U.S. Patent No. 5,584,039 relates to a selectable marker gene *ADE2* isolated from *Pichia methanolica*. Piontek *et al.*, Appl Microbiol. Biotechnol. 50:331-338 (1998) relates to novel gene expression systems in *Schwanniomyces occidentalis* and *Pichia stipitis*, which systems utilize vectors containing an *ADE2* marker and a putative replication sequence. However, no corresponding gene has previously been isolated from *Pichia pastoris*, and the effects of transformation with *ADE2* in *P. pastoris* have not previously been identified.

Accordingly, a need exists for improved methods of transformation, selection and expression of heterogeneous genes using the *Pichia pastoris* yeast as the host expression system.

## BRIEF SUMMARY

The present disclosure provides methods and materials for the use of lower eukaryotic cells such as yeast or filamentous fungi as an expression system for expressing recombinant proteins.

The method is based on constructing slower growing *ade2* auxotrophic strains of the lower eukaryote cells and using integration vectors that are capable of integrating into the genome of the *ade2* auxotrophic strain and which comprises nucleic acids encoding an *ADE2* marker gene or open reading frame (ORF) operably linked to a promoter and a recombinant protein, wherein the integration vector integrates into the genome of the *ade2* auxotrophic strain, the *ADE2* renders the auxotrophic strain prototrophic for adenine, and the recombinant protein is expressed.

Thus, provided is an expression system comprising (a) a *Pichia pastoris* host cell in which the endogenous *ADE2* gene encoding Ade2p has been removed from the genome of the host cell; and (b) an integration vector comprising (1) a nucleic acid encoding the Ade2p; (2) a nucleic acid having an insertion site for the insertion of one or more expression cassettes comprising a nucleic acid encoding one or more heterologous peptides, proteins, and/or functional nucleic acids of interest, and (3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the host cell by homologous recombination.

In one aspect, provided is an expression system comprising:

(a) a *Pichia pastoris* host cell in which the endogenous gene or the open reading frame thereof encoding an auxotrophic selectable marker protein encoded by the *ADE1* gene or the ADE2 gene has been removed from the genome of the host cell to render the cell auxotrophic and

(b) an integration vector comprising:

(1) a nucleic acid comprising an ADE1 or ADE2 gene operably linked to an endogenous truncated ADE1 or ADE2 promoter, in which the truncation renders the promoter with a transcription activity that is less than the native endogenous promoter or no promoter;

(2) a nucleic acid having an insertion site for the insertion of one or more expression cassettes comprising a nucleic acid encoding a peptide, protein, and/or functional nucleic acid of interest, and

(3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the host cell by roll-in single-crossover homologous recombination.

Also, provided is a method for producing a recombinant *Pichia pastoris* host cell that expresses a heterologous protein or peptide comprising (a) providing the host cell in which the endogenous *ADE2* gene encoding an Ade2p has been removed from the genome of the host cell; and (a) transforming the host cell with an integration vector comprising (1) a nucleic acid encoding the Ade2p; (2) a nucleic acid having one or more expression cassettes comprising a nucleic acid encoding one or more heterologous peptides, proteins, and/or functional nucleic acids of interest, and (3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the host cell by homologous recombination, wherein the transformed host cell produces the recombinant protein.

In another aspect, provided is a method for producing a *Pichia pastoris* host cell that has multiple copies of an integration vector integrated into the genome of the host cell, comprising:

(a) providing *Pichia pastoris* cells in which an endogenous *ADE2* genehas been removed from the genome of the host cell to render the cells auxotrophic for adenine;

(b) transforming the cells by roll-in single crossover homologous recombination with an integration vector comprising:

(1) a nucleic acid comprising an ADE2 gene operably linked to an endogenous truncated ADE2 promoter in which the truncation renders the promoter with a transcription activity that is less than the native endogenous promoter or no promoter;

(2) a nucleic acid having one or more expression cassettes comprising a nucleic acid encoding one or more heterologous peptides, proteins, and/or functional nucleic acids of interest, and

(3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the cells by homologous recombination; wherein the auxotrophy of said host cell is not fully complemented by the ADE2 gene in the integration vector when a single copy of the vector is in the host cell;

(c) cultivating the transformed cells in a medium lacking adenine.

Further provided is an isolated nucleic acid comprising the *ADE2* gene of *Pichia pastoris*. The nucleic acid comprises the open reading frame that encodes the Ade2p protein or the nucleic acid has a nucleotide sequence with 95% identity to the nucleic acid sequence shown in SEQ ID NO:60 from nucleotide 127 to nucleotide 1,815. Further provided is an isolated polypeptide comprising an amino acid sequence with 95% identity to the amino acid sequence shown in SEQ ID NO:61.

