



(51) International Patent Classification:

C12N 1/02 (2006.01) *C12P 7/64* (2006.01)
C12N 1/16 (2006.01) *C12N 15/81* (2006.01)
C12N 15/64 (2006.01)

(21) International Application Number:

PCT/US2014/071165

(22) International Filing Date:

18 December 2014 (18.12.2014)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/984,254 25 April 2014 (25.04.2014) US

(71) Applicant (for all designated States except US): **BP BIO-FUELS UK LIMITED** [GB/GB]; Chertsey Road, Subury-on-thames, Middlesex TW16 7BP (GB).

(72) Inventors; and

(71) Applicants (for US only): **BURJA, Adam** [AU/US]; 150 West Warrenville Road, MC 200-1W, Naperville, IL 60563 (US). **HANSEN, Jon** [US/US]; 150 West Warrenville Road, MC 200-1W, Naperville, IL 60563 (US). **SIMPSON, David** [US/US]; 150 West Warrenville Road, MC 200-1W, Naperville, IL 60563 (US). **APT, Kirk** [US/US]; 150 West Warrenville Road, MC 200-1W, Naperville, IL 60563 (US). **ZIRKLE, Ross** [US/US]; 150 West Warrenville Road, MC 200-1W, Naperville, IL 60563 (US).

(74) Agent: **CUMMINGS, Kelly, L.**; BP America Inc., 150 West Warrenville Road, MC 200-1W, Naperville, IL 60563 (US).

(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, BN, 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, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: DEVELOPMENT OF MOLECULAR BIOLOGY TOOLS FOR BASIDIOMYCETE 'RED' YEAST, *SPORIDILOBOLUS PARAROSEUS* ENGINEERING

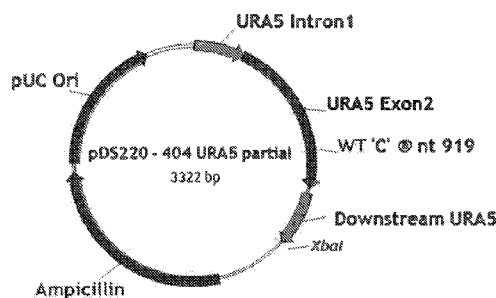


FIG. 1

(57) Abstract: A molecular biology toolkit including tools, systems, methods, and protocols for the metabolic engineering of Urediniomycetes, specifically *Sporidiobolus pararoseus*. These molecular biology tools includes extraction of genetic material such as RNA, DNA, and protein; transformation including chromosomal integration, gene targeting, and homologous recombination; and expression of both homologous and heterologous genes.

**DEVELOPMENT OF MOLECULAR BIOLOGY TOOLS
FOR BASIDIOMYCETE 'RED' YEAST, *SPORIDIOBOLUS PARAROSEUS*
ENGINEERING**

5 **CROSS-REFERENCE TO RELATED APPLICATIONS**

This application claims benefit of U.S. provisional patent application Serial No. 61/984,254 filed April 25, 2014.

NAMES OF THE PARTIES TO A JOINT RESEARCH AGREEMENT

10 For purposes of prior art determination, a joint research agreement was executed between BP Biofuels UK Limited and Martek Biosciences Corporation on December 18, 2008 in the field of biofuels. Also for purposes of prior art determination, a joint development agreement was executed between BP Biofuels UK Limited and Martek Biosciences Corporation on August 7, 2009 in the field of biofuels. Also for purposes of prior art determination, a joint development agreement
15 was executed between BP Biofuels UK Limited and DSM Biobased Products and Services B.V. on September 1, 2012 in the field of biofuels.

TECHNICAL FIELD

 The invention relates to tools, systems, methods, and protocols developed for the manipulation of recalcitrant Basidiomycete 'red' yeast, *Sporidiobolus pararoseus*,
20 for fermentation and lipid product improvement.

BACKGROUND

 A number of technologies for converting feedstocks into biofuels have been developed. A primary concern in the production of a bio-fuel / renewable diesel is the desire to improve the fermentation characteristics of the oleaginous yeast production
25 system. In particular, by improving fermentation performance, productivity and product yield, a maximum yield of lipids from a starting carbon source may be achieved.

 Genetic engineering of fermentation organisms has been a relatively common approach to the development of an economically viable, large-scale platform for the
30 production of biochemicals. However, this approach generally is taken with genetically established and tractable platform microorganisms, such as the bacterium *Escherichia coli* or the yeast *Saccharomyces cerevisiae*, which are well represented in the scientific and patent literature. Often an organism is selected for its established genetic tractability and then engineered for both product specifications and
35 fermentation yield and productivity requirements. Using such an approach, an

organism's inherent fermentative capability, namely yield and productivity, and product profile, namely fatty acids, are considered secondarily. Consequently, fermentation organisms having desirable inherent fermentative capabilities and product profiles can be overlooked if their genetic tractability is less workable than
5 other fermentation organisms.

One organism that is known to be a prolific lipid producing strain and capable of robust growth and production rates is a unicellular Basidiomycete 'red' yeast, *Sporidiobolus pararoseus*, which is a member of the class Urediniomycetes. Basidiomycetes are a unique group of fungi having a genetic architecture which can
10 make them exceedingly difficult to genetically manipulate. This species is not represented within the literature to any great extent. Although limited data is available, having only one publicly sequenced member of the group, namely *Sporobolomyces roseus* with a reduced 10 MB genome (compared to the 21.8MB genome of *Sporidiobolus pararoseus* MK29404), one atypical feature is that it is not
15 uncommon for a 3kbp gene to have upwards of 10 introns. This is an attribute which has previously all but excluded members of this group as a viable platform upon which to develop metabolic engineering approaches.

There is a need and a desire for tools, systems, methods, and protocols that can be used to select, manipulate and improve fermentation organisms for use in the
20 production of biofuels. There is a further need and desire for tools, systems, methods, and protocols that can be used to first select fermentation organisms having high yield and productivity characteristics and subsequently improve and optimize the selected fermentation organisms for use in the production of biofuels. There is an even further need and desire for molecular biology tools for the metabolic engineering of
25 Urediniomycetes, specifically, *S. pararoseus*, to provide optimal yields and production of middle distillate lipids.

SUMMARY

The invention relates to the development of molecular biology tools including parts, systems, methods, and protocols for the metabolic engineering of
30 Urediniomycetes, specifically *Sporidiobolus pararoseus*. Using these molecular biology tools, it is possible to successfully perform over-expression of native genes, integrate functions from other organisms, and perform targeted gene knock-outs and other genetic modifications. These tools enable the metabolic engineering of

Urediniomycetes for optimal yields and production of middle distillate lipids, resulting in improved fermentation and biodiesel production.

More particularly, according to certain embodiments, the methods herein may include DNA, RNA, and protein extraction. The methods may also include
5 protoplasting through enzyme activity, or through the development of improved protoplast formation cocktails.

In further embodiments, the molecular biology tools may include auxotrophic marker systems, such as URA3⁻ or URA5⁻ systems. Molecular biology tools may also include dominant selectable marker development, such as by providing a table of
10 all markers. Another component of the molecular biology tools may include one or more chromosomal integration or transformation vectors.

Methods herein may also include transformation of MK29404, such as through genetic material integration via biolistics, electroporation, and/or polyethylene glycol (PEG)-mediated transformation. These transformation
15 techniques may utilize URA3 or URA5 transformation vectors. Additionally or alternatively, these transformation techniques may utilize additional auxotrophic, reporter genes or dominant selectable marker vectors.

Molecular biology tools may further include both homologous over-expression as well as heterologous over-expression, with the incorporation of introns for
20 successful heterologous gene over-expression.

Methods herein may further include gene targeting or heterologous recombination, either through the use of a 'split marker' system or using targeted knockout of selected activities.

Molecular biology tools may also include development of a genetic
25 manipulation screening system to determine equivalency in subsequent strains.

According to certain embodiments, molecular biology tools may include a recombinant *Sporidiobolus* strain that includes at least one expression cassette functional in *Sporidiobolus* that is stably integrated into its genome for the production of homologous or heterologous proteins. The expression cassette may include,
30 operably linked in the direction of transcription, at least one promoter functional in *Sporidiobolus*, at least one nucleic acid sequence encoding a homologous or heterologous protein and at least one terminator region functional in *Sporidiobolus*. The promoter may be a promoter of a *Sporidiobolus* gene, such as a URA3 promoter.

The terminator may be a terminator of a *Sporidiobolus* gene, such as a URA3 terminator. The nucleic acid sequence encoding a heterologous protein may encode at least one protein chosen from a fungal protein, a bacterial protein, a plant protein, an animal protein, and a protein from unidentified origin. The nucleic acid sequence
5 encoding a homologous protein may encode at least one protein chosen from Ku70 protein, Ura3 protein, eGFP protein, Malic Enzyme Isoform 1, and Malic Enzyme Isoform 2, from *Sporidiobolus*.

In certain embodiments, a method for producing a recombinant *Sporidiobolus* strain includes the steps of: a) generating protoplasts of *Sporidiobolus*; b)
10 transforming the protoplasts generated in a) with at least one vector including an expression cassette encoding a homologous or heterologous protein; c) selecting recombinant *Sporidiobolus* transformed in b) under specific culture conditions; and d) recovering recombinant *Sporidiobolus* strains containing the expression cassette of the vector stably integrated into their genome. The transformation in b) may be
15 conducted in the presence of PEG. Additionally or alternatively, the transformation in b) may be by electroporation. In certain embodiments, the protoplast may be generated by an enzyme mixture, such as an enzyme mixture that includes a chitinase. The chitinase may be a *Streptomyces* chitinase, more particularly a *Streptomyces griseus* chitinase. In certain embodiments, the enzyme mixture may include a lysing
20 enzyme, such as a *Trichoderma* lysing enzyme, or more particularly a *Trichoderma harzianum* lysing enzyme. In certain embodiments, the enzyme mixture may include a driselase, such as a *Basidiomycetes* driselase. The enzyme mixture may include a lysing enzyme, a chitinase and a driselase.

According to certain embodiments, molecular biology tools may include a
25 recombinant cell derived from a protoplast of the species *Sporidiobolus* that is genetically transformed with a nucleic acid sequence. The cell may be transformed with either a homologous or heterologous nucleic acid sequence. The nucleic acid sequence may include a marker gene. In certain embodiments, protoplast may be generated by an enzyme mixture. The enzyme mixture may include a chitinase, such
30 as a *Streptomyces* chitinase, or more particularly a *Streptomyces griseus* chitinase. In certain embodiments, the enzyme mixture may include a lysing enzyme, such as a *Trichoderma* lysing enzyme, or more particularly a *Trichoderma harzianum* lysing enzyme. In certain embodiments, the enzyme mixture may include a driselase, such

as a *Basidiomycetes* driselase. The enzyme mixture may include a lysing enzyme, a chitinase and a driselase.

In certain embodiments, a recombinant cell derived from a protoplast of the species *Sporidiobolus* that is genetically transformed with a nucleic acid sequence may be obtained by: a) isolating a cell suspension comprising *Sporidiobolus* cells; b) 5 treating the cell suspension to obtain protoplasts; c) transforming the protoplasts; and d) regenerating cells from the transformed protoplasts. In step d), the protoplasts may be transformed using electroporation. Additionally or alternatively, the protoplasts in step d) may be transformed using PEG. The cell may be transformed with either a 10 homologous nucleic acid sequence or a heterologous nucleic acid sequence.

According to certain embodiments, an expression vector construct for the production of a polypeptide in *Sporidiobolus* includes a nucleic acid sequence encoding said polypeptide, and nucleic acid sequences allowing for expression of the polypeptide in *Sporidiobolus*. The nucleic acid sequence encoding the polypeptide 15 may be either homologous or heterologous. In certain embodiments, the nucleic acid sequence encoding the polypeptide may confer resistance to an antibiotic. In certain embodiments, the nucleic acid encoding the polypeptide may confer resistance to phleomycin. The nucleic acid sequence may encode a *Streptoalloteichus hindustanus* *Sh ble* gene.

20 More particularly, the vector construct may include a nucleic acid sequence encoding a polypeptide, a promoter in reading frame with the coding sequence, and a terminator. In certain embodiments, the polypeptide may confer resistance to an antibiotic. In certain embodiments, the polypeptide may confer resistance to phleomycin. The polypeptide may have 95% sequence identity to SEQ ID NO: 1. 25 The nucleic acid sequence may include an intron. The intron may include a nucleic acid sequence that is either homologous to a nucleic acid sequence encoding the polypeptide or heterologous to a nucleic acid sequence encoding the polypeptide. In certain embodiments, the intron may include a URA3 ORF. More particularly, the vector construct may include a polynucleotide coding sequence under the control of a 30 URA3 promoter and/or a URA3 terminator. For example, the vector construct may include a URA3 promoter, a phleomycin resistance gene, and a URA 3 terminator. As another example, the vector construct may include a promoter, a heterologous nucleic acid sequence encoding a polypeptide wherein the DNA sequence includes an

intron, and a terminator. The molecular biology tools may also include a *Sporidiobolus* cell that includes the vector construct.

According to certain embodiments, the molecular biology tools include a *Sporidiobolus* host cell wherein the host cell has been modified to reduce or eliminate
5 expression of a functional gene product of a nucleic acid sequence encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO: 72 or 82.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate embodiments of the invention and, together with the
10 description, serve to explain the features, advantages, and principles of the invention. In the drawings:

FIG. 1 is a diagram illustrating genetic targeting using a single cross-over homologous recombination technique in Example 2.

FIG. 2 is a diagram illustrating genetic targeting using a split-marker
15 homologous recombination technique in Example 2.

FIG. 3 is a diagram further illustrating the split-marker approach in Example 2.

FIGS. 4 and 5 are diagrams illustrating the replacement of PDH ORF with URA3 cassette as a result of homologous recombination of split markers in Example
20 2.

FIG. 6 shows the PCR products in the split-marker approach in Example 2.

FIG. 7 is a diagram that illustrates further testing of homologous recombination in Example 2.

FIG. 8 shows a banding pattern of the PCR on agarose gel in Example 2.

25 FIG. 9 is a diagram illustrating Southern blotting performed in Example 2.

FIG. 10 is a Southern blot analysis of the testing illustrated in FIG. 9.

FIGS. 11 and 12 are diagrams illustrating additional Southern blotting performed in Example 2.

FIG. 13 is a Southern blot analysis of the testing illustrated in FIGS. 11 and
30 12.

FIG. 14 shows a comparison between the sequence of the gene ku70 within a specific organism in Example 2 and three other related Basidiomycetes.

FIG. 15 is a diagram depicting the structure of the MK29404 ku70 homolog in Example 2.

FIG. 16 is a diagram depicting the structures of two DNA fragments, 495x546 and 545x500, in Example 2.

FIG. 17 is a diagram depicting integration occurring at the ku70 locus via homologous recombination within the Ku70 UP and Ku70 DN regions in Example 2.

5 FIG. 18 is a diagram depicting one approach to studying the role of circular and fragmented pLP159 DNA and its effect on gene copy number in Example 3.

FIG. 19 is a Southern blot analysis of the testing illustrated in FIG. 18.

FIG. 20 is a Southern blot analysis of phleomycin resistance in Example 3.

10 FIG. 21 is a reverse transcription PCR analysis of URA5 in the eGFP region of pLP180 in Example 3.

FIG. 22 is a Western blot analysis of MK29404-DAM6 transformed with pLP180 in Example 3.

FIG. 23 is a DNA sequence of the reverse transcription PCR product of eGFP in Example 3.

15 FIG. 24 is a reverse transcription PCR analysis of URA5 in the eGFP region of pLP184 and pLP185 in Example 3.

FIG. 25 is a Western blot analysis of MK29404-DAM6 transformed with pLP184 and pLP185 in Example 3.

20 FIG. 26 is a PCR analysis of transformants T001, T001A, T002, T002A, T005 and T005A, to confirm the successful recombination of the gene of interest into the DAM6 genome in Example 4.

FIG. 27 is a Western blot analysis using a monoclonal antibody specific to V5 tags in Example 4.

25 FIG. 28 is a PCR analysis of transformants T003, T003A, T004, T004A, T007 and T007A, to confirm the successful recombination of the gene of interest into the DAM6 genome in Example 4.

FIG. 29 is another Western blot analysis using a monoclonal antibody specific to V5 tags in Example 4.

30 FIG. 30 shows plate pictures from the transformations described in Example 4(iv).

FIG. 31 is a PCR analysis of genomic DNA obtained from red colony transformants in Example 4(iv).

DETAILED DESCRIPTION

The content of the electronically submitted sequence listing filed with the application is incorporated herein by reference in its entirety.

The invention provides molecular biology tools including systems, methods,
5 and protocols for the metabolic engineering of *Urediniomycetes*, specifically recalcitrant Basidiomycete 'red' yeast, *Sporidiobolus pararoseus* MK 29404. Using these molecular biology tools, it is possible to successfully perform over-expression of native genes, integrate functions from other organisms, and perform targeted gene knock-outs. These tools enable the metabolic engineering of *Urediniomycetes* for
10 optimal yields and production of middle distillate lipids, resulting in improved fermentation and biodiesel production. Additionally, these tools provide techniques to metabolically engineer this prolific oleaginous microorganism for optimal production at large-scale.

The term "biodiesel," as used herein, refers to components or streams suitable
15 for direct use and/or blending into a diesel pool and/or a cetane supply derived from renewable sources. Suitable biodiesel molecules can include fatty acid esters. Biodiesel can be used in compression ignition engines, such as automotive diesel internal combustion engines, truck heavy duty diesel engines, and/or the like. In the alternative, the biodiesel can also be used in gas turbines, heaters, boilers, and/or the
20 like. According to certain embodiments, the biodiesel and/or biodiesel blends meet or comply with industrially accepted fuel standards, such as B5, B7, B10, B15, B20, B40, B60, B80, B99.9, B100, and/or the like.

The term "oleaginous," as used herein, refers to oil bearing, oil containing and/or producing oils, lipids, fats, and/or other oil-like substances. Oleaginous may
25 include organisms that produce at least about 20 percent by weight of oils, at least about 30 percent by weight of oils, at least about 40 percent by weight oils, at least about 50 percent by weight oils, at least about 60 percent by weight oils, at least about 70 percent by weight oils, at least about 80 percent by weight oils, at least about 90 percent by weight oils and/or the like of the total weight of the organism. Oleaginous
30 may refer to a microorganism during culturing, lipid accumulation, at harvest conditions, and/or the like.

Rather than starting with a genetically tractable microorganism, a microorganism for use in tools, methods, and protocols for developing a microbial biodiesel production system was selected based primarily on its inherent fermentative

capability, namely its yield and productivity, as well as its product profile, namely its middle distillate fatty acids. While its genetic architecture would normally deter attempts to genetically engineer the selected microorganism, the selected microorganism, unicellular Basidiomycete 'red' yeast, which is a member of the class
5 Urediniomycetes, can in fact be successfully engineered to over-express both native and foreign genes, as well as to specifically target genes of interest for modification or inactivation "knock out".

The 'red' yeast *Sporidiobolus pararoseus* is a member of the Basidiomycetes family, a unique group of fungi, which is known to be a prolific lipid producing strain
10 and capable of growth and production rates similar to those of prokaryotic systems. This eukaryote is a member of the white-rot family of fungi (Dikarya > Basidiomycota > Puccinomycotina > Microbotrymycetes > *Sporidiobolus*). While Basidiomycetes are exceedingly difficult to genetically manipulate, the tools and methods described herein make possible the metabolic engineering of these organisms
15 for optimal yields and production of middle distillate lipids, resulting in improved fermentation and biodiesel production. Although limited data is available, having only one publicly sequenced member of the group (i.e. *Sporobolomyces roseus* with a reduced 10 MB genome [compared to the 21.8MB genome of MK29404]), one atypical feature is that it is not uncommon for a 3kbp gene to have upwards of 10
20 introns. This is an attribute which has previously all but excluded members of this group as a viable platform upon which to develop metabolic engineering approaches.

According to certain embodiments, molecular biology tools may include a recombinant *Sporidiobolus* strain that includes at least one expression cassette functional in *Sporidiobolus* that is stably integrated into its genome for the production
25 of homologous or heterologous proteins. The expression cassette may include, operably linked in the direction of transcription, at least one promoter functional in *Sporidiobolus*, at least one nucleic acid sequence encoding a homologous or heterologous protein and at least one terminator region functional in *Sporidiobolus*. The promoter may be a promoter of a *Sporidiobolus* gene, such as a URA3 promoter.
30 The terminator may be a terminator of a *Sporidiobolus* gene, such as a URA3 terminator. The nucleic acid sequence encoding a heterologous protein may encode at least one protein chosen from a fungal protein, a bacterial protein, a plant protein, an animal protein, and a protein from unidentified origin. The nucleic acid sequence encoding a homologous protein may encode at least one protein chosen from Ku70

protein, Ura3 protein, eGFP protein, Malic Enzyme Isoform 1, and Malic Enzyme Isoform 2, from *Sporidiobolus*.

Example 4 below demonstrates various *Sporidiobolus* gene expressions, both homologous and heterologous, in accordance with the molecular biology tools.

5 Molecular biology tools may include both homologous over-expression as well as heterologous over-expression, with the incorporation of introns for successful heterologous gene over-expression. This example also demonstrates auxotrophic marker systems, such as URA3 or URA5-minus systems.

The methods herein may include DNA, RNA, and protein extraction. The
10 methods may also include protoplasting through enzyme activity, such as through cryopreservation of protoplasts, or through the development of improved protoplast formation cocktails. Additionally, molecular biology tools may include one or more chromosomal integration or transformation vectors.

For example, a method for producing a recombinant *Sporidiobolus* strain may
15 include the steps of: a) generating protoplasts of *Sporidiobolus*; b) transforming the protoplasts generated in a) with at least one vector including an expression cassette encoding a homologous or heterologous protein; c) selecting recombinant *Sporidiobolus* transformed in b) with culture conditions appropriate for biomass generation (cell division); and d) recovering recombinant *Sporidiobolus* containing
20 the expression cassette of the vector stably integrated into their genome. The transformation in b) may be in the presence of PEG. Additionally or alternatively, the transformation in b) may be by electroporation. In certain embodiments, the protoplast may be generated by an enzyme mixture, such as an enzyme mixture that includes a chitinase. The chitinase may be a *Streptomyces* chitinase, more particularly
25 a *Streptomyces griseus* chitinase. In certain embodiments, the enzyme mixture may include a lysing enzyme, such as a *Trichoderma* lysing enzyme, or more particularly a *Trichoderma harzianum* lysing enzyme. In certain embodiments, the enzyme mixture may include a driselase, such as a *Basidiomycetes* driselase, available from Sigma-Aldrich of St. Louis, MO. The enzyme mixture may include a lysing enzyme, a
30 chitinase and a driselase.

Example 1 below demonstrates various *Sporidiobolus* transformation methods. More particularly, as shown in Example 1, *Sporidiobolus* transformation methods may include genetic material integration via biolistics, electroporation, and/or PEG-mediated transformation. These transformation techniques may utilize URA3 or

URA5 transformation vectors. Additionally or alternatively, these transformation techniques may utilize dominant selectable marker vectors.

According to certain embodiments, molecular biology tools may include a recombinant cell derived from a protoplast of the species *Sporidiobolus* that is
5 genetically transformed with a nucleic acid sequence. The cell may be transformed with either a homologous or heterologous nucleic acid sequence. The nucleic acid sequence may include a marker gene. In certain embodiments, protoplast may be generated by an enzyme mixture. The enzyme mixture may include a chitinase, such as a *Streptomyces* chitinase, or more particularly a *Streptomyces griseus* chitinase. In
10 certain embodiments, the enzyme mixture may include a lysing enzyme, such as a *Trichoderma* lysing enzyme, or more particularly a *Trichoderma harzianum* lysing enzyme. In certain embodiments, the enzyme mixture may include a driselase, such as a *Basidiomycetes* driselase. The enzyme mixture may include a lysing enzyme, a chitinase and a driselase.

15 In certain embodiments, a recombinant cell derived from a protoplast of the species *Sporidiobolus* that is genetically transformed with a nucleic acid sequence may be obtained by: a) isolating a cell suspension comprising *Sporidiobolus* cells; b) treating the cell suspension to obtain protoplasts; c) transforming the protoplasts; and
20 d) regenerating cells from the transformed protoplasts. In step d), the protoplasts may be transformed using electroporation. Additionally or alternatively, the protoplasts in step d) may be transformed using PEG. The cell may be transformed with either a homologous nucleic acid sequence or a heterologous nucleic acid sequence.

According to certain embodiments, an expression vector construct for the production of a polypeptide in *Sporidiobolus* includes a nucleic acid sequence
25 encoding said polypeptide, and nucleic acid sequences allowing for expression of the polypeptide in *Sporidiobolus*. The nucleic acid sequence encoding the polypeptide may be either homologous or heterologous. In certain embodiments, the nucleic acid sequence encoding the polypeptide may confer resistance to an antibiotic. In certain embodiments, the nucleic acid encoding the polypeptide may confer resistance to
30 phleomycin. The nucleic acid sequence may encode a *Streptoalloteichus hindustanus* *Sh ble* gene.

