

POLYNUCLEOTIDE SEQUENCES FROM
RHODOSPORIDIUM AND *RHODOTORULA* AND USE THEREOF

CROSS-REFERNCE TO RELATED APPLICATION

[0001] The present application is related to and claims priority to U.S. Provisional Patent Application Serial No. 61/782,832, filed 14 March 2013. This application is incorporated herein by reference.

SEQUENCE SUBMISSION

[0002] The present application is being filed along with a Sequence Listing in electronic format. The Sequence Listing is entitled 2577230PCTSequencListing.txt, was created on 18 February 2014 and is 39 kb in size. The information in the electronic format of the Sequence Listing is part of the present application and is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

[0003] The present invention relates to the field of fungal biotechnology, more particularly to strong gene expression systems in species in the Pucciniomycotina and Ustilaginomycotina subphyla.

[0004] The publications and other materials used herein to illuminate the background of the invention, and in particular, cases to provide additional details respecting the practice, are incorporated by reference, and for convenience are referenced in the following text by author and date and are listed alphabetically by author in the appended bibliography.

[0005] The Pucciniomycotina is a subphylum of fungi in the phylum of Basidiomycota (Kirk et al., 2008). It holds many species that have important industrial applications. For example, a number of species in the *Rhodosporidium* and *Sporidiobolus* genera, such as *Rhodosporidium toruloides* (also known as *Rhodotorula gracilis*, *Rhodosporidium glutinis*, *Rhodotorula glutinis*, *Torula koishikawensis* and *Torula rubescens*) and *Sporobolomyces salmonicolor*, are oil-rich single-cell yeasts capable of high density fermentation (Hu et al., 2009; Meng et al., 2009). These species hold great potential as a host for the production of long chain hydrocarbons, such as triacylglycerol (TAG, or fat), fatty acid esters (biodiesel), fatty alcohols, alcohols, lactones, terpenoids and vitamins (Wu et al., 2010a; Wu et al., 2010b; Zhao et al., 2010a; Zhao et al., 2010b). In another example, species in *Ustilaginomycotina* subphylum, in particular, *Ustilago* and *Pseudozyma* genera, are known to produce glycolipids, which may function as a surfactant or fungicide (Hewald et al., 2005; Teichmann et al., 2010).

[0006] Promoters that are able to drive strong gene expression, either constitutively or inducibly, are critical for the development of biotechnological applications of a microorganism. WO 2012/169969, incorporated by reference herein in its entirety, describes several polynucleotide sequences derived from the upstream region of *glyceraldehyde phosphate dehydrogenase* gene (*GPD1*), *translation initiation factor* gene (*TEF1*), and putative *stearyl-CoA-delta 9-desaturase* gene (*FAD1*) of selected fungal species that are able to function as a strong promoter of gene expression in *Pucciniomycotina* and *Ustilaginomycotina* subphyla. As repeated use of the identical or highly homologous promoter risks genome instability, epigenetic and genetic modification of chromatin resulted from repeat induced point mutation (RIP) or RNA silencing (Horns et al, 2012), an enlarged promoter pool is highly desirable for *Pucciniomycotina* and *Ustilaginomycotina* subphyla, wherein functionally verified promoters are scarce.

[0007] Promoters are DNA sequences located in the 5' region adjacent to the transcriptional start site. It houses a combination of *cis*-acting DNA elements that act to interact with transcription factors by activating or repressing transcription of RNA polymerase. To date, genome shotgun sequences have been published for *Rhodotorula glutinis* ATCC 204091 (GenBank Accession: GL989638.1), *Rhodospiridium toruloides* MTCC 457 (GenBank Accession: PRJNA112573), *Rhodospiridium toruloides* NP11 (GenBank: ALAU00000000.1) and draft genome sequences have been published for *Rhodotorula graminis* WP1 (http://genome.jgi-psf.org/Rhoba1_1/Rhoba1_1.home.html) and *Sporobolomyces roseus* (<http://genome.jgi-psf.org/Sporol1/Sporol1.home.html>). RNA-Seq, proteomic and genome shotgun data released for *Rhodospiridium toruloides* NP11 (Zhu, Z., et al, 2012) are not able to define the sequence of functional promoters because the activity of a promoter is influenced by several factors, such as the location of 5' and 3' ends, posttranscriptional silencing, influence of intron, etc. The activity of a promoter in a heterologous host species is even more unpredictable.

SUMMARY OF THE INVENTION

[0008] The present invention relates to the field of fungal biotechnology, more particularly to strong gene expression systems in species in the *Pucciniomycotina* and *Ustilaginomycotina* subphyla.

[0009] In a first aspect, the present invention provides polynucleotide sequences that function as strong promoters of gene expression in *Rhodospiridium*, *Rhodotorula*, *Sporobolomyces*, *Pseudozyma* and *Ustilago* genera. These polynucleotide sequences are sometimes referred to herein as polynucleotide promoter sequences. In one embodiment, the polynucleotide promoter

sequences comprises the sequence set forth in any one of SEQ ID NOs:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11. In another embodiment, the polynucleotide promoter sequences comprises the promoter sequence of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11, i.e., the sequence without the cloning sites. Each of the polynucleotide promoter sequences contains at least one GAGGAG sequence motif, which functions to enhance gene expression in said fungal species. Each of polynucleotide sequences is effective in performing strong gene expression in *Rhodospiridium*, *Rhodotorula*, *Sporobolomyces*, *Pseudozyma* and *Ustilago* genera. In addition, operable fragments of these polynucleotide promoter sequences can be isolated using convention promoter screening assays and can be screened for efficient selection of transformed fungal cells using the techniques described herein. In one embodiment, an operable fragment, also termed a promoter portion herein, is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon. As used herein "up to" refers to the length of the promoter portion of the promoters set forth in the disclosed SEQ ID NOs. Thus, "up to" refers to the maximal length of the promoter sequence if less than 1100 nucleotides of the promoters of the disclosed SEQ ID NOs.

[0010] In a second aspect, the present invention provides a DNA construct comprising the polynucleotide promoter sequences described herein, an operably linked polypeptide encoding sequence and an operably linked RNA transcriptional terminator sequence. Any eukaryotic transcriptional terminator, well known to the skilled artisan, may be used. Such a DNA construct allows strong expression of the polypeptide in a fungal species in which the genome is biased in C and G. Of particular relevance are species selected from the Pucciniomycotina and Ustilaginomycotina subphyla. The species of particular relevance are those in the *Rhodospiridium*, *Rhodotorula*, *Sporobolomyces*, *Ustilago* and *Pseudozyma* genera, in which reside a number of species with great potential for the bioconversion of renewable resources into high-value products, such as triglyceride, biodiesel, fatty alcohol, vitamins, lactone, terpenoids and biosurfactants.

BRIEF DESCRIPTION OF THE FIGURES

[0011] Fig. 1 shows cloning and transformation vectors for promoters. Both pPN007 and pRH2031 are based on pPZP200 and contain the *Umgpdl::hpt-3* hygromycin selection marker. *RtGFP* is a codon optimized GFP gene optimized for GC-rich genome.

[0012] Fig. 2 shows the location of GAGGAG motifs. Vertical lines with * indicate GAGGAG motifs in sense orientation the remaining lines are in the antisense orientation.

[0013] Fig. 3 shows the relative florescence of RtGFP driven by *ENO1* and *FAD1* promoters of various lengths, which is defined as the nucleotide number starting from the first nucleotide from the putative translation initiation codon (ATG) up to the point of 5' end, excluding the restriction sites added. The promoters showed similar trend in YNB and YNB nitrogen-free (N⁻) media. 35S-Ω is a basic promoter containing the *cauliflower mosaic virus* 35S gene promoter starting from the TATA box down to the -1 position of 5' UTR and the omega translation enhancer sequence of *Tobacco mosaic virus* (TMV) inserted immediate upstream of RtGFP. Background florescence of non-transformed strain (*R. toruloides* ATCC 10657) has been subtracted in all values.

[0014] Fig. 4A shows details of the 519 bp Rg2*ENO1* promoter sequence. Arrows indicate the positions of 5' ends of various deletions.

[0015] Fig. 4B shows the code and length of each truncated Rg2*ENO1* promoters.

[0016] Fig. 5A and Fig. 5B show the relative florescence of RtGFP driven by *ENO1* promoters of various lengths, which is defined as the nucleotide number starting from the first nucleotide from the putative translation initiation codon (ATG) up to the point of 5' end, excluding the restriction sites added. The promoters showed similar trend in YNB (Fig. 5A) and YNB nitrogen-free (N⁻) (Fig. 5B) media. Rg1 Wt refers to the background florescence of *R. glutinis* ATCC 90781.

