

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION EN MATIÈRE DE BREVETS (PCT)

(19) Organisation Mondiale de la
Propriété Intellectuelle
Bureau international



(43) Date de la publication internationale
24 septembre 2015 (24.09.2015)

WIPO | PCT

(10) Numéro de publication internationale
WO 2015/140455 A1

(51) Classification internationale des brevets :
C12N 15/90 (2006.01) *C12N 1/14* (2006.01)
C12R 1/885 (2006.01) *C12P 7/10* (2006.01)

(21) Numéro de la demande internationale :
PCT/FR2015/050629

(22) Date de dépôt international :
16 mars 2015 (16.03.2015)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
14 52191 17 mars 2014 (17.03.2014) FR

(71) Déposant : IFP ENERGIES NOUVELLES [FR/FR]; 1
& 4, avenue Bois Préau, F-92500 Rueil Malmaison (FR).

(72) Inventeurs : **POGGI PARODI, Dante**; 9 rue Pestalozzi,
F-75005 Paris (FR). **BIDARD-MICHELOT, Frédérique**;
7, chemin du clos Lardé, F-78620 l'Etang la Ville (FR).
MARGEOT, Antoine; 212 rue Marcadet, F-75018 Paris
(FR). **PORTNOY, Thomas**; 54 rue Notre Dame des
Champs, F-75006 Paris (FR). **LE CROM, Stéphane**; 12
rue Xaintrailles, F-75013 Paris (FR).

(74) Mandataire : **CABINET PLASSERAUD**; 52 rue de la
Victoire, F-75440 Paris Cedex 09 (FR).

(81) États désignés (sauf indication contraire, pour tout titre
de protection nationale disponible) : 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) États désignés (sauf indication contraire, pour tout titre
de protection régionale disponible) : ARIPO (BW, GH,
GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ,
TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, RU,
TJ, TM), européen (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).

Publiée :

- avec rapport de recherche internationale (Art. 21(3))
- avec la partie de la description réservée au listage des séquences (règle 5.2.a))

(54) Title : MUTANT STRAINS OF *TRICHODERMA REESEI*

(54) Titre : SOUCHES MUTANTES DE *TRICHODERMA REESEI*

(57) Abstract : The invention relates to mutant strains of *Trichoderma reesei*, characterized by an improvement in the enzyme production thereof relative to the non-modified *Trichoderma reesei* strain.

(57) Abrégé : Souches mutantes de *Trichoderma reesei* L'invention concerne des souches mutantes de *Trichoderma reesei* caractérisées par une amélioration de leur production d'enzymes par rapport à la souche *T. reesei* non modifiée.



WO 2015/140455 A1

Souches mutantes de *Trichoderma reesei*

La possibilité de produire de l'éthanol à partir de la cellulose a reçu beaucoup d'attention en raison de la disponibilité de grandes quantités de matière première ainsi que de l'intérêt de l'éthanol à titre de carburant. Les matières premières naturelles cellulosiques pour un tel processus sont désignées par le terme "biomasse". De nombreux types de biomasse, par exemple le bois, les résidus agricoles, les cultures herbacées et les déchets solides municipaux, ont été considérés comme des matières premières potentielles pour la production de biocarburant. Ces matières sont constituées principalement de cellulose, d'hémicellulose et de lignine.

La cellulose est un polymère constitué de molécules de glucose reliées par des liaisons beta 1-4, qui sont très résistantes à la dégradation ou à la dépolymérisation. Une fois la cellulose convertie en glucose, celui-ci est facilement fermenté en biocarburant, par exemple l'éthanol, en utilisant une levure.

Les plus anciennes méthodes étudiées pour convertir la cellulose en glucose sont basées sur l'hydrolyse acide. Ce processus peut se faire en présence d'acides concentrés ou dilués. Cependant, plusieurs inconvénients tels que la mauvaise récupération de l'acide lors de l'utilisation d'acides concentrés et la faible production de glucose dans le cadre de l'utilisation d'acides dilués nuisent à l'économie du processus d'hydrolyse acide.

Pour surmonter les inconvénients du processus d'hydrolyse acide, les processus de conversion de la cellulose ont porté plus récemment sur l'hydrolyse enzymatique, à l'aide d'enzymes de type cellulase. Cette hydrolyse enzymatique de la biomasse lignocellulosique (par exemple, la cellulose) présente cependant l'inconvénient d'être un procédé industriel coûteux. De ce fait, il est nécessaire d'utiliser des souches de microorganismes sécréteurs de cellulases de plus en plus performantes. A ce titre, beaucoup de microorganismes comportent des enzymes qui hydrolysent la cellulose, tels que les champignons *Trichoderma*, *Aspergillus*,

Humicola, *Fusarium* ainsi que les bactéries telles que *Thermomonospora*, *Bacillus*, *Cellulomonas* et *Streptomyces*.

Pour diminuer le coût associé à l'hydrolyse enzymatique de la lignocellulose, l'industrie est à la recherche constante de souches plus productives ou présentant un meilleur rendement.

Il existe donc un besoin d'optimisation de la production d'enzymes dans le processus industriel. Ainsi, pris dans son ensemble, l'art antérieur fait apparaître un besoin inassouvi et longtemps attendu de développer des souches aux performances de productions d'enzymes améliorées. C'est le problème que se propose de résoudre la présente invention.

A l'issue de long et fastidieux travaux sur le transcriptome de *Trichoderma reesei* les inventeurs ont montré que 592 gènes au total étaient induits ou réprimés au cours de la procédé de production d'enzymes dans une souche de ce champignon. Parmi ces 592 gènes, les inventeurs ont mis en évidence le rôle d'un d'entre eux, précisément PacC, dans la régulation de la production d'enzymes.

Les inventeurs ont ainsi développé une souche mutante de *Trichoderma reesei* caractérisée par une amélioration significative de la production d'enzymes, en particulier des enzymes cellulolytiques.

Aussi, l'invention concerne une souche de *Trichoderma reesei* dont le génome comprend une mutation choisie parmi:

- une délétion du gène codant pour PacC;
- une mutation entre les positions 1 à 735 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 249 à 495;
- une mutation entre les positions 736 à 744 de la séquence nucléotidique représentée en SEQ ID N°1; et
- une mutation entre les positions 1438 à 1515 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 1492 à 1494.

Cette souche mutée est caractérisée par une amélioration de la production d'enzymes cellulolytiques par rapport à la même souche de *T. reesei* non modifiée, ou par rapport à une souche de référence de *T. reesei*.

5

Trichoderma reesei est un champignon filamenteux cellulolytique. Compte tenu de la capacité de *T. reesei* à sécréter de grandes quantités de cellulases et des hémicellulases, cette souche est hautement intéressante pour la production d'enzymes pour la transformation des matières de la biomasse végétale en bioproduits utiles à l'industrie, tels que le bioéthanol.

10

Par "**souche de référence de *T. reesei***", on entend une souche de *Trichoderma reesei* choisie parmi les souches QM6a, NG14, RUTC30 et QM9414. Ces souches sont accessibles au public et ont notamment fait l'objet de dépôts, respectivement sous les numéros:

15

- ATCC 13631 (souche QM6a);
- ATCC 56767 (souche NG14);
- ATCC 56765 (souche RUT C30); et
- ATCC 26921 (souche QM9414).

20

Dans un mode de réalisation particulier, la souche est la souche CL847. Cette souche est une souche hyperproductive.

Typiquement, cette souche de référence ne présente aucune mutation sur le gène de PacC.

25

Dans le cadre de la présente invention, les termes "**PacC**" et "**RIM101**" font référence à un facteur de transcription dont la séquence protéique native est représentée en SEQ ID N°2. Cette séquence est disponible en ligne sous les numéros d'accès GenBank: EGR50864.1 (souche *T. reesei* QM6a) ou encore GenBank: ETS05643.1 (souche *T. reesei* RUT C-30) et sur le site dédié au génome de *T. reesei*, sous le numéro TR120698 à l'adresse suivante (<http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Trire2&id=120698>).

30

A moins que cela ne soit indiqué autrement, les termes PacC et RIM101 font référence à la protéine telle que retrouvée dans la souche *Trichoderma reesei*.

5 Le facteur PacC comprend un motif en doigt de zinc Cys2-His2. Les protéines présentant un motif en doigt de zinc sont caractérisées par leur capacité à se replier autour d'un ion Zn^{2+} , produisant ainsi un domaine compact. Le motif en doigt de zinc Cys2-His2 est le motif le plus courant de liaison à l'ADN. Il présente une séquence consensus de 23 à 26 résidus avec deux résidus cystéine et deux résidus histidine conservés, dont les chaînes latérales fixent un ion Zn^{2+} . Il s'ensuit que la
10 séquence consensus d'un doigt de zinc est comme suit:

Cys-X2-4-Cys-X-Phe-X5-Leu-X2-His-X3-His.

PacC a fait l'objet de plusieurs publications scientifiques du fait de son rôle en tant que senseur de pH, c'est-à-dire en tant que facteur dont l'expression est
15 modulée en fonction du pH ambiant de la cellule. L'activité de PacC est régulée par traitement protéolytique, comme décrit notamment dans la publication Diez et al., "*Activation of the Aspergillus PacC zinc finger transcription factor requires two proteolytic steps*", The EMBO Journal, Vol 21, No 6, pp1350-1359, 2002.

20 En effet, sous conditions acides, la forme non clivée de PacC prédomine. Cependant, dans des conditions alcalines, PacC fait l'objet d'un traitement protéolytique supprimant la queue C-terminale de PacC. La forme ainsi clivée de PacC agit alors comme un répresseur transcriptionnel et régule un ensemble de gènes cibles afin d'induire des modifications adaptatives en réponse à ce
25 changement de pH. Ce traitement protéolytique de PacC induit par le pH de la cellule est sous le contrôle de la voie biologique dite "voie de PacC" comme décrit dans la publication Maeda et al., "*The signaling mechanism of ambient pH sensing and adaptation in yeast and fungi*", FEBS Journal 279 (2012).

30 Les inventeurs ont maintenant montré, pour la première fois, que cette voie biologique n'est pas uniquement responsable d'un effet senseur de pH. En effet, l'inactivation de cette voie conduit à une augmentation significative de la production de protéines cellulolytiques.

Les inventeurs ont en outre déterminé, par comparaison avec l'homologue de PacC chez *Aspergillus nidulans*, que dans *T. reesei*, la maturation de la protéine PacC se fait en deux étapes essentielles de clivage aux positions suivantes:

- un premier clivage entre les positions 480 et 503 de la séquence protéique de PacC chez *T. reesei*; et
- un second clivage entre les positions 246 et 248 de la séquence protéique de PacC chez *T. reesei*.

Les inventeurs ont ainsi mis en évidence que des mutations sur ces zones de clivage spécifiques conduisent à la production d'une protéine PacC, dont la maturation post-traductionnelle est altérée et/ou empêchée. Une telle protéine est alors inactive, et la souche de *T. reesei* ainsi modifiée présente une amélioration de la production d'enzymes, en particulier d'enzymes cellulolytiques.

Lesdites enzymes cellulolytiques sont préférentiellement une cellulase ou une hémicellulase, encore plus préférentiellement il s'agit d'une cellulase.

Par "**séquence protéique native de PacC**", on entend désigner la séquence non mature telle qu'elle est produite, avant toute modification post-traductionnelle. La séquence protéique native de PacC est représentée en SEQ ID N°2 et compte 622 acides aminés.

Par "**protéine active ou protéine mature de PacC**", on entend désigner la protéine de PacC telle qu'obtenue à l'issue de modifications post-traductionnelles, en particulier d'un double clivage de ladite protéine native. La protéine mature de PacC comprend entre 246 et 248 acides aminés. Plus précisément, la séquence protéique de la protéine mature correspond respectivement aux 246 ou 248 premiers acides aminés de la séquence protéique de la protéine native (représentée en SEQ ID N°2).

