

(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
19 février 2004 (19.02.2004)

PCT

(10) Numéro de publication internationale
WO 2004/015134 A2

- (51) Classification internationale des brevets⁷ : C12Q 1/00 (74) Mandataires : BECKER, Philippe etc.; Cabinet Becker et Associés, 35, rue des Mathurins, F-75008 Paris (FR).
- (21) Numéro de la demande internationale : PCT/FR2003/002476 (81) États désignés (national) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) Date de dépôt international : 6 août 2003 (06.08.2003)
- (25) Langue de dépôt : français
- (26) Langue de publication : français
- (30) Données relatives à la priorité : 02/10048 7 août 2002 (07.08.2002) FR (84) États désignés (régional) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Déposant (pour tous les États désignés sauf US) : CEL-LECTIS [FR/FR]; 28, rue du Docteur Roux, F-75015 Paris (FR).
- (72) Inventeurs; et
- (75) Inventeurs/Déposants (pour US seulement) : ARNOULD, Sylvain [FR/FR]; 116, rue Lamarck, F-75018 Paris (FR). CHAMES, Patrick [FR/FR]; 123, rue du Château, F-75014 Paris (FR). CHOULIKA, André [FR/FR]; 164, rue de la Croix Nivert, F-75015 Paris (FR). LACROIX, Emmanuel [FR/FR]; 15, avenue Laumière, F-75015 Paris (FR).
- Publiée :
— sans rapport de recherche internationale, sera republiée dès réception de ce rapport
- En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(54) Title: METHOD OF DETERMINING THE CLEAVAGE ACTIVITY OF AN ENDONUCLEASE

(54) Titre : METHODE DE DETERMINATION DE L'ACTIVITE DE CLIVAGE D'UNE ENDONUCLEASE

(57) Abstract: The invention relates to a method of testing the cleavage activity of an endonuclease in relation to a DNA substrate. The inventive method is based on the discrimination of the cleaved DNA substrate and the intact substrate. The aforementioned discrimination can be performed by means of selective degradation or depletion of the intact substrate or by selective amplification of the cleaved substrate. Said method can be applied to the screening and the selection of an endonuclease or a cleavage site for an endonuclease.

(57) Abrégé : L'invention porte sur une méthode permettant de tester l'activité de coupure par une endonucléase d'un substrat AND. Cette méthode est basée sur la discrimination du substrat AND coupé et du substrat non coupé. La discrimination peut être faite par dégradation sélective ou déplétion du substrat non coupé, ou par amplification sélective du substrat coupé. Cette méthode peut être appliquée au criblage et à la sélection soit d'une endonucléase soit d'un site de coupure pour une endonucléase.

WO 2004/015134 A2

METHODE DE DETERMINATION DE L'ACTIVITE DE CLIVAGE D'UNE ENDONUCLEASE

L'invention porte sur des méthodes et compositions permettant de tester l'activité de
5 coupure d'une endonucléase vis-à-vis d'un substrat. Les méthodes de l'invention sont basées sur la
conception de constructions génétiques particulières et/ou sur des conditions et méthodes
avantageuses et efficaces de discrimination entre les substrats coupés et les substrats non coupés..
L'invention concerne également des outils et kits utilisables pour la mise en œuvre de ces
méthodes, ainsi que leurs utilisations. L'invention peut être appliquée au criblage, à la sélection
10 ou à l'amélioration soit d'endonucléases, soit de sites de coupure pour une endonucléase.

Les endonucléases sont des enzymes indispensables en biologie moléculaire et très utiles
pour l'ingénierie des génomes. Les endonucléases les plus connues sont les enzymes de restriction.
Les « homing endonucleases » constituent une autre classe d'endonucléases d'intérêt en raison de
15 leur large site de reconnaissance (Dalgaard et al, 1997, *Nucleic Acids Research*, **25**, 4626-463 ;
Chevalier et Stoddard, 2001, *Nucleic Acids Research*, **29**, 3757-3774).

L'avantage de ces endonucléases est leur spécificité de reconnaissance et de coupure de
l'ADN. Il existe une forte demande pour créer des endonucléases présentant de nouvelles
spécificités de séquence nucléotidique. Cette demande est particulièrement forte pour les
20 endonucléases présentant un site de coupure rare. Certaines nouvelles endonucléases ont été créées
soit par la modification des endonucléases existantes (Wenz et al, 1994, *Biochim Biophys Acta*,
1219, 73-80), soit par la génération d'endonucléases chimériques (par exemple, une tête
endonucléasique d'une endonucléase de restriction de type IIS fusionnée à une association de
plusieurs doigts de zinc; Smith et al, 2000, *Nucleic Acids Res*, **28**, 3361-3369 ; Kim et al, 1996,
25 *Proc Natl Acad Sci USA*, **93**, 1156-1160 ; Kim & Chandrasegaran, 1994, *Proc Natl Acad Sci USA*,
91, 883-887 ; WO 95/09233 ; WO 94/18313). Ces manipulations nécessitent de disposer de tests
de clivage pour estimer l'activité et la spécificité de la nouvelle protéine ou d'un ensemble de
protéines. Cependant, à ce jour, il n'existe pas de méthodes de sélection *in vitro* basée sur l'activité
et la spécificité de clivage.

30 Le grand intérêt des endonucléases justifie la nécessité du développement de méthodes
permettant de tester l'activité et la spécificité de clivage avec un haut rendement et/ou adapté pour
effectuer une sélection.

Les méthodes existantes de mesure de l'activité de clivage et de sa spécificité ne sont pas
adaptées à l'exigence de haut rendement pour les raisons exposées ci-dessous.

35 Un test est fait en incubant une macromolécule d'ADN avec une endonucléase pendant
une période de temps déterminée. La distribution des fragments d'ADN clivés et non-clivés est

analysée par une électrophorèse en gel d'agarose (Yolov et al., 1985, *Nucleic Acids Res*, **13**, 8983-8998). Ce test est l'un des plus anciens encore utilisé. Il est le plus commun pour l'analyse du clivage d'ADN par des endonucléases car il est peu coûteux, ne nécessite aucun matériel de haute technologie et est facile à mettre en oeuvre.

5 Une mesure de la coupure peut également être effectuée par plusieurs autres méthodes comme la chromatographie en couche fine (Jay et al., 1974, *Nucleic Acids Res*, **1**, 331-353), l'élution des produits de la réaction sur filtre de DEAE cellulose (McLaughlin et al., 1987, *Biochemistry*, **26**, 7238-7245), ou la chromatographie HPLC en mesurant l'absorbance dans l'UV (Alves et al., 1984, *Eur J Biochem*, **140**, 83-92 ; Newman et al., 1990, *Biochemistry*, **29**, 9891-10 9901 ; Newman et al., 1990, *Biochemistry*, **29**, 9902-9910). Ces méthodes permettent de mesurer la disparition du substrat ou l'apparition des produits.

Cependant tous ces tests présentent le désavantage d'être à la fois laborieux et long à mettre en oeuvre. Ils ne sont pas adaptés au haut rendement. Un désavantage supplémentaire de ces méthodes est que le niveau de sensibilité nécessite l'utilisation soit d'ADN radioactivement 15 marqué soit de fortes concentrations de substrat. En conséquence, il existe un besoin de méthodes précises et sensibles pour la mesure du clivage de l'ADN par des endonucléases.

Un test sensible utilisant la technique d'ELISA a été développé. Ce test utilise des substrats ADN double-brin marqués en 5' sur chaque brin ; l'un avec la biotine, l'autre avec soit la fluoréscéine soit la digoxygénine (Jettsch et al., 1993, *Anal Biochem*, **213**, 234-240). Ce test 20 nécessite un grand nombre d'étapes avant la détection du signal.

D'autres tests utilisant la fluorescence ont été développés. Par exemple, les produits de la réaction peuvent être analysés par électrophorèse capillaire (Wenz, 1998, *Biochemistry*, **37**, 2234-2242). Dans une autre alternative, le test est basé sur le transfert de fluorescence entre deux fluorophores (Li et al., 2000, *Nucleic Acids Res*, **28**, e52 ; Ghosh et al, 1994, *Nucleic Acids Res*, 25 **22**, 3155-3159 ; Eisenschmidt et al., 2002, *J Biotech*, **96**, 185-191). Un des désavantages de ces méthodes est le coût de l'appareillage nécessaire pour les mettre en oeuvre.

De plus, toutes les techniques utilisant la fluorescence présentent un problème de sensibilité car de fortes concentrations d'ADN sont nécessaires pour obtenir un signal. Ces techniques sont plus adaptées pour une étude approfondie que pour des tests à haut rendement ou 30 des méthodes de sélection.

S'il existe des méthodes d'analyse de l'activité de clivage d'une endonucléase, le manque de méthode de sélection *in vitro* basée sur la capacité de clivage d'un substrat ADN par une endonucléase est considérable.

La sélection s'applique à des systèmes permettant de lier le génotype au phénotype et elle 35 permet de sélectionner parmi un pool le génotype présentant le phénotype désiré. Ces systèmes comprennent, par exemple, le phage display (Smith et Petrenko, 1997, *Chem Rev*, **97**, 391-410), le

ribosome display et la fusion ARNm-peptide (Roberts, 1999, *Curr Opin Chem Biol*, **3**, 268-273). Ces systèmes sont utilisés pour faire des sélections essentiellement basées sur la capacité de fixation.

Des stratégies de sélection sur la base d'une activité catalytique peuvent cependant être
5 utilisées dans un système intramoléculaire à « turnover » unique (Pedersen et al., 1998, *Proc Natl Acad Sci USA*, **95**, 10523-10528 ; Demartis et al., 1999, *J Mol Biol*, **286**, 617-633 ; Neri et al., US 6,184,012). Dans ce système, le substrat doit être lié à la protéine à tester. Ce substrat est techniquement compliqué à préparer.

Une autre stratégie alternative pour une sélection sur la base d'une activité catalytique est
10 de tester la capacité de liaison de la molécule exposée à un intermédiaire de transition de la réaction (McCafferty et al., 1994, *Appl Biochem Biotechnol*, **47**, 157-173 ; Fujii et al., 1998, *Nat Biotechnol*, **16**, 463-467 ; Widersten et al., 1995, *J Mol Biol*, **250**, 115-222). Cependant, la conception d'un intermédiaire de transition est difficile et incertaine. De plus, cela permet uniquement une appréciation très imprécise et indirecte de l'activité catalytique. Cette technique
15 semble très mal adaptée au cas des endonucléases.

Un autre moyen de sélection consiste en l'utilisation de substrat suicide biotinylé (Soumillon et al, 1994, *J Mol Biol*, **237**, 415-422 ; Janda et al, 1994, *Proc Natl Acad Sci USA*, **91**, 2532-2536 ; Janda et al, 1997, *Science*, **275**, 945-948 ; Vanwetswinkel et al, 2000, *J Mol Biol*, **295**, 527-540). Les billes portant la streptavidine fixent les phages portant les enzymes liées au substrat
20 suicide. Cependant, le manque de substrats suicides adaptés pour les réactions catalytiques à tester limite l'utilisation de cette stratégie.

Enfin, il existe une possibilité de sélection par phage display basée sur l'activité catalytique par une chromatographie avec élution catalytique (Ponsard et al., 2001, *Chembiochem*, **2**, 253-259).

25 Le problème commun à toutes ces stratégies de sélection par l'activité est qu'elles sont basées sur la modification d'une unique molécule de substrat. Ainsi, elles ne permettent pas de sélectionner des enzymes selon leur efficacité de catalyse. En effet, ce système peut aboutir à la sélection d'enzymes très peu efficaces. De plus, ces stratégies n'ont jamais été appliquées à la sélection d'endonucléase.

30 Un système de compartimentation utilise une émulsion eau-huile pour isoler dans une microcapsule un génotype et son phénotype (Tawfik et Griffiths, 1998, *Nat Biotechnol*, **16**, 652-656 ; Griffiths et Tawfik, 2000, *Cur Opin Biotechnol*, **11**, 338-353 ; Griffiths et Duncan, 1998, *Cur Opin Biotechnol*, **9**, 102-108 ; WO 00/40712 ; WO 99/02671). Ce système permet de baser la sélection sur une fonction biologique. Cependant, il n'a jamais été adapté à une sélection pour la
35 capacité de clivage d'une séquence d'ADN.

A ce jour, la seule méthode de sélection basée sur l'activité de clivage est utilisée *in vivo* dans *E. coli* (Gruen et al., 2002, *Nucleic Acids Res*, **30**, e29). Cette méthode utilise la capacité de couper un plasmide contenant le site à tester et portant un gène codant pour une protéine toxique pour la cellule. La coupure, qui aboutit à la dégradation de ce plasmide, est nécessaire à la survie
5 de la cellule et permet donc d'effectuer la sélection. Elle n'est utilisable que pour des endonucléases présentant des sites de reconnaissance et de coupure rares ou absents dans la cellule. De plus, cette méthode, en raison de l'étape de transfection, est laborieuse et limitante pour la taille des banques de départ.

En conséquence, il existe un besoin pour des tests d'activité d'endonucléase adaptés au
10 criblage à haut rendement et pour des méthodes de sélection *in vitro* basées sur la capacité de clivage des endonucléases.

RESUME DE L'INVENTION

15 La présente invention fournit des compositions et méthodes permettant de tester de manière efficace l'activité de clivage d'endonucléases présentant un type de coupure décalée. Les méthodes de l'invention peuvent être utilisées pour cribler, sélectionner, améliorer ou identifier des endonucléases spécifiques et/ou actives sur un site de coupure particulier, ou des sites de coupures clivés par une ou des endonucléases déterminées. L'invention est particulièrement
20 adaptée au criblage de banques d'endonucléases ou de banques de sites de coupure.

D'une manière générale, les méthodes de l'invention comprennent la mise en présence, dans des conditions propices à l'activité de coupure, d'une endonucléase et d'un substrat ADN double-brin comprenant deux sites de coupure, et la détection de la présence de substrat coupé aux
25 deux sites. De préférence, le substrat comporte deux sites distaux, plus préférentiellement localisés aux extrémités du substrat, qui peuvent être identiques ou différents, constants ou variables. L'utilisation de tels substrats permet d'améliorer la sélectivité ou la sensibilité du test, de réduire les faux positifs et/ou d'améliorer l'étape de détection de l'activité de clivage. L'invention est également applicable à des substrats ne portant qu'un seul site de coupure, en utilisant des
30 conditions particulières de détection de l'activité de clivage. L'invention peut être utilisée pour tester, en parallèle ou en mélange, plusieurs types de substrats ou plusieurs types d'endonucléases. Ces méthodes peuvent être utilisées pour le criblage d'endonucléases ou de sites de coupure mais également pour la sélection d'endonucléases ou de sites de coupure.

Selon un premier aspect plus particulier, la présente invention concerne une méthode permettant de tester l'activité de clivage d'une endonucléase présentant un type de coupure décalée, comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin comprenant deux sites de coupure;
 - 5 2) fournir ladite endonucléase ;
 - 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
 - 4) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
 - 10 5) détecter le substrat ADN coupé aux deux sites;
- la détection dudit substrat ADN coupé révélant l'activité de clivage de ladite endonucléase.

Selon un autre aspect particulier, l'invention concerne une méthode de criblage permettant
15 d'identifier un ou des site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée, comprenant les étapes suivantes :

- 1) fournir une banque de substrats ADN double-brins comprenant chacun deux sites de coupure, au moins un des sites de coupure étant variable dans la banque;
- 2) fournir ladite endonucléase ;
- 20 3) distribuer les différents substrats ADN dans des récipients et mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
- 4) discriminer, dans chaque récipient, les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
- 5) détecter les substrats ADN coupés aux deux sites;
- 25 la présence des substrats ADN coupés révélant l'activité de clivage et permettant d'identifier un ou des sites de coupure coupé(s) par ladite endonucléase.

Un autre aspect de l'invention est relatif à une méthode de criblage permettant d'identifier
une ou des endonucléase(s) présentant un type de coupure décalée capable(s) de couper un site de
30 coupure déterminé, comprenant les étapes suivantes :

- 1) fournir une banque de séquences codant pour des endonucléases ;
- 2) fournir un substrat ADN double-brin comprenant deux sites de coupure déterminés;
- 3) distribuer chaque séquence codante de ladite banque dans des récipients;
- 4) exprimer l'endonucléase codée par les séquences codantes et mettre en présence ladite
35 endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;

- 5) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
- 6) détecter les substrats ADN coupés aux deux sites;
- la détection desdits substrats ADN coupés révélant l'activité de clivage et permettant
- 5 d'identifier une ou des endonucléase(s) capable(s) de couper lesdits sites de coupure.

L'invention concerne également une méthode de sélection d'endonucléases capables de couper un site de coupure déterminé, comprenant les étapes suivantes :

- 1) fournir une première banque comprenant des séquences codant pour des endonucléases
- 10 présentant un type de coupure décalée;
- 2) préparer une banque de substrats ADN double-brins comprenant deux sites de coupure déterminés encadrant une cassette d'expression pour une endonucléase de ladite première banque ;
- 3) compartimenter ladite banque de substrats;
- 4) mettre ladite banque compartimentée dans des conditions propices à l'expression de
- 15 l'endonucléase et à l'activité de coupure ;
- 5) rompre la compartimentation ;
- 6) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et, de manière facultative,
- 7) amplifier les substrats ADN coupés aux deux sites afin de restaurer le substrat ADN; et
- 20 8) reproduire les étapes 3) à 7) jusqu'à un stade de sélection suffisant.

Un autre aspect de l'invention concerne une méthode de sélection permettant d'identifier un ou des demi-site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée, comprenant les étapes suivantes :

- 25 1) fournir une banque de substrats ADN comprenant deux sites de coupure dont au moins un site présente un demi-site constant et un demi-site variable, ledit demi-site constant étant placé vers l'extrémité du substrat ADN et étant dérivé d'un site de coupure connu pour être coupé par ladite endonucléase ;
- 2) fournir ladite endonucléase ;
- 30 3) mettre en présence ladite endonucléase et ladite banque de substrats ADN dans des conditions propices à l'activité de coupure ;
- 4) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés; et, si nécessaire
- 5) amplifier lesdits substrats ADN coupés aux deux sites afin de restaurer les substrats
- 35 ADN initiaux ; et,
- 6) reproduire les étapes 3) à 5) jusqu'à un stade de sélection suffisant.

La discrimination entre les espèces clivées aux deux sites et celles non-clivées peut être avantageusement par dégradation différentielle, par déplétion du substrat ADN non coupé, ou par amplification sélective du substrat ADN coupé. Elle peut également être basée sur une combinaison de ces méthodes.

5

Un autre aspect de l'invention concerne des acides nucléiques utilisables comme substrats dans les réactions ci-avant, caractérisés en ce qu'il s'agit d'ADN double-brin, d'une taille comprise entre 100 et 3000 pb environ, comprenant au moins deux régions distales formant des sites de coupure pour une endonucléase. Des tels acides nucléiques peuvent être sous forme
10 linéaires ou circulaire (e.g., incorporés dans un vecteur). Ils sont préférentiellement sous forme linéaire et peuvent, dans ce cas, être marqués à une extrémité, ou aux deux.

LEGENDE DES FIGURES

15 **FIGURE 1** Discrimination par dégradation sélective du substrat ADN non coupé aux deux sites de coupure. Les boîtes grises et noires représentent les demi-sites de coupure. L'étape de PCR est facultative (flèche pointillée).

FIGURE 2 Discrimination par déplétion du substrat ADN non coupé aux deux sites de coupure. Les boîtes grises et noires représentent les demi-sites de coupure. L et C indiquent
20 respectivement le ligand et la cible. Le demi-cercle quadrillé représente un support solide portant la cible. L'étape de PCR est facultative (flèche pointillée).

FIGURE 3 Discrimination par amplification PCR sélective du substrat ADN coupé aux deux sites de coupure. Les boîtes grises et noires représentent les demi-sites de coupure. La boîte contenant des ronds dans l'adaptateur indique une séquence non-présente dans le substrat ADN.
25 Les flèches noires représentent les amorces servant pour l'amplification par PCR. **FIGURE 3A** : Amplification sélective pour des substrats ADN coupés présentant des extrémités non-cohésives. **FIGURE 3B** : Amplification sélective pour des substrats ADN coupés présentant des extrémités 5' sortantes. L'adaptateur est un oligonucléotide simple-brin présentant un didéoxynucléotide à l'extrémité 3' (3'dd). L'élongation est effectuée à basse température. **FIGURE 3C** : Amplification
30 sélective pour des substrats ADN coupés présentant des extrémités 3' sortantes. L'adaptateur est un oligonucléotide simple-brin présentant un groupement phosphate à l'extrémité 5' (5'P). Une étape de ligation permet de lier les adaptateurs aux substrats ADN.

FIGURE 4 Schéma représentant un criblage d'une banque de sites de coupure afin d'identifier les sites de coupure capables d'être coupé par une endonucléase déterminée. Une
35 banque de sites de coupure est clonée dans un vecteur. A partir de ce vecteur, le substrat ADN est

préparé par amplification par PCR. Ce substrat est ensuite incubé en présence d'une endonucléase déterminée. Le substrat ADN résultant de l'incubation en présence de l'endonucléase est ensuite soumis à la discrimination entre le substrat ADN coupé et non coupé. Le substrat ADN restant après discrimination est détecté directement ou après amplification par PCR.

- 5 **FIGURE 5** Schéma représentant un criblage d'une banque de séquences codantes pour une endonucléase afin d'identifier les endonucléases capables de couper un site de coupure déterminé. Une banque de séquences codantes pour une endonucléase est clonée dans un vecteur d'expression.

Alternative de gauche : L'endonucléase est exprimée et purifiée. Le substrat ADN est mis
10 en présence avec l'endonucléase.

Alternative de droite : Un substrat ADN portant une cassette d'expression pour une endonucléase est préparé par amplification par PCR. L'endonucléase est exprimée *in vitro*.

Le substrat ADN résultant de l'incubation en présence de l'endonucléase testée est ensuite soumis à la discrimination entre le substrat ADN coupé et non coupé. Le substrat ADN restant
15 après discrimination est détecté directement ou après amplification par PCR.

FIGURE 6 Schéma représentant une sélection des endonucléases capables de couper un site de coupure déterminé parmi une banque de séquences codantes pour une endonucléase. Des substrats ADN de départ portent deux sites de coupure encadrant une cassette d'expression pour une endonucléase (une représentée par un damier, l'autre par des rayures). Les boîtes grises et
20 noires représentent les demi-sites de coupure. Les cassettes d'expression contiennent un promoteur, un RBS/Kozak et un terminateur. Ces substrats sont compartimentés. L'endonucléase est exprimée *in vitro*. Les compartiments sont placés dans des conditions propices à l'activité des endonucléases. Les ronds représentent les endonucléases. La compartimentation est ensuite rompue et les endonucléases sont inactivées. Les substrats ADN résultant de l'incubation en
25 présence d'une endonucléase sont ensuite soumis à la discrimination entre le substrat ADN coupé et non coupé. Les substrats ADN restant après discrimination sont amplifiés par PCR avec des amorces permettant de régénérer les sites de coupure.

FIGURE 7 Schéma représentant une sélection des demi-sites de coupure capables d'être coupés par une endonucléase déterminée parmi une banque de demi-sites de coupure. DS1 et DS2
30 désignent respectivement les deux demi-sites de coupure. Les boîtes noires et grises représentent ces deux demi-sites. Les damiers, hachures et briques représentent des demi-sites variables. Le rond représente l'endonucléase testée. Les flèches épaisses noires représentent les amorces de PCR.

FIGURE 7A : Schéma général représentant une sélection parmi une banque de demi-sites
35 de coupure. Les substrats ADN comprennent un ou deux demi-sites variables. Après incubation de l'endonucléase en présence du pool de substrats ADN contenant les demi-sites variables, les

substrats ADN résultants sont soumis à la discrimination entre les substrats ADN coupés et non coupés. Les substrats ADN restants sont ensuite amplifiés par PCR de façon à restaurer le site de coupure initial. La flèche pointillée indique que l'étape d'amplification par PCR peut être facultative. En effet, lorsque la discrimination est faite par amplification sélective du substrat
5 coupé, l'étape d'amplification par PCR n'est pas nécessaire. Les substrats régénérés peuvent être utilisés pour un tour supplémentaire de sélection.

FIGURE 7B : Schéma d'une sélection des demi-sites de coupure capables d'être coupés par une endonucléase déterminée parmi une banque de demi-sites de coupure, cette endonucléase générant des extrémités 5' sortantes. Les substrats ADN peuvent contenir un ou deux demi-sites
10 variables. Après incubation de l'endonucléase en présence du pool de substrats ADN contenant les demi-sites variables, les substrats ADN résultants sont soumis à la discrimination entre les substrats ADN coupés et non coupés par déplétion ou digestion différentielle des substrats ADN non coupés. Les substrats ADN restants sont ensuite amplifiés. Pour cela, des adaptateurs sont ajoutés. Ces adaptateurs sont des oligonucléotides simple-brins présentant un groupement phosphate
15 à l'extrémité 5' (5'P). Une étape de ligation permet de lier les adaptateurs aux substrats ADN. Une amplification par PCR est ensuite effectuée de façon à restaurer les sites de coupure initiaux.

