

به نام خداوند بخشنده مهربان

روشها و  
پروتکل‌های  
ژنتیکی



# ژنتیکا

[www.genetica.ir](http://www.genetica.ir)

# Phylogeny

## فیلوژنی چیست؟

فیلوژنی<sup>۱</sup> عبارت است از طرح الگوهایی که بیولوژیست‌ها بوسیله آن تنوع و چگونگی پراکنش موجودات زنده را بیان می‌کند.

برای اینکه بیولوژیست‌ها در مورد هر کدام از ارگانیسم‌ها بحث و بررسی کنند باید آنها را به گروه‌هایی طبقه‌بندی نمایند. این طبقه‌بندی باید بر اساس تاریخچه زندگی بنا شده باشد. طبقه‌بندی فقط یکی از زمینه‌های بزرگ فیلوژنی است. متخصصین برای فهمیدن سیر تکامل مشترک موجودات زنده تلاش زیادی کردند و سعی در توضیح دادن راههای متغیر و متنوع زندگی را داشتند تا زمانیکه طبقه‌بندی بعنوان نامی برای گروه‌ها استفاده می‌شود، متخصصین بوسیله آن تئوری‌های سیر تکاملی ارگانیسم را توضیح می‌دهند.

سیستماتیک مطالعه الگوی ارتباطی در طبقه‌بندی گیاهان یا جانوران است و بیولوژیست‌ها از این روش برای ساختن فرضیات یا مدل تاریخی استفاده می‌کنند.

## مقدمه فیلوژنی<sup>۲</sup>

Paleontologists ها علاقه‌مند به فهمیدن زندگی در طول زمان هستند. نه فقط در یک لحظه، در گذشته، یا حال بلکه همه مدت گذشته، قبل از اینکه آنها تلاشی برای تجدید بنا (نوسازی) اشکال، عملکرد و زندگی یک ارگانیسم داشته باشند باید این ارگانیسم‌ها را در یک زمینه قرار بدهند.

زمینه بیولوژی تکامل فیلوژنی که به ارتباط بین موجودات زنده مه به نحوی با هم رابطه اجدادی دارند می‌پردازد و نه تنها فیلوژنی کمک می‌کند (فیلوژنی و دیرینه شناسی مکمل همدیگر هستند و هر دو جریان یکدیگر اتفاق می‌افتد)

---

<sup>1</sup>- Phylogeny

<sup>2</sup>- Introduction to Phylogeny

گروههای زیادی از ارگانیسمها منقرض شدند و بدون فسیلشان نمی‌توانیم به دقت نحوه زندگی آنها را به تصویر بکشیم.

ما نسبت بین گروههای ارگانیسم را در جدول کلادوگرام<sup>۱</sup> توضیح می‌دهیم که همانند ژئولوژی<sup>۲</sup> گونه‌هاست فیلوژنیک شاخه‌ای از فیلوژنی است که بخشی از آن در زمینه سیستماتیک و تاکسونومی<sup>۳</sup> می‌باشد.

### برنامه‌های فیلوژنی<sup>۴</sup>

فیلوژنی دارای برنامه‌های متعددی است که برای دستیابی بایستی در ابتدا از طریق ورود به سایتهای جستجو همانند yahoo, google, altavista وارد شبکه شد و در قسمت search عبارت Phylogeny and software را وارد نمود.



شکل -

همچنین برای دسترسی به برنامه‌ها می‌توان به آدرس زیر مراجعه نمود

<http://phylogeny.arizon.edu/tree/program/program.html>

---

<sup>1</sup> - Cladogram

<sup>2</sup> Genealogies

<sup>3</sup> Taxonomy

<sup>۴</sup> - تاکسونومی دانش نامگذاری و طبقه‌بندی ارگانیسم‌های متغیر است.

<sup>4</sup> - Phylogeny programs

این برنامه‌ها عبارتند از:

- General- purpose package -۱
- Parsimony programs -۲
- Distance matrix methods -۳
- Computation of distances -۴
- Maximum Likelihood and related methods -۵
- Quartets methods -۶
- Invariants (or Evolutionary parsimony) methods -۷
- Interactive tree manipulation -۸
- Biogeography analysis and host-Para site comparison -۹
- Bootstrapping -۱۰
- Compatibility analysis -۱۱
- Tree alignment -۱۲
- Comparative method analysis -۱۳
- Tree plotting/ drawing -۱۴
- Sequence management/ top submission -۱۵
- Teaching about phylogenies -۱۶
- miscellaneous -۱۷

**برنامه General purpose package**

این برنامه خود شامل قسمت‌های مختلفی است که می‌توان به موارد زیر اشاره کرد:

۱- PHYLIP:

این برنامه در طبقه بندی فیلوژنی بسیار وسیع است و بسیار کم هزینه می‌باشد.

۲- PAUP:

برنامه است که با گنجایش بیشتر متدها عریض می‌شود.

۳- MEGA:

این برنامه به آنالیز داده‌ها از سکانس‌های DNA و RNA و سکانس PRO می‌پردازد.

### **برنامه Parsimony programs:**

شامل برنامه‌های زیر می‌باشد:

۱- PAUP:

۲- Hennig 86: این برنامه هزینه بالایی دارد.

### **Distance matrix methods**

۱- Mact: برنامه‌ای است که محدوده (فاصله) ژنتیکی را حساب می‌کند.

۲- ODEN: برنامه‌ای است که سکانس‌های نوکلئوتید یا پروتئین را آنالیز می‌کند.

