

استفاده از تجزیه و تحلیل بیوانفورماتیکی برای شناسایی اهداف بالقوه پروتئینی در پاتوژن

Pseudomonas syringae به منظور کنترل این پاتوژن

کامران سمیعی^{۱*}، سید محسن سهرابی^۲

۱. گروه مهندسی کشاورزی، واحد کنگاور، دانشگاه آزاد اسلامی، کنگاور، ایران

۲. باشگاه پژوهشگران جوان و نخبگان، واحد خرم آباد، خرم آباد، ایران

چکیده

سابقه و هدف: شناسایی اهداف اختصاصی در هر پاتوژن برای جلوگیری از اثرهای سوء مصرف سموم شیمیایی ضروری است. در مطالعه حاضر، با استفاده از یک روش بیوانفورماتیکی مبتنی بر همولوژی، پروتئوم پاتوژن *P. syringae* با پروتئوم گیاه گوجه فرنگی و دیگر محصول های کشاورزی میزبان آن مقایسه و تعدادی پروتئین به عنوان اهداف بالقوه برای مبارزه در پاتوژن انتخاب شدند.

مواد و روش ها: پروتئوم پاتوژن با گیاهان میزبان مقایسه و پروتئین های غیرمشابه انتخاب شدند. این پروتئین ها با پروتئوم انسان و سایر موجودات غیرهدف مقایسه و پروتئین های غیرمشابه انتخاب و اهداف بالقوه تحت آنالیزهای کیفی مختلفی قرار گرفتند.

یافته ها: نتایج آنالیزها، ۹۵ پروتئین را به عنوان اهداف منحصر بفرد مشخص کرد. از مجموع پروتئین ها، ۵۹ پروتئین در سیتوپلاسم، ۱۵ پروتئین در غشاء داخلی، ۱۷ پروتئین در فضای پری پلاسمیک، ۳ پروتئین در غشاء خارجی و ۱ پروتئین در خارج از سلول پاتوژن *P. syringae* تجمع پیدا می کردند.

نتیجه گیری: پروتئین های مشخص شده در این مطالعه، اهدافی بسیار مؤثر برای مبارزه بوده و هدف قرار دادن چنین پروتئین هایی می تواند به طور مؤثری باعث کنترل پاتوژن *P. syringae* شود. هم چنین طراحی و تولید مواد شیمیایی علیه این پروتئین های هدف، در کنار کنترل مؤثر این پاتوژن، می تواند طیف وسیعی از پاتوژن ها را کنترل کرده و از اثرهای مضر بر انسان، موجودات غیرهدف و محیط زیست جلوگیری کند.

واژه های کلیدی: بیوانفورماتیک، گوجه فرنگی، پاتوژن *P. syringae*، پروتئین های هدف

مقدمه

گوجه فرنگی گیاهی از خانواده Solanaceae و از جنس *Solanum* و گونه اهلی آن *S. lycopersicum* است. گوجه فرنگی به طور ذاتی گیاهی علفی و چندساله است که در تمام نقاط جهان اغلب به صورت یکساله کشت می شود. گوجه فرنگی یکی از محصول های ارزشمند سبزی و صیفی در خاورمیانه به شمار می آید که پس از سیب زمینی از نظر اقتصادی در مقام دوم جهان قرار دارد. موطن اصلی گوجه فرنگی آمریکای مرکزی و جنوبی و به احتمال زیاد

نویسنده مسئول:

گروه مهندسی کشاورزی، واحد کنگاور، دانشگاه آزاد اسلامی، کنگاور، ایران.

پست الکترونیکی: kamransamiei@yahoo.com

تاریخ دریافت: ۱۳۹۶/۰۲/۲۴

تاریخ پذیرش: ۱۳۹۶/۱۲/۰۹

سواحل غربی آمریکای جنوبی است. گوجه فرنگی و محصول های تبدیلی آن یکی از مهم ترین تولیدات صنایع تبدیلی در جهان هستند به طوری که سالانه ۳۰ الی ۳۵ میلیون تن گوجه فرنگی تازه در کارخانجات فراهم می گردد. ایران از جنبه فرآوری گوجه فرنگی پس از کشورهای آمریکا، ایتالیا، چین، ترکیه در رتبه پنجم قرار دارد. بیماری های گیاهی یکی از مهم ترین عوامل محدودکننده کشت محصول های کشاورزی از جمله گوجه فرنگی هستند و سالانه خسارت قابل توجهی را به این محصول ها وارد می کنند. از مهم ترین بیماری های باکتریایی گوجه فرنگی می توان به لکه باکتریایی یا خال زدگی *P. syringae p.v tomato*، پژمردگی باکتریایی *Clavibacter solanacearum*، شانکر باکتریایی *Xanthomonas campestris michiganensis* و لکه باکتریایی *pv. vesicatoria* اشاره کرد (۱).

بین بردن آن است. با این روش علاوه بر کنترل مؤثر پاتوژن می توان از اثرهای سوء و زیان آور احتمالی مبارزه با پاتوژن بر روی انسان، موجودات غیرهدف و محیط زیست تا حد زیادی جلوگیری کرد (۱،۲۹). یکی از بهترین، سریع ترین و کم هزینه ترین روش های شناسایی اهداف در موجودات مختلف، روش های بیوانفورماتیکی هستند. بیوانفورماتیک شاخه ای از علم زیست شناسی است که به ذخیره، تجزیه و تحلیل و تفسیر اطلاعات آزمایشگاهی می پردازد (۱۰،۲۲). بیوانفورماتیک علم نوینی است که در آن با استفاده از کامپیوتر، نرم افزارهای کامپیوتری و بانک های اطلاعاتی سعی می گردد تا به مسائل بیولوژیکی به خصوص در زمینه های سلولی و مولکولی پاسخ داده شود (۱۸). در این علم با به کارگیری کامپیوتر سعی می گردد تا تحقیق های وسیع تری در خصوص پروتئین ها و ژن ها به عمل آید. بیوانفورماتیک توسط علوم می مانند زیست شناسی محاسبه ای، ریاضی، بیومتری، تکامل مولکولی، ژنومیکس، پروتئومیکس و زیست شناسی سلولی و مولکولی توسعه یافته است. وجه مشترکی که تمام این علوم در بیوانفورماتیک استفاده از کامپیوتر و برنامه های کامپیوتری است. در این مطالعه با استفاده از ابزار BLASTp و ماتریس BLOSUM62 و همچنین پایگاه های NCBI، KEGG، و VFDB، پروتئوم پاتوژن *P. syringae* با پروتئوم گیاهان میزبان، انسان و فلور میکروبی دستگاه گوارش و دیگر موجودات مانند مهره داران، بندپایان و نرم تنان مقایسه شد و در نهایت تعدادی پروتئین به عنوان اهداف منحصر به فرد برای مبارزه با پاتوژن *P. syringae* انتخاب شدند. اهداف انتخابی توسط ابزارهای DEG، CELLO، PSORT، Pfam، InterProScan و CDD مورد بررسی بیش تر قرار گرفتند (۳،۲۷).

