(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 April 2010 (29.04.2010)

PCT

(10) International Publication Number
WO 2010/045679 A1

(51) **International Patent Classification**:
C12N 15/11 (2006.01) ***C12N 15/82*** (2006.01)
A01H 5/00 (2006.01)

(21) **International Application Number**:
PCT/AU2009/001383

(22) **International Filing Date**:
20 October 2009 (20.10.2009)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
61/106,922 20 October 2008 (20.10.2008) US

(71) **Applicant** *(for all designated States except US)*: **THE UNIVERSITY OF QUEENSLAND** [AU/AU]; St Lucia, Queensland 4067 (AU).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **BIRCH, Robert George** [AU/AU]; 24 Yarrimbah Street, Jindalee, Queensland 4074 (AU). **MANNERS, John M.** [AU/AU]; The University Of Queensland, St Lucia, Queensland 4067 (AU). **MUDGE, Stephen R.** [AU/AU]; The University Of Queensland, St Lucia, Queensland 4067 (AU).

(74) **Agents**: **ARGAET, Victor P.** et al.; Davies Collison Cave, Level 3, 303 Coronation Drive, Milton, Queensland 4064 (AU).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2010/045679 A1

(54) **Title**: TRANSCRIPTIONAL CONTROL ELEMENTS AND USES THEREFOR

(57) **Abstract**: This invention discloses transcriptional control elements for use in plant genetic engineering. More particularly, the present invention discloses tissue-specific promoters for expression of heterologous nucleic acids in plants including monocotyledonous plants. The invention also discloses chimeric nucleic acid constructs comprising a promoter of the invention operably linked to a foreign or endogenous polynucleotide that codes for a protein of interest or a transcript that is capable of modulating expression of a target gene. The invention is further concerned with transformed plant cells, as well as differentiated plants and plant parts, containing the construct of the invention.

"TRANSCRIPTIONAL CONTROL ELEMENTS AND USES THEREFOR"

FIELD OF THE INVENTION

[0001] This invention relates generally to transcriptional control elements for use in plant genetic engineering. More particularly, the present invention relates to tissue-specific promoters for expression of heterologous nucleic acids in plants including monocotyledonous plants. The invention also relates to chimeric nucleic acid constructs comprising a promoter of the invention operably linked to a foreign or endogenous polynucleotide that codes for a protein of interest or a transcript that is capable of modulating expression of a target gene. The invention is further concerned with transformed plant cells, as well as differentiated plants and plant parts, containing the construct of the invention.

BACKGROUND OF THE INVENTION

[0002] A primary goal of genetic engineering is to obtain plants having improved characteristics or traits. Many different types of characteristics or traits are considered advantageous, but those of particular importance include enhanced yield or quality of harvestable plant products, production of additional compounds in plants, enhanced stability or shelf life of the ultimate consumer product obtained from a plant, improvement in the nutritional value of edible portions of a plant, and plant resistance to diseases, insects, herbicides, cold stress, water stress or soil salinity.

[0003] Recent advances in genetic engineering have enabled the incorporation of a selected gene (or genes) into plant cells to impart a desired quality (or qualities) to a plant of interest. The selected gene (or genes) may be derived from a source different from the plant of interest or may be native to the desired plant, but engineered to have different or improved qualities. This new gene (or genes) may then be expressed in cells of the regenerated plant to exhibit the new trait or characteristic.

[0004] In order for the newly incorporated gene to express the transcript and/or protein for which it codes in a plant cell, the proper regulatory signals must be present and in the proper location with respect to the gene. These regulatory signals include at least a promoter, which is a region involved in the regulation of transcription. Other regulatory signals that common in protein-coding genes for example include a 5' non-translated leader sequence and a 3' polyadenylation signal.

[0005] Generally, the efficiency of gene expression is governed largely by the promoter used to express the gene. A promoter is a DNA sequence that directs the cellular machinery of a plant to produce (transcribe) RNA (transcript) from a contiguous transcribable region downstream (3') of the promoter. By corollary, the promoter region lies upstream (5') of the transcribable region. DNA is typically comprised of two polynucleotide strands in anti-parallel orientation with complementary base pairing. The terminology for the location and orientation of DNA sequence elements, including non-transcribed regulatory signals, uses as a reference the sequence of the corresponding transcript, which is a single-stranded polynucleotide with unambiguous 5' and 3' ends. Nucleotides are commonly counted from the point at which transcription commences, with positive (+) numbering in the downstream direction and negative (-) numbering in the upstream direction.

[0006] Promoters typically comprise sequence modules that function in concert to determine the overall promoter activity. Conserved sequence motifs near the transcription start site are believed to function in the binding and orientation of RNA polymerase, whereas sequence motifs more distant from the transcription start site substantially modulate the level and developmental pattern of promoter activity. For example, many plant promoters include a motif with a TATA consensus in the vicinity of -30 nt , and a motif with a CAAT or AGGA consensus at about -70 nt upstream of the transcription start site (for reviews, see Messing, Geraghty, Heidecker, Hu, Kridl and Rubenstein, 1983, pp 211-21 in *Genetic Engineering of Plants* ed Kosuge, Meredith and Hollaender, Plenum; Waugh and Brown, 1991, pp 1-37 in *Plant Genetic Engineering*, ed Grierson, Blackie; Ferl and Paul, 2000, pp 312-57 in *Biochemistry and Molecular Biology of Plants*, ed Buchanan, Gruissem and Jones, ASPP.). Isolated regions of about 40-90 nt upstream of the transcription start site and containing such conserved sequence motifs commonly have little or no promoter activity alone. However, when combined with upstream regulatory sequences from the same promoter or from other promoters, they can show promoter activity in a pattern determined substantially by the upstream regulatory sequences (for example, see Puente, Wei and Deng, 1996, *EMBO J.* **15**: 3732).

[0007] The promoter influences the rate at which the transcript of the gene is made. Assuming the transcript includes a coding region with appropriate translational signals, the promoter also influences the rate at which the resultant protein product of the gene is produced. Promoter activity also can depend on the presence of several other *cis*-acting regulatory elements which, in conjunction with cellular factors, determine strength,

specificity, and transcription initiation site (for a review, see Zawel and Reinberg, 1992, *Curr. Opin. Cell Biol.* **4**:488).

[0008] It has been shown that certain promoters are able to direct RNA synthesis at a higher rate relative to other promoters. These are called "strong promoters." Certain other promoters have been shown to direct RNA production at higher levels only in particular types of cells or tissues and are often referred to as "tissue-specific promoters." Promoters that are capable of directing RNA production in many or all tissues of a plant are called "constitutive promoters". Thus, expression of a chimeric gene (or genes) introduced into a plant may potentially be controlled by identifying and using a promoter with the desired characteristics. The desired expression pattern depends on the nature of the gene product and the trait associated with expression of any particular gene. For example, in the case of an enzyme for conversion of an endogenous storage compound into an alternative industrial product, it may be useful to drive expression preferentially in mature storage tissues, to avoid interference with other plant functions in non-storage tissues.

[0009] There is currently a dearth of promoters that can be used for effective expression of foreign or endogenous coding sequences in desired developmental patterns in plants in general, and monocotyledonous plants in particular.

SUMMARY OF THE INVENTION

[0010] In work leading up to the present invention, studies were carried out to understand the molecular basis for developmental control of plant gene expression, and in particular to understand the basis for preferential expression of certain plant genes in sink tissues, sucrose-storage tissues and mature stem tissues. During the course of these studies, the present inventors surprisingly discovered a class of plant genes with developmental patterns of interest, which enabled the isolation of a structurally related class of promoter sequences that can be used to express heterologous genes in plants, including preferential expression in plant sink (*e.g.*, stem) tissue. Sequence analysis has revealed that these promoter sequences share a common segmental architecture, which can be represented by the following formula:

[0011] 5'-Θ–Δ–Z–Ψ–Φ–Ω-3' (I)

[0012] in which segment Ω is more conserved across the class of promoter sequences than segment Φ, segment Φ is more conserved than segment Z, segment Z is more conserved than segment Δ, segment Δ is more conserved than segment Θ, and Ψ is an optional spacer or intervening segment. Of interest, the present inventors have also determined that segment Ω is transcriptionally active in plant cells.

[0013] Accordingly, in one aspect of the present invention, isolated nucleic acid molecules are provided, which comprise a promoter sequence that is operable in plant cells, including monocotyledonous plant cells. The promoter sequence generally comprises a nucleotide sequence (also referred to herein as segment Ω sequence) selected from the group consisting of:

[0014] (a) RTCAAGTATAGYAGTGGCAGAGAYTCAAAAGATACCTKCGGA TTCTGAATACCACAACCWTAAAAAACAGGATGATGTTATACTTGTCCCCTTCCAT GATACAGGACTGTTTAGTAAYTYCCCAAACWGCCCATRATACATTCTGCACCCTT TATWAAACCTCTN$_a$AGCTACAWCATCTT [SEQ ID NO: 1];

[0015] (b) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 1 or a complement thereof; and

[0016] (c) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 1 or a complement thereof, under at least medium or high stringency conditions,

[0017] wherein:

[0018] A, C, G and T represent the nucleic acid bases adenine, cytosine, guanine and thymine, respectively;

[0019] M is A or C;

[0020] R is A or G;

[0021] W is A or T;

[0022] S is C or G;

[0023] Y is C or T

[0024] K is G or T;

[0025] V is A or C or G;

[0026] H is A or C or T;

[0027] D is A or G or T;

[0028] B is C or G or T;

[0029] N is G or A or C or T; and

[0030] *a* is an integer from 0-3.

[0031] In some embodiments, N_a comprises 0 bases or is the sequence ACT.

[0032] Illustrative examples of sequences according to SEQ ID NO: 1 are suitably selected from the group consisting of:

[0033] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTAT TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 2];

[0034] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA TACAGGACTGTTTAGTAATTCCCCAAACAGCCCATAATACATTCTGCACCCTTTAT TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 3];

[0035] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA

TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTAT
TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 4];

[0036] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT
CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTAT
TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 5];

[0037] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT
CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTAT
TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 6];

[0038] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT
CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTAT
AAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 7];

[0039] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT
CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAACTTCCCAAACAGCCCATGATACATTCTGCACCCTTTAT
TAAACCTCTAGCTACATCATCTT [SEQ ID NO: 8];

[0040] GTCAAGTATAGTAGTGGCAGAGATTCAAAAGATACCTTCGGATT
CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATGATACATTCTGCACCCTTTAT
TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 9];

[0041] GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATT
CTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTAT
TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 10]; and

[0042] ATCAAGTATAGCAGTGGCAGAGACTCAAAAGATACCTGCGGATT
CTGAATACCACAACCTTAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACTGCCCATAATACATTCTGCACCCTTTAT
TAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 11].

[0043] In some embodiments, the promoter sequence further comprises upstream of (a), (b) or (c) a nucleotide sequence (also referred to herein as segment Φ sequence) selected from the group consisting of:

[0044] (d) TKTSSRAMMMWSWKN$_b$WTRTDMYYWWAN$_c$YAHWWYAN$_d$TR TMKRAMCRARYACN$_e$AYSTKRTARTSWWKRTTYAWAWSCKRWGKWSTSAWRKY SDC [SEQ ID NO: 12];

[0045] (e) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 12 or a complement thereof; and

[0046] (f) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 12 or a complement thereof, under at least medium or high stringency conditions,

[0047] wherein:

[0048] A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

[0049] each of *b* and *c* is an integer from 0-2;

[0050] *d* is an integer from 0-4; and

[0051] *e* is an integer from 0-1.

[0052] In some embodiments, N_b is a sequence selected from GA or AA; N_c is the sequence GT; N_d is a sequence selected from ATA or CAAA; or N_e is T. In other embodiments, any one or more of N_b, N_c, N_d or N_e comprises 0 bases.

[0053] Representative examples of sequences according to SEQ ID NO: 12 are suitably selected from the group consisting of:

[0054] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 13];

[0055] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGGC [SEQ ID NO: 14];

[0056] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAGTGTAGATTTAAATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 15];

[0057] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 16];

[0058] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 17];

[0059] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 18];

[0060] TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAAT ACTATGTGGTAATGTAGATTTATAAGCTGTGTACTGAAAGTGAC [SEQ ID NO: 19];

[0061] TTTCCAAAAATCATATGTTACTAAACAAAACAATATATCTAACCA AATACTATGTGGTAGTGTAGATTTATAAGCTGTGTACTGAAAGTGAC [SEQ ID NO: 20];

[0062] TTTCGGAAACTCTTGCATATGCTCTTATATTTCATATCTAACCAAA TACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 21]; and

[0063] TGTGGAACCAAGTGAATTATAACCAAAGTCACTATACAAATGTA GGAACGAGCACACCTTATAATCATTGTTCATATCCGAAGGTGTCATGTCCTC [SEQ ID NO: 22].

[0064] Suitably, the promoter sequence further comprises upstream of (d), (e) or (f) a nucleotide sequence (also referred to herein as segment Z sequence) selected from the group consisting of:

[0065] (g) RBTN$_f$RDRVDBSSRWSDDYRWYKMBHRKRYKKBBN$_g$SAWYHK YTWSBVVVDYBYYSYWMSWRRRACVTRRN$_h$ARGAWAHYTWWMMN$_i$GYYDVDR WYMWYDRWSWWKYTCTDRHYVAMWMHMRS [SEQ ID NO: 23];

[0066] (h) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 23 or a complement thereof; and

[0067] (i) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 23 or a complement thereof, under at least medium or high stringency conditions,

[0068] wherein:

[0069] A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

[0070] *f* is an integer from 0 to 3;

[0071] *g* is an integer from 0 to 2;

[0072] *h* is an integer from 0 to 6; and

[0073] *i* is an integer from 0 to 1.

[0074] In some embodiments, N_f is a sequence selected from ATG, ATA or AAA; N_g is a sequence selected from TA or TC; N_h is the sequence CCCA; or N_i is W. In other embodiments any one or more of N_f, N_g, N_h, or N_i comprises 0 nucleotides.

[0075] Representative examples of sequences according to SEQ ID NO: 23 are suitably selected from the group consisting of:

[0076] ATTATGATAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTT AGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTTT GACATTCTCTGATTGAAACT [SEQ ID NO: 24];

[0077] ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTT AGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTCT GACATTCTCTGATTGAAACT [SEQ ID NO: 25];

[0078] ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTT AGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCT GACATTCTCTGATTGAAACT [SEQ ID NO: 26];

[0079] ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTT AGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTCT GACATTCTCTGATTGAAACT [SEQ ID NO: 27];

[0080] ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTT AGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTCT GACATTCTCTGATTGAAACT [SEQ ID NO: 28];

[0081] ATTATGATAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTT AGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTTT GACATTCTCTGATTGAAACT [SEQ ID NO: 29];

[0082] ATTATGATAAAGGCATCAGCAACTAGAGTACTTTGTAGAATCTTT AGTCCGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCT GACATTCTCTGATTGAAACA [SEQ ID NO: 30];

[0083] ACTATAGGGCACGCGTGGTCGACGCCCGGGCTGGTGAATCTTTA GTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTG ACATTCTCTGATTGAAACA [SEQ ID NO: 31];

[0084] GCTAAAAAAGTTGGGAGTATATTTATTGGGCTTCCTCCAACAGTT TCCAAAACCTCGCTCCTAAAACATAGAAGATAATTTTACGTTAGGAATCTTGAAC TATTTCTAAATCACTCC [SEQ ID NO: 32]; and

[0085] GGTGGACGCCCAAGGTCATCTCTTAGGTGTTCTCCATCTTCTTGG GGCGTGCTCCAACTTGGACCTGGGTCCCAAGGATATCTAACAAGCCTATGACAAC AGTGTAGCTCTTGCCAAAAAT [SEQ ID NO: 33].

[0086] In some embodiments, the promoter sequence further comprises upstream of (g), (h) or (i) a nucleotide sequence (also referred to herein as segment Δ sequence) selected from the group consisting of:

[0087] (j) CYVYRKKN$_j$WMDKSRYNVWRMMRN$_k$CYMRDVBHBRN$_l$YBRCN $_m$HHMSWMRHHMABRVRYARVWN$_n$RAVRMWRN$_o$BRVDWGRRRBRRCN$_p$RYTWYR GN$_q$TSDDTRDWWMVDMRN$_r$WARN$_s$KGGRSWYTWRHMHMYHGRKN$_t$VRRYVRSN$_u$ WRGKRSDTGN$_v$WRSCWAMYRWKHMYWTADCTTDHBBKN$_w$RKYCWAN$_x$AYCVCT MWTMHMN$_y$YYHWVWWGWKTCMYMAHRHKKYHMRWSRYYYRMMDSYCYMRH WN$_z$MRDTSBCN$_{aa}$YTRRBWRAWWATSVHCRN$_{ab}$TGHYTN$_{ac}$TYTYKWYN$_{ad}$AKYHVC N$_{ae}$YHRDTHMTGMADN$_{af}$RHYDTYKVKN$_{ag}$WMWMAN$_{ah}$VRYWHVVAWHBRN$_{ai}$YKCC YTRYYHAN$_{aj}$BWWSWWTN$_{ak}$ATCCWYN$_{al}$MTSYSRN$_{am}$YRWTKVWWAHGRMKHTCK KDWDYVYHYWSRBVWWVMRKTWWAYYAMN$_{an}$TYHTRAYN$_{ao}$YYRAN$_{ap}$GMAKA MWMVAMYWYRAN$_{aq}$AKTDGWYWKRRCMWYYVSMKTRGYWSBKAHDWN$_{ar}$YYT WGYN$_{as}$RGYMDWMYAMKGWRKTBKMCMWGBR [SEQ ID NO: 34];

[0088] (k) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 34 or a complement thereof; and

[0089] (l) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 34 or a complement thereof, under at least medium or high stringency conditions,

[0090] wherein:

[0091] A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

[0092] each of *k*, *l*, *p*, *q*, *v*, *ab*, *ad*, *ae*, *af*, *ag*, *ah*, *ai*, *al*, *ap*, *aq*, *ar* and *as* is an integer from 0 to 1;

[0093] each of *m*, *o*, *s*, *t*, *x*, *y*, *z*, *ac*, *aj*, *ak*, and *ao* is an integer from 0 to 2;

[0094] each of *r* and *w* is an integer from 0 to 3;

[0095] *am* is an integer from 0 to 5;

[0096] *an* is an integer from 0 to 7;

[0097] *aa* is an integer from 0 to 8;

[0098] *n* is an integer from 0 to 9;

[0099] *j* is an integer from 0 to 16; and

[0100] *u* is an integer from 0 to 42.

[0101] In some embodiments, N_j comprises C or a sequence selected from: TCTTAAGGTTGAATA [SEQ ID NO: 35] or CAATCCGGTCG [SEQ ID NO: 36]; N_k is K; N_l is B; N_m is a sequence selected from TG, TT or GA N_n is a sequence selected from GA, GGA, AGA, GGAGAAAGA or ACAAGATAA; N_o is a sequence selected from AG, TT or CT; N_p is T or A; N_q is A; N_r is a sequence selected from CTT or CTC; N_s is the sequence GT; N_t is C or a sequence selected from CC or AA; N_u is C or a sequence selected from TAGTGACTAAAGTTTAGTCCCTGAACTGAACTTTAATCC [SEQ ID NO: 37], TAGTGACTAAAGTTTAGTCCTTGAACTGAACTTTAATCC [SEQ ID NO: 38], TAGTGGACTGAACTTTAATCC [SEQ ID NO: 39] TTATCTCTAGAGGAGTCACTAAATTTTACACTAAACTT [SEQ ID NO: 40], or TGTTCAATAACATATTGATCTAAAACAAAACCCCAACCCTA [SEQ ID NO: 41]; N_v is G; N_w is T or a sequence selected from TT, GCT or AC; N_x is the sequence GG; N_y is the sequence GT; N_z is a sequence selected from AC or GA; N_{aa} is a sequence selected from GA or GCAGCATC; N_{ab} is C; N_{ac} is a sequence selected from GG or TG; N_{ad} is N; N_{ae} is G; N_{af} is R; N_{ag} is R; N_{ah} is M; N_{ai} is T; N_{aj} is a sequence selected from TA, AA or AT; N_{ak} is A or a sequence selected from TA or CC; N_{al} is K; N_{am} is a sequence selected from CAAT, GCAAT, TCA or AT; N_{an} is is a sequence selected from ACAAACA or AGACTGG; N_{ao} is the sequence GG; N_{ap} is K; N_{aq} is G; N_{ar} is A; or N_{as} is T.

[0102] In some embodiments, any one or more of N_j, N_k, N_l, N_m, N_n, N_o, N_p, N_q, N_r, N_s, N_t, N_u, N_v, N_w, N_x, N_y, N_z, N_{aa}, N_{ab}, N_{ac}, N_{ad}, N_{ae}, N_{af}, N_{ag}, N_{ah}, N_{ai}, N_{aj}, N_{ak}, N_{al}, N_{am}, N_{an}, N_{ao}, N_{ap}, N_{aq}, N_{ar} or N_{as} comprises 0 bases.

[0103] Representative sequences according to SEQ ID NO: 34 may be selected from the group consisting of:

[0104] CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCC TCATTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTAT

ATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTT
TAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCTTAACTT
TTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTG
CCGCCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAG
CAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTA
TAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTG
ACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGA
CAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCG
[SEQ ID NO: 42];

[0105] CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCC
TCATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTA
TATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGT
TTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAAC
TTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTT
GCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTAG
TAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTA
TAGTTATATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTG
ACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGA
CAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCG
[SEQ ID NO: 43];

[0106] CTATAGGGCACGCGTGGTGACGGCCCGGGCTGGTCTGCCCCCTC
ATTAATGCACAGAAGGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAG
TCCCTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGACCTTAACTTTT
TTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCC
ACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCA
ACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTGCTTGCTTATA
GTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGAC
GTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACA
ATCACCGTAGTACTTATTTACCTTGTAGTAGTATACTGTATTCGACCTGCG [SEQ
ID NO: 44];

[0107] CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCTGCTGCCCC
TCATTAATGCACAGAAAGAGAGAGAGAGGGGGAGGGAAGACATTTTGGTCTTTA

TATAGTAGCTTTAATGGACTTTAGTACCTAGGTCCAAACCGGTAGTGGACTGAAC
TTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTTTTTATTCTAATCCC
TCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCCTCACTACCAGT
GGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCAACCAAATACTGCAAG
CTATTGGGAAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCAT
GATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATC
AATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTAC
TTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCG [SEQ ID NO: 45];

[0108] CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCC
TCATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTA
TATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGT
TTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAAC
TTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTT
GCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTA
GTAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTT
ATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCT
GACGTAGAGTTTAATCAATTTTAGCTCAAGAATAATACACTATAAAGTGGATATG
ACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCG
[SEQ ID NO: 46];

[0109] CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCC
TCATTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTAT
ATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTT
TAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCTTAACTT
TTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTG
CCACCCTCACCACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAG
CAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTA
TAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTG
ACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGA
CAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCG
[SEQ ID NO: 47];

[0110] GTCTAGTTTCTTAAGGTTGAATAAAGGGACTAAACCATCTAAAA
GACACCCACTGCACCTCATTAATGCACAGAAGGAGAAAGAGAGAGAGAGTGAGA
GGGAGGGCATTTTGGTCTTTATATAGTAGCTTTAATGGACTTTAGTACATTGATCA

AATAGGCAGGGACTTGGAACCAAACATGCCCTTATCTTTTTTTTATTCTAATCGCT
CTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCCTCACTACCAGTG
GCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTAGCAACCAAATACTGCAAGCT
ATTGGGGAAAAAAGGCAAAAAAACATCTCCTTGCTTAGTTGTTTTAATCCATGAT
CCGGGCAATTGTTTGTTAAGGAGATCCTGAATCCTCTGACATAGAGTTTAATCAA
TTTTAACTCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTT
ATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGC
AACTAGAGTACTTTGTAGAATCTTTAGTCCGTTTCTGTAAGATGAACGTGACTAG
GAAACTTATACTGTTGCAATCATCTGACATTCTCTGATTGAAACACGGTTTCCAAA
AATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTGGTAATGTAGA
TTTATAAGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAA
AGATACCTGCG [SEQ ID NO: 48];

[0111] CCCCGGTTCTATGGACACTGACACTAAAGTTTAGTGGCTTTTAGT
CACCAAGGGGTAGCTAAAAGTGTTTACATGGGACAACTGCTTTAGTGGGTGGAAC
CAAATAGGGGGCACTAAACTCTAGAGGAGTCACTTATCTCTAGAGGAGTCACTAA
ATTTTACACTAAACTTTAGTGGGTGGAACCAAACATGCCCTTATCTTGCCCTGCTA
GTCAAACCACTAATCTCCCATGAAGATTCATCATGATGTCCAACACTTGCATGTCT
CGTTCATTCCCTTAAGTAATAATCACCATGCCTTCTTTATAATTTGCTCATTCATGC
ATCATCGTCTCGATCTCAAAGTTCCCAACCACGCCCTACCCAAACTAGAATATCC
TCTCTCTCATCATAATTCATATGACTCTCTTTTTTACATACATCTAAAAGTAAACC
AAACAAACATCATAACTTGATGAAGAAAAAAATATGAGAGTTGTTAGAGCATCT
CCAGTAGTTCCTAAAAATTTAGTTGGTAAACCAATGAGTTTTACAAGTG [SEQ ID
NO: 49]; and

[0112] CCGTGTTACAATCCGGTCGACCTCGCCAAGAAGGCTAATCCCTGC
TGCGAATCGAAGAACACAAGCAAGAACAAGATAAAACGCAACTCAATTGAAGTG
ACAATTACAGATGAATGATTAAGCACTCAAAGTTGGGGTCTTGCAAACCGATAAC
GGCGACTGTTCAATAACATATTGATCTAAAACAAAACCCCAACCCTAAAGTAGGT
GGTGGCTACTGATTATATAGCCTAAGGGACGTCCAAGGATCCCTCTTCACGTCCT
AAATGTGTCCCAACACGTTTCAAGGCCCAACGGCCCAAAAGAAGATGTCGCAGC
ATCCTGACAGATTATGGACGCTGACTTGTTTCGACGATTCCCGTTGGTTCTGAAGA
GATTTTGATGTCAAACCACTTGGATTGGTTTCCTTATCAAATTATCTTTCCATCCA
TATGTGGATCATTGAAAACGGAGTTCGGATGCGTCCTGGGCATCCGTTTTATTAC
AGACTGGTCCTGATGGCCGAGGCAGACTCGAACTCGAATTAGACTGGGCCTCTGG

CTTGGCTGGGACGTCCTTGCTGGCCTTCTAAGGTGGTGGCCAAGGA [SEQ ID NO: 50].

[0113] In some embodiments, the promoter sequence further comprises upstream of (j), (k) or (l) a nucleotide sequence (also referred to herein as segment Θ sequence) selected from the group consisting of:

[0114] (m) YDKHSN$_{at}$GKHHNDSRTDNHYDASN$_{au}$YN$_{av}$NKKHKRVNN$_{aw}$CKD SWBWYSGMYYYN$_{ax}$BYWYASHHHMYRRAKN$_{ay}$CWARWHSNGYRRMTWHARMDY VACKHVWN$_{az}$YYKRTTWWSWM [SEQ ID NO: 51];

[0115] (n) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 51 or a complement thereof; and

[0116] (o) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 51 or a complement thereof, under at least medium or high stringency conditions,

[0117] wherein:

[0118] A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

[0119] each of *at* and *av* is an integer from 0 to 1;

[0120] each of *aw* and *ax* is an integer from 0 to 2; and

[0121] each of *au*, *ay* and *az* is an integer from 0 to 3;

[0122] In some embodiments, N$_{at}$ comprises C, G or A; N$_{au}$ comprises a sequence selected from GGA, GGC or GAT; N$_{av}$ comprises C or T; N$_{aw}$ comprises a sequence selected from AC, CC or AG; N$_{ax}$ comprises the sequence AG; N$_{ay}$ comprises a sequence selected from TGC or ACC; and N$_{az}$ comprises A or a sequence selected from CGC or CCT. In other embodiments, any one or more of N$_{at}$, N$_{au}$, N$_{av}$, N$_{aw}$, N$_{ax}$, N$_{ay}$ or N$_{az}$ comprises 0 nucleotides.

[0123] Representative examples of sequences according to SEQ ID NO: 51 include:

[0124] TAGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTG GCCCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA TTTAGTC [SEQ ID NO: 52];

[0125] TGGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTG GCCTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA TTTAGTC [SEQ ID NO: 53];

[0126] TATAGGGNACNCGTGNTCGACGGCCGGGCTGGTCTGGTTTTGGCT CCTTTTANTCACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGATTT AGTC [SEQ ID NO: 54];

[0127] TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGC CTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGATT TAGTC [SEQ ID NO: 55];

[0128] TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGC CCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGATT TAGTC [SEQ ID NO: 56];

[0129] TGGCCCGGACGACATTTTTGAGGGACCCGTTTAGCACCTGGTTTT GGCCCCTTTTAGTCACTAAATTGCCAAAAAGGGTGACTAAAAAGTAACTAAAACT GATTTAGTC [SEQ ID NO: 57];

[0130] CGTAGGGTCATTCATACCCGAGCCAGTATACACCCTTGTGTCCGC TTTGCACACCATCCGGATCTAGACGCGCAGATACAAATTCACTCGTCGCTCTATTT AGAC [SEQ ID NO: 58]; and

[0131] CTGTCAGGTTAAGATAAACAACGATTTGTTCGAAGAGCGACACA CCGATCCAGCTTCAGATCATAAAGACCCAAATTCTGCAACTTTAGCACCACGTCT CCTCTGGTTATCAA [SEQ ID NO: 59].

[0132] In some embodiments, an intervening sequence (also referred to herein as "segment Ψ sequence") is interposed between the nucleotide sequence according to any one of (d), (e) or (f) and the nucleotide sequence according to the nucleotide sequence according to any one of (g), (h) or (i). Generally, the intervening sequence comprises from about 100 to about 150 nucleotides, typically from about 110 to about 140 nucleotides and suitably from about 115 to about 120 nucleotides. In illustrative examples of this type, the intervening sequence is selected from the group consisting of:

[0133] GGTGTGCTCACTTCGCGAAAAGTAGGCGAAAAGTTACTGTAGCA TTTTTCGTTGTTATTTGGCAAAATTTGTCCAATCATGGACTAACTAGGCTCAAAAG CTTCGTCCCAACATTTACAGACAAACTATGCAATTAGTTATCTTTTTTACCTACAT

TTAATGCTCCATGCATACGTCGAAAGATTTGATGTGATGGGTGTAAGAGGAAATA TTTTGGGTTGGGATGGGGAAGTGAACAAGGCCTCGG [SEQ ID NO: 60], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 60 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 60 or a complement thereof, under at least medium or high stringency conditions;

[0134] CACTTTTAGACTTTCTTTCTCTCTTATATATTTGTATGTACTAGGA TCCTGTCTTACTCTTTATTCTTCCTATATGCTTCCACCTCCAGATTGGCACATTGAT ACGTCCGTGTATATTTCCGCACGATTGGATTGATTTCCCAAATGCCGAGAGATTG AGAGTTGAGATAGGGAGTTCTTGGAGAGCAAATTTTTTCAACCTTGCTAAAAAAT CTAGATTGCGAGTGGG [SEQ ID NO: 61] , or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 61 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 61 or a complement thereof, under at least medium or high stringency conditions; and

[0135] GATAGAATATATTAGTGATTTGGAGGCCTTGCCTACGTGATATTG AGATGGGAACAGATCTATTTGAAAGCTTATCAAACAAGCTTTCCATCAAGTGCTC ATTGACCTCGATCGCACTCCGGATGGATGAGTTATGGCCTTCCCAATGAGGCACT GTCGGGCTGCCGACGAACCGAAAATTCAACTTTGATGAACTTCCATTATGAAACA CATTTGAACCTACAATCAAATAGTGACATGTACA [SEQ ID NO: 62] , or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 62 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 62 or a complement thereof, under at least medium or high stringency conditions.

[0136] In specific embodiments, the promoter sequence has a structure represented by formula (I):

[0137] 5'-Θ–Δ–Z–Ψ–Φ–Ω-3' (I)

[0138] wherein:

[0139] Θ is any one of (m), (n) or (o);

[0140] Δ is any one of (j), (k) or (l);

[0141] Z is any one of (g), (h) or (i);

[0142] Ψ is an optional spacer or intervening sequence;

[0143] Φ is any one of (d), (e) or (f); and

[0144] Ω is any one of (a), (b) or (c).

