RESEARCH REPORT
PORK SAFETY

Title: Identifying genes associated with *Salmonella* shedding to increase pork safety through improved genetic resistance - **NPB #08-034** revised

Investigator: Christopher K Tuggle[1]

Co-Investigators: Shawn Bearson[2] Jack Dekkers[1], Annette O'Connor[3], James McKean[3]

Graduate students: Jolita Uthe[1,4], Long Qu[1,5]
Undergraduate students: Yasi Rodriguez[1], Stephen Robinson[1]

Institution: [1]Department of Animal Science, Iowa State University, Ames, IA
[2] USDA, ARS, National Animal Disease Center, Ames, IA
[3]Department of Veterinary Microbiology and Production Animal Medicine, Iowa State University, Ames, IA
[4]Interdepartmental Genetics Graduate Program, Iowa State University, Ames, IA
[5]Department of Statistics, Iowa State University, Ames, IA

Industry summary

The long-term goals of this project are to use genomics tools to identify genetic variants that are associated with decreased *Salmonella* shedding on farms; and to use these variants to select for animals that shed fewer bacteria and so are less likely to cause abattoir carcass contamination. Thus, the aim is to directly improve pork safety, as well as decrease other on-farm contamination by contaminated fecal matter that may be used as fertilizer. To achieve the long term goals, we worked on 4 specific objectives: 1) generate a large bank of *Salmonella* fecal shedding data and matched DNA samples using US industry pigs; 2) create DNA tests for genetic variation at candidate genes, selected based on our previous data on porcine gut lymph node and whole blood gene expression patterns in response to *Salmonella* infection; 3) perform genotyping using new and

established DNA tests on US industry *Salmonella*-positive–control groups; and 4) test associations of genotypes with *Salmonella* incidence and integrate data from epidemiology and genomics projects to understand *Salmonella* contamination.

For objective 1, during a 2 year period a total of 462 belly flap tissue samples have been collected that include 163 pigs that shed *Salmonella* in their feces 7 days before marketing as well as 299 farm-mate negative controls matched for farm-visit cohort where a positive sample was found. Shedding of *Salmonella* was determined by qualitative fecal testing. Objective 1 was performed in collaboration with a large epidemiology study funded by the USDA-NRI Competitive Grants Program. We emphasize here the extraordinary contribution of this USDA-funded project; it would not have been possible to collect such a large number of *Salmonella* positive field samples within the current funding structure for standard NPB projects . In our laboratory, the DNA from *Salmonella*-positive and negative-control pigs belly flap tissue have been isolated and was used for genotyping. We also used other 3 porcine populations that we created using prior NPB funding or current funding from the USDA-NRI Functional Genomics program, such as the NADC-40 pig and NADC-77 pig *Salmonella* Typhimurium challenge populations; and the 228 pig IAH-Compton *Salmonella* Choleraesuis challenge population kindly provided by the Pig Improvement Company. In objective II we selected candidate genes for analysis of genetic variants (single nucleotide polymorphisms, SNPs). As candidates we selected genes that, based on our previous research, are involved in the porcine response to *Salmonella* infection. For objective III during the 1st-year period we genotyped 29-31 SNPs across the 4 populations and identified 18 SNPs with minor allele frequency (MAF) of 15% or higher in at least 2 populations. During the 2nd year period, we have genotyped a set of 31 additional SNPs across our pig populations and identified 10 new genetic variants with MAF of 15% or larger in at least 2 populations. In all, during the two year period we selected and genotyped across the 4 populations 62 SNPs in 53 genes that are involved in porcine response to *Salmonella* infection and confirmed 28 SNPs segregating in at least 2 populations. This output exceeded the original proposed goals of 10 segregating SNPs per year of funding. To genotype this larger number of SNPs, we adopted a Sequenom technology due to its increased throughput and reduced cost. In objective IV, we have newly applied a statistical approach to measuring *Salmonella*

shedding/tissue colonization levels over the 3-week period after infection. Statistical analyses revealed several SNPs associated with *Salmonella* fecal shedding or tissue colonization in pigs. Fecal shedding associated SNPs were CCT7 #3 (p=0.041), GNG3 (p=0.029), PGD (p=0.047) and HP #2 (p=0.0002) in the field population, AMT (p=0.005) in the NADC-40 pigs population and PGD (p=0.001) in the NADC-77 pigs population. In the IAH-Compton population three SNPs in the ACP2 gene were associated with *Salmonella* burden in spleen (ACP2#1, p=0.013, ACP2#2, p=0.026, ACP2#3, p=0.033); and in the NADC-40 pig population, a SNP in the EMP1 (p=0.002) gene was associated with bacterial load in ileo-cecal lymph node. The genotyping and statistical analysis data identified genetic markers that are potentially useful for selecting animals that shed fewer bacteria and are less likely to cause pen-mate *Salmonella* contamination in a farm and on a slaughter plant. Additional validation of these SNPs should be performed, but this project has provided novel information for the pig industry to implement in strategies for selecting pigs with reduced shedding and/or disease susceptibility.

Contact: Christopher Tuggle, 515-294-4252, email: cktuggle@iastate.edu.

Abstract

Salmonella shed from colonized swine can contaminate: slaughter plants and pork products during meat processing (Hurd et al, 2002, Appl Environ Microbiol, 68:2376-2381); edible crops when swine manure is used as a fertilizer; water supplies if manure used as crop fertilizer runs off into streams and waterways (Guan and Holley, 2003, J Environ Qual 32:383-92); and neighboring pigs resulting in a food safety problem and animal health issue. Control of food-borne *Salmonella* within the farm-retail continuum is a complex issue. An essential step in providing pathogen-free products to consumers is reduction of the risk of food-borne disease on the farm level. Reduction of on-farm *Salmonella* contamination by the use of antimicrobials is not a sustainable way of disease control for two reasons: one, antimicrobial overuse can lead bacteria to develop resistance to such antimicrobials which decreases their effectiveness for the producer and rises serious human health-care issues (Perron et al, 2008, PLoS One 3:e3749); and second, pigs can be infected with *Salmonella* not only at the farm, but also during transportation and/or lairage. A potentially more effective method of addressing pre-harvest food safety is through genetic improvement of disease resistance in animals. Thus, the long-term goals

of this project are to use the power of genomics to identify gene variants that are associated with decreased *Salmonella* shedding on farms; and to use these variants to select for animals that shed fewer bacteria and so are less likely to cause contamination in lairage, at the abattoir, or in carcasses during fabrication. To identify genetic components controlling disease resistance differences in pigs, we: 1) created a large resource population of pigs with known *Salmonella* shedding phenotypes; 2) created DNA tests for genetic variation in candidate genes that are involved in swine response to *Salmonella*, 3) genotyped our resource populations for single nucleotide polymorphisms (SNPs) in the candidate genes and 4) analyzed SNPs for associations with *Salmonella* shedding and/or tissue colonization phenotype. Previous work has been successful in identifying an important association between genetic variation in the CCT7 gene and *Salmonella* shedding (Uthe et al, 2009, Vet Microbiol 135:384-8). In the current work using a total of 750 pigs we genotyped 62 SNPs in 53 genes that are involved in porcine response to *Salmonella* infection. We found 28 segregating genetic variants with minor allele frequency of 15 % or higher in at least 2 of our 4 porcine populations. Statistical analysis revealed several SNPs associated with *Salmonella* fecal shedding or tissue colonization, with p-value <0.1 and q-value ≤ 0.2, where q-value is the minimum false discovery rate (FDR) such that the corresponding null hypothesis can be rejected. Fecal shedding associated SNPs were CCT7 #3 (p=0.041), GNG3 (p=0.029), PGD (p=0.047) and HP #2 (p=0.0002) in the field population, AMT (p=0.005) in the NADC-40 pig population and PGD (p=0.001) in the NADC-77 pigs population. In the IAH-Compton population three SNPs in the ACP2 gene were associated with *Salmonella* burden in spleen (ACP2#1, p=0.013, ACP2#2, p=0.026, ACP2#3, p=0.033); and in the NADC-40 pig population, a SNP in the EMP1 (p=0.002) gene was associated with bacterial load in ileo-cecal lymph node. The genotyping and statistical analysis data will be used to generate genetic markers useful in selecting animals that shed fewer bacteria and are less likely to cause pen-mate *Salmonella* contamination in a farm and on a slaughter plant, thus providing novel information for the pig industry to implement in strategies for selecting pigs with reduced shedding and/or disease susceptibility.

