Identification of target sequences for association studies - analysis of the pig *FABP3* and *FABP4 loci* using comparative genomics methods[*]

A. Chmurzynska[1,2], J. Cieslak[1], T. Jankowski[1], M. Szydlowski[1] and M. Switonski[1,3]

University of Life Sciences in Poznań,
[1]Department of Genetics and Animal Breeding,
[2] Department of Human Nutrition and Hygiene
Wołyńska 33, Poznań 60-637, Poland

ABSTRACT

Comparative genomics facilitates the identification of conserved and potentially functional elements. Their polymorphism can be responsible for phenotype variability. Two chromosomal regions were studied - one containing the *FABP* gene cluster and the second harbouring the *FABP3 locus*. Several highly conserved elements across four species (human, mouse, dog and cattle) were found, two of them were selected for association studies in the pig (the *FABP4* enhancer and a fragment of intron 2 of the *FABP3* gene). Four single nucleotide polymorphisms in the *FABP4* enhancer sequence were found: 9 T/G, 31 C/G, 186 G/A, and 189 C/A. One of them (9T/G) was tested but the statistical analysis revealed no association with fatness traits in Polish Large White and Polish Landrace pigs. These studies demonstrate an integrated approach to identify, classify and select target DNA sequences for further association studies.

KEY WORDS: pig, multispecies conserved sequence, enhancer, *FABP3*, *FABP4,* fatness, SNP

INTRODUCTION

Complete sequences of some mammalian genomes have provided new opportunities to understand their functioning. A promising approach for the

functional annotation of the genomes is to search for conserved orthologous sequences from multiple species, the so-called multispecies conserved sequences -MCSs (a group including coding and noncoding elements) or conserved noncoding sequences - CNSs (noncoding elements only) (Margulies et al., 2003; Thomas et al., 2003). Segments in the mammalian genome under purifying selection and thus likely functional are estimated to comprise about 5% (Mouse Genome Sequencing Consortium, 2002). The protein-coding regions occupy only ~1.2% of the human genome and the untranslated regions (UTRs) are estimated to cover ~ 0.7% of the genome (International Human Genome Sequencing Consortium, 2004). The remaining 3-3.5% is likely to contain regulatory regions such as promoters, enhancers, silencers, *locus* control elements, nuclear matrix attachment sites, RNA genes, origins of replication and perhaps some novel functional elements. Polymorphism of these regulatory regions should be taken into consideration when associations between genotype and phenotype are studied.

Obesity is one of the main health problems in developed countries and is a result of a complex interaction between genetic and environmental factors (Mutch and Clement, 2006). To understand how genetic background influences obesity-related phenotype many association studies have been conducted. The association studies (i.e. a test for linkage between a particular polymorphism and a trait of interest) are based on tests using candidate genes or alternatively on genome scans using evenly spread polymorphic markers. Genome scanning reveals the map location of a region harbouring a trait locus with a major effect. The chromosomal regions are then precisely analysed in order to identify a candidate gene and finally a causative mutation. Numerous genome scans have been performed and many quantitative trait *loci* (QTLs) and candidate genes for the porcine fatness traits have been found. The most promising seem to be QTLs located in the pig chromosomes 4 and 6 (http://www.animalgenome.org /QTLdb/pig.html; Hu et al., 2007). Among functional candidates for fatness traits, located in those QTL-harbouring chromosomal regions there are *FABP4* (*A-FABP*, adipocyte fatty acid-binding protein) - chromosome 4 and *FABP3* (*H-FABP* heart fatty acid-binding protein) - chromosome 6. FABPs are transport proteins involved in intracellular fatty acid movement. However, despite extensive studies no unambiguous causative mutations in these two FABP *loci* have been identified (Chmurzynska, 2006).

Therefore the aim of our study was the detailed annotation of the chromosomal regions containing *A-FABP* (*FABP4*) and *H-FABP* (*FABP3*) *loci* in order to select conserved and potentially functional elements for association studies.

MATERIAL AND METHODS

Comparative analysis of the genomes

The human, mouse, dog and cattle chromosome segments were retrieved from the NCBI database (http://www.ncbi.nlm.nih.gov) and masked for repeats using RepeatMasker. Global pair-wise alignments were performed with the mVISTA server (Frazer et al., 2004) - alignment program - AVID (Bray et al., 2003). Each of the four pair-wise sequence comparisons was searched for conserved elements ($\geq$100 bp and $\geq$75% sequence identity).