[0145] In illustrative examples of this type, the promoter sequence comprises a nucleotide sequence selected from the group consisting of:

[0146] TAGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTG
GCCCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA
TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCAT
TAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATAG
TAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAGT
CCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCTTAACTTTTTTT
TTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCGC
CCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAGCAAC
CAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGT
TGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGT
AGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACAAT
CACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGA
TAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGA
TGAACGTGACTAGGAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATTGAAA
CTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTAT
GTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACAG
GATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAA
CAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTT

[SEQ ID NO: 63], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 63 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 63 or a complement thereof, under at least medium or high stringency conditions;

[0147] TGGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTG
GCCTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA
TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCAT

TAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAG
TCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTT
TTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCA
CCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTAGTAAC
CAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGT
TATATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGT
AGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACAAT
CACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGA
TAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGA
TGAACGTGACTAGGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAA
ACTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTA
TGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGGCGTCAAGTATAGTAGT
GGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACA
GGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTCCCCAA
ACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTT

[SEQ ID NO: 64], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 64 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 64 or a complement thereof, under at least medium or high stringency conditions;

[0148] CTATAGGGCACGCGTGGTGACGGCCCGGGCTGGTCTGCCCCCTC
ATTAATGCACAGAAGGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAG
TCCCTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGACCTTAACTTTT
TTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCC
ACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCA
ACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTGCTTGCTTATA
GTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGAC
GTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACA
ATCACCGTAGTACTTATTTACCTTGTAGTAGTATACTGTATTCGACCTGCGATTAT
GATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAA
GATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTGACATTCTCTGATTG
AAACTCGGGGTGTGCTCACTTCGCGAAAAGTAGGCGAAAAGTTACTGTAGCATTT

TTCGTTGTTATTTGGCAAAATTTGTCCAATCATGGACTAACTAGGCTCAAAAGCTT
CGTCCCAACATTTACAGACAAACTATGCAATTAGTTATCTTTTTTACCTACATTTA
ATGCTCCATGCATACGTCGAAAGATTTGATGTGATGGGTGTAAGAGGAAATATTT
TGGGTTGGGATGGGGAAGTGAACAAGGCCTCGGTTTCCAAAAATCATATGTTACT
AAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTAAATGCTGTGT
ACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGA
TTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCAT
GATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTT
ATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 65], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 65 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 65 or a complement thereof, under at least medium or high stringency conditions;

[0149] TATAGGGNACNCGTGNTCGACGGCCGGGCTGGTCTGGTTTTGGCT
CCTTTTANTCACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGATTT
AGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCTGCTGCCCCTCATTAA
TGCACAGAAAGAGAGAGAGAGGGGGAGGGAAGACATTTTGGTCTTTATATAGTA
GCTTTAATGGACTTTAGTACCTAGGTCCAAACCGGTAGTGGACTGAACTTTAATC
CAGGGACATGGAACCAAACATGCCCTTAACTTTTTTTTATTCTAATCCCTCTTACA
TTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACT
GGTTAAATATCCTCATGTTTGGTTTTTTTTAGCAACCAAATACTGCAAGCTATTGG
GAAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGG
CAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAA
CTCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTAT
CTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAAACTA
GAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTAGGAAAC
TTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAACTCGGTTTCCAAAAATCA
TATGTTACTAAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTAT
ATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGAT
ACCTGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTC
CCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCT
GCACCCTTTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 66], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 66 or a

complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 66 or a complement thereof, under at least medium or high stringency conditions;

[0150] TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGC
CTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGATT
TAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCATTA
ATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATAGT
AGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAGTC
CTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTTTT
TATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACC
CTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGTAACC
AAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGTT
GTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGTA
GAGTTTAATCAATTTTAGCTCAAGAATAATACACTATAAAGTGGATATGACAATC
ACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGAT
AAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGAT
GAACGTGACTAGGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAA
CTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTAT
GTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACAG
GATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAA
CAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTT

[SEQ ID NO: 67], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 67 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 67 or a complement thereof, under at least medium or high stringency conditions;

[0151] TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGC
CCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGATT
TAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCATTA
ATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATAGTA
GCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAGTCC
CTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCTTAACTTTTTTTTT
ATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCC

TCACCACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAGCAACCA
AATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGTTG
TATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGTAG
AGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACAATCA
CCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATA
AAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATG
AACGTGACTAGGAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATTGAAACT
CGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGT
GGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGC
AGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACAGGA
TGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACA
GCCCATAATACATTCTGCACCCTTTATAAAACCTCTACTAGCTACAACATCTT

[SEQ ID NO: 68], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 68 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 68 or a complement thereof, under at least medium or high stringency conditions;

[0152] TGGCCCAGCATTGGTGTCCCGGTGGCTGCTAGCCGCCACTAGCCT
GCTGGGGTTCTGCGGCTAACCGGGAACATCTGCAGGCAAACTAGTGGCCCGGAC
GACATTTTTGAGGGACCCGTTTAGCACCTGGTTTTGGCCCCTTTTAGTCACTAAAT
TGCCAAAAAGGGTGACTAAAAAGTAACTAAAACTGATTTAGTCGTCTAGTTTCTT
AAGGTTGAATAAAGGGACTAAACCATCTAAAAGACACCCACTGCACCTCATTAAT
GCACAGAAGGAGAAAGAGAGAGAGAGTGAGAGGGAGGGCATTTTGGTCTTTATA
TAGTAGCTTTAATGGACTTTAGTACATTGATCAAATAGGCAGGGACTTGGAACCA
AACATGCCCTTATCTTTTTTTTATTCTAATCGCTCTTACATTCACTTGTCTCACAAA
GTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCAT
GTTTGGTTTTTTTAGCAACCAAATACTGCAAGCTATTGGGGAAAAAAGGCAAAAA
AACATCTCCTTGCTTAGTTGTTTTAATCCATGATCCGGGCAATTGTTTGTTAAGGA
GATCCTGAATCCTCTGACATAGAGTTTAATCAATTTTAACTCAAGAATAATACAC
TATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTG
TATTCGACCTGCGATTATGATAAAGGCATCAGCAACTAGAGTACTTTGTAGAATC
TTTAGTCCGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCAT
CTGACATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACTAAACAAAA
CATATCTAACCAAATACTATGTGGTAATGTAGATTTATAAGCTGTGTACTGAAAG

TGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTGAATA
CCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGA
CTGTTTAGTAACTTCCCAAACAGCCCATGATACATTCTGCACCCTTTATTAAACCT
CTAGCTACATCATCTT [SEQ ID NO: 69], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 69 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 69 or a complement thereof, under at least medium or high stringency conditions;

[0153] ACTATAGGGCACGCGTGGTCGACGCCCGGGCTGGTGAATCTTTA
GTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTG
ACATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACTAAACAAAACAA
TATATCTAACCAAATACTATGTGGTAGTGTAGATTTATAAGCTGTGTACTGAAAG
TGACGTCAAGTATAGTAGTGGCAGAGATTCAAAAGATACCTTCGGATTCTGAATA
CCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGA
CTGTTTAGTAATTTCCCAAACAGCCCATGATACATTCTGCACCCTTTATTAAACCT
CTACTAGCTACAACATCTT [SEQ ID NO: 70], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 70 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 70 or a complement thereof, under at least medium or high stringency conditions;

[0154] CGTAGGGTCATTCATACCCGAGCCAGTATACACCCTTGTGTCCGC
TTTGCACACCATCCGGATCTAGACGCGCAGATACAAATTCACTCGTCGCTCTATTT
AGACCCCCGGTTCTATGGACACTGACACTAAAGTTTAGTGGCTTTTAGTCACCAA
GGGGTAGCTAAAAGTGTTTACATGGGACAACTGCTTTAGTGGGTGGAACCAAATA
GGGGGCACTAAACTCTAGAGGAGTCACTTATCTCTAGAGGAGTCACTAAATTTTA
CACTAAACTTTAGTGGGTGGAACCAAACATGCCCTTATCTTGCCCTGCTAGTCAA
ACCACTAATCTCCCATGAAGATTCATCATGATGTCCAACACTTGCATGTCTCGTTC
ATTCCCTTAAGTAATAATCACCATGCCTTCTTTATAATTTGCTCATTCATGCATCA
TCGTCTCGATCTCAAAGTTCCCAACCACGCCCTACCCAAACTAGAATATCCTCTCT
CTCATCATAATTCATATGACTCTCTTTTTTACATACATCTAAAAGTAAACCAAACA
AACATCATAACTTGATGAAGAAAAAAATATGAGAGTTGTTAGAGCATCTCCAGTA
GTTCCTAAAAATTTAGTTGGTAAACCAATGAGTTTTACAAGTGGCTAAAAAAGTT
GGGAGTATATTTATTGGGCTTCCTCCAACAGTTTCCAAAACCTCGCTCCTAAAAC

ATAGAAGATAATTTTACGTTAGGAATCTTGAACTATTTCTAAATCACTCCAACCA
CTTTTAGACTTTCTTTCTCTCTTATATATTTGTATGTACTAGGATCCTGTCTTACTC
TTTATTCTTCCTATATGCTTCCACCTCCAGATTGGCACATTGATACGTCCGTGTAT
ATTTCCGCACGATTGGATTGATTTCCCAAATGCCGAGAGATTGAGAGTTGAGATA
GGGAGTTCTTGGAGAGCAAATTTTTTCAACCTTGCTAAAAAATCTAGATTGCGAG
TGGGTTTCGGAAACTCTTGCATATGCTCTTATATTTCATATCTAACCAAATACTAT
GTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACAG
GATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAA
CAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTT

[SEQ ID NO: 71], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 71 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 71 or a complement thereof, under at least medium or high stringency conditions; and

[0155] CTGTCAGGTTAAGATAAACAACGATTTGTTCGAAGAGCGACACA
CCGATCCAGCTTCAGATCATAAAGACCCAAATTCTGCAACTTTAGCACCACGTCT
CCTCTGGTTATCAACCGTGTTACAATCCGGTCGACCTCGCCAAGAAGGCTAATCC
CTGCTGCGAATCGAAGAACACAAGCAAGAACAAGATAAAACGCAACTCAATTGA
AGTGACAATTACAGATGAATGATTAAGCACTCAAAGTTGGGGTCTTGCAAACCGA
TAACGGCGACTGTTCAATAACATATTGATCTAAAACAAAACCCCAACCCTAAAGT
AGGTGGTGGCTACTGATTATATAGCCTAAGGGACGTCCAAGGATCCCTCTTCACG
TCCTAAATGTGTCCCAACACGTTTCAAGGCCCAACGGCCCAAAAGAAGATGTCGC
AGCATCCTGACAGATTATGGACGCTGACTTGTTTCGACGATTCCCGTTGGTTCTGA
AGAGATTTTGATGTCAAACCACTTGGATTGGTTTCCTTATCAAATTATCTTTCCAT
CCATATGTGGATCATTGAAAACGGAGTTCGGATGCGTCCTGGGCATCCGTTTTAT
TACAGACTGGTCCTGATGGCCGAGGCAGACTCGAACTCGAATTAGACTGGGCCTC
TGGCTTGGCTGGGACGTCCTTGCTGGCCTTCTAAGGTGGTGGCCAAGGAGGTGGA
CGCCCAAGGTCATCTCTTAGGTGTTCTCCATCTTCTTGGGGCGTGCTCCAACTTGG
ACCTGGGTCCCAAGGATATCTAACAAGCCTATGACAACAGTGTAGCTCTTGCCAA
AAATAGCGATAGAATATATTAGTGATTTGGAGGCCTTGCCTACGTGATATTGAGA
TGGGAACAGATCTATTTGAAAGCTTATCAAACAAGCTTTCCATCAAGTGCTCATT
GACCTCGATCGCACTCCGGATGGATGAGTTATGGCCTTCCCAATGAGGCACTGTC
GGGCTGCCGACGAACCGAAAATTCAACTTTGATGAACTTCCATTATGAAACACAT

TTGAACCTACAATCAAATAGTGACATGTACATGTGGAACCAAGTGAATTATAACC
AAAGTCACTATACAAATGTAGGAACGAGCACACCTTATAATCATTGTTCATATCC
GAAGGTGTCATGTCCTCATCAAGTATAGCAGTGGCAGAGACTCAAAAGATACCTG
CGGATTCTGAATACCACAACCTTAAAAAACAGGATGATGTTATACTTGTCCCCTT
CCATGATACAGGACTGTTTAGTAATTTCCCAAACTGCCCATAATACATTCTGCACC
CTTTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 72], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 72 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 72 or a complement thereof, under at least medium or high stringency conditions.

[0156] In other illustrative examples, the promoter sequence comprises a nucleotide sequence selected from the group consisting of:

[0157] ACTAGTGATTACTATAGGGCACGCGTGGTCGACGGCCCGGGCTG
GTCAGTTTCAGTGCAGGCATATGTTTTTGTTTCCCGTCTACTTCTATTAGGGTATC
ACCGAGAACTTATTAAAACAACTCCCTAATTCATTTTTTACAAAAAGGAACCGTA
AATCCTGTATCTGAACTTTCTAAAAGATTAGCATGCATTTGGTAAACTCTCCCCTC
CAATTCCCAGGTATGCGAATGCACGGTAAACTAGGATGCTGGCCAGCATTGGTGT
CCGGTGGCTGCTAGCCGCCACTAGCTTGCTGGGGTTCTACGGCTAACCGGGAACA
TCTGCACACAACTAGTAGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGT
TTTGGCCCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAA
CTGATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCC
TCATTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTAT
ATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTT
TAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCTTAACTT
TTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTG
CCGCCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAG
CAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTA
TAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTG
ACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGA
CAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATT
ATGATAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGT
AAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATT
GAAACTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATA

CTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGT
AGTGGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAA
ACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCC
CAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACAT
CTTACTCCATCTTGTCTAGTTGGACAGAAGTTCTCTCTTTCTTGGCTGACTCCAAC
TTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 73], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 73 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 73 or a complement thereof, under at least medium or high stringency conditions;

[0158] ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTAGTAATAG
TAAAATATCCGGTCTCTACTCACTCTATACACTTAGGACGTGTTTGGCTCCCGTCG
CACATGCATGTCGCCATAGCTATGGCTGCACCGTGAGTGCGGTGGATGAATCGTC
CGCCACACATTTGGCAAGTTGTGGTGTAAAAATGAATTGGAGCGTGTGGTGGCTG
CCAAAACTTCAGCGAGTTGCAGTGTGGGTGTTTGGCCGATGCGGCATGCCACACT
TGTGGCACAGCGAGCTATGGCGTGCTGTGGAGGTGAGCCAAACATGCCTTTATAT
TCTCTGATTGACAGTTTCAGTGCAGACATATGTTTTTGTTTCCCGTCTACTTCTATT
AGGGTATCACTGAGAACTTATTAAAACAACTCCCTAATTCATTTTTTACAAAAAG
GAACCGTAAATTATGTATCTGAACTTTCTAAAAGATTAGCATGCATTTGGTAAGC
TCTCCCCTCCAATTCCCAGGTATGCGAATGCACGGTAAACTAGGATGCTGGCCAG
CATTGGTGTCCGGTGGCTGCTAGCCGCCACTAGCCTGCTGGGGTTCTGCGGCTAA
CCGGGAACATCTGCACACAACTAGTGGTCGGATGACATTTTTTAGGGACCTGTTT
GGCACCTGGTTTTGGCCTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAA
AGTGACTAAACTGATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACAT
CCGCTGCCCCTCATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATT
TTGGTCTTTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGT
GACTAAAGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACAT
GCCCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGC
AAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGG
TTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCT
CCTTGCTTATAGTTATATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTG
AATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAG
TGGATATGACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGA
CCTGCGATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAGTC

AGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTCTGACA
TTCTCTGATTGAAACTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCT
AACCAAATACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGGCGTC
AAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAA
CCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTT
AGTAATTCCCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTA
GCTACAACATCTTACTCCATCTTGTCTAGTTGGACAAGTTCTCTCTTTCTTGGCTG
ACTCCAACTTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 74], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 74 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 74 or a complement thereof, under at least medium or high stringency conditions;

[0159] CTATAGGGCACGCGTGGTGACGGCCCGGGCTGGTCTGCCCCCTC
ATTAATGCACAGAAGGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAG
TCCCTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGACCTTAACTTTT
TTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCC
ACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCA
ACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTGCTTGCTTATA
GTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGAC
GTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACA
ATCACCGTAGTACTTATTTACCTTGTAGTAGTATACTGTATTCGACCTGCGATTAT
GATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAA
GATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTGACATTCTCTGATTG
AAACTCGGGGTGTGCTCACTTCGCGAAAAGTAGGCGAAAAGTTACTGTAGCATTT
TTCGTTGTTATTTGGCAAAATTTGTCCAATCATGGACTAACTAGGCTCAAAAGCTT
CGTCCCAACATTTACAGACAAACTATGCAATTAGTTATCTTTTTTACCTACATTTA
ATGCTCCATGCATACGTCGAAAGATTTGATGTGATGGGTGTAAGAGGAAATATTT
TGGGTTGGGATGGGGAAGTGAACAAGGCCTCGGTTTCCAAAAATCATATGTTACT
AAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTAAATGCTGTGT
ACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGA
TTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCAT
GATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTT
ATTAAACCTCTACTAGCTACAACATCTTACTCCATCTTGTCTAGTTGGACAGAAGT

TCTCTCTTTCTTGGCTGACTCCAACTTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 75], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 75 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 75 or a complement thereof, under at least medium or high stringency conditions;

[0160] ACTATAGGGNACNCGTGNTCGACGGCCGGGCTGGTCTGGTTTTG GCTCCTTTTANTCACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCTGCTGCCCCTCAT TAATGCACAGAAAGAGAGAGAGAGGGGGAGGGAAGACATTTTGGTCTTTATATA GTAGCTTTAATGGACTTTAGTACCTAGGTCCAAACCGGTAGTGGACTGAACTTTA ATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTTTTTATTCTAATCCCTCTT ACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCCTCACTACCAGTGGC GACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCAACCAAATACTGCAAGCTA TTGGGAAAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGAT CCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATT TTAACTCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTAT TTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAA ACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTAGG AAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAACTCGGTTTCCAAAA ATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGAT TTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAA GATACCTGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTT GTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACAT TCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCATCTTGTCTAG TTGGACAAGTTCTCTCTTTNTTGGCNGANTNCAACTTACTACACCGCAACTTCTTG TGCCCTTGTT [SEQ ID NO: 76], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 76 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 76 or a complement thereof, under at least medium or high stringency conditions;

[0161] ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTG GCCTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA

TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCAT
TAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAG
TCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTT
TTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCA
CCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGTAA
CCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAG
TTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACG
TAGAGTTTAATCAATTTTAGCTCAAGAATAATACACTATAAAGTGGATATGACAA
TCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATG
ATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAG
ATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGA
AACTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACT
ATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAG
TGGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAAC
AGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCA
AACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCT
TACTCCATCTTGTCTAGTTGGACAAGTTCTCTCTTTCTTGGCTGACTCCAACTTACT
ACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 77], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 77 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 77 or a complement thereof, under at least medium or high stringency conditions;

[0162] ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTG
GCCCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA
TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCAT
TAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATAG
TAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTTTAGT
CCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCTTAACTTTTTTT
TTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCAC
CCTCACCACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAGCAAC
CAAATACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGT
TGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGACGT
AGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATGACAAT

CACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGA
TAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGA
TGAACGTGACTAGGAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATTGAAA
CTCGGTTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTAT
GTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACAG
GATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAA
CAGCCCATAATACATTCTGCACCCTTTATAAAACCTCTACTAGCTACAACATCTTA
CTCCATCTTGTCTAGTTGGACAGAAGTTCTCTCTTTCTTGGCTGACTCCAACTTAC
TACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 78], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 78 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 78 or a complement thereof, under at least medium or high stringency conditions;

[0163] ACATCGTTTTTGTTTCGTTAACTCTATTAAGGTTCACGAGAACTAT
AAACCAACTCCCTAAATTTATTTTTTACAAAAAGGAACCATAAATCTGTATTTCG
ACCTTTCTAAAAGAATAGCATGGCATTGTAAACTCTCCCCTCCCAATTCCCAGGT
ATGCGAATGCATGGTAAACTAGGTGCTGGCCCAGCATTGGTGTCCCGGTGGCTGC
TAGCCGCCACTAGCCTGCTGGGGTTCTGCGGCTAACCGGGAACATCTGCAGGCAA
ACTAGTGGCCCGGACGACATTTTTGAGGGACCCGTTTAGCACCTGGTTTTGGCCC
CTTTTAGTCACTAAATTGCCAAAAAGGGTGACTAAAAAGTAACTAAAACTGATTT
AGTCGTCTAGTTTCTTAAGGTTGAATAAAGGGACTAAACCATCTAAAAGACACCC
ACTGCACCTCATTAATGCACAGAAGGAGAAAGAGAGAGAGAGTGAGAGGGAGG
GCATTTTGGTCTTTATATAGTAGCTTTAATGGACTTTAGTACATTGATCAAATAGG
CAGGGACTTGGAACCAAACATGCCCTTATCTTTTTTTTATTCTAATCGCTCTTACA
TTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACT
GGTTAAATATCCTCATGTTTGGTTTTTTTAGCAACCAAATACTGCAAGCTATTGGG
GAAAAAAGGCAAAAAAACATCTCCTTGCTTAGTTGTTTTAATCCATGATCCGGGC
AATTGTTTGTTAAGGAGATCCTGAATCCTCTGACATAGAGTTTAATCAATTTTAAC
TCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATC
TTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGCAACTAG
AGTACTTTGTAGAATCTTTAGTCCGTTTCTGTAAGATGAACGTGACTAGGAAACTT
ATACTGTTGCAATCATCTGACATTCTCTGATTGAAACACGGTTTCCAAAAATCATA
TGTTACTAAACAAAACATATCTAACCAAATACTATGTGGTAATGTAGATTTATAA

GCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATAC
CTGCGGATTCTGAATACCACAACCATAAAAAAACAGGATGATGTTATACTTGTCCC
CTTCCATGATACAGGACTGTTTAGTAACTTCCCAAACAGCCCATGATACATTCTGC
ACCCTTTATTAAACCTCTAGCTACATCATCTTACTCCATCTTGTCTAGTTGGACAG
AAGTTCTCTCTTTCTTGGCTGACTCCAACTACTACACCGCAACTTCTTGTGTCCTT
GTTCCAACCATCACAACTGAGC [SEQ ID NO: 79], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 79 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 79 or a complement thereof, under at least medium or high stringency conditions;

[0164] ACTATAGGGCACGCGTGGTCGACGCCCGGGCTGGTGAATCTTTA
GTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTG
ACATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACTAAACAAAACAA
TATATCTAACCAAATACTATGTGGTAGTGTAGATTTATAAGCTGTGTACTGAAAG
TGACGTCAAGTATAGTAGTGGCAGAGATTCAAAAGATACCTTCGGATTCTGAATA
CCACAACCATAAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGA
CTGTTTAGTAATTTCCCAAACAGCCCATGATACATTCTGCACCCTTTATTAAACCT
CTACTAGCTACAACATCTTACTCCATCTTGTCTAGTTGGACAGAAGTTCTCTCTTT
CTTGGCTGACTCCAACTACTGCACCGCAACTTCTTGTGCCCTTGTTCCAACCATCA
CAATTGAGC [SEQ ID NO: 80], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 80 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 80 or a complement thereof, under at least medium or high stringency conditions;

[0165] ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTCACTTC
GACGAGAAGACACAACACACACCACCTCTCAGAAGAGACCTAGGATCTGCTC
CCTCTCTCGAAAAGCTTGTGACCCCTACTACAAGTACTCCGGTGTAAGGCAACAC
GAGCAGCGGCAAAACCCACTGGACGTAGGGTCATTCATACCCGAGCCAGTATAC
ACCCTTGTGTCCGCTTTGCACACCATCCGGATCTAGACGCGCAGATACAAATTCA
CTCGTCGCTCTATTTAGACCCCCGGTTCTATGGACACTGACACTAAAGTTTAGTGG
CTTTTAGTCACCAAGGGGTAGCTAAAAGTGTTTACATGGGACAACTGCTTTAGTG
GGTGGAACCAAATAGGGGGCACTAAACTCTAGAGGAGTCACTTATCTCTAGAGG
AGTCACTAAATTTTACACTAAACTTTAGTGGGTGGAACCAAACATGCCCTTATCTT

GCCCTGCTAGTCAAACCACTAATCTCCCATGAAGATTCATCATGATGTCCAACAC
TTGCATGTCTCGTTCATTCCCTTAAGTAATAATCACCATGCCTTCTTTATAATTTGC
TCATTCATGCATCATCGTCTCGATCTCAAAGTTCCCAACCACGCCCTACCCAAACT
AGAATATCCTCTCTCTCATCATAATTCATATGACTCTCTTTTTTACATACATCTAA
AAGTAAACCAAACAAACATCATAACTTGATGAAGAAAAAAATATGAGAGTTGTT
AGAGCATCTCCAGTAGTTCCTAAAAATTTAGTTGGTAAACCAATGAGTTTTACAA
GTGGCTAAAAAAGTTGGGAGTATATTTATTGGGCTTCCTCCAACAGTTTCCAAAA
CCTCGCTCCTAAAACATAGAAGATAATTTTACGTTAGGAATCTTGAACTATTTCTA
AATCACTCCAACCACTTTTAGACTTTCTTTCTCTCTTATATATTTGTATGTACTAGG
ATCCTGTCTTACTCTTTATTCTTCCTATATGCTTCCACCTCCAGATTGGCACATTGA
TACGTCCGTGTATATTTCCGCACGATTGGATTGATTTCCCAAATGCCGAGAGATTG
AGAGTTGAGATAGGGAGTTCTTGGAGAGCAAATTTTTTCAACCTTGCTAAAAAAT
CTAGATTGCGAGTGGGTTTCGGAAACTCTTGCATATGCTCTTATATTTCATATCTA
ACCAAATACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCA
AGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAAC
CATAAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTA
GTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAG
CTACAACATCTTACTCCACCTTGTCTAGTTGGACAAGTTCTCTCTTTCTTGGCTGA
CTCCAACTTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 81], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 81 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 81 or a complement thereof, under at least medium or high stringency conditions; and

[0166] ACTAGTGATTACTATAGGGCACGCGTGGTCGACGGCCCGGGCTG
GTCTCAATTGAATAAAAGATGTGGCCTATAGAAAATTAGGACTGTGCTCTAAAAA
GCATACCAAAACTTGTGGGTCCAGAAACTCAATTTTCCTAATCGAATTTCGCAAA
TCTATTTTTTGCGAACCCAGCTATGCTGGGATGAAATAGATCTAAAATTTTCTGGC
GTTCTCCGTAGATATAAAATTTTCCCCGTTTCAAGTTCGGATGCAAAAGTTATGAC
TGTTTTAAGGAAGCTGCTCGAATCAGGGTTTTAGTGAACACAAGATGCAAACCGA
TGGTGATACGGAATAGCGGCGGGTGGAGGGATCAACACAAATTAGAAAAGAACA
CAATCTAAACCAGCAACAAGACTTTGACCTAGGACACAAACTCAACAAAGCGAA
CTCTGAAACTGAAATATGCAAAGGCTATGGCGCGGAAGGTTCTAGGACAGGAAA
AACAAGTATGGCACTGGACTATGGGACGAATGCGAAACACTCAAACGAGAGAAT
AAAATGTGAACCTAACGGTATACCTTAGCTCTGATTACCACTTTATGTAGACGGG

GGTCCCGATTTTCTGTCAGGTTAAGATAAACAACGATTTGTTCGAAGAGCGACAC
ACCGATCCAGCTTCAGATCATAAAGACCCAAATTCTGCAACTTTAGCACCACGTC
TCCTCTGGTTATCAACCGTGTTACAATCCGGTCGACCTCGCCAAGAAGGCTAATC
CCTGCTGCGAATCGAAGAACACAAGCAAGAACAAGATAAAACGCAACTCAATTG
AAGTGACAATTACAGATGAATGATTAAGCACTCAAAGTTGGGGTCTTGCAAACCG
ATAACGGCGACTGTTCAATAACATATTGATCTAAAACAAAACCCCAACCCTAAAG
TAGGTGGTGGCTACTGATTATATAGCCTAAGGGACGTCCAAGGATCCCTCTTCAC
GTCCTAAATGTGTCCCAACACGTTTCAAGGCCCAACGGCCCAAAAGAAGATGTCG
CAGCATCCTGACAGATTATGGACGCTGACTTGTTTCGACGATTCCCGTTGGTTCTG
AAGAGATTTTGATGTCAAACCACTTGGATTGGTTTCCTTATCAAATTATCTTTCCA
TCCATATGTGGATCATTGAAAACGGAGTTCGGATGCGTCCTGGGCATCCGTTTTA
TTACAGACTGGTCCTGATGGCCGAGGCAGACTCGAACTCGAATTAGACTGGGCCT
CTGGCTTGGCTGGGACGTCCTTGCTGGCCTTCTAAGGTGGTGGCCAAGGAGGTGG
ACGCCCAAGGTCATCTCTTAGGTGTTCTCCATCTTCTTGGGGCGTGCTCCAACTTG
GACCTGGGTCCCAAGGATATCTAACAAGCCTATGACAACAGTGTAGCTCTTGCCA
AAAATAGCGATAGAATATATTAGTGATTTGGAGGCCTTGCCTACGTGATATTGAG
ATGGGAACAGATCTATTTGAAAGCTTATCAAACAAGCTTTCCATCAAGTGCTCAT
TGACCTCGATCGCACTCCGGATGGATGAGTTATGGCCTTCCCAATGAGGCACTGT
CGGGCTGCCGACGAACCGAAAATTCAACTTTGATGAACTTCCATTATGAAACACA
TTTGAACCTACAATCAAATAGTGACATGTACATGTGGAACCAAGTGAATTATAAC
CAAAGTCACTATACAAATGTAGGAACGAGCACACCTTATAATCATTGTTCATATC
CGAAGGTGTCATGTCCTCATCAAGTATAGCAGTGGCAGAGACTCAAAAGATACCT
GCGGATTCTGAATACCACAACCTTAAAAAACAGGATGATGTTATACTTGTCCCCT
TCCATGATACAGGACTGTTTAGTAATTTCCCAAACTGCCCATAATACATTCTGCAC
CCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCATCTTGTCTAGTTGGACA
GAAGTTCTCTCTTTCTTGGCTGACTCCAACTTACTACACCGCAACTTCTTGTGCCC
TTGTT [SEQ ID NO: 82], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 82 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 82 or a complement thereof, under at least medium or high stringency conditions.

[0167] In a related aspect, the present invention provides isolated nucleic acid molecules, which comprise a promoter sequence that is operable in plant cells, including

monocotyledonous plant cells. The promoter sequence generally comprises a nucleotide sequence selected from the group consisting of:

[0168] (i) a nucleotide sequence set forth in any one of SEQ ID NO: 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81 and 82 or a complement thereof;

[0169] (ii) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in any one of SEQ ID NO: 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81 and 82 or a complement thereof; and

[0170] (iii) a nucleotide sequence that hybridizes to the sequence set forth in any one of SEQ ID NO: 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81 and 82 or a complement thereof, under at least medium or high stringency conditions.

[0171] In accordance with the present invention, it is predicted that promoter elements positioned upstream or 5' to the TATA box in segment Ω of the exemplified promoter sequences set forth in SEQ ID NO: 63-82 (*e.g.*, promoter elements in any one or more of segments Φ, Ψ, Z, Δ and Θ) influence the transcription initiation rate and/or tissue (*e.g.*, sink tissue) expression of operably connected nucleic acid sequences, and could be fused, therefore, to heterologous core promoter sequences to produce chimeric promoters with similar expression patterns to the exemplified promoter sequences. Accordingly, in another aspect, the present invention provides isolated nucleic acid molecules comprising chimeric promoter sequences that generally comprise at least one segment selected from the group consisting of a heterologous core promoter sequence in operable connection with a nucleotide sequence selected from the group consisting of :

[0172] (1) a nucleotide sequence corresponding to at least one segment selected from segments Φ, Ψ, Z, Δ and Θ of SEQ ID NO: 63-82, or

[0173] (2) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence of the at least one segment or a complement thereof, or

[0174] (3) a nucleotide sequence that hybridizes to the sequence s of the at least one segment or a complement thereof, under at least medium or high stringency conditions .

[0175] In some embodiments, the nucleotide sequence represented by (1) comprises segment Φ, and optionally Ψ. In other embodiments, the nucleotide sequence

represented by (1) comprises segments Φ and Z, and optionally Ψ. In yet other embodiments, the nucleotide sequence represented by (1) comprises segments Φ, Z and Δ, and optionally Ψ. In still other embodiments, the nucleotide sequence represented by (1) comprises segments Φ, Z, Δ and Θ, and optionally Ψ.

[0176] Generally, a promoter of the invention can be fused to a transcribable sequence to create a chimeric construct. This construct can then be introduced into a host cell, typically a plant cell or plant or plant part, by any method of choice. Accordingly, in another aspect of the present invention, there is provided chimeric nucleic acid constructs comprising an isolated nucleic acid comprising a promoter sequence as broadly described above, in operable connection with a heterologous (*e.g.*, foreign or endogenous) nucleic acid sequence to be transcribed.

[0177] In some embodiments, the chimeric construct further comprises a 3' non-translated sequence that is operably linked to the heterologous nucleic acid sequence and that functions in plant cells to terminate transcription and/or to cause addition of a polyadenylated nucleotide sequence to the 3' end of a transcribed RNA sequence.

[0178] The heterologous nucleic acid sequence is heterologous with respect to the plant cell in which it is or will be introduced. In some embodiments, the heterologous nucleic acid sequence encodes a structural or regulatory protein, or alternatively, a transcript capable of modulating expression of a corresponding target gene. In illustrative examples of this type, the transcript comprises an antisense RNA or a ribozyme or other transcribed region aimed at downregulation of expression of the corresponding target gene. For example, the other transcribed region may comprise a sense transcript aimed at sense suppression (co-suppression) of the corresponding target gene or a hairpin transcript aimed at RNAi-mediated downregulation of the target gene.

[0179] Depending upon the heterologous nucleic acid sequence selected, the chimeric construct in some embodiments may be further characterized in that the promoter sequence is capable of conferring transcription of the heterologous nucleic acid sequence in a specific tissue (*e.g.*, stem tissue) of the plant. In specific embodiments, the promoter sequence is capable of regulating transcription of the heterologous nucleic acid preferentially in stem tissue of a plant. In illustrative examples of this type, the chimeric construct may be further characterized in that the promoter sequence is capable of regulating transcription of the heterologous nucleic acid sequence in mature stem tissue of a plant.

[0180] Plants falling within the scope of the present invention encompass any taxonomic grouping, including angiosperms, gymnosperms, monocotyledons and dicotyledons. In some embodiments, the plant is selected from monocotyledonous plants such as cereals, sugarcane, bananas and pineapples. In specific embodiments, the plant is a sucrose-accumulating plant such as sugarcane, sugarbeet or sweet sorghum.

[0181] In yet another aspect, the present invention provides methods for expression of a heterologous nucleic acid sequence. These methods generally comprise introducing into a plant cell a chimeric construct as broadly described above.

[0182] In still another aspect, the present invention contemplates methods for producing a transformed plant cell, wherein the methods generally comprise introducing into a plant cell a chimeric construct as broadly described above.

[0183] In a further aspect, the present invention provides methods for producing transformed plant cells. These methods generally comprise introducing into regenerable plant cells a chimeric construct as broadly described above so as to yield transformed plant cells and identifying or selecting the transformed plant cells.

[0184] In still a further aspect, the present invention provides methods for selecting stable genetic transformants from transformed plant cells, wherein the methods generally comprise introducing into regenerable plant cells a chimeric construct as broadly described above so as to yield transformed plant cells and identifying or selecting a transformed plant cell line from the transformed plant cells.

[0185] In some embodiments, the regenerable cells are regenerable dicotyledonous plant cells. In other embodiments, the regenerable cells are monocotyledonous plant cells such as regenerable graminaceous or non-graminaceous monocotyledonous plant cells. In some embodiments, the expression of the chimeric construct in the transformed cells imparts a phenotypic characteristic to the transformed cells. Suitably, the imparted phenotype results from expression of the heterologous nucleic acid sequence.

[0186] According to another aspect of the present invention, transformed plant cells are provided, containing a chimeric construct as broadly described above.

[0187] In still another aspect, the present invention contemplates methods for producing a differentiated transgenic plant. These methods generally comprise introducing a chimeric construct as broadly described above into regenerable plant cells so as to yield

regenerable transformed cells, identifying or selecting a population of transformed cells, and regenerating a differentiated transgenic plant from the population.

[0188] In some embodiments, the expression of the chimeric construct renders the differentiated transgenic plant identifiable over the corresponding non-transgenic plant.

[0189] In still a further aspect, the invention provides differentiated transgenic plants comprising plant cells containing a chimeric construct as broadly described above. In some embodiments, the chimeric construct is transmitted through a complete cycle of the differentiated transgenic plant to its progeny so that it is expressed by the progeny plants. Accordingly, the present invention also provides cells, tissues, leaves, fruit, flowers, seeds and other reproductive material, material used for vegetative propagation, progeny plants including F1 hybrids, male-sterile plants and all other plants and plant products derivable from the differentiated transgenic plant.

