HUMAN MUTATION 15:57–61 (2000)

*MDI SPECIAL ARTICLE*

# Online Mendelian Inheritance in Man (OMIM)

**Ada Hamosh, Alan F. Scott,* Joanna Amberger, David Valle, and Victor A. McKusick**
*McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland*

**Online Mendelian Inheritance In Man (OMIM) is a public database of bibliographic information about human genes and genetic disorders. Begun by Dr. Victor McKusick as the authoritative reference *Mendelian Inheritance in Man*, it is now distributed electronically by the National Center for Biotechnology Information (NCBI). Material in OMIM is derived from the biomedical literature and is written by Dr. McKusick and his colleagues at Johns Hopkins University and elsewhere. Each OMIM entry has a full text summary of a genetic phenotype and/or gene and has copious links to other genetic resources such as DNA and protein sequence, PubMed references, mutation databases, approved gene nomenclature, and more. In addition, NCBI's neighboring feature allows users to identify related articles from PubMed selected on the basis of key words in the OMIM entry. Through its many features, OMIM is increasingly becoming a major gateway for clinicians, students, and basic researchers to the ever-growing literature and resources of human genetics. Hum Mutat 15:57–61, 2000.** © 2000 Wiley-Liss, Inc.

KEY WORDS: **MDI; OMIM; MIM; central database; medical genetics; electronic publication; mutation database; clinical genetics**

## INTRODUCTION

*Mendelian Inheritance in Man* (MIM) is a compendium of bibliographic material and observations on inherited disorders and genes maintained by geneticists and molecular biologists. Its online counterpart, Online Mendelian Inheritance In Man (OMIM), is freely available on the World Wide Web (http://www.ncbi.nlm.nih.gov/omim/). Unlike other databases that maintain primary sequence, mapping, or reference material, OMIM provides authoritative free text overviews about genetic disorders and gene loci that can be used by students, researchers, or clinical practitioners. Curation of the database and editorial decisions take place at the Johns Hopkins University School of Medicine. Distribution of OMIM and software development are provided by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM).

Dr. Victor A. McKusick conceived MIM in the early 1960s as a catalog of genetic traits. Although available only in print form until the 1980s, it has been maintained on computer since early in its development. In 1987, MIM was made internationally available online as OMIM by the NLM and the current Web interface debuted at NCBI in late 1995. MIM is organized by gene locus; however, its emphasis continues to be on content with medical relevance. A detailed history of MIM, its organization, and editorial policies are available in the 12th edition of the book [McKusick, 1997]. Figure 1 illustrates the growth of the database in terms of number of entries included in each of the editions of the book and through September 1999.

As a comprehensive, authoritative, and timely compendium of information in human genetics, OMIM is an important resource for clinicians, teachers, and researchers. Clinicians can use OMIM as an aid in differential diagnosis by searching the database using key clinical features of a patient. As a teaching tool, OMIM provides students a quick and easy way to find and review essential information about a given gene or genetic disorder. (A MEDLINE search for "cystic fibrosis," for example, yielded 16,973 references as of September, 1999, which would be difficult to prioritize.) In research, OMIM serves as a starting point for clinical researchers who want information about genes related to a particular disorder (e.g., what mutations are known for a gene and how they manifest themselves). For basic scientists, a search

Received 26 June 1999; accepted revised manuscript 7 October 1999.

*Correspondence to: Dr. Alan F. Scott, McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, 600 N. Wolfe Street, Blalock 1033, Baltimore, MD 21287. E-mail: afscott@welch.jhu.edu

©2000 WILEY-LISS, INC.

10981004, 2000, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-1004(20000)15:1<57::AID-HUMU12>3.0.CO;2-G by Cochrane Germany, Wiley Online Library on [01/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

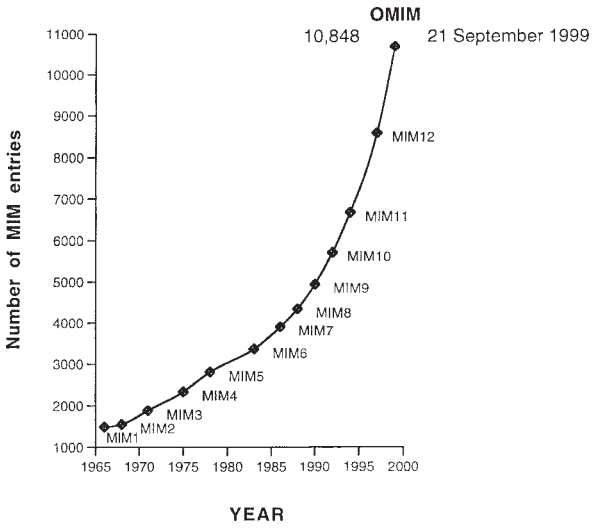



FIGURE 1. **Total number of entries in the printed editions of *Mendelian Inheritance in Man*.**

of OMIM may help in identifying disorders caused by a gene that they have characterized.