DETAILED DESCRIPTION OF THE INVENTION

[0017] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which the invention belongs.

[0018] The term "operably linked" or "operatively linked" is defined herein as a configuration in which a regulatory or control sequence is appropriately placed at a position relative to the nucleotide sequence of the nucleic acid construct such that the control sequence directs the expression of a polynucleotide of the present invention. Regulatory or control sequences may be positioned on the 5' side of the nucleotide sequence or on the 3' side of the nucleotide sequence as is well known in the art.

[0019] The term "strong expression" as used herein means expression of a marker protein or mRNA to a detectable level using detection methods known, for example, florescence for GFP, activity assay for GUS and lacZ genes.

[0020] The present invention relates to the field of fungal biotechnology, more particularly to strong gene expression systems in species in the Pucciniomycotina and Ustilaginomycotina subphyla.

[0021] In a first aspect, the present invention provides polynucleotide sequences that function as strong promoters of gene expression in *Rhodospiridium*, *Rhodotorula*, *Sporobolomyces*, *Pseudozyma* and *Ustilago* genera. These polynucleotide sequences are sometimes referred to herein as polynucleotide promoter sequences. In one embodiment, the polynucleotide promoter sequences comprises the sequence set forth in any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11. In another embodiment, the polynucleotide promoter sequences comprises the promoter sequence of any one of SEQ ID NOs: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11, i.e., the sequence without the cloning sites. Each of the polynucleotide promoter sequences contains at least one GAGGAG sequence motif, which functions to enhance gene expression in said fungal species. Each of polynucleotide sequences is effective in performing strong gene expression in *Rhodospiridium*, *Rhodotorula*, *Sporobolomyces*, *Pseudozyma* and *Ustilago* genera. In addition, operable fragments of these polynucleotide promoter sequences can be isolated using convention promoter screening assays and can be screened for efficient selection of transformed fungal cells using the techniques described herein. . In one embodiment, an operable fragment, also termed a promoter portion herein, is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon. As used herein "up to" refers to the length of the promoter portion of the promoters set forth in the disclosed SEQ ID NOs. Thus, "up to" refers to the maximal length of the promoter sequence if less than 1100 nucleotides of the promoters of the disclosed SEQ ID NOs.

[0022] In a second aspect, the present invention provides a DNA construct comprising the polynucleotide promoter sequences described herein, an operably linked polypeptide encoding sequence and an operably linked RNA transcriptional terminator sequence. Any eukaryotic transcriptional terminator, well known to the skilled artisan, may be used. Such a DNA construct allows strong expression of the polypeptide in a fungal species in which the genome is biased in C and G. Of particular relevance are species selected from the Pucciniomycotina and Ustilaginomycotina subphyla. The species of particular relevance are those in the *Rhodospiridium*, *Rhodotorula*, *Sporobolomyces*, *Ustilago* and *Pseudozyma* genera, in which reside a number of species with great potential for the bioconversion of renewable resources into high-value products, such as triglyceride, biodiesel, fatty alcohol, vitamins, lactone, terpenoids and biosurfactants.

[0023] Nucleic acid hybridization, a technique well known to those of skill in the art of DNA manipulation, can be used to identify other suitable polynucleotides. In accordance with the invention other suitable promoters for use may be obtained by the identification of polynucleotides that selectively hybridize to the promoters described above by hybridization under low stringency

conditions, moderate stringency conditions, or high stringency conditions. Selectively hybridizing sequences typically have at least 50% sequence identity, preferably at least 70%, 80% or 90% sequence identity, and most preferably 95%, 98% or 99% sequence identity with each other.

[0024] Database searches and homology searches of genome and nucleotide databases identify similar DNA or RNA molecules based on the alignment of nucleotides using algorithms or computer programs and these techniques well known to those of skill in the art. In accordance with the invention other suitable polynucleotides for use may be obtained by the *in silico* identification of polynucleotides for regulatory sequences with at least 50% sequence identity, preferably at least 70%, 80% or 90% sequence identity, and most preferably 95%, 98% or 99% sequence identity with each other.

[0025] The invention provides a polynucleotide promoter sequence selected from SEQ ID NOs:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11, or the promoter sequence thereof, i.e., the sequence without the cloning sites. In one embodiment, a polynucleotide promoter sequence is provided which has at least 60% identity with any one of these polynucleotide promoter sequences. In another embodiment, a polynucleotide promoter sequence is provided which has at least 70% identity with any one of these polynucleotide promoter sequences. In an additional embodiment, a polynucleotide promoter sequence is provided which has at least 80% identity with any one of these polynucleotide promoter sequences. In a further embodiment, a polynucleotide promoter sequence is provided which has at least 90% identity with any one of these polynucleotide promoter sequences. In another embodiment, a polynucleotide promoter sequence is provided which has at least 95% identity with any one of these polynucleotide promoter sequences. In another embodiment, a polynucleotide promoter sequence is provided which has at least 98% identity with any one of these polynucleotide promoter sequences. . In one embodiment, a promoter sequence herein, is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon. As used herein "up to" refers to the length of the promoter portion of the promoters set forth in the disclosed SEQ ID NOs. Thus, "up to" refers to the maximal length of the promoter sequence if less than 1100 nucleotides of the promoters of the disclosed SEQ ID NOs.

[0026] The invention provides a polynucleotide construct comprising an isolated promoter described herein, such as one selected from SEQ ID NOs:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11, or promoter portion thereof, operatively linked to a polypeptide-encoding sequence which is operatively linked to a transcriptional terminator. . In one embodiment, an operable fragment,

also termed a promoter portion herein, is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon. As used herein “up to” refers to the length of the promoter portion of the promoters set forth in the disclosed SEQ ID NOs. Thus, “up to” refers to the maximal length of the promoter sequence if less than 1100 nucleotides of the promoters of the disclosed SEQ ID NOs. In one embodiment, the polynucleotide construct enables efficient expression of a polypeptide in a fungal species selected from *Pucciniomycotina* and *Ustilaginomycotina* subphyla. The fungal species is preferably one selected from *Rhodosporidium*, *Rhodoturula*, *Ustilago*, *Pseudozyma*, or *Sporobolomyces* genus, the genome of which contains at least 50% C and G, preferably more than 60% C and G.

[0027] In one embodiment, the polynucleotide construct is inserted in a T-DNA vector, a shuttle vector, or in a fungal chromosome, wherein the polypeptide-encoding sequence contains at least 50% CG, preferably 60% CG and most preferably more than 80% CG.

[0028] In another embodiment, the polynucleotide promoter sequence contains at least one GAGGAG sequence, or its reverse complement sequence thereof. In another embodiment, the isolated promoter is operatively linked to a gene encoding antibiotic resistance enzyme, herbicide resistance enzyme, GFP, GUS, lacZ, terpene synthase, fatty acid desaturase, P450 cytochrome oxidase, glucanase, xylanase, mannanase, mannosidase, glucosidase, glucomannanase, xyluglucanase, hydroxymethylglutaryl-CoA synthase, hydroxymethylglutaryl-CoA reductase, acetyl-CoA C-acetyltransferase, mevalonate kinase, phosphomevalonate kinase, diphosphomevalonate decarboxylase, isopentenyl-diphosphate delta-isomerase, farnesyl diphosphate synthase, geranylgeranyl diphosphate synthase, methyl transferase, or glucosyl transferase, beta-carotenoid hydroxylase, beta-carotenoid oxidase.

[0029] In one embodiment, any transcriptional terminator operable in species of the fungi can be used. Terminators are typically located downstream (3') of the gene, after the stop codon (TGA, TAG or TAA). Terminators play an important role in the processing and stability of RNA as well as in translation. Most, but not all terminators, contain a polyadenylation sequence or cleavage site. Examples of specific polyadenylation sequences are AAUAAA or AAUAAU. These sequences are known as the near upstream elements (NUEs) (Nagaya et al., 2010). NUEs usually reside approximately 30 bp away from a GU-rich region (Mogen et al., 1990; Mogen et al., 1992; Rothnie et al. 1994), known as far upstream elements (FUEs). The FUEs enhance processing at the polyadenylation sequence or cleavage site, which is usually a CA or UA in a U-rich region (Bassett, 2007). Within the terminator, elements exist that increase the stability of the transcribed

RNA (Ohme-Takagi et al., 1993; Newman et al., 1993; Gutiérrez et al., 1999) and may also control gene expression (Ingelbrecht, 1989; An et al., 1989).

