DOI 10.1002/pmic.200600047

REVIEW

A high efficiency cloning and expression system for proteomic analysis

Xuan Z. Ding[1], Ian T. Paulsen[2], Apurba K. Bhattacharjee[1], Mikeljon P. Nikolich[1], Gary Myers[2] and David L. Hoover[1]

[1] Department of Bacterial Diseases, Division of Communicable Diseases and Immunology, Walter Reed Army Institute of Reseach, Silver Spring, MD, USA
[2] The Institute for Genomic Research, Rockville, MD, USA

Received: January 19, 2006
Revised: March 17, 2006
Accepted: March 19, 2006

The recent description of the complete genomes of the two most pathogenic species of *Brucella* opens the way for genome-based analysis of the antigenicity of their proteins. In the present report, we describe a bench-level high-efficiency cloning and expression system (HECES) that allow expression of large numbers of *Brucella* proteins based on genomic sequence information. Purified proteins are produced with high efficiency in a microarray format conducive to analysis of their sero-reactivity against serum from immunized animals. This method is applicable at either small or large scale of protein processing. While it does not require robotics, the format is amenable to robotic implementation for all aspects of the process and subsequent analysis of protein characteristics. This method will allow selection of new reagents for diagnosis of brucellosis and development of vaccine against *Brucella*, an important zoonotic disease and biothreat agent.

Keywords:
Brucella / Cloning / Protein array / Protein expression

1 Introduction

Since the first complete free-living organism genome sequence was disclosed in 1995 [1], full-length genome sequences from over 300 different living beings have been examined and published. Among these, the genome sequences of *Brucella melitensis* and *B. suis*, the two most pathogenic species of *Brucella* for humans, were recently described by two reaserch groups led by Dr. DelVecchio and Dr. Paulsen [2, 3]. The *B. suis* genome contains 3388 predicted ORF on two chromosomes. Comparison of these sequences revealed extensive similarity and gene synteny [3]. Indeed, these two species differ by only 74 genes [4]. *B. suis* contains 42 unique genes found in 22 chromosomal regions; *B. melitensis* contains 32 unique genes in 11 locations [2–4]. Development of new proteomic systems to understand and combat this fascinating facultative intracellular bacterium will provide examples of new tools to relate the burgeoning wealth of bp sequence information to the living world.

Ever-advancing biochemical technology will result in exponentially increasing amounts of primary genetic sequence information applied to all forms of life. Mushrooming bioinformatics power will continue to disclose previously unanticipated relationships among and within organisms. Even with the limited genetic information available in the last two decades, at least 160 drugs and vaccines have been successfully brought to the marketplace and more than 371 new protein products are currently in clinical trials [5]. These numbers can potentially skyrocket if the massive amount of new information can be effectively harnessed. A fundamental challenge in accelerating this process is development of systems to interface the related-

Correspondence: Dr. Xuan Z. Ding, Department of Bacterial Diseases, Division of Communicable Diseases and Immunology, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910-7500, USA
E-mail: tom.ding@na.amedd.army.mil
Fax: +1-301-319-9123

Abbreviations: HECES, high-efficiency cloning and expression system; **LB**, Luria broth

ness of primary structure with the dynamic biologic processes that occur in whole, living beings. Because all of these processes are mediated, directly or indirectly, by proteins whose code we can now list, a great deal of effort is being focused on describing the entire set of proteins that an organism expresses. This process can either begin with the whole organism, which expresses different subsets of proteins under different experimental conditions, or with the whole genome, which encodes all possible proteins. Each approach has its limitations and advantages. Whole-organism approaches conceptually involve subjecting the organism to different experimental conditions, then physically extracting, separating, identifying and quantitating all the proteins present in each condition. These schemas have the advantage of assessing the organism's response to its environment, and permit detection of numerous translational and PTM effected by that environment. In addition, they can provide native proteins and identify interesting proteins to express for further study based on sequence information. The primary limitations of these methods are quantitative: the detection threshold may be too high, thus critical molecules of low number but high value or specific activity may be overlooked, or quantitation may be inaccurate, thus information must be depicted qualitatively. Moreover, these approaches examine the effect of the environment on the organism, but do not examine the response of the environment (if the organisms exist in a community) to the organism. In the pre-genomic era, a number of studies compared proteins expressed under different conditions or by closely related strains of *Brucella* [6–10]. Proteins were radioactively labeled with [^{35}S]methionine, then subjected to 2-DE to separate proteins. Differences between conditions were analyzed either by comparison of different patterns of protein spot localization [6–8] or by N-terminal microsequencing [9, 10] followed by database searching to identify previously described *Brucella* proteins or orthologs. Since description of the *B. melitensis* genome, additional studies have examined expression of the *B. melitensis* proteome under typical culture conditions or have compared different strains of *B. melitensis* [11–13] using 2-DE followed by matrix-assisted laser adsorption MS and querying the genomic database. These studies have identified a number of differences between test conditions. Of note, in all those studies in which the total number of protein spots was recorded [9–13], only 513 [11] to 883 [14] spots were found on 2-D gels, out of a possible 3197 *B. melitensis* proteins predicted by genome sequencing, suggesting that some proteins may not have been identified. These studies demonstrate both the utility and limitations of the whole organism-based proteomics approach.