All previously known coding regions and MCSs overlapping repetitive sequences were excluded, and we focused on those conserved noncoding sequences, which aligned in all studied mammalian species. The next step of the studies was to classify the identified MCSs: all CNSs were analysed with BLAST programs (http://www.ncbi.nlm.nih.gov/BLAST/). Plots for the CpG dinucleotides were generated with the CpGplot program (http://www.ebi.ac.uk/ emboss/ cpgplot/index.html). Prediction of the noncoding RNAs was performed with the 'RNAZ' program (Washietl et al., 2005), which relies on a measure for RNA secondary structure conservation and a measure for thermodynamic stability. All putative ncRNAs were compared with databases of noncoding RNAs: Rfam (Griffiths-Jones et al., 2005), NONCODE (Liu et al., 2005) and RNAdb (Pang et al., 2005). Searching for microRNA targets were performed with miRanda (http://www.microrna.org/). All CNSs were also compared with the S/MARt database containing scaffold/matrix attachment regions - SAR/MAR (Liebich et al., 2002).

Polymorphism and association studies

Material. Animals representing nine breeds: Polish Large White, Polish Landrace, Duroc, Hampshire, Pietrain, Zlotnicka White, Zlotnicka Spotted, Pulawska and Vietnamese, a synthetic line - 990 and the wild pig were used for polymorphism searching. Polish Large White and Polish Landrace pigs were used for association studies. Pedigree structure was similar in each line: average size of sire and dam (fullsibs) group was 5 and 2 gilts, respectively. All pigs were fattened and slaughtered at a Test Station (Pawłowice, Poland). During the fattening period, from 25 to 100 kg liveweight, gilts were fed *ad libitum* with commercial mixed fodder. Nine fatness traits were considered: abdominal fat weight (AFW), intramuscular fat (IMF) and seven measurements of back fat thickness (BF).

Searching for polymorphism in conserved sequences. Searching for poly-morphism in two identified CNSs was performed with the use of PCR-SSCP. CNS localized in the *FABP* gene cluster region was identified as enhancer of the

FABP4 gene. PCR primers (ENH-F: 5' CAA GAA TCT GGT GGA AGG GGT AAT3', ENH-R: 5' CCC ATC CAA GGA CGA TTT TGA TT 3') were selected in order to amplify the highly conserved part of the enhancer (the expected length of the amplified product - 272 bp). Amplification was performed on genomic DNA (100 ng). Reaction mixture (a total of 20 μl) contained 100 ng of genomic DNA, 2 μl of 10×PCR buffer (700 mM Tris-HCl, pH 8.6; 166 mM $(NH_4)_2SO_4$; 25 mM $MgCl_2$), 7.5 pmol of each primer, 200 μM of dNTPs and 0.75U *Taq* polymerase (Novazym). The cycling profile was: 94ºC for 3 min followed by 35 cycles of 94ºC for 30 s; 45ºC for 45 s, 72ºC for 35 s and finally 72ºC for 10 min.

One CNS located in the second intron of the *FABP3* gene was analysed. PCR primers: INT-F: 5' GCA CCT TGA GGG GTA GGA TGT TAT 3', and INT-R: 5' AGG GGG AAG AAA GCC AGA GGT GTT 3' were designed. The PCR product was 251 bp long. Amplification was performed as described above. Annealing temperature was 57ºC.

All PCR products were screened with the use of the SSCP technique. Electrophoresis was run overnight at room temperature. Then the gel was stained with silver nitrate according to the standard protocol.

Sequencing. Selected DNA samples were sequenced at the Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw (Poland).

Genotyping. Genotyping was performed for one of the identified new polymorphic sites in the porcine *FABP4* enhancer - 9 T/G, for animals of the Polish Large White and Polish Landrace breeds. PCR products were digested overnight with the *FspB*I (*Mae*I) restriction enzyme and separated on 2.5% agarose gel. Digestion resulted in three genotypes: TT (258bp), GG (30 and 228bp) and TG (30, 228 and 258 bp).