[0030] Nucleic acid hybridization, a technique well known to those of skill in the art of DNA manipulation, can be used to identify other suitable terminators. In accordance with the invention other suitable promoters for use may be obtained by the identification of terminators that selectively hybridize to the promoters described above by hybridization under low stringency conditions, moderate stringency conditions, or high stringency conditions. Selectively hybridizing sequences typically have at least 50% sequence identity, preferably at least 70%, 80% or 90% sequence identity, and most preferably 95%, 98% or 99% sequence identity with each other.

[0031] Database searches and homology searches of genome and nucleotide databases identify similar DNA or RNA molecules based on the alignment of nucleotides using algorithms or computer programs and these techniques well known to those of skill in the art. In accordance with the invention other suitable terminators for use may be obtained by the *in silico* identification of terminators for regulatory sequences with at least 50% sequence identity, preferably at least 70%, 80% or 90% sequence identity, and most preferably 95%, 98% or 99% sequence identity with each other.

[0032] A DNA of interest can be added to the polynucleotide construct. The DNA of interest is operatively linked to promoter and a terminator. Any promoter and terminator operable in species of the Pucciniomycotina and Ustilaginomycotina subphyla can be used. In some embodiments, the DNA of interest may be used to insert or modify metabolic pathways, such as fatty acid biosynthesis, lipid biosynthesis, triglyceride biosynthesis, and the like. The DNA of interest may be inserted into the genome of the fungal cells to enhance the bioconversion of renewable resources into high-value products, such as triglycerides, biodiesel, fatty alcohol, vitamins, biosurfactants, lactone, terpenoid and the like.

[0033] A polynucleotide construct may be introduced directly into the genomic DNA of the fungal cell using techniques such as electroporation and microinjection of plant cell protoplasts, or the polynucleotide constructs can be introduced directly to fungal tissue using ballistic methods, such as DNA particle bombardment. Alternatively, the polynucleotide constructs may be combined with suitable T-DNA flanking regions and introduced into a conventional *Agrobacterium tumefaciens* host vector. The virulence functions of the *Agrobacterium tumefaciens* host will direct the insertion of the construct into the fungal cell DNA when the cell is infected by the bacteria. Thus, any method, which provides for effective transformation/transfection may be employed. See, for example, U.S. Patent Nos. 7,241,937,

7,273,966 and 7,291,765 and U.S. Patent Application Publication Nos. 2007/0231905 and 2008/0010704 and references cited therein. See also, International Published Application Nos. WO 2005/103271 and WO 2008/094127 and references cited therein.

[0034] The transformed fungi are transferred to standard growing media (e.g., solid or liquid nutrient media, grain, vermiculite, compost, peat, wood, wood sawdust, straw, etc.) and grown or cultivated in a manner known to the skilled artisan.

[0035] After the polynucleotide is stably incorporated into transformed fungi, it can be transferred to other fungi by sexual crossing. Any of a number of standard breeding techniques can be used, depending upon the species to be crossed.

[0036] It may be useful to generate a number of individual transformed fungi with any recombinant construct in order to recover fungi free from any positional effects. It may also be preferable to select fungi that contain more than one copy of the introduced polynucleotide construct such that high levels of expression of the recombinant molecule are obtained.

[0037] It may be desirable to produce fungal lines that are homozygous for a particular gene if possible in the particular species. In some species this is accomplished by the use monosporous cultures. By using these techniques, it is possible to produce a haploid line that carries the inserted gene and then to double the chromosome number either spontaneously or by the use of colchicine. This gives rise to a fungus that is homozygous for the inserted gene, which can be easily assayed for if the inserted gene carries with it a suitable selection marker gene for detection of fungi carrying that gene. Alternatively, fungi may be self-fertilized, leading to the production of a mixture of spores that consists of, in the simplest case, three types, homozygous (25%), heterozygous (50%) and null (25%) for the inserted gene. Although it is relatively easy to score null fungi from those that contain the gene, it is possible in practice to score the homozygous from heterozygous fungi by Southern blot analysis in which careful attention is paid to the loading of exactly equivalent amounts of DNA from the mixed population, and scoring heterozygotes by the intensity of the signal from a probe specific for the inserted gene. It is advisable to verify the results of the Southern blot analysis by allowing each independent transformant to self-fertilize, since additional evidence for homozygosity can be obtained by the simple fact that if the fungi was homozygous for the inserted gene, all of the subsequent fungal lines from the selfed individual will contain the gene, while if the fungus was heterozygous for the gene, the generation grown from the selfed seed will contain null fungal lines. Therefore, with simple selfing one can select homozygous fungal lines that can also be confirmed by Southern blot analysis.

[0038] Creation of homozygous parental lines makes possible the production of hybrid fungus and spores that will contain a modified protein component. Transgenic homozygous parental lines are maintained with each parent containing either the first or second recombinant DNA sequence operably linked to a promoter. Also incorporated in this scheme are the advantages of growing a hybrid crop, including the combining of more valuable traits and hybrid vigor.

[0039] The practice of the present invention employs, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA, genetics, immunology, cell biology, cell culture and transgenic biology, which are within the skill of the art. See, e.g., Maniatis *et al.*, 1982, *Molecular Cloning* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York); Sambrook *et al.*, 1989, *Molecular Cloning*, 2nd Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York); Sambrook and Russell, 2001, *Molecular Cloning*, 3rd Ed. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York); Ausubel *et al.*, 1992), *Current Protocols in Molecular Biology* (John Wiley & Sons, including periodic updates); Glover, 1985, *DNA Cloning* (IRL Press, Oxford); Russell, 1984, *Molecular biology of plants: a laboratory course manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.); Anand, *Techniques for the Analysis of Complex Genomes*, (Academic Press, New York, 1992); Guthrie and Fink, *Guide to Yeast Genetics and Molecular Biology* (Academic Press, New York, 1991); Harlow and Lane, 1988, *Antibodies*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York); *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Methods In Enzymology*, Vols. 154 and 155 (Wu *et al.* eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology*, Volumes I-IV (D. M. Weir and C. C. Blackwell, eds., 1986); Riott, *Essential Immunology*, 6th Edition, Blackwell Scientific Publications, Oxford, 1988; Fire *et al.*, *RNA Interference Technology: From Basic Science to Drug Development*, Cambridge University Press, Cambridge, 2005; Schepers, *RNA Interference in Practice*, Wiley-VCH, 2005; Engelke, *RNA Interference (RNAi): The Nuts & Bolts of siRNA Technology*, DNA Press, 2003; Gott, *RNA Interference, Editing, and Modification: Methods and Protocols (Methods in Molecular Biology)*, Human Press, Totowa, NJ, 2004; Sohail, *Gene Silencing by RNA Interference: Technology and Application*, CRC, 2004.

EXAMPLES

[0040] The present invention is described by reference to the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner. Standard techniques well known in the art or the techniques specifically described below were utilized.

EXAMPLE 1

Culture of Microbial Strains and Basic Molecular Methods

[0041] *Rhodospiridium toruloides* strain ATCC10657 (referred to as Rt1), and *Rhodotorula glutinis* strain ATCC204091 (referred to as Rg2), *Pseudozyma aphidis* ATCC32657, were sourced from the American Type Culture Collections (ATCC). *Sporobolomyces roseus* FGSC 10293 (IAM13481) and *Rhodotorula graminis* WP1 (FGSC WP1) (referred to as Rg3) were obtained from Fungal Genetics Stock Center (University of Missouri, USA). *Ustilago maydis* strain and *Agrobacterium tumefaciens* strain AGL-1 have been described (Ji et al., 2010; Lazo et al., 1991). *Escherichia coli* strain XL1-Blue was used for routine plasmid manipulation and amplification. Fungal strains were cultured at 28°C in YPD broth (1% yeast extract, 2% peptone, 2% glucose) or on solid potato-dextrose agar (PDA). *A. tumefaciens* was cultured at 28°C in either liquid or solid 2YT medium (1.6% tryptone, 1% yeast extract, 0.5% NaCl). *E. coli* was grown in LB broth or on solid LB agar.

[0042] Genomic DNA was extracted based on the method described for *U. maydis* (Ji et al., 2010) with some modifications. Briefly, the cell culture at exponential phase was collected and washed with 1 M sorbitol. The cells were resuspended in 0.1 ml of SCS buffer (1 M sorbitol, 20 mM sodium citrate, pH 5.8) and supplemented with glass beads (1 mm in diameter, Sigma-Aldrich, USA). Cells lysis made by vortexing and genomic DNA was isolated after phenol/chloroform extraction and ethanol precipitated. The extracted DNA was quantified with NanoDrop® ND-1000 Spectrophotometer (Nanodrop Technologies, USA) and DNA quality analyzed by agarose gel electrophoresis.

