

Simple database to select promoters for plant transgenesis

Olga G. Smirnova · Salmaz S. Ibragimova ·
Alex V. Kochetov

Received: 18 February 2011 / Accepted: 8 July 2011 / Published online: 3 August 2011
© Springer Science+Business Media B.V. 2011

Abstract The experiments with transgenic plants frequently demand selection of promoters providing appropriate transcription patterns. The set of promoters commonly used in vectors and genetic constructs is very limited, and these promoters provide only a few variants of gene expression patterns. Moreover, identical promoters in a complex construct can induce transgene silencing. This problem can be solved using a variety of plant gene promoters with experimentally verified characteristics. However, this requires a time-consuming analysis of literature data. Here, we describe a database of plant promoters (TransGene Promoters, TGP; <http://www.mgs.bionet.nsc.ru/mgs/dbases/tgp/home.html>). TGP contains the information on genomic DNA segments providing certain expression patterns of reporter genes in experiments with transgenic plants. TGP was constructed on the SRS platform, and its interface allows users to search for the promoters with particular characteristics.

Keywords Transgenic plants · Gene engineering · Promoter · Database

Introduction

To carry out experiments with transgenic plants, the researchers commonly need to design specific genetic constructs providing a certain expression pattern of the transgene(s). The most important element of any genetic construct is a promoter providing an appropriate transcription pattern. The requirements for promoter activity can vary considerably: a constitutive transcription in all tissues is demanded in some cases and tissue-specific, stage-specific, or inducible transcription is required in others. However, the set of commonly used promoters is very small, and these promoters provide only a few variants of transgene transcription pattern. It is quite evident that this limitation hampers experiments with transgenic plants (Furtado et al. 2008; Qu le et al. 2008).

Information on promoters of many plant genes is available in the literature. The gene expression control is frequently studied with the help of a reporter construct where the promoter DNA segment is located upstream of a reporter gene. Analysis of the reporter protein activity in transgenic plants allows for assessing the functional characteristics of the promoter being studied. The promoters described can, in principle, provide a wide range of choice of appropriate transcription patterns. For instance, if a foreign gene can be expressed at a high level but in a tissue-specific manner, a dozen variants could be considered. Thus, the tomato Lat52 promoter segment provides a high level of GUS expression in

O. G. Smirnova (✉) · S. S. Ibragimova · A. V. Kochetov
Institute of Cytology and Genetics, 10, Lavrentieva ave,
630090 Novosibirsk, Russia
e-mail: planta@bionet.nsc.ru

A. V. Kochetov
Novosibirsk State University, 2, Pirogova ave,
630090 Novosibirsk, Russia

tobacco pollen (Bate and Twell 1998); rice prolamin and glutenin promoter segments increase by four–sixfold the GUS gene activity in transgenic rice endosperm as compared with the maize ubiquitin promoter (Qu and Takaiwa 2004); oat globulin promoter segment directs a strong endosperm-specific GUS expression in barley seeds (up to 10% of soluble protein; Vickers et al. 2006); and sweet potato ADP-glucose pyrophosphorylase promoter provides a high-level expression of the GUS reporter gene in *Solanum tuberosum* tubers (Kim et al. 2009). In some cases, it has been demonstrated that such tissue-restricted expression is more beneficial as compared with the typical constitutive promoters (e.g., cauliflower mosaic virus 35S RNA or maize ubiquitin promoters; Qu and Takaiwa 2004; Tiwari et al. 2006; Eskelin et al. 2009). Development of new methods for increasing plant tolerance to various stresses also requires promoters with a specific transcription pattern, for example, RD29A, COR15A, and DREB1 (Yamaguchi-Shinozaki and Shinozaki 1993; Baker et al. 1994; Yang et al. 2011). In general, a toolbox of promoters with known specificities would be a valuable resource to control the expression of transgenes in an appropriate manner for both plant improvement and molecular farming (Furtado et al. 2008; Qu le et al. 2008). The usage of promoters with less sequence homology but similar specificities will also be crucial in avoiding homology-based gene silencing when expressing more than one transgene in the same tissue (Furtado et al. 2009).

