

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

(19) Organisation Mondiale de la  
Propriété Intellectuelle  
Bureau international



(10) Numéro de publication internationale  
**WO 2013/093324 A2**

(43) Date de la publication internationale  
27 juin 2013 (27.06.2013)

W I P O I P C T

- (51) Classification internationale des brevets :  
C12Q 1/68 (2006.01)
- (21) Numéro de la demande internationale :  
PCT/FR20 12/052970
- (22) Date de dépôt international :  
18 décembre 2012 (18.12.2012)
- (25) Langue de dépôt : français
- (26) Langue de publication : français
- (30) Données relatives à la priorité :  
1162027 20 décembre 2011 (20.12.2011) FR
- (71) Déposants : BIOMÉRIEUX [FR/FR]; F-69280 Marcy L'étoile (FR). HOSPICES CIVILS DE LYON [FR/FR]; 3 quai des Célestins, F-69002 Lyon (FR).
- (72) Inventeurs : PEROT, Philippe; 8 rue Jean Moulin, F-16710 saint-Yrieix (FR). MALLET, François; 36 cours de la République, F-69100 Villeurbanne (FR). MUGNIER, Nathalie; 20 rue Paul Bert, F-69003 Lyon (FR).
- (74) Mandataire : DORGET, Elisabeth; bioMérieux, Département Propriété Industrielle, Chemin de l'Orme, F-69280 Marcy L'étoile (FR).
- (81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,

DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), eurasiatique (AM, AZ, BY, KG, KZ, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Déclarations en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv))

Publiée :

- sans rapport de recherche internationale, sera republiée dès réception de ce rapport (règle 48.2.g))
- avec la partie de la description réservée au listage des séquences (règle 5.2.a))

(54) Title : METHOD FOR THE DIAGNOSIS OR PROGNOSIS, *IN VITRO*, OF PROSTATE CANCER

(54) Titre : PROCÉDE POUR LE DIAGNOSTIC OU LE PRONOSTIC, *IN VITRO*, DU CANCER DE LA PROSTATE

(57) Abstract : The subject matter of the present invention is a method for the diagnosis or prognosis, *in vitro*, of prostate cancer, which comprises a step of detecting at least one expression product of at least one HERV nucleic acid sequence, the use of said nucleic acid sequences, which have been isolated, as a molecular marker or molecular markers, and a kit comprising at least one binding partner specific for at least one of the expression products of the HERV nucleic acid sequences.

(57) Abrégé : La présente invention a pour objet un procédé pour le diagnostic ou le pronostic, *in vitro*, du cancer de la prostate, qui comprend une étape de détection d'au moins un produit d'expression d'au moins une séquence d'acide nucléique HERV, l'utilisation desdites séquences d'acide nucléique, isolées, comme marqueur(s) moléculaire(s) et un kit comprenant au moins un partenaire de liaison spécifique d'au moins un des produits d'expression des séquences d'acide nucléique HERV.



WO 2013/093324 A2

**Procédé pour le diagnostic ou le pronostic, *in vitro*, du cancer  
de la prostate**

Les rétrovirus endogènes constituent la descendance de rétrovirus infectieux s'étant intégrés sous leur forme provirale dans des cellules de la lignée germinale et ayant été transmis par ce biais dans le génome de la descendance de l'hôte.

Le séquençage du génome humain a permis de révéler l'extrême abondance des éléments transposables ou de leurs dérivés. De fait, les séquences répétées représentent près de la moitié du génome humain et les rétrovirus endogènes et les rétrotransposons en composent 8% avec un nombre s'élevant, à ce jour, à plus de 400 000 éléments .

L'abondance des éléments rétroviraux endogènes (ERVs) présents actuellement dans le génome humain est le résultat d'une centaine d'endogénisations réussies au cours de l'évolution de la lignée humaine. Les différentes vagues d'endogénisation s'étalent sur une période allant de 2 à 90 millions d'années avant notre ère et ont été suivies de l'expansion du nombre de copies par des phénomènes de type « copier/coller » avec possibilité d'apparition d'erreurs, conduisant à partir d'un provirus ancestral à la formation d'une famille de HERV, c'est à dire un ensemble d'éléments présentant des homologues de séquences. Les plus anciens éléments, ceux de la famille HERV-L, se seraient intégrés avant l'émergence des mammifères. Deux familles HERV-F et HERV-H sont apparues dans la période où les premiers primates faisaient leur apparition. Les familles HERV-FRD et HERV-K (HML-5) intégrées il y a 40 à 55 millions d'années, sont spécifiques des primates supérieurs. En revanche les familles HERV-W et HERV-E, par exemple, se sont intégrées 5 à 10 millions d'années plus tard après la séparation avec les singes du nouveau monde, et sont spécifiques des Catarhini (Hominoïdes et Cercopithèques) .

Les séquences ERVs sont représentées sur l'ensemble des chromosomes, avec une densité variant selon les familles et il

n'existe pas de corrélation entre la proximité physique des ERVs et leur proximité phylogénétique .

Les ERVs ont longtemps été considérés comme des parasites ou comme de simples déchets de l'ADN. Néanmoins, l'impact des ERVs sur l'organisme n'est pas uniquement limité à leur participation passée dans le modelage du génome ou à des recombinaisons délétères pouvant encore subvenir.

L'abondance et la complexité structurelle des ERVs rendent les analyses de leur expression très compliquées et souvent difficilement interprétables. La détection de l'expression de HERV peut refléter l'activation transcriptionnelle de un ou de plusieurs loci au sein d'une même famille. Le ou les loci activés peuvent de plus varier en fonction du tissu et/ou du contexte.

Les présents inventeurs ont maintenant découvert et démontré, que des séquences d'acides nucléiques correspondant à des loci précisément identifiés d'éléments rétroviraux endogènes, sont associés au cancer de la prostate et que ces séquences sont des marqueurs moléculaires de la pathologie. Les séquences identifiées sont soit des provirus, c'est à dire des séquences contenant tout ou partie des gènes *gag*, *pol* et *env* flanqués en 5' et 3' de longues terminaisons répétées (LTR ou « Long Terminal Repeat » selon la terminologie anglo-saxonne) , soit tout ou partie des LTR ou des gènes isolés. Les séquences ADN identifiées sont respectivement référencées en SEQ ID NO : 1 à 75 dans le listage de séquences, leur localisation chromosomique est identifiée dans le tableau ci-dessous (NCBI 36/hg18), ainsi que leur expression, surexpression ou sous-expression représentées par le « ratio d'expression » entre échantillon cancéreux et échantillon normal. Quand l'expression de l'acide nucléique ou le changement d'expression de l'acide nucléique est spécifique du tissu prostate on indique cette information par le symbole « x » dans la colonne tissu cible. Ceci signifie que si une expression ou un changement d'expression de l'acide nucléique concerné est déterminé dans un compartiment biologique autre que le

tissu prostate, ceci représente, à distance, une signature d'un cancer de la prostate. Les séquences ADN identifiées comme étant spécifiques du tissu prostate sont respectivement référencées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32. Les séquences ADN identifiées comme étant non spécifiques du tissu prostate sont respectivement référencées en SEQ ID NOs : 2, 5, 6, 7, 9, 12, 13, 14, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74 et 75.

