TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Tierzucht

QTL and candidate gene analysis of energy and lipid metabolism in swine

Anna Pertek

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Dr. h. c. J. Bauer

Prüfer der Dissertation:

1. Univ.-Prof. Dr. H.-R. Fries
2. Univ.-Prof- Dr. M. Klingenspor

Die Dissertation wurde am 25.11.2010 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum für Ernährung, Landnutzung und Umwelt am 13.06.2011 angenommen.

Table of contents:

Chapter 1

General introduction

Introduction

The domestic pig, *Sus scrofa*, is an important mammalian model organism for agricultural research due to its worldwide importance as a food source, as well as biomedical research on human disease conditions due to high similarity in terms of anatomy and physiology. Domestication and breeding of pigs have created a phenotypic diversity among breeds, which is characterised by a rich collection of mutations with favourable phenotypic effects (Andersson & Georges 2004). Several approaches have been applied to identify genes and mutations responsible for the phenotypic diversity (Andersson & Georges 2004), e.g. obesity in humans or fatness in the pig. In general, these approaches can be divided into two basic categories: whole genome scan, such as linkage mapping and genome-wide association study as well as conventional candidate gene studies (Hirschhorn & Daly 2005; Chen *et al.* 2007a). Each of these approaches has specific advantages and disadvantages. Genome-wide scanning usually proceeds disregarding the importance of specific functional features of the investigated traits, however basic disadvantages of this method are high costs and the fact that the procedure is resource-intensive. In principle, genome-wide scanning only locates quantitative trait loci at cM-level of the chromosomal regions to take advantage of DNA markers under family-based or population-based experimental designs, which usually comprise a large number of candidate genes. In comparison, the conventional candidate gene approach has been ubiquitously applied for gene-disease research and genetic association studies as an extremely powerful method for studying the genetic architecture of complex traits. In addition, this approach is more effective and economical for direct gene discovery. Nevertheless, the practicability of traditional candidate gene approach is largely limited to the existing knowledge about the known biological function of potentially contributing genes and knowledge of the phenotype under investigation (Zhu *et al.* 2007).

QTL study

Most of traits of economic importance in livestock, including growth, meat quality, fatness, fertility and behaviour, are controlled by an unknown number of genes, each segregating according to Mendel's laws, and by environmental factors. The regions of the genome that harbour one or more genes affecting quantitative traits are called quantitative trait loci (QTL) (Andersson 2001). QTL experiments are designed to study the genetic basis of phenotypic differences between different natural populations (Zeng

2005). Most of these studies were conducted on crosses between divergent breeds to create a hybrid. Subsequently, backcross of the hybrid to the parental population or an intercross between hybrids are produced to create the F2 population. QTL mapping analysis is performed on theses segregating populations to map QTLs exhibiting an effect on the trait of interest.

The current release of the PigQTLdb (http://www.animalgenome.org/QTLdb/pig.html) (Hu *et al.* 2007) contains 1,831 QTLs representing 316 different pig traits. The biggest numbers of QTLs were reported for fatness, in particular average backfat thickness. For other traits such as health or disease resistance few QTLs were discovered. However, only a limited number of these QTLs have been further investigated to the point that a known causative mutation has been implicated or proven (Rothschild *et al.* 2007). The identification of genes and mutations that underlie a QTL is problematic, because of difficulties with the determination of the exact chromosomal location of a QTL, a mild phenotypic QTL effect or problems with the discrimination between a causative mutation and neutral polymorphism. Moreover, a large proportion of QTL mutations are regulatory mutations, QTL effects might reflect the combined action of clusters of linked mutations or epigenetic inheritance might contribute to QTL variation (Andersson & Georges 2004). Interestingly, the first QTL found on SSC4 for fatness is still not identified (Andersson *et al.* 1994). The mutations that underlie mapped QTLs have been identified in the insulin-like growth factor 2 gene (*IGF2*) for muscling (Nezer *et al.* 1999; Van Laere *et al.* 2003) as well as the calpastatin gene (*CAST*) for tenderness on SSC2 (Ciobanu *et al.* 2004), and in the protein kinase, AMP-activated, gamma 3 non-catalytic subunit gene (*PRKAG3*) for glycogen content and meat quality on SSC15 (Milan *et al.* 2000).

Performance traits, such as fatness and growth traits, are relatively easy to measure and therefore numerous QTLs have been described for these traits. Traits, which are more difficult or expensive to measure, e.g. meat quality or fatty acid composition, were examined in a limited number of studies (Table 1). It is likely that some of these studies were carried out on a limited number of chromosomes and directed to QTL-rich genome regions reported in previously published studies. Thus, the number of QTLs on the most extensively studied chromosomes (SSC 1, 2, 4, 6, 7, X) could be over-represented.

Table 1 Number of QTL per chromosome and trait class obtained from the PigQTLDB.

Trait class	***Chromosome***																			***Total***
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**X**	**Y**	
Fatness[1]	51	31	8	40	18	52	67	4	7	8	9	6	17	10	3	3	4	6	32	0	376
Growth[2]	22	11	7	32	5	15	12	9	9	5	1	3	10	2	11	3	2	1	8	0	168
Meat quality[3]	12	14	7	7	2	26	8	1	2	0	2	3	5	2	2	3	2	5	7	0	110
Fatty acid	5	0	0	21	3	2	0	3	4	1	0	1	0	0	6	0	0	1	1	0	48
Total	90	56	22	100	28	95	87	17	22	14	12	13	32	14	22	9	8	13	48	0	702

[1]Fatness: backfat, external fat

[2]Growth: average daily gain, feed intake, feed conversion

[3]Meat quality: intramuscular fat content, pH, lean meat percentage

Candidate gene: *PDK4*

Pyruvate dehydrogenase kinase, isozyme 4 (PDK4) is a serine/threonine protein kinase that selectively inhibits the activity of the pyruvate dehydrogenase complex (PDC) (Rowles *et al.* 1996).

The mitochondrial PDC catalyzes the irreversible oxidative decarboxylation of pyruvate to form acetyl-CoA, NADH and CO_2 and thus links glycolysis to the tricarboxylic acid cycle and adenosine triphosphate (ATP) production (Sugden 2003). In addition, when the glucose supply is abundant, the combination of mitochondrial acetyl-CoA with oxaloacetate provides a precursor for malonyl-CoA production. Malonyl-CoA may limit the oxidation of long chain fatty acids and accelerate fatty acid biosynthesis from glucose in the lipogenic tissue . When glucose is scarce or long chain fatty acids supply and oxidation is sufficient to meet cellular energy requirements, mammalian PDC activity is suppressed. Under these conditions long chain fatty acid oxidation is facilitated (Holness & Sugden 2003). The mammalian PDC consists of three catalytic components: pyruvate dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and dihydrolipoamide dehydrogenase (E3); a structural component protein X and two regulatory enzymes: pyruvate dehydrogenase kinase (PDK) and pyruvate dehydrogenase phosphatase (PDP) (Patel & Roche 1990; Sugden & Holness 1994). Activity of the PDC is regulated by the phosphorylation (inactivation)-dephosphorylation (activation) cycle of three serine residues (site-1, serine-264; site-2, serine-271; site-3, serine-203) on a-subunits of the E1 component (Patel & Korotchkina 2001; Sugden & Holness 2003). In the human, phosphorylation is carried out by four PDK isoenzymes (PDK1, PDK2, PDK3, PDK4). Each has its unique tissue distribution, kinetic properties and different sensitivities to regulatory molecules (Bowker-Kinley *et al.* 1998). Of the four isoforms, PDK4 has been reported to be the most highly inducible (Wu *et al.* 1998) and therefore PDK4 is of special interest. For this reason, *PDK4* is considered as a candidate gene for energy and lipid metabolism in swine.

The PDK4 activity, similar to that of the other PDKs, is regulated in the short- and long-term mechanisms. In the short term, PDK4 is reversibly regulated by the NADH/NAD+ and acetyl-CoA/CoA concentration ratios. The accumulation of products of the pyruvate dehydrogenase reaction (NADH and acetyl-CoA) activates PDK4. The activity of kinase is inhibited by excess of NAD^+, CoA and pyruvate (Bowker-Kinley *et al.*

1998; Harris *et al.* 2002; Holness & Sugden 2003). In the long term, various nutritional conditions and pathological disorders (starvation, diabetes) increase PDK4 activity (Bowker-Kinley *et al.* 1998; Wu *et al.* 1999; Holness *et al.* 2000; Sugden *et al.* 2000a).

PDK4 is highly expressed in several tissues directly involved in fuel homeostasis, including the heart (Bowker-Kinley *et al.* 1998; Wu *et al.* 1998), skeletal muscles (Bowker-Kinley *et al.* 1998; Wu *et al.* 1999; Sugden *et al.* 2000a; Peters *et al.* 2001), the liver (Bowker-Kinley *et al.* 1998; Wu *et al.* 2000; Holness *et al.* 2003), kidneys (Bowker-Kinley *et al.* 1998; Wu *et al.* 2000; Sugden *et al.* 2001b) and pancreatic islets (Sugden *et al.* 2001a). Factors involved in the regulation of PDK4 expression include glucocorticoids and free fatty acids, which have positive effects, and insulin, which has a negative effect (Sugden *et al.* 2000b; Harris *et al.* 2002; Holness *et al.* 2003; Abbot *et al.* 2005). The peroxisome proliferators-activated receptor alpha (PPARα) and peroxisome proliferators-activated receptor γ (PPARγ) agonists have been found to up-regulate PDK4 mRNA expression in a tissue-specific manner (Holness *et al.* 2002; Abbot *et al.* 2005). PDK4 shows high activity for site-1 and has the highest activity toward site-2 of all PDK isoforms; however, it does not phosphorylate site-3 of free E1 (Korotchkina & Patel 2001).

The porcine *PDK4* has been assigned to a region on chromosome 9q12-(1/3)q21 (Davoli *et al.* 2000; Maak *et al.* 2001). This region has been reported to contain QTLs for backfat thickness, daily gain (Kim *et al.* 2006; Edwards *et al.* 2008), feed intake (Cepica *et al.* 2003a) and saturated fatty acids (Nii *et al.* 2006) (Table 2).

Table 2 QTLs on chromosome 9 reported in different pig populations.

Trait	*QTL Center Location*	*95% CI (cM)*	*Cross*	*Reference*
External Fat On Ham	77.0 (cM)	64.5–133.5	Hampshire x Landrace	KARLSKOV-MORTENSEN *et al.* 2006
Feed Intake	78.3 (cM)	66.7-80.9	Meishan x Pietrain x European Wild Boar	CEPICA *et al*. 2003
Total Saturated Fatty Acids	78.6 (cM)	61.5-83.3	Japanese Wild Boar x Large White	NII *et al.* 2006
Average Daily Gain	82.4 (cM)	n/a	Duroc x Yorkshire	KIM *et al*. 2006
Backfat (Average) Thickness	82.4 (cM)	n/a	Duroc x Yorkshire	KIM *et al*. 2006
Body Weight At 10 Weeks Of Age	87.0 (cM)	n/a	Large White x Meishan	QUINTANILLA *et al*. 2002
Backfat at tenth rib	88.4 (cM)	79-96	Duroc x Pietrain	EDWARDS *et al*. 2008
Shoulder Weight	91.0 (cM)	65.0-108.0	Large White x Meishan	MILAN *et al*. 2002

Candidate gene: *INSIG2*

The insulin induced gene 2 (*INSIG2*) encodes the second polytopic membrane protein of the endoplasmic reticulum (ER) (Insig-2), which binds in a sterol-regulated fashion the sterol regulatory element binding protein (SREBP), the cleavage-activating protein (Scap) and 3-hydroxy-3-methylglutaryl (HMG) CoA reductase (Goldstein *et al.* 2006). By controlling Scap and HMG CoA reductase, Insig proteins regulate cholesterol homeostasis (Goldstein *et al.* 2006).

In sterol-depleted cells, Scap transports SREBPs from the ER to the Golgi apparatus, where SREBPs undergo proteolytic activation (Gong *et al.* 2006). The active portions of SREBPs are released into the cytosol, hence they can enter the nucleus and regulate the expression of more than 30 genes involved in cholesterol, fatty acid, triglyceride and phospholipid metabolism (Yabe *et al.* 2002; McPherson & Gauthier 2004). In sterol-overloaded cells, Scap undergoes a conformational change mediated through its sterol-sensing domain. As a result, Scap binds to Insig proteins and this reaction leads to ER retention of Scap and prevents the delivery of SREBPs to the Golgi apparatus (Gong *et al.* 2006) (Figure 1).

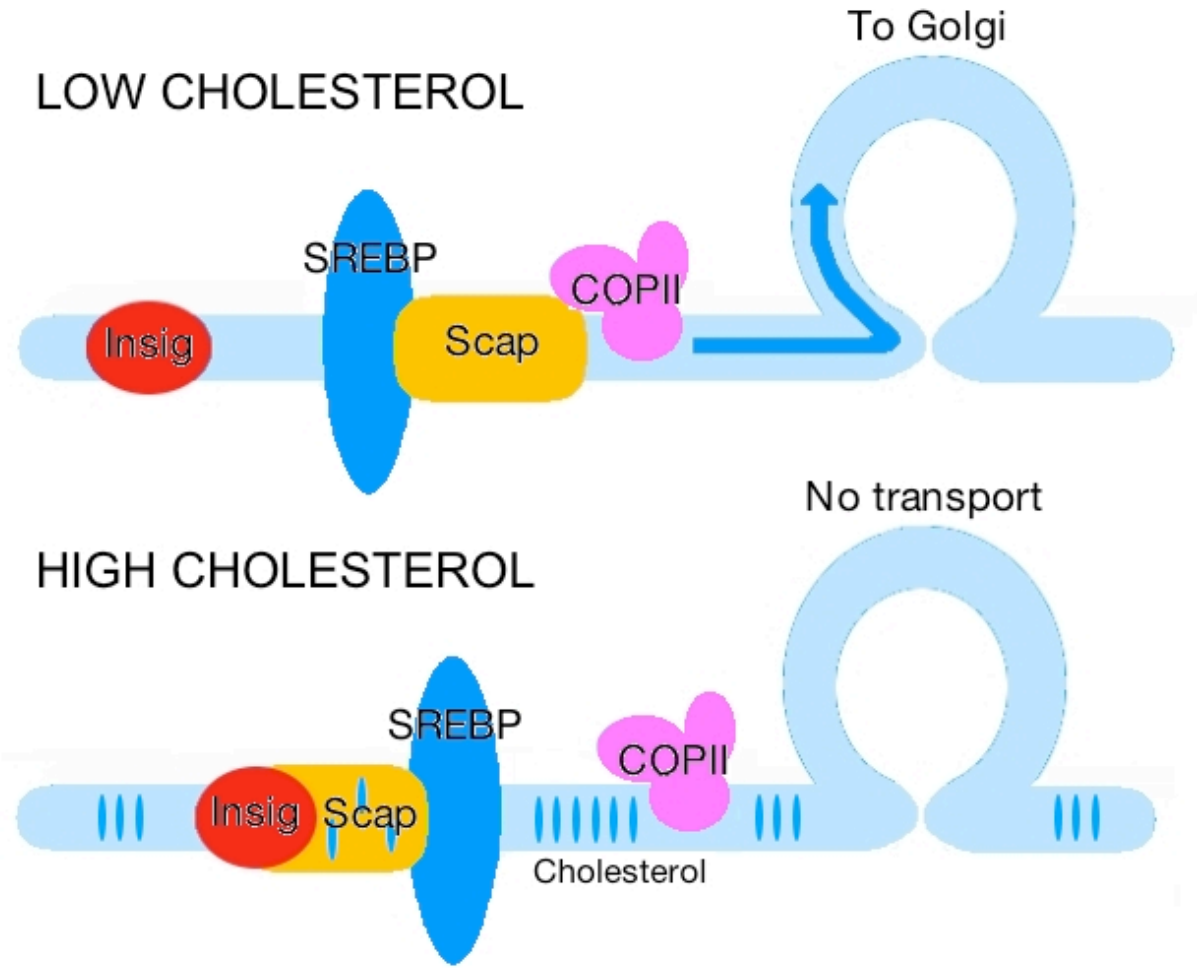


Figure 1 Low cholesterol: Scap escorts SREBPs to Golgi by binding COPII protein, nuclear forms of SREBPs activating genes of cholesterol synthesis and uptake are processed. High cholesterol: Scap/SREBP complex binds to Insig, halting transport of SREBPs to Golgi. Modified according to Goldstein *et al.* (Goldstein *et al.* 2006).

At the transcriptional level, HMG CoA reductase is controlled by the sterol-regulated transport mechanism of SREBPs, as it was described above. At the posttranscriptional level, HMG CoA reductase is regulated by products of the mevalonate pathway, in particular lanosterol. Lanosterol accumulation in ER membranes induces the reductase to bind to Insig proteins. This event leads to the ubiquitination of reductase by an Insig-bound ubiquitin ligase. Finally, the HMG CoA reductase is degraded in the proteasome (Sever *et al.* 2003; Goldstein *et al.* 2006; Gong *et al.* 2006).

Insig-2 was discovered through bioinformatics as a closely related isoform with attributes comparable to those of Insig-1, but differing in its pattern of regulation (Yabe *et al.* 2002; Goldstein *et al.* 2006). Insig-2 consists of six transmembrane helices

separated by short hydrophilic loops. The major differences between two Insigs are short sequences at the NH_2 and COOH termini (Gong *et al.* 2006). In Insig-2, Asp-149 is an essential residue required for interaction with Scap and HMG CoA reductase. This aspartic acid is adjacent to the fourth transmembrane helix of Insig protein at the cytosolic side of the ER (Gong *et al.* 2006). Insig-2 is expressed at a low, but constitutive level and is not regulated by SREBPs, in contrast to Insig-1, which requires nuclear SREBPs for its expression (Yabe *et al.* 2002). The highest expression of Insig-2 is reported in the murine liver and is selectively down-regulated by insulin (Yabe *et al.* 2003). Additionally, the transcript was found in murine testes, kidneys, heart, muscle, white fat and adrenal glands (Yabe *et al.* 2002). In human tissue *INSIG2* showed a ubiquitous expression pattern (Krapivner *et al.* 2008). Insig-2 can block SREBP processing only in the presence of sterols (Yabe *et al.* 2002). In the murine liver two transcripts with different first exons from the *INSIG2* gene were detected. Both transcripts splice into a common exon 2, which contains the start codon (Yabe *et al.* 2002).

INSIG2 was indicated through genome-wide association scan as a candidate gene for increasing the body mass index (BMI) in adults and children. Herbert *et al.* (Herbert *et al.* 2006) reported that a single nucleotide polymorphism (rs7566605), which lies 10 kb upstream from the transcription start site of the *INSIG2* gene, is associated with BMI in four separate samples composed of individuals of Western European ancestry, African Americans, and children. The association was replicated in several cohorts (Lyon *et al.* 2007; Hotta *et al.* 2008; Reinehr *et al.* 2008; Yang *et al.* 2008; Zhang *et al.* 2008; Le Hellard *et al.* 2009). However, no evidence for a major role of this variant in obesity was found in other studies (Hall *et al.* 2006; Dina *et al.* 2007b; Kumar *et al.* 2007; Kuzuya *et al.* 2007; Loos *et al.* 2007; Lyon *et al.* 2007; Rosskopf *et al.* 2007; Andreasen *et al.* 2008; Boes *et al.* 2008; Tabara *et al.* 2008; Yang *et al.* 2008). The possible explanations for observed divergences were widely discussed by Lyon et al. (Lyon *et al.* 2007). Nevertheless, *INSIG2* remains an attractive candidate gene for the QTL affecting BMI or fatness related traits in pigs.

The *INSIG2* region has been implicated as a factor in obesity by linkage studies in mice (Cheverud *et al.* 2004) and humans (Deng *et al.* 2002). The porcine *INSIG2* has been assigned to 15q14-q15 at a distance of ~49 cR from the most significantly linked

marker SW1118 (Qiu *et al.* 2005). In swine QTLs for backfat, fatty acids and meat quality were detected in this region (Table 3).

Table 3 QTLs on chromosome 15 reported in different pig populations.

Trait	*QTL Center Location*	*95% CI (cM)*	*Cross*	*Reference*
Backfat At Tenth Rib	34.6 (cM)	34.6-47.4	Meishan x Pietrain x European Wild Boar	KURYL *et al.* 2003
Backfat (Average) Thickness	34.6 (cM)	34.6-47.4	Meishan x Pietrain x European Wild Boar	KURYL *et al.* 2003
Backfat Depth At Last Rib	34.6 (cM)	34.6-47.4	Meishan x Pietrain x European Wild Boar	KURYL *et al.* 2003
Dressing	35 (cM)	34.6-57.09	Meishan x Pietrain x European Wild Boar	GELDERMANN *et al.* 2003
Palmitic Acid Percentage	35.9 (cM)	28.9-50.7	Japanese Wild Boar x Large White	NII *et al.* 2006
Subjective Juiciness Score	37 (cM)	29.5-50.7	Duroc x Landrace	ROHRER *et al.* 2006
Dressing Percentage	37.4 (cM)	34.6-47.4	Meishan x Pietrain x European Wild Boar	KURYL *et al.* 2003
Lean Percentage	38.3 (cM)	34.6-47.4	Meishan x Pietrain x European Wild Boar	KURYL *et al.* 2003
Total Saturated Fatty Acids	45.3 (cM)	28.9-50.7	Japanese Wild Boar x Large White	NII *et al.* 2006

Candidate gene: *FTO*

The fat mass and obesity associated gene (*FTO*) is an obesity susceptibility gene discovered in three independent genome-wide association studies (Dina *et al.* 2007a; Frayling *et al.* 2007; Scuteri *et al.* 2007). A cluster of variants in the first intron of FTO showed a highly significant association with obesity-related traits. In the first study Frayling *et al.* (Frayling *et al.* 2007) identified a common variant in the first intron of *FTO* that predisposes to diabetes through an effect on BMI. The association with BMI was replicated in 13 cohorts with more than 38 000 participants, including children and adults. The second study, by Dina *et al.* (Dina *et al.* 2007a), indicated a strong association of a set of single nucleotide polymorphisms (SNPs) in the first intron of *FTO* with early-onset and severe obesity in adults and children of European ancestry. In

the third study, Scuteri *et al*. (Scuteri *et al.* 2007) reported an association between variants in *FTO* and substantial changes in BMI, hip circumference and body weight in the genetically isolated population of more than 4 000 Sardinians. The results were replicated in European Americans and Hispanic Americans. Additionally, FTO SNPs were reported to be associated with various obesity-related traits, such as body weight (Andreasen *et al.* 2007; Peeters *et al.* 2007; Scuteri *et al.* 2007), leptin levels (Andreasen *et al.* 2007), subcutaneous fatness (Frayling *et al.* 2007; Peeters *et al.* 2007), fat mass (Andreasen *et al.* 2007; Frayling *et al.* 2007; Peeters *et al.* 2007), hip (Scuteri *et al.* 2007) and waist circumferences (Andreasen *et al.* 2007), but not with lean mass or body height (Andreasen *et al.* 2007; Frayling *et al.* 2007; Peeters *et al.* 2007).

The associations of FTO variants with BMI were not replicated in African Americans (Scuteri *et al.* 2007), Chinese Hans (Li *et al.* 2007), the Japanese (Horikoshi *et al.* 2007) or Oceanic populations (Melanesians, Micronesians and Polynesians) (Ohashi *et al.* 2007). A study on more than 6 000 middle-aged Danes carrying two *FTO* obesity-risk alleles revealed that genetic susceptibility could be suppressed through physical activity (Andreasen *et al.* 2007).

The *FTO* gene was first cloned after the identification of a fused-toe mutant mouse carrying a 1.6 Mb deletion on chromosome 8. This deletion eliminates six genes, including *FTO* (Peters *et al.* 1999). In these animals, no conclusive evidence for a role of *FTO* and other deleted genes in energy homeostasis has been detected (Anselme *et al.* 2007). *FTO* is well-conserved across vertebrates and marine algae, but no clear homologue was found in invertebrate animals, fungi, plants, heterotrophic protests, bacteria or archaea (Robbens *et al.* 2007). The bioinformatics analysis showed that *FTO* encodes a 2-oxoglutarate-dependent nucleic acid demethylase, a novel member of the non-heme dioxygenase superfamily (Gerken *et al.* 2007; Sanchez-Pulido & Andrade-Navarro 2007). *FTO* transcript is abundant in the brain, particularly in the hypothalamic nuclei responsible, for instance, for energy balance (Dina *et al.* 2007a; Gerken *et al.* 2007). Furthermore, *FTO* mRNA was observed in the murine and human skeletal muscles, pancreas, heart, kidney, liver, lung and white adipose tissues (Dina *et al.* 2007a; Gerken *et al.* 2007). In mice the *FTO* mRNA levels in the arcuate nucleus are up-regulated by feeding and down-regulated by fasting (Gerken *et al.* 2007; Stratigopoulos *et al.* 2008), but opposite expression patterns were found in rats

(Fredriksson *et al.* 2008). Tschritter *et al.* (Tschritter *et al.* 2007) reported that individuals' homozygous for the risk allele have a reduced cerebrocortical insulin response in the brain. Wahlen *et al.* (Wahlen *et al.* 2007) proposed a peripheral role for *FTO* by a study on healthy women, where *FTO* mRNA levels in subcutaneous and omental adipose tissues increases with BMI and carriers of the risk allele had a reduced lipolityic activity, independent of BMI. Recently, Fischer *et al.* (Fischer *et al.* 2009) presented that a loss of the mouse *FTO* gene product reduces adiposity and even a moderate reduction in *FTO* expression protects from diet-induced obesity. They demonstrated that FTO is functionally involved in energy homeostasis by the control of energy expenditure.

The porcine *FTO* has been placed between S0297 (49.1 cM) and S0502 (58.8 cM) on SSC6 (Fontanesi *et al.* 2008). This region has been reported to contain a QTL for several different traits (Table 4).

Table 4 QTLs on chromosome 6 reported in different pig populations.

Trait	*QTL Center Location*	*QTL Span (cM)*	*Cross*	*Reference*
Protein Content	42 (cM)	41.5-62.8	Large White x Landrace x Leicoma	MOHRMANN *et al.* 2006
Lipid Content	42.0 (cM)	41.5-62.8	Large White x Landrace x Leicoma	MOHRMANN *et al.* 2006
Backfat At Shoulder	50.8 (cM)	47.1-62.8	Meishan x Pietrain x European Wild Boar	YUE *et al.* 2003
Ham Weight	51.0 (cM)	41.5-62.8	Large White x Landrace x Leicoma	MOHRMANN *et al.* 2006
Abdominal Fat	58.0 (cM)	n/a	European Wild Boar x Large White	KNOTT *et al.* 1998
Marbling	58.0 (cM)	29.2-71.4	Pietrain/Large White x commercial sow cross	VAN WIJK *et al.* 2006
PH 24 Hours Post Mortem (Ham)	59 (cM)	51.5-65	Berkshire x Yorkshire	KIM *et al.* 2005
Intramuscular Fat Content	65.0 (cM)	0-153.0	Large White x Dutch Landrace	GERBENS *et al.* 2001
Dressing Percentage	68.4 (cM)	62.8-112.0	Meishan x Pietrain x European Wild Boar	YUE *et al.* 2003
Carcass Weight (Cold)	69.6 (cM)	62.8-112.0	Meishan x Pietrain x European Wild Boar	YUE *et al.* 2003
Shoulder Meat Weight	69.6 (cM)	62.8-112.0	Meishan x Pietrain x European Wild Boar	YUE *et al.* 2003
Ham Weight	70.0 (cM)	29.2-71.4	Pietrain/Large White x commercial sow cross	VAN WIJK *et al.* 2006
Ham Meat Weight	71.0 (cM)	29.2-71.4	Pietrain/Large White x commercial sow cross	VAN WIJK *et al.* 2006
Drip Loss	74.0 (cM)	47.1-88.0	Meishan x commercial Dutch pig	DE KONING *et al.* 2001
Total Body Fat Tissue Intercept	95.3 (cM)	14.4-113.0	Duroc x Pietrain	EDWARDS *et al.* 2008
Average Daily Gain (10-22 Weeks)	113.8 (cM)	55.0-115.0	Meishan x Large White	BIDANEL *et al.* 2001
Backfat Linear At Last Rib	123.3 (cM)	21.9-134.2	Duroc x Pietrain	EDWARDS *et al.* 2008
Lipid Content	127.2 (cM)	21.9-161.3	Duroc x Pietrain	EDWARDS *et al.* 2008
Protein Content	135.2 (cM)	28.0-149.0	Duroc x Pietrain	EDWARDS *et al.* 2008

Goals and outline of the thesis

This dissertation describes the detection of quantitative trait loci in an experimental cross of two highly divergent breeds, i.e. Mangalitsa and Piétrain, by the linkage-mapping approach and characterises three candidate genes for energy and lipid metabolism in swine.

The specific goals of this thesis were:

- Identification of quantitative trait loci for carcass composition, meat quality, fatty acid composition and growth traits in the Mangalitsa x Piétrain cross using QTLexpress which assumed that any QTL is fixed for alternative alleles in the two lines and Qxpak software which allows analysis of crosses between outbred lines, i.e. when there is genetic variation between, as well as within the line.
- Detection of porcine BACs containing candidate genes for the lipid deposition traits using *in silico* BAC library screening approach
- Genomic characterization of candidate genes for the lipid deposition traits by BAC sequencing and annotation
- Screening for sequence variants
- Evaluation of relationship between sequence variants in candidate genes and the lipid-related traits by association studies in Mangalitsa x Piétrain resource population and different commercial populations of pigs.

Results will be presented in the form of three chapters. Chapter 2 describes the QTL analysis in the Mangalitsa x Piétrain resource population in terms of carcass composition, growth, meat quality and fatty acid composition traits. In chapter 3 characterisation of candidate gene *PDK4* encoding key enzyme in energy metabolism is presented. Two of the genes associated with BMI in humans (*INSIG2* and *FTO*) were investigated with regards to lipid deposition traits in swine and described in chapter 4. The results presented in this study are discussed in Chapter 5 (General discussion).