Experiments with transgenic constructs are frequently used to clarify the structure–function organization of promoters. For this purpose, several types of experiments are commonly used, namely, deletion analysis, point mutagenesis, and detection of *cis*-acting transcription factor (TF) binding sites. Despite an ultimate goal of such research being a full reconstruction of promoter organization (i.e., detection of the full set of TF binding sites and combinatorial pattern of their activities), the corresponding papers frequently also contain other valuable data. While characterizing promoters, researchers usually create a set of genetic constructs carrying the genomic DNA segments from the studied promoter region, located upstream of the reporter gene coding sequence (CDS). The constructs are expressed in transgenic plants, where the pattern of reporter protein synthesis (commonly, GFP or GUS) reflects

the promoter characteristics. These data can further be used to detect TF binding sites and model the promoter structure and the mechanisms involved in transcription control. However, such information is evidently useful per se, since the genomic DNA segments with appropriate patterns of transcriptional activity represent potential promoters for plant transgenesis. These data are particularly interesting since different deletion variants can contain various combinations of TF binding sites and provide transcriptional patterns other than the full promoter regions. In our opinion, a specialized database on specific transcriptional activities of plant genomic DNA segments could be a valuable source of new candidate promoters for transgenic experiments.

A number of databases containing the information about promoter sequences and TF binding sites have been developed. However, most of them only accumulate the data on promoter organization and *cis*-acting regulatory elements. Here follows a short list of available www resources. The PLACE database (Higo et al. 1999) stores the consensus binding sequences for plant-specific transcription factors. Three interlinked databases—AtTFDB, AtcisDB, and AtRegNet in the Arabidopsis Gene Regulatory Information Server (AGRIS)—furnish comprehensive and updated information on the TFs, predicted and experimentally verified *cis*-regulatory elements, and their interactions (Yilmaz et al. 2011). The PlantProm database (PPDB) contains the nucleotide sequences of plant promoters with experimentally verified transcription start sites (Shahmuradov et al. 2003). PPDB contains information on *Arabidopsis* and rice promoter structures, and the transcription start sites predicted from full-length cDNA clones to TSS tag data. The core promoter structure, presence of *cis*-acting regulatory elements, and distribution of transcription start site clusters can also be viewed (Yamamoto and Obokata 2008). Athena and Osiris are the www resources for rapid visualization and systematic analysis of *Arabidopsis* (O'Connor et al. 2005) and rice (Morris et al. 2008) promoter sequences. Athena contains up to 3 kb of the promoter sequences for predicted *Arabidopsis* genes and the consensus sequences for 105 previously characterized TF binding sites imported from PLACE to AGRIS. Osiris contains the promoter sequences, predicted TF binding sites, gene ontology annotations, and microarray expression data for 24,209

genes of the rice genome. AthaMap database provides a genome-wide map of potential TF binding sites in *Arabidopsis thaliana*. The database contains the sites for 115 different TFs (Bülow et al. 2010). PlantPAN is the Plant Promoter Analysis Navigator for recognizing combinatorial *cis*-regulatory elements with a distance constraint in sets of plant genes (Chang et al. 2008). Thus, the available www resources mostly provide data on transcription factor binding sites and lack the information on experimentally verified promoter activities of plant DNA segments. In principle, some data (e.g., the presence of specific *cis*-acting regulatory elements) can be used to select candidate promoters. However, such predictions are most frequently ambiguous, as demonstrated by the following two examples: (1) despite the GluC promoter containing no endosperm-specific motifs (GCN4, AACA, or prolamin box), it directs a high-level transcription in this tissue (Qu le et al. 2008), and (2) the peach gene *Pptha1* was detected by its cold-inducibility, but the promoter region failed to determine such expression pattern in *A. thaliana* (Tittarelli et al. 2009), etc. In general, the information on the DNA segments with certain types of transcriptional activity in transgenic experiments looks to be the most reliable source of potential promoters (Potenza et al. 2004; Jones and Sparks 2009; Peremarti et al. 2010; Xiao et al. 2010). Thus, we developed a specialized database (TransGene Promoters, TGP) compiling information annotated from the literature. The TGP provides the data on candidate promoters with experimentally verified transcriptional patterns in transgenic plants of different species.