Introduction

Pigs shedding pathogenic bacteria such as *Salmonella* on a farm or at the slaughter plant are a major problem in pre-harvest food safety (Hurd et al., 2001, J. Food Prot. 64: 939-944.). Nearly 15% of all known causes of food-borne illness outbreaks in the United States can be tracked to consumption of contaminated pork (Guan and Holley, 2003, J Environ Qual 32: 383-92). A survey performed by the USDA's National Animal Health Monitoring System (NAHMS) indicated in 2006 that ~53% of pig farm sites sampled in the 17 states of the U.S. were positive for *Salmonella*. The difficulty in identifying *Salmonella* carrier animals due to sub-clinical status of the disease poses an additional challenge in scheming contamination in food chain. Control of food-borne *Salmonella* within the farm-retail continuum is a complex issue, as not only contaminated meat can be a food safety hazard. Manure from farms with *Salmonella* carrier animals that is used as fertilizer can contaminate edible crops and/or run into water streams posing additional food safety issues (Guan and Holley, 2003. J Environ Qual 32:383-92). Thus, reduction of the risk of food-borne disease at the farm level is an essential step in providing pathogen-free products to consumers, including both meat and non-meat products (i.e. vegetables and fruits). Antimicrobials are frequently used to control disease in pigs, however it is not a sustainable way of disease control for two reasons: one, antimicrobial overuse can lead to bacterial resistance to such antimicrobials which decreases their effectiveness for the producer; and two, pigs can be infected with *Salmonella* not only on the farm, but also during transportation and/or lairage (the holding pen of the abattoir). In fact, it was shown that high levels of *Salmonella* detected in carcasses are positively correlated with both *Salmonella* shedding by carrier swine and *Salmonella* prevalence on a farm (Nollet et al, 2005, Vet Res 53:645-56). Thus, a potentially more effective method of addressing *Salmonella* control and pre-harvest food safety issues is through genetic improvement of disease resistance in animals. Several reports have indicated that resistance to disease is under significant genetic control in pigs. However, few attempts to use pig breeding to improve genetic resistance have been reported. With few QTL and genetic variants associated with disease resistance, we are attempting to advance porcine disease resistance by identifying genes and genetic variations that can be used to accurately select more resistant pigs. To achieve our long-term goals of implementing genomics to identify genes that are associated with decreased *Salmonella* shedding on farms, we used a

collaborative effort to create a large porcine genetic resource that include four populations with known *Salmonella* shedding phenotypes or tissue colonization levels. The populations were then used to screen for genetic variations in genes identified by our group to be involved in transcriptional control of swine response to *Salmonella*. Statistical analyses assessed the associations of genetic polymorphism with the disease phenotype. Thus, this study provides important resource for genetic improvement of disease resistance in pigs.

Objectives

The long-term goals of this project are to use the power of genomics to identify genetic variants that are associated with decreased *Salmonella* shedding on farms; and to use these variants to select for animals that shed fewer bacteria and so are less likely to cause abattoir carcass contamination.

Objective I. Complete the creation of an expanded bank of *Salmonella* fecal shedding data and matched DNA samples based on a large cooperative epidemiology study that is surveying US industry pigs.

During the 2 year of the grant period about 3,200-3,600 porcine fecal samples were collected and tested from about 60 Iowa farms in a collaboration with Dr. A. O'Connor and J. McKean under a USDA-NRI funded project. Within this sample we identified which animals were shedding *Salmonella* seven days before marketing and which were therefore abattoir contamination hazards. During a 2-year period we identified 163 pigs that shed *Salmonella* in their feces seven days before marketing as well as 299 negative controls selected from within all farm-visit cohorts with positive animals. We isolated high quality DNA from the belly flaps of all the 462 pigs. We have created a DNA bank from the field population pigs with about 1.5 – 15 μg of DNA per sample. We can deliver the DNA bank aliquots to NPB for use in scientific research.

Objective II.

1st year of the grant: Create DNA tests for genetic variation at candidate genes, selected based on our previous data on gut lymph node gene expression patterns in response to *Salmonella* infection.

Previous research (Uthe et al, 2007, Mol Immunol 44:2900-14; Wang et al, 2007, Genomics 90:72-84; and 2008, BMC Genomics 9:437) identified genes differentially regulated in pigs at various stages during *Salmonella* infection from acute (8 hr post inoculation (p.i.)) to the chronic (21 days p.i.) stages. As candidates, we attempted to identify genes within these gene sets that would be involved in controlling porcine response to infection under different conditions at different farms. Using our SNP/gene selection criteria (see materials and methods) we selected 54 candidate SNPs which were then used to create a set of 29 SNPs (23 genes) for multiplex Sequenom genotyping.

2st year of the grant: Use gene expression data to develop DNA tests for functional candidate genes for *Salmonella* resistance.

For selection of functional candidate genes we used Affymetrix microarray analysis performed by Dr. Chris Tuggle group in collaboration with Dr. Shawn Bearson (NPB grant # 05-176) that profiled porcine response to *Salmonella* in whole blood. Analysis of transcriptional immune-inflammatory response in whole blood at 2 days post *Salmonella* inoculation revealed many genes that are responding to infection as well as genes that differentiate pigs that shed high numbers of bacteria compared to low-shedding pigs. We attempted to select candidate genes differentially regulated in persistent-shedder versus low-shedder pigs as such markers may be useful in determining *Salmonella* shedding status in a pig. We used our Anexdb database (www.anexdb.org, Couture et al, 2009, Mamm Genome, 11-12:768-77) to predict putative SNPs in candidate genes. Based on our SNP/gene selection criteria (see materials and methods) we selected 77 SNPs in candidate genes which were then used to create a set of 31 SNPs (28 genes) for multiplex Sequenom genotyping.

Objective III. Select case-control groups for SNP detection and perform genotyping using new and established DNA tests on US industry population bank as well as a published challenge study populations.

We used the fecal shedding data from Objective I to select the "cases" (*Salmonella* positives) and the "control" (negative farm lot-mates) for analysis. We isolated DNA from belly flap tissues of case and control pigs and then genotyped this field population using DNA markers from Objective II. As a complement to the

field population, for genotyping analysis we also used other 3 population with *Salmonella* shedding and/or tissue colonization data: we isolated DNA from the NADC-40 pig and NADC-77 pig populations as well as used the IAH-Compton population DNA, which was a gift from Pig Improvement Company.

Objective IV. Test associations of SNP genotypes with *Salmonella* incidence and integrate data from epidemiology and genomics projects to understand *Salmonella* contamination.

