Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vaccine 23 (2005) 3016–3025
Vaccine
www.elsevier.com/locate/vaccine

Screening the whole genome of a pathogen in vivo for individual protective antigens

Katherine Stemke-Hale[a], Bernhard Kaltenboeck[b], Fred J. DeGraves[b], Kathryn F. Sykes[a], Jin Huang[b], Chun-hui Bu[a], Stephen Albert Johnston[a,*]

[a] *Departments of Medicine and Microbiology, Center for Biomedical Inventions, University of Texas-Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-8573, USA*
[b] *Department of Pathobiology, College of Veterinary Medicine, 270 Greene Hall, Auburn University, Auburn, AL 36849-5519, USA*

Received 12 May 2004; received in revised form 1 December 2004; accepted 2 December 2004
Available online 13 January 2005

Abstract

We report the results of a general protocol that was used to screen the whole genome of *Chlamydophila abortus*, *type strain B*577 (*formerly Chlamydia psittaci strain B*577), in a mouse pneumonia model. Genetic immunization was used to functionally test the genes of *C. abortus* as vaccines in a mouse challenge system. Nine gene fragments were isolated that conferred protection, with five protecting as effectively as the live-vaccine positive control. Bioinformatics approaches were unable to reconstruct isolation of these antigens. These results suggest that pathogen genomes can be functionally screened for vaccine candidate antigens in a mouse model to reveal new classes of vaccine candidate antigens that may have therapeutic efficacy across host species, disease manifestations, and delivery platforms.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Chlamydia; Expression library immunization; Vaccine

1. Introduction

Subunit vaccines may offer many advantages over conventional live or killed formulations. Notably they may be significantly safer. In addition, by eliminating pathogen factors attenuating the host immune response it may be possible to create more effective vaccines [1]. The challenge is to find which factors in the pathogen should be included in the vaccine, particularly with bacterial pathogens carrying 5000 or more encoded proteins. There have been many approaches to this problem ranging from bioinformatic predictions of good antigens to immunological selections based on host responses [2–5]. These systems for selecting antigens either assume knowledge of the basis of the protective immune response or have biases in the sorting process. We have been developing technologies that allow us, in as unbiased a way as possible, to conduct complete screens of all of a pathogen's encoded proteins. We report here the application of these techniques to functionally screen the genome of *Chlamydophila abortus* [6] for protective antigens.

In 1995 we introduced the concept of expression library immunization (ELI) [7]. This approach to antigen discovery was based on our earlier demonstration of genetic immunization [8]. Genetic immunization simplified the immunization process by using plasmids encoding foreign proteins as the inoculum. ELI was based on the idea that pools of genes from a given pathogen could be introduced into the test animal by genetic immunization and then screened for whether they contained one or more protective antigens by directly challenging the host with the pathogen. We demonstrated that the *Mycoplasma pulmonis* genome could be fractionated into pools of ∼3000 plasmids and some pools conferred high levels of protection. Other groups have also now shown that pools representing a portion of different pathogen genomes

* Corresponding author. Tel.: +1 214 648 1415; fax: +1 214 648 1298.
E-mail address: stephen.johnston@utsouthwestern.edu (S.A. Johnston).

0264-410X/$ – see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.vaccine.2004.12.013

can confer protection [9–17]. The critical issue is whether protective pools can be reduced to individual genes and if so, what types of antigens are protective. Pools are not of practical value for most vaccine applications so this is a crucial question relative to the application of ELI.

We chose *C. abortus* to perform a complete ELI screen. Intracellular bacteria of the family *Chlamydiaceae* are important pathogens in both humans and animals. In humans, *Chlamydiaceae* bacteria are responsible for a wide variety of diseases, including trachoma-induced blindness, sexually transmitted disease, community-acquired pneumonia, and possibly even coronary heart disease [18–23]. In cattle, ubiquitous *C. abortus* infections cause infertility and abortion, resulting in a loss of calf and milk production [24,25]. Conventional approaches, such as biochemical purification of abundant or immune-reactive proteins have been used to search for vaccines against *Chlamydiaceae*-induced disease. These approaches have led to candidates with less than optimal responses, including inconsistent protection in animal models [26,27] and exacerbation of disease in trachoma field trials [27]. A leading candidate for a vaccine, the major outer membrane protein (MOMP), has been tested as a genetic vaccine in several host species with variable and inconclusive results [28–31].

We have now screened a library of DNA fragments representing most if not all of the *C. abortus* genome for their ability to protect in the mouse pneumonia model. Five gene fragments conferred protection at levels as good as or better than the live-vaccine control. Four of these five fragments were from gene categories not generally predicted to be good sources of vaccine candidates. These results demonstrate that a whole bacterial genome can be screened for single vaccine candidates by genetic immunization and that the screen produces novel candidates.

2. Materials and methods

2.1. Construction of the library

Genomic DNA of the ruminant *C. abortus* type strain B577 was physically sheared using a nebulizer (Glas Col, Terra Haute, IN), then size fractioned on a 1.5% TBE agarose gel. Fragments between 300 and 700 bp were excised and electroeluted. Adaptors (top strand 5′: GATCTGGATCCCGAT, bottom strand 5′: ATCGGGCTCCA) were ligated onto the fragments, then the fragments were cloned into the *Bgl*II site of pCMVi-UB and transformed into DH5α. The resulting library consisted of 27 individual pools of 2400–3400 plasmids, a total of approximately 82,000 clones, with 67% containing inserts. Orientation and frame constraints predict that only 1/6 of the plasmids would encode authentic *C. abortus* protein fragments. Given the approximately 1Mb-genome size of *C. abortus*, this library represents more than 6 genome expression equivalents, defined as the redundancy of authentic ORFs ligated in-frame with ubiquitin. To gauge the quality of the library, 300 clones were sequenced, of which 48 (∼1/6) were found to contain sequences that encoded in-frame proteins, as predicted. Of these 48 inserts, 3 were non-*C. abortus* DNA and matched mammalian DNA, which indicates a small contamination of genomic DNA from the host BGMK cells. We also searched the library to make sure that it contained all regions of MOMP (ompA). Further details on the library can be supplied on request.