Par "**protéine inactive**", ou "**protéine immature**", on entend désigner une protéine PacC dont l'activité est altérée.

Par "**souche mutante**", "**mutant**", "**souche modifiée**", "**souche variante**" ou encore "**transformants**", on entend désigner de manière indifférente dans la présente demande une souche de *Trichoderma reesei* selon l'invention, qui n'exprime pas la protéine PacC ou qui exprime une protéine immature. Cette

- 5 souche est caractérisée par la présence d'une mutation choisie parmi :
- une délétion du gène codant pour PacC, conduisant à l'absence complète de l'expression du gène, ou à la production d'une protéine immature ;
 - une mutation entre les positions 1 à 735 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 249
10 à 495;
 - une mutation entre les positions 736 à 744 de la séquence nucléotidique représentée en SEQ ID N°1; et
 - une mutation entre les positions 1438 à 1515 de la séquence
15 nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 1492 à 1494.

Ces souches sont caractérisées par:

- l'absence de production de protéine PacC; ou
- 20 - la production d'une protéine inactive PacC.

Dans ce dernier cas, ladite protéine inactive est préférentiellement tronquée. Elle se caractérise typiquement par une mutation au niveau protéique :

- entre les positions 1 à 245 de la séquence protéique native de PacC,
25 préférentiellement entre les positions 84 à 165 ;
- entre les positions 246 à 248 de la séquence protéique native de PacC ;
et/ou
- entre les positions 480 à 503 de la séquence protéique native de PacC,
préférentiellement en position 498, encore plus préférentiellement une
30 substitution de la leucine en position 498 en sérine.

Dans un mode de réalisation particulier, l'invention concerne une souche de *Trichoderma reesei* dont le génome comprend une mutation choisie parmi:

- une mutation entre les positions 1 à 735 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 249 à 495 ;
- une mutation entre les positions 736 à 744 de la séquence nucléotidique représentée en SEQ ID N°1 ; et
- une mutation entre les positions 1438 à 1515 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 1492 à 1494.

10 Les techniques de mutations sont connues de l'homme du métier. Préférentiellement, les mutations seront obtenues par les techniques de recombinaisons génétiques. Dans le cas d'introduction d'une séquence d'ADN, on préférera particulièrement l'introduction d'une cassette de délétion.

15 Bien entendu, il est possible de prévoir d'autres techniques assurant la mutation en cause, on pourra par exemple utiliser les systèmes de mutations connus de l'homme du métier, ainsi que les insertions ou autres modifications d'ADN obtenues par les techniques de recombinaison homologue. Par exemple des approches de mutagenèse dirigée par PCR ou de synthèse de gène pourraient être
20 utilisées pour créer un gène tronqué ou muté dans les séquences essentielles à la maturation. La connaissance de telles méthodes appartient aux connaissances de l'homme du métier.

La cassette ainsi obtenue pourrait être utilisée pour remplacer la copie non mutée
25 de PacC par recombinaison homologue dans le génome de la souche de référence.

Alternativement, l'homme du métier peut cibler l'activité de la protéine PacC, notamment à l'aide d'ARN interférent. Ces ARN permettent de réduire fortement l'expression d'une protéine. Pour ce faire, on introduit un petit ARN double brin
30 qui entraîne l'inactivation spécifique d'un ARN messager.

Par conséquent, dans un premier mode de réalisation, l'invention concerne une souche dans laquelle le gène codant pour PacC est délété. Une telle délétion

conduit indifféremment à l'absence de protéine PacC mature ou l'expression d'une protéine inactive, de préférence tronquée ou incomplète.

Préférentiellement, cette délétion est mise en œuvre par recombinaison homologue, typiquement à l'aide d'un vecteur.

5

Selon l'invention, on entend par « **vecteur** » toute séquence d'ADN dans laquelle il est possible d'insérer des fragments d'acide nucléique étranger, les vecteurs permettant d'introduire de l'ADN étranger dans une cellule hôte. Des exemples de vecteurs sont les plasmides, les cosmides, les chromosomes artificiels de levures (YAC), les chromosomes artificiels de bactéries (BAC) et les chromosomes artificiels dérivés du bactériophage P1 (PAC), les vecteurs dérivés de virus. Le vecteur selon l'invention permet l'introduction d'une mutation ou d'une délétion.

Selon l'invention, l'acide nucléique tel que décrit précédemment pourra être lié opérationnellement à un promoteur, un terminateur ou toute autre séquence nécessaire à son expression dans la cellule hôte. Le vecteur selon l'invention pourra également porter un marqueur de sélection. On entend par « marqueur de sélection » un gène dont l'expression confère aux cellules qui le contiennent une caractéristique permettant de les sélectionner. Il s'agit par exemple d'un gène de résistance aux antibiotiques.

20

Préférentiellement, ledit vecteur est un plasmide. L'homme du métier pourra utiliser un plasmide de type :

- YIP (Yeast Integrating Plasmid) : ce sont des vecteurs auxquels il manque l'origine de répllication dans la levure, donc qui doivent se propager comme un élément intégratif dans le chromosome de la levure et le plus souvent dans une seule copie par génome;
- YRP (Yeast Replicatif Plasmid) : ce sont des vecteurs qui présentent une séquence de répllication autonome (ARS), ce sont des plasmides instables et à répllication autonome;
- YCP (Yeast Centromer Plasmid) : ces vecteurs présentent une origine de répllication et une séquence centromérique, ce sont des plasmides stables et qui se propagent en un faible nombre de copies;

25

30

- YEP (Yeast Episomal Plasmid) : ces vecteurs présentent un fragment de plasmide de levure de 2um, ils se propagent en un grand nombre de copies et ils ont une réplication autonome et une ségrégation irrégulière.

5

Plus préférentiellement, le plasmide est un plasmide de type YEP. Encore plus préférentiellement, il s'agit du plasmide pRS426. Ainsi, pour déléter le gène de PacC, l'homme du métier pourra avoir recours à un plasmide modifié pRS426 comprenant une cassette de délétion. Cette cassette de délétion vise à supprimer le gène de PacC, préférentiellement tel qu'il est représenté en SEQ ID N°3. Cette séquence SEQ ID N°3 représente la séquence génique de PacC (et inclut les introns).

La cassette de délétion est obtenue en plaçant une séquence d'intérêt sous le contrôle transcriptionnel d'un promoteur et d'un terminateur provenant d'un gène de levure. Le gène d'intérêt peut être modifié à l'aide d'enzymes de restrictions, en y plaçant par exemple la séquence codant pour un gène de sélection. Préférentiellement, ledit gène de sélection peut être choisi parmi le gène de orotidine 5-phosphate décarboxylase (pyr4) et le gène hygromycine B phosphotransférase (Hph).

Ainsi, la traduction du gène d'intérêt donne une protéine inactive (de préférence tronquée ou incomplète) ou la protéine d'intérêt n'est pas exprimée, du moins sous forme mature. Le transgène peut alors s'intégrer par recombinaison homologue au locus visé. Il faut alors trier les cellules qui ont subi la recombinaison homologue. Pour cela, il suffit de digérer l'ADN génomique par des enzymes de restriction qui permettent de discriminer le gène endogène non touché par l'insertion et le gène endogène modifié par le transgène.

Lors d'une recombinaison homologue, deux éléments de recombinaison, un à chaque extrémité du gène, sont nécessaires pour cibler de façon précise le gène à déléter. Aussi, dans un mode de réalisation préféré, ladite cassette de délétion comprend trois fragments d'ADN:

- une région en amont du gène cible;

- un marqueur sélectif, préférentiellement choisi parmi le gène de orotidine 5-phosphate decarboxylase de *T.reesei* (pyr4) et le gène hygromycine B phosphotransférase (Hph); et
- une région en aval du gène cible.

5

Typiquement, les amorces utilisées pour l'amplification de la région en amont du gène cible chez *T.reesei* sont comme suit:

- l'amorce sens est constituée de 29 nucléotides homologues de la séquence du plasmide pRS426, à proximité du site de restriction XhoI. Cette amorce sens est représentée en SEQ ID N°4 (GTAACGCCAGGGTTTCCCAGTCACGACG); et
- l'amorce antisens contient une séquence de 21 nucléotides homologues de la partie 3' du gène du marqueur sélectif pyr4, tels que représentés en SEQ ID N°5 (CGACGATATCAGCTTCCATAT), et 8 nucléotides du site de restriction MmeI, tels que représentés en SEQ ID N°6 (TCCGACTA), le tout formant la séquence SEQ ID N°7 (CGACGATATCAGCTTCCATATTCCGACTA).

15

Typiquement, les amorces antisens utilisées pour l'amplification de la région en aval du gène cible chez *T. reesei* sont comme suit :

20

- l'amorce sens est constituée de 21 nucléotides homologues de la séquence en 5' du gène du gène du marqueur sélectif pyr4, tels que représentés en SEQ ID N°8 (AGAAAAGCACAAAGAAGAGGC), et 8 nucléotides du site de restriction MmeI, tels que représentés en SEQ ID N°9 (TCCA ACTA), le tout formant la séquence SEQ ID N°10 (AGAAAAGCACAAAGAAGAGGCTCCA ACTA); et
- l'amorce antisens contient une séquence de 29 nucléotides homologues de la séquence du plasmide pRS426, à proximité du site de restriction EcoRI. Cette amorce sens est représentée en SEQ ID N°11 (GCGGATAACAATTTACACAGGAAACAGC).

25

30

La cassette de délétion peut être ensuite amplifiée selon les techniques classiques bien connues de l'homme du métier, typiquement par une méthode choisie parmi

le clonage classique, la fusion PCR, ou encore le clonage *in vivo* dans la levure par PCR.

Préférentiellement, cette cassette de délétion est amplifiée par PCR.

- 5 La cassette de délétion est ensuite introduite par recombinaison dans une souche de *T.reesei* qui n'exprime pas un gène du marqueur de sélection, également appelée gène sélectif. L'homme du métier peut aisément identifier les gènes de marqueur de sélection pertinents pour la mise en œuvre de l'invention.
- 10 Préférentiellement, ladite souche recombinée est une souche préalablement déléetée pour le gène *pyr4* ou le gène *Hph*.

- Plus préférentiellement, ladite souche recombinée est une souche préalablement déléetée pour le gène *pyr4*. L'homme du métier pourra se référer notamment à la
- 15 publication Schuster, A., Bruno, K. S., Collett, J. R., Baker, S. E., Seiboth, B., Kubicek, C. P., & Schmoll, M. (2012). *A versatile toolkit for high throughput functional genomics with Trichoderma reesei*, Biotechnology for Biofuels, 5(1), 1 (doi:10.1186/1754-6834-5-1).

- 20 Après culture, les souches mutantes ayant incorporé la cassette de délétion sont sélectionnées en fonction de l'expression ou non du marqueur de sélection; les clones ayant été transformés exprimant ledit marqueur de sélection.

- Dans un second mode de réalisation, l'invention concerne une souche dans
- 25 laquelle une mutation est introduite, de telle sorte que la protéine PacC obtenue est inactive. Préférentiellement, ladite mutation est choisie parmi :

- une mutation entre les positions 1 à 735 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 249 à 495;
- 30 - une mutation entre les positions 736 et 744 de la séquence nucléotidique représentée en SEQ ID N°1; et
- une mutation entre les positions 1438 et 1515 de la séquence nucléotidique représentée en SEQ ID N°1.

La séquence SEQ ID N°1 correspond à la séquence nucléotidique codant pour la protéine PacC. Plus précisément, il s'agit d'un ADN complémentaire.

Ces mutations peuvent être insérées comme décrit précédemment à l'aide d'un plasmide, par exemple pUC19.

5

L'insertion de la mutation entre les positions 1 à 735 de la séquence nucléotidique représentée en SEQ ID N°1 est exemplifiée à l'exemple 9. Typiquement, cette mutation peut être mise en œuvre grâce à la séquence telle que représentée en SEQ ID N°30. Cette séquence est ensuite amplifiée, par exemple par PCR à l'aide d'amorces sens et antisens, respectivement représentées en SEQ ID N°31 et 32. Après amplification, le produit de PCR correspondant ainsi que le marqueur sélectif sont introduits dans une souche *T. reesei*.