FIGURE 7C : Schéma d'une sélection des demi-sites de coupure capables d'être coupés par une endonucléase déterminée parmi une banque de demi-sites de coupure, cette endonucléase générant des extrémités 3' sortantes. Les substrats ADN peuvent contenir un ou deux demi-sites
20 variables. Après incubation de l'endonucléase en présence du pool de substrats ADN contenant les demi-sites variables, les substrats ADN résultants sont soumis à la discrimination entre les substrats ADN coupés et non coupés par déplétion ou digestion différentielle des substrats ADN non coupés. Les substrats ADN restants sont ensuite amplifiés. Pour cela, des adaptateurs sont ajoutés. Ces adaptateurs sont des oligonucléotides simple-brins présentant un didéoxynucléotide à
25 l'extrémité 3' (3'dd). L'élongation est effectuée à basse température. Une amplification par PCR est ensuite effectuée de façon à restaurer les sites de coupure initiaux.

FIGURE 7D : Schéma d'une sélection des demi-sites de coupure capables d'être coupés par une endonucléase déterminée parmi une banque de demi-sites de coupure, cette endonucléase générant des extrémités 3' sortantes. Les substrats ADN peuvent contenir un ou deux demi-sites
30 variables. Après incubation de l'endonucléase en présence du pool de substrats ADN contenant les demi-sites variables, les substrats ADN résultants sont soumis à la discrimination entre les substrats ADN coupés et non coupés par amplification sélective. Pour cela, des adaptateurs sont ajoutés. Ces adaptateurs sont des oligonucléotides simple-brins présentant un didéoxynucléotide à l'extrémité 3' (3'dd) et comprenant une séquence non présente dans les substrats ADN.
35 L'élongation est effectuée à basse température. Une amplification par PCR est ensuite effectuée de

façon à restaurer les sites de coupure initiaux. ()n indique que la molécule est présente en un grand nombre d'exemplaires.

FIGURE 7E: Lorsque la sélection a été faite sur des substrats présentant deux demi-sites variables, ce schéma représente le principe d'une méthode permettant de tester les sites formés par les demi-sites sélectionnés. Les demi-sites portés par les différents substrats sélectionnés sont mélangés par amplification PCR (« suffire »). Les substrats mélangés sont clivés par l'endonucléase testée et circularisés par l'action d'une ligase. Les substrats circularisés peuvent être utilisés pour faire une sélection, un criblage ou un séquençage.

FIGURE 8: Réamplification du substrat ADN flanqué de sites I-SceI après dégradation sélective du substrat ADN non coupé aux deux sites par I-SceI

Le substrat ADN est incubé en absence (1) ou en présence (2) de l'endonucléase I-SceI et de magnésium, puis d'exonucléase III, et les fragments non digérés par l'exonucléase III sont réamplifiés. Seul les fragments digérés par I-SceI ou non soumis à l'action de l'exonucléase III sont réamplifiés. M: marqueur de poids moléculaires.

FIGURE 9: Schéma décrivant les éléments présents sur le substrat ADN portant une cassette d'expression pour I-Sce I permettant de programmer l'extrait de transcription/traduction *in vitro*.

FIGURE 10: Détection de l'activité de coupure de l'endonucléase I-SceI produite dans un extrait de transcription/traduction *in vitro*. Après une discrimination par dégradation sélective du substrat ADN non coupé aux deux sites par l'exonucléase III, seuls les fragments incubés en présence d'I-Sce I purifiée (puit 3) ou incubés dans l'extrait RTS (puits 4 et 5) donnent lieu à une réamplification. (système RTS100 (Roche))

FIGURE 11: Sélection d'un fragment d'ADN non biotinylé (CCC) à partir d'un mélange contenant un excès de fragment d'ADN biotinylé (bCSC) par déplétion sur billes magnétiques recouvertes de streptavidine. A: PCR sur un échantillon de chaque mélange, avant ou après déplétion. Les pistes 1 et 5, 2 et 6, 3 et 7, 4 et 8, 5 et 9, correspondent respectivement à des ratios bCSC:CCC de 1:1, 10:1, 100:1, 1000:1 et 10000:1 (la concentration de bCSC est invariable); B Même chose que pour A, mais en l'absence du fragment bCSC (les dilutions de CCC restent identiques). M: Marqueur de poids moléculaires.

30

TEST D'ACTIVITÉ ENDONUCLÉASIQUE

La présente invention concerne, de manière générale, une méthode permettant de tester l'activité de clivage d'une endonucléase présentant un type de coupure décalée, comprenant les étapes suivantes :

- mettre en présence, dans des conditions propices à l'activité de coupure, une endonucléase et un substrat, ledit substrat étant un ADN double-brin comprenant deux sites de coupure ; et
- détecter la présence de substrat ADN coupé aux deux sites.

5 Avantageusement, la détection du substrat coupé aux deux sites comprend une étape préalable de discrimination entre les substrats coupés aux deux sites et les substrats non coupés aux deux sites.

Ainsi, selon un mode particulier, la présente invention concerne une méthode permettant de tester l'activité de clivage d'une endonucléase présentant un type de coupure décalée,

10 comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin comprenant deux sites de coupure;
- 2) fournir ladite endonucléase ;
- 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
- 15 4) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
- 5) détecter le substrat ADN coupé aux deux sites;

la détection dudit substrat ADN coupé révélant l'activité de clivage de ladite endonucléase.

20 Par « discriminer » est entendu la séparation des deux types de substrat : les substrats coupés aux deux sites et les substrats non-coupés aux deux sites. La séparation peut être faite par séparation physique des deux types de substrats, par dégradation du substrat non coupé, par amplification sélective des substrats coupés aux deux sites, ou une combinaison de ces méthodes.

L'ordre des étapes 1) et 2) peut être interverti.

25 Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure.

30 De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus particulièrement, ils sont situés aux extrémités du substrat ADN. Lesdits sites de coupure sont de préférence identiques. Lorsqu'ils sont identiques, ils ont de préférence une orientation inverse.

Cependant les deux sites de coupures peuvent être différents : soit l'un constant et l'autre variable, soit les deux constants ou variables.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

Dans un mode de réalisation préféré, la discrimination consiste en la dégradation du
5 substrat ADN non coupé aux deux sites par une exonucléase incapable de dégrader le substrat ADN coupé aux deux sites. Facultativement, ladite exonucléase est inactivée après l'étape de discrimination, plus particulièrement avant l'étape 5) de détection. Lorsque ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 3' sortante, ladite exonucléase est de préférence l'exonucléase III. Lorsque ladite
10 endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 5' sortante, ladite exonucléase est de préférence l'exonucléase Lambda.

Dans un autre mode de réalisation, la discrimination consiste en la séparation des substrats ADN coupés aux deux sites et non coupés aux deux sites. De préférence, ledit substrat ADN initial porte à au moins une de ses extrémités, de préférence aux deux, un ligand (par exemple, une
15 biotine), ce ligand n'étant plus présent après la coupure du substrat ADN. Lors de l'étape de discrimination, le substrat ADN non coupé est éliminé par liaison non covalente avec une molécule ou cible présentant une affinité pour le ligand, cette molécule étant de préférence immobilisée sur un support solide. De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

20 Facultativement, l'étape de détection 5) comprend : a) amplifier par PCR le substrat ADN coupé; et, b) détecter le produit d'amplification ; ledit produit d'amplification constituant la mesure de l'activité de clivage de ladite endonucléase.

Dans un mode de réalisation additionnel, la discrimination consiste en l'amplification sélective des substrats ADN coupés. De préférence, cette amplification sélective est faite par une
25 PCR utilisant des amorces spécifiques d'une séquence présente uniquement sur des adaptateurs capables de s'hybrider à au moins une extrémité du substrat ADN coupé. Ainsi, des adaptateurs capables de s'hybrider à au moins une des extrémités sortantes générées par la coupure de l'endonucléase sur le substrat ADN, de préférence aux deux, sont ajoutés et sont soit liés covalamment par l'action d'une ligase soit intégrés dans le substrat ADN par élongation, lesdits
30 adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN utilisée pour les amorces de PCR.

Dans une alternative également envisagée, le substrat ADN comprend un site de coupure testé avec l'endonucléase étudiée et un autre site de coupure pour une autre endonucléase. Cette autre endonucléase doit présenter un type de coupure décalée. Lorsque l'étape de discrimination
35 implique une digestion par une exonucléase, un même type de coupure que l'endonucléase étudiée est nécessaire. Dans cette alternative, la méthode comprend en outre la fourniture de cette

deuxième endonucléase et sa mise en présence avec le substrat ADN. Ainsi, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 3).

Facultativement, une étape de purification du substrat ADN peut être ajoutée avant l'étape 5 de discrimination. Cette étape permet de séparer le substrat ADN des endonucléases utilisées. Cette étape est de préférence effectuée par une extraction Phénol/Chloroforme. Cette étape de purification peut se révéler nécessaire lorsque les endonucléases restent liées au site de coupure après le clivage.

10 La présente invention concerne une méthode permettant de tester l'activité de clivage d'une endonucléase présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin comprenant deux sites de coupure;
 - 2) fournir ladite endonucléase ;
 - 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
 - 15 4) ajouter au substrat ADN issu de l'étape 3) une exonucléase capable de dégrader spécifiquement le substrat ADN non coupé aux deux sites ; et,
 - 5) détecter le substrat ADN restant ;
- ledit substrat ADN restant révélant l'activité de clivage de ladite endonucléase.

20 L'étape de détection du substrat comprend de préférence deux étapes : a) amplifier par PCR le substrat ADN restant ; et, b) détecter le produit d'amplification ; ledit produit d'amplification constituant la mesure de l'activité de clivage de ladite endonucléase. L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression 25 pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus 30 particulièrement, ils sont situés aux extrémités du substrat ADN. Lesdits sites de coupure sont de préférence identiques. Lorsqu'ils sont identiques, ils ont de préférence une orientation inverse.

Cependant les deux sites de coupures peuvent être différents : soit l'un constant et l'autre variable soit les deux constants ou variables.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure 35 décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

De préférence, ladite exonuclease est inactivée avant l'étape de détection.

Lorsque ladite endonuclease possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 3' sortante, ladite exonuclease est de préférence l'exonuclease III. Lorsque ladite endonuclease possède un type de coupure décalée, plus
5 particulièrement d'au moins 4 nucléotides, générant une extrémité 5' sortante, ladite exonuclease est de préférence l'exonuclease Lambda.

Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonuclease que celle testée et cette autre endonuclease est ajoutée à l'étape 3).

10 La présente invention concerne une méthode permettant de tester l'activité de clivage d'une endonuclease présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin comprenant deux sites de coupure;
 - 2) fournir ladite endonuclease ;
 - 3) mettre en présence ladite endonuclease et ledit substrat ADN dans des conditions
15 propices à l'activité de coupure ;
 - 4) séparer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
 - 5) détecter le substrat ADN coupé aux deux sites;
- la détection dudit substrat ADN coupé révélant l'activité de clivage de ladite
20 endonuclease.

L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonuclease et l'étape 2) consiste en l'expression de ladite endonuclease. De préférence, ladite endonuclease est produite par un système de transcription/traduction couplée *in*
25 *vitro*.

Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus particulièrement, ils sont situés aux extrémités du substrat ADN. Lesdits sites de coupure sont de préférence identiques. Lorsqu'ils sont identiques, ils ont de préférence une orientation inverse.

30 Cependant les deux sites de coupures peuvent être différents : soit l'un constant et l'autre variable soit les deux constants ou variables.

Dans un mode de réalisation préféré, ladite endonuclease possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

35 De préférence, ledit substrat ADN initial porte à au moins une de ses extrémités un ligand et que ladite séparation est faite avec la cible du ligand immobilisée sur un support solide. De

préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique. De préférence, ledit substrat ADN porte un ligand à ses deux extrémités.

Facultativement, l'étape de détection 5) comprend : a) amplifier par PCR le substrat ADN coupé; et, b) détecter le produit d'amplification; ledit produit d'amplification constituant la

5 mesure de l'activité de clivage de ladite endonucléase.

Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 3).

La présente invention concerne une méthode permettant de tester l'activité de clivage

10 d'une endonucléase présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin comprenant deux sites de coupure;
 - 2) fournir ladite endonucléase ;
 - 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
 - 15 4) ajouter des adaptateurs capables de s'hybrider à au moins une des extrémités sortantes dudit substrat générées par la coupure de ladite endonucléase, de préférence aux deux extrémités, et soit les lier covalamment soit les intégrer audit substrat par élongation, lesdits adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN ;
 - 20 5) amplifier spécifiquement lesdits substrat ADN coupés aux deux sites avec une amorce s'hybridant sur ladite séquence non présente; et,
 - 6) détecter le substrat ADN amplifié;
- la détection dudit substrat ADN amplifié révélant l'activité de clivage de ladite endonucléase.

25 L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

30 Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus particulièrement, ils sont situés aux extrémités du substrat ADN. Lesdits sites de coupure sont de préférence identiques. Lorsqu'ils sont identiques, ils ont de préférence une orientation inverse.

Cependant les deux sites de coupures peuvent être différents: soit l'un constant et l'autre

35 variable soit les deux constants ou variables. Dans un mode de réalisation préféré, ladite

endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides.

La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

- 5 Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 3).

Dans le mode de réalisation alternatif où le substrat contient un site de coupure, la présente invention concerne une méthode permettant de tester l'activité de clivage d'une endonucléase
10 présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin contenant un site de coupure présentant une localisation distale sur ledit substrat ADN et portant un ligand à l'extrémité proche dudit site de coupure;
- 2) fournir ladite endonucléase ;
- 15 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
- 4) dépler les substrats ADN non coupés avec la cible du ligand immobilisée sur un support solide; et,
- 5) détecter le substrat ADN coupé;
- 20 la détection dudit substrat ADN coupé révélant l'activité de clivage de ladite endonucléase.

L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De
25 préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

- 30 De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Facultativement, l'étape de détection 5) comprend : a) amplifier par PCR le substrat ADN coupé; et, b) détecter le produit d'amplification ; ledit produit d'amplification constituant la mesure de l'activité de clivage de ladite endonucléase.

La présente invention concerne également une méthode permettant de tester l'activité de clivage d'une endonucléase présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir un substrat ADN double-brin contenant un site de coupure présentant une
5 localisation distale sur ledit substrat ADN;
 - 2) fournir ladite endonucléase ;
 - 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
 - 4) ajouter un adaptateur capable de s'hybrider à l'extrémité sortante dudit substrat générée
10 par la coupure de ladite endonucléase et soit le lier covalement soit l'intégrer audit substrat par élongation, ledit adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN ;
 - 5) amplifier spécifiquement lesdits substrat ADN coupés avec une amorce s'hybridant sur ladite séquence non présente; et,
 - 15 6) détecter le substrat ADN amplifié;
- la détection dudit substrat ADN amplifié révélant l'activité de clivage de ladite endonucléase.

L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression
20 pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de
25 préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

Le test d'activité d'une endonucléase selon la présente invention est à haut rendement. En effet, il est parfaitement adapté pour être mis en œuvre sur des plaques 96, 384, ou 1536 puits. Il nécessite peu de matériels. En effet, il ne nécessite aucun oligonucléotide marqué radioactivement
30 ou par fluorescence. Il nécessite uniquement un appareil de PCR, présent dans tous les laboratoires de nos jours. Il est beaucoup plus sensible que les technologies actuellement disponibles et requiert peu de substrat. En effet, la détection peut se faire sur le produit amplifié par PCR. De plus, il permet d'estimer l'efficacité catalytique de l'endonucléase testée.

Cette méthode permettant d'observer l'activité de clivage d'une endonucléase est
35 applicable pour des méthodes de criblage et/ou de sélection soit d'endonucléases soit de sites de coupure.

L'invention concerne donc une méthode de criblage permettant d'identifier parmi une banque d'endonucléases la ou les endonucléase(s) capable(s) de couper un site de coupure déterminé, la méthode de criblage étant basée sur la méthode d'observation de l'activité de clivage selon la présente invention.

- 5 L'invention concerne également une méthode de sélection permettant d'identifier parmi une banque d'endonucléases la ou les endonucléase(s) capable(s) de couper un site de coupure déterminé, la méthode de sélection étant basée sur la méthode d'observation de l'activité de clivage selon la présente invention.

- 10 L'invention concerne en outre une méthode de criblage permettant d'identifier parmi une banque de sites de coupure le ou les site(s) de coupure coupé(s) par une endonucléase déterminée, la méthode de criblage étant basée sur la méthode d'observation de l'activité de clivage selon la présente invention.

- 15 L'invention concerne de plus une méthode de sélection permettant d'identifier parmi une banque de sites de coupure le ou les site(s) de coupure coupé(s) par une endonucléase déterminée, la méthode de sélection étant basée sur la méthode d'observation de l'activité de clivage selon la présente invention.

ENDONUCLEASE

- 20 La méthode selon la présente invention peut être appliquée à toute endonucléase présentant un type de coupure décalée, de préférence décalée d'au moins 4 nucléotides. Plus particulièrement, le décalage est compris entre 4 et 8 nucléotides, de préférence il est de 4 nucléotides. Ces endonucléases peuvent être naturelles, mutées, mutantes, hybrides ou chimériques. Parmi les endonucléases, les plus intéressantes sont les enzymes de restriction et les
- 25 méganucléases.

Les enzymes de restriction sont des outils bien connus en biologie moléculaire. Elles possèdent une spécificité de reconnaissance du site de coupure. Les sites sont constitués par des séquences palindromiques de 4 à 8 nucléotides.

- 30 Les endonucléases de restriction peuvent être caractérisées par leur type de coupure. Il en existe plusieurs types: des coupures franches et des coupures décalées. Les coupures décalées présentent souvent un décalage de 1 à 5 nucléotides. Le type de coupures décalées le plus fréquent présente un décalage de 4 nucléotides. Les coupures décalées génèrent des extrémités sortantes qui peuvent être soit 3' sortantes soit 5' sortantes. Pour les enzymes de restriction, le type de coupure le plus fréquent pour les enzymes de restriction est décalé de 4 nucléotides.

- 35 Les méganucléases sont des endonucléases présentant des sites de reconnaissance de 10 à 40 pb, ou plus. Par la longueur de leur site de reconnaissance, elles présentent des sites de coupure

très rares. La rareté de leur site de coupure en fait des outils précieux, notamment pour l'ingénierie des génomes. Elles peuvent être classées suivant le type de coupure générée. Un exemple de méganucléases est constitué par les « homing endonucléases ». La grande majorité des « homing endonucléases » présente un type de coupure décalée de 4 nucléotides générant des extrémités 3' sortantes. Toutes les « homing endonucléases » à motif dodécapeptide ou LAGLIGDAG sont caractérisées par ce type de coupure (Dalgaard et al, 1997, *Nucleic Acids Research*, **25**, 4626-463 ; Chevalier et Stoddard, 2001, *Nucleic Acids Research*, **29**, 3757-3774 ; l'enseignement de ces documents étant incorporé par référence).

La méthode peut également être appliquée aux endonucléases chimériques présentant par exemple un domaine de reconnaissance de l'ADN lié à la tête catalytique d'une endonucléase de restriction de type IIS. Un exemple de domaine de liaison à l'ADN est le doigt à zinc. Un exemple de tête catalytique est celle de l'endonucléase *Fok I*. (Smith et al, 2000, *Nucleic Acids Res*, **28**, 3361-3369 ; Kim et al, 1996, *Proc Natl Acad Sci USA*, **93**, 1156-1160 ; Kim & Chandrasegaran, 1994, *Proc Natl Acad Sci USA*, **91**, 883-887 ; WO 95/09233 ; WO 94/18313).

Certaines endonucléases présentent des sites de reconnaissance et de clivage confondus. En effet, ces endonucléases effectuent la coupure double-brin dans leur site de reconnaissance. Ceci est le cas pour les enzymes de restriction du type II et pour les « homing endonucléases » à motif dodécapeptide. Les méthodes selon la présente invention sont particulièrement bien adaptées pour ce type d'endonucléase.

D'autres endonucléases présentent des sites de reconnaissance et des sites de clivage distincts. Ceci est le cas pour les enzymes de restriction de type IIS (*FokI*) et pour certaines « homing endonucléases » (*I-Tev I*). Les méthodes selon la présente invention sont également adaptées pour ce type d'endonucléase. Pour les méthodes de sélection, il est important que les sites de reconnaissance de l'endonucléase soient vers l'intérieur du substrat ADN par rapport aux sites de clivage qui ont une position plus distale.

Il est extrêmement intéressant de créer de nouvelles endonucléases ou de faire évoluer les endonucléases existantes pour avoir des endonucléases présentant de nouvelles spécificités. Dans ce but, il est nécessaire de générer des banques de séquences codant pour des endonucléases et de les tester. La méthode selon la présente invention peut être appliquée à la caractérisation, au criblage et/ou à la sélection de ces endonucléases.

SUBSTRAT ADN

Le substrat ADN pour la méthode selon la présente invention est double-brin et, de préférence, linéaire. Ce substrat peut présenter des extrémités franches ou 3' sortantes si la coupure par l'endonucléase engendre des extrémités 5' sortantes. Ce substrat peut présenter des

extrémités franches ou 5' sortantes si la coupure par l'endonucléase engendre des extrémités 3' sortantes. De préférence, le substrat ADN présente des extrémités franches.

Le substrat ADN présente au moins deux sites de coupure. De préférence, le substrat contient deux sites. Les sites ont de préférence une localisation distale sur le substrat ADN. Plus
5 particulièrement, ils sont placés aux extrémités du substrat ADN.

Lorsque l'endonucléase présente des sites de reconnaissance et de clivage distincts, on entend par « site », le site de reconnaissance associé au site de clivage. De préférence, les sites de reconnaissance ont une localisation proximale sur le substrat ADN par rapport à la localisation des sites de clivage. Le site peut être subdivisé en deux demi-sites, les demi-sites étant définis comme
10 la séquence se trouvant de chaque côté de la coupure double-brin. Par « position proximale » est entendue une position vers le centre du substrat ADN par rapport à l'autre demi-site du site de coupure. Par « position distale » est entendue une position vers l'extrémité du substrat ADN par rapport à l'autre demi-site du site de coupure.

Dans un premier mode de réalisation, les deux sites sont identiques. Ils peuvent avoir une
15 orientation directe ou inverse sur le substrat ADN. Dans un mode de réalisation préféré, les deux sites présentent une orientation inverse. Cette orientation inverse permet l'amplification par PCR du substrat ADN avec une seule amorce capable de s'hybrider au niveau du demi-site de coupure restant sur le substrat ADN coupé à amplifier. L'orientation inverse des sites est préférable pour l'utilisation de la méthode pour effectuer une sélection d'endonucléase basée sur une
20 discrimination utilisant une digestion différentielle ou une déplétion. En effet, cette orientation permet la régénération du substrat ADN, avec les deux sites de coupure entiers, après un tour de sélection.

Dans un deuxième mode de réalisation, le substrat ADN comprend deux sites de coupure différents. Un premier site est connu pour être coupé par l'endonucléase testée. Un deuxième site
25 est un site variable. Ce mode de réalisation permet de tester la spécificité de l'endonucléase étudiée et de faire un criblage ou une sélection d'une banque de sites de coupure.

Dans un troisième mode de réalisation, le substrat ADN comprend deux sites de coupure différents. Un des sites est coupé par une endonucléase autre que l'endonucléase testée. Le deuxième site est soit un site variable soit un site connu pour être coupé par l'endonucléase testée.
30 Par exemple, cette deuxième endonucléase peut être l'enzyme de restriction Sph I.

Dans un quatrième mode de réalisation, le substrat ADN comprend deux sites de coupure différents. Un substrat ADN particulièrement intéressant présente deux sites comportant chacun un demi-site variable, ce demi-site variable ayant une orientation proximale sur le substrat ADN. Les autres demi-sites sont de préférence des demi-sites présents dans des sites de coupure connus pour
35 être coupés par l'endonucléase testée, ce demi-site constant ayant une orientation distale sur le substrat ADN. De préférence, les sites de coupure présentent une orientation directe. Ce mode de

réalisation est particulièrement bien adapté pour tester la spécificité de l'endonucléase étudiée et faire un criblage ou une sélection d'une banque de sites de coupure.

Un site variable peut avoir soit une séquence déterminée, soit une séquence partiellement ou complètement aléatoire. La variation peut être contenue dans un demi-site ou concerner le site
5 entier.

Dans un mode de réalisation, le substrat comprend également la séquence codant pour l'endonucléase à tester. Plus particulièrement, il comprend une cassette d'expression pour l'endonucléase avec tous les éléments nécessaires à sa transcription en ARNm et à la traduction de cet ARNm en protéine. Cette séquence codante ou cassette d'expression est de préférence localisée
10 entre deux sites. Ce mode de réalisation est particulièrement avantageux pour l'utilisation de la méthode selon la présente invention pour la sélection d'une endonucléase. L'expression de cette endonucléase peut être faite dans une cellule eucaryote ou procaryote ou dans un système transcription/traduction couplée *in vitro*.