Statistical analysis. Association between the *FABP4* enhancer polymorphism and phenotypic traits was tested by the use of ANOVA including the fixed effects of genotype at the 9 T/G site (*FABP4* enhancer) and at the *RYR1 locus* (site 1843 C/T) and the random effect of a sire. Statistical analysis were performed for Polish Large White and Polish Landrace pigs separately and for a pooled group as well.

RESULTS

Comparative analysis of the genomes. To identify MCSs a global alignment program - AVID (Bray et al., 2003) visualized by the VISTA server (Frazer et al., 2004) was used and three pairwise alignments were performed: human-mouse, human-dog and human-cattle. Vista highlights regions with over 75% conservation in the 100 bp-window. The next step was the selection of MCSs for detailed analysis - all known coding regions were excluded. Because our aim was to search for functional elements of the genome common for mammalian species,

for further studies only those conserved noncoding sequences were selected, which were identified in all alignments.

The *FABP4* gene is located within a gene cluster containing three other FABP *loci*: *FABP5* (*E-FABP*, epidermal fatty acid-binding protein), *FABP8* (*PMP2*, myelin fatty acid-binding protein) and *FABP9* (*T-FABP*, testis fatty acid-binding protein), therefore the region spanning the whole cluster was compared. Altogether 45 conserved elements in this region were identified, but some of them overlapped with repetitive sequences (8 completely and 2 only partially) and thus they were not taken into consideration. The length of the MCSs ranged from 49 to 470 bp, with an average of about 170 bp, and they comprised approximately 2.74% of the entire analysed sequence length. Over 23% of the MCSs were coding and about 77% noncoding, namely 21% of them fall in introns, 10 and 6% in 5' and 3'UTRs, respectively, and the remaining 40% fall in the intergenic regions (Figure 1a). The effectiveness of the annotation was rather small, because analyses did not reveal any CpG islands, noncoding RNAs (ncRNAs), microRNA targets or SAR/MARs (scaffold/matrix attachment regions). Only two of the analysed CNSs were recognized as the *FABP9* promoter and the *FABP4* enhancer. The MCS containing the *FABP4* enhancer, consisting of 146 bp, was subjected for further polymorphism and association studies.

In the second analysed region, containing the *FABP3* gene, 18 conserved elements were identified, but two of them overlapped with repetitive elements and thus were excluded. The length of the 16 MCSs varied from 9 to 236 bp (mean of about 114 bp) and occupied approximately 3.8% of the whole region length.
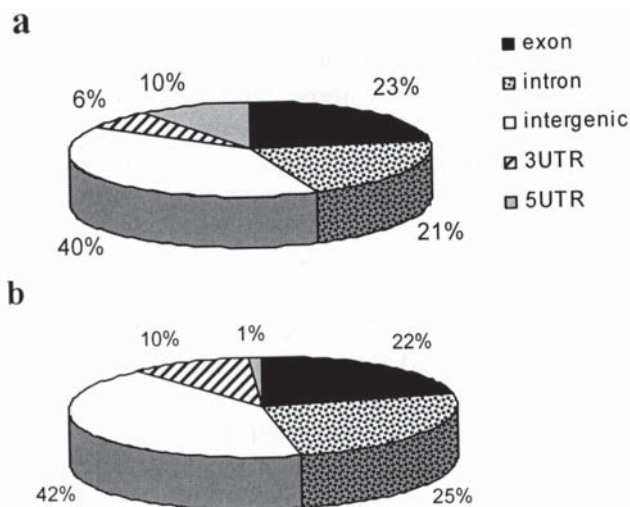

Figure 1. Fraction of bases of five annotation types (exon, intron, intergenic, 5'UTR and 3'UTR) in the two analysed regions - containing *FABP4* (a) and *FABP3* (b) genes. Annotations were performed for human genome

Almost 22% of the MCSs fall in coding regions, 25% in introns, 1 and 10% in 5'UTR and 3'UTRs, respectively, and 42% in the intergenic regions (Figure 1b). Three CNSs were predicted as ncRNAs, but a comparison with ncRNA sequences available in the Rfam, NONCODE and RNAdb databases did not reveal any similar sequence. A small part of the CNS located in 3'UTR of the *FABP3* gene appeared to be a microRNA target - hsa-miR-34b predicted by miRanda (http://www.microrna.org/). One CNS located upstream of transcription start site was recognized as CpG island. In the detected CNSs no SAR/MARs were found. The MCS located in intron 2 of the *FABP3* gene, consisting of 68 bp, was subjected to polymorphism studies.