مواد و روش ها

نمای کلی مراحل آنالیز در این مطالعه در شکل ۱ نشان داده شده است. آنالیز شامل ۳ مرحله هم ردیفی و چندین مرحله آنالیز کیفی است (شکل ۱). ابتدا، چند دسته از اطلاعات از منابع مختلف شامل مقالات، کتب و پایگاه ها استخراج و جمع آوری شد. اطلاعات مربوط به پاتوژن *P. syringae* شامل اندازه ژنوم، تعداد ژن ها، تعداد پروتئین ها، فاکتورهای بیماری زایی و اطلاعات مسیرهای بیوشیمیایی پاتوژن از پایگاه های NCBI، VFDB، KEGG و دریافت و دسته بندی شد. همچنین اطلاعات مربوط به میزبان های مختلف پاتوژن *P. syringae* شامل جنس های گیاهی *Solanum*، *Dysoxylum*، *Triticum*، *Beta*، *Aesculus*، *Actinidia*، *Acer*، *Prunus*، *Pisum*، *Phaseolus*، *Malus*، *Panicum*، *Hordeum*

پاتوژن *P. syringae* یک باکتری میله ای شکل، گرم منفی و تاژک دار از خانواده *Pseudomonadaceae* است. این پاتوژن دارای پاتووارهای زیادی بوده و در طیف وسیعی از گیاهان ایجاد بیماری می کند. این پاتوژن برای اولین بار در گیاه یاس بنفش یا *Syringa vulgaris* جداسازی شده و به همین علت *P. syringae* نام گرفته است. پاتووارهای مختلف این پاتوژن به گیاهانی از خانواده های *Sapindaceae*، *Actinidiaceae*، *Rosaceae*، *Meliaceae*، *Poaceae*، *Amaranthaceae*، *Fabaceae* و *Oleaceae* حمله و ایجاد بیماری می کنند. این پاتوژن بیماری های مختلفی مانند لکه برگ، شانکر، آتشک و سوختگی را در گیاهان میزبان خود ایجاد می کند. با توجه به طیف وسیع میزبانی و شدت ایجاد بیماری، پاتوژن *P. syringae* به عنوان یکی از مهم ترین باکتری های بیماری زای گیاهی مطرح بوده و سالیانه خسارت قابل توجهی را در سراسر دنیا به محصول های مختلف کشاورزی وارد می کند (۳۱،۹،۱).

ژنوم کامل پاتوژن *P. syringae* در سال ۲۰۰۳ توالی یابی و مستندسازی شد. پاتوژن *P. syringae* دارای ژنومی به طول ۶۰۹۳۶۹۸ جفت باز و فاقد پلاسمید است. این ژنوم دارای محتوای GC حدود ۵۹ درصد بوده و ۵۲۲۰ ژن دارد. این تعداد ژن، ۵۰۸۹ پروتئین، ۱۶ RNA ریبوزومی، ۶۴ tRNA و ۳ نوع دیگر RNA را کد کرده و ۴۷ ژن کاذب دارند. ژنوم پاتوژن *P. syringae* با شماره NC-007005.1 قابل دسترسی است (Buell et al., 2003). در مطالعه ای، پروتئوم گیاه گوجه فرنگی آلوده شده با پاتوژن *P. syringae* بررسی شد و نتایج ۴۷۷ را مشخص کرد که در پاسخ به آلودگی نقش داشتند (۲۱). در پژوهش دیگری، با استفاده از دو روش محاسبه ای، پروتئوم پاتوژن *P. syringae* و پروتئوم گیاه آراییدوپسیس با هم مقایسه و برهم کنش های بین آن ها بررسی شد. نتایج، وجود ۷۹۰ هزار برهم کنش را بین دو پروتئوم نشان داد (۲۴).

طبق یک تخمین محافظه کارانه، بیماری ها، حشرات و علف های هرز در سطح جهانی سالانه بین ۳۱ تا ۴۲ درصد محصول های کشاورزی را نابود یا از تولید آن ها جلوگیری می کنند. با توجه به این میزان خسارت مبارزه با بیماری ها به منظور حفظ امنیت غذایی و رونق اقتصادی امری اجتناب ناپذیر به نظر می رسد. از بین روش های مبارزه با بیماری های گیاهی دو روش مهندسی ژنتیک و استفاده از مواد شیمیایی غیر مضر برای انسان، موجودات غیرهدف و محیط زیست روش های مؤثرتری به نظر می رسند. هر دوی این روش ها نیازمند شناسایی و تعیین اهداف مؤثر در پاتوژن، به منظور جلوگیری از مکانیسم های آلودگی، رشد و تولیدمثل و همچنین از

Phaseolus, Malus, Panicum, Hordeum, Dysoxylum, Prunus, Pisum و Glycine، پروتئوم انسان و فلور میکروبی دستگاه گوارش، پروتئوم پاتوژن های باکتریایی گیاهی و پروتئوم مهره داران، بندپایان و نرم تنان چهار پایگاه داده اختصاصی در برنامه BLASTp پایگاه NCBI ایجاد شد. این پایگاه های داده اختصاصی بر اساس کدنویسی اختصاصی و برای محدود کردن هم ردیفی های بعدی و کاهش زمان هم ردیفی ها ایجاد شدند (۲).

و Glycine، اطلاعات مربوط به فلور میکروبی دستگاه گوارش انسان، اطلاعات سایر پاتوژن های باکتریایی گیاهی نیز از پایگاه های متنی شامل Pubmed، PMC، Google scholar دریافت و دسته بندی شد (۷،۱۳،۲۰).

در مرحله بعد، پروتئوم پاتوژن P.syringae (شامل ۵۰۸۹) و پروتئین از پایگاه NCBI دریافت شد. سپس، از پروتئوم میزبان های مختلف پاتوژن P.syringae شامل جنس های گیاهی Triticum, Beta, Aesculus, Actinidia, Acer, Solanum



شکل ۱- نمای کلی مراحل آنالیز

این مرحله از آنالیز برای جلوگیری از اثرهای ناخواسته هدف گیری پروتئین های پاتوژن بر انسان و سایر موجودات غیرهدف انجام می شود. کد پایگاه های داده سفارشی پروتئوم میزبان های گیاهی و پروتئوم انسان، فلور میکروبی دستگاه گوارش، مهره داران، بندپایان و نرم تنان در جداول ۱ و ۲ نشان داده شده است.

