



دانشگاه اصفهان
دانشکده مهندسی کامپیوتر

گزارش تمرین اول توالی ها

پدیدآورنده:

محمد امین کیانی

4003613052

دانشجوی کارشناسی، دانشکده مهندسی کامپیوتر، دانشگاه اصفهان، اصفهان،
Aminkianiworkeng@gmail.com

استاد: جناب آقای دکتر نقش نیلچی

نیمسال اول تحصیلی 1403-04

فهرست مطالب

3	مستندات
3	1-مسئله و تحلیل کلی آن:
4	2-تمرین اول:
5	3-شروع کار با برنامه :
6	4-بارگذاری توالی:
9	5-ذخیره‌ی توالی:
10	6-توالی چندگانه:
14	7- مراجع

مستندات

1- مسئله و تحلیل کلی آن:

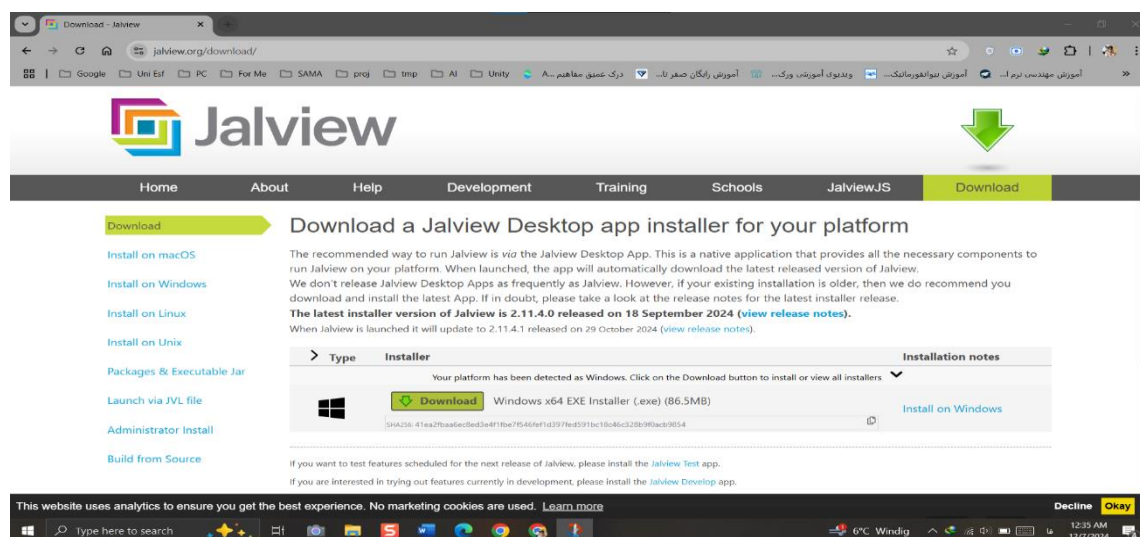
بخش اعظم سوالات مربوط به حیطة ی بیولوژی در دل مبحث تکامل نهفته است که از دیر باز مورد توجه زیست شناسان دوره های مختلف بوده است. یکی از ابزارهای مهم در جهت شناسایی شباهت های بین دو ژن و ارتباط آن با عملکرد مشابه ژن ها و به طور کلی پی بردن به فرضیه تکامل و یافتن جد مشترک، بررسی و مقایسه ی توالی های اسید نوکلئیک و پروتئین می باشد که تحت عنوان هم ردیفی توالی های اسید نوکلئیکی و پروتئینی شناخته می شود. هم ردیف سازی توالی ها به طور کلی به دو دسته تقسیم بندی می شود که یا از لحاظ تعداد توالی و یا از لحاظ طول توالی ها مورد بررسی قرار می گیرد. هم ردیف سازی توالی ها از نظر تعداد توالی ها به دو دسته هم ردیف سازی دوگانه و هم ردیف سازی چندگانه تقسیم بندی می شود. در هم ردیف سازی دوگانه تک تک توالی های مورد نظر با توالی الگو مقایسه می شود. در هم ردیفی چند گانه چندین توالی همزمان با هم مقایسه می شوند. هم ردیف سازی توالی ها از نظر طول به دو گروه محلی و عمومی تقسیم بندی می شود. در هم ردیف سازی محلی، محل های مشخصی از دو توالی با همدیگر هم ردیف می شود ولی در هم ردیفی کامل یا عمومی، طول کامل دو یا چند توالی با هم مقایسه می شود. از جمله مهم ترین ابزارهای مورد استفاده به این منظور می توان به بلاست، مومر، کوگ، آنالیز ژنوم، جلویو و کلوستال اومگا اشاره نمود.

