

# The GDB™ Human Genome Database Anno 1997

Kenneth H. Fasman\*, Stanley I. Letovsky, Peter Li, Robert W. Cottingham and David T. Kingsbury

Division of Biomedical Information Sciences, Johns Hopkins University School of Medicine, Baltimore, MD 21205-2236, USA

Received October 1, 1996; Accepted October 8, 1996

## ABSTRACT

**The value of the Genome Database (GDB) for the human genome research community has been greatly increased since the release of version 6.0 last year. Thanks to the introduction of significant technical improvements, GDB has seen dramatic growth in the type and volume of information stored in the database. This article summarizes the types of data that are now available in the Genome Database, demonstrates how the database is interconnected with other biomedical resources on the World Wide Web, discusses how researchers can contribute new or updated information to the database, and describes our current efforts as well as planned improvements for the future.**

## INTRODUCTION

The human Genome Database (refs 1-5; <http://gdbwww.gdb.org/>), an international collaboration hosted at the Johns Hopkins University, was established to address the information management needs of the Human Genome Project (HGP). GDB was created specifically to collect information from the mapping activities of the HGP. However, that mission is evolving along with the genome project itself. GDB's historical role of assembling data associated with low resolution cytogenetic maps has shifted to high resolution physical maps and is now moving toward the ultimate high resolution map, the human genome sequence.

The Genome Database provides a resource for the international biomedical research community that integrates all of the current scientific information on human genomics. GDB is becoming an encyclopedia of genome structure, content, diversity and evolution. GDB does not and cannot collect and manage all of this information in a single database, but it can provide the focal point for accessing data resources worldwide. Our goal is to provide a rich online model of the current scientific understanding of the human genome. To that end, the project's staff fosters collaborations with medical, biological and informatics groups to refine and improve the technologies and data.

## GDB ENHANCEMENTS

In the previous report in this series (5), we discussed the newly released GDB version 6.0. As detailed in the earlier article, GDB 6 introduced a number of significant new features: (i) direct

community data submission and curation, including ownership of GDB data by individuals or labs/collaborations and third-party annotation and enhancement of data submissions; (ii) improved representation of genes and related biological data; (iii) improved map representation and querying, including graphical map display; (iv) enhanced World Wide Web interface; and (v) an object-oriented data model using the Object-Protocol Model (6).

One of the fundamental technical principles of the GDB 6 architecture is its ability to easily and rapidly accommodate enhancements to the database schema without significant software modifications. Minor improvements to the data model and WWW interfaces are made several times a week. Significant improvements to the schema have been introduced every 4-8 weeks during the past year. This flexibility will allow the Genome Database to track much more swiftly advances in human genomics.

## CLASSES OF DATA IN GDB

The data model has been enhanced significantly during the past year, and Figure 1 summarizes the classes of information that can now be represented in the Genome Database. Note in particular that the types of genomic segments (mappable regions of the genome) and maps are being regularly expanded. New segment types have been added to support the integration of mapping and sequencing data (e.g., gene elements, repeats) and the construction of comparative maps (syntenic regions). New map types include comparative maps for representing conserved syntenies between species, and sequence feature maps for integrating high-resolution physical maps with sequencing data. Experimental observations of order, size/distance, chimerism, etc. established with various mapping reagents are available to enhance the process of map integration.

The representation of common polymorphisms and their molecular characterization has remained unchanged in the Genome Database since version 5.0. We are working with the members of the mutation database community ([http://ariel.ucs.unimelb.edu.au:80/~cotton/mut\\_database.html](http://ariel.ucs.unimelb.edu.au:80/~cotton/mut_database.html)) to enhance GDB's ability to serve as a central distribution venue for the information that is currently being curated in the many single-locus mutation databases (for examples, see 7-10).

Through the medium of the World Wide Web, information in the Genome Database is linked to myriad other biological data resources on the Internet. The 'external link' category in Figure 1 lists a growing collection of cross-references that can be established between genes and other entities in GDB and related

\* To whom correspondence should be addressed. Tel: +1 617 252 1919; Fax: +1 617 252 1902; Email: ken@gdb.org

<b>Genomic segment</b>	<b>Map</b>	<b>Biological objects</b>
Breakpoint	Comparative	Expression
Chromosome	Content contig	Gene family
Contig	Cytogenetic	Gene product
CpG island	Integrated	(protein/RNA)
Cytogenetic marker	Linkage	Organism
Fragile site	Radiation hybrid	
Gene	Sequence feature	<b>External link</b>
Gene element		Citation
Polymorphic marker	<b>Experimental data</b>	Enzyme
Regulatory region	Chimerism	Generic
Repeat	Detection method	Homology
Syndromic region	Distance	Nucleic acid seq.
Syntenic region	Order	Phenotype
	PCR condition	Protein sequence
<b>Reagent</b>	<b>Variation data</b>	Structure
ASO	Mutation	Taxonomy
Amplimer	Polymorphism	<b>Annotation</b>
Cell line	Population freq.	Misc. annotation
Clone		Object name (alias)
Library		Reagent source
Mapping panel		

Figure 1. Classes of data in the Genome Database.

```

#
# EDS Declaration (login information, etc.)
#
( declare {
( .user          "<login>" )          # user login name
} )
#
# EDS Insert
#
( insert {
( .db            hgd )
( .class        Clone )
( .declare      Clone1 ) # your ID for future reference

( displayName   "MyClone" ) # string value
( DNAType       "cDNA" )    # code value of DNATypeDict
( insertSize    85.5 )     # kilobases
( libraryAddresses [
( library       GDB:12345678 ) # key value to Library
( plateLoc     "345" ) # string value
( rowLoc       "A" ) # string value
( columnLoc    "1" ) # string value
] )
( multiCopy     "No" ) # code value of YesNoUnknown_UnkDict
( vectorType    "BAC" ) # code value of VectorTypeDict
} )
} )

```

Figure 2. Sample electronic data submission (EDS) file for inserting one clone into the database.