Statistical analysis revealed several SNPs associated with *Salmonella* fecal shedding or tissue colonization with p-value <0.1 and q-value ≤0.2. Fecal shedding associated SNPs were CCT7 #3 (p=0.041), GNG3 (p=0.029), PGD (p=0.047) and HP #2 (p=0.0002) in the field population, AMT (p=0.005) in the NADC-40 pig population and PGD (p=0.001) in the NADC-77 pigs population. In the IAH-Compton population three SNPs in the ACP2 gene were associated with *Salmonella* burden in spleen (ACP2#1, p=0.013, ACP2#2, p=0.026, ACP2#3, p=0.033); and in the NADC-40 pig population, a SNP in the EMP1 (p=0.002) gene was associated with bacterial load in ileo-cecal lymph node.

Materials and Methods

Collection of fecal and belly flap samples provided from O'Connor project. Fecal samples (20-30 grams) were collected into plastic bags with individual labeled identification as animals were tattooed at that time. The belly flap samples were collected at the abattoir from dressed carcasses identified to the on-farm fecal sample using a unique slap tattoo number. Belly flap samples were placed in plastic bags and frozen for DNA preparation later.

***Salmonella* fecal test.** To identify pigs shedding *Salmonella,* qualitative bacteriology was performed as follows: 10 gram samples of swine feces were assayed in duplicates using *Salmonella* enrichment and selective media as previously described (Hurd et al, 2002, Appl Environ Microbiol, 68:2376-2381; J T Gray et al, 1996, Appl Environ Microbiol, 6:141-6). Positive isolates were confirmed by serogroup antiserum agglutination assays.

DNA isolation. To isolate DNA from belly flap tissue samples, about 20 mg of the tissue was digested with proteinase K (Invitrogen, Carlsbad, CA). DNA from tissue lysates was extracted either using phenol-chloroform method or Wizard SV genomic DNA purification system (Promega, Madison, WI) according to the manufacturer's protocol.

Selection of genes for SNP analysis.

Criteria for selecting functionally relevant genes during the 1st year of the grant. Our preliminary research (Uthe et al, 2007, Mol Immunol 44:2900-14; Wang et al, 2007, Genomics 90:72-84; and 2008, BMC Genomics 9:437) identified genes differentially regulated in pigs at various stages during *Salmonella* infection from acute (8 hr post inoculation (p.i.)) to the chronic (24 hr p.i.) stages. We screened our list of differentially regulated genes to select candidate markers for SNP analysis. As candidates we attempted to identify genes that respond to both *Salmonella* Choleraesuis and *Salmonella* Typhimurium and thus would be more likely to be involved in controlling porcine response to infection under different conditions at different farms. In addition, we have recently enriched our target gene lists with genes that are differentially regulated in *Salmonella* persistent shedders versus low-shedder pigs' whole blood. Our latest analysis indicates that some gene lists from porcine whole blood microarray can differentiate pigs based on their shedding status (Tuggle, Huang, Uthe, and Bearson, unpublished data). So, our gene and putative SNP selection criteria were as follows:

1) Identify genes that are differentially regulated early, at 8 – 24 hr during *Salmonella* Typhimurium and *Salmonella* Choleraesuis infections in pigs.
    a. Select genes that, based on published data (Jenner and Young, 2005, Nat Rev Microbiol. 2005, 3:281-94), are involved in response of the organism to multiple bacterial pathogens.
    b. Identify genes which expression are known to be "non-tolerized" or non-diminished during continuing re-infection as such genes are most likely to be involved in clearance of bacterial pathogens (Foster et al, 2007, Nature 2007, 447:972-8)
2) Identify genes from whole blood microarray analysis that differentiate *Salmonella* persistent shedders from low-shedder pigs.

3) Identify genes that have a putative SNP based of PEDE (Pig Expression Data Explorer, http://pede.dna.affrc.go.jp/) or TIGR databases (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=pig)

Criteria for selecting functionally relevant genes during the 2st year of the grant. Using Affymetrix technology, our group previously profiled porcine response to *Salmonella* using whole blood RNA (NPB grant # 05-176). Analysis of transcriptional immune-inflammatory response in whole blood at 2 days post *Salmonella* inoculation revealed many genes that are responding to infection as well as genes that differentiate *Salmonella* persistent-shedder compared to low-shedding pigs. Persistent-shedders were pigs from the NADC-40 population that shed *Salmonella* in their feces throughout the 20 day challenge experiment and low-shedders were pigs that did not shed the bacteria after day 7 post inoculation (Uthe et al, 2009, Vet Microbiol, 135:384-8). We used the following criteria for selecting candidate SNPs (Fig 1):

1) We prioritized genes from our whole blood microarray analysis that can be used to differentiate *Salmonella* persistent-shedder versus low-shedder phenotype and thus can be used as markers identifying such phenotype on a farm.
2) We prioritized selection of non-synonymous SNPs.
3) An additional guiding resource was involvement of genes in host response to multiple bacterial pathogens as classified by Jenner and Young (2005, Nat Rev Microbiol, 3:281-94) as well as "non-tolerizable" expression pattern (non-diminished due to re-infection) as described by Foster et al (2007, Nature, 447:972-8).

Presence of putative SNPs in the candidate genes was predicted using AnexDB database (www.anexdb.org, Couture et al, 2009, Mamm Genome, 11-12:768-77).

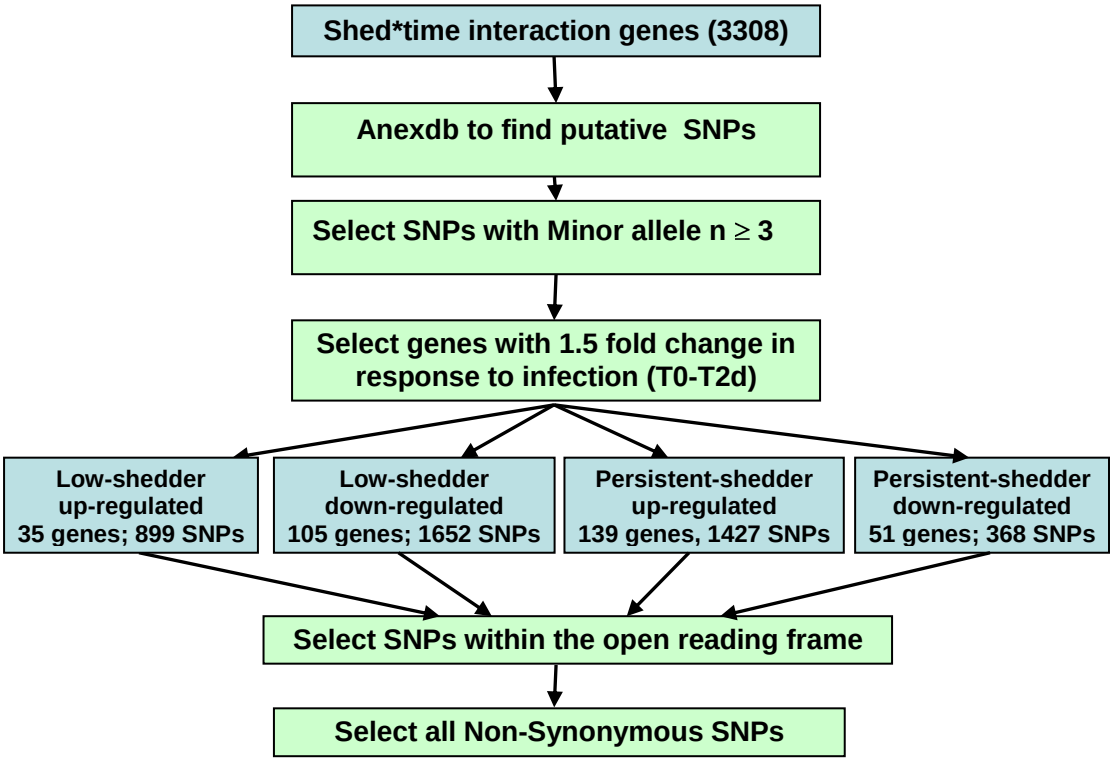

Fig 1. Selection of functional candidate genes for SNP genotyping during year 2.