La cassette d'expression est de préférence adaptée au système de transcription/traduction
15 couplé *in vitro*. Ce système de transcription/traduction couplé *in vitro* peut être, de manière non-exhaustive, le système S30 de *E. coli* (par exemple, le système RTS de Roche), un extrait de réticulocyte de lapin (Pelham et Jackson, 1976, *Eur J Biochem*, **67**, 247-256), ou un extrait de germe de blé (Anderson et al., 1983, *Methods Enzymol*, **101**, 635-644) (Promega, Pharmacia, Panvera). Cette cassette d'expression comprend, outre la séquence codante pour l'endonucléase,
20 un promoteur (par exemple, le promoteur T7), un site de liaison au ribosome (une séquence kozak ou shine-dalgarno suivant le système utilisé, eucaryote ou procaryote), un codon ATG en début de la séquence codante, un codon stop en fin de celle-ci et un terminateur de transcription (par exemple, celui du bactériophage T7).

La taille du substrat est de préférence suffisante pour obtenir une amplification par PCR
25 efficace lorsqu'une étape d'amplification par PCR précède l'étape de détection de l'ADN ou se révèle nécessaire. L'utilisation d'un très grand fragment d'ADN, sans raison particulière, n'est pas souhaitable car la grande taille de ce fragment pourrait entraîner une augmentation du coût de la méthode et gêner lors de l'étape de discrimination, par exemple gêner la dégradation différentielle par l'exonucléase. C'est pourquoi, dans un mode de réalisation préféré, la taille du substrat ADN
30 est comprise entre 25, 50 ou 100 et 3000 pb, de préférence entre 300 et 1000 pb.

Lorsqu'une étape d'amplification par PCR précède l'étape de détection de l'ADN ou se révèle nécessaire, la distance entre les deux sites est de préférence bien adaptée pour obtenir une amplification par PCR efficace. Cette distance est généralement comprise entre 25, 50 ou 100 pb et 3000 pb, de préférence entre 300 et 1000 pb. Dans l'alternative où le substrat comprend la
35 séquence codante pour l'endonucléase, cette distance dépend de la taille de la cassette d'expression.

Dans un mode de réalisation, le substrat peut contenir, en plus des deux sites testés, un ou plusieurs sites de reconnaissance et de coupure différents des sites testés. Ainsi, si l'endonucléase ne présente pas la spécificité requise, elle coupe dans les sites additionnels. Ainsi, le substrat ayant subi les coupures additionnelles ne pourra pas être amplifié par PCR. Les sites additionnels sont de
5 préférence placés entre les deux sites testés.

Le substrat ADN tel que décrit ci-dessus peut être utilisé dans la méthode selon la présente invention en combinaison avec d'autres molécules d'ADN contenant un site de coupure. Ce site peut être identique ou différent de ceux compris par le substrat ADN. Ces autres molécules permettent de faire une compétition avec les sites de coupure présents sur ledit substrat. Ainsi,
10 l'efficacité de l'endonucléase peut être estimée plus précisément. Ces autres molécules sont de préférence éliminées lors de l'étape de discrimination et/ou ne sont pas amplifiées lors de l'étape d'amplification par PCR. Cependant, les substrats initiaux peuvent être composés d'un mélange de substrat(s) portant une cassette d'expression pour une endonucléase et de substrat(s) sans la cassette d'expression. Ce mélange peut permettre une estimation plus fine de l'efficacité de
15 clivage de l'endonucléase. Ce mélange, de par la compétition entre divers substrats, augmente la rigueur du test, de la sélection ou du crible.

Facultativement, le substrat ADN peut également comprendre un gène de résistance à un antibiotique et/ou une origine de répllication. En effet, ces éléments permettent d'utiliser une forme circulaire du substrat ADN pour transformer une cellule hôte, de préférence une bactérie, plus
20 particulièrement E. coli.

Le substrat ADN peut être préparé par de nombreuses techniques bien connu de l'homme du métier. Parmi ces techniques, on trouve la synthèse d'oligonucléotides et l'amplification par PCR.

Par exemple, un substrat contenant deux sites constants peut être préparé à l'aide d'une
25 amplification par PCR avec des amorces contenant la séquence du site de coupure (voir Exemple I). Lorsque le substrat comprend en plus la séquence codante pour une endonucléase, ces amorces amplifient une matrice présentant cette séquence codante (voir Exemple II). Les éléments nécessaires à l'expression de l'endonucléase sont compris soit dans la séquence de la matrice amplifiée soit dans les amorces de PCR. De préférence, la matrice est un vecteur de clonage (par
30 exemple, un vecteur pET (Novagen)) dans lequel a été clonée une séquence codante pour une endonucléase.

Un substrat comprenant un site variable peut être préparé par amplification par PCR d'une matrice comprenant le site variable. Facultativement, cette matrice peut également comprendre la séquence codante pour une endonucléase, cette séquence pouvant être intégrée dans une cassette
35 d'expression. Les amorces sont conçues de telle manière que le fragment d'amplification comprenne les deux sites de coupure.

Dans un mode de réalisation alternatif, l'invention considère également l'usage d'un substrat ADN contenant un seul site de coupure. Le site de coupure présente de préférence une localisation distale sur le substrat ADN. Le substrat ADN contenant un site de coupure est compatible avec les modes de discrimination par déplétion du substrat ADN non coupé et/ou par
5 amplification spécifique du substrat ADN coupé.

Cependant, la présence d'au moins deux site de coupure permet de diminuer le nombre de faux positifs dans les méthodes selon la présente invention. En effet, l'utilisation d'un substrat ADN présentant deux sites de coupure nécessite que l'événement de coupure se produise deux fois. Ainsi, la spécificité du test est augmentée.

10

DISCRIMINATION

DISCRIMINATION PAR DEGRADATION DIFFERENTIELLE

La discrimination entre le substrat ADN coupé et non coupé est préférentiellement faite
15 par une digestion différentielle du substrat ADN non coupé par une exonucléase (FIGURE 1). Le choix de l'exonucléase utilisée dans la méthode selon la présente invention dépend du type de coupure généré par l'endonucléase testée.

Si la coupure par l'endonucléase génère des extrémités 3' sortantes, l'exonucléase doit être capable de dégrader des extrémités franches ou 5' sortantes mais incapable de digérer les ADN
20 présentant une extrémité 3' sortante. De préférence, ladite exonucléase est l'exonucléase III.

Si la coupure par l'endonucléase génère des extrémités 5' sortantes, l'exonucléase doit être capable de dégrader des extrémités franches ou 3' sortantes mais incapable de digérer les ADN présentant une extrémité 5' sortante. De préférence, ladite exonucléase est l'exonucléase Lambda.

Les exonucléases III et Lambda sont des enzymes largement utilisées en biologie
25 moléculaire, notamment pour le séquençage, la production de délétions dans un ADN double-brin ou l'analyse des interactions ADN-protéines.

L'exonucléase III présente une activité exonucléasique de 3' vers 5'. Ses substrats préférés sont les ADN à bout francs ou 5' sortants. Cette exonucléase n'agit pas sur les ADNs simple-brins. Par conséquent, les extrémités 3' sortantes sont assez résistantes au clivage. Le degré de résistance
30 dépend de la longueur de l'extrémité 3' sortante. Un ADN avec une extrémité 3' sortante de 4 base ou plus est pratiquement résistant au clivage.

L'exonucléase Lambda présente une activité exonucléasique de 5' vers 3'. Ses substrats préférés sont les ADN à bout francs ou 3' sortants. Cette exonucléase agit très mal sur les ADNs simple-brins. Par conséquent, les extrémités 5' sortantes sont plutôt résistantes au clivage. Le
35 degré de résistance dépend de la longueur de l'extrémité 5' sortante. Un ADN avec une extrémité 5' sortante de 4 base ou plus est pratiquement totalement résistante au clivage.

En considération des exonucléases actuellement disponibles, la méthode selon la présente invention est particulièrement bien adaptée pour des endonucléase présentant un type de coupure décalée d'au moins 4 nucléotides. Elle peut être adaptée à des endonucléases présentant un type de coupure avec un décalage de 2 ou 3 nucléotides avec une efficacité moindre.

- 5 Lors de l'étape de dégradation différentielle par l'exonucléase, il peut être nécessaire de mettre au point préalablement les conditions de cette dégradation différentielle. Parmi ces conditions, il y a le temps d'incubation, la température et la concentration d'exonucléase.

DISCRIMINATION PAR DEPLETION

- 10 La discrimination entre le substrat ADN coupé et non coupé peut également être faite par déplétion du substrat non coupé (FIGURE 2).

Cette déplétion est de préférence faite grâce à un couple ligand-cible, par exemple le couple biotine-streptavidine ou biotine-avidine. Dans ce mode de réalisation, le substrat comprend à au moins une extrémité, de préférence aux deux, un ligand. De préférence, le ligand est fixé aux
15 deux extrémités du substrat ADN. Le substrat ADN coupé perd l'extrémité portant le ligand. Ainsi, une incubation en présence d'un support solide portant la cible permet de fixer sélectivement le substrat ADN non coupé et de le séparer du substrat ADN coupé. Le support solide peut être un support couvert de cible, un gel d'agarose portant la cible ou une bille paramagnétique portant la cible. De préférence, le support solide est une bille paramagnétique
20 portant la cible. Ainsi, le substrat ADN coupé est séparé du substrat ADN non coupé en présence d'un aimant. De préférence, le ligand est la biotine et la cible est la streptavidine. Les oligonucléotides biotinylés, les gels agarose portant la streptavidine (Streptavidin-agarose from Prozyme, Pierce Chemical) ainsi que les billes paramagnétiques (par exemple #S1420S de new England Biolabs ; Dynabeads™ de Dynal Biotech) sont disponibles commercialement.

- 25 Dans le mode de réalisation alternatif où le substrat ADN contient un site de coupure, le ligand est porté par une extrémité du substrat ADN. Cette extrémité est celle où est localisé le site de coupure.

Cependant, la présence de ligand à chaque extrémité du ligand permet d'augmenter très fortement l'efficacité de la déplétion par effet d'avidité. En effet, cette augmentation de l'efficacité
30 permet d'obtenir des déplétions complètes.

DISCRIMINATION PAR AMPLIFICATION SELECTIVE

La discrimination entre le substrat ADN coupé et non coupé peut également être faite par amplification sélective du substrat ADN coupé (FIGURE 3).

- 35 Pour effectuer cette amplification sélective, le substrat ADN coupé est incubé en présence d'adaptateurs permettant d'introduire une ou des séquence(s) non présente(s) dans le substrat

ADN. Ces séquences non présentes dans le substrat servent pour l'hybridation des amorces utilisées pour l'amplification sélective des substrats ADN coupés. Les adaptateurs sont conçus de telle sorte qu'ils peuvent interagir uniquement avec les substrats ADN coupés.

Lorsque ce mode de discrimination est utilisé dans une méthode de sélection, la séquence
5 non-présente sur le substrat ADN de l'adaptateur doit varier à chaque tour de sélection.

Dans un premier mode de réalisation de l'amplification sélective, le substrat ADN coupé possède des extrémités non cohésives (FIGURE 3A). Plus particulièrement, les nucléotides des extrémités sortantes du substrat ADN ne sont pas être complémentaires. Les adaptateurs sont de préférence des oligonucléotides double-brins capables de s'hybrider aux extrémités sortantes
10 générées par la coupure de l'endonucléase. Ces adaptateurs sont liés covalamment au substrat ADN par l'action d'une ligase. Ces adaptateurs permettent d'introduire une séquence non présente sur le substrat ADN. Cette séquence s'hybride avec les amorces de PCR et permet donc une amplification sélective par PCR.

Il est important que les adaptateurs ne puissent s'hybrider entre eux pour former un produit
15 très court qui concurrencerait les substrats ADN pour l'amplification par PCR. Pour cela, les nucléotides des extrémités sortantes du substrat ADN ne doivent pas être complémentaires et donc cohésives.

Dans un deuxième mode de réalisation de l'amplification sélective, le substrat ADN coupé possède des extrémités 3' sortantes (FIGURE 3B). Dans ce mode de réalisation, les adaptateurs
20 sont des oligonucléotides simple-brins portant un didéoxynucléotide à son extrémité 3' et dont les extrémités 3' sont complémentaires des extrémités 3' sortantes du substrat ADN. Une extension des extrémités 3' à basse température est ensuite faite en utilisant l'adaptateur comme matrice. Cette extension permet l'introduction d'une séquence non contenue dans le substrat ADN sur laquelle s'hybride l'amorce utilisée pour l'amplification sélective du substrat ADN coupé.

25 Dans un troisième mode de réalisation de l'amplification sélective, le substrat ADN coupé possède des extrémités 5' sortantes (FIGURE 3C). Les adaptateurs sont dans ce cas des oligonucléotides simple-brins capables de s'hybrider aux extrémités générées par la coupure de l'endonucléase et présentant un phosphate à son extrémité 5'. Une ligase crée des liaisons covalentes entre l'extrémité 5' de l'adaptateur et l'extrémité 3' rentrante du substrat ADN coupé.
30 Les adaptateurs contiennent une séquence non présente dans le substrat ADN qui s'hybride avec les amorces utilisées pour l'amplification sélective du substrat ADN coupé. L'amplification par PCR doit être faite dans des conditions (principalement des conditions de température) qui ne permettent pas l'hybridation des adaptateurs entre eux.

Dans le mode de réalisation alternatif où le substrat ADN contient un site de coupure,
35 l'adaptateur est conçu de telle sorte qu'il s'hybride avec l'extrémité sortante générée par la coupure de l'endonucléase et il est soit lié au substrat ADN par l'action d'une ligase soit introduit

dans le substrat par élongation. L'amplification spécifique met en œuvre une amorce se liant spécifiquement sur l'adaptateur.

Les différentes méthodes de discrimination peuvent être combinées. Par exemple, il est possible de faire une première étape de discrimination par déplétion suivie d'une digestion différentielle par une exonucléase. Dans un autre exemple, on peut combiner une première étape de discrimination par déplétion suivie d'une amplification spécifique. En outre, il est possible de combiner les trois types de discrimination : une déplétion suivie de la digestion différentielle et de l'amplification spécifique. Ces combinaisons permettent d'augmenter la capacité de discrimination entre les substrats ADN coupés et non coupés.

AMPLIFICATION PAR PCR ET DETECTION

Le substrat ADN coupé peut être détecté directement ou amplifié par PCR avant détection, si nécessaire. L'avantage d'une détection sans étape d'amplification est une simplification de la méthode. Par contre, l'étape d'amplification par PCR peut permettre une meilleure spécificité pour la détection du substrat ADN coupé et est nécessaire dans les méthodes de sélection pour effectuer un tour de sélection supplémentaire. En effet, dans le cas où la discrimination des substrats ADN coupés et non coupés ne serait pas complète, l'amplification par PCR avec des amorces s'hybridant avec des séquences correspondant à la partie du site restant ou proches des sites de coupure permet de masquer la présence des substrats ADN non coupés restants.

Les substrats coupés peuvent être amplifiés par PCR. Lorsque la discrimination est faite par digestion différentielle ou par déplétion, les amorces utilisées pour la PCR s'hybrident avec des séquences présentes sur le substrat entre les deux sites de coupure. De préférence, les amorces s'hybrident avec des séquences localisées aux extrémités du substrat ADN coupé. Dans un mode de réalisation préféré lorsque les sites ne sont pas variables, les amorces de PCR s'hybrident avec les demi-sites restants sur le substrat ADN coupé. Ces amorces de PCR peuvent être conçues pour recréer un site de coupure entier après amplification. Ces amorces contiennent la séquence du demi-site absent ainsi que la séquence entière ou partielle du demi-site restant sur le substrat ADN coupé. De préférence, lorsque la méthode nécessite la régénération des sites de coupure et que les deux sites sont identiques, les sites de coupure présentent une orientation inverse et une unique amorce de PCR suffit. Lorsque la méthode nécessite la régénération des sites de coupure et que la discrimination est faite par déplétion des substrats ADN non coupés, au moins une des amorces utilisées pour l'amplification par PCR, de préférence les deux, porte de préférence le ligand. Ainsi, les substrats ADN régénérés présentent à au moins une de ses extrémités, de préférence aux deux, un ligand.

Dans le mode de réalisation où la discrimination est faite par amplification sélective du substrat ADN coupé, au moins une amorce de PCR, de préférence les deux amorces, s'hybride avec des séquences présentes dans les adaptateurs.

L'amplification par PCR est bien connue par l'homme du métier (Current Protocols in
5 Molecular Biology, Chapter 15).

Le substrat ADN coupé ou le produit de l'amplification peut être détecté par de nombreuses techniques disponibles et connues par l'homme du métier. Parmi ces techniques de détection, on trouve, de manière non-exhaustive, l'électrophorèse, des mesures d'absorption, des mesures de luminométrie, des mesures de fluorescence (par exemple, Picogreen® de Molecular
10 Probes ; Hoechst 33258 de TurnerBioSystems). Dans un mode de réalisation préféré, la détection est faite par des mesures de fluorescence.

METHODES DE CRIBLAGE

15 CRIBLAGE DE SITES DE COUPURE

Le test d'activité de clivage d'une endonucléase selon la présente invention permet facilement de caractériser une endonucléase. En effet, un des paramètres les plus importants pour caractériser une endonucléase est sa spécificité pour le site de coupure. Etant donné que ce test est bien adapté pour le haut rendement, il permet de tester la capacité et l'efficacité de coupure d'une
20 endonucléase donnée par rapport à une grande diversité de sites de coupure.

Le principe du criblage d'une banque de sites de coupure est illustré dans la FIGURE 4.

Les divers sites de coupure peuvent présenter une séquence déterminée, ou une séquence partiellement ou complètement aléatoire. De nombreuses techniques permettant de préparer cette diversité de sites de coupure sont disponibles pour l'homme du métier. Par exemple, les sites
25 peuvent être préparés par des méthodes de synthèse d'oligonucléotides. La diversité est introduite grâce à des mélanges de plusieurs nucléotides à des positions choisies lors de la synthèse. Ces oligonucléotides correspondant aux sites de coupure sont ensuite clonés. Ils sont ensuite insérés dans un substrat ADN tel que décrit ci-dessus.

Dans un mode de réalisation, les substrats ADN initiaux comprennent un site constant et
30 un site variable, le site constant étant connu pour être coupé par l'endonucléase testée ou par une autre endonucléase. Un mode de préparation de ces substrats consiste en une étape de clonage des oligonucléotides correspondant aux sites de coupure variable en amont ou en aval d'une séquence matrice pour le substrat ADN. La séquence matrice peut comprendre une cassette d'expression pour l'endonucléase étudiée. Une amplification par PCR de chaque clone utilisant des amorces
35 dont l'une est capable de s'hybrider en aval du site variable et permettant d'aboutir à la production

d'un substrat ADN où les sites constants et variables présentent une localisation distale sur le substrat ADN.

Alternativement, les substrats ADN initiaux comprennent deux sites variables. De préférence, les deux sites ne varient que pour un de leur demi-site. Plus particulièrement, le demi-
5 site variable est celui qui a une position proximale sur le substrat ADN.

Les substrats ADN sont distribués dans des récipients, de préférence dans les puits d'une microplaque ou d'un système équivalent (microsupport, microréseau, ou biopuce) de manière à avoir un type de substrat par récipient ou puit. Les microplaques utilisées peuvent présenter 96, 384 ou 1536 puits. Les substrats ADN sont ensuite mis en présence de l'endonucléase à tester.
10 Lorsque le substrat ADN porte un site de coupure constant pour une autre endonucléase, cette autre endonucléase doit être également ajoutée. On peut ajouter également un tampon de clivage adapté aux conditions nécessaires pour l'activité de coupure de l'endonucléase testée. Si le substrat porte une cassette d'expression pour l'endonucléase, cette endonucléase peut être produite par un système d'expression transcription/traduction couplée *in vitro*.

15 L'endonucléase peut également être distribuée dans les récipients ou puits avant la répartition des différents substrats ADN.

Après cette étape de coupure, les substrats ADN coupés et non coupés doivent être discriminés. La méthode de discrimination peut être choisie parmi la dégradation différentielle du substrat ADN non coupé, la déplétion du substrat ADN non coupé, ou l'amplification sélective du
20 substrat ADN coupé. Elle peut également être basée sur une combinaison de ces méthodes.

Lorsque la discrimination est faite par dégradation différentielle, l'exonucléase est ajoutée après l'incubation du substrat ADN en présence de la ou des endonucléase(s) et l'échantillon est placé dans les conditions appropriées à la dégradation différentielle du substrat ADN non coupé par l'endonucléase. De préférence, l'exonucléase est ensuite inactivée, par exemple par la chaleur.
25 Le substrat ADN non dégradé est détecté. L'étape de détection peut être précédée par une étape d'amplification par PCR.

Lorsque la discrimination est faite par déplétion du substrat ADN non coupé, le substrat ADN initial porte à au moins une de ses extrémités, de préférence aux deux, un ligand (par exemple, la biotine). Après l'incubation du substrat ADN en présence de la ou des
30 endonucléase(s), l'échantillon est mis en présence de la cible du ligand immobilisée sur un support solide (par exemple, bille paramagnétique portant la streptavidine). Le support solide portant le substrat ADN non coupé est séparé du reste de l'échantillon. Le substrat ADN coupé est ensuite détecté. L'étape de détection peut être précédée par une étape d'amplification par PCR.

Lorsque la discrimination est faite par amplification sélective du substrat ADN coupé, les
35 adaptateurs adéquats sont ajoutés après l'incubation du substrat ADN en présence de la ou des endonucléase(s) et ils sont soit liés covalamment au substrat ADN coupé soit intégrés au substrat

ADN par élongation. Le substrat ADN coupé est spécifiquement amplifié par PCR en utilisant des amorces reconnaissant les adaptateurs. La quantité d'ADN est ensuite détectée.

L'invention concerne donc une méthode de criblage permettant d'identifier parmi une banque de sites de coupure le ou les site(s) de coupure coupé(s) par une endonucléase déterminée

5 présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir une banque de substrats ADN double-brins comprenant au moins deux sites de coupure, au moins un des sites de coupure étant variable dans la banque;
- 2) fournir ladite endonucléase ;
- 3) distribuer les différents substrats ADN dans des récipients, ajouter ladite endonucléase
- 10 dans chaque récipient, dans des conditions propices à l'activité de coupure ;
- 4) discriminer dans chaque récipient les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
- 5) détecter le substrat ADN coupé aux deux sites;
- la présence du substrat ADN coupé révélant la capacité de coupure par l'endonucléase
- 15 d'un site .

L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in*

20 *vitro*.

Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus particulièrement, ils sont situés aux extrémités du substrat ADN.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure

25 décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

Dans un mode de réalisation, la discrimination consiste en la dégradation du substrat ADN non coupé aux deux sites par une exonucléase incapable de dégrader le substrat ADN coupé aux deux sites. Facultativement, ladite exonucléase est inactivée avant l'étape 5) de détection. Lorsque

30 ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 3' sortante, ladite exonucléase est de préférence l'exonucléase III. Lorsque ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 5' sortante, ladite exonucléase est de préférence l'exonucléase Lambda.

35 Dans un autre mode de réalisation, la discrimination consiste en la séparation des substrats ADN coupés aux deux sites et non coupés. De préférence, ledit substrat ADN initial porte à au

moins une de ses extrémités, de préférence aux deux, un ligand et ladite séparation est faite avec la cible du ligand immobilisée sur un support solide. De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Facultativement, l'étape de détection 5) comprend : a) amplifier par PCR le substrat ADN coupé aux deux sites; et, b) détecter le produit d'amplification ; ledit produit d'amplification révélant la capacité de clivage du site par ladite endonucléase.

Dans un mode de réalisation additionnel, la discrimination consiste en l'amplification sélective des substrats ADN coupés aux deux sites. Ainsi, des adaptateurs capables de s'hybrider à au moins une des extrémités, de préférence aux deux, sortantes générées par la coupure de ladite endonucléase sur le substrat ADN coupé sont ajoutés et sont soit liés covalamment par l'action d'une ligase soit intégrés au substrat ADN par élongation, lesdits adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN utilisée pour les amorces de PCR.

Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 3).

De préférence, ledit récipient est un puit de microplaque ou équivalent.

L'ensemble des résultats sur les différents substrats ADN présentant des sites de coupure variable permet de connaître la spécificité de la première endonucléase.

Facultativement, les substrats ADN peuvent contenir une cassette d'expression pour la première endonucléase et/ou la deuxième endonucléase. Dans ce cas, les étapes 2) et 4) de la méthode sont remplacées par l'expression de la ou des endonucléase(s).

Facultativement, une étape de purification du substrat ADN peut être ajoutée avant l'étape de discrimination. Cette étape permet de séparer le substrat ADN des endonucléases utilisées. Cette étape est de préférence effectuée par une extraction Phénol/Chloroforme. Cette étape de purification peut se révéler nécessaire lorsque les endonculéases restent fixées au site de coupure après le clivage.

Dans le mode de réalisation alternatif où le substrat contient un site de coupure, la présente invention concerne donc une méthode de criblage permettant d'identifier parmi une banque de sites de coupure le ou les site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir une banque de substrats ADN double-brins contenant un site de coupure présentant une localisation distale sur ledit substrat ADN et portant un ligand à l'extrémité proche dudit site de coupure, ledit site de coupure étant variable dans la banque;
- 2) fournir ladite endonucléase ;
- 3) distribuer les différents substrats ADN dans les récipients, ajouter ladite endonucléase dans chaque récipient et placer dans des conditions propices à l'activité de coupure ;

4) dépler les substrats ADN non coupés avec la cible du ligand immobilisée sur un support solide; et,

5) détecter le substrat ADN coupé;

la présence du substrat ADN coupé révélant la capacité de coupure par l'endonucléase

5 d'un site .