[0190] The 942-nt sequence set forth in SEQ ID NO: 83 is a transcribable nucleic acid sequence comprising an ORF, as set forth in SEQ ID NO: 83, which codes for a 313-amino acid sequence, as set forth in SEQ ID NO: 84. This ORF is transcribed at high levels in stem tissues, including mature stem tissues, of sugarcane (*Saccharum* sp.). Accordingly, nucleotide sequences that correspond or are complementary to at least a portion of the sequence set forth in SEQ ID NO: 83 may be useful as probes for isolating homologous transcribable sequences from other plants, especially from other sugarcane plants and more broadly other monocotyledonous plants such as cereals, bananas and pineapples to, in turn, permit the isolation of promoter sequences with analogous qualities to those described herein. Thus, in another aspect, the present invention contemplates isolated nucleic acid molecules comprising a promoter sequence or biologically active fragment thereof or variant of these, wherein the promoter sequence is located upstream of a transcribable nucleic acid sequence that hybridizes to a nucleic acid probe derived from the polynucleotide sequence set forth in SEQ ID NO: 83. Advantageously, the isolated promoter sequence is of sufficient length such that it is capable of initiating and/or regulating transcription of a DNA sequence to which it is coupled. The promoter sequence may be between about 150 nts and 1200 nts in length and usually greater than 250 nts in length. Suitably, analogous promoter sequences may be obtained from plants, especially from monocotyledonous plants such as cereals, sugarcane, bananas and pineapples, which contain a nucleotide sequence that is capable of hybridizing to a nucleic acid probe derived from the sequence set forth in SEQ ID NO: 83 under at least medium stringency conditions, and especially under high stringency conditions.

BRIEF DESCRIPTION OF THE DRAWINGS

[0191] Figure 1 is a photographic representation of various Northern analyses showing expression of *ScR1MYB1* in different sugarcane tissues. (a) Northern analysis using total RNA isolated from meristems (M), internodes 1 to 3 (1-3), internode 5 (5), internode 12 (12), internode 20 (20), expanding leaf (EL) and mature leaf (ML) from field-grown Q117 plants, plus root (R) from glasshouse-grown Q117 plants, probed with the 838bp *Hin*dIII fragment of the *ScR1MYB1* cDNA. (b) Northern analysis using total RNA from internode 20 (IN20), mature roots (MR) and root tips (RT) from glasshouse-grown Q117 plants, probed with the 838bp *Hin*dIII fragment of the *ScR1MYB1* cDNA. (c) Northern analysis using total RNA from dissected rind (R), storage parenchyma (P) and vascular strands (V) from internode 8 of field-grown Q117 plants, probed with the full insert of EST clone MCSA063B04.

[0192] Figure 2 is a photographic representation of a Southern analysis showing genetic complexity of *ScR1MYB1* in sugarcane. Genomic DNA from cultivar Q117 was digested with the restriction enzymes *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K), *Sac*I (S) and *Xba*I (X), and probed with the 838bp *Hin*dIII fragment of the *ScR1MYB1* cDNA.

[0193] Figure 3 is a diagrammatic representation illustrating an alignment of the proteins encoded by eight alleles of *ScR1MYB1* identified in sugarcane cultivar Q117. The MYB DNA binding domain is marked by a solid bar above, the CCHC-type zinc finger is marked by a dashed line, and the positions of introns 1 and 2 are marked by vertical arrows. "Z1 us" represents the sequence encoded by the partially spliced Z1 transcript.

[0194] Figure 4 is a graphical representation showing allelic expression patterns in mature stem (internode 12) of sugarcane cultivars Q117 and Q200, as revealed by cloning and sequencing RT-PCR products.

[0195] Figure 5 is a diagrammatic representation of a sequence alignment of the promoter sequences corresponding to alleles A1, A2, A3, A4, A5, A6, A7, A8, A9 and B of the ScR1MYB1 gene. The nucleotide sequences delineating segments Θ, Δ, Z, Ψ, Φ and Ω are shown.

TABLE A

BRIEF DESCRIPTION OF THE SEQUENCES

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 1	MYB promoter consensus sequence - segment Ω	165 nts
SEQ ID NO: 2	Allele A4 promoter sequence - segment Ω	181 nts
SEQ ID NO: 3	Allele A7 promoter sequence - segment Ω	181 nts
SEQ ID NO: 4	Allele A3 promoter sequence - segment Ω	181 nts
SEQ ID NO: 5	Allele A6 promoter sequence - segment Ω	181 nts
SEQ ID NO: 6	Allele A1 promoter sequence - segment Ω	181 nts
SEQ ID NO: 7	Allele A2 promoter sequence - segment Ω	181 nts
SEQ ID NO: 8	Allele A9 promoter sequence - segment Ω	178 nts
SEQ ID NO: 9	Allele A8 promoter sequence - segment Ω	181 nts
SEQ ID NO: 10	Allele A5 promoter sequence - segment Ω	181 nts
SEQ ID NO: 11	Allele B promoter sequence - segment Ω	181 nts
SEQ ID NO: 12	MYB promoter consensus sequence - segment Φ	92 nts
SEQ ID NO: 13	Allele A4 promoter sequence - segment Φ	89 nts
SEQ ID NO: 14	Allele A7 promoter sequence - segment Φ	89 nts
SEQ ID NO: 15	Allele A3 promoter sequence - segment Φ	89 nts
SEQ ID NO: 16	Allele A6 promoter sequence - segment Φ	89 nts
SEQ ID NO: 17	Allele A1 promoter sequence - segment Φ	18 nts
SEQ ID NO: 18	Allele A2 promoter sequence - segment Φ	18 nts
SEQ ID NO: 19	Allele A9 promoter sequence - segment Φ	18 nts
SEQ ID NO: 20	Allele A8 promoter sequence - segment Φ	92 nts
SEQ ID NO: 21	Allele A5 promoter sequence - segment Φ	91 nts
SEQ ID NO: 22	Allele B promoter sequence - segment Φ	96 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 23	MYB promoter consensus sequence - segment Z	119 nts
SEQ ID NO: 24	Allele A4 promoter sequence - segment Z	120 nts
SEQ ID NO: 25	Allele A7 promoter sequence - segment Z	120 nts
SEQ ID NO: 26	Allele A3 promoter sequence - segment Z	120 nts
SEQ ID NO: 27	Allele A6 promoter sequence - segment Z	120 nts
SEQ ID NO: 28	Allele A1 promoter sequence - segment Z	120 nts
SEQ ID NO: 29	Allele A2 promoter sequence - segment Z	120 nts
SEQ ID NO: 30	Allele A9 promoter sequence - segment Z	120 nts
SEQ ID NO: 31	Allele A8 promoter sequence - segment Z	118 nts
SEQ ID NO: 32	Allele A5 promoter sequence - segment Z	117 nts
SEQ ID NO: 33	Allele B promoter sequence - segment Z	121 nts
SEQ ID NO: 34	MYB promoter consensus sequence - segment Δ	504 nts
SEQ ID NO: 35	N_j sequence 1	15 nts
SEQ ID NO: 36	N_j sequence 2	11 nts
SEQ ID NO: 37	N_u sequence 1	39 nts
SEQ ID NO: 38	N_u sequence 2	39 nts
SEQ ID NO: 39	N_u sequence 3	21 nts
SEQ ID NO: 40	N_u sequence 4	38 nts
SEQ ID NO: 41	N_u sequence 5	41 nts
SEQ ID NO: 42	Allele A4 promoter sequence - segment Δ	540 nts
SEQ ID NO: 43	Allele A7 promoter sequence - segment Δ	540 nts
SEQ ID NO: 44	Allele A3 promoter sequence - segment Δ	538 nts
SEQ ID NO: 45	Allele A6 promoter sequence - segment Δ	524 nts
SEQ ID NO: 46	Allele A1 promoter sequence - segment Δ	531 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 47	Allele A2 promoter sequence - segment Δ	540 nts
SEQ ID NO: 48	Allele A9 promoter sequence - segment Δ	775 nts
SEQ ID NO: 49	Allele A5 promoter sequence - segment Δ	538 nts
SEQ ID NO: 50	Allele B promoter sequence - segment Δ	586 nts
SEQ ID NO: 51	MYB promoter consensus sequence - segment Θ	105 nts
SEQ ID NO: 52	Allele A4 promoter sequence - segment Θ	107 nts
SEQ ID NO: 53	Allele A7 promoter sequence - segment Θ	107 nts
SEQ ID NO: 54	Allele A6 promoter sequence - segment Θ	104 nts
SEQ ID NO: 55	Allele A1 promoter sequence - segment Θ	105 nts
SEQ ID NO: 56	Allele A2 promoter sequence - segment Θ	105 nts
SEQ ID NO: 57	Allele A9 promoter sequence - segment Θ	109 nts
SEQ ID NO: 58	Allele A5 promoter sequence - segment Θ	105 nts
SEQ ID NO: 59	Allele B promoter sequence - segment Θ	113 nts
SEQ ID NO: 60	Allele A3 intervening sequence Ψ	247 nts
SEQ ID NO: 61	Allele A9 intervening sequence Ψ	229 nts
SEQ ID NO: 62	Allele B intervening sequence Ψ	244 nts
SEQ ID NO: 63	Allele A4 promoter sequence – segments Θ–Ω	1040 nts
SEQ ID NO: 64	Allele A7 promoter sequence – segments Θ–Ω	1040 nts
SEQ ID NO: 65	Allele A3 promoter sequence – segments Θ–Ω	1178 nts
SEQ ID NO: 66	Allele A6 promoter sequence – segments Θ–Ω	1021 nts
SEQ ID NO: 67	Allele A1 promoter sequence – segments Θ–Ω	1039 nts
SEQ ID NO: 68	Allele A2 promoter sequence – segments Θ–Ω	1038 nts
SEQ ID NO: 69	Allele A9 promoter sequence – segments Θ–Ω	1111 nts
SEQ ID NO: 70	Allele A8 promoter sequence – segments Θ–Ω	394 nts

SEQUENCE ID NUMBER	SEQUENCE	LENGTH
SEQ ID NO: 71	Allele A5 promoter sequence – segments Θ–Ω	1264 nts
SEQ ID NO: 72	Allele B promoter sequence – segments Θ–Ω	1344 nts
SEQ ID NO: 73	Allele A4 promoter sequence – full-length	1460 nts
SEQ ID NO: 74	Allele A7 promoter sequence – full-length	1743 nts
SEQ ID NO: 75	Allele A3 promoter sequence – full-length	1262 nts
SEQ ID NO: 76	Allele A6 promoter sequence – full-length	1105 nts
SEQ ID NO: 77	Allele A1 promoter sequence – full-length	1123 nts
SEQ ID NO: 78	Allele A2 promoter sequence – full-length	1124 nts
SEQ ID NO: 79	Allele A9 promoter sequence – full-length Ω	1395 nts
SEQ ID NO: 80	Allele A8 promoter sequence – full-length	496 nts
SEQ ID NO: 81	Allele A5 promoter sequence – full-length	1522 nts
SEQ ID NO: 82	Allele B promoter sequence – full-length	2032 nts
SEQ ID NO: 83	MYB ORF	942 nts
SEQ ID NO: 84	Amino acid sequence encoded by MYB ORF	313 aa
SEQ ID NO: 85	*Nos* core promoter region	101 nts
SEQ ID NO: 86	CAMV core promoter region	65 nts
SEQ ID NO: 87	CAMV core promoter region with as-1 enhancer	90 nts

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

[0196] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For the purposes of the present invention, the following terms are defined below.

[0197] The articles "a" and "an" are used herein to refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

[0198] By "about" is meant a quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length that varies by as much as 15, 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 % to a reference quantity, level, value, number, frequency, percentage, dimension, size, amount, weight or length.

[0199] The term "amplicon" refers to a target sequence for amplification, and/or the amplification products of a target sequence for amplification. In certain other embodiments an "amplicon" may include the sequence of probes or primers used in amplification.

[0200] The term "biologically active fragment", as applied to promoter sequences, refers to a fragment that has at least about 0.1, 0.5, 1, 2, 5, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30% of the activity of a reference promoter sequence. It will also be understood that the phrase "biologically active fragment" refers to a part of an indicated DNA sequence that modulates RNA transcription or that, when fused to a particular gene and introduced into a plant cell, causes expression of the gene at a level higher than is possible in the absence of such part of the indicated DNA sequence. In specific embodiments, "biologically active fragments" encompass a portion of a promoter sequence that when added to a sequence including one or more 'core' promoter elements or motifs such as a TATA motif, promotes transcription in at least one tissue type to a greater extent than the 'core' sequence without the addition of the portion. Included within the scope of the present invention are biologically active fragments of at least about 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50 ,60 , 70, 80, 90, 100, 120, 140, 160, 180, 200, 250, 300, 350, 400, 450, 500, 550, 600, or 650

nucleotides in length, or almost up to the number of nucleotides present in a full-length promoter sequence.

[0201] The terms "chimeric construct," "chimeric nucleic acid" or "chimeric DNA" and the like are used herein to refer to a gene or nucleic acid sequence, either single- or double-stranded, comprising at least two nucleic acid sequences from species which do not combine nucleic acids such as DNA under natural conditions, or which nucleic acid sequences are positioned or linked in a manner which does not normally occur in the native genome of the untransformed plant.

[0202] Throughout this specification, unless the context requires otherwise, the words "comprise," "comprises" and "comprising" will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements.

[0203] By "corresponds to" or "corresponding to" is meant a polynucleotide (a) having a nucleotide sequence that is substantially identical or complementary to all or a portion of a reference polynucleotide sequence or (b) encoding an amino acid sequence identical to an amino acid sequence in a peptide or protein. This phrase also includes within its scope a peptide or polypeptide having an amino acid sequence that is substantially identical to a sequence of amino acids in a reference peptide or protein.

[0204] The terms "growing" or "regeneration" as used herein mean growing a whole, differentiated plant from a plant cell, a group of plant cells, a plant part (including seeds), or a plant piece (e.g., from a protoplast, callus, or tissue part).

[0205] As used herein, the term "heterologous" refers to a nucleic acid sequence linked to a nucleic acid sequence to which it is not naturally linked.

[0206] "Hybridization" is used herein to denote the pairing of complementary nucleotide sequences to produce a DNA-DNA hybrid or a DNA-RNA hybrid. Complementary base sequences are those sequences that are related by the base-pairing rules. In DNA, A pairs with T and C pairs with G. In RNA, U pairs with A and C pairs with G. In this regard, the terms "match" and "mismatch" as used herein refer to the hybridization potential of paired nucleotides in complementary nucleic acid strands. Matched nucleotides hybridize efficiently, such as the classical A-T and G-C base pair mentioned above. Mismatches are other combinations of nucleotides that do not hybridize efficiently.

[0207] By "isolated" is meant material that is substantially or essentially free from components that normally accompany it in its native state. For example, an "isolated polynucleotide", as used herein, refers to a polynucleotide, which has been purified from the sequences which flank it in a naturally-occurring state, *e.g.*, a DNA fragment which has been removed from the sequences that are normally adjacent to the fragment. Suitably, an "isolated" polynucleotide is free of sequences (*e.g.*, protein encoding sequences) that naturally flank the polynucleotide (i.e., sequences located at the 5' and 3' ends of the polynucleotide) in the genomic DNA of the organism from which the polynucleotide was derived. For example, in various embodiments, an isolated promoter polynucleotide can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the polynucleotide in genomic DNA of the cell from which the polynucleotide was derived.

[0208] By "marker gene" is meant a gene that imparts a distinct phenotype to cells expressing the marker gene and thus allows such transformed cells to be distinguished from cells that do not have the marker. A selectable marker gene confers a trait for which one can 'select' based on resistance to a selective agent (*e.g.*, a herbicide, antibiotic, radiation, heat, or other treatment damaging to untransformed cells). A screenable marker gene (or reporter gene) confers a trait that one can identify through observation or testing, *i.e.*, by 'screening' (*e.g.*, β-glucuronidase, neomycin phosphotransferase II, luciferase, or other enzyme activity not present in untransformed cells).

[0209] As used herein, a "naturally-occurring" nucleic acid molecule refers to a RNA or DNA molecule having a nucleotide sequence that occurs in nature. For example a naturally-occurring nucleic acid molecule can encode a protein that occurs in nature.

[0210] By "obtained from" is meant that a sample such as, for example, a nucleic acid extract is isolated from, or derived from, a particular source. For instance, the nucleic acid extract may be obtained from tissue isolated directly from a host plant.

[0211] The term "oligonucleotide" as used herein refers to a polymer composed of a multiplicity of nucleotide residues (deoxyribonucleotides or ribonucleotides, or related structural variants or synthetic analogues thereof, including nucleotides with modified or substituted sugar groups and the like) linked via phosphodiester bonds (or related structural variants or synthetic analogues thereof). Thus, while the term "oligonucleotide" typically refers to a nucleotide polymer in which the nucleotide residues and linkages between them are naturally-occurring, it will be understood that the term also includes within its scope various analogues including, but not restricted to, peptide nucleic acids (PNAs), phosphorothioate,

phosphorodithioate, phophoroselenoate, phosphorodiselenoate, phosphoroanilothioate, phosphoraniladate, phosphoroamidate, methyl phosphonates, 2-O-methyl ribonucleic acids, and the like. The exact size of the molecule can vary depending on the particular application. Oligonucleotides are a polynucleotide subset with 200 bases or fewer in length. Preferably, oligonucleotides are 10 to 60 bases in length and most preferably 12, 13, 14, 15, 16, 17, 18, 19, or 20 to 40 bases in length. Oligonucleotides are usually single stranded, e.g., for probes; although oligonucleotides may be double stranded, e.g., for use in the construction of a variant nucleic acid sequence. Oligonucleotides of the invention can be either sense or antisense oligonucleotides.

[0212] The term "operably connected" or "operably linked" as used herein means placing a transcribable nucleic acid sequence as defined herein under the regulatory control of a promoter sequence, which then controls the transcription and optionally translation of the gene. In the construction of heterologous promoter/ transcribable nucleic acid combinations, it is generally desirable to position the promoter sequence at a distance from the transcription start site that is approximately the same as the distance between the promoter sequence and the gene it controls in its natural setting; *i.e.*, the gene from which the promoter sequence was derived. As is known in the art, some variation in this distance can be accommodated without loss of function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous nucleic acid sequence to be placed under its control is defined by the positioning of the element in its natural setting; *i.e.*, the genes from which it is derived.

[0213] As used herein, "plant" and "differentiated plant" refer to a whole plant or plant part containing differentiated plant cell types, tissues and/or organ systems. Plantlets and seeds are also included within the meaning of the foregoing terms. Plants included in the invention are any plants amenable to transformation techniques, including angiosperms, gymnosperms, monocotyledons and dicotyledons. In specific embodiments, the plant is a monocotyledonous plant, illustrative examples of which include turf, turf grass, cereal, maize, rice, oat, wheat, barley, orchid, iris, lily, onion, banana, pineapple, sugarcane, sorghum, and palm.

[0214] The term "plant cell" as used herein refers to any plant cell or cell line including protoplasts, gamete-producing cells, and cells which regenerate into whole plants. Plant cells also include cells in plants as well as protoplasts in culture.

[0215] By "plant tissue" is meant differentiated and undifferentiated tissue derived from roots, shoots, pollen, seeds, tumour tissue, such as crown galls, and various forms of aggregations of plant cells in culture, such as embryos and calluses.

[0216] The term "polynucleotide" or "nucleic acid" as used herein designates mRNA, RNA, cRNA, cDNA or DNA. The term typically refers to polymeric form of nucleotides of at least 10 bases in length, either ribonucleotides or deoxynucleotides or a modified form of either type of nucleotide. The term includes single and double stranded forms of DNA.

[0217] The terms "polynucleotide variant" and "variant" refer to polynucleotides displaying substantial sequence identity with a reference polynucleotide sequence or polynucleotides that hybridize with a reference sequence under stringent conditions that are defined hereinafter. These terms also encompass polynucleotides in which one or more nucleotides have been added or deleted, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference polynucleotide whereby the altered polynucleotide retains a biological function or activity of the reference polynucleotide. The terms "polynucleotide variant" and "variant" also include naturally-occurring allelic variants.

[0218] "Polypeptide", "peptide" and "protein" are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues is a synthetic non-naturally-occurring amino acid, such as a chemical analogue of a corresponding naturally-occurring amino acid, as well as to naturally-occurring amino acid polymers.

[0219] By "primer" is meant an oligonucleotide which, when paired with a strand of DNA, is capable of initiating the synthesis of a primer extension product in the presence of a suitable polymerizing agent. The primer is usually single-stranded for maximum efficiency in amplification but can alternatively be double-stranded. A primer must be sufficiently long to prime the synthesis of extension products in the presence of the polymerization agent. The length of the primer depends on many factors, including application, temperature to be employed, template reaction conditions, other reagents, and source of primers. For example, depending on the complexity of the target sequence, the primer may be at least about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, 500, to one base shorter in length than the template sequence at the 3' end of the primer to allow extension of a nucleic acid chain, though the 5' end of the

primer may extend in length beyond the 3' end of the template sequence. In certain embodiments, primers can be large polynucleotides, such as from about 35 nucleotides to several kb or more. Primers can be selected to be "substantially complementary" to the sequence on the template to which it is designed to hybridize and serve as a site for the initiation of synthesis. By "substantially complementary", it is meant that the primer is sufficiently complementary to hybridize with a target polynucleotide. Desirably, the primer contains no mismatches with the template to which it is designed to hybridize but this is not essential. For example, non-complementary nucleotide residues can be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the template. Alternatively, non-complementary nucleotide residues or a stretch of non-complementary nucleotide residues can be interspersed into a primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridize therewith and thereby form a template for synthesis of the extension product of the primer.

[0220] "Probe" refers to a molecule that binds to a specific sequence or sub-sequence or other moiety of another molecule. Unless otherwise indicated, the term "probe" typically refers to a polynucleotide probe that binds to another polynucleotide, often called the "target polynucleotide", through complementary base pairing. Probes can bind target polynucleotides lacking complete sequence complementarity with the probe, depending on the stringency of the hybridization conditions. Probes can be labeled directly or indirectly and include primers within their scope.

[0221] The term "promoter" refers to a nucleic acid which directs expression of another nucleic acid to which it is operably linked, by initiating, regulating or otherwise controlling transcription of the nucleic acid. Promoters usually comprise a TATA box and often a "CAAT" box, capable of directing RNA polymerase to initiate RNA synthesis at the appropriate transcription initiation site for a particular coding sequence. They may additionally comprise other recognition sequences generally positioned upstream or 5' to the TATA box, referred to as upstream promoter elements, which may influence the transcription initiation rate, tissue expression and/or temporal expression of an operably connected nucleic acid sequence. As used herein, a "core promoter" is intended to mean a promoter sequence consisting only of all basal elements needed for transcription initiation, *e.g.*, a TATA box and/or an initiator, without ancillary (*e.g.*, upstream) promoter elements.

[0222] "Constitutive promoter" refers to a promoter that directs transcription in many or all tissues of a plant.

[0223] By "stem-specific promoter" is meant a promoter that preferentially directs transcription in stem tissue of a plant. "Preferentially directs transcription" means that the rate of transcription of an operably linked nucleic acid is higher in the nominated tissue or developmental stage than in another tissue or developmental stage used for comparison.

[0224] The term "recombinant" as used herein refers to a nucleic acid or polypeptide resulting from in vitro manipulation into a form not normally found in nature. As used in the art, "recombinant" usually refers to the products of recombinant DNA technology.

[0225] Terms used to describe sequence relationships between two or more polynucleotides or polypeptides include "reference sequence", "comparison window", "sequence identity", "percentage of sequence identity" and "substantial identity". These relationships are commonly analysed by use of sequence comparison programs such as BESTFIT (Deveraux et al. 1984, Nucleic Acids Research **12**, 387-395) which is incorporated herein by reference. Sequences of a similar or substantially different length may be aligned and compared by insertion of gaps into the alignment, such gaps being determined, for example, by the comparison algorithm used by BESTFIT.

[0226] As used herein, the term "regulatory element" refers to a sequence of DNA, usually, but not always, upstream (5') to the coding sequence of a structural gene, which includes sequences that control the expression of the coding region by providing the recognition for RNA polymerase and/or other factors required for transcription to start at a particular site. An example of a complex regulatory element that provides for the recognition for RNA polymerase or other transcriptional factors to ensure initiation at a particular site is a promoter. A promoter generally comprises a core promoter region, responsible for the initiation of transcription, and optionally other regulatory elements that modify gene expression. It is to be understood that nucleotide sequences, located within introns, or 3' of the coding region sequence may also contribute to the regulation of expression of a coding region of interest. Examples of suitable introns include, but are not limited to, the maize IVS6 intron, or the maize actin intron. A regulatory element may also include those elements located downstream (3') to the site of transcription initiation, or within transcribed regions, or both. In the context of this disclosure, a post-transcriptional regulatory element may include elements that are active following transcription initiation, for example translational and transcriptional enhancers, translational and transcriptional repressors, and mRNA stability determinants.

[0227] A "reference sequence" is at least 6 but frequently 15 to 18 and often at least 25 monomer units, inclusive of nucleotides and amino acid residues, in length. Because two

polynucleotides may each comprise (1) a sequence of nucleotide bases that is similar between the two polynucleotides, and (2) a sequence of nucleotide bases that is divergent between the two polynucleotides, sequence comparisons between two (or more) polynucleotides are typically performed by comparing sequences of the two polynucleotides over a "comparison window" to identify and compare local regions of sequence similarity. A "comparison window" refers to a conceptual segment of at least 6 contiguous positions, usually about 50 to about 100, more usually about 100 to about 150 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. Optimal alignment of sequences for comparison may be conducted by computerized implementations of algorithms (such as EMBOSS programs NEEDLE and WATER accessible on the EMBL-EBI website http://www.ebi.ac.uk/Tools/emboss/align/index.html, and GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA, incorporated herein by reference) or by inspection and the best alignment (i.e., resulting in the highest percentage sequence identity over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul et al., 1997, Nucl. Acids Res. **25**:3389, which is incorporated herein by reference. A detailed discussion of sequence analysis can be found in Unit 19.3 of Ausubel et al., "Current Protocols in Molecular Biology", John Wiley & Sons Inc, 1994-1998, Chapter 15.

[0228] The term "sequence identity" as used herein refers to the extent that sequences are identical on a nucleotide-by-nucleotide basis over a comparison window. Thus, a "percentage of sequence identity" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. The percentage of sequence identity can be determined using the EMBOSS programs NEEDLE or WATER or other algorithms such as those described above. In specific embodiments, "sequence identity" will be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA) using standard defaults as used in the reference manual accompanying the

software, which is incorporated herein by reference in its entirety. In relation to polypeptide sequences, "sequence similarity" refers to the extent that optimally aligned sequences show the same amino acid or conservatively substituted amino acids according to a matrix or algorithm specified in the sequence comparison program, for example the Blossum or PAM matrices used with multiple sequence alignment program Clustal.

[0229] "Stringency" as used herein, refers to the temperature and ionic strength conditions, and presence or absence of certain organic solvents, during hybridization. The higher the stringency, the higher will be the degree of complementarity between immobilized nucleotide sequences and the labeled polynucleotide sequences that remain bound to them following the hybridization procedure.

[0230] "Stringent conditions" refers to temperature and ionic conditions under which only nucleotide sequences having a high frequency of complementary bases will hybridize. The stringency required is nucleotide sequence dependent and depends upon the various components present during hybridization and subsequent washes, and the time allowed for these processes. Generally, in order to maximize the hybridization rate, non-stringent hybridization conditions are selected; about 20 to 25° C lower than the thermal melting point (T_m). The T_m is the temperature at which 50% of specific target sequence hybridizes to a perfectly complementary probe in solution at a defined ionic strength and pH. Generally, in order to require at least about 85% nucleotide complementarity of hybridized sequences, highly stringent washing conditions are selected to be about 5 to 15° C lower than the T_m. In order to require at least about 70% nucleotide complementarity of hybridized sequences, moderately stringent washing conditions are selected to be about 15 to 30° C lower than the T_m. Highly permissive (low stringency) washing conditions may be as low as 50° C below the T_m, allowing a high level of mis-matching between hybridized sequences. Those skilled in the art will recognise that other physical and chemical parameters in the hybridization and wash stages can also be altered to affect the outcome of a detectable hybridization signal from a specific level of similarity between target and probe sequences. Other examples of stringency conditions are described in Section 4 below.

[0231] The term "transcribable DNA sequence" or "transcribed DNA sequence", excludes the non-transcribed regulatory sequence that drives transcription. Depending on the aspect of the invention, the transcribable sequence may be derived in whole or in part from any source known to the art, including a plant, a fungus, an animal, a bacterial genome or episome, eukaryotic, nuclear or plasmid DNA, cDNA, viral DNA or chemically synthesized

DNA. A transcribable sequence may contain one or more modifications in either the coding or the untranslated regions which could affect the biological activity or the chemical structure of the expression product, the rate of expression or the manner of expression control. Such modifications include, but are not limited to, insertions, deletions and substitutions of one or more nucleotides. The transcribable sequence may contain an uninterrupted coding sequence or it may include one or more introns, bound by the appropriate splice junctions. The transcribable sequence may also encode a fusion protein. It is contemplated that introduction into plant tissue of chimeric nucleic acid constructs of the invention will include constructions wherein the transcribable sequence and its promoter are each derived from different species.

[0232] The term "transformation" means alteration of genotype by introduction of genetic material (*e.g.*, the chimeric construct of the present invention) into an organism.

[0233] As used herein, the term "transgenic" or "transformed" with respect to a plant cell, plant part (including seed), plant tissue or plant means a plant cell, plant part, plant tissue or plant which comprises an isolated chimeric DNA construct according to the invention which has been introduced into the genome of a plant cell, plant part, plant tissue or plant.

[0234] By "vector" is meant a polynucleotide molecule, suitably a DNA molecule derived, for example, from a plasmid, bacteriophage, yeast, virus, mammal, avian, reptile or fish into which a polynucleotide can be inserted or cloned. A vector preferably contains one or more unique restriction sites and can be capable of autonomous replication in a defined host cell including a target cell or tissue or a progenitor cell or tissue thereof, or be integrable with the genome of the defined host such that the cloned sequence is reproducible. Accordingly, the vector can be an autonomously replicating vector, i.e., a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, e.g., a linear or closed circular plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector can contain any means for assuring self-replication. Alternatively, the vector can be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. A vector system can comprise a single vector or plasmid, two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vector can also include a selection marker such as an antibiotic resistance gene that can be

used for selection of suitable transformants. Examples of such resistance genes are known to those of skill in the art.

[0235] The terms "wild type," "native" or "non-transgenic" refers to an untransformed plant cell, plant part, plant tissue or plant, *i.e.*, one where the genome has not been altered by the presence of a chimeric nucleic acid construct as defined herein.

2. Abbreviations

[0236] The following abbreviations are used throughout the application:

nt =nucleotide
nts =nucleotides
kb =kilobase(s) or kilobase pair(s)
ORF =open reading frame
SNPs =single nucleotide polymorphisms

3. Transcribable DNA sequences useful for isolating promoter sequences

[0237] The promoter sequences of the present invention were first isolated through their linkage to a transcribable sequence, which comprises an ORF whose sequence is set forth in SEQ ID NO: 83, and which was found to be transcribed at high levels in mature stem tissues of sugarcane (*Saccharum* sp.). In accordance with the present invention, nucleotide sequences that correspond or are complementary to at least a portion of the sequence set forth in SEQ ID NO: 83 may be useful as probes for isolating homologous transcribable sequences from other plants, especially from other sugarcane plants and more broadly other monocotyledonous plants such as cereals, turf, turf grass, rice, orchid, iris, lily, onion, banana, pineapples, sugarcane, sorghum, and palm and to, in turn, permit the isolation of promoter sequences with analogous qualities to those described herein.

[0238] The probes may be used in any suitable screening procedure. For example, a microarray screening procedure, as described below, may be used to identify genes expressed differentially in various tissues. However, it will be understood that the present invention is not restricted to use of any particular method for identifying such differentially expressed genes. For instance, alternative procedures for identifying genes expressed differentially in various tissues include, but are not restricted to: hybridization screening as for example described in patent specification PCT/AU99/01033; cDNA and genomic subtractive hybridization as for example described by Bulman and Neill (1996, In *"Plant Gene Isolation:*

Principles and Practice", G.D. Foster and D. Twell, eds Chichester, UK, Wiley, pp 369-397); multi-probe fluorescent analysis of microscopic cDNA arrays as for example described by Schena (1996 *BioEssays* **18**:427-431); mRNA differential display as for example described by Liang and Pardee (1992, *Science* **257**:967-970) and by Callard *et al.* (1994, *BioTechniques* **16**:1096-1103); computer analysis of mRNA abundance based on frequency of occurrence of identical sequences emerging from large-scale sequencing of cDNA ends (ESTs) as for example taught by Cooke *et al* (1996, *EST and genomic sequencing projects.* In *Plant Gene Isolation: Principles and Practice, supra*, pp. 410-419); or promoter tagging by insertional mutagenesis with promoterless reporter genes as for example disclosed by Lindsey and Topping (1996, *T-DNA-mediated insertional mutagenesis.* In *Plant Gene Isolation: Principles and Practice, supra*, pp. 275-300) and Mudge and Birch (1998, *Austral. J. Plant Physiol.* **25**:637-643).

4. Promoter sequences of the present invention

[0239] The present invention provides promoter sequences useful for expression of transcribable sequences in plants. In specific embodiments, stem-specific promoters for expression of chimeric or heterologous nucleic acid sequences in plants, especially monocotyledonous plants are provided. Representative examples of such promoter sequences may be selected from the sequences set forth in SEQ ID NOS: 63 to 82. Sequence analysis has revealed that these promoter sequences share a common segmental architecture, which can be represented by the following formula:

$$5'\text{-}\Theta\text{–}\Delta\text{–}Z\text{–}\Psi\text{–}\Phi\text{–}\Omega\text{-}3' \qquad (I)$$

in which segment Ω is more conserved across the class of promoter sequences than segment Φ, segment Φ is more conserved than segment Z, segment Z is more conserved than segment Δ, segment Δ is more conserved than segment Θ, and Ψ is an optional spacer or intervening segment, as illustrated in Figure 5. Segment Ω appears to comprise a core promoter region containing a TATA box and (AGGA)CAAT box and it has been experimentally determined that this segment is transcriptionally active in plant cells. In accordance with the present invention, it is predicted that segments Φ, Z, Δ and Θ comprise promoter elements, which preferentially direct transcription in plant sink tissue, including stem tissue. In specific embodiments, transcription is directed in stem tissue of monocotyledonous plants, illustrative examples of which include turf, turf grass, cereal, maize, rice, oat, wheat, barley, orchid, iris, lily, onion, banana, pineapple, sugarcane, sorghum, and palm.