Perhaps OMIM's greatest utility is to serve as a simple gateway to related information about genetics. Through its wealth of relevant links, OMIM serves as an easy access point to databases that contain more detailed information about sequence, map information, references, etc. The ability to easily navigate between databases whose structures and searching schemes are quite different has made OMIM especially popular with users whose need for more detailed data is sporadic and who may not be intimately familiar with other databases and how to use them.

## FEATURES OF THE OMIM DATABASE

OMIM entries have a standardized format which include a number of features, described below. Several of these are illustrated in Figure 2.

### The Entry

Each OMIM entry is assigned a unique number. Entries are prepared for each distinct gene or genetic disorder for which sufficient information exists and are not made for genes known only from expressed sequence tags (ESTs), pseudogenes, genetic markers, or even complete cDNAs for which nothing is known other than the sequence. Many disorders that are not yet characterized at the level of the gene or even at any meaningful biochemical level are included as entries if they show Mendelian inheritance. A major goal of OMIM is to help with the discovery process, whereby a gene sequence and a mapped phenotype can be associated.

Information that may be found in gene entries includes approved gene name and symbol (obtained from the Human Gene Nomenclature Committee), alternative names and symbols in common use, information about the map location in human and mouse, method of cloning, information about the protein and DNA sequence (e.g., the size of the gene, the type of product made), functional information, related genes in man or other species, and whether there are animal models. For entries where the gene causes a disease, information about key allelic variants is included and clinical details are presented.

For phenotypes, the distinctive characteristics of given clinical disorders are noted, including variations from the usual case. Clinical information, given in succinct form, is supplemented by that provided in selected citations that accompany each entry.

Text entries are generally diachronic, meaning that they are added in chronological order, with the most recent material at the end. This is done, in part, to minimize the effort of having to rewrite thousands of entries each time they are amended and to reflect the historical progression of the knowledge about the locus. Many of the larger entries, for which there is a wealth of information, have been restructured into topical sections. For these entries, new information is added in chronological order to the appropriate section.

### User Comments

OMIM encourages users to offer comments about existing entries and suggestions for improvements or additions of materials.