TGP database description

Although the set of very well-characterized promoters used in plant gene engineering is rather small, published experimental investigation of many plant genes has frequently provided some details on DNA segments transcriptional activities. For example, study of a promoter region commonly includes deletion analysis where the DNA segments of different lengths are placed upstream of the reporter gene and their expression characteristics are tested in experiments with the transgenic plants. This information may be found in the literature and further used for the selection of a promoter with potentially appropriate properties. Currently, this information

may not be automatically retrievable and may not have so far been accumulated in database format.

The database was constructed on the SRS (Sequence Retrieval System) platform and contains three cross-linked sections: TGP_PROMOTER, TGP_SEQUENCE, and TGP_GENE. Typically, study of the promoter region of a target gene involves experimental analysis of several deletion variants (i.e., genomic DNA segments of different lengths). If some of such DNA segments demonstrate certain promoter activities in experiments with reporter constructs in transgenic plants, they are selected for annotation in the TGP_PROMOTER section, while the data on their nucleotide sequences and the corresponding gene are annotated in the TGP_SEQUENCE, and TGP_GENE sections, respectively.

The TGP_PROMOTER (Table 1) contains information about the promoter size, positions in corresponding GenBank entry, position of transcription start site (TSS) or translation initiation site (TIS) as well as a general description of the promoter sequence used in transgene construct (fields LOCALIZATION and DESCRIPTION). [Example: the regulatory region upstream of the coding part of the reporter gene includes a promoter fragment (387 bp upstream of the transcription start site), 89 bp of the 5'UTR, and 68 bp of the coding sequence of the potato Ci21A gene.] This information is important for designing transgene constructs. The entry also contains data on the plant species used in experiments, reporter gene, and the factors influencing transcription. The field TARGET SPECIES provides a list of plant species where promoters were tested in transgenic experiments. For example, the P1 promoter of the *Arabidopsis* *SAG12* gene was evaluated in nine species. The field COMMENT gives a summary of the experimental data potentially useful for evaluation of promoter specificity and expression pattern. The nucleotide sequence of the described promoter can be retrieved from the corresponding GenBank entry according to the positions indicated in the LOCALIZATION field. Nonetheless, for the sake of convenience, we additionally compiled the promoter nucleotide sequences indicating the promoter positions relative to TSS or TIS (marked as +1) in the TGP_SEQUENCE section (Table 1). Finally, TGP_GENE (Table 1) contains the descriptions of native genes from which the promoter variants were annotated. Each entry contains the data on the

Table 1 Examples of the TGP_PROMOTER, TGP_SEQUENCE and TGP_GENE entries**TGP_PROMOTER****PROMOTER_ID** Ps:TOP2_P3**GENE_ID** Ps:TOP2**TARGET SPECIES** mouse-ear cress (*Arabidopsis thaliana*), tobacco (*Nicotiana tabacum*)**KEYWORDS** stress response, white light-induced, cold-induced, abscisic acid-induced, cotyledon**LOCALIZATION** from −362 to +74; GenBank; AF144649; TSS: 609; 247–682**DESCRIPTION** The regulatory region upstream of the coding part of the reporter gene includes a promoter sequence (262 bp upstream of the transcription start site) and 74 bp of the 5'UTR of the PsTOP2 gene.**SEQUENCE_ID** Ps:TOP2_P3S**REPORTER** GUS**STAGE_ORGAN_TISSUE** seedling, cotyledon**REGULATOR** white light, cold, abscisic acid**COMMENT** The minimal TOP2 promoter that is induced by cold.Transgenic *Arabidopsis*The −362 bp TOP2 promoter showed very little stimulation in GUS, if any, in white light, red light, far-red light and blue light as compared to dark-grown *Arabidopsis* seedlings (Hettiarachchi et al. 2003).

