Integrating Comparative Expression Profiling Data and Association of SNPs with Salmonella Shedding for Improved Food Safety and Porcine Disease Resistance

J. J. Uthe
Iowa State University

Shawn M. D. Bearson
United States Department of Agriculture

L. Qu
Iowa State University

Jack C. Dekkers
Iowa State University, jdekkers@iastate.edu

See next page for additional authors.

Follow this and additional works at: http://lib.dr.iastate.edu/vdpam_pubs
Part of the Epidemiology Commons, Food Processing Commons, Genetics and Genomics Commons, Large or Food Animal and Equine Medicine Commons, and the Veterinary Preventive Medicine, Epidemiology, and Public Health Commons

The complete bibliographic information for this item can be found at http://lib.dr.iastate.edu/vdpam_pubs/18. For information on how to cite this item, please visit http://lib.dr.iastate.edu/howtocite.html.
Integrating Comparative Expression Profiling Data and Association of SNPs with Salmonella Shedding for Improved Food Safety and Porcine Disease Resistance

Abstract

Salmonella in swine is a major food safety problem, as the majority of US swine herds are Salmonella-positive. Salmonella can be shed from colonized swine and contaminate (i) neighbouring pigs; (ii) slaughter plants and pork products; (iii) edible crops when swine manure is used as a fertilizer; and (iv) water supplies if manure used as crop fertilizer runs off into streams and waterways. A potentially powerful method of addressing pre-harvest food safety at the farm level is through genetic improvement of disease resistance in animals. In this research, we describe a successful strategy for discovering genetic variation at candidate genes associated with disease resistance in pigs. This involves integrating our recent global gene expression analysis of the porcine response to Salmonella with information from the literature about important candidate genes. We identified single-nucleotide polymorphisms (SNPs) in these functional candidate genes and genotyped three independent pig populations that had data on Salmonella faecal shedding or internal burden (total n = 377) at these loci. Of 31 SNPs genotyped, 21 SNPs segregated in at least two populations with a minor allele frequency of 15% or greater. Statistical analysis revealed thirteen SNPs associated with Salmonella faecal shedding or tissue colonization, with an estimated proportion of false positives (PFP) ≤0.2. The genes with associated SNPs included GNG3, NCF2, TAP1, VCL, AMT, CCR1, CD163, CCT7, EMP1 and ACP2. These associations provide new information about the mechanisms of porcine host response to Salmonella and may be useful in improving genetic resistance to this bacterium.

Keywords
disease resistance, Salmonella, shedding, single-nucleotide polymorphism

Disciplines
Epidemiology | Food Processing | Genetics and Genomics | Large or Food Animal and Equine Medicine | Veterinary Preventive Medicine, Epidemiology, and Public Health

Comments
This article is from Animal Genetics 42 (2011): 521, doi:10.1111/j.1365-2052.2010.02171.x.

Rights
Works produced by employees of the U.S. Government as part of their official duties are not copyrighted within the U.S. The content of this document is not copyrighted.

Authors
J. J. Uthe, Shawn M. D. Bearson, L. Qu, Jack C. Dekkers, Dan Nettleton Dr., Y. Rodriguez Torres, Annette M. O’Connor, James McKean, and Christopher K. Tuggle

This article is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/vdpam_pubs/18
Integrating comparative expression profiling data and association of SNPs with *Salmonella* shedding for improved food safety and porcine disease resistance

J. J. Uthe*, S. M. D. Bearson†, L. Qu*,+‡, J. C. Dekkers*,§, D. Nettleton‡, Y. Rodriguez Torres*, A. M. O’Connor‡, J. D. McKeen‡ and C. K. Tuggle*,*

*Department of Animal Science, Iowa State University, Ames, IA 50014, USA. †USDA, ARS, National Animal Disease Center, Ames, IA 50010, USA. ‡Department of Statistics, Iowa State University, Ames, IA 50014, USA. §Center for Integrated Animal Genomics, Iowa State University, Ames, IA 50014, USA. ¶Department of Veterinary Diagnostic and Production Animal Medicine (VDPAM), Iowa State University, Ames, IA 50014, USA.

Summary

*Salmonella* in swine is a major food safety problem, as the majority of US swine herds are *Salmonella*-positive. *Salmonella* can be shed from colonized swine and contaminate (i) neighbouring pigs; (ii) slaughter plants and pork products; (iii) edible crops when swine manure is used as a fertilizer; and (iv) water supplies if manure used as crop fertilizer runs off into streams and waterways. A potentially powerful method of addressing pre-harvest food safety at the farm level is through genetic improvement of disease resistance in animals. In this research, we describe a successful strategy for discovering genetic variation at candidate genes associated with disease resistance in pigs. This involves integrating our recent global gene expression analysis of the porcine response to *Salmonella* with information from the literature about important candidate genes. We identified single-nucleotide polymorphisms (SNPs) in these functional candidate genes and genotyped three independent pig populations that had data on *Salmonella* faecal shedding or internal burden (total n = 377) at these loci. Of 31 SNPs genotyped, 21 SNPs segregated in at least two populations with a minor allele frequency of 15% or greater. Statistical analysis revealed thirteen SNPs associated with *Salmonella* faecal shedding or tissue colonization, with an estimated proportion of false positives (PFP) \( \leq 0.2 \). The genes with associated SNPs included *GNG3*, *NCF2*, *TAP1*, *VCL*, *AMT*, *CCR1*, *CD163*, *CCT7*, *EMP1* and *ACP2*. These associations provide new information about the mechanisms of porcine host response to *Salmonella* and may be useful in improving genetic resistance to this bacterium.

Keywords disease resistance, *Salmonella*, shedding, single-nucleotide polymorphism.
significant genetic control (Sellwood 1979; Rothschild 2007). Control of foodborne Salmonella within the farm-retail continuum is a complex issue. In addition to meat contamination, when manure from farms with Salmonella carrier animals is used as fertilizer, it can contaminate edible crops and/or run into human water sources, posing additional food safety issues (Guan & Holley 2003). Thus, reduction in the risk of foodborne disease at the farm level is an essential step in providing pathogen-free products, including both meat and non-meat products, such as fruits and vegetables, to consumers. A variety of methods to reduce on-farm Salmonella prevalence have been employed, including feed management practices, vaccination, and the use of antibiotics; however, many such methods have economic or public health issues (Denagamage et al. 2007; O’Connor et al. 2008; Perron et al. 2008). A potentially effective method of addressing Salmonella control and pre-harvest food safety issues is through genetic improvement of disease resistance in animals to decrease the incidence of Salmonella at the beginning of the food chain.

