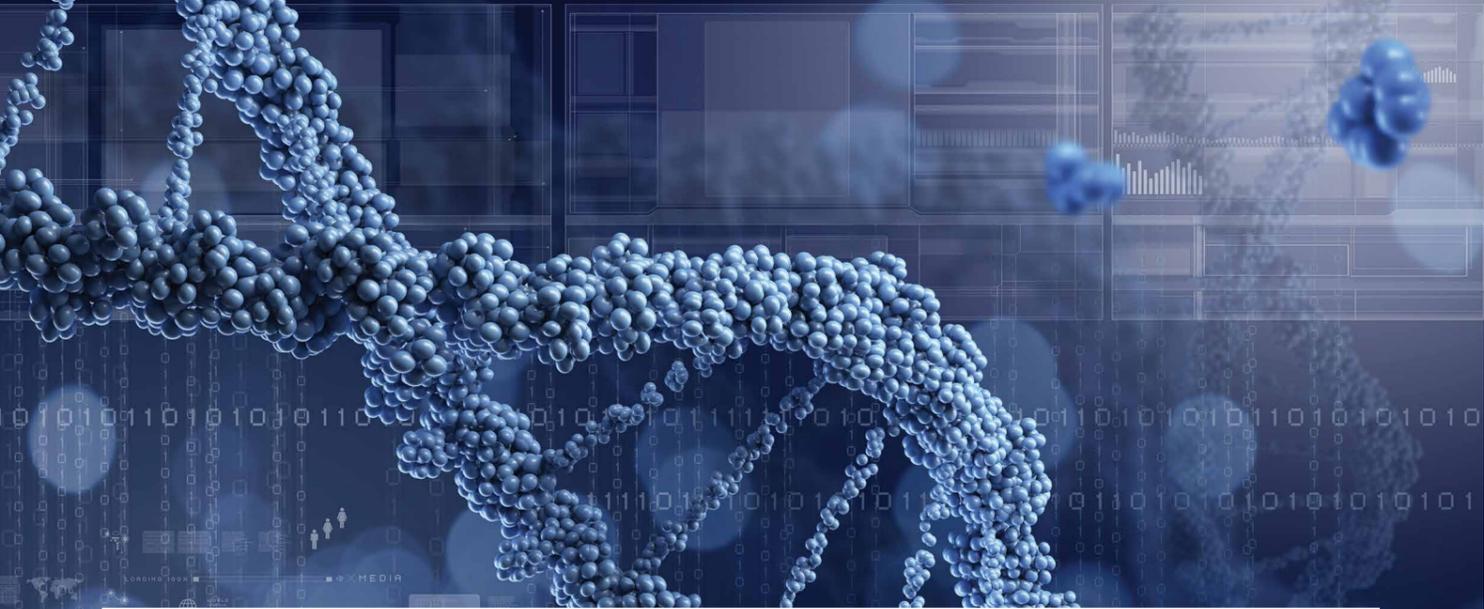




**OMIM® – The Online
Mendelian Inheritance in
Man Knowledgebase:
A Wardrobe Full of Genes**

Ada Hamosh, MD, MPH



OMIM® – THE ONLINE MENDELIAN INHERITANCE IN MAN KNOWLEDGEBASE: A WARDROBE FULL OF GENES

The OMIM knowledgebase provides a map to the complex world of human genetics, linking genes and their associated traits in an easy-to-understand system that is an essential resource used by researchers and clinicians worldwide.

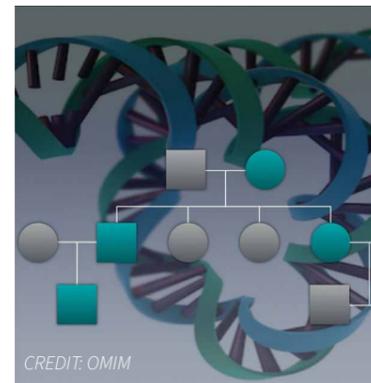
A commercial and technological revolution, the internet is often thought of as a source for fast communication, easy online shopping, and cute videos of baby elephants. However, the ability to access large amounts of information in public databases has led to vast changes in the manner in which science is performed. These changes are most noticeable in the field of life science, particularly genetics, where high-throughput biology yielded data that, when organised into databases, accelerated the understanding of the highly complex systems underlying biology and medicine.

Researchers and clinicians thus need a road-map to the complex pathways involved, such as a central database of genes and diseases and all of the pathways leading from one to the other. This is where the Online Mendelian Inheritance in Man (OMIM) knowledgebase comes into play, a continuously updated catalogue of human genes and their associated traits and disorders. Setting out to clarify the relationship between genes

and phenotypes, OMIM is a vital resource for everyone involved in the steadily-growing field of medical genetics.

