

(19) **DANMARK**

(10) **DK/EP 3510043 T3**



(12)

Oversættelse af europæisk patentskrift

Patent- og
Varemærkestyrelsen

-
- (51) Int.Cl.: **C 07 K 14/395 (2006.01)** **C 12 N 1/18 (2006.01)**
- (45) Oversættelsen bekendtgjort den: **2022-01-03**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2021-12-15**
- (86) Europæisk ansøgning nr.: **17771808.7**
- (86) Europæisk indleveringsdag: **2017-08-31**
- (87) Den europæiske ansøgnings publiceringsdag: **2019-07-17**
- (86) International ansøgning nr.: **FR2017052316**
- (87) Internationalt publikationsnr.: **WO2018046820**
- (30) Prioritet: **2016-09-08 FR 1658370**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **Lesaffre et Compagnie, 41, rue Etienne Marcel, 75001 Paris, Frankrig**
- (72) Opfinder: **PIGNEDE, Georges, 3 chemin des Dames, 59700 Marcq-en-Baroeul, Frankrig**
DESFOUGERES, Thomas, 95 Rue du Clos des Meuniers, 86130 DISSAY, Frankrig
- (74) Fuldmægtig i Danmark: **Budde Schou A/S, Dronningens Tværgade 30, 1302 København K, Danmark**
- (54) Benævnelse: **Anvendelse af MCM7 for opnåelse af eddikesyreresistente gærstammer**
- (56) Fremdragne publikationer:
WO-A1-2013/178915
M. J. FITCH ET AL: "Mcm7, a Subunit of the Presumptive MCM Helicase, Modulates Its Own Expression in Conjunction with Mcm1", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 278, no. 28, 3 juillet 2003 (2003-07-03), pages 25408-25416, XP055365698, US ISSN: 0021-9258, DOI: 10.1074/jbc.M300699200
NUNO P. MIRA ET AL: "Genomic Expression Program Involving the Haa1p-Regulon in Saccharomyces cerevisiae Response to Acetic Acid", OMICS: A JOURNAL OF INTEGRATIVE BIOLOGY, vol. 14, no. 5, 1 octobre 2010 (2010-10-01), pages 587-601, XP055202932, ISSN: 1536-2310, DOI: 10.1089/omi.2010.0048 cité dans la demande
LEE YEJI ET AL: "Transcriptome analysis of acetic-acid-treated yeast cells identifies a large set of genes whose overexpression or deletion enhances acetic acid tolerance", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, DE, vol. 99, no. 15, 11 juin 2015 (2015-06-11) , pages 6391-6403, XP035509970, ISSN: 0175-7598, DOI: 10.1007/S00253-015-6706-Y [extrait le 2015-06-11]

FIELD OF THE INVENTION

The present application relates to the use of regulation for the expression of the *MCM7* gene to confer resistance to an organic acid, advantageously acetic acid, to a *Saccharomyces cerevisiae* yeast strain during glucose fermentation.

5 Accordingly, the presence, at least at an allele, of a binding site for the transcription regulator Haa1p upstream of the *MCM7* gene would induce its expression, which translates by increased resistance to the organic acid.

Thus, the present invention offers a method allowing the genotypic
10 selection of a *Saccharomyces cerevisiae* yeast strain resistant to an organic acid, advantageously acetic acid, during the glucose fermentation. In the same way, the invention allows for substantial optimization of the yielded *Saccharomyces cerevisiae* yeast strains resistant to an organic acid, advantageously acetic acid, based on the genotypic selection of the spores.

15

BACKGROUND OF THE INVENTION

The capacity of genetically modified yeasts to ferment diverse substrates makes them a tool of choice in diverse industrial processes, particularly in the production of ethanol from lignocellulose biomass. Alcoholic fermentation is the
20 process yeasts use in anaerobic media during which sugars are transformed into alcohol. The yeast *Saccharomyces cerevisiae*, also known as “baker’s yeast,” remains the microorganism used most often for alcoholic fermentation.

However, some fermentation inhibitors are found naturally in these substrates
25 and negatively impact ethanol production. This is in particular the case for weak organic acids, particularly for acetic acid, a degradation product of hemicellulose. When yeasts are confronted with the presence of organic acid in their environment, they block their cellular cycle to be able to prepare to react to this new abiotic stress. The fermentation only starts once the cellular resistance machinery is in place. The
30 presence of a weak organic acid accordingly has the consequence of delaying the initiation of the fermentation on glucose, thereby increasing production costs.

The problem related to acetic acid is all the more crucial since it is a very powerful inhibitor for alcoholic fermentation by yeasts and it is found in high concentrations in some fermentation media.

5 Various means have been described to attempt to counter the effect of fermentation inhibitors, such as for instance detoxifying the fermentation medium, or adapting yeasts to fermentation inhibitors by acclimation or genetic modification. In the case of acetic acid, detoxifying the fermentation medium is a difficult option to implement, particularly industrially. It is therefore necessary to
10 attempt to modify the yeasts themselves.

In this context, the acclimation of yeasts can be achieved by adding the inhibitor into the culture medium, preferably at increasing doses. However, it has been observed that yeast adaptation according to this method is only
15 transient, and disappears quickly when they are again cultivated in a medium devoid of inhibitor. So the method proves to be of little industrial interest, as there phenotypically stable strains are necessary.

In the case of sensitivity to acetic acid, only the genetic modification of
20 yeasts can therefore be envisaged. This can be done either by modification by genetic engineering, targeting specific genes, or classically by crossing strains of interest. Currently, the molecular mechanisms related to the sensitivity or on the contrary the resistance to acetic acid are poorly understood, and insufficient for the methods targeted by genetic engineering.

25

Accordingly, the method of choice to improve resistance to acetic acid remains yielding yeasts by crossing. However, even though some methods have yielded acetic-acid resistant yeast strains, these are by definition random and cannot guarantee success.

30

Document WO 2013/178915 describes crossing processes for yeast strains that allow the production of yeasts that can metabolize glucose and are acetic-acid resistant. This method consists in crossing the yeast strain filed at the CNCM under the number I-4538 with the yeast strain filed at the CNCM under the number I-4627, then in selecting a hybrid that can metabolize xylose and, independently, resist acetic acid during the fermentation of glucose.

This hybridization method relies on the capacity of yeasts to reproduce either asexually, or sexually, according to the culture conditions in particular.

10

Yeast *S. cerevisiae* is an organism with a haplodiplontic reproductive cycle, i.e. an organism capable of actively multiplying both in the haploid and the polyploid, for example diploid, state.

As long as the medium is favorable, polyploid yeasts are capable of vegetative multiplication by sprouting giving rise to polyploid yeasts. In the case of a medium poor in nitrogen-containing nutrients and containing only a non-fermentable carbon source (for example glycerol, acetate, etc.), the heterozygote cells for the Mat locus enter into meiosis and form yeasts having lower ploidy (spores or segregants) by a mechanism called sporulation.

20

Segregants can multiply by sprouting, giving yeasts having the same genome. Among haploid yeasts two opposing sexual signs are distinguished, called *MATa* and *MATα*. Two haploid spores with opposite sexual type can fertilize to yield a diploid yeast.

25

The haplodiplontic cycle for *S. cerevisiae* has been widely used to cross sexually compatible segregants (*MATa* and *MATα*), particularly in the method called random recombination from mass sporulation and hybridization. In a classic manner, two parental diploid strains (different from the genomic point of view) are used. The sporulation of parental diploid strains is typically induced by cultivating them in

30

conditions where the nitrogen supply is limited and only in the presence of a non-fermentable carbon source. The meiosis operating during this step leads to a genetic cross-fertilization, creating spores with varied genotype. The spores (haploid) obtained for each of the parental strains are then put in contact, to produce diploid
5 (hybrid) strains by fusion. This last step is called the hybridization step.

This method is interesting in that it allows the creation of genetic cross-fertilization from which interesting phenotype traits can emerge. It does however require a step of selecting hybrids on the basis of desired phenotype traits. As an
10 example, in the case of strictly diploid parental strains each presenting a phenotype trait borne by 10 genes, the probability of obtaining the hybrid of interest is estimated at $1/2.097.10^6$. The final selection step is tedious, long and expensive, especially.

Therefore an obvious need exists for improved production methods for
15 acetic-acid resistant yeast strains.

DESCRIPTION OF THE INVENTION

The inventors have identified that the yeast strains of *S. cerevisiae* capable of having expression of *MCM7* induced by an organic acid, particularly acetic
20 acid, have a phenotype of resistance to this organic acid. As shown in the experimental section, the growth of these strains and their capacity to ferment the glucose in a medium rich in acetic acid are improved.

In a particularly interesting manner, the inventors have determined that the
25 expression of *MCM7* in these strains can be induced by acetic acid. In addition it appears that this expression of *MCM7*, induced by acetic acid, is mediated by the transcription factor Haa1p. This is particularly surprising, given that *MCM7* is not known as being a gene regulated by Haa1p. Accordingly, the work of Mira and collaborators published in 2010 (Mira *et al*, 2010, OMICS, 14: 587-601) proposes a
30 list of genes regulated by Haa1p in the strain BY4741, which does not include *MCM7*. The inventors did establish though that in the strains of *S. cerevisiae* that

resist acetic acid, the region upstream of the *MCM7* coding gene comprises a motif known to be a binding site for Haa1p. Without being bound to any theory, the presence of this binding site for Haa1p would be the fruit of single-nucleotide polymorphism (SNPs).

