



## Determining Difference in Evolutionary Variation of Bacterial *RecA* Proteins vs *16SrRNA* Genes Using 16S-Toxonomy Tree

Avisa Maleki<sup>1</sup>, Hossein Fahimi<sup>1</sup>, Mohammad Taghizadeh<sup>2\*</sup>

1. Department of Genetics, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran
2. Department of Biotechnology, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran

### Article Information

#### Article Subject:

Microbial Bioinformatics

DOI: 10.30699/ijmm.13.1.32

#### Corresponding author:

Mohammad Taghizadeh

Department of Biotechnology, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran

#### Email:

[mtaghizadeh@alumni.ut.ac.ir](mailto:mtaghizadeh@alumni.ut.ac.ir)

Use your device to scan  
and read the article online



### Abstract

**Background and Aims:** The rate of variation in various genes of a bacterial species is different during evolution. Therefore, in systematic bacterial studies many researchers compare the phylogenetic tree of a particular gene to the standard tree of an rRNA gene. Regarding the importance of *16SrRNA* gene and the evolutionary process of *RecA* protein family, we investigated the changes in sequences of the *RecA* protein family in selected bacterial phylums in comparing with their *16SrRNA* genes.

**Materials and Methods:** For this purpose, sequences of the *RecA* protein family were extracted from Uniprot database (with some help from ProSite) and then they were categorized by using CD-hit algorithm. One species was selected from each category. Then we found *16SrRNA* complete sequences for same species. After that, based on the Average Alignment Score (AAS), the 16S-taxonomic tree was obtained. Furthermore, Similar calculations were considered for corresponding *RecA* proteins phylums.

**Results:** By comparing amount of AAS in *16SrRNA* phylums and *RecA* phylums, we observed that the Actinobacteria phylum is the closest to the header phylum in the 16S-taxonomy tree, but this phylum in the *RecA* is the most distant to the header phylum, on the other hand, the position of the cyanobacteria phylum remains the same in both calculations, which indicates the least amount of changes in the genus and species of this phylum.

**Conclusion:** The 16S-taxonomy tree which has been compiled and presented in this study for the first time is different from the available bioinformatics algorithms for phylogenetic tree drawing. Finding the species with the highest and lowest rates of changes, can be a type of prediction method for indicating the reasons why bacteria become resistant to drugs over a long period of time.

**Keywords:** Taxonomy tree, Bacterial *RecA* protein family, *16SrRNA* gene, 16S-taxonomy tree, Drug resistance

Received: 2018/11/05 Accepted: 2019/05/15 Available online: 2019/06/20

Copyright © 2019 Iranian Journal of Medical Microbiology. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### How to cite this article:

Maleki A, Fahimi H, Taghizadeh M. Determining Difference in Evolutionary Variation of Bacterial *RecA* proteins vs *16SrRNA* Genes by using 16s-Toxonomy Tree. Iran J Med Microbiol. 2019; 13 (1) :32-43

## تعیین تفاوت تغییرات در طول تکامل برای پروتئین‌های باکتریایی *RecA* در مقایسه با ژن *16SrRNA* با استفاده از درخت *16S*-تاکسونومی

آویسا ملکی<sup>۱</sup>، حسین فهیمی<sup>۱</sup>، محمد تقی‌زاده<sup>۲\*</sup>

۱. گروه ژنتیک، دانشکده علوم و فناوری‌های نوین، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران  
۲. گروه بیوتکنولوژی، دانشکده علوم و فناوری‌های نوین، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران

### چکیده

### اطلاعات مقاله

**زمینه و اهداف:** میزان تغییرات در ژن‌های مختلف یک گونه باکتریایی طی تکامل متفاوت است. از این رو بسیاری از محققان برای مطالعه سیستماتیک باکتریایی، درخت فیلوژنتیک یک ژن خاص را با درخت استاندارد یک ژن *Arrn* مقایسه می‌کنند. با توجه به اهمیت ژن *16SrRNA* و روند تکاملی خانواده پروتئینی *RecA*، در این مطالعه تنوع تغییرات را در فایلوام‌های منتخب خانواده پروتئینی *RecA* باکتریایی در مقایسه با ژن‌های *16SrRNA* آنها بررسی کردیم.

**مواد و روش کار:** به این منظور توالی‌های باکتریایی خانواده پروتئینی *RecA* از پایگاه داده *UniProt* (و به کمک پروسایت) استخراج و با استفاده از نرم‌افزار *Cd-hit* دسته بندی شدند. از هر کلاس یک جنس و گونه انتخاب شد و توالی کامل ژن *16SrRNA* را برای جنس و گونه‌های منتخب به دست آوردیم. بعد از محاسبه میانگین امتیاز همترازسازی (AAS) برای فایلوام‌های ژن *16SrRNA*، درخت *16S*-تاکسونومی شکل گرفت. در ادامه، محاسبه *AAS* برای فایلوام‌های *RecA* نیز صورت گرفت.

**یافته‌ها:** با مقایسه مقدار AAS بین فایلوام‌های یکسان ژن *16SrRNA* و پروتئین *RecA*، مشاهده کردیم که فایلوام *Actinobacteria* در ژن *16SrRNA* نزدیک‌ترین فایلوام به سرگروه است، ولی همین فایلوام در پروتئین *RecA* بیشترین فاصله را با سرگروه دارد. جایگاه تکاملی فایلوام *cyanobacteria* در ژن *16SrRNA* و پروتئین *RecA* نسبت به سرگروه درخت *16S* - تاکسونومی تغییری نکرده است.

**نتیجه‌گیری:** درخت *16S*-تاکسونومی برای اولین بار در این پژوهش ارائه شده است و با الگوریتم‌های شناخته‌شده تفاوت دارد. به دست آوردن اطلاعات گونه‌هایی که کمترین و بیشترین میزان تغییرات تکاملی را دارند، می‌تواند یک روش پیش‌بینی باشد که چرا باکتری‌ها در زمان طولانی در برابر بعضی داروها مقاوم می‌شوند.

**کلمات کلیدی:** درخت تاکسونومی، خانواده پروتئینی *RecA* باکتریایی، ژن *16SrRNA*، درخت *16S*-تاکسونومی، مقاومت دارویی

### تاریخچه مقاله

دریافت: ۱۳۹۷/۱۰/۲۵

پذیرش: ۱۳۹۸/۰۳/۱۲

انتشار آنلاین: ۱۳۹۸/۳/۳۰

### موضوع:

بیوانفورماتیک میکروبی

IJMM1398;13(1): 32-43

### نویسنده مسئول:

دکتر محمد تقی‌زاده

گروه بیوتکنولوژی، دانشکده علوم و فناوری‌های نوین، علوم پزشکی تهران، دانشگاه آزاد اسلامی، تهران، ایران

پست الکترونیک:

[mtaghizadeh@alumni.ut.ac.ir](mailto:mtaghizadeh@alumni.ut.ac.ir)

کپی‌رایت © مجله میکروبیولوژی پزشکی ایران: دسترسی آزاد؛ کپی‌برداری، توزیع و نشر برای استفاده غیرتجاری با ذکر منبع آزاد است.

### مقدمه

سرعت تغییرات ژن‌های مختلف در طول تکامل با یکدیگر متفاوت است. به طور مثال ژن *16SrRNA* یکی از حفظ‌شده‌ترین ژن‌هایی است که کمترین تغییرات را در طول تکامل خود داشته است. اما سؤال اینجاست که در مقایسه با ژن *16SrRNA*، ژن‌های دیگر یا پروتئین آن‌ها، چه مقدار تغییرات را در فایلوام‌های مختلف جانداران طی تکامل نشان می‌دهند و یا تفاوت این تغییرات در فایلوام‌های مختلف موجودات زنده در مقایسه با ژن *16SrRNA* آن‌ها چگونه است. سؤالات به وجود آمده باعث انجام پژوهش‌هایی در این زمینه شده است که در ادامه بررسی برخی از آنها خواهیم پرداخت.

پروتئین *RecA* باکتریایی برای نوترکیبی همولوگ و ترمیم DNA آسیب‌دیده ضروری است. در نتیجه این خانواده پروتئینی در مطالعات مباحث مربوط به ژنتیک میکروبی نقش عمده‌ای دارد (۱). بر این اساس نبود خانواده پروتئینی *RecA* به علت تجمع جهش‌های مضر باعث مرگ باکتری می‌شود (۲). *RecA* فعالیت‌های زیادی دارد که شامل اتصال به DNA تکرشته‌ای و دو رشته‌ای، هیدرولیز ATP و ترمیم DNA با روش نوترکیبی همولوگ است. همچنین اعضای این خانواده پروتئینی که نوترکیبی همولوگ را ایجاد می‌کنند، برای محافظت از ژنوم و تنوع ژنتیکی به‌خصوص در یوکاریوت

ژن‌های *rRNA* برای زنده ماندن همه موجودات ضروری هستند و همچنین توالی بازهای آلی در این ژن‌ها به شدت حفظ شده است. به همین علت تعیین توالی بازهای آلی ژن *16SrRNA* به عنوان یک روش استاندارد برای شناسایی گونه‌ها، جنس‌ها و خانواده باکتری‌ها به کار می‌رود. امروزه روشن است که تجزیه و تحلیل توالی‌های ژن *16SrRNA* اهمیت زیادی دارد. تجزیه و تحلیل توالی ژن *rRNA* عمل قدرتمندی برای به دست آوردن تاکسونومی‌های میکروبی است (۱۰، ۱۱).

