Genetic Analysis of Fat Metabolism in Domestic Pigs and their Wild Ancestor

FRIDA BERG
Dissertation presented at Uppsala University to be publicly examined in B21, Uppsala biomedicinska centrum, Uppsala, Friday, September 29, 2006 at 13:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English.

Abstract

The domestication of the pig began about 9 000 years ago and many of the existing domestic breeds have been selected for phenotypic traits like lean meat and fast growth. Domestic pigs are phenotypically very different from the ancestral wild boar that has adapted to survive in their natural environment. Because of their divergence, crosses between domestic pigs and wild boars are suitable for constructing genetic maps and Quantitative trait locus (QTL) analyses. A cross between the Large White and the European wild boar was thus initiated in the late 1980s. A major QTL for fat deposition and growth, denoted FAT1, was found on chromosome 4. The aim of this thesis was to further characterise the FAT1 locus and to identify the causative gene(s) and mutation(s). We have identified new markers and constructed a high-resolution linkage and RH map of the FAT1 QTL interval. We also performed comparative mapping to the human genome and showed that the pig chromosome 4 is homologous to human chromosomes 1 and 8. The gene order is very well conserved between the two species. In parallel we have narrowed down the FAT1 QTL interval by repeated backcrossing to the domestic Large White breed for six generations. The QTL could be confirmed for fatness but not for growth. Furthermore, the data strongly suggested that there might be more than one gene underlying the FAT1 QTL. Depending on which hypothesis to consider, the one- or two-loci model, the FAT1 interval can be reduced to 3.3 or 20 centiMorgan (cM), respectively, based on the backcross experiments. In the last study we confirm the two-loci model with one locus primarily effecting abdominal fat and another locus primarily effecting subcutaneous fat. We have identified a missense mutation in the RXRG gene which is in strong association with the abdominal fat QTL and the mutation is a potential candidate for that locus.

Brown adipose tissue (BAT) is a specific type of fat essential for non-shivering thermogenesis in mammals. Piglets appear to lack BAT and rely on shivering as the main mechanism for thermoregulation. Uncoupling protein 1 (UCP1) gene is exclusively expressed in BAT and its physiological role is to generate heat by uncoupling oxidative phosphorylation. We show that the UCP1 gene has been disrupted in the pig lineage about 20 years ago. The inactivation of UCP1 provides a genetic explanation for the poor thermoregulation in piglets.

Keywords: Pig, Quantitative Trait Locus, Fat deposition, FAT1, Linkage map, RH map, Comparative map, Backcross, Identical by Descent mapping, Brown adipose tissue, UCP1

Frida Berg, Department of Medical Biochemistry and Microbiology, Box 582, Uppsala University, SE-75123 Uppsala, Sweden

© Frida Berg 2006

ISSN 1651-6206
ISBN 91-554-6623-0
urn:nbn:se:uu:diva-7089 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-7089)
List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


III  Berg F*, Bourneuf E*, Marklund S and Andersson L Further characterization of the \textit{FAT1} locus on pig chromosome 4 – evidence for two quantitative trait loci in the \textit{FAT1} region. \textit{Manuscript}.


* These two authors contributed equally to the work.

§ Reprints were made with kind permission of Springer Science and Business Media.
Introduction

Domestication

Domestication of plants and animals has been an important development in human history. It was essential for the inception of civilisation and it has influenced the size, structure and distribution of human populations (Diamond 2002). Domestication is a phenomenon whereby a wild biological organism is habituated to survive in the company of human beings. Domesticated animals have altered their behaviour, life cycle and physiology as a result of their breeding and living under careful human control for many generations. Animal domestication was an important part of the change in human behaviour from a hunter-gatherer life style to an agricultural life style. The agricultural life style started in the late Pleistocene (14 000-12 000 years ago) and was as a result of an unpredictable climate, a decline in favourable large-size preys and expansion of permanent communities (Diamond 2002). The controlling and management of herds were then followed by controlled breeding and subsequently replaced by artificial selection for desirable characteristics.

Pig domestication

Based on archaeological records the initiation of pig domestication has been estimated to ~9000 years ago in the Near East (Epstein et al. 1984). In a study by Larson et al. (2005), where they analysed the mitochondrial sequence, they provided evidence that pig domestication took place independently in multiple places around Europe and Asia. The Eurasian wild boar (Sus scrofa) is the ancestor of the domestic pig. The body size vary much between subspecies, the largest are the subspecies S.s. ussuricus and S.s. attila (males up to 300 and 275 kg respectively) and the smallest are from Asia (Ruvinsky et al. 1998). The Eurasian wild boars have thick necks, wedge-shaped heads and mobile snouts which enable them to root up the ground when feeding (Fig. 1). Since domestication, many pig breeds have evolved. A breed is defined as a group of animals having common ancestors and certain characteristics distinguishable from other groups of animals within the same species. Breed-specific phenotypes have developed by artificial selection and are maintained by controlled breeding. It is not exactly defined how many pig breeds that exist. The estimation is 200 to
300, though many of the breeds are extinct and have had no or little influence on today’s swine production (Jones 1998).

Figure 1. A European wild boar from Skansen zoo, Stockholm, Sweden. (Photo: Frida Berg)

Quantitative traits

Complex or quantitative traits, like size, weight and body composition, are influenced by many genes (polygenic), each with a small or moderate effect, and environmental factors. This means that there is no simple relationship between the phenotypic value and a particular genotype. The characteristics of a quantitative trait are i) continuous distribution; ii) controlled by many genes; and iii) affected by environmental factors. A Quantitative Trait Locus (QTL) is a chromosomal region harbouring one or several genes affecting a quantitative trait.

Pig as a model organism

Domestic pig breeds have been selected for various traits and have had to acclimatize to various environmental conditions. Many breeds are thus phenotypically very different from each other. Some of these traits are qualitative (simple, monogenic), such as coat colour, but most will be quantitative. Because of their divergence, crosses between different breeds or crosses between domestic pigs and their wild ancestor are suitable for constructing linkage maps and for performing QTL analyses in order to identify genes underlying the phenotypic variations (referred to as a positional cloning/positional candidate cloning, described below). Many widely used commercial breeds, such as the Landrace and Large White, have been intensively selected for lean meat and fast growth which makes them well-suited for studying genes controlling fat metabolism and growth.
Examples of experimental crosses that have been established to study these traits include wild boar vs. Large White (Andersson et al. 1994), Meishan vs Large White (Walling et al. 1998; Bidanel et al. 2001) and Iberian vs. Landrace (Perez-Enciso et al. 2000; Mercade et al. 2005).