2.2. Mouse vaccination and pathogen challenge

The DNA pools were tested for protection against *C. abortus* using a well-established *intranasal* mouse challenge model [32,33]. DNA from the pools was injected into 6-week old female NIH-Swiss mice in Round 1 and BALB/C mice for all subsequent rounds. Doses of 50 μg library DNA were delivered intramuscularly to the quadriceps and tibialis anterior muscles. Doses of 2.5 μg DNA were delivered to the ear skin of mice with a gene gun. *C. abortus* B577 was grown in BGMK cells and titrated for IFU in BGMK shell vial coverslip cultures by enumeration of chlamydial inclusions stained with FITC-labeled monoclonal antibody against chlamydial LPS as described [32,33]. For rounds 1 and 2 of ELI, mice were boosted 9 weeks after the prime inoculation in the same manner, and for rounds 3 and 4 of ELI the mice were given an additional boost 5 weeks after the prime. In all cases, mice were challenged at 13 weeks with a dose of 3×10^6 inclusion forming units (IFU) of *C. abortus* administered intranasally. The positive control group representing protection received a low dose intranasal inoculation of 3×10^4 IFU of the same live strain four weeks prior to the high-dose challenge. At the time of the high-dose challenge, *C. abortus* DNA from the low-dose live vaccine inoculation was not detectable in the lungs. Naïve, unchallenged animals served as unprotected controls for disease. Animals were sacrificed 12 days after receiving the challenge, and lungs were weighed. Lungs of mice from round 4 were frozen and stored at −80 °C. An irrelevant library pool, composed of *M. tuberculosis*, was used as a negative control for the first experiment (round 1), and non-protective *pools* were used in subsequent rounds.

2.3. Calculation of relative protection

The relative protection score was calculated by assigning the score of 1 to the average lung weight of the vaccinated positive control and 0 to the average lung weights of the challenged, unvaccinated control. These points define a line; protected animals with lower lung weight have a higher relative protection score. Animals with worse disease than challenged unvaccinated controls will have a negative relative protection score.

2.4. Quantitative PCR

Genomes of *C. abortus* B577 in lung tissue were detected by a fluorescence resonance energy transfer, real time quantitative PCR method as described [34].

3. Results

3.1. Library construction

To create the library of genetic immunization plasmids, genomic DNA of *C. abortus* strain B577 was physically sheared and cloned into the genetic immunization vector pCMVi-UB, which drives transcription using the strong mammalian CMV promoter [35,36]. The *C. abortus* sequences are fused to the murine ubiquitin gene which can enhance cellular immune responses, presumably by increasing proteasome targeting for MHC I loading, while still eliciting strong humoral responses [36]. This vector seemed appropriate for screening for a *Chlamydiaceae* vaccine since MHC I-restricted immunity appears to mediate protection in the early phase of chlamydial infection, and MHC II-restricted CD4 T-helper (Th1) immunity is indispensable for clearance of *Chlamydiaceae* [27,37,38].

3.2. Vaccination and challenge

The library was divided into 27 pools, each containing approximately 3000 plasmids, and injected into mice. This library represented approximately six genome equivalents. In the first round of screening, groups of four outbred, 6-week old, NIH-Swiss mice were immunized both intramuscularly (i.m.) by needle injection and epidermally with a gene gun. We used both delivery methods because some have argued that i.m. injections favor the development of Th1 cell responses whereas gun-delivered skin inoculations favor Th2 responses[39]. The mice were boosted once (for rounds 1 and 2) or twice (for rounds 3 and 4), then challenged with a dose of 3×10^6 inclusion forming units (IFU) of *C. abortus* administered intranasally. High levels of protection against chlamydial infection have been achieved through prior infection with low doses of live pathogen delivered mucosally [40]. Therefore the positive control group, which represents complete protection, received an intranasally administered vaccine consisting of a low dose (3×10^4 IFU) of live challenge material 4 weeks prior to experimental challenge. Naïve unchallenged animals served as unprotected controls for maximum disease. The animals were sacrificed 12 days after receiving the challenge infection, since disease intensity on day 12 after infection is well correlated with the chronic infection outcome.

An important aspect of the ELI approach is that the direct readout for protection is disease, not immune correlates. This is particularly relevant to developing a *Chlamydiaceae* vaccine since neither the mechanisms of *Chlamydiaceae*-induced disease nor the requirements for a protective immune response are well understood [26,27]. We scored directly for disease, and its prevention, by using the quantifiable disease-dependent parameter of lung weight [32,33]. Maximum disease resulted in a 2.5-fold average increase in lung weight on day 12 after challenge compared to naïve unchallenged animals. As an indirect measure of disease, the chlamydial lung burdens of the mice were also determined. Lung weights are very consistent among inbred mice within the age group used for this study, and any change in lung weight can be accurately quantified. In our hands, lung weight increase was a more accurate measure of protection than lung IFUs which showed higher inter- and intra-group variance, although the IFUs correlated with lung weight, such that live vaccinated mice with maximum protection (low lung weights) typically carried about 10^3–10^4 IFU. The diseased animals (high lung weights) carried about 10^4–10^6 IFU. Determination of chlamydial genomes per lung by quantitative PCR showed the same trend as IFU.

Mice in the positive control, live-vaccine group were completely protected from disease and had lung-weight increases of 10–30% compared to naïve, unchallenged controls. Lungs of the protected, positive control mice did not show any gross lung lesions, and examination revealed no interstitial infiltrates but did reveal prominent peribronchiolar lymphocytic cuffs, probably a sign of protective immune stimulation (data not shown). The lung weights were transformed to relative protection scores in a linear equation (see Section 2).The system used to deconvolute the sub-libraries into individual gene candidates is depicted in Fig. 1. In round 1, each sub-library consists of approximately 3000 different clones. Fourteen of the 27 plasmid sub-library pools inoculated into mice conferred protection as compared to non-immunized mice (Fig. 2). The approximately 40,000 colonies comprising these 14 libraries were individually picked and organized into a 2D array to derive the sub-library DNA inocula for the second round of screening. The colonies from each row (24) and column (24) were pooled and plasmid DNA was prepared from them. Each row and column represented a sub-library of 1700 colonies and each matrix intersection contained 96 colonies (Fig. 1).

