

(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
11 mai 2006 (11.05.2006)

PCT

(10) Numéro de publication internationale
WO 2006/048465 A2

(51) Classification internationale des brevets :
C12N 15/90 (2006.01) *A01K 67/027* (2006.01)

(21) Numéro de la demande internationale :
PCT/EP2005/055826

(22) Date de dépôt international :
8 novembre 2005 (08.11.2005)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :
0411871 8 novembre 2004 (08.11.2004) FR

(71) Déposant (pour tous les États désignés sauf US) :
GENOWAY [FR/FR]; 181 Avenue Jean-Jaurès, Immeuble
Châteaubriand, F-69007 LYON (FR).

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : **THIAM,**
Kader [FR/FR]; 144 Avenue de Saxe, F-69003 LYON
(FR). **FRAICHARD, Alexandre** [FR/FR]; 20 bis rue Lt
Colonel Girard, F-69007 LYON (FR).

(74) Mandataires : **CABINET REGIMBEAU** etc.; 129 rue
Servient, F-69326 LYON CEDEX 03 (FR).

(81) États désignés (sauf indication contraire, pour tout titre de
protection nationale disponible) : AE, AG, AL, AM, AT,
AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO,
CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG,
KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY,
MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO,
NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK,
SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, YU, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre
de protection régionale disponible) : ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT,
RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA,
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Publiée :

— sans rapport de recherche internationale, sera republiée
dès réception de ce rapport

En ce qui concerne les codes à deux lettres et autres abrévia-
tions, se référer aux "Notes explicatives relatives aux codes et
abréviations" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.

(54) Title: ANIMAL MODELS COMPRISING AT LEAST ONE MUTATION OR ONE CONDITIONAL KI AND ONE CON-
DITIONAL KO

(54) Titre : ANIMAUX MODÈLES COMPRENANT AU MOINS UNE MUTATION OU UN KI ET UN KO CONDITIONNEL

(57) Abstract: The invention relates to a combination of DNA sequences that makes it possible to develop animal models com-
prising, on the one hand, a modification of its genome in order to introduce either a new gene to be expressed (KI or Knock In) or a
mutation into a target gene and, on the other, a constitutive or conditional KO (Knock Out) into a target gene.

(57) Abrégé : La présente invention concerne une combinaison de séquences d'ADN permettant de développer des animaux mo-
dèles comprenant d'une part une modification de son génome afin d'introduire soit un nouveau gène pour être exprimé (KI ou Knock
In) soit une mutation dans un gène cible et d'autre part un KO (Knock Out), constitutif ou conditionnel, dans un gène cible.



WO 2006/048465 A2

Animaux modèles comprenant au moins une mutation ou un KI et un KO conditionnel

La présente invention concerne une combinaison de séquences d'ADN permettant de développer des animaux modèles comprenant d'une part une modification de son génome afin d'introduire soit un nouveau gène pour être exprimé (KI ou Knock In) soit une mutation dans un gène cible et d'autre part un KO (Knock Out), constitutif ou conditionnel, dans un gène cible.

La combinaison de séquences ADN selon l'invention permet de réaliser sur le même modèle animal plusieurs opérations nécessitant autrefois le développement et l'usage de plusieurs modèles distincts. Il permet par exemple, d'étudier avec le même modèle le phénotype obtenu par l'introduction d'une mutation dans la séquence du gène endogène ou de disposer d'un modèle transgénique par exemple exprimant le gène humain en lieu et place du gène endogène, puis après excision de l'ensemble effectuer un KO dans le gène d'insertion, permettant d'étudier le phénotype pour le modèle KO, ce KO pouvant être tissu spécifique ou ubiquitaire.

Différentes stratégies à base de recombinaisons et de vecteurs permettant d'insérer puis de supprimer des éléments de gènes dans les génomes de différents organismes sont décrits dans l'état de la technique, notamment dans EP 1 462 525, US 2003/165497, US 2004/241851, WO 02/40685, WO 2004/056964, WO 2004/044151, WO 2005/007877, Kasin & al (Nucleic Acid Research. 2004, vol 32, n° 7, pp E66.1-E66.8), Coumoul & al (Nucleic Acid Research 2004, vol 32, n° 10, pp E85.1-E85.7), Conrad & al. (Biotechnologies, vol 34, n° 6, juin 2003, 1136-1140), Kühn & al. (Methods in Molecular Biology 2003, vol 209, 159-185), Soukharev & al. (Nucleic Acid Research 1999, vol 27, n° 18, ppE21.1-E21.8), McMinn & al. (Mammalian Genome 2004, vol 15, n° 9, 677-685), Casanova & al (Genesis 2002, vol 32, n° 2, 15/8-160) et Ren & al (Genesis 2002, vol 32, n° 2, 105-108). Cependant, aucun de ces documents ne vient apporter de solution au problème technique associé à la présente invention, ni décrire la solution de la présente invention.