Tableau

SEQ ID NO:	Localisation chromosomique	Tissu cible	Ratio d'expression cancer/normal
1	(-) chr 8 : 125981185-125988649	x	6,5
2	(-) chr 11 : 60237235-60238528		5,1
3	(-) chr 19 : 20721466-20730278	x	3,5
4	(-) chr 1 : 46569499-46569788	x	2,7
5	(-) chr 2 : 165222667-165224367		2,4
6	(-) chr 5 : 146727162-146727562		2,4
7	(-) chr 7 : 79651365-79652053		2,3
8	(-) chr 3 : 75004674-75009922	x	2,2
9	(+) chr 19 : 60146916-60147844		2,1
10	(+) chr 3 : 74957891-74960634	x	2,1
11	(-) chr 19 : 58096473-58098768	x	2,1
12	(+) chr 1 : 46568022-46568774		2,1
13	(+) chr 6 : 142192789-142193227		2,1
14	(+) chr 8 : 8063655-8067207		-2,0
15	(+) chr 19 : 15807768-15807978	x	2,0
16	(-) chr 13 : 94759022-94759378	x	2,0
17	(+) chr 12 : 31851416-31851846		-2,0
18	(-) chr 4 : 92495874-92498563		2,0
19	(+) chr 4 : 69952306-69955060		1,9
20	(-) chr 2 : 157905798-157908183		-1,9
21	(+) chr 1 : 46558555-46559522	x	1,9
22	(-) chr 10 : 20449793-20453869		-1,9
23	(-) chr x : 135840667-135841473		1,9
24	(+) chr 20 : 24856581-24861663		1,9
25	(+) chr 4 : 153982431-153982932		1,9
26	(-) chr 1 : 144779633-144780605		1,9

27	(-) chr X: 15348 9882-1534 97212		- 1, 9
28	(+) chr 11: 117186039-1171 90257		- 1, 8
29	(-) chr 3: 117306894-117312765		1, 8
30	(+) chr 8: 8094180-8100651		- 1, 8
31	(+) chr 2: 188084458-188084785		1, 8
32	(-) chr 10: 93051085-93057066	x	1, 7
33	(-) chr 2: 54587807-54590183		1, 7
34	(-) chr 2: 188741658-188747 663		1, 7
35	(+) chr X: 92571323-92580146		1, 7
36	(-) chr 4: 92408723-92409131		1, 6
37	(+) chr 8: 90837193-90837630		1, 6
38	(+) chr 2: 201711 970-201712 935		- 1, 6
39	(-) chr 1: 15442071 9-154426128		1, 6
40	(+) chr 6: 15285321 9-152859441		- 1, 6
41	(-) chr 7: 139899253-139900211		- 1, 6
42	(+) chr 1: 146832410-14 6833382		1, 6
43	(-) chr 1: 14477 9633-144780605		1, 6
44	(+) chr 1: 148879269-14888088 9		- 1, 6
45	(-) chr 5: 34514 678-34514916		1, 6
46	(-) chr 3: 17687 9333-17 6879730		1, 6
47	(+) chr 8: 748 96654-74897392		1, 5
48	(-) chr 20: 159111118-15913833		1, 5
49	(-) chr 6: 14405150-14411033		1, 5
50	(-) chr 5: 92818136-9281 9135		- 1, 5
51	(-) chr 8: 54598330-54 600779		1, 5
52	(-) chr X: 78969339-78 970117		1, 5
53	(+) chr 3: 147554294-147559942		1, 5
54	(-) chr 1: 15334421-15335379		1, 5
55	(-) chr 8: 12395268-12398823		- 1, 5
56	(-) chr 3: 171872658-171878745		- 1, 5
57	(-) chr 2: 20737 9807-207385596		1, 5
58	(+) chr 6: 13168612 9-131689771		- 1, 4
59	(-) chr 4: 47707230-47708025		1, 4
60	(-) chr 2: 142 96371 6-142969364		1, 4
61	(+) chr 5: 130936343-130941430		- 1, 4
62	(-) chr 3: 18657458 9-18 6580188		1, 4
63	(+) chr 18: 70111304-7011724 9		- 1, 4
64	(-) chr 8: 56851123-56851350		1, 4
65	(+) chr 19: 63013510-6301474 6		1, 3
66	(+) chr 3: 752 69085-7527 6706		1, 3
67	(-) chr 19: 58067390-58068 685		1, 3
68	(+) chr 8: 91057 690-91058157		1, 3
69	(+) chr 7: 35702274-35703153		1, 3
70	(-) chr 13: 9029882 6-90304533		- 1, 2
71	(-) chr 13: 40347160-40352498		1, 2
72	(-) chr 7: 130691523-130692332		1, 2

73	(+) chr x : 9597773-9597824		-1,2
74	(+) chr 10: 92557026-92562997		1,2
75	(+) chr 5 : 43015565-43018176		1,2

La présente invention a donc pour objet un procédé pour le diagnostic, *in vitro*, du cancer de la prostate ou pour le pronostic de gravité, *in vitro*, du cancer de la prostate dans un échantillon  
 5 biologique prélevé chez un patient, qui comprend une étape de détection d'au moins un produit d'expression d'au moins une séquence d'acide nucléique, ladite séquence d'acide nucléique étant choisie parmi les séquences identifiées en SEQ ID NOS : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au  
 10 moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOS : I à 75.

Le diagnostic permet d'établir si un individu est malade ou non malade. Le pronostic permet d'établir un degré de gravité de la  
 15 maladie (grades et/ou stades) qui a une incidence sur la survie et/ou qualité de vie de l'individu. Dans le cadre de la présente invention, le diagnostic peut être très précoce.

Le pourcentage d'identité décrit ci-dessus a été déterminé en  
 20 prenant en considération la diversité nucléotidique dans le génome. Il est connu que la diversité nucléotidique est plus élevée dans les régions du génome riches en séquences répétées que dans les régions ne contenant pas de séquences répétées. A titre d'exemple, Nickerson D. A. et al. (1) ont montré une diversité d'environ 0,3% (0,32%)  
 25 dans des régions contenant des séquences répétées.

La capacité de discrimination d'un état cancéreux de chacune des séquences identifiées ci-dessus a été mise en évidence à l'aide d'une analyse statistique utilisant la procédure SAM (5) suivie  
 30 d'une correction par le taux de faux positif (6) et d'une suppression des valeurs inférieures à  $2^6$ . Par conséquent chacune des séquences identifiées ci-dessus présente une différence d'expression significative entre un état tumoral et un état normal. Il en résulte

qu'une différence d'expression observée pour une des séquences précitées constitue une signature de la pathologie. Bien entendu, il est possible de combiner les différences d'expression relevées pour plusieurs des séquences référencées ci-dessus par exemple par une ou des associations de 2, 3, 4, 5, 6, 7, 8, 9, 10 et plus voire jusqu'à 75 des séquences listées, de préférence par une ou des associations de 2, 3, 4, 5, 6, 7, 8, 9, 10 des séquences respectivement identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32. En particulier les séquences identifiées en SEQ ID NOS : 1, 4 et 10, prises seules ou en combinaison (2 à 2 ou 3) constituent une ou des signatures privilégiées.

