

Plant Promoters: An Approach of Structure and Function

Milena Silva Porto · Morganna Pollyne Nóbrega Pinheiro ·
Vandré Guevara Lyra Batista · Roseane Cavalcanti dos Santos ·
Péricles de Albuquerque Melo Filho · Liziane Maria de Lima

Published online: 12 October 2013
© Springer Science+Business Media New York 2013

Abstract With current advances in genomics, several technological processes have been generated, resulting in improvement in different segments of molecular research involving prokaryotic and eukaryotic systems. A widely used contribution is the identification of new genes and their functions, which has led to the elucidation of several issues concerning cell regulation and interactions. For this, increase in the knowledge generated from the identification of promoters becomes considerably relevant, especially considering that to generate new technological processes, such as genetically modified organisms, the availability of promoters that regulate the expression of new genes is still

limited. Considering that this issue is essential for biotechnologists, this paper presents an updated review of promoters, from their structure to expression, and focuses on the knowledge already available in eukaryotic systems. Information on current promoters and methodologies available for studying their expression are also reported.

Keywords Gene expression · *cis*-elements · Enhancers · Silencers · TATA box · CAAT box

Introduction

With the application of molecular biology in the development of advanced biotechnology techniques, genetic engineering has become an increasingly important tool in obtaining genetically modified plants, thereby strengthening the agricultural sector [1, 2]. Several research companies worldwide have greatly invested in transgenesis, a technique that addresses certain agricultural problems that are difficult to resolve using conventional breeding methods.

The availability of full genome sequences in GenBank has led to the elucidation of the function and regulation of several genes at the molecular level. However, application of these genes in transgenesis requires prior knowledge of their expression sites and levels.

The expression of a heterologous protein depends on the transcription induction of an introduced gene. Therefore, identifying the regulatory elements, i.e., the regulatory regions or promoters, is important. The structural analysis of promoters has wide implications in the scientific field because it allows the prediction the expression profiles and locations of genes in plants [3].

Recognition of plant promoters often involves the identification and characterization of genes expressed in

M. S. Porto · V. G. L. Batista
Departamento de Pós-Graduação em Ciências Agrárias,
Universidade Estadual da Paraíba, Campina Grande, PB, Brazil
e-mail: milenasporto@gmail.com

V. G. L. Batista
e-mail: vanguevara@gmail.com

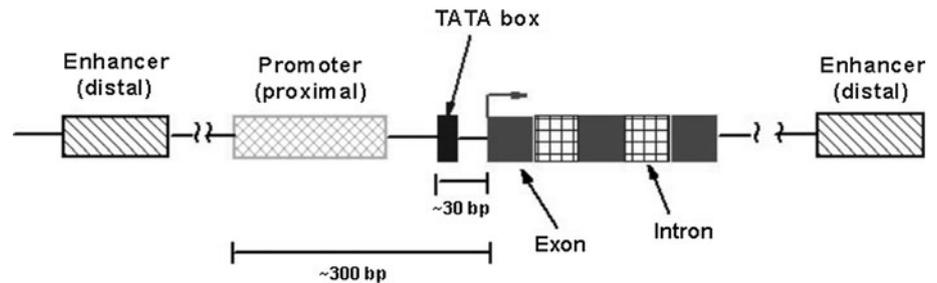
M. P. N. Pinheiro
Departamento de Pós-Graduação em Biotecnologia, Rede
Nordeste de Biotecnologia, Universidade Federal Rural de
Pernambuco, Recife, PE, Brazil
e-mail: morgannapollyne@yahoo.com.br

R. C. dos Santos · L. M. de Lima (✉)
Embrapa Algodão/Departamento de Pós-Graduação em Ciências
Agrárias, Universidade Estadual da Paraíba, Campina Grande,
PB, Brazil
e-mail: liziane.lima@embrapa.br

R. C. dos Santos
e-mail: roseane.santos@embrapa.br

P. de Albuquerque Melo Filho
Departamento de Agronomia, Universidade Federal Rural de
Pernambuco, Recife, PE, Brazil
e-mail: pericles@depa.ufrpe.br

Fig. 1 Common structure of a eukaryotic gene and transcription control regions. Source: Adapted by Klug and Cummings [12]



specific tissues or under physiological stress conditions. For the functional characterization of gene expression, isolation and structural analysis of the region upstream of a gene is required. After the promoter region is isolated, the regions that activate (i.e., enhancers) or repress (i.e., silencers) gene expression are identified by deleting the promoter region [4].

After a promoter is structurally characterized, gene constructs that contain only the regions of interest that are fused to a target gene can be obtained for further use in plant transformation.

Structure of Promoters in Eukaryotic Organisms

By definition, a promoter is the central processor of gene regulation, comprising the 5' region of the transcribed sequence located upstream from the transcription start site (TSS) of a gene, because it contains the binding sites for the protein complexes of RNA polymerases that are required for gene transcription [5, 6]. In eukaryotes, RNA polymerase I synthesizes the precursor of rRNA, RNA polymerase II produces the precursors for mRNA, and RNA polymerase III synthesizes the precursors of tRNAs, 5S-rRNAs and other miRNAs [7]. In plants, there are also RNA polymerase IV and V, which acting in the synthesis of miRNAs involved in gene silencing [8].

Structurally, a promoter is divided into proximal and distal regions (Fig. 1). The proximal region comprises the region adjacent to the TSS and approximately covers –250 to +250 nucleotides [9]. The least continuous sequence region of the DNA that is necessary to correctly guide the initiation of transcription by cell machinery is called the core promoter, which includes the TSS with –35 to +35 nucleotides [9]. This region usually contains a conserved sequence (e.g., T/A or A/T), which is located at ~25–30 base pairs (bp) from the TSS, called the TATA box. The proximal promoter elements called *cis*-elements are located ~100 (CCAAT-box) and 200 bp (GC-box) above the TSS [5]. Other elements such as the initiator (Inr), the element recognized by transcription factor (TF) IIB, B recognition element (BRE) and downstream promoter element (DPE) are generally conserved (A/G) G (A/T) CGTG). However,

these regions are not necessarily conserved since there are no universal elements in the core promoter [9]. The DPE is conserved in several eukaryotic organisms and is located at approximately 30 nucleotides downstream of the TSS of many TATA-less promoters [10], acting in conjunction with TSS to provide a binding site for TFIID. Studies carried out in transgenics *Drosophila* embryos support the hypothesis that there are three main recognition sites: TATA box, initiation site sequence/Inr and DPE, thus highlighting the relevance of this element in promoter activity [10, 11].

The TATA box was the first element identified in eukaryotes [13, 14]. It is the only element that has a relatively fixed location in relation to the TSS. The consensus sequence is approximately 8 bp and is entirely composed of A and T (the position and orientation of each nucleotide varies). The TATA box is often surrounded by sequences rich in G and C [15]. This element is similar to the –10 region (Pribnow box) found in prokaryotic promoters [16], although some authors believe that they are not homologous [9]. However, with the advancement in genomic studies this view has changed, and many genes have been found to lack the TATA box [17, 18].