یکی از نرم افزار های تحت وب به منظور هم ردیف سازی چندگانه توالی ها، نرم افزار کلوستال اومگا می باشد که برای هم ردیف سازی چندین توالی، از جعبه ی مربوطه نوع توالی مورد نظر را انتخاب نموده و در جعبه ی مربوط به توالی ها، توالی های مورد نظر را وارد نموده و گزینه ی تایید را انتخاب می نماییم.

جلویو که در ادامه به آن می پردازیم نیز نرم افزار بعدی است که به منظور هم ردیف سازی چندین توالی مورد استفاده قرار می گیرد و یک برنامه

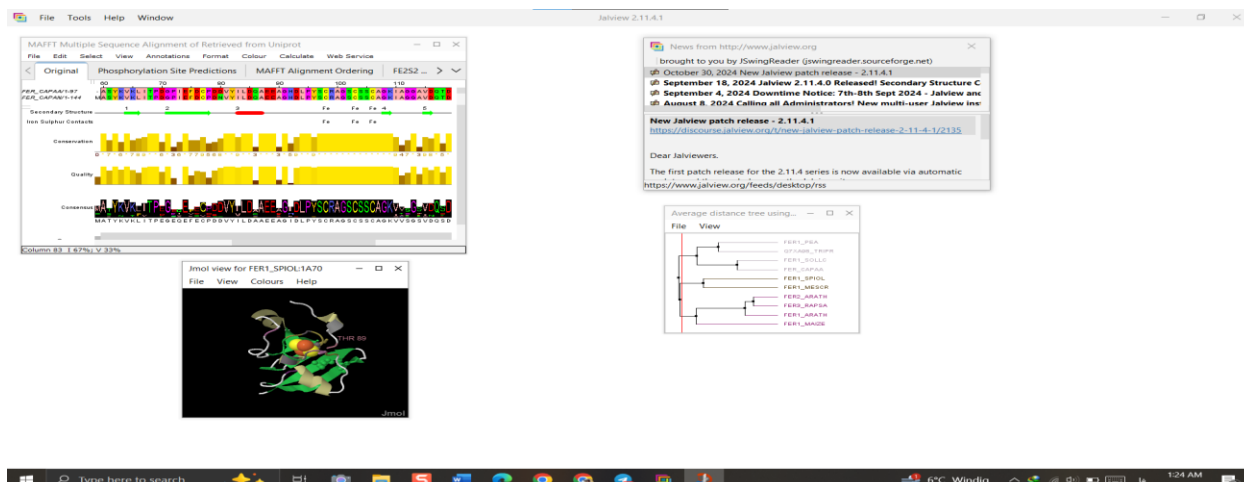
رایگان به منظور هم‌ردیف‌سازی چندگانه توالی‌ها می‌باشد. از آن می‌توان برای هم‌ردیف‌سازی چندگانه توالی‌ها، آنالیز توالی‌ها با درخت‌های فیلوژنتیک و نمودارهای پی‌سی‌ای و آنالیز ساختارهای مولکولی بهره‌برد.

برای استفاده از نرم‌افزار ابتدا باید توالی‌های مورد نظر را به صورت فست‌ا بدست می‌آوریم و از طریق منوی فایل، ورودی هم‌ردیف‌سازی را انتخاب و از جعبه‌ی متنی را انتخاب می‌کنیم و توالی مورد نظر را وارد نموده و اقدام به هم‌ردیف‌سازی توالی‌های مورد نظر می‌کنیم.



2-تمرین اول:

نصب برنامه و اجرای آن



3- شروع کار با برنامه :

جلویو دو حالت ناوبری و ویرایش دارد: حالت عادی (که ویرایش و ناوبری از طریق ماوس انجام می‌شود) و حالت مکان‌نما (که ویرایش و ناوبری از طریق صفحه‌کلید انجام می‌شود). جالووی همیشه در حالت عادی راه‌اندازی می‌شود. کلید اف دو برای جابجایی بین این دو حالت استفاده می‌شود. در سیستم عامل مک، برای استفاده از این کلید، باید از ترکیب کلید اف ان و اف دو استفاده کرد (چرا که کلید اف دو اغلب به روشنایی صفحه اختصاص داده شده است).

می‌توان پنجره‌ها را در دسکتاپ با قرار دادن ماوس روی نوار عنوان در بالای پنجره جابجا کرد، سپس کلیک و کشید.