در مرحله بعد، توالی پروتئینی اهداف اولیه، از پایگاه NCBI دریافت شد. توالی های مورد نظر پس از دسته بندی، با ارزش مورد انتظار ۰/۰۰۰۱ علیه پروتئین های موجود در پایگاه DEG10.0 تحت هم ردیفی قرار گرفتند. پروتئین هایی از پروتئوم پاتوژن که در ارزش مورد انتظار ۰/۰۰۰۱ با پروتئین های موجود در پایگاه DEG10.0 مشابه بودند به عنوان اهداف بالقوه نهایی در پاتوژن P.syringae انتخاب و بقیه پروتئین ها حذف شدند. اهداف بالقوه نهایی پاتوژن P.syringae برای بررسی های کیفی بیش تر به مراحل بعدی منتقل شدند. ژن های ضروری، ژن هایی هستند که برای بقا میکروارگانیسم حیاتی بوده و بدون وجود آنها میکروارگانیسم قادر به ادامه حیات، رشد و نمو، هموستازی و تولیدمثل نخواهد بود. پایگاه DEG10.0 مجموعه ای از ژن های ضروری در میکروارگانیسم های باکتریایی مختلف است (۱۶).

پروتئوم پاتوژن P. syringae با ارزش مورد انتظار ۰/۰۰۱ علیه پایگاه داده اختصاصی میزبان های مختلف پاتوژن تحت هم ردیفی قرار گرفت. پروتئین های مشابه با پروتئین های موجود در پایگاه داده اختصاصی میزبان ها حذف شده و بقیه پروتئین ها که هیچ شباهتی با پروتئین های موجود در این پایگاه داده اختصاصی نداشتند وارد مرحله بعدی هم ردیفی شدند. پروتئین های حاصل از مرحله قبل با ارزش مورد انتظار ۰/۰۰۱ علیه پایگاه داده اختصاصی انسان و فلور میکروبی دستگاه گوارش هم ردیف شدند و پروتئین های مشابه با پروتئین های موجود در پایگاه داده اختصاصی انسان و فلور میکروبی دستگاه گوارش حذف شده و بقیه پروتئین ها که هیچ شباهتی با پروتئین های موجود در این پایگاه داده اختصاصی نداشتند وارد مرحله بعدی هم ردیفی شدند. در مرحله آخر، پروتئین های باقی مانده از مرحله قبل با ارزش مورد انتظار ۰/۰۰۱ علیه پایگاه داده اختصاصی مهره داران، بندپایان و نرم تنان مورد هم ردیفی قرار گرفته و پروتئین های مشابه با پروتئین های موجود در پایگاه داده اختصاصی مهره داران، بندپایان و نرم تنان حذف شده و بقیه پروتئین ها که هیچ شباهتی با پروتئین های موجود در این پایگاه داده اختصاصی نداشتند به عنوان اهداف اولیه در پاتوژن P.syringae انتخاب و وارد مراحل بعدی آنالیز شدند.

جدول ۱- کد پایگاه داده سفارشی میزبان های گیاهی

txid49274 [ORGN] OR txid4022 [ORGN] OR txid3624 [ORGN] OR txid43363 [ORGN] OR txid3554 [ORGN] OR txid4564 [ORGN] OR txid15563 [ORGN] OR txid4512 [ORGN] OR txid4539 [ORGN] OR txid3749 [ORGN] OR txid3883 [ORGN] OR txid3887 [ORGN] OR txid3754 [ORGN] OR txid3846 [ORGN]

طیف وسیع و جستجوی مسیرهای متابولیکی تشکیل شده است. در جستجوی محل تجمع سلولی، جستجوی عملکردی، جستجوی

آنالیز کیفی اهداف بالقوه، از چهار جستجوی مختلف شامل جستجوی محل تجمع سلولی، جستجوی عملکردی، جستجوی

پاتوزن باکتریایی گیاهی مهم (مانند *Erwinia Agrobacterium*، *Ralstonia P. pseudomonas* و *Pectobacterium*) مورد هم-ردیفی قرار گرفتند. اهداف مشترک با بیش از ۴ پاتوزن به عنوان اهداف طیف وسیع مشخص شدند. در جستجوی مسیرهای متابولیکی، اهداف پروتئینی در پایگاه KEGG بررسی شدند و مسیرهای متابولیکی که این اهداف در آنها نقش دارند مشخص شدند (۱۳).

CELLO و PSORT محل تجمع اهداف شناسایی شده مشخص شد (۱۹،۳۲). محل های تجمع پروتئین در سلول باکتری های گرم منفی شامل سیتوپلاسم، غشاء داخلی، ناحیه پری پلاسمیک، غشاء خارجی و خارج سلول است. در جستجوی عملکردی، عملکرد اهداف پروتئینی با عملکرد ناشناخته با استفاده از ابزارهایی مانند Pfam، InterProScan و CDD شناسایی و مشخص شد (۸،۱۷). در جستجوی طیف وسیع، اهداف نهایی با ارزش مورد انتظار ۰/۰۰۱ علیه پایگاه داده اختصاصی متشکل از پروتئوم ۲۷ جنس

جدول ۲- کد پایگاه داده سفارشی انسان، فلور میکروبی دستگاه گوارش، مهره داران، بندپایان و نرم تنان