EXAMPLE 2

Agrobacterium tumefaciens-Mediated Transformation (ATMT)

[0043] ATMT was performed essentially as described previously for *R. toruloides*, *S. roseus*, *U. maydis* and *Pseudozyma aphidis* (Ji, et al., 2010; Liu, et al, 2012). Co-culturing of fungal cells and *Agrobacterium* donors were done at pH5.6, 24°C for 2.5-3 days on nylon membranes and

selection was done on YPD plates containing 300 µg/ml cefotaxime and 150-300 µg/ml (150 µg/ml for *R. toruloides*, *R. glutinis*, and 300 µg/ml for *P. aphidis* and *S. roseus*) hygromycin B at 28°C for 3-5 days.

EXAMPLE 3

Cloning of promoters

[0044] Based on the published EST sequence abundance in various media (Ho et al, 2007), we selected a number of genes (Table 1) as potential source of strong promoters for *Rhodospiridium* and *Rhodotorula*. Other candidate genes include those encoding proteins in the fatty acid biosynthesis in *Rhodospiridium* and *Rhodotorula*, eg, acetyl-CoA synthase (*ACCI*), acyl-CoA carrier protein (*ACPI*), pyruvate decarboxylase (*PDCI*) and nitrate regulated gene (*NARI*). *Ustilago maydis* CDS sequences were searched against the *Rhodospiridium* and *Rhodotorula* genome database.

TABLE 1

Candidate Genes with Strong Promoter

Complete Medium*	carbon starvation media*	Nitrogen starvation media*	dormant teliospore*	Description	MUMDB ID
7	18	22	3	glyceraldehyde 3-phosphate dehydrogenase	um02491
35	54	22	11	probable TAL1 - transaldolase	um04138
8	12	6	0	probable Actin	um11232
107	79	65	32	probable ubiquitin/ribosomal protein S27a fusion protein	um04588
146	54	16	8	probable FPR1 - peptidyl-prolyl cis-trans isomerase, FK506-binding protein	um11054
35	37	14	17	probable CPR1 - cyclophilin (peptidylprolyl isomerase)	um03726
48	13	29	4	probable translation elongation factor eEF-1 alpha chain	um00924
0	0	0	2	probable PDC1 - pyruvate decarboxylase, isozyme 1	um03994
17	6	11	1	probable ENO1 - enolase I (2-phosphoglycerate dehydratase)	um03356

Note: * Numbers indicate total hits in the EST library of the respective medium.

[0045] To define the 3' end of the promoters, 5' RACE were performed using BD SMARTer™ RACE cDNA Amplification Kit (Clontech, California, USA) according to the manufacturer's instruction. Promoter DNA fragments were obtained by PCR using a 3' end primer that is designed at the first ATG in the 5' untranslated regions, usually with an overlapping *NcoI* (CCATGG) or *BspHI* (TCATGA) site at the ATG codon. *BamHI* is used if the DNA sequence contains both *NcoI* and *BspHI* sites. 5' Primers were designed 1-2 kb from the ATG. The primers used are listed in Table 2. The PCR fragments were digested with corresponding enzymes, cloned in pPN006 or pRH2031 (Fig 1), which is a T-DNA vector containing the *RtGPD1::RtGFP:nos* cassette (Liu et al, 2012).

Table 2
Cloning of Selected Promoters

SEQ ID NO:	Promoter name	Origin	Length (bp)	5' PCR primers (SEQ ID NO:)	3' PCR primer (SEQ ID NO:)
1	Rg2TEF1	<i>R. glutinis</i> ATCC 204091 translation elongation factor	1017	AAAGGTACCGTGCG AGAAGAAGCGAGG C (13)	AAACCATGGTTCCTCC CCCAGTACACAGTAC AGTA (14)
2	Rg3TEF1	<i>R. graminis</i> strain WP1 translation elongation factor	946	AAAGGTACCCATGC TGCTGCTGCCCTC A (15)	ATACCATGGTGGATG AAGTGAGATTCGAGT G (16)
3	Rg3S27	<i>R. graminis</i> strain WP1 ribosomal protein S27	1464	AAAGGTACCCAAAG GGAGAGGAGCGGG CG (17)	AAACCATGGTGACCT ACGCCTACACAAGGG TGC (18)
4	Rg2ACP1	<i>R. glutinis</i> ATCC 204091 acyl-CoA carrier protein	1321	AAAGGTACCTGAGC GGGCGAGCCGCGAG (19)	aaaTCATGACTGAACA AAGTTTCTCTGCGGC GC (20)
5	Rg2NAR1	<i>R. glutinis</i> ATCC 204091 nitrate reductase	1499	AAAGGTACCGCTCA TCATCGAGCGAGGG CAG (21)	AAACCATGGCGGCGG GTGATTCTTCTTGTT C (22)
6	Rg2ENO1	<i>R. ATCC 204091</i> enolase(phosphopyruvate hydratase)	1596	gactagtcGCCAGGGAA CGCAGAGAAGG (23)	aaaccatgGCTGTGAGAG GAGTATCGCAG (24)
7	Rg2PDC1	<i>R. glutinis</i> ATCC 204091 pyruvate decarboxylase	1391	gactagtcGGCGAAGAG GAGGGGTGTTAGG (25)	aaaggatccgcatTGCGGT TCAAAGGCTGGGCG (26)
8	Rg3ENO1	<i>R. graminis</i> strain WP1 enolase(phosphopyruvate hydratase)	1500	gactagtcGTGCAGAAG GAACCCGAGGAG (27)	aaaccatgGCGAGGCGAG AGGGGTTGAG (28)
9	Rg3PDC1	<i>R. graminis</i> strain WP1 pyruvate decarboxylase	1500	gactagtcGGGAGGATT GATGATCGGGTTGC (29)	aaacCATGGTGGTTCGA GCTTGTGAGG (30)
10	Rg3PPI	<i>R. graminis</i> strain WP1 peptidylprolyl isomerase	1222	AAAGGTACCGCGAG GGGAAGGGCAGGA GAGTCG (31)	AAACCATGGTCGGAC GGTGGGAAGGGGGG GGA (32)

11	Rg3TAL1	<i>R. graminis strain WP1 transaldolase</i>	1507	AAAGGTACCACTAC CTCGTCACGACCCA GGGTG (33)	AAACCATGGTGTAGG TAGCTGCGGCGGGT (34)
12	Rg2ACC1	<i>R. glutinis ATCC 204091 acetyl-CoA synthase</i>	1646	AAAGGCGCGCCTGA AGCTGTACATCGAG GTGGAC (35)	aaaCCATGGTCCCACA ATCAGTAGTTGTCCT CGGAAG (36)

EXAMPLE 4

Promoter activity in *Rhodospiridium* in lipid production medium

[0046] The promoter GFP reporter constructs were transformed to *Rhodospiridium toruloides* ATCC 10657 by the ATMT method. The transformed colonies (>100) were pooled, cultured in YPD medium with 150 µg/ml hygromycin B and 300 µg/ml cefataxome and diluted to about 0.1 OD600 in lipid production medium [10 mM K₂HPO₄-KH₂PO₄, (pH6.13), 4 g/L yeast extract, 0.3 g/L urea, 0.1 g/L Na₂SO₄, 10 mg/L each of tyrosine, valine and vitamin B (B1+B6), 8% glucose] with no antibiotics added. Strains were cultured at 28°C with shaking (280rpm) 2days and then dilute to about 0.1 OD600 for 24 hours and the cultures were adjusted to 0.6 OD600 units before subjecting to florescence measurement in a Tecan M200 reader using 476 nM as excitation wavelength, and 509 nM as emission wavelength; gain value 100. The florescence intensity were normalized to OD600 and subtracted against non-transformed cell cultures under the same conditions. Rg3TAL1 is weak while Rg2ACC1 showed no obvious activity (Table 3). Transformants of the Rg2ACC1 reporter in *Rhodospiridium glutinis* ATCC 90781 showed no GFP florescence was cultured a nitrogen-limited medium (glucose 70g/l, yeast extract 0.75g/l, (NH₄)₂SO₄ 0.1g/l, KH₂PO₄ 1.0g/l, MgSO₄·7H₂O 1.5g/l, pH 5.6) for 24hr either (not shown). Promoter RtGPD1 is set forth in SEQ ID N:38.