Chapter 2

Mapping of QTL with effects on lipid deposition, growth, meat quality and fatty acid composition traits in a Mangalitsa x Piétrain cross

Summary

We performed a genome-wide quantitative trait locus (QTL) scan in a Mangalitsa x Piétrain cross. In the study 167 F2 animals produced from 9 full-sib families were genotyped with 109 informative markers covering 18 autosomes. Nineteen phenotypic traits, including carcass composition, growth, meat quality and fatty acid composition, were analysed using least squares regression interval mapping. The investigation revealed 16 QTLs for carcass composition, 6 QTLs for growth, 8 QTLs for meat quality and 7 QTLs for fatty acid composition. All QTLs were significant at the 5% chromosome-wide level and for 10 a 5% genome-wide significance level was recorded. SSC4 had the largest number of QTLs (12), of which five reached a 1% genome-wide threshold. Assuming that microsatellite alleles are segregating within and in the parental lines, we observed three additional QTLs on SSC2.

Keywords: QTL, pig, carcass composition, growth, meat quality, fatty acid composition

Introduction

Localization of chromosomal regions that significantly affect the variation of quantitative traits in a population is a primary goal of quantitative trait locus (QTL) mapping. This localization is important for many traits of agricultural, biological and medical significance (Andersson *et al.* 1994). Linkage mapping in informative families or experimental crosses is a well-established approach (Andersson & Georges 2004) and relationships between genotype at various genomic locations and phenotype for a set of quantitative traits are estimated. However, the identification of genes and mutations that underlie a QTL is still problematic (Andersson & Georges 2004). Nevertheless, segregation analysis in experimental crosses provides valuable information for further genetic analyses.

The Piétrain breed was developed in Europe and is known for its leanness and muscularity, although Piétrain animals show relatively poor growth performance and meat quality. The Mangalitsa has complementary features to Piétrain, including an excellent meat quality, but accompanied by high fatness, resistance to stress and diseases, good fodder utilization and lower prolificacy. These two highly divergent breeds were used to create a Mangalitsa x Piétrain resource population and QTL mapping was performed for carcass composition, growth, meat quality and fatty acid composition traits.

Material and methods

Resource population and traits

A three generation resource population was established, in which two Mangalitsa boars were crossed with nine Piétrain sows to generate the F1 population. Four F1 male animals and 9 F1 females were mated to produce 167 offspring, taken as a subject for QTL mapping. In the F2 generation 9 full-sib families were created. Pigs were raised at an experimental station in Thalhausen (Germany). Feeding was *ad libitum* and daily feed intake data per individual was recorded using an automatic feed intake monitoring system after pigs reached approx. 30 kg in weight. Daily gain was calculated as gained weight divided by the number of days at the experimental station. Feed efficiency was calculated as the average feed intake multiplied by days at the experimental station and divided by gained weight. Male pigs were castrated. Blood samples were taken when pigs reached a weight of 40 to 50 kg and genomic DNA was extracted via proteinase K

(Roche, Penzberg, Germany). Pigs were slaughtered at a weight of approximately 95 kg at the experimental slaughterhouse in Grub (Germany). Body length, weight at slaughtering, backfat thickness, muscle depth and other carcass traits were measured post mortem in accordance with German Pig Breeders Standards (Littmann *et al.* 2006). Longissimus muscle samples were taken from the 12th and 13th rib from every pig and intramuscular fat content (IMF) was determined by fat extraction according to the method proposed by Bligh and Dyer (Bligh & Dyer 1959) and modified by Hallermeyer (Hallermayer 1976). Fatty acid composition of intramuscular fat within the longissimus muscle was determined by gas chromatography according to a standard protocol (Arens *et al.* 1994). Fatty acids were combined in 3 groups: saturated fatty acids (SFA) monounsaturated fatty acids (MUFA) and polyunsaturated fatty acids (PUFA), for which contents [%] were estimated. These values were calculated based on the following equations: SFA= C14:0 (Myristic acid) + C16:0 (Palmitic acid) + C18:0 (Stearic acid), MUFA = C16:1 (Palmitoleic acid) + C18:1 (Oleic acid), PUFA = C18:2 (Linoleic acid) + C18:3 (Linolenic acid) + C20:4 (Arachidonic acid) + C20:5 (Eicosapentaenoic acid) + C22:4 (Docosatetraenoic acid) + C22:6 (Docosahexaenoic acid), where C is an abbreviation of a carbon atom followed by the number of C atoms and the number of double bonds of the particular fatty acid separated by a colon. Phenotypic traits are listed in Table 5. All examined animals were typed for the ryanodine receptor 1 (*RYR1*) (Fujii *et al.* 1991) and alpha-1-fucosyltransferase (*FUT1*) (Meijerink *et al.* 1997) mutation. *RYR1* is a major gene affecting meat quality and carcass traits in pig and *FUT1* is a susceptibility gene for the edema diseases and post-weaning diarrhoea in piglets. Both Mangalitsa boar founders were homozygous CC and all Piétrain sows were homozygous TT for the c.1843C>T polymorphism in *RYR1*. Alleles A and G in the M307 locus of *FUT1* were fixed both within the Mangalitsa and the Piétrain breed. The effects of the *RYR1* and *FUT1* loci in our data were estimated by fitting the genotype as an additional fixed effect in a QTL analysis.

Table 5 Traits, abbreviation, number of records (N), mean, standard deviation (SD), minimum (Min) and maximum (Max) values, fixed effects and covariates included in the QTL model of phenotypic traits investigated in QTL mapping study, numbers of identified QTLs per trait (No. of QTL).

Traits name [U]	*Trait*	*N*	*Mean*	*SD*	*Min*	*Max*	Fixed effects, covariates	No. QTL
				Carcass composition				
Backfat thickness at mid-back [mm]	***BFTM***	167	28.41	5.309	16.10	40.50	sex, days, weight	0
Backfat thickness at withers [mm]	***BFTW***	167	44.77	5.612	28.20	63.20	sire, sex, days, weight	1
Backfat thickness at loin [mm]	***BFTL***	167	25.95	6.040	9.90	45.00	dam, sex, days, weight, *RYR1*	3
Average backfat thickness [mm]	***BFT*** [1]	167	33.04	4.754	20.70	45.70	dam, sex, days, weight, *FUT1*	2
Thinnest backfat thickness above middle of backmuscle between shoulder and loin [mm]	***TBFTM***	166	84.45	5.148	65.70	94.90	sex, weight	2
Backfat thickness at site [mm]	***BFTS***	167	42.07	8.729	11.10	62.70	dam, sex, days, weight, *RYR1*	1
Backfat thickness (Hennessy grading probe) [%]	***BFTH***	163	27.28	5.259	14.80	38.80	sex, days, weight,	2
Muscle thickness (Hennessy grading probe) [%]	***MUTH***	163	60.30	5.911	47.20	76.00	dam, sex, *RYR1*	1
Meat content (Hennessy grading probe) [%]	***MCOH***	163	47.09	4.862	37.40	58.90	dam, sex, days, weight, *RYR1*	1
Carcass length [cm]	***LEN***	167	884.4	27.785	805.0	996.0	dam, weight, *RYR1*	3
				Growth				
Average feed intake [g/day]	***AFI***	167	2018	201.859	1491	2651	dam, sex, days, weight, *FUT1*	2
Daily gain [g/day]	***DG***	167	554.2	78.316	352.1	852.1	grouped, weight	1
Feed efficiency [-]	***FE***	167	3.737	0.400	2.815	4.892	dam, sex	3
				Meat quality				
Intramuscular fat content [mm]	***IMF***	167	2.102	0.666	1.040	4.560	dam, sex, *RYR1*	2
Lean meat after LPA-formula	***LM***	166	54.02	3.760	45.76	68.71	sex, days, *RYR1*	1
pH 45 minutes post mortem	***PH***	163	6.117	0.379	5.240	6.920	*RYR1*	5
				Fatty acid composition				
Saturated fatty acid [%]	***SFA***	162	36.47	1.541	32.42	40.05	dam, sex, days	2
Monounsaturated fatty acid [%]	***MUFA***	163	47.51	3.024	36.98	55.07	sire, sex, *RYR1*	2
Polyunsaturated fatty acid [%]	***PUFA***	167	14.52	3.881	6.88	28.40	dam, sex, days, *RYR1*	3

[1] Calculated as the average of three measurements (BFTM, BFTW, BFTL)

Marker selection

We selected 117 markers from the USDA-MARC v2 linkage map covering 18 porcine autosomes (Rohrer *et al.* 1996). However, 8 markers were non-informative, because of missing genotypes or a non-polymorphic character of markers, and 9 markers possesed only 2 alleles. The average interval between selected markers was 19.49 centimorgan (cM), although the distance between some markers was larger. On SSC 2 there were 13 markers with an average spacing of 11 cM. On SSC 18 only one marker was successfully genotyped. The first marker on SSC 3 and 5 was located about 45 cM from the chromosome start. On SSC 7, 8, 10 and 15 a gap was found between two informative markers of 53.5 cM, 41.4 cM, 46.2 cM, 80.1 cM, respectively. On SSC 6 two gaps with a length of 61.9 and 40.3 cM were found. Markers used in the genome scan are listed in Table 6, along with the relative map position, the number of alleles segregating for each marker, polymorphic information content for the additive effect and the percentage of genotyped F2 animals for each marker. It was not attempted in this study to confirm linkage mapping already reported by many other authors.

PCR and genotyping analysis

Polymerase chain reaction was performed in a volume of 20 ul per sample. In a reaction mixture containing 50 ng genomic DNA, 1x Qiagen PCR buffer, 1.5 mM $MgCl_2$, 200 μM of each nucleotide, 5 pmol of each forward and reverse primer (MWG Biotech, Ebersberg, Germany), 0.5 units of Qiagen Taq Polymerase (Qiagen, Hilden, Germany) PCR was performed using a T-Gradient Thermocycler (Biometra, Göttingen, Germany) under the following conditions: initial denaturation at 94°C for 3 minutes; followed by 30 cycles at 94°C for 30 seconds, melting temperature for 1 minute, 72°C for 1 minute; and final extension at 72°C for 5 minutes. Melting temperature varied depending on primers, ranging between 50°C and 60°C. The 5' end of each forward primer was labeled with fluorescent dye (Fam, Hex, Tet). PCR products were separated on a 1.5% agarose gel stained with ethidium bromide together with a DNA marker (Genruler 100 bp ladder: SM0241; MBI Fermentas, St. Leon-Rot, Germany). Genotyping was performed on an ABI Prism 377 sequencer based on the size-dependent electrophoretic mobility of fluorescence-labeled PCR products in polyacrylamide gel. Analysis of the resulting data was carried out with GeneScan Analysis software, version 3.1 and the Genotyper program, version 2.5 and stored in a MySQL database. Segregation distortion of Mendelian segregation was checked manually for all markers and

corrected or removed from the data. Genotypes were obtained for the F2 animals, their F1 parents, and the purebred Mangalitsa and Piétrain grandparents.

QTL analyses

Appropriate models for all investigated traits were established by stepwise backwards model selection in the R environment for statistical computing (http://www.r-project.org/). Traits listed in Table 5 were adjusted for the fixed effects of the dam, sire, sex, *RYR1* and *FUT1* and the continuous effects of weight at slaughter and days in the experimental station. Covariates were removed from the model if the P-value was ≥ 0.05. Because of substantial skewness data on intramuscular fat content were log-transformed.

QTL analysis was performed with QTL Express (Seaton *et al.* 2002) using the F2 analysis option. Marker information contents were estimated following Knott *et al.* (Knott *et al.* 1998). The statistical model in QTL Express assumes that a putative QTL is diallelic with alternative alleles fixed in each parental line (Haley *et al.* 1994), here QQ for the Mangalitsa genotype (with effect a), qq for the Piétrain genotype (with effect -a) and Qq for F1 (with effect d). Conditional probability of every F2 individual is calculated upon the markers at 1 cM intervals along the genome. The regression model was fitted at every centimorgan along each chromosome and an F-ratio was calculated by comparing the model including the QTL effect consisting of an additive and dominance effect, hence termed FIXad, or an additive effect, termed FIXa only, vs. a reduced model without a QTL. Chromosome and genome-wide levels of significance were estimated using a permutation procedure with 1000 permutations (Churchil & Doerge 1994), while 95% confidence intervals (CI) were assessed by 1000 random bootstraps (Visscher *et al.* 1996). Individuals with missing covariates or phenotypes were eliminated from analysis.

Table 6 Markers used in QTL analysis, relative map position in cM, number of alleles, polymorphic information content for additive effect (PIC) and percentage of genotyped F2 animals (GP) for each marker.

SSC	*Marker*	*Position*	*Alleles*	*PIC*	*GP*	*SSC*	*Marker*	*Position*	*Alleles*	*PIC*	*GP*
1	SW1824	3	6	1.0551	98		SW764	156	5	0.6697	97
	SW1515	16.4	3	0.5747	91	8	SW31	3	1	0.0	99
	SW64	23.5	5	0.8199	97		SW905	20.8	5	0.2485	98
	SW1851	44.6	5	0.1924	100		S0086	62.2	3	0.7812	98
	SW781	55.8	5	0.9221	98		S0225	82.8	2	0.3292	97
	SW2185	67.6	8	1.0811	98		SW763	92.4	1	0.0	68
	S0155	93.9	6	1.0312	99	9	SW21	11.1	4	0.4729	89
	SW1311	100.8	4	1.1633	94		S0024	23.4	6	0.5295	93
	SW1301	140.5	7	0.6409	97		SW911	32.8	5	0.7946	99
2	SWR2516	0	3	0.8597	98		SW54	62.7	3	0.2831	86
	IGF2	1	3	0.8725	100		SW2571	69.3	2	0.0539	100
	SW2623	9.8	5	0.7623	100		SW944	79.3	3	0.7151	100
	S0141	31.2	4	0.3079	97		SW2093	99.6	5	0.6612	94
	SW240	42	6	0.3232	98		SW749	135.4	3	0.2284	98
	SW1201	44.8	5	0.1894	98	10	S0038	4.3	4	0.2346	95
	SW1564	55	3	0.5239	100		SWC19	50.5	3	0.3671	95
	SW1026	60.6	5	1.11	98		SW173	56.1	3	0.6044	95
	S0010	77.9	3	0.8137	98		S0070	62.3	7	1.0159	98
	S0378	80	5	0.714	97		SW1041	67.5	4	0.8007	96
	S0226	85	4	0.9905	100	11	SW2008	14.1	4	0.9871	99
	SW2192	116.2	3	0.2903	98		S0071	43.7	3	1.1212	99
	S0036	132.1	5	0.7886	91		SW435	53.3	2	0.6615	98
3	S0206	42.3	7	0.9221	96		SW703	76.2	3	0.832	98
	SW2618	50.8	5	0.2783	99	12	SW2490	0	5	0.982	98
	SW902	58.4	4	0.1896	99		S0229	19.3	6	1.0475	96
	S0164	60.5	8	0.9426	94		SW957	33.4	6	0.4757	98
	SW2570	72.3	6	0.8133	97		SW874	64.7	4	0.3677	95
	SW2408	94.2	5	0.7942	95		S0090	80.2	4	0.9643	96
	S0002	102.2	6	1.0117	98		S0106	95.8	4	0.0239	11
	SW1327	109.6	6	0.9637	94		SWR1021	113.1	5	0.2552	88
	SW349	112.6	4	0.8876	100	13	S0219	1.6	2	0.5296	99
4	SW489	8	5	0.5269	99		SW935	25.4	5	0.5625	99
	S0001	41.8	5	1.0298	98		SW864	43.1	4	0.3737	99
	SW839	62.3	3	0.9773	100		S0068	62.2	4	0.6704	96
	S0217	69.6	2	0.1222	96		SW225	70.1	4	0.44	74
	SW524	99.3	2	0.3001	96		SW398	79.3	6	0.9163	98
	S0097	120	6	0.862	96		SW2440	102	4	0.3319	98
5	SW2	47.2	7	0.6766	92	14	SW857	7.4	4	0.5205	98
	S0005	56.7	8	0.8819	97		SW295	35.7	5	0.8323	95
	SW1200	82.8	2	0.0914	98		SW210	46.3	4	0.8169	95
	IGF1	87.2	5	1.0173	88		S0007	60	7	0.8726	95
	SW995	93.5	4	0.8242	99		SW77	69.4	7	0.9455	98
	SW378	102.9	3	0.7513	94		SW1557	87.9	4	0.933	95
6	S0099	0	5	0.9749	80		SWC27	111.5	2	0.0	100
	SW2406	21.4	6	0.7447	83	15	S0355	1.3	3	0.6406	98
	SW1067	71.4	0	0.0	0		SW1111	27.3	6	0.9819	95
	SW122	83.3	5	0.8732	94		SW1983	89	0	0.0	0
	SW316	89.3	0	0.0	0		SW1119	107.4	8	0.7638	99
	S0003	102	5	1.0179	98	16	S0111	0	6	0.7527	96
	SW1881	121.1	6	0.8293	98		SW742	9.3	6	0.7865	97
	SW2419	161.4	6	0.5239	97		S0298	33.2	3	0.2509	100
7	SW2564	0	3	0.2543	98		SW2517	55.7	6	0.9665	98
	SWR1343	12.2	5	0.5922	94		S0061	92.6	3	0.6001	99
	SW1354	22.3	4	0.5935	96	17	SW335	0	4	0.9922	96
	SW472	58.9	2	0.012	99		SWR1004	17.8	4	1.0652	98
	SW175	81.5	5	0.7686	97		SW2441	40.6	1	0.0	82
	S0066	82.8	6	0.838	93	18	SW1801	0	6	0.0262	13
	SW352	87.7	4	0.6086	94		SW1984	29.4	4	0.9853	97
	S0212	141.2	4	1.0774	90						

The F2 analysis module in QTL Express assumes that any QTL is fixed for alternative alleles in the two founder lines. However, this is often not the case in outbred populations, which reduces the power of standard analysis packages such as QTL Express (de Koning *et al.* 2000; Wang *et al.* 2006). Therefore, we applied an additional analysis using Qxpak (Perez-Enciso & Misztal 2004), which is based on mixed model methodology and maximum-likelihood estimation. The software calculates identity by descent probabilities using all animals (P, F1 and F2) in the pedigree simultaneously and thus makes it possible to model QTL alleles to be segregating also within lines as well as to include an infinitesimal genetic effect. Likelihood ratio tests were used for model selection comparing two hierarchical models at a time. In the first set of tests fixation of QTL alleles within lines was assumed: FIXad was compared with a null model without a QTL. If significant ($P<0.05$), FIXad was tested against FIXa. In the second set of tests models, that assumed segregation of QTL alleles within lines, were evaluated, hence termed MIXad and MIXa: depending on the most parsimonious model from the first set of tests MIXad was tested against FIXad, or MIXa versus FIXa.

Results and discussion

The investigation of 18 porcine autosomes revealed 16 QTLs for carcass composition, 6 QTLs for growth, 8 QTLs for meat quality and 7 QTLs for fatty acid composition. All QTLs were significant at the 5% chromosome-wide level and 10 achieved a 5% genome-wide significance level. QTLs for carcass composition traits were identified on chromosomes 4, 10, 11, 14, 15 and 17, for growth on chromosomes 1, 4, 7, 12, for meat quality on chromosomes 1, 4, 6, 10, 16 and 18, and for fatty acid composition on chromosomes 1, 5, 13, 16 and 17, respectively. On chromosomes 2, 3, 8 and 9 no QTLs for traits related to carcass composition, growth, meat quality or fatty acid composition were found with QTL Express. Up to five QTLs were detected for 18 of 19 traits, while no QTLs were identified for backfat thickness in the middle of the back. Results of QTL mapping are summarized in Table 7.

Table 7 QTLs significant at the 5% chromosome-wide level according to QTL Express analysis.

SSC[1]	Trait[2]	Abb[3]	QTL (cM)[4]	95% CI (cM)[5]	Flanking markers[6]	Additive (SE)[7]	Dominance (SE)[8]	LR[9]	F-ratio[10]		Var (%)[11]
					Carcass composition						
4	Average backfat thickness	**BFT**	55	31-108.5	S0001-SW839	2.30 (0.47)	-0.30 (0.70)	22.26	12.00	***	9.64
4	Backfat thickness (Hennessy)	**BFTH**	56	36.5-80	S0001-SW839	2.81 (0.57)	-0.60 (0.87)	22.16	11.91	***	10.72
4	Backfat thickness at loin	**BFTL**	57	32-116.5	S0001-SW839	2.63 (0.58)	-0.04 (0.83)	19.68	10.52	***	8.14
4	Backfat thickness at site	**BFTS**	59	43-83	S0001-SW839	5.01 (0.86)	-3.78 (1.28)	33.94	19.6	***	13.29
4	Backfat thickness at withers	**BFTW**	53	31-92.5	S0001-SW839	2.75 (0.63)	-1.00 (0.94)	18.35	9.74	**	9.18
4	Carcass length	**LEN**	46	20-70	S0001-SW839	-13.63 (2.62)	3.84 (4.00)	25.2	13.60	***	9.79
4	Meat content (Hennessy)	**MCOH**	62	41.5-98	SW839-S0217	-2.27 (0.50)	0.35 (0.74)	19.54	10.46	**	8.01
4	Thinnest backfat thickness	**TBFTM**	35	8-99	SW489-S0001	1.92 (0.56)	-0.55 (0.87)	11.47	5.94		5.93
10	Thinnest backfat thickness	**TBFTM**	42	14.3-64.3	S0038-SWC19	2.68 (0.82)	2.39 (1.67)	10.38	5.36		5.38
11	Muscle thickness (Hennessy)	**MUTH**	67	38.6-76.1	SW435-SW703	-1.82 (0.61)	1.89 (0.90)	13.28	6.94	*	5.81
14	Backfat thickness at loin	**BFTL**	17	7.4-78.4	SW857-SW295	0.39 (0.72)	3.88 (1.14)	11.18	5.81		4.76
15	Average backfat thickness	**BFT**	30	13.3-107.3	SW1111-SW1983	1.26 (0.53)	2.34 (0.81)	11.46	5.96		5.14
15	Backfat thickness at loin	**BFTL**	35	13.3-107.3	SW1111-SW1983	1.34 (0.64)	2.56 (0.92)	10.11	5.23		4.32
15	Carcass length	**LEN**	83	17.3-107.3	SW1111-SW1983	-15.16 (4.07)	-1.05 (8.18)	13.49	7.6*		5.48
17	Backfat thickness (Hennessy)	**BFTH**	6	0-34	SW335-SWR1004	-1.82 (0.55)	-0.40 (0.75)	10.85	5.62		5.44
17	Carcass length	**LEN**	18	0-40	SWR1004-SW2441	10.18 (2.58)	-4.34 (3.99)	15.97	8.42	**	6.43
					Growth						
1	Daily gain	**DG**	93	-	SW2185-S0155	-22.43 (7.93)	-34.54 (12.12)	13.37	7.11	*	5.14
1	Feed efficiency	**FE**	53	10-134	SW1851-SW781	0.13 (0.05)	-0.28 (0.09)	14.9	7.38		6.52
4	Average feed intake	**AFI**	54	18-92.5	S0001-SW839	44.41 (13.18)	5.46 (19.98)	11.49	5.97		1.84
4	Feed efficiency	**FE**	77	35.5-120	S0217-SW524	0.22 (0.05)	-0.12 (0.10)	16.75	8.85	**	7.69
7	Average feed intake	**AFI**	84	34.5-156	S0066-SW352	-17.39 (16.78)	107.19 (32.50)	10.84	5.62		1.74
12	Feed efficiency	**FE**	36	0-82.5	SW957-SW874	0.22 (0.06)	0.15 (0.11)	12.81	6.68		5.95
					Meat quality						
1	Intramuscular fat content	**IMF**	76	16-140	S0331-S0155	0.04 (0.01)	-0.02 (0.02)	10.95	5.68		4.43
1	pH 45 minutes post mortem	**PH**	22	3-84.5	SW1515-SW64	0.09 (0.03)	-0.07 (0.05)	12.10	6.29		3.92
4	Intramuscular fat content	**IMF**	75	8-118	S0217-SW524	-0.03 (0.02)	-0.07 (0.03)	10.66	5.52		4.06
4	Lean meat after LPA-formula	**LM**	76	42-99	S0217-SW524	-1.94 (0.54)	1.21 (1.07)	13.96	7.30	*	6.77
6	pH 45 minutes post mortem	**PH**	82	12.5-121	SW1067-SW122	0.19 (0.05)	-0.01 (0.07)	17.37	9.19	**	5.55
10	pH 45 minutes post mortem	**PH**	52	17.8-63.3	SWC19-SW173	-0.11 (0.03)	-0.04 (0.05)	10.90	5.65		3.57

16	pH 45 minutes post mortem	**PH**	59	0-92	SW2517-S0061	-0.10 (0.03)	-0.09 (0.05)	12.82	6.68	*	4.13
18	pH 45 minutes post mortem	**PH**	2	0-29	SW1808-SW1984	0.13 (0.05)	-0.23 (0.12)	9.28	4.78	*	3.06
					Fatty acid composition						
1	Polyunsaturated fatty acid	**PUFA**	47	7-100	SW1851-SW781	-1.39 (0.38)	-0.18 (0.54)	12.69	6.99		2.72
5	Saturated fatty acid	**SFA**	67	47.2-100.2	S0005-SW1200	0.01 (0.35)	-1.71 (0.54)	9.56	5.13	*	11.59
13	Monounsaturated fatty acid	**MUFA**	86	25.1-101.6	SW398-SW2440	1.00 (0.68)	-4.83 (1.38)	11.69	6.35		8.43
13	Polyunsaturated fatty acid	**PUFA**	73	25.6-101.6	SW225-SW398	-0.46 (0.47)	2.88 (0.88)	10.10	5.45		2.21
16	Monounsaturated fatty acid	**MUFA**	73	0-92	SW2517-S0061	-2.61 (0.77)	-2.02 (1.40)	11.72	6.37		8.45
16	Polyunsaturated fatty acid	**PUFA**	60	0-88.5	SW2517-S0061	1.79 (0.56)	0.23 (0.99)	9.64	5.19		2.12
17	Saturated fatty acid	**SFA**	22	0-40	SWR1004-SW2441	0.06 (0.45)	-2.73 (0.95)	7.82	4.14		9.59

[1] *Sus scrofa* chromosome
[2] Trait
[3] Trait abbreviation
[4] Position of QTL in cM
[5] 95% confidence intervals (CI) (Visscher *et al.* 1996) estimated by bootstrapping approach with 1000 replicates in QTL Express
[6] Flanking markers were those markers around the peak, as near as possible
[7 8] Estimated additive and dominance effects and their standard errors (SE)
[9] Likelihood ratio - the maximum probability of a result
[10] F-ratio, QTL significant at 5% chromosome-wide, * 1% chromosome-wide, ** 5% genome-wide, *** 1% genome-wide levels
[11] Percentage of the residual variance in the F2 population explained by the QTL

QTLs for carcass composition

A series of QTLs affecting several backfat thickness traits were found on SSC 4. These QTLs explained between 5.93 and 13.29% of phenotypic variance in the F2 animals; however, some backfat traits were highly correlated. Previously described QTLs for abdominal fat and backfat reported by Andersson *et al*. (Andersson *et al.* 1994) are in agreement with the QTL region for fatness detected in the Mangalitsa x Piétrain cross. The *FAT1* locus reported by Marklund *et al*. (Marklund *et al.* 1999) overlapped with the most significant region from our study. Knott *et al*. (Knott *et al.* 1998) described a QTL with additive effects for average backfat thickness consistent with our results, showing a similar profile of the QTL plot for average backfat thickness (BFT) as ours. A significant QTL for backfat thickness in this region reported by Wimmers *et al*. (Wimmers *et al.* 2002) gives additional support to our study. A QTL significant at the 5% genome-wide level for backfat thickness measured by the Hennessy grading probe (BFTH) was found within CI between 36.5 and 80 cM and overlapped with the QTL reported by de Koning *et al*. (de Koning *et al.* 2001). Our studies indicated an additional QTL for meat content (MCOH), backfat thickness at loin (BFTL), withers (BFTW) and site (BFTS), feed intake (AFI) in the *FAT1* region. In the central region of SSC 4, QTLs for carcass composition were repeatedly described in various intercrosses of different porcine breeds (Wang *et al.* 1998; Walling *et al.* 2000; Bidanel *et al.* 2001) and in a commercial population (Evans *et al.* 2003).

Several attempts were made to identify a gene(s) underlying the *FAT1* locus. A large QTL effect may be caused by mutations at several loci, each with a small effect (Marklund *et al.* 1999) or by a major causative gene with a causative mutation. Murani *et al*. (Murani *et al.* 2006a) proposed the corticotropin releasing hormone (*CRH)* gene as a positional and functional candidate gene for body composition and growth on porcine chromosome 4. Subsequently, an association between *CRH* with feed conversion ratio, carcass length, lean content and meat colour was detected in commercial pig lines (Murani *et al.* 2006b). Estelle *et al*. (Estelle *et al.* 2006) described an association between the fatty acid binding protein 5 (*FABP5)* gene with *FAT1* QTL and suggested more than one loci affecting fatness in this region. Szczerbal *et al*. (Szczerbal *et al.* 2007) proposed that a cluster of *FABP* genes is beyond the *FAT1* region; however, according to our study supported by Marklund *et al*. (Marklund *et al.* 1999) the cluster of *FABP* is located in the potential QTL region. Additionally, *FABP5*

was mapped on HAS 8, which corresponds to the region surrounding the porcine centromere and p arm of SSC 4 (Moller *et al.* 2004). The apolipoprotein A-II *(APOA2)* gene assigned to the QTL fatness region of SSC 4 was proposed by Knoll *et al*. (Knoll *et al.* 2003) as a candidate gene for some carcass and growth traits. However, according to a high-resolution comparative map of SSC 4 (Moller *et al.* 2004) *APOA2* is located more distally to the fatness region reported by Andersson *et al*. (Andersson *et al.* 1994) and is consistent with the results in our study. Nevertheless, despite numerous efforts the causative mutation responsible for *FAT1* QTL was not identified. Further studies considering a potential candidate gene should be conducted. Access to genome sequences and re-sequencing of the QTL interval will be an attractive approach to detect causative mutation(s) ascribed to *FAT1* effects (Andersson & Georges 2004).

