RESEARCH REPORT
SWINE HEALTH

Title: Reverse Vaccinology and Genomics Towards Controlling Post-Weaning Diarrhea – **NPB #11-081**

Investigator: Timothy Johnson and

Institution: University of Minnesota

Co-Investigator: Richard Isaacson

Date Submitted: 5/1/2013

Industry Summary.

Post-weaning diarrhea (PWD) in pigs is a continuing problem for the swine industry. The goal of this study was to utilize a concept known as reverse vaccinology towards the identification of novel vaccine candidates to be used towards the future development of a vaccine providing heterologous protection against enterotoxigenic *E. coli* causing (ETEC)-caused PWD in swine. The objectives used here to achieve this goal included genomic comparisons of K88+ and F18+ *E. coli* strains implicated in PWD, identification of vaccine candidates using a reverse genetics approach, and preliminary examination of a subset of proteins with the greatest antigenic potential. We used this approach to comprehensively analyze the porcine *E. coli* genomes and identify more than 50 antigenic candidates and defining loci that differentiate porcine ETEC from non-pathogenic *E. coli* strains. We then screened nearly 300 pathogenic and non-pathogenic *E. coli* from commercial pigs for the presence of these genes to identify the most discriminatory set of antigens/genes. Finally, we cloned and expressed eight of these genes to demonstrate feasibility in the development of a subunit vaccine that broadly

targets porcine ETEC. This promising antigen subset represents proteins that will be further characterized and assessed for their ability to elicit a mucosal immune response in weaned pigs. The incorporation of multiple antigens into a recombinant attenuated *Salmonella* strain will ultimately enable heterologous protection against PWD-associated *E. coli* in a cost effective and reasonable manner for the swine industry.

Keywords: diarrhea, *Escherichia coli*, antigen, vaccine, genome, reverse vaccinology

Scientific Abstract.

Enterotoxigenic *Escherichia coli* (ETEC) is the causative agent of post-weaning diarrhea (PWD) in production pigs, which is a disease that continues to be a major cause of morbidity for the swine industry. The primary ETEC types implicated in PWD are well known and characterized and mostly include those that harbor K88 or F18 fimbriae. However, traditional approaches at a vaccine providing protection against PWD have not proven effective. Here, we utilized a global approach known as reverse vaccinology towards the identification of novel vaccine candidates. This approach utilized the first completed genome sequences of K88+ and F18+ porcine ETEC isolates, subsequent mining of their predicted proteins for suitable antigenic candidates, and assessment of the prevalence of the gene sequences encoding for suitable antigenic candidates among porcine *E. coli* collections. Using this approach, we identified approximately 50 antigenic candidates from the K88+ and F18+ genome sequences. We screened a collection of 300 porcine *E. coli* isolates for the presence of these candidates, narrowing this subset to eight genes that are significantly associated with PWD ETEC isolates and possess antigenic potential. This subset of antigens will be used in future studies aimed at developing a recombinant attenuated *Salmonella* vaccine that will elicit a mucosal immune response and protect against ETEC strains implicated in PWD.

Introduction.

Post-weaning diarrhea (PWD) is perhaps the most regularly-occurring disease of commercial pig farms, and is caused by enterotoxigenic *Escherichia coli*, or ETEC [1]. This disease most often affects pigs that are

weaning at approximately 3-4 weeks of age, and the disease starts a few days after lacteal protection stops. Protection of pigs and other animals from *E. coli* pathogens is a constant challenge, as these strains have proven highly adaptable under environmental pressures. This is especially true for ETEC and other types harboring transmissible plasmids encoding virulence factors [2]. In fact, there are reports of recent surges in the occurrence of PWD in commercial farms around the world, and the ETEC implicated in these cases are often $K88^+$ or $F18^+$ multidrug resistant strains. Because of the genome diversity of ETEC, it has been challenging to respond to outbreaks involving these strains, and to develop strategies to prevent and treat diseases caused by them.

Great effort has been put forth towards controlling PWD in pigs. Potential control measures sought have included treatment with antibiotics, bacteriophage, probiotics, and vaccines. Of these, vaccination currently holds the greatest promise for a long-term solution. Unfortunately, active immunization against PWD in pigs has been met with very mixed results [3]. The primary reasons for this include an inability to stimulate mucosal immunity, the complexities of PWD itself, and hypersensitivity or tolerance to antigens encountered earlier in life. Most of the vaccines against ETEC-caused PWD have targeted the K88 and F18 fimbrial components. This has proven ineffective towards providing comprehensive protection against PWD. Therefore, we sought to use a global approach allowing us to identify novel vaccine candidates that are universal among PWD strains. This approach is coined "reverse vaccinology." Reverse vaccinology is based on the concept that secreted or extracellular proteins are more accessible to antibodies and are therefore the best vaccine candidates [4]. With the availability of whole genome sequences and advanced computer prediction programs that predict the destination of proteins based upon sequence, vaccine candidates can be identified using a global, reverse approach.

In organisms such as *E. coli*, it is desirable to avoid targeting proteins possessed by commensal strains because they would invoke a response against the normal flora in the host. By screening both pathogen and commensal populations using PCR, the subset of vaccine candidates can be further refined to include only pathogen-specific sequences. The resulting subset of genes encodes predicted proteins are those with the greatest potential as vaccine candidates, are pathogen-associated, and can be further evaluated [4]. Here,

modern reverse vaccinology was applied to porcine ETEC in our initial efforts to identify alternative approaches towards a vaccine protecting against PWD in pigs.

Project Objectives.

1. **The *in silico* identification of ETEC-specific proteins with vaccine potential**.
2. **Refinement of identified antigenic candidates using population genetics and protein expression.**

Materials and Methods.

Bacterial strains. The 286 *E. coli* strains used in this study were acquired from porcine diagnostic samples sent to the University of Minnesota, College of Veterinary Medicine's Veterinary Diagnostic Laboratory, and fecal samples collected from healthy pigs in Minnesota and Iowa. Diagnostic samples were obtained from sick young pigs with diarrhea, enteritis, or edema disease and were presumed to be predominantly porcine ETEC strains. Samples included 100 isolates classified as "neonatal diarrhea" ETEC, or ND; 100 isolates classified as "post-weaning diarrhea" ETEC, or PWD; and 86 isolates of fecal origin from healthy animals.

Bacterial sequencing. A K88-positive and F18-positive strain were selected for whole-genome sequencing. These strains were submitted to the University of Minnesota Veterinary Diagnostic Laboratory after collection from farms in Minnesota in 2007 (UMNK88) and Iowa in 2006 (UMNF18), and both strains were implicated in regional outbreaks of PWD. The K88 strain was isolated from a pig of post-weaning age (approximately 4 weeks old) whose clinical signs included enteritis. It was identified as an enterotoxigenic *E. coli* based on the presence of enterotoxin genes detected via multiplex PCR by the diagnostic laboratory. In addition to the K88ac adhesin, this strain was positive for heat labile enterotoxin and heat stable enterotoxins A and B. It belonged to the O149 serogroup (determined by slide agglutination test, performed at Pennsylvania State University's *E. coli* Reference Center) which is the most frequently occurring ETEC serogroup in K88-positive cases of post-weaning diarrhea [1]. This strain was designated UMNK88. The F18 strain, designated UMNF18,

was positive for F18, STa, STb, and Stx2e. Like the K88-positive strain, its antimicrobial susceptibility test reflected that it is multidrug resistant. The F18-strain contained five plasmids.