در پژوهشی که سال ۲۰۱۷ انجام شد، برای شناسایی و طبقه بندی دقیق گونه‌های مایکوباکتریوم از شباهت توالی ژن *16SrRNA* آن‌ها استفاده شد (۱۲). در آن پژوهش پس از جداسازی توالی‌های *16SrRNA* برای به دست آوردن بیشترین و کمترین مقدار شباهت بین تمام گونه‌های مایکوباکتریوم، از الگوریتم *muscle* برای همترازسازی دوگانه و از الگوریتم *likelihood* برای ترسیم درخت استفاده شد (۱۳، ۱۴). نتایج حاصل نشان داد گونه *poriferae* از جنس مایکوباکتریوم کمترین شباهت را به گونه‌های دیگر از این جنس باکتریایی دارد (۱۲).

در پژوهشی دیگر به مقایسه درخت فیلوژنی ژن *16SrRNA* و *RecA* از گونه‌های یکسان پرداخته اند. در این پژوهش آنالیز تکاملی ژن *RecA* با روش‌های فیلوژنی مولکولی صورت گرفته است. برای ترسیم درخت فیلوژنی 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 نزدیکشان استفاده کردند. در مواردی که توالی کامل برای ژن *16SrRNA* وجود نداشت، از توالی‌های ناقص (Partial) آن گونه استفاده شد. درخت‌های فیلوژنی تولیدشده، حاصل همترازسازی چندگانه توالی‌ها هستند و با استفاده از الگوریتم *Neighbor-joining* به دست آمده‌اند. بیشترین تفاوت میان درختان *RecA* و *rRNA* مربوط به پروتئوباکترها بوده است. طبق این مطالعات مشخص شد پروتئین *RecA* می‌تواند برای مطالعات مولکولی مفید باشد (۱۵).

در پژوهش حاضر نیز به علت اهمیت میکروبی‌شناسی پزشکی و مباحث تکاملی مربوط به خانواده پروتئینی *RecA*، سرعت تغییرات فایلوم‌های منتخب این خانواده پروتئینی در مقایسه با فایلوم‌های یکسان در ژن *16SrRNA* با استفاده از تعیین شاخص AAS (میانگین امتیاز همترازی) بررسی شد. به دست آوردن سرعت تغییرات می‌تواند کمک شایانی در زمینه مقاومت دارویی برای جنس و گونه‌های فایلوم‌های مذکور انجام دهد.

ها اهمیت دارند و برای میوز ضروری هستند. این پروتئین برای شروع نوترکیبی همولوگ، دومینی را که جایگاه اتصال ATP دارد، روی DNA تک‌رشته آسیب‌دیده قرار می‌دهد و با اتصال ATP به این جایگاه، ترمیم آغاز می‌شود (۳). *RecA* حدوداً ۳۵۰ اسید آمینه دارد که توالی آن در بین گونه‌های باکتریایی بسیار حفظ شده است. معادل همولوگ پروتئین *RecA* باکتریایی در یوکاریوت‌ها *DMC1*، *Rad51* و در آرکی‌باکترها *RadA*، *RadB* نامیده می‌شود (۴، ۵).

در پژوهشی که در سال ۱۹۹۶ در دانشگاه استنفورد آمریکا انجام شد، پروتئین‌های مختلف خانواده *RecA* را با یکدیگر مقایسه کردند. به این صورت که پروتئین *RecA* باکتریایی به عنوان گروه اول و *DMC1/Rad51* در یوکاریوت‌ها و *Rada* در آرکی‌باکترها را به عنوان گروه دوم در نظر گرفتند. برای راحت خوانده شدن پروتئین‌های گروه دوم، حروف اول هر پروتئین را در نظر گرفتند و گروه دوم را به اختصار RDR نام‌گذاری کردند. در آن پژوهش همترازسازی چندگانه برای توالی‌های انتخاب شده RDR انجام شد و نتایج را به صورت امتیاز ارائه دادند. نتایج این همترازسازی چندگانه نشان داد ۶ بلاک (قطعه‌ای از توالی که حفظ شده است) مشترک در میان توالی‌های RDR وجود دارد. سپس ۶ بلاک حفظ شده توالی‌ها را استخراج کردند. این کار برای همترازسازی توالی‌های RDR با توالی پروتئین *RecA* بسیار مناسب بود. در مرحله بعد یک همترازسازی چندگانه بین بلاک‌های مشترک پروتئین‌های RDR با پروتئین *RecA* صورت گرفت. نتایج حاصل از همترازسازی نشان داد پروتئین‌های RDR شباهت بیشتری به یکدیگر در مقایسه با پروتئین *RecA* دارند، اما همچنان این سه پروتئین با پروتئین *RecA* باکتریایی همولوگ هستند (۶).

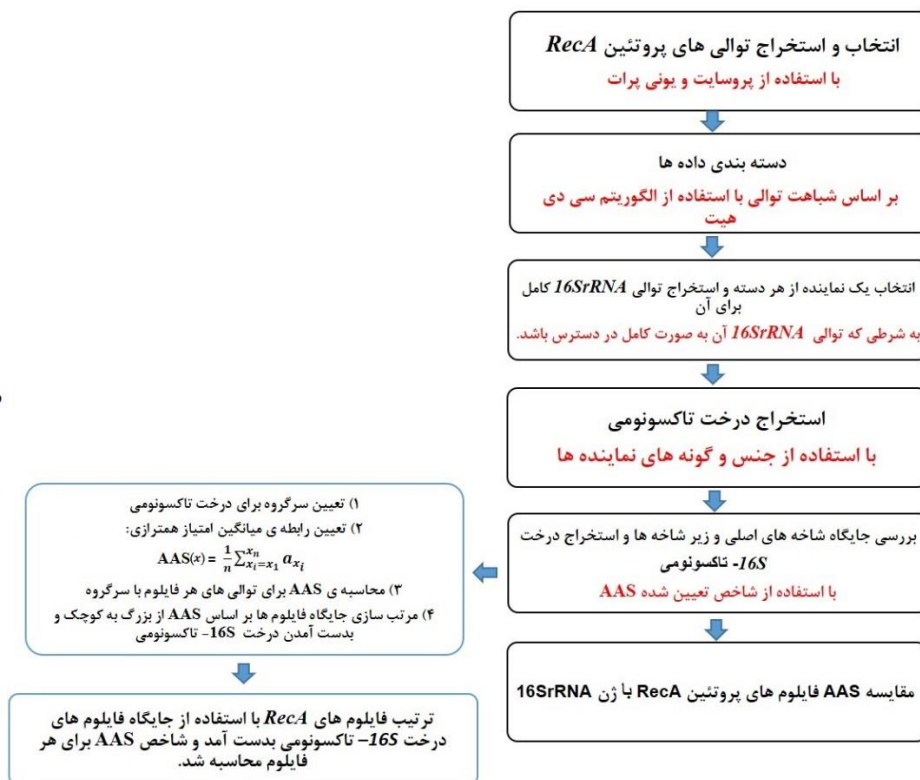
تاکسونومی باکتری‌ها در اواخر قرن ۱۹ آغاز شد؛ زمانی که ویژگی‌های فنوتیپی، محیط کشت، مورفولوژی، اندازه کلونی و رنگ واکنش‌های شیمیایی برای توصیف باکتری لازم بود (۷). به عبارت دیگر، رویکردهای اولیه برای ایجاد درخت تکاملی موجودات، بر اساس ویژگی‌های فیزیکی و متابولیکی بوده است. امروزه روش‌های مولکولی به سرعت گسترش یافته است و این گسترش شامل ایجاد تنوع در ارائه درخت زندگی (تکاملی) نیز می‌شود، زیرا سازماندهی درخت‌ها با مشاهده مستقیم و آزمایش روی توالی ژن‌ها صورت می‌گیرد (۸). درخت تاکسونومی اطلاعات ضروری را برای دانشمندان فراهم می‌کند تا با استفاده از این اطلاعات به رابطه درستی بین زندگی میکروارگانیسم‌ها و محیط زیست مختلف آن‌ها برسند (۹).