La combinaison de séquences d'ADN selon l'invention comprend une première séquence (S1) et une deuxième séquence (S2) destinées à être insérées dans l'ADN génomique d'une cellule eucaryote d'un organisme hôte.

Par séquence d'ADN on entend selon l'invention une molécule d'ADN constituée par une séquence d'acides désoxyribonucléiques liés de manière covalente.

Pour la combinaison selon l'invention, les première et deuxième séquences S1 et S2 sont supportées soit sur deux vecteurs, soit sur un même vecteur, soit encore dans le génome d'au moins une cellule eucaryote d'un organisme hôte.

5 Par cellule eucaryote, on entend plus particulièrement selon l'invention toutes cellules d'origine animale, préférentiellement à l'exception de l'homme, en particulier les cellules de mammifères, préférentiellement de mammifères rongeurs, plus préférentiellement de rat ou de souris.

De manière préférentielle, la première et deuxième séquences sont intégrées dans l'ADN génomique d'une cellule eucaryote au sein d'un même gène cible, la première
10 séquence S1 étant préférentiellement destinée à être intégrée en aval du promoteur du gène cible.

De manière préférentielle la deuxième séquence S2 est destinée à être intégrée en aval du site d'intégration de la première séquence S1 de préférence en aval d'exons important pour la fonction ou la fonctionnalité du gène cible.

15 La première séquence S1 de la combinaison selon l'invention comprend, dans le sens 5'-3', les éléments suivants :

- un premier site SR1 de reconnaissance d'une première recombinaise R1, et
- un acide nucléique hétérologue.

La deuxième séquence S2 de la combinaison selon l'invention comprend, dans le sens
20 5'-3', les éléments suivants :

- un premier site SR2 de reconnaissance d'une recombinaise R2,
- une séquence d'acide nucléique d'un gène codant pour un marqueur de sélection fonctionnel dans la cellule eucaryote dans le génome de laquelle il sera intégré,
- un second site SR2 de reconnaissance de la recombinaise R2, et
- 25 - le cas échéant, un second site SR1 de reconnaissance de la recombinaise R1.

Selon un premier mode particulier de réalisation de l'invention, l'acide nucléique hétérologue de la première séquence S1 est un acide nucléique codant pour un exon du gène cible, comprenant au moins une mutation. La séquence S1 est avantageusement intégrée dans le génome de la cellule cible de manière que l'exon muté de la séquence S1 vienne remplacer
30 l'exon non muté du gène cible.

Parmi les gènes cibles, on peut citer les gènes associés aux récepteurs nucléaires, aux récepteurs membranaires, aux facteurs de transcription, etc.

L'insertion de la combinaison de séquences dans le génome de l'organisme hôte peut conduire, après excision, à une délétion affectant un ou plusieurs gènes, par exemple lorsque le locus d'insertion porte plusieurs gènes ou parties de gènes, ou encore lorsque l'on laisse un grand fragment d'ADN génomique entre les deux séquences S1 et S2 selon l'invention.

5 Selon un deuxième mode particulier de réalisation de l'invention, l'acide nucléique hétérologue de la séquence S1 est un acide nucléique comprenant dans le sens 5'-3', les éléments suivants :

- une séquence d'acide nucléique codant pour une séquence d'intérêt, et
 - un deuxième site SR1 de reconnaissance de la troisième recombinaise R1, le site SR1
- 10 localisé en 3' de la séquence S2 n'est pas présent..

De manière préférentielle, la séquence d'acide nucléique codant pour une séquence d'intérêt dans la séquence S1 comprend, liés de manière fonctionnelle, des éléments fonctionnels nécessaires à l'arrêt de la transcription et de manière préférentielle des éléments fonctionnels nécessaires à l'initiation de la transcription.

