

①⑨ RÉPUBLIQUE FRANÇAISE
INSTITUT NATIONAL
DE LA PROPRIÉTÉ INDUSTRIELLE
COURBEVOIE

①① N° de publication :

3 142 490

(à n'utiliser que pour les
commandes de reproduction)

②① N° d'enregistrement national :

22 12597

⑤① Int Cl⁸ : C 12 N 1/14 (2023.01), C 12 N 15/11

⑫

DEMANDE DE BREVET D'INVENTION

A1

②② Date de dépôt : 30.11.22.

③③ Priorité :

④③ Date de mise à la disposition du public de la
demande : 31.05.24 Bulletin 24/22.

⑤⑤ Liste des documents cités dans le rapport de
recherche préliminaire : *Se reporter à la fin du
présent fascicule*

⑥⑥ Références à d'autres documents nationaux
apparentés :

○ Demande(s) d'extension :

⑦① Demandeur(s) : IFP Energies nouvelles Etablissement
Public à Caractère Industriel et Commercial (EPIC) —
FR.

⑦② Inventeur(s) : MATHIS Hugues, BIDARD-MICHELOT
Frédérique, AOUAM Thiziri et DALLONGEVILLE Anne-
Sophie.

⑦③ Titulaire(s) : IFP Energies nouvelles Etablissement
Public à Caractère Industriel et Commercial (EPIC).

⑦④ Mandataire(s) : Plasseraud IP.

⑤④ Procédé de translocation chromosomique chez un organisme eucaryote.

⑤⑦ La présente invention concerne un procédé de trans-
location chromosomique entre au moins deux régions cibles
chez un organisme eucaryote, ledit procédé étant caractéri-
sé en ce que chaque région cible comprend le même site de
reconnaissance de ladite nucléase. La présente invention
concerne également les organismes ainsi obtenus, et les
différentes utilisations de ces organismes.
(pas de figure).

FR 3 142 490 - A1



Description

Titre de l'invention : Procédé de translocation chromosomique chez un organisme eucaryote

[0001] La présente invention concerne un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance d'une nucléase. La présente invention concerne également les organismes ainsi obtenus, notamment les souches de champignons, et les différentes utilisations d'un tel organisme.

Contexte de l'invention

[0002] L'utilisation d'une nucléase pour modifier le génome d'un organisme a déjà été décrit à de nombreuses reprises. A titre d'exemple, on peut faire référence à la demande internationale PCT/IB2019/058760 qui décrit un système d'édition génomique, en particulier un système CRISPR, qui cible une région cible génomique dans une cellule eucaryote.

[0003] Les réarrangements induits par des nucléases ont également déjà été décrits chez différents organismes. Tang et al. ont notamment décrit l'induction de réarrangements chromosomiques chez des levures, à l'aide d'un système CRISPR-Cas9. Ainsi, à l'aide de deux ARN guides, chacun permettant de guider la nucléase CRISPR-Cas9 vers le site de coupure dans une région cible, une coupure double brin d'ADN est réalisée sur deux chromosomes. L'utilisation de deux séquences guides pour induire des translocations par CRISPR dans des cellules souches animales a également été décrite dans la demande de brevet CN104726494A.

[0004] Les réarrangements chromosomiques peuvent permettre de conférer des propriétés nouvelles ou améliorées aux génomes ainsi modifiés. Les différentes méthodes d'édition du génome peuvent être utilisées pour induire de tels réarrangements. Un inconvénient est cependant que chaque région cible à couper doit comprendre un site de reconnaissance de la nucléase. Pour induire une translocation chromosomique, a minima deux régions cibles et par conséquent deux sites de reconnaissance de nucléase doivent donc être présents (par exemple au moins deux sites sur le même chromosome ou au moins un site sur deux chromosomes).

[0005] Le système CRISPR-Cas9 permet de cibler simultanément plusieurs cibles, mais cela reste difficile à l'aide d'autres outils tels que les méganucléases ou les nucléases à doigt de zinc.

[0006] Il existe donc un besoin pour un nouvel outil permettant d'effectuer des translocations chromosomiques qui soit simple, rapide et efficace. De façon avantageuse cet

outil devrait pouvoir être utilisé chez tous les organismes eucaryotes, notamment chez les champignons.

[0007] De plus, la souche QM6a est la souche de champignon à partir de laquelle toutes les souches industrielles de *Trichoderma reesei* ont été issues. Des techniques de mutagénèses aléatoires ont été utilisées sur ces souches industrielles dans le but d'augmenter leur quantité de protéines cellulose produites et secrétées. Malheureusement ces mutagénèses aléatoires ont été réalisées sans chercher à identifier les événements génétiques effectivement survenus. Depuis les années 2010, les avancées dans les méthodes de séquençage haut débit (NGS) ont permis de connaître l'agencement exacte des 34 millions de paires de bases qui constituent le génome de *Trichoderma reesei*. Le séquençage complet du génome des souches QM6a, RutC30, CL847 ou NG14 (Le Crom et al.) a contribué à identifier des mutations génétiques ponctuelles et trois translocations dans les souches RutC30 et CL847. Les événements de translocation chromosomique ne sont pas rares dans les souches obtenues par mutagénèses aléatoires et des réarrangements chromosomiques ont été détectés dans d'autres souches telles que QM9414 et QM9978 (Jourdier et al. ; Ivanova et al. ; Vitikainen et al.). Cependant, en l'absence d'outils d'ingénierie génétique adaptés, il est néanmoins difficile d'établir un lien entre réarrangements chromosomiques et productivité de la souche. Dans les exemples de souche décrits ci-dessus, les réarrangements sont la conséquence d'événements mutagènes non maîtrisés.

[0008] L'ingénierie des génomes est une approche essentielle pour comprendre comment l'architecture des chromosomes impacte les phénotypes. A cet égard, il existe donc un besoin pour un outil permettant d'analyser et de réarranger de larges régions chromosomiques, notamment chez les champignons.

[0009] La présente invention permet de répondre à ces différents besoins.

Bref exposé de l'invention

[0010] La présente invention repose sur les résultats des Inventeurs montrant qu'il est possible d'avoir recours à un seul et unique site de reconnaissance d'une nucléase pour initier un brassage chromosomique, quel que soit le nombre de croisement souhaité.

[0011] De nombreux avantages découlent de l'utilisation de ce site unique. En effet, l'utilisation d'un site unique permet de s'affranchir de la nécessité d'identifier et de cibler un site différent pour chaque région cible. De plus, le site de reconnaissance de la nucléase étant unique, cela confère des coupures identiques pour chaque site de reconnaissance (naturellement présent dans le génome ou intégré) quel que soit la nucléase utilisée. Les effets indésirables (off target) sont par ailleurs limités puisqu'un seul site de reconnaissance et une seule nucléase sont utilisés. L'utilisation d'un seul site de reconnaissance et d'une seule nucléase permet par ailleurs de réduire les coûts (technique et humain) par rapport aux autres techniques classiquement utilisées.

- [0012] La présente invention permet ainsi de disposer d'un outil d'édition génétique permettant de réagencer de larges régions chromosomiques, ce qui peut ainsi permettre de 1) démêler les impacts phénotypiques des grands réarrangements chromosomiques de ceux des mutations ponctuelles, 2) modifier l'architecture de l'information génétique le long des chromosomes (par exemple le regroupement de gènes) et 3) réorganiser les génomes transloqués lors de mutagénèses aléatoires pour rétablir des cariotypes sauvages.
- [0013] Dans un premier aspect, l'invention concerne ainsi un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur le ou les chromosome(s) impliqué(s) dans la translocation au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase. Selon un mode de réalisation préféré, l'invention concerne ainsi un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur chaque chromosome impliqué dans la translocation au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase.
- [0014] Dans un second aspect, l'invention concerne aussi un organisme eucaryote, notamment une souche de champignon, transformé par le procédé de translocation chromosomique selon l'invention.
- [0015] Dans un autre aspect, l'invention concerne aussi un organisme eucaryote, notamment une souche de champignon, comprenant au moins deux fois le même site de reconnaissance d'une nucléase dans son génome, notamment dans au moins deux chromosomes différents.
- [0016] L'invention concerne aussi les différentes utilisations d'un organisme selon l'invention.
- [0017] L'invention concerne aussi un procédé pour identifier le site d'insertion d'un gène dans un organisme eucaryote, notamment un champignon, ledit procédé comprenant au moins les étapes suivantes :
- obtention d'un organisme eucaryote dans lequel un gène est inséré de manière ectopique, ledit gène comprenant un fragment de séquence similaire à un autre séquence du génome de l'organisme,
 - réalisation d'une translocation chromosomique entre ledit fragment du gène et la deuxième séquence possédant une similarité avec ledit fragment,
 - analyse génomique dudit organisme eucaryote, après la translocation chromosomique, pour identifier le site d'insertion dudit gène.

[0018] L'invention concerne aussi un procédé de croisement entre les souches de champignon selon l'invention.

Exposé de l'invention

[0019] L'invention concerne ainsi un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur le ou les chromosome(s) impliqué(s) dans la translocation au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase.

[0020] Selon un mode de réalisation, ledit procédé comprend au moins une étape de coupure dans le même chromosome ou entre au moins deux chromosomes, au niveau du site de reconnaissance d'une nucléase. Plus particulièrement, l'étape de coupure du double brin d'ADN sur le chromosome s'entend au moins de deux coupures du double brin d'ADN sur le même chromosome (ladite translocation peut donc être effectuée au sein d'un chromosome unique).

[0021] L'invention concerne ainsi un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, deux coupures du double brin d'ADN au sein du même chromosome, au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase.

[0022] L'invention concerne ainsi un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur chaque chromosome impliqué dans la translocation au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase.

[0023] Selon l'invention, une « translocation chromosomique » s'entend d'un réarrangement chromosomique ou brassage chromosomique, au sein d'un même chromosome ou entre au moins deux chromosomes. La translocation peut s'entendre d'un échange de matériel génétique entre chromosomes chez des organismes haploïdes, diploïdes ou hétérocaryotiques. La translocation chromosomique peut également s'entendre d'un échange de matériel génétique entre chromosomes homologues ou non-homologues, et/ou entre au moins deux chromosomes qui n'appartiennent pas à la même paire de chromosomes. Selon un mode de réalisation préféré, ladite « translocation chromosomique » s'entend d'un réarrangement chromosomique entre un ou deux chromosomes. Selon un mode de réalisation préféré, ladite « translocation chromosomique » s'entend d'un réarrangement chromosomique entre deux chromosomes

uniquement. Selon un mode de réalisation, l'invention concerne ainsi un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur chacun des deux chromosomes impliqués dans la translocation au niveau du site de reconnaissance d'une nucléase.

- [0024] Selon l'invention, une « région cible » s'entend d'une région chromosomique, plus particulièrement une région chromosomique à réarranger. Chaque région cible comprend un site de reconnaissance d'une nucléase, et pour une translocation donnée, toutes les régions cibles dans le génome comprennent le même site de reconnaissance de la nucléase. Selon un mode de réalisation, la région cible comprend naturellement le site de reconnaissance d'une nucléase, et/ou la région cible comprend un site de reconnaissance qui a été intégré au génome. Selon un mode de réalisation préféré, l'organisme comprend naturellement un site de reconnaissance d'une nucléase au sein d'une région cible et au moins un site de reconnaissance qui a été intégré au sein d'une autre région cible. Selon un mode de réalisation alternatif, l'organisme comprend uniquement des sites de reconnaissance intégrés dans les au moins deux régions cibles.
- [0025] Le double brin d'ADN est ainsi coupé au moins deux fois dans le procédé selon l'invention. Selon un mode de réalisation, la coupure du double brin d'ADN est réalisée par une cassure double brin ou deux cassures simple brin. A titre d'exemple, une étape de coupure du double brin d'ADN sur chaque chromosome impliqué dans la translocation peut ainsi s'entendre de deux cassures double brin (une sur chaque chromosome) ou de quatre cassures simples brin (deux sur chaque chromosome).
- [0026] Selon l'invention la translocation chromosomique s'effectue entre au moins deux régions cibles. Cela signifie que seules deux régions peuvent être réarrangées, mais cela signifie également que plus de deux régions, par exemple trois ou quatre, peuvent être réarrangées. Ladite translocation peut donc être simple (entre deux régions) ou complexe (entre au moins trois régions). Selon un mode de réalisation, la translocation chromosomique s'entend d'une translocation entre deux à quatre régions. Selon un mode de réalisation préféré, la translocation chromosomique s'entend d'une translocation entre deux régions uniquement. Selon un mode de réalisation préféré, la translocation chromosomique s'entend d'une translocation entre deux régions cibles et entre deux chromosomes.
- [0027] Selon un mode de réalisation de l'invention, la translocation chromosomique n'entraîne pas de perte de matériel génétique et/ou ladite translocation chromosomique entraîne une perte de matériel génétique.
- [0028] Selon un mode de réalisation, le procédé selon l'invention comprend au moins deux étapes de coupure du double brin d'ADN sur le même chromosome. A titre d'exemple, deux coupures sur le même chromosome sont donc réalisées, ce qui permet l'obtention

de trois fragments chromosomiques, et seulement deux sont réarrangés, avec perte du matériel génétique présent sur le troisième fragment.

