An Integrated Gene Mapping Database "HyperGenome System"

Nakashima, Susumu
Laboratory of Molecular Gene Technics, Graduate school of Genetic Resources and Technology, Faculty of Agriculture, Kyushu University

Takagi, Yoshihisa
Institute of Medical Science University of Tokyo

Kanehisa, Minoru
Institute for Chemical Research Kyoto University

Kuhara, Satoru
Laboratory of Molecular Gene Technics, Graduate school of Genetic Resources and Technology, Faculty of Agriculture, Kyushu University

https://doi.org/10.5109/24031

出版情報：九州大学大学院農学研究院紀要. 37 (3/4), pp.379-383, 1993-03. Kyushu University

J. Fac. Agr., Kyushu Univ., 37 (3•4), 379-383 (1993)

An Integrated Gene Mapping Database "HyperGenome System"

Susumu Nakashima, Yoshihisa Takagi*,Minoru Kanehisa* *
and Satoru Kuhara

Laboratory of Molecular Gene Technics,
Graduate school of Genetic Resources and Technology, Faculty of Agriculture,
Kyushu University 46, Fukuoka 812, Japan

(Received December 28, 1992)

Features of the HyperGenome System are described herein. This system was developed to integrate information from the human genomic map from Genome Data Base (GDB) to DNA sequences from GenBank. DNA sequence data were assigned to locus on genomic map by checking cross references. The OpenWindows has been adopted for display information.

INTRODUCTION

There are various facts of experimental information related to the human genome project. The top of reference information we focused on is the genomic map, and the bottom information we have focused on is DNA sequences. Information on the human genomic map was collected in Genome Data Base (GDB). The DNA sequence information was gathered in GenBank, EMBL, and DDBJ. And the amino acid sequences were gathered in Protein Identification Resources (PIR). The formats of databases are different from each other. In recent years, certain kinds of cross references which link two databases were created by some databases (Bilofsky and Burks, 1988). A researcher, however, needs access to numerous databases to acquire information of the genome related to a particular research. It would be difficult for a molecular biologist to handle a SQL system for retrieving information. Integrated database systems have been constructed, for example Chromosome Information System (Johnston et al., 1991) and Encyclopedia of the Mouse Genome (Epping and Ormsby, 1990). Almost of those systems involve use of a pointer device as a mouse and a window system to get the information that a researcher requires.

In attempts to integrate information from the human genomic map and DNA sequences, we set up an integrated database HyperGenome for ready access to information that the user needs.

SYSTEM CONFIGURATION

The integrated database HyperGenome supports the parallel display of information with each appearing in a separate window. This system works in OpenWindows on SUN SPARCstation 2.

* Institute of Medical Science University of Tokyo, Minato-ku, Tokyo 108, Japan.
** Institute for Chemical Research Kyoto University Uji, Kyoto 611, Japan.

In the genomic map, we gathered information on the locus, probe and reference in GDB. In the DNA sequence, GenBank entry which corresponds to each locus in GDB was chosen from GenBank. The map information was chosen according to assignment to each chromosome, and the input data were transformed to the HyperGenome. The selected entries from GenBank were collected according to assignment to each chromosome then were transformed into input data to the HyperGenome (Fig. 1).

The assignment of entry of GenBank to locus in GDB was confirmed by agreement between the accession number described in probe of GDB and the map position mentioned in features of GenBank. The input errors were picked up in cross references. Appropriate correction for input error was made by inspection of the definition of the gene.

We applied the OpenWindows system and graphical user interface to acquire information from Locus to sequence by pointing button or locus in a window.

HyperGenome has nine functions: Chromosome, Locus, Select, Full, Reference, Sequence, Probe, Keywords, and Superimpose. These functions are performed using a clicking button (Fig. 2).

In the Basic window, the banding pattern and assigned gene symbol will be displayed and any locus can be selected by clicking on a gene symbol. Information on reference, probe, and sequence of a selected locus will be shown in each window by clinking on the button (Fig. 3).

Clicking on any chromosome in the chromosome window will make it appear in the Basic window. The keywords are collected from the gene symbol, name, location, title of reference, probe name, and GenBank accession number. We can therefore retrieve a locus by keywords in a keyword window (Fig. 4). The chromosome on which the selected locus assigned will be displayed and the gene symbol will be

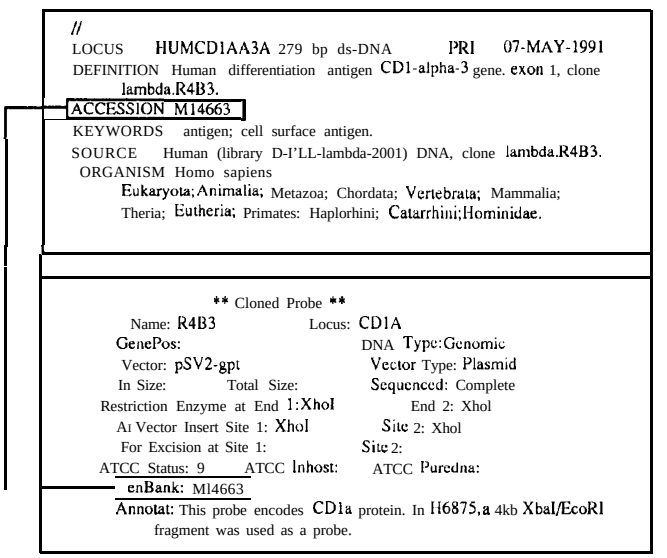

Fig. 1. Integration of GDB and GenBank.

