MAPPING PHENOTYPIC TRAITS IN SWINE

BY

ERIN KRISTEN WAGNER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Animal Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Associate Professor Jonathan E. Beever, Chair
Associate Professor Martin O. Bohn
Professor Floyd K. McKeith
Professor John Killefer
Professor Sandra L. Rodriguez-Zas

Abstract

The Illinois Meat Quality Pedigree (IMQP), a three-generation Berkshire x Duroc resource population, was created to discover quantitative trait loci (QTL) influencing carcass composition, growth, and meat quality traits in pigs. Each animal in the IMQP population was genotyped for 137 microsatellite markers and 43,486 single-nucleotide polymorphisms (SNP). Two different analyses were performed to detect QTL in this population. For the first analysis, a linkage map consisting of 137 microsatellite markers was used to perform a regression interval analysis with QTLexpress, a publicly available web-based software. In the second analysis, the SOLAR software package was used to perform a variance component analysis with a physical map created from the 43,486 SNP markers. Ninety-five QTL were detected using the QTLexpress software while 787 QTL were detected using the SOLAR software. Of these QTL, 40 of the QTL detected by QTLexpress were located in the same position as 119 of the QTL detected by SOLAR.

One of the QTL identified by the regression interval analysis was highly significant and associated with a relatively large effect on loin eye area. The QTL is located between microsatellite markers SW1129 and SW1647 on chromosome 6 encompassing a span of 73.2 million base pairs (Mb) with a 95% confidence interval spanning 98 Mb. To reduce the size of this interval and identify positional candidate genes, 1,600 SNP markers from the Illumina® PorcineSNP60 Genotyping BeadChip that were located within the 95% confidence interval were analyzed using available analysis software programs. Software included QTLexpress, Qxpak, R/qtl, SOLAR, PLINK and Bayes-C. Results from each analysis were consolidated and a consensus region was identified between 85.67 and 85.68 Mb.

To examine the power of high-density SNP genotyping platforms for the mapping of specific phenotypes, a unique coat color phenotype observed during the development of the IMQP was selected for analysis. Individuals of the F2 generation displayed a diverse set of coat

color phenotypes including several (n=9) solid white individuals. In pigs, a dominant mutation in the KIT gene has been shown to be responsible for white coat color, however neither of the founder breeds are known to have dominant white mutations. Based on matings between white F2 sows with Duroc, Berkshire and white F2 boars, it was concluded that the white coat color observed in this population is inherited as a recessive phenotype. A whole genome association study was performed to identify regions of the genome controlling recessive white coat color. Regions with significant associations were identified on chromosomes 1, 2, and 6. The locus on chromosome 6 contained melanocortin receptor 1 (MC1R), a gene known to control pigmentation in mammals. Each animal in the population was genotyped for the MC1R recessive e/MC1R*4 allele (A240T mutation), and all white F2 animals genotyped as homozygous 240T, or e/e. Tyrosine hydroxylase was identified as a candidate gene for the region on chromosome 2. Genomic sequence for tyrosine hydroxylase was generated and the population was genotyped for an A159T mutation. All solid white individuals except two genotyped as homozygous 159T for this mutation. Two significant regions associated with white coat color were identified on chromosome 1 and several positional candidate genes were identified but remain to be investigated.

Acknowledgements

I would like to express my gratitude to my advisor, Dr. Jonathan Beever, who allowed me to pursue my graduate studies and with whose knowledge, patience, encouragement, and understanding has allowed me to complete this graduate degree. I would like the thank my graduate committee, Dr. Martin Bohn, Dr. John Killefer, Dr. Floyd McKeith, and Dr. Sandra Rodriguez-Zas for their time, knowledge, and assistance throughout my graduate career. I would like to acknowledge my family for the love, patience, and emotional support they provided, especially my parents, Isabelle and Ronald Wagner; my oma, Eva Wagner; and my brother and sister-in-law, Adam and Millie Wagner. For all the help (inside and out of the lab) and friendship they have provided, I would like to thank Miri Cohen-Zinder, Alysta Markey, Brandy Marron, Shannon Maxey, Stacey Meyers, Audra Picard, Natalie Pierce, and Kristen Walker. Thanks to all of my friends especially Liz Kim, Cara McGee, Sarah Schultz, and Carol Bonin for their love, patience, and most importantly their ability to keep me sane. Lastly, I thank Colin Larson for his love and patience.

Table of Contents

List of Figures

List of Tables

Literature Review

Introduction

Pork producers have traditionally relied on both selective breeding based on phenotypic measurements and the improvement of management practices to yield enough pork to meet demands. According to the USDA, pork is the most consumed meat in the world (Figure 1), and if pork producers want to continue to meet the demands of the world's population, currently nearing seven billion [1], different strategies need to be implemented. In pigs, many of the traits producers are interested in are, to some degree, genetically controlled. Identification of the genes that control these traits can be achieved by performing mapping studies with the results leading to the discovery of polymorphisms that could ultimately be used as selection criteria for breeding programs.

Phenotypic traits can generally be characterized as either qualitative or quantitative. Qualitative traits are usually controlled by one or very few genes and environment has little effect on the development of the phenotype. These traits are relatively easy to study since the phenotype seen in the individual is directly associated with a genotype at the gene controlling the trait. In contrast, quantitative traits are genetically complex, i.e. they are controlled by many genes each of which has a relatively small effect on the phenotype and environment usually has an influence on the phenotype. Despite differences between qualitative and quantitative traits, the strategies used to identify the genes that control either type of phenotype are similar and can be divided into two categories, linkage mapping and association mapping.

Mapping Strategies

Linkage Mapping

During meiosis, chromosomes go through recombination, the process of exchanging genetic material between non-sister chromatids of homologous chromosomes. As the distance between two loci on a chromosome decrease, the probability of a recombination occurring between them decreases, thus the probability that alleles at these loci will be inherited together increases. The tendency of alleles at different loci to be inherited together is referred to as linkage and is the basis of linkage mapping strategies. Performing a linkage mapping analysis requires a population with several generations where the relationships between the individuals are known, a phenotype that displays variation in the population, and at least one polymorphic marker.

Though any population where relationships are known can be used, most linkage mapping studies in agricultural species have involved experimental population types such as F2 intercrosses (F2), backcrosses (BC) or recombinant inbred lines (RIL), particularly in plant species. Typically, these populations begin with the mating of two parental breeds or strains that differ from each other genetically and with regard to the phenotype that is being studied, thus producing an F1 generation that, in theory, will have both a genotype and phenotype that is intermediate to that of the parents. For F2 populations, individuals of the F1 generation are then intermated to produce a genetically and phenotypically heterogeneous F2 generation. Backcross populations are produced by mating the F1 individuals back to one of the two parental breeds, with the resulting offspring being heterogeneous, but not as heterogeneous as offspring from the F2 intercross. RIL populations are created by intermating the F1 to produce an F2 generation, subsequent generations are created by selfing individuals, or in the case of animals (such as mice and *Drosophila*) and self-incompatible plants, siblings from each generation are intermated. At

the end of several generations of inbreeding, the resulting offspring from the entire population will represent the genetic and phenotypic heterogeneity of the F2 generation though the offspring from each recombinant inbred line will be homogeneous due to inbreeding.

Also needed for a linkage mapping analysis are polymorphic markers. Some of the earliest linkage mapping analyses utilized morphological markers [2-4] but their usage was restricted by the limited number of observable phenotypes. The development of molecular markers facilitated linkage mapping analyses since molecular markers do not need to exhibit any noticeable change in phenotype and occur in abundance throughout the genome. Examples of molecular markers that have been used to create linkage maps include restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), microsatellites, and single nucleotide polymorphisms (SNP) [5-8]. Though not required, many linkage mapping analyses use genetic markers to create a linkage map and subsequently use the map for the mapping of phenotypes. Linkage maps are created by determining the recombination frequency between markers and establish the order of the markers via the calculated distance between the markers.

Since qualitative phenotypes can generally be placed into distinct categories, linkage mapping with qualitative data is very similar to the procedure used to create a linkage map, the difference being that a recombination frequency is calculated between the phenotype of interest and a marker on the linkage map, rather than between two markers. Some successful examples of linkage mapping of a qualitative trait include identifying genes controlling various human diseases, including Huntington's chorea [9], cystic fibrosis [10-11], and Duchenne muscular dystrophy [12].

The approach used to identify the genes that influence a quantitative trait using linkage mapping is referred to as quantitative trait locus (QTL) mapping. Since phenotypic values for quantitative traits cannot be placed into discrete categories, the strategy used for QTL mapping analyses involves grouping individuals by marker genotype and then calculating the mean phenotypic value for each marker genotype [13]. An analysis is then performed to determine if there is a statistical difference between the means at each marker genotype, with a difference indicating the marker is linked to the phenotype [13]. This approach can be done by analyzing one marker at a time, referred to as twopoint linkage analysis, or by analyzing all the markers in the linkage map simultaneously, referred to as multipoint or interval mapping [14-16]. Twopoint linkage analysis is most effective when markers are spaced very close together, while multipoint linkage analysis is effective even when markers are spaced relatively far apart, as it the case with many of the linkage maps created for QTL studies [17]. The first genome scan for QTL using molecular markers, published in 1988, identified 15 QTL affecting fruit size, pH and soluble solids of tomatoes [18]. Since then, there have been thousands of QTL studies carried out in hundreds of different species across the animal, plant and fungi kingdoms.

Quantitative Trait Locus Studies

To date, over 280 QTL studies have been published in pigs [19]. The first study performed in pigs, published in 1994 [20], involved a genome scan for QTL influencing growth and carcass traits using an F2 population derived from a European wild boar and a domestic white commercial line. This study identified four QTL on *Sus scrofa* 4 (SSC4) controlling ADG from birth to 70kg, abdominal fat percentage, backfat depth and small intestine length and one QTL on SSC13 controlling ADG from birth to 30kg. Since then over 6,300 QTL controlling almost 600 different phenotypic measures have been reported [19]. At least 4,600 of these QTL are influencing the same or similar traits as recorded in the extant population [19]. Of the over 280 QTL studies that have been performed, at least 35 are

whole genome analyses searching for QTL influencing carcass composition, growth and meat quality traits in various breed intercrosses (Table 1).

Quantitative Trait Nucleotide Case Studies

Genome scans have the potential to identify many QTL throughout a genome, but the low number of markers and large distances between them typically results in the identification of QTL intervals that can span tens of centiMorgans (cM); with a rough estimate of 1 cM per Mb [21], this translates into intervals spanning large regions encompassing thousands of genes. Thus, the identification of putative causative genes and mutations is tremendously difficult. Traditionally, fine mapping studies are needed to refine each of these intervals as much as possible.

Though many QTL have been mapped in livestock species, very few have resulted in the identification of a causative mutations or causative quantitative trait nucleotide (QTN). The few QTN examples that exist for agricultural species are found in the pig and the cow and include the porcine protein kinase and AMP-activated, gamma 3 non-catalytic subunit (PRKAG3) [22], bovine diacylglycerol O-acyltransferase 1 (DGAT1) [23], bovine ATP-binding cassette, sub family-G (White), member 2 (ABCG2) [24-25], and porcine insulin-like growth factor 2 (IGF2) [26] genes.

Porcine PRKAG3

It is well known that mutations in the porcine ryanodine receptor 1 (RYR1) gene, also known as the halothane gene, can produce pale, soft, and exudative (PSE) pork in addition to causing sudden death in response to stress or exposure to halothane. Studies conducted in the 1960's found that pork from the Hampshire breed was also pale, soft, and exudative [27-28], but subsequent studies documented that the occurrence of halothane susceptibility was almost zero in the Hampshire breed [29].

Monin and Sellier (1985) [30] showed that though Hampshire pork was pale, soft and exudative, it was generally darker and less exudative than pork from halothane positive pigs. These characteristics, termed the "Hampshire effect", were found to be the result of abnormally low ultimate pH, rather than an abnormal drop in pH immediately *post-mortem* as seen in PSE meat [30]. In addition, it was shown that the Hampshire breed had a significantly higher amount of glycogen and a higher glycolytic potential in fresh cuts of meat than the other breeds examined, suggesting this was the cause of the low ultimate pH. Le Roy et al. (1990) [31] investigated the inheritance of the Napole technological yield (RTN), a meat quality trait that is a ratio of cooked weight to fresh weight. Two Hampshire-crossbred populations were used to perform a segregation analysis that lead the researchers to conclude that the trait was under the control of a single major gene. The locus was designated RN with two alleles, RN- and rn+, affecting RTN in these populations. The authors theorized about the relationship between the RN gene and the "Hampshire effect" but suggested further experiments would need to be conducted for confirmation.

In 1996, two separate studies used linkage mapping to successfully map the RN gene to a region of SSC15 between two microsatellite markers, SW120 and SW936, a 9.2cM interval according to the USDA-MARC map [32-33]. Milan et al., (2000) narrowed the interval further by identifying additional markers including both microsatellites and SNPs. These new markers were used to create a high density linkage map of the region on SSC15 between SW120 and SW936 [22]. Using this high density map, the region was narrowed to between markers SLC11A1 and S1010, a distance of 1.7cM [22]. Since the corresponding human chromosome did not contain any obvious positional candidate genes, a shotgun library of the bacterial artificial chromosome (BAC) clone that was thought to harbour the RN gene was created and sequenced [22]. Sequence similarity searches of the non-redundant nucleotide database using the BLAST algorithm [34] revealed homology to three genes,

KIAA1073, CYP27A1, and Snf4. Two of the genes, KIAA1073 and CYP27A1, were poor candidate genes based on gene function; however, the third match showed sequence similarity to AMP-activated protein kinase (AMPK) γ subunits. AMPK is made up of three subunits, α, β, and γ and is known to be activated during times of energy deprivation. Furthermore, activated AMPK has been shown to increase the glucose uptake and glycogen content of skeletal muscle when in a state of ATP depletion [35] and is theorized to inhibit glycogen synthesis while activating glycogen degradation [22]. In humans, two genes encoding for the γ subunit of AMPK had previously been identified, PRKAG1 and PRKAG2; accordingly, this locus was designated as PRKAG3.

To further bolster the evidence that PRKAG3 was the potential causative gene, Northern blot analysis was performed and PRKAG3 was shown to be highly expressed in muscle tissue [22]. Sequencing of the PRKAG3 gene in both rn+/rn+ and RN-/RN- pigs revealed seven polymorphisms; four of which resulted in an amino acid substitutions, T30N, G52S, I199V, and R200Q [22]. Examining these polymorphisms in six different pig breeds revealed five PRKAG3 alleles, but only the R200Q allele was consistently associated with the RN- phenotype [22].

A QTL study involving a Berkshire x Yorkshire intercross identified QTL controlling glycogen content of skeletal muscle, glycolytic potential, Hunter L, pH, color, tenderness and flavour all in the same region of SSC15 [36]. Because of the proximity and known function, PRKAG3 was considered a positional candidate gene for these QTL. However, since none of the founders of the population carried the RN- allele, the effects of the other PRKAG3 alleles were investigated [37]. The results suggested that the three amino acid substitutions associated with rn+ alleles all have an effect on glycolytic potential, pH and color score with the largest effect associated with the I199V variant [37]. Subsequently, Lindahl et al. (2004) created a Hampshire x Finnish Landrace population to compare the

effects of the three PRKAG3 alleles; RN-, rn+, and rn* [38-39]. In addition to showing that the RN- allele was dominant over the rn+ and rn* alleles, the results revealed that the RN- allele decreased water holding capacity and pH while increasing glycogen content and color scores, as compared to the other two alleles. Moreover, the rn* allele was shown to reduce the glycogen content and increase the ultimate pH of the meat as compared to the rn+ allele, suggesting this allele may have a positive effect on pork quality.

Barnes et al. (2004) created three transgenic mice, one expressing wild type PRKAG3 ($PRKAG3^{wt}$), one expressing the 200Q variant ($PRKAG3^{200Q}$), and one PRKAG3 knock-out mouse (PRKAG3-/-). Phenotypic analyses of the mice showed that the PRKAG3-/- individuals were unable to synthesize glycogen after exercise while glycogen synthesis in the $PRKAG3^{200Q}$ mice was enhanced over the $PRKAG3^{wt}$ mice [40].

Bovine DGAT1

Genome scans for QTL controlling milk composition traits were performed using two populations of Holstein-Friesian cattle. Both studies detected a QTL with large effects on milk fat percentage, milk protein percentage and milk yield near the centromeric end of Bos taurus chromosome 14 (BTA14) [41-42]. In an effort to refine the map position of the QTL, ten additional markers were added to the map of BTA14, seven sires heterozygous for the QTL were genotyped for these new markers and haplotypes were constructed for each sire [43]. A haplotype-sharing analysis was performed and revealed a 5 cM segment common to all seven chromosomes that increased milk fat percentage [43]. Further refinement of the interval using linkage and linkage disequilibrium analysis was performed on additional populations. The results reduced the QTL interval to ~3 cM and also shifted the interval toward the end of the chromosome [44].

Grisart et al. (2002) identified DGAT1 as a candidate gene for the QTL controlling milk composition traits on BTA14 [23]. DGAT1 is an enzyme that uses acyl-CoA and diacylglycerol to catalyze the final step of triacylglycerol synthesis. DGAT1 has been shown to catalyze triglyceride synthesis in the mammary gland and other tissues while also causing inhibition of lactation in DGAT1 knockout mice [45-46]. The coding sequence of exons 1 and 2 and the sequence spanning exons 3-17 of DGAT1 were sequenced and four polymorphisms were identified, including two intronic base substitutions, a base substitution in the 3' untranslated region (UTR), and an AA to GC dinucleotide substitution in exon 8 causing a nonsynonymous amino acid substitution (K232A) [23]. A separate study also considered DGAT1 as a positional candidate gene and sequenced the gene in three different cattle breeds searching for polymorphisms associated with milk fat content [47]. This study identified a total of 19 polymorphisms and confirmed the previous association between the K232A mutation and milk fat content, thus strengthening the evidence that this may be the causative mutation for the QTL on BTA14 [47].

To rule out other mutations in this region of BTA14, Grisart et al. (2004) genotyped an additional 2,045 Holstein-Friesian bulls for 19 SNPs within the QTL interval, including the K232A mutation, and performed an association analysis using milk fat percentage as the phenotype [48]. The results showed the K232A SNP had the highest association with milk fat percentage. Although other SNPs in this interval were also highly associated with variation in milk fat percentage, the addition of K232A genotype to the statistical model showed no other SNPs remained significant [48]. In addition to the association analysis, Grisart et al. (2004) developed a baculovirus expression system for the expression of both DGAT1 alleles to assess potential differences in the rate of triglyceride production between the two alleles [48]. The results showed that the K allele, the allele that increases milk fat percentage *in vivo*, increased the rate of triglyceride production *in vitro*, further adding to the evidence that the K232A mutation is the causative mutation.

In yet another study, Kühn et al. (2004) analyzed markers on the distal end of BTA14 searching for QTL influencing milk fat percentage in 34 different German Holstein families [49]. As was expected based on previous studies, a QTL effect was detected in the families where the K232A mutation was segregating. Unexpectedly however, significant QTL effects were also detected in four families that were not segregating for the K232A mutation. In these families, a variable number of tandem repeats (VNTR) polymorphism in the 5' UTR identified by Winter et al. (2002) was examined for polymorphisms. A total of five VNTR alleles were found within these families, with the 232K allele being in almost complete linkage disequilibrium with the VNTR allele 3 while the 232A allele occurred in haplotypes with each of the VNTR alleles. The results also showed a varying effect of each VNTR allele, with allele 5 having the largest effect on milk fat percentage. Thus, the authors propose that the QTL effect seen in the families that were homozygous 232A is due to the effects of the different VNTR alleles and suggested that there is more than one polymorphism within DGAT1 influencing milk fat percentage on BTA14.

Bovine ABCG2

In one of the first genome scans using Holstein dairy cattle, Georges et al. (1995) reported a QTL influencing milk yield, milk fat percentage, and milk protein percentage between markers C3H3 and TGLA37, a 20cM interval on BTA6 [50]. Subsequently, at least six other studies have also examined BTA6 for QTL controlling similar milk production traits [51-57]. A summary of these QTL is as follows; milk protein percentage in Dutch Holstein-Friesians [51], milk protein percentage and milk fat percentage in Holsteins [52], milk protein percentage and milk yield in Finnish Ayrshires [53], milk fat yield and milk protein yield in German Holsteins [54], milk protein percentage in Holsteins [55], milk protein percentage, milk fat percentage, milk fat yield and milk protein yield in Israeli Holsteins [56], and milk

protein percentage and milk fat percentage in Norwegian dairy cattle [57]. Common to all of these analyses each identified a QTL located close to the microsatellite marker BM143.

According to the comparative map of the cattle and human genomes, the region flanking marker BM143 corresponds to human chromosome 4 (HSA4) [58] and contained 12 candidate genes, including KIAA0914, HERC3, CEB1, FLJ20637, ABCG2, PKD2, SSP1, MEPE, IBSP, DMP1, DSPP, and SPARCL1 [56]. Cohen et al. (2004) investigated the KIAA0914 gene. The gene is now renamed, family with sequence similarity 13, member A1 (FAM13A1) for polymorphisms controlling the QTL controlling milk production traits in BTA6, but concluded the SNPs identified within the gene were not causative [59].

In an effort to shorten the QTL interval, Olsen et al. (2004 and 2005) initially increased the number of markers in the interval by six microsatellites and later by 20 intragenic SNPs, followed by a combined linkage and linkage disequilibrium analysis to ultimately decrease the size of the QTL interval to 420kb (between the ABCG2 and LAP3 genes) [48-49]. A second candidate gene, osteopontin (SPP1), was sequenced to identify polymorphisms that may be causative for the QTL controlling milk production traits on BTA6 [60]. A total of nine mutations were identified and only an small insertion/deletion (indel) mutation, located ~1240bp upstream the transcriptional start site, was in complete LD with the QTL genotypes of the animals genotyped in this study [60].

A third candidate gene, ABCG2, was investigated for polymorphisms controlling the QTL on BTA6 by Cohen-Zinder et al. (2005) and Olsen et al. (2007) [24-25]. Both authors identified a base substitution within exon 14 that resulted in an amino acid substitution, Y581S [24-25] and suggested that this was the causative mutation controlling the milk production trait QTL on BTA6. In addition to identifying the ABCG2 Y581S polymorphism, both authors were able to exclude the indel mutation upstream SPP1 identified by Schnabel

et al. (2005) as the causative mutation. Cohen-Zinder et al. (2005) showed that the sequence surrounding the SPP1 indel showed variation in sires that were not segregating for the QTL [24]. And Olsen et al. (2007) showed a complete elimination of the QTL effect when adding the ABCG2 Y581S polymorphism to the analysis model, while addition of the SPP1 indel resulted in no change in QTL effect [25]. Since no functional studies have been performed for ABCG2, it is still uncertain whether the ABCG2 Y581S polymorphism is the causative mutation influencing milk production traits on BTA6 or if this polymorphism in tight LD with the causative mutation.

Porcine IGF2

A significant QTL controlling lean muscle mass and backfat thickness was identified on the distal end of the p arm of SSC2 in two different QTL studies, one using a Large White x wild boar F2 population and one using a Large White x Piétrain F2 population [61-63]. Based on synteny between this region of SSC2 and HSA11p, both groups proposed IGF2 as a candidate gene controlling this QTL [62-63]. Since IGF2 is known to be paternally imprinted in both human and mouse, the imprinting status of this QTL was tested as part of the QTL analyses. Results demonstrated that the paternal imprinting model was a better fit, thus adding to the evidence that IGF2 is the causative gene [62-63]. Nezer et al. (1999) sequenced the entire coding sequence and some of the 5' and 3' UTR of IGF2 in the Large White and Piétrain animals to look for polymorphisms associated with the QTL [63]. One polymorphism was identified, a G to A base substitution in the 5'UTR, but the authors failed to propose whether they felt this mutation could be causative [63].

Though IGF2 was considered a good candidate gene based on function and imprinting status, there are several other genes in this chromosomal region that are also paternally imprinted. Since previous studies were unable to find a causative mutation within IGF2, Nezer et al. (2003) further refined the QTL interval, effectively decreasing the size of the

interval to a 250kb region containing just two genes, insulin (INS) and IGF2 [64]. To refine the QTL interval, the authors first increased the number of markers in the chromosomal region to a total of 31 markers, genotyped 14 Large White x Piétrain F1 boars for each marker [64]. From these genotypes, two haplotypes were assembled for each boar. The authors then determined the QTL genotype carried by each of the boars (QQ, Qq, or qq), and pooled each of the haplotypes by QTL allele. The results showed a 250kb segment of the chromosome shared by all the Q-bearing chromosomes, suggesting the causative mutation is located within this segment [64]. To identify polymorphisms within this 250kb region, a 28.6kb segment containing IGF2, INS and the 3' end of tyrosine hydroxylase (TH) was sequenced in twelve animals with either known or assumed QTL genotypes [26]. A total of 258 polymorphisms were identified but only one was segregating perfectly with the QTL genotypes [26]. The polymorphism, a SNP in intron 3 of IGF2 (G to A), was located in a highly conserved CpG island [26]. The methylation level of CpG islands regulates the expression of genes, so disruptions or changes in the sequence of CpG islands could change the methylation level and expression level of the gene. To determine if this mutation has any functional effect, a Northern blot analysis was performed and found that there was an increase in the expression levels of IGF2 mRNA in samples that carried the A allele compared to those carrying the G allele, suggesting that the G to A transition is likely to be the causative mutation for the QTL controlling muscle mass and backfat thickness on SSC2 [26].

The above reviewed studies are good examples of the progression from QTL to QTN. These studies demonstrate various techniques that can be used to fine-map a region and narrow the interval containing potential genes while also revealing some of the complications associated with these techniques. While each of these studies has a unique approach to identifying causative mutations, each follows a general plan, including performing a genome scan to identify potential QTL, increasing the number of markers in the vicinity of the QTL to

narrow down the QTL interval, choosing a positional candidate gene, and finally sequencing the gene to identify potential causative mutations. The individual details of how each step was accomplished are unique for each study due to differences in population structure, QTL behavior (e.g. IGF2 is paternally imprinted), and size of interval and marker density within that interval.

Association Mapping/LD Mapping

Association mapping is another method of detecting genetic loci controlling phenotypic variation within a population. This method takes into account linkage disequilibrium (LD), the non-random association of alleles in a population, between the QTL and genetic markers. Similar to linkage mapping, association mapping can be performed on qualitative or quantitative traits. Association mapping of qualitative traits, commonly called a case-control association analysis, can be performed by using an allelic chi-square test. For the analysis of quantitative traits a phenotypic mean is calculated for each genotype at a SNP locus, and an association between the locus and the phenotype exists when there is a significant difference between the phenotypic means across the three different genotypes at a SNP locus. Both qualitative and quantitative association analyses can be performed on a genome-wide scale, referred to as genome wide association studies (GWAS) and with the development of high density SNP genotypic arrays, the opportunity to perform GWAS on various phenotypes has become possible in agricultural species such as swine, cattle, and sheep.

At least seven different GWAS studies have been published on phenotypic traits in pigs since 2010. All of these studies involve genotyping their populations using the Illumina® PorcineSNP60 BeadChip [65] and the association analysis was performed using either the association analysis option in PLINK [66] or the Bayes-C option of the Gensel program [67]. Three of the studies have looked at qualitative traits, including two studies on coat color

[68-69] and one on polydactyly [70]. The other studies published in pigs have all analyzed quantitative traits including boar infertility [71], boar taint levels [72], muscle area, body conformation, and backfat traits [73], and sow reproductive traits [74]. Several of these studies have identified candidate genes controlling the phenotype in question, but only one identified a potential causative mutation. Ren at al. (2010) performed a GWAS on pigs from several different breeds that display either a brown or black coat color phenotype [69]. The GWAS identified a 1.5Mb significant region on SSC1, a region surrounding the known coat color gene tyrosinase related protein 1 gene (TYRP1) [69]. Sequencing of this gene revealed a 6 base pair (bp) deletion that was in perfect linkage disequilibrium with the brown phenotype [69].

In cattle, there have been at least seven GWAS studies analyzing various phenotypic traits since 2009. All of these studies involve genotyping their populations using the Affymetrix® MegAllele™ Genotyping Bovine 10K SNP Panel [75] the Illumina® BovineSNP50 v2 DNA Analysis BeadChip [76], or both. The analysis methods used in these studies include using the association option in PLINK [66] and mixed model regression analyses implemented by several different software programs. The phenotypic traits studied in cattle using GWAS include carcass traits in beef breeds [77], fertility traits in dairy cattle [78], feed efficiency in beef and dairy cattle [79-80], milk production traits in dairy cattle [81-82], and resistance to *Mycobacterium avium* subspecies *paratuberculosis* (a bacterium that causes Johne's disease, a potentially fatal infection that affects the small intestine of cattle) [83-84]. Though, several of these studies have proposed candidate genes, none have investigated any genes for potential causative mutations. As discussed previously, there have been several causative mutations controlling milk production traits in dairy cattle including the DGAT1 K232A mutation. Both of the GWAS analyzing milk production traits identified SNPs surrounding the DGAT1 gene as significant at genome-wide levels,

demonstrating the power of GWAS on identifying regions of the genome containing known causative mutations.

To the best of our knowledge, there is only one GWAS that has been reported in sheep. Becker et al. (2010) genotyped sheep from a Texel population and performed an association analysis using the association option in PLINK and contrasted 23 lambs affected with microphthalmia, a recessive disorder characterized by small or absent eyes, with 23 unaffected animals [85]. The results showed one significant region of the genome, a 3.6Mb region from *Ovis aries* chromosome 22 (OAR22) and a homozygosity analysis of this region showed a 2.4Mb homozygous region shared by all the affected lambs [85]. This region corresponds to a region of HSA10 that contains at least 46 annotated genes including two genes with functions related to eye development, paired-like homeodomain 3 (PITX3) and paired box 2 (PAX2) [85]. Sequencing of both genes revealed no mutations in the PAX2 gene and only one in the PITX3 gene, a G to C base substitution in exon 4 resulting in an predicted amino acid substitution of arginine to proline at codon 113 (R113P) [85]. When testing all the animals in their population, the authors found that all the affected animals were homozygous for the C allele while all the known carrier animals were heterozygous and the other unaffected animals were either heterozygous or homozygous for the G allele, providing further evidence that this might be the causative mutation for this recessive disease [85].

Figures

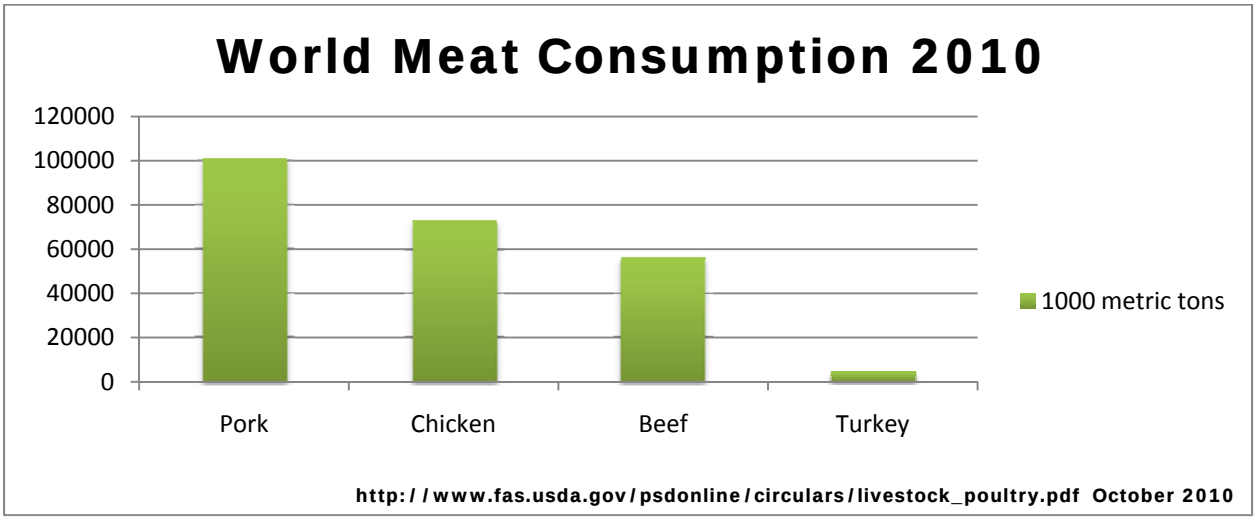

Figure 1. Summary of world meat consumption. Reported by the USDA as of October 2010. Each green bar shows the number of metric tons of meat consumed world wide for each of the four species, pig, chicken, cow, and turkey, with pork being the most consumed meat product in the world.

