



(51) International Patent Classification:

A61K 39/02 (2006.01)

(21) International Application Number:

PCT/EP2018/075634

(22) International Filing Date:

21 September 2018 (21.09.2018)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

17306235.7 21 September 2017 (21.09.2017) EP

(71) Applicant: **CENTRE DE COOPERATION INTERNATIONALE EN RECHERCHE AGRONOMIQUE POUR LE DEVELOPPEMENT (CIRAD)** [FR/FR]; 42 rue Scheffer, 75016 PARIS (FR).

(72) Inventors: **MEYER, Damien**; 13 res. l'orée de Bois Sergent Montebello, 97170 Petit Bourg GUADELOUPE (FR). **GORDON, Jonathan**; Mezenlaan 36 Heverlee, 3001 LEUVEN (BE). **VACHIERY, Nathalie**; 15 rue des Wisigoths, 34270 Saint Mathieu de Treviers (FR). **MARTINEZ, Dominique**; 5 route de Quissac, 30610 Sauve (FR).

(74) Agent: **GEVERS & ORES**; 41, avenue de Friedland, 75008 PARIS (FR).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: LIVE ATTENUATED BACTERIAL STRAIN AND ITS USE AS A VACCINE

(57) Abstract: The present invention relates to a vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene.



WO 2019/057909 A1

## LIVE ATTENUATED BACTERIAL STRAIN AND ITS USE AS A VACCINE

### FIELD OF THE INVENTION:

The present invention relates to a vaccine composition comprising a live attenuated bacterial strain and its use as a vaccine.

### 5 BACKGROUND OF THE INVENTION:

Vaccines have been one of the biggest success stories of modern medicine. There is arguably no single preventive health intervention more cost-effective than immunization. Time and again, the international community has endorsed the value of vaccines and immunization to prevent and control a large number of infectious  
10 diseases and, increasingly, several chronic diseases that are caused by infectious agents.

WHO estimates that at least 10 million human deaths were prevented between 2010 and 2015 thanks to vaccinations delivered around the world. Many millions more lives were protected from the suffering and disability associated with diseases such as  
15 pneumonia, diarrhea, whooping cough, measles, and polio.

Thank to vaccination global measles mortality has declined by 79% and the poliomyelitis is closer to be eradicated. Moreover, vaccines can help to limit the spread of antibiotic resistance.

In the last decades, significant advances in the human and veterinary vaccines and  
20 their production have been made. One of the main progresses is the improved production efficiency of live attenuated vaccines.

However, there is a still a need for new live attenuated vaccines against human or veterinary pathogens but many obstacles exist that have impeded their development. One of the obstacles to the design of live attenuated vaccine include notably  
25 understanding pathogen genetic and antigenic variability, variable host immune responses in order to identify protective antigens, and immunogenicity and identification of the antigens that stimulate protective immunity.

Therefore, it is an object of the invention to overcome at least some of the deficiencies of the existing vaccines and to provide new live attenuated vaccines.

- In particular, it is another object of the invention to provide, in at least one embodiment, a live attenuated strain which exhibits an attenuated pathogenicity while  
5 maintaining its ability to colonize and which induces a strong protective immunity with limited clinical signs after vaccination and no clinical sign after challenge.

It is also another object of the invention to provide, in at least one embodiment, a method for producing a vaccine which can be adapted to a large number of strains in the same family.

- 10 It is also another object of the invention to provide, in at least one embodiment, a method for producing a vaccine which is easy to implement and which is relatively low in production costs.

#### **SUMMARY OF THE INVENTION:**

- Now, the inventors have found that by inactivating or deleting the *ntrX* gene of a wild  
15 bacterial strain they obtain a live strain with an attenuated virulence. Moreover, the inventors have also shown that this live attenuated strain was able to induce immunity sufficient for ensuring protection. Indeed, in order to provide a vaccine, it is not sufficient to provide a live attenuated strain, it is necessary that this strain also induces a strong immune response.

- 20 A subject of the present invention is therefore a vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene.

The present invention also relates to a vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene for use to induce a protective immune response against the bacterial strain.

- 25 The present invention also relates to a vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene for use as a vaccine.

The present invention also relates to a vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene for use in preventing an infection caused by the bacterial strain.

5 In yet another aspect, the present invention relates to method for producing a vaccine composition for use against a bacterial strain comprising a step of:

- inactivating or deleting the *ntrX* gene of the bacterial strain, thereby obtaining an attenuated bacterial strain.

## DETAILED DESCRIPTION OF THE INVENTION

### Vaccine composition

10 The present invention relates to vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene. This bacterial strain with a deleted or inactive *ntrX* gene is a live attenuated strain.

15 *ntrX* is transcriptional regulator whose product is Nitrogen assimilation regulatory protein NtrX. It functions as a Signal transduction mechanism (T.COG2204) and contains a "Signal receiver domain; originally thought to be unique to bacteria (CheY, OmpR, NtrC, and PhoB), now recently identified in eukaryotes ETR1 *Arabidopsis thaliana*; this domain receives the signal from the sensor partner in a two-component systems; cd00156".

20

In one embodiment, the *ntrX* gene of the invention encodes a NtrX protein having the amino acid sequence of SEQ ID NO: 1.

25 The NtrX protein having the amino acid sequence of SEQ ID NO: 1 is the NtrX of *Ehrlichia ruminantum* Gardel. The NtrX protein of *Ehrlichia ruminantum* Welgevonden and *Ehrlichia ruminantum* Senegal have also the amino sequence of SEQ ID NO: 1.

The amino acid sequence of the NtrX protein of *Ehrlichia ruminantum* Gardel as well as the nucleic acid sequence of the *ntrX* gene of *Ehrlichia ruminantum* Gardel are given in the Table 1 below.

amino acid sequence of the NtrX protein of <i>Erllichia ruminantum</i> Gardel	SEQ ID NO: 1	MAQDFEMSKERLYI SEVLVVDDEVD IRNLIKDI LDDNYVTKLAVDGLSAIKMAYEKEP DVVLID IWLGRSDIDGLSVLEKLKERYPYLPVIMISGHGNIATAVKSLHMGAYDYIEKP FTEGRLLKLVVKRAIESGRLLRRENDELKSAFEDYEIVGNSPVIRNLRSMINKAATSSRI LITGSPGVGKEVVARLIHKKSKGYDTPFI SMYSMLPANNYLNVNIFGSEESNNILSHRV PPHIGLIEQANHGTFLF IDEVTDLRDYDTQLRLRLQLQEGKIYRENSKIPVSDVRIIVSS SKDIESEVKAGRFCEDLYYRLNVLP IRVPSLVEYCTDIPELCRYFMNSICKKIGLCTHV LSDEALIAMQSYEWPGNLRQLRNVIEWILIMKSPKEMITAKDLPVDIVSNSPINDVLSA KVI SVP L R K A R E E F E R Q Y L K T Q L S R F G G N V S R T A E F V G M E R S A L H R K L K I L G L C N V S E
nucleic acid sequence of the ntrX gene of <i>Erllichia ruminantum</i> Gardel	SEQ ID NO: 11	ATGGCAGGATTTTGAAAATGTCCAAGGAAAAGATTGTATATTTCTGAAAGTATTAGTTGT TGATGATGAAAGTTGATATCAGAAAATCTAATAAAAAGATATATTAAAGTGATGATAAATTATG TCACTAAATTAGCAGTTGATGGTTTATCCGCGATCAAGATGGCTTATGAAAAAGAGCCT GATGTGTATTATTGGATATATGGTTAAGAGGATCTGATATTGATGGATTAAAGTGTACT GGAGAACTTAAAGAAAAGGTAICCTTATTTGCCCTGTTATTATGATTAGTGGGCAATGGTA ATATTGCCACTGCTGTAAAGTCTCTGCATATGGTGGCTTATGATTATATAGAAAAGCCT TTTACAGAAAGGAAAGATTAAAGTTAGTTGTAAGAGAGAGCTATAGAGTCTGGTAGATTACG TAGAGAAAATGATGAGTTGAAAATCAGCATTTGAGGATTATGAAAATAGTCGGTAACTCCC CTGTTATACGTAATTTGAGAAATGATTAATAAAGCAGCTACTACATCGAGTCGTATA CTCATTTACTGGTTCGCCAGGTGTGAAAAGGAAAGTAGTTGCTAGGCTAATACATAAAAAA ATCCAAGGGGTATGATACTCCATTATATCTATGTACTCATCTATGCTACCAGCTAATA ATTACTTGGTTAAATATATTGGTAGTGAGGAAAAGTAATAATATTGTCTCATAGAGTA CCTCCTCATATTGGAATTATAGAGCAAGCAAAATCATGGTACGTTATTATAGATGAAGT AACAGATTACGATACGATACGCAATTAAAGATTACTCAGATTATTACAGGAGGAAAAA TATATAGGAAAAATAGTAAGATTCCCTGTTAGTATAGATGTGAGAAATTATTGTCTTCT TCCAAGATATTGAAAGTGAAAGTAAAGCTGGTAGGTTTGTGAGGATTATATTATAG ATTAAATGTCCTTCCAATTAGAGTACCGCTTTTAGTAGAATATTGTACAGATATACCGG AATTGTGTAGGTATTTTATGAATAGCATCTGTAAAAAAAATAGGTTTGTGTACTCATGTA TTAAGTGATGAAGCTTTAAATAGCAATGCAGTCATATGAATGGCCAGGTAACCTTAAGACA ATTACGTAATGTTATAGAAATGGATTTAATTAATGAAATCTCCTAAGGAGATGATTACAG CAAAAGATTTACCAGTAGATATAGTATCTAAATTCGCCTATTAAATGATGTTTTAAAGTGCT AAAATTATTTCTGTACCATTACGTAAAGCTCGTGAAGAAATTTGAAAAGACAGTATTTAAA AACTCAGTTATCTCGTTTTGGAGGTAATGTATCAGGAACTGCTGAAATTTGTTGGAATGG AACGTTACAGCATTACACCGTAAATTTGAAAAATTTCTTGGATTGTGTAATGTTTCTGAAATAA

Table 1

In another embodiment, the *ntrX* gene of the invention is an active homologue of the *ntrX* gene which encodes a NtrX protein having the amino acid sequence of SEQ ID NO: 1 ; said active homologue encodes a NtrX protein having an amino acid sequence with at least 50 % of identity with SEQ ID NO: 1.

- 5 Preferably, said active homologue encodes a NtrX protein having a amino acid sequence with at least 60%, more preferably at least 70%, more preferably at least 75%, more preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% of identity with SEQ ID NO: 1.
- 10 For the purposes of comparing two closely-related polynucleotide or polypeptide sequences, the "% identity" between a first sequence and a second sequence may be calculated using an alignment program, such as BLAST® (available at [blast.ncbi.nlm.nih.gov](http://blast.ncbi.nlm.nih.gov)) using standard settings. The % identity is the number of identical residues divided by the number of residues in the reference sequence,
- 15 multiplied by 100. The % identity in the above table is calculated by this methodology.

Alternatively, the % identity may also be calculated by dividing the number of identical residues by the number of aligned residues and multiplying by 100.

