分子進化の統計モデリングとモデル選択
演習

- Kakusan4 による分子進化モデルの選択
- モデル選択結果を見る
Kakusan4 を起動
Kakusan4 のウィンドウにドラッグ＆ドロップ

<table>
<thead>
<tr>
<th>名前</th>
<th>更新日時</th>
<th>種類</th>
<th>サイズ</th>
</tr>
</thead>
<tbody>
<tr>
<td>16S.fas</td>
<td>2019/10/22 0:33</td>
<td>FAS File</td>
<td>17 KB</td>
</tr>
<tr>
<td>16S.nuc.fas</td>
<td>2019/10/22 0:27</td>
<td>FAS File</td>
<td>18 KB</td>
</tr>
<tr>
<td>16S_unaligned_nuc.fas</td>
<td>2019/10/22 0:26</td>
<td>FAS File</td>
<td>17 KB</td>
</tr>
<tr>
<td>COX2.nuc.fas</td>
<td>2019/10/21 22:32</td>
<td>FAS File</td>
<td>10 KB</td>
</tr>
<tr>
<td>COX2_aligned_aa</td>
<td>2019/10/21 22:32</td>
<td>FAS File</td>
<td>10 KB</td>
</tr>
<tr>
<td>COX2_aligned_nuc.fas</td>
<td>2019/10/21 22:32</td>
<td>FAS File</td>
<td>10 KB</td>
</tr>
<tr>
<td>COX2.unaligned_aa.fasta</td>
<td>2019/10/21 22:32</td>
<td>FASTA File</td>
<td>4 KB</td>
</tr>
<tr>
<td>COX2.unaligned_nuc.fasta</td>
<td>2019/10/21 22:32</td>
<td>FASTA File</td>
<td>10 KB</td>
</tr>
<tr>
<td>sequence.gb</td>
<td>2019/10/21 22:32</td>
<td>GBファイル</td>
<td>478 KB</td>
</tr>
</tbody>
</table>
If you publish your study using Kakusan4, please cite the following. Tanabe AS (2011) "Kakusan4 and Aminosan: two programs for comparing nonpartitioned, proportional, and separate models for combined molecular phylogenetic analyses of multilocus sequence data", Molecular Ecology Resources, vol.11, pp.914-921.

Copyright (C) 2006-2012 Akifumi S. Tanabe

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301 USA.

Parsing command line options...
No input files are specified.
Entering interactive mode.
Specified options are ignored.
Specify an input file name.
Note that you can use wild card.
C:\Users\shimotsuki\Desktop\Drosophila16S.fas.
Kakusan4 のウィンドウにドラッグ＆ドロップ
This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301 USA.

Parsing command line options...
No input files are specified.
Entering interactive mode.
Specified options are ignored.
Specify an input file name.
Note that you can use wild card.
C:\Users\shimotsuki\Desktop\Drosophila\16S.fas
"C:/Users/shimotsuki/Desktop/Drosophila/16S.fas" was accepted.
Specify an input file name or just the specification.
C:\Users\shimotsuki\Desktop\Drosophila\COX2 P.fas
Parsing command line options...
No input files are specified.
Entering interactive mode.
Specified options are ignored.
Specify an input file name.
Note that you can use wild card.
C:\Users\shimotsuki\Desktop\Drosophila\16S.fas
"C:/Users/shimotsuki/Desktop/Drosophila/16S.fas" was accepted.
Specify an input file name or just press enter to leave input file specification.
C:\Users\shimotsuki\Desktop\Drosophila\COX2_P.fas
"C:/Users/shimotsuki/Desktop/Drosophila/COX2_P.fas" was accepted.
Specify an input file name or just press enter to leave input file specification.

OK. Input file specifications are completed.

Log, result and configuration files will be output to "C:\Users\shimotsuki\Desktop\Drosophila\16S.fas.kakusan".