Genetic variation in the porcine immune response has been reported in several studies, and it has been shown that porcine immunity and/or disease resistance is under significant genetic control (Sellwood 1979; Rothschild et al. 1984; Meijerink et al. 2000; Gibson & Bishop 2005; Galina-Pantoja et al. 2006; Petry et al. 2007; Clapperton et al. 2009; Reiner et al. 2007, 2010). Several quantitative trait loci (QTL) associated with immune capacity in the pig have been identified, including QTL for total white blood cell and leucocyte numbers, mitogen-induced proliferation, levels of pre-vaccination antibodies to Escherichia coli, and stress-induced immune response. In addition, QTL have been found for porcine humoral innate immune response and cytokine concentration (Edfors-Lilja et al. 1998, 2000; Reiner et al. 2008; Wimmers et al. 2009; Lu et al. 2010). Several immune traits have been suggested as predictors of swine health status, including acute phase protein alpha-1 acid glycoprotein levels and numbers of peripheral blood mononuclear leucocyte subsets such as CD11R1+ cells (Clapperton et al. 2009). In addition, several inherited immunological traits such as number and function of polymorphonuclear leucocytes (PMNs) were shown to affect porcine resistance to Salmonella (van Diemen et al. 2002). Further, chromosomal regions have been identified that are associated with S. enterica serovar Choleraesuis (SC) spleen and liver colonization at 7 days post-challenge in F2 offspring of a boar thought to be heterozygous for genes involved in Salmonella susceptibility (Galina-Pantoja et al. 2009). Effects of selection on porcine adaptive immune response have been analysed by developing high immune response and low immune response porcine lines selected for a combination of humoral and cellular response measures (Mallard et al. 1992; Wilkie & Mallard 1999; Crawley et al. 2005). Altogether, these studies indicate that there is measurable genetic control of immune traits that leads to variation in disease progression and pathogenesis. However, the specific genes, genetic variants and pathways that are involved in controlling such variation in porcine response and resistance to Salmonella have not been identified. Recent microarray gene expression profiling from pig mesenteric lymph node examined the porcine transcriptional response to SC and S. enterica serovar Typhimurium (ST) infections (Wang et al. 2007, 2008). In the research reported here, we used this global gene expression data to select candidate genes and analyse SNPs in genes whose expression levels respond to both Salmonella serovars. These genes and their variants are likely to be important in host response to Salmonella serovars found in the field. We also genotyped three porcine populations with a total of 377 animals and identified associations of the SNPs with Salmonella shedding and/or tissue colonization.

Materials and methods

Porcine populations

The three porcine populations with qualitative or quantitative Salmonella shedding and/or tissue colonization phenotypes that were used in this study are as follows:

Field population

A survey of swine farms in Iowa, USA, resulted in the collection of tissues as well as Salmonella faecal shedding data for >7000 animals (Wang et al. 2010). Faecal samples were collected seven days before marketing, and samples were tested for Salmonella presence as described later. Within a subset of approximately 1000 pigs of this population, we obtained tissue samples and isolated genomic DNA from 52 faecal positive pigs and 57 faecal negative control pigs. These controls were selected from the cohort of animals collected on that farm and day where a positive case was found; where possible, we used an animal collected near the positive sample. For genotyping, we used these 109 field population pigs.

IAH-Compton population

This experimental challenge population contained a total of 228 pigs that included multiple specific sire families of commercial pigs as previously described (van Diemen et al. 2002). Briefly, first-generation offspring of two boars that were potentially susceptible and resistant as well as two unknown boars were experimentally infected with S. choleraesuis followed by enumeration of bacteria in liver and spleen 7 days post-infection. The phenotypic data and DNA samples from this population were a gift from the Pig Improvement Company.
National Animal Disease Center (NADC) population

This experimental challenge population contained 40 mixed breed pigs that were experimentally inoculated with S. Typhimurium: details on this experiment have been reported previously (Uthe et al. 2009). Briefly, at days 2, 7, 14 and 21 post-inoculation (pi), quantitative Salmonella faecal shedding data were collected. At day 21 pi, ileo-caecal lymph nodes were collected, and qualitative Salmonella presence or absence in this tissue was determined as described (Utthe et al. 2009). All procedures involving animals were lawful and approved by the USDA, ARS, NADC Animal Care and Use Committee.

Sample collection and Salmonella bacteriology

For the field population, individual faecal samples (20–30 g) were collected into labelled plastic bags at the same time as animals were tattooed (Wang et al. 2010). Matched belly flap samples were collected at the abattoir from dressed carcasses using the unique slap tattoo number. Belly flap samples were placed in plastic bags and frozen for later DNA preparation. To identify pigs shedding Salmonella, qualitative bacteriology was performed as follows: 10 g samples of swine faeces were assayed in duplicate using Salmonella enrichment and selective media as previously described (Hurd et al. 2002). Positive isolates were confirmed by serogroup antisera agglutination assays.

DNA isolation

To isolate DNA from belly flaps (field) and liver tissues (NADC), about 20 mg of tissue samples was digested with proteinase K (Invitrogen, Carlsbad, CA, USA). DNA from tissue lysates was extracted using Wizard SV genomic DNA purification system (Promega, Madison, WI, USA) according to the manufacturer’s protocol. DNA was quantified by Nanodrop (Thermo Scientific, Wilmington, DE, USA), and DNA quality was checked by A260/A280 ratios and agarose gel electrophoresis.

Selection of functionally relevant genes for SNP analysis

Our gene and putative SNP selection criteria were as follows: 1 Genes that are differentially regulated early, at 8–48 h during both ST and SC infections in pigs (Wang et al. 2007, 2008). All annotated genes that showed a difference in expression of at least 1.5-fold when comparing infected animals to uninfected controls and with a q-value <0.24 were considered, as these criteria we used in the original studies. 2 Genes that, based on published data (Jenner & Young 2005), are involved in the response of the organism to multiple bacterial pathogens.

3 Genes whose expression is known to be ‘non-tolerizeable’ or non-diminished during continuing re-infection, as such genes are more likely to be involved in clearance of bacterial pathogens (Foster et al. 2007).