## مواد و روش‌ها

### انتخاب توالی

انتخاب کنیم. به این ترتیب تعداد توالی‌هایی که برای رسم درخت تکاملی یک خانواده پروتئینی لازم است، کاهش می‌یابد. این برنامه قابلیت کلاستر بندی میلیون‌ها توالی استخراج شده از پایگاه داده‌های اطلاعات را دارد که مربوط به یک خانواده بزرگ پروتئینی هستند. توالی‌ها در ابتدا با کاهش طول مرتب می‌شوند، و بلندترین توالی به عنوان نماینده یا سرگروه انتخاب می‌شود. سپس هر توالی باقی‌مانده با نماینده موجود مقایسه می‌شود. اگر شباهت توالی با اولین توالی نماینده بیشتر از حد آستانه تعیین شده باشد، بدون مقایسه با دیگر نماینده‌ها وارد یک کلاستر پروتئینی می‌شود؛ اگر یک توالی با هیچ کدام از نماینده‌های تعیین شده شباهت کافی نداشته، آن توالی خود به عنوان یک نماینده محسوب می‌شود. این نرم‌افزار با استفاده از الگوریتم پایه به نام فیلترینگ کلمه کوتاه (Short word filtering) میزان شباهت توالی‌ها را با توجه به حد آستانه تعیین شده، مشخص و کلاستر بندی می‌کند (۱۸).

در آغاز این پژوهش توالی‌های خانواده پروتئینی *RecA* با استفاده از پایگاه داده پروسایت (<https://prosite.expasy.org>) و یونی پرات (<https://www.uniprot.org/>) استخراج و ذخیره‌سازی شدند (۱۶). سپس جداسازی توالی‌های باکتریایی پروتئینی *RecA* از بین بقیه پروتئین‌های این خانواده صورت گرفت. به منظور کاهش خطا در رسم درخت تکاملی با دسته‌بندی کردن توالی‌ها، تعداد جنس و گونه‌های منتخب برای این پژوهش محدود شد. دسته‌بندی توالی‌ها با سرور سی دی هیت (<http://weizhongli-lab.org/cd->) انجام شد (۱۷). این ابزار توالی‌های مشابه را در یک کلاستر (دسته) قرار می‌دهد و به ما کمک می‌کند تا از هر کدام از این کلاسترها بتوانیم یک توالی یا یک جنس و گونه را به عنوان نماینده



شکل شماتیک، مراحل انجام کار به صورت خلاصه را نشان می‌دهد.

نداشتند، از دیگر اعضای کلاستر استفاده شد. در نتیجه ۳۰ توالی برای پروتئین *RecA* و ۳۰ توالی نوکلئوتیدی *16SrRNA* کامل از جنس و گونه یکسان به دست آمد. هر یک از این توالی‌ها به نمایندگی از دیگر اعضای کلاسترشان به دلیل شباهت زیاد و داشتن ژن کامل *16SrRNA* انتخاب شدند و باقی کلاسترها به دلیل نداشتن توالی کامل ژن *16SrRNA* یا تعداد اندک توالی در کلاسترشان در این

نتیجه استفاده از CD-HIT با حد آستانه ۰/۷۵ برای کل توالی‌های باکتریایی *RecA*، تعداد ۱۰۴ کلاستر بود که ۴۰ کلاستر از آنها به صورت تک‌عضوی بودند. در مرحله بعد سرگروه‌های مربوط به هر کلاستر انتخاب شد. به منظور یافتن توالی ژن *16SrRNA* کامل آن‌ها از پایگاه داده نوکلئوتیدی ان.سی.بی.آی. (NCBI) استفاده شد. برای کلاسترهای بزرگ‌تر که سرگروه آن‌ها توالی *16SrRNA* کامل

پژوهش استفاده نشدند.

## استخراج درخت تاکسونومی و محاسبه درخت ۱۶ اس-

### تاکسونومی برای ۳۰ گونه انتخابی

درخت تاکسونومی مربوط به ۳۰ جنس و گونه انتخابی از بخش [NCBI Taxonomy Browser](#) استخراج شد (۱۹). برای مشخص شدن صحت جایگاه تکاملی هر کدام از شاخه‌های اصلی و زیرشاخه‌های درخت تاکسونومی، همترازسازی‌های دوگانه میان توالی‌های ژن *16S rRNA* درون شاخه‌ای و بین شاخه‌ای صورت گرفت. برای این منظور در ابتدا با استفاده از نرم‌افزار جلیو (Jalview) برای توالی‌های هر یک از زیرشاخه‌های موجود در شاخه اصلی با توالی‌های همان شاخه همترازسازی دوگانه انجام شد. در ادامه مشخص شد هر کدام از شاخه‌های اصلی درخت تاکسونومی مربوط به یک فایلم (Phylum) باکتریایی هستند. در مرحله بعد به منظور بررسی جایگاه تکاملی صحیح ۸ شاخه اصلی در درخت تاکسونومی، از هر شاخه یک توالی به عنوان نماینده انتخاب شد و با هر یک از نماینده‌های ۷ شاخه دیگر همترازسازی صورت گرفت. توالی‌ای که کمترین مجموع امتیاز همترازسازی را نسبت به مجموع امتیاز همترازسازی دیگر نماینده‌ها گرفت، به عنوان توالی مرجع درخت انتخاب شد که در این پژوهش توالی شماره ۱۸ از درخت تاکسونومی است.

### محاسبه شاخص میانگین امتیاز همتراز به منظور تعیین

#### چینش شاخه‌های اصلی در درخت تاکسونومی

برای محاسبه شاخص میانگین امتیاز همترازسازی برای هر یک از ۷ شاخه اصلی درخت تاکسونومی از رابطه زیر استفاده شد.

$$AAS(x) = \frac{1}{n} \sum_{x_i=x_1}^{x_n} a_{x_i} \quad \text{رابطه:}$$

در این رابطه AAS مخفف Average Alignment Score و  $AAS(x)$  نشان‌دهنده میانگین امتیاز همترازی شاخه‌های فرعی  $i$  در هر شاخه اصلی  $x$  است. همچنین  $n$  نشان‌دهنده تعداد شاخه‌های فرعی  $i$  در هر شاخه اصلی  $x$  (از ۱ تا ۷) است. در این رابطه  $a_{x_i}$  نشان‌دهنده امتیاز همترازی هر شاخه فرعی  $i$  در یک شاخه اصلی  $x$  با توالی مرجع است.

توالی مرجع درخت، توالی ژن *16S rRNA* کامل از جنس و گونه *Bacteroides fragilis* است که بر اساس همترازی‌هایی که در بخش قبلی توضیح داده شد، انتخاب شد. از این رابطه به منظور چینش ۷ شاخه اصلی دیگر استفاده گردید. به این صورت که برای توالی‌های هر فایلم با توالی مرجع، همترازسازی دوگانه صورت گرفت و مقدار AAS برای هر فایلم محاسبه شد. سپس بر اساس بیشترین مقدار تا کمترین مقدار AAS، فایلم‌ها در ساختار درخت مرتب شدند. نتیجه به دست آمده درخت S16-تاکسونومی نامیده شد. در ادامه میانگین

امتیاز همترازسازی (AAS) برای توالی‌های پروتئینی RecA در فایلم‌ها محاسبه شد؛ با این تفاوت که به منظور بررسی مقدار تغییرات پروتئین RecA در مقایسه با ژن *16S rRNA*، جایگاه فایلم‌های پروتئین RecA همانند فایلم‌های درخت *16S*-تاکسونومی در نظر گرفته شد.

### یافته‌ها

همان‌طور که در بخش روش کار قسمت انتخاب توالی توضیح داده شد، با استفاده از پایگاه داده پروسایت توالی‌های خانواده پروتئینی RecA استخراج شدند و پس از دسته‌بندی از میان ۱۰۴ کلاستر حاصل شده، ۳۰ گونه از سی کلاستر (با تعداد توالی‌های بزرگ‌تر یا مساوی ۱۰) به عنوان گونه‌های نماینده کل خانواده، انتخاب شدند. هدف اصلی این پژوهش تعیین میزان تغییرات در پروتئین‌های RecA در طول زمان تکاملی آن‌ها در مقایسه با حفظ‌شده‌ترین ژن شناخته شده بود که به صورت کلاسیک ژن *16S rRNA* معرفی می‌شود. بر پایه این هدف می‌توان به گونه‌هایی که سرعت تغییرات پروتئین RecA آن‌ها بیشترین و کمترین مقدار بوده است، دست پیدا کرد. به این ترتیب می‌توان تعیین کرد سرعت تغییرات برای ژن مربوط به پروتئین RecA باکتریایی در کدام گونه بسیار زیاد و در کدام گونه بسیار کم است. این مسئله احتمالاً می‌تواند نوعی پیشگویی برای یکی از دلایل ایجاد مقاومت باکتری‌ها به داروهای طی زمان طولانی باشد. به طور اساسی تعیین میزان تغییرات در طول تکامل در پروتئین RecA باکتریایی در گونه‌های مختلف به طور مستقل اطلاعات پایه‌ای مهمی را در اختیار ما قرار می‌دهد.