OUTPUT OPTIONS

Which is a target analysis software? (MrBayes/Treefinder/PAUP/PHYML/RAxML)
(default: RAxML)
with this program; if not, write to the Free Software Foundation, Inc.,
51 Franklin Street, Fifth Floor, Boston, MA 02110-1301 USA.

Parsing command line options...
No input files are specified.
Entering interactive mode.
Specified options are ignored.
Specify an input file name.
Note that you can use wild card.
C:/Users/shimotsuki/Desktop/Drosophila/16S.fas
"C:/Users/shimotsuki/Desktop/Drosophila/16S.fas" was accepted.
Specify an input file name or just press enter to leave input file specification.
C:/Users/shimotsuki/Desktop/Drosophila/COX2_P.fas
"C:/Users/shimotsuki/Desktop/Drosophila/COX2_P.fas" was accepted.
Specify an input file name or just press enter to leave input file specification.

OK. Input file specification has terminated.

Log, result and configuration files will be output to "C:/Users/shimotsuki/Desktop/Drosophila/16S.fas.kakusan".

OUTPUT OPTIONS

Which is a target analysis software? (MrBayes/Treefinder/PAUP/PHYML/RAxML) (default: RAxML)
"C:/Users/shimotsuki/Desktop/Drosophila/COX2_P.fas" was accepted. Specify an input file name or just press enter to leave input file specification.

OK. Input file specification has terminated.

Log, result and configuration files will be output to "C:/Users/shimotsuki/Desktop/Drosophila/COX2_P.fas.kakusan".

OUTPUT OPTIONS

Which is a target analysis software? (MrBayes/Treefinder/PAUP/PHYML/RAxML) (default: RAxML)

OK. The model configuration files for RAxML will be output.

ANALYSIS OPTIONS

Because you enabled output option for MrBayes, Treefinder, or RAxML, partitioning of codon positions is forced to enable.

You enabled partitioning of codon positions.
Do you want to consider nonpartitioning of codon positions? (y/n)
If you say yes, applying nonpartitioned models to all-codon position-concatenated sequences will be considered on each locus. (default: y)
What is your target analysis software? (MrBayes/Treefinder/PAUP/PHYML/RAXML)
(default: RAXML)

OK. The model configuration files for RAXML will be output.

ANALYSIS OPTIONS

Because you enabled output option for MrBayes, Treefinder, or RAXML, partitioning of codon positions is forced to enable.

You enabled partitioning of codon positions.
Do you want to consider nonpartitioning of codon positions? (y/n)
If you say yes, applying nonpartitioned models to all-codon position-concatenated sequences will be considered on each locus.
(default: y)

OK. Nonpartitioning of codon positions is enabled.

You input multiple files.
Do you want to consider nonpartitioning of loci? (y/n)
If you say yes, applying nonpartitioned models to all-loci-concatenated sequences will be considered.
(default: n)

「y」と入力して Enter
OK. The model configuration files for RAxML will be output.

ANALYSIS OPTIONS

Because you enabled output option for MrBayes, Treefinder, or RAxML, partitioning of codon positions is forced to enable.

You enabled partitioning of codon positions.
Do you want to consider nonpartitioning of codon positions? (y/n)
If you say yes, applying nonpartitioned models to all-codon position-concatenated sequences will be considered on each locus.
(default: y)

OK. Nonpartitioning of codon positions is enabled.

You input multiple files.
Do you want to consider nonpartitioning of loci? (y/n)
If you say yes, applying nonpartitioned models to all-loci-concatenated sequences will be considered.
(default: n)

Y
OK. Nonpartitioning of loci is enabled.

You input multiple files and/or protein coding sequence.
Do you want to compare nonpartitioned, equalrate, and separate models on all-loci concatenated sequences? (y/n)
(default: y)
Do you want to consider nonpartitioning of codon positions? (y/n)
If you say yes, applying nonpartitioned models to all-codon position-concatenated sequences will be considered on each locus.
(default: y)

OK. Nonpartitioning of codon positions is enabled.