- Alternative methods include using a gapped method in which gaps in the alignment, for example deletions in one sequence relative to the other sequence, are accounted
- 20 for in a gap score or a gap cost in the scoring parameter (for more information, see the website [blast.ncbi.nlm.nih.gov](http://blast.ncbi.nlm.nih.gov)).

- In an embodiment, the inactive *ntrX* gene is an inactive mutant of a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ ID NO: 1 or of an active homologue thereof encoding a NtrX protein having a amino acid sequence with
- 25 at least 50 %, 60%, 70%, 75%, 80%, 85%, 90% or 95% of identity with SEQ ID NO: 1. In another embodiment, it is a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ ID NO: 1 or an active homologue thereof encoding a NtrX protein having a amino acid sequence with at least 50 %, 60%, 70%, 75%, 80%, 85%, 90% or 95% of identity with SEQ ID NO: 1 which is deleted.

Preferably, the bacterial strain is genetically engineered so as to delete or to inactivate the *ntrX* gene.

Advantageously, only the *ntrX* gene is inactivated or deleted.

5

In one embodiment, the vaccine composition comprises a bacterial strain with a deleted *ntrX* gene.

10 In this embodiment, the bacterial strain is a mutant strain from a bacterial strain, preferably a wild and/or pathogen strain, wherein the *ntrX* gene has been deleted.

The deletion of a gene can be carried out by any methods known by the skilled person.

In one embodiment, the *ntrX* gene which is a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ ID NO: 1 is deleted.

15 In another embodiment, the *ntrX* gene which is an active homologue thereof encoding a NtrX protein having a amino acid sequence with at least 50 %, 60%, 70%, 75%, 80%, 85%, 90% or 95% of identity with SEQ ID NO: 1 is deleted.

20 In another embodiment, the vaccine composition comprises a bacterial strain with an inactive *ntrX* gene. An inactive *ntrX* gene may be a *ntrX* gene which encodes a non functional NtrX protein or no NtrX protein at all.

In this embodiment, the bacterial strain is a mutant strain from a bacterial strain, preferably a wild and/or pathogen strain, wherein the *ntrX* gene has been inactivated.

25 The inactivation of a gene can be carried out by any methods known by the skilled person.

By way of example, inactivation of said *ntrX* gene can be obtained by mutagenesis and selection of the mutants having lost the NtrX protein activity. Mutagenesis can be performed for instance by targeted deletion of a portion of *ntrX* coding sequence or by  
30 targeted insertion of an exogenous sequence within said coding sequence. Mutagenesis may be also a deletion, an insertion or a replacement of at least one

nucleic acid. For example, the inactive mutant of a *ntrX* gene may be a *ntrX* gene wherein a codon stop is inserted or wherein a region is deleted and/or rearranged. The inactive mutant of a *ntrX* gene may also be a *ntrX* gene wherein at least one nucleic acid has been inserted, for example an inactivating point mutation or an  
5 insertion leading to a frameshift mutation.

The inactivation can also be performed by random chemical or physical mutagenesis, followed by screening of the mutants within the *ntrX* gene. Methods for high throughput mutagenesis and screening are available in the art.

Preferably, a strain with an inactive *ntrX* gene is a strain wherein the *ntrX* gene is not  
10 expressed.

In one embodiment, the inactive *ntrX* gene is an inactive mutant of a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ ID NO: 1.

In another embodiment, the inactive *ntrX* gene is an inactive mutant of an active homologue of a *ntrX* gene encoding a NtrX protein having the amino acid sequence of  
15 SEQ ID NO: 1 ; said active homologue encodes a NtrX protein having a amino acid sequence with at least 50 % of identity with SEQ ID NO: 1.

Preferably, said active homologue of a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ ID NO: 1 encodes a NtrX protein having a amino acid sequence with at least 60%, at least 70%, more preferably at least 75%, more  
20 preferably at least 80%, more preferably at least 85%, even more preferably at least 90%, and most preferably at least 95% sequence identity with SEQ ID NO: 1.

In one embodiment, the bacterial strain with a deleted or an inactive a *ntrX* gene is an *Alphaproteobacteria* strain, preferably a *Rickettsiales* strain, more preferably an *Anaplasmataceae* strain.

25 *Anaplasmataceae* are genetically related small Gram-negative pleomorphic cocci. *Anaplasmataceae* comprises the strains of genus selected from the group consisting of *Anaplasma*, *Ehrlichia*, *Neorickettsia*, and *Wolbachia*. *Anaplasmataceae* are obligatory intracellular bacteria which are mainly transmitted by arthropods, most



frequently ticks, lice and mites, and cause major illnesses such as ehrlichiosis and anaplasmosis.

Vaccines directed against strains of *Anaplasmataceae* remain limited. In particular, only few cases of live attenuated vaccine directed against *Anaplasmataceae* have  
5 been reported and the possible mechanism of attenuation of such strains has been studied in only few cases.

For example, in an attempt to produce live attenuated vaccines against *Ehrlichia ruminantium*, a tick-borne intracellular pathogen of ruminants that causes heartwater, 3 attenuated strains from Guadeloupe (Gardel), South Africa (Welgevonden) and  
10 Senegal have been generated *in vitro* by passaging the virulent bacterium in bovine endothelial cells, in addition to canine macrophage-monocyte cells for Welgevonden strain (Jongejan 1991; Zweggarth et al. 2005; Pilet et al. 2012; Marcelino et al. 2015).

However, usage of these attenuated strains as vaccines has remained limited due to constraints associated with storage (no cold chain disruption) and intravenous  
15 administration. Their mechanisms of attenuation are unknown which is problematic considering the risk of virulence recovery. Moreover, attenuated vaccines, as for any other vaccine against heartwater, protect against homologous challenges but confer limited protections against heterologous strains and co-occurrence of different strains in infected animals is common (Zweggarth et al. 2005; Frutos et al. 2006).

20 Thus, it is also an object of the invention to provide, in at least one embodiment, a vaccine against disease caused by *Anaplasmataceae* and in particular against ehrlichiosis.

In particular, it is an object of the invention to overcome at least some of the deficiencies of the existing therapies and prophylaxis of disease caused by  
25 *Anaplasmataceae*.

In order to overcome the drawbacks of the prior art vaccines against *Anaplasmataceae*, the present invention provides a vaccine composition comprising an *Anaplasmataceae* strain with a deleted or an inactive *ntrX* gene.

As shown in the table 2 below, the *NtrX* protein is highly conserved among the  
30 *Anaplasmataceae* family from 84% to 93% of identity.

<i>Anaplasmataceae</i> strain	Disease caused by the strain	SEQ ID NO of the NtrX protein	Amino acid sequence of NtrX	% of identity with SEQ ID NO : 1
<i>Ehrlichia ruminantium</i> Gard el	ehrlichiosis	SEQ ID NO : 1	MAQDFEMSKERLYISEVLVDDDEVDIRNLKIDILSDDDNYVTKLAVDGLSAIKMAY EKEPDVLLDIWLGSDIDGLSVLEKLERYPYLPVIMISGHGNIATAVKSLHMG AYDYIEKPFTEGRLLVVKRAIESGRLRRENDELKSAFEDYEIVGNSPVIRNLRS MINKAATTSSRILITGSPGVGKEVVARLIHKKSKGYDTPFISYSSMLPANNYL NIFGSEESSNLSHRVPPHIGIEQANHGTLFIDEVTDLYRNLRLRLQEGKI RENSKIPVSDVRIIVSSSKDIESEVAGKFCEDLYRNLVLPVPSLVEYCTDI ELCRYFMNSICKKIGLCTHLSDEALIAMQSYEWPGNLRQLRNVIEWILIMKSPK EMITAKDLPVDIVSNSPINDVLSAKVISVPLRQAREEFERQYLKQLSRFGGNSV RTAEFVGMSALHRKLVGLGNCVSE	100%
<i>Ehrlichia chaffeensis</i>	ehrlichiosis	SEQ ID NO : 2	MAQNFEMSKERLYISEVLVDDDEVDIRNLKIDILSDDDNYVTKLAVDGLSAIKMAY EKEPDVLLDIWLGSDIDGLSVLEKLERYPYLPVIMISGHGNIATAVKSLHMG AYDYIEKPFTEGRLLVVKRAIESGRLRRENDELKSTFEDYEIVGNSPVIRNLRS MINKAATTSSRILITGSPGVGKEVVARLIHKKSKGYDTPFISYSSMLPANNYL NIFGSEESSNLSHRVPPHIGIEQANHGTLFIDEVTDLYRNLRLRLQEGKI RENSKIPVSDVRIIVSSSKDIESEVAGKFCEDLYRNLVLPVPSLVEYCTDI PELCRYFMSSICKKIGLCTHLSDEALIAMQSYEWPGNLRQLRNVIEWILIMKSPK EIITAKDLPVDIVSNSPINDVLSAKVISVPLRQAREEFERQYLKQLSRFGGNSV TAEFVGMSALHRKLVGLGNCVSE	96%
<i>Ehrlichia</i> sp. HF	ehrlichiosis	SEQ ID NO : 3	MAQSFEMSKERLYISEVLVDDDEVDIRNLKIDILSDDDNYVTKLAVDGLSAIKMAY EKEPDVLLDIWLGSDIDGLSVLEKLERYPYLPVIMISGHGNIATAVKSLHMG AYDYIEKPFTEGRLLVVKRAIESGRLRRENDELKSTFEDYEIVGNSPVIRNLRS MINKAATTSSRILITGSPGVGKEVVARLIHKKSKGYDTPFISYSSMLPANNYL NIFGSEESSNLSHRVPPHIGIEQANHGTLFIDEVTDLYRNLRLRLQEGKI RENSKIPVSDVRIIVSSSKDIESEVAGKFCEDLYRNLVLPVPSLVEYCTDI PELCRYFMNSICKKIGLCTHLSDEALIAMQSYEWPGNLRQLRNVIEWILIMKSP KEIITAKDLPVDIVSNSPINDVLSAKVISVPLRQAREEFERQYLKQLSRFGGNSV KTAEFVGMSALHRKLVGLGNCVSE	96%
<i>Ehrlichia muris</i>	ehrlichiosis	SEQ ID NO : 4	MAQNFEMSKERLYISEVLVDDDEVDIRNLKIDILSDDDNYVTKLAVDGLSAIKMAY EKEPDVLLDIWLGSDIDGLSVLEKLERYPYLPVIMISGHGNIATAVKSLHMG AYDYIEKPFTEGRLLVVKRAIESGRLRRENDELKSTFEDYEIVGNSPVIRNLRS MINKAATTSSRILITGSPGVGKEVVARLIHKKSKGYDTPFISYSSMLPANNYL	96%