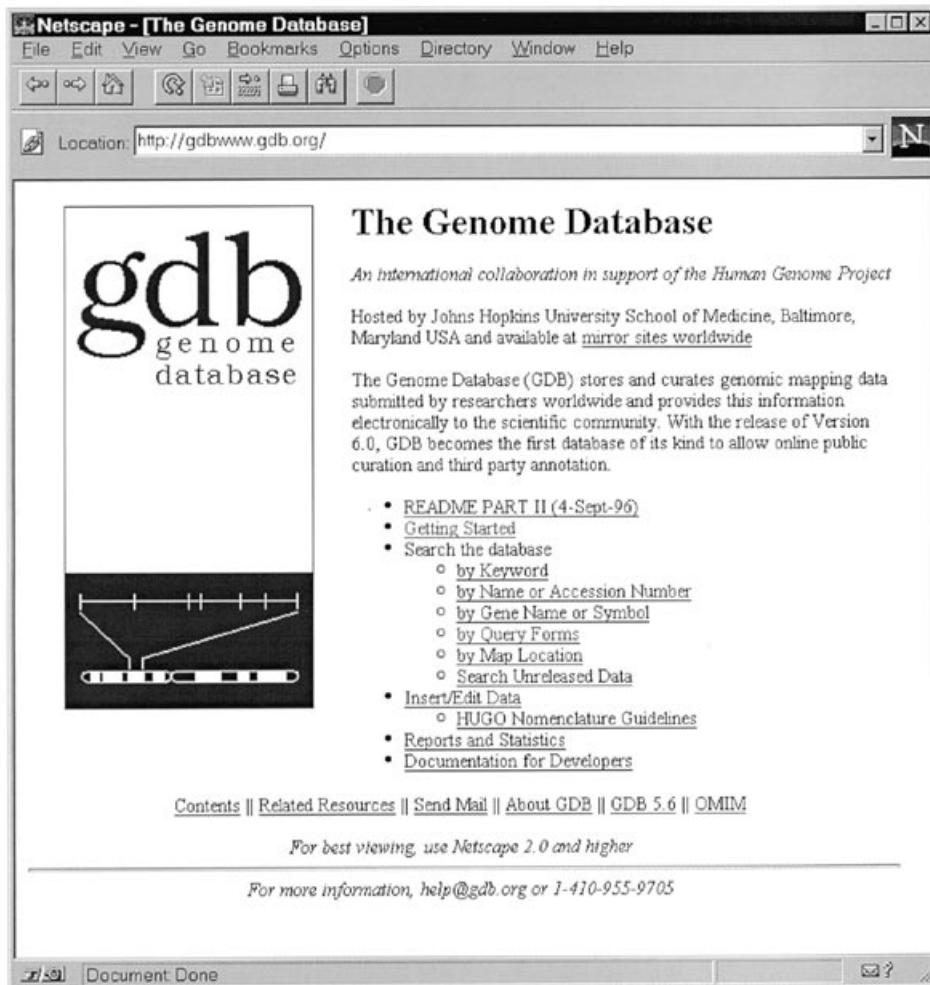
information in other databases. By providing a repository for such cross-references, the Genome Database can serve as a central point of inquiry for a growing number of research areas involving human genomics. In doing so, GDB does not duplicate the activities of the primary curators of such information, but enhances the accessibility of their efforts.

## CONTRIBUTING DATA TO GDB

There are currently two methods of contributing data to the Genome Database: electronic data submission (EDS) and interactive editing. GDB's system for bulk electronic submission and updating is designed primarily for genome centers and other large laboratories that are generating significant quantities of human genomic data and have some local informatics support. Figure 2 illustrates a sample GDB EDS file. The submission file contains a 'Declare' section for specifying the user's GDB account information, and

one or more 'Command' sections for inserting or updating data objects. The EDS syntax provides a mechanism for mapping the user's object names or symbols to objects in GDB, which is crucial for establishing the correct links to information already in the database. Detailed specifications on the Genome Database's EDS syntax, file templates, file submission, etc. can be found at <http://gdbwww.gdb.org/eds>.

Interactive editing of the database via the World Wide Web is available for submitting or updating small amounts of information. This system has been simplified since it was first introduced, and is still an area of active development for GDB. Two sources of complexity in this process are the unsuitability of HTML forms for interactive editing of a database and the unreliability of long-distance Internet communication outside of North America. The first issue has been addressed through enhancements to the editing interface which use the Javascript



**Figure 3.** GDB's home page on the World Wide Web.

programming language (and will likely include Java-based extensions in the future).