15 L'insertion de la mutation entre les positions 249 à 495 de la séquence nucléotidique représentée en SEQ ID N°1 est exemplifiée à l'exemple 10. Typiquement, cette mutation peut être mise en œuvre grâce à la séquence telle que représentée en SEQ ID N°40. Cette séquence est ensuite amplifiée, par exemple par PCR à l'aide d'amorces sens et antisens, respectivement représentées en SEQ ID N°31 et 32. Après amplification, le produit de PCR correspondant ainsi que le
20 marqueur sélectif sont introduits dans une souche *T. reesei*.

L'insertion de la mutation entre les positions 736 à 744 de la séquence nucléotidique représentée en SEQ ID N°1 est exemplifiée à l'exemple 11. Typiquement, cette mutation peut être mise en œuvre grâce à la séquence telle que
25 représentée en SEQ ID N°43. Cette séquence est ensuite amplifiée, par exemple par PCR à l'aide d'amorces sens et antisens, respectivement représentées en SEQ ID N°31 et 32. Après amplification, le produit de PCR correspondant ainsi que le marqueur sélectif sont introduits dans une souche *T. reesei*.

30 L'insertion de la mutation entre les positions 1438 à 1515 de la séquence nucléotidique représentée en SEQ ID N°1 est exemplifiée à l'exemple 12. Typiquement, cette mutation peut être mise en œuvre grâce à la séquence telle que représentée en SEQ ID N°47. Cette séquence est ensuite amplifiée, par exemple par PCR à l'aide d'amorces sens et antisens, respectivement représentées en SEQ

ID N°48 et 49. Après amplification, le produit de PCR correspondant ainsi que le marqueur sélectif sont introduits dans une souche *T. reesei*.

5 L'insertion de la mutation entre les positions 1492 à 1494 de la séquence nucléotidique représentée en SEQ ID N°1 est exemplifiée à l'exemple 13. Typiquement, cette mutation peut être mise en œuvre grâce à la séquence telle que représentée en SEQ ID N°54. Cette séquence est ensuite amplifiée, par exemple par PCR à l'aide d'amorces sens et antisens, respectivement représentées en SEQ ID N°48 et 49. Après amplification, le produit de PCR correspondant ainsi que le
10 marqueur sélectif sont introduits dans une souche *T. reesei*.

Ces mutations au niveau nucléotidique conduisent à la production d'une protéine inactive. Plus précisément, ces mutations sont respectivement responsables :

- d'une mutation entre les positions 1 à 245 de la séquence protéique de
15 PacC, préférentiellement entre les positions 84 à 165;
- d'une mutation entre les positions 246 et 248 de la séquence protéique de PacC; et
- d'une mutation entre les positions 480 et 503 de la séquence de la
séquence protéique de PacC, préférentiellement en la position 498.

20

La séquence protéique de PacC est représentée en SEQ ID N°2.

Ces mutations sont particulièrement pertinentes dans le cadre de l'invention comme détaillé ci-après.

25 Dans un premier aspect, la mutation concerne les positions 1 à 245 de la séquence protéique de PacC. Ces positions correspondent au domaine de fixation à l'ADN. Plus spécifiquement, la mutation concerne les positions 84 à 165 de la séquence protéique de PacC, qui comprennent les motifs à doigt de zinc.

30 Dans un second aspect, la mutation concerne les positions 246 à 248 de la séquence protéique. Ces positions correspondent à une zone de clivage de la protéine PacC, ledit clivage étant nécessaire à la maturation de ladite protéine. Une mutation sur ces positions conduit donc à une absence de clivage de la protéine PacC et ainsi son maintien à un état inactivé.

Dans un troisième aspect, la mutation concerne les positions 480 à 503 de la séquence protéique de PacC. Ces positions correspondent à une zone de clivage de la protéine PacC, ledit clivage étant nécessaire à la maturation de ladite protéine. Une mutation sur ces positions conduit donc également à une absence de clivage de la protéine PacC et ainsi son maintien à un état inactivé.

Par conséquent, l'invention concerne également une souche de *T. reesei* exprimant une protéine PacC inactive, qui présente au moins une mutation :

- 10 - entre les positions 1 à 245 de la séquence protéique de PacC, c'est à dire sur le domaine de fixation à l'ADN de la protéine PacC, préférentiellement entre les positions 84 à 165 de ladite séquence, qui correspondent au domaine à doigt de zinc;
- 15 - entre les positions 246 à 248 de la séquence protéique de PacC, qui correspondent aux positions auxquelles un clivage a lieu lors de la maturation post-traductionnelle de la protéine PacC; ou
- 20 - entre les positions 480 à 503 de la séquence protéique de PacC, préférentiellement en la position 498, plus préférentiellement une mutation Leu498Ser, qui correspond aux positions auxquelles un clivage a lieu lors de la maturation post-traductionnelle de la protéine PacC.

Dans un troisième mode de réalisation, l'invention concerne une souche de *T. reesei* dont le génome présente une mutation affectant l'expression d'une protéine choisie parmi RIM8, RIM9, RIM13, RIM20, RIM21. Les séquences de ces protéines sont accessibles en ligne.

RIM13 (également dénommé Palb) se retrouve notamment dans la souche *T. reesei* Tr 70560 et sa séquence est disponible sous le lien: <http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Trire2&id=70560>.

30 RIM20 (également dénommé PalA) se retrouve notamment dans la souche *T. reesei* Tr124132 et sa séquence est disponible sous le lien: <http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Trire2&id=124132>.

RIM9 (également dénommé PalL) se retrouve dans la souche *T. reesei* Tr21415 et sa séquence est disponible sous le lien: <http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Trire2&id=21415>.

5 RIM21 (également dénommé PalH) se retrouve dans la souche *T. reesei* Tr123296 et sa séquence est disponible sous le lien: <http://genome.jgi-psf.org/cgi-bin/dispGeneModel?db=Trire2&id=123296>.

Ces mutations ont pour effet l'absence de production desdites protéines ou l'expression de telles protéines inactives. Ces mutations dans une souche de
10 *T. reesei* conduisent à une absence du traitement protéolytique de PacC, ce qui améliore la production d'enzymes par la souche ainsi mutée. En effet, l'organisation et le fonctionnement de la voie RIM101 sont conservés dans de nombreuses levures et champignons filamenteux, notamment *S. cerevisiae* et *A. nidulans*. La voie se compose de plusieurs composants dont RIM8, RIM9,
15 RIM13, RIM20, RIM21. Des souches dont au moins l'une de ces composantes a été supprimée sont caractérisées par une absence du traitement protéolytique de PacC, ce qui confirme leur rôle dans le déclenchement de l'activité de PacC.

L'invention porte également sur l'utilisation de la souche selon l'invention pour la
20 production d'enzymes cellulolytiques. Aussi, dans un mode de réalisation particulier, l'invention concerne l'utilisation d'une souche de *Trichoderma reesei* pour la production d'enzymes cellulolytiques, ladite souche étant choisie parmi:

- une délétion du gène codant pour PacC;
- une mutation entre les positions 1 à 735 de la séquence nucléotidique
25 représentée en SEQ ID N°1, préférentiellement entre les positions 249 à 495;
- une mutation entre les positions 736 à 744 de la séquence nucléotidique représentée en SEQ ID N°1; et
- une mutation entre les positions 1438 à 1515 de la séquence
30 nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 1492 à 1494.

L'invention porte également sur l'utilisation de la souche selon l'invention pour l'hydrolyse du cellobiose en glucose.

L'invention a également pour objet, l'utilisation de la souche selon l'invention pour la production de biocarburant. Selon l'invention, le terme "**biocarburant**" peut être défini comme tout produit issu de la transformation de la biomasse et
5 pouvant être utilisé à des fins énergétiques. D'une part et sans vouloir se limiter, on peut citer à titre d'exemple des biogaz, des produits pouvant être incorporés (éventuellement après transformation ultérieure) à un carburant ou être un carburant à part entière, tels que des alcools (l'éthanol, le butanol et/ou l'isopropanol selon le type d'organisme fermentaire utilisé), des solvants
10 (acétone), des acides (butyrique), des lipides et leurs dérivés (acides gras à courtes ou longues chaînes, esters d'acides gras), ainsi que l'hydrogène.

De manière préférée, le biocarburant selon l'invention est un alcool, par exemple l'éthanol, le butanol et/ou l'isopropanol. Plus préférentiellement, le biocarburant
15 selon l'invention est l'éthanol. Dans un autre mode de réalisation, le biocarburant est du biogaz.

L'invention concerne également l'utilisation de la souche selon l'invention pour l'hydrolyse des beta-oligosaccharides.

20

Enfin, l'invention concerne l'utilisation de la souche selon l'invention pour l'hydrolyse du cellobiose en glucose.

FIGURES

Figure 1: Représentation du plasmide pRS426.

- 5 **Figure 2:** Production de protéines extracellulaires par des souches mutantes selon l'invention et des souches *T. reesei* non modifiées après 7 jours de culture.

Figure 3: Productivité spécifique des souches mutantes selon l'invention après 7 jours de culture.

10

Figure 4: Production de protéines extracellulaires pour un mutant selon l'invention par rapport à la souche de *T. reesei* non modifiée

15

EXEMPLES

Exemple 1 : Culture de *Trichoderma reesei* pour l'analyse du transcriptome et analyse des réponses cellulaires lors de l'induction de la production

d'enzymes cellulolytiques

Protocole de culture

Dans cette étude, les inventeurs ont utilisé les souches de *Trichoderma reesei* suivantes:

- 10 - NG14 (ATCC 56767), et
- RUT C30 (ATCC 56765).

Des spores congelés ont été utilisés pour ensemercer une fiole de Fernbach contenant 250 ml de milieu de culture ayant la composition suivante: glucose à 30
15 g.L⁻¹ ; cornsteep 2 g.L⁻¹ , (NH₄)₂SO₄ à 1,4 g L⁻¹ ; KOH à 0,8 g.L⁻¹ ; H₃PO₄ à 85 %
4 mL.L⁻¹ ; acide phthalique , sel dipotassique 5 g L⁻¹ ; MgSO₄.7H₂O à 0,3 g L⁻¹ ;
CaCl₂ à 0,3 g L⁻¹ ; FeSO₄.7H₂O à 5,0 mg L⁻¹ ; MnSO₄.H₂O à 1,6 mg.L⁻¹ ;
ZnSO₄.7H₂O à 1,4 mg.L⁻¹ ; CoCl₂.6H₂O à 2,0 mg.L⁻¹ ; à un pH de 6.

La culture a été effectuée à 30°C sous agitation à 150 tours par minute (rpm).

20 Après 72 h, le milieu contenant le mycélium a été utilisé comme inoculum pour la culture en bioréacteur.

Pour l'induction de la production des enzymes cellulolytiques, un bioréacteur de 4 litres a été utilisé durant un protocole de culture en deux étapes. Les souches ont d'abord été cultivées à 28 ° C dans 2 L de milieu décrit ci-dessus et le pH régulé à
25 4,8 avec du NH₃ à 5,5 M.

Le débit d'air a été ajusté à 0,5 volume par volume par min (VVM), et l'agitation initiale a été fixée à 400 tours par minute. Ce paramètre a été graduellement élevé pour maintenir une pression partielle d'O₂ (pO₂) au-dessus de 40% de la saturation en oxygène. Lorsque le glucose initial était proche de l'épuisement
30 (<20% de la teneur initiale en glucose), l'étape dite de "fed-batch" a été initiée.

Au cours de cette deuxième étape, une solution de source de carbone inductrice de la production d'enzymes (lactose) a été injectée à un débit de 0,98 g.h⁻¹ pour initier la production des enzymes cellulolytiques.