The other proteomic approach involves preparing all the organism's proteins individually and evaluating their structure or function in specific assay systems. These assay systems can examine not only physical and chemical properties of the proteins, but can also test the effect of proteins on the ecosystem. For example, immunoassays can determine host response to *Brucella*, or proteins could be screened to detect potential toxin or enzymatic activity or binding to small molecules as potential therapeutic agents. In the pre-genomic era, *Brucella* investigators tried to identify immunoreactive proteins by electrophoretically separating lysates of whole organisms [15–17]. Isolated proteins were either screened directly on blots for recognition by immune serum [17] or purified from the gels and tested for their ability to stimulate immune cells [15, 16]. These approaches had the same quantitative limitations described above for whole-organism studies, but yielded potentially useful immunogenic proteins. In the post-genomic era, the quantitative limitations of whole-organism methods can be overcome by expressing the entire orfeome (all the proteins predicted to be encoded by recognized ORF) with suitable tags to facilitate purification. Alternatively, subsets of the orfeome can be expressed based on pre-selected determinants, *e.g.* membrane proteins, cytoplasmic proteins, etc. This approach has the advantage of providing a larger range of well-defined proteins for study. Limitations include the efficiency of protein expression and purification, inability to faithfully express all the possible transcriptional and post-transcriptional variants, proper presentation of the expressed proteins to the recognition/response system and potential alterations of the expressed protein properties by attachment of purification handles. Despite their limitations both proteomics approaches are complementary, facilitating bidirectional analysis of the organism's composition and interaction with inanimate and animate components of the environment.

In this report, we describe cloning and expression of 20 of the 3388 predicted *Brucella suis* ORF using a high-throughput methodology based on the Gateway method (Invitrogen). This high efficiency cloning and expression system (HECES) results in synthesis and purification of predicted proteins arrayed in a configuration and platform appropriate for analysis of biological systems. The method we describe integrates several commercially available technologies in a working flow that can be tailored to express anywhere from a small number of selected genes to a complete orfeome. This process should be applicable to laboratories with little or no robotic equipment, but can also be scaled for high-throughput operations. Besides its advantages over methods that purify proteins from whole organisms, the system we describe has other positive features. Whether used to express multiple genes pre-selected by logical *in silico* or *in veritas* methods, or for expression of the entire orfeome, it results in a complete DNA library of ORF that can be used for generation of DNA vaccines, mobilization into other expression systems, or construction of microarrays. On an orfeome level, one can avoid the limitations inherent in libraries derived by shotgun cloning. These libraries may not represent all ORF, contain many incomplete and non-coding sequences, are not arrayed with unique products suitable for economical assays and require later sequencing of each individual clone.

2 Design of HECES

The following flow chart shows each step of the HECES and the time required to complete each step. The entire process requires 11 working days from the beginning of PCR synthesis of genes to final protein examination. This time is for using Gateway cloning system that required two sets of gene cloning and plasmid preparation. It will be reduced by 3 days if interested genes are directly cloned into an expression vector. We used two competent cells, TOP10 and BL21 (DE3), in the project. TOP10 competent cells were used for high efficiency amplification of recombinant plasmids. BL21 (DE3) cells that carry the gene for T7 RNA polymerase under control of the *lac*UV5 promoter were used to induce protein expression with IPTG. Since all procedures are conducted in a 96-well format, most of the work can be performed by robotic equipment. Even without this equipment, based on our experience, one experienced technician could easily simultaneously process two 96-well plates. During the 11 days of processing, two plasmid plates, one gene bank and one protein cell bank are developed and stored. Of the two plasmid plates, one provides genes in an entry vector and can be further recombined into multiple destination vector based on research requirements. The other provides recombinants in a destination vector used for protein expression. The gene bank and the protein cell bank can be amplified and repeatedly used for later studies. All the reagents are commercially available; some of them can also be prepared in the laboratory.

Flow Chart of HECES

PCR synthesis of genes **(Storage)**	**Day 1**	4 h
↓		
E-gel detection of PCR products		1 h
↓		
Ligation of genes into pENTR vector	**Day 2**	1 h
↓		
Transformation of TOP10 cells		3 h
↓		
Spread agar plates and culture		3 h → O/N
↓		
Pick 2 colonies per gene and culture **(Gene bank)**	**Day 3**	4 h → O/N
↓		
Turbo miniprep of plasmids (Plasmid **Storage** → DNA sequence)	**Day 4**	3 h
↓		
E-gel detection of plasmids		1 h
↓		
PCR verification of plasmids		2 h → O/N
↓		
E-gel detection of PCR products	**Day 5**	1 h
↓		
Ligation of entry clones into a destination vector		1 h
↓		
Transformation of TOP10 cells		3 h
↓		
Spread agar plates and culture		3 h → O/N
↓		
Pick 2 colonies per gene and culture (Gene bank 2)	**Day 6**	4 h → O/N
↓		
Turbo miniprep of plasmids **(Plasmid Storage)**	**Day 7**	3 h
↓		
E-gel detection of plasmids		1 h
↓		
PCR verification of plasmids		2 h → O/N
↓		
E-gel detection of PCR products	**Day 8**	1 h
↓		
Transformation of BL21 cells and culture **(Protein cell bank)**		3 h → O/N
↓		
IPTG Induction of protein expression	**Day 9**	7 h
↓		
Collection and lyses of cells	**Day 10**	4 h
↓		
Protein capture in His Grab plate		2 h → O/N
↓		
ELISA detection of proteins	**Day 11**	4 h

3 Gene construction

3.1 Selection of *Brucella* genes

To assess the reliability of the working system, we selected 20 *B. suis* genes with different molecular weights, cellular location and function (Table 1). The encoded molecules include major outer membrane proteins, cytosolic proteins, enzymes and virulence determinants. Besides the 20 genes, two controls were processed in parallel. The negative control was pET-DEST-42 vector alone and the positive control was plasmid pET101/D/lacZ containing the *lacZ* gene.