5

On the basis of these elements, the inventors have developed selection and production methods for acetic-acid resistant strains of *S. cerevisiae*.

DEFINITIONS

10 The term “yeast strain” denotes in the sense of the invention a yeast population strictly identical from a genetic point of view. This encompasses both strains referred to as laboratory strains and those referred to as industrial strains. This term is to be differentiated from the term “yeast,” a yeast being obtained by the culture of a strain as defined above. In the context of the invention, “yeast” is understood as a
15 commercial product obtained by implementation of a production method for a yeast strain. Thus, yeasts having different properties can be obtained from a single strain, where these differences are connected with the production method implemented.

 In the meaning of the invention, a “segregant” is the product of the meiosis
20 of a yeast strain, regardless of the ploidy level of said yeast. In the remainder of the application, the terms “segregant” and “spore” can be used interchangeably.

 The term “yeast strain able to metabolize glucose” denotes in the sense of the invention a yeast strain capable of converting glucose into ethanol, i.e.
25 capable of fermenting glucose. A yeast strain able to metabolize glucose within the meaning of the invention is a yeast strain that converts at least 70%, preferably at least 80%, and more preferably at least 90% of the glucose into ethanol in 60 hours in a fermentation medium comprising 150 g of glucose per kg of fermentation medium, in usual conditions for alcoholic fermentation.

30

Preferably, the method used to measure the percentage of glucose converted into ethanol is as follows:

The yeast strain used is inoculated in synthetic fermentation medium at 0.25 g
5 of yeast in dry matter / kg of fermentation medium. The 60 hour duration is
calculated from the inoculation of the fermentation medium with the yeast strain. A
synthetic fermentation medium is a medium whose exact chemical composition is
known. In the scope of the invention, a synthetic fermentation medium comprises a
carbon source, a nitrogen source, a phosphorus source, and the essential vitamins
10 and minerals for the growth of a yeast strain. Preferably, the fermentation medium
used to measure the percentage of glucose converted into ethanol is YF as defined
in the example embodiments (denoted YF Ac because of the presence of acetic
acid).

15 The fermentation is typically conducted at a temperature comprised
between 28 and 37 °C, or between 30 and 35 °C, advantageously equal to 32 °C,
with moderate stirring, for example at 90 or 100 rpm. The stirring is moderate
so as to not be oxygenating. The pH of the medium is preferably controlled, for
example by the buffering power of an acid/base pair (such as the acetic
20 acid/acetate pair), and acid, advantageously comprised between 3.5 and 6, or 4
and 5.5, even more advantageously equal to 4.4 or 5.

The amount of ethanol present in the fermentation medium is measured by
any appropriate means known to the person skilled in the art. It can be a direct
25 measurement of the ethanol produced or an indirect measurement through a
parameter correlated to ethanol production, such as CO₂ production determined
by measuring the mass lost. For example, the production of alcohol may be
measured by chromatography, including HPLC (High Performance Liquid
Chromatography), an enzymatic kit (for example the determination of ethanol by
30 Boehringer kit), or a determination by potassium dichromate. The amount of

glucose in the fermentation medium is measured by any appropriate means known to the person skilled in the art, preferably by chromatography, in particular HPLC.

In the context of the invention, “organic acid” or “weak organic acid” is understood to mean a carboxylic acid that can inhibit the fermentation of a sugar, advantageously glucose. It advantageously involves acetic acid, levulinic acid, or formic acid, still more advantageously acetic acid.

It should be noted that it is known that only the non-dissociated or non-ionized form of such acids have inhibition ability. In the context of the invention, “un-ionized or non-dissociated form” of a carboxylic acid is understood as the protonated form thereof. In practice, the form of such organic acids depends on the pH of the medium in which they are incorporated. At a pH greater than the pKa of the acid, the acid will be mostly found in dissociated form or COO^- ions. In contrast and at a lower pH, the majority form is the non-dissociated or unionized form (COOH). In the remainder of the invention, the quantities or concentrations stated refer to acetic acid added to the medium, containing dissociated and undissociated forms according to the pH of said medium.

The terms “resistant to an organic acid” or “acetic-acid resistant” denote a yeast strain that can ferment at least one sugar, particularly glucose, with the organic/acetic acid having limited impact on the alcoholic fermentation curve. The alcoholic fermentation curve representing the quantity of alcohol produced as a function of time generally includes three phases: a latency phase during which there is not ethanol production, an alcohol production phase, and a plateau phase, which corresponds to the end of the fermentation.

In a known manner, acetic acid inhibits glucose fermentation, this inhibition translating as a delay during initiation of fermentation with the kinetics subsequently remaining unchanged. It should be noted that in the

presence of both glucose and xylose in the medium, yeast strains ferment glucose first because of catabolite repression.

Accordingly, an “acetic-acid resistant strain” advantageously delays the initiation of alcoholic fermentation by less than 30 hours, preferably less than 20 hours, more preferably less than 15 hours, perhaps 10 hours. Preferably, reference is made to capacity to ferment glucose with a delay in initiation of alcoholic fermentation as indicated above.

The fermentation medium used to assess resistance to acetic acid is preferably a synthetic medium, more preferably the medium YFAc as illustrated in the embodiment examples. The composition of medium YFAc is as follows: 150 g/kg of glucose, 5 g/kg of yeast extract, 4.7 g/kg of DAP (diammonium phosphate), 11.4 g/kg of citric acid, 4 g/kg of acetic acid, 13.5 g/kg of sodium citrate, 1 mL/kg of Tween 80, 2 mL/kg of ZnSO₄ (at 10.6 g/L), 2.5 mL/kg of MgSO₄ 7H₂O (at 400 g/L), 1 mL/kg of thiamine (at 18.24 g/L), 1 mL/kg of pyridoxine (at 5.28 g/L), 1 mL/kg of biotin (at 1.76 g/L), 1 mL/kg of panthotenate (at 3.8 g/L), 2.5 mL/kg of nicotinic acid (at 8 g/L), 1 mL/kg of mesoinositol (at 50 g/L), 1 mL/kg of riboflavin (at 1 g/L), 1 mL/kg of para-aminobenzoate (at 1.2 g/L), pH adjusted to 4.4 with KOH. The inoculation of the yeast strain used to evaluate resistance to acetic acid is preferably 0.25 g dry matter /kg of fermentation medium. The time $t = 0$ of the alcoholic fermentation corresponds to the time when the fermentation medium is inoculated with the yeast strain. Alcoholic fermentation should be conducted preferably at 32 °C under medium stirring, for example 90 rpm.

Note that at the concentration of 2000 ppm, acetic acid does not inhibit the fermentation. In a suitable manner, acetic acid is added at from 1 to 10 g/kg of fermentation medium, for example 4 g/kg of fermentation medium

In the sense of the invention, “*MCM7*” denotes the gene coding the protein Mcm7p also called protein Cdc47p, and the product of the expression of this gene.

“Gene coding Mcm7p” or “*MCM7* gene” are understood in the sense of the invention the gene of the yeast *Saccharomyces cerevisiae* located on chromosome II between the 625767 and 628304 positions, corresponding to ORF (Open Reading Frame) coding Mcm7p. These positions are indicated in reference to the genome of the yeast strain *S. cerevisiae* S288c, particularly its complete sequence, available in the databases under reference GenBank GCA_000146045.2. (version of April 18, 2011), and whose NCBI reference is Gene ID: 852501. The sequence of chromosome II, used as a reference sequence for numbering, is that accessible under number NCBI NC_001134.7 (23/12/2010; SEQ ID NO: 1).

The term “upstream of gene” must be understood in its generally accepted meaning in molecular biology, i.e. as meaning the region located at 5’ (of the coding strand) of the initiation site for the gene transcription. In a classic manner, this region (also called “promoter/regulatory region”) is involved in the expression of *MCM7* gene, comprising in particular promoter and regulatory sequences, such as binding sites for transcription regulators. In the sense of the invention, this region is constituted of the 1200 5’ nucleotides at the initiation site of the transcription of the gene encoding Mcm7p.

“Induction of the expression of *MCM7* by or in the presence of an organic acid” or “overexpression of *MCM7* by or in the presence of an organic acid,” is understood in the presence of the invention to mean an increase in the level of expression of the gene encoding Mcm7p in the presence of organic acid, in comparison with the level of expression of this gene in the same strain in the absence of the organic acid. This increase in expression can translate to the nucleotide level (increased mRNA) or to the protein level. The person skilled in the art will therefore be able to choose the measurement method that seems the most appropriate to him and the simplest to implement among the well known methods of molecular biology and biochemistry (Northern blot, PCR, Western blot, etc.).

In the sense of the invention, the terms “induction of the expression of *MCM7* mediated by the transcription factor Haa1p” or “overexpression of *MCM7* mediated by the transcription factor Haa1p” mean that the induction of the expression or the overexpression of *MCM7* targeted in the present application
5 depend on transcription factor Haa1p. In other words, in the absence of transcription factor Haa1p, for example for the yeasts in which this gene is deleted or carries mutation(s) that make it non-functional, or even in conditions where the level of Haa1p is limited, yeasts are no longer capable of inducing the expression of *MCM7*, in particular in the presence of organic acid.