Improving the editing capabilities of human genome researchers outside of North America is more difficult. Currently, these scientists are encouraged to prepare electronic submissions or to send their data to Baltimore in any convenient form (preferably electronic) for us to enter there. We are hoping to develop software in the near future that will guide users through the task of creating EDS files on their Macintosh or PC computers, providing a form of 'off-line' editing.

It should be noted that both types of direct data submission require that the researcher have a user account with GDB in Baltimore. A Web form for obtaining an account can be found at <http://gdbwww.gdb.org/gdb-bin/gdb/regmail>. The Genome Database reserves the right to deny editing access to anyone who cannot demonstrate that they are a legitimate member of the genome research community, in order to minimize the potential for abuse (note that query access to the database is available to all). Researchers who wish to contribute information to GDB but who do not desire an editing account are encouraged to contact the GDB staff at [data@gdb.org](mailto:data@gdb.org) to arrange for the data to be entered in Baltimore.

## A TOUR THROUGH THE NEW GDB

As an introduction to the Genome Database's dramatically different interface, we will take a brief tour through the Web site, focusing on a single, detailed gene entry to show the wide variety of information that is now available from GDB. New users of the database, and experienced users who are coming in with novel inquiries, will usually start at the GDB home page (Fig. 3). The home page provides links to a variety of query interfaces, the interactive editing interface, important announcements, introductory material, assorted reports and GDB statistics, and documentation for programmers developing data submissions and other software that must interact with the database. Additional links lead to a complete catalogue of the Web site, an extensive list of other genomic information resources on the World Wide Web, and administrative and contact information for GDB.

There are a variety of ways to formulate database queries, including:

(i) search by keyword—users can search all or selected parts of GDB with words or phrases, including the use of simple Boolean expressions;

**Gene HPRT1**

Add Annotation | Add Alias | Link Citation | Add Generic Link | View History | Add Nucleic Acid Sequence Link | View Maps Containing this Segment | Add Homology Link | Add Phenotype Link | Add Protein Sequence Link

**Symbol:** HPRT1  
**Accession ID:** GDB:119317  
**Owner:** HUGO NC  
**Status:** Active

**Aliases:**

Name	Authority
HPRT	GDB EDIT GROUP
HPRT1	HUGO NC
hypoxanthine phosphoribosyltransferase 1 (Lesch-Nyhan syndrome)	GDB EDIT GROUP

**Genome:** Human

**Cytogenetic Localization:**

Chromosome	Left Marker	Right Marker	Map Element ID	Map
X	Xq26.1	Xq26.1	GDB:1088736	Cytogenetic Map HUGO - X - Cytogenetic Map
X	Xq26.1	Xq26.1	GDB:1197437	Cytogenetic Map Public Chromosome X Cytogenetic Map

**Other Localization:**

Chromosome	Map Element ID	Map	Coordinate	Units
X	GDB:3922317	Linkage Map EUROGEN Chromosome X comprehensive map (female)	165.800	Kosambi centiMorgans
X	GDB:3929516	Linkage Map CHLC Chromosome 23 Recombination Minimization (Female)	178.700	Kosambi centiMorgans

**Order Data:**

Relationship	Marker	Observation	Position	Order ID
Overlaps	Clone HPRT-600	Hybridizes with		GDB:768364
Overlaps	Clone HPRT-800	Hybridizes with		GDB:768365
Overlaps	Clone Hwlambda15	Hybridizes with exons 6-9, with intervening introns		GDB:772459

Figure 4. Entry for the HPRT1 gene in GDB: part 1.

(ii) search by name or accession number—for those who already know the name or database identifier of the object(s) of interest, this provides the fastest access to the data;

(iii) search by gene name or symbol—a specialized interface to simplify the process of finding gene entries and related information;

(iv) search by query forms—this is currently the most sophisticated Web-based interface to GDB. Queries are formulated by specifying values or ranges for one or more attributes of particular categories of data;

(v) search by map location—a specialized interface for finding genomic segments in a specified region of the genome.

In addition, direct programmatic access to the database is available, using both the standard SQL query language and that provided by the Object Protocol Model.

### Learning about HPRT1

Suppose that using one of these search methods, you retrieve the GDB record for the HPRT1 gene. A lengthy, scrollable document will be displayed in your Web browser, extensively hyperlinked both to other objects in GDB and to information elsewhere on the World Wide Web. The beginning of this document is shown in Figure 4.

Several features at the top of this figure are common to many GDB displays. For example, the miniature 'gdb home' logo is always a link back to the GDB home page. In the heading 'Gene

HPRT1', the category is hyperlinked to the database schema documentation describing the Gene class. Many GDB Web pages also include a text 'menu bar,' listing a variety of actions that can be taken in the current context. For example, clicking on 'Add Annotation' will take you to a form for adding a comment to the database that will be automatically linked to the HPRT1 record. This could be used, for example, to report the discovery of an error in the database.

Most objects in the Genome Database have several identifiers associated with them. These are displayed near the middle of Figure 4. Every object in the database has an official name or symbol (usually chosen by the object's owner), and an 'accession ID' assigned by GDB. That way, even if the owner decides to rename the object, cross-references to it from other databases and the literature can be maintained based on the accession ID, which is never changed or reused. Objects in GDB can also have other aliases associated with them. These record alternative names used by the owner or by other members of the genome community. Note that each alias is accompanied by the source of the alternate name. A 'Genome' attribute for genes allows GDB to store information from other organisms, particularly for purposes of representing homology data and comparative maps.