In promoters with the TATA box, the factors involved in transcription (i.e., transcription factors or TFs) are formed by proteins that bind to the DNA of eukaryotic cells to facilitate the binding between the RNA polymerase and DNA, which can then bind the preinitiation complex (PIC) as follows: TFIID, TFIIB, RNA polymerase II-TFIIF complex, TFIIE, and TFIIH. TFIID binds to the core promoter, TFIIB binds to the TFIID/TFIIA complex and recruits RNA polymerase II to the promoter, TFIIF tightly binds to the RNA polymerase II, and TFIIE and TFIIH bind to the polymerase/promoter complex. The TATA box is recognized by the TATA-binding protein (TBP), which is a subunit of TFIID, whereas TFIIB is a single polypeptide that interacts with TBP as well as the DNA upstream of the TATA box. Thus, these two factors might have an important role in the recognition of the core promoter elements [9].

In a comparative study of the promoter structure between plant and mammalian, using LDSS profiles, Yamamoto et al. [19] found that the TATA box element was most conserved among all identified elements in both organisms.

Another frequent *cis*-element is the CAAT box, named for its consensus sequence. It is usually located close to -80 bp, which might vary considerably in distance from the TSS, and acts in two directions. It is the major determinant of promoter efficiency due to its susceptibility to mutations [15]. Perhaps, it has no direct relationship with promoter specificity, but it is believed to determine the efficiency of transcription [15, 20, 21].

The GC box is a common element recognized by the transcription factor Sp1, which may be located at -90 bp or at varying distances from the TSS. It is represented by the sequence GGGCGG and often occurs in multiple copies in two orientations and can increase the activity of the promoter [15, 22]. Studies involving the Sp1 revealed that it is present in mammalian promoters; however, it is not associated with *Arabidopsis* and *Oryza* promoters, suggesting that this transcription factor is not linked to all plant promoters [19].

The Inr element, also present in the core promoter of some genes, usually covers the TSS [23]. It is found in promoters irrespective of whether the TATA box is present or not. It functionally acts as the TATA box once it is initially recognized by the TFIID, which might compensate for the absence of this element in promoters without the TATA box. However, when the Inr and TATA box are present on the same promoter, they work together [24]. In *Drosophila*, the consensus sequence is T-C-A₊₁-G/T-T-C/T [25–28], where A₊₁ refers to a position that is +1 bp from the start of transcription (i.e., the point where transcription is often initiated) [9]. In rice, *Arabidopsis* and mammals, a dimer motif named the YR Rule (C/T A/G) has been identified at the transcription start site (R +1), but there are moderate variations in the utilized sequences [19]. In mammalian promoters, the Inr consensus sequence is Py Py A₊₁ N T/A Py Py (where A₊₁ is the TSS) [29].

The BRE is located upstream of some TATA boxes. Studies of *in vitro* transcription analysis showed that BRE facilitates the incorporation of TFIIB into the transcription-start complexes.

The consensus sequence is G/C-G/C-G/A-C-G-C-C (where C, at 3' of BRE is followed by T at the 5' end of the TATA box) [30]. Additionally, a new TFIIB recognition site was found (i.e., BREd, downstream of the TATA box). This finding led to the renaming of the original BRE to BREu because it is located upstream from the TATA box [31]. Both BREu and BREd work together with the TATA box [30, 31]. In humans, studies have shown the absence of TATA box and presence of BRE element at $-81/-75$ bp of TSS [32]. However, this element is not present in plant and yeast promoters [24].

The element DPE is present in promoters without the TATA box. It is exactly located at +28 to +32 bp in relation to position A₊₁ of the Inr element. If any change occurs in 1 nucleotide between Inr and DPE, the TFIID

binding affinity and basal transcriptional activity will decrease [28, 33], and 1 DPE was identified as a posterior site of the core promoter to TFIID in *Drosophila*. The consensus sequence is A/G₊₂₈-G-A/T-C/T-G/A/C, and there is a lower preference for G at +24 bp [28]. DPE-dependent promoters typically contain only DPE and Inr motifs. However, in some cases, the TATA, Inr, and DPE motifs can be found in the same promoter [23]. In plants, DPE is functionally important in a range of promoters involved in different stimuli. It is present in different positions and in multiple copies upstream of the TSS [34]. The differences in the structure of the animals and plants promoters might be related to changes in the regulatory mechanisms as well as to composition of this region. In plants, for example, the upstream region is AT-rich (72.5 %), compared with mammals (52.5 %) [35, 36].

According to the PlantCARE database, other *cis*-elements are present in promoters but are not common, as mentioned above. Representative examples include the L and I boxes, which have TCTACCAACC and CTCTTATGCT as their consensus sequences at +58 to +68 bp and +95 to +103 bp from the TSS positions, respectively; ARE, an essential element for the induction of anaerobic respiration, with TGGTTT at -700 to -705 bp; 5'-UTR Py-rich stretch that ensures high levels of transcription, located at -533 to -547 bp, with a sequence of TTTCTCTCTCTC; GARE motif that responds to gibberellins, with sequence and localization of AACAGA and +139 to +145 bp, respectively; GT1 motif that is responsive to light, with sequence and localization of ATGGTGGTTGG and +168 to +178 bp, respectively; and TC-rich repeats, which are involved in the defense mechanism and respond to stressors, are characterized by an ATTTTCTCCA sequence located at +17 to +26 bp. Other *cis*-elements are shown in Table 1 (PlantCARE).

In some cases, eukaryotic promoters do not necessarily act alone (i.e., the transcription rates of a gene are significantly increased or decreased by elements that are located at a range of distances from the elements considered as a part of the core promoter). These elements include the distal part of the promoter and regulate the spatial and temporal expressions of a gene, so that proteins acting on these elements are combined for expression to occur only at required sites and within a specific time frame. They are known as the activation (enhancer; Fig. 2) and repressor (silencer) regions [37–39].

Unlike the TATA box, the enhancers do not need to have fixed positions in relation to the promoter, and thus they can operate in two directions [40]. They are approximately 100–200 bp and may be located hundreds or thousands of bp from the TSS, both upstream and downstream or in the introns [7, 41]. An enhancer needs to be identified along with its core promoter to verify its

Table 1 Description of some classes of transcription factors found in different plant species

Function	Motif	Position	Sequence
Core promoter	TATA Box	−33	TCCCTATAAATAA
	CAT Box	−49	GCCAAC
	CAAT Box	−80	GGCCAATCT
	G Box	−66	TGACGGTGT
Response to stress	ABRE	−76	TGGTTT
	AB14	−245	CACCG
Part of a light-responsive element	I Box	+95 to +103	CTCTTATGCT
	L box	+58 to +68	TCTCACCAACC
Defense against pathogen	W Box	−72	CTTCTTTGACGTGTCCA
Essential for the anaerobic induction	ARE	−700 to −705	TGGTTT
Confer high transcription levels	5' UTR Py-rich stretch	−533 to −547	TTTCTCTCTCTCTC
Gibberellin-responsive element	GARE-motif	+139 to +145	AAACAGA
Light-responsive element	GT1-motif	+168 to +178	ATGGTGGTTGG
Responsive to defense and stress	TC-rich repeats	+17 to +26	ATTTTCTCCA
MYB recognition site found in promoters of the dehydration	MBS	−425 to −430	CGGTCA
Responsive to heat stress	HSE	+79 to +87	AAAAAATTTTC
Involved in the MeJA-responsiveness	CGTCA-motif	+1 to +5	CGTCA
Required for endosperm expression	Skn-1_motif	−624 to 628	GTCAT
Responsible to vascular expression in xylem	AC-I	+58 to +68	TCTCACCAACC
Involved in response to salicylic acid	TCA-element	−101 to −110	CAGAAAAGGA

