Mitochondrion 11 (2011) 351–356

Contents lists available at ScienceDirect

Mitochondrion

journal homepage: www.elsevier.com/locate/mito

Short communication

MitoTool: A web server for the analysis and retrieval of human mitochondrial DNA sequence variations

Long Fan [a,b], Yong-Gang Yao [a,*]

[a] Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan 650223, China
[b] Graduate School of the Chinese Academy of Sciences, Beijing 100039, China

ARTICLE INFO

Article history:
Received 21 July 2010
Accepted 28 September 2010
Available online 7 October 2010

Keywords:
mtDNA
Haplogroup
Database
Statistics

ABSTRACT

MitoTool, a web-based bioinformatics platform, is designed for deciphering human mitochondrial DNA (mtDNA) data in batch mode. The platform has advantages in (i) parsing diverse types of mtDNA data; (ii) automatically classifying haplogroup according to mtDNA sequences or variants; (iii) discovering possibly missing variants of the samples with claimed haplogroups status; (iv) estimating the evolutionary conservation index, protein coding effect and potential pathogenicity of certain substitutions; (v) performing statistical analysis for haplogroup distribution frequency between case and control groups. Furthermore, it offers an integrated database for retrieving five types of mitochondrion-related information. The MitoTool is freely accessed at http://www.mitotool.org.

1. Introduction

Human mitochondrial DNA (mtDNA) is a 16,569-nt closed double-stranded circular molecule. It contains a 1121-nt non-coding region referred to as the control region or D-loop region, which harbors a series of transcription and replication regulatory segments. Human mtDNA encodes 37 genes including 12 S rRNA, 16 S rRNA, 22 tRNAs, and 13 proteins that are critical for oxidative phosphorylation (OXPHOS) (Anderson et al., 1981; Wallace, 2007).

Due to its prominent properties such as maternal inheritance, absence of recombination, high mutation rate and rich copy number per cell, mtDNA has been widely used as a genetic marker in various fields including population genetics, forensics and biomedicine (Pakendorf and Stoneking, 2005; Salas et al., 2007; Taylor and Turnbull, 2005; Torroni et al., 2006). With the rapid accumulation of mtDNA data during the past decades, a versatile workstation to handle these data and make them accessible for facilitating those studies is urgent to be founded.

There are several customized databases and online tools for mtDNA analysis. For instance, mtSNP (Tanaka et al., 2004), HmtDB (Attimonelli et al., 2005), MITOMAP (Brandon et al., 2005) and mtDB (Ingman and Gyllensten, 2006) are current repositories of human mtDNA genome sequences and their variations. MitoVariome (Lee et al., 2009) contains feature information of many variations in the human mitochondrial genome. MitoRes (Catalano et al., 2006) and MitoProteome (Cotter et al., 2004) provide information on human nuclear-encoded mitochondrial products. Mamit-tRNA (Putz et al., 2007) displays polymorphisms and pathology-related point mutations of human mitochondrial tRNA genes. However, none of these databases offer an integrated workflow to deal with bulk mtDNA data via massive processing, and supply comprehensive functions to analyze mtDNA data. In regards to these online tools, MtSNPscore (Bhardwaj et al., 2009) is beneficial for estimating pathogenicity of mutations in mtDNA; MitoWheel (Zsurka and Csordás, 2009) is a graphical representation of the circular human mitochondrial genome and is very helpful for searching nucleotide positions of the variations and changes of amino acids. MitoAnalyzer (Lee and Levin, 2002), V-MitoSNP (Lee et al., 2008), mtDNAmanager (Lee et al., 2008) and MITOMASTER (Brandon et al., 2009) are useful tools for the analysis of mtDNA. Yet, these online tools are limited to only one type of mtDNA data or are not suitable for parsing massive mtDNA sequences. Besides these online tools, other software has been developed, including eCOMPACT (Weissensteiner et al., 2010), a database-builder for storing and organizing mtDNA data, but not for analyzing it.

In this study, we established a web-based bioinformatics platform "MitoTool" that aims (1) to process diverse types of mtDNA data, (2) to automatically yield a list of variants in certain mtDNA relative to the revised Cambridge Reference Sequence (rCRS) (Andrews et al., 1999) and determine the haplogroup status of that lineage, (3) to appraise the accuracy of the haplogroups classified by users, (4) to display the location of the variant, interspecies conservation index and change of amino acid status, (5) to identify potentially pathogenic mutations based on the reported data, (6) to conduct statistical analysis for haplogroup distribution frequency between case and control groups and (7) to retrieve and batch download analytical

Abbreviations: mtDNA, mitochondrial DNA; rCRS, revised Cambridge Reference Sequence; SNP, single nucleotide polymorphisms; CI, Conservation index; OR, odds ratio.