The lower part of Figure 4 shows a variety of mapping information associated with the HPRT1 locus. The HUGO consensus cytogenetic location for the gene is shown, along with

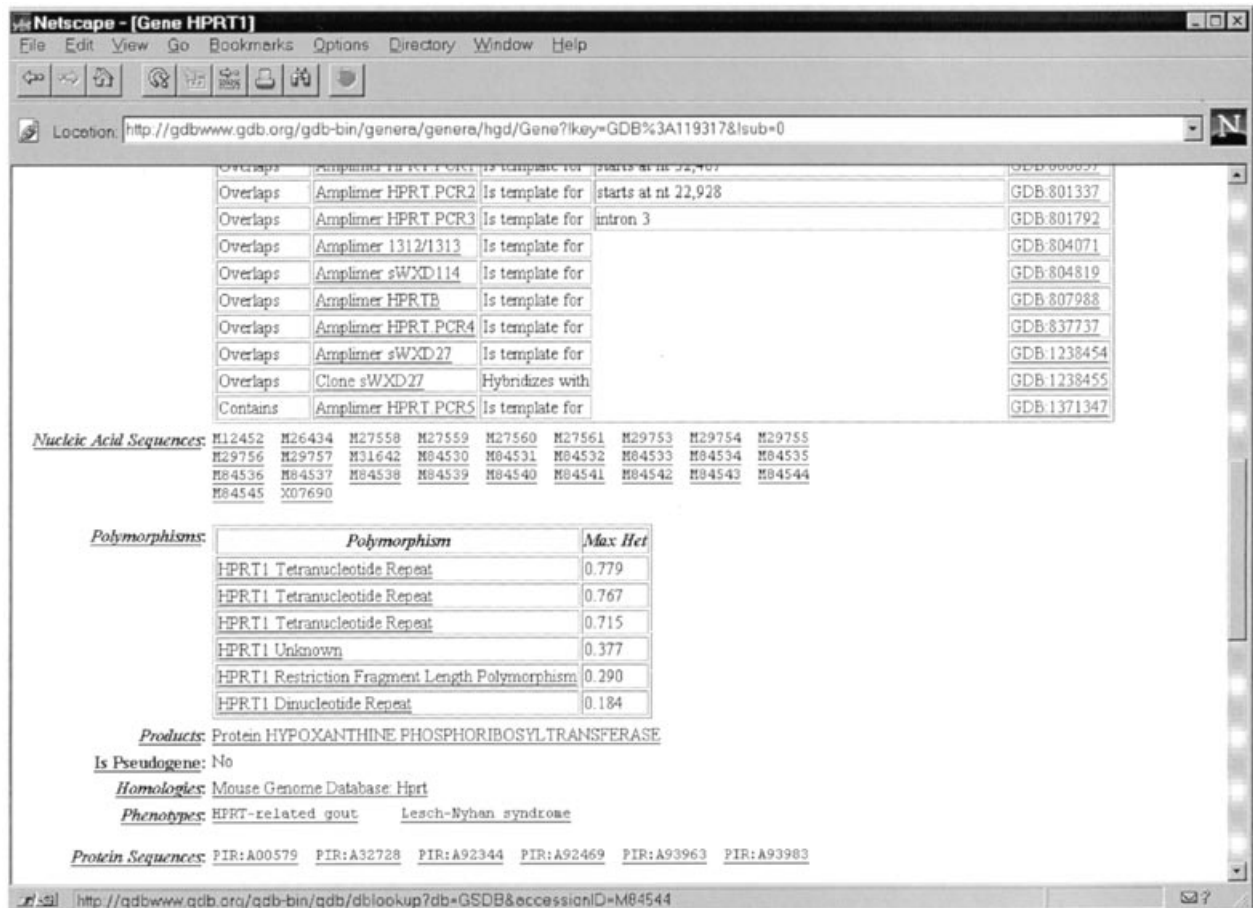


Figure 5. Entry for the HPRT1 gene in GDB: part 2.

any other cytogenetic localizations that may have been provided by the community. In addition, the figure shows that HPRT1 appears in one other map in GDB, a EUROGENE project linkage map of the X chromosome. At the bottom of this figure are a list of experimental observations indicating a variety of mapping reagents that have been shown to interact with this locus (this information is continued at the top of Fig. 5). These observations can be used to establish order and/or distance relationships in maps.

The remainder of Figure 5 shows a variety of molecular information for HPRT1, including known polymorphisms, and links to external databases describing the gene's nucleotide sequence, protein product and corresponding protein sequence data. Additional links to functional information are available, including associated homology data and phenotype descriptions.

Figure 6 shows the last part of the HPRT1 gene entry. It illustrates some of the variety of third-party annotations that can be added to objects in the Genome Database. These include editorial commentary (usually provided by GDB staff or HUGO editors), links to literature references and links to speciality databases on the World Wide Web (in this case, an HPRT mutation database). The latter is of particular importance; curators of special genome information resources are strongly encouraged to add 'External Links' to related items in the Genome Database, so that users of GDB can easily find their data collection. Also note that a variety of administrative information about a GDB entry can be

found, either at the end of the document or by selecting 'View History' from the menu bar at the top of the page.