Tables

Table 1. Summary of previous genome scans to detect QTL in swine

Author(s)	Year	No. of Markers	No. of animals used in analysis	Population structure	Breed Cross[a]	Type of Trait	Sug.[b]	5% Chr[b]	1% Chr[b]	5% Genome[b]	1% Genome[b]
Andersson et al.	1994	105	200	F2	WB x LW	Carcass			2		
						Growth		1	1		
Andersson-Ecklund et al.	1998	236	200	F2	WB x LW	Carcass				5	4
						Meat Quality					
Knott et al.	1998	240	199	F2	WB x LW	Carcass	11			2	
						Growth	5			2	
Rohrer and Keele	1998a,b	156	540	BC	M x LW	Carcass	24			4	13
de Koning et al.	1999	127	619	F2	M x L or M x LW	Carcass	7				1
Paszek et al.	1999	119	200	F2	M x Y	Growth			3	1	
de Koning et al.	2000	132	619	F2	M x L or M x LW	Carcass			5		
Rohrer	2000	157	706	BC	M x LW	Carcass	8			2	12
						Growth	6				4
Wada et al.	2000	318	265	F2	M x MP	Carcass					5
						Growth					3
Bidanel et al.	2001	137	1103	F2	M x LW	Carcass	7			3	21
						Growth	23			4	18
de Koning et al.	2001	132	785	F2	M x L or M x LW	Meat Quality	30			1	2
Malek et al.	2001a,b	125	525	F2	B x Y	Carcass		24		1	5
						Growth		4		1	
Paszek et al.	2001	119	116	F2	M x Y	Meat Quality		1	6		
Milan et al.	2002	137	488	F2	M x LW	Carcass	20			7	30
Nezer et al.	2002	137	525	F2	P x LW	Carcass		2	2		
						Growth			1		
Ovilo et al.	2002	92	369	F2	I x L	Meat Quality		1	4	2	2
Varona et al.	2002	92	321	F2	I x L	Carcass		27	1	7	13
Dekkers et al.	2003	125	525	F2	B x Y	Carcass		23	5	7	11
						Growth		14	1	2	
						Meat Quality		64	13	12	8
Sato et al.	2003	180	865	F2	M x D	Carcass				10	16
						Growth				4	2

Table 1 (cont.)

Geldermann et al.	2003	171	335	F2	WB x M	Carcass	82			
						Growth	25			
						Meat Quality	9			
		152	315	F2	WB x P	Carcass	102			
						Growth	24			
						Meat Quality	23			
		165	316	F2	M x P	Carcass	108			
						Growth	19			
						Meat Quality	19			
Thomsen et al.	2004	158	512	F2	B x Y	Carcass	7		2	
						Growth	5			
						Meat Quality	12		2	1
Nii et al.	2005	225	353	F2	WB x LW	Meat Quality			15	9
Harmegnies et al.	2006	198	1187	F1	(L x LW) x (LW x (LW x P))	Carcass	41		2	
						Growth	3		1	
						Meat Quality	33		8	
Karlskov-Mortensen et al.	2006	131	462	BC & F2	H x L	Carcass			12	3
Kim et al.	2006	183	525	F2	B x Y	Carcass			7	11
						Meat Quality			16	6
Rohrer et al.	2006	182	370	F2	D x L	Carcass			1	19
						Meat Quality			1	51
Wimmers et al.	2006	88	905	F2	D x MP	Carcass	8		1	1
						Meat Quality		1	1	2
van Wijk et al.	2006	73	715		(P x LW) x commercial sow	Carcass	16	3		
						Meat Quality	11	2		
Liu et al.	2007	106	585	F2	D x P	Carcass	27		8	8
						Growth	12		1	2
						Meat Quality	8		1	4
Edwards et al.	2008a,b	124	510	F2	D x P	Carcass	14	8	5	9
						Growth	23	6	4	10
						Meat Quality	36	8	9	5
Markljung et al.	2008	120	136	F2	L x H	Carcass				
						Growth				
						Meat Quality	37	8	8	
Guo et al.	2008	187	739	F2	M x LW	Growth	4		3	
						Carcass	10		3	5

Table 1 (cont.)

Liu et al.	2008	106	585	F2	D x P	Carcass	18	10	7	9
						Growth	20	3	4	2
						Meat Quality	11	3	2	
Duan et al.	2009	183	1030	F2	D x E	Meat Quality	58		6	10
Ma et al.	2009	194	1028	F2	D x E	Carcass			4	21
						Meat Quality			18	50

[a]B=Berkshire, D=Duroc, E=Erhualian, H=Hampshire, I=Iberian, L=Landrace, LW=Large White, M=Meishan, MP=Mini Pig, P=Pietrain, WB=Wild Boar, Y=Yorkshire, BC=Backcross

[b]Number of QTL found at each level of significance [Sug.=Suggestive significance, 5% Chr=5% Chromosome-wise significant, 1% CHR=1% Chromosome-wise significant, 5% Genome=5% Genome-wise significant, 1% Genome=1% Genome-wise significant.

Genome Wide Association Study Mapping Recessive White Coat Color in Pigs

Abstract

During the development of a Duroc x Berkshire F2 resource population, nine F2 animals with solid white coat color were produced. In pigs, mutations in the KIT gene have been shown to be responsible for dominant white coat color pattern. However since neither founder breed has the dominant white mutation, further study of the factors controlling white coat color in this population were investigated. Based on matings between the white F2 sows with Duroc, Berkshire and white F2 boars, it was concluded that the white coat color observed is inherited as a recessive phenotype. A whole genome association study using SNPs was performed to identify regions of the genome controlling recessive white coat color. Regions with significant associations were found on chromosomes 1, 2, and 6. The region on chromosome 6 contained MC1R, a gene known to control pigmentation in mammals. Each animal in this population was genotyped for the MC1R recessive e/MC1R*4 allele (A240T mutation), and all white F2 animals were homozygous 240T, or e/e. Tyrosine hydroxylase was identified as a candidate gene for a second locus, located within the significant region on chromosome 2. The genomic sequence for tyrosine hydroxylase was investigated and the entire population was genotyped for a A159T mutation, with all but two of the solid white animals genotyping homozygous 159T for this mutation. Two significant regions associated with white coat color were identified on chromosome 1 and several candidate genes were identified, however, sequence polymorphisms within these genes were not investigated during this study.

Introduction

As early as 1906, researchers have been investigating the genetics influencing coat color in pigs [86]. The first studies were able to propose both mode of inheritance and number of loci controlling various coat colors by analyzing phenotypic segregation ratios in experimental breeding trials [86-89]. With the development of porcine microsatellite maps, loci could be mapped to linkage groups or chromosomes [90-92]. Subsequent studies investigated the molecular basis of coat color, thus several genes have been well characterized as to their influence on coat color in mammalian species. Polymorphisms have been identified through sequence analysis of these genes in pigs, but only the polymorphisms in the v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT) and melanocortin 1 receptor gene (MC1R) have been associated with specific coat color phenotypes [69, 89, 91, 93-110].

The v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog gene (KIT) is a receptor tyrosine kinase necessary for the survival and migration of neural-crest-derived melanocyte precursors; melanocytes are the pigmented cells of the skin, hair and eye [111]. KIT is known as the Dominant white or "I" (I for inhibition of color) locus and has been mapped to pig chromosome 8 [90]. To date, there are a total of seven KIT alleles described in pigs. These seven alleles result from a combination of gene duplication, an point mutation resulting in a splice variant missing exon 17, and a 4 base pair deletion in exon 18 [95, 101, 105, 107]. The recessive "i" allele is the wild type allele and results in a colored phenotype as seen in wild boars, Durocs and other non-white breeds [101]. This allele is characterized as having a single copy of the gene that lacks the splice mutation. The semi-dominant "I^{P}" allele results in animals with colored patches on a white background [90]. This allele is the result of a complete duplication of the KIT coding sequence with neither copy having the splice mutation. The dominant "I^{Be}" allele causes the belted

phenotype as seen in the Hampshire breed. The causative mutation for this allele has yet to be identified, but individuals with the belted phenotype have been found to carry a single copy of the gene that does not contain the splice mutation. Giuffra et al. 1999, sequenced the entire coding sequence of this gene in belted animals and found no mutations within the coding sequence leading to theory that the phenotype is caused by a regulatory mutation [96]. Three different dominant "I" alleles resulting in solid white color have been identified in pigs. The "I^1" allele is characterized as having two complete copies of the gene in with one of the copies carrying the splice mutation [105]. The "I^2" allele is characterized as having three complete copies of the gene with one of the three copies carrying the splice mutation [107]. The "I^3" allele is characterized as also having three entire copies of the gene with two of the copies carrying the splice mutation [107]. Most recently the allele conferring the roan phenotype has been characterized as having one copy of the gene with a four base pair deletion in intron 18 [95]. Fontanesi et al. 2010, suggest this is the "I^d" allele described in earlier literature [95]. Though never documented, Pielberg et al. 2002, hypothesized that there is potential for an eighth allele occurring at low frequencies in some populations that would be characterized by a single copy of the gene containing the splice mutation. However, this allele would be lethal in the homozygous state [107] as are other loss-of-function mutations in other species [112].

The melanocortin receptor 1 (MC1R), also called the melanocyte stimulating hormone receptor, is a G-protein-coupled transmembrane receptor that regulates the amount of eumelanin (black/brown pigment) and pheomelanin (red/yellow pigment) produced by melanocytes. The effect of MC1R on skin, hair, and fur pigmentation has been investigated across many species including pigs [103, 113-124]. Originally referred to in pigs as the extension locus (E), the MC1R gene has been mapped to the distal end of the p arm of chromosome 6 in pigs and consists of a single exon containing seven conserved transmembrane domains (TM) [88, 92, 103]. To date, eight MC1R alleles associated with

coat color phenotypes in pigs have been identified. The brown skin and hair seen in wild boars is the result of the wild type MC1R allele E^+. The wild type allele produces a normal receptor that is able to bind with either melanocortin stimulating hormone (MSH) or agouti signalling protein (ASIP), resulting in melanocytes producing a mixture of eumelanin and pheomelanin. Two different wild type MC1R alleles have been identified in wild boars, European wild boars carry the E^+/MC1R*1 and Japanese wild boars carry the E^+/MC1R*5 with the two alleles differing by only a single synonymous base pair change (T354C) [97, 103]. To date, all of the MC1R mutations found in pigs lead to the melanocyte being able to only produce a single pigment. Most of the causative mutations identified, occur in one of the transmembrane domains where mutations can alter the protein conformation and disrupt ligand binding to the receptor; TM1, TM2, TM3, TM6, TM7 [125-126]. Many of the MC1R mutations characterized in pigs result in melanocytes that can only produce eumelanin. This phenotype is the result of mutations within the transmembrane domains of the gene leading to a constitutively active melanocyte due to the inability to bind ASIP. There are two different MC1R alleles that result in animals with a solid black coat phenotype. The solid black coat color in European black breeds, such as the British Large Black, is the result of the E^{D1}/MC1R*2 allele while the same phenotype in Asian breeds, such as the Chinese Meishan, is the result of the E^{D1}/MC1R*7 allele. The MC1R*2 allele is characterized by two amino acid substitutions V92M and L99P, both located within TM2 and MC1R*7 allele has the V92M and L99P substitutions along with a third amino acid substitution V119I located in TM3 [99, 103]. A third allele resulting in black coat color has been identified in Hampshires, a black pig with a white belt across the shoulders (the white belt is the result of the I^{Be} KIT allele). This allele, E^{D2}/MC1R*3, is characterized by the amino acid substitution D121N and is located within TM3 Kijas et al. 1998, originally reported that the Piétrain breed shared the MC1R*3 with the Hampshire breed, but further investigation showed that the E^P/MC1R*6 allele carried by Piétrain and other breeds with black spots contains a second mutation, an insertion of 2 base pairs within codon 23

(nt167insCC) [102]. The nt167insCC mutation results in a downstream stop codon at amino acid position 56 and is expected to cause melanocytes to produce only pheomelanin. The existence of the black spots on these animals is explained by spontaneous somatic reversions of the nt167insCC mutation resulting in a constitutively activated melanocyte [102] that produces eumelanin. A second E^P allele has been identified in a Piétrain sow, this allele designated $E^{P'}$ shares the D121N and nt167insCC mutations with the E^P allele, but also has another amino acid substitution, R163W [127]. Drögemüller et al. 2006, did not describe specific phenotypic differences in this sow and because this mutation does not occur within one of the transmembrane domains it may not have functional consequences [127]. The e/MC1R*4 allele causes the melanocyte to produce only pheomelanin [103]. This allele is characterized by the amino acid substitution A240T located within TM6, a transmembrane domain where mutations are known to cause conformational changes and disrupt the binding of MSH to the receptor [128].

The Illinois Meat Quality Pedigree (IMQP) is an F2 resource population that was originally developed to identify quantitative trait loci influencing meat quality traits [129]. The IMQP was produced by crossing Berkshire boars with Duroc sows followed by intermating of the resulting F1s. During the production of the F2 generation of this population, nine solid white pigs were born. As indicated previously, the occurrence of white coat color in pigs is more commonly associated with dominant mutations in KIT. However, the parental generation Duroc and Berkshire breeds are assumed to have the i/i (wild-type) genotype at the KIT locus. Additionally, the solid white pigs were not observed until the F2 generation and the ratio of white to non-white animals in the F2 generation was approximately 1 in 64. These data suggest that the solid white coat color seen in this population is inherited as a recessive phenotype and is potentially controlled by at least three independent loci.

To date, only two other studies have reported the recessive inheritance of white coat color in pigs, but neither study has identified the genes or causative mutations responsible for the lack of pigmentation [87, 130]. The first study, conducted in the 1920s, found that solid white piglets had been produced in the F2 generation of a Berkshire x Duroc-Jersey cross [87] similar to the IMQP founder breeds. The ratio of white to non-white F2 animals in this study was also very close to 1 in 64 leading the authors to suggest at least three segregating loci were controlling coat color in this population [87]. While the author never described a specific mode of inheritance, there were no white animals in either the founder or F1 generations suggesting the mode of inheritance was recessive. A second study involved Rongchang pigs, a Chinese breed of pigs characterized by their solid white coat color, although some animals have small black spots around the eyes and ears. Sequence analysis showed that the white coat color in this breed is not the result of any of the known dominant white KIT alleles, but the authors did find two amino acid substitutions, V84M and V893A, in the Rongchang sequence compared to the sequence of the dominant white KIT alleles [130]. Also, experimental matings between Rongchang pigs and colored pigs showed that the white coat color in this breed was recessive to non-white coat color [130]. This study aims to determine the genes controlling the recessive white coat color in the IMQP population by performing a genome wide association study (GWAS) and examining polymorphisms in potential candidate genes.

Materials and Methods

IMQP Reference Population

The Illinois Meat Quality Pedigree (IMQP) is an F2 Berkshire x Duroc intercross population consisting of 664 individuals. Semen from three purebred Berkshire boars, chosen based on progeny test performance values, was used to inseminate 17 purebred Duroc sows chosen from the breeding program at the University of Illinois Moorman Swine Research Unit. Each

boar was bred to an average of six different sows, but each sow was only mated to one of the three boars. All founder animals were genotyped for the halothane and Rendement Napole (R200Q) mutations and were found to be normal. An F1 population was produced, from which five boars and 44 sows were chosen to be mated. No mating of relatives occurred within the F1 intercross. Each of the F1 boars were mated to an average of 11 different F1 sows, while each F1 sow was mated to either one or two different F1 boars. The intercross between F1 individuals created an F2 generation consisting of 595 pigs.

All of the Duroc sows used to found the population displayed the characteristic solid red coloring (Figure 2A) while all of the Berkshire boars displayed the characteristic solid black coat with six white points (i.e. nose, tail and four feet) (Figure 2B). The F1 animals from this population were all solid red with small black spots (Figure 2C). The F2 generation consisted of animals with a variety of coat colors including solid red (ranging from pale red to orange to dark red), red with a white belly, red with six white points, red with small black spots, red with black spots and white underline, red with black spots and six white points, solid black, black with large white patches, black with six white points, and solid white.

It is these solid white animals (Figures 3A & B) that are of interest in this study. A series of matings involving these solid white pigs were conducted to determine the mode of inheritance for this phenotype, since it was assumed that these animals were not white due to any of the dominant white alleles at the KIT locus. Matings were performed between solid white sows and either solid white boars, purebred Duroc boars, purebred Berkshire boars, F1 boars from the IMQP population, or a white F2 x Berkshire crossbred boar.

SNP Genotyping and Genome Wide Association Study

Fifteen microliters of genomic DNA at a concentration of 100ng/µl was sent to GeneSeek® for SNP genotyping using the Illumina® PorcineSNP60 BeadChip [65]. The final list of SNPs

used for this analysis was created using PLINK v1.05 [66] by removing Mendelian errors, removing SNP markers with a minor allele frequency less than 0.01, and removing SNPs with a missing genotype frequency of 0.9 or greater.

To identify regions of the genome associated with solid white coat color in this population, a genome wide association study was done using PLINK v1.05 [66]. This was performed using a series of case-control, allelic Chi-square tests using the ASSOC command in PLINK v1.05. A Bonferroni adjustment was applied to the p-values obtained by using the ADJUST command in PLINK. Three GWASs involving solid white animals were performed including 1) solid white (n=19) versus solid red with MC1R genotype e/e (n=68), 2) solid white (n=19) versus purebred Berkshires and F2 individuals displaying the Berkshire phenotype (n=11), and 3) solid white, purebred Berkshires and F2 individuals displaying the Berkshire phenotype (n=30) versus solid red with MC1R genotype e/e (n=68) (Table 2). To narrow the regions further, a homozygosity analysis was performed using the HOMOZYG command in PLINK. To further refine the search for causative loci, sequences flanking the SNP markers were used to perform BLAST similarity searches [34] against the human genome and a list of candidate genes within these regions was compiled. Two genes from this list, MC1R and tyrosine hydroxylase (TH), were chosen for further investigation.

MC1R Genotyping

To confirm the MC1R genotypes for the founder and the F1 animals and to determine the MC1R genotypes of the F2 pigs, each animal was genotyped for the A240T MC1R mutation. Though not tested for directly, DNA fragments that did not have the A240T mutation were assumed have the D121N and nt167insCC mutations consistent with the Berkshire breed [102]. The primers used to amplify the fragment surrounding this mutation, (5'-TCTCCAGCACCCTCTTCATC-3') and (5'-CCTGCTTCCCTAGCAGTCAC-3'), were designed based on publically available Duroc MC1R sequence (GenBank accession number AY365250.1)

using Primer-3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). Each individual in the IMQP population was then genotyped for this mutation using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) genotyping. PCR was performed in a 20 μl reaction composed of 50 ng of DNA; 1X PCR buffer (containing 1.5 mM $MgCl_2$); 200 μM of each dNTP (Fermentas); 0.5 μM of each primer and 0.5 U of HotStar® Taq DNA polymerase (Qiagen®). PCR amplification conditions were as follows: initial denaturation at 95°C for 5 minutes, followed by 35 cycles of 94°C for 1 minute, 58°C for 1 minute and 72°C for 1.5 minutes, and a final extension step of 72°C for 5 minutes. After amplification, the PCR product was digested in a 30 μl reaction with 5 units of the restriction enzyme *BstU*I and 1.5X NEBuffer 4 at 60°C for one hour. Digested fragments were visualized by UV light following electrophoresis in a 2% agrose gel stained with ethidium bromide.

Porcine Tyrosine Hydroxylase Gene Sequence Analysis and Genotyping

Genomic sequence for the pig TH gene was available for exons 10-14 (Genbank accession number AY044828.1) and mRNA sequence was available for exons 6-14 (GenBank accession number CF367294). Primers were designed from the available sequence using Primer-3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and PCR was performed on several bacterial artificial chromosome (BAC) clones to identify a clone that contained the gene sequences. BAC clone CH242-392F17 was chosen as a good candidate and was sequenced by SeqWright© (Houston, TX).

The completed sequence contig generated from the BAC clone was used to design primers for sequencing of all exons, introns, and several select regions of the promoter (Table 3). The promoter of human TH has been found to extend 11kb from the 5' end of exon 1 and five regions conserved across human, rat, and mouse have been identified and are thought to be have a function associated with transcriptional activity and tissue specific expression

[131]. The Percent Identity Plot (PIP) software [132] identified four segments of the porcine TH promoter as having >50% identity with the human promoter surrounding these five conserved regions (Figure 4). Sections of the promoter sequenced to investigate these four segments include approximately 2.3Kb of the promoter directly 5' of the first exon, containing conserved region V (CR-V); the segment of promoter between 3.2Kb and 4.5Kb from the first exon, that contains the conserved regions II (CR-II), III (CR-III), and IV (CR-IV); and the segment of promoter between 4.8Kb and 7.8Kb from the first exon, that contains conserved region I (CR-I). Thirteen individuals from the IMQP population were selected for sequencing based on their coat color phenotype. The thirteen individuals included two purebred Berkshires, two purebred Durocs, three F1 individuals, three solid white F2 individuals, and three solid white F3 individuals.

PCR was performed for each primer set in a 20 μl reaction composed of 50 ng of DNA; 1X PCR buffer (containing 1.5mM $MgCl_2$) (Qiagen®); 200 μM of each dNTP (Fermentas); 0.5 μM of each primer and 0.5 U of HotStar® Taq DNA polymerase (Qiagen®). PCR amplification conditions were an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of 94°C for 1 minute, 57°C for 1 minute and 72°C for 1.5 minutes, and a final extension step of 72°C for 5 minutes. Residual primers and dNTPs were digested by incubating 10 μl of PCR product with 4 ul of a 1:12 dilution of ExoSAP-IT® (Affymetrix®) at 37°C for 45 minutes followed by denaturation at 80°C for 15 minutes.

Sequencing reactions were then performed in an 8 μl reaction composed of 2 μl of ExoSAP-IT® (Affymetrix®) PCR fragment, 3.62 μl of sequencing buffer, 0.25 μl of BigDye® Terminator v3.1 Ready Reaction Mix, 0.08 μl of 7-deaza-dGTP BigDye® Terminator v3.1 Ready Reaction Mix, and 1.3125 μM of the appropriate primer. Sequencing reaction cycle parameters were an initial denaturation at 96°C for 1.5 minutes, followed by 54 cycles of 96°C for fifteen seconds, 53°C for fifteen seconds, 60°C for three minutes, and a final

extension step of 60°C for ten minutes. Unincorporated nucleotides and primers were removed from sequencing reactions using illustra Sephadex™ G-50 Fine beads (GE Healthcare®) and then analyzed on an ABI Prism 3730xL Analyzer (Applied Biosystems™). All sequences were assembled and polymorphisms identified using CodonCode Aligner v1.5.2 (Dedham, MA).

The IMQP population was genotyped for a sequence polymorphism located within exon 3 of the TH gene using PCR-RFLP genotyping. PCR was performed using the TH 5 primer set (Table 3) in a 20 μl reaction composed of 50 ng of DNA; 1X PCR buffer (containing 1.5 mM $MgCl_2$) (Qiagen®); 200 μM of each dNTP (Fermentas); 0.5 μM of each primer and 0.5 U of HotStar® Taq DNA polymerase (Qiagen®). PCR amplification conditions were an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of 94°C for 1 minute, 57°C for 1 minute and 72°C for 1.5 minutes, and a final extension step of 72°C for 5 minutes. After amplification, PCR product was digested in a 30 μl reaction with 3 units of the restriction enzyme *BsaHI* and 1.5X NEBuffer 4 at 37°C for one hour.

Results

IMQP Reference Population

The segregation ratio of white to non-white animals in the F2 generation of the IMQP population was 9 white : 586 non-white, this is not statistically different (chi square=.0096, df=1, p=0.9218) from a 1 : 64 segregation ratio, the ratio of F2 individuals having a completely recessive genotype from a trihybrid cross.

A series of matings based on coat color phenotype was performed to determine the mode of inheritance of solid white coat color in the IMQP population (Table 4). One solid white F2 boar was mated to two different solid white F2 sows and the resulting litters consisted

exclusively of solid white piglets (Figures 5A-5D). Siblings from one of these litters were mated together producing a litter also composed solely of solid white piglets. The same two solid white F2 sows mentioned above were also mated to a purebred Berkshire boar resulting in litters consisting of piglets that were white with large black patches. A boar from one of these litters was then mated to 3 solid white F3 females, half of the offspring were solid white and the other half were white with large black patches. Two solid white F2 sows and 5 solid white F3 sows were bred to purebred Duroc boars and all of the resulting piglets displayed a solid red back with white legs, belly and chin (red "berkshire"). Two solid white F2 sows and 2 solid white F3 sows were mated to an F1 boar with the litters resulting an array of coat colors including solid white; white with large black spots; red "berkshires"; and red with black spots and white legs and bellies.

Genome Wide Association Study

The analysis comparing solid white animals with solid red animals with the MC1R genotype e/e identified a small region on the proximal end of the q arm of SSC1, a large region on the q arm of SSC1, and the distal end of the p arm of SSC2 as significantly (Bonferroni adjusted $p<0.01$) associated with solid white coat color (Figure 6). A group of 46 SNPs spanning the region between 18.99 and 41.70Mb on the p arm of SSC1 were identified as significantly associated with solid white coat color with the most significant SNP having an adjusted p-value of 6.98E-7 and being located at 35.94Mb (Appendix A: Coat color association analysis results). The q arm of SSC1 had a group of 261 SNPs, with an adjusted p-value below 0.01, spanning the region between 127.42 and 296.20Mb with the most significant SNP having an adjusted p-value of 8.25E-17 and being located at 240.36Mb (Appendix A: Coat color association analysis results). A total of 66 SNPs had adjusted p-values below $p=0.01$ on the distal end of the p arm of SSC2, these being located between 0.42 and 29.81Mb with the most significant SNP having an adjusted p-value of 2.60E-9 and being located at 2.64Mb (Appendix A: Coat color association analysis results). Eight other

SNPs were identified as being significant by this analysis, one was located at 115.30Mb on chromosome 3 (SSC3), five were located between 52.18 and 57.86MB on chromosome 12 (SSC12), one was located at 34.50Mb on chromosome 16 (SSC16) and the last SNP was located at 50.34Mb on chromosome 17 (SSC17) (Appendix A: Coat color association analysis results).

The analysis comparing solid white animals with purebred Berkshire and F2 animals displaying the Berkshire phenotype only identified the distal end of the p arm of SSC6 as being significantly associated with solid white coat color (Figure 7). Sixteen SNPs were identified as significant, these being located between 0.49 and 9.60Mb on SSC6 with the most significant SNP having an adjusted p-value of 3.23E-9 and being located at 1.77Mb (Appendix A: Coat color association analysis results).

The analysis comparing solid white, purebred Berkshire and F2 animals displaying the Berkshire phenotype with solid red animals with the MC1R genotype e/e identified a small region on the proximal end of the q arm of SSC1, a larger region on the distal end of the q arm of SSC1, the distal end of the p arm of SSC2 and the distal end of the p arm of SSC6 as significantly associated with solid white coat color (Figure 8). The p arm of SSC1 contains 30 SNPs with an adjusted p-value below 0.01 between 16.23 and 41.70Mb with the most significant SNP having an adjusted p-value of 1.05E-8 and being located at 27.67Mb (Appendix A: Coat color association analysis results). The q arm of SSC1 contains a group of 431 SNPs spanning the region between 122.09 and 296.20Mb were found to be significantly associated with white coat color on the distal end of the q arm of SSC1 (Appendix A: Coat color association analysis results). The most significant SNP in this region had an adjusted p-value of 2.07E-20 and was located at 240.36Mb. Sixty-four SNPs were identified as significant, between 0.42 and 18.56Mb on SSC2 with the most significant SNPs having an adjusted p-value of 1.63E-8 and being located at 0.42, 0.42, and 0.53Mb

(Appendix A: Coat color association analysis results). Forty-three SNPs were identified as significant between 0.49 and 7.51Mb on SSC6 with the most significant SNP having an adjusted p-value of 2.28E-8 and was located at 5.03Mb (Appendix A: Coat color association analysis results). Five other SNPs were found to be significant by this analysis with one SNP on each of the following chromosomes; SSC5 at 88.99Mb, SSC9 at 150.79Mb, SSC12 at 57.63Mb, SSC16 at 34.50Mb, and SSC17 at 50.34Mb (Appendix A: Coat color association analysis results).

The visual inspection of SNP genotypes on SSC6 (data not shown), revealed no regions of homozygosity consistent with the phenotypes of the animals analyzed. This was most likely due to the fact that, as mentioned above, the Sscrofa10 assembly is missing the most distal end of the chromosome, where MC1R is located. The visual inspection of SSC2 did identify several small regions of homozygosity within the significant region on SSC2 (Figure 9). These regions span from 0.4235 to 0.4237Mb, from 0.503 to 0.872Mb and from 1.06 to 1.07Mb.

BLAST [34] results show the significant region is located between 77.2 and 87.4Mb on HSA16, though no obvious candidate genes were identified within this region, just outside of this region is MC1R. The reason the significant region does not encompass MC1R is that the Sscrofa10 assembly map does not cover the most distal end of the p arm of SSC6, where MC1R is located.

Based on the results from the GWAS, the regions of the genome that are most likely to contain causative mutations associated with white coat color in this population include the distal end of the p arm of SSC1 from 16.23 to 41.91Mb, the q arm of SSC1 from 127.42 to 296.20Mb, the distal end of the p arm of SSC2 from 0.42 to 29.81Mb and the distal end of the p arm of SSC6 from 0.49 to 9.60Mb.

MC1R Genotyping

Digestion of PCR product with *BstU*I produces fragments of 114, 198, and 491 bp for the E^P/MC1R*6 allele and fragments of 114 and 689 bp for the e/MC1R*4 allele. As expected the three Berkshire boars were homozygous for the E^P/MC1R*6 allele, the 17 Duroc sows were homozygous for the e/MC1R*4 allele and the all 49 of the F1 animals were heterozygous. The F2 generation displayed a segregation ratio of 150 E^P/E^P : 263 E^P/e : 143 e/e, which is not statistically different (chi square=1.794, df=2, p=0.4078) from the 1:2:1 segregation ratio that was expected. While all the solid white animals from the F2 generation were homozygous recessive at the MC1R locus, not all F2 animals that were homozygous recessive were solid white in color.

Tyrosine Hydroxylase Sequence Analysis and Genotyping

Using the Spidey alignment tool available on the NCBI website (http://www.ncbi.nlm.nih.gov/spidey/) was used to align the porcine tyrosine hydroxylase genomic sequence with the human tyrosine hydroxylase, transcript variant 2 mRNA sequence (Genbank accession number NM_000360.3). The results predicted that the porcine tyrosine hydroxylase gene has a coding sequence of 1486 bp and contains 13 exons. All exons, introns, and several select regions of the promoter associated with tyrosine hydroxylase were sequenced in both the Duroc and Berkshire breeds and submission of the sequences to Genbank is in progress. A total of 20 sequence polymorphisms between the Duroc and Berkshire breeds were identified, including 19 base substitutions and an indel (Table 5). Ten of the base substitutions were located within the promoter sequence, but none of them occurred within any of the conserved regions so any regulatory effects they might have on tyrosine hydroxylase is unknown at this time. The other nine base substitutions and the indel occurred within the gene, with all but one base

substitution being located within introns. The one base substitution that occurred within an exon, was located within exon 3 and is predicted to result in an amino acid substitution, alanine to threonine at amino acid 159 (A159T). PCR-RFLP genotyping was performed to determine if there is an association with the polymorphism found in exon 3 with solid white coat color in the IMQP population. Digestion of the fragment PCR amplified by the TH 5 primer set with *BsaH*I produces fragments of 66 and 981 bp for the 159T allele and fragments of 66, 398, and 583 bp for the 159A allele. As expected, the Berkshire boars were homozygous for the 159A allele, the Duroc sows were homozygous for the 159T allele, and the F1 pigs were heterozygous. Of all the white animals, from both the F2 and F3 generations, 17 were homozygous for the 159A allele while the other 2 were heterozygous.

Discussion

Solid white coat color in pigs has routinely been found to be the result of a dominant mutation in the KIT gene [95, 98, 100-101, 105, 107-108]. Though two other studies have identified pigs displaying recessive white coat color, neither identified the causative mutations behind the phenotype [87, 130]. The objective of this study is to identify the genetics associated with a recessive white coat color phenotype occurring in the F2 generation of Duroc x Berkshire cross. Segregation analysis of coat color in the F2 generation and experimental matings involving these white animals suggest that the solid white coat color in these animals was recessive and controlled by at least three different genes and a GWAS identified four different genomic regions with significant associations to white coat color.

The association analyses performed by PLINK identified four different significant regions associated with coat color in the IMQP population including a 25.7Mb region at the distal

end of the p arm of SSC1, a 169Mb region on the q arm of SSC1, a 29.4Mb region at the distal end of the p arm of SSC2 and a 9Mb region at the distal end of the p arm of SSC6.