Source: PlantCARE (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)

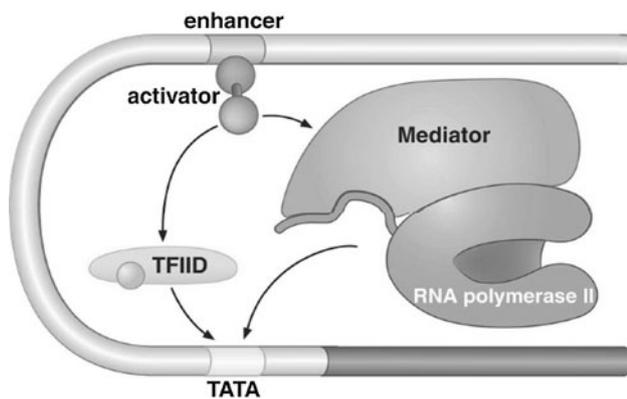


Fig. 2 Activation of transcription start in eukaryotes by recruitment of transcription machinery. Source: Watson et al. [42]

specificity and function in transcription. The presence of each element in a core promoter should be investigated separately as well as jointly [9].

Types of Promoters that Regulate Gene Expression

Promoters are classified according to their constitutive activity, which promotes gene expression throughout plant

phenology in all tissues, with low-intensity and uniform transcriptional activities in certain organs. Specifically, some promoters have elements that determine the intensity of transcription taking in account the tissue, physiological condition, age, and abiotic and biotic factors. These promoters are guided by specific transcription factors [43, 44].

Constitutive Promoters

Genes under the control of constitutive promoters are active in most cells throughout development, although their expression levels depend on the cell type [45]. In vegetal transgenesis, the Cauliflower mosaic virus promoter (CaMV35S) is highly used because it is active in most tissues and throughout the developmental stages of plants [46, 47]. Moreover, its subdomains are widely characterized, thus allowing for modulation of their temporal and spatial activities [48, 49]. Most commercial transgenic cultivars contain the constitutive 35S promoter (e.g., soybean, rapeseed, corn, and cotton Roundup Ready from Monsanto; corn, cotton, and soybean BT from Monsanto; corn (Aventis); and tomato from DNA Plant Technology). Although it is widely used in commercial cultivars, CaMV35S expression is often low in reproductive tissues, which limits its

use in cases, for example, where these structures require protection against insect or pathogen attacks.

To overcome this limitation, in recent years, researchers have conducted various studies to isolate and characterize constitutive promoters of plants, such as the promoters of polyubiquitin genes in rice [50, 51]; the APX, SCP1, PGD1, R1G1B, and EIF5 rice promoters [45, 46]; the *ibAGP1* potato promoter [52]; the promoter of the beta-carotene hydroxylase gene in *Arabidopsis* [53]; the MtHP *Medicago truncatula* promoter [54]; and the promoter of the H2B histone gene in corn [55]. Miranda et al. [56] identified a strong promoter in soybean (i.e., UceS8.3), which was able to direct a greater expression of the *GUS* gene in different tissues such as roots, stems, leaves, and flower buds. These are the examples of constitutive promoters that have been studied.

Tissue-specific Promoters

Understanding the functions of promoters acting in specific organs is essential for understanding the molecular mechanisms involved in gene expression and tissue differentiation. This produces advances in biotechnological processes because the availability of such promoters facilitates the acquisition of constructions that allow the expression of target genes in specific tissues where these promoters are active [57, 58].

Several research studies have identified potentially novel promoters that are expressed in specific tissues in plant species. Herein, we summarize some of the results for such promoters.

- Corn (*Zea mays*): Chen et al. [59] isolated and characterized the promoter of a small subunit of *AGPase* (ADP-glucose pyrophosphorylase) from the endosperm; this promoter regulates expression only in tobacco seeds
- Rice (*Oryza sativa* L. ssp.): Ye et al. [60] isolated and characterized the promoter PDX1, which is expressed only in green tissues. In this study, the authors characterized 2 new *cis*-regulatory elements, GSE1 and GSE2. GSE1 acts as a positive regulator in green tissues (e.g., leaves, sheaths, and panicle stems). GSE2 acts as a regulator only in the sheaths and stem tissues and has a weaker effect on gene expression. Thilmony and Cook [61] characterized the promoter OsGEX2, which is expressed in the sperm cells of mature rice pollen
- Leucaena (*Leucaena leucocephala*): Prashant et al. [62] isolated 2 specific promoters, LICCR and LICAD, that act in the vascular tissues of leaves, roots, and stamens
- Soybean (*Glycine max*): In the literature, several studies have shown specific promoters isolated from soybean; some of them have been used in transgenic research. Freitas [63] isolated the promoter SPB2 (sucrose-binding protein), which guides the expression of reporter genes in the vascular tissue of transgenic tobacco. Santana [64] isolated root- and leaf-specific promoters PsulfT0,5; PCit0,8; and PCit1,9. PCit0, which is expressed only in leaves, was also isolated. Other promoters that are expressed only in roots have also been described, such as isoflavone synthase (IFS), which is expressed in events associated with hormone production and nodulation [65]
- Coffee (*Coffea arabica*): Marraccini et al. [66] isolated the promoter of 11S (*csp1*), which encodes storage proteins in seeds. Brandalise et al. [67] identified a potential promoter of a gene, *CaIRL*, similar to isoflavone reductase, where the *GUS* gene confirmed the responsiveness of the putative promoter to abiotic stress in wounded leaves. These promoters are highly relevant to the development of transgenic plants due to the coverage they can achieve in the defense against environmental damage, thus ensuring plant health
- Sugarcane (*Saccharum* L.): Due to the broad importance of this crop on an industrial scale, the current literature offers promising results indicating potential uses for several genes and promoters in transgenic studies. Hoshino [68] isolated and characterized a root promoter, SCCCRT1004A07, which encodes a carrier lipid protein. Damaj et al. [69] reported that the promoters ProDIR16 and ProOMT are expressed in the stems, leaves, and roots and are induced by salicylic acid, jasmonic acid, and methyl jasmonate; all are key regulators of biotic and abiotic stressors. Moyle and Birch [70] analyzed the promoter ScLSG using bioinformatic and transgenic approaches and indicated the potential specificity of expression in stems. These studies will be useful for research and biotechnological advances in sugarcane, where the expression of transgenes in tailored stems is important for the accumulation of sugar to obtain value-added products and increase the use as bioenergy feedstock
- Eucalyptus (*Eucalyptus grandis*): Sasaki [71] isolated and characterized the promoter EGRT2, which is expressed in the roots and is associated with potassium transport and translocation. Costa [72] characterized the promoter EGJFLV3247C08.g that is expressed in the vascular bundles of leaves and roots. This promoter can be applied to eucalyptus transgenic studies to guide the expression of genes that confer improvements in cellulose accumulation or in defense against endogenous pathogens
- Cotton (*Gossypium hirsutum*): As one of the most important fiber crops worldwide, the findings from cotton genomic data have allowed the elucidation of

several cotton genes and promoters associated with oil production and fiber qualities; all are highly useful in biotechnological segments. Some examples are reported by Li et al. [73], who used the promoter GhACT1 to show that the *GhACT1* gene is involved in the elongation of cotton fiber. This promoter provides a powerful tool for improving fiber traits in future transgenic cultivars, since fiber is the main commodity in the cotton market. Another interesting result was reported by Sunilkumar et al. [74] who showed the potential of an antisense promoter that inhibits the activity of β -12 desaturase in seeds, resulting in the reduction of linoleic acid and increase in oleic oil. This finding is considerably relevant since high oleic acid is needed for oil stability. Lightfoot et al. [75] isolated and characterized 2 promoters, GhPRP3 and GhCHS1, which were shown to drive preferential transcription at high levels in boll wall tissues; GhPRP3 drove reporter expression in the boll wall between 5 and 25 days post-anthesis (DPA), and GhCHS1 drove the expression from 3 to 25 DPA in boll wall fiber and petal tissues