3.2 PCR synthesis of ORF

Primers used for synthesis of the 20 *B. suis* genes were designed as 18- or 22-mers (Table 2). The forward 22-mers of primer contained a cacc sequence at the 5'end followed by the first 18 bases of the ORF. The reverse 18-mer primer was comprised of the last 18 bases of the ORF prior to the stop codon. In preliminary experiments, we compared three DNA polymerases, pfx, pfu and ThermalAce (Invitrogen). All three polymerases produce blunt-end PCR products and perform 3' end proofreading. ThermalAce provided the best result and was used to produce the data depicted in Fig. 1. The PCR reaction was conducted in a final volume of 50 μL that contained 5 μL of 10× ThermalAce reaction buffer, 1 μL of 50 mM dNTP, 1 μL of 1:10 diluted *B. suis* template, 0.5 μL of each 50 μM primer, 41 μL of PCR water and 1 μL of ThermalAce. *B. suis* template without primers was used as nega-

Table 1. Genes and control plasmids used in demonstration expression experiment

No	Name	ID	Size bp	AA	Da
1.	Outer membrane protein	52011	2346	781	87 582
2.	Outer membrane protein	53076	1872	623	68 192
3.	Integral membrane protein	51300	1634	544	61 877
4.	Membrane proteins	51395	1350	449	49 312
5.	Hypothetical membrane spanning pro.	51245	1047	348	37 457
6.	Transposase	51382	822	273	30 911
7.	Outer membrane lipoproteins carrier protein precursor	51251	603	200	22 338
8.	Hypothetical membrane spanning pro.	51324	414	137	15 665
9.	Protein secretion chaperonin CSAA	51672	348	115	12 824
10.	Hypothetical membrane spanning pro.	52194	252	83	8 635
11.	DNA polymerase III, alpha chain	53057	3237	1078	121 112
12.	Iron-regulated outer membrane protein	53346	1965	654	72 056
13.	DNAK protein	53183	1926	641	68 742
14.	Outer membrane protein TOLC	52210	1371	456	48 490
15.	31-kDa immunogenic protein precursor	51977	990	329	34 272
16.	Outer membrane protein	52241	753	250	27 289
17.	Outer membrane protein	52742	594	197	21 254
18.	Small heat shock protein HSPA	52965	483	160	18 425
19.	Outer membrane lipoprotein	53258	381	126	13 259
20.	Heat shock protein 15	51458	402	133	14 731
	pET-DEST-42	Negative	0	0	0
	pET101/D/LacZ	Positive	3171	1056	120 000

Table 2. PCR primers used for synthesis of *B.* genes

No.	Forward	Reverse
1.	>5′-cacc atgacggcaa gttctaaatt	<5′-gaactttgtc gatacaccga
2.	>5′-cacc atggcgtttg agatttttgg	<5′-aaacgcc tgtcctatgc
3.	>5′-cacc atgcggtttc cacgcccaac	<5′-actctgcgat gccaccgccg
4.	>5′-cacc atgaattca ataggtcttt t	<5′-cgtatca ttttgtgtcc
5.	>5′-cacc atggggca gggggttact tt	<5′-tgtaaaatta aagtttcggc
6.	>5′-cacc atgcgagcgt taatcgcggc	<5′-tgctgtg agttcggcga
7.	>5′-cacc atgggcgtcg gcgttgcggg	<5′-ctggcctttg cgcttcatcg
8.	>5′-cacc atggagccgg gtgttctctc	<5′-cagaacc agagccagaa
9.	>5′-cacc atgagcgaag cagccaccat	<5′-aaaaagcttg ccaccattcg
10.	>5′-cacc atgggtgaag caggtattgg	<5′-cccgcct atcgcgcggc
11.	>5′-Cacc agtgatggttc cttattttga	<5′-atggaagtcg cgtgagttcg
12.	>5′-Cacc atgctcgcca gcacatctct	<5′-gaacttgaag gccgtctgga
13.	>5′-Cacc atggagagaa atatggctaa	<5′-cgacgacttc ttgttgtcgt
14.	>5′-Cacc atgaggtaca cggtgttcaa	<5′-gcggccatca ggcgtacgca
15.	>5′-Cacc atgaaattcg gaagcaaaat	<5′-tttcagcacg cccgcttcct
16.	>5′-Cacc atggttgcgc ttggtatcgg	<5′-cttctttgca gcgtcccggg
17.	>5′-Cacc atggccgcaa ccgcagccgg	<5′-cgtcttgcgc gcattcttga
18.	>5′-Cacc atgagaagat atgccgaaag	<5′-gtccttgacc gctatgtcga
19.	>5′-Cacc atgaaacgct tccgcatcgt	<5′-gccggcgttg cggcgggtga
20.	>5′-Cacc atggcg accg ctgga acaaa	<5′-ccagccttcg cgcaggcgat

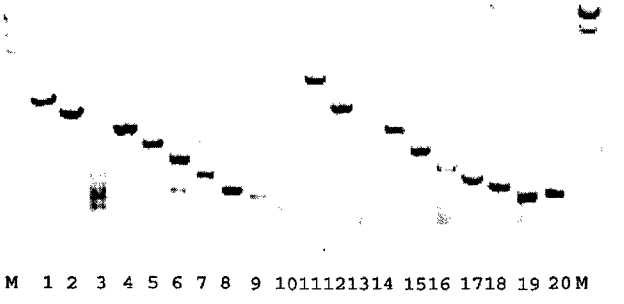

Figure 1. PCR products. Twenty genes of *B. suis* were PCR synthesized using ThermalAce. The molecular weights of the genes have been described in Table 1. The PCR products were detected on 1% agarose gel. M = λ DNA/Hind III fragments (Invitrogen).

tive control. The *lacZ* gene with primers from Invitrogen was used as a positive control. The PCR protocol was designed as 1 cycle of 95°C for 2 min, 30 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 2.5 min, 1 cycle of 72°C for 10 min and then refrigeration at 4°C. Eighteen out of 20 genes were successfully synthesized on the first attempt using this PCR protocol (Fig. 1). Genes 3 and 13 were obtained in replacement experiments by reducing the annealing temperature to 52 and 53°C, respectively.