Ainsi, dans le procédé de l'invention on détecte au moins deux produit d'expression respectivement d'au moins deux séquences d'acide nucléique, lesdites séquences d'acide nucléique étant choisies parmi les séquences identifiées en SEQ ID NOS : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOS : 1 à 75.

Dans un mode de réalisation du procédé selon l'invention, on détecte le produit d'expression d'au moins deux séquences d'acide nucléique, lesdites au moins deux séquences d'acide nucléique étant choisies parmi les séquences identifiées comme étant spécifiques du tissu prostate, c'est à dire choisies dans le groupe de séquences identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32.

Dans un autre mode de réalisation du procédé de l'invention, on détecte le produit d'expression d'au moins une séquence choisie parmi les séquences identifiées comme étant spécifiques du tissu prostate, c'est à dire choisies dans le groupe de séquences

identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 et le produit d'expression d'au moins une séquence choisie parmi les séquences identifiées comme étant non spécifiques du tissu prostate, c'est à dire choisies dans le groupe de séquences identifiées en SEQ ID NOs : 2, 5, 6, 7, 9, 12, 13, 14, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74 et 75 ou le produit d'expression d'au moins une séquence choisie parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 et le produit d'expression d'au moins une séquence choisie parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOs : 2, 5, 6, 7, 9, 12, 13, 14, 17, 18, 19, 20, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74 et 75.

De préférence dans le procédé de l'invention on détecte le produit d'expression d'au moins une séquence d'acide nucléique, de préférence d'au moins deux séquences d'acide nucléique ou de trois séquences d'acide nucléique, lesdites séquences d'acide nucléique étant choisies dans le groupe de séquences identifiées en SEQ ID NOs : 1, 4 et 10, ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences identifiées en SEQ ID NOs : 1, 4 et 10.

Le produit d'expression détecté est au moins un transcrit ARN, en particulier au moins un ARNm ou au moins un polypeptide.

Lorsque le produit d'expression est un transcrit ARNm, il est détecté par toute méthode appropriée, telle que l'hybridation, le séquençage ou l'amplification. L'ARNm peut être détecté directement par mise en contact avec au moins une sonde et/ou au moins une  
5 amorce qui sont conçues pour s'hybrider dans des conditions expérimentales prédéterminées aux transcrits ARNm, la mise en évidence de la présence ou de l'absence d'hybridation à l'ARNm et éventuellement la quantification de l'ARNm. Parmi les méthodes préférées on peut citer l'amplification (par exemple la RT-PCR, la  
10 NASBA, etc.), l'hybridation sur puce ou encore le séquençage. L'ARNm peut également être détecté indirectement à partir d'acides nucléiques dérivés desdits transcrits, comme les copies ADNc, etc.

Généralement le procédé de l'invention comprend une étape  
15 initiale d'extraction de l'ARNm de l'échantillon à analyser.

Ainsi, le procédé peut comprendre :

- (i) une étape d'extraction de l'ARNm de l'échantillon à analyser,
- (ii) une étape de détection et de quantification de l'ARNm de l'échantillon à analyser,
- 20 (iii) une étape d'extraction de l'ARNm dans un échantillon de référence, qui peut être un échantillon sain et provenant du même individu ou,
- (iv) une étape de détection et de quantification de l'ARNm de l'échantillon sain,
- 25 (iii) une étape de comparaison de la quantité d'ARNm exprimé dans l'échantillon à analyser et dans l'échantillon de référence ; la détermination d'une quantité d'ARNm exprimé dans l'échantillon à analyser différente de la quantité d'ARNm exprimé dans l'échantillon de référence sain pouvant être corrélée avec le diagnostic ou le  
30 pronostic de gravité d'un cancer de la prostate (la différence de la quantité d'ARNm dans le tissu prostate cancéreux par rapport à la quantité d'ARNm exprimé dans le tissu prostate sain étant

indifféremment une expression, une surexpression ou une sous-expression) ;

et en particulier :

- (i) une extraction de l'ARNm à analyser de l'échantillon,
- 5 (ii) une détermination, dans l'ARN à analyser, d'un niveau d'expression d'au moins une séquence ARN dans l'échantillon, de préférence d'au moins deux séquences ARN dans l'échantillon, la séquence ARN et les séquences ARN étant respectivement le produit de transcription d'au moins une séquence d'acide nucléique choisie  
10 parmi les séquences identifiées en SEQ ID NOS : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOS : 1 à 75, et
- 15 (iii) une comparaison du niveau d'expression de la ou des séquence (s) ARN définie (s) en (ii) avec un niveau d'expression de référence ; la détermination d'un niveau d'expression de l'ARN à analyser présentant une différence par rapport au niveau d'expression de référence pouvant être corrélé avec le diagnostic ou le pronostic  
20 d'un cancer de la prostate (comme déterminé ci-dessus) ; ou
  - (i) une étape d'extraction de l'ARNm de l'échantillon à analyser,
  - (ii) une étape de détection et de quantification de l'ARNm de l'échantillon à analyser,
  - (iii) une étape de comparaison de la quantité d'ARNm exprimé dans  
25 l'échantillon à analyser par rapport à une quantité d'ARNm de référence, la détermination d'une quantité d'ARNm exprimé dans l'échantillon à analyser différente de la quantité d'ARNm de référence pouvant être corrélée avec le diagnostic ou le pronostic d'un cancer de la prostate (la différence de la quantité d'ARNm dans  
30 l'échantillon à analyser par rapport à la quantité d'ARNm de référence étant indifféremment une expression, une surexpression ou une sous-expression) .

Dans un mode de réalisation du procédé de l'invention on prépare des copies ADN de l'ARNm, on met en contact les copies ADN avec au moins une sonde et/ou au moins une amorce dans des conditions prédéterminées permettant l'hybridation et en ce qu'on  
5 détecte la présence ou l'absence d'hybridation aux dites copies ADN.

Le produit d'expression qui est détecté peut aussi être un polypeptide qui est le produit de la traduction d'au moins un des transcrits décrits ci-dessus. Auquel cas, on détecte le polypeptide  
10 exprimé par mise en contact avec au moins un partenaire de liaison spécifique dudit polypeptide, en particulier un anticorps ou un analogue d'anticorps ou un aptamère. Le partenaire de liaison est de préférence un anticorps, par exemple un anticorps monoclonal ou un  
15 anticorps polyclonal hautement purifié ou un analogue d'un anticorps, par exemple une protéine d'affinité aux propriétés compétitives (nanofitine™) .

Les anticorps polyclonaux peuvent être obtenus par immunisation d'un animal avec l'immunogène approprié, suivie de la récupération des anticorps recherchés sous forme purifiée, par prélèvement du  
20 sérum dudit animal, et séparation desdits anticorps des autres constituants du sérum, notamment par chromatographie d'affinité sur une colonne sur laquelle est fixé un antigène spécifiquement reconnu par les anticorps .

Les anticorps monoclonaux peuvent être obtenus par la technique  
25 des hybridomes dont le principe général est rappelé ci-dessous.

