

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

(19) Organisation Mondiale de la Propriété  
Intellectuelle  
Bureau international



(43) Date de la publication internationale  
2 février 2006 (02.02.2006)

PCT

(10) Numéro de publication internationale  
WO 2006/010805 A2

(51) Classification internationale des brevets :  
G01N 33/68 (2006.01)

(74) Mandataire : SAUVAGE, Renée; Cabinet Sauvage, 65,  
boulevard Soult, F-75012 Paris (FR).

(21) Numéro de la demande internationale :  
PCT/FR2005/001548

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

(22) Date de dépôt international : 21 juin 2005 (21.06.2005)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :  
0407114 29 juin 2004 (29.06.2004) FR  
0408714 6 août 2004 (06.08.2004) FR  
0413063 8 décembre 2004 (08.12.2004) FR

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

(71) Déposants (pour tous les États désignés sauf US) : ABAG  
Société Anonyme [FR/FR]; 17, avenue du Parc, Tech-  
nopolis, F-91380 Chilly Mazarin (FR). UNIVERSITE  
CLAUDE BERNARD LYON 1 [FR/FR]; Etablissement  
public à Caractères Scientifique, Culturel et Professionnel,  
43, boulevard du 11 novembre 1918, F-69100 Villeur-  
banne (FR).

Publiée :

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

(72) Inventeurs; et

(75) Inventeurs/Déposants (pour US seulement) : ARIE,  
Jean-Philippe [FR/FR]; 65/67, rue Brillat Savarin,  
F-75013 Paris (FR). CYNOCYNATUS, Camille [FR/FR];  
72, rue de l'Amiral Roussin, F-75015 Paris (FR).

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

(54) Title: NOVEL POLYPEPTIDES FOR ISOLATING IN VITRO AND PREVENTING <I>LEGIONELLA PNEU-  
MOPHILA</I> INFECTIONS

(54) Titre : NOUVEAUX POLYPEPTIDES POUR LA MISE EN EVIDENCE IN VITRO ET LA PREVENTION DES INFEC-  
TIONS A LEGIONELLA PNEUMOPHILA

(57) Abstract: The invention concerns novel polynucleotides, comprising sequence ID SEQ N°1 or parts or variants of said se-  
quence, novel polypeptides encoded by said polynucleotides, expression vectors comprising said polynucleotides and host cells  
comprising said expression vectors. Said polynucleotides and said polypeptides are useful in the field of *in vitro* diagnosis and/or  
for production vaccines against *Legionella pneumophila*.

(57) Abrégé : L'invention a pour objet de nouveaux polynucléotides, comprenant la séquence ID SEQ N°1 ou des parties ou des  
variants de ladite séquence, de nouveaux polypeptides codés par ces polynucléotides, des vecteurs d'expression comprenant lesdits  
polynucléotides et des cellules hôtes comprenant lesdits vecteurs d'expression. Ces polynucléotides et ces polypeptides sont utili-  
sables dans le domaine du diagnostic *in vitro* et/ou la production de vaccins contre . *Legionella pneumophila*.

WO 2006/010805 A2

Nouveaux polypeptides pour la mise en évidence *in vitro* et la prévention des infections à *Legionella pneumophila*.

La présente invention concerne un outil de diagnostic sérologique multiparamétrique des infections à *Legionella pneumophila*.

L'invention comprend l'identification de nouveaux polynucléotides et polypeptides, leur production, leur optimisation et leur utilisation, d'une part, dans le domaine du diagnostic des infections impliquant *Legionella pneumophila* et, d'autre part, dans le domaine de la production de vaccins vis-à-vis de cette bactérie.

Les légionelloses sont des infections respiratoires sévères résultant de l'inhalation d'aérosols contaminés par des bactéries à croissance intracellulaire du genre *Legionella*, et qui se manifestent sous forme sporadique ou épidémique. Ces infections surviennent aussi bien en ville (infections communautaires) qu'en milieu hospitalier (infections nosocomiales). L'espèce *Legionella pneumophila* est responsable de 95% de l'ensemble des cas de légionellose.

Les légionelloses se présentent principalement comme des infections pulmonaires aiguës. Elles se caractérisent par la sévérité du tableau clinique : après une incubation de 2 à 10 jours et un début pseudo-grippal, le malade présente une température élevée, un malaise général, des douleurs abdominales, voire des troubles psychiques. L'atteinte radiologique est souvent bilatérale. La mortalité associée, particulièrement élevée chez certaines personnes dites à risque (sujets âgés ou souffrant d'un cancer, d'une hémopathie), varie de 10 à 20% (13% des 835 cas recensés en France en 2002).

Il est essentiel de pouvoir poser précocement le diagnostic de légionellose du fait de l'enjeu thérapeutique : en effet, les légionelles sont insensibles aux bêta-lactamines - qui sont les principaux antibiotiques prescrits en cas de pneumopathie - et il faut avoir recours

à un traitement spécifique reposant sur les macrolides (érythromycine) ou les fluoroquinolones.

A l'heure actuelle, le diagnostic est réalisé par trois méthodes générales : la mise en évidence des bactéries à partir de prélèvements broncho-pulmonaires (soit directement par immunofluorescence soit par culture sur milieux sélectifs) ; la détection d'antigènes urinaires spécifiques ; la mise en évidence d'une augmentation significative du titre des anticorps dans le sérum (sérologie).

L'examen direct des prélèvements réalisé par microscopie à fluorescence se pratique, pour la majorité des réactifs commercialisés, à l'aide d'anticorps monoclonaux ou polyclonaux qui reconnaissent l'ensemble des sérogroupes connus de *L. pneumophila*. Le principal inconvénient de cette technique est sa faible sensibilité. Sa spécificité est également imparfaite (~90%), du fait de réactions immunologiques croisées avec certaines bactéries comme *Pseudomonas aeruginosa*, *Bordetella pertussis* ou *Bacteroides fragilis*. La culture sur milieux spécifiques est la méthode de choix, notamment dans un but épidémiologique (typage des souches), mais donne des résultats tardifs puisque les légionelles cultivent en 3 à 4 jours, voire plus. La technique d'immunofluorescence indirecte (IFI) est la technique la plus employée pour le diagnostic sérologique (McDade JE, Shepard CC, Fraser DW, Tsai TR, Redus MA et Dowdle WR (1977), *Legionnaires' disease : isolation of a bacterium and demonstration of its role in other respiratory disease, N Engl J Med.* 297 (22) : 1197-203). Environ 70% des diagnostics au niveau européen sont réalisés par cette méthode.

Les tests urinaires sont rapides et très spécifiques. Néanmoins, leur sensibilité est très variable, de 50-90% ; cette forte variabilité est liée non seulement aux caractéristiques du patient mais aussi au fait que les tests actuels permettent seulement la détection de *L. pneumophila* séro groupe 1. De plus, la cinétique d'apparition et de disparition des antigènes de *Legionella*

dans les urines est aussi très variable selon les patients. Un test positif ne peut pas signer une légionellose en phase aiguë et la rechute ou la récurrence ne peuvent pas être évaluées par cette méthode. Enfin, en présence d'une  
5 recherche positive d'antigènes urinaires de légionelles, il reste nécessaire de réaliser en parallèle une culture de prélèvements respiratoires pour l'isolement de la bactérie dans un but épidémiologique.

Le sérodiagnostic est tout particulièrement utile  
10 chez les patients qui ne produisent pas d'expectoration et en cas d'enquête épidémiologique. Il repose sur la mise en évidence d'une augmentation significative du taux des anticorps entre un premier sérum et un sérum prélevé 4 à 6 semaines plus tard. Environ 30% des cas de légionellose  
15 sont identifiés par cette méthode à l'heure actuelle. Néanmoins, les antigènes utilisés dans les tests actuels sont constitués de simples extraits préparés directement à partir de culture de légionelles inactivées par la chaleur ou après ensemencement des bactéries dans le sac vitellin  
20 d'œufs embryonnés. Les anticorps détectés reconnaissent en majorité des déterminants du lipopolysaccharide (LPS), qui peuvent être proches de déterminants du LPS d'autres bactéries et provoquer des réactions croisées. De fait, des réactions de ce type ont été décrites avec d'autres  
25 bactéries à Gram négatif telles que *Pseudomonas sp.*, *Bacteroides sp.*, *Bordetella sp.* et certaines entérobactéries.

En résumé, les pratiques courantes répondent  
imparfaitement aux attentes médicales pour établir le  
30 diagnostic de légionellose. En particulier, les tests sérologiques actuels reposent sur des méthodes de détection anciennes, non automatisables, ou utilisent des antigènes non caractérisés ("sopes antigéniques"), dont la spécificité et la sensibilité sont peu satisfaisantes.  
35 L'utilisation de la sérologie souffre de très grandes variations selon les malades, et les tests basés sur la lecture d'un titre unique précoce possèdent des valeurs

prédictives positives très faible de l'ordre de 10-15% (Jarraud S, Reyrolle M et Etienne J. dans Freney J, Renaud F, Hansen W, Bollet C, Précis de bactériologie clinique, ed. ESKA, Paris, 2000).

Il existe donc un très grand besoin dans le domaine  
5 du diagnostic sérologique afin de réaliser des tests rapides, peu coûteux et non invasifs, et qui pourraient également permettre un diagnostic plus précoce. Enfin, dans l'avenir, en cas de développement d'approches vaccinales visant à prévenir les infections à *Legionella pneumophila*,  
10 de tels tests seront indispensables pour détecter les échecs de vaccination.

Les inventeurs ont ainsi identifié de nouveaux polynucléotides et de nouveaux polypeptides dont les utilisations possibles sont décrites ci-après.

15 La présente invention a comme objectifs additionnels de résoudre les déficiences des tests sérologiques actuels et de permettre une analyse plus précise et plus complète grâce à la détection des différents isotypes d'anticorps spécifiques des polypeptides et, éventuellement, de  
20 diagnostiquer différents sérogroupes et espèces de légionelles.

#### Définitions

Les définitions suivantes sont données afin de faciliter la compréhension de certains termes utilisés dans  
25 cette description.

On entend par "polynucléotide", un polyribonucléotide ou un polydésoxyribonucléotide qui peut être un ADN ou un ARN modifié ou non.

Le terme polynucléotide inclut, sans limitation, un  
30 ADN simple brin ou double brin, un ADN composé d'un mélange d'une ou plusieurs région(s) simple brin et d'une ou plusieurs région(s) double brin, un ADN qui est un mélange de régions simple brin, double brin et/ou triple brin, un ARN simple brin ou double brin, un ARN composé d'un mélange  
35 d'une ou plusieurs région(s) simple brin et d'une ou plusieurs région(s) double brin et les molécules hybrides comprenant un ADN et un ARN qui peuvent comprendre des

régions simple brin, double brin et/ou triple brin ou un mélange de régions simple brin et double brin. Le terme polynucléotide peut aussi comprendre un ARN et/ou un ADN comprenant une ou plusieurs régions triple brin. Les brins dans de telles régions peuvent provenir de la même molécule ou de molécules différentes. Par conséquent, les ADN ou ARN avec des squelettes modifiés pour la stabilité ou d'autres raisons sont inclus dans le terme polynucléotides. On entend également par polynucléotide, les ADN et ARN contenant une ou plusieurs bases modifiées. On entend par base modifiée, par exemple, les bases inhabituelles telles que l'inosine. Le terme polynucléotide vise également les polynucléotides de forme modifiée chimiquement, enzymatiquement ou métaboliquement. Les polynucléotides comprennent également les polynucléotides courts tels que les oligonucléotides.

On entend par "polypeptide", un peptide, un oligopeptide, un oligomère ou une protéine comprenant au moins deux acides aminés joints l'un à l'autre par une liaison peptidique normale ou modifiée.

Le terme polypeptide comprend les chaînes courtes, appelées peptides, oligopeptides et oligomères, et les chaînes longues, appelées protéines.

Un polypeptide peut être composé d'acides aminés autres que les 20 acides aminés codés par les gènes humains. Un polypeptide peut également être composé d'acides aminés modifiés par des processus naturels, tels que le processus de maturation post-traductionnel ou par des procédés chimiques, qui sont bien connus de l'homme du métier. Le même type de modification peut être présent à plusieurs endroits du polypeptide et n'importe où dans le polypeptide : dans le squelette peptidique, dans la chaîne d'acides aminés ou encore aux extrémités carboxy- ou amino-terminales.

Un polypeptide peut être ramifié suite à une ubiquitination ou être cyclique avec ou sans ramification. Ce type de modifications peut être le résultat de processus

de post-traduction naturels ou synthétiques, qui sont bien connus de l'homme du métier.

On entend, par exemple, par modifications d'un polypeptide, l'acétylation, l'acylation, l'ADP-  
5 ribosylation, l'amidation, la fixation covalente de flavine, la fixation covalente d'un hème, la fixation covalente d'un nucléotide ou d'un dérivé nucléotidique, la fixation covalente d'un lipide ou d'un dérivé lipidique, la fixation covalente d'un phosphatidylinositol, la  
10 réticulation covalente ou non-covalente, la cyclisation, la formation de pont disulfure, la déméthylation, la formation de cystéine, la formation de pyroglutamate, la formylation, la gamma-carboxylation, la glycosylation, la formation d'ancre au GPI, l'hydroxylation, l'iodisation, la  
15 méthylation, la myristoylation, l'oxydation, le processus protéolytique, la phosphorylation, la prénylation, la racémisation, la sénéloylation, la sulfatation, l'addition d'acides aminés telles que l'arginylation ou l'ubiquitination (*PROTEINS STRUCTURE AND MOLECULAR PROPERTIES*, 2<sup>nd</sup>  
20 Ed., T.E. Creighton, W.H. Freeman and Company, New York (1993) and Wold, F., *Posttranslational Protein Modifications: Perspectives and Prospects*, pgs 1-12 in *POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS*, B.C. Johnson, Ed., Academic Press, New York (1983); Seifter et al., *Meth. Enzymol.* 182:626-646 (1990) and Rattan et al., *Protein Synthesis: Posttranslational Modifications and Aging*, *Ann. N.Y. Acad. Sci.*  
25 663:48-62 (1992)).

On entend par "pourcentage d'identité" entre deux séquences polynucléotidiques ou polypeptidiques le pourcentage de nucléotides ou d'acides aminés identiques entre les deux séquences à comparer, obtenu après le  
30 meilleur alignement, ce pourcentage étant purement statistique et les différences entre les deux séquences étant réparties au hasard et sur toute leur longueur. Les comparaisons entre deux séquences polynucléotidiques ou polypeptidiques sont traditionnellement réalisées en  
35 comparant ces séquences après les avoir alignées de manière optimale, ladite comparaison étant réalisée par segment ou par "fenêtre de comparaison" pour identifier et comparer

les régions locales de similarité de séquence. Cette comparaison peut être réalisée grâce à un programme, par exemple le programme EMBOSS-Needle (alignement global Needleman-Wunsch) à l'aide de la matrice BLOSUM62/Open Gap  
5 10.0 et Extension Penalty de 0,5 (*Needleman et Wunsch (1970). J. Mol. Biol. 48, 443-453 et Kruskal, J.B. (1983), An overview of sequence comparison, In D. Sankoff and J.B. Kruskal, (ed), Time warps, strind edits and macromolecules : the theory and practice of sequence comparison, pp. 1-44 Addison Wesley*).