SNP confirmation and genotyping. To confirm putative SNPs, four pools of 4 PCR product DNA samples, pooled according to pig breed, were created and sequenced at the ISU DNA sequencing facility. Allele frequency estimates based on the sequencing data was confirmed by restriction fragment length polymorphism (RFLP) analysis. To improve the efficiency, throughput and reduce the cost of SNP genotyping we have adopted the Sequenom MassARRAY genotyping technology (MassARRAY Compact System, Sequenom Inc., San Diego, CA) which has been installed at the ISU Center for Plant Genomics. In Sequenom technology allele-specific products with distinct masses are generated using simple primer-extension chemistry. Analysis allows multiplexing up to 40 SNP assays simultaneously. Alleles are then accurately discriminated using Matrix-Assisted Laser Desorption / Ionization Time-Of-Flight mass spectrometry (MALDI-TOF MS). Adoption of this new technology in our laboratory took more time initially, but proved to be more efficient in the long term.

Using Sequenom technology we genotyped the SNPs selected in Objective II. A total of 750 pigs from our 4 populations were included for genotyping: 405 field (qualitative shedding data), 228 IAH-Compton (quantitative tissue colonization data) and 40 as well as 77 NADC population pigs (quantitative shedding across 4-5 time points post challenge). SNPs that had less than 80% of allele call rate in Sequenom analysis were completed through manual genotyping by using either RFLP or tetra primer ARMS (amplification refractory mutation system)-PCR technologies.

Association analyses. Statistical association analysis of genotype with *Salmonella* shedding and / or tissue colonization phenotype was performed as follows:

Field population: *Salmonella* shedding status (positive/negative) data were analyzed using nominal logistic regression in the JMP software, with farm and the SNP of interest treated as fixed effects. The analysis was performed for each SNP separately. Parameters were estimated by the method of maximum likelihood. As all estimated parameters were finite and the sample size was reasonably large, an asymptotic likelihood ratio test was performed to test the effect of each SNP of interest.

IAH-Compton population: Log transformed spleen and liver counts were analyzed using mixed linear model in the JMP software, with group, sire within group, the SNP of interest as fixed effects, and dam within group-by-sire interaction treated as random effects. The analysis was performed for each tissue and each SNP separately. Parameters were estimated using the method of residual maximum likelihood (REML). Kenward-Roger (1997; Biometrics, 53:983-997) corrected Wald-type *F*-test was conducted for the effect of SNP of interest.

NADC-40 pig and NADC-77 pig population statistical analysis. Due to small sample sizes from these two populations, we conducted the exact permutation tests in place of asymptotic tests applied to the field population and IAH-Compton population. For the binary ileo-cecal lymph node (ICLN) trait in NADC-40 pig population, a logistic regression model was used. Each SNP was analyzed separately. Log likelihood ratio statistic was calculated by comparing the maximized log likelihood from the model with sow as the only fixed effect and the model with both sow and the SNP being analyzed as fixed effects. The null distribution of the log likelihood ratio statistic was obtained by randomly permuting the SNP genotype within each sow for 19,999 times. The 4-dimensional and 5-dimentional shedding count trait of the NADC-40 and the NADC-77 pigs

populations respectively was transformed by taking natural logarithms after adding one to each count to avoid infinite logarithms. The multi-response permutation procedure (MRPP) (Mielke and Berry, 2007, New York: Springer) was then applied to the transformed counts for each SNP and each population separately. The MRPP test statistic was calculated as the weighted mean of within-"treatment" average pair-wise Euclidean distances of pigs, with weights being the number of pigs within the same treatment. For the NADC-40 population, the "treatment" was simply defined as the SNP genotype; for the NADC-70 population, the "treatment" was defined as SNP-by-sow interaction due to presence of influential sow effects. However, the null distribution of the MRPP test statistics was always obtained by randomly permuting the SNP genotypes *within* each sow for 19,999 times to account for sow effects. For both the binary ICLN trait and the multidimensional shedding trait, a raw p-value for each SNP was calculated as the proportion of permutations that resulted in as or more extreme test statistics than the one observed without permutation.

Due to multiple hypothesis tests performed in the analyses, we control SNP false discovery rates (FDR) for each trait in each population separately. For each set of *p*-values, we applied the procedure of Storey and Tibshirani (2003, Proc Natl Acad Sci USA, 100:9440-5), except that we fixed lambda to be the observed *p*-value such that difference between empirical distribution function evaluated at this *p*-value and the *p*-value itself is maximum, because the smoother method of Storey and Tibshirani (2003) is not stable when there are only dozens of *p*-values in each set. If there are much more *p*-values in each set, our method is more conservative than the smoother method, and hence it controls FDR.

All analyses were performed in the statistical computing environment R, with code available upon request.

Results

Objective I.

In collaboration with O'Connor project, belly flap samples were collected from farms with *Salmonella* positive animals. During a 2-year period a total of 462 belly flap samples from about 60 farms have been collected that includes 163 *Salmonella* positive and 299 negative samples as determined by qualitative fecal

testing (Table 1). It should be noted that NPB did not fund this part of the project; all work and costs associated with farm enrollment, sample collection and travel were USDA-NRI funded, and this was an extremely costly and important contribution to the NPB project.

Table 1. Farms with positive *Salmonella* shedder pigs for genotyping and association analysis

ID of Farm Sampled	*Salmonella* positives	*Salmonella* negatives
2-3	1	2
2-4	4	12
3-5	3	6
7-3	1	3
8-3	5	10
10-3	8	12
10-4	5	10
11-2	1	2
13-3	3	6
14-1	3	6
14-2	4	8
14-3	1	2
14-4	3	6
15-1	3	6
15-2	6	12
16-1	1	2
20-2	1	2
21-1	6	12
25-2	2	4
26-1	10	24
27-1	1	2
27-2	9	18
28-1	1	2
29-1	30	26
30-1	4	8
30-2	4	11
31-1	11	22
33-1	6	12
34-1	4	8
35-1	1	2
36-1	8	16
37-1	4	8
38-1	4	7
39-1	1	2
40-1	4	8
Total	**163**	**299**

Three other resource populations with quantitative *Salmonella* shedding and / or tissue colonization data are also available and were used by our group for genotyping analysis. These include the NADC-40 pig (NPB Project # 05-176) and the NADC-77 pig (newly funded USDA-NRI Functional Genomics project) populations

as well as the IAH-Compton 228 pigs population (van Diemen et al, Vet Immunol Immunopathol. 2002, 88:183-96).

Objective II and Objective III.

Year 1 selection and genotyping of functionally relevant genes.