This approach identified 223 candidate genes, with 32 of them found in the Jenner and Young list of common host response genes, two genes reported as non-tolerizeable, as well as seven genes coming from the combination of the lists above (Fig. 1 and Table S2). This candidate gene list was expanded by adding genes that are differentially expressed in porcine whole blood RNA 2 days after inoculation with ST (Huang et al., manuscript in preparation, Table S2). In this experimental ST challenge study, pigs that stopped shedding the bacteria after day 7 of the experiment were called low shedders, and pigs that continued shedding until the end of the 21-day study were called persistent shedders. The 62 genes were chosen based on Affymetrix analysis and were those genes with q-value <0.1 for shed class by infection interaction and ≥1.5-fold change because of infection, expressed in the opposite direction for low shedder versus persistent shedder pigs (Huang et al., manuscript in preparation).

Identification of SNPs in functional candidate genes and SNP genotyping

Putative SNPs in the selected genes were identified using the Pig Expression Data Explorer (PDE, http://pede.dna.affrc.go.jp/) or Dana-Farber Cancer Institute (DFCI) Pig Gene Index databases (http://compbio.dfc.fraction.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=pig). From the databases, we selected sequences with putative SNPs, requiring a minimum of three counts of the minor allele that was different from the allele in the consensus sequence reported by the database.

DNA genotyping was performed by PCR–restriction fragment length polymorphism (PCR–RFLP), tetra primer

© 2011 The Authors, Animal Genetics © 2011 Stichting International Foundation for Animal Genetics, 42, 521–534
amplification refractory mutation system (ARMS)-PCR and Sequenom technologies. Initially, we tested 13 predicted SNPs by PCR amplification and product sequencing of representative DNA pools from the NADC population (DNA Sequencing and Synthesis facility, Iowa State University, Ames, IA, USA). Analysis by RFLP was used to genotype the confirmed SNPs in the NADC and IAH-Compton populations (Table 1). The PCR was prepared in 15 μl volumes, and the PCR conditions were as follows: 5 min at 95 °C followed by 35 cycles of 1 min at 94 °C, 30 s at 55 °C and 1 min at 72 °C, and a final extension of 10 min at 72 °C. Enzymatic digestions followed by agarose gel electrophoresis to determine genotypes were carried out using standard protocols.

To increase genotyping throughput and efficiency, we used Sequenom MassARRAY technology (MassARRAY Compact System; Sequenom Inc., San Diego, CA, USA) for the remaining SNPs. Selection of SNPs for multiplex reaction, primer design (Table S1) and genotyping was performed using iPLEX reagent kit and the Sequenom Typer 3.1 software according to the manufacturer’s instructions (Sequenom, http://www.sequenom.com). Assays that had a minimum of 80% of genotyping calls and minor allele frequency (MAF) of at least 15% were subsequently statistically analysed. For quality control, within the set of SNPs genotyped by Sequenom, we included 13 markers with known genotypes based on PCR–RFLP analysis of the NADC population. The accuracy of Sequenom genotype data compared to PCR–RFLP results was 99%. Finally, tetra primer ARMS-PCR was used to genotype the CD163 #2 SNP (NM_213976.1:c.2685C>A), as data for this SNP from Sequenom analysis were incomplete. Each PCR contained 30 ng of DNA template. 10 pmol of each inner primer, including the C allele-specific forward (5′GTGTTCACTGTCC TAAAGGACCTGACGCC) and the A allele-specific reverse primer (5′ATGGAGATGAAGGCGCAGTGGCCTGATG) and 1 pmol of each outer primer (forward: 5′ GGACATCGACCCTCGATCCTCCAGACAG and reverse: 5′ATGTTGATCCGTGCTTCCTGAGGCGGT). The above PCR primers were designed using tetra primer ARMS-PCR primer design software at http://cedar.genetics.soton.ac.uk/public_html/primer1.html (Ye et al. 2001). PCR reactions were prepared in 10 μl volumes using GoTaq Hot Start Green Master Mix (Promega, Madison, WI, USA). Touchdown PCR conditions were as follows: 2 min at 95 °C followed by 35 cycles of 1 min at 95 °C, 1 min at 75 °C for the first cycle, decreasing by 1 °C each cycle until the temperature of 60 °C was reached and continuing at that annealing temperature for the rest of the cycles; then 1 min at 72 °C and final extension 10 min at 72 °C. PCR products were detected by electrophoresis on an agarose gel consisting of 1:1 low EEO (Fisher Scientific, Pittsburg, PA, USA) to 1:1 TAE agarose. Enzymatic digestions were carried out using tetra primer ARMS-PCR primer design software at http://cedar.genetics.soton.ac.uk/public_html/primer1.html (Ye et al. 2001). PCR reactions were prepared in 10 μl volumes using GoTaq Hot Start Green Master Mix (Promega, Madison, WI, USA). Touchdown PCR conditions were as follows: 2 min at 95 °C followed by 35 cycles of 1 min at 95 °C, 1 min at 75 °C for the first cycle, decreasing by 1 °C each cycle until the temperature of 60 °C was reached and continuing at that annealing temperature for the rest of the cycles; then 1 min at 72 °C and final extension 10 min at 72 °C. PCR products were detected by electrophoresis on an agarose gel consisting of 1:1 low EEO (Fisher Scientific, Pittsburg, PA, USA) to MetaPhor (Cambrex, East Rutherford, NJ, USA) agaroses. The presence of a 99-bp PCR product was indicative of the C allele.

Table 1 RFLP analysis of selected polymorphisms.