* Corresponding author. Tel./fax: +86 871 5180085.
E-mail addresses: mitotool@gmail.com (L. Fan), ygyaozh@gmail.com (Y.-G. Yao).

1567-7249/$ – see front matter © 2010 Elsevier B.V. and Mitochondria Research Society. All rights reserved.
doi:10.1016/j.mito.2010.09.013

output and data of interest. While some functions are essentially incorporated into tools such as MitoWheel (Zsurka and Csordás, 2009), MitoAnalyzer (Lee and Levin, 2002), mtDNAmanager (Lee et al., 2008) and others, MitoTool uniquely offers the full array of these functions. Moreover, MitoTool provides processing of various types of mtDNA data in batch mode and does not require user login for access.

2. Material and methods

2.1. Implementation

MitoTool, which runs on a dual-processor server with Ubuntu linux operating system, is implemented under LAMP (Linux-Apache-MySQL-Perl) software stack. The application of mod_perl increases the security and efficiency of our Apache server. Data stored in MySQL are administrated with the help of phpMyAdmin. Web interfaces are shaped by Javascript and Cascading Style Sheet. A set of scripts for data processing, database connectivity, and dynamic generation of web pages are written in Perl with Bioperl modules (http://www.bioperl.org/), DBI module and CGI module.

2.2. Modules

For the purpose of improving the performance of MitoTool, we established four functional modules (Fig. 1): (1) Database module, (2) Haplogroup classification module, (3) Detailed parsing module