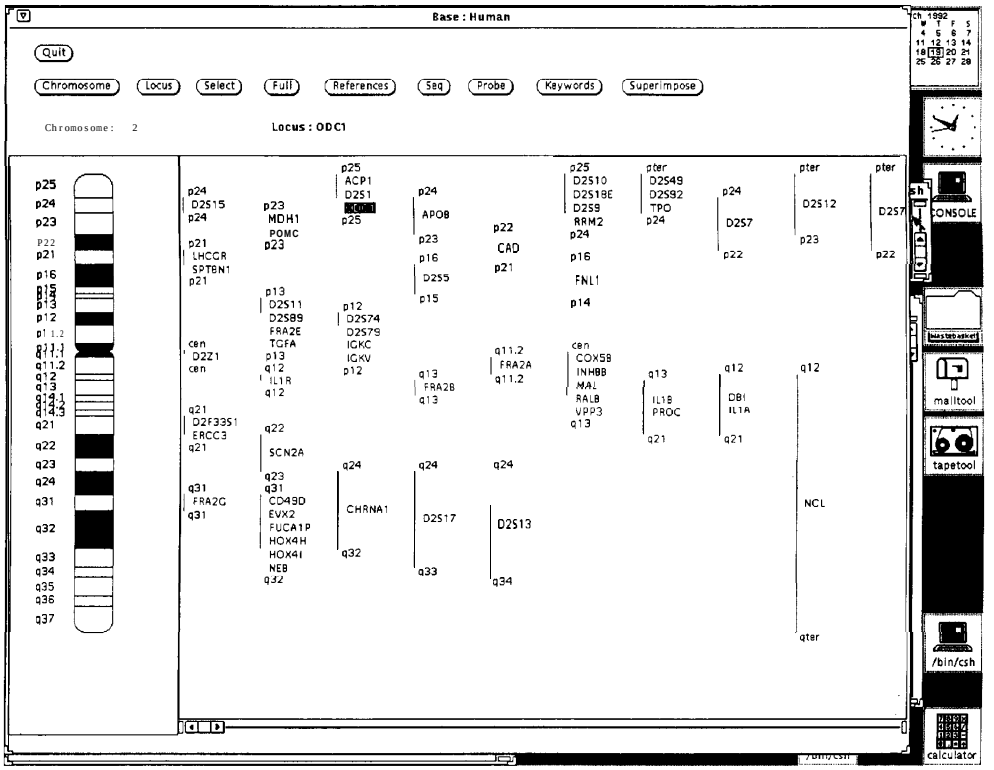

Fig. 2. Basic Window in the HyperGenome System.