De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Facultativement, l'étape de détection 5) comprend : a) amplifier par PCR le substrat ADN coupé; et, b) détecter le produit d'amplification ; ledit produit d'amplification révélant la capacité
10 de clivage du site par ladite endonucléase.

L'invention concerne donc une méthode de criblage permettant d'identifier parmi une banque de sites de coupure le ou les site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée comprenant les étapes suivantes :

1) fournir une banque de substrats ADN double-brins contenant un site de coupure
15 présentant une localisation distale sur ledit substrat ADN, ledit site de coupure étant variable dans la banque;

2) fournir ladite endonucléase ;

3) ajouter dans les différents récipients un adaptateur capable de s'hybrider à l'extrémité sortante dudit substrat générée par la coupure de ladite endonucléase et soit le lier
20 covalement soit l'intégrer audit substrat par élongation, ledit adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN ;

5) amplifier spécifiquement lesdits substrat ADN coupés avec une amorce s'hybridant sur ladite séquence non présente; et,

6) détecter le substrat ADN amplifié;

la détection dudit substrat ADN amplifié révélant la capacité de coupure par
25 l'endonucléase d'un site .

L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De
30 préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

35 De préférence, ledit récipient est un puit de microplaque ou équivalent.

CRIBLAGE D'ENDONUCLÉASES

Cette méthode de mesure de l'activité de clivage d'une endonucléase est bien adaptée au haut rendement. Elle permet d'effectuer un criblage d'une banque d'endonucléases pour identifier
5 celles qui sont capables de couper un site de coupure déterminé. Le principe du criblage d'une banque d'endonucléases est illustré dans la FIGURE 5.

Une banque d'endonucléases telles que définies ci-dessus peut être préparée par plusieurs méthodes permettant de générer de la diversité. Ces méthodes sont bien connu de l'homme du métier (Current Protocol in Molecular Biology ; Chapter 8). Ces méthodes comprennent, de
10 manière non-exhaustive, la mutagenèse aléatoire par PCR (Error Prone PCR), la mutagenèse par des oligonucléotides dégénérés, la mutagenèse par des souches mutatrice (par exemple E. coli mutD5), des irradiations ionisantes ou UV, des agents chimiques mutagènes ou le mélange de gènes (WO98/27230 ; WO97/20078 ; WO95/22625).

La banque est ensuite clonée dans un vecteur, de préférence dans un vecteur d'expression.
15 Par exemple, elle peut être clonée dans un vecteur pET (Novagen). Ce vecteur contient tous les éléments nécessaires pour une expression dans un système T7 (promoteur T7, terminateur T7, etc...).

Ces vecteurs peuvent être utilisés pour exprimer l'endonucléase afin d'effectuer le criblage ou pour préparer un substrat ADN comprenant une cassette d'expression pour une endonucléase de
20 la banque.

A partir de ces clones, les substrats ADN contenant une cassette d'expression peuvent être préparés par amplification par PCR en utilisant deux amorces contenant le site de coupure à tester. La première amorce contient une séquence lui permettant de s'hybrider en amont de la séquence codante pour l'endonucléase. La deuxième amorce s'hybride en aval de la séquence codante.

25 Dans une première possibilité, l'amorce « amont » comprend, dans l'ordre, le site de coupure, le promoteur, le site de liaison au ribosome, et s'hybride au début de la séquence codante. Dans une autre, l'amorce « amont » comprend le site de coupure et s'hybride en amont du promoteur. Une combinaison intermédiaire de ces deux possibilités est également envisagée dans le cadre de la présente invention.

30 Dans une première possibilité, l'amorce « aval » comprend, dans l'ordre, le site de coupure et le terminateur et s'hybride à la fin de la séquence codante pour l'endonucléase. Dans une autre, l'amorce « aval » comprend le site de coupure et s'hybride en aval du terminateur. Une combinaison intermédiaire de ces deux possibilités est également envisagée dans le cadre de la présente invention.

35 Le fragment d'amplification par PCR comprenant les deux sites de coupures entourant la cassette d'expression pour l'endonucléase constitue un substrat ADN de départ pour l'expérience

de criblage (par exemple, FIGURE 9). Ainsi, on obtient des substrats ADN comprenant toujours les mêmes sites de coupure et flanquant une banque de séquence codante pour les endonucléases.

Le principe du criblage d'une banque d'endonucléases pour identifier celles qui sont capables de couper un site de coupure déterminé est d'isoler chaque endonucléase de la banque et de la mettre en présence du substrat ADN comprenant le site de coupure à cliver. Après incubation de l'endonucléase et du substrat ADN, l'échantillon est soumis à la discrimination des substrats ADN coupés aux deux sites et des substrats ADN non coupés aux deux sites. Cette discrimination peut se faire par dégradation différentielle des substrats ADN non coupés, par déplétion des substrats ADN non coupés aux deux sites ou par amplification sélective des substrats ADN coupés aux deux sites. La présence du substrat ADN restant est révélatrice de l'activité de l'endonucléase et donc de sa capacité à cliver le site de coupure déterminé.

De façon plus détaillée, la méthode de criblage d'une banque d'endonucléases pour identifier celles qui sont capables de couper un site de coupure déterminé comprend les étapes suivantes.

Chaque séquence codante de la banque d'endonucléases est clonée dans un vecteur d'expression et isolée, par exemple dans les puits d'une plaque à 96, 384 ou 1536 puits. Elles peuvent également être isolées dans tout système de biopuce compatible avec la présente méthode.

L'endonucléase est ensuite exprimée.

Dans une mode de réalisation, le vecteur de clonage est introduit dans une cellule, par exemple une bactérie comme *E. coli*, et l'endonucléase est exprimée. Ensuite, les cellules sont lysées et l'endonucléase est purifiée. Pour faciliter la purification, l'endonucléase peut présenter une étiquette, par exemple une étiquette polyhistidine, GST, MBP (maltose binding protein). Ainsi, elle peut facilement être purifiée avec un anticorps spécifique de l'étiquette utilisée (par exemple, un anticorps anti-polyhistidine pour une étiquette polyhistidine) ou avec un composé capable de lier non-covalamment l'étiquette (par exemple, le nickel pour l'étiquette polyhistidine). Pour cette étape de purification, il est préférable d'utiliser des plaques présentant des puits recouverts directement ou indirectement soit de ces anticorps soit par ce composé.

Dans un autre mode de réalisation, l'endonucléase est produite à l'aide d'un système de transcription/traduction couplé *in vitro*. Chaque vecteur d'expression codant pour une endonucléase est distribué, par exemple, dans un puit d'une microplaque 96, 384 ou 1536 puits. Du mélange d'expression est ajouté dans chaque puit. Ainsi l'endonucléase est exprimée et l'étape de purification de cette endonucléase devient optionnelle.

Le substrat ADN selon la présente invention présentant le site de coupure à tester est distribué dans chaque puit. Chaque puit doit présenter des conditions compatibles avec l'activité de l'endonucléase criblée. Le substrat peut être ajouté accompagné du tampon de clivage.

Lorsque le substrat ADN porte la cassette d'expression pour l'endonucléase, il suffit de distribuer les différents substrats ADN codant pour une endonucléase dans les puits, de produire l'endonucléase par un système d'expression *in vitro*, et de permettre la coupure des sites par l'endonucléase produite.

- 5 Une discrimination est ensuite faite entre les substrats ADN coupés ou non aux deux sites. Cette discrimination peut être faite par dégradation différentielle ou déplétion du substrat ADN non coupé ou par amplification sélective du substrat ADN coupé.

Le substrat ADN restant est ensuite détecté directement ou après une étape d'amplification par PCR.

- 10 En conséquence, l'invention concerne une méthode de criblage permettant d'identifier parmi une banque de séquences codantes pour des endonucléases, présentant un type de coupure décalée, la ou les endonucléase(s) capable(s) de couper un site de coupure déterminé comprenant les étapes suivantes :

- 1) fournir ladite banque de séquences codantes pour des endonucléases ;
 - 15 2) fournir un substrat ADN double-brin comprenant deux sites de coupure;
 - 3) distribuer chaque séquence codante de ladite banque dans des récipients;
 - 4) exprimer l'endonucléase codée par la séquence codante en présence dudit substrat ADN dans des conditions propices à l'activité de coupure ;
 - 5) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés
 - 20 aux deux sites; et,
 - 6) détecter le substrat ADN coupé aux deux sites;
- la détection dudit substrat ADN coupé révélant l'activité de clivage et permettant d'identifier la ou les endonucléase(s) capable(s) de couper ledit site de coupure déterminé.

- L'étape 4) d'expression de l'endonucléase peut être faite *in vivo* ou *in vitro*. Lorsqu'elle
- 25 est faite *in vivo*, une étape de purification de l'endonucléase est nécessaire avant l'étape 5). De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation, ledit substrat ADN comprend en outre une cassette d'expression pour une séquence codante de ladite banque.

- 30 Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus particulièrement, ils sont situés aux extrémités du substrat ADN. Lesdits sites de coupure sont de préférence identiques. Lorsqu'ils sont identiques, ils ont de préférence une orientation inverse. Cependant les deux sites de coupures peuvent également être différents.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

Dans un mode de réalisation préféré, la discrimination consiste en la dégradation du
5 substrat ADN non coupé aux deux sites par une exonucléase incapable de dégrader le substrat ADN coupé aux deux sites. Facultativement, ladite exonucléase est inactivée avant l'étape 6) de détection. Lorsque ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 3' sortante, ladite exonucléase est de préférence l'exonucléase III. Lorsque ladite endonucléase possède un type de coupure décalée, plus
10 particulièrement d'au moins 4 nucléotides, générant une extrémité 5' sortante, ladite exonucléase est de préférence l'exonucléase Lambda.

Dans un autre mode de réalisation, la discrimination consiste en la séparation des substrats ADN coupés et non coupés. De préférence, ledit substrat ADN initial porte à au moins une de ses extrémités, de préférence aux deux, un ligand et ladite séparation est faite avec la cible du ligand
15 immobilisée sur un support solide. De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Facultativement, l'étape de détection 6) comprend : a) amplifier par PCR le substrat ADN coupé; et, b) détecter le produit d'amplification ; ledit produit d'amplification constituant la mesure de l'activité de clivage de ladite endonucléase.

20 Dans un mode de réalisation additionnel, la discrimination consiste en l'amplification sélective des substrats ADN coupés aux deux sites. Ainsi, des adaptateurs capables de s'hybrider à au moins une des extrémités sortantes générées par la coupure de ladite endonucléase sur le substrat ADN coupé, de préférence aux deux extrémités, sont ajoutés et sont soit liés covalamment par l'action d'une ligase soit intégrés au substrat par élongation, lesdits adaptateurs permettant
25 d'introduire une séquence non présente sur le substrat ADN utilisée pour les amorces de PCR.

Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 4).

De préférence, ledit récipient est un puit de microplaque ou équivalent.

Facultativement, une étape de purification du substrat ADN peut être ajoutée avant l'étape
30 de discrimination. Cette étape permet de séparer le substrat ADN des endonucléases utilisées. Cette étape est de préférence effectuée par une extraction Phénol/Chloroforme. Cette étape de purification peut se révéler nécessaire lorsque les endonculéases restent fixées au site de coupure après le clivage.

35 Dans le mode de réalisation alternatif où le substrat contient un site de coupure, l'invention concerne une méthode de criblage permettant d'identifier parmi une banque de séquences codantes

pour des endonucléases, présentant un type de coupure décalée, la ou les endonucléase(s) capable(s) de couper un site de coupure déterminé comprenant les étapes suivantes :

- 1) fournir ladite banque de séquences codantes pour des endonucléases ;
 - 2) fournir un substrat ADN double-brin contenant un site de coupure présentant une localisation distale sur ledit substrat ADN et portant un ligand à l'extrémité proche dudit site de coupure;
 - 3) distribuer chaque séquence codante de ladite banque dans des récipients;
 - 4) exprimer l'endonucléase codée par la séquence codante en présence dudit substrat ADN dans des conditions propices à l'activité de coupure ;
 - 5) dépler les substrats ADN non coupés avec la cible du ligand immobilisée sur un support solide; et,
 - 6) détecter le substrat ADN coupé;
- la détection dudit substrat ADN coupé révélant l'activité de clivage et permettant d'identifier la ou les endonucléase(s) capable(s) de couper ledit site de coupure déterminé.
- De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Facultativement, l'étape de détection 6) comprend : a) amplifier par PCR le substrat ADN coupé; et, b) détecter le produit d'amplification ; ledit produit d'amplification constituant la mesure de l'activité de clivage de ladite endonucléase.

- L'invention concerne une méthode de criblage permettant d'identifier parmi une banque de séquences codantes pour des endonucléases, présentant un type de coupure décalée, la ou les endonucléase(s) capable(s) de couper un site de coupure déterminé comprenant les étapes suivantes :

- 1) fournir ladite banque de séquences codantes pour des endonucléases ;
- 2) fournir un substrat ADN double-brin contenant un site de coupure présentant une localisation distale sur ledit substrat ADN;
- 3) distribuer chaque séquence codante de ladite banque dans des récipients;
- 4) exprimer l'endonucléase codée par la séquence codante en présence dudit substrat ADN dans des conditions propices à l'activité de coupure ;
- 5) ajouter dans chaque récipient un adaptateur capable de s'hybrider à l'extrémité sortante dudit substrat générée par la coupure de ladite endonucléase et soit le lier covalamment soit l'intégrer audit substrat par élongation, ledit adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN ;
- 6) amplifier spécifiquement lesdits substrat ADN coupés avec une amorce s'hybridant sur ladite séquence non présente; et,
- 7) détecter le substrat ADN amplifié;

la détection dudit substrat ADN amplifié révélant l'activité de clivage et permettant d'identifier la ou les endonucléase(s) capable(s) de couper ledit site de coupure déterminé.

Dans un mode de réalisation, ledit substrat ADN comprend en outre une cassette d'expression pour une séquence codante de ladite banque.

- 5 Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

De préférence, ledit récipient est un puit de microplaque ou équivalent.

10

SELECTION D'ENDONUCLEASES

- La sélection permet d'identifier dans un mélange d'endonucléases, plus précisément une banque de séquences codantes pour des endonucléases, la séquence codante pour une endonucléase présentant l'activité de coupure désirée. Plus particulièrement, la sélection permet
- 15 d'enrichir une banque ou un mélange de séquences codantes pour des endonucléases en séquences codantes pour des endonucléases présentant l'activité de coupure désirée.

- La méthode selon la présente invention est bien adaptée pour faire une sélection. Pour cela, chaque substrat ADN porte de préférence au moins deux sites de coupure entourant une cassette d'expression pour une endonucléase. La sélection nécessite que chaque substrat ADN soit
- 20 isolé dans un compartiment durant l'expression et la coupure éventuelle du substrat par l'endonucléase exprimée.

COMPARTIMENTATION

- La méthode de compartimentation préférée pour la présente invention est la technique
- 25 d'émulsion (Tawfik et Griffiths, 1998, *Nat Biotechnol*, **16**, 652-656 ; Griffiths et Tawfik, 2000, *Cur Opin Biotechnol*, **11**, 338-353 ; Griffiths et Duncan, 1998, *Cur Opin Biotechnol*, **9**, 102-108 ; WO 99/02671 (plus particulièrement p14-19 et p37-38; l'enseignement de ces documents étant incorporé par référence). Une émulsion est un mélange hétérogène de deux liquides non miscibles dont l'un forme des gouttelettes microscopiques en suspension dans l'autre. Cependant, toute
- 30 méthode équivalente de compartimentation est également envisagée dans la présente invention (WO 99/02671 p13-14 ; Benita, 1996, *Microencapsulation : methods and industrial applications. Drug and pharmaceutical sciences*. Ed. Swarbrick, J. New York : Marcel Dekker ; l'enseignement de ces documents étant incorporé par référence). Dans la présente invention, les termes « goutte », « compartiment », « microcapsule » sont interchangeables.

- 35 Une émulsion peut être préparée à partir d'une combinaison adéquate de liquides non miscibles. De préférence, l'émulsion présente un milieu aqueux dans les gouttes formées et un

liquide non miscible hydrophobe (une huile) comme matrice dans laquelle les gouttes sont suspendues. Ce type d'émulsion présente l'avantage que la phase aqueuse contenant les composants biochimiques est compartimentée dans des gouttes séparées. La phase externe (huile hydrophobe) ne contient aucun composant biochimique et est inerte. Les microcapsules aqueuses
5 sont compatibles avec des procédés biochimiques complexes tels que la transcription et la traduction.

L'émulsion peut être stabilisée par un ou plusieurs surfactants. Ces surfactants sont appelés agents émulsifiants et agissent sur l'interface eau/huile pour éviter la séparation des deux phases en deux phases uniques. Une grande variété d'huiles et d'émulsifiants peuvent être utilisés
10 pour la préparation d'émulsion. L'huile préférée est l'huile minérale (Sigma ; ref M-3516). Les émulsifiants adéquats comprennent le monooléate de sorbitan (Span 80, Fluka ; ref 85548) ou de polyoxyéthylènesorbitan (Tween 80, Sigma Ultra ; ref P-8074). Par exemple, le mélange initial comprend l'huile minérale avec 4,5 % de Span 80 (vol/vol) et 0,5 % de Tween 80 (vol/vol).

L'utilisation de surfactants anioniques peut être bénéfique. Les surfactants adéquats
15 incluent le cholate de sodium et le taurocholate de sodium, de préférence le desoxycholate de sodium (Fluka ; ref 30970). L'ajout de ces surfactants peut parfois augmenter l'expression et/ou l'activité des enzymes. L'addition de surfactants anioniques à un mélange non-émulsifié abolit complètement la traduction. Pendant l'émulsification, le surfactant est transféré de la phase aqueuse vers l'interface et l'activité est restaurée. L'addition de surfactant anionique au mélange
20 assure que la réaction ne commence pas avant la compartimentation.

La création d'une émulsion nécessite généralement l'application d'une énergie mécanique pour forcer les phases à se mélanger. De nombreux outils mécaniques peuvent être utilisés, par exemple agitateur, homogénéisateur, réducteur colloïdal, sonicateur, outil d'émulsification de membrane. De préférence, l'émulsion est préparée à l'aide d'un agitateur magnétique. Par
25 exemple, on peut utiliser une constante d'agitation de 1150 rpm.

La taille préférée des compartiments varie selon la taille de la banque et la concentration nécessaire des composés dans le compartiment pour effectuer l'expression et la réaction de clivage du substrat ADN. De préférence, le volume moyen des compartiments est inférieur à $5,2 \cdot 10^{-16} \text{ m}^3$ (diamètre inférieur à $10 \text{ }\mu\text{m}$), plus particulièrement inférieur à $6,5 \cdot 10^{-17} \text{ m}^3$ (diamètre inférieur à $5 \text{ }\mu\text{m}$). De préférence, les compartiments présentent un volume d'environ $4 \text{ à } 9 \cdot 10^{-18} \text{ m}^3$ (diamètre d'environ $2 \text{ à } 2,6 \text{ }\mu\text{m}$).
30

La taille des compartiments peut être changée simplement par contrôle des conditions d'émulsification utilisées. Plus le compartiment est grand, plus le volume nécessaire pour compartimenter le substrat ADN est grand. Le rapport phase aqueuse/phase huile est, par exemple,
35 1/19.

Le procédé de compartimentation doit être adapté de telle sorte que chaque compartiment doit contenir une molécule de substrat ADN ou moins. Eventuellement, si la banque de séquences codantes pour une endonucléase est très large ou trop complexe pour avoir une molécule de substrat ADN par compartiment, plusieurs substrats ADN peuvent être contenus dans un même

5 compartiment lors d'un premier tour de sélection.

L'émulsion peut être rompue par divers moyens connus de l'homme du métier. De préférence, elle est rompue par centrifugation (par exemple 3000 G pendant 5 min) suivie d'une extraction. Cette extraction peut être faite par l'éther, par phénol/chloroforme ou par huile minérale pure. Ces méthodes sont bien connues de l'homme du métier.

10 Plus précisément, après la première centrifugation, la phase huile est enlevée. Un mélange éther/eau, de préférence éther/tampon Tris NaCl (rapport 10/1 par exemple), est ajouté. Le mélange est agité puis centrifugé. La phase éther est enlevée. Une deuxième extraction par l'éther peut éventuellement être faite. Facultativement, le culot est évaporé par speed vac.

Dans une autre alternative, après la première centrifugation, un grand volume d'huile

15 minérale pure est ajouté. Le mélange est agité puis centrifugé. Ce procédé est moins violent. Il permet d'extraire les surfactants/émulsifiant jusqu'à ce que les microcapsules fusionnent en une unique phase aqueuse.

SELECTION D'ENDONUCLEASES

20 L'invention concerne une méthode de sélection des endonucléases capables de couper un site de coupure donné parmi une banque de séquences codantes pour des endonucléases, ladite méthode de sélection étant basée sur la méthode d'observation de l'activité d'une endonucléase selon la présente invention en combinaison de la compartimentation. Une méthode de sélection selon la présente invention est illustrée dans la FIGURE 6.

25 Dans un premier temps, une banque de substrats ADN présentant au moins deux sites de coupure encadrant une cassette d'expression pour une endonucléase de la banque est préparée. La banque de substrats ADN peut être préparé par amplification par PCR en utilisant deux amorces contenant le site de coupure à tester. La première amorce contient une séquence lui permettant de s'hybrider en amont de la séquence codante pour l'endonucléase. La deuxième amorce s'hybride

30 en aval de la séquence codante. Les méthodes d'obtention de ce substrat ont été décrites plus en détail ci-dessus. De préférence, le substrat ADN comprend deux sites de coupure placés aux extrémités de ce substrat. Ces sites sont de préférence identiques et présentent une orientation inverse. Cette structure de substrat permet de régénérer facilement le substrat ADN par amplification par PCR après un premier tour de sélection.

35 Une émulsion est ensuite préparée de telle sorte que chaque compartiment comprenne une molécule de substrat et les composants nécessaires pour la transcription/traduction couplée *in*

vitro. De préférence, la préparation de l'émulsion est faite dans des conditions ne permettant pas la production de l'endonucléase lors de cette phase de préparation. Par exemple, elle est maintenue à basse température. Par exemple, l'émulsion est préparée en ajoutant graduellement 50 µl de mélange comprenant lesdits substrats ADN et un extrait de transcription/traduction *in vitro* à 950
5 µl du mélange huile/émulsifiants froid sous une agitation constante de 1150 rpm.

L'endonucléase à tester est produite au sein du compartiment par transcription/traduction couplée *in vitro* et peut agir sur le substrat. Les conditions d'expression doivent de préférence permettre l'activité des endonuclease. Les compartiments de l'émulsion sont ensuite rompus et les endonuclease sont inactivées. Les substrats ADN sont ensuite purifiés. L'ensemble des substrats
10 ADN est soumis à l'étape de discrimination. La discrimination peut être faite par séparation physique des deux types de substrats, par dégradation du substrat non coupé, par amplification sélective des substrats coupés aux deux sites, ou une combinaison de ces méthodes. Les substrats ADN présentant deux sites coupés peuvent être détectés directement ou après amplification par PCR. Cependant, dans le cas où plusieurs cycles de sélection doivent se succéder, ces substrats
15 ADN présentant deux sites coupés sont amplifiés par PCR de manière à régénérer les substrats originaux. Pour cela, les amorces contiennent la séquence du demi-site absent ainsi que la séquence du demi-site restant sur le substrat ADN coupé.

Le nombre de tours en cycles de sélection nécessaire dépend généralement de la taille de la banque de départ. Le nombre de tours de sélection est de préférence compris entre 1 et 10, plus
20 particulièrement entre 2 et 5. Un cycle de sélection comprend : 1- préparation du substrat adapté pour l'expression ; 2- compartimentation ; 3- expression/activité ; 4- discrimination.

En conséquence, l'invention concerne une méthode de sélection d'endonuclease capables de couper un site de coupure parmi une banque de séquences codantes pour des endonuclease, présentant un type de coupure décalée, comprenant les étapes suivantes :

- 25 1) fournir ladite banque de séquences codantes pour des endonuclease ;
- 2) préparer une banque de substrats ADN double-brins comprenant deux sites de coupure encadrant une cassette d'expression pour une endonuclease de ladite banque ;
- 3) préparer un mélange comprenant lesdits substrats ADN et un extrait de transcription/traduction *in vitro* ;
- 30 4) compartimenter ledit mélange dans des microcapsules ;
- 5) mettre dans des conditions propices à l'expression de l'endonuclease et à son activité de coupure ;
- 6) rompre la compartimentation ;
- 7) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés
35 aux deux sites; et, de manière facultative,
- 8) amplifier le substrat ADN coupé aux deux sites afin de restaurer le substrat ADN initial;

9) reproduire les étapes 3) à 8) jusqu'à un stade de sélection suffisant.

L'avancement de la sélection est généralement déterminé par la quantification du taux d'enrichissement à l'aide d'une méthode de criblage. La méthode de criblage peut éventuellement être une des méthodes de la présente invention.