For the first round genotyping based on our selection criteria we selected a set of 54 SNPs in genes differentially regulated during *Salmonella* infection in pigs. Using a Sequenom technology a set of 29 SNPs have been created for multiplex genotyping, that included 14 novel putative SNPs and 15 SNPs previously confirmed by our group. All 29 SNPs were genotyped using the 4 populations that total 750 DNA samples including 405 field pigs, 228 IAH-Compton as well as 40 and 77 NADC population pigs (Table 2). A set of previously established SNPs served as internal quality control in our Sequenom genotyping analysis; the agreement between our manual RFLP analysis and Sequenom genotyping was 99%. For further statistical analysis we used Sequenom assays (SNPs) that had a minimum of 80% of genotyping calls and minor allele frequency (MAF) of 15%. Majority of our assays fell within this category with a few of the SNPs, such as TYROBP, CD163 #2 and CCT7#3 requiring additional manual genotyping to improve the allele call rate.

Table 2. Summary of genotyping and associations analysis, Sequenom set #1.

	Porcine populations			
	NADC-40	NADC-77	Field	IAH-Compton
SNPs tested by Sequenom and /or RFLP	31	29	29	31
Genes tested	25	23	23	25
Informative SNPs found (MAF ≥ 15%)	18	12	18	13
Informative genes found	14	8	15	13
# of SNPs associated with shedding /tissue colonization	2	0	2	2

We identified 18 segregating SNPs in 15 genes with MAF of ≥ 15% in at least 2 populations. Of those SNPs 10 were new markers and 8 were confirmed markers (NPB project #05-176) that this time we used to genotype larger populations of pigs (Table 3).

Table 3. Sequenom SNP set #1: segregating SNPs genotyped across the 4 porcine populations.

SNP	Geno-typing method**	NADC-40 population, n=40				NADC-77 population, n=77			Field population, n=405			Compton population, n=228				Total # of pigs geno-typed
		MAF %	% pigs geno-typed	Fecal counts p-value	ICLN counts p-value	MAF %	% pigs geno-typed	Fecal counts p-value	MAF %	% pigs geno-typed	Shedding status p-value	MAF %	% pigs geno-typed	Spleen counts p-value	Liver counts p-value	
ACP2 #1	Sequenom	36	93	0.36	0.933	31	100		17	94	0.575	25	89	0.013*	0.226	699
ACP2 #3	Sequenom	40	90	0.48	0.872	39	97		21	90	0.537	29	83	0.033*	0.147	664
ACP2 #1+3				0.413	0.992									0.012*	0.108	
AMT	Sequenom	31	100	0.005*	0.115	18	99	0.329	27	96	0.465	21	100	0.056	0.189	732
BM88	Sequenom	9	93			5	96		5	95		1	100			722
CCR1	RFLP+ S	44	100	†	†	38	100	0.788	29	96	0.955	42	100	0.127	0.062	733
CCT7 #2	RFLP+ S	25	100	†	†	41	99	0.747	29	96	0.186	10	99			730
CCT7 #3	RFLP+ S	29	100	0.877	0.996	38	78		38	93	0.041*	4	100			705
CCT7 #2+3		-	-	0.691	0.414	-	-		-	-	0.087	-	-			
CD163 #2	TP-ARMS-PCR+S	25	100	0.308	0.611	45	97	0.649	32	79		18	100	0.119	0.085	663
CD163 #3	RFLP+ S	26	100	†	†	41	83	0.751	23	86	0.634	15	99	0.467	0.168	677
CDC123	Sequenom	24	100	0.546	0.999	0	5		34	33		40	99	0.829	0.876	404
CD47/IAP	RFLP	8	90			not typed			9	8		48	100	0.777	0.895	300
EMP1	Sequenom	15	98	0.347	0.002*	0	100		19	97	0.475	0	99			734
GNG10	Sequenom	21	85	0.875	0.679	not typed			26	33		36	96	0.682	0.508	387
GNG3	Sequenom	9	98			3	100		33	96	0.029*	15	99	0.343	0.621	731
GSTA3 #3	Sequenom	13	50			11	95		-	-		23	93	0.199	0.527	306
LCP1 #1	RFLP+ S	15	100	†	†	0	79		6	61		0	98			570
MGP #2	Sequenom	0	60			0	92		3	63		2	43			450
NCF2 #1	RFLP+ S	24	100	†	†	42	100	0.104	24	93	0.117	32	99	0.446	0.112	720
PDXK #1	Sequenom	8	95			not typed			9	29		7	78			334
TAP1 #2	RFLP+ S	14	100	0.215	0.097	not typed			28	26		35	78			323
TLR4 #4	RFLP+ S	40	100	0.356	0.067	26	100	0.834	9	98		0	98			735
TLR4 #6	RFLP+ S	41	100	0.436	0.065	26	99		19	93	0.834	0	99			719
TYROBP	RFLP	34	100	†	†	not typed			30	26		35	99	0.654	0.288	370
VCL	Sequenom	34	85	0.697	0.456	not typed			46	41		11	85			394
VCP #3	RFLP+ S	31	100	†	†	20	99	0.074	26	96	0.356	25	99	0.733	0.496	732

† Uthe JJ et al, Vet Microbiol. 2009;135:384-8.

** RFLP+S – marker genotyped using RFLP and Sequenom;
TP-ARMS-PCR+S – marker genotyped using tetra primer ARMS-PCR and Sequenom;

* Significant p-value with q-value ≤ 0.2; only p values for SNPs with MAF≥15% and pigs genotyped ≥80% are reported.

Year 2 selection and genotyping of functionally relevant genes.

An additional set of functionally relevant genes was selected for genotyping by Sequenom during the year 2 of the grant. Based on criteria described in materials and methods we selected a set of 77 SNPs in genes that by microarray analysis were are differentially regulated in pigs that shed high numbers of *Salmonella* (persistent- shedders) versus pigs that were low-shedders. Using Sequenom technology a set of 31 SNPs (28 genes) was created for multiplex genotyping (Table 4). Again, for statistical analysis we used Sequenom assays (SNPs) that had a minimum of 80% of genotyping calls and minor allele frequency (MAF) of 15%. Several segregating SNPs are currently being manually genotyped to improve the allele call rate.

Table 4. Summary of genotyping and associations analysis, Sequenom set #2.

	Porcine populations			
	NADC-40	**NADC-77**	**Field**	**IAH-Compton**
SNPs* tested	31	31	31	31
Genes tested	28	28	28	28
# of non-synonymous SNPs	26	26	26	26
Informative SNPs found (MAF ³ 15%)	9	10	9	11
Informative genes found	8	9	8	10
# of SNPs associated with shedding /tissue colonization	0	1	2	2

This work identified 10 new additional segregating SNPs in 9 genes with MAF of ≥ 15% in at least 2 populations (Table 5).

Table 5. Sequenom SNP set #2: segregating SNPs genotyped across the 4 porcine populations.

