Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

مقاله پژوهشی

مجله انفورماتیک سلامت و زیست پزشکی
مرکز تحقیقات انفورماتیک پزشکی
دوره هفتم شماره سوم. ص ۳۳۶-۳۲۶، ۱۳۹۹

مقایسه کمی زوج درخت‌های حاصل از درخت‌های فیلوژنتیکی ژن و پروتئین برای آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا، درخت‌های 5SrRNA و تاکسونومی در گونه‌های منتخب باکتریایی

الهام ترحمی[۱]، حسین فهیمی [۲]، محمد تقی‌زاده[۳]*

• **دریافت مقاله:** ۹۸/۹/۸ • **پذیرش مقاله:** ۹۸/۱۰/۲۱

مقدمه: خانواده پروتئینی FAD-FR دارای کوفاکتور FAD می‌باشد. یکی از آنزیم‌های مهم این خانواده سولفیت ردوکتاز فلاووپروتئینی آلفا می‌باشد. هدف این مطالعه توجه به کاربردهای این آنزیم در بیوتکنولوژی و صنعت این آنزیم در ۱۹ گونه مشخص در ارتباط با تکامل است.

روش: توالی‌های ژن و پروتئین سولفیت ردوکتاز فلاووپروتئینی آلفا، توالی‌های 5SrRNA و درخت تاکسونومی برای ۱۹ گونه انتخابی باکتریایی استخراج شد. سپس به مقایسه کمی درخت‌های فیلوژنتیکی توالی‌های 5SrRNA، ژن و پروتئین با یکدیگر و با درخت تاکسونومی پرداخته شد. درخت‌های فیلوژنتیکی به کمک نرم‌افزار Mega7 و با استفاده از الگوریتم Neighbor joining رسم شد و درخت تاکسونومی با استفاده از Taxonomy Browser NCBI استخراج شد.

نتایج: با مقایسه زوج درخت‌های مربوطه، درصد گونه‌های هم ارز و متوسط امتیاز هم ارزی گونه‌ها در هر زوج درخت محاسبه شد. زوج درخت ژن-پروتئین بالاترین امتیاز را در هر دو کمیت به خود اختصاص داد. در مقایسه درخت تاکسونومی با سه درخت دیگر، زوج درخت ژن-تاکسونومی بالاترین درصد را در متوسط امتیاز هم ارزی به دست آورد و زوج درخت پروتئین-تاکسونومی بالاترین درصد گونه‌های هم ارز را کسب کرد.

نتیجه‌گیری: بر اساس پژوهش حاضر، می‌توان دریافت که مناسب‌ترین جایگزین برای هرکدام از درخت‌هایی که در این پژوهش برای بررسی روابط تکاملی محاسبه شده است کدام است. به بیان دیگر، کدام درخت تکاملی را می‌توان جایگزین درخت تکاملی دیگری کرد.

کلیدواژه‌ها: درخت فیلوژنتیکی، درخت تاکسونومی، آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا، ژن 5SrRNA، زوج درخت

• **ارجاع:** ترحمی الهام، فهیمی حسین، تقی‌زاده محمد. مقایسه کمی زوج درخت‌های حاصل از درخت‌های فیلوژنتیکی ژن و پروتئین برای آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا، درخت‌های 5SrRNA و تاکسونومی در گونه‌های منتخب باکتریایی. مجله انفورماتیک سلامت و زیست‌پزشکی ۱۳۹۹؛ ۷(۳): ۳۶-۳۲۶.

۱. کارشناس ارشد ژنتیک مولکولی، گروه ژنتیک، دانشکده علوم و فناوری‌های نوین، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران
۲. دکتری ژنتیک مولکولی، استادیار، گروه ژنتیک، دانشکده علوم و فناوری‌های نوین، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران
۳. دکتری بیوانفورماتیک، استاد مدعو، گروه بیوتکنولوژی، دانشکده علوم و فناوری‌های نوین، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران

* **نویسنده مسئول:** محمد تقی‌زاده

آدرس: تهران، خیابان شریعتی، خیابان خاقانی، دانشگاه علوم پزشکی آزاد اسلامی تهران، دانشکده علوم و فناوری‌های نوین، گروه بیوتکنولوژی

• **شماره تماس:** ۰۹۸۲۲۰۰۶۶۶۰+ • **Email:** mtaghizadeh@Alumni.ut.ac.ir

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

مقدمه

یکی از آنزیم‌های مهم در خانواده فلاوین آدنین دی نوکلئوتید- یکی از آنزیم‌های مهم در خانواده فلاوین آدنین دی نوکلئوتید- فرودوکسین ردوکتاز (FAD-FR)، سولفیت ردوکتاز فلاووپروتئینی آلفا می‌باشد. این آنزیم از NADPH به عنوان کوفاکتور استفاده می‌کند. سولفیت ردوکتازها آنزیم‌هایی هستند که در متابولیسم‌های سولفور شرکت می‌کنند [۱] که طی این متابولیسم‌ها سولفیت به سولفید، هیدروژن و آب تبدیل می‌شود [۲،۳]. سولفیت ردوکتاز که متعلق به خانواده اکسیدو ردوکتاز می‌باشد، در آرکی‌ها، باکتری‌ها، قارچ‌ها و گیاهان یافت می‌شود [۴-۶]. یکی از کاربردهای علمی و صنعتی آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا سنتز نانو ذرات طلا به کمک این آنزیم می‌باشد [۷]. خانواده پروتئینی FAD-FR دارای کوفاکتور FAD می‌باشد. کوفاکتور FAD اهمیت ویژه‌ای در بین موجودات زنده دارد و نقش حیاتی در انتقال انرژی دارد و محصول تراکم ریبوفلاوین و آدنوزین دی فسفات می‌باشد [۸].

در طبقه‌بندی زیست‌شناختی، هر جانداری در جایگاه ویژه‌ای از نظر دسته‌بندی قرار می‌گیرد. به هر سطح از دسته‌بندی، یک آرایه (Taxon) یا زیر مجموعه گفته می‌شود [۹]. تاکسونومی باکتری‌ها در اواخر قرن ۱۹ آغاز شد؛ زمانی که ویژگی‌های فنوتیپی، محیط کشت، مورفولوژی، اندازه کلونی و رنگ واکنش‌های شیمیایی مبنای اصلی در تاکسونومی باکتری‌ها به حساب می‌آمد. به طور کلی، در زیست‌شناسی تاکسونومی به معنای استفاده از ویژگی‌های مختلف موجودات زنده برای توسعه دسته‌بندی آن‌ها است [۱۰]. فیلوژنی بررسی تکامل گروهی از موجودات زنده است که از نظر ژنتیکی با هم مرتبط هستند. در مطالعات فیلوژنی معمولاً از توالی‌های ژنی یا پروتئینی استفاده می‌شود که با یکدیگر هومولوگ هستند. رابطه هومولوژی به معنای داشتن جد مشترک است. حال اگر نتایج مطالعات فیلوژنی به صورت یک درخت ارائه شود، به چنین درختی، درخت فیلوژنی یا درخت فیلوژنتیکی گفته می‌شود [۱۱].