Le pourcentage d'identité est calculé en déterminant  
10 le nombre de positions identiques pour lesquelles le nucléotide ou l'acide aminé est identique entre les deux séquences, en divisant ce nombre de positions identiques par le nombre total de positions dans la fenêtre de comparaison et en multipliant le résultat obtenu par 100.

15 Un premier polynucléotide ayant, par exemple, une identité d'au moins 95% avec un second polynucléotide est donc un polynucléotide comportant, au plus, 5 nucléotides modifiés sur 100 nucléotides, par rapport à la séquence dudit second polynucléotide. En d'autres termes jusqu'à 5%  
20 des nucléotides dudit second polynucléotide peuvent être délétés ou substitués par un autre nucléotide, ou ledit premier polynucléotide peut comporter jusqu'à 5% de nucléotides en plus par rapport au nombre total de nucléotides du second polynucléotide. Ces modifications  
25 peuvent être positionnées aux extrémités 3' et/ou 5', ou à un endroit quelconque entre ces extrémités, en un seul ou plusieurs emplacements (*Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D.W., ed., Academic Press, New York, 1993; Computer Analysis of Sequence Data, Part I, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987*).

Par analogie, un polypeptide ayant, par exemple, une identité d'au moins 95% avec un second polypeptide est un polypeptide comportant, au plus, 5 acides aminés modifiés  
35 sur 100 acides aminés, par rapport à la séquence dudit second polypeptide. En d'autres termes jusqu'à 5% des acides aminés dudit second polynucléotide peuvent être

délétés ou substitués par d'autres acides aminés, ou ledit premier polypeptide peut comporter jusqu'à 5% d'acides aminés en plus par rapport au nombre total d'acides aminés du second polypeptide. Ces modifications peuvent être positionnées aux extrémités amino et/ou carboxy-terminales, ou à un endroit quelconque entre ces positions terminales, en un seul ou plusieurs emplacements (*Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M., and Griffin, H.G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987).

On entend par "cellule hôte", une cellule qui a été transformée ou transfectée, ou est capable de transformation ou de transfection, par une séquence polynucléotidique exogène.

On entend par "milieu de culture", le milieu dans lequel on purifie le polypeptide de l'invention. Ce milieu peut être constitué par le milieu extracellulaire et/ou le lysat cellulaire. Des techniques bien connues de l'homme du métier permettent également à ce dernier de redonner la conformation active au polypeptide, si la conformation dudit polypeptide a été altérée lors de l'isolation ou de la purification.

On entend par "fonction", l'activité biologique d'un polypeptide ou d'un polynucléotide.

La fonction d'un polypeptide conforme à l'invention est celle d'un antigène de *Legionella pneumophila*, et la fonction d'un polynucléotide conforme à l'invention est celle de coder ce polypeptide. La fonction, selon l'invention, d'une combinaison d'antigènes est de permettre d'établir un diagnostic d'une infection à *L. pneumophila*, mais aussi de réaliser des suivis thérapeutiques et, enfin, de préciser s'il s'agit d'une infection active ou aiguë, par exemple, par l'utilisation de la détection de différents isotypes d'anticorps.

On entend par "antigène", tout composé qui, seul ou en association avec un adjuvant ou porteur, est capable

d'induire une réponse immunitaire spécifique. Cette définition comprend également tout composé, présentant une analogie structurale avec ledit antigène, capable d'induire une réponse immunologique dirigée contre ledit antigène.

5 On entend par "analogie structurale" une analogie aussi bien de la structure primaire (séquence) que de la structure secondaire (éléments structuraux), ou que de la structure tertiaire (structure tridimensionnelle) ou que de la structure quaternaire (association de plusieurs  
10 polypeptides dans un même complexe)(*BIOCHEMISTRY, 4<sup>ème</sup> Ed, L. Stryer, New York, 1995*).

On entend par "variant" d'un polynucléotide dit initial ou d'un polypeptide dit initial, respectivement, un polynucléotide ou un polypeptide qui en diffère par au  
15 moins un nucléotide ou un acide aminé, mais qui garde les mêmes propriétés intrinsèques, c'est-à-dire la même fonction.

Une différence dans la séquence polynucléotidique du variant peut altérer ou non la séquence d'acides aminés du  
20 polypeptide qu'il code, par rapport à un polypeptide initial. Néanmoins, par définition, ces variants doivent conférer la même fonction que la séquence polynucléotidique initiale, par exemple, coder un polypeptide ayant une fonction antigénique.

25 Le polynucléotide ou le polypeptide variant diffère généralement du polynucléotide initial ou du polypeptide initial par une (ou plusieurs) substitution(s), addition(s), délétion(s), fusion(s) ou troncation(s) ou plusieurs de ces modifications, prises en combinaison. Un  
30 variant non-naturel d'un polynucléotide initial ou d'un polypeptide initial peut être obtenu, par exemple, par mutagenèse dirigée ou par synthèse directe.

On définit comme "séquence polynucléotidique complémentaire à la séquence polynucléotidique", un  
35 polynucléotide qui peut être hybridé avec cette séquence polynucléotidique dans des conditions stringentes.

On entend généralement, mais pas nécessairement, par "conditions stringentes" les conditions chimiques qui permettent une hybridation lorsque les séquences polynucléotidiques ont une identité d'au moins 80%.

5 Ces conditions peuvent être obtenues selon les méthodes bien connues de l'homme du métier.

On entend par "anticorps", les anticorps monoclonaux, polyclonaux, chimériques, simple chaîne ou humanisés, ainsi que les fragments Fab, incluant les produits d'un Fab ou  
10 d'une banque d'expression d'immunoglobulines. Il est aussi possible d'utiliser, à la place des anticorps, d'autres molécules immunospécifiques, par exemple des récepteurs de cellules T (TCR) comme récemment décrit (*Li et al. 2005. Directed evolution of human T-cell receptors with picomolar affinities by phage display. Nat Biotechnol.*  
15 *23:349-54*) ou des molécules sélectionnées pour leur fixation spécifique, par exemple par évolution dirigée (voir, par exemple, *Conrad et Scheller. 2005. Considerations on antibody-phage display methodology. Comb Chem High Throughput Screen. 8:117-26*).

Un anticorps immunospécifique peut être obtenu par  
20 administration à un animal d'un polypeptide donné suivie d'une récupération des anticorps produits par ledit animal par extraction depuis ses fluides corporels. On peut également administrer à l'animal un variant dudit polypeptide, ou des cellules hôtes exprimant ce  
25 polypeptide.

Le terme "immunospécifique" appliqué au terme anticorps, vis-à-vis d'un polypeptide donné, signifie que l'anticorps possède une meilleure affinité pour ce polypeptide que pour d'autres polypeptides connus de l'art  
30 antérieur.

Par sérum "positif", on entend un sérum contenant des anticorps, produits à la suite d'une infection à *L. pneumophila*, identifiés par leur liaison avec les polypeptides (antigènes) de l'invention.

35 On entend par "sensibilité" la proportion de malades infectés, identifiés selon l'art antérieur et donnés positifs par le diagnostic selon l'invention.

On entend par "spécificité" la proportion de donneurs de sang, testés comme témoins, soumis au diagnostic selon l'invention et donnés négatifs par le diagnostic selon l'invention.

5 On entend par "amorces spécifiques" de courtes séquences nucléotidiques capables de s'hybrider de façon spécifique, grâce à la complémentarité des bases, sur le brin d'ADN ou sur son brin complémentaire ; ces amorces sont choisies par l'homme de l'art de façon à encadrer la  
10 séquence d'ADN et à l'amplifier spécifiquement.

On entend par "fixation spécifique" une interaction entre deux molécules ayant une affinité de l'ordre de  $10^6$  M,  $5.10^6$  M,  $10^7$  M,  $5.10^7$  M,  $10^8$  M,  $5.10^8$  M,  $10^9$  M ou supérieure.

La présente invention répond aux objectifs qu'elle  
15 s'est fixés plus haut en apportant de nouveaux polynucléotides et de nouveaux polypeptides.

L'invention a pour objet l'utilisation, *in vitro*, de l'une au moins des protéines de séquence ID SEQ N°4, 6, 8, 10, 12, 14 ou 16, d'une partie ou de variants de ces  
20 protéines, de cellules hôtes comprenant des vecteurs incluant un polynucléotide codant pour l'une au moins desdites protéines ou un variant desdites protéines, dans la production d'anticorps et le domaine du diagnostic *in vitro* de *L. pneumophila*. Selon une forme d'exécution  
25 particulière de l'invention, celle-ci a pour objet l'utilisation, *in vitro*, de l'une au moins des protéines de séquence ID SEQ N°4, 6, 8, 10, 12, 14 ou 16, en combinaison avec la protéine de séquence ID SEQ N°2, dans la production d'anticorps et le domaine du diagnostic *in vitro* de  
30 *L. pneumophila*. L'invention a également pour objet un kit de diagnostic et une composition pharmaceutique.

#### Utilisation des polypeptides

Certains polypeptides issus de *L. pneumophila* ne permettent pas de diagnostiquer une infection à cette  
35 bactérie, par exemple les protéines Pal ou FlaA de *Legionella pneumophila*. La Déposante a cependant identifié, de façon tout à fait inattendue, qu'un tel diagnostic était

possible en utilisant d'autres polypeptides issus de *L. pneumophila* qu'elle a spécifiquement identifiés et, même, des fragments de polypeptides issus de *L. pneumophila*.

5 L'identification des polypeptides selon l'invention est le résultat d'un crible et d'études approfondies que les séquences issues des programmes de génomique de *L. pneumophila* ne laissaient pas envisager.

Ainsi, la présente invention a pour objet  
10 l'utilisation, dans la production d'anticorps et dans le domaine du diagnostic *in vitro* d'infections à *L. pneumophila*, d'au moins un polypeptide comprenant :

- la séquence d'acides aminés ID SEQ N°4 (appelée protéine G5), codée par la séquence polynucléotidique  
15 ID SEQ N°3 ; ou

- la séquence d'acides aminés ID SEQ N°6 (appelée protéine 2E1), codée par la séquence polynucléotidique ID SEQ N°5 ; ou

- la séquence d'acides aminés ID SEQ N°8 (appelée  
20 protéine 4A6), codée par la séquence polynucléotidique ID SEQ N°7 ; ou

- la séquence d'acides aminés ID SEQ N°10 (appelée protéine 4B6), codée par la séquence polynucléotidique ID SEQ N°9 ; ou

25 - la séquence d'acides aminés ID SEQ N°12 (appelée protéine 4C6), codée par la séquence polynucléotidique ID SEQ N°11 ; ou

- la séquence d'acides aminés ID SEQ N°14 (appelée  
30 protéine 4D6), codée par la séquence polynucléotidique ID SEQ N°13 ; ou

- la séquence d'acides aminés ID SEQ N°16 (appelée protéine 4E6), codée par la séquence polynucléotidique ID SEQ N°15.

La présente invention a également pour objet  
35 l'utilisation, dans la production d'anticorps et dans le domaine du diagnostic *in vitro* d'infections à *L. pneumophila*, d'au moins un polypeptide, tel que défini

ci-dessus, en combinaison avec le polypeptide comprenant la séquence d'acides aminés ID SEQ N°2 (appelée 2A1), codée par la séquence polynucléotidique ID SEQ N°1.

La présente invention vise également l'utilisation  
5 d'au moins un polypeptide comprenant :

a) une partie de la séquence d'acides aminés ID SEQ N°4, 6, 8, 10, 12, 14 ou 16 ayant la même fonction que ladite séquence, ou

b) une séquence d'acides aminés ayant au moins 60%  
10 d'identité, de préférence au moins 80% d'identité, et mieux au moins 90% d'identité, avec l'une des séquences d'acides aminés ID SEQ N° ID SEQ N°4, 6, 8, 10, 12, 14 ou 16, , ou avec la partie de séquence définie sous a), et ayant la même fonction que ladite séquence.

15 Ainsi, des polypeptides conformes à l'invention peuvent comprendre des variants ou des fragments des séquences d'acides aminés ID SEQ N°2, 4, 6, 8, 10, 12, 14 ou 16.

L'invention vise aussi l'utilisation d'au moins un  
20 polypeptide, tel que défini ci-dessus, en combinaison avec un polypeptide comprenant :

a) une partie de la séquence d'acides aminés ID SEQ N°2, et ayant la même fonction que ladite séquence, ou

25 b) une séquence d'acides aminés ayant au moins 60% d'identité, de préférence au moins 80% d'identité, et mieux au moins 90% d'identité, avec la séquence d'acides aminés ID SEQ N°2 ou avec la partie de séquence définie sous a), et ayant la même fonction que ladite séquence.

30 Les échantillons biologiques testés peuvent être du sang, de l'urine, de la salive, du liquide de ponction de sérologie (par exemple liquide céphalo-rachidien, liquide pleural ou liquide articulaire) ou l'un de leurs constituants (par exemple, le sérum).

35 Le diagnostic *in vitro* d'une infection à *L. pneumophila* est effectué, par exemple, par des tests classiques de réactions immunologiques, tels que ELISA ou

Western Blot. Ces tests utilisent l'un au moins des polypeptides selon l'invention qui vient se fixer, de manière spécifique, aux anticorps sériques contre *L. pneumophila* présents dans les échantillons biologiques.

5 Grâce aux polypeptides définis ci-dessus, l'invention permet la détection d'anticorps naturellement produits lors d'infections à *L. pneumophila*.

L'invention vise, en outre, l'utilisation des polypeptides selon l'invention, pour détecter, 10 périodiquement, *in vitro*, les anticorps dirigés contre *L. pneumophila* et ainsi suivre l'évolution de la pathologie et de l'effet d'un traitement appliqué à un patient.

#### Utilisation de vecteurs d'expression et de cellules hôtes

15 La présente invention a aussi pour objet l'utilisation d'au moins un polypeptide, selon l'invention, préparé par culture d'une cellule hôte comprenant un vecteur recombinant ayant, inséré, un polynucléotide codant pour ledit polypeptide.