10

In the context of the invention, “Haa1p binding site” is understood to mean the nucleotide sequence recognized by the Haa1p transcription factor, allowing it to bind at the target gene whose transcription level is then regulated by Haa1p. According to Mira *et al.* (*Nucleic Acid Research*, 2011, 39 (16):6896-6907), the
15 minimum motif recognized by Haa1 has the following sequence:

5'- (G/C)(A/C)GG(G/C)G - 3'

instead of motif -5'-GNN(G/C)(A/C)(A/G)G(A/G/C)G-3', determined previously *in silico*.

20 DETAILED DESCRIPTION OF THE INVENTION

The present application relates to the use of the *MCM7* gene to confer resistance to an organic acid on or to a yeast strain.

Advantageously, the resistance conferred is resistance to acetic acid, whose
25 presence in high concentrations in the lignocellulose hydrolysates is intrinsically related to that of acetyl groups associated in a covalent manner with the hemicellulose molecules.

Advantageously, the resistance to the organic acid in the yeast strain
30 translates during the glucose fermentation, for which the delay in initiation of fermentation is lessened or reduced.

Advantageously, the use of the *MCM7* gene is of interest when it is induced, advantageously in the presence of the organic acid in question, in particular of acetic acid.

5 Even more advantageously, the expression of the *MCM7* gene induced by the presence of the organic acid is mediated by the transcription factor Haa1p. According to the invention, the sequence upstream of the *MCM7* gene comprises a binding site for Haa1p.

10 According to a specific embodiment, the sequence upstream of the *MCM7* gene comprises the following sequence:

GAGGGG or

GAGGAGGGG or

SEQ ID NO: 2 or

15 SEQ ID NO: 3 or

SEQ ID NO: 4.

According to another embodiment of the invention, the sequence upstream of the *MCM7* gene has at least the following characteristic:

20 - a G at the 624794 position 624794.

As already stated and in the scope of the invention, the position number corresponds to that of the reference strain S288c. Therefore “position” must be understood as the position, in the strain studied, that corresponds to the given position.

25

According to an embodiment of the invention, the sequence upstream of the *MCM7* gene has at least one G at position 624794. According to another embodiment, it has a C at position 624758, a G at position 624794 and an A at position 624801, or all of the following characteristics:

30 - a T at position 624536;

- a T at position 624732;

- a G at position 624736;
- a C at position 624758;
- a G at position 624794;
- an A at position 624801;
- 5 - an A at position 624832;
- a C at position 625073;
- a G at position 625146;
- an A at position 625199.

10 The yeast strain targeted by the present invention belongs to the Hemiascomycetes group, to the *Saccharomyces* genera. It advantageously concerns *Saccharomyces cerevisiae*.

Thus and according to a first aspect, the invention concerns a selection
15 process for a *Saccharomyces cerevisiae* yeast strain resistant to an organic acid comprising:

- the demonstrated presence, at least at one allele of the strain, of a Haa1p binding site in the sequence upstream of the *MCM7* gene, advantageously the sequence GAGGGG or GAGGAGGGG or SEQ ID NO: 2 or SEQ ID
20 NO: 3 or SEQ ID NO: 4; and/or
- the demonstrated presence, at least at one allele of the strain, of the presence of a base G at position 624794 on chromosome II.

As already stated, the organic acid is advantageously acetic acid.

25

According to another preferred embodiment, the resistance to the organic acid of the resulting selected strain is observed during its glucose fermentation.

Accordingly, the present invention offers a genotypic screening method for
30 strains of interest, here resistant to an organic acid. This approach is much less

expensive, in terms of both time and money, than phenotypic screening traditionally used.

As set out, the following criteria can be evaluated:

5

The first criterion relies on the presence, at 5' of *MCM7*, of a Haa1p binding site. Accordingly, and in a suitable manner, the region located upstream of *MCM7* comprises the sequence (G/C)(A/C)GG(G/C)G, perhaps GNN(G/C)(A/C)(A/G)G(A/G/C)G (where N is a nucleotide chosen from A, C,
10 G and T).

According to a specific embodiment, this region contains at least one sequence chosen from:

- GAGGGG corresponding to a minimum Haa1p binding site;
- 15 - GAGGAGGGG corresponding to the motif recognized by Haa1p, determined *in silico*;
- The sequence SEQ ID NO: 2;
- The sequence SEQ ID NO: 3;
- The sequence SEQ ID NO: 4;
- 20 advantageously the sequence SEQ ID NO: 3 or 4.

These sequences can be demonstrated by any technique known to the person skilled in the art such as sequencing, PCR, hybridization.

25 As is shown in the example embodiments, this may relate only to one allele of this strain, perhaps to several, perhaps even to all the alleles (two in the case of a diploid yeast).

According to a second criterion, the strain presents, at least at one allele, a
30 base G at the position corresponding to the 624794 position of chromosome II, indicated in bold in the corresponding sequence in the minimum Haa1p binding

site (GAGGGG). Note that an A is observed in this position in non-acetic-acid resistant strains. Without wanting to be bound by any category, the replacement of a base A by a base G allows the creation of a functional Haa1p binding site, allowing the induction of the expression of *MCM7* in the presence of acetic acid.

5

Note that other mutations can be found in this region, advantageously chosen from:

- a T at the 624536 position;
- a T at the 624732 position;
- 10 - a G at the 624736 position;
- a C at the 624758 position;
- an A at the 624801 position;
- an A at the 624832 position;
- a C at the 625073 position;
- 15 - a G at the 625146 position;
- an A at the 625199 position.

According to a specific embodiment, a strain targeted by the invention is selected because of the presence of at least one base G at the position
20 corresponding to the 624794 position of chromosome II, perhaps of at least one C at the 624758 position, a G at the 624794 position and an A at the 624801 position, perhaps even because of the 10 nucleotides mentioned above in relation with the specific positions stated, at least in one of its alleles.

25 The demonstration or identification of these mutations is easily achieved by the person of skill in the art, for example by sequencing the positions of interest.

The interest of a strain identified using this method can of course be confirmed by a phenotypic approach, consisting of evaluating the capacity of the strain selected
30 to ferment glucose in the presence of acetic acid, for example as described above.

These various genotypic screening criteria can also be implemented to produce or obtain yeast strains resistant to an organic acid, advantageously acetic acid, particularly in the context of glucose fermentation.

5 Accordingly and using the available molecular biology tools, it is possible to do mutagenesis (directed or random) on yeast strains to obtain the desired phenotype. As already stated, the presence of the genotypic characteristics stated above is only *a priori* necessary at one allele. Alternatively, this mutagenesis can therefore be achieved on spores (or segregants) that are then hybridized with other
10 spores, optionally from another strain presenting another phenotypic trait of interest.

Advantageously, the invention proposes a method for obtaining a yeast strain resistant to an organic acid, based on a genotypic screening conducted in
15 haploid spores or segregants.

Accordingly and according to another feature, the invention proposes a method for obtaining a *S. cerevisiae* yeast strain resistant to an organic acid comprising:

- 20 - a sporulation step for two parental strains having different genomes or divergent phenotypic traits;
- a mass hybridization step for the spores or segregants obtained, said process comprising at least one selection step of spores or segregants because of the presence of a Haa1p binding site in the
25 sequence upstream of *MCM7*, advantageously the demonstration of the sequence SEQ ID NO: 3 or SEQ ID NO: 4, and/or the presence of a base G at the 624794 position of chromosome II.

In a characteristic manner, the method of the invention comprises a step of
30 sporulation of the parental strains. This technique is well known to the person skilled in the art and does not therefore require further description. As an

example, the sporulation step can be conducted by cultivating the parental strains in appropriate culture conditions, such as for example in a deprived medium.

Among the parental strains, at least one of them has the phenotype of interest, here resistance to an organic acid, advantageously acetic acid, in the sense of the invention. Advantageously, this delays initiation of alcoholic fermentation less than 30 h, preferably less than 20 hours, more preferably less than 15 hours perhaps 10 hours in the fermentation medium YFAc with an inoculum of 0.25 g of dry matter of yeast /kg of medium. Such strains are well known to the person skilled in the art. If need be, the person skilled in the art will be able to produce the parental yeast strain resistant to the organic/acetic acid in the sense of the invention, for example by the usual techniques using selection pressure.

Note that a parental strain of interest can also be identified using a selection method, the subject-matter of the present invention.

In the specific case of the use of *S. cerevisiae*, it may in particular be strain EGAc1 filed at the CNCM (Collection Nationale de Cultures de Microorganismes, Institut Pasteur, 25 rue du Doctor Roux, 75724 Paris Cedex 15) under number I-4839 on March 13, 2014.

As shown in the context of the present application, where this phenotypic trait is optionally carried by only one allele, sporulation gives rise to spores not bearing this genotypic trait, here half of the spores in the case of a diploid parental strain bearing only one allele with resistance to organic/acetic acid.

In a specific case, the two parental strains used in the context of the method of the invention have a phenotype of resistance to an organic acid in the sense of the invention.

In a suitable manner, at least one of the parental yeast strains, particularly that resistant to organic/acetic acid, advantageously both parental strains, can ferment glucose in the sense of the invention.

5 According to a preferred embodiment, the second parental yeast strain, advantageously the one not having resistance to organic/acetic acid, has a second phenotypic trait of interest. This is for example the capacity to metabolize pentoses, particularly glucose present in large quantities in lignocellulose hydrolysates.

