

(12) DEMANDE DE BREVET D'INVENTION BELGE

(41) Date de publication : 23/07/2024

(21) Numéro de demande : BE2022/6100

(22) Date de dépôt : 28/12/2022

(62) Divisée de la demande de base :

(62) Date de dépôt demande de base :

(51) Classification internationale : C12N 15/11, C12N 15/63, C12N 15/85, C12N 15/90

(30) Données de priorité :

(71) Demandeur(s) :

QUIDDITAS SA
SA
6900, MARCHE-EN-FAMENNE
Belgique

(72) Inventeur(s) :

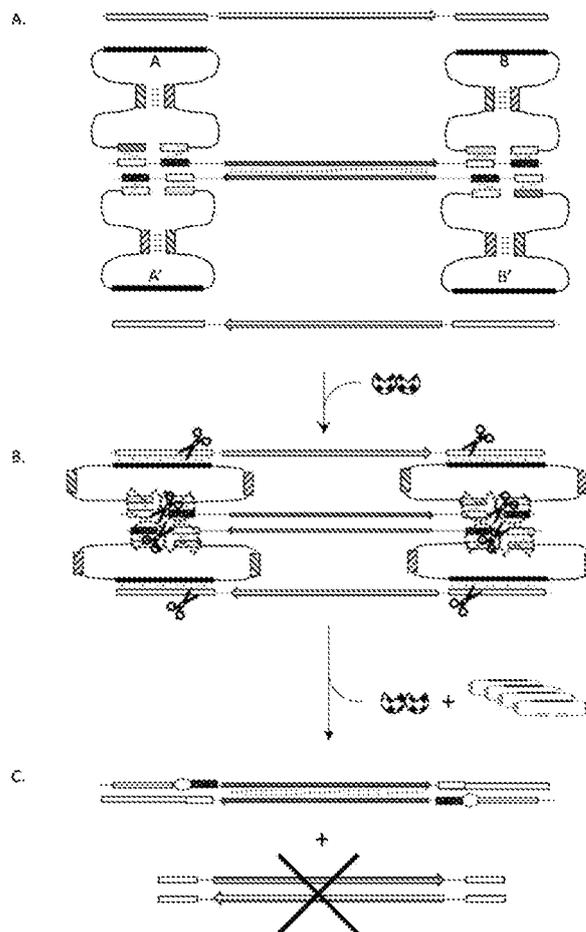
CHERBONNEAU François
6900 MARCHE-EN-FAMENNE
Belgique

CHERBONNEAU PRUNEVIEILLE Aurore
6900 MARCHE-EN-FAMENNE
Belgique

(54) Composition et son utilisation pour le traitement des maladies héréditaires

(57) L'invention concerne une composition comprenant - une première molécule d'acides nucléiques simple brin comprenant une région complémentaire d'une cible et des sites de reconnaissance d'une transposase, - une deuxième molécule d'acides nucléiques simple brin comprenant une région complémentaire d'une cible et des sites de reconnaissance d'une transposase, et - une troisième molécule simple brin comprenant une séquence codant une partie du récepteur gamma de l'interleukine 2 ou IL2-R γ , ou codant le récepteur IL2-R γ complet, et des sites de liaison à une transposase.

[Fig. 3]



Description

Titre de l'invention : Composition et son utilisation pour le traitement des maladies héréditaires

- 5 L'invention concerne une composition et son utilisation pour le traitement des maladies héréditaires.
- La DICS-X ou Déficit Immunitaire Combiné Sévère lié à l'X est causé par des mutations du gène $IL2R\gamma$ induisant un dysfonctionnement du système immunitaire causant des infections sévères, des fièvres et des éruptions cutanées.
- 10 La sous-unité gamma du récepteur à l'interleukine-2 ($IL2-R\gamma$), régulée par le gène $IL2-R\gamma$, est commune à beaucoup de récepteurs à l'interleukine (IL-2, IL-4, IL-7, IL-9, IL-15) et est présente à la surface des cellules souches hématopoïétiques. Cette protéine, en association avec d'autres, est essentielle pour la formation des lymphocytes. Ainsi, la mutation du gène $IL2R\gamma$ entraînant un blocage de la production du récepteur $IL2-R\gamma$ entrainera par la même un blocage
- 15 de la différenciation cellulaire en lymphocytes, principalement les lymphocytes T et Natural Killer (NK). De même, l'absence de lymphocytes T et NK dans l'organisme entraîne une inactivation des lymphocytes B formés.
- Ce gène $IL2R\gamma$ se situe sur le chromosome X. Les mâles ne présentant qu'une seule copie de ce chromosome et donc du gène $IL2R\gamma$ sont automatiquement atteints par la maladie en cas
- 20 de mutation du gène. En revanche, les femelles possédant 2 copies du gène, sont dites « porteuses saines » et ne présenteront des symptômes de la maladie que dans des cas extrêmement rares.
- Les symptômes de la DICS-X apparaissent généralement à l'âge de 3 à 6 mois. La plupart des enfants atteints présentent des retards de croissance, des éruptions cutanées au niveau
- 25 oral et génital, ainsi que de multiples infections persistantes entraînant la mort au bout d'1 ou 2 ans en l'absence de traitement.
- La prévalence de la DICS-X est très difficilement évaluable due à un manque de diagnostic, mais a été estimée à environ 1 naissance sur 50 000.
- Le traitement de référence est la reconstitution du système immunitaire par transplantation de
- 30 moelle osseuse. Le taux de succès de cette approche est directement lié au niveau de compatibilité entre le donneur de moelle et l'enfant malade. Si la compatibilité est élevée (généralement la moelle provenant du frère ou de la sœur du patient), une rémission totale sera atteinte dans 20% des cas. En revanche, dans 80% des cas, le patient développera des symptômes plus ou moins intenses de GVHD (Graft Versus Host Disease), c'est-à-dire que
- 35 les cellules immunitaires de la moelle transplantée attaqueront le patient.
- Le traitement alternatif est la thérapie génique basée sur l'utilisation de virus (lentivirus ou AAV), où le gène $IL2R\gamma$ est réintroduit dans un échantillon prélevé des cellules souches

hématopoïétiques du patient puis réinjectées pour repeupler le système immunitaire. Cette approche, toujours en cours d'essais cliniques, se veut efficace dans 80% des cas, mais nécessite plusieurs séances de traitements extrêmement coûteux et entraînent dans 20% des cas le développement de leucémies dû au fait que cette approche virale n'est pas ciblée et n'apporte pas le nouveau gène IL2R γ dans son locus originel.

La demande de brevet US20150152436A1 décrit de nouveaux développements thérapeutiques qui tendent à adapter l'approche d'édition du génome CRISPR/CAS9 au traitement de la DICS-X. Toutefois, le polymorphisme associé à cette pathologie, ainsi que la possible insertion de l'ADNc du gène IL2R γ de plus de 600 paires de bases, demeure un frein technique notable en vue d'une utilisation clinique.

Aussi, le besoin de fournir un traitement efficace demeure.

L'un des buts de l'invention est de pallier les inconvénients de l'art antérieur.

Un but de l'invention est de proposer une composition capable de permettre le remplacement du gène IL2R γ de manière efficace et complète.

Un autre but de l'invention est de fournir une méthode ou un médicament capable de traiter des pathologies liées au déficit dudit gène.

L'invention concerne une composition comprenant :

- une première molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence A permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence complémentaire étant liée en 5' à une première séquence riches en T de 40 à 60 nucléotides de long et en 3' à une deuxième séquence riches en T de 40 à 60 nucléotides de long, lesdites première et deuxième séquences riches en T comprenant respectivement un premier et un deuxième domaine de 6 à 12 nucléotides riches en G/C, la séquence du premier domaine étant complémentaire de la séquence du deuxième domaine, lesdits premier et deuxième domaines étant positionnés de 15 à 52 nucléotides de ladite séquence A, ladite première molécule comprenant en son extrémité 5' au moins une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite transposase,

- une deuxième molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence B permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence B complémentaire étant liée en 5' à une troisième séquence riche en T de 40 à 60 nucléotides de long et en 3' à une quatrième séquence riche en T de 40 à 60 nucléotides de long, lesdites troisième et quatrième séquences riche en T comprenant respectivement un troisième et un quatrième domaine de 6 à 12 nucléotides riches en G/C, la

séquence du troisième domaine étant complémentaire de la séquence du quatrième domaine, lesdits troisième et quatrième domaines étant positionnés de 15 à 52 nucléotides de ladite séquence B, ladite deuxième molécule comprenant en son extrémité 5' au moins la première séquence orientés 5'-3' de reconnaissance de ladite transposase et en son extrémité 3' la

5 deuxième séquence de reconnaissance de ladite transposase, ladite séquence B étant une séquence complémentaire dudit acide nucléique d'intérêt, la séquence A étant positionnée en 5' d'une région d'intérêt dudit acide nucléique d'intérêt et la séquence B positionnée en 3' de la région d'intérêt dudit acide nucléique d'intérêt, et

- une troisième molécule simple brin comprenant

10 * dans sa partie 5', au moins une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule,

* dans sa partie 3' au moins une séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, et

15 * une région intermédiaire située entre séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, ladite région intermédiaire comprenant une séquence codant une partie du récepteur gamma de l'interleukine 2 ou IL2-R γ , ou codant le récepteur IL2-R γ complet,

20 les première et troisième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase et les deuxième et troisième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase.

L'invention repose sur la constatation faite par les inventeurs que l'utilisation d'un complexe

25 molécule tel que décrit ci-dessus, et faisant intervenir une transposase bactérienne permet un remplacement ciblé et efficace du gène IL2R γ , qui code l'IL2-R γ .

Cela signifie que l'invention concerne un complexe moléculaire comprenant une première molécule d'acides nucléiques simple brin, une deuxième molécule d'acides nucléiques simple brin et une troisième molécule d'acides nucléiques simple brin, ladite troisième molécule

30 d'acides nucléiques simple brin comprenant ou étant constituée essentiellement en son extrémité 5' d'au moins une séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase, et ladite troisième molécule d'acides nucléiques simple brin comprenant ou étant constituée essentiellement en son extrémité 3' d'au moins une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite

35 transposase,

ledit complexe étant tel que les première, deuxième et troisième molécules d'acides nucléiques simple brin sont appariées selon la complémentarité des bases définie par Watson

et Crick de sorte à définir deux sites double brin de liaison de ladite transposase dans l'extrémité 5' de la troisième molécule et deux sites double brin de liaison de ladite transposase dans l'extrémité 3' de la troisième molécule.

5 L'invention repose sur l'observation inattendue faite par l'inventeur que l'utilisation de guides simples brins spécifiques capables de cibler une région d'un acide nucléique d'intérêt permettent de mobiliser de manière contrôlée et « site spécifique » des transposases, et ainsi utiliser les propriétés de recombinaison desdites transposases pour réaliser du remplacement de séquences dans des molécules d'intérêt.

10 Le complexe moléculaire susmentionné est en fait l'unité de base de la technologie définie dans l'invention. Cette unité de base sert à guider les recombinases sur un site spécifique où doit avoir lieu la recombinaison, et donc le remplacement de séquence. Contrairement au système CRISPR/Cas9 qui requiert la présence de séquences de type PAM (NGG), l'outil moléculaire défini ici peut être utilisé sur n'importe quelle séquence cible, quelle qu'en soit sa séquence.

15 Le complexe moléculaire susmentionné est donc l'unité de base qui devra être complétée par :

- une région homologie d'une région 5' de la séquence cible, ladite région d'homologie étant représentée par la séquence A et

20 - une région homologie d'une région 3' de la séquence cible, ladite région d'homologie étant représentée par la séquence B.

Il s'agit donc ici d'un produit intermédiaire de l'outil tel que décrit ci-après.

Le complexe moléculaire est constitué de trois molécules d'acides nucléiques simples brins, qui peuvent être des molécules d'ADN, des molécules d'ARN ou des molécules mixtes d'ARN et d'ADN.

25 Ces trois molécules sont partiellement complémentaires entre elles deux à deux, selon la complémentarité des bases d'acides nucléiques définie par Watson et Crick, c'est-à-dire qu'une Adénine s'apparie à un Thymidine ou un Uracile, et une Cytosine s'apparie à une Guanine, et réciproquement.

30 Plus particulièrement les trois molécules formant le complexe susmentionné comprennent chacune la séquence d'un des brins d'une molécule double brin correspondant à la séquence de fixation d'une transposase. Aussi, chaque molécule simple brin comprend donc une « demie séquence » de fixation de transposase et ne peut donc pas permettre une interaction avec ladite transposase correspondante. En revanche, lorsque les trois molécules du complexe interagissent ensemble, par appariement des bases tel que défini ci-dessus, une
35 molécule double brin est ainsi formée, reconstituant un site de liaison double brin de ladite transposase, celle-ci pouvant ainsi interagir avec la molécule formée.

La première molécule.

La première molécule du complexe est la molécule qui comprendra, une fois modifiée, une séquence d'acides nucléiques qui permettra de cibler spécifiquement une région d'intérêt d'une molécule d'acide nucléique d'intérêt. Cette séquence d'intérêt sera choisie par l'utilisateur du système selon la cible choisie. Cette séquence d'intérêt sera insérée dans la première molécule dudit complexe au niveau de la région A. Cette région A correspond à minima à deux acides nucléiques entre lesquels sera inséré la séquence permettant de cibler la molécule cible. Au vu de la structure orientée des acides nucléiques (sens 5'-3'), il est important que la séquence permettant de cibler la région d'intérêt soit positionnée dans le bon sens, afin que l'appariement soit possible avec la séquence cible.

10 Aussi, avantageusement, la région A comprend un ou plusieurs sites reconnaissant des enzymes de restrictions afin de favoriser une insertion orientée. Un ou plusieurs des sites suivants peuvent être présents dans la région A :

[Table 1]

		AA/CGTT	AclI
		A/AGCTT	HindIII
		AAT/ATT	SspI
		/AATT	MluCI
		A/CATGT	PciI
		A/CCGGT	AgeI
		ACCTGC(4/8)	BfuAI BspMI
		A/CCWGGT	SexAI
		A/CGCGT	MluI
		ACGGC(12/14)	BceAI
		A/CGT	HpyCH4IV
		ACN/GT	HpyCH4III
SEQ ID NO :	4	(10/15)ACNNNNGTAYC(12/7)	BaeI
SEQ ID NO :	5	(9/12)ACNNNNNCTCC(10/7)	BsaXI
		A/CRYGT	AflIII
		A/CTAGT	SpeI
		ACTGG(1/-1)	BsrI
		ACTGGG(5/4)	BmrI
		A/GATCT	BglII
		AGC/GCT	AfeI
		AG/CT	AluI
		AGG/CCT	StuI
		AGT/ACT	Scal-

		AT/CGAT	Clal BspDI
SEQ ID NO :	6	ATCTATGTCGGGTGCGGAGAAAGAGGTA AT(-15/-19)	PI-SceI
		ATGCA/T	Nsil
		AT/TAAT	Asel
		ATTT/AAAT	Swal
SEQ ID NO :	7	(11/13)CAANNNNNGTGG(12/10)	CspCI
		C/AATTG	Mfel
		CACCTGC(4/8)	PaqCI
		CACGAG	Nb.BssSI
		CACGAG(-5/-1)	BssSI-v2
		CACGTC(-3/-3)	BmgBI
		CAC/GTG	PmlI
		CACNNN/GTG	DrallI
SEQ ID NO :	8	CACNN/NNGTG	AleI-v2
		CAGCAG(25/27)	EcoP15I
		CAG/CTG	PvuII
		CAGNNN/CTG	AlwNI
		CAGTG(2/0)	BtsIMutI
		CA/TATG	NdeI
		CATG/	NlaIII
		/CATG	FatI
		C/ATG	CviAII
SEQ ID NO :	9	CAYNN/NNRTG	MslI
		CC(12/16)	FspEI
SEQ ID NO :	10	CCANNNNN/NNNNTGG	XcmI
SEQ ID NO :	11	CCANNNNN/NTGG	BstXI
SEQ ID NO :	12	CCANNNN/NTGG	PfiMI
		CCATC(4/5)	BccI
		C/CATGG	NcoI
		CCCAGC(-5/-1)	BseYI
		CCCGC(4/6)	FauI
		CCC/GGG	SmaI
		C/CCGGG(0/-1)CCD	TspMI XmaI Nt.CviPII

		CCDG(10/14)	LpnPI
		CCGC(-3/-1)	Acil
		CCGC/GG	SacII
		CCGCTC(-3/-3)	BsrBI
		C/CGG	MspI HpaII
		CC/NGG	ScrFI
		/CCNGG	StyD4I
		C/CNNGG	BsaJI
SEQ ID NO :	13	CCNNNNN/NGG	BsII
		C/CRYGG	BtgI
		CC/SGG	NciI
		C/CTAGG	AvrII
		CCTC(7/6)	MnII
		CCTCAGC	Nb.BbvCI
		CCTCAGC(-5/-7)	Nt.BbvCI
		CCTCAGC(-5/-2)	BbvCI
		CCTGCA/GG	SbfI
		CCTNAGC(-5/-2)	Bpu10I
		CC/TNAGG	Bsu36I
SEQ ID NO :	14	CCTNN/NNNAGG	EcoNI
		CCTTC(6/5)	HpyAV
		/CCWGG	PspGI
		CC/WGG	BstNI
		C/CWWGG	StyI
SEQ ID NO :	15	(10/12)CGANNNNNNTGC(12/10)	BcgI
		CGAT/CG	PvuI
		CG/CG	BstUI
		C/GGCCG	EagI
		CG/GWCCG	RsrII
		CGRY/CG	BsiEI
		C/GTACG	BsiWI
		CGTCTC	BsmBI-v2
		CGTCTC(1/5)	Esp3I
		CGWCG/	Hpy99I
		CMG/CKG	MspA1I

SEQ ID NO :	16	CNNNNNNNNNNN/NNNNNNNNNG	AbaSI
		CNNR(9/13)	MspJI
		CR/CCGGYG	SgrAI
		C/TAG	BfaI
		CTCAG(9/7)	BspCNI
		C/TCGAG	XhoI PaeR7I
		CTCTTC(1/4)	EarI
		CTGAAG(16/14)	AcuI
		CTGCA/G	PstI
		CTGGAG(16/14)	BpmI
		C/TNAG	DdeI
		C/TRYAG	SfcI
		C/TTAAG	AflII
		CTTGAG(16/14)	BpuEI
		C/TYRAG	SmlI
		C/YCGRG	BsoBI Aval
		GAAGA(8/7)	MbolI
		GAAGAC(2/6)	BbsI
SEQ ID NO :	17	GAANN/NN TTC	XmnI
		GAATGC(1/-1)	BsmI
		GAATGC	Nb.BsmI
		G/AATTC	EcoRI
		GACGC(5/10)	HgaI
		GACGT/C	AatII
		GAC/GTC	ZraI
		GACN/NGTC	PfiFI Tth111I
SEQ ID NO :	18	GACNN/NGTC	PshAI
SEQ ID NO :	19	GACNNN/NGTC	AhdI
SEQ ID NO :	20	GACNNNN/NGTC	DrdI
		GAG/CTC	Eco53kI
		GAGCT/C	SacI
		GAGGAG(10/8)	BseRI
		GAGTC(4/-5)	Nt.BstNBI
		GAGTC(4/5)	PleI
		GAGTC(5/5)	MlyI