The whole genomes of these strains were sequenced to a depth of 30x coverage using pyrosequencing on Roche 454 GS-FLX with Titanium chemistry. Two libraries were generated for each genome, consisting of a shotgun library and a 3-kb paired-end library. Sequencing was performed on a Titanium plate. Sequence reads were assembled into contigs with Newbler Assembler (454 Life Sciences) to resolve full chromosomal and plasmid sequences. Homopolymers were examined for sequence quality and manually edited in SeqMan (DNASTAR, Lasergene). The genome and plasmids were fully assembled and verified for accuracy using PCR.

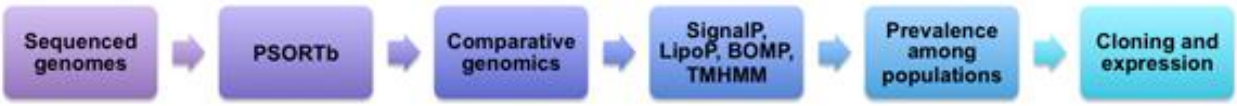
Gene prediction and annotation. Annotation was automated using previously annotated *E. coli* genomes to orthologous genes, and then manually curated. Annotation was carried out using the genome viewers RAST (Rapid Annotation using Subsystem Technology; http://rast.nmpdr.org/) and Artemis [5]. Coding sequences were predicted using the gene prediction program Glimmer3 [6], then selectively checked using BLAST [7]. Plasmids were also curated manually. The annotated genome sequence of UMNK88 has been deposited in NCBI (accession numbers: CP002729 for the UMNK88 complete chromosome; CP002730, CP002731, CP002732, CP002733, and HQ023862 for the UMNK88 plasmids pUMNK88_K88, pUMNK88_IncI1, pUMNK88_Ent, pUMNK88_Hly, and pUMNK88_IncA/C). Draft sequence accession numbers for UMNF18 are also deposited in NCBI (CP002890 for the chromosome, and CP002891, CP002892, CP002893, CP002894, and CP002895 for UMNF18 plasmids pUMNF18_IncI1, pUMNF18_87, pUMNF18_IncFV, pUMNF18_P7, and pUMNF18_32). These genomes are described in detail in a recent publication [8].

***In silico* analysis.** The predicted proteins in UMNK88 and UMNF18 were analyzed using the PSORTb 3.0 subcellular localization prediction program to identify putative outer membrane and exported proteins [9]. These proteins were compared to the non-pathogenic, commensal lab K-12 strain MG1655 using protein BLAST. Proteins that matched (>90% amino acid homology) between the strains were excluded from further

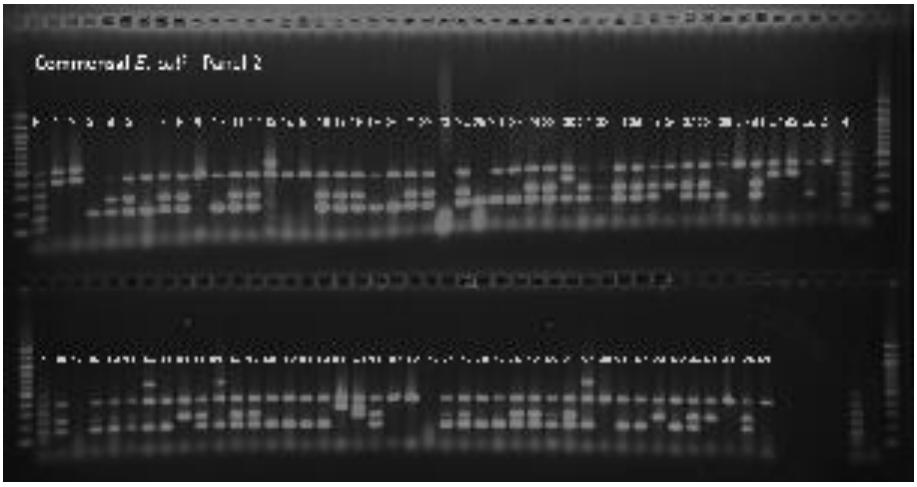
analysis, resulting in the identification of predicted outer membrane or exported proteins that were unique to UMNK88 and/or UMNF18.

The remaining proteins were analyzed according to predicted immunogenicity and suitability as vaccine targets. Using SignalP 3.0 [10] and LipoP [11], proteins were identified that contained predicted signal peptides. Predicted beta-barrel structures, which indicate an association with the bacterial outer membrane, were identified by BOMP [12]. TMHMM (Center for Biological Sequence Analysis, Technical University of Denmark) was used to predict transmembrane alpha-helices [13]. While the presence of alpha-helix structures suggests association with a cell membrane, which is favorable for vaccine candidate selection, proteins with multiple transmembrane regions have been demonstrated to pose problems in the cloning and expression steps necessary for vaccine development [4]. For this reason, we excluded any proteins with 3 or more predicted transmembrane alpha-helices. The criteria for antigenic candidates was as follows: the presence of beta-barrels or predicted signal peptides, with no more than one transmembrane alpha-helix.

Gene prevalence determination. Genes of interest representing antigenic candidates or genes of interest were



sought among the porcine *E. coli* collections using multiplex PCR. Initially, runs with different annealing temperatures were completed to determine the ideal annealing time for each panel. The final PCR thermocycler protocol was as follows: 10 minutes at 95°C; 25 cycles of: 94°C for 30 seconds, 58-60°C 30 seconds, 72°C for 1 minute; then 7 minutes at 72°C and holding at 4°C. The PCR reaction contained 2 µl boiled bacterial template, 2 µl of pooled primer mix, 15.5 µl water, 2.5 µl 5x PCR buffer with MgCl2, 1 µl dNTP, and 2 µl Taq polymerase for 25 µl total volume used per reaction. After amplification, samples were electrophoresed in 2% agarose gels run at 150V for 40 min, stained with ethidium bromide, and visualized under ultraviolet light to detect PCR amplification products. A 100 bp size standard was used to confirm appropriate PCR product size. An example of multiplex PCR results is seen below.