3.3. Reduction to single clones

In round 2 each row and column DNA was used to vaccinate a group of five mice that were challenged as described for round 1. Seven rows and eight columns were scored as protection-positive sub-libraries (Fig. 1, and data not shown). The designated 41 intersections contained a total of 3936 colonies. These colonies were re-arrayed in a second 2D grid, with 24 rows and 36 columns containing four to five colonies at each intersection (Fig. 1, round 3). Again, DNA was prepared from the colonies comprising each row and column. After inoculation of the DNA and subsequent challenge with the pathogen, three rows and three columns were scored as positive. Data on the libraries for round 3 are shown in Fig. 2.

Using the third round data, plasmids at each intersection of protective rows and columns were sequenced (46 plasmids from nine intersections) in order to eliminate those not likely to encode protective antigens. Only those ubiquitin-fused gene fragments that encoded peptides longer than 50 amino acids were tested as single clones in the round 4 chal-

lenge. Fourteen plasmids met this criterion. Additional test groups included a pool of all 14 plasmids, a pool of the 32 plasmids encoding proteins fewer than 50 amino acids, and the positive and negative controls for protection. The results are presented in Fig. 3.

3.4. Analysis of protection level

Of the 14 individually tested clones, CP #1 through CP #9 had positive relative protection scores (Fig. 3). Three of the clones (CP #1–3) elicited protection that was statistically higher than the unvaccinated control, which has high variance, and five clones were significantly more protective than the genes encoding fewer than 50 amino acids (CP #1–5 and 7, *p*-value of less than 0.05 when comparing lung weights). CP #1 was more protective than the live-vaccine, positive control. Since some of the intersections of round 3 contained more than one plasmid it was expected that some plasmids would not confer protection, as was the case for CP #10 and 11. Interestingly, CP #12–14 appear to condition increased disease severity. The pool of all 14 plasmids, including the negative effectors, conferred protection that was better than the positive control but not better than CP #1. The pool of plasmids encoding proteins fewer than 50 amino acids conferred no significant protection (Fig. 3). Chlamydial loads were also determined for each group (Fig. 3). While the chlamydial loads generally tracked with protection and the most protective genes were significantly lower than in unvaccinated controls ($p < 0.05$ in the Mann–Whitney *U*-test for genes CP #1, 2, 4–7, 9, 10), the organism load, unlike lung weight, did not allow for clear distinction among these 10 genes. The sequences of these 14 *C. abortus* gene fragments are available through our website (http://cbi.swmed.edu/).

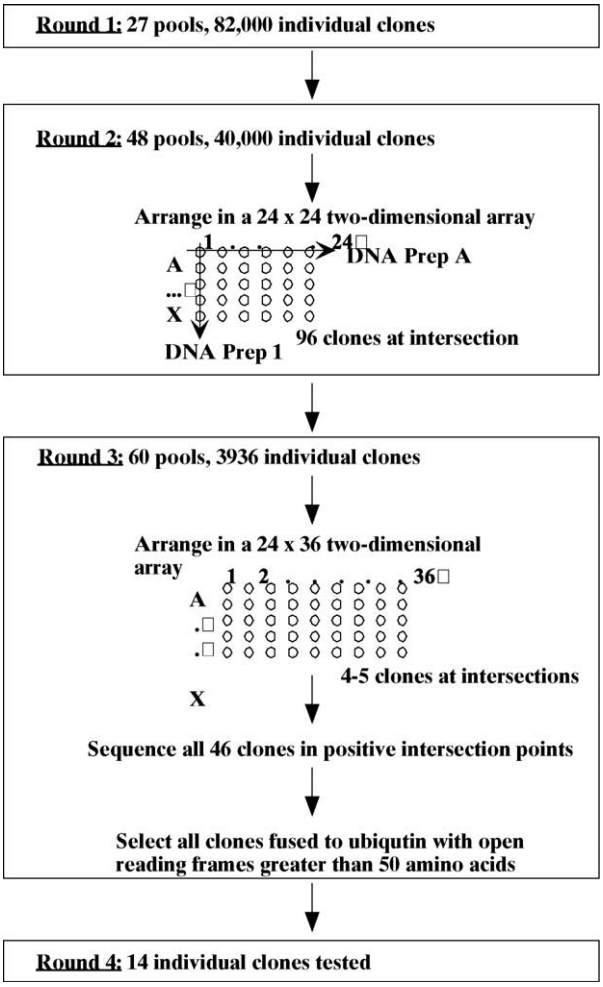

Fig. 1. Flowchart of the library deconvolution. Each round consisted of preparation of DNA samples, vaccination of mice, challenge and determination of the relative protection in each group.

3.5. Validation of the ELI approach

Having individual protective genes allowed us to test some specific aspects of the deconvolution process. First, we asked whether individual genes could confer the protection we observed in the libraries. Protective clone CP #4 was diluted 1/2000 in a non-protective *sublibrary pool of clones*. As observed in Fig. 4, this CP #4-*spiked sublibrary* conferred protection. Another informative observation was that clone CP #11 was the only ORF (of a length long enough to suggest it is expressed) present in a protective round 3 pool, yet it did not confer protection when delivered separately (Fig. 3). However, when the complex environment was reconstituted by spiking CP #11 into a non-protective pool, with an independent mixture of clones, this pool was now able to confer protection (Fig. 4). These results indicate that components of the ELI libraries can immunologically interact. At least in this case the interaction was non-specific, presumably T cell help, since two independent pools both transformed CP #11 into a protective clone.

3.6. Informatic comparison of genes selected

The sequences of the 14 selected *C. abortus* fragments were compared to genomic databases and to *C. pneumoniae* [41]. The results are depicted in Fig. 5. Three comparisons are noteworthy. First, six of the nine genes conferring protection are not outer membrane proteins (clone CP #9 is a hypothetical protein and not likely to be an outer membrane protein). It is generally believed that membrane bound or secreted proteins would produce the best protective antigens against chlamydial infections [42]. However, the most protective clone, CP #1, is a fragment of a DNA polymerase III subunit. Interestingly, the second and third most protective clones (CP #2 and 3) were independent derivatives from two of the three subunits of an amido-transferase.