		G/ANTC	Hinfl
		GAT/ATC	EcoRV
		GA/TC	DpnI
		/GATC	Sau3AI DpnI I MboI
SEQ ID NO :	21	GATNN/NNATC	BsaBI
		G/AWTC	TfiI
		GCAATG	Nb.BsrDI
		GCAATG(2/0)	BsrDI
		GCAGC(8/12)	BbvI
		GCAGTG(2/0)	BtsI-v2
		GCAGTG	Nb.BtsI
SEQ ID NO :	22	GCANNNN/NTGC	BstAPI
		GCATC(5/9)	SfaNI
		GCATG/C	SphI
		GCCC/GGGC	SrfI
		GCCGAG(21/19)	NmeAIII
		G/CCGGC	NgoMIV
		GCC/GGC	NaeI
SEQ ID NO :	23	GCCNNNN/NGGC	BglI
		GCGAT/CGC	AsiSI
		GCGATG(10/14)	BtgZI
		GCG/C	HhaI
		G/CGC	HinP1I
		G/CGCGC	BssHII
		GC/GGCCGC	NotI
		GC/NGC	Fnu4HI
		GCN/NGC	Cac8I
SEQ ID NO :	24	GCNNNNN/NNGC	MwoI
		G/CTAGC	NheI
		GCTAG/C	BmtI
		GCTCTTC(1/-7)	Nt.BspQI
		GCTCTTC(1/4)	SapI BspQI
		GC/TNAGC	BplI
		G/CWGC	ApeKI Tsel

		GDGCH/C	Bsp1286I
		GGATC(4/5)	AlwI
		GGATC(4/-5)	Nt.AlwI
		G/GATCC	BamHI
		GGATG(9/13)	FokI
		GGATG(2/0)	BtsCI
		GG/CC	HaeIII
		GGCCGG/CC	FseI
SEQ ID NO :	25	GGCCNNNN/NGGCC	SfiI
		G/GCGCC	KasI
		GG/CGCC	NarI
		GGCGC/C	PluTI
		GGC/GCC	SfoI
		GG/CGCGCC	AscI
		GGCGGA(11/9)	EciI
		GGGAC(10/14)	BsmFI
		GGGCC/C	ApaI
		G/GGCCC	PspOMI
		G/GNCC	Sau96I
		GGN/NCC	NlaIV
		G/GTACC	Acc65I
		GGTAC/C	KpnI
		GGTCTC(1/5)	BsaI v2
		GGTGA(8/7)	HphI
		G/GTNACC	BstEII
		G/GWCC	Avall
		G/GYRCC	BanI
		GKGC/M/C	BaeGI
		GR/CGYC	BsaHI
		GRGCY/C	BanII
		GT/AC	RsaI
		G/TAC	CviQI
		GTATAC	BstZ17I
		GTATCC(6/5)	BciVI
		G/TCGAC	Sall

		GTCTC(1/5)	BsmAI BcoD I
		GTCTC(1/-5)	Nt.BsmAI
		G/TGCAC	ApaLI
		GTGCAG(16/14)	BsgI
		GT/MKAC	Accl
		GTN/NAC	Hpy166II
		/GTSAC	Tsp45I
		GTT/AAC	HpaI
		GTTT/AAAC	PmeI
		GTY/RAC	HincII
		GWGCW/C	BsiHKAII
		NNCASTGNN/	TspRI
		R/AATTY	ApoI
		RCATG/Y	NspI
		R/CCGGY	BsrFI-v2
		R/GATCY	BstYI
		RGCGC/Y	HaeII
		RG/CY	CviKI-1
		RG/GNCCY	EcoO109I
		RG/GWCCY	PpuMI
SEQ ID NO :	26	TAACTATAACGGTCCTAAGGTAGCGAA(- 9/-13)	I-CeuI
		TAC/GTA	SnaBI
SEQ ID NO :	27	TAGGGATAACAGGGTAAT(-9/-13)	I-SceI
		T/CATGA	BspHI
		T/CCGGA	BspEI
		TCCRAC(20/18)	MmeI
		T/CGA	TaqI-v2
		TCG/CGA	NruI
		TCN/GA	Hpy188I
		TC/NNGA	Hpy188III
		T/CTAGA	XbaI
		T/GATCA	BclI
		TG/CA	HpyCH4V

		TGC/GCA	Fspl
SEQ ID NO :	28	TGGCAAACAGCTATTATGGGTATTATGGG T(-13/-17)	PI-Pspl
		TGG/CCA	MscI
		T/GTACA	BsrGI
		T/TAA	MseI
		TTAAT/TAA	PacI
		TTA/TAA	PsiI-v2
		TT/CGAA	BstBI
		TTT/AAA	DraI
		VC/TCGAGB	PspXI
		W/CCGGW	BsaWI
		YAC/GTR	BsaAI
		Y/GGCCR	EaeI

Évidemment, dans le cadre d'une synthèse chimique de la première molécule, il n'est pas nécessaire de disposer de sites de clonage (ou d'insertion) de la séquence permettant de cibler la région cible, mais de bien prendre la précaution de donner une séquence
5 correctement orientée. Cela est bien évidemment toutefois possible.

La première molécule est en outre constituée de part et d'autre de la région A, de séquences riches en A/T, ou s'il s'agit d'ARN en A/U, afin de permettre une certaine flexibilité de la structure. Par riche en A/T, ou en A/U, on entend dans l'invention une séquence qui comprend plus de 50% de A ou T, ou U, de préférence plus de 50% de T ou de U, par rapport
10 au nombre total de nucléotides constituant la séquence. Ces séquences de part et d'autre de la région A ont une taille en nucléotides variant de 10 nucléotides à 60 nucléotides.

La flexibilité de ces séquences flanquantes de la région A, du fait de la présence de nombreux A, T ou U, peut avoir pour effet de permettre une recombinaison par l'intermédiaire des recombinases trop peu contrôlées, voire même lors que le complexe n'a pas encore
15 reconnu la molécule cible.

Aussi, afin de pallier ce problème, des séquences riches en G/C sont introduites dans chacune des séquences riches en A/T, notamment riches en T, ou A/U, bordant la région A. Ces régions riches en G/C sont constituées de 6 à 12 nucléotides, dont la quantité de C ou de G est supérieure à 50% des nucléotides contenus dans ladite séquence riche en G/C.

Pour stabiliser la structure de la première molécule, et comme décrit ci-dessus éviter une recombinaison intempestive, les régions riches en G/C sont positionnées de 15 à 52
20 nucléotides par rapport à l'extrémité de la région A.

Pour plus de clarté, si la région A consiste en 3 nucléotides, le nucléotide central correspondant à la position 0, la région riche en A/T ou A/U débutera à gauche en position -2, et à droite en position +2. Dès lors, à gauche, la région riche en G/C sera positionnée de la position -17 à la position -54, et du côté droit de la position +17 à la position +54.

5 Autre élément important, la séquence riche en G/C à droite (ou 5') de la région A est nécessairement complémentaire (selon le principe d'appariement de Watson et Crick) de la région riche en G/C de la région à droite (ou 3') de la région A. Aussi, la première molécule simple brin s'apparie-t-elle avec elle-même au niveau des régions riches en G/C, ce qui empêche toute recombinaison par les transposases, tant qu'il n'y aura pas d'interaction avec
10 la séquence cible complémentaire de la région qui sera insérée dans la région A de la première molécule.

Enfin la première molécule comprend à son extrémité 5' une séquence correspondant à un premier site de liaison à une transposase, et à son extrémité 3' un deuxième site de liaison à ladite transposase.

15 Le premier site de liaison et le deuxième site de liaison sont avantageusement les mêmes, et surtout correspondent tous les deux au même brin du site de liaison double brin de ladite transposase. Cela signifie que le premier site de liaison à la transposase présent dans la région 5' de la première molécule ne peut s'apparier intégralement, et donc de manière stable avec le site de liaison à la transposase présent dans la région 3'.

20 Le premier site de liaison et le deuxième site de liaison sont avantageusement les mêmes, mais correspondent chacun à un brin différent du site double brin de liaison à la transposase. Aussi, par exemple, si le premier site de liaison à la transposase correspond au brin sens, le deuxième site de liaison à la transposase correspond à la séquence du brin complémentaire. Il est alors possible d'avoir deux configurations : i) soit le deuxième site de liaison qui
25 correspond au brin complémentaire est orienté dans le sens 3'-5', dans ce cas il pourra s'apparier avec le premier site de liaison à la transposase et former le site double brin, ii) soit le deuxième site de liaison qui correspond au brin complémentaire est orienté dans le sens 5'-3', et dans ce cas ne pourra pas s'apparier avec la première séquence de liaison à la transposase, les séquences, du fait de leur orientation n'étant pas complémentaires. Dans le
30 cas i) susmentionné, si la première molécule simple brin s'auto-apparie au niveau des premier et deuxième sites de liaison, il ne sera pas possible de former le complexe susmentionné, car il n'y aura plus de région complémentaire simple brin disponible pour s'apparier avec la deuxième molécule de sorte à former deux sites de liaison double brin à la transposase.

35 Aussi, la première molécule, lorsqu'elle est dépourvue de séquence complémentaire de la région cible dans la partie A, ou lorsqu'elle contient une telle séquence cible mais que cette dernière n'interagit pas (ne s'apparie pas) avec ladite séquence cible, forme une structure

tridimensionnelle où toute la molécule est simple brin à l'exception de la région correspondant aux régions riches en G/C qui s'apparient entre elles.

Une représentation schématique de sous forme appariée est représentée aux **Figures 1A à 1E**.

5 La deuxième molécule.

La deuxième molécule du complexe est de structure similaire à la première molécule. Les explications susmentionnées s'appliquent donc mutatis mutandis. Toutefois, dans la deuxième molécule la région B (ou séquence B) ainsi que les sites de liaison à la transposase sont organisés de manière différente.

10 Tout d'abord la région B est différente de la région A de la première molécule. En effet, il ne peut, dans l'objectif d'une recombinaison orientée, y avoir de compétition pour la même cible d'intérêt entre la première et la deuxième molécule.

Aussi, est-il nécessaire que la région B corresponde à une deuxième séquence complémentaire de la séquence cible, cette deuxième séquence complémentaire de la
15 séquence cible étant positionnée en 3' par rapport à la première séquence reconnue par la séquence complémentaire correspondant à la séquence A de la première molécule.

Par conséquent, lorsque la première molécule et la deuxième molécule seront appariées à la séquence cible, la séquence cible sera bordée par ces deux molécules, la première molécule étant localisée en 5' et la deuxième molécule étant positionnée en 3'. La région de
20 la séquence cible localisée entre la région d'interaction avec la première molécule et la région d'interaction avec la deuxième molécule correspond à la séquence qui sera remplacée par celle de la troisième molécule.

La troisième molécule

La troisième molécule du complexe susmentionné est plus simple que les deux premières
25 (la première et la deuxième molécule). La troisième molécule comprend, dans sa partie 5', un site de liaison à la transposase qui est complémentaire du site de liaison à ladite transposase présent dans la partie 3' de la première molécule. De ce fait, lorsque le complexe sera formé, le ½ site de liaison à la transposase (simple brin) localisé en 3' de la première molécule pourrait s'apparier avec le ½ site de liaison (simple brin) à la transposase localisé en 5' de la deuxième
30 molécule de sorte à former un site de liaison double brin à la transposase, site double brin sur lequel la transposase pourra venir se fixer.

Dans la partie 3' de la troisième molécule se trouve un site de liaison à la transposase qui est complémentaire du site de liaison à ladite transposase présent dans la partie 5' de la
35 deuxième molécule. De ce fait, lorsque le complexe sera formé, le ½ site de liaison à la transposase (simple brin) localisé en 3' de la troisième molécule pourrait s'apparier avec le ½ site de liaison (simple brin) à la transposase localisé en 5' de la deuxième molécule de sorte

à former un site de liaison double brin à la transposase, site double brin sur lequel la transposase pourra venir se fixer.

Entre la partie 5' qui comprend un ½ site de liaison à la transposase et la partie 3' qui comprend un ½ site de liaison à la transposase, la troisième molécule comprend une séquence de remplacement, c'est à dire la séquence qui à terme remplacera la séquence cible. cette
5 séquence de remplacement est bordée en 5' par au moins un site de restriction et en 3' par au moins un site de restriction ; ces deux sites de restriction étant distincts. Aussi, la troisième molécule comprend dans le sens 5' vers 3' : un ½ site de liaison à la transposase complémentaire du site de liaison à la transposase de la première molécule, suivi d'au moins
10 un site de restriction, suivi de la séquence de remplacement, suivie d'au moins un site de restriction différent du site de restriction en amont de la séquence de remplacement, suivi enfin d'un ½ site de liaison à la transposase complémentaire du site de liaison à la transposase de la deuxième molécule.

La séquence de remplacement correspond quant à elle à tout ou partie du gène codant la
15 protéine IL2-R γ , i.e. tout ou partie du gène ILR2 γ .

La chaîne gamma commune (γ c) (ou CD132), est également connue sous le nom de sous-unité gamma du récepteur de l'interleukine-2 ou IL-2RG, ou IL2-R γ . Il s'agit d'une sous-unité du récepteur des cytokines qui est commune aux complexes récepteurs d'au moins six récepteurs d'interleukines différents : interleukine 2 (IL-2), interleukine 4 (IL-4), interleukine 7 (IL-7), interleukine 9 (IL-9), interleukine 15 (IL-15) et le récepteur de l'interleukine-21. Cette
20 chaîne est une glycoprotéine qui fait partie de la famille des récepteurs de cytokines de type I exprimés sur la plupart des populations de lymphocytes. Le gène ILR2 γ se trouve sur le chromosome X des mammifères.

IL2-R γ est exprimée à la surface des cellules sanguines immatures de la moelle osseuse.
25 Une extrémité de la protéine réside à l'extérieur de la cellule où elle se lie aux cytokines et l'autre extrémité de la protéine réside à l'intérieur de la cellule où elle transmet des signaux au noyau de la cellule. La chaîne gamma commune s'associe à d'autres protéines pour diriger les cellules hématopoïétiques vers la formation de lymphocytes. Le récepteur dirige également la croissance et la maturation des sous-types de lymphocytes : les lymphocytes T, les
30 lymphocytes B et les cellules tueuses naturelles (en anglais *natural killer* ou NK).

La composition selon l'invention permet d'obtenir le complexe sus-décrit. La **Figure 2** illustre schématiquement le complexe formé entre les première, deuxième et troisième molécules selon l'invention.

Le complexe formé à partir des molécules de la composition de l'invention, lorsque les trois
35 molécules sont correctement appariées, comprend deux paires de sites de liaison double brin de reconnaissance d'une transposase :

- la première paire étant obtenue par l'hybridation de la première molécule avec la troisième molécule, et

- la seconde paire étant obtenue par l'hybridation de la deuxième molécule avec la troisième molécule.

5 La séquence A contenue dans la première molécule est complémentaire du même brin de l'acide nucléique qui est complémentaire de la séquence B contenu dans la deuxième molécule. En d'autres termes, la séquence A contenue dans la première molécule et la séquence B contenue dans la deuxième molécule sont capables de s'hybrider au même acide nucléique simultanément, car les deux séquences A et B ne reconnaissent pas la même
10 séquence.

Pour clarifier encore plus le propos, l'intérêt dans la présente invention est de proposer une première molécule et une deuxième molécule, toutes deux telles que définies ci-dessus, les séquences A et B respectives étant telles qu'elles sont capables de reconnaître pour l'une, une séquence localisée en 5' de la séquence cible de l'acide nucléique d'intérêt et pour l'autre,
15 une séquence localisée en 3' de la même séquence cible de l'acide nucléique d'intérêt. Les séquences A et B sont donc complémentaires de régions bordant la séquence d'intérêt, que l'on souhaite remplacer, de la molécule d'acides nucléiques d'intérêt.

Les première et deuxième molécules d'intérêt sont donc essentielles pour cibler spécifiquement la molécule d'intérêt, afin d'encadrer la séquence à remplacer.

20 La troisième molécule de l'ensemble, quant à elle, est celle qui apporte la molécule d'acides nucléiques qui contient la séquence de remplacement, i.e. une séquence correspondant à tout ou partie du gène codant l'IL2-R γ .

D'un point de vue mécanistique, le complexe selon l'invention est tel qu'il est constitué de ses trois molécules, la première et la deuxième molécule étant structurellement organisées
25 dans l'espace de sorte que leurs régions riches en G/C soient appariées.

De part et d'autre de la troisième molécule, c'est à dire en 5' et en 3', du fait de l'hybridation avec les première et deuxième molécules, deux paires de sites de liaison à une transposase permettent, lorsqu'elles sont présentes, à des dimères de transposases de se lier à l'ensemble.

On notera que les sites de liaison à la transposase de la première molécule peuvent être
30 les mêmes que ceux de la deuxième molécule, ou être différents. Dans le cas où les séquences de liaison sont les mêmes, les dimères de transposases en 5' de la troisième molécule (par hybridation de la partie 5' de la troisième molécule avec la première molécule), et en 3' de la troisième molécule (par hybridation de la partie 3' de la troisième molécule avec la deuxième molécule) seront les mêmes. Aussi, à titre d'exemple, si les sites de liaison sont
35 tous des sites de liaison de la transposase de Tn5, l'ensemble sera associé à 2 dimères de transposases de Tn5.

Il est également possible que les sites de liaison de la première molécule et de la deuxième molécule ne reconnaissent pas la même transposase. Dans ce cas, et selon la définition donnée ci-dessus, la partie 5' de la troisième molécule formera par hybridation avec la première molécule deux sites double brin de liaison à une première transposase, et partie 3' de la troisième molécule formera par hybridation avec la deuxième molécule deux sites double brin de liaison à une deuxième transposase.