Polymorphism and association studies. Two MCSs were selected for association studies. One was located in the *FABP* gene cluster region, upstream of the *FABP4* gene, and identified as its enhancer. The second was located in the second intron of the *FABP3* gene. The most conserved part of the *FABP4* enhancer was amplified and the obtained sequence - 211 bp was deposited in GenBank (DQ372683). Moreover, enhancer sequences of other studied mammalian species were compared. The overall sequence identity was: 63.4 67.3, 85.0 and 88.2% for the pig-human, pig-mouse, pig-dog and pig-cattle homologues, respectively. All transcription factor binding sites (TFBSs) characterized by Graves et al. (1992) in the murine *FABP4* enhancer were found: ARE1 (the NF-1 transcription factor site), ARE6, ARE2, ARE7 and ARE4 (Figure 2).

```
                10        20        30        40        50        60        70        80        90        100
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
mouse  GAATTCCAGCAGGAATCAGGTAGCTGGAG--AATCGCACAGAGCCATGC-GATTCTTGGCAAGCCATG-CGACAAAGGCAGAAATG---------------
pig    ----------------------------------CATATCTAGATTTGGAGAATCTTGGCAAACGCCT-GAATAAGTGCAGGAATG--------------
human  --------------------CAGCAGAAGGAATAACATCTAGATGTGGAGAATCTTGGCAAGCCACAGTAATAAATGCAGTAATTTATTCTGCATTTATT
dog    -----TTATCAAGAATCCTGTTTCAGGGATAATAATATCTAGGTTTGGAGAATCTTGGCAAAGCACT-GAATGAATACAGGAATG--------------
cattle ---CATTATCAAGAATCTGGTGGAAGGGGTAATAATACCCAGATTTGGAGAATCTTGGCAAACCAC--TGAATAATACAGGAATG--------------

                110       120       130       140       150       160       170       180       190       200
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
mouse  CACATTTCACCCAGAGAGAAGGGATTGATGTCAGCAG-GAAGTCACCACCCAGAGAGCAAATGGAGTTCCCAGATGCCTGACATTTG--CCTTCTTACTG
pig    TACATTTCACCCAGTGAGAGAGG---------------GAAGTTAACACCCAGATGGCAAATAAAGTTCACAGATGCTTGACATTTGGTGCATTTCTTACTC
human  CACATTTCACCTAGTGAGAGAGGATTTTCGTCAAGAGGGAAGTTAACACCCAGATGGCAAATGTATTTCACAGATGCCTGACGTTTGTGGTTTCTTACTC
dog    CACATTTCACCCAGTGAGAGAGGATTTTCGTCAGGAGGGAAGTTAACACCCAGATGGCAAATATAGTTCACAGATGCCTGACATTTGTGCTTTCTTACTC
cattle CACATTTCACCCAGTGAGAGAGGATTTTCGTCAAGAGGGAAGTTAACACCCACGCGGCAAACATAGTTTGCAGATGCCTGACATTTGTGCTTTCTTTCTC

                210       220       230       240       250       260       270       280       290       300
       ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
mouse  GATCAGAGTTCACTAG-TGGAAGTG----TCACAGCCCAAACACTCCCCAAAGCTCAGCCCTTCCTTGCCTTGTAACAATCAAGCCGCTCCTGGATGAA
pig    GATCAGAGTTCACTGGAGGGAAGTC----CCACAGCAAAAGTGCCTCCCAGAAAGCATACTTAGCGTTTCCTTGTAATC-------------------
human  GATCAGAGTTCACTGG-GGGAAATC----TCACAGCAAAAGTGCCTCCCAGAAAGAATACTCAGTGTTTCCTTGTAATCA--AAATTGTTCCTGGATGAG
dog    GATCAGAGTTCACTGGGGGAGAAGTC----TCACAGCAAAAGTGCCTCCCAGAAAGAATACTCAGTGCCTCCTTGTAATCT--AAATTATTCCTGGATGAG
cattle GATCAGAGTTCACTGGAGGGAAGTCATGGCCACAGCAAAAGGGCCCCACAGAAAGAATACTCAGTGTTTCCTCGTAATCA--AAATCGTCCTTGGATGGG
```
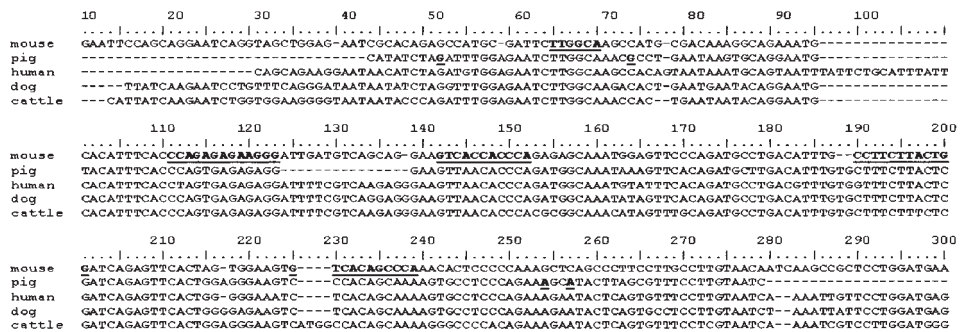
Figure 2. Alignment of the *FABP4* enhancer sequences from the mouse, pig, human, dog and cattle. Transcription factor binding sites (ARE-1, ARE-6, ARE-2, ARE-7, ARE-4) according to Graves et al. (1992) are underlined and bolded. In the porcine enhancer sequence polymorphic sites are underlined and bolded