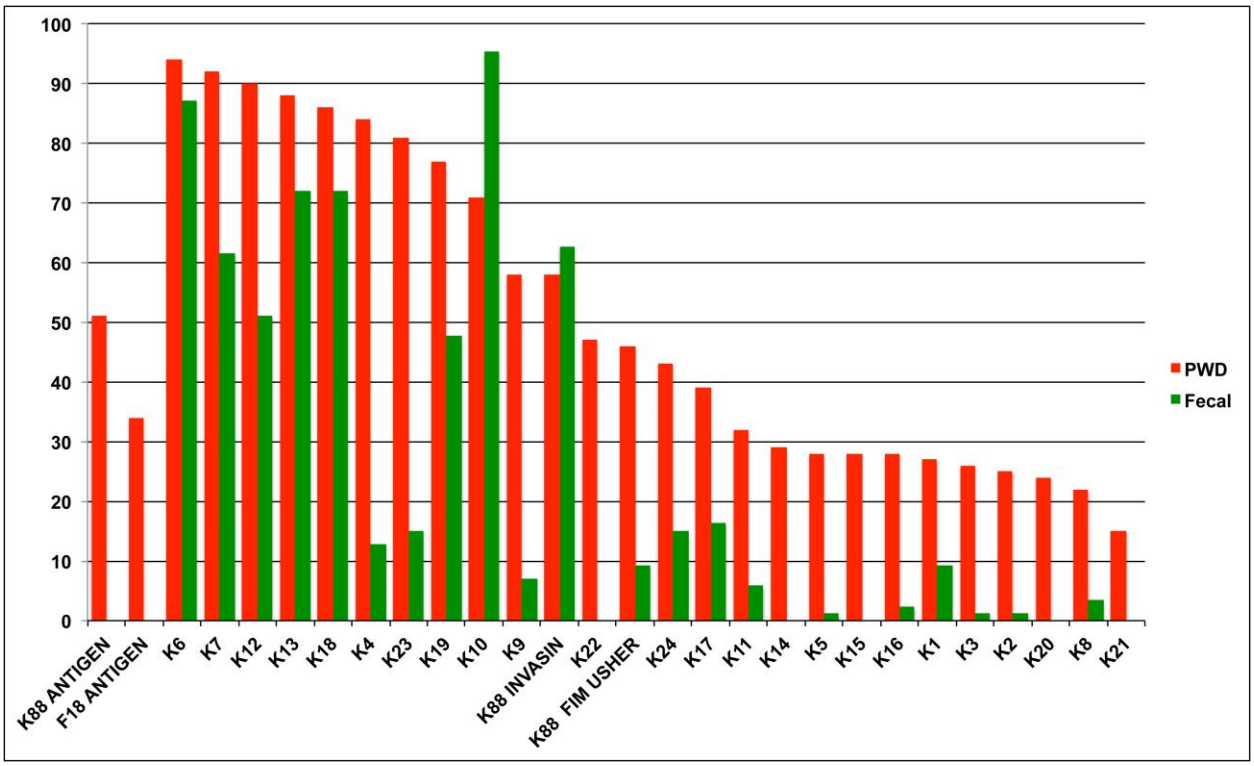


Protein expression. Reverse vaccine candidate genes (Table 1) were cloned into pQE-70 Vector (Qiagen) and electroporated into M15 cells. Transformation cultures were plated on LB Amp_{100} plates and incubated overnight at 37C. The presence of insert was confirmed with PCR and sequencing. The 6xHis-tagged proteins were expressed following the QIAexpressionist (QIAGEN) protocol for rapid screening of small expression cultures. Samples were collected at T=0, (no-IPTG), T=3 hrs (IPTG and no-IPTG) and T=16hrs(IPTG). Crude lysates of IPTG induced and non – induced cultures were analyzed on SDS-PAGE gel (PIERCE). Samples that did not show expression of the 6xHis-tagged proteins in crude lysates were purified with MagneHis purification kit (Promega) and analyzed by Western Blotting using SuperSignal West HisProbe Kit (Pierce).

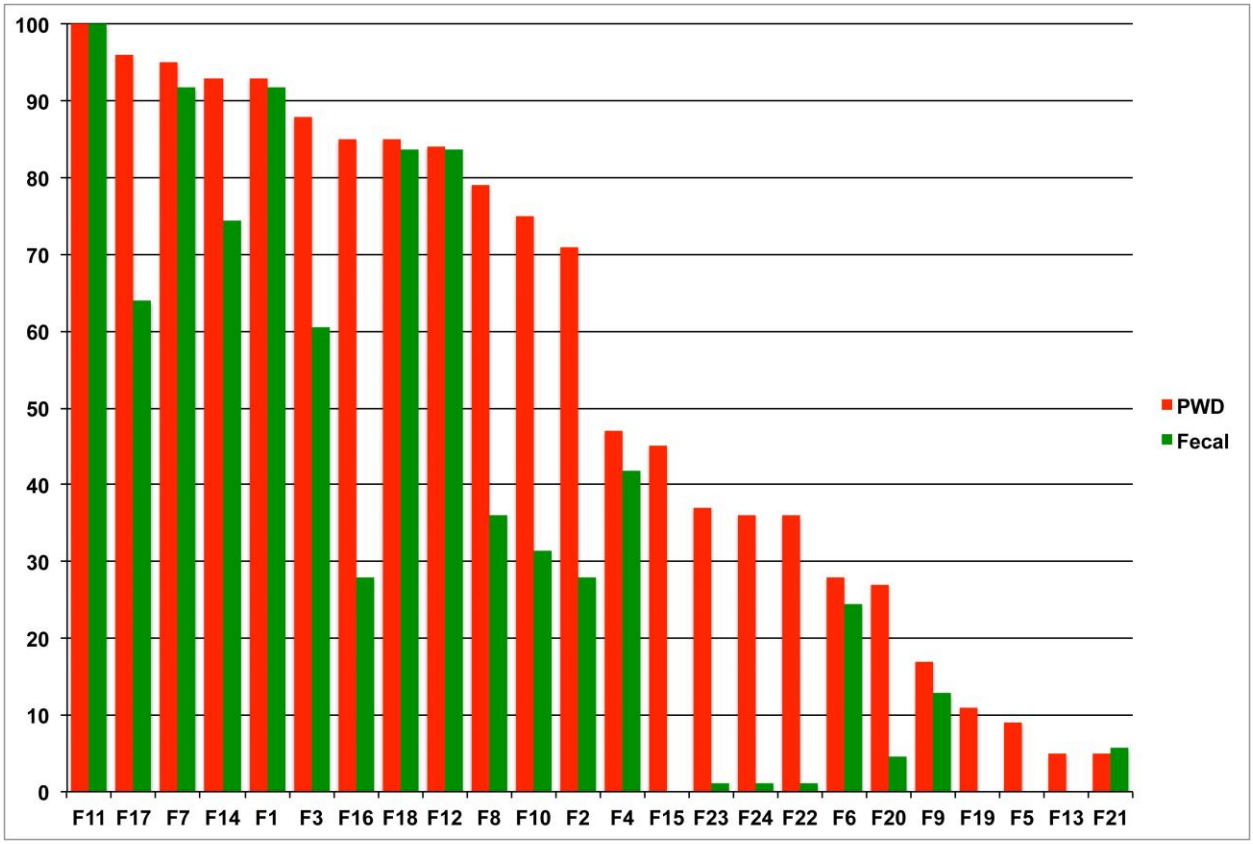
Results.