### **Computation of distance**

۱- RAP Distance: برنامه‌ای است که برای محاسبه محدوده Matrices و آنالیزهای Rapd

۲- Multicomp: یک برنامه‌ای برای محدودیت‌های متفاوت از داده‌های سکانس

### **:Maximum Likelihood and related methods**

۱- Fast DNA ml: برنامه‌ای است که درخت‌های خیلی بزرگ را با اندازه بزرگ و سرعت بالا به

صورت اجتماعات (مجموع) موازی بررسی می‌کند.

۲- STATGEOM: برنامه‌ای است که حساب می‌کند هندسه آماری را در محدوده (فاصله) و

مکان سکانس را در DNA و RNA و اسیدهای آمینه

## Invariants methods

۱- Evomeny: یک برنامه Macintosh است که برای استفاده از روش‌های تکاملی و برای

استنباط کردن فیلوژنی از سکانس DNA و RNA

## Interactive tree manipulation

۱- Mcclade: یک برنامه Macintosh است که برای مته‌های کم هزینه در حالت داخواه است.

## Biogeography analysis comparison

۱- COMPONENT: یک برنامه‌ای است که برای محاسبه Cladogram، برای استفاده در

فیلوژنی و مطالعات Biogeraphy است.

۲- TREE MAP: برنامه‌ای است برای محاسبه فیلوژنی میزبان و انگل

## Bootstrapping

۱- Zhorkikh program: برنامه‌ای است برای Bootstrapping سکانس‌های نوکلئوئید که

شامل نوآوری در تکمیل و حالت جزیی Bootstrapp می‌باشد.

## Compatibility analysis

۱- CLINCH: احتمالاً در مقایسه با برنامه آنالیزی توسعه بیشتری پیدا کرده است.

## Tree alignment

۱- Tree Align: برنامه تنظیمی است که چندین سکانس شاخه‌ای شبیه DNA یا سکانس‌های

پروتئین را می‌سازد.

۲- Clusta LW: برنامه‌ای است که ترتیب چند سکانس را در کاربرد درخت به صورت ترتیب

سکانس‌های متعدد تخمین می‌زند.

## Comparative method analysis

۱- COMPARE: یک برنامه است برای تکمیل متدهای آنالیز و شمایل برنامه‌های مختلفی برای

رفتار آماری، داده نسبی آنالیزها در زمینه فیلوژنی

۲- PA: یک برنامه Macintosh است که ابزار فیلوژنی است.

:Tree plotting/ drawing

۱- Tree View: برنامه‌ای است که برای نمایش درخت‌ها در Apple Macs و Windows  
Pcs است.

۲- Njplot: برنامه‌ای است که نمایش می‌دهد فیلوژنی را در Macintosh و آن را در فایل‌ها  
Pict نگهداری می‌کند.

:Sequence management

۱- Random Cladistics: مجموعه برنامه‌ای است می‌تواند Bootstrapping و Jackkfin و  
انواع مختلف پیش موتاسیون را تست می‌کند.

:Teaching about phylogenies

۱- Phylogenic Investigator: یک برنامه تفکیکی است که به فراگیران اجازه می‌دهد که با  
اتصال به شبکه یک فیلوژنی بسازند و آن را آزمایش کنند.

۲- Systack: یک برنامه تکنیکی است که طراحی می‌کند به طور اساسی Synapomorphy و  
آنالیز homology را در زمینه فیلوژنی

:Miscellaneous

۱- OSA: این برنامه یک سکانس بزرگ را پیدا کرده و نواحی را از نظر اطلاعاتی بسیار شبیه به  
هم هستند نمایش می‌دهد.

از میان این برنامه‌ها، برنامه فیلیپ را انتخاب کرده و بررسی و کاربرد این برنامه می‌پردازیم.

PHYLIP

یک دسته برنامه آزاد در مجموعه فیلوژنی محسوب می‌شود که به صورت کد، فایل‌های  
Document و انواع مختلفی از فایل‌های قابل اجرا توزیع شده است.

برای دسترسی به این برنامه آدرس مورد نظر

<http://evolution.gentics.washington>

Philip program

DNA/RNA Sequence – ۱

Gene Frequencies – ۲

Distance matrices – ۳

Restriction sites – ۴

Continuous quantitative character – ۵

Protein Sequences – ۶

Discrete (0.1) characters – ۷

### **Type of algorithm**

Heuristic search for best tree – ۱

Interactive tree man/ pulation – ۲

Brach- and- bound exact for tree – ۳

DNA and RNA sequence data

۱- DNA PARS: این برنامه با استفاده از سکانس‌های اسید نوکلئیکی و با روش Parsimony

به بیان فیلوژنی می‌پردازد در این برنامه با استفاده از تمام کدهای مبهم IUB حالت نوکلئیدی اجدادی را تخمین می‌زند که این حالت‌های اجدادی مربوط به GAP می‌باشد.

۲- DNA MOVE: این برنامه به تاثیر متقابل فیلوژنی از سکانس‌های اسیدنوکلئیکی می‌پردازد و آنها را با روش‌های Parsimony و سازشی ارزیابی کرده و بازهای اجدادی ساختار بازسازی شده را نشان می‌دهد که این کار را به روش دستی هم انجام می‌دهند.

۳- DNA PENNY: این برنامه اکثراً به روش parsimony برای سکانس‌های اسیدنوکلئیک با جستجوی شاخه و ریشه تشخیص داده می‌شود این ممکن است برای ۱۰ تا ۱۱ گونه باشد.