Investigation of the region on the distal end of the p arm of SSC1 did not yield any potential candidate genes, so this region was not examined further during this study. However, the region on the q arm of SSC1 does include three genes with known functions in melanogenesis; the guanine nucleotide binding protein (G protein), q polypeptide (GNAQ), tyrosinase-related protein 1 (TYRP1), and melan-A (MLANA). The association analysis identified the most significant SNP on the q arm of SSC1 to be located at 240.3Mb and GNAQ is located at 240.7Mb. GNAQ is an alpha subunit of a heterotrimeric G protein that in melanocytes is associated with the seven transmembrane domain receptor, Frizzled. In mice this gene has been found to regulate the amount of both eumelanin and pheomelanin produced in the melanocyte. Mutations in this gene have been found to cause an over production of melanin resulting in mice with dark skin and hair color, specifically on the ears, tail and feet [133-134]. Though the two known mutations in this gene result in a darkening of the skin and hair, the gene regulates the amount of melanin produced by the melanocyte therefore it may be possible for other mutations to result in a reduction in the amount of melanin produced resulting in a lighter skin and hair. The darkening of the skin and hair on the feet and tail may also be important to consider due to the similarity of the locations to the white coat color seen in the Berkshire breed. Therefore GNAQ could be considered a candidate for controlling either the color of the skin and hair in these pigs or the location of the color on their bodies or both. While the most significant SNP identified on the q arm of SSC1 was located at 240.3 Mb, the groups of SNPs within this significant region spanned a very large portion of the chromosome. When visually inspected, the SNPs in this region can be broken into two distinct groups, with the second group containing the most significant SNP. If considered a separate group of SNPs, the most significant SNP in the first group is located at 216.0 Mb and both TYRP1 and MLANA fall within this first group

of SNPs. TYRP1 is located at 219.1 Mb and MLANA is located at 225.9 Mb. TYRP1 encodes an enzyme that oxidizes 5, 6-dihydroxyindole-2-carboxylic acid (DHICA) to 5,6-Indolequinone-2-carboxylic acid (5,6-IQCA) and then incorporates this molecule into the eumelanin polymer, resulting brown eumelanin rather than black eumelanin [135-136]. Several different sequence polymorphisms have been identified within the TYRP1 gene, but only a six base pair deletion in exon 8 has shown an association with a specific coat color phenotype [69, 110]. This six base pair deletion in the TYRP1 gene was associated with brown coat color phenotypes in Chinese breeds, and though animals from both the Berkshire and Duroc breeds have been genotyped for this deletion, neither breed carried the deletion [69]. Mutations in the human TYRP1 gene are known to cause oculocutaneous albinism type 3, a disorder characterized as a dilution in the coloration of the skin, hair and iris [137-140]. In many mammals, including mice, cats, dogs, cattle, and sheep, mutations within the TYRP1 gene are associated with a brown coat color with the exception of a dominant mutation found in mice that results in white hairs that pigment only at the very tips [141-147]. Mice that are homozygous for this mutation and are also homozygous for the nonagouti "a" allele at the agouti signaling protein locus (ASIP), have much lighter ventral hairs compared to their dorsal hairs [148]. Though most of the mutations identified within the TYRP1 gene in animals result in a brown coat color phenotype, the dominant mutation in mice does result in the production of partially white hairs and the mutations found in humans result in an albino phenotype, so there is a possibility that uncharacterized mutations in the porcine TYRP1 gene may contribute to the white coat color phenotype seen in the IMQP population. MLANA is a protein that complexes with and regulates the activity of the melanocyte protein 17 precursor protein (Pmel17), whose function is essential for melansome maturation [149]. To date, no studies have identified any sequence polymorphisms in any species, nor have any studies associated this gene with any specific pigmentation phenotype, but because of its regulation of Pmel17 and melanosome

maturation, mutations in this gene may result in a white coat color phenotype similar to the phenotypes seen with dominant white KIT mutations.

The significant region identified on the distal end of the p arm of SSC2 spans from 0.42 to 29.81Mb. Though no good candidate genes were identified within this region, the TH gene was located just upstream from the most 5' end of this region. One reason the association analysis may not have identified the region encompassing TH as significant is that Sscrofa10 is missing approximately 2.40 Mb of sequence within this region.

Tyrosine hydroxylase is the rate limiting enzyme responsible for catalyzing the conversion of tyrosine to dopamine, this being the first step in the biosynthesis of melanin. Though no mutations, associated with a change in pigmentation, have been identified within TH in any mammalian species, mutations at the *ple* locus, the Drosophila melanogaster TH homolog, result in unpigmented cuticle and head skeletons [150-151]. Also, human TH isoenzyme I has been found to be expressed in human melanocytes [152]. The known enzymatic activity, Drosophila mutations and the human expression data of TH make this gene a good candidate for producing white coat color even though it is not within the significant region identified on SSC2.

We investigated TH for sequence polymorphisms within the Duroc and Berkshire breeds to identify any mutations that may be associated with white coat color as seen in some of the F2 animals from a cross between these two breeds. A total of 20 sequence polymorphisms were identified between the Duroc and Berkshire breeds, suggesting this gene may have some influence on differences seen between these two breeds. Only one of these polymorphisms is located within the coding sequence; this base substitution is located within exon 3 and is predicted to result in change from alanine in the Berkshires to threonine in the Durocs at amino acid 159. When the rest of the IMQP population was

genotyped for this mutation, it was expected that all the white animals in both the F2 and F3 generations would be homozygous for the 159T mutation, but this was not the case. Two of the white animals were heterozygous for this mutation, and when these two animals were sequenced, it was found that they were heterozygous for all 20 sequence polymorphisms within TH, so it is unlikely that the heterozygous genotype for the mutation in exon 3 is an error. The number of white animals being homozygous for the 159T TH allele is statistically significant (chi square=19.66, df=2, $p<0.0001$), though because not all the white animals are homozygous there is a chance that mutations within TH do not control coat color. A more likely hypothesis is that since coat color is controlled by more than a single locus there may be an epistatic relationship between TH and one of the other loci controlling coat color which could explain why there are white animals that are heterozygous for the A159T mutation.

The significant region identified on the distal end of SSC6, spans from 0.49 to 9.60Mb. Similar to the region on SSC2, there are no good candidate genes within this region, but there is a good candidate gene just outside the region. MC1R is a G-protein-coupled transmembrane receptor that regulates the amount of eumelanin (black/brown pigment) and pheomelanin (red/yellow pigment) produced by melanocytes. Mutations within MC1R have been well characterized in both the Duroc and Berkshire breeds. Durocs carry the e/MC1R*4 allele, this allele corresponds to the MC1R A240T mutation that causes the melanocyte to produce only pheomelanin resulting in red coat color as seen in the Duroc and Tamworth breeds [103]. Berkshires carry the E^P/MC1R*6 allele, the allele that is characterized as having both the D121N mutation and a 2bp insertion (nt167insCC) that that is somatically unstable and undergoes spontaneous somatic reversions resulting in a spotted coat as seen in the Berkshire and Piétrain breeds [102, 127]. The entire IMQP population was genotyped for the A240T mutation to identify if there was any association between either of the MC1R alleles and solid white coat color. The results concluded that all

the solid white animals were homozygous for the e/MC1R*4 allele, but the white animals were not the only animals that were homozygous for the e/MC1R*4 allele, adding to the evidence that more than one loci are controlling this phenotype.

Conclusion

In conclusion, this study identified one of the recessive alleles controlling coat color, the MC1R e/MC1R*4; identified another potential allele, the TH 159T allele; and proposed several other candidate genes that may play a role in producing recessive white coat color in the IMQP population. Future projects will focus on further investigation of the region on SSC2 for other potential mutations and on examining the candidate genes on SSC1 for sequence polymorphisms associated with the white coat color phenotype in this population.

Figures

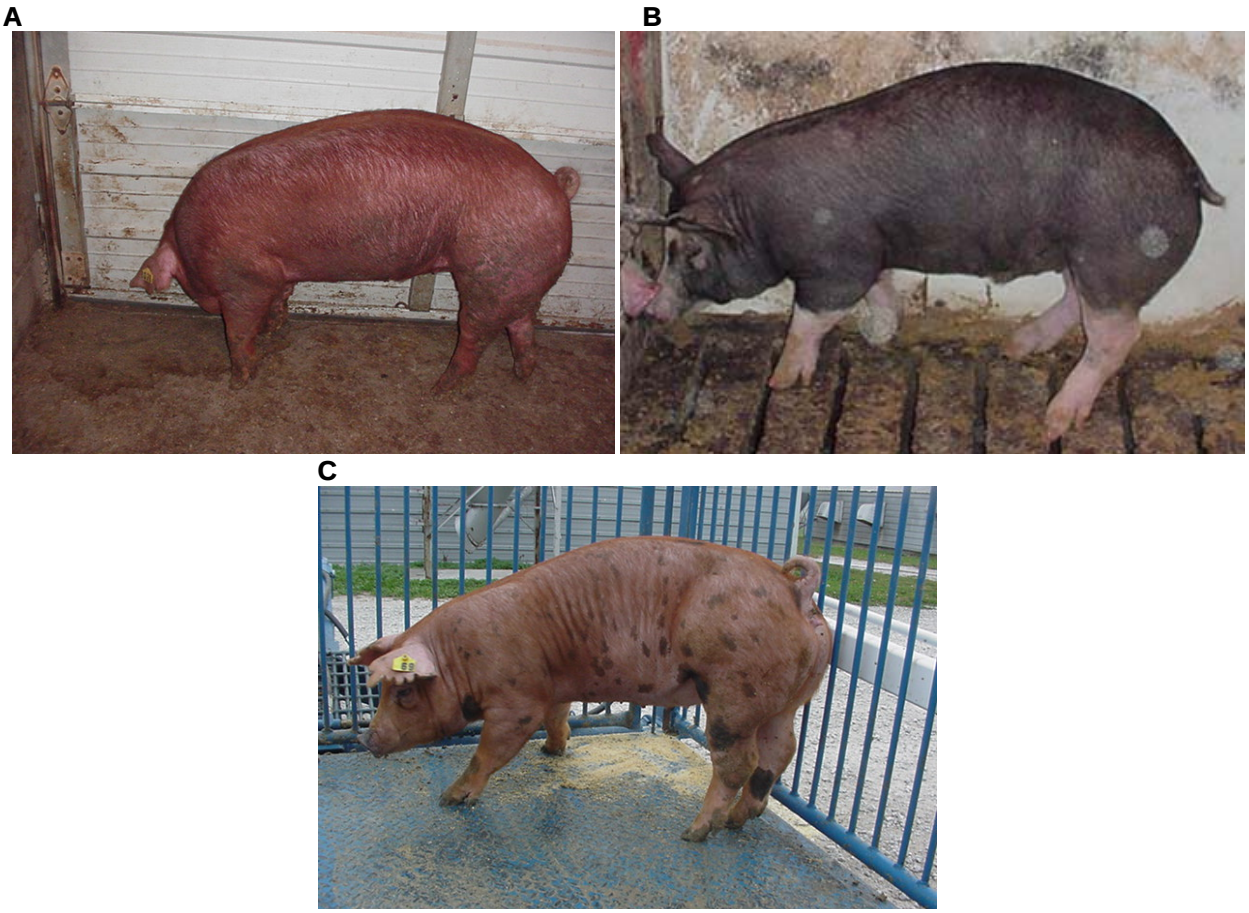

Figure 2. Photographs Duroc, Berkshire, and F1 animals. A) solid red color of the Duroc sows, B) the coat color and patterning of the Berkshire boars (solid black coat with the exception of a white nose, tail and feet), and C) the coat color and patterning of the F1 animals; solid red with small black spots.

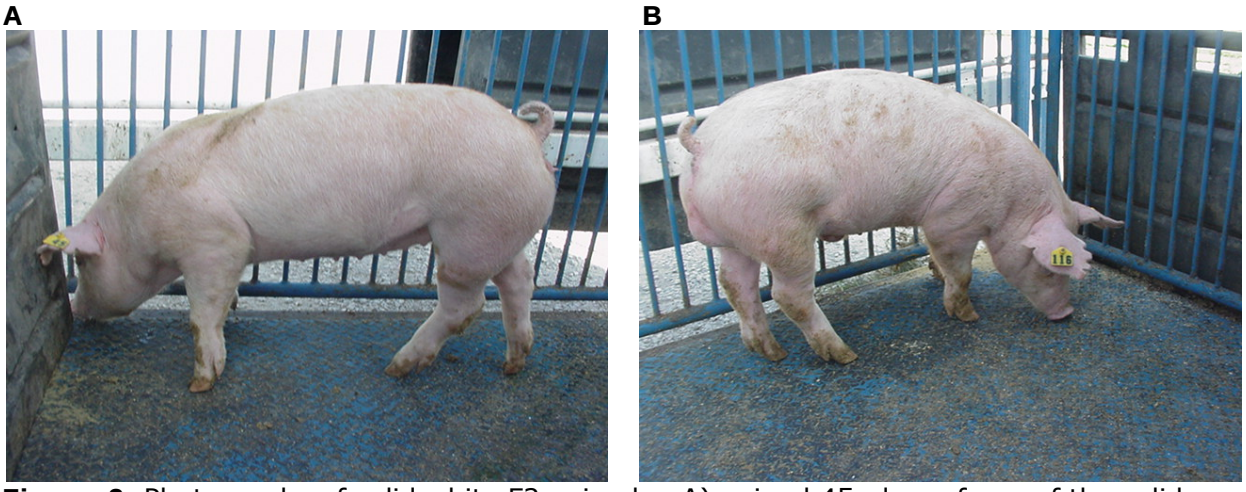

Figure 3. Photographs of solid white F2 animals. A) animal 45, dam of one of the solid white F3 litters, B) animal 116, sire of all solid white F3 animals.

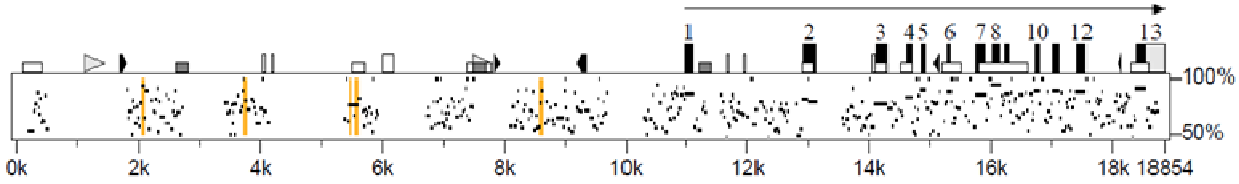

Figure 4. Percent Identity Plot (PIP) of porcine tyrosine hydroxylase. Identify compared against human tyrosine hydroxylase. The thin solid black arrow above the plot shows the location and direction of transcription, the vertical solid black blocks above the plot represent the 13 exons, orange vertical lines within the plot indicate the location of the five conserved regions with CR-I located closest to the 5' end of the promoter, the other shapes above the plot represent regions of the sequence containing repeats.

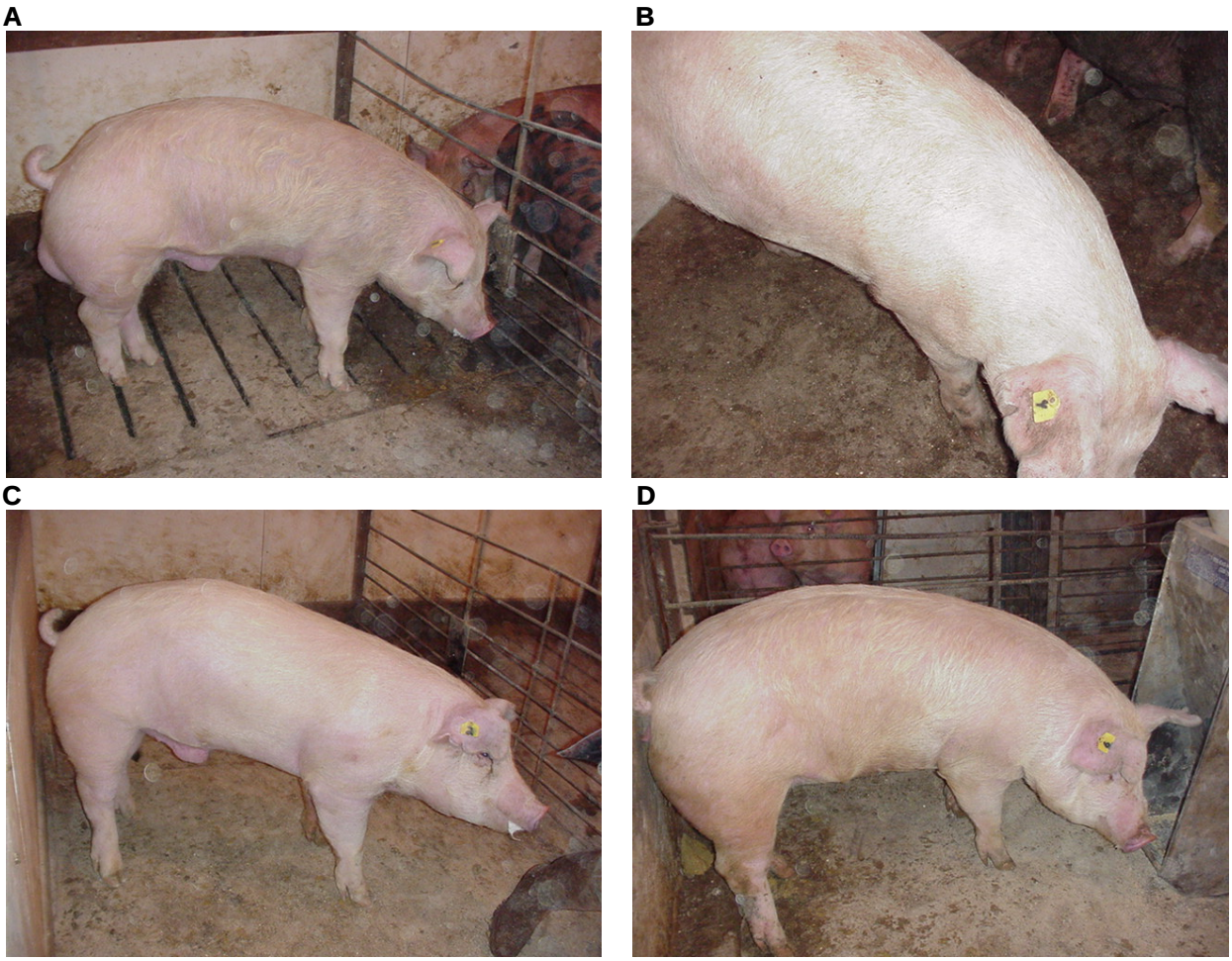

Figure 5. Photographs of solid white F3 animals. A) animal 1270B B) animal 1271S C) animal 1272B D) animal 1273.

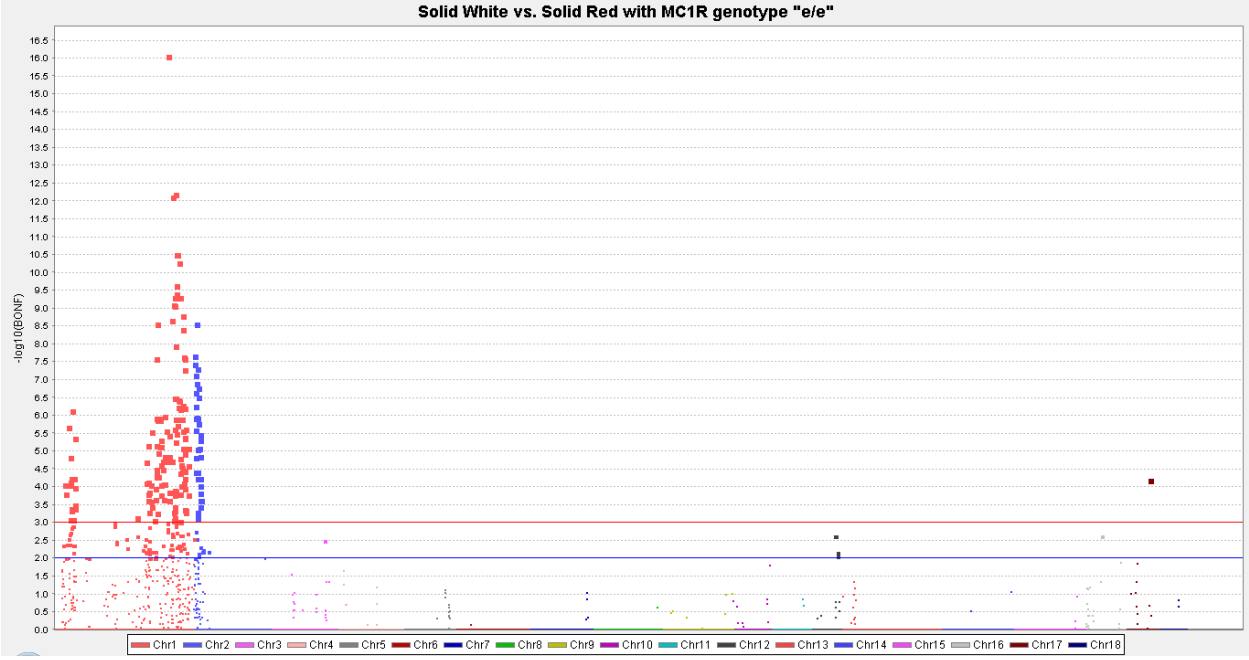

Figure 6. PLINK plot contrasting solid white with solid red pigs. Bonferroni corrected p-values from allelic chi square test contrasting solid white pigs and solid red pigs carrying the MC1R "e/e" genotype as performed by PLINK. Each autosomal chromosome coded in a different color along the x-axis and –log10 of the p-values is displayed along the y-axis. The horizontal blue line indicates p=0.01 and the horizontal red line indicates p=0.001.

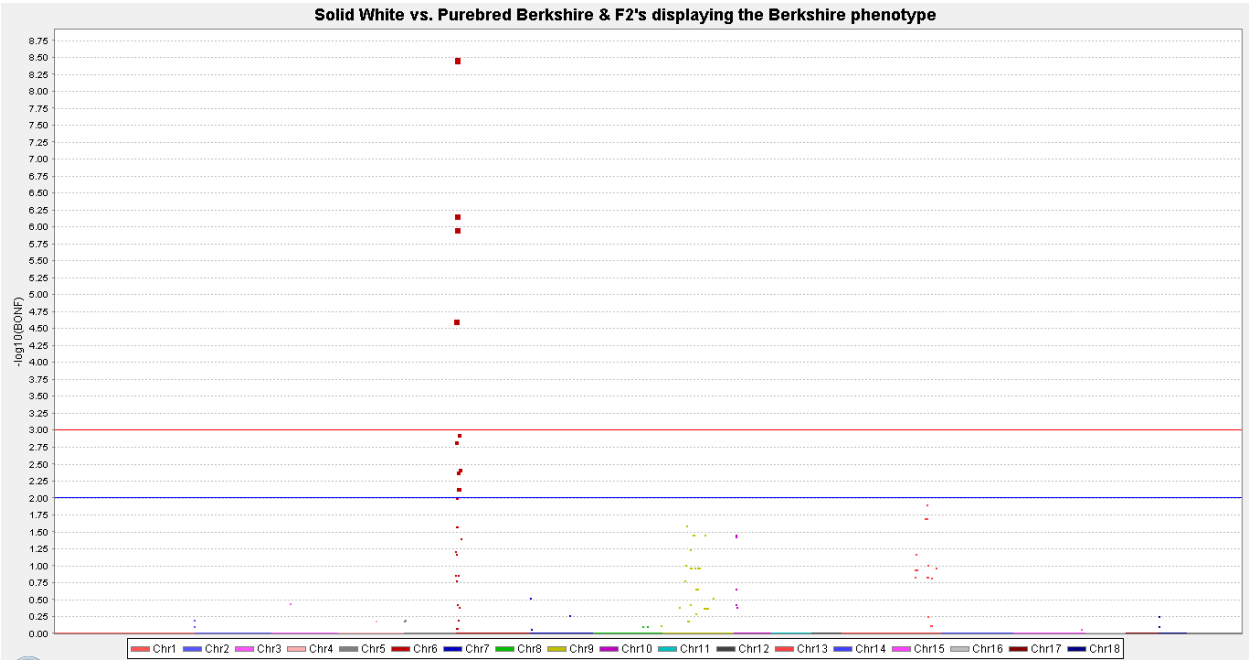

Figure 7. PLINK plot contrasting solid white with purebred Berkshire or F2 pigs displaying the Berkshire coat color. Bonferroni corrected p-values from allelic chi square test contrasting solid white pigs with purebred Berkshire and F2 pigs displaying the Berkshire coat color phenotype as performed by PLINK. Each autosomal chromosome coded in a different color along the x-axis and –log10 of the p-values is displayed along the y-axis. The horizontal blue line indicates p=0.01 and the horizontal red line indicates p=0.001.

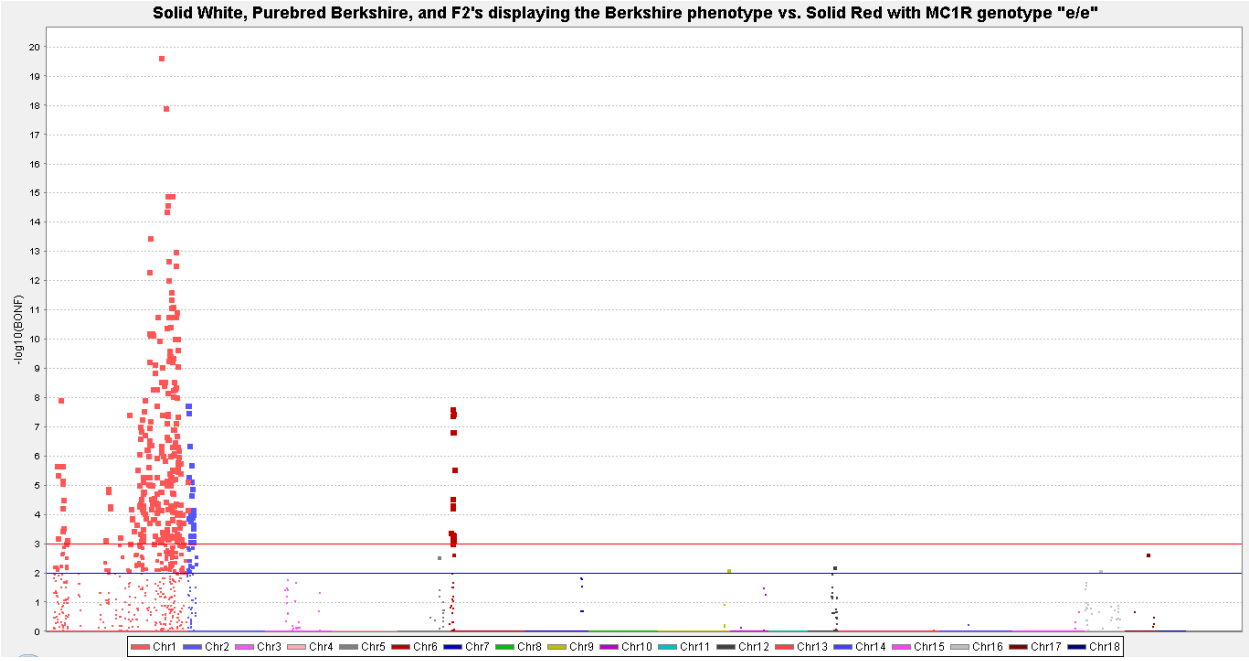

Figure 8. PLINK plot contrasting solid white, purebred Berkshire or F2 pigs displaying the Berkshire coat color with solid red pigs. Bonferroni corrected p-values from allelic chi square test contrasting solid white, purebred Berkshire, and F2 pigs displaying the Berkshire phenotype with solid red pigs carrying the MC1R genotype "e/e" as performed by PLINK. Each autosomal chromosome coded in a different color along the x-axis and –log10 of the p-values is displayed along the y-axis. The horizontal blue line indicates p=0.01 and the horizontal red line indicates p=0.001.

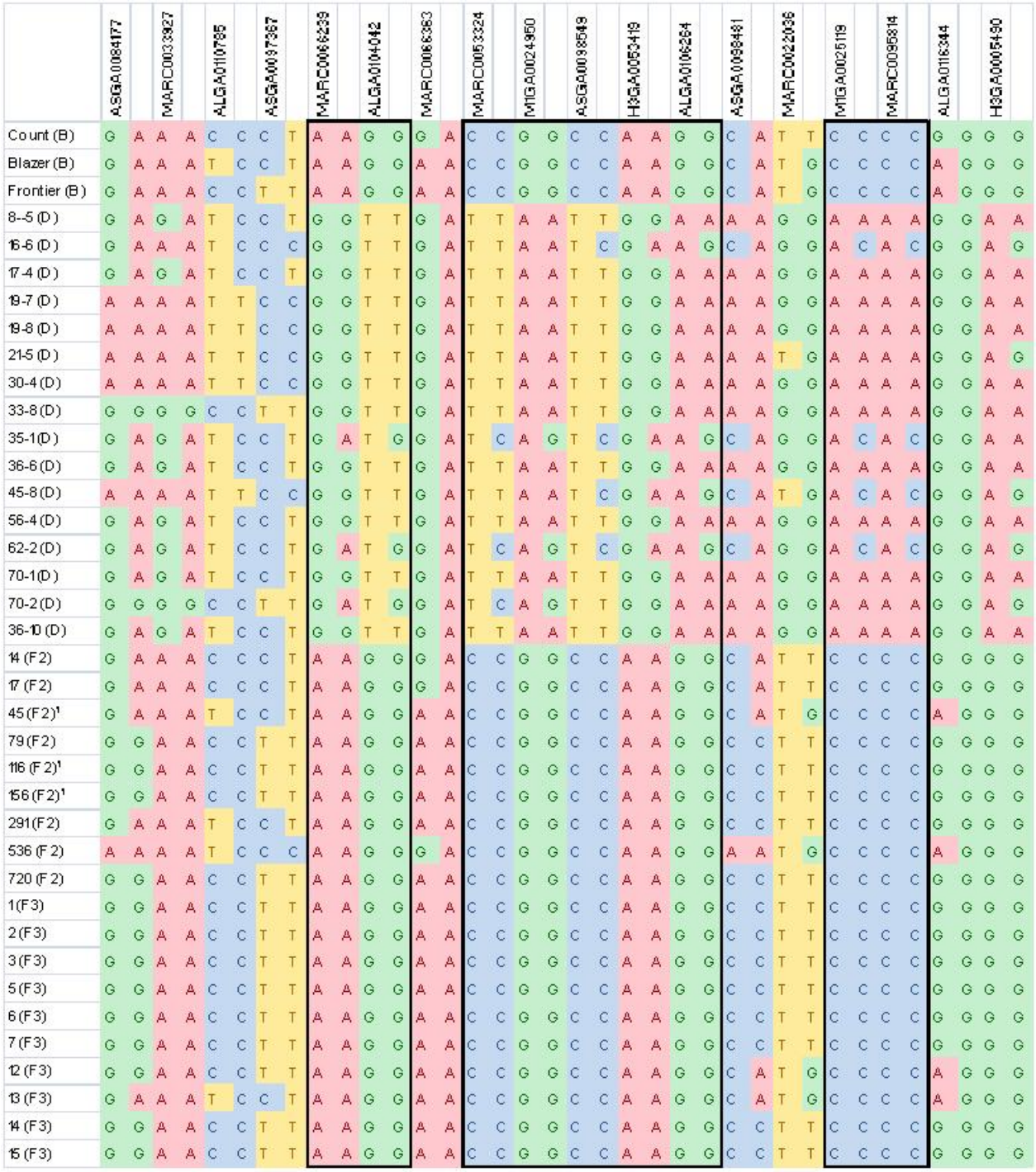

Figure 9. Homozygosity analysis of SSC2 between 0.16 and 1.19Mb. Showing SNP genotypes between 0.16 and 1.19Mb on SSC2 with each allele color coded and SNP names listed above each column. Individuals listed on the right include (in order from the top), the 3 Berkshire boars labeled with a "(B)", 16 Duroc sows labeled with a "(D)" (one sow was never sent for SNP genotyping), 9 solid white F2 animals labeled with a "(F2)", and 10 solid white F3 animals labeled with a "(F3)". SNPs surrounded by the black border show regions of homozygosity shared by the Berkshire boars and solid white animals. [1]parents of the solid white F3 pigs.

Tables

Table 2. Genome-wide association analyses performed in PLINK.

Contrast	Phenotype 1	Phenotype 2
1	Solid white (n=19)	Solid red with MC1R genotype e/e (n=68)
2	Solid white (n=19)	Purebred Berkshires and F2 individuals displaying the Berkshire phenotype (n=11)
3	Solid white, purebred Berkshires and F2 individuals displaying the Berkshire phenotype (n=30)	Solid red with MC1R genotype e/e (n=68)

Table 3. Primers used to amplify segments of Tyrosine Hydroxylase for sequencing.