- [0029] Selon un mode de réalisation, le procédé selon l'invention comprend au moins une étape de coupure du double brin d'ADN sur chaque chromosome impliqué dans la translocation. A minima, deux coupures sur deux chromosomes sont donc réalisées : une première coupure sur un premier chromosome et une deuxième coupure sur un deuxième chromosome. A titre exemple, trois ou quatre coupures peuvent également être réalisées : par exemple une coupure sur un premier chromosome et deux autres coupures sur un deuxième chromosome, ou encore deux coupures sur un premier chromosome et deux autres coupures sur un deuxième chromosome.
- [0030] Selon un mode de réalisation, le procédé selon l'invention peut comprendre à la fois des étapes de coupure du double brin d'ADN au sein du même chromosome et entre des chromosomes différents.
- [0031] Selon l'invention, chaque coupure du double brin est effectuée au niveau d'un site de reconnaissance d'une nucléase. Chaque chromosome impliqué dans la translocation comprend donc au moins un site de reconnaissance d'une nucléase (qui est identique). Tous les chromosomes impliqués dans la translocation comprennent donc au moins un site de reconnaissance d'une nucléase (qui est identique).
- [0032] Le site de reconnaissance d'une nucléase correspond à une séquence d'acides nucléiques qui est spécifiquement ciblée par la nucléase. Selon un mode de réalisation, le site de reconnaissance est spécifique d'une seule nucléase. Cela permet avantageusement d'assurer la spécificité de la coupure du double brin à un site spécifiquement déterminé et identifiable dans le génome de l'organisme.
- [0033] Selon un mode de réalisation, l'invention concerne un procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN au niveau de deux sites de reconnaissance d'une nucléase au sein du même chromosome, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase.
- [0034] Selon un mode de réalisation préféré, l'invention concerne un procédé de translocation chromosomique entre deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur chacun des deux chromosomes impliqués dans la translocation au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase.
- [0035] Selon l'invention, un « organisme eucaryote » s'entend de tout organisme, qu'il soit unicellulaire ou multi-cellulaire, dont la/les cellule(s) possède(nt) un noyau. Plus particulièrement, cela comprend, de façon non-exhaustive, les levures, les champignons, les

plantes, les animaux vertébrés y compris les mammifères. Lorsque l'organisme est un mammifère :

- le procédé selon l'invention n'est pas une méthode de traitement chirurgical ou thérapeutique du corps humain ou animal, et/ou
- le procédé selon l'invention n'est pas un procédé de modification de l'identité génétique, notamment identité génétique germinale, de l'être humain, et/ou
- le procédé selon l'invention n'est pas un procédé de modification de l'identité génétique des animaux de nature à provoquer chez eux des souffrances sans utilité médicale substantielle pour l'homme ou l'animal.

[0036] Selon un mode de réalisation préféré, l'organisme eucaryote selon l'invention est un champignon, notamment un champignon filamenteux. A titre d'exemple on peut citer les champignons du phylum des Ascomycètes (Ascomycota), des Basidiomycètes (Basidiomycota) et des Zygomycètes (Zygomycota). Encore plus préférentiellement, le champignon filamenteux est choisi parmi les classes des orbiliomycètes, des péziomycètes, des dothideomycètes, des eurotiomycètes, des lecanoromycètes, des léotiomycètes, des sordariomycètes et des saccharomycètes. De façon avantageuse, le champignon filamenteux selon l'invention appartient à la classe des sordariomycètes, notamment au genre *Trichoderma*, et plus particulièrement à l'espèce *Trichoderma reesei*.

[0037] Selon l'invention, la nucléase permet d'entraîner une coupure du double brin d'ADN. Tout enzyme ayant cette activité hydrolytique peut être utilisée. La nucléase permet de couper le double brin d'ADN à un site de coupure précis et identifié en amont de la coupure. Cette coupure du double brin peut être directe (il y a une cassure double brin) ou bien indirecte (deux cassures simple brin sont réalisées). La nucléase peut être de type sauvage, artificielle ou mutée ou encore un variant d'une nucléase.

[0038] Lorsqu'une nucléase entraînant directement une cassure double brin est utilisée, alors a minima il y a deux coupures qui sont réalisées (au moins deux coupures au sein du même chromosome ou bien une coupure sur deux chromosomes différents).

[0039] Lorsqu'une nucléase entraînant seulement une cassure simple brin est utilisée, alors a minima il y a quatre coupures qui sont réalisées (au moins quatre coupures au sein du même chromosome ou bien deux coupures sur au moins deux chromosomes différents).

[0040] Selon un mode de réalisation, la nucléase est choisie parmi une méganucléase, une nucléase à doigt de zinc, une nucléase effectrice de type activateur de transcription (nucléase TALE), une nickase et une nucléase Cas. L'avantage des outils de type nucléase à doigt de zinc, nucléase TALE, CRISPR/nickase ou CRISPR/Cas est qu'ils peuvent être construits à façon afin de reconnaître une séquence d'ADN particulière. Selon un mode de réalisation préféré, la nucléase est ainsi choisie parmi une nucléase à

doigt de zinc, une nucléase TALE et une nucléase Cas. Avantageusement, la nucléase est une nucléase Cas, de préférence CRISPR/Cas9.

- [0041] Les méganucléases sont des endodésoxyribonucléases. A titre d'exemple les méganucléases I-SceI, I-CreI, I-DmoI peuvent être utilisées. Le site de reconnaissance d'une méganucléase comprend typiquement entre 12 et 40 paires de bases.
- [0042] Les nucléases à doigts de zinc ou les nucléases Tale sont des enzymes de restriction artificielles. Elles sont créées par la fusion d'un domaine de liaison à l'ADN, respectivement de type « doigt de zinc » ou TALE (Transcription activator-like effector), et d'un domaine catalytique de coupure de l'ADN (FokI ou un variant). L'utilisation d'un variant du domaine FokI peut être envisagée afin d'améliorer l'activité ou la spécificité de l'outil (e.g. Guo et al.).
- [0043] Les nucléases Cas sont des nucléases associées à l'outil CRISPR. L'outil CRISPR/Cas se compose (i) d'une protéine appelée Cas doté d'une activité nucléase permettant de couper un double brin d'ADN et (ii) d'un domaine de recrutement d'un ARN spécifique dit ARN guide. Ce dernier est composé d'une séquence variable d'environ 16-23 nucléotides qui permet de pointer un locus très précis de l'ADN cible, grâce à la complémentarité entre les environ 16-23 (par exemple 20) bases de l'ARN guide et les environ 16-23 (par exemple 20) bases cibles dans le génome de l'organisme. L'ARN guide est également composé d'une séquence nucléotidique invariable permettant la liaison entre l'ARN guide et la protéine Cas. Une fois le guide en place, l'activité nucléase de la protéine Cas s'active et coupe l'ADN à l'endroit précis et déterminé à l'avance (en choisissant la séquence d'ARN guide à façon). A titre d'exemple les nucléases Cas9 et Cas12 peuvent être utilisées.
- [0044] Les nickases sont également des enzymes associées à l'outil CRISPR. Il s'agit d'enzyme de type Cas dans laquelle un des deux domaines de type nucléase est non fonctionnel (par mutation ou autre modification génétique). Par exemple, une nucléase Cas9 dans laquelle un des deux domaines HNH ou RuvC est inactivé est une nickase. De préférence la nickase est une nickase Cas9.
- [0045] Toutes ces nucléases permettent de couper l'ADN double brin à un site de coupure précis, identifié et connu à l'avance. Selon un mode de réalisation, l'étape de coupure du double brin d'ADN sur chaque chromosome impliqué dans la translocation est ainsi effectuée au niveau du site de reconnaissance identifié d'une nucléase.
- [0046] Selon un mode de réalisation de l'invention, le site de reconnaissance de ladite nucléase comprend au moins 12 paires de base. L'expression « au moins 12 paires de bases » s'entend avantageusement de 12 paires de base consécutives. Selon un mode de réalisation de l'invention, le site de reconnaissance de ladite nucléase comprend entre 12 et 40 paires de base, par exemple entre 12 et 20. L'expression « entre 12 et 40 paires de base » s'entend de toutes les valeurs comprises entre 12 et 40, soit 12, 13, 14,

15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39 et 40.

- [0047] La région cible selon l'invention doit contenir un site de reconnaissance d'une nucléase. Le choix de ce site de reconnaissance est déterminé en fonction de l'expérience à mener, du type de translocation chromosomique à réaliser, ... A titre d'exemple, le site de reconnaissance d'une nucléase peut être choisi afin que la translocation permette d'invalider un gène, identifier la position d'un transgène, ...
- [0048] Selon un mode de réalisation, la séquence du site de reconnaissance de la nucléase est uniquement présente dans le génome au niveau des sites de coupure d'intérêt. Cela signifie que le site de reconnaissance est spécifique, ce qui permet de s'assurer de la spécificité de la coupure aux sites souhaités et identifiés (et donc éviter les coupures indésirables).
- [0049] De préférence, la translocation chromosomique selon l'invention s'entend d'une translocation chromosomique dirigée. Elle est réalisée après identification des régions cibles à réarranger, et éventuellement insertion d'au moins une cassette contenant le site de reconnaissance de la nucléase choisie aux sites chromosomiques souhaités.
- [0050] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend, avant l'étape de coupure du double brin d'ADN, une étape d'identification du site de reconnaissance de ladite nucléase.
- [0051] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend après l'étape d'identification du site de reconnaissance de ladite nucléase une étape de vérification que la séquence n'existe pas ailleurs dans le génome. Cette étape vise à confirmer que le site de reconnaissance est unique (présent uniquement au niveau des sites de coupure d'intérêt) et n'entraînera pas de coupure indésirable de l'ADN.
- [0052] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend avant l'étape de coupure du double brin d'ADN, une étape d'identification de la région cible. Cette étape peut être concomitante, ou non, à l'identification du site de reconnaissance de la nucléase.
- [0053] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend avant l'étape de coupure du double brin d'ADN, une étape d'insertion dans le génome d'un organisme, notamment un champignon, d'un ou plusieurs site(s) de reconnaissance de ladite nucléase dans les régions chromosomiques d'intérêt (e.g. les cibles de la translocation chromosomique). Cette étape est avantageusement suivie d'une étape d'identification et/ou sélection des clones d'intérêt (e.g. ceux dans lesquels la cassette a bien été intégrée dans le génome). Avant la translocation chromosomique, chaque région cible comprend ainsi le même site de reconnaissance de la nucléase.

- [0054] Selon un mode de réalisation, dans le procédé de translocation chromosomique selon l'invention, l'insertion dans le génome d'un organisme d'au moins un site de reconnaissance d'une nucléase est effectuée à l'aide d'une cassette. La cassette selon l'invention comprend tous les éléments nécessaires à son intégration dans le génome. De préférence, ladite cassette comprend au moins un site de reconnaissance de ladite nucléase et un marqueur. Ledit marqueur peut comprendre le site de reconnaissance de la nucléase. Selon un mode de réalisation, le marqueur s'entend d'un marqueur de sélection, par exemple un gène dont l'expression confère aux cellules qui le contiennent une caractéristique permettant de les sélectionner, ou un gène dont la non-expression confère aux cellules qui contiennent une version défectueuse de ce gène une caractéristique permettant de les sélectionner par rapport aux cellules contenant une version fonctionnelle de ce même gène. L'utilisation d'un marqueur de sélection permet en effet d'identifier les cellules ayant intégré une modification génétique par rapport à celles qui ne l'ont pas intégrée. Il s'agit par exemple d'un gène de résistance aux antibiotiques, ou d'un gène modifiant l'auxotrophie d'un organisme, par exemple le gène URA3 ou URA5.
- [0055] Plus particulièrement, une cassette selon l'invention comprend :
- une région flanquante en amont du site de reconnaissance d'une nucléase,
 - ledit site de reconnaissance d'une nucléase,
 - une région flanquante en aval du site de reconnaissance d'une nucléase.
- [0056] Selon un mode de réalisation particulier, une cassette selon l'invention comprend :
- une région flanquante en amont du site de reconnaissance d'une nucléase,
 - optionnellement un promoteur,
 - ledit site de reconnaissance d'une nucléase, par exemple la séquence d'un gène comprenant un site de reconnaissance d'une nucléase,
 - optionnellement un terminateur,
 - une région flanquante en aval du site de reconnaissance d'une nucléase.
- [0057] Une cassette selon l'invention peut ainsi être standardisée, seules les régions flanquantes en amont et aval changent pour cibler l'intégration de la cassette aux locus souhaités dans le génome.
- [0058] A titre d'exemple, le site de reconnaissance est choisi au sein d'un gène rapporteur au sein de l'organisme.
- [0059] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend, avant l'étape de coupure du double brin d'ADN, une étape d'insertion dans le génome d'un organisme, notamment un champignon, d'un ou plusieurs site(s) de reconnaissance de ladite nucléase, notamment dans au moins une des régions cibles d'un chromosome. A titre d'exemple, dans ce mode de réalisation, une région cible sur un chromosome comprend déjà un site de re-