Using the reverse vaccinology and comparative genomics approaches, 50 targets were identified that were further screened for their presence among nearly 300 porcine-source isolates. The figure below shows the prevalence of UMNK88-associated antigens among PWD and fecal *E. coli* collections. Compared to fecal isolates from healthy pigs, a number of antigens were identified that were significantly more prevalent among PWD isolates than fecal isolates (Figure below). Using a cut-off of at least 50% prevalence and significant difference between populations, the following antigenic candidates were identified: K17, K12, K4, K23, K19, and K9. K12 is a putative outer membrane protein of unknown function; K12 is a putative effector protein of a type VI secretion system; K4 is an AIDA-1-like adhesion with similarity to antigen 43; K23 is the HlyA

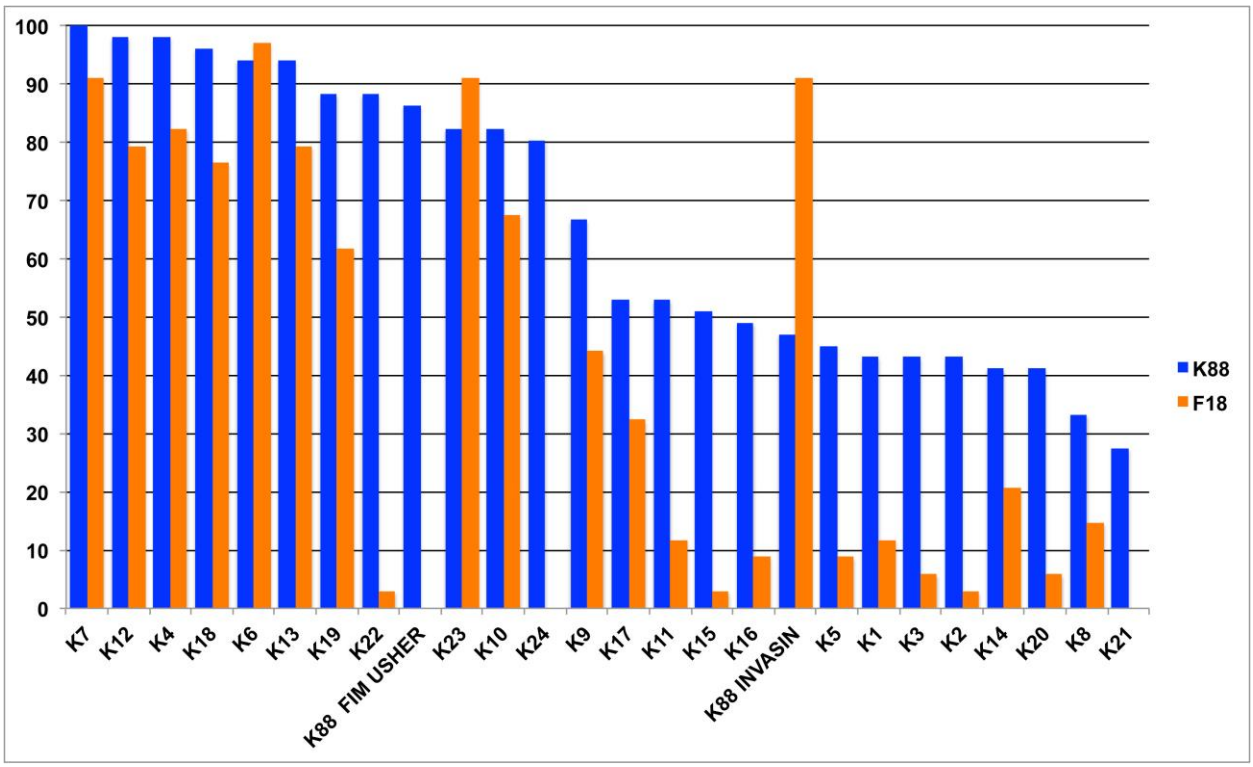
hemolysin; K19 is a glucoside-specific outer membrane porin; and K9 is DNase. Of these K4 and K23 had the most discrimination between PWD and fecal isolates.



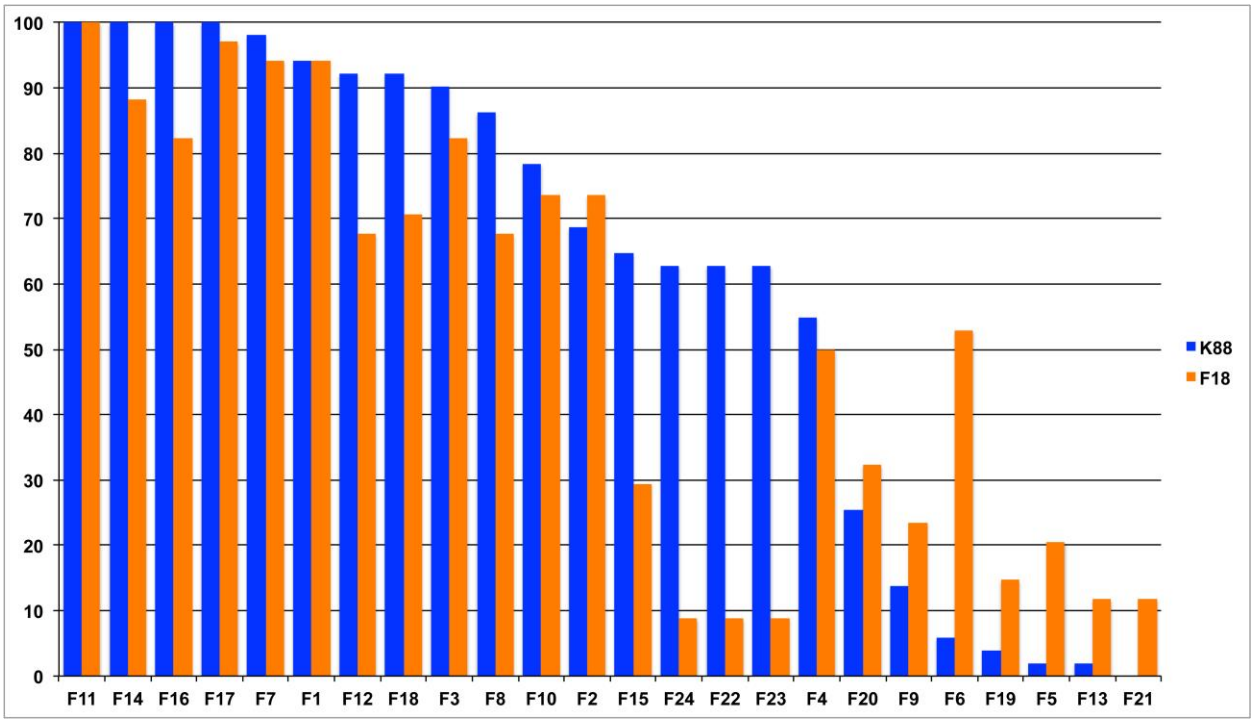
We then examined the presence of UMNF18-associated antigenic candidates among the same collection of bacterial isolates (Figure below). Using the same criteria, the following antigenic candidates were identified: F17, F3, F16, F8, F10, and F2. F17 is the GspD protein of a type II secretion system; F3, F16, and F10 are putative autotransporters with similarity to antigen 43; and F8 and F2 are outer membrane porins. Of these, F16, F8, F10, and F2 had the most discrimination between PWD and fecal isolates.