3.3 Preparation of recombinant constructs in entry vector

We selected the Gateway cloning system (Invitrogen) to provide versatility for additional cloning for future applications. The traditional cloning method requires site-specific restriction of both insert and vector with enzymes and then ligation with ligase. The topo cloning method uses the properties of vaccinia DNA topoisomerase I to develop a ligase-free tech-

nology for the covalent joining of DNA fragments to suitable plasmid vectors [18]. Reactions were performed according to the manufacturer's instructions. PCR product (1 μL) plus 1 μL of water was mixed with 0.5 μL of each salt solution and pENTR/SD/D-TOPO vector. Ligation was conducted in a 96-well plate at room temperature for 5 min. The ligation reaction (3 μL) was then added to 40 μL of TOP10 competent cells and incubated at 4°C for 30 min. After heating at 42°C for 30 s, the cells were added with 250 μL culture medium and shaken at 37°C for 1 h. After the incubation, 20 μL of TOP10 cells were streaked in wells of a 12-well tissue culture plate containing trypticase soy agar with 50 μg/mL kanamycin for clone selection. Plates were incubated at 37°C overnight. All 20 samples resulted in colonies on the selection medium (Fig. 2). The vector control well was empty and the LacZ positive control had six to eight colonies. Gene 15 had only one colony. In a replacement experiment, there were enough colonies obtained for this gene. Thus, 19/20 genes were successfuly cloned on the first attempt using the PCR products. We attempted to save time at this labor-intensive step by selecting recombinant cells in liquid instead of plating on solid medium as previously suggested [19], but experienced an unsatisfactorily lower cloning eficiency. Moreover, we have concerns that liquid culture may result in a mixed culture of recombinant and native cells, which could lead to confusion after long-term culture or successive passages. We do, however, use liquid medium to select cells transformed by destination vector for protein expression, as inclusion of a minority of cells that have lost the plasmid should have a negligible effect on expression and reduced expression would be readily detected.

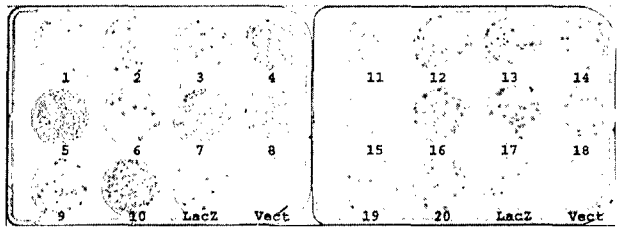

Figure 2. Colonies of pENTR-*Brucella* genes on 12-well agar plate. *B. suis* genes were cloned into pENTR vector and amplified in Top 10 *E. coli* cells. LacZ was used as the positive control and vector alone was used as the negative control.

3.4 preparation of plasmids from entry clones

Four colonies of each gene were picked and cultured in 1.3 mL Luria broth (LB) per well in a 96-well plate for propagation of the plasmid. On the next day, the plasmids were purified using the turbo miniprep method (QIAgen). Twelve microliter of the purified plasmid were electrophoresed on an E-gel system (Invitrogen). Figure 3 shows plasmids isolated from the four colonies (rows A to D or E to H) of each gene (columns 1 to 10). Column 13 was λDNA/HindIII (Invitrogen) as a molecular weight marker. Plasmids of

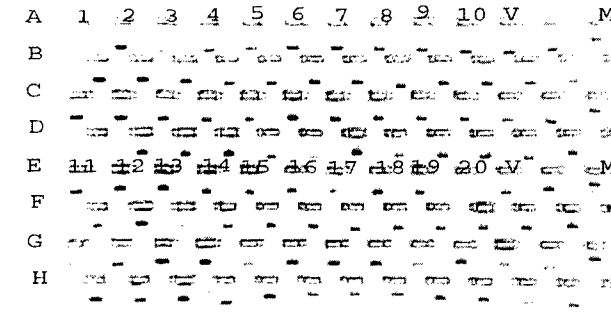

Figure 3. pENTR recombinant constructs. Four colonies for each pENTR recombinant constructs were selected and amplified. The plasmids (A–D or E–H) were purified and detected in E-gel. The 20 *B. suis* genes are placed as indicated in columns 1 to 10. Column 11 was pENTR vector alone and column 13 was λDNA/HindIII (Invitrogen) as a molecular weight marker.

pENTR vector alone (column 11) showed a lower molecular weight than the plasmids derived from the colonies of 20 genes, although the resolution of the gel limited interpretation of molecular weights.

3.5 Confirmation of pENTR plasmids

The presence of inserted genes in the plasmids was confirmed by PCR, using primers designed according to the M13 sequence of the pENTR vector. Their sequences are M13 forward 5'-TGTAA AACGA CGGCC AGT CT and M13 reverse 5'- CA GGAAA CAGCT ATGAC C. A negative control used the vector alone as the PCR template. Electrophoresis of PCR products on E-gel showed that all the 20 genes had at least one positive clone, whose plasmid contained a gene insert with correct molecule size (data not shown). As noted above, molecular weights were not reliably determined with the E-gel method, but the method was sufficient to demonstrate the presence of specific gene products. A second confirmation step taken in parallel was to end-sequence the inserts in pENTR using M13 forward and reverse primers. Sequence verification confirmed the presence of the correct insert in all cases and identified at least one positive clone for all the *B. suis* genes. These results indicate that efficiency of entry cloning was 100%. After the confirmation, positive cell clones were prepared in 20% glycerol stock and labeled as **GENE BANK 1** (Entry bank). The positive plasmids were also stored for future studies.

3.6 Cloning of genes into destination vectors

One positive plasmid for each gene was selected for ligation into Gateway destination vector pET-DEST42. Each ligation reaction (LR) contained 2 μL of buffer, 4 μL of pENTR-gene, 2 μL of pET-DEST42 vector and 2 μL of LR clonase. The reactions were incubated at 25°C for 60 min. Proteinase K (1 μL) was added and the reaction mixture incubated at 37°C for another 10 min. For transformation, 2 μL of each LR was

transformed into 50 μL of TOP10 competent *Escherichia coli* cells and incubated at 4°C for 30 min. After heating at 42°C for 30 s, the cells were added with 250 μL culture medium and shaken at 37°C for 1 h. TOP10 cells (20 μL) were then streaked in wells of 12-well LB agar culture plates containing 100 μg/mL ampicillin for clone selection. Plates were incubated at 37°C overnight. On the next day, four colonies of each gene were picked and cultured in 1.3 mL LB solution in a 96-well plate at 37°C overnight. Plasmids were purified from these wells using turbo miniprep and purity verified on E-gels. Insertion of *B. suis* genes in the destination constructs was confirmed by PCR (Fig. 4), using primers designed according to the attl sequence located on the pET-DEST42 vector. Their sequences were attL1 5'- G TAC AAA AAA GCA GGC T, attL2 5'- GTA CAA GAA AGC TGG GT. Positive recombinant cells were prepared in 20% glycerol stock and labeled as **GENE BANK 2** (destination bank). The positive plasmids were also stored for future studies.