<i>Ehrlichia canis</i>	ehrlichiosis	SEQ ID NO : 5	NIFGSESSNILSHRVPPHIGIEQANHGTLFIDEVTDRLRYDTQLRLLRLQEGKIY RENSKIPVGVVDVRIIVSSSKDIENEVKAGKFCEDLYYRLNVLPPIRVPSLVEYCTDI PELCRYFMNSICKIGLCTYILSDEALIAMQSYEWPGNLRQLRNVIEWILIMKSPK EITAKDLPVDIVSNSPINDVLSAKVISIPLRQAREEFKQYLTQLSRFGGNSVK TAEFVGMEERSALHRKLVGLGNISE	94%
<i>Ehrlichia minasensis</i>	ehrlichiosis	SEQ ID NO : 6	MAQDFEMSKERLYISEVLVDDDEGDIRNLKIDILSDDNHGTCLAVDGLSAIKMA YEKEPDVLLDWLKGSDIDGLSVLEKLERYPHLPVIVISGHGNIATAVKSLHIG AYDIEKPTTESRLKLVKRAIESGRLRRENDELKSAFEDYEVGSSSVIKNLR MVNKAATSSRLITGSPGVGKEVVARLIHKKSQGYDTPFISMYSSMLPANNYL NIFGSESSNILSHRVPPHIGIEQANHGTLFIDEVTDRLRYDTQLRLLRLQEGKIY RENSKIPVGVVDVRIIVSSSKDIENEVRAGKFCEDLYYRLNVLPPIRVPSLVEYCTDI PELCRYFMNSICKKMGCLTHILSDEALIAMQSYEWPGNLRQLRNVIEWILIMKSP KEVITAKDLPVDIVSNSPINDVLSAKVISIPLRQAREEFERQYLTQLSRFGGNSV SKTAEFVGMEERSALHRKLVGLGNISE	94%
<i>Neoehrlichia lotoris</i>	ehrlichiosis	SEQ ID NO : 7	MLKAKRFGCISELIVDDEADIRNLIRDILSDENYATKSSVDGLSAIKLAYEKEPDV VLLDWLKGSDVDGLSVLEKLERYPYLPVIMISGHGNAVAVKSLHKGAYDYIE KPFTENRLKLVKRAIESGRLRRENDELKSAFEDYEFIGNSPVIRNLRNIEKASI TSSRVLITGASGTGKEVVARLIHKKSQGYDTPFISMYSSMLPANNYL EYNSLPLNKLIPNIGIEQANHGTLFIDEVTEVRYDVQLRLLRLQEGKIYRENG KIPIDARVIAASSRVIKDEVKLGRFCEDLYYRLNVLPVPSLCEYCSIDIPEICR YFMKSICKKMGCLADHVISNDIAIAMQSYWPGNLRQLRNVLEWVLMKSSGIIIT VEDLPAEISGSVNDLTSARMVSVPLRQAREEFERQYLTQLSRFGGNSVSKTA EFGVMEERSALHRKLVGLGNIS	78%
<i>Anaplasma phagocytophilum</i>	anaplasmosis	SEQ ID NO : 8	MSKVKRFYIPEVLIVDDEADIRNLIRDILSDENYATKSSVDGLSAIKLAYEKEPDV DVLDDWLKGSDIDGLSVLEKLERYPYLPVIMISGHGNIATAVKSLHIGAYDYI EKPTENRLKLVKRAIESGRLRRENDELKSAFEDYEFIGNSPVIRNLRNIEKASI AASFTRILITGAPGTGKEVVARLIHKKSQGYDTPFISMYSSMLPANNYL GENSSLOHVPVSHVIGIEQANHGTLFIDEVTDRLRYDAQLRMLRLQEGKIYRES GKIPVIVDTRVIASSSKIMEDEVRLKFCEDLYYRLNVFPPIRVPSLSEYCADLPEI CEYMMKSICKKMLFPRAISEEIVAMQSYWPGNLRQLRNVLEWVLMQTKH DVTIVDDLPVIAEIVNGAPINSVFTHIVSAPLRKAREEFERQYLTQLSRFGGNSVR	74%

<i>Anaplasma centrale</i>	anaplasmosis	SEQ ID NO : 9	<p>TAEFIGMERSALHRKLMGLFTG</p> <p>MGFMSRAKRFYTPPEVLIVDDDEADLRAMVQDILSDDNYVTNVSHDGLTAIKLAYE          REPDWLLDIWLGSDIDGLSVLEKLRCRYPHLPVVISGHGNIATAVKSLHIGAY          DYIEKPFTENRLKLVKRAVESGRLKRENRELRSLFEDYEIVGSSPQIKNLRSTI          SKAASTCSRILITGAPGTGKEVVARFVHRKFRGYNSSFVSCFPSI          LPESYLENIFGSESGSEALPNVVPHSVGIIEQANHGTGFIDEVTNLRYDAQMRL          MRLFQEGRIYREHGKSPVTVTRVIASSSRAIEEEVKLGRFCEDLYYRLGVFFPIR          VPSLSEYCVDLPEVCEYLMRSICRKMGLLPRIPSEDAIVAMQSYSWPGNLRQL          RNVLEWILIMQSSKDIITVNDLPAEIASGSPINNAFTNIVSAPLREAREEFERHYL          RTQLSRFGGSYSKTAEFIKMERSALHRKLMGLCGS</p>	70%
<i>Anaplasma marginale</i>	anaplasmosis	SEQ ID NO : 10	<p>MGFMSRAKRFYTPPEVLIVDDDEADLRAMVQDILSDDNYVTNVSHDGLTAIKLAYE          REPDWLLDIWLGSDIDGLSVLEKLRYPHLPVVISGHGNIATAVKSLHIGAY          DYIEKPFTESRLKLVKRAVESGRLKRENRELRSLFEDYEIVGSSPQIKNLRSTIS          KAASTCSRILITGAPGTGKEVVARFVHRKFRGYNSSFVSCFPSI          GSESGNGALPNVVPHSVGIIEQANHGTGFIDEVTNLRYDAQMRLMRLQEGRIY          RENGKNPVTIDTRVIASSSRAIEEEVKLGRFCEDLYYRLGVFFPIRVPSEYCVDI          PEVCEYMMRSICRKMGLSPRISEDAIVAMQSYSWPGNLRQLRNVLEWILIMQ          SSKDIITIDDLPAEIVSGSPINNVFTHIVSAPLRKAREEFERHYLRTQLSRFGGSV          SKTAEFIGMERSALHRKLMGLCGS</p>	72%

Table 2: % of identity of NtrX protein between different strains of *Anaplasmataceae*

Thus, the inactive *ntrX* gene may be an inactive mutant of a *ntrX* gene encoding a NtrX protein having the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 and SEQ ID NO:10.

In one embodiment, the inactive *ntrX* gene may be an inactive mutant of a *ntrX* gene encoding a NtrX protein having at least 50%, 60%, 65%, 70%, 75%, 80%, 85% 90% or 95% of the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 and SEQ ID NO:10.

In another embodiment wherein the *ntrX* gene is deleted, it is a *ntrX* gene encoding a NtrX protein having the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 and SEQ ID NO:10 which is deleted.

In one embodiment, the deleted *ntrX* gene may be a *ntrX* gene encoding a NtrX protein having at least 50%, 60%, 65%, 70%, 75%, 80%, 85% 90% or 95% of the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 and SEQ ID NO:10.

In a preferred embodiment, the bacterial strain is selected from the group consisting of *Anaplasma* strain, *Ehrlichia* strain, *Wolbachia* strain and *Neorickettsia* strain.

In a preferred embodiment, the bacterial strain is an *Ehrlichia* strain.

The *Ehrlichia* strain may be selected from the group consisting of *E. canis*, *E. chaffeensis*, *E. ewingii*, *E. muris*, and *E. ruminantium*.

Among the species of the genus *Ehrlichia*, some are species pathogenic for humans and animals such as *E. chaffeensis*, responsible for Human monocytic ehrlichiosis or *E. canis*, the causing agent of canine monocytic ehrlichiosis.

- Ehrlichia ruminantium*, is also pathogenic. This tick-borne intracellular pathogen of ruminants that causes heartwater. Heartwater is an economically important tropical disease of ruminants. It can cause up to 80% mortality in susceptible animals. This disease is present in Sub-saharan Africa, islands in the Indian Ocean and the Caribbean, and is threatening the American mainland. It is listed within the 12<sup>th</sup> most important transboundary animal diseases for US homeland security department.

Preferably, the *Ehrlichia* strain is an *E. ruminantium* strain.

*E. ruminantium* strain may be selected from the group consisting of Senegal, Gardel and Welgevonden *E. ruminantium* strains.

- 10 Preferably, the *E. ruminantium* strain is a Senegal *E. ruminantium* strain.

In one embodiment, the vaccine composition comprises a bacterial strain which is not an *E. ruminantium* strain.

In one embodiment, the vaccine composition comprises a bacterial strain which is not a Senegal *E. ruminantium* strain.

- 15 In one embodiment, the vaccine composition comprises a bacterial strain which is not an *E. canis* strain.

In one embodiment, the vaccine composition comprises a bacterial strain which is not a *E. canis Israel* strain.

In another embodiment, the bacterial strain is not an *Anaplasmataceae* strain.

- 20 The bacterial strain may also be selected from the group consisting of a *Rickettsia* strain, an *Orientia* strain, a *Bartonella* strain and a *Brucella* strain.

Typically, the vaccine composition may comprise a pharmaceutically acceptable excipient.

- 25 If desired, the vaccine composition according to the invention may also comprise an adjuvant. Examples of suitable compounds and compositions with adjuvant activity are well known in the art. Adjuvant may be for example selected from the group

consisting of: inorganic adjuvants, organic adjuvants, oil-based adjuvants, cytokines, particulate adjuvants, liposomes, virosomes, bacterial adjuvants, synthetic adjuvants, synthetic polynucleotides adjuvants, immunostimulatory oligonucleotides containing unmethylated CpG dinucleotides and a combination thereof.

- 5 In addition to the live attenuated bacterial strain, the vaccine composition can also comprise one or more additional active immunizing agent in order to produce a vaccine composition capable of inducing immunity against a number of different pathogens. The vaccine composition may comprise several live attenuated bacterial strains.

10 **Method of treatment**

The present invention also relates to a vaccine composition according to the invention for use to induce an immune response against a bacterial strain, in particular an *Anaplasmataceae* strain.

- 15 As discussed above, the inventors have found that by inactivating the *ntrX* gene in a wild bacterial they obtain a live strain with an attenuated virulence which induces a protective immune response against the wild bacterial strain in the subject in which it is administered.

As used herein, "subject" refers to a human or animal that may benefit from the administration of a vaccine composition as recited herein.

- 20 The subject may be a human or a non human animal. For example, the non human animal is selected from the group consisting of a bovine, an ovine, an equine, a feline, a caprine or a canine.

- 25 The present invention also relates to a method for inducing an immune response against a bacterial strain, preferably an *Anaplasmataceae* strain, in a subject in need thereof comprising administering to said subject an effective amount of the vaccine composition of the invention.

The present invention also relates to a vaccine composition according to the invention for use as a vaccine, preferably a prophylaxis vaccine.

The inventors have shown that administering to a subject a vaccine composition comprising bacterial strain attenuated by the inactivation or the deletion of the *ntrX* gene prevents or alleviates the effects of an infection caused by the wild bacterial strain.

- 5 Also provided, a method for protecting or treating a subject against disease caused by a bacterial strain comprising administering to said subject a bacterial strain with a deleted or inactive *ntrX* gene.

In a preferred embodiment, the vaccine composition is for use in preventing and/or treating an infection caused by bacterial strain, preferably an *Anaplasmataceae* strain.

- 10 The infection caused by an *Anaplasmataceae* strain is preferably ehrlichiosis or anaplasmosis, more preferably ehrlichiosis.