More particularly, the vector construct may include a nucleic acid sequence encoding a polypeptide, a promoter in reading frame with the coding sequence, and a terminator. In certain embodiments, the polypeptide may confer resistance to an

antibiotic. In certain embodiments, the polypeptide may confer resistance to phleomycin. The polypeptide may have 95% sequence identity to SEQ ID NO: 1. The nucleic acid sequence may include an intron. The intron may include a nucleic acid sequence that is either homologous to a nucleic acid sequence encoding the polypeptide or heterologous to a nucleic acid sequence encoding the polypeptide. In certain embodiments, the intron may include a URA3 ORF. More particularly, the vector construct may include a polynucleotide coding sequence under the control of a URA3 promoter and/or a URA3 terminator. For example, the vector construct may include a URA3 promoter, a phleomycin resistance gene, and a URA 3 terminator. As another example, the vector construct may include a promoter, a heterologous nucleic acid sequence encoding a polypeptide wherein the DNA sequence includes an intron, and a terminator. Molecular biology tools may also include a *Sporidiobolus* cell that includes the vector construct.

According to certain embodiments, molecular biology tools includes a *Sporidiobolus* host cell wherein the host cell has been modified to reduce or eliminate expression of a functional gene product of a nucleic acid sequence encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO: 72 or 82.

Methods herein may further include gene targeting or heterologous recombination, either through the use of a 'split marker' system or using targeted knockout of selected activities. Example 2 demonstrates various methods of gene targeting and recombination.

The effect of media on genomic DNA preparation and analysis is also demonstrated in Example 1.

Molecular biology tools may also include development of a genetic manipulation screening system to determine equivalency in subsequent *Sporidiobolus pararoseus* strains. Example 2 demonstrates such a screening system.

Further, molecular biology tools may also include dominant selectable marker development, such as by providing a table of all markers. Example 3 describes the dominant selectable marker development and provides a resulting table.

Example 3 also shows the removal of a native splice site for transcription, and Example 5 provides an example of protoplast fusion.

According to certain embodiments, a method of extracting recombinant cells suitable in production of biofuels from an aqueous fermentation broth includes extracting recombinant cells from the aqueous fermentation broth, wherein the broth

contains oleaginous microorganisms or sugarcane or lignocellulosic feedstocks, or a combination of oleaginous microorganisms, sugarcane, and/or lignocellulosic feedstocks, leaving residual biomass solids of recombinant cells that expressed one or more of malic enzyme, acetyl-CoA synthetase, ATP citrate lyase, and phytoene dehydrogenase genes after oil extraction, and residual broth water. The method may also include using at least part, or all, of the residual broth water as imbibition water for washing a process feedstock to extract sugar. The aqueous fermentation broth may be pasteurized, such as by heating the aqueous fermentation broth to about 40°C to about 80°C for about 1 minute up to almost 3 hours. The aqueous fermentation broth may be thermally pre-treated by heating the broth at a temperature between about 90°C and about 150°C, or between about 100°C and about 150°C, or between about 110°C and about 150°C, or between about 120°C and about 150°C, or between about 130°C and about 150°C for about 30 minutes to about 18 hours, or more than 3 hours to about 18 hours, or more than 3 hours to about 8 hours. The aqueous fermentation broth may be stirred during the heating interval. An acid, a base, or both an acid and a base may be added to the aqueous fermentation broth. Additionally or alternatively, a salt may be added to the aqueous fermentation broth. The aqueous fermentation broth may be passed through a bead mill, a homogenizer, an orifice plate, a high-shear mixer, a press, an extruder, pressure disruption, wet milling, dry milling, or other shear or mechanical disruption device at least once, or at least twice, or more. The aqueous fermentation broth may be stirred in a vessel at about 70°C to about 100°C or under reflux for about 1 to about 60 hours. A salt, such as up to about 2% by weight of the salt, such as NaCl, KCl, K₂SO₄, or Na₂SO₄, may be added to the aqueous fermentation broth in the vessel or alternatively may be produced *in situ*, for example, by adding NaOH or KOH, plus H₂SO₄. By producing such salts exothermically, the heat generated can help to reduce steam requirements for pasteurization. In certain embodiments, a salt may be added prior to lysis or further downstream in the process. An acid or a base may be added to adjust a pH of the aqueous fermentation broth in the vessel to between about 3 and about 11. The recombinant cells may be separated from the aqueous fermentation broth through an appropriate solid-liquid-liquid separation scheme that may include one or more steps such as gravity separation, hydrocyclones, filters, and/or centrifuges, leaving biomass solids and residual broth water. The residual broth water can be used as imbibition water for washing the process feedstock to extract sugar. Additionally, the biomass

solids can be recycled with the residual broth water. Residual biomass from the recombinant cell may be either wet biomass, or the residual biomass may have been dried, such as via spray drying, drum drying, rotary vacuum filter, fluidized-bed drying, oven drying, freeze drying/lyophilization, pneumatic drying, or combinations thereof.

The recombinant cells may be converted into a biofuel through the use of hydrotreating or transesterification, for example.

According to certain embodiments, the invention may be directed to a manufacturing facility for producing biofuels. According to certain embodiments, the manufacturing facility may include a recombinant cells extraction unit. Additionally, the manufacturing facility may include a thermal pre-treatment unit. In certain embodiments, the manufacturing facility may include equipment that enables recycle of residual broth water.

According to certain embodiments, the invention may be directed to a renewable material or a biofuel, or both a renewable material and a biofuel, made according to any of the methods described herein.

According to certain embodiments, the methods described herein may result in an increase in the oil extraction yield of the microorganism. For example, the method may result in an increase in the oil extraction yield of the microorganism of at least about 10 weight percent. According to certain embodiments, the increase in oil extraction yield may be at least about 10 weight percent, at least about 15 weight percent, or at least about 20 weight percent.

EXAMPLES

Materials and Methods

Yeast Culture

Unless stated otherwise for a particular example, yeast transformants were grown on yeast peptone dextrose (YPD) media containing about 2% glucose (dextrose) at 27°C for 2 days (from patch plate(s) or cryovial) or overnight (thereafter) in 250 ml flat-bottom shake flasks set to about 200 rpm. A variety of dilutions for sub-cultures could then be performed which generated the desired optical densities for particular molecular biology procedures. For example (Table 1), following overnight growth cultures would reach a desired optical density before

being concentrated via centrifugation (4,000 rcf for 5 minutes), washed twice with sterile diH₂O and prepared for subsequent manipulation.

5

Table 1: Dilution to OD_{A600} correlations for various molecular biology approaches using *Sporidiobolus pararoseus*.

Procedure	Dilution (pre-culture: YPD)	OD _{A600}
Protoplast formation	1:2500	1 - 3
Genomic DNA isolation	1:250	4 - 8

'MPPE' Protoplasting Enzyme Mix

Unless stated otherwise for a particular example, the enzyme mixture used to generate *Sporidiobolus pararoseus* protoplasts for genetic manipulations, known as the 'MPPE' protoplasting enzyme mix, was prepared as follows - in 100 mL 0.7M NaCl mix 5 mg of chitinase from *Streptomyces griseus* (Sigma C6137), 0.5 g Lysing enzymes from *Trichoderma harzianum* (Sigma L1412) and 2.5 g driselase from *Basidiomycetes* sp. (Sigma D9515 or Sigma D8037). Stir for about 10 minutes to dissolve as well as possible, centrifuge at 1,500 rcf for 10 minutes and filter sterilize.

Protoplast/Polyethylene glycol (PEG)-mediated transformation

Unless stated otherwise for a particular example, yeast transformations were conducted using a polyethylene glycol (PEG) mediated approach, which first involves the generation of protoplasts, followed by PEG-mediate transformation. *For Protoplast generation:* Cell pellets were generated as described within the yeast culture method above. Once generated, this pellet was then resuspended in 25 mM β-mercaptoethanol/5 mM EDTA and gently mixed for 20 minutes at room temperature, before centrifugation as above. The resulting supernatant was discarded and the pellet resuspended in 25 mL 0.7 M NaCl, before again being centrifuged as above. Supernatant was again discarded and the pellet resuspended in 25 mL 'MPPE' protoplasting enzyme mix. The cell suspension was then incubated for 2 hours and agitated at 100 rpm at 30°C. Cellular morphology was then monitored via microscopy to determine degree of protoplasting, with incubation complete when single cells were observed. *For PEG-mediated transformation:* The cells were centrifuged as above and the supernatant was discarded. The cell pellet was washed with 1M sorbitol,

resuspended and centrifuged as before. This step was repeated. Supernatant was again discarded and the cell pellet was gently resuspended in 0.5 mL STC (1M sorbitol, 10mM Tris-Cl, pH 8.0, 25mM CaCl₂). 50 µL aliquots of cells were added to 1 mM aurintricarboxylic acid and 10mM β-mercaptoethanol and mixed by gently
5 pipetting. 5 µg of experimental plasmid DNA (in less than 10µL) was then added, mixed gently and further incubated on ice for 20 minutes. 500µL of PEG solution (40% PEG4000, 10mM Tris-Cl, pH 8.0, 25mM CaCl₂) was then added and incubated at room temperature for 5 minutes. A 1 mL aliquot of STC at 4°C was subsequently added and mixed by inversion. The entire mixture was transferred to a 25 mL shaker
10 flask containing 4.5 mL YNB (Difco) media with 0.8M sucrose and incubated at 27°C for 2 hours with shaking at 100 rpm. The cells were then concentrated by centrifuging as described previously, resuspended in 250 µL of YNB/0.8M sucrose media and plated on 2 YNB/0.8M sucrose plates. Plates were then incubated at 27°C until colonies appeared (i.e. about 3-5 days).

15 **Genomic DNA (gDNA) isolation method**

Unless stated otherwise, this method of isolating genomic DNA (gDNA) from all strains of *Sporidiobolus pararoseus* involved the use of a combination of enzymes, in addition to reagents available in a variety of commercially available DNA extraction kits (i.e. EpiCentre MasterPure Yeast DNA Purification Kit [Cat No.
20 MPY80010]), to digest the cell wall of *S. pararoseus* and significantly improved DNA extraction via mechanical and chemical means. Specifically, the method involved growing *S. pararoseus* on YPD media to generate biomass for this particular molecular biology approach. Subsequent to generation of 4-8 units of OD_{A600} cellular suspension as described previously. A concentrated pellet representing approximately
25 OD_{A600} units of cells was then processed via centrifugation, washed with sterile diH₂O, supernatant removed and resuspended again in sterile diH₂O. The resulting pellet was then resuspended in 25 mM β-mercaptoethanol/5 mM EDTA, gently mixed for 20 minutes at room temperature, and centrifuged as above. The pellet was then washed twice with diH₂O, a 0.7M NaCl and a 'MPPE' protoplasting enzyme mix
30 centrifugation / decanting / resuspension step. The cell suspension was then placed at 30°C with shaking at 100 rpm for 1.5 hours, before centrifuging as above. Next, the supernatant was discarded and the pellet resuspended in a commercially available yeast cell lysis solution (e.g. EpiCentre), incubated at 65°C again with agitation at 100

rpm for 1 hour and cooled on ice for 5 minutes. Subsequently, a commercially available protein precipitation reagent (e.g. EpiCentre) was added, vortexed for 10 seconds and centrifuged at 20,800 rcf for 10 minutes at room temperature. The resulting supernatant was transferred to a microcentrifuge tube, 500 μ L isopropanol was added, mixed by inversion, and centrifuged at 20,800 rcf for 15 minutes at 4°C. The supernatant was then discarded and the resulting pellet, which contained the desired gDNA, was washed with 500 μ L 70% ethanol and centrifuged at 20,800 rcf for 2 minutes at 4°C. The resulting supernatant was then discarded and the gDNA was air dried and then resuspended in 50 μ L EB buffer (10 mM TRIS-Cl, pH 8.5) and 0.5 μ L RNaseA (10 mg/mL). Samples were then stored at 4°C before use.

**Small-scale Genomic DNA preparation and PCR method
to confirm identity of transformants**

Unless stated otherwise, this method was used to analyze transformants post transformation in order to determine the presence and/or successful recombination of the gene of interest within the *Sporidiobolus pararoseus* genome. This method consisted of patching individual colony transformants onto YNB 2% dextrose media to confirm prototrophy and then growing these colonies on YPD plates for a further 2-5 days at 27°C. This step has been experimentally shown to be necessary for good quality genomic DNA. Next a small swab of biomass from the patch from YPD plate was picked with an inoculation loop and suspended in 500 μ L of lysis buffer (400 mM Tris-HCl, pH 7.5, 60 mM EDTA, pH 8.0, 150 mM NaCl, 1% SDS) in 2.0 mL microcentrifuge tube. Biomass in lysis buffer was vortexed and incubated for 30 minutes at room temperature. Following the incubation 150 μ L of 3M NaOAc (pH 5.2) was added to the lysate and vortexed briefly. Tubes were spun at 14,000 rpm for 1 min, after the spin supernatants were transferred to a new 1.7 mL microcentrifuge tube and 1 μ L of glycogen (20 mg/mL) was added to each tube. An equal volume of 100% isopropanol was then added to each tube and mixed with the supernatant by inversion. Tubes were spun again at 14,000 rpm for 10 min, supernatant discarded and the resulting gDNA pellet washed with 500 μ L of 70% isopropanol. Finally, the tubes were spun at 14,000 rpm for 1 min, supernatant was discarded and gDNA pellet was air dried and then resuspended in 25 μ L of EB (10 mM tris-HCl, pH 8.0). PCR was set up using 1 μ L of quick prep gDNA that resulted from the procedure described above as a template.

Sporidiobolus pararoseus strains used throughout this work

Table 2: List of *S. pararoeseus* strains cited within this application and used throughout

Name	Feature
MK29404 (WT)	Wild-type <i>S. pararoseus</i>
MK29404-Dry-1	Lower viscosity derivative of MK29404 WT strain
MK29404-DAM-6	MK29404 Uracil (Ura3) auxotroph
MK29404-DAM7	MK29404-Dry1 Uracil (Ura5) auxotroph
MK29404-AM-6	MK29404 Uracin (Ura5) auxotroph
MK29404-W3A	Spontaneous white mutant of DAM6

5 The *S. pararoeseus* strain MK29404(WT) was deposited as ATCC Deposit No. PTA-11616. The *S. pararoeseus* strain MK29404-Dry-1 was deposited as ATCC Deposit No. PTA-12513. The remainder of the MK29404 strains identified herein are single gene manipulations of either MK29404(WT) or MK29404-Dry-1, as described in the sequence files and examples.

10 PCR reaction parameters

Unless otherwise stated, the PCR conditions employed for genomic DNA template PCRs were as follows – 300 μ M dNTPs, 0.4 μ M each primer, 50 ng genomic DNA, 1 U KAPA® HiFi HotStart polymerase (KAPA Biosystems), and 1 $\times$ KAPA® HiFi buffer (KAPA Biosystems) in a 25 μ L total volume reactions. The PCR Protocol included the following steps: (1) 95° C. for 2 minutes; (2) 98° C. for 20 seconds; (3) 55° C. for 30 seconds; (4) 72° C. for 35 seconds; (5) repeat steps 2-4 for 25 cycles; (6) 72° C. for 3 minutes; and (7) hold at 4° C. Further, unless otherwise stated, the PCR conditions employed for plasmid DNA template PCRs were as follows – 500 μ M dNTPs, 0.4 μ M each primer, 10 ng plasmid DNA, 5 U Herculaase II Fusion polymerase (Agilent product no. 600675) and 1x Herculaase II Fusion Buffer for 50 μ L total volume plasmid DNA reactions. The corresponding PCR protocol included the following steps: (1) 95°C for 3 min; (2) 95°C for 20 sec; (3) 55°C for 20 sec; (4) 72°C for 1 min; (5) repeat steps 2-4 for 30 cycles; (6) 72°C for 5 min; and (7) hold at 4°C. Finally, unless otherwise stated, the PCR conditions employed for colony PCR were as follows – 12.5 μ l of 2X GoTaq Green Master Mix (Promega), 0.4 μ M of each primer, 1 μ l of quick prep. gDNA and water for 25 μ l total volume

final reactions. Cycling conditions were used as follows: (1) 95°C for 2 min; (2) 95°C for 30 sec; (3) 59°C for 30 sec; (4) 72°C for 1 min; (5) repeat steps 2-4 for 40 cycles; (6) 72°C for 5 min; and (7) hold at 4°C. Primer pairs used in the investigations listed here are shown in Table 3 below.

5

Primers used throughout this work

Table 3: List of forward and reverse primers used throughout this work; and including SEQID No., primer name and sequence from 5' to 3'.

SEQID No.	Primer Name	Sequence (5' → 3')
SEQID No. 001	prLP485 (Forward)	ATGCCCTCAATCACTCACCGTAC
SEQID No. 002	prLP486 (Reverse)	CTACTCCTTCAACCTCTGTTCG
SEQID No. 003	prLP487 (Forward)	GGATAAATTCCTGGACCCCA
SEQID No. 004	prLP488 (Reverse)	TCCCCTTTTIGGTCTGAGTCGTG
SEQID No. 005	prLP566	GCTAAGCTCACTTCGGCTGT
SEQID No. 006	prLP567	CGACAACCTTAATCTTGCTCTTCTG
SEQID No. 007	prDS387 (sense)	TCACAATCAATTGAAAGAGCGACCC
SEQID No. 008	prDS388 (antisense)	CAGAAACAAGAGCCGACAAGACGAC
SEQID No. 009	prDS389 (sense)	ATGGATCTCGTATCCTCGGTATCGG
SEQID No. 010	prDS390 (antisense)	CGGAACGAGCGAGGTACTCTTTGT
SEQID No. 011	prURA3 F (forward)	CACAGTTCGCAACGAACTTCTC
SEQID No. 012	prYB11 R (reverse)	GTCACCTTCGACGTTCTGGAAC
SEQID No. 013	prYB12 R (reverse)	GATGGACAAGGGATTCTTGAC
SEQID No. 014	prYB13 R (reverse)	GTCCAAGAGAGTCTCAGCGTGC
SEQID No. 015	prYB14 R (reverse)	CCTGTCAATTCAACTTCTCTCGAG
SEQID No. 016	prDS438 (sense)	TCACAATCAATTGAAAGAGCGACCC
SEQID No. 017	prDS440 (antisense)	CAGAAACAAGAGCCGACAAGACGAC
SEQID No. 018	prLP499 (forward)	ATGATTACGCCAAGCTCCTTCC
SEQID No. 019	prLP500 (reverse)	TCTCCCTTCTCTACCGACTCCA
SEQID No. 020	DHD532	CGCAACGAACTTCTCTCGC
SEQID No. 021	DHD550	CCTCCTCTTGTGTGGTTTCG
SEQID No. 022	DHD494	CAAACCTCTTCCACTGACGAG
SEQID No. 023	DHD552	GGTCGAGTCCTGTATAAGTG
SEQID No. 024	DHD518	GAGTACTGTCGAGCTGCTG
SEQID No. 025	DHD520	GGTGAAAGCAAATTAACAGAGG
SEQID No. 026	DHD508	GGTGCTCATAGTGACTATCG
SEQID No. 027	DHD498	GCAGTACCTTGAATAGGATCG
SEQID No. 028	DHD528	CCTCAATCACTCACCGTACC
SEQID No. 029	DHD529	GACGTGGACTTCTCCTTCTC
SEQID No. 030	DHD507	GGAAGAATGAGACTCACCTC
SEQID No. 031	DHD506	GAGGTGAGTCTCATTCTTCC
SEQID No. 032	DHD555	GTCAGTCCCTTCCGTAAACG
SEQID No. 033	DHD530	TGCTGACCCATTCCATCACC
SEQID No. 034	DHD531	CGATAGTCACTATGAGCACC
SEQID No. 035	DHD496	CCCTTAGAGTCTTGAGTTGC
SEQID No. 036	DHD509	CGATAGTCACTATGAGCACC
SEQID No. 037	prLP637	TCAACAACCTGCCACAATCA

SEQID No. 038	prLP638	CCCTTTTGGTCTGAGTCGT
SEQID No. 039	prLP631	GGTGTGTCCCTATCCTCGTAG
SEQID No. 040	prLP632	TAAAGTTCATCCATCCCAAGG
SEQID No. 041	prDS402	ACAGATCTTTAGAATGGGAGAACCC
SEQID No. 042	prDS403	TCTCTATCTATGGGAATGCGACG
SEQID No. 043	prDS406	TGAGGATCTTCTCAAATTAGGTGGG
SEQID No. 044	prDS407	TCTAAACCAACACCAGGGGTCA
SEQID No. 045	prDS398	GACGTGTCTACACTCGACTCGCA
SEQID No. 046	prDS399	CGGCGAGAATGAAGAGAATAGTAGA
SEQID No. 047	prDS418	AATCAAAAATACGTGATCCGGG
SEQID No. 048	prDS419	TTGGGTGGCTCCTTTCCTCG

Outgrowth, selection and propagation of transformed *S. pararoseus* strains

Unless otherwise stated, the following method was used prior to confirmation of presumed transformants from protoplast/PEG-mediated transformations in all cases. Specifically, after 3 - 5 days of outgrowth on YNB/0.8M sucrose/phleo150 (150 µg/ml) selection plates following transformation of Dry-1 with pLP159 (fragment), colonies were picked, patched onto fresh YNB/2% dextrose plates/phleo150, wrapped in parafilm and placed at 27°C for 2-5 days. With sterile loop, small amount of patches (once they were ~5 mm in size) were picked and inoculated into 250 mL flat-bottom flask containing 25 mL YPD medium and incubated at 27°C with shaking at 200 rpm for 48-72 hrs. Once cultures have reached late-log to stationary phase (OD₆₀₀ 12-25), they were sub-cultured at 1:500 into 250 mL flat-bottom flask containing 25 mL YPD medium and incubate at 27°C with shaking at 200 rpm overnight (~16 hrs.). For genomic DNA isolation, a cell culture with OD₆₀₀ of 5-7 OD units/ml was used to concentrate cells to approximately 25 OD₆₀₀ units by harvesting cells in 15 mL conical tube by spinning at 4,000 x g for 5 min and genomic DNA isolation/plasmid preparation was performed following protocol outlined below.

Southern blotting

Southern blotting was conducted using techniques known to those skilled in the art (see e.g. Sambrook et. al., Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press. 1989). Specifically, gels with gDNA separated by electrophoresis were depurinated via crosslinking in a UV box. Gel was denatured by submersion and ~30 min incubation in denaturing buffer (0.5M NaOH, 1.5M NaCl) with gentle rocking. Gel was washed 1-2 times with copious amounts of diH₂O. Gel was submersed in neutralization buffer (1.5M NaCl, 0.5M Tris-HCl, pH 7.0) and

incubated for ~30 min with gentle rocking. Vacuum blotter (Bio-Rad) was prepared by pre-warming the vacuum pump for ~10 min (following steps are adapted from the manufacturer's protocol). Nylon membrane was cut to size of the gel and pre-wet in diH₂O and then transferred in a container with 10X SSC (Sigma). A piece of
5 Whatmann paper slightly larger than the dimensions of the membrane was cut, pre-wet in 10X SSC and placed in the middle of vacuum stage. Wet Nylon membrane was placed atop of Whatmann paper, bubbles were rolled out. Pre-cut window gasket was laid over Nylon membrane followed by a gel on top, making sure that gel borders overlap the window gasket slightly. Reservoir O-ring seal was wetted with diH₂O,
10 sealing frame was placed on top of vacuum stage and locked in place. Vacuum source was turned on and the bleeder valve was adjusted to 5 in. Hg. Amount of 10X SSC sufficient to cover the top of the gel was poured on top, gel transfer was set for 90 minutes at constant pressure of 5 in Hg. Once the transfer was complete, nylon membrane was removed and soaked in 2X SSC for 5 minutes. Nylon membrane was
15 cross-linked in a UV box, and then air-dried between two pieces of Whatmann paper. Fragments for labeling were generated specific to each example and described below. Once generated, the probe was labeled using following approach. 100 ng of probe was denatured with 0.45 ng of diluted λ fragment in total volume of 20 μ l for 5 minutes in boiling water and cooled on ice. While on ice 2 μ l of 0.5 mM dATP, 2 μ l
20 of 0.5 mM dGTP, 2 μ l of 0.5 mM dTTP solutions, 15 μ l Random Primers Buffer Mix, 44 μ l of H₂O (to final volume of 50 μ l) and 5 μ l (50 μ Ci) [α -³²P]dCTP and 1 μ l Klenow Fragment were added. Reaction was incubated at room temperature for 1 hr and stopped by addition of 5 μ l of 0.5 M EDTA, pH 8.0. Dried nylon membrane with gDNA cross-linked to it was pre-wet in 6X SSC. Membrane was prehybridized with
25 40 mLs of pre-hybridization solution (0.5M Na₂HPO₄ pH 7, 1 mM EDTA, 7% SDS, 1% BSA) in a roll bottle at 65°C for 30 minutes. The probe was denatured by boiling it for 5 minutes, then the probe was added to hybridization solution (same as pre-hybridization solution). Membrane was hybridized with the probe overnight. The following day the membrane was washed with 40 mLs of pre-warmed wash solutions
30 1 (40 mM Na₂HPO₄ pH 7, 1 mM EDTA, 5% SDS, 0.5% BSA) for 1 minute. The membrane was washed with wash solution 2 (40 mM Na₂HPO₄ pH 7, 1 mM EDTA, 1% SDS) for 5 minutes, this step was repeated 3-4 times, membrane was removed from the roll bottle, wrapped in Saran Wrap and imaged. In some instances, southern hybridization blots were carried out by use of a radioactive-based system, whereby

the probe was labeled with ^{32}P -dCTP using the Random Labeling kit from Invitrogen (Catalog #18187013) and following the Invitrogen protocol. The hybridization protocol is from Church and Gilbert (G. M. Church and W. Gilbert (1984), Genomic sequencing, PNAS Vol. 81 (7):1991-5). Pre-hybridization of the membrane was
5 carried out at 65°C . for 1 hour in hybridization buffer (500 mM sodium phosphate buffer, pH 7.0, 1 mM EDTA, 7% SDS, 1% BSA). Hybridization was carried out at 65°C . for 18 hours in hybridization buffer containing the URA5 gene probe that had been heat-denatured for 5 min at 94°C . The membrane was then washed once for 5 minutes with 50 mL of Wash Solution 1 (40 mM sodium phosphate buffer, pH 7.0, 1
10 mM EDTA, 5% SDS, 0.5% BSA). The membrane was then washed 3 times with Wash Solution 2 (40 mM sodium phosphate buffer, pH 7.0, 1 mM EDTA, 1% SDS). The image was visualized on a Kodak phosphor screen.