### درخت تاکسونومی حاصل برای ۳۰ گونه انتخابی

برای به دست آوردن رابطه تکاملی میان گونه‌های انتخابی، در قدم اول درخت تاکسونومی از Taxonomy browser پایگاه داده NCBI استخراج شد. شکل ۱ درخت تاکسونومی مربوط به ۳۰ گونه انتخابی را نشان می‌دهد. در این درخت ۸ شاخه اصلی دیده می‌شود که هر کدام از آن‌ها مربوط به یک فایلم باکتریایی مجزاست. سؤال مهمی که اینجا مطرح است این است که ترتیب شاخه‌های اصلی باید چگونه باشد؛ یعنی کدام شاخه قدیمی‌تر (یا جدیدتر) است و کدام شاخه‌ها باید پس از آن قرار داده شوند؟

برای اصلاح ترتیب شاخه‌های اصلی درخت تاکسونومی، باید یک قاعده مناسب تعیین شود. به این منظور از توالی کامل *16S rRNA* مربوط به ۳۰ گونه انتخابی استفاده شد. برای پیدا کردن و استفاده از یک قاعده مناسب مراحل زیر به ترتیب انجام شد:

۱. کسب اطمینان از صحیح بودن جایگاه تکاملی شاخه‌های فرعی درون هر شاخه اصلی:

به این منظور دو بررسی صورت گرفت. در اولین بررسی، رده‌بندی‌های تاکسونومی ۳۰ گونه انتخابی با استفاده از مرورگر تاکسونومی ان.سی.بی.آی. (NCBI Taxonomy Browser) استخراج و به تفکیک شاخه‌های اصلی، دسته‌بندی شدند. نتیجه این بررسی در شکل ۲ قابل مشاهده است. همان‌طور که در شکل ۲ دیده می‌شود، گونه‌های موجود در هر شاخه اصلی، حداقل از بخش فایلوم با یکدیگر مشابه هستند. البته در بعضی از شاخه‌های اصلی میزان این شباهت تا رده‌های پایین‌تر ادامه پیدا می‌کند.

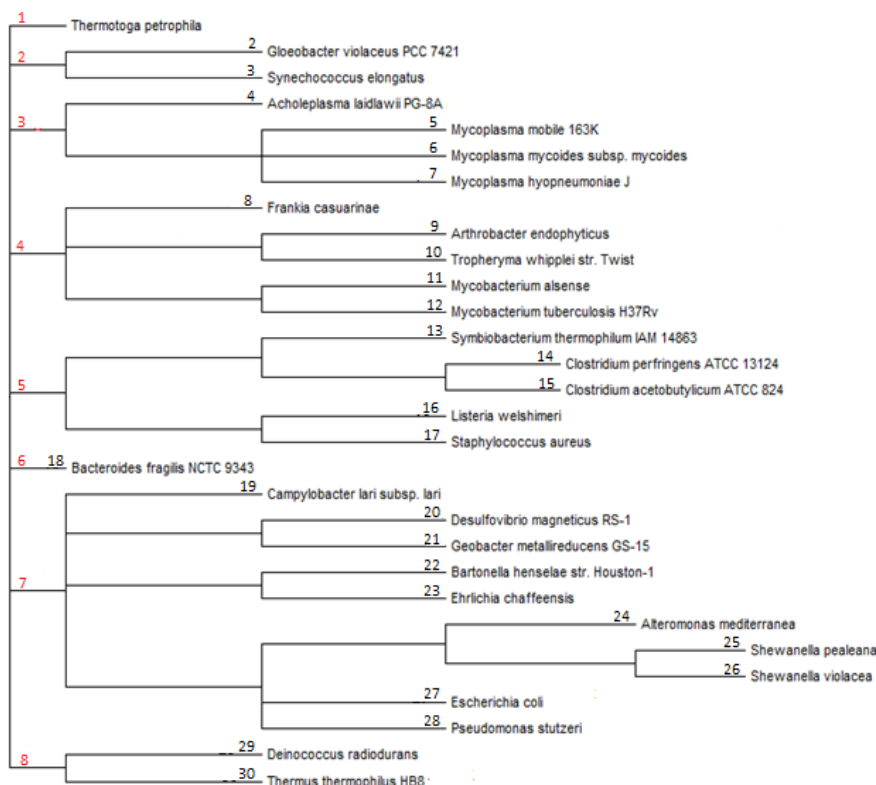
در دومین بررسی میزان شباهت توالی‌های ژن *16SrRNA* شاخه‌های فرعی در هر شاخه اصلی، بعد از انجام ۷۳ همترازسازی محاسبه شد. نتایج این بررسی به این صورت به دست آمد که میانگین امتیاز همترازسازی توالی‌های درون فایلومی شاخه‌های اصلی با شماره شاخه‌های ۲، ۳، ۴، ۵، ۷ و ۸ برابر شد با ۲۱۱۷، ۱۵۳۷، ۲۲۰۲، ۱۷۱۱، ۱۴۸۱ و ۱۵۰۰. شاخه‌های اصلی ۱ و ۶ به

دلیل تک‌توالی بودن در این محاسبه استفاده نشد.

همچنین میانگین امتیاز همترازسازی خارج فایلومی برای هر یک از شاخه‌های ذکر شده برابر شد با ۱۱۴۹، ۱۰۲۳، ۱۲۵۵، ۱۲۰۲، ۱۱۸۴ و ۱۱۱۳. همان‌طور که قابل ملاحظه است میانگین امتیاز همترازسازی درون فایلومی هر شاخه اصلی عددی بیشتر را نسبت به میانگین امتیاز همترازسازی خارج فایلومی نشان می‌دهد. لذا بنا بر نتایج اولین و دومین بررسی، به طور نسبی می‌توان گفت که در این درخت تاکسونومی، شاخه‌های فرعی در هر شاخه اصلی، به طور صحیحی قرار گرفتند.

۲. بررسی صحت جایگاه تکاملی هر شاخه اصلی در درخت تاکسونومی

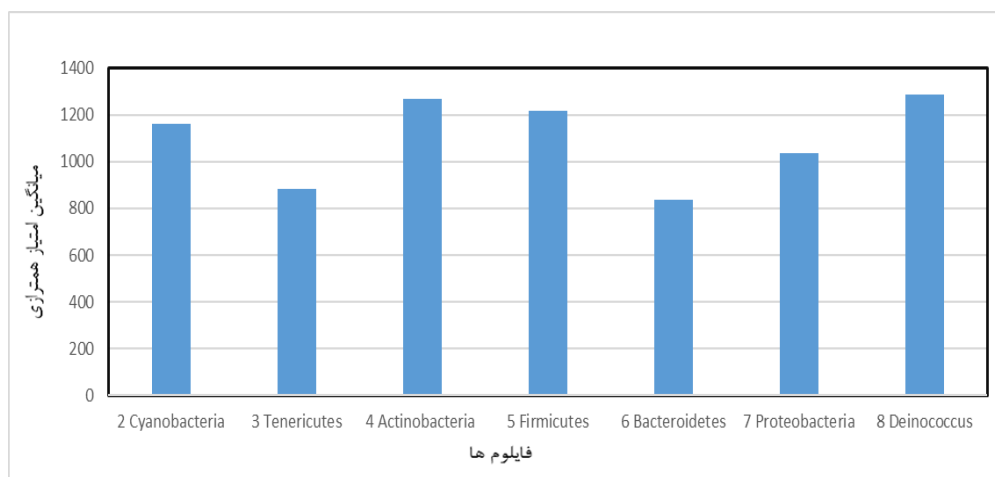
با توجه به درخت تاکسونومی (شکل ۱)، به منظور بررسی جایگاه شاخه‌های اصلی که آیا از نظر تکاملی ترتیب درستی دارند یا خیر، میانگین امتیاز همترازسازی برای توالی‌های هر فایلوم با توالی شماره ۱ درخت تاکسونومی که گونه *Thermotoga* است، محاسبه شد و نتیجه آن در شکل ۳ قابل ملاحظه است.