- In order to prevent and/or treat an infection caused by a given wild bacterial strain, it would be preferred to use a vaccine composition comprising the given bacterial strain which has been attenuated by deleting or inactivating its *ntrX* gene thereby providing  
15 an attenuated mutant strain of the given bacterial. For example, for preventing or treating ehrlichiosis, it would be preferred to use an *Ehrlichia* strain with a deleted or an inactive *ntrX* gene. In the same way, for preventing or treating ehrlichiosis caused by Senegal Ehrlichia strain, it would be preferred to use a Senegal Ehrlichia strain with a deleted or an inactive *ntrX* gene.

- 20 In another embodiment, in order to prevent and/or treat an infection caused by a given bacterial strain, it could be use a vaccine composition comprising a bacterial strain close from the given bacterial strain, for example a strain from the same genus or species, with a deleted or inactive *ntrX* gene.

- In yet another aspect, the present invention relates to a method for preventing and/or  
25 treating a subject in need thereof against an infection caused by a bacterial strain comprising administering to said subject an effective amount of the vaccine composition of the invention.

The vaccine composition may be administered by intramuscular, intradermal, subcutaneous or intranasal inoculation or injection. The vaccine composition is



administered in an amount, which is effective to protect the subject against challenge, by a virulent bacterial strain. This amount may vary according to the subject being inoculated, taking into consideration the size and weight of the subject. The vaccine composition according to the invention comprises preferably an effective dosage of the live attenuated bacterial strain as the active component, i.e. a sufficient amount of the bacterial live attenuated strain that will induce immunity in the vaccinated subjects, against challenge by the corresponding virulent bacterial strain.

**Method for producing a vaccine composition**

In yet another aspect, the present invention provides the use of a bacterial strain, preferably an *Anaplasmataceae* strain, with a deleted or inactive *ntrX* gene for the manufacture of a vaccine.

The present invention also relates to a method for producing a vaccine composition for use as a vaccine against a bacterial strain comprising a step of:

- inactivating or deleting the *ntrX* gene of the bacterial strain, thereby obtaining an attenuated bacterial strain.

Preferably, the bacterial strain is an *Anaplasmataceae* strain.

Examples of *Anaplasmataceae* strain are given herein above. In a preferred embodiment, the *Anaplasmataceae* strain is an *Ehrlichia* strain.

The *Ehrlichia* strain may be selected from the group consisting of: *E. canis*, *E. chaffeensis*, *E. ewingii*, *E. muris*, and *E. ruminantium*.

The *ntrX* gene of the bacterial strain may be deleted or inactivated for example by mutation. Thus, the *ntrX* gene may be mutated by point mutation, by insertion or by deletion of one or more nucleotides so as to for example interrupting the open reading frame of the *ntrX* gene.

Techniques of inactivation of a gene are well known from the person skilled in the art. Techniques which may be used to inactivate the *ntrX* gene, may be for example site directed mutagenesis or homologous recombination. It also includes the use of molecules that trigger preferentially the *ntrX* gene for inactivation.

The method for producing a vaccine composition may also comprise a step of mixing the attenuated strain with a pharmaceutically acceptable excipient.

The method for producing a vaccine composition according to the invention can also comprise steps conventionally used in the manufacture of the commercially available vaccine composition based on live attenuated strains.

The invention will be further illustrated by the following figures and examples. However, these examples should not be interpreted in any way as limiting the scope of the present invention.

## FIGURES

Figures 1 and 2 show an evidence of gene conversion between *ntrX* and its inverted duplicate.

Figure 1 shows an alignment showing the duplicated region (top four rows) of *ntrX* gene and the matching parental *ntrX* region (next four rows) in *Welgevonden*, *Gardel* and *Senegal* virulent and attenuated strains and in *E. chaffeensis* (last row) where there is no inverted duplicate. Alignment positions that show a closer relationship between *ntrX* and the duplicate within a strain than between strains, including the 4bp deletion in *Senegal* strain are bordered by boxes.

Figure 2 shows a phylogram (left) depicting the expected relationship between *ntrX* and its duplicate (star) given that it is present in all *E. ruminantium* strains. The structure of the phylogram is based on the interstrain relationship reconstructed in the maximum likelihood phylogeny (right) by concatenating 54 orthologous ribosomal proteins from the strains and *E. chaffeensis*.

## EXAMPLES

### **Material and Methods:**

#### **Genomics**

*DNA extraction and purification*

*Ehrlichia ruminantium* Senegal strain passage 7 (80% lysis at day 21 p.i.) and 63 and 64 (80% lysis at day 7 p.i.) were cultured in Bovine Aortic endothelial cells as previously described for Gardel strain (Marcelino et al. 2005). When 80% lysis had occurred, supernatant and cellular debris were collected and centrifuged for 15 minutes at 4,000g at 4°C to remove cellular debris. The supernatant was then centrifuged for 30 minutes at 20,000g at 4°C in order to collect elementary bodies and remove the supernatant. The pellet was resuspended in 2ml of cold PBS and homogenized gently then 20 ml of PBS was added to wash it, followed by centrifugation for 30 minutes at 20,000g at 4°C. The pellet was resuspended in 350µl PBS and 10µl of RNase at 10mg/ml (SIGMA, Lyon, France) and 150µl of DNase I (Roche, Boulogne-Billancourt, France) were added. The pellet was incubated at 37°C for 90 minutes and the reaction was stopped by adding 25µl 0.5M EDTA, pH8. The DNase and Rnase were removed by centrifugation for 15 minutes at 20,000g at 12°C followed by a wash with 900µl of sterile DNase and RNase free water. Cells were centrifuged under the same conditions and the washing step was repeated twice. Lysis of elementary bodies was performed by adding 500µl of lysis solution (0.1M TRIS-HCl (pH8), 0.15M NaCl, 0.025M EDTA (pH8), 1.5% SDS, 0.3 mg/ml Proteinase K) followed by Incubation for 120 minutes at 55°C. DNA extraction was performed using phenol/chloroform (Perez et al. 1997) as follows: 500µl of phenol (Eurobio, Courtaboeuf, France) was added to 500µl of sample and mixed gently to homogenize followed by centrifugation for 5 minutes at 8000 g. The aqueous phase was collected and an equivalent volume of phenol was added before centrifugation for 5 minutes at 8,000 g. The aqueous phase was collected and an equivalent volume of phenol (Eurobiotech, France), chloroform and isoamylalcohol (Prolabo, Normapur, France) in 24:24:1 proportion was added and mixed gently. The mixture was centrifuged for 5 minutes at 8,000g and the aqueous phase collected and mixed gently with an equal volume of chloroform/isoamylalcohol in 48:1 proportion. The mixture was centrifuged for 5 minutes at 8,000g and the supernatant was collected and precipitated in ethanol (Prolabo, Normapur) by adding 2 volumes of absolute ethanol for 1 volume of the collected sample. The sample was stored at 4°C for one hour in order to obtain a white precipitate which was then centrifuged for 10 minutes at 10,000g at 10°C. The supernatant was removed and the pellet resuspended in 1 ml of 75% ethanol followed by centrifugation for 5 minutes at 10,000g. The supernatant was removed and the

pellet air dried. The pellet was resuspended in 25µl of TE buffer (10mM TRISpH8, 1mM EDTA pH8). DNAs from Senegal passage 7, 63 and 64 in TE buffer were stored at –20°C before being used for further sequencing.

#### *Genome sequencing and assembly*

- 5 The genome of the virulent Senegal strain (passage 7) was sequenced using 454 GS FLX technology. The attenuated strain (passage 63 & 64) was sequenced using 454 GS FLX technology and Sanger sequencing. Adapters for the 454 sequences were clipped using NextGen Sequence Workbench v3.2.3 ("NextGen Sequence Workbench" 2015), and reads were clipped for adaptors and quality scores according
- 10 to the information in the SFF sequence files as well as removing reads less than 25bp long or those with an average quality score less than 16. Quality trimming was performed for the Sanger sequences using NextGen Sequence Workbench v3.2.3 ("NextGen Sequence Workbench" 2015) with default settings (clip the reads in a 14bp window until >63% have a quality score >= 20) and removing reads less than 25bp
- 15 long or those with an average quality score less than 16. The virulent strain was de novo assembled using Mira 4.0.2 (Chevreux, Wetter, and Suhai 1999), resulting in 43 contigs larger than 1kb. The contigs were ordered according to the previously published Welgevonden strain genome (Collins et al. 2005) using CONTIGuator (Galardini et al. 2011), which mapped 42 of the 43 contigs to the Welgevonden
- 20 genome. The assembly was checked by realigning the reads onto the concatenated assembly using Bowtie 2 (Langmead and Salzberg 2012) and checking any predicted variants against the Welgevonden genome sequence (Collins et al. 2005). Duplicate reads were removed using Picard ("Picard Tools" 2016), conversion between SAM and BAM formats, sorting and mpileup was done using Samtools (Li et al. 2009) and
- 25 variants were called using BCFtools (Li et al. 2009) and viewed with Tablet (Milne et al. 2013). Variant calls were manually examined in Tablet (Milne et al. 2013) and the assembly was edited accordingly. The attenuated strain was mapped onto the assembled virulent strain as described for remapping the virulent strain. Indels and SNPs were identified with quality cutoffs of 50 and 20 respectively and manually
- 30 examined in the read alignment. Both genome assemblies were submitted to Genbank (Accessions MQUJ000000000 for *Ehrlichia ruminantium* Senegal Virulent, MRDC000000000 for *Ehrlichia ruminantium* Senegalp63 attenuated).

### *E. ruminantium* RNA purification for RT-PCR

*E. ruminantium* Senegal attenuated (passage 63 and 66) and virulent (passage 7 and 11) were inoculated in 25 cm<sup>2</sup> TC flask containing BAE cells as previously described (Marcelino et al. 2005). The medium was changed at 24h and 72h for the attenuated strain and at 24h and every two days, thereafter, for the virulent strain. Cell layers were allowed to reach around 80-90% lysis and were mechanically harvested with a scraper. Re-suspended cells were centrifuged at 4,500 x g for 30 minutes at 4 °C. Supernatant was discarded and cells were re-suspended in 1 ml PBS. Cells were then centrifuged at 10,000 x g for 10 minutes and PBS was removed. Pellets were stored at -70°C until RNA purification. Cell pellets were allowed to thaw for 5 minutes in ice and RNA was purified using the SV total RNA isolation system (Promega Corporation, Wisconsin, USA). An additional DNase treatment was added by using the rigorous DNase treatment with Turbo DNA-free (Ambion, Fisher Scientific, Illkirch, France), which consisted in adding 0.5 µl of the DNase, incubating for 30 minutes, and repeating this procedure. RNA was immediately stored at -70 °C after inactivation of the DNase before ntrX RT-PCR.

### *RT-PCR of ntrX*

The expression of the *ntrX* in the virulent (passage 7 and 11) and attenuated (passage 63 and 66) Senegal strains was determined using the primers *ntrX* qRT F1 (5'-GGAAAGATTGTATATTTCTG-3') (SEQ ID NO: 21) and *ntrX* qRT2 R1 (5'-ACCAAGTAATGAGTATACGAC-3') (SEQ ID NO: 22) that amplify a 517 bp piece of the *ntrX* gene. Amplifications were done using the OneStep RT-PCR kit (QIAGEN, California, USA) with the following conditions: one Reverse transcriptase cycle at 50 °C for 30 minutes, a denaturing cycle at 95 °C for 15 minutes for activation of HotStart Taq, 35 cycles with a denaturing step 95 °C for 1 minute, an annealing step at 50 °C for 1 minute, and an amplification step at 72 °C for 1 minute, followed by an amplification cycle of 72 °C for 10 minutes. DNA from *E. ruminantium* Senegal passage 6 was used as positive control. Products were run in agarose gel and bands were visualized with SYBR safe.