connaissance d'une nucléase, et une cassette comprenant le même site de reconnaissance de la nucléase est insérée afin qu'au moins une deuxième région cible d'intérêt sur un autre chromosome comprenne le même site de reconnaissance de la nucléase. A titre d'exemple, un chromosome peut être porteur d'un marqueur de sélection comprenant un site de reconnaissance d'une nucléase et la région comprenant ce marqueur peut ainsi caractériser une région cible selon l'invention ; à l'aide d'une cassette comprenant le même marqueur de sélection comprenant le même site de reconnaissance d'une nucléase, un deuxième site de reconnaissance d'une nucléase peut ainsi être inséré au niveau d'un autre chromosome. A titre d'exemple, une région cible sur un chromosome comprend déjà un site de reconnaissance d'une nucléase, et une cassette comprenant le même site de reconnaissance de la nucléase est insérée afin qu'au moins une deuxième région cible d'intérêt sur le même chromosome comprenne le même site de reconnaissance de la nucléase. Dans ce cas, un chromosome peut être porteur d'un marqueur de sélection comprenant un site de reconnaissance d'une nucléase et la région comprenant ce marqueur peut ainsi caractériser une région cible selon l'invention ; à l'aide d'une cassette comprenant le même marqueur de sélection comprenant le même site de reconnaissance d'une nucléase, un deuxième site de reconnaissance d'une nucléase peut ainsi être inséré au niveau du même chromosome.

[0060] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend avant l'étape de coupure du double brin d'ADN, une étape d'insertion dans le génome d'un organisme, notamment un champignon, d'un ou plusieurs site(s) de reconnaissance de ladite nucléase dans au moins une région cible sur le chromosome impliqué dans la translocation. A titre d'exemple, dans ce mode de réalisation, aucune région cible sur le chromosome ne comprend un site de reconnaissance d'une nucléase, et une cassette comprenant un site de reconnaissance d'une nucléase est insérée afin que chaque région cible d'intérêt, sur le même chromosome, comprenne le même site de reconnaissance de la nucléase.

[0061] Selon un mode de réalisation, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend avant l'étape de coupure du double brin d'ADN, une étape d'insertion dans le génome d'un organisme, notamment un champignon, d'un ou plusieurs site(s) de reconnaissance de ladite nucléase dans au moins une région cible sur chaque chromosome impliqué dans la translocation. A titre d'exemple, dans ce mode de réalisation, aucune région cible ne comprend un site de reconnaissance d'une nucléase, et une cassette comprenant un site de reconnaissance d'une nucléase est insérée afin que chaque région cible d'intérêt, sur au moins deux chromosomes, comprenne le même site de reconnaissance de la nucléase.

[0062] Selon un mode de réalisation préféré, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend les étapes suivantes :

- optionnellement, l'identification d'un site de reconnaissance d'une nucléase,
- l'insertion dans le génome d'un organisme, notamment un champignon, d'au moins un site de reconnaissance de ladite nucléase,
- la sélection d'un organisme issu de l'étape précédente comprenant le site de reconnaissance de ladite nucléase dans chaque région cible,
- l'introduction dans l'organisme issu de l'étape précédente de ladite nucléase ou d'un vecteur comprenant au moins un acide nucléique codant pour ladite nucléase,
- optionnellement, la vérification de la translocation chromosomique.

[0063] Selon un mode de réalisation préféré, le procédé de translocation chromosomique selon l'invention est caractérisé en ce qu'il comprend les étapes suivantes :

- optionnellement, l'identification d'un site de reconnaissance d'une nucléase,
- l'insertion dans le génome d'un organisme, notamment un champignon, d'au moins un site de reconnaissance de ladite nucléase, à l'aide d'une cassette selon l'invention comprenant: (i) une région flanquante en amont du site de reconnaissance d'une nucléase, (ii) ledit site de reconnaissance d'une nucléase, et (iii) une région flanquante en aval du site de reconnaissance d'une nucléase,
- la sélection d'un organisme issu de l'étape précédente comprenant le site de reconnaissance de ladite nucléase dans chaque région cible, sur au moins deux chromosomes,
- l'introduction dans l'organisme issu de l'étape précédente de ladite nucléase ou d'un vecteur comprenant au moins un acide nucléique codant pour ladite nucléase,
- optionnellement, la vérification de la translocation chromosomique.

[0064] Ladite vérification de la translocation chromosomique peut être réalisée par tout outil d'analyse génomique approprié, par exemple le séquençage ou par analyse PCR.

[0065] Selon un mode de réalisation, la nucléase (par exemple Cas9) peut être introduite directement dans l'organisme. Un complexe RNP (Ribonucléoprotéine) comprenant l'ARN guide et la nucléase est ainsi formé in vitro puis introduit dans l'organisme.

[0066] Selon un autre mode de réalisation, la nucléase (par exemple Cas9) et l'ARN guide sont produits in vivo, dans l'organisme. Un vecteur (par exemple un plasmide) comprenant au moins un acide nucléique codant pour ladite nucléase et l'ARN guide est introduit dans l'organisme. Un « vecteur » peut s'entendre de 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 un organisme. 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.

[0067] Selon un mode de réalisation, dans le procédé de translocation chromosomique selon l'invention, ladite translocation chromosomique peut être réalisée avec une matrice de

réparation. Cela peut permettre d'augmenter le réarrangement ciblé. Une matrice de réparation peut comprendre notamment la région flanquante en amont de la séquence du site de reconnaissance d'une nucléase et la région flanquante en aval de la séquence du site de reconnaissance d'une nucléase, tel que présent dans le génome avant la translocation. De préférence, la matrice de réparation comprend uniquement la région flanquante en amont de la séquence du site de reconnaissance d'une nucléase et la région flanquante en aval de la séquence du site de reconnaissance d'une nucléase.

- [0068] Selon un mode de réalisation, dans le procédé de translocation chromosomique selon l'invention, ladite translocation chromosomique est réalisée sans utilisation d'une matrice de réparation. En effet, les coupures réalisées par la nucléase visent à initier un brassage chromosomique, et ne visent pas à insérer dans le génome la version modifiée d'un gène (par exemple remplacer un gène défectueux par la copie d'un gène fonctionnel), ou bien insérer un transgène.
- [0069] Selon un mode de réalisation, dans le procédé de translocation chromosomique selon l'invention, ladite coupure est effectuée dans les régions cibles.
- [0070] Selon un mode de réalisation, dans le procédé de translocation chromosomique selon l'invention, chaque coupure sur le chromosome ou chaque chromosome est effectuée au niveau d'un seul site de reconnaissance de ladite nucléase.
- [0071] Dans un deuxième aspect, l'invention concerne un organisme eucaryote, notamment une souche de champignon, transformé par le procédé de translocation chromosomique selon l'invention.
- [0072] Selon l'invention, un « organisme eucaryote » s'entend de tout organisme, qu'il soit unicellulaire ou multi-cellulaire, dont la/les cellule(s) possède(nt) un noyau. Selon un mode de réalisation, l'organisme eucaryote s'entend d'une cellule eucaryote ou d'une cellule eucaryote hôte. L'organisme ou cellule eucaryote selon l'invention n'est pas une cellule de lignée germinale humaine.
- [0073] L'invention concerne aussi la descendance d'un organisme ou cellule eucaryote selon l'invention. La descendance de l'organisme ou cellule eucaryote selon l'invention n'est pas une cellule de lignée germinale humaine et/ou ne comprend pas de corps humain.
- [0074] L'invention concerne aussi un organisme eucaryote, notamment une souche de champignon, comprenant au moins deux fois le même site de reconnaissance d'une nucléase dans son génome, notamment dans au moins deux chromosomes différents. Le même site de reconnaissance d'une nucléase peut également être présent au moins deux fois au sein du même chromosome dans le génome de l'organisme. De tels organismes sont ainsi prêts à réaliser une translocation chromosomique selon l'invention. Selon un mode de réalisation, un tel organisme a été modifié par l'insertion dans son génome, d'un ou plusieurs site(s) de reconnaissance d'une nucléase dans les régions chromosomiques d'intérêt (e.g. les cibles de la translocation

chromosomiques). De préférence, au moins un site de reconnaissance d'une nucléase a été insérée dans un organisme à l'aide d'une cassette. Selon un mode de réalisation, un tel organisme eucaryote comprend au moins deux fois le même site de reconnaissance d'une nucléase dans au moins deux régions cibles situées sur au moins un chromosome. Selon un mode de réalisation particulier, un tel organisme eucaryote comprend au moins deux fois le même site de reconnaissance d'une nucléase dans au moins deux régions cibles situées sur au moins deux chromosomes différents.

[0075] Selon un mode de réalisation, ledit organisme eucaryote est isolé et/ou purifié.

[0076] Dans un troisième aspect, l'invention concerne aussi l'utilisation d'un organisme tel que défini précédemment pour diverses applications. L'organisme est tel que défini précédemment. Selon un mode de réalisation préféré, il s'agit d'un champignon.

[0077] Selon un mode de réalisation, l'invention concerne l'utilisation d'un organisme tel que défini précédemment (notamment une souche de champignon), pour la production de molécules ou de protéines d'intérêt. Selon l'invention, les molécules ou protéines d'intérêt sont toutes les molécules ou protéines pouvant être produites par un organisme eucaryote de façon naturelle (par exemple certaines enzymes) ou bien par modification génétique (par exemple après transformation à l'aide d'un vecteur approprié). La molécule ou protéine d'intérêt peut être endogène ou exogène. La molécule ou protéine d'intérêt peut être native (telle que retrouvée dans la nature), chimérique ou synthétisée. La molécule ou protéine d'intérêt peut également être un variant d'une molécule ou protéine, présentant notamment des propriétés améliorées et/ou modifiées par rapport à la molécule ou protéine de référence (par exemple une propriété biologique améliorée par rapport à une protéine native). Les molécules ou protéines d'intérêt peuvent être des protéines utiles à tout type d'industrie (la production de biocarburant, alimentaire, cosmétique ou pharmaceutique, ...)

[0078] Selon un mode de réalisation, un organisme selon l'invention est modifié afin de comprendre la séquence d'un gène d'intérêt, d'exprimer le gène d'intérêt de façon constitutive ou bien de surexprimer ledit gène d'intérêt. Un gène d'intérêt code pour une protéine d'intérêt selon l'invention. Ledit gène d'intérêt peut comprendre des éléments additionnels appropriés comme, par exemple, une séquence codant pour un signal de sécrétion.

[0079] Selon un mode de réalisation préféré, les protéines d'intérêt selon l'invention sont des enzymes cellulolytiques ou des protéines d'intérêt hétérologues. Une protéine hétérologue s'entend de toute protéine produite par un organisme dont le matériel génétique a été modifié afin de produire ladite protéine. Selon un mode de réalisation, les enzymes cellulolytiques sont des cellulases ou des hémicellulases. De façon préférée, les enzymes sont des cellulases. Selon l'invention, le terme « cellulases » s'entend plus particulièrement des enzymes choisies parmi les endoglucanases, les

exoglucanases et les glucosidases. Le terme « cellulase » fait plus particulièrement référence à une enzyme adaptée à l'hydrolyse de la cellulose et permettant aux microorganismes (tels que *Trichoderma reesei*) qui les produisent d'utiliser la cellulose comme source de carbone, en hydrolysant ce polymère en sucres simples (glucose).

[0080] Selon un mode de réalisation, l'invention concerne aussi un procédé de production d'une protéine d'intérêt comprenant au moins une étape de culture d'un organisme modifié selon le procédé de translocation chromosomique selon l'invention dans des conditions permettant la production de la protéine d'intérêt.

[0081] Selon un mode de réalisation, l'invention concerne aussi un procédé de production d'une protéine d'intérêt comprenant les étapes suivantes :

- (i) une étape de transformation d'un organisme selon l'invention à l'aide d'un vecteur comprenant une cassette d'expression comprenant (1) un promoteur, (2) un gène d'intérêt codant pour ladite protéine d'intérêt et (3) un terminateur,
- (ii) une étape de culture de l'organisme transformé à l'étape (i) dans des conditions permettant la production de la protéine d'intérêt. Selon un mode de réalisation particulier, ledit procédé de production d'une protéine d'intérêt comprend une étape de sélection d'un organisme transformé entre l'étape de transformation et l'étape de culture.