Primer name	Features within fragment	Forward Primer	Reverse Primer	Annealing Temperature (C°)	Fragment Length
TH_promoter 1	Promoter	5'-CACCTGCCTCCGTCTCATTA-3'	5'-TGGGAACCTCCATAAGCTGT-3'	66	1082
TH_promoter 2	Promoter	5'-CCGCATGAATGTGCTTAGTG-3'	5'-TGGAGACTGTGGTCCAGATG-3'	61	863
TH_promoter 3	Promoter	5'-CTGTGTCCCTTCACTTCACG-3'	5'-ACGTTCTCGCTCTCTGCTTG-3'	59	1020
TH_promoter 4	Conserved Region I	5'-ACGGCAGTGACCTTCCAC-3'	5'-TGCCTCACTCGCTGACAT-3'	59	1021
TH_promoter 5	Conserved Region II	5'-CCAGATGTCCAGCTGAGAGG-3'	5'-ATTCCAGGCACTCAGGGATA-3'	59	886
TH_promoter 6	Conserved Regions III & IV	5'-GCCATAGCCGAGACCTCCTA-3'	5'-TTGCTACTGTCACCGTGTGC-3'	59	893
TH_promoter 7	Conserved Region V	5'-GTGCAGAACCAGAGGAGAGG-3'	5'-CAAACTCCCTGAGGACAGGA-3'	59	896
TH_promoter 8	Promoter	5'-CAGGGAGGCAGAGGATTAGA-3'	5'-GCCCACCGTGTTATCTCAAG-3'	58	945
TH_promoter 9	Promoter	5'-TCCCAGAGCAGGAAATAGGA-3'	5'-GTGATGGGGGTACATGGAAA-3'	58	902
TH 1	Promoter, Exon 1	5'-ATAGCACTCCCGCTGAGG-3'	5'-GGAGGCCCAGGGTTACTG-3'	68	970
TH 2		5'-GCCGAGGCCATCATGGTAAG-3'	5'-AGCTGACTGTTCTGGGCCCT-3'	65	558
TH 3	Exon 2	5'-CCGAGGCCATCATGGTAAGA-3'	5'-ACCTCGCAGCGCACAAAGTA-3'	65	2241
TH 4	Exon 2	5'-AGCTCCCTGCCCGTTGAAA-3'	5'-TCCACTGCCAGCTCCATCA-3'	59	1181
TH 5	Exons 3 & 4	5'-TGACAGCTGCTGTGAGACAA-3'	5'-ATCTGGGCAATCAGCTTCCT-3'	65	1047
TH 6	Exons 5, 6, 7	5'-CGGATGTTGGCCTGCTGTGT-3'	5'-GCGGATGTACTGGGTGCACT-3'	67	1162
TH 7	Exons 6 & 7	5'-ACAGGCAGTAAGTCCCTCCA-3'	5'-GTACTGGGTGCACTGGAACA-3'	58	942
TH 8	Exons 8, 9, 10	5'-GGACAGGAAGACCTCAGAAT-3'	5'-ACCGTGGACAGCTTCTCAAT-3'	57	1232
TH 9	Exons 10, 11, 12	5'-CCACTCACAAGGTGCCCAAG-3'	5'-CCTCTGCTGGGAAATGGGCT-3'	65	892
TH 10		5'-ACCAGCCCGTCTACTTCGTG-3'	5'-AGCGTGTACGGGTCGAACTT-3'	65	894
TH 11	Exon 13	5'-CCCATCCCTCCCACTGTAAA-3'	5'-AAGACATCTGCAGGCAGAGC-3'	59	535
TH 12		5'-AAGTTCGACCCGTACACGCT-3'	5'-ACCAAGGTGGGACAGTGCAG-3'	68	563

Table 4. Matings to determine mode of inheritance of solid white coat color.

Sire	Dam	Offspring
White	White	100% solid white
Berkshire	White	100% white with black spots
Duroc	White	100% Red "berkshire"
IMQP F1	White	25% Red "berkshire", 25% white with black spots, 37.5% red with black spots and white legs and belly, 12.5% white
White x Berkshire[1]	White	50% white, 50% white with black spots

[1] This animal is an offspring from the white dam x purebred Berkshire sire mating listed in this table.

Table 5. Polymorphisms within tyrosine hydroxylase. Each polymorphism is numbered according to the distance from the first nucleotide of the start codon. All polymorphisms were base changes with the exception of an 11 base pair indel at positions 5548-5558. The highlighted polymorphism, nucleotide 2401, is located in exon 3 and is the only polymorphism located within a coding sequence. This polymorphism causes an alanine to threonine mutation at amino acid 159 (A159T).

	Distance from TH start codon																			
	-8064	-7865	-7682	-7303	-7299	-5245	-5075	-4975	-1794	-1419	433	470	1984	2118	2401	2854	5159	5419	5548-5558	6153
Berkshire	T/T	G/G	C/C	T/T	C/C	G/G	G/G	G/G	G/G	A/A	A/A	C/C	G/G	T/T	C/C	G/G	G/G	T/T	-/-	A/A
Duroc	C/C	C/C	T/T	C/C	T/T	A/A	A/A	A/A	A/A	C/C	G/G	T/T	A/A	G/G	T/T	A/A	A/A	A/A	GACCTGGCCTT/GACCTGGCCTT	G/G
F1	T/C	G/C	C/T	T/C	C/T	G/A	G/A	G/A	G/A	A/C	A/G	C/T	G/A	T/G	C/T	G/A	G/A	T/A	-/GACCTGGCCTT	A/G
Solid White (homozygous)	T/T	G/G	C/C	T/T	C/C	G/G	G/G	G/G	G/G	A/A	A/A	C/C	G/G	T/T	C/C	G/G	G/G	T/T	-/-	A/A
Solid White (heterozygous)	T/C	G/C	C/T	T/C	C/T	G/A	G/A	G/A	G/A	A/C	A/G	C/T	G/A	T/G	C/T	G/A	G/A	T/A	-/GACCTGGCCTT	A/G

Mapping Quantitative Trait Loci Controlling Carcass Composition, Growth, and Meat Quality Traits in an F2 Berkshire x Duroc Cross

Abstract

The Illinois Meat Quality Pedigree, a three generation Berkshire x Duroc cross, was created to discover quantitative trait loci controlling 30 carcass composition, growth, and meat quality traits across the swine genome. Each animal in the IMQP population was genotyped for 137 microsatellite markers and 43,486 SNP markers. Two different QTL analyses were performed to detect QTL in this population. For the first analysis, the 137 microsatellite markers were used to create a linkage map that was used by the QTLexpress software to perform a regression interval analysis. For the second analysis, the SOLAR software was used to perform a variance component analysis using a physical map created from the 43,486 SNP markers. Nintety-five QTL were detected using the QTLexpress software while 787 QTL were detected using the SOLAR software. Of these QTL, 40 of the QTL detected by QTLexpress were located at the same position 119 of the QTL detected by SOLAR. These results will serve as the starting point for fine mapping efforts to be conducted in the future to determine the causative mutations behind the significant regions obtained in this study.

Introduction

Pork is an important food source for the majority of the world and with the world's population nearing seven billion [1], the swine industry needs to find a way to produce enough pork while continuing to improve the pork quality. Many of the phenotypes involved in reaching these goals in pigs including, growth, carcass composition and meat quality traits, are quantitative and are often controlled by multiple genes that have a small impact

on the trait. These characteristics make it difficult to locate and identify every gene responsible for a quantitative trait. One technique used to help identify these genes is a quantitative trait locus (QTL) analysis. To date, the results for over 35 whole genome QTL analyses controlling carcass composition, growth and meat quality traits performed on various pig breed intercrosses have been published (Table 1) [20, 36, 61, 153-187].

The majority of studies performing QTL analyses in pigs use linkage maps consisting of less than 200 microsatellite markers spanning the entire genome. With the development of the Illumina® PorcineSNP60 BeadChip [65], the ability to obtain thousands of genotypes for many animals has become economically feasible. For this study we created two different maps, one consisting of 137 microsatellite markers and the other consisting of 43,486 SNP markers obtained from genotyping our population using the Illumina® PorcineSNP60 BeadChip [65]. The aim of this study was to use these maps and perform an analysis to identify QTL controlling carcass composition, growth, and meat quality phenotypes in a Berkshire x Duroc cross.

Materials and Methods

IMQP Reference Population

The Illinois Meat Quality Pedigree (IMQP) is an F2 Berkshire x Duroc intercross population consisting of 664 individuals. Semen from three purebred Berkshire boars, chosen based on progeny test performance values, was used to inseminate 17 purebred Duroc sows chosen from the breeding program at the University of Illinois Moorman Swine Research Unit. Each boar was bred to an average of six different sows, but each sow was only mated to one of the three boars. All founder animals were genotyped for the halothane and Rendement Napole (R200Q) mutations and were found to be normal. An F1 population was produced, from which five boars and 44 sows were chosen to be mated. No mating of relatives

occurred within the F1 intercross. Each of the F1 boars were mated to an average of 11 different F1 sows, while each F1 sow was mated to either one or two different F1 boars. The intercross between F1 individuals created an F2 generation consisting of 595 pigs.

Phenotypic Data

Thirty different phenotypic values were collected on 591 F2 individuals (Table 6). Each phenotype could be classified as either a meat quality, carcass composition, or growth trait. Growth trait data included weight at birth, weight when entering the nursery, and weight when entering the grower/finisher. Nineteen different meat quality values were collected on each animal, including thirteen objective measurements (pH at 45 minutes; pH at 24 hours; average glucose, glycogen, and glucose-6-phosphate; average lactate; glycolytic potential; Hunter L*; Hunter a*; Hunter b*; shear force; fat percentage; cook loss percentage; moisture percentage; and drip loss percentage) and six subjective scores (marbling, juiciness, tenderness, off flavor, color and firmness). Eight different carcass composition measurements were taken on each animal including carcass length, loin eye area, dressing percentage, average back fat thickness, and back fat thicknesses measured at the first rib, 10^{th} rib, last rib, last lumbar vertebrae. In addition to these phenotypic measurements, sex, birth year month, hot carcass weight, harvest group, parity of dam, litter size, and age were also recorded for each individual to be used as covariates in the QTL statistical analysis.

Genetic Marker Maps

Microsatellite Markers

The microsatellite markers used in this study were chosen based on the heterozygosity, number of alleles, number of informative meioses and position according to the USDA-MARC swine genome database (www.marc.usda.gov). Markers with publicly available sequences had their primer sequences re-designed using Primer-3 (http://frodo.wi.mit.edu/cgi-

bin/primer3/primer3_www.cgi) so all markers had a length of 20-22 base pairs, product size between 150 and 300 base pairs, GC content between 45 and 60%, and annealing temperature of 58°C. To check for heterozygosity and fragment size in the IMQP population, the forward primer of each microsatellite marker was radioactively labeled with ^{32}P and then amplified individually in a 10 μl polymerase chain reaction (PCR) composed of 30 ng of DNA; 1X PCR buffer (containing 1.5mM $MgCl_2$) (Qiagen®); 200 μM of each dNTP (Fermentas); 0.5 μM of each primer and 0.25 U of HotStar® Taq DNA polymerase (Qiagen®). PCR amplification conditions were as follows: initial denaturation at 95°C for 15 min, followed by 35 cycles of 94°C for 30 sec, 58°C for 45 sec and 72°C for 45 sec, and finally an extension step of 72°C for 45 min. PCR product was run on 6% acrylamide gels and the gels were then exposed to X-ray film overnight. Microsatellite markers found to be heterozygous in the F1 boars had their forward primers fluorescently labeled with one of four dyes (6-FAM™, VIC®, NED™ and PET®) from Applied Biosystems™. Each microsatellite marker was then multiplexed together with between one and four other markers and the multiplexes were amplified together in a 10 μl polymerase chain reaction composed of 30 ng of DNA and 1X Qiagen® Multiplex PCR Master Mix (Qiagen®). PCR amplification conditions were as follows: initial denaturation at 95°C for 15 min followed by between 29 and 24 cycles of 94°C for 30 sec, 58°C for 1 min 30 sec and 72°C for one min and finally an extension step of 60°C for 45 min. PCR products were then treated with ExoSAP-IT® (USB®) to remove remaining primers and dNTPS. Five μl of the ExoSAP-IT® product was then purified using the Promega plasmid DNA purification system (Promega). Each sample was mixed with 25 μl of isopropanol, added to a 96-well purification plate and dried using the purification system. Next the sample was washed with 200 μl of 80% ethanol three times. Samples were then provided to the W.M. Keck Center for Comparative and Functional Genomics where capillary electrophoresis was done using ABI Prism 3730xL Analyzers (Applied Biosystems™). Genotypes were viewed and inheritance of alleles was checked using GeneMarker® (Softgenetics®).

A total of 137 markers with twopoint linkage LOD scores greater than 3 were used to create genetic maps of each porcine autosomal chromosome using CRIMAP v2.4 [188]. Mendelian errors and pedigree errors were checked using the PREPARE option of CRIMAP and calculation of twopoint recombination frequencies between each marker pair was performed using the TWOPOINT option of CRIMAP. Though centiMorgan positions along each chromosome were calculated, for ease of comparison with results from the SNP marker analyses, a BLAST similarity search [34] against the Sscrofa10 genome assembly was performed for each microsatellite to obtain physical map position. The BLAST similarity search [34] was unsuccessful for some of the markers, thus a BLAST similarity search [34] was performed using microsatellite markers that were at similar positions according to the MARC linkage map and their physical positions were used to approximate the location of the markers.

SNP Markers

Fifteen microliters of genomic DNA at a concentration of 100ng/µl was sent to GeneSeek® to SNP genotype the IMQP population using the Illumina® PorcineSNP60 BeadChip [65]. The list of SNPs used in the subsequent analyses was pruned down from the initial list of SNPs using PLINK v1.05 [66] by removing Mendelian errors, removing SNPs with a minor allele frequency less than 0.01, and removing SNPs with a missing genotype frequency of 0.9 or greater. Correlation coefficients (r^2) between all markers was calculated using PLINK v1.05 [66].

Quantitative Trait Loci Analysis

Microsatellite Markers

Identification of QTL using the microsatellite marker dataset was performed using the F2 regression analysis option of QTLexpress [189]. Physical maps estimated by a BLAST

similarity search [34] against the Sscrofa10 genome assembly performed for each microsatellite marker were used in place of linkage maps for these analyses. A one QTL model analysis was run for each phenotype on each chromosome, with F-values, additive and dominance estimates calculated every 1.0 Mb. A permutation test, using 5000 iterations, was performed using QTLexpress to calculate the 0.05 and 0.01 chromosome-wise empirical p-value thresholds for each QTL [190].

For each phenotype, fixed effects and covariates were tested for significance using the POLYGENIC command in SOLAR. Though QTLexpress has the option of categorizing fixed effects and covariates separately, SOLAR treats all fixed effects and covariates the same. For the QTLexpress analyses; sex, parity, birth year month and harvest group were categorized as fixed effects while age, hot carcass weight, litter size, nursery-in weight, grower/finisher-in weight, and age at harvest were categorized as covariates. Ninety-five percent confidence intervals (95%CI) for QTL were defined as where the LOD score is decreased by one [16].

SNP Markers

Identification of QTL using SNP markers was performed using variance component linkage analysis. The Sequential Oligogenic Linkage Analysis Routine software package (SOLAR) was used to do all of the subsequent calculations and variance component twopoint linkage analysis. Allele frequencies were estimated using pedigree based maximum likelihood estimates (MLE). These estimates were computed using the FREQ MLE command in SOLAR, a command that runs an external program Allfreq, an extension of the MENDEL program [191]. The IBD command in SOLAR was used to compute marker-specific identity-by-descent (IBD) matrices using a Monte Carlo algorithm. The POLYGENIC command in SOLAR calculates the heritability (H2r) of each phenotype, the significance of the H2r, creates a model to be used during the twopoint linkage analysis, tests each covariate for significance

in the model, and calculates the proportion of variance contributed by each covariate. Covariates were removed from the model if they did not reach a statistical significance ($p<0.05$). The POLYGENIC command also calculates the residual kurtosis for each phenotype and determines if a LOD score adjustment is needed. LOD score adjustment values were calculated using the LODADJ command and applied to the twopoint linkage results for each analysis regardless of the recommendation by the POLYGENIC command. A twopoint linkage analysis was performed using the TWOPOINT command in SOLAR. For each marker, the TWOPOINT command computes a LOD score, the amount of genotypic variance controlled by the marker, and the residual genetic variance. SNPs were considered significant if a LOD score of 3 was reached within an analysis. Significant regions were defined as any group of at least two SNPs that have a LOD score of 3 and are within 1Mb of each other. SNPs further apart then 1Mb were considered part of a significant region as long as the SNPs had an r^2 of 0.2 or greater.

Results

To investigate the effect of marker density on the ability to identify QTL within a genome, whole genome maps were created using two different molecular marker types. A sparse map of the 18 porcine autosomal chromosomes was created using microsatellite markers while a more dense map was created using SNP markers. A total of 137 microsatellite markers were used to create linkage maps for each chromosome. Each chromosome map contained between four and ten markers with an average spacing of 17.2Mb (Table 7). A total of 56,046 SNP markers had genomic positions according to the Sscrofa10 genome assembly and these positions were used to create physical maps using the SNP markers. Pruning out SNPs with a minor allele frequency of less than 0.01 and a missing genotype frequency of 0.9 or greater left a final physical map consisting of 43,486 SNP markers (Supplementary Files).

The POLYGENIC command in SOLAR was used to test the significance of the covariates for each phenotype and found that each phenotype had between zero and four significant covariates (Table 8). The POLYGENIC command also calculated the heritability of the phenotypes (H2r) used in this study ranged between 0.01 and 0.83 and the proportion of variance contributed by each covariate ranged between 0 and 0.34 (Table 8).

To identify chromosomal regions controlling meat quality, carcass composition, and growth traits in the IMQP population, QTL analyses were performed on both the microsatellite marker and SNP marker data. Due to limitations within many of the publically available QTL analysis softwares, QTLexpress was chosen to analyze the microsatellite data while the Sequential Oligogenic Linkage Analysis Routine software package (SOLAR) [192] was chosen to analyze the SNP data.

QTLexpress Analyses

QTLexpress identified a total of 95 QTL with a significance level at or above a chromosome-wise p value of 0.05 that control carcass composition, growth and meat quality phenotypes across the porcine autosomal chromosomes (Table 9). Of the 95 significant QTL, 59 were identified as significant at the 0.05 chromosome-wise level, 26 were identified as significant at the 0.01 chromosome-wise level, three were identified at the 0.05 genome-wide level, three were identified as significant at the 0.01 genome-wide level and 4 were identified as significant at the 0.001 genome-wide level. At least one significant QTL was identified on each of the chromosomes, but the number of QTL per chromosome varied greatly, with chromosomes 10 and 13 having only one significant QTL and chromosome 2 having 13 significant QTL. At least one QTL was identified for each of the 30 phenotypes analyzed with the number of significant QTL per phenotype varying between 1 and 7 per phenotype. Ninety-five percent confidence intervals were calculated for all QTL with interval sizes

ranging from 8 to 148Mb and an average interval size of approximately 38Mb (Table 9). Twenty six of the QTL were identified near the end of their respective linkage maps, making calculation of the true size of the intervals unfeasible due to the fact that the LOD score for the markers did not decrease by one before the end of the linkage map was reached (Table 9).

SOLAR Analyses

Over 1.3 million twopoint linkage analyses were performed for this study (Appendix B: SOLAR whole genome QTL scan results files). Out of these analyses, SOLAR identified a total of 10,309 SNPs with a LOD score over 3.0 across the porcine autosomal chromosomes (Appendix C: SOLAR whole genome QTL scan; SNPs with LOD scores greater than 3). The 10,309 SNPs were assembled into 778 different significant regions, with a significant region defined as at least two SNPs with a LOD score greater than 3.0 that are within 1Mb of each other or SNPs further than 1Mb that have an r^2 greater than 0.2 (Appendix C: SOLAR whole genome QTL scan; SNPs with LOD scores greater than 3). Of the 10,309 significant SNPs, only 821 SNPs were not part of a significant region. Each phenotype analyzed had between 10 and 1661 SNPs grouped into 3 to 222 significant regions (Appendix C: SOLAR whole genome QTL scan; SNPs with LOD scores greater than 3).

Discussion

Significant Results Found by Both QTLexpress and SOLAR

Of the 95 QTL identified by QTLexpress, the 95%CI of 39 of them overlapped with at least one significant region as identified by the SOLAR analyses (Table 9). These 39 regions overlapped with 119 of the 787 regions identified by SOLAR (Appendix C: SOLAR whole genome QTL scan; SNPs with LOD scores greater than 3).

Chromosome 1 had only one QTL, the 95%CI for the QTL controlling DRESS% ranges from 50.9 to 130.9Mb with the most significant position located at 59.9Mb and overlaps with three significant regions as identified by SOLAR. These three regions (DRESS%, regions 1-3) range from 111.2 to 117.3Mb and in total, contain 10 SNPs. Even though the 95%CI from QTLexpress and the significant regions identified by SOLAR are overlapping, the most significant position from the QTL express analysis was located over 60Mb from the significant regions identified by SOLAR, suggesting the two programs are identifying different loci controlling the DRESS% QTL on SSC1.

In total, ten QTL were identified on SSC2, including QTL controlling LEA, BFTLL, BFTTR, BFTAV, MARB, JUICY, FAT%, SHEAR, TENDER, and COLOR. The most significant positions of the QTL controlling LEA, BFTLL, BFTTR, and BFTAV were all identified at 25.3Mb and all shared the same 95%CI ranging from 25.3 to approximately 35Mb, suggesting they may all be controlled by the same loci. Though no good candidate genes are located within the 95%CIs, the 95%CIs for these traits have been truncated at 25.3Mb because the linkage map for SSC2 does not extend any further up the p arm of SSC2. Based on the location of the most significant SNPs identified by the SOLAR analysis for these traits and considering the type of traits sharing this QTL, the insulin-like growth factor 2 gene (IGF2) would be an excellent candidate gene. The IGF2 gene is an imprinted gene that encodes a polypeptide growth factor involved in growth and development and a well characterized mutation inherited from the sire, IGF2 intron3-g.3072G>A, has been shown to increase muscle mass while decreasing backfat thickness in pigs [26]. Though not genotyped for the intron3-g.3072G>A mutation in IGF2, it is assumed that the Duroc animals in this population carry the intron3-g.3072A allele. The additive estimates calculated by QTLexpress for all of these traits are consistent with this assumption, showing that the Duroc allele decreases backfat thickness as it increases the size of the loin eye muscle, adding to the evidence that IGF2 may be the causative gene controlling the variation in these traits on SSC2. The most

significant positions for the QTL controlling MARB and JUICY were locate close together, at positions 64.3 and 69.3Mb respectively, and they had overlapping 95%CI suggesting they may be controlled by the same locus. The additive estimates suggest that the Berkshire allele increases both the marbling and juiciness of the meat in this region of the genome. The QTL controlling TENDER and SHEAR on SSC2 shared similar 95%CIs and the most significant positions identified by QTLexpress were very close together, 115.3and 111.3 respectively. For both TENDER and SHEAR, the 95%CIs identified by QTLexpress corresponded to a single significant region (TENDER, region 8 and SHEAR, region 13) as identified by SOLAR. The additive estimates for these traits show that the Duroc allele increases shear force and decreases the tenderness of the meat, this is consistent with the relationship between these traits, suggesting that these two QTL may be controlled by the same locus. To narrow the QTL intervals for these traits, this region of SSC2 was fine mapped by increasing the number of markers in the region and repeating the QTLexpress analysis [193]. The calpastatin gene (CAST), was identified as a candidate gene and sequenced to discover causative polymorphisms associated with the tenderness and shear force phenotypes [194]. Almost 400 polymorphisms, distributed throughout the gene, were identified and it was concluded that the causative mutation was likely to be regulatory in nature [194]. The 95%CI for the QTL controlling FAT% was found to be overlapping with the 95%CI for SHEAR and TENDER, but the most significant position was at 91.3Mb suggesting that this QTL may be controlled by a different locus than the locus controlling SHEAR and TENDER. The QTL controlling COLOR is located at 145.3Mb and though the 95%CI does overlap with 95%CIs of some of the other traits, the most significant position is far enough away from the other traits that it is very unlikely it is controlled by the same loci as any of the other traits.

Chromosome 3 had two QTL identified as significant by both QTLexpress and SOLAR. The QTL controlling GGG and GP have very similar 95%CIs and the most significant positions

are very close together. The SOLAR analysis also showed overlapping significant regions (GGG, region 25 and GP, regions 5-7) and in fact SOLAR identified the same SNP as being most significant for both traits. Since the values for GP are calculated using the GGG values, it is likely that both these traits are being controlled by the same locus on SSC3. Also, the additive effects calculated by QTLexpress show that both traits are being increased by the Duroc allele, adding to the evidence that both traits may be controlled by the same locus.

Chromosome 4 had no QTL that overlapped according to the results from QTLexpress and SOLAR, but this is most likely due to the fact that the linkage map used by QTLexpress does not cover the first 21.0Mb of the p arm of the chromosome. The QTLexpress results show that the upper boundary of the 95%CIs for two QTL on SSC4 coincided with the last marker used to create the linkage map while the significant regions identified by SOLAR were located just upstream of the upper boundary of the linkage map. Had the linkage map covered more of the p arm of the chromosome, there would most likely be an overlapping area identified by QTLexpress and SOLAR. The two QTL identified by on SSC4 included one controlling DRESS% and one controlling LENGTH. The QTLexpress results show the 95%CIs for both QTL span from 21.0 to 67.0Mb but indicate that the most significant positions were located more than 20.0Mb away from each other. The SOLAR results show the significant region for the DRESS% QTL (DRESS%, region 7) spans from 11.4 to 15.9Mb with the most significant SNP being located at 11.9Mb while the significant region for the LENGTH QTL (LENGTH, region 2) spans from 15.9 to 16.1Mb with the most significant SNP being located at 15.9Mb. Together these results suggest there may be more than one locus controlling these two traits on SSC4, but further investigation is needed to confirm this hypothesis.

Chromosomes 5 and 6 each had a single QTL identified by both QTLexpress and SOLAR. The QTL identified on SSC5 is a LENGTH QTL while the QTL on SSC6 is a GROW QTL. The

95%CI for the LENGTH QTL on SSC5 spanned 55Mb and corresponded to 6 different significant SOLAR regions (LENGTH, regions 3-8). The most significant position identified by the QTLexpress analysis was located at 63.7 Mb and one significant region (LENGTH, region 8) contained this position, but this region did not contain the most significant SNP in this part of the chromosome. The region that contained the most significant SNP (LENGTH, region 3) ranged from 26 to 36Mb with the most significant SNP located at 33Mb. Since the most significant position identified by QTLexpress and the most significant SNP identified by SOLAR were located at such different points on the chromosome it may suggest more than one loci is controlling LENGTH on SSC5. The 95%CI for the GROW QTL on SSC6 spans almost 20Mb but the significant region identified by SOLAR (GROW, region 5) contains only two SNPs and spans less than 1Mb. The most significant position identified by QTLexpress is located at 140.1Mb and is very close to the most significant SNPs identified by SOLAR, at 141.7Mb, suggesting the locus controlling LENGTH on SSC6 may be located near this position.

Chromosome 7 had four significant QTL identified as significant by both QTLexpress and SOLAR including a QTL controlling LEA, LENGTH, DRESS%, and HUNTA. Even though all four traits had overlapping 95%CI it is unlikely that these traits are being controlled by the same locus on SSC7. The 95%CI of the QTL controlling LEA corresponds to seven different significant regions (LEA, regions 27-33) as identified by SOLAR, but the most significant position identified by QTL is not contained within any of these regions. The most significant SNP identified by SOLAR is located approximately 45Mb away from the most significant position identified by QTLexpress, so further investigation is needed to identify the locus controlling LEA on SSC7. The 95%CI for the QTL controlling LENGTH on SSC7 corresponded to seven different significant regions (LENGTH, regions 16-22) but the most significant position identified by QTLexpress was not located within the same significant region as the most significant SNP as identified by SOLAR (LENGTH, region 16). Though the most

significant position identified by QTLexpress and the most significant SNP identified by SOLAR were only 8Mb apart, it is unlikely to be the same locus, so further investigation is needed to determine where the candidate gene may be located. The 95%CI for the DRESS% QTL on SSC7 spans from 22.7 to 105.7Mb, while the two significant regions identified by SOLAR (DRESS%, regions 8-9) only cover from 28.0 to 74.0Mb. Similar to the LEA QTL on SSC7, the most significant position identified by QTLexpress is not contained within either of the significant regions identified by SOLAR, but rather it is located between these two regions. Further investigation is needed to identify the location containing the locus controlling this QTL. The QTL controlling HUNTA was identified by QTLexpress at 56.7Mb, while this position was within one of the three significant regions corresponding to the 95%CI for this QTL, it was not located in the same significant region (HUNTA, regions 3-5) as the most significant SNP as identified by SOLAR. The most significant position identified by QTLexpress was located at 56.7Mb and the most significant SNP identified by SOLAR was located at the opposite end of the 95%CI at 32.8Mb, suggesting the two programs may be identifying two different loci controlling the HUNTA QTL on SSC7.

Chromosome 8 contains two QTL, one controlling BFTLL and the other controlling LENGTH. The 95%CI for these two traits overlap and the most significant positions identified by QTLexpress are close together, 120.3 and 128.3Mb respectively, suggesting these two QTL may be controlled by the same loci. Genetic correlations suggest a negative correlation between these two traits and this is in agreement with the additive effects seen in the IMQP population, that show the Berkshire allele increases the length of the animal as it decreases the backfat at the last lumbar vertebrae [195]. However, the results from the SOLAR analysis show the most significant SNP for the BFTLL QTL is located at 98.3Mb while the most significant SNP for the LENGTH QTL is located at 133.3Mb. Though the two traits have overlapping 95%CI, there is a possibility that the two traits are controlled by two different

loci due to the fact that both the most significant position and most significant SNP for the BFTLL QTL are not contained within the 95%CI identified for the LENGTH QTL.

Chromosomes 11 and 12 each had only one significant QTL identified by both QTLexpress and SOLAR, with a FAT% QTL identified on SSC11 and a MOIST QTL identified on SSC12. The significant regions (FAT%, regions 55-57) for the FAT% QTL on SSC11 span the entire 95%CI, but the most significant position identified by QTLexpress is not contained within one of the significant regions due to a gap between two of the regions. Though the most significant position is not contained within one of the regions, the most significant position is less than 1Mb from the most significant SNP, suggesting the two programs are identifying the same locus controlling this QTL. The results from QTLexpress show the most significant position for the MOIST% QTL on SSC12 is located at 28.4Mb while the most significant SNP as identified by SOLAR is located at 41.9Mb. Since QTLexpress and SOLAR show such a difference between the most significant positions, the programs may be identifying different loci controlling this QTL.

Three QTL were identified by QTLexpress and SOLAR on SSC14. The 95%CI for the LENGTH QTL on SSC14 spanned from 103.3 to 126.3Mb and corresponded to a single significant region (LENGTH, region 50) spanning from 120.8 to 123.6Mb. Though the most significant position identified by QTLexpress, 120.3Mb, was not within this significant region, it is less than 1Mb from the edge of the significant region and the most significant SNP within the region is located at 122.1Mb. The results from both programs seem to be in agreement regarding a small region of the chromosome where a locus controlling this QTL may be located, suggesting between 120 and 122Mb should be where a search for a candidate gene should begin. The other two QTL on this chromosome, one controlling GGG and one controlling GP, shared the same 95%CI and the most significant positions identified by QTLexpress were also at the same position. The QTLexpress analysis showed the most

significant position was located at 135.3Mb and the SOLAR analysis identified the most significant SNP at 132.9Mb for GGG and 134.7 for GP. The additive effects calculated by QTLexpress show that the Duroc increases the GGG at the same time it increases the GP, this is consistent with the biological relationship between these two traits, adding to the evidence that these two traits are being controlled by the same locus in this region of SSC14.

Four different QTL were identified on SSC15; BFTFR, BFTLL, BFTAV, and GGG. According to the QTL analyses, all four share the same 95%CI, spanning from approximately 104.6 to 134.6, and the most significant positions for each QTL were also the same, at approximately 130.0Mb. The SOLAR analyses showed that significant regions for the BFTFR (BFTFR, region 5), BFTAV (BFTAV, region 16), and GGG (GGG, regions 188-189) QTL all spanned from approximately 120.0 to 122.0Mb. Though the significant region for BFTLL (BFTLL, region 35), spanning from 133.63 to 133.67Mb, was not overlapping with the significant regions for the other three QTL, it was located close to the most significant position identified by QTLexpress. Though it does not seem intuitive that these four traits would be controlled by the same locus, this region of the genome contains the protein kinase, AMP-activated, gamma 3 non-catalytic subunit (PRKAG3) gene, also known as the Rendement Napole gene. This gene encodes a regulatory subunit of adenosine monophosphate-activated protein kinase (AMPK), an enzyme that phosphorylates both glycogen synthase and hormone-sensitive lipase. Glycogen synthase is an enzyme involved in the conversion of glucose to glycogen, explaining the presence of the GGG QTL in this region of the chromosome and hormone-sensitive lipase is involved in the hydrolysis of fatty acids from adipose tissue, explaining the presence of the backfat QTL in this region of the chromosome. The IMQP population was genotyped for several sequence variations within PRKAG3 to make sure no animals carried the R200Q mutation. The R200Q mutation causes an increase in glycogen levels in meat and results in poor meat quality [22]. Though no animals carried the R200Q

mutation, there was variation in the PRKAG3 genotypes in the IMQP population, so the genotypes were used as a marker on the linkage maps used by the QTLexpress analysis, and in fact the most significant positions identified by QTLexpress for all four traits are located at the same location as the PRKAG3 marker. The additive effects for these four QTL as calculated by QTLexpress show the favorable allele comes from the Berkshire breed, resulting in lower backfat values and lower GGG values. This is in agreement with previous evidence that shows the Berkshire breed having a high frequency of the PRKAG3 haplotype corresponding to good meat quality values [37]. Based on the evidence above, PRKAG3 is a good candidate gene controlling the BFTFR, BFTLL, BFTAV, and GGG QTLs on SSC15.

A QTL controlling FAT% was the only QTL identified on SSC16. The QTLexpress analysis identified the 95%CI as spanning from 40.6 to 57.6Mb, while the SOLAR analysis showed significant regions (FAT%, regions 85-86) spanning from 32.7 to 55.2Mb. Even though the most significant position according to QTLexpress, 57.6Mb, is not contained within the significant regions identified by SOLAR, the most significant position is very close to the edge of one of the regions, suggesting the two programs are identifying the same locus controlling the FAT% QTL.