10 Accordingly, yeast strains that can ferment glucose and also metabolize pentoses are available:

As an example, document WO 2010/000464 reports obtaining yeast strains able to ferment pentoses because of a bacterial gene coding for a xylose
15 isomerase (XI) which converts xylose into xylulose which can be metabolized by the yeast. As an alternative, a eukaryotic pathway should be noted comprising xylose reductase (*XR* or *XYL1*) generating xylitol and a xylitol dehydrogenase (*XDH* or *XYL2*) that can also produce xylulose.

20 Thus, document WO 2012/072793 describes improved yeast strains combining exogenous genes coding a xylose isomerase and a xylitol dehydrogenase with which to eliminate xylitol which proves to be an inhibitor of xylose isomerase. Such strains, in particular the strain registered with the CNCM (Collection
Nationale de Cultures de Microorganismes) on October 5, 2011 under number
25 I-4538, have improved yields and therefore proven industrial utility for the production of ethanol.

In that respect and according to a specific embodiment, a second parental strain used in the context of the method of the invention is the strain registered with the
30 CNCM (Collection Nationale de Cultures de Microorganismes, Institut Pasteur, 25

rue du Doctor Roux, 75724 Paris Cedex 15) under number I-4538 on October 5, 2011.

The method of the invention further comprises a mass hybridization step
5 for the spores or segregants obtained.

This step is easily carried out according to conventional methods used in the field and described in detail in chapter 7 “Sporulation and Hybridization of Yeast” by R.R. Fowell, in the reference work “The Yeasts”, Volume 1, edited
10 by A.H. Rose and J. S. Harrison, 1969-Academic Press. In brief, the hybridization is achieved by adding the spores in question to a suitable culture in the hybridization step. Typically, the person skilled in the art will be able to use for this step complete culture medium of the YPG type (containing 10 g/L yeast extract, Bactopeptone 20 g/L, glucose 20 g/L and demineralized water qsp
15 1 L).

The selection step for spores or segregants of interest achieved as described previously on the basis of at least one of the two following criteria:

20 The first criterion relies on the presence, at 5' of *MCM7* of the spore, of a Haa1p binding site. Accordingly, and in a suitable manner, the region located upstream of *MCM7* comprises the sequence (G/C)(A/C)GG(G/C)G, perhaps GNN(G/C)(A/C)(A/G)G(A/G/C)G (where N is a nucleotide chosen from A, C, G and T).

25

According to a specific embodiment, this region contains at least one sequence chosen from:

- GAGGGG corresponding to the minimum Haa1p binding site;
- GAGGAGGGG corresponding to the motif recognized by Haa1p,
30 determined *in silico*;
- The sequence SEQ ID NO: 2 or

- The sequence SEQ ID NO: 3 or
 - The sequence SEQ ID NO: 4)
- advantageously the sequence SEQ ID NO: 3 or 4.

5 These sequences can be demonstrated by any technique known to the person skilled in the art such as sequencing, PCR, hybridization.

According to another criterion, the spore has a base G at the position corresponding to the 624794 position of chromosome II, indicated in bold in the
 10 corresponding sequence in the minimum Haa1p binding site (GAG**GGG**). Note that an A is observed in this position in non-acetic-acid resistant spores. Without wanting to be bound by any category, the replacement of a base A by a base G allows the creation of a functional Haa1p binding site, allowing the expression or overexpression of *MCM7* in response to acetic acid.

15

Note that other mutations can be found in this region, advantageously chosen from:

- a T at the 624536 position;
- a T at the 624732 position;
- 20 - a G at the 624736 position;
- a C at the 624758 position;
- an A at the 624801 position;
- an A at the 624832 position;
- a C at the 625073 position;
- 25 - a G at the 625146 position;
- an A at the 625199 position.

According to a specific embodiment, a segregant targeted by the invention is selected because of the presence of at least one base G at the 624794 position
 30 of chromosome II, perhaps of at least one C at the 624758 position, a G at the

624794 position and an A at the 624801 position, perhaps even because of the 10 nucleotides mentioned above in relation with the specific positions stated.

The demonstration or identification of these mutations is easily achieved by
5 the person of skill in the art, for example by sequencing the positions of interest.

A selection like this can for example be made on the basis of total or partial sequencing according to the techniques of molecular biology well known to the person of skill in the art, or by PCR techniques. The sequencing can accordingly
10 be achieved by sequencing by hybridization or by high throughput sequencing techniques such as pyrosequencing, synthetic sequencing or ligation sequencing. Alternatively, the selection can be made on the basis of PCR techniques by looking for the targeted polymorphisms. In this context mention may be made of, for example, multiplex PCR techniques, which will look for several
15 polymorphisms in a single test, nested PCR, which produces highly sensitive results, or colony PCR, which does not require DNA extraction.

To achieve the sequencing or PCR, a DNA multiplication step may be necessary, to have sufficient biological material available. It will then be possible
20 to proceed with amplification by cultivating each strain or spore in a culture medium suitable for its reproduction. Further, a DNA extraction step for yeasts may be necessary, and can be done according to molecular biology methods well known in the field of the invention.

25 The method of the invention can be implemented in diverse ways: the yeast selection step can be done from parental strains, and/or from spores that come from it, and/or from strains obtained after hybridization. As already stated, it is however particularly advantageous to proceed with the selection step from spores coming from the parental strains.

Thus, according to one preferred embodiment, a method according to the invention comprises the following steps:

- a) prepare segregants from a first parental strain and segregants from a second parental strain;
- 5 b) select from among the segregants of step a) those having a Haa1p binding site in the sequence upstream of *MCM7*, advantageously the sequence SEQ ID NO: 3 or SEQ ID NO: 4, and/or having a base G at the 624794 position of chromosome II;
- 10 c) hybridize the segregants from the first parental strain and selected in step b) with segregants from the second parental strain, optionally selected in step b);
- d) select among the hybrids from step c) those resistant to organic acid.

Step a) of the method corresponds to a preparation step for segregants from
15 two different parental yeast strains, advantageously *S. cerevisiae*, i.e. a sporulation step. The person skilled in the art will easily be able to obtain segregants from parental strains defined above, according to the methods well known in the field of the invention.

20 In an advantageous embodiment, this step comprises the culture of the first parental yeast strain firstly and the second parental yeast strain secondly, in a medium deprived of nitrogen or sugar.

 According to the invention, the parental strains used are yeast strains
25 belonging to *Saccharomyces cerevisiae*.

 According to a preferred embodiment, the first strain is chosen for its resistance capacity to organic acid, advantageously acetic acid, particularly during glucose fermentation. It may in particular be strain EGAc1 registered
30 with the CNCM under number I-4839 on March 13, 2014. The segregants of this strain will as a priority be the subject of the selection step b) of the method.

According to another preferred embodiment, the second parental yeast strain is chosen for another phenotypic trait of interest, for example its capacity to metabolize xylose. It can for example be the strain registered with the CNCM under number I-4538 on October 5, 2011.

5

In a specific case, the second parental yeast may also have resistance capacity to organic acid, advantageously acetic acid, particularly during glucose fermentation. In that case, the spores of this strain will also be the subject of the selection step b) of the method.

10

Step b) of the method of the invention is used as described and allows the removal of spores that do not have the desired phenotypic trait, here resistance to an organic acid in the sense of the invention. Accordingly and thanks to this step, the probability of obtaining from the hybridization a strain having this phenotypic trait is greatly increased.

15

Step c) of the method of the invention is a step of hybridization of segregants from a first parental yeast strain with segregants from a second parental strain, where at least one of the segregant populations has been previously selected in step b), advantageously that from the first parental yeast strain.

20

Preferably, a mass hybridization is conducted in step c). In other words, step c) corresponds preferably to a hybridization step for all of the spores from step b).

25

Optional step d) of the method consists in selecting among the hybrids from step c) those capable of alcoholic fermentation and having a phenotype of resistance to organic acid, advantageously acetic acid. As already stated, this step is easily achieved by traditional selection methods, using usual culture techniques. This step is advantageously used on a medium comprising glucose.

30

The present invention is going to be illustrated more ahead using the following example embodiments, supported by the attached figures. However, they have no limiting scope.

5 LEGENDS FOR THE FIGURES:

Figure 1 shows the flow chart used to determine the proportion of alleles from EGAc1 (I-4839) in yeast populations that fermented in the presence of acetic acid (popB) or in yeast populations that fermented without acetic acid (popC) and along all of chromosome II.

10 S288c: strain of *S. cerevisiae* whose genome serves as GenBank reference GCA_000146045.2., version of April 18, 2011; NCBI Gene ID: 852501; chromosome II: NCBI NC_001134.7)

EGAc1: strain registered with the CNCM (Collection Nationale de Cultures de Microorganismes, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex
15 15) on March 13, 2014, under number I-4839.

I-4538: strain registered with the CNCM (Collection Nationale de Cultures de Microorganismes, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15) under number I-4749.

EGAc2: strain registered with the CNCM (Collection Nationale de Cultures
20 de Microorganismes, Institut Pasteur, 25 rue du Docteur Roux, 75724 Paris Cedex 15) on March 13, 2014, under number I-4840.