*Distance of E. muris and Senegal strain intergenic regions*

The Senegal strain genome was aligned to the *Ehrlichia muris* genome using Mauve 2.3.1 (Darling et al. 2004) and intergenic regions were output from the alignment based on the *E. muris* annotation. Distances were calculated between the aligned intergenic regions using the TN93 model (Tamura and Nei 1993) in the APE package (Paradis, Claude, and Strimmer 2004) in R version 3.2.3 (R Core Team 2015) where at least 30 bases could be aligned, resulting in 364 intergenic distances between Senegal strain and *E. muris*. The distance between the *ntrX* gene from *E. muris* and the duplicate from Senegal strain was measured in the same way for comparison with the intergenic regions.

## 10 Vaccination

### *Preparation of the inoculum*

#### A. Preparation of the purified supernatant

As soon as 80% of lysis of cells is reached with a synchronous infection corresponding to 5 days of culture of Senegal passage 68, the supernatant with the cellular debris was passed in a syringe 26G3/8. The totality of the supernatant was collected and centrifuged for 15 minutes at 3000rpm. Then, the supernatant was recovered without removing the cellular pellet.

500µl of the supernatant was tested in order to evaluate the viability of the sample. One part of the purified supernatant has been used to infect endothelial cell TC flask in order to control the infectivity of the inoculum 14 ml were at 4°C before inoculation.

#### B Evaluation of the viability using The LIVE/DEAD BacLight™ Bacterial Viability Kits (ThermoFisher Scientific)

The purified supernatant was washed in 15ml of physiological serum and then centrifuged 30min at 20,000g at 4°C. It has been suspended again in 500µl of physiological serum, passed in a syringe. 1.5µl of Syto9 and Propidium Iodide was added. It was incubated 15min then counted on a slide and passed in a flow cytometer. It was counted in a Neubaur chamber in triplicate 8 squares each time. The final concentration was obtained by multiplication by the factor  $1.6 \times 10^5$ . There was no dilution factor.

*Infection of the animals*

- Naïve goats were injected intravenously with the following doses of elementary bodies of live attenuated Senegal strain passage 64. The calibrated doses which reproduce natural challenge for virulent strain (between 10 and 12 days before hyperthermia and dead between day 12 and 15 after infection) is comprised between  $3 \cdot 10^4$  and  $9 \cdot 10^4$  live elementary bodies per goat (Vachieri et al, 2006). For this experiment, we decided to use a lethal dose and ten times the lethal dose using  $9 \cdot 10^4$  and  $9 \cdot 10^5$  live elementary bodies per goat.

Goat number	Challenge Doses (Live elementary body number)
0615	$9 \cdot 10^4$
0636	$9 \cdot 10^4$
0803	$9 \cdot 10^5$

- 10 Preparation of inoculum depending on the viabilities and concentrations which were found.

*Viability*

- From the cell culture supernatant we measured  $5.89 \cdot 10^6$  elementary bodies /ml on neaubeaur counting cell. The percentage of viability was measured by flow cytometry and we obtained 40% of viability. The number of live elementary bodies was  $2.35 \cdot 10^6$  CE live/ml. The supernatant was diluted in fresh cell culture medium to get  $9 \cdot 10^4$  and  $9 \cdot 10^5$  elementary bodies in a final volume of 2ml.

*Monitoring of animals:*

- 20 Clinical signs were checked every day in order to score the severity of the disease. A serology targeting MAP-1 antibodies was done one time a week and one blood sample were taken daily.

**Results:**Virulent and attenuated Senegal strains genomic differences: SNPs and Indels

The variants found between the virulent and attenuated strains are shown in Table below.

5

Passage	Days to lyse	Virulence
4	5	Virulent
4	8	Virulent
4	>8	Virulent
4	>10	Virulent
5	6	Virulent
5	>7	Virulent
5	5	Virulent
6	14	Virulent
6	14	Virulent
6	>7	Virulent
6	7	Virulent
6	>5	Virulent
7	4	Virulent
7	4	Virulent



7	6	Virulent
7	6	Virulent
8	6	Virulent
8a	6	Virulent
8b	6	Attenuated
65	4	Attenuated
68	5	Attenuated
69	5	Attenuated
70	5	Attenuated
70	4	Attenuated
71	4	Attenuated
73	4	Attenuated
74	4	Attenuated
74	4	Attenuated
75	5	Attenuated
75	4	Attenuated
76	4	Attenuated

Table: Days for lysis of virulent and attenuated Senegal strain passages

We identified only two SNPs and three indels between the virulent and attenuated Senegal strains. The SNPs occur in the *glyA* (ERGA\_CDS\_07110 ortholog), a serine hydroxymethyltransferase involved in the interconversion of serine and glycine as well as tetrahydrofolate production, and the ERGA\_CDS\_07720 ortholog, a putative M16

protease. Both SNPs are non synonymous. The indels occur in a hypothetical gene (ERGA\_CDS\_01780 ortholog), the response regulator of the putative nitrogen-sensing two-component system, *ntrX* (ERGA\_CDS\_06840 ortholog) and *map1-2* (ERGA\_CDS\_09130 ortholog), a member of the *map1* family of outer membrane proteins. The hypothetical gene and *ntrX* both contain a 4bp deletion, while the *map1-2* gene contains a 2bp insertion. The *map1-2* gene appears to be a pseudogene in the virulent Senegal strain and the result of the insertion is to restore the open reading frame of the gene in the attenuated strain, making it a possible gain of function mutation. The hypothetical protein contains a Patatin domain and has sequence similarity to other patatin-like phospholipase family proteins that have lipolytic activity (Banerji and Flieger 2004).

#### Candidate mutations for Senegal strain attenuation

The two nonsynonymous SNPs and the three indels identified between the virulent and attenuated Senegal strains provide a small number of candidates to explain the attenuation process in this strain.

The *glyA* gene is involved in the interconversion of serine and glycine, and is necessary for virulence in *Salmonella* Typhimurium and *Brucella* species (Köhler et al. 2002; Xiang, Zheng, and He 2006; Jelsbak et al. 2014). However, the attenuation in *Brucella suis* was obtained by Tn5 insertion to knock out the gene, and in the case of the SNP in our data, the protein is still probably produced. The substituted amino acid lies near the end of the protein (position 388/421) and outside of any predicted domains. An alignment of *glyA* in several Rickettsial species shows that while most of the protein is well conserved, the portion where the mutation occurs is variable across the species suggesting that the region is not vital for protein function.

The *map1-2* gene is a member of a multigene family of Major Antigenic Proteins in the *Anaplasmataceae*, Pfam PF01617 (Dunning Hotopp et al. 2006), which are surface exposed and have been identified as potential vaccine targets (Bekker et al. 2002). Some members of these map genes have been experimentally characterized as porins (Huang et al. 2007) and are suspected to be involved in host cell adhesion (Garcia-Garcia et al. 2004; Park, Choi, and Dumler 2003). Thus, its mutation appears to be a potential candidate for attenuation. However, a study of the *map1-2* gene in several different strains of *E. ruminantium* failed to provide evidence of transcription of this gene in any of the strains they tested (Senegal virulent and attenuated, Gardel,

Welgevonden and Sankat 430) in either ticks or bovine endothelial cells (Bekker et al. 2002). A different study did however report the transcription of the *map1-2* gene in Welgevonden (van Heerden et al. 2004). The *map1-2* gene contains a 2bp deletion in the Senegal virulent strain (presumably rendering it non-functional) that is reverted in  
5 the attenuated strain. Curiously, the deletion is not present in a previous study using a different isolate of the virulent Senegal strain (Bekker et al. 2005). Examination of the reads covering this indel in our data revealed that they fully support the deletion in the virulent Senegal, while it is not present in any reads in the attenuated strain. This result suggests that this deletion may be unique to the virulent strain sequenced in  
10 this study. The lack of detectable transcription of the gene in some strains (Bekker et al. 2002) and the fact that the complete *map1-2* sequence is present in several other virulent strains of *E. ruminantium* including another Senegal isolate makes the reversion insertion unlikely to be the cause of attenuation of the Senegal strain.

The patatin domain-containing protein (ERGA\_CDS\_01780 ortholog) may play a role  
15 in the virulence of *E. ruminantium*, as patatin-like phospholipase proteins are putatively involved in host cell entry in Rickettsia (Rahman et al. 2010). Phospholipase proteins have been identified as virulence factors secreted by the Type III secretion system in *Pseudomonas aeruginosa* (ExoU), and the Type IV-B secretion system in *Legionella pneumophila* (VipD) (Rahman et al. 2010). Bacterial pathogens also tend  
20 to contain more patatin-like-proteins than non-pathogens (Banerji and Flieger 2004), suggesting possible roles in virulence. However, their presence in non-pathogens also shows that patatins are not always involved in virulence functions. The indel in the attenuated strain is close to, but outside of the predicted patatin domain, suggesting that the phospholipase activity might be maintained, but making the result of the indel  
25 unsure in terms of its effect on the function of the protein. Proteolysis can play roles in virulence at various levels in bacterial pathogens (Frees, Brøndsted, and Ingmer 2013), and although we could find no evidence for a known role of proteases from the M16 family in bacterial virulence, proteases from this family in *Toxoplasma gondii* have been suggested to play a possible role in host invasion by the parasite (Laliberté  
30 and Carruthers 2011).

Consequently, the only attenuation candidate is the *ntrX* gene containing the 4bp deletion because it is the response regulator of one of only three two-component systems in *E. ruminantium* that are responsible for global regulation of various

bacterial systems (Cheng et al. 2006; Kumagai et al. 2006; Cheng, Lin, and Rikihisa 2014). Two-component systems are involved in sensing of environmental or cellular signals and the downstream expression of genes, allowing bacteria to coordinate their gene expression in response to their environment or cellular state. During infection, bacterial pathogens need to efficiently coordinate their metabolic activities with their virulence to allow for maximization of their growth and successful attack on the host cells, while avoiding host defences. Signals such as the availability of nutrients or metabolites may inform the bacteria of the optimal time to produce virulence factors, and the regulation of many bacterial virulence factors is linked with nutrient availability (Somerville and Proctor 2009; Barbier, Nicolas, and Letesson 2011). NtrX lies at the crossroads of environmental sensing and gene expression and is therefore the ideal candidate to explain attenuation because perturbations to the NtrY/X system are likely to have major consequences for the growth, survival and the coordination of bacterial metabolism and virulence. Furthermore, *ntrX* as the attenuator explains the apparently biased nature of attenuation in the Senegal strain as compared to other *E. ruminantium* strains due to its genomic context, which will be discussed later.