20 De nombreux systèmes d'expression peuvent être utilisés comme, par exemple, les chromosomes, les épisomes, les virus dérivés. Plus particulièrement, les vecteurs recombinants utilisés peuvent être dérivés de plasmides bactériens, de transposons, d'épisome de levure, d'éléments 25 d'insertion, d'éléments chromosomiques de levures, de virus tels que les baculovirus, les *papillonna virus* comme SV40, les *vaccinia virus*, les adénovirus, les *fox pox virus*, les *pseudorabies virus*, les rétrovirus.

Ces vecteurs recombinants peuvent également être des 30 dérivés de cosmides ou de phagemides. La séquence polynucléotidique peut être insérée dans le vecteur recombinant d'expression par les méthodes bien connues de l'homme du métier.

Le vecteur recombinant peut comprendre des séquences 35 polynucléotidiques de contrôle de la régulation de l'expression du polynucléotide ainsi que des séquences polynucléotidiques permettant l'expression et la

transcription d'un polynucléotide selon l'invention et la traduction d'un polypeptide selon l'invention, ces séquences étant choisies en fonction des cellules hôtes mises en oeuvre.

5 L'introduction du vecteur recombinant dans une cellule hôte peut être effectuée selon les méthodes bien connues de l'homme du métier telles que la transfection par phosphate de calcium, la transfection par lipides cationiques, l'électroporation, la transduction ou  
10 l'infection.

Les cellules hôtes peuvent être, par exemple, des cellules bactériennes, telles que des cellules de streptocoques, de staphylocoques, d'*Escherichia coli* ou de *Bacillus subtilis* ; des cellules de champignons, telles que  
15 des cellules de levure et des cellules d'*Aspergillus* ; des cellules de *Streptomyces* ; des cellules d'insectes, telles que des cellules de *Drosophila S2* et de *Spodoptera Sf9* ; des cellules animales, telles que les cellules CHO, COS, HeLa, C127, BHK, HEK 293 ou encore des cellules végétales.

20 Le polypeptide peut être purifié à partir des cellules hôtes, selon les méthodes bien connues de l'homme du métier, telles que la précipitation à l'aide d'agents chaotropiques comme les sels, en particulier le sulfate d'ammonium, l'éthanol, l'acétone ou l'acide  
25 trichloroacétique, ou de moyens tels que l'extraction à l'acide, la chromatographie échangeuse d'ions, la chromatographie sur phosphocellulose, la chromatographie par interaction hydrophobe, la chromatographie d'affinité, la chromatographie sur hydroxylapatite ou les  
30 chromatographies d'exclusion.

Utilisation des polynucléotides

La présente invention a également pour objet l'utilisation, dans la production d'anticorps et dans le diagnostic d'une infection à *L. pneumophila*, d'au moins un polynucléotide comprenant la séquence polynucléotidique ID SEQ N°3, 5, 7, 9 11, 13 ou 15, codant, respectivement, pour la protéine ID SEQ N°4, 6, 8, 10, 12, 14 ou 16, ou d'une combinaison d'au moins l'un de ces polynucléotides en combinaison avec le polynucléotide de séquence ID SEQ N°1, codant pour la protéine de séquence ID SEQ N°2.

L'invention vise également l'utilisation d'au moins un polynucléotide comprenant :

- a) une partie de la séquence ID SEQ N°3, 5, 7, 9 11, 13 ou 15 et ayant la même fonction que ladite séquence, ou
- b) une séquence polynucléotidique ayant au moins 60% d'identité, de préférence au moins 80% d'identité, et mieux au moins 90% d'identité, avec la séquence polynucléotidique ID SEQ N°3, 5, 7, 9 11, 13 ou 15, ou avec la partie de séquence telle que définie sous a), et ayant la même fonction que ladite séquence, ou
- c) une séquence polynucléotidique complémentaire à la séquence polynucléotidique ID SEQ N°3, 5, 7, 9 11, 13 ou 15, ou à la partie de séquence définie sous a) ou à la séquence définie sous b).

Ainsi, des polynucléotides conformes à l'invention peuvent comprendre des variants ou des fragments de l'une des séquences polynucléotidiques ID SEQ N°1, 3, 5, 7, 9 11, 13 ou 15.

L'invention vise également l'utilisation d'au moins un polynucléotide, tel que défini ci-dessus, en combinaison avec un polynucléotide comprenant :

- a) une partie de la séquence ID SEQ N°1 et ayant la même fonction que ladite séquence, ou
- b) une séquence polynucléotidique ayant au moins 60% d'identité, de préférence au moins 80% d'identité, et mieux au moins 90% d'identité, avec la séquence polynucléotidique

ID SEQ N°1 ou avec la partie de séquence telle que définie sous a), et ayant la même fonction que ladite séquence, ou  
c) une séquence polynucléotidique complémentaire à la séquence polynucléotidique ID SEQ N°1 ou à la partie de  
5 séquence définie sous a) ou à la séquence définie sous b).

Les polynucléotides de l'invention peuvent être obtenus par les méthodes standard de synthèse d'ADN ou d'ARN.

10 Les polynucléotides conformes à l'invention peuvent également comprendre des séquences polynucléotidiques comme les séquences 5' et/ou 3' non-codantes, telles que, par exemple, des séquences transcrites, des séquences non-traduites, des séquences signal d'épissage, des séquences  
15 polyadénylées, des séquences de liaison avec des ribosomes ou encore des séquences qui stabilisent l'ARNm.

#### Utilisations des anticorps selon l'invention

L'invention vise également l'utilisation d'anticorps, selon l'invention, pour détecter, *in vitro*, dans des  
20 échantillons biologiques la présence d'antigènes de *L. pneumophila*.

L'invention vise, en outre, l'utilisation d'anticorps selon l'invention, pour détecter, périodiquement, *in vitro*, les antigènes de *L. pneumophila* et ainsi suivre l'évolution  
25 de la pathologie et de l'effet d'un traitement appliqué à un patient.

Les anticorps immunospécifiques peuvent être obtenus par administration d'un polypeptide selon l'invention, d'un de ses fragments, d'un analogue ou d'un fragment épitopique  
30 ou d'une cellule exprimant ce polypeptide, à un mammifère, de préférence non humain, selon les méthodes bien connues de l'homme du métier.

Pour la préparation d'anticorps monoclonaux, on peut utiliser des méthodes usuelles de production d'anticorps, à  
35 partir de lignées cellulaires, telles que la technique des hybridomes, la technique des triomes, la technique des

hybridomes de cellules B humaines et la technique des hybridomes EBV.

D'autres molécules, notamment protéines, suffisamment spécifiques peuvent remplacer les anticorps dans ces utilisations. Ces molécules sont, par exemple, d'autres molécules immunospécifiques comme les récepteurs de cellules T (TCR) (*Li et al. 2005. Directed evolution of human T-cell receptors with picomolar affinities by phage display. Nat Biotechnol. 23:349-54*). Ces molécules peuvent également être des protéines sélectionnées pour leur fixation spécifique des polypeptides selon l'invention, par exemple, par évolution dirigée (voir par exemple *Conrad et Scheller. 2005. Considerations on antibody-phage display methodology. Comb Chem High Throughput Screen. 8:117-26*).

#### Isotypes d'anticorps et affinité

Les immunoglobulines (ou anticorps) sont constituées de deux chaînes polypeptidiques différentes : deux chaînes légères (L) d'isotypes [kappa] ou [lambda], et deux chaînes lourdes (H) d'isotypes [gamma], [alpha], [mu], [delta] ou [epsilon]. Ces quatre chaînes sont liées covalamment par des ponts disulfures. Les deux types de chaînes H ou L possèdent des régions qui contribuent à la fixation des antigènes et sont très variables d'une immunoglobuline à une autre. Ces régions déterminent l'affinité des anticorps pour leur ligand.

Les isotypes des chaînes lourdes permettent de définir la classe de l'immunoglobuline ; il existe donc 5 isotypes d'anticorps (IgA, IgM, IgD, IgE et IgG). Des sous-classes, par exemple IgG1, IgG2, IgG3 et IgG4, sont définies par des différences de séquence de la chaîne lourde. Les différentes classes d'immunoglobulines ont des propriétés physicochimiques propres et leur synthèse dépend directement des phases et des niveaux d'activation de la réponse immunitaire.

Chez l'homme, les IgG représentent la principale classe d'immunoglobulines : leur concentration sérique chez l'adulte varie de 8 à 16 g/l. Leur demi-vie plasmatique est d'environ 3 semaines.

Les IgA constituent la deuxième classe d'immunoglobulines sériques, après les IgG, en terme de concentration (2 à 4 g/l). En revanche, elles représentent la classe prépondérante des immunoglobulines dans les  
5 sécrétions (sécrétions respiratoires, salivaires, digestives..., lait, colostrum, larmes). Sur le plan structural, les IgA ont la particularité de se présenter sous plusieurs formes moléculaires :

- dans le sérum, les IgA peuvent se présenter sous  
10 forme de monomères (forme prépondérante) ou sous forme de dimères associés à une chaîne J (chaîne de jonction). La chaîne J est un peptide riche en cystéine de 137 acides aminés (poids moléculaire : 15 000 Da), d'origine plasmocytaire ; la sous-classe IgA1 est prépondérante dans  
15 le sérum.

- dans les sécrétions, les IgA, appelées IgA sécrétoires, sont sous forme de dimères ; elles sont donc associées à une chaîne J, mais elles comportent également un composant sécrétoire. Les IgA produites par les  
20 lymphocytes B sont captées par les cellules épithéliales, par un récepteur (polyIgR) situé au pôle basal de la cellule. C'est pendant le transfert de l'IgA à travers la cellule épithéliale, vers le pôle apical de la cellule, que le composant sécrétoire (poids moléculaire : 70 000 Da), ou  
25 pièce sécrétoire, est ajouté. Le rôle de la pièce sécrétoire est de protéger les IgA sécrétoires vis-à-vis des enzymes protéolytiques présents dans les sécrétions.

Comme les IgA, les IgM peuvent se présenter sous  
2 formes moléculaires distinctes :

30 - une forme monomérique : c'est la forme sous laquelle les IgM sont synthétisées et insérées dans la membrane des lymphocytes B ;

- une forme pentamérique : c'est la forme sous laquelle les IgM sont sécrétées. Les 5 monomères de bases  
35 sont reliés par des ponts disulfures. De plus, une chaîne J relie l'extrémité de 2 monomères.

Les IgD représentent moins de 1 % des immunoglobulines sériques. Elles sont habituellement coexprimées avec les IgM à la surface des lymphocytes B où elles semblent jouer un rôle de récepteur pour les antigènes.

5 Les IgE ne sont présentes, chez un sujet normal, qu'à l'état de traces.

#### Cinétique d'apparition des anticorps et affinité

La cinétique d'apparition d'anticorps spécifiques et d'un isotype donné est aujourd'hui encore mal comprise.  
10 D'une façon générale les cellules B, dites naïves, synthétisent différents IgM qui constituent un large répertoire de molécules capables de fixer des antigènes (Steven A. Frank, *Immunology and Evolution of Infectious Disease*. (2002), Princeton University Press, Princeton, USA). Ainsi, lors de la première exposition à un  
15 antigène, celui-ci va se fixer avec une faible affinité à différents IgM du système immunitaire. Cette interaction va néanmoins stimuler la division de la cellule B naïve correspondante. La division est d'autant plus rapide que l'affinité de l'IgM est forte, ce qui permet une sélection  
20 initiale des meilleurs clones. De plus, lors de la division, les réarrangements génétiques permettent l'augmentation de l'affinité des anticorps. La région constante des immunoglobulines change également afin de produire des IgG dans le système circulatoire et des IgA à  
25 la surface des muqueuses (Steven et Frank 2002. *Immunology and Evolution of Infectious Disease*. Princeton University Press, Princeton, USA).

#### Utilisations des isotypes d'anticorps selon l'invention

L'homme de l'art peut donc déduire de ce qui est  
30 explicité ci-dessus que la présence d'une ou plusieurs classes d'anticorps et l'affinité particulière des anticorps ont différentes implications d'un point de vue du diagnostic médical dans la différenciation des infections récurrentes ou non, d'une infection active, aiguë,  
35 chronique, latente, et/ou récente.

### Les kits

L'invention a également pour objet des kits de diagnostic *in vitro* comprenant au moins l'un des polypeptides conformes à l'invention, ou au moins l'un des polynucléotides codant pour lesdits polypeptides, ainsi que des kits de diagnostic *in vitro* comprenant au moins l'un des anticorps conformes à l'invention.

### Les vaccins

La présente invention concerne également une composition pharmaceutique, utilisable comme vaccin, renfermant, à titre de principe actif, au moins un polypeptide selon l'invention ou un polynucléotide ou un vecteur recombinant ou une cellule hôte selon l'invention.

## 15 Partie expérimentale

### A) Protocoles d'obtention des antigènes

#### A.1. Clonage de la séquence codant pour les polypeptides ID SEQ N°2, 4, 6, 8, 10, 12, 14 ou 16.

Les gènes codant pour les séquences des polypeptides, qui sont des antigènes, sont obtenus par amplification par PCR à partir de l'ADN génomique de la bactérie *Legionella pneumophila* (souche Philadelphia-1, ATCC 33152), en utilisant, respectivement, des amorces spécifiques des séquences ID SEQ N°1, 3, 5, 7, 9 11, 13 ou 15.

Le fragment correspondant, ainsi amplifié, est cloné dans un vecteur selon des techniques classiques bien connues de l'homme du métier. Ce vecteur permet la production des protéines clonées sous contrôle d'un promoteur inductible par l'isopropyl-thiogalactoside (IPTG). Les protéines clonées correspondent aux polypeptides ID SEQ N°2, 4, 6, 8, 10, 12, 14 ou 16.

#### A.2. Expression des protéines

Une souche d'*Escherichia coli* est transformée par les vecteurs d'expression précédemment décrits. Les bactéries sélectionnées sont mises à cultiver une nuit à 30°C sous agitation dans 30 ml de milieu Luria Bertani (LB, J. Miller, "A short Course in Bacterial Genetics", Cold Spring Harbor Laboratory Press, 1992) contenant

de l'ampicilline à une concentration finale de 100 µg/ml. Le lendemain, la culture est diluée au 1/50<sup>ème</sup> dans un volume final de 1 litre de milieu LB additionné d'ampicilline à une concentration finale de 100 µg/ml qui  
5 est incubée à 30°C sous agitation. Lorsque la turbidité de la culture atteint une valeur d'absorbance à 600 nm (en abrégé, A600) d'environ 0,7, la production de la protéine est induite par l'isopropyl-thiogalactoside (en abrégé, IPTG) à une concentration finale de 0,1 mM. Les bactéries  
10 sont récoltées par centrifugation (10 minutes à 1400xg et à 4°C) lorsque la turbidité de la culture atteint une A600 d'environ 1,5.