ژن‌های rRNA برای زنده ماندن همه موجودات ضروری می‌باشند و همچنین توالی بازهای آلی در این ژن‌ها به شدت حفظ شده است. 5SrRNA یک مولکول نسبتاً کوچک RNA است و با توجه به ساختار و عملکرد بیولوژیکی آن، مطالعات مختلفی انجام شده است. این گونه RNA یک عنصر جامع از ریبوزوم‌های موجود در تمام موجودات زنده است. توالی نوکلئوتید 5SrRNA به شدت حفظ شده است و برای سالیان طولانی به عنوان یکی از توالی‌های مطرح در مطالعه تکامل مولکولی مورد استفاده قرار گرفته است [۱۲،۱۳].

در پژوهشی که سال ۲۰۰۱ در آلمان روی ۵' آدنوزین فسفوسولفات ردوکتاز [APS] انجام شد، به مقایسه فیلوژنتیکی APS ردوکتاز با 16SrRNA و dsrAB پرداخته شده است. به کمک PCR قطعه‌ای از ژن APS ردوکتاز برای مطالعه فیلوژنتیکی استخراج شد. درخت‌های مورد استفاده براساس محاسبه ماتریکس فاصله با استفاده از الگوریتم‌های اتصال همسایگی (NJ)، Maximum likelihood و پارسیمونی به دست آمد [۱۴]. هدف در این پژوهش، تعیین وضعیت انتقال جانبی ژن (Lateral gene transfer) در ژن‌های APS ردوکتاز در میکروارگانیسم‌های احیا کننده سولفات بوده است. مقایسه درخت فیلوژنتیکی که با استفاده از APS ردوکتاز و dsrAB در این پژوهش به دست آمده است، نشان می‌دهد که انتقال جانبی ژن در باکتری‌های ترموفیلیک (گرمادوست) و گرم مثبت SRB، زیاد اتفاق می‌افتد [۱۵].

در پژوهشی دیگر به مقایسه درخت فیلوژنتیکی ژن 16SrRNA و پروتئین RecA از گونه‌های یکسان پرداخته شده است. در این پژوهش آنالیز تکاملی ژن RecA با روش‌های فیلوژنتیکی مولکولی صورت گرفته است. برای ترسیم درخت فیلوژنتیکی از توالی‌های باکتریایی پروتئین RecA و توالی‌های 16SrRNA از جنس و گونه یکسان استفاده شد و برای گونه‌هایی که توالی شناخته شده نداشتند، از توالی گونه‌های نزدیک استفاده شد. همچنین در مواردی که توالی کامل برای ژن 16SrRNA وجود نداشت، از توالی‌های ناقص (Partial) در آن گونه استفاده شد. پس از هم ترازسازی چندگانه توالی‌ها، با استفاده از الگوریتم اتصال همسایگی (Neighbor joining) برای این توالی‌ها درخت فیلوژنتیکی رسم شد [۱۶]. در این پژوهش نتایج نهایی به این صورت بود که بیشترین تفاوت بین زوج درخت RecA و 16SrRNA مربوط به پروتئوباکترها بوده است. طبق این مطالعات مشخص شد پروتئین RecA می‌تواند برای مطالعات مولکولی مفید باشد؛ لذا از نتایج دیگری که در این پژوهش به دست آمد این بود که بررسی زوج درخت‌های فیلوژنی در گونه‌های خاص می‌تواند به ما در شناسایی ژن‌های مفید برای مطالعات تکاملی کمک کند [۱۷].

در این پژوهش نیز به علت کاربردهای آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا در بیوتکنولوژی و صنعت که جزء خانواده پروتئینی FAD-FR به حساب می‌آید، به مقایسه

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

درخت‌های فیلوژنتیکی توالی‌های 5SrRNA، ژن و پروتئین این آنزیم با یکدیگر و با درخت تاکسونومی پرداخته شد. در واقع، یکی از کاربردهای صنعتی آنزیم سولفیت ردوکتاز فلاوودوکسین آلفا، سنتز نانو ذرات طلا به کمک این آنزیم می‌باشد. یون‌های طلا در حضور این آنزیم کاهش می‌یابد که منجر به تشکیل هیدروسول طلای پایدار می‌شود. در چند دهه اخیر، مطالعات بر روی آنزیم سولفات ردوکتازها بسیار گسترش یافته است. از آنجایی که محصولات سولفات ردوکتاز میکروبی نقش مهمی را در محیط‌زیست، بیوتکنولوژی و صنعت ایفا می‌کنند، مدل‌سازی سولفور بحث مهمی در تحقیقات می‌باشد. سنتز بیومولکول‌های حاوی گوگرد از نظر صنعتی پراهمیت هستند، بنابراین اغلب تحقیقات روی کاربردهای بیوتکنولوژی متمرکز شده است که براساس ارزش بازار، مهم‌ترین این مولکول‌ها متیونین می‌باشد. در پژوهش حاضر، درخت‌های فیلوژنتیکی برای گونه‌های ثابتی از آنزیم سولفیت ردوکتاز فلاوودوکسین آلفا رسم شد و سپس به مقایسه کمی آن‌ها پرداخته شده است [۱۸].

روش

استخراج توالی‌های ژنی و پروتئینی سولفیت ردوکتاز فلاوودوکسین آلفا و توالی‌های 5SrRNA برای گونه‌های باکتریایی

با استفاده از پروسایت، خانواده پروتئینی FAD-FR با شناسه پروسایتی PS51384 شناسایی شد [۱۹] و توالی‌های باکتریایی موجود در این خانواده پروتئینی که مربوط به آنزیم سولفیت ردوکتاز فلاوودوکسین [جزء آلفا] بود، جداسازی شد. تعداد گونه‌ها و سوش‌های باکتریایی که در این مرحله به دست آمد، ۵۸ مورد بود که پس از استخراج توالی‌های پروتئینی و ژنی برای این آنزیم، همچنین استخراج توالی‌های 5SrRNA برای بزرگ‌ترین زیرمجموعه ممکن از این ۵۸ مورد و حذف سوش‌های مشترک، ۱۹ گونه باقی ماند. از عدد ۵۸ تا ۱۹ علت حذف ۳۹ گونه [یا سوش]، رسیدن به گونه‌های باکتریایی بود که برای آنزیم سولفیت ردوکتاز فلاوودوکسین دارای توالی پروتئینی و ژنی و همچنین برای گونه‌های باکتریایی مربوطه دارای توالی 5SrRNA باشد. توالی‌های پروتئینی از پایگاه داده یونی پرات مربوط به وب سایت اکسپسی (Expasy) (https://www.expasy.org/) و توالی‌های ژنی از پایگاه داده نوکلئوتید [Nucleotide] متعلق NCBI (https://www.ncbi.nlm.nih.gov/) استخراج شد.