Also available from the menu bar on all types of genomic segments in GDB is an item labeled 'View Maps Containing this Segment.' Selecting this link will allow you to retrieve one or more maps from the database which contain the marker in question. Figure 7 illustrates the EUROGENE linkage map of X mentioned above, displayed in the original GDB Mapview program. This program, which is available as a helper application for Web browsers on Macs, PCs and Sun workstations, provides an interactive graphical display of GDB maps. Objects in the map can be clicked on, and the Web browser will then retrieve the details of the marker from the database.

## WORK IN PROGRESS

### Map alignment viewer

As this article is being written, a much-improved version of GDB's map viewer is in final preparation, to be released in the late fall of 1996. The new map viewer is being developed using the Java programming language, which will make it available in many more computing environments than the original Mapview program. More importantly, the new viewer can display multiple aligned maps, rather than a single map at a time. Figure 8 shows a display

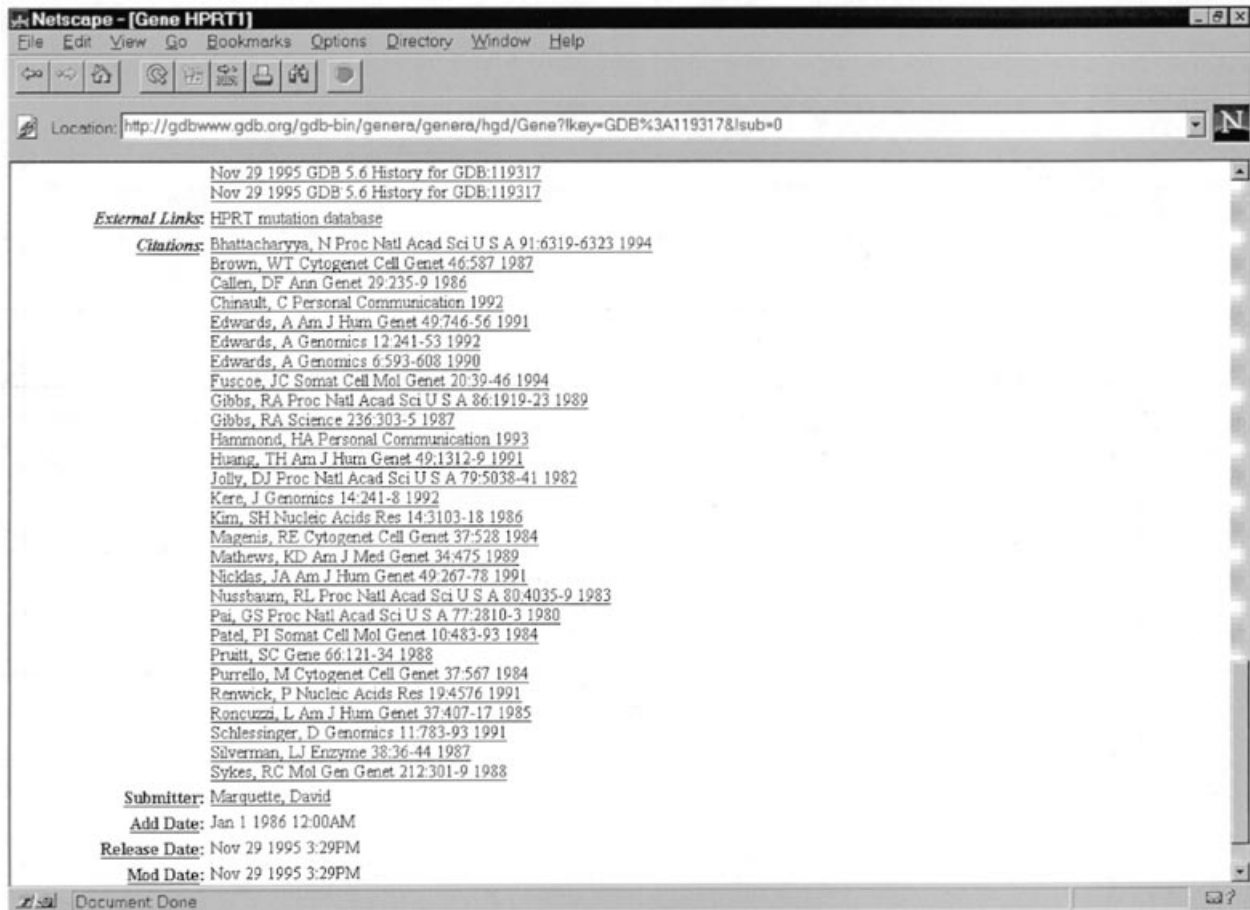


Figure 6. Entry for the HPRT1 gene in GDB: part 3.

generated by this program. Several maps are displayed side-by-side in the window, with alignment lines indicating common markers in neighboring maps. As with the single map viewer, the user can select individual markers to retrieve more information about them from the database.

The Genome Database staff are working very closely with the other members of the bioWidget consortium (including, but not limited to, the University of Pennsylvania, Lawrence Berkeley National Laboratory, and the Jackson Laboratory; see <http://info.gdb.org/biowidgets>) so that the Java-based map alignment viewer will become part of a growing collection of freely available software tools for displaying and manipulating biological data.