۴- DNA COMP: در این برنامه کاربرد فیلوژنی از سکانس‌های اسیدنوکلئیکی به عنوان یک معیار با ارزش در جستجوی بیشترین جایگاهی که هر یک از نوکلئوتیدها به تنهایی بر روی شاخه به



خودشان اختصاص می‌دهند استفاده می‌کنند روش سازشی به طور اختصاصی به حالتیکه جایگاهها در نسبت تکاملی خود بسیار متغیر می‌باشد اختصاص می‌گیرد.

۵- DNA INVAR: این برنامه برای داده‌های اسید نوکلئیکی در ۴ گونه ثابت فیلوژنتیک Covenders و Lokes را محاسبه می‌کند که بر این اساس محاسبه توپولوژی درخت را به طور متناوب تست می‌کند این برنامه فرکانس وقوع الگوهای نوکلئوتیدی مختلف را جدول بندی می‌کند.

۶- DNA ML: این برنامه فیلوژنی سکانس‌های نوکلئیکی را با ماکزیمم احتمالات تخمین می‌زند مدل بکار گرفته شده برای فرکانس مورد انتظار ۴ نوکلئوتید برای نسبتهای نامساوی Transversions, Transitions برای نسبت مختلف تغییر در دسته‌بندی مختلف جایگاهها منظور می‌شود.

۷- DNA MLK: این برنامه همان DNA ML است با این فرض که یک سلول مولکولی است استفاده از دو برنامه با هم اجازه می‌دهد که یک احتمال برای آزمون فرضیه ساعت مولکولی ایجاد کنیم.

۸- DNA DISK: این برنامه ۴ محدوده مختلف بین گونه‌ای از توالی از اسید نوکلئیک را محاسبه می‌کند سپس این محدوده‌های بدست آمده را می‌توان در برنامه Mutrix distance مورد استفاده قرار داد. این محدوده‌ها عبارتند از روش Kimuras 2-parameter محدوده jin and Neis که اختلاف از جایگاهی به جایگاه دیگر مد نظر قرار می‌دهد و روش ماکزیمم احتمالات از مدل به کار رفته شده در DNA ML استفاده می‌کند. این روش محاسبه فاصله می‌تواند خیلی کند باشد.

۹- SEQ BOOT: این برنامه یک مجموعه داده را می‌خواند و از آن مجموعه داده مضاعف با روش boot starp تولید می‌کند از آنجاییکه اغلب برنامه‌ها را در Version متداول این Package از مجموع داده‌های مضاعف استفاده می‌کنند این برنامه می‌تواند در برنامه درختی Consense برای

محاسبه آنالیز bootstrap استفاده شود این برنامه از تکنیک Archie/Fahth که باعث تغییر گونه‌ها در داخل صفات می‌شود نیز استفاده می‌کند.

۱۰- CONSENE: این برنامه درخت consensus را با روش قانون درختی Consensus محاسبه می‌کند. این برنامه می‌تواند در آنالیز boots trap برای بسیاری از متدها در طبقه‌بندی استفاده شود.

۱۱- MAKEINF: این برنامه برای ارجمه فایل‌های خروجی از برنامه تنظیمی چندین سکانس jotun.Heins، فایل‌های ورودی Phylip استفاده می‌شود که توسط Arend Sidow طراحی شده است این برنامه به شما اجازه می‌دهد که به طور انتخابی به آنالیز جایگاه‌های کدونی و ارگانسیم‌های مختلف بپردازید. خروجی سایر برنامه‌های تنظیمی به راحتی به یک شکل قابل خواندن تنظیم می‌شود.

۱۲- COALLIKE: این برنامه در Version اخیر PHYLIP بوده اما امروزه از این package برداشته شده است چون روش این برنامه دارای نقص بوده است بنابراین این برنامه را استفاده نمی‌کنند.

### Gene frequencies

۱- CONTML: این برنامه با استفاده از داده فرکانس ژنی با حداکثر احتمالات فیلوژنی را برآورد می‌کند این برنامه نشان می‌دهد که تمام این انشعابات ناشی از diff ژنتیکی در غیاب موتاسیون جدید می‌باشد برای این برنامه ساعت مولکولی در نظر نمی‌گیرند. در این برنامه با محاسبه کردن فاصله ژنتیکی Neis، داده‌ها را به طور متناوب آنالیز می‌کنند و برای این هدفها از یکی از برنامه‌های ماتریکس فاصله‌ای استفاده می‌کنند.

۲- GENDIST: در این برنامه یکی از سه فرمول ژنتیکی مربوط به داده‌های فرکانس ژنی را حساب می‌کند این فرمول‌ها شامل ۱- فاصله ژنتیکی neis ۲- اندازه Cavalli-Sforza و فاصله ژنتیکی Reynolds است که اولی برای داده‌های موتاسیون جدیدی است که به صورت موتاسیون

خنثی ایزوآلل است اما دو مدل بعدی بدون موتاسیون و با یک drift ژنتیکی خالص است این فاصله‌ها در یک فایل با فورمت متناوب با داده‌های برنامه ماتریکس فاصله‌ای نوشته می‌شود.

۳- SEQBOOT: این برنامه یک مجموعه داده را می‌خواند و با استفاده از برنامه boots trap یک مجموعه داده مضاعف ایجاد می‌کند. از آنجاییکه اغلب برنامه‌ها را در Version متداول این package از مجموع داده‌های مضاعف استفاده می‌کنند این برنامه می‌تواند در برنامه درختی CONSENSE برای محاسبه آنالیز boots trap استفاده شود. این برنامه از تکنیک Archie/faith که باعث تغییر گونه‌ها در داخل صفات بیشتر نیز استفاده می‌کند.