SNP	Geno-typing method	NADC-40 population, n=40				NADC-77 population, n=77			Field population, n=405			Compton population, n=228				Total # of pigs geno-typed
		MAF %	% pigs geno-typed	Fecal counts p-value	ICLN counts p-value	MAF %	% pigs geno-typed	Fecal counts p-value	MAF %	% pigs geno-typed	Shedding status p-value	MAF %	% pigs geno-typed	Spleen counts p-value	Liver counts p-value	
ACP2#2	Sequenom	25	100	0.565	1.000	38	100	0.247	19	94	0.537	18	99	0.026*	0.297	723
ACP2 #1+2+3		-	-	-	-	-	-	-	-	-	-	-	-	0.012*	0.108	
DDRGK1	Sequenom	26	98	0.953	0.248	48	66		32	48		33	44			385
ELF3	Sequenom	15	100	0.713	0.489	16	100	0.829	12	96		32	99	0.514	0.520	732
GLS2	Sequenom	19	98	0.376	0.080	28	100	0.127	23	98	0.474	18	99	0.634	0.852	739
HLA-A	Sequenom	9	95			6	83		12	76		49	63			553
HP #2	Sequenom	12	98			39	95	0.782	16	88	0.0002*	36	99	0.915	0.533	694
ITIH1#2	Sequenom	13	30			0	18		8	12		13	34			152
NCF1#1	Sequenom	49	98	0.904	0.139	29	100	0.585	50	91	0.739	25	95	0.600	0.288	701
NCF1#2	Sequenom	33	53	0.652		31	35		41	67		34	42			415
NCF2 #2	Sequenom	38	100		0.126	45	96	0.377	41	94	0.368	45	99	0.586	0.126	720
PGD	Sequenom	42	93	0.300	0.365	24	100	0.001*	50	89	0.047*	11	98			698
RNH1	Sequenom	41	85	0.848	0.595	48	66		44	82	0.695	23	83	0.519	0.665	606
SERPINB1	Sequenom	10	100			5	100		4	96		10	99			732
SULT1A1 #2	Sequenom	6	100			4	100		7	95		20	94	0.909	0.878	716

* Significant p-value with q-value ≤0.2; only p values for SNPs with MAF≥15% and pigs genotyped ≥80% are reported.

In summary:

During year 1 and 2 of the grant we genotyped a total of 62 SNPs in 53 genes involved in porcine response to *Salmonella* infection. We identified 28 segregating markers with MAF of ≥ 15% in two or more populations. Eight markers were established (NPB#05-176) and this time used to genotype larger populations of pigs while 20 markers were novel SNPs identified in this study.

Objective IV.

Statistical association analysis revealed several SNPs that were associated with *Salmonella* shedding or tissue colonization phenotypes with p-value < 0.1 and q-value ≤ 0.2, where *q*-value is the minimum false discovery rate (FDR) such that the corresponding null hypothesis can be rejected. In the field population SNPs associated with fecal *Salmonella* shedding were CCT7 #3, GNG3, PGD and HP #2 (Fig. 2).

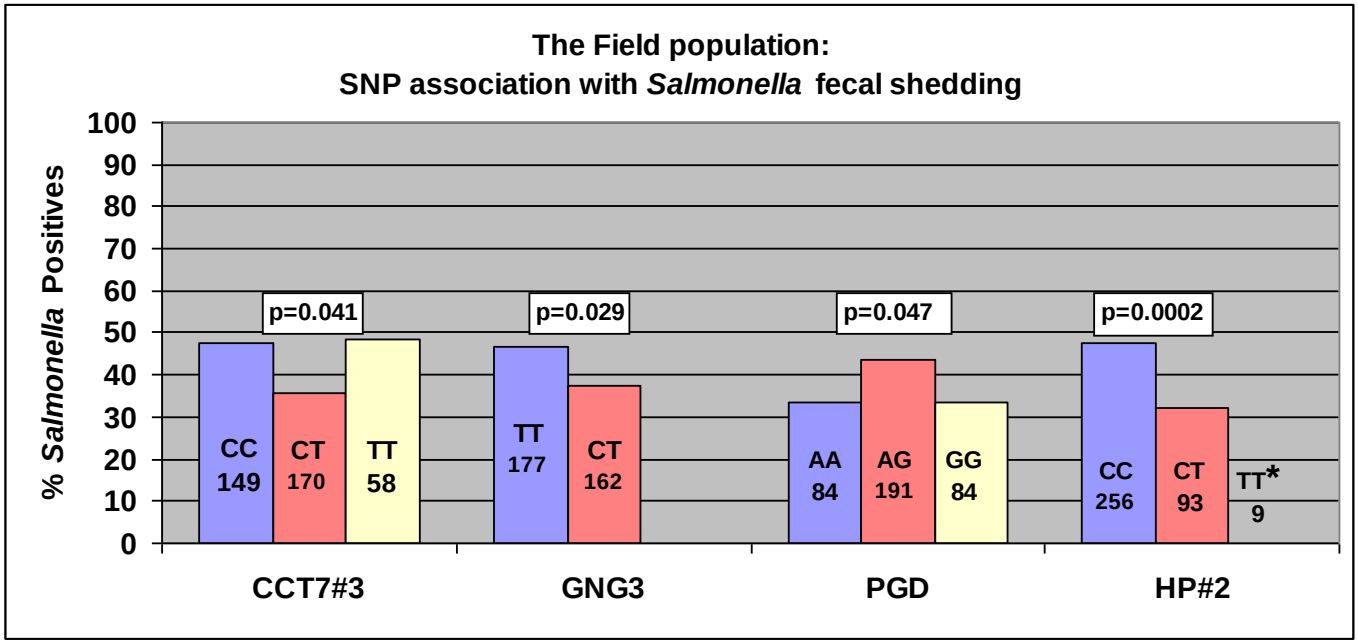

Fig. 2. Association of SNPs with *Salmonella* fecal shedding in the field population. Genotype and the number of pigs having the particular genotype are indicated on the bar graphs. Expected random percentage of positives is 40%. * Pigs with the T/T genotype for HP#2 SNP were all *Salmonella* negative.

Association analysis in IAH-Compton population identified three SNPs in the ACP2 gene associated with *Salmonella* burden in spleen (Fig. 3). In this population animals having homozygous GG (ACP2 #1), TT (ACP2 #2) and AA (ACP2 #3) genotypes had the highest number of bacteria in the spleen tissue at 7 days post inoculation.