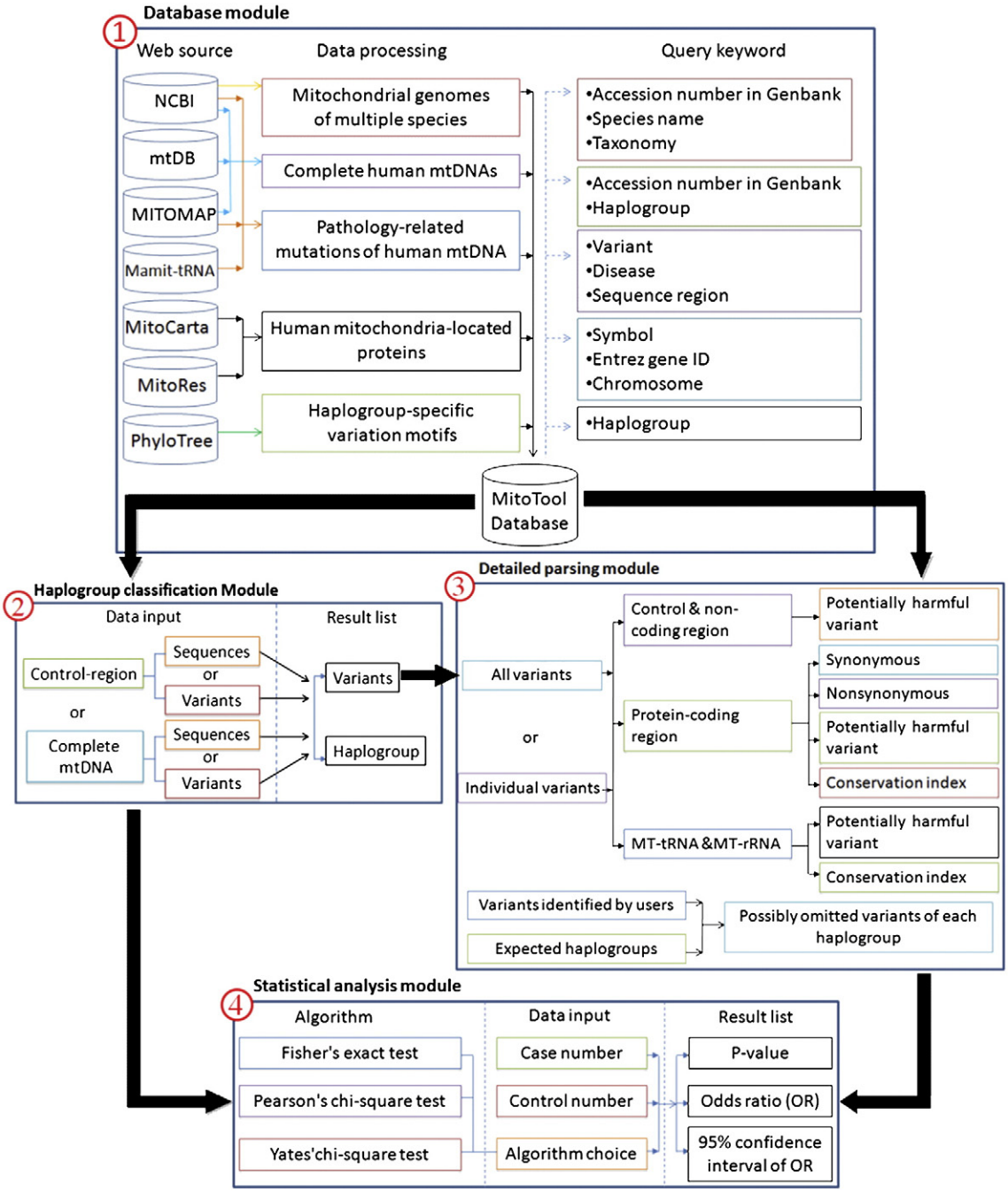

Fig. 1. Schematic view of the structure of MitoTool and its computational pipeline. (1) the database module, (2) the haplogroup classification module, (3) the detailed parsing module and (4) the statistical analysis module. The black arrow indicates the direction of data transmission between two modules.

and (4) Statistical analysis module. These modules can be executed independently or collaboratively as an integrated system.

(1) Database module
As the foundation of the computational system, the database module is crucial for the other three modules. The data gathered from various websites are processed, integrated and meticulously edited before being pooled into the MitoTool database (Table 1). Some rewarding information was extracted as query keywords. The whole module is a semiautomatic process that consists of a suite of perl scripts and manual corrections.

(2) Haplogroup classification module
This module is responsible for determining the haplogroup status of certain mtDNA. First, the users are allowed to upload their data in batch mode. Four distinct types of submitted mtDNA data, including the control–region sequences, complete mtDNA sequences, variants of the control–region sequences or variants of complete mtDNA sequences are acceptable (Fig. 2). Second, each sequence is aligned with the rCRS (Andrews et al., 1999) backstage by Clustal W software (Larkin et al., 2007). The module will decipher the result from Clustal W and then export the variants in each sequence in batch mode. Finally, on the basis of the variants and haplogroup-specific variation motifs (van Oven and Kayser, 2009) stored in the MitoTool database, the haplogroup status of each mtDNA is determined in accordance with the principle of optimal exact matching and fuzzy or near matching. This process allows the user to gather samples' names, variants in each sequence and estimated haplogroup status of each sequence (Fig. 3). If the submitted data are composed of mtDNA variants, the second procedure of the overall process would be bypassed, and only haplogroup status information would be reported.

(3) Detailed parsing module
When the data are transmitted into the third module, all of the variants are first assorted into three classes according to their locations in control region and non-coding region, protein coding region and tRNA and rRNA coding region. Afterwards, each variant in the three different classes gains a distinct and concrete description (see Fig. 1 for details). For instance, entering data for a single nucleotide polymorphism in protein coding region will generate a report about its non-synonymous or synonymous status (Fig. 4) and its conservation index (CI) (Ruiz-Pesini et al., 2004) which is based on the count of mitochondrial genomes of 43 species of primates (Table 2). If the variation has been reported to be associated with any disorders, the related references and diseases will be demonstrated. The auxiliary duty of this module is to scan the missing variants, thereby promises users to evaluate the accuracy of haplogroups that they themselves classified, and to detect the possibly missing variants. For instance, if the variants in certain mtDNA control region sequence contained 73, 150, 199, 263, 489 and 16129, and this sample was classified into haplogroup M7b by researchers, this program would remind that two variants (16223 and 16297) are possibly omitted in the queried mtDNA.

(4) Statistical analysis module
Statistical tools play an important role in handling raw data and drawing rational conclusions. Here, a statistical module is programmed for data mining, especially for data from the second and third module. It embodies three algorithms (Fisher's exact test, Pearson's Chi-square test, and Yates' Chi-square test), which are quite frequently used for case–control studies and for deciphering data in batch-input mode. The submodules used in these methods stem from CPAN (http://www.cpan.org/). After this module performs computations on the inputted data, three meaningful parameters including *P*-value, odds ratio (OR) and 95% confidence interval of OR will be exported to the final spreadsheet which can be saved to the office software (for details see Fig. 5).