#### A.3. Purification des protéines

Après centrifugation, les cellules sont remises en  
15 suspension dans un tampon Tris-HCl 20 mM à pH 8,0 contenant du saccharose à 0,5 mM, puis traitées par du lysozyme (0,2 g/l) en présence d'acide éthylènediaminotétraacétique (EDTA) 12,5 mM, de DNase, RNase et PMSF. La suspension est incubée 30 minutes à 4°C puis centrifugée 10 minutes à 4°C  
20 à 15500xg. Le culot est congelé à -20°C pendant au moins une nuit.

#### A.4. Exemple : purification de 2E1 (ID SEQ N°6)

Les protocoles de purification diffèrent pour chaque protéine en fonction de sa solubilité et de ses caractères  
25 physicochimiques propres. La purification de la protéine 2E1 est détaillée ci-après, à titre d'exemple. Après décongélation, les bactéries sont reprises dans un tampon Mes [acide 2-(N-morpholino) éthane sulfonique] 25 mM à pH 6,0, puis soniquées 4 fois 20 secondes dans la glace. Après  
30 centrifugation à 15500xg à 4°C pendant 30 minutes, le surnageant est filtré sur membrane de porosité 0,22 µm. Le filtrat est alors déposé sur une colonne échangeuse de cation (par exemple, SP-Sépharose 12 ml, Amersham Biosciences). Après lavage de la colonne, la protéine est  
35 éluée par un gradient linéaire de 0 à 1 M de NaCl dans du tampon Mes 25 mM à pH 6,0 en 20 volumes de colonne. Les fractions contenant la protéine sont rassemblées, dialysées

sur la nuit contre du tampon Mes 25 mM à pH 6,0 puis redéposées sur cette même colonne. La protéine est ensuite éluée par un nouveau gradient de NaCl optimisé pour la protéine 2E1. De nouveau, les fractions contenant la protéine sont rassemblées et les protéines sont précipitées par du sulfate d'ammonium à une concentration finale de 0,6 g/l. La solution est laissée au moins une nuit à 4°C puis centrifugée 30 minutes à 20800xg. Le culot est alors repris dans un volume le plus faible possible (en général 300 µl de tampon Na<sub>2</sub>HPO<sub>4</sub>/NaH<sub>2</sub>PO<sub>4</sub> 50 mM pH 8,0 contenant du NaCl 100 mM puis déposé sur une colonne de gel filtration, par exemple Superdex HR75-10/30, Amersham). Les fractions éluées contenant la protéine sont rassemblées et du glycérol est ajouté jusqu'à une concentration finale de 20%. La protéine purifiée est alors stockée à -20°C jusqu'à son utilisation dans les tests.

Les concentrations en protéine sont déterminées spectrophotométriquement à partir des coefficients d'absorption calculés par la méthode de Pace (*Pace et al. 1995. How to measure and predict the molar absorption coefficient of a protein. Protein Science 4, 2411-2423*). La pureté de la protéine est vérifiée par analyse par électrophorèse SDS-PAGE et par spectrométrie de masse.

#### B) Test de diagnostic *in vitro*

Il a été utilisé des sérums provenant de patients ayant eu une infection documentée à *Legionella pneumophila* (collection du laboratoire). L'infection a pu être mise en évidence soit par l'isolement/culture des bactéries à partir de prélèvements broncho-pulmonaires, soit par une séroconversion, ou encore par un test urinaire positif.

Les sérums témoins correspondaient à des sérums de donneurs de sang (collection du laboratoire).

La fixation, sur les protéines recombinantes purifiées (obtenue comme décrit précédemment), des anticorps présents dans les sérums peut être évaluée soit par des tests en Western Blot soit par la technique ELISA.

Exemple B1 : Protocole de test pour les polypeptides selon l'invention en ELISA

La fixation des anticorps présents dans les sérums a été évaluée par des tests ELISA. Les plaques ELISA sont  
5 laissées pendant une nuit à 4°C en présence de 0,5 µg d'antigène purifié dans du "tampon salin phosphate" (PBS). Après quatre lavages par du PBS contenant du polyoxyéthylène sorbitan (Tween) à 0,05%, les plaques sont saturées une heure à 37°C par du PBS-Tween contenant 5% de  
10 lait demi-écrémé (250 µl par puits). Quatre nouveaux lavages sont réalisés, puis 100 µl de chaque sérum positif, à la dilution adéquate dans du tampon PBS-Tween contenant 5% de lait demi-écrémé, sont ajoutés dans chaque puits. La plaque est ensuite laissée à 25°C pendant 30 minutes. Après  
15 quatre nouveaux lavages, des anticorps (secondaires) anti-immunoglobulines G, M, ou A ou simultanément anti-immunoglobulines G et/ou A et/ou M, humaines de chèvre marqués à la phosphatase alcaline (par exemple 170-6462, Biorad) sont ajoutés pendant 30 minutes à 25°C après avoir  
20 été dilués, selon le protocole du fournisseur, dans du tampon PBS-Tween contenant 5% de lait demi-écrémé. Quatre nouveaux lavages sont réalisés puis 100 µl de substrat (phosphate de p-nitrophényle), pNPP, par exemple A-3469, Sigma) sont ajoutés. L'absorbance à 405 nm de chacun des  
25 puits est mesurée après une incubation de 30 minutes à 37°C.

Sont considérés comme "positifs" en ELISA, les sérums identifiés par leur liaison avec les polypeptides (antigènes) de l'invention.

30 Exemple B2 : Protocole de test pour les polypeptides selon l'invention en Western Blot

La fixation éventuelle des anticorps présents dans les sérums peut également être évaluée par des tests en Western Blot, par exemple, sur des extraits totaux de  
35 bactéries produisant les polypeptides ou sur les protéines purifiées.

Pour cela, la membrane de nitrocellulose sur laquelle les protéines de l'extrait sont transférées est saturée pendant 30 minutes à l'aide d'une solution de "tampon salin phosphate" (PBS) contenant 3% de lait demi-écrémé. Après  
5 trois lavages par du PBS contenant du polyoxyéthylène sorbitan (Tween) à 0,05%, la membrane est mise en présence du sérum testé à la dilution adéquate dans du tampon PBS contenant 3% de lait demi-écrémé, pendant 45 minutes. Après  
10 trois nouveaux lavages, des anticorps (secondaires) anti-immunoglobulines G, M ou A ou simultanément anti-immunoglobulines G et/ou A et/ou M humaines de chèvre, marqués à la phosphatase alcaline (par exemple 170-6462, Biorad), dilués, selon le protocole du fournisseur, dans du tampon PBS contenant 3% de lait demi-écrémé, sont ajoutés  
15 puis laissés pendant 30 minutes. Trois nouveaux lavages sont réalisés puis du 5-bromo-4-chloro-3-indolyl-phosphate et du nitrobleu tétrazolium sont ajoutés selon les indications du fournisseur jusqu'à l'obtention du résultat. Un résultat "positif" correspond à une précipitation du  
20 substrat sur la membrane à la position du polypeptide étudié.

### C) Résultats et interprétation

Des résultats types sont présentés dans les différents exemples ci-après.

25 Exemple C1 : Résultats (ELISA) obtenus en utilisant des polypeptides selon l'invention pour la détection d'anticorps spécifiques dans les sérums

Le tableau 1 présente les résultats obtenus selon l'invention pour les polypeptides 2E1 (ID SEQ N°6) et 2A1  
30 (ID SEQ N°2) avec des anticorps secondaires reconnaissant les immunoglobulines A présentes dans des sérums.

Tableau 1 : Résultats (ELISA) obtenus en utilisant des anti-IgA comme anticorps secondaires

	Polypeptide testé	
	2E1	2A1
Proportion de sérums "positifs" parmi 49 sérums de malades infectés par <i>L. pneumophila</i> diagnostiqués selon l'art antérieur et soumis au diagnostic selon l'invention	51,0%	45,0%
Proportion de sérums "négatifs" parmi 96 sérums de donneurs de sang testés comme témoins	97,0 %	98,0%

5 D'après le tableau, on constate que, par le test ELISA, on identifie *in vitro* des infections à *Legionella pneumophila*, grâce aux antigènes selon l'invention. Il est donc démontré, d'une part, l'existence chez l'homme d'une réponse anticorps significative (la probabilité associée à un test de  $\chi^2$  est inférieure à 0,05) vis-à-vis de la protéine 2E1 ou 2A1 au cours des infections à *Legionella pneumophila* et, d'autre part, que les protéines 2E1 ou 2A1 sont pertinents pour le diagnostic sérologique de ce type d'infection. Les résultats exposés ci-dessus, à propos des protéines 2E1 et 2A1 sont également valables pour les autres polypeptides selon l'invention.

De plus l'utilisation des polypeptides selon l'invention permet, dans certains cas, l'identification de malades non détectés à ce stade de l'infection (mais confirmés dans des analyses ultérieures) par les techniques de sérologie de l'art antérieur.

Exemple C2 : Résultats (ELISA) en utilisant des polypeptides selon l'invention pour la détection d'isotype d'anticorps dans les sérums

L'utilisation de certains polypeptides, par exemple  
 5 2A1 (ID SEQ N°2) (tableau 2), pour la détection d'immunoglobulines d'un isotype particulier est particulièrement pertinente.

Tableau 2 : Résultats (ELISA) obtenus en utilisant  
 10 2A1 (ID SEQ N°2) comme polypeptide et en détectant les anticorps grâce à différents anticorps secondaires

Anticorps secondaire utilisés	IgA	IgG	IgG ou A
Proportion de sérums "positifs" parmi 49 sérums de malades infectés par <i>L. pneumophila</i> diagnostiqués selon l'art antérieur et soumis au diagnostic selon l'invention	45,0%	43,0%	59,0%
Proportion de sérums "négatifs" parmi 96 sérums de donneurs de sang testés comme témoins	98,0%	98,0%	96,0%

Il en découle que les polypeptides peuvent être  
 15 utilisés pour mettre en évidence la présence de certains isotypes d'anticorps lors d'infections à *L. pneumophila*.

De plus il existe une complémentarité des résultats donnés par les différents isotypes, des patients possédant des anticorps IgA contre les polypeptides selon l'invention  
 20 ne possédant pas toujours des anticorps IgG contre ce même polypeptide. Par conséquent, l'utilisation de différents isotypes d'anticorps pour un même antigène, permet

d'augmenter la sensibilité du diagnostic sérologique (tableau 2).

Exemple C3 : combinaison de polypeptides pour le diagnostic *in vitro* des infections à *L. pneumophila*

5 Il existe une complémentarité des résultats donnés par les différents antigènes ; des patients possédant des anticorps contre l'un des polypeptides selon l'invention ne possédant pas toujours des anticorps contre l'autre polypeptide.

10 Ainsi, l'utilisation conjointe de plusieurs antigènes permet d'augmenter la sensibilité de la méthode (Tableau 3).

15 Tableau 3 : Résultats (ELISA) obtenus en utilisant des combinaisons de différents polypeptides selon l'invention, pour la détection d'immunoglobulines de type IgA dans les sérums de patients

Polypeptides testés	2E1	2A1	2E1 et 2A1
Proportion de sérums "positifs" parmi 49 sérums de malades infectés par <i>L. pneumophila</i> diagnostiqués selon l'art antérieur et soumis au diagnostic selon l'invention	51,0%	45,0%	70,0%
Proportion de sérums "négatifs" parmi 96 sérums de donneurs de sang testés comme témoins	97,0 %	98,0%	95,0%

Exemple C4 : Utilisation des polypeptides selon l'invention pour le suivi thérapeutique

Chez des patients en convalescence, la constatation de la diminution ou de l'augmentation, au cours du temps, des valeurs ELISA obtenues en détectant les immunoglobulines (anticorps) et les isotypes d'anticorps présents dans les sérums, par leur fixation aux polypeptides selon l'invention, indique la guérison du patient ou l'évolution de la maladie. L'utilisation des polypeptides selon l'invention permet donc de réaliser des suivis thérapeutiques, c'est-à-dire (i) d'évaluer l'effet d'un traitement et (ii) de suivre l'évolution de l'infection chez un patient.

15 D) Conclusions

Les résultats des tests précédents démontrent l'existence d'une réponse anticorps significative au cours des infections à *Legionella pneumophila* vis-à-vis des polypeptides de l'invention et la pertinence de ces polypeptides et de leurs associations/combinaisons pour le diagnostic sérologique de ce type d'infection. De même, la combinaison des résultats obtenus par la détection des différents isotypes d'anticorps dans les sérums, pour chacun des polypeptides, est elle aussi pertinente.

25

Ainsi, l'utilisation conjointe de plusieurs antigènes et/ou la détection de plusieurs isotypes d'anticorps permet en outre :

- 30 (i) d'augmenter la sensibilité de la méthode ;  
(ii) de déterminer l'évolution de la pathologie (par l'utilisation de certains antigènes et d'isotypes d'anticorps) et ;  
(iii) de réaliser des suivis thérapeutiques à moindre coût.

REVENDEICATIONS

1. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila*, d'au  
5 moins un polypeptide de séquence d'acides aminés ID SEQ N°4, 6, 8, 10, 12, 14 ou 16.

2. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila*, d'une  
10 combinaison d'au moins un polypeptide, tel que défini selon la revendication 1, avec un polypeptide de séquence d'acides aminés ID SEQ N°2.

3. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits  
15 à la suite d'une infection à *Legionella pneumophila*, d'au moins un polypeptide comprenant :

a) une partie d'une des séquences d'acides aminés telles que définies selon la revendication 1 et ayant la même fonction que ladite séquence, ou

20 b) une séquence d'acides aminés ayant au moins 60% d'identité avec l'une des séquences d'acides aminés telles que définies selon la revendication 1 ou avec une partie d'une des séquences telles que définies sous a), et ayant la même fonction que ladite séquence.

25 4. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila*, d'au moins un polypeptide, tel que défini selon la revendication 3, en combinaison avec un polypeptide comprenant :

30 a) une partie de la séquence d'acides aminés ID SEQ N°2 et ayant la même fonction que ladite séquence, ou

b) une séquence d'acides aminés ayant au moins 60% d'identité avec la séquence d'acides aminés ID SEQ N°2 ou  
35 avec une partie de la séquence telle que définie sous a), et ayant la même fonction que ladite séquence.

5. Utilisation selon l'une quelconque des revendications 1 à 4, caractérisée en ce que le ou les polypeptide(s) est(sont) préparé(s) par culture d'une cellule hôte comprenant un vecteur recombinant ayant, 5 inséré, un polynucléotide codant pour ledit ou lesdits polypeptide(s) et par isolement dudit ou desdits polypeptide(s) du milieu de culture.

6. Utilisation selon la revendication 5, quand elle dépend de la revendication 1, caractérisée en ce que le 10 polynucléotide est de séquence ID SEQ N°3, 5, 7, 9 11, 13 ou 15.

7. Utilisation selon la revendication 5, quand elle dépend de la revendication 2, caractérisée en ce que le polynucléotide est de séquence ID SEQ N°1.