The applicants further discovered that operably linking an auxotrophic marker gene or ORF to a minimal promoter in the integration vector, that is a promoter that has low transcriptional activity, enabled the production of recombinant host cells that contain a sufficient number of copies of the integration vector integrated into the genome of the auxotrophic host cell to render the cell prototrophic and which render the cells capable of producing amounts of the recombinant protein or functional nucleic acid of interest that are greater than the amounts that would be produced in a cell that contained only one copy of the integration vector integrated into the genome.

Therefore, provided is a method in which an auxotrophic strain of a lower eukaryote cell is obtained or constructed and an integration vector is provided that is capable of integrating into the genome of the auxotrophic strain and which comprises nucleic acids encoding a marker gene or ORF that compliments the auxotrophy and is operably linked to a weak promoter, an attenuated endogenous or heterologous promoter, a cryptic promoter, or a truncated endogenous or heterologous promoter and a recombinant protein. Host cells in which a number of the integration vectors have been integrated into the genome to compliment the auxotrophy of the host cell are selected in medium that lacks the metabolite that compliments the auxotrophy and maintained by propagating the host cells in medium that lacks the metabolite that compliments the auxotrophy or in medium that contains the metabolite because in that case, cells that evict the plasmids including the marker will grow more slowly.

In a further embodiment, provided is an expression system comprising (a) a host cell in which the endogenous gene encoding an auxotrophic selectable marker protein has been removed from the genome of the host cell; and (b) an integration vector comprising (1) a nucleic acid comprising an open reading frame (ORF) encoding a function that is complementary to the function of the endogenous gene encoding the auxotrophic selectable marker protein and which is operably linked to a weak promoter, an attenuated endogenous or heterologous promoter, a cryptic promoter, a truncated endogenous or heterologous promoter, or no promoter; (2) a nucleic acid having an insertion site for the insertion of one or more expression cassettes comprising a nucleic acid encoding one or more heterologous peptides, proteins, and/or functional nucleic acids of interest, and (3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the host cell by homologous recombination.

In a further still embodiment, provided is a method for expression of a recombinant protein in a host cell comprising (a) providing the host cell in which the endogenous gene encoding an auxotrophic selectable marker protein has been removed from the genome of the host cell; and (a) transforming the host cell with an integration vector comprising (1) a nucleic acid comprising an open reading frame (ORF) encoding a function that is complementary to the function of the endogenous gene encoding the auxotrophic selectable marker protein and which is operably linked to a weak promoter, an attenuated endogenous or heterologous promoter, a cryptic promoter, a truncated endogenous or heterologous promoter, or no promoter; (2) a nucleic acid having one or more expression cassettes comprising a nucleic acid encoding one or more heterologous peptides, proteins, and/or functional nucleic acids of interest, and (3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the host cell by homologous recombination, wherein the transformed host cell produces the recombinant protein.

In another aspect, provided is a method for expressing a recombinant peptide, protein and/or nucleic acid of interest in a *Pichia pastoris* host cell using the expression system as defined herein wherein the transformed host cell produces the recombinant peptide, protein and/or nucleic acid of interest.

In another aspect, provided is a *Pichia pastoris* cell comprising a deletion or disruption of the endogenous *ADE2* gene, or any other modification, rendering the *ADE2* gene or Ade2p non-functional and more than one copy of a vector comprising an *ADE2* open reading frame operably linked to an endogenous truncated ADE1 or ADE2 promoter in which the truncation renders the promoter with a transcription activity that is less than the native endogenous promoter or no promoter; wherein the auxotrophy of said host cell is not fully complemented by the ADE2 gene in the integration vector when a single copy of the vector is in the host cell.

The auxotrophic selectable marker protein is encoded by a gene selected from the group consisting of *ADE, URA*, and *LYS*. The auxotrophic selectable marker protein is encoded by the *ADE1* gene or the *ADE2* gene.

The integration vector comprises multiple insertion sites for the insertion of one or more expression cassettes encoding the one or more heterologous peptides, proteins and/or functional nucleic acids of interest. The integration vector comprises more than one expression cassette. The integration vector comprises little or no homologous DNA sequence between the expression cassettes. The integration vector comprises a first expression cassette encoding a light chain of a monoclonal antibody and a second expression cassette encoding a heavy chain of a monoclonal antibody.