می‌توان پنجره‌ها را با قرار دادن ماوس روی گوشه پایین سمت راست پنجره تغییر اندازه داد، سپس کلیک و کشید.

برای پیمایش در هم‌تراز از نوار پیمایش عمودی و نوار پیمایش افقی استفاده می‌کنیم.

پنجره نمای کلی را باز کرده و از منوی نمایش، گزینه پنجره نمای کلی را در پنجره هم‌تراز انتخاب و پنجره را جابجا و تغییر اندازه می‌دهیم.

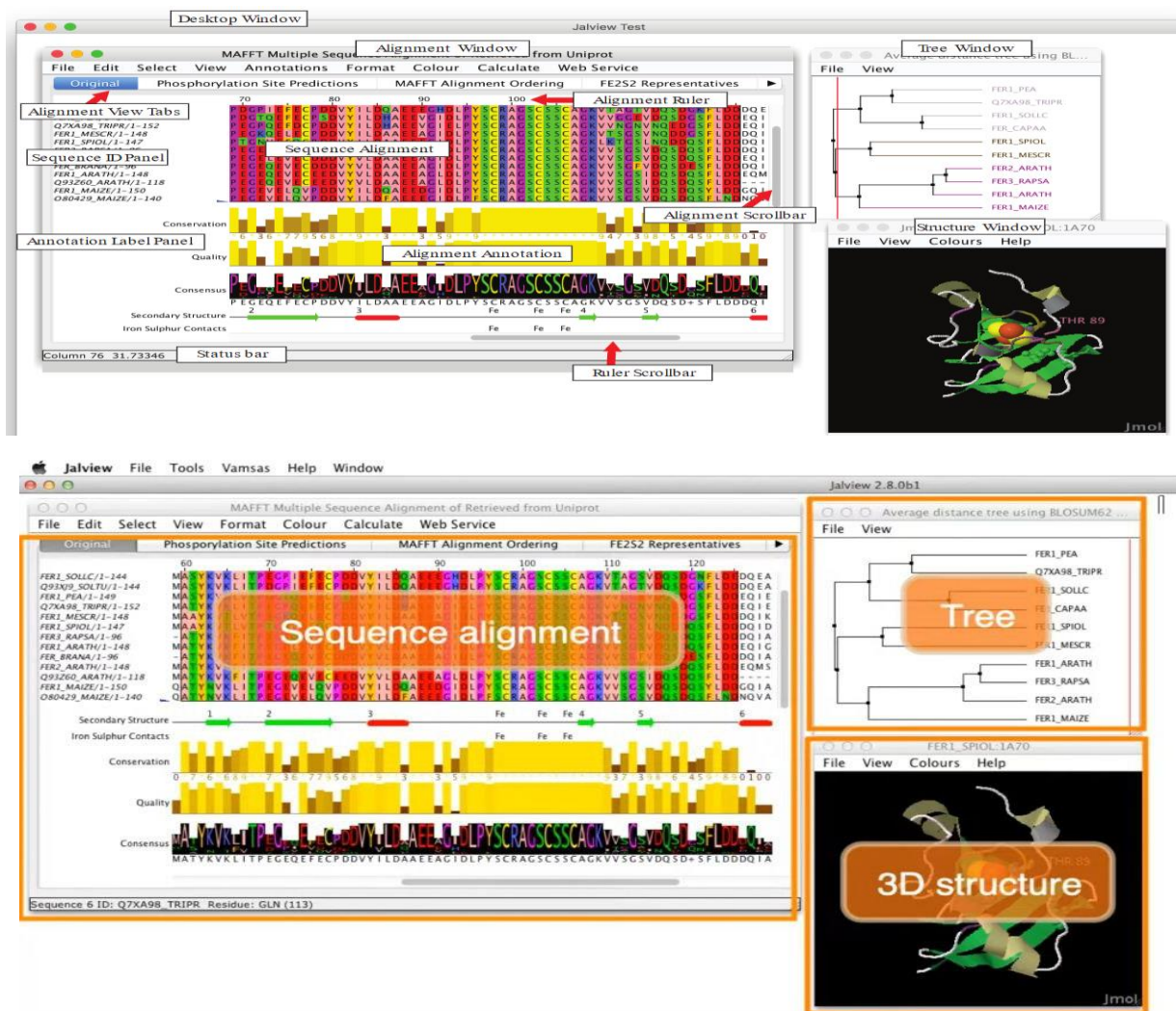
مستطیل قرمز را در پنجره نمای کلی کشیده و مشاهده می‌کنیم که در پنجره هم‌تراز چه تأثیری دارد.