You input multiple files.
Do you want to consider nonpartitioning of loci? (y/n)
If you say yes, applying nonpartitioned models to all-loci-concatenated sequences will be considered.
(default: n)

Y

OK. Nonpartitioning of loci is enabled.

You input multiple files and/or protein coding sequence.
Do you want to compare nonpartitioned, equalrate, and separate models on all-loci-concatenated sequences? (y/n)
(default: y)

OK. Comparison among nonpartitioned, equalrate, and separate models is enabled.

Because you enabled output option for RAxML, the likelihoods will be calculated by RAxML.

Do you want to optimize the parameters of base composition? (y/n)
(default: n)
Because you enabled output option for RAxML, the likelihoods will be calculated by RAxML.

Do you want to optimize the parameters of base composition? (y/n) (default: n)

OK. The empirical values will be used.

Because you enabled output option for RAxML, the number of rate categories of discrete gamma rate heterogeneity is set to 4.

Because you specify BASEML or RAxML as likelihood calculator, invariant model for among-site rate variation will not be considered.

Because you did not specify BASEML as likelihood calculator, because you disabled the output option for MrBayes, or because you enabled comparison among nonpartitioned, proportional and separate models, N-GAM and autocorrelated discrete gamma models for among-site rate variation will not be considered.

Because you enabled comparison among nonpartitioned, proportional and separate models, a common tree topology will be used for parameter optimization on each locus.

If you want to give tree(s) for parameter optimization, specify an input file name for me. Otherwise, just press enter.
OK. The empirical values will be used.

Because you enabled output option for RAxML, the number of rate categories of discrete gamma rate heterogeneity is set to 4.

Because you specify BASEML or RAxML as likelihood calculator, invariant model for among-site rate variation will not be considered.

Because you did not specify BASEML as likelihood calculator, because you disabled the output option for MrBayes, or because you enabled comparison among nonpartitioned, proportional and separate models, N-GAM and autocorrelated discrete gamma models for among-site rate variation will not be considered.

Because you enabled comparison among nonpartitioned, proportional and separate models, a common tree topology will be used for parameter optimization on each locus.

If you want to give tree(s) for parameter optimization, specify an input file name before me.
Otherwise, just press enter.

OK. Tree(s) for parameter optimization will be made by neighbor-joining based on JC69 distances.

How many processes do you want to run simultaneously? (integer) (default: 1) 4
Because you specify BASEML or RAxML as likelihood calculator, invariant model for among-site rate variation will not be considered.

Because you did not specify BASEML as likelihood calculator, because you disabled the output option for MrBayes, or because you enabled comparison among nonpartitioned, proportional and separate models, N-GAM and autocorrelated discrete gamma models for among-site rate variation will not be considered.

Because you enabled comparison among nonpartitioned, proportional and separate models, a common tree topology will be used for parameter optimization on each locus.

If you want to give tree(s) for parameter optimization, specify an input file name to me. Otherwise, just press enter.

OK. Tree(s) for parameter optimization will be made by neighbor-joining based on JC69 distances.

How many processes do you want to run simultaneously? (integer) (default: 1)

4

OK. The number of processes is set to 4.

Which instruction do your processor support? (SSE2/SSE3/AVX/AVX2) (default: SSE2)

SSE3
Because you enabled comparison among nonpartitioned, proportional and separate models, N-GAM and autocorrelated discrete gamma models for among-site rate variation will not be considered.

If you want to give tree(s) for parameter optimization, specify an input file name. Otherwise, just press enter.

OK. Tree(s) for parameter optimization will be made by neighbor-joining based on JC69 distances.

How many processes do you want to run simultaneously? (integer)  
(default: 1)  
4  
OK. The number of processes is set to 4.