The host cell is a lower eukaryote. The host cell is from a species selected from the group consisting of *Pichia pastoris*, *Pichia finlandica*, *Pichia trehalophila*, *Pichia koclamae*, *Pichia membranaefaciens*, *Pichia minuta* (*Ogataea minuta*, *Pichia lindneri*), *Pichia opuntiae*, *Pichia thermotolerans*, *Pichia salictaria*, *Pichia guercuum*, *Pichia pijperi*, *Pichia stiptis*, *Pichia methanolica*, *Pichia* sp., *Saccharomyces cerevisiae*, *Saccharomyces* sp., *Hansenula polymorpha*, *Kluyveromyces* sp., *Kluyveromyces lactis*, *Candida albicans*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Chrysosporium lucknowense*, *Fusarium* sp., *Fusarium gramineum*, *Fusarium venenatum*, *Physcomitrella patens*, and *Neurospora crassa*. The expression system as defined herein, wherein the host cell is *Pichia pastoris* or a *Pichia pastoris* cell that has been modified to be capable of producing glycoproteins having hybrid or complex *N*-glycans.

WHAT IS CLAIMED:

1. An expression system comprising:

(a) a *Pichia pastoris* host cell in which the endogenous gene or the open reading frame thereof encoding an auxotrophic selectable marker protein encoded by the *ADE1* gene or the ADE2 gene has been removed from the genome of the host cell to render the cell auxotrophic and

(b) an integration vector comprising:

(1) a nucleic acid comprising an ADE1 or ADE2 gene operably linked to an endogenous truncated ADE1 or ADE2 promoter, in which the truncation renders the promoter with a transcription activity that is less than the native endogenous promoter or no promoter;

(2) a nucleic acid having an insertion site for the insertion of one or more expression cassettes comprising a nucleic acid encoding a peptide, protein, and/or functional nucleic acid of interest, and

(3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the host cell by roll-in single-crossover homologous recombination.

2. The expression system of claim 1, wherein the ADE1 gene or ADE2 gene is the ADE1 gene.

3. The expression system of claim 1, wherein the ADE1 gene or ADE2 gene is the ADE2 gene.

4. The expression system of claim 1, wherein the integration vector comprises a first expression cassette encoding a light chain of a monoclonal antibody and a second expression cassette encoding a heavy chain of a monoclonal antibody.

5. The expression system of claim 1, wherein the *Pichia pastoris* cell has been modified to be capable of producing glycoproteins having hybrid or complex N-glycans.

6. A method for expressing a recombinant peptide, protein and/or nucleic acid of interest in a *Pichia pastoris* host cell using the expression system according to any of claims 1-5 wherein the transformed host cell produces the recombinant peptide, protein and/or nucleic acid of interest.

7. A *Pichia pastoris* cell comprising a deletion or disruption of the endogenous *ADE2* gene, or any other modification, rendering the *ADE2* gene or Ade2p non-functional and more than one copy of a vector comprising an *ADE2* open reading frame operably linked to an endogenous truncated ADE1 or ADE2 promoter in which the truncation renders the promoter with a transcription activity that is less than the native endogenous promoter or no promoter; wherein the auxotrophy of said host cell is not fully complemented by the ADE2 gene in the integration vector when a single copy of the vector is in the host cell.

8. The *Pichia pastoris* of Claim 7 wherein a vector comprising an *ADE2* open reading frame is integrated into a location in the *Pichia pastoris* genome other than the location where the endogenous *ADE2* gene had been located.

9. A method for producing a *Pichia pastoris* host cell that has multiple copies of an integration vector integrated into the genome of the host cell, comprising:

(a) providing *Pichia pastoris* cells in which an endogenous *ADE2* genehas been removed from the genome of the host cell to render the cells auxotrophic for adenine;

(b) transforming the cells by roll-in single crossover homologous recombination with an integration vector comprising:

(1) a nucleic acid comprising an ADE2 gene operably linked to an endogenous truncated ADE2 promoter in which the truncation renders the promoter with a transcription activity that is less than the native endogenous promoter or no promoter;

(2) a nucleic acid having one or more expression cassettes comprising a nucleic acid encoding one or more heterologous peptides, proteins, and/or functional nucleic acids of interest, and

(3) a targeting nucleic acid that directs insertion of the integration vector into a particular location of the genome of the cells by homologous recombination; wherein the auxotrophy of said host cell is not fully complemented by the ADE2 gene in the integration vector when a single copy of the vector is in the host cell;

(c) cultivating the transformed cells in a medium lacking adenine.

10. The method of claim 9, wherein the integration vector comprises multiple insertion sites for the insertion of one or more expression cassettes encoding the one or more heterologous peptides or proteins.

11. The method of claim 9, wherein the integration vector comprises more than one expression cassette.