Identification of *Ehrlichia ruminantium* proteins that activate cellular immune responses using a reverse vaccinology strategy

J. Liebenberg[1,2,*], A. Pretorius[1], F. E. Faber[1], N. E. Collins[2], B. A. Allsopp[2] and M. van Kleef[1,2]

[1]ARC-Onderstepoort Veterinary Institute, Private Bag X05, Onderstepoort, 0110, South Africa

[2]Department of Veterinary Tropical Diseases, University of Pretoria, Private Bag X04, Onderstepoort, 0110, South Africa

[*]**Corresponding Author**

Mailing address: ARC-Onderstepoort Veterinary Institute, Private Bag X5, Onderstepoort, 0110, South Africa

Phone: +27 12 529 9425

Fax: +27 12 529 9417

E-mail: LiebenbergJ@arc.agric.za

Abstract

Ehrlichia ruminantium is an obligate intracellular bacterial pathogen which causes heartwater, a serious tick-borne disease of ruminants throughout sub-Saharan Africa. The development of promising recombinant vaccines has been reported previously, but none has been as effective as immunisation with live organisms. In this study we have used reverse vaccinology to identify proteins that elicit an *in vitro* cellular immune response similar to that induced by intact *E. ruminantium*. The experimental strategy involved four successive steps: (i) *in silico* selection of the most likely vaccine candidate genes from the annotated genome; (ii) cloning and expression of the selected genes; (iii) *in vitro* screening of the expressed proteins for their ability to induce interferon-gamma (IFN-γ) production in *E. ruminantium*–immune lymphocytes; and (iv) further examination of the cytokine response profiles of those lymphocytes which tested positive for IFN-γ induction. Based on their overall cytokine induction profiles the recombinant proteins were divided into four distinct groups. Eleven recombinant proteins induced a cytokine profile that was similar to the recall immune response induced by immune peripheral blood mononuclear cells (PBMC) stimulated with intact *E. ruminantium*. This response comprised the upregulation of cytokines associated with adaptive cellular immune responses as well as innate immunity. A successful vaccine may therefore need to contain a combination of recombinant proteins which induce both immune pathways to ensure protection against heartwater.

Keywords: *Ehrlichia ruminantium*, reverse vaccinology, Th1 cytokines.

Abbreviations: GM-CSF, granulocyte-macrophage colony-stimulating factor; iNOS, inducible nitric oxide synthase; MAP, major antigenic protein; MFT, multifunctional T; MW, molecular weight; NO, nitric oxide; ORFs, open reading frames; qPCR, quantitative real-time PCR; SFC, spot forming cells; SPMC, spots per million cells; TLR, toll-like receptor; th, transmembrane helix.

Introduction

E. ruminantium, the causative agent of heartwater, is a Gram-negative α-proteobacterium, belonging to the family Anaplasmataceae, order Rickettsiales. All organisms in the order Rickettsiales are obligate intracellular bacteria, and members of the family Anaplasmataceae are located within membrane-bound vacuoles. Heartwater is considered to be one of the most important endemic diseases of domestic livestock in southern Africa, but despite this no safe and effective vaccine is currently available. The only commercially available heartwater immunisation procedure employs the infection and treatment method, using live virulent bacteria, and it has several practical disadvantages as well as offering only limited protection against some common virulent genotypes (Allsopp, 2009).

Heartwater vaccine development has been hindered by technical difficulties, many of which derive from the fact that obligate intracellular bacteria such as *E. ruminantium* are inherently difficult to study at the molecular genetic level. The availability of the extensively annotated *E. ruminantium* genome sequence (Collins *et al.*, 2005) allowed us to apply reverse vaccinology (Rappuoli, 2000) in our attempts to develop a vaccine against heartwater. In this approach, potential vaccine candidates with selected criteria can be identified from the whole genome data with the aid of suitable computer algorithms. The most likely candidates are then screened for their ability to induce the appropriate immune responses. Reverse vaccinology was first used to identify vaccine candidates of serogroup B *Neisseria meningitidis* (Pizza *et al*., 2000). Since then the technique has been applied in developing vaccines for several other pathogens, some of which are currently in clinical development (Muzzi *et al.*, 2007; Sette & Rappuoli, 2010).

It has been documented that $CD8^+$ and $CD4^+$ lymphocytes play a role in protection against heartwater in adoptive transfer (Du Plessis *et al*., 1991) and knock out mice studies (Byrom *et al*., 2000). Furthermore, when PBMC were obtained from heartwater immune animals an increase in $CD4^+$ and $CD8^+$ lymphocyte proliferation was observed in response to *in vitro* stimulation with *E. ruminantium* antigens (Totté *et al*., 1999; Mwangi *et al*., 2002). In addition T cell growth factors (Mahan *et al*.,

1994) and IFN-γ (Totté *et al*., 1996) have been shown to inhibit *E. ruminantium* growth *in vitro*. The effect of IFN-γ may be due to upregulation of MHC class I and II expression on monocytes leading to increased antigen presentation to immune cells, or by increased phagosytosis, reactive oxygen intermediates, nitric oxide and lysosomal enzyme production. To date the potential roles of antibodies and cytotoxic T cells have not yet been demonstrated. We hypothesise that heartwater vaccine candidates should be able to elicit a protective cell mediated immune response similar to that induced by live organisms. It has been found that *E. ruminantium* proteins in the molecular weight ranges 13-18 kDa (Van Kleef *et al*., 2002) and 22-32 kDa (Esteves *et al*., 2004) induce IFN-γ production, but the specific antigens responsible for this effect have not been identified. Two recombinant *E. ruminantium* proteins, major antigenic protein (MAP) 1 and MAP2, have been shown to induce T cell lines to produce IFN-γ (Mwangi *et al*., 2002) and protection was observed in mice after DNA immunisation using the *map*1 gene (Nyika *et al*., 2002). Similarly, the protective immune response induced in sheep by four other *E. ruminantium* genes corresponded with increased IFN-γ expression (Pretorius *et al.*, 2008). We therefore aimed to identify antigens that induce strong cell mediated immune responses in *E. ruminantium*–immune T cells characterised by the expression of IFN-γ and other Th1 cytokines.

This communication describes the use of reverse vaccinology to identify potential vaccine candidates for protection against *E. ruminantium*. Open reading frames (ORFs) with selected criteria were identified from the annotated *E. ruminantium* (Welgevonden) genome sequence (Collins *et al.*, 2005). The corresponding recombinant proteins were expressed in *E. coli* and assessed for their ability to induce recall T cell responses *in vitro* using *E. ruminantium*-immune PBMC.

Materials and methods

Subcellular localisation prediction of putative gene products

The *E. ruminantium* genome annotation (Collins *et al*., 2005) contains predictions for the subcellular localisation of each putative gene product obtained using the algorithms SignalP (Nielsen *et al.*, 1997) and TMHMM2.0 (Krogh *et al*., 2001). We refined these predictions using the algorithms Phobius (Käll *et al*., 2004), PSORTb2.0 (Gardy *et al*., 2005) and CELLO (Yu *et al*., 2004) and the new results were used in the subsequent selection procedure.

In silico selection strategy

The annotation data for each putative *E. ruminantium* gene product were used as the starting point for the selection procedure. All products classified in the following categories were considered as possible vaccine candidates: surface-associated or secreted proteins, transporters, proteins putatively involved in the adaptation of bacteria to heat shock and other environmental stresses, and proteins of unknown function. We also included proteins containing tandem, tetratricopeptide or ankyrin repeats, as well as those products described as adhesins, proteases, iron-binding proteins, methyltransferases or GTPases. Homologs of proteins identified as vaccine candidates in other pathogens by means of functional genomics were also included. All ORFs with more than four predicted transmembrane helices, genes tested previously in vaccine formulations, and patented ORFs were removed from the dataset.

Expression of recombinant proteins

Directional cloning into the pET vector

Protein expression was performed using the pET102/TOPO® expression system (Invitrogen) according to the instructions of the manufacturer. Sequence specific amplification primers were designed for each of the selected ORFs to facilitate directional cloning into the pET vector by adding the sequence CACC to the 5' end of the forward primer and removing the stop codon sequence for the reverse primer. In the case of ORFs having signal peptide coding sequences the 5' primers were designed so as to omit the signal sequences. ORFs larger than 2,000 bp were divided into smaller

subfragments and we also made sure that primer sequences did not overlap large tandem repeat sequences. The ORFs were amplified in a 50 µl reaction containing 25 ng *E. ruminantium* (Welgevonden) genomic DNA, 1.25 U *Pfu* polymerase (Promega), 0.2 µM of each primer, 0.2 mM dNTPs, and 1x reaction buffer (containing 2 mM Mg^{2+}). Reaction conditions were: one cycle at 95°C for 2 min, 35 cycles of 95°C for 30 s, 50°C for 30 s and 72°C for 3 min, and a final extension at 72°C for 7 min. The amplicons were cloned into the TOPO® pET vector following the manufacturer's protocols. Clones containing inserts of the correct size were sequenced, using the TrxFus forward and T7 reverse primers, to verify the orientation and sequences of inserts and to ensure that the His-tag was in-frame.

Expression and purification of recombinant proteins

Recombinant proteins were expressed using the Overnight Express™ Autoinduction system 1 (Novagen). The recombinant proteins were extracted from the cell pellets using BugBuster® Protein Extraction Reagent (Novagen) and purified using Protino® Ni 1000 prepacked columns (Macherey-Nagel) following the manufacturer's instructions. The purified proteins were separated on Criterion™ XT precast gels (4-12% gradient, Bio-Rad) and analysed by Anti-His_6 Western blot analysis using standard procedures.

Selection of recombinant proteins that induce cellular immunity

Immune Animals

Heartwater naïve Merino sheep (s6050) and Nguni cattle (B8460, B8404 and B8347) were immunised by infection with virulent organisms (10 LD_{50}) of the Welgevonden stock of *E. ruminantium* (Brayton *et al.,* 2003). The sheep developed heartwater symptoms and were treated with tetracycline while the cattle showed a mild temperature reaction, but did not show any heartwater symptoms after the initial immunisation, and recovered without treatment. These animals were challenged four months later with the same stock to confirm their immunity to heartwater. The immunity of the cattle was boosted by an additional subcutaneous inoculation of 50 µg of inactivated organisms of the Welgevonden

stock in Montanide ISA50 adjuvant (SEPPIC, France) nine months after the challenge. All animal research was performed in accordance with the stipulations of the animal ethics committee at the ARC-Onderstepoort Veterinary Institute and the University of Pretoria animal use and care committee.

Collection of blood and isolation of PBMC

Recall immune assays were performed using purified immune PBMC. Blood (50 ml (ovine) or 200 ml (bovine)) was collected in BD Vacutainer®-EDTA tubes (Becton, Dickinson) and PBMC were isolated by density gradient centrifugation (Histopaque®-1077; Sigma-Aldrich®) according to the instructions of the manufacturer. The PBMC were washed three times, live cells were counted using GIBCO® trypan blue stain (Invitrogen) and the cells resuspended (4 x 10^6 cells/ml) in cRPMI (GIBCO® RPMI+GlutaMAX™-I (Invitrogen) supplemented with 55 mM 2-mercaptoethanol and 1% GIBCO® Pen Strep (Invitrogen)).

