MITOMAP: an update on the status of the human mitochondrial genome database

Andreas M. Kogelnik^{1,2}, Marie T. Lott¹, Michael D. Brown¹, Shamkant B. Navathe² and Douglas C. Wallace^{1,*}

¹Department of Genetics and Molecular Medicine, Emory University School of Medicine, 1462 Clifton Road, Suite 446, Atlanta, GA 30322, USA and ²Bioengineering Program, College of Computing, Georgia Institute of Technology, Atlanta, GA 30332-0280, USA

Received October 1, 1996; Accepted October 8, 1996

ABSTRACT

We have continued to develop MITOMAP, a comprehensive database for the human mitochondrial DNA (mtDNA). MITOMAP uses the mtDNA sequence as the unifying element for bringing together information on mitochondrial genome structure and function, pathogenic mutations and their clinical characteristics, population associated variation and gene-gene interactions. As increasingly larger regions of the human genome are sequenced and characterized, the need for integrating such information will grow. Consequently, MITOMAP not only provides a valuable reference for the mitochondrial biologist, it will also provide a model for the development of comprehensive, multi-media information storage and retrieval systems for other components of the human genome. This paper is an update of the changes which have occurred to MITOMAP over the past year.

INTRODUCTION

MITOMAP (1) is a database of all published data relating to the human mitochondrial genome. It is available through the World Wide Web and is managed using our own database software which we plan to make available to the public later this year. The purpose of this paper is to document additions to the database and describe significant changes which have been implemented within the past year.

BACKGROUND

The human mitochondrial DNA (mtDNA) is a 16 569 nucleotide pair (np) closed, circular molecule located within the cytoplasmic mitochondria (Fig. 1), the first component of the human genome to be completely sequenced (2). Each of the several thousand mtDNAs per cell encodes a control region encompassing a replication origin and the promoters, a large (16S) and small (12S) rRNA, 22 tRNAs and 13 polypeptides. All of the mtDNA polypeptides are components of the mitochondrial energy

generating pathway, oxidative phosphorylation (OXPHOS), which is functionally essential and evolutionarily constrained (4).

The maternally inherited mtDNA has a very high mutation rate (3). This has resulted in a wide variety of pathologic mutations and neutral polymorphisms. MITOMAP attempts to integrate the broad spectrum of available molecular, genetic, functional and clinical information into a single unified entity which can be queried from a variety of different perspectives.

MITOMAP ARCHITECTURE

MITOMAP is currently implemented using our own in-house database software. MITOMAP was initially implemented as a relational database (Sybase) and subsequently as an object-oriented database (ObjectStore). However, due to limitations inherent in these technologies we have created our own information management system. MITOMAP utilizes an ASN.1 data format. It is both a self-contained information system for the mitochondrial biologist and is linked to the Genome Database (4) of the international Human Genome Organization (HUGO) and to On-Line Mendelian Inheritance of Man (5), thus providing mtDNA map and clinical resources to these entities.

The architecture of the five overlapping data sets of MITOMAP remains unchanged (Fig. 2). The first set, the 'standard' mtDNA sequence (2), is utilized as the key unifying element for interrelating the remaining four elements: the functional genetic element data set, the clinical mutation data set, the population variation data set and the gene-gene interaction data set.

The functional genetic element data set provides the genomic location of the known functional domains of the mtDNA, defined by nucleotide position. It also provides information on the amino acid sequence of proteins, structure of RNAs and sequences of the regulatory elements.