TABLE 3

Relative GFP Fluorescence of Selected Promoters in *Rhodospiridium toruloides*

Promoter	Rg2TEF1	Rg3TEF1	Rg3TAL1	Rg3S27	Rg3PPI	Rg2ENO1	Rg2ACC1	RtGPD1
Fluorescence reading	183	616	83	316	566	383	0	1100

EXAMPLE 5

Promoter Activity in N-rich and N-free Medium
in *Rhodospiridium*, *Pseudozyma*, *Ustilago* and *Sporobolomyces*

[0047] A selected set of promoter GFP reporter constructs were transformed by the ATMT method to *Rhodospiridium toruloides* ATCC 10657, *Ustilago maydis* L8, *Pseudozyma aphidis* ATCC32657 and *Sporobotomyces roseus* FGSC 10293. The transformed colonies (>100) were

pooled, cultured in YPD medium with 150 µg/ml (or 300 µg/ml for *Pseudozyma aphidis*) hygromycin B and 300 µg/ml cefataxome for 2 days at 28°C and diluted to about 0.1 OD600 in YNB Medium and also in YNB N- Medium (both medium with 5% glucose.) and continued culture for 1-3 days with shaking. OD600 and GFP florescence were measured with Tecan infinite200. The GFP fluorescence intensity (normalized against the OD600) is listed in Tables 4-11. Promoters Umgpd1, RtGPD1, Rg3GPD1, Rg2FAD1 and SrGPD1 are set forth in SEQ IN NOs:37, 38, 39, 40 and 41, respectively. The isolation of these promoters is described in WO 2012/169969.

TABLE 4
Fluorescence Intensity in *R. toruloides* ATCC 10657 in YNB Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACP1	Rg2NAR1	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umppd1	RtGPD1	Rg3GPD1	Rg2FAD1	SrGPD1
Day1	67098.1 ±4810.2	493314.2 ±17949.6	315280.7 ±13486.8	822714.3 ±29975.4	76162.8 ±5196.2	1229511.5 ±38057.0	439906.8 ±20026.2	310853.1 ±11379.8	524990.1 ±19316.2	185894.4 ±5956.8	647267.9 ±17825.0	482273.0 ±17080.6	790204.6 ±26581.6	94929.5 ±3664.8
Day2	90051.8 ±5580.0	324001.3 ±15528.4	138885.1 ±5578.8	328422.2 ±13168.4	77194.7 ±5304.6	960118.3 ±35167.0	195357.6 ±14536.8	49159.7 ±2473.8	318157.8 ±13705.8	23430.4 ±11392.6	470829.4 ±19049.2	302113.7 ±27780.8	622505.2 ±24257.0	57888.5 ±3134.6
Day3	125629.1 ±5884.6	109829.4 ±5871.2	71034.5 ±3793.6	391254.0 ±15650.6	72916.9 ±5191.4	1080769.4 ±37277.4	186096.4 ±15397.2	69658.8 ±2717.4	179610.3 ±7379.4	52504.2 ±2197.4	417595.6 ±16078.8	294063.5 ±16996.2	678405.8 ±25364.6	72393.3 ±3358.4

TABLE 5
Fluorescence Intensity in *R. toruloides* ATCC 10657 in YNB N-free Medium

	Rg2TEF	Rg3TEF1	Rg3S27	Rg2ACP1	Rg2NAR1	Rg2ENO1	Rg2PDC1	Rg3ENO	Rg3PDC1	Umppd1	RtGPD1	Rg3GPD1	Rg2FAD1	SrGPD1
Day1	70881.8 ±5110.9	380951.5 ±18050.3	365547.6 ±13587.5	815919.4 ±30076.1	272301.0 ±5296.9	1424622.6 ±38157.7	446260.0 ±20126.9	146494.1 ±11480.5	521271.4 ±19416.9	179467.0 ±6057.5	798469.1 ±17925.7	543617.9 ±17181.3	1046580.2 ±26682.3	208402.0 ±3765.5
Day2	85225.0 ±5880.7	154658.6 ±15629.1	229116.8 ±5979.5	194767.0 ±13269.1	96375.2 ±5405.3	1023479.6 ±35267.7	189373.9 ±14637.5	87625.2 ±2574.5	229565.0 ±13806.5	70298.9 ±11493.3	559941.0 ±19149.9	325267.7 ±27881.5	729764.3 ±24357.7	74769.4 ±3235.3
Day3	101884.6 ±6085.3	75616.7 ±5971.9	131082.7 ±3894.3	274152.1 ±15751.3	49724.9 ±5292.1	1045808.5 ±37378.1	175901.0 ±15497.9	41721.8 ±2818.1	116966.8 ±7480.1	42892.0 ±2298.1	240704.2 ±16179.5	143840.6 ±17096.9	495414.9 ±25564.6	27930.9 ±3459.1

TABLE 6
Fluorescence Intensity in *P. aphidis* in YNB Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACP1	Rg2NAR1	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umcpd1	RtGPD1	Rg3GPD1	Rg2FAD1	StGPD1
Day1	20710.7	24954.1	17864.7	25694.5	11741.6	15870.6	21807.7	19768.3	27733.8	22180.7	29319.0	37495.4	16309.0	23103.0
	±535.5	±1347.7	±903.2	±1084.7	±577.1	±783.5	±790.4	±978.4	±986.7	±1009.0	±465.9	±1774.8	±215.5	±1455.1
Day2	5617.3	7829.7	4804.2	7541.3	3854.7	4634.5	6430.8	5748.1	8123.9	6537.3	9876.9	9713.0	4859.7	7830.9
	±268.1	±612.5	±403.4	±629.4	±311.2	±306.2	±595.9	±519.0	±617.0	±584.8	±803.1	±709.4	±615.0	±752.6
Day3	3738.0	3980.9	3106.8	3793.1	1679.0	2504.6	3504.2	3165.2	4602.1	3260.7	4542.3	5781.1	2671.1	3900.5
	±173.8	±298.1	±240.7	±699.3	±137.9	±301.5	±410.4	±396.5	±780.2	±324.1	±154.2	±478.1	±167.1	±290.0

TABLE 7
Fluorescence Intensity in *P. aphidis* YNB N-free Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACP1	Rg2NAR1	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umcpd1	RtGPD1	Rg3GPD1	Rg2FAD1	StGPD1
Day1	4030.5	17201.3	4757.5	14666.9	5485.1	13203.8	10218.7	12705.6	11338.6	9034.5	11511.6	9322.7	13410.0	21144.1
	±455.2	±1145.5	±367.8	±322.0	±490.5	±466.0	±271.8	±831.7	±338.7	±257.7	±396.1	±208.6	±383.1	±936.9
Day2	5247.5	14281.5	7093.2	7835.6	6845.8	6734.8	5801.6	3302.7	9382.1	13731.3	10564.5	10390.4	16069.0	16013.3
	±214.5	±290.0	±322.7	±503.6	±549.0	±245.0	±476.7	±415.2	±193.6	±467.8	±142.5	±567.5	±492.0	±202.1
Day3	5762.8	12371.9	1983.5	2564.9	2014.3	2791.4	2013.8	5876.3	5874.1	5004.3	7856.3	9851.7	12774.0	14045.7
	±530.3	±223.6	±180.5	±224.5	±203.4	±226.1	±187.8	±297.4	±585.2	±243.1	±215.7	±233.6	±225.3	±317.5

TABLE 8
Fluorescence Intensity in *U. maydis* L8 in YNB Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACPI	Rg2NARI	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umcpd1	Rg3GPD1	Rg2FAD1	SrGPD1
	4497.4	31255.7	8744.3	16705.1	5663.0	10351.3	17224.0	18699.4	30749.5	12977.4	16539.8	35361.2	20121.3
Day1	±405.1	±1974.8	±743.4	±987.7	±598.1	±608.5	±813.1	±689.9	±1658.1	±978.4	±540.3	±1290.8	±932.4
	2237.4	20291.8	12961.1	7046.8	1899.9	3508.0	1948.5	4185.3	10692.2	14498.0	5729.0	13483.6	13639.8
Day2	±190.0	±1064.2	±789.4	±584.2	±252.3	±183.5	±268.4	±236.9	±452.9	±1096.3	±290.4	±728.5	±567.3
	728.7	29221.7	2132.4	1651.0	6453.6	1885.2	1610.0	10261.0	13449.8	10667.8	4328.0	19667.2	7559.5
Day3	±142.3	±1935.6	±196.8	±125.3	±295.7	±138.7	±198.6	±858.7	±689.7	±108.7	±498.1	±1282.3	±219.2

TABLE 9
Fluorescence Intensity in *U. maydis* L8 in YNB N-free Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACPI	Rg2NARI	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umcpd1	Rg3GPD1	Rg2FAD1	SrGPD2
	2736.3	23444.3	4848.0	14420.4	4965.9	3393.2±	5099.7	13592.6	20887.7	8457.1	36784.6	27948.4	19682.9
Day1	±195.3	±1674.3	±743.4	±785.7	±291.7	258.4	±363.5	±907.9	±1043.2	±275.2	±909.4	±1073.6	±1136.5
	4635.9	22335.2	11189.0	12643.8	8929.3	8782.8±	6278.8	16443.4	18643.1	15615.6	31362.0	19059.7	10507.5
Day2	±200.5	±984.7	±628.3	±784.1	±332.9	384.2	±435.4	±1023.5	±853.6	±1001.7	±931.9	±839.1	±509.5
	4778.9	21945.0	5243.9	3563.5	1052.9	402.1	2136.7	3832.4	10741.0	5416.3	22029.0	14070.9	5509.8
Day3	±242.3	±1531.6	±171.1	±105.6	±99.3	±138.7	±108.1	±251.7	±619.2	±157.7	±791.3	±812.8	±293.1