On SSC 10 suggestive QTLs were detected for thinnest backfat thickness above middle of the backmuscle between the shoulder and loin (TBFTM) within the CI between 14 and 64 cM. The CI partially overlapped with a more distal QTL for ultrasound fat depth reported by Rohrer *et al.* (Rohrer *et al.* 2005). Kim *et al.* (Kim *et al.* 2006) reported significantly different backfat thickness at the S0038 marker in the proximal region of SSC 10. A 1% chromosome-wide QTL for muscle thickness measured by the Hennessy grading probe (MUTH) was detected on SSC 11 and Piétrain alleles caused increased muscle thickness. In the proximal region of SSC 14 a 5% chromosome-wide QTL was identified for BFTL with a dominant effect. The wide CI overlapped with a QTL for last-lumbar and average backfat reported by Rohrer *et al.* (Rohrer & Keele 1998a); however, the most significant peak was located more distally than in our study. Within the same marker interval (SW857-SW295) de Koning *et al*. (de Koning *et al.* 2001) detected a suggestive QTL for ultrasonic backfat thickness with additive effects.

On SSC 15 5% chromosome-wide QTLs for two backfat thickness traits (BFT, BFTL) were identified in marker interval SW1111-SW1983 and they correspond to QTLs presented by Kuryl *et al.* (Kuryl *et al.* 2003) in Wild Boar x Piétrain and Wild Boar x Meishan families.

In the proximal region of SSC 17, a suggestive QTL with additive effects for BFTH was detected. Interestingly, Piétrain alleles were associated with higher backfat thickness. Pierzchala *et al.* (Pierzchala *et al.* 2003a) described a QTL for fatness more distally on

SSC17 in a Meishan x Piétrain resource population.

QTLs for carcass length achieved a significant 5% genome-wide level on SSC 4 and 17, a 1% chromosome-wide level on SSC 15 and jointly explained 21.7% of the phenotypic variance in the F2 population. All QTLs exhibited an additive effect. On SCC 4 and 15, Piétrain QTL alleles were associated with a longer carcass, while on SSC 17 Piétrain QTL alleles were associated with a shorter carcass. The CI for a QTL on SSC 4 overlapped with a QTL reported by Rohrer and Keele (Rohrer & Keele 1998b) in a Meishan x White composite pigs and by Wimmers *et al.* (Wimmers *et al.* 2002) in a Berlin Miniature Pig x Duroc population. The QTL on SSC 17 was detected between 0 and 40 cM, which is in close proximity to the 1% genome-wide significant QTL described previously by Karlskov-Mortensen *et al.* (Karlskov-Mortensen *et al.* 2006). This QTL is not identical to the QTL detected in our study, because marker SW2441 in the Mangalitsa x Piétrain cross was non-informative and more distal markers were not genotyped. The 1% chromosome-wide QTL on SSC 15 was not reported in previous studies.

QTLs for growth

Three QTLs for feed efficiency (FE) were detected on SSC 1, 4 and 12, which jointly explained 20.16% of phenotypic variance in the F2 animals. QTLs on SSC 4 and 12 exhibited additive effects and Mangalitsa alleles were associated with higher FE when compared with Piétrain alleles.

The CI for the 5% genome-wide significant QTL for FE on SSC 4 overlapped with the fatness QTL found in our study, but the highest peak was located more distally. These results confirmed the report of Andersson *et al.* (Andersson *et al.* 1994), who described a fatness region more proximally compared with the region responsible for growth traits.

A QTL significant at the 1% chromosome-wide level was found on SSC 1 for daily gain (DG) and was bracketed by markers SW2185-S0155. In the close proximity QTLs for DG in the weight ranges from 30 to 60 kg and 90 to 120 kg were reported by Mohrmann *et al.* (Mohrmann *et al.* 2006). Piétrain alleles tended to be associated with the lower daily gain, which was consistent with our study. The melanocortin 4 receptor

MC4R gene located in this region shows an association of the Asp298Asn Piterain-derived polymorphism with DG in the Mangalitsa x Piétrain cross (Meidtner *et al.* 2006).

Two suggestive QTLs for average feed intake (AFI) were detected on SSC 4 and 7, which jointly explained 3.58% of phenotypic variance in the F2 population. QTLs on SSC 4 located in the fatness region exhibited additive effects, in contrast to QTLs on SSC 7, which exhibited strong overdominance effects. In both cases Mangalitsa alleles tended to be associated with a higher AFI. Cepica *et al.* (Cepica *et al.* 2003b) described a QTL significant at the 5% genome-wide level on SSC 4 for food consumption in a Meishan x Piétrain population.

QTLs for meat quality

Suggestive evidence for QTLs affecting intramuscular fat content (IMF) was found on SCC 1 and 4. The CI for QTLs on SSC 4 was between flanking markers S0217-SW524, which is in close proximity to the suggestive QTL region described by de Koning *et al.* (de Koning *et al.* 1999) and Rattink *et al.* (Rattink *et al.* 2000) in Meishan x Dutch pigs. The QTLs significant at the 5% chromosome-wide level on SSC 1 have not been reported previously in literature.

Five QTLs for pH were obtained on SSC 1, 6, 10, 16, and 18, which jointly explained 20.23% of phenotypic variance in the F2 population. The QTLs on SSC 6, despite including the *RYR1* effect in the model, still show significance at the 5% genome-wide level centred on the *RYR1* locus. This could indicate that other loci close to *RYR1* or an unknown *RYR1* allele contribute to the variation in the trait. Without adjustment for *RYR1* genotypes, in the vicinity of the *RYR1* gene (83 cM), a QTL analysis showed a highly significant QTL for pH with an F-ratio of 57.91, which explained 40.58% of phenotypic variance. The position of this QTL corresponds with the QTL for pH reported by Yue *et al.* (Yue *et al.* 2003) in Meishan x Piétrain and Wild Boar x Piétrain crosses. However, we did not detect any additional effect on meat quality (Yue *et al.* 2003), backfat thickness (de Koning *et al.* 1999; Bidanel *et al.* 2001; Malek *et al.* 2001), carcass length (Malek *et al.* 2001) proximally and distally to the *RYR1* locus on SSC6, when phenotypic data was adjusted for the *RYR1* genotypes. The strong QTL for MUTH at 78 cM (F-ratio 8.23, 1% chromosome-wide threshold 7.15), lean meat after

the LPA-formula (LM) at 79 cM (F-ratio 8.04, 1% chromosome-wide threshold 7.05), carcass length (LEN) at 84 cM (F-ratio 5.27, 5% chromosome-wide threshold 4.95) in the immediate vicinity of the *RYR1* gene were detected without *RYR1* adjustment in the models.

A suggestive QTL for pH between markers SW1515-SW64 was detected on SSC 1, which confirmed findings by Ponsuksili *et al.* (Ponsuksili *et al.* 2005). On SSC 16, a QTL for pH significant at the 5% chromosome-wide level was found at position 61 cM. Pierzchala *et al.* (Pierzchala *et al.* 2003b) described a QTL for pH45min in *Musculus longissimus dorsi* in a more proximal region of SSC 16. Edwards *et al.* (Edwards *et al.* 2007) reported a suggestive QTL for 24-h pH in the proximal part of SSC 18, which was consistent with our QTL for 45-min pH. Nevertheless, because only one marker was informative on SSC 18, we cannot exclude the possibility that results from our study for this chromosome show low precision. In the present study a novel suggestive QTL was identified on SSC 10.

A QTL for LM significant at the 1% chromosome-wide level was found on SSC 4 with Piétrain alleles associated with higher lean meat compared with Mangalitsa alleles. The QTL overlapped with the 5% genome-wide significant QTL for bone/lean meat in ham reported by Andersson-Eklund *et al.* (Andersson-Eklund *et al.* 1998).

QTLs for fatty acid composition

Two suggestive QTLs with the overdominance effect for SFA were found on SSC 5 and 17. Piétrain alleles were associated with an increasing level of SFA in intramuscular fat content compared to Mangalitsa alleles. Nii *et al.* (Nii *et al.* 2006) reported QTLs for SFA significant at the 5% genome-wide level on SSC5 more proximally when compared with our results. Interestingly, Wild Boar alleles in the Nii *et al.* study, alike Mangalitsa alleles in our study, decrease SFA.

QTLs for MUFA were located on SSC13 and 16 and jointly explained 16.88% of phenotypic variance. Three 5% chromosome-wide QTLs for PUFA, which jointly explained 7.05% of phenotypic variance, were located on SSC 1, 13 and 16; however, for the QTLs on SSC13 and 16 the significant threshold was exceeded minimally.

Multiple QTLs on SSC4

Prominent effects were found for fatness, meat quality and growth on SSC4 in the Mangalitsa x Piétrain cross. Chromosome 4 had the largest number of QTLs (12), of which 5 QTLs reached the 1% genome-wide threshold and 3 QTLs reached the 5% genome-wide threshold. The greatest impact of QTLs on BFTS was in the S0001-SW839 interval. Within the same markers interval, the 1% genome-wide significant QTL for several correlated fatness traits (BFT, BFTH, BFTL) and LEN, the 5% genome-wide significant QTL for BFTW and the 5% chromosome-wide significant QTL for AFI were mapped. A more distal 5% chromosome-wide significant QTL for meat quality traits (IMF, LM) and the 5% genome-wide significant QTL for MCOH and FE were detected. QTLs for thinnest backfat thickness (TBFTM) were found more proximally to the most significant region, but CI encompassed the region responsible for fatness traits. At the SW839 marker locus the Mangalitsa allele increases the amount of fat measured as backfat thickness at site, at loin, at withers, and ultrasound backfat thickness (BFTH), while it increases AFI. The MCOH, LM and LEN were increased by Piétrain alleles. The QTLs confirmed the results described previously (Andersson *et al.* 1994; Knott *et al.* 1998; Rohrer & Keele 1998b; Wang *et al.* 1998; Marklund *et al.* 1999; Rattink *et al.* 2000; Walling *et al.* 2000; Bidanel *et al.* 2001; de Koning *et al.* 2001; Wimmers *et al.* 2002; Edwards *et al.* 2008) and were discussed widely beforehand.

QTLs with qxpak

The model of analysis in QTL Express assumes that alternative QTL alleles are fixed in the parental lines. However, in the Mangalitsa x Piétrain cross alleles at the QTLs are still segregating in either or both of the parental lines, thus the power of QTL detection will be highly reduced, and its effect will be underestimated. Therefore, we carried out an additional analysis where segregation of alleles within parental lines was taken into account. Only three additional QTLs were detected on SSC 2 between S0001-SW839 markers (Table 8).

Table 8 Chromosome-wide and genome-wide significant QTL according to qxpak analysis.

SSC[1]	Trait[2]	Abb[3]	Position[4]	Flanking markers[5]	df[6]	Add (SE)[7]	Dom (SE)[8]	LR[9]	Nominal P-value[10]	Chromosome significant P-value[11]	Genome significant P-value[12]
				Carcass composition							
2	Average backfat thickness	**BFT**	46	S0001-SW839	1	2.99 (0.55)	-	27.8	1.95e-07	<0.001	0.003
2	Backfat thickness (Hennessy)	**BFTH**	53	S0001-SW839	1	3.60 (0.69)	-	25.1	5.70e-07	<0.001	0.002
2	Backfat thickness at loin	**BFTL**	45	S0001-SW839	1	3.41 (0.68)	-	23.30	1.39e-06	0.002	0.013

[1] *Sus scrofa* chromosome
[2] Trait
[3] Trait abbreviation
[4] Position of QTL in cM
[5] Flanking markers
[6] Degree of freedom
[7 8] Estimated additive and dominance effects and their standard errors (SE)
[9] Likelihood ratio - the maximum probability of a result
[10] Nominal P-value
[11] Chromosome significant P-value after 1000 permutation test (Churchil & Doerge 1994)
[12] Genome significant P-value after 1000 permutation test (Churchil & Doerge 1994)

Chapter 3

Characterisation of the porcine PDK4 gene

Summary

The pyruvate dehydrogenase kinase, isozyme 4 (PDK4) is involved in three important energy pathways: glycolysis, fatty acid oxidation and synthesis; and its gene (*PDK4)* is a functional candidate gene for lipid deposition and energy homeostasis-related traits. Additionally, *PDK4* is located in a QTL region for backfat thickness, feed intake, daily gain and saturated fatty acid content. We have annotated the porcine *PDK4* gene and have re-sequenced it in the parental generation of a Mangalitsa x Piétrain cross. We have identified 24 variants, including a non-synonymous mutation in exon 10, a promoter substitution within the estrogen-related receptor alpha (ERRα) factor binding site and an insertion-deletion polymorphism of short interspersed element (SINE) in intron 9. These three potentially functional variants were subjected to association studies. A significant association is observed between the non-synonymous substitution and saturated fatty acid (SFA) content (P-value=0.046) and a suggestive association with backfat thickness at loin (P-value=0.083). The insertion-deletion polymorphism of SINE was found to be associated with SFA content (P-value=0.039) and average feed intake (P-value=0.011). The promoter substitution shows a suggestive association with daily gain (P-value=0.100). In the Piétrain population the non-synonymous mutation is associated with the estimated breeding value of IMF (P-value=0.030). In the German Landrace breed the association of average feed intake with the insertion-deletion polymorphism of SINE could not be replicated (P-value=0.381).

Keywords: *PDK4*, pig, gene structure, SINE, association study

Introduction

The pyruvate dehydrogenase kinase, isozyme 4 (PDK4) is a key regulatory enzyme controlling the activity of the mitochondrial puryvate dehydrogenase complex, which catalyzes the conversion of pyruvate to acetyl-CoA. Thus, PDK4 is involved in the control of three important energy pathways: glycolysis, fatty acid oxidation and synthesis. Additionally, *PDK4* has been mapped to a region on porcine chromosome 9q12-(1/3)q21 (Davoli *et al.* 2000; Maak *et al.* 2001), which harbours quantitative traits loci (QTLs) for backfat thickness, daily gain (Kim *et al.* 2006; Edwards *et al.* 2008), feed intake (Cepica *et al.* 2003a) and saturated fatty acid content (Nii *et al.* 2006). Therefore, the purpose of this study was to explore genetic variations of porcine *PDK4* in a Mangalitsa x Piétrain cross as a functional and positional candidate gene for lipid deposition and energy homeostasis-related traits in swine.

Material and methods

Animals

We studied 524 animals from the F2 generation of the Mangalitsa x Piétrain resource population. The pigs were fed *ad libitum* up to approximately 95 kg. Feed intake per individual was obtained by an automatic feed intake monitoring system. Male pigs were castrated. Backfat thickness, muscle depth and other traits were measured post mortem (Table 9) according to German Pig Breeders Standards (Littmann *et al.* 2006). The longissimus muscle sample was taken from the 12th and 13th rib from every pig and measurement of intramuscular fat content was determined by fat extraction according to the method proposed by Bligh and Dyer (Bligh & Dyer 1959) and modified by Hallermeyer (Hallermayer 1976). Both Mangalitsa boar founders were homozygous CC and all Piétrain sows were homozygous TT for c.1843C>T polymorphism in the ryanodine receptor 1 gene *(RYR1)*. Thus, all investigated animals were typed for the c.1843C>T *RYR1* polymorphism.

Samples of the confirmation study were provided by the Landesanstalt für Landwirtschaft Bayern, Grub (Germany) and contained 504 German Landrace animals from the performance testing that were slaughtered in the period 2002 – 2005 in Grub (Germany) and Schwarzenau (Germany), and 1013 Piétrain animals from the performance testing that were slaughtered in the years 2004 – 2006 in Grub (Germany) and Schwarzenau (Germany) (Table 9). German Landrace, German Large White and

Piétrain DNA samples for the estimation of allele frequencies were collected from the artificial insemination centres in Bergheim (Germany), Landshut (Germany) and Neustadt/Aisch (Germany) from 2000 to 2002.

Table 9 Means and standard deviations (SD) of investigated phenotypes in the F2 (n=524), Piétrain (n=1013) and German Landrace (n=504) populations.

Trait	Abb.	Mean	SD
F2 animals			
Average backfat thickness [mm]	BFT[1]	33.43	5.22
Backfat thickness at loin [mm]	BFTL	26.26	6.34
Intramuscular fat content [%]	IMF	2.05	0.63
Saturated fatty acid [%]	SFA[2]	36.60	2.61
Average feed intake [kg/day]	AFI	2.03	0.21
Daily gain [kg/day]	DG	0.56	0.08
Piétrain			
Average backfat thickness [mm]	BFT	0.57	0.17
Backfat thickness at loin [mm]	BFTL	0.79	0.24
Intramuscular fat	IMF	0.01	0.16
Daily gain [kg/day]	DG	0.76	0.08
German Landrace			
Average feed intake [kg/day]	AFI	2.32	0.24

[1] The average of three measurements: backfat thickness at shoulders, backfat thickness at mid-back, backfat thickness at loin

[2] SFA = C14:0 (Myristic acid) + C16:0 (Palmitic acid) + C18:0 (Stearic acid)

Five animals carrying two longer alleles and ten homozygous animals of the shorter allele of c.967-139_967-138ins183,GenBank FJ853993 the SINE polymorphism in *PDK4* were selected for the gene expression experiment. All animals were descended from the F3 generation of the Mangalitsa x Piétrain cross with a mean weight at slaughter of 88.93 kg. Homozygous animals with the longer allele came from one litter, homozygous animals with the shorter allele - from four litters. Tissue samples (approximately 0.5 g) for RNA isolation from the longissimus muscle and backfat were collected and stored in 5 ml of RNAlater™ (Qiagen, Hilden, Germany).

Prediction of the genomic organization of *PDK4*

The full region of porcine *PDK4* was obtained using an internet-based *in silico* BAC library (http://www.sanger.ac.uk/Projects/S_scrofa/BES.shtml). The exon-intron structure was identified using the GenomeThreader tool (Gremme *et al.* 2005) and Genome Annotation and Curation Tool Apollo (Lewis *et al.* 2002).

Re-sequencing

Primers were designed based on the porcine genomic *PDK4* sequence using the Primer3 software (Rozen & Skaletsky 2000) and purchased from MWG (Ebersberg, Germany) (Table 10). A PCR (20 µl) sample contained 50 ng genomic DNA, 1x Qiagen PCR buffer, 1.5 mM $MgCl_2$, 200 µM of each nucleotide, 5 pmol of each forward and reverse primer (MWG Biotech, Ebersberg, Germany), 0.5 units of Qiagen Taq Polymerase (Qiagen, Hilden, Germany) was performed using a T-Gradient Thermocycler (Biometra, Göttingen, Germany) under the following conditions: initial denaturing at 94°C for 3 minutes; followed by 30 cycles at 94°C for 30 s, melting temperature for 1 min, 72°C for 1 min; and the final extension at 72°C for 5 min. Melting temperature varied depending on the primer and ranged between 56°C and 60°C. PCR products were separated on a 1.5% agarose gel stained with ethidium bromide together with a DNA marker (Genruler 100 bp ladder: SM0241; MBI Fermentas, St. Leon-Rot, Germany). PCR products were purified using MultiScreen PCRµ96 Filter Plates (Millipore, Schwalbach, Germany) in combination with a Millipore vacuum manifold. The sequencing reaction was performed using BigDye Terminator v1.1 Sequencing Kit (Applied Biosystems Applera, Darmstadt, Germany) under the following conditions: initial denaturation for 15 seconds at 96°C; followed by 35 cycles at 96°C for 10 s, 51°C for 5 s and 60°C for 4 min, and purified with Sephadex columns in a MultiScreen 96 well filtration plate (Millipore, Schwalbach, Germany). Sequencing was performed on an ABI Prism 377 automated sequencer (Applied Biosystems Applera, Darmstadt, Germany), the obtained sequences were analysed using the Phred/Phrap/Polyphred/Consed software suite (Nickerson *et al.* 1997; Ewing & Green 1998; Ewing *et al.* 1998) and the results were viewed using Consed (Gordon *et al.* 1998).

Genotyping

Genotyping of c.1049G>A in the F2 animals was performed using restriction fragment length polymorphism (RFLP). Amplified fragments were digested with *Mva*I according to the manufacturer's protocol. The restriction enzyme was selected with the help of the Hypertext DNA Sequence Display (http://pga.mgh.harvard.edu/web_apps/web_map/start) and ordered from Ferments (St.Leon-Rot, Germany). For the short interspersed element polymorphism, PCR amplification was followed by separation of the products on a 1.5 % agarose gel.

Genotypes of c.1049G>A in Piétrain and c.-560T>C in the F2 and Piétrain animals were determined by TaqMan genotyping assays (Table 10) according to manufacturer's protocols.

Table 10 Primers used for re-sequencing, genotyping and expression study of porcine *PDK4*.

Locus	Primer	Pair	Sequence	Region	Product size	Tm °C
PDK4	4416	4417	GGGAGGGGAAAATAGTGAGTC	5'UTR	678	56
PDK4	4417	4416	CGCCTCCCAAGGTTTTATTT			
PDK4	4744	4745	CCCACCCCTTTATCCCTTAC	5'UTR	496	58
PDK4	4745	4744	GAGAAATGCTCGACCTCTCG			
PDK4	4418	4419	GACTCCTCCCCTTTGGAGTT	Exon 1	592	56
PDK4	4419	4418	AAGCAAATCTAACAAAAGGCAAA			
PDK4	4515	4516	AAGTAGTAACCCAACTGAGAGTGAAA	Exon 2	564	56
PDK4	4516	4515	GGATTCCTTTGTAATATCAAACTGTC			
PDK4	4420	4421	GACAACGGTTTTCCAAATGTC	Exon 3	578	60
PDK4	4421	4420	GCCGAAACTCTGTTACTCTGAAA			
PDK4	4517	4518	AATCTGCAGTGTGGGAACCT	Exon 4	492	58
PDK4	4518	4517	AAGTGTGACCCACCAAGGAG			
PDK4	4519	4751	GATGGGCGTGGAATATAAGC	Exon 5	1097	60
PDK4	4751	4519	GCAGTCTTTCTTTGATCTACGG			
PDK4	4750	4751	TTGATGGAAGATCACTTGAAAA	Exon 6	496	58
PDK4	4751	4750	GCAGTCTTTCTTTGATCTACGG			
PDK4	4422	4423	GAGTGAGGAGGATGGGGAGT	Exon 7	541	60
PDK4	4423	4422	TGGGGGTGGAAGTTAACAAG			
PDK4	4890	4891	TGAATTACCCTCTGTGCTGCT	Exon 8	379	60
PDK4	4891	4890	GCCACATGTTAAGAATAATGCAC			
PDK4	4424	4425	TGGCCTTCTCTGACACCAAT	Exon 9	499	56
PDK4	4425	4424	TCATATCCCCTATGCCCAGT			
PDK4	3664	3665	GACAGAGGAGGTGGTGTTCC	Intron 9	510/693	59
PDK4	3665	3664	GATGATAGCATCTGTTCCATATCC			
PDK4	4426	4427	ATGGCTGCCTCTTTGTCAGT	Exon 10	566	56
PDK4	4427	4426	GAATGTGTATGCGTGACTGGA			
PDK4	4428	4429	CCATGGGCTCTAGTATGCTTTT	Exon 11	695	60
PDK4	4429	4428	TCTTTTGACATCAACCAGCAG			
PDK4	4738	4739	TGGCTGGCCAGTTCTTTATT	3'UTR	691	58
PDK4	4739	4738	CTGAGTTCAGTTCCGAATGC			
PDK4	4740	4741	TTGGAAAAACCAGTGTTACAGAA	3'UTR	692	58
PDK4	4741	4740	TGGTTATTTCATTAACAAGAAAATGC			
PDK4	4742	4743	CTGTCAAATTGATGTTTTAAACTGG	3'UTR	691	58
PDK4	4743	4742	ACAGATGTGATAAATTCATTTGAGTAA			
GAPDH	4128	4129	TCCCACGGCACAGTCAA	Expression study	536	58
GAPDH	4129	4128	GCAGGTCAGGTCCACAA			
PDK4	3664	4271	GACAGAGGAGGTGGTGTTCC	Expression study	234	58
PDK4	4271	3664	ATTCAGAAGACAAAGCCTTTAGG			
PDK4	1086F	1086R	TGCCAAGTACTTCCAAGGAGATCTA	c.1049G>A	78	
PDK4	1086R	1086F	ACCTTTAGGTAGATGATAGCATCCGT			
PDK4	1086A1	-	VIC-TCTTTGCCAGAATATG	TagMan probes		
PDK4	1086A2	-	FAM-TCTTTGCCAGGATATG			
PDK4	1099F	1086R	GGAGCCCAGATTCTTGGAAATAGTT	c.-560T>C	133	
PDK4	1099R	1086F	GGAGGAGGACTCCGGATAGC			
PDK4	1099A1	-	VIC-AAGCTGTGACAGTGTCACG	TagMan probes		
PDK4	1099A2	-	FAM-AGCTGTGACAATGTCACG			

RNA isolation

Total RNA was isolated form skeletal muscle using an RNeasy® Fibrous Tissue Mini Kit (Qiagen, Hilden, Germany) according to the standard protocol, omitting DNase digestion steps. Homogenisation of the 30 mg tissue was achieved by processing the sample in a FastPrep Instrument (Qbiogen, Inc, CA) for 30 seconds at a speed setting of 5.0 m/s using Lysing Martix D (Qbiogen).

The method for RNA isolation from fat tissue was optimised and established. Total RNA from backfat was isolated from 80 mg adipose tissue using the RNeasy® Lipid Tissue Mini Kit (Qiagen, Hilden, Germany) with a modified protocol. Adipose tissue was placed directly into 1.3 ml RTL buffer with ß-Mercaptoethanol and homogenised using Lysing Matrix D (Qbiogen) and the FastPrep Instrument (Qbiogen, Inc, CA) for 30 seconds at a speed setting of 4.0 m/s. The homogenate was incubated for 5 minutes at room temperature. The supernatant was transferred to a new reaction tube, mixed vigorously with 300 ml chloroform and incubated for 5 minutes at room temperature. The upper aqueous phase was transferred to a new reaction tube after centrifugation at 6000 rpm for 15 minutes at 4°C. The volume of 650 µl of 96% Isopropanol was added and the reaction mix was incubated for 2 h on ice before it was transferred to an RNeasy Mini Spin Column. The further extraction was carried out according to the manufacturer's protocol starting with step 10. Total RNA was stored at -80°C. The quality and quantity of RNA were evaluated only on a 1.5% denaturing, ethidium bromide stained agarose gel containing 1% formaldehyde. RNA was not measured with a spectrophotometer due to the small volume of samples.

cDNA synthesis

The reverse transcription reaction was performed using a First Strand cDNA Synthesis Kit (Fermentas, St.-Leon Roth, Germany) with random hexamer primers according to manufacturer's protocol. cDNA was measured by an Ultraspec III UV/Visible Photometer (Pharmacia LKB Biotechnology, Uppsala, Sweden) at 260, 280 and 320 nm and 400 ng/µl of cDNA was used in a subsequent quantitative Real-Time PCR.

Quantitative Real-Time PCR

Real-Time PCR was carried out on an ABI PRISM® 7000 Sequence Detection System (Applied Biosystems, Foster City, CA, USA) using a Power SYBER® Green PCR

Master Mix (Applied Biosystems) under the following conditions: 10 min incubation at 95°C: followed by 40 cycles at 95°C for 30 s and 58°C for 1 min. The 20 µl reaction mixture was composed of 2X SYBER® Green, 5 pmol of each forward and reverse primer and 400 ng cDNA. Specific primers (one border primer) for *PDK4* and *GAPDH* are listed in Table 10. Samples were set up in three technical replicates. The threshold cycle (Ct) value of the experimental gene was divided by the Ct value of the reference gene and non-parametric Wilcoxon Rank Sum Test (Wilcoxon 1945) was applied.

Statistical analysis

Statistical analysis was carried out using the R environment for the statistical computing (http://www.r-project.org/). Appropriate models were established by stepwise backwards model selection. Linear models accounted for the fixed effects of genotypes, dam, sire, sex, *RYR1* and the continuous effects: weight at slaughter and days at the experimental station. Models included the *RYR1* C1843T genotype, since *RYR1* variation is known to explain phenotypic variations in meat and carcass traits (Fujii *et al.* 1991). Covariates were removed from the model if the P-value was ≥ 0.05. The nominal P-values were obtained by ANOVA analysis in the R.package 'car' (version1.1-1). If necessary, the data was transformed or the P-value was permuted. Least square means (LSM) and standard errors (SE) implemented in the R.package 'effects' (version1.0-8) were estimated with untransformed data. Outliers were removed if they were still present after the transformation of phenotypic data. For each commercial population linear models accounted for the fixed effects: genotypes, *RYR1,* performance testing station and the random effect of weight at slaughter. The effect of sex was not included because animals in each commercial population were of the same gender. The sire effect was not included in the model, due to the 270 sires used in the Piétrain population and 193 sires in German Landrace. The Hardy-Weinberg Equilibrium was tested by Fisher's exact test statistic (Wigginton *et al.* 2005) implemented in the R.package 'genetics', version 1.3.2. Sequential Bonferroni correction was applied to account for multiple testing (significance threshold 1.51×10^{-3}).

