MEGA软件的使用初探

胡亚平

82101061040

一、MEGA的安装和
主操作界面
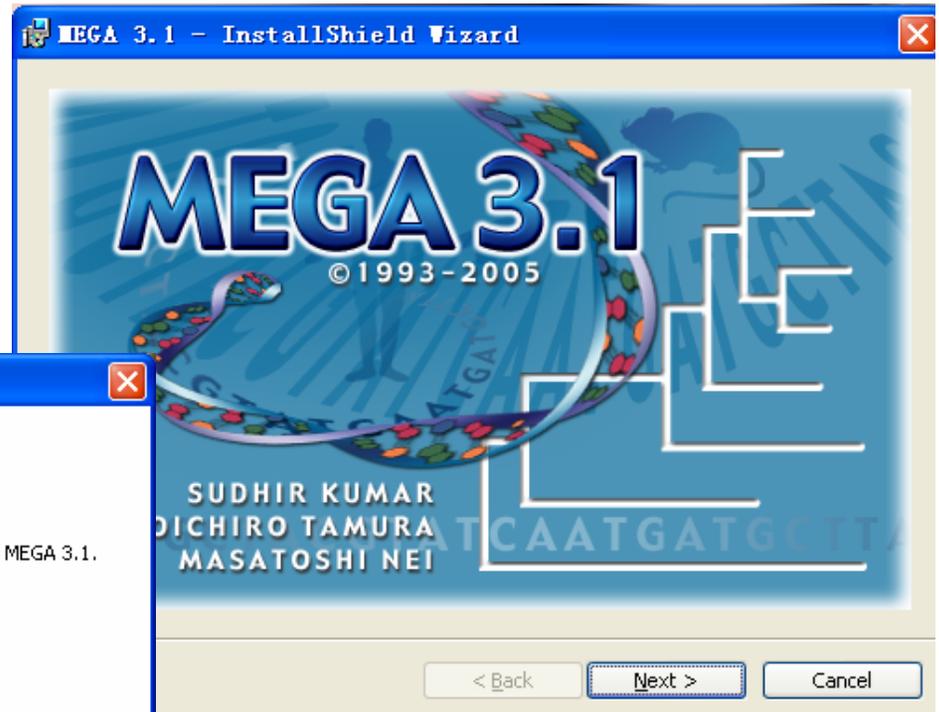
- MEGA软件的安装过程与其他在windows下使用的软件一样，比较方便。

Molecular Evolutionary Genetics Analysis, Version ...

File   Phylogeny   Alignment   Windows   Help

Tutorial on How to Use MEGA
Click me to activate a data file

Citing MEGA in publications
Go to the MEGA web page

ACGTATCAATGAT
ATGCTACGTATC ATGATGCTT

20:25:52

Introduction to Walk Through MEGA

This section provides a MEGA tutorial. The data files for these examples can be found in the **EXAMPLES** folder, located in the MEGA installation directory (example in **C:\Program Files\MEGA 3.1\Examples**). In these example files, data are deliberately written in different input formats. We recommend that you study the examples in the order presented because the techniques explained in the initial examples are used again in the subsequent ones.

In the following write-up, **highlighted** words indicate the keys you must press on the keyboard. If you must press two keys simultaneously, they are shown with a + sign between them (*e.g.*, **Alt + F3** means that the **Alt** and **F3** keys should be pressed simultaneously). Italicized letters are used to mark the commands found in menus, submenus, and other locations as they appear on the computer screen. In every example, we discuss many procedures introducing analytical techniques. For ease of reference in later examples, these procedures are numbered in the Ex *u.v.w* format, where *u* is the example number, *v* is the procedure number, and *w* is the step number. For instance, Ex 1.3.2 refers to the 2nd step of the 3rd procedure in example 1.

A list of tutorials is as follows:
1. Aligning Sequences
2. Estimating Evolutionary Distances
3. Building Trees
4. Testing Tree Reliability
5. Marking Genes/Domains
6. Testing for Selection
7. Grouping Sequences
8. Computing Sequence Statistics
9. Trees from Distance Data