<table>
<thead>
<tr>
<th>Gene name and SNP</th>
<th>PCR primers, 5′-3′</th>
<th>Enzyme</th>
<th>PCR product (bp)</th>
<th>Digestion product sizes (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD47/IAP, NM_213982.1:c.2581G&gt;A</td>
<td>f: GTGCACCTGTGAAGTATGCCAC; r: CAGCAAAACCTTGTTGCCAGGAGT</td>
<td>Tsp509 I</td>
<td>370</td>
<td>A: 103, 93, 74, 48, 42, 10 bp; G: 151, 93, 74, 42, 10 bp</td>
</tr>
<tr>
<td>VCP #3, NM_214280.1:c.1318G&gt;A</td>
<td>f: GTGCCTTGACGGAGGAGT; r: TGGTCAAGGGCAGCATCTCG</td>
<td>BstY I</td>
<td>112</td>
<td>G: 64, 48 bp; A: no digestion</td>
</tr>
<tr>
<td>LCP1 #1, XM_001929138.1:c.644T&gt;C</td>
<td>f: CTGCCCTGCTGCTCCTG; r: GATAAACCTGCTAAAGCTTAGC</td>
<td>BtsC I</td>
<td>107</td>
<td>T: 82, 24; C: no digestion</td>
</tr>
<tr>
<td>CCT7 #3, NM_001170522:c.1520C&gt;T</td>
<td>f: GAGGACATTGCTGACGCC; r: GTGAGAAGACAGGGCTCAAAAC</td>
<td>BtsC I</td>
<td>299</td>
<td>T: 188, 111 bp; C: no digestion</td>
</tr>
<tr>
<td>CCR1, NM_001001621.1:c.2878G&gt;A</td>
<td>f: CACATCAGCAGAACACACACAC; r: GATTATTGTTCTGGAGAAATGAT</td>
<td>Tsp509 I</td>
<td>474</td>
<td>G: 224, 127, 123 bp; A: 186, 127, 123, 38 bp</td>
</tr>
<tr>
<td>CD163 #3, NM_213976.1:c.3066T&gt;C</td>
<td>f: ATATGGCTCAAAAGCTTAGAAGT; r: GGGATTCCCTGCGGCTCCCTG</td>
<td>Afl III</td>
<td>504</td>
<td>T: 109, 207, 184 bp; C: 316, 184 bp</td>
</tr>
<tr>
<td>NCF2 #1, NM_001123142.1:c.1419C&gt;T</td>
<td>f: ACCAAGAACACTCCTAAG; r: ATAGTGTCCTCCCTGCTGGAAGA</td>
<td>Acl I</td>
<td>101</td>
<td>T: 59, 30, 12 bp; C: 71, 30 bp</td>
</tr>
<tr>
<td>TYROBP, NM_214202.1:c.232C&gt;T</td>
<td>f: TGTTGCTGACCCCTTCTC; r: CTCAGCGATGTGTTCTGCTC</td>
<td>Btg I</td>
<td>253</td>
<td>T: 70, 183 bp; T: no digestion.</td>
</tr>
<tr>
<td>TAP1 #2, NM_001044581.1:c.448A&gt;G</td>
<td>f: CTCTCCTGCTAAGGCGCCCTTCC; r: CAGAGCCTCCTGACCTTCCTG</td>
<td>HpyCH4V</td>
<td>702</td>
<td>A: 587, 53, 40, 22 bp; G: 587, 93, 22 bp</td>
</tr>
<tr>
<td>ACP2 #3, ITC.59388.1004:c.644G&gt;A</td>
<td>f: TTGGACAGCTGCAAGAGCAG; r: GAGTGGTTCATAGCCATCACCC</td>
<td>Sau96 I</td>
<td>226</td>
<td>A: 153, 74, 42 bp; G: 111, 74, 42 bp</td>
</tr>
</tbody>
</table>

The PCR–RFLP assay for CCT7 #2 SNP was reported in Uthe et al. 2009.
allele, a 110-bp product indicated the A allele, and the outer primers always produced a 155-bp product.

**Statistical analysis**

Statistical association analysis of genotype and *Salmonella* shedding and/or tissue colonization phenotype was performed in the statistical computing environment R, with all code available upon request. Because of small sample sizes and non-normal distributions, permutation-based hypothesis testing was used to assess the significance of association between each SNP and each phenotype in each population separately. Within each population, the test statistic used for permutation testing was carefully chosen based on the nature of the phenotype and the structure of block factors unique to each population, as described in later subsections. Regardless of the test statistic selected, permutation testing was carried out for each population, phenotype and SNP as follows. First, the test statistic was computed for the original data. Then, SNP genotypes were randomly shuffled 19,999 times within the levels of the available block factor(s). The test statistic was recomputed for each of these 19,999 data sets. The resulting 20,000 test statistics provided a reference distribution under the null hypothesis of no association. A permutation P-value was computed as the proportion of these 20,000 statistics that were as extreme as or more extreme than the statistic observed for the original data.

Besides analysing each SNP of interest separately, SNPs that lie within the same gene were also assessed for joint association. Each joint SNP was first treated as a single multilevel factor (up to nine levels) and was analysed in the same way as other SNPs. To further improve statistical power by recovering some inter-block information, the permutation P-value of this joint SNP together with permutation P-values from the individual SNPs in the same gene were then combined using Fisher’s combination function, i.e. the geometric mean function. The same combination procedure was applied to all other 19,999 permutations. The final permutation P-value for the joint SNP was computed as the proportion of permutations that had a combined P-value no larger than the one obtained in the original data (Pesarin 2001).

Because of multiple hypothesis tests performed in the analyses, we controlled SNP false discovery rates (FDR) from the set of P-values for each trait in each population separately, using a procedure similar to Storey & Tibshirani (2003). The default smoother method used by Storey & Tibshirani (2003) was not stable because we had only dozens of P-values. Thus, we chose to set the lambda parameter discussed by Storey & Tibshirani (2003) to be the observed P-value such that the difference between the P-value itself and the empirical distribution function evaluated at this P-value was maximum. When the number of P-values is large, this modification results in more conservative FDR control, i.e. it controls FDR by providing a wider safety margin. Note that this FDR estimator also provides a natural estimate of the proportion of false positives (PFP) (Fernando et al. 2004). Hence, even though the multiple testing procedure was applied to each phenotype-by-population combination separately, the procedure also provides a reasonable estimate of PFP when all hypotheses in the currently study are considered together.

Details specific to each population are as follows:

### Field population

The binary *Salmonella* shedding status trait was used as the response variable. A logistic regression model was used with farm-and-visit combination as a block factor and SNP as a fixed factor of interest. Because of complete separation of the binary trait with either block or the SNP of interest, not all maximum likelihood estimates of parameters were finite. Hence, Firth’s penalized likelihood method was used for parameter estimation (Firth 1993). The modified (Rao’s) score statistic was computed for the SNP of interest, based on Firth’s modified score equation (Heinze & Schemper 2002).