15 8. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila*, d'au moins un polynucléotide de séquence ID SEQ N°3, 5, 7, 9 11, 13 ou 15.

20 9. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila* d'une combinaison d'au moins un polynucléotide, tel que défini selon la revendication 8, avec un polynucléotide de 25 séquence ID SEQ N°1.

10. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila*, d'au moins un polynucléotide comprenant

30 a) une partie d'une des séquences telles que définies selon la revendication 8 et ayant la même fonction que ladite séquence, ou

b) une séquence polynucléotidique ayant au moins 60% d'identité avec une séquence polynucléotidique telle que 35 définie selon la revendication 8, ou avec une partie d'une des séquences telles que définies sous a), et ayant la même fonction que ladite séquence, ou

c) une séquence polynucléotidique complémentaire à une des séquences polynucléotidiques telles que définies selon la revendication 8, ou selon la revendication 10 a) ou b).

5 11. Utilisation pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'anticorps produits à la suite d'une infection à *Legionella pneumophila*, d'au moins un polynucléotide, telle que défini selon la revendication 10, en combinaison avec un polynucléotide  
10 comprenant :

a) une partie de la séquence ID SEQ N°1 et ayant la même fonction que ladite séquence, ou

b) une séquence polynucléotidique ayant au moins 60% d'identité avec la séquence polynucléotidique ID SEQ N°1,  
15 ou avec une partie de cette séquence, et ayant la même fonction que ladite séquence, ou

c) une séquence polynucléotidique complémentaire à la séquence polynucléotidique ID SEQ N°1 ou la séquence telle que définie selon la revendication 11 a) ou b).

20 12. Utilisation d'anticorps pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'antigènes issus d'une infection par *Legionella pneumophila*, caractérisée en ce que lesdits anticorps sont immunospécifiques pour un polypeptide tel que défini selon  
25 l'une quelconque des revendications 1 à 5.

13. Utilisation de molécules pour détecter, *in vitro*, dans des échantillons biologiques, la présence d'antigènes issus d'une infection par *Legionella pneumophila*, caractérisée en ce que lesdites molécules fixent  
30 spécifiquement un polypeptide tel que défini dans les revendications 1 à 5.

14. Kit pour l'utilisation selon l'une quelconque des revendications 1 à 11, caractérisé en ce qu'il comprend au moins l'un desdits polypeptides ou l'un desdits  
35 polynucléotides codant pour lesdits polypeptides.

15. Kit pour l'utilisation selon la revendication 12, caractérisé en ce qu'il comprend au moins l'un desdits anticorps.

16. Kit pour l'utilisation selon la revendication 13, 5 caractérisé en ce qu'il comprend au moins l'une desdites molécules.

17. Composition pharmaceutique, notamment vaccin, comprenant, à titre de principe actif, au moins :

a) un polypeptide de séquence d'acides aminés 10 ID SEQ N°4, 6, 8, 10, 12, 14 ou 16, ou

b) un polypeptide dont la séquence est une partie de l'une des séquences d'acides aminés telles que définies sous a) et ayant la même fonction que ladite séquence, ou

c) un polypeptide dont la séquence est une séquence 15 d'acides aminés ayant au moins 60% d'identité avec l'une des séquences d'acides aminés telles que définies sous a) ou avec la partie de séquence telle que définie sous b), et ayant la même fonction que ladite séquence

ainsi qu'un excipient pharmaceutiquement acceptable.

20 18. Composition pharmaceutique, notamment vaccin, comprenant, à titre de principe actif, au moins l'un des polypeptides, tel que défini selon la revendication 17, en combinaison avec :

a) un polypeptide de séquences d'acides aminés 25 ID SEQ N°2, ou

b) un polypeptide dont la séquence est une partie de la séquence ID SEQ N°2 et ayant la même fonction que ladite séquence, ou

c) un polypeptide dont la séquence d'acides aminés 30 est une séquence ayant au moins 60% d'identité avec la séquence d'acides aminés telle que définie sous a) ou avec une partie de la séquence telle que définie sous b) et ayant la même fonction que ladite séquence

ainsi qu'un excipient pharmaceutiquement acceptable.

35 19. Composition pharmaceutique, notamment vaccin, comprenant, à titre de principe actif, au moins :

a) un polynucléotide de séquence ID SEQ N°3, 5, 7, 9  
11, 13 ou 15, ou

b) un polynucléotide dont la séquence est une partie  
de l'une des séquences polynucléotidiques telles que  
5 définies sous a) et ayant la même fonction que ladite  
séquence, ou

c) un polynucléotide dont la séquence est une  
séquence polynucléotidique ayant au moins 60% d'identité  
avec l'une des séquences polynucléotidiques telles que  
10 définies sous a) ou avec une partie de l'une des séquences  
telles que définies sous b), et ayant la même fonction que  
ladite séquence, ou

d) un polynucléotide dont la séquence est une  
séquence polynucléotidique complémentaire à l'une des  
15 séquences polynucléotidiques telles que définies sous a),  
b) ou c)

ainsi qu'un excipient pharmaceutiquement acceptable.

20. Composition pharmaceutique, notamment vaccin,  
comprenant, à titre de principe actif, au moins l'un des  
20 polynucléotides, tel que défini selon la revendication 19,  
en combinaison avec :

a) un polynucléotide de séquence ID SEQ N°1, ou

b) un polynucléotide dont la séquence est une partie  
de la séquence polynucléotidique telle que définie sous a)  
25 et ayant la même fonction que ladite séquence, ou

c) un polynucléotide dont la séquence est une  
séquence polynucléotidique ayant au moins 60% d'identité  
avec la séquence polynucléotidique telle que définie sous  
a) ou avec une partie de la séquence telle que définie sous  
30 b), et ayant la même fonction que ladite séquence, ou

d) un polynucléotide dont la séquence est une  
séquence polynucléotidique complémentaire à la séquence  
polynucléotidique telle que définie sous a), b) ou c)

ainsi qu'un excipient pharmaceutiquement acceptable.



aca tct gag cag aac atc atg ttt tac aga gga tta aac aaa ttc tta	384
Thr Ser Glu Gln Asn Ile Met Phe Tyr Arg Gly Leu Asn Lys Phe Leu	
115 120 125	
aat ttc att tat gaa agc aat gat tct cgt aaa ggc tta aaa aag gag	432
Asn Phe Ile Tyr Glu Ser Asn Asp Ser Arg Lys Gly Leu Lys Lys Glu	
130 135 140	
cat ttc ctg caa gtc ctt tcg tta aaa aag ata tac tct tta gcc aaa	480
His Phe Leu Gln Val Leu Ser Leu Lys Lys Ile Tyr Ser Leu Ala Lys	
145 150 155 160	
tta agt tat gag cag gaa gag gca gct gaa aat aat gct ttg gca aaa	528
Leu Ser Tyr Glu Gln Glu Glu Ala Ala Glu Asn Asn Ala Leu Ala Lys	
165 170 175	
tta act gct gat gga aaa acc aaa gcc aat gcg aac agc ttc cat gtg	576
Leu Thr Ala Asp Gly Lys Thr Lys Ala Asn Ala Asn Ser Phe His Val	
180 185 190	
gaa aaa cca atc gat tca tcc att gtt gag caa ttc aaa tcc tgg gat	624
Glu Lys Pro Ile Asp Ser Ser Ile Val Glu Gln Phe Lys Ser Trp Asp	
195 200 205	
gaa atg aaa ggc gct ctt cat caa tta att ctc gat gaa ctt tct gat	672
Glu Met Lys Gly Ala Leu His Gln Leu Ile Leu Asp Glu Leu Ser Asp	
210 215 220	
aag aac gtt gct aaa att tca gct tta agt caa gcc cgc tcc gca caa	720
Lys Asn Val Ala Lys Ile Ser Ala Leu Ser Gln Ala Arg Ser Ala Gln	
225 230 235 240	
ctg aaa ttt tta caa acg atg gca gaa caa cta gac aag atc cct aac	768
Leu Lys Phe Leu Gln Thr Met Ala Glu Gln Leu Asp Lys Ile Pro Asn	
245 250 255	
caa tct ctt gag ccg tct gag aaa atg gcc att ctt gct gga gca atg	816
Gln Ser Leu Glu Pro Ser Glu Lys Met Ala Ile Leu Ala Gly Ala Met	
260 265 270	
tat atc gtt cgt ggc caa ata gcc cag gag tat gga aaa gat cca tta	864
Tyr Ile Val Arg Gly Gln Ile Ala Gln Glu Tyr Gly Lys Asp Pro Leu	
275 280 285	
agc aac gat aaa atc agc gct act gtg att cat aca ggt tta agc acg	912
Ser Asn Asp Lys Ile Ser Ala Thr Val Ile His Thr Gly Leu Ser Thr	
290 295 300	
ata ctt cat gcc aat gct gat tgc tgt gaa gac aag gaa gta ctg att	960
Ile Leu His Ala Asn Ala Asp Cys Cys Glu Asp Lys Glu Val Leu Ile	
305 310 315 320	
gct gct gcg aat aaa ttt att cgt cat atg gtt att gag cgt cct gaa	1008
Ala Ala Ala Asn Lys Phe Ile Arg His Met Val Ile Glu Arg Pro Glu	
325 330 335	
caa tca aat aaa aaa atc act aag gaa tcc gtt cga gaa aac aac atg	1056
Gln Ser Asn Lys Lys Ile Thr Lys Glu Ser Val Arg Glu Asn Asn Met	
340 345 350	

ttt tcc gac atc gcc ggc ttc caa ttg atc tct gtc ttg acg tta ata 1104  
 Phe Ser Asp Ile Ala Gly Phe Gln Leu Ile Ser Val Leu Thr Leu Ile  
 355 360 365

caa aac atg atc aaa aca tgt cgt act gat gcc att gaa gct tgt gtc 1152  
 Gln Asn Met Ile Lys Thr Cys Arg Thr Asp Ala Ile Glu Ala Cys Val  
 370 375 380

acc aag cgt aag gaa gaa ctc gaa gca tta aaa ccc aaa aaa gag ggt 1200  
 Thr Lys Arg Lys Glu Glu Leu Glu Ala Leu Lys Pro Lys Lys Glu Gly  
 385 390 395 400

tat tcc att gcg agt tca gtc act gga tat gta ggc agc tgg ttt aaa 1248  
 Tyr Ser Ile Ala Ser Ser Val Thr Gly Tyr Val Gly Ser Trp Phe Lys  
 405 410 415

aaa gca cca agc atg tct gaa gaa gac gaa gaa gat gac tta aaa gat 1296  
 Lys Ala Pro Ser Met Ser Glu Glu Asp Glu Glu Asp Asp Leu Lys Asp  
 420 425 430

caa aac aca gca gaa gag acc agc aaa ccg acc gta 1332  
 Gln Asn Thr Ala Glu Glu Thr Ser Lys Pro Thr Val  
 435 440

<210> 2  
 <211> 444  
 <212> PRT  
 <213> Legionella pneumophila

<400> 2

Thr Phe Val Leu Lys Glu Phe Asp Ala Leu Lys Ser His Phe Asn Asp  
 1 5 10 15

Thr Val Lys Ile Ile Leu Gln Arg Glu Lys Lys Asp Lys Ile Glu Asp  
 20 25 30

Leu Pro Asn Pro Arg Lys Glu Glu Leu Gln Phe Leu Thr Ala Val Leu  
 35 40 45

Asn Gln Leu Glu Ala Lys Ile Asp Glu Leu Lys Pro Arg Ser Leu Ala  
 50 55 60

Ser Tyr Val His Val Phe Tyr Gly Ala Met Leu Leu Val Cys Lys Asp  
 65 70 75 80

Val Glu Asn Asn Leu Arg Val Met Glu Lys Lys Glu Asn Ser Leu Leu  
 85 90 95

Phe Thr Arg Leu Met Asp Gly Met Gly Ile Ser Asp Glu Asn Ile Pro  
 100 105 110

Thr Ser Glu Gln Asn Ile Met Phe Tyr Arg Gly Leu Asn Lys Phe Leu  
 115 120 125

Asn Phe Ile Tyr Glu Ser Asn Asp Ser Arg Lys Gly Leu Lys Lys Glu  
 130 135 140

His Phe Leu Gln Val Leu Ser Leu Lys Lys Ile Tyr Ser Leu Ala Lys  
 145 150 155 160

Leu Ser Tyr Glu Gln Glu Glu Ala Ala Glu Asn Asn Ala Leu Ala Lys  
 165 170 175

Leu Thr Ala Asp Gly Lys Thr Lys Ala Asn Ala Asn Ser Phe His Val  
 180 185 190

Glu Lys Pro Ile Asp Ser Ser Ile Val Glu Gln Phe Lys Ser Trp Asp  
 195 200 205

Glu Met Lys Gly Ala Leu His Gln Leu Ile Leu Asp Glu Leu Ser Asp  
 210 215 220

Lys Asn Val Ala Lys Ile Ser Ala Leu Ser Gln Ala Arg Ser Ala Gln  
 225 230 235 240

Leu Lys Phe Leu Gln Thr Met Ala Glu Gln Leu Asp Lys Ile Pro Asn  
 245 250 255

Gln Ser Leu Glu Pro Ser Glu Lys Met Ala Ile Leu Ala Gly Ala Met  
 260 265 270

Tyr Ile Val Arg Gly Gln Ile Ala Gln Glu Tyr Gly Lys Asp Pro Leu  
 275 280 285

Ser Asn Asp Lys Ile Ser Ala Thr Val Ile His Thr Gly Leu Ser Thr  
 290 295 300

Ile Leu His Ala Asn Ala Asp Cys Cys Glu Asp Lys Glu Val Leu Ile  
 305 310 315 320

Ala Ala Ala Asn Lys Phe Ile Arg His Met Val Ile Glu Arg Pro Glu  
 325 330 335

Gln Ser Asn Lys Lys Ile Thr Lys Glu Ser Val Arg Glu Asn Asn Met  
 340 345 350

Phe Ser Asp Ile Ala Gly Phe Gln Leu Ile Ser Val Leu Thr Leu Ile  
 355 360 365

Gln Asn Met Ile Lys Thr Cys Arg Thr Asp Ala Ile Glu Ala Cys Val  
 370 375 380

Thr Lys Arg Lys Glu Glu Leu Glu Ala Leu Lys Pro Lys Lys Glu Gly  
 385 390 395 400

Tyr Ser Ile Ala Ser Ser Val Thr Gly Tyr Val Gly Ser Trp Phe Lys  
 405 410 415

Lys Ala Pro Ser Met Ser Glu Glu Asp Glu Glu Asp Asp Leu Lys Asp  
 420 425 430

Gln Asn Thr Ala Glu Glu Thr Ser Lys Pro Thr Val  
 435 440

<210> 3  
 <211> 1008  
 <212> DNA  
 <213> Legionella pneumophila

<220>  
 <221> CDS  
 <222> (1)..(1008)