Table 1
Summary of MitoTool database.

Data type	Description	Resources	References
(1)	1968 (sub)haplogroups and their corresponding variation motifs	PhyloTree (Build 9)	van Oven and Kayser, 2009
(2)	2130 species' mitochondrial genomes	NCBI	
(3)	6159 human complete mtDNAs	mtDB	Ingman and Gyllensten, 2006
		MITOMAP	Brandon et al., 2005
		NCBI	
(4)	436 pathogenic mutations in human mtDNA and 1141 references	MITOMAP	Brandon et al., 2005
		Mamit-tRNA	Putz et al., 2007
		NCBI	
(5)	1012 mitochondria-located proteins	MitoCarta	Pagliarini et al., 2008
		MitoRes	Catalano et al., 2006

(a)
>sample name 1
146, 249d,263,309+C,315+C,750,1438, 3010 ,4769
> sample name 2
249del, A263G, 309insCC, A750G, A1842G, T4592C
…

(b)
>sample name 1
ATGCCAGTCGTTATCTATAGCTTATCTATCATTATTTACTATTCGTT
TAGCAGTCGATGTCATGTCATCGTAGCTAGTGCTAGTCGTATTCG
C…
>sample name 2
ATGCCCGTAGCGTTATCTATAGCTTATCTATCATTATTTACTATTCG
CTAGTGCTAGTTGTCATGTCATCGTAGCTAGTGCTAGTCGTATTC
A…

Fig. 2. MitoTool sample input. (a) Variants in fasta-like format. All of the variations indentified by users should follow International Society for Forensic Genetics guidelines (Carracedo et al., 2000); however, one difference is that the insertions are recorded with "+" or "ins" in MitoTool. Each variant needs to be separated using a comma or a space. (b) Sequences in fasta format.

3. Results

3.1. Database summary

MitoTool currently contains five types of mtDNA data (Table 1): (1) haplogroup-specific variation motifs, (2) mitochondrial genomes of multiple species, (3) human complete mtDNA sequences, (4) human mtDNA pathogenic mutations and (5) human mitochondria-located proteins. All of these data are collected from various public databases, e.g. NCBI, MITOMAP (Brandon et al., 2005), and mtDB (Ingman and Gyllensten, 2006). The data were processed exhaustively and incorporated systematically, then imported into MitoTool. All of the data can be retrieved easily and efficiently through keyword searches (for details see Fig. 1).

3.2. Performance of each module

To test the efficiency of the haplogroup classification module, we downloaded 6159 human complete mtDNA sequences from GenBank and classified them by MitoTool. Among these mtDNAs, the haplogroups of 99.8% of the sequences, which are utilized to construct the human mtDNA tree (van Oven and Kayser, 2009), have been classified correctly through exact matching and fuzzy or near

