



## ره یافت بیشینه‌ی درست‌نمایی (maximum likelihood) در فیلوژنتیک

بنفشه صفایی فراهانی<sup>۱</sup>، رضا مستوفی‌زاده قلمفرسا<sup>۲</sup>، آزاده حبیبی<sup>۳</sup>

<sup>۱</sup> استادیار، بخش تحقیقات گیاهپزشکی، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان فارس، سازمان تحقیقات، آموزش و ترویج کشاورزی، شیراز، ایران، b.safae@areeo.ac.ir

<sup>۲</sup> استاد، بخش گیاه‌پزشکی، دانشگاه شیراز، rmostofi@shirazu.ac.ir

<sup>۳</sup> استادیار، گروه تنوع زیستی، پژوهشگاه علوم و تکنولوژی پیشرفته و علوم محیطی، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کرمان، a.habibi@kgut.ac.ir

### چکیده

درختان فیلوژنتیکی روابط خویشاوندی را بین موجودات مختلف به صورت گرافیکی نشان می‌دهند. روش‌های مختلفی برای رسم تبارنماها وجود دارد که برخی از آن‌ها از معیارهای بهینگی برای انتخاب بهترین تبارنما در بین مجموعه‌ی تبارنماهای رسم شده استفاده می‌کنند. یکی از این معیارهای بهینگی، بیشینه‌ی درست‌نمایی است. سه جزء اصلی رهیافت بیشینه‌ی درست‌نمایی شامل داده، مدلی برای توصیف احتمال مشاهده‌ی داده و معیار بیشینه‌ی درست‌نمایی است. در این روش، تبارنمایی که بیش‌ترین احتمال وقوع را داشته است از میان مجموعه‌ی تبارنماهای رسم شده انتخاب می‌شود. مزیت بزرگ استفاده از رهیافت بیشینه‌ی درست‌نمایی در رسم درختان فیلوژنتیکی، قابل اعتماد بودن نتایج حاصل از آن است، اما علیرغم کارایی استفاده از این رهیافت در واکاوی‌های فیلوژنتیکی، بزرگ‌ترین مشکل استفاده از رهیافت بیشینه‌ی درست‌نمایی محاسبات پیچیده و وقت‌گیر آن است. به همین دلیل گاهی به جای بیشینه‌ی درست‌نمایی از سایر رهیافت‌ها مثل بیشینه‌ی پارسمونی استفاده می‌شود.

### واژه‌های کلیدی

تبارنما، معیار بهینگی، بیشینه‌ی درست‌نمایی

### مقدمه

روابط فیلوژنتیکی بین موجودات اغلب با رسم درختان فیلوژنتیکی نشان داده می‌شود. این درختان با استفاده از هم ردیف‌سازی داده‌های حاصل از توالی‌سنجی دی‌ان‌ای به دست می‌آید. تبارنماها را می‌توان با استفاده از روش‌های دسته‌بندی<sup>۱</sup> و روش‌های جست‌وجو<sup>۲</sup> رسم کرد. روش‌های دسته‌بندی مراحل را دنبال کرده و به یک تبارنمای واحد می‌رسند اما بزرگ‌ترین محدودیت این روش‌ها این است که آن‌ها فرضیه‌های مختلف را در مورد تبارنماهای مختلف ارزیابی نمی‌کنند زیرا فقط یک تبارنما رسم می‌کنند. روش‌های جست‌وجو از معیارهای بهینگی مثل بیشینه‌ی پارسمونی<sup>۳</sup> و بیشینه-

ی درست‌نمایی<sup>۴</sup> استفاده کرده، بهترین تبارنما را از بین گروهی از تبارنماها انتخاب می‌کنند. در روش‌های جست‌وجو برای هر تبارنما یک رتبه تعیین می‌شود که در اصل نشان دهنده‌ی رابطه‌ی بین تبارنما و داده‌های نوکلئوتیدی است. مهم‌ترین مانع استفاده از روش‌های جست‌وجو این است که عملیات محاسباتی بسیار پیچیده‌ای دارند. در این روش‌ها باید به دو پرسش پاسخ داد: اول این که برای یک گروه تبارنمای ارائه شده چه چیزی معیار ارزش بهینگی است و دوم این که کدام تبارنما از بین تبارنماهای ارائه شده، حداکثر ارزش معیار بهینگی را دارد. در علوم کشاورزی، رسم درختان فیلوژنتیکی در تشخیص و رده‌بندی گونه‌های بیمارگر گیاهی، گونه‌های بیمارگر دامی، گونه‌های مفید خاک، گونه‌های همزیست و غیره کاربرد دارد [۷].

### فرآیند جایگزینی در توالی دی‌ان‌ای

در اولین نوع طبقه‌بندی، شش حالت مختلف برای جایگزینی در توالی دی‌ان‌ای وجود دارد که عبارتند از جایگزینی منفرد، جایگزینی چندگانه، جایگزینی تصادفی، جایگزینی موازی، جایگزینی همگرا و جایگزینی برگشتی (شکل ۱).

جایگزینی منفرد: در این حالت یک نوکلئوتید توسط یک نوکلئوتید دیگر جایگزین می‌شود. بنابراین یک تغییر سبب مشاهده‌ی یک تفاوت می‌شود.

جایگزینی چندگانه: در این حالت یک نوکلئوتید توسط یک نوکلئوتید دیگر جایگزین می‌شود سپس همان نوکلئوتید توسط یک نوکلئوتید دیگر جایگزین می‌شود. بنابراین دو تغییر سبب مشاهده‌ی یک تفاوت می‌شود.

جایگزینی تصادفی: در این حالت هر نسل یک جایگزینی متفاوت دارد؛ اما علیرغم ایجاد دو تغییر تنها یک تفاوت بین دو نسل دیده می‌شود.

جایگزینی موازی: در این حالت در هر یک از نسل‌ها یک جایگزینی رخ داده اما نوکلئوتیدهایی که در هر نسل جایگزین

<sup>۱</sup> clustering methods

<sup>۲</sup> search methods

<sup>۳</sup> maximum parsimony

<sup>۴</sup> maximum likelihood