استخراج درخت‌های فیلوژنتیکی پروتئین، ژن و 5SrRNA و تاکسونومی برای ۱۹ گونه باکتریایی منتخب

توالی‌های ژن، پروتئین و 5srRNA از ۱۹ جنس و گونه یکسان انتخاب شد، سپس به کمک نرم‌افزار Mega7 و با استفاده از الگوریتم اتصال همسایگی (joining Neighbor) برای این سه توالی درخت فیلوژنتیکی رسم شد [۲۰]. رسم درخت‌ها با الگوریتم Maximum likelihood نیز صورت گرفت، با مقایسه بین این دو الگوریتم مشخص شد تفاوت درخت‌ها با الگوریتم NJ و Maximum likelihood بسیار اندک است. بر اساس مطالعه و جستجوی مقالات مختلف مشخص شد که الگوریتم NJ یکی از بهترین الگوریتم‌های موجود برای رسم درخت فیلوژنی می‌باشد.

برای مقایسه این درخت‌ها دو به دو روبه‌رو یکدیگر قرار داده شده‌اند و گونه‌های مشابه با خط به یکدیگر وصل شدند. برای به دست آوردن ماکزیمم شباهت بین درخت‌ها، از طریق نرم‌افزار Mega7 چرخش زیر شاخه‌ها (Flip) انجام شد. در این حالت هر کجا یک خط موازی به دست آمده بود، به عنوان یک موقعیت صحیح در دو درخت شمارش شد.

سپس درخت تاکسونومی مربوط به ۱۹ گونه انتخابی با استفاده از ابزار مرورگر تاکسونومی (Taxonomy Browser) متعلق به وب‌سایت https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi NCBI استخراج شد [۲۱] و در ادامه به مقایسه دو به دوی درخت تاکسونومی با درخت های فیلوژنتیکی پروتئین، ژن و 5SrRNA پرداخته شد. در ادامه، این مقایسه ها با استفاده از روابطی که در بخش بعدی مطرح شده است، کمی‌سازی شد.

محاسبات کمی به منظور مقایسه درخت‌های فیلوژنتیکی پروتئین، ژن، 5SrRNA و تاکسونومی برای ۱۹ گونه باکتریایی منتخب

برای مقایسه کمی داده‌ها دو شیوه مورد استفاده قرار گرفت: در روش اول، تعداد خطوط موازی مشاهده شده ناشی از هم ارزی گونه‌ها در دو درخت، تقسیم بر خطوط موازی هم ارزی مورد انتظار شد و مقدار در عدد صد ضرب شد. این محاسبه، درصد گونه‌های هم ارز در دو درخت نامیده شد (رابطه ۱). همان‌طور که اشاره شد، زوج درخت‌ها به شکل مقایسه‌ای به دست آمد و بین گونه‌های مشابه یک خط رسم شد. سپس خطوط موازی به دست آمده شمارش شد. با توجه به این که در هر درخت ۱۹

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

گونه وجود داشت و تعداد کل خطوط رسم شده بین گونه‌های یکسان برابر ۱۹ بود؛ لذا خطوط موازی هم ارزی مورد انتظار برابر ۱۹ در نظر گرفته شد.

در روش دوم به هر یک از خطوطی که بین دو درخت مورد مقایسه رسم شده بود یک امتیاز نسبت داده شد. امتیاز خطوط موازی ۱۸ در نظر گرفته شد و به ازای هر گام (یک گونه در درخت) پایین رو یا بالا رو یک امتیاز کم شد. در این حالت، کمترین امتیاز به گونه‌ای تعلق می‌گیرد که یکی بالاترین انشعاب یک درخت و دیگری پایین‌ترین انشعاب درخت مقابل را ایجاد کند. امتیاز چنین گونه‌ای برابر یک خواهد بود و این امتیاز نشان دهنده گونه‌ای است که عامل بیشترین تفاوت بین دو درخت تکاملی است. به این ترتیب کم‌ترین فاصله که متعلق به هر یک از خطوط موازی است، برابر امتیاز ۱۸ به دست می‌آید و این امتیاز نشان دهنده گونه‌ای است که عامل ایجاد بیشترین شباهت بین دو درخت می‌باشد. با استفاده از امتیازدهی به خطوط که در بالا توضیح داده شد، برای هر زوج درخت که شامل پروتئین-ژن، پروتئین-5SrRNA، ژن-5SrRNA، پروتئین-تاکسونومی، ژن-تاکسونومی و 5SrRNA-تاکسونومی می‌شود، رابطه ۲ محاسبه شد. در این رابطه، صورت کسر مجموع امتیاز هم ارزی همه خطوط برای همه گونه‌های زوج درخت و مخرج کسر مجموع امتیاز هم ارزی مورد انتظار می‌باشد و نام رابطه امتیاز هم ارزی گونه‌ها در زوج درخت در نظر گرفته شده است.

رابطه [۱]:

$$\text{در صد گونه های هم ارز در زوج درخت} = \frac{\text{تعداد خطوط موازی مشاهده شده ناشی از هم ارزی گونه ها در زوج درخت}}{\text{خطوط موازی هم ارزی مورد}} \times ۱۰۰$$

رابطه [۲]:

$$\text{متوسط امتیاز هم ارزی گونه ها در زوج درخت} = \frac{\text{مجموع امتیاز هم ارزی همه خطوط برای همه گونه های زوج درخت}}{\text{مجموع امتیاز هم ارزی مورد انتظار}} \times ۱۰۰$$

نتایج

با استفاده از پایگاه داده پروسایت توالی آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا از خانواده FAD-FR استخراج شد، تعداد گونه‌ها و سوش‌های باکتریایی که در این مرحله به دست آمد، ۵۸ مورد بود که پس از حذف سوش‌های مشترک، ۱۹ گونه باقی ماند.

برای مقایسه درخت‌های پروتئین، ژن، 5SrRNA و تاکسونومی میان گونه‌های منتخب که هدف اصلی در این پژوهش به حساب می‌آید، در قدم اول به مقایسه زوج درخت‌های پروتئین-ژن، پروتئین-5SrRNA و ژن-5SrRNA پرداخته شد (شکل ۱ قسمت الف، ب و ج).

اولین زوج درخت، پروتئین-ژن می‌باشد. چنان که در شکل ۱-الف قابل مشاهده است، هم ارزی گونه‌های مشابه در این زوج درخت بسیار زیاد است. چنان که تعداد خطوط موازی در این زوج درخت به عدد ۱۴ می‌رسد؛ لذا می‌توان پذیرفت که زوج درخت ژن-پروتئین برای آنزیم سولفیت ردوکتاز فلاوو پروتئین آلفا کامپوننت، علارغم وجود تفاوت‌هایی دارای شباهت قابل توجه هستند. البته باید توجه کرد که توپولوژی دو درخت که در شکل قابل مشاهده است، دارای تفاوت‌هایی بیش از آنچه خطوط غیرموازی به ما نشان می‌دهد، است؛ اما تفاوت‌هایی که در توپولوژی درخت‌ها وجود دارد و با خطوط غیر موازی قابل نمایش و اندازه‌گیری نیست، نسبتاً اندک است. همان‌طور که مشاهده می‌شود، اغلب خطوط در این زوج درخت به صورت موازی می‌باشد و تنها گونه‌های Sodalis glossinidius، Photorhabdus laumondii، Baumannia cicadellinicola، Buchnera aphidicola و Zymomonas mobilis دارای خطوط غیرموازی می‌باشند؛ لذا می‌توان گفت، این گونه‌ها مهم‌ترین عامل ایجاد تفاوت بین دو درخت در زوج درخت ژن-پروتئین به حساب می‌آیند.