Retro Denim

OMIM is an extension of Mendelian Inheritance in Man (MIM), a catalogue of genetic disorders and genes which was published as a rather hefty book in 12 editions between 1966 and 1998. The original author and eventual chief editor of the book was Dr Victor McKusick, one of the pioneers of the medical genetics field. A man with an impressive degree of talent, he was the founder of the well-respected journal *Genomics*, recipient of the Lasker Award, and at one point held no less than four professorships simultaneously. Mendelian Inheritance in Man was one of his side projects, beginning as a catalogue of genetic traits and research thereof in the early 1960s but rapidly growing into a weighty tome which no self-respecting geneticist would be seen without.



CREDIT: OMIM

The advent of the internet meant that information could be shared far more easily than before, and the scientific world jumped at the chance. Because MIM had been stored in electronic form since 1964, it was an ideal resource to be used in pioneering efforts by the National Library of Medicine to create a full-text search engine. The result was OMIM, which became open to internet visitors around the world in 1987, 10 years before the founding of Google. In 1995, OMIM was moved to the World Wide Web by NCBI, the National Center for Biotechnology Information, where it was brought alongside their many different medical and biological databases – the most well-known of these is the literature database PubMed, used by many researchers who need to keep up with publications in their field.

OMIM established their own website (the appropriately named OMIM.org) in 2011 to provide a greater level of control over the presentation of information. Since going online in 1987, OMIM has been updated daily

With well-referenced entries for over 15500 genes and every known Mendelian disease, OMIM has become a cornerstone knowledgebase for the medical genetics field and increasingly all of medicine.



by a small group of expert science writers and curators now under the guidance of Professor Ada Hamosh, of the Johns Hopkins University. Currently, OMIM has over 23,900 well-referenced entries describing over 15500 genes and every known Mendelian disease. Each entry includes copious targeted links to other online databases, and OMIM.org has become a cornerstone resource for almost everyone in the fields of medicine and genomics.

Blue, Grey, Skinny, Flared?

OMIM contains information on both genes and phenotypes, two related but nonetheless differing subjects. A phenotype is a constellation of observable clinical features describing an individual or a disorder – think of the blue eyes and blonde hair of your partner, the black hair of your teacher, your highly specific blood group, or your family member's early-onset Alzheimer disease. All of these observable attributes are related to genetic factors – variations or other alterations in specific genes that are inherited from your parents, i.e. your genotype. The genotype thus leads to the phenotype,

although this process is exceptionally complex and spread across many different interacting factors. A Mendelian trait, the focus of OMIM, is one in which the phenotype is controlled by a single gene – mutations in one gene will lead to an altered state such as a disease (cystic fibrosis is a classic example here – mutations within the CFTR gene inherited from both parents leads directly to the disease).

The evolving knowledge of genes and phenotypes is evident in the content in the MIM books. The first edition of *Mendelian Inheritance in Man*, printed in 1966, had almost 1500 entries. The science of genetics was still in its infancy at that stage; the structure of DNA had been found a mere 13 years prior, while the first genetic sequence was still 6 years in the future. As such, almost every entry referred to a phenotype. By contrast the current database at OMIM has around 7,500 phenotypes alongside 15,500 genes – a clear indication of the rapid growth of genetic sequencing and associated information.

Sorting the Collection

Every entry within OMIM follows a standard format that enables researchers to easily find the required information. Particular attention is given to the relationship between genes and phenotypes. If mutations in a gene are known to be associated with a disease or trait, this is displayed at the top of the entry. Sometimes different mutations in a gene underlie different phenotypes. These allelic phenotypes are listed together in gene-phenotype tables. In other instances, mutations in different genes can lead to phenotypes with the same or highly overlapping clinical features. These phenotypes are brought together in phenotypic series. OMIM entries describing phenotypes are accompanied by a clinical synopsis – an anatomical overview of the clinical features associated with the phenotype. Cataloguing clinical features through both discursive text and synopses provides clinicians with a unique and powerful way to arrive at a diagnosis.

OMIM also has a tabular listing of genes and phenotypes. This genemap provides a way



to view all Mendelian phenotypes within a specified genomic region. Locations of genes and phenotypes are given in cytogenetic or genomic coordinates. The cytogenetic terminology refers to regions observed when viewing a chromosome under a microscope, and the genomic coordinates refer to the number of bases counting along the length of a chromosome. This is comparable to a street mailing address and the corresponding GPS coordinates.