Five meat quality QTL were identified on SSC17, including one each controlling 24PH, GGG, GP, LAC, and FIRM. The 24PH, GGG, GP, and LAC QTL all shared a similar 95%CI spanning from approximately 33.5 to 47.5, also the most significant positions identified by QTLexpress for these four QTL were located in a similar location, at approximately 40.5Mb. The SOLAR analyses also showed an overlap of the significant regions for these four QTL (24PH, regions 31-35; GGG, regions 207-209; GP, regions 72-77; LAC, regions 6-9), spanning from approximately 34.9 to 45.8Mb. Considering the overlapping 95%CI and significant regions for these four traits it is reasonable to suggest that these traits could be controlled by the same loci, and this hypothesis is reinforced by the biological relationship

between these traits coupled with the additive effect values calculated by QTLexpress. There is a positive relationship between glucose, glycogen, glucose-6-phosphate (GGG), lactate (LAC) and glycolytic potential (GP) biologically. In muscle both glucose and glycogen can be converted into glucose-6-phosphate, this is then converted to pyruvate through the glycolysis pathway, and during anaerobic conditions, including the time period after a pig is harvested, pyruvate is converted to lactate. Glycolytic potential values are calculated using the formula; glycolytic potential = 2 x (glucose + glycogen + glucose-6-phosphate) + lactate. In addition to the positive relationship between GGG, GP and LAC, there is a negative relationship between these three traits and 24PH, this can be explained biologically in that as glucose is converted into lactate post harvest, the pH of the meat will decrease. The additive effects calculated by QTLexpress show that the Duroc animals carry the favorable allele on SSC17, in that the Duroc allele increases the 24PH values while decreasing the GGG, GP and LAC values. The other QTL identified on SSC17, FIRM, has a 95%CI that overlaps with the other four traits, but extends further upstream than the other four. Also, the 95%CI shared by the other four traits does not encompass the most significant position for the FIRM QTL. The SOLAR analysis also identifies the most significant region (FIRM, region 7) further upstream than the regions for the other traits. Despite both analyses indicating the FIRM QTL is located upstream from the other traits, FIRM is biologically related to the other traits, as seen when firmness of meat decreases as pH of meat decreases. This is in agreement with the additive effects calculated by QTLexpress that shows the Duroc allele increases FIRM as it increases 24PH and decreases GGG, GP, and LAC. Considering the evidence above, further analysis will be needed to determine if the FIRM QTL is being controlled by the same locus as the 24PH, GGG, GP and LAC QTLs.

Chromosome 18 had a GGG QTL and a FAT% QTL identified by both QTLexpress and SOLAR. The 95%CI for the GGG QTL spans from 38.7 to 52.7Mb and the most significant

position according to QTLexpress was identified at 47.7Mb. The SOLAR analysis identified the significant regions (GGG, regions 218-221) as spanning from 41.0 to 53.0Mb with the most significant SNP being located at 51.1Mb. The QTLexpress analysis identified the 95%CI for the FAT% QTL as spanning from 46.7 to 52.7Mb with the most significant position located at 50.7Mb while the results from the SOLAR analysis showed the significant region (FAT%, region 92) spanned from 51.7 to 51.8Mb with the most significant SNP located at 51.8Mb. The 95%CIs for these two QTL have extensive overlap and both QTL show the most significant SNP as being located at 51Mb, so it is possible that these two QTL could be controlled by the same locus, though further analysis is needed to confirm this hypothesis.

Significant Results Found by QTLexpress Only

Of the 95 QTL identified by QTLexpress, 55 QTL did not overlap with any significant regions as identified by SOLAR. There are several explanations why SOLAR may not have identified the SNPs in these regions as being significant, including a difference in the LOD score thresholds used by the two programs and a difference in how the two programs analyze the data. For this study we imposed a LOD score threshold of 3 on all the results from the SOLAR analyses while QTLexpress calculates an empirical LOD score threshold. For this study, it was found that the empirical LOD score threshold for many of the analyses was lower than 3, suggesting that QTL with LOD scores less than 3 could be considered significant according to QTLexpress. In fact, 45 of the 55 QTL not identified by SOLAR had LOD scores less than 3 at the most significant position, suggesting that the difference in LOD score thresholds may be part of the explanation why there are QTL identified by QTLexpress but not by SOLAR. Another reason for the discrepancy in results between programs is the difference in analysis methods implemented by the two programs; QTLexpress performs a regression interval mapping analysis while SOLAR performs a variance component linkage analysis.

The differences in analysis methods may also explain the slight discrepancy in the location of the most significant position for a QTL controlling HUNTL on SSC17. Both QTLexpress and SOLAR identified HUNTL as significant on SSC17, but the two programs disagreed regarding the position of the QTL. QTLexpress identified the 95%CI as spanning from 15.5 to 18.5Mb with the most significant position being located at 15.5Mb. This is in contrast to the significant region identified by SOLAR that spanned from 22.7 to 23.0Mb with the most significant SNP being located at 22.7Mb. The two programs may be identifying the same locus controlling HUNTL, but since there is no overlap in the significant areas, further analysis is needed to resolve this question.

Significant Results Found by SOLAR Only

Of the 787 significant regions identified by SOLAR, 119 of them overlapped with a QTLexpress 95%CI as discussed above. There are several reasons to explain why the remaining 669 to not correspond with the QTL identified by QTLexpress, including the fact that the physical maps used in the SOLAR analyses cover more of the genome than do the linkage maps used by QTLexpress and as stated above, QTLexpress and SOLAR differ in the way they analyze data. When comparing the coverage of the QTLexpress linkage maps with the coverage of the SOLAR physical maps, several of the linkage maps fail to cover large areas at the telomeric ends of the chromosome. Some of the larger areas not covered include 25.3Mb at the top of SSC2, 21.0Mb at the top of SSC4, 41.9Mb at the bottom of SSC9, and 28.4Mb at the bottom of SSC16. Of the 669 significant regions identified by SOLAR but not by QTLexpress, 100 of these regions were located in areas of the genome that were not covered by the linkage maps used in the QTLexpress analyses. Since the linkage maps used in the QTLexpress analyses cover approximately 315Mb less of the genome than the physical maps used in the SOLAR analyses there is a definite possibility that the QTLexpress analyses failed to identify a number of QTL. One example of this is the

25.3Mb missing from the top of SSC2. This chromosomal segment that is of interest in this study since it contains the IGF2 gene, a gene known to affect both muscle mass and backfat thickness. As discussed earlier, the SOLAR and QTLexpress analyses both identified several significant QTL controlling backfat near the top of the p arm of SSC2 including; BFTLL, BFTAV, and BFTTR. Due to the linkage map not extending to the end of the p arm the other backfat traits, including BFTFR and BFTLR, are not identified as being significant on SSC2, when in fact they both show very significant signals at the telomeric end of the p arm of SSC2. This shows the importance of examining the significant regions identified by SOLAR that fall outside the boundaries of the QTLexpress linkage maps for additional chromosomal regions that may contain loci controlling the phenotypes investigated in this study.

Conclusion

To conclude, QTLexpress identified a total of 95 QTL controlling growth, carcass composition, and meat quality traits using a sparse map composed of microsatellite markers while SOLAR found 787 significant regions controlling the same phenotypes using a dense SNP map. Comparing the results from the two programs, it was found that 40 of the QTL identified by QTLexpress corresponded to 119 of the significant regions identified by SOLAR. These results will serve as the starting point for fine mapping efforts to be conducted in the future to determine the causative mutations behind the significant regions obtained in this study.

Tables

Table 6. Abbreviations of phenotypes. Collected on each F2 animal listed with the corresponding units when appropriate.

Phenotype	Abbreviation
Birth Weight, kg	BW
Nursery-in Weight, kg	NURSERY
Grower/Finisher-in Weight, kg	GROW
45 Minute pH	45PH
24 Hour pH	24PH
Average Glucose, Glycogen, Glucose-6-Phosphate, mmol/g of fresh tissue	GGG
Average Lactate, mmol/g of fresh tissue	LAC
Glycolytic Potential, mmol/g of fresh tissue	GP
Hunter L*	HUNTL
Hunter a*	HUNTA
Hunter b*	HUNTB
Shear Force, kg/cm^2	SHEAR
Fat %	FAT%
Cook Loss Percentage	COOK%
Moisture Percentage	MOIST%
Drip Loss Percentage	DRIP%
Marbling score	MARB
Juiciness score	JUICY
Tenderness score	TENDER
Off-Flavor score	OFFLAV
Color score	COLOR
Firmness score	FIRM
Carcass Length, cm	LENGTH
Loin Eye Area, cm^2	LEA
Dressing Percentage	DRESS%
Average Backfat, cm	BFTAV
1st Rib Backfat, cm	BFTFR
10th Rib Backfat, cm	BFTTR
Last Rib Backfat, cm	BFTLR
Last Lumbar Backfat, cm	BFTLL
Age at harvest, days	AGE
Birth Year Month	BYM
Grower/Finisher-in Age, days	GF_AGE
Hot carcass weight, kg	HCW
Litter Size	LITTER
Nursery-in Age, days	N_AGE
Parity of sow	PARITY
Sex	SEX

Table 7. Microsatellite marker linkage maps.

Chromosome	Locus ID	Linkage map cM position[1]	Sscrofa10 Mb position
SSC1	SW552	0.0	10.9
	SW1515	21.0	15.3
	SW1332	26.2	21.4
	SWR2300	39.2	23.2
	SW1417	52.2	48.0[2]
	ALPHA	62.5	59.3
	SW780	90.2	212.2
	SW1462	95.4	245.4
	SW974	104.3	261.7
	S0112	123.9	277.1[2]
SSC2	SW1201	0.0	25.3
	SW1686	6.5	31.8[2]
	SW1026	20.1	45.4
	SW1517	89.3	114.6
	SW1844	118.1	143.4
	SW2192	120.5	145.8
	SWR308	126.9	152.2[2]
	S0036	133.4	158.7
SSC3	SW2021	0.0	5.7
	SW2429	11.6	13.7
	SE47329	32.9	18.7[2]
	SW2527	37.7	21.8[2]
	SW2618	46.6	37.5
	SW271	68.2	84.4
	S0352	79.6	100.9[2]
	SW314	95.3	117.9
	S0002	109.0	125.3
	SW349	117.0	141.2
SSC4	SW2409	0.0	21.0
	SW752	9.1	36.0[2]
	SW839	22.5	67.9
	SW841	33.1	80.8[2]
	SW512	45.6	102.5
	SW58	73.1	121.3
	MP77	91.7	128.9
	SW856	102.4	135.5
SSC5	ACR	0.0	1.7[2]
	SW2425	62.7	20.4
	SW2003	76.4	54.5
	SW1468	89.1	65.9
	S0018	100.5	74.0
	SW1982	112.7	84.0
	SW1954	126.8	92.6
	SW378	129.8	96.8
	SW967	142.3	105.3
SSC6	S0035	0.0	2.1
	SW1841	39.1	15.6[2]
	SW492	62.7	33.4
	RYR1	75.7	42.1
	SW1129	77.5	51.4
	SW1647	106.6	124.6[2]
	S0121	113.7	132.9[2]
	SW322	142.3	145.5
	SW2419	150.2	153.8[2]

Table 7 (cont.)

Chromosome	Locus ID	Linkage map cM position[1]	Sscrofa10 Mb position
SSC7	S0383	0.0	0.7[2]
	SW1873	12.4	6.1
	SJ040	26.4	9.1
	TNF	57.8	27.7
	SW1701	73.0	48.2
	SW147	94.9	99.0[2]
	SW632	106.5	106.8
	SW2108	136.9	123.5
	SW1303	151.6	129.3
SSC8	S0098	0.0	6.3
	SW905	5.5	9.5
	SWR1101	24.6	20.0
	SW211	35.6	29.6
	SW29	47.5	40.7
	KS133	64.8	98.0
	SW2160	70.8	118.2[2]
	KS192	73.8	120.3
	OPN	105.2	138.5
	S0178	116.5	141.5[2]
SSC9	SW983	0.0	0.1[2]
	SW911	31.0	23.8
	SWR1848	42.9	35.1
	APOA1	57.0	50.1
	SW1677	68.1	65.5
	SW1	82.9	110.3
SSC10	SWR136	0.0	6.4
	SW767	19.4	11.1
	SWR1849	53.5	35.9
	SW2043	77.3	56.0
	SW1626	100.0	65.7[2]
SSC11	S0385	0.0	5.0[2]
	SW1460	13.0	11.9
	SW1632	25.0	18.0
	SW435	54.8	51.6
	SW1465	84.6	75.6
	SW2413	100.3	82.2
SSC12	SW2490	0.0	2.9
	S0143	9.4	4.7
	SW957	33.5	14.2[2]
	SE77921	42.2	20.0
	S0090	69.1	39.0
	SWC62	76.9	42.4
	SWC23	100.6	47.7
SSC13	S0282	0.0	3.3
	SWR1941	12.0	9.3
	SW1378	21.7	12.2
	SW344	46.7	26.7
	SW1495	68.5	89.7
	SW225	84.3	155.3
	SW520	91.8	186.6
	SW2440	112.2	201.1
	S0289	125.6	204.2
SSC14	SW857	0.0	7.3
	S0089	10.5	11.7
	SW1125	22.9	19.8
	SW2519	49.2	93.7
	S0007	58.8	121.1
	SW1081	68.0	134.9[2]
	SWC55	81.6	140.2
	SW2515	112.7	149.6

Table 7 (cont.)

Chromosome	Locus ID	Linkage map cM position[1]	Sscrofa10 Mb position
SSC15	KS169	0.0	2.6
	S0148	30.8	24.2
	SW1111	42.4	32.1
	SW964	51.8	39.6
	SW1989	57.2	54.7[2]
	SW2129	64.7	84.8
	PRKAG3	81.1	128.9
	SW936	85.0	134.3[3]
	SW2608	92.6	131.9[3]
	SWR2121	134.2	142.6[2]
SSC16	S0111	0.0	6.6[2]
	SW742	16.7	6.9
	SW419	32.5	22.8
	SW1809	48.5	40.5
	SW977	63.5	58.3
SSC17	SWR1120	0.0	15.5
	SW2142	30.3	21.5[2]
	SW1920	56.7	43.3
	SW2427	96.0	66.4[2]
SSC18	SW1808	0.0	4.7
	SW1984	31.7	27.7
	S0120	49.8	46.9
	S0177	70.6	52.7[2]

[1]Maps calculated by CRIMAP v2.4 with the corresponding Sscrofa10 Mb positions. [2]no BLAST [34] position was found so Mb positions for microsatellites at the same cM positions according to the MARC linkage map are reported. [3]Discrepancy between the order of the markers in the linkage maps compared to the physical map.

Table 8. Results from performing the POLYGENIC command in SOLAR. Each phenotype is listed with the corresponding heritability (H2r), significance of H2r (H2r p-value), and standard error of this p-value, the covariates identified as being significant ($p \leq 0.05$), and the proportion of phenotypic variance attributed to the covariates.

Trait	H2r	H2r p value	H2r Standard error	Significant covariates	Proportion of variance due to all significant covariates
BW	0.8345957	2.59E-23	0.1473708	LITTER	0.1432133
NURSERY	0.5891868	1.68E-16	0.1432405	N_AGE, PARITY	0.1825132
GROW	0.7360290	1.20E-25	0.1530637	GF_AGE	0.2765319
45PH	0.0737811	6.16E-03	0.0487083	SEX, AGE, HG	0.0426468
24PH	0.5150695	1.74E-19	0.1282431	-	-
GGG	0.6086354	9.34E-18	0.1346139	SEX, HG	0.0365168
LAC	0.4602476	3.26E-16	0.1299205	SEX, HG	0.0358052
GP	0.6765360	3.79E-21	0.1382231	-	-
HUNTL	0.4176461	4.75E-14	0.1210163	SEX, AGE	0.0255113
HUNTA	0.4361111	1.87E-05	0.1475761	SEX, AGE, HG	0.0348093
HUNTB	0.5214924	6.47E-15	0.1429990	HG	0.0404486
SHEAR	0.3352480	6.24E-09	0.1231459	SEX	0.0255372
FAT%	0.4125274	5.16E-15	0.1238185	SEX, AGE	0.0825679
COOK%	0.2599385	1.08E-09	0.1026565	-	-
MOIST%	0.3704329	2.15E-12	0.1234466	SEX, AGE, HG	0.0714494
DRIP%	0.4961957	7.52E-20	0.1314384	SEX, AGE, HG	0.0342743
MARB	0.0918357	7.34E-03	0.0738301	SEX, AGE	0.0814231
JUICY	0.1749887	5.70E-06	0.0859054	SEX	0.0129392
TENDER	0.2182282	1.25E-08	0.0880968	SEX	0.0383166
OFFLAV	0.0186900	2.47E-01	0.0356100	-	-
COLOR	0.4003674	4.81E-12	0.1287117	-	-
FIRM	0.4005362	4.40E-14	0.1272684	SEX	0.0108557
LENGTH	0.2731351	6.00E-07	0.1245745	SEX, AGE, HCW	0.3394119
LEA	0.4658328	1.58E-22	0.1263682	SEX, BYM, AGE, HCW	0.2052534
DRESS%	0.4451246	9.53E-14	0.1397523	BYM, AGE, HCW	0.1882819
BFTAV	0.3974506	1.02E-11	0.1372465	SEX, HCW	0.1351888
BFTFR	0.3478243	5.56E-10	0.1301565	SEX, HCW	0.0718423
BFTTR	0.3399769	1.00E-07	0.1385902	SEX, HCW	0.1251063
BFTLR	0.1282873	5.03E-04	0.0807845	SEX, HCW	0.1428518
BFTLL	0.5321409	3.50E-18	0.1403953	SEX, BYM, HCW	0.1171675

Table 9. Significant QTL identified in the microsatellite marker dataset by QTLexpress.

CHR	Trait	Mb[1]	F value[2]	LR[3]	LOD score[4]	Additive est.[5]	Add S.E.[6]	Dominance est.[7]	Dom S. E.[8]	95% CI[9]
0.001 genome-wise significant										
2	LEA	25.3	18.90	36.61	7.95	-1.645	0.280	0.794	0.429	25.3-39.3[10,11]
2	SHEAR	111.3	25.64	49.15	10.67	-0.225	0.033	-0.106	0.052	86.3-121.3[11]
2	TENDER	115.3	26.96	51.57	11.20	0.649	0.096	0.360	0.144	86.3-123.3[11]
7	LENGTH	39.7	16.54	32.18	6.99	-0.705	0.123	0.186	0.194	28.7-78.7[11]
0.01 genome-wise significant										
8	LENGTH	128.3	13.59	26.56	5.77	-0.626	0.123	0.274	0.212	121.3-135.3[11]
14	LENGTH	120.3	11.85	23.23	5.04	-0.526	0.111	0.157	0.149	103.3-126.3[11]
17	GP	42.5	13.71	26.80	5.82	7.368	1.464	-2.996	2.120	33.5-47.5[11]
0.05 genome-wise significant										
2	BFTLL	25.3	10.28	20.20	4.39	0.126	0.028	-0.014	0.043	25.3-32.3[10,11]
2	FAT%	91.3	10.90	21.39	4.64	0.350	0.087	-0.372	0.166	68.3-123.3[11]
2	JUICY	69.3	11.21	22.00	4.78	0.488	0.107	-0.195	0.207	39.3-99.3[11]
0.01 chromosome-wise significant										
1	DRESS%	59.9	7.35	14.51	3.15	0.256	0.070	-0.093	0.107	50.9-130.9[11]
2	BFTTR	25.3	9.17	18.05	3.92	0.126	0.031	-0.053	0.047	25.3-35.3[10,11]
3	GP	12.7	7.30	14.41	3.13	-3.473	0.148	6.737	2.240	6.7-16.7[11]
3	GGG	11.7	7.45	14.71	3.19	-1.323	0.576	2.669	0.876	5.7-15.7[10,11]
4	BW	109	6.86	13.56	2.95	-0.114	0.061	0.329	0.100	92-118
5	HUNTA	70.7	7.43	14.67	3.19	-0.111	0.077	0.449	0.124	62.7-80.7
6	DRESS%	37.1	7.14	14.10	3.06	-0.127	0.073	-0.410	0.122	23.1-73.1
7	HUNTA	56.7	7.55	14.90	3.24	0.379	0.098	-0.095	0.167	32.7-84.7[11]
12	LEA	2.9	7.21	14.25	3.09	0.930	0.273	0.640	0.409	2.9-12.9[10]
14	FIRM	136.3	7.32	14.47	3.14	-0.141	0.065	0.287	0.097	127.3-144.3
14	45PH	134.3	8.20	16.16	3.51	0.019	0.066	0.397	0.098	122.3-142.3
14	GGG	135.3	8.30	16.37	3.55	2.527	0.627	-0.312	0.936	129.3-143.3[11]
15	DRESS%	91.6	7.34	14.50	3.15	-0.314	0.087	-0.220	0.149	68.6-119.6
15	HUNTB	37.6	8.39	16.53	3.59	-0.320	0.082	-0.140	0.116	30.6-70.6
16	FAT%	57.6	6.20	12.26	2.66	0.263	0.075	0.078	0.131	40.6-57.6[10,11]
16	COOK%	34.6	6.36	12.59	2.73	-1.024	0.318	0.834	0.557	17.6-57.6[10]
16	LEA	57.6	6.39	12.64	2.75	-1.223	0.343	-0.056	0.600	33.6-57.6[10]
16	DRIP%	49.6	6.77	13.37	2.90	-0.387	0.173	0.862	0.306	27.6-57.6[10]
17	24PH	38.5	6.60	13.06	2.84	-0.198	0.063	0.188	0.099	24.5-51.5[11]
17	FIRM	32.5	6.66	13.17	2.86	-0.177	0.067	0.310	0.113	22.5-55.5[11]
17	DRIP%	37.5	7.12	14.06	3.05	0.404	0.158	-0.733	0.252	27.5-54.5
17	HUNTL	15.5	7.95	15.69	3.41	1.168	0.295	-0.370	0.434	15.5-18.5[10]
17	GGG	40.5	8.10	15.96	3.47	2.127	0.592	-1.688	0.902	29.5-48.5[11]
17	COLOR	33.5	8.62	16.99	3.69	-0.159	0.062	0.357	0.104	24.5-42.5
18	GGG	47.7	6.98	13.79	3.00	2.074	0.562	0.532	0.893	38.7-52.7[10,11]
18	FAT%	50.7	7.89	15.56	3.38	0.230	0.069	0.215	0.109	46.7-52.7[10,11]
0.05 chromosome-wise significant										
1	SHEAR	169.9	6.03	11.95	2.59	0.213	0.071	-0.371	0.224	119.9-267.9
2	MARB	64.3	5.43	10.77	2.34	0.135	0.079	-0.384	0.144	25.3-115.3[10,11]
2	HUNTB	87.3	5.46	10.81	2.35	0.252	0.108	-0.484	0.214	52.3-137.3
2	DRIP%	145.3	5.56	11.01	2.39	0.464	0.159	-0.384	0.220	125.3-151.3
2	BFTAV	25.3	5.70	11.29	2.45	0.190	0.057	-0.057	0.088	25.3-36.3[10,11]
2	COLOR	145.3	6.18	12.24	2.66	-0.188	0.060	0.146	0.083	115.3-149.3[11]
2	HUNTL	146.3	6.40	12.67	2.75	0.841	0.301	-1.003	0.429	140.3-151.3
3	LAC	18.7	5.77	11.42	2.48	-1.273	0.627	2.479	0.902	14.7-37.7
3	COOK%	16.7	6.97	13.77	2.99	-0.414	0.278	1.519	0.440	8.7-20.7
3	COLOR	94.7	7.05	13.92	3.02	-0.039	0.054	0.326	0.088	84.7-104.7
4	LENGTH	53	5.35	10.60	2.30	-0.387	0.119	-0.008	0.202	21-79[10,11]
4	NURSERY	117	5.47	10.84	2.35	-0.236	0.076	0.197	0.120	46-124
4	DRESS%	30	6.79	13.42	2.92	0.275	0.075	-0.013	0.117	21-67[10]
5	OFFLAV	14.7	5.51	10.92	2.37	-0.018	0.070	-0.407	0.124	8.7-43.7
5	LENGTH	63.7	6.52	12.90	2.80	0.347	0.110	-0.340	0.182	15.7-70.7[11]
6	GROW	140.1	5.75	11.38	2.47	-1.014	0.339	0.868	0.567	132.1-151.1[11]
6	LEA	87.1	5.78	11.45	2.49	-1.543	0.454	-0.251	1.027	42.1-140.1

Table 9 (cont.)

7	24PH	66.7	5.48	10.86	2.36	0.212	0.081	0.287	0.147	21.7-100.7
7	LEA	27.7	5.69	11.27	2.45	0.857	0.299	-0.766	0.412	19.7-103.7[11]
7	HUNTL	60.7	5.93	11.73	2.55	-1.286	0.374	0.091	0.656	24.7-95.7
7	DRESS%	40.7	6.08	12.04	2.61	0.290	0.083	-0.008	0.132	22.7-105.7[11]
8	45PH	19.3	5.86	11.60	2.52	-0.198	0.062	0.134	0.089	11.3-28.3
8	BFTLL	120.3	5.88	11.64	2.53	0.099	0.029	-0.010	0.040	87.3-132.3[11]
9	GGG	110.1	4.91	9.74	2.12	-1.542	0.599	-1.712	0.858	95.1-110.1
9	BFTFR	47.1	5.26	10.43	2.27	-0.081	0.042	-0.185	0.069	35.1-59.1
9	BFTTR	49.1	5.69	11.27	2.45	-0.041	0.033	-0.168	0.053	36.1-64.1
9	BFTAV	49.1	5.71	11.30	2.46	-0.058	0.061	-0.323	0.099	38.1-62.1
10	LENGTH	24.4	6.45	12.76	2.77	0.304	0.160	-0.653	0.238	6.4-40.4[10]
11	JUICY	38	5.27	10.44	2.27	-0.221	0.101	-0.468	0.195	18-57
11	MOIST%	6	5.57	11.03	2.40	0.169	0.052	0.073	0.078	5-18[10]
11	FAT%	8	5.67	11.24	2.44	-0.208	0.064	-0.119	0.099	5-16[10,11]
11	BFTLR	10	5.95	11.79	2.56	-0.115	0.059	-0.277	0.091	5-17[10]
12	MOIST%	41.9	4.91	9.73	2.11	0.138	0.056	0.138	0.083	14.9-46.9[11]
12	BFTLR	46.9	5.06	10.04	2.18	-0.122	0.065	-0.252	0.104	41.9-46.9[10]
12	COLOR	5.9	5.20	10.31	2.24	0.115	0.055	0.206	0.082	2.9-12.9
12	BFTLL	46.9	5.93	11.74	2.55	-0.033	0.031	-0.159	0.050	41.9-46.9[10]
12	BFTAV	46.9	6.15	12.17	2.64	-0.105	0.063	-0.298	0.101	41.9-46.9[10]
12	24PH	4.9	6.55	12.96	2.81	0.148	0.057	0.210	0.082	2.9-10.9
13	DRESS%	118.3	6.17	12.21	2.65	-0.032	0.119	-0.974	0.277	102.3-132.3
14	DRESS%	132.3	5.34	10.59	2.30	0.190	0.080	0.284	0.120	13.3-146.3
14	HUNTL	136.3	5.57	11.04	2.40	0.556	0.319	-1.271	0.477	122.3-147.3
14	SHEAR	144.3	5.71	11.31	2.46	-0.125	0.039	-0.093	0.061	129.3-149.3[10]
14	GP	135.3	6.37	12.61	2.74	5.301	1.664	-3.290	2.452	128.3-143.3[11]
15	BFTLL	131.6	5.12	10.14	2.20	-0.090	0.028	-0.005	0.041	104.6-134.6[11]
15	BFTFR	128.6	5.38	10.66	2.32	-0.124	0.038	0.012	0.057	86.6-134.6[11]
15	GGG	131.6	5.59	11.08	2.41	-1.768	0.552	-0.790	0.813	56.5-134.6[11]
15	HUNTA	38.6	5.64	11.17	2.43	-0.269	0.081	0.062	0.113	9.6-83.6
15	BFTTR	131.6	5.86	11.60	2.52	-0.104	0.031	0.023	0.045	110.6-136.6
15	COOK%	134.6	6.23	12.32	2.68	0.074	0.253	-1.327	0.378	104.6-137.6
15	BFTAV	131.6	6.30	12.47	2.71	-0.201	0.057	-0.037	0.083	106.6-134.6[11]
15	DRIP%	32.6	6.49	12.83	2.79	-0.293	0.144	-0.593	0.200	25.6-44.6
17	LAC	42.5	5.07	10.05	2.18	1.900	0.625	-0.873	0.909	27.5-52.5[11]
17	BW	28.5	5.63	11.15	2.42	0.135	0.061	0.229	0.102	18.5-40.5
17	HUNTB	15.5	6.04	11.95	2.59	0.255	0.083	0.158	0.120	15.5-18.5[10]
18	LENGTH	27.7	4.73	9.38	2.04	-0.122	0.103	-0.420	0.147	14.7-36.7
18	MOIST%	51.7	5.74	11.36	2.47	-0.136	0.056	-0.186	0.083	47.7-52.7[10]
18	BFTLL	35.7	5.81	11.51	2.50	0.097	0.031	0.068	0.051	11.7-49.7
18	24PH	46.7	5.98	11.83	2.57	-0.174	0.055	-0.121	0.085	37.7-51.7
18	GP	46.7	6.00	11.88	2.58	4.751	1.402	1.710	2.187	37.7-51.7

[1]Chromosomal position in Mb, [2]calculated F-values, [3]likelihood ratio, [4]LOD score, [5]additive estimates (negative additive estimates indicate the Duroc allele increased the trait value, while positive estimates indicate the Berkshire allele increase the trait value), [6]standard errors for the additive estimate, [7]dominance estimates, [8]standard error for the dominance estimate, and the [9]95%CI. [10]95% confidence intervals that are not the true size due to being located close to the end of the linkage map. [11]QTL that correspond to a significant region as identified by SOLAR.

Fine Mapping of a Quantitative Trait Locus Controlling Loin Eye Area on *Sus scrofa* Chromosome 6

Abstract

A previous study used an outbred F2 interval mapping analysis to scan pig chromosomes 2, 6, 13, and 18 for quantitative trait loci controlling carcass composition traits and identified a highly significant quantitative trait locus controlling loin eye area on chromosome 6 within the Illinois Meat Quality Pedigree. The quantitative trait locus is located between microsatellite markers SW1129 and SW1647, a span of 73.2 Mb, and has a 95% confidence interval spanning 98 Mb. To reduce this interval to a more manageable size for candidate gene selection, 1600 single nucleotide polymorphism markers from the Illumina® PorcineSNP60 BeadChip [65] that are located within the 95% confidence interval were analyzed using various software programs, including QTLexpress, Qxpak, R/qtl, SOLAR, PLINK, Bayes-C. Results from each program were consolidated and a consensus region was identified between 85.67 and 85.68Mb.

Introduction

The Illinois Meat Quality Pedigree (IMQP) is a three generation Berkshire x Duroc population created to identify quantitative trait loci controlling carcass composition, growth, and meat quality traits across the swine genome. Previously, an outbred interval mapping analysis performed on chromosomes 2 (SSC2), 6 (SSC6), 13 (SSC13), and 18 (SSC18) discovered 14 significant QTL controlling carcass composition, growth and meat quality traits in the IMQP population [129]. Of these 14 QTL, two were located on SSC6, with the more significant QTL controlling loin eye area (LEA). It is this LEA QTL that will be the focus of this study.

Loin eye area is a measurement of the loin muscle (*Longissimus dorsi*) at the 10th rib. The loin is one of the most valuable cuts of a pig carcass, behind the ribs and belly, and makes up approximately 12% of the pig's carcass weight. Meat packers pay a premium for loins that are within a certain size range, too large or to small are discounted, so narrowing down the region to find the causative mutation would aid in marker assisted selection programs aiming to control muscle size and provide insight to the underlying mechanisms of muscle growth. The LEA QTL on SSC6 is located between markers SW1129 and SW1647, a 73.2Mb interval, and is significant at the 5% genome-wise level. The additive effect of this QTL suggests that the Duroc allele is increasing the size of the loin eye area, this being the breed we expected to carry favorable alleles relating to carcass composition traits.

To date, there have been five other scans of the porcine genome that have identified a significant LEA QTL on SSC6 [168, 178, 182, 196-197]. Of these, two of the studies [178, 196] had QTL intervals that overlapped completely with the interval identified by Stearns at al. (2005) [129] while the other three studies had QTL intervals that overlapped with the 95%CI of the QTL identified by Stearns at al. (2005) [168, 182, 197]. Similar to the IMQP population, both Wimmers et al. (2006) and Edwards et al. (2008) used Durocs as one of the breeds in their crosses, but unlike our study they found the Duroc allele decreased the size of the loin eye area [178, 182]. Though all five studies identified a LEA QTL on SSC6, only one has made an effort to narrow the QTL interval and propose candidate genes [168, 198].

The aim of this study is to analyze additional SNP markers, within the 95%CI of the LEA QTL on SSC6 found by Stearns et al. 2005 [129], using various QTL analysis programs including QTLexpress [189], QxPak [199], R/qtl [200], SOLAR [192], PLINK [66], and Bayes-C [67]

to narrow the QTL interval and identify a small consensus region to be examined further for candidate genes.