Identifying genes or chromosomal regions influencing quantitative traits such as fat deposition and growth are of course important for the pig meat industry. But the pig resembles humans both physiologically and genetically and is therefore also a useful model organism for medical research. The ability to carry out somatic cell nuclear transfer in pig (Onishi et al. 2000) and genetic modifications (Lai et al. 2002a, b, 2006) will in the future be valuable for the pig meat industry, for pig as a model of human diseases and for pig as an organ donor in human medicine (Prather et al. 2003). Two major disadvantages working with the pig are the expensive husbandry costs and their long generation time of about ten months. Compare that to for example the mouse and chicken, which are both cheap to maintain and have generation times of approximately ten and 20-25 weeks, respectively.

The pig genome

The karyotype of the domestic pig is 2n=38; 18 autosomal chromosome pairs and two sex chromosomes. Some wild boars have a karyotype of 2n=36 which is due to a centric fusion of chromosomes 15 and 17 (Western wild boars) or 16 and 17 (Asian wild boars) (Chowdhary 1998).

In the mid 1990s three porcine linkage maps were reported: a map from the PiGMaP consortium (Pig Gene Mapping Project, (Archibald et al. 1995), a Scandinavian map (Ellegren et al. 1994; Marklund et al. 1996) and a map from USDA-MARC (United States Department of Agriculture - Meat Animal Research Centre, (Rohrer et al. 1994, 1996)). The three maps altogether comprised more than 1500 genetic markers (Archibald et al. 1998). The size of the linkage map is estimated to approximately 2500 centiMorgan (cM) (NCBI Map Viewer, http://www.ncbi.nlm.nih.gov/mapview).

The first porcine radiation hybrid (RH) cell map was constructed using the INRA-University of Minnesota RH panel (IMpRH, 7,000-rad of X rays, (Yerle et al. 1998)) and contained 757 markers (Hawken et al. 1999). A 12,000-rad RH panel has also been constructed (Yerle et al. 2002) and both panels are available to the research community through the IMpRH server (Milan et al. 2000). The server can be used for high-resolution mapping (IMpRH mapping tool) and to search for officially submitted results (IMpRH database). The methods for linkage mapping and RH mapping are described in the section “Positional cloning and positional candidate cloning”.

9
The porcine linkage and RH maps together with other information and valuable links can be assessed at the NAGRP (National Animal Genome Research Program) Pig Genome Coordination website (http://www.animalgenome.org/pigs/).

The pig genome is about 3 billion base-pairs (Schmitz et al. 1992; Wernersson et al. 2005). A 0.66X coverage of the pig genome sequence has been released (Wernersson et al. 2005). A comparison of conserved sequences among human, mouse and pig, reveals that the pig sequences are more similar to the human than to the mouse sequences. The mean GC content is similar between the three species but for the distribution of the GC content the pig and human is more alike with a wider distribution than the mouse (Wernersson et al. 2005). The complete genome sequence is underway (Sanger Institute, http://www.sanger.ac.uk/Projects/S_scrofa/) which will further facilitate the use of the pig as a model organism.

Identification of genes controlling traits and disorders
Candidate gene approach
A candidate gene is defined on the basis of its known function i.e. the gene product is known to contribute to the phenotype (trait/disease) of interest. A candidate gene may be identified by searching databases for model organisms where a similar phenotype has been well-studied. The gene will most probably have an ortholog in the species of interest. If a candidate gene is identified it is possible to find information about its function. Such data might include the pattern of expression and the phenotype of mutants. The candidate gene may be investigated by sequencing and if a mutation is found an association test can be performed using pedigree material or using cases and controls. If the phenotype is mapped to a chromosomal region it is crucial to find out if the candidate gene maps to that locus.

An example of a gene studied using a candidate gene approach is the Melanocortin 1 Receptor (MC1R) gene. MC1R mutations are associated with coat colour variations in many species including cattle, horse, pig, sheep, dog and chicken (Klungland et al. 1995; Marklund et al. 1996; Våge et al. 1997; Kijas et al. 1998; Våge et al. 1999; Everts et al. 2000; Newton et al. 2000; Kerje et al. 2003). MC1R was first shown to influence coat colour in mice (Robbins et al. 1993) and the gene was then analysed for sequence variations in different domestic animals with great success. The results reported in paper IV is an example of a study where a candidate gene approach has been used.
Positional cloning and positional candidate cloning

A candidate gene approach may be rather inefficient unless there is an existing animal model for the particular phenotype with a distinct candidate gene(s). Most biochemical pathways are complex with many genes involved and it can be very time-consuming to identify every transcript in a biochemical pathway, and excessively laborious to screen them all for mutations. When using a positional cloning strategy the trait is mapped to a chromosomal region. For monogenic traits the position of the trait can be determined by linkage analysis but for polygenic traits the chromosomal positions are determined by QTL analysis. The size of this region depends on the resolution of the map; a dense map will define the position more exact. The strategy is to clone and sequence the defined region, to identify transcripts and possible candidate genes and finally to search the candidate genes for causative mutation(s). However, if the genome has been sequenced candidate genes can be identified by a combination of their map position and their known properties, this is referred to as positional candidate cloning. Positional cloning and positional candidate cloning approaches are being sped up by new high-throughput genotyping methods (reviewed in Syvänen 2005) and the many sequenced genomes. This will make it possible to construct very dense linkage maps and the access to whole genome sequences of interest will facilitate the search for candidate genes/mutations. The strategy for a positional cloning/positional candidate cloning is illustrated in Figure 2 and the experimental outline is described below.