ماوس را روی یک باقی‌مانده در پنجره هم‌تراز قرار داده و نوار وضعیت را در پایین سمت چپ پنجره می‌بینیم. این نوار اطلاعاتی درباره توالی و باقی‌مانده انتخاب شده را نشان می‌دهد.

کلید اف دو را فشار داده تا به حالت مکان‌نما برویم. موقعیت مکان‌نما با پس‌زمینه سیاه و متن سفید نشان داده می‌شود. از کلیدهای جهت‌دار (↑، ↓، ←، →) برای جابجایی مکان‌نما در هم‌تراز استفاده می‌کنیم.

با فشار دادن س هفت به توالی هفتم می‌رود. با فشار دادن ز هجده به ستون هجدهم می‌رود. با فشار دادن ح هجده به باقی‌مانده هجده می‌رود. توجه شود که شماره باقی‌مانده و توالی ممکن است در موقعیت‌های متفاوتی قرار داشته باشند اگر فضاها خالی در توالی درج شده باشد.

کلید اف دو را فشار می‌دهیم، تا به حالت عادی برگردد و همچنین تمامی پنجره‌ها را با استفاده از گزینه «بستن» می‌توان بست.



4- بارگذاری توالی:

دنباله‌ها ممکن است از پایگاه‌های داده عمومی بازیابی شوند یا از یک فایل بارگذاری شوند.

- بارگذاری دنباله‌ها از پایگاه داده عمومی:

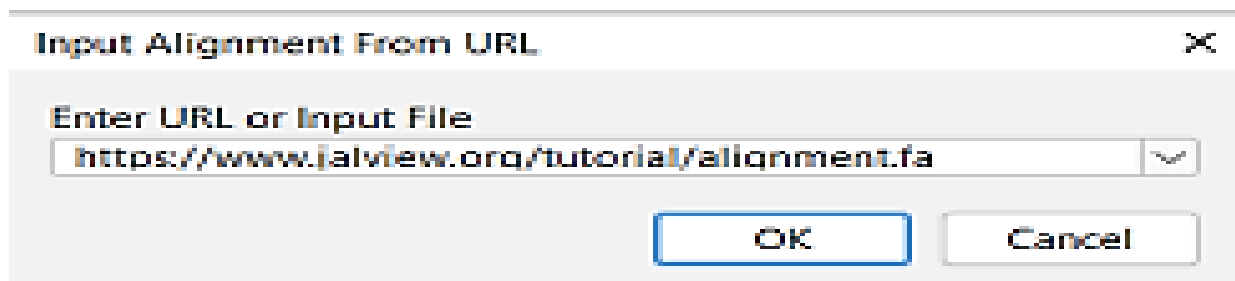
از منوی فایل گزینه واکنشی توالی را انتخاب کرده و در کادر گفتگوی جدیدی که باز می‌شود، روی زبانه پایگاه داده را انتخاب کنید و پایگاه داده (Seed) PFAM را انتخاب می‌کنیم. شماره شناسایی PF03460 را وارد و روی تایید کلیک می‌کنیم. یک هم‌ترازی از حدود 170 دنباله باید بارگذاری شود.



- بارگذاری دنباله‌ها از URL:

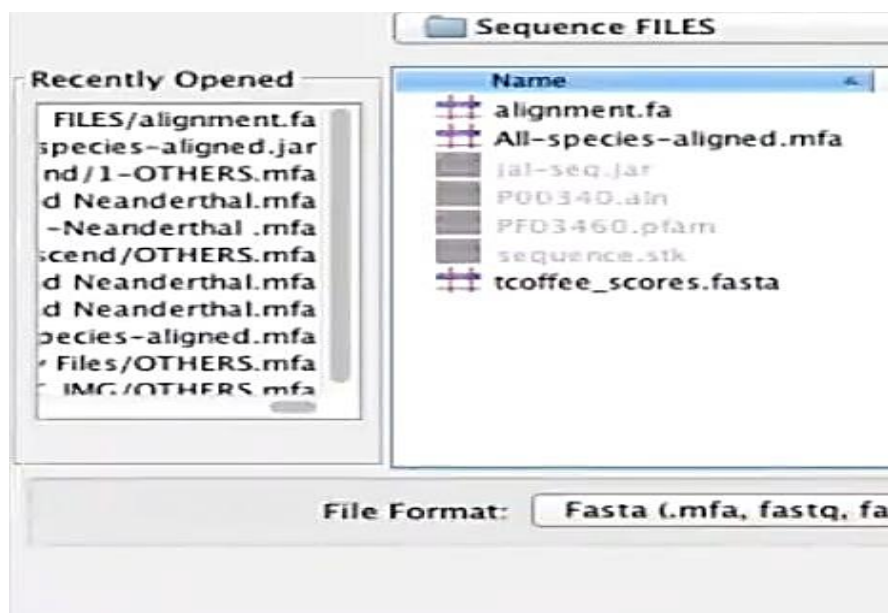
از منوی فایل گزینه URL From Input Alignment را انتخاب کرده و <https://www.jalview.org/tutorial/alignment.fasta> را در کادر گفتگوی باز شده وارد کنید و روی تایید کلیک می‌کنیم.