Western Blotting

Unless otherwise stated, the following method was used for Western Blot
15 analysis. 25 mL YPD medium in a 250 mL flat-bottom shaker flask was inoculated with AM6 and/or DAM7 +pDS228 or +pDS229 transformants from a YNB/2% dextrose patch plate or parental strains, MK29404 (WT) and MK29404-Dry1 from a cryovial. Cultures were incubated for 2 days at 27°C with shaking at 200 rpm, sub-cultured 1:1500 into 25 mL YPD in 250 mL flat-bottom shaker flask and incubated at
20 27°C overnight with shaking at 200 rpm. Following this overnight growth, cultures optical densities at A_{600} were typically between 4-8. Based on the A_{600} measurement, 40 A_{600} units (calculated by: $40/A_{600}$ measurement) were harvested by centrifugation in a 50 mL conical tube at $\sim 4,000$ rcf for 5 minutes, washed once with 25 mL sterile dH_2O , resuspended and again centrifuged at $\sim 4,000$ rcf for 5 minutes. All traces of
25 supernatant were carefully removed from the resulting pellet. Cell pellet was then resuspended with 10 mL sterile dH_2O , cell suspension transferred into 15 mL conical tube, and centrifuged at $\sim 4,000$ rcf for 5 minutes. All traces of supernatant were carefully removed from the resulting pellet. Cell pellet was resuspended with 250 μL Y-PER buffer + protease inhibitors [per 10 mL: 10 mL Y-PER protein extraction
30 buffer (Thermo Scientific product no. 78990) + 1 Complete, mini, EDTA-free protease inhibitor cocktail tablet (Roche product no. 11836170001)], and the cell suspension was transferred into 1.7 mL screw-cap tubes containing 150 mg of 0.5 mm glass beads (Sigma G8772). Tubes were fixed horizontally on a vortex adapter (Ambion) and vortexed at high speed for 10 min. Cell debris and glass beads were

pelleted by spinning at ~14,000 rcf for 10 min. Total protein-containing supernatant was pipetted away from the resulting cell/glass bead pellet and transferred into a 1.7 mL centrifuge tube, then stored at -20°C until use. For denaturation, 10 µL of total protein from each sample was mixed with 10 µL 2X NuPAGE LDS Sample Buffer (Life Technologies product no. NP007) containing 100 mM DTT. Samples were heated at 70 °C for 5 min, then placed on ice. 10 µL of each denatured sample was then loaded on a 4-12% Bis-Tris polyacrylamide gel (Life Technologies product no. NP0336PK2) along with an unstained protein standard (Benchmark protein ladder; Life Technologies, cat. no. 10747-012). Samples were electrophoresed for 35 min at 200 V in NuPAGE MES SDS Running Buffer (Life Technologies product no. NP0002-02) using the Novex Xcell SureLock Mini-Cell system (Life Technologies product no. EI1001). After electrophoresis, protein was transferred from the polyacrylamide gel to PVDF membrane (Life Technologies LC2002) using NuPAGE transfer buffer (Life Technologies product no. NP0006) containing 20% methanol via the Novex XCell II Blot Module (Life Technologies product no. EI9051) for 1 hr at 30 V. PVDF membrane was removed from module and blocked with 25 mL 5% milk-TBST (20 mM TRIS-Cl, 140 mM NaCl, 0.1% Tween-20, pH 7.5; dried milk to 5% w/v) at room temperature with gentle rocking for 1 hr. Membrane was washed 3 times with TBST (see recipe above; without milk); 5 min per wash with gentle rocking at room temperature. Mouse monoclonal V5-AP conjugated antibody (Life Technologies product no. R962-25) at 1:2000 in 10 mL 1% Milk-TBST (20 mM TRIS-Cl, 140 mM NaCl, 0.1% Tween-20, pH 7.5; dried milk to 1% w/v) was added to membrane, and incubation with antibody was carried out at room temperature with gentle rocking for 2 hrs. Membranes were washed at room temperature 4 times with 20 mL TBST (see recipe above), 5 min per wash. Membranes were transferred into a new gel box containing 20 mL TBS (20 mM TRIS-Cl, 140 mM NaCl, pH 7.5) and washed at room temperature for 5 min with gentle rocking. Membranes were washed twice briefly with 10 mL substrate-free alkaline phosphate (AP) buffer (100 mM Tris-Cl, 100 mM NaCl, 5 mM MgCl₂; pH 9.5). 10 mL substrate-containing AP buffer [AP buffer – see recipe above plus 66 µL NBT (50 mg/mL), 33 µL BCIP (50 mg/mL)] was added to each membrane and the reaction allowed to proceed at room temperature until sufficient colorimetric signal was observed (~10 min).

**Effect of media on genomic DNA preparation and analysis
and development of a PCR-based, medium-throughput screening technique**

PCR amplification of genomic sequences is a useful tool for the analysis of chromosomal structure and DNA sequencing in fungi such as *Sporidiobolus*
5 *pararoseus* strain MK29404. In many instances, it is desirable to be able to analyze dozens or hundreds of fungal colonies for a particular genetic feature by PCR assays. For example, as described here in Examples 3a-d, in gene targeting experiments, it can be necessary to screen hundreds of transformants in order to identify the particular transformant(s) of interest. Depending on the phenotype caused by a given
10 targeted genetic manipulation, it may be necessary to screen transformants by a PCR assay to detect the transformants of interest. Therefore, it is desirable to develop simple, rapid and robust procedures for "medium-throughput" PCR-based assays of hundreds of colonies grown in petri plates on solid growth media. Such a PCR-based medium-throughput screening technique, that is suitable for the *S. pararoseus* strain,
15 has been developed.

Any population of several hundred MK29404 colonies, such as, for example, transformants arising from transformation of MK29404 by a DNA molecule (or molecules) that confers a selectable trait to the transformed MK29404 cell can be rapidly subjected to PCR analysis to identify a particular feature of the genomic
20 structure of the transformants as follows: Colonies from the population to be analyzed are patched onto agar plates containing YPD medium in an ordered array or grid. If desired, the same colony can be patched onto multiple plates containing different growth media, such as, for example, YNB + glucose, YNB + glucose + uracil, or YNB plus any desired nutritional supplement or antibiotic. Similarly, the
25 plate media might be YPD + any desired nutritional supplement or antibiotic, or any other medium of interest. Five hundred to 1,000 or more colonies can be patched and gridded in a day. These patch plates or grid plates are incubated under suitable conditions for growth, such as, for example, 27°C for 4 to 8 days. After patches of cell growth have reached a sufficient size, DNA templates for PCR assays can be
30 prepared from any of the plates.

PCR templates are prepared by picking a swab of cell mass off the YPD plate with a sterile toothpick and transferring the cells to a sterile 1.7 ml eppendorf tube containing 400 µl of lysis buffer [400 mM Tris-HCl (8.0), 60 mM EDTA (8.0), 150 mM NaCl, 1% SDS] and vortexing vigorously to suspend the cells in the lysis buffer.

Following incubation of the cell suspension at room temperature for approximately 20 – 30 minutes, 120 µl of 3M KOAc (4.8) is added and the sample is then mixed by vortexing and centrifuged at ~18,000 x g for 5 minutes at room temperature. A 400 µl aliquot of the supernatant is carefully removed by pipetting and added to a fresh
5 sterile 1.7 ml eppendorf tube containing 20 µl of glycogen at 1 mg/ml. 400 µl of IPA is then added and the solution vortexed briefly and centrifuged at ~18,000 x g for 10 minutes at room temperature. The resulting supernatant is discarded and the pellet washed with 180 µl 70% EtOH and centrifuged at ~18,000 x g for 2 minutes at room temperature. The resulting supernatant is discarded. The pellet is air-dried at room
10 temperature for 10-15 minutes and resuspended in 25 µl 10 mM Tris-HCl (8.0) 0.1 mM EDTA (8.0). Typically 1 µl of this DNA preparation is used in a 25 µl PCR reaction. Samples can be processed in sets of 12 or more and a total of 48 – 72, or more, samples can be processed readily in a single day by one individual.

PCR assays with these templates are typically performed using either of two
15 commercial DNA Polymerases, KAPA™2G Robust HotStart (“KAPA2G”) or KAPA HiFi™ HotStart (“KAPAHiFi”), according to the vendor (KAPABiosystems, Cape town, South Africa) protocols. Reactions with the KAPA2G enzyme are performed in a volume of 25 µl which contains: 12.3 µl H₂O, 5 µl 5X KAPA2G Buffer B with MgCl₂ (final concentration = 1.5 mM MgCl₂), 0.5 µl 10 mM dNTP mix (final
20 concentration = 200 µM each), 3 µl each oligonucleotide primer at 4 µM, 0.2 µl KAPA2G (1 U) and 1 µl DNA template (produced as described above). Reactions can be performed in a BIO-RAD C1000™ Thermal Cycler (Bio-Rad, Hercules CA) using the following parameters: 1 cycle @ 95°C for 5 minutes, 35 cycles {95°C for 30 seconds, X°C for 15 seconds, 72°C for Y seconds}, cool to 4°C and hold. X°C, the
25 annealing temperature, is chosen based on calculated T_m of the primers; Y seconds, the extension time, is chosen based on the length of the target amplicon.

Reactions with the KAPAHiFi enzyme are performed in a volume of 25 µl which contains: 13.75 µl H₂O, 5 µl 5X KAPAHiFi Fidelity Buffer (contains 10 mM MgCl₂, final concentration = 2 mM), 0.75 µl 10 mM dNTP mix (final concentration =
30 300 µM each), 2 µl each oligonucleotide primer at 4 µM, 0.5 µl KAPAHiFi (0.5 U) and 1 µl DNA template (produced as described above). Reactions can be performed in a BIO-RAD C1000™ Thermal Cycler using the following parameters: 1 cycle @ 95°C for 5 minutes; 35 cycles {98°C for 20 seconds, X°C for 15 seconds, 72°C for Y

seconds}; 1 cycle @ 72°C for Z minutes; cool to 4°C and hold. X°C, the annealing temperature, is chosen based on calculated T_m of the primers; Y seconds, the extension time, is chosen based on the length of the target amplicon; Z minutes, final extension time, is chosen based on the length of the target amplicon. The parameters of the PCR reactions used to assay these templates can be varied in ways that will be obvious to one skilled in the art. For example, a wide variety of thermostable DNA polymerases are commercially available and could be tested for suitability in these assays. Similarly, there are numerous commercially available thermocyclers which could also be tested for suitability in these assays. It will also be obvious to one skilled in the art that the specifics of time, temperature and number of cycles, for each of the steps in the PCR reaction can be varied for any specific application.

Aliquots (5 – 10 μ l) of the PCR reactions can be analyzed by agarose gel electrophoresis to screen for the presence of the expected reaction products. Surprisingly, we have found that the media upon which cells are grown significantly affects the performance of the DNA template in PCR assays. Optimal results in PCR assays have been obtained when templates are prepared from cells cultured on YPD-based media. These templates consistently yield robust positive signals (i.e. strong bands of expected sizes on agarose gels) in reactions with a variety of oligonucleotide primer pairs. In contrast, when templates are prepared from cells cultured on YNB-based media are used poor results (i.e. very faint bands of expected sizes detected on agarose gels) and/or negative results (i.e. no bands detected on agarose gels) are the outcome. The suitability of other media for cell growth prior to template preparation can readily be tested by following the above protocols.

25 Example 1: *Sporidiobolus pararoseus* transformation methods

(i) Generation of Uracil (URA3 and URA5) auxotrophic mutants

(a) The following example demonstrates that MK29404 is capable of generating uracil auxotrophs (URA3 and URA5) for use in a number of auxotrophic transformation methods. Namely, uracil auxotrophs were generated as follows: MK29404 was streaked on YPD plates (with 2 % glucose) for single colony isolation and incubated at 27°C for 3 days. A single colony was picked and inoculated into 2 mL of YPD media and incubated overnight at 27°C with shaking at 200 rpm. The culture was then concentrated via centrifugation at 2,000 rcf for 5 minutes, the supernatant removed and 3 mL sterile diH₂O added the cells to resuspend. The

resuspended cells were centrifuged as above and the resulting pelleted cells then resuspended in 2 mL YNB (Difco)/2% glucose, diluted 1:2 with YNB/2% glucose, and 100 μ L aliquots spread on 6 YNB/2% glucose/5-fluoroorotic acid/uracil (20mg/L) plates. Plates were incubated at 27°C until single colonies were visible (~5-7 days). Single, isolated, colonies were picked and patched onto YNB/2% glucose/5-FOA/uracil plates as well as YNB/2% glucose/5-FOA plates and incubated at 27°C for about 3 days. Colonies that grew in the presence of uracil but not in its absence were then chosen for genetic characterization. Specifically, genomic DNA isolation and PCR reactions were conducted as described above, in order to identify mutations of URA3 and URA5 loci. PCR primers were designed to amplify the MK29404 URA3 (SEQID No. 001 and SEQID No. 002) and URA5 loci (SEQID No. 003 and SEQID No. 004). Sequencing involved cloning the respective PCR products into the vector pJET1.2/blunt (Fermentas) according to the manufacturer's instructions, and the insert sequence using standard methods known in the art. Results indicated that all mutations in MK29404 were mapped to the URA5 gene; no mutations observed in URA3. The identified MK29404 uracil mutants are detailed below.

Table 4: MK29404 Uracil Mutants

<u>Strain name</u>	<u>Mutated uracil gene</u>	<u>Mutation(s)</u>
AM3	URA5	Δ A792
AM5	URA5	$\downarrow$ A793
AM6	URA5	C919T (STOP)
AM7	URA5	T388C, Δ C392
AM9	URA5	T538G, T545G
AM10	URA5	T416A
AM11	URA5	Δ T425

(b) The following example further demonstrates that uracil auxotrophs from the MK29404 sub-isolate, MK29404-Dry-1, can also be developed and characterized using the method described above. The identified MK29404-Dry-1 auxotrophic mutants are detailed below.

Table 5: MK29404-Dry-1 Auxotrophic Mutants

Strain name	Mutated uracil gene	Mutation
DAM2	<i>URA3</i>	Δ T140
DAM5	<i>URA3</i>	C161A
DAM6	<i>URA3</i>	A148G, Δ 153-419
DAM10	<i>URA3</i>	Δ 135-601
DAM11	<i>URA3</i>	Δ 154-246
DAM1	<i>URA5</i>	T536C
DAM7	<i>URA5</i>	$\downarrow$ A793

(ii) Transformation using electroporation (with enzyme pretreatment)

5 The following example demonstrates that MK29404 can be transformed through the use of electroporation. Electroporation of MK29404 was used to evaluate whether prototrophy could be restored in MK29404 AM-6, a uracil auxotroph. Specifically, MK29404 AM-6, lacking functional *URA5*, was transformed via electroporation with enzyme pretreatment (see below) with pLP149 containing the

10 *URA5* gene. Cells were grown in 25 mL YPD medium on a shaker at 200 rpm for 2 days at 27° C. The cells were diluted at 1:2500 into 2 volumes of 50 mL YPD medium and grown overnight (16-18 hours), attempting to reach early to mid-log phase growth (OD600 of 1-3). The cells were centrifuged in a 50 mL conical tube for 5 minutes at about 1,500 rcf. The supernatant was removed and the second culture

15 was poured on top of the cell pellet. The tube was centrifuged as described previously. The supernatant was removed and the cell pellet was washed with 25 mL sterile deionized water. The cells were centrifuged as described previously and the supernatant was removed. The cells were washed with 25 mL 25mM mM β -mercaptoethanol/5 mM EDTA and placed on a gyrator for 20 minutes at room

20 temperature, and centrifuged as above. The supernatant was discarded and the cells were washed with sterile water and centrifuged as previously described. The supernatant was discarded, the cells were washed with 25 mL 0.7M NaCl and centrifuged again as above. The resulting supernatant was discarded and the pellet was either resuspended in 25 mL 'MPPE' protoplasting enzyme mix (in the case of

25 enzyme assisted transformations) or directly resuspended in STC media and the remaining method carried out as described below. The cell suspension was then transferred to a 250 mL round-bottom flask and placed at 30°C with shaking at 100

rpm for 2 hours. Cells were monitored under the microscope to determine the degree of protoplasting, with single cells desired. The cells were centrifuged as above and the supernatant was discarded. The cells were resuspended in 25 mL ice cold STC (1M sorbitol, 10mM Tris-Cl, pH 8.0, 25mM CaCl₂) and centrifuged as before. The
5 supernatant was discarded, the cells were resuspended in 25 mL ice cold 1M sorbitol and centrifuged as above. The supernatant was removed and the cells were washed again with 25 mL ice cold 1M sorbitol, centrifuged as before and the supernatant removed. The cells were gently resuspended in 0.6 mL 1M sorbitol using a wide-bore pipet tip. 100 µL of cells were aliquoted into a prechilled electro-cuvette (Gene
10 Pulser® cuvette—0.2 cm gap, Bio-Rad, Hercules, Calif.). Six µg of DNA (in less than or equal to a 10 µL volume) was added to the cuvette, mixed gently with a pipette tip, and placed on ice for 10 minutes. Cells were electroporated at 200 ohms (resistance), 25 µF (capacitance), and 1500 volts. 0.5 mL of YNB (Difco)/0.8M sucrose media was added immediately to the cuvette. The cells were then transferred to 4.5 mL of
15 YNB/0.8M sucrose media in a 25 mL shaker flask and incubated for 3 hours at 27°C. and 100 rpm on a shaker. The cells were centrifuged for 5 min at about 1,500 rcf in round bottom tubes. The supernatant was removed and the cell pellet was resuspended in 0.25 mL of YNB/0.8M sucrose media. Cells were plated onto 2 YNB/0.8M sucrose plates and incubated at 27°C until colonies appeared, 5-7 days. This protocol typically
20 results in the generation of 2-4 prototrophic strains per electroporation. Colonies that appeared to be prototrophic were then patched onto fresh YNB/0.8M sucrose plates to confirm prototrophy. PCR was then used to confirm the presence of plasmid sequences in these putatively transformed strains. Briefly, the quick prep and colony PCR methods were used as described previously. To detect the presence of pLP149
25 (SEQID No. 50) sequences in the MK29404 AM-6 gDNA, primers SEQID No. 018 and SEQID No. 019 were used; prLP499 anneals to the Blue Heron vector sequence and prLP500 anneals to the URA5 sequence in plasmid pLP149. These primers amplify a 912-bp DNA fragment. PCR products were analyzed by standard agarose gel electrophoresis, followed by staining with SYBR® Safe DNA gel stain (Life
30 Technologies). The results of these analyses confirm that all of the strains selected under these conditions are true transformants that contain plasmid DNA. No PCR products were generated using template DNA from control MK29404 AM-6 cells that had not been electroporated with the transformation plasmid.

(iii) Transformation using biolistics

The following example demonstrates that MK29404 can be transformed through the use of biolistics. Biolistics, also known as particle bombardment, of MK29404 was used to evaluate whether prototrophy could be restored in MK29404 AM-6. Genetic transformation of MK29404 cells was performed by particle bombardment (J.C. Sanford et. al (1993) Optimizing the biolistic process for different biological applications, Meth. Enzymol. Vol. 217:483-509) using a Bio-Rad Biolistic PDS-1000/He Particle Delivery System (Bio-Rad Laboratories, Hercules, Calif.). MK29404 AM-6 cells were grown in liquid YPD medium at 27°C to an optical density of 8-9 at A600. The culture was collected by centrifugation at 1,500 rcf for 5 minutes. The pellet was washed with 25 mL sterile water and centrifuged as described previously. The washing of the cell pellet with water was repeated 2 more times. The cell pellet was then resuspended with Na₂SO₄ to bring the OD₆₀₀ of the culture to 30 units. A volume of 150 µL of the resuspended cells was then spread in a 5 to 6 cm circle onto a Petri plate containing agar-solidified YNB medium and allowed to sit for 30 to 60 minutes so that the excess water could be absorbed into the solid medium; this is referred to as the target plate.

A 30 mg aliquot of tungsten M10 microcarriers (0.7µm diameter, available from Bio-Rad Laboratories, Inc., Hercules, Calif.) was coated with 5 µg of transformation plasmid DNA (i.e. plasmid pLP149 (SEQID No. 50)) as per the manufacturer's instructions (Biolistic® PDS-1000/He Particle Delivery System Instruction Manual; Bio-Rad Laboratories, Hercules, Calif.). The cells were bombarded with the DNA-coated tungsten microcarriers using the following conditions: 1100 psi burst disk, chamber vacuum of 25" Hg, microcarrier launch assembly placed on the top shelf and the target plates placed on the middle shelf, giving a burst disk-to-stopping screen distance of 1.5-2 cm and a stopping screen-to-target distance of approximately 7 cm. Complementation of the auxotrophic phenotype was expected and so the bombarded plates were incubated at 27°C until colonies appeared, 5-7 days. This protocol typically results in the generation of 200-500 prototrophic strains per bombardment. Colonies that appeared to be prototrophic were then patched onto fresh YNB plates to confirm prototrophy. PCR was then used to confirm the presence of plasmid sequences in these putatively transformed strains. Briefly, the quick prep and colony PCR methods were used as described previously. To detect the presence of pLP149 (SEQID No. 50) sequences in the MK29404 AM-6

gDNA, primers SEQID No. 018 and SEQID No. 019 were used; prLP499 anneals to the Blue Heron vector sequence and prLP500 anneals to the URA5 sequence in plasmid pLP149. These primers amplify a 912-bp DNA fragment. PCR products were analyzed by standard agarose gel electrophoresis, followed by staining with SYBR®
 5 Safe DNA gel stain (Life Technologies). The results of these analyses confirm that the vast majority of strains selected under these conditions are true transformants that contain plasmid DNA. No PCR products were generated using template DNA from control MK29404 AM-6 cells that had not been electroporated with the transformation plasmid. To further confirm these results, Southern hybridization blots
 10 were conducted, as described previously, using DNA isolated from parental MK29404 cells and several putative transformants, and digested with BamHI and EcoRI restriction enzymes, in order to confirm the presence of transformation vector DNA sequences within the transformed cells. The results were as follows: The BamHI/EcoRI-digested DNA from transformed MK29404 AM-6 cells contained a
 15 2.1-kb DNA fragment that hybridized to the URA5 gene probe and confirmed that the pLP149 plasmid was integrated into the MK29404 genome, thus rescuing the uracil auxotroph. The non-recombinant MK29404 cells contained an 8.1-kb DNA fragment, which is consistent with the wild type sequence.

Example 2: Targeted gene disruption (knock-out) in *Sporidiobolus pararoseus*

20 (i) Demonstration of genetic targeting using single cross-over homologous recombination technique

Plasmid pDS220 was constructed and contains a partial URA5 locus from MK29404. The URA5 portion of pDS220 contains the first intron within the URA5, the second exon and includes the region downstream of the locus. A XbaI restriction
 25 site were then added to the 3' end of URA5 locus, allowing for differentiation between introduced URA5 (via pDS220 plasmid) and native URA5. A diagram illustrating this genetic targeting is provided in FIG. 1.

Standard PEG-mediated transformation was then carried out on MK29404-AM6 using pLP149 and pDS220; with colonies from AM6+pDS220 picked and
 30 patched to confirm prototrophy on YNB/2% glucose medium. Genomic DNA was then prepared from 10 AM6+pDS220 transformants, following standard gDNA isolation protocol described previously. Southern blotting was then carried out using an XbaI-digested AM6+pDS220 transformant gDNA as well as AM6 parental gDNA, using a URA5 probe. Results show expected size URA5-hybridized band at 4,885 bp

for AM6 control lane; with all AM6+pDS220 transformants display the same 4,885 bp URA5-hybridized band. All AM6+pDS220 transformants analyzed contain an unmodified URA5 allele representing the mutant *ura5* within the AM6 background. The restoration of URA pathway with the partial URA5 construct and genetic analysis of the locus demonstrates successful homologous recombination and gene targeting.