شکل ۱. درخت تاکسونومی استخراج شده از پایگاه داده ان.سی.بی.آی.، ۸ شاخه اصلی دارد. اعداد قرمز در شکل نشان‌دهنده شماره گذاری شاخه اصلی و اعداد سیاه نشان‌دهنده شاخه‌های فرعی هستند. همان‌طور که از روی درخت تاکسونومی می‌توان تشخیص داد، در بعضی از فایلوم‌ها چند گونه و در بعضی از آن‌ها فقط یک گونه قرار گرفته است

- 1) Bacteria; Thermotogae; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga
- 2) Bacteria; Terrabacteria group; Cyanobacteria/Melainobacteria group; cyanobacteria; Gloeobacteria; Gloeobacteriales; Gloeobacteraceae; Gloeobacter
- 3) Bacteria; Terrabacteria group; cyanobacteria/Melainobacteria group; cyanobacteria; Synechococcales; Synechococcaceae; Synechococcus
- 4) Bacteria; Terrabacteria group; Tenericutes; Mollicutes; Achaeplastales; Achaeplastaceae; Achaeplasma
- 5) Bacteria; Terrabacteria group; Tenericutes; Mollicutes; Mycoplasmatales; Mycoplasmataceae; Mycoplasma
- 6) Bacteria; Terrabacteria group; Tenericutes; Mollicutes; Mycoplasmatales; Mycoplasmataceae; Mycoplasma; Mycoplasma mycoides group
- 7) Bacteria; Terrabacteria group; Tenericutes; Mollicutes; Mycoplasmatales; Mycoplasmataceae; Mycoplasma
- 8) Bacteria; Terrabacteria group; Actinobacteria; Actinobacteria; Frankiales; Frankiaceae; Frankia
- 9) Bacteria; Terrabacteria group; Actinobacteria; Actinobacteria; Micrococcales; Micrococcaceae; Arthrobacter
- 10) Bacteria; Terrabacteria group; Actinobacteria; Actinobacteria; Micrococcales; unclassified Micrococcales; Tropheryma
- 11) Bacteria; Terrabacteria group; Actinobacteria; Actinobacteria; Corynebacteriales; Mycobacteriaceae; Mycobacterium
- 12) Bacteria; Terrabacteria group; Actinobacteria; Actinobacteria; Corynebacteriales; Mycobacteriaceae; Mycobacterium; Mycobacterium tuber
- 13) Bacteria; Terrabacteria group; Firmicutes; Clostridia; Clostridiales; Symbiobacteriaceae; Symbiobacterium
- 14) Bacteria; Terrabacteria group; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium
- 15) Bacteria; Terrabacteria group; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium
- 16) Bacteria; Terrabacteria group; Firmicutes; Bacilli; Bacillales; Listeriaceae; Listeria
- 17) Bacteria; Terrabacteria group; Firmicutes; Bacilli; Bacillales; Staphylococcaceae; Staphylococcus
- 18) Bacteria; FCB group; Bacteroidetes/Chlorobi group; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae; Bacteroides
- 19) Bacteria; Proteobacteria; delta/epsilon subdivisions; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter
- 20) Bacteria; Proteobacteria; delta/epsilon subdivisions; Deltaproteobacteria; Desulfuovibrionales; Desulfuovibrionaceae; Desulfuovibrio
- 21) Bacteria; Proteobacteria; delta/epsilon subdivisions; Deltaproteobacteria; Desulfuomonadales; Geobacteraceae; Geobacter
- 22) Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales; Bartonellaceae; Bartonella
- 23) Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales; Anaplasmataceae; Ehrlichia; canis group
- 24) Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Alteromonadaceae; Alteromonas
- 25) Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Shewanellaceae; Shewanella
- 26) Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales; Shewanellaceae; Shewanella
- 27) Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Escherichia
- 28) Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas stutzeri group; Pseudomonas
- 29) Bacteria; Terrabacteria group; Deinococcus-Thermus; Deinococci; Deinococcales; Deinococcaceae; Deinococcus
- 30) Bacteria; Terrabacteria group; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae; Thermus

شکل ۲. رده‌بندی تکاملی برای ۳۰ گونه انتخابی. در این شکل اعضای هر شاخه اصلی با ایجاد فاصله از یکدیگر جدا شده‌اند. شماره‌گذاری‌ای که در سمت چپ رده‌بندی هر گونه دیده می‌شود، با شماره‌گذاری درخت شکل ۱ مطابقت دارد. در این پژوهش شاخه‌های اصلی را فایلوم نامیده‌ایم



شکل ۳. نمودار میانگین امتیاز همترازی برای توالی‌های هر فایلوم با توالی شماره ۱ (گونه *Thermotoga*) در درخت تاکسونومی (شکل ۱).

درخت تاکسونومی مرتب شدند. حاصل این چینش درخت 16S- تاکسونومی نامیده شد (شکل ۳).

طبق چینش انجام شده فایلوم Bacteroidetes به عنوان سرگروه درخت انتخاب شد و پس از آن به ترتیب فایلوم‌های شماره دو: Actinobacteria، سه: Firmicutes، چهار: Cyanobacteria، پنج: Deinococcus-thermus، شش: Tenericutes، هفت: proteobacteria و آخرین فایلوم Thermotogae قرار گرفتند. فایلوم یا شاخه اصلی proteobacteria خود شامل Epsilon subdivision و Gammaproteobacteria می‌شود.

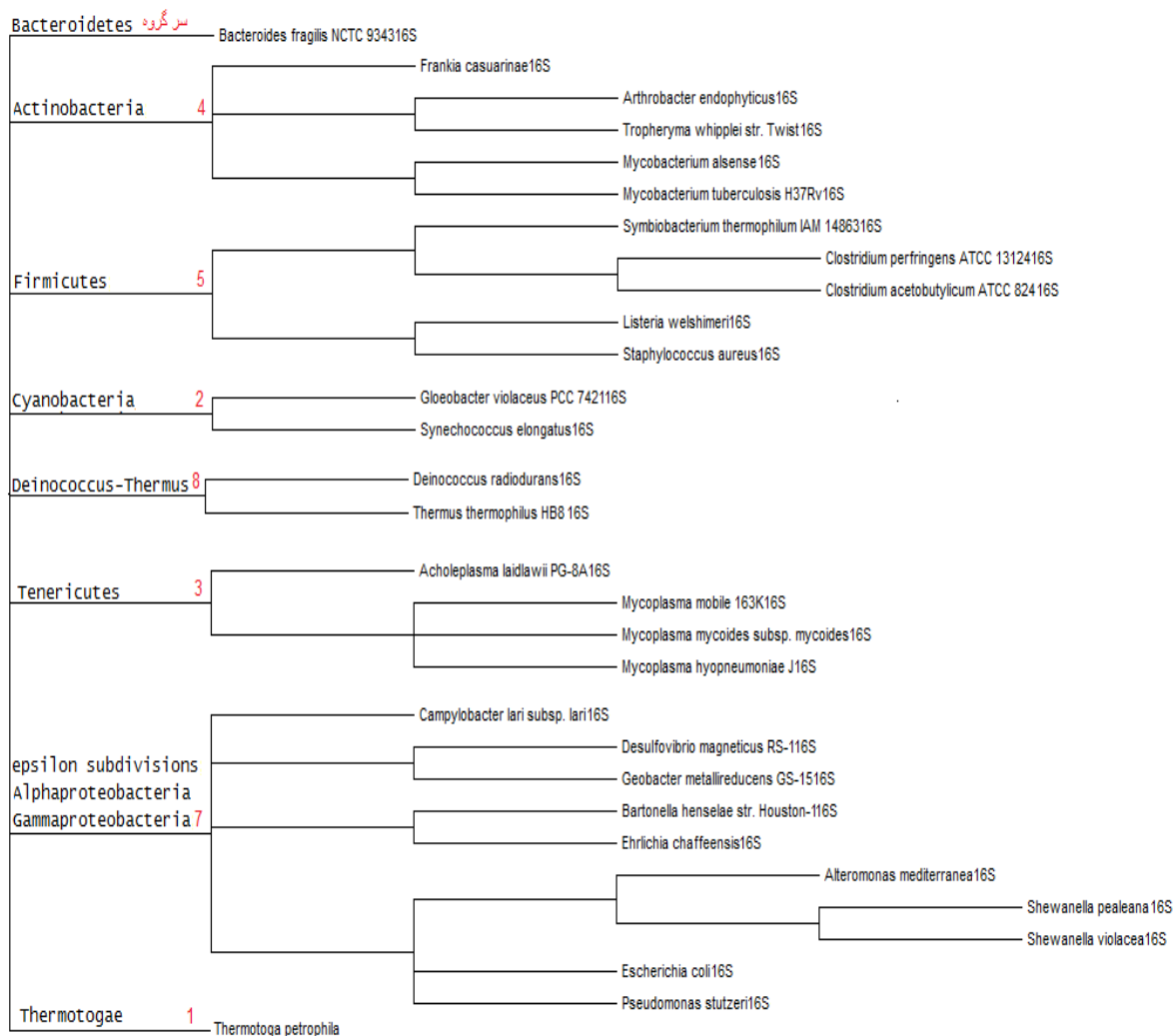
تعیین روشی دقیق و مناسب برای چینش شاخه‌های اصلی درخت تاکسونومی پیچیده است، اما ترتیبی که در شکل ۴ ارائه داده شده است، بر مبنای ژن *16S rRNA* است و می‌تواند نتیجه‌ای بهتر از خروجی درخت تاکسونومی ان.سی.بی.آی. در ترتیب شاخه‌های