### **Distance matrix data**

FITCH: این برنامه فیلوژنی از داده ماتریکس فاصله‌ای و توسط مدل درختی additive تخمین می‌زند که برآورد فاصله‌های بدست آمده با مجموع طول شاخه‌های بین گونه‌ای برابر است که در این برنامه از معیار Fitch-Morgoliush و معیار وابسته به کمترین توان استفاده می‌کند در اینجا ساعت تکاملی در نظر گرفته نمی‌شود این برنامه برای فاصله‌های محاسبه شده از سکانس DNA، اندازه گیری هیبریداسیون DNA و فاصله ژنتیکی محاسبه شده از فرکانس ژنی مفید خواهد بود.

۲- KITSCH: این برنامه فیلوژنی را از داده ماتریکس فاصله‌ای و توسط مدل Ultrametric تخمین می‌زند این روش شبه مدل درختی additive است به جز اینکه در اینجا ساعت تکاملی در نظر گرفته می‌شود در اینجا هم معیار Fitch-Margoliush و سایر معیارهای کمترین توان در نظر گرفته می‌شود این برنامه برای محاسبه فاصله‌ها از روی توالی DNA، اندازه‌گیر هیبریداسیون DNA و فاصله ژنتیکی بدست آمده از فرکانس ژنی مفید خواهد بود.

۳- NEIGHBOR: این برنامه توسط john yamoto, Mary kuhner, Neis neighbor روش ماتریکس محدود است که یک درخت غیر رشته‌ای را بدون در نظر گرفتن یک ساعت ایجاد می‌کند. UPGMA را با ساعت مد در نظر می‌گیرند. طول شاخه‌ها به وسیله معیار

کمترین توان بهینه نمی‌شود اما روشهایی هستند که بسیار سریع عمل کرده و بنابراین قادرند مجموعه زیادی از اعداد را بکار گیرند.

### **Restriction sites data**

۱- RESTM: این برنامه شبیه به RESTML در Heuristic Search می‌باشد.

۲- SEQBOOT: این برنامه شبیه به SEQBOOT در DNA/RNA Sequen می‌باشد.

۳- CONSENSE: این برنامه شبیه به CONSENSE در DNA/RNA می‌باشد.

اگر شما قادر به محاسبه ماتریکس محدوده‌ای با روش  $neis, Li$  باشید شما بعداً می‌توانید از روش ماتریکس محدوده‌ای برای آنالیز جایگاههای آنزیم محدود الاثر استفاده کنید.

### **Continuous quantitive characters**

۱- CONTM: این برنامه برآورد فیلوژنی ناشی از مجموعه صفات مداوم با روش حرکت Brownian می‌پردازد این برنامه برای تغییر شکل صفات مناسب می‌باشد این تغییرات به این صورت است که مجموعه صفات از نظر ژنتیکی مستقل‌اند اما تحت یک فشار انتخابی هم ارز قرار می‌گیرند.

۲- CONTRAST: این برنامه یک درخت از فایل‌های درختی را می‌خواند و یک مجموعه داده با صفات دائمی ایجاد می‌کند بطوریکه این صفات از هم مستقل هستند که این صفات در هر یک از Package آماری استفاده می‌شود و برای این صفات مستقل Correlations با ضریب همبستگی و Covarinces را بیان می‌کند.

### **Protein Sequences**

۱- PROTPARS: این برنامه فیلوژنی را از سکانس‌های پروتئینی با روش parsimony (از یک کد حرفی استاندارد برای آمینواسیدها استفاده می‌کنند) برای این امر تنها آن نوکلئوتیدهایی را

محاسبه می‌کند که تغییر در آنها منجر به تغییر در یک آمینو اسید می‌شود با این فرض که تغییرات خاموش به راحتی ترمیم می‌شود.

۲- PROTDIST: این برنامه یک اندازه محدود را برای سکانسهای پروتئین حساب می‌کند که احتمالاً بیشترین استفاده بر اساس Dayhoff PAM می‌باشد که Kimuras 1983 شباهت زیادی به آن دارد یافته‌هایی که بر اساس کدهای ژنتیکی تغییرات مختلفی را در طبقه اسیدهای آمینه ایجاد می‌کند.

۳- SEQ BOOT: این برنامه شبیه به برنامه SEQ BOOT در gene frequensis می‌باشد

۴- CONSENSE: این برنامه شبیه به برنامه CONSENSE در برنامه DNA/RNA Sequen می‌باشد.

۵- MAKEINF: این برنامه شبیه به برنامه MAKEINF در برنامه DNA/RNA Sequen می‌باشد.

۶- PROTML: این برنامه بزرگ pascal که توسط Masami Hasegawa , Jun Adachi طراحی شده فیلوژنی را از سکانس پروتئینی با حداکثر احتمالات تخمین می‌زند این برنامه آنالوگ DNAML محسوب می‌شود اما به جای مدل تکاملی DNA، مدل ماتریکسی DNA طراحی شده توسط Margaret Dayhoff را دارا می‌باشد سرعت این برنامه در مقایسه با برنامه DNAML به دلیل افزایش تعداد داده‌ها (۲۰ به جای ۴) کمتر می‌باشد با وجود این، طراحان این برنامه تغییراتی در آن ایجاد کرده و سرعت آن را افزایش دادند و آن را مجعز به استراتژی بازآرایی نموده و به این وسیله بهترین حالت از میان توپولوژی درختی انتخاب می‌کنند.

