

PLACE: a database of plant *cis*-acting regulatory DNA elements

Kenichi Higo*, Yoshihiro Ugawa, Masao Iwamoto and Hiromi Higo

Department of Genetic Resources, National Institute of Agrobiological Resources, 2-1-2 Kannondai, Tsukuba, Ibaraki 305, Japan

Received September 11, 1997; Revised and Accepted October 10, 1997

ABSTRACT

PLACE (<http://www.dna.affrc.go.jp/htdocs/PLACE/>) is a database of motifs found in plant *cis*-acting regulatory DNA elements, all from previously published reports. It covers vascular plants only. In addition to the motifs originally reported, their variations in other genes or in other plant species reported later are also compiled. The PLACE database also contains a brief description of each motif and relevant literature with PubMed ID numbers. This report summarizes the present status of this database and available tools.

INTRODUCTION

Recent rapid progress in gene cloning techniques for higher plants as well as other higher eukaryotes has provided an opportunity to isolate plant genes and to determine their nucleotide sequences without the conventional laborious steps, which include purification of their gene products from plants and their characterization. Using such cloned DNAs, we can now perform various studies on the genes and their expression, or can create artificial mutants. Thus 'reverse genetics', the term coined for these types of approaches, has become applicable to higher plants. It is therefore desirable that biological information can be extracted from the nucleotide sequences by computer analyses. A database of nucleotide sequence motifs found in plant *cis*-acting regulatory DNA elements (*cis*-elements) and a tool for homology searches within such a database will be helpful in estimating the mode of gene regulation, regions involved in such regulation, and other pertinent regions in the DNA sequence. Previous databases created for similar purposes such as TFD (1–3) and TRANSFAC (4,5) contain many *cis*-elements, but only a very small number of those from plant genes were compiled.

PLACE is based on a personal simple database on plant DNA *cis*-elements, the compilation of which was started in 1991. The *cis*-elements have been extracted from the original literature, but review articles on the regulatory regions of some plant genes (6,7) were also very useful for compilation of specific groups of motifs. The database was transferred in March 1997 to a server located in the MAFF (Ministry of Agriculture, Forestry and Fisheries, Japan) DNA Bank at the NIAR (National Institute of Agrobiological Resources), and has been maintained thereafter in collaboration

with colleagues at the DISC (DNA Information and Stock Center), NIAR. In this brief report, we describe the present structure of the PLACE database, its content and the database analysis tools which are presently available.

DATABASE ORGANIZATION

The PLACE database is a compilation of motifs found in plant *cis*-acting regulatory DNA elements, extracted from previously published reports. Not only the motifs originally reported, but also their variations in other genes or other plant species reported later are also included. PLACE also includes some motifs in non-plant *cis*-elements which may have plant homologues. PLACE also contains a brief description of each motif and relevant literature with PubMed ID numbers of MEDLINE database at NCBI-NIH, USA. We intend to add PubMed ID numbers to all references. Identification of most of the *cis*-elements compiled was based upon several different types of experiments. We intend to add indications of the experimental evidence to all the motifs in the Reference Criteria (RC) field in each document. The symbols used, in addition to A, G, C or T, for the sequences of *cis*-acting element motifs are according to the recommendation by IUPAC-IUB.

HOW TO MAKE USE OF PLACE

Access to PLACE database

Presently, the PLACE database is accessible via WWW: <http://www.dna.affrc.go.jp/htdocs/PLACE/>. It is also available by anonymous ftp from the ftp server: ftp://ftp.dna.affrc.go.jp/pub/dna_place/. Information on *cis*-elements can also be obtained by keyword search through the WWW page. Almost any query word can be used as the 'keyword', such as motif name, name of inducer or plant hormone involved, type of stress, names of tissues or organs in which the gene is expressed, name of author of the original report, motif nucleotide sequence, plant species, etc. The result report will show the list of motifs which match the query word. By clicking the PLACE accession number in the list, a brief description of the motif and the reference with PubMed and/or MEDLINE ID numbers will be presented. We intend to link the PLACE database to the MEDLINE database Web site at NCBI-NIH, USA, so that clicking the ID number will allow reading of the abstract of the paper, other information, and the accession number of the sequence in the DDBJ/EMBL/GenBank nucleotide sequence databases.

*To whom correspondence should be addressed. Tel: +81 298 38 7016; Fax: +81 298 38 7408; Email: kenhigo@abr.affrc.go.jp

Homology search against PLACE database

The query sequence can be searched for the presence of motifs identical with or similar to the previously reported *cis*-element motifs in the PLACE database by homology search tools. Presently the Web version of the SIGNAL SCAN program (8) is available at our Web site. The query sequence can be entered by copying and pasting into the window. Various forms of result presentation can be chosen: motif group order, sequence order or sequence map. The result report will show the site (motif) name, location, motif sequence and the PLACE accession number. Information can be obtained on the motifs with the WAIS program described above by using the accession number as the keyword. We intend to link the result report to the WAIS program so that clicking the accession number will allow viewing of information on the motifs in the document.

Notes on PLACE Releases 1.0 and 1.1, and future prospects

The PLACE database Release 1.0, which was opened to the public via the Internet in March 1997, contained 114 entries. Some information, e.g. original literature in the reference and PubMed ID numbers, was however missing. Most of these and other empty fields were filled in Release 1.1, which was opened to the public at the end of August, 1997. Release 1.1 contains 103 motifs, because although new motifs were added some were removed because the original literature were not available by the time of release. We will add these motifs back to the database as soon as the sequences and

their descriptions are confirmed in the original literature. We intend to add more motifs, and plan to update the database twice a year.

Reference to PLACE database

Users are asked to cite this article when publishing results which have been obtained with the database described here.

ACKNOWLEDGEMENTS

Authors wish to thank Ms Meena Sakharkar, National University of Singapore for providing the source code of the WEB version of Signal Scan. We are indebted to our colleagues at the DISC for their excellent technical support. Part of this work was supported through a grant from the Ministry of Agriculture, Forestry and Fisheries (MAFF) of Japan in the framework of the Research Project for New Recombinant DNA Techniques.

REFERENCES

- 1 Ghosh,D. (1990) *Nucleic Acids Res.* **18**, 1749–1756.
- 2 Ghosh,D. (1991) *Trends Biochem. Sci.* **16**, 445–447.
- 3 Ghosh,D. (1993) *Nucleic Acids Res.* **21**, 3117–3118 [see also this issue (1998) *Nucleic Acids Res.* **26**, 360–361].
- 4 Wingender,E., Dietze,P., Karas,H. and Knüppel,R. (1996) *Nucleic Acids Res.* **24**, 238–241 [see also this issue (1998) *Nucleic Acids Res.* **26**, 362–367].
- 5 Wingender,E., Kel,A.E., Kel,O.V., Karas,H., Heinemeyer,T., Dietze,P., Knüppel,R., Romaschenko,A.G. and Kolchanov,N.A. (1996) *Nucleic Acids Res.* **25**, 265–268.
- 6 Gilmartin,P.M., Sarokin,L., Memelink,J. and Chua,N.-H. (1990) *Plant Cell* **2**, 369–378.
- 7 Thomas,T.L. (1993) *Plant Cell* **5**, 1401–1410.
- 8 Prestridge,D.S.(1991) *Comput. Appl. Biosci.* **7**, 203–206.