[0082] L'étape de culture s'entend d'une culture dans un milieu approprié. Il s'agit d'un milieu adapté à la survie de l'organisme transformé (tel que le champignon) et/ou à la croissance de l'organisme et/ou à la production de la protéine d'intérêt, voire sa sécrétion. Il peut s'agir d'un milieu adapté à la culture de l'organisme, complété au besoin pour devenir un milieu adapté à la production et/ou sécrétion de la protéine d'intérêt.

[0083] Selon un mode de réalisation, le procédé de production d'une protéine d'intérêt peut comprendre une étape de sécrétion de ladite protéine d'intérêt. La sécrétion peut être consécutive ou concomitante à la production de la protéine d'intérêt.

[0084] De façon avantageuse, plus particulièrement lorsque le gène d'intérêt code pour une enzyme cellulolytique et que l'organisme est un champignon, ledit procédé de production d'une protéine d'intérêt comprend aussi une phase de croissance d'une souche de champignon obtenue selon l'invention, puis une phase de croissance et de production de protéines d'intérêt. Encore plus préférentiellement, ladite phase de croissance est réalisée en présence d'un substrat de croissance et ladite phase de croissance et de production de protéines d'intérêt est réalisée en présence d'un substrat inducteur. Le substrat de croissance et le substrat inducteur sont de préférence des substrats carbonés.

[0085] Selon un mode de réalisation, l'invention concerne l'utilisation d'un organisme tel que défini précédemment (notamment une souche de champignon modifiée selon le

procédé de translocation chromosomique selon l'invention), pour l'hydrolyse de la cellulose ou de la lignocellulose en sucre, notamment les sucres C5 (à 5 atomes de carbone) et/ou C6 (à 6 atomes de carbone), tel que le glucose.

[0086] Selon un mode de réalisation, la présente invention concerne un procédé de production de sucre à partir de substrats cellulotiques ou lignocellulosiques, comprenant les étapes suivantes :

- (i) une étape de prétraitement d'un substrat cellulotique ou lignocellulosique afin d'obtenir un substrat prétraité,
- (ii) une étape d'utilisation/culture d'un organisme selon l'invention, optionnellement transformé à l'aide d'un vecteur comprenant une cassette d'expression comprenant (1) un promoteur, (2) un gène d'intérêt codant pour ladite protéine d'intérêt, notamment une enzyme cellulolytique et (3) un terminateur,
- (iii) une étape d'hydrolyse enzymatique du substrat prétraité obtenu à l'étape i), en présence des enzymes cellulolytiques obtenues à l'étape ii), afin d'obtenir un hydrolysate.

[0087] La culture dudit organisme est réalisée dans un milieu approprié. Elle permet notamment la survie, la croissance de l'organisme et la production d'une protéine d'intérêt, notamment une enzyme cellulolytique.

[0088] Les sucres (ou jus sucré) sont ainsi obtenus à l'issue de l'étape d'hydrolyse. De préférence ce sont des sucres en C₅-C₆.

[0089] Selon un mode de réalisation, ledit procédé de production de sucres peut également comprendre un ou plusieurs étapes de séparation. Ceci permet ainsi de séparer, notamment par distillation, les différents produits d'intérêt dilués dans l'eau.

[0090] Selon un mode de réalisation, l'invention concerne l'utilisation d'un organisme tel que défini précédemment (notamment une souche de champignon), pour la production de produits biosourcés à partir de substrats cellulotiques ou lignocellulosiques.

[0091] Selon un mode de réalisation, l'invention concerne l'utilisation d'un organisme tel que défini précédemment (notamment une souche de champignon modifiée selon le procédé de translocation chromosomique selon l'invention), pour la production de biocarburant à partir de substrats cellulotiques ou lignocellulosiques.

[0092] Selon l'invention, le terme « biocarburant » s'entend de préférence d'un biocarburant de seconde génération, c'est-à-dire issu de ressources non alimentaires. Selon l'invention, le terme « biocarburant » peut également être défini comme étant tout produit issu de la transformation de la biomasse et 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 (acétone), des acides (butyrique), des lipides et leurs dérivés (acides gras à courtes ou longues chaînes, esters d'acides gras), ainsi que l'hydrogène. Le bio-carburant peut ainsi être un alcool ou un biogaz.. De manière préférée, l'alcool est par exemple l'éthanol, le butanol, l'isopropanol, le 1,2-propane diol, le 1,3-propane diol, le 1,4-butane diol, et/ou le 2,3-butane diol. Plus préférentiellement, il s'agit de l'éthanol. L'alcool ainsi obtenu peut aussi être utilisé comme solvant ou comme produit intermédiaire dans l'industrie chimique.

[0093] Selon un mode de réalisation, l'invention concerne un procédé de production d'un alcool à partir de substrats cellulotiques ou lignocellulosiques, comprenant les étapes suivantes :

- (i) une étape de prétraitement d'un substrat cellulotique ou lignocellulosique afin d'obtenir un substrat prétraité,
- (ii) une étape d'utilisation/culture d'un organisme selon l'invention, optionnellement transformé à l'aide d'un vecteur comprenant une cassette d'expression comprenant (1) un promoteur, (2) un gène d'intérêt codant pour ladite protéine d'intérêt, notamment une enzyme cellulolytique, et (3) un terminateur,
- (iii) une étape d'hydrolyse enzymatique du substrat prétraité obtenu à l'étape i), en présence des enzymes cellulolytiques obtenues à l'étape ii), afin d'obtenir un hydrolysate,
- (iv) une étape de fermentation alcoolique de l'hydrolysate obtenu,
- (v) une étape de séparation, notamment par distillation.

[0094] Dans un quatrième aspect, la présente invention concerne aussi l'utilisation du même site de reconnaissance d'une nucléase pour effectuer une translocation chromosomique entre un ou plusieurs chromosomes, avec ladite nucléase. Selon un mode de réalisation particulier, l'invention concerne ainsi l'utilisation du même site de reconnaissance d'une nucléase pour effectuer une translocation chromosomique avec ladite nucléase entre les chromosomes possédant ledit site de reconnaissance. Selon un mode de réalisation particulier, l'invention concerne l'utilisation du site de reconnaissance unique d'une nucléase pour effectuer une translocation chromosomique, avec ladite nucléase, entre le ou les chromosome(s) porteur(s) de ce site de reconnaissance unique.

[0095] Dans un cinquième aspect, la présente invention vise aussi à identifier l'impact des translocations chromosomiques sur le génotype et/ou phénotype des organismes.

[0096] Ainsi, la présente invention concerne aussi un procédé pour identifier le site d'insertion d'un gène dans un organisme eucaryote, notamment un champignon, ledit procédé comprenant au moins les étapes suivantes :

- obtention d'un organisme eucaryote dans lequel un gène est inséré de manière ectopique, ledit gène comprenant un fragment de séquence similaire à un autre séquence du génome de l'organisme,

- réalisation d'une translocation chromosomique entre ledit fragment et la deuxième séquence possédant une similarité avec ledit fragment,
- analyse génomique dudit organisme eucaryote, après la translocation chromosomique, pour identifier le site d'insertion dudit gène.

- [0097] Selon l'invention, l'expression « gène est inséré de manière ectopique » s'entend d'un gène qui a été intégré de façon aléatoire dans le génome. Son site d'insertion dans le génome n'est pas connu.
- [0098] Selon l'invention, un fragment de séquence similaire à une autre séquence du génome s'entend d'une similarité de séquence ou d'une identité de séquence. Plus particulièrement, cela s'entend d'une similarité ou d'une identité de séquence d'au moins 80%, notamment 90%, de préférence 95% et encore plus préférentiellement 99% ou 100%. Le pourcentage de similarité ou d'identité peut être déterminé par les méthodes classiques tels que les outils d'alignements de séquences et comparaisons des acides similaires ou identiques.
- [0099] Selon un mode de réalisation, ledit fragment et la deuxième séquence possédant une similarité avec ledit fragment s'entendent ici des régions cibles selon l'invention. Selon un mode de réalisation, ledit fragment et la deuxième séquence possédant une similarité avec ledit fragment s'entendent ici des régions cibles selon l'invention possédant un site de reconnaissance d'une nucléase.
- [0100] Ladite analyse génomique peut être effectuée à l'aide de tout outil conventionnel, tel que le séquençage ou la PCR. Ladite analyse génomique permet ainsi d'identifier le lieu de réarrangement chromosomique. Ladite analyse génomique peut également être couplée à une analyse phénotypique afin d'évaluer comment l'arrangement des gènes dans le génome impacte les phénotypes des organismes.
- [0101] Selon un mode de réalisation, la présente invention concerne aussi un procédé pour identifier l'impact d'une modification génétique sur le phénotype d'organisme eucaryote, notamment un champignon, ledit procédé comprenant au moins les étapes suivantes :
- obtention d'un organisme eucaryote génétiquement modifié par un procédé de translocation chromosomique tel que défini ci-dessus,
 - analyse génotypique et/ou phénotypique dudit organisme eucaryote génétiquement modifié.
- [0102] Lesdites analyses, génomique et/ou phénotypique, peuvent permettre d'identifier les régions cibles d'intérêt pour une translocation chromosomique chez d'autres organismes.
- [0103] La translocation chromosomique peut par exemple être réalisée en comparant différentes régions cibles, afin de déterminer le(s) sites de réagencement chromosomique d'intérêt. En effet, la comparaison de deux organismes qui diffèrent uniquement par le

site de translocation chromosomique peut permettre d'identifier un site de réarrangement chromosomique d'intérêt pour conférer un phénotype avantageux, lorsque l'un des organismes -après translocation chromosomique- présente un phénotype amélioré ou nouveau par rapport à l'autre organisme.

[0104] Selon un mode de réalisation, la présente invention concerne ainsi un procédé pour identifier l'impact d'une translocation chromosomique sur le phénotype d'organisme eucaryote, notamment un champignon, ledit procédé comprenant au moins les étapes suivantes :

- (i) obtention d'un organisme eucaryote génétiquement modifié par un procédé de translocation chromosomique entre deux régions cibles tel que défini ci-dessus,
- (ii) obtention d'un deuxième organisme eucaryote génétiquement modifié par un procédé de translocation chromosomique entre deux régions cibles tel que défini ci-dessus, lesdites régions cibles étant différentes de celles définies à l'étape (i) ci-dessus,
- (iii) analyse phénotypique des organismes eucaryote génétiquement modifié obtenus aux étapes (i) et (ii),
- (iv) optionnellement, détermination des régions cibles d'intérêt pour une translocation chromosomique.

[0105] Dans un sixième aspect, la présente invention concerne également un procédé de croisement sexuel entre deux souches de champignon obtenues par un procédé de translocation chromosomique selon l'invention. Avantageusement, cela pourrait permettre de générer des caryotypes identiques et/ou obtenir des souches avec des propriétés améliorées (Chuang et al.). Plus particulièrement, ledit procédé et réalisé dans des conditions adaptées au croisement sexuel des champignons.

[0106] Selon un mode de réalisation, les souches ainsi croisées sont compatibles, c'est-à-dire que leur croisement permet de générer une nouvelle souche. Selon un mode de réalisation, la présente invention concerne un procédé de reproduction comprenant les étapes suivantes :

- (i) optionnellement l'obtention d'une première souche de champignon obtenue par un procédé de translocation chromosomique selon l'invention,
- (ii) optionnellement l'obtention d'une deuxième souche de champignon obtenue par un procédé de translocation chromosomique selon l'invention,
- (iii) un croisement entre deux souches obtenus par un procédé de translocation chromosomique entre deux régions cibles selon l'invention, notamment la première et la deuxième obtenues aux étapes (i) et (ii),
- (iv) optionnellement l'obtention d'une nouvelle souche.

[0107] Les définitions et modes de réalisation relatif au premier aspect de l'invention s'appliquent *mutatis mutandis* aux autres aspects. Par exemple, toutes les définitions et préférences indiquées dans le premier aspect de l'invention ci-dessus s'appliquent

également à tous les autres aspects de l'invention.