- MEGA软件同样具有纤细的使用说明，我们可以按照使用说明一步步进行操作。本次作业主要使用它的序列比对和构建系统发育树的功能。

二、利用MEGA进行序列比对

- 在ABC的网站中可以直接用MEGA打开后缀为*.fasta格式的序列。如果我们使用自己的序列，可以建一个TXT的文档，输入fasta格式的一系列序列，保存时命名为*.fasta文件就可以进行被MEGA所正确识别。我们先选取SBPD-50Seq进行练习。
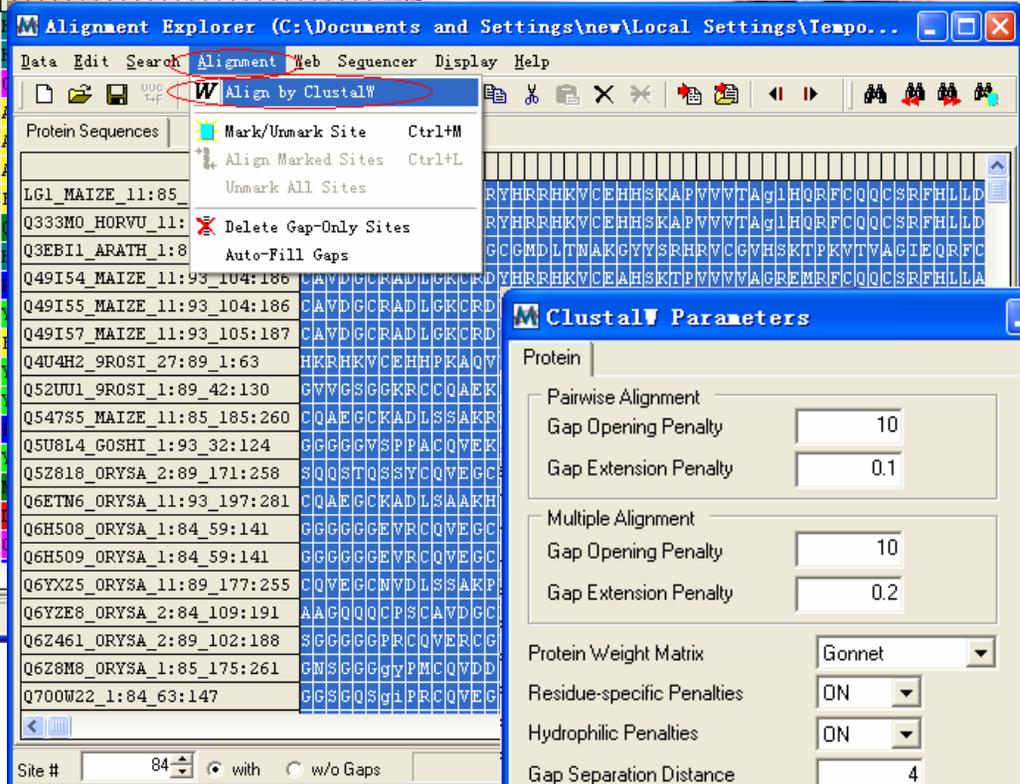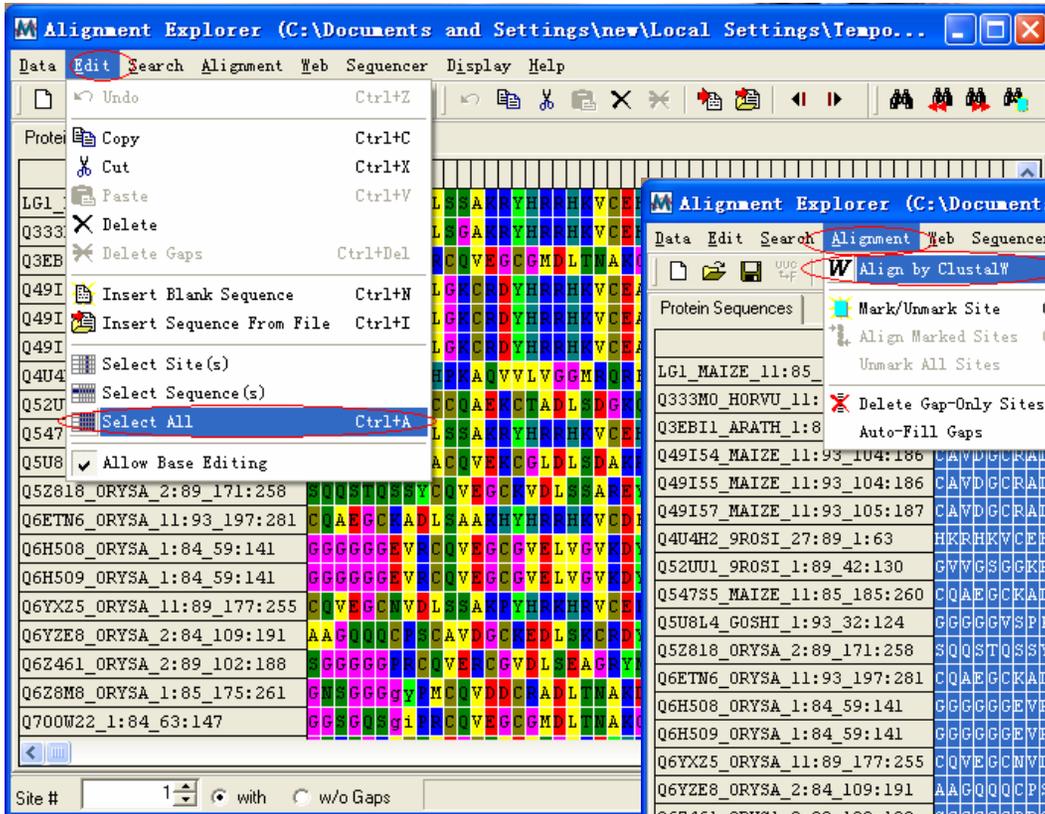
