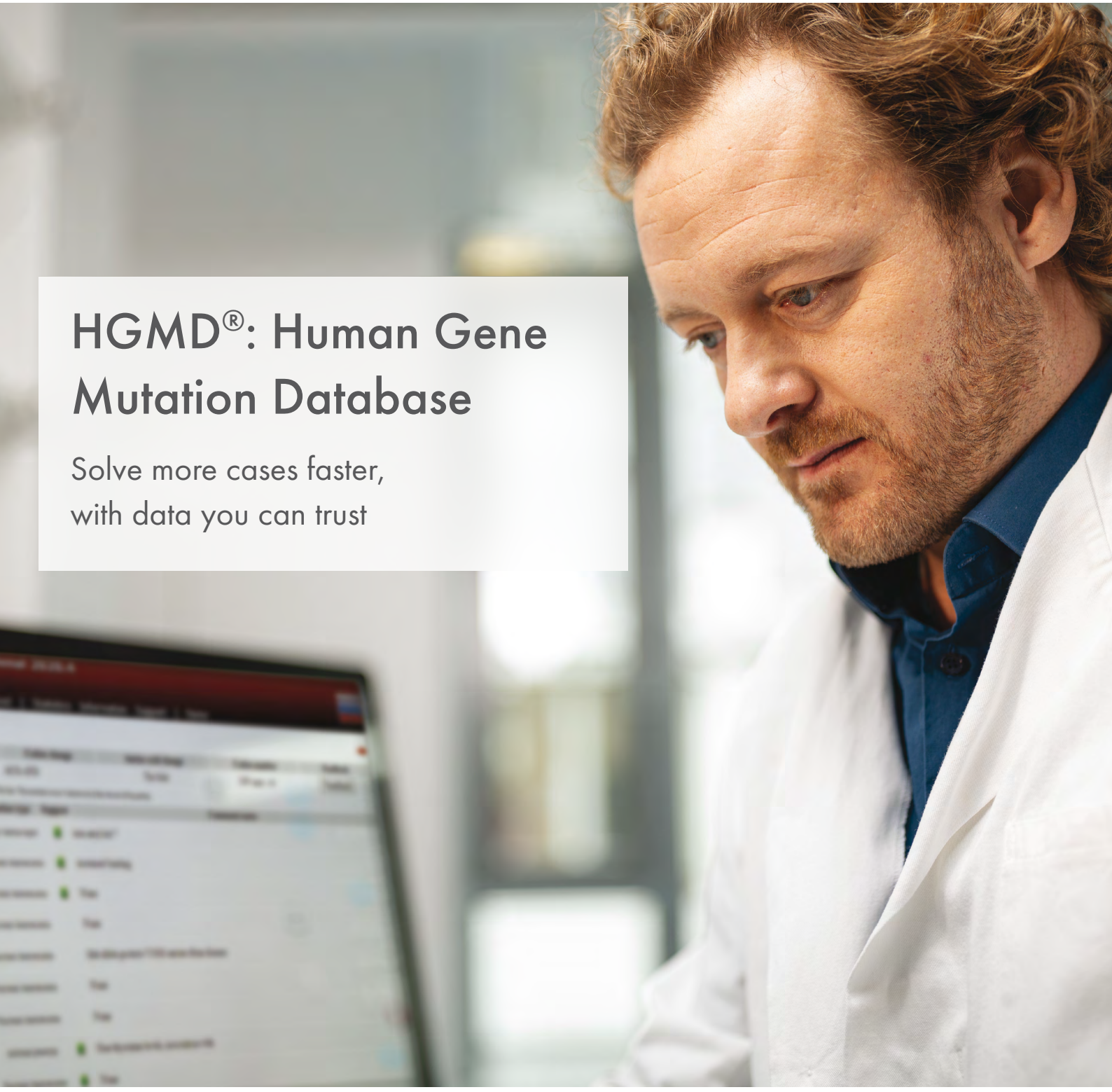
QIAGEN



HGMD®: Human Gene Mutation Database

Solve more cases faster,
with data you can trust

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 15,000+ 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