Références

- [0108] Chuang, YC., Li, WC., Chen, CL. et al. *Trichoderma reesei* meiosis generates segmentally aneuploid progeny with higher xylanase-producing capability. *Biotechnol Biofuels* 8, 30 (2015). <https://doi.org/10.1186/s13068-015-0202-6>
- [0109] Guo J, Gaj T, Barbas CF 3rd. Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. *J Mol Biol.* 2010 Jul 2;400(1):96-107. doi: 10.1016/j.jmb.2010.04.060. Epub 2010 May 4. PMID: 20447404; PMCID: PMC2885538.
- [0110] Jourdier, E., Baudry, L., Poggi-Parodi, D. et al. Proximity ligation scaffolding and comparison of two *Trichoderma reesei* strains genomes. *Biotechnol Biofuels* 10, 151 (2017). <https://doi.org/10.1186/s13068-017-0837-6>
- [0111] Ivanova, C., Ramoni, J., Aouam, T. et al. Genome sequencing and transcriptome analysis of *Trichoderma reesei* QM9978 strain reveals a distal chromosome translocation to be responsible for loss of *vib1* expression and loss of cellulase induction. *Biotechnol Biofuels* 10, 209 (2017). <https://doi.org/10.1186/s13068-017-0897-7>
- [0112] Le Crom S, Schackwitz W, Pennacchio L, Magnuson JK, Culley DE, Collett JR, Martin J, Druzhinina IS, Mathis H, Monot F, Seiboth B, Cherry B, Rey M, Berka R, Kubicek CP, Baker SE, Margeot A. Tracking the roots of cellulase hyperproduction by the fungus *Trichoderma reesei* using massively parallel DNA sequencing. *Proc Natl Acad Sci U S A.* 2009 Sep 22;106(38):16151-6. doi: 10.1073/pnas.0905848106. Epub 2009 Sep 2. PMID: 19805272; PMCID: PMC2752593.
- [0113] Tang, X.-X.; Wen, X.-P.; Qi, L.; Sui, Y.; Zhu, Y.-X.; Zheng, D.-Q. Origin, Regulation, and Fitness Effect of Chromosomal Rearrangements in the Yeast *Saccharomyces cerevisiae*. *Int. J. Mol. Sci.* 2021, 22, 786. <https://doi.org/10.3390/ijms22020786>
- [0114] Vitikainen, M., Arvas, M., Pakula, T. et al. Array comparative genomic hybridization analysis of *Trichoderma reesei* strains with enhanced cellulase production properties. *BMC Genomics* 11, 441 (2010). <https://doi.org/10.1186/1471-2164-11-441>
- [0115] Les séquences de la présente invention sont décrites dans le Tableau 1.

[0116] [Tableaux1]

Nom de la séquence	Séquence
SEQ ID NO : 1 Séquence de la cassette génétique ura3+ (la séquence codante du gène ura3 est en gras , les séquences promotrice (<i>en</i> <i>italique</i>) et terminatric e (<i>en</i> <i>italique</i> <u>souligné</u>) et les séquences flanquante s au gène cbh1 en italique gras)	CCAAGAACAATAGCCGATAAAGATAGCCTCATTAACGGAA TGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATATAAAGG TTCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCA TCAACTCAGATCCTCCAGGAGACTTGTACACCATCTTTTGA GGCACAGAAACCCAATAGTCAACCGCGGACTGCGCACATAC CTACCGACTGCATTGCATTGGTACAGCTGCAGGCACTTCCAGGCAC GGCCACCAAATTGCAGCGGCCCTTGCTTGCTTGCTTGGTTGAGAA CTAGGATCTGTGTCTTTTGCCTTGCCTTGTCTTGTCTGGGTTCTGC TCGTCTGCGGCAATCGGAACGCCGCCAGTGCGGTGCCAAGCAAAC CAGCCAAGGTAGGTACCTACCACTAGGCTTCTTTTCTCGTTGTCTCA CTCTCTCTTTTCTCTTTGTCCTCTCTTATCCCCATCTTTTCTCTCTCT CTCTGCTCCTTTCCCTAACCACTTCCCTACCTTTCTCTTTTCTTTTC TTGTCATCTCCATCTTGGCTGACGAAAAAGGTCTGACTGGGTAGGT ATTATCTGGCAGACTTGTGTGTATCATTACCCTATTTCTGCTTCATA GTACATGTACTGTACCTGAACGGCTCAACCGCTATTTACGACTCTTA TTTTTTTGTGGCGTTGGTCACGTTTGCCAGCTGTTGTCCGTCTTTCT AGGGCTCCTCAAACCTTGACCTGACCGAGCTCCCTTTCTGGACCCGG TGGGCTTCACTTCCAGCTGCTGAGCGACCTGAGCCGAACATCCTCA GTCTTGTCCAGCGCAATTCATTTTCTTTTCTTTTCTTTTTTTTATTC CTTTCTTTACTTTTATTCTCTCTTTTCTCCTCTTCCTCTTCTTCTTCTT TCTCCTCCTCCTCCATATCCTCACTCTCGTCTCCCTCATTACTACCC TCTCGGCTCCTCAGGTCCACCAACCCTCCCGCACCCAAACCTCTGC CGCTGAAACCCATTGGGTGGTCGCCGTTTTTTTTTTTTTTTTTCTC ACCCCCAAAGTCGCAATATCGGGTATCGCCGCCGGCATTGAATCGC CTTCTCCGCTAGCATCGACTACTGCTGCTCTGCTCTCGTTGCCAGC GCTGCTCCCTAGAATTTTGACCAGGGGACGAGCCCGACATTAAAGC AACTCCCTCGCCTCGAGACGACTCGGATCGCACGAAATTCTCCCAA TCGCCGACAGTTCCTACTCCTCTTCTCCTCCCGCACGGCTGTGCGCT TCCAACGTCATTGCGACAGCAGAATTGTGCCATCTCTCTTTTTTTT CCCCCCTCTAAACCGCCACAACGGCACCCCTAAGGGTTAAACTATC CAACCAGCCGCAGCCTCAGCCTCTCTCAGCCTCATCAGCCATGGC ACCACACCCGACGCTCAAGGCCACCTTCGCGGCCAGGAGC GAGACGGCGACGCACCCGCTGACGGCTTACCTGTTCAAGC TCATGGACCTCAAGGCGTCCAACCTGTGCCTGAGCGCCGA

CGTGCCGACAGCGCGGAGCTGCTGTACCTGGCCGACAAG
 ATTGGCCCGTCGATTGTCGTGCTCAAGACGCACTACGACAT
 GGTCTCGGGCTGGGACTTCCACCCGGAGACGGGCACGGGA
 GCCCAGCTGGCGTCGCTGGCGCGCAAGCACGGCTTCCTCA
 TCTTCGAGGACCGCAAGTTTGGCGACATTGGCCACACCGTC
 GAGCTGCAGTACACGGGCGGGTCGGCGCGCATCATCGACT
 GGGCGCACATTGTCAACGTCAACATGGTGCCCGGCAAGGC
 GTCGGTGGCCTCGCTGGCCCAGGGCGCCAAGCGCTGGCTC
 GAGCGCTACCCCTGCGAGGTCAAGACGTCCGTCACCGTCG
 GCACGCCCACCATGGACTCGTTTGACGACGACGCCGACTC
 CAGGGACGCCGAGCCCGCCGGCGCCGTCAACGGCATGGGC
 TCCATTGGCGTCCTGGACAAGCCCATCTACTCGAACC GGTC
 CGGCGACGGCCGCAAGGGCAGCATCGTCTCCATCACCACC
 GTCACCCAGCAGTACGAGTCCGTCTCCTCGCCCCGGTTAAC
 AAAGGCCATCGCCGAGGGCGACGAGTCGCTCTTCCCGGGC
 ATCGAGGAGGGCGCCGCTGAGCCGCGGCCTCCTGATCCTCG
 CCCAAATGTCCAGCCAGGGCAACTTCATGAACAAGGAGTAC
 ACGCAGGCCTGCGTCGAGGCCGCCCGGGAGCACAAGGACT
 TTGTCATGGGCTTCATCTCGCAGGAGACGCTCAACACCGAG
 CCCGACGATGCCTTTATCCACATGACGCCCGGCTGCCAGCT
 GCCCCCGAAGACGAGGACCAGCAGACCAACGGATCGGTC
 GGTGGAGACGGCCAGGGCCAGCAGTACAACACGCCGCACA
 AGCTGATTGGCATCGCCGGCAGCGACATTGCCATTGTGGG
 CCGGGGCATCCTCAAGGCCTCAGACCCCGTAGAGGAGGCA
 GAGCGGTACCGATCAGCAGCGTGGAAGCCTACACCGAGA
 GGCTGCTGCGATAGGGGGAGGGGAAGGGGAAGAAAGAAGTAAAGAA
AGGCATTTAGCAAGAAGGGGGGAAAAAGGGAGGGAGGACAAACGGAG
CTGAGAAAGAGCTCTTGTCCAAAGCCCCGGCATCATAGAATGCAGCT
GTATTTAGGCGACCTCTTTTTCCATCTTGTCGATTTTTGTTATGACGT
ACCAGTTGGGATGATGGATGATTGTACCCAGCTGCGATTGATGTG
TATCTTTGCATGCAACAACACGCGATGGCGGAGGGCGAACTGCACAT
TGGAAGGTTCATATATGGTCCTGACATATCTGGTGGATCTGGAAGC
ATGGAATTGTATTTTTGATTTGGCATTTGCTTTTGCGCGTGGAGGGA
ACATATCACCCCTCGGGCATTTTTCATTTGGTAGGATGGTTTGGATGC
AGTTGTCGACGATATCAGCTTCCATATATACGGAATTGTTCAATTGA
GTATTACCTGATGGCTATTTGTGTCGATCCTTTGTGACATTGACATG
TGATGGCGCCTGTACCTCTAAATGTATGCAGCGTAACATCTTCAGA
GCACGGAGCTCCGTGGCGAAAGCCTGACGCACCGGTAGATTCTTG

	<u>GTGAGCCCCGTATCATGACGGCGGGAGCTACATGGCCCCGGG</u> <u>TGATTATTTTTTTTGTATCTACTTCTGACCCTTTTCAAATATACGGTC</u> <u>AACTCATCTTTCACTGGAGATGCGGCCTGCTTGGTATTGCGATGTT</u> <u>GTCAGCTTGGCAAATTGTGGCT</u>
SEQ ID NO : 2 Séquence de l'ARN guide	GGA CCA GCA CAA CGG AU
SEQ ID NO : 3 Séquence de la matrice de réparation	ATCGCACGAAATTCTCCCAATCGCCGACAGTTCCTACTCCTCTT CCTCCCGCACGGCTGTCGCGCTTCCAACGTCATTGCGACAGCA GAATTGTGCCATCTCTCTCTTTTTTTTCCCCCCTCTAAACCGCC ACAACGGCACCCCTAAGGGTTAACTATCCAACCAGCCGCAGCC TCAGCCTCTCTCAGCCTCATCAGCCGGGAGGGAAGGGAAGAAA GAAGTAAAGAAAGGCATTTAGCAAGAAGGGGGAAAAGGGAGG GAGGACAAACGGAGCTGAGAAAGAGCTCTTGTCCAAAGCCCCG GCATCATAGAATGCAGCTGTATTTAGGCGACCTCTTTTTCCATC TTGTCGATTTTTGTTATGACGTACCAGTTGGGATGATGGATGAT TGTACCCAG
SEQ ID NO : 4 Primer 1	TGGCCTCGCTGGCTCGA
SEQ ID NO : 5 Primer 2	CAGCCGGGCGTCATGTGG
SEQ ID NO : 6 Primer 3	AGGAGCGAGACGGCGACG
SEQ ID NO : 7 Primer 4	CTCGAGCCAGCGCTTGGC
SEQ ID NO : 8 Primer 5	GCTCAGCCTTAGTAGGTAGG
SEQ ID NO : 9	TCTCGTTCAGCTCGTAATCG

Primer 6	
SEQ ID NO : 10 Primer 7	TTGTCACCCAGTTCGAGACG
SEQ ID NO : 11 Primer 8	CAGAGGCATTTCATCTGAAGCC
SEQ ID NO : 12 200pb en amont du gène CBH1 (région flanquante)	CCAAGAACAATAGCCGATAAAGATAGCCTCATTAACGGAA TGAGCTAGTAGGCAAAGTCAGCGAATGTGTATATATAAAGG TTCGAGGTCCGTGCCTCCCTCATGCTCTCCCCATCTACTCA TCAACTCAGATCCTCCAGGAGACTTGTACACCATCTTTTGA GGCACAGAAACCCAATAGTCAACCGCGGACTGCGCA
SEQ ID NO : 13 Promoteur natif du gène ura3	CATACCTACCGACTGCATTGCATTGGTACAGCTGCAGGCACTTCCA GGCACGGCCACCAAATTGCAGCGGCCCTTGCTTGCTTGCTTGGTTC GAGAACTAGGATCTGTGTCTTTTGCCTTGCCTTGTCTTGTCTGGGTT CCTGCTCGTCTGCGGCAATCGGAACGCCGCCAGTGCGGTGCCAAG CAAACCAGCCAAGGTAGGTACCTACCACTAGGCTTCTTTTCTCGTTG TCTCACTCTCTCTTTTCTCTTTGTCTCTCTTATCCCCATCTTTTCTC TCTCTCTCTGCTCCTTTTCTTAACCACTTCCCTACCTTTCTCTTTTCC TTTTCTTGTCATCTCCATCTTGGCTGACGAAAAAGGTCTGACTGGGT AGGTATTATCTGGCAGACTTGTGTGTATCATTACCCCTATTTCTGCTT CATAGTACATGTA CTGTACCTGAACGGCTCAACCGCTATTTACGACT CTTATTTTTTTGTGGCGTTGGTCACGTTTGCCAGCTGTTGTCCGTCT TTCTAGGGCTCCTCAAACCTTGACCTGACCGAGCTCCCTTTCTGGAC CCGGTGGGCTTCACTTCCAGCTGCTGAGCGACCTGAGCCGAACAT CCTCAGTCCTTGTCCAGCGCAATTCATTTTCTTTCTTTTCTTTTTTTT TATTCTTTCTTTACTTTTATTCTCTCTTTTTCTCCTCTTCCCTCTTCTTC TTCTTTCTCCTCCTCCTCCATATCCTCACTCTCGTCTCCCTCACTACT ACCTCTCGGCTCCTCAGGTCCACCAACCCTCCCGCACCCAAACCT CTGCCGCTGAAACCCATTCGGTGGTCGCCGTTTTTTTTTTTTTTTTT TCTACCCCCAAAGTCGCAATATCGGGTATCGCCGCCGGCATTGAA TCGCCTTCTCCGCTAGCATCGACTACTGCTGCTCTGCTCTCGTTGC CAGCGCTGCTCCCTAGAAATTTTGACCAGGGGACGAGCCCGACATTA