### **Discrsts(0,1) Characters**

۱- MIX: این برنامه شبیه به برنامه MIX در برنامه Heuristic Search می‌باشد.

۲- MOVE: این برنامه تشکیلات متقابل فیلوژنی را با روش سازشی و Parsimony ارزیابی کرده و حالت‌های بازسازی شده و تمام درخت را نشان می‌دهد این برنامه، برآورد فیلوژنی با روش دستی هم استفاده می‌شود.

۳- این برنامه شبیه به برنامه PENNY در روش branch-and-bound می‌باشد.

۴- DOLLOP: این برنامه شبیه DOLLOP در برنامه Heuristic Search می‌باشد.

۵- DOLMOVE: در این برنامه ساختارهای متقابل فیلوژنی ناشی از مجموعه صفات مورد نظر در ۲ حالت (۱ و ۰) را با استفاده از معیار Dollo، Parsimony و Polymorphism نشان می‌دهد و معیار سازشی و Parsimony را برای فیلوژنی ارزیابی می‌کند.

۶- DOLPENNY: این برنامه شبیه Dolpenny در branch-and-bound می‌باشد.

### **:Plot trees, edit trees, or make consensus**

۱- ORAWGRAN: این برنامه در یک محدوده وسیع از فرصت‌های قابل کنترلی فیلوژنی، Cladogram و فتوگرام را رسم می‌کند. این برنامه یک برنامه با تاثیر متقابل است و قادر است ساخت شجره‌ای را بر روی Tekrinx, screen یا گرافیک‌های DEC, PC نمایش دهد.

۲- DRAWTREE: این برنامه، مشابه برنامه DRAWTREE است اما با این تفاوت که فیلوژنی درخت‌های غیر ریشه‌ای را رسم می‌کند.

۳- CONSENSE: این برنامه شبیه CONSENSE در DNA/RNA می‌باشد.

۴- RETREE: این برنامه شبیه RETREE در برنامه Interactive tree می‌باشد.

### **Heuristic Search for best tree**

**PROTPARS**: این برنامه فیلوژنی را از سکانس‌های پروتئینی با روش Parsimony (از یک کد حرفی استاندارد برای آمینو اسیدها استفاده می‌کنند) برای این امر تنها آن نوکلئوتیدهایی را محاسبه می‌کند که تغییر در آنها منجر به تغییر در یک آمینو اسید می‌شود با این فرض که تغییرات خاموش به راحتی ترمیم می‌شود.

**DNAPARS:** این برنامه با استفاده از روش Pursimony فیلوژنی را از سکانس‌های اسید نوکلئیکی محاسبه می‌کند در اینجا از تمام کدهای مبهم IUB استفاده می‌کند و حالت‌های نوکلئوتید اجدادی

را محاسبه می‌کند در اینجا 1/5 حالت‌های نوکلئوتیدی به صورت GAP در نظر گرفته شود.  
**DNACOMP:** این برنامه برآورد فیلوژنی از داده‌های سکانس اسید نوکلئیک استفاده از یک معیار سازشی است که بزرگترین تعداد جایگاه‌هایی که هر یک از نوکلئوتیدها می‌توانند بر روی همان شاخه به خود اختصاص دهند جستجو می‌کند این حالت سازش خصوصاً در زمانی که در یک نسبت تکاملی خیلی بزرگ هستند بسیار مناسب است اما اعتبار آن را نمی‌دانیم.

**DNAML:** این برنامه فیلوژنی از داده‌های سکانس‌های نوکلئیکی را با ماکزیمم احتمالات تخمین می‌زند مدل به کار گرفته شده برای فرکانس‌های مورد انتظار ۴ نوکلئوتید برای نسبت‌های نامساوی transversion transition برای نسبت مختلف تغییر در دسته بندی مختلف جایگاهها منظور می‌شود.

**DNAMLK:** این برنامه همان DNAML است اما با این فرض که یک ساعت مولکولی استفاده از دو برنامه با هم اجازه می‌دهد که یک نسبت احتمالی برای آزمون فرضیه ساعت مولکولی ایجاد کنیم.

**RESTML:** این برنامه با استفاده از جایگاه‌های آنزیم محدودالثر به برآورد فیلوژنی با حداکثر احتمالات می‌پردازد ( در اینجا قطعات تحدیدی وجود ندارد اما حضور یا عدم حضور جایگاه‌های خاص مورد نظر است در اینجا قرینه‌ای (jukes – Contor) تغییر نوکلئوتیدی به کار گرفته می‌شوند و به تفاوت نسبت بین transition , transversion ربطی ندارد.

**FITCH:** این برنامه فیلوژنی از داده ماتریکس فاصله‌ای و توسط مدل درختی additive تخمین می‌زند که این برآورد با فاصله‌های به دست آمده با مجموع طول شاخه‌های بین گونه‌ای برابر است که در این برنامه از معیار Fitch – Margoliash و معیار وابسته به کمترین توان استفاده

می‌کنند در اینجا ساعت تکاملی در نظر گرفته نمی‌شود این برنامه برای فاصله‌های محاسبه شده از سکانس DNA، اندازه‌گیری هیبریداسیون DNA و فاصله ژنتیکی محاسبه شده از فرکانس ژنی مفید خواهد بود.

