

# Supporting Genome Information Processing by MetaCommander

Yasuhiko Kitamura<sup>1</sup>

kitamura@info.eng.osaka-cu.ac.jp

Shoji Tatsumi<sup>1</sup>

tatsumi@info.eng.osaka-cu.ac.jp

Tetsuya Nozaki<sup>1</sup>

noz@kdel.info.eng.osaka-cu.ac.jp

Akira Tanigami<sup>2</sup>

atanigam@gan2.res.ncc.go.jp

<sup>1</sup> Department of Information and Communication Engineering  
Faculty of Engineering, Osaka City University  
3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558 Japan

<sup>2</sup> Radiobiology Division, National Cancer Center Research Institute  
5-1-1, Tsukiji, Chuo-ku, Tokyo 104, Japan

## Abstract

*MetaCommander is developed as a generic tool to retrieve and integrate information from WWW servers by interpreting a script. By using MetaCommander, we can support genome information processing on WWW browsers in various ways.*

## 1 MetaCommander

With the advance of genome informatics and Internet technologies, an overwhelming large amount of genome data has been stored in a number of databases which are distributed over the world, and nowadays it is accessible through the WWW system. As the number and size of genome databases increase, we have more demands to retrieve and collect a large amount of data through WWW browsers, but to meet the demands we have tiresome operations such as repeating similar operations many times and inputting various combinations of parameters.

We developed MetaCommander [1, 2] as a remedy for this problem. By preparing a script, MetaCommander can automates many of tiresome operations on the WWW. Figure 1 show the overview of MetaCommander system. A user gives commands to MetaCommander by writing a MetaCommander (MC) script. MetaCommnader interprets the script and executes the commands. It gets access to WWW servers through the Internet by talking HTTP (Hyper Text Transfer Protocol) and is also able to read and write local data files. It rearranges retrieved data and outputs the result as an HTML file, so the user can read it by using a WWW browser.

MetaCommander is a public domain software which is written in Java language, so it runs on most of platforms such as Macintosh, Windows95, and UNIX. It is available by down-loading the code from our MetaCommander homepage (<http://www.kdel.info.eng.osaka-cu.ac.jp/MetaCommander/>).

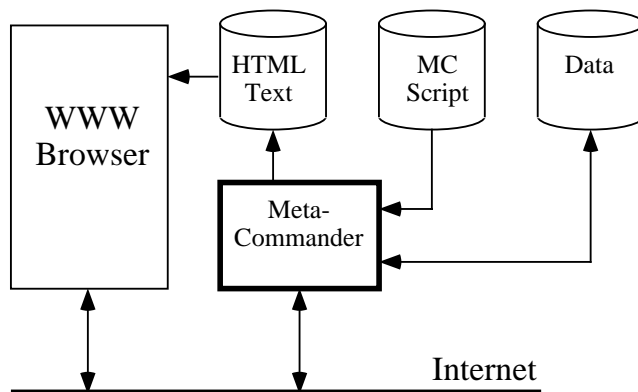


Figure 1: Overview of MetaCommander.

## 2 MetaCommander Scripts

```
1: set( url, "http://www.genome.ad.jp/htbin/www_bfind_sub" ) /* URL of Genome Net */
2: set( dbkey, "genbank" ) /* database name */
3: set( mode, "bfind" ) /* data retrieve method */
4: set( max_hit, 10 ) /* max number of data hits */
5: set( keywords, "hiv human mRNA") /* search keywords */

6: getURL( $url, "dbkey"=$dbkey,"keywords"=$keywords, "mode"=$mode,
  "max_hit"=$max_hit) { /* get access to genbank with parameters *
7:   getAnchor($max_hit) { /* get each of hyperlinks from index page */
8:     print /* print the hyperlink */
9:     tag("BR") /* break the line */
10:    getURL($_) { /* follow the hyperlink and retrieve the page */
11:      cutString("REFERENCE","FEATURES") {
12:        /* cut data between REFERENCE and FEATURES */
13:        tag("PRE") { /* print the data as a plain text */
          print }}}}
```

Figure 2: An example of script.

Figure 2 shows an example of MetaCommander script. This script is for collecting reference data concerning keywords “hiv human mRNA” from GenBank and integrating them as an HTML text. We have made a number of MetaCommander scripts for supporting genome data processing as follows, and some of them are available through the script library of our MetaCommander homepage.

- Automating parameter inputting for BLAST<sup>1</sup>, GDB Insert Amplimer<sup>2</sup>, Grail<sup>3</sup>, and Hexon<sup>4</sup>.
- Integrating reference data from GenBank<sup>5</sup> and abstracts from Medline<sup>6</sup>.
- Integrating Grail<sup>3</sup> and Hexon<sup>4</sup> for better gene prediction.
- Extracting updated data from Genbank-upd<sup>7</sup> into a personalized database.

## Acknowledgments

The authors would like thank Toru Ishida, Teruhisa Miura, Hideyuki Nakanishi, Minoru Kanehisa, Shinsei Minoshima, and Katsuhiko Murakami for their helpful comments. This work was supported in part by a Grant-in-Aid for Scientific Research on Priority Areas “Genome Science” from The Ministry of Education, Science, Sports and Culture of Japan.

## References

- [1] Y. Kitamura, et al., “MetaViewer and MetaCommander: Applying WWW Tools to Genome Informatics,” Proceedings of the Seventh Workshop on Genome Informatics, 137-146, 1996.
- [2] Y. Kitamura, et al., “A Script-based WWW Information Integration Support System and Its Application to Genome Databases,” Proceedings of IEICE Symposium on Software Agents and Their Applications, 87-92, 1997.

<sup>1</sup><http://www.genome.ad.jp/SIT/BLAST.html>

<sup>2</sup><http://gdbwww.gdb.org/gdb-bin/genera/restrict/genera/hgd/Amplimer?!action=insertform>

<sup>3</sup><http://avalon.epm.ornl.gov/Grail-bin/EmptyGrailForm>

<sup>4</sup><http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html>

<sup>5</sup>[http://www.genome.ad.jp/htbin/www\\_bfind?genbank-today](http://www.genome.ad.jp/htbin/www_bfind?genbank-today)

<sup>6</sup>[http://www.genome.ad.jp/dbget-bin/www\\_bfind?medline](http://www.genome.ad.jp/dbget-bin/www_bfind?medline)

<sup>7</sup>[http://www.genome.ad.jp/htbin/www\\_bfind?genbank-upd](http://www.genome.ad.jp/htbin/www_bfind?genbank-upd)