Transgenic tobacco

The expression of the −362 bp TOP2 promoter was detectable only in the tobacco cotyledons (Hettiarachchi et al. 2003). The −362 bp TOP2 promoter showed about a twofold higher level of activity at 12 h, with very little induction, if any, up to 4 h of cold treatment of transgenic tobacco seedlings. The −362 bp TOP2 promoter had no detectable induction by salt or salicylic acid treatment. The −362 bp TOP2 promoter showed very little induction, if any, at 100 microM abscisic acid and about twofold induction at 150 microM abscisic acid (Hettiarachchi et al. 2005)

REFERENCE Hettiarachchi G.H., Yadav V., Reddy M.K., Chattopadhyay S., Sopory S.K. Light-mediated regulation defines a minimal promoter region of TOP2. *Nucleic Acids Res.*, 2003, 31(18), 5256–5265.**PUBMED** 12954761**REFERENCE** Hettiarachchi G.H., Reddy M.K., Sopory S.K., Chattopadhyay S. Regulation of TOP2 by various abiotic stresses including cold and salinity in pea and transgenic tobacco plants. *Plant Cell Physiol.*, 2005, 46, 1154–1160.**PUBMED** 15879449**TGP_SEQUENCE****SEQUENCE_ID** Ps:TOP2_P3S**PROMOTER_ID** Ps:TOP2_P3**GENE_ID** Ps:TOP2**SEQUENCE**

−362 ttaa	taaccctagt	ttgacactat	aaatactaaa	gatgctggtg	aatgaaagaa	gaaaaccaac
agatgcctag	cgcgtagecc	cgaaatgecc	tctctctca	ctctccacct	accaacaacc	ggataccccc
acgtgtatgc	caacaaaaac	attaaaagac	acactgccag	aactgataca	acaacacaca	ctcacaaaat
caccatcctc	ctcaccctcc	atctctctcc	acctgctccc	tccactctca	atccgccgaa	aaagcaccac
gccggcaacc	acaaacctaa	tgcttctca	acctccaatc	tccacccttc	attcttctcc	accgttcgtg
	+1					
ttttatcgt	tggtgctcCt	cacctccacc	caacgacacc	aaaaatcctg	cgagcaaacc	ctcaagccgc
cgcaagccca	ccatctaccg	tc +74				

TGP_GENE**GENE_ID** Ps:TOP2**GENE** top2**PRODUCT** DNA topoisomerase 2, TOP2, TOPII

Table 1 continued**TGP_GENE**

TAXON Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; *Pisum*.

SPECIES pea (*Pisum sativum*)

KEYWORDS stress response, salt-induced, cold-induced, abscisic acid-induced, salicylic acid-induced, white light-induced, red light-induced, far-red light-induced, blue light-induced, cell proliferation

ACTIVITY The topo II transcript level was maximal in actively growing tissues such as root tips and young leaves from the apical region as compared to the differentiated tissues such as the upper part of the root and internodal region of the stem of 7- to 10-day old seedlings. The topo II transcript was also present in the young and mature flower buds and in immature pods. The topo II transcript was abundant in proliferative tissues. The level of topo II transcripts could be stimulated by exogenous application of growth factors that induced proliferation in vitro cultures. The increase in the topo II transcript was seen within an hour after irradiation of etiolated seedlings (Reddy et al. 1999).

The expression of TOP2 was more than tenfold higher in constant white light grown seedlings as compared to the darkness. The expression of TOP2 was induced up to eightfold by a broad spectrum of light (red, far-red and blue). The rate of blue light-mediated induction of TOP2 was slower when compared to the red light- and far-red light-mediated inductions. The expression of TOP2 was detected to be at the highest level with about tenfold more as compared to dark in 7-day-old plants. The transcript level significantly decreased in 14-day-old plants and showed only about threefold more expression than the dark in 21-day-old plants (Hettiarachchi et al. 2003).

In 10-day-old pea seedlings, 2.5-fold higher expression of TOP2 mRNA was detected in shoots after 2 h of cold treatment and this increased level remained the same up to 12 h exposure to cold. Although the cold-mediated induction of TOP2 in roots was slower, more than six-fold induction was observed in roots after 12 h exposure to cold. TOP2 mRNA is up-regulated by salinity stress and the response was stronger in roots (seven-fold) compared with shoots (three-fold). The expression of TOP2 was up-regulated in response to exogenous abscisic acid (ABA) and the maximum level of induction was observed at 150 mM ABA in both root (fourfold) and shoot (2.5-fold) tissues. Salicylic acid could induce the maximum level of expression of TOP2 at 100 mM in both shoot (2.5-fold) and root (six-fold) tissues. No significant change in the expression of TOP2 in 10-day-old pea seedlings under dehydration stress was detected (Hettiarachchi et al. 2005).