5 Un stade de sélection est suffisant lorsque le rapport taux d'enrichissement par rapport à la diversité est correct. En effet, la sélection doit permettre d'enrichir la population en endonucléase présentant l'activité de coupure désirée tout en gardant une bonne diversité. De préférence, la sélection est suffisante lorsque la diversité a été réduite à une taille compatible pour le criblage. Par exemple, une sélection est suffisante lorsque les endonucléases capables de couper le site testé
10 représentent plus de 5% de la population, de préférence plus de 10 %, de manière privilégiée plus de 20%. De préférence, une sélection est suffisante lorsqu'une endonucléase représente plus de 5% de la population, de préférence plus de 10 %, de manière privilégiée plus de 20%.

Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus
15 particulièrement, ils sont situés aux extrémités du substrat ADN. Lesdits sites de coupure sont de préférence identiques. Lorsqu'ils sont identiques, ils présentent une orientation inverse. Lesdits sites de coupure peuvent également être différents.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de
20 préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

Dans un mode de réalisation préféré, la discrimination consiste en la dégradation du substrat ADN non coupé par une exonucléase incapable de dégrader le substrat ADN coupé aux deux sites. Facultativement, ladite exonucléase est inactivée avant l'étape 8). Lorsque ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides,
25 générant une extrémité 3' sortante, ladite exonucléase est de préférence l'exonucléase III. Lorsque ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 5' sortante, ladite exonucléase est de préférence l'exonucléase Lambda.

Dans un autre mode de réalisation, la discrimination consiste en la séparation des substrats
30 ADN coupés aux deux sites et non coupés aux deux sites. De préférence, ledit substrat ADN initial porte à au moins une de ses extrémités, de préférence aux deux, un ligand et ladite séparation est faite avec la cible du ligand immobilisée sur un support solide. De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Dans un mode de réalisation additionnel, la discrimination consiste en l'amplification
35 sélective des substrats ADN coupés aux deux sites. Ainsi, des adaptateurs capables de s'hybrider à au moins une des extrémités sortantes générées par la coupure de ladite endonucléase sur le

substrat ADN coupé, de préférence aux deux extrémités, sont ajoutés et soit liés covalamment par l'action d'une ligase soit intégrés au substrat ADN par élongation, lesdits adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN utilisée pour les amorces de PCR. Ladite séquence non présente sur le substrat ADN varie à chaque tour de sélection.

- 5 Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 3).

De préférence, ladite compartimentation est faite par une émulsion eau/huile. Dans ce mode de réalisation, la rupture de la compartimentation est faite par centrifugation suivie d'une extraction.

- 10 La compartimentation peut amener à devoir déterminer et/ou optimiser les conditions expérimentales nécessaires à l'expression et/ou l'activité de l'endonucléase.

Les séquences codantes pour les endonucléases présentant les caractéristiques requises et issues de la sélection sont ensuite clonées et séquencées. De préférence, ces séquences codantes sont soumises à une méthode de criblage d'endonucléase.

15

Dans le mode de réalisation alternatif où le substrat contient un site de coupure, l'invention concerne une méthode de sélection d'endonucléases capables de couper un site de coupure parmi une banque de séquences codantes pour des endonucléases, présentant un type de coupure décalée, comprenant les étapes suivantes :

- 20 1) fournir ladite banque de séquences codantes pour des endonucléases ;
- 2) préparer une banque de substrats ADN double-brins contenant un site de coupure présentant une localisation distale sur ledit substrat ADN et bordant une cassette d'expression pour une endonucléase de ladite banque, lesdits substrats portant un ligand à l'extrémité proche dudit site de coupure ;
- 25 3) préparer un mélange comprenant lesdits substrats ADN et un extrait de transcription/traduction *in vitro* ;
- 4) compartimenter ledit mélange dans des microcapsules ;
- 5) mettre dans des conditions propices à l'expression de l'endonucléase et à son activité de coupure ;
- 30 6) rompre la compartimentation ;
- 7) dépler les substrats ADN non coupés avec la cible du ligand immobilisée sur un support solide; et, de manière facultative,
- 8) amplifier le substrat ADN coupé afin de restaurer le substrat ADN initial;
- 9) reproduire les étapes 3) à 8) jusqu'à un stade de sélection suffisant.
- 35 De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

La présente invention concerne également une méthode de sélection d'endonucléases capables de couper un site de coupure parmi une banque de séquences codantes pour des endonucléases, présentant un type de coupure décalée, comprenant les étapes suivantes :

- 1) fournir ladite banque de séquences codantes pour des endonucléases ;
- 5 2) préparer une banque de substrats ADN double-brins contenant un site de coupure présentant une localisation distale sur ledit substrat ADN et bordant une cassette d'expression pour une endonucléase de ladite banque;
- 3) préparer un mélange comprenant lesdits substrats ADN et un extrait de transcription/traduction *in vitro* ;
- 10 4) compartimenter ledit mélange dans des microcapsules ;
- 5) mettre dans des conditions propices à l'expression de l'endonucléase et à son activité de coupure ;
- 6) rompre la compartimentation ;
- 7) dépler les substrats ADN non coupés avec la cible du ligand immobilisée sur un
- 15 support solide; et, de manière facultative,
- 8) amplifier spécifiquement lesdits substrat ADN coupés avec une amorce s'hybridant sur ladite séquence non présente de manière à restaurer le substrat ADN initial;
- 9) reproduire les étapes 3) à 8) jusqu'à un stade de sélection suffisant.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure
20 décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

SELECTION DE SITES DE COUPURE

25 La méthode d'observation de l'activité de coupure par une endonucléase selon la présente invention peut également être utilisée pour effectuer une sélection des sites de coupure capables d'être coupés par une endonucléase déterminée. Cette sélection permet l'exploration des sites de coupure par demi-site. En effet, la coupure par l'endonucléase du site de coupure entraîne la perte des informations sur la séquence du demi-site qui sera détaché du substrat ADN.

30 Dans un premier mode de réalisation, l'endonucléase testée coupe dans le site de reconnaissance et la sélection de sites de coupure est pratiquée soit sur une banque des substrats ADN comprenant un site constant connu pour être coupé par l'endonucléase testée ou par une autre endonucléase et un autre site composé d'un demi-site variable et d'un demi-site constant. Dans un deuxième mode de réalisation, l'endonucléase testée coupe dans le site de reconnaissance
35 et la sélection de sites de coupure est pratiquée soit sur une banque des substrats ADN comprenant deux sites composés d'un demi-site variable et d'un demi-site constant. Le demi-site constant a

une position distale sur le substrat. En effet, il est placé vers l'extrémité du substrat ADN. Le demi-site variable présente une position proximale sur le substrat. Il est placé vers l'intérieur du substrat ADN de telle sorte que le clivage par l'endonuclease fournit un substrat ADN conservant le demi-site variable. De préférence le demi-site constant est un demi-site dérivé d'un site de
5 coupure connu pour être coupé. De préférence, le substrat contient deux sites. Les sites ont de préférence une location distale sur le substrat ADN. Plus particulièrement, ils sont placés aux extrémités du substrat ADN.

Dans un troisième mode de réalisation, l'endonuclease testée coupe en dehors de son site de reconnaissance. Dans ce cas, les substrats ADN de la banque comprennent un ou deux sites
10 variables. Ces sites peuvent être partiellement ou totalement variables. De préférence, la position du site de reconnaissance est proximale sur le substrat par rapport au site de clivage.

Les substrats ADN peuvent également comprendre un gène de résistance à un antibiotique et/ou une origine de répllication.

Les diverses demi-sites de coupure peuvent présenter une séquence déterminée, ou une
15 séquence partiellement ou complètement aléatoire.

De nombreuses techniques permettant de préparer cette diversité de demi-sites de coupure sont disponibles pour l'homme du métier. Par exemple, les sites peuvent être préparés par des méthodes de synthèse d'oligonucléotides. La diversité est introduite en insérant des mélanges de plusieurs nucléotides à des positions choisies lors de la synthèse. Ces oligonucléotides
20 correspondant aux demi-sites de coupure sont ensuite clonés. Ils sont ensuite insérés dans un substrat ADN pour la préparation de la banque de substrats ADN avec le ou les demi-site(s) variable(s). La banque de substrats ADN comprenant un ou deux demi-site(s) variable(s) peut être préparé avec une amplification par PCR de matrices comprenant les demi-sites variables. Facultativement, ces matrices peuvent également comprendre une cassette d'expression pour
25 l'endonuclease testée. Les amorces sont conçues de telle manière que les fragments d'amplification comprennent les deux sites de coupure.

Dans une alternative où le substrat comprend un site de coupure constant, le site de coupure constant peut être prédigéré avant la mise en présence de la banque de substrats ADN avec l'endonuclease testée. Dans ce cas, le site de coupure constant est clivé par une endonuclease
30 autre que celle testée.

La banque de substrats ADN est incubée en présence de l'endonuclease testée dans des conditions propices à l'activité de coupure de l'endonuclease. Lorsque les substrats ADN portent un site de coupure constant pour une autre endonuclease, cette autre endonuclease doit également être ajoutée. Si les substrats portent une cassette d'expression pour l'endonuclease, cette
35 endonuclease peut être produite par un système d'expression transcription/traduction couplée *in vitro*.

Après cette étape de coupure, les substrats ADN coupés et non coupés doivent être discriminés. La méthode de discrimination peut être choisie parmi la dégradation différentielle ou la déplétion du substrat ADN non coupé (FIGURE 7B et C), ou l'amplification sélective du substrat ADN coupé (FIGURE 7D). La discrimination peut également être basée sur une
5 combinaison de ces méthodes.

Les substrats ADN sont ensuite amplifiés. De préférence, cette amplification permet de restaurer les sites de coupure initiaux. Ainsi, si nécessaire, il peut être procédé à un nouveau tour de sélection à partir de ces substrats ADN régénérés.

La FIGURE 7 illustre des méthodes de sélection sur une banque de sites de coupure.

10 Lorsque la discrimination est faite par dégradation différentielle, l'exonucléase est ajoutée après l'incubation des substrats ADN en présence de la ou des endonucléase(s) et l'échantillon est placé dans les conditions appropriées à la dégradation sélective du substrat ADN non coupé par l'endonucléase. De préférence, l'exonucléase est ensuite inactivée, par exemple par la chaleur. Les substrats ADN coupés sont ensuite amplifiés par PCR de façon à régénérer les sites de coupure
15 initiaux.

Lorsque la discrimination est faite par déplétion du substrat ADN non coupé, le substrat ADN initial porte à au moins une de ses extrémités, de préférence aux deux, un ligand (par exemple, la biotine). Après l'incubation du substrat ADN en présence de la ou des endonucléase(s), l'échantillon est mis en présence de la cible du ligand immobilisée sur un support
20 solide (par exemple, bille paramagnétique portant la streptavidine). Le support solide portant le substrat ADN non coupé est séparé du reste de l'échantillon. Les substrats ADN coupés sont ensuite amplifiés par PCR de façon à régénérer les sites de coupure initiaux.

La méthode d'amplification après l'étape de discrimination par digestion différentielle ou déplétion des substrats non coupés dépend des extrémités générées par l'endonucléase testée. En
25 effet, il est nécessaire d'utiliser des adaptateurs pour obtenir la régénération des sites de coupure initiaux. La nature de ces adaptateurs dépend des extrémités des substrats coupés générées par l'endonucléase testée.

Lorsque les substrats ADN coupés possèdent des extrémités non cohésives, des adaptateurs, de préférence double-brins et capables de s'hybrider aux extrémités cohésives
30 générées par la coupure de l'endonucléase sont ajoutés et liés covalamment aux substrats ADN par l'action d'une ligase. Ces adaptateurs contiennent la séquence du demi-site constant qui a été éliminé lors du clivage de l'endonucléase testée. Ainsi, ils permettent de régénérer le site initial. Des amorces s'hybridant sur l'adaptateur, de préférence sur le demi-site constant, permettent d'effectuer l'amplification par PCR.

35 Lorsque les substrats ADN coupés possèdent des extrémités 5' sortantes, les adaptateurs sont des oligonucléotides simple-brins capables de s'hybrider aux extrémités sortantes générées

par la coupure de l'endonucléase et présentant un phosphate à son extrémité 5'. Une ligase crée des liaisons covalentes entre l'extrémité 5' de l'adaptateur et l'extrémité 3' rentrante des substrats ADN coupés. Les adaptateurs contiennent la séquence du demi-site constant qui a été éliminé lors du clivage de l'endonucléase testée qui s'hybride, de préférence avec les amorces utilisées pour
5 l'amplification des substrats ADN coupés. L'amplification par PCR doit être faite de préférence dans des conditions (principalement des conditions de température) qui ne permettent pas l'hybridation des adaptateurs entre eux. (FIGURE 7B)

Lorsque les substrats ADN coupés possèdent des extrémités 3' sortantes, les adaptateurs utilisés sont des oligonucléotides simple-brins portant un didéoxynucléotide à son extrémité 3' et
10 dont les extrémités 3' sont complémentaires des extrémités 3' sortantes des substrats ADN. Une extension des extrémités 3' à basse température est ensuite faite en utilisant l'adaptateur comme matrice. Cette extension permet l'introduction de la séquence du demi-site constant qui a été éliminé lors du clivage de l'endonucléase testée sur laquelle s'hybride, de préférence, l'amorce utilisée pour l'amplification des substrats ADN coupés. (FIGURE 7C)

15 Dans un mode de réalisation préféré, lorsque la discrimination est faite par déplétion des substrats ADN non coupés, au moins une des amorces utilisées pour l'amplification par PCR, de préférence les deux, porte de préférence le ligand. Ainsi, les substrats ADN régénérés présentent à au moins une de ses extrémités, de préférence aux deux, un ligand.

Lorsque la discrimination est faite par amplification sélective des substrats ADN coupés,
20 les adaptateurs adéquats sont ajoutés après l'incubation du substrat ADN en présence de la ou des endonucléase(s) et ils sont soit liés covalamment au substrat ADN coupé soit intégrés au substrat ADN par élongation. La nature des adaptateurs dépend des extrémités des substrats ADN générées par le clivage de l'endonucléase. Les adaptateurs contiennent la séquence du demi-site constant qui a été éliminé lors du clivage de l'endonucléase testée ainsi qu'une séquence non-présente dans
25 les substrats ADN. L'agencement de ces deux séquences permet à la régénération des sites de coupures initiaux. Les substrats ADN coupés sont spécifiquement amplifiés par PCR en utilisant des amorces reconnaissant les nouvelles séquences introduites par les adaptateurs. La séquence non-présente doit varier à chaque tour de sélection. (FIGURE 7D)

Une méthode de discrimination additionnelle lorsque la coupure des deux sites du substrat
30 ADN par l'endonucléase génère des extrémités cohésives et complémentaires sur les substrats ADN est une circularisation des substrats ADN coupés aux deux sites. Dans cette alternative, les substrats ADN comprennent, outre entre les deux sites de coupure, un gène de résistance à un antibiotique et une origine de répllication. Des cellules (de préférence procaryote, par ex E. coli) sont ensuite transformées par les substrats ADN circularisés. L'étape de transformation constitue
35 également une étape de discrimination entre les substrats ADN coupés aux deux sites et les autres. En effet, les substrats ADN non-coupés aux deux sites ne sont pas capables de se circulariser.

Ainsi, ils restent linéaires. Or l'ADN linéaire présente une capacité très faible de transformation des bactéries. Facultativement, cette étape de transformation peut constituer l'unique étape de discrimination entre les substrats ADN coupés aux deux sites et les autres ou être combinée avec d'autres méthodes de discrimination décrite dans la présente invention.

5 En conséquence, la présente invention concerne une méthode de sélection permettant d'identifier parmi une banque de sites de coupure le ou les demi-site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée comprenant les étapes suivantes :

1) fournir une banque de substrats ADN comprenant deux sites de coupure dont au moins
10 un site présente un demi-site constant et un demi-site variable, ledit demi-site constant étant placé vers l'extrémité du substrat ADN et étant dérivé d'un site de coupure connu pour être coupé par ladite endonucléase ;

2) fournir ladite endonucléase ;

3) mettre en présence ladite endonucléase et ladite banque de substrats ADN dans des
15 conditions propices à l'activité de coupure ;

4) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés ;

5) si nécessaire, amplifier lesdits substrats ADN coupés aux deux sites afin de restaurer les
substrats ADN initiaux ; et,

6) reproduire les étapes 3) à 5) jusqu'à un stade de sélection suffisant.

20 L'avancement de la sélection est généralement déterminé par la quantification du taux d'enrichissement à l'aide d'une méthode de criblage. La méthode de criblage est de préférence une méthode de la présente invention.

Un stade de sélection est suffisant lorsque le rapport taux d'enrichissement par rapport à la diversité diversité est correct. De préférence, la sélection est suffisante lorsque la diversité a été
25 réduite à une taille compatible pour le criblage. Par exemple, une sélection est suffisante lorsqu'un substrat ADN représente plus de 5% de la population, de préférence plus de 10 %, de manière privilégiée plus de 20%. Le nombre de tours en cycles de sélection nécessaire dépend généralement de la taille de la banque de départ. Le nombre de tours de sélection est de préférence compris entre 1 et 10, plus particulièrement entre 2 et 5.

30 L'ordre des étapes 1) et 2) peut être interverti.

Dans un mode de réalisation, ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et l'étape 2) consiste en l'expression de ladite endonucléase. De préférence, ladite endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

Dans un mode de réalisation préféré, ledit substrat ADN contient deux sites de coupure. De préférence, lesdits sites de coupure ont une localisation distale sur ledit substrat ADN. Plus particulièrement, ils sont situés aux extrémités du substrat ADN.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure
5 décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

Facultativement, lesdits substrats comportent un gène de résistance et/ou une origine de répllication.

Dans un mode de réalisation préféré, la discrimination consiste en la dégradation des
10 substrats ADN non coupés par une exonucléase incapable de dégrader les substrats ADN coupés aux deux sites. Facultativement, ladite exonucléase est inactivée avant l'étape 5). Lorsque ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 3' sortante, ladite exonucléase est de préférence l'exonucléase III. Lorsque
15 ladite endonucléase possède un type de coupure décalée, plus particulièrement d'au moins 4 nucléotides, générant une extrémité 5' sortante, ladite exonucléase est de préférence l'exonucléase Lambda.

Dans un autre mode de réalisation, la discrimination consiste en la séparation des substrats ADN coupés aux deux sites et non coupés aux deux sites. De préférence, ledit substrat ADN initial porte à au moins une de ses extrémités, de préférence aux deux, un ligand et ladite séparation est
20 faite avec la cible du ligand immobilisée sur un support solide. De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

Dans un mode de réalisation additionnel, la discrimination consiste en l'amplification sélective des substrats ADN coupés aux deux sites. Ainsi, des adaptateurs capables de s'hybrider aux extrémités sortantes générées par la coupure de ladite endonucléase sur le substrat ADN coupé
25 sont ajoutés et sont soit liés covalement par l'action d'une ligase soit intégrés au substrat par élongation, lesdits adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN utilisée pour les amorces de PCR. Lesdits adaptateurs portent de préférence le demi-site constant qui a été éliminé lors du clivage de l'endonucléase testée. Ladite séquence non présente sur le substrat ADN varie à chaque tour de sélection.

30 L'étape d'amplification est facultative lorsque la discrimination est faite par amplification sélective ou lorsqu'un seul tour de sélection est désiré. Cette étape est nécessaire lorsqu'un deuxième tour de sélection est nécessaire et lorsque la discrimination est effectuée par digestion différentielle ou déplétion des substrats ADN non coupés.

L'étape d'amplification comprend les étapes suivantes : a) ajouter des adaptateurs capables
35 de s'hybrider aux extrémités sortantes générées par la coupure de ladite endonucléase sur le substrat ADN coupé ; b) soit lier covalement lesdits adaptateurs audit substrat coupé par l'action

d'une ligase soit intégrer la séquence des adaptateurs audit substrat coupé par élongation, lesdits adaptateurs permettant d'introduire la séquence du demi-site constant qui a été éliminé lors du clivage de l'endonucléase testée ; et, c) amplifier par PCR avec des amorces capables de s'hybrider sur lesdits adaptateurs.

- 5 Dans une alternative également envisagée, le substrat ADN comprend un site de coupure pour une autre endonucléase que celle testée et cette autre endonucléase est ajoutée à l'étape 3).

Facultativement, une étape de purification du substrat ADN peut être ajoutée avant l'étape de discrimination. Cette étape permet de séparer le substrat ADN des endonucléases utilisées. Cette étape est de préférence effectuée par une extraction Phénol/Chloroforme. Cette étape de
10 purification peut se révéler nécessaire lorsque les endonculéases restent fixées au site de coupure après le clivage.

Des sites de coupure entiers peuvent être déduits de l'étude des deux demi-sites. En effet, lorsque les sites de coupure présentent une orientation directe, il est possible d'étudier les deux demi-sites formant le site complet. A partir d'un substrat ADN comprenant deux sites variables,
15 après digestion du substrat par l'endonucléase testée et ligation, il est possible de préparer un substrat circulaire contenant un site formé des deux demi-sites variables (FIGURE 7E). Il est avantageux d'effectuer un mélange des sites entre les différents substrats par amplification PCR (« shuffling »). Ces substrats ADN circulaires peuvent être utilisés en pool pour une sélection ou séparément pour un criblage de sites de coupure. Les substrats ADN sélectionnés sont ensuite
20 séquencés. La séquence des demi-sites permet une analyse fine des sites de coupure capables d'être coupés par l'endonucléase testée.

Dans le mode de réalisation alternatif où le substrat contient un site de coupure, la présente invention concerne une méthode de sélection permettant d'identifier parmi une banque de sites de
25 coupure le ou les demi-site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir une banque de substrats ADN un site de coupure ayant une localisation distale sur ledit substrat et présentant un demi-site constant et un demi-site variable, ledit demi-site constant étant placé vers l'extrémité du substrat ADN et étant dérivé d'un site de coupure connu
30 pour être coupé par ladite endonucléase, lesdits substrats portant un ligand à l'extrémité proche dudit site de coupure ;
- 2) fournir ladite endonucléase ;
- 3) mettre en présence ladite endonucléase et ladite banque de substrats ADN dans des conditions propices à l'activité de coupure ;
- 35 4) dépléter les substrats ADN non coupés avec la cible du ligand immobilisée sur un support solide;

- 5) amplifier lesdits substrats ADN coupés afin de restaurer les substrats ADN initiaux ; et,
- 6) reproduire les étapes 3) à 5) jusqu'à un stade de sélection suffisant.

De préférence, le ligand, la cible et le support solide sont respectivement la biotine, la streptavidine et une bille paramagnétique.

- 5 La présente invention concerne également une méthode de sélection permettant d'identifier parmi une banque de sites de coupure le ou les demi-site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée comprenant les étapes suivantes :

- 1) fournir une banque de substrats ADN un site de coupure ayant une localisation distale
10 sur ledit substrat et présentant un demi-site constant et un demi-site variable, ledit demi-site constant étant placé vers l'extrémité du substrat ADN et étant dérivé d'un site de coupure connu pour être coupé par ladite endonucléase ;
- 2) fournir ladite endonucléase ;
- 3) mettre en présence ladite endonucléase et ladite banque de substrats ADN dans des
15 conditions propices à l'activité de coupure ;
- 4) ajouter un adaptateur capable de s'hybrider à l'extrémité sortante dudit substrat générée par la coupure de ladite endonucléase et soit le lier covalamment soit l'intégrer audit substrat par élongation, ledit adaptateurs permettant d'introduire une séquence non présente sur le substrat ADN ;
- 20 5) amplifier spécifiquement lesdits substrat ADN coupés avec une amorce s'hybridant sur ladite séquence non présente de manière à restaurer les substrats ADN initiaux ; et,
- 6) reproduire les étapes 3) à 5) jusqu'à un stade de sélection suffisant.

Dans un mode de réalisation préféré, ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides, de préférence de 4 nucléotides. La taille du substrat ADN est de
25 préférence comprise entre 100 et 3000 pb, plus particulièrement entre 300 et 1000 pb.

EXEMPLES

Exemple I: Discrimination par dégradation sélective du substrat ADN non coupé aux 30 **deux sites par une exonucléase.**

Le but de l'expérience était de démontrer qu'il est possible de discriminer des substrats ADN intacts de ceux qui ont été digérés par l'endonucléase I-Sce I, l'endonucléase I-Sce I générant des extrémités 3' sortantes de 4 nucléotides. Ces substrats ADN ont ensuite été amplifiés

par PCR après une digestion enzymatique par l'exonucléase III des substrats ADN laissés intact lors du traitement par I-SceI.

Méthode

5 Un fragment d'ADN double brin correspondant à un fragment de vecteur pET24d+ (Novagen) et flanqué à ses extrémités de sites de restriction I-SceI inversés a été produit par amplification PCR (Accucore, Sigma) à l'aide des amorces Iscepro et Isceter. Ce fragment d'ADN constituait le substrat ADN.

Les conditions de PCR ont été standard dans les exemples. Elles ont été les suivantes : une
10 dénaturation initiale à 92°C pendant 1 min, suivi par 26 cycles (94°C pendant 45 sec; 45°C pendant 45 sec; 72°C pendant 1 min 30), et une extension finale à 72 °C pendant 5 min.