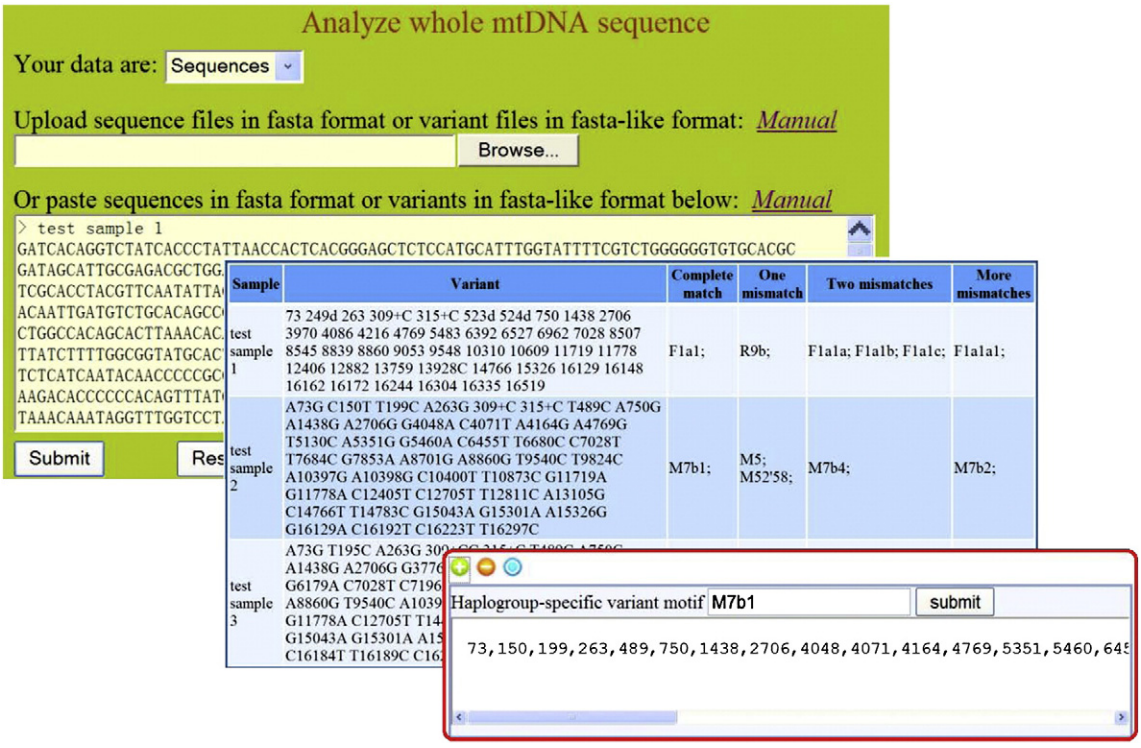

Fig. 3. Input form and result page of haplogroup classification module. The floating window, emphasized here by the red box, is mobile and can be opened and closed at the user's convenience to query haplogroup variation motifs.

matching. Moreover, after repeated checking, we can confirm that the incorrect classifications result from the absence of many expected variants in the tested sequences. The detailed parsing module has been verified by parallel running of the MitoWheel (Lee and Levin, 2002) and MitoAnalyzer (Lee and Levin, 2002) and subsequently checked with MITOMAP (Brandon et al., 2005). The results from the statistical analysis module have been double checked using the R version 2.11.0 (http://www.r-project.org/) and SPSS 11.5 (SPSS Inc., Chicago, Illinois, USA). In brief, we validated all three modules by comparing analytical results which were exported from different software, online tools and MitoTool: these verifications indicated that the execution of our platform is reliable and consistent.

4. Discussion

To date, there are several resources that are similar to certain component of MitoTool. MitoAnalyzer (Lee and Levin, 2002), for example, can only individually determine the effect of single variation as non-synonymous or synonymous variant in certain mtDNA coding gene. V-MitoSNP (Chuang et al., 2006) has an advantage for RFLP genotyping but not for the analysis of variants and haplogroup classification. mtDNAmanager (Lee et al., 2008) is an application for haplogroup estimation according to the variants of control–region sequences which need to be manually identified by the users; additionally, the accuracy of its haplogroup classification is limited

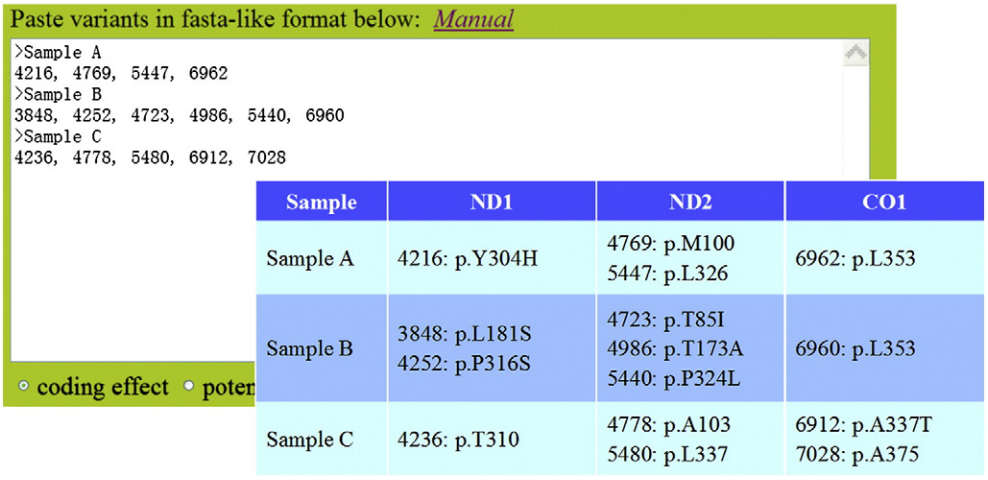

Fig. 4. A screenshot of the report from the detailed parsing module. It shows the result of the non-synonymous or synonymous status of the variants.