15 Parmi les éléments nécessaires à l'initiation de la transcription, on peut citer ceux décrits notamment dans Baker et al ((1972) J Mol Biol.; 69:89-102.) et dans Georgiev GP.((1972) Curr Top Dev Biol. 7, 1-60).

Parmi les éléments nécessaires à l'arrêt de la transcription, on peut citer ceux décrits notamment dans Proudfoot et al (Cell. (1991) 64,:671-4) et dans Beaudoin et al. (Genome Res. (2001) 11, 520-526).

20

La séquence d'intérêt de la séquence S1 selon l'invention est de préférence choisie parmi les séquences suivantes :

- une séquence d'acide nucléique codant pour un polypeptide, en particulier une protéine d'intérêt, ou
- 25 - une séquence d'acide nucléique codant pour un RNA, en particulier un RNA antisens ou un RNAi, miRNA, shRNA ou siRNA.

De manière avantageuse, la séquence d'acide nucléique codant pour un polypeptide est un ADNc, plus préférentiellement un ADNc codant pour un polypeptide d'origine humaine. On citera notamment ceux décrits dans Li, et al. ((1997). J Cell Biol 139, 129-44), Schneider-
30 Maunoury et al. ((1993). Cell 75, 1199-214) et Tajbakhsh et al. ((1996) Nature 384, 266-70).

De manière avantageuse, le RNA antisens ou interférent permet d'inhiber la traduction d'ARNm cibles, préférentiellement choisis parmi les gènes suivants : GPCR, récepteurs nucléaires, récepteurs membranaires, cytokines, canaux ionique, etc.

De manière avantageuse, les sites de reconnaissance de recombinaisons SR1 et SR2 sont choisis, indépendamment l'un de l'autre, parmi les sites Lox P, Lox P modifiés et FRT.

Les combinaisons préférées sont représentées sur le tableau ci-dessous.

SR1	SR2
Lox P	Lox P modifié
FRT	FRT
Lox P muté	FRT
FRT	Lox P muté

5 En fonction des sites de reconnaissance choisis, les recombinaisons R1 ou R2 sont choisies, indépendamment l'une de l'autre, parmi les recombinaisons suivantes : Cre ou FLP.

Les sites de reconnaissance Lox P, Lox P modifié, FRT et les recombinaisons Cre et FLP correspondantes sont bien connus de l'homme du métier, notamment décrits dans Kilby, et al. ((1993). Trends Genet 9, 413-21), Sauer, et Henderson ((1988). Proc Natl Acad Sci U S
10 A 85, 5166-70), Gu et al. ((1994). Science 265, 103-6), Cohen-Tannoudjiet Babinet ((1998). Mol Hum Reprod 4, 929-38), Shibata, et al. ((1997). Science 278, 120-3), Schlake et Bode ((1994). Biochemistry 33, 12746-51).

La présente invention concerne également un procédé de transformation de cellules eucaryotes pour intégrer dans son génome la combinaison de vecteurs selon l'invention.

15 Les méthodes d'intégration de séquences dans le génome de cellules eucaryotes sont bien connues de l'homme du métier, notamment décrites dans Smithies et al. ((1985). Nature 317, 230-4) et Wong et al. ((1986). Somat Cell Mol Genet 12, 63-72).

De manière préférentielle, les cellules sont des cellules animales transformées par les méthodes telles que décrites dans Vasquez et al ((2001) PNAS 98, 8403-8410) et Yanez et al
20 ((1999), somatic cell and molecular genetics 25, 27-31).

La présente invention concerne aussi les cellules eucaryotes dans le génome desquelles est intégré la combinaison de vecteurs selon l'invention, plus préférentiellement les cellules animales telles que définies précédemment, en particulier souris ou rat.

Ces cellules selon l'invention peuvent être des cellules obtenues par intégration de la
25 combinaison de séquences par le procédé selon l'invention. Elles peuvent également comprendre des cellules obtenues par prélèvement sur des animaux obtenus à partir desdites cellules obtenues par le procédé selon l'invention.

L'invention concerne également des embryons d'animaux obtenus à partir des cellules selon l'invention, ainsi que les animaux obtenus à partir de ces embryons et la descendance desdits animaux ayant conservé le vecteur selon l'invention. Les animaux selon l'invention sont définis précédemment à l'exception de l'homme, préférentiellement des mammifères, plus préférentiellement des mammifères rongeurs, encore plus préférentiellement des rats ou des souris.