چنان‌که در شکل ۱-ب در مقایسه با شکل ۱-الف قابل مشاهده است، هم ارزی گونه‌های مشابه در زوج درخت ژن-

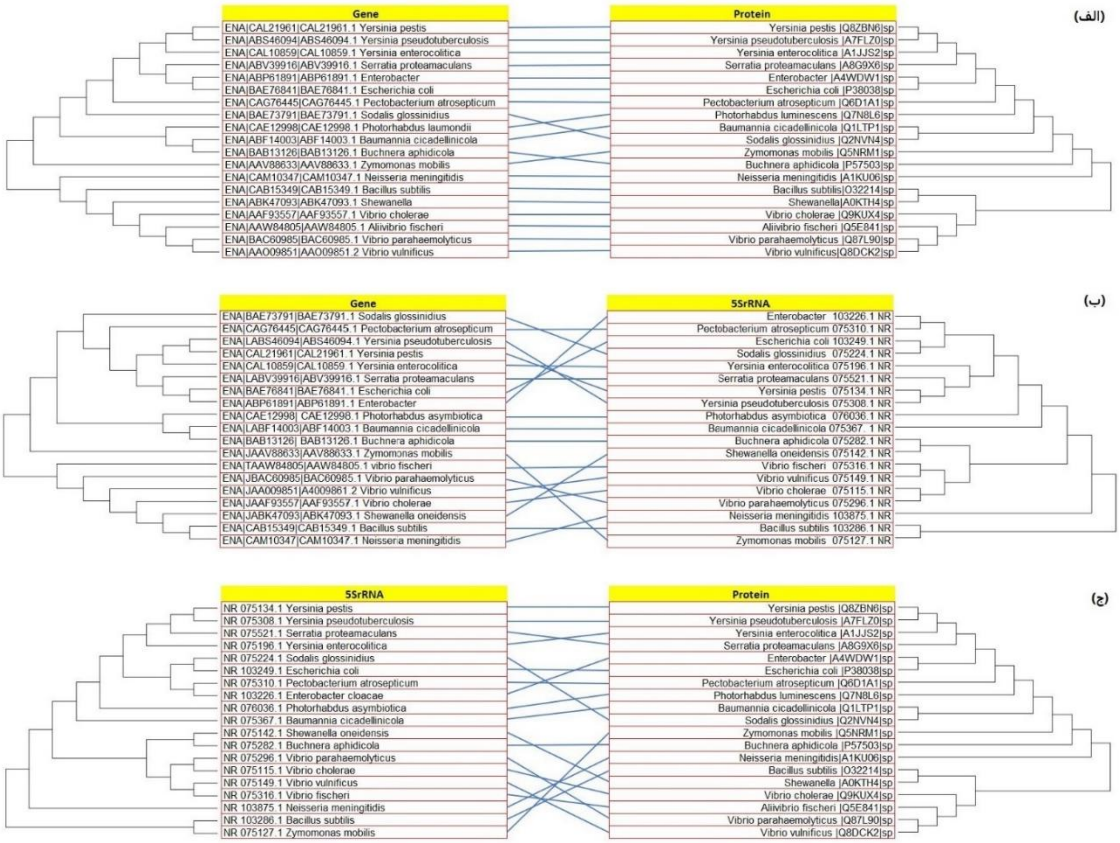
www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

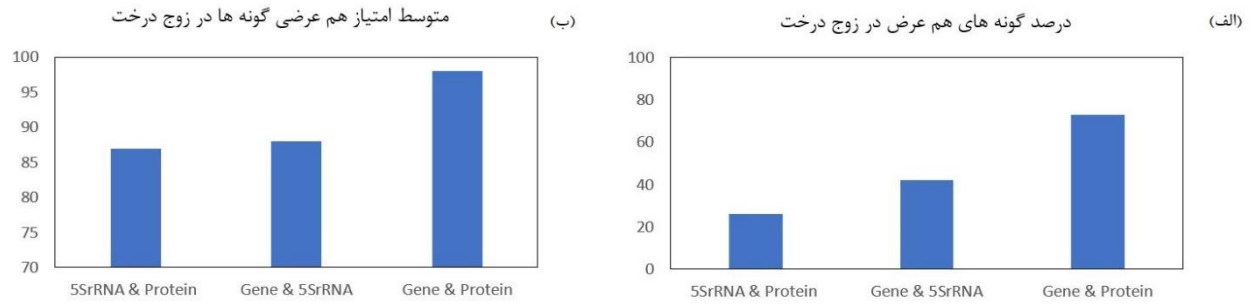
5SrRNA کمتر از زوج درخت ژن–پروتئین می‌باشد. چنان که تعداد خطوط موازی در این زوج درخت به عدد ۸ می‌رسد؛ لذا می‌توان پذیرفت که شباهت دو درخت در زوج درخت ژن–پروتئین بیشتر از دو درخت در زوج درخت ژن–5SrRNA می‌باشد.

همان‌طور که شکل ۱–ج نشان می‌دهد، هم ارزی گونه‌های مشابه در زوج درخت پروتئین–5SrRNA به زوج درخت ژن–5SrRNA نزدیک بوده و از زوج درخت ژن–پروتئین فاصله قابل توجهی دارد. چنان که تعداد خطوط موازی در این زوج درخت به عدد ۵ می‌رسد.



شکل ۱: الف، زوج درخت پروتئین–ژن، ب، پروتئین–5SrRNA، ج، ژن–5SrRNA

مقایسه زوج درخت‌های ذکر شده به صورت دو به دو انجام گرفته است و گونه‌های مشابه با خطوط به یکدیگر وصل شده است.



نمودار ۱: الف، درصد گونه‌های هم ارز در زوج درخت‌های ژن–پروتئین، ژن–5SrRNA و پروتئین–5SrRNA، ب، امتیاز هم ارزی گونه‌ها در زوج درخت‌های ذکر شده

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

نمودار۱-الف درصد گونه‌های هم ارز را در هر زوج درخت نشان می‌دهد. چنان‌که در نمودار ۱ قابل مشاهده است، هم ارزی گونه‌های مشابه در زوج درخت ژن-پروتئین بسیار زیاد است. چنان که این امتیاز در زوج درخت مربوطه به ۷۳ درصد می‌رسد؛ لذا می‌توان پذیرفت که زوج درخت ژن-پروتئین برای آنزیم سولفیت ردوکتاز فلاوو پروتئین آلفا کامپوننت، دارای شباهت قابل توجهی می‌باشد. درصد گونه‌های هم ارز در زوج درخت ژن-5SrRNA، ۴۲ می‌باشد که این درصد نسبت به زوج درخت ژن-پروتئین کمتر است. این کمیت در زوج درخت پروتئین-5SrRNA، ۲۶ درصد می‌باشد که کمترین امتیاز را دارد؛ لذا می‌توان پذیرفت که شباهت زوج درخت‌ها در ژن-پروتئین بیشترین مقدار را به خود اختصاص می‌دهد و رتبه‌های بعدی به ترتیب مربوط به زوج درخت‌های ژن-5SrRNA و پروتئین-5SrRNA می‌باشد.

