Brochure

HGMD®:
Human Gene Mutation Database

The community resource for comprehensive coverage of published human hereditary disease mutations, licensed exclusively through QIAGEN

Sample to Insight

Introduction

The human gene mutation database (HGMD®) represents an up-to-date and comprehensive collection of known and published pathogenic gene lesions responsible for human inherited disease.

HGMD is a database which provides information of practical importance to medical and clinical geneticists, bioinformaticians, researchers in human and molecular genetics and physicians and genetic counselors interested in a particular inherited condition in a given patient or family. HGMD is a widely used, trusted resource that has been cited in over 5000 publications in leading scientific journals.

HGMD is available as a free public version with restricted content and limited search options for academic use only and as a fully functional professional version that requires annual subscription through QIAGEN.

Key Capabilities

- Easily verify whether an observed mutation has been previously described to be responsible for causing human inherited disease
- Obtain an overview of the pathogenic mutational spectrum of a particular gene or disease
- Quickly access detailed reports for disease-associated human inherited mutations

Figure 1. HGMD Professional sample mutation report.

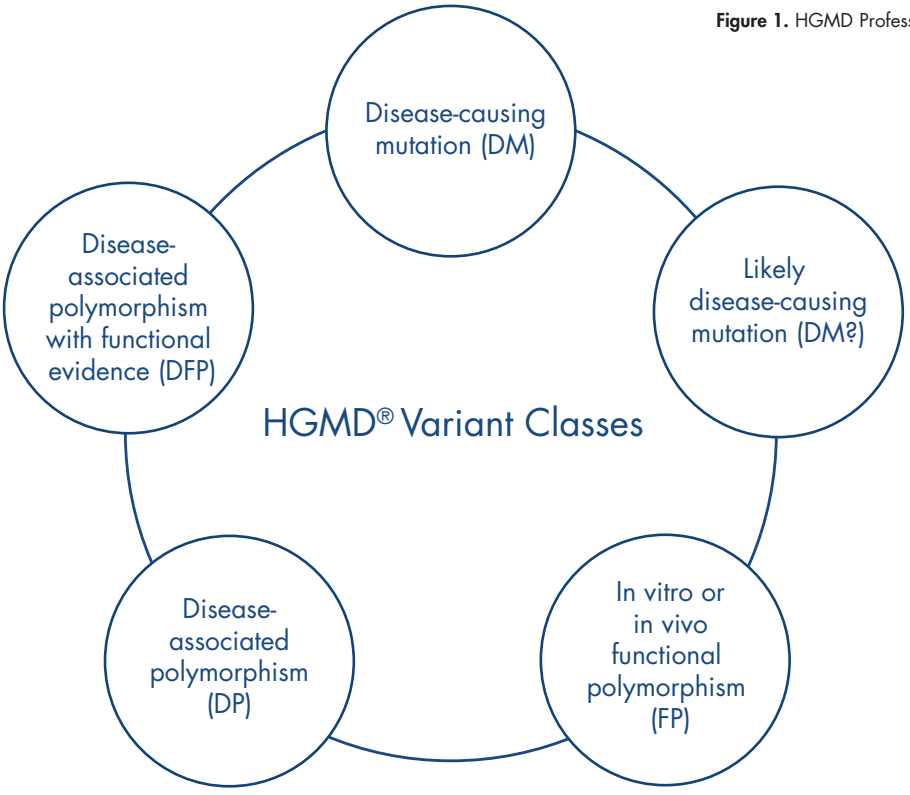


Figure 2. Types of mutation within HGMD.

"HGMD professional provides the most comprehensive database of human disease associations and is an invaluable resource in both clinical and research-grade genetics and genomics activties."

Dr. Ali Torkamani
CSO at Cypher Genomics

"We rely on HGMD professional heavily for reporting our clinical tests. We are currently working on next genertion sequencing projects, identifying genes for disease-causing mutations and disease-associated / functional polymorphisms."

We Yaping Yang, PhD
Baylor College of Medicine

A view into comprehensive coverage

HGMD is widely accepted as the gold standard for information about published inherited disease mutations, but what makes it such a great resource? In short, it's HGMD's comprehensive literature coverage. But what does comprehensive mean, and is it really a big deal?