Si maintenant le complexe susmentionné, lié à deux dimères de transposase, est mis en présence d'une molécule d'acides nucléiques d'intérêt dont la partie 5' est complémentaire de la séquence A de la première molécule de l'ensemble et dont la partie 3' est complémentaire de la séquence B de la deuxième molécule de l'ensemble, alors il y aura appariement entre la molécule d'acides nucléiques d'intérêt et l'ensemble au niveau des régions A et B sus-décrites. Cette interaction aura pour conséquence de rompre l'interaction des deux séquences riches en G/C de chacune des première et deuxième molécules de l'ensemble. Dès lors, la troisième molécule et la molécule d'acides nucléiques d'intérêt seront rapprochées spatialement et les transposases pourront exercer leur activité de tagmentation, dont le résultat sera le remplacement de la séquence de la molécule d'intérêt, bordée par les séquences complémentaires des séquences A et B, par la séquence de la troisième molécule de l'ensemble, qui se trouve entre les demi-sites de liaison à la /aux transposases en 5' et en 3'.

Dans l'invention, les première, deuxième et troisième molécules de l'ensemble sont avantageusement des molécules constituées de désoxyribonucléotides, afin de former des molécules d'ADN simple brin.

De manière encore plus avantageuse, les première, deuxième et troisième molécules de l'ensemble sont des molécules hybrides ADN/ARN, où le « squelette » des molécules est de l'ADN, et les séquences de reconnaissance A et B de la molécule d'acides nucléiques d'intérêt, et la région centrale de la troisième molécule sont de l'ARN. Ceci est particulièrement avantageux lorsque le remplacement de séquence que permet l'invention doit se faire directement sur une molécule d'ARN.

La **Figure 2-A** illustre l'interaction entre la molécule d'intérêt et l'ensemble selon l'invention.

De manière avantageuse, l'invention concerne le complexe susmentionné, où ladite première molécule ou ladite deuxième molécule, ou les deux est/sont couplée(s) à une enzyme, notamment par le biais d'un nucléotide modifié. Cette enzyme a pour but de favoriser le remplacement de la molécule d'intérêt. Il peut s'agir :

- d'une hélicase, enzyme capable d'ouvrir une molécule double brin qui serait superenroulée ou encore associée à des protéines telles que des histones,
- une topoisomérase, enzyme agissant sur la structure topologique de l'ADN en générant des coupures transitoires,

- une ligase qui permet de former une liaison phosphodiester entre une extrémité 5' phosphate d'un nucléotide et l'extrémité 3' OH d'un autre nucléotide,

- une polymérase, qui synthétisera une molécule d'acides nucléiques à partir d'un site d'initiation, 3'OH libre, dans le sens 5'-> 3', notamment en recopiant un brin complémentaire antiparallèle selon le modèle de Watson et Crick.

Il est également possible de combiner deux ou plusieurs desdites enzymes pour disposer de tout le matériel enzymatique nécessaire pour permettre le remplacement de séquence prévu dans le cadre de l'invention.

Dans un aspect particulier, de l'invention, la première molécule de l'ensemble peut contenir dans sa partie 5', plus précisément entre le premier site de liaison à la transposase et la région A, un ou plusieurs nucléotides modifiés. De la même manière, la troisième molécule de l'ensemble peut contenir dans sa partie 3', plus précisément entre le premier site de liaison à la transposase et la région 3, un ou plusieurs nucléotides modifiés.

Ce nucléotide modifié est notamment modifié par greffage d'une chaîne carbonée substituée avec une étiquette protéique, ou *tag* en anglais, ou encore une molécule permettant une interaction spécifique telle que la streptavidine ou la biotine.

De telles modifications permettent alors de lier spécifiquement, à la première molécule de l'ensemble, des enzymes qui peuvent servir à favoriser la tagmentation, et le remplacement de séquence. Il sera particulièrement avantageux d'avoir par exemple un greffage streptavidine, qui permettra de greffer une hélicase biotinylée (inversement hélicase greffée à la streptavidine et nucléotide biotinylé) servant à ouvrir une molécule double brin. Il est également possible d'envisager le greffage avec une ligase biotinylée (ou couplé à la streptavidine), afin de relier le brin recombiné coté 3'.

Il est également possible d'associer à la première molécule ou à la deuxième molécule de l'ensemble un oligonucléotide, de sorte que ledit oligonucléotide s'appariera sur une région prédéterminée de la dite première ou deuxième molécule. Cet oligonucléotide est alors avantageusement couplé à une molécule de greffage comme expliqué ci-dessus.

De manière avantageuse, l'invention concerne la composition susmentionnée, où ladite séquence A comprend une séquence complémentaire de la séquence d'une première région du gène codant l'IL2-R γ , et où ladite séquence B comprend une séquence complémentaire de la séquence d'une deuxième région du gène codant l'IL2-R γ , ladite séquence A et la dite séquence B étant deux séquences différentes, lesdites première et deuxième régions du gène codant l'IL2-R γ , encadrant une région comprenant une séquence codant une partie de l'IL2-R γ , ou codant le récepteur IL2-R γ complet.

Comme expliqué ci-dessus, afin d'opérer le remplacement du gène cible par la troisième molécule, il est avantageux que les séquences de la partie A de la première molécule et de la partie B de la deuxième molécule soient capables :

- de reconnaître la même molécule cible, et

- 5 - d'encadrer la région cible à remplacer, ce qui signifie que la première molécule doit, par l'intermédiaire de sa région A, reconnaître une séquence en amont de la séquence à remplacer de la molécule cible et la deuxième molécule doit, par l'intermédiaire de sa région B, reconnaître une séquence en aval de la séquence à remplacer de la molécule cible, ou l'inverse, i.e. la première molécule doit, par l'intermédiaire de sa région A,
10 reconnaître une séquence en aval de la séquence à remplacer de la molécule cible et la deuxième molécule doit, par l'intermédiaire de sa région B, reconnaître une séquence en amont de de la séquence à remplacer de la molécule cible.

Il est ainsi possible de remplacer n'importe quelle séquence cible par n'importe quelle séquence de remplacement, du moment que les séquences A et B respectivement des
15 première et deuxième molécules reconnaissent la région cible et l'encadre.

Dans l'invention il est particulièrement avantageux que tout ou partie du gène codant l'IL2-R γ soit remplacé par la séquence de la troisième molécule. Ceci est particulièrement avantageux lorsque la séquence cible est une séquence qui comprend une ou plusieurs mutations par rapport à la séquence sauvage de référence, par exemple, une substitution d'un ou plusieurs
20 nucléotides, contigus ou espacés, une délétion d'un ou plusieurs nucléotides, contigus ou espacés, ou encore une insertion d'un ou plusieurs nucléotides, contigus ou espacés.

Par tout le gène codant l'IL2-R γ on entend toute la séquence dudit gène comprenant les éléments de régulation 5', 3' les exons et les introns.

Par « parties du gène codant l'IL2-R γ » on entend dans l'invention un élément du gène qui ne
25 permet pas de coder intégralement la protéine IL2-R γ . Il peut s'agir, sans être limitatif, d'un intro, d'un exon, de plusieurs introns et exons sans que la totalité de l'unité traductionnelle soit complète... Il peut être particulièrement avantageux de ne remplacer qu'une partie du gène codant l'IL2-R γ , afin par exemple de ne remplacer qu'un seul exon au sein duquel se trouve une mutation. Dans ce cas, il pourrait être avantageux de choisir une séquence
30 complémentaire de la région A de la première molécule dans l'intron précédent la séquence à remplacer, et une partie complémentaire de la région B dans l'intron suivant l'exon à remplacer. Bien évidemment, l'homme du métier saura, fonction de la substitution qui souhaitera opérer, choisir entre le gène entier ou une partie du gène, et ainsi définir les régions A et B les plus pertinentes.

35 De manière avantageuse, l'invention concerne la composition telle que définie ci-dessus, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3'

de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence orientée 5'-3' de reconnaissance de ladite transposase et

où la troisième molécule comprend en son extrémité 5' une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule suivie d'une deuxième séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule.

De manière plus avantageuse, l'invention concerne la composition telle que définie ci-dessus, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence orientée 5'-3' de reconnaissance de ladite transposase,

ladite deuxième molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence orientée 5'-3' de reconnaissance de ladite transposase, et

où la troisième molécule comprend en son extrémité 5' une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule suivie d'une deuxième séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et en son extrémité 3' une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule suivie d'une deuxième séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la deuxième molécule.

Ici il est défini une possibilité de format de première et deuxième molécule de l'invention. Cette configuration est schématisée en figure 1B. Bien évidemment, dans le cadre de cet exemple de configuration des première, deuxième et troisième molécules, entre les deux demis sites de transposase des parties 5' et 3' de la troisième molécule se trouve la séquence de remplacement, ici, tout ou partie du gène codant l'IL2-R γ .

De manière avantageuse, l'invention concerne la composition susmentionnée, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite transposase, suivie d'une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule et

où la troisième molécule comprend en son extrémité 5' une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule.

De manière encore plus avantageuse, l'invention concerne la composition susmentionnée, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de

reconnaissance de ladite transposase, suivie d'une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule, la deuxième molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite transposase, ladite première séquence de reconnaissance de ladite transposase de l'extrémité 5' étant précédé d'une deuxième séquence complémentaire de la ladite deuxième séquence reconnaissance de ladite transposase de la première molécule, et

où la troisième molécule comprend en son extrémité 5' une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et en son extrémité 3' une séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule.

Ici il est défini une autre possibilité de format de première et deuxième molécules de l'invention. Cette configuration est schématisée en figure 1C et en figure 1D. Bien évidemment, dans le cadre de cet exemple de configuration des première, deuxième et troisième molécules, entre les deux demis sites de transposase des parties 5' et 3' de la troisième molécule se trouve la séquence de remplacement, ici, tout ou partie du gène codant l'IL2-R γ .

De manière avantageuse, l'invention concerne la composition susmentionnée, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' d'une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule suivi d'une deuxième séquence de reconnaissance de ladite transposase, et

où la troisième molécule comprend en son extrémité 5' une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule.

De manière encore plus avantageuse, l'invention concerne la composition susmentionnée, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' d'une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule suivie d'une deuxième séquence de reconnaissance de ladite transposase, la deuxième molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' d'une deuxième séquence de reconnaissance de ladite transposase de la première molécule, ladite première séquence étant suivie d'une deuxième séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule, et

où la troisième molécule comprend en son extrémité 5' une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule

et en son extrémité 3' une séquence complémentaire de la première séquence de reconnaissance de la transposase de la deuxième molécule.

Cette configuration est schématisée en figure 1E. Bien évidemment, dans le cadre de cet exemple de configuration des première, deuxième et troisième molécules, entre les deux
5 demis sites de transposase des parties 5' et 3' de la troisième molécule se trouve la séquence de remplacement, ici, tout ou partie du gène codant l'IL2-R γ .

Avantageusement, l'invention concerne la composition susmentionnée, où ladite transposase est une transposase bactérienne, notamment une transposase choisie parmi Tn5, Tn9, Tn10 ou Tc1/mariner.

10 De manière avantageuse, la transposase susmentionnée est une transposase de type bactérienne choisie parmi la transposase du transposon Tn5, la transposase du transposon Tn9, la transposase du transposon Tn10, Tn903, Tn602 ou encore la transposase du transposon Tc1, ou plus généralement de la superfamille des transposons mariner.

D'autres exemples de transposases qui peuvent être utilisées dans le cadre de l'invention
15 sont : la transposase de *Vibrio harveyi* (transposase caractérisée par Agilent et utilisée dans le produit SureSelect QXT), la transposase MuA et un site de reconnaissance de la transposase Mu comprenant les séquences terminales R1 et R2, la transposase du transposon Tn552 de *Staphylococcus aureus*, la transposase du transposon Tn7, la transposase Tn/O et IS10, la transposase du transposon Tn3.

20 La transposase de Tn5 est la plus connue. Elle est codée par le gène Tnp du transposon Tn5. La transposase initie la transposition en formant un dimère de transposases qui se fixe sur ses séquences cibles. Dans le contexte de ce complexe, la transposase catalyse ensuite quatre réactions de transfert de phosphoryle (coupure d'ADN, formation en épingle à cheveux d'ADN, résolution en épingle à cheveux et transfert de brin dans l'ADN cible) entraînant l'intégration
25 du transposon dans son nouveau site d'ADN : il s'agit de la tagmentation.

C'est sur le principe de cette tagmentation que l'invention repose. En utilisant les propriétés de tagmentation des transposases, il est possible d'insérer une séquence dans une autre de manière ciblée, grâce au complexe susmentionné.

Aussi dans le cadre de l'invention, lorsqu'il est fait référence à une transposase, il est fait
30 référence à l'une des transposases susmentionnées, à savoir les transposases des transposons Tn5, Tn9, Tn10 ou Tc1/mariner (ou transposases mutées afin d'augmenter leur activité de transposition ou de tagmentation).

Dans l'invention, lorsque plusieurs transposases sont utilisées simultanément, l'une se liant au complexe formé par la première molécule et à la troisième molécule, et l'autre se liant au
35 complexe formé par la deuxième molécule et la troisième molécule, on privilégiera les couples de transposases issus de transposons Tn5 et Tn10.

De manière avantageuse, la première et la deuxième séquence de reconnaissance de la transposase sont des séquences de reconnaissance de la transposase de Tn5 ayant l'une des séquences suivantes :

- CTGtCTCTTataCAcAtcT (SEQ ID NO : 29),
- 5 - CTGACTCTTataCACAagT (SEQ ID NO : 30), et
- CTGtCTCTTgatCAgATCT (SEQ ID NO : 31).

En conséquence les séquences complémentaires correspondantes sont les suivantes :

- AgaTgTGtatAAGAGaCAG (SEQ ID NO : 32), complémentaire de la séquence SEQ ID NO: 29,
- 10 - ActTGTGtatAAGAGTCAG (SEQ ID NO : 33), complémentaire de la séquence SEQ ID NO: 30, et
- AGATcTGatcAAGAGaCAG (SEQ ID NO : 34), complémentaire de la séquence SEQ ID NO: 31.

D'autres séquences de reconnaissance d'une transposase sont les suivantes :

- 15 Tn5MErev,
5'-[phos]CTGTCTCTTATACACATCT-3' (SEQ ID NO : 35)
Tn5ME-A (Illumina FC-121-1030),
5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3'; (SEQ ID NO : 36)
et Tn5ME-B (Illumina FC-121-1031),
- 20 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3' (SEQ ID NO : 37)

Encore d'autres séquences de reconnaissance de la transposase sont les suivantes :

- séquence sens SEQ ID NO : i
- séquence antisens SEQ ID NO : i+1,
où i varie de 38 à 192.
- 25 Cela signifie par exemple que les couples de séquences sens et antisens respectifs suivants sont envisagés : SEQ ID NO : 38 et SEQ ID NO: 39, SEQ ID NO : 40 et SEQ ID NO: 41, SEQ ID NO : 42 et SEQ ID NO: 43, SEQ ID NO : 44 et SEQ ID NO: 45, SEQ ID NO : 46 et SEQ ID NO: 47, SEQ ID NO : 48 et SEQ ID NO: 49, SEQ ID NO : 50 et SEQ ID NO: 51, SEQ ID NO : 52 et SEQ ID NO: 53, SEQ ID NO : 54 et SEQ ID NO: 55, SEQ ID NO : 56 et SEQ ID NO: 57,
- 30 SEQ ID NO : 58 et SEQ ID NO: 59, SEQ ID NO : 60 et SEQ ID NO: 61, SEQ ID NO : 62 et SEQ ID NO: 63, SEQ ID NO : 64 et SEQ ID NO: 65, SEQ ID NO : 66 et SEQ ID NO: 67, SEQ ID NO : 68 et SEQ ID NO: 69, SEQ ID NO : 70 et SEQ ID NO: 71, SEQ ID NO : 72 et SEQ ID NO: 73, SEQ ID NO : 74 et SEQ ID NO: 75, SEQ ID NO : 76 et SEQ ID NO: 77, SEQ ID NO : 78 et SEQ ID NO: 79, SEQ ID NO : 80 et SEQ ID NO: 81, SEQ ID NO : 82 et SEQ ID NO: 83,
- 35 SEQ ID NO : 84 et SEQ ID NO: 85, SEQ ID NO : 86 et SEQ ID NO: 87, SEQ ID NO : 88 et SEQ ID NO: 89, SEQ ID NO : 90 et SEQ ID NO: 91, SEQ ID NO : 92 et SEQ ID NO: 93, SEQ ID NO : 94 et SEQ ID NO: 95, SEQ ID NO : 96 et SEQ ID NO: 97, SEQ ID NO : 98 et SEQ ID

NO: 99, SEQ ID NO : 100 et SEQ ID NO: 101, SEQ ID NO : 102 et SEQ ID NO: 103, SEQ ID NO : 104 et SEQ ID NO: 105, SEQ ID NO : 106 et SEQ ID NO: 107, SEQ ID NO : 108 et SEQ ID NO: 109, SEQ ID NO : 110 et SEQ ID NO: 111, SEQ ID NO : 112 et SEQ ID NO: 113, SEQ ID NO : 114 et SEQ ID NO: 115, SEQ ID NO : 116 et SEQ ID NO: 117, SEQ ID NO : 118 et
5 SEQ ID NO: 119, SEQ ID NO : 120 et SEQ ID NO: 121, SEQ ID NO : 122 et SEQ ID NO: 123, SEQ ID NO : 124 et SEQ ID NO: 125, SEQ ID NO : 126 et SEQ ID NO: 127, SEQ ID NO : 128 et SEQ ID NO: 129, SEQ ID NO : 130 et SEQ ID NO: 131, SEQ ID NO : 132 et SEQ ID NO: 133, SEQ ID NO : 134 et SEQ ID NO: 135, SEQ ID NO : 136 et SEQ ID NO: 137, SEQ ID NO : 138 et SEQ ID NO: 139, SEQ ID NO : 140 et SEQ ID NO: 141, SEQ ID NO : 142 et SEQ ID
10 NO: 143, SEQ ID NO : 144 et SEQ ID NO: 145, SEQ ID NO : 146 et SEQ ID NO: 147, SEQ ID NO : 148 et SEQ ID NO: 149, SEQ ID NO : 150 et SEQ ID NO: 151, SEQ ID NO : 152 et SEQ ID NO: 153, SEQ ID NO : 154 et SEQ ID NO: 155, SEQ ID NO : 156 et SEQ ID NO: 157, SEQ ID NO : 158 et SEQ ID NO: 159, SEQ ID NO : 160 et SEQ ID NO: 161, SEQ ID NO : 162 et SEQ ID NO: 163, SEQ ID NO : 164 et SEQ ID NO: 165, SEQ ID NO : 166 et SEQ ID NO:
15 167, SEQ ID NO : 168 et SEQ ID NO: 169, SEQ ID NO : 170 et SEQ ID NO: 171, SEQ ID NO : 172 et SEQ ID NO: 173, SEQ ID NO : 174 et SEQ ID NO: 175, SEQ ID NO : 176 et SEQ ID NO: 177, SEQ ID NO : 178 et SEQ ID NO: 179, SEQ ID NO : 180 et SEQ ID NO: 181, SEQ ID NO : 182 et SEQ ID NO: 183, SEQ ID NO : 184 et SEQ ID NO: 185, SEQ ID NO : 186 et SEQ ID NO: 187, SEQ ID NO : 188 et SEQ ID NO: 189, SEQ ID NO : 190 et SEQ ID NO: 191,
20 et SEQ ID NO : 192 et SEQ ID NO: 193.