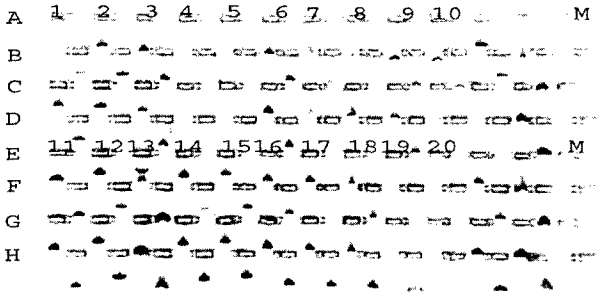

Figure 4. Confirmation of *B. suis* genes in destination vector. Destination constructs were confirmed by PCR using Attl primers located on the vector. As described in Fig. 3, column 11 is negative control (vector alone as template) and column 13 is λDNA/HindIII (Invitrogen) as a molecular weight marker.

4 Protein expression and purification

4.1 Transformation of expression cells

The purified pDEST-*B.* plasmid of each gene (2 μL) was mixed with 50 μL of expression host cells, BL21Star(DE3) and incubated at 4°C for 30 min. After heating at 42°C for 30 s, the cells were added with 250 μL culture medium and shaken at 37°C for 1 h. Then the host cells were added with 1 mL of LB solution containing 100 μg/mL of ampicillin and shaken at 37°C overnight.

4.2 Induction of protein expression

On the next day, 6 μL of culture from each sample were transferred into 1.3 mL of LB solution containing 1% glucose and 100 μg/mL ampicillin and shaken at 37°C for 2 h. Bsaed on our previouse experiments, the OD600 measurements of the cell cultures would be within 0.5–0.8 U after 2 h of incubation. The cells were then induced with 1 μM IPTG (Invitrogen) for 5 h. The optimal time was also obtained in preliminary experiments. After protein induction, cells were centrifuged and dessolved in PBS for detection.

4.3 Western blot analysis of the expressed recombinant proteins

Although not part of the planned HECES, we performed Western blots using anti-6-His mAb (Invitrogen) to verify protein size. Protein sample (7 μL) for each gene was mixed with an equal amount of 2X SDS loading buffer. The samples were boiled for 10 min and purified on 13% acrylamide/bis gel. The protein samples were transfered onto cellulose membrane and hybridized overnight with 1:2000 alkaline (AP)-conjugated anti-6-His. The membranes were developed with 1 mg/mL of Naphthol AS-MX phosphate (Sigma) and 2 mg/mL Fast red TR salt (Sigma) in 50 mM Tris buffer. Sixteen out of the 20 genes were expressed in the first round of induction following the HECES. Proteins 3, 10, 19 and 20 were not detected in this experiment. Protein 10, 19 were induced and detected in the second round of experiment. Examination of the expressed protein by anti-6-His indicates that the recombinant proteins were expressed in their correct ORF (Fig. 5).

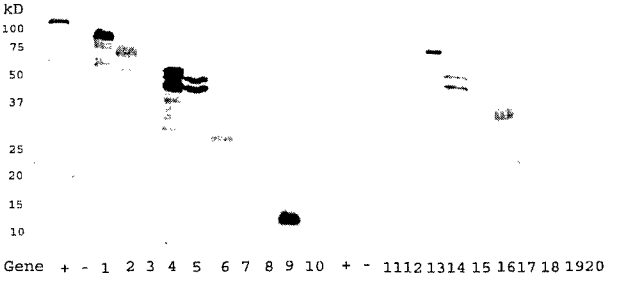

Figure 5. Western blot analysis of expressed recombinant *B. suis* proteins. The recombinant cells were induced by 1 M IPTG for 5 h. Samples were collected after induction and detected by 1:2000 anti 6-His /VP mAb. The pDEST-LacZ was used as positive control (+) and vector alone transformed cells were used as negative control (−).

4.4 Purification of recombinant proteins

Western blot is a conventional method to purify and detect proteins, but it is limited for large-scale processing proteins due to its working load, time and cost. These recombinant proteins have been designed to be fused with His and V5 tags at C-terminal for further purification. For high throughput purposes, and to array material in a convenient format, we used HisGrab (Pierce) 96-well plates to purify and immobilize proteins for later analysis [20]. HisGrab plates are prepared with versatile metal chelate chromatography material, Ni-NTA silica at the bottom of each well and used for rapid purification of proteins from crude cell lysates under either native or denaturing conditions. The experiment was con-

ducted using ten expressed proteins and two controls in one row (12 wells) of 96-well plate. Samples from induced cells of ten selected proteins were dissolved in lysis buffer, kept at room temperature for 30 min, then frozen and thawed three times. Cell lysate (100 μL) was added to triplicate wells, and plate was incubated at room temperature overnight. After washing six times with 1× PBS, 100 μL of anti-V5 mAb conjugated with AP (1:2000 dilution) was added to each well and incubated at 4°C for 16 h. The samples were detected with a plate reader after substrate development. Figure 6 shows that ten out of ten processed recombinant proteins were captured and detected by the method. The negative control was cell lysate prepared from cells transformed with vector alone. The positive control was LacZ-transformed cells (Invitrogen). The positive reactions were consistent with Western blot data (Fig. 5), indicating the suitability of this method to capture 6His-tagged proteins in a microtiter plate format with our lysis method.