Figure 2 shows the frequency averages for alleles in strain EGAc1 (I-4839) in both populations: PopC = not stressed, dotted; PopB = stressed, line

Figure 3 shows

25 (A) the frequency averages for alleles in strain EGAc1 (I-4839) in the yeast populations that fermented in the presence of acetic acid (popB: stressed, line) or in the yeast populations that fermented without acetic acid (popC: not stressed, dotted) and along all of chromosome II.

(B) the LOD score value (similarity index) as a function of the position
30 along chromosome II for the 3 series of experiments conducted.

Figure 4 corresponds to

(A) a graphic representation of assembly of genome reads for strain EGAc2 (I-4840) in the 624000 to 626000 bp area of chromosome II of the reference strain. The area where the LOD scores are the highest is shown (624500 to 625200 bp; SEQ ID NO: 3).

- 5 (B) a focus on the impact of SNP (A > G) that affects 50% of the alleles in this strain in the 624794 position. The arrows represent the presence of a site recognized by the Haa1p transcription factor in the sequence SEQ ID NO: 2.

Figure 5 relates the average change in mass loss during a fermentation on YFAc medium at 32 °C with an inoculum of 0.25 g/kg (dry matter eq.). The strains used

10 are:

strain I-4538, sensitive to acetic acid (S);

strain EGAc1 (I-4839), acetic-acid resistant (R);

strain EGAc2 (I-4840) heterozygous, denoted “EGAc2-AlleleR/AlleleS”;

strain EGAc2 (I-4840) made homozygous for the allele from EGAc1

- 15 (resistant allele) at locus II-624794, denoted “EGAc2-AlleleR/AlleleR”;

strain EGAc2 (I-4840) made homozygous for the allele from I-4538 (sensitive allele) at locus II-624794, denoted “EGAc2-AlleleS/AlleleS”.

The error bars correspond to standard deviations calculated on the basis of measurements made with the 7 homozygous strains for the resistant allele and the

20 5 homozygous strains for the sensitive allele.

EXAMPLE EMBODIMENTS

Example 1: Identification of genetic traits and mutations related to the phenotype of resistance to weak organic acids such as acetic acid

25

In these first experiments, the goal is to produce a diverse yeast population from a genetic point of view (a), so as not only to select the strains resistant to weak organic acids but also to be able to analyze the genetic traits involved in this phenotype (b).

a. Production of a yeast population comprising strains resistant to weak organic acids:

The yeast population was obtained by random recombination from mass sporulation and hybridization. This strategy is inspired by the work of Leo Parts
5 *et al.* (2011, *Genome Res*, 21(7):1131–8). In brief, a segregant (also called spore) for an acetic-acid resistant strain (strain EGAc1/I-4839) is crossed with another segregant, from strain I-4538, thereby creating a first hybrid (strain EGAc2/I-4840), as described in WO2013/178915.

10 Secondly, the genome of strain EGAc2 (I-4840) was recombined randomly to obtain a very diverse yeast population from a genetic point of view. In practice, hybrid EGAc2 (I-4840) was set to sporulate then the spores obtained were left to freely rehybridize among themselves, as described in WO2013/178915. The cycle was reproduced 4 times, thereby generating a 24
15 reduction in the genetic distance in centiMorgan (cM).

b. Selection of strains resistant to weak organic acids:

Strains resistant to weak organic acids were selected following the principles of population genetics, particularly the Hardy-Weinberg principle,
20 which states that in an isolated population with unlimited count, not subject to selection, and in which there are no mutations, allele and genotype frequencies remain stable from generation to generation.

Accordingly, in the absence of selection, in the case of 2 alleles “A” and “B,”
25 where only “A” can play a role in adapting the population at a given selection pressure (for example resistance to weak organic acids), the frequencies of allele “A” and allele “B” in the population remain stable. By contrast, if the environment changes, and the medium is enriched with weak organic acid, then the less adapted strains will disappear (B) for the benefit of the more adapted strains (A). According
30 to this principle, in the case where selection pressure exists, we then observe a deviation in this equilibrium over several generations. Accordingly, by comparing

the allele frequency variations between a population not subjected to the selection and a population subjected to a selection pressure, one can determine the alleles that may be involved in the resistance or adaptation to the selection applied.

- 5 In practice, to have a control population, a sample of the population obtained in point a) was cultivated in a medium devoid of acetic acid (no selection pressure). The resulting population is called “population C.”

10 In parallel, a sample of the population obtained in point a) was subjected to a high selection pressure by adding acetic acid at the start of the alcoholic fermentation in the presence of glucose. The resulting population is called “population B.” In practice, acetic acid was added to the culture medium so as to obtain a concentration of 4 g/L.

- 15 c. Determination of genetic traits involved in the phenotype for resistance to weak organic acids:

20 Strain EGAc1 (I-4839) being acetic-acid resistant, it is expected that the genetic traits associated with this resistance are present in this strain. Consequently, to limit the number of alleles to analyze, one firstly identifies the alleles specific to strain EGAc1 (I-4839) that are present in strain EGAc2 (I-4840), and in populations C and B. Secondly, it is possible to determine the frequency of each of these alleles appearing in the stressed (B) or unstressed (C) populations.

25 The study of allele frequency variations along the genome was conducted in the following manner:

After the fermentations, the genomic DNA coming from strain EGAc1 (I-4839), from strain EGAc2 (I-4840) and from populations B and C were extracted then sequenced by the “Paired End” method using an Illumina HiSeq 2000.

The results were then processed according to the approach illustrated in Figure 1. Note that the reference genome is that for strain S288c (GenBank GCA_000146045.2., version of April 18, 2011; NCBI Gene ID: 852501; NCBI NC_001134.7).

5

In the case of a study conducted on complex populations, the allele frequency reflects the number of individuals carrying the allele in question. So as an example and in direct application of the Hardy-Weinberg principle, if an allele is present with a frequency of 70%, it is possible to deduce that 91% of individuals in the population carry at least one copy of this allele.

10

Figure 2 shows how the allele frequency from strain EGAc1 (I-4839) changes along chromosome II and present in populations B and C.

15

The results in Figure 2 show that in the first portion of chromosome II, the allele frequencies from EGAc1 (I-4839) are very close in both types of populations and up to about 530 kb. This observation suggests that the genes present in this part of chromosome II (from 0 to 530 kb) would not be involved in the acetic-acid resistance process transmitted by the segregant from strain EGAc1 (I-4839) to strain EGAc2 (I-4840).

20

The second part of chromosome II (from 530 kb to 660 kb) shows a dissociation in curves corresponding to allele frequencies from EGAc1 (I-4839) in both types of populations. This result shows that there is an allele that is more represented in the populations subjected to selection pressure than in those that fermented without this pressure. This overrepresentation suggests that this allele from EGAc1 (I-4839) would favor the multiplication of cells in the presence of acetic acid. In other words, the second part of chromosome II (from 530 kb to 660 kb) is identified as being a quantifiable genetic trait (also called QTL), i.e. a genetic region involved in acetic-acid resistance.

25
30

QTL are generally large sequences. One sub-region seems to stand out: this is the region from bases 624000 to 626000 of chromosome II. This region seems to include the highest allele frequency variations. Next, it is this sub-region that was analyzed.

5

d. Analysis of genetic traits involved in the phenotype for resistance to weak organic acids:

Within the QTL, these are more specifically frequency variations for punctual polymorphisms (SNP for “Single-Nucleotide Polymorphism”), in comparison with the strains that do not present the phenotypic trait studied, which can be the most relevant.

10

The base pair region 624000 to 626000 of chromosome II was therefore analyzed in more detail. The similarity index values for the SNPs that it comprises were analyzed to target the region comprising the most relevant SNPs in the phenotype for resistance to weak organic acids (i). Once this region was identified, the SNPs that it comprises were analyzed more deeply (ii).

15

i. Determination of the most region the plus pertinente:

20

The approach chosen was that of LOD scores, as described by Lander and Botstein (1989, *Genetics*, 121:185-99).

The first step consists in determining the inference:

$$fB = \frac{n_{1,B}}{n_{1,B} + n_{2,B}}$$

$$fC = \frac{n_{1,C}}{n_{1,C} + n_{2,C}}$$

25

In the first equation, “n1” is the number of reads bearing the SNP from strain EGAc1 (I-4839) in population B and “n2” the number of reads bearing the

other SNP. The next equation relates to the same calculation but applied to the results of unstressed populations.

The second step consists in calculating the similarity:

- 5 The similarity (L) is defined as being a conditional probability function. Accordingly, this is the probability of having the allele of strain EGAc1 (I-4839), either in the population that fermented in the presence of acetic acid, or in the population that did not ferment in the presence of acetic acid, calculated according to the equations below:

$$L(B|fB) = \prod_{i=1}^3 (B_i|fB)$$

$$L(C|fC) = \prod_{i=1}^3 (C_i|fC)$$

10

The third step consists in calculating the LOD score:

- From these determinations of similarity for each allele, it is now possible to calculate the LOD for each SNP and to combine it with the allele from strain
15 EGAc1. The equation used is that published by Lander and Botstein in 1989:

$$LOD = \log \left(\frac{L(B|fB) \cdot L(C|fC)}{L(B \cup C|f)} \right)$$

- The LOD score was analyzed 3 times on the area of interest for chromosome II (3 independent experiments). These results are shown in Figure 3. These results
20 show that when the difference in allele frequency is high, the dispersion of LOD scores is also high (principally because of their mathematical and biological links).