The *ntrX* gene is disrupted by segmental gene conversion from a nearby inverted partial *ntrX* duplication in attenuated Senegal strain

A segment of the *ntrX* gene has been duplicated in all *E. ruminantium* genomes for which there is available genome sequence data. The duplicated section is inverted and covers 421bp close to the 5' end of *ntrX* (not including the start codon), and lies roughly 2kb downstream of the *ntrX* gene itself. The alignment between the *ntrX* and the duplicate region reveals that the 4bp deletion identified in the *ntrX* gene in the Senegal attenuated strain is also present in the duplicated segment in both the attenuated and virulent strains of Senegal, but not in any of the other strains. The 4bp deletion in *ntrX* in the attenuated strain causes a frameshift between the domain regions for the response regulator receiver domain and the sigma factor interaction domain, which introduces a stop codon that disrupts the gene. There are also seven other mutations across the sequenced strains between *ntrX* and its duplicate that exhibit a pattern that is incongruent with the expected phylogeny. At these residues, the *ntrX* gene and its duplicate are more similar within a strain than they are to their orthologous positions in the other strains, a pattern that indicates gene conversion (figures 1 and 2).

	SEQ ID NO:	Nucleic acid sequences of regions shown at figure 1
Welgevonden Duplicate	12	CTAATAATGATATATTAGGTGTGATAATTATACCTAAATAGCAGTTGATGGTTATCCCGCATCAAGATG GCTTATGAAAAGAGCCTGATGTTGTTATTTGGATATATGGTTAAGAGGATCTGATATTGATGGATTAAAGT GTACTGAAAAGCTTAAAGAAAGGTATCCTTATTTGCCCTGTTATTATGATTAGTGGCATGGTAAATATTGCCA CTGCTGTAAGTCTCTGCATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAGT TAGTTGTAAGAGAGCTATAGAGTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGAG GATTATGAAATAGTCGGTAACCTCCCTGTTATACGTAATTTGAGAAGTATGGTTAATAA
Gardel Duplicate	13	CTAATAATGATATATTAAAGTGATGATAATTATGTCACATAATTAGCAGTTGATGGTTATCCCGCATCAAGAT GGCTTATGAAAAGAGCCTGATGTTGTTATTTGGATATATGGTTAAGAGGATCTGATATTGATGGATTAAAG TGTACTGGAGAAGCTTAAAGAAAGGTATCCTTATTTGCCCTGTTATTATGATTAGTGGCATGGTAAATATTGCC ACTGCTGTAAGTCTCTGCATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAG TTAGTTGTAAGAGAGCTATAGAGTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGAG GATTATGAAATAGTCGGTAACCTCCCTGTTATACGTAATTTGAGAAGTATGATTAAATA
Senegal Virulent Duplicate	14	CTAATAATGATATATTAGGCAATGATAATTATGTAAACCAATTAGCAGTTGATTATTTATGAAAAGAGCCTG ATGTTGATTATTGGATATATAGTTAAAGAGGATCTGATATTGATGGATTAAAGCTACTGGAAGAAAGCTTAAA GAAAGGTATCCTTATTTGCCCTGTTATTATGATTAGTGGCATGGTAAATATTGCTACTGCTGTAAGTCTTTGC ATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAGTTGTAAGAGAGCTATAGA GTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGAGGATTATGAGATAGTGGGTAACCTC GCCTGTTATACGTAATTTGAGAAGTATGATTAAATA
Senegal Attenuated Duplicate	15	CTAATAATGATATATTAGGCAATGATAATTATGTAAACCAATTAGCAGTTGATTATTTATGAAAAGAGCCTG ATGTTGATTATTGGATATATAGTTAAAGAGGATCTGATATTGATGGATTAAAGCTACTGGAAGAAAGCTTAAA GAAAGGTATCCTTATTTGCCCTGTTATTATGATTAGTGGCATGGTAAATATTGCTACTGCTGTAAGTCTTTGC ATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAGTTGTAAGAGAGCTATAGA GTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGAGGATTATGAGATAGTGGGTAACCTC GCCTGTTATACGTAATTTGAGAAGTATGATTAAATA
Welgevonden <i>ntrX</i>	16	CTAATAAAGATATATTAAAGTGATGATAATTATGTCACATAATTAGCAGTTGATGGTTATCCCGCATCAAGA TGGCTTATGAAAAGAGCCTGATGTTGTTATTTAGATATATGGTTAAGAGGATCTGATATTGATGGATTAAAG TGTACTGAAAAGCTTAAAGAAAGGTATCCTTATTTGCCCTGTTATTATGATTAGTGGCATGGTAAATATTGCC ACTGCTGTAAGTCTCTGCATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAG TTAGTTGTAAGAGAGCTATAGAGTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGAG GATTATGAAATAGTCGGTAACCTCCCTGTTATACGTAATTTGAGAAGTATGATTAAATA
Gardel <i>ntrX</i>	17	CTAATAAAGATATATTAAAGTGATGATAATTATGTCACATAATTAGCAGTTGATGGTTATCCCGCATCAAGA TGGCTTATGAAAAGAGCCTGATGTTGTTATTTAGATATATGGTTAAGAGGATCTGATATTGATGGATTAA GGTACTGGAGAAGCTTAAAGAAAGGTATCCTTATTTGCCCTGTTATTATGATTAGTGGCATGGTAAATATTGC CACTGCTGTAAGTCTCTGCATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAG GTTAGTTGTAAGAGAGCTATAGAGTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGGA GGATTATGAAATAGTCGGTAACCTCCCTGTTATACGTAATTTGAGAAGTATGATTAAATA
Senegal Virulent <i>ntrX</i>	18	CTAATAAAGATATATTAAAGTGATGATAATTATGTCACATAATTAGCAGTTGATGGTTATCTCCGATCAAGA TGGCTTATGAAAAGAGCCTGATGTTGTTATTTAGATATATGGTTAAGAGGATCTGATATTGATGGATTAA GCGTACTGGAAGCTTAAAGAAAGGTATCCTTATTTACCCTGTTATTATGATTAGTGGCATGGTAAATATTGC TACTGCTGTAAGTCTTTGCATATGGGTGCTTATGATTATAGAAAAGCCCTTTACAGAAGGAAGATTAAAG GTTAGTTGTAAGAGAGCTATAGAGTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTGGA GGATTATGAGATAGTGGTAACCTCCCTGTTATACGTAATTTGAGAAGTATGATTAAATA

Senegal Attenuated <i>ntrX</i>	19	CTAATAAAGATATATTAAGTGATGATAATTATGTCACATAATAGCAGTTGATGGTTTATCTCGCATCAAGA TGGCTTATGAAAAAGAGCCTGATGTTGTATTATTGGATATATGGTTAAGAGGATCTGATATTGATGGATTAA GCGTACTGGAAAAGCTTAAGAAAGGATCCTTATTACCTGTTATTATGATTAGTGGGCATGGTAATATTGC TACTGCTGTAAGTCCTTGCATATGGTGCTTATGATTATAGAAAAGCCTTTTACAGAAGGAAGATTAAA GTTGTAAGAGAGCTATAGAGTCTGGTAGATTACGTAGAGAAAATGATGAGTTGAAATCAGCATTTGAGGA TTATGAGATAGTGGTAACCTCGCCTGTTATACGTAATTTGAGAAGTATGATTAAATAA TTAATAAAGGATATATTAAGTGATGATAATTATGTCACAAAATTAGCAGTTGATGGTTGCTGCTATTAAAGA TGGCTTATGAGAAAGAACCCAGATGTTGTTTTACTAGATATATGGTTAAAAGGATCAGATATTGATGGGTTAA GTGTTTTAGAGAAAACCTAAAGGAAAGGTATCCATATTACCTGTGATTATGATTAGTGACATGGTAATATTGC TACTGCTGTAAGTCCTTGCACATGGGAGCTTATGATTATATAGAGAAACCTTTACAGAAGGTAGATTAAA GTTAGTAGTTAAGAGAGCGATAGAATCTGGTAGATTGCGTAGAGAAAATGACGAATTAATAATCAACATTGGA AGATTACGAAATAGTTGGCAACTCTCCTGTAATAAAAAATCTAAGGAGTATGATTAAATAA
<i>E. chaffeensis</i> <i>ntrX</i>	20	

NtrX is expressed in Senegal virulent strain but not in the attenuated strain

To confirm the pseudogenisation of *ntrX* in the attenuated strain, we performed a RT-PCR on RNA samples from the virulent (passage 7 and 11) and avirulent (passage 63 and 66) strains. RT-PCR resulted in the amplification of a 517 bp in one (Senegal passage 11 and not Senegal passage 7) of the two virulent samples confirming the expression of *ntrX*. No band was observed in either of the samples from the attenuated strain indicating that *ntrX* gene is not expressed in the attenuated Senegal strain.

10 Vaccination with the attenuated strain

After inoculation, the goats infected with the attenuated strain suffered from a slight hyperthermia with no other clinical signs and all survived even with 10 fold lethal dose of challenge. Consequently, the mutant *E. ruminantium* passage 64 is attenuated.

100% of goats infected with the attenuated strain and which were submitted to a challenge with the virulent strain (*E. ruminantium* passage 7) survived, whereas all the control goats, which were submitted to a challenge with the virulent strain without having been previously infected with the attenuated strain, died.

These results show that a *E. ruminantium* with an inactive or deleted *ntrX* gene provides an efficient vaccine against ehrlichiosis.

20 Owing to the very high conservation degree of the *ntrX* gene in the *Anaplasmataceae* family (see table 1), these results apply to the other strains of the *Anaplasmataceae* family.

## REFERENCES

Throughout this application, various references describe the state of the art to which this invention pertains. The disclosures of these references are hereby incorporated by reference into the present application.

- Atack, John M., Yogitha N. Srikhanta, Karrera Y. Djoko, Jessica P. Welch, Norain H. M. Hasri, Christopher T. Steichen, Rachel N. Vanden Hoven, et al. 2013. "Characterization of an ntrX Mutant of *Neisseria Gonorrhoeae* Reveals a Response Regulator That Controls Expression of Respiratory Enzymes in Oxidase-Positive Proteobacteria." *Journal of Bacteriology* 195 (11): 2632–41. doi:10.1128/JB.02062-12.
- 5 Banerji, Sangeeta, and Antje Flieger. 2004. "Patatin-like Proteins: A New Family of Lipolytic Enzymes Present in Bacteria?" *Microbiology (Reading, England)* 150 (Pt 3): 522–25. doi:10.1099/mic.0.26957-0.
- Bekker, Cornelis P. J., Lesley Bell-Sakyi, Edith A. Paxton, Dominique Martinez, Albert Bensaid, and Frans Jongejan. 2002. "Transcriptional Analysis of the Major Antigenic Protein 1 Multigene Family of *Cowdria Ruminantium*." *Gene* 285 (1–2): 193–201.
- 10 Bekker, Cornelis P. J., Milagros Postigo, Amar Taoufik, Lesley Bell-Sakyi, Conchita Ferraz, Dominique Martinez, and Frans Jongejan. 2005. "Transcription Analysis of the Major Antigenic Protein 1 Multigene Family of Three In Vitro-Cultured *Ehrlichia Ruminantium* Isolates." *Journal of Bacteriology* 187 (14): 4782–91. doi:10.1128/JB.187.14.4782-4791.2005.
- 15 Camus, E., and N. Barré. 1988. "[Diagnosis of heartwater from brain ecrasement]." *Revue*
- Cheng, Zhihui, Yumi Kumagai, Mingqun Lin, Chunbin Zhang, and Yasuko Rikihisa. 2006. "Intra-Leukocyte Expression of Two-Component Systems in *Ehrlichia Chaffeensis* and *Anaplasma Phagocytophilum* and Effects of the Histidine Kinase Inhibitor Closantel." *Cellular Microbiology* 8 (8): 1241–52. doi:10.1111/j.1462-5822.2006.00704.x.
- 20 Cheng, Zhihui, Mingqun Lin, and Yasuko Rikihisa. 2014. "*Ehrlichia Chaffeensis* Proliferation Begins with NtrY/NtrX and PutA/GlnA Upregulation and CtrA Degradation Induced by Proline and Glutamine Uptake." *mBio* 5 (6). doi:10.1128/mBio.02141-14.
- 25 Chevreux, B., T. Wetter, and S. Suhai. 1999. "Genome Sequence Assembly Using Trace Signals and Additional Sequence Information." In *Proceedings of the German Conference on Bioinformatics*, 99:45–56.