[0240] The promoter sequences of the present invention can be used to prepare biologically active fragments that have promoter activity, to isolate corresponding sequences from other organisms, particularly other plants and more particularly other monocotyledonous plants, or to synthesize synthetic sequences. They can also be used in combination with native or heterologous core promoter regions, control elements or other regulatory sequences to modulate transcription and/or translation.

[0241] Accordingly, the present invention contemplates that biologically-active fragments of any one of SEQ ID NOS: 63 to 82, which comprise less than the entire promoter sequences disclosed herein, may be utilized to drive expression of an operably linked nucleotide sequence of interest, such as a nucleotide sequence encoding a heterologous protein. It is within skill in the art to determine whether such fragments decrease or increase expression levels or alter the nature of expression, *i.e.*, constitutive or inducible expression. Such fragments should retain promoter activity, or the ability to modulate the activity of a coupled core promoter region, particularly the ability to control expression of operably linked nucleotide sequences. For example, the 5' portion of a promoter up to the TATA box near the transcription start site can be deleted without abolishing promoter activity, as described by Opsahl-Sorteberg, H-G. *et al.* (2004, *Gene* **341**:49-58).

[0242] Biologically active fragments of promoters can be readily identified by randomly preparing and assaying deletion mutants of the promoter sequences of the invention (*e.g.*, SEQ ID NOS: 63 to 82). With this strategy, a series of constructs is prepared, wherein each construct contains a different portion of the clone (a subclone), and these constructs are then screened for activity. The activity of a promoter can be determined by standard methods known in the art, as disclosed for example in Medberry *et al.* (1992, *Plant Cell* **4**:185; 1993, *The Plant J.* **3**:619), Sambrook *et al.* (1989, *supra*) and McPherson *et al.* (U.S. Patent No. 5,164,316). For instance, a suitable means for screening for activity is to operably link a deleted promoter construction to a selectable or screenable marker, and to isolate only those cells or tissues or plants which express the marker gene. In this way, a number of different, deleted promoter constructs are identified which still retain the desired, or even enhanced, activity. The smallest segment which is required for activity is thereby identified through comparison of the selected constructs. This segment may then be used for the construction of vectors for the expression of heterologous genes. Alternatively, biologically active fragments may be identified by fusion to a core promoter region that is coupled to a reporter gene, and screening transformed plant cells for reporter gene expression at developmental stages of

interest, as disclosed for example in Puente *et al.* (1996, *EMBO J.* **15**: 3732). The combination of segments most effective for particular developmental patterns of promoter activity is thereby identified through comparison of the selected constructs. These segments may then be used for the construction of vectors for the expression of heterologous genes.

[0243] Based on sequence analysis of the promoter sequences set forth in SEQ ID NO: 62-83, it is predicted that promoter elements positioned upstream or 5' to the TATA box in segment Ω (*e.g.*, promoter elements in any one or more of segments Φ, Ψ, Z, Δ and Θ) influence the transcription initiation rate and/or tissue (*e.g.*, stem-specific) expression of operably connected nucleic acid sequences. Accordingly, the present invention also contemplates fusing any one or more of segments Φ, Ψ, Z, Δ and Θ to a heterologous core promoter (*e.g.*, a core promoter region of any promoter operative in the host organism into which it is desired to express a transcribable sequence of interest) to thereby produce a chimeric promoter sequence. Suitable core promoter sequences are well known in the art and are generally derived from plant-operative promoters, such as but not limited to the CaMV35S, nopaline synthase, ferrodoxin-RolD, maize ubiquitin and rice actin promoters. Reference also may be made to Puente *et al.* (1996, *EMBO J.* **15**: 3732) and Klimyuk *et al.* (1995 *Molecular & General Genetics* **249**: 357), who disclose illustrative plant core promoters. Alternatively the core promoter may correspond to a consensus core promoter sequence. Non-limiting examples of core promoter sequences are set forth in SEQ ID NO: 85-87.

[0244] Representative biologically-active fragments of the promoter sequences set forth in SEQ ID NOS: 63 to 82 may comprise at least about 125, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, or 650 nucleotides, or almost up to the number of nucleotides present in a full-length promoter sequence. Non-limiting examples of such biologically active fragments are set forth in SEQ ID NO: 2 to 11, which correspond to segment Ω, respectively, of alleles A4, A7, A3, A6, A1, A2, A9, A8, A5 and B, as described herein. Longer fragments can be produced by adding corresponding upstream sequences such as but not limited to Φ, Ψ, Z, Δ and Θ. Alternatively, biologically active fragments may include any one or more of segments Φ, Ψ, Z, Δ and Θ, which can be fused to core promoter sequences, to create chimeric promoter sequences with expression patterns, in some embodiments, reflecting the activity of the promoter sequences set forth in SEQ ID NO: 62-83. Illustrative examples of biologically active fragments according to the present invention are listed in Table 1 below.

TABLE 1

Fragment	A4	A7	A3	A6	A1	A2	A9	A8	A5	B
A	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω	Ω
B	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω	Φ–Ω
C	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω	Z–Φ–Ω
D	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω	–	Δ–Z–Φ–Ω	Δ–Z–Φ–Ω
E	Θ–Δ–Z–Φ–Ω	Θ–Δ–Z–Φ–Ω	–	Θ–Δ–Z–Φ–Ω	Θ–Δ–Z–Φ–Ω	Θ–Δ–Z–Φ–Ω	Θ–Δ–Z–Φ–Ω	–	Θ–Δ–Z–Φ–Ω	Θ–Δ–Z–Φ–Ω
F			Ψ–Φ–Ω						Ψ–Φ–Ω	Ψ–Φ–Ω
G			Z–Ψ–Φ–Ω						Z–Ψ–Φ–Ω	Z–Ψ–Φ–Ω
H			Δ–Z–Ψ–Φ–Ω						Δ–Z–Ψ–Φ–Ω	Δ–Z–Ψ–Φ–Ω
I									Θ–Δ–Z–Ψ–Φ–Ω	Θ–Δ–Z–Ψ–Φ–Ω

F	Φ	Φ	Φ	Φ	Φ	Φ	Φ	Φ	Φ	Φ
G	Z–Φ	Z–Φ	Z–Φ	Z–Φ	Z–Φ	Z–Φ	Z–Φ	Z–Φ	Z–Φ	Z–Φ
H	Δ–Z–Φ	Δ–Z–Φ	Δ–Z–Φ	Δ–Z–Φ	Δ–Z–Φ	Δ–Z–Φ	Δ–Z–Φ		Δ–Z–Φ	Δ–Z–Φ
I	Θ–Δ–Z–Φ	Θ–Δ–Z–Φ		Θ–Δ–Z–Φ	Θ–Δ–Z–Φ	Θ–Δ–Z–Φ	Θ–Δ–Z–Φ		Θ–Δ–Z–Φ	Θ–Δ–Z–Φ

[0245] The present invention also encompasses promoter sequence variants that are substantially complementary to a reference sequence. Such variants are identified by blotting techniques that include a step whereby nucleic acids are immobilized on a matrix (preferably a synthetic membrane such as nitrocellulose), followed by a hybridization step, and a detection step. Southern blotting is used to identify a complementary DNA sequence; northern blotting is used to identify a complementary RNA sequence. Dot blotting and slot blotting can be used to identify complementary DNA/DNA, DNA/RNA or RNA/RNA polynucleotide sequences. Such techniques are well known by those skilled in the art, and have been described in Ausubel *et al.* (1994-1998, *supra*) at pages 2.9.1 through 2.9.20.

[0246] According to such methods, Southern blotting involves separating DNA molecules according to size by gel electrophoresis, transferring the size-separated DNA to a synthetic membrane, and hybridizing the membrane-bound DNA to a complementary nucleotide sequence labeled radioactively, enzymatically or fluorochromatically. In dot blotting and slot blotting, DNA samples are directly applied to a synthetic membrane prior to hybridization as above.

[0247] An alternative blotting step is used when identifying complementary polynucleotides in a cDNA or genomic DNA library, such as through the process of plaque or colony hybridization. A typical example of this procedure is described in Sambrook *et al.* ("Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, 1989) Chapters 8-12.

[0248] Typically, the following general procedure can be used to determine hybridization conditions. Polynucleotides are blotted/transferred to a synthetic membrane, as described above. A reference polynucleotide such as a polynucleotide of the invention is labeled as described above, and the ability of this labeled polynucleotide to hybridize with an immobilized polynucleotide is analyzed.

[0249] A skilled person will recognize that a number of factors influence hybridization. The specific activity of radioactively labeled polynucleotide sequence should typically be greater than or equal to about 10^8 dpm/mg to provide a detectable signal. A radiolabeled nucleotide sequence of specific activity 10^8 to 10^9 dpm/mg can detect approximately 0.5 pg of DNA. It is well known in the art that sufficient DNA must be immobilized on the membrane to permit detection. It is desirable to have excess immobilized DNA, usually 10 μg. Adding an inert polymer such as 10% (w/v) dextran sulfate (MW

500,000) or polyethylene glycol 6000 during hybridization can also increase the sensitivity of hybridization (see Ausubel *supra* at 2.10.10).

[0250] To achieve meaningful results from hybridization between a polynucleotide immobilized on a membrane and a labeled polynucleotide, a sufficient amount of the labeled polynucleotide must be hybridized to the immobilized polynucleotide following washing. Washing ensures that the labeled polynucleotide is hybridized only to the immobilized polynucleotide with a desired degree of complementarity to the labeled polynucleotide.

[0251] It will be understood that polynucleotide sequence variants according to the invention will hybridize to a reference polynucleotide under at least low stringency conditions. Reference herein to low stringency conditions include and encompass from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridization at 42°C, and at least about 1 M to at least about 2 M salt for washing at 42°C. Low stringency conditions also may include 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (i) 2 × SSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at room temperature.

[0252] Suitably, the polynucleotide variants hybridize to a reference polynucleotide under at least medium stringency conditions. Medium stringency conditions include and encompass from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridization at 42°C, and at least about 0.1 M to at least about 0.2 M salt for washing at 55°C. Medium stringency conditions also may include 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (i) 2 × SSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at 60-65°C.

[0253] In specific embodiments, the polynucleotide variants hybridize to a reference polynucleotide under high stringency conditions. High stringency conditions include and encompass from at least about 31% v/v to at least about 50% v/v formamide and from about 0.01 M to about 0.15 M salt for hybridization at 42°C, and about 0.01 M to about 0.02 M salt for washing at 55°C. High stringency conditions also may include 1% BSA, 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (i) 0.2 × SSC, 0.1% SDS; or (ii) 0.5% BSA, 1mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 1% SDS for washing at a temperature in excess of 65°C.

[0254] Other stringent conditions are well known in the art. A skilled addressee will recognize that various factors can be manipulated to optimize the specificity of the hybridization. Optimization of the stringency of the final washes can serve to ensure a high degree of hybridization. For detailed examples, see Ausubel *et al.*, *supra* at pages 2.10.1 to 2.10.16 and Sambrook *et al.* (1989, *supra*) at sections 1.101 to 1.104, which are incorporated herein by reference.

[0255] While stringent washes are typically carried out at temperatures from about 42°C to 68°C, one skilled in the art will appreciate that other temperatures may be suitable for stringent conditions. Maximum hybridization typically occurs at about 20°C to 25°C below the T_m for formation of a DNA-DNA hybrid. It is well known in the art that the T_m is the melting temperature, or temperature at which two complementary polynucleotide sequences dissociate. Methods for estimating T_m based on the length and sequence composition of duplex DNA and the composition of the hybridization solution are well known in the art (see Ausubel *et al.*, *supra* at page 2.10.8).

[0256] In general, the T_m of a perfectly matched duplex of DNA may be predicted as an approximation by the formula:

[0257] $T_m = 81.5 + 16.6 (\log_{10} M) + 0.41 (\%G+C) - 0.63 (\%\text{ formamide}) - (600/\text{length})$

[0258] wherein: M is the concentration of Na^+, preferably in the range of 0.01 molar to 0.4 molar; %G+C is the sum of guanosine and cytosine bases as a percentage of the total number of bases, within the range between 30% and 75% G+C; % formamide is the percent formamide concentration by volume; length is the number of base pairs in the DNA duplex.

[0259] The T_m of a duplex DNA decreases by approximately 1°C with every increase of 1% in the number of randomly mismatched base pairs. Washing is generally carried out at $T_m - 15$ °C for high stringency, or $T_m - 30$ °C for moderate stringency.

[0260] In a preferred hybridization procedure, a membrane (*e.g.*, a nitrocellulose membrane or a nylon membrane) containing immobilized DNA is hybridized overnight at 42°C in a hybridization buffer (50% deionized formamide, 5 × SSC, 5 × Denhardt's solution (0.1% ficoll, 0.1% polyvinylpyrollidone and 0.1% bovine serum albumin), 0.1% SDS and 200 mg/mL denatured salmon sperm DNA) containing labeled probe. The membrane is then subjected to two sequential medium stringency washes (*i.e.*, 2 × SSC, 0.1% SDS for 15 min at

45°C, followed by 2 × SSC, 0.1% SDS for 15 min at 50°C), followed by two sequential higher stringency washes (*i.e.*, 0.2 × SSC, 0.1% SDS for 12 min at 55°C followed by 0.2 × SSC and 0.1%SDS solution for 12 min at 65-68°C).

[0261] Methods for detecting a labeled polynucleotide hybridized to an immobilized polynucleotide are well known to practitioners in the art. Such methods include autoradiography, phosphorimaging, chemiluminescent, fluorescent and colorimetric detection.

[0262] In general, variants will comprise regions that show at least 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99% sequence identity over a reference promoter sequence of identical size ("comparison window") or when compared to an aligned sequence in which the alignment is performed by a computer program known in the art, as described in the definitions section on sequence relationships. What constitutes suitable variants may be determined by conventional techniques. For example, polynucleotides according to any one of SEQ ID NO: 62-82 can be mutated using random mutagenesis (*e.g.*, transposon mutagenesis), oligonucleotide-mediated (or site-directed) mutagenesis, PCR mutagenesis and cassette mutagenesis of an earlier prepared variant or non-variant version of an isolated natural promoter according to the invention. Thus, in addition to the unmodified promoter sequences set forth in SEQ ID NO: 63-82, the current invention encompasses functional variants of these sequences. For example, based on the information provided herein with respect to functional elements, one of skill in the art could derive functional promoters by deleting or modifying non-essential portions of the sequence without disrupting the essential functional elements. Such mutants may potentially have enhanced or altered function relative to the native sequence or alternatively, may be silent with regard to function.

[0263] Oligonucleotide-mediated mutagenesis is a preferred method for preparing nucleotide substitution variants of a promoter of the invention. This technique is well known in the art as, for example, described by Adelman *et al.* (1983, *DNA* **2**:183). Briefly, promoter DNA is altered by hybridizing an oligonucleotide encoding the desired mutation to a template DNA, where the template is the single-stranded form of a plasmid or bacteriophage containing the unaltered or native DNA sequence of the promoter of interest. After hybridization, a DNA polymerase is used to synthesize an entire second complementary strand of the template that will thus incorporate the oligonucleotide primer, and will code for the selected alteration in the promoter of interest.

[0264] Generally, oligonucleotides of at least 25 nucleotides in length are used. An optimal oligonucleotide will have 12 to 15 nucleotides that are completely complementary to the template on either side of the nucleotide(s) coding for the mutation. This ensures that the oligonucleotide will hybridize properly to the single-stranded DNA template molecule.

[0265] The DNA template can be generated by those vectors that are either derived from bacteriophage M13 vectors, or those vectors that contain a single-stranded phage origin of replication as described by Viera *et al.* (1987, *Methods Enzymol.* **153**:3). Thus, the DNA that is to be mutated may be inserted into one of the vectors to generate single-stranded template. Production of single-stranded template is described, for example, in Sections 4.21-4.41 of Sambrook *et al.* (1989, *supra*).

[0266] Alternatively, the single-stranded template may be generated by denaturing double-stranded plasmid (or other DNA) using standard techniques.

[0267] For alteration of the native DNA sequence, the oligonucleotide is hybridized to the single-stranded template under suitable hybridization conditions. A DNA polymerizing enzyme, usually the Klenow fragment of DNA polymerase I, is then added to synthesize the complementary strand of the template using the oligonucleotide as a primer for synthesis. A heteroduplex molecule is thus formed such that one strand of DNA encodes the mutated form of the promoter under test, and the other strand (the original template) encodes the native unaltered sequence of the promoter under test. This heteroduplex molecule is then transformed into a suitable host cell, usually a prokaryote such as *E. coli*. After the cells are grown, they are plated onto agarose plates and screened using the oligonucleotide primer having a detectable label to identify the bacterial colonies having the mutated DNA. The resultant mutated DNA fragments are then cloned into suitable expression hosts such as *E. coli* using conventional technology and clones that retain the desired promoter activity are detected. Where the clones have been derived using random mutagenesis techniques, positive clones would have to be sequenced in order to detect the mutation.

[0268] Alternatively, linker-scanning mutagenesis of DNA may be used to introduce clusters of point mutations throughout a sequence of interest that has been cloned into a plasmid vector. For example, reference may be made to Ausubel *et al.*, *supra*, (in particular, Chapter 8.4, incorporated herein by reference) which describes a first protocol that uses complementary oligonucleotides and requires a unique restriction site adjacent to the region that is to be mutagenized. A nested series of deletion mutations is first generated in the region. A pair of complementary oligonucleotides is synthesized to fill in the gap in the

sequence of interest between the linker at the deletion endpoint and the nearby restriction site. The linker sequence actually provides the desired clusters of point mutations as it is moved or "scanned" across the region by its position at the varied endpoints of the deletion mutation series. An alternate protocol is also described by Ausubel *et al.*, *supra*, which makes use of site directed mutagenesis procedures to introduce small clusters of point mutations throughout the target region. Briefly, mutations are introduced into a sequence by annealing a synthetic oligonucleotide containing one or more mismatches to the sequence of interest cloned into a single-stranded M13 vector. This template is grown in an *Escherichia coli dut⁻ ung⁻* strain, which allows the incorporation of uracil into the template strand. The oligonucleotide is annealed to the template and extended with T4 DNA polymerase to create a double-stranded heteroduplex. Finally, the heteroduplex is introduced into a wild-type *E. coli* strain, which will prevent replication of the template strand due to the presence of apurinic sites (generated where uracil is incorporated), thereby resulting in plaques containing only mutated DNA.

[0269] Region-specific mutagenesis and directed mutagenesis using PCR may also be employed to construct promoter variants according to the invention. In this regard, reference may be made, for example, to Ausubel *et al.*, *supra*, in particular Chapters 8.2A and 8.5.

[0270] The activity of promoter sequence variants thus obtained can be determined by conventional methods known to those of skill in the art, illustrative examples of which are discussed above.

5. Chimeric nucleic acid constructs and expression vectors

[0271] An isolated nucleic acid promoter sequence or variant according to the invention can be fused to a heterologous nucleic acid to form a chimeric construct. The heterologous nucleic acid may be a foreign or endogenous DNA sequence. For the purposes of transformation and expression of the heterologous nucleic acid in plants, it is generally desirable that the chimeric construct includes regulatory sequences which influence expression of the heterologous nucleic acid in plants. Suitably, the chimeric construct is present in an expression vector which includes regulatory sequences that enable selective propagation in bacteria.

5.1 3' Non-translated region

[0272] A 3' non-translated sequence refers to that portion of a gene comprising a DNA segment that contains a polyadenylation signal and any other regulatory signals capable

of effecting mRNA processing or gene expression. The polyadenylation signal is characterized by effecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. Polyadenylation signals are commonly recognized by the presence of homology to the canonical form 5' AATAAA-3' although variations are not uncommon.

[0273] The 3' non-translated regulatory DNA sequence typically includes from about 50 to 1,000 base pairs and contains plant transcriptional and translational termination sequences. Examples of suitable 3' non-translated sequences are the 3' transcribed non-translated regions containing a polyadenylation signal from the nopaline synthase (*nos*) gene of *Agrobacterium tumefaciens* (Bevan *et al.*, 1983, *Nucl. Acid Res.*, **11**:369) and the terminator for the T7 transcript from the octopine synthase gene of *Agrobacterium tumefaciens*. Alternatively, suitable 3' non-translated sequences may be derived from plant genes such as the 3' end of the protease inhibitor I or II genes from potato or tomato, the soybean storage protein genes and the pea E9 small subunit of the ribulose-1,5-bisphosphate carboxylase (ssRUBISCO) gene, although other 3' elements known to those of skill in the art can also be employed. Alternatively, 3' non-translated regulatory sequences can be obtained *de novo* as, for example, described by An (1987, *Methods in Enzymology*, **153**:292), which is incorporated herein by reference.

5.2 Optional sequences

[0274] The chimeric construct of the present invention can further include enhancers, either translation or transcription enhancers, as may be required. These enhancer regions are well known to persons skilled in the art, and can include an ATG initiation codon and adjacent sequences. The initiation codon must be in phase with the reading frame of the coding sequence relating to the foreign or endogenous DNA sequence to ensure translation of the entire sequence. The translation control signals and initiation codons can be of a variety of origins, both natural and synthetic. Translational initiation regions may be provided from the source of the transcriptional initiation region, or from the foreign or endogenous DNA sequence. The sequence can also be derived from the source of the promoter selected to drive transcription, and can be specifically modified so as to increase translation of the mRNA.

[0275] Examples of transcriptional enhancers include, but are not restricted to, elements from the CaMV 35S promoter and octopine synthase genes as for example described by Last *et al.* (U.S. Patent No. 5,290,924, which is incorporated herein by reference). It is proposed that the use of an enhancer element such as the *ocs* element, and particularly

multiple copies of the element, will act to increase the level of transcription from adjacent promoters when applied in the context of plant transformation.

[0276] As the DNA sequence inserted between the transcription initiation site and the start of the coding sequence, *i.e.*, the untranslated leader sequence, can influence gene expression, one can also employ a particular leader sequence. Preferred leader sequences include those that comprise sequences selected to direct optimum expression of the foreign or endogenous DNA sequence. For example, such leader sequences include a preferred consensus sequence which can increase or maintain mRNA stability and prevent inappropriate initiation of translation as for example described by Joshi (1987, *Nucl. Acid Res.*, **15**:6643), which is incorporated herein by reference. However, other leader sequences, *e.g.*, the leader sequence of RTBV, have a high degree of secondary structure that is expected to decrease mRNA stability and/or decrease translation of the mRNA. Thus, leader sequences (i) that do not have a high degree of secondary structure, (ii) that have a high degree of secondary structure where the secondary structure does not inhibit mRNA stability and/or decrease translation, or (iii) that are derived from genes that are highly expressed in plants, will be most preferred.

[0277] Regulatory elements such as the sucrose synthase intron as, for example, described by Vasil *et al.* (1989, *Plant Physiol.*, **91**:5175), the Adh intron I as, for example, described by Callis *et al.* (1987, *Genes Develop.*, II), or the TMV omega element as, for example, described by Gallie *et al.* (1989, *The Plant Cell*, **1**:301) can also be included where desired. Other such regulatory elements useful in the practice of the invention are known to those of skill in the art.

[0278] Additionally, targeting sequences may be employed to target a protein product of the foreign or endogenous DNA sequence to an intracellular compartment within plant cells or to the extracellular environment. For example, a DNA sequence encoding a transit or signal peptide sequence may be operably linked to a sequence encoding a desired protein such that, when translated, the transit or signal peptide can transport the protein to a particular intracellular or extracellular destination, respectively, and can then be post-translationally removed. Transit or signal peptides act by facilitating the transport of proteins through intracellular membranes, *e.g.*, vacuole, vesicle, plastid and mitochondrial membranes, whereas signal peptides direct proteins through the extracellular membrane. For example, the transit or signal peptide can direct a desired protein to a particular organelle such as a plastid (*e.g.*, a chloroplast), rather than to the cytoplasm. Thus, the chimeric DNA construct can

further comprise a plastid transit peptide encoding DNA sequence operably linked between a promoter region or promoter variant according to the invention and the foreign or endogenous DNA sequence. For example, reference may be made to Heijne *et al.* (1989, *Eur. J. Biochem.*, **180**:535) and Keegstra *et al.* (1989, *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, **40**:471), which are incorporated herein by reference.

[0279] The chimeric construct can also be introduced into a vector, such as a plasmid. Plasmid vectors include additional DNA sequences that provide for easy selection, amplification, and transformation of the expression cassette in prokaryotic and eukaryotic cells, *e.g.*, pUC-derived vectors, pSK-derived vectors, pGEM-derived vectors, pSP-derived vectors, or pBS-derived vectors. Additional DNA sequences include origins of replication to provide for autonomous replication of the vector, selectable marker genes, preferably encoding antibiotic or herbicide resistance, unique multiple cloning sites providing for multiple sites to insert DNA sequences or genes encoded in the chimeric construct, and sequences that enhance transformation of prokaryotic and eukaryotic cells.

[0280] In general, the vector suitably contains an element(s) that permits stable integration of the vector into the host cell genome or autonomous replication of the vector in the cell independent of the genome of the cell. The vector may be integrated into the host cell genome when introduced into a host cell. For integration, the vector may rely on the foreign or endogenous DNA sequence or any other element of the vector for stable integration of the vector into the genome by homologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location in the chromosome. To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a sufficient number of nucleic acids, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences.

[0281] For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are the origins of replication of plasmids pBR322,

pUC19, pACYC177, and pACYC184 permitting replication in *E. coli*, and pUB110, pE194, pTA1060, and pAM.beta.1 permitting replication in *Bacillus*. The origin of replication may be one having a mutation to make its function temperature-sensitive in a *Bacillus* cell (see, *e.g.*, Ehrlich, 1978, *Proc. Natl. Acad. Sci. USA* **75**:1433).

5.3 Marker genes

[0282] To facilitate identification of transformants, the chimeric nucleic acid construct desirably comprises a selectable or screenable marker gene as, or in addition to, the expressible foreign or endogenous DNA sequence. The actual choice of a marker is not crucial as long as it is functional (*i.e.*, selective) in combination with the plant cells of choice. The marker gene and the foreign or endogenous DNA sequence of interest do not have to be linked, since co-transformation of unlinked genes as, for example, described in U.S. Pat. No. 4,399,216 is also an efficient process in plant transformation.

[0283] Included within the terms selectable or screenable marker genes are genes that encode a "secretable marker" whose secretion can be detected as a means of identifying or selecting for transformed cells. Examples include markers that encode a secretable antigen that can be identified by antibody interaction, or secretable enzymes that can be detected by their catalytic activity. Secretable proteins include, but are not restricted to, proteins that are inserted or trapped in the cell wall (*e.g.*, proteins that include a leader sequence such as that found in the expression unit of extensin or tobacco PR-S); small, diffusible proteins detectable, *e.g.* by ELISA; and small active enzymes detectable in extracellular solution (*e.g.*, α-amylase, β-lactamase, phosphinothricin acetyltransferase).

5.3.1 Selectable markers

[0284] Examples of bacterial selectable markers are the *dal* genes from *Bacillus subtilis* or *Bacillus licheniformis*, or markers that confer antibiotic resistance such as ampicillin, kanamycin, erythromycin, chloramphenicol or tetracycline resistance. Exemplary selectable markers for selection of plant transformants include, but are not limited to, a *hyg* gene which encodes hygromycin B resistance; a neomycin phosphotransferase (*neo*) gene conferring resistance to kanamycin, paromomycin, G418 and the like as, for example, described by Potrykus *et al.* (1985, *Mol. Gen. Genet.* **199**:183); a glutathione-S-transferase gene from rat liver conferring resistance to glutathione derived herbicides as, for example, described in EP-A 256 223; a glutamine synthetase gene conferring, upon overexpression, resistance to glutamine synthetase inhibitors such as phosphinothricin as, for example,

described WO87/05327, an acetyl transferase gene from *Streptomyces viridochromogenes* conferring resistance to the selective agent phosphinothricin as, for example, described in EP-A 275 957, a gene encoding a 5-enolshikimate-3-phosphate synthase (EPSPS) conferring tolerance to N-phosphonomethylglycine as, for example, described by Hinchee *et al.* (1988, *Biotech.*, **6**:915), a *bar* gene conferring resistance against bialaphos as, for example, described in WO91/02071; a nitrilase gene such as *bxn* from *Klebsiella ozaenae* which confers resistance to bromoxynil (Stalker *et al.*, 1988, *Science*, **242**:419); a dihydrofolate reductase (DHFR) gene conferring resistance to methotrexate (Thillet *et al.*, 1988, *J. Biol. Chem.*, **263:**12500); a mutant acetolactate synthase gene (ALS), which confers resistance to imidazolinone, sulfonylurea or other ALS-inhibiting chemicals (EP-A-154 204); a mutated anthranilate synthase gene that confers resistance to 5-methyl tryptophan; or a dalapon dehalogenase gene that confers resistance to the herbicide.

5.3.2 Screenable markers

[0285] Representative screenable markers include, but are not limited to, a *uidA* gene encoding a β-glucuronidase (GUS) enzyme for which various chromogenic substrates are known; a β-galactosidase gene encoding an enzyme for which chromogenic substrates are known; an aequorin gene (Prasher *et al.*, 1985, *Biochem. Biophys. Res. Comm.*, **126**:1259), which may be employed in calcium-sensitive bioluminescence detection; a green fluorescent protein gene (Niedz *et al.*, 1995 *Plant Cell Reports*, **14**:403); a luciferase (*luc*) gene (Ow *et al.*, 1986, *Science*, **234**:856), which allows for bioluminescence detection; a β-lactamase gene (Sutcliffe, 1978, *Proc. Natl. Acad. Sci. USA* **75**:3737), which encodes an enzyme for which various chromogenic substrates are known (*e.g.*, PADAC, a chromogenic cephalosporin); an R-locus gene, encoding a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.*, 1988, in *Chromosome Structure and Function*, pp. 263-282); an α-amylase gene (Ikuta *et al.*, 1990, *Biotech.*, **8**:241); a tyrosinase gene (Katz *et al.*, 1983, *J. Gen. Microbiol.*, **129**:2703) which encodes an enzyme capable of oxidizing tyrosine to dopa and dopaquinone which in turn condenses to form the easily detectable compound melanin; or a *xylE* gene (Zukowsky *et al.*, 1983, *Proc. Natl. Acad. Sci. USA* **80**:1101), which encodes a catechol dioxygenase that can convert chromogenic catechols.

6. Uses of the promoter sequences of the invention

[0286] The isolated promoters sequences of the invention may be used, *inter alia*, to drive expression of a foreign or endogenous nucleic acid sequence. Illustrative agronomic

properties encoded by the foreign or endogenous sequence include, but are not limited to, increased yield of a desired endogenous plant component, production of additional compounds by reactions involving endogenous plant components, decreased production of an undesired plant component, insect resistance or tolerance, herbicide resistance or tolerance, disease resistance or tolerance, tolerance to other stresses (*e.g.* drought, salinity, cold).

[0287] The foreign or endogenous nucleic acid sequence may comprise a region transcribed into a molecule that modulates the expression of a corresponding target gene. The molecule may be an antisense RNA or a ribozyme or other transcript aimed at downregulation of expression of the corresponding target gene.

[0288] Anti-sense regulation, co-suppression and the use of ribozymes and hairpin RNA in plants are well known in the art. However, the skilled person is referred to United States Patent 5,759,829 for an example of antisense technology; and to U. S. Patent 5,283,184, U. S. Patent 5,686,649, and WIPO PCT specification WO9853083 for examples of co-suppression technology; and to U.S. patent 5,707,835, U.S. patent 5,747,335 and U.S. patent 5,840,874 which each provide examples of ribozyme technology; and to USA patent application 2008/0104732 for examples of hairpin RNA technology for RNAi induction. Each of these patent documents is incorporated herein by reference.

[0289] Alternatively, the foreign or endogenous nucleic acid sequence may encode a molecule which is readily detectable or measurable, *e.g.* β-glucuronidase or luciferase; a selectable product, *e.g.*, neomycin phosphotransferase (nptII) conferring resistance to aminoglycosidic antibiotics such as geneticin and paramomycin; a product conferring herbicide tolerance, *e.g.* glyphosate resistance or glufosinate resistance; a product affecting starch biosynthesis or modification *e.g.* starch branching enzyme, starch synthases, ADP-glucose pyrophosphorylase; a product involved in fatty acid biosynthesis, *e.g.* desaturase or hydroxylase; a product conferring insect resistance, *e.g.* crystal toxin protein of *Bacillus thuringiensis*; a product conferring viral resistance, *e.g.* viral coat protein; a product conferring fungal resistance, *e.g.* chitinase, β-1,3-glucanase or phytoalexin; a product altering sucrose metabolism, *e.g.* invertase or sucrose synthase; a gene encoding valuable pharmaceuticals, *e.g.* antibiotics, secondary metabolites, pharmaceutical peptides or vaccines.

7. Introduction of chimeric genes into plant cells

[0290] There are many plant transformation techniques well known to workers in the art, and new techniques are continually becoming known. The particular choice of a

transformation technology will be determined by its efficiency to transform certain plant species as well as the experience and preference of the person practising the invention with a particular methodology of choice. It will be apparent to the skilled person that the particular choice of a transformation system to introduce a chimeric DNA construct into plant cells is not essential to or a limitation of the invention, provided it achieves an acceptable level of nucleic acid transfer. Guidance in the practical implementation of transformation systems for plant improvement is provided by Birch (1997, *Annu. Rev. Plant Physiol. Plant Molec. Biol.* **48**: 297-326), which is incorporated herein by reference.

[0291] In principle both dicotyledonous and monocotyledonous plants that are amenable to transformation, can be modified by introducing a chimeric DNA construct according to the invention into a recipient cell and growing a new plant that harbors and expresses the foreign or endogenous DNA sequence.

[0292] Introduction and expression of foreign or chimeric nucleic acid sequences in dicotyledonous (broad-leafed) plants such as tobacco, potato and alfalfa has been shown to be possible using the T-DNA of the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* (See, for example, Umbeck, U.S. Patent No. 5,004,863, and International application PCT/US93/02480). A construct of the invention may be introduced into a plant cell utilizing *A. tumefaciens* containing the Ti plasmid. In using an *A. tumefaciens* culture as a transformation vehicle, it is most advantageous to use a non-oncogenic strain of the *Agrobacterium* as the vector carrier so that normal non-oncogenic differentiation of the transformed tissues is possible. It is preferred that the *Agrobacterium* harbors a binary Ti plasmid system. Such a binary system comprises (1) a first Ti plasmid having a virulence region essential for the introduction of transfer DNA (T-DNA) into plants, and (2) a chimeric plasmid. The chimeric plasmid contains at least one border region of the T-DNA region of a wild-type Ti plasmid flanking the nucleic acid to be transferred. Binary Ti plasmid systems have been shown effective to transform plant cells as, for example, described by De Framond (1983, *Biotechnology*, **1**:262) and Hoekema *et al.* (1983, *Nature*, **303**:179). Such a binary system is preferred *inter alia* because it does not require integration into the Ti plasmid in *Agrobacterium*.

[0293] Methods involving the use of *Agrobacterium* include, but are not limited to: (a) co-cultivation of *Agrobacterium* with cultured isolated protoplasts; (b) transformation of plant cells or tissues with *Agrobacterium*; or (c) transformation of seeds, apices or meristems with *Agrobacterium*.