The purity of recombinant proteins is very important for further examination for their structure and function. To determine how well these proteins were purified with His-Grab plate, we washed the well extensively and eluted the protein samples from each well. We then examined pooled samples by both Western blot and CBB staining. We saw a single, but very faint band for each protein (data not shown). To further confirm this result, we purified four expressed recombinant proteins using Ni-NTA spin columns. The purification principle of the column was the same as the His-Grab plate, but it provided higher yield. After elution, the purified proteins were analyzed by both Western blot and CBB staining. In contrast to the multiple bands noted on whole-cell lysates, only one band was present in the elutes (Fig. 7A and B). This result further confirms that the expressed recombinant *B. suis* proteins are highly purified by the Ni-NTA purification method.

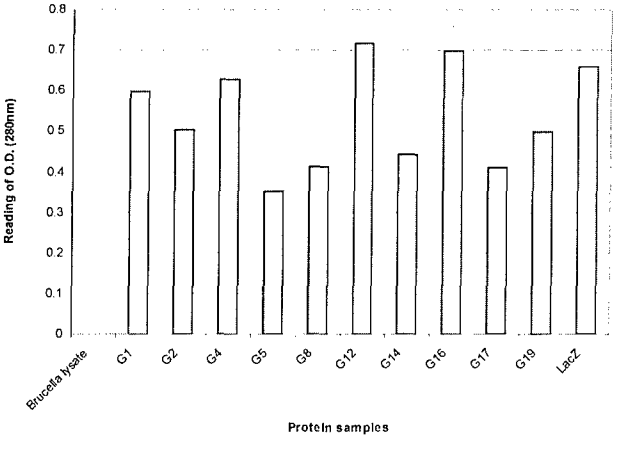

Figure 6. Purification of *B. suis* recombinant protein using His-Grab plate. Ten *B. suis* recombinant proteins were treated with cell lysate buffer and loaded on the plate. The protein samples were detected using 1:2000 mAb against V5. The negative control was vector alone transformed cells and the positive control was lacZ gene-transformed cells.

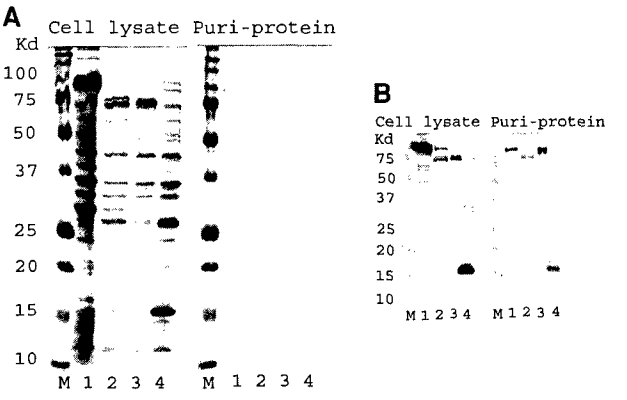

Figure 7. Analysis of recombinant protein purity. Four expressed *B. suis* recombinant proteins were examined before (whole-cell lysate) and after (purified protein) purification on Ni-NTA spin columns. The protein samples were examined by CBB staining (A) and Western blot (B). M represents protein standard (Bio-Rad).

5 Screening of recombinant proteins using immunized animal serum

In previous studies, we showed that mice immunized with live, attenuated *Brucella* vaccines made serum antibody to both LPS and to non-LPS protein components of *Brucella* [21]. Anti-LPS antibody titers were far higher than ani-protein titers. To obtain high titer anti-protein antibodies to test the immunogenicity of recombinant proteins, rabbits were immunized with a dialyzed whole-cell lysate (RFBL) of rough mutant *Brucella* strain WRR51 [21]. This strain lacks *wboA*, which encodes a glycosyltransferase required for synthesis of the O-polysaccharide (OPS) side chain on LPS. Rabbits were immunized intramuscularly with two doses of vaccine (25 μg protein per dose) given 4 weeks apart. Serum was obtained before immunization (pre-bleed) and 2 weeks after the second dose of vaccine (immune). IgG was prepared from a portion of the immune serum by affinity chromatography on Protein G-Sepharose (Pharmacia-LKB Biotech) and final concentration was 1.8 mg/mL (Bhattacharjee *et al.*, unpublished data). Research was conducted in compliance with the Animal Welfare Act and other federal statutes and regulations relating to animals and experiments involving animals and adheres to principles stated in the *Guide for the Care and Use of Laboratory Animals*, NRC Publication, 1996 edition.

Eighteen recombinant *Brucella* proteins and recombinant LacZ as a negative control were captured on 96-well HisGrab plates as described above. Additional wells were coated with 100 μL of RFBL at 1 μg/mL as a positive control. Quintuplicate wells were used for each protein. ELISA was then performed using 1:2000 of either normal rabbit serum or imunized rabbit serum. A second anitbody, AP-conjugated goat anti-rabbit IgG was used instead of anti-V5 as the detecting antibody.

Immune rabbit serum, but not pre-bleed serum, strongly bound to RFBL (Fig. 8). In contrast, binding of both sera to LacZ was weak. Immune serum also recognized recombinant *Brucella* protein numbers 1, 2, and 16, with 11.2-fold, 10.6-fold and 5.1-fold increased binding compared to pre-bleed serum. The experiments were repeated four times. The differences between OD readings of pre-bleed and immune serum for these proteins were highly significant and were reproducible in another, identical experiment (data not shown). In the experiment, proteins 1, 2 and 16 were also specifically recognized. None of the individual proteins bound as much antibody as the *Brucella* lysate. Although a number of the other proteins in the experiment presented in Fig. 8 also had statistically significantly increased binding to immune serum, the magnitude of increased binding was small and was not consistently reproducible in repeat experiments. These data suggested that at least proteins 1, 2 and 16, all outer membrane proteins, would be of interest as potential target antigens for serologic diagnosis of brucellosis.