- Collins, Nicola E., Junita Liebenberg, Etienne P. de Villiers, Kelly A. Brayton, Elmarié Louw, Alri Pretorius, F. Erika Faber, et al. 2005. "The Genome of the Heartwater Agent *Ehrlichia Ruminantium* Contains Multiple Tandem Repeats of Actively Variable Copy Number." *Proceedings of the National Academy of Sciences of the United States of America* 102 (3): 838–43. doi:10.1073/pnas.0406633102.
- Darling, Aaron C. E., Bob Mau, Frederick R. Blattner, and Nicole T. Perna. 2004. "Mauve: Multiple Alignment of Conserved Genomic Sequence with Rearrangements." *Genome Research* 14 (7): 1394–1403. doi:10.1101/gr.2289704.
- Dunning Hotopp, Julie C., Mingqun Lin, Ramana Madupu, Jonathan Crabtree, Samuel V. Angiuoli, Jonathan A. Eisen, Jonathan Eisen, et al. 2006. "Comparative Genomics of Emerging Human Ehrlichiosis Agents." *PLoS Genetics* 2 (2): e21. doi:10.1371/journal.pgen.0020021.
- Frees, Dorte, Lone Brøndsted, and Hanne Ingmer. 2013. "Bacterial Proteases and Virulence." *Sub-Cellular Biochemistry* 66: 161–92. doi:10.1007/978-94-007-5940-4\_7.
- Frutos, Roger, Alain Viari, Conchita Ferraz, Anne Morgat, Sophie Eychenié, Yane Kandassamy, Isabelle Chantal, et al. 2006. "Comparative Genomic Analysis of Three Strains of *Ehrlichia Ruminantium* Reveals an Active Process of Genome Size Plasticity." *Journal of Bacteriology* 188 (7): 2533–42. doi:10.1128/JB.188.7.2533-2542.2006.
- Galardini, Marco, Emanuele G. Biondi, Marco Bazzicalupo, and Alessio Mengoni. 2011. "CONTIGuator: A Bacterial Genomes Finishing Tool for Structural Insights on Draft Genomes." *Source Code for Biology and Medicine* 6: 11. doi:10.1186/1751-0473-6-11.
- Garcia-Garcia, Jose C., José de la Fuente, Gianna Bell-Eunice, Edmour F. Blouin, and Katherine M. Kocan. 2004. "Glycosylation of *Anaplasma Marginale* Major Surface Protein 1a and Its Putative Role in Adhesion to Tick Cells." *Infection and Immunity* 72 (5): 3022–30. doi:10.1128/IAI.72.5.3022-3030.2004.
- Jansen, Gunther, Lena L. Crummenerl, Felix Gilbert, Timm Mohr, Roxana Pfefferkorn, Robert Thänert, Philip Rosenstiel, and Hinrich Schulenburg. 2015. "Evolutionary

- Transition from Pathogenicity to Commensalism: Global Regulator Mutations Mediate Fitness Gains through Virulence Attenuation.” *Molecular Biology and Evolution* 32 (11): 2883–96. doi:10.1093/molbev/msv160.
- Jelsbak, Lotte, Hassan Hartman, Casper Schroll, Jesper T. Rosenkrantz, Sebastien  
5 Lemire, Inke Wallrodt, Line E. Thomsen, et al. 2014. “Identification of Metabolic Pathways Essential for Fitness of *Salmonella Typhimurium* In Vivo.” *PLoS ONE* 9 (7): e101869. doi:10.1371/journal.pone.0101869.
- Jongejan, F. 1991. “Protective Immunity to Heartwater (*Cowdria Ruminantium* Infection) Is Acquired after Vaccination with in Vitro-Attenuated *Rickettsiae*.” *Infection and Immunity* 59 (2): 729–31.  
10
- Köhler, Stephan, Vincent Foulongne, Safia Ouahrani-Bettache, Gisèle Bourg, Jacques Teyssier, Michel Ramuz, and Jean-Pierre Liautard. 2002. “The Analysis of the Intramacrophagic Virulome of *Brucella Suis* Deciphers the Environment Encountered by the Pathogen inside the Macrophage Host Cell.” *Proceedings of the*  
15 *National Academy of Sciences of the United States of America* 99 (24): 15711–16. doi:10.1073/pnas.232454299.
- Kumagai, Yumi, Zhihui Cheng, Mingqun Lin, and Yasuko Rikihisa. 2006. “Biochemical Activities of Three Pairs of *Ehrlichia Chaffeensis* Two-Component Regulatory System Proteins Involved in Inhibition of Lysosomal Fusion.” *Infection and Immunity* 74 (9):  
20 5014–22. doi:10.1128/IAI.00735-06.
- Laliberté, Julie, and Vern B. Carruthers. 2011. “*Toxoplasma Gondii* Toxolysin 4 Is an Extensively Processed Putative Metalloproteinase Secreted from Micronemes.” *Molecular and Biochemical Parasitology* 177 (1): 49–56. doi:10.1016/j.molbiopara.2011.01.009.
- 25 Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59. doi:10.1038/nmeth.1923.
- Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, Richard Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. “The Sequence Alignment/Map Format and SAMtools.”

- Bioinformatics (Oxford, England) 25 (16): 2078–79.  
doi:10.1093/bioinformatics/btp352.
- Marcelino, Isabel, Célia Veríssimo, Marcos F. Q. Sousa, Manuel J. T. Carrondo, and Paula M. Alves. 2005. "Characterization of Ehrlichia Ruminantium Replication and Release Kinetics in Endothelial Cell Cultures." Veterinary Microbiology 110 (1–2): 87–96. doi:10.1016/j.vetmic.2005.07.012.
- Milne, Iain, Gordon Stephen, Micha Bayer, Peter J. A. Cock, Leighton Pritchard, Linda Cardle, Paul D. Shaw, and David Marshall. 2013. "Using Tablet for Visual Exploration of Second-Generation Sequencing Data." Briefings in Bioinformatics 14 (2): 193–202. doi:10.1093/bib/bbs012.
- Nene, Vishvanath, and Chittaranjan Kole. 2008. Genome Mapping and Genomics in Animal-Associated Microbes. Springer Science & Business Media.
- Park, Jinho, Kyoung Seong Choi, and J. Stephen Dumler. 2003. "Major Surface Protein 2 of Anaplasma Phagocytophilum Facilitates Adherence to Granulocytes." Infection and Immunity 71 (7): 4018–25. doi:10.1128/IAI.71.7.4018-4025.2003.
- Parkinson, J. S., and E. C. Kofoed. 1992. "Communication Modules in Bacterial Signaling Proteins." Annual Review of Genetics 26: 71–112. doi:10.1146/annurev.ge.26.120192.000443.
- Paradis, Emmanuel, Julien Claude, and Korbinian Strimmer. 2004. "APE: Analyses of Phylogenetics and Evolution in R Language." Bioinformatics 20 (2): 289–90. doi:10.1093/bioinformatics/btg412.
- "Picard Tools." 2016. Accessed March 22. <http://broadinstitute.github.io/picard/>.
- Pilet, Héloïse, Nathalie Vachiéry, Moez Berrich, Rim Bouchouicha, Benoît Durand, Ludovic Pruneau, Valérie Pinarello, et al. 2012. "A New Typing Technique for the Rickettsiales Ehrlichia Ruminantium: Multiple-Locus Variable Number Tandem Repeat Analysis." Journal of Microbiological Methods 88 (2): 205–11. doi:10.1016/j.mimet.2011.11.011.

- Rahman, M. Sayeedur, Nicole C. Ammerman, Khandra T. Sears, Shane M. Ceraul, and Abdu F. Azad. 2010. "Functional Characterization of a Phospholipase A2 Homolog from *Rickettsia Typhi*." *Journal of Bacteriology* 192 (13): 3294–3303. doi:10.1128/JB.00155-10.
- 5 Tanner, Jennifer R., Laam Li, Sebastien P. Faucher, and Ann Karen C. Brassinga. 2016. "The CpxRA Two-Component System Contributes to *Legionella Pneumophila* Virulence." *Molecular Microbiology*, March, n/a-n/a. doi:10.1111/mmi.13365.
- Tamura, K., and M. Nei. 1993. "Estimation of the Number of Nucleotide Substitutions in the Control Region of Mitochondrial DNA in Humans and Chimpanzees." *Molecular*  
10 *Biology and Evolution* 10 (3): 512–26.
- Vachiéry N., Lefrançois, T., Esteves, I., Molia, S., Sheikboudou, C., KandassamyY., Martinez, D., Optimisation of the inactivated vaccine dose against heartwater and in vitro quantification of *Ehrlichia ruminantium* challenge material. 2006. *Vaccine*, 24: 4747-4756.
- 15 Van Heerden H1, Steyn HC, Allsopp MT, Zwegarth E, Josemans AI, Allsopp BA. 2004 "Characterization of the pCS20 region of different *Ehrlichia ruminantium* isolates." *Vet Microbiol.* 101(4):279-91.
- Xiang, Zuoshuang, Wenjie Zheng, and Yongqun He. 2006. "BBP: *Brucella* Genome Annotation with Literature Mining and Curation." *BMC Bioinformatics* 7: 347.  
20 doi:10.1186/1471-2105-7-347.
- Zwegarth, Erich, Antoinette I. Josemans, M. Fransie Van Strijp, Laura Lopez-Rebollar, Mirinda Van Kleef, and Basil A. Allsopp. 2005. "An Attenuated *Ehrlichia Ruminantium* (Welgevonden Stock) Vaccine Protects Small Ruminants against Virulent Heartwater Challenge." *Vaccine* 23 (14): 1695–1702.  
25 doi:10.1016/j.vaccine.2004.09.030.