Mutations associated with several monogenic traits have been identified by positional cloning approach e.g. the mutation in the PRKAG3 gene which explain the excess glycogen content in skeletal muscle of the RN phenotype in pig (Milan et al. 2000) and the CLPG mutation associated with the Callipyge muscle hypertrophy phenotype in sheep (Freking et al. 2002). The first QTL to be identified was the paternally expressed QTL affecting muscle growth, fat deposition and the size of the heart in pig (Jeon et al. 1999; Nezer et al. 1999). This QTL is caused by a regulatory mutation in the insulin-like growth factor 2 (IGF2) gene affecting muscle growth (Van Laere et al. 2003). The IGF2-“story” is an example of a positional candidate cloning approach. Papers I-III use different parts of a positional cloning approach.
Figure 2. A flowchart showing a positional cloning/positional candidate cloning strategy. The experimental outline is explained briefly in the text.
Experimental outline for a positional cloning/positional candidate cloning (Fig. 2):

**Pedigree**
When using family material, the trait or disorder of interest must segregate within the pedigree. In an experimental cross it is therefore important that the two parental lines differ in the phenotype(s) of interest. The first generation (F1) between the two lines is intercrossed to make a segregating F2 generation (Fig. 3A) or it may be backcrossed to one of the parental lines (Fig. 3B). The number of animals in the cross is critical for the resolution of the linkage map and for the power of the QTL analysis (Doerge 2002).

![Figure 3. A schematic illustration of two experimental crosses. A) intercross and B) backcross, BC1 = backcross generation 1](image)

**Markers**
It is important to select markers that are informative in the pedigree, i.e. the two parental lines are fixed or nearly fixed for different alleles at the marker locus. The two most commonly used type of markers are single nucleotide polymorphisms (SNPs) and microsatellites. SNPs are most often biallelic and are the most abundant genetic variation in the genome. A common definition of an SNP is that the less common allele should have a minimum allele frequency of >1% in the population. Microsatellites are sequence length polymorphisms comprising tandem copies of usually dinucleotide repeats. The advantage of microsatellites is that they often are highly polymorphic and thus more informative than biallelic markers.

**Linkage map**
Linkage or genetic maps provide information on the relative position for each marker, but not the exact location on the chromosome. Markers located
on different chromosomes or far apart on the same chromosome segregate independently. In contrast, markers that are in close proximity on the same chromosome show genetic linkage and do not segregate independently. Two markers that are close together will be separated by crossovers less frequently than two markers that are more distant from each other. The frequency with which the markers are separated by crossovers will be proportional to how far apart they are on their chromosome. Though, the recombination rate varies throughout the genome (Nachman 2002), there are genomic regions with high and low recombination rate. These are alleged recombination hot and cold spots. Thus, there is no simple linear relationship between genetic and physical map distances. A linkage map uses the frequency of recombination between two loci as a measure of the distance between them. The map distance is measured in centiMorgan (cM); a recombination frequency of 1% equals a genetic distance of 1 cM. To be able to do rapid and automated constructions of linkage maps computer programs are needed. A frequently used software package is the CRI-MAP package (Green et al. 1990).

Phenotypic data
Phenotypic measures are collected in the F₂-animals for the traits of interest, e.g. different behaviour, coat colour or body composition. It is essential to measure the phenotypic traits as precisely as possible and to collect information on factors that may influence the trait. Depending on the measured trait different factors, such as sex, age, place and time point, may influence the trait. For example, when measuring body weight it is crucial that the animals studied are of the same age.

QTL analysis
A quantitative trait locus (QTL) is a chromosomal region, which constitutes part of the genetic variation for a given quantitative trait. The goal of QTL analysis is to determine the loci that are responsible for the variation in a complex trait. The objective is to determine the number, the locations and the interaction (i.e. epistasis) of the loci and finally identify the causative genes and mutations. The localization of QTLs to small regions within chromosomes requires that the markers are closely spaced along the chromosome. The power to detect a QTL will depend on how much of the phenotypic variation that is explained by a given locus. QTL mapping is preferably performed in experimental crosses between genetically divergent populations and where sufficient number of animals can be obtained (Doerge 2002). Mapping QTLs for complex trait in humans is more difficult as many families segregating for the trait of interest are necessary in order to obtain an adequate amount of material. Instead, association analyses are often performed where correlation between a genetic marker and the phenotype is measured between two groups; cases vs. controls.
Fine mapping
One approach to narrow down a QTL interval is to build a more comprehensive linkage map of the QTL region. A dense map will position the QTL more precisely and this can also be used to identify animals that are recombinant for the QTL interval. The genome of many species has been sequenced which thus facilitate the search for new SNPs and microsatellites. If a genome has not yet been sequenced a comparative map (see below) to a species where the genome has been sequenced can be created, e.g. to the human genome sequence which was released 2001 (Lander et al. 2001; Venter et al. 2001).

Another approach to determine the order of genes is to develop a radiation hybrid map (RH map). The RH map is not, as the linkage map, limited by the need of polymorphic markers. RH mapping take advantage of somatic cell approach first described by Goss et al. (1975) and further developed by Cox et al. (1990). Briefly, a genome of interest from a donor is irradiated and fused to a rodent recipient cell line. The larger the X-ray doses, the smaller the donor fragments will be and the resolution will be higher. A panel of radiation hybrids is established and analysed for the presence or absence of DNA sequences from the donor genome, usually by PCR. The frequency with which two markers co-segregate reflects the distance and allows ordering of markers. The distance is measured in centiRay (cR) and one cR is equivalent to a 1% probability that a chromosome break has occurred between two markers after a given X-ray dose (Cox et al. 1990). The cR distances have a more linear relationship to the physical distance than cM.

Comparative mapping
Genes that are found close to each other in one species are also often found close to each other in other species. This is a result of shared evolutionary history. The human genome sequence was released in 2001 (Lander et al. 2001; Venter et al. 2001) and the latest release (March 2006, NCBI Build 36.1) is considered to be "finished" with less than one error per 10,000 bases and with few remaining gaps. Comparing a genomic region from a species with the corresponding region in the human genome provides a first preliminary indication of the gene content of the species of interest. Furthermore, if a region in another mammal has a conserved segment in the human genome, the region can be used for developing new markers in the species of interest. When a new marker has been identified linkage analysis or RH mapping can be performed to establish the relative position of that marker and to conclude if the gene order is conserved between the two species or not.
**Backcrossing**
By performing backcrossing to one of the parental lines the background genetic variation is reduced, the proportion of genes from the recipient parental line is reduced by 50% in each generation. In each backcross, boars that are recombinant between the flanking markers of the QTL interval are selected and progeny testing is performed to test if the trait is segregating (Fig. 4). Depending on the QTL result either side of the recombination point can be excluded (Fig. 4). In order to identify recombinants a dense map in the region has to be constructed. By backcrossing the QTL region can be reduced, but the smaller the interval the more animals are needed in each backcross generation in order to generate recombinants. Thus, at some point the backcrossing approach becomes very expensive.