Dans un premier temps, on immunise un animal, généralement une souris avec l'immunogène approprié, dont les lymphocytes B sont alors capables de produire des anticorps contre cet antigène. Ces lymphocytes producteurs d'anticorps sont ensuite fusionnés avec des  
30 cellules myéломateuses "immortelles" (murines dans l'exemple) pour donner lieu à des hybridomes. A partir du mélange hétérogène des cellules ainsi obtenu, on effectue alors une sélection des cellules capables de produire un anticorps particulier et de se multiplier indéfiniment. Chaque hybridome est multiplié sous la forme de clone,

chacun conduisant à la production d'un anticorps monoclonal dont les propriétés de reconnaissance vis-à-vis de la protéine pourront être testées par exemple en ELISA, par immunotransfert (Western blot) en une ou deux dimensions, en immunofluorescence, ou à l'aide d'un biocapteur. Les anticorps monoclonaux ainsi sélectionnés, sont par la suite purifiés notamment selon la technique de chromatographie d'affinité décrite ci-dessus.

Les anticorps monoclonaux peuvent être également des anticorps recombinants obtenus par génie génétique, par des techniques bien connues de l'homme du métier.

Les nanofitines™ sont de petites protéines qui comme les anticorps sont capables de se lier à une cible biologique permettant ainsi de la détecter, de la capturer ou tout simplement de la cibler au sein d'un organisme. On les présente, entre autres, comme des analogues d'anticorps.

Les aptamères sont des oligonucléotides synthétiques capables de fixer un ligand spécifique.

L'invention concerne également l'utilisation d'au moins une séquence d'acide nucléique, isolée, comme marqueur moléculaire pour le diagnostic ou le pronostic *in vitro* du cancer de la prostate, caractérisée en ce que ladite séquence d'acide nucléique consiste en :

(i) au moins une séquence ADN choisie parmi les séquences SEQ ID NOS : 1 à 75, ou

(ii) au moins une séquence ADN complémentaire d'une séquence choisie parmi les séquences SEQ ID NOS : 1 à 75, ou

(iii) au moins une séquence ADN qui présente au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une séquence telle que définie en (i) et (ii), ou

(iv) au moins une séquence ARN qui est le produit de transcription d'une séquence choisie parmi les séquences telles que définies en (i), ou

(v) au moins une séquence ARN qui est le produit de transcription d'une séquence choisie parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une séquence telle que définie en (i).

10 Dans un mode de réalisation, on utilise au moins deux séquences d'acide nucléique qui consistent en :

(i) au moins deux séquences ADN choisies parmi les séquences SEQ ID NOs : 1 à 75, de préférence choisies parmi les séquences identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 et en particulier les séquences SEQ ID Nos : 1, 4 et 10, ou

(ii) au moins deux séquences ADN respectivement complémentaires d'au moins deux séquences choisies parmi les séquences SEQ ID NOs : 1 à 75, de préférence choisies parmi les séquences identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 et en particulier choisies parmi les séquences SEQ ID Nos : 1, 4 et 10, ou

(iii) au moins deux séquences ADN qui présentent respectivement au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec deux séquences telles que définies en (i) et (ii), ou

25 (iv) au moins deux séquences ARN qui sont respectivement le produit de transcription de deux séquences choisies parmi les séquences telles que définies en (i), ou

(v) au moins deux séquences ARN qui sont le produit de transcription de deux séquences choisies parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et

avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences telles que définies en (i) .

L'invention a aussi pour objet un kit pour le diagnostic ou le pronostic, *in vitro*, du cancer de la prostate dans un échantillon biologique prélevé chez un patient qui comprend au moins un partenaire de liaison spécifique d'au moins un produit d'expression d'au moins une séquence d'acide nucléique choisie parmi les séquences identifiées en SEQ ID NOs : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité, avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences d'acide nucléique identifiées en SEQ ID NOs 1 à 75 et pas plus de 75 partenaires de liaisons spécifiques des produits d'expression des séquences d'acide nucléique identifiées en SEQ ID NOs 1 à 75 ou des séquences d'acides nucléiques qui présentent au moins 99% d'identité avec les séquences d'acides nucléiques identifiées en SEQ ID NOs 1 à 75, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec une des séquences identifiées en SEQ ID NOs : 1 à 75.

Dans un mode de réalisation le kit comprend au moins deux partenaires de liaison respectivement spécifiques d'au moins deux produits d'expression d'au moins deux séquences d'acide nucléique choisies parmi les séquences identifiées en SEQ ID NOs : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences d'acide nucléique identifiées en SEQ ID NOs 1 à 75 et pas plus de 75 partenaires de liaisons spécifiques des produits d'expression des séquences d'acide nucléique identifiées en SEQ ID NOs 1 à 75 ou des séquences d'acides nucléiques qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins

99,6% ou au moins 99,7% d'identité avec les séquences d'acides nucléiques identifiées en SEQ ID NOS 1 à 75.

Par exemple, le kit comprend au moins deux partenaires de liaison respectivement spécifiques du produit d'expression d'au moins deux séquences d'acide nucléique choisies dans le groupe de séquences identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 ou des séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou 99,7% d'identité avec les séquences identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32.

De préférence, le kit comprend un partenaire de liaison spécifique du produit d'expression d'au moins une séquence d'acide nucléique choisie dans le groupe de séquences identifiées en SEQ ID NOS : 1, 4 et 10 ou des séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences identifiées en SEQ ID NOS : 1, 4 et 10.

20

En particulier le kit comprend 1, 2 ou 3 partenaire (s) de liaison spécifique (s) du ou des produit (s) d'expression des séquences d'acide nucléique identifiées en SEQ ID NOS : 1, 4 et 10 ou des séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences identifiées en SEQ ID NOS : 1, 4 et 10.

Le au moins partenaire de liaison spécifique du produit d'expression répond aux définitions données ci-dessus.

L'invention concerne encore un procédé pour évaluer l'efficacité d'un traitement et/ou une progression dans un cancer de la prostate qui comprend un étape d'obtention d'une série d'échantillons biologiques, une étape de détection d'au moins un

35

produit d'expression d'au moins une séquence d'acide nucléique dans ladite série d'échantillons biologiques, ladite séquence d'acide nucléique étant choisie parmi les séquences identifiées en SEQ ID NOs : 1 à 75 avec une des séquences identifiées en SEQ ID NOs : 1 à 5 75 ou des séquences qui présentent au moins 99% d'identité, de préférence au moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences identifiées en SEQ ID NOs 1 à 75.

10 Dans un mode de réalisation, on détecte au moins deux produits d'expression d'au moins deux séquences d'acide nucléique, lesdites deux séquences d'acide nucléique étant choisies parmi les séquences identifiées en SEQ ID NOs : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité respectivement, de préférence au 15 moins 99,5% d'identité et avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences identifiées en SEQ ID NOs : 1 à 75.

Dans un autre mode de réalisation du procédé, on détecte le produit d'expression d'au moins une séquence d'acide nucléique, de 20 préférence d'au moins deux séquences d'acides nucléiques ou de trois séquences d'acide nucléique, lesdites séquences d'acide nucléiques étant choisies dans le groupe de séquences identifiées en SEQ ID NOs 1, 4 et 10 ou parmi les séquences qui présentent au moins 99% d'identité respectivement, de préférence au moins 99,5% d'identité et 25 avantageusement au moins 99,6% ou au moins 99,7% d'identité avec les séquences identifiées en SEQ ID NOs : 1, 4 et 10.