Les animaux modèles ainsi obtenus peuvent être employés de la manière décrite ci-dessous.

Les méthodes de préparation d'animaux modèles sont bien connues de l'homme du métier, et notamment décrites pour le mouton (Wilmot *et al.* Nature **385**, 810-813, 1997 ; WO 97 07669), la souris (Wakayama *et al.* Nature **394**, 369-374, 1998 ; WO 99 37143), les bovins (Wells *et al.* Biol. Reprod. **60**, 996-1005, 1999), la chèvre (Baguisi *et al.* Nature Biotechnol. **17**, 456-461, 1999 ; WO 00 25578), le porc (Polejaeva *et al.* Nature **407**, 86-90, 2000), le lapin (Chesne *et al.* Nature Biotechnol. **20**, 366-369, 2002) et le rat (Zhou & al., Science, 25 septembre 2003 ; PCT/FR04/001275). Les méthodes de transfert nucléaire sont notamment décrites par Campbell & al. (*Nuclear transfer in practice*, School of Biosciences, University of Nottingham, Leicestershire, United Kingdom).

La première application concerne le développement d'un modèle porteur d'une mutation dans un exon porteur d'un domaine fonctionnel (S1). Ce modèle permet d'étudier le phénotype créé par cette mutation.

Dans un second temps, il est possible de le comparer au phénotype obtenu par une inactivation complète de la fonction du gène (KO) en croisant cet animal avec un animal exprimant une recombinaise (Cre ou Flp selon les sites SR utilisés) de façon ubiquitaire. Cette comparaison de phénotype peut être par la suite raffinée en inactivant le gène de façon tissu spécifique en utilisant des animaux exprimant la recombinaise de façon tissu spécifique ou en délivrant la recombinaise de façon ectopique (virus, ...).

La seconde application concerne le développement d'un modèle dans lequel a été substitué le gène humain au gène de souris (S1). Ce modèle humanisé fournit un outil d'étude in vivo de molécules et de drogues qui est plus relevant que l'animal de départ. Plus l'écart entre la protéine Humaine et la protéine de souris est importante et plus ce modèle est pertinent.

Dans un second temps, il est possible de disposer d'un modèle KO pour cette forme humaine en croisant cet animal humanisé avec un animal exprimant une recombinaise (Cre ou

Flp suivant les sites SR utilisés) de façon ubiquitaire. Ce nouveau modèle est particulièrement adapté pour étudier la toxicité et les effets non spécifiques des molécules testées. Cette étude peut être par la suite raffinée en inactivant le gène de façon tissu spécifique en utilisant des animaux exprimant la recombinaise de façon tissu spécifique ou en délivrant la recombinaise de façon ectopique (virus, ...).

Exemple 1 : Souris KO exempte d'expression du gène *Prpf31*

La stratégie employée permet de générer un modèle de souris KO exempt d'expression du gène *Prpf31* par la génération d'un modèle qui a à la fois un point de mutation Kin qui mime la mutation A216P chez l'homme et une délétion conditionnelle du domaine NOP.

Ce modèle sera obtenu en introduisant un point de mutation dans l'exon 7 en introduisant deux sites LoxP aux deux extrémités de l'exon 7. La stratégie de développement du modèle est exposée sur la figure 2. Sur cette figure, les triangles rouges représentent les sites LoxP, les doubles triangles jaunes représentent les sites FRT, la boîte pourpre représente la cassette de résistance Neo, la boîte orange représente la cassette de sélection négative DTA, les boîtes noires représentent les exons et l'astérisque au dessus de l'exon 7 schématise la mutation qui y est introduit.

L'injection des cellules cibles ES 2 dans les blastocystes va permettre de générer le point de mutation Kin du modèle. L'accouplement de la lignée cellulaire Kin avec une souris exprimant Cre de manière tissu-spécifique va permettre la génération de souris modèles KO du domaine NOP tissu spécifique.

La stratégie choisie comprend les étapes suivantes :

- Clonage et séquençage de la région ciblée du gène murin 129Sv *Prpf31*
- Construction d'un vecteur de ciblage permettant à la fois de générer un point de mutation et un KO conditionnel du domaine NOP
- Electroporation du vecteur de ciblage dans les cellules ES et sélection des événements de recombinaison homologue
- Excision in vitro via Flp de la cassette de sélection
- Injection des cellules ES dans les blastocystes
- Génération de souris chimères et génotypage pour la caractérisation des mutants hétérozygotes.