	AAGCAACTCCCTCGCCTCGAGACGACTCGGATCGCACGAAATTCTC CCAATCGCCGACAGTTCTACTCCTCTTCTCTCCCGCACGGCTGTGCG CGCTTCCAACGTCATTTCGCACAGCAGAATTGTGCCATCTCTCTCTTT TTTTCCCCCCTCTAAACCGCCACAACGGCACCCCTAAGGGTTAAA CTATCCAACCAGCCGCGAGCCTCAGCCTCTCTCAGCCTCATCAGCC
SEQ ID NO : 14 Gène ura3	<u>ATGGCACCAACCCGACGCTCAAGGCCACCTTCGCGGCCA</u> <u>GGAGCGAGACGGCGACGCACCCGCTGACGGCTTACCTGTT</u> <u>CAAGCTCATGGACCTCAAGGCGTCCAACCTGTGCCTGAGC</u> <u>GCCGACGTGCCGACAGCGCGCGAGCTGCTGTACCTGGCCG</u> <u>ACAAGATTGGCCCGTCGATTGTCTGTCTCAAGACGCACTAC</u> <u>GACATGGTCTCGGGCTGGGACTTCCACCCGAGACGGGCA</u> <u>CGGGAGCCCAGCTGGCGTCGCTGGCGCGCAAGCACGGCTT</u> <u>CCTCATCTTCGAGGACCGCAAGTTTGGCGACATTGGCCACA</u> <u>CCGTCGAGCTGCAGTACACGGGCGGGTCGGCGCGCATCAT</u> <u>CGACTGGGCGCACATTGTCAACGTCAACATGGTGCCCGGC</u> <u>AAGGCGTCGGTGGCCTCGCTGGCCCAGGGCGCCAAGCGCT</u> <u>GGCTCGAGCGCTACCCCTGCGAGGTCAAGACGTCCGTCAC</u> <u>CGTCGGCACGCCCACCATGGACTCGTTTGACGACGACGCC</u> <u>GACTCCAGGGACGCCGAGCCCGCCGGCGCCGTCAACGGCA</u> <u>TGGGCTCCATTGGCGTCCTGGACAAGCCCATCTACTCGAAC</u> <u>CGGTCCGGCGACGGCCGCAAGGGCAGCATCGTCTCCATCA</u> <u>CCACCGTCACCCAGCAGTACGAGTCCGTCTCCTCGCCCCGG</u> <u>TTAACAAGGCCATCGCCGAGGGCGACGAGTCGCTCTTCC</u> <u>CGGGCATCGAGGAGGCGCCGCTGAGCCGCGGCCTCCTGAT</u> <u>CCTCGCCCCAAATGTCCAGCCAGGGCAACTTCATGAACAAGG</u> <u>AGTACACGCAGGCCTGCGTCGAGGCGCCCGGGAGCACAA</u> <u>GGACTTTGTCATGGGCTTCATCTCGCAGGAGACGCTCAACA</u> CCGAGCCCGACGATGCCTTTATCCACATGACGCCCGGCTGC CAGCTGCCCCCGAAGACGAGGACCAGCAGACCAACGGAT CGGTTCGGTGGAGACGGCCAGGGCCAGCAGTACAACACGCC GCACAAGCTGATTGGCATCGCCGGCAGCGACATTGCCATT GTGGGCGGGGCATCCTCAAGGCCTCAGACCCCGTAGAGG AGGCAGAGCGGTACCGATCAGCAGCGTGGAAGCCTACAC CGAGAGGCTGCTGCGATAG
SEQ ID NO : 15 Terminate	<u>GGGAGGGAAGGGAAGAAAGAAGTAAAGAAAGGCATTAGCAAGAA</u> <u>GGGGGAAAAGGGAGGGAGGACAAACGGAGCTGAGAAAGAGCTCTT</u> <u>GTCCAAAGCCCGGCATCATAGAATGCAGCTGTATTAGGCGACCTC</u>

ur natif d'ura3	<u>TTTTCCATCTTGTCGATTTTTGTTATGACGTACCAGTTGGGATGATG</u> <u>GATGATTGTACCCCAGCTGCGATTGATGTGTATCTTTGCATGCAACA</u> <u>ACACGCGATGGCGGAGGGCGAACTGCACATTGGAAGGTTTCATATATG</u> <u>GTCTTGACATATCTGGTGGATCTGGAAGCATGGAATTGTATTTTTGA</u> <u>TTTGGCATTGCTTTTGCGCGTGGAGGGGAACATATCACCTCGGGC</u> <u>ATTTTTCATTGTTAGGATGGTTTGGATGCAGTTGTCGACGATATCA</u> <u>GCTTCCATATATACGGAATTGTTCAATTGAGTATTACCTGATGGCTAT</u> <u>TTGTGTCGATCCTTTGTGACATTGACATGTGATGGCGCCTGTACCTC</u> <u>TAAATGTATGCAGCGTAACATCTTCAGAGCACGG</u>
SEQ ID NO : 16 200pb en aval du gène CBH1 (région flanquante)	AGCTCCGTGGCGAAAGCCTGACGCACCGGTAGATTCTTGG TGAGCCCGTATCATGACGGCGGGAGCTACATGGCCCC GGGTGATTTATTTTTTTTGTATCTACTTCTGACCCTTTTCAA ATATACGGTCAACTCATCTTTCCTGGAGATGCGGCCTGCT TGGTATTGCGATGTTGTCAGCTTGGCAAATTGTGGCT
SEQ ID NO : 17 Région flanquante, 200pb en amont du gène ura3	ATCGCACGAAATTCTCCCAATCGCCGACAGTTCCTACTCCTCTT CCTCCCGCACGGCTGTGCGGCTTCCAACGTCATTGCGACAGCA GAATTGTGCCATCTCTCTCTTTTTTTTCCCCCCTCTAAACCGCC ACAACGGCACCCCTAAGGGTTAACTATCCAACCAGCCGCAGCC TCAGCCTCTCTCAGCCTCATCAGCC
SEQ ID NO : 18 Région flanquante, 200pb en aval du gène ura3	GGGAGGGAAGGGAAGAAAGAAGTAAAGAAAGGCATTTAGCAA GAAGGGGGAAAAGGGAGGGAGGACAAACGGAGCTGAGAAAG AGCTCTTGTCCAAAGCCCGGCATCATAGAATGCAGCTGTATTT AGGCGACCTCTTTTCCATCTTGTCGATTTTGTATGACGTACC AGTTGGGATGATGGATGATTGTACCCAG

[0117] Séquences de l'invention

[0118] D'autres caractéristiques, détails et avantages de l'invention apparaîtront à la lecture des Figures annexées et des exemples qui visent uniquement à illustrer l'invention, sans la limiter.

Brève description des Figures

Fig. 1

[0119] [Fig.1] représente un schéma d'un exemple de cassette contenant le site de reconnaissance (SR) qui est intégré au préalable dans le génome de l'organisme hôte avant l'action de la nucléase. Ledit site de reconnaissance est intégré au sein d'un rapporteur d'intégration, par exemple un gène modifiant l'auxotrophie de l'organisme. Les termes « 1/2 marqueur » et « 2/2 marqueur » représentent le même marqueur, la séquence étant scindée en deux par l'insertion de la séquence du site de reconnaissance. Dans un tel cas, le gène n'est plus fonctionnel et permet de sélectionner des organismes contenant le site de reconnaissance de la nucléase, par comparaison avec des organismes qui ne contiennent pas la séquence du site de reconnaissance de la nucléase aux locus souhaités et qui possèdent une copie fonctionnelle du gène. Les deux chromosomes ici ciblés sont nommés X et Y, qui doivent simplement être compris comme chromosomes 1 et 2, et non pas comme des chromosomes sexuels chez l'humain. Les cibles Kx et Ky, en amont et en aval du site de reconnaissance correspondent aux régions flanquantes en 5' et 3' pour insérer le site de reconnaissance de la nucléase aux sites chromosomiques souhaitée, respectivement sur les chromosomes X et Y.

Fig. 2

[0120] [Fig.2] représente deux chromosomes, dans le cas hors invention (à gauche) et dans le cas de l'invention (à droite). Le site de reconnaissance (locus cible) de la nucléase Cas9 est indiqué. Chacun des deux chromosomes est porteur de deux sites de reconnaissance. Dans le cas hors-invention, les quatre locus cibles sont différents, de même que les ARN guide associés à la nucléase (le locus cible 1 sera coupé à l'aide la nucléase Cas9 associé à l'ARN guide 1, le locus cible 2 sera coupé à l'aide la nucléase Cas9 associé à l'ARN guide 2, ...). Au contraire, dans le cas de l'invention, les quatre locus cibles sont identiques, de même que les ARN guide associés à la nucléase : tous les locus cibles seront coupés à l'aide la nucléase Cas9 associé à l'ARN guide 1.

Fig. 3

[0121] [Fig.3] représente une cassette génétique pour l'insertion d'un gène ura3 fonctionnel au site du gène cbh1. En haut, la cassette génétique est représentée, et en bas le site d'insertion localisé sur le chromosome VII (CHRVII).

Fig. 4

[0122] [Fig.4] représente la localisation chromosomique des deux versions du gène ura3 . La version mutée et native ura3- est localisée sur le chromosome V (CHRV) et la version fonctionnelle et additionnelle ura3+ est localisée sur le chromosome VII (CHRVII).

Fig. 5

[0123] [Fig.5] représente une matrice de réparation pour l'excision de la partie codante du gène *ura3*. Elle est composée de 180 pb en 5' du gène et de 180 pb en 3' du gène.

Fig. 6

[0124] [Fig.6] représente un schéma des amplifications PCR pour l'identification de la présence ou non des gènes *ura3* et la présence ou non d'un brassage chromosomique. CHRV représente le chromosome V, CHRVII représente le chromosome VII, pURA représente un promoteur pour le gène *ura3*, *ura* natif représente le gène *ura3*, tUra représente le terminateur pour le gène *ura3*.

Fig. 7

[0125] [Fig.7] représente la translocation dirigée obtenue entre les extrémités des chromosomes V et VII (CHRV et CHRVII). Les régions promotrices et terminatrices sont les restes de l'excision des gènes *ura3*⁻ et *ura3*⁺. Avant translocation, un chromosome est uni, l'autre hachuré, et après translocation, un réarrangement entre une région unie et une région hachurée est obtenu.

[0126] Exemples

[0127] **Exemple 1 : Translocation à l'aide d'un seul site de reconnaissance répété au sein du génome**

[0128] L'étape 1 vise à identifier un site de reconnaissance de la nucléase, optionnellement vérifier si celui-ci est déjà présent une fois dans le génome sur un chromosome.

[0129] L'étape 2 vise à construire au moins une cassette pour insérer le site de reconnaissance de la nucléase dans le génome sur un autre chromosome ou au sein du même chromosome.

[0130] Lorsque le site de reconnaissance de la nucléase est présent une fois sur un chromosome, une seule cassette peut être utilisée pour insérer le même site de reconnaissance sur le même ou un autre chromosome.

[0131] Au contraire, lorsque le site de reconnaissance de la nucléase n'est pas présent naturellement sur un chromosome, au moins deux cassettes peuvent être utilisées pour insérer le même site de reconnaissance sur le même ou deux chromosomes différents. Dans un tel cas les cassettes utilisées sont identiques, sauf en ce qui concerne les régions flanquantes qui diffèrent pour cibler l'intégration du site de reconnaissance aux locus souhaités sur le chromosome ou sur chacun des deux chromosomes. La [Fig.1] illustre le cas où il y a des sites de reconnaissance de la nucléase sur deux chromosomes.