De manière avantageuse, le premier domaine riche en G/C de la première molécule (ou de la deuxième molécule) correspond à la séquence suivante GGCGATCGC (SEQ ID NO : 194) de sorte que le deuxième domaine riche en G/C sera le même. En effet, du fait du repliement de la molécule sur elle-même, le deuxième domaine riche en G/C sera dans une orientation
25 complémentaire et antiparallèle par rapport au premier domaine riche en G/C, et l'interaction se fera au niveau de la région palindromique (souligné dans la séquence ci-dessus).

Le premier et le deuxième domaines riches en G/C peut également être la séquence suivante GCGGCGATCGGC (SEQ ID NO : 195). Les explications ci-dessus s'appliquent *mutatis mutandis*.

30 D'autres séquences des domaines riches en G/C de la première molécule ou de la deuxième molécule, peuvent être les suivantes :

- premier domaine riche en G/C de séquence GGTCGC (SEQ ID NO : 196) et le second domaine riche en G/C de séquence GCGACC (SEQ ID NO : 197).

Ces exemples sont donnés à titre illustratif seulement et ne sauraient limiter la portée de
35 l'invention.

Dans un mode de réalisation avantageux, les séquences riches en A/T de la première molécule dudit complexe sont constituées essentiellement, ou constituées de A ou de T.

De manière encore plus avantageuse, la séquence riche en A/T de la première molécule dudit complexe est constituée de T.

De manière avantageuse, l'invention concerne la composition susmentionnée où le gène codant l'IL2-R γ comprend la séquence SEQ ID NO : 1, ou la séquence SEQ ID NO : 2 ou encore la séquence SEQ ID NO : 3.

La séquence SEQ ID NO : 1 correspond au gène IL2R γ humain complet de référence, référencé sous le numéro GenBank: AY692262.1. Ce gène est également connu sous les noms suivants : P64, CIDX, IMD4, CD132, SCIDX, IL-2RG ou encore SCIDX1. Il fait 7130 bases.

10 Les exons du gène correspondent aux séquences délimitées de la manière suivante :

Exon 1 : de 1956 (premier nucléotide du codon initiateur ATG) à 2070, et intron 1 : de 2071 à 2448,

Exon 2 : de 2449 à 2602, et intron 2 : de 2603 à 2810,

Exon 3 : de 2811 à 2995, et intron 3 : de 2996 à 3203,

15 Exon 4 : de 3204 à 3343, et intron 4 : de 3344 à 4108,

Exon 5 : de 4109 à 4271, et intron 5 : de 4272 à 4803,

Exon 6 : de 4804 à 4900, et intron 6 : de 4901 à 5152,

Exon 7 : de 5153 à 5222, et intron 7 : de 5223 à 5577, et

Exon 8 : de 5578 à 5763 (dernier nucléotide du codon stop),

20 La séquence SEQ ID NO : 2 correspond à un fragment de la SEQ ID NO : 1, puisque sa séquence commence en position 1956 de la séquence SEQ ID NO : 1 et se termine au codon stop en position 5763 de la SEQ ID NO : 1.

La séquence SEQ ID NO : 3 correspond à la région codante (ou CDS en anglais) du gène. Celle correspondant à la réunion des exons susmentionnés. Cette séquence est référencée sous le numéro GenBank: AK314932.1.

25 Au vu du découpage susmentionné, l'homme du métier pourra choisir une partie du gène tel que défini ci-dessus, dans le but de réaliser une substitution partielle du gène. Par exemple, et sans être limitatif, s'il est souhaité de remplacer l'exon 1, il sera possible de cibler avec la première molécule l'intron 1, avec la deuxième molécule l'intron 2 tandis que la troisième molécule comprendra la séquence de l'exon 2.

De manière encore plus avantageuse, l'invention concerne la composition telle que définie ci-dessus, où les premières deuxième et troisième molécules sont choisies parmi les triplets tels que définis dans le tableau 2 ci-dessous.

35 De manière avantageuse, l'invention concerne une composition comprenant l'un des triplets de première deuxième et troisième molécules décrits dans le tableau 2 suivant

[Table 2]

Triplet	Première molécule SEQ ID NO :	Deuxième molécule SEQ ID NO :	Troisième molécule SEQ ID NO :
#1	198	199	200
#2	201	202	203
#3	204	205	206
#4	207	208	209
#5	210	211	212
#6	213	214	215
#7	216	217	218
#8	219	220	221
#9	222	223	224
#10	225	226	227
#11	228	229	230
#12	231	232	233
#13	234	235	236
#14	237	238	239
#15	240	241	242
#16	243	244	245
#17	246	247	248
#18	249	250	251
#19	252	253	254
#20	255	256	257
#21	258	259	260
#22	261	262	263
#23	264	265	266
#24	267	268	269
#25	270	271	272
#26	273	274	275
#27	276	277	278
#28	279	280	281
#29	282	283	284
#30	285	286	287
#31	288	289	290
#32	291	292	293
#33	294	295	296
#34	297	298	299
#35	300	301	302
#36	303	304	305
#37	306	307	308
#38	309	310	311
#39	312	313	314
#40	315	316	317

#41	318	319	320
#42	321	322	323
#43	324	325	326
#44	327	328	329
#45	330	331	332
#46	333	334	335
#47	336	337	338
#48	339	340	341
#49	342	343	344
#50	345	346	347
#51	348	349	350
#52	351	352	353
#53	354	355	356
#54	357	358	359
#55	360	361	362
#56	363	364	365
#57	366	367	368
#58	369	370	371
#59	372	373	374
#60	375	376	377
#61	378	379	380
#62	381	382	383
#63	384	385	386
#64	387	388	389
#65	390	391	392
#66	393	394	395
#67	396	397	398
#68	399	400	401
#69	402	403	404
#70	405	406	407
#71	408	409	410
#72	411	412	413
#73	414	415	416
#74	417	418	419
#75	420	421	422
#76	423	424	425
#77	426	427	428
#78	429	430	431
#79	432	433	434
#80	435	436	437
#81	438	439	440
#82	441	442	443

#83	444	445	446
#84	447	448	449
#85	450	451	452
#86	453	454	455
#87	456	457	458
#88	459	460	461
#89	462	463	464
#90	465	466	467
#91	468	469	470
#92	471	472	473
#93	474	475	476
#94	477	478	479
#95	480	481	482
#96	483	484	485
#97	486	487	488
#98	489	490	491
#99	492	493	494
#100	495	496	497
#101	498	499	500
#102	501	502	503
#103	504	505	506
#104	507	508	509
#105	510	511	512
#106	513	514	515
#107	516	517	518
#108	519	520	521
#109	522	523	524
#110	525	526	527
#111	528	529	530
#112	531	532	533
#113	534	535	536
#114	537	538	539
#115	540	541	542
#116	543	544	545
#117	546	547	548
#118	549	550	551
#119	552	553	554
#120	555	556	557
#121	558	559	560
#122	561	562	563
#123	564	565	566
#124	567	568	569

#125	570	571	572
#126	573	574	575
#127	576	577	578
#128	579	580	581
#129	582	583	584
#130	585	586	587
#131	588	589	590
#132	591	592	593
#133	594	595	596
#134	597	598	599
#135	600	601	602
#136	603	604	605
#137	606	607	608
#138	609	610	611
#139	612	613	614
#140	615	616	617
#141	618	619	620
#142	621	622	623
#143	624	625	626
#144	627	628	629
#145	630	631	632
#146	633	634	635
#147	636	637	638
#148	639	640	641
#149	642	643	644
#150	645	646	647
#151	648	649	650
#152	651	652	653
#153	654	655	656
#154	657	658	659
#155	660	661	662
#156	663	664	665
#157	666	667	668
#158	669	670	671
#159	672	673	674
#160	675	676	677
#161	678	679	680
#162	681	682	683
#163	684	685	686
#164	687	688	689
#165	690	691	692
#166	693	694	695

#167	696	697	698
#168	699	700	701
#169	702	703	704
#170	705	706	707
#171	708	709	710
#172	711	712	713
#173	714	715	716
#174	717	718	719
#175	720	721	722
#176	723	724	725
#177	726	727	728
#178	729	730	731
#179	732	733	734
#180	735	736	737
#181	738	739	740
#182	741	742	743
#183	744	745	746
#184	747	748	749
#185	750	751	752
#186	753	754	755
#187	756	757	758
#188	759	760	761
#189	762	763	764
#190	765	766	767
#191	768	769	770
#192	771	772	773
#193	774	775	776
#194	777	778	779
#195	780	781	782
#196	783	784	785
#197	786	787	788
#198	789	790	791
#199	792	793	794
#200	795	796	797
#201	798	799	800
#202	801	802	803
#203	804	805	806
#204	807	808	809
#205	810	811	812
#206	813	814	815
#207	816	817	818
#208	819	820	821

#209	822	823	824
#210	825	826	827
#211	828	829	830
#212	831	832	833
#213	834	835	836
#214	837	838	839
#215	840	841	842
#216	843	844	845
#217	846	847	848
#218	849	850	851
#219	852	853	854
#220	855	856	857
#221	858	859	860
#222	861	862	863
#223	864	865	866
#224	867	868	869
#225	870	871	872
#226	873	874	875
#227	876	877	878
#228	879	880	881
#229	882	883	884
#230	885	886	887
#231	888	889	890
#232	891	892	893
#233	894	895	896
#234	897	898	899
#235	900	901	902
#236	903	904	905
#237	906	907	908
#238	909	910	911
#239	912	913	914
#240	915	916	917
#241	918	919	920
#242	921	922	923
#243	924	925	926
#244	927	928	929
#245	930	931	932
#246	933	934	935
#247	936	937	938
#248	939	940	941
#249	942	943	944
#250	945	946	947

#251	948	949	950
#252	951	952	953
#253	954	955	956
#254	957	958	959
#255	960	961	962
#256	963	964	965
#257	966	967	968
#258	969	970	971
#259	972	973	974
#260	975	976	977
#261	978	979	980
#262	981	982	983
#263	984	985	986
#264	987	988	989
#265	990	991	992
#266	993	994	995
#267	996	997	998
#268	999	1000	1001
#269	1002	1003	1004
#270	1005	1006	1007
#271	1008	1009	1010
#272	1011	1012	1013
#273	1014	1015	1016
#274	1017	1018	1019
#275	1020	1021	1022
#276	1023	1024	1025
#277	1026	1027	1028
#278	1029	1030	1031
#279	1032	1033	1034
#280	1035	1036	1037
#281	1038	1039	1040
#282	1041	1042	1043
#283	1044	1045	1046
#284	1047	1048	1049
#285	1050	1051	1052
#286	1053	1054	1055
#287	1056	1057	1058
#288	1059	1060	1061
#289	1062	1063	1064
#290	1065	1066	1067
#291	1068	1069	1070
#292	1071	1072	1073

#293	1074	1075	1076
#294	1077	1078	1079
#295	1080	1081	1082
#296	1083	1084	1085
#297	1086	1087	1088
#298	1089	1090	1091
#299	1092	1093	1094
#300	1095	1096	1097
#301	1098	1099	1100
#302	1101	1102	1103
#303	1104	1105	1106
#304	1107	1108	1109
#305	1110	1111	1112
#306	1113	1114	1115
#307	1116	1117	1118
#308	1119	1120	1121
#309	1122	1123	1124
#310	1125	1126	1127
#311	1128	1129	1130
#312	1131	1132	1133

Aussi avantageusement, l'invention concerne la composition susmentionnée, ladite composition comprenant un triplet choisi parmi les triplets #1 à #312 susmentionnés.

De manière avantageuse, l'invention concerne la composition décrite ci-dessus, comprenant en outre

- 5 ** une quatrième molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence A' permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence complémentaire étant liée en 5' à une cinquième
- 10 T de 40 à 60 nucléotides de long, lesdites cinquième et sixième séquences riches en T comprenant respectivement un cinquième et un sixième domaine de 6 à 12 nucléotides riches en G/C, la séquence du cinquième domaine étant complémentaire de la séquence du sixième domaine, lesdits cinquième et sixième domaines étant positionnés de 15 à 52 nucléotides de ladite séquence A', ladite quatrième molécule comprenant en son extrémité 5' au moins une
- 15 première séquence orientés 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite transposase,
- ** une cinquième molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence B' permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide

nucléique d'intérêt, ladite séquence B' complémentaire étant liée en 5' à une septième séquence riche en T de 40 à 60 nucléotides de long et en 3' à une huitième séquence riche en T de 40 à 60 nucléotides de long, ladite septième et huitième séquence riche en T comprenant respectivement un septième et un huitième domaine de 6 à 12 nucléotides riches en G-C, la séquence du septième domaine étant complémentaire de la séquence du huitième domaine, ledit septième et huitième domaine étant positionnés de 15 à 52 nucléotides de ladite séquence B', ladite cinquième molécule comprenant en son extrémité 5' au moins la première séquence orientée 5'-3' de reconnaissance de ladite transposase et en son extrémité 3' la deuxième séquence de reconnaissance de ladite transposase,

5

10 ladite séquence B' étant une séquence complémentaire dudit acide nucléique d'intérêt, la séquence A' étant positionnée en 5' d'une région d'intérêt dudit acide nucléique d'intérêt et la séquence B' positionnée en 3' de la région d'intérêt dudit acide nucléique d'intérêt, et

- une sixième molécule simple brin comprenant
 - * dans sa partie 5', au moins une séquence complémentaire de ladite cinquième séquence de reconnaissance de ladite transposase de la première molécule,
 - * dans sa partie 3' au moins une séquence complémentaire de ladite quatrième séquence de reconnaissance de ladite transposase de la deuxième molécule, et
 - * une région intermédiaire située entre séquence complémentaire de ladite quatrième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite cinquième séquence de reconnaissance de ladite transposase de la deuxième molécule, ladite région intermédiaire comprenant une séquence complémentaire, et antiparallèle, de la séquence codant le récepteur gamma de l'interleukine 2 ou IL2-R γ contenue dans la troisième molécule de la première composition,

20 les quatrième et sixième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase et les cinquième et sixième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase.

25 De manière avantageuse, la composition selon l'invention comprend 6 molécules, soit deux triplets qui permettent un remplacement double brin d'une molécule cible. Ce remplacement pourra se faire seulement si les régions A, B et A' et B' sont correctement choisies de sorte que la région d'intérêt à remplacer soit correctement encadrée.

30 Bien évidemment, toutes les définitions données pour la composition comprenant 3 molécules s'appliquent mutatis mutandis pour une composition comprenant 6 molécules.

35 De manière avantageuse, l'invention concerne la composition décrite ci-dessus, où

- ladite séquence A comprend une séquence complémentaire de la séquence d'une première région du gène codant l'IL2-R γ ,

- ladite séquence B comprend une séquence complémentaire de la séquence d'une deuxième région du gène codant l'IL2-R γ ,

ladite séquence A et la dite séquence B étant deux séquences différentes, lesdites première et deuxième régions du gène codant l'IL2-R γ , encadrant une région comprenant une séquence

5 codant une partie de l'IL2-R γ , ou codant le récepteur IL2-R γ complet, où

- ladite séquence A' comprend une séquence complémentaire de la séquence d'une troisième région du gène codant l'IL2-R γ ,

- ladite séquence B' comprend une séquence complémentaire de la séquence d'une quatrième 10 région du gène codant l'IL2-R γ ,

ladite séquence A' et la dite séquence B' étant deux séquences différentes, lesdites troisième et quatrième régions du gène codant l'IL2-R γ , encadrant une région comprenant une séquence complémentaire d'une séquence codant une partie de l'IL2-R γ , ou codant le récepteur IL2-R γ complet,

15 et où

la séquence A et la séquence A' sont au plus en partie complémentaires,

la séquence A et la séquence B' sont au plus en partie complémentaires,

la séquence B et la séquence A' sont au plus en partie complémentaires, et

la séquence B et la séquence B' sont au plus en partie complémentaires.

20 Par « au plus en partie complémentaires » on entend dans l'invention que les séquences possèdent des séquences qui sont capables de s'apparier selon la complémentarité des bases définie par Watson et Crick, mais pas sur la totalité des séquences. En d'autres termes, seule une partie des séquences est complémentaire. De manière avantageuses les séquences A, A', B et B' sont en partie complémentaire sur moins de 50% des séquences, notamment moins 25 de 30% des séquences, en particulier moins de 10% des séquences, notamment ne sont pas du tout complémentaires entre elles.

De cette manière il est possible d'encadrer le gène codant le récepteur IL2-R γ de manière spécifique et orientée, pour les deux brins, afin d'éviter des recombinaisons (i.e. des tagmentations) désordonnées ou incohérentes de sorte que le résultat ne serait pas celui 30 escompté.