Bioinformatics analysis

A putative porcine transcription start site was predicted using the open source software Neutral Network Promoter Prediction software v.2.2

(http://www.fruitfly.org/seq_tools/promoter.html) (Reese 2001). To predict conserved promoter sites of *PDK4* the 900 kb porcine and human sequences upstream of the translation start site were submitted into the Consite software (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/) (Sandelin *et al.* 2004). The conservation cutoff was adjusted to sequence identity of 80% and analyses were carried out with a set of transcription factors characteristic for human and vertebrates. Additionally, transcription factor binding sites were predicted with the Cister software (http://zlab.bu.edu/~mfrith/cister.shtml) (Frith *et al.* 2001) with default setting. To investigate putative CpG islands the porcine, human and bovine 5'-regions were analysed with the program cpgplot from the Emboss web service (http://www.ebi.ac.uk/emboss/cpgplot/) (Larsen *et al.* 1992) with default setting. A sequence fragment of 900 kb upstream of the translation start site was used as a query. A polyadenylation signal were identified based on the human polyadenylation signal database by submitting the last exon and flanking distal sequence of *PDK4* to the polyadq software (http://rulai.cshl.org/tools/polyadq/polyadq_form.html) (Tabaska & Zhang 1999).

Human (ENST00000005178), cattle (ENSBTAT00000038879), mouse (ENSMUST00000019721) and rat (ENSRNOT00000012760) *PDK4* reference sequences were obtained from the Ensembl database (www.ensembl.org).

Results

The genomic sequence of the porcine *PDK4* gene was obtained by sequencing a porcine CH242-147O16 BAC clone and it was deposited in the GenBank database under accession number FJ853993. Porcine *PDK4* consists of 11 exons with the translation start codon located in exon 1 and stop codon in exon 11 (Table 11). The entire locus spans approximately 11 kb and the coding sequence extends over 1221 bp. The human and the porcine coding sequences of *PDK4* show 90% identity. The predicted transcription start site is located at position -226 bp with a score of 0.95, relative to adenine of the translation start site (TSS). A TATA-like motif (AAATAAAA) was identified, beginning at position -249 in relation to adenine of the TSS (Figur 2). The putative CpG island with a length of 203 bp is located -347 bp relative to the TSS. In the highly conserved part of the 5'-region transcription binding sites for the estrogen-related receptor alpha (ERRα), the forkhead transcription factor and an incompletely

preserved SP1 binding site as well as a CAAT box were found (Wende *et al.* 2005; Araki & Motojima 2006) (Figure 3).

Porcine *PDK4* encodes 407 amino acids like the bovine gene, but in contrast to the human and mouse genes, which encode 411 and 412 amino acids, respectively. Similarities between the amino acid sequences of the porcine PDK4 and human (NP_002603, NCBI), cattle (NP_001095353, NCBI), rat (NP_446003, NCBI) and mouse (NP_038771, NCBI) proteins are 93, 93, 92 and 92%, respectively.

Table 11 Exon-intron organization of the porcine *PDK4* gene.

	Exon						**Intron**		
No.	Size (bp)	Position in cDNA	aa	3'-splice acceptor		5'-splice donor[b]	No.	Size (bp)	Human (bp)
1	340	1-115	38	ggtgacAGGACA	...	ACTTCG**gt**gagt	*1*	953	999
2	142	116-257	47	ttac**ag**GTTCAG	...	GAGCTG**gt**aagt	*2*	1155	1252
3	72	258-329	24	ttac**ag**GTATAT	...	AGCAGA**gt**aagt	*3*	748	754
4	185	330-514	62	gctt**ag**TTTTGT	...	AACACA**gt**aagt	*4*	165	162
5	87	515-601	29	caac**ag**TCCTGA	...	TCCAAG**gt**aatt	*5*	455	445
6	78	602-679	26	ttcc**ag**ATGCCT	...	TGAATG**gt**aagc	*6*	2189	2271
7	77	680-756	26	ttct**ag**GAAAAC	...	TTCAAG**gt**aaaa	*7*	1770	1814
8	99	757-855	33	ctttag**AA**TGCA	...	ATTAAG**gt**aaac	*8*	204	198
9	111	856-966	37	tataag**AT**TTCA	...	CCTTTG**gt**aaga	*9*	299	294
10	114	967-1080	38	ttct**ag**GCTGGT	...	CTAAAG**gt**atcc	*10*	1695	1228
11	144[a]/-	1081-1224	47	tttc**ag**CTTTG	...				

[a] Exon contains stop codon

[b] Exon sequences are shown in capital letters, intron sequences in lowercase letters. The consensus splice site sequences are bolded.

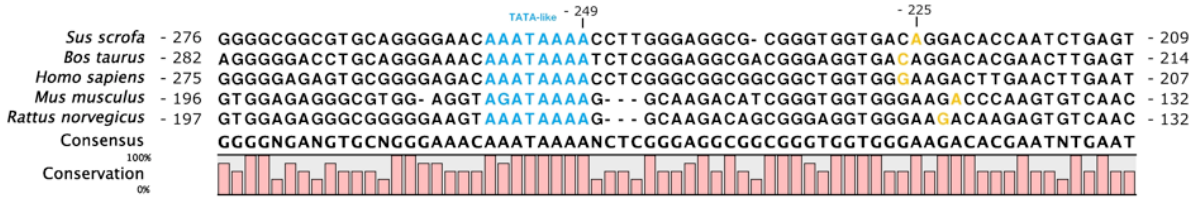


Figure 2 Predicted transcription start site in partial alignment of *PDK4* 5'-region from several species.

Sequence alignment of the*PDK4* 5'-region from pig, cattle (ENSBTAT00000038879), human (ENST00000005178), mouse (ENSMUST00000019721) and rat (ENSRNOT00000012760). Putative transcription start sites predicted by Neutral Network Promoter Prediction software v.2.2 (Reese 2001) in the pig, cattle, human, mouse and rat sequences with a score of 0.95, 0.89, 0.99, 0.96 and 0.98, respectively, are indicated in yellow. TATA-like motifs are shown in blue. Numbers on the left and right indicate nucleotide positions relative to adenine of the translation start site. Numbers above sequence alignment are relevant to the *Sus scrofa* sequence. Ambiguous nucleotides in the consensus sequence are indicated by the N symbol.

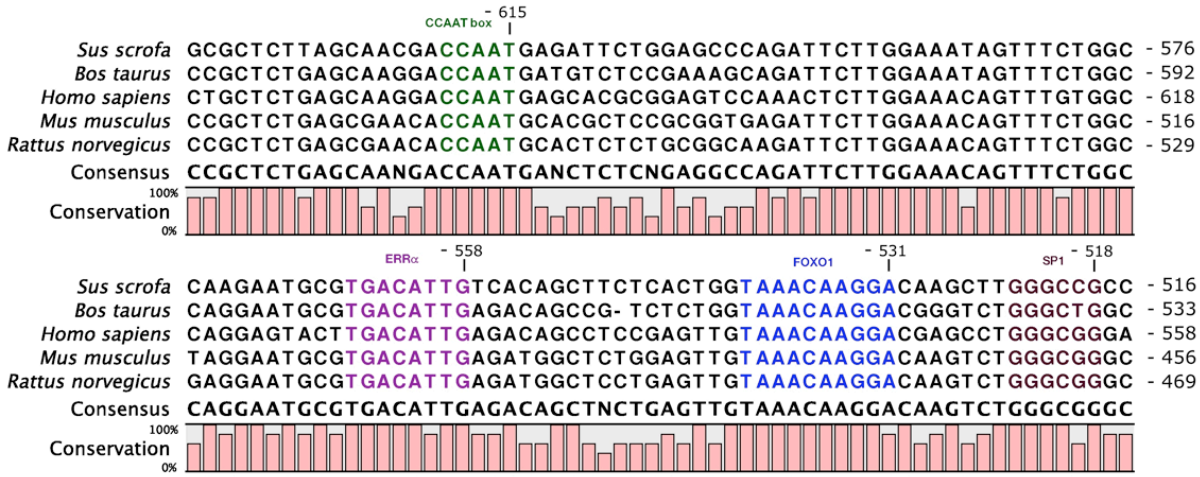


Figure 3 The highly conserved part of *PDK4* 5'-region in alignment from several species.

Sequence alignment of the *PDK4* 5'-region from the pig, cattle (ENSBTAT00000038879), human (ENST00000005178), mouse (ENSMUST00000019721) and rat (ENSRNOT00000012760). Putative CCAAT boxes are indicated in green, the estrogen-related receptor alpha (ERRα) binding site in purple, the forkhead transcription factor (FOXO1) in dark blue, and the SP1 binding site in brown (Araki & Motojima 2006). Numbers on the left indicate nucleotide positions relative to adenine of the translation start site. Numbers above sequence alignments are relevant to the *Sus scrofa* sequence. Ambiguous nucleotides in the consensus sequence are indicated by the N symbol.

The coding sequence, flanking splice junctions and the putative promoter region of *PDK4* were screened for polymorphisms in the parental animals of the cross. In total, 10440 bp were re-sequenced resulting in the detection of 24 variants, including 21 SNPs and 3 insertion-deletion polymorphisms (Table 12).

We have identified a non-synonymous substitution c.1049G>A at codon Gly350 in exon 10. This substitution causes a conservative amino acid exchange from nonpolar glycine to the negatively charged glutamic acid. In the Mangalitsa boars the G allele was fixed, whereas in the Piétrain founders both alleles were segregating. In the putative promoter region the SNP c.-560T>C has been identified within the ERRα factor binding site, the partner of the peroxisome proliferators-activated receptor γ co-activator 1α (PGC-1α) in the PPARα-independent activation of the *PDK4* gene expression (Ma *et al.* 2005; Wende *et al.* 2005; Araki & Motojima 2006; Zhang *et al.* 2006). Both alleles were segregating in the Mangalitsa as well as Piétrain founders. In intron 9 an insertion-deletion polymorphism of a short interspersed element (SINE) was detected (Figure 4). The inserted allele with a length of 183 bp derived from heterozygous Mangalitsa boars

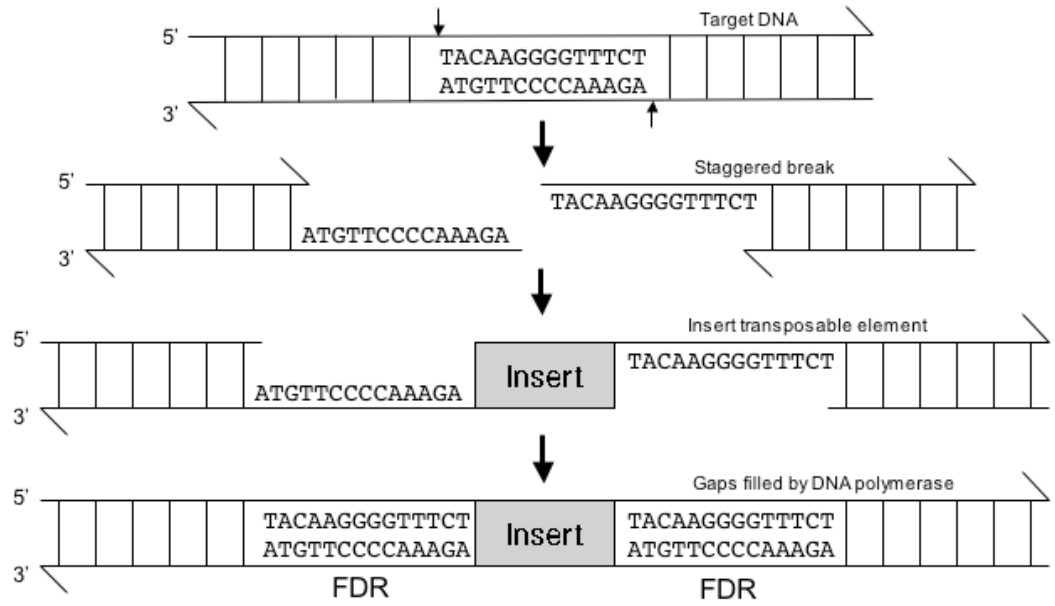
increases the length of the intron by 61%. Potentially the longer intron length reduces the level of gene expression (Vinogradov 2001; Swinburne *et al.* 2008). These three functionally substantiated polymorphisms were selected to an association study in F2 animals.

Table 12 Sequence variants in porcine *PDK4*.

Polymorphisms of *PDK4* with localisation, indication of minor allele (MA), minor allele frequency (MAF) in the parental animals of Mangalitsa x Piétrain cross and description. The position of polymorphisms is given according to the sequence deposited in the GenBank FJ853993.

No.	Variant[1]	Localisation	Type	MA	MAF	Description
1	c.-815T>G	Promoter	SNP	G	0.179	
2	c.-754C>G	Promoter	SNP	G	0.179	
3	c.-574G>A	Promoter	SNP	G	0.214	
4	c.-560T>C	Promoter	SNP	C	0.429	Putative ERRα binding site
5	c.-525T>C	Promoter	SNP	C	0.179	
6	c.-471G>A	Promoter	SNP	G	0.464	
7	c.-310C>T	Promoter	SNP	T	0.214	
8	c.-284delG	Promoter	DEL		-	
9	c.116-29C>T	Intron 1	SNP	T	0.393	
10	c.257+32A>G	Intron 2	SNP	G	0.393	
11	c.329+100A>G	Intron 3	SNP	G	0.179	
12	c.601+71A>G	Intron 5	SNP	G	0.286	
13	c.679+12A>G	Intron 6	SNP	G	0.286	
14	c.679+56T>G	Intron 6	SNP	T	0.071	
15	c.679+206T>C	Intron 6	SNP	C	0.286	
16	c.679+279G>A	Intron 6	SNP	A	0.071	
17	c.756+47A>C	Intron 7	SNP	C	0.286	
18	c.756+286C>T	Intron 7	SNP	T	0.286	
19	c.855+44_855+45insT	Intron 8	INS	I	0.071	
20	c.967-139_967-138ins183,GenBank FJ853993	Intron 9	INS	I	0.071	Short Interspersed Element Polymorphism
21	c.967-40A>G	Intron 9	SNP	G	0.071	
22	c.1049G>A	Exon 10	SNP	A	0.107	Non-synonymous substitution
23	*1252G>A	3' UTR	SNP	A	0.179	
24	*1898G>A	3' UTR	SNP	A	0.179	

a)



b)

```
0001 TACAAGGGGT TTCTTTTTTT TTTTTTTTTT TTTTTTTAAA AAAAATTTTA
0051 TTTTCCCACT GTACAGCAAG GGGGTCTGGT TATCCTTACA TGTATACATT
0101 ACAATTACAG TTTTTCCCCC ACCATTTCTT CTGTTGCAAC ATGAGTATCT
0151 AGACATAGTT CTCAATGCTA TTCAGCGGGA TCT
```

Figure 4 Insertion-deletion polymorphism of interspersed element (SINE) in *PDK4*.

a) Insertion mechanism and generation of flanking direct repeats (FDRs); b) sequence of SINE with FDR sequence indicated in grey.

In the investigated F2 animals the frequency of c.1049A, c.-560C alleles and the inserted allele of SINE is 0.07, 0.35 and 0.34, respectively. We observed a significant association between c.1049G>A and saturated fatty acid (SFA) contents and a suggestive association with backfat thickness at loin (BFTL) (Table 13). The insertion-deletion polymorphism of SINE was found to be associated with SFA content and average feed intake (AFI) (Table 13). The c.-560T>C substitution shows a suggestive association with daily gain weight (Table 13). None of the associations are significant after Bonferroni correction for multiple testing.

Table 13 Association of three *PDK4* polymorphisms (1) c.1049G>A (2) SINE (3) c.-560T>C with lipid deposition and energy homeostasis-related traits in F2 population of Mangalitsa x Pitérain cross.

The A sign corresponds to c.1049G, the deleted SINE allele, c.-560T allele and B indicates the second allele of investigated SNPs. Investigated traits: average backfat thickness (BFT), backfat thickness at loin (BFTL), intramuscular fat content (IMF), saturated fatty acids (SFA) content, average feed intake (AFI) and daily gain (DG); fixed effect and covariates used in the linear model, outliers (Outl), transformation of data (Transf), polymorphism number (Pol), least square means (LSM) with standard errors (SE) and P-value are present. Significant results are indicated in bold.

Trait	**Fixed effect Covariates**	**Outl**	**Transf**	**Pol**	**LSM (SE)**			**P-value**
					AA	AB	BB	
Numbers of observation				1	GG=453	AG=71	AA=0	
				2	DD=226	DI=239	II=59	
				3	TT=221	CT=237	CC=66	
BFT [mm]	dam, sire, sex, weight, days	>11	-	1	33.420 (0.230)	34.113 (0.731)	-	0.395
				2	33.857 (0.326)	33.263 (0.307)	33.232 (0.675)	0.406
				3	33.785 (0.386)	33.183 (0.335)	33.799 (0.672)	0.438
BFTL [mm]	dam, sire, sex, days, RYR1	-	-	1	26.020 (0.283)	27.755 (0.899)	-	0.083
				2	26.392 (0.401)	26.255 (0.379)	25.735 (0.838)	0.794
				3	26.600 (0.476)	25.875 (0.413)	26.465 (0.830)	0.508
IMF [%]	dam, sire, sex, weight, RYR1	-	log	1	2.066 (0.027)	1.979 (0.086)	-	0.234
				2	2.056 (0.039)	2.039 (0.036)	2.107 (0.080)	0.946
				3	2.075 (0.046)	2.018 (0.039)	2.115 (0.081)	0.376
SFA [%]	dam, sire, sex, days, RYR1	<30 >60	-	1	36.746 (0.063)	36.303 (0.199)	-	**0.046**
				2	36.816 (0.08)	36.665 (0.083)	36.273 (0.186)	**0.039**
				3	36.535 (0.106)	36.764 (0.092)	36.918 (0.189)	0.221
AFI [kg/day]	dam, sire, sex, weight, days	>2.8	-	1	2.021 (0.006)	2.036 (0.019)	-	0.497
				2	2.043 (0.008)	2.008 (0.008)	2.007 (0.018)	**0.011**
				3	2.017 (0.010)	2.024 (0.009)	2.043 (0.018)	0.489
DG [kg/day]	dam, sex, weight, days	-	-	1	0.555 (0.002)	0.558 (0.006)	-	0.602
				2	0.557 (0.003)	0.553 (0.002)	0.555 (0.006)	0.434
				3	0.554 (0.003)	0.558 (0.003)	0.547 (0.005)	0.100

Because of the extensive linkage disequilibrium (LD) blocks existing in the F2 population further investigations of the significantly associated c.1049G>A and an insertion-deletion polymorphism of SINE were performed. The Piétrain-derived non-synonymous substitution c.1049G>A was investigated in additional 1013 unrelated Piétrain pigs. The frequency of the c.1049A allele is 0.06 and the SNP is in the Hardy-Weinberg equilibrium (P-value= 0.253). Only one animal was homozygous AA, thus we excluded it from the further study. A significant difference was found in the breeding value of IMF (Table 14). The AFI and SFA data were not available in the Piétrain population, thus we could not investigate these traits.

Table 14 Association of the non-synonymous c.1049G>A *PDK4* polymorphism with lipid deposition and daily gain in the Piétrain population.

Investigated traits: average backfat thickness (BFT), backfat thickness at loin (BFTL), breeding value for intramuscular fat (IMF) and daily gain (DG); fixed effect and covariates used in the linear model, transformation of data (Transf), least square means (LSM) with standard errors (SE) and P-value are presented. Significant results are indicated in bold.

Trait	Fixed effect Covariates	Transf	LSM (SE)		P-value
			GG (n=893)	AG (n=119)	
BFT [mm]	*RYR1*, station, weight	log	0.573 (0.005)	0.568 (0.015)	0.954
BFTL [mm]	*RYR1*, station, weight	sqrt	0.795 (0.008)	0.776 (0.021)	0.505
IMF [ebv]	*RYR1*	-	0.016 (0.005)	-0.018 (0.014)	**0.030**
DG [kg/day]	*RYR1*, station, weight	-	0.758 (0.003)	0.770 (0.007)	0.111

The Mangalitsa-derived insertion-deletion polymorphism of SINE was tested in 253 F3 animals. A suggestive association was observed between this polymorphism and AFI with the same tendency as in the F2 animals (Table 15).

Table 15 Association analysis of the *PDK4* insertion-deletion polymorphism of SINE with average feed intake (AFI) in F3 animals and German Landrace.

Trait	Fixed effect Covariates	Outl	LSM (SE)			P-value
			DD	DI	II	
F3 animals						
AFI [kg/day]	sire, sex, weight, days	>2.8	n=115 2.078 (0.014)	n=92 2.049 (0.015)	n=46 2.026 (0.022)	0.119
German Landrace						
AFI [kg/day]	weight, days	-	n=466 2.321 (0.009)	n=36 2.363 (0.031)	n=2 2.384 (0.131)	0.381

Based on observations reported by Vinogradov *et al*. and Swinburne *et al*. (Vinogradov 2001; Swinburne *et al.* 2008) we speculated that homozygous animals without the SINE, hence with the shorter transcript, would represent a higher PDK4 expression level. To test this hypothesis we selected five F3 animals carrying two longer alleles and ten homozygous F3 animals of the shorter allele of the SINE polymorphism. RNA was isolated from muscle and fat tissue, reverse transcribed and PDK4 expression was analysed in a relative quantification approach with *GAPDH* as the reference gene. We observed a lower PDK4 expression in muscle and fat tissue of pigs with the SINE

compared to the pigs without the SINE (Table 16). The differences were non-significant; however, the observed results are in an agreement with our assumption.

Table 16 mRNA expression of PDK4 regarding the insertion-deletion polymorphism of SINE.

	Ratio (SE)		*P-value*
	II (n=5)	**DD (n=10)**	
Muscle	1.123 (0.028)	1.178 (0.133)	0.5941
Fat	0.934 (0.093)	0.946 (0.068)	0.7679

In the Piétrain founders the inserted allele of the SINE polymorphism was not segregating, therefore its frequency was examined in additional commercial pig breeds (Table 17). The highest frequency of the inserted allele is observed in German Landrace, thus 504 animals with an average weight at slaughter of 107 kg were genotyped regarding the insertion-deletion polymorphism of SINE. The frequency of the inserted allele is 0.04 and the polymorphism is in the Hardy-Weinberg equilibrium (P-value= 0.178). The association with AFI observed in the F2 animals was not reproduced in the commercial breed (Table 15).

Table 17 Allele frequency of the *PDK4* insertion-deletion polymorphism of SINE in different pig breeds.

	Number of animals			*Allele frequency*		*HWE*
	DD	**DI**	**II**	**D**	**I**	*P-value*
Piétrain	70	2	0	0.99	0.01	1
German Landrace	43	6	2	0.90	0.10	0.057
Large White	12	0	0	1	0	-

Discussion

PDK4 was selected as a candidate gene for lipid deposition and energy homeostasis-related traits because of the role of PDK4 in the regulation of energy and lipid metabolism and its localisation in a QTL region for backfat thickness, daily gain, feed intake and saturated fatty acids. The candidate gene analysis of porcine *PDK4*, presented here, identified the non-synonymous mutation in exon 10, the insertion-deletion polymorphism of SINE in intron 9 and the promoter mutation in the ERRa factor binding site. These three potentially functional variants were selected for an association study in the F2 animals.

The association of the insertion-deletion polymorphism of SINE with AFI as observed in the F2 animals could not be confirmed in the German Landrace pigs, possibly due to

the low minor allele frequency (0.04) in the German Landrace animals. The suggestive association of the non-synonymous polymorphism with BFTL observed in the F2 population could not be replicated in the Piétrain population, which might be explained by the phenotypic variation of BFTL among investigated populations (Table 9). The association with breeding value of IMF and the non-synonymous substitution was found in Piétrain, while in the F2 population the association with IMF content was not detected. The estimated breeding value is a more accurate indicator and thus we were able to observe the significant association in Piétrain and not in the F2 animals. It should be noted that inconsistent results in the F2 and commercial populations might indicate that the two analysed variants were unlikely casual and the observed association could be due to linkage disequilibrium with casual mutation(s).

Recently, Lan *et al.* (Lan *et al.* 2008) showed that an A/G mutation in intron 9, which correspond to c.967-40A>G detected in our resource population, is significantly associated with IMF and muscle water content with a dominant effect. The c.967-40A>G polymorphism is in LD with the insertion-deletion polymorphism of SINE in our cross. The association studies of insertion-deletion polymorphism did not show a significant effect on IMF in the F2 (Table 13) and F3 (data not show) animals. Due to a lack of phenotypic data we could not test the association in the German Landrace population.

Chapter 4

Characterisation of the porcine INSIG2 and FTO genes

Summary

We have annotated the porcine *INSIG2* and *FTO* genes and re-sequenced all exons, flanking splice junctions and putative promoter region in the parental generation of the Mangalitsa x Piétrain cross. The coding sequence of the porcine *INSIG2* and *FTO* genes is highly conserved across human, cattle, mouse, and rat. Moreover, all the functionally important domains and amino acids are conserved. A total of 13 variants were identified in *INSIG2*, including a synonymous substitution in exon 2 and a microsatellite polymorphism in the 3' untranslated region. In *FTO* 53 variants were detected, including 49 single nucleotide polymorphisms and 4 deletions. In the F2 animals significant associations were observed between the microsatellite polymorphism of *INSIG2* and thinnest backfat thickness (P-value=0.025), saturated fatty acid content (P-value=0.033) and polyunsaturated fatty acid content (P-value=0.008). The synonymous substitution c.594G>C of *FTO* shows an association with backfat thickness at shoulders (P-value=0.020) in the F2 animals and with loin muscle area (P-value=0.023) in the Piétrain population.

Keywords: INSIG2, FTO, pig, genomic characterisation, association study

Introduction

The insulin induced gene 2 (*INSIG2*) and the fat mass and obesity associated gene (*FTO*) were indicated through a genome-wide association scan as strong candidate genes for increasing the body mass index (BMI) in children and adults. Herbert *et al.* (Herbert *et al.* 2006) presented an association between a single nucleotide polymorphism (SNP) (rs7566605) upstream of the human *INSIG2* gene and BMI. However, the association has been reproduced only in several, but not in all analysed cohorts (Dina *et al.* 2007b; Loos *et al.* 2007; Lyon *et al.* 2007; Rosskopf *et al.* 2007). The possible explanations for the observed divergence are widely discussed by Lyon *et al.* (Lyon *et al.* 2007). A clear association between the cluster of SNPs in the first intron of *FTO* and BMI was reported by Frayling *et al*. (Frayling *et al.* 2007). This finding was replicated in several other studies (Andreasen *et al.* 2007; Dina *et al.* 2007a; Hinney *et al.* 2007; Peeters *et al.* 2007; Scuteri *et al.* 2007; Price *et al.* 2008). Our interest was aimed at investigating an association of the possible human obesity candidate genes with lipid deposition traits in a Mangalitsa x Piétrain cross.

Material and Methods

Animals

We studied F2 animals from the Mangalitsa x Piétrain cross. The pigs were fed *ad libitum* up to approximately 95 kg. Male pigs were castrated. Backfat thickness, muscle depth and other carcass traits were measured post mortem. Fatty acid composition of intramuscular fat within longissimus muscle was determined by gas chromatography according to a standard protocol (Arens *et al.* 1994). Fatty acids were combined in several groups: saturated fatty acids SFA= C14:0 (Myristic acid) + C16:0 (Palmitic acid) + C18:0 (Stearic acid) [%], monounsaturated fatty acids MUFA = C16:1 (Palmitoleic acid) + C18:1 (Oleic acid) [%], polyunsaturated fatty acids PUFA = C18:2 (Linoleic acid) + C18:3 (Linolenic acid) + C20:4 (Arachidonic acid) + C20:5 (Eicosapentaenoic acid) + C22:4 (Docosatetraenoic acid) + C22:6 (Docosahexaenoic acid) [%] and n-3 polyunsaturated fatty acid n-3PUFA = C18:3 + C20:3 + C20:5 + C22:6, for which contents [%] were estimated. Both Mangalitsa boar founders were homozygous CC and all Piétrain sows were homozygous TT for c.1843C>T polymorphism in the ryanodine receptor 1 gene *(RYR1)*. Thus, all investigated animals were typed for the c.1843C>T *RYR1* polymorphism.

Samples of the confirmation study were provided by the Landesanstalt für Landwirtschaft Bayern, Grub (Germany) and contained 1019 Piétrain animals from the performance testing that were slaughter within 2004 – 2006 in Grub (Germany) and Schwarzenau (Germany). German Landrace, German Large White and Piétrain DNA samples for estimation of allele frequencies were collected from the artificial insemination centres in Bergheim (Germany), Landshut (Germany) and Neustadt/Aisch (Germany) from 2000 to 2002. Duroc samples were provided by the SUISAG (Sempach, Switzerland) and were used for estimating allele frequencies.

Prediction of the genomic organization of *INSIG2* and *FTO*

The full region of porcine *INSIG2* and *FTO* was obtained using an internet-based *in silico* BAC library (http://www.sanger.ac.uk/Projects/S_scrofa/BES.shtml). The exon-intron structure was determined manually based on alignment with the Human Reference Sequence (GenBank, NM_002612, NM_001080432.1) according to the GT/AG rule (Breathnach & Chambon 1981) as well as semi-automatically using the GenomeThreader tool (Gremme *et al.* 2005) and visualised by Genome Annotation and Curation Tool Apollo (Lewis *et al.* 2002).