(ii) Demonstration of genetic targeting using split-marker homologous recombination technique – knocking out the Phytoene dehydrogenase gene

The following example demonstrates successful homologous recombination technique using a split-marker approach. Phytoene dehydrogenase is an enzyme that catalyzes conversion of phytoene (colorless carotenoid precursor) to lycopene (first colored carotenoid) in carotenoid synthesis pathway in MK29404. Successful knock-out of Phytoene Dehydrogenase will result in colorless colonies in otherwise red-pigmented organism and will serve as a primary screening tool in selection of transformants. Briefly, the vector pDS233 (SEQID 81) was constructed by having 2 kb of chromosomal DNA sequence upstream of “ATG” “start” codon of Phytoene Dehydrogenase (SEQID 82), and 2 kb downstream of “TGA” “stop” codon of the gene synthesized by an external vendor with an additional “polylinker” sequence in between encoding two restriction enzyme sites (PvuII and BamHI). URA3 cassette was then subcloned off of pLP148 using ScaI and BamHI restriction sites generating final vector pDS233. Notably, the following vector, pDS233 (SEQID 81), was constructed for this experiment, and included the features listed in Table 6 and shown in FIG. 2.

Table 6: Features of the Vector pDS233 (SEQID 81)

Feature	Base pair coordinates
PDH Upstream homology	17-2022
MK29404 URA3 cassette	2023-3585
PDH Downstream homology	3586-5599
Kanamycin resistance gene	6047-6844
pUC origin of replication	6893-7696

The “split marker” approach was used to transform MK29404-DAM6 strain. Briefly, two fragments approximately 3kb in size with overlapping region of homology (~500 bp) were generated by PCR, with the two resulting PCR products (from primer pairs prDS402+prDS407 [SEQID No. 041 + SEQID No. 044] and prDS406+prDS403 [SEQID No. 043 + SEQID No. 042]) were gel purified, and used either both together or each separately for transformation of MK29404-DAM6 using PEG-mediated transformation method outlined previously (PMX21). A diagram

illustrating this approach is provided in FIG 3. Corresponding data is provided in Table 7.

Table 7: Data Corresponding to the “Split Marker” Approach

PCR product used:	# of Transformants	# White colonies/% White
No DNA control	0	N/A
UP PCR fragment 1	0	N/A
DOWN PCR fragment 2	0	N/A
UP+DOWN fragments 1+2	178	2/~1%

5 The resulting two white colonies and 8 red colonies were re-patched on YNB/2% glucose to confirm restored prototrophy, white colonies maintained their phenotype. Following several days of growth on YNB/2% glucose plates gDNA was isolated from PMX21-1, -4, -7, and -10 transformants following quick gDNA prep method. Replacement of PDH ORF with URA3 cassette as a result of homologous recombination of split markers was tested by PCR.

Two primers (SEQID No. 045 and SEQID No. 046) flanking the URA3 or PDH ORF were used to amplify DNA in between. In case of wild-type locus the resulting product will be 2,807 bp and in case of recombination the product will be 2,046 bp. FIG. 4 and FIG. 5 provide diagrams illustrating these results.

15 As evidenced by PCR, shown in FIG. 6, PMX21-1 and PMX21-4 both have PCR product of size indicative of homologous recombination, but not PMX21-7 or -10. To further test homologous recombination and successful integration of both UP and DOWN homology fragments, a series of PCR reactions were carried out using primers internal to URA3 and external to regions of PDH locus homology (primer pairs prDS418+prDS406 [SEQID No. 047+SEQID No. 043] and prDS407+prDS419 [SEQID No. 044+SEQID No. 048]). The diagram in FIG. 7 illustrates this further testing.

25 Banding pattern of the PCR on agarose gel, shown in FIG. 8, further confirms that both PMX21-1 and PMX21-4 contain the unique DNA arrangement that could have only resulted from homologous recombination (both lanes have either expected 3,310 bp product for Upstream region or 3,118 bp product for Downstream region), but this banding pattern is not observed in PMX21-7 or -10 transformants, suggesting that split marker recombined and inserted in another location without disrupting Phytoene Dehydrogenase locus.

30 Furthermore, Southern blotting has been done with a probe specific to Phytoene Dehydrogenase ORF. Here, genomic DNA has been extracted from DAM6,

PMX21-1, -4, -7, and -10 and digested with XbaI, run on an agarose gel, hybridized to a nylon membrane and membrane was then incubated with a probe specific to PDH ORF only. A diagram illustrating this example is provided in FIG. 9. Both PMX21-1 and -4 have no band in their respective lanes, as shown in FIG. 10, suggesting that
5 PDH ORF has been replaced and is no longer a part of the genome of these transformants.

Another Southern blot was carried out to test for URA3 locus presence. In addition to having native URA3 gene (with a deletion that creates auxotrophy) these transformants will have full-length URA3 gene that was a result of homologous
10 recombination and allowed the transformants to grow on minimal medium (YNB/2% glucose) without supplementation. Diagrams illustrating this example are provided in FIG. 11 and FIG. 12.

The genomic DNA was isolated and digested with either XbaI or SacI, run on an agarose gel, transferred to a nylon membrane and hybridized with URA3 locus probe. The Southern blot analysis of this example is provided in FIG. 13. URA3
15 locus that is a part of "native" DAM6 will have 3,358 bp band as a result of SacI digestion, while restored URA3 in PDH context will have a band of 8,872 bp, which is exactly what is observed for PMX21-1. PMX21-4 has a higher molecular weight band above 8,872 bp size, further experiments need to be carried out to determine the
20 exact identity of that fragment. XbaI digestion will either produce a 9,256bp product due to URA3 native to DAM6 or a 6,222 bp product due to URA3 in PDH context. Similarly, the predicted banding pattern is observed in PMX21-1, but PMX21-4 has an additional lower molecular weight band that requires further studies.

To further confirm knock-out of Phytoene Dehydrogenase, extraction of
25 carotenoids from biomass of the white transformants was carried out. The red oil extracted from the biomass of *Sporidiobolus pararoseus* is known to contain 16-184 ppm of the carotenoids torularhodin, torulene, gamma-carotene and beta-carotene. Freeze-dried samples (10 g) of Dry1, DAM6, PMX21-1 and PMX21-4 biomass were submitted for determination of carotenoids and phytoene. Samples were extracted
30 with chloroform/methanol with a blender. Extracts were filtered and evaporated to dryness, producing crude oil. Phytoene and carotenoids were isolated from crude oil in duplicate by solid phase extraction on silica gel and measured by reversed-phase HPLC with diode array detection; average results were reported. Five-point calibration curves were used to quantitate each analyte at the maximum wavelength

under the conditions of the assay. Analytical standards for phytoene, gamma-carotene, beta-carotene and lycopene were obtained from DSM or Carotenature, while standards for torularhodin and torulene were isolated in-house from extraction of *S. pararoseus* strains (100 g, Lot NBS30 0128M13) and further purified by flash column chromatography and semi-preparative HPLC. The concentrations of analytical standards used for calibration were determined spectrophotometrically in hexane by the following equation:

$$\text{Concentration (mg/mL)} = \frac{(\text{Absorbance}) \times 1000}{A^{1\%} \times 10}$$

Absorbance was measured using a bench top UV/VIS spectrophotometer; published values of specific absorption coefficients ($A^{1\%}$) for each analyte were used. The correlation coefficient of the calibration curve for torularhodin was less than 0.990. Therefore, the response factor of torulene was used to measure torularhodin in all samples. These carotenoids have similar chromophores, and the approximation was not expected to introduce significant error in the measurement of torularhodin.

Table 8: Carotenoids and phytoene measured in parent strain and mutant biomass samples of *S. pararoseus*. Results are expressed in ppm ($\mu\text{g/g}$).

Analyte	21-1	21-4	Dry1	DAM6
Torularhodin	ND	ND	18	21
Lycopene	ND	ND	9	10
Torulene	ND	ND	75	79
Gamma-Carotene	ND	ND	34	30
Beta-Carotene	ND	ND	38	32
Phytoene	1565	1400	127	141
Total Red Carotenoids	ND	ND	174	171

ND: Not detected

While phytoene was detected in all biomass samples, levels in PDH-KO strains were substantially greater than those measured in carotenoid-enriched yeasts. This observation is expected, since phytoene is the precursor to gamma-carotene, beta-carotene, torulene and torularhodin in the biosynthetic pathway proposed for carotenoids and would be consumed during conversion to these carotenoids. In addition to the major carotenoids previously identified in red yeasts, lycopene can also be detected in extracts of *S. pararoseus* biomass. Identification of lycopene, gamma-carotene, beta-carotene and phytoene was based on retention time matching with authentic standards and spectral characteristics of each analyte, including the

wavelength at the maximum absorbance and shape of the spectrum in comparison to published results. The identities of torulene and torularhodin are assumed, based on literature reports and the elution characteristics of these compounds under the conditions of the HPLC assay.

5 Biomass samples 21-1 and 21-4 were white solids in contrast to the cherry-pink precursors DAM6 and Dry1. Moreover, the solid phase extracts of the sister colonies were colorless, whereas those of DAM6 and Dry1 were dark red/orange. Colorless/white products are not expected to contain measurable levels of orange or red carotenoids, since the chromophores impart detectable color to products in the
10 lower ppm range. Results of both visual inspection and HPLC analysis demonstrate that the carotenoid pathway was successfully knocked out in the sister colonies 21-1 and 21-4. The amount of total red carotenoids measured in DAM6 and Dry1 is consistent with levels previously reported in red yeasts. [References used in this method: (P. Buzzini et al. (2007) Carotenoid profiles of yeasts belonging to the genera
15 *Rhodotorula*, *Rhodospiridium*, *Sporobolomyces*, and *Sporidiobolus*, Can. J. Microbiol. Vol. 53:1024-1031; Microbial Carotenoids from Fungi: Methods and Protocols. Edited by Jose Luis Barredo (2012), Chapter 19: Production of torularhodin, torulene and beta-carotene by *Rhodotorula* yeasts, M. Moline et al, pg 275-283; Carotenoids, Volume 1A: Isolation and Analysis. Edited by G. Britton et al. (1995), Chapter 5: Isolation and Analysis. K. Scheidt and S. Liaaen-Jensen, pg
20 104; Carotenoids. Volume 1B: Spectroscopy. Edited by G. Britton et al. (1995) Chapter 2: UV/Visible Spectroscopy. G. Britton, pg 57-61.]

(iii) Demonstration of genetic targeting using "split-marker" homologous recombination technique - modifications to the Non-homologous end-joining machinery (knock-out of ku70)
25

Targeted gene deletion, inactivation or replacement, via homologous recombination, can be a useful tool for genetic manipulation of fungi (Goswami, R.S., 2012. Methods in Molecular Biology 835: 255-269; Krappmann, S., 2007. Fungal Biology Reviews 21: 25-29). Such targeted genetic manipulations typically involve
30 transformation of an organism with a DNA molecule that has homology to one or more segments of the genome of the organism to be transformed, and the maintenance of which can be selected for in the particular strain that is being transformed. Linear DNA molecules in which a selectable element is flanked by segments of DNA having homology to the chromosome of the recipient organism are often used in such

targeted genetic manipulations. Typically, the linear DNA molecule is not capable of being replicated in the recipient cell and therefore is not maintained within the cell unless it is incorporated into the genome. Such linear DNA molecules, when introduced into a recipient cell, can be incorporated into the genome by two mechanisms: homologous recombination (HR) and non-homologous end joining (NHEJ). HR requires stretches of DNA homology and results in targeted insertion of the linear DNA molecule in the corresponding region of chromosomal homology. In contrast NHEJ, also sometimes referred to as “illegitimate recombination” or “ectopic integration”, does not require homology for insertion events. In many fungi, including *Sporidiobolus pararoseus* strain MK29404, NHEJ occurs far more frequently than HR. As a result, the efficiencies of gene targeting are very low. For example, in gene targeting experiments described above, NHEJ appears to be approximately 20 to 100-fold more frequent than HR. Depending on the phenotype caused by a given targeted genetic manipulation, it can be difficult, tedious, and time-consuming to identify the infrequent transformants that have arisen via HR amongst the vast majority of transformants that have arisen from NHEJ.

The biochemical complex responsible for NHEJ is known to be distinct from that which is responsible for HR. Certain proteins, such as the Ku70 protein, are known to be involved in NHEJ within fungi and in several instances (e.g., *Neurospora crassa*, *Aspergillus fumigatus*, *Aspergillus oryzae* and *Sordaria macrospora*) it has been shown that mutations in the ku70 gene which eliminate Ku70 protein activity result in a significant increase in the relative frequency of HR vs. NHEJ (Krappmann, S., 2007. Fungal Biology Reviews 21: 25-29). Therefore in order to facilitate targeted genetic manipulations of MK29404 and derivatives thereof, a deletion mutation that deletes ~50% of the ku70 gene was constructed as detailed in FIG. 14. Specifically, a ku70 homolog was identified in the genome of *Sporidiobolus pararoseus* strain MK29404 by homology to known ku70 genes of other fungi such as *Cryptococcus neoformans*, *Coprinus cinerea* and *Ustilago maydis*. FIG. 14 shows a comparison between the sequence of the gene ku70 within the organism of interest and three other related Basidiomycetes. Shading highlights the degree to which amino acids are conserved between all four strains, with the lightest shading indicating amino acids that are strongly conserved, medium shading indicating amino acids that are conserved in the majority of cases, and the darkest shading indicating amino acids that

are related. The sequence of the region of the MK29404 ku70 homolog is given in SEQ ID No 72 and the structure of the gene is depicted in FIG. 15.

A 1729 bp deletion mutation that eliminates ~50% of the MK29404 ku70 gene was constructed using a “split marker” selection procedure (Goswami, R.S., Methods
5 in Molecular Biology, 835, pp. 255-269. 2012) to transform the ura3 mutant DAM-6 to Ura⁺ by inserting a copy of the wild-type MK29404 ura3 gene into the ku70 locus and, as a consequence of the insertion, deleting over 50% of the ku70 coding sequence. The ura3 mutant strain DAM-6 contains a 267 bp deletion within the ura3
10 gene and this deletion renders the strain auxotrophic, i.e., unable to grow on minimal media, such as YNB + sucrose or YNB + glucose unless supplemented with uracil. In the split marker approach, DAM-6 is simultaneously transformed with two DNA fragments, 495x546 and 545x500. The structures of 495x546 and 545x500 are shown schematically in FIG. 16 and the sequences are given in SEQ ID No 73 and SEQ ID
No 74 respectively.

15 As shown in FIG. 16, 495x546 includes ~ 1 kb of DNA from the 5' end of the ku70 locus fused to ~ 1 kb of DNA from the 5'-region of the wild-type ura3 locus, while 545x500 includes ~ 1 kb of DNA from the 3' end of the ku70 locus fused to ~ 1 kb of DNA from the 3'-region of the wild-type ura3 locus. Each individual fragment contains an incomplete copy of the ura3 locus and these truncated ura3 genes are
20 predicted to be incapable of producing active Ura3 protein. However, the 5' and 3' segments of ura3 overlap by 522 bp so that a full length, functional, ura3 gene can be generated by homologous recombination within this overlap region. Selection for Ura⁺ prototrophic transformants selects for isolates in which such a recombination has occurred and, additionally, the resulting linear molecule, comprising ku70 UP-
25 ura3-ku70 DN, has been integrated into the DAM-6 genome. In MK29404, in most instances, integration occurs ectopically, via the NHEJ pathway, but at low frequency, integration occurs at the ku70 locus via homologous recombination within the Ku70 UP and Ku70 DN regions. As depicted in FIG. 17, this latter event results in deletion of a major portion of the ku70 gene and coding sequence. Such events can be
30 identified by PCR assays that detect transformants in which the ura3 gene is closely linked to ku70 sequences located 3' to the ku70 DN sequence and 5' to the ku70 UP sequence. The construction of such a ku70 deletion mutation in DAM-6 is detailed below. DAM-6 was grown for transformation and protoplasted for PEG-mediated transformation according to the protocol described above. Specifically, using one or

both of the DNA fragments 495x546 and 545x500, a buffer-only negative control, or a positive control plasmid (pDD109) that contains an intact full length *ura3* gene and had, in previous transformation experiments with DAM-6, generated approximately 50-100 Ura⁺ transformants per µg of plasmid DNA.

- 5 Transformation 23: 9 µl EB buffer (no DNA)
 Transformation 24: 5 µl pDD109 (2.5 µg) + 4 µl EB buffer
 Transformation 25: 4.2 µl PCR 495x546 (2.8 µg) + 4.7 µl EB buffer
 Transformation 26: 4.7 µl PCR 545x500 (2.8 µg) + 4.2 µl EB buffer
 Transformation 27: 4.2 µl PCR 495x546 (2.8 µg) + 4.7 µl PCR 545x500 (2.8
 10 µg)
 Transformation 28: 4.2 µl PCR 495x546 (2.8 µg) + 4.7 µl PCR 545x500 (2.8 µg)

Following transformation and grow-out according to the procedures described within, transformations 23, 24, 25 and 26 were plated on two YNB 0.8M sucrose
 15 plates and transformations 27 and 28 were plated on four YNB 0.8M sucrose plates. Plates were sealed with parafilm and incubated at 27°C for 7 days. Transformations 23, 25 & 26 did not yield any colonies. Transformations 24, 27 & 28 each yielded several hundred colonies. Isolated colonies obtained from transformations 27 and 28 were patched and gridded onto YNBglc, YNBglc+ura, and YPD plates which were
 20 incubated at 27°C for 6 days and then stored at 4°C. Ura⁺ transformants from transformations 27 and 28 were screened for insertion of the *ura3* gene at the *ku70* locus by PCR assays. DNA templates prepared from Ura⁺ transformants generated by Transformations 27 and 28 were initially screened by PCR reactions with primers
 25 (DHD532 x DHD550; SEQID No. 020 and SEQ ID No. 021, respectively) or (DHD494 x DHD552; SEQID No. 022 and SEQ ID No. 023, respectively). As depicted in FIG. 17, only those transformants that have inserted *ura3* at the *ku70* locus are expected to be positive in reactions with these primer pairs. Aliquots (5 – 10 µl) of these reactions were analyzed by gel electrophoresis on 1% agarose gels to screen for the presence of the expected products of DHD532 x DHD550 (1277 bp) or
 30 DHD494 x DHD552 (1309 bp). A total of 264 Ura⁺ transformants from transformations 27 and 28 (132 from each transformation) were tested in this primary PCR screen.

One transformant (designated 27/76) from transformation 27 was positive in the primary screen with the DHD532xDHD550 primer pair. Two transformants (designated 28/162 and 28/198) from transformation 28 were positive in the primary screen with the DHD532xDHD550 primer pair. Transformants 27/76 and 28/198
5 were identified as positive in assays using the KAPAHiFi enzyme assay protocol. Transformant 28/162 was identified as positive in an assay using the KAPA2G enzyme assay protocol. These 3 transformants were further tested with additional diagnostic PCR assays to confirm the insertion of the *ura3* gene at the *ku70* locus and consequent deletion of ~1700 bp of the *ku70* gene sequence. These secondary PCR
10 screens were all performed using the KAPAHiFi enzyme assay protocol. The primer pairs used were: DHD532xDHD550 (SEQID No. 020 x SEQID No. 021), DHD494xDHD552 (SEQID No. 022 x SEQID No. 023; FIG. 17), DHD518xDHD520 (SEQID No. 024 x SEQID No. 025; FIG. 16) and DHD508xDHD498 (SEQID No. 026 x SEQID No. 027; FIG. 17). In these secondary PCR screens a wild type gDNA
15 control template (CB19-E12 gDNA @ 25 ng/μl) and a “no DNA” control template (1 μl 10 mM Tris-HCl (8.0) 0.1 mM EDTA (8.0)) were also tested. These PCR reactions with 27/76, 28/262 and 28/198 all gave the expected results for insertion of the *ura3* wild type gene at the *ku70* locus and consequent deletion of a significant portion of the *ku70* gene. The repeat of the primary screen reaction
20 (DHD532xDHD550) again produced a band consistent with the expected size of 1277 bp. Reactions with DHD494xDHD552 also produced a product of a size consistent with that expected (1309 bp) for the insertion of the *ura3* gene at the *ku70* locus. Both the wild type gDNA and “no DNA” controls did not produce any visible products in PCR reactions with these primer pairs. In the reaction of 27/76, 28/262 and 28/198
25 with DHD508xDHD498 no products were observed. This result is consistent with the expected deletion of *ku70* sequences (including the DHD498 sequence) in these three transformants. The reaction with wild type gDNA produced a band consistent with the size expected (1199 bp) for wild type *ku70*. The “no DNA” control did not produce any visible products with these primer pairs. In the reaction of 27/76, 28/262
30 and 28/198 with DHD518xDHD520, two product bands were observed. The sizes of these products were consistent with the sizes expected for the DAM-6 *ura3* gene, which contains a 267 bp deletion, and the full length wild type *ura3* gene introduced by the split-marker transformation. The reaction with wild type gDNA control produced a single band consistent with the wild type *ura3* gene and the “no DNA”

control did not produce any visible products with these primer pairs. Thus the PCR analyses indicate that transformants 27/76, 28/262 and 28/198 all have the expected chromosomal structure for *ura3* inserted at the *ku70* locus as depicted in FIG. 17.

To further confirm the chromosomal structures of the *ku70* locus in each of these three strains, the entire chromosomal region spanning the inserted DNA was amplified by PCR using primers DHD494xDHD550 and the nucleotide sequence of this 3885 bp PCR product was determined. As noted above, and depicted in FIG. 17, primers DHD494 and DHD550 lie outside the *ku70* regions of homology used to drive insertion of the *ura3* gene at the *ku70* locus via homologous recombination. In these reactions, the DNA template for 27/76 was the preparation described above. For 28/162 and 28/198, more highly purified gDNA was prepared by the procedures described previously. PCR reactions with primers DHD494xDHD550 were performed using the KAPAHifi enzyme assay protocol as described above. The DNA sequences of the DHD494xDHD550 PCR products for all three transformants (27/76, 28/162 & 28/198) matched exactly the expected sequence (SEQ ID No. 75). The sequencing results confirm the findings and conclusions from the PCR analyses of the structure of the *ku70* locus in transformants 27/76, 28/162 and 28/198. These three strains contain an insertion of the wild type *ura3* gene at the *ku70* locus and are deleted for 1729 bp of the *ku70* gene. It is predicted that these strains will not produce functional Ku70 protein. It is also predicted, that when these strains in which the Ku70 protein is not functional are transformed, the relative proportion of HR events vs. NHEJ events will be significantly increased.

Example 3: Gene expression in *Sporidiobolus pararoseus*

(i) Expression of the native MK29404 *URA3* in a MK29404-Dryl uracil auxotroph (DAM6)

The plasmid pLP148 (SEQID No. 49), containing the native *URA3* locus, was used to restore the uracil amino acid production auxotrophy in the *URA3*⁻ strain DAM-6. Features of plasmid pLP148 are presented in Table 9.

Table 9: Features of Plasmid pLP148

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 Origin of replication (Ori)	1380-1686
<i>URA3</i> 5' UTR	2308-2402
<i>URA5</i> Exon 1	2403-2588
<i>URA5</i> Exon 2	2655-2766
<i>URA5</i> Exon 3	2831-3372
<i>URA3</i> 3' UTR	3373-3570
pBR322 Origin of replication (Ori)	4073-4692

Specifically, MK29404 DAM-6 was transformed via the protoplast/PEG-mediated system detailed above. This protocol typically results in the generation of 50-100 prototrophic strains per transformation. Colonies that appeared to be prototrophic were then patched onto fresh YNB/0.8M sucrose plates to confirm prototrophy. Genomic DNA was isolated from the resulting putative transformants and analyzed by Southern blot as described above. The probe used for the Southern blot was the *URA3* locus PCR-amplified from pLP148. The primers used for the PCR reaction were SEQID No. 001 and SEQID No. 002. The Southern blot analysis confirmed that the pLP148 plasmid DNA was integrated into the DAM-6 genome, thus rescuing the auxotroph. The non-recombinant MK29404 cells contained a fragment that was consistent with the wild type sequence.

(ii) Expression of the native MK29404 *URA5* in a MK29404-Dry1 uracil auxotroph (AM6)

The plasmid pLP149 (SEQID No. 50), containing the native *URA3* locus, was used to restore the uracil amino acid production auxotrophy in the *URA5*⁻ strain AM-6. Features of plasmid pLP149 are presented in Table 10.

Table 10: Features of Plasmid pLP149

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 Origin of replication (Ori)	1016-1635
<i>URA5</i> Exon 1	2985-3006
<i>URA5</i> Exon 2	3228-3940
<i>URA5</i> 3' UTR	3941-3966
pBR322 Origin of replication (Ori)	4073-4692

Specifically, MK29404 AM-6 was transformed via the protoplast/PEG-mediated system detailed above. This protocol typically results in the generation of 50-100 prototrophic strains per transformation. Colonies that appeared to be prototrophic were then patched onto fresh YNB/0.8M sucrose plates to confirm

prototrophy. Genomic DNA was isolated from the resulting putative transformants and analyzed by Southern blot as described above. The probe used for the Southern blot was the URA5 locus PCR-amplified from pLP149. The primers used for the PCR reaction were SEQID No. 003 and SEQID No. 004. The Southern blot analysis confirmed that the pLP149 plasmid DNA was integrated into the AM-6 genome, thus rescuing the auxotroph. The non-recombinant MK29404 cells contained a fragment that was consistent with the wild type sequence.