نمودار ۱-ب، متوسط امتیاز هم ارزی گونه‌ها را در هر زوج درخت نشان می‌دهد. چنان که در این نمودار مشاهده می‌شود، امتیاز هم ارزی گونه‌های مشابه در زوج درخت ژن-پروتئین بسیار زیاد است و ۹۸ درصد می‌باشد. این امتیاز در زوج درخت ژن-5SrRNA و پروتئین-5SrRNA به ترتیب ۸۸ و ۸۷ درصد شد که نشان دهنده آن است که دو کمیت ارزیابی ارائه شده برای زوج درخت‌ها به طور کلی با یکدیگر همخوانی دارد؛ اما در کمیت دوم مقادیر به یکدیگر نزدیک ترند.

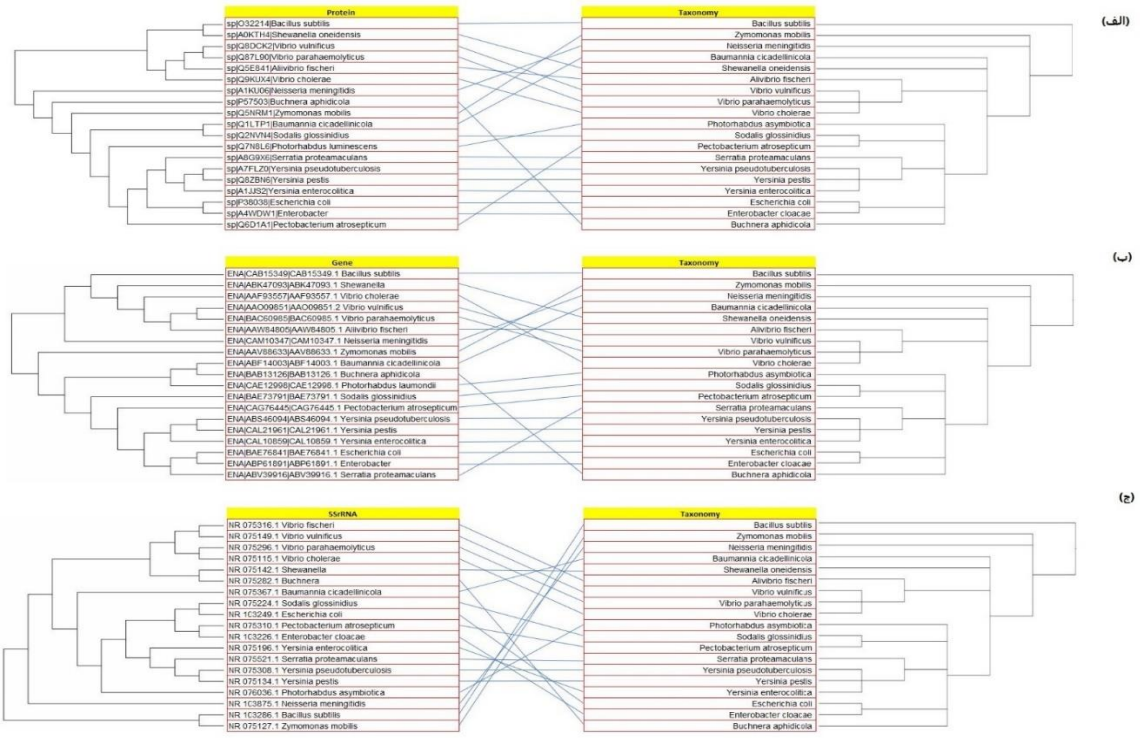
یکی از مهم‌ترین کاربردهایی که نتایج به دست آمده تا این بخش برای ما فراهم می‌کند، این است که اگر مجبور شویم درخت‌های تکاملی را با هم جایگزین کنیم این مسئله شامل چه مقدار تقریب می‌شود. به طور مثال، اگر در پژوهشی هدف مطالعه روابط تکاملی در سطح یک ژن خاص باشد و توالی‌های ژنی به سادگی و به طور کامل برای گونه‌های موردنظر در دسترس نباشد به جای آن‌ها می‌توان کدام درخت رسم شود درخت پروتئینی یا درخت 5SrRNA؟ بر اساس نتایجی که در نمودار ۱ بخش الف و ب قابل مشاهده است، جواب درخت پروتئینی است، حال اگر هدف به دست آوردن درخت 5SrRNA باشد؛ اما توالی‌های مورد نیاز برای گونه‌های انتخابی در دسترس نباشد پاسخ چه می‌شود؟ بر اساس نتایجی که در نمودار ۱ قابل مشاهده است، درخت ژن به درخت 5SrRNA نزدیک‌تر است. البته این پاسخ‌ها به طور قطعی درمورد پروتئین/ژن سولفیت ردوکتاز فلاوو پروتئین آلفا کامپوننت باکتریایی صحیح هستند؛ اما اگر پروتئین/ژن مورد دیگری باشد، این پاسخ‌ها می‌تواند به صورت فرضی مطرح باشد؛ اما وقتی مسئله پیچیده‌تر می‌شود که بخواهیم درخت تاکسونومی هم مطرح شود. در این صورت پاسخ‌ها چگونه‌اند؟ این پرسشی است که تلاش شده است در ادامه نتایج به آن پاسخ داده شود.

برای مقایسه درخت‌های پروتئین، ژن و 5SrRNA با درخت تاکسونومی میان گونه‌های منتخب، ابتدا به بررسی زوج درخت پروتئین-تاکسونومی پرداخته شده است. چنان که در شکل ۲-الف قابل مشاهده است، هم ارزی گونه‌های مشابه در این زوج درخت کمی بیشتر از زوج درخت ژن-تاکسونومی می‌باشد؛ اما این مقدار هم ارزی با زوج درخت 5SrRNA-تاکسونومی فاصله بیشتری دارد. چنان که تعداد خطوط موازی در این زوج درخت به عدد ۸ می‌رسد. در حالی که طبق شکل ۲-ب تعداد خطوط موازی در زوج درخت ژن-تاکسونومی به عدد ۷ می‌رسد. هم ارزی گونه‌های مشابه در زوج درخت ژن-تاکسونومی به زوج درخت پروتئین-تاکسونومی نزدیک بوده و از زوج درخت 5SrRNA-تاکسونومی فاصله قابل‌توجهی دارد. چنان که در شکل ۲-ج قابل مشاهده است، تعداد خطوط موازی در زوج درخت 5SrRNA-تاکسونومی به عدد ۴ می‌رسد؛ لذا می‌توان پذیرفت که شباهت دو درخت در زوج درخت پروتئین-تاکسونومی و ژن-تاکسونومی بیشتر از دو درخت در زوج درخت 5SrRNA-تاکسونومی می‌باشد. تنها ۴ گونه Yersinia pseudotuberculosis، Serratia proteamaculans، Yersinia pestis و Shewanella با خطوط موازی به یکدیگر وصل شده‌اند که نشان دهنده آن است که علی‌رغم وجود تفاوت‌هایی دارای شباهت‌هایی نیز هستند.

