریع می‌کند. با توجه به نتایج ارزیابی کیفی مدل‌های ایجاد شده، می‌توان با ضریب اطمینان بالایی از این مدل‌ها در آنالیزها و طراحی‌های بعدی استفاده نمود. ساختار سه‌بعدی روبان پروتئین‌های CCD4a و CCD4b با نرم‌افزار SWISS-MODEL به روش Homology modeling ساخته

شد (شکل ۴a).

۱ - Global Model Quality Estimation

۲ - Native structure

آنزیم CCD4 در مسیر سنتز آپوکاروتنوئیدهای زعفران بسیار مفید باشد و به فهم چگونگی نحوه عمل این آنزیم‌ها با سایر کاروتنوئیدها کمک کرده و زمینه‌ساز روشن‌تر شدن نحوه فعالیت آن در گیاه زعفران باشد. همچنین وجود تفاوت بین نتایج حاصل از این تحقیق با تحقیقات انجام شده در سایر کشورها در خصوص تعداد اینترون در این ژن‌ها و با توجه به این که اینترون‌ها به ویژه اینترون اول، نقش مهمی در جای‌گیری صحیح سیتوپلاسمی برخی mRNA ها و همچنین در تنظیم رونویسی و ترجمه دارند و از طرفی اینترون‌ها به عنوان یکی از عوامل مهم در ایجاد تنوع اپی‌ژنتیکی شناخته می‌شوند، مطالعه بیشتر در خصوص تنوع اپی‌ژنتیکی و منشأ جغرافیایی زعفران ایران یک امر استراتژیک و ملی در راستای ثبت هویت ایرانی این محصول در جهان خواهد بود.

سپاسگزاری

از ستاد توسعه زیست فناوری به دلیل حمایت مالی از این پروژه تحقیقاتی، گروه زیست فناوری مواد غذایی مؤسسه پژوهشی علوم و صنایع غذایی به دلیل در اختیار قراردادن امکانات آزمایشگاهی و تحقیقاتی و کلیه بزرگوارانی که در انجام این پژوهش همکاری نموده‌اند، صمیمانه تشکر و قدردانی می‌گردد.

مقبولی برخوردارند. با توجه به نتایج ارزیابی کیفی مدل‌های ایجاد شده می‌توان از این مدل‌ها در بررسی‌ها و طراحی‌های بعدی استفاده نمود.

نتیجه‌گیری

نتایج تحقیق حاضر نشان داد که دو ژن *CCD4a* و *CCD4b* دو فرم از خانواده *CCD4*، با درصد تشابه می‌باشند که اختلاف آنها به دلیل حضور تک نوکلئوتیدهای متفاوت در بین دو توالی این دو ژن است. این تک نوکلئوتیدها می‌توانند به عنوان نشانگرهای تک نوکلئوتیدی (SNP's) معرف حضور این ژن‌ها در تحقیقات مورد استفاده قرار گیرند.

مطالعه ژن‌های مسیر تولید آپوکاروتنوئیدها در زعفران و نحوه تولید آن‌ها در این گیاه ضروری به نظر می‌رسد و از آنجا که اهمیت درمانی، اقتصادی و صادراتی زعفران تنها به چرخه تولید گلدهی و به‌طور خاص به کلاله آن خلاصه می‌شود، نیاز است تا با استفاده از روش‌های مولکولی و بخصوص بیوتکنولوژی، ژن‌ها و پروتئین‌های حاصل از رونوشت و ترجمه مسیر تولید آپوکاروتنوئیدها و به‌طور خاص کروسین در این گیاه راهبردی مورد توجه و مطالعه بیشتری واقع شود. زیرا با مطالعه و تحقیق بر روی ژن‌ها و آنزیم‌های زعفران می‌توان گام مؤثری در افزایش و بهبود تولید متابولیت‌های ثانویه در زعفران ایران برداشت. نتایج این تحقیق می‌تواند در مطالعه رفتار و کنش

منابع

- Ahrazem, O., Trapero, A., Gomez, M.D., Rubio-Moraga, A., and Gomez-Gomez, L. 2010. Genomic analysis and gene structure of the plant carotenoid dioxygenase 4 family: A deeper study in *Crocus sativus* and its allies. *Genomics* 96: 239-250.
- Ahrazem, O., Rubio-Moraga, A., Berman, J., Capell, T., Christou, P., Changfu, Z., and Gomez-Gomez, L. 2015. The carotenoid cleavage dioxygenase CCD2 catalysing the synthesis of crocetin in spring crocuses and saffron is a plastidial enzyme. *New Phytologist* 209 (2): 650-663.
- Ahrazem, O., Rubio-Moraga, A., Argandola-Picazo, J., Castillo, R., and Gomez-Gomez, L. 2016. Intron retention and rhythmic diel