txid9606 [ORGN] OR txid411466 [ORGN] OR txid349741 [ORGN] OR txid445970 [ORGN] OR txid445971 [ORGN] OR txid411490 [ORGN] OR txid445972 [ORGN] OR txid411467 [ORGN] OR txid537012 [ORGN] OR txid470145 [ORGN] OR txid483217 [ORGN] OR txid483216 [ORGN] OR txid483215 [ORGN] OR txid471870 [ORGN] OR txid411476 [ORGN] OR txid483218 [ORGN] OR txid484018 [ORGN] OR txid449673 [ORGN] OR txid411479 [ORGN] OR txid367928 [ORGN] OR txid411481 [ORGN] OR txid518635 [ORGN] OR txid500634 [ORGN] OR txid518634 [ORGN] OR txid473819 [ORGN] OR txid205913 [ORGN] OR txid206672 [ORGN] OR txid1682 [ORGN] OR txid498739 [ORGN] OR txid511680 [ORGN] OR txid451640 [ORGN] OR txid518636 [ORGN] OR txid445973 [ORGN] OR txid411902 [ORGN] OR txid500633 [ORGN] OR txid428125 [ORGN] OR txid537013 [ORGN] OR txid500632 [ORGN] OR txid445974 [ORGN] OR txid411468 [ORGN] OR txid411489 [ORGN] OR txid411486 [ORGN] OR txid411484 [ORGN] OR txid428126 [ORGN] OR txid471871 [ORGN] OR txid411472 [ORGN] OR txid411903 [ORGN] OR txid521003 [ORGN] OR txid445975 [ORGN] OR txid470146 [ORGN] OR txid411474 [ORGN] OR txid411461 [ORGN] OR txid411462 [ORGN] OR txid479437 [ORGN] OR txid500639 [ORGN] OR txid428127 [ORGN] OR txid411469 [ORGN] OR txid428128 [ORGN] OR txid411463 [ORGN] OR txid411483 [ORGN] OR txid411485 [ORGN] OR txid362948 [ORGN] OR txid420247 [ORGN] OR txid521001 [ORGN] OR txid521002 [ORGN] OR txid483214 [ORGN] OR txid500635 [ORGN] OR txid387661 [ORGN] OR txid411477 [ORGN] OR txid411465 [ORGN] OR txid243265 [ORGN] OR txid537011 [ORGN] OR txid520999 [ORGN] OR txid521000 [ORGN] OR txid500637 [ORGN] OR txid451638 [ORGN].

نتایج

انتخاب کرد و به مرحله بعد انتقال داد. به علت عدم تشابه بین ۳۴۷۷ پروتئین از پاتوزن با پروتئوم میزبان ها، هدف قرار دادن این پروتئین ها در پاتوزن هیچ گونه اثر مضر بر میزبان ها نخواهد گذاشت. با این وجود در میان این ۳۴۷۷ پروتئین ممکن است پروتئین های وجود داشته باشند که دارای مشابهت با پروتئین های انسان و موجودات غیرهدف باشند (۳۰،۱۴). برای پیدا کردن این پروتئین ها و حذف اثرهای ناخواسته هدف-گیری آن ها روی انسان و موجودات غیرهدف چند مرحله دیگر از هم ردیفی انجام گرفت. ۳۴۷۷ پروتئین مرحله قبل با ارزش مورد انتظار ۰/۰۰۱ علیه پروتئوم انسان و فلور میکروبی دستگاه گوارش هم ردیف شدند. از ۳۴۷۷ پروتئین ورودی به هم ردیفی ۲۱۷۷ پروتئین به علت مشابهت قابل توجه به پروتئوم انسان و فلور میکروبی دستگاه گوارش حذف و ۱۳۰۰ پروتئین باقی مانده هیچ شباهتی با پروتئوم انسان و فلور میکروبی دستگاه گوارش نشان ندادند و برای مرحله بعدی آنالیز انتخاب شدند. تا این مرحله ۱۳۰۰ پروتئین از پاتوزن هیچ شباهتی به پروتئوم میزبان ها، انسان و فلور میکروبی دستگاه گوارش نشان ندادند. هدف گیری این ۱۳۰۰ پروتئین با اطمینان بالایی هیچ اثر مضر بر میزبان ها، انسان و فلور میکروبی دستگاه گوارش آن نخواهد داشت. با این وجود در میان این ۱۳۰۰ پروتئین ممکن است پروتئین هایی مشابه با پروتئین های موجودات غیرهدف وجود داشته باشند (۱۴،۳۰).

ژنوم *P. syringae* ۵۲۲۰ ژن داشته و ۵۰۸۹ پروتئین تولید می کند. از ۵۰۸۹ پروتئین پاتوزن *P. syringae* تعداد ۲۵۳۶ پروتئین توسط رشته مثبت و ۲۵۵۳ پروتئین توسط رشته منفی کد می شوند. طول پروتئین های موجود در پاتوزن *P. syringae* از ۲۴ تا ۱۳۵۳۷ اسید آمینه متغیر بوده و تعداد عمده پروتئین ها در محدوده ۱۰۰ تا ۵۰۰ اسید آمینه قرار دارند. همچنین این پاتوزن دارای ۲۵۹ مسیر متابولیکی و غیرمتابولیکی برای فرآیندهای حیاتی مختلف است. مسیرهای متابولیکی شامل متابولیسم انرژی، کربوهیدرات ها، لیپیدها، نوکلئوتیدها، آمینواسیدها، سایر آمینوها، گلیکان ها، ویتامین ها، کوفاکتورها و متابولیت های ثانویه هستند. مسیرهای غیرمتابولیکی شامل مسیرهای پردازش اطلاعات ژنتیکی، پردازش اطلاعات محیطی و پردازش سلولی هستند. تمامی مسیرهای متابولیکی و غیر متابولیکی اصلی یاد شده دارای مسیرهای فرعی و مسیرهای زیر مجموعه ای هستند (۲۳،۲۷). از ۵۰۸۹ پروتئین ورودی، ۱۶۱۲ پروتئین پاتوزن دارای مشابهت قابل توجهی با پروتئوم میزبان ها بودند که از آنالیز حذف شدند و ۳۴۷۷ پروتئین هیچ مشابهتی با پروتئوم میزبان ها نشان ندادند و برای مرحله بعدی آنالیز انتخاب شدند. در این مرحله پروتئین های مشابه با پروتئین های میزبان حذف شدند، بنابراین با اطمینان بالایی می توان بقیه پروتئین ها را به عنوان اهداف اولیه برای مبارزه