![Progeny testing of two recombinant chromosomes](image)

**IBD mapping**
Identical by Descent (IBD) mapping is an approach that can be used to refine the map position of a QTL and also to identify causative mutations. Breeds that have been selected for a common phenotype will share one or a limited number of mutations. Each mutation has most likely occurred only once in a common ancestor of the animals carrying the particular mutation. Consequently, these animals will share haplotypes that are IBD and carry the causative mutation. A dense set of genetic markers is used to determine the minimum shared haplotype i.e. a minimum chromosomal segment, among a breed or a group of animals that carry a specific mutation. The size of the IBD segment depends on the recombination rate in the interval and on the number of generations since the mutation arose.
Fat metabolism

Fat metabolism comprises a huge and complex field. This brief introduction therefore focuses on parts that are relevant for this thesis.

Visceral fat

Obesity and overweight is a growing problem in the world today. The World Health Organization (WHO) has estimated that over 300 million people suffer from obesity (http://www.who.int/dietphysicalactivity/media/en/gsfs_obesity.pdf). It has been shown that visceral or abdominal fat constitutes larger health risks than subcutaneous fat (Despres 2006). Obese people with a high amount of visceral adipose tissue are in the risk zone to develop for instance insulin resistance, high blood pressure and elevation of lipids in the blood. All these metabolic abnormalities are features of the metabolic syndrome (Despres 2006). Insulin resistance often leads to Type 2 diabetes. Insulin resistance means that cells become resistant to the effect of insulin and higher amount of insulin is needed in order for the insulin to have its effect. When the pancreas can no longer produce enough insulin, glucose levels increase. Eventually the glucose levels are high even in fasting state and type 2 diabetes has developed.

Brown adipose tissue

Brown adipose tissue (BAT) is an organ expected to be found in all mammals. It is abundant in small animals as well as in newborns of larger mammals. BAT produces heat in the absence of muscular activity and ensures that the animal is kept warm under cold conditions, an adaptive process termed non-shivering thermogenesis.

Morphology

Brown adipocytes have multilocular lipid vacuoles and are rich in large mitochondria. BAT is mostly found in the neck, between and underneath the shoulder blades, in the area around the heart and oesophagus and around the kidney. Islets of brown adipocytes can also be found in white adipose tissue depots. The organ possesses plastic properties; cold exposure induces the development of brown cells whereas a positive energy balance induces transformation of brown fat cells into cells similar to white adipocytes (i.e. a large lipid vacuole and less and smaller mitochondrion). In general the number of brown adipocytes tends to decrease with age in all depots (Cinti 2005).

Uncoupling protein 1

Uncoupling protein 1 (UCP1) is exclusively expressed in BAT and is involved in nonshivering thermogenesis. UCP1 is located in the inner
membrane of mitochondria where it uses the proton gradient generated during respiration to produce heat instead of adenosine triphosphate (ATP). The brown adipose uncoupling is regulated by sympathetic nervous system through noradrenaline release, which activates the β-adrenergic receptors and a cascade of events leading to heat production in BAT (Cannon et al. 2004). The control and mechanism behind the UCP1 proton transport is not fully understood. Free fatty acids are involved in the activation of UCP1 and/or the transport mechanism but the question is still how; as allosteric regulators, as cofactors, or as proton shuttles (Cannon et al. 2004).

The key features differentiating white (i.e. visceral and subcutaneous fat) and brown adipose tissue are the presence of UCP1 in BAT and the proton conductance pathway.
Present investigations

Research aims

1. To construct a high-resolution linkage and RH map of the \textit{FAT1} QTL interval and a comparative map to the human and mouse genomes.

2. To refine the map of the \textit{FAT1} QTL region on pig chromosome 4 by repeated backcrossing.

3. To identify the mutation(s) influencing the \textit{FAT1} by IBD mapping of the QTL interval.

4. To perform genetic analysis of the pig Uncoupling protein 1 gene.

Characterization of a QTL for fatness on pig chromosome 4 (paper I, II and III)

Background

The Large White (or Yorkshire) breed originated in Yorkshire County, England. The breed was first recognised as a distinct breed in England in the year 1868 (Jones 1998). The Large White has white colour and is large in size. Their ability to grow fast and not lay down excess fat is a favourable trait (Briggs 1983). The Large White is phenotypically very different from the ancestral wild boar that has adapted to survive in their natural environment and thus put on excess energy as fat.

A cross between the Large White and the European wild boar was initiated in the late 1980s (Andersson et al. 1994). The founders were two European wild boars and 8 Large White sows. Four sires and 22 dams from the F\textsubscript{1} generation were reciprocally intercrossed to generate 200 F\textsubscript{2} progenies. About 240 linked markers, mostly microsatellites, were used to establish a linkage map covering almost the entire pig genome. The marker data combined with phenotypic measurements on the F\textsubscript{2} generation were
used for QTL analyses and a number of QTLs for different traits, such as carcass and immunological traits, were revealed (Andersson-Eklund et al. 1998; Edfors-Lilja et al. 1998, 2000). A major QTL for fatness and growth was identified on chromosome 4, denoted FAT1 (Andersson et al. 1994; Knott et al. 1998). Progeny that carried the wild pig chromosome 4 segment had higher fat deposition, shorter length of carcass, and reduced growth. To confirm the QTL and narrow down the interval a backcross was generated between two F2 sows and a Large White boar. One of the sows carried the wild haplotype and the other sow carried a recombinant haplotype in the QTL interval. Two BC1 boars, with the same QTL haplotypes as the corresponding mothers, were then backcrossed to ten Large White sows, generating 85 BC2 progenies. The position and the estimated effects of the QTL for growth and fatness were confirmed in the BC2 population (Marklund et al. 1999). Papers I-III describe further studies of the characterization of the FAT1 QTL on pig chromosome 4.