**KITSCH**: این برنامه شبیه به برنامه KITSCH در matrize, distance می‌باشد.

**NEIGHBOR**: این برنامه شبیه به برنامه NEIGHBOR در distance matrize می‌باشد.

**CONTML**: این برنامه شبیه به برنامه CONTML در gene frequencies می‌باشد.

**MIX**: این برنامه با استفاده از روش Parsimony، فیلوژنی را برای صفات مورد نظر با دو حالت (۰ و ۱) محاسبه می‌کند در اینجا از روش Wanger Parsimony و روش Comin – sokal یا مخلوطی از این دو روش استفاده می‌شود در این برنامه داده‌های اجدادی را بازسازی کرده و اندازه صفات را هم مشخص می‌کند.

**DOLLOP**: در این برنامه با استفاده از Parsimony, Polymorphism یا Dollo، فیلوژنی را برای صفات مورد نظر در حالت (۰ و ۱) برآورد می‌کند در اینجا هم اجدادی را بازسازی کرده و اندازه صفات را مشخص می‌کند روش Parsimony, Dollo برای جایگاههای حالات محدودالاطر اختصاصی شده، در مورد داده‌های اجدادی نامشخص این روش قطعات حاصل از آنزیم محدودالاطر را مشخص می‌کند.

**PROTML**: این برنامه بزرگ pascal که توسط Masami Hasegawa, Jun Adachi طراحی شده فیلوژنی را از سکانس پروتئینی با حداکثر احتمالات تخمین می‌زند این برنامه آنالوگ DNAML محسوب می‌شود اما به جای مدل تکاملی DNA، مدل ماتریکسی DNA طراحی شده توسط Margaret Dayhoff را دارا می‌باشد سرعت این برنامه در مقایسه با برنامه DNAML به دلیل افزایش تعداد داده‌ها (۲۰ به جای ۴) کمتر می‌باشد با وجود این، طراحان این برنامه تغییراتی در آن ایجاد کرده و سرعت آن را افزایش دادند و آن را مجعز به استراتژی بازآرایی نموده و به این وسیله بهترین حالت از میان توپولوژی درختی انتخاب می‌کنند.



## Interactive tree manipulation

۱- **DNAMOVE**: در این برنامه به تاثیر متقابل فیلوژنی از سکانس‌های اسیدنوکلیدی با روش‌های Parsimony و سازشی ارزیابی کرده و بازهای اجدادی بازسازی شده را نشان می‌دهد.

۲- **MOVE**: این برنامه تاثیر متقابل ساختارهای فیلوژنی از صفات مورد نظر با دو حالت ۰ و ۱ را بیان می‌کند و آنها را با معیار سازشی و Parsimony ارزیابی کرده و حالت‌های بازسازی شده را به صورت یک درخت نشان می‌دهد این روش برای تخمین نشان می‌دهد که این کار را به روش دستی هم انجام می‌دهند Parsimony و سازشی به صورت دستی و دلخواه هم استفاده می‌شود.

۳- **Dolomove**: در این برنامه تاثیر متقابل ساختارهای فیلوژنی ناشی از صفات مورد نظر با ۲ حالت (۱ و ۰) را بررسی می‌کند و این بررسی روش Parsimony، پلی مورفیسم یا Dollo انجام می‌شود در این جا هم داده‌های بازسازی شده به صورت درختی نشان داده می‌شود این برنامه هم به روش دستی انجام می‌شود.

۴- **RETREE**: این برنامه درخت را خوانده و اگر لازم باشد طول شاخه‌ها را هم بررسی می‌کند و به شما اجازه می‌دهد تا درخت را دوباره ریشه یابی کنید. شاخه‌ها را تغییر داده، نام گونه و طول شاخه را تغییر دهید و نتیجه را به عنوان فایل خروجی برای شما می‌نویسد.

## Branch-and-bound exact for tree

۱- **DANAPENNY**: در این برنامه اکثر فیلوژنی‌ها به روش Parsimony اسیدهای نوکلیدی را با جستجوی شاخه و ریشه پیدا می‌کند.

۲- **PENNY**: این برنامه با روش تحقیقی ریشه و شاخه Branch-and-bound و با استفاده از معیارهای Camin-Sokal parsimony و Wagner و یا هر دو به بررسی فیلوژنی، مدل Parsimony صفات مورد نظر در ۲ حالت می‌پردازد.

۳- **DOLPENNY**: این برنامه با روش دقیق، شاخه و ریشه و با استفاده از معیار Parsimony, Dollo یا Polymorphism به بررسی فیلوژنی مدل Parsimony در صفات مورد نظر در دو حالت می‌پردازد.

۴- **CLQUE**: این برنامه بزرگترین دسته صفات سازشی متقابل را پیدا کرده و برای صفات مورد نظر در دو حالت فیلوژنی را بیان می‌کند برای این امر از روش تحقیقی بسیار سریع Branch bound استفاده می‌کنند. این روش برای داده‌های اشتباه صدق نمی‌کند.

مثال: شجره نامه دو خانواده از Segment Rattus norvegicus genomy B.burgdorferi را می‌کشیم

### نرم‌افزار ClustalW

این نرم‌افزار یکی از نرم‌افزارهای آن‌لاین<sup>۱</sup> تطبیق سکانسهای ژنی و یا پروتئینی<sup>۲</sup> است. در منوی آبی رنگ بالا نام ۳ نرم‌افزار آن‌لاین ملاحظه می‌شود که در ابتدا به بررسی نرم‌افزار ClustalW پرداخته می‌شود. در حالی که نرم‌افزار clustalW باز است، اطلاعات سکانسهای ژنی یا پروتئینی را بصورت فایل‌های پذیرفته شده<sup>۳</sup> بطور پشت هم در قسمت مورد نظر وارد می‌کنیم (این قسمت در شکل زیر توسط علامت پیکان مشخص شده‌است). به این صورت که در ابتدا یک سکانس را وارد می‌کنیم، از طریق صفحه کلید، کلید Inter را می‌زنیم تا چشم‌کزن در سر سطر قرار گیرد. به این ترتیب می‌توان چند سکانس را پشت سر هم قرار داد. همچنین می‌توان برای وارد کردن یک سکانس بر روی گزینه Brows کلیک کرده و فایل مورد نظر خود را وارد نمائید.