- Palm oil (*Elaeis guineensis*): Kamaladini et al. [76] isolated and identified the promoter MT3-A, which is expressed only in the reproductive organs, including anthers, fruits, and seeds

The usefulness of all these promoters, especially those related to the reproductive organs, is widely exploited in genetic engineering because grains are the main marketable products. Stored grain pests and pathogens that affect these organs lead to irreparable damage, which is often controlled using chemical pesticides that raise the cost of production. The possibility of controlling pests by developing transgenic plants opens a range of opportunities for farmers who can adopt more agroecological and safe management strategies, since most of the current transgenic species are harmless to man and the environment [59, 77, 78].

Strategies for Promoter Isolation

Promoters can be isolated by screening genomic libraries or using genome walking or thermal asymmetric interlaced polymerase chain reaction (TAIL-PCR) strategies. Previously, identifying genes with specific techniques that focused on differential expression, such as reverse-transcription polymerase chain reaction (RT-PCR, semi-quantitative PCR), real-time RT-qPCR (quantitative PCR), electronic northern, and/or northern blotting, was necessary. The sequences of the genes selected were used to isolate genomic sequences containing the promoter regions. The principle of any strategy for the isolation of promoters is to amplify regions located upstream of a

known sequence using PCR with primers complementary to the adapters and a known sequence, such as expressed sequence tags (ESTs) [79, 80]. PCR-based methods have frequently been used since they are rapid and efficient. However, the limitation of these techniques is that the restriction site is located upstream of the specific primer for the gene of interest. Therefore, different restriction enzymes are used to isolate large upstream fragments [81].

Several kits for genome walking technology are available commercially, which were developed to identify and isolate putative regulatory regions in different plant species, in addition to humans and rats [82–86]. The first step is to construct genomic libraries using the genomic DNA of interest and specific restriction enzymes. The protocol consists of a primary PCR using primers that are more externally connected to the adapter (AP1), which is supplied by the kit, and a specific primer of the target gene (GSP1). Subsequently, a secondary PCR is performed using more internal primers, the second adapter (AP2), and other, more internal-specific primers of the target gene (GSP2). In this phase, a template from the primary PCR products is used.

The criteria recommended for the primers are as follows: [1] minimum, optimum, and maximum sizes (bp) of 26, 28, and 30, respectively; [2] minimum, optimum, and maximum annealing temperatures ($^{\circ}\text{C}$) of 66, 67, and 68, respectively; and [3] minimum and maximum GC contents (%) of 40 and 60, respectively.

Another strategy for promoter isolation is TAIL-PCR, described by Liu and Whittier [87], which is simpler and faster than the genome walking techniques. The advantage of this method is the absence of the formation of chimeric artifacts because it does not involve connections to the adapters. Another advantage is the rapid confirmation of specific products since they can be readily sequenced [88].

This technique uses two primers. The first primer anneals to the initial region of the target sequence, whereas the second anneals upstream to other random or degenerate regions, allowing amplification of several fragments; some of these are associated with the desired sequence. Next, they are reamplified in an additional two selective reactions using a diluted product of the first reaction, a random primer, and two internal-specific primers. Terauchi and Kahl [89] optimized this technique using three specific gene primers that were positioned consecutively for the third amplification reaction, thus ensuring higher selectivity.

Considering that several genes might exhibit constitutive expression, regulatory sequences can be made more specific by performing deletions of these sequences via PCR using specific primers containing the restriction sites [4]. The products of these deletions might be cloned into

Table 2 Database available for identification of regulatory sequences

Database	Web site	Reference
CONREAL	http://conreal.niob.knaw.nl/description.html	[93]
ConSite	http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite	[94]
JASPAR	http://jaspar.genereg.net	[95]
JPREDictor	http://bibiserv.techfak.uni-bielefeld.de/jpredictor	[96]
MAPPER	http://mapper.chip.org	[97]
MatInspector	http://www.genomatix.de/online_help/help_matinspector/matinspector_help.html	[98]
PHYLONET	http://stormo.wustl.edu/PhyloNet	[99]
PLACE	http://www.dna.affrc.go.jp/PLACE/signalscan.html	[91]
PlantCARE	http://bioinformatics.psb.ugent.be/webtools/plantcare/html	[92]
RSAT ^a	http://rsat.ulb.ac.be	[100]
RegulonDB	http://www.ccg.unam.mx/en/projects/collado/regulondb	[101]
Seqmotifs	http://seqmotifs.stanford.edu	[102]
TRRD ^b	http://www.mgs.bionet.nsc.ru/mgs/gnw/trrd/	[103]
TRANSFAC	http://www.gene-regulation.com/pub/databases.html	[104]

^a Regulatory Sequence Analysis Tools

^b Transcription regulatory region database

commercial vectors and further sequenced for in silico analysis.

Analysis of Promoters

Promoter sequences can be analyzed using several computational tools currently available, allowing the identification of putative TFs, which are analyzed by alignment using databases available online and further characterization in silico [60, 62]. Some database sites are shown in Table 2. Among them, PLACE [90, 91] and/or PlantCARE are considerably useful [92].

These platforms contain several identified and characterized *cis*-elements, a brief description, and links to publications. Additionally, the platform PlantCARE generates a dynamic page where *cis*-elements are highlighted in a sequence, thus facilitating their location. In addition, a detailed characterization of these elements and a similarity matrix of the sequences are provided [84, 105–107]. The sequence results can then be used for subcloning into expression vectors and further transformation into model plants to estimate the expression level of the promoter isolated.

Procedures for Estimating the Expression of Promoters

First, conducting plant transformation using any method available for the model plant is necessary. Next, a reliable procedure should be adapted to estimate the expression levels, taking in account the costs related to components such as ease of processing and availability of reagents. Herein, we describe some procedures available to estimate the expression of promoters.

Histochemical Assay

The histochemical assay is a qualitative method based on the cleavage of the substrate 5-bromo-4-chloro-3-indolyl- β -D-glucuronide (X-gluc) by β -glucuronidase, generating a blue-colored precipitate that allows the identification of cells expressing the recombinant protein [108]. The *GUS* gene, which encodes β -glucuronidase, was isolated from *Escherichia coli* and is a reporter widely used in gene expression tracking, particularly with regard to specific promoter sequences. Moreover, it does not exhibit expressive, endogenous activity in most plants [109–111].

This assay detects transient expression, which can be detected in protoplasts or in specific tissues and is essential for studies associated with subcellular localization and interactions with other biomolecules. The transient expression assays are advantageous because gene expression is estimated over a short period and does not depend on the regeneration of transformed cells. This allows the establishment of cells whose calli are recalcitrant to regeneration.