[0294] Recently, rice, corn, pineapple and sugarcane, which are monocots, have been shown to be susceptible to transformation by *Agrobacterium,* for example as described in United States Patent No. 6,037,522, International Publication WO99/36637 and Arencibia *et al.* (1998, *Transgenic Res.* 7:213). However, some monocot crop plants have not yet been successfully transformed using *Agrobacterium*-mediated transformation. The Ti plasmid, however, may be manipulated in the future to act as a vector for these other monocot plants. Additionally, using the Ti plasmid as a model system, it may be possible to artificially construct transformation vectors for these plants. Ti plasmids might also be introduced into monocot plants by artificial methods such as microinjection, or fusion between monocot protoplasts and bacterial spheroplasts containing the T-region, which can then be integrated into the plant nuclear DNA.

[0295] In addition, gene transfer can be accomplished by *in situ* transformation by *Agrobacterium*, as described by Bechtold *et al.* (1993, *C.R. Acad. Sci. Paris*, **316**:1194). This approach is based on the vacuum infiltration of a suspension of *Agrobacterium* cells.

[0296] Alternatively, nucleic acids may be introduced using root-inducing (Ri) plasmids of *Agrobacterium* as vectors.

[0297] Cauliflower mosaic virus (CaMV) may also be used as a vector for introducing of exogenous nucleic acids into plant cells (U.S. Pat. No. 4,407,956). CaMV DNA genome is inserted into a parent bacterial plasmid creating a recombinant DNA molecule that can be propagated in bacteria. After cloning, the recombinant plasmid again may be cloned and further modified by introduction of the desired nucleic acid sequence. The modified viral portion of the recombinant plasmid is then excised from the parent bacterial plasmid, and used to inoculate the plant cells or plants.

[0298] Nucleic acids can also be introduced into plant cells by electroporation as, for example, described by Fromm *et al.* (1985, *Proc. Natl. Acad. Sci., U.S.A*, **82**:5824) and Shimamoto *et al.* (1989, *Nature* **338**:274-276). In this technique, plant protoplasts are electroporated in the presence of vectors or nucleic acids containing the relevant nucleic acid sequences. Electrical impulses of high field strength reversibly permeabilize membranes allowing the introduction of nucleic acids. Electroporated plant protoplasts reform the cell wall, divide and form a plant callus.

[0299] Another method for introducing nucleic acids into a plant cell is high velocity ballistic penetration by small particles (also known as particle bombardment or

microprojectile bombardment) with the nucleic acid to be introduced contained either within the matrix of small beads or particles, or on the surface thereof as, for example described by Klein *et al.* (1987, *Nature* **327**:70). Although typically only a single introduction of a new nucleic acid sequence is required, this method particularly provides for multiple introductions.

[0300] Alternatively, nucleic acids can be introduced into a plant cell by contacting the plant cell using mechanical or chemical means. For example, a nucleic acid can be mechanically transferred by microinjection directly into plant cells by use of micropipettes. Alternatively, a nucleic acid may be transferred into the plant cell by using polyethylene glycol which forms a precipitation complex with genetic material that is taken up by the cell.

[0301] Also contemplated are silicon carbide or tungsten whiskers, for example as described in United States Patent No. 5,302,523.

[0302] There are a variety of methods known currently for transformation of monocotyledonous plants. Presently, preferred methods for transformation of monocots are microprojectile bombardment of explants or suspension cells, and direct DNA uptake or electroporation as, for example, described by Shimamoto *et al.* (1989, *supra*). Transgenic maize plants have been obtained by introducing the *Streptomyces hygroscopicus bar* gene into embryogenic cells of a maize suspension culture by microprojectile bombardment (Gordon-Kamm, 1990, *Plant Cell*, **2**:603-618). The introduction of genetic material into aleurone protoplasts of other monocotyledonous crops such as wheat and barley has been reported (Lee, 1989, *Plant Mol. Biol.* **13**:21-30). Wheat plants have been regenerated from embryogenic suspension culture by selecting only the aged compact and nodular embryogenic callus tissues for the establishment of the embryogenic suspension cultures (Vasil, 1990, *Bio/Technol.* **8**:429-434). The combination with transformation systems for these crops enables the application of the present invention to monocots. These methods may also be applied for the transformation and regeneration of dicots. Transgenic sugarcane plants have been regenerated from embryogenic callus as, for example, described by Bower et al. (1996, *Molecular Breeding* **2**:239-249).

[0303] Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process, *e.g.*, bombardment with *Agrobacterium* coated microparticles (EP-A-486234) or microprojectile bombardment to induce wounding followed by co-cultivation with *Agrobacterium* (EP-A-486233).

8. Strategies to avoid transgene silencing

[0304] Transgenes driven by various promoter sequences can be efficiently silenced during the regeneration and growth of mature plants. To avoid undesired transgene silencing, constructs are usually designed and transferred so as to minimize insofar as possible features likely to contribute to the production of aberrant transcripts and associated sequence-specific gene silencing responses in plants.

8.1 Simple integration events and targeted integration sites

[0305] Simple integrations at a known favorable site for stable transgene expression may be obtained by homologous recombination in plants. Accessory components such as site-specific recombinases, transposases or rare cutting nucleases can be used to facilitate simple and/or targeted integration as described below.

8.2 Recombinase-mediated resolution of complex integration events

[0306] Site-specific recombinase systems, such as Cre-*lox* from bacteriophage P1, can be employed for resolution of complex integration events into simple patterns (Srivastava 1999, 2001; De Buck 2001a, 2001b, 2007).

8.3 Recombinase-mediated site-specific integration

[0307] Site-specific recombination systems can also be used to bring about targeted integration of transgenes into predetermined genomic loci through a two-round transformation procedure. The first round of transformation introduces a construct containing a *lox* site. Single-copy lines are selected as recipients for the next step, in which Cre recombinase activity integrates a *lox*-flanked gene of interest specifically into the single-copy *lox* 'landing pad'. The three necessary components (favorable *lox* landing site, Cre recombinase enzyme activity and *lox*-flanked gene for integration) can be brought together either by crossing transgenic plant lines with separate components, or by a second round of transformation into the selected recipient lines. Several site-specific recombinase systems including FLP-*FRT* and R-*RS* have shown promise in plants, but Cre-*lox* is the best characterized (Ow 2002; Ow 2007).

[0308] Direct gene transfer and *Agrobacterium*-mediated gene transfer systems have been applied for Cre-*lox* site-specific integration (Albert 1995, Vergunst 1998, Srivastava 2002). The excision by site-specific recombinases is intrinsically more favorable than integration. Therefore, the strategy to obtain stable integrants typically involves use of (i)

mutant *lox* sites that are compatible for targeted integration but less effective for excision of the integrated molecule, (ii) transient recombinase activity and (iii) selection based on reconstitution of a promoter - selectable marker expression unit through targeted integration. Conditions (ii) and (iii) can be accomplished together if the landing pad is a promoter-*lox-cre* cassette, so that targeted integration abolishes Cre production (provided the recipient line was hemizygous for the landing pad) and commences integrated marker gene expression. Counter-selection has been essential to eliminate lines with non-targeted integration events in most attempts at homologous recombination in plants, but not when using an efficient recombinase strategy. With an inducible *cre* gene and appropriate vector design it should be possible to achieve both single-copy site-specific integration and removal of unwanted selectable marker genes following selection of transformed cells (Ow 2007).

[0309] *Agrobacterium*-mediated gene transfer can be used with a recombinase-mediated cassette exchange (RMCE) strategy in which (i) the landing pad and replacement gene are each flanked by a wild–type and incompatible mutant *lox* site, (ii) Cre is provided by transient or integrative transformation using a separate T-DNA, and (iii) stringent selection for site-specific integration is based on expression of a replacement selectable marker gene from a promoter in the landing pad (Nanto 2005; Louwerse 2007).

8.4 Transposon-mediated integration

[0310] In contrast with random integration following *Agrobacterium*-mediated or direct gene transfer, maize *Ac/Ds* transpositions show a bias to regions that might be more amenable to transgene expression (Zhao 2006). When barley transformed using T-DNA that contained a *bar* selectable marker within a non-autonomous *Ds* element was crossed with *Ac* transposase-expressing plants, the *bar*-containing *Ds* element transposed to different sites in the genome and segregated away from the initial T-DNA insertions in subsequent generations. The resulting plants that harboured a single copy of the *bar*-containing *Ds* element showed greatly enhanced *bar* transgene expression stability over several generations relative to the source lines (Koprek 2001). The requirement for crossing has subsequently been avoided by including the transposase gene and a negative selectable marker (cytosine deaminase) in the same T-DNA as the *Ds ::* positive selectable marker element, and selecting against cells that incorporate the full T-DNA (Yan and Rommens 2007).

8.5 Restriction-site-directed integration

[0311] Co-expression of a rare-cutting restriction enzyme or a zinc finger nuclease, which induces double-stranded breaks in DNA, has been developed for site-specific integration of transgenes into a corresponding recognition site following protoplast or microprojectile or *Agrobacterium*- mediated transfer. The recognition site can be native to the recipient genome, or in a previously engineered 'landing pad' (Wright 2005, D'Halluin 2008).

8.6 Use of S/MAR elements

[0312] Nuclear scaffold or matrix attachment regions (S/MARs) are sequences of about 300 base pairs to several kilobases that play a structural role in anchoring chromatin to the framework of the nuclear scaffold (Allen 2000, Chernov 2004). The inclusion of flanking S/MARs in constructs has been reported to substantially increase transgene expression in plants (Allen 1996, Ulker 1999), Vain 1999,,Brouwer 2002, Mankin 2003).

8.7 Elimination of silencing triggers in the transgene sequence

[0313] Duplicated T-DNA borders can be used to minimise the incidence of integrated sequences from the flanking binary or Ti vector (Thole 2007). Constructs employing tandem terminators can be used to reduce read-through (Luo 2007).

8.8 Minimal expression cassettes

[0314] It is useful to transfer the minimum DNA sequence needed to confer the desired transgenic trait. This eliminates the possibility that vector sequences might be unintended sources of signals to trigger silencing responses that can spread into the transgene. In the case of *Agrobacterium*-mediated transfer this is accomplished by use of minimal distance between the T-DNA borders. In the case of DGT, the transgene expression cassette may be isolated without vector backbone by either restriction digestion or PCR amplification (Agrawal 2005; Kumar 2006).

8.9 Codon Optimisation

[0315] Although the genetic code is universal, the frequency of use of particular codons in the redundant set encoding any amino acid varies between organisms and even between protein classes (for example those expressed at high or low levels). This is thought to be related to the abundance of corresponding tRNAs in cells, which might in turn limit translation of transgenes containing rarely used codons. Synonymous mutations may be

introduced into transgene sequences, either to use the most abundant codons in recipient cell genes or to more closely match the overall codon usage frequencies of the recipient (De Rocher 1998). Potential poly(A) addition signals can be removed simultaneously (Diehn 1998, Misztal 2004). Recently, Li (2007) reported that replacement of uncommon codons while maintaining codon diversity gave superior expression relative to use of the most abundant codon throughout a synthetic gene.

[0316] Heterologous genes may simultaneously be modified to eliminate signals for other processing events that limit transgene expression, for example intron splice signals, sequence context around the start codon, or mRNA destabilizing elements (Gutierrez 1999, Haseloff 1997, Holmberg 2001, Khanna 2006).

8.10 Inhibiting RNAi pathways

[0317] One role for RNAi in plants is in antiviral defense, and many viruses encode proteins that interfere with RNAi mechanisms. Diverse viral suppressors act at several different steps in the RNAi pathway (Roth 2004; Voinnet 2005; Scholthof 2007). These suppressors can inhibit silencing of transgenes in addition to viral RNAs, as first demonstrated for the potyvirus helper component-protease, HC-Pro. Co-transformation with silencing suppressors or crossing to combine the transgene with a silencing suppressor may be used to inhibit silencing (Anandalakshmi 1998, Johansen 2001, Lim 2005, Lewsey 2007).

9. Production and characterization of differentiated transgenic plants

9.1 Regeneration

[0318] The methods used to regenerate transformed cells into differentiated plants are not critical to this invention, and any method suitable for a target plant can be employed. Normally, a plant cell is regenerated to obtain a whole plant following a transformation process.

[0319] Regeneration from protoplasts varies from species to species of plants, but generally a suspension of protoplasts is first made. In certain species, embryo formation can then be induced from the protoplast suspension, to the stage of ripening and germination as natural embryos. The culture media will generally contain various amino acids and hormones, necessary for growth and regeneration. Examples of hormones utilized include auxins and cytokinins. It is sometimes advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Efficient regeneration will depend on the

medium, on the genotype, and on the history of the culture. If these variables are controlled, regeneration is reproducible. Regeneration also occurs from plant callus, explants, organs or parts. Transformation can be performed in the context of organ or plant part regeneration as, for example, described in *Methods in Enzymology*, Vol. 118 and Klee *et al.* (1987, *Annual Review of Plant Physiology*, **38**:467), which are incorporated herein by reference. Utilizing the leaf disk-transformation-regeneration method of Horsch *et al.* (1985, *Science*, **227**:1229, incorporated herein by reference), disks are cultured on selective media, followed by shoot formation in about 2-4 weeks. Shoots that develop are excised from calli and transplanted to appropriate root-inducing selective medium. Rooted plantlets are transplanted to soil as soon as possible after roots appear. The plantlets can be repotted as required, until reaching maturity.

[0320] In vegetatively propagated crops, the mature transgenic plants are propagated by the taking of cuttings or by tissue culture techniques to produce multiple identical plants. Selection of desirable transgenics is made and new varieties are obtained and propagated vegetatively for commercial use.

[0321] In seed propagated crops, the mature transgenic plants can be self-crossed to produce a homozygous inbred plant. The inbred plant produces seed containing the newly introduced foreign gene(s). These seeds can be grown to produce plants that would produce the selected phenotype, *e.g.*, early flowering.

[0322] Parts obtained from the regenerated plant, such as flowers, seeds, leaves, branches, fruit, and the like are included in the invention, provided that these parts comprise cells that have been transformed as described. Progeny and variants, and mutants of the regenerated plants are also included within the scope of the invention, provided that these parts comprise the introduced nucleic acid sequences.

[0323] It will be appreciated that the literature describes numerous techniques for regenerating specific plant types and more are continually becoming known. Those of ordinary skill in the art can refer to the literature for details and select suitable techniques without undue experimentation.

9.2 Characterization

[0324] To confirm the presence of the heterologous nucleic acid in the regenerating plants, a variety of assays may be performed. Such assays include, for example, "molecular biological" assays well known to those of skill in the art, such as Southern and Northern

blotting and PCR; a protein expressed by the heterologous DNA may be analysed by western blotting, high performance liquid chromatography or ELISA (*e.g.*, nptII) as is well known in the art.

[0325] Representative examples of various methods applicable to characterization of transgenic plants are provided in Chapters 9 and 11 of PLANT MOLECULAR BIOLOGY A Laboratory Manual Ed. M.S. Clark (Springer-Verlag, Heidelberg, 1997.

[0326] In order that the invention may be readily understood and put into practical effect, particular preferred embodiments will now be described by way of the following non-limiting examples.

EXAMPLES

EXAMPLE 1

IDENTIFICATION OF A GENE THAT IS UP-REGULATED IN MATURE STEM

[0327] Sugarcane EST clones potentially up-regulated in mature stems have been indicated by analysis of microarray data (Casu *et al.* 2004). There are many uncertainties in the interpretation of such data, with the consequence that further experimentation is required to determine which of the many candidate genes from the microarray analysis indeed show the particular expression patterns of interest.

[0328] For example, cDNA clone MCSA063B04 was identified as putatively up-regulated in mature vs. immature stem based on microarray data (Casu *et al.* 2004). EST sequencing of this clone revealed a putative methionine sulfoxide reductase (GenBank Accession number CF572917). However, the present inventors found by Northern analysis that this gene is not up-regulated in mature stems. Surprisingly, further sequence determination and analysis of the full insert in clone MCSA063B04 revealed the presence of likely concatenated cDNAs. A previously unknown region within this chimeric cDNA clone comprised a full open reading frame with a predicted translation product similar to SHAQKYF R1MYB transcription factors. The sequence of this region is provided as SEQ ID NO: 83.

[0329] Transcripts hybridising to a probe comprising SEQ ID NO: 83 were shown, by northern analysis, to be strongly up-regulated in mature *vs.* immature stem (Figure 1a), and also detectable in mature roots (Figure 1b), but not in leaf tissue. Northern analysis using RNA from vascular, parenchyma and rind tissue dissected from internode eight of field-grown Q117 (Rae *et al.*, 2005) revealed the presence of transcript in all three tissue types within the stem (Figure 1c). Multiple bands in the northern analysis appear to reflect alternative splicing of the transcript (see below).

[0330] Southern analysis using the same probe revealed a relatively simple banding pattern by polyploid sugarcane standards (Figure 2), given that the restriction enzymes *Eco*RI, *Hin*dIII and *Sac*I all cut once within the region covered by the probe. The sugarcane gene locus corresponding to this isolated nucleotide sequence is hereby designated as *ScR1MYB1.*

EXAMPLE 2

SUGARCANE IS HIGHLY HETEROZYGOUS AT THE ScR1MYB1 LOCUS

[0331] Allelic variation within sugarcane cultivar Q117 was analysed by cloning and sequencing multiple amplicons obtained from genomic PCR and RT-PCR. These amplicons contained either the full ScR1MYB1 coding region (~ 1.3 kb and 0.9 kb products for genomic and RT-PCR, respectively), or a product lacking the first 220 nts of exon 1. Thirty two sequence variants were detected among 51 genomic PCR products from five independent amplifications, and 106 RT-PCR products from four amplifications on RNA from mature stem (internode 12). However, in control reactions on a known template we detect approximately 1 error per 1.3 kb amplicon under the same PCR conditions. PCR recombination can also lead to an overestimation of allelic diversity (Cronn *et al.*, 2002; Yu *et al.*, 2006). Using a conservative approach of counting only sequence variants detected multiple times in independent PCRs, we obtained a total of eight haplotypes from Q117. Evidence from genomic Southern analysis, BAC library screening, and analysis of the corresponding locus in *Sorghum bicolor* (see below), indicates that these haplotypes most likely represent alleles at a single locus. The frequency of single nucleotide polymorphisms (SNPs) detected in introns and exons was similar (perhaps reflecting the alternative splicing of the larger intron 1; see below). Approximately 40% of SNPs lead to an amino acid substitution, and the eight putative alleles encode different predicted proteins (Figure 3). Substantial variability towards the downstream end of the coding region resulted in predicted protein lengths ranging from 292-313 amino acids.

[0332] As an alternative measure of allelic variation in sugarcane, we isolated and sequenced multiple clones containing the full coding region from a genomic BAC library of cultivar Q200 (thereby avoiding any PCR-derived artefacts). We screened approximately 95000 clones with an average insert size of 115 kb (S. Brumbley, BSES, Brisbane, Queensland, pers. comm.). Assuming a non-replicated somatic cell (2C) genome size of 6400 Mb (Edme *et al.*, 2005) and using the formula of Clarke and Carbon (1976), there is an 82% probability of any Q200 gene being present this library, but only a 14% probability of 10 homo(eo)logous copies of this sequence being present. The present inventors identified six alleles amongst 13 hybridizing clones in this library. An additional allele (X2) is represented several times in the pool of cloned RT-PCR products from mature stem of this cultivar. Thus, at least seven alleles are present in Q200. Several alleles appear to be specific to each cultivar.

EXAMPLE 3

MULTIPLE SCR1MYB1 ALLELES ARE EXPRESSED IN SUGARCANE

[0333] The frequency of occurrence among sequenced RT-PCR products indicates that transcripts of the Z1 allele predominate (40%) in internode 12 tissue of Q117, with three other alleles each contributing ≥10% (Fig 4). Interestingly, the same alleles predominate among Q200 transcripts. These cultivars do not have the same parents, but do share common ancestors in earlier generations.

[0334] Sequence analysis of the RT-PCR products also reveals the presence of alternatively spliced transcripts. In ScR1MYB1 and closely related sequences from the rice and *Arabidopsis* genomes, there are two introns, with the position of intron 1 being highly conserved and the position of intron 2 being less conserved. In approximately one third (34/106) of RT-PCR products from Q117, intron 1 was unspliced, and in rare cases (2/106 clones), a small part of intron 1 remained. Similarly, in Q200 40% (26/65) of all RT-PCR products were unspliced. In all cases, intron 2 was fully spliced. The presence of all or part of intron 1 in these transcripts results in a smaller predicted protein (152 and 124 amino acids, respectively), that does not include any of the MYB domain but does include the N-terminal zinc finger and serine-rich domains (Figure 3). The presence of both spliced and unspliced transcript explains the double band observed on the northern blot probed with the ScR1MYB1 probe (Figure 1).

EXAMPLE 4

SCR1MYB1 PROMOTER ALLELES

[0335] The presence of a full open reading frame for ScR1MYB1 in the cDNA clone facilitated amplification of upstream promoter sequences. Using the GenomeWalker single-specific-primer PCR procedure, we identified eight promoter alleles among 41 cloned products of at least 1 kb in size. Three of these (A1, A2 and B) were identified multiple times in multiple PCRs. An additional three (A3, A6 and A5) had unique insertions or deletions (of 240 bp, 18 bp and at least 1.2 kb, respectively) unlikely to arise from PCR errors. The remaining two alleles (A4 and A7) differ only by a few SNPs. They were recovered because of rare events in which the adapter primer ligated to a random blunt end rather than the blunt restriction site used to construct the library.

[0336] These eight putative promoter alleles show 99% identity for approximately 280 bp upstream of the start codon. A putative TATA box predicted using the program HCtata (Milanesi *et al.*, 1996) occurs approximately 133-135 bp upstream of the start codon, and 23 bp upstream of the transcription start site determined by 5'-RACE. Beyond 280 bp upstream, the sequence of the B allele diverges and approximately 60 bp further upstream, the A5 sequence diverges, suggesting that major rearrangements have occurred in these promoter alleles. The sequences of ten representative promoter alleles from the sugarcane *ScR1MYB1* locus are provided as SEQ ID NOS: 73- 82.

[0337] To investigate the relationship between promoter alleles and coding sequence alleles, the inventors have amplified across the promoter:coding sequence junction using either conserved or allele-specific PCR primers in both the promoter and coding sequence. Among 16 clones generated with a Z1 specific primer, all contained the Z1 coding sequence linked to the A1 promoter. All five clones containing Z1 from amplifications using conserved primers were also linked to A1, indicating that the A1 promoter allele is uniquely linked to the most strongly expressed (Z1) coding sequence in Q117. In Q200 BAC clones, the Z1 and Z2 coding sequences were always linked to the A1 promoter sequence. The A1 promoter:Z1 coding sequence was identified 5 times among the 13 BAC clones, suggesting that the high proportion of Z1 transcripts corresponds with a high genomic copy number of this allele in Q200. It can be concluded that the A1 promoter is functional for *ScR1MYB1* expression in both cultivars. The analysis also indicates that the X1 transcribed sequence is linked to both the A4 and B promoter sequences in Q117, so these promoters are likely to be functional in their native state.

[0338] In the course of PCR analysis of the promoter:coding sequence links, the present inventors revealed two additional promoters (A8 and A9) linked to weakly expressed alleles (T and U) which comprise ~1% of the transcripts from the *ScR1MYB1* locus in Q117. The A8 and A9 upstream sequences are provided as SEQ ID NOS: 10 and 11. The same process can be continued to identify promoter(s) linked to other transcribed alleles at this locus in sugarcane. It is advantageous to be able to access promoters that drive different levels of expression in this desired developmental pattern, because the optimal expression level for commercial utility will differ between genes and traits.

[0339] At least ten independent transgenic sugarcane lines containing each of eight promoter alleles fused to the LUC and/or GUS*Plus* reporter genes were assayed for reporter

gene activity. All eight promoters drove GUS activity in some regenerating shoots in tissue culture.

[0340] Additionally, the A1 promoter [SEQ ID NO.77] was fused to the coding region for a vacuole-targeted sucrose isomerase that converts sucrose into isomers isomaltulose and trehalulose that are not normally present in sugarcane. Transgenic plants obtained after bombardment using this construct were grown in the field for 8 months, and then juice was extracted from mature stem internodes. Among 21 analysed plants, 15 produced detectable levels of isomaltulose and trehalulose. In the lines with strongest expression, the novel sucrose isomers made up three quarters of the total sugar in these basal internodes, showing that the A1 promoter drives strong transgene expression in mature stem internodes of sugarcane.

Materials and Methods

CHOICE OF TARGET LOCUS

[0341] Sugarcane cDNA clones that are potentially up-regulated in mature stems were identified initially based on microarray data (Casu *et al.*, 2004). Further analysis of many such clones led us to a particular interest in a sugarcane gene locus designated below as *ScR1MYB1*.

RNA EXTRACTION AND NORTHERN ANALYSIS

[0342] RNA was extracted from field- or glasshouse-grown sugarcane cultivar Q117 tissues by CsCl ultracentrifugation (Chirgwin *et al.*, 1979). Following denaturation, approximately 10 µg RNA was fractionated on a 1.2% TBE agarose gel, and transferred onto Hybond N^+ membrane (Amersham) using capillary transfer with 10×SSC. Hybridization was done for 16 h at 65°C in 0.5 M NaH_2PO_4, 1 mM EDTA, 7% SDS; using ^{32}P-dCTP-labelled probes prepared using a Megaprime kit (Amersham). Following hybridization, membranes were washed to high stringency (0.1×SSC; 0.1% SDS at 65°C), and analyzed using a PhosphorImager SI (Molecular Dynamics).

DNA EXTRACTION AND SOUTHERN ANALYSIS

[0343] Sugarcane genomic DNA was isolated using a CTAB-based method (Rogers and Bendich 1988). Ten µg of each sample was digested with restriction enzymes, fractionated on a 0.8% agarose TAE gel, and transferred onto Hybond N^+ membrane

(Amersham) using capillary transfer with either 0.4 M NaOH or 10 × SSC. Labeling of probes and hybridization conditions were as described for northern analysis.

DNA SEQUENCING

[0344] DNA sequencing of plasmid templates was done using a BigDye Terminator 3.1 DNA sequencing kit (Applied Biosystems), and separations by the Australian Genomic Research Facility (Brisbane, Queensland, Australia).

ANALYSIS OF ALLELIC VARIATION AND EXPRESSION

[0345] The degree of allelic variation for ScR1MYB1 was assessed by cloning and sequencing multiple PCR products amplified from Q117 using a forward primer beginning at the start codon (MYB F: 5'-ATGGCTAGGAAATGCTCCAG-3') and a reverse primer beginning at the stop codon (MYB R: 5'-TTATTCTCCTAGACGCCAGT-3'). Similarly, RT-PCR products derived from mature stem (amplified using the above primer set or the MYB R primer in combination with MYBF220: 5'-GCCTACTAYGGAGCTGCTGC-3') were sequenced to determine which alleles are expressed. For RT-PCR, cDNA was synthesized using Superscript III (Invitrogen) and oligo dT primer using RNA from internodes 11-12 or 30. In all cases, the PCR was done using Expand High Fidelity polymerase (Roche) according to the manufacturer's instructions, and all amplified products were cloned into pGEM-T Easy (Promega) prior to sequencing.

ISOLATION OF UPSTREAM PROMOTER REGIONS AND PCR AMPLIFICATION

[0346] Promoter alleles corresponding to *ScR1MYB1* were amplified using a PCR-based Genome Walker strategy (Clontech), using primers GSP1 (5'-ATTGTTTCCACAACTGGAGCATTTCCTAGC-3') and GSP2 (5'-AACAAGGGCACAAGAAGTTGCGGTGTAGTA -3') derived from the 5' end of the *ScR1MYB1* transcript. To maximize the chances of amplifying all promoter alleles in the polyploidy genome of cultivar Q117, we used a total of 8 GenomeWalker libraries (prepared using restriction enzymes *Dra*I, *Eco*RV, *Pvu*II, *Sca*I, *Stu*I, *Bsr*BI, *Ecl*136II, and *Ssp*I). These and all other PCR amplification reactions used Expand High Fidelity polymerase (Roche) according to the manufacturer's instructions. Amplified products were cloned into pGem-T Easy (Promega) prior to sequencing. The transcriptional start site was determined by 5'RACE using the GeneRacer Kit (Invitrogen) according to the manufacturer's instructions.

CONSTRUCTION OF PROMOTER-REPORTER GENE FUSIONS

[0347] Promoter sequences with native 5'UTRs were amplified from Genome Walker clones, by use of PCR primers designed to incorporate *Pst*I and *Nco*I restriction sites. These restriction sites enabled cloning of the promoters in front of the GUS*Plus* (CAMBIA) and *luc*$^{+}$NF (Promega) reporter genes. Constructs containing the same reporter genes driven by the maize *Ubi-1* promoter (Christensen et al. 1992) were included as positive controls. All constructs included the *nos* terminator.

SUGARCANE TRANSFORMATION

[0348] Transformation and regeneration of sugarcane cultivar Q117 (obtained from BSES Meringa Sugar Experiment Station, Gordonvale, Qld) was done as previously described (Bower et al. 1996). Plants were grown in a containment glasshouse under natural light intensity, at 28°C and watered twice a day. Each plant was grown as a single stalk in a pot of 20 cm diameter and a density of 18 pots/m^2, fertilized with Osmocote® at 5 g/month for the first and the second months, then 10 g/month. Leaves were numbered from one for the top visual dewlap (TVD) with higher numbers for older leaves. Internodes were numbered according to the leaf attached to the node immediately above.

REPORTER GENE ASSAYS

[0349] For histochemical GUS staining, tissues were vacuum infiltrated in assay buffer (50 mM sodium phosphate, pH 7.0; 0.1% Triton X-100; 0.5 mM potassium ferrocyanide; 0.5 mM potassium ferricyanide; 10 mM EDTA; 0.05% 5-bromo-4-chloro-3-indolyl glucuronide) and incubated at 37 °C for 18 h. Green tissues were destained with ethanol. For fluorimetric GUS assays (Jefferson 1987), fluorescence measurements were taken at 0, 60, 120 and 180 minute timepoints using a Labsystems Fluoroskan Ascent microplate reader (Pathtech Diagnostics, Melbourne, Victoria, Australia). LUC assays were done as described (Luehrsen and Walbot 1993), using a BMG POLARstar OPTIMA luminometer.

[0350] The disclosure of every patent, patent application, and publication cited herein is hereby incorporated herein by reference in its entirety.

[0351] The citation of any reference herein should not be construed as an admission that such reference is available as "Prior Art" to the instant application.

[0352] Throughout the specification the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. Those of skill in the art will therefore appreciate that, in light of the instant disclosure, various modifications and changes can be made in the particular embodiments exemplified without departing from the scope of the present invention. All such modifications and changes are intended to be included within the scope of the appended claims.

BIBLIOGRAPHY

Agrawal P.K., Kohli A., Twyman R. M.and Christou P. 2005. Transformation of plants with multiple cassettes generates simple transgene integration patterns and high expression levels. *Molecular Breeding* 16: 247-260.

Albert, H., Dale, E. C., Lee, E. and Ow, D. W. 1995. Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome. *Plant Journal* 7(4):649-659.

Allen, G. C., Hall, G., Michalowski, S., Newman, W., Spiker, S., Weissinger, A. K. and Thompson, W. F. 1996. High-level transgene expression in plant cells: Effects of a strong scaffold attachment region from tobacco. *Plant Cell* 8(5):899-913.

Allen, G. C., Spiker, S. and Thompson, W. F. 2000. Use of matrix attachment regions (MARS) to minimize transgene silencing. *Plant Molecular Biology* 43(2-3):361-376.

Anandalakshmi, R., Pruss, G. J., Ge, X., Marathe, R., Mallory, A. C., Smith, T. H. and Vance, V. B. 1998. A viral suppressor of gene silencing in plants. *Proceedings of the National Academy of Sciences of the United States of America* 95(22):13079-13084.

Bower, R., Elliott, A.R., Potier, B.A.M., and Birch R.G. 1996. High-efficiency, microprojectile-mediated cotransformation of sugarcane, using visible or selectable markers. *Moecularl Breeding* 2: 239-249.

Brouwer, C., Bruce, W., Maddock, S., Avramova, Z. and Bowen, B. 2002. Suppression of transgene silencing by matrix attachment regions in maize: A dual role for the maize 5 ' *Adh1* matrix attachment region. *Plant Cell* 14(9):2251-2264.

Casu, R.E., Dimmock, C.M., Chapman, S.C., Grof, C.P.L., McIntyre, C.L., Bonnett, G.D, and Manners, J.M. 2004. Identification of differentially expressed transcripts from maturing stem of sugarcane by in silico analysis of stem expressed sequence tags and gene expression profiling. *Plant Molecular Biology* 54: 503-517.

Chernov, I. P., Akopov, S. B. and Nikolaev, L. G. 2004. Structure and functions of nuclear matrix associated regions (S/MARS). *Russian Journal of Bioorganic Chemistry* 30(1):1-11.

Chirgwin, J.M., Przybyla, A.E., Macdonald, R.J., and Rutter, W.J. 1979. Isolation of biologically-active ribonucleic-acid from sources enriched in ribonuclease. *Biochemistry* 18: 5294-5299.

Clarke, L., and Carbon, J. 1976. Colony bank containing synthetic Col El hybrid plasmids representative of entire *Escherichia Coli* genome. *Cell* 9: 91-99.

Christensen, A.H., Sharrock, R.A., and Quail, P.H. 1992. Maize polyubiquitin genes - Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Molecular Biology* 18: 675-689.

Cronn, R., Cedroni, M., Haselkorn, T., Grover, C., and Wendel, J.F. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* 104: 482-489.

D'Halluin, K., Vanderstraeten, C., Stals, E., Cornelissen, M. and Ruiter, R. 2008. Homologous recombination: A basis for targeted genome optimization in crop species such as maize. *Plant Biotechnology Journal* 6(93-102.

De Buck, S., Van Montagu, M. and Depicker, A. 2001a. Transgene silencing of invertedly repeated transgenes is released upon deletion of one of the transgenes involved. *Plant Molecular Biology* 46(4):433-445.

De Buck, S. and Depicker, A. 2001b. Disruption of their palindromic arrangement leads to selective loss of DNA methylation in inversely repeated *gus* transgenes in *Arabidopsis*. *Molecular Genetics and Genomics* 265(6):1060-1068.

De Buck, S., Peck, I., De Wilde, C., Marjanac, G., Nolf, J., De Paepe, A. and Depicker, A. 2007. Generation of single-copy T-DNA transformants in Arabidopsis by the *cre/loxP* recombination-mediated resolution system. *Plant Physiology* 145(4):1171-1182.

De Rocher, E. J., Vargo-Gogola, T. C., Diehn, S. H. and Green, P. J. 1998. Direct evidence for rapid degradation of *Bacillus thuringiensis* toxin mRNA as a cause of poor expression in plants. *Plant Physiology* 117(4):1445-1461.