Prélèvement des échantillons et traitement des échantillons

Des échantillons ont été prélevés périodiquement pour déterminer la production
5 de biomasse, la consommation de glucose et la production de protéines extracellulaires.

Pour l'isolement de l'ARN, des échantillons ont été prélevés avant l'injection de lactose et respectivement 1h, 3h, 6h et 24h après l'injection de lactose.

Les concentrations de biomasse ont été déterminées par analyse gravimétrique.
10 Dix millilitres de milieu de culture contenant du mycélium recueilli a été filtré à l'aide de membrane en microfibres de verre GF/C ayant des pores à 1,2 μm . La biomasse sèche a été déterminée 24h après l'incubation de la membrane à 105 ° C. La concentration de glucose au cours de la phase dite de "batch" a été évaluée par la réaction enzymatique à l'aide d'un analyseur de glucose Analox GM10 (Imlab)
15 pour déterminer le moment opportun pour initier l'étape de fed-batch.

Toutes les concentrations des sources de carbone ont été quantifiées *a posteriori* par HPLC à l'aide une colonne HPX- 87p (Biorad) maintenue à 85°C. De l'eau distillée dégazée a été utilisée comme éluant à un débit de 0,4 mL.L⁻¹.

Les concentrations de protéines extracellulaires ont été quantifiées par le kit de
20 dosage de protéine Bradford (Bio-Rad) avec de l'albumine de sérum bovin (BSA) comme étalon (0 à 2,0 g L⁻¹ de gamme avec une régression du deuxième ordre).

Exemple 2 : Isolement et purification des ARNs et hybridation aux puces à ADN

25 Pour l'isolement des ARNs, des échantillons de mycélium ont été filtrés, lavés avec de l'eau froide distillée, immédiatement congelés dans de l'azote liquide et broyés jusqu'à l'obtention d'une poudre de mycélium.

La poudre a été soumise à un traitement par le phénol en utilisant la solution de
30 réactif TRI (Applied Biosystems). L'ARN total extrait a ensuite été isolé à l'aide du bromochloropropane (Sigma), précipité avec de l'isopropanol (Sigma), lavé avec de l'éthanol, et solubilisé dans de l'eau exempte de nucléase (Qiagen) selon les instructions du fabricant. Les échantillons ont été lavés en suivant la procédure

RNeasy de Qiagen et ont été soumis à une digestion DNase sur colonne à l'aide du kit RNase-Free DNase set (Qiagen).

Les ARNs de chaque expérience ont fait l'objet d'une transcription inverse et ont été marqués à l'aide des fluorochromes Cy3 ou Cy5 (Invitrogen) en utilisant la

5 procédure de marquage indirect.

Ensuite, 1 µg d'ADNc marqué a été hybridé à l'aide de la puce à ADN de *T. reesei* 8x60k fabriqué par Agilent et conçue en utilisant le logiciel Teloenn comme décrit notamment par Jourdren et al., 2010.

10 **Exemple 3 : Traitement des données obtenues des puces à ADN et analyse l'analyse de l'expression des gènes des cultures induites**

Les essais ont été analysés à l'aide d'un scanner Agilent G2565CA et les signaux analysés par le logiciel GenePix Pro 6.1.

15 Les données brutes d'intensités mesurées sur les puces à ADN sont normalisées par la méthode Lowess globale (locally weighted polynomial regression) au moyen du programme Goulphar [Lemoine et al, 2006].

L'analyse de l'expression des gènes a été réalisée au moyen de scripts spécifiquement développés pour le traitement des fichiers de données

20 normalisées. Pour chaque condition expérimentale, les spots présentant une intensité de fluorescence inférieure à la moyenne ajoutée de deux écarts types du bruit de fond sur l'ensemble de la puce ont été considérés comme non détectés et éliminés de l'analyse.

Pour chaque sonde le rapport log2 d'hybridation a été associé à l'annotation du

25 génome selon le génome *T. reesei* (v2.0) (<http://genome.jgi-psf.org/Trire2/Trire2.home.html>). Le rapport log2 final pour chaque transcrit a été obtenu en moyennant les valeurs d'hybridation détectées de l'ensemble des sondes positionnées sur le brin codant et à l'intérieur de la séquence codante correspondante. Les transcrits ne possédant aucune ou une seule sonde considérée
30 comme détectée ont été éliminés. Les résultats issus des réplicats biologiques et techniques ont été traités parallèlement et indépendamment.

À l'issue de la procédure, l'analyse différentielle a été effectuée à l'aide d'une modélisation linéaire (lmFit) et de tests bayésiens (eBayes), implémentés dans le

package R limma [Smyth, GK et al, 2004] pour chaque condition. Les p-values ont été ajustées par la méthode de correction des tests multiples de Benjamini-Hochberg, et considérés comme significatives quand elles sont inférieures à 0,05. Enfin, seuls les transcrits présentant un rapport log2 d'hybridation supérieur à 1 ou inférieur à -1, ont été conservés comme les cibles les plus fortement régules.

Une matrice d'expression regroupant les rapports log2 d'hybridation de chaque gène pour chacune des conditions expérimentales testées est réalisée. Dans cette matrice, seuls les gènes qui étaient différentiellement exprimés dans au moins une condition expérimentale testée ont été retenus.

La matrice d'expression finale obtenue par les inventeurs contenait alors 592 gènes exprimés de manière différentielle pendant l'induction par le lactose de la production d'enzyme cellulolytique.

Une liste de 16 gènes cibles a alors été établie après une analyse détaillée sélectionnant les gènes présentant l'expression différentielle la plus élevée et susceptibles d'avoir une fonction dans la production d'enzymes cellulolytiques.

Parmi cette liste, les inventeurs ont sélectionné le gène TRIRE120698 (orthologue de RIM101 de *S. cerevisiae* et de PacC de *A. niger*) pour des analyses supplémentaires.

Afin d'évaluer l'effet de ce gène sur la production de protéines, et plus particulièrement la production d'enzymes cellulolytiques telles que les cellulases et les hémicellulases, une cassette de délétion a été construite de manière à supprimer *in vivo* le gène de PacC par recombinaison homologue.

Exemple 4: Construction de la cassette de délétion de gène pour PacC

La cassette de délétion est composée de trois fragments d'ADN :

- une région en amont du gène cible, d'environ 1 kb
- le gène orotidine -5'-phosphate décarboxylase de *T. reesei* (pyr4, TRIRE0074020) comme marqueur de sélection, et
- la région en aval du gène cible, d'environ 1 kb.

Le plasmide pRS426 a été utilisé comme vecteur pour la construction de la cassette de délétion par recombinaison *in vivo* dans *S. cerevisiae* selon une

méthode inspirée de la littérature (Schuster, A., Bruno, K. S., Collett, J. R., Baker, S. E., Seiboth, B., Kubicek, C. P., & Schmoll, M. (2012). *A versatile toolkit for high throughput functional genomics with Trichoderma reesei*. Biotechnology for Biofuels, 5(1), 1. doi:10.1186/1754-6834-5-1). Pour ce faire, elle a été digérée
5 avec EcoRI et XhoI (New England Biolabs) et purifiée par électrophorèse sur gel en utilisant QIAquick Gel Extraction Kit (Qiagen). Ce plasmide utilisé pour la modification est représentée en figure 1.

Amplification de la région en amont du gène cible chez *T. reesei* QM9414

10 Les amorces utilisées pour l'amplification de la région en amont du gène cible à partir de *T. reesei* QM9414 sont décrits ci-après.

L'amorce sens est constituée d'une partie générale de 29 nucléotides (nt) homologues de la séquence du vecteur pRS426 à proximité du site de restriction XhoI (GTAACGCCAGGGTTTTCCCAGTCACGACG) (SEQ ID NO : 4).

15

La partie générale de l'oligonucléotide a été suivie de la séquence spécifique correspondant à 20 nt de la région amont du gène de PacC.

L'amorce antisens contient une partie générale constituée de 21 nucléotides
20 homologues à l'extrémité 3' du gène du marqueur sélectif pyr4 (CGACGATATCAGCTTCCATAT) (SEQ ID NO: 5) et de 8 nucléotides du site de restriction MmeI (TCCGACTA) (SEQ ID NO: 6), le tout formant la séquence CGACGATATCAGCTTCCATATTCCGACTA (SEQ ID NO: 7).

25 La partie générale de l'oligonucléotide a été suivie de la séquence spécifique correspondant à 20 nucléotides de la région amont du gène de PacC.

Les parties de l'amorce spécifiques du gène cible ont été conçues sur la base de la prédiction ORF dans la version v2.0 du génome (<http://genome.jgi-psf.org/Trire2/Trire2.home.html>).
30

Les amorces utilisées pour l'amplification de la région en amont du gène cible à partir de l'ADN *T. reesei* QM9414 comme matrice ont donc été les suivantes :

- 120698 amont-amorce sens: SEQ ID NO: 8

GTAACGCCAGGGTTTTCCCAGTCACGACGGAAGAACTAGAAGACGGAGC

- 120698 amont-amorce antisens: SEQ ID NO: 9

5 CGACGATATCAGCTTCCATATTCCGACTAATGAGGAAGACTGTAGGAGG

Ces amorces permettent d'amplifier un fragment de zone intergénique de taille adéquate pour la recombinaison.

10 Amplification de la région en aval du gène cible PacC

Les amorces utilisées pour l'amplification de la région en aval du gène cible à partir de *T. reesei* QM9414 sont décrits ci-après.

- 15 L'amorce sens contient une partie générale constituée de 21 nucléotides homologues à l'extrémité 5' du gène du marqueur sélectif pyr4 (AGAAAAGCACAAAGAAGAGGC) (SEQ ID NO: 10) et de 8 nucléotides du site de restriction MmeI (TCCAATA) (SEQ ID NO: 11), le tout formant ainsi la séquence AGAAAAGCACAAAGAAGAGGCTCCAATA (SEQ ID NO: 12).

20

La partie générale de l'oligonucléotide a été suivie de la séquence génique spécifique correspondant à 20 nt de la région en aval du gène de PacC.

- 25 L'amorce antisens contient une séquence de 29 nucléotides homologues de la séquence de vecteur pRS426 à proximité du site de restriction EcoRI (GCGGATAACAATTTACACAGGAAACAGC) (SEQ ID NO: 13).

La partie générale de l'oligonucléotide a été suivie de la séquence génique spécifique correspondant à 20 nucléotides de la région en aval du gène de PacC.

30

Les parties de l'amorce spécifiques du gène cible ont été conçues sur la base de la prédiction ORF dans la version v2.0 du génome (<http://genome.jgi-psf.org/Trire2/Trire2.home.html>).

Les amorces utilisées pour l'amplification de la région en aval du gène cible à partir de l'ADN de *T. reesei* QM9414 comme matrice ont été les suivantes :

➤ 120698 aval-amorce sens : SEQ ID NO: 14

5 5'AGAAAAGCACAAAGAAGAGGCTCCAACTAACGACTGACTCTCAGTGACG

➤ 120698 aval-amorce antisens: SEQ ID NO: 15

5'GCGGATAACAATTTACACAGGAAACAGCACACTTGTCTGCTTGTCTCG

10 Ces amorces permettent d'amplifier un fragment de zone intergénique de taille adéquate pour la recombinaison.

Amorces utilisées pour l'amplification du gène de marqueur sélectif

15 Les amorces utilisées pour l'amplification du gène de pyr4 à partir de *T. reesei* QM9414 ont été les suivantes :

L'amorce sens contient une partie générale constituée de 8 nucléotides du site de restriction MmeI (TAGTTGGA) (SEQ ID NO : 16) et 21 nucléotides de la
20 séquence du gène spécifique de l'extrémité 5 ' de TRIRE0074020 (pyr4) (GCCTCTTCTTTGTGCTTTTCT), (SEQ ID NO: 17), le tout formant ainsi la séquence TAGTTGGAGCCTCTTCTTTGTGCTTTTCT (SEQ ID NO: 18).