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021



شکل ۱: الف، زوج درخت پروتئین–تاکسونومی، ب، ژن–تاکسونومی، ج، 5SrRNA–تاکسونومی

مقایسه زوج درخت‌های ذکر شده به صورت دو به دو انجام گرفته است و گونه‌های مشابه با خطوط به یکدیگر وصل شده است.

نمودار ۲–الف درصد گونه‌های هم ارز را در هر زوج درخت نشان می‌دهد، چنانچه در نمودار قابل مشاهده است، درصد گونه‌های هم ارز در زوج درخت‌های پروتئین–تاکسونومی، ژن–تاکسونومی و 5SrRNA–تاکسونومی به ترتیب ۴۲، ۳۶ و ۲۱ می‌باشد؛ لذا می‌توان پذیرفت که درصد گونه‌های هم ارز در زوج درخت پروتئین–تاکسونومی به زوج درخت ژن–تاکسونومی نزدیک بوده و از زوج درخت 5SrRNA–تاکسونومی فاصله قابل‌توجهی دارد.

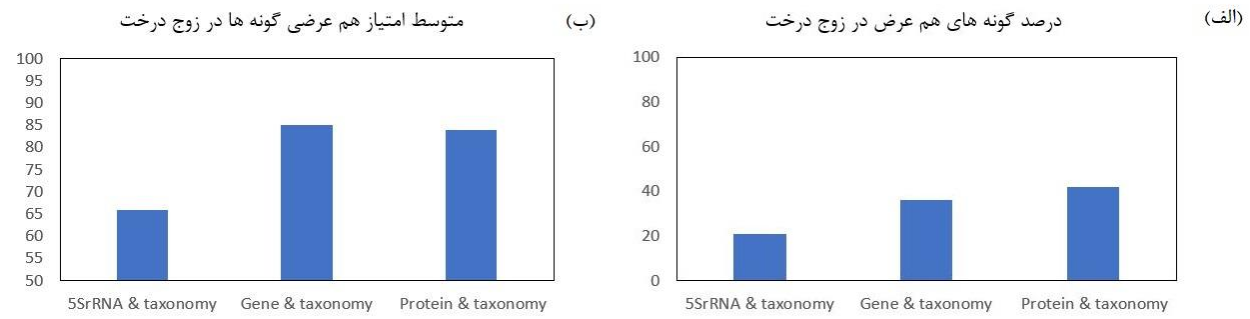
نمودار ۲–ب متوسط امتیاز هم ارزی گونه‌ها را در هر زوج درخت نشان می‌دهد. چنان که در این نمودار مشاهده می‌شود، امتیاز هم‌ارزی گونه‌های مشابه در زوج درخت ژن–تاکسونومی و پروتئین–تاکسونومی به ترتیب ۸۵ و ۸۴ درصد می‌باشد. این امتیاز در زوج درخت 5SrRNA–تاکسونومی به ۶۶ درصد می‌رسد.

همان‌طور که اشاره شد، یکی از مهم ترین کاربردهایی که نتایج به دست آمده تا این بخش برای ما فراهم می‌کند، این است که مناسب‌ترین جایگزین برای هرکدام از درخت‌هایی که در این پژوهش برای بررسی روابط تکاملی محاسبه شده است کدام است؟ به بیان دیگر، کدام درخت تکاملی را می‌توان جایگزین درخت تکاملی دیگری کرد؟ اگر در پژوهشی هدف مطالعه روابط تکاملی در سطح یک ژن خاص باشد و توالی‌های ژنی به سادگی و به طور کامل برای گونه‌های مورد نظر در دسترس نباشد به جای آن‌ها آیا می‌توان از درخت تاکسونومی استفاده نمود؟ همین سؤال به صورت بالعکس نیز مطرح است. به این معنی که اگر برای یک ژن خاص توالی‌های کامل ژنی یا پروتئنی در دسترس نباشد برای پیدا کردن روابط تکاملی در سطح ژن مربوطه، در صورتی که از درخت تاکسونومی استفاده شود چقدر می‌توان انتظار داشت که درخت تاکسونومی به درخت تکاملی ژن یا پروتئین مربوطه نزدیک باشد؟ بر اساس نتایجی که در نمودار ۲ بخش الف و ب قابل مشاهده است، درخت پروتئین و ژن در مقایسه با درخت 5SrRNA شباهت بیشتری نسبت به درخت تاکسونومی دارند. این موضوع نشان دهنده آن است که درخت پروتئین و ژن در مقایسه با درخت 5SrRNA می‌تواند جایگزین بهتری برای درخت تاکسونومی باشد و بالعکس. البته باید در نظر داشت که در این مطالعه از توالی‌های ژن و پروتئین آنزیم سولفیت ردوکتاز فلاوو پروتئین آلفا کامپوننت باکتریایی استفاده شده است؛ لذا این که نتایج به دست آمده در این پژوهش را بخواهیم به ژن‌ها و پروتئین‌های دیگر بسط بدهیم یک تقریب به حساب می‌آید.

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021



نمودار ۲: الف، درصد گونه‌های هم ارز در زوج درخت های پروتئین-تاکسونومی، ژن-تاکسونومی و 5SrRNA-تاکسونومی، ب، امتیاز هم ارزی گونه‌ها در زوج درخت‌های ذکر شده

بحث و نتیجه‌گیری

در این پژوهش حاضر برای ۱۹ گونه مشخص، به مقایسه درخت‌های فیلوژنتیکی ژن و پروتئین در آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا و درخت فیلوژنتیکی 5SrRNA با یکدیگر و با درخت تاکسونومی پرداخته شد. یکی از مهم‌ترین کاربردهای این پژوهش این است که مناسب‌ترین جایگزین برای هرکدام از درخت‌هایی که در این پژوهش برای بررسی روابط تکاملی محاسبه شده است کدام است؟ برای پیدا کردن جایگزین مناسب به مقایسه کمی زوج درخت‌ها پرداخته شده است. یکی دیگر از خروجی‌های منحصر به فرد این پژوهش، تعیین میزان شباهت و تفاوت بین چهار درخت ذکر شده برای ۱۹ گونه انتخابی در ارتباط با آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا به صورت کمی می‌باشد.

در دیگر مطالعات انجام شده برای بررسی توالی‌های 16SrRNA در مقایسه با ژن RecA، تمام توالی‌های موجود حتی توالی‌های ناقص استخراج‌شده از پایگاه داده نوکلئوتیدی NCBI بررسی شد؛ اما در پژوهش حاضر از توالی‌های کامل و گونه‌های یکسان از 5SrRNA، پروتئین و ژن استفاده شد تا نتایج کامل‌تری ارائه شود. همچنین در تحقیقات دیگر از الگوریتم‌های مختلفی برای ترسیم درخت فیلوژنتیکی استفاده شده است، ولی در این پژوهش برای رسم درخت فیلوژنتیکی تنها از روش اتصال همسایگی استفاده شد و همچنین از درخت خام تاکسونومی استخراج شده از NCBI، برای مقایسه زوج درخت‌ها استفاده شد. این مسئله یعنی استفاده از درخت خام تاکسونومی در مطالعات مشابه مشاهده نشد.