Re-sequencing

Primers were designed based on the porcine genomic *INSIG2* and *FTO* sequence using the Primer3 software (Rozen & Skaletsky 2000) and purchased from MWG (Ebersberg, Germany) (Table 18). A PCR (20 µl) contained 50 ng genomic DNA, 1x Qiagen PCR buffer, 1.5 mM $MgCl_2$, 200 µM of each nucleotide, 5 pmol of each forward and reverse primer (MWG Biotech, Ebersberg, Germany), 0.5 units of Qiagen Taq Polymerase (Qiagen, Hilden, Germany) was performed using T-Gradient Thermocycler (Biometra, Göttingen, Germany) under the following conditions: initial denaturing at 94 °C for 3 minutes; followed by 30 cycles at 94 °C for 30 seconds, melting temperature for 1 minute, 72 °C for 1 minute; and the final extension at 72 °C for 5 minutes. The melting temperature varied in accordance with the primers range between 56 °C and 60 °C. The PCR products were separated on a 1.5% agarose gel stained with ethidium bromide together with a DNA marker (Genruler 100 bp ladder: SM0241; MBI Fermentas, St. Leon-Rot, Germany). PCR products were purified using MultiScreen PCRµ96 Filter Plates (Millipore, Schwalbach, Germany) in combination with a Millipore vacuum manifold. The sequencing reaction was performed using BigDye Terminator v1.1

Sequencing Kit (Applied Biosystems Applera, Darmstadt, Germany) under following conditions: initial denaturation for 15 seconds at 96 °C; followed by 35 cycles at 96 °C for 10 seconds, 51 °C for 5 seconds and 60 °C for 4 minute, and purified with Sephadex columns in a MultiScreen 96 well filtration plate (Millipore, Schwalbach, Germany). Sequencing for *INSIG2* was performed on an ABI Prism 377 automated sequencer (Applied Biosystems Applera, Darmstadt, Germany) and for *FTO* on an ABI 3130xl Genetic Analyser (Applied Biosystems Applera, Darmstadt, Germany), obtained sequences were analysed using the Phred/Phrap/Polyphred/Consed software suite (Nickerson *et al.* 1997; Ewing & Green 1998; Ewing *et al.* 1998) and the results viewed using Consed (Gordon *et al.* 1998).

Genotyping

Genotyping of the F2 animals were performed using restriction fragment length polymorphism (RFLP) method. Amplified fragments were digested with *Alu*I, *Mse*I, *BstU*1, *Hpa*2 regarding to the *INSIG2* c.237A>G, and the *FTO* c.46-143C>T, c.594G>C, *1176C>T polymorphisms, respectively. Restriction enzymes were chosen with help of the Hypertext DNA Sequence Display (http://pga.mgh.harvard.edu/web_apps/web_map/start) and ordered from Ferments (St.Leon-Rot, Germany) or New England BioLabs GmbH (Frankfurt am Main, Germany). Genotypes of *INSIG2* c.237A>G and *FTO* c.594G>C polymorphisms in Piétrain were determined by TaqMan genotyping assays (Table 18) according to manufacture's protocols.

Table 18 Primers used for re-sequencing and genotyping of porcine *INSIG2* and *FTO*.

Locus	Primer	Pair	Sequence	Description	Product size [bp]	Tm [°C]
INSIG2	4408	4409	CCATTCTGTATAATCCACCCACA	5'UTR	846	60
INSIG2	4409	4408	GTGGAGGCTACCAAATCAGC			
INSIG2	4410	4411	ATTCGTCACAACCCTCTGCT	Exon 1	665	60
INSIG2	4411	4410	TGACAGCTCTGTGCTGGAGA			
INSIG2	4137	4138	TCTCGTCATTTCATTATTTTGTCC	Exon 2	571	56
INSIG2	4138	4137	TTGCTCCGACAATAACTACCC			
INSIG2	4206	4207	TCCATCTGGATGGCATTACA	Exon 3	555	56
INSIG2	4207	4206	CTACGCACCGCATACACTG			
INSIG2	4412	4413	TTTCAAATGCTTATGCTGGTTG	Exon 3	594	56
INSIG2	4413	4412	CCCATCCTTACTGGCAAATC			
INSIG2	4139	4140	GAAGGCACAAAGCTAAAGCA	Exon 4	492	56
INSIG2	4140	4139	ACTCCACAGCTCTTGGCAAC			
INSIG2	4141	4142	TGAGGATTCGGGTTCTACC	Exon 5	500	56
INSIG2	4142	4141	AAATCTGTCTTACATCTCAACAACA			
INSIG2	4414	4415	TGAGGAATAAAGGGAGCCAAT	Exon 6	772	56
INSIG2	4415	4414	GCAAGTTTGAAATGTTTCATGG			
INSIG2	4731	4732	6-FAM-GCCAGGATGCAGTTTTTC	Microsatellite	115/117	60

				in 3'UTR		
INSIG2	4732	4731	TGATTTACAGATGAAGCAATCACA			
INSIG2	1059F	1059R	AGCATCTTTTCTTCGGCATGGT	c.237A>G	79	
INSIG2	1059R	1059F	GGCATTACAGAAATAGCCCACACAT			
INSIG2	1059A1	-	VIC-CCTGAAGCTGTGCCAC	TagMan		
INSIG2	1059A2	-	FAM-CTGAAGCCGTGCCAC	probes		
FTO	5526	5527	AACTTTTCCCTGAGAGAATTTGG	5'UTR	447	53
FTO	5527	5526	GGTGCATAATGAGCCAGGAG			
FTO	5528	5529	GGCTGCAATGTCTCTCCAAT	5'UTR	564	53
FTO	5529	5528	TGTCCTGGGTTAGGAGTTCG			
FTO	5530	5531	TGGGGTAATCACGATAAACAAA	5'UTR	596	53
FTO	5531	5530	AGAATTTCCCAGGTCCAACA			
FTO	5425	5426	CGTTGACTCCTGGGAAATGT	Exon 1	450	60
FTO	5426	5425	CCTTGTCATCCCTTGGCTTA			
FTO	5427	5428	TGCTTCATGAAATGCGAGAG	Exon 2	473	56
FTO	5428	5427	GGAGTCACTGGAGGACCAAA			
FTO	5429	5430	CCAAGGAAAGACATACTCTTGGA	Exon 3	492	60
FTO	5430	5429	CCTGGGGAAATCTGGACCTA			
FTO	5431	5432	CCTTCCTCAAGCTCAACGAC	Exon 3	549	60
FTO	5432	5431	GAAAGAGGCAGGTTCTGCAC			
FTO	5433	5434	CCCAAAAGAAATGGCATGTTA	Exon 4	599	56
FTO	5434	5433	CAACTCTGGAAGTTGGCAGT			
FTO	5500	5501	TGAATGGTGGGCTTACTGTTT	Exon 5	397	56
FTO	5501	5500	TGCTGTGGGCACACTTTTTA			
FTO	5435	5436	AGGTGAGTAGCCCAGGTGAG	Exon 6	482	60
FTO	5436	5435	GCCCTGGCTACTGCTACTGT			
FTO	5437	5438	GAAAGGCTCCAGAGACTCCA	Exon 7	472	60
FTO	5438	5437	CATGGAACTGAGCTGGTCCT			
FTO	5502	5503	GGCCATTTTTGACCATCATT	Exon 8	450	56
FTO	5503	5502	GGAAGGTACCACGTTCAGGA			
FTO	5504	5505	ACCATCAAGGCGGATCGT	Exon 9	599	58
FTO	5505	5504	AGGCCTGAGAAGGAGGAAAG			
FTO	5506	5507	AAATCTCCCCGCAAACTCTT	3'UTR	795	58
FTO	5507	5506	TCAAAGCCCAGGCTTCTTAC			
FTO	5508	5509	TTGTTTTCCACGTGTCTGGA	3'UTR	647	56
FTO	5509	5508	CCTACCGCAGCTCAGCTTAC			
FTO	1928F	1928R	GACGAGGTGGACAGGAAGAG	c.594G>C	63	
FTO	1928R	1928F	GGGATCCATGAAGTTCAACAAAGTT			
FTO	1928A1	-	VIC-CAGAGCCGCCTACAAC	TagMan		
FTO	1928A2	-	FAM-AGAGCCGCGTACAAC	probes		

Microsatellite genotyping was performed on an ABI Prism 377 sequencer (Applied Biosystems Applera, Darmstadt, Germany) based on the size dependent electrophoretic mobility of the fluorescence-labeled PCR products in polyacrylamid gel. Analysis of resulting data was carried out with the GeneScan Analysis software, version 3.1 and Genotyper program, version 2.5. The genotyping procedure was repeated independently on ABI 3130xl Genetic Analyser (Applied Biosystems Applera, Darmstadt, Germany) and analysis was performed with GeneMapper Software, version 4.0.

Statistical analysis

Statistical analysis was carried out using the R environment for the statistical computing (http://www.r-project.org/). Analysis of variance was applied using a linear model,

established by stepwise backwards model selection, accounted for the fixed effects of genotypes, dam, sire, gender, *RYR1* genotype and the continuous effects: weight at slaughter in kg and days at the experimental station. Models included the *RYR1* C1843T genotype, since *RYR1* variation is known to explain phenotypic variations in meat and carcass traits (Fujii *et al.* 1991). Intramuscular fat content was additionally included as the covariable to the model for fatty acid pattern. The insignificant effects were removed from the model (P-value $\geq$ 0.05). In case of necessity, the data was transformed or P-value was permuted. Outliers were removed if they were still present after transformation of the phenotypic data. Least square means (LSM) and standard errors (SE) implemented in the R.package 'effects' (version1.0-8) were estimated with untransformed data. For commercial population the linear models were accounted for the fixed effects: genotypes, *RYR1,* performance testing station and the random effect weight at the slaughter. The effect of sex was not included because animals in each commercial population were of the same gender. The sire effect was included in the model for the microsatellite polymorphism in *INSIG2*, because from 149 sires only 12 transmitted the longer allele to progeny. Hardy-Weinberg Equilibrium was tested by Fisher's exact test statistic (Wigginton *et al.* 2005) implemented in the R.package 'genetics', version 1.3.2. Sequential Bonferroni correction was applied to account for multiple testing. The significance thresholds of 5% were divided by the product of traits and number of tested variants (16 traits x 4 variants, significance threshold 7.81×10^{-4}).

Bioinformatics analysis

A putative porcine transcription start site was foreseen using the open source software the Neutral Network Promoter Prediction software v.2.2 (http://www.fruitfly.org/seq_tools/promoter.html) (Reese 2001). To predict conserved promoter sites of *INSIG2* the 1 kb porcine and human sequences upstream of the 3' splice site of exon 1 and of *FTO* the 1 kb porcine and human sequences upstream of the translation start site were submitted into the Consite software (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/) (Sandelin *et al.* 2004). The conservation cutoff was adjusted to sequence identity of 80% for *INSIG2*, 90% for *FTO* and analysis was carried out with a set of transcription factors characteristic for humans and vertebrates. Additionally, transcription factor binding sites were predicted with the Cister software (http://zlab.bu.edu/~mfrith/cister.shtml) (Frith *et al.* 2001) with default setting. To investigate putative CpG islands the porcine, human and bovine 5'-region

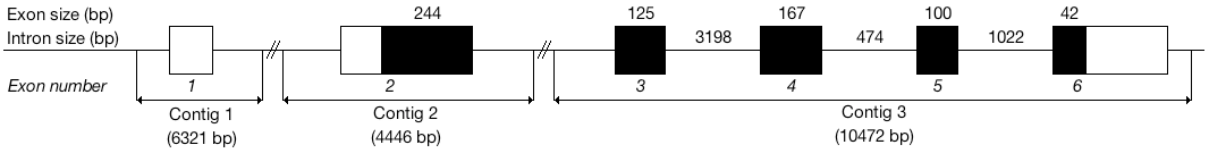
were analysed with the program cpgplot from the Emboss web service (http://www.ebi.ac.uk/ emboss/cpgplot/) (Larsen *et al.* 1992). A sequence fragment of 1 kb upstream of the 3' splice site of *INSIG2* exon 1 and 1 kb upstream of the translation start site of *FTO* was used as a query. The default length of a CpG block was decreased to 100 nucleotides for *INSIG2*. A polyadenylation signal was identified based on the human and mouse polyadenylation signal database for *FTO* and on the mouse polyadenylation signal database for *INSIG2* by submitting the last exon and flanking distal sequence of investigated genes to the polyadq software (http://rulai.cshl.org/tools/polyadq/polyadq_form.html) (Tabaska & Zhang 1999).

Human (ENST00000245787), cattle (ENSBTAT00000002725), mouse (ENSMUST00000003818) and rat (ENSRNOT00000003391) *INSIG2* reference sequences and human (ENSG00000140718), cattle (ENSBTAG00000012501), mouse (ENSMUSG00000055932) and rat (ENSRNOG00000011728) *FTO* reference sequences were obtained from the Ensembl database (www.ensembl.org).

Results

The genomic sequence of the porcine *INSIG2* gene was obtained by sequencing a porcine PigE-40J14 BAC clone. The assembly of the BAC shotgun sequences resulted in three genomic contigs containing *INSIG2* (Figure 5). The genomic porcine *INSIG2* sequence was deposited in GenBank (FJ853992). Porcine *INSIG2* is organised in 6 exons with the translation start site assigned to exon 2 and the stop codon to exon 6 (Table 19). The derived coding sequence of porcine *INSIG2* shows 94% identity with the human coding sequence. Only one non-coding exon 1 was detected. The predicted exon 1 sequence is in 81% identical to the sequence of human exon 1. No TATA or CAAT boxes were identified. The putative CpG island with a length of 158 bp partially overlaps the predicted exon 1 and is located 57 bp upstream of the 3' splice site of exon 1. The putative SP1 binding sites are located -24 bp, -28 bp and -68 bp from the 3' splice site of exon 1. Based on the human polyadenylation signal database polyadq the polyadenylation signal (ATTAAA) is predicted at 742 bp downstream of the stop codon.

a)



b)

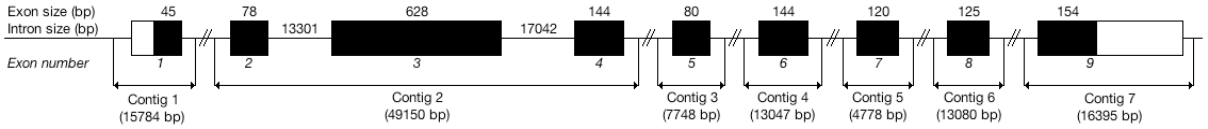


Figure 5 Exon-intron structure of porcine *INSIG2* and *FTO*.

The coding sequences are shown as black rectangles, numbers on top indicated the size (in basepare) of coding sequences. The 3' and 5' untranslated regions are marked as white rectangles. Numbers of exons are indicated below black/white boxes. Introns are shown as a thin line with sizes (in basepare) indicated on top. Genomic contigs with a size (bp in parentheses) are indicated below a gene structure.

Table 19 Exon-intron organization of the porcine *INSIG2* and *FTO* genes.

Exon						Intron		
No.	Size (bp)	Position in cDNA	aa	3'-splice accteptor[b]	5'-splice donor[b]	No.	Size (bp)	Human (bp)
				INSIG2				
1b	66	-	-	ggagttACGTGA ...	ACAGGG**gt**gagt	*1b*	-	7877
2	390	1-244	81	ttac**ag**GATTTC ...	CTTCAG**gt**atgt	*2*	-	6396
3	125	245-369	42	ccac**ag**CTGTGA ...	AGCGCT**gt**atcc	*3*	3198	3415
4	167	370-536	56	tctc**ag**AAAGTG ...	TTATCA**gt**aagt	*4*	474	186
5	100	537-636	33	ttct**ag**ATATAC ...	GCAATG**gt**atgt	*5*	1022	1091
6	42[a]/-	637-678	13	ttac**ag**TATGAA ...				
				FTO				
1	74	1-45	15	cggctaGCAGTG ...	GCTAAG**gt**atct	*1*	-	105910
2	78	46-123	26	caac**ag**AAACTG ...	CAGCAG**gt**aagg	*2*	13301	15646
3	628	124-751	209	ttttag**TGGCAG** ...	GTGAAG**gt**acag	*3*	17042	17663
4	14	752-895	48	cgac**ag**GCCCTG ...	TGCTGG**gt**aatt	*4*	-	29487
5	80	896-975	27	ctgc**ag**ATGATC ...	GCCGAG**gt**aagt	*5*	-	5978
6	144	976-1119	48	ctgc**ag**TGCTCG ...	AACGAG**gt**aagg	*6*	-	8844
7	120	1120-1239	40	cccc**ag**GTCGAG ...	GGTGCG**gt**aagt	*7*	-	45033
8	125	1240-1364	42	ttcc**ag**ACCCAT ...	TGCCAG**gt**tagt	*8*	-	177652
9	154[a]/-	1365-1518	50	tccc**ag**GTGCCA ...				

[a] Exon contains stop codon

[b] Exon sequences are shown in capital letters, intron sequences in lowercase letters. The consensus splice site sequences are bolded.

An interrupted microsatellite with a $(GT)_{17}AT(AC)_4ATAC(AT)_3$ motif was identified in the predicted 3' untranslated region (UTR). The microsatellite starts 121 nucleotides after the stop codon and it is located before a putative poly(A) signal. The alignment of the porcine microsatellite sequence with the corresponding sequences of other

vertebrate species indicates a high variability of the GT and AT repeats across all compared species and a high variability of the AC repetition within the *Artiodactyla* species (Figure 6).

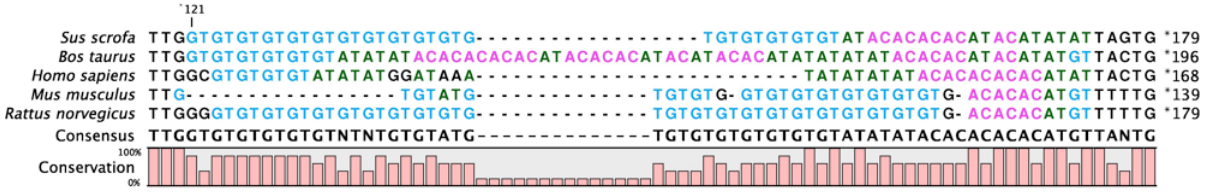


Figure 6 Microsatellite structure in alignment.

Alignment of microsatellites within the 3'UTR of *INSIG2* between swine, cattle (ENSBTAT00000002725, Ensembl), human (NM_016133, NCBI), mouse (NM_178082, NCBI) and rat (ENSRNOT00000003391, Ensembl). GT repeats are highlighted in blue, AT in green, AC in rosé. Numbers on the right site indicate nucleotide position in the 3'UTR. Number above sequence alignment is relevant to the *Sus scrofa* sequence. Ambiguous nucleotides in the consensus sequence are indicated by the N symbol.

The derived amino acid sequence is in 99% identical to the sequence of the human INSIG2 and it is organized in six transmembrane domains (Figure 7).

The genomic sequence of porcine *FTO* was obtained by sequencing of porcine BAC clones (CH242-240L15, RP44-263I3BAC). The assembly of the BAC shot gun sequences resulted in seven genomic contigs containing exons of *FTO* with adjacent parts of intron sequences (Figure 5). Full intron sequences were obtained for introns 2 and 3. The entire gene spans approximately 410 kb in the human. The coding sequence of both human and porcine *FTO* is 1515 bp and shows 88% identity. *FTO* is structured into 9 exons (Table 19) with the translation start site located within exon 1 and the translation codon stop within exon 9. The genomic sequence of porcine *FTO* was submitted to GenBank (FJ853994). No TATA or CAAT boxes were detected. Alignment with a human sequence locates the putative transcription start at position -29 bp (±3 bp) of the ATG codon. The putative SP1 binding site is located at -296 bp from the translation start site (TSS). The putative CpG island with a length of 566 bp is located -57 bp relative to the TSS. Based on the human and mouse polyadenylation signal database polyadq the polyadenylation signals (AATAAA) were computationally predicted at 932 bp and at 1055 bp downstream of the stop codon, respectively.

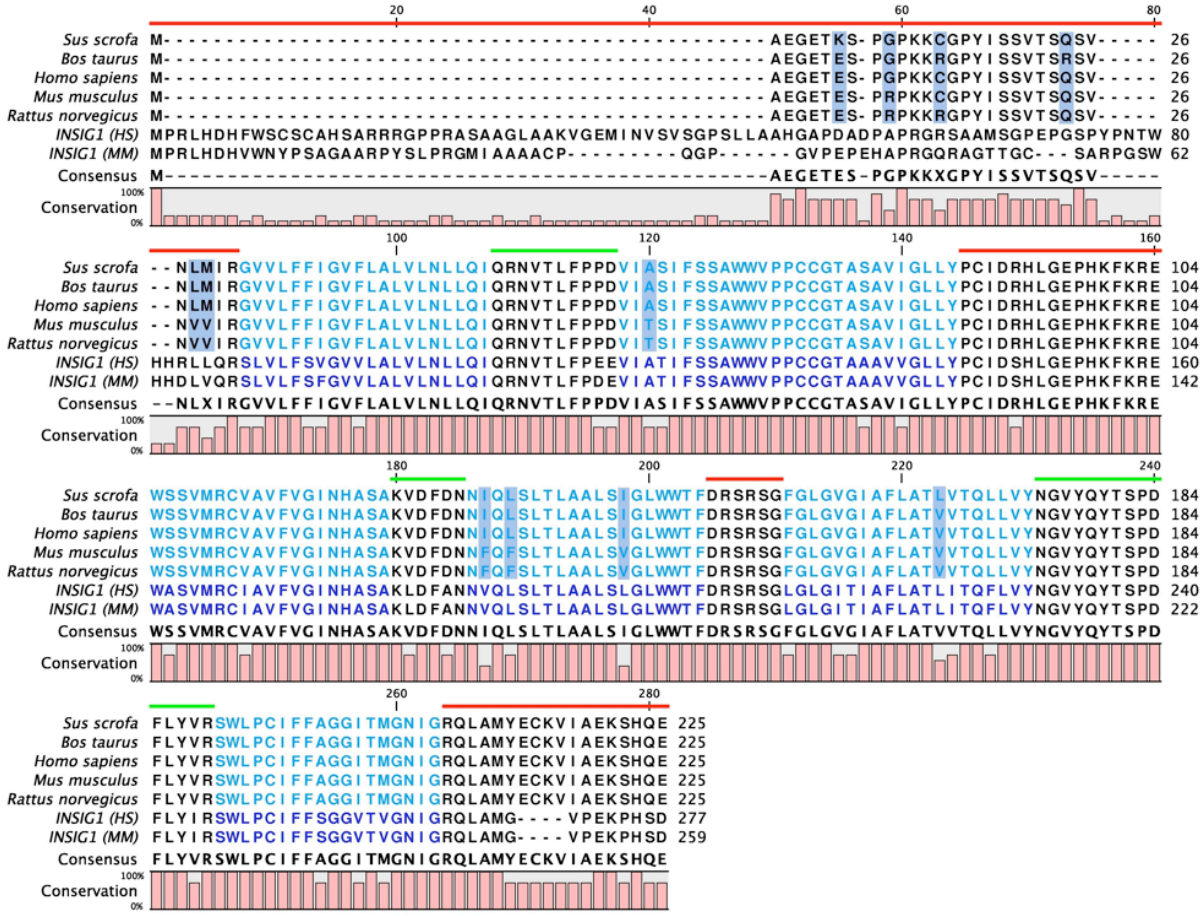


Figure 7 Alignment of INSIG2 protein sequences from several species.

Sequence alignment of INSIG2 protein from *Sus scrofa* (translated from predicted mRNA sequence), *Bos taurus* (XP_614207, NCBI), *Homo sapiens* (NP_057217, NCBI), *Mus musculus* (NP_598509, NCBI), *Rattus norvegicus* (NP_835192, NCBI) and of INSIG1 protein from *Homo sapiens* (abbreviation HS, NP_005533, NCBI) and *Mus musculus* (abbreviation MM, NP_705746, NCBI). Numbers on the right side indicate amino acid position. Light blue and dark blue letters show transmembrane domains in INSIG2 and INSIG1, respectively. Red lines above alignment indicate cytoplasmatic domains, green lines luminal domains. Blue backgrounds of letters indicate different residues of INSIG2 alignment. Ambiguous residues in the consensus sequence are indicated by the X symbol.

The derived amino acid sequence is in 89% identical to the sequence of human FTO. All conserved residues of the Fe(II) and 2-oxoglutarate binding site reported by Gerken *et al*. as well as Sanchez-Pulido and Andrade-Navarro (Gerken *et al.* 2007; Sanchez-Pulido & Andrade-Navarro 2007) were found in the porcine protein sequence (Figure 8).

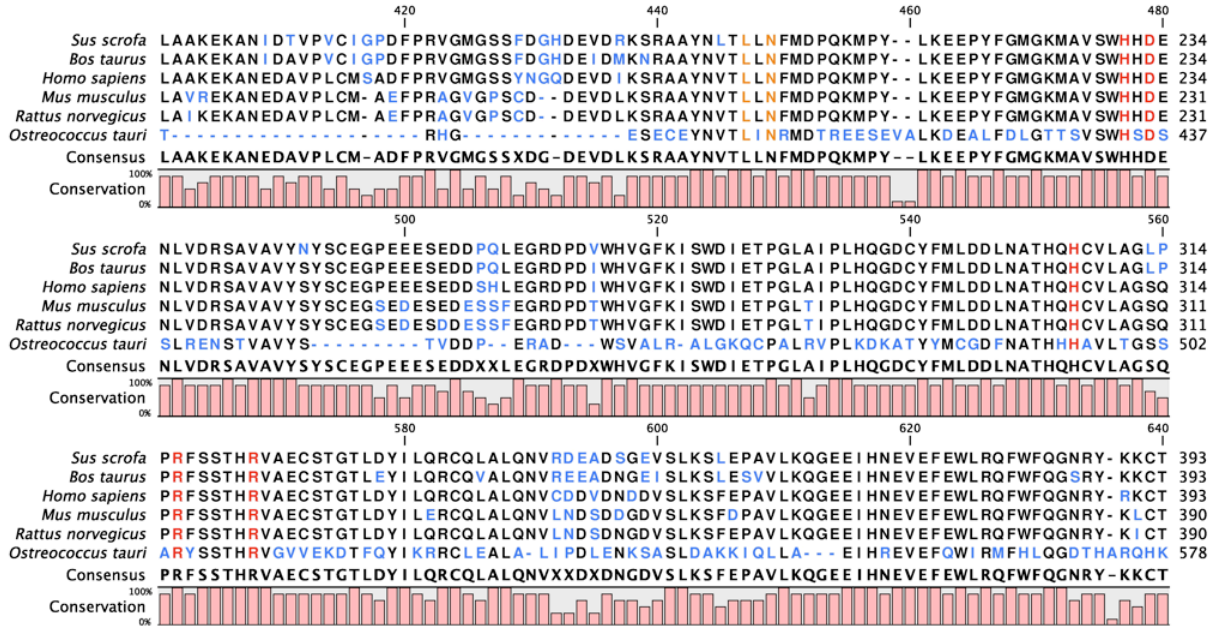


Figure 8 Partial alignment of *FTO* protein sequences from several species.

Sequence alignment of FTO protein from *Sus scrofa* (translated from predicted mRNA sequence), *Bos taurus* (NP_001091611, NCBI), *Homo sapiens* (NP_001073901, NCBI), *Mus musculus* (NP_036066, NCBI*), Rattus norvegicus* (NP_001034802, NCBI) and *Ostreococcus tauri* (CAL57236, NCBI). Numbers on the right side indicate amino acid position. Blue letters indicate different residues in alignment. Conserved residues according to Gerken *et al*. (2007) are highlighted in red, according to Sanchez-Pulido and Andrade-Navarro (2007) in red and orange. Ambiguous residues in the consensus sequence are indicated by the X symbol.

The coding sequence, flanking splice junctions and the putative promoter region of *INSIG2* were screened for polymorphisms in the parental animals of the cross. In total, 4995 bp were re-sequenced resulting in the detection of 13 variants, including 11 SNPs and two insertion-deletion polymorphisms (Table 20). A synonymous substitution (c.237A>G) has been detected in the coding region at codon Thr79. In the parental animals the synonymous substitution is in linkage disequilibrium (LD) with c.-180C>T transition located within the putative SP1 binding site. In the Piétrain sows the c.237A and c-180C alleles were fixed, whereas in the Mangalitsa founders both alleles were segregating.

Within the microsatellite located in the 3' UTR GT repetitions are polymorphic (*121_122insGT). Two alleles were found with 16 or 17 GT repeats. The longer allele (L) with a length of 54 bp descended from a heterozygous Mangalitsa boar and a heterozygous Piétrain sow. However, only the paternal allele was transmitted to the next generation. The shorter allele is designated as S.

Table 20 Sequence variants in porcine *INSIG2* and *FTO*.

Polymorphisms of *INSIG2* and *FTO* with localisation, indication of minor allele (MA), minor allele frequency (MAF) in the parental animals of Mangalitsa x Piétrain cross and description. The position of polymorphisms is given according to the sequence deposited in the GenBank FJ853992 for *INSIG2* and FJ853994 for *FTO*.