IFN-γ ELISPOT assays

IFN-γ expression was measured by ELISPOT assays as described previously (Sebatjane *et al.,* 2010). Briefly, ELISPOT plates (Millipore) were coated with mouse anti-bovine IFN-γ mAb CC302 (1 µg/ml). After blocking the plates with RPMI-1640 supplemented with 10% FCS, PBMC (2 x 10^5 cells/well) were incubated with the recombinant proteins (10 µg/ml; 1 µg/ml or 0.1 µg/ml), or partially purified intact *E. ruminantium* (Welgevonden) organisms isolated from infected bovine endothelial cells (1 µg/well, positive antigen), or uninfected bovine endothelial cell extract (1 µg/well, negative antigen) in a total volume of 100 µl. PBMC stimulated with ConA (5 µg/ml, Sigma) were included as a positive control, while wells containing PBMC without antigen were used as a negative control. The plates were developed as described by Sebatjane *et al.* (2010) and spot forming cells (SFC) were enumerated using an automated ELISPOT reader (Zeiss KS ELISPOTCompact 4.5). The number of SFC produced after stimulation of immune PBMC with the recombinant proteins was compared to the number of SFC produced by immune PBMC without antigen. ELISPOT samples

with 4x the number of spots per million cells (SPMC) compared to the non-stimulated cells were considered positive.

Cytokine profiling using quantitative real-time PCR (qPCR)

Isolated PBMC (sheep or bovine) were resuspended in complete RPMI-1640 medium and cells (4×10^6 cells/ml, final concentration) were stimulated with the expressed recombinant proteins (10 µg/ml; 1 µg/ml or 0.1 µg/ml) or incubated in medium alone for approximately 18 h (overnight). Cells were harvested by centrifugation at 300 *g* and TRI®-Reagent (Sigma-Aldrich®) was added to each sample. These samples were kept at –70°C until total RNA was isolated according to the TRI®-Reagent protocol. The mRNA quality was assessed by agarose gel electrophoresis and contaminating genomic DNA was removed by using the DNA-free kit (Ambion) according to the instructions of the manufacturer. Single stranded cDNA was generated using Expand reverse transcriptase (2.5 U/µl final concentration; Roche) and random hexamer primers (12.5 µg/µl final concentration) (Promega). PCR was performed using the LightCycler® FastStart DNA Master^{PLUS} SYBR Green 1 kit (Roche) according to the manufacturer's instructions. Cytokine primers used were previously optimised for bovine mRNA: IL-1α, IL-18, toll-like receptor (TLR) 2 and TLR4 (Lahmers *et al.,* 2006); IL-4, IL-10, IFN-γ, TNF-α and inducible nitric oxide synthase (iNOS) (Thacker *et al.,* 2007); IL-6, IL-8, IL-12p40 and granulocyte-macrophage colony-stimulating factor (GM-CSF) (Lee *et al.,* 2006) and GAPDH (Claerebout *et al.,* 2005). Ovine qPCR primers were: IFN-γ, IL-4 and GAPDH (Pretorius *et al.,* 2008); IL-8 and IL-1α (Smeed *et al.,* 2007); IL-18 (Markus *et al.,* 2007) and TNF-α (Budhia *et al.,* 2006). qPCR cycling conditions for all the cytokines were: 10 min at 94°C; 50 cycles of 15 s at 94°C, 30 s at 58°C and 20 s at 72°C; followed by melting starting at 65°C rising to 95°C at 0.3°C per second; and cooling to 40°C. Relative quantitation was determined with the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen, 2001). Gene expression was reported as the normalised cycle threshold (ΔCt = Ct target gene - Ct housekeeping gene and $\Delta\Delta Ct = \Delta Ct_{stimulated} - \Delta Ct_{medium}$). A one-fold increase indicates that the cytokine mRNA concentration of the stimulated sample is twice that of non-stimulated samples, and the cytokines were considered to be significantly upregulated if the mRNA increase difference was more than one-fold.

Results

Subcellular localisation of ORFs

The algorithm SignalP predicted that 66 *E. ruminantium* ORFs showed 5' sequences coding for N-terminal signal peptides, of which 13 also contained one or two predicted transmembrane helices. According to the TMHMM results 28% (247) of all ORFs in the *E. ruminantium* genome are predicted to contain at least one transmembrane helix. Forty-eight of these transmembrane helices were also predicted to be signal sequences by SignalP. When compared with the results of another algorithm, Phobius, 15 of the 48 transmembrane helices were in fact predicted to code for signal peptides. Two additional algorithms, pSORTb and CELLO, were utilised to assist in the assignment of proteins to subcellular locations (Figure 1). However the results varied significantly between the two algorithms with only 39% of the putative proteins being assigned to the same location by both programs. The majority of the shared predictions were for allocations to the cytoplasm (217 ORFs) and inner membrane (109 ORFs). Only 20 of the 888 putative proteins were predicted by both algorithms to be in the outer membrane.

In silico selection of possible vaccine candidates

ORFs coding for proteins with functional or structural similarity to proven protective antigens or known bacterial virulence factors, plus all putative proteins of unknown function, were identified from the annotated *E. ruminantium* (Welgevonden) genome sequence. In the *E. ruminantium* annotation database, 521 out of 888 putative protein coding genes have an assigned biological function. From these we selected all ORFs in the functional categories transporters, proteins involved in protection responses and adaptation, chaperones and proteases (98 ORFs in total). In addition we searched the rest of the annotation data for genes coding for adhesins, iron-binding proteins, methyltransferases and GTPases. From the ORFs that did not show any sequence similarity to well-characterised proteins we selected all ORFs classified as membrane-associated or exported proteins (175 in total), as well as ORFs for which no functional or structural information was available (141 ORFs). The selection was aided by the subcellular localisation predictions of the programs CELLO

and PSORTb. ORFs containing tandem, tetratricopeptide or ankyrin repeats were specifically included. Finally five homologs of proteins which have been described as being immunogenic or protective in other organisms were included in the list of 419 ORFs (Table 1, round 1). The number of candidates was reduced to 272 (Table 1, round 2; Table S1, Supporting information) by eliminating patented genes (United States Patent 6,593,147; Barbet *et al*., 2001) and ORFs tested previously (Louw *et al*., 2002; Nyika *et al*., 2002; Pretorius *et al*., 2002; 2007). ORFs with more than four predicted transmembrane helices were also excluded for purely practical reasons, since these are often difficult to express (Pizza *et al.*, 2000; Ariel *et al.*, 2003).

Expression and selection of recombinant proteins that induce specific Th1 cellular immune responses

Several of the selected 272 ORFs were amplified as smaller subfragments to facilitate expression in *E. coli*, giving a total of 283 genes plus subfragments. The products of 228 of these were successfully expressed, purified either in a water-soluble form or as insoluble inclusion bodies, and used in IFN-γ ELISPOT assays. Initially the PBMC obtained from a heartwater immune sheep (s6050) were used to screen 38 of the recombinant proteins. PBMC isolated from 50 ml of blood from this animal (the maximum amount of blood ethically permitted per bleed) did not yield enough cells to screen all 228 recombinant proteins. The remainder of the recombinant proteins (190) were therefore screened at three different concentrations using PBMC isolated from three immune cattle. IFN-γ producing cells were induced by 3 out of 38 recombinant proteins tested using immune sheep PBMC, and by 32 of the 190 recombinant proteins in PBMC isolated from cattle. The number of IFN-γ SPMC for positive samples were significantly higher than those for non-stimulated PBMC and for those that did not induce IFN-γ in immune PBMC. The recombinant proteins that did not induce detectable IFN-γ protein were further screened using qPCR for their ability to induce the transcription of IFN-γ mRNA using the same bovine PBMC and protein concentrations as those used in the ELISPOT assay. This yielded an additional 11 recombinant proteins that induced low levels of IFN-γ mRNA.

All 46 of the recombinant proteins that induced either IFN-γ or its transcript were screened for their ability to induce additional cytokines in the same sheep or bovine PBMC samples which had shown induced IFN-γ at optimum concentration. Additional cytokines were induced by 26 of these IFN-γ-inducing proteins (Table 2). Based on their overall cytokine induction profiles the proteins could be divided into four separate groups designated A to D. Proteins which induced cytokine profile A in bovine PBMC induced early increase in IFN-γ mRNA levels (1-8 fold after 4 h incubation) which significantly increased to 16-809 fold overnight. With some exceptions, the upregulation of IFN-γ mRNA coincided with increased mRNA levels of iNOS; TNF-α; IL-1α; IL-18 and GM-CSF. In addition, increased levels of IL-4 and IL-10 mRNA were detected after stimulation with recombinant proteins coded by Erum7280, -7620 and -8010, but the IFN-γ:IL-4 mRNA ratios (246:10; 809:7; and 33:5 respectively) remained polarised towards Th1 cellular immunity. Cytokine profile A was, except for the absence of TLR2 expression, similar to that induced by intact *E. ruminantium* in two of the cattle tested. PBMC stimulated with intact *E. ruminantium* had high levels of overnight IFN-γ mRNA (64 and 232 fold increase), that corresponded with high levels of IFN-γ protein (200 and 985 SPMC). In contrast, in spite of the high levels of IFN-γ mRNA induced by Erum7280, and -7620, (246 and 809 fold increase respectively) low levels of IFN-γ (20 and 48 SPMC) were detected with ELISPOT. The recombinant proteins inducing cytokine profile A were the only proteins shown to induce IFN-γ mRNA transcripts after overnight incubation. Although not measured after 4 h, increased levels of IFN-γ mRNA were also detected after overnight incubation when three recombinant proteins, coded by Erum5270, -5400 and -8050, were tested using sheep PBMC. All three of these proteins also induced upregulation of IFN-γ, TNF-α, IL-1α, and IL-18 mRNA in immune PBMC stimulated overnight with 1 μg/ml protein (Table 2). Hence, these three proteins were also grouped into cytokine profile A. Characteristics of all the proteins in cytokine profile A are summarised in Table 3.

Four recombinant proteins (coded by Erum1040_3; -1050; -1990; -8510) induced cytokine profile B. IFN-γ protein expression levels were similar to those induced by proteins that induced cytokine profile A. However, cytokine profile B recombinant proteins induced increased IFN-γ mRNA transcription at 4 h after stimulation, but IFN-γ mRNA expression was much lower, or completely

absent, after overnight incubation (Table 2). A decrease in mRNA levels between the 4 h and overnight samples was also detected for TNF-α; IL-18; iNOS and GM-CSF.

Proteins that induced cytokine profiles C and D did not induce IFN-γ but did induce IFN-γ mRNA transcripts. These two profiles were also similar to cytokine profile B in that IFN-γ mRNA expression could be measured after 4 h but was decreased or absent after overnight incubation. Proteins that induced cytokine profile C also induced high levels of IL-18; GM-CSF; iNOS and IL-8, while proteins in cytokine profile D induced TNF-α; GM-CSF; iNOS and IL-1α after 4 h incubation.