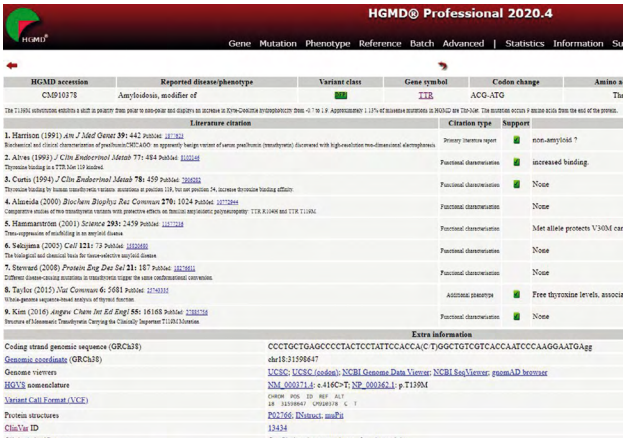


Figure 1. HGMD Professional sample mutation report.

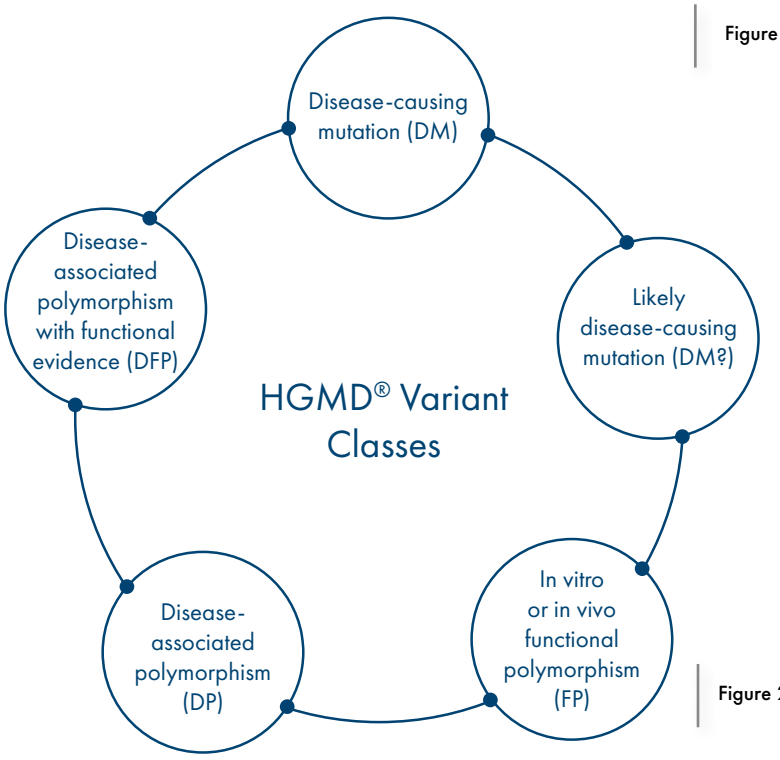


Figure 2. Types of mutation within HGMD.

HGMD: Quality versus quantity

Source of variant data

Unlike other competitors who offer little to no data curation or overload users with unhelpful literature and volumes of conflicting data, HGMD combines electronic and human search procedures during data curation in order to provide high-quality information. For more than 30 years, a team of expert curators has consistently screened peer-reviewed biomedical literature in over 250 journals to update HGMD.

Discrepancy in variant reporting

The existence of a significant discrepancy in variant reporting requires additional scrutiny. Some of the discrepancies could only be resolved manually by scanning conflicting journals, by assessing and reviewing supplementary materials in the articles, or by direct contact with the authors. None of these could be done automatically and require manual input by the experienced scientific experts that HGMD offers.

Utilization

The high quality HGMD data can be used in both research and clinical settings. The public version is available for free but is kept 3 years behind, while professional version users can access the most up-to-date database with additional literature reports, chromosomal coordinates, population frequency data, and functional prediction.