No.	Variant[1]	Localisa-tion	Type	MA	MAF	Description
		INSIG2				
1	c.-212-669T>C	Promoter	SNP	C	0.036	
2	c.-212-340G>T	Promoter	SNP	T	0.036	
3	c.-212-84G>C	Promoter	SNP	C	0.036	
4	c.-180C>T	Exon 1	SNP	T	0.036	5' UTR, Putative SP1 binding site
5	c.237A>G	Exon 2	SNP	G	0.036	Synonymous substitution
6	c.245-303A>G	Intron 2	SNP	G	0.036	
7	c.245-266A>G	Intron 2	SNP	G	0.286	
8	c.245-207C>G	Intron 2	SNP	G	0.036	
9	c.245-165C>G	Intron 2	SNP	G	0.036	
10	c.537-159delA	Intron 4	DEL	-	-	
11	*121_122delGT	3' UTR	DEL	I	0.071	Microsatellite
12	*423G>A	3' UTR	SNP	A	0.286	
13	*442A>G	3' UTR	SNP	G	0.036	
		FTO				
1	c.-1034T>C	Promoter	SNP	T	0.036	
2	c.-678A>G	Promoter	SNP	G	0.036	
3	c.-449C>G	Promoter	SNP	G	0.036	
4	c.-167T>G	Promoter	SNP	T	0.393	
5	c.28C>A	Exon 1	SNP	G	0.036	Synonymous substitution
6	c.46-143C>T	Intron 1	SNP	T	0.179	
7	c.46-43G>A	Intron 1	SNP	A	0.179	
8	c.46-12C>A	Intron 1	SNP	A	0.179	
9	c.213G>A	Exon 3	SNP	A	0.107	Synonymous substitution
10	c.528C>T	Exon 3	SNP	T	0.036	Synonymous substitution
11	c.594G>C	Exon 3	SNP	C	0.357	Synonymous substitution
12	c.751+91A>G	Intron 3	SNP	G	0.357	
13	c.751+157G>T	Intron 3	SNP	T	0.357	
14	c.752-45C>G	Intron 3	SNP	G	0.036	
15	c.752-42G>A	Intron 3	SNP	A	0.393	
16	c.752-36C>T	Intron 3	SNP	T	0.357	
17	c.858G>A	Exon 4	SNP	A	0.036	Synonymous substitution
18	c.895+22G>T	Intron 4	SNP	T	0.357	
19	c.895+58G>T	Intron 4	SNP	T	0.036	
20	c.895+128G>T	Intron 4	SNP	T	0.357	
21	c.895+191C>T	Intron 4	SNP	T	0.357	
22	c.895+216C>A	Intron 4	SNP	A	0.357	
23	c.895+232T>C	Intron 4	SNP	C	0.357	
24	c.895+233G>A	Intron 4	SNP	A	0.357	
25	c.976-102C>T	Intron 5	SNP	T	0.036	
26	c.976-98T>C	Intron 5	SNP	C	0.036	
27	c.976-22A>G	Intron 5	SNP	G	0.036	
28	c.1080G>A	Exon 6	SNP	A	0.036	Synonymous substitution
29	c.1119+70A>G	Intron 6	SNP	G	0.036	
30	c.1119+75G>A	Intron 6	SNP	A	0.036	
31	c.1119+98G>A	Intron 6	SNP	A	0.036	
32	c.1239+38A>T	Intron 7	SNP	T	0.179	
33	c.1239+60C>G	Intron 7	SNP	G	0.036	
34	c.1240-26C>T	Intron 7	SNP	T	0.036	
35	c.1240-25_1240-24delTT	Intron 7	DEL	A	0.036	
36	c.1395T>C	Exon 9	SNP	C	0.036	Synonymous substitution

37	*79G>A	3' UTR	SNP	A	0.143	
38	*100C>T	3' UTR	SNP	T	0.143	
39	*140G>A	3' UTR	SNP	A	0.143	
40	*158A>C	3' UTR	SNP	C	0.036	
41	*243C>A	3' UTR	SNP	A	0.036	
42	*275C>G	3' UTR	SNP	G	0.071	
43	*456_466delAATTTAAATTC	3' UTR	DEL	A	0.143	
44	*469delT	3' UTR	DEL	A	0.143	
45	*981A>T	3' UTR	SNP	T	0.036	
46	*988C>G	3' UTR	SNP	G	0.143	
47	*1057_1060delTAAA	3' UTR	DEL	A	0.179	Putative AAATAAA
48	*1176C>T	3' UTR	SNP	T	0.179	
49	*1213G>T	3' UTR	SNP	T	0.036	
50	*1257C>T	3' UTR	SNP	T	0.036	
51	*1288T>C	3' UTR	SNP	C	0.143	
52	*1305T>C	3' UTR	SNP	C	0.143	
53	*1430T>G	3' UTR	SNP	G	0.036	

For the association study we selected the synonymous substitution and the microsatellite polymorphism as two potentially functional variants. Synonymous substitutions as well as microsatellite polymorphisms in the 3' UTR are described as possible factors influencing mRNA stability (Tabor *et al.* 2002; Citores *et al.* 2004; Lee *et al.* 2004; Nackley *et al.* 2006; Chen *et al.* 2007b).

We studied 306 F2 animals. The synonymous substitution and microsatellite polymorphism are in linkage disequilibrium (LD) in the F2 animals and thus in the presentation of the results we concentrated only on the microsatellite mutation. In the investigated F2 animals the frequency of the longer allele of the microsatellite is 0.14. We observed a significant association with thinnest backfat thickness, saturated fatty acid content and polyunsaturated fatty acid content (Table 21). The shorter allele increases thinnest backfat thickness and saturated fatty acid content, while it decreases polyunsaturated fatty acid content.

The allele frequency of the microsatellite polymorphism was examined in four commercial pig breeds (Table 22). The longer allele is quite infrequent in the Piétrain; nevertheless, we genotyped 544 animals with an average weight at slaughter of 103 kg. In the investigated Piétrain animals the frequency of the longer allele is 0.02 and the polymorphism is in the Hardy-Weinberg equilibrium (P-value=1). No association was found between the microsatellite polymorphism and the investigated traits in the Piétrain population (Table 21). The c.237A allele of the synonymous substitution was fixed in 1204 Piétrain animals. None of the associations are significant after the Bonferroni correction for multiple testing.

Table 21 Association of the *INSIG2* microsatellite polymorphism with lipid deposition and fatty acid metabolism in the F2 population of Mangalitsa x Piétrain cross and in Piétrain.

The shorter allele is designated as S, the longer as L. Investigated traits: backfat thickness in middle of back (BFTM), backfat thickness at shoulder (BFTS), thinnest backfat thickness (TBFTM), loin muscle area (LMA), intramuscular fat content (IMF), saturated fatty acids (SFA) content, monounsaturated fatty acids (MUFA) content, polyunsaturated fatty acids (PUFA) content and n-3 polyunsaturated fatty acids (n-3PUFA) content; fixed effect and covariates used in linear model, outliers (Outl), transformation of data (Transf), least square means (LSM) with standard errors (SE) and nominal P-value are presented. In Piétrain breeding value of IMF, MUFA and n-3PUFA are shown. Significant results are indicated in bold.

Trait	Fixed effect Covariates	Outl	Transf	LSM (SE)		P-value
			F2 animals			
				SS (n=232)	SL (n=92)	
BFTM [mm]	dam, sex, weight, days	<0	-	29.211 (0.365)	29.282 (0.674)	0.932
BFTS [mm]	sire, sex, weight, days	<22	log	45.300 (0.356)	45.475 (0.618)	0.811
TBFTM [mm]	dam, sex, weight	<4	-	85.092 (0.344)	83.340 (0.633)	**0.025**
LMA [mm^2]	dam, sex, weight, days, *RYR1*	<15 >96	log	46.813 (0.316)	46.005 (0.556)	0.135
IMF [%]	dam, sex, father, *RYR1*	<0	log	2.041 (0.040)	1.948 (0.073)	0.101
SFA [%]	dam, sire, IMF, sex, weight, days	<30 >60	-	36.643 (0.082)	36.245 (0.151)	**0.033**
MUFA [%]	sex, IMF, *RYR1*	<30 >60	-	47.140 (0.143)	46.639 (0.237)	0.073
PUFA [%]	sire, sex, days, IMF	<0	log	14.803 (0.154)	15.653 (0.263)	**0.008**
n-3PUFA [%]	dam, sire, IMF, sex, weight, days	<0 >3.7	log	1.283 (0.024)	1.307 (0.044)	0.662
			Piétrain			
				SS (n=521)	SL (n=23)	
BFTM [mm]	station, *RYR1*	>2.5	-	1.419 (0.013)	1.403 (0.060)	0.799
BFTS [mm]	sire, station, weight, *RYR1*	>5.0	-	2.793 (0.016)	3.002 (0.120)	0.093
TBFTM [mm]	sire, station, weight	-	-	8.444 (0.136)	8.588 (0.018)	0.306
LMA [mm^2]	sire, station, weight, *RYR1*	-	-	62.587 (1.482)	64.856 (0.200)	0.140
IMF [ebv]	sire, *RYR1*			0.003 (0.005)	-0.012 (0.038)	0.702
MUFA [ebv]	sire, IMF			0.091 (0.025)	0.155 (0.183)	0.735
n-3PUFA [%]	sire, IMF, *RYR1*	< -0.52		-0.012 (0.002)	-0.027 (0.014)	0.328

Table 22 Allele frequency of the *INSIG2* microsatellite polymorphism in different pig breeds.

	Number of animals			*Allele frequency*		*HWE*
	SS	**SL**	**LL**	**S**	**L**	*P-value*
Piétrain	52	3	0	0.97	0.03	1
German Landrace	56	5	0	0.96	0.04	1
Large White	8	0	0	1.00	0.00	1
Duroc	10	4	0	0.86	0.14	1

FTO was re-sequenced by following the same procedure as for *INSIG2*. In total, 8012 bp were re-sequenced, resulting in the detection of 53 polymorphisms, including 49 SNPs and 4 deletions (Table 20). We have identified seven synonymous substitutions in codons Arg10, Ala71, Phe176, Ala198, Ala286, Glu360 and Pro465. In the 5'-region 4 SNPs and in the 3'-region 17 variants were detected. Polymorphisms within the 1kb region upstream of the translation start site do not affect binding sites of any putative transcription factors. *1057_1060delTAAA is located within a putative poly(A) signal. The Mangalitsa boars were monomorphic regarding all detected variants. The majority of the detected polymorphisms are located in the non-coding regions as intronic or 5' and 3' -regions. Thus, we gave priority to the polymorphisms regarding the information about the location and type of the sequence variants (Tabor *et al.* 2002). We selected three SNPs for genotyping 508 F2 animals: c.46-143C>T in intron 1, c.594G>C in exon 3 and *1176C>T in the 3'-region.

The c.46-143T allele of the c.46-143C>T variant was fixed in the Mangalitsa founders and the c.46-143C allele was overrepresented in Piétrain sows. Only one Piétrain sow was heterozygous CT. The c.594G>C synonymous mutation in exon 3 is linked with nine other variants adjacent to exons 3 and 4. The c.594G allele was fixed in the Mangalitsa boars, while both alleles were segregating in Piétrain sows. The *1176C>T polymorphism is in LD with *1057_1060delTAAA located in the putative poly(A) signal. Mangalitsa boars were homozygous *1176CC.

Both *RYR1* and *FTO* genes were mapped to SSC6 with a distance of approximately 20 cM between each other. None of the genotyped SNPs were in LD with the Cys614 *RYR1* polymorphism (data not shown); nevertheless, *RYR1* with a major effect on meat and carcass traits (Fujii *et al.* 1991) was included in the model as a cofactor to decrease the residual error variance and thereby increase the statistical power in this study.

Table 23 Association of *FTO* polymorphisms (1) c.46-143C>T, (2) c.594G>C and (3) *1176C>T with lipid deposition and fatty acid metabolism in the F2 population of Mangalitsa x Piétrain cross.

The A sign correspond to c.46-143C, c.594C, *1176C alleles and B indicated the second allele of investigated SNPs. Investigated traits: backfat thickness at mid-back (BFTM), backfat thickness at shoulders (BFTS), thinnest backfat thickness (TBFTM), loin muscle area (LMA), intramuscular fat content (IMF), saturated fatty acids (Hughes *et al.*) content, monounsaturated fatty acids (MUFA) content, polyunsaturated fatty acids (PUFA) content and n-3 polyunsaturated fatty acids (n-3PUFA); fixed effect and covariates used in linear model, outliers (Outl), transformation of data (Transf), SNP number, least square means (LSM) with standard errors (SE) and nominal P-value are presented. Significant results are indicated in bold.

Trait	**Fixed effect Covariates**	**Outl**	**Tran-sf**	**SNP**	**LSM (SE)**			**P-value**
					AA	AB	BB	
Numbers of observation				1	CC=250	CT=84	TT=174	
				2	CC=37	CG=189	GG=282	
				3	CC=482	CT=26	-	
BFTM [mm]	dam, sex, weight, days	<12 >50	-	1	29.205 (0.333)	29.164 (0.576)	28.883 (0.393)	0.820
				2	29.685 (0.897)	29.108 (0.385)	28.996 (0.319)	0.783
				3	29.147 (0.236)	27.985 (1.410)	-	0.429
BFTS [mm]	dam, sex, weight, days	<16	log	1	45.419 (0.369)	44.897 (0.634)	44.437 (0.437)	0.248
				2	47.350 (0.997)	45.410 (0.423)	44.420 (0.353)	**0.020**
				3	44.866 (0.261)	47.346 (1.563)	-	0.128
TBFTM [mm]	dam, sex, weight	<4	-	1	84.926 (0.320)	83.901 (0.555)	83.862 (0.380)	0.080
				2	83.684 (0.858)	84.505 (0.372)	84.409 (0.310)	0.664
				3	84.328 (0.227)	85.620 (1.359)	-	0.360
LMA [mm^2]	sire, sex, weight, days, *RYR1*	<20 >76	log	1	46.520 (0.345)	48.182 (0.556)	47.175 (0.437)	**0.048**
				2	45.583 (0.883)	47.357 (0.388)	46.985 (0.326)	0.201
				3	47.033 (0.224)	46.727 (0.987)	-	0.867
IMF [%]	dam, sire, sex, weight, *RYR1*	<0	log	1	2.058 (0.041)	2.106 (0.066)	2.044 (0.051)	0.909
				2	1.929 (0.111)	2.051 (0.046)	2.085 (0.039)	0.366
				3	2.071 (0.026)	1.873 (0.152)	-	0.349
SFA [%]	dam, sire, sex, days, imf, *RYR1*	<30 >60	-	1	36.636 (0.093)	36.545 (0.149)	36.783 (0.115)	0.435
				2	37.066 (0.250)	36.700 (0.103)	36.600 (0.089)	0.271
				3	36.680 (0.058)	36.513 (0.354)	-	0.651
MUFA [%]	sex, IMF, *RYR1*	<30 >60	-	1	47.317 (0.144)	47.071 (0.242)	47.334 (0.179)	0.634
				2	46.466 (0.372)	47.354 (0.162)	47.344 (0.134)	0.075
				3	47.297 (0.099)	47.010 (0.436)	-	0.521
PUFA [%]	sire, sex, IMF	>0	log	1	14.792 (0.174)	14.974 (0.297)	14.313 (0.213)	0.199
				2	15.650 (0.456)	14.832 (0.201)	14.411 (0.167)	0.111
				3	14.659 (0.121)	14.634 (0.526)	-	0.983
n-3PUFA [%]	dam, sire, sex, IMF	>0 <3.7	log	1	1.280 (0.024)	1.270 (0.042)	1.213 (0.029)	0.300
				2	1.290 (0.066)	1.284 (0.028)	1.231 (0.024)	0.526
				3	1.256 (0.017)	1.238 (0.098)	-	0.769

We observed a significant association between the synonymous substitution in exon 3 c.594G>C and backfat thickness at shoulders (Table 23) in the F2 population of the Mangalista x Piétrain cross. Subsequently, we evaluated the effect of the porcine FTO c.594G>C gene polymorphism on carcass and meat quality traits in Piétrain pigs. The observed association in the F2 animals was not replicated in the Piétrain population. We observed a significant association between the c.594G>C polymorphism with loin

muscle area in the Piétrain pigs (Table 24).

Table 24 Association of c.594G>C *FTO* polymorphisms with lipid deposition and fatty acid metabolism in the Piétrain population.

Investigated traits: backfat thickness at mid-back (BFTM), backfat thickness at shoulders (BFTS), thinnest backfat thickness (TBFTM), loin muscle area (LMA), breeding value for intramuscular fat (IMF), breeding value for monounsaturated fatty acids (MUFA), breeding value for n-3 polyunsaturated fatty (n-3PUFA); fixed effect and covariates used in linear model, transformation of data (Transf), SNP number, least square means (LSM) with standard errors (SE) and nominal P-value are presented. Significant results are indicated in bold.

Trait	Fixed effect Covariates	Transf	LSM (SE)			P-value
			CC (n=266)	CG (n=527)	GG (n=226)	
BFTM [mm]	station, weight, *RYR1*	-	1.446 (0.017)	1.450 (0.012)	1.428 (0.018)	0.604
BFTS [mm]	station, weight, *RYR1*	-	2.827 (0.023)	2.809 (0.016)	2.825 (0.024)	0.769
TBFTM [mm]	station, weight, *RYR1*	-	8.612 (0.028)	8.604 (0.020)	8.588 (0.030)	0.832
LMA [mm^2]	station, weight, *RYR1*	log	64.013 (0.301)	64.800 (0.213)	63.929 (0.327)	**0.023**
IMF (ebv)	*RYR1*	-	0.019 (0.010)	0.011 (0.007)	0.005 (0.010)	0.670
MUFA (ebv)	IMF	-	0.077 (0.031)	0.138 (0.022)	0.143 (0.034)	0.234
n-3PUFA (ebv)	IMF, *RYR1*	-	-0.023 (0.003)	-0.026 (0.002)	-0.021 (0.003)	0.675

Discussion

INSIG2 and *FTO* were selected as candidate genes for fatness related traits in pigs because of the significant association with BMI in the human. Additionally, both genes are located in QTL regions for lipid deposition and fatty acid metabolism in swine. The candidate analysis of porcine *INSIG2* and *FTO* presented here identified 11 polymorphisms in *INSIG2* and 53 mutations in *FTO*. For association studies the sequence variants were prioritized according to their functional and positional significance. The association of the microsatellite polymorphism in 3' UTR of *INSIG2* with backfat thickness, as observed in the F2 animals, could not be replicated in the Piétrain pigs, possibly due to a low minor allele frequency (0.02) in the Piétrain animals. In the Piétrain population we could not test the association between the microsatellite polymorphism and SFA and PUFA content due to a lack of phenotypic data. The association of the synonymous substitution c.594G>C of FTO and backfat thickness at shoulders in the F2 animals could not be confirmed in the Piétrain breed. In the Piétrain population a significant association was observed between the c.594G>C polymorphism with loin muscle area.

Chapter 5

General discussion

General approach

In this study two main strategies were applied for finding trait loci involved in lipid and energy metabolism in swine: QTL analysis and the candidate gene approach. The analysis of QTLs affecting lipid deposition, growth, meat quality and fatty acid composition traits was performed in the Mangalitsa x Piétrain cross. Parallel with the QTL study, a functional candidate gene (*PDK4*), which product is involved in the control of three energy pathways, as well as two human obesity susceptibility genes (*INSIG2* and *FTO*) were investigated. When this work was initiated, no prior reports on the porcine *PDK4*, *INSIG2* and *FTO* genes were available. Therefore, firstly candidate genes were characterized and screened for sequence variants. Finally, association studies were carried out in the cross and commercial populations.

QTL study

The QTL approach will reveal a map location of a trait with a major effect if an accurate genetic model has been applied, a reasonable sample size has been used and the marker set provides full genome coverage. A trait with a minor effect will not be detected if it does not reach the stringent significant threshold (Andersson 2001). Although not all susceptibility genes were identified in linkage studies, the genes that account for reproducible linkage peaks in genome scans probably represent major susceptibility genes (Hattersley & McCarthy 2005).

The genome scan reported in this thesis was conducted within a crossbreed between two highly divergent breeds relative to lipid and energy traits. The population of genetically and phenotypically divergent ancestors provides a valuable source for QTL mapping. However, QTLs identified in experimental crosses between divergent breeds are not necessarily relevant or may not segregate in commercial populations, therefore a successful implementation of marker assisted selection (MAS) in pig breeding remains limited. Only major genes, such as *RYR1*, were implemented in MAS.

The investigated F2 population size was relatively low, which might prevent the detection of additional QTLs with minor effects. Small sample sizes have limited power to detect minor QTLs and are more suitable to find QTLs with a major effect, such as *RYR1* (Andersson *et al.* 1994; Wondji *et al.* 2007). In fact, this study has confirmed the already described QTLs affecting backfat thickness, growth and meat quality traits on

several chromosomes. Only few novel 5% chromosome-wide QTLs were reported in the Mangalitsa x Piétrain cross: a QTL for average feed intake on SSC 7, a QTL for intramuscular fat content on SSC 1, a QTL for pH on SSC 10, a QTL for saturated fatty acid content on SSC 17 and all QTLs for monounsaturated and polyunsaturated fatty acid contents as well as the 1% chromosome-wide QTL for carcass length on SSC 15.

A total of 117 genetic markers selected from the USDA-MARC v2 linkage map covering 18 porcine autosomes (Rohrer *et al.* 1996) were used. The informativeness of a genetic marker for linkage studies is often presented by the value of polymorphism information content (PIC). In QTL analysis it is important for the PIC of markers to be high. In our study only 11.96% markers were fully informative and 69.23% exceed the 0.5 threshold of PIC. From the set of 117 markers 5.29% markers were uninformative (PIC=0). Including all markers the average information content was 0.64. Additionally, six gaps with an average length of 53.9 cM appeared on SSCs 6, 7, 10 and 15. Thus, the density of the marker map could greatly reduce the power of QTL detection and the effect could be underestimated.

In an ideal situation all markers were fixed for different alleles in the parental lines. However, in this study the fixation assumption of alleles is violated and the alleles at the QTL are segregating in either or both of the parental lines. To deal with this problem the qxpak analysis was carried out. Only three additional QTLs were detected on SSC 2. Additionally, in order to increase the power of an unknown QTL detection a major gene (*RYR1*) affecting meat quality and carcass traits as well as edema and post-weaning diarrhea susceptibility gene (*FUT1*) were taken into account (Fujii *et al.* 1991; Meijerink *et al.* 1997).

A comparison of results between studies appears to be difficult. The main reasons for this include low consistency of trait nomenclature, different marker maps, poor precision in QTL mapping and thus large CI sometimes covering all chromosome length. Additionally, the extended CI of the QTLs greatly affect search for candidate genes. The low accuracy of the estimated QTL position increased the number of potential candidate genes. Therefore, fine mapping of the regions is necessary to restrict the region of interest and thus the number of possible candidate genes, unless the causative mutation occurs in an obvious and well-studied candidate gene within a CI

(Andersson 2001). A further characterisation of many identified QTLs will be facilitated by the access to the genome sequences and the re-sequencing of QTL intervals if the map resolution is sufficiently good (Andersson & Georges 2004). To refine the map position of a QTL further breeding experiments can be used. In each subsequent generation, the block of linkage disequilibrium that surrounds the QTL is gradually reduced by recombination (Andersson & Georges 2004). However, a disadvantage of this approach is the high cost of maintaining the experimental population and genotyping of a large number of markers. Additionally, individuals with missing genotypes are excluded from the study.

When the number of the reported QTLs in livestock genomes with the number of identified quantitative trait nucleotides (QTNs) is compared, the suggestive reflection is not encouraging. This opinion is supported by the fact that in the swine genome there are almost 1700 QTLs for fatness, meat quality and production traits and only a limited number of these QTLs have been further investigated to the point that a known causative mutation has been implicated or proven (Rothschild *et al.* 2007). Therefore, as an alternative to the genome-wide linkage mapping strategy, the candidate gene approach has also been widely used to identify the genetic basis for complex quantitative traits.

Candidate gene approach

Candidate genes are selected based on previous knowledge on risk factors for traits of interest as well as biological understanding of genes and proteins that are likely to be involved in the investigated phenotype (Tabor *et al.* 2002). The candidate gene strategy for finding trait loci can be powerful and can detect loci even with small effects. However, this approach to evaluate all of candidate genes for each trait of interest might be more time-consuming. Furthermore, the knowledge on gene function is still poor and hence the candidate gene strategy might fail to identify a major locus. Additionally, the presence of linkage disequilibrium to a linked or non-linked causative gene and an improperly adjusted threshold for multiple testing increase the risk of false positive associations in candidate gene studies (Andersson 2001).

Candidate gene: *PDK4*

PDK4 was chosen as a candidate gene for energy and lipid metabolism because of the role of PDK4 in three important energy pathways: glycolysis, fatty acid oxidation and synthesis, as well as its localisation in a QTL region for backfat thickness, daily gain, feed intake and saturated fatty acids. When this work was initiated, no a priori association analysis in the porcine *PDK4* gene was available. The candidate gene analysis presented here revealed an association of the porcine *PDK4* variants with average feed intake (AFI) and saturated fatty acid (SFA) content in the F2 population of the Mangalitsa x Piétrain cross. The insertion-deletion polymorphism of SINE shows an association with AFI as well as SFA content and the non-synonymous substitution with SFA content. A further association between the insertion-deletion polymorphism of SINE and AFI was studied in the F3 animals and the German Landrace population. In the F3 animals a suggestive association was observed, while in the German Landrace no association was detected, possibly due to a lower minor allele frequency in the commercial population.

The homozygous F2 animals with the insertion of SINE exhibited a lower AFI and a lower PDK4 expression in muscle and fat in the F3 generation. The homozygous F2 animals without the SINE exhibited a higher AFI and a higher PDK4 expression in muscle and fat in the F3 animals. Findings from association studies in the cross, when considered together with the results from the PDK4 expression study, are contradictory. The higher amount of PDK4 causes an inactivation of PDC due to the higher phosphorylation of the complex. Subsequently, when PDC is inactive, acetyl-CoA is derived from ß-oxidation, thus glucose level is preserved. Hence, we could assume that animals with a higher expression of PDK4 would exhibit a lower AFI. However, it should be stressed that differences in the expression showed only a tendency and were not even suggestive.

Lan *et al.* (Lan *et al.* 2008) reported on a SNP c.967-40A>G detected in four individuals representing the Large White and Chinese Meishan pig breeds, which shows an association with the ratio of lean to fat, intramuscular fat and muscle water content. In the parental generation of the Mangalitsa x Piétrain cross c.967-40A>G polymorphism is in LD with the investigated insertion-deletion polymorphism of SINE, which did not exhibit an association with any investigated backfat and meat quality traits.

Candidate gene: *INSIG2*

INSIG2 was chosen as a candidate gene for fat deposition because it was previously indicated as a human obesity susceptibility gene. Identification of genes influencing related traits in the human offers candidates for testing in animals. The product of *INISG2* plays an important role in the cholesterol homeostasis by binding to Scap or HMG CoA reductase in sterol-overloaded cells. When this work was initiated, no prior association analysis in the porcine *INSIG2* gene was available. The systematic search for sequence variants in the *INSIG2* revealed possible functional polymorphisms in the gene: the microsatellite polymorphism in 3' UTR and the synonymous substitution c.237A>G in LD with the promoter polymorphism c.-180C>T located within the putative SP1 binding site. Additionally, both potentially functional variants are in LD in the F2 animals.

The association of the microsatellite polymorphism with the thinnest backfat thickness was observed in the F2 animals of the Mangalitsa x Piétrain cross. The association could not be replicated in the Piétrain population, possibly due to the low minor allele frequency (0.02) in the Piétrain breed. We could not test the association between microsatellite polymorphism and SFA as well as PUFA content, as observed in the F2 animals, due to a lack of phenotypic data in the Piétrain population. In the present study, it remains unclear whether or not *INSIG2* causes the observed association. Thus, further investigations are advisable in a population where the microsatellite allele is segregating with a higher frequency. However, it is often problematic to collect fatty acid data due to the high cost and time-consuming procedure and therefore the validation of the results could be difficult.

The extensive linkage in QTL mapping populations could not exclude the possibility that the observed association of the microsatellite polymorphism with fatty acid composition traits and thinnest backfat thickness is due to QTLs located at some distance from *INSIG2* (Zhao *et al.* 2003). However, the mutation within the microsatellite could be responsible for mRNA stability (Citores *et al.* 2004; Lee *et al.* 2004; Chen *et al.* 2007b) and it is an intriguing possibility that the microsatellite motifs in the 3' UTR of porcine *INSIG2* might participate in the pre- and post-transcriptional regulation of INSIG2 expression. Additionally, the conservation across all Artiodactyles

and the convergent position in the 3' UTR indicated a probably important function in this region of the gene. Thus, the functional evidence as the analysis of mRNA levels could strongly support the hypothesis.

No QTLs on SSC 15 for SFA, MUFA and PUFA were discovered in the Mangalitsa x Piétrain cross. Only 5% chromosomal significant QTLs for average backfat thickness and backfat thickness at loin, but no QTLs for the thinnest backfat thickness, were detected in a proximal part of SSC15, where *INSIG2* is located. Nevertheless, the small number of investigated animals in the QTL study could prevent the detection of QTLs with minor effects (Wondji *et al.* 2007).

Candidate gene: *FTO*

FTO was chosen as a candidate gene for fat deposition because recently it was indicated through a genome-wide association scan as a human obesity susceptibility gene (Dina *et al.* 2007a; Frayling *et al.* 2007; Scuteri *et al.* 2007). The candidate gene analysis presented in this study identified only synonymous substitutions in the coding sequence and numerous variants in the non-coding region, such as intronic or 5' and 3' -regions. Theoretically, any of these polymorphisms could be located in unknown regulatory elements and have potential functional effects. However, it is often difficult to assess their function on the basis of the nucleotide sequence only (Tabor *et al.* 2002). On the other hand, the synonymous SNPs were found to modulate the protein expression by altering the mRNA secondary structure (Nackley *et al.* 2006). Thus, the sequence variants were prioritized according to their possible function, position in the investigated gene sequence and regarding to allelic segregation in the parental lines of the cross (Tabor *et al.* 2002).

An association was identified for a synonymous mutation c.594G>C with backfat thickness at shoulder in the F2 animals of the Mangalitsa x Piétrain cross. Allele C originating from the Piétrain breed is associated with increased backfat thickness. The observed association in the F2 population was not replicated in the Piétrain pigs. However, the allele frequency in both investigated populations is highly divergent. The frequency of allele C in the F2 animals is 0.26, while in the Piétrain pigs it is 0.52. The inconsistent effects observed in the cross and the Piétrain population could indicate both that the investigated sequence variant is in LD with the putative casual mutation and

that the different LD structure might exist among the investigated pig populations.