برای حذف اثرهای ناخواسته هدف گیری این پروتئین های مشابه بر موجودات غیرهدف، مرحله دیگری از هم ردیفی انجام شد و در آن ۱۳۰۰ پروتئین مرحله قبل با ارزش مورد انتظار ۰/۰۰۱ علیه پروتئوم موجودات غیرهدف شامل مهره داران، بندپایان و نرم تنان هم ردیف شدند. از ۱۳۰۰ پروتئین هم ردیف شده ۲۲۳ پروتئین دارای مشابهت زیادی با پروتئوم موجودات غیرهدف بودند و حذف شدند. ۱۰۷۷ پروتئین باقی مانده پروتئین باقی مانده هیچ شباهتی با پروتئوم موجودات غیرهدف نداشتند و برای مرحله بعدی آنالیز انتخاب شدند. تا این مرحله ۱۰۷۷ پروتئین در پاتوژن شناسایی شده که هیچ شباهتی به پروتئوم میزبان ها، انسان، فلور میکروبی دستگاه گوارش و موجودات غیرهدفی مانند مهره داران، بندپایان و نرم تنان ندارند و هدف گیری آن ها هیچ اثر مضر روی این موجودات نخواهد داشت. با این حال، هدف گیری هر پروتئینی نمی تواند به طور مؤثری پاتوژن را کنترل کند، زیرا بسیاری از پروتئین ها برای پاتوژن ضروری نبوده و پاتوژن بدون آن ها نیز قابلیت زنده ماندن، رشد و تکثیر را خواهد داشت. برای کنترل مؤثر پاتوژن، باید پروتئین های حیاتی که ارگانیزم بدون آن ها توانایی زنده ماندن، رشد و تکثیر را از دست خواهد داد را هدف قرار داده شود. به این نوع پروتئین های حیاتی، پروتئین های ضروری گویند و ژن های کدکننده چنین پروتئین هایی را ژن های ضروری می نامند. بنابراین در مرحله آخر، ۱۰۷۷ پروتئین هدف با ارزش مورد انتظار ۰/۰۰۱ علیه پایگاه ژن های ضروری باکتریایی مورد هم ردیفی قرار گرفتند. از ۱۰۷۷ پروتئین مورد بررسی، ۹۸۲ پروتئین ضروری نبودند و از آنالیز حذف شدند و ۹۵ پروتئین که مشابهت بالایی با پروتئین های موجود در پایگاه ژن های ضروری نشان دادند، به عنوان اهداف بالقوه نهایی در پاتوژن *P.syringae* انتخاب شدند (جدول ۳ و ۴).

از ۹۵ هدف شناسایی شده، ۵۳ هدف توسط رشته مثبت و ۴۲ هدف توسط رشته منفی کد می شوند. هم چنین عمده اهداف شناسایی شده در محدوده طولی بین ۱ تا ۵۰۰ اسید آمینه ای قرار گرفته و روندی مشابه با تمام پروتئین های موجود در پاتوژن *P.syringae* دارند. ۹۵ هدف شناسایی شده دارای نقش های مختلفی در سلول پاتوژن *P.syringae* هستند. برخی اهداف در متابولیسم کربوهیدرات، انرژی، چربی، نوکلئوتیدها، آمینواسیدها، سایر آمین ها، گلیکان ها، ویتامین ها و کوفاکتورها و متابولیت های ثانویه دخالت دارند. برخی اهداف دارای نقش هایی در پردازش اطلاعات ژنتیکی و محیطی هستند و برخی در فرآیندهای سلولی نقش دارند. اهداف شناسایی شده در قسمت های مختلفی از سلول پاتوژن *P.syringae* تجمع پیدا می کنند. برخی اهداف دارای محل

تجمع سیتوپلاسمی هستند، برخی در غشاء داخلی مجتمع می شوند و برخی در فضای پری پلاسمیک تجمع دارند. غشاء خارجی و خارج سلول هم محل های تجمع برخی از اهداف هستند. به طور کلی از ۹۵ هدف شناسایی شده، ۵۹ هدف در سیتوپلاسم، ۱۵ هدف در غشاء داخلی، ۱۷ هدف در فضای پری پلاسمیک، ۳ هدف در غشاء خارجی و ۱ هدف در خارج سلول پاتوژن *P.syringae* تجمع پیدا می کنند (شکل ۲). تعیین محل استقرار پروتئین هدف یکی از مهم ترین مراحل در مبارزه با پاتوژن است. اهداف موجود در سیتوپلاسم و بخش های داخلی سلول پاتوژن اهداف مناسبی برای مبارزه شیمیایی هستند، در حالی که اهداف موجود در غشاء برای طراحی پلنتی بادی ها و تولید گیاهان تراریخت مقاوم بیان کننده این پلنتی بادی ها مناسب هستند. با توجه به نتایج آنالیز محل تجمع سلولی، بیش تر اهداف شناسایی شده به علت تجمع سیتوپلاسمی برای مبارزه شیمیایی مناسب هستند. عوامل شیمیایی طراحی شده علیه اهداف سیتوپلاسمی باید به سلول وارد شوند تا بتوانند هدف مورد نظر را نابود کنند. در حالی که آنتی بادی های تولید و ترشح شده توسط گیاهان تراریخت می توانند در فضای بین سلولی گیاه با اهداف موجود در بخش های خارجی سلول پاتوژن برهم کنش کرده و پاتوژن را غیرفعال کنند یا از بین ببرند. برخی از پروتئین های هدف شناسایی شده دارای عملکرد ناشناخته هستند، بنابراین، با استفاده از ابزارهایی مانند Pfam و CDD عملکرد آن ها مشخص شد. بیست و یک پروتئین هدف دارای دامنه DUF بودند و عملکرد آن ها هم چنان ناشناخته باقی ماند. سایر دامنه های عملکردی شناسایی شده Zinc- or iron-chelating, Lipopolysaccharide kinase, domain, Cupin, Winged helix-turn helix, domain, Sigma-factor, ATPgrasp, ATPase, TarH domain, MacB-like periplasmic core, FtsX-like permease, TonB, Lipopolysaccharide-assembly, domain, Membrane protein, ZapA و GAF domain هستند. وجود دامنه های عملکردی مختلف در این پروتئین ها نشان دهنده فعالیت این پروتئین ها در فرآیندهای مختلف سلولی است. پروتئین هایی که عملکرد آن ها در این آنالیز مشخص شد دارای نقش های مختلفی در فرآیندهای متابولیسم کربوهیدرات ها، انرژی، چربی، نوکلئوتیدها، آمینواسیدها، سایر آمین ها، گلیکان ها، ویتامین ها، کوفاکتورها و متابولیت های ثانویه و پردازش اطلاعات ژنتیکی، محیطی و دیگر فرآیندهای سلولی بودند. دو هدف پروتئینی در سیستم دو جزئی باکتریایی دخیل هستند. با استفاده از سیستم دو جزئی، پاتوژن به محرک های محیطی پاسخ داده و بر اساس آن

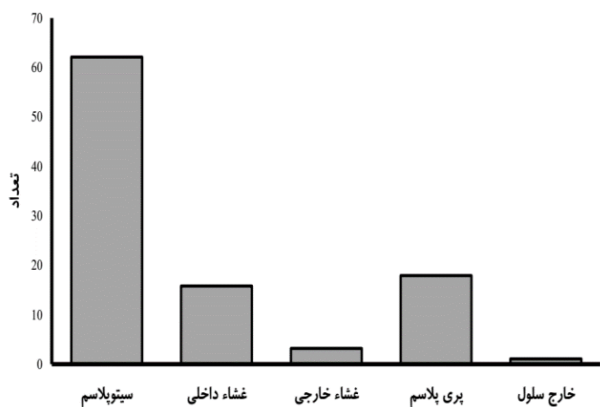
پروتئینی دارای نقش تنظیم کنندگی در بیان ژن هستند و بیان گروه های خاصی از ژن ها را تنظیم می کنند. هدف گیری پروتئین های تنظیمی ممکن است باعث مختل کردن بیان تعداد زیادی از ژن هایی شود که به وسیله آن پروتئین خاص تنظیم می شوند. سه هدف پروتئینی در تقسیم سلولی پاتوژن دخیل هستند و هدف قرار دادن آن ها ممکن است باعث اختلال در تولید مثل سلول باکتری شده و به طور مؤثری باعث کنترل پاتوژن شود.