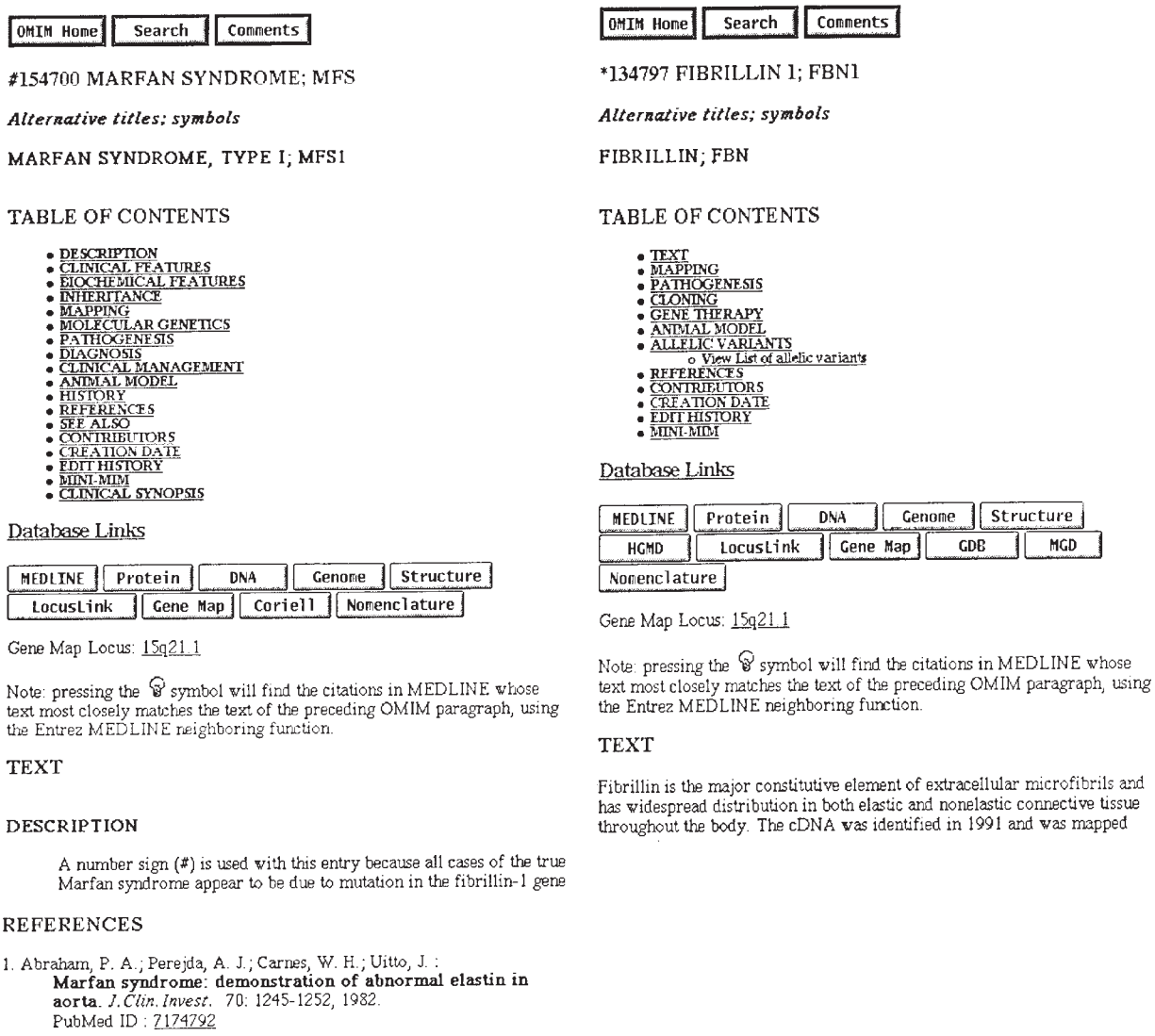
### Clinical Synopses

Clinical synopses are available for entries that describe a phenotype. These single-word listings of signs, symptoms, laboratory tests, and inheritance are written using a controlled vocabulary and provide a quick survey of features of a given disorder. The synopses are particularly useful in creating lists for differential diagnosis.

### Allelic Variants

A valuable feature of OMIM is its lists of noteworthy "allelic variants" for a given gene. These are most often disease-causing mutations, but can also include common polymorphisms that do not produce disease. In each case, the variants are inherited and are distinct from somatic mutations, as seen in cancer, which are generally not included. OMIM does not try to exhaustively document all known variants at a locus, but rather focuses on those that are rela-

10981004, 2000, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-1004(20000)15:1<57::AID-HUMU12>3.0.CO;2-G by Cochrane Germany, Wiley Online Library on [01/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



OMIM Home | Search | Comments

#154700 MARFAN SYNDROME; MFS

*Alternative titles; symbols*

MARFAN SYNDROME, TYPE I; MFS1

TABLE OF CONTENTS

- DESCRIPTION
- CLINICAL FEATURES
- BIOCHEMICAL FEATURES
- INHERITANCE
- MAPPING
- MOLECULAR GENETICS
- PATHOGENESIS
- DIAGNOSIS
- CLINICAL MANAGEMENT
- ANIMAL MODEL
- HISTORY
- REFERENCES
- SEE ALSO
- CONTRIBUTORS
- CREATION DATE
- EDIT HISTORY
- MINI-MIM
- CLINICAL SYNOPSIS

Database Links

MEDLINE | Protein | DNA | Genome | Structure | LocusLink | Gene Map | Coriell | Nomenclature

Gene Map Locus: 15q21.1

Note: pressing the symbol will find the citations in MEDLINE whose text most closely matches the text of the preceding OMIM paragraph, using the Entrez MEDLINE neighboring function.