فایل (.fasta) را روی کامپیوتر خود با استفاده از منوی فایل گزینه Save as در منوی همترازی ذخیره و فایل را در یک پوشه مناسب قرار می‌دهیم.



- بارگذاری دنباله‌ها از یک فایل:

فایل را که اکنون ذخیره کرده‌اید، با انتخاب منوی فایل گزینه Input Alignment From File را انتخاب و باز می‌کنیم.



- بارگذاری دنباله‌ها با روش کشیدن و رها کردن:

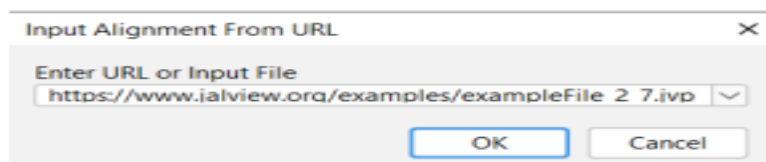
فایل fasta را که از پوشه‌اش ذخیره کرده‌ایم، کشیده و روی پنجره دسکتاپ جالویو رها تا هم‌ترازی در یک پنجره جدید باز می‌شود. این فرآیند را دوباره تکرار اما فایل را روی پنجره هم‌ترازی جالو وی رها کرده تا این دنباله‌ها را در زیر دنباله‌های دیگر در هم‌ترازی اضافه کند و می‌دانیم که در بارگذاری پروژه نمی‌توان به تنهایی یک هم ردیف سازی را آورد!

5-ذخیره‌ی توالی:

- انواع روش های ذخیره و بارگذاری فایل به فرمت عادی یا فستا:

-- از URL از منوی دسکتاپ، لینک زیر را در کادر محاوره‌ای وارد کرده و روی تایید کلیک می‌کنیم.

https://www.jalview.org/examples/exampleFile_2_7.jvp



پروژه را به عنوان ... ذخیره و از منوی دسکتاپ و فایل را روی کامپیوتر ذخیره می‌کنیم. تمام پنجره‌ها را بسته و پروژه را از طریق فایل بارگذاری پروژه در منوی دسکتاپ بارگذاری می‌کنیم.