Materials and Methods

IMQP Reference Population

The Illinois Meat Quality Pedigree (IMQP) is an F2 Berkshire x Duroc intercross population consisting of 664 individuals. Semen from three purebred Berkshire boars, chosen based on progeny test performance values, was used to inseminate 17 purebred Duroc sows chosen from the breeding program at the University of Illinois Moorman Swine Research Unit. Each boar was bred to an average of six different sows, but each sow was only mated to one of the three boars. All founder animals were genotyped for the halothane and Rendement Napole (R200Q) mutations and were found to be normal. An F1 population was produced, from which five boars and 44 sows were chosen to be mated. No mating of relatives occurred within the F1 intercross. Each of the F1 boars were mated to an average of 11 different F1 sows, while each F1 sow was mated to either one or two different F1 boars. The intercross between F1 individuals created an F2 generation consisting of 595 pigs.

SNP Genotyping

Fifteen microliters of genomic DNA at a concentration of 100ng/µl was sent to GeneSeek® to SNP genotype the IMQP population using the Illumina® PorcineSNP60 BeadChip [65]. The final list of SNPs used for this analysis was created using PLINK v1.05 [66] by removing Mendelian errors, removing SNPs with a minor allele frequency less than 0.01, and removing SNPs with a missing genotype frequency of 0.9 or greater. The SNPs located within the 95%CI of the LEA QTL on SSC6 were then analyzed for QTL with the following programs.

QTLexpress

A linkage analysis using regression interval mapping was performed on the 1600 SNPs within the 95%CI of the LEA QTL on SSC6 using QTLexpress [189]. Instead of a linkage map, a physical map of the SNPs with corresponding Mb positions was used. Using the F2 population option of QTLexpress, a one QTL model analysis using the Haley-Knott regression [201] was run on the LEA phenotype with F-values, additive effects, and dominance deviations calculated every 1.0Mb. For this analysis, sex and birth year month were fitted as fixed effects, while age and hot carcass weight were fitted as covariates. A permutation test, using 5,000 permutations, was performed to calculate the 0.05 and 0.01 chromosome-wise empirical p-value thresholds.

QxPAK

QxPak [199] was used to perform an association analysis on the 1600 SNPs within the 95%CI of the LEA QTL on SSC6. The program did not allow for true spacing of markers due to the SNPs being so close together, so for the purposes of this analysis, each SNP was spaced 0.1Mb apart. For this analysis, sex and birth year month were fitted as fixed effects, while age and hot carcass weight were fitted as covariates. Pedigree information was able to be uploaded to the program, but was not taken into account for this analysis. Using the SNP_AD command, QxPak calculated the likelihood of the effect of each SNP on the trait in addition to the additive effects and dominance deviations for each SNP.

R/qtl

A linkage analysis using regression interval mapping was performed on the 1600 SNPs within the 95%CI of the LEA QTL on SSC6 using R/qtl [200]. Instead of a linkage map, a physical map of the SNPs with corresponding Mb positions was used. The F2 pedigree structure was able to be accounted for by coding the genotypes accordingly, but this coding

was unable to conserve the true relationships between each animal in the pedigree. For this analysis; sex, birth year month, age and hot carcass weight were fitted as covariates. Using the SCANONE command, a one QTL model analysis using the Haley-Knott regression [201] was run on the LEA phenotype with LOD scores, additive estimates, and dominance estimates calculated every 1.0Mb.

SOLAR

The Sequential Oligogenic Linkage Analysis Routine software package (SOLAR) [192] was used to identify the QTL region controlling loin eye area on SSC6 by performing a variance component twopoint linkage analysis. Allele frequencies were estimated using pedigree based maximum likelihood estimates (MLE). These estimates were computed using the FREQ MLE command in SOLAR, a command that runs an external program Allfreq, an extension of the MENDEL program. The IBD command in SOLAR was used to compute marker-specific identity-by-descent (IBD) matrices using a Monte Carlo algorithm. The POLYGENIC command in SOLAR calculates the additive polygenic heritability (H2r) of each phenotype, the significance of the H2r, creates a model to be used during the twopoint linkage analysis, tests each covariate for significance in the model, and calculates the proportion of variance contributed by each covariate. Covariates were removed from the model if they did not reach a statistical significance value p=0.05. The POLYGENIC command also calculates the residual kurtosis for each phenotype and determines if a LOD score adjustment is needed. LOD score adjustment values were calculated using the LODADJ command and applied to the twopoint linkage results for each analysis regardless of the recommendation by the POLYGENIC command. A twopoint linkage analysis was performed using the TWOPOINT command in SOLAR. For each marker, the TWOPOINT command computes a LOD score, the amount of genotypic variance controlled by the marker, and the residual genetic variance.

PLINK

PLINK [66] was used to perform an association analysis using a linear regression of LEA on the 1600 SNP genotypes within the 95%CI of the LEA on SSC6. For this analysis; sex, birth year month, age and hot carcass weight were fitted as covariates. Using the LINEAR command, PLINK performs a test that includes each of the covariates, an additive effect, and a dominance deviation in the model, this was followed by 10,000 permutations, performed using the MPERM command in PLINK, to calculate empirical p-values for each SNP. Pedigree was not accounted for in this analysis.

BAYES-C

An association analysis was performed using the Bayes-C option of the Gensel program [67]. Genetic and residual variance were estimated using the Bayes-C method with a Markov chain Monte Carlo (MCMC) consisting of a burn-in period of 1,000 cycles followed by 9,000 cycles. The Bayes-CPi method was then used to estimate a value for Pi. Bayes-C was then used to estimate all 1600 SNP marker effects simultaneously using estimated genetic variance, residual variance, Pi values with a Markov chain Monte Carlo (MCMC) consisting of a burn-in period of 1,000 cycles followed by 9,000 cycles. Pedigree was not accounted for in this analysis.

Results

The 95%CI for the LEA QTL on SSC6 spans the region between 42.1 and 140.1Mb and contains a total of 1600 SNPs from the Illumina® PorcineSNP60 BeadChip [65] according to the Sscrofa10 assembly. For this study we imposed a LOD score threshold of 3.0 on the results from the programs that reported results as LOD scores including; QxPak, R/qtl, and SOLAR. While QTLexpress reports results as a LOD score, a permutation test is performed within the program and an empirical LOD score threshold is calculated, this empirical

threshold was used for this study. PLINK reports results as empirical p-values, and a p-value of 0.99 was used as a significance threshold for this study. Bayes-C reports results as SNP effects and an effect size of 0.01 and greater was considered significant for this study.

QTLexpress

Permutations test resulted in a 0.05 chromosome-wise LOD score threshold of 2.785 and a 0.01 chromosome-wise LOD score threshold of 3.635. The results from the QTL analysis, showed the most significant position is located at 78.05 Mb with a LOD score of 5.18 and an additive effect of -1.226 (Figure 10) (Appendix D: SSC6 LEA fine mapping results). Though this is the most significant position, there were three other segments with LOD scores that exceed both the 0.05 and 0.01 chromosome-wise significance thresholds. For this analysis, a significant region is defined as two or more positions with LOD scores that exceed the 0.05 chromosome-wise threshold. The first significant region spans from 47.17 to 63.10 Mb with the most significant position within this region located at 61.15 Mb and has a LOD score of 4.205. The second significant region spans from 75.08 to 87.08 Mb, this being the significant region that contains the most significant position identified in this analysis as reported above. The last significant region spans from 92.15 to 103.27 Mb, with the most significant position located at 91.10 Mb and has a LOD score of 3.442.

QxPAK

A total of 113 SNPs had LOD scores above 3.0 according to the association analysis performed by QxPak (Figure 11) (Appendix D: SSC6 LEA fine mapping results). The most significant SNP is located at 88.28 Mb, has a LOD score of 6.48, and has an additive effect of -1.8912. For this analysis, a significant region is defined as at least two SNPs with a LOD score over 3.0 that have a correlation coefficient (r^2) of 0.2 or greater. One hundred and two (the other 11 SNPs did not cluster with any other SNP) of the significant SNPs clustered

into 21 significant regions spanning from 65.67 to 131.10 Mb. The first group of regions fell between 60.0 and 70.0 Mb; 65.67 to 65.82 Mb, 66.73 to 66.79 Mb, and 68.0253 to 68.0255 Mb. The next group of regions were located between 70.0 and 80.0 Mb; 74.22 to 75.65 Mb, 76.47 to 78.63 Mb, 78.69 to 79.03 Mb, and 79.51 to 80.07 Mb. The group containing the most significant SNP from this analysis spanned from 80.0 to 90.0 Mb; 80.19 to 80.31 Mb, 84.01 to 84.16 Mb, 85.67 to 86.26 Mb, 86.91 to 87.30 Mb, and 88.28 to 88.84 Mb (containing the most significant SNP). The next group contained four significant regions and spanned from 100.0 to 110.0 Mb; 101.42 to 101.49 Mb, 103.00 to 103.98 Mb, 104.56 to 104.66 Mb, and 109.96 to 110.38 Mb. The next four regions grouped together and spanned from 114.91 to 116.31 Mb, 116.92 to 117.03 Mb, and 118.26 to 118.31 Mb. Finally, the last two regions spanned from 122.50 to 122.79 Mb and from 130.38 to 131.10 Mb.

R/qtl

A total of 214 SNPs had LOD scores over 3.0 according to the regression interval mapping analysis performed by R/qtl (Figure 12) (Appendix D: SSC6 LEA fine mapping results). The most significant LOD score identified in this analysis was 7.14 and was shared by four adjacent SNPs located between 85.67 and 85.68 Mb. For this analysis, a significant region is defined as at least two SNPs with a LOD score over 3.0 that have a correlation coefficient (r^2) of 0.2 or greater. One hundred and ninety-six (the other 18 SNPs did not cluster with any other SNP) SNPs clustered into 32 significant regions spanning from 57.64 to 140.04 Mb. The first group of significant regions spanned from 50.0 to 60.0 Mb; 57.64 to 58.00 Mb, 58.64 to 58.85 Mb, and 59.03 to 59.76 Mb. The next two significant regions were located between 60.0 and 70.0 Mb and spanned from 64.80 to 64.88 Mb and 68.01 to 68.09 Mb. The third group of significant regions spanned from 70.0 to 80.0 Mb and contained 5 significant regions; 72.48 to 72.65 Mb, 74.22 to 75.65 Mb, 76.47 to 78.63 Mb, 78.69 to 78.71 Mb, and 79.51 to 79.94 Mb. The group containing the most significant SNPs from this analysis spanned from 80.0 to 90.0 Mb and contained four significant regions; 80.81 to

80.83 Mb, 85.43 to 85.80 Mb (containing the most significant SNPs), 86.91 to 87.30 Mb, and 88.28 to 88.35 Mb. Another group of significant regions spanned from 100.0 to 110.0 Mb and contained five significant regions; from 100.99 to 101.15 Mb, from 101.16 to 101.18 Mb, from 101.24 to 102.23 Mb, from 103.00 to 103.98 Mb, from 109.96 to 109.97 Mb. The next group of significant regions spanned from 110.0 to 120.0 Mb and contained eight significant regions; 111.53 to 112.35 Mb, 114.87 to 114.88 Mb, 115.22 to 115.83 Mb, 116.92 to 117.00 Mb, 117.49 to 117.52 Mb, 118.31 to 118.33 Mb, 118.45 to 118.48 Mb, 119.33 to 119.70 Mb. And the final group of significant regions spanned from 130.0 to 140.0 Mb and containing six significant regions; 130.38 to 131.92 Mb, 136.68 to 137.22 Mb, 137.66 to 137.69, 137.99 to 138.15 Mb, 139.99 to 140.00 Mb, and 140.03 to 140.04 Mb.

SOLAR

Using a LOD score threshold of 3.0, there were no SNPs identified as being significant when analyzed by using the variance component linkage analysis performed by SOLAR (Figure 13) (Appendix D: SSC6 LEA fine mapping results). Though there were no SNPs with LOD scores above 3.0, there were 26 SNPs with LOD scores between 2.0 and 2.9, and these will be used to make a comparison to the other analysis programs. The SNP with the highest LOD score is located at 80.07 and has a LOD score of 2.99. For this analysis, a region used for comparison with the other programs is defined as at least two SNPs with a LOD score over 2.0 that have a correlation coefficient (r^2) of 0.2 or greater. The four regions for comparison span from 79.51 to 79.54 Mb, 80.00 to 82.61 Mb, 94.96 to 95.68 Mb, and 96.89 to 97.55 Mb.

PLINK

A total of 28 SNPs had an empirical p-value of less than 0.01 according to the association analysis performed by PLINK (Figure 14) (Appendix D: SSC6 LEA fine mapping results).

The most significant p-value identified in this analysis was 0.0001 and was shared by four adjacent SNPs located between 85.67 and 85.68 Mb. For this analysis, a significant region is defined as at least two SNPs with a p-value less than 0.01 that have a correlation coefficient (r^2) of 0.2 or greater. Twenty-two (the other six SNPs did not cluster with any other SNP) SNPs clustered into 12 significant regions spanning from 63.79 to 131.53Mb. The first group of regions spanned from 60.0 to 70.0 Mb; 63.79 to 63.85 Mb, 64.87 to 64.88 Mb, and 65.67 to 65.82 Mb. The next two regions were located between 70.0 and 80.0 Mb; 74.89 to 75.14 Mb and 78.95 to 79.03 Mb. The group of regions containing the most significant SNPs from this analysis spanned from 80.0 to 90.0 and included two significant regions; 80.19 to 80.31 Mb and 85.67 to 85.68 Mb (containing the most significant SNPs). The last two regions spanned from 115.22 to 115.70 Mb and 130.38 to 131.53 Mb.

BAYES-C

A total of 40 SNPs had an effect of 0.01 or greater according to the association analysis performed by Bayes-C (Figure 15) (Appendix D: SSC6 LEA fine mapping results). The SNP with the highest effect in this analysis was located at 138.09 Mb and had an effect of 0.020570. For this analysis, a significant region is defined as at least two SNPs with an effect of 0.01 or greater that have a correlation coefficient (r^2) of 0.2 or greater. Thirty-one (the other none SNPs did not cluster with any other SNP) SNPs clustered into eight significant regions spanning from 76.52 to 137.12 Mb. The first two regions were located between 70.0 and 80.0Mb; 76.52 to 78.6 Mb and 78.97 to 79.03 Mb. The group of regions containing the second most significant SNP from this analysis spanned from 80.0 to 90.0 Mb and included two significant regions; 84.01 to 84.16 Mb and 85.67 to 85.68 Mb (containing the SNP with the second highest effect). The next significant region spanned from 111.55 to 112.3 Mb. And finally the last group or significant regions spanned from 130.0 to 140.0 Mb; 130.38 to 131.53 Mb, 132.50 to 132.81 Mb, and 136.86 to 137.12 Mb.

Discussion

A LEA QTL on SSC6 has been previously identified by several different studies aiming to examine the pig genome for QTL controlling carcass composition phenotypes, with the exception of the study by Ovilo et al. (2002), no other fine mapping efforts for this QTL have been published [129, 168, 178, 182, 196-197]. This study further examined the QTL identified by Stearns et al. (2005) by analyzing 1600 SNP markers to narrow the LEA QTL interval and propose candidate genes.

Comparison of Analysis Methods

This study used six different analysis programs to examine the 1600 SNPs located within the 95%CI of the LEA QTL on SSC6. The use of multiple programs is due to the current lack of a program that will appropriately perform a fine mapping QTL analysis on the LEA phenotype in the IMQP population. To properly analyze a QTL for a trait measured in the IMQP population, several important pieces of information need to be accounted for including the F2 family structure of the population, any covariates or fixed effects influencing the phenotype, and the order and distance between markers used in the analysis. Each of the six programs used was unable to account for at least one of these pieces of information within their analyses (Table 10).

Stearns et al. (2005) performed the initial QTL analysis on the LEA phenotype within the IMQP population using QTLexpress, this analysis program performs a regression interval mapping analysis using the Haley-Knott method [201] and accounts for each of the pieces of information needed for a QTL analysis as mentioned above. Fine mapping of the 95%CI of the LEA QTL with QTLexpress was able to account for the F2 family structure, the covariates and fixed effects inflecting loin eye area, and the order of the markers. Since the markers used for the analysis were spaced so close together, centiMorgan (cM) distances

between the markers were unable to be calculated, so a physical map of the markers was used and Megabase distances were used instead of cM distances. This resulted in some markers not being included in the analysis and several warning messages alerting us of questionable recombination probabilities. This was not surprising since the average distance between any two SNPs in this dataset was 0.06Mb, and QTLexpress tries to estimate the probability of linkage with the QTL every 1.0Mb.

Of the other programs used in this study, R/qtl is the only other program that utilizes the same method of analysis as QTLexpress. Though the analysis method was similar, R/qtl was unable to account for the same pieces of information as QTLexpress. R/qtl accounts for family structure by coding the genotypes of the individuals according to the type of population the animals are from, but was unable to account for the true relationships between individuals in the IMQP population. As for the covariates and fixed effects influencing loin eye area, R/qtl treats them all as covariates, depending on how these parameters are fit into the model, the two fixed effects, sex and birth year month, may not be used appropriately. And similar to QTLexpress, since the SNP markers were spaced so close together, a physical map of the markers was used and Megabase distances were used instead of cM distances. Unlike QTLexpress, no warning messages were produced when using such small distances between markers, but since the markers were spaced less than 1Mb apart, R/qtl was only able to calculate the probability of linkage with the QTL for each marker.

The SOLAR software has the capability of performing both a twopoint and a multipoint variance component linkage analysis, but due to time and computer resource constraints only the twopoint analysis was performed for this study. SOLAR is able to account for the F2 family structure by calculating an identity-by-descent (IBD) value for each allele between each individual in the population. Similar to R/qtl, SOLAR fits all of the covariates and fixed

effects influencing loin eye area as covariates. Since we performed the twopoint instead of the multipoint linkage analysis, the order and distance between the SNP markers was not taken into account for this analysis. After the analysis was completed, the SNP markers were ordered and LOD score results for individual markers were plotted according to physical positions along the chromosome.

The QxPak software has the capability of accounting for family structure when performing a QTL analysis, but the amount of data and close spacing of the SNP markers used in this study prohibitied us from using this option. Instead a quantitative trait association analysis was performed using QxPak, and while relationship information could be entered into the program, this information was ignored during our analysis. Fixed effects and covariates were able to be entered as separate variables into QxPak, with the assumption that these values will be treated appropriately by the program. Association analyses examine one SNP marker at a time, so order and distance between markers was not preserved by the program, but rather after the analysis was finished, the SNPs were plotted in order and spaced according to their physical positions.

The PLINK software was used to also perform a quantitative association analysis and similar to the QxPak analysis, family structure was unable to be accounted for by the PLINK software. Use of the --linear option in PLINK allowed for the inclusion of covariates and fixed effects into the model, but no distinction is made between fixed effects and covariates so the fixed effects may not be fit in the model appropriately. Again, since association analyses examine markers individually, the order and distance between the markers was not taken into account during the analysis, but the SNPs were plotted in order and spaced according to the physical positions after analysis.

The Bayes-C analysis performs an association analysis, and similar to the QxPak and PLINK association analyses, no family structure information was able to be accounted for by this analysis. As for fixed effects and covariates, these factors influencing loin eye area were treated as separate variable types within this analysis. Unlike the QxPak and PLINk association analyses, Bayes-C fits all the SNP markers in the model at the same time and considers other markers when calculating the effect a marker has on the phenotype, thus the order of the markers is important for this analysis. Though the order of the markers is considered, the distance between the markers is not taken into account, so the SNP markers were plotted according to their physical positions after completion of the analysis.

Until an appropriate analysis method is developed to fine map QTL using densely spaced SNP data, this study has used the combined results from six different analysis programs to identify regions containing potential candidate genes controlling LEA on SSC6.

Consensus Region

Despite the differences in analysis methods and information used by the programs used in this study, there was one small region identified as significant by all five programs that had significant results. This consensus region contains four SNPs and spans from 85.67 and 85.68Mb. PLINK and R/qtl identified these SNPs as the four most significant SNPs in their analyses and in fact the results showed that all four SNPs had the same p-values or LOD scores, respectively. QxPak only identified one SNP more significant than these SNPs and though the LOD scores for the four SNPs were not identical, they are very similar. QTLexpress does not report individual LOD scores for each of these SNPs, but all four were located within the significant region containing the most significant position. As for the Bayes-C analysis, these four SNPs were not the most significant, with the most significant of the four being the third most significant SNP in the analysis, and though they did not have identical effect values, the values are very similar.

A BLAST similarity search [34] against the human genome was performed for each of the four SNPs to search for candidate genes that could be controlling loin eye area size. The BLAST [34] results showed these four SNPs were located between 37.56 and 37.58Mb on human chromosome 1 (HSA1). Though no genes were located within these coordinates, there were several potential candidate genes located within a Megabase of either side of this region, including zinc finger CCCH-type domain containing protein 12A (ZC3H12A), SMAD nuclear interacting protein 1 (SNIP1), guanine nucleotide binding protein-like 2 (GNL2) and four and a half LIM domains 3 (FHL3). Zhou et al. (2006) showed that ZC3H12A was involved in the cell death of cardiomyocytes, though cardiomyocytes are different than skeletal myocytes, it is possible that ZC2H12A plays a similar role in skeletal myocytes and mutations within the gene could change the number of myocytes present in the muscle, thus changing its size [202]. Kim et al. (2000) overexpressed SNIP1 leading to the inhibition of NODAL, a member of the TGF-β gene family that is essential in the formation of the mesoderm, the embryonic tissue that gives rise to skeletal muscle, so there is a possibility that mutations in SNIP1 could affect the formation of skeletal muscle during development resulting in changes in size of adult muscle [203]. As for the other two possible candidate genes, GNL2 and FHL3, both show very high expression in skeletal muscle compared to other tissues, but little else is known about the genes so further investigation into these two genes is needed [204-205].

Other Regions

Since loin eye area is a quantitative trait and since the 95%CI for this QTL spans such a large portion of the chromosome, it would not be unusual to consider the possibility that there may be more than loci controlling loin eye area on SSC6. Though the consensus region discussed above is the one region that all five programs identified as significant, there were two other regions that were found to be highly significant in four of the five

programs. The first region spans from 76.52 to 78.63Mb and contains a total of eight significant SNPs. These SNPs were all found to be significant according to the QTLexpress, QxPak, R/qtl and Bayes-C analyses. The BLAST [34] results show this region is located between 27.26 and 29.36Mb on HSA1, a region containing two potential candidate genes including, mitogen-activated protein kinase kinase kinase 6 (MAP3K6) and erythrocyte membrane protein band 4.1 (elliptocytosis 1, RH-linked) (EPB41). Little is known about these two genes with the exception that expression levels of EPB41 is highest in skeletal muscle and MAP3K6 has been shown to be only expressed in skeletal muscles [192-193]. Just upstream of the region containing the eight significant SNPs, located at 26.37 and 26.73Mb on HSA1 respectively, are two strong candidate genes, tripartite motif containing 63 (TRIM63) and lin-28 homolog A (C. elegans) (LIN28A). TRIM63, a gene thought to regulate protein degradation in muscle, has been found by Northern blot to be expressed exclusively in muscle tissue [206-207]. In addition, Witt et al. (2008) saw an increase in the size of skeletal muscle fibers in TRIM63 knockout mice and found that these knockout mice were resistant to muscle atrophy, a finding substantiating results from Bodine et al. (2001) [207-208]. These findings suggest that TRIM63 would be a good candidate gene for a QTL controlling the size of a muscle, in that a mutation in the gene could cause an increase in the size of the muscle by increasing the size of the individual muscle fibers. LIN28A, shown to be essential for differentiation of skeletal muscle cells, is expressed in both embryonic and adult muscle tissue [209]. In addition, Polesskaya et al. (2007) showed, through both loss and gain-of-function assays, that LIN28A is essential for differentiation of myoblasts and that LIN28A is a post-transcriptional regulator of insulin-like growth factor 2 (IGF2), a gene known to be important for the growth and development of skeletal muscle (Polesskaya, Cuvellier et al. 2007). This evidence suggests that mutations within LIN28A could affect differentiation, differentiation and growth of skeletal muscle as early as during embryonic development and this could lead to a change in the amount of muscle tissue accumulated in an animal.

The second region spans from 130.38 to 131.53 and contains three significant SNPs. The SNPs in this region were all found to be significant according to the analyses performed by QxPak, R/qtl, PLINK and Bayes-C. The BLAST [34] results show the region is located between 70.59 and 71.79Mb on HSA1, a region containing one potential candidate gene, the cystathionine gamma-lyase (CTH) gene. CTH is responsible for transforming cystathionine into cysteine and Ishii et al. (2010) found that CTH-null mice that were fed a diet of low in cysteine developed acute skeletal muscular atrophy, suggesting that this gene could play a role in muscle mass [210].

Conclusion

Comparing the results from the analyses performed using QTLexpress, Qxpak, R/qtl, SOLAR, PLINK, and Bayes-C, showed numerous differences as to which of the 1600 SNPs within the 95%CI of the LEA QTL on SSC6 were found to be significant; but despite differences in analysis methods, the results also showed three consensus regions that contain several potential candidate genes. These candidate genes include ZC3H12A, SNIP1, TRIM63, and LIN28A; and will be the focus of future studies concentrating on finding causative polymorphisms controlling the size of the loin eye area on SSC6 in pigs.

Figures

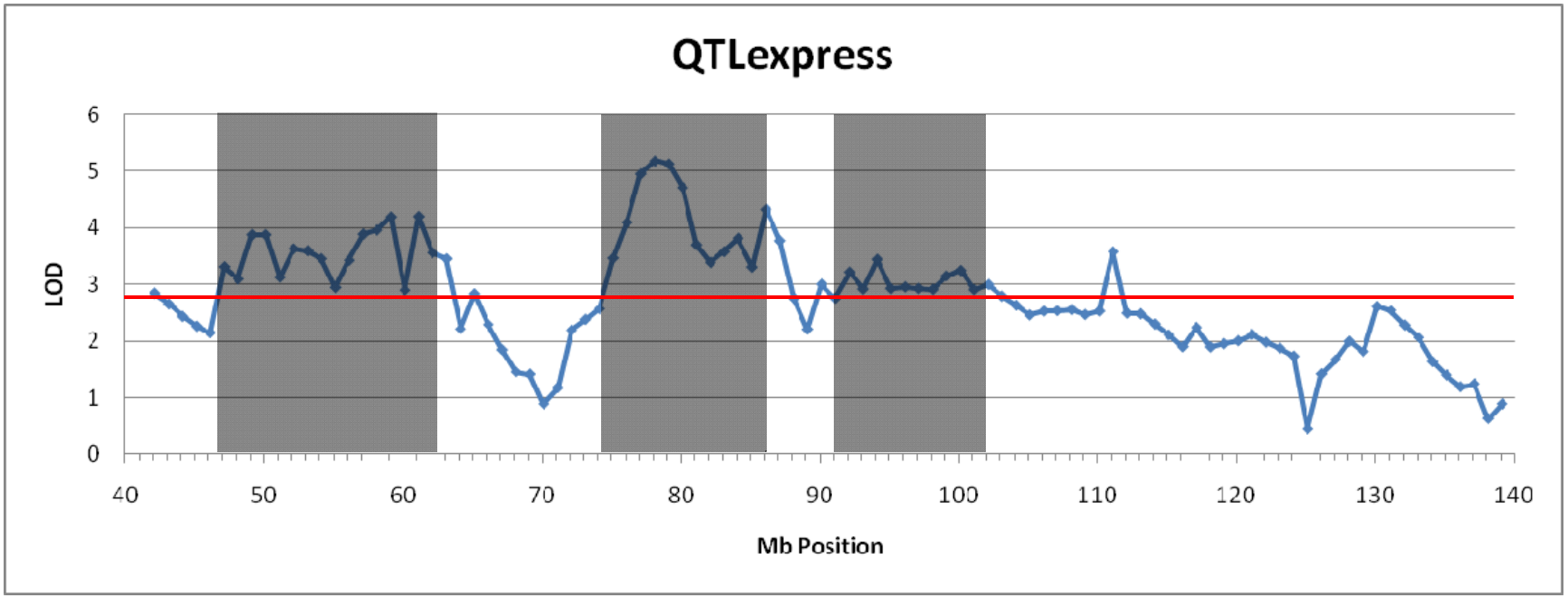

Figure 10. Plot of results from the QTLexpress SNP analysis of the 95%CI of the LEA QTL on SSC6. LOD scores plotted against Mb position. The horizontal red line indicates the significance threshold. The vertical grey bars highlight significant regions.

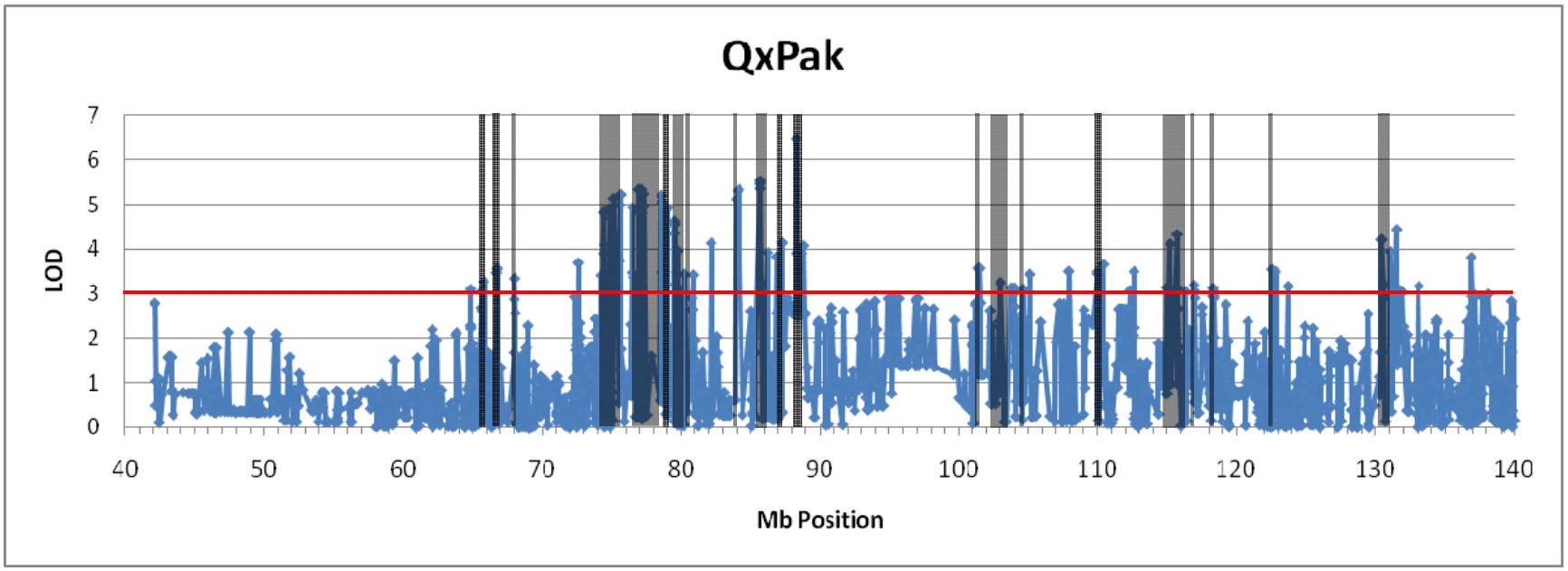

Figure 11. Plot of results from the QxPak SNP analysis of the 95%CI of the LEA QTL on SSC6. LOD scores plotted against Mb position. The horizontal red line indicates the significance threshold. The vertical grey bars highlight significant regions.

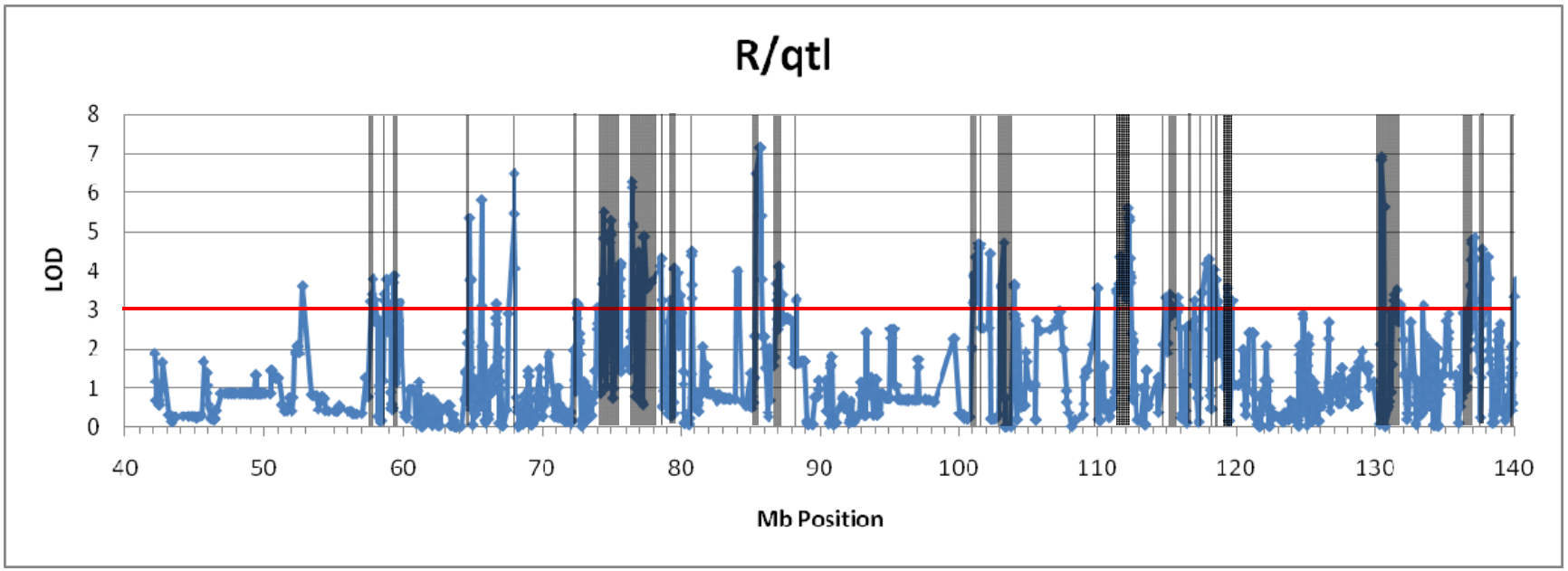

Figure 12. Plot of results from the R/qtl SNP analysis of the 95%CI of the LEA QTL on SSC6. LOD scores plotted against Mb position. The horizontal red line indicates the significance threshold. The vertical grey bars highlight significant regions.