Discussion

The purpose of this work was to identify *E. ruminantium* vaccine candidates that induce immune responses similar to those induced by whole organisms. Our approach involved four steps: (i) *in silico* selection from the annotated genome; (ii) cloning of selected genes, and the expression and purification of the corresponding recombinant proteins; (iii) identification of recombinant proteins that induce IFN-γ production in ELISPOT and/or qPCR assays; and (iv) cytokine profiling of IFN-γ-inducing recombinant proteins using qPCR.

A large part of our initial *in silico* selection consisted of putative proteins of unknown function, as well as surface-associated and exported proteins. Although surface-exposed antigens are associated with humoral responses, surface-associated proteins of intracellular pathogens have been reported to elicit $CD4^+$ and/or $CD8^+$ responses (Sampson, 2011). We employed several programs to predict the subcellular location of putative proteins, but there was very little agreement between the different algorithms. Other workers have reported this situation (Sprenger *et al.*, 2006), and without experimental evidence it is not possible to determine which algorithm is the superior predictor. In addition, it is currently impossible to predict how changing conditions *in vivo* will affect the expression of different membrane components (Grandi, 2003). One must therefore be cautious in the interpretation of localisation predictions, but algorithms are constantly improving (Choo *et al.*, 2009;

Yu *et al*., 2010) and more experimental data are becoming available, hence future predictions ought to be more reliable.

Many *E. ruminantium* proteins of unknown function, and some of the membrane-associated proteins, contained tetratricopeptide or ankyrin repeat domains or tandem repeats. All three repeat elements have been implicated in host-pathogen interactions (Core & Perego, 2003; De la Fuente *et al*., 2004; Wakeel *et al*., 2010; Zhu *et al*., 2009), hence these genes may be considered as vaccine candidates. Other possibly important categories for protection against bacterial infection include type IV secretion system proteins (Lopez *et al.*, 2007; Juhas *et al*., 2008), transporters, particularly the ABC transport system (Brown *et al*., 2001; Pretorius *et al*., 2007; Roset *et al*., 2004) and proteases (Miyoshi & Shinoda, 2000; Ariel *et al*., 2003, Myers *et al*., 2007).

T cell responses characterised by the expression of IFN-γ are essential in protection against *E. ruminantium* infection (Totté *et al*. 1999; Mwangi *et al.,* 2002) which was the rationale behind attempting to determine whether any of our *E. ruminantium* recombinant proteins induced IFN-γ production in *E. ruminantium*–immune lymphocytes *in vitro*. Initially ELISPOT and qPCR assays used sheep PBMC, but later PBMC isolated from three immune cattle were used as more cells could be harvested from the cattle and the bovine PBMC gave lower background stimulation values. In general, the PBMC isolated from one of the animals responded to a specific protein tested at a given time point. Outbred cattle and sheep were used in this study each with unique MHC alleles (personal communication, N. Thema), thus these genetic differences may result in immune response variations as has been found in other studies using outbred animals (Babiuk *et al.*, 2003). Another factor to consider is that the *E. ruminantium* specific lymphocytes may have been present in the blood at concentrations too low to allow effective measurement, or were not circulating at the time of isolation (Kennedy *et al.*, 2002). A better correlation between the three bovine PBMC and the selection of additional IFN-γ producing proteins may be achieved by repeating the immune assays at more than one time point. The use of T

cell lines specific to each recombinant protein would yield more sensitive and specific data and will in future be used for testing the most promising vaccine candidates. A total of 228 expressed recombinant proteins were assayed using immune sheep or bovine PBMC. Either soluble and/or insoluble fractions were included because insolubility and protein denaturation usually do not affect the outcome of cellular immunological assays (Leung *et al.,* 2004). Although the recombinant proteins had been purified to differing degrees all indications were that the positive responses obtained were *E. ruminantium* specific. A total of 35 recombinant proteins could induce IFN-γ mRNA and protein in either sheep or bovine PBMC, while 11 recombinant proteins induced IFN-γ mRNA only (bovine PBMC). The cytokine profiles of these 46 IFN-γ-inducing recombinant proteins were determined and compared to those induced after stimulation with intact *E. ruminantium*.

The cytokine profile induced in immune bovine PBMC by intact *E. ruminantium* (cytokine profile A) was characterised by increased IFN-γ; TNF-α and iNOS mRNA from 4 h to overnight. iNOS expression is induced after IFN-γ and/or TNF-α /receptor interaction (Goff *et al.,* 2002; Bogdan, 2001) and regulates the production of nitric oxide (NO) that is detrimental to intracellular pathogens (Kobayashi, 2010). The toll-like receptors TLR4 and TLR2 mRNAs are also upregulated in the presence of IFN-γ and TNF-α (Winder *et al.,* 2009) and induce innate immunity (Werling *et al.,* 2006). Intact *E. ruminantium* also induced myeloid-specific cytokine mRNAs including IL-18; IL-1α; IL-6 and GM-CSF. TNF-α, IL-12 and IL-18 may also, in synergy with IL-2 produced by $CD4^+$ T cells after interaction with antigen presenting cells, induce a recall NK cell response that is essential for early pathogen clearance (Horowitz *et al.,* 2010). Activated NK cells are an important source of IFN-γ during early recall immune responses in both viral (Horowitz *et al.,* 2010) and intracellular bacterial infections (Haeberlein *et al.,* 2010; Humann & Lenz, 2010; Waters *et al.,* 2011). IL-4 and IL-10 cytokine mRNAs were also upregulated by intact *E. ruminantium*. The IL-4 secretion might indicate humoral immune activation, but it is more likely that IL-10 and IL-4 are produced to regulate high levels of NO (Bogdan, 2001; Goff *et al.,* 2002). Pro-inflammatory cytokines and chemokines (IL-6 and IL-8), also detected in intact *E. ruminantium*-stimulated PBMC, assist in the recruitment

and activation of immune cells (Lee *et al.,* 2004). Protective immunity against *E. ruminantium* may therefore be mediated not only by $CD4^+$ and $CD8^+$ T cells producing IFN-γ and TNF-α (Totté *et al.,* 1999), but may also require activation of other immune pathways, such as recall NK cell responses and innate immunity mediated by specific cytokines and chemokines. Thus, the vaccine candidates selected should induce these appropriate cytokines and the final vaccine formulation should activate both innate and cellular immunity.

Cytokine profile A was measured for a total of 11 recombinant proteins (Table 3), three that were selected using sheep PBMC and eight that were selected using bovine PBMC. These were the only recombinant proteins which induced IFN-γ protein production that correlated with relatively low levels of IFN-γ mRNA at 4 h that increased markedly overnight. These recombinant proteins also induced elevated TNF-α, GM-CSF, iNOS and TLR4 expression. IFN-γ protein expression was however at lower levels than that induced by intact *E. ruminantium*. All of these IFN-γ inducing proteins, with the exception of those coded by Erum0660_1 and Erum8050, were of low molecular weight, correlating with previous studies indicating that low molecular weight *E. ruminantium* proteins induced IFN-γ production (Van Kleef *et al.,* 2002; Esteves *et al.,* 2004; Sebatjane *et al.,* 2010). Cytokine profile A proteins could also induce TLR4, probably via IFN-γ and TNF-α induction. Proteins that induce the activation of these receptors may thus be useful in the promotion of innate immunity when included as part of a vaccine against heartwater. Recombinant proteins inducing cytokine profiles C and D may also be useful in vaccine formulations. Cytokine profile C proteins induced high levels of IL-18, a cytokine important for IFN-γ production and the development of memory CTLs (Iwai *et al.,* 2008). In contrast, relatively low levels of IL-18 mRNA were detected in PBMC stimulated with intact *E. ruminantium*; this may be attributed to the presence of IL-18 inhibitory proteins similar to those found for *E. chaffeensis* (Wakeel *et al.,* 2010). This down regulation will possibly reduce any CTL response directed against *E. ruminantium* antigen, thus promoting survival of the pathogen. Recombinant proteins that induced cytokine profile C and D also induced GM-CSF, a cytokine used in DNA vaccines targeted to dendritic cells (DC-based vaccines) (Le *et al.,* 2010).

It should be noted that IFN-γ expression alone is not always indicative of protection against *E. ruminantium* infection *in vivo* (Vachiéry *et al.,* 2006). Recent studies on protective immunity against *Mycobacterium tuberculosis* have shown a better correlation between protection and the number of multifunctional T (MFT) cells which were detected in immune pulmonary and spleen cell preparations (Derrick *et al.,* 2011). MFT cells are single cells that produce high concentrations of the cytokines IFN-γ; TNF-α and IL-2. Although we did not determine multiple cytokines produced by a single T cell using flow cytometry, the recombinant proteins in cytokine profile A did show simultaneous upregulation of IFN-γ and TNF-α. These proteins are therefore the most likely to induce protective immunity to heartwater. Future studies should focus on determining whether both of these cytokines are produced by the same $CD4^+$ T cell and whether the activation of these MFT cells is linked to protection against heartwater.

In conclusion, using a reverse vaccinology strategy, we could select recombinant proteins that induced a cellular immune response in immune bovine PBMC that was similar to the response induced by intact *E. ruminantium.* It remains to be determined whether these antigens are capable of inducing protective immune responses against heartwater infection *in vivo*. It is likely that a future recombinant vaccine formulation should contain genes and/or proteins that induce the different immune pathways induced by intact organisms.

Acknowledgements

This work was supported by the South African Department of Agriculture OV9/23/C167 grant and the FP6 EU INCO-DEV EPIGENEVAC FP6-003713 grant.

References

Allsopp, B.A., 2009. Trends in the control of heartwater. Onderstepoort J. Vet. Res. 76, 81-88.

Ariel, N., Zvi, A., Makarova, K.S., Chitlaru, T., Elhanany, E., Velan, B., Cohen, S., Friedlander, A.M., Shafferman, A., 2003. Genome-based bioinformatic selection of chromosomal *Bacillus anthracis* putative vaccine candidates coupled with proteomic identification of surface-associated antigens. Infect. Immun. 71, 4563-4579.

Babiuk, L.A., Pontarollo, R., Babiuk, S., Loehr, B., Van Drunen Littel-Van Den Hurk, S., 2003. Induction of immune responses by DNA vaccines in large animals. Vaccine. 21, 649-658.

Barbet, A.F., Whitmire, W.M., Kamper, S.M., Simbi, B.H., Ganta, R.R., Moreland, A.L., Mwangi, D.M., McGuire, T.C., Mahan, S.M., 2001. A subset of *Cowdria ruminantium* genes important for immune recognition and protection. Gene 275, 287-298.

Bogdan, C., 2001. Nitric oxide and the immune response. Nature Immunol. 2, 907-916.