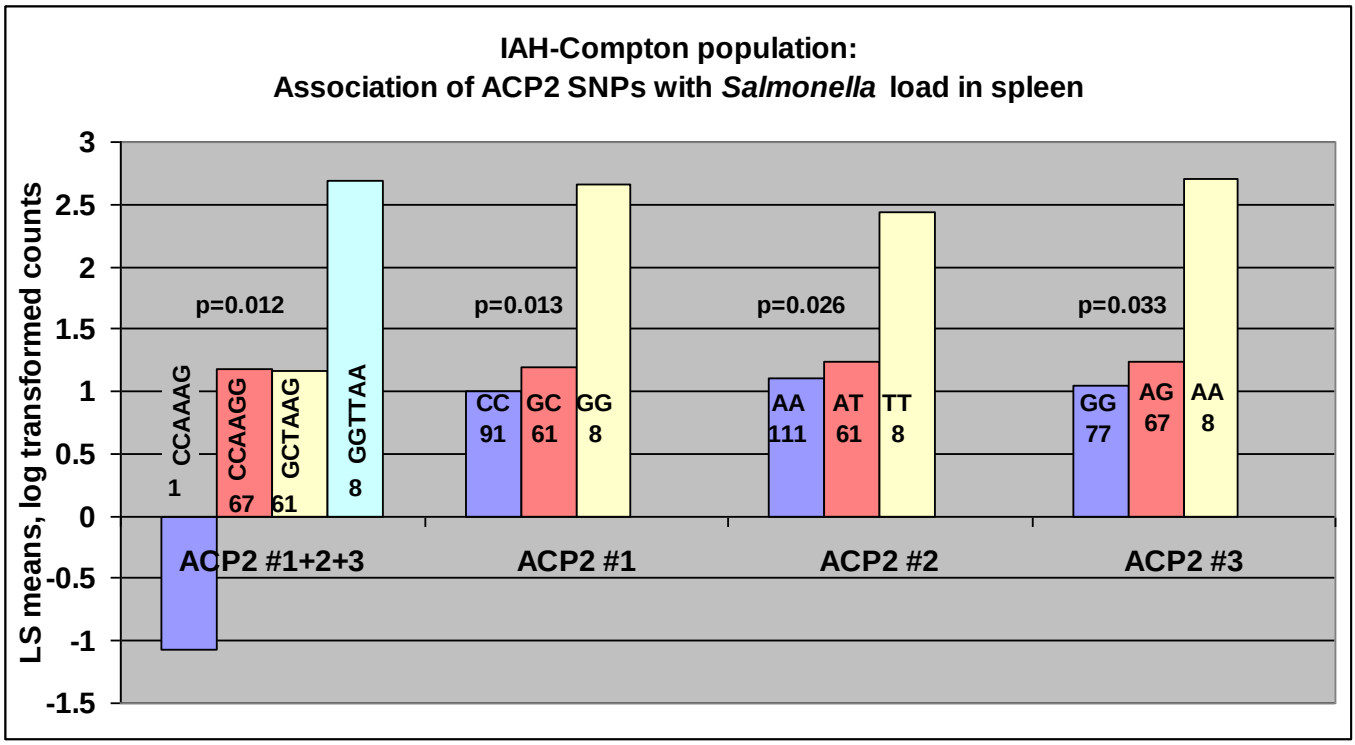

Fig. 3. Association of SNPs in the ACP2 gene with spleen *Salmonella* load in IAH-Compton population. Genotype and the number of pigs having the particular genotype are indicated on the bar graphs. Note: negative LS mean of the *Salmonella* count in spleen arises from the mean count being very small (0.989)

In the NADC-40 pigs population, 2 new SNPs were statistically associated: SNP in the AMT gene was associated (p=0.005) with *Salmonella* fecal shedding over time course of the infection (Fig. 4) and SNP in the EMP1 (p=0.002) gene was associated with bacterial load in ileo-cecal lymph node (ICLN) at 21 days post inoculation (Fig. 5)

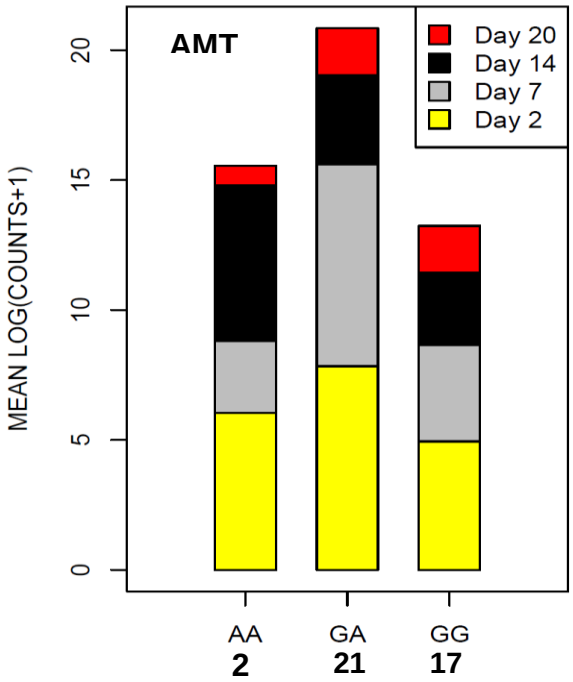

Fig. 4. Association of AMT with *Salmonella* fecal shedding in the NADC-40 pigs population, p=0.005, q=0.095. Number below the genotype is indicating the number of pigs with that genotype in the population.

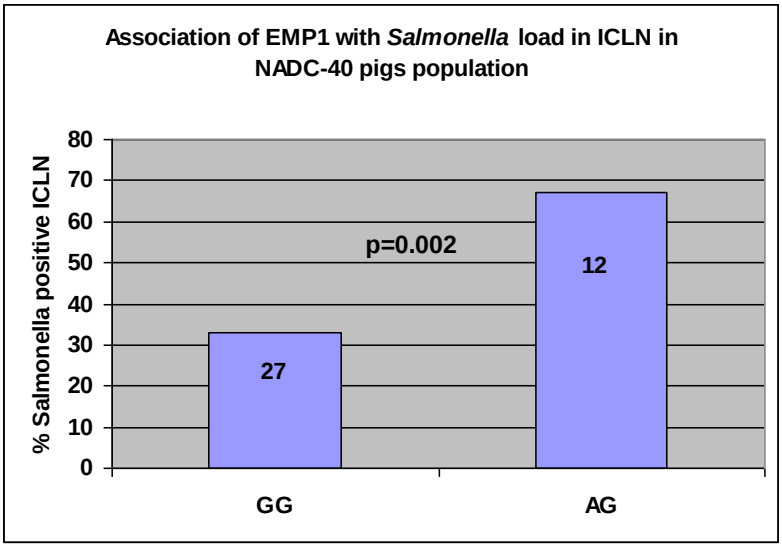

Fig. 5. Association of EMP1 with *Salmonella* load in ICLN in the NADC-40 pigs population, p=0.002, q=0.04. Number of pigs having the particular genotype is indicated on the bar graphs.

In the NADC-77 pig population, a SNP in the PGD gene was associated (p=0.001) with *Salmonella* fecal shedding over time (Fig. 6).

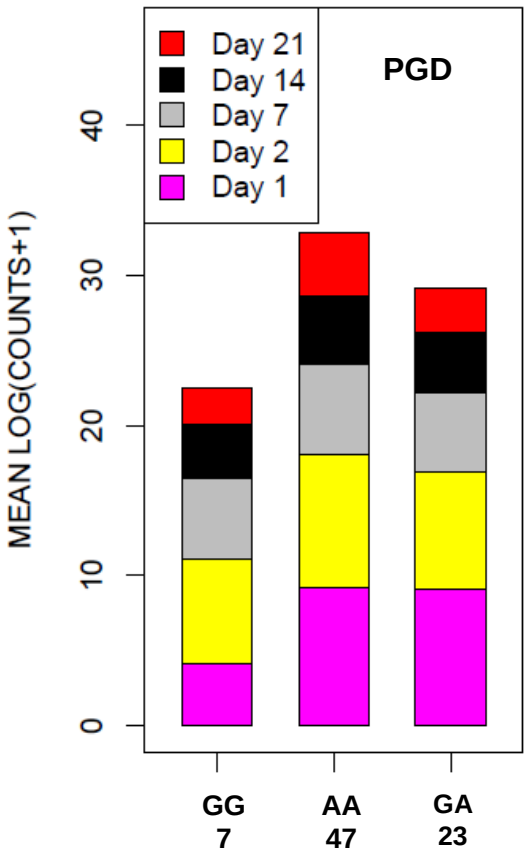

Fig. 6. Association of SNP in PGD gene with *Salmonella* shedding in the NADC-77 pigs population, p=0.001, q=0.021. Number below the genotype is indicating the number of pigs with that genotype in the population.

VIII. Discussion

Salmonella in a serious animal health and food-safety issue with salmonellosis being one of the 10 most common diseases among weaning age and grower-finished pigs in the farms in the United States (NAHMS. 2000. See: http://www.aphis.usda.gov/vs/ceah/ncahs/nahms/swine/swine2000/swine2kPt2.pdf).