Comprehensive means identifying every published article that describes a germline mutation and assessing whether the mutation has been convincingly demonstrated to be associated with a specific disease or phenotype. If the association is convincing, and the mutation has not been previously reported in HGMD, a new entry will be created. Likewise, if the article provides information that calls a previously reported association into doubt or provides information about a new associated phenotype for an existing mutation, the information is captured and added to HGMD.

Just picture the effort that it has taken to curate the more than 85,000 published articles describing the 197,000+ (Table 1) mutation entries cataloged in HGMD today. And then consider that the work is never ending. New papers describing new mutations are published every week. And in fact the rate of publication of new mutations is only increasing as NGS technologies have helped advance the pace of discovery. As Figure 3 shows, in less than a decade the number of new mutations described in the literature for a single year has more than doubled. That's a lot more articles to read than ever before.

What does this mean for you? It means that you can confidently use HGMD to substantially decrease the time it takes to search for and collect information about inherited disease mutations in the published scientific literature. It takes less than five minutes to search HGMD for a disease and return the list of associated mutations. Compare that to the time it would take to comb through a typical list of articles returned by a PubMed search. At a very conservative five minutes per article, the savings for a moderately well studied disease such as Bloom Syndrome could be more than 7 hours (Table 2). For a very well studied disease like Cystic Fibrosis the savings could be upwards of 280 hours (Table 2). That's the advantage of comprehensive coverage.

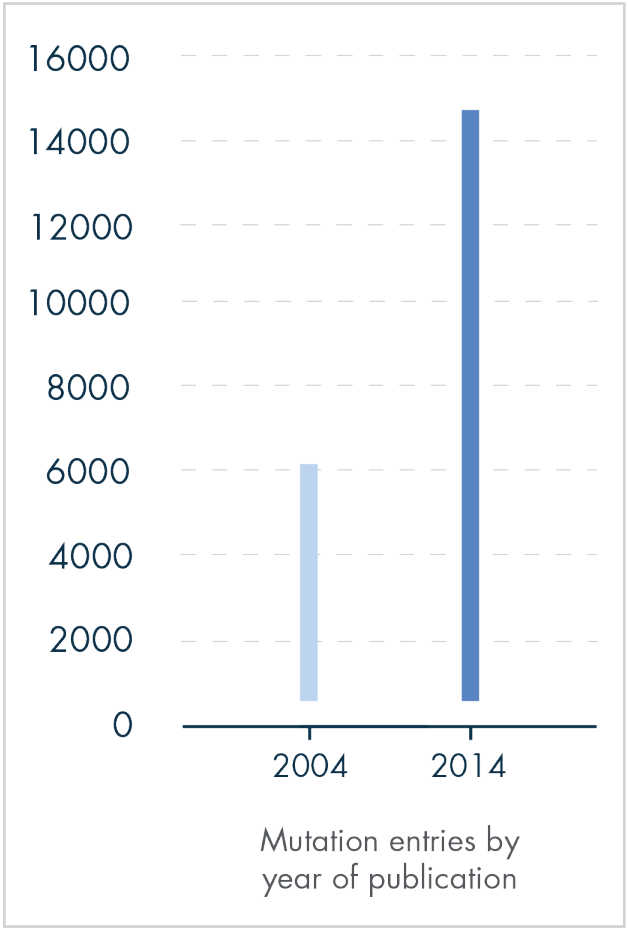


Figure 3. Number of mutations described in the literature for a given year

Mutations Type	Number of entries
Micro Lesions	
Missense/Nonsense	111,135
Splicing	17,924
Regulatory	3,725
Small Deletions	29,381
Small Insertions/Duplications	12,245
Small Indels	2,784
Gross Lesions	
Gross deletions	14,781
Gross Insertions/Duplications	3,670
Complex Rearrangements	1,802
Repeat Variations	505
Total	**197,952**

Table 1. HGMD database statistics

Disease	PubMed articles*	Time to read**
Bloom Syndrome	92	7.7 hours
Cystic Fibrosis	3,415	284 hours

* Based on Title/Abstract search of "disease" AND mutation, performed 08/01/2016

** Based on very conservative assumption of 5 minutes per article

Table 2. Number of articles available through a PubMed search and the estimated time to read all the articles for Bloom Syndrome and Cystic Fibrosis