Disease associated nucleotide positions and base changes are listed in the clinical mutation data set which has been extended to cover >60 base substitutions (Table 1). In addition >200 mtDNA rearrangements are included, documenting nucleotide positions of breakpoint junctions and sequences of associated repeat elements. The clinical characteristics associated with the mutations are accessible both through associated data sets of MITOMAP as well

* To whom correspondence should be addressed. Tel: +1 404 727 5624; Fax: +1 404 727 3949; Email: dwallace@gmm.gen.emory.edu

Figure 1. Human mtDNA structural and morbid maps. The tRNAs are indicated by their cognate amino acid single letter code and gene designations are defined in (3). The genes encoded by the G-rich heavy (H) strand are on the outside of the circle, while those for the C-rich light (L) strand are on the inside. The H- and L-strand origins (O_H and O_L) and promoters (P_H and P_L) are shown. The positions of representative pathologic base substitutions are shown on the inside of the circle. ADPD, Alzheimer and Parkinson disease; DEAF, neurosensory hearing loss; LHON, Leber's hereditary optic neuropathy; LDYT, Leber's hereditary optic neuropathy and dystonia; MELAS, mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes; MERRF, myoclonic epilepsy and ragged red muscle fibers; NARP, neurogenic muscular weakness, ataxia, and retinitis pigmentosa. Two representative rearrangements are shown outside the circle, the common 5 kb deletion associated with ocular myopathy and Pearson marrow/pancreas syndrome and the 6.1 kb insertion observed in the maternally inherited diabetes and deafness family harboring the trimolecular heteroplasmy (1,3,4). Reproduced with permission of the American Society for Human Genetics.

as through linkage to OMIM. In addition, a multimedia library of pathogenic and normal phenotypes is now available.

The population variation data set provides access to known polymorphic sites (Table 1). These include restriction site polymorphisms, small insertion-deletion variants and identified sequence changes. The population associations of highly informative variants are provided through available information of mtDNA haplotypes and the continental distributions and population frequencies.

The gene-gene interactions data set catalogs known information on the polypeptide associations within the OXPHOS enzymes. It also provides information on nuclear genes which impinge on mtDNA structure and function.

Table 1. Summary of mutations among the major regions of mtDNA

mtDNA regions	Total no. genes	Total no. nucleotides	No. of mutations with disease	Total no. of polymorphisms
mRNA	13	11388	28	393
tRNA	22	1509	32	37
rRNA	2	2513	2	64
control region ^a	16	1121	0	390

^aThe control region spans roughly from between tRNA phenylalanine to tRNA proline.

Figure 2. The primary database areas and their interrelationship with examples of the types of information which each currently documents. nDNA, nuclear DNA; mtDNA, mitochondrial DNA.

MITOMAP ACCESS

The mitochondrial database is available to the general public through the World Wide Web:
<http://www.gen.emory.edu/mitomap.html> (Fig. 3).

The interface provides both browsing and querying capabilities. Users can browse through the database in its published flat file format [adapted from (4)]. In addition a query interface is provided to perform searches on specific aspects of the mitochondrial genome. For example, the database can be used to answer questions such as 'are there any polymorphisms or mutations within a particular gene, and if so, have they been reported to cause disease?'

CITING MITOMAP

To reference MITOMAP in the published research, please cite this paper and include the version number and date/time stamp from the WWW page.

DATA ACQUISITION

Data is taken from published works on the mitochondrial genome. The committee regularly searches the literature for new publications. The database is updated as new data is obtained, WWW pages and query responses are generated on-the-fly, providing users with the most up-to-date information possible. Submissions should be sent to Dr Douglas C. Wallace, attn: Mitochondrial Genome Committee, Department of Genetics and Molecular Medicine, Emory University, 1462 Clifton Road NE, Atlanta, GA 30322, USA or mitomap@infinity.gen.emory.edu.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Emory-Georgia Tech Center for Biotechnology (S.B.N., M.D.B. and D.C.W.) by a National Library of Medicine Fellowship (A.M.K.) and by NIH grants GM46915, NS21328, HL30164, AG10130, AG13154 and a Muscular Dystrophy Foundation clinical research grant (D.C.W.).