Par échantillon biologique on entend un tissu, un fluide, des composants desdits tissu et fluide, tels que des cellules ou des 30 corps apoptotiques, des vésicules excrétées, comprenant notamment des exosomes et des microvésicules. A titre d'exemple, l'échantillon biologique peut être issu d'une biopsie de prostate réalisée au préalable chez un patient suspecté d'être atteint d'un cancer de la prostate ou être issu d'une biopsie pratiquée sur un organe autre 35 que la prostate chez un patient présentant des métastases. Dans ce

deuxième cas, quand le changement d'expression de l'acide nucléique (marqueur moléculaire) est spécifique de l'organe prostate, il est possible de remonter au cancer primaire, c'est à dire au cancer du prostate. L'échantillon biologique peut être également un fluide  
5 biologique, tel que du sang ou une fraction sanguine (sérum, plasma), de l'urine, de la salive, du liquide céphalo-rachidien, de la lymphe, du lait maternel, du sperme, de même que des composants desdits fluides, en particulier des vésicules excrétées telles que définies ci-dessus. Par exemple, la détection d'un transcrit spécifique du  
10 tissu prostate dans un exosome ou une microvésicule, originaire d'une cellule épithéliale, signe soit la présence d'un cancer primaire, soit des métastases, sans qu'il soit nécessaire de faire un prélèvement au niveau de l'organe.

15

**Figures :**

Les figures 1 et 2 représentent le différentiel d'expression observé dans le cancer de la prostate pour un ensemble de séquences HERV. Plus précisément, la figure 1 (clustering) regroupe de manière  
20 exploratoire les éléments HERV qui ont un tropisme d'expression associé au cancer de la prostate par rapport à l'ensemble des tissus contrôles, et la figure 2 montre les différences statistiques d'expression d'éléments HERV entre prostate normale et prostate tumorale .

25 Les figures 3 et 4 montrent la détection de séquences HERV dans deux fluides biologiques : les urines et les sérums .

Les figures 5 à 7 montrent trois exemples de séquences HERV présentant un différentiel d'expression dans les urines en association avec le statut clinique cancéreux des patients.

30

**Exemples :**

**Exemple 1 : Identification de séquences HERV présentant un différentiel d'expression dans le cancer de la prostate**

**Méthode :**

L'identification de séquences HERV présentant un différentiel d'expression dans le cancer de la prostate repose sur la conception et l'utilisation d'une puce à ADN haute densité au format GeneChip, dénommée HERV-V2, conçue par les inventeurs et dont la fabrication a été sous-traitée à la société Affymetrix. Cette puce contient des sondes qui correspondent à des séquences HERV distinctes au sein du génome humain. Ces séquences ont été identifiées à partir d'un ensemble de références prototypiques découpées en régions fonctionnelles (LTR, *gag*, *pol* et *env*) puis, par une recherche de similarité à l'échelle du génome humain entier (NCBI 36/hgl8), 10 035 loci HERV distincts ont été identifiés, annotés, et finalement regroupés dans une banque de données nommée HERVgDB3 .

Les sondes entrant dans la composition de la puce ont été définies à partir de HERVgDB3 et sélectionnées en appliquant un critère de spécificité d'hybridation, dont le but est d'exclure du procédé de création les sondes présentant un risque d'hybridation élevé avec une cible non recherchée. Pour cela, les séquences de HERVgDB3 ont d'abord été segmentées par ensemble de 25 nucléotides (25-mers) chevauchants, aboutissant à un ensemble de sondes candidates. Le risque d'hybridation aspécifique a ensuite été évalué pour chaque sonde candidate en réalisant des alignements sur l'ensemble du génome humain à l'aide de l'algorithme KASH (2) . Un score expérimental sanctionne le résultat de l'hybridation, addition de l'impact du nombre, du type et de la position des erreurs dans l'alignement. La valeur de ce score est corrélée au potentiel d'hybridation cible/sonde. La connaissance de tous les potentiels d'hybridation d'une sonde candidate sur l'ensemble du génome humain permet d'évaluer sa spécificité de capture. Les sondes candidates qui présentent une bonne affinité de capture sont conservées puis regroupées en « probesets » et enfin synthétisées sur la puce HERV-V2 .

Les échantillons analysés à l'aide de la puce haute densité HERV-V2 correspondent à des ARN extraits de tumeurs et aux ARN

extraits des tissus sains adjacents à ces tumeurs. Les tissus analysés sont la prostate, avec en contrôles le sein, l'ovaire, l'utérus, le côlon, le poumon, le testicule et le placenta. Dans le cas du placenta, seuls des tissus sains ont été utilisés. Pour chaque échantillon, 50ng d'ARN ont servi à la synthèse de cDNA en utilisant le protocole d'amplification connu sous le nom de WTO. Le principe de l'amplification WTO est le suivant : des amorces aléatoires, ainsi que des amorces ciblant l'extrémité 3' du transcrit ARN, sont ajoutées, avant une étape de transcription inverse suivie d'une amplification linéaire et simple brin désignée SPIA. Les cDNA sont ensuite dosés, caractérisés et purifiés, puis 2µg sont fragmentés, marqués à la biotine en extrémité 3' par l'action de l'enzyme *terminal transferase*. Le produit cible ainsi préparé est mélangé à des oligonucléotides de contrôle, puis l'hybridation est réalisée selon le protocole recommandé par la société Affymetrix. Les puces sont alors révélées et lues pour acquérir l'image de leur fluorescence. Un contrôle qualité se basant sur les contrôles standards est réalisé, et un ensemble d'indicateurs (MAD, MAD-Med plots, RLE) servent à exclure les puces non conformes à une analyse statistique.

L'analyse des puces consiste d'abord en un pré-traitement des données par l'application d'une correction du bruit de fond basée sur l'intensité des signaux des sondes tryptophanes, suivie d'une normalisation RMA (3) basée sur la méthode des quantiles. Puis une double correction des effets liés aux lots d'expériences est réalisée en appliquant la méthode COMBAT (4) afin de garantir que les différences d'expression observées sont d'origine biologique et non techniques. A ce stade, une analyse exploratoire des données est conduite à l'aide d'outils de regroupement de données par partitionnement euclidien (*clustering*), et enfin une analyse statistique utilisant la procédure SAM (5) suivie d'une correction par le taux de faux positif (6) et d'une suppression des valeurs inférieures à  $2^6$  est appliquée pour la recherche de séquences présentant un différentiel d'expression entre l'état normal et l'état tumoral d'un tissu.

**Résultats :**

Le traitement des données générées par l'analyse des puces à ADN HERV-V2 à l'aide de cette méthode a permis d'identifier un ensemble de « probesets » présentant une différence d'expression statistiquement significative entre la prostate normale et la prostate tumorale. Les résultats du *clustering* ainsi que la recherche d'expression différentielle au sein des échantillons contrôles a par ailleurs mis en évidence des éléments HERV dont l'expression différentielle est spécifiquement associée à la prostate tumorale.