De manière encore plus avantageuse, la composition susmentionnée comprend l'un au moins des sextuplets mentionnés dans le tableau 3 suivant :

[Table 3]

Sextuplet	Première molécule SEQ ID NO :	Deuxième molécule SEQ ID NO :	Troisième molécule SEQ ID NO :	Cinquième molécule SEQ ID NO :	Quatrième molécule SEQ ID NO :	Sixième molécule SEQ ID NO :
-----------	-------------------------------------	-------------------------------------	--------------------------------------	--------------------------------------	--------------------------------------	------------------------------------

#1	198	199	200	201	202	203
#2	204	205	206	207	208	209
#3	210	211	212	213	214	215
#4	216	217	218	219	220	221
#5	222	223	224	225	226	227
#6	228	229	230	231	232	233
#7	234	235	236	237	238	239
#8	240	241	242	243	244	245
#9	246	247	248	249	250	251
#10	252	253	254	255	256	257
#11	258	259	260	261	262	263
#12	264	265	266	267	268	269
#13	270	271	272	273	274	275
#14	276	277	278	279	280	281
#15	282	283	284	285	286	287
#16	288	289	290	291	292	293
#17	294	295	296	297	298	299
#18	300	301	302	303	304	305
#19	306	307	308	309	310	311
#20	312	313	314	315	316	317
#21	318	319	320	321	322	323
#22	324	325	326	327	328	329
#23	330	331	332	333	334	335
#24	336	337	338	339	340	341
#25	342	343	344	345	346	347
#26	348	349	350	351	352	353
#27	354	355	356	357	358	359
#28	360	361	362	363	364	365
#29	366	367	368	369	370	371
#30	372	373	374	375	376	377
#31	378	379	380	381	382	383
#32	384	385	386	387	388	389
#33	390	391	392	393	394	395
#34	396	397	398	399	400	401
#35	402	403	404	405	406	407
#36	408	409	410	411	412	413
#37	414	415	416	417	418	419
#38	420	421	422	423	424	425
#39	426	427	428	429	430	431
#40	432	433	434	435	436	437
#41	438	439	440	441	442	443
#42	444	445	446	447	448	449

#43	450	451	452	453	454	455
#44	456	457	458	459	460	461
#45	462	463	464	465	466	467
#46	468	469	470	471	472	473
#47	474	475	476	477	478	479
#48	480	481	482	483	484	485
#49	486	487	488	489	490	491
#50	492	493	494	495	496	497
#51	498	499	500	501	502	503
#52	504	505	506	507	508	509
#53	510	511	512	513	514	515
#54	516	517	518	519	520	521
#55	522	523	524	525	526	527
#56	528	529	530	531	532	533
#57	534	535	536	537	538	539
#58	540	541	542	543	544	545
#59	546	547	548	549	550	551
#60	552	553	554	555	556	557
#61	558	559	560	561	562	563
#62	564	565	566	567	568	569
#63	570	571	572	573	574	575
#64	576	577	578	579	580	581
#65	582	583	584	585	586	587
#66	588	589	590	591	592	593
#67	594	595	596	597	598	599
#68	600	601	602	603	604	605
#69	606	607	608	609	610	611
#70	612	613	614	615	616	617
#71	618	619	620	621	622	623
#72	624	625	626	627	628	629
#73	630	631	632	633	634	635
#74	636	637	638	639	640	641
#75	642	643	644	645	646	647
#76	648	649	650	651	652	653
#77	654	655	656	657	658	659
#78	660	661	662	663	664	665
#79	666	667	668	669	670	671
#80	672	673	674	675	676	677
#81	678	679	680	681	682	683
#82	684	685	686	687	688	689
#83	690	691	692	693	694	695
#84	696	697	698	699	700	701

#85	702	703	704	705	706	707
#86	708	709	710	711	712	713
#87	714	715	716	717	718	719
#88	720	721	722	723	724	725
#89	726	727	728	729	730	731
#90	732	733	734	735	736	737
#91	738	739	740	741	742	743
#92	744	745	746	747	748	749
#93	750	751	752	753	754	755
#94	756	757	758	759	760	761
#95	762	763	764	765	766	767
#96	768	769	770	771	772	773
#97	774	775	776	777	778	779
#98	780	781	782	783	784	785
#99	786	787	788	789	790	791
#100	792	793	794	795	796	797
#101	798	799	800	801	802	803
#102	804	805	806	807	808	809
#103	810	811	812	813	814	815
#104	816	817	818	819	820	821
#105	822	823	824	825	826	827
#106	828	829	830	831	832	833
#107	834	835	836	837	838	839
#108	840	841	842	843	844	845
#109	846	847	848	849	850	851
#110	852	853	854	855	856	857
#111	858	859	860	861	862	863
#112	864	865	866	867	868	869
#113	870	871	872	873	874	875
#114	876	877	878	879	880	881
#115	882	883	884	885	886	887
#116	888	889	890	891	892	893
#117	894	895	896	897	898	899
#118	900	901	902	903	904	905
#119	906	907	908	909	910	911
#120	912	913	914	915	916	917
#121	918	919	920	921	922	923
#122	924	925	926	927	928	929
#123	930	931	932	933	934	935
#124	936	937	938	939	940	941
#125	942	943	944	945	946	947
#126	948	949	950	951	952	953

#127	954	955	956	957	958	959
#128	960	961	962	963	964	965
#129	966	967	968	969	970	971
#130	972	973	974	975	976	977
#131	978	979	980	981	982	983
#132	984	985	986	987	988	989
#133	990	991	992	993	994	995
#134	996	997	998	999	1000	1001
#135	1002	1003	1004	1005	1006	1007
#136	1008	1009	1010	1011	1012	1013
#137	1014	1015	1016	1017	1018	1019
#138	1020	1021	1022	1023	1024	1025
#139	1026	1027	1028	1029	1030	1031
#140	1032	1033	1034	1035	1036	1037
#141	1038	1039	1040	1041	1042	1043
#142	1044	1045	1046	1047	1048	1049
#143	1050	1051	1052	1053	1054	1055
#144	1056	1057	1058	1059	1060	1061
#145	1062	1063	1064	1065	1066	1067
#146	1068	1069	1070	1071	1072	1073
#147	1074	1075	1076	1077	1078	1079
#148	1080	1081	1082	1083	1084	1085
#149	1086	1087	1088	1089	1090	1091
#150	1092	1093	1094	1095	1096	1097
#151	1098	1099	1100	1101	1102	1103
#152	1104	1105	1106	1107	1108	1109
#153	1110	1111	1112	1113	1114	1115
#154	1116	1117	1118	1119	1120	1121
#155	1122	1123	1124	1125	1126	1127
#156	1128	1129	1130	1131	1132	1133

La composition selon l'invention, propose ainsi de manière avantageuse, 156 sextuplets de molécules permettant le remplacement au locus du gène IL2R γ d'une séquence potentiellement mutée ou anormale par la séquence sauvage de référence.

Cela permet donc en fonction de la cellule cible de restaurer une fonction perdue par la mutation, et de réaliser une complémentation fonctionnelle.

5

Dans un autre aspect, l'invention concerne une composition pharmaceutique comprenant la composition susmentionnée, en association avec un véhicule pharmaceutiquement acceptable.

La composition susmentionnée, en association avec un véhicule pharmaceutiquement acceptable peut être utilisée dans le cadre du traitement de pathologies. Le véhicule est un véhicule communément admis et connu de l'homme du métier, par exemple de l'eau distillée,

10

ou encore un tampon physiologique. Ce véhicule doit permettre la formation d'un complexe comme mentionné ci-dessus, mais aussi être acceptable pour une cellule ou un être vivant. Dans un autre aspect, l'invention concerne la composition susmentionnée, pour son utilisation en tant que médicament.

5 Il est d'ailleurs avantageux que la composition comprenant l'un quelconque des triplets du tableau 2, ou l'un quelconque des sextuplets du tableau 3 soit utilisée en tant que médicament. Dans encore un autre aspect, l'invention concerne une composition pour son utilisation pour le traitement des pathologies liées à une mutation du gène codant l'IL2-R γ .

10 Les mutations du gène IL2R γ sont largement décrites dans la littérature, et peuvent correspondre à des substitutions, des insertions ou des délétions touchant soit un exon, un intron ou encore les régions régulatrices en 5' ou en 3' du gène.

Plus particulièrement, l'invention concerne une composition pour son utilisation pour le traitement des pathologies liées à une mutation du gène codant l'IL2-R γ , ladite composition comprenant un triplet tel que mentionné dans le tableau 2 ou un sextuplet tel que mentionné
15 dans le tableau 3.

Avantageusement, l'invention concerne la composition pour son utilisation susmentionnée, où la maladie associée à une mutation du gène codant l'IL2-R γ est la maladie DICS-X ou encore le syndrome d'Omenn.

20 Le déficit immunitaire combiné sévère (DICS) T-B+ par déficit en chaîne gamma, aussi appelé SCIDX1 ou DICS-X, est une forme de DICS caractérisée par des infections sévères et récurrentes associées à des diarrhées et un retard de croissance staturo-pondérale.

Le SCIDX1 se manifeste durant les premiers mois de vie par des infections virales, bactériennes ou fongiques sévères et souvent fatales (par ex pneumonitis à *Pneumocystis jiroveci*, infection à BCG disséminée secondaire à une vaccination) et un retard de croissance
25 staturo-pondérale. Des diarrhées chroniques sont une caractéristique fréquente. Certains patients présentent des éruptions cutanées et des anomalies de la fonction hépatique. La maladie du greffon contre l'hôte liée à la transmission materno-foetale est aussi associée à la maladie. Les résultats immunologiques révèlent une lymphopénie avec absence de lymphocytes T ou NK, une hypogammaglobulinémie, et un nombre de lymphocytes B normal
30 ou élevé.

Le syndrome d'Omenn est un trouble inflammatoire caractérisé par une érythrodermie, une desquamation, une alopécie, des diarrhées chroniques, un retard de croissance staturo-pondérale, adénopathies, et une hépatosplénomégalie, associés à un déficit immunitaire combiné sévère. Le syndrome d'Omenn apparaît au cours de la première année de vie avec
35 des manifestations caractéristiques du déficit immunitaire combiné sévère incluant des diarrhées chroniques, une pneumonite et un retard de croissance staturo-pondérale. Les

patients présentent aussi des symptômes inflammatoires tels qu'adénopathies, une hépatosplénomégalie et une érythrodermie généralisée, souvent à l'origine d'une alopecie avec chute des sourcils et des cils. La perte de protéines peut entraîner un oedème généralisé et des troubles métaboliques. Les signes et symptômes peuvent évoluer avec le temps et peuvent apparaître séparément. Certains patients manifestent seulement certains symptômes et peuvent alors être décrits comme des cas de Syndrome atypique d'Omenn. Plus qu'une forme de déficit immunitaire combiné sévère distincte, c'est davantage un phénotype inflammatoire qui peut être associé à différents types de déficit immunitaires combinés sévères. La plupart des cas rapportés à ce jour présente des mutations hypomorphes des gènes *RAG1* et *RAG2* (11p13). Les autres cas de mutations ont lieu dans les gènes *RMRP*, *ADA* et *IL2R γ* , ainsi que d'autres.

Dans un autre aspect, l'invention concerne une méthode de traitement du DICS-X ou encore du syndrome d'Omenn, ladite méthode comprenant l'administration d'une quantité efficace d'une composition susmentionnée, à un individu dans le besoin, ladite composition comprenant notamment un triplet tel que défini dans le tableau 2 ou un sextuplet tel que défini dans le tableau 3.

Dans encore un autre aspect, l'invention concerne l'utilisation d'une composition telle que définie ci-dessus, pour la substitution, dans une cellule somatique, d'une séquence mutée du gène codant l'IL2-R γ par une séquence sauvage du gène codant l'IL2-R γ sauvage, sous réserve que l'utilisation ne comprenne pas un procédé pour la modification de l'identité génétique germinale d'êtres humains et que ladite utilisation ne soit pas une méthode pour le traitement du corps humain ou animal par une chirurgie ou une thérapie.

Comme expliqué ci-dessus la composition permet de cibler spécifiquement une région cible et de la remplacer par une séquence d'intérêt, en utilisant les propriétés de tagmentation des transposases.

La composition selon l'invention est notamment avantageuse pour réaliser des modifications géniques ciblées (et en particulier des remplacements de gène, ou de séquences non codantes) chez l'homme et en particulier pour remplacer le gène muté IL2R γ par sa séquence sauvage de référence.

Dans encore un autre mode de réalisation avantageuse, l'invention concerne une méthode de remplacement, notamment in vitro, d'une séquence mutée du gène codant l'IL2-R γ par une par une séquence sauvage codant l'IL2-R γ , ladite méthode comprenant :

- la mise en contact d'une composition telle que définie précédemment, avec l'acide nucléique comprenant la séquence mutée codant le gène de l'IL2-R γ ,

ladite composition étant telle que

la séquence A de ladite première molécule comprend une séquence complémentaire de la région immédiatement en 5' de la séquence mutée codant le gène de l'IL2-R γ ,
la séquence B de ladite deuxième molécule comprend une séquence complémentaire de la région immédiatement en 3' de la séquence mutée codant le gène de l'IL2-R γ , et

- 5 la troisième molécule comprend la séquence sauvage du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, afin d'obtenir un complexe de remplacement,
- 10 - la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ledit ensemble, pour obtenir un complexe de recombinaison, et
- la recombinaison du complexe de combinaison pour obtenir la molécule d'acide nucléique hybride la séquence sauvage du gène codant l'IL2-R γ , à la place de la séquence mutée du gène codant l'IL2-R γ .
- 15

De manière avantageuse, l'invention concerne une méthode de remplacement, notamment in vitro, d'une séquence mutée du gène codant l'IL2-R γ par une séquence sauvage codant l'IL2-R γ , ladite méthode comprenant :

- la mise en contact d'une composition comprenant un triplet tel que défini dans le tableau 2 ou un sextuplet tel que défini dans le tableau 3, avec l'acide nucléique comprenant la séquence mutée codant le gène de l'IL2-R γ , afin d'obtenir un complexe de remplacement,
- 20 - la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ledit ensemble, pour obtenir un complexe de recombinaison, et
- 25 - la recombinaison du complexe de combinaison pour obtenir la molécule d'acide nucléique hybride la séquence sauvage du gène codant l'IL2-R γ , à la place de la séquence mutée du gène codant l'IL2-R γ .

Dans encore un autre aspect, l'invention concerne une méthode de remplacement in vitro d'une séquence double brin mutée du gène codant l'IL2-R γ d'une cellule souche, notamment une cellule souche hématopoïétique, par une séquence sauvage double brin codant l'IL2-R γ , ladite méthode comprenant :

- la mise en contact d'une composition telle que définie ci-dessus, avec la cellule souche hématopoïétique comprenant la séquence mutée codant le gène de l'IL2-R γ , ledit ensemble étant tel que
- 30 la séquence A de ladite première molécule comprend une séquence complémentaire de la région immédiatement en 5' de la séquence mutée codant le gène de l'IL2-R γ ,

la séquence A' de ladite quatrième molécule comprend une séquence complémentaire de la région immédiatement en 5' de la séquence complémentaire mutée codant le gène de l'IL2-R γ ,

5 la séquence B de ladite deuxième molécule comprend une séquence complémentaire de la région immédiatement en 3' de la séquence mutée codant le gène de l'IL2-R γ ,

la séquence B' de ladite cinquième molécule comprend une séquence complémentaire de la région immédiatement en 3' de la séquence complémentaire mutée codant le gène de l'IL2-R γ , et

10 la troisième molécule comprend la séquence sauvage du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule,

15 la sixième molécule comprend la séquence complémentaire de la séquence sauvage du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, afin d'obtenir un complexe de remplacement,

20 - la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ladite composition, pour obtenir un complexe de recombinaison, et

- la recombinaison du complexe de combinaison pour obtenir la molécule d'acide nucléique hybride la séquence sauvage du gène codant l'IL2-R γ , à la place de la séquence mutée du gène codant l'IL2-R γ ,

25 - et éventuellement la purification de la cellule souche hématopoïétique ayant un remplacement du gène muté.

De manière avantageuse, l'invention concerne une méthode de remplacement in vitro d'une séquence double brin mutée du gène codant l'IL2-R γ d'une cellule souche, notamment une cellule souche hématopoïétique, par une séquence sauvage double brin codant l'IL2-R γ , ladite méthode comprenant :

30 - la mise en contact d'une composition telle que définie ci-dessus, avec la cellule souche hématopoïétique comprenant la séquence mutée codant le gène de l'IL2-R γ , afin d'obtenir un complexe de remplacement, ladite composition comprenant un triplet tel que défini dans le tableau 2 ou un sextuplet tel que défini dans le tableau 3,

35 - la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ladite composition, pour obtenir un complexe de recombinaison, et

- la recombinaison du complexe de combinaison pour obtenir la molécule d'acide nucléique hybride la séquence sauvage du gène codant l'IL2-R γ , à la place de la séquence mutée du gène codant l'IL2-R γ ,

5 - et éventuellement la purification de la cellule souche hématopoïétique ayant un remplacement du gène muté.

Dans encore un autre aspect, l'invention concerne une méthode d'insertion, notamment in vitro, d'une séquence double brin du gène sauvage codant l'IL2-R γ dans une cellule, notamment une cellule souche, en particulier une cellule souche hématopoïétique, au locus du gène codant l'IL2-R γ , où ledit locus du gène codant l'IL2-R γ comprend une mutation

10 affectant l'expression, la fonction, ou les deux, de l'IL2-R γ ,

où ladite séquence double brin du gène sauvage codant l'IL2-R γ correspond à l'ADNc dudit gène, en particulier la séquence SEQ ID NO : 3,

ladite méthode comprenant :

15 - la mise en contact d'une composition telle que définie dans l'une quelconque des revendications 1 à 9, avec la cellule comprenant une mutation affectant l'expression, la fonction, ou les deux, de l'IL2-R γ , au niveau du locus du gène codant l'IL2-R γ

ladite composition étant telle que

la séquence A de ladite première molécule comprend une séquence complémentaire de la région en 5' de l'exon 1 du gène codant l'IL2-R γ ,

20 la séquence A' de ladite quatrième molécule comprend une séquence complémentaire de la séquence complémentaire de la région en 5' de l'exon 1 du gène codant l'IL2-R γ ,

la séquence B de ladite deuxième molécule comprend une séquence complémentaire en 3' la région en 5' de l'exon 1 du gène codant l'IL2-R γ ,

la séquence B' de ladite cinquième molécule comprend une séquence complémentaire en 3'

25 la séquence complémentaire de la région en 5' de l'exon 1 du gène codant l'IL2-R γ , et

la troisième molécule comprend la séquence sauvage de l'ADNc du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule,

30 la sixième molécule comprend la séquence complémentaire de la séquence sauvage de l'ADNc du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule,

35 afin d'obtenir un complexe de remplacement,

- la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ledit ensemble, pour obtenir un complexe de recombinaison, et
- la recombinaison du complexe de combinaison pour obtenir l'insertion de l'ADNc du gène codant l'IL2-R γ en 5' de l'exon 1 du gène codant l'IL2-R γ ,
- 5 - et éventuellement la purification de la cellule ayant l'insertion.

Brève description des figures

L'invention sera mieux comprise à la lecture des exemples ci-après et des figures suivantes :

- 10 [Fig. 1A] est une première représentation schématique d'une première forme d'appariement des première, deuxième et troisième molécules de la composition selon l'invention.
- [Fig. 1B] est une deuxième représentation schématique d'une première forme d'appariement des première, deuxième et troisième molécules de la composition selon l'invention.
- [Fig. 1C] est une troisième représentation schématique d'une première forme d'appariement
- 15 des première, deuxième et troisième molécules de la composition selon l'invention.
- [Fig. 1D] est une quatrième représentation schématique d'une première forme d'appariement des première, deuxième et troisième molécules de la composition selon l'invention.
- [Fig. 1E] est une cinquième représentation schématique d'une première forme d'appariement des première, deuxième et troisième molécules de la composition selon l'invention.
- 20 [Fig. 2] est un schéma montrant le principe et les étapes du remplacement simple brin en utilisant les première, deuxième et troisième molécules de la composition selon l'invention.
- [Fig. 3] est un schéma montrant le principe et les étapes du remplacement double brin en utilisant les première, deuxième, troisième, quatrième, cinquième et sixième molécules de la composition selon l'invention.
- 25 [Fig. 4] est un schéma montrant l'association des molécules contenues dans le tube 1 de l'exemple 1. La boule noire représente la biotine.
- [Fig. 5] est un schéma montrant l'association des molécules contenues dans le tube 2 de l'exemple 1. La boule noire représente la biotine.