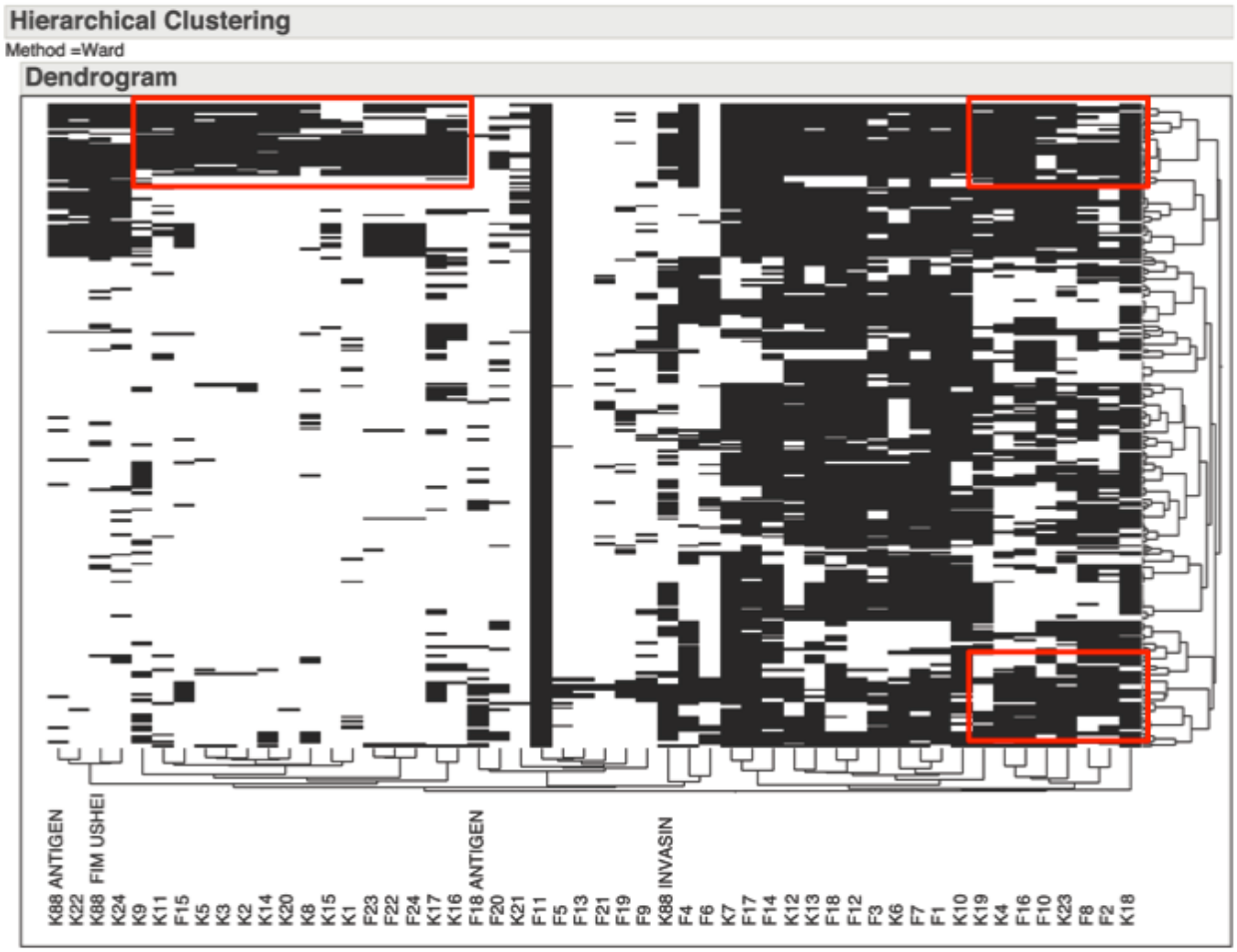
We further examined the presence of UMNK88-associated antigens relative to their presence in F18+ and K88+ PWD isolates (Figure below). Considering the previously identified antigenic candidates that were significantly associated with PWD isolates (K17, K12, K4, K23, K19, and K9), several of these candidates were highly prevalent among F18+ and/or K88+ isolates. Among these were K12, K4, and K23, present at >75% among both F18+ and K88+ isolates, and K19, present among greater than 75% of K88+ isolates and >60% of F18+ isolates.



We also examined the presence of UMNF18-associated antigenic candidates among F18+ and K88+ PWD isolates (Figure below). Considering the previously identified antigenic candidates that were significantly associated with PWD isolates (F17, F3, F16, F8, F10, and F2), several of these candidates were highly prevalent among F18+ and/or K88+ isolates. F17, F3, and F16 were present in >75% of both F18+ and K88+ isolates, while F8, F10, and F2 were present among both populations at greater than 60% prevalence.