Results and Discussion

High-resolution map (paper I)

We used two different methods, linkage and RH mapping, to develop pig chromosome 4 (SSC4) maps. Pig expressed sequenced tags (ESTs) were mapped on the INRA-University of Minnesota RH panel (IMpRH, 7,000-rad of X rays (Yerle et al. 1998)) covering the entire chromosome. New markers developed around the conserved synteny breakpoint to human chromosomes (HSA) 1 and 8 and across the FAT1 QTL region were mapped by linkage and RH mapping.

Seventeen new markers located within and in the proximity of the FAT1 region were added to our previous linkage map of SSC4 (Marklund et al. 1999). The linkage map comprises altogether 34 markers and spans 135.9 cM on the sex averaged map (Fig. 5A). The order could be obtained for most of the markers, but four clusters were formed where there were no recombination events between the markers (Fig. 5A). These genes/markers are apparently located very close to each other, consistent with the position of these genes in the human genome. As a consequence, crossing over is a rare event and our pedigree is too small (200 F2 animals) to resolve the order of very closely linked markers. However, the orders of these loci were resolved by RH analysis (Fig. 5B).

Previous results from cytogenetic mapping indicate that SSC4 corresponds primarily to two human chromosomes, HSA1 and HSA8 (http://www.toulouse.inra.fr/lgc/pig-compare/SSCHTML/SSC4S.HTM). For the RH map, EST-sequences were selected based on the significance of BLAST scores to human orthologs, agreement of human cytogenetic and RH mapping localizations, and on equal distribution across human
chromosomes. Altogether, 77 new genes/markers were placed on the SSC4 RH map and 26 loci could be localized to their most probable interval. Part of the SSC4 RH map, emphasizing on the *FAT1* interval, is presented in Figure 5B.

The linkage and RH maps in the pig are in very good agreement. However, comparison of the genetic and physical map distances reveals that the recombination frequency across SSC4 varies significantly. We observe higher recombination rate towards the telomeric regions of the chromosome and very low recombination rate around the centromere, a phenomenon known to occur on many chromosomes in several species (Nachman 2002). The average ration of cR vs. cM across the chromosome was estimated at 30 cR\textsubscript{7000}/cM.
Figure 5. High-resolution maps of SSC4. A) The sex averaged linkage map of SSC4 includes 34 markers and has a length of 135.9 cM. B) Part of the SSC4 RH map emphasizing the FAT1 interval. The region is homologous to HSA8 (yellow) and HSA1 (green). Framework markers in black (Hawken et al. 1999) or grey (Milan et al., unpublished). Coloured marker names indicate likelihood of the position; additional orders are possible with no additional breaks required (orange); one order only but a second order with less than three breaks is possible (green); one order only, as other orders require more than three additional breaks (blue). Loci for which the two-point and multipoint localizations were not in agreement are positioned next to the comprehensive map. For these a thick bar indicates the interval calculated by multipoint analysis and an arrow shows the loci with highest two-point score.

* = EST

**Comparative map** (paper I)
The dense comparative gene map between SSC4 and HSA1/8 does not reveal any major rearrangements. There is some evidence that suggests a small inversion of the gene order compared to HSA8 around the centromere on SSC4. The gene order within the large synteny blocks is very well conserved on SSC4. There are some indicated inconsistencies between the human and pig gene orders but they all involve closely linked markers and the lod score values in the multipoint mapping analysis imply that the order needs to be confirmed before further conclusions can be made regarding the definite order. The breakpoint of conserved synteny with HSA1 and HSA8 is defined between UBE2V2 (HSA8q11.21), located close to the centromere, and SELE (HSA1.q24.2) and the distance between these markers is estimated to only 0.5 cM (Fig. 5B).

Comparisons of distances in the pig and human genomes reveal a close to linear relationship for the pig RH map (Rays) vs. the physical distances (Mbp) in human. The overall ratio on HSA8/SSC4 is 4.9 Mbp/Ray and on HSA1/SSC4 it is lower with 2.2 Mbp/Ray over the p arm on HSA1 and 3.5 Mbp/Ray across the q arm. The centromere has not been conserved, but repositioning of centromeres has been observed by many species suggesting that localisation of the centromere does not display the same degree of conservation as other parts of the genome (Ventura et al. 2001).

SSC4 corresponds to at least 13 chromosomal regions of the mouse genome. The HSA8 synteny block is divided into 9 mouse chromosome blocks and the HSA1 block into 4 mouse blocks which can be derived from the result from the human/mouse comparative map (Waterston et al. 2002).

**Characterization of the FAT1 locus on pig chromosome 4** (paper II and III)

*Backcrossing (Paper II)*

In paper II we traced the inheritance of the wild boar QTL allele for six generations in order to narrow down the FAT1 interval. For each backcross generation new boars, with a smaller and smaller portion of the wild boar
SSC4 segment, were selected (Fig. 6). The chosen boars were subsequently backcrossed to Large White sows and approximately 50 progeny were generated from each recombinant.

Boar BC3_{311} and BC5_{160} were highly significant for both abdominal and subcutaneous fat and the \textit{FAT1} QTL was clearly segregating in these boars. BC3_{65}, on the other hand, harbours a smaller proportion of the wild chromosome (Fig. 6) and was highly significant for the lean meat plus bone in ham only but showed a clear tendency for an effect on subcutaneous fat depth (p=0.07). Both the BC7_{161} and BC7_{333} boars were significant for abdominal fat and side fat at the last rib. BC7_{161} was also highly significant for lean meat content. None of them were significant for subcutaneous fat depth but the expected trend of higher subcutaneous fat associated with the wild boar allele was present.

The QTL analyses for the recombinants BC5_{157}, BC5_{162} and BC7_{328} were considered inconclusive since there was a tendency for a QTL effect (the wild boar haplotype associated with higher fat deposition) but it did not reach statistical significance for any trait. We cannot exclude the possibility that the wild type allele remains segregating at a low frequency in the domestic line in which case the lack of QTL segregation could sometimes occur because a backcross sire is homozygous for the wild type allele at \textit{FAT1}. Thus, haplotype data obtained from segregating sires should be given more weight than haplotype data from non-segregating sires.