### IAH-Compton population

Two log-transformed count traits, spleen counts (splC) and liver counts (livC), were analysed separately for each SNP. Each of 39 combinations of group, sire and dam was treated as a distinct level of a block factor. The test statistic was chosen to be the weighted average of the within-SNP-genotype pairwise Euclidean distances (see the MRBP statistic proposed by Mielke & Berry 2007).

### NADC population

The phenotypes for this population included a univariate binary ileo-caecal lymph node (ICLN) trait and a 4-dimensional (four time points pi) shedding count trait. Sow was the only block factor. For the binary ICLN trait, a permutation-based likelihood ratio test was used to separately test the association of each SNP with the binary response in a logistic regression model. For each SNP being analysed, log likelihood ratio statistic was calculated by comparing the maximized log likelihood from the model with the sow as the only fixed factor and the model with both the sow and the SNP being analysed as fixed factors. The 4-dimensional shedding count trait was transformed by taking natural logarithms after adding one to each count to avoid infinite logarithms. The test statistic for the 4-dimensional shedding count trait was calculated as the weighted mean of within-genotype average pairwise Euclidean distances of pigs, with weights being the number of pigs within the same SNP genotype (see the MRBP statistic of Mielke & Berry 2007) (Appendix S1).
Results

Identification of functionally relevant candidate genes for improving resistance to Salmonella

Expression profiling by our research team of the porcine response to Salmonella (Wang et al. 2007, 2008) has identified genes differentially regulated in pig mesenteric lymph nodes at various stages during Salmonella infection from acute (8 hr post-inoculation (p.i.)) to the chronic (21 days p.i.) stages. This microarray analysis of mesenteric lymph node gene expression was chosen as the most comprehensive published study on porcine transcriptional response to Salmonella and provides large lists of genes responding to two different Salmonella serovars (SC and ST) to be screened to select candidate markers for SNP analysis. Aiming to identify genes that would be involved in controlling porcine response to infection under different conditions at different farms, we also attempted to select genes that commonly respond to bacterial pathogens using information from the literature. Global gene expression analysis in response to Salmonella and other pathogens is available in other animal species including chicken (Zhou & Lamont 2007; Chiang et al. 2008; Zhang et al. 2008). Cross-species comparison of microarray data is complex because of the variety of tissues or cells and various microarray platforms used, as well as the logistics of identifying orthologs to microarray elements in other species. Thus, we focused on using our Salmonella response transcriptions data that is directly relevant to porcine resistance. However, as an additional level for selecting genes, we opted to use a well-described common host response gene list by Jenner & Young (2005). The authors analysed a wide range of microarray experiments in a variety of human cells, thus encompassing genes generally responsive to pathogenic stimulus in mammalian cells. As an additional level in choosing genes, we used a non-tolerizable gene list described by Foster et al. (2007), attempting to prioritize genes that may be involved in clearance of the bacteria. Figure 1 shows the total number of genes found under each of the above criteria and the intersections of each. In addition, we selected 62 genes that by Affymetrix microarray analysis were found to be differentially regulated in porcine whole blood two days after inoculation with ST in two different classes of pigs that differ in faecal shedding phenotypes (Huang et al., manuscript in preparation, Table S2; see Methods). In total, we selected 285 genes for SNP identification and analysis.

SNP genotyping in functionally relevant genes

From the selected candidates, we chose genes that, based on PEDE or DFCI databases, had a putative SNP with a minimum of three counts of the minor allele that was different from the allele on the consensus sequence. In many cases, the SNP was confirmed by initial sequencing. Thus, we selected for genotyping a set of 54 SNPs in 41 genes responding to acute Salmonella infection in pigs. A subset of 29 SNPs was selected for multiplex genotyping by Sequenom software, and genotyping calls were obtained for 23 SNPs through Sequenom analysis. As a check on Sequenom call accuracy, we checked eleven SNPs in this set that had been genotyped earlier in our populations by PCR–RFLP. For ten SNPs, Sequenom and previous genotype agreement frequency was 99.8%, while the assay for one SNP did not produce Sequenom results. Several SNPs that had MAF ≥15% but low Sequenom call rate were manually completed to achieve at least 80% of allele call rates by either RFLP or tetra primer ARMS-PCR technology. In total, we genotyped 31 SNPs and confirmed 28 putative SNPs in 23 functionally relevant genes (Table 2). Twenty-four SNPs in 21 genes had a MAF of 15% or greater in one of the three populations, and 21 SNPs in eighteen genes had MAF of at least 15% in at least two of the three populations (Table 3).

Association of SNPs with Salmonella shedding or tissue colonization

Several SNPs were associated with Salmonella shedding or tissue colonization phenotypes. SNPs with an estimated q-value ≤0.2 were selected as statistically significant, where q-value is the estimated PFP; P-values for significant SNPs ranged from 0.002 to 0.088. In the NADC-40 population, the SNP in AMT was associated with quantitative faecal shedding of S. Typhimurium over the time course of infection with a raw P-value 0.005 and estimated q-value 0.095 (Fig. 2). In the field population, qualitative Salmonella shedding status was associated with SNPs in GNG3 (P = 0.07, q = 0.19), NCF2 #1 (P = 0.023, q = 0.16), TAP1 #2 (P = 0.027, q = 0.16), VCL (P = 0.059, q = 0.19) (Fig. 3) and CCT7 #3 (P = 0.088, q = 0.20, data not shown) as well as combined SNPs in ACP2 #1 + 2 + 3 (P = 0.031, q = 0.16) and CCT7 #2 + 3 (P = 0.065, q = 0.19) (Fig. 4). The association of SNPs with Salmonella tissue colonization phenotype was assessed using the NADC

Table 2 Summary of genotyping and association analyses.

<table>
<thead>
<tr>
<th>Porcine populations</th>
<th>NADC-40</th>
<th>Field</th>
<th>IAH-Compton</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNPs tested by Sequenom and/or RFLP</td>
<td>31</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td>Genes tested</td>
<td>25</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Informative SNPs found (MAF ≥15%)</td>
<td>18</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Informative genes found</td>
<td>14</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td># of SNPs associated with shedding/tissue colonization</td>
<td>2</td>
<td>5</td>
<td>8</td>
</tr>
</tbody>
</table>

1SNP #2 for CCT7 was found previously (Uthe et al. 2009) and is not counted in this list.