- 序列被正确识别后，我们可以利用MEGA进行序列比对，先edit/select all选中要比对的序列，然后在alignment/align by clustalw进行比对，在弹出的界面可以更改参数，点击OK得到结果。
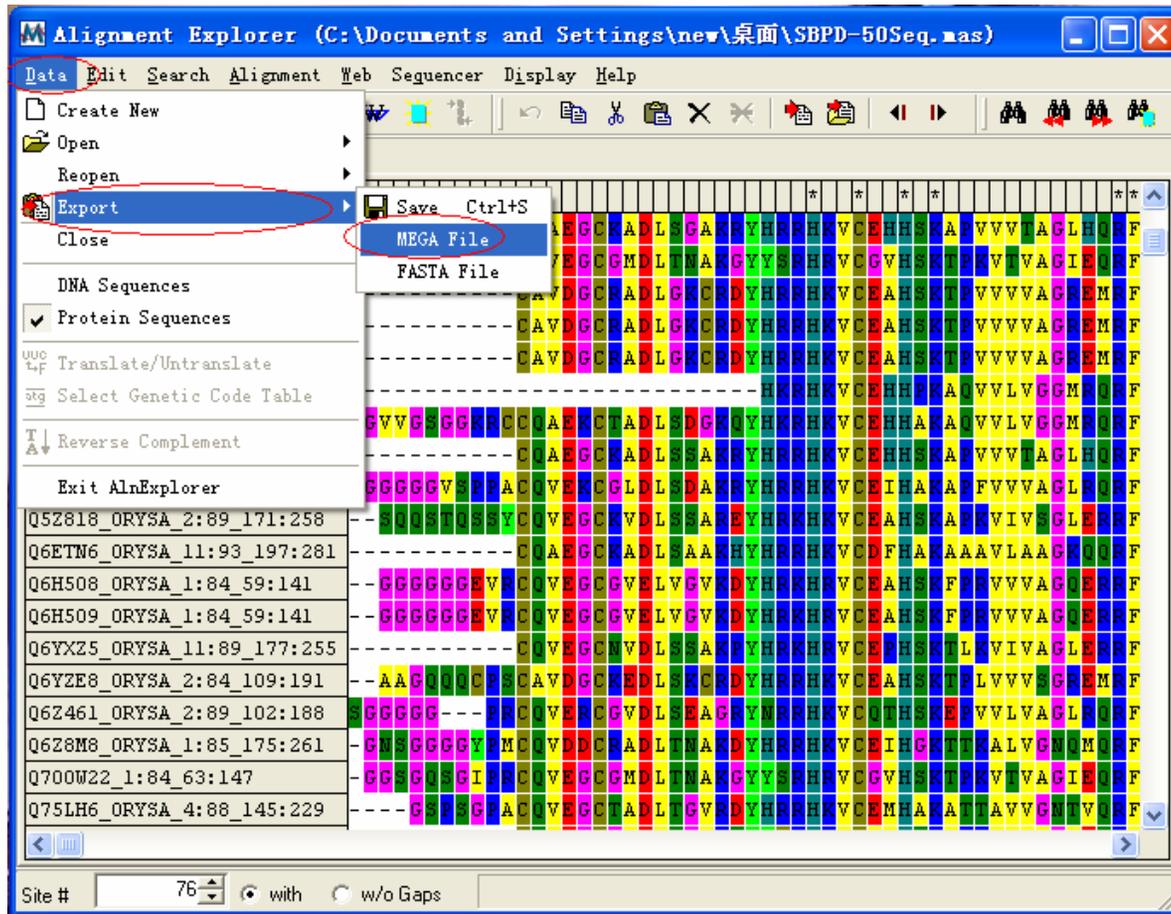
- 得到的序列比对结果可以保存为MEG格式的文件，以便进行后续的建系统发育树等操作。以*.fasta格式保存的文件进行序列比对的过程也基本相同。