<400> 3  
 gaa aaa gtt caa gca aaa ggg atg ggc ttt ggt gga aat cgt aaa att 48  
 Glu Lys Val Gln Ala Lys Gly Met Gly Phe Gly Gly Asn Arg Lys Ile  
 1 5 10 15  
 ggt gaa tat caa ttt ggc aag gac ttg cca tta tta gaa atc aca cgt 96  
 Gly Glu Tyr Gln Phe Gly Lys Asp Leu Pro Leu Leu Glu Ile Thr Arg  
 20 25 30  
 gat tca agt gtt gaa atg tgc ttt atg gaa aat acc gat gtt aaa gta 144  
 Asp Ser Ser Val Glu Met Cys Phe Met Glu Asn Thr Asp Val Lys Val  
 35 40 45  
 gtc gac atg gga cat aag tac tat tca aat aac aaa cct atg caa ttt 192  
 Val Asp Met Gly His Lys Tyr Tyr Ser Asn Asn Lys Pro Met Gln Phe  
 50 55 60  
 act tgt aaa gag act ccc gac act caa tcg act aaa act tat tat acc 240  
 Thr Cys Lys Glu Thr Pro Asp Thr Gln Ser Thr Lys Thr Tyr Tyr Thr  
 65 70 75 80  
 gga tat tcg gct gat ggt tat gat aga gat aat gga gcc gct tct cct 288  
 Gly Tyr Ser Ala Asp Gly Tyr Asp Arg Asp Asn Gly Ala Ala Ser Pro  
 85 90 95

acc aat gac gca ttg tat gca ggt tat gtc att aaa cac atg tac cat	336
Thr Asn Asp Ala Leu Tyr Ala Gly Tyr Val Ile Lys His Met Tyr His	
100 105 110	
gat tgg tat ggg gtg gag gca ttg act aaa tca gat gga tcg cca atg	384
Asp Trp Tyr Gly Val Glu Ala Leu Thr Lys Ser Asp Gly Ser Pro Met	
115 120 125	
caa tta gtt atg aga gtg cat tat ggt caa ggc tat gaa aac gcc tac	432
Gln Leu Val Met Arg Val His Tyr Gly Gln Gly Tyr Glu Asn Ala Tyr	
130 135 140	
tgg gac ggc aaa caa atg acc ttt ggt gat ggt gac acc atg atg tat	480
Trp Asp Gly Lys Gln Met Thr Phe Gly Asp Gly Asp Thr Met Met Tyr	
145 150 155 160	
cct ttg gtt tct cta ggt gta ggt ggc cat gaa atc agt cac ggg ttc	528
Pro Leu Val Ser Leu Gly Val Gly Gly His Glu Ile Ser His Gly Phe	
165 170 175	
acc gaa caa cat tct gga ctt gaa tat ttc ggt cag tct ggt ggc atg	576
Thr Glu Gln His Ser Gly Leu Glu Tyr Phe Gly Gln Ser Gly Gly Met	
180 185 190	
aat gag tcg ttt tca gat atg gct gcc caa gct gct gaa tat tat tcg	624
Asn Glu Ser Phe Ser Asp Met Ala Ala Gln Ala Ala Glu Tyr Tyr Ser	
195 200 205	
gtt gga aaa aac agc tgg caa atc ggc ccg gaa ata atg aaa gaa gac	672
Val Gly Lys Asn Ser Trp Gln Ile Gly Pro Glu Ile Met Lys Glu Asp	
210 215 220	
agt ggt tat gac gca ttg cgt tac atg gat aaa cca agc cgt gat ggt	720
Ser Gly Tyr Asp Ala Leu Arg Tyr Met Asp Lys Pro Ser Arg Asp Gly	
225 230 235 240	
atg tct att gat gta gct gat gat tat tat ggt ggg tta gat gtt cat	768
Met Ser Ile Asp Val Ala Asp Asp Tyr Tyr Gly Gly Leu Asp Val His	
245 250 255	
tac tcc agt gga gtg tat aac cat tta ttc tac ata tta gcc aat caa	816
Tyr Ser Ser Gly Val Tyr Asn His Leu Phe Tyr Ile Leu Ala Asn Gln	
260 265 270	
cct aat tgg aat ctt cga atg gct ttt gat gtc atg gtg aaa gcc aat	864
Pro Asn Trp Asn Leu Arg Met Ala Phe Asp Val Met Val Lys Ala Asn	
275 280 285	
atg gat tat tgg aca cct tat tca aca ttt gat gag ggt ggt tgc ggt	912
Met Asp Tyr Trp Thr Pro Tyr Ser Thr Phe Asp Glu Gly Gly Cys Gly	
290 295 300	
atg ttg agt gct gcc aaa gac ctg ggt tac aat ctg gat gat gtt aag	960
Met Leu Ser Ala Ala Lys Asp Leu Gly Tyr Asn Leu Asp Asp Val Lys	
305 310 315 320	
aag tca ttg agt gaa gtg act ata aat tat caa tct tgt tat gtc gat	1008
Lys Ser Leu Ser Glu Val Thr Ile Asn Tyr Gln Ser Cys Tyr Val Asp	
325 330 335	

<210> 4  
 <211> 336  
 <212> PRT  
 <213> Legionella pneumophila

<400> 4

Glu Lys Val Gln Ala Lys Gly Met Gly Phe Gly Gly Asn Arg Lys Ile  
 1 5 10 15

Gly Glu Tyr Gln Phe Gly Lys Asp Leu Pro Leu Leu Glu Ile Thr Arg  
 20 25 30

Asp Ser Ser Val Glu Met Cys Phe Met Glu Asn Thr Asp Val Lys Val  
 35 40 45

Val Asp Met Gly His Lys Tyr Tyr Ser Asn Asn Lys Pro Met Gln Phe  
 50 55 60

Thr Cys Lys Glu Thr Pro Asp Thr Gln Ser Thr Lys Thr Tyr Tyr Thr  
 65 70 75 80

Gly Tyr Ser Ala Asp Gly Tyr Asp Arg Asp Asn Gly Ala Ala Ser Pro  
 85 90 95

Thr Asn Asp Ala Leu Tyr Ala Gly Tyr Val Ile Lys His Met Tyr His  
 100 105 110

Asp Trp Tyr Gly Val Glu Ala Leu Thr Lys Ser Asp Gly Ser Pro Met  
 115 120 125

Gln Leu Val Met Arg Val His Tyr Gly Gln Gly Tyr Glu Asn Ala Tyr  
 130 135 140

Trp Asp Gly Lys Gln Met Thr Phe Gly Asp Gly Asp Thr Met Met Tyr  
 145 150 155 160

Pro Leu Val Ser Leu Gly Val Gly Gly His Glu Ile Ser His Gly Phe  
 165 170 175

Thr Glu Gln His Ser Gly Leu Glu Tyr Phe Gly Gln Ser Gly Gly Met  
 180 185 190

Asn Glu Ser Phe Ser Asp Met Ala Ala Gln Ala Ala Glu Tyr Tyr Ser  
 195 200 205

Val Gly Lys Asn Ser Trp Gln Ile Gly Pro Glu Ile Met Lys Glu Asp  
 210 215 220

Ser Gly Tyr Asp Ala Leu Arg Tyr Met Asp Lys Pro Ser Arg Asp Gly  
 225 230 235 240

Met Ser Ile Asp Val Ala Asp Asp Tyr Tyr Gly Gly Leu Asp Val His  
 245 250 255

Tyr Ser Ser Gly Val Tyr Asn His Leu Phe Tyr Ile Leu Ala Asn Gln  
 260 265 270

Pro Asn Trp Asn Leu Arg Met Ala Phe Asp Val Met Val Lys Ala Asn  
 275 280 285

Met Asp Tyr Trp Thr Pro Tyr Ser Thr Phe Asp Glu Gly Gly Cys Gly  
 290 295 300

Met Leu Ser Ala Ala Lys Asp Leu Gly Tyr Asn Leu Asp Asp Val Lys  
 305 310 315 320

Lys Ser Leu Ser Glu Val Thr Ile Asn Tyr Gln Ser Cys Tyr Val Asp  
 325 330 335

- <210> 5
- <211> 432
- <212> DNA
- <213> Legionella pneumophila

- <220>
- <221> CDS
- <222> (1)..(432)

<400> 5  
 gat tgg gca aag ata ggc gta gtt gat tta caa aaa atc atg caa acc 48  
 Asp Trp Ala Lys Ile Gly Val Val Asp Leu Gln Lys Ile Met Gln Thr  
 1 5 10 15  
 tca aat caa atg aag gaa atc caa cag aag ttg gaa aaa gaa ttt aag 96  
 Ser Asn Gln Met Lys Glu Ile Gln Gln Lys Leu Glu Lys Glu Phe Lys  
 20 25 30  
 cct cgt cga gac aag ctt gtc gca atg gaa gca agc ttg aaa agc gac 144  
 Pro Arg Arg Asp Lys Leu Val Ala Met Glu Ala Ser Leu Lys Ser Asp  
 35 40 45  
 atg gaa aaa ttc aag cgt gat agt gct ata atg agc gcg agc caa aag 192  
 Met Glu Lys Phe Lys Arg Asp Ser Ala Ile Met Ser Ala Ser Gln Lys  
 50 55 60  
 aaa gaa ttg gaa aag aaa att gtt gcg tca caa caa cag ttt gaa cgt 240  
 Lys Glu Leu Glu Lys Lys Ile Val Ala Ser Gln Gln Gln Phe Glu Arg  
 65 70 75 80



<210> 7  
 <211> 285  
 <212> DNA  
 <213> Legionella pneumophila

<220>  
 <221> CDS  
 <222> (1)..(285)

<400> 7  
 aaa att cgt cct tta cac gat cgc gtt gtt gtt cgt cgt atg gaa gaa 48  
 Lys Ile Arg Pro Leu His Asp Arg Val Val Val Arg Arg Met Glu Glu  
 1 5 10 15  
 gag cgt acc act gct ggt ggt att gtt att cca gat agt gct acc gag 96  
 Glu Arg Thr Thr Ala Gly Gly Ile Val Ile Pro Asp Ser Ala Thr Glu  
 20 25 30  
 aaa cct atg cgt ggt gag atc att gct gtc ggc gct ggc aaa gta ttg 144  
 Lys Pro Met Arg Gly Glu Ile Ile Ala Val Gly Ala Gly Lys Val Leu  
 35 40 45  
 gaa aac ggt gat gtt cgt gct tta gca gtt aaa gta ggt gat gtg gta 192  
 Glu Asn Gly Asp Val Arg Ala Leu Ala Val Lys Val Gly Asp Val Val  
 50 55 60  
 ttg ttt ggc aag tac tct ggt act gaa gtt aaa gtt gat gga aaa gaa 240  
 Leu Phe Gly Lys Tyr Ser Gly Thr Glu Val Lys Val Asp Gly Lys Glu  
 65 70 75 80  
 tta gtt gtg atg cgc gaa gac gac atc atg ggt gta atc gag aag 285  
 Leu Val Val Met Arg Glu Asp Asp Ile Met Gly Val Ile Glu Lys  
 85 90 95

<210> 8  
 <211> 95  
 <212> PRT  
 <213> Legionella pneumophila

<400> 8  
 Lys Ile Arg Pro Leu His Asp Arg Val Val Val Arg Arg Met Glu Glu  
 1 5 10 15  
 Glu Arg Thr Thr Ala Gly Gly Ile Val Ile Pro Asp Ser Ala Thr Glu  
 20 25 30  
 Lys Pro Met Arg Gly Glu Ile Ile Ala Val Gly Ala Gly Lys Val Leu  
 35 40 45  
 Glu Asn Gly Asp Val Arg Ala Leu Ala Val Lys Val Gly Asp Val Val  
 50 55 60  
 Leu Phe Gly Lys Tyr Ser Gly Thr Glu Val Lys Val Asp Gly Lys Glu  
 65 70 75 80

Leu Val Val Met Arg Glu Asp Asp Ile Met Gly Val Ile Glu Lys  
 85 90 95

<210> 9  
 <211> 711  
 <212> DNA  
 <213> Legionella pneumophila

<220>  
 <221> CDS  
 <222> (1)..(711)

<400> 9  
 gat acg gct agc gct tct tta tca gat gct caa aaa aag gaa att gaa 48  
 Asp Thr Ala Ser Ala Ser Leu Ser Asp Ala Gln Lys Lys Glu Ile Glu  
 1 5 10 15  
 aaa gtc att cat gat tat ctg att aat aac ccg gaa gtt tta ttg gag 96  
 Lys Val Ile His Asp Tyr Leu Ile Asn Asn Pro Glu Val Leu Leu Glu  
 20 25 30  
 gct tcg caa gca ttg caa caa aaa cag caa caa aat atg caa cag caa 144  
 Ala Ser Gln Ala Leu Gln Gln Lys Gln Gln Gln Asn Met Gln Gln Gln  
 35 40 45  
 gct caa gcc gca att cag gaa aat gct gaa caa gtt ttt caa ggc aaa 192  
 Ala Gln Ala Ala Ile Gln Glu Asn Ala Glu Gln Val Phe Gln Gly Lys  
 50 55 60  
 tta act act gtt ggt aat cct aaa ggg aat gta aca ctg gtt gag ttt 240  
 Leu Thr Thr Val Gly Asn Pro Lys Gly Asn Val Thr Leu Val Glu Phe  
 65 70 75 80  
 ttt gat tat caa tgc att cat tgc aaa aaa atg gca tct acc att gaa 288  
 Phe Asp Tyr Gln Cys Ile His Cys Lys Lys Met Ala Ser Thr Ile Glu  
 85 90 95  
 aat ctg gtt aaa aaa gac agc gga tta cga gtt ata tac aaa gaa ttt 336  
 Asn Leu Val Lys Lys Asp Ser Gly Leu Arg Val Ile Tyr Lys Glu Phe  
 100 105 110  
 cca atc ttt ggc aag act tct gat ctg gct tcc aga gtt gct ctg gcg 384  
 Pro Ile Phe Gly Lys Thr Ser Asp Leu Ala Ser Arg Val Ala Leu Ala  
 115 120 125  
 gct ggc atg caa ggc aaa tat cag gct atg cat aat gct ttg att acg 432  
 Ala Gly Met Gln Gly Lys Tyr Gln Ala Met His Asn Ala Leu Ile Thr  
 130 135 140  
 ata gac aaa cgt ttg gat gaa aaa aca gtg atg gat gcc gca aaa tcc 480  
 Ile Asp Lys Arg Leu Asp Glu Lys Thr Val Met Asp Ala Ala Lys Ser  
 145 150 155 160  
 ata ggt ttg gat atg caa aag ctg aaa aaa gat atg gat agt cag gaa 528  
 Ile Gly Leu Asp Met Gln Lys Leu Lys Lys Asp Met Asp Ser Gln Glu  
 165 170 175

gtg acc gat atc ctg gat gcg aat cgc caa ttg gct gaa aaa tta cat 576  
 Val Thr Asp Ile Leu Asp Ala Asn Arg Gln Leu Ala Glu Lys Leu His  
 180 185 190

ttg atg ggt aca cca gcc ttt ata att ggc tca aca cca gat ggt caa 624  
 Leu Met Gly Thr Pro Ala Phe Ile Ile Gly Ser Thr Pro Asp Gly Gln  
 195 200 205

tat aaa aag ggc agt gaa ata tca ttt ata cca ggc gct aca agc gag 672  
 Tyr Lys Lys Gly Ser Glu Ile Ser Phe Ile Pro Gly Ala Thr Ser Glu  
 210 215 220

caa tca cta cgt gag tta att aaa aag gcg gct ggc aat 711  
 Gln Ser Leu Arg Glu Leu Ile Lys Lys Ala Ala Gly Asn  
 225 230 235