Variant reclassification

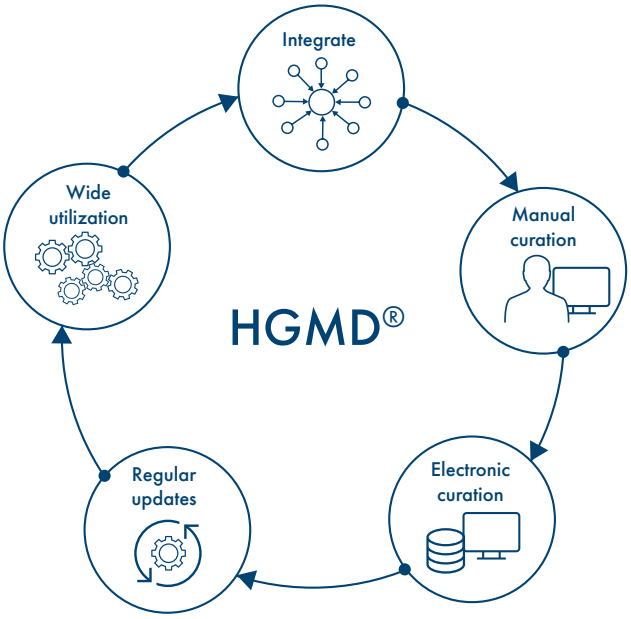
The HGMD curators adhere to a policy of continual content curation, commenting on and annotating new information. When new evidence suggests a benign nature of a variant, it may be removed from the database at the discretion of experienced curators. Variant reclassification continually takes place in the HGMD system providing high-quality data to users and ensuring they don't waste time on benign variants or polymorphisms... HGMD is the only database that continuously curates and reclassifies wherever necessary — not relying solely on the original submitter updating their submission.

Variants of Unknown Significance

Many of the variants reported in the literature as disease-causing may, in fact, be variants of unknown significance and may not cause disease. Unlike other sources that include practically all submitted VUS without further inspection, HGMD reclassifies VUS when there's enough manually curated evidence. The decision on whether to include a variant is not arbitrary either. It involves an exercise of expert curation, and these variants are also specifically marked to indicate that some degree of uncertainty exists.

The importance of quality

HGMD Pro provides scientists and clinicians with the most up to date information about potential disease-causing variants. The manual curation process ensures that little time is wasted going through polymorphisms, unrelated literature, or unverified information.



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 127,000+ published articles describing the 307,000+ (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.

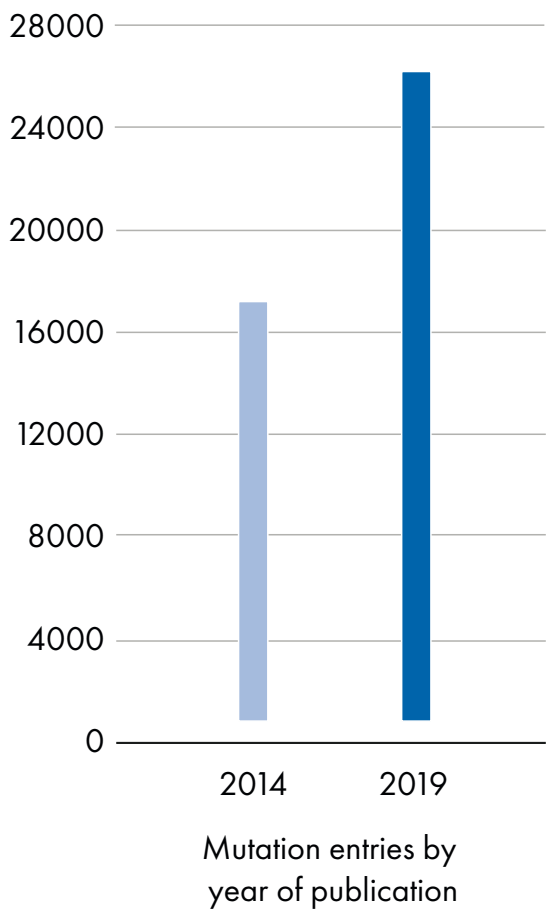


Figure 3. Number of mutations described in the literature for a given year

Mutations Type	Number of entries
Micro Lesions	
Missense/Nonsense	179,772
Splicing	26,610
Regulatory	4,909
Small Deletions	44,074
Small Insertions/Duplications	18,530
Small Indels	3,975
Gross Lesions	
Gross deletions	21,166
Gross Insertions/Duplications	5,391
Complex Rearrangements	2,341
Repeat Variations	598
Total	307,366

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 QCI Interpret Translational, 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 QCI Interpret Translational, visit **https://digitalinsights.qiagen.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 to identify published mutations by 50% or more with HGMD compared with previous methods.