Screening for polymorphism in the enhancer of the *FABP4* gene was performed for 35 pigs of nine breeds, one synthetic line and the wild pig. Four different SSCP patterns were found and the comparison of the sequencing results revealed four polymorphic sites: 9 T/G, 31 C/G, 186 G/A, and 189 C/A. Only the first one (9 T/G) could be analysed by the RFLP test (*FspB*I (*Mae*I) restriction enzyme),

therefore genotyping was performed only for this site. No polymorphism was identified in the MCS located in intron 2 of the *FABP3* gene.

An association between the 9 T/G polymorphism at the enhancer sequence of the *FABP4* gene and production traits were analysed in the Polish Large White and Polish Landrace breeds. Altogether 166 animals of Polish Large White and 191 animals of the Polish Landrace were genotyped. The frequencies were as follows: TT-0.34, GG-0.22, and TG-0.44 in Polish Large White and TT-0.17, GG-0.32 and TG-0.51 in Polish Landrace. All the animals were also genotyped at *RYR1* locus. The analysis gave no statistically significant results. Thus it can be concluded that the new polymorphism (9 T/G) is not a QTN (quantitative trait nucleotide).

DISCUSSION

In spite of the exponential development of structural genomics, knowledge of causative mutations that bring about phenotypic differences within and among species is still unsatisfactory. Human geneticists have been very successful in the identification of gene mutations causing monogenic disorders, but the recognition of major genes influencing variability of multifactorial traits (diseases) is rather limited. The great potential lies in studies of domestic animals, which have some advantages when compared with model organisms. Firstly, their phenotypes and pedigrees are monitored very precisely and, what is more, among breeds there are large phenotypic differences and a limited genetic variation.

Fatness of livestock and human obesity are complex traits, resulting from interactions between genetic and environmental factors. The pig seems to be a good animal model for unravelling the molecular basis of obesity. Numerous genome scanning experiments revealed QTLs for fatness traits in many porcine chromosomal regions. Unquestionably, chromosomes 4 and 6 harbour genes with major effects on fatness. However, in spite of many attempts, no such gene was pointed out. Good functional candidates are *FABP loci* which encode proteins involved in the transmembrane and intracellular fatty acid movement (Zimmerman and Veerkamp, 2002). The porcine *FABP4* is localized in 4q12 and *H-FABP* in 6q26, in chromosomal regions where QTLs for fatness traits were found (Szczerbal et al., 2007).