TEXT

DESCRIPTION

A number sign (#) is used with this entry because all cases of the true Marfan syndrome appear to be due to mutation in the fibrillin-1 gene

REFERENCES

1. Abraham, P. A.; Perejda, A. J.; Carnes, W. H.; Uitto, J. :
**Marfan syndrome: demonstration of abnormal elastin in aorta.** *J. Clin. Invest.* 70: 1245-1252, 1982.
PubMed ID : 7174792

OMIM Home | Search | Comments

*134797 FIBRILLIN 1; FBN1

*Alternative titles; symbols*

FIBRILLIN; FBN

TABLE OF CONTENTS

- TEXT
- MAPPING
- PATHOGENESIS
- CLONING
- GENE THERAPY
- ANIMAL MODEL
- ALLELIC VARIANTS
  - View List of allelic variants
- REFERENCES
- CONTRIBUTORS
- CREATION DATE
- EDIT HISTORY
- MINI-MIM

Database Links

MEDLINE | Protein | DNA | Genome | Structure | HGMD | LocusLink | Gene Map | GDB | MGD | Nomenclature

Gene Map Locus: 15q21.1

Note: pressing the symbol will find the citations in MEDLINE whose text most closely matches the text of the preceding OMIM paragraph, using the Entrez MEDLINE neighboring function.

TEXT

Fibrillin is the major constitutive element of extracellular microfibrils and has widespread distribution in both elastic and nonelastic connective tissue throughout the body. The cDNA was identified in 1991 and was mapped

FIGURE 2. **The Marfan Syndrome and Fibrillin entries illustrate several features of the OMIM database. At the top of the page are buttons that allow the user to perform another search, return to the home page, or make comments about the entry. The OMIM number is shown along with the approved gene name and symbol. Next are listed alternative names and symbols. The Table of Contents, a feature of the longer entries, takes users to specific regions of the text. Where available, links to other resources are presented. Next, the cytogenetic gene map position is given followed by an explanation of the neighboring feature button. MEDLINE references, including abstracts, are available through the PubMed ID link associated with most references.**

tively common, represent a novel mechanism of mutation, or have historic significance.

### The OMIM Gene Map

The OMIM gene map is maintained as a convenience to users and focuses on the "morbid map," i.e., the mapping of disorders. In chromosome-by-chromosome tabular form, the OMIM synopsis of the human gene map gives, for each gene, the chromosomal location, gene symbol, method(s) of mapping, and disorder(s) related to the specific gene. Links to the human/mouse homology maps are also provided. For the mapped disorders, the OMIM gene map also indicates whether specific mutations have been identified. The OMIM gene map is not the official map curated by the genetics mapping community and additional, more detailed, mapping information is available through the links to NCBI and GDB.

### Citations

Citations are highlighted in the text and, if they are cited in PubMed (MEDLINE), the PubMed ID linked to the abstract is listed after the refer-

10981004, 2000, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-1004(200001)15:1<57::AID-HUMU12>3.0.CO;2-G by Cochrane Germany, Wiley Online Library on [01/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ence. References, through their associated PubMed number, have been used to generate most of the links in OMIM (see "Links," below). More recently, OMIM has collaborated with the Human Gene Nomenclature Committee and the LocusLink/RefSeq projects at NCBI to create a table that directly associates an MIM number to a particular gene symbol and GenBank accession number. This will result in more accurate links.

### Credits and Edit History

The Creation Date field lists the date that the entry was created and the name of the person who created it. Authors who contribute significantly to the updating or editing of an entry are given credit in the Contributors field. When changes are made to an entry by the OMIM editing staff, they are documented in the Edit History field with the date and name of the person who made the changes.

### Searching OMIM

Searching OMIM is as simple as typing in the words of interest. In addition, there are a number of ways to restrict a search (i.e., Boolean operators and field restriction). Entries are retrieved and ranked in order of the number of times the search terms occur in the entry and where they occur. For example, terms that appear in the title are given greater weight than those occurring in the text or references.

### Neighboring

A key feature of OMIM is neighboring (indicated by a light bulb icon) which allows users to perform MEDLINE searches using keywords selected from the text of the preceding paragraph. Each paragraph in the text of an OMIM entry is processed against the entire MEDLINE set of citations using the NEIGHBOR algorithm developed at NCBI by B. Brylawski. NEIGHBOR finds the MEDLINE articles that seem most closely related to the OMIM paragraph and links them to the paragraph on the OMIM entry's WWW page. This feature permits the user to find MEDLINE references that are germane to a specific area of interest in addition to the references in the OMIM entry itself. The NEIGHBOR algorithm is updated against MEDLINE regularly.