Diehn, S. H., Chiu, W. L., De Rocher, E. J. and Green, P. J. 1998. Premature polyadenylation at multiple sites within a *Bacillus thuringiensis* toxin gene-coding region. *Plant Physiology* 117(4):1433-1443.

Edme, S.J., Comstock, J.C., Miller, J.D., and Tai, P.Y.P. 2005. Determination of DNA content and genome size in sugarcane. *Journal of the American Society of Sugar Cane Technologists* 25: 1-16.

Gutierrez, R. A., MacIntosh, G. C. and Green, P.J. 1999. Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control. *Trends in Plant Science* 4: 429-438.

Haseloff, J., Siemering, K. R., Prasher, D. C. and Hodge, S. 1997. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proceedings of the National Academy of Sciences of the USA* 94: 2122-2127.

Holmberg, N., Farres, J., Bailey, J. E. and Kallio, P. T. 2001. Targeted expression of a synthetic codon optimized gene, encoding the spruce budworm antifreeze protein, leads to accumulation of antifreeze activity in the apoplasts of transgenic tobacco. *Gene* 275(1):115-124.

Jefferson, R. 1987. Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Moecularl Biology Reporter* 5: 387-405.

Johansen, L. K. and Carrington, J. C. 2001. Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. *Plant Physiology* 126(3):930-938.

Khanna, H. K. and Daggard, G. E. 2006. Targeted expression of redesigned and codon optimised synthetic gene leads to recrystallisation inhibition and reduced electrolyte leakage in spring wheat at sub-zero temperatures. *Plant Cell Reports* 25(12):1336-1346.

Koprek, T., Rangel, S., McElroy, D., Louwerse, J. D., Williams-Carrier, R. E. and Lemaux, P. G. 2001. Transposon-mediated single-copy gene delivery leads to increased transgene expression stability in barley. *Plant Physiology* 125(3):1354-1362.

Kumar, S., Arul, L., Talwar, D. and Raina, S. K. 2006. PCR amplification of minimal gene expression cassette: an alternative, low cost and easy approach to 'clean DNA' transformation. *Current Science* 91: 930-934.

Lewsey, M., Robertson, F. C., Canto, T., Palukaitis, P. and Carr, J. P. 2007. Selective targeting of miRNA-regulated plant development by a viral counter-silencing protein. *Plant Journal* 50: 240-252.

Li, X. Q., Wei, J. Z., Tan, A. and Aroian, R. V. 2007. Resistance to root-knot nematode in tomato roots expressing a nematicidal *Bacillus thuringiensis* crystal protein. *Plant Biotechnology Journal* 5(4):455-464.

Lim, H. S., Ko, T. S., Lambert, K. N., Kim, H. G., Korban, S. S., Hartman, G. L. and Domier, L. L. 2005. Soybean mosaic virus helper component-protease enhances somatic embryo production and stabilizes transgene expression in soybean. *Plant Physiology and Biochemistry* 43(10-11):1014-1021.

Louwerse, J. D., van Lier, M. C. M., van der Steen, D. M., de Vlaam, C. M. T., Hooykaas, P. J. J. and Vergunst, A. C. 2007. Stable recombinase-mediated cassette exchange in *Arabidopsis* using *Agrobacterium tumefaciens*. *Plant Physiology* 145(4):1282-1293.

Luehrsen, K., and Walbot, V. 1993. Firefly luciferase as a reporter for plant gene expression studies. *Promega Notes* 44: 24-29.

Luo, Z. H. and Chen, Z. X. 2007. Improperly terminated, unpolyadenylated mRNA of sense transgenes is targeted by *RDR6*-mediated RNA silencing in *Arabidopsis*. *Plant Cell* 19(3):943-958.

Mankin, S. L., Allen, G. C., Phelan, T., Spiker, S. and Thompson, W. F. 2003. Elevation of transgene expression level by flanking matrix attachment regions (MAR) is promoter dependent: A study of the interactions of six promoters with the *RB7* 3 ' MAR. *Transgenic Research* 12(1):3-12.

Milanesi, L., Muselli, M., and Arrigo, P. 1996. Hamming Clustering method for signals prediction in 5' and 3' regions of eukaryotic genes. *Computing Applications in the Biosciences* 12: p399-404.

Misztal, L. H., Mostowska, A., Skibinska, M., Bajsa, J., Musial, W. G. and Jarmolowski, A. 2004. Expression of modified *cry1ac* gene of *Bacillus thuringiensis* in transgenic tobacco plants. *Molecular Biotechnology* 26(1):17-26.

Nanto, K., Yamada-Watanabe, K., Ebinuma, H. 2005. Agrobacterium-mediated RMCE approach for gene replacement. *Plant Biotechnology Journal* 3: 203-214.

Ow, D. W. 2002. Recombinase-directed plant transformation for the post-genomic era. *Plant Molecular Biology* 48: 183-200.

Ow, D. W. 2007. GM maize from site-specific recombination technology, what next? *Current Opinion in Biotechnology* 18: 115-120.

Rae, A.L., Grof, C.P.L., Casu, R.E., and Bonnett, G.D. 2005. Sucrose accumulation in the sugarcane stem: pathways and control points for transport and compartmentation. *Field Crops Research* 92: 159-168.

Rogers, S., and Bendich, A. 1988. Extraction of DNA from plant tissues. In: *Plant molecular biology manual*. Kluwer Academic Publishers, Dordrecht, pp 1-10.

Roth, B. M., Pruss, G. J. and Vance, V. B. 2004. Plant viral suppressors of RNA silencing. *Virus Research* 102(1):97-108.

Scholthof, H. B. 2007. Heterologous expression of viral RNA interference suppressors: Risc management. *Plant Physiology* 145:1110-1117.

Srivastava, V., Anderson, O. D. and Ow, D. W. 1999. Single-copy transgenic wheat generated through the resolution of complex integration patterns. *Proceedings of the National Academy of Sciences of the United States of America* 96(20):11117-11121.

Srivastava, V. and Ow, D. W. 2001. Single-copy primary transformants of maize obtained through the co-introduction of a recombinase-expressing construct. *Plant Molecular Biology* 46(5):561-566.

Srivastava,V.and Ow, D. W. 2002. Biolistic mediated site-specific integration in rice. *Molecular Breeding* 8: 345-350.

Thole, V., Worland, B., Snape, J. W.and Vain P 2007. The pCLEAN dual binary vector system for *Agrobacterium*-mediated plant transformation. *Plant Physiology* 145: 1211-1219.

Ulker, B., Allen, G. C., Thompson, W. F., Spiker, S. and Weissinger, A. K. 1999. A tobacco matrix attachment region reduces the loss of transgene expression in the progeny of transgenic tobacco plants. *Plant Journal* 18(3):253-263.

Vain, P., Worland, B., Kohli, A., Snape, J. W., Christou, P., Allen, G. C. and Thompson, W. F. 1999. Matrix attachment regions increase transgene expression levels and stability in transgenic rice plants and their progeny. *Plant Journal* 18(3):233-242.

Vergunst, A. C., Jansen, L. E. T. and Hooykaas, P. J. J. 1998. Site-specific integration of Agrobacterium T-DNA in *Arabidopsis thaliana* mediated by *cre* recombinase. *Nucleic Acids Research* 26(11):2729-2734.

Voinnet, O. 2005. Induction and suppression of RNA silencing: Insights from viral infections. *Nature Reviews Genetics* 6(3):206-U201.

Wright, D. A., Townsend, J. A., Winfrey, R. J., Irwin, P. A., Rajagopal, J., Lonosky, P. M., Hall, B. D., Jondle, M. D. and Voytas, D. F. 2005. High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant Journal* 44: 693-705.

Yan, H. and Rommens, C. M. 2007. Transposition-based plant transformation. *Plant Physiology* 143(2):570-578.

Yu, W., Rusterholtz, K.J., Krummel, A.T., and Lehman, N. 2006. Detection of high levels of recombination generated during PCR amplification of RNA templates. *Biotechniques* 40: 499-507.

Zhao, T., Palotta, M., Langridge, P., Prasad, M., Graner, A., Schulze-Lefert, P. and Koprek, T. 2006. Mapped ds/T-DNA launch pads for functional genomics in barley. *Plant Journal* 47(5):811-826.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid molecule comprising a promoter sequence that is operable in plant cells, wherein the promoter sequence comprises a nucleotide sequence selected from the group consisting of:

(a) RTCAAGTATAGYAGTGGCAGAGAYTCAAAAGATACCTKCGGA TTCTGAATACCACAACCWTAAAAAACAGGATGATGTTATACTTGTCCCCTTC CATGATACAGGACTGTTTAGTAAYTYCCCAAACWGCCCATRATACATTCTG CACCCTTTATWAAACCTCTN$_a$AGCTACAWCATCTT [SEQ ID NO: 1];

(b) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 1 or a complement thereof; and

(c) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 1 or a complement thereof, under at least medium or high stringency conditions,

wherein:

A, C, G and T represent the nucleic acid bases adenine, cytosine, guanine and thymine, respectively;

M is A or C;

R is A or G;

W is A or T;

S is C or G;

Y is C or T

K is G or T;

V is A or C or G;

H is A or C or T;

D is A or G or T;

B is C or G or T;

N is G or A or C or T; and

a is an integer from 0-3.

2. A nucleic acid sequence according to claim 1, wherein N_a comprises 0 bases or is the sequence ACT.

3. A nucleic acid sequence according to claim 1, which is selected from the group consisting of:

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCT
TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 2];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTCCCCAAACAGCCCATAATACATTCTGCACCCT
TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 3];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCT
TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 4];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCT
TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 5];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCT
TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 6];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA
TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCT
TTATAAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 7];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG
AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA

TACAGGACTGTTTAGTAACTTCCCAAACAGCCCATGATACATTCTGCACCCT TTATTAAACCTCTAGCTACATCATCTT [SEQ ID NO: 8];

GTCAAGTATAGTAGTGGCAGAGATTCAAAAGATACCTTCGGATTCTG AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATGATACATTCTGCACCCT TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 9];

GTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG AATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA TACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCT TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 10]; and

ATCAAGTATAGCAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTG AATACCACAACCTTAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGA TACAGGACTGTTTAGTAATTTCCCAAACTGCCCATAATACATTCTGCACCCT TTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 11].

4. A nucleic acid sequence according to claim 1, wherein the promoter sequence further comprises upstream of (a), (b) or (c) a nucleotide sequence selected from the group consisting of:

(e) TKTSSRAMMMWSWKN$_b$WTRTDMYYWWAN$_c$YAHWWYAN$_d$TRTMK RAMCRARYACN$_e$AYSTKRTARTSWWKRTTYAWAWSCKRWGKWSTSAWRKY SDC [SEQ ID NO: 12];

(f) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 12 or a complement thereof; and

(g) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 12 or a complement thereof, under at least medium or high stringency conditions,

wherein:

A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

each of *b* and *c* is an integer from 0-2;

d is an integer from 0-4; and

e is an integer from 0-1.

5. A nucleic acid sequence according to claim 4, wherein N_b is a sequence selected from GA or AA.

6. A nucleic acid sequence according to claim 4, wherein N_c is the sequence GT.

7. A nucleic acid sequence according to claim 4, wherein N_d is a sequence selected from ATA or CAAA.

8. A nucleic acid sequence according to claim 4, wherein N_e is T.

9. A nucleic acid sequence according to claim 4, wherein any one or more of N_b, N_c, N_d or N_e comprises 0 bases.

10. A nucleic acid sequence according to claim 4, which is selected from the group consisting of:

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 13];

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGGC [SEQ ID NO: 14];

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAGTGTAGATTTAAATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 15];

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 16];

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 17];

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 18];

TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATAC TATGTGGTAATGTAGATTTATAAGCTGTGTACTGAAAGTGAC [SEQ ID NO: 19];

TTTCCAAAAATCATATGTTACTAAACAAAACAATATATCTAACCAAA TACTATGTGGTAGTGTAGATTTATAAGCTGTGTACTGAAAGTGAC [SEQ ID NO: 20];

TTTCGGAAACTCTTGCATATGCTCTTATATTTCATATCTAACCAAATA CTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGAC [SEQ ID NO: 21]; and

TGTGGAACCAAGTGAATTATAACCAAAGTCACTATACAAATGTAGGA ACGAGCACACCTTATAATCATTGTTCATATCCGAAGGTGTCATGTCCTC [SEQ ID NO: 22].

11. A nucleic acid sequence according to claim 4, wherein the promoter sequence further comprises upstream of (d), (e) or (f) a nucleotide sequence selected from the group consisting of:

(g) RBTN$_f$RDRVDBSSRWSDDYRWYKMBHRKRYKKBBN$_g$SAWYHKYT WSBVVVDYBYYSYWMSWRRRACVTRRN$_h$ARGAWAHYTWWMMN$_i$GYYDVD RWYMWYDRWSWWKYTCTDRHYVAMWMHMRS [SEQ ID NO: 23];

(h) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 23 or a complement thereof; and

(i) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 23 or a complement thereof, under at least medium or high stringency conditions,

wherein:

A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

f is an integer from 0 to 3;

g is an integer from 0 to 2;

h is an integer from 0 to 6; and

i is an integer from 0 to 1.

12. A nucleic acid sequence according to claim 11, wherein N_f is a sequence selected from ATG, ATA or AAA.

13. A nucleic acid sequence according to claim 11, wherein N_g is a sequence selected from TA or TC.

14. A nucleic acid sequence according to claim 11, wherein N_h is the sequence CCCA.

15. A nucleic acid sequence according to claim 11, wherein N_i is W.

16. A nucleic acid sequence according to claim 11, wherein any one or more of N_f, N_g, N_h, or N_i comprises 0 nucleotides.

17. A nucleic acid sequence according to claim 11, which is selected from the group consisting of:

ATTATGATAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAG TCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTT TGACATTCTCTGATTGAAACT [SEQ ID NO: 24];

ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAG TCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTC TGACATTCTCTGATTGAAACT [SEQ ID NO: 25];

ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAG TCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATC TGACATTCTCTGATTGAAACT [SEQ ID NO: 26];

ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAG TCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTC TGACATTCTCTGATTGAAACT [SEQ ID NO: 27];

ATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAATCTTTAG TCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTC TGACATTCTCTGATTGAAACT [SEQ ID NO: 28];

ATTATGATAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAG TCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTT TGACATTCTCTGATTGAAACT [SEQ ID NO: 29];

ATTATGATAAAGGCATCAGCAACTAGAGTACTTTGTAGAATCTTTAG TCCGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATC TGACATTCTCTGATTGAAACA [SEQ ID NO: 30];

ACTATAGGGCACGCGTGGTCGACGCCCGGGCTGGTGAATCTTTAGTC AGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTG ACATTCTCTGATTGAAACA [SEQ ID NO: 31];

GCTAAAAAAGTTGGGAGTATATTTATTGGGCTTCCTCCAACAGTTTCC AAAACCTCGCTCCTAAAACATAGAAGATAATTTTACGTTAGGAATCTTGAA CTATTTCTAAATCACTCC [SEQ ID NO: 32]; and

GGTGGACGCCCAAGGTCATCTCTTAGGTGTTCTCCATCTTCTTGGGGC GTGCTCCAACTTGGACCTGGGTCCCAAGGATATCTAACAAGCCTATGACAA CAGTGTAGCTCTTGCCAAAAAT [SEQ ID NO: 33].

18. A nucleic acid sequence according to claim 11, wherein the promoter sequence further comprises upstream of (g), (h) or (i) a nucleotide sequence selected from the group consisting of:

(j) CYVYRKKN$_j$WMDKSRYNVWRMMRN$_k$CYMRDVBHBRN$_l$YBRCN$_m$HH MSWMRHHMABRVRYARVWN$_n$RAVRMWRN$_o$BRVDWGRRRBRRCN$_p$RYTWYR GN$_q$TSDDTRDWWMVDMRN$_r$WARN$_s$KGGRSWYTWRHMHMYHGRKN$_t$VRRYV RSN$_u$WRGKRSDTGN$_v$WRSCWAMYRWKHMYWTADCTTDHBBKN$_w$RKYCWAN$_x$ AYCVCTMWTMHMN$_y$YYHWVWWGWKTCMYMAHRHKKYHMRWSRYYYRM MDSYCYMRHWN$_z$MRDTSBCN$_{aa}$YTRRBWRAWWATSVHCRN$_{ab}$TGHYTN$_{ac}$TYT YKWYN$_{ad}$AKYHVCN$_{ae}$YHRDTHMTGMADN$_{af}$RHYDTYKVKN$_{ag}$WMWMAN$_{ah}$VR YWHVVAWHBRN$_{ai}$YKCCYTRYYHAN$_{aj}$BWWSWWTN$_{ak}$ATCCWYN$_{al}$MTSYSRN$_{am}$YRWTKVWWAHGRMKHTCKKDWDYVYHYWSRBVWWVMRKTWWAYYAM N$_{an}$TYHTRAYN$_{ao}$YYRAN$_{ap}$GMAKAMWMVAMYWYRAN$_{aq}$AKTDGWYWKRRC MWYYVSMKTRGYWSBKAHDWN$_{ar}$YYTWGYN$_{as}$RGYMDWMYAMKGWRKTB KMCMWGBR [SEQ ID NO: 34];

(k) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 34 or a complement thereof; and

(l) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 34 or a complement thereof, under at least medium or high stringency conditions,

wherein:

A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

each of *k*, *l*, *p*, *q*, *v*, *ab*, *ad*, *ae*, *af*, *ag*, *ah*, *ai*, *al*, *ap*, *aq*, *ar* and *as* is an integer from 0 to 1;

each of *m*, *o*, *s*, *t*, *x*, *y*, *z*, *ac*, *aj*, *ak*, and *ao* is an integer from 0 to 2;

each of *r* and *w* is an integer from 0 to 3;

am is an integer from 0 to 5;

an is an integer from 0 to 7;

aa is an integer from 0 to 8;

n is an integer from 0 to 9;

j is an integer from 0 to 16; and

u is an integer from 0 to 42.

19. A nucleic acid sequence according to claim 18, wherein N_j comprises C or a sequence selected from: TCTTAAGGTTGAATA [SEQ ID NO: 35] or CAATCCGGTCG [SEQ ID NO: 36].

20. A nucleic acid sequence according to claim 18, wherein N_k is K.

21. A nucleic acid sequence according to claim 18, wherein N_l is B.

22. A nucleic acid sequence according to claim 18, wherein N_m is a sequence selected from TG, TT or GA.

23. A nucleic acid sequence according to claim 18, wherein N_n is a sequence selected from GA, GGA, AGA, GGAGAAAGA or ACAAGATAA.

24. A nucleic acid sequence according to claim 18, wherein N_o is a sequence selected from AG, TT or CT.

25. A nucleic acid sequence according to claim 18, wherein N_p is T or A.

26. A nucleic acid sequence according to claim 18, wherein N_q is A.

27. A nucleic acid sequence according to claim 18, wherein N_r is a sequence selected from CTT or CTC.

28. A nucleic acid sequence according to claim 18, wherein N_s is the sequence GT.

29. A nucleic acid sequence according to claim 18, wherein N_t is C or a sequence selected from CC or AA.

30. A nucleic acid sequence according to claim 18, wherein N_u is C or a sequence selected from TAGTGACTAAAGTTTAGTCCCTGAACTGAACTTTAATCC [SEQ ID NO: 37], TAGTGACTAAAGTTTAGTCCTTGAACTGAACTTTAATCC [SEQ ID NO: 38], TAGTGGACTGAACTTTAATCC [SEQ ID NO: 39] TTATCTCTAGAGGAGTCACTAAATTTTACACTAAACTT [SEQ ID NO: 40], or TGTTCAATAACATATTGATCTAAAACAAAACCCCAACCCTA [SEQ ID NO: 41].

31. A nucleic acid sequence according to claim 18, wherein N_v is G.

32. A nucleic acid sequence according to claim 18, wherein N_w is T or a sequence selected from TT, GCT or AC.

33. A nucleic acid sequence according to claim 18, wherein N_x is the sequence GG.

34. A nucleic acid sequence according to claim 18, wherein N_y is the sequence GT.

35. A nucleic acid sequence according to claim 18, wherein N_z is a sequence selected from AC or GA.

36. A nucleic acid sequence according to claim 18, wherein N_{aa} is a sequence selected from GA or GCAGCATC.

37. A nucleic acid sequence according to claim 18, wherein N_{ab} is C; N_{ac} is a sequence selected from GG or TG.

38. A nucleic acid sequence according to claim 18, wherein N_{ad} is N.

39. A nucleic acid sequence according to claim 18, wherein N_{ae} is G.

40. A nucleic acid sequence according to claim 18, wherein N_{af} is R.

41. A nucleic acid sequence according to claim 18, wherein N_{ag} is R,

42. A nucleic acid sequence according to claim 18, wherein N_{ah} is M.

43. A nucleic acid sequence according to claim 18, wherein N_{ai} is T

44. A nucleic acid sequence according to claim 18, wherein N_{aj} is a sequence selected from TA, AA or AT.

45. A nucleic acid sequence according to claim 18, wherein N_{ak} is A or a sequence selected from TA or CC; N_{al} is K.

46. A nucleic acid sequence according to claim 18, wherein N_{am} is a sequence selected from CAAT, GCAAT, TCA or AT.

47. A nucleic acid sequence according to claim 18, wherein N_{an} is a sequence selected from ACAAACA or AGACTGG.

48. A nucleic acid sequence according to claim 18, wherein N_{ao} is the sequence GG.

49. A nucleic acid sequence according to claim 18, wherein N_{ap} is K.

50. A nucleic acid sequence according to claim 18, wherein N_{aq} is G.

51. A nucleic acid sequence according to claim 18, wherein N_{ar} is A.

52. A nucleic acid sequence according to claim 18, wherein or N_{as} is T.

53. A nucleic acid sequence according to claim 18, wherein any one or more of N_j, N_k, N_l, N_m, N_n, N_o, N_p, N_q, N_r, N_s, N_t, N_u, N_v, N_w, N_x, N_y, N_z, N_{aa}, N_{ab}, N_{ac}, N_{ad}, N_{ae}, N_{af}, N_{ag}, N_{ah}, N_{ai}, N_{aj}, N_{ak}, N_{al}, N_{am}, N_{an}, N_{ao}, N_{ap}, N_{aq}, N_{ar} or N_{as} comprises 0 bases.

54. A nucleic acid sequence according to claim 18, which is selected from the group consisting of:

CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCA
TTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTAT
ATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAA
GTTTAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCT
TAACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGC
AAGTCATTTGCCGCCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGT
TTGGTTTTTTTCAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAA
AAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTT
ACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAA
TAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCTTGT
AGTAGTATACTGTATTCGACCTGCG [SEQ ID NO: 42];

CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCA
TTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTA
TATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAA
AGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGC
CCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTG

GCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCAT
GTTTGGTTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAA
AAAATTATCTCCTTGCTTATAGTTATATAATCCATGATCCGGCAATTGTTTGT
TACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGA
ATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCTTG
TAGTAGTATACTGTATTCGACCTGCG [SEQ ID NO: 43];

CTATAGGGCACGCGTGGTGACGGCCCGGGCTGGTCTGCCCCCTCATT
AATGCACAGAAGGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTT
TAGTCCCTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGACCTT
AACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCA
AGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTT
TGGTTTTTTTTAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAA
AATTATCTGCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTA
CGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAAT
AATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTACCTTGTA
GTAGTATACTGTATTCGACCTGCG [SEQ ID NO: 44];

CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCTGCTGCCCCTCA
TTAATGCACAGAAAGAGAGAGAGAGGGGGAGGGAAGACATTTTGGTCTTTA
TATAGTAGCTTTAATGGACTTTAGTACCTAGGTCCAAACCGGTAGTGGACTG
AACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTTTTTATTC
TAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCC
TCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCAA
CCAAATACTGCAAGCTATTGGGAAAAAAAGGCAAAAAATTATCTCCTTGCT
TATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAAT
CCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAG
TGGATATGACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTAT
TCGACCTGCG [SEQ ID NO: 45];

CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCA
TTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTA
TATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAA
AGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGC

CCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTG
GCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCAT
GTTTGGTTTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAA
AAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGT
TACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAGCTCAAGA
ATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCTTG
TAGTAGTATACTGTATTCGACCTGCG [SEQ ID NO: 46];

CTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCTCA
TTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTAT
ATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAA
GTTTAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGCCCT
TAACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGC
AAGTCATTTGCCACCCTCACCACCAGTGGCGACTGGTTAAATATCCTCATGT
TTGGTTTTTTTCAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAA
AAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTT
ACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAA
TAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCTTGT
AGTAGTATACTGTATTCGACCTGCG [SEQ ID NO: 47];

GTCTAGTTTCTTAAGGTTGAATAAAGGGACTAAACCATCTAAAAGAC
ACCCACTGCACCTCATTAATGCACAGAAGGAGAAAGAGAGAGAGAGTGAG
AGGGAGGGCATTTTGGTCTTTATATAGTAGCTTTAATGGACTTTAGTACATT
GATCAAATAGGCAGGGACTTGGAACCAAACATGCCCTTATCTTTTTTTTATT
CTAATCGCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACC
CTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTAGCA
ACCAAATACTGCAAGCTATTGGGGAAAAAAGGCAAAAAAACATCTCCTTGC
TTAGTTGTTTTAATCCATGATCCGGGCAATTGTTTGTTAAGGAGATCCTGAA
TCCTCTGACATAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAA
GTGGATATGACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTA
TTCGACCTGCGATTATGATAAAGGCATCAGCAACTAGAGTACTTTGTAGAAT
CTTTAGTCCGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCA
ATCATCTGACATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACT
AAACAAAACATATCTAACCAAATACTATGTGGTAATGTAGATTTATAAGCT

GTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATAC CTGCG [SEQ ID NO: 48];

CCCCGGTTCTATGGACACTGACACTAAAGTTTAGTGGCTTTTAGTCAC CAAGGGGTAGCTAAAAGTGTTTACATGGGACAACTGCTTTAGTGGGTGGAA CCAAATAGGGGGCACTAAACTCTAGAGGAGTCACTTATCTCTAGAGGAGTC ACTAAATTTTACACTAAACTTTAGTGGGTGGAACCAAACATGCCCTTATCTT GCCCTGCTAGTCAAACCACTAATCTCCCATGAAGATTCATCATGATGTCCAA CACTTGCATGTCTCGTTCATTCCCTTAAGTAATAATCACCATGCCTTCTTTAT AATTTGCTCATTCATGCATCATCGTCTCGATCTCAAAGTTCCCAACCACGCC CTACCCAAACTAGAATATCCTCTCTCTCATCATAATTCATATGACTCTCTTTT TTACATACATCTAAAAGTAAACCAAACAAACATCATAACTTGATGAAGAAA AAAATATGAGAGTTGTTAGAGCATCTCCAGTAGTTCCTAAAAATTTAGTTGG TAAACCAATGAGTTTTACAAGTG [SEQ ID NO: 49]; and

CCGTGTTACAATCCGGTCGACCTCGCCAAGAAGGCTAATCCCTGCTG CGAATCGAAGAACACAAGCAAGAACAAGATAAAACGCAACTCAATTGAAG TGACAATTACAGATGAATGATTAAGCACTCAAAGTTGGGGTCTTGCAAACC GATAACGGCGACTGTTCAATAACATATTGATCTAAAACAAAACCCCAACCC TAAAGTAGGTGGTGGCTACTGATTATATAGCCTAAGGGACGTCCAAGGATC CCTCTTCACGTCCTAAATGTGTCCCAACACGTTTCAAGGCCCAACGGCCCAA AAGAAGATGTCGCAGCATCCTGACAGATTATGGACGCTGACTTGTTTCGAC GATTCCCGTTGGTTCTGAAGAGATTTTGATGTCAAACCACTTGGATTGGTTT CCTTATCAAATTATCTTTCCATCCATATGTGGATCATTGAAAACGGAGTTCG GATGCGTCCTGGGCATCCGTTTTATTACAGACTGGTCCTGATGGCCGAGGCA GACTCGAACTCGAATTAGACTGGGCCTCTGGCTTGGCTGGGACGTCCTTGCT GGCCTTCTAAGGTGGTGGCCAAGGA [SEQ ID NO: 50].

55. A nucleic acid sequence according to claim 18, wherein the promoter sequence further comprises upstream of (j), (k) or (l) a nucleotide sequence selected from the group consisting of:

(m) YDKHSN$_{at}$GKHHNDSRTDNHYDASN$_{au}$YN$_{av}$NKKHKRVNN$_{aw}$CKDSWB WYSGMYYYN$_{ax}$BYWYASHHHMYRRAKN$_{ay}$CWARWHSNGYRRMTWHARMDY VACKHVWN$_{az}$YYKRTTWWSWM [SEQ ID NO: 51];

(n) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 51 or a complement thereof; and

(o) a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 51 or a complement thereof, under at least medium or high stringency conditions,

wherein:

A, C, G, T, M, R, W, S, Y, K, V, H, D, B and N are as defined above;

each of *at* and *av* is an integer from 0 to 1;

each of *aw* and *ax* is an integer from 0 to 2; and

each of *au*, *ay* and *az* is an integer from 0 to 3;

56. A nucleic acid sequence according to claim 55, wherein N_{at} comprises C, G or A.

57. A nucleic acid sequence according to claim 55, wherein N_{au} comprises a sequence selected from GGA, GGC or GAT.

58. A nucleic acid sequence according to claim 55, wherein N_{av} comprises C or T.

59. A nucleic acid sequence according to claim 55, wherein N_{aw} comprises a sequence selected from AC, CC or AG.

60. A nucleic acid sequence according to claim 55, wherein N_{ax} comprises the sequence AG.

61. A nucleic acid sequence according to claim 55, wherein N_{ay} comprises a sequence selected from TGC or ACC.

62. A nucleic acid sequence according to claim 55, wherein N_{az} comprises A or a sequence selected from CGC or CCT.

63. A nucleic acid sequence according to claim 55, wherein any one or more of N_{at}, N_{au}, N_{av}, N_{aw}, N_{ax}, N_{ay} or N_{az} comprises 0 nucleotides.

64. A nucleic acid sequence according to claim 55, which is selected from the group consisting of:

TAGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTGGCC CCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTG ATTTAGTC [SEQ ID NO: 52];

TGGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTGGCC TCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTG ATTTAGTC [SEQ ID NO: 53];

TATAGGGNACNCGTGNTCGACGGCCGGGCTGGTCTGGTTTTGGCTCC TTTTANTCACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGAT TTAGTC [SEQ ID NO: 54];

TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGCCT CTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA TTTAGTC [SEQ ID NO: 55];

TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGCCC CTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA TTTAGTC [SEQ ID NO: 56];

TGGCCCGGACGACATTTTTGAGGGACCCGTTTAGCACCTGGTTTTGGC CCCTTTTAGTCACTAAATTGCCAAAAAGGGTGACTAAAAAGTAACTAAAAC TGATTTAGTC [SEQ ID NO: 57];

CGTAGGGTCATTCATACCCGAGCCAGTATACACCCTTGTGTCCGCTTT GCACACCATCCGGATCTAGACGCGCAGATACAAATTCACTCGTCGCTCTATT TAGAC [SEQ ID NO: 58]; and

CTGTCAGGTTAAGATAAACAACGATTTGTTCGAAGAGCGACACACCG ATCCAGCTTCAGATCATAAAGACCCAAATTCTGCAACTTTAGCACCACGTCT CCTCTGGTTATCAA [SEQ ID NO: 59].

65. A nucleic acid sequence according to claim 18, wherein an intervening sequence is interposed between the nucleotide sequence according to any one of (d), (e) or (f) and the nucleotide sequence according to the nucleotide sequence according to any one of (g), (h) or (i).

66. A nucleic acid sequence according to claim 65, wherein the intervening sequence is selected from the group consisting of:

GGTGTGCTCACTTCGCGAAAAGTAGGCGAAAAGTTACTGTAGCATTT TTCGTTGTTATTTGGCAAAATTTGTCCAATCATGGACTAACTAGGCTCAAAA GCTTCGTCCCAACATTTACAGACAAACTATGCAATTAGTTATCTTTTTTACCT ACATTTAATGCTCCATGCATACGTCGAAAGATTTGATGTGATGGGTGTAAGA

GGAAATATTTTGGGTTGGGATGGGGAAGTGAACAAGGCCTCGG [SEQ ID NO: 60], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 60 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 60 or a complement thereof, under at least medium or high stringency conditions;

CACTTTTAGACTTTCTTTCTCTCTTATATATTTGTATGTACTAGGATCC TGTCTTACTCTTTATTCTTCCTATATGCTTCCACCTCCAGATTGGCACATTGA TACGTCCGTGTATATTTCCGCACGATTGGATTGATTTCCCAAATGCCGAGAG ATTGAGAGTTGAGATAGGGAGTTCTTGGAGAGCAAATTTTTTCAACCTTGCT AAAAAATCTAGATTGCGAGTGGG [SEQ ID NO: 61] , or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 61 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 61 or a complement thereof, under at least medium or high stringency conditions; and

GATAGAATATATTAGTGATTTGGAGGCCTTGCCTACGTGATATTGAG ATGGGAACAGATCTATTTGAAAGCTTATCAAACAAGCTTTCCATCAAGTGCT CATTGACCTCGATCGCACTCCGGATGGATGAGTTATGGCCTTCCCAATGAGG CACTGTCGGGCTGCCGACGAACCGAAAATTCAACTTTGATGAACTTCCATTA TGAAACACATTTGAACCTACAATCAAATAGTGACATGTACA [SEQ ID NO: 62] , or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 62 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 62 or a complement thereof, under at least medium or high stringency conditions.

67. A nucleic acid sequence according to any one of claim 1 to 66, wherein the promoter sequence has a structure represented by formula (I):

$$\Theta\text{–}\Delta\text{–}Z\text{–}\Psi\text{–}\Phi\text{–}\Omega \qquad (I)$$

wherein:

Θ is any one of (m), (n) or (o);

Δ is any one of (j), (k) or (l);

Z is any one of (g), (h) or (i);

Ψ is an optional spacer or intervening sequence;

Φ is any one of (d), (e) or (f); and

Ω is any one of (a), (b) or (c).