- Figure 3B shows that an area comprised between 624500 bp and 625200 bp has high LOD scores that appear reproducible. The most relevant region
25 identified is therefore the area between 624500 bp and 625200 bp of chromosome II (SEQ ID NO: 3).

ii. *Analysis of SNPs in the most relevant region (between 624500 bp and 625200 bp of chromosome II):*

The genetic structures present and the SNPs found most frequently in the area considered (between 624500 bp and 625200 bp of chromosome II) in strain
5 EGAc2 (I-4840) were analyzed.

These results are shown in Figure 4.

10 SNPs were identified in the area between 624500 bp and 625200 bp
10 (SEQ ID NO: 3) of chromosome II:

Position (with reference to strain S288c: NC_001134.7; SEQ ID NO: 1)	Base in the reference strain	Mutation
624536	A	T
624732	C	T
624736	A	G
624758	T	C
624794	A	G
624801	G	A
624832	C	A
625073	T	C
625146	C	G
625199	C	A

The impact of these SNPs on the sequence of coding areas (open reading frames) did not show mutations causing changes to the protein sequence.

15

By contrast, polymorphism (A ->G) in the 624794 position, which affects 50% of the alleles in strain EGAc2 (I-4840), reconstitutes a site recognized by the Haa1p transcription factor located upstream of *MCM7*, and codes a DNA helicase.

DNA helicases replicate DNA and therefore regulate the cell cycle progression and, consequently, the production of biomass. In other words, yeasts having this SNP have a binding site for Haa1p transcription factor that the other yeasts do not have, upstream of a gene known to play an important role in cellular division.

- 5 Accordingly, the fact that the presence of a site recognized Haa1p upstream of *MCM7* makes the strains more resistant could hold to the fact that in the presence of acetic acid, yeasts carrying this site would multiply more quickly.

Example 2: Validation of the interest of the allele denoted “EGAc1-II-

10 **624794”**

To take another approach to analyzing the identified area of interest (particularly bearing mutations T-> C in the 624758 position, A -> G in the 624794 position and G -> A in the 624801 position; SEQ ID NO: 4), it was chosen to achieve the loss of heterozygosity in strain EGAc2 (I-4840). Secondly, the performances of

15 strains made homozygous for one or other of the two alleles were compared.

a) Construction of homozygous EGAc2 strains for the wild allele or for the allele comprising the SNP identified as of interest:

- Strain EGAc2 (I-4840) is heterozygous for the QTL identified as of interest, i.e.
- 20 bearing the SNP (A ->G) in the 624794 position. Indeed, it has an allele denoted “II-624794” of strain EGAc1/I-4839 (i.e. having a G in the 624794 position) and an allele “I4538-II-624794” from strain I-4538 (i.e. having an A in the 624794 position). As a reminder, strain EGAc1 (I-4839) is acetic-acid resistant whereas strain I-4538 is not.

- 25 To better study the role of this mutation, homozygous strains either for allele II-624794 of strain EGAc1 (denoted “EGAc1-II-624794”) or for allele II-624794 of strain I-4538 (denoted “I4538-II-624794”) were prepared. To do this, a cassette called LOH was used. In its principle, the cassette LOH is constituted of a gene (KanMX4) conferring resistance to geneticin to yeasts that express it.
- 30 Another part of the cassette carries the sequence GIN11m86. The latter is toxic for cells that express it (Akada *et al.*, 1997, *Mol Gen Genet* 254, 267-74;

Kawahata *et al.*, 1999, *Yeast*, 15, 1-10 and Akada *et al.*, 2002, *Yeast*, 19, 393-402). This system is qualified as dominant negative because a single copy of this sequence is necessary to confer the lethal genotype on the cells. In so far as the sequence GIN11m86 is placed under the dependence of promoter *GAL2*, it is possible to select on YNB + Galactose the cells that would have lost the cassette.

The strategy used to build homozygous EGAc2 strains for allele II-624794 is as follows:

The LOH cassette was flanked by recombinogenic sequences capable of deleting one of the two alleles. The selection of transformants is conducted on a YNB-G418 medium (containing geneticin). To determine which of the two alleles was kept, the locus was amplified by PCR using seeds, called validation seeds, outside the LOH cassette. The smallest of the resulting PCR fragments (650 bp) was cloned in a plasmid pTOPO then sequenced, from which it was determined which of the two alleles was kept. The strains were retransformed with a sequence identical to the allele kept. The new transformants were selected on a YNB medium containing galactose as only carbon source, allowing the selection of those that lost the LOH cassette. Finally, using validation seeds, PCR was again done. The product of this reaction is unique and was directly sequenced.

Using this strategy, 7 homozygous EGAc2 strains for allele “EGAc1-II-624794” (conferring resistance to acetic acid), denoted “EGAc2-AlleleR/AlleleR”, and 5 homozygous EGAc2 strains for the allele denoted “I4538-II-624794” (conferring sensitivity to acetic acid), denoted “EGAc2-AlleleS/AlleleS”, were built and validated.

b) Analysis of the impact of homozygosity on resistance to acetic acid:

After having obtained the 12 previously cited homozygous strains, their capacity to ferment glucose in the presence of acetic acid was tested. These

performances were measured on a YFAc medium (4000 ppm; pH 4.4), defined as follows:

Glucose	150 g/kg
EXL type J100 (Yeast extract)	5 g/kg
DAP (Diammonium phosphate)	4.7 g/kg
Citric acid	11.4 g/kg
Na citrate (sodium citrate)	13.5 g/kg
Acetic acid	4 g/kg
Tween 80	1 mL/kg
ZnSO ₄ 10.6 g/L;	2 mL/kg
MgSO ₄ 7H ₂ O 400 g/L;	2.5 mL/kg
Thiamine 18.24 g/L;	1 mL/kg
Pyridoxine 5.28 g/L;	1 mL/kg
Biotin (1.76 g/L) + KOH	1 mL/kg
Pantothenate (3.8 g/L)	1 mL/kg
Niacin (8 g/L)	2.5 mL/kg
Myo-inositol (50 g/L)	1 mL/kg
Riboflavin (1 g/L)	1 mL/kg
para-Aminobenzoate (1.2 g/L)	1 mL/kg

- 5 The mass loss results observed during the fermentation on this medium are shown in figure 5. The curves shown are the averages of the values obtained for all strains in the same genotype.

10 The results presented in figure 5 show that strains EGAc2 made homozygous for the sensitive allele (S) at locus II-624794 have a delay initiation in fermentation by more than EGAc2 strain (I-4840) on YFAc. This shows that the loss of allele “EGAc1-II-624794” in a strain EGAc2 for the benefit of a homozygosity of allele “I4538-II-624794” makes the strains obtained more sensitive to acetic acid during the glucose fermentation.

By contrast, the EGAc2 strains made homozygous for allele “EGAc1-II-624794” at locus II-624794 have fermentation kinetics on this medium that is not different than that of EGAc2 strain (I-4840). This result suggests that the homozygosity does not add anything to the strains in this medium.

5

In conclusion, this work reveals that the allele “EGAc1-II-624794” does indeed contribute to resistance to strain EGAc2 (I-4840) regarding acetic acid. However, a single copy seems to be sufficient to confer a resistance phenotype on acetic acid.

10

PATENTKRAV

1. Fremgangsmåde til selektering af en *Saccharomyces cerevisiae* gærstamme, som er resistent overfor en organisk syre, fortrinsvis eddikesyre, under glucose-fermentering, omfattende følgende trin:
 - detektering af tilstedeværelsen, i det mindste ved en allel af stammen, af en Haa1p bindingsposition i sekvensen opstrøms for *MCM7* genet, fortrinsvis sekvensen SEQ ID NO: 3 eller SEQ ID NO: 4; og/eller
 - detektering af tilstedeværelsen, i det mindste ved en allel af stammen, af en base G ved position 624794 på chromosomet II.
2. Fremgangsmåde for opnåelse af en *Saccharomyces cerevisiae* gærstamme, som er resistent overfor en organisk syre, fortrinsvis eddikesyre, under glucose-fermentering, omfattende følgende trin:
 - et sporulationstrin fra to forældrestammer, som har forskellige genomer,
 - et massehybridiseringstrin for de opnåede segregeranter,
 - i det mindste ét selektionstrin for segregeranter baseret på tilstedeværelsen af en Haa1p bindingsposition i sekvensen opstrøms for *MCM7*, fortrinsvis tilstedeværelsen af sekvensen SEQ ID NO: 3 eller SEQ ID NO: 4, og/eller baseret på tilstedeværelsen af en base G ved position 624794 på chromosomet II.
3. Fremgangsmåde ifølge krav 2, **kendetegnet ved, at** den omfatter følgende trin:
 - a) fremstilling af segregeranter ud fra en første forælderstamme og segregeranter ud fra en anden forælderstamme;
 - b) selektering, ud fra segregeranterne fra trin a), af de, som har en Haa1p bindingsposition i sekvensen opstrøms for *MCM7*, fortrinsvis sekvensen SEQ ID NO: 3 eller SEQ ID NO: 4, og/eller af de, som har en base G ved position 624794 på chromosomet II;
 - c) hybridisering af segregeranterne fra den første forælderstamme og selekteret i trin b) med segregeranter fra den anden forælderstamme, eventuelt selekteret i trin b);
 - d) selektering, blandt hybriderne fra trin c), af de, som er resistente overfor den organiske syre.