---

1 online

2 Multiple sequence alignment

3 FASTA (Pearson), NBRF/PIR, EMBL/Swiss Prot, GDE, CLUSTAL, and GCG/MSF

شکل –

### : Output Format –

: ClustalW

: GCG (MSF)

: GDE

: PIR

: Phylip

### : Pairwise Alignment –

: FAST/APPROXIMATE

: SLOW/ACCURATE

در قسمت Enter your sequence مشخص می‌کنیم نوع سکانس را مشخص می‌کنیم که DNA باشد و یا پروتئین. برای ریست کردن داده‌ها بر روی گزینه Reset کلیک نمائید.

در قسمت بعد با More Detail Parameters برخورد می‌کنیم. می‌توان پارامترهای مورد نظر را در این قسمت تنظیم نمود.

**More Detail Parameters...**

**Pairwise Alignment Parameters:**

*For FAST/APPROXIMATE:*  
 K-tuple(word) size: , Window size: , Gap Penalty:   
 Number of Top Diagonals: , Scoring Method:

*For SLOW/ACCURATE:*  
 Gap Open Penalty: , Gap Extension Penalty:   
 Select Weight Matrix:

(Note that only parameters for the algorithm specified by the above "Pairwise Alignment" are valid.)

**Multiple Alignment Parameters:**

Gap Open Penalty: , Gap Extension Penalty:   
 Weight Transition: ☒ YES (Value: ) , ☐ NO  
 Hydrophilic Residues for Proteins:   
 Hydrophilic Gaps: ☒ YES ☐ NO  
 Select Weight Matrix:

Type **additional options** (delimited by whitespaces) below:  
 [-options for help]

در قسمت Pairwise Alignment Parameters پارامترهای زیر قرار دارند:

**: For FAST/APPROXIMATE-**

K-tuple (word) size

Window size

Gap Penalty

Number of Top Diagonals

Scoring Method

: **For SLOW/ACCURATE-**

Gap Open Penalty

Gap Extension Penalty

Select Weight Matrix

: **Multiple Alignment Parameters-**

Gap Open Penalty

Gap Extension Penalty

Weight Transition

Hydrophilic Residues for Proteins

Hydrophilic Gaps

Select Weight Matrix

Type additional options

در پایان گزینه Execute Multiple Alignment را کلیک می‌کنیم.

## نرم افزار MAFFT

(ویژگی و کاربرد)

Multiple sequence alignment by MAFFT ver.5

### Basic settings

#### Input:

Paste protein or DNA sequences in fasta format

or upload a file:

Browse...

Submit

Reset

شکل -

**Advanced settings**

**Strategy:**

- ☐ FFT-NS-1 (Very fast; progressive method)
- ☐ FFT-NS-2 (Fast; progressive method)
- ☒ (Medium; iterative refinement method, two cycles only)
- ☐ FFT-NS-i (Slow; iterative refinement method)
- ☐ F-INS-i (Very slow; iterative refinement method, pairwise alignment information by FASTA is included)

**Parameters:**

Scoring matrix (for amino acid sequences):

Gap opening penalty:  (1.0-3.0)

Offset value:  (0.0-1.0; functions like gap extension penalty)

**mafftE (aligns input sequences together with their homologues collected from SwissProt by BLAST):**

☐ On

☐ Show collected homologues also (if any)

Number of homologues:  (5-200)

Threshold:  ( $1e-5$  -  $1e-40$ )

شکل -

### Strategy:-

FFT-NS-1

FFT-NS-2

Medium

FFT-NS-i

F-INS-i

### Parameterers:-

Scorning matrix

Gap opening penalty

Offset value

**mafftE-:**

On

Show collected homologues also

Number of homologues

Threshold

**برنامه PRRN**

توضیح و تفسیر

| CLUSTALW                              | MAFFT                                                              | PRRN |
|---------------------------------------|--------------------------------------------------------------------|------|
| Help                                  |                                                                    |      |
| <b>Specify Parameters :</b>           |                                                                    |      |
| Sequence Type :                       | <input checked="" type="radio"/> PROTEIN <input type="radio"/> DNA |      |
| Score Matrix :                        | PAM250                                                             |      |
| Gap Penalty (PROTEIN) :               | Open: 9.0 , Extend: 2.0 , Background: 1.0 (0 .. 100)               |      |
| Gap Penalty (DNA) :                   | Open: 5.0 , Extend: 2.0 , Background: 2.0 (0 .. 100)               |      |
| Output Format :                       | Native                                                             |      |
| Method :                              | Progressive (pairwise alignment) + Iterative refinement            |      |
| Window Size :                         | 100 (0 .. 1000)                                                    |      |
| Score Matrix (Conserved Regions) :    | PAM100                                                             |      |
| Threshold Value (Conserved Regions) : | 25.0 (0 .. 1000)                                                   |      |