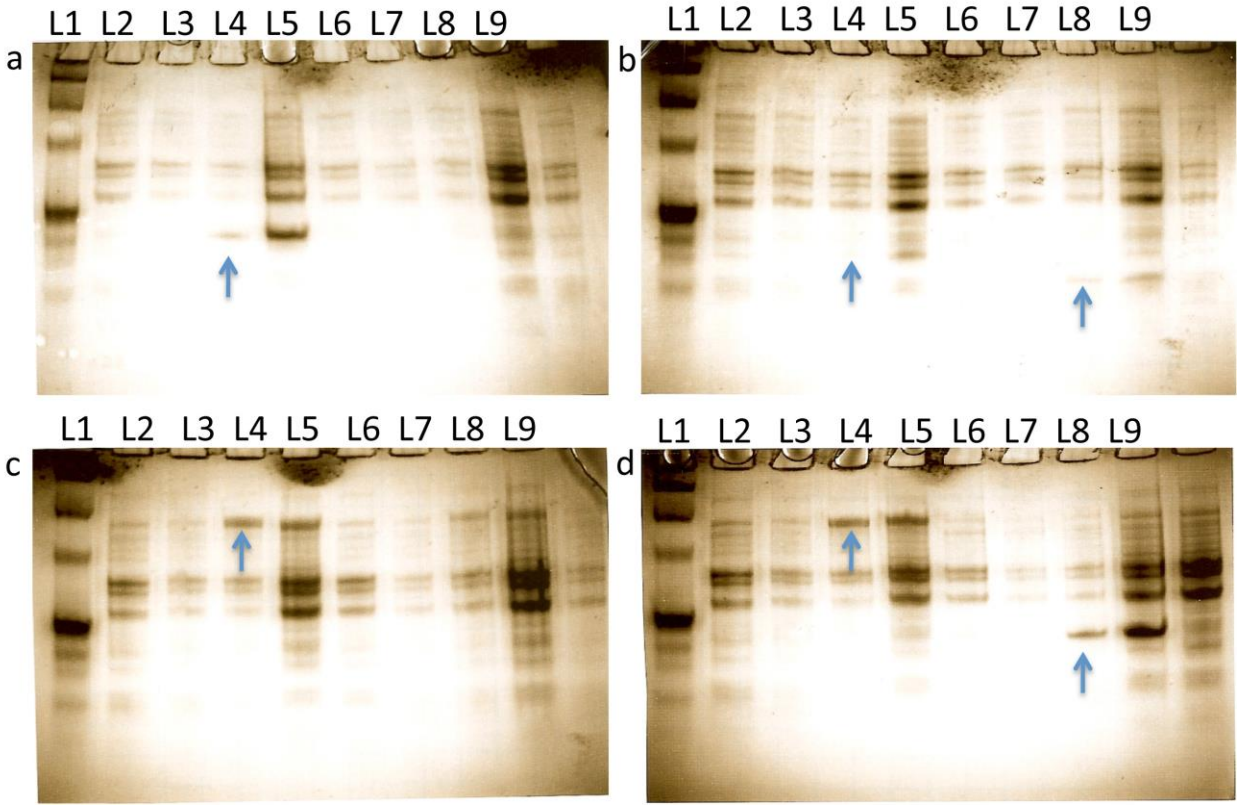


Hierarchical clustering was then performed to examine patterns of antigenic candidate prevalence among the isolates examined (Figure below). Several clusters were identified that would provide enhance coverage against most of the F18+ and K88+ isolates studied (boxed in red). One cluster (top left) contained antigenic candidates that were present in most K88+ isolates but were absent in most non-K88+ isolates. This cluster included K11, F15, K5, K3, K2, K14, K20, K8, K15, K1, F23, F22, F24, K17, and K16. Particularly interesting is that some of these antigenic candidates were plasmid-associated genes. For example, F22-F24 are IncFV conjugal transfer proteins. Also, F15 is a virulence-associated membrane protein found in the bacterial chromosome. K2 is an invasion-like outer membrane protein found in the bacterial chromosome. A second cluster of antigenic candidates were found in both K88+ and F18+ isolates, and included K10, K19, K4, F16, F10, K23, F8, F2, and K18. Many of these proteins were identified above as being strongly associated with PWD isolates as compared to fecal isolates, and several were highly prevalent among K88+ and F18+ isolates.



As proof-of-concept for expression of antigenic candidates, eight candidates were cloned and expressed in a His-tagged vector in *E. coli* (Figure below). We were able to successfully express all of the attempted proteins in E. coli strain M15. We also examined cell fractions of these purified proteins, and all were localized to the outer membrane fraction.

Figure below. SDS-PAGE gel analysis of crude lysates from IPTG induced and non-induced cultures. Lane 1 = ladder; Lanes 2 and 6 = time zero; Lanes 3 and 7 = time 3 hours, no IPTG; Lanes 4 and 8 = time 3 hours, IPTG, Lanes 5 and 9 = time 16 hours, IPTG. a) Lanes 1-5 = F21; Lanes 6-9 = F23; b) Lanes 1-5 = K21; Lanes 6-9 = K22; c) Lanes 1-5 = F3; Lanes 6-9 = F10; d) Lanes 1-5 = K4; Lanes 6-9 = mouse DHFR positive control.



Discussion.

The purpose of this project was to identify novel antigenic candidates among PWD-associated ETEC in weaned pigs. A great deal of work has already been performed towards identifying suitable vaccines targeting PWD ETEC, and the purpose of this study was not to discount but instead to supplement those efforts to gain greater knowledge about the distribution of potential conserved antigens among PWD ETEC. For example, numerous

studies have identified and refined approaches targeting the K88 and F18 fimbrial antigens and ETEC enterotoxins [14-22]. Advances are being made towards the development of an effective vaccine. However, these efforts have largely ignored the remainder of the porcine ETEC genome. Our work here has described the genome sequences of K88+ and F18+ porcine ETEC and performed comparative genome analyses of the shared and unique components of these genomes [8]. We then sought to identify novel antigens that may be useful in the development of future subunit vaccines, acknowledging that previously identified antigens would also undoubtedly be useful. This study demonstrated that a combinatorial strategy could be used to develop a subunit vaccine that provides broad coverage against the diverse array of porcine ETEC capable of causing PWD.

Some noteworthy antigens identified included several autotransporter-like proteins with similarity to proteins previously described as "antigen 43." These proteins have been shown to play roles in colony morphology and autoaggregation [23] and cell-to-cell interactions in biofilms [24]. They are phase variable proteins existing in multiple alleles [25]. These antigens have been demonstrated to afford protection against human ETEC challenge, therefore they have potential as subunit candidates among PWD ETEC [26]. Additionally, several novel fimbrial-associated proteins and outer membrane proteins of unknown function were identified here that could supplement autotransporter proteins in the development of a multiple subunit vaccine. These, combined with previously characterized antigens, could afford broad protection against PWD ETEC.

Because vaccines targeting the fimbrial components of common PWD-associated strains have not provided comprehensive protection against the full repertoire of porcine ETEC causing this disease, we hypothesize that a more effective approach is to utilize reverse genomics to identify a subset of proteins that will elicit an immune response against most of these strains. Usually, these subsets of proteins are subsequently used to design true protein subunit vaccines. This approach has worked well for many types of bacterial disease. However, since a primary problem with PWD protection involves the stimulation of a mucosal immune response, logic would suggest that a protein subunit vaccine won't work well towards this endeavor. Therefore, we believe a more effective approach will be the integration and expression of these proteins in a bacterium known to effectively stimulate mucosal immunity. Certainly, one such candidate for this approach is

S. Choleraesuis strain CS54, which is licensed for use as a live vaccine for pigs in the United States. Recently, it has been shown that *E. coli* antigens can be incorporated into the chromosome of this strain and ultimately expressed. In the future, we hope to use such an approach with our subset of proteins to ultimately protect against the diverse collection of strains causing PWD in weaned pigs.

Table 1. Characteristics of antigenic candidates identified in this study. Designations: "K" = K88 genome origin, "F" = F18 genome origin.