Two polymorphisms detected in Duroc pigs in intron 4 of *FTO* (Fontanesi *et al.* 2008) were also identified in the parental animals of our cross (g.170T>G correspond to c.895+22G>T, g.276T>G correspond to c.895+128G>T) and were in LD to c.594G>C genotyped in our study. We have not identified polymorphism in the 3'-region mentioned by Fontanesi *et al.* (Fontanesi *et al.* 2008); however, 17 different variations were described. Recently, Fontanesi *et al.* (Fontanesi *et al.* 2008), described the association between g.276T>G polymorphism in intron 4 and visible intermuscular fat in the Italian Duroc population. In our study we have not detected the association with IMF and synonymous substitution c.594G>C, which is in LD with reported g.276T>G (Fontanesi *et al.* 2008), only with backfat thickness at shoulders. Interestingly, Yue *et al*. (Yue *et al.* 2003) reported a QTL for fat thickness at shoulder at the 5% chromosome-wide level in a Meishan x Piértain cross, where *FTO* is located. We have not identified a QTL for backfat thickness at shoulders in the Mangalitsa x Piétrain cross on SSC 6, irrespective of the fact whether the *RYR1* genotypes were included or not in the QTL model. However, in an F2 intercross a minor effect of the candidate gene fixed for different alleles may be entirely confounded with the effect of the linked major QTL. Additionally, a high phenotypic variance of QTLs with large effects such as *RYR1* could reduce the power to detect minor loci (Park *et al.* 2002; Hirschhorn & Daly 2005). Furthermore, the investigated F2 population size was relatively low, which might prevent the detection of additional QTLs with minor effects. Small sample sizes have limited power to detect minor QTLs and are more suitable to find QTLs with a major effect, such as *RYR1* (Wondji *et al.* 2007). Additionally, in our QTL mapping studies in the putative location of the *FTO* gene the 61.9 cM interval between two flanking markers (SW2406-SW122) was described. Therefore, the performed linkage study could fail to detect the effect of this locus due to the lack of power. Furthermore, a QTL study on SSC 6 in a Wild Boar x Meishan cross, where only the *RYR1* C allele was present, indicated that additional loci in the *RYR1* region or possibly unrecognized alleles of this locus were involved in the variation of carcass and growth traits (Yue *et al.* 2003). Without the mutant *RYR1* T allele segregating in the Wild Boar x Meishan cross, only minor QTL effects were found (Yue *et al.* 2003). These findings support the importance of the investigated polymorphism and more extensive association studies are recommendable.

Recently, Fan *et al.* (Fan *et al.* 2009) reported a significant association of c.594G>C polymorphism with lumbar backfat, total lipid percentage in muscle and a suggestive association with tenth rib backfat and average backfat in the ISU Berkshire x Yorkshire pig resource family. In the Mangalitsa x Piétrain cross we only detected an association of c.594G>C polymorphism with backfat thickness at shoulders and no association with growth traits, such as feed intake or daily gain (data not shown).

Conclusion

Several associations between investigated candidate genes and lipid deposition and energy homeostasis-related traits were detected in the Mangalitsa x Piétrain cross. However, none of these associations were replicated in commercial populations. Replication studies provide insurance against errors and biases that can afflict any individual study and strengthen confidence that any associations are biologically interesting rather than methodologically inadequate (Hattersley & McCarthy 2005). The inconsistent effects observed across populations indicated that the investigated polymorphisms could be in LD with the putative casual mutations. Especially in the F2 population of the Mangalitsa x Piétrain cross, LD could be a plausible explanation for the observed associations. In F2 extensive LD blocks are found, while in purebreds short LD blocks are expected. These long LD blocks are desirable for a linkage study, but can cause difficulties in an association study. The investigation of a candidate gene in an intercross is reasonable when the candidate gene is selected from genes suited within the CI of the QTL founded in the studied reference family and has physiological relevance to the trait (Ron & Weller 2007; Park *et al.* 2002; Bruun *et al.* 2006). However, it does not mean directly that the results reported here represent false positive results. QTLs with mild effects could be missed because of the small sample size in the represented genome-wide linkage mapping as well as the marker density map shows for some chromosomes small resolution.

It should be noted that some of the associations could not be investigated in commercial populations because of a lack of phenotypic data. In addition, allele frequency of the investigated sequence variants was relatively different in the experimental cross and the commercial population, which could also affect the observed results. Nevertheless, the main conclusions that can be drawn after the analysis of the obtained results are as

follows:

- Significant evidences for QTLs were obtained for carcass composition, growth, meat quality and fatty acid composition traits in the Mangalitsa x Piétrain cross.
- Many polymorphisms were found in the non-coding region of candidate genes without obvious functional implications. However, in theory any of the non-coding polymorphisms could be located in an unknown regulatory region and have functional consequences.
- The insertion-deletion polymorphism of SINE and the non-synonymous substitution in the *PDK4* gene are significantly associated with SFA content. The SFA data were not available in the Piétrain population, hence the confirmation studies were not applied. The insertion-deletion polymorphism of SINE shows an association with average feed intake, which association could not be replicated in the German Landrace population.
- The microsatellite polymorphism in *INSIG2* shows an association with fatness and fatty acid composition of intramuscular fat content in the Mangalitsa x Piétrain cross. However, whether or not the observed association in swine is caused by the *INSIG2* gene remains unclear.
- The synonymous substitution c.594G>C of *FTO* shows an association with backfat thickness at shoulders in the F2 animals of the Mangalitsa x Piétrain cross and with loin muscle area in the Piétrain population.

Acknowledgments

Firstly I would like to express my gratitude to my supervisor Prof. Dr. Ruedi Fries for giving me the opportunity to work in his group and for his encouragement, patience and support during this work.

I would like to thank the Deutsche Forschungsgemeinschaft for supporting this project financially (FR 1284/2-1 and FR 1284/2-2) and Landesanstalt für Landwirtschaft (LfL, Grub), Suisag and Besamungsstation Bergheim that provided samples.

A special thank you to Hermann Schwarzenbacher and Mahdi Osman for their help with the statistical analyses, as well as Kristof Flisikowski for the fruitful discussions on laboratory techniques. I want to extend my thanks to Bettina Hayn, Rudi Antes and Anne Keller for their support and guidance regarding the laboratory work, Theresia Böhm for DNA isolation and Hermine Kienberger for analyses of meat samples. I would like to thank Birgit Hoffman for her friendly help in administrative matters.

I warmly thank Karina, Li and Daniela for their help and friendship. Thanks to all colleagues not mentioned by name at the Chair of Animal Breeding.

Finally, I would like to thank my parents. You always support me wherever I go and whatever I do. Thank you.

Anna

October 2010

Summary

Aims of this thesis were to analyze quantitative trait loci (QTLs) and characterize candidate genes for lipid metabolism in a Mangalitsa x Piétrain resource population. The studies of traits related to lipid metabolism in livestock are important for many aspects of agricultural and biomedical sciences.

For the QTL study offspring of the Mangalitsa x Piétrain cross were extensively phenotyped for carcass composition, growth, meat quality and fatty acid composition traits. A total of 167 F2 animals were genotyped using 109 microsatellite markers covering 18 autosomes. Using least squares regression interval mapping, significant evidences for QTLs were obtained for all analysed group of traits.

The candidate genes include pyruvate dehydrogenase kinase, isozyme 4 (*PDK4)* involved in three important energy pathways: glycolysis, fatty acid oxidation and synthesis; and two genes previously indicated as human obesity susceptibility genes: insulin induced gene 2 (*INSIG2*) and fat mass and obesity associated gene (*FTO*). Additionally, all candidate genes were assigned to regions containing QTLs for lipid-related traits in pig. The genomic sequence of the investigated candidate genes was obtained by sequencing BAC clones. *PDK4, INSIG2* and *FTO* were screened for genetic variants in putative functional regions (5' end, coding regions, intron-exon boundaries and 3' end) by re-sequencing animals of the parental population of the Mangalitsa x Piétrain cross. Potentially functional polymorphisms of each gene were genotyped in the F2 generation of the intercross and investigated for an association with lipid deposition traits such as backfat thickness and other traits.

Several significant associations between investigated polymorphisms of candidate genes and fat related traits were found in the Mangalitsa x Piétrain cross. The main findings of the candidate gene studies are:

- The insertion-deletion polymorphism of short interspersed element (SINE) and the non-synonymous substitution c.1049G>A in the *PDK4* gene are significantly associated with saturated fatty acid content of intramuscular fat. In addition, the

insertion-deletion polymorphism of the SINE shows an association with average feed intake.

- The microsatellite polymorphism *121_122delGT in *INSIG2* shows an association with thinnest backfat thickness as well as saturated- and polyunsaturated- fatty acids content of intramuscular fat content.
- The synonymous substitution c.594G>C in *FTO* is associated with backfat thickness at shoulders in the F2 animals of the Mangalitsa x Piétrain cross and with loin muscle area in the Piétrain population.

This study was conducted as an effort to elucidate the genetic background of lipid-related traits. Further research, e.g. extensive association analyses of sequence variants in each candidate gene in different independent populations and functional studies, are necessary to clarify the importance of the investigated variants as genetic determinants of fatness.

Zusammenfassung

Ziel dieser Dissertation war es QTL (quantitative trait loci) zu kartieren und Kandidatengene des Fettstoffwechsels in einer Mangalitsa x Piétrain Kreuzung zu analysieren. Die Untersuchung von Merkmalen des Fettstoffwechsels in Nutztieren ist für viele Aspekte der landwirtschaftlichen und biomedizinischen Forschung von Bedeutung.

Für die QTL-Studien wurden die Nachkommen aus einer Mangalitsa x Piétrain Kreuzung umfassend für Merkmale der Schlachtkörperbeschaffenheit, des Wachstum, der Fleischqualität und der Fettsäurezusammensetzung phänotypisiert. In der Studie wurden insgesamt 167 F2-Tiere mit Hilfe von 109 Mikrosatelliten-Markern, die 18 Autosomen abdecken, genotypisiert. Intervall-Kartierung nach der Methode der Kleinsten Quadrate wurde angewendet um QTLs aller analysierten Merkmalsgruppen zu detektieren.

Die Kandidatengenanalyse umfasste die Pyruvat Dehydrogenase Kinase, Isozym 4 (*PDK4*), welches in 3 wichtige Energiestoffwechselwege involviert ist: Glycolyse, Oxidation und Synthese von Fettsäuren; und zwei weitere Gene die in Zusammenhang mit Adipositas beim Menschen diskutiert werden*:* "insulin induced gene 2" (*INSIG2*)

und "fat mass and obesity associated gene" (*FTO*). Alle analysierten Gene liegen in chromosomalen Regionen, in denen QTLs für Merkmale der Fetteinlagerung beim Schwein berichtet worden sind. Weiterhin wurden für die Kandidatengene *PDK4*, *INSIG*2 und *FTO* BAC-Klone sequenziert, um die genomische Sequenz zu erhalten. In der Parentalpopulation der Mangalitsa x Piétrain Kreuzung wurde dann durch Resequenzierung nach genetischen Variationen in potentiell funktionellen Regionen (5' Ende kodierende Regionen, Intron - Exon Grenzen und 3' Ende) gesucht. Potenziell funktionelle Polymorphismen jedes Gens wurden in der F2-Generation genotypisiert und auf Assoziation mit Merkmalen der Fetteinlagerung, wie beispielsweise Rückenspeckdicke, untersucht.

Es wurden einige bedeutende Assoziationen zwischen den untersuchten Polymorphismen in den Kandidatengenen und Merkmalen der Fetteinlagerung in der Mangalitsa x Piétrain Kreuzung gefunden. Die Hauptergebnisse der Kandidatengenstudien waren:

- Der SINE Insertion-Deletions Polymorphismus und der nicht-synonyme Aminosäureaustausch c.1049G>A im *PDK4* Gen waren signifikant mit dem Gehalt an gesättigten Fettsäuren im intramuskulären Fett assoziiert. Zusätzlich zeigte der Insertion-Deletions Polymorphismus des SINE eine Assoziation mit der durchschnittlichen Futteraufnahme.
- Der Mikrosatelliten-Polymorphismus *121_122delGT in *INSIG2* zeigte eine Assoziation mit der Fetteinlagerung und der Fettsäurezusammensetzung des intramuskulären Fettes.
- Der synonyme Basenaustausch c.594G> C im FTO Gen war in F2-Tieren der Kreuzung Mangalitsa x Piétrain mit der Rückenspeckdicke an den Schultern und in einer Piétrain Population mit der Rückenspeckdicke im Lendenmuskelbereich assoziiert.

Diese Studie wurde durchgeführt um die genetische Ursache von Merkmalen der Fetteinlagerung zu untersuchen. Weitere Untersuchungen, wie zum Beispiel umfassende Assoziationsstudien der Sequenzvarianten in jedem der Kandidatengene in verschiedenen unabhängigen Populationen oder funktionelle Studien sind notwendig, um die Bedeutung der untersuchten Varianten als genetische Determinanten der Fetteinlagerung zu klären.

Bibliography

Abbot E.L., McCormack J.G., Reynet C., Hassall D.G., Buchan K.W. & Yeaman S.J. (2005) Diverging regulation of pyruvate dehydrogenase kinase isoform gene expression in cultured human muscle cells. *Febs J* **272**, 3004-14.

Andersson L. (2001) Genetic dissection of phenotypic diversity in farm animals. *Nat Rev Genet* **2**, 130-8.

Andersson L. & Georges M. (2004) Domestic-animal genomics: deciphering the genetics of complex traits. *Nat Rev Genet* **5**, 202-12.

Andersson L., Haley C.S., Ellegren H., Knott S.A., Johansson M., Andersson K., Andersson-Eklund L., Edfors-Lilja I., Fredholm M., Hansson I. & et al. (1994) Genetic mapping of quantitative trait loci for growth and fatness in pigs. *Science* **263**, 1771-4.

Andersson-Eklund L., Marklund L., Lundstrom K., Haley C.S., Andersson K., Hansson I., Moller M. & Andersson L. (1998) Mapping quantitative trait loci for carcass and meat quality traits in a wild boar x Large White intercross. *J Anim Sci* **76**, 694-700.

Andreasen C.H., Mogensen M.S., Borch-Johnsen K., Sandbaek A., Lauritzen T., Sorensen T.I., Hansen L., Almind K., Jorgensen T., Pedersen O. & Hansen T. (2008) Non-replication of genome-wide based associations between common variants in INSIG2 and PFKP and obesity in studies of 18,014 Danes. *PLoS One* **3**, e2872.

Andreasen C.H., Stender-Petersen K.L., Mogensen M.S., Torekov S.S., Wegner L., Andersen G., Nielsen A.L., Albrechtsen A., Borch-Johnsen K., Rasmussen S.S., Clausen J.O., Sandbaek A., Lauritzen T., Hansen L., Jorgensen T., Pedersen O. & Hansen T. (2007) Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes*.

Anselme I., Laclef C., Lanaud M., Ruther U. & Schneider-Maunoury S. (2007) Defects in brain patterning and head morphogenesis in the mouse mutant Fused toes. *Dev Biol* **304**, 208-20.

Araki M. & Motojima K. (2006) Identification of ERRalpha as a specific partner of PGC-1alpha for the activation of PDK4 gene expression in muscle. *Febs J* **273**, 1669-80.

Arens M., Schulte E. & Weber K. (1994) Fettsaeuremethylester, Umesterung mit Trimethylsulfoniumhydroxid (Schnellverfahren)- Gemeinschaftsarbeitn Der DGF, 138.Mitteilung: Deutsche Einheitsmethoden Zur Untersuchung Von Fetten, Fettproducten, Tensiden Und Verwandetenstoffen, 105.Mitt.: Analyse Von fetten Xxxi. *Fat Science Technology* **96**, 67-8.

Bidanel J.P., Milan D., Iannuccelli N., Amigues Y., Boscher M.Y., Bourgeois F., Caritez J.C., Gruand J., Le Roy P., Lagant H., Quintanilla R., Renard C., Gellin J., Ollivier L. & Chevalet C. (2001) Detection of quantitative trait loci for growth and fatness in pigs. *Genet Sel Evol* **33**, 289-309.

Bligh E.G. & Dyer W.J. (1959) A rapid method of total lipid extraction and purification. *Can J Biochem Physiol* **37**, 911-7.

Boes E., Kollerits B., Heid I.M., Hunt S.C., Pichler M., Paulweber B., Coassin S.,

Adams T.D., Hopkins P.N., Lingenhel A., Wagner S.A. & Kronenberg F. (2008) INSIG2 polymorphism is neither associated with BMI nor with phenotypes of lipoprotein metabolism. *Obesity (Silver Spring)* **16**, 827-33.

Bowker-Kinley M.M., Davis W.I., Wu P., Harris R.A. & Popov K.M. (1998) Evidence for existence of tissue-specific regulation of the mammalian pyruvate dehydrogenase complex. *Biochem J* **329 (Pt 1)**, 191-6.

Breathnach R. & Chambon P. (1981) Organization and expression of eucaryotic split genes coding for proteins. *Ann. Rev. Biochem* **50**, 349-83.

Bruun C.S., Jorgensen C.B., Nielsen V.H., Andersson L. & Fredholm M. (2006) Evaluation of the porcine melanocortin 4 receptor (MC4R) gene as a positional candidate for a fatness QTL in a cross between Landrace and Hampshire. *Anim Genet* **37**, 359-62.

Cepica S., Schröffel Jr J., Stratil A., Hojny J., Pierzchala M., Kuryl J., Brunsch C., Sternstein I., Davoli R., Fontanesi L., Reiner G., Bartenschlager H., Moser G. & Geldermann H. (2003a) Linkage and QTL mapping for Sus scrofa chromosome 9. *J Anim Breed Genet.* **120**, 74-81.

Cepica S., Stratil A., Kopecny M., Blazkova P., Schröffel Jr J., Davoli R., Fontanesi L., Reiner G., Bartenschlager H., Moser G. & Geldermann H. (2003b) Linkage and QTL mapping for Sus scrofa chromosome 4. *J. Anim. Breed. Genet.* **120**, 28-37.

Chen K., Baxter T., Muir W.M., Groenen M.A. & Schook L.B. (2007a) Genetic resources, genome mapping and evolutionary genomics of the pig (Sus scrofa). *Int J Biol Sci* **3**, 153-65.

Chen T.M., Kuo P.L., Hsu C.H., Tsai S.J., Chen M.J., Lin C.W. & Sun H.S. (2007b) Microsatellite in the 3' untranslated region of human fibroblast growth factor 9 (FGF9) gene exhibits pleiotropic effect on modulating FGF9 protein expression. *Hum Mutat* **28**, 98.

Cheverud J.M., Ehrich T.H., Hrbek T., Kenney J.P., Pletscher L.S. & Semenkovich C.F. (2004) Quantitative trait loci for obesity- and diabetes-related traits and their dietary responses to high-fat feeding in LGXSM recombinant inbred mouse strains. *Diabetes* **53**, 3328-36.

Churchil G.A. & Doerge R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963-71.

Ciobanu D.C., Bastiaansen J.W., Lonergan S.M., Thomsen H., Dekkers J.C., Plastow G.S. & Rothschild M.F. (2004) New alleles in calpastatin gene are associated with meat quality traits in pigs. *J Anim Sci* **82**, 2829-39.

Citores M.J., Rua-Figueroa I., Rodriguez-Gallego C., Durantez A., Garcia-Laorden M.I., Rodriguez-Lozano C., Rodriguez-Perez J.C., Vargas J.A. & Perez-Aciego P. (2004) The dinucleotide repeat polymorphism in the 3'UTR of the CD154 gene has a functional role on protein expression and is associated with systemic lupus erythematosus. *Ann Rheum Dis* **63**, 310-7.

Davoli R., Bigi D., Fontanesi L., Zambonelli P., Yerle M., Zijlstra C., Bosma A.A., Robic A. & Russo V. (2000) Mapping of 14 expressed sequence tags (ESTs) from porcine skeletal muscle by somatic cell hybrid analysis. *Anim Genet* **31**, 400-3.

de Koning D.J., Harlizius B., Rattink A.P., Groenen M.A., Brascamp E.W. & van Arendonk J.A. (2001) Detection and characterization of quantitative trait loci for

meat quality traits in pigs. *J Anim Sci* **79**, 2812-9.

de Koning D.J., Janss L.L., Rattink A.P., van Oers P.A., de Vries B.J., Groenen M.A., van der Poel J.J., de Groot P.N., Brascamp E.W. & van Arendonk J.A. (1999) Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (Sus scrofa). *Genetics* **152**, 1679-90.

de Koning D.J., Rattink A.P., Harlizius B., van Arendonk J.A., Brascamp E.W. & Groenen M.A. (2000) Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proc Natl Acad Sci U S A* **97**, 7947-50.

Deng H.W., Deng H., Liu Y.J., Liu Y.Z., Xu F.H., Shen H., Conway T., Li J.L., Huang Q.Y., Davies K.M. & Recker R.R. (2002) A genomewide linkage scan for quantitative-trait loci for obesity phenotypes. *Am J Hum Genet* **70**, 1138-51.

Dina C., Meyre D., Gallina S., Durand E., Korner A., Jacobson P., Carlsson L.M., Kiess W., Vatin V., Lecoeur C., Delplanque J., Vaillant E., Pattou F., Ruiz J., Weill J., Levy-Marchal C., Horber F., Potoczna N., Hercberg S., Le Stunff C., Bougneres P., Kovacs P., Marre M., Balkau B., Cauchi S., Chevre J.C. & Froguel P. (2007a) Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet*.

Dina C., Meyre D., Samson C., Tichet J., Marre M., Jouret B., Charles M.A., Balkau B. & Froguel P. (2007b) Comment on "A common genetic variant is associated with adult and childhood obesity". *Science* **315**, 187; author reply

Edwards D.B., Ernst C.W., Raney N.E., Doumit M.E., Hoge M.D. & Bates R.O. (2007) QTL mapping in an F2 Duroc x Piétrain resource population: II. Carcass and meat quality traits. *J Anim Sci*.

Edwards D.B., Ernst C.W., Tempelman R.J., Rosa G.J., Raney N.E., Hoge M.D. & Bates R.O. (2008) Quantitative trait loci mapping in an F2 Duroc x Piétrain resource population: I. Growth traits. *J Anim Sci* **86**, 241-53.

Estelle J., Perez-Enciso M., Mercade A., Varona L., Alves E., Sanchez A. & Folch J.M. (2006) Characterization of the porcine FABP5 gene and its association with the FAT1 QTL in an Iberian by Landrace cross. *Anim Genet* **37**, 589-91.

Evans G.J., Giuffra E., Sanchez A., Kerje S., Davalos G., Vidal O., Illan S., Noguera J.L., Varona L., Velander I., Southwood O.I., de Koning D.J., Haley C.S., Plastow G.S. & Andersson L. (2003) Identification of quantitative trait loci for production traits in commercial pig populations. *Genetics* **164**, 621-7.

Ewing B. & Green P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* **8**, 186-94.

Ewing B., Hillier L., Wendl M.C. & Green P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res* **8**, 175-85.

Fan B., Du Z.Q. & Rothschild M.F. (2009) The fat mass and obesity-associated (FTO) gene is associated with intramuscular fat content and growth rate in the pig. *Anim Biotechnol* **20**, 58-70.

Fischer J., Koch L., Emmerling C., Vierkotten J., Peters T., Bruning J.C. & Ruther U. (2009) Inactivation of the Fto gene protects from obesity. *Nature* **458**, 894-8.

Fontanesi L., Scotti E., Buttazzoni L., Davoli R. & Russo V. (2008) The porcine fat mass and obesity associated (FTO) gene is associated with fat deposition in Italian Duroc pigs. *Anim Genet*.

Frayling T.M., Timpson N.J., Weedon M.N., Zeggini E., Freathy R.M., Lindgren C.M., Perry J.R., Elliott K.S., Lango H., Rayner N.W., Shields B., Harries L.W., Barrett J.C., Ellard S., Groves C.J., Knight B., Patch A.M., Ness A.R., Ebrahim S., Lawlor D.A., Ring S.M., Ben-Shlomo Y., Jarvelin M.R., Sovio U., Bennett A.J., Melzer D., Ferrucci L., Loos R.J., Barroso I., Wareham N.J., Karpe F., Owen K.R., Cardon L.R., Walker M., Hitman G.A., Palmer C.N., Doney A.S., Morris A.D., Davey-Smith G., Hattersley A.T. & McCarthy M.I. (2007) A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science*.

Fredriksson R., Hagglund M., Olszewski P.K., Stephansson O., Jacobsson J.A., Olszewska A.M., Levine A.S., Lindblom J. & Schioth H.B. (2008) The obesity gene, FTO, is of ancient origin, up-regulated during food deprivation and expressed in neurons of feeding-related nuclei of the brain. *Endocrinology* **149**, 2062-71.

Frith M.C., Hansen U. & Weng Z. (2001) Detection of cis-element clusters in higher eukaryotic DNA. *Bioinformatics* **17**, 878-89.

Fujii J., Otsu K., Zorzato F., de Leon S., Khanna V.K., Weiler J.E., O'Brien P.J. & MacLennan D.H. (1991) Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science* **253**, 448-51.

Geldermann H., Muller E., Moser G., Reiner G., Bartenschlager H., Cepica S., Strait J., Kuryl J., Moran C., Davoli R. & Brunsch C. (2003) Genome-wide linkage and QTL mapping in porcine F2 families generated from Piertain, Meishan and Wild Boar cosses. *J Anim Breed Genet.* **120**, 363-93.

Gerbens F., Verburg F.J., Van Moerkerk H.T., Engel B., Buist W., Veerkamp J.H. & te Pas M.F. (2001) Associations of heart and adipocyte fatty acid-binding protein gene expression with intramuscular fat content in pigs. *J Anim Sci* **79**, 347-54.

Gerken T., Girard C.A., Tung Y.C., Webby C.J., Saudek V., Hewitson K.S., Yeo G.S., McDonough M.A., Cunliffe S., McNeill L.A., Galvanovskis J., Rorsman P., Robins P., Prieur X., Coll A.P., Ma M., Jovanovic Z., Farooqi I.S., Sedgwick B., Barroso I., Lindahl T., Ponting C.P., Ashcroft F.M., O'Rahilly S. & Schofield C.J. (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* **318**, 1469-72.

Goldstein J.L., DeBose-Boyd R.A. & Brown M.S. (2006) Protein sensors for membrane sterols. *Cell* **124**, 35-46.

Gong Y., Lee J.N., Brown M.S., Goldstein J.L. & Ye J. (2006) Juxtamembranous aspartic acid in Insig-1 and Insig-2 is required for cholesterol homeostasis. *Proc Natl Acad Sci U S A* **103**, 6154-9.

Gordon D., Abajian C. & Green P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res* **8**, 195-202.

Gremme G., Brendel V., Sparks M.E. & Kurtz S. (2005) Engineering a software tool for gene structure prediction in higher organisms. *Information and Software Technology* **47**, 965-78.

Haley C.S., Knott S.A. & Elsen J.M. (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics* **136**, 1195-207.

Hall D.H., Rahman T., Avery P.J. & Keavney B. (2006) INSIG-2 promoter polymorphism and obesity related phenotypes: association study in 1428

members of 248 families. *BMC Med Genet* **7**, 83.

Hallermayer R. (1976) Eine Schnellmethode zur Bestimmung des Fettgehaltes in lebensmitteln. I. Die Fettbestimmung in Fleischwaren. *Deutsche Lebensmittelrundschau* **10**, 356-9.

Harris R.A., Bowker-Kinley M.M., Huang B. & Wu P. (2002) Regulation of the activity of the pyruvate dehydrogenase complex. *Adv Enzyme Regul* **42**, 249-59.

Hattersley A.T. & McCarthy M.I. (2005) What makes a good genetic association study? *Lancet* **366**, 1315-23.

Herbert A., Gerry N.P., McQueen M.B., Heid I.M., Pfeufer A., Illig T., Wichmann H.E., Meitinger T., Hunter D., Hu F.B., Colditz G., Hinney A., Hebebrand J., Koberwitz K., Zhu X., Cooper R., Ardlie K., Lyon H., Hirschhorn J.N., Laird N.M., Lenburg M.E., Lange C. & Christman M.F. (2006) A common genetic variant is associated with adult and childhood obesity. *Science* **312**, 279-83.

Hinney A., Nguyen T.T., Scherag A., Friedel S., Bronner G., Muller T.D., Grallert H., Illig T., Wichmann H.E., Rief W., Schafer H. & Hebebrand J. (2007) Genome Wide Association (GWA) Study for Early Onset Extreme Obesity Supports the Role of Fat Mass and Obesity Associated Gene (FTO) Variants. *PLoS One* **2**, e1361.

Hirschhorn J.N. & Daly M.J. (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* **6**, 95-108.

Holness M.J., Bulmer K., Smith N.D. & Sugden M.C. (2003) Investigation of potential mechanisms regulating protein expression of hepatic pyruvate dehydrogenase kinase isoforms 2 and 4 by fatty acids and thyroid hormone. *Biochem J* **369**, 687-95.

Holness M.J., Kraus A., Harris R.A. & Sugden M.C. (2000) Targeted upregulation of pyruvate dehydrogenase kinase (PDK)-4 in slow-twitch skeletal muscle underlies the stable modification of the regulatory characteristics of PDK induced by high-fat feeding. *Diabetes* **49**, 775-81.

Holness M.J., Smith N.D., Bulmer K., Hopkins T., Gibbons G.F. & Sugden M.C. (2002) Evaluation of the role of peroxisome-proliferator-activated receptor alpha in the regulation of cardiac pyruvate dehydrogenase kinase 4 protein expression in response to starvation, high-fat feeding and hyperthyroidism. *Biochem J* **364**, 687-94.

Holness M.J. & Sugden M.C. (2003) Regulation of pyruvate dehydrogenase complex activity by reversible phosphorylation. *Biochem Soc Trans* **31**, 1143-51.

Horikoshi M., Hara K., Ito C., Shojima N., Nagai R., Ueki K., Froguel P. & Kadowaki T. (2007) Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia* **50**, 2461-6.

Hotta K., Nakamura M., Nakata Y., Matsuo T., Kamohara S., Kotani K., Komatsu R., Itoh N., Mineo I., Wada J., Masuzaki H., Yoneda M., Nakajima A., Miyazaki S., Tokunaga K., Kawamoto M., Funahashi T., Hamaguchi K., Yamada K., Hanafusa T., Oikawa S., Yoshimatsu H., Nakao K., Sakata T., Matsuzawa Y., Tanaka K., Kamatani N. & Nakamura Y. (2008) INSIG2 gene rs7566605 polymorphism is associated with severe obesity in Japanese. *J Hum Genet* **53**, 857-62.

Hu Z.L., Fritz E.R. & Reecy J.M. (2007) AnimalQTLdb: a livestock QTL database tool

set for positional QTL information mining and beyond. *Nucleic Acids Res* **35**, D604-9.

Hughes A.E., Orr N., Esfandiary H., Diaz-Torres M., Goodship T. & Chakravarthy U. (2006) A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. *Nat Genet* **38**, 1173-7.

