

What mtPhyl can do

The program was initially designed to compare any mtDNA sequence or set of sequences with rCRS and to export the results into Excel table. With passage of time it became clear that other functions necessary to make us happier. Now it can perform tasks listed below:

1. Search sequences which contain particular mutation
2. Search similar sequences
3. Analysis of mutations
4. Export list of mutations in particular mtDNAs into Excel table
5. Phylogenetic tree reconstruction based on maximum parsimony method
6. Calculate coalescence time of cluster
7. Calculate natural selection
8. Make reference list
9. Automatically download human mtDNA complete sequences from GenBank

Data types handled by mtPhyl

mtPhyl can handle files with extension .txt which contains human mtDNA nucleotide sequences. It can be a partial or complete sequence. White spaces and numbers are discarded automatically. Data can be entered in upper or lower register. You can find examples of input files in the folder Examples\DNAs_txt

System requirements

Windows XP

To work properly mtPhyl needs Microsoft Office Excel and PowerPoint (2000, 2003, XP) MtPhyl doesn't require installation. Just unzip it after downloading and run mtPhyl.exe.

List of the files included in the package

mtPhyl file	Description	Required by mtPhyl to run properly
mtPhyl.exe	mtPhyl executable file	+
Cambridge.txt	A text file containing revised Cambridge Reference Sequence (Andrews et al. 1999)	+
ND1.txt, ND2.txt, ND3.txt, ND4.txt, ND4l.txt, ND5.txt, ND6.txt, ATP6.txt, ATP8.txt, Cytb.txt, COI.txt, COII.txt, COIII.txt	A text files containing Index of Conservation for every amino acid in each of 13 mitochondrial protein genes.	+
mtPhylManual.doc	Short user manual	-
Examples	Folder containing examples of input files	-
qtintf70.dll	Library	+

The main algorithm of the mtPhyl that make all computing intensive tasks is in the executable mtPhyl.exe. This program is written in a Delphi.

Bug reports and comments

Please report any bug through the eltsovnp@mail.ru with a note “mtPhyl bugs and comments”. Other comments and suggestions will be also appreciated and can be communicated with us using the same e-mail.

Getting started

Preparing input files. The first step for the analysis of your data is to prepare an input data file for mtPhyl. mtPhyl can handle files with extension .txt which contains human mtDNA nucleotide sequences. It can be a partial or complete sequence. White spaces and numbers are discarded automatically. Data can be entered in upper or lower register.

Also, we have human mtDNA database on our website <http://eltsov.org/mtPhyl.aspx> which contains all human mtDNAs placed in GenBank and is updated monthly.

Once you have your files ready, you must load it into mtPhyl. You can do this by activating the menu File|Add files/Add files from directory. At this point you have to choose which analysis to perform on your data.

Search

You may perform different types of search in this menu

- a) Mutation search – Allows to perform search of sequences which contain or don't contain mutations specified
- b) Search identical sequences from the list of sequences you added
- c) Search similar DNAs from the list of sequences. mtPhyl will ask you to specify sequence of interest and you may specify how many sequences to display. The results will be displayed in the form of table with sequences in order of increasing nucleotide differences and the number of differences between sequences will be displayed in first row.

Show mutations

Shows mutations of sequence selected from Opened files window.

Draw tree

Constructs phylogenetic tree of all sequences in Opened Files and represent it in Microsoft PowerPoint file. Coalescence time may be also calculated – see menu “Coalescent time”.

Different fonts and markers can be used. And if you have large tree which doesn't fit one page, you may choose page puzzle in Geometric Parameters menu to represent your tree on several pages and glue printed pages together. In menu “Not used mutations”, mutations which are not used for phylogenetic tree reconstruction might be designated. In “Pathogenic mutations” menu you may add pathogenic mutations, which will not be used in tree reconstruction.

Example:

- 1) Open files from the folder Examples\DNAs_txt. Push Draw tree button and then OK button, you will have phylogenetic tree (like one in Tree01.ppt file). See main description in Tree01.ppt file in Examples folder.
- 2) Open files from the folder Examples\DNAs_txt. Push Draw tree button and then chose Geometric Parameters menu. In Tree layout select Puzzle and than OK. Now you have tree like the Tree02.ppt. You may easily print this tree and glue pieces together.
- 3) Open files from the folder Examples\DNAs_txt. Push Draw tree button and then chose Fonts and markers menu. In mutation Feature check Protein synonymous, and Show Index of conservation. This will show additional mutation information (like gene, amino acid change and Index of Conservation (Ruiz-Pesini et al, 2004). To highlight mutations in different regions (like protein, tRNA, rRNAs etc) uncheck use same marking for all regions. Additional menus will appear. Double click on the window

with 12345 will allow you to change font. Now click OK button. Now you have tree like the Tree03.ppt. See main description in Tree03.ppt file in Examples folder.

Mutations to Excel menu

By checking the box “Show mutation list in Excel”, mtPhyl will output mutations in Excel table. You may specify the number of sequences per spreadsheet and also use different fonts and markings for displayed mutations in “Fonts and labels” menu appearing after checking “Marking mutation in Excel table” box.