"Facilitates the clinical Interpretation of variants. Very useful for diagnostics. Provides variant specific links to several other databases. Links to disease specific databases would further facilitate the clinical Interpretation of variants.."

Associate Professor in Austria Educational Institution

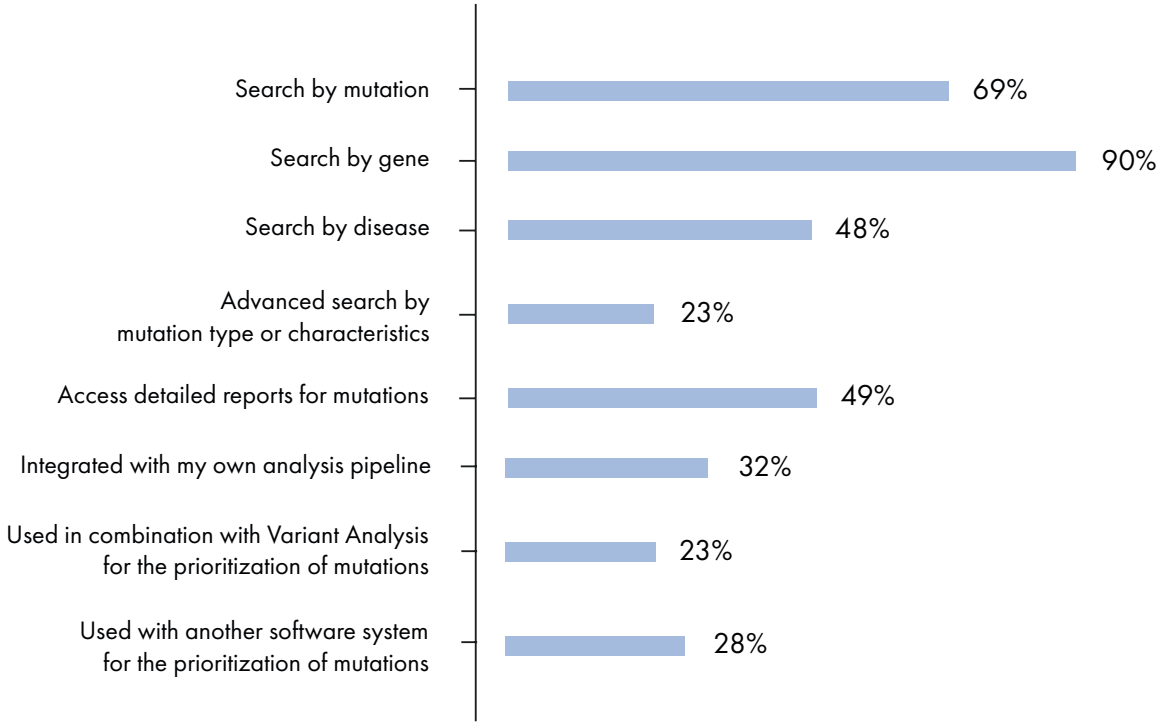
HGMD Survey Results

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.