The screening experiments will allow quantitative comparison of the intensity of serologic recognition elicited by each protein on a molar basis, because, ideally, all proteins will be represented in equimolar amounts by virtue of their ligation to the plate by their 6-His moiety. Immobilization to the plate by anti-6-His is expected to maintain the large majority of proteins in their native conformation. Maintenance of conformational integrity may not be important for analysis of the cellular response, which depends on recognition of peptides presented on antigen presenting cells in the context of MHC molecules. It is, however, crucial for screening the protein bank for recognition by antibodies, which may recognize not only linear epitopes but also conformational epitopes that depend on proper protein folding. AS we expect those proteins identified by sero-reactivity to be useful for development of diagnostic tests rather than as vaccine antigens, this approach favors selection of antigens particularly relevant for this purpose. The serologic screening approach could also be used in future studies to evaluate the maturation of antibody avidity or class switching against all or some *Brucella* antigens during the course of natural disease or after vaccination. Comparison of the reactivity of mouse, human and nonhuman primate sera against all proteins in the bank may provide insight into conformation- or sequence-specific differences in antigen processing and presentation.

6 Discussion

These studies describe a method to use information from the *B. suis* genome sequence to clone, express and purify, in an assay-friendly multiwell format, multiple genes that express proteins of different sizes, cellular locations, and function. This fast, dependable and inexpensive method concatenates commercially available high-throughput and high-efficiency technologies in a versatile system. The approach can be used either for a relatively small number of proteins, as shown here, for expression of larger sets of proteins selected according to a particular function or structure, or for an entire orfeome. Based on our experience, the system could easily be used at a bench scale to produce up to 192 proteins in 11 working days. A major advantage of this method over traditional libraries based on shotgun cloning or derived from analysis of lysates of whole organisms is that essentially all of the predicted genes and 80% of their products can be produced in high purity. Since this preliminary study included a wide variety of genes, we expect that this level of effi-

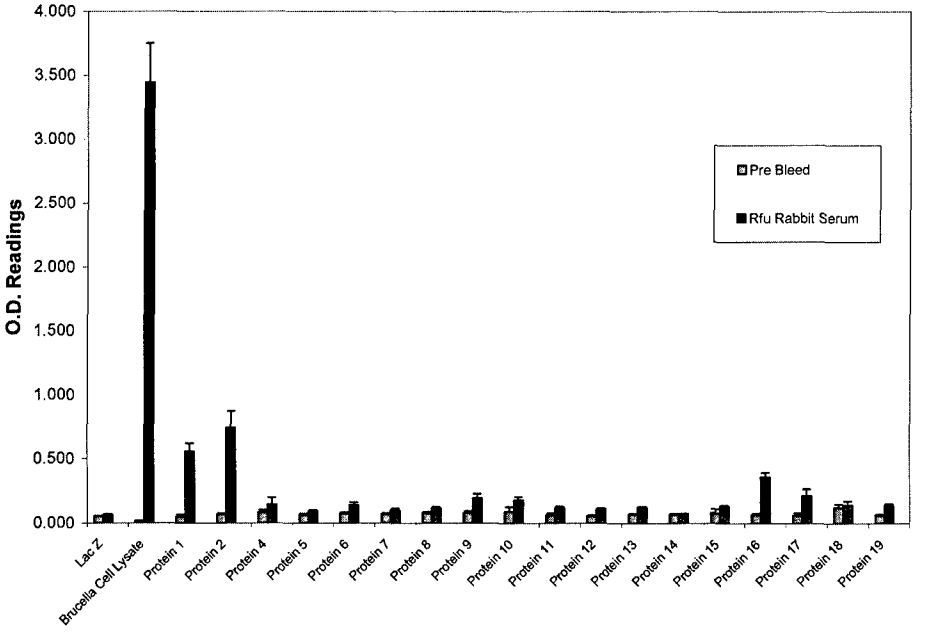

Figure 8. Screening of *B.* recombinant proteins using immunized rabbit serum. Eighteen *B.* recombinant proteins were purified in HisGrab 96-well plates. Antiserum binding activity was evaluated using 1:2000 diluted rabbit pre-bleeding and immunized serum. Negative control was recombinant LacZ protein and positive was *B.* cell lysate. The experiments were repeated four times. Data were processed with analysis of variance and Student's *t*-test for comparison of groups. $p < 0.05$ was used as significance level for the study.

ciency would be applicable over the entire orfeome. One should note that the success of this method is critically dependent on high-quality sequence data. Our studies confirm the quality of the *B. suis* genomic data, as all 20 genes had the correct sequence. They also validate the algorithms used by TIGR to delineate the ORF; predicted reading frames were all correct, as the proteins were in frame with the 6His and V5 tags at the C terminus.

In addition to the excellent efficiency, use of the 96-well format allows easy confirmation of gene sequence and the presence of proteins; moreover, expressed proteins are in a convenient format for further high-throughput assays. In the present study, we have identified three new antigenic proteins that are recognized by immune rabbit serum. Interestingly, all of these proteins are predicted to be components of the outer bacterial membrane. These proteins may be useful for development of diagnostic tests for brucellosis. Although *Brucella* LPS is the immunodominant antigen for elicitation of antibody responses in infected humans and animals antibodies to it cross-react with the OPS of *Francisella tularensis*, *Yersinia enterocolitica* O:9, *Vibrio cholerae*, and *E. coli* O157 [22]. Because of this cross-reactivity, interpretation of low antibody titers in individuals potentially exposed to these organisms is difficult. It is possible that use of a large number of recombinant *Brucella* proteins as targets in ELISA or other serologic assays would overcome these difficulties. Attempts to use *Brucella* protein antigens as targets in serologic tests, however, have been hampered by lack of sensitivity and specificity. In part, these attempts failed because of incomplete removal of LPS from bacterial lysates used as a source of antigen, leading to false-positive results [23, 24]. Use of individual recombinant proteins as targets in ELISA increases specificity, but titers are lower and delayed relative to the anti-OPS response [25]. Interestingly, a previous study showed that cytoplasmic proteins are better than outer membrane proteins as target antigens [25]. In the present study, we conclusively identified only outer membrane proteins as targets of antibodies from animals immunized with a whole-cell lysate. It is possible that, by screening a large number of membrane and non-membrane proteins, we will be able to identify sero-reactive antigens from a number of protein classes. We have, for example, initiated studies to examine the ability of a set of *Brucella* proteins to be recognized by cells and proteins from immunized animals. This approach may permit discovery of new candidates for development of vaccines and diagnostics. Moreover, analysis of these arrayed data by evolving bioinformatics methods should greatly enhance our understanding of the impact that *Brucella* has on the host cells it encounters. Besides the protein library, the gene libraries constructed by this method are also suitable for development of microarrays, for moving into other expression vectors, or for use as DNA vaccines. We are extending our current studies to express all predicted membrane proteins of *Brucella* and screening them with immune cells and sera. This approach may permit discovery of new candidates for development of anti-*Brucella* vaccines and diagnostics.