<210> 10

<211> 237

<212> PRT

<213> Legionella pneumophila

<400> 10

Asp Thr Ala Ser Ala Ser Leu Ser Asp Ala Gln Lys Lys Glu Ile Glu  
 1 5 10 15

Lys Val Ile His Asp Tyr Leu Ile Asn Asn Pro Glu Val Leu Leu Glu  
 20 25 30

Ala Ser Gln Ala Leu Gln Gln Lys Gln Gln Gln Asn Met Gln Gln Gln  
 35 40 45

Ala Gln Ala Ala Ile Gln Glu Asn Ala Glu Gln Val Phe Gln Gly Lys  
 50 55 60

Leu Thr Thr Val Gly Asn Pro Lys Gly Asn Val Thr Leu Val Glu Phe  
 65 70 75 80

Phe Asp Tyr Gln Cys Ile His Cys Lys Lys Met Ala Ser Thr Ile Glu  
 85 90 95

Asn Leu Val Lys Lys Asp Ser Gly Leu Arg Val Ile Tyr Lys Glu Phe  
 100 105 110

Pro Ile Phe Gly Lys Thr Ser Asp Leu Ala Ser Arg Val Ala Leu Ala  
 115 120 125

Ala Gly Met Gln Gly Lys Tyr Gln Ala Met His Asn Ala Leu Ile Thr  
 130 135 140

Ile Asp Lys Arg Leu Asp Glu Lys Thr Val Met Asp Ala Ala Lys Ser  
 145 150 155 160

Ile Gly Leu Asp Met Gln Lys Leu Lys Lys Asp Met Asp Ser Gln Glu  
 165 170 175

Val Thr Asp Ile Leu Asp Ala Asn Arg Gln Leu Ala Glu Lys Leu His  
 180 185 190

Leu Met Gly Thr Pro Ala Phe Ile Ile Gly Ser Thr Pro Asp Gly Gln  
 195 200 205

Tyr Lys Lys Gly Ser Glu Ile Ser Phe Ile Pro Gly Ala Thr Ser Glu  
 210 215 220

Gln Ser Leu Arg Glu Leu Ile Lys Lys Ala Ala Gly Asn  
 225 230 235

<210> 11  
 <211> 744  
 <212> DNA  
 <213> Legionella pneumophila

<220>  
 <221> CDS  
 <222> (1)..(744)

<400> 11  
 ttg ata ttc ctt atg caa ata aat act ttc ccc atc act cta gaa gag 48  
 Leu Ile Phe Leu Met Gln Ile Asn Thr Phe Pro Ile Thr Leu Glu Glu  
 1 5 10 15  
 tct ttc atg att tct ccc aag gta aag cac ttt ata ttc aat tgt gaa 96  
 Ser Phe Met Ile Ser Pro Lys Val Lys His Phe Ile Phe Asn Cys Glu  
 20 25 30  
 tta tca ccc cct ttt gaa tat tta cca gga caa ttt atc act att cat 144  
 Leu Ser Pro Pro Phe Glu Tyr Leu Pro Gly Gln Phe Ile Thr Ile His  
 35 40 45  
 ttt gaa cat gat gga aaa aac tta aag cgt agt tac agt att gcc aat 192  
 Phe Glu His Asp Gly Lys Asn Leu Lys Arg Ser Tyr Ser Ile Ala Asn  
 50 55 60  
 gaa ccc aaa aag gat aat cgc att gaa ttt gct gcc ggg tat ttc gaa 240  
 Glu Pro Lys Lys Asp Asn Arg Ile Glu Phe Ala Ala Gly Tyr Phe Glu  
 65 70 75 80  
 gga gga cca ggc aca gaa cta ctc tac aat tta aag cca ggt gat gtg 288  
 Gly Gly Pro Gly Thr Glu Leu Leu Tyr Asn Leu Lys Pro Gly Asp Val  
 85 90 95



Phe Glu His Asp Gly Lys Asn Leu Lys Arg Ser Tyr Ser Ile Ala Asn  
 50 55 60  
 Glu Pro Lys Lys Asp Asn Arg Ile Glu Phe Ala Ala Gly Tyr Phe Glu  
 65 70 75 80  
 Gly Gly Pro Gly Thr Glu Leu Leu Tyr Asn Leu Lys Pro Gly Asp Val  
 85 90 95  
 Ile His Ile Asn Gly Pro Phe Gly Arg Leu Ile Phe Arg Asp Glu Thr  
 100 105 110  
 Pro Gly Arg Tyr Ile Leu Val Ala Thr Ser Thr Gly Thr Thr Pro Tyr  
 115 120 125  
 Arg Ala Met Leu Asn Glu Leu Gly Gln Arg Ile Glu Lys His Pro Asp  
 130 135 140  
 Leu Gln Val Val Ile Leu Gln Gly Val Gln Arg Ser Glu Glu Ile Leu  
 145 150 155 160  
 Tyr Pro Asp Asp Phe Gln Ala Phe Ala Lys Lys Tyr Pro Gln Ala Ser  
 165 170 175  
 Phe Arg Pro Tyr Leu Ser Arg Val Gln Lys Gln Asp Leu Lys Asp Asn  
 180 185 190  
 Glu Tyr Ser Gly Tyr Val Gln His Ala Phe Pro Glu Leu Asn Leu Asn  
 195 200 205  
 Pro Thr Arg Asp Met Val Tyr Leu Cys Gly Asn Pro Gly Met Ile Asp  
 210 215 220  
 Glu Ala Phe Asn Tyr Leu Lys Glu His Gly Phe Ser Met Gln His Ile  
 225 230 235 240  
 Val Arg Glu Lys Tyr Ile Ser Arg  
 245

<210> 13  
 <211> 1947  
 <212> DNA  
 <213> Legionella pneumophila

<220>  
 <221> CDS  
 <222> (1)..(1947)

<400> 13  
 ttg gag caa aat aga atg gct aaa ata ata ggt ata gac ctg ggt act 48  
 Leu Glu Gln Asn Arg Met Ala Lys Ile Ile Gly Ile Asp Leu Gly Thr  
 1 5 10 15

acg aac tct tgt gtt gct gtc atg gaa ggt gac aaa ccg aag gta att 96  
 Thr Asn Ser Cys Val Ala Val Met Glu Gly Asp Lys Pro Lys Val Ile  
 20 25 30

gaa aat agt gaa ggt cac aga aca act cct tct atc gta gca ttt act 144  
 Glu Asn Ser Glu Gly His Arg Thr Thr Pro Ser Ile Val Ala Phe Thr  
 35 40 45

gat gat aat gaa ata ttg gta ggc caa tca gcc aag cgt caa tcc gtg 192  
 Asp Asp Asn Glu Ile Leu Val Gly Gln Ser Ala Lys Arg Gln Ser Val  
 50 55 60

act aac cca gag aaa acg ctg ttt gcg att aaa cga tta att ggg cgt 240  
 Thr Asn Pro Glu Lys Thr Leu Phe Ala Ile Lys Arg Leu Ile Gly Arg  
 65 70 75 80

cga ttt gat gat cct atc gtg caa aaa gat att aag atg gtt ccc tac 288  
 Arg Phe Asp Asp Pro Ile Val Gln Lys Asp Ile Lys Met Val Pro Tyr  
 85 90 95

aaa att atg aaa gcg gac aat ggg gat gct tgg gtt aga gtc aaa gat 336  
 Lys Ile Met Lys Ala Asp Asn Gly Asp Ala Trp Val Arg Val Lys Asp  
 100 105 110

caa gat aag gca cca cca caa att tct gct gaa gtg tta aga aaa atg 384  
 Gln Asp Lys Ala Pro Pro Gln Ile Ser Ala Glu Val Leu Arg Lys Met  
 115 120 125

aaa aaa aca gca gaa gat tat ttg ggt gaa gaa gta aaa gaa gca gtg 432  
 Lys Lys Thr Ala Glu Asp Tyr Leu Gly Glu Glu Val Lys Glu Ala Val  
 130 135 140

att act gta cct gcc tat ttc aac gat tcg caa cgc cag gca acc aaa 480  
 Ile Thr Val Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys  
 145 150 155 160

gat gct ggt cgt att gct ggc ctg gaa gta aaa cga atc att aat gag 528  
 Asp Ala Gly Arg Ile Ala Gly Leu Glu Val Lys Arg Ile Ile Asn Glu  
 165 170 175

cct acc gct gct gca ctg gct tat ggt atg gat aaa aaa cgc ggt gac 576  
 Pro Thr Ala Ala Ala Leu Ala Tyr Gly Met Asp Lys Lys Arg Gly Asp  
 180 185 190

tct gtt att gcg gta tat gat tta ggg gga gga act ttt gat att tca 624  
 Ser Val Ile Ala Val Tyr Asp Leu Gly Gly Gly Thr Phe Asp Ile Ser  
 195 200 205

att att gaa att gcc gaa gtg gat ggt gaa cat caa ttt gaa gta ctt 672  
 Ile Ile Glu Ile Ala Glu Val Asp Gly Glu His Gln Phe Glu Val Leu  
 210 215 220

gca acc aat ggt gat acc ttc ctc ggt gga gaa gat ttt gat ttg gct 720  
 Ala Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Ala  
 225 230 235 240

tta att gag tat ttg gct tct gag ttc aag aaa gac act ggt att gat	768
Leu Ile Glu Tyr Leu Ala Ser Glu Phe Lys Lys Asp Thr Gly Ile Asp	
245 250 255	
ttg cat aat gat cct ctt gct ctg caa cgc ttg aaa gag gct gcc gag	816
Leu His Asn Asp Pro Leu Ala Leu Gln Arg Leu Lys Glu Ala Ala Glu	
260 265 270	
aaa gca aaa att gaa ttg tca tct gca caa caa act gat gtc aat ttg	864
Lys Ala Lys Ile Glu Leu Ser Ser Ala Gln Gln Thr Asp Val Asn Leu	
275 280 285	
cct tac att aca gca gat gct tct ggg cca aaa cat tta aat atc aaa	912
Pro Tyr Ile Thr Ala Asp Ala Ser Gly Pro Lys His Leu Asn Ile Lys	
290 295 300	
ttg act cgc gca aaa ctg gaa tct tta gtg gag aaa ctg gtt gag aga	960
Leu Thr Arg Ala Lys Leu Glu Ser Leu Val Glu Lys Leu Val Glu Arg	
305 310 315 320	
act att gag cca tgt aaa aca gct tta aaa gat gca ggg ctt act gtc	1008
Thr Ile Glu Pro Cys Lys Thr Ala Leu Lys Asp Ala Gly Leu Thr Val	
325 330 335	
tca caa atc aac gaa gtt att ctg gtc ggt ggg caa act cgt atg cct	1056
Ser Gln Ile Asn Glu Val Ile Leu Val Gly Gly Gln Thr Arg Met Pro	
340 345 350	
ctg gta caa aaa act gtg gaa gaa ttc ttt ggc aaa gaa cct cgt aag	1104
Leu Val Gln Lys Thr Val Glu Glu Phe Phe Gly Lys Glu Pro Arg Lys	
355 360 365	
gat gta aac cct gat gag gct gtg gca gta ggt gct gca att cag gca	1152
Asp Val Asn Pro Asp Glu Ala Val Ala Val Gly Ala Ala Ile Gln Ala	
370 375 380	
gca gta ttg tcc ggt gaa gtg aaa gac att ctt tta ctc gat gtc act	1200
Ala Val Leu Ser Gly Glu Val Lys Asp Ile Leu Leu Leu Asp Val Thr	
385 390 395 400	
cct ttg tct ctc ggt ata gaa acc atg ggt ggt gta atg acc aaa ctc	1248
Pro Leu Ser Leu Gly Ile Glu Thr Met Gly Gly Val Met Thr Lys Leu	
405 410 415	
att gaa aaa aac aca act att cct aca aag gct aca caa gtg ttt tca	1296
Ile Glu Lys Asn Thr Thr Ile Pro Thr Lys Ala Thr Gln Val Phe Ser	
420 425 430	
aca gcc gat gat aat cag aca gca gtg act gtg cat gta ttg cag ggt	1344
Thr Ala Asp Asp Asn Gln Thr Ala Val Thr Val His Val Leu Gln Gly	
435 440 445	
gaa aga gaa cag gct tca gca aac aag tca ctg gga cgg ttt gat ttg	1392
Glu Arg Glu Gln Ala Ser Ala Asn Lys Ser Leu Gly Arg Phe Asp Leu	
450 455 460	
cgt gac atc cct cct gct cct cgt ggt gtt cct caa att gaa gta act	1440
Arg Asp Ile Pro Pro Ala Pro Arg Gly Val Pro Gln Ile Glu Val Thr	
465 470 475 480	



Glu Asn Ser Glu Gly His Arg Thr Thr Pro Ser Ile Val Ala Phe Thr  
 35 40 45

Asp Asp Asn Glu Ile Leu Val Gly Gln Ser Ala Lys Arg Gln Ser Val  
 50 55 60

Thr Asn Pro Glu Lys Thr Leu Phe Ala Ile Lys Arg Leu Ile Gly Arg  
 65 70 75 80

Arg Phe Asp Asp Pro Ile Val Gln Lys Asp Ile Lys Met Val Pro Tyr  
 85 90 95

Lys Ile Met Lys Ala Asp Asn Gly Asp Ala Trp Val Arg Val Lys Asp  
 100 105 110

Gln Asp Lys Ala Pro Pro Gln Ile Ser Ala Glu Val Leu Arg Lys Met  
 115 120 125

Lys Lys Thr Ala Glu Asp Tyr Leu Gly Glu Glu Val Lys Glu Ala Val  
 130 135 140

Ile Thr Val Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys  
 145 150 155 160