[0132] La [Fig.2] (à droite) représente l'exemple où chacun des deux chromosomes, est porteur de deux sites de reconnaissance. La même nucléase peut ainsi couper sur chacun de ces quatre sites et un brassage chromosomique peut être initié entre les fragments chromosomiques ainsi obtenus.

- [0133] Lorsqu'un au moins deux sites de reconnaissance de la même nucléase est présent sur au moins deux chromosomes différents, le brassage chromosomique peut être initié, par transformation de l'organisme avec la même nucléase adapté au site de reconnaissance.
- [0134] **Exemple 2 : Translocation dirigée obtenue à partir d'une souche CL847 (ura3+)**
- [0135] L'exemple décrit ici une translocation dirigée obtenue à partir d'une souche dérivée de la souche CL847 (CL847-ura3+) possédant deux versions du gène ura3, une version fonctionnelle (ura3+) et une version non fonctionnelle (ura3-). Le gène ura3 est ici le marqueur de sélection.
- [0136] **1. Construction de la souche CL847-ura3+**
- [0137] Il est possible de sélectionner des souches auxotrophes pour l'uracile en utilisant un analogue de pyrimidine, l'acide 5-fluoro-orotique (5FOA) dans le milieu. En effet, les souches présentant une copie sauvage des gènes URA3 (ura3+) ou URA5 (ura5+) ne peuvent pas pousser sur un tel milieu tandis que les souches mutantes ura3- ou ura5- pousseront normalement. La toxicité du 5FOA vient de sa conversion en 5-fluoro-UMP par l'OMP pyrophosphorylase codé par le gène URA5 ou/et par l'OMP decarboxylase codé par le gène URA5.
- [0138] La souche dérivée de la souche CL847 est un mutant spontané sélectionné sur le milieu MM additionné de 5FOA et présentant une mutation dans le gène ura3 (au niveau du chromosome V). Cette mutation est une délétion de 20 paires de bases.
- [0139] L'auxotrophie pour l'uracile de cette souche a été complétée par réinsertion d'une cassette génétique contenant le gène ura3 (SEQ ID NO : 14) avec 1217 pb en amont de l'ATG (région promotrice, SEQ ID NO : 13) et 546 pb en aval du codon stop (région terminatrice, SEQ ID NO : 15) et 200 pb des régions flanquantes du gène cbh1 (en amont représentée par SEQ ID NO : 12, et en aval représentée par SEQ ID NO : 16). La séquence de la cassette est représentée en SEQ ID NO : 1, et la représentation d'une telle construction est présentée en [Fig.3].
- [0140] L'insertion de la cassette génétique au site du gène cbh1 a été vérifié par PCR. La localisation des deux gènes ura3 (ura3+ et ura3-) est représentée sur la [Fig.4].
- [0141] **2. Transformation de la souche CL847-ura3+**
- [0142] La souche CL847-ura3+ a été transformée avec une matrice de réparation et un complexe ribonucléoprotéique composée d'un ARN guide et de l'endonucléase Cas9. Les séquences de la matrice de réparation et de l'ARN guide sont indiquées dans les SEQ ID NO : 3 et SEQ ID NO : 2.
- [0143] Un schéma de la matrice de réparation est représenté en [Fig.5].
- [0144] Toute méthode de transformation peut être utilisée, par exemple celle décrite dans la demande internationale PCT/FR2016/053601. Les transformants ont été sélectionnés sur un milieu contenant du 5FOA sur lequel seules les souches auxotrophes pour

l'uracile peuvent pousser.

- [0145] Les souches ainsi obtenues ont été purifiées et l'élimination des deux gènes *ura3* a été vérifiée par PCR. Un schéma des amplifications PCR pour l'identification de la présence ou non des gènes *ura3* et la présence ou non d'un brassage chromosomique est représentée en [Fig.6]. Sur cette [Fig.6], les 2 paires d'oligonucléotides (primer 1-2) et (primer 3-4) couplées à une PCR d'un gène rapporteur permettent de vérifier la présence ou l'absence des gènes *ura3* de ladite souche CL847. Les deux paires d'oligonucléotides (primer 5-6) et (primer 7-8) permettent de vérifier la bonne excision des deux gènes *ura3* et le bon rapprochement entre les deux brins chromosomiques frères. Les deux paires d'oligonucléotides (6-7) et (5-8) permettent de vérifier la bonne excision des deux gènes *ura3* et d'identifier un brassage chromosomique.
- [0146] Comme indiqué sur la [Fig.6], après coupure des deux chromosomes (V et VII) au niveau du site de reconnaissance de la nucléase, présent sur chacun des deux chromosomes, quatre fragments chromosomiques sont obtenus : un fragment 5' du chromosome V, un fragment 3' du chromosome V, un fragment 5' du chromosome VII, et un fragment 3' du chromosome VII. Différents cas peuvent alors se produire :
- lorsque les fragments 5' et 3' du chromosome V se rapprochent, il n'y a pas de brassage chromosomique,
 - lorsque les fragments 5' et 3' du chromosome VII se rapprochent, il n'y a pas de brassage chromosomique,
 - lorsque le fragment 5' et du chromosome V et le fragment 3' du chromosome VII se rapprochent, il y a un brassage chromosomique,
 - lorsque le fragment 3' et du chromosome V et le fragment 5' du chromosome VII se rapprochent, il y a un brassage chromosomique.
- [0147] **3. Identification de la translocation dirigée**
- [0148] L'ADN des souches présentant le phénotype et génotype attendus a été extrait et séquencé en paired end 2X150pb. L'alignement des lectures a été effectué sur le génome de référence QM6a en utilisant l'outil d'alignement Geneious (Geneious Prime® 2022.2.1) qui permet aussi d'identifier les variations structurales.
- [0149] Une souche présentant une translocation entre les deux sites d'insertion a été sélectionnée ([Fig.7]). La translocation a été vérifiée par PCR.
- [0150] Le milieu de culture MM additionné de 5FOA utilisé dans cet exemple est le suivant :

[0151] [Tableaux2]

	Pour 1 litre
KH ₂ PO ₄	5g
(NH ₄)SO ₄	5g
Tris-Sodium Citrate	11.7 g
Glucose	20 g
Agar	20 g
H ₂ O (eau distillée)	Qsp 1l

[0152] Composition du milieu MM agar + 5FOA

[0153] Autoclave à 20 minutes à 120°C, puis après l'autoclave, addition de :

[0154] [Tableaux3]

	Solution stock	Pour 1 litre
MgSO ₄	200X (120g/l)	5 ml
CaCl ₂	200X (120g/l)	5 ml
Éléments Traces	1000X	1 ml
Uridine	100X (1M : 24,4g dans 100mL)	10 ml
5'FOA	-	1,5 mg

[0155] Ajout à la composition du milieu MM agar + 5FOA après autoclave

Revendications

- [Revendication 1] Procédé de translocation chromosomique entre au moins deux régions cibles chez un organisme eucaryote, ledit procédé comprenant, au moins, une étape de coupure du double brin d'ADN sur le ou les chromosome(s) impliqué(s) dans la translocation au niveau du site de reconnaissance d'une nucléase, et ledit procédé étant caractérisé en ce que chaque région cible comprend le même site de reconnaissance de ladite nucléase, sous réserve que ledit organisme eucaryote n'est pas un organisme humain, un embryon humain, une cellule de lignée germinale humaine ou une cellule embryonnaire humaine.
- [Revendication 2] Procédé de translocation chromosomique selon la revendication 1, caractérisé en ce que l'étape de coupure du double brin d'ADN est :
- une étape de coupure sur chaque chromosome impliqué dans la translocation au niveau du site de reconnaissance d'une nucléase, ou
 - une étape de coupure dans le même chromosome au niveau du site de reconnaissance d'une nucléase.
- [Revendication 3] Procédé de translocation chromosomique selon l'une quelconque des revendications précédentes, caractérisé en ce que l'étape de coupure du double brin d'ADN est réalisée par une cassure double brin ou deux cassures simple brin.
- [Revendication 4] Procédé de translocation chromosomique selon l'une quelconque des revendications précédentes, caractérisé en ce que l'organisme eucaryote est un champignon, notamment un champignon filamenteux, notamment appartenant au genre *Trichoderma*.
- [Revendication 5] Procédé de translocation chromosomique selon l'une quelconque des revendications précédentes, caractérisé en ce que la nucléase est choisie parmi une méganucléase, une nucléase à doigt de zinc, une nucléase TALE, une nickase, et une nucléase Cas, de préférence CRISPR/Cas9 ou CRISPR/Cas12.
- [Revendication 6] Procédé de translocation chromosomique selon l'une quelconque des revendications précédentes, caractérisé en ce qu'il comprend les étapes suivantes :
- optionnellement, l'identification d'un site de reconnaissance d'une nucléase,
 - l'insertion dans le génome d'un organisme, notamment un champignon, d'au moins un site de reconnaissance de ladite nucléase,

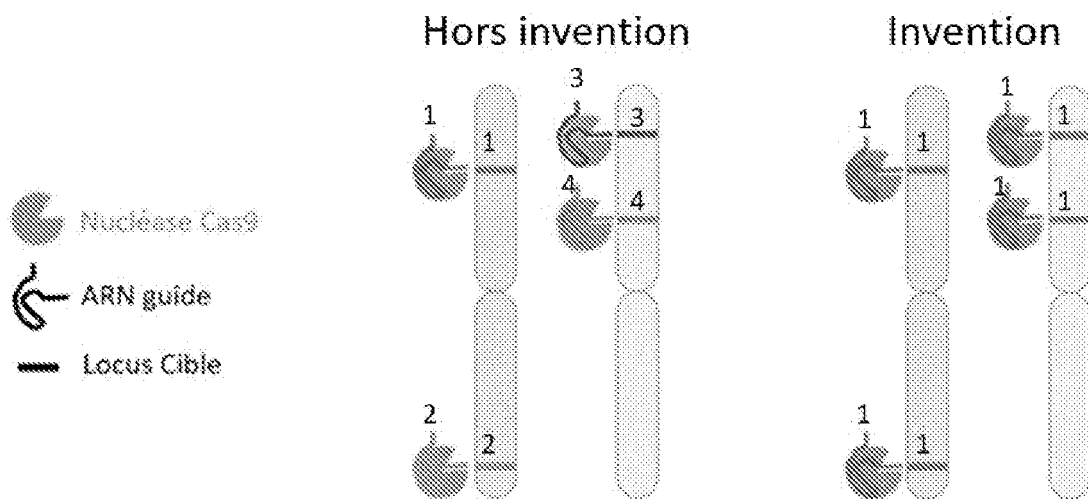
- la sélection d'un organisme issu de l'étape précédente comprenant le site de reconnaissance de ladite nucléase dans chaque région cible,
 - l'introduction dans l'organisme issu de l'étape précédente de ladite nucléase ou d'un vecteur comprenant au moins un acide nucléique codant pour ladite nucléase,
 - optionnellement, la vérification de la translocation chromosomique.
- [Revendication 7] Organisme eucaryote, notamment une souche de champignon, transformé par le procédé de translocation chromosomique selon l'une quelconque des revendications précédentes, ou la descendance d'un tel organisme, sous réserve que ledit organisme eucaryote n'est pas un organisme humain, un embryon humain, une cellule de lignée germinale humaine ou une cellule embryonnaire humaine.
- [Revendication 8] Organisme eucaryote, notamment une souche de champignon, comprenant au moins deux fois le même site de reconnaissance d'une nucléase dans son génome, notamment dans au moins deux chromosomes différents, sous réserve que ledit organisme eucaryote n'est pas un organisme humain, un embryon humain, une cellule de lignée germinale humaine ou une cellule embryonnaire humaine.
- [Revendication 9] Utilisation d'un organisme, notamment une souche de champignon, selon l'une quelconque des revendications 7-8, pour la production de molécules ou de protéines d'intérêt, notamment des enzymes cellulolytiques ou des protéines d'intérêt d'origine hétérologue.
- [Revendication 10] Utilisation d'un organisme, notamment une souche de champignon, selon l'une quelconque des revendications 7-8, pour l'hydrolyse de la cellulose ou de la lignocellulose en sucres C5 et/ou C6 et/ou pour la production de produits biosourcés ou de biocarburant à partir de substrats cellulotiques ou lignocellulosiques.
- [Revendication 11] Utilisation du même site de reconnaissance d'une nucléase sur un ou plusieurs chromosomes pour effectuer une translocation chromosomique avec ladite nucléase, sous réserve que l'utilisation n'est pas sur un organisme humain, un embryon humain, une cellule de lignée germinale humaine ou une cellule embryonnaire humaine.
- [Revendication 12] Procédé de reproduction entre deux souches comprenant une étape de croisement entre deux souches de champignon obtenues par un procédé de translocation chromosomique selon l'une quelconque des reven-

- dications 1-6, et optionnellement l'obtention d'une nouvelle souche.
- [Revendication 13] Procédé pour identifier le site d'insertion d'un gène dans un organisme eucaryote selon la revendication 8, notamment un champignon, ledit procédé comprenant au moins les étapes suivantes :
- obtention d'un organisme eucaryote dans lequel un gène est inséré de manière ectopique, ledit gène comprenant un fragment de séquence similaire à une autre séquence du génome de l'organisme,
 - réalisation d'une translocation chromosomique entre ledit fragment et la deuxième séquence possédant une similarité avec ledit fragment,
 - analyse génomique dudit organisme eucaryote, après la translocation chromosomique, pour identifier le site d'insertion dudit gène.