Example:

Open files from the folder Examples\DNAs_txt. Press Mutation to Excel button or activate Sequence/Mutation to Excel. Mutation list will appear in Excel table

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
	Position	CRS	N mut	DQ112837(EU37)	DQ904330(BG-14)	EF667269(MITDNA124)	EF667274(MITDNA138)	EF667412(MITDNA262)	EF667626(MITDNA164)	EF667667(MITDNA163)	EU490797	EU1694385	Map locus	Amino-acidic Change	Conservation Index
1															
2		C	1									+cc	12S rRNA		
3	960	C	2						+C	+C			12S rRNA		
4	1438	A	9	G	G	G	G	G	G	G	G	G	12S rRNA		
5	1700	T	5		C	C	C	C		C			16S rRNA		
6	1721	C	3						T		T	T	16S rRNA		
7	2706	A	9	G	G	G	G	G	G	G	G	G	16S rRNA		
8	3197	T	9	C	C	C	C	C	C	C	C	C	16S rRNA		
9	3861	A	2								G	G	ND1	syn	
10	4769	A	9	G	G	G	G	G	G	G	G	G	ND2	syn	
11	4908	C	1			T							ND2	P147S	64
12	5319	A	2			G	G						ND2	T284A	5
13	5351	A	1						G				ND2	syn	
14	5495	T	3		C			C		C			ND2	syn	
15	5836	A	2								G	G	Y		
16	6023	G	1						A				COI	syn	

Neutrality test

Perform neutrality tests by Elson et al. 2004 and by Ruiz-Pescini et al. 2004 and displays results into excel file with two sheets - Neutrality test (Elson) and Neutrality test (Ruiz-Pesini) respectively. If you want to treat similar sequences as one, check Similar sequences as one

Example:

Open files from the folder Examples\DNAs_txt, then OK. You will be asked to divide mutations into two groups – used and not used ones. After doing so, click OK. You will have Excel book with two spreadsheets:

1) Neutrality test (Elson)

	A	B	C	D	E	F	G	H	I	J	K
1											
2	Gene	No. of haplogroup-associated substitutions			No. of private substitutions			Fisher's Exact Test		Ni	
3		NS	S	NS/S	NS	S	NS/S	one-tailed	two-tailed		
4	ND1	0	1	0	0	1	0	1	1	0	
5	ND2	1	1	1	1	1	1	0,833	1	1	
6	ND3	0	1	0	0	0	0	1	1	0	
7	ND4	0	0	0	0	1	0	1	1	0	
8	ND4L	0	0	0	0	1	0	1	1	0	
9	ND5	2	1	2	0	3	0	0,2	0,2	0	
10	ND6	0	1	0	0	0	0	1	1	0	
11	COI	0	2	0	0	3	0	1	1	0	
12	COII	0	1	0	0	1	0	1	1	0	
13	COIII	0	0	0	2	1	2	1	1	0	
14	Cytb	2	0	0	1	0	0	1	1	0	
15	ATP6	0	0	0	0	0	0	1	1	0	
16	ATP8	0	0	0	0	0	0	1	1	0	
17	Total	5	8	0,62	4	12	0,33	0,57	0,688	0,53	
18											

See Elson et al (2004) for explanation.

2) Neutrality test (Ruiz-Pesini)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
1				Internal nonsynonymous mutations				Internal synonymous mutations			Terminal nonsynonymous mutations				Terminal synonymous mutations							
2	Summary		Position	MutFreq	Map locus	Amino-acidic Change	Conservati Index	Position	MutFreq	Map locus	Position	MutFreq	Map locus	Amino-acidic Change	Conservati Index	Position	MutFreq	Map locus	Position	MutFreq	Map locus	
3	N	9	12346	2	ND5	H4Y	5,13	10262	2	ND3	9333	1	COIII	L43F	89,74	8216	1	COII				
4	RF(I)	0,6	5319	2	ND2	T284A	5,13	6719	2	COI	9667	1	COIII	N154S	23,08	9869	1	COIII				
5	RF(T)	0,3	13637	3	ND5	Q434R	58,97	6629	2	COI	4908	1	ND2	P147S	64,1	3816	1	ND1				
6	CI(I)	41(30)	15218	6	Cytb	T158A	79,49	3861	2	ND1	15314	1	Cytb	A190T	33,33	6023	1	COI				
7	CI(T)	53(26)	14793	6	Cytb	H16R	53,85	13017	3	ND5						6293	1	COI				
8	I/T	1,9						5495	3	ND2						5351	1	ND2				
9								7768	3	COII						6920	1	COI				
10								14182	3	ND6						13434	1	ND5				
11																12582	1	ND5				
12																12064	1	ND4				
13																13353	1	ND5				
14																10685	1	ND4L				
15																						

See Ruiz-Pesini et al (2004) for explanation.

Create sequence

Using this menu you may create any sequence from rCRS (by default) or you may specify initial sequence (by choosing it in OpenFile menu) on the basis of which other sequence can be created. By clicking Add mutation, additional window will appear. In this window you will be asked to specify position and type of mutation. You may "Save" edited sequences or "Save as" newly created sequence under new name or in a new place.

MutaAnalyzer

Will show characteristics of mutation specified.

Reference list

With mtPhyl you may extract references in which sequences have been published.

Example:

Open files from the folder Examples\DNAs_txt. In List Menu activate Show references. Chose folder where sequences in html format from GenBank are located. In our case it is the Folder Examples\DNAs_html. Then click OK. References will be extracted into excel file which is self-explanatory.

References

- 1) Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., and Howell, N. (1999). Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23, 147.
- 2) Elson, J.L., Turnbull, D.M., and Howell, N. (2004). Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection. *Am J Hum Genet* 74, 229-238.
- 3) Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., and Wallace, D.C. (2004). Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* 303, 223-226

We will be glad to help with your data analysis and to hear any suggestions and comments. Together we can make mtPhyl better.