Le fragment obtenu par PCR a été purifié par précipitation éthanol puis resuspendu dans du Tris 10 mM pH 8.

15

Table I

Nom de l'oligonucléotide	Séquence
Iscepro	acgcTAGGGATAACAGGGTAATatagcTAATACGACTCACTATAGG
20 Isceter	acgcTAGGGATAACAGGGTAATatagcCAAAAAACCCCTCAAGACC
Reampsce	ACGCTAGGGATAACAGGGTAAT

Iscepro (SEQ ID No 1) ; Isceter (SEQ ID No 2) ; Reampsce (SEQ ID No 3) ; La séquence en gras
25 correspond au site I-Sce I et celle soulignée à la séquence complémentaire au vecteur.

400 ng du substrat ADN purifié ont ensuite été incubés dans un tampon de coupure I-SceI (Diethanolamine 0,1 M pH 9.5, DTT 2 mM, BSA 0,2 mg/ml, MgCl₂ 0,8 mM) en présence ou en absence (contrôle négatif) de 10 unités de l'endonucléase I-SceI pendant 30 min à 37°C dans 20
30 µl. Comme contrôle supplémentaire, la réaction a aussi été effectuée en absence de MgCl₂, nécessaire à l'activité d'I-SceI.

2 µl de cette digestion ont été directement incubés dans un tampon de digestion Exo III d'E. coli (NE Biolabs) dans 10 µl final en présence ou en absence (contrôle négatif) de l'enzyme Exo III (100 U) à 37°C pendant 10 min. La réaction a été arrêtée par congélation dans l'azote
35 liquide, suivi d'une dénaturation pendant 10 min à 65 °C.

Finalement 1 µl de ce mélange de réaction a été utilisé comme matrice dans une réaction de PCR (Accucore, Sigma) dans 25 µl final, en utilisant l'amorce Reampsce hybridant sur la partie restante du site I-SceI après digestion.

10 µl de chaque PCR ont été chargés sur un gel agarose 1% et colorés au BET.

5

Résultats

Comme montré sur la Figure 8, toutes les conditions d'incubation ont donné lieu à une amplification quand l'étape d'incubation avec l'ExoIII est omise. Après une étape d'incubation avec l'ExoIII, seules les réactions faites en présence de I-SceI ont donné lieu à une réamplification par PCR efficace. Il est intéressant de remarquer que la PCR était légèrement positive lorsque le MgCl₂ avait été omis dans la réaction de digestion. Ceci suggère que cette méthode de détection est capable de mettre en évidence une activité résiduelle d'I-SceI en présence d'une très faible concentration de Mg²⁺ co-purifié avec I-SceI ou une activité d'I-SceI durant l'étape d'incubation avec l'ExoIII (le tampon ExoIII contenant du MgCl₂).

15 On peut donc conclure que la réamplification du substrat ADN dépend de l'activité de l'endonuclease I-SceI qui génère des extrémités 3' sortantes de 4 paires de bases. Ce type d'extrémités est pratiquement insensible à l'action de l'exoIII.

Exemple II: Expression de l'endonuclease I-SceI dans un extrait de transcription/traduction in vitro et protection du gène servant à programmer l'extrait par 20 génération d'extrémité 3' sortante

Le but de l'expérience était de démontrer, en utilisant une dégradation exoIII suivie d'une PCR, que la protéine I-SceI pouvait être produite dans un extrait de transcription/traduction couplée *in vitro* et était capable de digérer son substrat dans ces conditions.

Méthode

Le gène I-SceI (SEQ ID No 4) a été cloné dans le vecteur d'expression pET24d+ (Novagen) en utilisant les sites de restriction NcoI et EagI.

30 Ce vecteur pET24d+Iscel a ensuite été utilisé comme matrice dans une amplification par PCR utilisant les amorces iscepro et isceter (voir table I). Cette amplification a produit un fragment d'ADN correspondant au gène I-SceI, sous la dépendance du promoteur T7 et comprenant un site de liaison au ribosome et un terminateur, et flanqué à ses extrémités de sites de restriction I-SceI inversés (voir figure 9). Ce fragment d'ADN constituait le substrat ADN.

Le fragment obtenu a été purifié par précipitation éthanol/NaAc puis resuspendu dans du Tris 10 mM pH 8.

400 ng de ce substrat ADN ont été utilisés pour programmer 25 µl d'un extrait de transcription/traduction in vitro (RTS100, Roche). Le mélange a été incubé pendant 1h à 30°C
5 pour permettre l'expression de I-SceI puis 30 min à 37°C de façon à permettre à l'endonucléase de digérer les sites du substrat ADN.

En contrôle, ce même substrat a été incubé dans le tampon de digestion d'I-SceI contenant du MgCl₂, en présence ou en l'absence de 10 unités de la protéine I-SceI purifiée indépendamment.

10 L'ADN résultant de ces différentes réactions a été ensuite purifié par une extraction phénol:chloroforme:isoamylalcool suivi d'une précipitation éthanol/sodiumAcétate, puis resuspendu dans 25 µl de Tris 10 mM pH 8.

2 µl de chaque préparation ont directement été incubés dans un tampon de digestion ExoIII d'E. coli (NE Biolabs) dans 10 µl final en présence de l'enzyme ExoIII (100 u) à 37°C pendant 10
15 min. La réaction a été arrêtée par congélation dans l'azote liquide, suivi d'une dénaturation pendant 10 min à 65 °C.

Finalement 1 µl de ce mélange de réaction a été utilisé comme matrice dans une réaction de PCR (Accucore, Sigma) dans 25 µl final, en utilisant l'amorce Reampsce hybridant sur la partie restante du site I-SceI après digestion.

20 10 µl de chaque PCR ont été chargés sur un gel agarose 1% et colorés au BET.

Résultats

Comme démontré sur la figure 10, la PCR n'a été capable de réamplifier efficacement le substrat ADN que dans le cas où la protéine I-SceI était présente, soit après addition de
25 l'endonucléase (puits 3), soit lors d'une production de l'endonucléase dans un extrait RTS (puits 4). L'ajout de 2 mM de MgCl₂ dans l'extrait n'a pas augmenté l'efficacité de la réaction (puits 5). L'endonucléase I-SceI a donc été produite dans l'extrait RTS et a été capable de digérer les sites de restriction du substrat ADN. Cette digestion par l'endonucléase a protégé le substrat ADN d'une dégradation par l'exonucléase ExoIII.

30

Exemple III: Discrimination par déplétion sur bille magnétique du substrat ADN non coupé aux deux sites par une exonucléase

Le but de l'expérience était de démontrer qu'il est possible de discriminer un fragment d'ADN biotinylé à ses deux extrémités d'un autre non biotinylé et d'enrichir ce dernier dans un mélange par déplétion sur billes superparamagnétiques recouvertes de streptavidine.

5 Méthode

Deux fragments de taille respective 1300 pb et 800 pb appelés bSSS et SCS ont été préparés par PCR en utilisant les amorces iscepro et isceter (Table 1). Le fragment SCS a été préparé en utilisant des amorces biotinylées.

Une quantité constante de fragment biotinylé de 1300 pb bSSS (500 ng) a été mélangée à
10 différentes quantités de fragments non biotinylés SCS (500 ng, 50 ng, 5 ng, 0.5 ng, 0.05 ng) dans un volume final de 100 µl de PBS. 2µl de chaque mélange ont été prélevés. Le reste a été incubé avec l'équivalent de 300 µl de billes magnétique recouvertes de streptavidine (DYNAL) préalablement lavées avec du PBS contenant 0,1% de Tween20, pendant 1h à 37°C.

2µl de chaque surnageant, ainsi que les 2 µl d'échantillons prélevés avant la déplétion ont
15 été utilisés comme matrice dans une réaction de PCR de 25 µl (Accucore) en utilisant l'amorce Reampsce. 5 µl de chaque PCR ont été chargés sur gel d'agarose 1% et colorés au BET.

Résultats

Comme démontré sur la figure 11A, un rapport bSSS:SCS 1:1 a donné lieu à une
20 amplification équivalente de chaque fragment (puits 1). Cependant un rapport bSSS:SCS 100:1 (puits 3) n'a donné lieu qu'à une amplification du fragment bSSS. Après déplétion, les premiers rapports (1:1, 10:1, 100:1, respectivement puits 5, 6 et 7) n'ont donnée lieu qu'à une amplification du fragment SCS et il fallait employer un rapport entre 1000:1 et 10000:1 pour obtenir à nouveau une amplification équivalente de chaque fragment (puits 8 et 9). En contrôle, une expérience
25 identique mais en l'absence de la quantité constante du fragment bSSS a démontré que la PCR était positive jusqu'à une dilution 10^{-5} du fragment SCS (figure 11B, puits 6 et 12) et que l'incubation sur bille streptavidine ne changeait rien dans ce cas là (le fragment n'étant pas biotinylé).

Ceci démontre donc qu'il est possible de discriminer entre un fragment biotinylé et un
30 fragment ayant perdu les extrémités biotinylées (après digestion de sites de restriction placés aux extrémités par exemple) par déplétion sur bille de streptavidine, et que l'enrichissement obtenu par cette méthode est de l'ordre de 5000.

(Dans les exemples, µl et µg désignent micro-litre et micro-gramme.)

REVENDICATIONS

1- Méthode permettant de tester l'activité de clivage d'une endonucléase présentant un type de coupure décalée, comprenant les étapes suivantes :

- 5 1) fournir un substrat ADN double-brin comprenant deux sites de coupure;
 2) fournir ladite endonucléase ;
 3) mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
 4) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés
10 aux deux sites; et,
 5) détecter les substrats ADN coupés aux deux sites;
 la détection desdits substrats ADN coupés révélant l'activité de clivage de ladite endonucléase.

- 15 2- Méthode de criblage permettant d'identifier un ou des site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée, comprenant les étapes suivantes :

- 1) fournir une banque de substrats ADN double-brins comprenant chacun deux sites de coupure, au moins un des sites de coupure étant variable dans la banque;
20 2) fournir ladite endonucléase ;
 3) distribuer les différents substrats ADN dans des récipients et mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;
 4) discriminer, dans chaque récipient, les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,
25 5) détecter les substrats ADN coupés aux deux sites;
 la présence des substrats ADN coupés révélant l'activité de clivage et permettant d'identifier un ou des sites de coupure coupé(s) par ladite endonucléase.

- 3- Méthode de criblage permettant d'identifier une ou des endonucléase(s) présentant un
30 type de coupure décalée capable(s) de couper un site de coupure déterminé, comprenant les étapes suivantes :

- 1) fournir une banque de séquences codant pour des endonucléases ;
 2) fournir un substrat ADN double-brin comprenant deux sites de coupure déterminés;
 3) distribuer chaque séquence codante de ladite banque dans des récipients;
35 4) exprimer l'endonucléase codée par les séquences codantes et mettre en présence ladite endonucléase et ledit substrat ADN dans des conditions propices à l'activité de coupure ;

5) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et,

6) détecter les substrats ADN coupés aux deux sites;

la détection desdits substrats ADN coupés révélant l'activité de clivage et permettant
5 d'identifier une ou des endonucléase(s) capable(s) de couper lesdits sites de coupure.

4- Méthode de sélection d'endonucléases capables de couper un site de coupure déterminé, comprenant les étapes suivantes :

1) fournir une première banque comprenant des séquences codant pour des endonucléases
10 présentant un type de coupure décalée;

2) préparer une banque de substrats ADN double-brins comprenant deux sites de coupure déterminés encadrant une cassette d'expression pour une endonucléase de ladite première banque ;

3) compartimenter ladite banque de substrats;

4) mettre ladite banque compartimentée dans des conditions propices à l'expression de
15 l'endonucléase et à l'activité de coupure ;

5) rompre la compartimentation ;

6) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés aux deux sites; et, de manière facultative,

7) amplifier les substrats ADN coupés aux deux sites afin de restaurer le substrat ADN; et

20 8) reproduire les étapes 3) à 7) jusqu'à un stade de sélection suffisant.

5- Méthode de sélection permettant d'identifier un ou des demi-site(s) de coupure coupé(s) par une endonucléase déterminée présentant un type de coupure décalée, comprenant les étapes suivantes :

25 1) fournir une banque de substrats ADN comprenant deux sites de coupure dont au moins un site présente un demi-site constant et un demi-site variable, ledit demi-site constant étant placé vers l'extrémité du substrat ADN et étant dérivé d'un site de coupure connu pour être coupé par ladite endonucléase ;

2) fournir ladite endonucléase ;

30 3) mettre en présence ladite endonucléase et ladite banque de substrats ADN dans des conditions propices à l'activité de coupure ;

4) discriminer les substrats ADN coupés aux deux sites et les substrats ADN non coupés; et, si nécessaire

5) amplifier lesdits substrats ADN coupés aux deux sites afin de restaurer les substrats
35 ADN initiaux ; et,

6) reproduire les étapes 3) à 5) jusqu'à un stade de sélection suffisant.

6- La méthode selon l'une des revendications 1, 2 et 5 caractérisée en ce que ledit substrat comprend en outre une cassette d'expression pour ladite endonucléase et en ce que les étapes 2) et 3) consistent en l'expression de ladite endonucléase.

5

7- La méthode selon la revendication 3 caractérisée en ce que ledit substrat ADN comprend en outre une cassette d'expression pour une séquence codante de ladite banque.

8- La méthode selon l'une des revendications 3, 6 et 7 caractérisée en ce que ladite
10 endonucléase est produite par un système de transcription/traduction couplée *in vitro*.

9- La méthode selon l'une des revendications 1, 3 et 4 caractérisée en ce que lesdits deux sites de coupure du substrat sont identiques.

15 10- La méthode selon la revendication 9 caractérisée en ce que lesdits sites de coupure ont une orientation inverse.

11- La méthode selon l'une des revendications 1 à 5 caractérisée en ce que lesdits sites de coupure du substrat sont différents.

20

12 - La méthode selon l'une des revendications 1 à 5, caractérisée en ce que un des deux sites dudit substrat ADN est un site de coupure pour une endonucléase autre que celle testée et en ce que cette autre endonucléase est ajoutée à l'étape de mise en présence de l'endonucléase testée et dudit substrat ADN.

25

13- La méthode selon l'une des revendications 1 à 12 caractérisée en ce que ledit substrat ADN contient deux sites de coupure.

14- La méthode selon l'une des revendications 1 à 13 caractérisée en ce que lesdits sites de
30 coupure ont une localisation distale sur ledit substrat ADN.

15- La méthode selon la revendication 14 caractérisée en ce que lesdits sites de coupure sont situés aux extrémités du substrat ADN.

35 16- La méthode selon l'une des revendications 1 à 15 caractérisée en que la taille dudit substrat ADN est comprise entre 100 et 3000 pb.

17- La méthode selon la revendication 16 caractérisée en ce que la taille dudit substrat est comprise entre 300 et 1000 pb.

5 18- La méthode selon la revendication 5, caractérisée en ce que ledit substrat ADN comprend en outre un gène de résistance à un antibiotique et une origine de réplication.

19- La méthode selon l'une des revendications 1 à 3 caractérisée en ce que la détection comprend : a) amplifier par PCR le substrat ADN coupé aux deux sites; et, b) détecter le produit
10 d'amplification ; ledit produit d'amplification révélant la capacité de clivage de ladite endonucléase.

20- La méthode selon l'une des revendications 1 à 19 caractérisée en ce que la discrimination comprend la dégradation du substrat ADN non coupé aux deux sites par une
15 exonucléase incapable de dégrader le substrat ADN coupé aux deux sites.

21- La méthode selon la revendication 20 caractérisée en ce que ladite exonucléase est inactivée après l'étape de discrimination.

20 22- La méthode selon l'une des revendications 20 à 21, caractérisée en ce que ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides générant une extrémité 3' sortante et en ce que ladite exonucléase est l'exonucléase III.

23- La méthode selon l'une des revendications 20 à 21, caractérisée en ce que ladite
25 endonucléase possède un type de coupure décalée d'au moins 4 nucléotides générant une extrémité 5' sortante et en ce que ladite exonucléase est l'exonucléase Lambda.

24- La méthode selon l'une des revendications 1 à 19, caractérisée en ce que la discrimination comprend la séparation du substrat ADN coupé aux deux sites et du substrat ADN
30 non coupé aux deux sites.

25- La méthode selon la revendication 24, caractérisée en ce que le substrat ADN initial porte à au moins une de ses extrémités un ligand et en ce que ladite séparation est faite avec la cible du ligand immobilisée sur un support solide.

26- La méthode selon la revendication 25, caractérisée en ce que ledit ligand est la biotine, ladite cible est la streptavidine, et ledit support solide est une bille paramagnétique.

27- La méthode selon l'une des revendications 1 à 18, caractérisée en ce que la
5 discrimination comprend une amplification sélective des substrats ADN coupés aux deux sites.

28- La méthode selon la revendication 27, caractérisée en ce que des adaptateurs capables de s'hybrider à au moins une des extrémités sortantes générées par la coupure du substrat ADN par l'endonucléase sont ajoutés et liés covalamment par l'action d'une ligase, lesdits adaptateurs
10 permettant d'introduire une séquence non présente sur le substrat ADN et utilisée pour les amorces de PCR.

29- La méthode selon l'une des revendications 1 à 28, caractérisée en ce que ladite endonucléase possède un type de coupure décalée d'au moins 4 nucléotides.
15

30- La méthode selon la revendication 29, caractérisée en ce que ladite endonucléase possède un type de coupure décalée de 4 nucléotides.

31- La méthode selon l'une des revendications 2 et 3, caractérisée en ce que ledit récipient
20 est un puit d'une microplaque.

32- La méthode selon la revendication 4, caractérisée en ce que ladite compartimentation est faite par une émulsion eau/huile.

25 33- La méthode selon la revendication 32, caractérisée en ce que la rupture de la compartimentation est faite par centrifugation suivie d'une extraction.

34- La méthode selon la revendication 4, caractérisée en ce que chaque compartiment comprend au plus une molécule de substrat ADN.
30

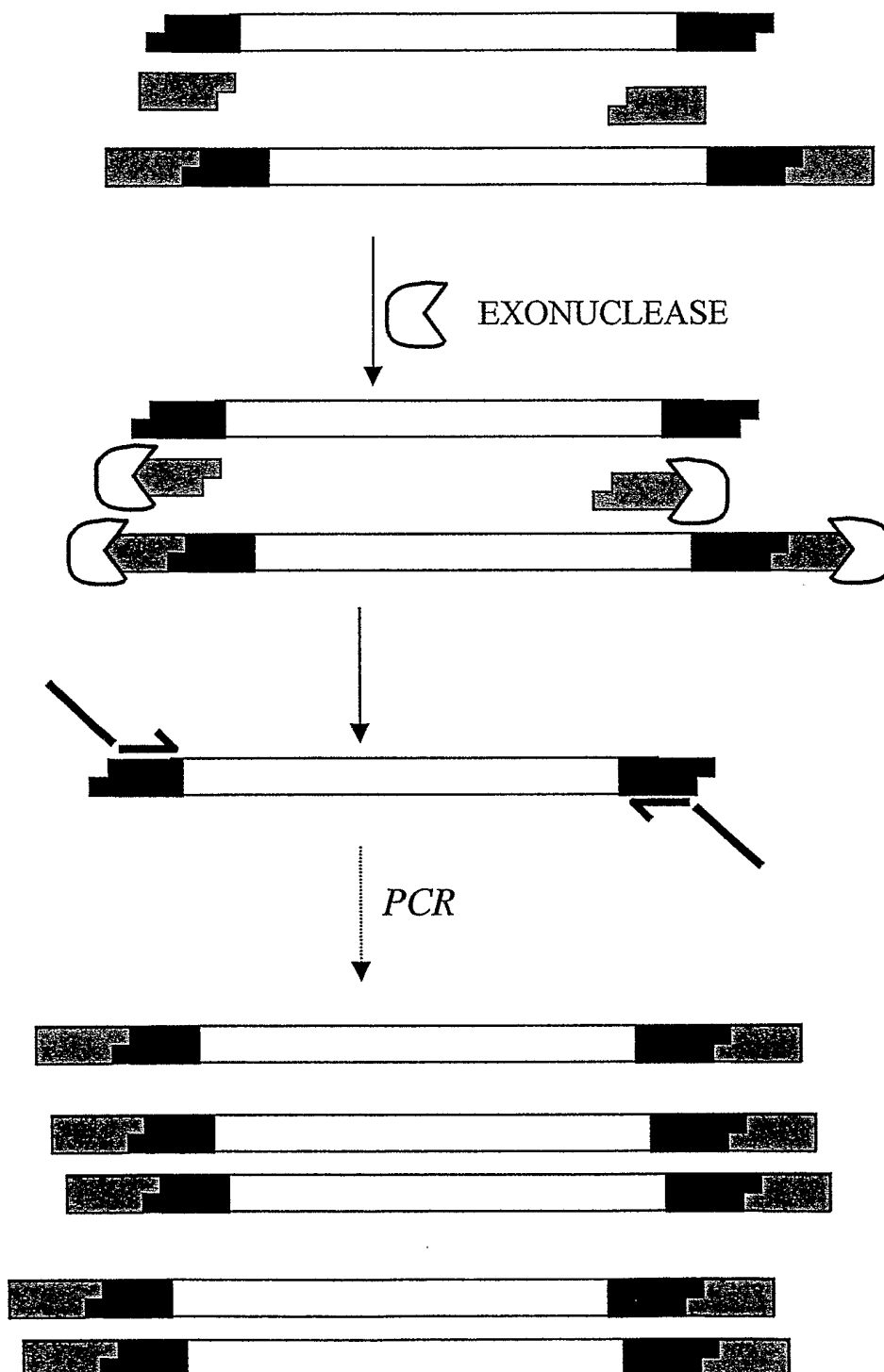
35- Méthode selon la revendication 4, caractérisée en ce que, préalablement à la compartimentation, les substrats ADN sont mis en présence d'un extrait de transcription/traduction in vitro, dans des conditions ne permettant pas une expression et/ou une traduction de la séquence codant l'endonucléase.

35

36- Molécule d'acide nucléique, caractérisée en ce qu'il s'agit d'ADN double-brin, d'une taille comprise entre 100 et 3000 pb environ, comprenant au moins deux régions distales formant des sites de coupure pour une endonucléase.