Determining which gene causes a particular disease can be a difficult and time-consuming job. Years ago, geneticists would spend long hours taking samples from extended families with particular diseases and then trying to determine how those samples related to each other, over time building up a map of related data points that could be used to pick out where on the genome the disease-causing mutation must lie. The advent of full-genome sequencing changed this immensely – geneticists now sequence DNA from individuals with a disorder and compare it to online databases such as OMIM to identify genes with a high probability of causing the phenotype. Additional research is performed to validate the relationship between the gene and disease. As in many fields, the existence of a shared knowledgebase helps scientists to find results far faster than otherwise possible.

OMIM also plays a central role in providing a unified system of classifying Mendelian diseases and genes – essentially providing a unifying structure with stable number identifiers which allows researchers to unambiguously refer to each gene/locus/marker or phenotype. The system is fairly simple: each entry is given a unique six-digit number as an identifier: sex-chromosome-linked entries begin with 3 (X-linked) or 4 (Y-linked), and mitochondrial entries begin with 5. Autosomal entries created before 1995 begin with 1 or 2, those made afterwards begin with 6. Each entry number is preceded by a symbol (*, #, %) to indicate whether the entry is a gene or a phenotype and whether the phenotype has a known molecular basis or has not yet been associated with mutations in a specific gene. These numbers are regularly used in scientific literature, even in journals that do not directly relate to genetics, simply to ensure that everyone understands exactly what is being discussed.

A Pair for Everyone

OMIM is intended to be useful to a wide range of professionals, from researchers to clinicians, and students of many disciplines. From a clinical point of view, OMIM comes into its own as a reference for diagnoses – doctors can look at summaries of research on diseases associated with particular genes and then use that to decide how their own patients should be evaluated. OMIM also links to many clinical resources, including databases of laboratories across the world that provide genetic testing services, as well as databases for clinical management and clinical trials. For researchers, OMIM provides a comprehensive resource that can be surveyed for potential scientific correlation – either computationally via bioinformatics programs or by an individual leisurely browsing the entries and maps at OMIM.org. To facilitate research connections and stay current on a topic, OMIM users can tag entries and receive a notification whenever an update occurs. This service, MIMmatch, also connects researchers with similar interests together, as a sort of genetics-focused social media network.

The information in OMIM is sourced from the peer-reviewed biomedical literature accessed through journals and publication databases such as PubMed. Curators also keep a close eye on news feeds and life-science articles. This information is then organised and incorporated into OMIM by a team of science writers and curators headed by Professor Ada Hamosh. Only selected articles are included in OMIM. Additional articles on a topic are readily available through the 'reference plus' links at the end of each paragraph in an OMIM entry. Funding for the project comes from grants, in particular the National Human Genome Research Institute, from licenses to industry, and from public donations.

A Weave for the Future

Over the course of 50 years, OMIM has grown from a single printed book into a vast knowledgebase with thousands of entries describing genes and phenotypes. As research pushes the boundaries of genetic knowledge ever further, OMIM will grow alongside it, providing a map for scientists and medical professionals around the world. How will it look after the next 50 years? We'll just have to wait to find out.



Meet the researcher

Ada Hamosh, MD, MPH

Dr Frank V. Sutland Professor of Pediatric Genetics
McKusick-Nathans Institute of Genetic Medicine
Johns Hopkins University
Baltimore, USA

Professor Ada Hamosh began her medical genetics career with a series of publications covering cystic fibrosis, a prime example of a Mendelian disease. Her interests then broadened to cover the general field of Mendelian diseases, from which she progressed to the role of scientific director of one of the leading genetic knowledgebases, OMIM. Currently located at the Johns Hopkins University, where she is simultaneously a Professor of Paediatrics and Genetic Medicine, she mixes her work at OMIM with further research on discovery of novel Mendelian disease genes and the integration of genetics into general clinical medicine by building tools to facilitate the process.

CONTACT:

T: (+1) 410 614 3313

E: ahamosh@jhmi.edu

W: www.hopkinsmedicine.org/profiles/results/directory/profile/0002818/ada-hamosh

KEY COLLABORATORS

Joanna Amberger, Johns Hopkins University
Carol Bocchini, Johns Hopkins University
Alan F Scott, Johns Hopkins University
Francois Schiettecatte, Johns Hopkins University
David Valle, Johns Hopkins University
Nara Sobreira, Johns Hopkins University

FUNDING

National Human Genome Research Institute

REFERENCES

JS Amberger, CA Bocchini, F Schiettecatte, AF Scott and A Hamosh, OMIM.org: Online Mendelian Inheritance in Man (OMIM), an online catalog of human genes and genetic disorders, Nucleic Acids Research, 2015, 43, Database issue D789–D798.

JS Amberger, CA Bocchini and A Hamosh, A New Face and New Challenges for Online Mendelian Inheritance in Man (OMIM®) Human Mutation, 2011, 32, 564–7.