### Links

The integration of OMIM with other Entrez databases increases the utility of OMIM by allowing direct access from the entry to related information. Other Entrez databases include LocusLink, Unigene, RefSeq, PubMed, protein and nucleic acid databases, mapping databases, etc. Detailed information about these can be found at http://www.ncbi.nlm.nih.gov/Database/index.html. OMIM also provides links to other resources, including Human Gene Nomenclature, Human Gene Mutation Database [HGMD; see Krawczak et al., 2000], locus-specific mutation databases (such as the CF and PAH [Scriver et al., 2000] Mutation Databases), MitoMap (mitochondrial sequence and disorders), Online Mendelian Inheritance in Animals (OMIA) and the Coriell database of cell lines and probes, the Genome Database [GDB; see Cuticchia, 2000], the Mouse Genome Database (MGD), and the Alliance of Genetic Support Groups. Additional links are reviewed and added as deemed appropriate.

### Update Log

An update log is available from the OMIM home page that takes users to a list of entries that have been changed. The list is arranged by the month and within the month by day. New entries are separated from updated entries; this is particularly helpful for users who only want to browse the database for new entries.

## HOW OMIM IS CURATED

The primary source material for OMIM is the published literature. The scientific staff review several leading journals that publish major articles in clinical and molecular genetics. In addition, tables of contents of dozens of additional journals are scanned for relevant articles. First, a search of OMIM is made to assure that the material is not already entered and whether the new reference contains sufficiently important information to add to an existing entry.

If the article constitutes a new gene, then additional processing occurs. For new genes, DNA or protein sequence included in the article is compared to GenBank using the NCBI blast service [Altschul et al., 1990]. Based on the results of that search, additional articles may be identified that have reported the same sequence, often using a different name. If an approved symbol is not available in GDB, then information about the gene is submitted to the HUGO Nomenclature Committee and a request for an approved gene symbol and name is made. The latter are then added to the new OMIM entry. Only senior clinicians decide whether an article describes a new genetic disorder. Unique case reports are not included. OMIM science writers are typically PhDs or MDs with

10981004, 2000, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-1004(2000)15:1<57::AID-HUMU12>3.0.CO;2-G by Cochrane Germany, Wiley Online Library on [01/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

training in genetics. They read the articles and abstract the salient points in the article. Importantly, the science writers attempt to put the work in context and provide sufficient background so that a reader can understand the significance of the work.

OMIM obtains articles prior to publication for several major scientific journals and adds to entries or creates new entries that are released online on the date of publication. In the past year, the database averaged about 100 new entries and 850 changes in existing entries per month. OMIM does not attempt to be an exhaustive review of the literature. Not only would this be an impossible task, given the enormous volume of the world's genetics literature, but it would defeat the useful winnowing process that attempts to select "key" information. Also, the "neighboring" feature of OMIM easily allows the use of the text of a preceding paragraph to search MEDLINE for related articles.

## THE FUTURE OF OMIM

Perhaps the best measure of a database is how often people use it. At the end of 1995, OMIM received about 4,000 user queries per week. By the middle of 1999, OMIM received over 22,000 searches per day (4,000 unique users per day). Unlike some databases, whose growth will diminish as the human genomic sequence is completed and the map locations of all genes are known with precision, OMIM is likely to continue to grow at an expanded rate. The reason for this is that the windfall of sequence will identify tens of thousands of new genes that will become the basis for genetic discovery for scientists worldwide. As the community reports information about the function and medical significance of these sequences, there will be an ongoing need to supplement OMIM with that knowledge.

A challenge to OMIM will be to maintain standards of authoritativeness, thoroughness, and timeliness as it deals with the increasing torrent of new genetic discoveries. Ultimately, we hope that OMIM will be seen as a Rosetta Stone for the genetics community, allowing scientists, students, and clinicians to identify the relationships between genes and diseases.

## REFERENCES

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. J Mol Biol 215:403–410.

Cuticchia AJ. 2000. Future vision of the GDB Human Genome Database. Hum Mutat 15:62–67.

Krawczak M, Ball EV, Fenton I, Stenson PD, Abeysinghe S, Thomas N, Cooper DN. 2000. Human Gene Mutation Database: a biomedical information and research resource. Hum Mutat 15:45–51.

McKusick VA. 1997. Mendelian inheritance in Man. Catalogs of human genes and genetic disorders, 12th ed. Baltimore: Johns Hopkins University Press.

Scriver CR, Waters PJ, Sarkissian C, Ryan S, Prevost L, Cote D, Novak J, Nowacki PM. 2000. PAHdb: a locus-specific knowledgebase. Hum Mutat 15:99–104.

10981004, 2000, 1, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1098-1004(20000)15:1<57::AID-HUMU12>3.0.CO;2-G by Cochrane Germany, Wiley Online Library on [01/12/2022]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License