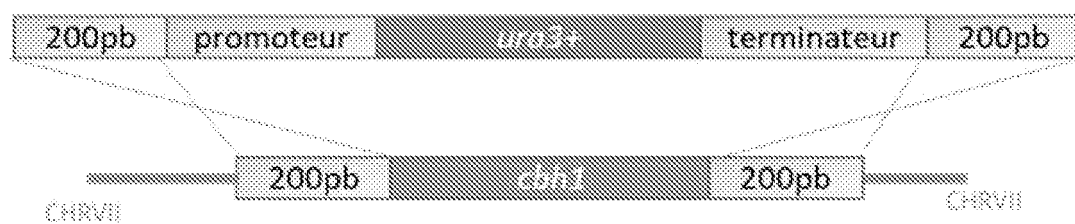
[Fig. 1]



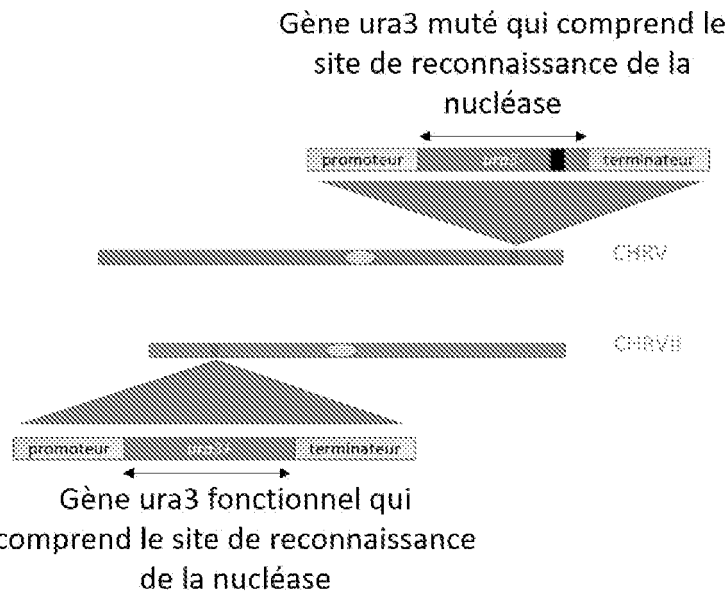
[Fig. 2]



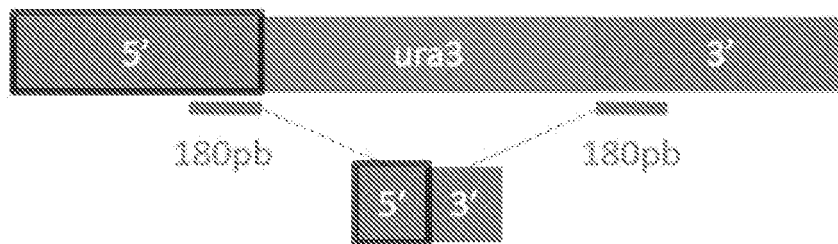
[Fig. 3]



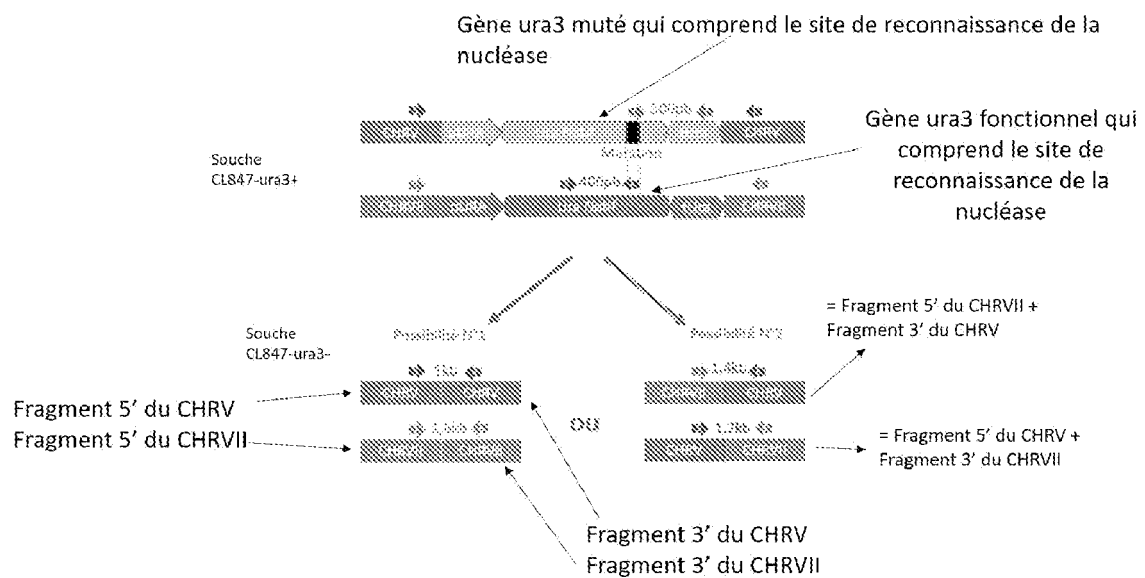
[Fig. 4]



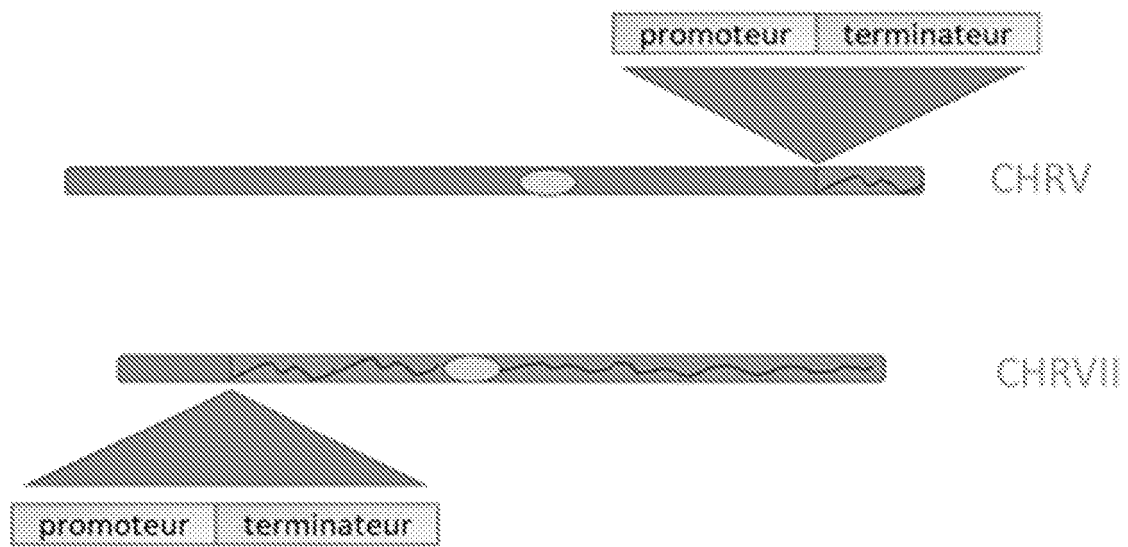
[Fig. 5]



[Fig. 6]



[Fig. 7]



RAPPORT DE RECHERCHE PRÉLIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 919515
FR 2212597

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
X	TANG, X.-X.WEN, X.-P.QI, L.SUI, Y.ZHU, Y.-X.ZHENG, D.-Q: "Origin, Re-gulation, and Fitness Effect of Chromosomal Rearrangements in the Yeast Saccharomyces cerevisiae.", Int. J. Mol. Sci., vol. 22 2021, page 786, XP002810197, Extrait de l'Internet: URL:https://www.ncbi.nlm.nih.gov/pmc/artic les/PMC7830279/pdf/ijms-22-00786.pdf * le document en entier *	7, 8	DOMAINES TECHNIQUES RECHERCHÉS (IPC)
X	VITIKAINEN MARIKA ET AL: "Array comparative genomic hybridization analysis of Trichoderma reesei strains with enhanced cellulase production properties", BMC GENOMICS, BIOMED CENTRAL LTD, LONDON, UK, vol. 11, no. 1, 19 juillet 2010 (2010-07-19), page 441, XP021072738, ISSN: 1471-2164, DOI: 10.1186/1471-2164-11-441 * le document en entier *	7, 8	
Y	----- -/--	9, 10	
Date d'achèvement de la recherche		Examineur	
28 septembre 2023		Seroz, Thierry	
CATÉGORIE DES DOCUMENTS CITÉS X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant			

RAPPORT DE RECHERCHE PRÉLIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 919515
FR 2212597

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
Y	CHEN YUMENG ET AL: "Engineering of Trichoderma reesei for enhanced degradation of lignocellulosic biomass by truncation of the cellulase activator ACE3", BIOTECHNOLOGY FOR BIOFUELS, vol. 13, no. 1, 1 avril 2020 (2020-04-01), XP055976462, DOI: 10.1186/s13068-020-01701-3 Extrait de l'Internet: URL: http://link.springer.com/article/10.1186/s13068-020-01701-3/fulltext.html > * le document en entier * -----	9,10	DOMAINES TECHNIQUES RECHERCHÉS (IPC)
Y	FONSECA LUCAS MIRANDA ET AL: "Rational engineering of the Trichoderma reesei RUT-C30 strain into an industrially relevant platform for cellulase production", BIOTECHNOLOGY FOR BIOFUELS, vol. 13, no. 1, 22 mai 2020 (2020-05-22), XP093037486, DOI: 10.1186/s13068-020-01732-w Extrait de l'Internet: URL: https://link.springer.com/article/10.1186/s13068-020-01732-w/fulltext.html > * le document en entier * -----	9,10	
Y	ALEXANDER V GUSAKOV: "Alternatives to biofuel production", TRENDS IN BIOTECHNOLOGY, vol. 29, no. 9, 18 avril 2011 (2011-04-18) , pages 419-425, XP028383328, ISSN: 0167-7799, DOI: 10.1016/J.TIBTECH.2011.04.004 [extrait le 2011-04-18] * le document en entier * -----	9,10	
Date d'achèvement de la recherche Examineur 28 septembre 2023 Seroz, Thierry			
CATÉGORIE DES DOCUMENTS CITÉS X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant			

1

EPO FORM 1503 12.99 (P04C14)

RAPPORT DE RECHERCHE PRÉLIMINAIRE

établi sur la base des dernières revendications
déposées avant le commencement de la recherche

N° d'enregistrement
national

FA 919515
FR 2212597

DOCUMENTS CONSIDÉRÉS COMME PERTINENTS		Revendication(s) concernée(s)	Classement attribué à l'invention par l'INPI
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes		
A	US 2022/340935 A1 (ZHENG QI [US] ET AL) 27 octobre 2022 (2022-10-27) -----	1-13	DOMAINES TECHNIQUES RECHERCHÉS (IPC)
Date d'achèvement de la recherche		Examineur	
28 septembre 2023		Seroz, Thierry	
CATÉGORIE DES DOCUMENTS CITÉS X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : arrière-plan technologique O : divulgation non-écrite P : document intercalaire T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons & : membre de la même famille, document correspondant			

ANNEXE AU RAPPORT DE RECHERCHE PRÉLIMINAIRE **RELATIF A LA DEMANDE DE BREVET FRANÇAIS NO. FR 2212597 FA 919515**

La présente annexe indique les membres de la famille de brevets relatifs aux documents brevets cités dans le rapport de recherche préliminaire visé ci-dessus.
 Les dits membres sont contenus au fichier informatique de l'Office européen des brevets à la date du **28-09-2023**
 Les renseignements fournis sont donnés à titre indicatif et n'engagent pas la responsabilité de l'Office européen des brevets, ni de l'Administration française

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
CN 104726494 B	23-10-2018	AUCUN	
US 2022340935 A1	27-10-2022	CN 114008201 A	01-02-2022
		US 2022340935 A1	27-10-2022
		WO 2020257545 A1	24-12-2020