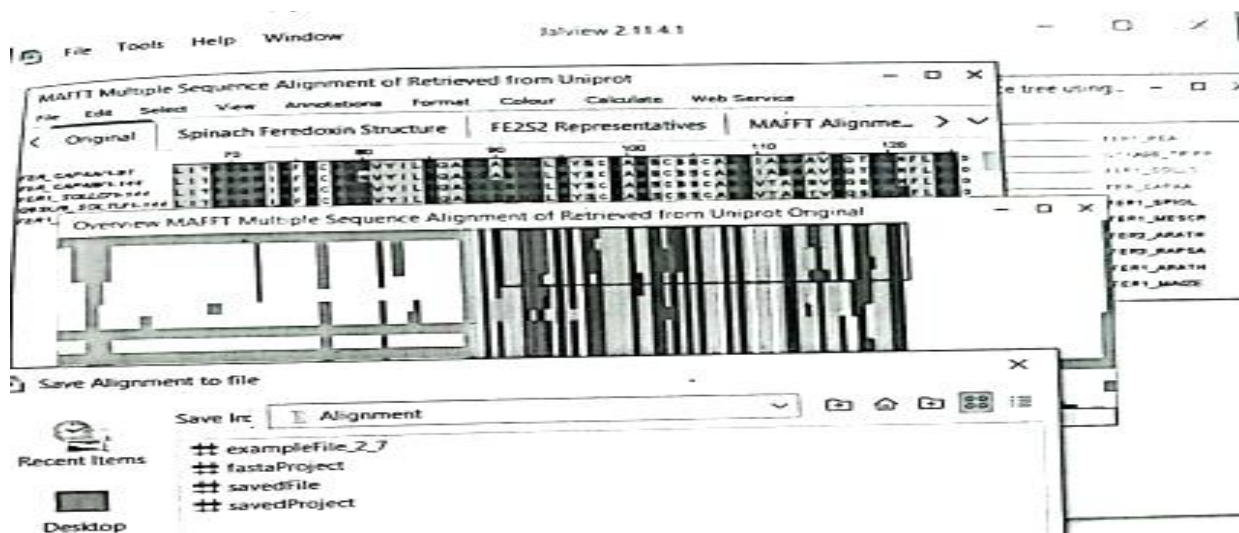


-- انتخاب فایل به عنوان ذخیره و فایل را در فرمت fasta (بالای لیست) در یک پوشه روی کامپیوتر ذخیره می‌کنیم، نام فایل را وارد کرده، فرمت فایل را انتخاب کرده و روی ذخیره کلیک می‌کنیم. حال فایل fasta را از طریق فایل در منوی دسکتاپ باز می‌کنیم.

حال در ادامه باید مقایسه کرد که در هر دو روش چند پنجره باز می‌شود؟

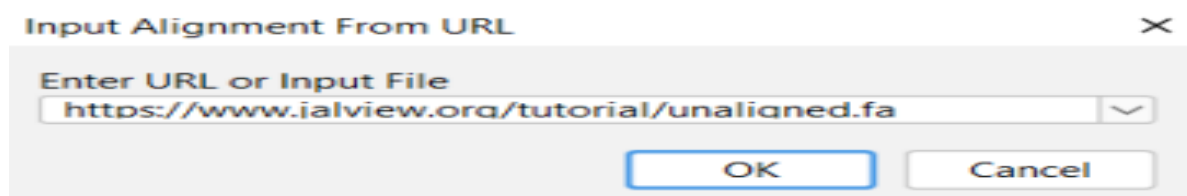
هنگام بارگذاری پروژه با روش اول یک پنجره برای هم‌ردیف سازی باز می‌شود و هنگام بازکردن فایل fasta، 3 صفحه‌ی هم‌ردیف سازی و درخت و دید جمول باز می‌شود.

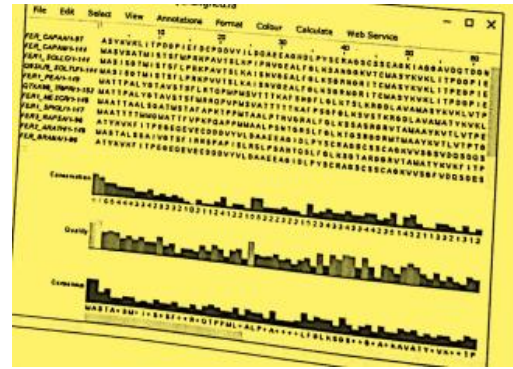
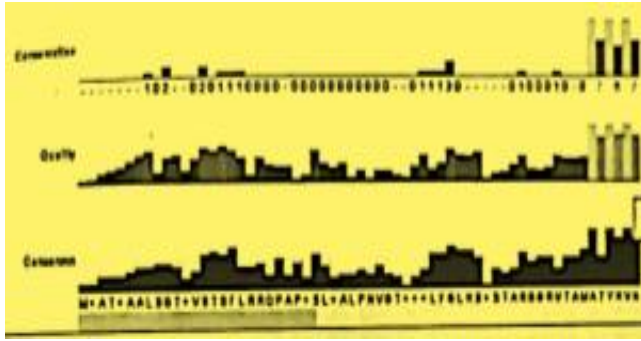
برای ذخیره کردن نیز با منوی اصلی و ذخیره ی پروژه به عنوان... به فرمت عادی jvp می باشد ولی با صفحه ی همردیف سازی در منوی فایل ها پروژه فستار را با فرمت های jvp و jar ذخیره می کند.