Table 2
mtDNA information of forty-three primate species for estimating conservation index.

Species name	GenBank accession number
Cebus albifrons	NC_002763
Chlorocebus aethiops	NC_007009
Chlorocebus pygerythrus	NC_009747
Chlorocebus sabaeus	NC_008066
Chlorocebus tantalus	NC_009748
Colobus guereza	NC_006901
Daubentonia madagascariensis	NC_010299
Eulemur fulvus fulvus	NC_012766
Eulemur fulvus mayottensis	NC_012769
Eulemur macaco macaco	NC_012771
Eulemur mongoz	NC_010300
Galago senegalensis	NC_012761
Gorilla gorilla	NC_001645
Gorilla gorilla gorilla	NC_011120
Homo sapiens	NC_012920
Homo sapiens neanderthalensis	NC_011137
Hylobates lar	NC_002082
Lemur catta	NC_004025
Loris tardigradus	NC_012763
Macaca fascicularis	NC_012670
Macaca mulatta	NC_005943
Macaca sylvanus	NC_002764
Macaca thibetana	NC_011519
Nasalis larvatus	NC_008216
Nycticebus coucang	NC_002765
Otolemur crassicaudatus	NC_012762
Pan paniscus	NC_001644
Pan troglodytes	NC_001643
Papio hamadryas	NC_001992
Perodicticus potto	NC_012764
Piliocolobus badius	NC_008219
Pongo abelii	NC_002083
Pongo pygmaeus	NC_001646
Presbytis melalophos	NC_008217
Propithecus coquereli	NC_011053
Pygathrix nemaeus	NC_008220
Rhinopithecus roxellana	NC_008218
Saimiri sciureus	NC_012775
Semnopithecus entellus	NC_008215
Tarsius bancanus	NC_002811
Tarsius syrichta	NC_012774
Trachypithecus obscurus	NC_006900
Varecia variegata variegata	NC_012773

by the data existing in its database. MITOMASTER (Brandon et al., 2009), a tool for the analysis of mtDNA sequences, resembles MitoTool essentially in data procedure. Nevertheless, it does not include the statistical function, the database query function and the missing variant inspection function. In addition, MITOMASTER is incapable of assigning haplogroups on the basis of control–region sequences and the manually inputted mtDNA variants. Because MITOMASTER is undergoing a complete revision at the present time, we were unable to test it further to compare with MitoTool. Taken together, the available tools for mtDNA are restricted to dealing with only a small amount of data or only one data type. Since MitoTool possesses an auxiliary statistical tool, a valuable database and the additional advantages of the capability to analyze four kinds of mtDNA data and to automatically process data in batch mode, it is versatile enough to cope with the emerging problem of massive mtDNA data analysis.

Coupled with the expansion of mtDNA data, a large number of errors are arising. Fortunately, phylogenetic analysis of these available sequences could help to tag most of the errors and maintain high-quality data (Bandelt et al., 2001; Salas et al., 2005; Yao et al., 2004; Yao et al., 2009). Hence, we enable MitoTool to perform the precise and fault-tolerant haplogroup classification according to control–region sequences or whole mtDNA genome sequences. Inconsistency from exact matching and fuzzy matching reveals that some errors might occur during sequencing or data entry, so it is suggested that researchers recheck their sequences. Moreover, MitoTool enables users to download similar mtDNA sequences whose haplogroups have already been classified.

To regularly bring the database up to date is essential for improving the reliability and timeliness of MitoTool. Therefore, we plan to renew and store more data onto the MitoTool database every four months. Transplanting this system to other species, enhancing its functions and developing a stand-alone version of MitoTool have also been put on our agenda.

5. Conclusion

MitoTool serves to comprehend mtDNA data and browse online information concerning mitochondrion. Compared with existing resources, it is outstanding in the following four aspects: (1) batch analysis of diverse input contents, (2) agile multifunction and higher accuracy, (3) easy manipulation without login requirement and (4) regular updates. Furthermore, with the advancement of dynamic visualization and its functional extensions, it is highly promising that MitoTool will be a powerful and versatile tool for mtDNA studies.