Second, secreted or outer membrane proteins are not necessarily protective. While one clone, CP #4, is a fragment of a previously sequenced *C. abortus* outer membrane protein OMP90A [42,43] that does confer protection, three other putative outer membrane protein fragments (CP #10–12) pro-

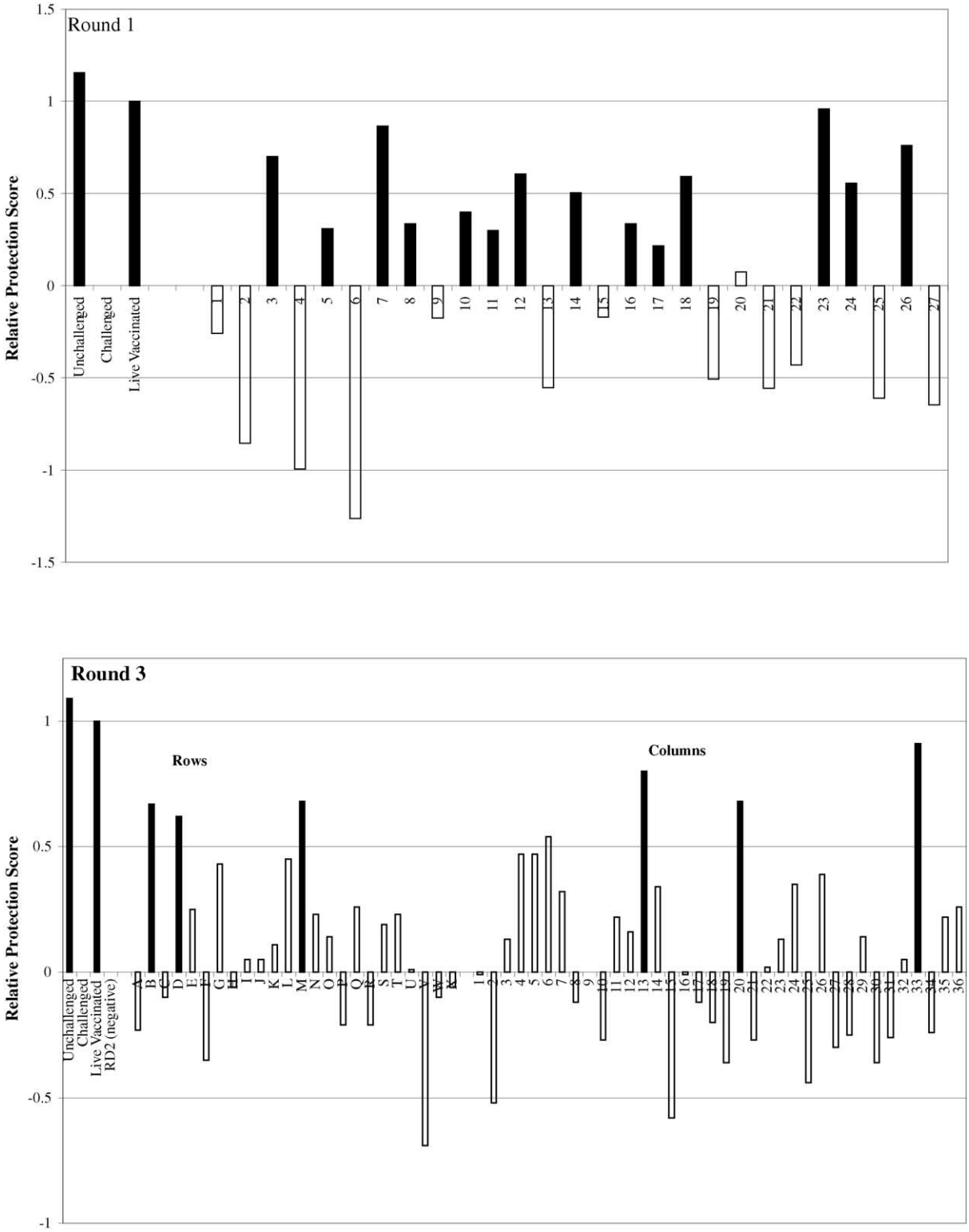

Fig. 2. Results of protection assays in rounds 1 and 3. Protection was scored as lung weight relative to the average of the vaccinated, maximum protection positive control (vaccinated, relative protection score = 1) and the non-vaccinated, challenged, maximum disease negative control (challenged = 0). There is a limit to the value that can be obtained with this readout, since no mouse can have a lower score than the naïve, unchallenged animals. The unchallenged naïve group consistently had lung weights slightly lower than the best protected positive controls (live vaccinated) due to the peribronchiolar accumulation of lymphatic cells in these animals. In rounds 2 and 3 the pools of plasmids from columns in the two-dimensional (2D) arrays are assigned numbers and the rows assigned letters. The solid bars indicate pools that were designated as protective and entered into the subsequent round.