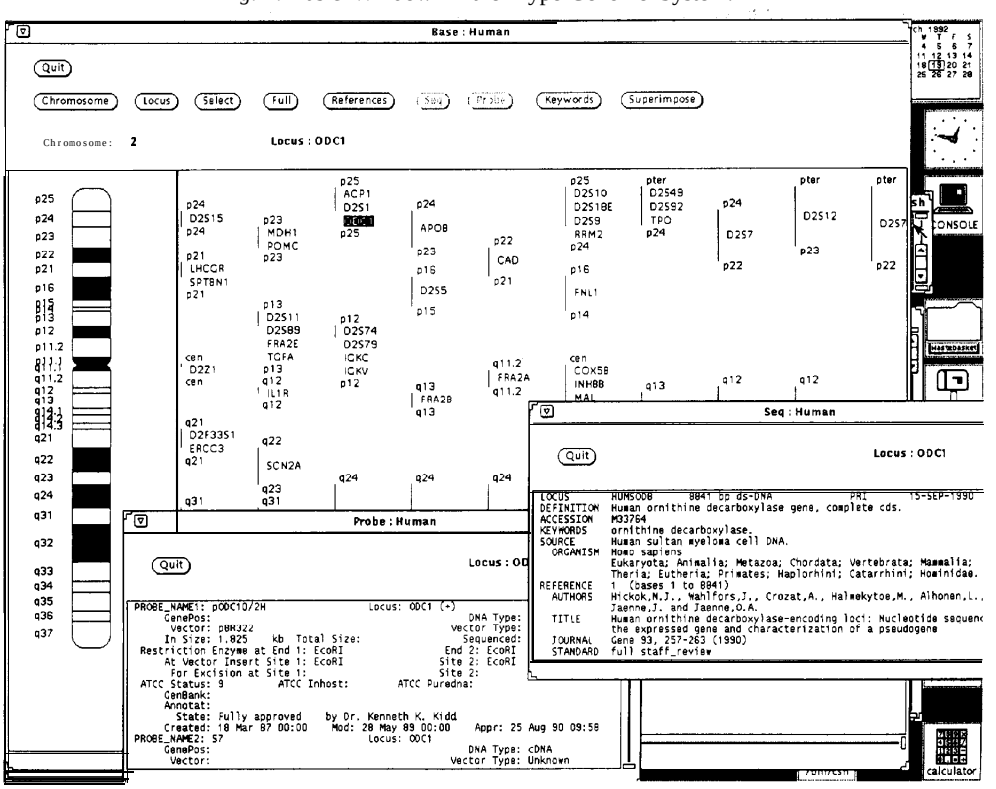

Fig. 3. Sequence and Probe Windows in the HyperGenome System.

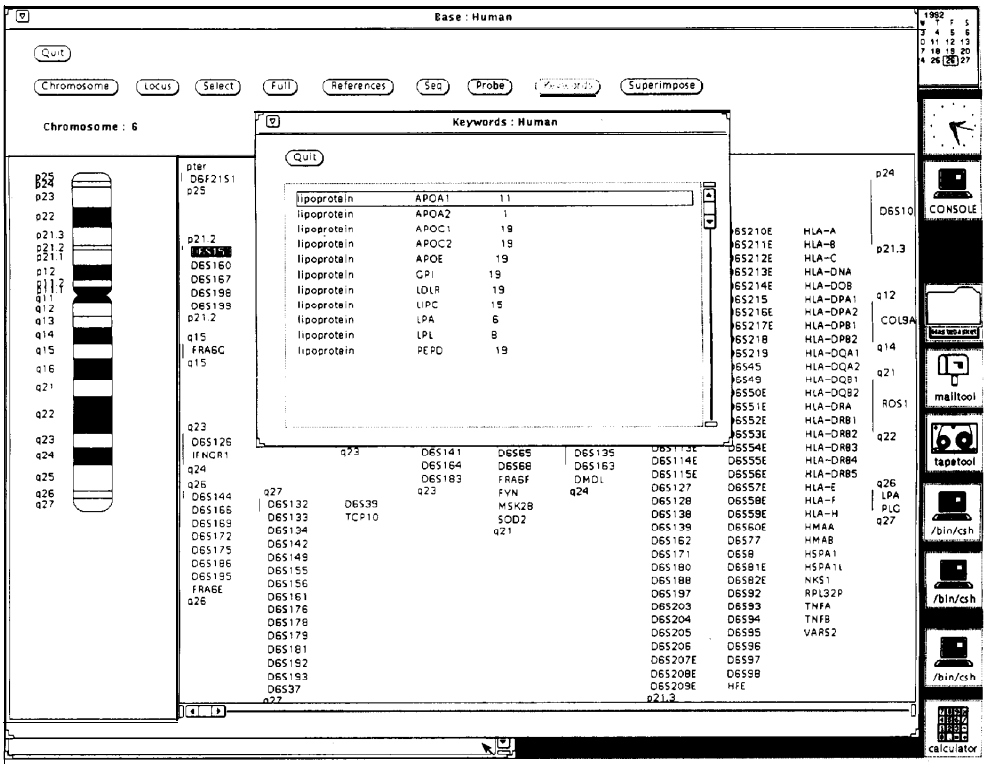

Fig. 4. Keywords Window in the HyperGenome System.

darkened in the Basic window.
The function of superimpose is to overlap a chromosome on the selected chromosome to determine the difference between chromosomes of different species.

CONCLUSION

We have registered over 4000 gene symbols and DNA fragments from GDB and over 1500 DNA sequence entries have been registered. Almost all entries related to the human genome project are incorporated into HyperGenome. This system can provide ready retrieval by graphical user interface. Thus, molecular biologists with little computer experience, can search for any kind of information on the genomic map to the DNA sequence.

ACKNOWLEDGEMENTS

This work was supported by a Grant-in Aid for Scientific Research on Priority Areas, “Genome Informatics”, from the Ministry of Education, Science and Culture of Japan.

REFERENCES

Bilofsky, H. S. and C. Burks 1988 The GenBank (R) Genetic Sequence Data Bank, Nuc. Acids Res, Vol. 16 1861-1864

Eppig, J. T. and J. E. Ormsby 1990 Encyclopedia of the mouse genome ver. 1.0, the Jackson Laboratory 600 Main Street Bar Harbor, ME

Johnston, W., S. Lewis, V. Markowitz, J. MaCathy, F. Olken and M. Zorn 1991 Chromosome Information System, Lawrence Berkeley Laboratory Technical Report LBL 29675