Acknowledgements

We are grateful to the contributors and maintainers of Comprehensive Perl Archive Network and Bioperl project, to Yuan Hong for his

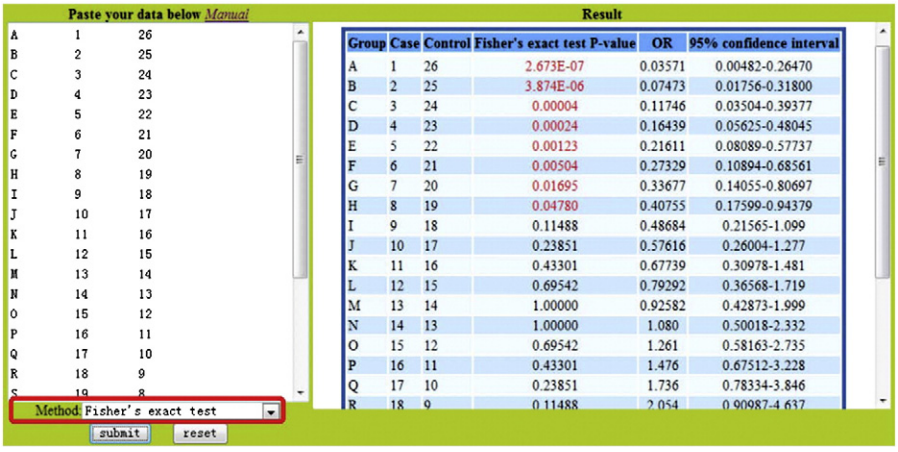

Fig. 5. Interface of statistical analysis module. The left section of this graph is the input form; the right is the result export form. The researchers can choose one algorithm from the pull-down list marked in the red rectangle. *P*-value with significant difference (*P*-value<0.05) is labeled in red.

advice on programming, to Deng-feng Zhang, Fu-hui Xiao and Yu Fan for helping with data input and to the other members of Yao's laboratory for helpful discussions. We also appreciate the language editing done by Miss Diana Chen. This work was supported by the National Science Foundation of China (30925021), Yunnan Province (云南省高端人才计划2009CI119) and the Chinese Academy of Sciences.

References

Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J., Staden, R., Young, I.G., 1981. Sequence and organization of the human mitochondrial genome. Nature 290, 457–465.

Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowlers, R.N., Turnbull, D.M., Howell, N., 1999. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat. Genet. 23, 147.

Attimonelli, M., Accetturo, M., Santamaria, M., Lascaro, D., Scioscia, G., Pappada, G., Russo, L., Zanchetta, L., Tommaseo-Ponzetta, M., 2005. HmtDB, a human mitochondrial genomic resource based on variability studies supporting population genetics and biomedical research. BMC Bioinform. 6 (Suppl 4), S4.

Bandelt, H.-J., Lahermo, P., Richards, M., Macaulay, V., 2001. Detecting errors in mtDNA data by phylogenetic analysis. Int. J. Leg. Med. 115, 64–69.

Bhardwaj, A., Mukerji, M., Sharma, S., Paul, J., Gokhale, C.S., Srivastava, A.K., Tiwari, S., 2009. MtSNPscore: a combined evidence approach for assessing cumulative impact of mitochondrial variations in disease. BMC Bioinform. 10 (Suppl 8), S7.

Brandon, M.C., Lott, M.T., Nguyen, K.C., Spolim, S., Navathe, S.B., Baldi, P., Wallace, D.C., 2005. MITOMAP: a human mitochondrial genome database — 2004 update. Nucleic Acids Res. 33, D611–D613.

Brandon, M.C., Ruiz-Pesini, E., Mishmar, D., Procaccio, V., Lott, M.T., Nguyen, K.C., Spolim, S., Patil, U., Baldi, P., Wallace, D.C., 2009. MITOMASTER: a bioinformatics tool for the analysis of mitochondrial DNA sequences. Hum. Mutat. 30, 1–6.

Carracedo, A., Bar, W., Lincoln, P., Mayr, W., Morling, N., Olaisen, B., Schneider, P., Budowle, B., Brinkmann, B., Gill, P., Holland, M., Tully, G., Wilson, M., 2000. DNA commission of the international society for forensic genetics: guidelines for mitochondrial DNA typing. Forensic Sci. Int. 110, 79–85.

Catalano, D., Licciulli, F., Turi, A., Grillo, G., Saccone, C., D'Elia, D., 2006. MitoRes: a resource of nuclear-encoded mitochondrial genes and their products in Metazoa. BMC Bioinform. 7, 36.

Chuang, L.-Y., Yang, C.-H., Cheng, Y.-H., Gu, D.-L., Chang, P.-L., Tsui, K.-H., Chang, H.-W., 2006. V-MitoSNP: visualization of human mitochondrial SNPs. BMC Bioinform. 7, 379.