Which instruction do your processor support? (SSE2/SSE3/AVX/AVX2)  
(default: SSE2)  
SSE3  
OK. The likelihoods will be calculated by "raxmlHPC-PTHREADS-SSE3".

All configurations have been completed. Just press enter to run!
Calculating likelihood under EqualRate_CodonEqualRate model which selected by BIC C2 on whole partition...
Calculating likelihood under EqualRate_CodonNonpartitioned model which selected by BIC3 on whole partition...
Calculating likelihood under EqualRate_CodonEqualRate model which selected by BIC C3 on whole partition...
Calculating likelihood under EqualRate_CodonNonpartitioned model which selected by BIC4 on whole partition...
Calculating likelihood under EqualRate_CodonEqualRate model which selected by BIC C4 on whole partition...
Calculating likelihood under EqualRate_CodonNonpartitioned model which selected by BIC5 on whole partition...
Calculating likelihood under EqualRate_CodonEqualRate model which selected by BIC C5 on whole partition...
Calculating likelihood under EqualRate_CodonNonpartitioned model which selected by BIC6 on whole partition...
Calculating likelihood under EqualRate_CodonEqualRate model which selected by BIC C6 on whole partition...
done.
Outputting results...
done.
The model selection has been finished.
The results was output to "C:¥Users¥shimotsuki¥Desktop¥Drosophila¥16S.fas.kakusan¥n".
Please press enter to close.
ダブルクリック
デスクトップのサクラエディタのアイコンへドラッグ&ドロップ
実行時のオプションが保存してある
論文執筆時に必要ならこれを見る
デスクトップのサクラエディタのアイコンへドラッグ＆ドロップ
<table>
<thead>
<tr>
<th>model</th>
<th>BIC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separate_CodonSeparate</td>
<td>1.191732191115e+004</td>
</tr>
<tr>
<td>Separate_CodonNonpartitioned</td>
<td>1.193311948357e+004</td>
</tr>
<tr>
<td>EqualRate_CodonEqualRate</td>
<td>1.203128135019e+004</td>
</tr>
<tr>
<td>EqualRate_CodonNonpartitioned</td>
<td>1.206428898392e+004</td>
</tr>
<tr>
<td>Nonpartitioned</td>
<td>1.211431226779e+004</td>
</tr>
</tbody>
</table>
アライメント長をサンプルサイズとした場合の BIC の値（BIC4 なるものがあるわけではないので注意）
クリック
ダブルクリック
デスクトップのサクラエディタのアイコンへドラッグ＆ドロップ

デスクトップのサクラエディタのアイコンへドラッグ＆ドロップ
OTU 間の塩基組成の均一性が棄却されていないことを確認

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>G</th>
<th>T</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila_mercatorum</td>
<td>524</td>
<td>85</td>
<td>151</td>
<td>559</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_melanogaster</td>
<td>538</td>
<td>79</td>
<td>145</td>
<td>557</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_formosana</td>
<td>541</td>
<td>81</td>
<td>144</td>
<td>553</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_incompta</td>
<td>535</td>
<td>81</td>
<td>150</td>
<td>553</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_albomicans</td>
<td>537</td>
<td>81</td>
<td>147</td>
<td>549</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_santomea_strain_san_BS11</td>
<td>536</td>
<td>80</td>
<td>142</td>
<td>549</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_simulans</td>
<td>540</td>
<td>80</td>
<td>143</td>
<td>556</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_mauritiana</td>
<td>542</td>
<td>83</td>
<td>142</td>
<td>552</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_sechellia</td>
<td>537</td>
<td>81</td>
<td>145</td>
<td>556</td>
<td>131</td>
</tr>
<tr>
<td>Drosophila_yakuba</td>
<td>549</td>
<td>79</td>
<td>142</td>
<td>549</td>
<td>131</td>
</tr>
</tbody>
</table>
同様に内容を確認
遺伝子座間非区分・コドン位置間非区分モデルは最適モデルではなかったので今回これらはどうでもいい