三、利用MEGA在NCBI等网站下载并保存数据

MEGA可以作为浏览器查看NCBI等网站的页面，可以在其中查找感兴趣的基因序列并保存在MEGA中

- 在得到需要的序列后可以点击add to alignment保存序列。
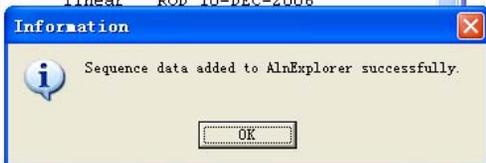
- 保存的序列可以进行比对后保存为*.MEG，以便进行后续操作，也可以保存为*.fasta格式的文件。
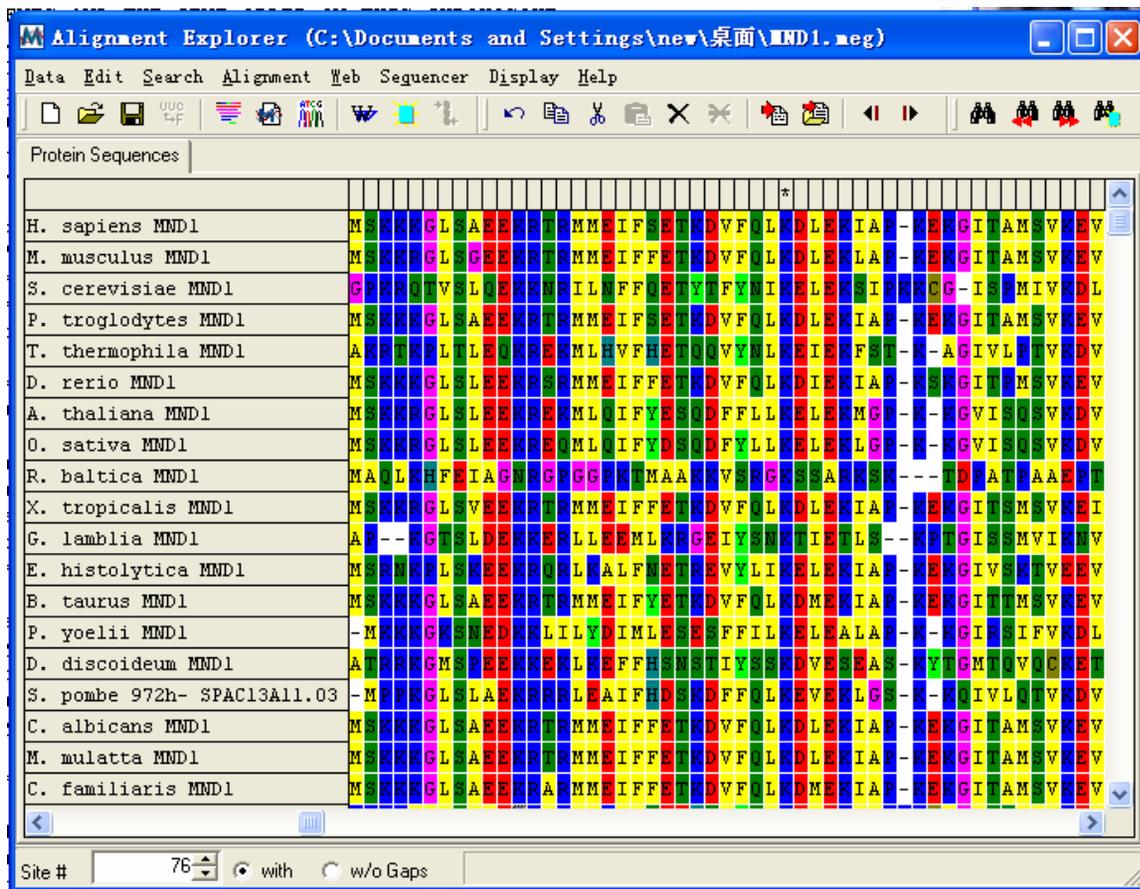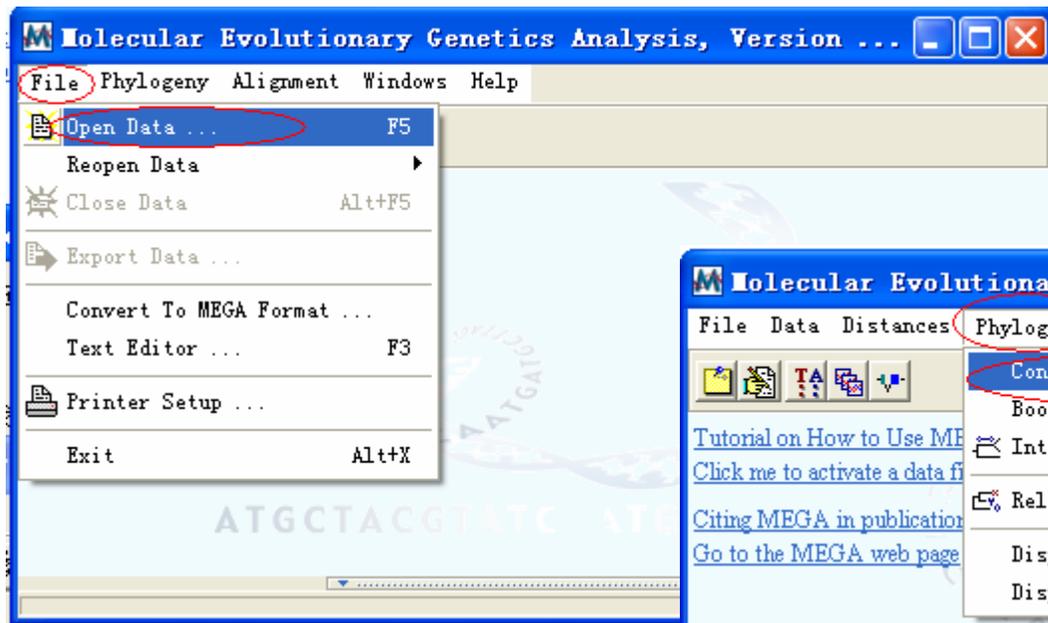