TABLE 10
Fluorescence Intensity in *S. roseus* for in YNB Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACP1	Rg2NAR1	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umcpd1	RiGPD1	Rg3GPD1	Rg2FAD1	SrGPD2
Day1	34433.7 ±2961.3	62781.0 ±5013.6	49396.6 ±3965.7	35726.2 ±3012.6	44850.3 ±7639.8	54393.9 ±2013.6	45625.4 ±3968.1	75257.5± 3086.3	75080.8 ±3015.7	61607.3 ±2345.1	115602.4 ±4598.7	36185.9 ±2845.2	55143.4 ±5423.9	76783.0 ±1946.3
Day2	7223.1 ±1025.8	7997.9 ±1169.4	5486.4 ±1032.6	11336.9 ±698.1	7453.9 ±854.1	16829.3 ±852.3	7401.2± 897.5	7108.0± 1087.6	7110.1 ±521.7	13471.3 ±369.3	7832.5 ±496.8	6434.0 ±963.4	10613.7 ±1015.8	24326.0 ±876.5
Day3	9874.2 ±1356.9	6212.6 ±1598.7	7978.3 ±1648.1	6425.8 ±963.7	8653.9 ±1587.5	9632.8 ±897.6	9013.8± 1258.4	9898.4± 1698.7	9985.6 ±765.1	8013.7 ±965.4	11231.1 ±1065.8	4931.2 ±1259.8	7469.8 ±1345.9	16036.7 ±1897.4

TABLE 11
Fluorescence Intensity in *S. roseus* in YNB N-free Medium

	Rg2TEF1	Rg3TEF1	Rg3S27	Rg2ACP1	Rg2NAR1	Rg2ENO1	Rg2PDC1	Rg3ENO1	Rg3PDC1	Umcpd1	RiGPD1	Rg3GPD1	Rg2FAD1	SrGPD2
Day1	30415.8 ±2111.3	38195.6 ±2163.4	34075.9 ±3115.7	37998.4 ±2162.6	39594.0 ±4678.3	38138.8 ±1163.6	25733.8 ±3118.1	43722.6 ±2236.3	42986.2 ±2165.7	53744.0 ±1495.1	60102.2 ±2365.9	49641.6 ±1995.2	45162.3 ±4573.9	47127.7 ±1096.3
Day2	6135.6 ±905.8	9358.5 ±1049.4	10507.2 ±912.6	13226.8 ±578.1	9190.2 ±734.1	11034.9 ±732.3	5662.3 ±777.5	11550.1 ±967.6	11959.0 ±401.7	7691.8 ±249.3	12012.4 ±376.8	10303.5 ±843.4	10948.1 ±895.8	14795.5 ±756.5
Day3	2698.7 ±1206.9	5401.2 ±1448.7	3012.9 ±1498.1	6968.6 ±813.7	2165.8 ±1437.5	2857.4 ±747.6	3541.9 ±1108.4	4986.2 ±1548.4	4879.3 ±615.1	5897.1 ±815.4	7854.3 ±915.8	4013.7 ±1109.8	3946.5 ±1195.9	9154.6 ±1747.4

EXAMPLE 6

Identification of Critical Elements for Strong Promoters

[0048] Promoter sequences of Rg3TEF1, Rg3S27, Rg2ACP1, Rg2ENO1, Rg2PDC1, Rg3PDC1 and Rg2FAD1 (stearoyl-CoA delta-9 desaturase) were subjected to promoter motif scanning using the Gibbs Motif Sampler at http://ccmbweb.ccv.brown.edu/cgi-bin/gibbs.12.pl?data_type=DNA&layout=advancedprgm&restore=/var/www/cgi-bin/euk.def.txt.

[0049] A conserved motif sharing the GAGGAG core sequence were found in each promoter. Noticeably, Rg2FAD1 promoter, which is among the strongest promoters contains the largest number of the motif. (Fig 2).

EXAMPLE 7

Nested Deletion of Rg2FAD1 and Rg2ENO1 Promoters

[0050] The full length Rg2ENO1 and Rg2FAD1 (stearoyl-CoA delta-9 desaturase gene) promoter GFP reporter constructs in pRH2031 were modified to have serially shortened promoters. This was done replacing the promoter with PCR fragments about 300, 500, 1000 and 1500 bp version of the promoter. All 5' primers included a *SpeI* cutting site while the 3' primer contains a *NcoI* cutting site. Constructs were transformed by the ATMT method to *Rhodospiridium toruloides* ATCC 10657. The transformed colonies (>500) were pooled, cultured in YNB medium with 150 µg/ml hygromycin B and 300 µg/ml cefataxome for 2 days at 28°C and diluted to about 0.1 OD₆₀₀ in YNB Medium and also in YNB N⁻ Medium (both medium with 5% glucose.) and continued culture for 24 hours with shaking. The cultures reached OD₆₀₀ 0.2 units in YNB N⁺ and YNB N⁻ media. GFP florescence were measured with Tecan infinite M200. Gain parameter is consistently set at 85; Excitation and Emission wavelength are 476, 509. The GFP fluorescence intensity (normalized against the OD600) is listed in Fig. 3, which shows that minimal length of the ENO1 promoter for the optimal expression of reporter gene is approximately 320 to 520 bp, whereas the FAD1 promoter requires approximately 570 to 1120 bp.

EXAMPLE 8

Nested Deletion of 519 bp Rg2ENO1 Promoter

[0051] Primers were designed at various locations in the 519 bp Rg2ENO1 promoter sequence, which were used for PCR in combination with the reverse primer targeting the 3' end of the promoter (Fig. 4A). All 5' primers included a *SpeI* cutting site while the 3' primer contains a *NcoI* cutting site. The lengths of the PCR products (excluding the extra linker sequence at 5' end and

the ATG codon at the 3' site) are summarized in Fig. 4B. The PCR fragments were individually digested with *SpeI* and *NcoI* and cloned in pRH2031-Rg2ENO1-RtGFP at the same sites, replacing the full-length *ENO1* promoter. Constructs were transformed by the ATMT method to *Rhodotorula glutinis* ATCC 90781, which is the diploid parent of *Rhodospiridium toruloides* ATCC 10657 and ATCC 10788. The transformed colonies (>500) were pooled, cultured in YNB medium with 150 µg/ml hygromycin B and 300 µg/ml cefataxome for 2 days at 28°C and diluted to about 0.1 OD₆₀₀ in YNB Medium and also in YNB N⁻ Medium (both medium with 5% glucose.) and continued culture for 12 hours with shaking. The cultures reached OD₆₀₀ 0.5~0.7 in YNB and YNB N⁻ media. GFP florescence was measured with Tecan infinite M200. Gain parameter is consistently set at 85; Excitation and Emission wavelength are 476, 509. The GFP fluorescence intensity (normalized against the OD₆₀₀) is listed in Fig. 5A and 5B. The promoter showed similar trends in the two media tested. The biggest drop in activity was seen promoter M6 and M7. Another significant drop was observed between M3 and M6, where several GAGGAG -related motifs can be found (Fig. 4A).

[0052] The use of the terms “a” and “an” and “the” and similar referents in the context of describing the invention (especially in the context of the following claims) are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms “comprising,” “having,” “including,” and “containing” are to be construed as open-ended terms (i.e., meaning “including, but not limited to,”) unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any non-claimed element as essential to the practice of the invention.

[0053] Embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the

invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.

[0054] The term “efficient expression” refers to expression of a reporter protein to a level that is detectable for fluorometry, photomicroscopy or phenotypic selection of transformants by antibiotics, such as hygromycin.