We tested for the possible existence of a second QTL, as indicated by Marklund \textit{et al.} (1999), proximal to the wild/domestic breakpoint of BC3_{65}. The BC4_{672} boar was selected but clearly not significant for the QTL and we could exclude the region proximal to marker \textit{S0107} (Fig. 6). Furthermore, the effect on growth seen by Marklund \textit{et al.} (1999) could not be confirmed in this study.

If there is a single underlying locus for \textit{FAT1} we can reduce the critical interval to only 3.3 cM with \textit{RXRG} and \textit{SDHC/S0832} as flanking markers, indicated by arrows in Figure 6. Though, the two sires carrying the largest haplotype block from the wild boar, BC3_{311} and BC5_{160}, also showed the largest QTL effects which imply that it may be multiple genes underlying this QTL. Assuming a two-loci model, BC3_{311} and BC5_{160} should be segregating at both loci whereas BC7_{161} and BC7_{333} should only be segregating for a proximal locus located in the interval \textit{S0107-Sw714} and BC3_{65} should only be segregating for a locus in the interval \textit{Sw714-S0214}.
Figure 6. Summary of the genetic constitution as regards the \textit{FAT1} region of the backcross animals used for QTL analysis. The QTL status for each animal are presented; ++ = sire showing highly significant QTL effect; + = sire showing significant QTL effect; - = sire deduced to be not segregating for \textit{FAT1}; ? = QTL data inconclusive; n.t. = not tested for QTL segregation. The refined \textit{FAT1} interval, if assuming a one locus model, is indicated by vertical arrows and determined by the boars BC3\textsubscript{165}, BC7\textsubscript{161} and BC7\textsubscript{333}, all segregating for the QTL. The map distances are from the linkage map in Paper I. BCX\textsubscript{y}: BCX = backcross generation X, y = pig identity number. (This figure is modified from Fig. 1 in paper II)
**IBD-mapping (paper III)**

A reasonable hypothesis is that the mutations reducing fat deposition have been under strong selection in domestic lines and gone through selective sweeps. In this study we have thus performed an IBD-mapping of the \textit{FAT1} interval in order to search for alleles and/or haplotypes that are shared among the domestic lines. About 80 SNPs were genotyped in different pig breeds and wild boars. We identified two candidate regions, one defined by a polymorphism in the Retinoid X receptor-gamma (\textit{RXRG}) gene and the other by a haplotype block formed by four SNPs located in the middle of the \textit{FAT1} region.

The SNP in \textit{RXRG} (identified in Paper I) showed an interesting pattern where the \textit{G} allele is nearly fixed in the domestic breeds and allele \textit{A} is fixed among the wild boars. Interestingly, one of the two founder wild boars, W2, carry a wild/domestic recombinant haplotype and appeared to be heterozygous for a haplotype also found in domestic pigs for the region including \textit{ATP6V1H}, \textit{Sw1364} and \textit{RXRG} (Fig. 6). W2 is thus heterozygous \textit{A/G} for the \textit{RXRG} SNP which gave us an opportunity to test if the haplotype carrying the \textit{G} allele and transmitted from the W2 wild boar showed a significantly different QTL effect than wild boar haplotypes carrying the \textit{A} allele. The result show that a QTL model based on the \textit{RXRG} genotype is significantly better than a model based on the population of origin (wild boar versus domestic) for abdominal fat but not for subcutaneous fat. This indicates a close association between this haplotype and a causative mutation.

The defined haplotype block in the second region, denoted \textit{HAP1}, was almost fixed among domestic breeds. The haplotype is ~118kb and located in a gene desert comprising 1.25 Mb between \textit{CDCAI} and \textit{PBX1} in the human genome. \textit{HAP1} is also present in European wild boars but at a significantly lower frequency (\(p=0.63\)). A haplotype, denoted \textit{HAP2}, was found in the European wild boar population tested (\(p=0.32\)) and our two parental European wild boars were homozygous for \textit{HAP2}. Two of our eight dams in the parental line were heterozygous \textit{HAP1/HAP2} and the dams transmitted \textit{HAP2} to one \(F_1\) sire and three \(F_1\) dams. Thus, we could test for differences in QTL effects associated with \textit{HAP1} and \textit{HAP2} and QTL analysis showed that the model based on the population of origin (wild versus domestic) was significantly better than a model based on the \textit{HAP1/HAP2} locus ignoring population of origin for both fat traits.

**Summary paper II and III**

The results presented in paper III provide evidence of two QTLs in the \textit{FAT1} region: the \textit{RXRG} SNP, with a primary effect on abdominal fat (\textit{FAT1a}) and a locus, located in the interval \textit{RXRG} – \textit{S0214}, with a primary effect on subcutaneous fat (\textit{FAT1sc}). The two-loci model was suggested in paper II.
but the data were not conclusive. The backcross study was designed to
distinguish between segregation at a QTL with major effects on fatness
versus no QTL segregation and the progeny groups were not sufficiently
large to resolve multiple loci. Also, repeated backcrossing change the
genetic background which may influence epistatic interaction, i.e. interaction
between loci. Epistasis often contributes to the phenotype of a complex trait
and should be taken into consideration (Carlberg et al. 2004).

As the wild boar derived segment is split up we cannot distinguish a clear
division between boars segregating for abdominal or subcutaneous fat. An
explanation for this may be epistatic interactions within the \textit{FAT1} interval
and a certain segment of the wild chromosome is required to obtain the full
QTL effect. Boar BC5\textsubscript{160} shows the same strong effect for both fat traits as
BC3\textsubscript{311} and has a large wild chromosomal segment (Fig. 6). For BC3\textsubscript{65},
BC7\textsubscript{161} and BC7\textsubscript{333} the effect is there but despite the fact that the two
segments should combine the full QTL effect we cannot observe a clear
relationship. This indicates that there may be a more complex genetic pattern
behind the \textit{FAT1} phenotype.