Van Bostel et al. [112] showed that the *GUS* gene showed transient expression under the control of different promoters by bombarding various tissues of coffee (*C. arabica*). Transient expression was also observed in transgenic potato plants (*Solanum tuberosum*). Torres et al. [113] used the method of co-culture with *A. tumefaciens* and found that the transient expression and stability of the *GUS* marker gene driven by the promoter rol A (pBRA3) were tissue-specific, localized mainly in the vascular system of internodes, and absent in the roots and leaves.

In order to confirm the functionality of the promoter region of the candidate EGLV1 in eucalyptus (*E. grandis*), Sasaki [71] performed transient expression assays. The

results showed that *GUS* activity could be detected in the cotyledons of eucalyptus seedlings infected with *Agrobacterium tumefaciens* and transformed with the expression cassette *EGLVI:GUS*.

Southern Blot, Western Blot, and Enzyme-linked Immunosorbent Assay

The Southern blot technique is based on the detection of specific DNA fragments in samples of complex compositions, such as genomic DNA. This methodology was first described by Edwin M. Southern in 1975. It allows researchers to analyze exogenous DNA sequences in organisms using any method of genetic transformation [110, 114].

Western blot enables detection of small amounts of protein-by-protein extract scans, characterization of recombinant polypeptides, detection of protein degradation products, etc. The detection is performed using antibodies reacting specifically with the epitopes of the protein of interest and is accompanied by radiographic or colorimetric reactions [115–118].

The enzyme-linked immunosorbent assay (ELISA) technique involves the use of non-radioactive reagents and enzyme reagents. It identifies a protein from among a population of other proteins. This method can be used to detect a protein encoded by a foreign gene in a transgenic plant. Antibodies covalently combined to different enzymes, such as peroxidase, alkaline phosphatase, β -galactosidase, urease, penicillinase, and glucoamylase, have been successfully used in ELISA assays [119].

All the techniques mentioned above can be used to study the actions of a promoter in driving a gene's constitutive expression or its actions in specific tissues.

Final Remarks

In recent years, the application of transgenic techniques has led to improvement in several crop species because of the identification of a large number of genes. Because most studies focus on the expression of genes in specific organs, molecular researchers have made efforts to isolate tissue-specific promoters to add value to transgenes.

Because of the importance of transcriptional regulation, a major goal in the post-genomic era is the understanding of the functions of TFs in the promoter regions. This will allow the construction of a transcriptional network model that might provide a basis for the analysis of regulatory sequences. Because of the databases that are currently available, analyzing, identifying, and characterizing promoters from different species have become possible, thus facilitating molecular advancement in the field of bioinformatics.

With the progress that has been achieved in the agricultural sector through current biotechnological techniques, the availability of new tissue- and stage-specific promoters might contribute to the advancement of transcription machinery in new GM cultivars, considering that this process is frequently required in cells. Despite the indisputable economic advantage of current GM crops, which contain relevant genes controlled by constitutive promoters, to farmers, additional efforts are required to transcribe an exogenous gene, which often results in the reduction of transgene expression throughout the crop cycle [120–122]. Gene banks often receive hundreds of thousands of new deposits of genes, some of which have unknown functions. Opportunities to invest in identifying new promoters are broad, and investments will be offset by biotechnological advances and the applicability of AT-rich genomics for the benefit of agricultural growth worldwide.

References

1. Borém, A. (1998). *Melhoramento de plantas*. Viçosa: UFV.
2. Mittler, R., & Blumwald, E. (2010). Genetic engineering for modern agriculture: challenges and perspectives. *Annual Review of Plant Biology*, 61, 443–462.
3. Shelenkov, A., & Korotkov, E. (2009). Search of regular sequences in promoters from eukaryotic genomes. *Computational Biology and Chemistry*, 33, 196–204.
4. Cai, M., Wei, J., Li, X., Xu, C., & Wang, S. (2007). A rice promoter containing both novel positive and negative *cis*-elements for regulation of green tissue-specific gene expression in transgenic plants. *Plant Biotechnology Journal*, 5, 664–674.
5. Griffiths, A. J. F., Suzuki, D. T., Lewontin, R. C. & Gelbart, W. M. (2000). Regulation of gene transcription, in *Modern Genetic Analysis*, 3rd edition, (W.H. Freeman and Company), New York, USA.
6. Périer, R. C., Junier, T., & Bucher, P. (1998). The eukaryotic promoter database EPD. *Nucleic Acids Research*, 26, 353–357.
7. Lodish, H., Berk, A., Zipursky, L. S., Matsudaira, P., Baltimore, D., & Darnell, J. (2000). Regulation of transcription initiation. In W. H. Freeman (Ed.), *Molecular Cell Biology*. New York: W. H. Freeman.
8. Haag, J. R., & Pikaard, C. S. (2011). Multisubunit RNA polymerases IV and V: purveyors of noncoding RNA for plants gene silencing. *Nature Reviews Molecular Cell Biology*, 2, 483–492.
9. Butler, J. E. F., & Kadonaga, J. T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & Development*, 6, 2583–2592.
10. Burke, T. W., & Kadonaga, J. T. (1996). *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes & Development*, 10, 711–724.
11. Burke, T. W., & Kadonaga, J. T. (1997). The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII60 of *Drosophila*. *Genes & Development*, 11, 3020–3031.
12. Klug, W. S., & Cummings, M. R. (1997). *Concepts of Genetics*. New Jersey: Prentice-Hall.
13. Goldberg, M. L. (1979). Sequence analysis of *Drosophila* histone genes, PhD thesis, Stanford University, California, USA.