Brayton, K.A., Collins, N.E., van Strijp, F., Allsopp, B.A., 2003. Preparation of *Ehrlichia ruminantium* challenge material for quantifiable and reproducible challenge in mice and sheep. Vet. Parasitol. 112, 63-73.

Brown, J.S., Ogunniyi, A.D., Woodrow, M.C., Holden, D.W., Paton, J.C., 2001. Immunization with components of two iron uptake ABC transporters protects mice against systemic *Streptococcus pneumoniae* infection. Infect. Immun. 69, 6702-6706.

Budhia, S., Haring, L.F., McConnell, I., Blacklaws, B.A., 2006. Quantitation of ovine cytokine mRNA by real-time RT-PCR. J. Immunol. Methods 309, 160-172.

Byrom, B., Barbet, A.F., Obwolo, M., Mahan, S.M., 2000. CD8(+) T cell knockout mice are less susceptible to *Cowdria ruminantium* infection than athymic, CD4(+) T cell knockout, and normal C57BL/6 mice. Vet. Parasitol. 93, 159-172.

Choo, K.H., Tan, T.W., Ranganathan, S., 2009. A comprehensive assessment of N-terminal signal peptides prediction methods. BMC Bioinformatics 10, S2.

Claerebout, E., VerCauteren, I., Geldhof, P., Olbrechts, A., Zarlenga, D.S., Goddeeris, B.M., Vercruysse, J., 2005. Cytokine responses in immunized and non-immunized calves after *Ostertagia ostertagi* infection. Parasite Immunol. 27, 325-331.

Collins, N.E., Liebenberg, J., de Villiers, E.P., Brayton, K.A., Louw, E., Pretorius, A., Faber, F.E., van Heerden, H., Josemans, A., van Kleef, M., Steyn, H.C., van Strijp, M.F., Zweygarth, E., Jongejan, F., Maillard, J.C., Berthier, D., Botha, M., Joubert, F., Corton, C.H., Thomson, N.R., Allsopp, M.T., Allsopp, B.A., 2005. The genome of the heartwater agent *Ehrlichia ruminantium* contains multiple tandem repeats of actively variable copy number. Proc. Natl. Acad. Sci. U. S. A. 102, 838-843.

Core, L., Perego, M., 2003. TPR-mediated interaction of RapC with ComA inhibits response regulator-DNA binding for competence development in *Bacillus subtilis*. Mol. Microbiol. 49, 1509-1522.

De la Fuente, J., Garcia-Garcia, J.C., Barbet, A.F., Blouin, E.F., Kocan, K.M., 2004. Adhesion of outer membrane proteins containing tandem repeats of *Anaplasma* and *Ehrlichia* species (Rickettsiales: Anaplasmataceae) to tick cells. Vet. Microbiol. 98, 313-322.

Derrick, S.C., Yabe, I.M., Yang, A., Morris, S.L., 2011. Vaccine-induced anti-tuberculosis protective immunity in mice correlates with the magnitude and quality of multifunctional CD4 T cells. Vaccine 29, 2902-2909.

Du Plessis, J.L., Berche, P., Van Gas, L., 1991. T cell-mediated immunity to *Cowdria ruminantium* in mice: the protective role of Lyt-2+ T cells. Onderstepoort J. Vet. Res. 58, 171-179.

Esteves, I., Martinez, D., Totté, P., 2004. Identification of *Ehrlichia ruminantium* (Gardel strain) IFN-gamma inducing proteins after vaccination with a killed vaccine. Vet. Microbiol. 100, 233-240.

Gardy, J.L., Laird, M.R., Chen, F., Rey, S., Walsh, C.J., Ester, M., Brinkman, F.S., 2005. PSORTb v.2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. Bioinformatics 21, 617-623.

Goff, W.L., Johnson, W.C., Parish, S.M., Barrington, G.M., Elsasser, T.H., Davis, W.C., Valdez, R.A., 2002. IL-4 and IL-10 inhibition of IFN-gamma- and TNF-alpha-dependent nitric oxide production from bovine mononuclear phagocytes exposed to *Babesia bovis* merozoites. Vet. Immunol. Immunopathol. 84, 237-251.

Grandi, G., 2003. Rational antibacterial vaccine design through genomic technologies. Int. J. Parasitol. 33, 615-620.

Haeberlein, S., Sebald, H., Bogdan, C., Schleicher, U., 2010. IL-18, but not IL-15, contributes to the IL-12-dependent induction of NK-cell effector functions by *Leishmania infantum* in vivo. Eur. J. Immunol. 40, 1708-1717.

Horowitz, A., Behrens, R.H., Okell, L., Fooks, A.R., Riley, E.M., 2010. NK cells as effectors of acquired immune responses: effector CD4+ T cell-dependent activation of NK cells following vaccination. J. Immunol. 185, 2808-2818.

Humann, J., Lenz, L.L., 2010. Activation of naive NK cells in response to *Listeria monocytogenes* requires IL-18 and contact with infected dendritic cells. J. Immunol. 184, 5172–5178.

Iwai, Y., Hemmi, H., Mizenina, O., Kuroda, S., Suda, K., Steinman, R.M., 2008. An IFN-gamma-IL-18 signaling loop accelerates memory CD8+ T cell proliferation. PLoS One 3, e2404.

Juhas, M., Crook, D.W., Hood, D.W., 2008. Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. Cell Microbiol. 10, 2377-2386.

Käll, L., Krogh, A., Sonnhammer, E.L., 2004. A combined transmembrane topology and signal peptide prediction method. J. Mol. Biol. 338, 1027-1036.

Kennedy, H.E., Welsh, M.D., Bryson, D.G., Cassidy, J.P., Forster, F.I., Howard, C.J., Collins, R.A., Pollock, J.M., 2002. Modulation of immune responses to *Mycobacterium bovis* in cattle depleted of $WC1^+$ γδ T cells. Infect. Immun. 70, 1488-1500.

Kobayashi, Y., 2010. The regulatory role of nitric oxide in proinflammatory cytokine expression during the induction and resolution of inflammation. J. Leukoc. Biol. 88, 1157-1162.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305, 567-580.

Lahmers, K.K., Hedges, J.F., Jutila, M.A., Deng, M., Abrahamsen, M.S., Brown, W.C., 2006. Comparative gene expression by $WC1^+$ γδ and $CD4^+$ αβ T lymphocytes, which respond to *Anaplasma marginale,* demonstrates higher expression of chemokines and other myeloid cell-associated genes by $WC1^+$ γδ T cells. J. Leukoc. Biol. 80, 939-952.

Le, D.T., Pardoll, D.M., Jaffee, E.M., 2010. Cellular vaccine approaches. Cancer J. 16, 304-310.

Lee, J-K., Kim, S-H., Lewis, E.C., Azam, T., Reznikov, L.L., Dinarello, C.A., 2004. Differences in signaling pathways by IL-1β and IL-18. Proc. Natl. Acad. Sci. U. S. A. 101, 8815-8820.

Lee, J-W., Bannerman, D.D., Max J. Paape, M.J., Huang, M-K., Zhao, X., 2006. Characterization of cytokine expression in milk somatic cells during intramammary infections with *Escherichia coli* or *Staphylococcus aureus* by real-time PCR. Vet. Res. 37, 219-229.

Leung, W.H., Meng, Z.Q., Hui, G., Ho, W.K., 2004. Expression of an immunologically reactive merozoite surface protein ($MSP\text{-}1_{42}$) in *E. coli*. Biochim. Biophys. Acta 1675, 62-70.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C (T)) method. Methods 25, 402-408.

Lopez, J.E., Palmer, G.H., Brayton, K.A., Dark, M.J., Leach, S.E., Brown, W.C., 2007. Immunogenicity of *Anaplasma marginale* type IV secretion system proteins in a protective outer membrane vaccine. Infect. Immun. 75, 2333-2342.

Louw, E., Brayton, K.A., Collins, N.E., Pretorius, A., Van Strijp, F., Allsopp, B.A., 2002. Sequencing of a 15-kb *Ehrlichia ruminantium* clone and evaluation of the *cpg1* open reading frame for protection against heartwater. Ann. N. Y. Acad. Sci. 969, 147-150.

Mahan, S.M., Smith, G.E. and Byrom, B., 1994. Concanavalin A stimulated bovine T-cell supernatants inhibit growth of *Cowdria ruminantium* in bovine endothelial cells in vitro. Infect. Immun. 62, 747–750.

Markus, T., Hansson, S., Amer-Wåhlin, I., Hellström-Westas, L., Saugstad, O.D., Ley, D., 2007. Cerebral inflammatory response after fetal asphyxia and hyperoxic resuscitation in newborn sheep. Pediatr. Res. 62, 71-77.

Miyoshi, S., Shinoda, S., 2000. Microbial metalloproteases and pathogenesis. Microbes Infect. 2, 91-98.

Muzzi, A., Masignani, V., Rappuoli, R., 2007. The pan-genome: towards a knowledge-based discovery of novel targets for vaccines and antibacterials. Drug Discov. Today 12, 429-439.

Mwangi, D.M., McKeever, D.J., Nyanjui, J.K., Barbet, A.F., Mahan, S.M., 2002. Immunisation of cattle against heartwater by infection with *Cowdria ruminantium* elicits T lymphocytes that recognise major antigenic proteins 1 and 2 of the agent. Vet. Immunol. Immunopathol. 85, 23-32.

Myers, G.S., Parker, D., Al-Hasani, K., Kennan, R.M., Seemann, T., Ren, Q., Badger, J.H., Selengut, J.D., Deboy, R.T., Tettelin, H., Boyce, J.D., McCarl, V.P., Han, X., Nelson, W.C., Madupu, R., Mohamoud, Y., Holley, T., Fedorova, N., Khouri, H., Bottomley, S.P., Whittington, R.J., Adler, B., Songer, J.G., Rood, J.I., Paulsen, I.T., 2007. Genome sequence and identification of candidate vaccine antigens from the animal pathogen *Dichelobacter nodosus*. Nat. Biotechnol. 25, 569-575.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1-6.

Nyika, A., Barbet, A.F., Burridge, M.J., Mahan, S.M., 2002. DNA vaccination with *map1* gene followed by protein boost augments protection against challenge with *Cowdria ruminantium*, the agent of heartwater. Vaccine 20, 1215-1225.

Pizza, M., Scarlato, V., Masignani, V., Giuliani, M.M., Aricò, B., Comanducci, M., Jennings, G.T., Baldi, L., Bartolini, E., Capecchi, B., Galeotti, C.L., Luzzi, E., Manetti, R., Marchetti, E., Mora, M., Nuti, S., Ratti, G., Santini, L., Savino, S., Scarselli, M., Storni, E., Zuo, P., Broeker, M., Hundt, E., Knapp, B., Blair, E., Mason, T., Tettelin, H., Hood, D.W., Jeffries, A.C., Saunders, N.J., Granoff, D.M., Venter, J.C., Moxon, E.R., Grandi, G., Rappuoli, R., 2000. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 287, 1816-1820.

Pretorius, A., Van Strijp, F., Brayton, K.A., Collins, N.E., Allsopp, B.A., 2002. Genetic immunization with *Ehrlichia ruminantium* GroEL and GroES homologues. Ann. N. Y. Acad. Sci. 969, 151-154.