L'amorce antisens contient une partie générale constituée de 8 nucléotides du site
25 de restriction MmeI (TAGTCGGA) (SEQ ID NO : 19) et 21 nucléotides de la séquence du gène spécifique de l'extrémité 3' de pyr4 (ATATGGAAGCTGATATCGTCG), (SEQ ID NO: 20), le tout formant la séquence TAGTCGGAATATGGAAGCTGATATCGTCG (SEQ ID NO: 21).

30 Les parties de l'amorce spécifiques du gène cible ont été conçues sur la base de la prédiction ORF dans la version v2.0 du génome (<http://genome.jgi-psf.org/Trire2/Trire2.home.html>).

Les amorces utilisées pour l'amplification de la région en aval du gène cible à partir de l'ADN de *T. reesei* QM9414 comme matrice ont été les suivantes :

- pyr4 amorce sens : SEQ ID NO: 22

5 5' – TAGTTGGAGCCTCTTCTTTGTGCTTTTCT

- pyr4 amorce antisens : SEQ ID NO: 23

5' – TAGTCGGAATATGGAAGCTGATATCGTCG

10 Les PCR mises en œuvre à l'aide des amorces ci-dessus ont donné lieu à des fragments d'ADN à extrémités homologues. Ces fragments d'ADN ainsi que le plasmide pRS426 digéré ont été transformés dans une souche *Saccharomyces cerevisiae* W303 compétente.

15 **Exemple 5: Construction d'une souche QM9414 de *T. reesei* délétée du gène PacC**

Les plasmides pRS426 contenant la cassette de délétion du gène de PacC ont été extraits de *S. cerevisiae* et utilisés comme matrice pour amplifier par PCR la
20 cassette de délétion en utilisant les amorces 120698 amont-amorce sens (SEQ ID NO : 8) et 120 698 en aval-amorce antisens (SEQ ID NO: 15).

Le produit de la PCR a ensuite été transformé dans une souche *T. reesei* QM9414 Δtku70 - pyr4 (Schuster, A., Bruno, K. S., Collett, J. R., Baker, S. E., Seiboth, B.,
25 Kubicek, C. P., & Schmoll, M. (2012). A versatile toolkit for high throughput functional genomics with *Trichoderma reesei*. *Biotechnology for Biofuels*, 5(1), 1. doi:10.1186/1754-6834-5-1).

Les transformants ont été sélectionnés sur la base de la fonction de gène marqueur
30 de sélection pyr4, présent dans la cassette de délétion. Les mutants ont été purifiés à partir des colonies issues de spores individuelles et, au final, six transformants ont été retenus.

L'intégration de la cassette a été confirmée par l'absence d'une amplification par PCR du gène cible et par l'amplification par PCR des séquences flanquantes du site d'insertion. La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une
5 amorce qui se lie au gène marqueur. Pour la séquence en aval du site d'insertion, une amorce qui se lie au gène marqueur et une amorce qui se lie en aval du site d'insertion ont été utilisés.

Les amorces utilisées pour l'amplification ;

- 10 - du gène cible,
 - du site d'insertion en région en amont dudit gène cible; et
 - du site d'insertion de la région en aval du gène cible,

à l'aide de l'ADN génomique de *T. reesei* à partir des mutants purifiés comme matrice ont été les suivantes :

15

- Gène PacC – amorce sens : SEQ ID NO: 24

5' – CCCAGCCCCGAGCTATTATG

- Gène PacC – amorce antisens : SEQ ID NO: 25

20

5' – GACCGGTTGACTGAGGGAAG

- Région en amont du gène PacC - amorce sens : SEQ ID NO: 26

5' – ATTCGGCTGCTGGAGATATG

25

- Marqueur de sélection – antisens : SEQ ID NO: 27

5' – CCTCTTTTCCATCTTGTC

- Marqueur de sélection – sens : SEQ ID NO: 28

5' – ATCCGAGCGTGAGCATCAAA

30

- Région en aval du gène PacC – amorce antisens : SEQ ID NO: 29

5' – AATGGCGGGCTTTATTTGCG

Exemple 6: Culture de la souche de *T. reesei* QM9414 dépourvue du gène PacC et analyse des cultures pour la production de protéines et la croissance.

Les spores provenant de six mutants qui présentent une délétion du gène PacC et la souche de *T. reesei* QM9414 Δ tku70 - pyr4 non modifiée ont été utilisés pour ensemencer une plaque de culture à 6 puits contenant 5 ml de milieu de culture par puits.

Le milieu était composé de K₂HPO₄ 8,7 g.L⁻¹ ; (NH₄)₂SO₄ 4,2 g.L⁻¹ ; MgSO₄.7H₂O 0,3 g.L⁻¹ ; cornsteep 1,5 g.L⁻¹ ; Lactose 10 g.L⁻¹ ; Cellulose 10 g.L⁻¹ ; acide maléique 11,6 g.L⁻¹ ; CaCl₂ 0,3 g.L⁻¹ ; FeSO₄.7H₂O 5,0 mg.L⁻¹ ; MnSO₄.H₂O 1,6 mg.L⁻¹ ; ZnSO₄.7H₂O 1,4 mg.L⁻¹ ; CoCl₂.6H₂O 2,0 mg.L⁻¹ ; pH 6 .

La culture a été effectuée à 30°C sous agitation à 150 rpm, en double exemplaire. Après 7 jours de culture, le surnageant a été recueilli pour mesurer la concentration en protéines dans le milieu. La production de protéines extracellulaires des cultures est présentée dans la Fig. 2. Cette figure montre une augmentation de la production de protéines pour les six mutants par rapport à la souche de *T. reesei* non modifiée.

Une culture de trois transformants et la souche *T. reesei* QM9414 Δ tku70 - pyr4 non modifiés a été réalisée en fioles de 250 ml contenant 50 ml du même milieu de culture pour suivre la production de protéines extracellulaires obtenir une mesure de la vitesse spécifique de production de la protéine (production de protéine par mg de biomasse fongique par unité de temps, ou qP) à la fin de la culture.

Ces valeurs sont présentées en Fig. 3.

Les données montrent des valeurs qP améliorées chez les mutants par rapport à la souche de *T. reesei* non modifiée.

Exemple 7 : construction d'une cassette de délétion pour le gène PacC pour une souche hyper-productrice et construction de la souche.

Les souches hyper-productrices telles que RUT C30 et CL847 (Durand, H., Clanet, M., & Tiraby, G. (1988). *Genetic Improvement of Trichoderma reesei for large scale cellulase production. Enzyme and Microbial Technology*, 10, 341–346.) ne sont pas auxotrophes pour l'uracile.

5

Un autre marqueur de sélection doit être utilisé.

La cassette de délétion est composée de trois fragments d'ADN :

- une région en amont du gène cible, d'environ 1 kb ;
- 10 - une cassette HygroR contenant le gène Hph (hygromycine B phosphotransférase) responsable de la résistance à l'hygromycine B isolé d'*Escherichia coli* placé sous le contrôle du promoteur du gène *cpc-1* de *Neurospora crassa* (NCU04050) et le terminateur du gène *trpC* d'*Aspergillus nidulans* ANID_00648) ; et
- 15 - la région en aval du gène cible, d'environ 1 kb.

Les séquences suivantes ont été amplifiées indépendamment à l'aide d'amorces ayant des régions communes avec les fragments à fusionner :

- la région 5' du gène PacC (SEQ ID N°55) à l'aide des amorces sens SEQ ID N°58 et antisens SEQ ID N°59 ;
- 20 - de la région 3' du gène PacC (SEQ ID N°56) à l'aide des amorces sens SEQ ID N°60 et antisens SEQ ID N°61 ;
- la cassette hygroR (SEQ ID N°57) à l'aide des amorces sens SEQ ID N°62 et antisens SEQ ID N°63.

25

Ainsi, la SEQ ID N°55 possède en amont 20 nt en commun avec le plasmide pUC19 et 20 nt en aval commun avec la cassette hygroR, la SEQ ID N°57 possède en amont 20nt en commun avec la région 5' en amont de RIM101 et en aval 20nt en commun avec la région 3' en aval du gène RIM101, la SEQ ID N°56 possède en amont 20 nt en commun avec la cassette hygroR et en aval 20 nt en commun avec le plasmide pUC19.

30

Les fragments ont ensuite été assemblés à l'aide du kit d'assemblage Gibson Assembly® Master Mix (NEB) dans le plasmide pUC19 et amplifiés dans *E. Coli*.

- 5 Le fragment contenant les régions flanquantes 3' et 5' du gène RIM101 avec au milieu la cassette hygR a été amplifié à l'aide des amorces sens SEQ ID N°64 et antisens SEQ ID N°65.

- 10 Les transformants ont été sélectionnés sur la base de la fonction de gène marqueur de sélection pyr4, présent dans la cassette de délétion. Les mutants ont été purifiés à partir des colonies issues de spores individuelles et, au final, six transformants ont été retenus.

- 15 L'intégration de la cassette a été confirmée par l'absence d'une amplification par PCR du gène cible et par l'amplification par PCR des séquences flanquantes du site d'insertion. La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie au gène marqueur. Pour la séquence en aval du site d'insertion, une amorce qui se lie au gène marqueur et une amorce qui se lie en aval du site
20 d'insertion ont été utilisées.

- Un clone de la souche CL847 a ensuite été transformé avec 5µg du fragment obtenu selon la méthode connue de l'homme de l'art (Penttilä, M., Nevalainen, H., Rättö, M., Salminen, E., & Knowles, J. (1987). A versatile transformation
25 system for the cellulolytic filamentous fungus *Trichoderma reesei*. *Gene*, 61, 155–164.).

- Les transformants ont été sélectionnés sur la base de la fonction du marqueur de sélection Hph (hygromycine B phosphotransférase) responsable de la résistance à
30 l'hygromycine B placé sous le contrôle du promoteur du gène *cpc-1* de *Neurospora crassa* (NCU04050) et le terminateur du gène *trpC* d'*Aspergillus nidulans* ANID_00648), par sélection en milieu riche PDA (Difco) additionné de 50 µg d'hygromycine. Les mutants résistants à l'hygromycine ont été purifiés à

partir des colonies issues de spores individuelles. Un transformant a été sélectionné : CL847-Drim101-C8.

Le remplacement du gène PacC par la cassette HygroR dans le transformant a été

5 vérifié par :

- par l'absence d'une amplification par PCR du gène cible à l'aide des amorces sens SEQ ID N°66 et antisens SEQ ID N°67 ;

- par l'amplification par PCR des séquences flanquantes du site d'insertion. La séquence adjacente en amont du site d'insertion a été
10 amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie au gène marqueur (amorces sens SEQ ID N°68 et antisens SEQ ID N°69). Pour la séquence en aval du site d'insertion, une amorce qui se lie au gène marqueur et une amorce qui se lie en aval du site d'insertion ont été utilisés (amorces sens SEQ
15 ID N°70 et antisens SEQ ID N°71).

Exemple 8 : Culture de la souche de *T. reesei* CL847 dépourvue du gène PacC et analyse des cultures pour la production de protéines

20 Les spores provenant du mutant qui présentant une délétion du gène PacC et la souche d'origine *T. reesei* CL847 non modifiée ont été utilisés pour ensemercer une plaque de culture à 24 puits contenant 2 ml de milieu de culture par puits.

Le milieu était composé de K₂HPO₄ 8,7 g.L⁻¹ ; (NH₄)₂SO₄ 4,2 g.L⁻¹ ;
25 MgSO₄.7H₂O 0,3 g.L⁻¹ ; cornsteep 1,5 g.L⁻¹ ; Lactose 10 g.L⁻¹ ; Cellulose 10 g.L⁻¹ ; acide maléique 11,6 g.L⁻¹ ; CaCl₂ 0,3 g.L⁻¹ ; FeSO₄.7H₂O 5,0 mg.L⁻¹ ; MnSO₄.H₂O 1,6 mg.L⁻¹ ; ZnSO₄.7H₂O 1,4 mg.L⁻¹ ; CoCl₂.6H₂O 2,0 mg.L⁻¹ ; pH 6 .