14. Breathnach, R., & Chambon, P. (1981). Organization and expression of eukaryotic split genes coding for proteins. *Annual Review of Biochemistry*, *50*, 349–383.
15. Lewin, B. (2001). *Genes VII*. Porto Alegre: Artmed Editora LTDA.
16. Pribnow, D. (1975). Nucleotide sequence of an RNA polymerase binding site at an early T7 promoter. *Proceedings of the National Academy of Sciences*, *72*, 784–788.
17. Basehoar, A. D., Zanton, S. J., & Pugh, B. F. (2004). Identification and distinct regulation of yeast TATA box-containing genes. *Cell*, *116*, 699–709.
18. Shi, W., & Zhou, W. (2006). Frequency distribution of TATA Box and extension sequences on human promoters. *BMC Bioinformatics*, *7*, S2.
19. Yamamoto, Y. Y., Ichida, H., Matsui, M., Obokata, J., Sakurai, T., Satou, M., et al. (2007). Identification of plant promoter constituents by analysis of local distribution of short sequences. *BMC Genomics*, *8*, 1–23.
20. Chen, L., Tu, Z., Hussain, J., Cong, L., Yan, Y., Jin, L., et al. (2010). Isolation and heterologous transformation analysis of a pollen-specific promoter from wheat (*Triticum aestivum* L.). *Molecular Biology Reports*, *37*, 737–744.
21. Tiwari, S. B., Shen, Y., Chang, H. C., Hou, Y., Harris, A., Ma, S. F., et al. (2010). The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. *New Phytologist*, *187*, 57–66.
22. Bucher, P. (1990). Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *Journal of Molecular Biology*, *212*, 563–578.
23. Juven-Gershon, T., & Kadonaga, J. T. (2010). Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Developmental Biology*, *15*, 225–229.
24. Smale, S. T., & Kadonaga, J. T. (2003). The RNA polymerase II core promoter. *Annual Review of Biochemistry*, *72*, 449–479.
25. Hultmark, D., Klemenz, R., & Gehring, W. J. (1986). Translational and transcriptional control elements in the untranslated leader of the heat-shock gene hsp22. *Cell*, *44*, 429–438.
26. Purnell, B. A., Emanuel, P. A., & Gilmour, D. S. (1994). TFIID sequence recognition of the initiator and sequences farther downstream in *Drosophila* class II genes. *Genes & Development*, *8*, 830–842.
27. Arkhipova, I. R. (1995). Promoter elements in *Drosophila melanogaster* revealed by sequence analysis. *Genetics*, *139*, 1359–1369.
28. Kutach, A. K., & Kadonaga, J. T. (2000). The downstream promoter element DPE appears to be as widely used as the TATA box in *Drosophila* core promoters. *Molecular and Cellular Biology*, *20*, 4754–4764.
29. Javahery, R., Khachi, A., Lo, K., Zenzie-Gregory, B., & Smale, S. T. (1994). DNA sequence requirements for transcriptional initiator activity in mammalian cells. *Molecular and Cellular Biology*, *14*, 116–127.
30. Lagrange, T., Kapanidis, A. N., Tang, H., Reinberg, D., & Ebright, R. (1998). New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB. *Genes & Development*, *12*, 34–44.
31. Deng, W., & Roberts, S. G. (2005). A core promoter element downstream of the TATA box that is recognized by TFIIB. *Genes & Development*, *19*, 2418–2423.
32. Iacobazzi, V., Infantino, V., & Palmieri, F. (2013). Regulação da transcrição do citrato mitocondrial e transportadores carnitina/acilcarnitina: dois genes envolvidos na biossíntese de ácidos graxos e β -oxidação. *Biology*, *2*, 284–303.
33. Kadonaga, J. T. (1998). Eukaryotic transcription: an interlaced network of transcription factors and chromatin-modifying machines. *Cell*, *92*, 307–313.
34. Sawant, S. V., Singh, P. K., Madanala, R., & Tuli, R. (2001). Designing of an artificial expression cassette for high level expression of transgenes in plants. *Theoretical and Applied Genetics*, *102*, 635–644.
35. Kanhere, A., & Bansal, M. (2005). Structural properties of promoters: similarities and differences between prokaryotes and eukaryotes. *Nucleic Acids Research*, *33*, 3165–3175.
36. Zhang, G., Lukoszek, R., Mueller-Roeber, B., & Ignatova, Z. (2011). Different sequence signatures in the upstream regions of plant and animal tRNA genes shape distinct modes of regulation. *Nucleic Acids Research*, *39*, 3331–3339.
37. Fessele, S., Maier, H., Zischek, C., Nelson, P. J., & Werner, T. (2002). Regulatory context is a crucial part of gene function. *Trends in Genetics*, *18*, 60–63.
38. Riethoven, J. J. M. (2010). Regulatory regions in DNA: promoters, enhancers, silencers, and insulators. *Methods in Molecular Biology*, *674*, 33–42.
39. Kolovos, P., Knoch, T. A., Grosveld, F. G., Cook, P. R., & Papanonis, A. (2012). Enhancers and silencers: an integrated and simple model for their function. *Epigenetics Chromatin*, *5*, 3–8.
40. Bulger, M., & Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. *Cell*, *144*, 327–339.
41. He, H. H., Meyer, C. A., Shin, H., Bailey, S. T., Wei, G., Wang, Q., et al. (2010). Nucleosome dynamics define transcriptional enhancers. *Nature Genetics*, *42*, 343–347.
42. Watson, J. D. (2008). *Molecular biology of the gene*. San Francisco: Benjamin-Cummings.
43. Fickett, J. W., & Hatzigeorgiou, A. G. (1997). Eukaryotic promoter recognition. *Genome Research*, *7*, 861–878.
44. Hochheimer, A., & Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. *Genes & Development*, *17*, 1309–1320.
45. Park, S. H., Yi, N., Kim, Y. S., Jeong, M. H., Bang, S. W., Choi, Y. D., et al. (2010). Analysis of five novel putative constitutive gene promoters in transgenic rice plants. *Journal of Experimental Botany*, *61*, 2459–2467.
46. Park, S. H., Bang, S. W., Jeong, J. S., Jung, H., Felipe, M. C., Redillas, R., et al. (2012). Analysis of the APX, PGD1 and R1G1B constitutive gene promoters in various organs over three homozygous generations of transgenic rice plants. *Planta*, *235*, 1397–1408.
47. Ranjan, R., Patro, S., Kumari, S., Kumar, D., Dey, N., & Maiti, I. B. (2011). Efficient chimeric promoters derived from full-length and sub-genomic transcript promoters of Figwort mosaic virus (FMV). *Journal of Biotechnology*, *152*, 58–62.
48. Benfey, P. N., & Chua, N. H. (1990). The Cauliflower Mosaic Virus 35S Promoter: combinatorial regulation of transcription in plants. *Science*, *250*, 959–966.
49. Kay, R., Chau, A., & Daly, M. (1987). Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. *Science*, *236*, 1299–1302.
50. Lu, J., Sivamani, E., Li, X., & Qu, R. (2008). Activity of the 5' regulatory regions of the rice polyubiquitin rubi3 gene in transgenic rice plants as analyzed by both GUS and GFP reporter genes. *Plant Cell Reports*, *27*, 1587–1600.
51. Hernandez-Garcia, C. M., Adriana, P., Martinelli, A. P., Bouchard, R. A., & Finer, J. J. (2009). A soybean (*Glycine max*) polyubiquitin promoter gives strong constitutive expression in transgenic soybean. *Plant Cell Reports*, *28*, 837–849.
52. Kwak, M. S., Oh, M. J., Lee, S. W., Shin, J. S., Paek, K. H., & Bae, J. M. (2007). A strong constitutive gene expression system