Exemples

Exemple 1 – Obtention de cellules B exprimant le gène ILR2G sauvage

Afin de traiter les patients atteints de DICS-X, impliquant une mutation dans le gène ILR2G, il peut être avantageux de proposer une thérapie cellulaire visant à greffer chez des patients des cellules souches hématopoïétiques CD34+ ayant été modifiées à l'aide d'une composition selon l'invention, notamment pour remplacer le gène ILR2G muté par la séquence sauvage dudit gène.

A - Purification des cellules souches hématopoïétiques (HSC) CD34+ à partir de sang

La purification des cellules CD34+ est bien connue de l'homme du métier, et peut-être résumée comme suit :

1- Des échantillons de sang de donneurs (adultes ou du sang de cordon ombilical riche en cellules CD34+) sont collectés, frais ou congelés (il est également possible de purifier les cellules CD34+ d'un patient atteint de DISC-X) ;

2- Les échantillons sont centrifugés à 1000g pendant 8min sans frein sur milieu de centrifugation à gradient de densité stérile (par exemple Ficoll en présence d'EDTA) pour isoler les cellules mononucléaires du sang périphérique (en anglais PBMC pour *Peripheral Blood Mononuclear Cells*) ;

3- Les anneaux de globules blancs sont collectés et lavés au tampon phosphate PBS (en anglais pour *Phosphate Buffered Saline*) ;

4- Les cellules souches hématopoïétiques (i.e. les cellules CD34+) sont ensuite isolées au trieur cellulaire par sélection positive grâce à un marquage par anticorps anti-CD34 ou par sélection négative grâce à un cocktail d'anticorps anti-CD2, anti-CD3, anti-CD14, anti-CD16, anti-CD19, anti-CD24, anti-CD56n anti-CD66b et anti-CD61 ;

5- Les cellules CD34+ purifiées sont centrifugées à 300g pendant 5min et les culots cellulaires sont suspendus à une densité cellulaire de 100 000 à 300 000 cellules/mL et cultivées à 37°C et 5% de CO₂ pendant 48h en plaques 24 puits à raison de 1mL par puits dans un milieu de culture spécifiquement dédié à la culture des cellules souches et supplémenté avec un cocktail contenant de l'IL-3 humaine, de l'IL-6 humaine, de la thrombopoïétine recombinante, du hCSF (ligand Kit) recombinant et du ligand Flt-3 recombinant, à raison de 100ng/mL pour chaque.

6- les cellules souches CD34+ sont ainsi prêtes à être transfectées avec la composition selon l'invention.

B- Préparation des molécules de la composition selon l'invention

Afin de cibler le gène IL2R γ , et vérifier l'insertion au locus IL2R γ chrX:70,327,254-70,331,958 GRCh37/hg19, une construction comprenant la GFP a été préparée.

Deux tubes ont été préparés comme suit :

** Tube 1 (10 μ L) :

- l'oligo Great X1 (10µM) correspondant à la molécule 1 de séquence SEQ ID : 198 suivante
5'-

AgaTgTGtatAAGAGaCAGGTAGTGTATTTTTTTTTTTTTTTTTATCATCCTGtCTCTTataCAcAtcTTTTT
TTTTTTTTTTTTTTTTTTGGCGATCGCTTTTTTTTTTTTTTTTTGCTAGAAAGAGTACTGTTCTGGAAAC
5 TGACTTTTTTTTTTTTTTTTTGCGATCGCCTTTTTTTTTTTTTTGATACATTTAgaTgTGtatAAGAGaCAG
GATGAT-3'

- l'oligo Great X3 (10µM) correspondant à la molécule 4, de séquence SEQ ID : 201 suivante
5'-

CACGTGCTGtCTCTTataCAcAtcTTTTCTCGATCATTATTATTTTTGGCGATCGCTTTTTTTTTTTTTTT
10 TTAGCAAATTGGTAATACTCCTGCCTCCACAGTTTTTTTTTTTTTTTTGCGATCGCCTTTTTTTTTTTT
TTTTTTTTTTAgaTgTGtatAAGAGaCAGCACGTGTTTTTTTTTTTTTTTTTTTTTCAGCTATACTGtCTCT
TataCAcAtcT-3'

- l'oligo reverse GREAT X1 Rev (10µM) correspondant à la partie 5' de la molécule 3 de
séquence SEQ ID NO : 1134 suivante

15 5'-TACACTACCTGtCTCTTataCAcAtcTTTTTTTTTTTTTTTTTTTTTTTTTggttacATGCATCGTACA-3'

- l'oligo reverse GREAT X3 Rev (10µM) correspondant à la partie 3' de la molécule 6 de
séquence SEQ ID NO : 1135 suivante

5'-TGTACGATGCATgtaaccTTTTTTTTTTTTTTTTTTTTTTTTAgaTgTGtatAAGAGaCAGTATAGCTG-3'

20 - l'oligo Reverse Helix Rev correspondant à un brin additionnel permettant de lier une hélicase
par l'extrémité 3' biotinylée, l'oligo ayant la séquence SEQ ID NO : 1136 suivante

5'-ATAATAATGATCGAGAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT[Biot]-3'

** Tube 2 (10µL) :

- l'oligo Great X2 (10µM) correspondant à la molécule 2, de séquence SEQ ID NO : 199
suivante

25 5'-
CACGTGCTGtCTCTTataCAcAtcTTTTCTCGATCATTATTATTTTTGGCGATCGCTTTTTTTTTTTTTTT
TTGTTGAGAATGGTGCTAGTGGTAGTGAACAGTTTTTTTTTTTTTTTTGCGATCGCCTTTTTTTTTTTT
TTTTTTTTTTAgaTgTGtatAAGAGaCAGCACGTGTTTTTTTTTTTTTTTTTTTTTGACGAATACTGtCTCT
TataCAcAtcT-3'

30 - l'oligo Great X4 (10µM) correspondant à la molécule 5, de séquence SEQ ID NO : 202
suivante :

5'-
AgaTgTGtatAAGAGaCAGAGTATGAATTTTTTTTTTTTTTTTTATCATCCTGtCTCTTataCAcAtcTTTTT
TTTTTTTTTTTTTTTTTTGGCGATCGCTTTTTTTTTTTTTTTTTCTTCTCCTCTAAATCATTACCTTCTA
35 TAATTTTTTTTTTTTTTTTTGCGATCGCCTTTTTTTTTTTTTTGATACATTTAgaTgTGtatAAGAGaCAGGA
TGAT-3'

- l'oligo reverse GREAT X2 Rev (10µM) correspondant à la partie 3' de la molécule 3 de
séquence SEQ ID NO : 1137 suivante

5'-CGATACgcgcccatgttcTTTTTTTTTTTTTTTTTTTTTTTTAgaTgTGtatAAGAGaCAGTATTCGTC-3'

- l'oligo reverse GREAT X4 Rev (10 μ M) correspondant à la partie 5' de la molécule 6 SEQ ID NO : 1138 suivante

5'-TTCATACTCTGtCTCTTataCAcAtcTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTgaacatgCGGCCGCATATCG-3'

5 - l'oligo Reverse Helix Rev correspondant à un brin additionnel permettant de lier une hélicase par l'extrémité 3' biotinylée, l'oligo ayant la séquence SEQ ID NO : 1136.

Les deux tubes sont alors chauffés à 95°C pendant 5 min puis laissés 1h à température ambiante puis chaque tube est alors digéré avec les enzymes de restriction correspondantes : Tube 1 : Nsil et Tube 2 : NotI puis purifié par PCR purification kit (élution 20 μ L).

10 Les figures 4 et 5 représentent schématiquement le contenu des tubes 1 et 2 respectivement avant la digestion enzymatique.

En parallèle, une séquence amplifiée GFP-ILR2 γ a été obtenue via lyse de cellules HEK293T et protocole d'extraction d'ADN génomique et amplification du locus d'intérêt via PCR à Taq hyper processive. Cette séquence a été amplifiée (à l'aide d'oligonucléotides possédant en 5' un site de restriction Nsil et en 3' un site de restrictions NotI). Le produit de PCR a été digéré par les deux enzymes et purifié par kit PCR purification élué à un volume de 20 μ L. La fusion GFP-ILR2 γ est alors à bouts flanquants (i.e. demi-site Nsil et NotI libres) dans le tube 3.

20 Le contenu des 3 tubes (tube 1, tube 2 et tube 3) est alors mélangé et de la ligase de phage T4 est ajoutée au mélange (4 μ L, 100 U) en présence de tampon approprié (tampon T4 ; 5 μ L 10X) et 11 μ L d'eau distillée. Le mélange est laissé 1h à température ambiante. Puis le mélange est purifié sur gel (30 μ L d'élution) afin d'isoler le fragment le plus important, GFP-ILR2 γ , de taille supérieure à 1kbases.

25 Le fragment purifié est prêt à être utilisé et comprend les première, deuxième, quatrième et cinquième molécules aux extrémités, les troisième et sixième correspondant à la fusion GFP ILR2 γ (tube 4).

A ce stade, l'oligonucléotide simple brin modifié complémentaire d'une partie des première et quatrième molécules, la partie 3' de l'oligonucléotide simple brin modifié complémentaire étant couplée à de la biotine (oligo biotinylé), a été ajouté. Cette molécule biotinylée n'est pas obligatoire.

C- Transfection de cellules CD34+ avec la composition selon l'invention.

35 Le contenu du tube 4, contenant l'oligo biotinylé ou non, est alors utilisé pour transfecter par électroporation les cellules CD34+ isolées.

Les cellules sont centrifugées à 300g pendant 5min et lavées au PBS avant d'être suspendues dans du tampon électrolytique à une densité cellulaire de 75 000 à 100 000 cellules par chambre d'électroporation. Le contenu du tube 4, ainsi qu'un plasmide codant la protéine Tn5, et éventuellement un plasmide codant une fusion hélicase-streptavidine (UVRD-mSA) est
5 ajouté dans la chambre d'électroporation selon le protocole « *Episomal iPSC Reprogramming Vectors* » du système de transfection Neon® Transfert System proposé par Thermofisher.

Le mélange est électroporé à 1650V par trois séquences de 10 ms.

Les cellules sont ensuite immédiatement centrifugées à 300g pendant 5min et suspendues dans du milieu de culture pour cellules souches supplémenté pendant 48 à 96h à 37°C et 5%
10 de CO₂.

La tagmentation (i.e. le remplacement génique) est alors possible. Les cellules sont entretenues dans un milieu de culture approprié.

D- Vérification de la transfection

15 Les cellules transfectées ayant inséré au locus la molécule d'intérêt sont sélectionnées au moyen de la sélection par un antibiotique dont le gène de résistance est contenu dans la séquence de remplacement (par exemple la néomycine, la bléomycine, blasticidine S etc...). L'ADN génomique des cellules transfectées, et résistantes, est extrait puis fragmenté par sonication, puis on ajoute les adaptateurs de séquençage. Une amplification par PCR de tous
20 les fragments d'ADN génomiques est réalisée. Cette librairie a été séquencé sur séquenceur (50 × 8 × 50 *reads* sur un appareil HiSeq 2500, paired-end) et les résultats ont été compilé par cartographie et alignement via module bwa-mem (banque hg19). Analyse *on-target/off-target* ont été réalisé par sélection spécifique et *cluster read* via comparaison avec génome de référence (hg19) par cartographie par le module Bwa-mem en mode *short reads* et le module
25 minimap2. Les données de pic d'insertion ont quant à elles été assemblées via le module macs2.

Il est également possible de réaliser une simple PCR en utilisant un oligonucléotide sens en 5' de l'insertion et un oligonucléotide antisens en 3' de l'insertion (*Single-Tail Adapter/Tag (STAT)-PCR based method*). Le produit de PCR est alors séquencé selon les techniques
30 classiques, afin de détecter dans la région d'intérêt au moins la présence de la GFP, signe de de l'insertion de la fusion.

E- Analyse fonctionnelle des cellules CD34+ modifiées par la composition selon l'invention

35 La différenciation des cellules CD34+ modifiées est suivie in vitro grâce au modèle de coculture sur cellules stromales murines OP9-DL1 ou DL4 comme décrit dans la littérature.

Brièvement, les cellules OP9-DL1 ou DL4 sont cultivées avec les cellules CD34 modifiées pendant 14 jours à 37°C à 5% CO₂ et 10% O₂ en milieu MEM (Minimum Essential Medium) supplémenté avec 10% de SVF (Sérum de veau fœtal), le cocktail de cytokines nécessaire pour la culture des cellules CD34 (comme décrit ci-dessus) et un cocktail de cytokines favorisant l'activation des voies de différenciation myélo-érythroïde et lymphoïde.

Après 2 semaines de culture, les cellules CD34+ différenciées sont recueillies et analysées par cytométrie de flux pour observer leur profil de différenciation hématopoïétique grâce aux anticorps anti-CD19 (lignage lymphoïde B), anti-CD3 (lignage lymphoïde T, anti-CD56 (lignage NK), anti-CD16 (lignage monocytes/macrophages), anti-CD14 (lignage macrophage/neutrophiles, anti-CD11c (lignage myéloïde) et anti-CD235a (lignage érythroïde).

La différenciation des cellules CD34+ modifiées est également suivie in vivo grâce au modèle murin NSG ou NOD SCID (souris immunodéficientes).

Les cellules CD34+ modifiées sont injectées IH (intra hépatique) ou IF (intra fémorale) à des souris de 3-4 jours ou 6-8 semaines.

La différenciation est suivie à 8 semaines, 12 semaines et 16 semaines après injection, dans la moelle osseuse, la rate et le sang des souris. Les cellules humaines sont analysées par cytométrie de flux grâce aux anticorps anti-CD45 et anti-HLA ABC, et leur profil cellulaire est analysé grâce au même panel d'anticorps qu'utilisé in vitro.

Pour chaque type cellulaires l'expression de la GFP peut être détectée par cytométrie de flux. Sur les cellules différenciées, il peut également être réalisé une transcription inverse sur cellules isolées afin d'obtenir une banque d'ADNc. Une amplification de la séquence GFP et IL2R γ est alors réalisée par PCR Oligos pour la PCR et le séquençage IL2R γ : AGTGAACAGATCCTTCCCAGG (SEQ ID NO : 1139) et GFP Rev : CACGAACTCCAGCAGGACCATG (SEQ ID NO : 1140) et le fragment amplifié est séquençé en séquençage de Sanger mais aussi migré sur gel d'agarose. Comme attendu, après analyse (NCBI blast), on retrouve une similitude très proche avec le fragment GFP-IL2R γ théorique de remplacement de Locus IL2R γ .

Ces résultats montrent que la technologie selon l'invention permet de remplacer le gène IL2R γ endogène de manière efficace par une séquence exogène.

Bien évidemment, l'exemple donné ci-dessus a pour but de démontrer l'efficacité de la technologie, et utilise des gènes tels que la GFP pour faciliter la détection de l'insertion.

Dans le cadre d'une thérapie cellulaire, le gène codant la GFP n'est pas utilisé et le séquençage ou encore la détection fonctionnelle du récepteur sont alors utilisés pour vérifier l'efficacité.

Exemple 2 – Obtention de cellules B exprimant le gène ILR2G sauvage à partir de cellules souches pluripotentes.

Dans la mesure où il peut être difficile d'obtenir des cellules souches CD34+ chez le patient atteint de DICS-X, il est possible d'utiliser la technologie des cellules pluripotentes induites (iPSc).

Les cellules iPSc peuvent être issues de banques de cellules, dont le profil immunologique est compatible avec les patients.

Il est possible également de prélever des cellules différenciées chez le patient (par exemple des fibroblastes ou encore des cellules épithéliales), et de forcer la dédifférenciation par l'expression des gènes Oct4 et Socs2, ainsi que d'autres gènes tels que Klf4, Dub3, c-Myc, Nanog. De telles techniques sont désormais bien connues de l'homme du métier, et il pourra adapter les protocoles décrits dans l'art antérieur selon le type cellulaire qu'il souhaitera engager dans un modèle de dédifférenciation.

A ce stade il est possible de transfecter les cellules iPSc avec de la lipofectamine en présence du tube 4 comme décrit à l'exemple 1.

Les cellules transfectées sont sélectionnées par utilisation de l'antibiotique approprié, et l'insertion est vérifiée comme décrit à l'exemple 1.

Les cellules iPSc sont cultivées sur plaques enduites de gel de matrice cellulaire, en coculture avec des fibroblastes embryonnaires murins inactivés pendant 24h, en milieu de culture complet pour les cellules souches et supplémenté en bFGF et d'inhibiteur de ROCK.

Les cellules iPSc sont ensuite cultivées pendant 7 jours dans milieux de culture de 24h contenant des cocktails de différenciation distincts chaque jour :

- o J0-1 : milieu supplémenté en hBMP-4, hVEGF, hWnt3a et KOSR ;
- o J2 : milieu supplémenté en hBMP-4, hVEGF et KOSR ;
- o J3 : milieu supplémenté en hBMP-4, hVEGF et bFGF ;
- o J4-5 : milieu supplémenté en hVEGF et bFGF ;
- o J6 : milieu IMDM supplémenté en F12, B27, N2, BSA, hVEGF, bFGF, human stem cell factor et hFlt3 ligand ;
- o J7 : milieu IMDM supplémenté en F12, B27, N2, BSA, hVEGF, bFGF, human stem cell

factor, hFlt3 ligand, TPO, IL-6, hEPOgen, et FICZ (6-formylindolo [3,2-b]carbazole)

Au bout de 7 jours, les cellules sont ensuite maintenues en culture pendant 3 à 7 jours supplémentaires avant de poursuivre les expérimentations sur les cellules non adhérentes récoltées. Ces cellules sont testées pour vérifier qu'elles expriment le marqueur CD34+, par cytométrie de flux en utilisant un anticorps anti-CD34.

Si les cellules n'ont pas été transfectées au stade iPSc, elles peuvent dès lors être transfectées au stade CD34+ comme décrit à l'exemple 1. Leur différenciation est alors induite comme décrite à l'exemple 1.