### Improved positional querying

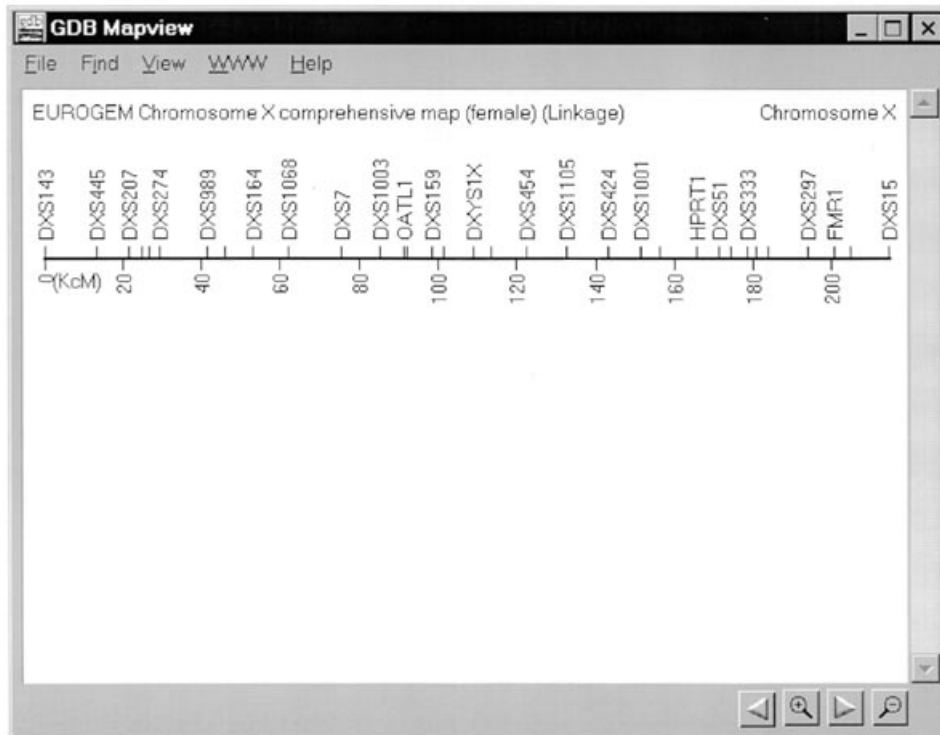
Over the past year GDB has put in place a set of flexible, schema-driven software tools for editing and accessing data via the Web. 'Schema-driven' means that, for the most part, the knowledge of how the database is organized is not built into the software; rather it is stored in a schema file which the programs consult as they need. The result is that it is now much easier to extend the schema into different subject areas as the need arises. Our efforts over the next few years will be focused on extending GDB into those subject areas that will provide the greatest enhancement to the utility of the database.

In recent months we have been developing a new capability, soon to be released, which might be termed 'Integrated Positional and Relational Querying with Optional Graphic Display,' or just 'positional querying,' for short. Positional querying will allow users to frame questions of the form:

find all loci in region-of-interest R,  
satisfying additional requirements Q

The region of interest R can be specified by a variety of methods, including flanking markers, map coordinates, and so on. By default, such queries will search all maps of the specified region in GDB. The results of such a query will be a set of loci which will be viewed either in tabular format or via our new Java multiple map drawing program. Queries of this type are important to the hunt for disease genes, among other applications. Suppose for example that you had genetic evidence of a schizophrenia locus in a region of 6p, and a theory suggesting the involvement of a neurotransmitter receptor. A query one might want to pose in searching for candidate genes is: find things in this region which look like they might be neurotransmitter receptors. The schema revisions we have put in place over the last year have given us a handle on the 'find things in this region' (R) part of the query. The extensions over the next few years will expand the sorts of things we can say in the Q part of the query, such as:

having homology with a given sequence  
having a specified level of polymorphism



**Figure 7.** The original GDB map viewer, displaying a single linkage map containing HPRT1.

expressed in fetal liver  
having transmembrane domains

Note that in order to search all maps, we will need to prealign all maps of a given region, using functions that transform the coordinates of one map to those of another. We are experimenting with different types of functions, including simple linear functions produced by regression of common marker coordinates, to more complex piecewise-linear and nonlinear mappings. The optimal transformation will minimize the scatter in the positions of all markers when they are mapped into the common coordinate system.

#### Other data model enhancements

The reorganization of the Genome Database that began with the release of version 6.0 in January 1996 will continue well into the foreseeable future. GDB's evolution follows that of the Human Genome Project itself, both in terms of tracking the transition from mapping to sequencing, and the more long-term trend of dissemination of the fruits of human genomics throughout all of biomedical research.

In the near term, sequence data will come to play a larger role in the Genome Database. While there are a number of public nucleotide sequence databases in existence (DDBJ, EMBL, GenBank and GSDB), GDB must insure that the human DNA sequence they contain is well integrated with mapping data and other information about genomic regions of interest. The GDB staff are working with various groups undertaking large-scale sequence analysis, so that putative genes and putative function assignments can be placed in the context of genomic mapping data and put to use in the hunt for disease genes. We also plan to provide a central point for assessing the progress of human

genomic sequencing, BAC-end sequencing, EST sequencing and other large scale efforts.

Other areas of the GDB data model that will be expanded in the near future, in collaboration with specialists in each subject area, are variation data (mutations and polymorphisms), phenotypes, function and homology and comparative mapping information.