- derived from ibAGP1 promoter and its transit peptide. *Plant Cell Reports*, 26, 1253–1262.
53. Liang, Y. S., Bae, H. J., Kang, S. H., Lee, T., Kim, M. G., Kim, Y. M., et al. (2009). The Arabidopsis beta-carotene hydroxylase gene promoter for a strong constitutive expression of transgene. *Plant Biotechnol Report*, 3, 325–331.
 54. Xiao, K., Khang, C., Harrison, M., & Wang, Z. Y. (2005). Isolation and characterization of a novel plant promoter that directs strong constitutive expression of transgenes in plants. *Molecular Breeding*, 2, 221–231.
 55. Rasco, G., Rasco-Gaunt, S., Liu, D., Li, C., Doherty, A., Hagemann, K., et al. (2003). Characterization of the expression of a novel constitutive maize promoter in transgenic wheat and maize. *Plant Cell*, 6, 569–576.
 56. Miranda, V. J. (2011). Caracterização da expressão do gene codificador da enzima de conjugação a ubiquitina (E2) em soja inoculada com *Meloidoyne incognita* e infestada com *Anticarsia gemmatalis*. Dissertação, Universidade de Brasília, Brasília, DF.
 57. Ortiz, R. (1998). Critical role of plant biotechnology for the genetic improvement of food crops: perspectives for the next millennium. *Electronic Journal of Biotechnology*, 1, 1–7.
 58. Nain, V., Verma, A., Kumar, N., Sharma, P., Ramesh, B., & Kumar, A. A. (2008). Cloning of an ovule specific promoter from *Arabidopsis thaliana* and expression of B-glucuronidase. *Indian Journal of Experimental Biology*, 46, 207–211.
 59. Chen, X., Wang, Z., Wang, J., Wang, M., Zhao, L., & Wang, G. (2007). Isolation and characterization of Brittle2 promoter from *Zea mays* and its comparison with Ze19 promoter in transgenic tobacco plants. *Plant Cell, Tissue and Organ Culture*, 88, 11–20.
 60. Ye, R., Zhou, F., & Lin, Y. (2012). Two novel positive cis-regulatory elements involved in green tissue-specific promoter activity in rice (*Oryza sativa* L. ssp.). *Plant Cell*, 31, 1159–1172.
 61. Cook, M., & Thilmony, R. (2012). The OsGEX2 gene promoter confers sperm cell expression in transgenic rice. *Plant Molecular Biology Reporter*, 30, 1138–1148.
 62. Prashant, S., Sunita, M. S. L., Sirisha, V. L., Bhaskar, V. V., Rao, A. M., Narasu, M. L., et al. (2012). Isolation of cinnamoyl CoA reductase and cinnamyl alcohol dehydrogenase gene promoters from *Leucaena leucocephala*, a leguminous tree species, and characterization of tissue-specific activity in transgenic tobacco. *Plant Cell, Tissue and Organ Culture*, 108, 421–436.
 63. Freitas, R. L. (2007). Identificação de regiões no promotor do gene SBP2 (Sucrose Binding Protein) de soja que conferem expressão espacial específica. Dissertação, Universidade Federal de Viçosa, Minas Gerais, Brazil.
 64. Santana, R. H. (2012). Isolamento e caracterização de promotores órgão-específicos de plantas de soja (*Glycine max*), Dissertação, Universidade de Brasília, Brasília, DF.
 65. Subramanian, S., Hu, X., Lu, G., Odelland, J. T., & Yu, O. (2004). The promoters of two isoflavone synthase genes respond differentially to nodulation and defense signals in transgenic soybean roots. *Plant Molecular Biology*, 54, 623–639.
 66. Marraccini, P., Deshayes, A., Petiard, V., & Rogers, W. J. (1999). Molecular cloning of the complete 11S seed storage protein gene of *Coffea arabica* and promoter analysis in transgenic tobacco plants. *Plant Physiology and Biochemistry*, 37, 273–282.
 67. Brandalise, M., Severino, F. E., Maluf, M. P., & Maia, I. G. (2009). The promoter of a gene encoding an isoflavone reductase-like protein in coffee (*Coffea arabica*) drives a stress responsive expression in leaves. *Plant Cell Reports*, 28, 1699–1708.
 68. Hoshino, A. A. (2007). Isolamento e caracterização de promotores tecido-específicos a partir das informações do SUCEST (“Sugarcane expressed sequence TAGS”), thesis, Universidade Estadual Paulista, SP, BR.
 69. Damaj, M. B., Kumpatla, S. P., Emani, C., Beremand, P. D., Reddy, A. S., Rathore, K. S., et al. (2010). Sugarcane DIRIGENT and O-methyltransferase promoters confer stem-regulated gene expression in diverse monocots. *Planta*, 231, 1439–1458.
 70. Moyle, R. L., & Birch, R. G. (2013). Sugarcane loading stem Gene promoters drive transgene expression preferentially in the stem. *Plant Molecular Biology*, 82, 51–58.
 71. Sasaki, F. T. (2008). Isolamento e caracterização de promotores órgão-específicos a partir de informações do banco forests (Eucalyptus Genome Sequencing Project Consortium). Thesis, Universidade Estadual Paulista, SP, Brazil.
 72. Costa, C. S. (2011). Caracterização de promotores de eucalipto com expressão tecido-específica: raiz e folha. Dissertação, Universidade Estadual Paulista, Campus de Botucatu, SP.
 73. Li, X., Fan, X., Lanwang, X., Cai, L., & Yang, A. (2005). The cotton actin1 gene is functionally expressed in fibers and participates in fiber elongation. *Plant Cell*, 17, 859–875.
 74. Sunilkumar, G., Campbell, L. M., Hossen, M., Connell, J. P., Hernandez, E., Reddy, A. S., et al. (2005). A comprehensive study of the use of a homologous promoter in antisense cotton lines exhibiting a high seed oleic acid phenotype. *Plant Biotechnology Journal*, 3, 319–330.
 75. Lightfoot, D. J., Orford, S. J., & Timmis, J. N. (2013). Identification and characterization of cotton boll wall-specific gene promoters for future transgenic cotton varieties. *Plant Molecular Biology Reporter*, 31, 174–184.
 76. Kamaladini, H., Abdullah, N. A. S., Aziz, M. A., Ismail, I. B., & Haddadi, F. (2013). Breaking-off tissue specific activity of the oil palm metallothionein-like gene promoter in T1 seedlings of tomato exposed to metal ions. *Journal of Plant Physiology*, 170, 346–354.
 77. Daniell, H. (2002). Molecular strategies for gene containment in transgenic crops. *Nature Biotechnology*, 20, 581–586.
 78. Sisterson, M. S., Carriere, Y., Dennehy, T. J., & Tabashnik, B. E. (2007). Nontarget effects of transgenic insecticidal crops: implications of source-sink population dynamics. *Environmental Entomology*, 36, 121–127.
 79. Siebert, P. D., Chenchik, A., Kellogg, D. E., Lukyanov, K. A., & Lukyanov, S. A. (1995). An improved method for walking in uncloned genomic DNA. *Nucleic Acids Research*, 23, 1087–1088.
 80. Devic, M., Albert, S., Delseny, M., & Roscoe, T. J. (1997). Efficient PCR walking on plant genomic DNA. *Plant Physiology and Biochemistry*, 35, 1–9.
 81. Rishi, A. S., & Nelson, N. D. (2004). Arun Goyal Genome walking of large fragments: an improved method. *Journal of Biotechnology*, 111, 9–15.
 82. Almasia, N., Narhirñak, V., Hopp, H. E., & Vazquez-Rovere, C. (2010). Isolation and characterization of the tissue and development-specific potato snakin-1 promoter inducible by temperature and wounding. *Electronic Journal of Biotechnology*, 13, 1–12.
 83. Damaj, M. B., Beremand, P. D., Buenrostro-Nava, M. T., Ivy, J., Kumpatla, S. P., Jifon, J., et al. (2010). Isolating promoters of multigene family members from the polyploid sugarcane genome by PCR-based walking in BAC DNA. *Genome*, 53, 840–847.
 84. Tajrishi, M. M., & Tuteja, N. (2011). Isolation and in silico analysis of promoter of a high salinity stress-regulated pea DNA helicase 45. *Plant Signaling & Behavior*, 6, 1447–1450.
 85. Xiang, J., Liu, Z., Yang, Y., & Li, X. (2012). Cloning and analysis of the ascorbate peroxidase gene promoter from *Brassica napus*. *African Journal of Biotechnology*, 11, 6428–6433.