(iii) Heterologous expression of the Phleomycin resistance gene, in MK29404 and MK29404-Dry1

Phleomycin is an example of a dominant selectable marker and its expression can serve as an example of heterologous gene expression and transformation selection by strains of *S. pararoseus*. Due to the complex nature of the genome of *S. pararoseus* MK29404 strain, the phleomycin open reading frame had the first intron from the URA3 gene inserted after the first two codons of the phleomycin gene (*experimentally determined, data not shown*). The promoter and terminator regions of MK29404 URA3 were used to drive expression of phleomycin gene and lead to the construction of the vector, pLP159 (SEQID No. 51). This finding is consistent with results observed in other Basidiomycete yeast (e.g. *Copriniopsis cinerea*) (S. Kilaru et. al. (2009), Investigating dominant selection markers for *Copriniopsis cinerea*: a carboxin resistance system and re-evaluation of hygromycin and phleomycin resistance vectors, Curr. Genet. Vol 55(5):543-550). Features of plasmid pLP159 are presented in Table 11.

Table 11: Features of Plasmid pLP159

Feature	Base pair coordinates
Kanamycin resistance gene	24-821
<i>URA3</i> promoter	1255-1634
<i>sh ble</i> exon 1	1635-1640
<i>URA3</i> intron 1	1641-1706
<i>sh ble</i> exon 2	1707-2075
<i>URA3</i> terminator	2076-2289
pUC origin of replication (Ori)	2944-3747

MK29404 and MK29404-Dry1 were transformed with pLP159 via PEG-mediated transformation and plated on YNB/0.8M sucrose/phleomycin (150 µg/mL) plates. Putative transformants were patched onto fresh YNB/0.8M sucrose/phleomycin (150µg/mL) plates to confirm phleomycin resistance. Southern

blot analyses of transformant DNA confirmed the integration of pLP159 plasmid DNA into the MK29404 and MK29404-Dry1 genomes.

(iv) Heterologous expression of multiple copies of the Phleomycin resistance gene, in MK29404 and MK29404-Dry1 – Circular vs. Fragmented DNA transformations

5 The role of circular and fragmented pLP159 (SEQID No. 51) DNA and its effect on gene copy number was studied. This approach resulted in the integration of multiple copies of a heterologously expressed gene into the genome. Specifically, the MK29404-Dry1 strain was transformed (using the protoplast/PEG-mediated transformation protocol as described above), with circular and fragmented plasmid
10 DNA (both from plasmid pLP159). Transformants resulting from circular pLP159 transformation were analyzed by PCR and Southern-blotting methods described in detail previously. Briefly, isolated genomic DNA was digested with ApaLI and SacI (A/S), this excises Phleomycin resistance gene, or with XmaI (X), digested gDNA was run on an agarose gel, transferred to a nylon membrane and membrane was
15 incubated with Phleomycin specific probe. Probe was made specific to phleomycin exon after the intron, as shown in FIG. 18. Results from Southern Blotting showed bands of expected sizes but of higher intensity than that observed for single copy integrants, leading to the conclusion that phleomycin gene is present in multiple copies within the MK29404-Dry1 transformant genome.

20 Six transformants (CPX7-1 to CPX7-6) which were confirmed as containing the phleomycin cassette by PCR, resulted from transformation with fragmented pLP159. Southern blotting following the method described previously was also conducted to further confirm these results using fragment for labeling from vector pLP155 (phleomycin construct with URA3 promoter region and no URA3 intron1,
25 data not shown) and primers SEQID No. 005 and SEQID No. 006. Results are shown in FIG. 19. The expected 924 bp band was noted for all of the transformants except CPX7-4, which contained no bands and suggests that a fragmented phleomycin resistance gene construct integrates in multiple copies in the genome.

(v) Heterologous expression of multiple copies of the Phleomycin resistance gene, in MK29404 and MK29404-Dry1 – Promoter region testing and the effect on expression of Phleomycin and copy number

30 An evaluation of a series of potential promoter regions (SEQID No. 52-56) was conducted and reviewed for their ability to influence copy number and function in MK29404 Dry1 transformants. This approach resulted in the integration of multiple

copies of a heterologously expressed gene into the genome. Specifically, five variant plasmids were created by subcloning promoter regions (SEQID No. 52-56) into pLP159 and transformed into MK29404-Dry1 using the protoplast/PEG-mediated transformation protocol, as described previously. Results from these transformations are listed in Table 12.

Table 12: Promoter Region Testing

Strain	Transformation Construct	Number of Colonies
Dry1	No DNA	1
Dry1	pLP159 (URA3 phleo)—Pos. Control	21
Dry1	pLP174 (Beta Tubulin phleo)	98
Dry1	pLP175 (GAPDH phleo)	35
Dry1	pLP176 (Phytoene Synthase)	16
Dry1	pLP177 (Malic Enzyme Isoform 2 phleo)	49
Dry1	pLP179 (Malic Enzyme Isoform 1 phleo)	24

Phleomycin resistance was confirmed by restreaking of colonies on YNB/2% dextrose/Phleo150 plates and by PCR (data not shown). Genomic DNA was prepared as described above and served as PCR template. This was further digested with NdeI (this cut linearizes pLP159 backbone and that of its descendants) or SacI and ApaLI (yields fragments of 2.1 kb for pLP159, 2.2 kb for pLP174, 2.2 k for pLP175, 1.8 kb for pLP176, 2.4 kb for pLP177 and 2.2 kb for pLP179), run on a gel and transferred to the nylon membrane for Southern Blotting Analysis as described above. Phleomycin probe from above was used to detect phleomycin in the genome of Dry1 transformants.

Results are shown in FIG. 20. Both Southern Blots showed multiple bands in most lanes and bands of high intensity suggesting multiple copies of Phleomycin in the genome of transformants in the study either due to integration in several locations or concatemers.

(vi) Heterologous expression of other antibiotic agents

The sensitivity of MK29404 to selection agents was determined by including these inhibitors in YPD agar medium at various concentrations and spreading cells on the plates at densities similar to those that are present during procedures used for selection of recombinant or transformed cells. This level is known as the minimal growth inhibitory concentration (MIC). MK29404 was grown in 50 mL of YPD medium for 2 days at 27°C with shaking at 200 rpm. After this initial growth phase, cells were sub-cultured at 1:100 into 50 mL of YPD medium and grown overnight at 27°C with shaking at 200 rpm. 1 mL of the overnight culture was removed and

diluted at 1:4 into YPD medium. 25 μ L of diluted culture was plated on YPD agar containing a single antibiotic at a various concentrations from 0 to 200 μ g/mL. Plates were incubated at 27 °C for ~3 days and assayed for growth. Table 13 shows agents agent that successfully inhibited growth in MK29404 and MIC.

5

Table 13: Selection Agents Against MK29404

Antibiotic	MIC (μ g/mL)
Hygromycin B	75
Zeocin	75
Geneticin	75
Carboxin	100

The expression of a gene to confer resistance a selection agent is needed for transformation. Expression of the phleomycin resistance gene and selection of transformed, resistance cells is described above. In a similar way, established resistance genes for hygromycin B, geneticin, and carboxin can be expressed and transformed cells selected. For successful expression of the resistance genes for hygromycin B and geneticin one or more of the following techniques can be utilized with the resistance gene: codon optimized, inserted with an intron of a homologous or heterologous gene, inserted behind a promoter element of a homologous or heterologous gene, targeted or random insertion into the genome.

For resistance to carboxin the native succinate dehydrogenase (SDH1; SEQID No. 76; SEQID No. 77) locus can be cloned and highly conserved histidine residue within the third cysteine-rich cluster of the gene replaced by a leucine residue. This mutated copy of the succinate dehydrogenase is resistance to carboxin and can be transformed into MK29404 and a resistance gene.

20

(vii) Extra-chromosomal plasmid expression

The following example demonstrates extra-chromosomal plasmid expression. The *Ustilago maydis* plasmid, pUXV1, is an autonomously replicating plasmid, and has demonstrated cross-functionality in several *Basidiomycetes* (Kinal et. al., 1993; Ortiz-Alvarado et. al., 2006). The MK29404 URA3 locus is inserted into this plasmid creating pLP160 (SEQ ID No. 71) and is used to transform a URA3 auxotrophic mutant, DAM-6. The resulting transformants were analyzed to determine whether they contained extra-chromosomal replication of the plasmid. Table 14 lists feature coordinates of pLP160.

30

Table 14: pLP160 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
<i>URA3</i> 5' UTR	2334-2428
<i>URA3</i> Exon 1	2429-2614
<i>URA3</i> Exon 2	2681-2792
<i>URA3</i> Exon 3	2857-3398
<i>URA3</i> 3' UTR	3399-3596
<i>Ustilago Maydis</i> autonomously replicating sequence	5692-6068

MK29404 DAM-6 is transformed with pLP160 via the PEG-mediated
 5 transformation method described earlier, and the putative transformants are analyzed
 by plasmid extraction. Namely, 25 mL YPD medium in a 250 mL flat-bottom shaker
 flask is inoculated with putative transformant cells from patch plate (or cryovial).
 Cultures are incubated for 2 days at 27°C with shaking at 200 rpm, sub-cultured 1:250
 into 25 mL YPD in 250 mL flat-bottom shaker flask and incubated at 27°C overnight
 10 with shaking at 200 rpm. Following this overnight growth, cultures ODs at A600 are
 typically between 4-8. From such cultures, 25 OD₆₀₀ units of cells are centrifuged in
 a 15 mL conical tube at ~4,000 rcf for 5 minutes, washed once with 5 mL sterile
 diH₂O, resuspended by vortexing, and again centrifuged at ~4,000 rcf for 5 minutes.
 All traces of supernatant are carefully removed from the resulting pellet. This pellet
 15 was resuspended in 5 mL 25 mM β-mercaptoethanol/5 mM EDTA and placed on a
 gyrator for 20 minutes at room temperature, and centrifuged as above. The
 supernatant was discarded and the pellet was resuspended in 5 mL sterile diH₂O by
 vortexing, and centrifuged again as above. The resulting supernatant is discarded and
 the pellet is resuspended in 5 mL 0.7 M NaCl by vortexing, and centrifuged again as
 20 above. The resulting supernatant is discarded and the pellet is resuspended in 1.8 mL
 'MPPE' protoplasting enzyme mix. The reaction is placed at 30°C with shaking at
 100 rpm for ~1.5 hours. The cells are then centrifuged cells as before and
 resuspended cell pellet in 250 μL of Buffer P1 (Qiagen, QIAprep Spin Miniprep Kit
 (Catalog #27104)) containing 0.1 mg/mL RNase A. The cell pellet is transferred to a
 25 1.5 mL microfuge tube and added 250 μL lysis buffer P2 (Qiagen) to the tube and
 invert gently 4-5 times to mix and incubated at room temperature for 5 minutes. After
 incubation, 350 μL of neutralization buffer N3 (Qiagen, QIAprep Spin Miniprep Kit
 (Catalog #27104)) is added to the tube and inverted immediately but gently 4-6 times.
 The lysate is centrifuged for 10 minutes at 16,100 rcf. The supernatant cleared lysate

is transferred from step 6 to QIAprep Spin column by decanting or pipetting and centrifuged as described previously for 1 minute. The QIAprep Spin Column is washed by adding 0.75 mL of buffer PE (Qiagen, QIAprep® Spin Miniprep Kit (Catalog #27104)) and centrifuged as before for 1 minute. The flow-through is discarded and DNA eluted by adding 25 µL of Buffer EB (Qiagen, QIAprep® Spin Miniprep Kit (Catalog #27104)) to the center of each QIAprep Spin column and centrifuged as above for 1 minute. The resulting elutant from the plasmid preps from the putative transformants is analyzed on a 0.7% agarose gel by standard agarose gel electrophoresis, followed by staining with SYBR Safe DNA gel stain (Life Technologies). Bands corresponding to the 6.6 kb plasmid would then be visualized on the gel, transformed into *E. coli* 10-beta chemically competent cells by standard procedures with 5 µL of each plasmid preparation and plated on LB Amp100 plates and resulting in a number of colonies for further manipulation.

(viii) Removal of native splice sites for transcription

To investigate the expression of the reporter gene eGFP, the plasmid pLP180 (SEQID No. 78), containing the native URA3 locus, was used to reconstitute the uracil amino acid production auxotrophy in the URA3⁻ strain DAM-6. Table 15 lists features of pLP180.

Table 15: pLP180 Feature Coordinates

Feature	Base pair coordinates
Kanamycin resistance gene (Kan)	1-807
High copy Origin of replication	849-1652
URA3 locus	1825-3388
URA5 upstream	3891-4938
eGFP exon 1	4939
URA5 intron 1	4940-5018
eGFP exon 2	5019-5737
URA5 downstream	5738-6013

MK29404 DAM-6 was transformed via the protoplast/PEG-mediated system detailed above. Colonies that appeared to be prototrophic were then patched onto fresh YNB/0.8M sucrose plates to confirm prototrophy. Fluorescence was not detected from transformed cells. RNA transcripts of eGFP were then analyzed by reverse transcription PCR of RNA isolated from MK29404 cells transformed with plasmid pLP180, as shown in FIG. 21. RNA was isolated (protocols followed Applied Biosystem's Ribopure Yeast RNA Purification kit; part number AM1926). Cells were grown in 250 mL flat-bottom flask containing 25 mL YPD medium and incubated at 27°C with shaking at 200 rpm for 48 hrs. These cells were inoculated at

1:200 dilution factor into 250 mL flat-bottom flask containing 25 mL YPD medium and incubated at 27°C with shaking at 200 rpm for 7 hours. These cells were used to inoculate a culture to OD₆₀₀ of 0.05 in 25 mL YPD medium contained within 250 mL flat-bottom flask and incubated at 27°C with shaking at 200 rpm overnight for 16 hrs.

5 1.8 mL cells were harvested from these cultures in 2.0 mL microcentrifuge tube at 12,000 x g for 2 minutes and then removed all traces of supernatant. 480 µL of lysis buffer, 48 µL of 10% SDS, and 480 µL Phenol:Chloroform:Isoamyl alcohol was added to the cell pellet and vortexed vigorously at maximum speed for 10-15 sec until cell pellet is fully re-suspended. The resulting mixture was transferred into 1.5 mL

10 screw-cap tube containing 750 µL ice-cold Zirconia bead and vortex at for 10 minutes. The mixture was centrifuged at 16,100 x g at room temperature for 5 minutes and the aqueous phase transferred into a 15 mL screw-cap conical tube. 1.9 mL binding buffer was added and mixed thoroughly by inversion. 1.25 mL 100% ethanol was added and mixed thoroughly by inversion. 700 µL of sample mixture was applied

15 to a filter cartridge assembled in a collection tube and centrifuged at high speed for 30 seconds. The filter cartridge was then washed per manufacturer's instructions. The RNA was eluted by applying 50 µL elution solution to the center of the filter and centrifuged at high speed for 1 minute. DNase I treatment was by adding 10 µL 10X DNase I buffer and 4 µL DNase I to RNA solution and mixed well and incubated at

20 37°C for 10 minutes. The DNase I was then inactivated by adding 11 µL DNase inactivation reagent to the sample, mixed well by vortexing briefly and allowed to stand at room temperature for 5 minutes. The samples were then centrifuged at high speed for 3 minutes to pellet DNase inactivation reagent and then the RNA-containing supernatant transferred to fresh 1.7 mL tube and stored at -80°C.

25 cDNA was synthesized from RNA using the iScript cDNA Synthesis Kit from Bio-Rad (Cat # 170-8890). Per reaction: 4 µL of 5X iScript reaction mix, 1 µL iScript reverse transcriptase, 800 ng of RNA template and nuclease-free water to total 20µL. Incubated the complete reaction mix for 5 minutes at 25°C, 30 minutes at 42°C and 5 minutes at 85°C. Stored cDNA at -20°C. The reverse transcriptase PCR conditions

30 employed with cDNA template were as follows: 300 µM dNTPs, 0.4 µM each primer, 2 µL cDNA, 1 U KAPA® HiFi HotStart polymerase (KAPA Biosystems), and 1× KAPA® HiFi buffer (KAPA Biosystems) in a 25 µL total volume. The PCR Protocol included the following steps: (1) 95° C. for 2 minutes; (2) 98° C. for 20 seconds; (3) 55° C. for 15 seconds; (4) 72° C. for 30 seconds; (5) repeat steps 2-4 for 25 cycles; (6)

72° C. for 3 minutes; and (7) hold at 4° C. The primers used in the amplification were specific to the 5' and 3' URA5 untranslated region and flanked the eGFP gene. They were prLP637 (SEQID No. 037) and prLP638 (SEQID No. 038). PCR products were analyzed by standard agarose gel electrophoresis, followed by staining with SYBR®
 5 Safe DNA gel stain (Life Technologies). The results showed a DNA fragment of approximately 764 bp corresponding to the predicted full length eGFP gene as well as the native URA5 gene. Two additional DNA fragments of approximately 260 bp and 450 bp were also seen. These fragments were excised and gel purified using the GeneJET gel extraction kit (# K0691). Sequence analysis of the smaller fragments
 10 showed sequence internal to the eGFP gene and indicated the eGFP transcript was being processed or spliced by MK29404's RNA processing enzymes (see FIG. 23 - underlined sequences are the sequence of the two smaller, spliced products seen in FIG. 21). The results shown in FIG. 22 indicate no detectable expression of eGFP protein in MK29404 DAM6 cells transformed with pLP180.

15 Plasmids pLP184 (SEQ ID No. 79) and pLP185 (SEQID No. 80) were constructed to mutate the putative splice sites indicated from the data above. In both constructs the potential splice sites (bold in FIG. 22) were mutated from GGT to GGA, AAG to AAA and CAG to CAA. These changes conserved the amino acid sequence. Additionally, the second codon in the eGFP sequence was changed from a
 20 serine to a valine pLP185. Tables 16 and 17 list features of pLP184 and pLP185, respectively.

Table 16: pLP184 Feature Coordinates

Feature	Base pair coordinates
Kanamycin resistance gene (Kan)	1-807
High copy Origin of replication	849-1652
<i>URA3</i> locus	1825-3388
<i>URA5</i> upstream	3891-4938
<i>eGFP</i> exon 1	4939
<i>URA5</i> intron 1	4940-5018
<i>eGFP</i> exon 2	5019-5737
<i>URA5</i> downstream	5738-6013

Table 17: pLP185 Feature Coordinates

Feature	Base pair coordinates
Kanamycin resistance gene (Kan)	1-807
High copy Origin of replication	849-1652
<i>URA3</i> locus	1825-3388
<i>URA5</i> upstream	3891-4938
<i>eGFP</i> exon 1	4939
<i>URA5</i> intron 1	4940-5018
<i>eGFP</i> exon 2	5019-5737
<i>URA5</i> downstream	5738-6013

The uracil auxotrophic MK29404 strain, DAM6 was transformed with pLP184 and pLP185 using the protoplast/PEG-mediated transformation protocol (see Example 1a). The presence of the plasmids was confirmed via PCR using the primers prLP631 (SEQID No. 039) and prLP632 (SEQID No. 040). RNA was extracted from the transformants and RT-PCR was performed as described above. PCR products were analyzed as above, followed by staining with SYBR® Safe DNA gel stain (Life Technologies, FIG. 24). The smaller DNA fragments, representing the spliced fragments of eGFP are no longer detected. The ~ 760 bp product was excised from the agarose gel, purified (GeneJET gel extraction kit (# K0691)) and sequenced and confirmed to be full-length eGFP. Protein isolation and Western Blot analysis was carried out as described in section x on MK29404 cells transformed with pLP184 (a) and pLP185 (b), as shown in FIG. 25. Low-levels of florescence were observed in these cells demonstrating that the mutation of the splice sites allowed expression of eGFP.

Example 4: Over-expression of the native malic enzyme, acetyl-CoA synthetase,

ATP Citrate lyase and phytoene dehydrogenase genes in *S. pararoseus*

(i) *Malic Enzyme*:

The following example demonstrates the successful gene expression of the native MK29404 Malic Enzyme Isoform 1 and Malic Enzyme Isoform 2 genes in the MK29404-ura5 auxotrophic background (AM6) and the MK29404-Dry1-ura5 auxotrophic background (DAM7) using expression plasmids that were integrated into the respective strain genomes via genetic transformation. Two versions of the expression plasmids were analyzed with one in its native context and one containing a V5 tag from the Simian virus 5. Plasmids used for transformation and subsequent expression in the aforementioned strains are listed in Table 18.

Table 18: Plasmids Used for Transformation and Subsequent Expression

Name	SEQID	Description
pDS223	SEQID No. 57	Malic Enzyme-1 expression plasmid
pDS224	SEQID No. 58	Malic Enzyme-2 expression plasmid
pDS228	SEQID No. 59	Malic Enzyme-1+V5 expression plasmid
pDS229	SEQID No. 60	Malic Enzyme-2+V5 expression plasmid

Plasmid features and sequences for each Malic Enzyme expression plasmid are listed in Tables 19-22.

5

Table 19: pDS223 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
pBR322 origin of replication (Ori)	1016-1635
MK29404 URA5 locus	2094-4222
MK29404 Malic Enzyme Isoform 1 locus	4319-8889
f1 origin of replication (Ori)	9170-9476

Table 20: pDS224 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
pBR322 origin of replication (Ori)	1016-1635
MK29404 URA5 locus	2094-4222
MK29404 Malic Enzyme Isoform 2 locus	4319-7638
f1 origin of replication (Ori)	7919-8225

Table 21: pDS228 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
pBR322 origin of replication (Ori)	1016-1635
MK29404 URA5 locus	2094-4222
MK29404 Malic Enzyme Isoform 1 locus	4319-8931
V5 Tag (GKPIPNPLLGLDST)	8137-8178
f1 origin of replication (Ori)	9062-9368

10

Table 22: pDS229 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
pBR322 origin of replication (Ori)	1016-1635
MK29404 URA5 locus	2094-4222
MK29404 Malic Enzyme Isoform 2 locus	4319-7680
V5 Tag (GKPIPNPLLGLDST)	7268-7309
f1 origin of replication (Ori)	7961-8267

Plasmids pDS223, pDS224, pDS228 and pDS229 were transformed into strains AM6 and/or DAM7 via the protoplast/polyethylene glycol-mediated system detailed above. Genomic DNA was isolated from the resulting putative AM6 and/or DAM7 +pDS223 and +pDS224 transformants and analyzed by Southern blot. The probes used for the Southern blots were specific to either Malic Enzyme Isoform 1 or to Malic Enzyme Isoform 2. The Malic Enzyme specific probes were generated by

PCR using primers SEQID No 007 and 008 for Isoform 1 and SEQID No. 009 and 010 for Isoform 2.

Southern blot analysis revealed a hybridized band consistent in size with the native Malic Enzyme Isoform 1 gene in all transformants analyzed. In a sub-set of transformants, additional hybridized band(s) were observed and sizes were consistent with what would be expected for the presence of additional copy/copies of the full-length Malic Enzyme Isoform 1 gene, thus confirming the insertion of additional copy/copies of Malic Enzyme Isoform 1 gene in these select transformants. Southern blot analysis revealed a hybridized band consistent in size with the native Malic Enzyme Isoform 2 gene in all transformants analyzed. In a sub-set of transformants, additional hybridized band(s) were observed and sizes were consistent with what would be expected for the presence of additional copy/copies of the full-length Malic Enzyme Isoform 2 gene, thus confirming the insertion of additional copy/copies of Malic Enzyme Isoform 2 gene in these select transformants.

In order to detect expression of the Malic Enzyme isoform 1 and Malic Enzyme isoform 2 genes introduced via transformation with pDS228 and pDS229, respectively, Western blotting was carried out using a monoclonal antibody specific to the V5 tag on the 3' end of each Malic Enzyme gene contained within these expression constructs. Total protein was isolated and purified from putative AM6 and/or DAM7 +pDS228 and +pDS229 transformants using the following method: Western blot analysis on AM6 and/or DAM7 +pDS228 and +pDS229 transformants revealed a positive signal for most transformants within the size range predicted for both the Malic Enzyme Isoform 1 and Malic Enzyme Isoform 2 proteins (with a V5 tag). No signal was observed in the lanes containing total protein derived from the untransformed parental strains, MK29404 (WT) or MK29404-Dry1. These data confirm successful expression of Malic Enzyme Isoform 1+V5 and Malic Enzyme Isoform 2+V5 via the introduced nucleotide sequence derived from pDS228 and pDS229, respectively.

(ii) Acetyl-CoA Synthetase:

The following example outlines the successful gene expression of the native MK29404 Acetyl-CoA Synthetase (ACS) in the MK29404-Dry1-ura3 auxotrophic background strain (DAM6) using expression plasmids that were integrated into the DAM6 strain genome via PEG-mediated genetic transformation. There are 2 isoforms and 2 versions of the expression plasmids – each isoform in its native context and

each containing a "V5 tag" from the Simian virus 5. All the genes were synthesized and subcloned into pLP148 (SEQID No. 49) vector (URA3 cassette-expressing vector) via a unique PciI restriction site. Plasmids used for transformation and subsequent expression in DAM6 are listed in Table 23.

5 Table 23: Plasmids Used for Transformation and Subsequent Expression

Plasmid	SEQID	Description
pYB11	SEQID No. 61	Acetyl-CoA Synthetase-1 expression plasmid
pYB12	SEQID No. 62	Acetyl-CoA Synthetase -2 expression plasmid
pYB11A	SEQID No. 63	Acetyl-CoA Synthetase -1+V5 expression plasmid
pYB12A	SEQID No. 64	Acetyl-CoA Synthetase -2+V5 expression plasmid

Plasmid Features and sequences for each Acetyl-CoA Synthetase expression plasmid are listed in Tables 24-27.