Pretorius, A., Collins, N.E., Steyn, H.C., van Strijp, F., van Kleef, M., Allsopp, B.A., 2007. Protection against heartwater by DNA immunisation with four *Ehrlichia ruminantium* open reading frames. Vaccine 25, 2316-2324.

Pretorius, A., van Kleef, M., Collins, N.E., Tshikudo, N., Louw, E., Faber, F.E., van Strijp, M.F., Allsopp, B.A., 2008. A heterologous prime/boost immunisation strategy protects against virulent *E. ruminantium* Welgevonden needle challenge but not against tick challenge. Vaccine 26, 4363-4371.

Rappuoli, R., 2000. Reverse vaccinology. Curr. Opin. Microbiol. 3, 445-450.

Roset, M.S., Ciocchini, A.E., Ugalde, R.A., Iñón de Iannino, N., 2004. Molecular cloning and characterization of *cgt*, the *Brucella abortus* cyclic beta-1,2-glucan transporter gene, and its role in virulence. Infect. Immun. 72, 2263-2271.

Sampson, S.L., 2011. Mycobacterial PE/PPE proteins at the host-pathogen interface. Clin. Dev. Immunol. 2011, 497203.

Sebatjane, S.I., Pretorius, A., Liebenberg, J., Steyn, H., Van Kleef, M., 2010. *In vitro* and *in vivo* evaluation of five low molecular weight proteins of *Ehrlichia ruminantium* as potential vaccine components. Vet. Immunol. Immunopathol. 137, 217-225.

Sette, A., Rappuoli, R., 2010. Reverse vaccinology: developing vaccines in the era of genomics. Immunity 33, 530-541.

Smeed, J.A., Watkins, C.A., Rhind, S.M., Hopkins, J., 2007. Differential cytokine gene expression profiles in the three pathological forms of sheep paratuberculosis. BMC Vet. Res. 3, 18.

Sprenger, J., Fink, J.L., Teasdale, R.D., 2006. Evaluation and comparison of mammalian subcellular localization prediction methods. BMC Bioinformatics 7 Suppl 5, S3.

Thacker, T.C., Palmer, M.V., Waters, W.R., 2007. Associations between cytokine gene expression and pathology in *Mycobacterium bovis* infected cattle. Vet. Immunol. Immunopathol. 119, 204-213.

Totté, P., Vachiery, N., Martinez, D., Trap, I., Ballingall, K.T., MacHugh, N.D., Bensaid, A., Wérenne, J., 1996. Recombinant bovine interferon gamma inhibits the growth of *Cowdria ruminantium* but fails to induce major histocompatibility complex class II following infection of endothelial cells. Vet. Immunol. Immunopathol. 53, 61-71.

Totté, P., Bensaid, A., Mahan, S.M., Martinez, D., McKeever, D.J., 1999. Immune responses to *Cowdria ruminantium* infections. Parasitol. Today 15, 286-290.

Vachiéry, N., Lefrançois, T., Esteves, I., Molia, S., Sheikboudou, C., Kandassamy, Y., Martinez, D., 2006. Optimisation of the inactivated vaccine dose against heartwater and in vitro quantification of *Ehrlichia ruminantium* challenge material. Vaccine 24, 4747-4756.

Van Kleef, M., Macmillan, H., Gunter, N.J., Zweygarth, E., Allsopp, B.A., Shkap, V., Du Plessis, D.H., Brown, W.C., 2002. Low molecular weight proteins of *Cowdria ruminantium* (Welgevonden isolate) induce bovine CD4+-enriched T-cells to proliferate and produce interferon-gamma. Vet. Microbiol. 85, 259-273.

Wakeel, A., Zhu, B., Yu, X.J., McBride, J.W., 2010. New insights into molecular *Ehrlichia chaffeensis*-host interactions. Microbes Infect. 12, 337-345.

Waters, W.R., Palmer, M.V., Thacker, T.C., Davis, W.C., Sreevatsan, S., Coussens, P., Meade, K.G., Hope, J.C., Estes, D.M., 2011. Tuberculosis immunity: opportunities from studies with cattle. Clin. Dev. Immunol. 2011, 768542.

Werling, D., Piercy, J., Coffey, T.J., 2006. Expression of TOLL-like receptors (TLR) by bovine antigen-presenting cells—Potential role in pathogen discrimination? Vet. Immunol. Immunopathol. 112, 2-11.

Winder, A.A., Wohlford-Lenane, C., Scheetz, T.E., Nardy, B.N., Manzel, L.J., Look, D.C., McCray, P.B. Jr., 2009. Differential effects of cytokines and corticosteroids on toll-like receptor 2 expression and activity in human airway epithelia. Respir. Res. 10, 96.

Yu, C.S., Lin, C.J., Hwang, J.K., 2004. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions. Protein Sci. 13, 1402-1406.

Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., Dao, P., Sahinalp, S.C., Ester, M., Foster, L.J., Brinkman, F.S., 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics 26, 1608-1615.

Zhu, B., Nethery, K.A., Kuriakose, J.A., Wakeel, A., Zhang, X., McBride, J.W., 2009. Nuclear translocated *Ehrlichia chaffeensis* ankyrin protein interacts with a specific adenine-rich motif of host promoter and intronic Alu elements. Infect. Immun. 77, 4243-4255.

Table 1. Number of ORFs identified as possible vaccine candidates grouped according to their putative function.

	Round 1	**Round 2**
Unknown function	80	70
Unknown, some miscellaneous information	61	23
Membrane-associated	148	98
Exported	27	24
Type IV secretion system	14	9
ABC transporters	16	8
Other transporters	32	12
Proteases	19	12
Chaperones	13	7
Protection responses and adaptation	4	4
Other*	5	5
Total	**419**	**272**

* ORFs from other categories shown to be protective or immunogenic in other organisms.

Table 2. Cytokine profiles induced by recombinant proteins in PBMC from immune sheep or cattle determined by ELISPOT and qPCR assays. The qPCR was done using the same sheep or bovine PBMC that produced IFN-γ spot producing cells and were stimulated with protein at the same concentration indicated for the ELISPOT assay.

						qPCR: Cytokine mRNA fold increase Sheep (overnight) **and Bovine: (4 h;** overnight)												
		Protein used to stimulate PBMC (Erum ID)	**Immune sheep/ bovine number**	**Protein conc. (µg/ml)**	**ELISPOT[a] SPMC**	**IFN-γ**	**TNF-α**	**IL-18**	**IL-12-p40**	**GM-CSF**	**TLR4**	**TLR2**	**iNOS**	**IL-8**	**IL-1α**	**IL-4**	**IL-6**	**IL-10**
Sheep PBMC	Cytokine profile A	Erum Ag	6050	1	244	469	n/d	n/d	n/d	n/d	n/d	n/d	n/d	n/d	n/d	-	n/d	n/d
		Erum5270	6050	1	27	6	9	3	n/d	n/d	n/d	n/d	n/d	-	5	2	n/d	n/d
		Erum5400	6050	1	35	31	7	2	n/d	n/d	n/d	n/d	n/d	-	7	-	n/d	n/d
		Erum8050	6050	1	27	9	7	3	n/d	n/d	n/d	n/d	n/d	-	10	-	n/d	n/d
Bovine PBMC	Cytokine profile A	Erum Ag	8347	1	493	n/d[c]	n/d	n/d	-	n/d	n/d	n/d	n/d	n/d	n/d	n/d	n/d	n/d
		Erum Ag	8404	1	200	**2**; 64	**1**; 10	**2**; 1	**-**	**1**; 11	**1**; 9	**1;** 4	**1**; 18	**1**; 7	**1**; 16	**1**; 11	**1**; 7	**1**; 9
		Erum Ag	8460	1	985	**1**; 232	**2**; 18	**1**; 3	**-**	**1**; 9	**1**; 11	**-;** 12	**1**; 21	-; 21	**1**; 32	**1**; 3	**2**; 48	**-** ;10
		Erum0660_1	8460	10	85	**3**; 16	**6**; 10	**3**; 3	**4;** 3	**11**;58	-;5	**-;** 3	**6**; 39	-; 2	**4**; 11	**-**	**2**; 1	**6**; 4
		Erum1150	8347; 8404;8460	1; 10; 1	320; 85; 77[b]	**2**; 8	**2**; 5	**-**	**-**	-; 8	**-**	**-**	**-**	**-**	**-**	**-**	-; 18	**-**
		Erum7130	8404	1	5	**9**; 13	**7**; 15	**-**	**-**	**11**; 7	-; 8	**-**	**24**; 100	-; 10	**4**; 2	**-**	**-**	-; 7
		Erum7140	8347; 8460	1; 10	52; 27	**8**; 49	-; 2	-; 4	**-**	-; 13	-; 4	**-**	**14**; 132	-; 21	-; 13	**-**	**-**	**-**
		Erum7280	8460	0.1	48	**3**; 246	**8**; 19	-; 10	**-**	**13**; 11	-; 11	**-**	**24**; 87	**-**	-; 4	-; 10	**-**	**5**; 9
		Erum7620	8460	10	20	**1**; 809	-; 24	**-**	**-**	-; 9	-; 10	**-**	-; 170	**-**	-; 4	-; 7	**-**	-; 11
		Erum8010	8460	0.1	42	**4**; 33	-; 10	-; 8	**-**	-; 18	-; 9	**-**	-; 516	-; 6	-; 5	-; 5	**-**	-; 7
		Erum8460	8347	0.1	52	**6**; 42	-; 6	-; 4	**-;** 10	**20**; 17	-; 3	**-**	**21**; 76	-; 19	-; 5	**-**	**-**	-; 2
	Cytokine profile B	Erum1040_3	8460	1	37	**4**; -[d]	**7**; -	**3**; -	**-**	-; 11	**-**	**-**	**10**; 2	**-**	**4**; -	**-**	**4**; -	**5**; -
		Erum1050	8347; 8404; 8460	10; 10;1	158; 73; 20	**2**; -	**6**; -	**2**; -	**4**; -	**10**; 2	**-**	**-**	**3**; 4	**-**	**3**; 2	**-**	**-**	**3**; -
		Erum1990	8460	10	58	**4**; 1	**8**; -	**-**	**4; -**	**10**; 4	**-**	**-**	**11**; 3	**-**	**4**; 2	**-**	**2**; -	**3**; -
		Erum8510	8460	0.1	35	**14; -**	**8; -**	**2; -**	**-**	**12; -**	**5; -**	**-**	**16**; 5	**-**	**-**	**-**	**-**	**-**
	Cytokine profile C	Erum2330	8404	1	0	**4**; -	**-**	**4**; -	**-**	**29**; 3	**-**	**-**	**-**; 5	-; 2	**-**	-; 2	**-**	**-**
		Erum4450	8460	10	0	**3**; 2	**-**	**118**; 19	**-**	**7**; 3	**-**	**-**	**-**; 38	-; 8	**-**	-; 2	**-**	**-**
		Erum4530	8460	10	0	**5**; 2	**-**	**184**; 15	**-**	**7**; 3	**-**	**-**	**14**; 26	**25**; 4	**-**	**-**	**-**	**5**; 2
		Erum4640	8460	1	0	**2**; -	**-**	**11**; -	**-**	**6**; 2	**4**; -	**-**	**4**; 7	**-**	**-**	**4**; -	**-**	-; 12
		Erum5420	8460	1	0	**4**; 2	**-**	**138**; 33	**-**	**8**; 6	**-**	**-**	**19**; 10	**28**; 6	-; 3	-; 3	**-**	-; 2
		Erum6200	8460	10	0	**3**; -	**-**	**87**; 8	**-**	**7**; -	**-**	**-**	**21**; 2	**16**; -	-; 3	-; 3	**-**	-; 2
		Erum7110	8404	1	0	**3**; -	**-**	**46;** 5	**-**	**4;** 2	**-**	**-**	**18;** 5	-; 3	**-**	**-**	**-**	**-**
	Cytokine profile D	Erum1851	8404	0.1	0	**3**; -	**9**; -	**-**	**-**	**12**; -	**-**	**-**	**7**; -	**-**	**4**; -	**-**	**-**	**-**
		Erum1900	8404	0.1	0	**3**; -	**5**; -	**-**	**-**	**7**; -	**-**	**-**	**8**; 3	**-**	**4**; -	**-**	**-**	**5**; -
		Erum2440	8460	0.1	0	**3**; -	**4**; -	**-**	**-**	**16**; -	**4**; -	**-**	**6**; -	**-**	**-**	**-**	**-**	**2**; -
		Erum3221	8347	10	0	**3**; -	**-**	**-**	**-**	**16**; -	**-**	**-**	**4**; -	**4**; -	**12**; -	**-**	**-**	**-**