Several association studies concerning the *FABP4* and *FABP3* genes have been conducted. Gerbens et al. (1998) reported that polymorphism of microsatellite sequence in intron 1 of the *FABP4* was associated with fatness traits in a Duroc population, but studies on other pig breeds did not confirm this observation (Nechtelberger et al., 2001; Chmurzyńska et al., 2004). Moreover, Mercade et al. (2006) suggested that the InDel polymorphism in the first intron (position

2653) is tightly associated with fatness traits. SNPs in the promoter region and in intron 2 of the *FABP3* gene were analysed by Gerbens et al. (1999) who detected significant associations between three polymorphisms (one in the promoter region and two in the second intron) and fatness traits. Again, these results were not confirmed by other groups (Nechtelberger et al., 2001; Chmurzynska et al., 2007). It should be underlined that in all mentioned experiments genotype frequencies were dissimilar, thus comparison of the results is rather problematic. However, the test for an association between a new T(-158)G polymorphism in the promoter and fatness traits gave some significant results, but not in all the three studied breeds (Chmurzynska et al., 2007).

   All results mentioned above are rather inconsistent and for this reason we decided to look for conserved noncoding elements with a putative role for the functioning of FABP genes. The procedure for the detection of multispecies conserved orthologous sequences was identical as presented by Frazer et al. (2003). However, our approach was supplemented with the stage of CNSs classification. In the chromosome interval containing the *FABP* gene cluster all identified MCSs represented 2.74% of its length, while in the second analysed region it was almost 3.8%. It is rather little when comparing to the 4.3% predicted by Siepel et al. (2005) for the human genome or 5% estimated by Thomas et al. (2003) for the *CFTR* region. Our results could be affected by the arbitrary cutoff criteria to define conserved sequences (>75% identity over 100 bp), but it is also obvious that the saturation of the genome in conserved sequences is *locus*-specific. In both regions nearly the same proportion of coding to noncoding bases was found (23 to 77% and 22 to 78%), almost identical with previous observations of Siepel et al. (2005) and also Margulies et al. (2003). Conservation thresholds allowed the detection of all known coding sequences. In terms of UTRs, not all of them were entirely recognized as MCSs, but in spite of that fact, the relatively large fraction of MCS bases fall in UTRs - 16 and 11%, for both studied regions respectively (Figure 1). Siepel et al. (2005) observed for example that 1.1 and 3.6% of the conserved elements in the human genome belonged to the 5' and 3'UTRs, respectively. Similarly, Margulies et al. (2003) with the use of three different methods found that 3.9-4.9% of MCS bases overlapped with UTRs. In our study the remaining conserved sequences, found outside the coding exons, fall in introns (21 and 25% in the gene cluster and *FABP3* containing regions, respectively) and in the intergenic regions (40 and 42%). A similar distribution of the conserved elements was observed for the entire human genome -28.5% of the conserved bases fall in introns and 41.2% in the intergenic regions (Siepel et al., 2005).

   The next step of the study was the classification of identified CNSs. Searching the nr (non-redundant) database with BLAST programs made it possible to recognize two of them as the *FABP4* enhancer and the *FABP9* promoter. In the region containing the *FABP* gene cluster 9 CNSs were found, located within

1000 bp upstream of the transcription start site (TSS). Altogether, they represented over 15% of the all highly conserved bases and almost 20% of the CNSs. In the vicinity of the TSS of the *FABP3* gene, a highly conserved region was also found. This data may suggest that a substantial part of the conserved sequences serve as promoters. In fact, we did not search for TFBSs, due to huge number of the false positives produced by programs for TFBS prediction. However, the localization of these CNSs alone allows us to anticipate their functions. Our hypothesis is supported by the results obtained by Thomas et al. (2003). In the *CFTR* region they found that 1.7% of the non-exonic MCSs fall within 1-kb upstream of the TSS, which was 2.3 times more frequent than it would be expected if the MCS bases were randomly distributed. A total of 3 CNSs were found to contain candidates for ncRNA, but only in the *FABP3* containing region. All of them were located in the intergenic region. One CNS in the 3'UTR of the *FABP3* gene contained a putative microRNA target - hsa-miR-34b. MicroRNA targets are rather very limited in their length, thus it is most likely that microRNA binding is not the only role of this particular CNS. Both studied regions contained no SAR/MARs. A great majority of the MCSs remained unclassified. In the *FABP* cluster region the fraction of unannotated bases is about 72 or 60%, if we exclude CNSs with putative promoter functions. Similarly, in the second analysed region most of the MCSs (about 60% of bases) are of an unknown role.