Les séquences nucléotidiques des éléments HERV présentant un différentiel d'expression dans la prostate tumorale sont identifiées par les SED ID N°1 à 75, la localisation chromosomique de chaque séquence est donnée dans le référentiel NCBI 36/hgl8, et la mention « tissu cible » (une croix) pointe les éléments dont l'expression différentielle n'a été observée que dans la comparaison entre prostate normale et prostate tumorale (par rapport aux comparaisons au sein des tissus contrôles). Une valeur indicative du ratio d'expression entre état normal et état tumoral est également communiquée, et sert à ordonner les séquences dans un soucis de présentation uniquement.

**Exemple 2 : Détection de séquences HERV dans les fluides biologiques****Principe :**

Les inventeurs ont montré que des séquences HERV sont détectées dans les fluides biologiques, ce qui permet entre autres de caractériser un cancer de la prostate par recours à une détection à distance de l'organe primaire. Une étude a été menée sur 20 échantillons d'urine et 38 échantillons de sérum provenant d'individus différents.

Les sérums et les urines ont été centrifugés dans les conditions suivantes :

Sérums : 500 g pendant 10 minutes à 4°C. Le surnageant a été récupéré et centrifugé de nouveau à 16500 g pendant 20 minutes à 5 4°C. Le surnageant de cette deuxième centrifugation, dépourvu de cellules, mais comprenant encore des exosomes, des microvésicules, des acides nucléiques, des protéines, a été analysé sur des puces. La puce est la puce HERV-V2 utilisée selon les modalités précédemment décrites .

10 Urines : après recueil, centrifugation à 800 g pendant 4 minutes à 4°C. Le culot a été récupéré avec du RNA protect cell reagent™. Puis, centrifugation à 5000 g pendant 5 minutes avant ajout du tampon de lyse sur le culot. La puce est la puce HERV-V2 utilisée selon les modalités précédemment décrites.

#### 15 **Résultats :**

Un nombre important de signaux positifs, incluant les signaux d'expression correspondant aux séquences listées dans le tableau ci-dessus, a été détecté à la fois dans les surnageants de sérums et dans les culots cellulaires provenant d'urines, comme illustré dans 20 les figures 3 et 4. Ceci confirme que les fluides biologiques, en particulier le sérum et l'urine, sont une source utilisable de matériel biologique pour la détection de séquences HERV. Il est communément accepté que le seuil de positivité est de l'ordre de 2<sup>6</sup> soit 64.

25

#### ***Exemple 3 : Mise en évidence d'une expression différentielle de séquences HERV dans des fluides biologiques dans le cas du cancer de la prostate***

##### **Principe :**

30 Deux classes cliniques ont été définies : (PBPNeg) absence de cancer de la prostate établie par références biopsiques ; (CAPR)

cancer de la prostate établi après analyse anatomopathologique des pièces de prostatectomies du patient. Les urines des patients ont été recueillies et traitées selon le protocole décrit plus haut. La puce HERV-V2 a été utilisée selon les modalités précédemment  
5 décrites afin de mettre en évidence les séquences HERV présentant un différentiel d'expression entre les deux classes cliniques dans une étude incluant 20 patients.

## 10 **Résultats :**

Un ensemble de séquences HERV présentant un différentiel d'expression statistiquement significatif entre les classes cliniques a été identifié. Trois exemples parmi ces séquences HERV sont montrés en figures 5 à 7. Chaque point représente la valeur  
15 d'expression de la séquence considérée chez un individu. La barre horizontale indique la médiane des valeurs. Les trois exemples montrent le caractère discriminant du niveau d'expression des séquences considérées dans la mesure où les variances des groupes PBPNeg et CAPR sont significativement différentes (test de Fisher,  
20 pvalue inférieure à 0,05) .

## Références bibliographiques

1. Nickerson, D.A., Taylor, S.L., Weiss, K.M., Clark, A.G.,  
Hutchinson, R.G., Stengard, J., Salomaa, V., Vartiainen, E.,  
5 Boerwinkle, E. and Sing, C.F. (1998) DNA séquence diversity in a  
9.7-kb région of the human lipoprotein lipase gene. *Nat. Genét.*,  
**19**, 233-240.
2. Navarro, G. and Raffinot, M. (2002) Flexible Pattern Matching in  
Strings : Practical On-Line Search Algorithms for Texts and  
10 Biological Séquences. Cambridge University Press.
3. Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D.,  
Antonellis, K.J., Scherf, U. and Speed, T.P. (2003) Exploration,  
normalization, and summaries of high density oligonucleotide  
array probe level data. *Biostatistics (Oxford, England)*, **4**, 249-  
15 264.
4. Johnson, W.E., Li, C. and Rabinovic, A. (2007) Adjusting batch  
effects in microarray expression data using empirical Bayes  
methods. *Biostatistics (Oxford, England)*, **8**, 118-127.
5. Tusher, V.G., Tibshirani, R. and Chu, G. (2001) Significance  
20 analysis of microarrays applied to the ionizing radiation  
response. *Proceedings of the National Academy of Sciences of the  
United States of America*, **98**, 5116-5121.
6. Storey, J.D. and Tibshirani, R. (2003) Statistical significance  
for genomewide studies. *Proceedings of the National Academy of  
25 Sciences of the United States of America*, **100**, 9440-9445.

**REVENDICATIONS**

1. Procédé pour le diagnostic ou le pronostic, *in vitro*, du cancer  
5 de la prostate dans un échantillon biologique prélevé chez un patient qui comprend une étape de détection d'au moins deux produits d'expression respectivement d'au moins deux séquences d'acide nucléique, lesdites séquences d'acide nucléique étant choisies parmi les séquences identifiées en SEQ ID NOs : 1 à 75 ou parmi les  
10 séquences qui présentent au moins 99% d'identité avec une des séquences identifiées en SEQ ID NOs : 1 à 75.

2. Procédé selon la revendication 1, dans lequel est détecté le produit d'expression d'au moins deux séquences d'acide nucléique,  
15 lesdites au moins deux séquences d'acide nucléique étant choisies dans le groupe de séquences identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 ou parmi les séquences qui présentent au moins 99% d'identité avec les séquences identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32.

20 3. Procédé selon la revendication 2, dans lequel est détecté le produit d'expression d'au moins deux séquences d'acide nucléique, de préférence de trois séquences d'acide nucléique, lesdites séquences d'acide nucléique étant choisies dans le groupe de séquences  
25 identifiées en SEQ ID NOs 1, 4 et 10, ou parmi les séquences qui présentent au moins 99% d'identité avec les séquences identifiées en SEQ ID NOs : 1, 4 et 10.

4. Procédé selon la revendication 1, dans lequel le produit  
30 d'expression détecté est au moins un transcrit ARN ou au moins un polypeptide .

5. Procédé selon la revendication 4, caractérisé en ce que le transcrit ARN est au moins un ARNm.

35

6. Procédé selon la revendication 4 ou 5, dans lequel le transcrit ARN, en particulier l'ARNm est détecté par hybridation, par amplification ou par séquençage .