2SNPs in the ACP2 and AMT genes were associated in two populations.
Table 3: Genotyping analysis using the 3 porcine populations.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Nomenclature</th>
<th>NADC-40 population, n = 40</th>
<th>Field population, n = 109</th>
<th>Compton population, n = 228</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Allele frequency %</td>
<td>% pigs genotyped</td>
<td>Faecal counts P-value</td>
<td>ICLN counts P-value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACP2 #1</td>
<td>ITC.59388.1004:c.195G&gt;C 1</td>
<td>G-36; C-64 93 0.359 0.933</td>
<td>G-22; C-78 91 0.189</td>
<td>G-25; C-75 89 0.011</td>
</tr>
<tr>
<td>GNG10</td>
<td>TC251588:c.659A&gt;C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
<tr>
<td>GNG3</td>
<td>TC26007:1.619T&gt;C-C</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>GSA3 #3</td>
<td>TC238734:c.621T&gt;G-C</td>
<td>C-31; T-69 100 0.308 0.611</td>
<td>C-85; A-15 93 0.291</td>
<td>C-82; A-18 94 0.002</td>
</tr>
</tbody>
</table>

1The ITC sequence comes from annexdb.org database.
2SNP was analysed previously.
*Statistically significant, with q ≤ 0.2.
and IAH-Compton populations. In the NADC population, a SNP in the \textit{EMP1} gene was associated with qualitatively measured \textit{Salmonella} load in the ileo-caecal lymph node at 21 days post-inoculation with a raw $P$-value of 0.002 and an estimated $q$-value of 0.041 (Fig. 5). In the IAH-Compton population, three compound SNPs in the ACP2 gene were associated with quantitatively measured \textit{Salmonella} burden in spleen ($P = 0.013$, $q = 0.138$) and liver ($P = 0.089$, $q = 0.15$) (Fig. 6 and data not shown). In this population, animals having homozygous G/G (ACP2, SNP#1, spleen $P = 0.011$, $q = 0.035$; liver $P = 0.093$, $q = 0.15$), T/T (ACP2, SNP #2, spleen $P = 0.024$, $q = 0.035$) and A/A (ACP2, SNP#3, spleen $P = 0.022$, $q = 0.035$; liver $P = 0.079$, $q = 0.15$) genotypes had the highest tissue colonization at 7 days post-inoculation with \textit{S. choleraesuis}. A SNP in the AMT gene ($P = 0.054$, $q = 0.068$, data not shown) was associated with \textit{Salmonella} burden in spleen, while a SNP in the CCR1 gene ($P = 0.021$, $q = 0.11$) was associated with bacterial numbers in liver in the IAH-Compton population (Fig. 7). Bacterial numbers in liver and spleen were associated with \textit{CD163} SNP #2 (liver: $P = 0.047$, $q = 0.13$) and SNP #3 (liver: $P = 0.014$, $q = 0.11$; spleen: $P = 0.002$, $q = 0.018$) (Fig. 7). The combined \textit{CD163} SNP #1 + 3 genotypes were associated with \textit{Salmonella} burden in both liver and spleen (liver: $P = 0.038$, $q = 0.13$; spleen: $P = 0.017$, $q = 0.035$; data not shown).

**Discussion**

This research focuses on the long-term goal of using the power of genomics to identify genetic variants associated with decreased tissue colonization or \textit{Salmonella} shedding in

---

**Figure 2** Association of \textit{AMT} with \textit{Salmonella} faecal shedding in the NADC population, $P = 0.005$, $q = 0.095$. The number below the genotype indicates the number of pigs with that genotype in the population. Y-axis indicates LS means of natural log-transformed \textit{Salmonella} counts+1.

**Figure 3** Association of SNPs with \textit{Salmonella} faecal shedding in the field population. The number of pigs with each genotype is indicated on the bar graphs. Y-axis represents the fitted probability of positive faecal shedding (see Methods).

**Figure 4** Association of combined SNPs in ACP2 and CCT7 genes with \textit{Salmonella} faecal shedding in the field population. The number of pigs with each genotype is indicated on the bar graphs. Genotype i.e. CCTAAG means SNP 1 C/C + SNP#2 T/A + SNP#3 A/G. Y-axis represents the fitted probability of positive faecal shedding (see Methods).

© 2011 The Authors, Animal Genetics © 2011 Stichting International Foundation for Animal Genetics, 42, 521–534
live pigs. In addition, this is the first study to use global gene expression data and literature information on important immune response genes to select genetic polymorphisms for association with Salmonella colonization and shedding phenotypes in pigs. Selection of candidate genes for SNP analysis was an essential step in an investigation of this scale that aimed to understand porcine genetic resistance to Salmonella. Our previous analysis of global transcriptional regulation of swine response to Salmonella (Wang et al. 2007, 2008) was particularly useful in the gene selection process. By focusing on genes that are (i) differentially regulated in both S. Choleraesuis and S. Typhimurium infections, (ii) known to respond to many types of pathogens in human cells (Jenner & Young 2005) and/or (iii) genes recently shown to be insensitive to suppression (non-tolerizable) by repeated lipopolysaccharide (LPS) stimulation, we attempted to identify candidate genes involved in Salmonella as well as general porcine disease resistance. The criteria we applied were valuable in the selection of candidate genes for SNP and association analysis. In particular, four genes with confirmed SNPs were selected as belonging to a common host response gene list by Jenner & Young (2005), and SNPs in three of those genes were associated with Salmonella-related phenotypes, indicating the usefulness of this gene selection criterion. However, SNPs in genes that were non-tolerizable (nine genes out of 223) were either not found or not confirmed in our pig populations by initial pre-screening. In all, while global gene expression data provided novel lists of genes potentially controlling Salmonella colonization and shedding in pigs, general immune response pathways reported by Jenner & Young (2005) were useful to focus on particular immune genes. Of similar importance to defining the most appropriate candidate genes is developing test populations with relevant...
phenotypic records. It is very difficult to create substantive data sets for infectious disease work in large food animals, because of the great expense and substantial logistical work needed to collect data from natural or experimental challenge populations. We were able to gain access to the largest experimental challenge population samples that have been published for Salmonella infection in pigs, the IAH-Compton study (van Diemen et al. 2002), as well as develop an additional challenge population (NADC). We also sampled a field population, in the process creating a unique data set for genetic association with naturally occurring infections in the farm setting. These three porcine populations provided a total of 377 pigs with quantitative as well as qualitative phenotypes related to exposure to a variety of Salmonella, including S. Typhimurium (NADC), S. Choleraesuis (IAH-Compton) and untyped Salmonella from Iowa farms (field population).