The SNP in \textit{RXRG} is a protein coding mutation (Ala403Thr) at a highly
conserved position. The normal variant, Alanine, is found in domestic lines
and thus the European wild boar carries the “mutant” allele. The results
imply that the mutation is a recent event that has occurred in the European
wild boar, maybe providing an adaptation to colder environment. The \textit{RXRG}
gene is a potential candidate for \textit{FAT1a}. It encodes a transcription factor
which binds retinoid 9-cis-retinoic acid and forms heterodimers with many
different partners, e.g. retinoic acid receptors (RAR) and peroxisome
proliferator-activated receptors-\textgamma (PPAR\textgamma). It has previously been shown that
\textit{RXR}\textgamma-deficient mice have a higher metabolic rate than the wild-type and are
resistant to weight gain when fed with high fat diet (Brown et al. 2000;
Haugen et al. 2004). Furthermore, PPAR\textgamma and RXRs are both expressed in
adipocytes and Metzger et al. (2005) show that PPAR\textgamma/RXR heterodimers
are important for adipocyte differentiation and function.

Future prospects
The pig genome will be released shortly (Sanger Institute, http://www.sanger.ac.uk/Projects/S_scrofa). This will greatly facilitate the
identification of new markers to fill in the gaps of the recombination
intervals of the backcross animals (Fig 6). This would help characterize the
\textit{FAT1} interval by comparing the QTL results and a more exact genetic
constitution of each animal. The wild/domestic haplotype (\textit{ATP6V1H-Sw1364-RXRG}, Fig. 6), present in one of the parental wild boars, W2,
showed a high correlation with abdominal fat but not with subcutaneous fat.
Thus, the locus primarily controlling subcutaneous fat should be located in
the interval where this boar is homozygous for the wild type (RXRG-S0214, Fig. 6).

Is the coding mutation of RXRG in strong association with FAT1a or in fact the causative mutation? Functional studies of the RXRG gene is required to investigate if the coding mutation changes the function of the protein in some way. It would indeed be relevant to characterise the FAT1a more exact. Our Large White intercross may be a model for mild obesity and type II diabetes. Type II diabetes has shown to be in strong association with abdominal fat (Despres 2006).

Even if the defined haplotype, HAP1, does not explain our FAT1 QTL it would be interesting to investigate. The haplotype indicates a selective sweep in domestic pig breeds and is positioned in a gene desert. The haplotype could for example harbour a SNP in a cis-acting element influencing transcription far away or in a yet unidentified transcript, which is favourable for the domestic pigs and thus has lead to fixation.

Genetic analysis of the uncoupling protein 1 (UCP1) gene in the pig (paper IV).

Background

Brown adipose tissue (BAT) has never been isolated from pig and it has been shown that newborn pigs use shivering as the main thermogenic mechanism (Mount 1968; Berthon et al. 1995) in contrast to other mammalian newborns which use non-shivering thermogenesis. In their natural environment, pigs, as the only hoofed animals, make nests for their piglets in order to keep them warm (Jensen 2002). In farms, newborn pigs are housed under infrared lamps for the same reason (Fig.7). The non-shivering thermogenesis is mediated by the uncoupling protein 1 (UCP1), which is exclusively expressed in BAT (Cannon et al. 2004). No conclusive evidence for the presence of BAT has yet been demonstrated in pigs (Dauncey et al. 1981) and there are two studies published analysing the presence of UCP1 protein in piglets and they provide inconsistent results. Henningfield et al. (1987) have reported UCP1 protein in adipose tissue from piglets younger than 1 week, but Trayhurn et al. (1989) could not find any evidence of UCP1 in pig adipose tissue. The study by Trayhurn et al. (1989) is more extensive; they have examined pigs at three different ages (4 days, 4 weeks and 8 weeks of age) as well as the effect of pigs acclimatized to cold. In paper IV we perform a genetic analysis of the pig UCP1 gene.
Results and Discussion

In this study we show that the \textit{UCP1} gene is disrupted in the pig lineage. Alignment with the human \textit{UCP1} sequence revealed two gaps in the pig sequence reducing the total size from 10.1 kb in human to 4.3 kb in the pig. The first gap is located in intron 2 and represents an insertion in the human lineage or a deletion that occurred before the split of the pig and cattle lineages. The second gap identifies one unique deletion in the pig sequence which eliminates exons 3 to 5. The deletion was confirmed in different pig breeds, as well as in European wild boars and in Bornean bearded pig (\textit{Sus barbatus}), wart hog (\textit{Phacochoerus africanus}) and Red river hog (\textit{Potamochoerus porcus}). Furthermore, the \textit{UCP1} coding sequence in pigs contains three additional disrupting mutations: a two bp insertion in exon 1, a 16 bp deletion in exon 2 (both causing frameshifts) and a mutation in exon 6 leading to an early stop codon.

Pigs and cattle diverged from the human lineages at the same time point (estimated to ~ 94 million years ago (Springer \textit{et al.} 2003)). Thus the frequency of synonymous substitutions (\(d_s\)) and non-synonymous substitutions (\(d_N\)) should have accumulated in approximately the same rate since the separation. The frequency of \(d_s\) was nearly identical when comparing human with pig and human with cattle (Fig. 8). In contrast, \(d_N\) between human and pig (13.9±2.4%) was significantly higher than between
human and cattle (8.8±1.9%) (Fig. 8). This implies that porcine *UCP1* was inactivated sometime subsequent to the divergence from the cattle lineage and that it has accumulated $d_S$ and $d_N$ at the same rate since then. The increased $d_N$ rate was used to estimate that the inactivation of *UCP1* happened about 20 million years ago.

![Figure 8. Rate of nonsynonymous and synonymous substitutions in pairwise comparisons of *UCP1* sequences from humans, mice, cattle, and pigs. The data are based on sequences from exon 1, 2, and 6. B, *B. taurus*; H, *H. sapiens*; M, *Mus musculus*; and S, *S. scrofa*. (Figure 1C from Paper IV).](image)

The inactivation of *UCP1* provides an explanation of the poor thermoregulation in piglets. *UCP1* knockout mice are cold sensitive but fully viable (Enerback *et al.* 1997). Furthermore, these *UCP1*-null mice develop brown adipose tissue which implies that disruption of *UCP1* in the pig lineage cannot by itself explain the absence of BAT.