[a]Proteins that induced IFN-γ (ELISPOT) but did not induce other cytokines were, 1) for PBMC isolated from B8347: Erum1070 (72 SPMC), -3380 (50 SPMC), -3750_2 (95 SPMC), and -7400 (57 SPMC) 2) for PBMC from B8404: Erum1860 (93 SPMC) and 3) PBMC from B8460: Erum0590_2 (137 SPMC), -1070 (47 SPMC), -3050 (37 SPMC), -3380 (101 SPMC), -3750_2 (138 SPMC), -4010 (60 SPMC), -5550 (32 SPMC), -6560 (50 SPMC), -6830 (42 SPMC), -7100 (35 SPMC), -7190 (35 SPMC), -8390 (67 SPMC), -8760 (53 SPMC), -8770 (37 SPMC) and -8790 (30 SPMC).

[b]More than one number indicates IFN-γ production was measured in PBMC from all immune bovines indicated in column 4 stimulated with the amount of protein indicated in column 5. For samples in which PBMC from more than one immune bovine were positive, the bovine PBMC that gave the highest ELISPOT value was used in the qPCR, for example for Erum1050 we used bovine 8347 PBMC stimulated with 10 µg/ml protein.

[c](n/d) not done.

[d](-) value less than 1 fold.

Table 3. Characteristics of the recombinant proteins, and the ORFS encoding them, inducing cytokine profile A.

Erum ID	Predicted protein product	Length (aa)	MW[a] (kDa)	TMHMM & SignalP[b]	Phobius[b]	CELLO[c]	PSORTb[c]	Solubility	Yield (µg/ml)
Erum5270	superoxide dismutase [Fe]	210	24.2	–	–	extra cellular	unknown	Insoluble	231
Erum5400	unknown	173	19.8	–	1 th	outer membrane	unknown	Soluble	369
Erum8050	exported serine protease	476	51.3	signal	signal	outer membrane	periplasmic	Soluble	1015
Erum0660_1[d]	unknown	422[d]	48	–	–	outer membrane	unknown	Insoluble	75
Erum1150	unknown	179	19.5	–	–	cytoplasmic	cytoplasmic	Soluble	400
Erum7130	membrane protein	186	21.1	1 th	1 th	cytoplasmic	cytoplasmic	Soluble	534
Erum7140	membrane protein	197	21.9	1 th	1 th	outer membrane	cytoplasmic	Soluble	209
Erum7280	membrane protein	181	21.1	1 th	1 th	cytoplasmic	unknown	Insoluble	233
Erum7620	integral membrane protein	120	14.4	3 th	3 th	inner membrane	unknown	Insoluble	133
Erum8010	integral membrane protein	118	12.3	3 th	signal, 2 th	inner membrane	inner membrane	Soluble	252
Erum8460	unknown	56	6.4	–	–	cytoplasmic	cytoplasmic	Soluble	450

[a]Molecular weight (MW) was predicted using the program Protein Molecular Weight of the Sequence Manipulation Suite (http://www.bioinformatics.org/sms/prot_mw.html).

[b]The transmembrane helices (th) and signal sequences predicted by TMHMM2.0, SignalP3.0 and Phobius.

[c]Subcellular localisation predictions by CELLO and pSORTb2.0.

[d]1,266 bp fragment amplified from the 5' end of a large ORF (11,145 bp encoding 3,715 aa).

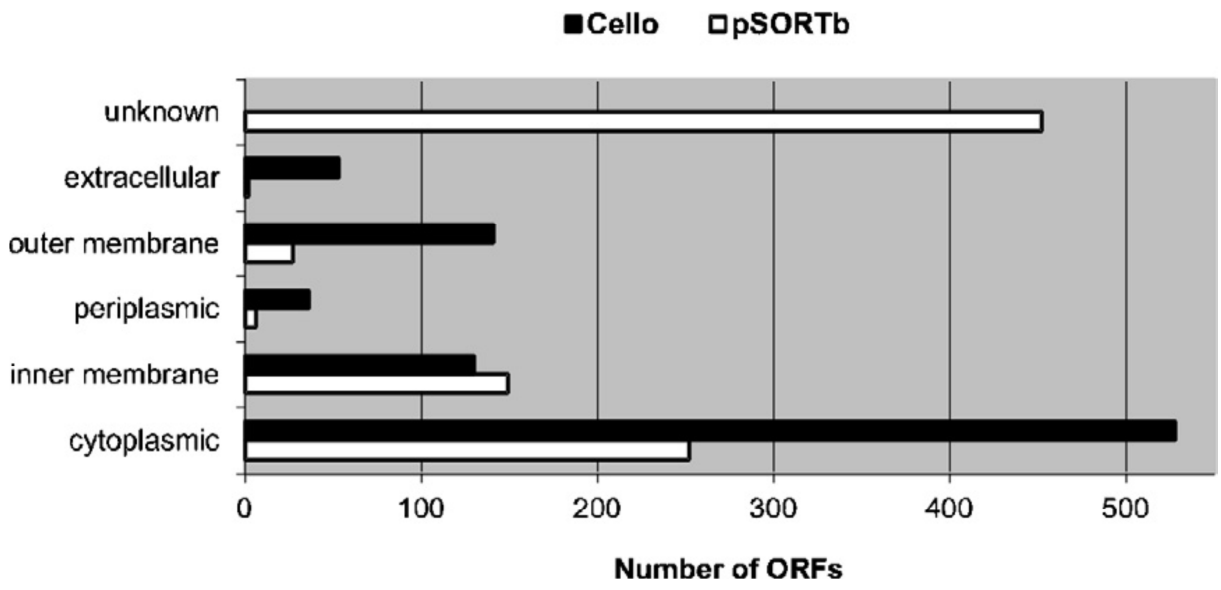

Figure 1. Predicted compartmentalisation of putative proteins by pSORTb and CELLO.

Supporting information

Table S1. The 272 selected ORFs grouped according to their putative function (th = transmembrane helices; C = cytoplasmic, P = periplasmic, IM = inner membrane, OM = outer membrane, E = extra cellular, U = unknown; TR = tandem repeat, TPR = tetratricopeptide repeat).

Erum ID	gene name	putative protein product	length (aa)	pI[a]	Mw[a] (kDa)	TMHMM/ SignalP	Phobius	CELLO	pSORTb	repeats[b]
Type IV secretion system										
0260	*virD4*	type IV secretion system protein VirD4	801	4.58	93.5	3 th	2 th	C	IM	TR
0270	*virB11*	type IV secretion system protein VirB11	332	6.62	36.9			C	U	
0280	*virB10*	type IV secretion system protein VirB10	448	5.65	48.7	1 th	1 th	P	U	TR
0290	*virB9*	type IV secretion system protein VirB9	267	8.88	31.0	signal	signal	E	U	
0300	*virB8*	type IV secretion system protein VirB8	232	9.60	26.9	1 th	1 th	OM	U	
4410		type IV secretion system protein	232	5.97	27.0	1 th	1 th	OM	OM	
5210		type IV secretion system protein	2455	4.35	267.7	4 th, signal	6 th, signal	OM	OM	TR
5260	*virB3*	type IV secretion system protein VirB3	97	9.37	11.2	2 th	2 th	IM	U	
7980		type IV secretion system protein	790	5.90	90.2			C	C	
ABC transporters										
0580		ABC transporter, ATP binding protein	239	8.66	26.4			C	U	
0860	*lolE*	lipoprotein releasing system transmembrane protein LolE	411	7.00	45.5	4 th	4 th	IM	IM	
1190	*lolD*	lipoprotein releasing system ATP-binding protein LolD	228	9.10	25.1		signal	C	IM	
3110	*uvrA*	uvrABC system protein A (ABC transporter domain, ATP/GTP-binding)	959	8.62	107.2			OM	OM	
5760	*pstB*	phosphate ABC transporter, ATP-binding protein	253	9.02	28.5			C	U	
6270		ABC transporter, ATP-binding protein	593	9.33	64.6	5 th	5 th	IM	IM	
6820		ABC transporter, ATP-binding and membrane-spanning protein	583	8.98	63.9	5 th	6 th	OM	IM	
7050	*ccmA*	heme exporter protein A	213	7.64	24.5			C	U	
Other transporters										
0190	*corC*	magnesium and cobalt efflux protein	288	4.96	32.6		1 th	C	C	
1170	*secG*	protein-export membrane protein SecG	110	5.39	12.0	2 th	1 th, signal	P/IM	U	
1590		secretion protein	514	9.37	57.8	1 th	2 th	OM	IM	
1780		Na+/H+ antiporter subunit	172	9.68	19.5	2 th	2 th	IM	IM	
2560	*tatA*	Sec-independent protein translocase membrane protein	56	5.04	6.2	1 th	1 th	C	U	
4600		magnesium transporter	456	5.6	51.2	4 th	5 th	IM	IM	
5430	*ffh*	signal recognition particle protein (GTPase)	450	9.40	49.6			C	C	