6-توالی چندگانه:

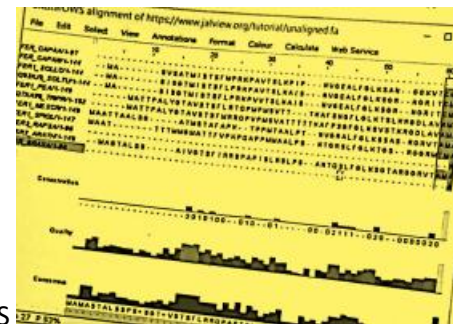
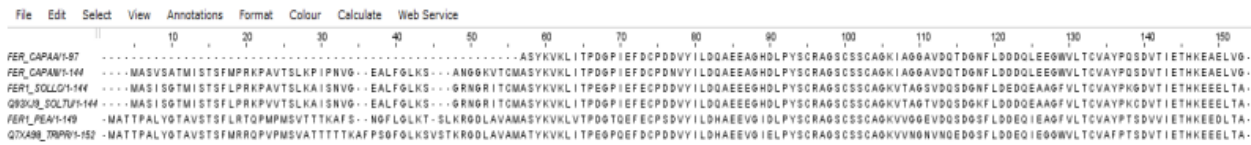
حال مواردی از چند الگوریتم هم ردیف سازی muscle and mafft و clustalO را اجرا بر روی یک فایل بارگزاری شده هم ردیف سازی می کنیم.
<https://www.jalview.org/tutorial/unaligned.fasta> را وارد کرده و روی «تأیید» کلیک می کنیم.





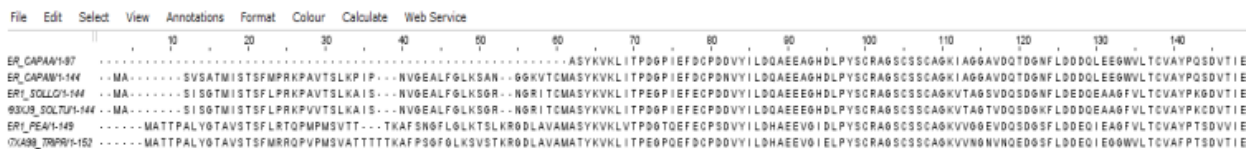
MuscleWS روشی برای هم ردیف کردن توالی ها با تقلیل پیچیدگی های زمانی

MuscleWS alignment of <https://www.jalview.org/tutorial/unaligned.fa>



ClustalOWS روشی برای هم ردیف کردن توالی ها با دقت بیشتر

ClustalOWS alignment of <https://www.jalview.org/tutorial/unaligned.fa>



تفاوت الگوریتم های اجرایی؟

Muscle: که بهینه ساز هم ردیف (تراز) جفتی حین فرایند هم ترازى پیش روی از الگوریتم تراز پیشرونده استفاده اما از روند پیچیده تری به نام "روش پالیش تکراری" استفاده می کند. این یک هم ترازى تدریجی اولیه را انجام و سپس به طور تکراری هم ترازى را با استفاده از تکنیک تراز پروفایل-پروفایل دقیق تر مورد تایید می نماید. پس نواحی کوچک با شباهت بالا را با دقتی بالا بدون توجه به هدر رفت هم ردیف سازی عمومی هم تراز می کند.

Clustal: دارای درخت راهنما با ماتریس فاصله ای قوی که با کمک مدل های نهان مارکوف از الگوریتم تراز پیشرونده، مانند روش همترازی زوجی پیشرونده، برای ساختن همترازی چند توالی نهایی استفاده که یک درخت راهنما بر اساس شباهت توالی دوتایی می سازد و سپس به تدریج توالی ها را بر اساس این درخت تراز می کند ولی اینجا تمرکز بر حفظ تراز عمومی و یکنواخت سازی نواحی کمتر محافظت شده برای هم ترازى است. پس در ابتدا ممکن است تغییرات جزئی مانند جابجایی و درج یا حذف را بر اساس سازگاری کلی موثر باشند.

ClustalOWS alignment of htt				MuscleWS alignment of htt			
File	Edit	Select	View	File	Edit	Select	View
FER_CAPAA/1-97	----	MA	-----	ER_CAPAA/1-97	----	MASVS	-----
FER_CAPAN/1-144	----	MA	-----	ER_CAPAN/1-144	----	MASIS	-----
FER1_SOLLG/1-144	----	MA	-----	ER1_SOLLG/1-144	----	MASIS	-----
Q93XJ9_SOLTU/1-144	----	MA	-----	Q93XJ9_SOLTU/1-144	----	MATTPAL	-----
FER1_PEA/1-149	----	MAT	-----	ER1_PEA/1-149	----	MATTPAL	-----
Q7XA98_TRIPR/1-152	----	MAT	-----	Q7XA98_TRIPR/1-152	----	MATTPAL	-----
FER1_MESCR/1-148	MAATTAALS	ER1_MESCR/1-148	MAATTAALS				
FER1_SPIOL/1-147	MAAT	ER1_SPIOL/1-147	MAATTTT				
FER3_RAPSA/1-96	-----	ER3_RAPSA/1-96	-----				
FER2_ARATH/1-148	MASTALS	ER2_ARATH/1-148	MASTALS				
FER_BRANA/1-96	-----	ER_BRANA/1-96	-----				
FER1_ARATH/1-148	MASTALS	ER1_ARATH/1-148	MASTALS				
Q93Z60_ARATH/1-118	MASTALS	Q93Z60_ARATH/1-118	MASTALS				
FER1_MAIZE/1-150	MA	TYLOS	ER1_MAIZE/1-150	MA	TYLOS	MA	-----
Q80429_MAIZE/1-140	MAAT	Q80429_MAIZE/1-140	MA				