PROMOTER_ID Ps:TOP2_P1 Ps:TOP2_P2 Ps:TOP2_P3 Ps:TOP2_P4

SEQUENCE_ID Ps:TOP2_P1S Ps:TOP2_P2S Ps:TOP2_P3S Ps:TOP2_P4S

REFERENCE Reddy M.K., Nair S., Tewari K.K., Mudgil Y., Yadav B.S., Sopory S.K. Cloning and characterization of a cDNA encoding topoisomerase II in pea and analysis of its expression in relation to cell proliferation. *Plant Mol. Biol.*, 1999, 41(1), 125–137.

PUBMED 10561074

REFERENCE Hettiarachchi G.H., Yadav V., Reddy M.K., Chattopadhyay S., Sopory S.K. Light-mediated regulation defines a minimal promoter region of TOP2. *Nucleic Acids Res.*, 2003, 31(18), 5256–5265.

PUBMED 12954761

REFERENCE Hettiarachchi G.H., Reddy M.K., Sopory S.K., Chattopadhyay S. Regulation of TOP2 by various abiotic stresses including cold and salinity in pea and transgenic tobacco plants. *Plant Cell Physiol.*, 2005, 46, 1154–1160.

PUBMED 15879449

corresponding protein and species as well as some information about gene activities. All these sections are cross-linked.

Currently, the TGP database contains description of 224 promoters with their nucleotide sequences corresponding to 114 genes. They belong to 26 plant species (mostly to *A. thaliana*, *Oryza sativa*, and *Nicotiana tabacum*). The transgenic experiments have been made with 28 plant species. The database describes the promoters whose expression is sensitive to 37 different exogenous and endogenous stimuli, such as heavy metals, elicitors, hormones, cold,

drought, salt, dehydration, infection, light, senescence, etc. According to the annotated information, these promoters are active in 40 different tissues and cell types, namely, seeds (51 promoters), roots (47 promoters), and pollen (12 promoters); more information is listed in Table 2.

How to search in TGP database

The TGP database allows the user to search for candidate promoters whose expression was

Table 2 Number of tissue-specific and μ regulated promoters in the TGP database

Regulators	Number of promoters	Organs and tissues	Number of promoters
Salicylic acid	23	Seed	51
Abscisic acid	23	Root	47
Jasmonic acid	14	Leaf	44
Ethylene	11	Embryo	24
Auxin	7	Endosperm	22
Auxin inhibitor	15	Inflorescence	10
Brassinolide	2	Stem	20
Gibberellin	5	Vascular	16
Cytokinin	1	Shoot	14
Cold	22	Flower	11
Salt	13	Pollen	12
Heat shock	3	Fruit	4
Anaerobic conditions	2	Aleurone	12
Dehydration	2	Epiderm	12
Flooding	3	Cotyledon	10
Drought	7	Parenchyma	10
Osmotic stress	2	Meristem	7
Elicitor	34	Phloem	12
Tobacco mosaic virus	14	Anther	6
Wounding	13	Tuber	6
Heavy metal	12	Hypocotyl	5
Oxidative stress	6	Nodule	5
Ozone	1	Xylem	7
Hydrogen peroxide	3	Grain	4
Paraquat	3	Sepal	4
Mitomycin C	2	Tapetum	4
White light	10	Xylem	4
Red light	8	Endoderm	3
Far-red light	8	Ovary	3
Blue light	2	Primordia	1
UV light	2	Coleoptile	2
Gamma irradiation	1	Carpel	2
High light	2	Lemma	2
Sucrose	8	Pericarp	2
Glucose	6	Stamen	2
Sugar deprivation	2	Cortex	1
Senescence	7	Silique	1

experimentally studied in particular species. TGP has a user-friendly SRS interface. Detailed tutorial (how-to-use) is available at the database www site. It is possible to use a combined search with the help of

different logic operators (AND, OR, etc.). Below, we have provided a brief description of a few typical queries. TGP (based on the SRS system) allows for various types of queries, for example:

- finding promoters working in a particular plant species;
- finding promoters influenced by a particular regulator;
- finding promoters working in a particular plant species and influenced by a particular regulator;
- finding promoters isolated from particular plant species;
- finding promoters influenced by several different regulators;
- finding promoters active in certain organ or tissue; and
- finding the tissue-specific promoters responsible to a particular regulator.