REFERENCES

- 1 Kogelnik, A.M., Lott, M.T., Brown, M.D., Navathe, S.B. and Wallace, D.C. (1996) MITOMAP: a human mitochondrial genome database. *Nucleic Acids Res.* **24**, 177–179.
- 2 Anderson, S., Bankier, A.T., Barrell, B.G., de Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C. *et al.* (1981). Sequence and organization of the human mitochondrial genome. *Nature* **290**, 457–465.
- 3 Wallace, D.C. (1992) Diseases of the Mitochondrial DNA. *Annu. Rev. Biochem.* **61**, 1175–1212.
- 4 Fasman, K.H., Letovsky, S.I., Cottingham, R.W. and Kingsbury, D.T. (1996) Improvements to the GDB Human Genome Data Base. *Nucleic Acids Res.* **24**, 57–63.
- 5 Pearson, P., Francomano, C., Foster, P., Bocchini, C., Li, P. and McKusick, V. (1994) The status of online Mendelian inheritance in man (OMIM) medio 1994. *Nucleic Acids Res.* **22**, 3470–3473.
- 6 Wallace, D.C., Lott, M.T., Brown, M.D., Huoponen, K. and Torroni, A. (1995) Report of the committee on human mitochondrial DNA. In Cuticchia, A.J. (ed.) *Human Gene Mapping 1995: A Compendium*. Johns Hopkins University Press, Baltimore, pp. 910–954 (also available at <http://www.gen.emory.edu/mitomap.html>)
- 7 Wallace, D.C. (1995) Mitochondrial DNA Variation in Human Evolution, Degenerative Disease, and Aging. *Am. J. Hum. Genet.* **57**, 201–223.

MITOMAP v3.0

A human mitochondrial genome database

A. M. Kogelnik ^{1,2}, M.T. Lott ¹, M.D. Brown ¹, S.B. Navathe ², D.C. Wallace ¹

¹. Department of Genetics and Molecular Medicine, Emory University, Atlanta, Georgia

². Bioengineering Program, College of Computing, Georgia Institute of Technology, Atlanta, Georgia

Search MITOMAP for information on:

Gene, disease, enzyme names may be abbreviated, truncated, etc.

Adapted from: The Report of the Committee on the Human Mitochondrial Genome 1995

D.C. Wallace, M. T. Lott, and M. D. Brown

- [Mitochondrial DNA Functional Map](#)
- [Mitochondrial DNA Variation](#)
- [Mitochondrial DNA Mutations and Disease](#)
- [Somatic Mitochondrial DNA Mutations in Aging and Degenerative Diseases](#)
- [Figure 1: Mitochondrial DNA Function Map](#)
- [Figure 2: Eleven pathological mutations in the MTTL1 gene](#)
- [Table 1: Mitochondrial DNA Function Locations](#)
- [Table 2: Mitochondrial DNA Polypeptide Assignments](#)
- [Table 3: Polymorphic MtDNA Restriction Sites Detected by Low Resolution Screening of Human Populations](#)
- [Table 4: Polymorphic MtDNA Restriction Sites Detected by High Resolution Screening of Human Populations](#)
- [Table 5: Common Continent-Specific mtDNA Variants](#)
- [Table 6: MtDNA Control Region Sequence Polymorphisms](#)
- [Table 7: MtDNA Coding Region Sequence Polymorphisms](#)
- [Table 8: Reported Mitochondrial DNA Base Substitution Diseases: Missense mutations](#)
- [Table 9: Reported Mitochondrial DNA Base Substitution Diseases: rRNA and tRNA mutations](#)
- [Table 10: Reported mtDNA Deletions](#)
- [Table 11: Reported Multiple mtDNA Deletions Within Individuals](#)
- [Table 12: mtDNA Simple Insertions](#)
- [Table 13: mtDNA Complex Rearrangements](#)
- [References](#)

MITOMAP Information

- [About MITOMAP](#)
- [NAR Vol 24\(1\):177-9 - MITOMAP: A human mitochondrial genome database](#)
- [How to cite MITOMAP](#)
- [The Human Mitochondrial Sequence](#) - browsing tools coming soon!!

[Go to the Genetics and Molecular Medicine homepage](#)

Figure 3. The MITOMAP homepage on the World Wide Web. <http://www.gen.emory.edu/mitomap.html>