5 7. Procédé selon la revendication 6, dans lequel l'ARNm est mis en contact avec au moins une sonde et/ou au moins une amorce dans des conditions prédéterminées permettant l'hybridation et en ce qu'on détecte la présence ou l'absence d'hybridation à l'ARNm.

10 8. Procédé selon la revendication 5, caractérisé en ce qu'on prépare des copies ADN de l'ARNm, on met en contact les copies ADN avec au moins une sonde et/ou au moins une amorce dans des conditions prédéterminées permettant l'hybridation et en ce qu'on détecte la présence ou l'absence d'hybridation aux dites copies ADN.

15

9. Procédé selon la revendication 4, dans lequel on détecte le polypeptide exprimé par mise en contact avec au moins un partenaire de liaison spécifique dudit polypeptide, en particulier un anticorps ou un analogue d'anticorps, une protéine d'affinité ou un aptamère.

20

10. Utilisation d'au moins deux séquences d'acide nucléique, isolées, comme marqueur moléculaire pour le diagnostic ou le pronostic *in vitro* du cancer de la prostate, caractérisée en ce que les deux séquences d'acide nucléique consistent en :

25 (i) au moins deux séquences ADN choisies parmi les séquences SEQ ID NOs : 1 à 75, de préférence choisies parmi les séquences identifiées en SEQ ID NOs : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 et en particulier parmi les séquences identifiées en SEQ ID Nos : 1, 4 et 10, ou

30 (ii) au moins deux séquences ADN respectivement complémentaires d'au moins deux séquences choisies parmi les séquences SEQ ID NOs : 1 à 75, de préférence choisies parmi les séquences identifiées en SEQ ID

NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 et en particulier choisies parmi les séquences identifiées en SEQ ID Nos : 1, 4 et 10, ou

(iii) au moins deux séquences ADN qui présentent au moins 99% d'identité avec deux séquences telles que définies en (i) et (ii),  
5 ou

(iv) au moins deux séquences ARN qui sont respectivement le produit de transcription de deux séquences choisies parmi les séquences telles que définies en (i), ou

10 (v) au moins deux séquences ARN qui sont le produit de transcription d'au moins deux séquences choisies parmi les séquences qui présentent au moins 99% d'identité avec les séquences telles que définies en (i).

15 11. Kit pour le diagnostic ou le pronostic, *in vitro*, du cancer de la prostate dans un échantillon biologique prélevé chez un patient qui comprend au moins deux partenaires de liaison respectivement spécifiques d'au moins deux produits d'expression d'au moins deux séquences d'acide nucléique choisies parmi les séquences identifiées  
20 en SEQ ID NOS : 1 à 75 ou parmi les séquences qui présentent au moins 99% d'identité avec les séquences d'acide nucléique identifiées en SEQ ID NOS 1 à 75 et pas plus de 75 partenaires de liaison spécifiques des produits d'expression des séquences d'acide nucléique identifiées en SEQ ID NOS 1 à 75 ou des séquences d'acides  
25 nucléiques qui présentent au moins 99% d'identité avec les séquences d'acides nucléiques identifiées en SEQ ID NOS 1 à 75.

12. Kit selon la revendication 11, qui comprend au moins deux partenaires de liaison respectivement spécifiques du produit  
30 d'expression d'au moins deux séquences d'acide nucléique choisies dans le groupe de séquences identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 ou parmi les séquences qui présentent au

moins 99% d'identité avec les séquences identifiées en SEQ ID NOS :  
1, 3, 4, 8, 10, 11, 15, 16, 21 et 32.

13. Kit selon la revendication 12, qui comprend au moins deux  
5 partenaires de liaison respectivement spécifiques du produit  
d'expression d'au moins deux séquences d'acide nucléique choisies  
dans le groupe de séquences identifiées en SEQ ID NOS : 1, 4 et 10  
ou parmi les séquences qui présentent au moins 99% d'identité avec  
les séquences identifiées en SEQ ID NOS : 1, 4 et 10.

10

14. Kit selon la revendication 12 qui comprend 2 ou 3 partenaires de  
liaison respectivement spécifiques des produits d'expression des  
séquences d'acide nucléique identifiées en SEQ ID NOS : 1, 4 et 10  
ou des séquences d'acide nucléique qui présentent au moins 99%  
15 d'identité, avec les séquences identifiées en SEQ ID NOS : 1, 4 et  
10.

15. Kit selon l'une quelconque des revendications 11 à 14, dans  
lequel les au moins deux partenaires de liaison respectivement  
20 spécifiques des produits d'expression sont respectivement au moins  
une sonde d'hybridation et/ou au moins une amorce d'amplification,  
ou au moins un anticorps, ou au moins un analogue d'anticorps, ou au  
moins une protéine d'affinité, ou au moins un aptamère.

25 16. Procédé pour évaluer l'efficacité d'un traitement et/ou une  
progression dans un cancer de la prostate qui comprend une étape de  
détection d'au moins deux produits d'expression respectivement d'au  
moins deux séquences d'acide nucléique, lesdites deux séquences  
d'acide nucléique étant choisies parmi les séquences identifiées en  
30 SEQ ID NOS : 1 à 75 ou parmi les séquences qui présentent au moins 99%  
d'identité respectivement avec les séquences identifiées en SEQ ID NOS :  
1 à 75.

17. Procédé selon la revendication 16, dans lequel on détecte le  
35 produit d'expression d'au moins deux séquences d'acide nucléique,

lesdites au moins deux séquences d'acide nucléique étant choisies dans le groupe de séquences identifiées en SEQ ID NOS : 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32 ou parmi les séquences qui présentent au moins 99% d'identité avec les séquences identifiées en SEQ ID NOS :

5 1, 3, 4, 8, 10, 11, 15, 16, 21 et 32.

18. Procédé selon la revendication 17, dans lequel est détecté le produit d'expression d'au moins deux séquences d'acide nucléique, de préférence de trois séquences d'acide nucléique, lesdites séquences

10 d'acide nucléiques étant choisies dans le groupe de séquences identifiées en SEQ ID NOS 1, 4 et 10 ou parmi les séquences qui présentent au moins 99% d'identité avec les séquences identifiées en SEQ ID NOS 1, 4 et 10.