Cotter, D., Guda, P., Fahy, E., Subramaniam, S., 2004. MitoProteome: mitochondrial protein sequence database and annotation system. Nucleic Acids Res. 32, D463–D467.

Ingman, M., Gyllensten, U., 2006. mtDB: human mitochondrial genome database, a resource for population genetics and medical sciences. Nucleic Acids Res. 34, D749–D751.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23, 2947–2948.

Lee, M.S., Levin, B.C., 2002. MitoAnalyzer, a computer program and interactive web site to determine the effects of single nucleotide polymorphisms and mutations in human mitochondrial DNA. Mitochondrion 1, 321–326.

Lee, H.Y., Song, I., Ha, E., Cho, S.-B., Yang, W.I., Shin, K.-J., 2008. mtDNAmanager: a web-based tool for the management and quality analysis of mitochondrial DNA control–region sequences. BMC Bioinform. 9, 483.

Lee, Y.S., Kim, W.-Y., Ji, M., Kim, J.H., Bhak, J., 2009. MitoVariome: a variome database of human mitochondrial DNA. BMC Genomics 10 (Suppl 3), S12.

Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.-E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., Hill, D.E., Vidal, M., Evans, J.G., Thorburn, D.R., Carr, S.A., Mootha, V.K., 2008. A mitochondrial protein compendium elucidates complex I disease biology. Cell 134, 112–123.

Pakendorf, B., Stoneking, M., 2005. Mitochondrial DNA and human evolution. Annu. Rev. Genomics Hum. Genet. 6, 165–183.

Putz, J., Dupuis, B., Sissler, M., Florentz, C., 2007. Mamit-tRNA, a database of mammalian mitochondrial tRNA primary and secondary structures. RNA 13, 1184–1190.

Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., Wallace, D.C., 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303, 223–226.

Salas, A., Yao, Y.-G., Macaulay, V., Vega, A., Carracedo, A., Bandelt, H.-J., 2005. A critical reassessment of the role of mitochondria in tumorigenesis. PLoS Med. 2, e296.

Salas, A., Bandelt, H.-J., Macaulay, V., Richards, M.B., 2007. Phylogeographic investigations: the role of trees in forensic genetics. Forensic Sci. Int. 168, 1–13.

Tanaka, M., Cabrera, V.M., Gonzalez, A.M., Larruga, J.M., Takeyasu, T., Fuku, N., Guo, L.J., Hirose, R., Fujita, Y., Kurata, M., Shinoda, K., Umetsu, K., Yamada, Y., Oshida, Y., Sato, Y., Hattori, N., Mizuno, Y., Arai, Y., Hirose, N., Ohta, S., Ogawa, O., Tanaka, Y., Kawamori, R., Shamoto-Nagai, M., Maruyama, W., Shimokata, H., Suzuki, R., Shimodaira, H., 2004. Mitochondrial genome variation in eastern Asia and the peopling of Japan. Genome Res. 14, 1832–1850.

Taylor, R.W., Turnbull, D.M., 2005. Mitochondrial DNA mutations in human disease. Nat. Rev. Genet. 6, 389–402.

Torroni, A., Achilli, A., Macaulay, V., Richards, M., Bandelt, H.-J., 2006. Harvesting the fruit of the human mtDNA tree. Trends Genet. 22, 339–345.

van Oven, M., Kayser, M., 2009. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. 30, E386–E394.

Wallace, D.C., 2007. Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. Annu. Rev. Biochem. 76, 781–821.

Weissensteiner, H., Schonherr, S., Specht, G., Kronenberg, F., Brandstatter, A., 2010. eCOMPACT integrates mtDNA: import, validation and export of mitochondrial DNA profiles for population genetics, tumour dynamics and genotype-phenotype association studies. BMC Bioinform. 11, 122.

Yao, Y.-G., Bravi, C.M., Bandelt, H.-J., 2004. A call for mtDNA data quality control in forensic science. Forensic Sci. Int. 141, 1–6.

Yao, Y.-G., Salas, A., Logan, I., Bandelt, H.-J., 2009. mtDNA data mining in GenBank needs surveying. Am. J. Hum. Genet. 85, 929–933.

Zsurka, G., Csordás, A., 2009. MitoWheel, visualizing the human mitochondrial genome Nature Precedings http://dx.doi.org/10.1038/npre.2009.3167.1.