#### Electronic data submission

Improvements to the Genome Database electronic data submission system are focusing on tools and file formats for the joint submission of mapping and sequencing data. The ultimate goal of this effort will be the creation of a universal data submission format that will enable genome centers and other large laboratories to prepare a single data file to be sent to GDB and their sequence database(s) of choice. Not only will this significantly ease the burden on data contributors (who must now prepare two or more submissions), but it will also insure that the mapping and sequencing information are integrated from the outset, by the group best qualified to establish the connections.

#### COMMUNITY OUTREACH

The Genome Database receives direct community feedback through a variety of mechanisms. GDB's activities are overseen by an International Scientific Advisory Committee that meets annually, and a smaller Review Committee that confers quarterly with the database staff to help track rapidly evolving areas of human genomics and informatics technology.

GDB staff also interact frequently with the Human Genome Organisation's chromosome, nomenclature and other editorial committees. The HUGO editors have special curatorial privileges

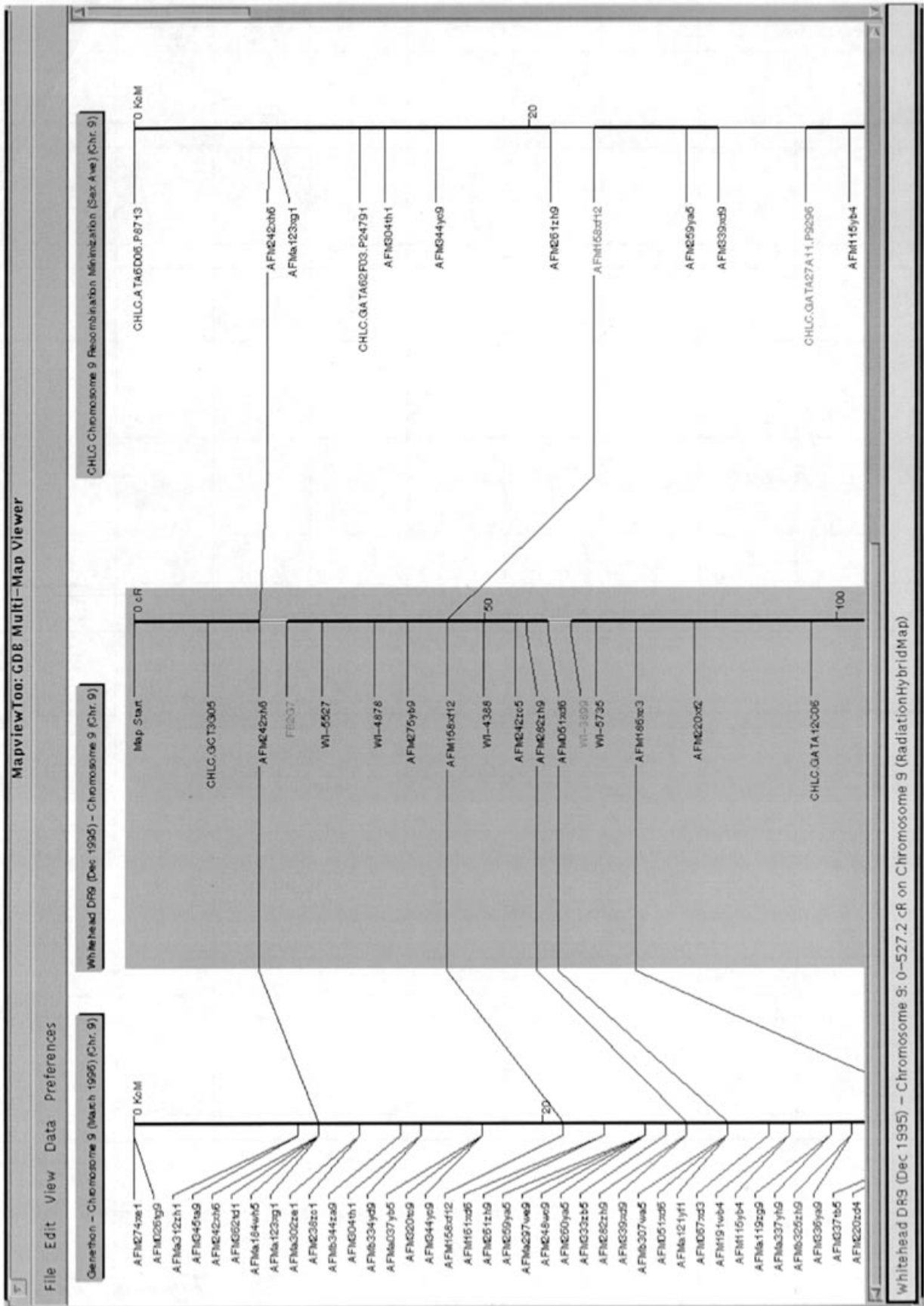


Figure 8. The new GDB map alignment viewer.



in the database, allowing them to establish official nomenclature for genes and other significant genomic landmarks, provide consensus maps of each chromosome, and perform additional activities to insure the high quality of GDB's content.

Copies of the Genome Database are maintained at 10 mirror sites around the world. These GDB Nodes help to make the database more easily accessible to international researchers, and provide documentation, training and technical support for local research communities in their native languages. GDB staff meets with the Node managers on an annual basis to facilitate our interaction and benefit from their insights and those of their regional users.