Karlskov-Mortensen P., Bruun C.S., Braunschweig M.H., Sawera M., Markljung E., Enfalt A.C., Hedebro-Velander I., Josell A., Lindahl G., Lundstrom K., von Seth G., Jorgensen C.B., Andersson L. & Fredholm M. (2006) Genome-wide identification of quantitative trait loci in a cross between Hampshire and Landrace I: carcass traits. *Anim Genet* **37**, 156-62.

Kim C.W., Hong Y.H., Yun S.I., Lee S.R., Kim Y.H., Kim M.S., Chung K.H., Jung W.Y., Kwon E.J., Hwang S.S., Park D.H., Cho K.K., Lee J.G., Kim B.W., Kim J.W., Kang Y.S., Yeo J.S. & Chang K.T. (2006) Use of microsatellite markers to detect quantitative trait loci in Yorkshire pigs. *J Reprod Dev* **52**, 229-37.

Kim J.J., Zhao H., Thomsen H., Rothschild M.F. & Dekkers J.C. (2005) Combined line-cross and half-sib QTL analysis of crosses between outbred lines. *Genet Res* **85**, 235-48.

Knoll A., Stratil A., Vykoukalova Z., Van Poucke M., Bartenschlager H., Peelman L.J. & Geldermann H. (2003) Polymorphism analysis and mapping to SSC4 of the porcine apolipoprotein A2 (APOA2) gene. *Anim Genet* **34**, 384-6.

Knott S.A., Marklund L., Haley C.S., Andersson K., Davies W., Ellegren H., Fredholm M., Hansson I., Hoyheim B., Lundstrom K., Moller M. & Andersson L. (1998) Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. *Genetics* **149**, 1069-80.

Korotchkina L.G. & Patel M.S. (2001) Site specificity of four pyruvate dehydrogenase kinase isoenzymes toward the three phosphorylation sites of human pyruvate dehydrogenase. *J Biol Chem* **276**, 37223-9.

Krapivner S., Popov S., Chernogubova E., Hellenius M.L., Fisher R.M., Hamsten A. & van't Hooft F.M. (2008) Insulin-induced gene 2 involvement in human adipocyte metabolism and body weight regulation. *J Clin Endocrinol Metab* **93**, 1995-2001.

Kumar J., Sunkishala R.R., Karthikeyan G. & Sengupta S. (2007) The common genetic variant upstream of INSIG2 gene is not associated with obesity in Indian population. *Clin Genet* **71**, 415-8.

Kuryl J., Pierzchala M., Hojny J., Reiner G., Bartenschlager H., Moser G. & Geldermann H. (2003) Linkage and QTL mapping for Sus scrofa chromosome 15. *J Anim Breed Genet.* **120**, 119-25.

Kuzuya M., Ando F., Iguchi A. & Shimokata H. (2007) No association between rs7566605 variant and being overweight in Japanese. *Obesity (Silver Spring)* **15**, 2531-4.

Lan J., Lei M.G., Zhang Y.B., Wang J.H., Feng X.T., Xu D.Q., Gui J.F. & Xiong Y.Z. (2008) Characterization of the porcine differentially expressed PDK4 gene and association with meat quality. *Mol Biol Rep*.

Larsen F., Gundersen G., Lopez R. & Prydz H. (1992) CpG islands as gene markers in the human genome. *Genomics* **13**, 1095-107.

Le Hellard S., Theisen F.M., Haberhausen M., Raeder M.B., Ferno J., Gebhardt S., Hinney A., Remschmidt H., Krieg J.C., Mehler-Wex C., Nothen M.M., Hebebrand J. & Steen V.M. (2009) Association between the insulin-induced gene 2 (INSIG2) and weight gain in a German sample of antipsychotic-treated schizophrenic patients: perturbation of SREBP-controlled lipogenesis in drug-related metabolic adverse effects? *Mol Psychiatry* **14**, 308-17.

Lee J.H., Jeon M.H., Seo Y.J., Lee Y.J., Ko J.H., Tsujimoto Y. & Lee J.H. (2004) CA repeats in the 3'-untranslated region of bcl-2 mRNA mediate constitutive decay of bcl-2 mRNA. *J Biol Chem* **279**, 42758-64.

Lewis S.E., Searle S.M., Harris N., Gibson M., Lyer V., Richter J., Wiel C., Bayraktaroglir L., Birney E., Crosby M.A., Kaminker J.S., Matthews B.B., Prochnik S.E., Smithy C.D., Tupy J.L., Rubin G.M., Misra S., Mungall C.J. & Clamp M.E. (2002) Apollo: a sequence annotation editor. *Genome Biol* **3**, RESEARCH0082.

Li H., Wu Y., Loos R.J., Hu F.B., Liu Y., Wang J., Yu Z. & Lin X. (2007) Variants in FTO gene are not associated with obesity in a Chinese Han population. *Diabetes*.

Littmann E., Götz K.U. & Dodenhoff J. (2006) *Schweinezucht und Schweineproduktion*. Bayrische Landesanstalt für Lanwirtschaft, Freising.

Loos R.J., Barroso I., O'Rahilly S. & Wareham N.J. (2007) Comment on "A common genetic variant is associated with adult and childhood obesity". *Science* **315**, 187; author reply

Lyon H.N., Emilsson V., Hinney A., Heid I.M., Lasky-Su J., Zhu X., Thorleifsson G., Gunnarsdottir S., Walters G.B., Thorsteinsdottir U., Kong A., Gulcher J., Nguyen T.T., Scherag A., Pfeufer A., Meitinger T., Bronner G., Rief W., Soto-Quiros M.E., Avila L., Klanderman B., Raby B.A., Silverman E.K., Weiss S.T., Laird N., Ding X., Groop L., Tuomi T., Isomaa B., Bengtsson K., Butler J.L., Cooper R.S., Fox C.S., O'Donnell C.J., Vollmert C., Celedon J.C., Wichmann H.E., Hebebrand J., Stefansson K., Lange C. & Hirschhorn J.N. (2007) The association of a SNP upstream of INSIG2 with body mass index is reproduced in several but not all cohorts. *PLoS Genet* **3**, e61.

Ma K., Zhang Y., Elam M.B., Cook G.A. & Park E.A. (2005) Cloning of the rat pyruvate dehydrogenase kinase 4 gene promoter: activation of pyruvate dehydrogenase kinase 4 by the peroxisome proliferator-activated receptor gamma coactivator. *J Biol Chem* **280**, 29525-32.

Maak S., Jaesert S., Neumann K., Yerle M. & von Lengerken G. (2001) Isolation of expressed sequence tags of skeletal muscle of neonatal healthy and splay leg piglets and mapping by somatic cell hybrid analysis. *Anim Genet* **32**, 303-7.

Malek M., Dekkers J.C., Lee H.K., Baas T.J. & Rothschild M.F. (2001) A molecular genome scan analysis to identify chromosomal regions influencing economic traits in the pig. I. Growth and body composition. *Mamm Genome* **12**, 630-6.

Marklund L., Nystrom P.E., Stern S., Andersson-Eklund L. & Andersson L. (1999) Confirmed quantitative trait loci for fatness and growth on pig chromosome 4. *Heredity* **82 (Pt 2)**, 134-41.

McPherson R. & Gauthier A. (2004) Molecular regulation of SREBP function: the Insig-SCAP connection and isoform-specific modulation of lipid synthesis.

Biochem Cell Biol **82**, 201-11.

Meidtner K., Wermter A.K., Hinney A., Remschmidt H., Hebebrand J. & Fries R. (2006) Association of the melanocortin 4 receptor with feed intake and daily gain in F2 Mangalitsa x Piétrain pigs. *Anim Genet* **37**, 245-7.

Meijerink E., R. Fries R., Vögeli P., Masabanda J., Wigger G., Stricker C., Neuenschwander S., Bertschinger H. U & Stranzinger G. (1997) Two α(1,2) fucosyltransferase genes on porcine Chromosome 6q11 are closely linked to the blood group inhibitor (S) and Escherichia coli F18 receptor (ECF18R) loci. Mamm Genome **8**, 736-741.

Milan D., Bidanel J.P., Iannuccelli N., Riquet J., Amigues Y., Gruand J., Le Roy P., Renard C. & Chevalet C. (2002) Detection of quantitative trait loci for carcass composition traits in pigs. *Genet Sel Evol* **34**, 705-28.

Milan D., Jeon J.T., Looft C., Amarger V., Robic A., Thelander M., Rogel-Gaillard C., Paul S., Iannuccelli N., Rask L., Ronne H., Lundstrom K., Reinsch N., Gellin J., Kalm E., Roy P.L., Chardon P. & Andersson L. (2000) A mutation in PRKAG3 associated with excess glycogen content in pig skeletal muscle. *Science* **288**, 1248-51.

Mohrmann M., Roehe R., Knap P.W., Looft H., Plastow G.S. & Kalm E. (2006) Quantitative trait loci associated with AutoFOM grading characteristics, carcass cuts and chemical body composition during growth of Sus scrofa. *Anim Genet* **37**, 435-43.

Moller M., Berg F., Riquet J., Pomp D., Archibald A., Anderson S., Feve K., Zhang Y., Rothschild M., Milan D., Andersson L. & Tuggle C.K. (2004) High-resolution comparative mapping of pig Chromosome 4, emphasizing the FAT1 region. *Mamm Genome* **15**, 717-31.

Murani E., Muraniova M., Ponsuksili S., Schellander K. & Wimmers K. (2006a) Molecular characterization and evidencing of the porcine CRH gene as a functional-positional candidate for growth and body composition. *Biochem Biophys Res Commun* **342**, 394-405.

Murani E., Ponsuksili S., Schellander K. & Wimmers K. (2006b) Association of corticotropin-releasing hormone gene variation with performance and meat quality traits in commercial pig lines. *Anim Genet* **37**, 509-12.

Nackley A.G., Shabalina S.A., Tchivileva I.E., Satterfield K., Korchynskyi O., Makarov S.S., Maixner W. & Diatchenko L. (2006) Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* **314**, 1930-3.

Nezer C., Moreau L., Brouwers B., Coppieters W., Detilleux J., Hanset R., Karim L., Kvasz A., Leroy P. & Georges M. (1999) An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. *Nat Genet* **21**, 155-6.

Nickerson D.A., Tobe V.O. & Taylor S.L. (1997) PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res* **25**, 2745-51.

Nii M., Hayashi T., Tani F., Niki A., Mori N., Fujishima-Kanaya N., Komatsu M., Aikawa K., Awata T. & Mikawa S. (2006) Quantitative trait loci mapping for fatty acid composition traits in perirenal and back fat using a Japanese wild boar

x Large White intercross. *Anim Genet* **37**, 342-7.

Ohashi J., Naka I., Kimura R., Natsuhara K., Yamauchi T., Furusawa T., Nakazawa M., Ataka Y., Patarapotikul J., Nuchnoi P., Tokunaga K., Ishida T., Inaoka T., Matsumura Y. & Ohtsuka R. (2007) FTO polymorphisms in oceanic populations. *J Hum Genet* **52**, 1031-5.

Park H.B., Carlborg O., Marklund S. & Andersson L. (2002) Melanocortin-4 receptor (MC4R) genotypes have no major effect on fatness in a Large White x Wild Boar intercross. *Anim Genet* **33**, 155-7.

Patel M.S. & Korotchkina L.G. (2001) Regulation of mammalian pyruvate dehydrogenase complex by phosphorylation: complexity of multiple phosphorylation sites and kinases. *Exp Mol Med* **33**, 191-7.

Patel M.S. & Roche T.E. (1990) Molecular biology and biochemistry of pyruvate dehydrogenase complexes. *FASEB J* **4**, 3224-33.

Peeters A., Beckers S., Verrijken A., Roevens P., Peeters P., Van Gaal L. & Van Hul W. (2007) Variants in the FTO gene are associated with common obesity in the Belgian population. *Mol Genet Metab*.

Perez-Enciso M. & Misztal I. (2004) Qxpak: a versatile mixed model application for genetical genomics and QTL analyses. *Bioinformatics* **20**, 2792-8.

Peters S.J., Harris R.A., Heigenhauser G.J. & Spriet L.L. (2001) Muscle fiber type comparison of PDH kinase activity and isoform expression in fed and fasted rats. *Am J Physiol Regul Integr Comp Physiol* **280**, R661-8.

Peters T., Ausmeier K. & Ruther U. (1999) Cloning of Fatso (Fto), a novel gene deleted by the Fused toes (Ft) mouse mutation. *Mamm Genome* **10**, 983-6.

Pierzchala M., Cieslak D., Reiner G., Bartenschlager H., Moser G. & Geldermann H. (2003a) Linkage and QTL mapping for Sus scrofa chromosome 17. *J Anim Breed Genet.* **120**, 132-7.

Pierzchala M., Kuryl J., Reiner G., Bartenschlager H., Moser G. & Geldermann H. (2003b) Linkage and QTL mapping for Sus scrofa chromosome 16. *J Anim Breed Genet.* **120**, 126-31.

Ponsuksili S., Chomdej S., Murani E., Blaser U., Schreinemachers H.J., Schellander K. & Wimmers K. (2005) SNP detection and genetic mapping of porcine genes encoding enzymes in hepatic metabolic pathways and evaluation of linkage with carcass traits. *Anim Genet* **36**, 477-83.

Price R.A., Li W.D. & Zhao H. (2008) FTO gene SNPs associated with extreme obesity in cases, controls and extremely discordant sister pairs. *BMC Med Genet* **9**, 4.

Qiu H., Xia T., Chen X.D., Feng S.Q., Gan L., Lei T., Peng Y., Zhang G.D., Nie T., Yue G.P., Zhao X.L. & Yang Z.Q. (2005) Sequencing and chromosome mapping of pig INSIG 2 and a related pseudogene. *Anim Genet* **36**, 188-9.

Quintanilla R., Milan D. & Bidanel J.P. (2002) A further look at quantitative trait loci affecting growth and fatness in a cross between Meishan and Large White pig populations. *Genet Sel Evol* **34**, 193-210.

Rattink A.P., De Koning D.J., Faivre M., Harlizius B., van Arendonk J.A. & Groenen M.A. (2000) Fine mapping and imprinting analysis for fatness trait QTLs in pigs. *Mamm Genome* **11**, 656-61.

Reese M.G. (2001) Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. *Comput Chem* **26**, 51-6.

Reinehr T., Hinney A., Nguyen T.T. & Hebebrand J. (2008) Evidence of an influence of a polymorphism near the INSIG2 on weight loss during a lifestyle intervention in obese children and adolescents. *Diabetes* **57**, 623-6.

Richtlinie für die Stationsprüfung auf Mastleistung, Schlachtkörperwert und Fleischbeschaffenheit beim Schwein. URL http://www.zds-bonn.de/download.php/790/rlnkp04.pdf.

Robbens S., Rouze P., Cock J.M., Spring J., Worden A.Z. & Van de Peer Y. (2007) The FTO Gene, Implicated in Human Obesity, Is Found Only in Vertebrates and Marine Algae. *J Mol Evol*.

Rohrer G.A., Alexander L.J., Hu Z., Smith T.P., Keele J.W. & Beattie C.W. (1996) A comprehensive map of the porcine genome. *Genome Res* **6**, 371-91.

Rohrer G.A. & Keele J.W. (1998a) Identification of quantitative trait loci affecting carcass composition in swine: I. Fat deposition traits. *J Anim Sci* **76**, 2247-54.

Rohrer G.A. & Keele J.W. (1998b) Identification of quantitative trait loci affecting carcass composition in swine: II. Muscling and wholesale product yield traits. *J Anim Sci* **76**, 2255-62.

Rohrer G.A., Thallman R.M., Shackelford S., Wheeler T. & Koohmaraie M. (2005) A genome scan for loci affecting pork quality in a Duroc-Landrace F2 population. *Anim Genet* **37**, 17-27.

Rohrer G.A., Thallman R.M., Shackelford S., Wheeler T. & Koohmaraie M. (2006) A genome scan for loci affecting pork quality in a Duroc-Landrace F population. *Anim Genet* **37**, 17-27.

Ron M. & Weller J.I. (2007) From QTL to QTN identification in livestock--winning by points rather than knock-out: a review. *Anim Genet* **38**, 429-39.

Rosskopf D., Bornhorst A., Rimmbach C., Schwahn C., Kayser A., Kruger A., Tessmann G., Geissler I., Kroemer H.K. & Volzke H. (2007) Comment on "A common genetic variant is associated with adult and childhood obesity". *Science* **315**, 187; author reply

Rothschild M.F., Hu Z.L. & Jiang Z. (2007) Advances in QTL mapping in pigs. *Int J Biol Sci* **3**, 192-7.

Rowles J., Scherer S.W., Xi T., Majer M., Nickle D.C., Rommens J.M., Popov K.M., Harris R.A., Riebow N.L., Xia J., Tsui L.C., Bogardus C. & Prochazka M. (1996) Cloning and characterization of PDK4 on 7q21.3 encoding a fourth pyruvate dehydrogenase kinase isoenzyme in human. *J Biol Chem* **271**, 22376-82.

Rozen S. & Skaletsky H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**, 365-86.

Sanchez-Pulido L. & Andrade-Navarro M.A. (2007) The FTO (fat mass and obesity associated) gene codes for a novel member of the non-heme dioxygenase superfamily. *BMC Biochem* **8**, 23.

Sandelin A., Wasserman W.W. & Lenhard B. (2004) ConSite: web-based prediction of regulatory elements using cross-species comparison. *Nucleic Acids Res* **32**, W249-52.

Scuteri A., Sanna S., Chen W.M., Uda M., Albai G., Strait J., Najjar S., Nagaraja R., Orru M., Usala G., Dei M., Lai S., Maschio A., Busonero F., Mulas A., Ehret G.B., Fink A.A., Weder A.B., Cooper R.S., Galan P., Chakravarti A., Schlessinger D., Cao A., Lakatta E. & Abecasis G.R. (2007) Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. *PLoS Genet* **3**, e115.

Seaton G., Haley C.S., Knott S.A., Kearsey M. & Visscher P.M. (2002) QTL Express: mapping quantitative trait loci in simple and complex pedigrees. *Bioinformatics* **18**, 339-40.

Sever N., Song B.L., Yabe D., Goldstein J.L., Brown M.S. & DeBose-Boyd R.A. (2003) Insig-dependent ubiquitination and degradation of mammalian 3-hydroxy-3-methylglutaryl-CoA reductase stimulated by sterols and geranylgeraniol. *J Biol Chem* **278**, 52479-90.

Sjogren P., Sierra-Johnson J., Gertow K., Rosell M., Vessby B., de Faire U., Hamsten A., Hellenius M.L. & Fisher R.M. (2008) Fatty acid desaturases in human adipose tissue: relationships between gene expression, desaturation indexes and insulin resistance. *Diabetologia* **51**, 328-35.

Stratigopoulos G., Padilla S.L., LeDuc C.A., Watson E., Hattersley A.T., McCarthy M.I., Zeltser L.M., Chung W.K. & Leibel R.L. (2008) Regulation of Fto/Ftm gene expression in mice and humans. *Am J Physiol Regul Integr Comp Physiol* **294**, R1185-96.

Sugden M.C. (2003) PDK4: A factor in fatness? *Obes Res* **11**, 167-9.

Sugden M.C., Bulmer K., Augustine D. & Holness M.J. (2001a) Selective modification of pyruvate dehydrogenase kinase isoform expression in rat pancreatic islets elicited by starvation and activation of peroxisome proliferator-activated receptor-alpha: implications for glucose-stimulated insulin secretion. *Diabetes* **50**, 2729-36.

Sugden M.C., Bulmer K., Gibbons G.F. & Holness M.J. (2001b) Role of peroxisome proliferator-activated receptor-alpha in the mechanism underlying changes in renal pyruvate dehydrogenase kinase isoform 4 protein expression in starvation and after refeeding. *Arch Biochem Biophys* **395**, 246-52.

Sugden M.C. & Holness M.J. (1994) Interactive regulation of the pyruvate dehydrogenase complex and the carnitine palmitoyltransferase system. *FASEB J* **8**, 54-61.

Sugden M.C. & Holness M.J. (2003) Recent advances in mechanisms regulating glucose oxidation at the level of the pyruvate dehydrogenase complex by PDKs. *Am J Physiol Endocrinol Metab* **284**, E855-62.

Sugden M.C., Kraus A., Harris R.A. & Holness M.J. (2000a) Fibre-type specific modification of the activity and regulation of skeletal muscle pyruvate dehydrogenase kinase (PDK) by prolonged starvation and refeeding is associated with targeted regulation of PDK isoenzyme 4 expression. *Biochem J* **346 Pt 3**, 651-7.

Sugden M.C., Lall H.S., Harris R.A. & Holness M.J. (2000b) Selective modification of the pyruvate dehydrogenase kinase isoform profile in skeletal muscle in hyperthyroidism: implications for the regulatory impact of glucose on fatty acid oxidation. *J Endocrinol* **167**, 339-45.

Swinburne I.A., Miguez D.G., Landgraf D. & Silver P.A. (2008) Intron length increases oscillatory periods of gene expression in animal cells. *Genes Dev* **22**, 2342-6.

Szczerbal I., Chmurzynska A. & Switonski M. (2007) Cytogenetic mapping of eight genes encoding fatty acid binding proteins (FABPs) in the pig genome. *Cytogenet Genome Res* **118**, 63-6.

Tabara Y., Kawamoto R., Osawa H., Nakura J., Makino H., Miki T. & Kohara K. (2008) No association between INSIG2 Gene rs7566605 polymorphism and being overweight in Japanese population. *Obesity (Silver Spring)* **16**, 211-5.

Tabaska J.E. & Zhang M.Q. (1999) Detection of polyadenylation signals in human DNA sequences. *Gene* **231**, 77-86.

Tabor H.K., Risch N.J. & Myers R.M. (2002) Candidate-gene approaches for studying complex genetic traits: practical considerations. *Nat Rev Genet* **3**, 391-7.

Tschritter O., Preissl H., Yokoyama Y., Machicao F., Haring H.U. & Fritsche A. (2007) Variation in the FTO gene locus is associated with cerebrocortical insulin resistance in humans. *Diabetologia* **50**, 2602-3.

Van Laere A.S., Nguyen M., Braunschweig M., Nezer C., Collette C., Moreau L., Archibald A.L., Haley C.S., Buys N., Tally M., Andersson G., Georges M. & Andersson L. (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* **425**, 832-6.

van Wijk H.J., Dibbits B., Baron E.E., Brings A.D., Harlizius B., Groenen M.A., Knol E.F. & Bovenhuis H. (2006) Identification of quantitative trait loci for carcass composition and pork quality traits in a commercial finishing cross. *J Anim Sci* **84**, 789-99.

Vinogradov A.E. (2001) Intron length and codon usage. *J Mol Evol* **52**, 2-5.

Visscher P.M., Thompson R. & Haley C.S. (1996) onfidence intervals in QTL mapping by bootstrap. *Genetics* **143**, 1013-20.

Wahlen K., Sjolin E. & Hoffstedt J. (2007) The common rs9939609 gene variant of the fat mass and obesity associated gene (FTO) is related to fat cell lipolysis. *J Lipid Res*.

Walling G.A., Visscher P.M., Andersson L., Rothschild M.F., Wang L., Moser G., Groenen M.A., Bidanel J.P., Cepica S., Archibald A.L., Geldermann H., de Koning D.J., Milan D. & Haley C.S. (2000) Combined analyses of data from quantitative trait loci mapping studies. Chromosome 4 effects on porcine growth and fatness. *Genetics* **155**, 1369-78.

Wang J., Song L., Gonder M.K., Azrak S., Ray D.A., Batzer M.A., Tishkoff S.A. & Liang P. (2006) Whole genome computational comparative genomics: A fruitful approach for ascertaining Alu insertion polymorphisms. *Gene* **365**, 11-20.

Wang L., Yu T.P., Tuggle C.K., Liu H.C. & Rothschild M.F. (1998) A directed search for quantitative trait loci on chromosomes 4 and 7 in pigs. *J Anim Sci* **76**, 2560-7.

Wende A.R., Huss J.M., Schaeffer P.J., Giguere V. & Kelly D.P. (2005) PGC-1alpha coactivates PDK4 gene expression via the orphan nuclear receptor ERRalpha: a mechanism for transcriptional control of muscle glucose metabolism. *Mol Cell Biol* **25**, 10684-94.

Wigginton J.E., Cutler D.J. & Abecasis G.R. (2005) A Note on Exact Tests of Hardy-

Weinberg Equilibrium. *Am J Hum Genet* **76**, 887-93.

Wilcoxon F. (1945) Individual Comparisons by Ranking Methods. *Biometrics Bulletin* **1**, 80-3.

Wimmers K., Murani E., Ponsuksili S., Yerle M. & Schellander K. (2002) Detection of quantitative trait loci for carcass traits in the pig by using AFLP. *Mamm Genome* **13**, 206-10.

Wondji C.S., Morgan J., Coetzee M., Hunt R.H., Steen K., Black W.C.t., Hemingway J. & Ranson H. (2007) Mapping a quantitative trait locus (QTL) conferring pyrethroid resistance in the African malaria vector Anopheles funestus. *BMC Genomics* **8**, 34.

Wu P., Blair P.V., Sato J., Jaskiewicz J., Popov K.M. & Harris R.A. (2000) Starvation increases the amount of pyruvate dehydrogenase kinase in several mammalian tissues. *Arch Biochem Biophys* **381**, 1-7.

Wu P., Inskeep K., Bowker-Kinley M.M., Popov K.M. & Harris R.A. (1999) Mechanism responsible for inactivation of skeletal muscle pyruvate dehydrogenase complex in starvation and diabetes. *Diabetes* **48**, 1593-9.

Wu P., Sato J., Zhao Y., Jaskiewicz J., Popov K.M. & Harris R.A. (1998) Starvation and diabetes increase the amount of pyruvate dehydrogenase kinase isoenzyme 4 in rat heart. *Biochem J* **329 (Pt 1)**, 197-201.

Yabe D., Brown M.S. & Goldstein J.L. (2002) Insig-2, a second endoplasmic reticulum protein that binds SCAP and blocks export of sterol regulatory element-binding proteins. *Proc Natl Acad Sci U S A* **99**, 12753-8.

Yabe D., Komuro R., Liang G., Goldstein J.L. & Brown M.S. (2003) Liver-specific mRNA for Insig-2 down-regulated by insulin: implications for fatty acid synthesis. *Proc Natl Acad Sci U S A* **100**, 3155-60.

Yang L., Wu Y., Li H., Yu Z., Li X., Liu Y., Lin X. & Chen Y. (2008) Potential association of INSIG2 rs7566605 polymorphism with body weight in a Chinese subpopulation. *Eur J Hum Genet* **16**, 759-61.

Yue G., Stratil A., Kopecny M., Schröffelova D., Schröffel Jr J., Hojny J., Cepica S., Davoli R., Zambonelli P., Brunsch C., Sternstein I., Moser G., Bartenschlager H., Reiner G. & Geldermann H. (2003) Linkage and QTL mapping for Sus scrofa chromosome 6. *J Anim Breed Genet.* **120**, 45-55.

Zeng Z.B. (2005) QTL mapping and the genetic basis of adaptation: recent developments. *Genetica* **123**, 25-37.

Zhang J., Lin R., Wang F., Lu M., Lin R.Y., Wang S.Z., Wen H., Jin L. & Wang X.F. (2008) A common polymorphism is associated with body mass index in Uyghur population. *Diabetes Res Clin Pract* **81**, e11-3.

Zhang Y., Ma K., Sadana P., Chowdhury F., Gaillard S., Wang F., McDonnell D.P., Unterman T.G., Elam M.B. & Park E.A. (2006) Estrogen-related receptors stimulate pyruvate dehydrogenase kinase isoform 4 gene expression. *J Biol Chem* **281**, 39897-906.

Zhao H., Rothschild M.F., Fernando R.L. & Dekkers J.C. (2003) Tests of candidate genes in breed cross populations for QTL mapping in livestock. *Mamm Genome* **14**, 472-82.

Zhu M, Zhao S. (2007) Candidate Gene Identification Approach: Progress and

Challenges. *Int J Biol Sci* **3**,420-427.

List of Tables:

List of Figures:

Abbreviations

A	adenine
ABI	Applied Biosystems
AFI	Average feed intake
ATP	Adenosine-5'-triphosphate
BAC	bacterial artificial chromosome
BLAST	basic local alignment search tool
BMI	body mass index
bp	base pair
C	cytosine
cDNA	complementary deoxyribonucleic acid
cM	centi Morgan
DNA	deoxyribonucleic acid
dNTP	nucleotides
EDTA	ethylendiamintetraacetat
ER	Endoplasmic reticulum
FTO	fat mass and obesity associated gene
FUT1	fucosyltransferase 1 (galactoside 2-alpha-L-fucosyltransferase, H blood group)
G	guanine
HMG-CoA	3-hydroxy-3-methylglutaryl-coenzyme A
IMF	intramuscular fat content
INSIG2	insulin induced gene 2
kb	kilo base pairs
LD	Linkage Disequilibrium
MAS	Marker assisted selection
Mb	mega base pairs
mRNA	messenger ribonucleic acid
MUFA	monounsaturated fatty acid
N	A, C, G, T, U
NCBI	National Center for Biotechnology Information
PCR	polymerase chain reaction

PDK	pyruvate dehydrogenase kinase
PDP	pyruvate dehydrogenase phosphatase
PDC	pyruvate dehydrogenase complex
PDK4	pyruvate dehydrogenase kinase, isozyme 4
PUFA	polyunsaturated fatty acid
QTL	quantitative trait locus/loci
RFLP	restriction length polymorphism
RNA	ribonucleic aid
RT-PCR	reverse transcription PCR
RYR1	ryanodine receptor 1 (skeletal) gene, calcium release channel gene
SFA	saturated fatty acid
SINE	short interspersed nuclear element
SNP	single nucleotide polymorphism
SREBP	Sterol Regulatory Element Binding Protein
SSC	porcine chromosome
T	thymine
TSS	transcription start site
UTR	untranslated region