La culture a été effectuée à 30°C sous agitation à 150 rpm, en double exemplaire.

30 Après 7 jours de culture, le surnageant a été recueilli pour mesurer la concentration en protéines dans le milieu. La production de protéines extracellulaires des cultures est présentée dans la Figure. 4. Cette figure montre une augmentation de la production de protéines pour le mutant (CL847-Drim101-C8) par rapport à la souche de *T. reesei* non modifiée (CL847).

Exemple 9 : Construction d'une cassette de délétion des positions 1 à 735 de SEQ ID N°1

5

Une séquence nucléotide synthétique été dessinée, synthétisée puis a été introduite dans le plasmide pUC19. Cette séquence (SEQ ID N°30) contient :

- la séquence de SEQ ID N°1 avec une délétion des 735 premières bases ;
- 10 - une région en amont du gène cible d'environ 1 kb.

Le plasmide pUC19-RUC19-RIM101-735-4152 contenant la séquence synthétique a été utilisé comme matrice pour amplifier la SEQ ID N°30 par PCR en utilisant les amorces sens 31-R (SEQ ID N° 31) et antisens 30-F (SEQ ID N°
15 32).

La cassette de sélection *PYR4* contenant le gène *PYR4* avec environ 1,2kb en amont du gène et 0.6kb en aval du gène (SEQ ID N° 33) a été amplifiée à partir de l'ADN génomique de la souche Cl847 (Durand, H., M. Clanet, and G. Tiraby.
20 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enzyme Microb. Technol. 10:341–346.) en utilisant les amorces sens 34-R (SEQ ID N° 34) et antisens 35-F (SEQ ID N° 35).

Les produits de PCR correspondant à SEQ ID N°30 et SEQ ID N° 33 ont ensuite
25 été co-transformés dans une souche *T. reesei* QM9414 Δ tku70 - pyr4 (Guangtao, Z., Hartl, L., Schuster, A., Polak, S., Schmoll, M., Wang, T., Seidl, V., et al. (2009). Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. Journal of biotechnology, 139(2), 146-51. doi:10.1016/j.jbiotec.2008.10.007).

30

Les transformants ont été sélectionnés sur milieu minimum sans uridine. Les mutants ont été purifiés à partir des colonies issues de spores individuelles.

L'intégration de la cassette a été confirmée :

- par l'absence d'une amplification de la région de 735 bp situé entre les positions 1 et 735 de SEQ ID N°1 : amorce sens- SEQ ID N°36 et amorce antisens SEQ ID N° 37 ;
- par l'amplification par PCR des séquences flanquantes de la cassette d'insertion (SEQ ID N°30). La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie en aval du site d'insertion. Le fragment amplifié présente une taille inférieure de 735 paires de base par rapport celui de la souche d'origine (amorce sens- SEQ ID N° 38 et amorce antisens SEQ ID N° 39).

Exemple 10 : Construction d'une cassette de délétion des positions 249 à 495 de SEQ ID N°1

- 15 Une séquence nucléotide synthétique été dessinée, synthétisée puis a été introduite dans le plasmide pUC19. Cette séquence (SEQ ID N°40) contient :
- la séquence de SEQ ID N°1 avec une délétion des positions 249 à 495 nucléotides ;
 - une région en amont et en val de la délétion d'environ 1 kb.

20

Le plasmide pUC19-RUC19-RIM101-del249-495 contenant la séquence synthétique a été utilisé comme matrice pour amplifier la SEQ ID N°40 par PCR en utilisant les amorces sens 31-R (SEQ ID N° 31) et antisens 32-F (SEQ ID N° 32).

25

- La cassette de sélection PYR4 contenant le gène *PYR4* avec environ 1,2kb en amont du gène et 0.6kb en aval du gène (SEQ ID N° 33) a été amplifiée à partir de l'ADN génomique de la souche Cl847 (Durand, H., M. Clanet, and G. Tiraby. 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. *Enzyme Microb. Technol.* 10:341–346.) en utilisant les amorces sens 34-R (SEQ ID N° 34) et antisens 35-F (SEQ ID N° 35).

30

Les produits de PCR correspondant à SEQ ID N°40 et SEQ ID N°33 ont ensuite été co-transformés dans une souche *T. reesei* QM9414 Δ tku70 - pyr4 (Guangtao,

Z., Hartl, L., Schuster, A., Polak, S., Schmoll, M., Wang, T., Seidl, V., et al. (2009). Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. Journal of biotechnology, 139(2), 146-51. doi:10.1016/j.jbiotec.2008.10.007).

5

Les transformants ont été sélectionnés sur milieu minimum sans uridine. Les mutants ont été purifiés à partir des colonies issues de spores individuelles.

L'intégration de la cassette a été confirmée :

- 10 - par l'absence d'une amplification de la région de 246 bp situé entre les positions 249 et 495 de SEQ ID N°1 : amorce sens- SEQ ID N°40 et amorce antisens SEQ ID N° 41 ; et
- 15 - par l'amplification par PCR des séquences flanquantes de la cassette d'insertion (SEQ ID N°30). La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie en aval du site d'insertion. Le fragment amplifié présente une taille inférieure de 246 paires de base par rapport celui de la souche d'origine (amorce sens- SEQ ID N° 38 et amorce antisens SEQ ID N° 39).

20

Exemple 11 : Construction d'une cassette de délétion des positions 736 à 744 de SEQ ID N°1

25 Une séquence nucléotide synthétique été dessinée, synthétisée puis a été introduite dans le plasmide pUC19. Cette séquence (SEQ ID N°43) contient :

- la séquence de SEQ ID N°1 avec une délétion des positions 736 à 744 nucléotides ; et
- une région en amont et en aval de la délétion d'environ 1 kb.

30 Le plasmide pUC19-RUC19-RIM101-del736-744 contenant la séquence synthétique a été utilisé comme matrice pour amplifier la SEQ ID N°43 par PCR en utilisant les amorces sens 31-R (SEQ ID N° 31) et antisens 32-F (SEQ ID N° 32).

La cassette de sélection *PYR4* contenant le gène *PYR4* avec environ 1,2kb en amont du gène et 0.6kb en aval du gène (SEQ ID N°33) a été amplifiée à partir de l'ADN génomique de la souche C1847 (Durand, H., M. Clanet, and G. Tiraby. 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. *Enzyme Microb. Technol.* 10:341–346.) en utilisant les amorces sens 34-R (SEQ ID N° 34) et antisens 35-F (SEQ ID N° 35).

Les produits de PCR correspondant à SEQ ID N°43 et SEQ ID N°33 ont ensuite été co-transformés dans une souche *T. reesei* QM9414 Δ tku70 - pyr4 (Guangtao, Z., Hartl, L., Schuster, A., Polak, S., Schmoll, M., Wang, T., Seidl, V., et al. (2009). Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. *Journal of biotechnology*, 139(2), 146-51. doi:10.1016/j.jbiotec.2008.10.007).

Les transformants ont été sélectionnés sur milieu minimum sans uridine. Les mutants ont été purifiés à partir des colonies issues de spores individuelles.

L'intégration de la cassette a été confirmée :

- par l'absence d'une amplification de la région de contenant la délétion 736-744 en utilisant une amorce localisée sur le site de la délétion : amorce sens- SEQ ID N°44 et amorce antisens SEQ ID N° 46 (amorce sur le site de la délétion) ;
- par l'amplification par PCR des séquences flanquantes de la cassette d'insertion (SEQ ID N°43). La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie en aval du site d'insertion (amorce sens- SEQ ID N° 38 et amorce antisens SEQ ID N° 39). La présence de la délétion 736-744 est vérifiée par séquençage en utilisant les amorces SEQ ID N° 44 et amorce antisens SEQ ID N° 45.

Exemple 12 : Construction d'une cassette de délétion des positions 1438 à 1515 de SEQ ID N°1

- 5 Une séquence nucléotide synthétique été dessinée, synthétisée puis a été introduite dans le plasmide pUC19. Cette séquence (SEQ ID N°47) contient :
- la séquence de SEQ ID N°1 avec une délétion des positions 1438 à 1515 nucléotides ; et
 - 10 - une région en amont et en val de la délétion d'environ 1 kb.

Le plasmide pUC19-RUC19-RIM101-del-1438-1515 contenant la séquence synthétique a été utilisé comme matrice pour amplifier la SEQ ID N°47 par PCR en utilisant les amorces sens 48-R (SEQ ID N° 48) et antisens 49-F (SEQ ID N° 49).

15 49).

La cassette de sélection *PYR4* contenant le gène *PYR4* avec environ 1,2kb en amont du gène et 0.6kb en aval du gène (SEQ ID 33) a été amplifiée à partir de l'ADN génomique de la souche Cl847 (Durand, H., M. Clanet, and G. Tiraby. 20 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. *Enzyme Microb. Technol.* 10:341–346.) en utilisant les amorces sens 34-R (SEQ ID N° 34) et antisens 35-F (SEQ ID N° 35).

Les produits de PCR correspondant à SEQ ID N°47 et SEQ ID N°33 ont ensuite été co-transformés dans une souche *T. reesei* QM9414 Δ tku70 - pyr4 (Guangtao, Z., Hartl, L., Schuster, A., Polak, S., Schmoll, M., Wang, T., Seidl, V., et al. (2009). Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. *Journal of biotechnology*, 139(2), 146-51, doi:10.1016/j.jbiotec.2008.10.007).

25

30

Les transformants ont été sélectionnés sur milieu minimum sans uridine. Les mutants ont été purifiés à partir des colonies issues de spores individuelles.

L'intégration de la cassette a été confirmée :

- 5 - par l'amplification par PCR des séquences flanquantes de la cassette d'insertion (SEQ ID N°47). La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie en aval du site d'insertion (amorce sens- SEQ ID N° 31 et amorce antisens SEQ ID N° 50). La présence de la délétion 1438-1515 est vérifiée par séquençage en utilisant les amorces SEQ ID N° 51 et amorce antisens SEQ ID N° 52 ;
- 10 - par l'absence d'une amplification de la région de contenant la délétion 1438 à 1515 en utilisant une amorce localisée sur le site de la délétion : amorce sens- SEQ ID N°44 et amorce antisens SEQ ID N° 53 (dans la délétion).

15 **Exemple 13 : Construction d'une cassette de délétion des positions 1492 à 1494 de SEQ ID N°1**

Une séquence nucléotide synthétique été dessinée, synthétisée puis a été introduite dans le plasmide pUC19. Cette séquence (SEQ ID N°54) contient :

- 20 - la séquence de SEQ ID N°1 avec une délétion des positions 1492 à 1494 nucléotides, et
- une région en amont et en val de la délétion d'environ 1 kb.

25 Le plasmide pUC19-RUC19-RIM101-del-1492-1494 contenant la séquence synthétique a été utilisé comme matrice pour amplifier la SEQ ID N°54 par PCR en utilisant les amorces sens 48-R (SEQ ID N° 48) et antisens 49-F (SEQ ID N° 49).

30 La cassette de sélection PYR4 contenant le gène *PYR4* avec environ 1,2kb en amont du gène et 0.6kb en aval du gène (SEQ ID 33) a été amplifiée à partir de l'ADN génomique de la souche Cl847 (Durand, H., M. Clanet, and G. Tiraby. 1988. Genetic improvement of *Trichoderma reesei* for large scale cellulase production. Enzyme Microb. Technol. 10:341–346.) en utilisant les amorces sens 34-R (SEQ ID N° 34) et antisens 35-F (SEQ ID N° 35).

Les produits de PCR correspondant à SEQ ID N°54 et SEQ ID 33 ont ensuite été co-transformés dans une souche *T. reesei* QM9414 Δ tku70 - pyr4 (Guangtao, Z., Hartl, L., Schuster, A., Polak, S., Schmoll, M., Wang, T., Seidl, V., et al. (2009).