Revendications

1. Composition comprenant

- 5 - une première molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence A permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence complémentaire étant liée en 5' à une première séquence riches en T de 40 à 60 nucléotides de long et en 3' à une deuxième séquence riches en T de 40 à 60 nucléotides de long, lesdites première et deuxième séquences riches en T comprenant respectivement un premier et un deuxième domaine de 6 à 12
- 10 nucléotides riches en G/C, la séquence du premier domaine étant complémentaire de la séquence du deuxième domaine, ledit premier et deuxième domaine étant positionnés de 15 à 52 nucléotides de ladite séquence A, ladite première molécule comprenant en son extrémité 5' au moins une première séquence orientés 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite
- 15 transposase,
- une deuxième molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence B permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence B complémentaire étant liée en 5' à une troisième
- 20 séquence riche en T de 40 à 60 nucléotides de long et en 3' à une quatrième séquence riche en T de 40 à 60 nucléotides de long, lesdites troisième et quatrième séquences riches en T comprenant respectivement un troisième et un quatrième domaine de 6 à 12 nucléotides riches en G/C, la séquence du troisième domaine étant complémentaire de la séquence du quatrième domaine, lesdits troisième et quatrième domaines étant
- 25 positionnés de 15 à 52 nucléotides de ladite séquence B, ladite deuxième molécule comprenant en son extrémité 5' au moins la première séquence orientée 5'-3' de reconnaissance de ladite transposase et en son extrémité 3' la deuxième séquence de reconnaissance de ladite transposase,
- ladite séquence B étant une séquence complémentaire dudit acide nucléique d'intérêt, la
- 30 séquence A étant positionnée en 5' d'une région d'intérêt dudit acide nucléique d'intérêt et la séquence B positionnée en 3' de la région d'intérêt dudit acide nucléique d'intérêt, et
- une troisième molécule simple brin comprenant
- * dans sa partie 5', au moins une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule,
- 35 * dans sa partie 3' au moins une séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, et

* une région intermédiaire située entre séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, ladite région intermédiaire comprenant une séquence codant une partie du récepteur gamma de l'interleukine 2 ou IL2-R γ , ou codant le récepteur IL2-R γ complet,

les première et troisième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase et les deuxième et troisième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase.

2. Composition selon la revendication 1, où ladite séquence A comprend une séquence complémentaire de la séquence d'une première région du gène codant l'IL2-R γ , et où ladite séquence B comprend une séquence complémentaire de la séquence d'une deuxième région du gène codant l'IL2-R γ , ladite séquence A et la dite séquence B étant deux séquences différentes, lesdites première et deuxième régions du gène codant l'IL2-R γ , encadrant une région comprenant une séquence codant une partie de l'IL2-R γ , ou codant le récepteur IL2-R γ complet.

3. Composition selon l'une quelconque des revendications 1 à 2, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence orientée 5'-3' de reconnaissance de ladite transposase et où la troisième molécule en comprend en son extrémité 5' une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la première molécule suivie d'une deuxième séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule.

4. Composition selon l'une quelconque des revendications 1 à 3, où ladite première molécule comprend en son extrémité 5' une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite transposase, suivie d'une première séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase et

où la troisième molécule comprend en son extrémité 5' une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase.

5. Composition selon l'une quelconque des revendications 1 à 4, où ladite transposase est une transposase bactérienne, notamment une transposase choisie parmi Tn5, Tn9, Tn10 ou Tc1/mariner.
5
6. Composition selon l'une quelconque des revendications 1 à 5, où le gène codant l'IL2-R γ comprend une séquence choisie parmi SEQ ID NO : 1, SEQ ID NO : 2 et SEQ ID NO : 3.
10
7. Composition selon l'une quelconque des revendications 1 à 6, où les première, deuxième et troisième molécules sont choisies parmi les triplets tels que définis dans le tableau 2.
8. Composition selon l'une quelconque des revendications 1 à 7, comprenant en outre
15
** une quatrième molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence A' permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence complémentaire étant liée en 5' à une cinquième séquence riche en T de 40 à 60 nucléotides de long et en 3' à une sixième séquence riche en T de 40 à 60 nucléotides de long, lesdites cinquième et sixième séquences riches en T comprenant respectivement un cinquième et un sixième domaine de 6 à 12 nucléotides riches en G/C, la séquence du cinquième domaine étant complémentaire de la séquence du sixième domaine, lesdits cinquième et sixième domaine étant positionnés de 15 à 52 nucléotides de ladite séquence A', ladite quatrième molécule comprenant en son extrémité
20
5' au moins une première séquence orientée 5'-3' de reconnaissance d'une transposase et en son extrémité 3' une deuxième séquence de reconnaissance de ladite transposase,
25
** une cinquième molécule d'acides nucléiques simple brin comprenant ou constituée essentiellement d'une séquence B' permettant l'insertion d'une séquence complémentaire d'un acide nucléique d'intérêt, ou comprenant une séquence complémentaire d'un acide nucléique d'intérêt, ladite séquence B' complémentaire étant liée en 5' à une septième séquence riche en T de 40 à 60 nucléotides de long et en 3' à une huitième séquence riche en T de 40 à 60 nucléotides de long, ladite septième et huitième séquence riche en T comprenant respectivement un septième et un huitième domaine de 6 à 12 nucléotides riches en G-C, la séquence du septième domaine étant complémentaire de la séquence du huitième domaine, lesdits septième et huitième domaines étant positionnés de 15 à 52 nucléotides de ladite séquence B', ladite cinquième molécule comprenant en son extrémité
30
35

5' au moins la première séquence orientée 5'-3' de reconnaissance de ladite transposase et en son extrémité 3' la deuxième séquence de reconnaissance de ladite transposase, ladite séquence B' étant une séquence complémentaire dudit acide nucléique d'intérêt, la séquence A' étant positionnée en 5' d'une région d'intérêt dudit acide nucléique d'intérêt et la séquence B' positionnée en 3' de la région d'intérêt dudit acide nucléique d'intérêt, et - une sixième molécule simple brin comprenant

* dans sa partie 5', au moins une séquence complémentaire de ladite cinquième séquence de reconnaissance de ladite transposase de la première molécule,

* dans sa partie 3' au moins une séquence complémentaire de ladite quatrième séquence de reconnaissance de ladite transposase de la deuxième molécule, et

* une région intermédiaire située entre séquence complémentaire de ladite quatrième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite cinquième séquence de reconnaissance de ladite transposase de la deuxième molécule, ladite région intermédiaire comprenant une séquence complémentaire, et antiparallèle, de la séquence codant le récepteur gamma de l'interleukine 2 ou IL2-R γ contenue dans la troisième molécule de la première composition, les quatrième et sixième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase et les cinquième et sixième molécules d'acides nucléiques simple brin étant appariées selon la complémentarité des bases définie par Watson et Crick de sorte à définir deux sites double brin de liaison de ladite transposase.

9. Composition selon la revendication 8, où

- ladite séquence A comprend une séquence complémentaire de la séquence d'une première région du gène codant l'IL2-R γ ,

- ladite séquence B comprend une séquence complémentaire de la séquence d'une deuxième région du gène codant l'IL2-R γ ,

ladite séquence A et la dite séquence B étant deux séquences différentes, lesdites première et deuxième régions du gène codant l'IL2-R γ , encadrant une région comprenant une séquence codant une partie de l'IL2-R γ , ou codant le récepteur IL2-R γ complet,

où

- ladite séquence A' comprend une séquence complémentaire de la séquence d'une troisième région du gène codant l'IL2-R γ ,

- ladite séquence B' comprend une séquence complémentaire de la séquence d'une quatrième région du gène codant l'IL2-R γ ,

ladite séquence A' et la dite séquence B' étant deux séquences différentes, lesdites troisième et quatrième régions du gène codant l'IL2-R γ , encadrant une région comprenant une séquence complémentaire d'une séquence codant une partie de l'IL2-R γ , ou codant le récepteur IL2-R γ complet,

5 et où

la séquence A et la séquence A' sont au plus en partie complémentaires,
la séquence A et la séquence B' sont au plus en partie complémentaires,
la séquence B et la séquence A' sont au plus en partie complémentaires, et
la séquence B et la séquence B' sont au plus en partie complémentaires.

10

10. Composition pharmaceutique comprenant une composition selon l'une quelconque des revendications 1 à 9, en association avec un véhicule pharmaceutiquement acceptable.

15

11. Composition selon l'une quelconque des revendications 1 à 9, pour son utilisation en tant que médicament.

12. Composition selon l'une quelconque des revendications 1 à 9, pour son utilisation pour le traitement d'une maladie associée à une mutation du gène codant l'IL2-R γ .

20

13. Composition pour son utilisation selon la revendication 12, où la maladie associée à une mutation du gène codant l'IL2-R γ est la maladie DICS-X ou encore le syndrome d'Omenn.

25

14. Utilisation d'une composition selon l'une quelconque des revendications 1 à 9, pour la substitution, dans une cellule somatique, d'une séquence mutée du gène codant l'IL2-R γ par une séquence sauvage du gène codant l'IL2-R γ sauvage, sous réserve que l'utilisation ne comprenne pas un procédé pour la modification de l'identité génétique germinale d'êtres humains et que ladite utilisation ne soit pas une méthode pour le traitement du corps humain ou animal par une chirurgie ou une thérapie.

30

15. Méthode de remplacement in vitro d'une séquence mutée du gène codant l'IL2-R γ par une par une séquence sauvage codant l'IL2-R γ , ladite méthode comprenant :

- la mise en contact d'une composition telle que défini dans l'une quelconque des revendications 1 à 9, avec l'acide nucléique comprenant la séquence mutée codant le gène de l'IL2-R γ ,

35

ladite composition étant telle que

la séquence A de ladite première molécule comprend une séquence complémentaire de la région immédiatement en 5' de la séquence mutée codant le gène de l'IL2-R γ ,

la séquence B de ladite deuxième molécule comprend une séquence complémentaire de la région immédiatement en 3' de la séquence mutée codant le gène de l'IL2-R γ , et

5 la troisième molécule comprend la séquence sauvage du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, afin d'obtenir un complexe de remplacement,

10 - la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ledit ensemble, pour obtenir un complexe de recombinaison, et

- la recombinaison du complexe de combinaison pour obtenir la molécule d'acide nucléique hybride la séquence sauvage du gène codant l'IL2-R γ , à la place de la séquence mutée
15 du gène codant l'IL2-R γ .

16. Méthode de remplacement in vitro d'une séquence double brin mutée du gène codant l'IL2-R γ d'une cellule souche, notamment une cellule souche hématopoïétique, par une séquence sauvage double brin codant l'IL2-R γ , ladite méthode comprenant :

20 - la mise en contact d'une composition telle que définie dans l'une quelconque des revendications 1 à 9, avec la cellule souche hématopoïétique comprenant la séquence mutée codant le gène de l'IL2-R γ ,

ledit ensemble étant tel que

25 la séquence A de ladite première molécule comprend une séquence complémentaire de la région immédiatement en 5' de la séquence mutée codant le gène de l'IL2-R γ ,

la séquence A' de ladite quatrième molécule comprend une séquence complémentaire de la région immédiatement en 5' de la séquence complémentaire mutée codant le gène de l'IL2-R γ ,

30 la séquence B de ladite deuxième molécule comprend une séquence complémentaire de la région immédiatement en 3' de la séquence mutée codant le gène de l'IL2-R γ ,

la séquence B' de ladite cinquième molécule comprend une séquence complémentaire de la région immédiatement en 3' de la séquence complémentaire mutée codant le gène de l'IL2-R γ , et

35 la troisième molécule comprend la séquence sauvage du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite

transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule,

la sixième molécule comprend la séquence complémentaire de la séquence sauvage du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la

5

séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule, afin d'obtenir un complexe de remplacement,

- la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ledit ensemble, pour obtenir un complexe de recombinaison, et

10

- la recombinaison du complexe de combinaison pour obtenir la molécule d'acide nucléique hybride la séquence sauvage du gène codant l'IL2-R γ , à la place de la séquence mutée du gène codant l'IL2-R γ ,

15

- et éventuellement la purification de la cellules souche hématopoïétique ayant un remplacement du gène muté.

17. Méthode d'insertion in vitro d'une séquence double brin du gène sauvage codant l'IL2-R γ dans une cellule, notamment une cellule souche, en particulier une cellule souche

20

hématopoïétique, au locus du gène codant l'IL2-R γ , où ledit locus du gène codant l'IL2-R γ comprend une mutation affectant l'expression, la fonction, ou les deux, de l'IL2-R γ , où ladite séquence double brin du gène sauvage codant l'IL2-R γ correspond au ADNc dudit gène,

ladite méthode comprenant :

- la mise en contact d'une composition telle que définie précédemment, avec la cellule comprenant une mutation affectant l'expression, la fonction, ou les deux, de l'IL2-R γ , au niveau du locus du gène codant l'IL2-R γ

25

ladite composition étant telle que

la séquence A de ladite première molécule comprend une séquence complémentaire de la région en 5' de l'exon 1 du gène codant l'IL2-R γ ,

30

la séquence A' de ladite quatrième molécule comprend une séquence complémentaire de la séquence complémentaire de la région en 5' de l'exon 1 du gène codant l'IL2-R γ ,

la séquence B de ladite deuxième molécule comprend une séquence complémentaire en 3' la région en 5' de l'exon 1 du gène codant l'IL2-R γ ,

35

la séquence B' de ladite cinquième molécule comprend une séquence complémentaire en 3' la séquence complémentaire de la région en 5' de l'exon 1 du gène codant l'IL2-R γ , et

la troisième molécule comprend la séquence sauvage de l'ADNc du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule,

5

la sixième molécule comprend la séquence complémentaire de la séquence sauvage de l'ADNc du gène codant l'IL2-R γ , située entre une séquence complémentaire de ladite deuxième séquence de reconnaissance de ladite transposase de la première molécule et la séquence complémentaire de ladite première séquence de reconnaissance de ladite transposase de la deuxième molécule,

10

afin d'obtenir un complexe de remplacement,

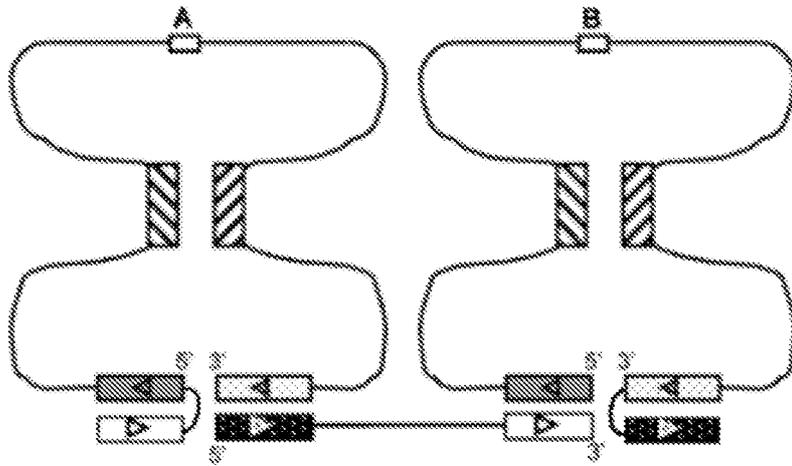
- la mise en présence du complexe de remplacement avec une transposase reconnaissant les sites double brin de liaison de ladite transposase contenus dans ledit ensemble, pour obtenir un complexe de recombinaison, et

15

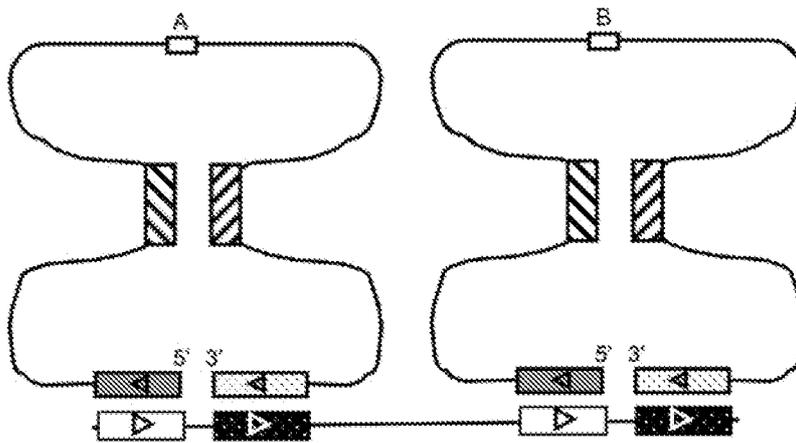
- la recombinaison du complexe de combinaison pour obtenir l'insertion de l'ADNc du gène codant l'IL2-R γ en 5' de l'exon 1 du gène codant l'IL2-R γ ,

- et éventuellement la purification de la cellule ayant l'insertion.