شرایط درونی سلول خود را تنظیم می کند. این سیستم هم چنین در بیماری زایی پاتوژن نقش دارد و به علت عدم وجود آن در یوکاریوت ها و انسان هدف مناسبی برای مبارزه است (۶). دوازده هدف پروتئینی دارای نقش در انتقال مواد به داخل و خارج سلول پاتوژن هستند و به عنوان پمپ، کانال و ناقل عمل می کنند. هدف گیری چنین پروتئین هایی باعث اختلال در نقل و انتقال مواد در سلول پاتوژن شده و باعث از بین رفتن آن خواهد شد. پنج هدف

جدول ۳- تعداد پروتئین های ورودی، خروجی و حذف شده در هر مرحله آنالیز

تعداد خروجی	تعداد حذف شده	تعداد ورودی	نوع آنالیز
۳۴۷۷	۱۶۱۲	۵۰۸۹	هم ردیفی با پایگاه داده اختصاصی میزبان ها
۱۳۰۰	۲۴۰۰	۳۴۷۷	هم ردیفی با پایگاه داده اختصاصی انسان و فلور میکروبی دستگاه گوارش
۱۰۷۷	۲۲۳	۱۳۰۰	هم ردیفی با پایگاه داده اختصاصی مهره داران، بند پایان و نرم تنان
۹۵	۹۸۲	۱۰۷۷	آنالیز ضروری بودن

در حالی که هدف گیری هم زمان چند مسیر احتمال ایجاد مقاومت در پاتوژن را تا حد زیادی کاهش می دهد. پروتئین هایی که در چند مسیر متابولیکی مختلف نقش دارند اهداف مناسب تری برای مبارزه هستند زیرا هدف گیری چنین پروتئین هایی باعث ایجاد اختلال در چند مسیر متابولیکی سلول شده و به طور مؤثرتری باعث از بین رفتن پاتوژن خواهد شد (۱۱، ۲۶).



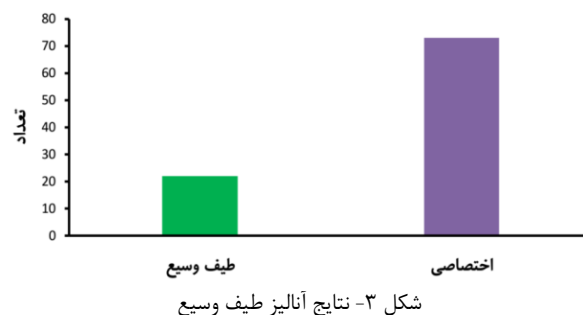
شکل ۲- محل تجمع اهداف بالقوه پیش بینی شده

در آنالیز طیف وسیع ۲۲ مورد از اهداف شناسایی شده در دسته بندی طیف وسیع قرار گرفتند و ۷۳ مورد باقی مانده به عنوان اهداف اختصاصی پاتوژن *P.syringae* در نظر گرفته شدند (شکل ۳).

یکی از مهم ترین و سودمندترین ویژگی های یک هدف برای مبارزه، طیف وسیع بودن آن است به این معنی که این هدف در طیف وسیعی از دیگر پاتوژن ها با شباهت بالا موجود بوده و محافظت شده باشد. از ۹۵ هدف شناسایی شده ۲۲ هدف طیف وسیع و ۷۳ هدف اختصاصی پاتوژن *P.syringae* بودند. مزیت پروتئین های هدف طیف وسیع این است که با هدف گیری آن ها می توان طیف وسیعی از پاتوژن ها را هدف قرار داد و از حمله و صدمه آن ها به گیاه جلوگیری کرد. هم چنین هدف گیری چنین اهدافی باعث کاهش قابل توجهی در وقت و هزینه مبارزه با دیگر پاتوژن ها خواهد شد. از طرفی در مبارزه شیمیایی می توان با سم تولید شده علیه این اهداف طیف وسیع، پاتوژن های مختلفی را از بین برد بدون اینکه نیاز به تولید سم اختصاصی برای هر پاتوژن باشد.

جستجو در پایگاه KEGG نشان داد که اهداف پیش بینی شده در مسیرهای متابولیکی حیاتی مختلفی از جمله متابولیسم کربوهیدرات ها، متابولیسم انرژی، متابولیسم اسیدهای آمینه دخیل هستند. از طرف دیگر برخی اهداف در مسیرهای تنظیمی به عنوان فاکتورهای تنظیمی بیان ژن نقش دارند و برخی در سیستم های نقل و انتقال مواد به درون و بیرون سلول نقش اساسی ایفا می کنند (۱۳).

از ۹۵ پروتئین هدف شناسایی شده تنها یک پروتئین در بیش از یک مسیر متابولیکی نقش داشت. در سلول، برخی پروتئین ها تنها در یک مسیر متابولیکی نقش دارند و برخی دیگر در بیش از یک مسیر متابولیکی نقش بازی می کنند. هدف قرار دادن یک مسیر متابولیکی خاص باعث ایجاد سریع مقاومت در پاتوژن ها خواهد شد



بحث و نتیجه گیری

بیماری های گیاهی همواره در طول تاریخ تهدید مهمی علیه محصول های کشاورزی تولیدی انسان ها بوده و کوچک ترین غفلت بشر از حضور آن ها در مقطعی باعث از بین رفتن تا صد درصد محصول شده است. لذا کنترل این عوامل خسارت زا اجتناب ناپذیر است. بیماری های باکتریایی یکی از مهم ترین عوامل ایجاد خسارت بر محصول های کشاورزی هستند. مبارزه با بیماری های باکتریایی به طور معمول پیچیده تر است و در مبارزه با این عوامل بیماری زا کمابیش از چندین روش مبارزه به صورت تلفیقی استفاده می شود. با این وجود، مبارزه شیمیایی هنوز قوی ترین و مؤثرترین روش مبارزه است. طی یک صد سال گذشته، کنترل بیماری ها و سایر آفات گیاهی به طور فزاینده ای به استعمال مواد شیمیایی سمی وابسته بوده است. کنترل بیماری های گیاهی استعمال این مواد سمی را نه تنها بر روی گیاه و محصول های گیاهی مورد مصرف ما، بلکه در داخل خاک جایی که بسیاری از میکروارگانیسم های بیماری زا زندگی کرده و به ریشه گیاه حمله می کنند، نیز ضروری ساخته است. ثابت شده که بسیاری از این مواد شیمیایی برای میکروارگانیسم های غیر هدف، حیوانات و حتی برای انسان نیز ممکن است سمی باشند. دشوار است که بتوان هزینه های بلند مدت و کوتاه مدت آلودگی محیط زیست بر سلامت و بهزیستی انسان را که بر اثر تلاش ما برای کنترل بیماری های گیاهی و سایر آفات به وجود می آید، تخمین زد.