On the home page of the TGP_PROMOTER section, the field's names are present at the left column (Fig. 1). State “ok” at the right column indicates that the corresponding fields are searchable. Column “No of Keys” reflects the number of terms, for example, field “Target species” contains 27 different species; field “Keywords” contains 94 different terms, etc. To find the promoters working in a particular plant species, the user may click the field “Target species” at the home page of the TGP_PROMOTER table (marked by arrow 1 in the Fig. 1). On the next page clicking the button “List Values” results in a list of transgenic species (currently, 27).

For instance, to find the promoters whose activities have been verified in barley, click barley (*Hordeum vulgare*). This will result in a list of links (ID contains info on the corresponding species; e.g., Hv means *Hordeum vulgare* or Ta, *Triticum aestivum*). Check the list of promoters the expression of which has been evaluated in the specified organism.

To browse the full list of regulators, the user may click the field “Regulator” on the home page of the TGP_PROMOTER table (Fig. 1, see above). On the next page, clicking on the button “List Values” will give a current list of regulators of the TGP database. To find the promoters influenced by low temperature, click “cold” to get the list of corresponding entries.

To find the promoters that are active in a particular plant species and are influenced by a particular

Fig. 1 The home page of the TGP_PROMOTER

TGP_PROMOTER Search

Status The current release has 224 entries and was indexed 17-Jun-2011.

Description This database accumulates information on functionally active promoters described in annotated scientific sources

Data-fields in SRS

Name	Short Name	Type	No of Keys	No of Entry References	Indexing Date	Status
PromoterID	pri	id	224	224	17-Jun-2011	ok
GeneID	gen	index	114	224	17-Jun-2011	ok
Target_species	tspec	index	27	262	17-Jun-2011	ok
Keywords	kw	index	94	796	17-Jun-2011	ok
Localization	loc	show	0	0		not indexed
Description	dsc	show	0	0		not indexed
SequenceID	seq	index	224	224	17-Jun-2011	ok
Reporter	rep	index	7	238	17-Jun-2011	ok
Transgene	transgene	index	4	7	17-Jun-2011	ok
Stage_organ_tissue	stg	index	50	426	17-Jun-2011	ok
Regulator	regulator	index	58	338	17-Jun-2011	ok
Comment	cmt	index	2597	18609	17-Jun-2011	ok
Reference	ref	index	1324	5890	17-Jun-2011	ok
Pubmed	pmb	index	113	254	17-Jun-2011	ok

Links

From Databank	Entries Linked	To Databank	Entries Linked	Links Total	Indexing Date
TGP_GENE	114	TGP_PROMOTER	224	224	17-Jun-2011
TGP_PROMOTER	224	TGP_GENE	114	224	17-Jun-2011
TGP_PROMOTER	224	TGP_SEQUENCE	224	224	17-Jun-2011

regulator, the user has to click the button “Search” on the home page of the TGP_PROMOTER table (Fig. 1). This will result in a Standard query form for the TGP_PROMOTER table (Fig. 2).

From the drop-down menu “combine searches with”, select AND (marked by arrow 1 in Fig. 2). Select the field “Target_species” from the drop-down menu (arrow 2). In the text box, type the species name (e.g., tobacco, arrow 3). Next, select the field “Regulator” from the drop-down menu (arrow 2) and type the corresponding term in the text box (e.g., cold, arrow 3). Click the button “Submit Query” (arrow 4). This will give the list of promoters active in tobacco and influenced by cold.