- 5 Gene targeting in a nonhomologous end joining deficient *Hypocrea jecorina*. Journal of biotechnology, 139(2), 146-51. doi:10.1016/j.jbiotec.2008.10.007).

Les transformants ont été sélectionnés sur milieu minimum sans uridine. Les mutants ont été purifiés à partir des colonies issues de spores individuelles.

10

L'intégration de la cassette a été confirmée par l'amplification par PCR des séquences flanquantes de la cassette d'insertion (SEQ ID N°54). La séquence adjacente en amont du site d'insertion a été amplifiée par PCR en utilisant une amorce qui se lie en amont du site d'insertion et une amorce qui se lie en aval du

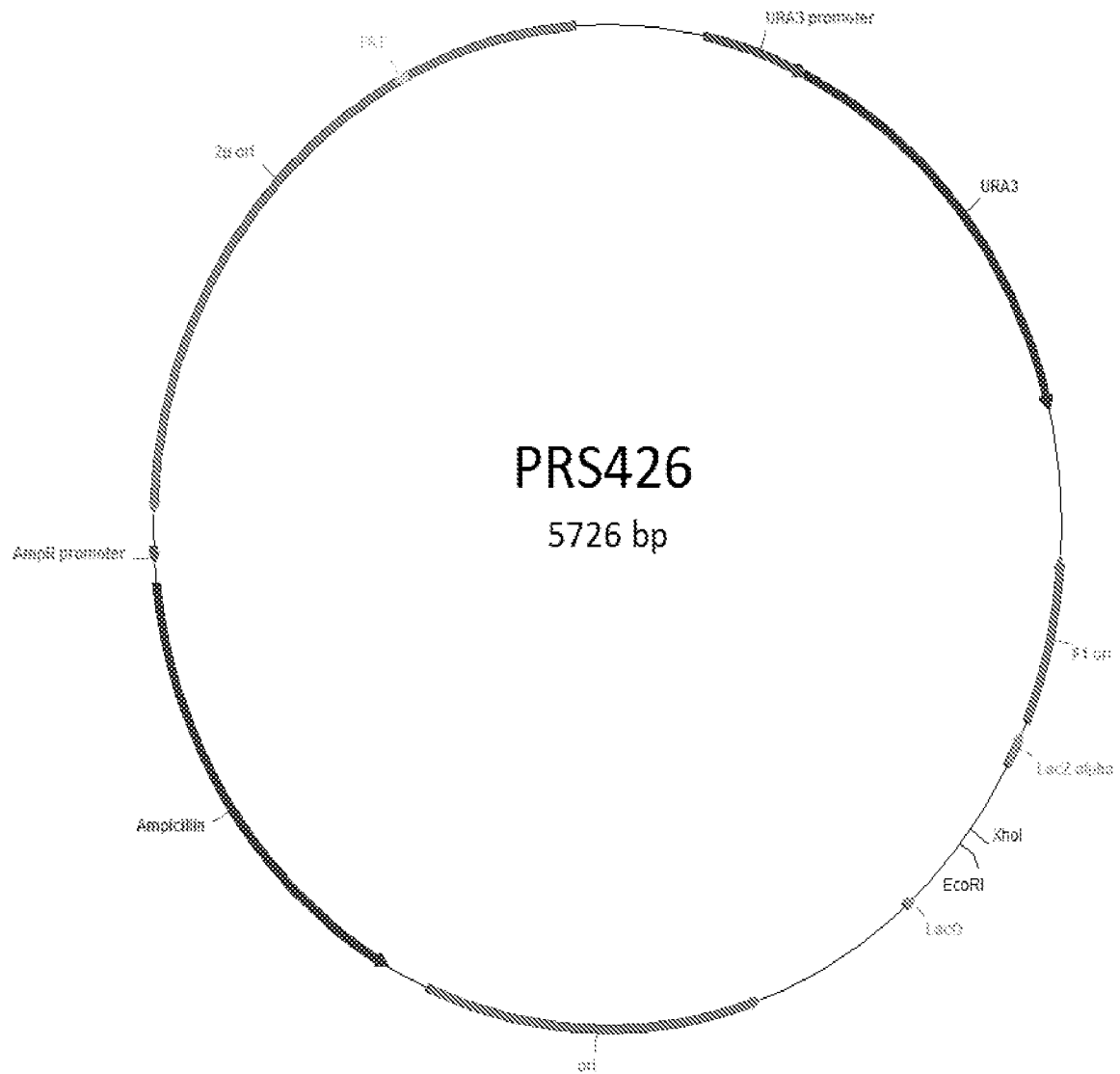
15 site d'insertion (amorce sens- SEQ ID N° 31 et amorce antisens SEQ ID N° 50).

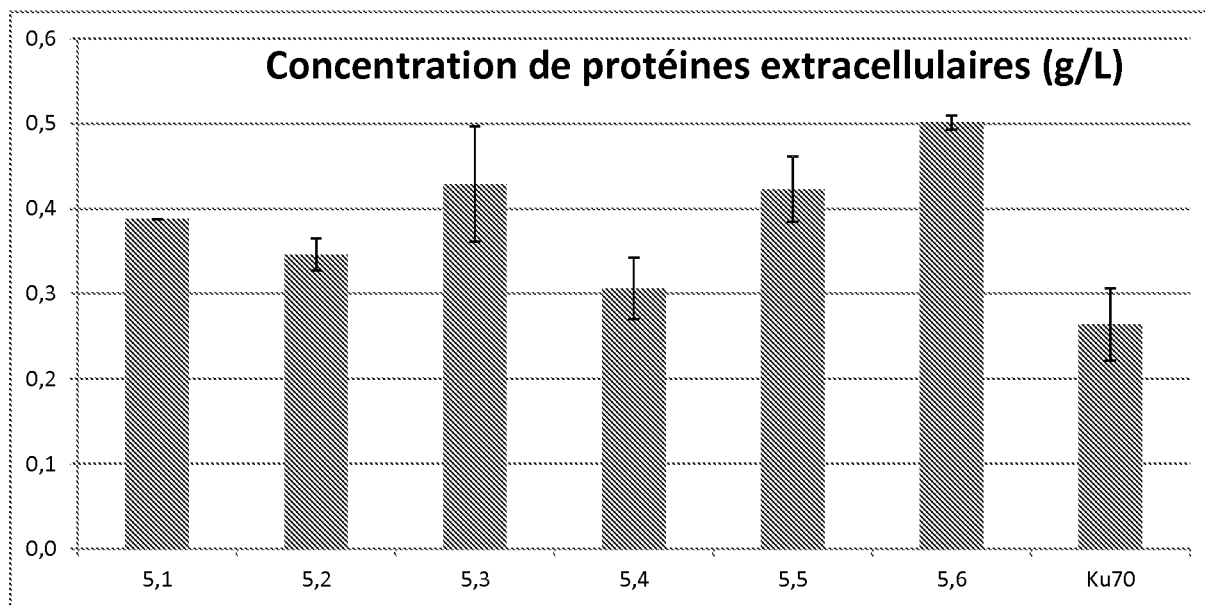
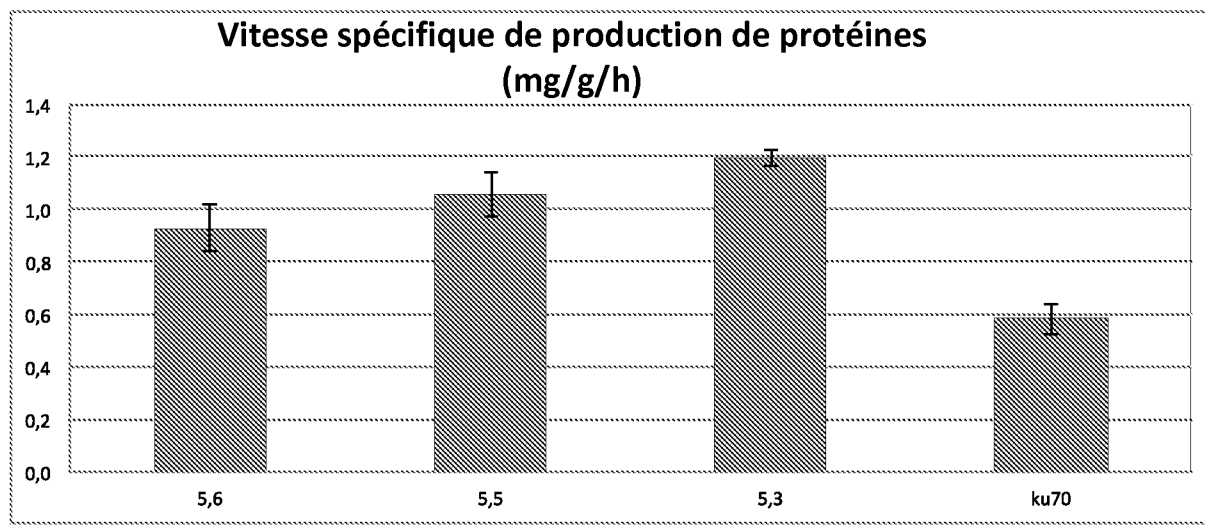
La présence de la délétion 1492-1494 est vérifiée par séquençage en utilisant les amorces SEQ ID N° 51 et amorce antisens SEQ ID N° 52.

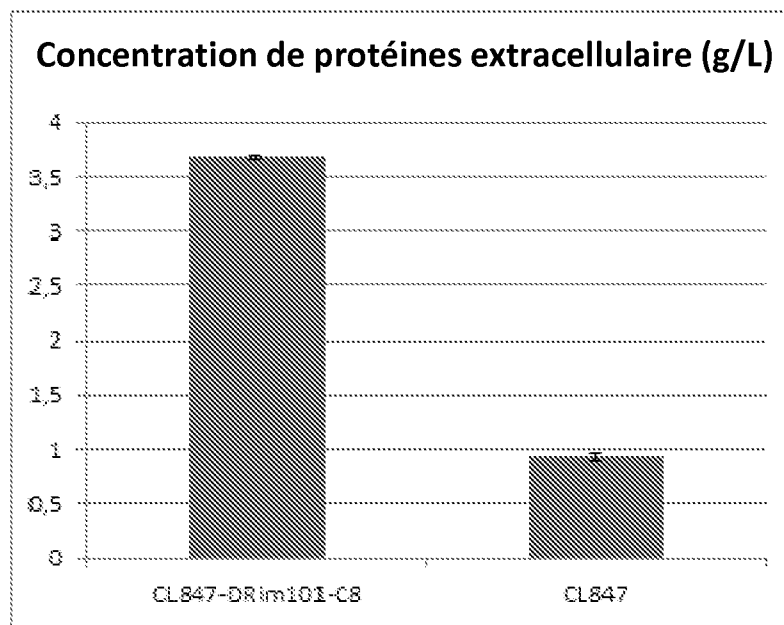
20

REVENDICATIONS

1. Souche de *Trichoderma reesei* dont le génome comprend une mutation choisie parmi:
- 5 - une délétion du gène codant pour PacC;
- une mutation entre les positions 1 à 735 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 249 à 495;
- une mutation entre les positions 736 à 744 de la séquence
- 10 nucléotidique représentée en SEQ ID N°1; et
- une mutation entre les positions 1438 à 1515 de la séquence nucléotidique représentée en SEQ ID N°1, préférentiellement entre les positions 1492 à 1494.
- 15 2. Souche selon la revendication 1, caractérisée en ce qu'elle présente une capacité de production d'enzymes améliorée par rapport à la même souche *T. reesei* non modifiée ou par rapport à une souche de référence de *T. reesei*.
- 20 3. Souche selon la revendication 1 ou 2, caractérisé en ce que lesdites enzymes sont choisies parmi les enzymes cellulolytiques, préférentiellement les cellulases et les hémicellulases, encore plus préférentiellement les cellulases.
- 25 4. Souche selon l'une des revendications 1 à 3, caractérisée en ce que ladite mutation est induite par mutagenèse dirigée, préférentiellement à l'aide d'une cassette de délétion.
5. Utilisation d'une souche modifiée selon l'une quelconque des revendications 1 à 4 pour la production d'enzymes cellulolytiques.
- 30 6. Utilisation d'une souche selon l'une quelconque des revendications 1 à 4 pour l'hydrolyse des beta-oligosaccharides.
7. Utilisation d'une souche selon l'une quelconque des revendications 1 à 4 pour l'hydrolyse du cellobiose en glucose.