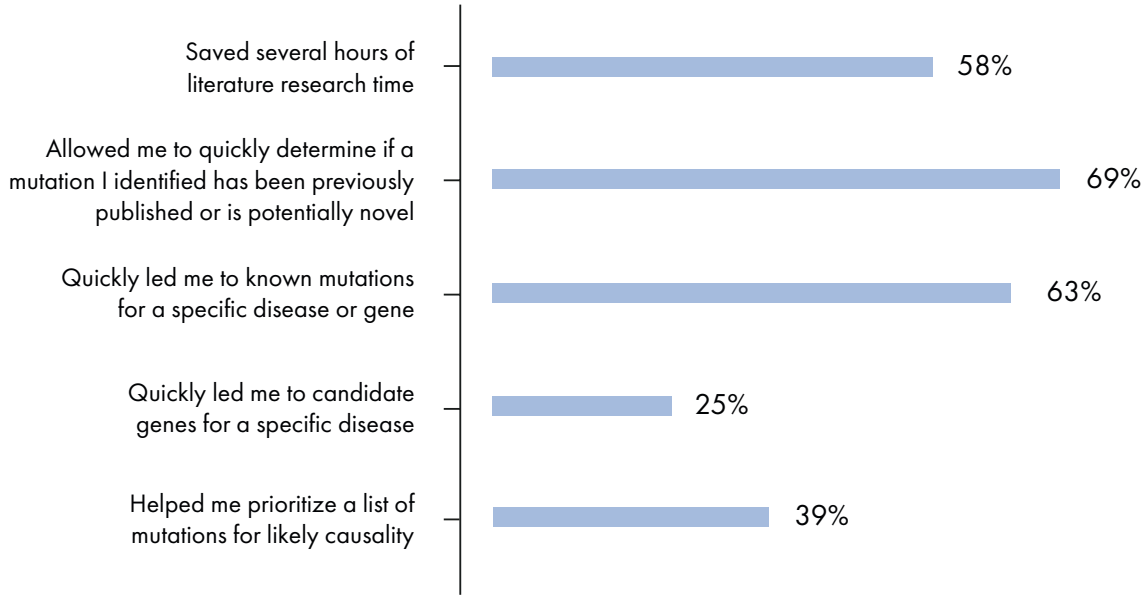
Primary Application Focus

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.



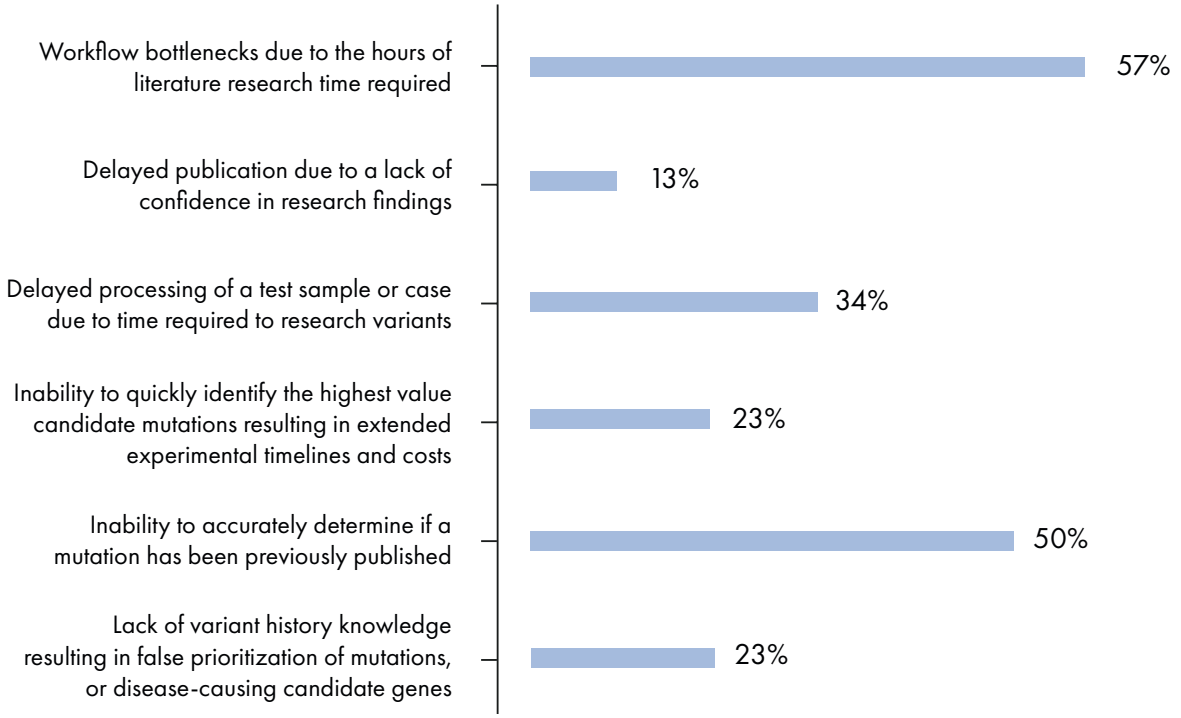
HGMD Contribution

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 Analyst, Medical College

Experience HGMD Pro with a free trial. Click **here**

Ordering www.qiagen.com/shop | Technical Support ts-bioinformatics@qiagen.com | Website digitalinsights.qiagen.com