Asp Ala Gly Arg Ile Ala Gly Leu Glu Val Lys Arg Ile Ile Asn Glu  
 165 170 175

Pro Thr Ala Ala Ala Leu Ala Tyr Gly Met Asp Lys Lys Arg Gly Asp  
 180 185 190

Ser Val Ile Ala Val Tyr Asp Leu Gly Gly Gly Thr Phe Asp Ile Ser  
 195 200 205

Ile Ile Glu Ile Ala Glu Val Asp Gly Glu His Gln Phe Glu Val Leu  
 210 215 220

Ala Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Ala  
 225 230 235 240

Leu Ile Glu Tyr Leu Ala Ser Glu Phe Lys Lys Asp Thr Gly Ile Asp  
 245 250 255

Leu His Asn Asp Pro Leu Ala Leu Gln Arg Leu Lys Glu Ala Ala Glu  
 260 265 270

Lys Ala Lys Ile Glu Leu Ser Ser Ala Gln Gln Thr Asp Val Asn Leu  
 275 280 285

Pro Tyr Ile Thr Ala Asp Ala Ser Gly Pro Lys His Leu Asn Ile Lys  
 290 295 300

Leu Thr Arg Ala Lys Leu Glu Ser Leu Val Glu Lys Leu Val Glu Arg  
 305 310 315 320

Thr Ile Glu Pro Cys Lys Thr Ala Leu Lys Asp Ala Gly Leu Thr Val  
 325 330 335

Ser Gln Ile Asn Glu Val Ile Leu Val Gly Gly Gln Thr Arg Met Pro  
 340 345 350

Leu Val Gln Lys Thr Val Glu Glu Phe Phe Gly Lys Glu Pro Arg Lys  
 355 360 365

Asp Val Asn Pro Asp Glu Ala Val Ala Val Gly Ala Ala Ile Gln Ala  
 370 375 380

Ala Val Leu Ser Gly Glu Val Lys Asp Ile Leu Leu Leu Asp Val Thr  
 385 390 395 400

Pro Leu Ser Leu Gly Ile Glu Thr Met Gly Gly Val Met Thr Lys Leu  
 405 410 415

Ile Glu Lys Asn Thr Thr Ile Pro Thr Lys Ala Thr Gln Val Phe Ser  
 420 425 430

Thr Ala Asp Asp Asn Gln Thr Ala Val Thr Val His Val Leu Gln Gly  
 435 440 445

Glu Arg Glu Gln Ala Ser Ala Asn Lys Ser Leu Gly Arg Phe Asp Leu  
 450 455 460

Arg Asp Ile Pro Pro Ala Pro Arg Gly Val Pro Gln Ile Glu Val Thr  
 465 470 475 480

Phe Asp Ile Asp Ala Asn Gly Ile Leu Asn Val Ser Ala Lys Asp Lys  
 485 490 495

Ala Thr Gly Lys Ala Gln Ser Ile Val Ile Lys Ala Ser Ser Gly Leu  
 500 505 510

Ser Glu Glu Glu Val Ala Ala Met Val Lys Asp Ala Gln Ser His Ala  
 515 520 525

Glu Glu Asp Lys Lys Phe Lys Glu Met Ala Glu Leu Arg Asn Gln Ala  
 530 535 540

Asp Ser Leu Ile His Ser Cys Glu Lys Ser Met Lys Asp Leu Ala Asp  
 545 550 555 560

Glu Leu Ser Glu Asp Glu Lys Arg Gly Ile Glu Thr Ala Ile Ser Glu  
 565 570 575

Leu Lys Glu Ala Val Gln Gly Thr Asp Lys Ala Arg Ile Glu Asp Lys  
 580 585 590

Leu Lys Val Leu Thr Asp Ala Ser Ala Lys Met Ala Glu Arg Ile Tyr  
 595 600 605

Ala Lys Lys Ser Ser Glu Gly Gln Ala Ala Gln Gly Gln Thr Gln Ser  
 610 615 620

Gln Glu Ser Thr Lys Pro Ala Glu Glu Gly Val Val Asp Ala Glu Phe  
 625 630 635 640

Glu Glu Val Lys Glu Glu Asp Lys Lys  
 645

<210> 15  
 <211> 1371  
 <212> DNA  
 <213> Legionella pneumophila

<220>  
 <221> CDS  
 <222> (1)..(1371)

<400> 15  
 ttg gag caa aat aga atg gct aaa ata ata ggt ata gac ctg ggt act 48  
 Leu Glu Gln Asn Arg Met Ala Lys Ile Ile Gly Ile Asp Leu Gly Thr  
 1 5 10 15

acg aac tct tgt gtt gct gtc atg gaa ggt gac aaa ccg aag gta att 96  
 Thr Asn Ser Cys Val Ala Val Met Glu Gly Asp Lys Pro Lys Val Ile  
 20 25 30

gaa aat agt gaa ggt cac aga aca act cct tct atc gta gca ttt act 144  
 Glu Asn Ser Glu Gly His Arg Thr Thr Pro Ser Ile Val Ala Phe Thr  
 35 40 45

gat gat aat gaa ata ttg gta ggc caa tca gcc aag cgt caa tcc gtg Asp Asp Asn Glu Ile Leu Val Gly Gln Ser Ala Lys Arg Gln Ser Val 50 55 60	192
act aac cca gag aaa acg ctg ttt gcg att aaa cga tta att ggg cgt Thr Asn Pro Glu Lys Thr Leu Phe Ala Ile Lys Arg Leu Ile Gly Arg 65 70 75 80	240
cga ttt gat gat cct atc gtg caa aaa gat att aag atg gtt ccc tac Arg Phe Asp Asp Pro Ile Val Gln Lys Asp Ile Lys Met Val Pro Tyr 85 90 95	288
aaa att atg aaa gcg gac aat ggg gat gct tgg gtt aga gtc aaa gat Lys Ile Met Lys Ala Asp Asn Gly Asp Ala Trp Val Arg Val Lys Asp 100 105 110	336
caa gat aag gca cca cca caa att tct gct gaa gtg tta aga aaa atg Gln Asp Lys Ala Pro Pro Gln Ile Ser Ala Glu Val Leu Arg Lys Met 115 120 125	384
aaa aaa aca gca gaa gat tat ttg ggt gaa gaa gta aaa gaa gca gtg Lys Lys Thr Ala Glu Asp Tyr Leu Gly Glu Glu Val Lys Glu Ala Val 130 135 140	432
att act gta cct gcc tat ttc aac gat tcg caa cgc cag gca acc aaa Ile Thr Val Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys 145 150 155 160	480
gat gct ggt cgt att gct ggc ctg gaa gta aaa cga atc att aat gag Asp Ala Gly Arg Ile Ala Gly Leu Glu Val Lys Arg Ile Ile Asn Glu 165 170 175	528
cct acc gct gct gca ctg gct tat ggt atg gat aaa aaa cgc ggt gac Pro Thr Ala Ala Leu Ala Tyr Gly Met Asp Lys Lys Arg Gly Asp 180 185 190	576
tct gtt att gcg gta tat gat tta ggg gga gga act ttt gat att tca Ser Val Ile Ala Val Tyr Asp Leu Gly Gly Gly Thr Phe Asp Ile Ser 195 200 205	624
att att gaa att gcc gaa gtg gat ggt gaa cat caa ttt gaa gta ctt Ile Ile Glu Ile Ala Glu Val Asp Gly Glu His Gln Phe Glu Val Leu 210 215 220	672
gca acc aat ggt gat acc ttc ctc ggt gga gaa gat ttt gat ttg gct Ala Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Ala 225 230 235 240	720
tta att gag tat ttg gct tct gag ttc aag aaa gac act ggt att gat Leu Ile Glu Tyr Leu Ala Ser Glu Phe Lys Lys Asp Thr Gly Ile Asp 245 250 255	768
ttg cat aat gat cct ctt gct ctg caa cgc ttg aaa gag gct gcc gag Leu His Asn Asp Pro Leu Ala Leu Gln Arg Leu Lys Glu Ala Ala Glu 260 265 270	816
aaa gca aaa att gaa ttg tca tct gca caa caa act gat gtc aat ttg Lys Ala Lys Ile Glu Leu Ser Ser Ala Gln Gln Thr Asp Val Asn Leu 275 280 285	864

cct tac att aca gca gat gct tct ggg cca aaa cat tta aat atc aaa 912  
 Pro Tyr Ile Thr Ala Asp Ala Ser Gly Pro Lys His Leu Asn Ile Lys  
 290 295 300

ttg act cgc gca aaa ctg gaa tct tta gtg gag aaa ctg gtt gag aga 960  
 Leu Thr Arg Ala Lys Leu Glu Ser Leu Val Glu Lys Leu Val Glu Arg  
 305 310 315 320

act att gag cca tgt aaa aca gct tta aaa gat gca ggg ctt act gtc 1008  
 Thr Ile Glu Pro Cys Lys Thr Ala Leu Lys Asp Ala Gly Leu Thr Val  
 325 330 335

tca caa atc aac gaa gtt att ctg gtc ggt ggg caa act cgt atg cct 1056  
 Ser Gln Ile Asn Glu Val Ile Leu Val Gly Gly Gln Thr Arg Met Pro  
 340 345 350

ctg gta caa aaa act gtg gaa gaa ttc ttt ggc aaa gaa cct cgt aag 1104  
 Leu Val Gln Lys Thr Val Glu Glu Phe Phe Gly Lys Glu Pro Arg Lys  
 355 360 365

gat gta aac cct gat gag gct gtg gca gta ggt gct gca att cag gca 1152  
 Asp Val Asn Pro Asp Glu Ala Val Ala Val Gly Ala Ala Ile Gln Ala  
 370 375 380

gca gta ttg tcc ggt gaa gtg aaa gac att ctt tta ctc gat gtc act 1200  
 Ala Val Leu Ser Gly Glu Val Lys Asp Ile Leu Leu Leu Asp Val Thr  
 385 390 395 400

cct ttg tct ctc ggt ata gaa acc atg ggt ggt gta atg acc aaa ctc 1248  
 Pro Leu Ser Leu Gly Ile Glu Thr Met Gly Gly Val Met Thr Lys Leu  
 405 410 415

att gaa aaa aac aca act att cct aca aag gct aca caa gtg ttt tca 1296  
 Ile Glu Lys Asn Thr Thr Ile Pro Thr Lys Ala Thr Gln Val Phe Ser  
 420 425 430

aca gcc gat gat aat cag aca gca gtg act gtg cat gta ttg cag ggt 1344  
 Thr Ala Asp Asp Asn Gln Thr Ala Val Thr Val His Val Leu Gln Gly  
 435 440 445

gaa aga gaa cag gct tca gca aac aag 1371  
 Glu Arg Glu Gln Ala Ser Ala Asn Lys  
 450 455

<210> 16  
 <211> 457  
 <212> PRT  
 <213> Legionella pneumophila

<400> 16

Leu Glu Gln Asn Arg Met Ala Lys Ile Ile Gly Ile Asp Leu Gly Thr  
 1 5 10 15

Thr Asn Ser Cys Val Ala Val Met Glu Gly Asp Lys Pro Lys Val Ile  
 20 25 30

Glu Asn Ser Glu Gly His Arg Thr Thr Pro Ser Ile Val Ala Phe Thr  
 35 40 45

Asp Asp Asn Glu Ile Leu Val Gly Gln Ser Ala Lys Arg Gln Ser Val  
 50 55 60

Thr Asn Pro Glu Lys Thr Leu Phe Ala Ile Lys Arg Leu Ile Gly Arg  
 65 70 75 80

Arg Phe Asp Asp Pro Ile Val Gln Lys Asp Ile Lys Met Val Pro Tyr  
 85 90 95

Lys Ile Met Lys Ala Asp Asn Gly Asp Ala Trp Val Arg Val Lys Asp  
 100 105 110

Gln Asp Lys Ala Pro Pro Gln Ile Ser Ala Glu Val Leu Arg Lys Met  
 115 120 125

Lys Lys Thr Ala Glu Asp Tyr Leu Gly Glu Glu Val Lys Glu Ala Val  
 130 135 140

Ile Thr Val Pro Ala Tyr Phe Asn Asp Ser Gln Arg Gln Ala Thr Lys  
 145 150 155 160

Asp Ala Gly Arg Ile Ala Gly Leu Glu Val Lys Arg Ile Ile Asn Glu  
 165 170 175

Pro Thr Ala Ala Ala Leu Ala Tyr Gly Met Asp Lys Lys Arg Gly Asp  
 180 185 190

Ser Val Ile Ala Val Tyr Asp Leu Gly Gly Gly Thr Phe Asp Ile Ser  
 195 200 205

Ile Ile Glu Ile Ala Glu Val Asp Gly Glu His Gln Phe Glu Val Leu  
 210 215 220

Ala Thr Asn Gly Asp Thr Phe Leu Gly Gly Glu Asp Phe Asp Leu Ala  
 225 230 235 240

Leu Ile Glu Tyr Leu Ala Ser Glu Phe Lys Lys Asp Thr Gly Ile Asp  
 245 250 255

Leu His Asn Asp Pro Leu Ala Leu Gln Arg Leu Lys Glu Ala Ala Glu  
 260 265 270

Lys Ala Lys Ile Glu Leu Ser Ser Ala Gln Gln Thr Asp Val Asn Leu  
275 280 285

Pro Tyr Ile Thr Ala Asp Ala Ser Gly Pro Lys His Leu Asn Ile Lys  
290 295 300

Leu Thr Arg Ala Lys Leu Glu Ser Leu Val Glu Lys Leu Val Glu Arg  
305 310 315 320

Thr Ile Glu Pro Cys Lys Thr Ala Leu Lys Asp Ala Gly Leu Thr Val  
325 330 335

Ser Gln Ile Asn Glu Val Ile Leu Val Gly Gly Gln Thr Arg Met Pro  
340 345 350

Leu Val Gln Lys Thr Val Glu Glu Phe Phe Gly Lys Glu Pro Arg Lys  
355 360 365

Asp Val Asn Pro Asp Glu Ala Val Ala Val Gly Ala Ala Ile Gln Ala  
370 375 380

Ala Val Leu Ser Gly Glu Val Lys Asp Ile Leu Leu Leu Asp Val Thr  
385 390 395 400

Pro Leu Ser Leu Gly Ile Glu Thr Met Gly Gly Val Met Thr Lys Leu  
405 410 415

Ile Glu Lys Asn Thr Thr Ile Pro Thr Lys Ala Thr Gln Val Phe Ser  
420 425 430

Thr Ala Asp Asp Asn Gln Thr Ala Val Thr Val His Val Leu Gln Gly  
435 440 445

Glu Arg Glu Gln Ala Ser Ala Asn Lys  
450 455