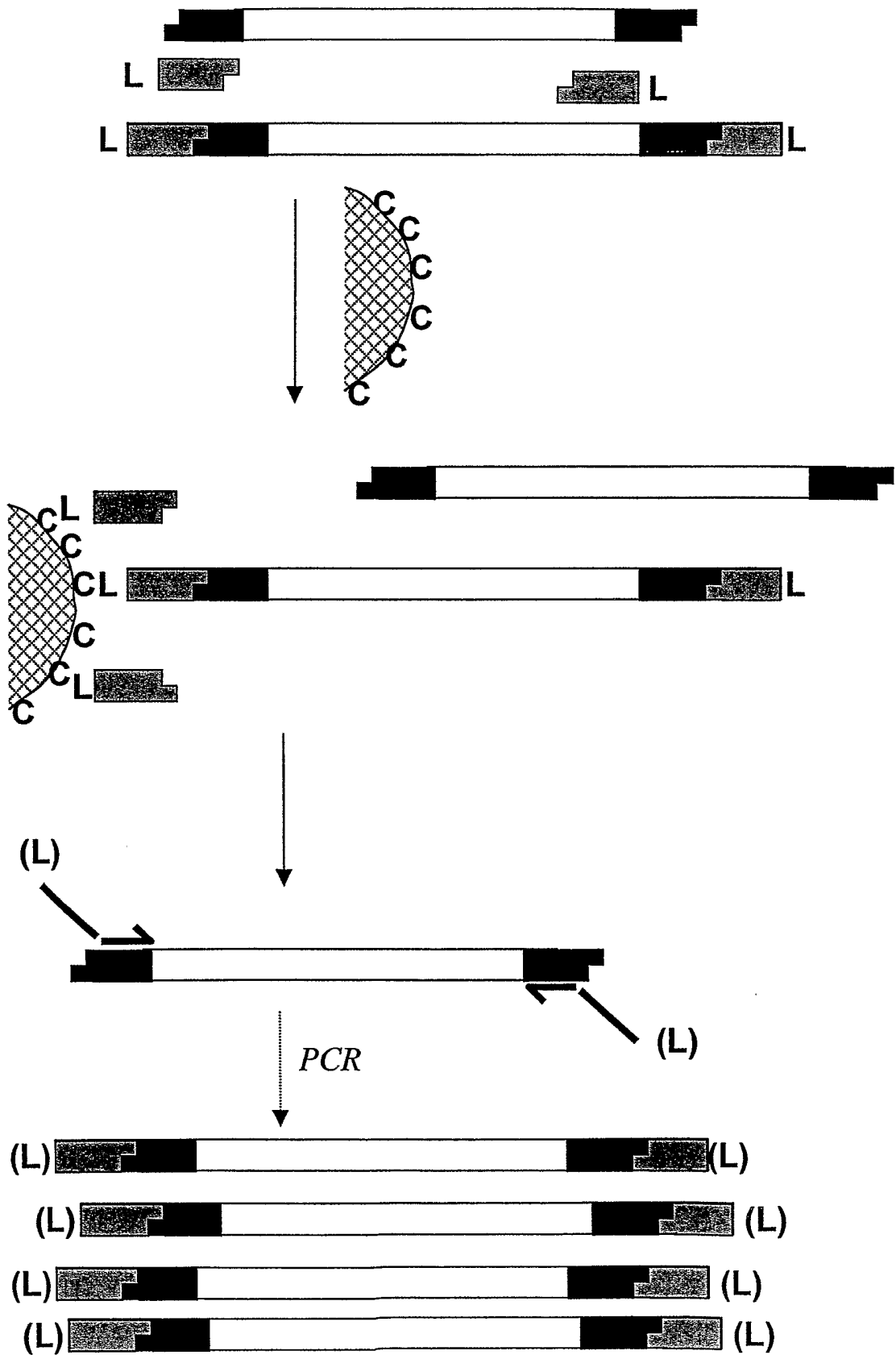
1/17

FIGURE 1



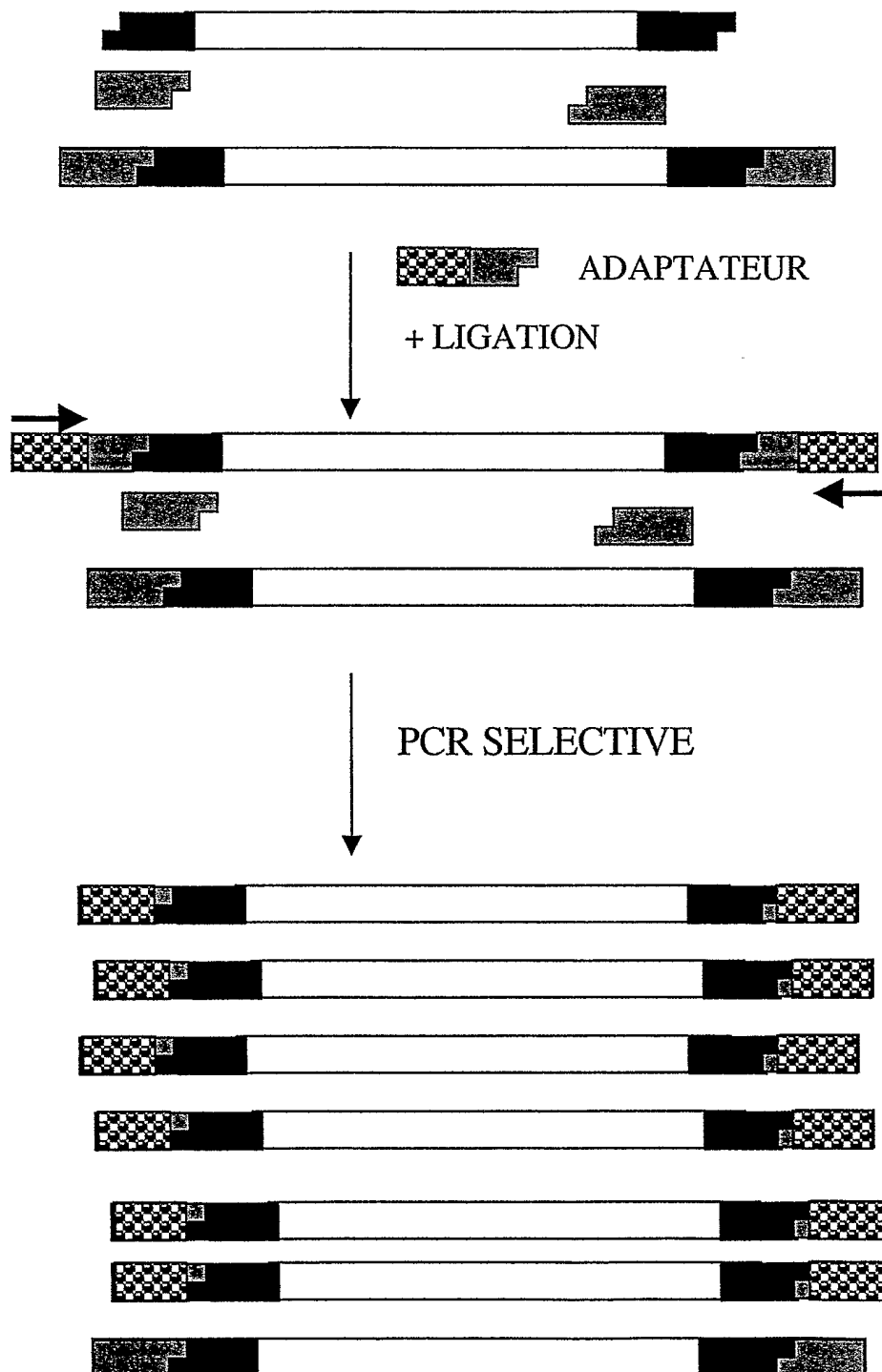
2/17

FIGURE 2

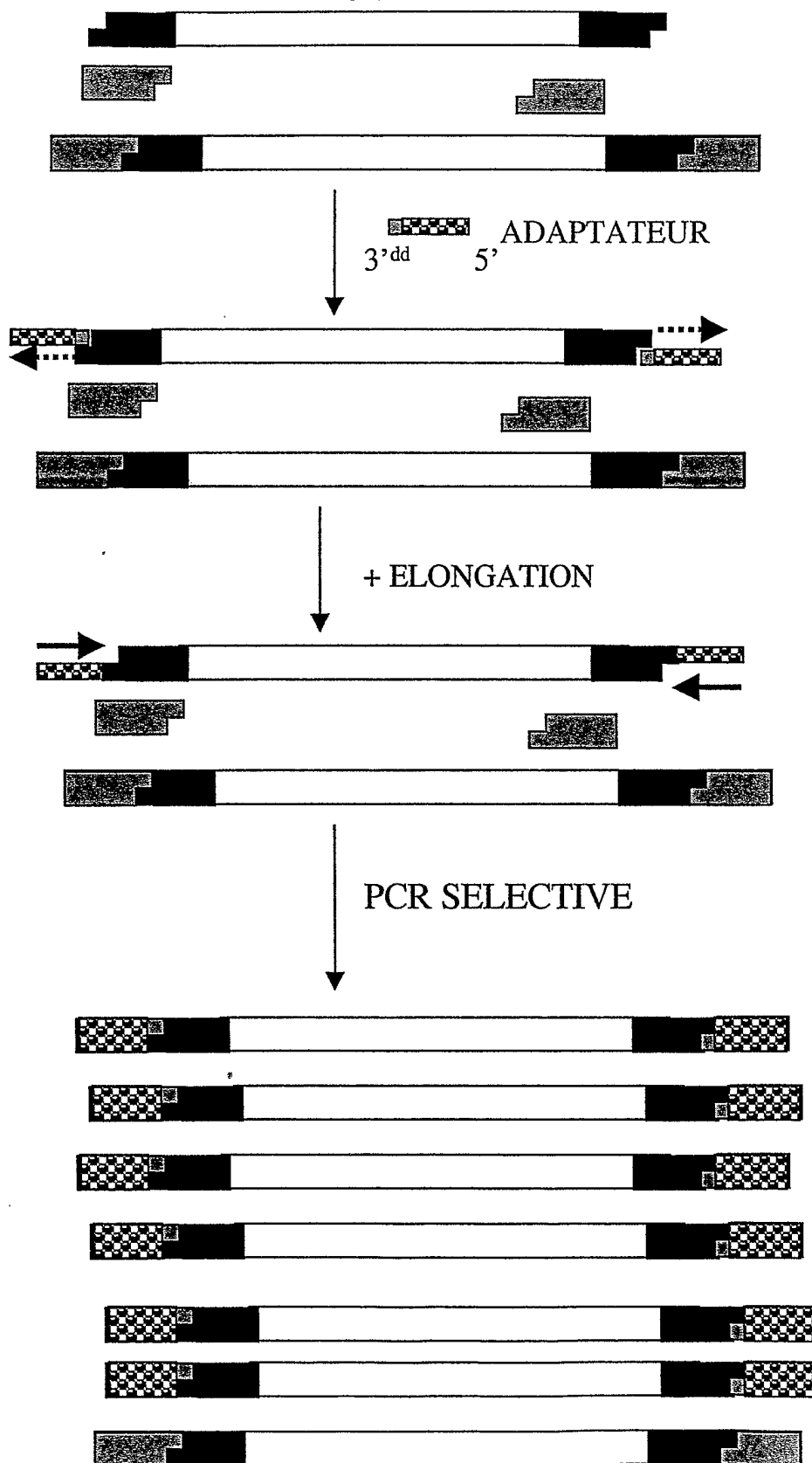


3/17

FIGURE 3A

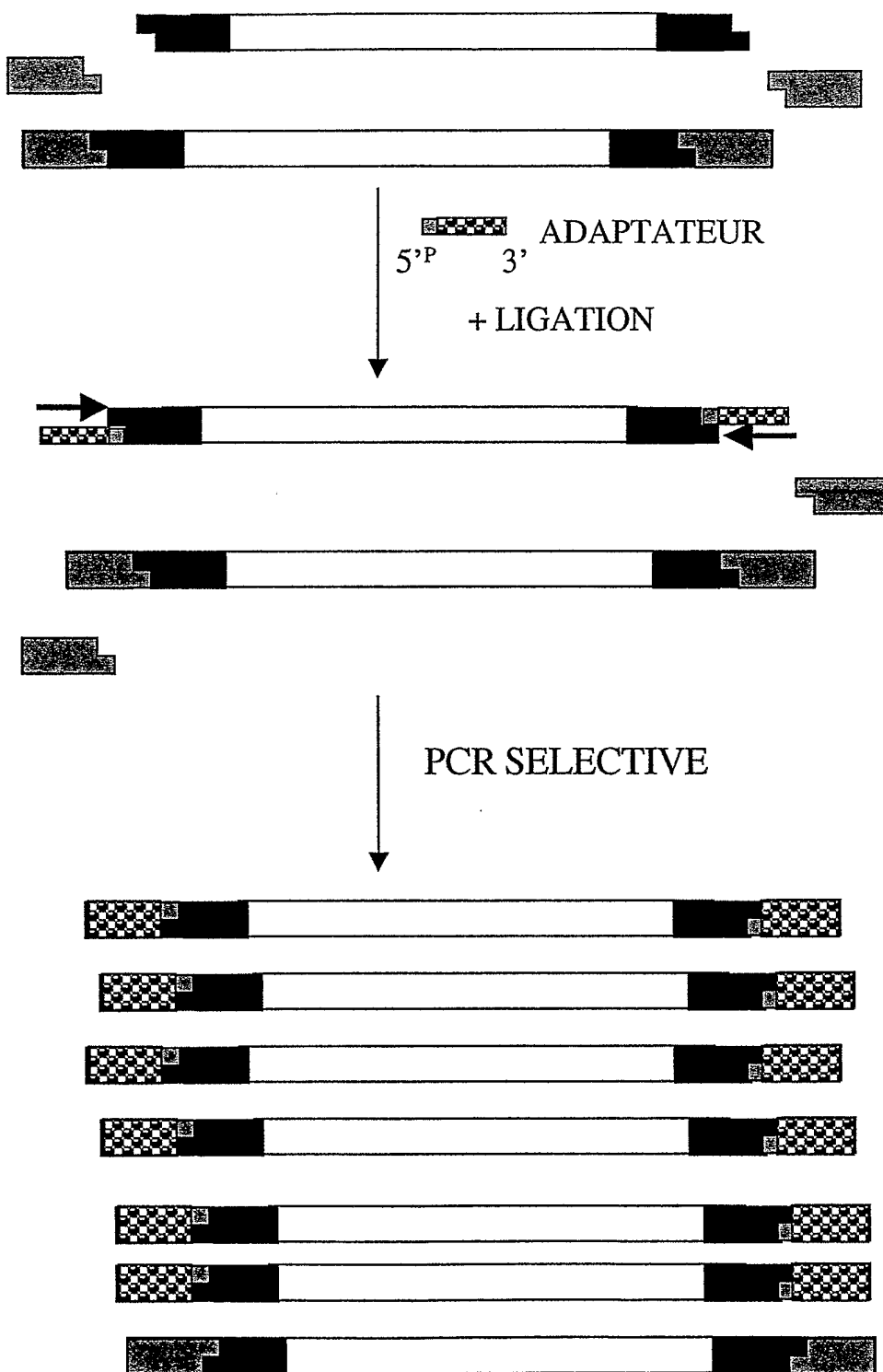


4/17

FIGURE 3B

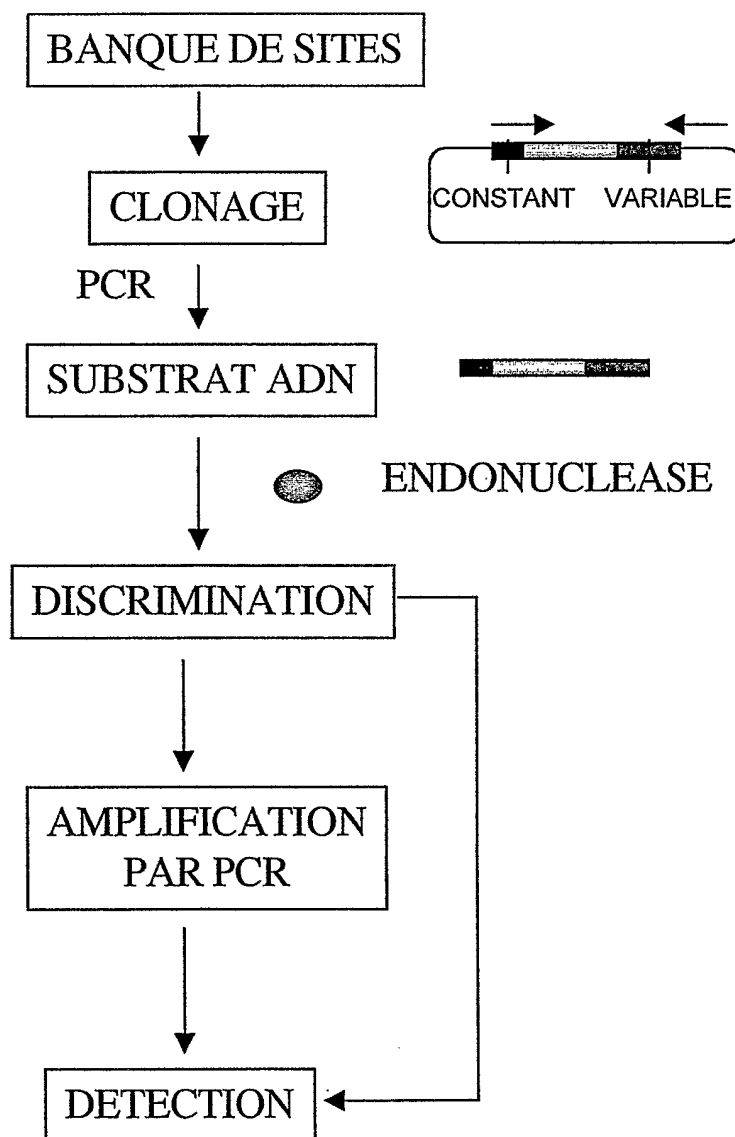
5/17

FIGURE 3C



6/17

FIGURE 4



7/17

FIGURE 5

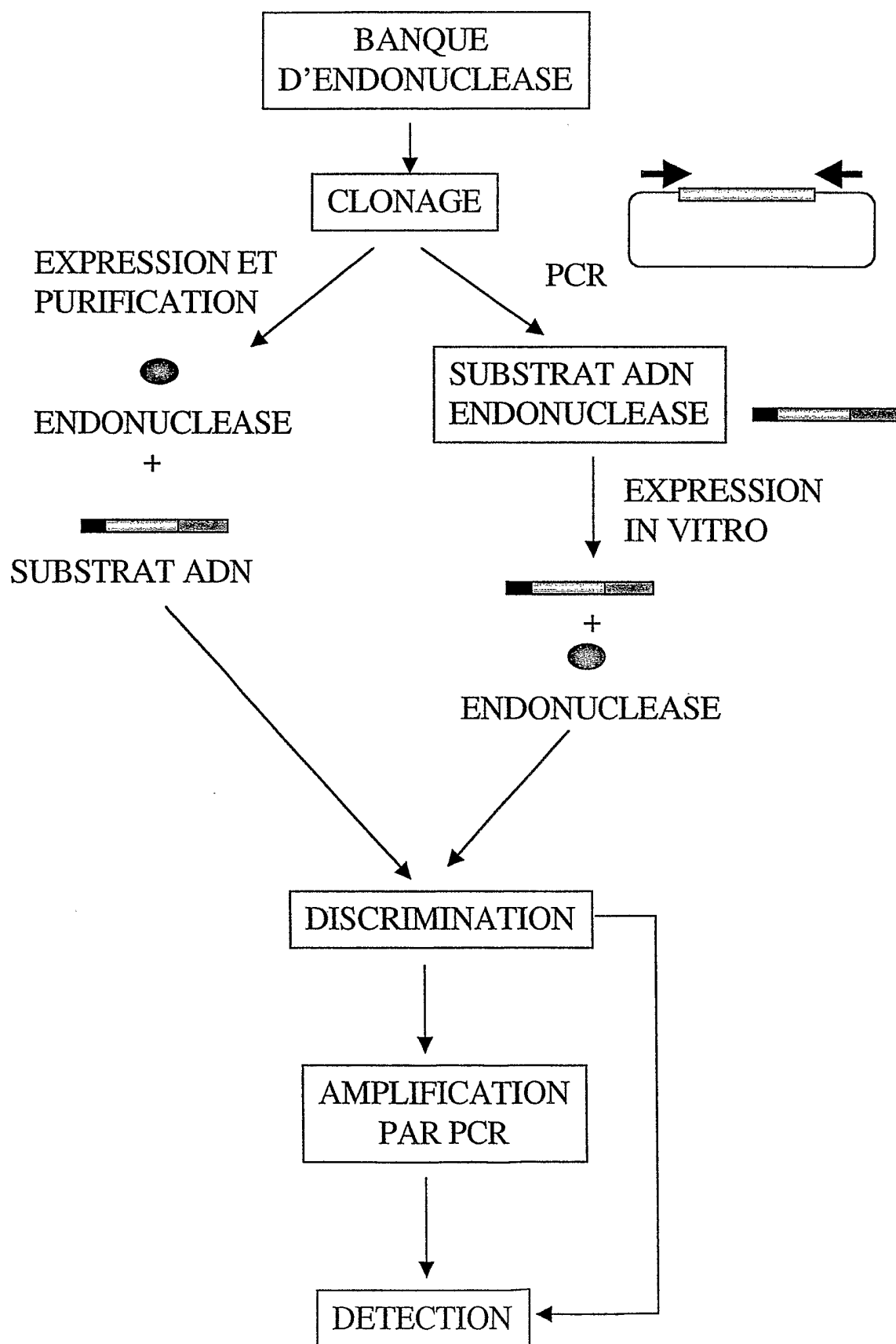
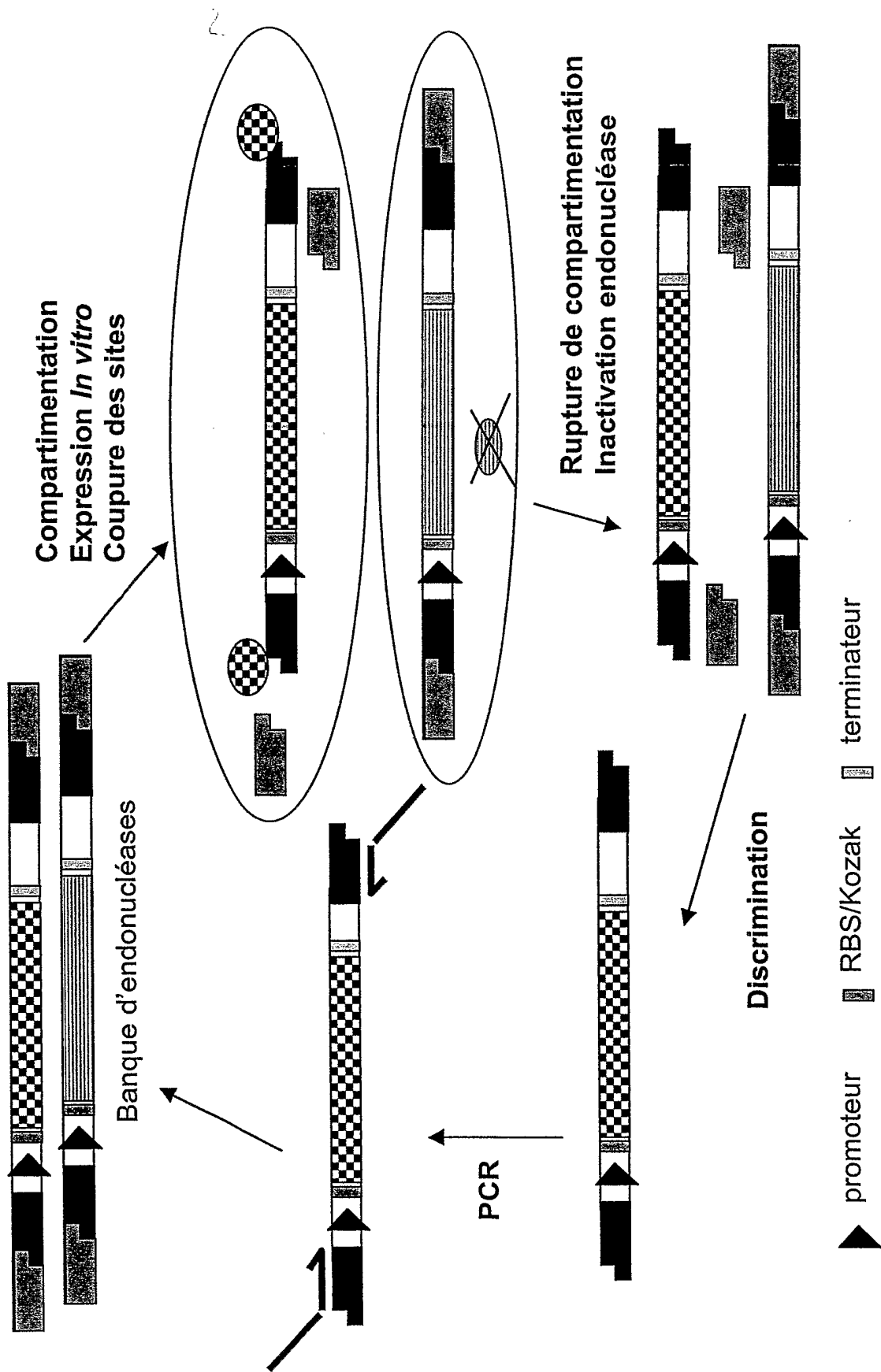
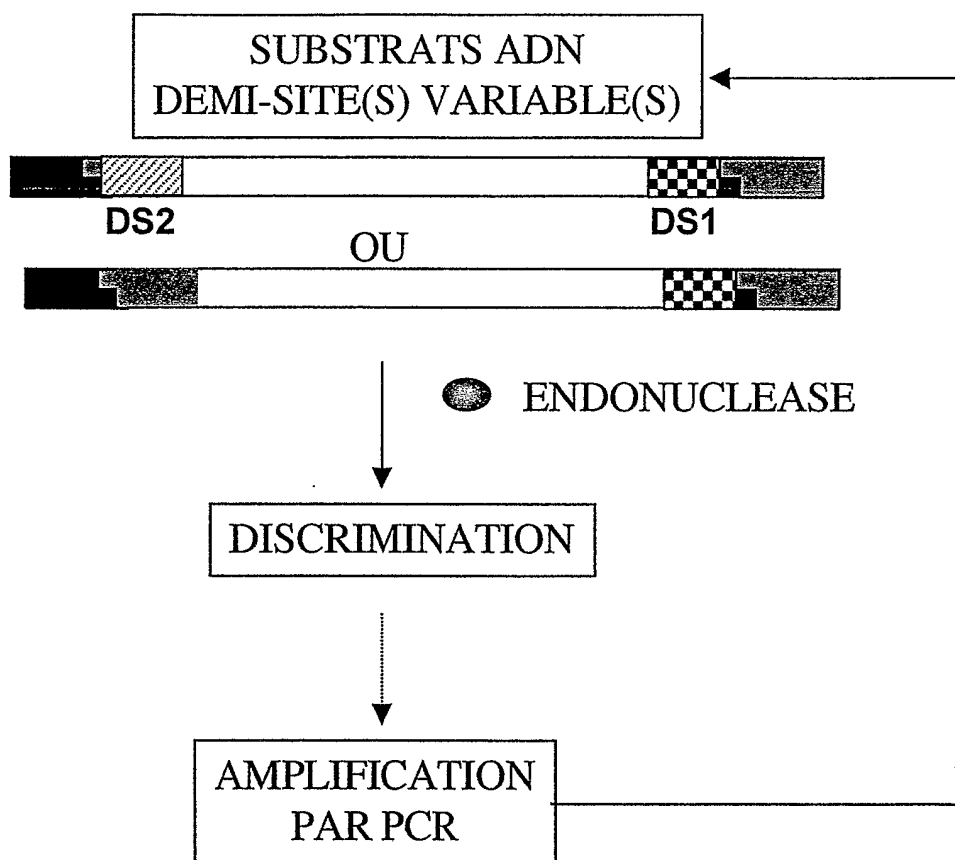