Table 24: pYB11 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 Acetyl-CoA Synthetase Isoform 1 locus	4032-10334
pBR322 origin of replication (Ori)	10380-10999

10

Table 25: pYB12 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 Acetyl-CoA Synthetase Isoform 2 locus	4032-7003
pBR322 origin of replication (Ori)	7049-7668

Table 26: pYB11A Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 Acetyl-CoA Synthetase Isoform 1 locus	4028-10376
V5 Tag (GKPIPNPLLGLDST)	8118-8159
pBR322 origin of replication (Ori)	10422-11041

15

Table 27: pYB12A Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 Acetyl-CoA Synthetase Isoform 2 locus	4032-7045
V5 Tag (GKPIPNPLLGLDST)	6809-6850
pBR322 origin of replication (Ori)	7091-7710

Plasmids pYB11, pYB11A, pYB12 and pYB12A were transformed into strain DAM6 via the protoplast/PEG-mediated approach as described previously. Transformation results are listed in Table 28.

Table 28: Transformation Results

Strain	Transformation	Construct	# of Colonies
DAM6	No DNA	Negative Control	0
DAM6	pLP148	Positive Control	43
DAM6	T001	pYB11A circular	10
DAM6	T001A	pYB11A linear	14
DAM6	T002	pYB12A circular	14
DAM6	T002A	pYB12A linear	20
DAM6	T005	pYB11 circular	52
DAM6	T005A	pYB11 linear	51

5 A quick gDNA preparation and PCR method (described above) was then used for transformants T001, T001A, T002, T002A, T005 and T005A, to confirm the successful recombination of the gene of interest into the DAM6 genome. PCR primer prURA3 F (SEQID No. 011) is specific to URA3 cassette in the backbone of all the plasmids; prYB11 R (SEQID No. 012) and prYB12 R (SEQID No. 013) are specific to Acetyl-CoA Synthetase Isoform 1 and 2 respectively. Results from the PCR reaction (FIG. 26), following conditions set out previously showed that T005 lanes 2-6, and in T005A lanes 1-3, 5 have the expected size band of ~1kb, while bands in lanes T005-1, T005A-4 and T005A-6 are smaller in size and could indicate incomplete recombination event or other issues with the construct in the genome.

Further, in order to detect expression of the Acetyl-CoA Synthetase Isoform 1 and Acetyl-CoA Synthetase Isoform 2 genes introduced via transformation with pYB11A and pYB12A, respectively, Western blotting was carried out using a monoclonal antibody specific to the V5 tag on the 3' end of each Acetyl-CoA Synthetase gene contained within these expression constructs. Results are shown in FIG. 27. Total protein was isolated and purified from putative T001/T001A and T002/T002A transformants (DAM6 with pYB11A and pYB12A, respectively) as described previously.

Western Blot analysis on DAM6 transformants with Acetyl-CoA Synthetase Isoform 1 revealed a positive signal for all four transformants (T001/T001A) in the range predicted for the protein (73-75 kDa) as well as in Positive control-2 lane. Multiple bands in V5 tag-positive lanes are a possible indication of protein degradation. No signal was observed in T002/T002A lanes or in parental strain

DAM6. No signal in T002/T002A lanes—DAM6 transformed with Acetyl-CoA Synthetase Isoform2 were detected.

(iii) *ATP Citrate Lyase:*

The following example demonstrates the successful gene expression of the native MK29404 ATP Citrate Lyase (ACL) in the MK29404-Dry1-ura3 auxotrophic background strain (DAM6) using expression plasmids that were integrated into the DAM6 strain genome via PEG-mediated genetic transformation. There are 2 isoforms and 2 versions of the expression plasmids – each isoform in its native context and each containing a “V5 tag” from the Simian virus 5. All the genes were synthesized and subcloned into pLP148 (SEQID No. 49) vector (URA3 cassette-expressing vector) via a unique PciI restriction site. Table 29 lists the plasmids used for transformation and subsequent expression in DAM6.

Table 29: Plasmids Used for Transformation and Subsequent Expression

Plasmid	SEQID	Description
pYB13	SEQID No. 65	ATP Citrate Lyase-1+2 expression plasmid
pYB14	SEQID No. 66	ATP Citrate Lyase-1 expression plasmid
pYB13A	SEQID No. 67	ATP Citrate Lyase-1+2-V5 expression plasmid
pYB14A	SEQID No. 68	ATP Citrate Lyase-1+V5 expression plasmid

Plasmid Features and sequences for each ATP Citrate Lyase expression plasmid are listed in Tables 30-33.

Table 30: pYB13 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 ATP Citrate Lyase-1+2	4032-13055
pBR322 origin of replication (Ori)	13101-13720

Table 31: pYB14 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 ATP Citrate Lyase-1	4032-11207
pBR322 origin of replication (Ori)	11253-11872

Table 32: pYB13A Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 ATP Citrate Lyase-1+2-V5	4032-13139
V5 Tag (GKPIPNNPLLGLDST)	4897-4938
V5 Tag (GKPIPNNPLLGLDST)	12867-12908
pBR322 origin of replication (Ori)	13185-13804

Table 33: pYB14A Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1-861
f1 origin of replication (Ori)	1380-1686
MK29404 URA3 locus	2018-3587
MK29404 ATP Citrate Lyase-1	4032-11249
V5 Tag (GKPIPNNPLLGLDST)	10977-11018
pBR322 origin of replication (Ori)	11295-11914

5

Plasmids pYB13, pYB13A, pYB14 and pYB14A were transformed into strain DAM6 via the protoplast/polyethylene glycol-mediated approach as described previously. Transformation results are shown in Table 34 (Transformation of DAM6 strain with pYB14 (ATP Citrate Lyase-1) was not done due to lack of V5 signal in transformants with pYB12A plasmid as determined by Western Blot results in FIG. 29).

Table 34: Transformation Results

Strain	Transformation	Construct	# of Colonies
DAM6	No DNA	Negative Control	0
DAM6	pLP148	Positive Control	43
DAM6	T003	pYB14A circular	9
DAM6	T003A	pYB14A linear	9
DAM6	T004	pYB13A circular	12
DAM6	T004A	pYB13A linear	11
DAM6	T007	pYB13 circular	15
DAM6	T007A	pYB13 linear	15

A quick gDNA preparation and PCR method (described above) was then used for transformants T003, T003A, T004, T004A, T007 and T007A, to confirm the successful recombination of the gene of interest into the DAM6 genome. PCR primer prURA3 F (SEQID No. 011) is specific to URA3 cassette in the backbone of all the plasmids; prYB13 R (SEQID No. 014) and prYB14 R (SEQID No. 015) are specific to ATP Citrate Lyase 1 and 1+2, respectively. Results from the PCR reaction (FIG. 28), following conditions set out previously, showed that T004/T004A lanes 1, 4-6, 9, 12-13 have the expected size band of ~1kb, while bands in lanes 2, 3, 7, 8, 10, 11 are

smaller in size or nonexistent and could indicate incomplete recombination event, lack of recombination or other issues with the construct in the genome.

Further, in order to detect expression of the ATP Citrate Lyase-1+2-V5 and ATP Citrate Lyase-1+V5 genes introduced via transformation with pYB13A and pYB14A, respectively, Western blotting was carried out using a monoclonal antibody specific to the V5 tag on the 3' end of each ATP Citrate Lyase gene contained within these expression constructs. Results are shown in FIG. 29. Total protein was isolated and purified from putative T003/T003A and T004/T004A transformants (DAM6 with pYB14A and pYB13A, respectively) as described previously.

Western Blot analysis on DAM6 transformants with ATP Citrate Lyase Isoforms-1+2-V5 revealed a positive signal for four out of twelve transformants (T004/T004A) in the range predicted for the proteins (~146 kDa for Isoform1, ~19 kDa for Isoform 2) as well as in Positive control-2 lane. Isoform 2-specific size bands are not observed. The detection of V5 epitope in 4/12 T004/T004A transformants could be due to integration issues of the full-length construct, with full-length construct being ~9 kb (~14 kb for full-length plasmid). Multiple bands in lanes with V5 tag expression could be the result of protein degradation and need to be studied further. No signal was observed in T003/T003A (ATP Citrate Lyase Isoform-1+V5) lanes or in parental strain DAM6.

(iv) Phytoene Dehydrogenase Complementation:

The following example demonstrates the successful complementation of a strain of MK29404 (PMX21-1) which lacks the Phytoene Dehydrogenase (PDH) gene (e.g. a Phytoene Dehydrogenase “knockout” – *following both the single and split-marker homologous recombination methods*). This was accomplished by re-introduction of the native PDH gene via genetic transformation, and is evidenced by: (a) Pigment restoration, and (b) PCR data. Table 35 lists the plasmids used to demonstrate PDH complementation in strain PMX21-1 (MK29404-DAM6, PDH:URA3).

Table 35: Plasmids Used to Demonstrate PDH Complementation in Strain PMX21-1

Plasmid	SEQID	Description
pDS243	SEQID No. 69	PDH complementation plasmid with adapter
pDS245	SEQID No. 70	PDH complementation plasmid without adapter

Plasmid Features and sequences for each PDH complementation plasmid (with and without adapter sequence) are listed in Tables 36 and 37.

Table 36: pDS243 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1 – 861
Origin of replication (pMB1)	1021 – 1635
Phleomycin resistance cassette (antisense strand)	2164 – 3198
Phytoene Dehydrogenase upstream seq.	3678 – 5615
Phytoene Dehydrogenase ORF	5616 – 7945
Phytoene Dehydrogenase downstream seq.	7946 – 9930
Adapter Sequence	8342 – 8359

5

Table 37: pDS245 Feature Coordinates

Feature	Base pair coordinates
Ampicillin resistance gene (Amp)	1 – 861
Origin of replication (pMB1)	1021 – 1635
Phleomycin resistance cassette (antisense strand)	2164 – 3198
Phytoene Dehydrogenase upstream seq.	3678 – 5615
Phytoene Dehydrogenase ORF	5616 – 7945
Phytoene Dehydrogenase downstream seq.	7946 – 9908

Plasmids pDS243 and pDS245 were transformed separately into MK29404 strain PMX21-1 via the protoplast/polyethylene glycol-mediated approach as described previously, along with a negative control transformation. FIG. 30 shows plate pictures of these transformations. Transformation #2 and #3 resulted in mostly white colonies, similar to that observed in the “negative control” transformation; however, a small portion of the colonies observed on these plates (17 out of 691 and 26 out of 789, respectively) were red in coloration. These results, provided in Table 38, indicate that the PDH deletion in the PMX21-1 strain had been complemented by introduction of either pDS243 (transformation #2) or pDS245 (transformation #3) in the case of the red pigmented colonies.

Table 38: Transformation Results

Transformation	Plasmid transformed	No. of Colonies	No. of Red Colonies
1	-	673	0
2	pDS243	691	17
3	pDS245	789	26

Further, to confirm the presence of the pDS243 plasmid within the red colony transformants observed in Transformation No.2, genomic DNA obtained from 16 of these colonies was isolated and analyzed by PCR as described previously. Results are shown in FIG. 31. Specifically, SEQID No. 016 and SEQID No. 017, corresponding to sense and antisense primers specific to the adapter sequence present in the pDS243 plasmid, and is therefore unique to this plasmid and should not be present in the

MK29404 PMX21-1 genome, respectively were used. Templates used in PCR reactions included: PMX21-1 “parental” genomic DNA, genomic DNA isolated from Transformation No.2, red colonies 1-16. A “no DNA, negative control” reaction was included as well. Amplification of a 1,046 base pair product by PCR in all of the Transformation No.2 PCR reactions indicates successful introduction of the pDS243 plasmid in these red colonies. The lack of a PCR product of this size in the No DNA negative control reaction or the PMX21-1 parental reaction confirms the specificity of this PCR product to colonies containing the pDS243 plasmid only.

Example 5: Development of a Protoplast fusion technique in *S. pararoensis*

The following describes a method developed for quickly combining traits from various strains, which arise from either classical strain improvement or metabolic engineering. The use of mating strains is one standard way of combining traits and has been demonstrated in various yeast-based systems. A functional mating system allows independent selection of dispersed traits to occur simultaneously and breeding those traits into a single cell-line. For protoplast-mediated fusion, protoplasts of the cells are generated and these cells are forced to artificially fuse. Surviving fused cells divide and the chromosomes (and traits) are distributed between the cells. The Protoplast generation developed with MK29404 for genetic transformation and nucleic acid preparation was used to develop a protoplast fusion protocol. Specifically, MK29404 strains of interest to be fused were transferred from cryovials into 250 mL flat-bottom flasks containing 25 mL YPD medium and incubated at 27°C with shaking at 200 rpm until cultures reached a density of $OD_{600} > 20$ (~48 hrs of growth). Sub-cultures were then generated from dense (~2-day) cultures at 1:1500 – 1:5000 (depending on strain growth rate) into 25 mL YPD and incubated at 27°C with shaking at 200 rpm overnight (~16 hrs) for intended early-log phase growth and OD_{600} readings taken (should be @ $OD_{600} \sim 1 - 2.5$). Harvested 60-100 OD_{600} units of each strain, ensuring equal quantities to one another, by centrifuging at 1,250 x g for 5 min and supernatant removed. Cells were then washed with 25 mL sterile diH_2O and centrifuged at 1,250 x g for 5 min and supernatant removed. Followed by another washing step with 25 mL 25 mM β -mercaptoethanol, 5 mM EDTA solution, vortexed to resuspend, and placed on gyrator/rocker at room temperature for 20 min and centrifuged at 1,250 x g for 5 min and removed supernatant. Cells were then washed with 25 mL 0.7 M NaCl, vortexed to resuspend, and centrifuged at 1,250 x g for 5 min and supernatant removed. Resuspended cell pellets with 12 mL ‘MPPE’

protoplasting enzyme mix by pipetting, and transferred entire volume of each strain into separate 250 mL flat-bottom flasks; placed at 30°C with shaking at 100 rpm for 1.5 - 2 hrs. Transferred cell mixtures into separate 50 mL conical tubes, and spun as in step 4; removed all traces of supernatant by pipetting. Washed cell pellets with 12 mL 1 M sorbitol by pipetting, centrifuged at 1,250 x g for 5 min and removed the supernatant. Resuspended cell pellets by pipetting with 250 µL STC media and set up fusion reactions in 15 ml conical tubes as follows: (i) Parental control-1: 150 µL Strain 'A' only; (ii) Parental control-2: 150 µL Strain 'B' only; and (iii) Fusion reaction: 150 µL Strain 'A' + 150 µL Strain 'B'. Added 8-10 volumes of PEG solution to each tube (1.25 mL for controls; 2.5 mL for fusion reaction) and mixed well by pipetting gently. Incubated at room temperature for 20 min and harvested cell pellet at 800 x g for 5 min in swinging bucket rotor; carefully removed all traces of supernatant by pipetting. Resuspended cell pellet with 4 mL YNB/0.8 M sucrose, transferred to 25 mL flat-bottom shaker flask and incubated at 27°C with shaking at 100 rpm for 1 – 2 hrs and repeated. Carefully removed all but ~400 µL supernatant by pipetting. Resuspended cell pellet with remaining ~400 µL YNB/0.8 M sucrose supernatant and spread equal volumes onto appropriate solid medium for selection (e.g. YNB/0.8 M sucrose plates for fusing 2 uracil auxotrophs - see recipe above; use 2-3 plates per reaction); parafilm and incubated at 27°C until colonies appeared (3 – 5 days). Once colonies reached a workable size (~1-2 mm), picked isolated colonies and streaked for singles onto appropriate selection medium (same as used for initial selection) and incubated at 27°C until single colonies appeared on streakouts (3-5 days). From streakout plates, picked single isolated colonies and patched onto appropriate selection medium and incubated at 27°C until enough biomass exited (~3 days) to inoculate into liquid media for further evaluation.

It will be apparent to those skilled in the art that various modifications and variations can be made in the disclosed structures and methods without departing from the scope or spirit of the invention. Particularly, descriptions of any one embodiment can be freely combined with descriptions or other embodiments to result in combinations and/or variations of two or more elements or limitations. Other embodiments of the invention will be apparent to those skilled in the art from consideration of the specifications and practice of the invention disclosed herein. It is intended that the specification and examples be considered exemplary only, with a true scope and spirit of the invention being indicated by the following claims.

CLAIMS

What is claimed is:

1. A recombinant *Sporidiobolus* strain comprising at least one expression
5 cassette functional in *Sporidiobolus* that is stably integrated into its genome for the
production of homologous or heterologous proteins.
2. The recombinant *Sporidiobolus* strain of claim 1, wherein the at least
one expression cassette comprises, operably linked in the direction of transcription, at
10 least one promoter functional in *Sporidiobolus*, at least one nucleic acid sequence
encoding a homologous or heterologous protein and at least one terminator region
functional in *Sporidiobolus*.
3. The recombinant *Sporidiobolus* strain according to claim 2, wherein
15 the promoter is a promoter of a *Sporidiobolus* gene.
4. The recombinant *Sporidiobolus* strain according to claim 2, wherein
the terminator region is a terminator of a *Sporidiobolus* gene.
- 20 5. The recombinant *Sporidiobolus* strain according to any one of claims
2-4, wherein the promoter is a URA3 promoter.
6. The recombinant *Sporidiobolus* strain according to any one of claims
2-5, wherein the terminator is a URA3 terminator.
25
7. The recombinant *Sporidiobolus* strain according to any one of claims
2-6, wherein said nucleic acid sequence encoding a homologous protein and the
homologous protein encodes at least one protein chosen from Ku70 protein, Ura3
protein, eGFP protein, Malic Enzyme Isoform 1, and Malic Enzyme Isoform 2, from
30 *Sporidiobolus*.
8. The recombinant *Sporidiobolus* strain according to any one of claims
2-7, wherein said nucleic acid sequence encoding a heterologous protein encodes at

least one protein chosen from a fungal protein, a bacterial protein, a plant protein, an animal protein, and a protein from unidentified origin.

9. A method for producing a recombinant *Sporidiobolus* strain
5 comprising:
a) generating protoplasts of *Sporidiobolus*,
b) transforming the protoplasts generated in step a) with at least one vector comprising an expression cassette encoding a homologous or heterologous protein,
c) selecting recombinant *Sporidiobolus* transformed in step b) under specific
10 culture conditions, and
d) recovering recombinant *Sporidiobolus* containing the expression cassette of the vector stably integrated into their genome.
10. The method of claim 9, wherein the transformation in step b) is in the
15 presence of PEG.
11. The method of claim 9, wherein the transformation in step b) is by electroporation.
- 20 12. The method of any one of claims 9-11, wherein the protoplasts are generated by an enzyme mixture.
13. The method of claim 12, wherein the enzyme mixture comprises a chitinase.
25
14. The method of claim 13, wherein the chitinase is a *Streptomyces* chitinase.
15. The method of claim 14, wherein the *Streptomyces* chitinase is a
30 *Streptomyces griseus* chitinase.
16. The method of claim 12, wherein the enzyme mixture comprises a lysing enzyme.

17. The method of claim 16, wherein the lysing enzyme is a *Trichoderma* lysing enzyme.

18. The method of claim 17, wherein the *Trichoderma* lysing enzyme is a
5 *Trichoderma harzianum* lysing enzyme.

19. The method of claim 12, wherein the enzyme mixture comprises a driselase.

10 20. The method of claim 19, wherein the driselase is a *Basidiomycetes* driselase.

21. The method of claim 12, wherein the enzyme mixture comprises a lysing enzyme, a chitinase, and a driselase.

15

22. The method of any one of claims 9-21, wherein the specific culture conditions comprise culture conditions used in outgrowth, selection, and propagation of transformed *Sporidiobolus pararoseus* strains.

20 23. The method of any one of claims 9-22, wherein the specific culture conditions comprise a yeast culture.

24. The method of any one of claims 9-23, wherein the specific culture conditions comprise patching and gridding transformations onto at least one of
25 YNBglc, YNBglc+ura, and YPD plates, and incubating the transformations.

25. The method of any one of claims 9-24, wherein the specific culture conditions comprise screening transformations using PCR assays.

30 26. A recombinant cell derived from a protoplast of the species *Sporidiobolus* and genetically transformed with a nucleic acid sequence.

27. The cell of claim 26 wherein the cell is transformed with a homologous nucleic acid sequence.

28. The cell of claim 26 wherein the cell is transformed with a heterologous nucleic acid sequence.

5 29. The cell as in either of claims 27 and 28, wherein the nucleic acid sequence includes a marker gene.

30. The cell of any one of claims 26-29, wherein the cell expresses genes that have been codon-optimized or optimized for expression in *Sporidiobolus*.

10

31. The cell of any one of claims 26-30, wherein protoplast is generated by an enzyme mixture.

15 32. The cell of claim 31, wherein the enzyme mixture comprises a chitinase.

33. The cell of claim 32, wherein the chitinase is a *Streptomyces* chitinase.

20 34. The cell of claim 33, wherein the *Streptomyces* chitinase is a *Streptomyces griseus* chitinase.

35. The cell of any one of claims 31-34, wherein the enzyme mixture comprises a lysing enzyme.

25 36. The cell of claim 35, wherein the lysing enzyme is a *Trichoderma* lysing enzyme.

37. The cell of claim 36, wherein the *Trichoderma* lysing enzyme is a *Trichoderma harzianum* lysing enzyme.

30

38. The cell of any one of claims 31-37, wherein the enzyme mixture comprises a driselase.

39. The cell of claim 38, wherein the driselase is a *Basidiomycetes* driselase.

40. The cell of any one of claims 31-39, wherein the enzyme mixture
5 comprises a lysing enzyme, a chitinase and a driselase.

41. Residual biomass from the recombinant cell of any one of claims 26 to
40.

10 42. The residual biomass of claim 41, wherein the residual biomass is wet biomass.

43. The residual biomass of claim 41, wherein the residual biomass has
been dried via spray drying, drum drying, rotary vacuum filter, fluidized-bed drying,
15 oven drying, freeze drying/lyophilization, pneumatic drying, or any combination
thereof.

44. A recombinant cell derived from a protoplast of the species
Sporidiobolus and genetically transformed with a nucleic acid sequence, wherein the
20 cell is obtained by:

- a) isolating a cell suspension comprising *Sporidiobolus* cells,
- b) treating the cell suspension to obtain protoplasts,
- c) transforming the protoplasts, and
- d) regenerating cells from the transformed protoplasts.

25 45. The cell of claim 44, wherein in step c), the protoplasts are transformed
using electroporation.

46. The cell of claim 44, wherein in step c), the protoplasts are transformed
30 using PEG.

47. The cell of claim 44, wherein the protoplasts are transformed with a
homologous nucleic acid sequence.

48. The cell of claim 44, wherein the protoplasts are transformed with a heterologous nucleic acid sequence.

49. The cell of claim 48, wherein the heterologous nucleic acid sequence
5 comprises MK29404 URA3.

50. Oil produced from the recombinant cell of claim 44.

51. An expression vector construct for the production of a polypeptide in
10 *Sporidiobolus* comprising:

- a) a nucleic acid sequence encoding said polypeptide; and
- b) nucleic acid sequences allowing for expression of the polypeptide in *Sporidiobolus*.

52. The vector construct according to claim 51 wherein the nucleic acid
15 sequence is homologous.

53. The vector construct according to claim 51 wherein the nucleic acid
20 sequence is heterologous.

54. The vector construct according to any one of claims 51-53, wherein the
nucleic acid sequence confers resistance to an antibiotic.

55. The vector construct according to any one of claims 51-54, wherein the
25 nucleic acid sequence confers resistance to phleomycin.

56. The vector construct according to any one of claims 51-55, wherein the
nucleic acid sequence encodes a *Streptoalloteichus hindustanus ble* gene.

57. The vector construct according to any one of claims 51-56, further
30 comprising:

- a) the nucleic acid sequence encoding a polypeptide;
- c) a promoter in reading frame with the coding sequence; and
- d) a terminator.

58. The vector construct of claim 57, wherein said polypeptide confers resistance to an antibiotic.

5 59. The vector construct of claim 57, wherein said polypeptide confers resistance to phleomycin.

60. The vector construct of any one of claims 57-59, wherein said polypeptide has 95% sequence identity to SEQ ID NO: 1.

10

61. The vector construct of any one of claims 57-60, wherein said nucleic acid sequence comprises an intron.

62. The vector construct of claim 61, wherein said intron comprises a nucleic acid sequence that is homologous to said nucleic acid sequence encoding the polypeptide.

15

63. The vector construct of claim 61, wherein said intron comprises a nucleic acid sequence that is heterologous to said nucleic acid sequence encoding the polypeptide.

20

64. The vector construct of any one of claims 61 to 63 wherein said intron comprises a URA3 ORF.

25 65. The vector construct according to any one of claims 51-64 comprising a polynucleotide coding sequence under the control of a URA3 promoter.

66. The vector construct according to any one of claims 51-64 comprising a polynucleotide coding sequence under the control of a URA3 terminator.

30

67. The vector construct according to any one of claims 51-66, comprising:

c) URA3 promoter;

a) a phleomycin resistance gene; and a

d) URA 3 terminator.

68. The vector construct according to any one of claims 51-66, comprising:

- 5 c) a promoter;
a) a heterologous nucleic acid sequence encoding a polypeptide wherein said DNA sequence comprises an intron; and
d) a terminator.

10 69. A *Sporidiobolus* cell comprising the vector construct according to any one of claims 51-68.

70. A *Sporidiobolus* host cell wherein the host cell has been modified to reduce or eliminate expression of a functional gene product of a nucleic acid sequence
15 encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO: 72.