BIBLIOGRAPHY

- [0055] An, G., et al., 1989. Functional analysis of the 3' control region of the potato wound-inducible proteinase inhibitor II gene. *Plant Cell*, 1: 115-122.
- [0056] Bassett, C.L., 2007. Regulation of Gene Expression in Plants: The Role of Transcript Structure and Processing. New York: Springer Press.
- [0057] Bölker, M., et al., 1995. Tagging pathogenicity genes in *Ustilago maydis* by restriction enzyme-mediated integration (REMI). *Mol Gen Genet*. 248, 547-52.
- [0058] Boulton, S., et al., 1999. Interactive effects of inhibitors of poly(ADP-ribose) polymerase and DNA-dependent protein kinase on cellular responses to DNA damage. *Carcinogenesis*. 20, 199-203.
- [0059] Boulton, S., et al., 1996. Wortmannin is a potent inhibitor of DNA double strand break but not single strand break repair in Chinese hamster ovary cells. *Carcinogenesis*. 17, 2285-90.
- [0060] Bundock, P., et al., 1995. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J*. 14, 3206-14.
- [0061] Choi, J., et al., 2007. Genome wide analysis of T DNA integration into the chromosomes of *Magnaporthe oryzae*. *Molecular Microbiology*. 66, 371-382.
- [0062] Comai, L., et al., 1983. An altered *aroA* gene product confers resistance to the herbicide glyphosate. *Science*. 221, 370.
- [0063] De Groot, M. J. A., et al., 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nature Biotechnology*. 16, 839-842.
- [0064] de Oliveira, M. L. P., et al., 2009. High-efficiency *Agrobacterium*-mediated transformation of citrus via sonication and vacuum infiltration. *Plant Cell Reports*. 28, 387-395.
- [0065] Durant, S., Karran, P., 2003. Vanillins--a novel family of DNA-PK inhibitors. *Nucleic Acids Res*. 31, 5501-12.
- [0066] Gietz, R.D. and Woods, R. A., 2002. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol*. 350, 87-96.
- [0067] Goldstein, A. L. and McCusker, J. H., 1999. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast*. 15, 1541-1553.
- [0068] Gutiérrez, R.A., et al., 1999. Current perspectives on mRNA stability in plants: multiple levels and mechanisms of control. *Trends Plant Sci*, 4: 429-438.

- [0069] Hajdukiewicz, P., et al., 1994. The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Molecular Biology*. 25, 989-994.
- [0070] Haughn, G. W., et al., 1988. Transformation with a mutant *Arabidopsis* acetolactate synthase gene renders tobacco resistant to sulfonylurea herbicides. *Molecular and General Genetics*. 211, 266-271.
- [0071] Heiser, W. C., 2000. Optimizing electroporation conditions for the transformation of mammalian cells. *Methods in Molecular Biology*. 130, 117-34.
- [0072] Hentges, P., et al., 2005. Three novel antibiotic marker cassettes for gene disruption and marker switching in *Schizosaccharomyces pombe*. *Yeast*. 22, 1013-9.
- [0073] Hewald, S., et al., 2005. Genetic analysis of biosurfactant production in *Ustilago maydis*. *Applied and Environmental Microbiology*. 71, 3033.
- [0074] Hill, J., et al., 1991. DMSO-enhanced whole cell yeast transformation. *Nucleic Acids Research*. 19, 5791.
- [0075] Ho, Eric CH, et al (2007). Gene discovery and transcript analyses in the corn smut pathogen *Ustilago maydis*: expressed sequence tag and genome sequence comparison. *BMC genomics* 8:334.
- [0076] Horns, F., et al. 2012. Patterns of repeat-induced point mutation in transposable elements of basidiomycete fungi. *Genome biology and evolution*, 4, 240-247.
- [0077] Hu, C., et al., 2009. Effects of biomass hydrolysis by-products on oleaginous yeast *Rhodospiridium toruloides*. *Bioresour Technol*. 100, 4843-7.
- [0078] Ianiri, G., et al., 2011. Development of resources for the analysis of gene function in *Pucciniomycotina* red yeasts. *Fungal Genetics and Biology*. 48, 685-695.
- [0079] Ingelbrecht, I.L., et al., 1989. Different 3' end regions strongly influence the level of gene expression in plant cells. *Plant Cell*, 1: 671-680.
- [0080] Ito, H., et al., 1983. Transformation of intact yeast cells treated with alkali cations. *Journal of bacteriology*. 153, 163-8.
- [0081] Ji, L., et al., 2010. A Simplified and efficient method for transformation and gene tagging of *Ustilago maydis* using frozen cells. *Fungal Genet Biol*. 47, 279-87.
- [0082] Kämper, J., 2004. A PCR-based system for highly efficient generation of gene replacement mutants in *Ustilago maydis*. *Mol Genet Genomics*. 271, 103-10.
- [0083] Khanna, H. K., et al., 2007. Inhibition of *Agrobacterium*-induced cell death by antiapoptotic gene expression leads to very high transformation efficiency of banana. *Molecular Plant-Microbe Interactions*. 20, 1048-1054.
- [0084] Kirk, M. P., et al. Dictionary of the Fungi. CABI, Wallingford, 2008, pp. 716.
- [0085] Krugel, H., et al., 1988. Analysis of the nourseothricin-resistance gene (nat) of *Streptomyces noursei*. *Gene*. 62, 209-217.
- [0086] Kück, U. and Hoff, B., 2010. New tools for the genetic manipulation of filamentous fungi. *Appl Microbiol Biotechnol*. 86, 51-62.
- [0087] Kuo, C. Y., et al., 2004. Cloning of glyceraldehyde-3-phosphate dehydrogenase gene and use of the *gpd* promoter for transformation in *Flammulina velutipes*. *Appl Microbiol Biotechnol*. 65, 593-9.

- [0088] Lazo, G. R., et al., 1991. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. *Biotechnology (N Y)*. 9, 963-7.
- [0089] Liu, Y. G., Chen, Y., 2007. High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *BioTechniques*. 43, 649-50, 652, 654 passim.
- [0090] Liu et al, 2012, Characterization of glyceraldehyde-3-phosphate dehydrogenase gene *RtGPD1* and development of genetic transformation method by dominant selection in oleaginous yeast *Rhodospiridium toruloides*. *Applied Microbiology and Biotechnology*. 10.1007/s00253-012-4223-9.
- [0091] Liu, Y. G., Whittier, R. F., 1995. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*. 25, 674-81.
- [0092] Maehara, T., et al., 2010. Improvement of the Transformation Efficiency of *Flammulina velutipes* Fv-1 Using the Glyceraldehyde-3-phosphate Dehydrogenase Gene Promoter. *Bioscience, Biotechnology, and Biochemistry*. 74, 2523-2525.
- [0093] Maier, F. J., Schafer, W., 1999. Mutagenesis via insertional- or restriction enzyme-mediated-integration (REMI) as a tool to tag pathogenicity related genes in plant pathogenic fungi. *Biol Chem*. 380, 855-64.
- [0094] Meng, X., et al., 2009. Biodiesel production from oleaginous microorganisms. *Renewable Energy*. 34, 1-5.
- [0095] Meyer, V., 2008. Genetic engineering of filamentous fungi--Progress, obstacles and future trends. *Biotechnology Advances*. 26, 177-185.
- [0096] Meyer, V., et al., 2003. Comparison of different transformation methods for *Aspergillus giganteus*. *Curr Genet*. 43, 371-7.
- [0097] Mogen, B.D., et al., 1990. Upstream sequences other than AAUAAA are required for efficient messenger RNA 3'-end formation in plants. *Plant Cell*, 2: 1261-1272.
- [0098] Mogen, B.D., et al., 1992. Several distinct types of sequence elements are required for efficient mRNA 3' end formation in a pea *rbcS* gene. *Molecular and Cellular Biology*, 12: 5406-5414.
- [0099] Nagaya, S., et al., 2010. The HSP terminator of *Arabidopsis thaliana* Increases Gene Expression in Plant Cells. *Plant Cell Physiol*, 51: 328-332.
- [0100] Newman, T.C., et al., 1993. DST sequences, highly conserved among plant SAUR genes, target reporter transcripts for rapid decay in tobacco. *Plant Cell*, 5: 701-14.
- [0101] Ochman, H., et al., 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics*. 120, 621-3.
- [0102] Ohme-Takagi, M., et al., 1993. The effect of sequences with high AU content on mRNA stability in tobacco. *Proc Natl Acad Sci U S A*, 90: 11811-5.
- [0103] Pfeifer, T. A., et al., 1997. Baculovirus immediate-early promoter-mediated expression of the Zeocin (TM) resistance gene for use as a dominant selectable marker in Dipteran and Lepidopteran insect cell lines. *Gene*. 188, 183-190.
- [0104] Punt, P. J., et al., 1987. Transformation of *Aspergillus* based on the hygromycin B resistance marker from *Escherichia coli*. *Gene*. 56, 117-124.