Exemple 2 : Souris KO exempte d'expression du gène *C3Ra*

La stratégie employée permet de générer un modèle de souris KO exempt d'expression du gène murin **C3Ra** par la génération d'un modèle qui a à la fois un Kin de la forme Humaine et une délétion conditionnelle du gène humain.

La stratégie de développement du modèle est exposée sur la figure 1. Sur cette figure,
5 les triangles bleus représentent les sites LoxP, les doubles triangles rouges représentent les sites FRT, la boîte blanche représente la cassette de résistance Neo.

L'action de la recombinaise Flip permet d'obtenir un modèle humanisé propre.

L'action de la recombinaise Cre tissue spécifique ou ubiquitaire permet d'obtenir un KO tissue spécifique ou ubiquitaire.

REVENDICATIONS

1. Combinaison de séquences d'ADN comprenant une première séquence S1 et une deuxième séquence S2 destinées à être insérées dans un gène cible d'une cellule eucaryote, caractérisée en ce que :
- la première séquence S1 comprend, dans le sens 5'-3', les éléments suivants :
- un premier site SR1 de reconnaissance d'une première recombinaise R1, et
 - un acide nucléique hétérologue,
- et
- la deuxième séquence S2 comprend, dans le sens 5'-3', les éléments suivants :
- un premier site SR2 de reconnaissance d'une recombinaise R2,
 - une séquence d'acide nucléique d'un gène codant pour un marqueur de sélection fonctionnel dans la cellule eucaryote dans le génome de laquelle il sera intégré
 - un second site SR2 de reconnaissance de la recombinaise R2, et
 - le cas échéant, un second site SR1 de reconnaissance de la recombinaise R1.
2. Combinaison selon la revendication 1, caractérisée en ce que l'acide nucléique hétérologue de la séquence S1 est un acide nucléique codant pour un exon du gène cible, comprenant au moins une mutation, destiné à être dans le génome de la cellule cible de manière que l'exon muté de la séquence S1 vienne remplacer l'exon non muté du gène cible.
3. Combinaison selon la revendication 1, caractérisée en ce que la séquence d'acide nucléique dans la séquence S1 est une séquence d'acide nucléique comprenant, liés de manière fonctionnelle, des éléments fonctionnels nécessaires à l'initiation de la transcription et/ou des éléments fonctionnels nécessaires à l'arrêt de la transcription.
4. Combinaison selon la revendication 3, caractérisée en ce que l'acide nucléique hétérologue de la séquence S1 comprend une séquence d'intérêt choisie parmi les séquences suivantes :
- une séquence d'acide nucléique codant pour un polypeptide, en particulier une protéine d'intérêt, ou
 - une séquence d'acide nucléique codant pour un RNA, en particulier un RNA antisens ou un RNAi, miRNA, shRNA ou siRNA.

5. Combinaison selon la revendication 4, caractérisée en ce que la séquence d'acide nucléique codant pour un polypeptide est un ADNc, plus préférentiellement un ADNc codant pour un polypeptide d'origine humaine.

5

6. Combinaison selon l'une des revendications 1 à 5, caractérisée en ce que les sites de reconnaissance de recombinaisons SR1 et SR2 sont choisis, indépendamment l'un de l'autre, parmi les sites Lox P, Lox P modifiés et FRT.

10 7. Cellules eucaryotes dans le génome desquelles est intégré la combinaison de séquences selon l'une des revendications 1 à 6.

8. Cellules selon la revendication 7, caractérisées en ce qu'elles sont choisies parmi les cellules de mammifères, à l'exception de l'homme, préférentiellement de mammifères rongeurs, de préférence rat ou souris.

15

9. Embryons d'animaux, à l'exception de l'homme, obtenus à partir des cellules selon l'une des revendications 7 ou 8.

20 10. Animaux obtenus à partir des embryons selon la revendication 9, comprenant des cellules selon l'une des revendications 7 ou 8, à l'exception de l'homme.