71. A *Sporidiobolus* host cell wherein the host cell has been modified to reduce or eliminate expression of a functional gene product of a nucleic acid sequence encoding a polypeptide having the amino acid sequence set forth in SEQ ID NO: 82.

20

72. A method for combining traits from various strains into a single cell-line, comprising:

- a) generating protoplasts of each strain of *Sporidiobolus* of interest,
b) forcing the protoplasts generated in step a) to artificially fuse, and
25 c) harvesting surviving fused cells.

73. The method of claim 72, wherein the strains of *Sporidiobolus* are obtained from classical strain improvement, metabolic engineering, or combinations thereof.

30

74. The method of claim 72, wherein the protoplasts are generated by an enzyme mixture.

75. The method of claim 74, wherein the enzyme mixture comprises a chitinase.

76. The method of claim 75, wherein the chitinase is a *Streptomyces* chitinase.

77. The method of claim 76, wherein the *Streptomyces* chitinase is a *Streptomyces griseus* chitinase.

78. The method of claim 74, wherein the enzyme mixture comprises a lysing enzyme.

79. The method of claim 78, wherein the lysing enzyme is a *Trichoderma* lysing enzyme.

80. The method of claim 79, wherein the *Trichoderma* lysing enzyme is a *Trichoderma harzianum* lysing enzyme.

81. The method of claim 74, wherein the enzyme mixture comprises a driselase.

82. The method of claim 81, wherein the driselase is a *Basidiomycetes* driselase.

83. The method of claim 74, wherein the enzyme mixture comprises a lysing enzyme, a chitinase, and a driselase.

84. The method of any one of claims 72-83, wherein the fusion reaction in step b) comprises an equal amount of each of the strains of *Sporidiobolus* of interest.

85. The method of any one of claims 72-84, wherein the fusion reaction in step b) is in the presence of PEG.

86. The method of any one of claims 72-85, wherein the harvesting in step c) comprises spreading the artificially fused protoplasts onto YPD plates, and incubating the artificially fused protoplasts.

5 87. A method of extracting recombinant cells suitable in production of biofuels from an aqueous fermentation broth, comprising:

extracting recombinant cells from the aqueous fermentation broth, wherein the broth contains oleaginous microorganisms, leaving residual biomass solids of recombinant cells that expressed one or more genes selected from the group
10 consisting of malic enzyme, acetyl-CoA synthetase, ATP citrate lyase, and phytoene dehydrogenase genes after oil extraction, and residual broth water.

88. The method of claim 87, further comprising using at least part of the residual broth water as imbibition water for washing a process feedstock to extract
15 sugar.

89. The method of claim 87 or 88, wherein the recombinant cells are derived from a protoplast of the species *Sporidiobolus* and are genetically transformed with a nucleic acid sequence.

20

90. The method of any one of claims 87 to 89, further comprising pasteurizing the aqueous fermentation broth.

91. The method of claim 90, comprising pasteurizing the aqueous
25 fermentation broth by heating the aqueous fermentation broth to between about 40°C and about 80°C for about 1 minute up to about 3 hours.

92. The method of any one of claims 87 to 91, further comprising heating the aqueous fermentation broth for about 30 minutes to about 18 hours, or more than 3
30 hours to about 18 hours, or more than 3 hours to about 8 hours.

93. The method of claim 92, further comprising holding the aqueous fermentation broth at a temperature between about 90°C and about 150°C, or between about 100°C and about 150°C, or between about 110°C and about 150°C, or

between about 120°C and about 150°C, or between about 130°C and about 150°C for about 30 minutes to about 18 hours, or more than 3 hours to about 18 hours, or more than 3 hours to about 8 hours.

5 94. The method of claim 93, further comprising stirring the aqueous fermentation broth during the heating interval.

 95. The method of any one of claims 87 to 94, further comprising adding an acid to the aqueous fermentation broth.

10

 96. The method of any one of claims 87 to 95, further comprising adding a base to the aqueous fermentation broth.

 97. The method of any one of claims 87 to 96, further comprising adding a
15 salt to the aqueous fermentation broth.

 98. The method of any one of claims 87 to 97, further comprising passing the aqueous fermentation broth through a bead mill, a homogenizer, an orifice plate, a high-shear mixer, a press, an extruder, pressure disruption, wet milling, dry milling, or
20 other shear or mechanical disruption device at least once to achieve a lysed fermentation broth.

 99. The method of claim 98, comprising passing the aqueous fermentation broth through a bead mill, a homogenizer, an orifice plate, a high-shear mixer, a press, an extruder, pressure disruption, wet milling, dry milling, or other shear or mechanical
25 disruption device at least twice to achieve a lysed fermentation broth.

 100. The method of claims 98 or 99, further comprising stirring the lysed fermentation broth in a vessel at about 70°C to about 100°C for about 1 to about 60
30 hours.

 101. The method of claim 100, further comprising adding a salt to the lysed fermentation broth in the vessel.

102. The method of claim 101, comprising adding up to about 2% by weight of the salt to the lysed fermentation broth in the vessel.

103. The method of any one of claims 97, 101, or 102, wherein the salt
5 includes at least one of the group consisting of NaCl, Na₂SO₄, KCl, and K₂SO₄.

104. The method of any one of claims 100 to 103, further comprising adding a base to adjust a pH of the lysed fermentation broth in the vessel to between about 3 and about 11.

10

105. The method of any one of claims 87 to 104, further comprising separating the recombinant cells from the lysed fermentation broth through centrifugation.

15 106. The method of any one of claims 87 to 105, wherein the aqueous fermentation broth comprises sugarcane extract.

107. The method of any one of claims 87 to 105, further comprising recycling the biomass solids with the residual broth water.

20

1 / 15

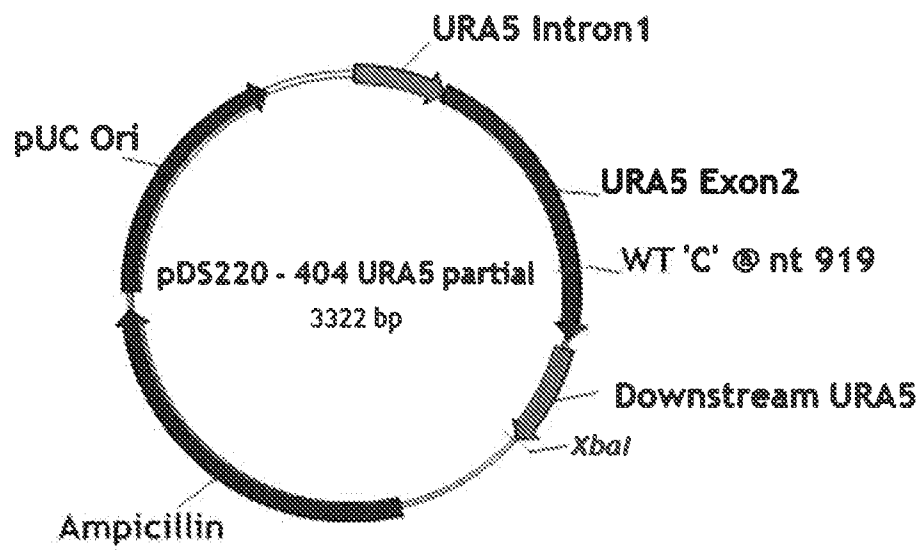


FIG. 1

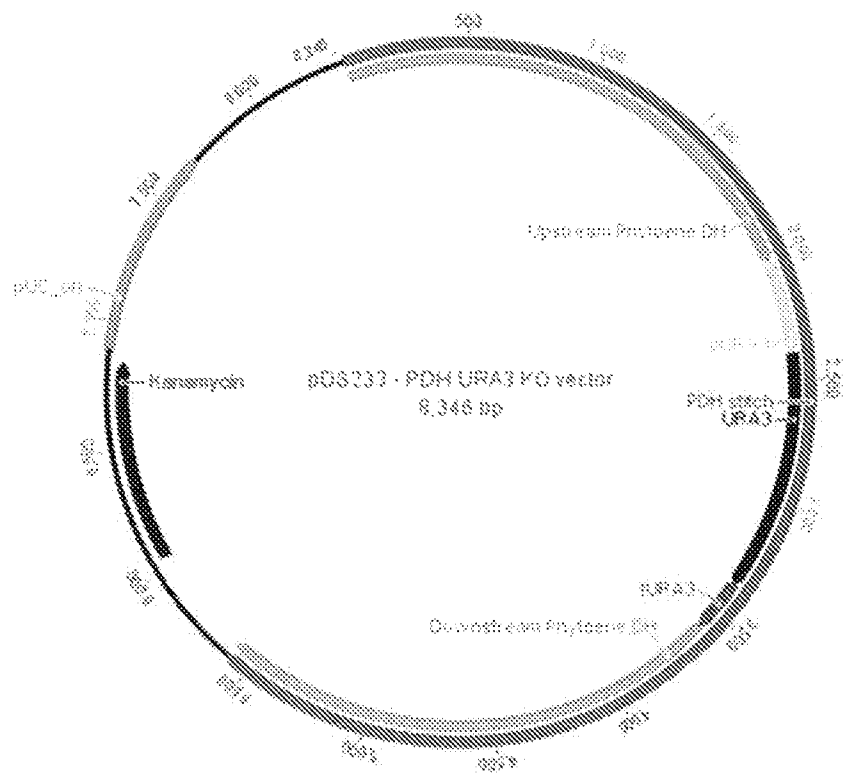


FIG. 2

2 / 15

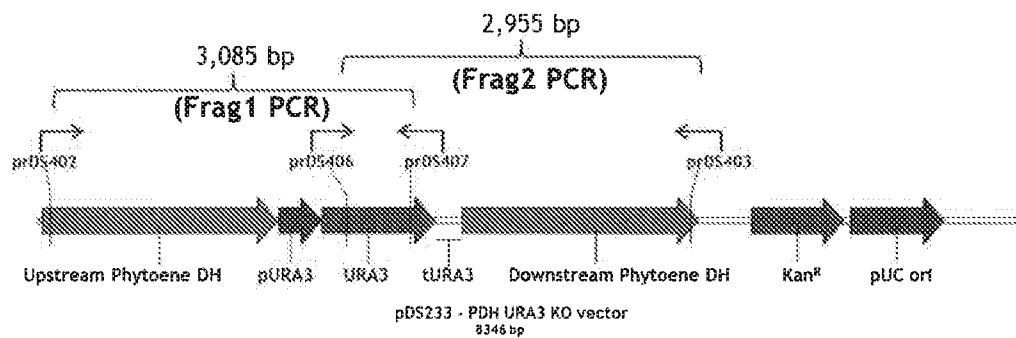


FIG. 3

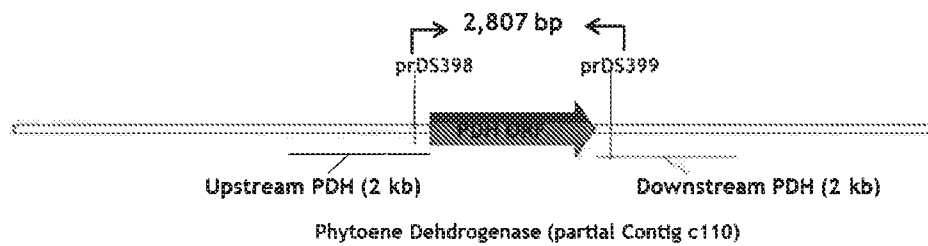


FIG. 4

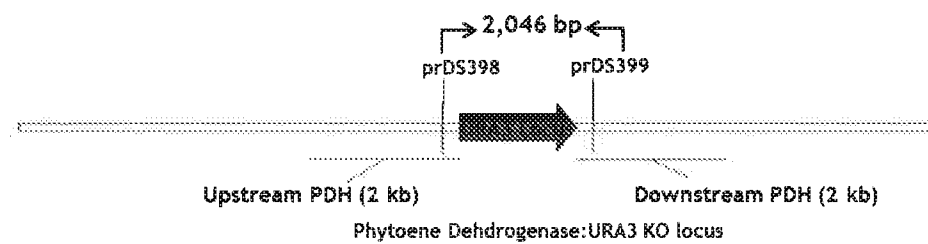


FIG. 5

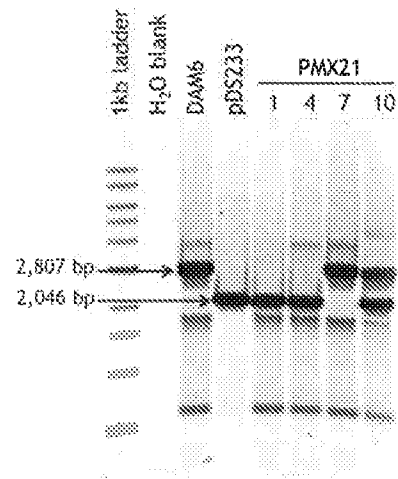


FIG. 6

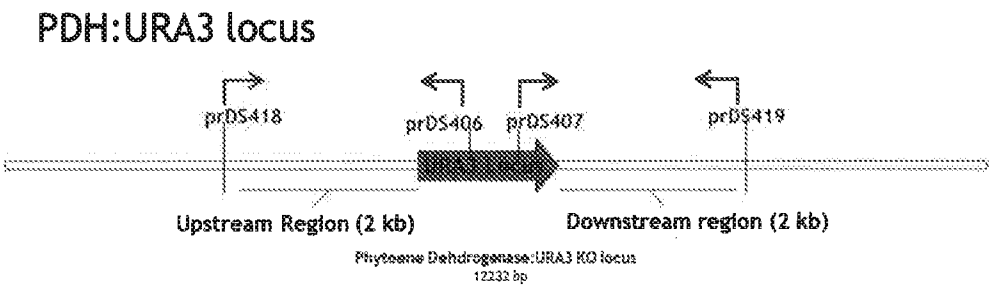


FIG. 7

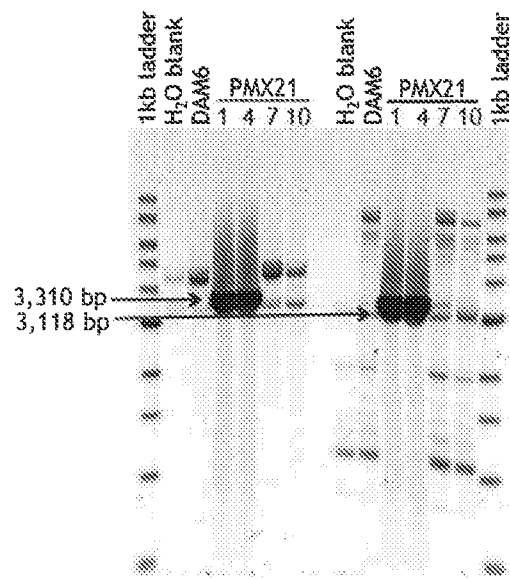


FIG. 8

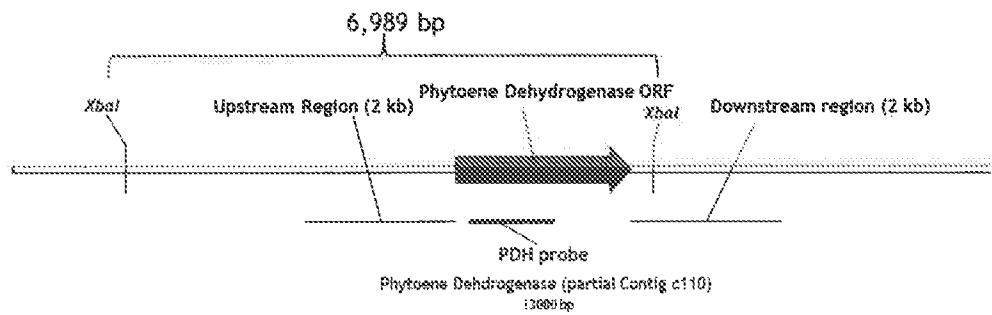


FIG. 9

5 / 15

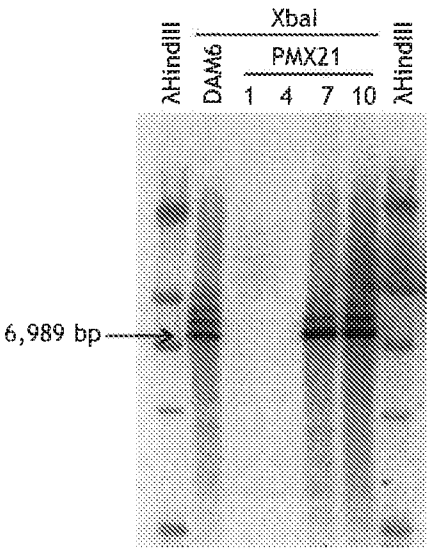


FIG. 10

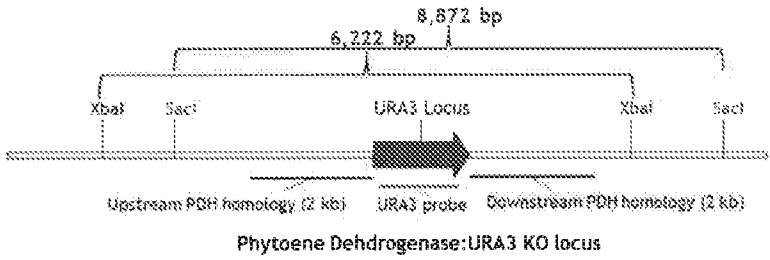


FIG. 11

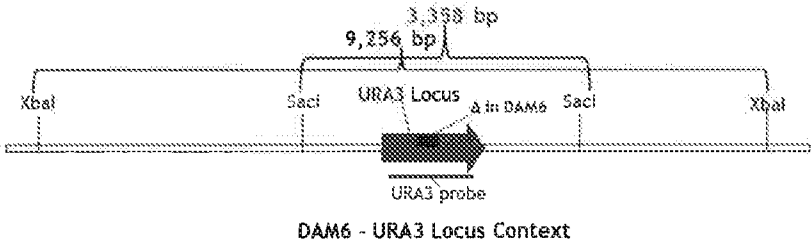


FIG. 12

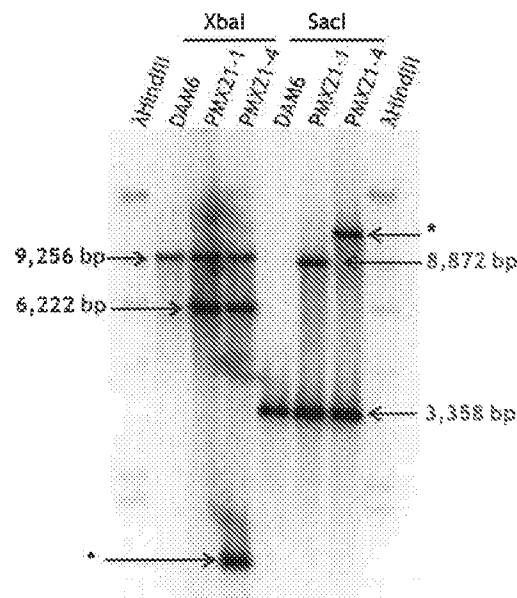


FIG. 13

7/15

	50	
Coprinus cinerea Ku70 / XP_001836717	(34)	
Crpto neoformans Ku70	(33)	
S para Ku70 putative AA seq	(31)	
Ustilago mydis Ku70 / NW_101153	(1)	MPKAYFVHKRE.....
	51	
Coprinus cinerea Ku70 / XP_001836717	(40)	
Crpto neoformans Ku70	(47)	
S para Ku70 putative AA seq	(39)	
Ustilago mydis Ku70 / NW_101153	(51)	
	101	
Coprinus cinerea Ku70 / XP_001836717	(53)	
Crpto neoformans Ku70	(54)	
S para Ku70 putative AA seq	(59)	
Ustilago mydis Ku70 / NW_101153	(64)	
	151	
Coprinus cinerea Ku70 / XP_001836717	(93)	
Crpto neoformans Ku70	(133)	
S para Ku70 putative AA seq	(139)	
Ustilago mydis Ku70 / NW_101153	(93)	
	201	
Coprinus cinerea Ku70 / XP_001836717	(133)	
Crpto neoformans Ku70	(153)	
S para Ku70 putative AA seq	(178)	
Ustilago mydis Ku70 / NW_101153	(140)	
	251	
Coprinus cinerea Ku70 / XP_001836717	(146)	
Crpto neoformans Ku70	(203)	
S para Ku70 putative AA seq	(222)	
Ustilago mydis Ku70 / NW_101153	(181)	
	301	
Coprinus cinerea Ku70 / XP_001836717	(184)	
Crpto neoformans Ku70	(241)	
S para Ku70 putative AA seq	(263)	
Ustilago mydis Ku70 / NW_101153	(231)	
	351	
Coprinus cinerea Ku70 / XP_001836717	(218)	
Crpto neoformans Ku70	(268)	
S para Ku70 putative AA seq	(295)	
Ustilago mydis Ku70 / NW_101153	(231)	
	401	
Coprinus cinerea Ku70 / XP_001836717	(251)	
Crpto neoformans Ku70	(302)	
S para Ku70 putative AA seq	(328)	
Ustilago mydis Ku70 / NW_101153	(331)	
	451	
Coprinus cinerea Ku70 / XP_001836717	(278)	
Crpto neoformans Ku70	(338)	
S para Ku70 putative AA seq	(364)	
Ustilago mydis Ku70 / NW_101153	(381)	
	501	
Coprinus cinerea Ku70 / XP_001836717	(316)	
Crpto neoformans Ku70	(379)	
S para Ku70 putative AA seq	(403)	
Ustilago mydis Ku70 / NW_101153	(428)	
	551	
Coprinus cinerea Ku70 / XP_001836717	(335)	
Crpto neoformans Ku70	(429)	
S para Ku70 putative AA seq	(425)	
Ustilago mydis Ku70 / NW_101153	(448)	
	601	
Coprinus cinerea Ku70 / XP_001836717	(350)	
Crpto neoformans Ku70	(479)	
S para Ku70 putative AA seq	(446)	
Ustilago mydis Ku70 / NW_101153	(463)	
	651	
Coprinus cinerea Ku70 / XP_001836717	(388)	
Crpto neoformans Ku70	(529)	
S para Ku70 putative AA seq	(499)	
Ustilago mydis Ku70 / NW_101153	(512)	
	701	
Coprinus cinerea Ku70 / XP_001836717	(430)	
Crpto neoformans Ku70	(576)	
S para Ku70 putative AA seq	(539)	
Ustilago mydis Ku70 / NW_101153	(556)	
	751	
Coprinus cinerea Ku70 / XP_001836717	(478)	
Crpto neoformans Ku70	(635)	
S para Ku70 putative AA seq	(588)	
Ustilago mydis Ku70 / NW_101153	(605)	
	801	
Coprinus cinerea Ku70 / XP_001836717	(521)	
Crpto neoformans Ku70	(678)	
S para Ku70 putative AA seq	(633)	
Ustilago mydis Ku70 / NW_101153	(650)	
	851	
Coprinus cinerea Ku70 / XP_001836717	(546)	
Crpto neoformans Ku70	(722)	
S para Ku70 putative AA seq	(688)	
Ustilago mydis Ku70 / NW_101153	(696)	
	901	
Coprinus cinerea Ku70 / XP_001836717	(560)	
Crpto neoformans Ku70	(752)	
S para Ku70 putative AA seq	(731)	
Ustilago mydis Ku70 / NW_101153	(714)	