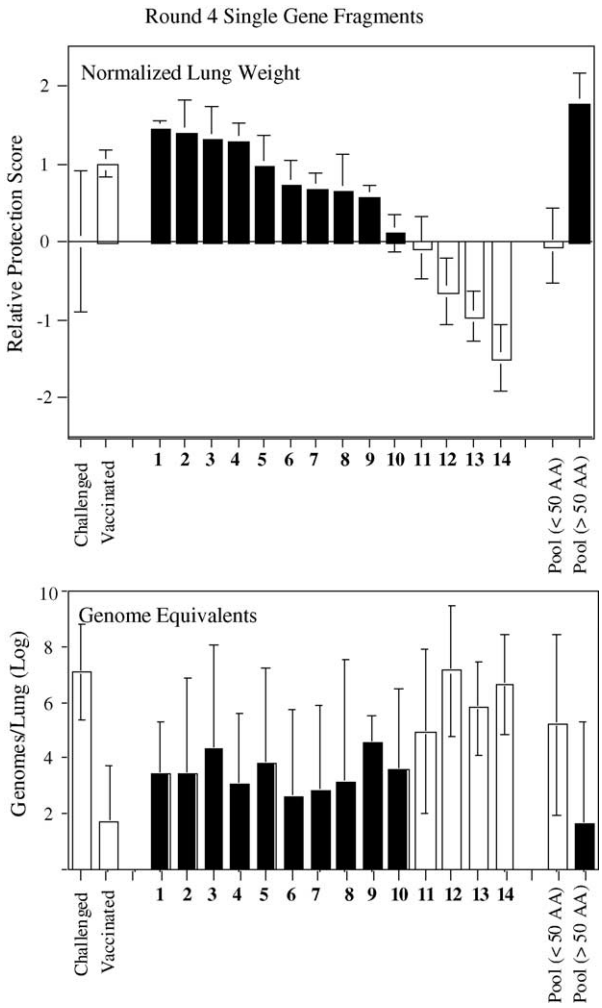

Fig. 3. Results of protection assays of individual gene fragments in round 4. Groups 1–14 are individual plasmids with *C. abortus* inserts encoding in-frame peptides of more than 50 amino acids. The numbers of each individual gene fragment tested correspond to the numbers in Fig. 4. The pool >50AA consists of all plasmids 1–14. Pool <50AA is the DNA of all 32 plasmids from Round 3 with predicted ORFs of fewer than 50 amino acids. (A) Relative protection scores (±S.D.) and (B) *C. abortus* B577 genomes per lung (±S.D.).

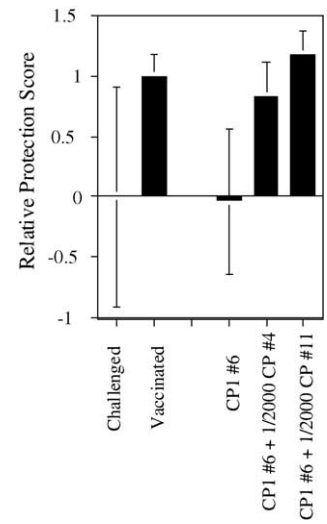

Fig. 4. Protection data from DNA pools. RD1 pool #6 is a non-protective pool from round 1. To test whether a single protective gene could be detected in a negative pool, 25 ng of either CP #4 or CP #11 were added to 50 μg of RD1 pool #6.

vide little or no protection. In fact, fragment CP #12, a fragment of the *C. abortus* 98 kDa putative outer membrane protein [43], appears to enhance disease, as does a second surface exposed protein fragment, CP #13 (protein translocase). The implication is that using the criterion of surface exposure or secretion is not sufficient to predict a vaccine candidate and that certain members in this class may be disease enhancing. Note that the isolation of a previously demonstrated protective antigen, CP #4, by this screen argues for the validity of this screening method.

Third, this protocol can differentiate parts of the same gene that have opposite effects on protection. Clones CP #6 and 13 are non-overlapping sections of a gene predicted to encode a protein translocase. Yet clone CP #6 is protective while CP #13 is disease enhancing. This implies that libraries based on fragmentation of genes may reveal antigens that would not have been evident from using whole genes, which is consistent with previous findings [36].

3.7. Retrospective informatic comparison

With the first completion of an ELI screen, we can ask if the same genes could have been arrived at more simply, namely by informatic approaches. In one sense the answer is no, since protective secreted and outer membrane proteins could not be distinguished [3]. However, we wanted to know if other informatic approaches could have predicted the same genes. After completion of this work, the genome sequence of the *C. caviae* genome sequence became available. This genome is the closest of all known *Chlamydiaceae* genome sequences to *C. abortus* B577, and was used to make sequence-based predictions [44]. We divided the 14 selected genes into three classes, protective, neutral and adverse, and applied available algorithms for predicting immune responses to them. The question was whether any would show a significant difference between these three classes. We tested algorithms predicting (1) density of peptides binding the MHC I of BALB/C mice [45], (2) density of peptides binding to the MHC II of BALB/C mice [45], (3) density of antigenic peptides (hydrophobicity) [46], and (4) density of coiled peptides [47]. Only the density of coiled peptides demonstrated significant differences between the three classes (Table 1). This description generally correlates with water exposed surfaces, so not surprisingly the membrane proteins score low by this criterion. However, none of these criteria were able to significantly separate the protective clones from non-protective starting from the whole genome sequence (data not shown).

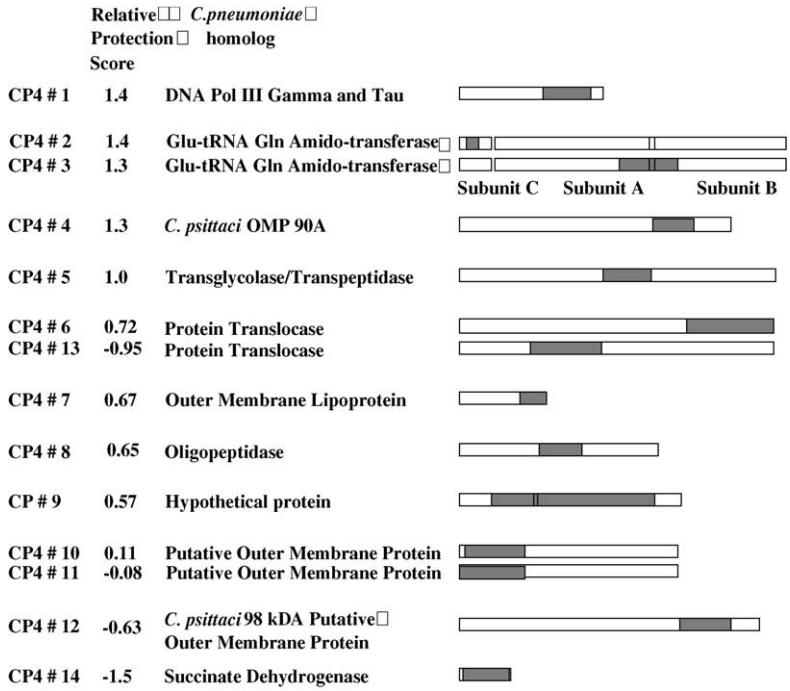

Fig. 5. Summary of characterization of the single gene fragments of round 4. The relative protection score of each *C. abortus* (CP) gene fragment is provided along with the designation of the gene in *C. pneumoniae* that has the highest similarity (*C. pneumoniae* homolog). In two cases, gene fragments CP #4 and 12, the *C. abortus* gene could be identified. On the right is a linear map showing the location in each gene of the fragment that conferred protection (shaded).