86. Wee, C. C., & Roslan, H. A. (2012). Isolation of alcohol dehydrogenase cDNA and basal regulatory region from *Metroxylon sagu*. *ISRN Molecular Biology*, 2012, 1–10.
87. Liu, Y. G., & Whittier, R. F. (1995). Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, 25, 674–681.
88. Wan, X. L., Yang, R. H., & Yao, Y. J. (2011). Development of microsatellite markers for *Ophiocordyceps sinensis* (Ophiocordycipitaceae) using an ISSR-TAIL-PCR method. *American Journal of Botany*, 98, 391–394.
89. Terauchi, R., & Kahl, G. (2000). Rapid isolation of promoter sequences by TAIL-PCR: the 5'-flanking regions of Pal and Pgi genes from yams (*Dioscorea*). *Molecular Genetics and Genomics*, 263, 554–560.
90. Prestridge, D. S. (1991). SIGNAL SCAN: a computer program that scans DNA sequences for eukaryotic transcriptional elements. *Bioinformatics*, 7, 203–206.
91. Higo, K., Ugawa, Y., Iwamoto, M., & Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Research*, 27, 297–300.
92. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Peer, Y. V., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, 30, 325–327.
93. Berezikov, E., Guryev, V., & Cuppen, E. (2005). CONREAL Web Server: identification and visualization of conserved transcription factor binding sites. *Nucleic Acids Research*, 33, W447–W450.
94. Sandelin, A., Wasserman, W. W., & Lenhard, B. (2004). ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Research*, 32, 249–252.
95. Portales-Casamar, E., Thongjuea, S., Kwon, A. T., Arenillas, D., Zhao, X., Valen, E., et al. (2010). JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. *Nucleic Acids Research*, 38, D105–D110.
96. Fiedler, T., & Rehmsmeier, M. (2006). jPREdictor: a versatile tool for the prediction of cis-regulatory elements. *Nucleic Acids Research*, 34, W546–W550.
97. Marinescu, V. D., Kohane, I. S., & Riva, A. (2005). MAPPER: a search engine for the computational identification of putative transcription factor binding sites in multiple genomes. *BMC Bioinformatics*, 7, 1–36.
98. Cartharius, K., Frech, K., Grote, K., Klocke, B., Haltmeier, M., Klingenhoff, A., et al. (2005). MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics*, 21, 2933–2942.
99. Wang, T., & Stormo, G. D. (2005). Identifying the conserved network of cis-regulatory sites of a eukaryotic genome. *Proceedings of the National Academy of Sciences*, 102, 17400–17405.
100. Thomas-Chollier, M., Sand, O., Turatsinze, J. V., Janky, R., Defrance, M., Vervisch, E., et al. (2008). RSAT: regulatory sequence analysis tools. *Nucleic Acids Research*, 36, W119–W127.
101. Karp, P. D., Riley, M., Saier, M., Paulsen, I. T., Collado-Vides, J., Paley, S. M., et al. (2002). The EcoCyc Database. *Nucleic Acids Research*, 30, 56–58.
102. Liu, X., Brutlag, D. L., Liu, J. S. (2001). BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. Pacific Symposium on Biocomputing pp. 127–38.
103. Kolchanov, N. A., Ignatieva, E. V., Ananko, E. A., Podkolodnaya, O. A., Stepanenko, I. L., Merkulova, T. I., et al. (2002). Transcription regulatory regions database (TRRD): its status in 2002. *Nucleic Acids Research*, 30, 312–317.
104. TRANSFAC 7.0 Public 2005 and TRANSCOMP 7.0 Public 2005. Available in: http://www.gene-regulation.com/pub/data_bases.html. Accessed 9 December 2013.
105. Jaiswal, R., Nain, V., Abdin, M. Z., & Kumar, P. A. (2007). Isolation of pigeon pea (*Cajanus cajan* L.) legumin gene promoter and identification of conserved regulatory elements using tools of bioinformatics. *Indian Journal of Experimental Biology*, 6, 495–503.
106. Lindlöf, A., Bräutigam, M., Chawade, A., Olsson, O., & Olsson, B. (2009). In silico analysis of promoter regions from cold-induced genes in rice (*Oryza sativa* L.) and *Arabidopsis thaliana* reveals the importance of combinatorial control. *Bioinformatics*, 25, 1345–1348.
107. Nain, V., Sahi, S., & Kumar, P. A. (2011). In silico identification of regulatory elements in promoters, in Computational Biology and Applied Bioinformatics, (Lopes, H., ed), InTech China.
108. Lacorte, C. (1998). β -glucuronidase (GUS), in manual de transformação genética de plantas, (Brasileiro, A. C. M. Carneiro, V. T. D., ed), Brasília, Embrapa: pp. 127–141.
109. Jefferson, R. A., Burgess, S. M., & Hirsh, D. (1986). β -glucuronidase from *Escherichia coli* as a gene fusion marker, Proceedings of the National Academic of Science, Washington, DC, USA.
110. Brasileiro, A. C. M., & Carneiro, V. T. C. (1995). *Manual de Transformação Genética de Plantas*. Brasília: Embrapa.
111. Gilissen, L. J. W., Metz, P. L. J., Stiekema, W. J., & Nap, J. (1998). Biosafety of *E. coli* β -glucuronidase (GUS) in plants. *Transgenic Research*, 7, 157–163.
112. Van, B. J., Berthouly, M., Carasco, C., Dufour, M., & Eskes, A. (1995). Transient expression of β -glucuronidase following biolistic delivery of foreign DNA into coffee tissue. *Plant Cell Reports*, 14, 748–752.
113. Torres, A. C., Ferreira, A. T., Widholzer, C. F. N., Romano, E., & Peters, J. A. (2003). Expressão eficiente do gene repórter β -glucuronidase nos tecidos vasculares de batata (*Solanum tuberosum* L.) utilizando de um promotor específico (BRA3) de *Agrobacterium rhizogenes*. *Horticultura Brasileira*, 21, 176–179.
114. Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology*, 98, 503–517.
115. Burnette, W. N. (1981). Western blotting: electrophoretic transfer of proteins from sodium dodecyl sulfate-polyacrylamide gels to unmodified nitrocellulose and radiographic protein A. *Analytical Biochemistry*, 112, 195–203.
116. Burridge, K. (1976). Changes in cellular glycoproteins after transformations: identification of specific glycoproteins and antigens in sodium dodecyl sulphatase gels. *Proceedings of the National Academy of Sciences USA*, 73, 4447–4461.
117. Showe, M. K., Isobe, E., & Onorato, L. (1976). Bacteriophage T4 pre head proteinase II. Its cleavage from the product of gene 21 and regulation in phage-infected cells. *Journal of Molecular Biology*, 107, 55–69.
118. Towbin, H., Staehelin, T., & Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Sciences USA*, 76, 4350–4354.
119. Hornbeck, P. (1991). Enzyme-linked immunosorbent assays, in Current protocols in immunology (Coligan, J. E., Kruisbeek, A. M., Margulies, D. H., Shevach, E. T. & Strober, W., eds), John Wiley & Sons, NY.
120. Greenplate, J. G. (1999). Quantification of *Bacillus thuringiensis* insect control protein (Cry1Ac) over time in Bollgard cotton

- fruit and terminals. *Journal of Economic Entomology*, 92, 1377–1383.
121. Gore, J., Leonard, B. R., & Adamczyk, J. J. (2001). Bollworm (Lepidoptera: Noctuidae) survival and 'Bollgard II' cotton flower bud and flower components. *Journal of Economic Entomology*, 94, 1445–1451.
122. Kranthi, K. R., Naidu, S., Dhawad, C. S., Tatwawadi, A., Mate, K., Patil, E., et al. (2005). Temporal and intra-plant variability of Cry1Ac expression in Bt-cotton and its influence on the survival of the cotton bollworm, *Helicoverpa armigera* (Hübner) (Noctuidae: Lepidoptera). *Current Science*, 89, 291–298.