FIG. 14

8 / 15

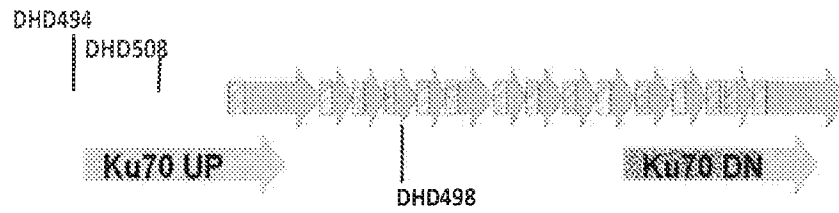


FIG. 15

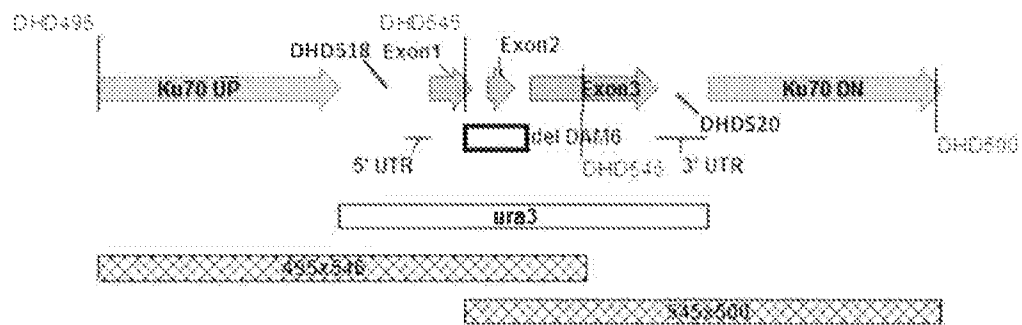


FIG. 16

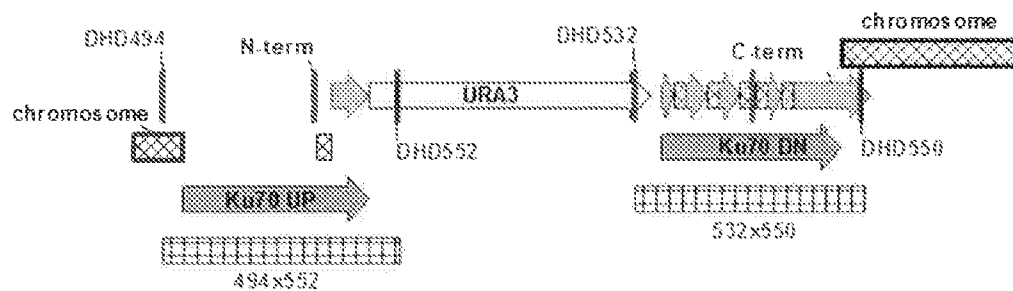


FIG. 17

9/15

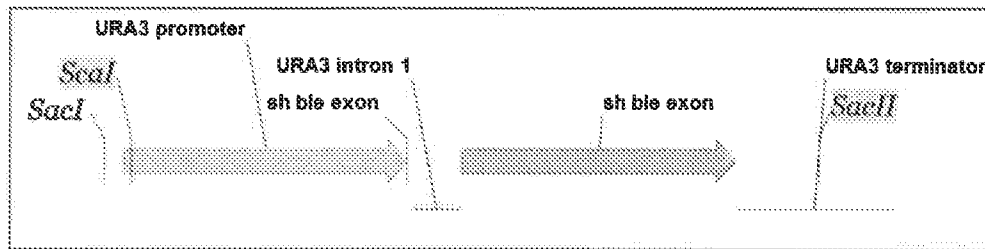


FIG. 18

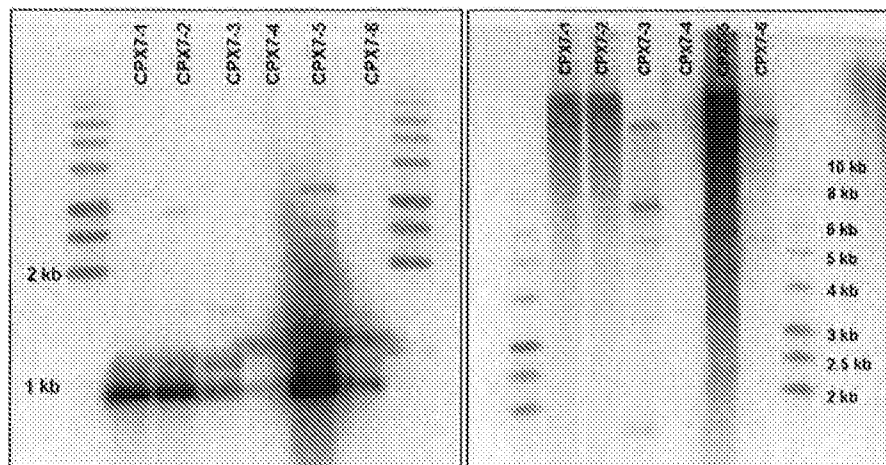


FIG. 19

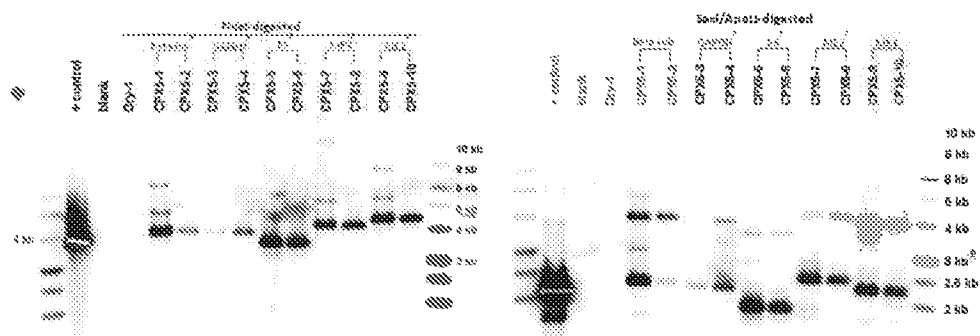


FIG. 20

10 / 15

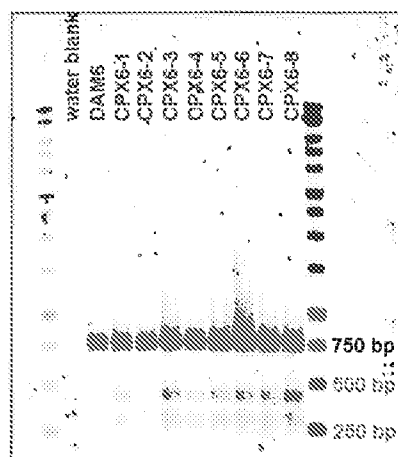


FIG. 21

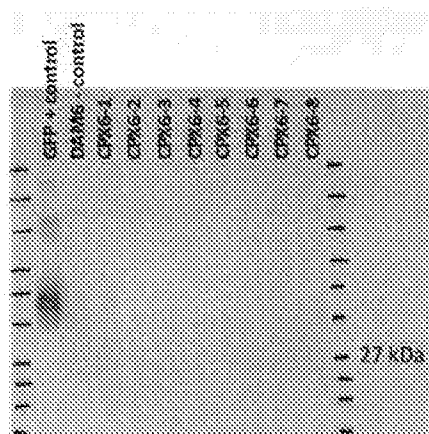


FIG. 22

001	GGTGTCTTCG	CTATCTCTGT	AGAGCTGTAT	GCTCATCTCA	AGGATCTTAA	GTTCTTCGTT	CTCTGGAGAG	CTGAGAGAGA	TGCTACTTTC	GAAGAGCTCA
101	GGCTGCAATT	CGATCGGACT	AGGAGGAGAT	TCCTCTCTCC	TTCCTCTACT	CGCTGATCCA	CGCTGATCCA	CGGAGTCCCA	GGTCTCTGAG	GGTCTCTGAG
201	CGAGGAGAGG	CAGACATGAT	TCCTCTCTCT	GGGAGGCTTC	GGGAGGCTTC	TTCAGAGGCT	TTCAGAGGCT	TTCAGAGGCT	AGGAGAGCTA	CAGACATGAT
301	CGGAGGATTC	AGGAGGAGAT	AGGAGGAGAT	GTTAGGCTTC	TTCAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	AGGAGGCTTC
401	AGGAGGATTC	CTCTCTCTCT	CTCTCTCTCT	TTCAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	AGGAGGCTTC
501	TGAGCTCTCT	CGGAGGCTTC	AGGAGGCTTC	CTCTCTCTCT	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	AGGAGGCTTC
601	GGGAGGCTTC	AGGAGGCTTC	CTCTCTCTCT	CTCTCTCTCT	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	AGGAGGCTTC
001	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC
101	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC
201	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC
301	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC
401	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC
501	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC
601	GGGAGGCTTC	CTATCTCTGT	AGGAGGCTTC	GGGAGGCTTC	AGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC	GGGAGGCTTC

FIG. 23

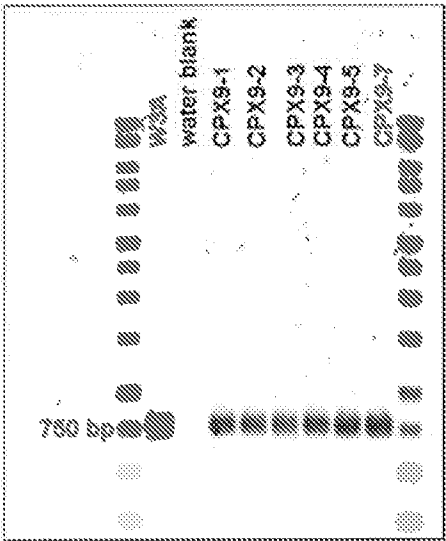


FIG. 24

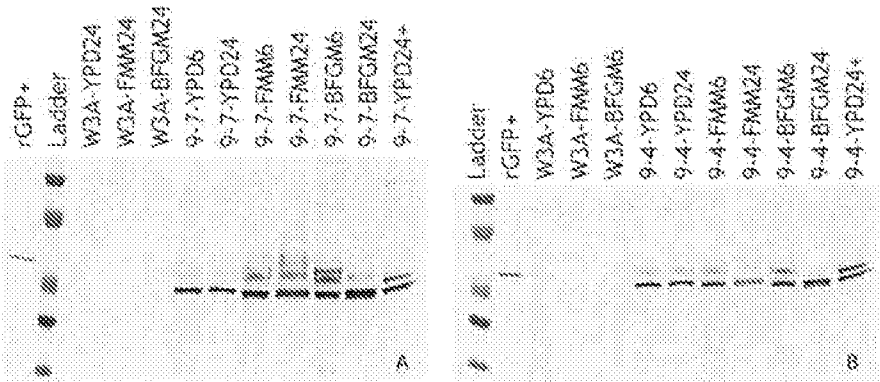


FIG. 25

12 /15

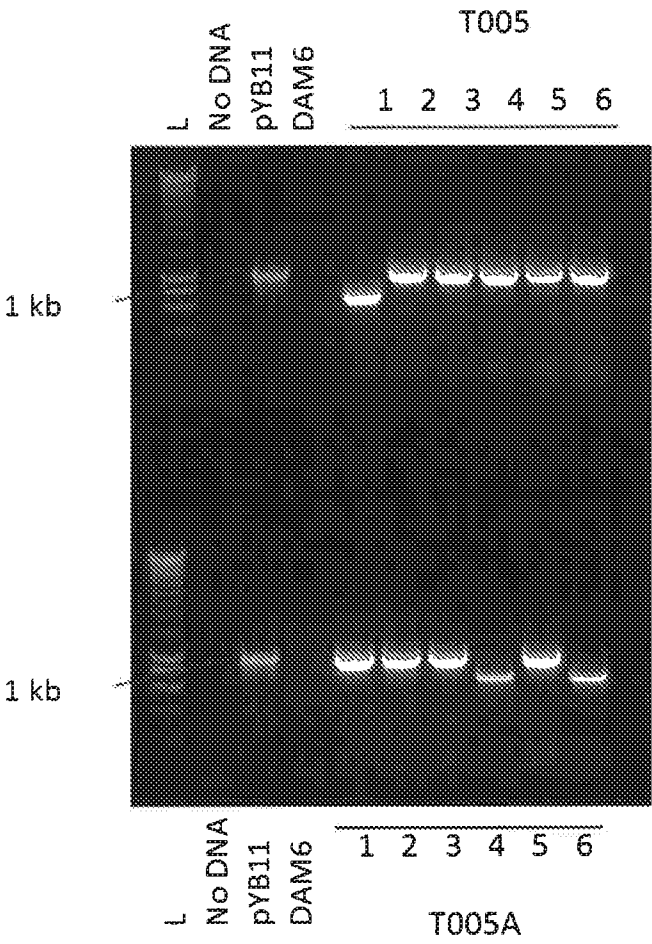


FIG. 26

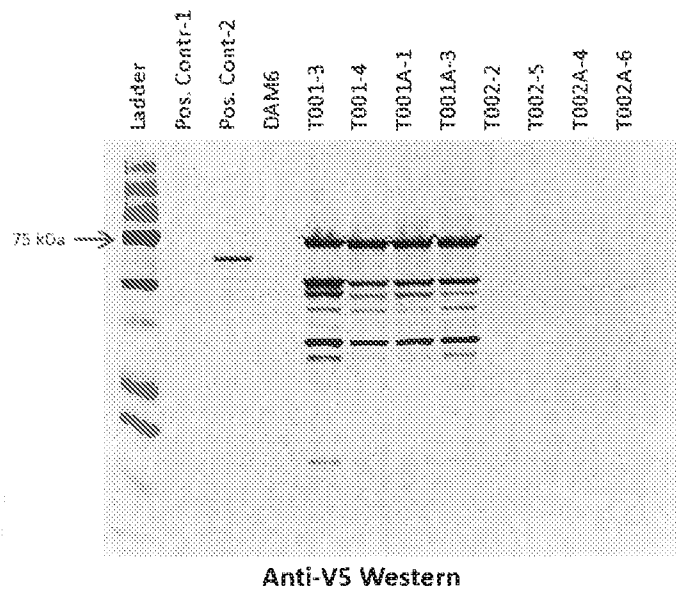


FIG. 27

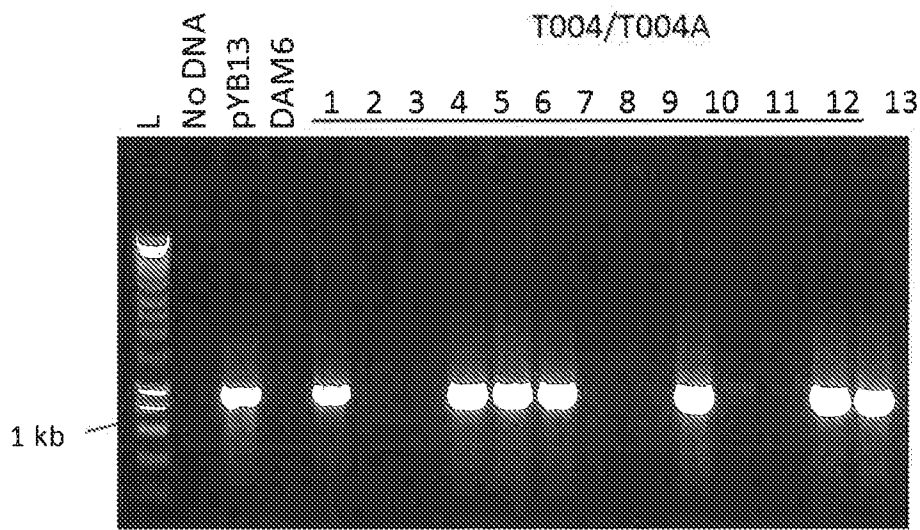


FIG. 28

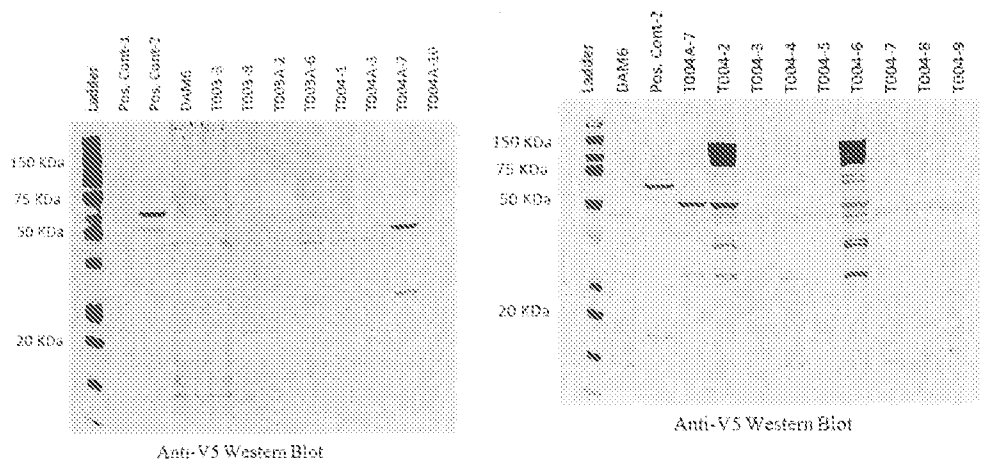


FIG. 29

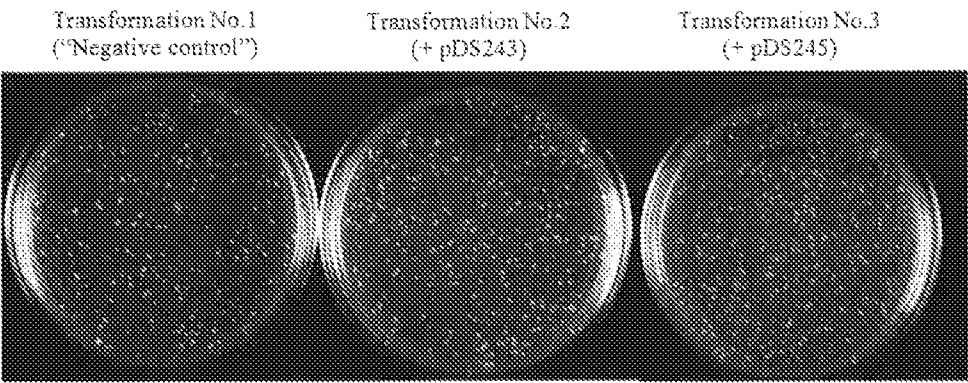


FIG. 30

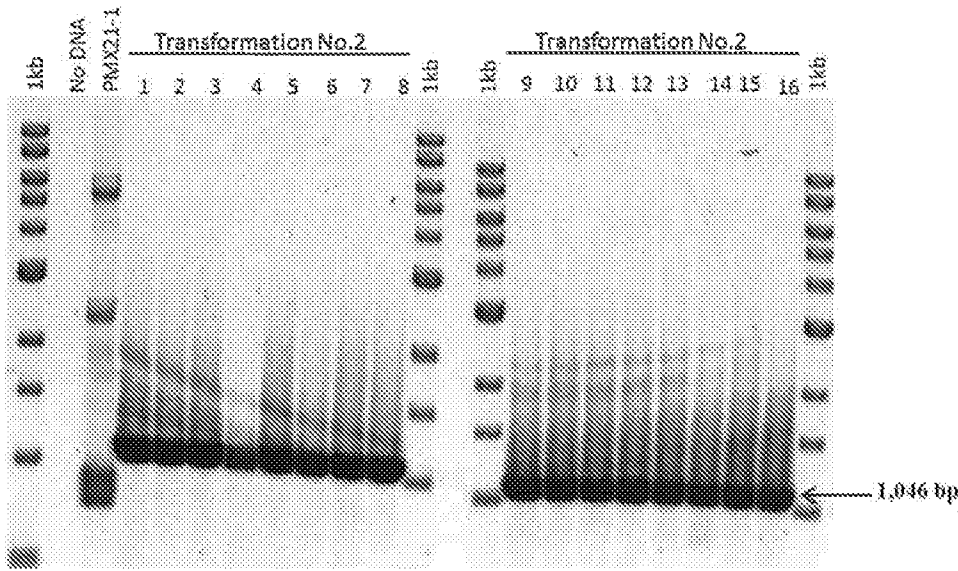


FIG. 31

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2014/071165

A. CLASSIFICATION OF SUBJECT MATTER INV. C12N1/02 C12N1/16 C12N15/64 C12P7/64 C12N15/81 ADD.		
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) C12N C12P		
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) EPO-Internal, WPI Data, Sequence Search, BIOSIS		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/112948 A1 (MARTEK BIOSCIENCES CORP [US]; APT KIRK E [US]; BARCLAY WILLIAM R [US];) 15 September 2011 (2011-09-15)	50
Y	in particular see claims 1-5, 22-23, 41,43 and pages 4-5, 6-12; the whole document	1-107
X	----- DE 10 2005 029170 A1 (CONSORTIUM ELEKTROCHEM IND [DE]) 28 December 2006 (2006-12-28)	1-4,9, 11, 51-55, 57-59
Y	in particular see [016]-[020] as well as claims 1-4 and 14 and [052] and [054] and examples ----- -/-	1-107
☒ 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 March 2015		Date of mailing of the international search report 07/04/2015
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Vix, Olivier

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/071165

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2012/169969 A1 (TEMASEK LIFE SCIENCES LAB LTD [SG]; JI LIANG HUI [SG]; LIU YAN BIN [SG] 13 December 2012 (2012-12-13)	1-4
Y	in particular see pages 4-5 and 7-10 and claims 1,9,; the whole document	1-107
X	----- EP 1 726 637 A1 (KANEKA CORP [JP]) 29 November 2006 (2006-11-29)	1
Y	in particular see claim 1, 7, 17-18, 20 and pages 7-8; the whole document	1-107
Y	----- WO 2011/112627 A1 (BP BIOFUELS UK LTD [GB]; PFEIFER JOSEPH W [US]; BEHRENS PAUL W [US]; J) 15 September 2011 (2011-09-15)	1-107
	in particular see pages 2-6, claims 1, 6, 11; the whole document	
Y	----- LIANG MING-HUA ET AL: "Advancing oleaginous microorganisms to produce lipid via metabolic engineering technology", PROGRESS IN LIPID RESEARCH, vol. 52, no. 4, 16 May 2013 (2013-05-16), pages 395-408, XP028769972, on line ISSN: 0163-7827, DOI: 10.1016/J.PLIPRES.2013.05.002 the whole document	1-107
Y	----- RADZIO R ET AL: "Synthesis of biotechnologically relevant heterologous proteins in filamentous fungi", PROCESS BIOCHEMISTRY, ELSEVIER SCIENCE PUBLISHERS LTD, GB, vol. 32, no. 6, 1 January 1997 (1997-01-01), pages 529-539, XP009137292, ISSN: 0032-9592, DOI: 10.1016/S0032-9592(97)00004-6 the whole document	1-107
Y	----- ZHIWEI ZHU ET AL: "A multi-omic map of the lipid-producing yeast Rhodosporidium toruloides", NATURE COMMUNICATIONS, vol. 3, 9 October 2012 (2012-10-09), page 1112, XP055178339, DOI: 10.1038/ncomms2112 the whole document	1-107
	----- -/--	

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/071165

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YANBIN LIU ET AL: "Characterization of glyceraldehyde-3-phosphate dehydrogenase gene RtGPD1 and development of genetic transformation method by dominant selection in oleaginous yeast <i>Rhodospiridium toruloides</i>", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 97, no. 2, 22 June 2012 (2012-06-22), pages 719-729, XP055120693, ISSN: 0175-7598, DOI: 10.1007/s00253-012-4223-9 see page 719,722-720,724-725 -----	1-107
Y	SEIKI TAKENO ET AL: "Establishment of an overall transformation system for an oil-producing filamentous fungus, <i>Mortierella alpina</i> 1S-4", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, vol. 65, no. 4, 12 May 2004 (2004-05-12), pages 419-425, XP055178299, ISSN: 0175-7598, DOI: 10.1007/s00253-004-1622-6 the whole document -----	1-107
Y	WYNN JAMES P ET AL: "The role of malic enzyme in the regulation of lipid accumulation in filamentous fungi", MICROBIOLOGY, SGM, GB, vol. 145, no. 8, 1 August 1999 (1999-08-01), pages 1911-1917, XP002537164, ISSN: 0002-5972, DOI: 10.1099/13500872-145-8-1911 illustrates the role of the malic enzyme in regulating the extent of lipid accumulation in filamentous fungi.; the whole document -----	1-107
Y	GOSWAMI RUBELLA S: "Targeted Gene Replacement in Fungi Using a Split-Marker Approach", METHODS IN MOLECULAR BIOLOGY, HUMANA PRESS INC, NJ, US, vol. 835, 1 January 2012 (2012-01-01), pages 255-269, XP009180950, ISSN: 1064-3745, DOI: 10.1007/978-1-61779-501-5_16 the whole document ----- -/--	1-107

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2014/071165

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DE BEKKER C ET AL: "An enzyme cocktail for efficient protoplast formation in <i>Aspergillus niger</i>", JOURNAL OF MICROBIOLOGICAL METHODS, ELSEVIER, AMSTERDAM, NL, vol. 76, no. 3, 1 March 2009 (2009-03-01), pages 305-306, XP025946281, ISSN: 0167-7012, DOI: 10.1016/J.MIMET.2008.11.001 [retrieved on 2009-02-13] the whole document -----	1-107
Y	MA ZHENG ET AL: "Formation, regeneration, and transformation of protoplasts of <i>Streptomyces diastatochromogenes</i> 1628", FOLIA MICROBIOLOGICA, PRAQUE, CZ, vol. 59, no. 2, 31 July 2013 (2013-07-31), pages 93-97, XP035323958, ISSN: 0015-5632, DOI: 10.1007/S12223-013-0271-5 [retrieved on 2013-07-31] the whole document -----	9-25

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2014/071165

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011112948 A1	15-09-2011	AU 2011224181 A1 CA 2792494 A1 CN 103210080 A EA 201201278 A1 EP 2545159 A1 JP 2013521781 A US 2012034190 A1 US 2013122557 A1 US 2014199755 A1 WO 2011112948 A1	11-10-2012 15-09-2011 17-07-2013 28-06-2013 16-01-2013 13-06-2013 09-02-2012 16-05-2013 17-07-2014 15-09-2011
DE 102005029170 A1	28-12-2006	DE 102005029170 A1 WO 2006136311 A1	28-12-2006 28-12-2006
WO 2012169969 A1	13-12-2012	AU 2012267241 A1 CN 104245938 A EP 2718442 A1 SG 195304 A1 US 2014113377 A1 WO 2012169969 A1	30-01-2014 24-12-2014 16-04-2014 30-12-2013 24-04-2014 13-12-2012
EP 1726637 A1	29-11-2006	EP 1726637 A1 US 2008233620 A1 WO 2005085415 A1	29-11-2006 25-09-2008 15-09-2005
WO 2011112627 A1	15-09-2011	AR 080486 A1 AU 2011224449 A1 CA 2791986 A1 CN 102971430 A EA 201201105 A1 EP 2545180 A1 JP 2013521778 A KR 20130016268 A US 2012323029 A1 WO 2011112627 A1	11-04-2012 27-09-2012 15-09-2011 13-03-2013 29-03-2013 16-01-2013 13-06-2013 14-02-2013 20-12-2012 15-09-2011