با توجه به نمودار قابل مشاهده در شکل ۳ می‌توان دریافت که نمودار، روند کاهشی (یا افزایشی) را دنبال نمی‌کند و این نشانگر آن است که جایگاه شاخه‌های اصلی به صورت تصادفی تعیین شده است و بر اساس روند تغییرات تکاملی نیست. ۳. مرحله اصلی در اصلاح جایگاه شاخه‌های اصلی درخت تاکسونومی

همان‌طور که در بخش روش کار توضیح داده شد، با استفاده از شاخص AAS میانگین امتیاز همترازی توالی‌های هر فایلوم با توالی سرگروه محاسبه شد. چگونگی انتخاب سرگروه درخت که توالی شماره ۱۸ (با جنس و گونه *Bacteroides fragilis*) درخت تاکسونومی است در بخش سوم روش کار توضیح داده شده است. فایلومی که بیشترین مقدار AAS را داشت، بعد از توالی سرگروه قرار گرفت. فایلوم‌ها بر اساس بیشترین تا کمترین مقدار AAS در

آوردن درخت‌های فیلوژنی مانند الگوریتم neighbour joining و UPGMA روی ۳۰ گونه انتخابی ژن *16S rRNA*، نتایجی می‌دهد که هم با یکدیگر و هم با درخت تاکسونومی ان.سی.بی.آی. متفاوت است.

اصلی به ما بدهد. ترسیم یک درخت تکاملی صحیح که ارتباط دقیق تکاملی جنس و گونه‌های مختلف را به ما بدهد مسئله‌ای حل نشده است. بر اساس نتایجی که در این پژوهش به دست آمده و بررسی‌هایی که انجام شده است، الگوریتم‌های موجود برای به دست



شکل ۴. درخت *16S*-تاکسونومی: این درخت بر اساس میانگین امتیاز همترازی (AAS) شاخه‌های فرعی هر فایلوم با فایلوم سرگروه به دست آمده است.

از سرگروه فاصله می‌گیرند، میزان شباهت و امتیاز آنها کمتر می‌شود. همچنین میزان شباهت از فایلوم دوم به سوم به طور چشمگیری کاهش یافته است که این امر بیانگر تغییرات شدید تکاملی در این فایلوم‌ها است. در ادامه برای توالی‌های پروتئین *RecA* محاسبات مشابه با موارد توضیح داده شده برای ژن *16S rRNA* انجام شد.

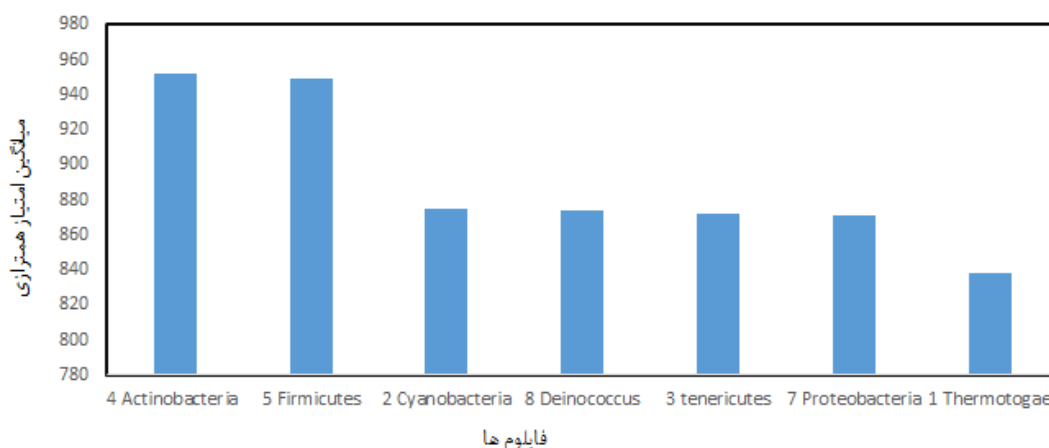
**آنالیز توالی‌های پروتئینی *RecA* برای تعیین فاصله تکاملی**  
هدف اصلی این پژوهش به دست آوردن فاصله تکاملی توالی‌های منتخب *RecA* از خانواده پروتئینی *RecA* در مقایسه با

نمودار شکل ۵، متوسط امتیاز همترازی توالی‌های هر فایلوم با فایلوم سرگروه یعنی AAS را برای درخت *16S*-تاکسونومی نشان می‌دهد. همان‌طور که انتظار می‌رفت نمودار سیر نزولی را دنبال می‌کند و فایلوم Thermotogae با کمترین میزان میانگین (۸۳۸) در دورترین فاصله تکاملی نسبت به سرگروه قرار گرفته است و فایلوم Actinobacteria با میانگین امتیاز ۹۵۲/۵ نزدیک ترین فایلوم به سرگروه است که نشان‌دهنده شباهت بیشتر این فایلوم‌ها از نظر تکاملی با سرگروه است. این روند نشان داده است هرچه توالی‌ها

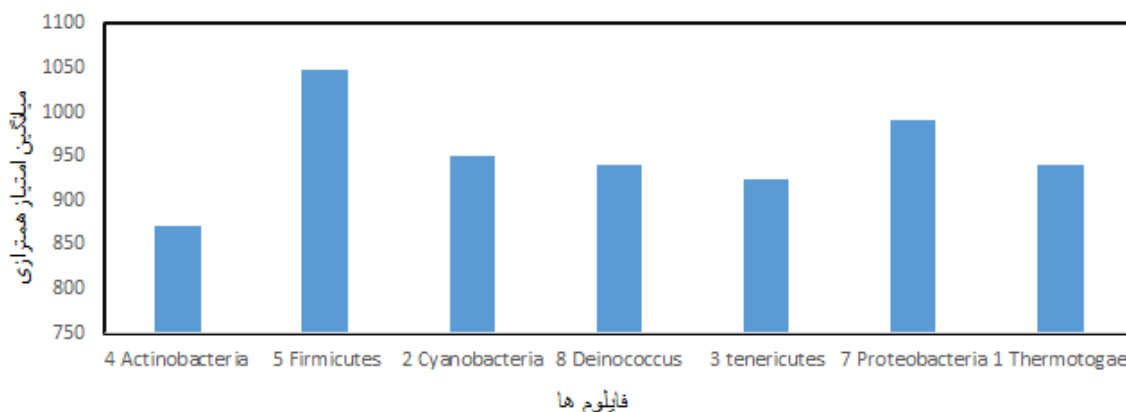


فاصله تکاملی ۳۰ گونه منتخب پروتئین *RecA* (البته در قالب ۸ فایلوم) در مقایسه با ژن *16SrRNA* تغییرات زیادی را نشان می‌دهد. این مسئله بیانگر آن است که میزان تغییراتی که در پروتئین *RecA* و به طبع ژن مربوط به این پروتئین در طول تکامل در این ۳۰ گونه انتخابی در سطح فایلم رخ داده است، با تغییرات در ژن‌های *16SrRNA* نظیر برای گونه‌ها (در سطح فایلم) مطابقت ندارد و مقدار متفاوتی را نشان می‌دهد. میزان این تغییرات در جدول ۱ و نمودار شکل ۶ مشخص شده است.

توالی‌های ژن *16SrRNA* همان جنس و گونه‌ها بوده است. به این منظور توالی‌های پروتئینی *RecA* با گونه یکسان با ژن‌های *16SrRNA* در درخت *16S*-تاکسونومی جایگزین شدند. سپس توالی‌های آمینواسیدی هر فایلم با توالی فایلم سرگروه (*Bacteroides*) همترازسازی و مقدار AAS برای هر فایلم محاسبه شد. نتایج حاصل از این محاسبات برای نماینده‌های خانواده *RecA* (بخش باکتریایی) در شکل ۶ قابل ملاحظه است.