For association studies we selected two MCSs: one located upstream of the *FABP4* gene, identified as its enhancer, and one MCS in *FABP3* intron 2, with an unknown role. The most conserved region of the *FABP4* enhancer was amplified. In the porcine enhancer TFBSs characterized by Graves et al. (1992) ARE1, ARE6, ARE2, ARE7, and ARE4 were found. All TFBSs are highly conserved among species (Figure 2). The ARE1 is the NF-1 site and its mutations resulted in a significant reduction of the enhancer activity (Graves et al., 1992). The ARE6 and ARE7 sites are target sequences for a heterodimeric complex of mPPAR gamma 2 and RXR alpha (Tontonoz et al., 1994). Searching for polymorphism of the porcine *FABP4* gene enhancer resulted in the identification of four SNPs: 9T/G, 31C/G, 186C/G, and 189C/A. Unfortunately, all of them were located outside the TFBSs. Only the first one (9T/G) was recognized by restriction enzyme *FspB*I (*Mae*I), therefore genotyping was performed for this site. Statistical analysis did not reveal any significant association between 9T/G SNP and fatness traits. No polymorphic variants of the MCS located in the second intron of the *FABP3* gene were found.

The performed procedure is a new variant of association studies. We implemented comparative genomics analyses of highly conserved sequences (MCS) in four mammalian species in order to select targets at known QTL regions for further polymorphism and association studies. A detailed annotation of the chromosomal QTL-harbouring region brings valuable information, which might be

very useful. Firstly, such an interspecies characteristic of the chromosome region reflects its evolution. Secondly, the identification of the conserved and thus putatively functional elements of the genome may promote an understanding of their function. Finally, such an approach improves the selection of potential candidate sequences for association studies. These conclusions are consistent with a recent report that SNPs in MCSs are useful markers in association studies and can be integrated with the current approaches for studies of genes contributing to complex diseases (McCauley, 2007).

REFERENCES

Bray N., Dubchak I., Pachter L., 2003. AVID: A Global Alignment Program. Genome Res. 13, 97
Chmurzynska A., 2006. The multigene family of fatty acid-binding proteins (FABPs): function, structure and polymorphism. J. Appl. Genet. 47, 39-48
Chmurzyńska A., Maćkowski M., Szydłowski M., Melonek J., Kamyczek M., Eckert R., Różycki M., Świtoński M., 2004. Polymorphism of intronic microsatellites in the *A-FABP* and *LEPR* genes and their association with productive traits in the pig. J. Anim. Feed Sci. 13, 615-624
Chmurzynska A., Szydlowski M., Stachowiak M., Stankiewicz M., Switonski M., 2007. Association of a new SNP in promoter region of the porcine FABP3 gene with fatness traits in a Polish synthetic line. Anim. Biotechnol. 18, 37-44
Frazer K.A., Elnitski L., Church D.M., Dubchak I., Hardison R.C., 2003. Cross-species sequence comparisons: a review of methods and available resources. Genome Res. 13, 1-12
Frazer K.A., Pachter L., Poliakov A., Rubin E.M., Dubchak I., 2004. VISTA: computational tools for comparative genomics. Nucl. Acid. Res. 32, W273-W279
Gerbens F., Jansen A., VanErp A.J.M., Harders F., Meuwissen T.H.E., Rettenberger G., Veerkamp J.H., TePas M.F.W., 1998. The adipocyte fatty acid-binding protein locus: characterization and association with intramuscular fat content in pigs. Mamm. Genome 9, 1022-1026
Gerbens F., van Erp A.J.M., Harders F.L., Verburg F.J., Meuwissen T.H.E., Veerkamp J.H., te Pas M.F.W., 1999. Effect of genetic variants of the heart fatty acid-binding protein gene on intramuscular fat and performance traits in pigs. J. Anim. Sci. 77, 846-852
Graves R.A., Tontonoz P., Spiegelman B.M., 1992. Analysis of a tissue-specific enhancer: ARF6 regulates adipogenic gene expression. Mol. Cell Biol. 12, 1202-1208
Griffiths-Jones S., Moxon S., Marshall M., Khanna A., Eddy S.R., Bateman A., 2005. Rfam: annotating noncoding RNAs in complete genomes. Nucl. Acid. Res. 33, D121-D124
Hu Z., Fritz E.R., Reecy J.M., 2007. AnimalQTLdb: a livestock QTL database tool set for positional QTL information mining and beyond. Nucl. Acid. Res. 35, D604-D609
International Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. Nature 43, 1931-1945
Liebich I., Bode J., Frisch M., Wingender E., 2002. S/MARt DB: a database on scaffold/matrix attached regions. Nucl. Acid. Res. 30, 372-374
Liu C., Bai B., Skogerbo G., Cai L., Deng W., Zhang Y., Bu D., Zhao Y., Chen R., 2005. NONCODE: an integrated knowledge database of non-coding RNAs. Nucl. Acid. Res. 33, D112-D115
Margulies E.H., Blanchette M., Haussler D., Green E.D., 2003. Identification and characterization of multi-species conserved sequences. Genome Res. 13, 2507-2518