Table 1
Correlation between level of protection and coiled peptide density

C. pneumonia homolog	Position	ELI result	Density of coiled peptide (Number aa/100 aa)
DNA pol III Gamma and Tau (CP #1)	258–406	Protective	30
Glu-tRNA Gln Amidotransferase (CP #2)	25–65	Protective	32
Glu-tRNA Gln Amidotransferase (CP #3)	390–491	Protective	30
C. abortus OMP 90A (CP #4)	600–725	Protective	17
Transglycolase/transpeptidase (CP #5)	432–583	Protective	39
Average ± standard error			30 ± 4.0
Protein translocase (CP #6)	712–969	Neutral	13
Outer membrane lipoprotein (CP #7)	192–267	Neutral	21
Oliogopeptidae (CP #8)	252–385	Neutral	23
Hypothetical protein (CP #9)	78–222	Neutral	0
Putative outer membrane protein (CP #10)	27–185	Neutral	26
Average ± standard error			17 ± 5.2
Putative outer membrane protein (CP #11)	68–264	Negative	19
C. abortus 98 kDa putative outer membrane protein (CP #12)	672–833	Negative	9.3
Protein translocase (CP #13)	195–426	Negative	9.5
Succinate dehydrogenase (CP #14)	26–211	Negative	11
Average ± standard error			12 ± 2.7

4. Discussion

We have demonstrated that the total genome of the pathogen, *C. abortus*, can be screened in the mouse infection model for individual vaccine candidates in a relatively unbiased format. This is the first report of resolving genomic ELI libraries to single genes. Five gene fragments were capable of conferring as much protection as the gold standard, a live-vaccine. There was no propensity for outer membrane proteins to be isolated, in contrast to common prediction. Surprisingly, household genes provided the best protection. Different fragments of the same gene could confer either protection or exacerbation of disease symptoms. In fact, three of the 14 individual genes tested actually conferred more susceptibility to the disease. Of several bioinformatics algorithms that were tested retrospectively, only one could distinguish signif-

icant differences between the protective and non-protective groups and none could do so even in combinations from the whole genome.

The fact that high levels of protection were mediated by housekeeping genes presumably not exposed on the chlamydial surface can be rationalized if we consider the requirement of MHC II-restricted Th1 immunity for clearance of *Chlamydiaceae* [27,37,38]. The peptide fragments of these chlamydial proteins would have the opportunity to bind to the MHC II complex and be recognized by the immune system in this context. We chose to screen the genome with a library of fragments intended to favor partial gene clones. This decision was based on the often greater number of epitopes found to be reactive when genes are split, and also the dominant role of MHC II presentation in chlamydial protection over humoral-based immunity. We were more interested in optimizing the screen for identifying cell-based immunity than concerned about missing any conformational antibody determinant.

Expression library immunization (ELI) enables comprehensive genomic analysis of protective genes. The process demonstrated here should be readily applicable to any pathogen for which a suitable model exists. The major limitation is the size of the libraries needed if the genome is made of randomly sheared DNA. However, now that the genomes of many pathogens have been sequenced, the ELI process is greatly facilitated. Using the genome sequence to PCR or synthesize the ORFs reduces the complexity of the library ∼20-fold compared to the technology described herein. The development of linear expression element (LEE) technology now obviates the need for cloning, making the creation of libraries much faster [48]. ELI can be done by directly PCR-amplifying or chemically synthesizing all the genes of a pathogen, adding mammalian promoter- and terminator-encoding DNA fragments to create LEEs, and directly administering them in pools to the test animal via genetic immunization. The process described in this paper took over two years to complete. With the technological improvements described above, the process takes one year or less (KS, SAJ, unpublished results).

Since the 1995 description of ELI, several groups have reported creating genomic libraries that conferred protection [9,10,12–17]. Recently, a fraction of a fungal pathogen's genome, *Coccidioides immitis*, was reduced to a single protective gene by ELI [11]. In this case, ∼15% of the genome was surveyed using cDNA rather than genomic DNA. The gene isolated encodes a protein of unknown function and may be surface exposed. Other high-throughput approaches to vaccine discovery have been reported. Two noteworthy studies have used bioinformatics to identify proteins that should protect. In one study, in silico genomics was used to select 570 possible group B meningococcal vaccine candidates [4]. For technical reasons only a subset of these ORFs were cloned, expressed as recombinant proteins, purified, and tested, yielding 25 candidates that stimulated bactericidal antibodies in mice. In another study, informatics was used to produce surface proteins among *C. pneumoniae* ORFs, 141 of which were then tested [5]. The surface location of 28 of these antigens was confirmed from their ability to stimulate *Chlamydiaceae*-surface binding antibodies. While informative, these bioinformatics-based approaches have limitations, since as yet there are no criteria for accurately predicting what constitutes a good vaccine candidate. ELI as an unbiased and comprehensive approach may afford the databases on which to create better predictive algorithms.Finally, a critical question for identifying vaccine candidates using a model system is whether the candidates will be useful in the relevant host. As a first step towards answering this question, a preliminary vaccination experiment with *C. abortus* B577 was performed in cattle, the natural and economically important host [49]. To increase the probability that the correct antigenic epitope was included in the vaccine, we used the 5 most protective full-length genes and each was codon-optimized for mammals [50,51]. Five heifers were vaccinated with this pool. Six heifers were vaccinated with an Alum-Quil A based vaccine containing 50 μg of each of the affinity-purified protein fragments for these genes. Twelve control heifers were vaccinated with a pool of plasmids expressing unrelated *Salmonella typhimurium* genes. All groups received a booster vaccination. The protocol for this challenge system has been described [52]. Six of the twelve control heifers became pregnant (50% fertility), while the genetic vaccine group had 4/5 (80% fertility) pregnant animals, and the protein vaccine group had 5/6 (83% fertility) pregnant animals. Thus, 9 out of 11 animals in both vaccine groups were pregnant (82% fertility). This corresponds very well with typical data of bovine herds with and without fertility problems. The *p*-value for the pooled vaccines is 0.122 that vaccination is effective to improve *Chlamydophila*-induced reduction of fertility. In contrast to the screen in the mouse model, all heifers had previously been exposed to chlamydiae and experienced low-level herd infection [53]. This study is preliminary but required to justify a larger field trial, which it has. We present these results to underscore the notion that products of this ELI screen can be protective in the relevant host.