شکل ۵. میانگین امتیاز همترازی توالی‌های *16srRNA* هر فایلم با توالی سرگروه درخت *16S*-تاکسونومی



شکل ۶. میانگین امتیاز همترازی توالی‌های *RecA* هر فایلم با توالی فایلم سرگروه

*Actinobacteria* کمتر بوده است. بنابراین می‌توان گفت که تغییرات *RecA* در طول تکامل در فایلم‌های مختلف سرعت‌های متنوعی داشته است و این سرعت‌های متنوع در بیشتر موارد به طور کلی تفاوت‌های اساسی با الگوی تغییرات در ژن‌های *16SrRNA* نظیر دارد. فایلم بعدی که دستخوش بیشترین تغییرات شده است، فایلم *Proteobacteria* است. فایلم سیانوباکتر *Cyanobacteria* در هر دو درخت جایگاه ثابت دارد. فایلم‌های *Deniococcus*

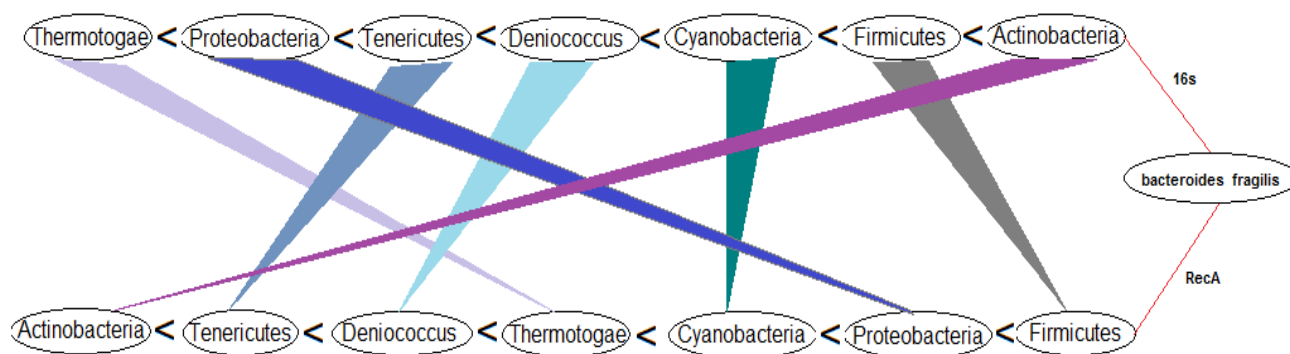
همان‌طور که در شکل ۶ قابل ملاحظه است، فایلم *Actinobacteria* در درخت *16S*-تاکسونومی نزدیک‌ترین فایلم به فایلم سرگروه است، اما در مورد *RecA* دورترین فایلم است. این مسئله نشان می‌دهد ژن *RecA* در فایلم *Actinobacteria* در مقایسه با ژن *16SrRNA* تغییرات بسیار بیشتری در طول تکامل خود داشته است. در حالی که در باقی فایلم‌هایی که در شکل دیده می‌شوند، میزان تغییرات در طول دوره تکاملی در مقایسه با فایلم

شباهت با سرگروه در ژن 16SrRNA و پروتئین RecA در شکل ۷ نشان داده شده است.

Tenericutes هر دو به اندازه یک فایلوم در درخت RecA نسبت به درخت 16S-تاکسونومی از سرگروه فاصله گرفته اند. توالی میزان

جدول ۱. مقدار AAS محاسبه شده برای فایلوم‌های ژن 16SrRNA و پروتئین RecA

مقدار AAS برای توالی‌های پروتئینی RecA در فایلوم‌های منتخب	مقدار AAS برای توالی‌های 16SrRNA در فایلوم‌های منتخب	اسامی فایلوم‌ها
۸۷۰	۹۵۲	Actinobacteria
۱۰۴۷	۹۴۹	Firmicutes
۹۴۹	۸۷۴	Cyanobacteria
۹۴۰	۸۷۴	Deinococcus
۹۲۴	۸۷۲	Tenericutes
۹۹۱	۸۷۱	Proteobacteria
۹۴۱	۸۳۸	Thermotogae



شکل ۷. نامعادله تعیین شده بر اساس محاسبات AAS در فایلوم‌های منتخب برای توالی‌های 16SrRNA و پروتئینی RecA

جدید گونه‌های میکوباکتریوم از شباهت توالی‌های نوکلئوتیدی 16SrRNA استفاده کردند. برای این منظور توالی‌های 16SrRNA از گونه‌های مختلف میکوباکتریوم را همترازسازی کردند که نتیجه حاصل نشان‌دهنده میزان شباهت ۸۵ الی ۹۵ درصدی بین توالی‌ها بود. به همین علت برای تفسیر دقیق‌تر به بررسی توالی‌ها در سطح درون‌گونه‌ای و بین‌گونه‌ای پرداختند. پس از انجام همترازسازی دوگانه توالی‌های 16SrRNA در سطح بین‌گونه‌ای، متوجه شباهت ۹۵ الی ۹۸ درصدی در خانواده میکوباکتریوم شدند. همچنین میزان شباهت برای توالی‌های 16SrRNA درون‌گونه‌ای ۹۸ درصد به بالا بود (۱۱).

در این پژوهش نیز ما به بررسی روند تکاملی خانواده پروتئینی RecA در مقایسه با توالی‌های 16SrRNA نظیر پرداختیم. پس از

## بحث و نتیجه‌گیری

مقایسه نتایج فیلوژنتیک برای جانداران مشخص با استفاده از ژن‌های مختلف می‌تواند تعیین کند کدام ژن‌ها برای مطالعات تکاملی مفید هستند؛ لذا اینکه آیا تاریخچه ژن‌های متنوع با یکدیگر متفاوت است یا خیر در این مورد سوال بسیار مهمی به حساب می‌آید. با این حال، برای مقایسه‌های مستقیم مهم است که متغیرهای زیادی را در مطالعات ژن‌های مختلف حذف کنیم (۲۰). بسیاری از محققان برای مطالعه سیستماتیک باکتریایی، درخت فیلوژنتیک یک ژن خاص را با درخت استاندارد یک ژن rRNA مقایسه می‌کنند (۲۱). در دهه گذشته، توالی‌های ژن 16SrRNA برای مشخص کردن تفاوت‌های درون و بین‌گونه‌ای در بعضی از خانواده‌های باکتریایی کاربرد داشته است (۲۲). در سال‌های اخیر به منظور دسته‌بندی

به این صورت که وقتی نرخ تغییرات طی تکامل برای گونه‌ای بیشتر باشد، داروهای موجود برای از بین بردن این گونه باکتریایی ممکن است (طی مدت طولانی) در مقایسه با سایر گونه‌ها که در طول تکامل تغییراتشان کندتر بوده، زودتر ناکارآمد شوند.

با توجه به اهمیت روند تکاملی و تغییرات باکتری‌ها و ژن‌های آن‌ها که باعث ایجاد مقاومت دارویی می‌شوند، در این پژوهش برای بررسی تغییرات پروتئین *RecA* در مقایسه با ژن حفظ‌شده *16SrRNA* در سطح فایلووم‌های باکتریایی، درخت تاکسونومی نوینی برای جنس و گونه‌های منتخب ژن *16SrRNA* و خانواده پروتئینی *RecA* ارائه شده است که با الگوریتم‌های موجود بیوانفورماتیکی متفاوت است. درنهایت بر اساس محاسبات انجام شده در این پژوهش می‌توان شاهد تغییرات چشمگیر *RecA* در مقایسه با *16SrRNA* در طول تکامل در فایلووم *Actinobacteria* بود. جنس و گونه‌هایی که در این فایلووم قرار دارند از نظر پزشکی اهمیت زیادی دارند و برای بررسی مقاومت دارویی باکتریایی در اثر تغییرات و جهش‌های عمده گزینه‌های مهمی به حساب می‌آیند. یکی از جنس و گونه‌های شناخته‌شده در این فایلووم *Mycobacterium Tuberculosis* است. از جهتی دیگر می‌توان به فایلووم *Cyanobacteria* اشاره کرد که کمترین میزان تغییرات را دارد. در نتیجه احتمال ایجاد مقاومت دارویی مربوط به پروتئین *RecA* در جنس و گونه‌های این فایلووم کمتر از دیگر فایلووم‌های موجود در این پژوهش است. نتایج به‌دست‌آمده در این پژوهش ممکن است به شیوه‌های دیگری هم تفسیر شود. بنابراین می‌خواهیم به استنباط خود از این نتایج به صورت یک فرضیه اشاره کنیم. البته بررسی‌های بیشتر در کارهایی برای مطالعات آینده می‌تواند به شفاف‌سازی بیشتر این استنباط کمک کند.