68. A nucleic acid sequence according to claim 67, wherein the promoter sequence comprises a nucleotide sequence selected from the group consisting of:

TAGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTGGCC
CCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTG
ATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCC
CTCATTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCT
TTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGAC
TAAAGTTTAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACAT
GCCCTTAACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAA
GTGGCAAGTCATTTGCCGCCCTCACTACCAGTGGCGACTGGTTAAATATCCT
CATGTTTGGTTTTTTTCAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGG
CAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGT
TTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTC
AAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTAT
CTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAAA
CAAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTA
GGAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATTGAAACTCGGTTT
CCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTGG
TAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGG
CAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAAC
AGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTT
CCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTA
CAACATCTT [SEQ ID NO: 63], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 63 or a complement thereof, or a

nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 63 or a complement thereof, under at least medium or high stringency conditions;

TGGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTGGCC
TCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTG
ATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCC
CTCATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGT
CTTTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTG
ACTAAAGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAA
CATGCCCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACA
AAGTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATC
CTCATGTTTGGTTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAAAG
GCAAAAAATTATCTCCTTGCTTATAGTTATATAATCCATGATCCGGCAATTG
TTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACT
CAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTT
ATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGA
AACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGAC
TAGGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAACTCGGT
TTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTG
GTAGTGTAGATTTATATGCTGTGTACTGAAAGTGGCGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAA
CAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATT
CCCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCT
ACAACATCTT [SEQ ID NO: 64], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 64 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 64 or a complement thereof, under at least medium or high stringency conditions;

CTATAGGGCACGCGTGGTGACGGCCCGGGCTGGTCTGCCCCCTCATT
AATGCACAGAAGGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTT
TAGTCCCTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGACCTT
AACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCA
AGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTT

TGGTTTTTTTTAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAA
AATTATCTGCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTA
CGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAAT
AATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTACCTTGTA
GTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAAACTAGA
GTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTAGGAAA
CTTATACTGTTGCAATCATCTGACATTCTCTGATTGAAACTCGGGGTGTGCT
CACTTCGCGAAAAGTAGGCGAAAAGTTACTGTAGCATTTTTCGTTGTTATTT
GGCAAAATTTGTCCAATCATGGACTAACTAGGCTCAAAAGCTTCGTCCCAA
CATTTACAGACAAACTATGCAATTAGTTATCTTTTTTACCTACATTTAATGCT
CCATGCATACGTCGAAAGATTTGATGTGATGGGTGTAAGAGGAAATATTTT
GGGTTGGGATGGGGAAGTGAACAAGGCCTCGGTTTCCAAAAATCATATGTT
ACTAAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTAAAT
GCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGA
TACCTGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATAC
TTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATA
ATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 65], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 65 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 65 or a complement thereof, under at least medium or high stringency conditions;

TATAGGGNACNCGTGNTCGACGGCCGGGCTGGTCTGGTTTTGGCTCC
TTTTANTCACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGAT
TTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCTGCTGCCCCTC
ATTAATGCACAGAAAGAGAGAGAGAGGGGGAGGGAAGACATTTTGGTCTTT
ATATAGTAGCTTTAATGGACTTTAGTACCTAGGTCCAAACCGGTAGTGGACT
GAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTTTTTATT
CTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACC
CTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTAGCA
ACCAAATACTGCAAGCTATTGGGAAAAAAAGGCAAAAAATTATCTCCTTGC
TTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAA
TCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAA

GTGGATATGACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTA
TTCGACCTGCGATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAGAAT
CTTTAGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCA
ATCCTCTGACATTCTCTGATTGAAACTCGGTTTCCAAAAATCATATGTTACT
AAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTATATGCTG
TGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACC
TGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGT
CCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATAC
ATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 66], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 66 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 66 or a complement thereof, under at least medium or high stringency conditions;

TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGCCT
CTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA
TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCT
CATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCT
TTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGAC
TAAAGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAAACA
TGCCCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAA
GTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCT
CATGTTTGGTTTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAAAGG
CAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGT
TTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAGCTC
AAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTAT
CTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAAA
CTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTA
GGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAACTCGGTTT
CCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTGG
TAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGG
CAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAAC
AGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTT

CCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTA
CAACATCTT [SEQ ID NO: 67], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 67 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 67 or a complement thereof, under at least medium or high stringency conditions;

TATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGCCC
CTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTGA
TTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCCCCT
CATTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTT
ATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTA
AAGTTTAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAACATGC
CCTTAACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGT
GGCAAGTCATTTGCCACCCTCACCACCAGTGGCGACTGGTTAAATATCCTCA
TGTTTGGTTTTTTTCAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGGCA
AAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTT
GTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAA
GAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCT
TGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAAAC
AAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTAG
GAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATTGAAACTCGGTTTCC
AAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTGGTA
GTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCA
GAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACAG
GATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCC
AAACAGCCCATAATACATTCTGCACCCTTTATAAAACCTCTACTAGCTACAA
CATCTT [SEQ ID NO: 68], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 68 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 68 or a complement thereof, under at least medium or high stringency conditions;

TGGCCCAGCATTGGTGTCCCGGTGGCTGCTAGCCGCCACTAGCCTGC
TGGGGTTCTGCGGCTAACCGGGAACATCTGCAGGCAAACTAGTGGCCCGGA

CGACATTTTTGAGGGACCCGTTTAGCACCTGGTTTTGGCCCCTTTTAGTCACT
AAATTGCCAAAAAGGGTGACTAAAAAGTAACTAAAACTGATTTAGTCGTCT
AGTTTCTTAAGGTTGAATAAAGGGACTAAACCATCTAAAAGACACCCACTG
CACCTCATTAATGCACAGAAGGAGAAAGAGAGAGAGAGTGAGAGGGAGGG
CATTTTGGTCTTTATATAGTAGCTTTAATGGACTTTAGTACATTGATCAAATA
GGCAGGGACTTGGAACCAAACATGCCCTTATCTTTTTTTTATTCTAATCGCT
CTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCACCCTCACTACC
AGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTAGCAACCAAATAC
TGCAAGCTATTGGGGAAAAAAGGCAAAAAAACATCTCCTTGCTTAGTTGTT
TTAATCCATGATCCGGGCAATTGTTTGTTAAGGAGATCCTGAATCCTCTGAC
ATAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATATG
ACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCTG
CGATTATGATAAAGGCATCAGCAACTAGAGTACTTTGTAGAATCTTTAGTCC
GTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTGA
CATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACTAAACAAAAC
ATATCTAACCAAATACTATGTGGTAATGTAGATTTATAAGCTGTGTACTGAA
AGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTC
TGAATACCACAACCATAAAAAAACAGGATGATGTTATACTTGTCCCCTTCCAT
GATACAGGACTGTTTAGTAACTTCCCAAACAGCCCATGATACATTCTGCACC
CTTTATTAAACCTCTAGCTACATCATCTT [SEQ ID NO: 69], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 69 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 69 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGCACGCGTGGTCGACGCCCGGGCTGGTGAATCTTTAGTC
AGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTG
ACATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACTAAACAAAA
CAATATATCTAACCAAATACTATGTGGTAGTGTAGATTTATAAGCTGTGTAC
TGAAAGTGACGTCAAGTATAGTAGTGGCAGAGATTCAAAAGATACCTTCGG
ATTCTGAATACCACAACCATAAAAAAACAGGATGATGTTATACTTGTCCCCTT
CCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATGATACATTCTG
CACCCTTTATTAAACCTCTACTAGCTACAACATCTT [SEQ ID NO: 70], or a

nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 70 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 70 or a complement thereof, under at least medium or high stringency conditions;

CGTAGGGTCATTCATACCCGAGCCAGTATACACCCTTGTGTCCGCTTT
GCACACCATCCGGATCTAGACGCGCAGATACAAATTCACTCGTCGCTCTATT
TAGACCCCCGGTTCTATGGACACTGACACTAAAGTTTAGTGGCTTTTAGTCA
CCAAGGGGTAGCTAAAAGTGTTTACATGGGACAACTGCTTTAGTGGGTGGA
ACCAAATAGGGGGCACTAAACTCTAGAGGAGTCACTTATCTCTAGAGGAGT
CACTAAATTTTACACTAAACTTTAGTGGGTGGAACCAAACATGCCCTTATCT
TGCCCTGCTAGTCAAACCACTAATCTCCCATGAAGATTCATCATGATGTCCA
ACACTTGCATGTCTCGTTCATTCCCTTAAGTAATAATCACCATGCCTTCTTTA
TAATTTGCTCATTCATGCATCATCGTCTCGATCTCAAAGTTCCCAACCACGC
CCTACCCAAACTAGAATATCCTCTCTCTCATCATAATTCATATGACTCTCTTT
TTTACATACATCTAAAAGTAAACCAAACAAACATCATAACTTGATGAAGAA
AAAAATATGAGAGTTGTTAGAGCATCTCCAGTAGTTCCTAAAAATTTAGTTG
GTAAACCAATGAGTTTTACAAGTGGCTAAAAAAGTTGGGAGTATATTTATT
GGGCTTCCTCCAACAGTTTCCAAAACCTCGCTCCTAAAACATAGAAGATAAT
TTTACGTTAGGAATCTTGAACTATTTCTAAATCACTCCAACCACTTTTAGACT
TTCTTTCTCTCTTATATATTTGTATGTACTAGGATCCTGTCTTACTCTTTATTC
TTCCTATATGCTTCCACCTCCAGATTGGCACATTGATACGTCCGTGTATATTT
CCGCACGATTGGATTGATTTCCCAAATGCCGAGAGATTGAGAGTTGAGATA
GGGAGTTCTTGGAGAGCAAATTTTTTCAACCTTGCTAAAAAATCTAGATTGC
GAGTGGGTTTCGGAAACTCTTGCATATGCTCTTATATTTCATATCTAACCAA
ATACTATGTGGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAG
TATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAA
CCATAAAAAAACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTG
TTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACC
TCTACTAGCTACAACATCTT [SEQ ID NO: 71], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 71 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth

in SEQ ID NO: 71 or a complement thereof, under at least medium or high stringency conditions; and

CTGTCAGGTTAAGATAAACAACGATTTGTTCGAAGAGCGACACACCG
ATCCAGCTTCAGATCATAAAGACCCAAATTCTGCAACTTTAGCACCACGTCT
CCTCTGGTTATCAACCGTGTTACAATCCGGTCGACCTCGCCAAGAAGGCTAA
TCCCTGCTGCGAATCGAAGAACACAAGCAAGAACAAGATAAAACGCAACTC
AATTGAAGTGACAATTACAGATGAATGATTAAGCACTCAAAGTTGGGGTCT
TGCAAACCGATAACGGCGACTGTTCAATAACATATTGATCTAAAACAAAAC
CCCAACCCTAAAGTAGGTGGTGGCTACTGATTATATAGCCTAAGGGACGTC
CAAGGATCCCTCTTCACGTCCTAAATGTGTCCCAACACGTTTCAAGGCCCAA
CGGCCCAAAAGAAGATGTCGCAGCATCCTGACAGATTATGGACGCTGACTT
GTTTCGACGATTCCCGTTGGTTCTGAAGAGATTTTGATGTCAAACCACTTGG
ATTGGTTTCCTTATCAAATTATCTTTCCATCCATATGTGGATCATTGAAAACG
GAGTTCGGATGCGTCCTGGGCATCCGTTTTATTACAGACTGGTCCTGATGGC
CGAGGCAGACTCGAACTCGAATTAGACTGGGCCTCTGGCTTGGCTGGGACG
TCCTTGCTGGCCTTCTAAGGTGGTGGCCAAGGAGGTGGACGCCCAAGGTCA
TCTCTTAGGTGTTCTCCATCTTCTTGGGGCGTGCTCCAACTTGGACCTGGGTC
CCAAGGATATCTAACAAGCCTATGACAACAGTGTAGCTCTTGCCAAAAATA
GCGATAGAATATATTAGTGATTTGGAGGCCTTGCCTACGTGATATTGAGATG
GGAACAGATCTATTTGAAAGCTTATCAAACAAGCTTTCCATCAAGTGCTCAT
TGACCTCGATCGCACTCCGGATGGATGAGTTATGGCCTTCCCAATGAGGCAC
TGTCGGGCTGCCGACGAACCGAAAATTCAACTTTGATGAACTTCCATTATGA
AACACATTTGAACCTACAATCAAATAGTGACATGTACATGTGGAACCAAGT
GAATTATAACCAAAGTCACTATACAAATGTAGGAACGAGCACACCTTATAA
TCATTGTTCATATCCGAAGGTGTCATGTCCTCATCAAGTATAGCAGTGGCAG
AGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCTTAAAAAACAGG
ATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCA
AACTGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAAC
ATCTT [SEQ ID NO: 72], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 72 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 72 or a complement thereof, under at least medium or high stringency conditions.

69. A nucleic acid sequence according to claim 67, wherein the promoter sequence comprises a nucleotide sequence selected from the group consisting of:

ACTAGTGATTACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTC
AGTTTCAGTGCAGGCATATGTTTTTGTTTCCCGTCTACTTCTATTAGGGTATC
ACCGAGAACTTATTAAAACAACTCCCTAATTCATTTTTTACAAAAAGGAACC
GTAAATCCTGTATCTGAACTTTCTAAAAGATTAGCATGCATTTGGTAAACTC
TCCCCTCCAATTCCCAGGTATGCGAATGCACGGTAAACTAGGATGCTGGCC
AGCATTGGTGTCCGGTGGCTGCTAGCCGCCACTAGCTTGCTGGGGTTCTACG
GCTAACCGGGAACATCTGCACACAACTAGTAGTCGGATGACATTTTTTAGG
GACCTGTTTGGCACCTGGTTTTGGCCCCTTTTAGTTACTAAATTGCCAAAAA
GAGTGACTAAAAAGTGACTAAACTGATTTAGTCCTCTAGTCAAGGGACTAA
ACCAGCTAAAAGACATCCGCTGCCCCTCATTAATGCACAGGAGAGAGAGAG
AGGGAGAGGGAGGACATTTTGGTCTTTATATAGTAGCTTTAATGGACTTTAG
TACCTAGATCCAAACCGGTAGTGACTAAAGTTTAGTCCCTGAACTGAACTTT
AATCCAGGGACATGAACCAAACATGCCCTTAACTTTTTTTTTATTCTAATCC
CTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCCGCCCTCACTA
CCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTCAGCAACCAAA
TACTGCAAGCTATTGGGAAAAAAGGCAAAAAATTATCTCCTTGCTTATAGTT
GTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCTGAATCCTCTGA
CGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTATAAAGTGGATAT
GACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACTGTATTCGACCT
GCGATTATGATAAAGGCATCAGAAACAAGAGTACTTTCTAGAATCTTTAGT
CAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCCTTT
GACATTCTCTGATTGAAACTCGGTTTCCAAAAATCATATGTTACTAAACAAA
ACATATCTAACCAAATACTATGTGGTAGTGTAGATTTATATGCTGTGTACTG
AAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATACCTGCGGAT
TCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTTCC
ATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAATACATTCTGCA
CCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCATCTTGTCTAGTTG
GACAGAAGTTCTCTCTTTCTTGGCTGACTCCAACTTACTACACCGCAACTTC
TTGTGCCCTTGTT [SEQ ID NO: 73], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 73 or a complement thereof, or a

nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 73 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTAGTAATAGTAA
AATATCCGGTCTCTACTCACTCTATACACTTAGGACGTGTTTGGCTCCCGTC
GCACATGCATGTCGCCATAGCTATGGCTGCACCGTGAGTGCGGTGGATGAA
TCGTCCGCCACACATTTGGCAAGTTGTGGTGTAAAAATGAATTGGAGCGTGT
GGTGGCTGCCAAAACTTCAGCGAGTTGCAGTGTGGGTGTTTGGCCGATGCG
GCATGCCACACTTGTGGCACAGCGAGCTATGGCGTGCTGTGGAGGTGAGCC
AAACATGCCTTTATATTCTCTGATTGACAGTTTCAGTGCAGACATATGTTTTT
GTTTCCCGTCTACTTCTATTAGGGTATCACTGAGAACTTATTAAAACAACTC
CCTAATTCATTTTTTACAAAAAGGAACCGTAAATTATGTATCTGAACTTTCT
AAAAGATTAGCATGCATTTGGTAAGCTCTCCCCTCCAATTCCCAGGTATGCG
AATGCACGGTAAACTAGGATGCTGGCCAGCATTGGTGTCCGGTGGCTGCTA
GCCGCCACTAGCCTGCTGGGGTTCTGCGGCTAACCGGGAACATCTGCACAC
AACTAGTGGTCGGATGACATTTTTTAGGGACCTGTTTGGCACCTGGTTTTGG
CCTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACT
GATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCC
CCTCATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGG
TCTTTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGT
GACTAAAGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAA
ACATGCCCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCAC
AAAGTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATAT
CCTCATGTTTGGTTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAAA
GGCAAAAAATTATCTCCTTGCTTATAGTTATATAATCCATGATCCGGCAATT
GTTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAAC
TCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTT
ATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGA
AACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGAC
TAGGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAACTCGGT
TTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTG
GTAGTGTAGATTTATATGCTGTGTACTGAAAGTGGCGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAA
CAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATT

CCCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGCT
ACAACATCTTACTCCATCTTGTCTAGTTGGACAAGTTCTCTCTTTCTTGGCTG
ACTCCAACTTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 74], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 74 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 74 or a complement thereof, under at least medium or high stringency conditions;

CTATAGGGCACGCGTGGTGACGGCCCGGGCTGGTCTGCCCCCTCATT
AATGCACAGAAGGAGAGAGAGGGAGAGGGAGGACATTTTGGTCTTTATATA
GTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTGACTAAAGTT
TAGTCCCTGAACTGAACTTTAATCCAGGGACATGGAACCAAACATGACCTT
AACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCA
AGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTT
TGGTTTTTTTTAGCAACCAAATACTGCAAGCTATTGGGAAAAAAGGCAAAA
AATTATCTGCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTA
CGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAAT
AATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTACCTTGTA
GTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGAAACTAGA
GTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTAGGAAA
CTTATACTGTTGCAATCATCTGACATTCTCTGATTGAAACTCGGGGTGTGCT
CACTTCGCGAAAAGTAGGCGAAAAGTTACTGTAGCATTTTTCGTTGTTATTT
GGCAAAATTTGTCCAATCATGGACTAACTAGGCTCAAAAGCTTCGTCCCAA
CATTTACAGACAAACTATGCAATTAGTTATCTTTTTTACCTACATTTAATGCT
CCATGCATACGTCGAAAGATTTGATGTGATGGGTGTAAGAGGAAATATTTT
GGGTTGGGATGGGGAAGTGAACAAGGCCTCGGTTTCCAAAAATCATATGTT
ACTAAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTAAAT
GCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGA
TACCTGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATAC
TTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATA
ATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCAT
CTTGTCTAGTTGGACAGAAGTTCTCTCTTTCTTGGCTGACTCCAACTTACTAC
ACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 75], or a nucleotide sequence

that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 75 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 75 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGNACNCGTGNTCGACGGCCGGGCTGGTCTGGTTTTGGCT
CCTTTTANTCACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACTG
ATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCTGCTGCCCC
TCATTAATGCACAGAAAGAGAGAGAGAGGGGGAGGGAAGACATTTTGGTC
TTTATATAGTAGCTTTAATGGACTTTAGTACCTAGGTCCAAACCGGTAGTGG
ACTGAACTTTAATCCAGGGACATGGAACCAAACATGCCCTTAACTTTTTTTT
ATTCTAATCCCTCTTACATTCACTTGTCTCACAAAGTGGCAAGTCATTTGCC
ACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCATGTTTGGTTTTTTTTA
GCAACCAAATACTGCAAGCTATTGGGAAAAAAAGGCAAAAAATTATCTCCT
TGCTTATAGTTGTATAATCCATGATCCGGCAATTGTTTGTTACGGAGATCCT
GAATCCTCTGACGTAGAGTTTAATCAATTTTAACTCAAGAATAATACACTAT
AAAGTGGATATGACAATCACCGTAGTACTTATTTATCTTGTAGTAGTATACT
GTATTCGACCTGCGATTATGATAAAGGCATCAGAAACTAGAGTACTTTCTAG
AATCTTTAGTCAGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTT
GCAATCCTCTGACATTCTCTGATTGAAACTCGGTTTCCAAAAATCATATGTT
ACTAAACAAAACATATCTAACCAAATACTATGTGGTAGTGTAGATTTATATG
CTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGAT
ACCTGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACT
TGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAA
TACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCATC
TTGTCTAGTTGGACAAGTTCTCTCTTTNTTGGCNGANTNCAACTTACTACAC
CGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 76], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 76 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 76 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGC
CTCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACT
GATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCC
CCTCATTAATGCACAGAAGGAGAGAGAGAGGGAGAGGGAGGACATTTTGG
TCTTTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGT
GACTAAAGTTTAGTCCTTGAACTGAACTTTAATCCAGGGACATGGAACCAA
ACATGCCCTTAACTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCAC
AAAGTGGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATAT
CCTCATGTTTGGTTTTTTTTAGTAACCAAATACTGCAAGCTATTGGGAAAAA
AGGCAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAAT
TGTTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAG
CTCAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATT
TATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAG
AAACTAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGA
CTAGGAAACTTATACTGTTGCAATCCTCTGACATTCTCTGATTGAAACTCGG
TTTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGT
GGTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGT
GGCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAA
ACAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAAT
TTCCCAAACAGCCCATAATACATTCTGCACCCTTTATTAAACCTCTACTAGC
TACAACATCTTACTCCATCTTGTCTAGTTGGACAAGTTCTCTCTTTCTTGGCT
GACTCCAACTTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 77],

or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 77 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 77 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTGGTTTTGGC
CCCTTTTAGTTACTAAATTGCCAAAAAGAGTGACTAAAAAGTGACTAAACT
GATTTAGTCCTCTAGTCAAGGGACTAAACCAGCTAAAAGACATCCGCTGCC
CCTCATTAATGCACAGGAGAGAGAGAGAGGGAGAGGGAGGACATTTTGGT
CTTTATATAGTAGCTTTAATGGACTTTAGTACCTAGATCCAAACCGGTAGTG
ACTAAAGTTTAGTCCCTGAACTGAACTTTAATCCAGGGACATGAACCAAAC

ATGCCCTTAACTTTTTTTTTATTCTAATCCCTCTTACATTCACTTGTCTCACAA
AGTGGCAAGTCATTTGCCACCCTCACCACCAGTGGCGACTGGTTAAATATCC
TCATGTTTGGTTTTTTTCAGCAACCAAATACTGCAAGCTATTGGGAAAAAAG
GCAAAAAATTATCTCCTTGCTTATAGTTGTATAATCCATGATCCGGCAATTG
TTTGTTACGGAGATCCTGAATCCTCTGACGTAGAGTTTAATCAATTTTAACT
CAAGAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTT
ATCTTGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGA
AACAAGAGTACTTTCTAGAATCTTTAGTCAGTTTCTGTAAGATGAACGTGAC
TAGGAAACTTATACTGTTGCAATCCTTTGACATTCTCTGATTGAAACTCGGT
TTCCAAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTG
GTAGTGTAGATTTATATGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTG
GCAGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAA
CAGGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATT
TCCCAAACAGCCCATAATACATTCTGCACCCTTTATAAAACCTCTACTAGCT
ACAACATCTTACTCCATCTTGTCTAGTTGGACAGAAGTTCTCTCTTTCTTGGC
TGACTCCAACTTACTACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 78],
or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 78 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 78 or a complement thereof, under at least medium or high stringency conditions;

ACATCGTTTTTGTTTCGTTAACTCTATTAAGGTTCACGAGAACTATAA
ACCAACTCCCTAAATTTATTTTTTACAAAAAGGAACCATAAATCTGTATTTC
GACCTTTCTAAAAGAATAGCATGGCATTGTAAACTCTCCCCTCCCAATTCCC
AGGTATGCGAATGCATGGTAAACTAGGTGCTGGCCCAGCATTGGTGTCCCG
GTGGCTGCTAGCCGCCACTAGCCTGCTGGGGTTCTGCGGCTAACCGGGAAC
ATCTGCAGGCAAACTAGTGGCCCGGACGACATTTTTGAGGGACCCGTTTAG
CACCTGGTTTTGGCCCCTTTTAGTCACTAAATTGCCAAAAAGGGTGACTAAA
AAGTAACTAAAACTGATTTAGTCGTCTAGTTTCTTAAGGTTGAATAAAGGGA
CTAAACCATCTAAAAGACACCCACTGCACCTCATTAATGCACAGAAGGAGA
AAGAGAGAGAGAGTGAGAGGGAGGGCATTTTGGTCTTTATATAGTAGCTTT
AATGGACTTTAGTACATTGATCAAATAGGCAGGGACTTGGAACCAAACATG
CCCTTATCTTTTTTTTATTCTAATCGCTCTTACATTCACTTGTCTCACAAAGT

GGCAAGTCATTTGCCACCCTCACTACCAGTGGCGACTGGTTAAATATCCTCA
TGTTTGGTTTTTTTAGCAACCAAATACTGCAAGCTATTGGGGAAAAAAGGCA
AAAAAACATCTCCTTGCTTAGTTGTTTTAATCCATGATCCGGGCAATTGTTT
GTTAAGGAGATCCTGAATCCTCTGACATAGAGTTTAATCAATTTTAACTCAA
GAATAATACACTATAAAGTGGATATGACAATCACCGTAGTACTTATTTATCT
TGTAGTAGTATACTGTATTCGACCTGCGATTATGATAAAGGCATCAGCAACT
AGAGTACTTTGTAGAATCTTTAGTCCGTTTCTGTAAGATGAACGTGACTAGG
AAACTTATACTGTTGCAATCATCTGACATTCTCTGATTGAAACACGGTTTCC
AAAAATCATATGTTACTAAACAAAACATATCTAACCAAATACTATGTGGTA
ATGTAGATTTATAAGCTGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGC
AGAGACTCAAAAGATACCTGCGGATTCTGAATACCACAACCATAAAAAACA
GGATGATGTTATACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAACTTC
CCAAACAGCCCATGATACATTCTGCACCCTTTATTAAACCTCTAGCTACATC
ATCTTACTCCATCTTGTCTAGTTGGACAGAAGTTCTCTCTTTCTTGGCTGACT
CCAACTACTACACCGCAACTTCTTGTGTCCTTGTTCCAACCATCACAACTGA
GC [SEQ ID NO: 79], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 79 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 79 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGCACGCGTGGTCGACGCCCGGGCTGGTGAATCTTTAGTC
AGTTTCTGTAAGATGAACGTGACTAGGAAACTTATACTGTTGCAATCATCTG
ACATTCTCTGATTGAAACACGGTTTCCAAAAATCATATGTTACTAAACAAAA
CAATATATCTAACCAAATACTATGTGGTAGTGTAGATTTATAAGCTGTGTAC
TGAAAGTGACGTCAAGTATAGTAGTGGCAGAGATTCAAAAGATACCTTCGG
ATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTTGTCCCCTT
CCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATGATACATTCTG
CACCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCATCTTGTCTAGT
TGGACAGAAGTTCTCTCTTTCTTGGCTGACTCCAACTACTGCACCGCAACTT
CTTGTGCCCTTGTTCCAACCATCACAATTGAGC [SEQ ID NO: 80], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 80 or a complement thereof, or a nucleotide sequence that hybridizes to the

sequence set forth in SEQ ID NO: 80 or a complement thereof, under at least medium or high stringency conditions;

ACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTCTCACTTCGAC
GAGAAGACACAACACACACCACCTCTCAGAAGAGACCTAGGATCTGCTC
CCTCTCTCGAAAAGCTTGTGACCCCTACTACAAGTACTCCGGTGTAAGGCAA
CACGAGCAGCGGCAAAACCCACTGGACGTAGGGTCATTCATACCCGAGCCA
GTATACACCCTTGTGTCCGCTTTGCACACCATCCGGATCTAGACGCGCAGAT
ACAAATTCACTCGTCGCTCTATTTAGACCCCCGGTTCTATGGACACTGACAC
TAAAGTTTAGTGGCTTTTAGTCACCAAGGGGTAGCTAAAAGTGTTTACATGG
GACAACTGCTTTAGTGGGTGGAACCAAATAGGGGGCACTAAACTCTAGAGG
AGTCACTTATCTCTAGAGGAGTCACTAAATTTTACACTAAACTTTAGTGGGT
GGAACCAAACATGCCCTTATCTTGCCCTGCTAGTCAAACCACTAATCTCCCA
TGAAGATTCATCATGATGTCCAACACTTGCATGTCTCGTTCATTCCCTTAAG
TAATAATCACCATGCCTTCTTTATAATTTGCTCATTCATGCATCATCGTCTCG
ATCTCAAAGTTCCCAACCACGCCCTACCCAAACTAGAATATCCTCTCTCTCA
TCATAATTCATATGACTCTCTTTTTTACATACATCTAAAAGTAAACCAAACA
AACATCATAACTTGATGAAGAAAAAAATATGAGAGTTGTTAGAGCATCTCC
AGTAGTTCCTAAAAATTTAGTTGGTAAACCAATGAGTTTTACAAGTGGCTAA
AAAAGTTGGGAGTATATTTATTGGGCTTCCTCCAACAGTTTCCAAAACCTCG
CTCCTAAAACATAGAAGATAATTTTACGTTAGGAATCTTGAACTATTTCTAA
ATCACTCCAACCACTTTTAGACTTTCTTTCTCTCTTATATATTTGTATGTACT
AGGATCCTGTCTTACTCTTTATTCTTCCTATATGCTTCCACCTCCAGATTGGC
ACATTGATACGTCCGTGTATATTTCCGCACGATTGGATTGATTTCCCAAATG
CCGAGAGATTGAGAGTTGAGATAGGGAGTTCTTGGAGAGCAAATTTTTTCA
ACCTTGCTAAAAAATCTAGATTGCGAGTGGGTTTCGGAAACTCTTGCATATG
CTCTTATATTTCATATCTAACCAAATACTATGTGGTAGTGTAGATTTATATGC
TGTGTACTGAAAGTGACGTCAAGTATAGTAGTGGCAGAGACTCAAAAGATA
CCTGCGGATTCTGAATACCACAACCATAAAAAACAGGATGATGTTATACTT
GTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACAGCCCATAAT
ACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCCACCT
TGTCTAGTTGGACAAGTTCTCTCTTTCTTGGCTGACTCCAACTTACTACACCG
CAACTTCTTGTGCCCTTGTT [SEQ ID NO: 81], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in

between) sequence identity with the sequence set forth in SEQ ID NO: 81 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 81 or a complement thereof, under at least medium or high stringency conditions; and

ACTAGTGATTACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGTC
TCAATTGAATAAAAGATGTGGCCTATAGAAAATTAGGACTGTGCTCTAAAA
AGCATACCAAAACTTGTGGGTCCAGAAACTCAATTTTCCTAATCGAATTTCG
CAAATCTATTTTTTGCGAACCCAGCTATGCTGGGATGAAATAGATCTAAAAT
TTTCTGGCGTTCTCCGTAGATATAAAATTTTCCCCGTTTCAAGTTCGGATGCA
AAAGTTATGACTGTTTTAAGGAAGCTGCTCGAATCAGGGTTTTAGTGAACAC
AAGATGCAAACCGATGGTGATACGGAATAGCGGCGGGTGGAGGGATCAAC
ACAAATTAGAAAAGAACACAATCTAAACCAGCAACAAGACTTTGACCTAGG
ACACAAACTCAACAAAGCGAACTCTGAAACTGAAATATGCAAAGGCTATGG
CGCGGAAGGTTCTAGGACAGGAAAAACAAGTATGGCACTGGACTATGGGA
CGAATGCGAAACACTCAAACGAGAGAATAAAATGTGAACCTAACGGTATAC
CTTAGCTCTGATTACCACTTTATGTAGACGGGGGTCCCGATTTTCTGTCAGG
TTAAGATAAACAACGATTTGTTCGAAGAGCGACACACCGATCCAGCTTCAG
ATCATAAAGACCCAAATTCTGCAACTTTAGCACCACGTCTCCTCTGGTTATC
AACCGTGTTACAATCCGGTCGACCTCGCCAAGAAGGCTAATCCCTGCTGCG
AATCGAAGAACACAAGCAAGAACAAGATAAAACGCAACTCAATTGAAGTG
ACAATTACAGATGAATGATTAAGCACTCAAAGTTGGGGTCTTGCAAACCGA
TAACGGCGACTGTTCAATAACATATTGATCTAAAACAAAACCCCAACCCTA
AAGTAGGTGGTGGCTACTGATTATATAGCCTAAGGGACGTCCAAGGATCCC
TCTTCACGTCCTAAATGTGTCCCAACACGTTTCAAGGCCCAACGGCCCAAAA
GAAGATGTCGCAGCATCCTGACAGATTATGGACGCTGACTTGTTTCGACGAT
TCCCGTTGGTTCTGAAGAGATTTTGATGTCAAACCACTTGGATTGGTTTCCTT
ATCAAATTATCTTTCCATCCATATGTGGATCATTGAAAACGGAGTTCGGATG
CGTCCTGGGCATCCGTTTTATTACAGACTGGTCCTGATGGCCGAGGCAGACT
CGAACTCGAATTAGACTGGGCCTCTGGCTTGGCTGGGACGTCCTTGCTGGCC
TTCTAAGGTGGTGGCCAAGGAGGTGGACGCCCAAGGTCATCTCTTAGGTGT
TCTCCATCTTCTTGGGGCGTGCTCCAACTTGGACCTGGGTCCCAAGGATATC
TAACAAGCCTATGACAACAGTGTAGCTCTTGCCAAAAATAGCGATAGAATA
TATTAGTGATTTGGAGGCCTTGCCTACGTGATATTGAGATGGGAACAGATCT

ATTTGAAAGCTTATCAAACAAGCTTTCCATCAAGTGCTCATTGACCTCGATC
GCACTCCGGATGGATGAGTTATGGCCTTCCCAATGAGGCACTGTCGGGCTG
CCGACGAACCGAAAATTCAACTTTGATGAACTTCCATTATGAAACACATTTG
AACCTACAATCAAATAGTGACATGTACATGTGGAACCAAGTGAATTATAAC
CAAAGTCACTATACAAATGTAGGAACGAGCACACCTTATAATCATTGTTCAT
ATCCGAAGGTGTCATGTCCTCATCAAGTATAGCAGTGGCAGAGACTCAAAA
GATACCTGCGGATTCTGAATACCACAACCTTAAAAAACAGGATGATGTTAT
ACTTGTCCCCTTCCATGATACAGGACTGTTTAGTAATTTCCCAAACTGCCCA
TAATACATTCTGCACCCTTTATTAAACCTCTACTAGCTACAACATCTTACTCC
ATCTTGTCTAGTTGGACAGAAGTTCTCTCTTTCTTGGCTGACTCCAACTTACT
ACACCGCAACTTCTTGTGCCCTTGTT [SEQ ID NO: 82], or a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in SEQ ID NO: 82 or a complement thereof, or a nucleotide sequence that hybridizes to the sequence set forth in SEQ ID NO: 82 or a complement thereof, under at least medium or high stringency conditions.

70. An isolated nucleic acid molecule, which comprise a promoter sequence that is operable in plant cells, wherein the promoter sequence comprises a nucleotide sequence selected from the group consisting of:

(i) a nucleotide sequence set forth in any one of SEQ ID NO: 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81 and 82 or a complement thereof;

(ii) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence set forth in any one of SEQ ID NO: 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81 and 82 or a complement thereof; and

(iii) a nucleotide sequence that hybridizes to the sequence set forth in any one of SEQ ID NO: 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81 and 82 or a complement thereof, under at least medium or high stringency conditions.