				Locus tag:	
Designation	**F primer**	**R primer**	**Protein description**	**UMNK88**	**UMNF18**
K1	CCTGACTGGACAGCAACAGA	TCCCATTCGAAACCATAAGC	TonB-dependent vitamin B12 receptor BtuB	AEE59297	
K2	AGTACACAGGCCGGTAATGC	GTCTGTTTGCTTCCCGGATA	Putative outer membrane protein/invasin	AEE59544	AEE60055
K3	GCACGGTGACAGACCCTTAT	GACTTACGGGTGCCACTGAT	Putative exported protein from ICE	AEE59513	
K4	CAACACCACTGTCGGACATC	GAGGTTGGTGGTGGTCAGTT	Antigen 43 precursor, AIDA-1 like adhesin	AEE55888	AEJ55945
K5	GGAGCAGCCAGAAACAGAAC	CAGTTGCTGGCAAAGGTGTA	MltA-interacting protein MipA	AEE59489	
K6	TGCCAAAACACTTCACCAAA	GGCGCAATAATTTGACGATT	Fimbrial usher protein	AEE54684	AEJ54677
K7	GGCTGGACTCAGAAATGGAA	AACTGCATTAATCCGGCATC	General secretion pathway protein GspD	AEE58253	AEJ58346
K8	CCGGCTATGAGAAAAAGCTG	GTGACTCCGTGGGATAAGGA	Bifunctional enterobactin receptor/adhesin protein	AEE55853	
K9	CTGAACAAGCCGGTTATCGT	GGCGAAGTTATCGGTGGTTA	Dnase TatD	AEE55982	AEJ55896
K10	CTGATGCGTTACCACACTGG	ATGTCCCGTAAATCCAGCAG	Putative hemolysin activator protein	AEE55873	
K11	CGACGGTGATGTTGATAACG	AGCATCACCTACAGCCCAAC	Putative flagellin structural protein	AEE55068	
K12	AGGTTGCTCAACGCTTGATT	CCGCTTTACTGGTTCGGTAG	Putative type VI secretion system effector protein	AEE54880	AEJ54872
K13	TCCGATTCTCAACAACACCA	GTTCCATCCAGGCAAAGAAA	Putative type I fimbrial protein	AEE55388	
K14	TGATGATTGCACTGGTCAGC	CAGGTTTCACTCATCTGGGTTT	Complement resistance protein TraT	AEE55823	
K15	AGTTCTTCGCTGAGGT	CTTCGACGGCTTAACGCTA	Putative phage tail fiber protein	AEE57668	

	GCAT	C			
K16	TAGCAAATCTGATCGCACCA	ACACGGTCAGCGTCAACATA	Putative outer membrane protein	AEE57023	
K17			Putative outer membrane protein	AEE57023	
K18	AACGCATTACAGCGATACCC	CGTTTTCCCCTTGTTGTCAG	Putative outer membrane precursor Lom	AEE56572	
K19	TGAAAGTGGCGCTCATAGTG	TCACACCCTGCTGAGTTTTG	Glucoside-specific outer membrane porin BglH	AEE57668	
K20	AGTACACAGGCCGGTAATGC	TTTGCTTCCCGGATAAACAC	Putative invasin on Hly plasmid	AEE59544	AEE60055
K21	TCAACAATTGCTGTGGCATT	CCTTCAGGGGAAGTGAAGGT	Putative fimbrial protein on K88 plasmid	AEE59840	
K22	GGCGGTGTTGCTAAAAGTCT	ACTCGCCTGACCTTCAGAGA	Putative fimbrial protein on K88 plasmid	AEE59842	
K23	CTTTTTATCTGCCGGCTCAG	CTGCTCCCCCAATCAACTTA	Hemolysin HlyA on Hly plasmid	AEE60090	
K24	TAGGTTACTGGCCCGGTATG	CTTTGGGAACGCCTGTTAAA	Putative fimbrial usher protein on K88 plasmid	AEE59839	
F1	TGCCAGGCGTTTTTAGAATGACTA	GGGGCGAAGTGGAAATGCTA	Putative fimbrial usher protein	AEE54684	AEJ54677
F2	ACCTGCGGCCCAAACTTCTG	CCGCGCTGAATCTCAAGGTTC	Outer membrane porin protein NmpC		AEJ55180
F3	CCCCCGGCATTAACCGTAAA	AAACCGCCTGGGGCATTTC	Putative autotransporter, antigen 43-like		AEJ55945
F4	CCCCGGGTTTAGTTTGGGACTTA	CACTGACTACTGCGCTGGCTGTTA	Putative autotransporter	AEE56083	AEJ56083
F5	CTGCGGGAATGTCACAAACTGAA	ACGCCGTTATTCTCGCTGTTGA	Conserved hypothetical protein, tail fiber protein		AEJ56153
F6	TTTCTGGCGATGCGAATACAGTG	TCTCTGCGCTGCCATCAAATAAAT	Putative autotransporter		AEJ56442
F7	CTCGCGTTCGGGTTGATAACATT	CCTGCGCGTTGCTACTCCACT	Putative autotransporter	AEE56524	AEJ56560
F8	TGCCGGGCACTACTTCTCCTCTA	TGCCCATACTTCAGCATTTTGACC	Outer membrane porin protein NmpC		AEJ56947
F9	GCCCGAAGGCTCTAA	TCAGCGGAAGCCCAGGAA	Outer membrane protein OmpN		AEJ57043

	TCAGGAA	TC			
F10	ATACCGGCGGGCAATGGATA	CTCACCGCCGTTCAGAGTGGT	Putative autotransporter, antigen 43-like	AEE58235	AEJ57110
F11	GCGCGGAATGACGGAACAG	ATGTCGCGGGGAAAATGAAAAT	Multidrug resistance outer membrane protein MdtQ	AEE57251	AEJ57277
F12	CTTATCCGCCCGGTTGTCATACT	CGCGGTCGCGTTTGATGT	Outer membrane usher protein, PapC-like	AEE57455	AEJ57557
F13	CTCGAAGAAGCCGGAATAACACC	CTGCGGGCCGTCTGAAGTC	Conserved hypothetical protein, tail fiber protein		AEJ57808
F14	CAGCGGTTACGCACGGAATCT	CGCCTGCATGGGCATCTG	Outer membrane protein Lom	AEE56572	AEJ57943
F15	CGCCAGCCGCATACAGTGAC	CCCCGGTTAGTATTATGGGCAGTT	Virulence membrane protein pagC		AEJ58306
F16	CTGCCGGGGGAAGTCTGAGT	CTGCGCAGGTACCAGTCTTCATC	Putative autotransporter, antigen 43-like	AEE55888	AEJ55945
F17	GCGCGTCCGGTGAGCATA	CTGGGCGGAAGAAGCCACTT	Type II secretion protein GspD	AEE58253	AEJ58346
F18	CACTACGCGCCAGTTATCAGCATA	ACCCCAGCGTAAAGCAAGTTTGT	Putative outer membrane usher protein YqiG	AEE58333	AEJ58430
F19	TGCGCTCCCACTTAATGCTGTAA	TGCCGCCCTTGTAGATTCGTC	Phage tail fibre repeat family protein		AEJ59650
F20	ACTGCCTGCCGGGCTGTTA	CTTGCGGTGGGATACGATTTTTAT	Tia invasion determinant		AEJ59746
F21	AATGCGCCTTCGGGATTACTG	GAAGCCGCAAGATGGACAGAAC	IncX type IV secretion system protein VirB5		AEJ60272
F22	TGCCGGCGTTAAAAGCAACTT	GACCCGACCGGGCTGATG	IncFV conjugal transfer protein TraV		AEJ60141
F23	AGCGCGGTTTACTCCTGGAATAA	CTCCCCGCCGCAAATCTG	IncFV conjugal transfer protein TraN		AEJ60147
F24	CCGGAAAGCGCATATTACTGGATA	CTAACGCCGCCCCCATTG	IncFV conjugal transfer protein TraT		AEJ60153