5530		Na+/H+ antiporter subunit	139	5.16	15.0	4 th	4 th	IM	IM	
5550		Na+/H+ antiporter subunit	99	7.8	10.8	3 th	3 th	IM	IM	
7780		preprotein translocase subunit YajC	122	9.07	13.0	1 th	1 th ,signal	C	U	
7800		outer membrane efflux protein	415	6.52	45.9	signal	signal	OM	OM	
8780	*secA*	preprotein translocase SecA subunit	870	6.46	99.9			C	C	
Chaperones										
0130	*dnaJ*	chaperone protein DnaJ	382	8.81	42.2			OM	C	
1990	*tig*	trigger factor	446	5.26	50.8			C	U	
3500	*ppiD*	peptidyl-prolyl cis-trans isomerase D	630	6.39	71.8	1 th	signal	OM	OM	
4010	*pmbA*	PmbA protein	455	5.78	49.5			OM	C	
4190	*hscA*	chaperone protein HscA	616	5.94	68.8			C	C	
5500	*dnaK*	chaperone protein DnaK	645	5.17	69.9			C	U	
7430	*secB*	protein-export protein SecB	174	4.53	19.4			C	U	
Protection responses and adaptation										
3050	*surE*	acid phosphatase SurE	252	6.16	27.6			E	U	
3350	*cutA*	periplasmic divalent cation tolerance protein CutA	109	5.02	12.8			C	C	
3480		peroxiredoxin	205	5.84	23.3			C	C	
5270	*sodB*	superoxide dismutase [Fe]	210	5.53	24.2			E	U	
Proteases/peptidases										
3510		glycoprotease	193	9.06	21.5			C	U	
4660	*clpA*	ATP-dependent Clp protease, ATP-binding subunit	764	6.13	85.8			C	C	
5610		carboxypeptidase	491	6.83	56.7			C	U	
6130		peptidase	289	8.45	32.3	1 th	1 th	OM	U	
7410		zinc protease	421	5.28	47.9			E	U	
8050		exported serine protease	476	8.35	51.3	signal	signal	OM	P	
8090		exported peptidase	438	5.08	49.8	signal	signal	C	P	
8100		exported M16 family peptidase	455	4.90	51.4	signal	signal	E/OM	U	
8120	*lpsA*	lipoprotein signal peptidase	149	9.10	17.2	3 th	4 th	IM	IM	
8220		exported D-alanyl-D-alanine carboxypeptidase	290	5.46	32.8	signal	signal	IM	U	
8250		membrane-associated zinc metalloprotease	379	9	42.2	4 th	4 th	IM	IM	
8430	*ftsH*	cell division protein FtsH (ATP-dependent zinc metallopeptidase)	611	7.24	67.6	2 th	1 th, signal	IM	IM	
Unknown with some miscellaneous information										
0050		unknown	101	6.72	11.5			C	U	
0320		unknown (TPR domain)	354	9.38	41.0			C	U	TPR
1300		unknown (zinc metallopeptidases, zinc-binding region signature)	1334	5.37	153.1			E	OM	
1840		unknown (patatin-like)	267	5.61	30.1			C	U	
2690		unknown	352	8.43	38.5			E	U	

2980		unknown	186	9.49	21.3			E	U	
3700	*typA*	GTP-binding protein TypA/BipA (GTPase)	612	5.42	67.4			C	C	
3730		unknown	153	3.92	17			C	U	
3750		unknown (Ankyrin repeats)	1674	5.09	179.8			OM	C	Ankyrin
3980		unknown (Ankyrin repeats)	3002	5.82	330.3			OM	E	Ankyrin, TR
4950		competence protein	492	9.19	54.1			C	C	
5120		secretion protein	363	6.02	40.68	1 th	1 th	OM	U	
5400		unknown	173	8.92	19.8		1 th	OM	U	
5420	*era*	GTP-binding protein ERA (GTPase)	296	9.08	33.7			C	U	
5620		unknown (OmpA or pal)	217	6.08	24.4		signal	P	OM	
6220		unknown (Ankyrin repeats)	125	5.76	14.2			C	U	Ankyrin
6540		zinc metallopeptidase	433	6.03	49.8	2 th	signal	IM	IM	
6670		haloacid dehalogenase-like hydrolase	210	5.7	23.5			E	U	
6970		unknown	95	8.87	11.0			C	C	
7850		unknown	209	8.41	23.4			C	U	
7960		unknown	1304	7.70	150.1		signal	OM	OM	
8150		methyltransferase	280	9.19	32.3			C	U	
8240		conserved hypothetical protein (TPR repeat region)	93	6.28	11.1			C	C	TPR
Membrane and surface-associated										
0090		membrane protein	193	9.66	23.0	1 th, signal	2 th	P	U	
0330		integral membrane protein	159	8.55	18.7	2 th	2 th	C	C	
0590		integral membrane protein	613	8.73	70.7	3 th	3 th	OM	U	
0700		integral membrane protein	547	9.11	61.4	2 th	2 th, signal	OM	U	
0831		integral membrane protein	84	4.72	9.6	2 th	2 th	IM	U	
0840		integral membrane protein	413	8.62	49.2	2 th	1 th, signal	C	C	
0850		membrane protein (lipoprotein lipid attachment site)	258	6.96	30.1	1 th	1 th	C	C	
0970		integral membrane protein	155	9.00	18.3	4 th	4 th	IM	IM	
0990		integral membrane protein	607	6.15	68.6	2 th	2 th	OM	OM	
1040		integral membrane protein	1165	4.63	128.8	2 th	2 th	E	C	TR
1050		integral membrane protein	454	6.54	51.6	2 th	2 th	OM	U	
1440		membrane protein	482	6.28	55.0	1 th	1 th	C	IM	
1450		membrane protein	208	8.94	23.7	1 th	1 th	OM	U	
1620		integral membrane protein	197	8.47	22.1	4 th	4 th	IM	IM	
1750		integral membrane protein	142	9.15	16.1	4 th	4 th	IM	IM	
1790		membrane protein	205	8.28	23.4	1 th	1 th	C	U	
1860		membrane protein	270	8.47	31.3	1 th	signal	C	U	
2070		integral membrane protein	431	6.12	48.1	4 th	2 th, signal	IM	IM	

2080	integral membrane protein	94	9.45	11.2	2 th	3 th	IM	U	
2180	integral membrane protein (Ankyrin repeats)	876	6.24	98.8	2 th	2 th	OM	C	Ankyrin
2240	membrane protein	369	9.17	42.5	1 th	1 th	C	U	
2250	membrane protein	347	8.45	39.6	1 th	1 th	IM	U	
2260	membrane protein	313	7.70	35.9	1 th	1 th	E	U	
2270	membrane protein	384	9.54	44.5	1 th	1 th	E	U	
2280	membrane protein	341	9.56	39.3	1 th	1 th	C	U	
2290	membrane protein	342	9.14	39.6	1 th	1 th	E/C	U	
2300	membrane protein	370	8.99	43.6	1 th	1 th	C	U	
2330	membrane protein	306	8.62	34.9	1 th	1 th	IM	U	
2340	membrane protein	326	8.22	37.4	1 th	1 th	C	U	
2400	membrane protein	391	5.32	44.8	1 th	1 th	E	U	TR
2410	membrane protein	326	5.56	37.1	1 th	1 th	C	U	
2440	integral membrane protein	220	5.09	25.0	2 th	2 th	IM	U	
2470	integral membrane protein	358	8.22	41.6	2 th	4 th	C	IM	
2750	membrane protein	527	6.57	60.8	1 th, signal	2 th	E	IM	
2760	membrane protein	519	9.09	60.2	1 th, signal	2 th	OM	U	
2770	membrane protein	526	8.6	60.4	1 th, signal	2 th	C	U	
2780	membrane protein	524	7.06	59.9	1 th, signal	2 th	C	U	
2790	integral membrane protein	653	5.82	75.3	2 th	2 th	C	OM	
2800	membrane protein	520	6.4	60.9	1 th, signal	2 th	C	U	TR
2900	integral membrane protein	331	5.96	37.5	2 th	2 th	C	C	
3240	integral membrane protein	210	9.84	23.2	4 th	4 th	IM	IM	
3570	integral membrane protein	376	5.45	40.9	2 th	2 th	E	U	TR
3580	integral membrane protein	188	6.83	21.3	2 th	signal, 1 th	OM	U	
3590	integral membrane protein	389	4.43	41.3	2 th	2 th	E	U	TR
3600	integral membrane protein	585	5.87	67.2	2 th	2 th	OM	IM	TR
3610	membrane protein	513	5.94	59.5	1 th, signal	2 th	C	U	
3620	integral membrane protein	537	8.92	60.6	2 th	2 th	OM	U	
3630	membrane protein	519	8.67	59.1	1 th, signal	signal, 1 th	OM	U	
3860	membrane protein	171	8.19	20.1	1 th	1 th	E	U	
4070	integral membrane protein	193	9.7	22.7	3 th	3 th	IM	IM	
4210	membrane protein	356	8.67	40.4	1 th	signal	OM	U	
4230	integral membrane protein	135	9.37	15.7	2 th	signal	C	U	
4440	integral membrane protein	195	8.78	22.4	4 th	4 th	C	IM	
4610	membrane protein	124	6.49	14.4	1 th	1 th	C	C	
4620	membrane protein	134	5.27	15.1	1 th	1 th	IM	U	

4630		membrane protein	125	8.43	14.2	1 th	1 th	P	U	
4640		membrane protein	123	7.61	13.8	1 th	1 th	P	U	
4960		integral membrane protein	129	8.54	13.7	2 th	2 th	P/IM	U	
5310		integral membrane protein	1392	7.83	162.9	2 th	3 th	C	C	
5470		membrane protein	158	8.95	18.3	1 th	1 th	C	U	
5480		membrane protein	111	4.64	11.8	1 th	signal	C	IM	
5520		integral membrane protein	111	9.87	12.3	3 th	3 th	IM	IM	
5560		integral membrane protein	88	5.45	9.7	3 th	2 th, signal	IM	IM	
5700		membrane protein	142	8.87	16.9	1 th	1 th	C	U	
6210		integral membrane protein	285	5.84	31.9	3 th	3 th	IM	IM	
6240		membrane protein	81	4.74	9.1	1 th	signal	P	U	
6300		integral membrane protein	352	8.61	40.3	2 th	2 th	C	U	
6680		integral membrane protein	170	8.68	19.7	3 th	3 th	IM	IM	
6880		integral membrane protein	203	4.97	23.1	3 th	3 th	IM	IM	
7090		membrane protein	228	5.37	26.0	1 th	1 th	C	C	
7100		membrane protein	250	6.49	28.3	1 th	1 th	C	U	
7130		membrane protein	186	5.57	21.1	1 th	1 th	C	C	
7140		membrane protein	197	4.38	21.9	1 th	1 th	OM	C	
7250		membrane protein	999	6.14	113.1	1 th	1 th	OM	OM	
7270		membrane protein	198	8.54	22.6	1 th	1 th	P	U	
7280		membrane protein	181	4.55	21.1	1 th	1 th	C	U	
7300		integral membrane protein	157	4.24	16.4	2 th	signal, 1 th	E	U	
7310		integral membrane protein	202	4.20	22.1	2 th	signal, 1 th	E	U	
7320		integral membrane protein	266	3.97	28.1	2 th	1 th	OM	U	
7330		membrane protein	291	4.30	31.8	1 th	signal, 1 th	OM	U	
7340		membrane protein	122	5.35	13.3	1 th	1 th	C	U	
7350		membrane protein	145	4.90	16.2	1 th	1 th	E	U	
7360		membrane protein	147	4.62	16.1	1 th	1 th	IM	U	
7370		integral membrane protein	169	4.26	18.5	2 th	2 th	E	U	
7380		membrane protein	157	4.16	17.2	1 th	1 th	C	C	
7600		membrane protein	425	5.54	48.5	1 th	signal	OM	U	TR
7620		integral membrane protein	120	9.4	14.4	3 th	3 th	IM	U	
7950		ATP/GTP-binding membrane protein	735	7.02	85.4	1 th	1 th	OM	U	
8000		integral membrane protein	112	8.74	11.9	3 th	signal, 2 th	IM	IM	
8010		integral membrane protein	118	9.18	12.3	3 th	signal, 2 th	IM	IM	
8020		integral membrane protein	124	8.8	13.2	3 th	signal, 2 th	IM	IM	
8040	*hflC*	HflC membrane protein	290	9.44	33.1	1 th	1 th	C	U	