The analysis revealed several SNPs in our selected genes that were associated with Salmonella shedding or tissue colonization phenotypes. The genes with associated SNPs are involved in a variety of cellular functions. The AMT gene was associated with faecal Salmonella shedding over time in the NADC population, where pigs with G/G genotype had about 64-fold less bacteria compared to G/A pigs (Fig. 2). Consistent with this finding, we observed an association of AMT with Salmonella load in spleen in the IAH-Compton population as well; however, this association is based on a single pig with the A/A genotype that had about 120- to 300-fold higher numbers of bacteria compared with G/A and G/G pigs, respectively (data not shown). The AMT gene codes for aminomethyltransferase, which is involved in the mitochondria-specific enzyme system for cleavage of glycine (Applegarth & Toone 2006). Mutations in the AMT gene have been implicated in accumulation of glycine in body fluids, leading to a pathology known as non-ketotic hyperglycinemia (Kure et al. 2006). A role of AMT or its polymorphisms in infection has not been identified, yet we showed that this gene was upregulated 1.5-fold in whole blood from pigs that did not shed Salmonella in their faeces past 7 days of experimental inoculation and was 1.6-fold downregulated in pigs that shed bacteria throughout the 21-day experiment (Huang et al., manuscript in preparation).

Genotypes in several genes were associated with the frequency of Salmonella-positive animals in the field population (Figs 3 & 4). The GNG3 gene product (guanine nucleotide exchange factor gamma 3) is involved in a variety of receptor signalling pathways (Kelly et al. 2008). GNG3 also stimulates the effectiveness of CD4-T-cell dependent immune response, as shown by an increase in antibody production in wild type compared to GNG3 knock-out mice (Dubeykovskiy et al. 2006). Thus, GNG3 may play a role in Salmonella infection, as CD4-T cells are known to be involved in defence against Salmonella. In our research, the expression of GNG3 was diminished 1.5-fold in pigs that shed Salmonella throughout the 21-day experiment, while it was twofold increased in pigs that stopped shedding bacteria after 7 days p.i. (Huang et al. manuscript in preparation). Another Salmonella shedding-associated SNP was found in neutrophil cytotoxic factor 2 (NCF2), a gene encoding a cytotoxic subunit of neutrophil NADPH oxidase, which produces superoxide anions that are delivered to the neutrophil phagosome for their microbicidal effects (Ammons et al. 2007). Oxidative burst capacity of PMNs was shown to affect resistance of pigs to infection with S. Choleraesuis (van Diemen et al. 2002), and thus, NCF2 gene function may be directly involved in this process. In mice, NCF2 was suggested as a candidate gene for the Ity3 (immunity to Typhimurium) locus-mediated Salmonella resistance (Sancho-Shimizu & Malo 2006). Researchers found that variation in NCF2 sequence in mice resulted in reduction of superoxide production in response to infection with S. Typhimurium. While our permutation analysis indicated significant association of NCF2 SNPs with Salmonella shedding, only two pigs with a rare A/A genotype were 100% positive for Salmonella. This may reflect an insufficient sampling of the population or, potentially, some deleterious effect of the A/A genotype on swine health: the latter possibility is less likely. A SNP associated with Salmonella shedding was found in the transporter associated with antigen processing 1 (TAP1) gene that belongs to a superfamily of ATP-binding cassette transporters and is involved in transporting cytosolic peptides across the endoplasmic reticulum where they are loaded on major histocompatibility complex class I molecules (Schölz & Tampé 2009). The DNA for TAP1 was up-regulated about twofold at 24–48 h pi with S. Typhimurium and S. Choleraesuis. In addition, TAP1 is known to be involved in response to multiple pathogens in human cells. Polymorphisms in the TAP1 gene have been linked to effects on the outcome of viral or parasitic infections as well as cancers (Soundravally & Hoti 2008; Aquino-Galvez et al. 2008). We also found genotypes in the VCL gene associated with shedding. The VCL gene codes for vinculin, a cytoskeletal protein that, through integrins, can couple the extracellular matrix with the actin-myosin cytoskeleton and is therefore implicated in cell adherence and motility (Mierke 2009). As actin polymerization is an essential step in Salmonella-mediated intracellular invasion, vinculin may play an important role in this process by affecting actin dynamics (Finlay & Brumell 2000; Wen et al. 2009). We have found that VCL was upregulated twofold in pigs shedding Salmonella throughout the 21-day experiment and downregulated fivefold in pigs that cleared the bacteria after day seven pi (Huang et al., manuscript in preparation); however, a direct link between Salmonella infection and vinculin has not been made previously. Finally, CCT7 (chaperonin containing TCP1, subunit seven, encoded by CCT7) is a chaperone involved in cytoskeleton protein folding (Valpuesta et al. 2002). In our previous
SNP #2 (NM_001170522:c.1026G>A) of the CCT7 gene was associated with Salmonella shedding in the NADC population (Uthe et al. 2009). Here, we extended this observation by confirming an additional SNP #3 in the coding sequence of CCT7 and identified association of the combined CCT7 SNP #2 + #3 genotype with Salmonella incidence in the field population (Fig. 4).

A SNP associated with Salmonella load in ileo-caecal lymph node was found in the EMP1 gene, where pigs with the G/A EMP1 genotype were twice as likely to be ICLN-positive for Salmonella compared to G/G pigs (Fig. 5). In addition, polymorphism in EMP1 results in a non-synonymous amino acid substitution from alanine to threonine at position 130. EMP1 was chosen for this study because gene expression was induced in pigs by both S. Typhimurium (Wang et al. 2007) and S. Choleraesuis (Wang et al. 2008) 1.6-fold (48 h pi) and twofold (24 h pi), respectively. The EMP1 gene encodes epithelial membrane protein 1, which is involved in regulation of cell adhesion, cell signalling and cell communication (Jetten & Suter 2000; Wang et al. 2003). Cross-talk between EMP1 and epithelial growth factor receptor signalling has been described (Jain et al. 2005). Furthermore, EMP1 is a gene known to be involved in response to multiple bacterial pathogens (Jenner & Young 2005). However, a role of EMP1 in infectious disease is unknown.