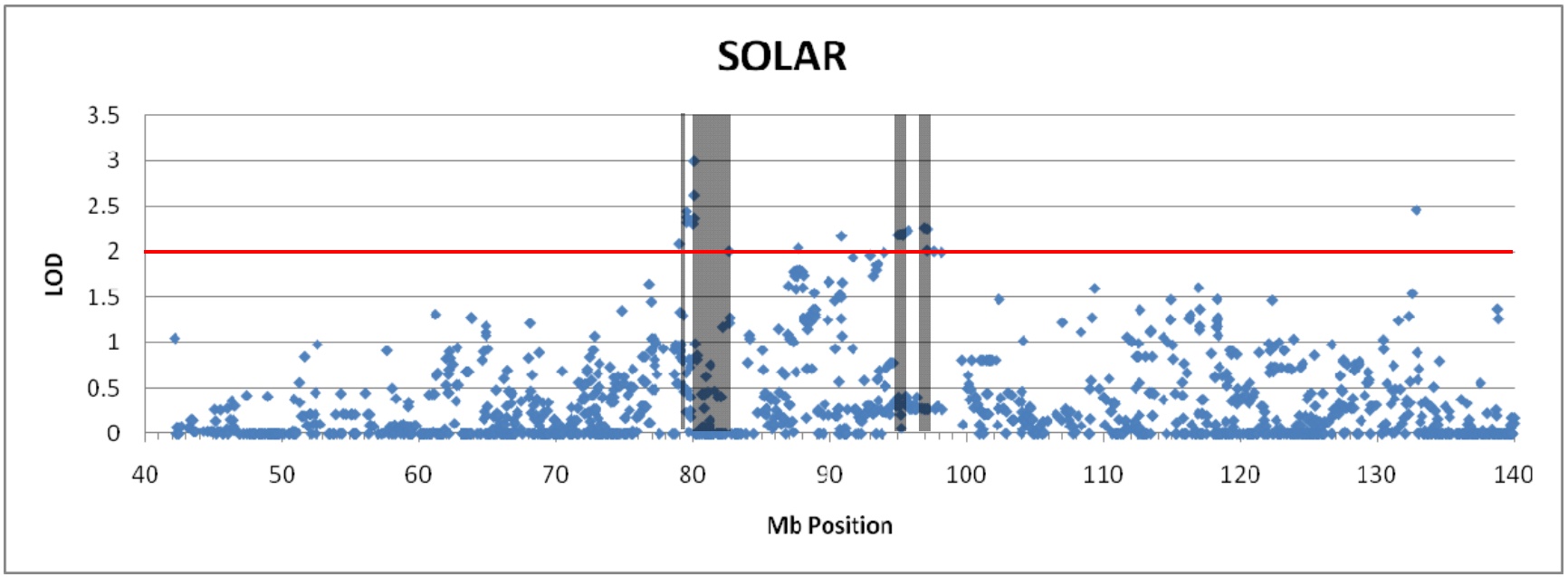

Figure 13. Plot of results from the SOLAR SNP analysis of the 95%CI of the LEA QTL on SSC6. LOD scores plotted against Mb position. The horizontal red line indicates the significance threshold. The vertical grey bars highlight significant regions.

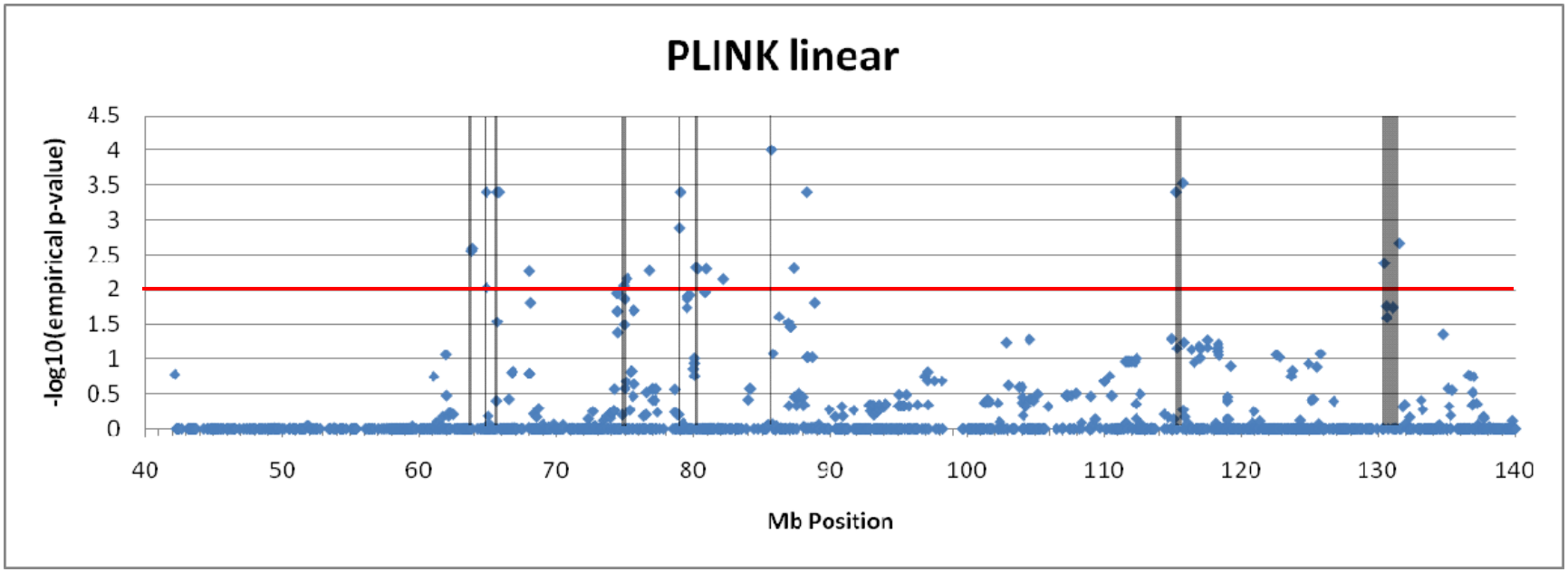

Figure 14. Plot of results from the PLINK SNP analysis of the 95%CI of the LEA QTL on SSC6. –log10 of the empirical p-value plotted against Mb position. The horizontal red line indicates the significance threshold. The vertical grey bars highlight significant regions.

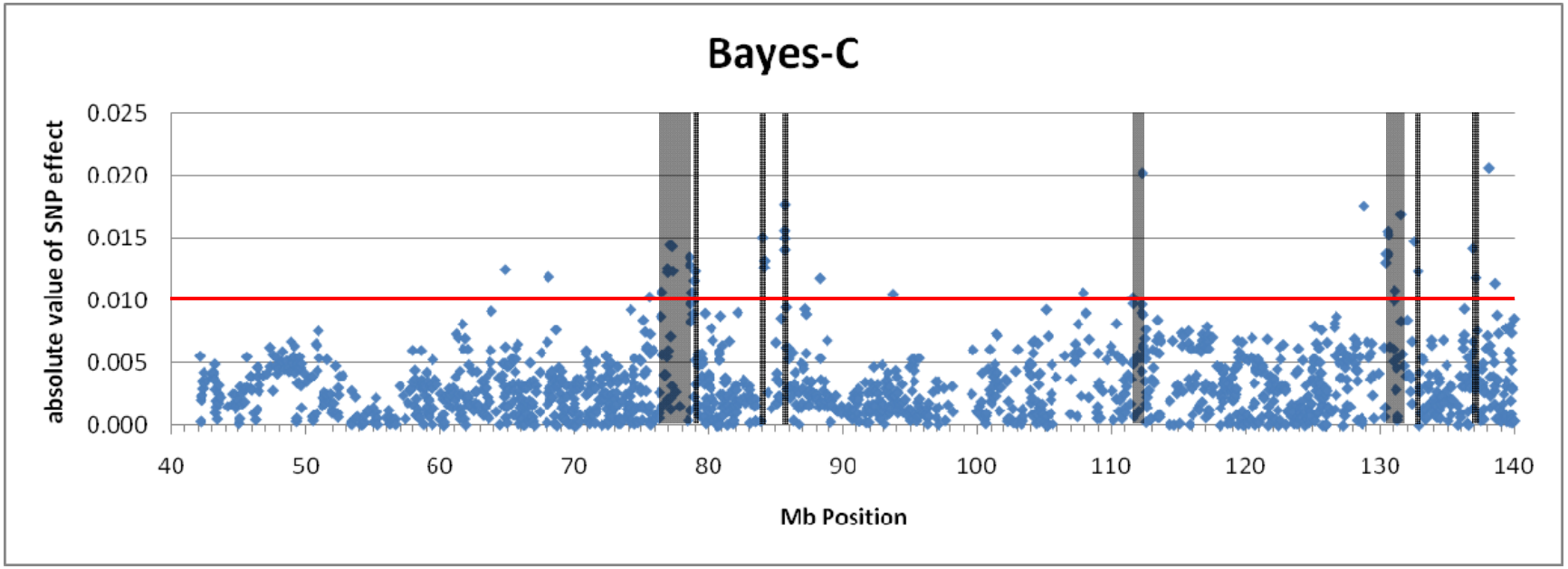

Figure 15. Plot of results from the Bayes-C SNP analysis of the 95%CI of the LEA QTL on SSC6. Absolute values of the effect of each SNP plotted against Mb position. The horizontal red line indicates the significance threshold. The vertical grey bars highlight significant regions.

Tables

Table 10. Summary of analysis program differences.

Program	Pedigree[1]	Map[1]	Covariates[1]	Analysis method
QTLexpress	YES	YES[3]	YES[4]	linkage analysis using regression interval mapping
QxPak	NO	NO	YES[4]	association analysis
R/qtl	YES[2]	YES[3]	YES[5]	linkage analysis using regression interval mapping
SOLAR	YES	NO	YES[5]	twopoint linkage analysis using variance components
PLINK	NO	NO	YES[5]	association analysis
Bayes-C	NO	NO	YES[4]	association analysis

[1]"YES"/"NO" indicates if this information was/was not used in each analysis program. [2]The F2 population was accounted for, but actual relationships between animals were not. [3]A physical map was used in place of the linkage map. [4]Fixed effects and covariates treated as separate variable types in the model. [5]Fixed effects and covariates treated the same in the model.

References

1. **U.S. & World Population Clocks** [http://www.census.gov/main/www/popclock.html]
2. Butler L: **The Linkage Map of the Tomato**. *The Journal of Heredity* 1952, **43**:25-35.
3. Sax K: **The Association of Size Differences with Seed-Coat Pattern and Pigmentation in PHASEOLUS VULGARIS**. *Genetics* 1923, **8**:552-560.
4. Breese EL, Mather K: **The organisation of polygenic activity within a chromosome in Drosophila**. *Heredity* 1957, **11**:373-395.
5. Botstein D, White RL, Skolnick M, Davis RW: **Construction of a genetic linkage map in man using restriction fragment length polymorphisms**. *Am J Hum Genet* 1980, **32**:314-331.
6. Thomas CM, Vos P, Zabeau M, Jones DA, Norcott KA, Chadwick BP, Jones JD: **Identification of amplified restriction fragment polymorphism (AFLP) markers tightly linked to the tomato Cf-9 gene for resistance to Cladosporium fulvum**. *Plant J* 1995, **8**:785-794.
7. Ellegren H, Chowdhary BP, Johansson M, Marklund L, Fredholm M, Gustavsson I, Andersson L: **A primary linkage map of the porcine genome reveals a low rate of genetic recombination**. *Genetics* 1994, **137**:1089-1100.
8. Sawera M, Cirera S, Jorgensen CB, Gorodkin J, Fredholm M: **Linkage mapping of gene-associated SNPs to pig chromosome 11**. *Anim Genet* 2006, **37**:199-204.
9. Gusella JF, Wexler NS, Conneally PM, Naylor SL, Anderson MA, Tanzi RE, Watkins PC, Ottina K, Wallace MR, Sakaguchi AY *et al*: **A polymorphic DNA marker genetically linked to Huntington's disease**. *Nature* 1983, **306**:234-238.
10. Tsui LC, Buchwald M, Barker D, Braman JC, Knowlton R, Schumm JW, Eiberg H, Mohr J, Kennedy D, Plavsic N *et al*: **Cystic fibrosis locus defined by a genetically linked polymorphic DNA marker**. *Science* 1985, **230**:1054-1057.
11. Knowlton RG, Cohen-Haguenauer O, Van Cong N, Frezal J, Brown VA, Barker D, Braman JC, Schumm JW, Tsui LC, Buchwald M *et al*: **A polymorphic DNA marker linked to cystic fibrosis is located on chromosome 7**. *Nature* 1985, **318**:380-382.
12. Davies KE, Pearson PL, Harper PS, Murray JM, O'Brien T, Sarfarazi M, Williamson R: **Linkage analysis of two cloned DNA sequences flanking the Duchenne muscular dystrophy locus on the short arm of the human X chromosome**. *Nucleic Acids Res* 1983, **11**:2303-2312.
13. Tanksley SD: **Mapping polygenes**. *Annu Rev Genet* 1993, **27**:205-233.
14. Lander ES, Botstein D: **Strategies for studying heterogeneous genetic traits in humans by using a linkage map of restriction fragment length polymorphisms**. *Proc Natl Acad Sci U S A* 1986, **83**:7353-7357.
15. Lander ES, Botstein D: **Mapping complex genetic traits in humans: new methods using a complete RFLP linkage map**. *Cold Spring Harb Symp Quant Biol* 1986, **51 Pt 1**:49-62.
16. Lander ES, Botstein D: **Mapping mendelian factors underlying quantitative traits using RFLP linkage maps**. *Genetics* 1989, **121**:185-199.
17. Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES: **Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers**. *Genetics* 1992, **132**:823-839.
18. Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD: **Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms**. *Nature* 1988, **335**:721-726.
19. **PigQTLdb** [http://www.animalgenome.org/cgi-bin/QTLdb/SS/index]

20. Andersson L, Haley CS, Ellegren H, Knott SA, Johansson M, Andersson K, Anderssoneklund L, Edforslilja I, Fredholm M, Hansson I *et al*: **Genetic-Mapping of Quantitative Trait Loci for Growth and Fatness in Pigs**. *Science* 1994, **263**:1771-1774.
21. Amaral AJ, Megens HJ, Crooijmans RP, Heuven HC, Groenen MA: **Linkage disequilibrium decay and haplotype block structure in the pig**. *Genetics* 2008, **179**:569-579.
22. Milan D, Jeon JT, Looft C, Amarger V, Robic A, Thelander M, Rogel-Gaillard C, Paul S, Iannuccelli N, Rask L *et al*: **A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle**. *Science* 2000, **288**:1248-1251.
23. Grisart B, Coppieters W, Farnir F, Karim L, Ford C, Berzi P, Cambisano N, Mni M, Reid S, Simon P *et al*: **Positional candidate cloning of a QTL in dairy cattle: identification of a missense mutation in the bovine DGAT1 gene with major effect on milk yield and composition**. *Genome Res* 2002, **12**:222-231.
24. Cohen-Zinder M, Seroussi E, Larkin DM, Loor JJ, Everts-van der Wind A, Lee JH, Drackley JK, Band MR, Hernandez AG, Shani M *et al*: **Identification of a missense mutation in the bovine ABCG2 gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle**. *Genome Res* 2005, **15**:936-944.
25. Olsen HG, Nilsen H, Hayes B, Berg PR, Svendsen M, Lien S, Meuwissen T: **Genetic support for a quantitative trait nucleotide in the ABCG2 gene affecting milk composition of dairy cattle**. *BMC Genet* 2007, **8**:32.
26. Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M *et al*: **A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig**. *Nature* 2003, **425**:832-836.
27. Hedrick HB, Leavitt RK, Alexander MA: **Variation in porcine muscle quality of Duroc and Hampshire barrows**. *J Anim Sci* 1968, **27**:48-52.
28. Sayre RN, Briskey EJ, Hoekstra WG: **Comparison of Muscle Characteristics and Post-Mortem Glycolysis in three Breeds of Swine**. *J Anim Sci* 1963, **22**:1012-1020.
29. Kukoyi EA, Addis PB, McGrath CJ, Rempel WE, Martin FB: **Porcine stress syndrome and postmortem muscle characteristics of two purebreds and three specific terminal crosses**. *J Anim Sci* 1981, **52**:278-284.
30. Monin G, Sellier P: **Pork of Low Technological Quality with a Normal Rate of Muscle Ph Fall in the Immediate Post-Mortem Period - the Case of the Hampshire Breed**. *Meat Science* 1985, **13**:49-63.
31. Le Roy P, Naveau J, Elsen JM, Sellier P: **Evidence for a new major gene influencing meat quality in pigs**. *Genet Res* 1990, **55**:33-40.
32. Mariani P, Lundstrom K, Gustafsson U, Enfalt AC, Juneja RK, Andersson L: **A major locus (RN) affecting muscle glycogen content is located on pig chromosome 15**. *Mamm Genome* 1996, **7**:52-54.
33. Milan D, Woloszyn N, Yerle M, Le Roy P, Bonnet M, Riquet J, Lahbib-Mansais Y, Caritez JC, Robic A, Sellier P *et al*: **Accurate mapping of the "acid meat" RN gene on genetic and physical maps of pig chromosome 15**. *Mamm Genome* 1996, **7**:47-51.
34. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs**. *Nucleic Acids Res* 1997, **25**:3389-3402.
35. Bergeron R, Russell RR, 3rd, Young LH, Ren JM, Marcucci M, Lee A, Shulman GI: **Effect of AMPK activation on muscle glucose metabolism in conscious rats**. *Am J Physiol* 1999, **276**:E938-944.
36. Malek M, Dekkers JCM, Lee HK, Baas TJ, Prusa K, Huff-Lonergan E, Rothschild MF: **A molecular genome scan analysis to identify chromosomal regions**

influencing economic traits in the pig. II. Meat and muscle composition. *Mammalian Genome* 2001, **12**:637-645.
37. Ciobanu D, Bastiaansen J, Malek M, Helm J, Woollard J, Plastow G, Rothschild M: **Evidence for new alleles in the protein kinase adenosine monophosphate-activated gamma(3)-subunit gene associated with low glycogen content in pig skeletal muscle and improved meat quality**. *Genetics* 2001, **159**:1151-1162.
38. Lindahl G, Enfalt AC, Seth Gv, Joseli A, Hedebro-Velander I, Andersen HJ, Braunschweig M, Andersson L, Lundstrom K: **A second mutant allele (V199I) at the PRKAG3 (RN) locus-II. Effect on colour characteristics of pork loin**. *Meat Science* 2004, **66**:621-627.
39. Lindahl G, Enfalt AC, von Seth G, Josell A, Hedebro-Velander I, Andersen HJ, Braunschweig M, Andersson L, Lundstrom K: **A second mutant allele (V199I) at the PRKAG3 (RN) locus- I. Effect on technological meat quality of pork loin**. *Meat Science* 2004, **66**:609-619.
40. Barnes BR, Marklund S, Steiler TL, Walter M, Hjalm G, Amarger V, Mahlapuu M, Leng Y, Johansson C, Galuska D *et al*: **The 5'-AMP-activated protein kinase gamma3 isoform has a key role in carbohydrate and lipid metabolism in glycolytic skeletal muscle**. *J Biol Chem* 2004, **279**:38441-38447.
41. Coppieters W, Riquet J, Arranz JJ, Berzi P, Cambisano N, Grisart B, Karim L, Marcq F, Moreau L, Nezer C *et al*: **A QTL with major effect on milk yield and composition maps to bovine chromosome 14**. *Mamm Genome* 1998, **9**:540-544.
42. Heyen DW, Weller JI, Ron M, Band M, Beever JE, Feldmesser E, Da Y, Wiggans GR, VanRaden PM, Lewin HA: **A genome scan for QTL influencing milk production and health traits in dairy cattle**. *Physiol Genomics* 1999, **1**:165-175.
43. Riquet J, Coppieters W, Cambisano N, Arranz JJ, Berzi P, Davis SK, Grisart B, Farnir F, Karim L, Mni M *et al*: **Fine-mapping of quantitative trait loci by identity by descent in outbred populations: application to milk production in dairy cattle**. *Proc Natl Acad Sci U S A* 1999, **96**:9252-9257.
44. Farnir F, Grisart B, Coppieters W, Riquet J, Berzi P, Cambisano N, Karim L, Mni M, Moisio S, Simon P *et al*: **Simultaneous mining of linkage and linkage disequilibrium to fine map quantitative trait loci in outbred half-sib pedigrees: revisiting the location of a quantitative trait locus with major effect on milk production on bovine chromosome 14**. *Genetics* 2002, **161**:275-287.
45. Cases S, Smith SJ, Zheng YW, Myers HM, Lear SR, Sande E, Novak S, Collins C, Welch CB, Lusis AJ *et al*: **Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis**. *Proc Natl Acad Sci U S A* 1998, **95**:13018-13023.
46. Smith SJ, Cases S, Jensen DR, Chen HC, Sande E, Tow B, Sanan DA, Raber J, Eckel RH, Farese RV: **Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat**. *Nat Genet* 2000, **25**:87-90.
47. Winter A, Kramer W, Werner FA, Kollers S, Kata S, Durstewitz G, Buitkamp J, Womack JE, Thaller G, Fries R: **Association of a lysine-232/alanine polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat content**. *Proc Natl Acad Sci U S A* 2002, **99**:9300-9305.
48. Grisart B, Farnir F, Karim L, Cambisano N, Kim JJ, Kvasz A, Mni M, Simon P, Frere JM, Coppieters W *et al*: **Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition**. *Proc Natl Acad Sci U S A* 2004, **101**:2398-2403.
49. Kuhn C, Thaller G, Winter A, Bininda-Emonds OR, Kaupe B, Erhardt G, Bennewitz J, Schwerin M, Fries R: **Evidence for multiple alleles at the DGAT1 locus better**

explains a quantitative trait locus with major effect on milk fat content in cattle. *Genetics* 2004, **167**:1873-1881.
50. Georges M, Nielsen D, Mackinnon M, Mishra A, Okimoto R, Pasquino AT, Sargeant LS, Sorensen A, Steele MR, Zhao X *et al*: **Mapping quantitative trait loci controlling milk production in dairy cattle by exploiting progeny testing**. *Genetics* 1995, **139**:907-920.
51. Spelman RJ, Coppieters W, Karim L, van Arendonk JA, Bovenhuis H: **Quantitative trait loci analysis for five milk production traits on chromosome six in the Dutch Holstein-Friesian population**. *Genetics* 1996, **144**:1799-1808.
52. Zhang Q, Boichard D, Hoeschele I, Ernst C, Eggen A, Murkve B, Pfister-Genskow M, Witte LA, Grignola FE, Uimari P *et al*: **Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred pedigree**. *Genetics* 1998, **149**:1959-1973.
53. Velmala RJ, Vilkki HJ, Elo KT, de Koning DJ, Maki-Tanila AV: **A search for quantitative trait loci for milk production traits on chromosome 6 in Finnish Ayrshire cattle**. *Anim Genet* 1999, **30**:136-143.
54. Kuhn C, Freyer G, Weikard R, Goldammer T, Schwerin M: **Detection of QTL for milk production traits in cattle by application of a specifically developed marker map of BTA6**. *Anim Genet* 1999, **30**:333-340.
55. Nadesalingam J, Plante Y, Gibson JP: **Detection of QTL for milk production on Chromosomes 1 and 6 of Holstein cattle**. *Mamm Genome* 2001, **12**:27-31.
56. Ron M, Kliger D, Feldmesser E, Seroussi E, Ezra E, Weller JI: **Multiple quantitative trait locus analysis of bovine chromosome 6 in the Israeli Holstein population by a daughter design**. *Genetics* 2001, **159**:727-735.
57. Olsen HG, Gomez-Raya L, Vage DI, Olsaker I, Klungland H, Svendsen M, Adnoy T, Sabry A, Klemetsdal G, Schulman N *et al*: **A genome scan for quantitative trait loci affecting milk production in Norwegian dairy cattle**. *J Dairy Sci* 2002, **85**:3124-3130.
58. Band MR, Larson JH, Rebeiz M, Green CA, Heyen DW, Donovan J, Windish R, Steining C, Mahyuddin P, Womack JE *et al*: **An ordered comparative map of the cattle and human genomes**. *Genome Res* 2000, **10**:1359-1368.
59. Cohen M, Reichenstein M, Everts-van der Wind A, Heon-Lee J, Shani M, Lewin HA, Weller JI, Ron M, Seroussi E: **Cloning and characterization of FAM13A1--a gene near a milk protein QTL on BTA6: evidence for population-wide linkage disequilibrium in Israeli Holsteins**. *Genomics* 2004, **84**:374-383.
60. Schnabel RD, Kim JJ, Ashwell MS, Sonstegard TS, Van Tassell CP, Connor EE, Taylor JF: **Fine-mapping milk production quantitative trait loci on BTA6: analysis of the bovine osteopontin gene**. *Proc Natl Acad Sci U S A* 2005, **102**:6896-6901.
61. Andersson-Eklund L, Marklund L, Lundstrom K, Haley CS, Andersson K, Hansson I, Moller M, Andersson L: **Mapping quantitative trait loci for carcass and meat quality traits in a wild boar x large white intercross**. *Journal of Animal Science* 1998, **76**:694-700.
62. Jeon JT, Carlborg O, Tornsten A, Giuffra E, Amarger V, Chardon P, Andersson-Eklund L, Andersson K, Hansson I, Lundstrom K *et al*: **A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus**. *Nat Genet* 1999, **21**:157-158.
63. Nezer C, Moreau L, Brouwers B, Coppieters W, Detilleux J, Hanset R, Karim L, Kvasz A, Leroy P, Georges M: **An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs**. *Nature Genetics* 1999, **21**:155-156.
64. Nezer C, Collette C, Moreau L, Brouwers B, Kim JJ, Giuffra E, Buys N, Andersson L, Georges M: **Haplotype sharing refines the location of an imprinted quantitative trait locus with major effect on muscle mass to a 250-kb**

chromosome segment containing the porcine IGF2 gene. *Genetics* 2003, **165**:277-285.
65. Ramos AM, Crooijmans RP, Affara NA, Amaral AJ, Archibald AL, Beever JE, Bendixen C, Churcher C, Clark R, Dehais P *et al*: **Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology**. *PLoS One* 2009, **4**:e6524.
66. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al*: **PLINK: a tool set for whole-genome association and population-based linkage analyses**. *Am J Hum Genet* 2007, **81**:559-575.
67. **GenSel—User manual for a portfolio of genomic selection related analyses** [http://taurus.ansci.iastate.edu/gensel]
68. Cho I-C, Zhong T, Seo B-Y, Jung E-J, Yoo C-K, Kim J-H, Lee J-B, Lim H-T, Kim B-W, Lee J-H *et al*: **Whole-genome association study for the roan coat color in an intercrossed pig population between Landrace and Korean native pig**. *Genes & Genomics* 2011, **33**:17-23.
69. Ren J, Mao H, Zhang Z, Xiao S, Ding N, Huang L: **A 6-bp deletion in the TYRP1 gene causes the brown colouration phenotype in Chinese indigenous pigs**. *Heredity* advance online publication, 27 October 2010 (doi: 10.1038/hdy.2010.129).
70. Gorbach D, Mote B, Totir L, Fernando R, Rothschild M: **Polydactyl Inheritance in the Pig**. *Journal of Heredity* 2010, **101**:469-475.
71. Sironen A, Uimari P, Nagy S, Paku S, Andersson M, Vilkki J: **Knobbed acrosome defect is associated with a region containing the genes STK17b and HECW2 on porcine chromosome 15**. *BMC Genomics* 2010, **11**:699.
72. Duijvesteijn N, Knol EF, Merks JW, Crooijmans RP, Groenen MA, Bovenhuis H, Harlizius B: **A genome-wide association study on androstenone levels in pigs reveals a cluster of candidate genes on chromosome 6**. *BMC Genet* 2010, **11**:42.
73. Fan B, Onteru SK, Du ZQ, Garrick DJ, Stalder KJ, Rothschild MF: **Genome-wide association study identifies Loci for body composition and structural soundness traits in pigs**. *PLoS One* 2011, **6**:e14726.
74. Onteru SK, Fan B, Nikkila MT, Garrick DJ, Stalder KJ, Rothschild MF: **Whole-genome association analyses for lifetime reproductive traits in the pig**. *J Anim Sci* 2011, **89**:988-995.
75. Hardenbol P, Yu F, Belmont J, Mackenzie J, Bruckner C, Brundage T, Boudreau A, Chow S, Eberle J, Erbilgin A *et al*: **Highly multiplexed molecular inversion probe genotyping: over 10,000 targeted SNPs genotyped in a single tube assay**. *Genome Res* 2005, **15**:269-275.
76. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TP, Sonstegard TS *et al*: **Development and characterization of a high density SNP genotyping assay for cattle**. *PLoS One* 2009, **4**:e5350.
77. Kim Y, Ryu J, Woo J, Kim JB, Kim CY, Lee C: **Genome-wide association study reveals five nucleotide sequence variants for carcass traits in beef cattle**. *Animal Genetics*, 6 February 2011 (DOI: 10.1111/j.1365-2052.2010.02156.x).
78. Schulman NF, Sahana G, Lund MS, Viitala SM, Vilkki JH: **Quantitative trait loci for fertility traits in Finnish Ayrshire cattle**. *Genet Sel Evol* 2008, **40**:195-214.
79. Barendse W, Reverter A, Bunch RJ, Harrison BE, Barris W, Thomas MB: **A validated whole-genome association study of efficient food conversion in cattle**. *Genetics* 2007, **176**:1893-1905.
80. Bolormaa S, Hayes BJ, Savin K, Hawken R, Barendse W, Arthur PF, Herd RM, Goddard ME: **Genome-wide association studies for feedlot and growth traits in cattle**. *J Anim Sci*, 14 January 2011 (doi:10.2527/jas.2010-3079).