McCauley J.L., Kenealy S.J., Margulies E.H., Schnetz-Boutaud N., Gregory S.G., Hauser S.L., Oksenberg J.R., Pericak-Vance M.A., Haines J.L., Mortlock D.P., 2007. SNPs in Multi-species Conserved Sequences (MCS) as useful markers in association studies: a practical approach. BMC Genomics 8, 266

Mercade A., Perez-Enciso M., Varona L., Alves E., Noguera J.L., Sanchez A., Folch J.M., 2006. Adipocyte fatty-acid binding protein is closely associated to the porcine FAT1 locus on chromosome 4. J. Anim. Sci. 84, 2907-2913

Mouse Genome Sequencing Consortium, 2002. Initial sequencing and comparative analysis of the mouse genome. Nature 420, 520-562

Mutch D.M., Clement K., 2006. Unraveling the genetics of human obesity. PLoS Genet. 2, e188

Nechtelberger D., Pires V., Solkner J., Stur I., Brem G., Mueller M., Mueller S., 2001. Intramuscular fat content and genetic variants at fatty acid-binding protein loci in Austrian pigs. J. Anim. Sci. 79, 2798-2804

Pang K.C., Stephen S., Engstrom P.G., Tajul-Arifin K., Chen W., Wahlestedt C., Lenhard B., Hayashizaki Y., Mattick Y.S., 2005. RNAdb--a comprehensive mammalian noncoding RNA database. Nucl. Acid. Res. 33, D125-D130

Siepel A., Bejerano G., Pedersen J.S., Hinrichs A.S., Hou M., Rosenbloom K., Clawson H., Spieth J., Hillier L.W., Richards S., Weinstock G.M., Wilson R.K., Gibbs R.A., Kent W.J., Miller W., Haussler D., 2005. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034-1050

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, 63-66

Thomas J.W., Touchman J.W., Blakesley R.W., Bouffard G.G., Beckstrom-Sternberg S.M., Margulies E.H., Blanchette M., Siepel A.C., Thomas P.J., McDowell J.C., Maskeri B., Hansen N.F., Schwartz M.S., Weber R.J., Kent W.J., Karolchik D., Bruen T.C., Bevan R., Cutler D.J., Schwartz S., Elnitski L., Idol J.R., Prasad A.B., Lee-Lin S.Q., Maduro V.V., Summers T.J., Portnoy M.E., Dietrich M.L., Akhter N., Ayele K., Benjamin B., Cariaga K., Brinkley C.P., Brooks S.Y., Granite S., Guan X., Gupta J., Haghighi P., Ho S.L., Huang M.C., Karlins E., Laric P.L., Legaspi R., Lim M.J., Maduro Q.L., Masiello C.A., Mastrian S.D., McCloskey J.C., Pearson R., Stantripop S., Tiongson E.E., Tran J.T., Tsurgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.S., Young A.C., Zhang L.H., Osoegawa K., Zhu B., Zhao B., Shu C.L., De Jong P.J., Lawrence C.E., Smit A.F., Chakravarti A., Haussler D., Green P., Miller W., Green E.D., 2003. Comparative analyses of multi-species sequences from targeted genomic regions. Nature 424, 788-793

Tontonoz P., Hu E., Graves R.A., Budavari A.I., Spiegelman B.M., 1994. mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer. Gene. Develop. 8, 1224-1234

Washietl S., Hofacker I.L., Stadler P.F., 2005. Fast and reliable prediction of noncoding RNAs. Proc. Nat. Acad. Sci. USA 102, 2454-2459

Zimmerman A.W., Veerkamp J.H., 2002. New insights into the structure and function of fatty acid-binding proteins. Cell. Mol. Life Sci 59, 1096-1116