4. Fremgangsmåde ifølge ethvert af kravene 2 til 3, **kendetegnet ved, at** den første stamme er valgt for dennes evne til at modstå en organisk syre og den anden stamme for en anden egenskab af interesse, eksempelvis dens evne til at metabolisere xylose.
- 5 5. Fremgangsmåde ifølge ethvert af kravene 2 til 4, **kendetegnet ved, at** den første stamme er stammen EGAc1 deponeret ved CNCM under nummeret I-4839 den 13. marts 2014 og den anden stamme er stammen deponeret ved CNCM under nummeret I-4538 den 5. oktober 2011.
- 10 6. Fremgangsmåde ifølge ethvert af kravene 2 til 5, **kendetegnet ved, at** evnen til at modstå den organiske syre afprøves på et medie omfattende glucose.

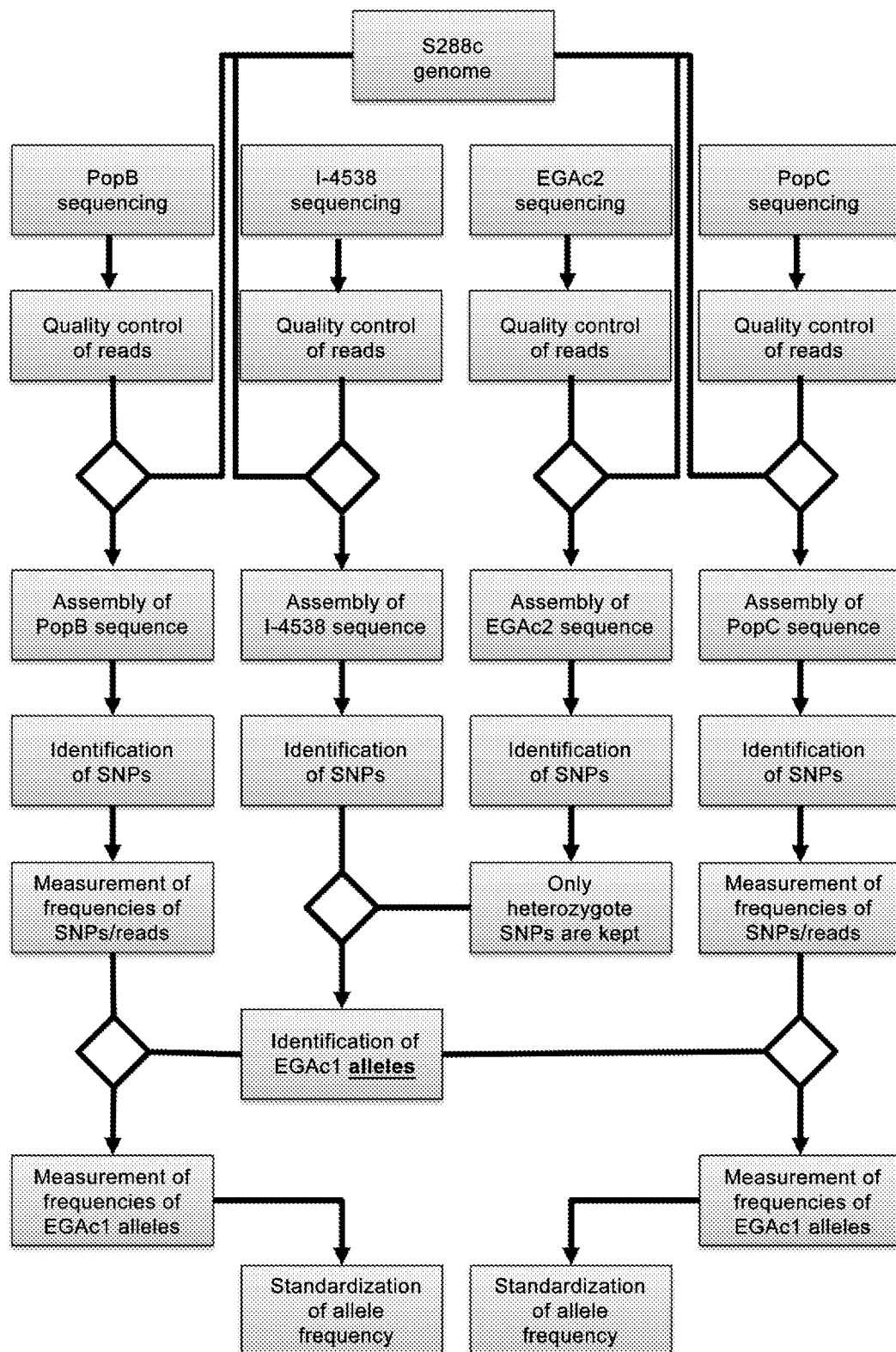
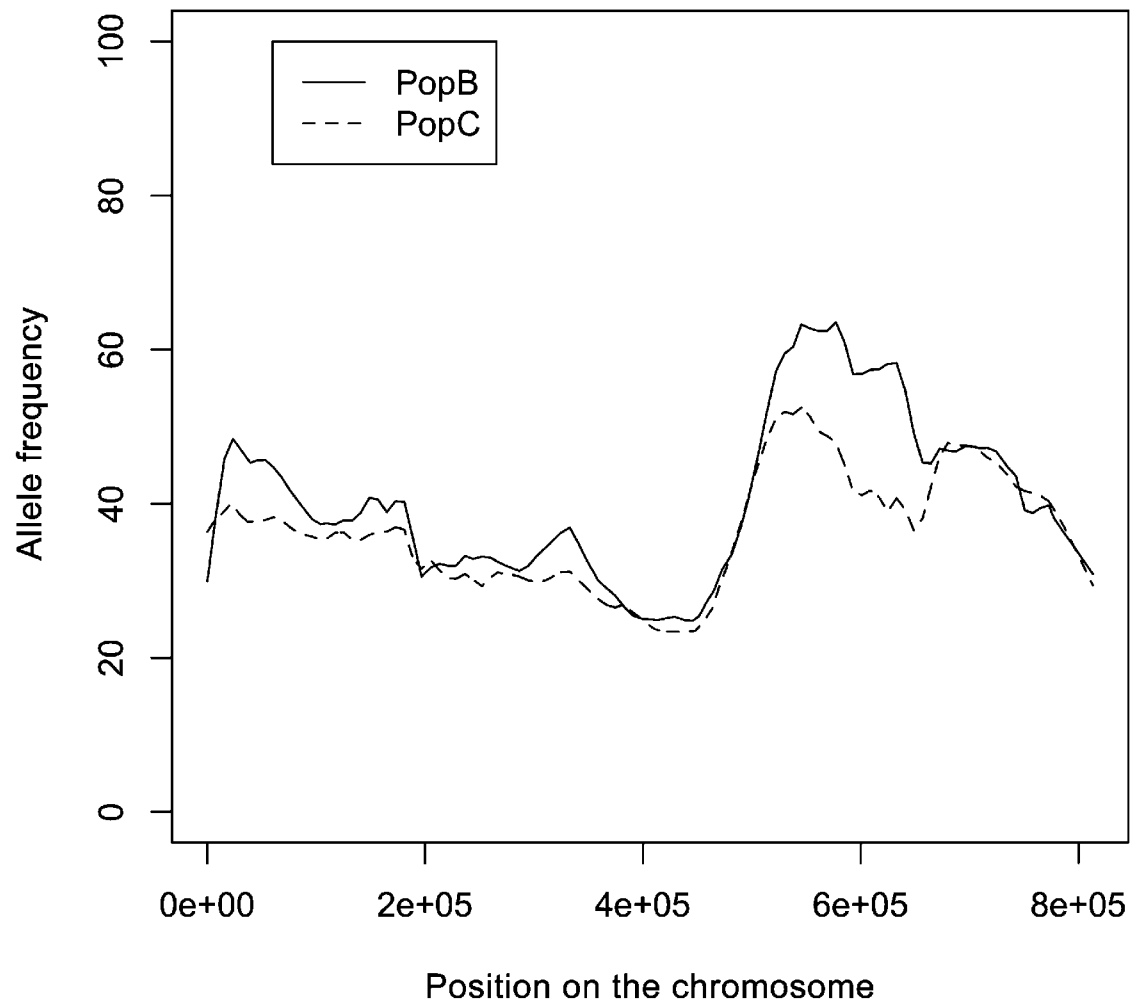
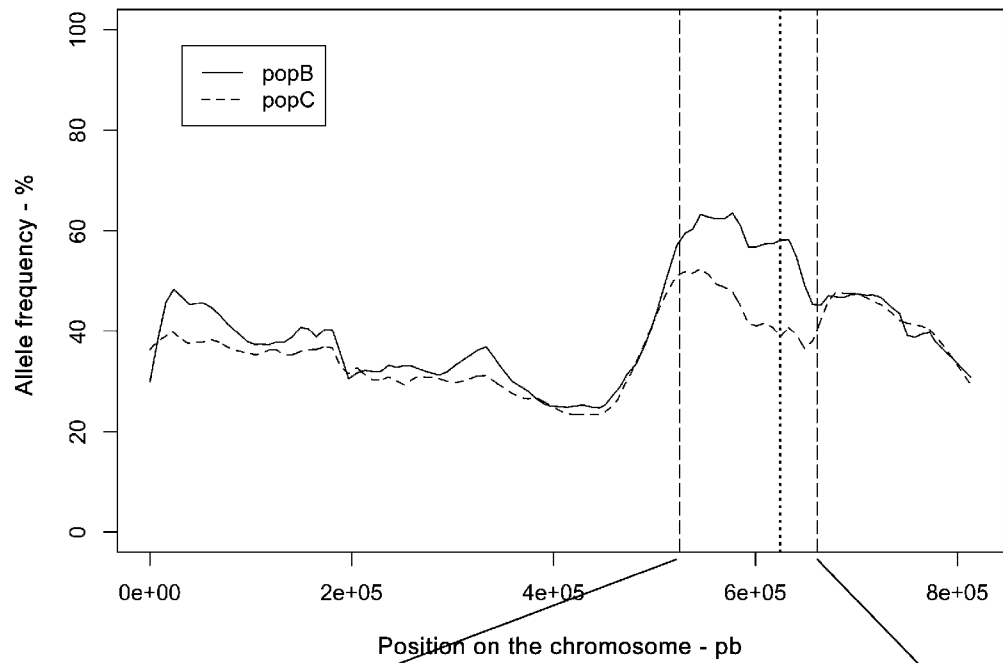
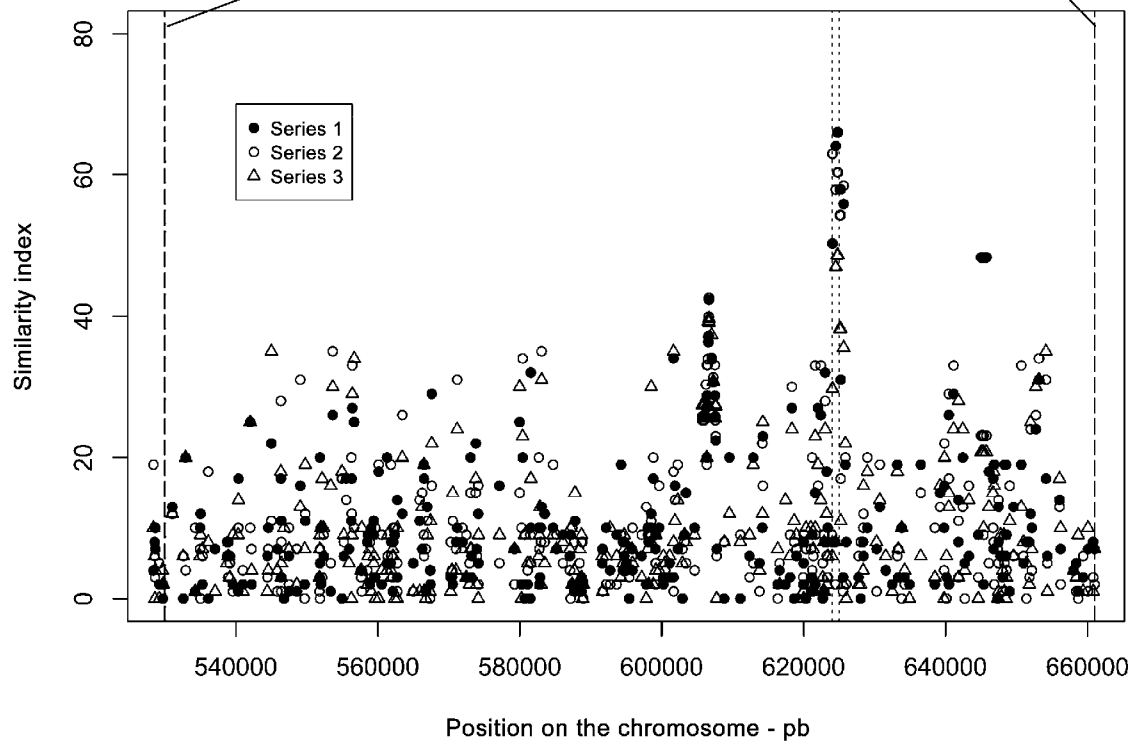
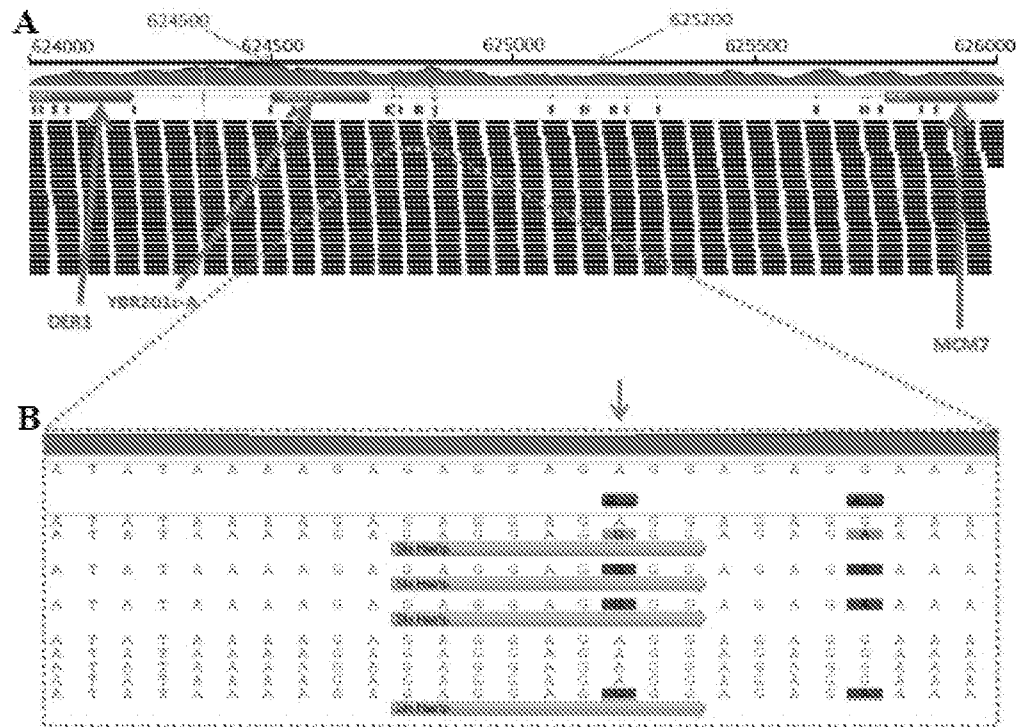
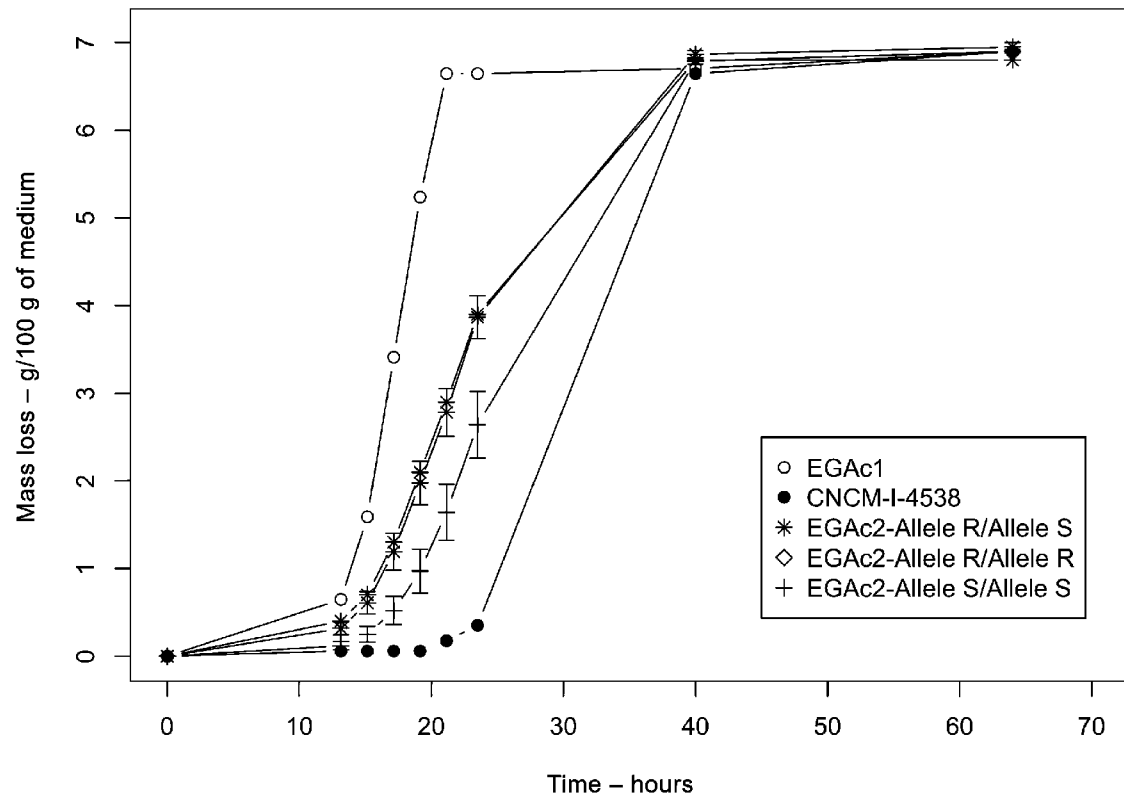


Figure 1

**Figure 2**

A/**B/****Figure 3**

**Figure 4**

**Figure 5**