الگوریتم MAFFT:

فقط دنباله های قابل مشاهده به سرویس ارسال می شوند و در آن برخی از آمینو اسید ها جابجا شده تا شکاف ها را کاهش دهند و با پنهان سازی ستون های صورت سوال و اجرای آن در هم ترازى، تغییرات جزئی می دهد و با اصلاح تدریجی هم ردیف ها، نواحی محافظت شده را بهتر می کند. پس نسبت به دو روش قبلی هم ترازى روان تر و تعدا شکاف کمتر دارد یعنی با جایگاه باقی مانده ها درست می شود. نتایج هر کار فرعی با مناطق پنهان در داده های ورودی قبل از نمایش آنها در یک پنجره جدید الحاق می شود. در نهایت این روش جدید برای هم ترازى سریع توالی چندگانه بر اساس تبدیل فوریه است.



MafftV25 alignment of ClustalOWS alignment of https://www.jalview.org/tutorial/unaligned .										
File	Edit	Select	View	Annotations	Format	Colour	Calculate	Web Service		
FER_CAPAN1-197										
FER_CAPAN1-144	MAS	-	-	-	VSATMISTSFMPRKPAVTSL	-	-	-	KPIP-NVG-EALFGLKS	-A
FER1_SOLLO1-144	MAS	-	-	-	ISGTMISTSFLLPRKPAVTSL	-	-	-	KAIS-NVG-EALFGLKS	-G
QR3XJB_SOLTL1-144	MAS	-	-	-	ISGTMISTSFLLPRKPAVTSL	-	-	-	KAIS-NVG-EALFGLKS	-G
FER1_PAV1-149	MAT	-	-	-	PAISGTVSFLRTPMPMS	-	-	-	TTTAF-NVG-EALFGLKS	-G
QTXA98_TRIMP1-152	MATT	-	-	-	PALYGTAVSTSFMRGRQFVMSVAT	-	-	-	TTTTKAF-NVG-EALFGLKS	-SVST
FER1_MESCR1-148	MAAT	-	-	-	TAALSGATMSTAFAPKTPPMTA	-	-	-	ALPTNVG-RALFGLKSSA	-A
FER1_SPOL1-147	MAAT	-	-	-	TTTMMG-MATTFVFKPQAFPM	-	-	-	AALPSNTG-RSLFGLKTO	-S
FER3_RAS1-190	MAT	-	-	-	PAISGTVSFLRTPMPMS	-	-	-	TTTAF-NVG-EALFGLKS	-G
FER1_RAS1-190	MAST	-	-	-	ALSSAIVGSTFIRRAPISL	-	-	-	RSLP-SANTQSLFGLKSSA	-TA
FER_BRAN1-98	MAT	-	-	-	ALSSAIVGSTFIRRAPISL	-	-	-	RSLP-SANTQSLFGLKSSA	-TA
FER1_ARATH1-148	MAST	-	-	-	ALSSAIVGSTFIRRAPISL	-	-	-	RSLP-SANTQSLFGLKSSA	-TA
QR3260_ARATH1-118	MAT	-	-	-	ALSSAIVGSTFIRRAPISL	-	-	-	RSLP-SANTQSLFGLKSSA	-TA
FER1_PAV1-149	MAT	-	-	-	PAISGTVSFLRTPMPMS	-	-	-	TTTAF-NVG-EALFGLKS	-G
QR329_MALVE1-140	MAAT	-	-	-	ALSSAIVGSTFIRRAPISL	-	-	-	RSLP-SANTQSLFGLKSSA	-TA

7- مراجع

- [1] <https://www.jalview.org/tutorial/exercises/>
- [2] <https://www.geniranlab.ir>