The lost *UCP1* function and the ability to use BAT for thermoregulation ought to have occurred in a warm climate where no or only weak selection for this mechanism are expected. The wild boar is the only porcine species that has adapted to temperate climates whereas all other Suidae live in tropical or sub-tropical environments (Ruvinsky *et al.* 1998).

Pigs are the only hoofed animals that build nests for their young (Jensen 2002) and it has been shown that winter farrowing nests have a temperature of around 20 °C even with an outdoor temperature down to -20 °C (Algers *et al.* 1990). The wild boar sows give birth to up to 10 gilts (Clutton-Brock 1999; Jensen 2002) which is many more offspring compared to other hoofed mammals where the young are well developed at birth. Horse, cow and sheep have one or two offspring per litter. These observed phenomenons are possible compensatory mechanisms for the loss of *UCP1* and consequently the poor thermoregulation in piglets.
Future prospects

We have shown that the UCP1 gene was disrupted in the pig lineage about 20 million years ago which leads us to ask whether the loss of UCP1 and BAT are associated with other genetic changes during the evolution of the pig lineage? When the pig genome sequence is released it will be possible to perform bioinformatic analyses of genes involved in BAT metabolism.

It will be of interest to date the inactivation of the pig UCP1 gene more exact by analysing sequences from closely related species. For example, the Tayassuidae family (peccaries), which separated from pigs approximately 40 million years ago, or the Hippopotamidae family (though, close relationships of Hippopotamus with Suidae and Tayassuidae are debated (Boisserie et al. 2005)). The disruption has been confirmed in two of the three Suidae subfamilies; Phacochoerinae (wart hog) and Suinae (Red river hog, wild boar and bearded pig). We have not analysed the UCP1 in Babirusa from the subfamily Barbyrousinae. This subfamily is less related with the other two subfamilies and differs significantly from other Suidae (Ruvinsky et al. 1998; Niebert et al. 2005). Furthermore, a Babirusa sow often delivers two offspring compared to 3-4 in most Suidae (Ruvinsky et al. 1998) and up to 10 for the wild boar (Clutton-Brock 1999; Jensen 2002). An interesting study would be to investigate i) if the loss of UCP1 function closely correlates with maternal behaviour and litter size and ii) if there is an association between large litter size and a cold climate.
En del av grisens kromosom 4 kan förklara fettansättning


Grisen saknar brun fettvävnad

Brun fettvävnad finns i stor utsträckning hos nyfödda däggdjur, gnagare och hos djur som går i ide. Denna vävnad hjälper till att bibehålla kroppstemperaturen. Vid fettnedbrytningen i vanlig fettvävnad produceras molekyler som lagrar energi, ATP. I brun fettvävnad frigörs nästan all energi i form av värme. Man har inte identifierat brun fettvävnad hos gris eller påvisat existensen av proteinet uncoupling protein 1 (UCP1) som är specifikt för brun fettvävnad. UCP1-proteinen är nödvändigt för att bilda värme i stället för ATP. Vi har sekvenserat *UCP1*-genen hos gris och visat att genen inaktiverades (slogs ut) för ca 20 miljoner år sedan. Avsaknaden av UCP1-
proteinet och brun fettvävnad kan förklara varför griskultingar huttrar för att hålla värmen. Vid svinuppfödning används värmelampon för att hålla nyfödda griskultingar varma. Grisen är också det enda klövdjur som bygger bo åt sina ungar.
Acknowledgements

The studies presented in this thesis were carried out at the Department of Animal Breeding and Genetics at the Swedish University of Agricultural Sciences and at the Department of Medical Biochemistry and Microbiology at Uppsala University.

The project was supported by the Foundation for Strategic Research and the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning.

I would like to express my gratitude to all people that have supported and encouraged me during these years.

I would particularly like to thank:

Leif Andersson, main supervisor, for your guidance during this time. Your expertise and passion for science is very inspiring! I have learnt very much from you.

Maria Moller, co-supervisor, for introducing me to the lab and for all help and support during these years. Thank you for being an active supervisor even after moving from Uppsala. I really appreciate that.

Stefan Marklund, co-supervisor, for always taking the time to answer my questions, for all assistance in the lab and for valuable discussions. See you at “Vasaloppet” next year?!?

All co-authors for fruitful collaborations, especially Susanne Stern and Kjell Andersson for statistical analyses, Susan Ansell for all the BAC’s and Emmanuelle Bourneuf, it has been struggling and stimulating, good luck with your new position in France.

All former and present members of the lab for a good research atmosphere. Keep it up!! Thank you for all your encouragement and many laughs. I have really enjoyed the time.
Ulle, my roommate, it has been so much fun sharing the office with you, thanks for putting up with me 😊. We have had many talks about science and life. (What do we not know about each other?!..)

Lina and Susanne, for many valuable discussions about science and life, pep talks and for generously sharing your experiences and knowledge. You have been a tremendous support during these years. Lina, it was a great idée to join the mentor-program, thanks.

Ulla Gustafson, for all sequences and lab support and Gudrun Wieslander, for help with administrative matters and nice talks. The lab wouldn’t manage without you!

Ulla Schmidt and others at Lövsta for taking care of the pigs and providing all the blood samples.

All friends and colleagues from the UGSBR year, what a fantastic year! Thanks to the SNiB-2005 committee, it was productive and fun working with you.

Helena Mannerfelt, my mentor, for all support, encouragement and guidelines for my future career. I always felt full with self-confidence after our meetings.

To ALL my friends outside the “lab world”, for keeping my mind of science and for being fantastic friends. You all mean so much to me!

My dearest family, mum, dad and my sister Karin, thank you for always being there, you are the best family one can wish for.

Oskar, for your love and understanding and for always looking at life from the bright side. When I am with you I feel happy and confident. You are my love, my friend and everything…
References


Freking, B. A., Murphy, S. K., *et al.* (2002). Identification of the single base change causing the callipyge muscle hypertrophy phenotype, the only known example of polar overdominance in mammals. *Genome Res* **12**: 1496-506.


Green, P., Falls, K., *et al.* (1990). *Documentation of CRIMAP, ver. 2.4*. St. Louis, Washington University School of Medicine, St. Louis.


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 164

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)