References Cited.

1. Fairbrother JM, Nadeau E, Gyles CL (2005) *Escherichia coli* in postweaning diarrhea in pigs: an update on bacterial types, pathogenesis, and prevention strategies. Anim Health Res Rev 6: 17-39.
2. Nagy B, Fekete PZ (2005) Enterotoxigenic *Escherichia coli* in veterinary medicine. Int J Med Microbiol 295: 443-454.
3. Haesebrouck F, Pasmans F, Chiers K, Maes D, Ducatelle R, et al. (2004) Efficacy of vaccines against bacterial diseases in swine: what can we expect? Vet Microbiol 100: 255-268.
4. Mora M, Donati C, Medini D, Covacci A, Rappuoli R (2006) Microbial genomes and vaccine design: refinements to the classical reverse vaccinology approach. Curr Opin Microbiol 9: 532-536.
5. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA (2012) Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28: 464-469.
6. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27: 4636-4641.
7. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.
8. Shepard SM, Danzeisen JL, Isaacson RE, Seemann T, Achtman M, et al. (2012) Genome sequences and phylogenetic analysis of K88- and F18-positive porcine enterotoxigenic *Escherichia coli*. J Bacteriol 194: 395-405.
9. Gardy JL, Spencer C, Wang K, Ester M, Tusnady GE, et al. (2003) PSORT-B: Improving protein subcellular localization prediction for Gram-negative bacteria. Nucleic Acids Res 31: 3613-3617.
10. Zhang Z, Henzel WJ (2004) Signal peptide prediction based on analysis of experimentally verified cleavage sites. Protein Sci 13: 2819-2824.
11. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, et al. (2003) Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12: 1652-1662.
12. Berven FS, Flikka K, Jensen HB, Eidhammer I (2004) BOMP: a program to predict integral beta-barrel outer membrane proteins encoded within genomes of Gram-negative bacteria. Nucleic Acids Res 32: W394-399.
13. Chen Y, Yu P, Luo J, Jiang Y (2003) Secreted protein prediction system combining CJ-SPHMM, TMHMM, and PSORT. Mamm Genome 14: 859-865.
14. Ruan X, Zhang W (2013) Oral immunization of a live attenuated *Escherichia coli* strain expressing a holotoxin-structured adhesin-toxoid fusion (1FaeG-FedF-LTA(2):5LTB) protected young pigs against enterotoxigenic *E. coli* (ETEC) infection. Vaccine 31: 1458-1463.
15. Hur J, Lee JH (2012) Comparative evaluation of a vaccine candidate expressing enterotoxigenic *Escherichia coli* (ETEC) adhesins for colibacillosis with a commercial vaccine using a pig model. Vaccine 30: 3829-3833.
16. Liu M, Zhang C, Mateo K, Nataro JP, Robertson DC, et al. (2011) Modified heat-stable toxins (hSTa) of enterotoxigenic *Escherichia coli* lose toxicity but display antigenicity

after being genetically fused to heat-labile toxoid LT(R192G). Toxins (Basel) 3: 1146-1162.
17. Ruan X, Liu M, Casey TA, Zhang W (2011) A tripartite fusion, FaeG-FedF-LT(192)A2:B, of enterotoxigenic *Escherichia coli* (ETEC) elicits antibodies that neutralize cholera toxin, inhibit adherence of K88 (F4) and F18 fimbriae, and protect pigs against K88ac/heat-labile toxin infection. Clin Vaccine Immunol 18: 1593-1599.
18. Zhang C, Zhang W (2010) *Escherichia coli* K88ac fimbriae expressing heat-labile and heat-stable (STa) toxin epitopes elicit antibodies that neutralize cholera toxin and STa toxin and inhibit adherence of K88ac fimbrial *E. coli*. Clin Vaccine Immunol 17: 1859-1867.
19. Zhang W, Francis DH (2010) Genetic fusions of heat-labile toxoid (LT) and heat-stable toxin b (STb) of porcine enterotoxigenic *Escherichia coli* elicit protective anti-LT and anti-STb antibodies. Clin Vaccine Immunol 17: 1223-1231.
20. Zhang W, Zhang C, Francis DH, Fang Y, Knudsen D, et al. (2010) Genetic fusions of heat-labile (LT) and heat-stable (ST) toxoids of porcine enterotoxigenic *Escherichia coli* elicit neutralizing anti-LT and anti-STa antibodies. Infect Immun 78: 316-325.
21. Yahong H, Liang W, Pan A, Zhou Z, Wang Q, et al. (2006) Protective immune response of bacterially-derived recombinant FaeG in piglets. J Microbiol 44: 548-555.
22. Verdonck F, Cox E, Van der Stede Y, Goddeeris BM (2004) Oral immunization of piglets with recombinant F4 fimbrial adhesin FaeG monomers induces a mucosal and systemic F4-specific immune response. Vaccine 22: 4291-4299.
23. Henderson IR, Meehan M, Owen P (1997) Antigen 43, a phase-variable bipartite outer membrane protein, determines colony morphology and autoaggregation in *Escherichia coli* K-12. FEMS Microbiol Lett 149: 115-120.
24. Danese PN, Pratt LA, Dove SL, Kolter R (2000) The outer membrane protein, antigen 43, mediates cell-to-cell interactions within *Escherichia coli* biofilms. Mol Microbiol 37: 424-432.
25. Roche A, McFadden J, Owen P (2001) Antigen 43, the major phase-variable protein of the *Escherichia coli* outer membrane, can exist as a family of proteins encoded by multiple alleles. Microbiology 147: 161-169.
26. Harris JA, Roy K, Woo-Rasberry V, Hamilton DJ, Kansal R, et al. (2011) Directed evaluation of enterotoxigenic *Escherichia coli* autotransporter proteins as putative vaccine candidates. PLoS Negl Trop Dis 5: e1428.