1/1

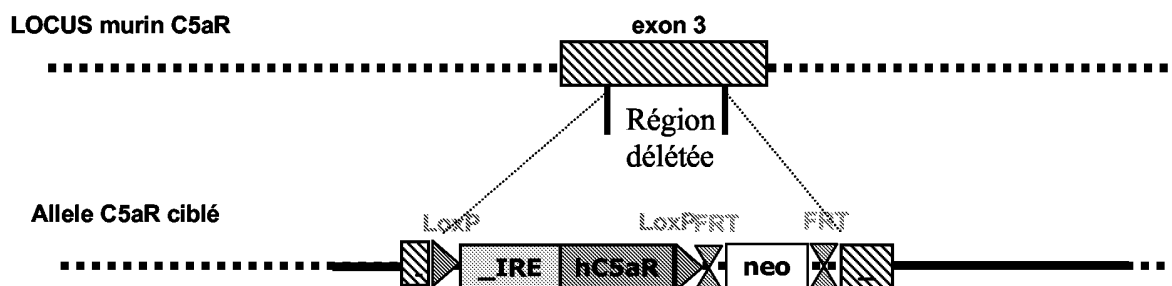


Fig. 2

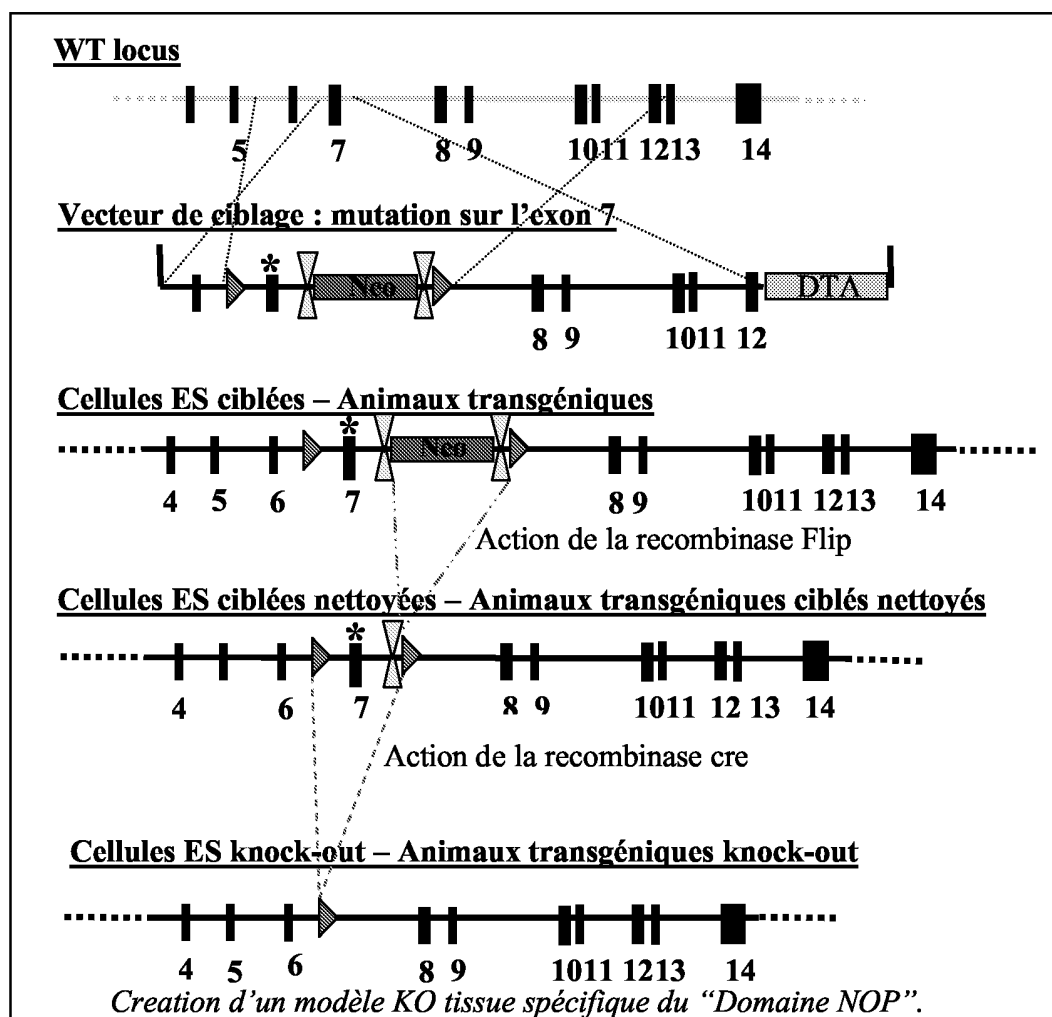


Fig. 1