- *.MEG格式的文件可以用来构建系统发育树，我可以构建四种类型的树，分别为 neighbor joining、maximum parsimony、minimum evolution和UPGMA四种算法。
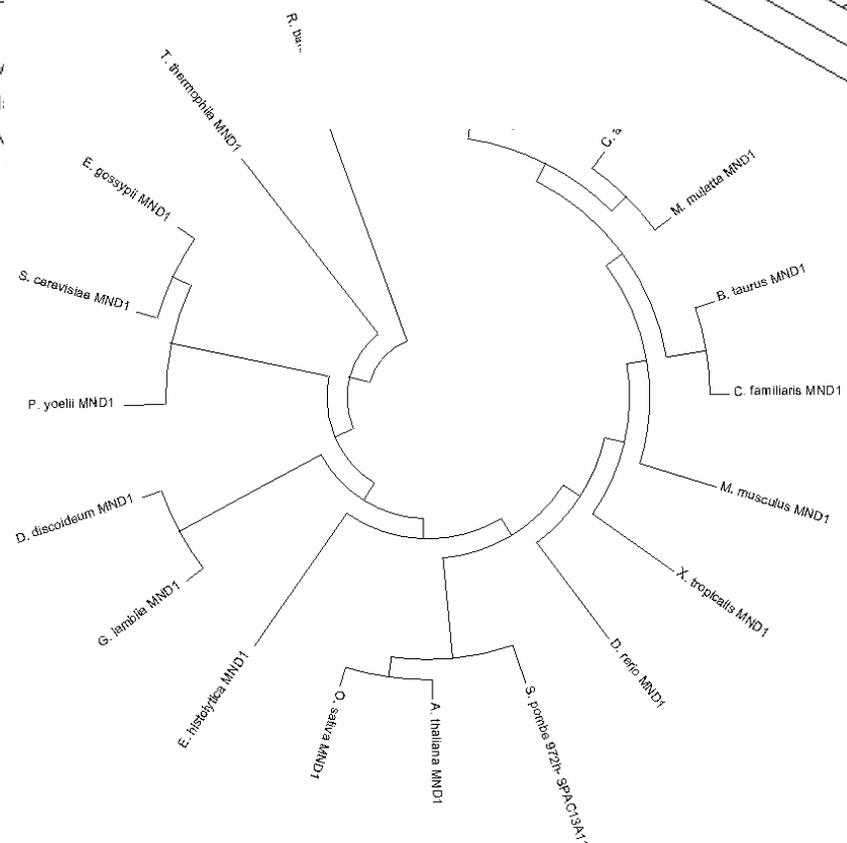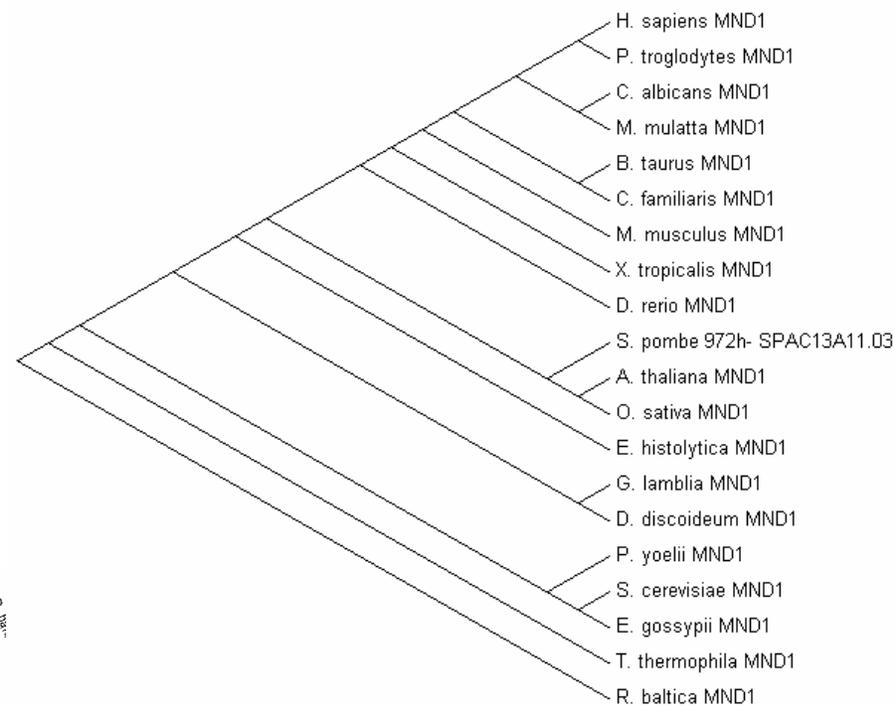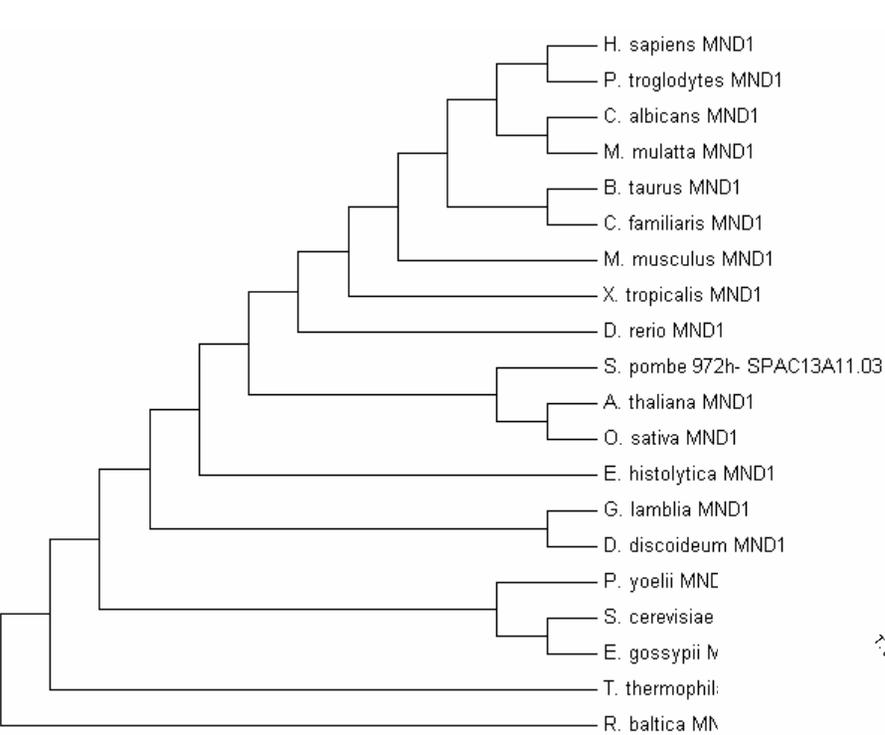
我用neighbor joining
法构建的减数分裂基
因mnd1在不同物种中
同源基因的系统发育
树。可以用不同的模
式显示。