در پژوهشی دیگر، توالی ژن‌های gyrB، 16SrRNA و hsp65 برای ده گونه نوکاریا از بانک ژنی دریافت شد و بعد از هم‌ترازسازی توالی‌ها، درخت فیلوژنتیکی با استفاده از نرم‌افزار Mega به طور جداگانه برای این سه ژن رسم شد. در تحلیل یکی از درخت‌ها نشان داده شد که دو گونه به اشتباه کنار یکدیگر قرار گرفته‌اند، درحالی که شباهت دو گونه دورتر در درخت بیشتر بود. در نهایت نشان داده شد که ژن gyrB نسبت به بقیه بهتر عمل کرده و بهترین گزینه برای ترسیم روابط فیلوژنتیک گونه نوکاردیا می‌باشد [۲۲]. در این پژوهش، توالی ژن‌های gyrB، 16SrRNA و hsp65 برای ده گونه مورد بررسی قرار گرفته است، در حالی که در پژوهش حاضر توالی‌های پروتئین و ژن از آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا، 5SrRNA و درخت تاکسونومی برای ۱۹ گونه انتخابی بررسی شده است. مقایسه درخت‌ها در این پژوهش به صورت تک به تک انجام شده؛ اما در پژوهش حاضر این مقایسه دو به دو صورت گرفته است. به‌طورکلی، نحوه مقایسه درخت‌ها، گونه‌ها و تعداد آن‌ها در این دو پژوهش متفاوت است.

در پژوهشی دیگر، سوپراکسید دیسموتاز منگنز در ۴۸ گونه مختلف مورد مطالعه قرار گرفته است. به این منظور به مقایسه درخت مرجع که بر اساس استفاده از توالی‌های پروتئین MNSOD با ترکیب چند الگوریتم مختلف رسم درخت فیلوژنی به دست آمده است، با درختی که بر مبنای ماتریس فاصله حاصل از آنالیز کلاسترهای هیدروفوب (HCA) به دست آمده است، پرداخته شده است [۲۳]. مقایسه این دو درخت به گونه‌ای تفاوت روند تغییرات تکاملی در خصوصیات مربوط به توالی در مقایسه با HCA که یک خصوصیت ساختاری است را نشان می‌دهد. این پژوهش بیشتر به دنبال میزان معرفی همگرایی در خصوصیت HCA بوده است و در نهایت به این نتیجه رسیده است که در برخی گیاهان، قارچ ها،

www.SID.ir

Archive of SID

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

سیانوباکترها و فوتوباکترها، هم‌گرایی در HCA مشاهده می‌شود. بر اساس مطالعاتی که در این پژوهش ارائه شده است، HCA خصوصیتی بوده است که در طول تکامل در آنزیم MNSOD تمایل به هم‌گرایی از خود نشان داده است. چنان که مشاهده می‌شود، تقریباً در همه ابعاد به جزء در مسئله مقایسه درخت تفاوت‌های اساسی بین این پژوهش و پژوهش حاضر دیده می‌شود.

در مطالعه دیگری که در مجله میکروبیولوژی پزشکی به چاپ رسیده است، تنوع تغییرات را در فایلوم‌های منتخب خانواده پروتئینی RecA باکتریایی در مقایسه با ژن‌های 16SrRNA مورد بررسی قرار داده است. توالی‌های باکتریای خانواده پروتئینی RecA استخراج و با استفاده از نرم‌افزار CD-Hit دسته‌بندی شدند. بر اساس محاسبه میانگین امتیاز هم ترازسازی [ASS] برای فایلوم‌های ژن 16SrRNA، درخت 16SrRNA–تاکسونومی رسم شد. محاسبه ASS برای فایلوم‌های RecA نیز صورت گرفت. با مقایسه مقدار AAS بین فایلوم‌های یکسان ژن 16SrRNA و پروتئین RecA، مشاهده شد که فایلوم Actinobacteria در ژن 16SrRNA در مقایسه با RecA بیشترین فاصله را از یکدیگر دارند [۲۴].

در این پژوهش، بر پایه یک هدف بنیادی توانسته‌ایم به شباهت درختی که از توالی پروتئین آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا به دست می‌آید با درخت‌های ژن، 5SrRNA و تاکسونومی برای ۱۹ گونه انتخابی پی برد. آیا در عمل باید انتظار داشته باشیم همه درخت‌های رسم شده مثل هم باشند یا با هم تفاوت داشته باشند؟ اگر با هم تفاوت دارند، این میزان تفاوت چقدر می‌باشد؟ این‌ها سؤالاتی است که در این پژوهش به طور کمی به پاسخ آن‌ها رسیده شد. در نهایت، این محاسبات به ما در درک ساز و کار بخش‌هایی از مباحث تکامل کمک می‌کند؛ لذا یکی از نتایج کلی که از پژوهش حاضر به دست آمده است، جزئیاتی کمی است که گوشه‌هایی از مباحث تکاملی مرتبط با آنزیم سولفیت ردوکتاز فلاووپروتئینی آلفا و ۱۹ گونه انتخابی را روشن تر می‌کند.

تعارض منافع

این مقاله هیچ‌گونه تضاد منافعی ندارد.

References

1. Parey K, Warkentin E, Kroneck PM, Ermler U. Reaction cycle of the dissimilatory sulfite reductase from Archaeoglobus fulgidus. Biochemistry 2010;49(41):8912-21. doi: 10.1021/bi100781f

2. Parey K, Warkentin E, Kroneck PM, Ermler U. Reaction cycle of the dissimilatory sulfite reductase from Archaeoglobus fulgidus. Biochemistry 2010;49(41):8912-21. doi: 10.1021/bi100781f

3. Pinto R, Harrison JS, Hsu T, Jacobs WR, Leyh TS. Sulfite reduction in mycobacteria. J Bacteriol 2007;189(18):6714-22. doi: 10.1128/JB.00487-07

4. Brychkova G, Yarmolinsky D, Ventura Y, Sagi M. A novel in-gel assay and an improved kinetic assay for determining in vitro sulfite reductase activity in plants. Plant Cell Physiol 2012;53(8):1507-16. doi: 10.1093/pcp/pcs084

5. Yarmolinsky D, Brychkova G, Kurmanbayeva A, Bekturova A, Ventura Y, Khozin-Goldberg I, et al. Impairment in sulfite reductase leads to early leaf senescence in tomato plants. Plant Physiol 2014; 165(4): 1505–20. doi: 10.1104/pp.114.241356

6. Schnell R, Sandalova T, Hellman U, Lindqvist Y, Schneider G. Siroheme-and [Fe4-S4]-dependent 1.NirA from Mycobacterium tuberculosis is a sulfite reductase with a covalent Cys-Tyr bond in the active site. J Biol Chem 2005;280(29):27319-28. doi: 10.1074/jbc.M502560200

7. Gholami-Shabani M, Shams-Ghahfarokhi M, Gholami-Shabani Z, Akbarzadeh A, Riazi G, Ajdari S, et al. Enzymatic synthesis of gold nanoparticles using sulfite reductase purified from Escherichia coli: a green eco-friendly approach. Process Biochemistry 2015;50(7):1076-85.

8. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32(5):1792-7. doi: 10.1093/nar/gkh340

9. Delphi Group. "Towards on integrated information architecture"; 2004 [cited 2006 Nov 25]. Available from: http://www.delphigroup. Com.

10. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ. A new view of the tree of life. Nat Microbiol 2016;1:16048. doi: 10.1038/nmicrobiol.2016.48

11. Waikagul J, Thaekham U. Approaches to Research on the Systematics of Fish-Borne Trematodes. Academic Press; USA: Elsevier; 2014.

12. Dayhoff M, Schwartz R, Orcutt B. A model of evolutionary change in proteins. Atlas of Protein Sequence and Structure 1978;5:345-52.

13. Friedrich MW. Phylogenetic analysis reveals multiple lateral transfers of adenosine-5′-phosphosulfate reductase genes among sulfate-reducing microorganisms. Journal of Bacteriology 2002;184(1):278-89. doi: 10.1128/JB.184.1.278-289.2002

www.SID.ir

Archive of SID

14. Pearson WR, Robins G, Zhang T. Generalized neighbor-joining: more reliable phylogenetic tree reconstruction. Mol Biol Evol 1999;16(6):806-16. doi: 10.1093/oxfordjournals.molbev.a026165
15. Friedrich MW. Phylogenetic analysis reveals multiple lateral transfers of adenosine-5′-phosphosulfate reductase genes among sulfate-reducing microorganisms. Journal of Bacteriology 2002;184(1):278-89. doi: 10.1128/JB.184.1.278-289.2002
16. Eisen JA. The RecA protein as a model molecule for molecular systematic studies of bacteria: comparison of trees of RecAs and 16S rRNAs from the same species. J Mol Evol 1995;41(6):1105-23. doi**:** 10.1007/BF00173192
17. Berson AE, Peters MR, Waleh NS. Nucleotide sequence of recA gene of Aquaspirillum magnetotacticum. Nucleic Acids Res 1990; 18(3): 675. doi: 10.1093/nar/18.3.675
18. Rückert C. Sulfate reduction in microorganisms—recent advances and biotechnological applications. Current Opinion in Microbiology 2016;33:140-6. https://doi.org/10.1016/j.mib.2016.07.007
19. Hofmann K, Bucher P, Falquet L, Bairoch A. The PROSITE database, its status in 1999. Nucleic Acids Res 1999;27(1):215-9. doi: 10.1093/nar/27.1.215
20. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987;4(4):406-25. doi: 10.1093/oxfordjournals.molbev.a040454
21. Federhen S. The NCBI taxonomy database. Nucleic Acids Res 2012;40(Database issue):D136-43. doi: 10.1093/nar/gkr1178
22. Keikha M. Phylogenetic Analysis of Nocardia spp. using 16S rRNA, hsp65 and gyrB Genes. https://www.sid.ir/en/journal/ViewPaper.aspx?id=605484
23. Xiang H, Zhang R, Li N, Vossbrinck CR. Searching for convergent evolution in manganese superoxidase dismutase using hydrophobic cluster analysis. Genet Mol Biol 2014;37(2):460-74. doi: 10.1590/s1415-47572014005000008
24. Maleki A, Fahimi H, Taghizadeh M. Determining Difference in Evolutionary Variation of Bacterial RecA proteins vs 16SrRNA Genes by using 16s_Toxonomy Tree. Iranian Journal of Medical Microbiology 2019;13(1):32-43. [In Persian] doi: 10.30699/ijmm.13.1.32

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

www.SID.ir

Archive of SID

Original Article

Journal of Health and Biomedical Informatics
Medical Informatics Research Center
2020; 7(3): 326-336

Quantitative Comparison of Tree Pairs Resulted from Gene and Protein Phylogenetic Trees for Sulfite Reductase Flavoprotein Alpha-Component and 5S rRNA and Taxonomic Trees in Selected Bacterial Species

Tarahomi Elham[1], Fahimi Hossein[2], Taghizadeh Mohammad[3]*

• **Received:** 29 Nov 2019 • **Accepted:** 11 Jan 2020

Introduction: FAD is the cofactor of FAD-FR protein family. Sulfite reductase flavoprotein alpha-component is one of the main enzymes of this family. Based on applications of this enzyme in biotechnology and industry, it was chosen as the subject of evolutionary studies in 19 specific species.

Method: Gene and protein sequences of sulfite reductase flavoprotein alpha-component, 5S rRNA sequences, and taxonomic tree were extracted from 19 selected bacterial species. Then, phylogenetic trees of 5S rRNA and gene and protein sequences were compared with each other and with taxonomic tree. Phylogenetic trees were drawn by Mega7 software using neighbor-joining algorithm and taxonomic tree was extracted using NCBI taxonomy browser.

Results: By comparing the corresponding tree pairs, the percentage of equivalent species and the mean equivalence score of species were calculated for each tree pair. The gene-protein tree was allocated the highest scores in both quantities. In comparing the taxonomic tree with three other trees, gene-taxonomy tree achieved the highest percentage in the mean equivalence score and protein-taxonomy tree obtained the highest percentage of equivalent species.

Conclusion: Based on the results of the present research, the best replacement for each of the trees investigated in this study regarding evolutionary relations was identified. In other words, this study helps detect which evolutionary tree can be replaced for another evolutionary tree.

Keywords: Phylogenetic Tree, Taxonomic Tree, Sulfite Reductase Flavoprotein Alpha-Component Enzyme, 5S rRNA Gene, Tree Pair

• **Citation:** Tarahomi E, Fahimi H, Taghizadeh M. Quantitative Comparison of Tree Pairs Resulted from Gene and Protein Phylogenetic Trees for Sulfite Reductase Flavoprotein Alpha-Component and 5S rRNA and Taxonomic Trees in Selected Bacterial Species. Journal of Health and Biomedical Informatics 2020; 7(3): 332-36. [In Persian]

1. M.Sc. in Molecular Genetics, Genetics Dept., Islamic Azad University, Tehran Medical Sciences, Tehran, Iran
2. Ph.D. in Molecular Genetics, Assistant Professor, Genetics Dept., Islamic Azad University, Tehran Medical Sciences, Tehran, Iran
3. Ph.D. in Bioinformatics, Lecturer, Biotechnology, Islamic Azad University Dept., Tehran Medical Sciences, Tehran, Iran

***Corresponding Author:** Mohammad Taghizadeh
Address: Biotechnology Dept., Faculty of Advanced Sciences and Technology, Islamic Azad University, Tehran Medical Sciences, Khaghani St, Shariati St, Tehran, Iran
• **Tel:** 021-22006660 • **Email:** mtaghizadeh@Alumni.ut.ac.ir

Downloaded from jhbmi.ir at 11:37 +0330 on Monday January 11th 2021

www.SID.ir