شکل -

| CLUSTALW                              | MAFFT                                                              | PRRN |
|---------------------------------------|--------------------------------------------------------------------|------|
| Help                                  |                                                                    |      |
| <b>Specify Parameters :</b>           |                                                                    |      |
| Sequence Type :                       | <input checked="" type="radio"/> PROTEIN <input type="radio"/> DNA |      |
| Score Matrix :                        | PAM250                                                             |      |
| Gap Penalty (PROTEIN) :               | Extend: 2.0 , Background: 1.0 (0 .. 100)                           |      |
| Gap Penalty (DNA) :                   | Extend: 2.0 , Background: 2.0 (0 .. 100)                           |      |
| Output Format :                       | Native                                                             |      |
| Method :                              | Progressive (pairwise alignment) + Iterative refinement            |      |
| Window Size :                         | 100 (0 .. 1000)                                                    |      |
| Score Matrix (Conserved Regions) :    | PAM100                                                             |      |
| Threshold Value (Conserved Regions) : | 25.0 (0 .. 1000)                                                   |      |

شکل -

| CLUSTALW                              | MAFFT                                                              | PRRN |
|---------------------------------------|--------------------------------------------------------------------|------|
| Help                                  |                                                                    |      |
| <b>Specify Parameters :</b>           |                                                                    |      |
| Sequence Type :                       | <input checked="" type="radio"/> PROTEIN <input type="radio"/> DNA |      |
| Score Matrix :                        | PAM250                                                             |      |
| Gap Penalty (PROTEIN) :               | Open: 9.0 , Extend: 2.0 , Background: 1.0 (0 .. 100)               |      |
| Gap Penalty (DNA) :                   | Open: 5.0 , Extend: 2.0 , Background: 2.0 (0 .. 100)               |      |
| Output Format :                       | Native                                                             |      |
| Method :                              | Progressive (pairwise alignment) + Iterative refinement            |      |
| Window Size :                         | 100 (0 .. 1000)                                                    |      |
| Score Matrix (Conserved Regions) :    | PAM100                                                             |      |
| Threshold Value (Conserved Regions) : | 25.0 (0 .. 1000)                                                   |      |

شکل -



CLUSTALW

MAFFT

PRRN

Help

**Specify Parameters :**

Sequence Type : ☒ PROTEIN ☐ DNA

Score Matrix : PAM250

Gap Penalty (PROTEIN) : Open: 9.0 , Extend: 2.0 , Background: 1.0 (0 .. 100)

Gap Penalty (DNA) : Open: 5.0 , Extend: 2.0 , Background: 2.0 (0 .. 100)

Output Format : Native

Method : Progressive (pairwise alignment) + Iterative refinement

Window Size : Iterative refinement

Score Matrix (Conserved Regions) : Sequential (input order) + Iterative refinement

Threshold Value (Conserved Regions) : Sequential (from shorter) + Iterative refinement

Sequential (from longer) + Iterative refinement

Sequential (random) + Iterative refinement

Progressive (input alignment) + Iterative refinement

Progressive (amino acid content) + Iterative refinement

Progressive (pairwise alignment) + Iterative refinement

Enter your sequences (FASTA format) :

شکل -

CLUSTALW

MAFFT

PRRN

Help

**Specify Parameters :**

Sequence Type : ☒ PROTEIN ☐ DNA

Score Matrix : PAM250

Gap Penalty (PROTEIN) : Open: 9.0 , Extend: 2.0 , Background: 1.0 (0 .. 100)

Gap Penalty (DNA) : Open: 5.0 , Extend: 2.0 , Background: 2.0 (0 .. 100)

Output Format : Native

Method : Progressive (pairwise alignment) + Iterative refinement

Window Size : 100 (0 .. 1000)

Score Matrix (Conserved Regions) : PAM100

Threshold Value (Conserved Regions) : PAM0 1000

PAM50

PAM100

PAM150

PAM200

PAM250

PAM300

Enter your sequences (FASTA format) :

Restrictions of the sequence size:

1. [The number of sequences] <= 200

شکل -

Enter your sequences (FASTA format) :

Restrictions of the sequence size:

1. [The number of sequences] <= 200
2. [The maximum length] <= 2000
3. [The number of sequences] x [The maximum length] <= 50000

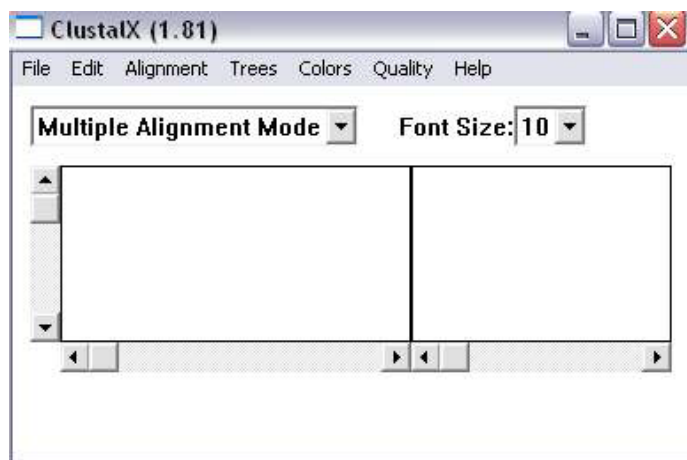
File Upload

Copy & Paste

شکل -

## نرم افزار ClustalX

این برنامه محیط رایانه را جهت تنظیم تعیین توالی سکانس و همچنین تجزیه و تحلیل نتایج آن آماده می‌سازد.



شکل -