مهم ترین مشکل مواد شیمیایی مانند آنتی بیوتیک ها و سایر سموم شیمیایی که برای کنترل بیماری های باکتریایی در مقیاس وسیع استفاده می شوند، وجود مولکول های هدف مشابه با میزبان و سایر موجودات غیر هدف است. مولکول های هدف این مواد شیمیایی در پاتوژن که به طور عمومی مولکول های پروتئینی هستند، در اغلب موارد دارای همولوگ های بسیار مشابهی در میزبان پاتوژن و موجودات غیر هدف هستند. هدف گیری این اهداف توسط مواد شیمیایی متداول باعث آسیب ناخواسته به میزبان و موجودات

غیر هدفی مانند انسان خواهد شد. کمابیش تمام مواد شیمیایی و آنتی بیوتیک هایی که برای مبارزه با بیماری های باکتریایی گیاهی به کار می روند دارای اهدافی در پاتوژن هستند که شباهت بالایی با میزبان، انسان و سایر موجودات غیر هدف دارد.

در پژوهش حاضر از روشی استفاده شد که در آن با چند مرحله هم ردیفی متوالی اهداف پروتئینی به دست آمد که هیچ شباهتی با میزبان و موجودات غیر هدف نداشتند (۵،۱۵،۲۵،۲۷،۲۸). در مرحله اول پروتئین هایی انتخاب شدند که هیچ شباهتی با پروتئین های میزبان های گیاهی نداشتند. تا این مرحله میزبان ها از اثرهای ناخواسته هدف گیری در امان می ماندند اما دیگر موجودات غیر هدف مانند انسان هم چنان در معرض اثرهای مضر ناخواسته هدف گیری هستند. در مرحله بعدی، پروتئین هایی باقی ماندند که هدف گیری آن ها هیچ اثر جانبی مضر بر میزبان ها، انسان، فلور میکروبی دستگاه گوارش، مهره داران، بندپایان و نرم تنان نخواهد داشت. اما هدف گیری این پروتئین ها ممکن است اثر قابل توجهی بر پاتوژن نداشته باشد، که به علت ضروری و حیاتی نبودن چنین پروتئین هایی است. برای حل این مشکل، آنالیز ضروری بودن انجام و تنها ۹۵ پروتئین ضروری و حیاتی تشخیص داده شدند. این ۹۵ پروتئین اهداف نهایی را تشکیل داده و مورد آنالیزهای تکمیلی قرار گرفتند. آنالیزهای تکمیلی محل تجمع، عملکرد پروتئین های با عملکرد ناشناخته و مسیرهای متابولیکی که پروتئین در آن ها دخیل هستند را مشخص کرد. علاوه بر این پروتئین های هدفی که علاوه بر پاتوژن *P. syringae* پاتوژن های گیاهی دیگر حضور داشتند را معین نمود. در نهایت این ۹۵ پروتئین، اهدافی به-طور کامل اختصاصی و مشترک در بسیاری از پاتوژن های گیاهی بوده که در فرآیندهای بسیار حیاتی سلول نقش دارند و برای بقاء پاتوژن ضروری هستند. این پروتئین ها در بخش های مختلفی از سلول تجمع دارند که باعث آزادی عمل در انتخاب نوع سموم شیمیایی می شود. این اهداف در مسیرهای متابولیکی حیاتی پاتوژن نقش داشته و هدف گیری آن ها باعث مختل شدن این مسیرها و از بین رفتن پاتوژن می شود. این اهداف در محل های مختلفی مانند سیتوپلاسم، غشاء داخلی، غشاء خارجی، پری پلاسم و خارج سلول تجمع پیدا می کنند. اهداف موجود در پری پلاسم، غشاء خارجی، غشاء داخلی و خارج سلول برای هدف گیری توسط مولکول هایی مانند پلنتی بادی ها مناسب هستند و می توان گیاهان تراریختی با بیان آنتی بادی های اختصاصی علیه این اهداف تولید کرد. چنین گیاهانی می توانند به طور مؤثری در مقابل پاتوژن های تولیدکننده این اهداف مقاوم باشند. اهداف موجود در سیتوپلاسم برای هدف گیری توسط عوامل ضد میکروبی مناسب هستند. عوامل

غیرهدف و همچنین محیط زیست خواهد شد و سلامت آن ها را به خطر خواهد انداخت. باتوجه به گسترش روز افزون اطلاعات و داده های بیولوژیک و پیشرفت نرم افزارها و روش های محاسبه ای استفاده از روش های بیوانفورماتیکی می تواند به طور مؤثری زمان و هزینه های مطالعه های آزمایشگاهی را کاهش دهد.

سپاسگزاری

مقاله حاضر بخشی از نتایج طرح پژوهشی است که بودجه آن از طرف دانشگاه آزاد اسلامی واحد کنگاور تأمین شده است که بدین ترتیب تشکر و قدردانی می گردد.

ضدمیکروبی که علیه این اهداف تولید می شوند، با توجه به عدم مشابهت این پروتئین ها با پروتئین های موجودات غیرهدف، ترکیب هایی کامل اختصاصی بوده و علاوه بر کنترل مؤثر پاتوژن برای محیط زیست و موجودات غیرهدف نیز اثر مضرى ندارند. اهداف شناسایی شده به دو گروه طیف وسیع و اختصاصی تقسیم می شوند. اهداف طیف وسیع موارد مناسب تری برای هدف گیری هستند. از آنجای که اهداف طیف وسیع پاتوژن های بیش تری را هدف گیری می کنند، اهداف مؤثر تری برای مبارزه به نظر می رسند و با هدف گیری آن ها با یک بار صرف هزینه می توان گیاه را نسبت به طیف وسیعی از پاتوژن ها مقاوم کرد. همچنین سمومی که علیه این اهداف تولید می شوند طیف وسیعی از پاتوژن ها را از بین می برند. از طرفی دیگر، سمومی که علیه اهداف بالقوه تولید می شوند، سمومی بی خطر برای انسان، موجودات غیرهدف و محیط زیست هستند زیرا پروتئین هدف آن ها تنها در پاتوژن های ویژه وجود دارد.