**CLAIMS**

1. A vaccine composition comprising a bacterial strain with a deleted or inactive *ntrX* gene.
- 5 2. The vaccine composition according to claim 1 wherein said *ntrX* gene:
- encodes a NtrX protein having the amino acid sequence of SEQ ID NO: 1
- or
- is an active homologue thereof encoding a NtrX protein having a amino acid sequence with at least 50 % of identity with SEQ ID NO: 1.
- 10 3. The vaccine composition according to claim 1 or 2 wherein:
- the inactive *ntrX* gene is an inactive mutant:
    - of a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ ID NO: 1
- or
- 15 -of an active homologue thereof encoding a NtrX protein having a amino acid sequence with at least 50 % of identity with SEQ ID NO: 1
- or
- the *ntrX* gene which is:
    - a *ntrX* gene encoding a NtrX protein having the amino acid sequence of SEQ
- 20 ID NO: 1
- or
- an active homologue thereof encoding a NtrX protein having a amino acid sequence with at least 50 % of identity with SEQ ID NO: 1

is deleted.

4. The vaccine composition according to any of claims 1 to 3 wherein:

- the inactive *ntrX* gene is an inactive mutant of a *ntrX* gene encoding a NtrX protein having the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 and SEQ ID NO:10

or

- the *ntrX* gene encoding a NtrX protein having the amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9 and SEQ ID NO:10 is deleted.

5. The vaccine composition according to any of claims 1 to 4 wherein said vaccine composition comprises a pharmaceutically acceptable excipient.

6. The vaccine composition according to any of claims 1 to 5 for use to induce an immune response against the bacterial strain.

7. The vaccine composition according to any of claims 1 to 6 for use as a vaccine.

8. The vaccine composition according to any of claims 1 to 7 for use in preventing and/or treating an infection caused by the bacterial strain.

9. The vaccine composition for use according to claim 8 wherein the bacterial strain is an *Ehrlichia* strain and the infection caused by the bacterial strain is ehrlichiosis.

10. The vaccine composition according to any of claims 1 to 5 or for use according to any of claims 6 to 9, wherein the bacterial strain is an *Alphaproteobacteria* strain, preferably a *Rickettsiales* strain, more preferably an *Anaplasmataceae* strain.

11. The vaccine composition according to any of claims 1 to 5 or for use according to any of claims 6 to 9, wherein the bacterial strain is selected from the group consisting of an *Ehrlichia* strain, an *Anaplasma* strain, a *Rickettsia* strain, an *Orientia* strain, a *Bartonella* strain and a *Brucella* strain, preferably *Ehrlichia* strain.

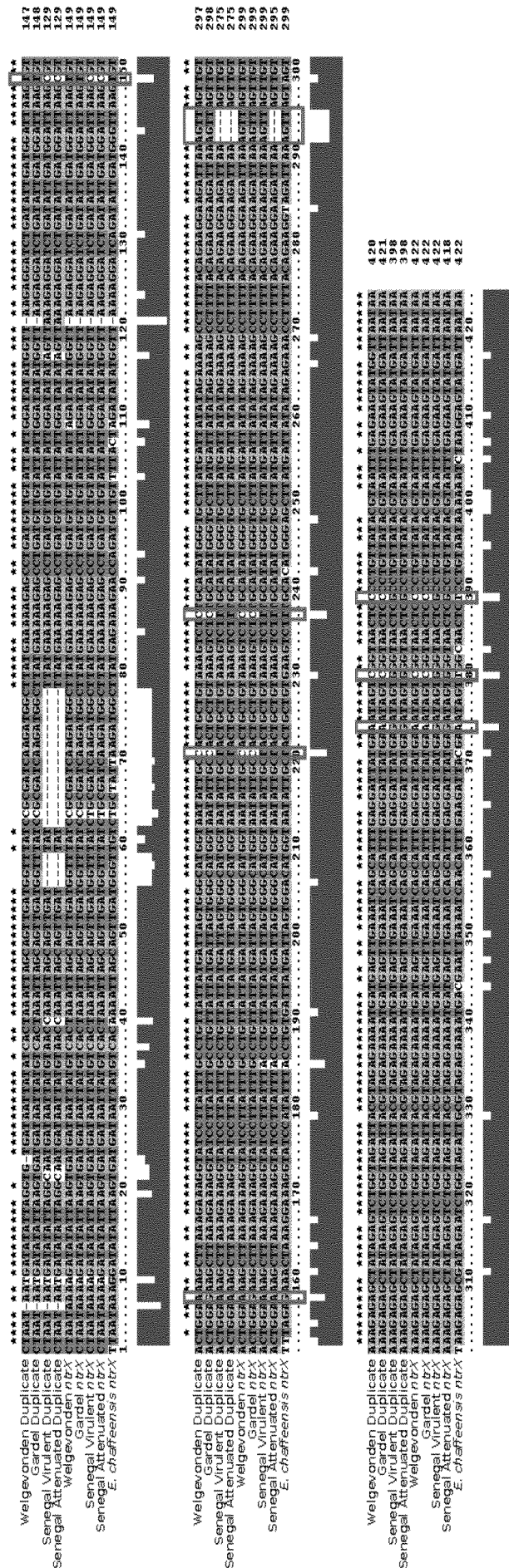
5 12. The vaccine composition according to any of claims 1 to 5 or for use according to any of claims 6 to 9 wherein the bacterial strain is selected from the group consisting of: *E. canis*, *E. chaffeensis*, *E. ewingii*, *E. muris*, and *E. ruminantium*, preferably an *E. ruminantium*.

10 13. A method for producing a vaccine composition for use against a bacterial strain comprising a step of:

- inactivating or deleting the *ntrX* gene of the bacterial strain, thereby obtaining an attenuated bacterial strain.

14. The method for producing a vaccine composition according to claim 13 wherein the bacterial strain is as defined according to any of claims 10 to 12.

15





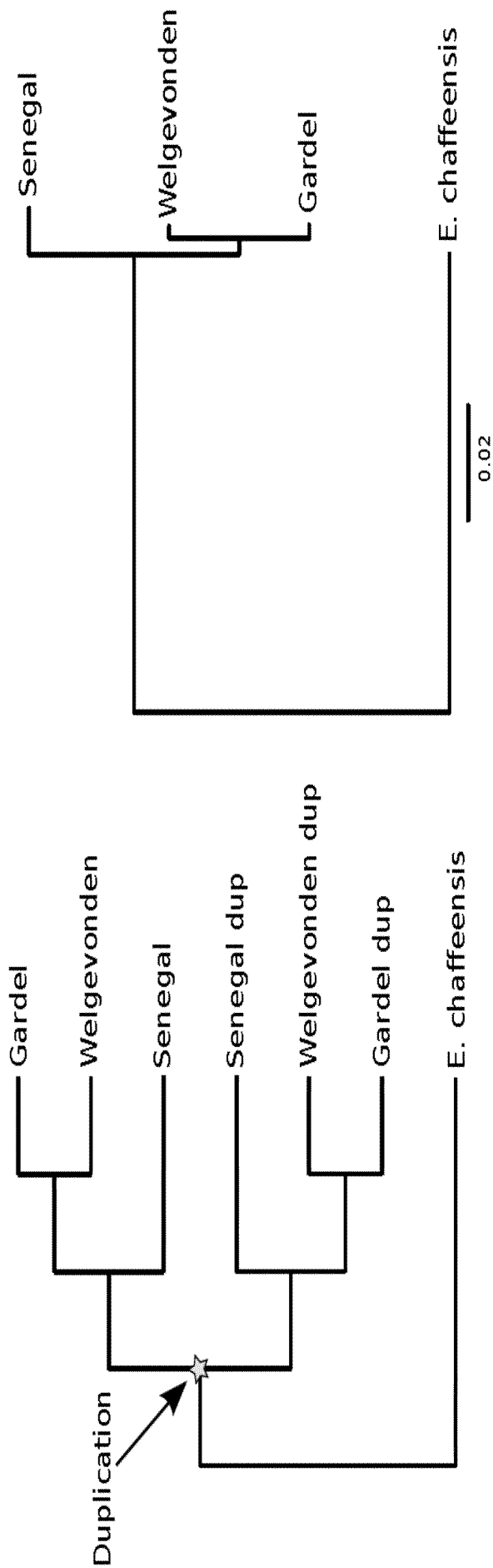


Fig. 2

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/EP2018/075634

A. CLASSIFICATION OF SUBJECT MATTER  
INV. A61K39/02  
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2016/206720 A1 (PURTLE LISA [US] ET AL) 21 July 2016 (2016-07-21) the whole document	1-8,10
A	----- MINGQUN LIN ET AL: "Analysis of complete genome sequence and major surface antigens of Neorickettsia helminthoeca, causative agent of salmon poisoning disease", MICROBIAL BIOTECHNOLOGY, vol. 10, no. 4, 6 June 2017 (2017-06-06), pages 933-957, XP055425651, GB ISSN: 1751-7915, DOI: 10.1111/1751-7915.12731 ----- -/-	1-8,10



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

10 October 2018

Date of mailing of the international search report

23/10/2018

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Marinoni J-C

## INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2018/075634

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	J. M. ATTACK ET AL: "Characterization of an ntrX Mutant of Neisseria gonorrhoeae Reveals a Response Regulator That Controls Expression of Respiratory Enzymes in Oxidase-Positive Proteobacteria", JOURNAL OF BACTERIOLOGY, vol. 195, no. 11, 5 April 2013 (2013-04-05), pages 2632-2641, XP055425662, US ISSN: 0021-9193, DOI: 10.1128/JB.02062-12 -----	1-8
A	ISABEL MARCELINO ET AL: "Comparative Proteomic Profiling of Ehrlichia ruminantium Pathogenic Strain and Its High-Passaged Attenuated Strain Reveals Virulence and Attenuation-Associated Proteins", PLOS ONE, vol. 10, no. 12, 21 December 2015 (2015-12-21), page e0145328, XP055425821, DOI: 10.1371/journal.pone.0145328 -----	1-14
A	JERE W. MCBRIDE ET AL: "Molecular and cellular pathobiology of Ehrlichia infection: targets for new therapeutics and immunomodulation strategies", EXPERT REVIEWS IN MOLECULAR MEDICINE, vol. 13, 1 October 2011 (2011-10-01), XP055425829, DOI: 10.1017/S1462399410001730 -----	1-14
A	Bonto J. Y. Faburay: "MOLECULAR EPIDEMIOLOGY OF HEARTWATER (EHRlichia RUMINANTIUM INFECTION) IN THE GAMBIA", 1 January 2011 (2011-01-01), pages 1-155, XP055425894, Utrecht Retrieved from the Internet: URL:https://www.researchgate.net/profile/Bonto_Faburay/publication/46687069_Molecular_epidemiology_of_heartwater_Ehrlichia_ruminantium_infection_in_The_Gambia/links/00463524b6699d36eb000000/Molecular-epidemiology-of-heartwater-Ehrlichia-ruminantium-infection-in-The-Gambia.pdf [retrieved on 2017-11-16] -----	1-14
A	WO 2009/034575 A1 (YISSUM RES DEV CO [IL]; HARRUS SHIMON [IL]; BANETH GAD [IL]) 19 March 2009 (2009-03-19) -----	1-14

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2018/075634

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016206720 A1	21-07-2016	US 2012052091 A1	01-03-2012
		US 2014186385 A1	03-07-2014
		US 2016206720 A1	21-07-2016
-----			
WO 2009034575 A1	19-03-2009	EP 2187958 A1	26-05-2010
		ES 2618842 T3	22-06-2017
		PT 2187958 T	17-03-2017
		US 2010239613 A1	23-09-2010
		US 2013122041 A1	16-05-2013
		WO 2009034575 A1	19-03-2009
-----			