四、用MEGA估算进化距离

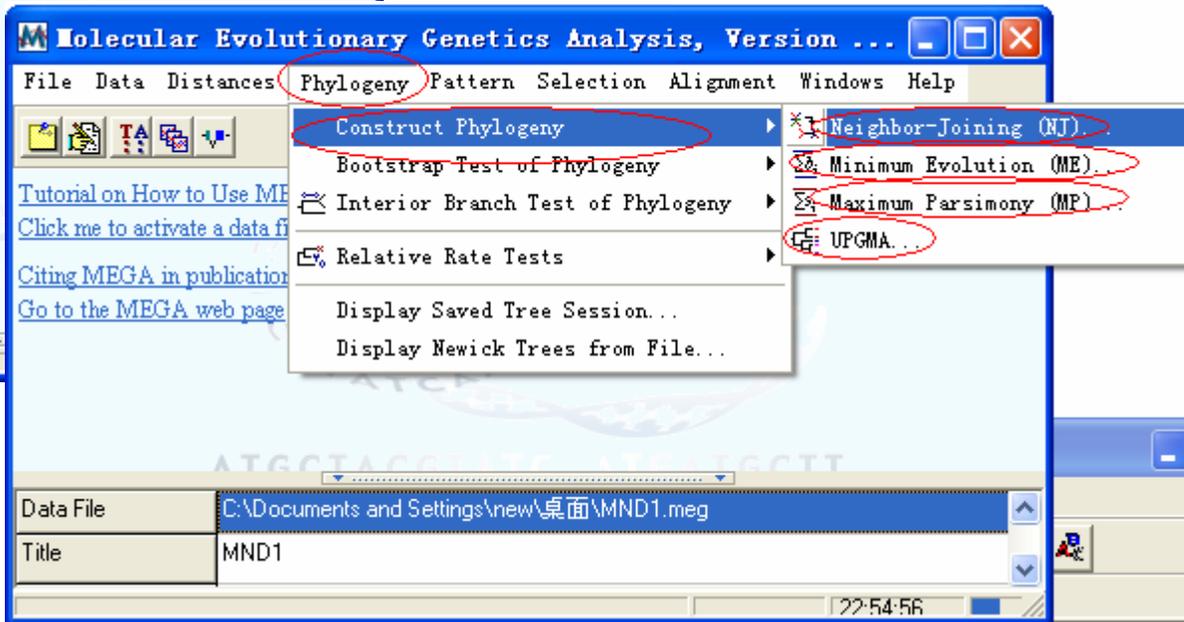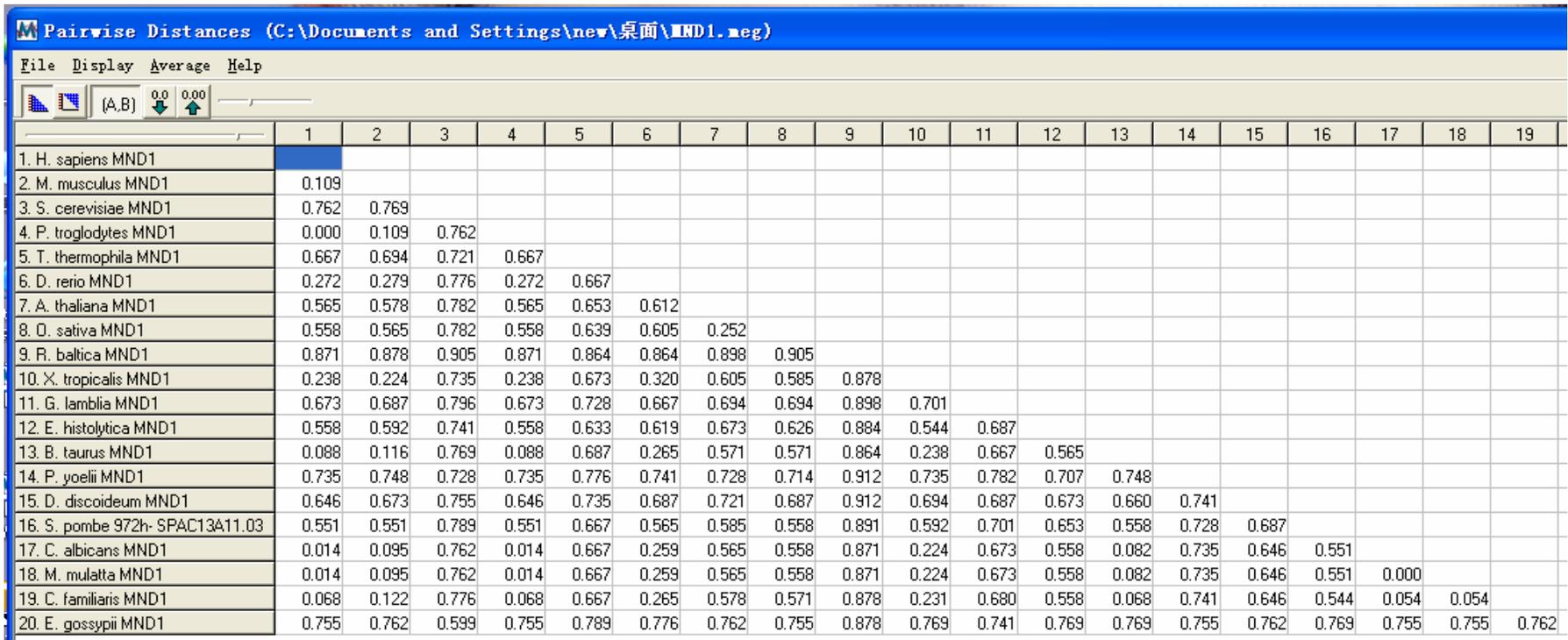|	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. H. sapiens MND1																			
2. M. musculus MND1	0.115																		
3. S. cerevisiae MND1	1.435	1.464																	
4. P. troglodytes MND1	0.000	0.115	1.435																
5. T. thermophila MND1	1.099	1.184	1.277	1.099															
6. D. rerio MND1	0.318	0.327	1.494	0.318	1.099														
7. A. thaliana MND1	0.832	0.863	1.525	0.832	1.059	0.947													
8. O. sativa MND1	0.816	0.832	1.525	0.816	1.020	0.930	0.290												
9. R. baltica MND1	2.046	2.100	2.351	2.046	1.995	1.995	2.282	2.351											
10. X. tropicalis MND1	0.272	0.254	1.327	0.272	1.119	0.385	0.930	0.880	2.100										
11. G. lamblia MND1	1.119	1.162	1.589	1.119	1.302	1.099	1.184	1.184	2.282	1.206									
12. E. histolytica MND1	0.816	0.896	1.353	0.816	1.001	0.965	1.119	0.983	2.157	0.786	1.162								