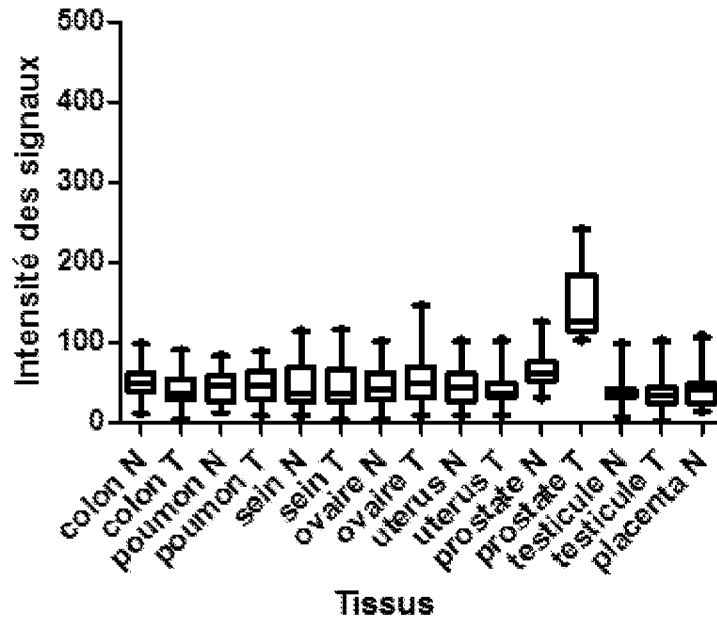


Figure 1

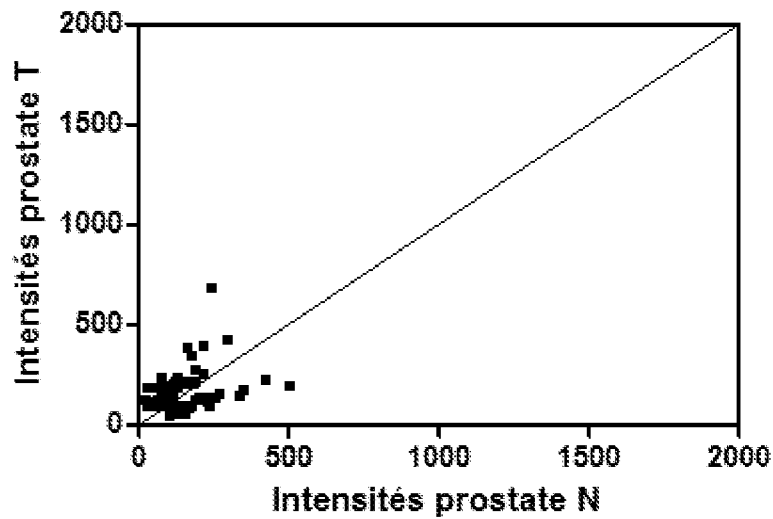


Figure 2

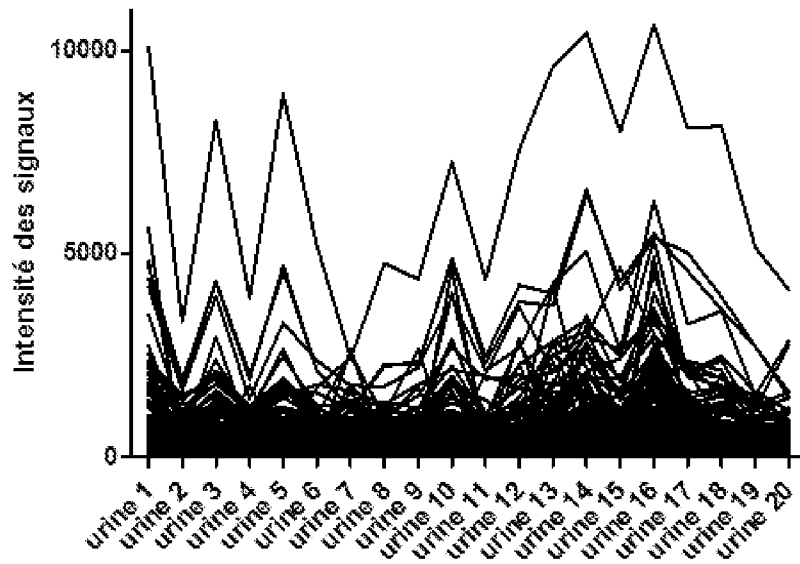


Figure 3

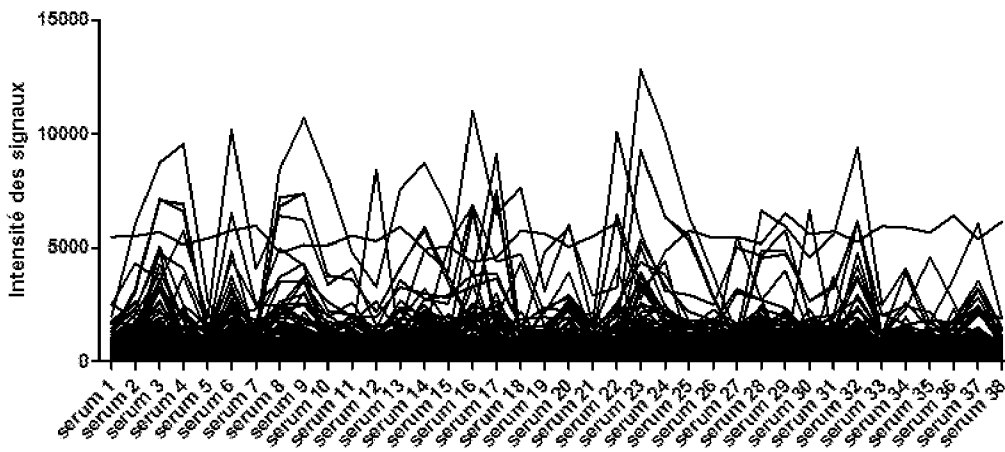


Figure 4

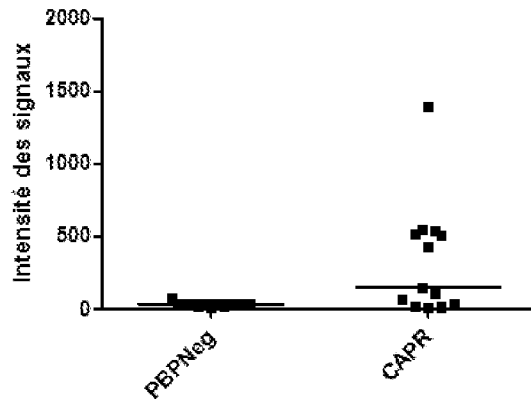


Figure 5

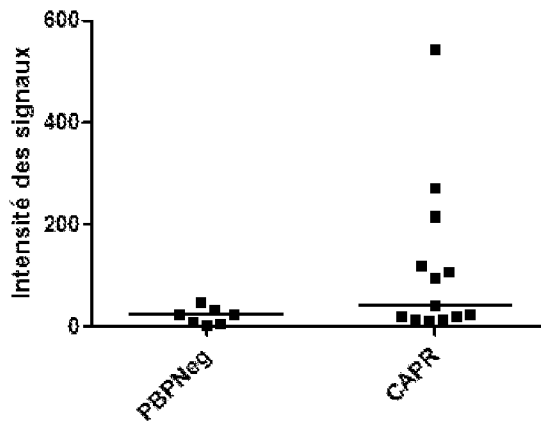


Figure 6

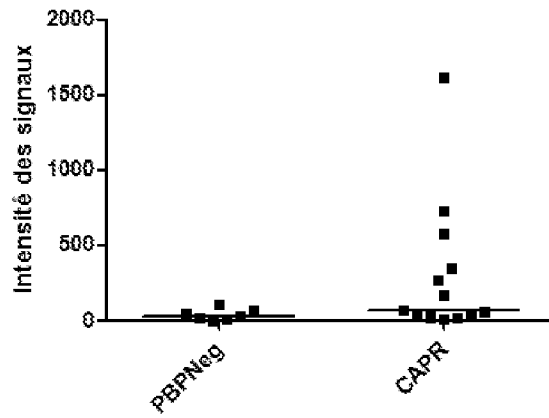


Figure 7