- [0105] Rosenzweig, K. E., et al., 1997. Radiosensitization of human tumor cells by the phosphatidylinositol3-kinase inhibitors wortmannin and LY294002 correlates with inhibition of DNA-dependent protein kinase and prolonged G2-M delay. Clin Cancer Res. 3, 1149-56.
- [0106] Rothnie, H.M., et al., 1994. The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3'-end formation in plants. EMBO (Eur Mol Biol Organ) J, 13: 2200-2210.
- [0107] Schulz, B., et al., 1990. The b alleles of *U. maydis*, whose combinations program pathogenic development, code for polypeptides containing a homeodomain-related motif. Cell. 60, 295-306.
- [0108] Scorer, C. A., et al., 1994. Rapid Selection Using G418 of High Copy Number Transformants of *Pichia pastoris* for High level Foreign Gene Expression. Nature Biotechnology. 12, 181-184.
- [0109] Skalitzky, D. J., et al., 2003. Tricyclic benzimidazoles as potent poly(ADP-ribose) polymerase-1 inhibitors. J Med Chem. 46, 210-3.
- [0110] Smith, G. and Jackson, S., 1999. The DNA-dependent protein kinase. Genes & development. 13, 916.
- [0111] Smith, T. L. and Leong, S. A., 1990. Isolation and characterization of a *Ustilago maydis* glyceraldehyde-3-phosphate dehydrogenase-encoding gene. Gene. 93, 111-7.
- [0112] Soltani, J., et al., Agrobacterium-mediated transformation of non-plant organisms. In: T. Tzfira, V. Citovsky, Eds.), Agrobacterium: from biology to biotechnology. Springer press, New York, USA, 2008, pp. 649-675.
- [0113] Spellig, T., et al., 1996. Green fluorescent protein (GFP) as a new vital marker in the phytopathogenic fungus *Ustilago maydis*. Mol Gen Genet. 252, 503-9.
- [0114] Steiner, S. and Phillippsen P, 1994. Sequence and promoter analysis of the highly expressed TEF gene of the filamentous fungus *Ashbya gossypii*. Mol Gen Genet 242, 263-271.
- [0115] Sweigard, J. A., et al., 1998. Magnaporthe grisea pathogenicity genes obtained through insertional mutagenesis. Mol Plant Microbe Interact. 11, 404-12.
- [0116] Takeno, S., et al., 2005. Transformation of oil-producing fungus, *Mortierella alpina* 1S-4, using Zeocin, and application to arachidonic acid production. Journal of bioscience and bioengineering. 100, 617-622.
- [0117] Teichmann, B., et al., 2010. Molecular characterization of the biocontrol activity of *Pseudozyma flocculosa*. Phytopathology. 100.
- [0118] Tentori, L., et al., 2002. Potential clinical applications of poly(ADP-ribose) polymerase (PARP) inhibitors. Pharmacol Res. 45, 73-85.
- [0119] Tully, M. and Gilbert, H. J., 1985. Transformation of *Rhodosporidium toruloides*. Gene. 36, 235-40.
- [0120] Turgeon, B. G., et al., 2010. Protoplast transformation of filamentous fungi. Methods in molecular biology. 638, 3-19.
- [0121] van Attikum, H., et al., 2001. Non-homologous end-joining proteins are required for Agrobacterium T-DNA integration. Embo J. 20, 6550-8.

- [0122] van Attikum, H., et al., 2003. The Arabidopsis AtLIG4 gene is required for the repair of DNA damage, but not for the integration of Agrobacterium T-DNA. *Nucleic Acids Res.* 31, 4247-55.
- [0123] Vega, J. M., et al., 2008. Improvement of Agrobacterium-mediated transformation in Hi-II maize (*Zea mays*) using standard binary vectors. *Plant cell reports.* 27, 297-305.
- [0124] Veuger, S. J., et al., 2003. Radiosensitization and DNA repair inhibition by the combined use of novel inhibitors of DNA-dependent protein kinase and poly(ADP-ribose) polymerase-1. *Cancer Res.* 63, 6008-15.
- [0125] Willmore, E., et al., 2004. A novel DNA-dependent protein kinase inhibitor, NU7026, potentiates the cytotoxicity of topoisomerase II poisons used in the treatment of leukemia. *Blood.* 103, 4659-65.
- [0126] Wu, S. and Letchworth, G. J., 2004. High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol. *BioTechniques.* 36, 152-155.
- [0127] Wu, S., et al., 2010a. Phosphate-limitation mediated lipid production by *Rhodospiridium toruloides*. *Bioresour Technol.* 101, 6124-9.
- [0128] Wu, S., et al., 2010b. Microbial lipid production by *Rhodospiridium toruloides* under sulfate-limited conditions. *Bioresour Technol.*
- [0129] Ye, X. and Gilbertson, L., use of multiple transformation enhancer sequences to improve plant transformation efficiency. EP Patent 2,038,420, 2009.
- [0130] Zhao, X., et al., 2010a. Lipid production by *Rhodospiridium toruloides* Y4 using different substrate feeding strategies. *J Ind Microbiol Biotechnol.*
- [0131] Zhao, X., et al., 2010b. Lipid production from Jerusalem artichoke by *Rhodospiridium toruloides* Y4. *J Ind Microbiol Biotechnol.* 37, 581-5.
- [0132] Zhu, Z., et al, 2012, A multi-omic map of the lipid-producing yeast *Rhodospiridium toruloides*, *Nature Communications*, 3, 1112.

WHAT IS CLAIMED IS:

1. A DNA construct comprising an isolated nucleotide sequence selected from SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or 11 or promoter portion thereof operatively linked to a polypeptide-encoding sequence operatively linked to a transcriptional terminator, wherein the DNA construct enables efficient expression of the polypeptide in a fungal species selected from *Pucciniomycotina* and *Ustilaginomycotina* subphyla.
2. The DNA construct of claim 1, wherein the promoter portion is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon.
3. The DNA construct of claim 1 or 2, wherein the DNA construct is inserted in a T-DNA vector, a shuttle vector, or in a fungal chromosome.
4. The DNA construct of claim 1 or 2, wherein the fungal species is selected from *Rhodospiridium*, *Rhodoturula*, *Pseudozyma*, *Ustilago* or *Sporobolomyces* genus, wherein the genome contains at least 50% C and G.
5. The DNA construct of claim 1 or 2, wherein the polypeptide-encoding sequence contains at least 50% CG, preferably 60% CG and most preferably more than 70% CG.
6. The DNA construct of claim 1 or 2, wherein the isolated nucleotide sequence contains at least one GAGGAG sequence motif, or its reverse complement sequence.
7. A fungal cell containing at least one DNA construct of any one of claims 1 to 6.
8. The fungal cell of claim 7, wherein the cell preferably contains no more than 3 identical copies of a nucleotide sequence set forth in SEQ ID NO:1-11 or promoter portion thereof, preferably not more than 2 identical copies of a nucleotide sequence set forth in SEQ ID NO:1-11 or promoter portion thereof.
9. The fungal cell of claim 8, wherein the promoter portion is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon.

10. The fungal cell of claim 7, 8 or 9, wherein the fungal cell is species selected from *Pucciniomycotina* and *Ustilaginomycotina* subphyla, preferably selected from *Rhodosporidium*, *Rhodoturula*, *Pseudozyma*, *Ustilago* or *Sporobolomyces* genus.
11. An isolated promoter selected from the group consisting of:
 - (i) a promoter comprising the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof;
 - (ii) a promoter comprising a nucleotide sequence having at least 50% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof;
 - (iii) a promoter comprising a nucleotide sequence having at least 60% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof;
 - (iv) a promoter comprising a nucleotide sequence having at least 70% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof;
 - (v) a promoter comprising a nucleotide sequence having at least 80% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof;
 - (vi) a promoter comprising a nucleotide sequence having at least 90% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof;
 - (vii) a promoter comprising a nucleotide sequence having at least 95% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof; and
 - (viii) a promoter comprising a nucleotide sequence having at least 98% identity with the nucleotide sequence of SEQ ID NO:1, 2, 3, 4, 5, 6, 7, 8, 9, 10 or 11 or promoter portion thereof,wherein the promoter enables efficient expression of the polypeptide in a fungal species selected from *Pucciniomycotina* and *Ustilaginomycotina* subphyla.

12. The isolated promoter of claim 11, wherein the promoter portion is about 400 base pairs up to about 1100 base pairs in length starting from the -1 position from the ATG codon.
13. The isolated promoter of claim 11 or 12, wherein the isolated promoter is inserted in a T-DNA vector, a shuttle vector, or in a fungal chromosome.
14. The isolated promoter of claim 11, 12 or 13, wherein the fungal species is selected from *Rhodospiridium*, *Rhodoturula*, *Pseudozyma*, *Ustilago* or *Sporobolomyces* genus, wherein the genome contains at least 50% C and G.