13. B. taurus MND1	0.093	0.123	1.464	0.093	1.162	0.308	0.847	0.847	1.995	0.272	1.099	0.832							
14. P. yoelii MND1	1.327	1.380	1.302	1.327	1.494	1.353	1.302	1.253	2.425	1.327	1.525	1.229	1.380						
15. D. discoideum MND1	1.039	1.119	1.407	1.039	1.327	1.162	1.277	1.162	2.425	1.184	1.162	1.119	1.078	1.353					
16. S. pombe 972h- SPAC13A11.03	0.801	0.801	1.556	0.801	1.099	0.832	0.880	0.816	2.218	0.896	1.206	1.059	0.816	1.302	1.162				
17. C. albicans MND1	0.014	0.100	1.435	0.014	1.099	0.299	0.832	0.816	2.046	0.254	1.119	0.816	0.085	1.327	1.039	0.801			
18. M. mulatta MND1	0.014	0.100	1.435	0.014	1.099	0.299	0.832	0.816	2.046	0.254	1.119	0.816	0.085	1.327	1.039	0.801	0.000		
19. C. familiaris MND1	0.070	0.131	1.494	0.070	1.099	0.308	0.863	0.847	2.100	0.263	1.140	0.816	0.070	1.353	1.039	0.786	0.056	0.056	
20. E. gossypii MND1	1.407	1.435	0.913	1.407	1.556	1.494	1.435	1.407	2.100	1.464	1.353	1.464	1.464	1.407	1.435	1.464	1.407	1.407	1.435