با استفاده از روش های بیوانفورماتیکی مبتنی بر همولوژی می توان فرآیند تولید گیاهان تراریخت مقاوم به بیماری ها و آفات و هم چنین فرآیندهای تولید سموم شیمیایی را به طور قابل توجهی کوتاه کرد و هزینه های آن را کاهش داد. با استفاده از این روش ها همچنین می توان ایمنی سموم تجاری تولید شده را بررسی و ارزیابی کرد. مهم ترین مزیت استفاده از روش های بیوانفورماتیکی و محاسبه ای، پیش بینی اهدافی در پاتوژن هاست که برای بقاء پاتوژن ها ضروری بوده و در عین حال هیچ شباهتی با میزبان و موجودات غیرهدف ندارند و هدف قرار دادن آن ها اثر سوء بر میزبان و سایر موجودات غیرهدف نخواهد داشت. در صورتی که سایر اهداف که با روش های غیر بیوانفورماتیکی انتخاب می شوند ممکن است برای بقاء پاتوژن ضروری نبوده و از طرفی دارای توالی مشابه در میزبان و موجودات غیرهدف باشند. هدف گیری چنین پروتئین هایی باعث اثرهای مضر بر میزبان، انسان و موجودات

منابع

1. Agrios, G.N., Plant pathology, 2005. Elsevier Academic Press Burlington, MA.
2. Altschul, S.F., et al., Basic local alignment search tool. Journal of molecular biology, 1990. 215: p. 403-410.
3. Anishetty, S., et al., Potential drug targets in Mycobacterium tuberculosis through metabolic pathway analysis. Computational biology and chemistry, 2005. 29: p. 368-378.
4. Buell, C.R., et al., The complete genome sequence of the Arabidopsis and tomato pathogen *P. syringae* pv. tomato DC3000. Proceedings of the National Academy of Sciences, 2003. 100: p. 10181-10186.
5. Butt, A.M., et al., Correction: Comparative Genomics Analysis of Mycobacterium ulcerans for the Identification of Putative Essential Genes and Therapeutic Candidates. PloS one, 2013. 8.
6. Capra, E.J. and Laub, M.T. The Evolution of Two-Component Signal Transduction Systems. Annual review of microbiology, 2012. p. 66: 325.
7. Chen, L., et al., update: toward the genetic diversity and molecular evolution of bacterial virulence factors. Nucleic acids research, 2012. gkr989.
8. Finn, R.D., et al., Pfam: the protein families database. Nucleic acids research, 2012. gkt1223.
9. Hirano, S.S. and Upper, C.D., Population biology and epidemiology of *Pseudomonas syringae*. Annual review of phytopathology, 1990. 28: 155-177.
10. Horner, D.S., et al., Bioinformatics approaches for genomics and post genomics applications of next-generation sequencing. Briefings in bioinformatics, 2009. 11: p. 181-197.
11. Huthmacher, C., et al., Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. BMC systems biology, 2010. 4: 1.
12. Jones, P., et al., InterProScan 5: genome-scale protein function classification. Bioinformatics, 2014. 30: p. 1236-1240.
13. Kanehisa, M. and Goto, S., KEGG: kyoto encyclopedia of genes and genomes. Nucleic acids research, 2000. 28: p. 27-30.
14. Koch, L., Microbiome: Shaping the gut microbiome. Nature Reviews Microbiology, 2005. p. 13: 4-4.
15. Lewis, K., Platforms for antibiotic discovery. Nature reviews Drug discovery, 2013. 12: p. 371-387.
16. Luo, H., et al., DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic acids research, 2014. 42: D574-D580.
17. Marchler-Bauer, A., et al., CDD: NCBI's conserved domain database. Nucleic acids research, 2014. gku1221.
18. Mochida, K. and Shinozaki, K., Genomics and bioinformatics resources for crop improvement. Plant and Cell Physiology, 2010. 51: p. 497-523.
19. Nancy, Y.Y., et al., PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics, 2010. 26: p. 1608-1615.
20. NCBI, R.C., Database resources of the National Center for Biotechnology Information. Nucleic acids research, 2015. 43: D6.
21. Parker, J., et al., Quantitative proteomics of tomato defense against *P. syringae* infection. Proteomics, 2013. p. 13: 1934-1946.
22. Pevsner, J., Bioinformatics and functional genomics, 2015. John Wiley & Sons.
23. Raman, K., et al., targetTB: a target identification pipeline for Mycobacterium tuberculosis through an interactome, reactome and genome-scale structural analysis. BMC systems biology, 2008. 2: 1.
24. Sahu, S.S., et al., Predicting genome-scale Arabidopsis-*P. syringae* interactome using domain and interolog-based approaches. BMC bioinformatics, 2014. 15: S13.
25. Sakharkar, K.R., et al., A novel genomics approach for the identification of drug targets in pathogens, with special reference to *Pseudomonas aeruginosa*. In silico biology, 2004. 4: p. 355-360.
26. Sarkar, M., et al., In silico quest for putative drug targets in *Helicobacter pylori* HPAG1: molecular modeling of candidate enzymes from lipopolysaccharide biosynthesis pathway. Journal of molecular modeling, 2012. p. 18: 1855-1866.
27. Shanmugham, B. and Pan, A., Identification and characterization of potential therapeutic candidates in emerging human pathogen Mycobacterium abscessus: a novel hierarchical in silico approach, 2013. PloS one, 8.

28. Shanmugham, B. and Pan, A., Identification and characterization of potential therapeutic candidates in emerging human pathogen *Mycobacterium abscessus*: a novel hierarchical in silico approach, 2012. PloS one, 8: e59126.
29. Strange, R.N. and Scott, P.R., Plant disease: a threat to global food security. *Phytopathology*, 2005. 4: 3.
30. Walter, J. and Ley, R., The human gut microbiome: ecology and recent evolutionary changes. *Annual review of microbiology*, 2011. 65: 411-429.
31. Yaish, M.W., et al., Genetic mapping of quantitative resistance to race 5 of *P. syringae* pv. *phaseolicola* in common bean. *Euphytica*, 2006. 152: p. 397-404.
32. Yu, C.-S., et al., CELLO2GO: a web server for protein subcellular localization prediction with functional gene ontology annotation. 2014.