8110	integral membrane protein	224	8.03	24.6	4 th	4 th	IM	IM	
8260	outer membrane protein	771	8.03	87.3	signal	signal	OM	OM	
8270	outer membrane protein	182	6.01	21.0	signal	signal	C	U	
8320	Surf1-like protein	213	8.9	25.1	2 th	2 th	IM	IM	
8390	membrane protein	163	6.53	18.3	1 th	1 th	C	C	
8510	membrane protein	223	6.29	25.2	1 th		C	U	
Exported proteins									
0470	exported protein	208	8.65	24.6	signal	signal	C	C	
1070	exported protein	180	8.99	20.9	signal	1 th	C	U	
1210	exported protein	877	5.71	97.1	signal	signal	OM	OM	
1460	exported protein	180	5.36	20.8	signal	signal	C	C	
1960	exported protein	383	6.14	45.4	signal	signal	OM/C	U	
2310	exported protein	317	6.56	36.0	signal	1 th	C	C	
2320	exported protein	307	8.69	35.0	signal	1 th	C	U	
3450	exported protein	327	6.1	36.7	signal	signal	C	U	
3780	exported protein	223	8.73	25.4	signal	signal	OM	IM	
3790	exported protein	235	6.29	27.2	signal	signal	OM	OM	
4470	exported protein	385	6.10	41.8	signal	signal	OM	OM	
4740	exported protein	639	7.61	71.9	signal	signal	E	OM	TR
5000	exported protein	490	8.36	53.6	signal	signal	OM	OM	
5010	exported protein	564	8.50	63.0	signal	signal	OM	OM	TR
5140	exported protein	389	5.82	43.4	signal	signal	OM	U	
5370	exported protein	325	5.92	37.0	signal	signal	C	U	
6230	exported protein	134	4.16	15.6	signal	signal	P	U	
6650	exported protein	200	9.62	22.0	signal	signal	IM	P	
7110	exported protein	182	6.88	21.5	signal	1 th	C	U	
7120	exported protein	204	5.33	23.3	signal	1 th	C	C	
7760	exported lipoprotein	250	6.15	29.3	signal	signal	C	U	
7970	exported protein	1710	6.38	196.7	signal	signal	OM	OM	
8060	exported protein	204	5.92	23.5	signal	signal	OM	U	
8610	exported protein	236	9.70	27.8	signal	signal	OM	U	
Unknown									
0250	unknown	457	4.89	50.5			C	C	TR
0350	unknown	143	9.38	16.5			C	U	
0500	unknown	102	8.75	11.8			C	U	
0660	unknown	3715	5.35	420.2			OM	U	TR
0680	unknown	540	5.53	60.7			OM	U	

0690	unknown	470	8.07	54.8	1 th	OM	U	
0710	unknown	123	4.55	14.0	signal	C	U	
0720	unknown	931	7.73	105.6	1 th, signal	OM	U	
0730	unknown	93	5.41	10.6		C	U	
1110	unknown	661	3.97	63.5		E	C	TR
1150	unknown	179	3.54	19.5		C	C	
1430	unknown	951	9.13	106.7	1 th	OM	U	TR
1600	unknown	204	6.08	23.9		C	U	
1770	unknown	1529	6.53	170.8		OM	OM	
1851	unknown	92	7.07	10.8		C	C	
1900	unknown	417	9.34	44.9		OM	U	
2170	unknown	1073	5.89	121.8		E	OM	TR
2370	unknown	417	6.56	46.9		E	U	
2380	unknown	332	6.11	37.6	signal	OM	U	
2630	unknown	1202	6.54	139.6		OM/E	OM	TR
2730	unknown	912	5.97	106.1		OM	C	
3221	unknown	93	8.59	10.8		C	U	
3290	unknown	194	8.43	22.3		E	U	
3380	unknown	94	9.7	11.4		C	U	
3410	unknown	119	6.72	14.2		C	U	
3640	unknown	111	9.72	12.4		C	U	
3701	unknown	106	8.98	12.7		C	U	
3890	unknown	126	10.64	14.7		P	U	
3900	unknown	189	9.52	21.9		C	U	
3910	unknown	129	7.77	15.2		C	U	
3920	unknown	136	9.04	15.7		C	U	
3930	unknown	188	6.16	22.0		C	C	
3940	unknown	115	5.3	13.6		C	C	
4261	unknown	84	9.81	9.3		P	U	
4320	unknown	425	6.03	48.9		C	C	
4340	unknown	392	5.79	44.8		OM	U	
4350	unknown	409	5.44	47.1		OM	U	
4360	unknown	157	9.68	17.5		OM	U	
4390	unknown	240	8.5	28.1		C	C	
4400	unknown	994	5.62	112.7		C	U	
4450	unknown	280	8.3	31.4		OM	U	
4530	unknown	199	8.97	23.0		E	U	TR

4650		unknown	771	5.47	89.8	C	C	
4930		unknown	186	4.70	19.7	E	U	
5300		unknown	464	5.50	52.9	OM	U	
5450		unknown	264	8.12	29.5	OM	U	
5460		unknown	258	8.52	29.1	C	U	
5570		unknown	552	9.32	61.9	OM	U	TR
5580		unknown	344	5.96	38.3	E	U	
5590		unknown	213	8.71	23.9	C	C	
6150		unknown	97	4.57	11.1	C	C	
6160		unknown	121	7.07	13.9	C	C	
6200		unknown	101	5.02	11.9	C	U	
6320		unknown	105	5.88	11.9	C	U	
6560		unknown	295	6.18	33.7	C	U	
6570		unknown	212	4.74	23.8	C	U	
6830		unknown	109	4.53	12.3	C	U	
7060		unknown	546	5.81	62.7	C	C	
7190		unknown	281	9.06	32.3	OM	C	
7200		unknown	360	6.2	40.7	C	C	
7400		unknown	467	8.61	52.3	C	U	
7650		unknown	468	9.82	54.8	E	U	
7790		unknown	234	5.10	26.4	OM	U	
8170		unknown	372	5.46	42.7	C	IM	
8180		unknown	150	8.34	16.3	OM	U	
8340		unknown	622	8.64	70.6	OM	U	
8460		unknown	56	4.82	6.4	C	C	
8760		unknown	111	4.09	12.4	C	U	
8770		unknown	177	4.05	19.8	C	C	
8790		unknown	143	4.69	16.3	C	U	
Other (immunogenic/protective against other organisms)								
0010	*gapB*	NAD(P)-dependent glyceraldehyde 3-phosphate dehydrogenase[c,d]	335	7.15	37.3	C	C	
1710	*rplL*	50S ribosomal protein L7/L12[e,f]	131	5.09	14.3	C	C	
4840	*eno*	Enolase[d,g,h]	421	5.70	46.6	C	C	
4860	*mraW*	S-adenosyl-methyltransferase MraW (methyltransferase)[i]	301	9.37	33.7	OM	C	
5160	*engA*	GTP binding protein EngA (GTPase)[i]	439	9.26	50.2	C	U	

[a]The theoretical isoelectric point (pI) and molecular weight (Mw) as calculated with the Compute pI/Mw tool on the ExPASy Proteomics Server (http://us.expasy.org/tools/pi_tool.html)

[b]Repeat sequences identified as described previously (Collins *et al.*, 2005)

[c]Rosinha, G.M., Myioshi, A., Azevedo, V., Splitter, G.A., Oliveira, S.C., 2002. Molecular and immunological characterisation of recombinant *Brucella abortus* glyceraldehyde-3-phosphate-dehydrogenase, a T- and B-cell reactive protein that induces partial protection when co-administered with an interleukin-12-expressing plasmid in a DNA vaccine formulation. J. Med. Microbiol. 51, 661-671.

[d]Pitarchm A., Pardo, M., Jiménez, A., Pla, J., Gil, C., Sánchez, M., Nombela, C., 1999. Two-dimensional gel electrophoresis as analytical tool for identifying *Candida albicans* immunogenic proteins. Electrophoresis 20, 1001-1010.

[e]Luo, D., Ni, B., Li, P., Shi, W., Zhang, S., Han, Y., Mao, L., He, Y., Wu, Y., Wang, X., 2006. Protective immunity elicited by a divalent DNA vaccine encoding both the L7/L12 and Omp16 genes of *Brucella abortus* in BALB/c mice. Infect. Immun. 74, 2734-2741.

[f]Pontes, D.S., Dorella, F.A., Ribeiro, L.A., Miyoshi, A., Le Loir, Y., Gruss, A., Oliveira, S.C., Langella, P., Azevedo, V., 2003. Induction of partial protection in mice after oral administration of *Lactococcus lactis* producing *Brucella abortus* L7/L12 antigen. J. Drug Target 11, 489-493.

[g]Finco, O., Bonci, A., Agnusdei, M., Scarselli, M., Petracca, R., Norais, N., Ferrari, G., Garaguso, I., Donati, M., Sambri, V., Cevenini, R., Ratti, G., Grandi, G., 2005. Identification of new potential vaccine candidates against *Chlamydia pneumoniae* by multiple screenings. Vaccine 23, 1178-1188.

[h]Montigiani, S., Falugi, F., Scarselli, M., Finco, O., Petracca, R., Galli, G., Mariani, M., Manetti, R., Agnusdei, M., Cevenini, R., Donati, M., Nogarotto, R., Norais, N., Garaguso, I., Nuti, S., Saletti, G., Rosa, D., Ratti, G., Grandi, G., 2002. Genomic approach for analysis of surface proteins in *Chlamydia pneumoniae*. Infect. Immun. 70, 368-379.

[i]Garbom, S., Forsberg, A., Wolf-Watz, H., Kihlberg, B.M. 2004. Identification of novel virulence-associated genes via genome analysis of hypothetical genes. Infect Immun. 72(3):1333-40.