FIGURES**Figure 1**

**Figure 2****Figure 3**

**Figure 4**

INTERNATIONAL SEARCH REPORT

International application No

PCT/FR2015/050629

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N15/90 C12R1/885 C12N1/14 C12P7/10
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 C12R 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, BIOSIS, CHEM ABS Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 103 060 341 A (TIANJIN INST IND BIOTECHNOLOGY) 24 April 2013 (2013-04-24)	1-4
Y	example 6; sequences 1,2	5-7
Y	----- WO 00/46375 A2 (NOVO NORDISK BIOTECH INC [US]; YAYER DEBBIE S [US] NOVO NORDISK BIOTEC) 10 August 2000 (2000-08-10) page 1 - page 4 page 8 page 20 claims 1-21	5-7
Y	----- WO 2009/098365 A2 (INST FRANCAIS DU PETROLE [FR]; MARGEOT ANTOINE [FR]; MONOT FREDERIC [F] 13 August 2009 (2009-08-13) claims 1-7; figure 1 ----- -/-	5-7



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

5 June 2015

Date of mailing of the international search report

19/06/2015

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

Deleu, Laurent

INTERNATIONAL SEARCH REPORT

International application No

PCT/FR2015/050629

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ARO N ET AL: "Transcript regulation of plant cell wall degradation by filamentous fungi", FEMS MICROBIOLOGY REVIEWS, ELSEVIER, AMSTERDAM; NL, vol. 29, no. 4, 2005, pages 719-739, XP008114070, ISSN: 0168-6445, DOI: 10.1016/J.FEMSRE.2004.11.006 [retrieved on 2004-12-22] page 729 - page 730 -----	5-7
Y	M. A. PENALVA ET AL: "Regulation of Gene Expression by Ambient pH in Filamentous Fungi and Yeasts", MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, vol. 66, no. 3, September 2002 (2002-09), pages 426-446, XP055150781, ISSN: 1092-2172, DOI: 10.1128/MMBR.66.3.426-446.2002 page 427 - page 433 -----	5-7
Y	NAOMI TRUSHINA ET AL: "PacC and pH-dependent transcriptome of the mycotrophic fungus Trichoderma vi", BMC GENOMICS, BIOMED CENTRAL LTD, LONDON, UK, vol. 14, no. 1, 28 February 2013 (2013-02-28), page 138, XP021146386, ISSN: 1471-2164, DOI: 10.1186/1471-2164-14-138 abstract figure 10; table 1 -----	5-7
A	ELIECER DÍEZ ET AL: "Activation of the Aspergillus PacC zinc finger transcription factor requires two proteolytic steps", THE EMBO JOURNAL, vol. 21, no. 6, 15 March 2002 (2002-03-15), pages 1350-1359, XP055150796, ISSN: 0261-4189, DOI: 10.1093/emboj/21.6.1350 cited in the application figures 1,8 ----- -/--	1

INTERNATIONAL SEARCH REPORT

International application No

PCT/FR2015/050629

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SAMOLSKI ILANIT ET AL: "Gene expression analysis of the biocontrol fungus <i>Trichoderma harzianum</i> in the presence of tomato plants, chitin, or glucose using a high-density oligonucleotide microarray", BMC MICROBIOLOGY, BIOMED CENTRAL LTD, GB, vol. 9, no. 1, 13 October 2009 (2009-10-13), page 217, XP021062321, ISSN: 1471-2180, DOI: 10.1186/1471-2180-9-217 page 5 -----	1-7
T	LILIAN DOS SANTOS CASTRO ET AL: "Comparative metabolism of cellulose, sophorose and glucose in <i>Trichoderma reesei</i> using high-throughput genomic and proteomic analyses", BIOTECHNOLOGY FOR BIOFUELS, BIOMED CENTRAL LTD, GB, vol. 7, no. 1, 21 March 2014 (2014-03-21), page 41, XP021181143, ISSN: 1754-6834, DOI: 10.1186/1754-6834-7-41 -----	1-7
X,P	HE RONGLIN ET AL: "Trpacl, a pH response transcription regulator, is involved in cellulase gene expression in <i>Trichoderma reesei</i> ", ENZYME AND MICROBIAL TECHNOLOGY, STONEHAM, MA, US, vol. 67, 6 September 2014 (2014-09-06), pages 17-26, XP029084273, ISSN: 0141-0229, DOI: 10.1016/J.ENZMICTEC.2014.08.013 page 20 - page 25 -----	1-5

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/FR2015/050629

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CN 103060341	A	24-04-2013	NONE

WO 0046375	A2	10-08-2000	AU 2755300 A 25-08-2000
			CN 1342204 A 27-03-2002
			EP 1151106 A2 07-11-2001
			JP 2002535990 A 29-10-2002
			WO 0046375 A2 10-08-2000

WO 2009098365	A2	13-08-2009	AT 508196 T 15-05-2011
			BR PI0819641 A2 04-11-2014
			CA 2704215 A1 13-08-2009
			CN 101868549 A 20-10-2010
			DK 2222865 T3 15-08-2011
			EP 2222865 A2 01-09-2010
			ES 2363863 T3 18-08-2011
			FR 2923840 A1 22-05-2009
			RU 2010125127 A 27-12-2011
			US 2010297717 A1 25-11-2010
			WO 2009098365 A2 13-08-2009

RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/FR2015/050629

A. CLASSEMENT DE L'OBJET DE LA DEMANDE INV. C12N15/90 C12R1/885 C12N1/14 C12P7/10 ADD.		
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB		
B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE Documentation minimale consultée (système de classification suivi des symboles de classement) C12N C12R C12P		
Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche		
Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés) EPO-Internal, WPI Data, BIOSIS, CHEM ABS Data, Sequence Search		
C. DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	CN 103 060 341 A (TIANJIN INST IND BIOTECHNOLOGY) 24 avril 2013 (2013-04-24)	1-4
Y	exemple 6; séquences 1,2	5-7
Y	WO 00/46375 A2 (NOVO NORDISK BIOTECH INC [US]; YAVER DEBBIE S [US] NOVO NORDISK BIOTEC) 10 août 2000 (2000-08-10) page 1 - page 4 page 8 page 20 revendications 1-21	5-7
Y	WO 2009/098365 A2 (INST FRANCAIS DU PETROLE [FR]; MARGEOT ANTOINE [FR]; MONOT FREDERIC [F] 13 août 2009 (2009-08-13) revendications 1-7; figure 1	5-7
	-/-	
☒ Voir la suite du cadre C pour la fin de la liste des documents ☒ Les documents de familles de brevets sont indiqués en annexe		
* Catégories spéciales de documents cités: "A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent "E" document antérieur, mais publié à la date de dépôt international ou après cette date "L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée) "O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens "P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée "T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention "X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément "Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier "&" document qui fait partie de la même famille de brevets		
Date à laquelle la recherche internationale a été effectivement achevée 5 juin 2015		Date d'expédition du présent rapport de recherche internationale 19/06/2015
Nom et adresse postale de l'administration chargée de la recherche internationale Office Européen des Brevets, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Fonctionnaire autorisé Deleu, Laurent

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	ARO N ET AL: "Transcript regulation of plant cell wall degradation by filamentous fungi", FEMS MICROBIOLOGY REVIEWS, ELSEVIER, AMSTERDAM; NL, vol. 29, no. 4, 2005, pages 719-739, XP008114070, ISSN: 0168-6445, DOI: 10.1016/J.FEMSRE.2004.11.006 [extrait le 2004-12-22] page 729 - page 730 -----	5-7
Y	M. A. PENALVA ET AL: "Regulation of Gene Expression by Ambient pH in Filamentous Fungi and Yeasts", MICROBIOLOGY AND MOLECULAR BIOLOGY REVIEWS, vol. 66, no. 3, septembre 2002 (2002-09), pages 426-446, XP055150781, ISSN: 1092-2172, DOI: 10.1128/MMBR.66.3.426-446.2002 page 427 - page 433 -----	5-7
Y	NAOMI TRUSHINA ET AL: "PacC and pH-dependent transcriptome of the mycotrophic fungus Trichoderma vi", BMC GENOMICS, BIOMED CENTRAL LTD, LONDON, UK, vol. 14, no. 1, 28 février 2013 (2013-02-28), page 138, XP021146386, ISSN: 1471-2164, DOI: 10.1186/1471-2164-14-138 abrégé figure 10; tableau 1 -----	5-7
A	ELIECER DÍEZ ET AL: "Activation of the Aspergillus PacC zinc finger transcription factor requires two proteolytic steps", THE EMBO JOURNAL, vol. 21, no. 6, 15 mars 2002 (2002-03-15), pages 1350-1359, XP055150796, ISSN: 0261-4189, DOI: 10.1093/emboj/21.6.1350 cité dans la demande figures 1,8 ----- -/--	1

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	SAMOLSKI ILANIT ET AL: "Gene expression analysis of the biocontrol fungus <i>Trichoderma harzianum</i> in the presence of tomato plants, chitin, or glucose using a high-density oligonucleotide microarray", BMC MICROBIOLOGY, BIOMED CENTRAL LTD, GB, vol. 9, no. 1, 13 octobre 2009 (2009-10-13), page 217, XP021062321, ISSN: 1471-2180, DOI: 10.1186/1471-2180-9-217 page 5 -----	1-7
T	LILIAN DOS SANTOS CASTRO ET AL: "Comparative metabolism of cellulose, sophorose and glucose in <i>Trichoderma reesei</i> using high-throughput genomic and proteomic analyses", BIOTECHNOLOGY FOR BIOFUELS, BIOMED CENTRAL LTD, GB, vol. 7, no. 1, 21 mars 2014 (2014-03-21), page 41, XP021181143, ISSN: 1754-6834, DOI: 10.1186/1754-6834-7-41 -----	1-7
X,P	HE RONGLIN ET AL: "Trpacl, a pH response transcription regulator, is involved in cellulase gene expression in <i>Trichoderma reesei</i>", ENZYME AND MICROBIAL TECHNOLOGY, STONEHAM, MA, US, vol. 67, 6 septembre 2014 (2014-09-06), pages 17-26, XP029084273, ISSN: 0141-0229, DOI: 10.1016/J.ENZMICTEC.2014.08.013 page 20 - page 25 -----	1-5

RAPPORT DE RECHERCHE INTERNATIONALE

Renseignements relatifs aux membres de familles de brevets

Demande internationale n°

PCT/FR2015/050629

Document brevet cité au rapport de recherche		Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
CN 103060341	A	24-04-2013	AUCUN	

WO 0046375	A2	10-08-2000	AU 2755300 A	25-08-2000
			CN 1342204 A	27-03-2002
			EP 1151106 A2	07-11-2001
			JP 2002535990 A	29-10-2002
			WO 0046375 A2	10-08-2000

WO 2009098365	A2	13-08-2009	AT 508196 T	15-05-2011
			BR PI0819641 A2	04-11-2014
			CA 2704215 A1	13-08-2009
			CN 101868549 A	20-10-2010
			DK 2222865 T3	15-08-2011
			EP 2222865 A2	01-09-2010
			ES 2363863 T3	18-08-2011
			FR 2923840 A1	22-05-2009
			RU 2010125127 A	27-12-2011
			US 2010297717 A1	25-11-2010
			WO 2009098365 A2	13-08-2009