FIGURE 6

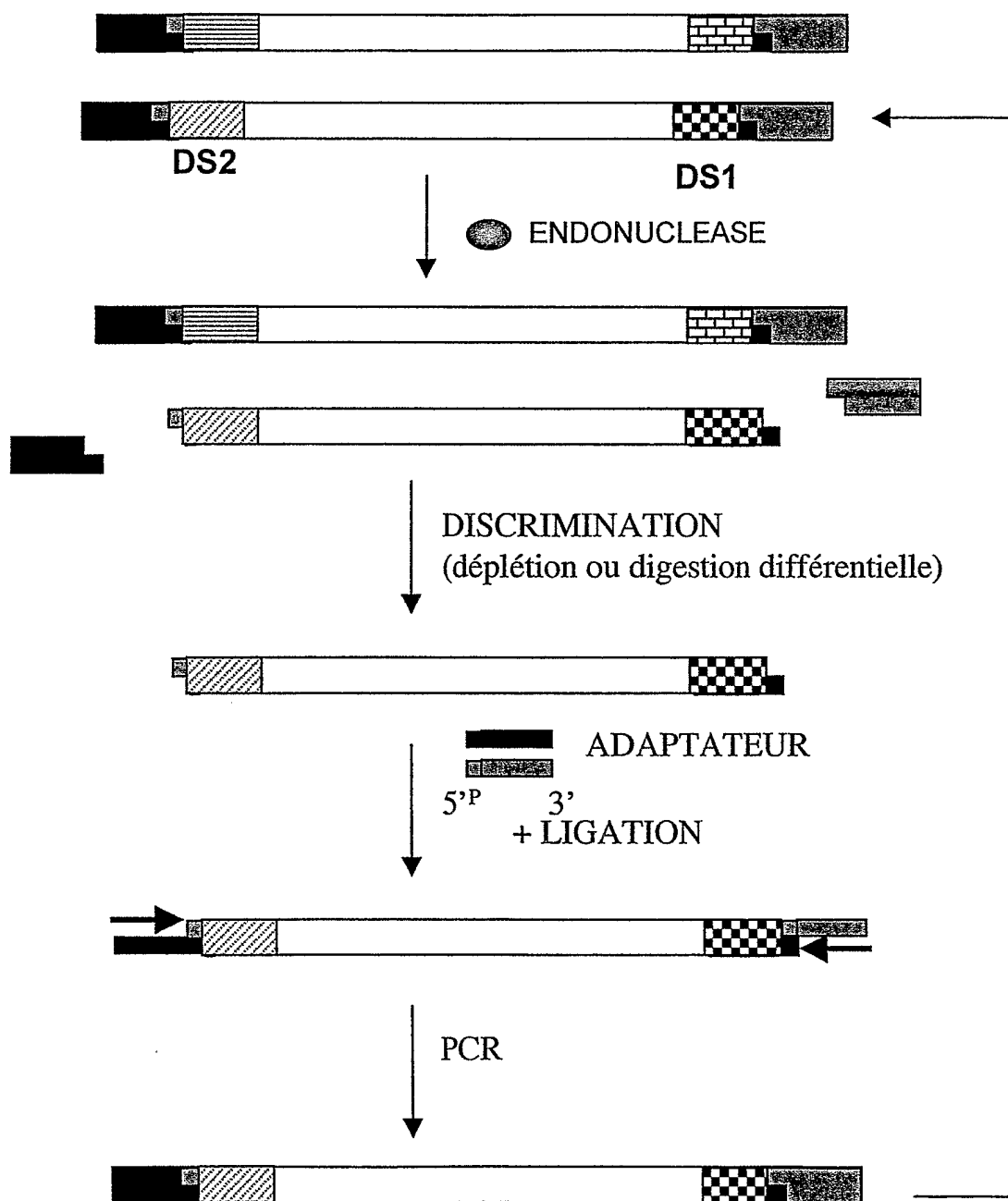


9/17

FIGURE 7A

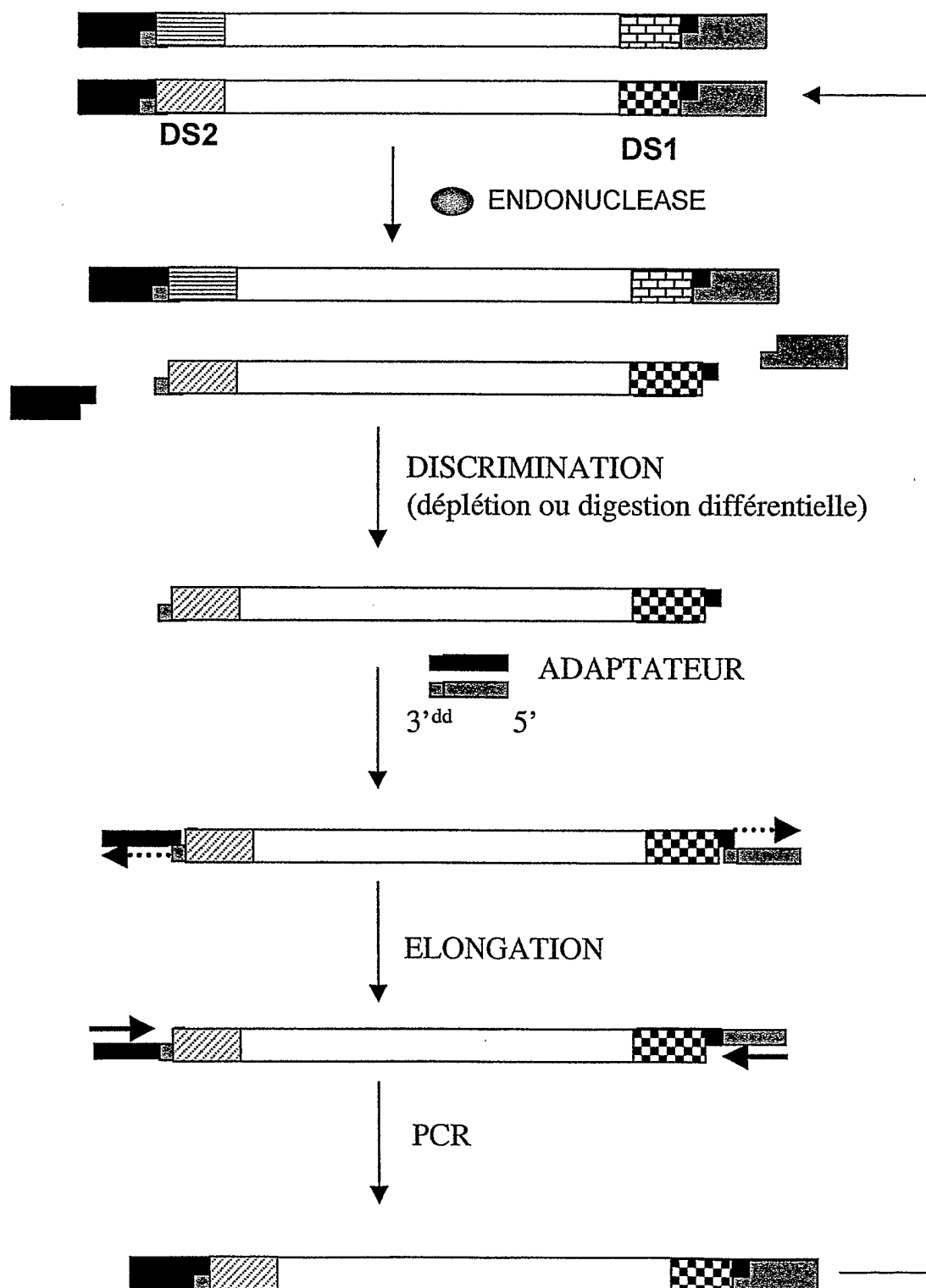
10/17

FIGURE 7B



11/17

FIGURE 7C



12/17
FIGURE 7D

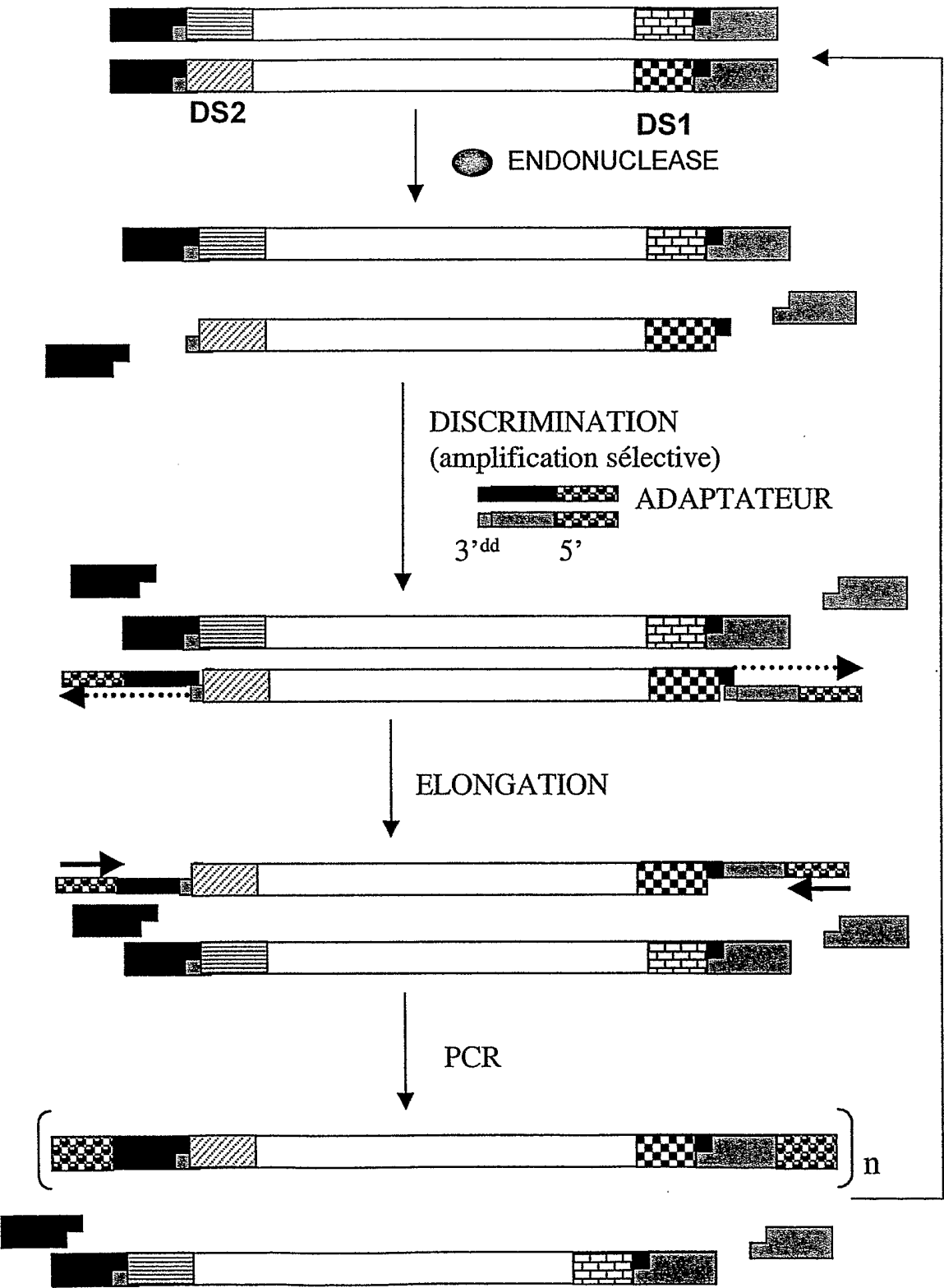
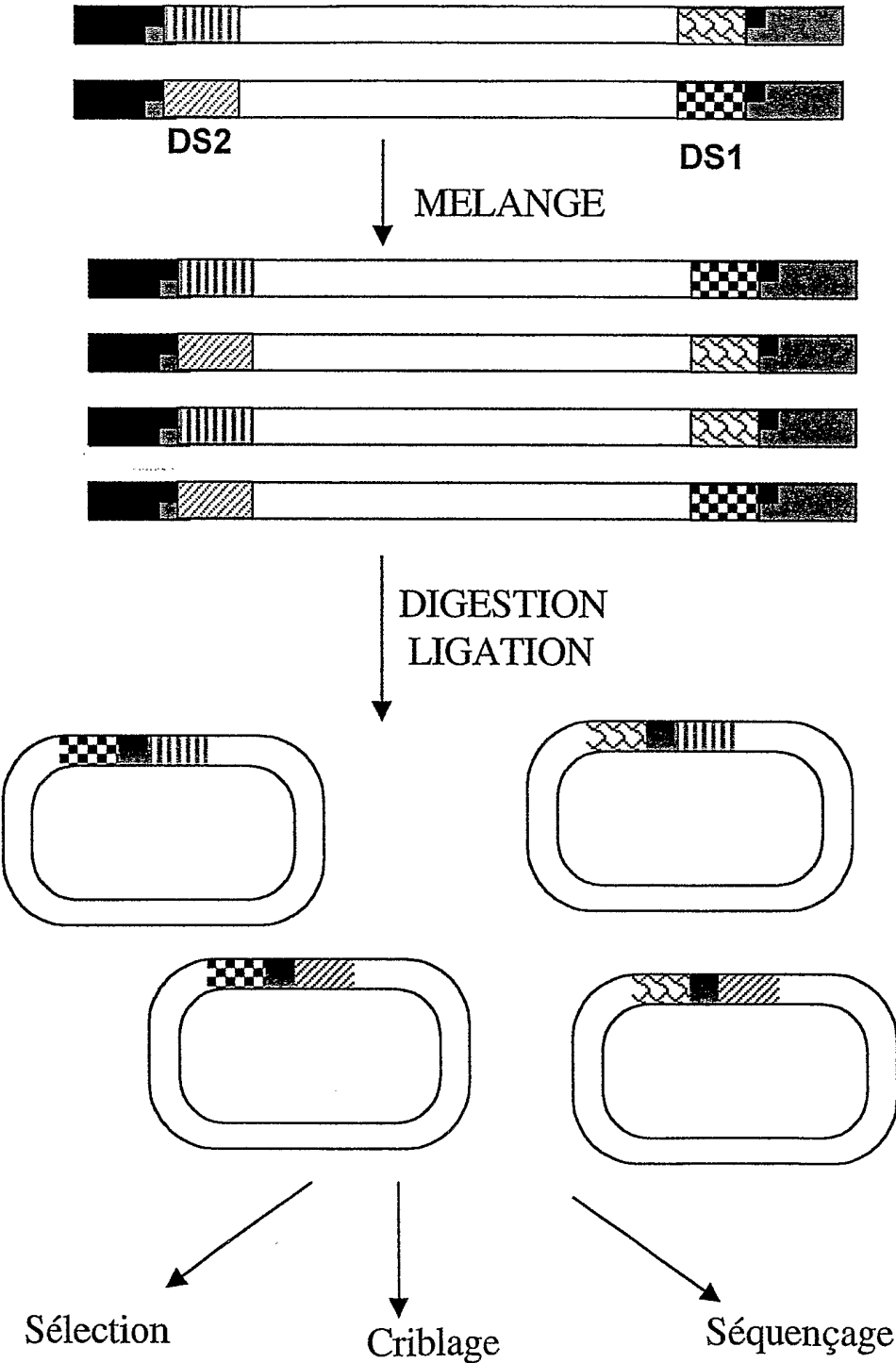


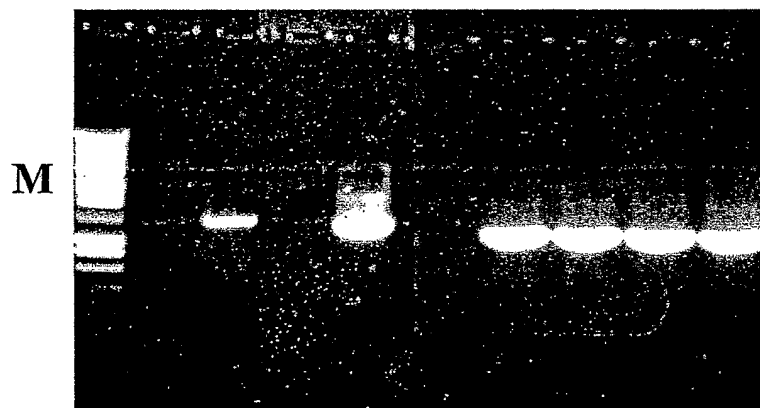
FIGURE 7E



14/17

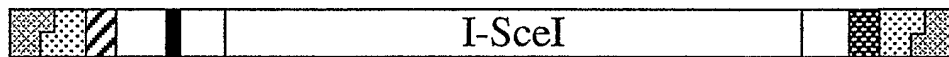
FIGURE 8

ExoIII	+	+	+	+	-	-	-	-
MgCl ₂	-	-	+	+	-	-	+	+
I-SceI	-	+	-	+	-	+	-	+



15/17

FIGURE 9



Site I-SceI



Promoteur T7



Site de liaison au ribosome



T7 terminateur

16/17

FIGURE 10

	1	2	3	4	5
Buffer	X	X	X	-	-
RTS	-	-	-	X	X
MgCl ₂	-	-	X	-	X
I-SceI	-	X	X	-	-

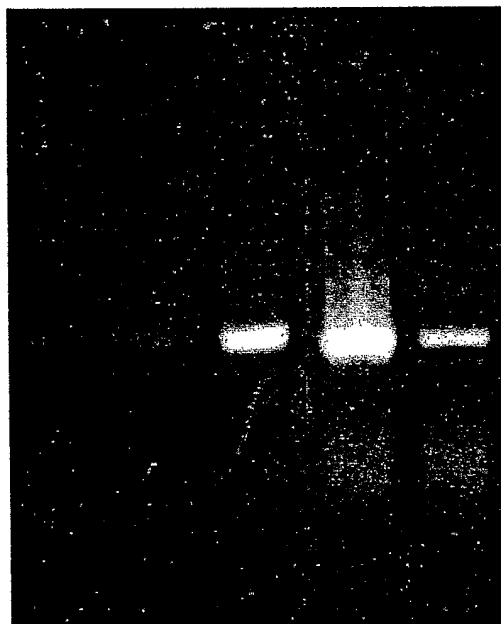


FIGURE 11A

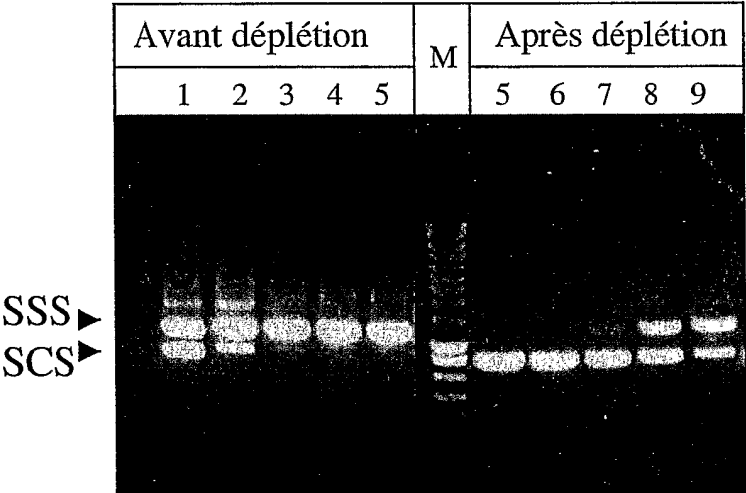
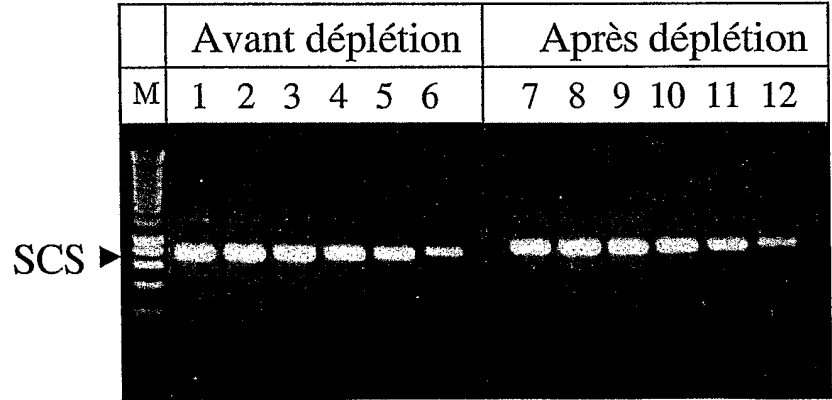


FIGURE 11B



SEQUENCE LISTING

<110> CELLECTIS

<120> METHODE DE DETERMINATION DE L'ACTIVITE DE CLIVAGE D'UNE
ENDONUCLEASE

<130>

<160> 5

<170> PatentIn version 3.1

<210> 1

<211> 45

<212> DNA

<213> Artificial

<220>

<221> primer-bind

<222> (1)..(45)

<223>

<400> 1

acgctaggga taacaggga atatagctaa tacgactcac tatag
45

<210> 2

<211> 45

<212> DNA

<213> artificial

<220>

<221> primer-bind

<222> (1)..(45)

<223>

<400> 2

cgctagggat aacagggtaa tatagccaaa aaaccctca agacc
45

<210> 3

<211> 22

<212> DNA

<213> artificial

<220>

<221> primer-bind

<222> (1)..(22)

<223>

<400> 3

acgctaggga taacaggga at
22

<210> 4

<211> 768

<212> DNA

<213> artificial

<220>

<221> CDS

<222> (1) .. (768)

<223>

<400> 4

atg ggc cat cat cat cat cat cat cat cat cac agc agc ggc cat
48

Met Gly His His His His His His His His His Ser Ser Gly His
1 5 10 15

atc gaa ggt cgt cat atg aaa aac atc aaa aaa aac cag gta atg aac
96

Ile Glu Gly Arg His Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn
20 25 30

ctg ggt ccg aac tct aaa ctg ctg aaa gaa tac aaa tcc cag ctg atc
144

Leu Gly Pro Asn Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile
35 40 45

gaa ctg aac atc gaa cag ttc gaa gca ggt atc ggt ctg atc ctg ggt
192

Glu Leu Asn Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly
50 55 60

gat gct tac atc cgt tct cgt gat gaa ggt aaa acc tac tgt atg cag
240

Asp Ala Tyr Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln
65 70 75 80

ttc gag tgg aaa aac aaa gca tac atg gac cac gta tgt ctg ctg tac
288

Phe Glu Trp Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr
85 90 95

gat cag tgg gta ctg tcc ccg ccg cac aaa aaa gaa cgt gtt aac cac
336

Asp Gln Trp Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His
100 105 110

ctg ggt aac ctg gta atc acc tgg ggc gcc cag act ttc aaa cac caa
384

Leu Gly Asn Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln
115 120 125

gct ttc aac aaa ctg gct aac ctg ttc atc gtt aac aac aaa aaa acc
432

Ala Phe Asn Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr
130 135 140

atc ccg aac aac ctg gtt gaa aac tac ctg acc ccg atg tct ctg gca
480

Ile Pro Asn Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala
145 150 155 160

tac tgg ttc atg gat gat ggt ggt aaa tgg gat tac aac aaa aac tct
 528
 Tyr Trp Phe Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser
 165 170 175
 acc aac aaa tcg atc gta ctg aac acc cag tct ttc act ttc gaa gaa
 576
 Thr Asn Lys Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu
 180 185 190
 gta gaa tac ctg gtt aag ggt ctg cgt aac aaa ttc caa ctg aac tgt
 624
 Val Glu Tyr Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys
 195 200 205
 tac gta aaa atc aac aaa aac aaa ccg atc atc tac atc gat tct atg
 672
 Tyr Val Lys Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met
 210 215 220
 tct tac ctg atc ttc tac aac ctg atc aaa ccg tac ctg atc ccg cag
 720
 Ser Tyr Leu Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln
 225 230 235 240
 atg atg tac aaa ctg ccg aac act atc tcc tcc gaa act ttc ctg aaa
 768
 Met Met Tyr Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
 245 250 255

<210> 5
 <211> 256
 <212> PRT
 <213> artificial

<400> 5

Met Gly His His His His His His His His His His Ser Ser Gly His
 1 5 10 15
 Ile Glu Gly Arg His Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn
 20 25 30
 Leu Gly Pro Asn Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile
 35 40 45
 Glu Leu Asn Ile Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly
 50 55 60
 Asp Ala Tyr Ile Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln
 65 70 75 80
 Phe Glu Trp Lys Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr
 85 90 95
 Asp Gln Trp Val Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His
 100 105 110

Leu Gly Asn Leu Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln
115 120 125

Ala Phe Asn Lys Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr
130 135 140

Ile Pro Asn Asn Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala
145 150 155 160

Tyr Trp Phe Met Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser
165 170 175

Thr Asn Lys Ser Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu
180 185 190

Val Glu Tyr Leu Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys
195 200 205

Tyr Val Lys Ile Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met
210 215 220

Ser Tyr Leu Ile Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln
225 230 235 240

Met Met Tyr Lys Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys
245 250 255