71. An isolated nucleic acid molecule comprising a chimeric promoter sequence that comprises an heterologous core promoter sequence in operable connection with a nucleotide sequence selected from the group consisting of :

(1) a nucleotide sequence corresponding to at least one segment selected from segments Φ, Ψ, Z, Δ and Θ of SEQ ID NO: 63-82, or

(2) a nucleotide sequence that shares at least 80% (and at least 81% to at least 99% and all integer percentages in between) sequence identity with the sequence of the at least one segment or a complement thereof, or

(3) a nucleotide sequence that hybridizes to the sequence s of the at least one segment or a complement thereof, under at least medium or high stringency conditions.

72. An isolated nucleic acid molecule according to any one of claims 1 to 71, comprising plant transcriptional regulatory elements able to confer expression preferentially in a plant sink tissue.

73. A chimeric nucleic acid construct comprising an isolated promoter according to any one of claims 1 to 72 in operable connection with a heterologous nucleic acid sequence to be transcribed.

74. A chimeric construct according to claim 73, further comprising a 3' non-translated sequence that is operably connected to the heterologous nucleic acid sequence and that functions in plant cells to terminate transcription and/or to cause addition of a polyadenylated nucleotide sequence to the 3' end of a transcribed RNA sequence.

75. A method for expression of a heterologous nucleic acid sequence, the method comprising introducing into a plant cell a chimeric construct according to claim 73 or claim 74.

76. A method for producing a transformed plant cell, the method comprising introducing into a plant cell a chimeric construct according to claim 73 or claim 74.

77. A method for producing transformed plant cells, the method comprising introducing into regenerable plant cells a chimeric construct according to claim 73 or claim 74 so as to yield transformed plant cells and identifying or selecting the transformed plant cells.

78. A method for selecting stable genetic transformants from transformed plant cells, the method comprising introducing into regenerable plant cells a chimeric construct according to claim 73 or claim 74 so as to yield transformed plant cells and identifying or selecting a transformed plant cell line from the transformed plant cells.

79. A method for producing a differentiated transgenic plant, the method comprising introducing a chimeric construct according to claim 73 or claim 74 into regenerable plant cells so as to yield regenerable transformed cells, identifying or

selecting a population of transformed cells, and regenerating a differentiated transgenic plant from the population.

80. A method according to claim 79, wherein expression of the chimeric construct renders the differentiated transgenic plant identifiable over the corresponding non-transgenic plant.

81. A differentiated transgenic plant comprising plant cells containing a chimeric construct according to claim 73 or claim 74.

82. Cells, tissues, leaves, fruit, flowers, seeds or other plant reproductive material, material used for vegetative propagation, progeny plants including F1 hybrids, male-sterile plants and all other plants and plant products derived from the differentiated transgenic plant of claim 81.

83. An isolated nucleic acid molecule comprising a promoter sequence or biologically active fragment thereof or variant of these, wherein the promoter sequence is located upstream of a transcribable nucleic acid sequence that hybridizes to a nucleic acid probe derived from the polynucleotide sequence set forth in SEQ ID NO: 83.

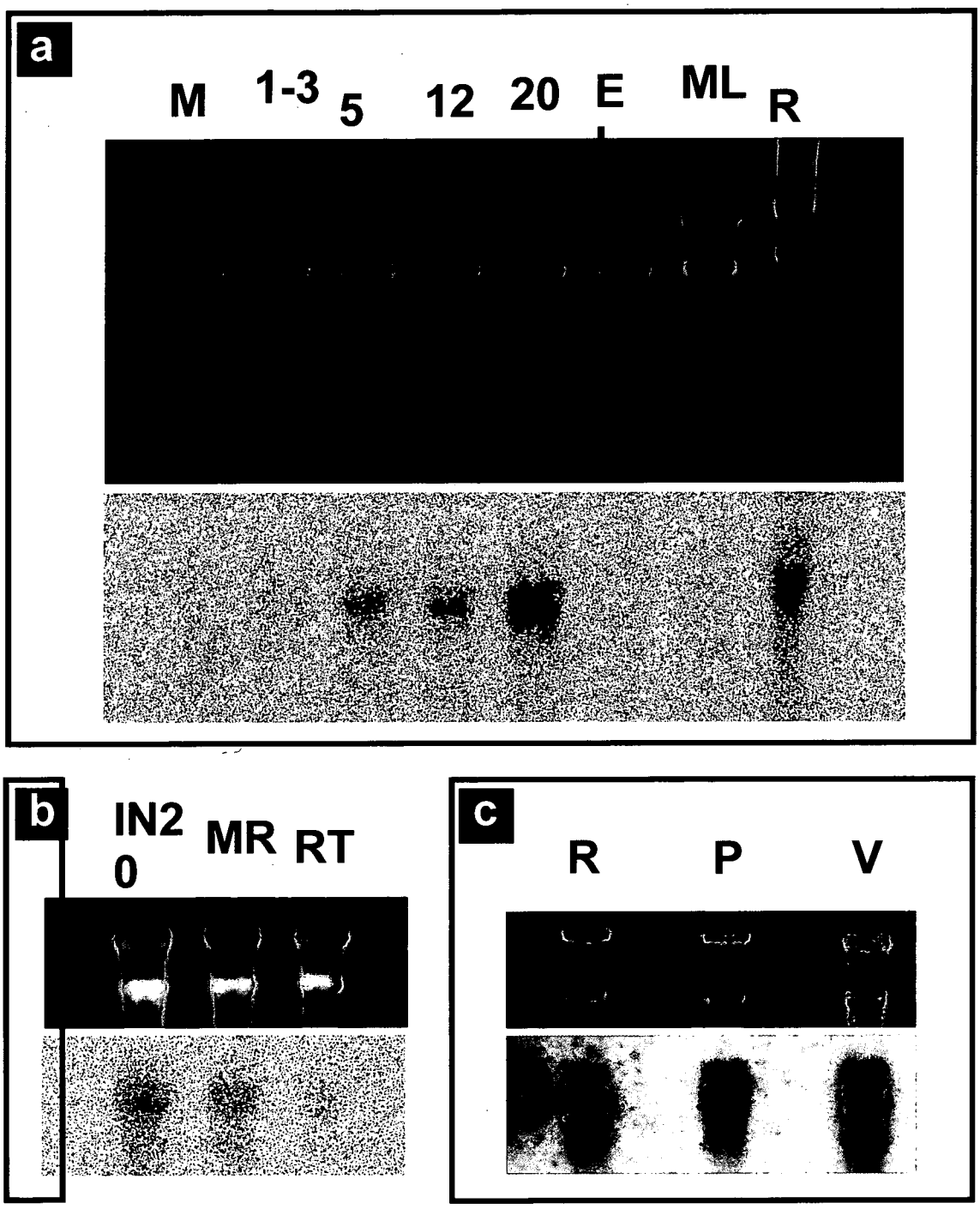


FIGURE 1

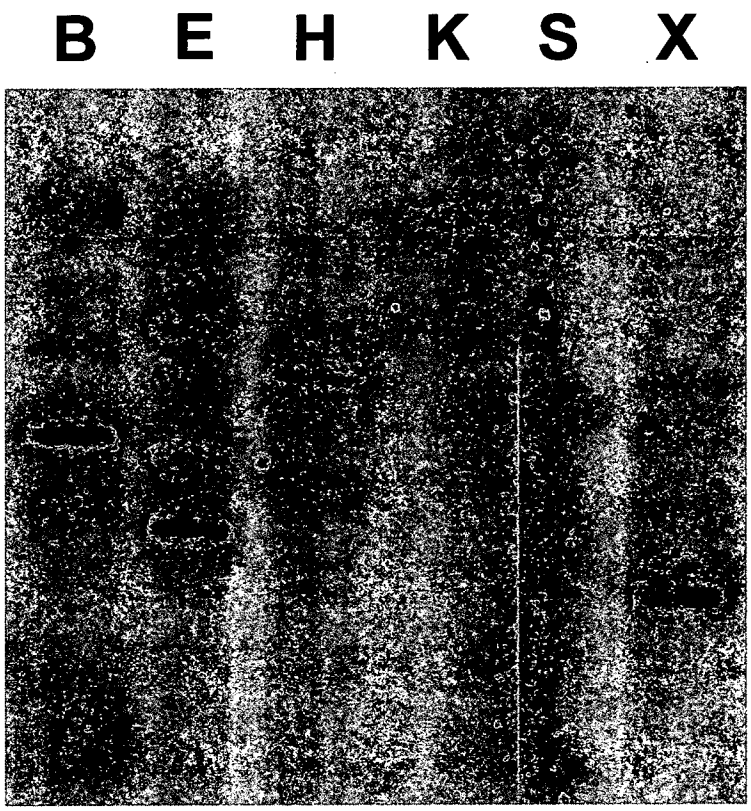


FIGURE 2

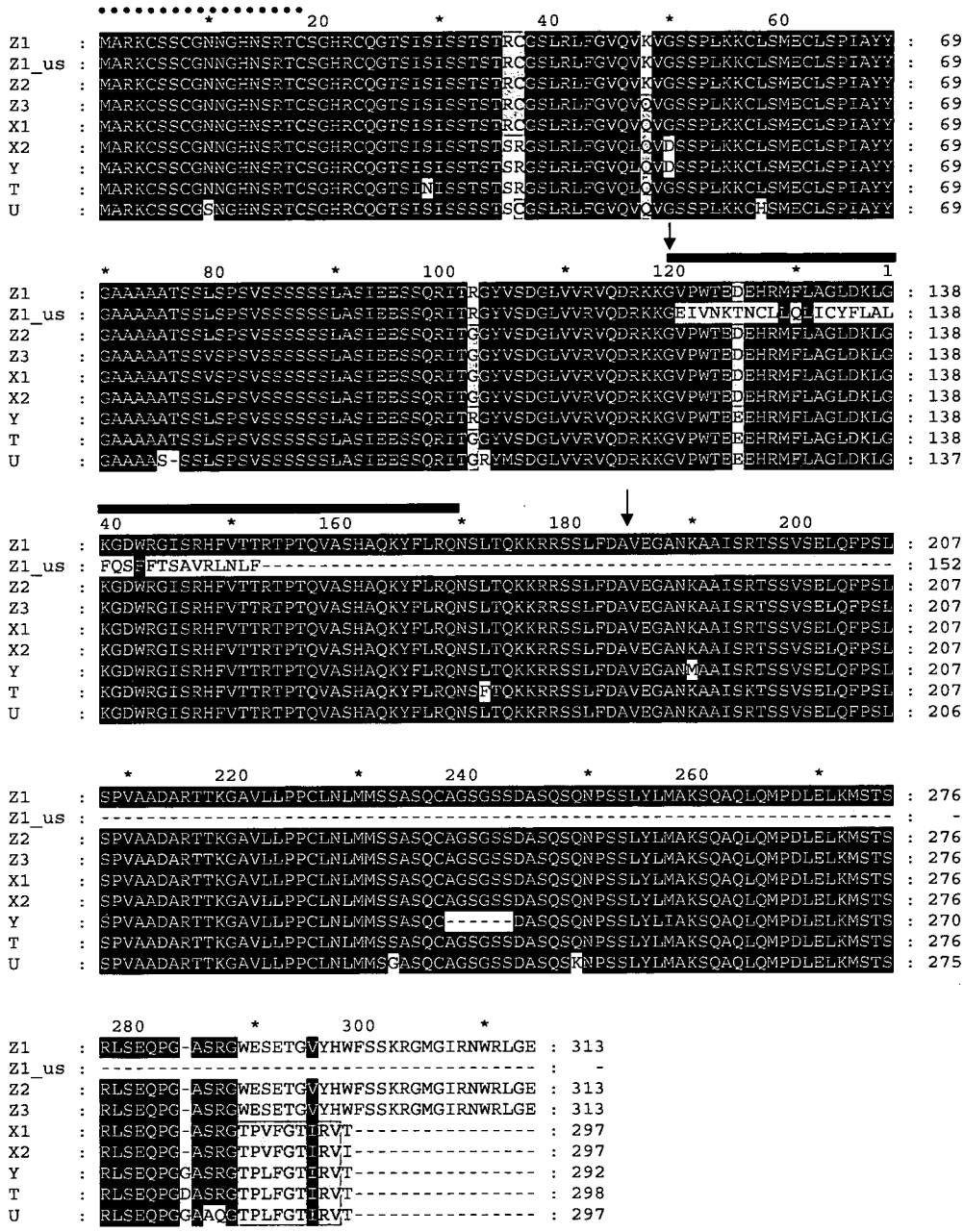
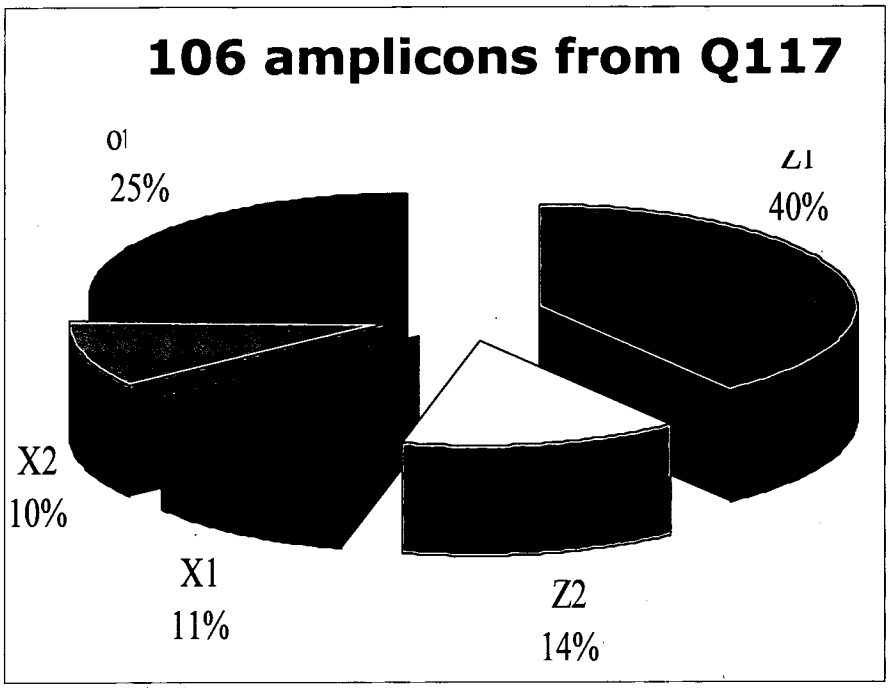

Z1 : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTRCGSLRLFGVQVKVGSSPLKKCLSMECLSPIAYY : 69
Z1_us : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTRCGSLRLFGVQVKVGSSPLKKCLSMECLSPIAYY : 69
Z2 : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTRCGSLRLFGVQVKVGSSPLKKCLSMECLSPIAYY : 69
Z3 : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTRCGSLRLFGVQVQVGSSPLKKCLSMECLSPIAYY : 69
X1 : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTRCGSLRLFGVQVQVGSSPLKKCLSMECLSPIAYY : 69
X2 : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTSRGSLRLFGVQLQVDSSPLKKCLSMECLSPIAYY : 69
Y : MARKCSSCGNNGHNSRTCSGHRCQGTSISISSTSTSRGSLRLFGVQLQVDSSPLKKCLSMECLSPIAYY : 69
T : MARKCSSCGNNGHNSRTCSGHRCQGTSINISSTSTSRGSLRLFGVQLQVGSSPLKKCLSMECLSPIAYY : 69
U : MARKCSSCGSNGHNSRTCSGHRCQGTSISISSSSTSCGSLRLFGVQVQVGSSPLKKCHSMECLSPIAYY : 69
Z1 : GAAAAATSSLSPSVSSSSSSLASIEESSQRITRGYVSDGLVVRVQDRKKGVPWTEDEHRMFLAGLDKLG : 138
Z1_us : GAAAAATSSLSPSVSSSSSSLASIEESSQRITRGYVSDGLVVRVQDRKKGEIVNKTNCLLQLICYFLAL : 138
Z2 : GAAAAATSSLSPSVSSSSSSLASIEESSQRITGGYVSDGLVVRVQDRKKGVPWTEDEHRMFLAGLDKLG : 138
Z3 : GAAAAATSSVSPSVSSSSSSLASIEESSQRITGGYVSDGLVVRVQDRKKGVPWTEDEHRMFLAGLDKLG : 138
X1 : GAAAAATSSVSPSVSSSSSSLASIEESSQRITGGYVSDGLVVRVQDRKKGVPWTEDEHRMFLAGLDKLG : 138
X2 : GAAAAATSSLSPSVSSSSSSLASIEESSQRITGGYVSDGLVVRVQDRKKGVPWTEDEHRMFLAGLDKLG : 138
Y : GAAAAATSSLSPSVSSSSSSLASIEESSQRITRGYVSDGLVVRVQDRKKGVPWTEEEHRMFLAGLDKLG : 138
T : GAAAAATSSLSPSVSSSSSSLASIEESSQRITGGYVSDGLVVRVQDRKKGVPWTEEEHRMFLAGLDKLG : 138
U : GAAAAS-SSLSPSVSSSSSSLASIEESSQRITGRYMSDGLVVRVQDRKKGVPWTEEEHRMFLAGLDKLG : 137
Z1 : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANKAAISRTSSVSELQFPSL : 207
Z1_us : FQSFFTSAVRLNLF---------- : 152
Z2 : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANKAAISRTSSVSELQFPSL : 207
Z3 : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANKAAISRTSSVSELQFPSL : 207
X1 : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANKAAISRTSSVSELQFPSL : 207
X2 : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANKAAISRTSSVSELQFPSL : 207
Y : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANMAAISRTSSVSELQFPSL : 207
T : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSFTQKKRRSSLFDAVEGANKAAISKTSSVSELQFPSL : 207
U : KGDWRGISRHFVTTRTPTQVASHAQKYFLRQNSLTQKKRRSSLFDAVEGANKAAISRTSSVSELQFPSL : 206
Z1 : SPVAADARTTKGAVLLPPCLNLMMSSASQCAGSGSSDASQSQNPSSLYLMAKSQAQLQMPDLELKMSTS : 276
Z1_us : ---------- : -
Z2 : SPVAADARTTKGAVLLPPCLNLMMSSASQCAGSGSSDASQSQNPSSLYLMAKSQAQLQMPDLELKMSTS : 276
Z3 : SPVAADARTTKGAVLLPPCLNLMMSSASQCAGSGSSDASQSQNPSSLYLMAKSQAQLQMPDLELKMSTS : 276
X1 : SPVAADARTTKGAVLLPPCLNLMMSSASQCAGSGSSDASQSQNPSSLYLMAKSQAQLQMPDLELKMSTS : 276
X2 : SPVAADARTTKGAVLLPPCLNLMMSSASQCAGSGSSDASQSQNPSSLYLMAKSQAQLQMPDLELKMSTS : 276
Y : SPVAADARTTKGAVLLPPCLNLMMSSASQC------DASQSQNPSSLYLIAKSQAQLQMPDLELKMSTS : 270
T : SPVAADARTTKGAVLLPPCLNLMMSSASQCAGSGSSDASQSQNPSSLYLMAKSQAQLQMPDLELKMSTS : 276
U : SPVAADARTTKGAVLLPPCLNLMMSGASQCAGSGSSDASQSKNPSSLYLMAKSQAQLQMPDLELKMSTS : 275
Z1 : RLSEQPG-ASRGWESETGVYHWFSSKRGMGIRNWRLGE : 313
Z1_us : ---------- : -
Z2 : RLSEQPG-ASRGWESETGVYHWFSSKRGMGIRNWRLGE : 313
Z3 : RLSEQPG-ASRGWESETGVYHWFSSKRGMGIRNWRLGE : 313
X1 : RLSEQPG-ASRGTPVFGTIRVT---------- : 297
X2 : RLSEQPG-ASRGTPVFGTIRVI---------- : 297
Y : RLSEQPGGASRGTPLFGTIRVT---------- : 292
T : RLSEQPGDASRGTPLFGTIRVT---------- : 298
U : RLSEQPGGAAQGTPLFGTIRVT---------- : 297

FIGURE 3


106 amplicons from Q117
O1
25%
Z1
40%
X2
10%
X1
11%
Z2
14%

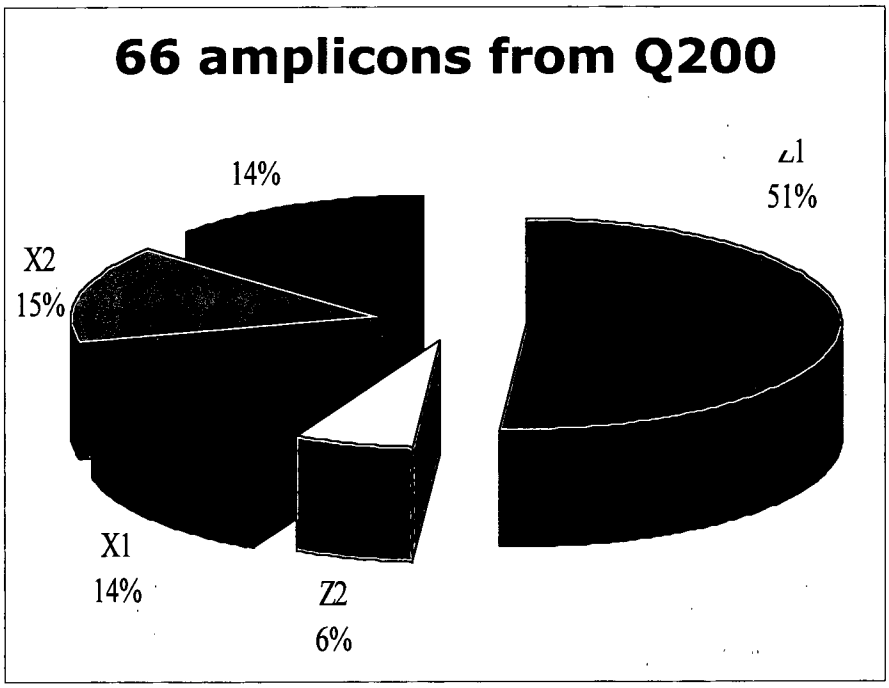

66 amplicons from Q200
14%
Z1
51%
X2
15%
X1
14%
Z2
6%

FIGURE 4

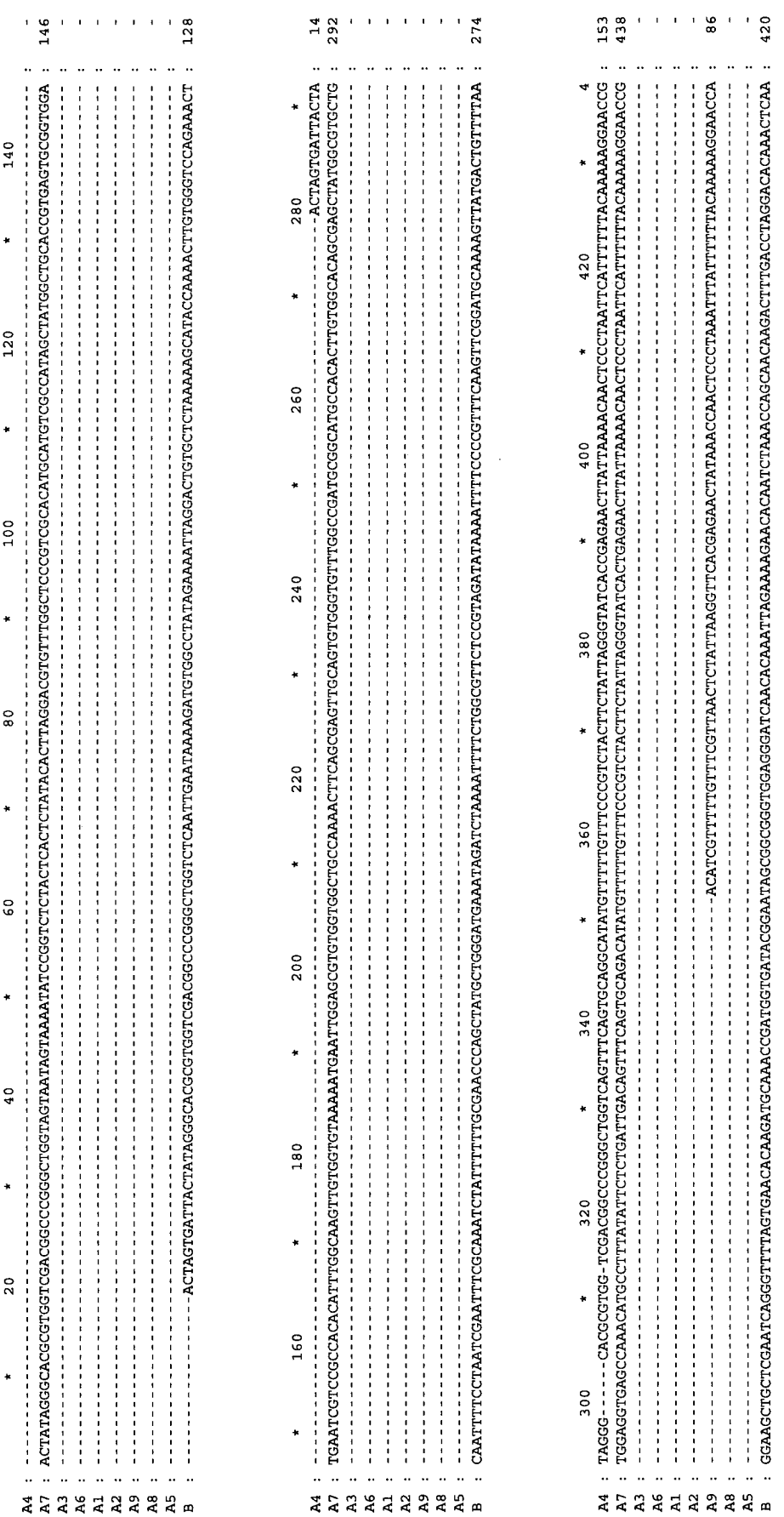


FIGURE 5

FIGURE 5 cont'd

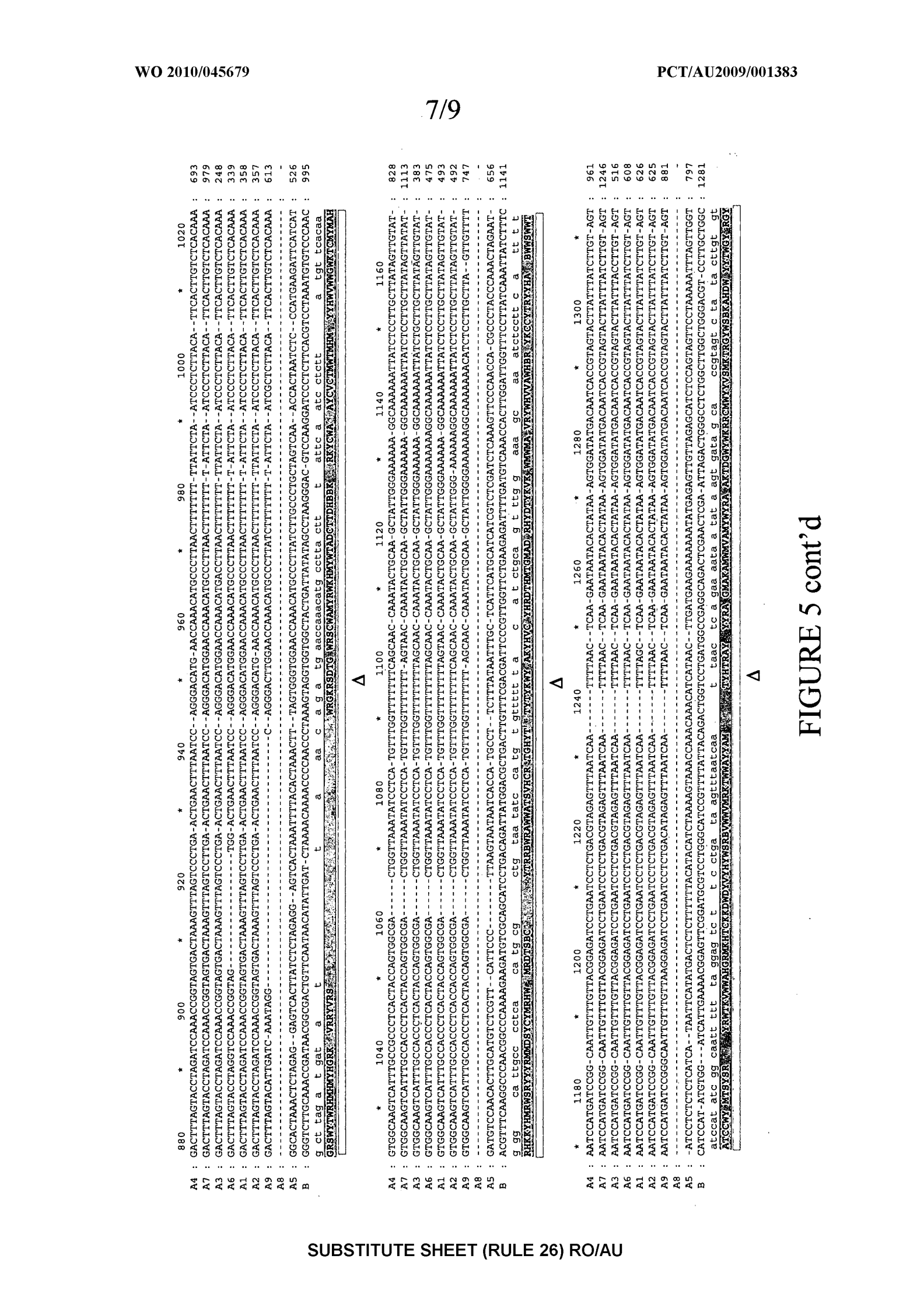


FIGURE 5 cont'd

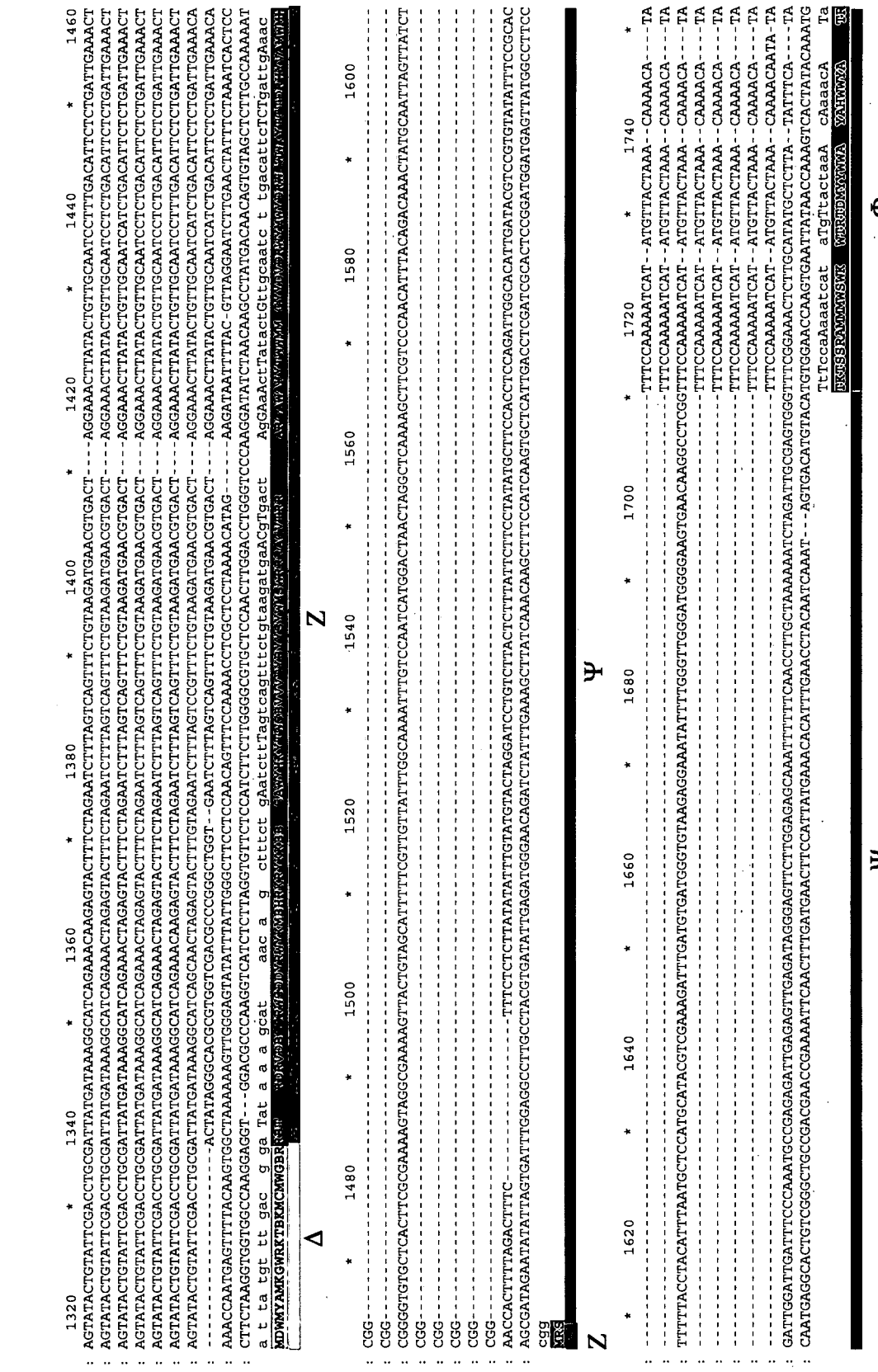


FIGURE 5 cont'd

FIGURE 5 cont'd

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU2009/001383

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl.

C12N 15/11 (2006.01) ***A01H 5/00*** (2006.01) ***C12N 15/82*** (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
MEDLINE, CAPLUS, BIOSIS, AGRICOLA, WPI, EPODOC: promoter, sugarcane, saccharum, stem, sink
GENOMEQUEST SEQ ID NOS: 1-11, 13-22

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MUDGE S.R., et al., "Efficient silencing of reporter transgenes coupled to known functional promoters in sugarcane, a highly polyploid crop species", Planta, 2009. Vol. 229, No. 3, pages 549-558. Published electronically 15 November 2008. Whole document	
A	GENBANK Accession EU719199, Saccharum hybrid cultivar Q117 R1MYB1 protein (R1MYB1) gene, R1MYB1-Z1 allele, 11 February 2009 Whole document	
A	WO 2004/062366 A2 (DAMAJ, M.B. et al) 29 July 2004 Whole document	

☐ Further documents are listed in the continuation of Box C ☒ See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
25 November 2009

Date of mailing of the international search report
02 DEC 2009

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
E-mail address: pct@ipaustralia.gov.au
Facsimile No. +61 2 6283 7999

Authorized officer
SUSAN MACKENZIE
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No : +61 2 6283 2690

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU2009/001383

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report	Patent Family Member					
WO 2004062366	BR	PI0406620	EP	1590447	US	7253276
	US	2005034192	US	2007016979		

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

END OF ANNEX