Using HGMD Data for NGS Applications

In addition to the easy lookup access for individual mutations and genes provided by HGMD Online, HGMD data can be licensed for download as a MySQL relational database, as well as in .bed, .gff and VCF formats, enabling more advanced querying and mining of the content as well as integration into local pipelines and tools. HGMD is also available, pre-integrated, in two software platforms optimized for NGS data analysis: QIAGEN Clinical Insight (QCI™) Interpret, a clinical decision support platform for facilitating test interpretation and reporting, and Ingenuity Variant Analysis™, a platform for annotation, filtering and interpretation of causal variants. Both platforms work directly with the QIAGEN Knowledge Base which integrates HGMD mutations and their disease associations with additional content sources. The QIAGEN Knowledge Base also provides manually curated clinical case counts and contextual details such as zygosity, observed co-occurring mutations, ethnicity and functional studies for many HGMD mutations, enabling more accurate variant scoring in clinical contexts.

To learn more about QCI Interpret or Ingenuity Variant Analysis, visit **http://www.qiagenbioinformatics.com**.

QIAGEN HGMD Online Customer Statistics

93% of surveyed research organizations rely on links to the mutation source (for example, the PubMed abstract) from HGMD, which contributes to their trust in the resource. 75% of surveyed research organizations have reduced the amount of time needed for identifying published mutations by 50% or more with HGMD compared with previous methods.

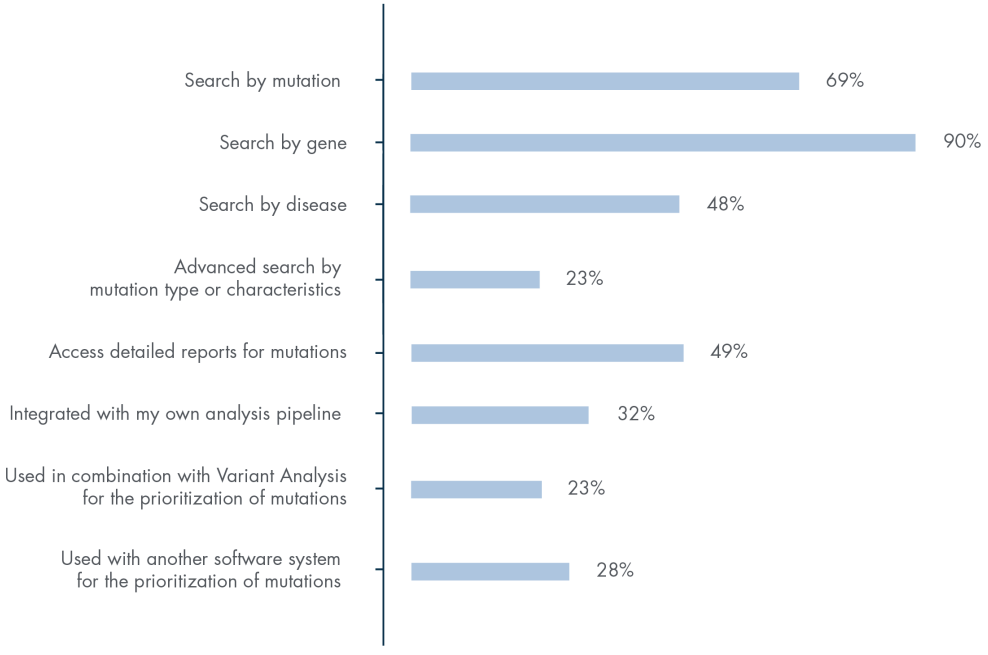
HGMD Survey Results

How is HGMD used by current customers in their work?

We conducted a survey of more than 200 users of the online and download offerings. The results provide a summary of viewpoints from customers spanning diverse backgrounds and institutions, giving insight into the many ways that HGMD can be applied and the challenges that it helps to resolve.