In summary, we have demonstrated that an unbiased, functional genomic search for bacterial vaccine candidates in vivo is feasible and produces vaccine candidates that would not have been predicted by contemporary knowledge-based approaches. Particularly with the current improvements in the protocol, ELI should be applicable to any pathogen.

Acknowledgements

We thank Qihua Sun and Mary Chien for excellent technical help and Irene Rombel for comments. This work was supported by grants from Bayer to BK and SAJ, a grant from DARPA to SAJ and unrestricted CBI funds to SAJ.

Appendix A

A.1. Data deposition

The following accession numbers are valid for swissprot (http://us.expasy.org/sprot/sprot-retrieve-list.html):

CP4 #1	Q823S5
CP4 #2	Q823W7
CP4 #3	Q823W6
CP4 #4	Q822Q5
CP4 #5	Q823N4
CP4 #6	Q821L5
CP4 #7	Q823C2
CP4 #8	Q822N8
CP4 #9	Q824V4
CP4 #10	Q823X2
CP4 #11	Q823X2
CP4 #12	Q823W9
CP4 #13	Q821L5
CP4 #14	Q821H1

References

[1] Piguet V, Trono D. Living in oblivion: HIV immune evasion. Semin Immunol 2001;13:51–7.

[2] Grifantini R, Bartolini E, Muzzi A, et al. Previously unrecognized vaccine candidates against group B meningococcus identified by DNA microarrays. Nat Biotechnol 2002;20:914–21.

[3] Grandi G. Antibacterial vaccine design using genomics and proteomics. Trends Biotechnol 2001;19:181–8.

[4] Pizza M, Scarlato V, Masignani VEA. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science 2000;287:1816–20.

[5] Montigiani S, Falugi F, Scarselli M, et al. Genomic approach for analysis of surface proteins in *Chlamydia pneumoniae*. Infect Immun 2002;70:368–79.

[6] Everett KD, Bush RM, Anderson AA. Emended description of the order *Chlamydiales*, proposal of *Parachlemydiaceae* fam. nov. and *Simkaniaceae* fam. nov., each containing one monotypic genus, revised taxonomy of the family *Chlamydiaceae*, including a new genus and five new species, and standards for the identification of organisms. Int J Syst Bacteriol 1999;49:415–40.

[7] Barry MA, Lai WC, Johnston SA. Protection against mycoplasma infection using expression-library immunization. Nature 1995;377(6550):632–5.

[8] Tang DC, DeVit M, Johnston SA. Genetic immunization is a simple method for eliciting an immune response. Nature 1992;356(6365):152–4.

[9] Alberti E, Acosta A, Hidalgo C, et al. Specific cellular and humoral immune response in balb/c mice immunised with an expression genomic library of *Trypanosoma cruzi*. Vaccine 1998;16:608–12.

[10] Brayton KA, Vogel SW, Allsopp BA. Expression library immunization to identify protective antigens from *Cowdria ruminantium*. Ann N Y Acad Sci 1998;849:369–71.

[11] Ivey FD, Magee DM, Woitaske MD, Johnston SA, Cox RA. Identification of a protective antigen of *Coccidioides immitis* by expression library immmunization. Vaccine 2003;21:4359–67.

[12] Manoutcharian K, Terrazas LI, Gevorkian G, Govezensky T. Protection against murine cysticercosis using cDNA expression library immunization. Immunol Lett 1998;62(3):131–6.

[13] Moore RJ, Lenghaus C, Sheedy SA, Doran TJ. Improved vectors for expression library immunization—applications to *Mycoplasma hyopneumoniae* infection in pigs. Vaccine 2001;20:115–20.

[14] Piedrafita D, Xu D, Hunter D, Harrison RA, Liew FY. Protective immune responses induced by vaccination with an expression genomic library of *Leishmania major*. J Immunol 1999;163(3):1467–72.

[15] Rainczuk A, Scorza T, Smooker PM, Spithill TW. Induction of specific T-cell responses, opsonizing antibodies, and protection against *Plasmodium chabaudi* adami infection in mice vaccinated with genomic expression libraries expressed in targeted and secretory DNA vectors. Infect Immun 2003;71:4506–15.

[16] Smooker PM, Setiady YY, Rainczuk A, Spithill TW. Expression library immunization protects mice against a challenge with virulent rodent malaria. Vaccine 2000;18(23):2533–40.

[17] Wizemann TM, Heinrichs JH, Adamou JE, et al. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. Infect Immun 2001;69:1593–8.

[18] Danesh J, Collins R, Peto R. Chronic infections and coronary heart disease: is there a link? Lancet 1997;350(9075):430–6 [see comments].

[19] Campbell LA, Kuo CC, Grayston JT. *Chlamydia pneumoniae* and cardiovascular disease. Emerg Infect Dis 1998;4(4):571–9.

[20] Muhlestein JB. Bacterial infections and atherosclerosis. J Invest Med 1998;46(8):396–402.

[21] Schachter J. Infection and disease epidemiology. In: Stephens RS, editor. Chlamydia: intracellular biology, pathogenesis, and immunity. Washington, DC: American Society for Miocrbiology; 1999. p. 139–70.

[22] Schachter JS, Dawson CR. Human chlamydial infections. Littleton, MA: PSG Publishing Company Inc.; 1978.

[23] Quinn, 1998. Does *Chlamydia pneumoniae* cause coronary heart disease? Curr Opin Infect Dis 11, 301–307.

[24] Kaltenboeck B, Kousoulas KG, Storz J. Structures of and allelic diversity and relationships among the major outer membrane protein (*ompA*) genes of the four chlamydial species. J Bacteriol 1993;175(2):487–502.

[25] Storz J. Antigenic structures and interrelations of PL agents associated with polyarthritis, enzootic abortion, intrauterine and latent intestinal infections. J Comp Pathol 1966;76:351.

[26] Rank RG. Models of immunity. In: Stephens RS, editor. Chlamydia: intracellular biology, pathogenesis, and immunity. Washington, DC: American Society for Microbiology; 1999. p. 239–96.

[27] Ward ME. Mechanism of *Chlamydia*-induced disease. In: Stephens RS, editor. Chlamydia: intracellular biology, pathogenesis, and immunity. Washington, DC: American Society for Microbiology; 1999. p. 171–210.