The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

7 References

[1] Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A. *et al.*, *Science* 1995, *269*, 496–512.

[2] DelVecchio, V. G., Kapatral, V., Redkar, R. J., Patra, G. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *99*, 443–448.

[3] Paulsen, I. T., Seshadri, R., Nelson, K. E., Eisen, J. A. *et al.*, *Proc. Natl. Acad. Sci. USA* 2002, *99*, 13148–13153.

[4] Tsolis, R. M., *Proc. Natl. Acad. Sci. USA* 2002, *99*, 12503–12505.

[5] Das, R. C., *American BioTech Lab.* 2003, *21*, 812.

[6] Lin, J., Ficht, T. A., *Infect. Immun.* 1995, *63*, 1409–1414.

[7] Rafie-Kolpin, M., Essenberg, R. C., Wyckoff, J. H., 3rd, *Infect. Immun.* 1996, *64*, 5274–5283.

[8] Sowa, B. A., Kelly, K. A., Ficht, T. A., Adams, L. G., *Appl. Theor. Electrophor.* 1992, *3*, 33–40.

[9] Teixeira-Gomes, A. P., Cloeckaert, A., Bezard, G., Dubray, G., Zygmunt, M. S., *Electrophoresis* 1997, *18*, 156–162.

[10] Teixeira-Gomes, A. P., Cloeckaert, A., Zygmunt, M. S., *Infect. Immun.* 2000, *68*, 2954–2961.

[11] Eschenbrenner, M., Wagner, M. A., Horn, T. A., Kraycer, J. A. *et al.*, *J. Bacteriol.* 2002, *184*, 4962–4970.

[12] Mujer, C. V., Wagner, M. A., Eschenbrenner, M., Horn, T. *et al.*, *Ann. NY Acad. Sci.* 2002, *969*, 97–101.

[13] Wagner, M. A., Eschenbrenner, M., Horn, T. A., Kraycer, J. A. *et al.*, *Proteomics* 2002, *2*, 1047–1060.

[14] DelVecchio, V. G., Wagner, M. A., Eschenbrenner, M., Horn, T. A. *et al.*, *Vet. Microbiol.* 2002, *90*, 593–603.

[15] Brooks, W. B., Splitter, G. A., *Infect. Immun.* 1992a, *60*, 2459–2464.

[16] Brooks, W. B., Splitter, G. A., *Infect. Immun.* 1992b, *60*, 2136–2138.

[17] Teixeira-Gomes, A. P., Cloeckaert, A., Bezard, G., Bowden, R. A. *et al.*, *Electrophoresis* 1997, *18*, 1491–1497.

[18] Heyman, J. A., Cornthwaite, J., Foncerrada, L., Gilmore, J. R. *et al.*, *Genome Res.* 1999, *9*, 383–392.

[19] Dieckman, L., Gu, M., Stols, L., Donnelly, M. I. *et al.*, *Protein Expr. Purif.* 2002, *25*, 1–7.

[20] Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R. A., *et al.*, *Proc. Natl. Acad. Sci. USA* 1991, *88*, 8972–8976.

[21] Hoover, D. L., Crawford, R. M., Van De Verg, L. L., Izadjoo, M. J. *et al.*, *Infect. Immun.* 1999, *67*, 5877–5884.

[22] Chart, H., Jenkins, C., *J. Appl. Microbiol.* 1999, *86*, 731–740.

[23] Baldi, P. C., Miguel, S. E., Fossati, C. A., Wallach, J. C., *Clin. Infect. Dis.* 1996 *22*, 446–455.

[24] Goldbaum, F. A., Rubbi, C. P., Wallach, J. C., Miguel, S. E. *et al.*, *J. Clin. Microbiol.* 1992, *30*, 604–607.

[25] Letesson, J. J., Tibor, A., Van, E. G., Wansard, V. *et al.*, *Clin. Diagn. Lab. Immunol.* 1997, *4*, 556–564.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
13-09-2006		

4. TITLE AND SUBTITLE

A high efficiency cloning and expression system for proteomic analysis

(Review paper)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Xuan Z. Ding1, Ian T. Paulsen2, Apurba K. Bhattacharjee1, Mikeljon P. Nikolich1, Gary Myers2 and David L. Hoover1

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

1Department of Bacterial Diseases, Division of CD&I
Walter Reed Army Institute of Research, Washington, DC
503 Robert Grant Ave, Silver Spring, MD20910
2The Institute for Genomic Research, Rockville, MD 20850

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

This work was supported by the Department of Defense, JSTO/DTRA.

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The recent description of the complete genomes of the two most pathogenic species of Brucella opens the way for genome-based analysis of the antigenicity of their proteins. In the present paper, we describe a bench-level High-Efficiency Cloning and Expression System (HECES) that allow expression of large numbers of Brucella proteins based on genomic sequence information. Purified proteins are produced with high efficiency in a microarray format conducive to analysis of their seroreactivity against serum from immunized animals. This method is applicable at either small or large scale of protein processing. While it does not require robotics, the format is amenable to robotic implementation for all aspects of the process and subsequent analysis of protein characteristics. This method will allow selection of new reagents for diagnosis of brucellosis and development of vaccine against Brucella, an important zoonotic disease and biothreat agent.

15. SUBJECT TERMS

proteomics, Brucellar surface proteins, high throughput.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
Unclass	Unclass	Unclass		10	