- pattern regulation of carotenoid cleavage dioxygenase 2 during crocetin biosynthesis in saffron. *Plant Molecular Biology* 91 (3): 355-374.
- Almagro Armenteros, J.J., Salvatore, M., Emanuelsson, O., Winther, O., von Heijne, Arne Elofsson, G., and Nielsen, H. 2019. Detecting sequence signals in targeting peptides using deep learning. *Life Science Alliance* 2 (5): 1-14.
- Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., de Castro, E., Duvaud, S., Flegel, V., Fortier, A., Gasteiger, E., Grosdidier, A., Hernandez, C., Ioannindis, V., Kuznetsov, D., Liechti, R., Moretti, S., Mostaguir, K., Redaschi, N., Rossier, G., Xenarios, I., and Stockinger, H. 2012. ExPASy: SIB bioinformatics resource portal, *Nucleic Acids Research* 40 (W1): W597-W603.
- Baba, Sh.A, Mohiuddin, T., Basu, S., Swarnkar, M.K., Malik, A.H., Wani1, Z.A., Abbas, N., Singh, A.K., and Ashraf, N. 2015. Comprehensive transcriptome analysis of *Crocus sativus* for discovery and expression of genes involved in apocarotenoid biosynthesis. *BMC Genomics* 16 (698): 1-14.
- Bathaie, S.Z., Ashrafi, M., Bolhasani, A., Etemadikia, B., and Moosavi-movahedi, A.A. 2006 Purification of carotenoids and monoterpen aldehydes from Iranian saffron and investigation of their effect on the structure of DNA, histone H1 and H1-DNA complex. *Iranian Journal of Medicinal and Aromatic Plants* 22 (2): 85-97. (In Persian with English Summary).
- Beiki, A.H., Keify, F., and Mozafari, J. 2011. Rapid genomic DNA isolation from corm of *Crocus* species for genetic diversity analysis. *Journal of Medicinal Plants Research* 5 (18): 4596-4600.
- Bendtsen, J.D., Nielsen, H., Von Heijne, G., and Brunak, S. 2004. Improved prediction of signal 3.0. *Journal Molecular Biology* 340 (4): 783-795.
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., and Schwede, T. 2017. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Scientific Reports* 7 (1): 10480.
- Bhat, A., Mishra, S., Kaul, S., and Dhar, M.K. 2018. Elucidation and functional characterization of *CsPSY* and *CsUGT* promoters in *Crocus sativus* L.. *Plos One* 13 (4): 1-15.
- Bienert, S., Waterhouse, A., de Beer, T.A.P., Tauriello, G., Studer, G., Bordoli, L., and Schwede, T. 2017. The SWISS-MODEL Repository - new features and functionality. *Nucleic Acids Research* 45: D313-D319.
- Bilas, R., Szafran, K., Hnatuszko-Konka, K., and Kononowicz, A.K. 2016. Cis-regulatory elements used to control gene expression in plants. *Plant cell, Tissue and Organ Culture (PCTOC)* 127 (2): 269-286.
- Brown, AK., Sridharam, S., Kremer, L., Lindenberg, S., Dover, LG., Sacchettini, JC., and Besra, GS., 2005. Probing the mechanism of the Mycobacterium tuberculosis beta-ketoacyl-acyl carrier protein synthase III mtFabH: factors influencing catalysis and substrate specificity. *The Journal of Biological Chemistry* 280 (37): 32539-32547.
- Brown, T.A., 2007. *Genome 3*, 3rd ed, c2006. Khosravi Publication, Tehran. (In Persian).
- Busconi, M., Colli, L., Sanchez, R. A., Santaella, M., Pascual, M., Santana, O., Roldan, M., Fernandez, A. 2015. AFLP and MS-AFLP analysis of the variation within Saffron *Crocus (Crocus sativus* L.) germplasm. *PLoS ONE* 10 (4): e0123434.
- Combet, C., Blanchet, C., Geourjon, C., and Deleage, G. 2000. NPS@: Network protein sequence analysis. *Trends in Biochemical Science* 25 (3): 147-150.
- Cunningham, F.X., and Gantt, E. 1998. Genes and enzymes of carotenoid biosynthesis in plants. *Annual Review of Plant Physiology and Molecular Biology* 49: 557-583.
- El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, J.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E., and Finn, R.D. 2019. The Pfam protein families database in 2019. *Nucleic Acids Research* 48 (D1): D427-D432.
- Emanuelsson, O., Nielsen, H., and von Heijne, G. 1999. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Science* 8 (5): 978-984.
- Gasteiger, E., Hoogland, Ch., Gattiker, A.,

- Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. 2005. Protein Identification and Analysis Tools on the ExPASy Server. The Proteomics Protocols Handbook. Humana Press, New York. pp. 571-607.
- Fallah Ziarani, M., Tohidfar, M., and Aminfar, Z. 2017. Bioinformatic analysis of Acyl Carrier Protein (ACP) in eukaryotes and prokaryotes. *Crop Biotech* 17: 15-29 (In Persian with English Summary).
- Giuliano, G., Al-Babili, S., and Lintig, J. 2003. Carotenoid oxygenases: cleave it or leave it. *Plant Science* 8 (4): 145-148.
- Goli, S.A.H., Mokhtari, F., and Rahimmalek, M. 2012. Phenolic Compounds and antioxidant activity from saffron (*Crocus sativus* L.) petal. *Journal of Agricultural Science* 4 (10): 175-181.
- Golovnina, K.A., Glushkov., S.A., Blinov., A.G., Adkison, L.R., and Goncharov, N.P. 2007. Molecular phylogeny of the genus *Triticum*. *Plant Systematics and Evolution* 264 (3-4): 195-216.
- Gomez-Gomez, L., Rubio-Moraga, A., and Ahrazem, O. 2010. Understanding carotenoid metabolism in saffron stigmas: unravelling aroma and colour formation. *Functional Plant Science and Biotechnology* 4 (2): 56-63.
- Grilli, M.C., and Canini, A. 2004. Ultrastructure of chromoplasts and other plastids in *Crocus sativus* L. (Iridaceae). *Plant Biosystems* 138 (1): 43-52.
- Guex, N., Peitsch, M.C., and Schwede, T. 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30 (S1): S162-S173.
- Hosseinpour azad, N., Nematzadeh, G.H., Gouliano, G., Ranjbar, G.A., and Yamch, A. 2016. Identification of Apo-Carotenoids' crocin and crocetin isomers in saffron crude extracts by HPLC coupled to atmospheric pressure chemical ionization and high resolution orbitrap mass spectrometry. *Saffron Agronomy and Technology* 4 (4): 291-299. (In Persian with English Summary).
- Huang, FC., Horváth, G., Molnár, P., Turcsi, E., Deli, J., Schrader, J., Sandmann, G., Schmidt, H., and Schwab, W. 2009. Substrate promiscuity of RdCCD1, a carotenoid cleavage oxygenase from *Rosa damascene*. *Phytochemistry* 70 (4): 457-464.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305 (3): 567-580.
- Kyte, J., and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular biology* 157 (1): 105-132.
- Ma, G., Zhang, L., Matura, A., Matura, K., Yamawaki, K., Yahata, M., Wahyudi, A., Motohashi, R., and Kato, M. 2013. Enzymatic formation of beta-citraurin from beta-cytophanthol and zeaxanthin by carotenoid cleavage dioxygenase4 in the flavedo of citrus fruit. *Plant Physiology* 163: 682-695.
- Maniatis, T., Fritsch, E., and Sambrook, F. 1995. *Molecular Cloning. A laboratory Manual*. Cold Spring Harbor Laboratory, New York .
- Messing, S.A., Gabelli, S.B., Echeverria, I., Vogel, T.J., Guan, J.C., Cai Tan, B., Klee, H.J., McCarty, R.D., and Amzel, L.M. 2010. Structural insights into maize viviparous14, a key enzyme in the biosynthesis of the phytohormone abscisic acid. *Plant Cell* 22 (9): 2970-2980.
- Mirhoseini, S.Z., Pezeshkian, Z., and Ghovvati, Sh. 2016. Phylogenetic and *In Silico* Analysis of Interferon Beta-1b Protein. *Journal Mazandaran University Medical Science* 26 (145): 70-82. (In Persian).
- Naghavi, M R., Malboobi, M.A., and Rashidi, S. 2009. *Bioinformatics*. University of Tehran Press. (In Persian).
- Nassaj Hoseini, S.M., and Shamsbakhsh, M. 2010. *Phylogenetic Analysis Methods*. Haghshenass Publication, Rasht. (In Persian).
- Nogueira, M., Mora, L., Enfissi, E.M., Bramley, P.M., and Fraser, P.D. 2013. Subchromoplast sequestration of carotenoids affects regulatory mechanisms in tomato lines expressing different carotenoid gene combinations. *Plant Cell* 25: 4560-4579.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. 2004. UCSF Chimera-visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25 (13): 1605-1612.
- Petersen, T.N., Brunak, S., Von Heijne, G., and

- Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane region. *Nat Metthod* 8 (10): 785-786.
- Rubio, A., Rambla, J.L., Santaella, M., Gomez, D., Orzaez, D., Granell, A., and Gomez-Gpmez, L. 2008. Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in β -Ionone release. *The Journal of Biological Chemistry* 283 (36): 24816–24825.
- Rubio, A., Rambla, J.L., Ahrazem, O., Granell, A., and Gomez-Gomez, L. 2009. Metabolic and target transcript analysis during *Crocus sativus* stigma development. *Phytochemistry* 70 (8): 1009-1016.
- Schwartz, S.H., Tan, B.C., Gage, D., Zeevaart, J.A.D., and McCarty, D.R. 1997. Specific oxidative cleavage of carotenoids by VP14 of Maize. *Science* 276 (5320): 1872- 1874.
- Softberry Protcomp 9.0. 2016. Available at web site <http://www.softberry.com/berry.phtml?topic=protcomppl&group=programs&subgroup=proloc>
- Messing, S.A., Gabelli, S.B., Echeverria, I., Vogel, J.T., Guan, J.C., Tan, B.C., Klee, H.J., McCarty, D.R., and Amzel, L.M. 2010. Structural insights into maize viviparous14, a key enzyme in the biosynthesis of the phytohormone abscisic acid. *Plant Cell* 22 (9): 2970-2980.
- Sivakumar, K. 2005. WWW.Databases tools and prediction server for protein sequence analysis and characterization. *Advance Biotechnology* 27-31.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: Moist lecular evolutionary genetics analyze version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.
- Tarmontano, A., Leplae, R., and Morea, V. 2001. Analysis and assessment of comparative modeling predictions in proteins. *Proteins* 45 (S5): 22-38.
- Varjonen, E., Vainio, E., Kalimo, K., and Juntunen-Backman, K. 2002. Clinical importance of non-specific lipid transfer proteins as food allergens. *Biochemical Society Transactions* 30 (6): 910-913.
- Wang, Y., You, F.M., Lazo, G.R., Lou, M-C., Thilmony, R., Gordon, S., Kianian, Sh.F., and Gu, Y.Q. 2013. PIECE: a database for plant gene structure comparison and evolution. *Nucleic Acids Research* 41:1159–1166.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* 46 (W1): W296-W303.
- Yousefi Javan, I., and Gharari, F. 2017. The structure of the protein and gene expression of PIC2 affecting blooming flowers (*Crocus sativus* L.). *Saffron Agronomy and Technology* 5 (1): 73-90.

Cloning and Bioinformatics Investigation on CCD4a and CCD4b Genes from Iranian Saffron (*Crocus sativus* L.)

Mohammad Javad Habibzadeh¹, Ebrahim Dorani-Uliaie^{2*}, Seyed Mahdi Ziaratnia³ and Mostafa Valizadeh⁴

Submitted: 20 April 2019

Accepted: 13 October 2019

Habibzadeh, M.J., Dorani-Uliaie, E., Ziaratnia, S.M., and Valizadeh, M. 2020. Cloning and Bioinformatics investigation on CCD4a and CCD4b genes from Iranian saffron (*Crocus sativus* L.). Saffron Agronomy & Technology, 8(2): 211-229.

Abstract

Saffron is one of the most expensive spices and natural colors used in various food, pharmaceutical and cosmetic industries. In recent years, a family of enzymes that digest carotenoid substrates into double bonds are identified and introduced in plants. This family is of enzymes Carotenoid Cleavage Dioxygenase (CCD) enzymes. In this study, two isoforms of this gene were cloned and sequenced due to the importance of CCD genes in biosynthesis of apocarotenoids. Bioinformatics analyses including phylogenetic relationships and protein structures were evaluated. 3D modeling of these proteins was done by homologous modeling and using the Swiss Model database after selecting the appropriate pattern. The Ramachandran plot was drawn in order to validate the structure of the 3D model. The results show that the two CCD4a and CCD4b isoforms have both exons and one intron. *In silico* analysis, the physicochemical properties of CsCCD4a and CsCCD4b proteins also show that the proteins derived from these two isoforms are similar in terms of molecular weight, amino acids, isoelectric points, aliphatic index, instability index and solubility. The results of study of 3D structures resulted in proposal of similar structures for two isoforms. The results of this study can provide valuable information on the behavior and response of CCD4 enzyme in the pathway for synthesis of apocarotenins in saffron, and these results can be useful in future protein engineering programs.

Keywords: *Crocus sativus*, CCD4, Phylogenetic, Bioinformatics, Modeling.

1 - Ph.D. Student, Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran.

2 - Associate Professor, Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz

3 - Assistant Professor, Department of Food Biotechnology, Research Institute of Food Science and Technology (RIFST), Mashhad, Iran.

4 - Professor, Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran.

(* - Corresponding author. Email: dorani@tabrizu.ac.ir)

DOI: 10.22048/jsat.2019.182490.1343