باید اشاره کنیم نزدیک‌ترین مقاله به پژوهش حاضر کاری است که به وسیله Eisen JA در سال ۱۹۹۵ در مجله تکامل مولکولی چاپ شده است و در بخش مقدمه هم به این مقاله اشاره شد (۱۵). تفاوت‌های اساسی این کار با پژوهش حاضر قابل توجه می‌باشد. به طور مثال حجم داده‌های اولیه به دلیل قدیمی بودن کمتر بوده است. برای کوچک‌سازی تعداد گونه‌ها کار مناسبی انجام نشده است. کل روش کار آنها در مقایسه با پژوهش حاضر کاملاً متفاوت بوده و در نهایت حرفی از *Actinobacteria* که یکی از مهم‌ترین نتایج نهایی پژوهش حاضر می‌باشد یا کمپایل درخت ۱۶S-تاکسونومی زده نشده است. در کل علاوه بر اینکه نتایج نهایی در این دو کار متفاوت است، کمی‌سازی نتایج به هیچ وجه به صورتی که در

استخراج درخت تاکسونومی و قرار دادن توالی‌های *16SrRNA* برای گونه‌های منتخب در قالب این درخت، مقایسه و همترازسازی توالی‌ها در سطح درون فایلومی و بین فایلومی صورت گرفت. نتیجه‌ای که از همترازسازی‌های دوگانه حاصل شد این است که توالی‌های *16SrRNA* درون یک فایلووم شباهت‌های بسیاری با یکدیگر دارند و این شباهت بین ۷۷ تا ۹۸ درصد است. همچنین با مقایسه صورت گرفته (همترازسازی دوگانه) بین نماینده‌های انتخابی هر فایلووم با فایلووم‌های دیگر شاهد کاهش میزان شباهت بین فایلومی هستیم (میزان شباهت بین ۶۹ تا ۸۱ درصد است).

در دیگر مطالعات انجام شده برای بررسی توالی‌های *16SrRNA* در مقایسه با ژن *RecA*، تمام توالی‌های موجود حتی توالی‌های ناقص استخراج‌شده را بررسی کردند (۲۳)، اما در این پژوهش ما از توالی‌های کامل و گونه‌های یکسان برای *16SrRNA* و خانواده پروتئینی *RecA* استفاده کرده ایم تا نتایج کامل‌تری ارائه شود. همچنین در تحقیقات دیگر از الگوریتم‌های شناخته‌شده برای ترسیم درخت فیلوژنی استفاده شده است (۲۴)، ولی در این پژوهش برای رسم درخت تکاملی با انجام بیش از ۲۰۰ همترازسازی برای ۳۰ توالی و همچنین با کمک درخت خام تاکسونومی استخراج‌شده از آن‌سی.بی.آی.، توانسته‌ایم روشی جدید را برای رسم درخت فیلوژنی برای تعداد توالی‌های محدود ارائه کنیم.

در این پژوهش، پس از انجام محاسبات لازم و اصلاح جایگاه فایلووم‌ها در درخت تاکسونومی و ایجاد درخت ۱۶S-تاکسونومی جایگاه‌های جدیدی برای فایلووم‌ها مشخص شد. سپس به مقایسه و بررسی جایگاه فایلووم‌ها در درخت *RecA*-تاکسونومی پرداخته شد و با درخت رفرنس *16S*-تاکسونومی مقایسه شد. یکی از تفاوت‌های عمده این پژوهش با کارهای مشابه در نحوه استخراج درخت فیلوژنی و بررسی روند تکاملی خانواده پروتئینی *RecA* در مقایسه با ژن ریبوزومی بود. بر اساس محاسبات جدیدی که در این پژوهش ارائه شده است، *RecA* در مقایسه با *16SrRNA* در بعضی از فایلووم‌ها طی روند تکاملی تغییرات بسیار زیادی داشته است.

جایگاه فایلووم *Cyanobacteria* در شکل ۷ ثابت است که نشان‌دهنده ثابت ژنی و کمترین جهش (یا تفاوت تغییرات بین *RecA* و *16SrRNA* در این فایلووم) طی تکامل است. همچنین بر اساس محاسباتی که در این پژوهش انجام شده است، بیشترین تغییر مربوط به فایلووم *Actinobacteria* است. تعیین بیشترین و کمترین میزان تغییرات گونه‌های استفاده‌شده در این پژوهش طی روند تکاملی برای بررسی مقاومت‌های دارویی نیز حائز اهمیت است.

پژوهش حاضر انجام شده است. دنبال نشده و نتایج عددی زیادی ارائه نشده است.

## تعارض منافع

بین نویسندگان تعارض در منافع گزارش نشده است.

## سپاسگزاری

نویسندگان از تمامی کسانی که آنها را در انجام این پژوهش یاری کردند، تشکر و قدردانی می‌کنند.

## References

1. Taghizadeh M, Khoei S, Nikoofar AR, Ghamsari L, Goliaei B. The Role of Rad51 Protein in Radioresistance of Spheroid Model of DU145 Prostate Carcinoma Cell Line. *International Journal of Radiation Research*. 2009; 7(1):19.
2. Chintapalli SV, Bhardwaj G, Babu J, Hadjiyianni L, Hong Y, Todd GK, et al. Reevaluation of the Evolutionary Events Within RecA/RAD51 Phylogeny. *BMC Genomics*. 2013; 14(1):240.
3. Wang TF, Chen LT, Wang AH. Right or Left Turn? RecA Family Protein Filaments Promote Homologous Recombination Through Clockwise Axial Rotation. *Bioessays*. 2008; 30(1):48-56.
4. Shinohara A, Ogawa H, Ogawa T. Rad51 Protein Involved in Repair and Recombination in *S. Cerevisiae* is a RecA-Like Protein. *Cell*. 1992; 69(3):457-470.
5. Bishop DK, Park D, Xu L, Kleckner N. DMC1: A Meiosis-Specific Yeast Homolog of *E. Coli* RecA Required for Recombination, Synaptonemal Complex Formation, and Cell Cycle Progression. *Cell*. 1992; 69(3):439-456.
6. Brendel V, Brocchieri L, Sandler SJ, Clark AJ, Karlin S. Evolutionary Comparisons of RecA-Like Proteins Across All Major Kingdoms of Living Organisms. *Journal of Molecular Evolution*. 1997; 44(5):528-541.
7. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A New View of the Tree of Life. *Nature Microbiology*. 2016; 1(5):16048.
8. Gevers D, Cohan FM, Lawrence JG, Spratt BG, Coenye T, Feil EJ, et al. Re-Evaluating Prokaryotic Species. *Nature Reviews Microbiology*. 2005; 3(9):733.
9. Godfray HC. Challenges for Taxonomy. *Nature*. 2002; 417(6884):17.
10. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools. *Nucleic Acids Research*. 2012; 41(D1):D590-6.
11. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, et al. Uniting the Classification of Cultured and Uncultured Bacteria and Archaea Using t rRNA Gene Sequences. *Nature Reviews Microbiology*. 2014; 12(9):635.
12. Beye M, Fahsi N, Raoult D, Fournier PE. Careful Use of 16S rRNA Gene Sequence Similarity Values for the Identification of Mycobacterium Species. *New Microbes and New Infections*. 2018; 22:24-9.
13. Edgar RC. MUSCLE: A Multiple Sequence Alignment Method With Reduced Time and Space Complexity. *BMC Bioinformatics*. 2004; 5(1):113.
14. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution*. 2011; 28(10):2731-2739.
15. 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. *Journal of Molecular Evolution*. 1995; 41(6):1105-1123.
16. Hofmann K, Bucher P, Falquet L, Bairoch A. The PROSITE Database, Its Status in 1999. *Nucleic Acids Research*. 1999; 27(1):215-219.
17. Li W, Godzik A. Cd-hit: A Fast Program for Clustering and Comparing Large Sets of Protein or Nucleotide Sequences. *Bioinformatics*. 2006; 22(13):1658-1659.
18. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: Accelerated for Clustering the Next-Generation Sequencing Data. *Bioinformatics*. 2012; 28(23):3150-3152.
19. Federhen S. The NCBI Taxonomy Database. *Nucleic Acids Research*. 2011; 40(D1):D136-143.
20. Coenye T, Vandamme P. Use of the Genomic Signature in Bacterial Classification and Identification. *Systematic and Applied Microbiology*. 2004; 27(2):175-185.
21. Stackebrandt E, Goebel BM. Taxonomic Note: A Place for DNA-DNA Reassociation and 16S rRNA Sequence Analysis in the Present Species Definition in Bacteriology. *International Journal of Systematic and Evolutionary Microbiology*. 1994; 44(4):846-849.
22. Woo PC, Tsoi HW, Leung KW, Lum PN, Leung AS, Ma CH. Identification of Mycobacterium Neoaurum Isolated From a Neutropenic Patient With Catheter-Related Bacteremia by 16S rRNA Sequencing. *Journal of Clinical Microbiology*. 2000; 38(9):3515-3517.
23. Berson AE, Peters MR, Waleh NS. Nucleotide Sequence of RecA Gene of *Aquaspirillum Magnetotacticum*. *Nucleic Acids Research*. 1990; 18(3):675.
24. Chintapalli SV, Bhardwaj G, Babu J, Hadjiyianni L, Hong Y, Todd GK, et al. Reevaluation of the Evolutionary Events Within RecA/RAD51 Phylogeny. *BMC Genomics*. 2013; 14(1):240.