- 打开前面保存的*.meg文件，点击Distance|Compute Pairwise command (F7) 显示操作参数；设定好Models等参数后点击Compute启动计算，得到如上图的结果。

MEGA也可以用来显示不同序列

氨基酸组成比例的差异

Pairwise Distances (C:\Documents and Settings\new\桌面\MND1.meg)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. H. sapiens MND1																			
2. M. musculus MND1	0.109																		
3. S. cerevisiae MND1	0.762	0.769																	
4. P. troglodytes MND1	0.000	0.109	0.762																
5. T. thermophila MND1	0.667	0.694	0.721	0.667															
6. D. rerio MND1	0.272	0.279	0.776	0.272	0.667														
7. A. thaliana MND1	0.565	0.578	0.782	0.565	0.653	0.612													
8. O. sativa MND1	0.558	0.565	0.782	0.558	0.639	0.605	0.252												
9. R. baltica MND1	0.871	0.878	0.905	0.871	0.864	0.864	0.898	0.905											
10. X. tropicalis MND1	0.238	0.224	0.735	0.238	0.673	0.320	0.605	0.585	0.878										
11. G. lamblia MND1	0.673	0.687	0.796	0.673	0.728	0.667	0.694	0.694	0.898	0.701									
12. E. histolytica MND1	0.558	0.592	0.741	0.558	0.633	0.619	0.673	0.626	0.884	0.544	0.687								
13. B. taurus MND1	0.088	0.116	0.769	0.088	0.687	0.265	0.571	0.571	0.864	0.238	0.667	0.565							
14. P. yoelii MND1	0.735	0.748	0.728	0.735	0.776	0.741	0.728	0.714	0.912	0.735	0.782	0.707	0.748						
15. D. discoideum MND1	0.646	0.673	0.755	0.646	0.735	0.687	0.721	0.687	0.912	0.694	0.687	0.673	0.660	0.741					
16. S. pombe 972h- SPAC13A11.03	0.551	0.551	0.789	0.551	0.667	0.565	0.585	0.558	0.891	0.592	0.701	0.653	0.558	0.728	0.687				
17. C. albicans MND1	0.014	0.095	0.762	0.014	0.667	0.259	0.565	0.558	0.871	0.224	0.673	0.558	0.082	0.735	0.646	0.551			
18. M. mulatta MND1	0.014	0.095	0.762	0.014	0.667	0.259	0.565	0.558	0.871	0.224	0.673	0.558	0.082	0.735	0.646	0.551	0.000		
19. C. familiaris MND1	0.068	0.122	0.776	0.068	0.667	0.265	0.578	0.571	0.878	0.231	0.680	0.558	0.068	0.741	0.646	0.544	0.054	0.054	
20. E. gossypii MND1	0.755	0.762	0.599	0.755	0.789	0.776	0.762	0.755	0.878	0.769	0.741	0.769	0.769	0.755	0.762	0.769	0.755	0.755	0.762

- MEGA还可以计算氨基酸比例的差异，具体操作如下：打开前面保存的*.meg文件，点击Distance / Compute Pairwise command (F7) 显示操作参数；设定好Models为Amino Acid / p-distance等参数后点击Compute启动计算，得到如上图的结果。