[28] Pal S, Barnhart KM, Wei Q, Abai AM, Peterson EM, de la Maza LM. Vaccination of mice with DNA plasmids coding for the *Chlamydia trachomatis* major outer membrane protein elicits an immune response but fails to protect against a genital challenge. Vaccine 1999;17(5):459–65.

[29] Vanrompay D, Cox E, Kaiser P, et al. Protection of turkeys against *Chlamydophila psittaci* challenge by parenteral and mucosal inoculations and the effect of turkey interferon-gamma on genetic immunization. Immunology 2001;103(1):106–12.

[30] Sandbulte J, TerWee J, Wigington K, Sabara M. Evaluation of *Chlamydia psittaci* subfraction and subunit preparations for their protective capacities. Vet Microbiol 1996;48(3–4):269–82.

[31] Tan TW, Herring AJ, Anderson IE, Jones GE. Protection of sheep against *Chlamydia psittaci* infection with a subcellular vaccine containing the major outer membrane protein. Infect Immun 1990;58(9):3101–8.

[32] Huang J, Wang MD, Lenz S, Gao D, Kaltenboeck B. Interleukin-12 administered during *Chlamydia psittaci* lung infection in mice confers immediate and long-term protection and reduces MIP-2 level and neutrophil infiltration in lung tissue. J Immunol 1999;162(4):2217–26.

[33] Huang J, DeGraves FJ, Lenz SD, et al. The quantity of nitric oxide released by macrophages regulates *Chlamydia*-induced disease. Proc Natl Acad Sci USA 2002;99:3914–9.

[34] Huang J, DeGraves FJ, Gao D, Feng P, Schlapp T, Kaltenboeck B. Quantitative detection of *Chlamydia* spp. by fluorescent PCR in the LightCycler. Biotechniques 2001;30(1):150–7.

[35] Cheng L, Ziegelhoffer PR, Yang NS. In vivo promoter activity and transgene expression in mammalian somatic tissues evaluated by using particle bombardment. Proc Natl Acad Sci USA 1993;90(10):4455–9.

[36] Sykes KF, Johnston SA. Genetic live vaccines mimic the antigenicity but not pathogenicity of live viruses. DNA Cell Biol 1999;18(7):521–31.

[37] Rottenberg ME, Gigliotti Rothfuchs AC, Gigliotti D, Svanholm C, Bandholtz L, Wigzell H. Role of innate and adaptive immunity in the outcome of primary infection with *Chlamydia pneumoniae*, as analyzed in genetically modified mice. J Immunol 1999;162(5):2829–36.

[38] Morrison RP, Feilzer K, Tumas DB. Gene knockout mice establish a primary protective role for major histocompatibility complex class II-restricted responses in *Chlamydia trachomatis* genital tract infection. Infect Immun 1995;63(12):4661–8.

[39] Feltquate DM, Heaney S, Webster RG, Robinson HL. Different T helper cell types and antibody isotypes generated by saline and gene gun DNA immunization. J Immunol 1997;158(5):2278–84.

[40] Kelly KA, Robinson EA, Rank RG. Initial route of antigen administration alters the T-cell cytokine profile produced in response to the mouse pneumonitis biovar of *Chlamydia trachomatis* following genital infection. Infect Immun 1996;64(12):4976–83.

[41] Kalman S, Mitchell W, Marathe R, et al. Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*. Nat Genet 1999;21(4):385–9.

[42] Longbottom D, Russell M, Dunbar SM, Jones GE, Herring AJ. Molecular cloning and characterization of the genes coding for the highly immunogenic cluster of 90-kilodalton envelope proteins from the *Chlamydia psittaci* subtype that causes abortion in sheep. Infect Immun 1998;66(4):1317–24.

[43] Longbottom D, Russell M, Jones GE, Lainson FA, Herring AJ. Identification of a multigene family coding for the 90 kDa proteins of the ovine abortion subtype of *Chlamydia psittaci*. FEMS Microbiol Lett 1996;142(2–3):277–81.

[44] Read TD, Myers GS, Brunham RC, Nelson EA. Genome sequence of *Chlamydophila caviae* (*Chlamydia psittaci* GPIC): examining the role of niche-specific genes in the evolution of the *Chlamydiaceae*. Nucleic Acids Res 2003;31:2134–47.

[45] Reche PA, Glutting JP, Reinherz EL. Prediction of MHC Class I binding peptides using profile motifs. Hum Immunol 2002;63:701–9.

[46] MacVectorTM, 6.5 User Guide, Oxford Molecular Ltd., 1999 [Section 14–19].

[47] McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server. Bioinfomatics 2000;16:404–5.

[48] Sykes KF, Johnston SA. Linear expression elements: a rapid, in vivo, method to screen for gene functions. Nat Biotechnol 1999;17(4):355–9.

[49] DeGraves FJ, Kim TY, Jee JB, Schlapp T, Hehnen H-R, Kaltenboeck B. Re-infection with *Chlamydophila abortus* by uterine and indirect cohort routes reduces fertility in cattle pre-exposed to *Chlamydophila*. Infect Immun 2004;72(5):2538–45.

[50] Stemmer WP, Crameri A, Ha KD, Brennan TM, Heyneker HL. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. Gene 1995;164(1):49–53.

[51] Nagata T, Uchijima M, Yoshida A, Kawashima M, Koide Y. Codon optimization effect on translational efficiency of DNA vaccine in mammalian cells: analysis of plasmid DNA encoding a CTL epitope derived from microorganisms. Biochem Biophys Res Commun 1999;261(2):445–51.

[52] DeGraves FJ, Stemke-Hale K, Huang J, et al. Vaccine identified by in vivo genomic screening enhances fertility in cattle during environmental challenge with Chlamydia. In: Schachter J, e.a., editor. Chlamydial Infections: Proceedings of the 10th International Symposium on Human Chlamydial Infection, Antalya, Turkey, 2002. p. 265–68.

[53] DeGraves FJ, Gao D, Hehnen H-R, Schlapp T, Kaltenboeck B. High prevalence of *Chlamydia psittaci* and *C. pecorum* vaginal infection in cattle quantitatively detected by high-sensitivity real-time PCR. J Clin Microbiol 2003;41:1726–9.