These activities are complemented by GDB staff's regular attendance at single chromosome workshops worldwide, as well as annual conferences such as the Cold Spring Harbor Laboratory's Genome Mapping and Sequencing meeting, TIGR's Genome Sequencing and Analysis conference and the meetings of the American and European Societies for Human Genetics.

## SUMMARY OF GDB SERVICES

The Genome Database provides the following services:

**World Wide Web** <http://gdbwww.gdb.org/> Complete searching and editing of human genomic data, documentation

**Anonymous FTP** <ftp://ftp.gdb.org/> Data files, documentation, standard reports, software

## CONTACTING GDB

### Baltimore

Many questions about GDB can be answered by documents on our Web server (<http://gdbwww.gdb.org>) or those of GDB's international sites (see below). Additional inquiries can be directed to:

GDB Users Services  
Johns Hopkins University School of Medicine  
2024 E. Monument Street, Suite 1-200  
Baltimore, MD 21205-2236, USA  
Tel: +1 410 955 9705  
Fax: +1 410 614 0434  
E-mail: [help@gdb.org](mailto:help@gdb.org)

(Similar services are provided by each of the international sites. Please check their Web servers for local contact information.)

For information regarding the submission of data to GDB, please address inquiries to Data Acquisition and Curation at the above mailing address, telephone and fax numbers or preferably via e-mail to [data@gdb.org](mailto:data@gdb.org).

### GDB international sites

The Genome Database provides access to the database at the following 10 international nodes:

Australia: ANGIS, University of Sydney;  
<http://mogan.angis.su.oz.au/gdb/gdbtop.html>  
WEHI, Melbourne;  
<http://wehih.wehi.edu.au/gdb/gdbtop.html>  
France: INSERM, INFOBIOGEN, Villejuif;  
<http://gdb.infobiogen.fr/>  
Germany: DKFZ, Heidelberg;  
<http://gdbwww.dkfz-heidelberg.de/>  
Israel: Weizmann Institute of Science, Rehovot;  
<http://gdb.weizmann.ac.il/>  
Italy: TIGEM, Milan;  
<http://www.tigem.it/>  
Japan: JICST, Tsukuba Science City;  
<http://www.gdb.gdbnet.ad.jp/>  
Netherlands: CAOS/CAMM, University of Nijmegen;  
<http://www-gdb.caos.kun.nl/gdb/gdbtop.html>  
Sweden: Uppsala Biomedical Center, Uppsala;  
<http://gdb.embnet.se/gdb/>  
United Kingdom: MRC HGMP Resource Center, Hinxton;  
<http://www.hgmp.mrc.ac.uk/gdb/gdbtop.html>  
All of the data and documentation discussed in this article are available at these URLs as well.

## CITING THE GENOME DATABASE

When citing the Genome Database in the literature, please reference this article as:

Fasman, K.H., Letovsky, S.I., Li, P., Cottingham, R.W. and Kingsbury, D.T. (1997) The GDB™ Human Genome Database Anno 1997. *Nucleic Acids Res.*, Vol. **25**, 72–80.

## ACKNOWLEDGEMENTS

The GDB Human Genome Database is an international project funded by a grant from the US Department of Energy (DE-FC02-9ER6130) with additional support from the US National Institutes of Health, the Science and Technology Agency of Japan, the Medical Research Council of the United Kingdom, the INSERM of France, and the European Union.

## REFERENCES

- Pearson, P.L. (1991) *Nucleic Acids Res.*, **19**, 2237–2239.
- Pearson, P.L., Matheson, N.W., Flescher, D.C. and Robbins, R.J. (1992) *Nucleic Acids Res.*, **20**, 2201–2206.
- Cuticchia, A.J., Fasman, K.H., Kingsbury, D.T., Robbins, R.J. and Pearson, P.L. (1993) *Nucleic Acids Res.*, **21**, 3003–3006.
- Fasman, K.H., Cuticchia, A.J. and Kingsbury, D.T. (1994) *Nucleic Acids Res.*, **22**, 3462–3469.
- Fasman, K.H., Letovsky, S.I., Cottingham, R.W. and Kingsbury, D.T. (1996) *Nucleic Acids Res.*, **24**, 57–63.
- Chen, I.A. and Markowitz, V.M. (1995) *Information Systems*, **20**, 393–418. [http://gizmo.lbl.gov/DM\\_TOOLS/OPM/opm.html](http://gizmo.lbl.gov/DM_TOOLS/OPM/opm.html)
- Wacey, A.I., Kembell-Cook, G., Kazazian, H.H., Antonarakis, S.E., Schwaab, R., Lindley, P. and Tuddenham, E.G.D. (1996) *Nucleic Acids Res.*, **24**, 100–103.
- Giannelli, F., Green, P.M., Sommer, S.S., Poon, M.-C., Ludwig, M., Schwaab, R., Reitsma, P.H., Goossens, M., Yoshioka, A. and Brownlee, G.G. (1996) *Nucleic Acids Res.*, **24**, 103–119.
- Cariello, N.F., Douglas, G.R. and Soussi, T. (1996) *Nucleic Acids Res.*, **1**, 119–120.
- Bérout, C. and Soussi, T. (1996) *Nucleic Acids Res.*, **1**, 121–124.