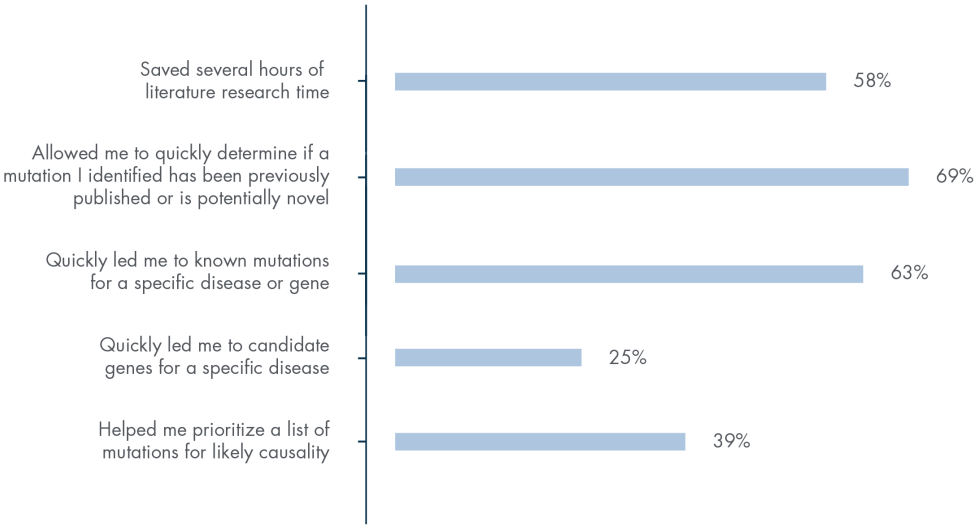
Primary Application Focus

Listed below are examples of how HGMD can be applied. Users were asked to select all that they had used in the previous 12 months.



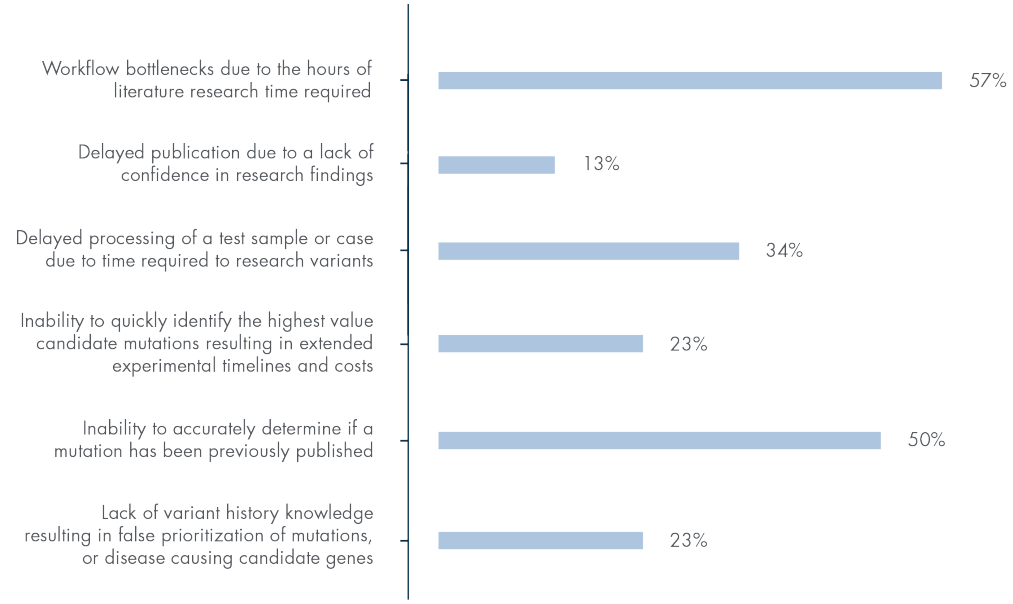
HGMD Contribution

How has the use of HGMD contributed to your work, or the work of others that you support?



Challenges Solved

What bottlenecks has HGMD solved for your laboratory or organization?



"Fast reliable searching that we can use to integrate either into our own pipeline or being able to use in conjunction with other tools for variant analysis."

Research Analyst
Federal Government

"It has provided up-to-date information about genes and mutations that help facilitate the interpretation of test results. It is an excellent tool and saves me a lot of time."

Lab Director
Health Care Company

"The product has provided a fast, convenient way to prioritize previously described variants in relation to an exome or genome's worth of variant data."

Research Analys
Medical College

To learn more from a sales or support solution specialist, contact us using the information below:

QIAGEN Bioinformatics

EMEA
Silkeborgvej 2 · Prismet
8000 Aarhus C
Denmark
Phone: +45 8082 0167
E-mail: bioinformaticssales@qiagen.com

Americas
1700 Seaport Boulevard #3
Redwood City · CA 94063
USA
Phone Toll Free: +1 866 464 3684
E-mail: bioinformaticssales@qiagen.com

qiagenbioinformatics.com