81. Jiang L, Liu J, Sun D, Ma P, Ding X, Yu Y, Zhang Q: **Genome wide association studies for milk production traits in Chinese Holstein population**. *PLoS One* 2010, **5**:e13661.
82. Mai MD, Sahana G, Christiansen FB, Guldbrandtsen B: **A genome-wide association study for milk production traits in Danish Jersey cattle using a 50K single nucleotide polymorphism chip**. *J Anim Sci* 2010, **88**:3522-3528.
83. Pant SD, Schenkel FS, Verschoor CP, You Q, Kelton DF, Moore SS, Karrow NA: **A principal component regression based genome wide analysis approach reveals the presence of a novel QTL on BTA7 for MAP resistance in holstein cattle**. *Genomics* 2010, **95**:176-182.
84. Settles M, Zanella R, McKay SD, Schnabel RD, Taylor JF, Whitlock R, Schukken Y, Van Kessel JS, Smith JM, Neibergs H: **A whole genome association analysis identifies loci associated with Mycobacterium avium subsp. paratuberculosis infection status in US holstein cattle**. *Animal Genetics* 2009, **40**:655-662.
85. Becker D, Tetens J, Brunner A, Bürstel D, Ganter M, Kijas J, Drögemüller C, for the International Sheep Genomics C: **Microphthalmia in Texel Sheep Is Associated with a Missense Mutation in the Paired-Like Homeodomain 3 PITX3 Gene**. *PLoS One* 2010, **5**:e8689.
86. Spillman WJ: **Inheritance of Color Coat in Swine**. *Science* 1906, **24**:441-443.
87. Lush JL: **Inheritance in Swine**. *Journal of Heredity* 1921, **12**:57-71.
88. Hetzer HO: **Inheritance of coat color in swine**. *Journal of Heredity* 1945, **36**:121-128.
89. Ollivier L, Sellier P: **Pig genetics: a review**. *Genetics Selection Evolution* 1982, **14**:481-544.
90. Johansson M, Ellegren H, Marklund L, Gustavsson U, Ringmar-Cederberg E, Andersson K, Edfors-Lilja I, Andersson L: **The gene for dominant white color in the pig is closely linked to ALB and PDGRFRA on chromosome 8**. *Genomics* 1992, **14**:965-969.
91. Kijas JM, Tornsten A, Chowdhary B, Andersson L: **Porcine Agouti gene map position SSC 17q21**. *Chromosome Res* 1998, **6**:243.
92. Mariani P, Moller MJ, Hoyheim B, Marklund L, Davies W, Ellegren H, Andersson L: **The extension coat color locus and the loci for blood group O and tyrosine aminotransferase are on pig chromosome 6**. *J Hered* 1996, **87**:272-276.
93. Andersson L: **Melanocortin receptor variants with phenotypic effects in horse, pig, and chicken**. *Ann N Y Acad Sci* 2003, **994**:313-318.
94. Dun G, Li X, Cao H, Zhou R, Li L: **Variations of melanocortin receptor 1 (MC1R) gene in three pig breeds**. *J Genet Genomics* 2007, **34**:777-782.
95. Fontanesi L, D'Alessandro E, Scotti E, Liotta L, Crovetti A, Chiofalo V, Russo V: **Genetic heterogeneity and selection signature at the KIT gene in pigs showing different coat colours and patterns**. *Anim Genet* 2010, **41**:478-492.
96. Giuffra E, Evans G, Tornsten A, Wales R, Day A, Looft H, Plastow G, Andersson L: **The Belt mutation in pigs is an allele at the Dominant white (I/KIT) locus**. *Mamm Genome* 1999, **10**:1132-1136.
97. Giuffra E, Kijas JM, Amarger V, Carlborg O, Jeon JT, Andersson L: **The origin of the domestic pig: independent domestication and subsequent introgression**. *Genetics* 2000, **154**:1785-1791.
98. Giuffra E, Tornsten A, Marklund S, Bongcam-Rudloff E, Chardon P, Kijas JM, Anderson SI, Archibald AL, Andersson L: **A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking KIT**. *Mamm Genome* 2002, **13**:569-577.
99. Gustafsson AC, Kijas JM, Alderborn A, Uhlen M, Andersson L, Lundeberg J: **Screening and scanning of single nucleotide polymorphisms in the pig**

melanocortin 1 receptor gene (MC1R) by pyrosequencing. *Anim Biotechnol* 2001, **12**:145-153.
100. Hadjiconstantouras C, Sargent CA, Skinner TM, Archibald AL, Haley CS, Plastow GS: **Characterization of the porcine KIT ligand gene: expression analysis, genomic structure, polymorphism detection and association with coat colour traits**. *Anim Genet* 2008, **39**:217-224.
101. Johansson Moller M, Chaudhary R, Hellmen E, Hoyheim B, Chowdhary B, Andersson L: **Pigs with the dominant white coat color phenotype carry a duplication of the KIT gene encoding the mast/stem cell growth factor receptor**. *Mamm Genome* 1996, **7**:822-830.
102. Kijas JM, Moller M, Plastow G, Andersson L: **A frameshift mutation in MC1R and a high frequency of somatic reversions cause black spotting in pigs**. *Genetics* 2001, **158**:779-785.
103. Kijas JM, Wales R, Tornsten A, Chardon P, Moller M, Andersson L: **Melanocortin receptor 1 (MC1R) mutations and coat color in pigs**. *Genetics* 1998, **150**:1177-1185.
104. Leeb T, Doppe A, Kriegesmann B, Brenig B: **Genomic structure and nucleotide polymorphisms of the porcine agouti signalling protein gene (ASIP)**. *Anim Genet* 2000, **31**:335-336.
105. Marklund S, Kijas J, Rodriguez-Martinez H, Ronnstrand L, Funa K, Moller M, Lange D, Edfors-Lilja I, Andersson L: **Molecular basis for the dominant white phenotype in the domestic pig**. *Genome Res* 1998, **8**:826-833.
106. Okumura N, Hayashi T, Sekikawa H, Matsumoto T, Mikawa A, Hamasima N, Awata T: **Sequencing, mapping and nucleotide variation of porcine coat colour genes EDNRB, MYO5A, KITLG, SLC45A2, RAB27A, SILV and MITF**. *Anim Genet* 2006, **37**:80-82.
107. Pielberg G, Olsson C, Syvanen AC, Andersson L: **Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig**. *Genetics* 2002, **160**:305-311.
108. Shi KR, Wang AG, Yuan XF, Deng XM, Li N: **Analysis of the MC1R, KIT and ASIP loci in Chinese and European pigs**. *Anim Genet* 2006, **37**:300-302.
109. Okumura N, Hayashi T, Uenishi H, Fukudome N, Komatsuda A, Suzuki A, Shibata M, Nii M, Yamaguchi T, Kojima-Shibata C *et al*: **Sequence polymorphisms in porcine homologs of murine coat colour-related genes**. *Anim Genet* 2010, **41**:113-121.
110. Okumura N, Hayashi T, Sekikawa H, Matsumoto T, Mikawa A, Hamasima N, Awata T: **Sequences and mapping of genes encoding porcine tyrosinase (TYR) and tyrosinase-related proteins (TYRP1 and TYRP2)**. *Anim Genet* 2005, **36**:513-516.
111. Goding CR: **Melanocytes: the new Black**. *Int J Biochem Cell Biol* 2007, **39**:275-279.
112. Jackson IJ: **Molecular and developmental genetics of mouse coat color**. *Annu Rev Genet* 1994, **28**:189-217.
113. Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS, O'Brien SJ: **Molecular genetics and evolution of melanism in the cat family**. *Curr Biol* 2003, **13**:448-453.
114. Fontanesi L, Beretti F, Riggio V, Dall'Olio S, Gonzalez EG, Finocchiaro R, Davoli R, Russo V, Portolano B: **Missense and nonsense mutations in melanocortin 1 receptor (MC1R) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences**. *BMC Genet* 2009, **10**:47.
115. Fontanesi L, Tazzoli M, Beretti F, Russo V: **Mutations in the melanocortin 1 receptor (MC1R) gene are associated with coat colours in the domestic rabbit (Oryctolagus cuniculus)**. *Anim Genet* 2006, **37**:489-493.

116. Kerje S, Lind J, Schutz K, Jensen P, Andersson L: **Melanocortin 1-receptor (MC1R) mutations are associated with plumage colour in chicken**. *Anim Genet* 2003, **34**:241-248.
117. Klungland H, Vage DI, Gomez-Raya L, Adalsteinsson S, Lien S: **The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination**. *Mamm Genome* 1995, **6**:636-639.
118. Newton JM, Wilkie AL, He L, Jordan SA, Metallinos DL, Holmes NG, Jackson IJ, Barsh GS: **Melanocortin 1 receptor variation in the domestic dog**. *Mamm Genome* 2000, **11**:24-30.
119. Raudsepp T, Kijas J, Godard S, Guerin G, Andersson L, Chowdhary BP: **Comparison of horse chromosome 3 with donkey and human chromosomes by cross-species painting and heterologous FISH mapping**. *Mamm Genome* 1999, **10**:277-282.
120. Robbins LS, Nadeau JH, Johnson KR, Kelly MA, Roselli-Rehfuss L, Baack E, Mountjoy KG, Cone RD: **Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function**. *Cell* 1993, **72**:827-834.
121. Vage DI, Fleet MR, Ponz R, Olsen RT, Monteagudo LV, Tejedor MT, Arruga MV, Gagliardi R, Postiglioni A, Nattrass GS *et al*: **Mapping and characterization of the dominant black colour locus in sheep**. *Pigment Cell Res* 2003, **16**:693-697.
122. Vage DI, Lu D, Klungland H, Lien S, Adalsteinsson S, Cone RD: **A non-epistatic interaction of agouti and extension in the fox, Vulpes vulpes**. *Nat Genet* 1997, **15**:311-315.
123. Valverde P, Healy E, Jackson I, Rees JL, Thody AJ: **Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans**. *Nat Genet* 1995, **11**:328-330.
124. Vidal O, Vinas J, Pla C: **Variability of the melanocortin 1 receptor (MC1R) gene explains the segregation of the bronze locus in turkey (Meleagris gallopavo)**. *Poult Sci* 2010, **89**:1599-1602.
125. Frandberg PA, Muceniece R, Prusis P, Wikberg J, Chhajlani V: **Evidence for alternate points of attachment for alpha-MSH and its stereoisomer [Nle4, D-Phe7]-alpha-MSH at the melanocortin-1 receptor**. *Biochem Biophys Res Commun* 1994, **202**:1266-1271.
126. Schioth HB, Yook P, Muceniece R, Wikberg JE, Szardenings M: **Chimeric melanocortin MC1 and MC3 receptors: identification of domains participating in binding of melanocyte-stimulating hormone peptides**. *Mol Pharmacol* 1998, **54**:154-161.
127. Drogemuller C, Giese A, Martins-Wess F, Wiedemann S, Andersson L, Brenig B, Fries R, Leeb T: **The mutation causing the black-and-tan pigmentation phenotype of Mangalitza pigs maps to the porcine ASIP locus but does not affect its coding sequence**. *Mamm Genome* 2006, **17**:58-66.
128. Garcia-Borron JC, Sanchez-Laorden BL, Jimenez-Cervantes C: **Melanocortin-1 receptor structure and functional regulation**. *Pigment Cell Res* 2005, **18**:393-410.
129. Stearns TM, Beever JE, Southey BR, Ellis M, McKeith FK, Rodriguez-Zas SL: **Evaluation of approaches to detect quantitative trait loci for growth, carcass, and meat quality on swine chromosomes 2, 6, 13, and 18. I. Univariate outbred F-2 and sib-pair analyses**. *Journal of Animal Science* 2005, **83**:1481-1493.
130. Lai F, Ren J, Ai H, Ding N, Ma J, Zeng D, Chen C, Guo Y, Huang L: **Chinese white Rongchang pig does not have the dominant white allele of KIT but has the dominant black allele of MC1R**. *J Hered* 2007, **98**:84-87.

131. Romano G, Suon S, Jin H, Donaldson AE, Iacovitti L: **Characterization of five evolutionary conserved regions of the human tyrosine hydroxylase (TH) promoter: implications for the engineering of a human TH minimal promoter assembled in a self-inactivating lentiviral vector system**. *J Cell Physiol* 2005, **204**:666-677.
132. Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W: **PipMaker--a web server for aligning two genomic DNA sequences**. *Genome Res* 2000, **10**:577-586.
133. Van Raamsdonk CD, Barsh GS, Wakamatsu K, Ito S: **Independent regulation of hair and skin color by two G protein-coupled pathways**. *Pigment Cell Melanoma Res* 2009, **22**:819-826.
134. Van Raamsdonk CD, Fitch KR, Fuchs H, de Angelis MH, Barsh GS: **Effects of G-protein mutations on skin color**. *Nat Genet* 2004, **36**:961-968.
135. Jimenez-Cervantes C, Solano F, Lozano JA, Garcia-Borron JC: **The DHICA oxidase activity of the melanosomal tyrosinases LEMT and HEMT**. *Pigment Cell Res* 1994, **7**:298-304.
136. Kobayashi T, Urabe K, Winder A, Jimenez-Cervantes C, Imokawa G, Brewington T, Solano F, Garcia-Borron JC, Hearing VJ: **Tyrosinase related protein 1 (TRP1) functions as a DHICA oxidase in melanin biosynthesis**. *EMBO J* 1994, **13**:5818-5825.
137. Boissy RE, Zhao H, Oetting WS, Austin LM, Wildenberg SC, Boissy YL, Zhao Y, Sturm RA, Hearing VJ, King RA *et al*: **Mutation in and lack of expression of tyrosinase-related protein-1 (TRP-1) in melanocytes from an individual with brown oculocutaneous albinism: a new subtype of albinism classified as "OCA3"**. *Am J Hum Genet* 1996, **58**:1145-1156.
138. Manga P, Kromberg JG, Box NF, Sturm RA, Jenkins T, Ramsay M: **Rufous oculocutaneous albinism in southern African Blacks is caused by mutations in the TYRP1 gene**. *Am J Hum Genet* 1997, **61**:1095-1101.
139. Forshew T, Khaliq S, Tee L, Smith U, Johnson CA, Mehdi SQ, Maher ER: **Identification of novel TYR and TYRP1 mutations in oculocutaneous albinism**. *Clin Genet* 2005, **68**:182-184.
140. Rooryck C, Roudaut C, Robine E, Musebeck J, Arveiler B: **Oculocutaneous albinism with TYRP1 gene mutations in a Caucasian patient**. *Pigment Cell Res* 2006, **19**:239-242.
141. Berryere TG, Schmutz SM, Schimpf RJ, Cowan CM, Potter J: **TYRP1 is associated with dun coat colour in Dexter cattle or how now brown cow?** *Anim Genet* 2003, **34**:169-175.
142. Gratten J, Beraldi D, Lowder BV, McRae AF, Visscher PM, Pemberton JM, Slate J: **Compelling evidence that a single nucleotide substitution in TYRP1 is responsible for coat-colour polymorphism in a free-living population of Soay sheep**. *Proc Biol Sci* 2007, **274**:619-626.
143. Jackson IJ, Chambers D, Rinchik EM, Bennett DC: **Characterization of TRP-1 mRNA levels in dominant and recessive mutations at the mouse brown (b) locus**. *Genetics* 1990, **126**:451-459.
144. Lyons LA, Foe IT, Rah HC, Grahn RA: **Chocolate coated cats: TYRP1 mutations for brown color in domestic cats**. *Mamm Genome* 2005, **16**:356-366.
145. Schmidt-Kuntzel A, Eizirik E, O'Brien SJ, Menotti-Raymond M: **Tyrosinase and tyrosinase related protein 1 alleles specify domestic cat coat color phenotypes of the albino and brown loci**. *J Hered* 2005, **96**:289-301.
146. Schmutz SM, Berryere TG, Goldfinch AD: **TYRP1 and MC1R genotypes and their effects on coat color in dogs**. *Mamm Genome* 2002, **13**:380-387.
147. Johnson R, Jackson IJ: **Light is a dominant mouse mutation resulting in premature cell death**. *Nat Genet* 1992, **1**:226-229.

148. Quevedo WC, Chase HB: **An analysis of the light mutation of coat color in mice**. *Journal of Morphology* 1958, **102**:329-345.
149. Hoashi T, Watabe H, Muller J, Yamaguchi Y, Vieira WD, Hearing VJ: **MART-1 is required for the function of the melanosomal matrix protein PMEL17/GP100 and the maturation of melanosomes**. *J Biol Chem* 2005, **280**:14006-14016.
150. Jürgens G, Wieschaus E, Nüsslein-Volhard C, Kluding H: **Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster***. *Development Genes and Evolution* 1984, **193**:283-295.
151. Neckameyer WS, White K: **Drosophila tyrosine hydroxylase is encoded by the pale locus**. *J Neurogenet* 1993, **8**:189-199.
152. Marles LK, Peters EM, Tobin DJ, Hibberts NA, Schallreuter KU: **Tyrosine hydroxylase isoenzyme I is present in human melanosomes: a possible novel function in pigmentation**. *Exp Dermatol* 2003, **12**:61-70.
153. Knott SA, Marklund L, Haley CS, Andersson K, Davies W, Ellegren H, Fredholm M, Hansson I, Hoyheim B, Lundstrom K *et al*: **Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs**. *Genetics* 1998, **149**:1069-1080.
154. Rohrer GA, Keele JW: **Identification of quantitative trait loci affecting carcass composition in swine: I. Fat deposition traits**. *Journal of Animal Science* 1998, **76**:2247-2254.
155. Rohrer GA, Keele JW: **Identification of quantitative trait loci affecting carcass composition in swine: II. Muscling and wholesale product yield traits**. *Journal of Animal Science* 1998, **76**:2255-2262.
156. de Koning DJ, Janss LLG, Rattink AP, van Oers PAM, de Vries BJ, Groenen MAM, van der Poel JJ, de Groot PN, Brascamp EW, van Arendonk JAM: **Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa)**. *Genetics* 1999, **152**:1679-1690.
157. Paszek AA, Wilkie PJ, Flickinger GH, Rohrer GA, Alexander LJ, Beattie CW, Schook LB: **Interval mapping of growth in divergent swine cross**. *Mammalian Genome* 1999, **10**:117-122.
158. de Koning DJ, Rattink AP, Harlizius B, van Arendonk JAM, Brascamp EW, Groenen MAM: **Genome-wide scan for body composition in pigs reveals important role of imprinting**. *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**:7947-7950.
159. Rohrer GA: **Identification of quantitative trait loci affecting birth characters and accumulation of backfat and weight in a Meishan-White Composite resource population**. *Journal of Animal Science* 2000, **78**:2547-2553.
160. Wada Y, Akita T, Awata T, Furukawa T, Sugai N, Inage Y, Ishii K, Ito Y, Kobayashi E, Kusumoto H *et al*: **Quantitative trait loci (QTL) analysis in a Meishan x Gottingen cross population**. *Animal Genetics* 2000, **31**:376-384.
161. Bidanel JP, Milan D, Iannuccelli N, Amigues Y, Boscher MY, Bourgeois F, Caritez JC, Gruand J, Le Roy P, Lagant H *et al*: **Detection of quantitative trait loci for growth and fatness in pigs**. *Genetics Selection Evolution* 2001, **33**:289-309.
162. de Koning DJ, Harlizius B, Rattink AP, Groenen MAM, Brascamp EW, van Arendonk JAM: **Detection and characterization of quantitative trait loci for meat quality traits in pigs**. *Journal of Animal Science* 2001, **79**:2812-2819.
163. Malek M, Dekkers JCM, Lee HK, Baas TJ, Rothschild MF: **A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. I. Growth and body composition**. *Mammalian Genome* 2001, **12**:630-636.
164. Paszek AA, Wilkie PJ, Flickinger GH, Miller LM, Louis CF, Rohrer GA, Alexander LJ, Beattie CW, Schook LB: **Interval mapping of carcass and meat quality traits in a divergent swine cross**. *Animal Biotechnology* 2001, **12**:155-165.

165. Milan D, Bidanel JP, Iannuccelli N, Riquet J, Amigues Y, Gruand J, Le Roy P, Renard C, Chevalet C: **Detection of quantitative trait loci for carcass composition traits in pigs**. *Genetics Selection Evolution* 2002, **34**:705-728.
166. Nezer C, Moreau L, Wagenaar D, Georges M: **Results of a whole genome scan targeting QTL for growth and carcass traits in a Pietrain x Large White intercross**. *Genetics Selection Evolution* 2002, **34**:371-387.
167. Ovilo C, Clop A, Noguera JL, Oliver MA, Barragan C, Rodriguez C, Silo L, Toro MA, Coll A, Folch JM *et al*: **Quantitative trait locus mapping for meat quality traits in an Iberian x Landrace F-2 pig population**. *Journal of Animal Science* 2002a, **80**:2801-2808.
168. Varona L, Ovilo C, Clop A, Noguera JL, Perez-Enciso M, Coll A, Folch JM, Barragan C, Toro MA, Babot D *et al*: **QTL mapping for growth and carcass traits in an Iberian by Landrace pig intercross: additive, dominant and epistatic effects**. *Genetical Research* 2002, **80**:145-154.
169. Dekkers JCM, Kim JJ, Malek M, Thomsen H, Lee H, Zhao HH, Rothschild M: **A genome scan to detect QTL affecting growth, composition, and meat quality traits in a Berkshire x Yorkshire cross**. In: *28th National Swine Improvement Federation: 2003; Des Moines, Iowa*; 2003.
170. Geldermann H, Muller E, Moser G, Reiner G, Bartenschlager H, Cepica S, Stratil A, Kuryl J, Moran C, Davoli R *et al*: **Genome-wide linkage and QTL mapping in porcine F-2 families generated from Pietrain, Meishan and Wild Boar crosses**. *Journal of Animal Breeding and Genetics* 2003, **120**:363-393.
171. Sato S, Oyamada Y, Atsuji K, Nade T, Sato S, Kobayashi E, Mitsuhashi T, Nirasawa K, Komatsuda A, Saito Y *et al*: **Quantitative trait loci analysis for growth and carcass traits in a Meishan x Duroc F2 resource population**. *Journal of Animal Science* 2003, **81**:2938-2949.
172. Thomsen H, Lee HK, Rothschild MF, Malek M, Dekkers JCM: **Characterization of quantitative trait loci for growth and meat quality in a cross between commercial breeds of swine**. *Journal of Animal Science* 2004, **82**:2213-2228.
173. Nii M, Hayashi T, Mikawa S, Tani F, Niki A, Mori N, Uchida Y, Fujishirna-Kanaya N, Komatsu M, Awata T: **Quantitative trait loci mapping for meat quality and muscle fiber traits in a Japanese wild boar x Large White intercross**. *Journal of Animal Science* 2005, **83**:308-315.
174. Harmegnies N, Davin F, De Smet S, Buys N, Georges M, Coppieters W: **Results of a whole-genome quantitative trait locus scan for growth, carcass composition and meat quality in a porcine four-way cross**. *Anim Genet* 2006, **37**:543-553.
175. Karlskov-Mortensen P, Bruun CS, Braunschweig MH, Sawera M, Markljung E, Enfalt AC, Hedebro-Velander I, Josell A, Lindahl G, Lundstrom K *et al*: **Genome-wide identification of quantitative trait loci in a cross between Hampshire and Landrace I: carcass traits**. *Animal Genetics* 2006, **37**:156-162.
176. Kim CW, Hong YH, Yun SI, Lee SR, Kim YH, Kim MS, Chung KH, Jung WY, Kwon EJ, Hwang SS *et al*: **Use of microsatellite markers to detect quantitative trait loci in Yorkshire pigs**. *Journal of Reproduction and Development* 2006, **52**:229-237.
177. Rohrer GA, Thallman RM, Shackelford S, Wheeler T, Koohmaraie M: **A genome scan for loci affecting pork quality in a Duroc-Landrace F-2 population**. *Animal Genetics* 2006, **37**:17-27.
178. Wimmers K, Fiedler I, Hardge T, Murani E, Schellander K, Ponsuksili S: **QTL for microstructural and biophysical muscle properties and body composition in pigs**. *BMC Genetics* 2006, **7**:15.
179. van Wijk HJ, Dibbits B, Baron EE, Brings AD, Harlizius B, Groenen MAM, Knol EF, Bovenhuis H: **Identification of quantitative trait loci for carcass composition and pork quality traits in a commercial finishing cross**. *Journal of Animal Science* 2006, **84**:789-799.

180. Liu G, Jennen DG, Tholen E, Juengst H, Kleinwachter T, Holker M, Tesfaye D, Un G, Schreinemachers HJ, Murani E *et al*: **A genome scan reveals QTL for growth, fatness, leanness and meat quality in a Duroc-Pietrain resource population**. *Anim Genet* 2007, **38**:241-252.
181. Edwards DB, Ernst CW, Tempelman RJ, Rosa GJ, Raney NE, Hoge MD, Bates RO: **Quantitative trait loci mapping in an F2 Duroc x Pietrain resource population: I. Growth traits**. *J Anim Sci* 2008, **86**:241-253.
182. Edwards DB, Ernst CW, Raney NE, Doumit ME, Hoge MD, Bates RO: **Quantitative trait locus mapping in an F2 Duroc x Pietrain resource population: II. Carcass and meat quality traits**. *J Anim Sci* 2008, **86**:254-266.
183. Markljung E, Braunschweig MH, Karlskov-Mortensen P, Bruun CS, Sawera M, Cho IC, Hedebro-Velander I, Josell A, Lundstrom K, von Seth G *et al*: **Genome-wide identification of quantitative trait loci in a cross between Hampshire and Landrace II: meat quality traits**. *BMC Genet* 2008, **9**:22.
184. Guo YM, Lee GJ, Archibald AL, Haley CS: **Quantitative trait loci for production traits in pigs: a combined analysis of two Meishan x Large White populations**. *Anim Genet* 2008, **39**:486-495.
185. Liu G, Kim JJ, Jonas E, Wimmers K, Ponsuksili S, Murani E, Phatsara C, Tholen E, Juengst H, Tesfaye D *et al*: **Combined line-cross and half-sib QTL analysis in Duroc-Pietrain population**. *Mamm Genome* 2008, **19**:429-438.
186. Duan YY, Ma JW, Yuan F, Huang LB, Yang KX, Xie JP, Wu GZ, Huang LS: **Genome-wide identification of quantitative trait loci for pork temperature, pH decline, and glycolytic potential in a large-scale White Duroc x Chinese Erhualian resource population**. *J Anim Sci* 2009, **87**:9-16.
187. Ma J, Ren J, Guo Y, Duan Y, Ding N, Zhou L, Li L, Yan X, Yang K, Huang L *et al*: **Genome-wide identification of quantitative trait loci for carcass composition and meat quality in a large-scale White Duroc x Chinese Erhualian resource population**. *Anim Genet* 2009, **40**:637-647.
188. Green P, Falls K, Crooks S: **Documentation for CRIMAP, version 2.4**. In., 2.4 edn. St. Louis, MO: Washington University School of Medicine; 1989.
189. Seaton G, Haley CS, Knott SA, Kearsey M, Visscher PM: **QTL Express: mapping quantitative trait loci in of simple and complex pedigrees**. *Bioinformatics* 2002, **18**:339-340.
190. Churchill GA, Doerge RW: **Empirical threshold values for quantitative trait mapping**. *Genetics* 1994, **138**:963-971.
191. Lange K: **Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets**. *American journal of human genetics* 2001, **69**:A1886.
192. Almasy L, Blangero J: **Multipoint quantitative-trait linkage analysis in general pedigrees**. *Am J Hum Genet* 1998, **62**:1198-1211.
193. Meyers SN, Rodriguez-Zas SL, Beever JE: **Fine-mapping of a QTL influencing pork tenderness on porcine chromosome 2**. *BMC Genet* 2007, **8**:69.
194. Meyers SN, Beever JE: **Investigating the genetic basis of pork tenderness: genomic analysis of porcine CAST**. *Anim Genet* 2008, **39**:531-543.
195. Oh SH, Lee DH, See MT: **Genetic Correlations Between Growth and Carcass Traits**. In: *Animal Science Departmental Report.* Edited by Lovercamp K, See T. Raleigh: North Carolina State University; 2004-2005.
196. Mohrmann M, Roehe R, Knap PW, Looft H, Plastow GS, Kalm E: **Quantitative trait loci associated with AutoFOM grading characteristics, carcass cuts and chemical body composition during growth of Sus scrofa**. *Animal Genetics* 2006, **37**:435-443.

197. Yue G, Stratil A, Kopecny M, Schroffelova D, Schroffel J, Hojny J, Cepica S, Davoli R, Zambonelli P, Brunsch C *et al*: **Linkage and QTL mapping for Sus scrofa chromosome 6**. *Journal of Animal Breeding and Genetics* 2003, **120**:45-55.
198. Ovilo C, Oliver A, Noguera JL, Clop A, Barragan C, Varona L, Rodriguez C, Toro M, Sanchez A, Perez-Enciso M *et al*: **Test for positional candidate genes for body composition on pig chromosome 6**. *Genetics Selection Evolution* 2002b, **34**:465-479.
199. Perez-Enciso M, Misztal I: **Qxpak: a versatile mixed model application for genetical genomics and QTL analyses**. *Bioinformatics* 2004, **20**:2792-2798.
200. Arends D, Prins P, Jansen RC, Broman KW: **R/qtl: high-throughput multiple QTL mapping**. *Bioinformatics* 2010, **26**:2990-2992.
201. Haley CS, Knott SA: **A simple regression method for mapping quantitative trait loci in line crosses using flanking markers**. *Heredity* 1992, **69**:315-324.
202. Zhou L, Azfer A, Niu J, Graham S, Choudhury M, Adamski FM, Younce C, Binkley PF, Kolattukudy PE: **Monocyte chemoattractant protein-1 induces a novel transcription factor that causes cardiac myocyte apoptosis and ventricular dysfunction**. *Circ Res* 2006, **98**:1177-1185.
203. Kim RH, Wang D, Tsang M, Martin J, Huff C, de Caestecker MP, Parks WT, Meng X, Lechleider RJ, Wang T *et al*: **A novel smad nuclear interacting protein, SNIP1, suppresses p300-dependent TGF-beta signal transduction**. *Genes Dev* 2000, **14**:1605-1616.
204. Racevskis J, Dill A, Stockert R, Fineberg SA: **Cloning of a novel nucleolar guanosine 5'-triphosphate binding protein autoantigen from a breast tumor**. *Cell Growth Differ* 1996, **7**:271-280.
205. Lee SM, Tsui SK, Chan KK, Kotaka M, Li HY, Chim SS, Waye MM, Fung KP, Lee CY: **Chromosomal mapping of a skeletal muscle specific LIM-only protein FHL3 to the distal end of the short arm of human chromosome 1**. *Somat Cell Mol Genet* 1998, **24**:197-202.
206. Dai KS, Liew CC: **A novel human striated muscle RING zinc finger protein, SMRZ, interacts with SMT3b via its RING domain**. *J Biol Chem* 2001, **276**:23992-23999.
207. Witt CC, Witt SH, Lerche S, Labeit D, Back W, Labeit S: **Cooperative control of striated muscle mass and metabolism by MuRF1 and MuRF2**. *EMBO J* 2008, **27**:350-360.
208. Bodine SC, Latres E, Baumhueter S, Lai VK, Nunez L, Clarke BA, Poueymirou WT, Panaro FJ, Na E, Dharmarajan K *et al*: **Identification of ubiquitin ligases required for skeletal muscle atrophy**. *Science* 2001, **294**:1704-1708.
209. Yang DH, Moss EG: **Temporally regulated expression of Lin-28 in diverse tissues of the developing mouse**. *Gene Expr Patterns* 2003, **3**:719-726.
210. Ishii I, Akahoshi N, Yamada H, Nakano S, Izumi T, Suematsu M: **Cystathionine gamma-Lyase-deficient mice require dietary cysteine to protect against acute lethal myopathy and oxidative injury**. *J Biol Chem* 2010, **285**:26358-26368.

Appendix A: Coat color association analysis results

Table 1a: List of SNPs found significant (Bonferroni adjusted p<0.01) by allelic chi square test performed by PLINK contrasting solid white animals with solid red animals with MC1R genotype of "e/e". Table 1b: List of SNPs found significant (Bonferroni adjusted p<0.01) by allelic chi square test performed by PLINK contrasting solid white animals with purebred Berkshire and F2 animals displaying the Berkshire phenotype. Table 1c: List of SNPs found significant (Bonferroni adjusted p<0.01) by allelic chi square test performed by PLINK contrasting solid white animals, purebred Berkshire and F2 animals displaying the Berkshire phenotype with solid red animals with MC1R genotype of "e/e". These tables can be found in the supplementary file called "Supplementary Tables 1a, 1b, 1c - Coat color association analysis results.xlsx" located on the accompanying CD.

Appendix B: SOLAR whole genome QTL scan results files

Folder on accompanying CD labeled “Supplementary Files” contains 30 subfolders each labeled according to the phenotype analysis results it contains. Each subfolder will contain the results files for all 18 chromosomes. The files within these subfolders contain the results from the analyses performed by SOLAR. Each .xlsx file contains at least one tab, labeled "SOLAR", that contains both the raw and adjusted LOD scores in addition to a plot of the adjusted LOD scores. Files ending in " - sig" contain results from a chromosome where either QTLexpress, SOLAR or both programs identified at least one significant quantitative trait locus for this phenotype. Each of these files will contain three tabs regardless of which program identified the significant locus. The "QTLexpress" tab contains both the raw LOD score data from QTLexpress and a plot of the LOD scores. The "Both" tab contains an overlaid plot of LOD scores from the QTLexpress analysis and the adjusted LOD scores from the SOLAR analysis, with the QTLexpress LOD scores in blue and the SOLAR adjusted LOD scores in red.

Appendix C: SOLAR whole genome QTL scan; SNPs with LOD scores greater than 3

Results from the SOLAR analysis, table containing all the SNPs with LOD scores greater than 3. For each phenotype, the SNPs are listed in order of Megabase position. Listed for each SNP is the LOD score calculated by SOLAR, the adjusted LOD score, the log likelihood of the model used, the amount of the variance not explained by the SNP (H2r), the amout of the variance that is explained by the SNP, and the significant region the SNP belongs to. This table can be found in the supplementary file called "Supplementary Table 2 – SOLAR whole genome QTL scan; SOLAR SNPs w LOD greater than 3.xlsx located on the accompanying CD.

Appendix D: SSC6 LEA fine mapping results

Results from the fine mapping of the 95%CI of the LEA QTL on SSC6. Listed for each SNP is the Megabase position on SSC6, the LOD score from the QxPak analysis, the LOD score from the R/qtl analysis, the LOD score from the SOLAR analysis, the empirical p-value from the PLINK analysis, and the absolute value of the SNP effect from the Bayes-C analysis. The third column in the file contains the LOD scores from the QTLexpress analysis, these LOD scores correspond to certain Megabase positions across the interval rather than to a specific SNPs, so the LOD scores are listed next to the SNP with the nearest Megabase position to the QTLexpress position. SNPs with significant LOD scores, p-values or effects, are highlighted in green, with the most significant SNP or SNPs highlighted in blue. The SOLAR analysis did not identify any significant SNPs, so to compare the SNPs with the highest LOD scores with the other analyses, SNPs with LOD scores between 2 and 2.9 are highlighted in gold. The consensus regions discussed in chapter 4 are highlighted in either dark or light red. This table can be found in the supplementary file called "Supplementary Table 3 – SSC6 LEA Fine Mapping Results.xlsx" located on the accompanying CD.