Specific genotypes associated with both Salmonella burden in spleen and liver in the IAH-Compton population as well as Salmonella faecalis shedding in the field population were found in the ACP2 gene (Fig. 4 & 6). Approximately 20- to 50-fold more bacteria in spleen were found in animals that were homozygous for specific ACP2 SNPs when compared to the other homozygous genotype (Fig. 6). Several haplotypes could be predicted from the three observed compound homozygous genotypes in ACP2. Thus, pigs with GGTTAA genotype (GTA haplotype) are estimated to have a much higher spleen burden when compared to pigs without the GTA haplotype. For example, GGTTAA animals in the IAH-Compton population have about 32-fold more bacteria compared to CCAAGG animals (Fig. 6). Interestingly, the GGTTAA animals in the field population also have the highest estimated incidence of bacterial shedding, suggesting that the GTA haplotype could predict pigs with a diminished ability to control Salmonella infections. Microarray analysis revealed the upregulation of ACP2 expression by about twofold at 48 h post-inoculation with both S. Typhimurium and S. Choleraesuis (Wang et al. 2007, 2008). The gene encodes acid lysosomal phosphatase, which is expressed in lysosomal compartments in all tissues and is involved in phosphomonoester cleavage (Moss et al. 1995). One of our associated SNPs in this gene, ACP2 SNP #2, causes an amino acid substitution from glutamate to aspartate at position 187. A mutation in the ACP2 gene has been previously described in mice (Mannan et al. 2004; Sallig et al. 1997). It was found that the non-synonymous SNP causing a glycine to glutamate substitution (p.Gly244Glu) in murine ACP2 rendered the protein inactive while causing a variety of pathologies in mice, including growth retardation, ataxia-like phenotype, delayed hair appearance and lysosomal storage bodies in nucleated cells (Mannan et al. 2004). Even though ACP2 belongs to a list of genes involved in response of human cells to multiple bacterial pathogens (Jenner & Young 2005), an association of ACP2 polymorphisms with infection has not been previously described, and the effect of this novel identified non-synonymous SNP on ACP2 function in pigs remains to be revealed.

Several other SNPs were associated with Salmonella burden in liver (CCR1) or both liver and spleen (CD163) in the IAH-Compton population (Fig. 7). CCR1 encodes chemokine (C-C motif) receptor 1, which was sixfold upregulated in pigs inoculated with S. Choleraesuis at 48 h pi and twofold downregulated in pigs inoculated with S. Typhimurium at 8 h pi (Wang et al. 2007, 2008). CCR1 is expressed on T cells, monocytes, neutrophils, eosinophils, basophils (Charo & Ransohoff 2006) and has chemoattractant as well as inflammatory modulating functions (Hartl et al. 2008; Khan et al. 2001). Polymorphisms in the CCR1/CCR3/CCR2 region have been associated with coeliac disease in humans (Amundsen et al. 2010); this study is the first to demonstrate an association of CCR1 with bacterial infection in pigs.

Macrophage scavenger receptor CD163 is another gene associated with Salmonella infection phenotype, with pigs of A/A genotype at SNP #2 having about tenfold more bacteria in spleen and liver compared to C/C genotype animals (Fig. 7). Gene expression analysis revealed that CD163 was induced sixteen-fold in S. Choleraesuis-inoculated pigs at 48 h pi (Wang et al. 2008) and about twofold in S. Typhimurium-inoculated pigs at 24 h pi (Uthe et al. 2007). CD163 plays a homeostatic function by binding haemoglobin-haptoglobin complexes and protecting tissues from oxidative damage (Fabriek et al. 2005). It was found that CD163 serves as a receptor for porcine reproductive and respiratory syndrome virus, which is needed for viral invasion (Calvert et al. 2007). In addition, CD163 has been shown to bind gram-positive and gram-negative bacteria as well as respond to LPS inducing/regulating pro-inflammatory cytokine synthesis (Fabriek et al. 2009).

This research demonstrates a successful strategy for analysing genetic variation associated with the response of swine to Salmonella. Our unique approach for selecting candidate genes as well as initial pre-screening of putative SNPs resulted in twelve SNPs associated with Salmonella shedding or tissue colonization in pigs. This research is the first to link polymorphisms in the NGN3, TAP1, NCF2, VCL, AMT, EMP1, ACP2, CCR1 and CD163 genes with Salmonella infection traits, and we extended our investigations into the association of polymorphisms in CCT7 with porcine resistance to Salmonella. We found that only one of these genes
had SNPs with association across populations (ACP2). However, because the populations had different phenotypes measured, it is unclear whether we should expect such across-population agreement. While these differences in Salmonella-related phenotypes as well as different genetic background of the populations may have provided more opportunities to find association, these characteristics may have reduced our ability to find SNPs with association across the populations. As additional Salmonella-challenged pig populations become available, validation of the association of these markers with Salmonella colonization and shedding in swine will be performed. In addition, further studies may reveal the effect of these SNPs on the function of the genes and their role in porcine disease resistance. The SNPs in our phenotype-associated genes can now be evaluated as markers for selecting animals that shed fewer bacteria and are less likely to cause pen-mate Salmonella contamination on-farm and in the slaughter plant. Thus, this research provides novel information for the pig industry regarding strategies for selecting pigs with reduced shedding and/or disease susceptibility.

Acknowledgements

NPB grant #07-062 has funded the creation of the field pig DNA bank; the collection of the tissues and all Salmonella incidence data was funded by the USDA-NRI under a grant to A. O’Connor, H.S. Hurd and J. McKeen; USDA-CSREES Grant 2005-35212-15929. NPB grant #05-176 funded the NADC 40 pig experiment; NPB grant #08-034 funded genotyping and completion of field population collections. The phenotypic data and DNA samples from the IAH-Compton population were a gift from Pig Improvement Company. We thank Katherine Steffensmeier and Stephen Robinson for isolation of genomic DNA from the field population.

References


Sellwood R. (1979) Escherichia coli diarrhoea in pigs with or without the K88 receptor. The Veterinary Record 105, 228–30.


Supporting information

Additional supporting information may be found in the online version of this article.

Appendix S1 Explanation of MRBP and AULC methods of statistical analysis for the NADC population.

Table S1 Design of Sequenom assays.

Table S2 Gene lists for SNP analysis.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.