Ability of *Salmonella* to establish a carrier state in pigs provides a reservoir for the pathogen, where upon stressful conditions (i.e. transportation, mixing, etc.) the pathogen can re-emerge from the animal and be shed in the feces to contaminate/infect pen mates as well as the environment. Thus, control of *Salmonella* in the farm and on the slaughter plant is a complicated issue. Therefore, to minimize *Salmonella* shedding it is not only important to identify the pigs that are shedders and may become long-term carriers, but it is also important to prevent animals from becoming carrier in the first place. The use of genomics has great potential power in addressing the *Salmonella* control problem, as genetic markers in animals differently responding to the pathogen can be identified and genetic tests can be developed to improve shedding phenotypes. One of the first steps in this process in identifying target genetic variants that potentially differentiate swine in their response to *Salmonella* is to create large resource populations to be screened for SNPs in functionally relevant genes. In this study with collaborative effort from the O'Connor and McKean USDA-funded epidemiology project, we have created a large field population by sampling pigs from Iowa farms. We have identified 462 tissue samples and prepared DNA from *Salmonella* positive and negative pigs that came from about 60 Iowa farms. In addition, we are using the 40 pig and the 77 pigs NADC challenge as well as the 228 pigs Compton populations to increase the power of our genotyping and genotype-phenotype association analysis. Selecting candidate genes and SNPs is another essential step in understanding porcine genetic resistance to disease. Our previous research in characterizing transcriptional response of pigs to infection with *Salmonella* (Uthe et al, 2007, Mol Immunol 44:2900-14; Wang et al, 2007, Genomics 90:72-84; and 2008, BMC Genomics 9:437) and profiling differences in transcriptional response of *Salmonella* persistent-shedder and low-shedder pigs was particularly important in a gene selection process, as global transcriptional regulation of swine response to *Salmonella* has not been available before. Through the means of computational analysis we genotyped a total of 62 novel target SNPs in 53 genes involved in porcine response to *Salmonella*. To

increase throughput and reduce the costs of genotyping, we have adopted Sequenom technology for SNP genotyping. Using this new high throughput technology we have genotyped 750 DNA samples that include of 405 field population, 40 and 77 pigs NADC as well as 228 Compton population pigs. The genotyping information was used to statistically analyze any associations between genotype and *Salmonella* shedding and /or tissue colonization phenotype in these pigs. Statistical analysis revealed several associated SNPs. In the field population, 4 SNPs (CCT7#3, GNG3, PGD and HP #2) were associated with *Salmonella* shedding. Protein functions of the associated genes involve chaperoning and cytoskeleton protein folding (chaperonin containing TCP1, subunit 7, CCT7); receptor signaling (guanine nucleotide exchange factor gamma 3, GNG3); enzymatic activity (phosphogluconate dehydrogenase, PGD); and a multifunctional protein involved in hemoglobin metabolism and inflammation (haptoglobin, HP). HP is also one of the genes that is known to be not tolerized, (expression of this gene is not diminished during repeated LPS challenge) as reported by Foster et al. (Nature 2007, 447:972-8) and thus is likely to directly contribute to anti-bacterial defense function of the host. In our data, pigs having the rare homozygous T/T genotype for the HP #2 SNP, were all negative for *Salmonella* by fecal testing. These *Salmonella* negative pigs came from 7 different farms with a percentage of *Salmonella* positive pigs on the farm ranging between 1.4 – 20% at the time of sampling. Further research is needed to reveal the role of HP#2 SNP in porcine resistance to *Salmonella*.

In the IAH-Compton population 3 SNPs including one non-synonymous SNP in the ACP2 gene were associated with *Salmonella* burden in spleen. ACP2 codes for enzyme acid lysosomal phosphatase that is expressed in lysosomal compartment in all tissues and involved phosphomonoesters cleavage. The identified non-synonymous ACP2 #2 SNP is causing an amino acid substitution from glutamate to aspartate in the protein sequence.

In the NADC-40 pigs population, one SNPs was statistically associated one with *Salmonella* shedding over time (AMT) and another with bacterial burden in ileo-cecal lymph node (EMP1). The AMT gene codes for aminomethyltransferase involved in mitochondria-specific enzyme system for cleavage of glycine and EMP1 is epithelial membrane protein 1 involved in regulation of cell adhesion, cell signaling and cell communication. In the NADC-77 pigs population, a non-synonymous SNP in gene coding for enzyme phosphogluconate dehydrogenase (PGD) was associated with *Salmonella* shedding over time course of infection. The associated SNP in the PGD gene causes a non-synonymous amino acid change from histidine to arginine.. Further research could reveal an effect of our identified SNPs on a function of these genes and their role in disease resistance in pigs. We are moving forward with such mechanistic follow-up in the near future, as well as completing the genotyping of SNP group #2 through manual genotyping as done for SNP group #1.

This initial genotyping and statistical analysis generated several markers that can be further investigated to potentially discriminate porcine response to *Salmonella.* SNPs in our phenotype-associated genes can be useful in selecting animals that shed fewer bacteria and are less likely to cause pen-mate *Salmonella* contamination in a farm and on a slaughter plant. Thus, this research provides novel information for the pig industry to implement in strategies for selecting pigs with reduced shedding and/or disease susceptibility.

Lay interpretation

Pigs that shed *Salmonella* in their feces are an animal health issue, a food safety problem and an environmental contamination risk. *Salmonella* can enter the food chain not only through contaminated pork, but also through edible crops where bacteria-contaminated manure has been used as a fertilizer. Thus, an essential step in providing pathogen-free products to consumers is reduction of the risk of food-

borne disease on the farm level. Reduction of on-farm *Salmonella* contamination by the use of antimicrobials alone is not a sustainable way to control disease for two reasons: one, bacteria can become resistant to approved antimicrobials rendering them less effective in controlling disease, and second, pigs can be infected with *Salmonella* not only at the farm, but also during transportation and/or lairage. A potentially more effective method of addressing pre-harvest food safety is through genetic improvement of disease resistance in animals. This project aimed to identify genetic variants in porcine candidate genes that potentially control the differences in disease response seen in individual animals. To analyze genetic variants, four porcine populations with *Salmonella* shedding data were used: 1) a field pig population of 462 pigs including 163 *Salmonella* shedding pigs and 299 their farm-mate negative controls; 2) NADC-40 pigs and 3) NADC-77 pigs *Salmonella* Typhimurium challenge populations; and 3) 228 pigs IAH-Compton *Salmonella* Choleraesuis challenge population. Based on our previous research, multiple genetic variants were selected in genes that respond/regulate porcine response to *Salmonella*. In this research we genotyped 62 new genetic variants or single nucleotide polymorphisms (SNPs) in 53 genes using the four populations. We found 28 segregating genetic variants with minor allele frequency of 15 % or higher in at least 2 of the populations. Statistical analysis revealed several SNPs associated with *Salmonella* fecal shedding or tissue colonization. Fecal shedding associated SNPs were CCT7 #3 ($p=0.041$), GNG3 ($p=0.029$), PGD ($p=0.047$) and HP #2 ($p=0.0002$) in the field population, AMT ($p=0.005$) in the NADC-40 pig population and PGD ($p=0.001$) in the NADC-77 pigs population. In the IAH-Compton population three SNPs in the ACP2 gene were associated with *Salmonella* burden in spleen (ACP2#1, $p=0.013$, ACP2#2, $p=0.026$, ACP2#3, $p=0.033$); and in the NADC-40 pig population, a SNP in the EMP1 ($p=0.002$) gene was associated with bacterial load in ileo-cecal lymph node.

The molecular variation in our phenotype-associated genes can be useful in selecting animals that shed fewer bacteria and are less likely to cause pen-mate *Salmonella* contamination in a farm and on a

slaughter plant. Thus, this research provided novel information for the pig industry to implement in strategies for selecting pigs with reduced shedding and/or disease susceptibility.

Contact: Christopher Tuggle, 515-294-4252, email: cktuggle@iastate.edu.