TGP was constructed as a tool for selection of candidate promoters with appropriate characteristics. Thus, we did not try to make a full annotation of various experimental data concerning the promoters described. The main idea of TGP was to provide the link between a DNA segment and its ability to

direct transcription of a reporter gene (level of expression, specificity, and species of transgenic plants). Thus, the deletion variants of full-length promoters were also annotated, since they could provide a transcription pattern distinct from their full-length variant (and sometimes more interesting for gene engineering). For example, a twofold induction was observed in the case of the –608 bp pea TOP2 promoter (Ps:TOP2_P1) after salicylic acid treatment and more than a threefold induction was observed in the case of the –468 bp TOP2 promoter (Ps:TOP2_P2) with the same treatment (Hettiarachchi et al. 2005). This provides an opportunity to select which level of induction is more suitable to solve a particular experimental problem. The experimentally measured activities of promoters with various lengths compiled in TGP are a useful supplement to the data on the promoter structure and can be used for basic research in molecular and computer biology. The information on promoters is

Fig. 2 Standard SRS query form for the TGP_PROMOTER

The screenshot shows the SRS query interface. At the top, there are navigation tabs: TOP PAGE, QUERY, RESULTS, PROJECTS, VIEWS, DATABANKS, and HELP. Below this is a search bar containing 'TGP_PROMOTER' and a 'Reset' button. To the right of the search bar is an 'Info' button and a dropdown menu for 'PromoterID'. Below the search bar is a 'Submit Query' button. To the left of the main form is a sidebar with options: 'append wildcards to words' (checked), 'combine searches with' (set to AND), and 'Number of entries to display per page' (set to 30). There is also an 'Extended query form' button. The main form area has a header 'separate multiple values by & (and), | (or), ! (and not)'. It contains several input fields: 'Target_species' (tobacco), 'Regulator' (cold), and two 'PromoterID' fields. Below these fields is a 'Use view' dropdown set to '* Names only *' and a 'Create your own view' section. A 'Select fields to display:' list includes PromoterID, GeneID, Target_species, Keywords, Localization, Description, and SequenceID. Red circles and arrows are overlaid on the image: circle 1 points to the 'Extended query form' button; circle 2 points to the 'Regulator' field; circle 3 points to the first 'PromoterID' field; and circle 4 points to the search input field.

updated on a regular basis from the information contained in published scientific papers.

Acknowledgments We thank D. Grigorovich and D. Rasskazov for their help in database support. This work was supported by Russian Ministry of Science & Education (2.1.1/10551 (2.1.1/6382); 02.740.11.0705), and the Program of Russian Academy of Sciences (Biodiversity, 26.22). The authors are also grateful to SB RAS Complex Integration Program (28) and RFBR (grant 10-04-90411) for partial support.

References

- Baker SS, Wilhelm KS, Thomashow MF (1994) The 5'-region of *Arabidopsis thaliana* cor15a has cis-acting elements that confer cold-, drought- and ABA-regulated gene expression. *Plant Mol Biol* 24:701–713
- Bate N, Twell D (1998) Functional architecture of a late pollen promoter: pollen-specific transcription is developmentally regulated by multiple stage-specific and co-dependent activator elements. *Plant Mol Bio* 37:859–869
- Bülow L, Brill Y, Hehl R (2010) AthaMap-assisted transcription factor target gene identification in *Arabidopsis thaliana*. *Database* (Oxford) 2010: baq034
- Chang WC, Lee TY, Huang HD, Huang HY, Pan RL (2008) PlantPAN: plant promoter analysis navigator, for identifying combinatorial cis-regulatory elements with distance constraint in plant gene groups. *BMC Genomics* 9:561
- Eskelin K, Ritala A, Suntio T, Blumer S, Holkeri H, Wahlström EH, Baez J, Mäkinen K, Maria NA (2009) Production of a recombinant full-length collagen type I alpha-1 and of a 45-kDa collagen type I alpha-1 fragment in barley seeds. *Plant Biotechnol J* 7:657–672
- Furtado A, Henry RJ, Takaiwa F (2008) Comparison of promoters in transgenic rice. *Plant Biotechnol J* 6:679–693
- Furtado A, Henry RJ, Pellegrineschi A (2009) Analysis of promoters in transgenic barley and wheat. *Plant Biotechnol J* 7:240–253
- Hettiarachchi GH, Reddy MK, Sopory SK, Chattopadhyay S (2005) Regulation of TOP2 by various abiotic stresses including cold and salinity in pea and transgenic tobacco plants. *Plant Cell Physiol* 46:1154–1160
- Hettiarachchi GH, Yadav V, Reddy MK, Chattopadhyay S, Sopory SK (2003) Light-mediated regulation defines a minimal promoter region of TOP2. *Nucleic Acids Res* 31:5256–5265
- Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* 27:297–300
- Jones HD, Sparks CA (2009) Promoter sequences for defining transgene expression. *Methods Mol Biol* 478:171–184
- Kim TW, Goo YM, Lee CH, Lee BH, Bae JM, Lee SW (2009) The sweet potato ADP-glucose pyrophosphorylase gene (ibAGP1) promoter confers high-level expression of the GUS reporter gene in the potato tuber. *C R Biol* 332:876–885
- Morris RT, O'Connor TR, Wyrick JJ (2008) Osiris: an integrated promoter database for *Oryza sativa* L. *Bioinformatics* 24:2915–2917