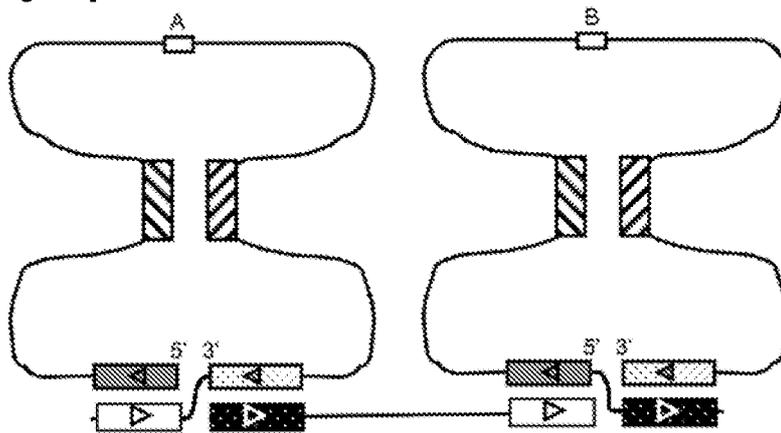
[Fig. 1A]



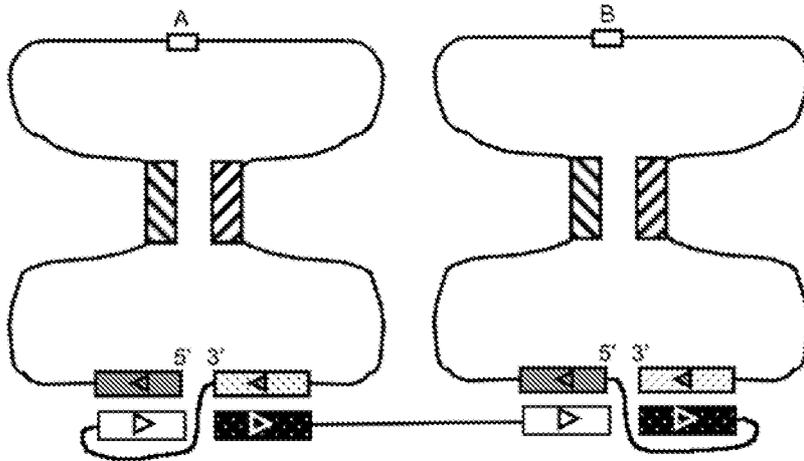
[Fig. 1B]



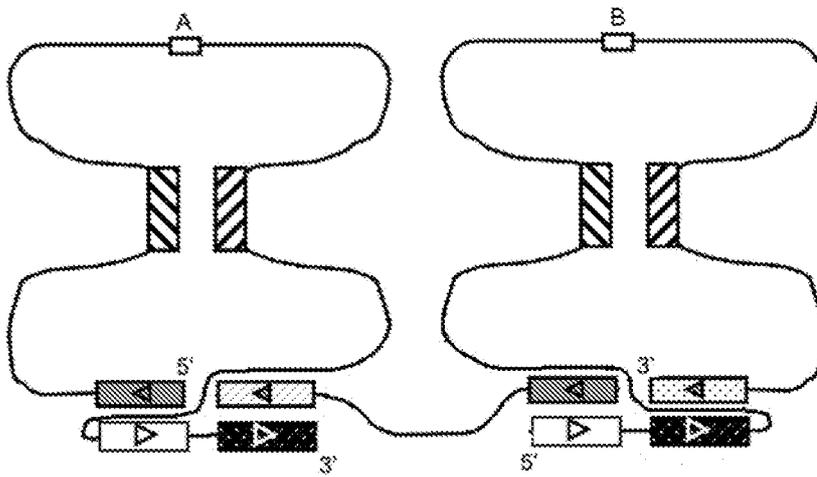
5 [Fig. 1C]



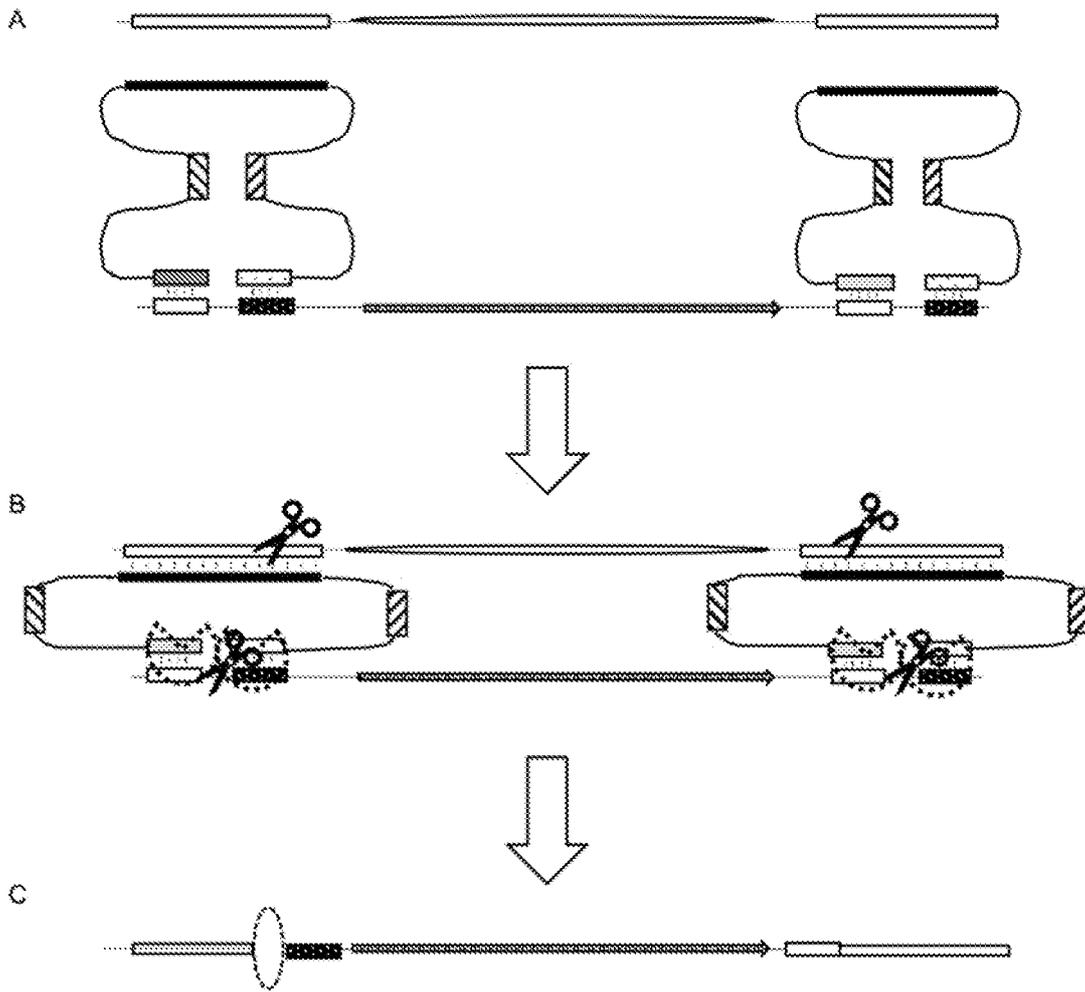
[Fig. 1D]



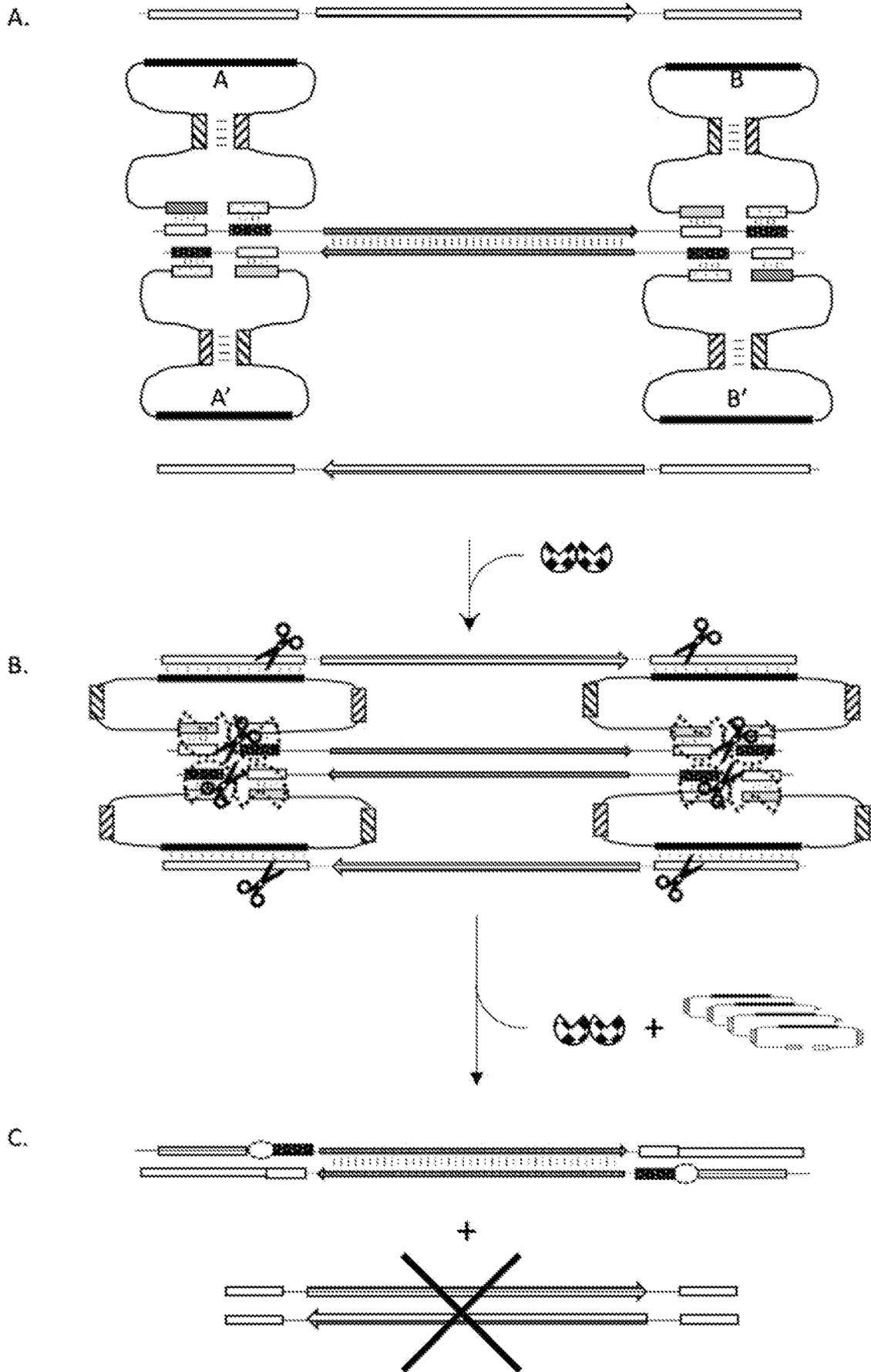
[Fig. 1E]



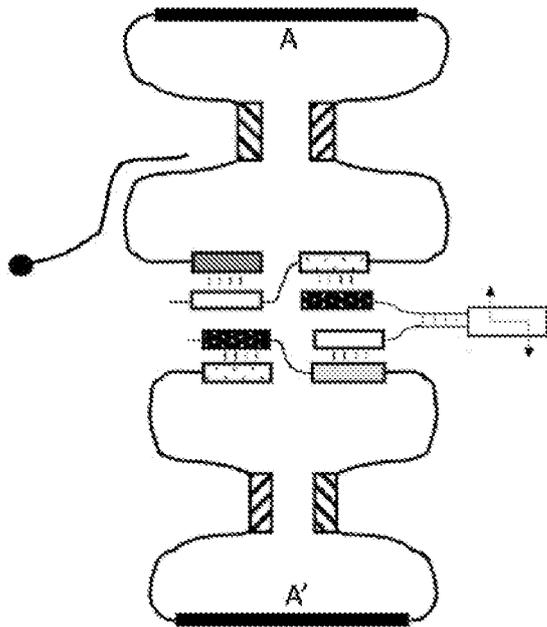
[Fig. 2]



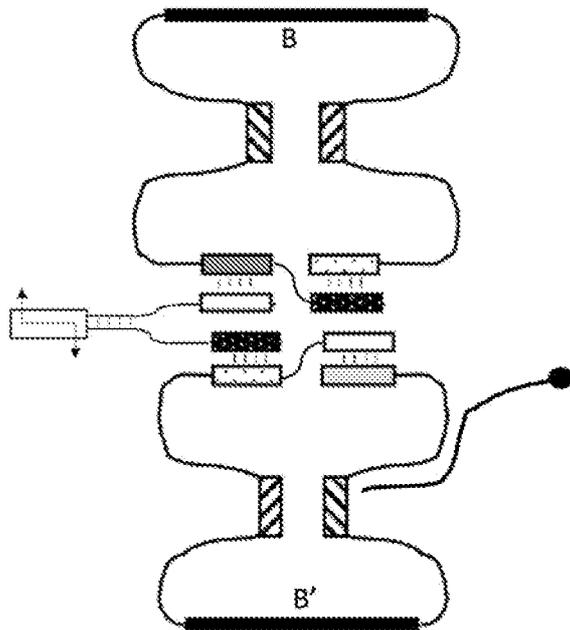
[Fig. 3]



[Fig. 4]



[Fig. 5]



TRAITE DE COOPERATION EN MATIERE DE BREVETS

RAPPORT DE RECHERCHE DE TYPE INTERNATIONAL ÉTABLI EN VERTU DE L'ARTICLE XI.23., §10 DU CODE DE DROIT ÉCONOMIQUE BELGE

IDENTIFICATION DE LA DEMANDE INTERNATIONALE	REFERENCE DU DEPOSANT OU DU MANDATAIRE BR 141768/RIM/EVG
Demande nationale belge n° 202206100	Date du dépôt 28-12-2022
	Date de priorité revendiquée
Déposant (Nom) QUIDDITAS SA	
Date de la requête d'une recherche de type international 14-01-2023	Numéro attribué par l'administration chargée de la recherche internationale à la requête d'une recherche de type international SN83010
I. CLASSEMENT DE L'OBJET DE LA DEMANDE (en cas de plusieurs symboles de la classification, les indiquer tous)	
Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB Voir rapport de recherche	
II. DOMAINES RECHERCHES	
Documentation minimale consultée	
Système de classification	Symboles de la classification
IPC	Voir rapport de recherche
Documentation consultée autre que la documentation minimale dans la mesure où ces documents font partie des domaines consultés	
III. <input type="checkbox"/> IL A ÉTÉ ESTIMÉ QUE CERTAINES REVENDICATIONS NE POUVAIENT FAIRE L'OBJET D'UNE RECHERCHE (Observations sur la feuille supplémentaire)	
IV. <input type="checkbox"/> ABSENCE D'UNITÉ DE L'INVENTION ET/OU CONSTATATION RELATIVE À L'ÉTENDUE DE LA RECHERCHE (Observations sur la feuille supplémentaire)	

RAPPORT DE RECHERCHE DE TYPE INTERNATIONAL

Demande de recherche No

BE 202206100
A. CLASSEMENT DE L'OBJET DE LA DEMANDE
INV. C12N15/11 C12N15/63 C12N15/85 C12N15/90
ADD.

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

 Documentation minimale consultée (système de classification suivi des symboles de classement)
C12N C40B

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si réalisable, termes de recherche utilisés)

EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data
C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie °	Documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
A	WO 2021/257997 A2 (BROAD INST INC [US]; MASSACHUSETTS INST TECHNOLOGY [US] ET AL.) 23 décembre 2021 (2021-12-23) * le document en entier * -----	1-17
A	WO 2021/146666 A2 (ARIDIS PHARMACEUTICALS INC [US]; NGO KATHY ET AL.) 22 juillet 2021 (2021-07-22) * le document en entier * -----	1-17

 Voir la suite du cadre C pour la fin de la liste des documents

 Les documents de familles de brevets sont indiqués en annexe

° Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche de type international a été effectivement achevée

28 septembre 2023

Date d'expédition du rapport de recherche de type international

 Nom et adresse postale de l'administration chargée de la recherche internationale
 Office Européen des Brevets, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040,
 Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Chavanne, Franz

RAPPORT DE RECHERCHE DE TYPE INTERNATIONAL

Renseignements relatifs aux membres de familles de brevets

Demande de recherche n

BE 202206100

Document brevet cité au rapport de recherche	Date de publication	Membre(s) de la famille de brevet(s)	Date de publication
WO 2021257997	A2	23-12-2021	AU 2021293587 A1 15-12-2022
			CA 3178165 A1 23-12-2021
			CN 116096880 A 09-05-2023
			EP 4168540 A2 26-04-2023
			US 2023265420 A1 24-08-2023
			WO 2021257997 A2 23-12-2021

WO 2021146666	A2	22-07-2021	CA 3161872 A1 22-07-2021
			CN 115279910 A 01-11-2022
			EP 4090754 A2 23-11-2022
			WO 2021146666 A2 22-07-2021



OPINION ÉCRITE

Dossier N° SN83010	Date du dépôt(<i>jour/mois/année</i>) 28.12.2022	Date de priorité (<i>jour/mois/année</i>)	Demande n° BE202206100
Classification internationale des brevets (CIB) INV. C12N15/11 C12N15/63 C12N15/85 C12N15/90			
Déposant QUIDDITAS SA			

La présente opinion contient des indications et les pages correspondantes relatives aux points suivants :

- Cadre n° I Base de l'opinion
- Cadre n° II Priorité
- Cadre n° III Absence de formulation d'opinion quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle
- Cadre n° IV Absence d'unité de l'invention
- Cadre n° V Déclaration motivée quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration
- Cadre n° VI Certains documents cités
- Cadre n° VII Irrégularités dans la demande
- Cadre n° VIII Observations relatives à la demande

Formulaire BE237A (feuille de couverture) (Juillet 2022)	Examineur Chavanne, Franz
--	------------------------------

Cadre n° I Base de l'opinion

1. Cette opinion a été établie sur la base des revendications déposées avant le commencement de la recherche.
2. En ce qui concerne **la ou les séquences de nucléotides ou d'acides aminés** divulguées dans la demande, la présente opinion a été effectuée sur la base d'un listage des séquences
 - a. faisant partie de la demande telle que déposée.
 - b. remis postérieurement à la date du dépôt aux fins de la recherche,
 - accompagné d'une déclaration selon laquelle le listage des séquences ne va pas au-delà de la divulgation faite dans la demande telle que déposée.
3. En ce qui concerne la ou les séquences de nucléotides ou d'acides aminés divulguées dans la demande, la présente opinion a été effectuée dans la mesure où une opinion valable pouvait être formulée en l'absence d'un listage des séquences conforme à la norme ST.26 de l'OMPI.
4. Commentaires complémentaires :

Cadre n° V Opinion motivée quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle; citations et explications à l'appui de cette déclaration

1. Déclaration

Nouveauté	Oui : Revendications	1-17
	Non : Revendications	
Activité inventive	Oui : Revendications	1-17
	Non : Revendications	
Possibilité d'application industrielle	Oui : Revendications	1-17
	Non : Revendications	

2. Citations et explications

voir feuille séparée

Ad point V

Déclaration motivée quant à la nouveauté, l'activité inventive et la possibilité d'application industrielle ; citations et explications à l'appui de cette déclaration

1 Il est fait référence aux documents suivants :

- D1 WO 2021/257997 A2 (BROAD INST INC [US]; MASSACHUSETTS INST TECHNOLOGY [US] ET AL.) 23 décembre 2021 (2021-12-23)
- D2 WO 2021/146666 A2 (ARIDIS PHARMACEUTICALS INC [US]; NGO KATHY ET AL.) 22 juillet 2021 (2021-07-22)

2 L'objet revendiqué n'est pas décrit dans l'état de la technique. Cet objet est donc nouveau.

3 Le document de l'état de la technique le plus proche pour évaluer l'inventivité de la présente demande est D1.

D1 décrit une composition pour le remplacement d'une séquence d'acide nucléique par une séquence nucléique d'intérêt par la provision de site de recombinaison et une transposase par le system CRISP-Cas. D1 mentionne le gène codant pour le récepteur gamma de l'interleukine 2 (IL2-R γ) comme cible dans le cadre d'un traitement thérapeutique.

L'objet revendiqué diffère de D1 par le system de recombinaison mettant en oeuvre une transposase.

Le problème à résoudre est donc de fournir une autre composition pour le remplacement de la séquence codant pour tout ou partie du récepteur gamma de l'interleukine 2 par recombinaison homologue et action d'une transposase.

La présente demande résout ce problème en proposant une composition comprenant 3 acides nucléiques simples brins comprenant des séquences complémentaires pour l'insertion de la séquence codant pour tout ou partie du récepteur gamma de l'interleukine 2 sauvage, permettant ainsi de restaurer substituer une séquence mutée du gène codant l'IL2-R γ par la séquence sauvage correspondante du gène codant l'IL2-R γ , dans le cadre du traitement de la maladie DICS-X ou syndrome d'Omenn.

L'art antérieur ne décrit pas, ni ne suggère, la provision de trois acides nucléiques simples brins comprenant des séquences complémentaires et permettant le remplacement de tout ou partie du gène codant l'IL2-R γ par recombinaison homologue et action d'une transposase. La composition revendiquée ne peut être déduite de D1 ni des enseignements de l'état de la technique. L'objet revendiqué est donc inventif.