- O'Connor TR, Dyreson C, Wyrick JJ (2005) Athena: a resource for rapid visualization and systematic analysis of *Arabidopsis* promoter sequences. *Bioinformatics* 21:4411–4413
- Peremarti A, Twyman RM, Gyme-Galera S, Naqvi S, Farrı G, Sabalza M, Miralpeix B, Dashevskaya S, Yuan D, Ramessar K, Christou P, Zhu C, Bassie L, Capell T (2010) Promoter diversity in multigene transformation. *Plant Mol Biol* 73:363–378
- Potenza C, Aleman L, Sengupta-Gopalan C (2004) Targeting transgene expression in research, agricultural, and environmental applications: promoters used in plant transformation. *In Vitro Cell Dev Biol Plant* 40:1–22
- Qu le Q, Xing YP, Liu WX, Xu XP, Song YR (2008) Expression pattern and activity of six glutelin gene promoters in transgenic rice. *J Exp Bot* 59:2417–2424
- Qu LQ, Takaiwa F (2004) Evaluation of tissue specificity and expression strength of rice seed component gene promoters in transgenic rice. *Plant Biotechnol J* 2:113–125
- Reddy MK, Nair S, Tewari KK, Mudgil Y, Yadav BS, Sopory SK (1999) Cloning and characterization of a cDNA encoding topoisomerase II in pea and analysis of its expression in relation to cell proliferation. *Plant Mol Biol* 41:125–137
- Shahmuradov IA, Gammerman AJ, Hancock JM, Bramley PM, Solovyev VV (2003) PlantProm: a database of plant promoter sequences. *Nucleic Acids Res* 31:114–117
- Tittarelli A, Santiago M, Morales A, Meisel LA, Silva H (2009) Isolation and functional characterization of cold-regulated promoters, by digitally identifying peach fruit cold-induced genes from a large EST dataset. *BMC Plant Biol* 9:121
- Tiwari S, Spielman M, Day RC, Scott RJ (2006) Proliferative phase endosperm promoters from *Arabidopsis thaliana*. *Plant Biotechnol J* 4:393–407
- Vickers CE, Xue G, Gresshoff PM (2006) A novel cis-acting element, ESP, contributes to high-level endosperm-specific expression in an oat globulin promoter. *Plant Mol Biol* 62:195–214
- Xiao YL, Redman JC, Monaghan EL, Zhuang J, Underwood BA, Moskal WA, Wang W, Wu HC, Town CD (2010) High throughput generation of promoter reporter (GFP) transgenic lines of low expressing genes in *Arabidopsis* and analysis of their expression patterns. *Plant Methods* 6:18
- Yamaguchi-Shinozaki K, Shinozaki K (1993) Characterization of the expression of a desiccation-responsive rd29 gene of *Arabidopsis thaliana* and analysis of its promoter in transgenic plants. *Mol Gen Genet* 236:331–340
- Yamamoto YY, Obokata J (2008) ppdb: a plant promoter database. *Nucleic Acids Res* 36:D977–D981
- Yang W, Liu XD, Chi XJ, Wu CA, Li YZ, Song LL, Liu XM, Wang YF, Wang FW, Zhang C, Liu Y, Zong JM, Li HY (2011) Dwarf apple MbDREB1 enhances plant tolerance to low temperature, drought, and salt stress via both ABA-dependent and ABA-independent pathways. *Planta* 233: 219–229
- Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E (2011) AGRIS: the arabidopsis gene regulatory information server, an update. *Nucleic Acids Res* 39(Database issue):D1118–D1122